

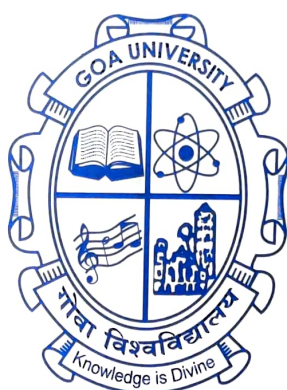
CHARACTERIZATION OF BACTERIOPHAGE FROM A HOT SPRING

A THESIS SUBMITTED IN PARTIAL FULFILMENT FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
in Biotechnology

IN THE SCHOOL OF BIOLOGICAL SCIENCES AND BIOTECHNOLOGY

GOA UNIVERSITY



By

Mrs. NICOLA FARIA

School of Biological Sciences and Biotechnology

Goa University

Taleigao Plateau, Goa-403206

JUNE 2023

Dr. P. A.
Dr. Parvati A.
12/12/2024

DECLARATION

I, **Nicola Faria** hereby declare this thesis represents work which has been carried out by me and that it has not been submitted, either in part or full, to any other University or Institution for the award of any research degree.

Place: Taleigao Plateau

Date: 30/06/2023


Nicola Faria

CERTIFICATE

I hereby certify that the work was carried out under my supervision and may be placed for evaluation.



Prof. Sanjeev C. Ghadi

Professor (Biotechnology)

School of Biological Sciences and Biotechnology

Goa University, Taleigao Plateau, Goa-403 206,

India.

ACKNOWLEDGEMENTS

Firstly, I thank the Lord Almighty to whom I owe my very existence, for providing me with this opportunity and granting me the capability to proceed successfully.

A doctoral thesis is undeniably a collaborative process. I would want to express my sincerer gratitude to everyone who assisted me during the years leading to my thesis.

I would like to express my deepest gratitude to my Ph.D. supervisor **Prof. Sanjeev. C. Ghadi** for his valuable and constructive guidance during the planning and development of this research work. He gave me total freedom to pursue my ideas and created the right environment for these ideas to come to fruition. Being one of the few first researchers in India to study viruses from hot springs was a challenging experience, but ultimately, I have learnt and expanded my knowledge of hot spring viruses. As it is said,

“We learn knowledge from failure much more than success”

My supervisor has supported me when I failed in my experiments and encouraged me to discover solutions to the problem. His unwavering support and attention persuaded me to make every effort to provide quality research and writing for this thesis.

I take this opportunity to acknowledge **Prof. Harilal B. Menon, the Vice-Chancellor, of Goa University and Prof. Varun Sahni (former Vice-Chancellor)** of Goa University for providing the necessary infrastructure and permission to carry out my research. I express my sincere gratitude to the present and former Deans of the School of Biological Sciences and Biotechnology, **Prof. Savita Kekar, Prof. P.K. Sharma, Prof. M. K. Janarthanam** and Vice-Dean **Prof. S. Krishnan** for their valuable guidance and encouragement.

I gratefully acknowledge the financial assistance received from University Grants Commission, India through the **UGC-Maulana Azad National Fellowship for Minorities (2015-20)**.

A very special thanks to **Dr. Samir Damare**, Senior Scientist, National Institute of Oceanography (NIO), Goa and **Prof. Usha Muraleedharan** for their valuable suggestion and constructive critiques as Vice Chancellor's nominee. I would like to offer my special thanks to the entire faculty of the Biotechnology program in the School of Biological

Sciences and Biotechnology; **Prof. Savita Kerkar, Prof. Usha D. Muraleedharan, Prof. Urmila Barros, Dr. Meghanath Prabhu, Dr. Dharmendra Tiwari, and Dr. Sanika Samant** all past and present faculty members for their insightful questions, enthusiastic encouragement and useful critiques for this research work.

I thank the sponsors and the organizers of the Second South Asian Thematic School in Aquatic Microbial Ecology (**ASIAME2**) workshop for selecting me among the 25 candidates to attend the workshop, which was held at NIOT in Port Blair, India, in 2017. Funded by CNRS, DBT, IRD, CEFIPRA, Betty and MOORE Foundation, and NIOT. Late **Dr. Hervé Moreau**, former Research director of CNRS, France with whom I interacted more about challenges faced while studying viruses of microorganisms from extreme environments.

Federation of European Microbiological Societies (**FEMS**) for Early Career Scientist-FEMS grant for the VoM2022, at Guimarães, Portugal. The International Society for Viruses of Microorganisms-(**ISVM**) for Early Career Scientist-ISVM travel grant for the VoM2022, at Guimarães, Portugal. Department of Biotechnology, **DBT-CTEP** for the travel grant for VoM2022, Guimarães, Portugal. Grateful for the insightful conversations with **Prof. Curtis Suttle** on the future of RNA-virus studies and **Prof. Mart Krupovic** on archaeal viruses for their small interactions which gave me future insights into studying Viruses of microorganisms on a global platform. **Dr Simon Roux**, group lead of Viral genomics, JGI (Joint genome institute) tremendous support to clear my doubts about Viral-metagenomics during the VEGA 2021 symposium through webinar sessions.

Virus-girls: Dr. Judith Noronha and Dr. Preethi Poduval as we refer to ourselves as Virus girls-we have evolved along with our respective virus research. A significant portion of the research could not have been carried out without their moral support and constant guidance. During our tea breaks, we talk more about how obsessed we are with the terms "viral discovery," "bacteriophages," and "TEM images."

I would also like to thank *Dr. Abhishek Mishra, Dr. Meghanath Prabhu, Dr. Dharmendra Tiwari, Dr. Sanika Samant, Dr. Trupti Asolkar, and Dr. Manisha Tiwari* for their encouragement and suggestions. The valuable assistance provided by the non-teaching staff of our department, Late. Mr Martinho, Mr Serrao, Mr Tulsidas, Mr Ulhas, Mrs. Ruby, Mrs. Neelima, Mrs. Sonam, Mr. Parijat, Mr. Sameer, Mr. Rahul, Dr. Sandhya, Ms. Jaya and Mr. Ashish is greatly appreciated. I would like to thank them for their kind help, ever-willing

support and encouragement throughout the tenure of my research.

I would like to thank all my Ph.D colleagues; Dr. Imran, Dr. Surya, Dr. Amruta, Dr. Kirti, Dr. Michelle, Dr. Srijay, Dr. Ruchira, Pingal, Dr. Manasi, Dr. Sreekala, Perantho, Delicia, Priti, Noha, Deepti, Devika, Veda, Heytika, and Elaine for their practical suggestions, helpful advice, unwavering support and motivation during the entire tenure of this research. My earnest thanks to Dr. Samantha Fernandes with whom I have shared moments of anxiety as well as great excitement. They were of immense support in deliberating over my problems and findings and their presence was very important in a process that is often filled with tremendous solitude. I also want to express my gratitude to the former M.Sc. students Srikanth, Sania, Trupti, Devika, Tushar, Sakshi, and Lakshmi with whom I worked on research projects and who in some way contributed to expanding my knowledge.

I thank Dr. *Srikant Mutnuri* for the ultracentrifugation facility at the Dept. of Biological Sciences, BITS Pilani Goa Campus. Dr. *Meghanath Prabhu* (Asst professor SBSB, Goa University) for helping me with ultracentrifugation during my initial days of this research. Special thanks to Dr. *Manguesh* who provided the facility flow cytometry at NIO, Goa. Thanks to Dr. Amara for also helping me out with flow cytometry analysis. I wish to express my gratitude to SAIF-IIT, Bombay for providing the facility for elemental analysis and AIIMS, New Delhi and SAIF-CUSAT, Cochin for the Transmission electron microscope imaging studies.

I would like to offer my sincere gratitude to my beloved parents **Late. Joseph Macdonald Faria** and **Stella Faria**, have played an important role in moulding me into the person that I am today. Daddy and Mummy- my in-laws, **Prof. Marian Pinheiro** and **Joyce Pinheiro**, deserve the most profound gratitude for their unending love, support, and aid throughout my Ph.D journey, including during my pregnancy. The constant love and care you provided for your Grandson when I was away at work was simply beyond words. Numerous thanks to **Rommel Pinheiro**, my husband, and our son **Ethan**, who have been my strength and have provided me with unwavering support during the period of my PhD. My twin sister **Ricola Mendez**, who always motivated me since childhood to take up research. My sisters-in-law, brothers-in-law, nieces and nephew for being my motivation and for supporting me spiritually throughout writing this thesis and my life in general. They have always been a happy distraction to rest my mind outside of my research.

I remember all my teachers during graduation and post-graduation, especially **Late. Prof. PC Sebastian** who had motivated me into scientific research.

Special thanks to my friends **Roy, Munna, and Gaurav Pandit** for driving me down to the hot springs and helping me to collect the water samples for my research work. I am also grateful to all those who have directly or indirectly helped me through their prayers and good wishes for the completion of my research work.

Nicola Faria

Dedicated to Dada.....

***In loving memory of my father, who was unable to
witness this incredible journey***

Joseph Macdonald Faria

Contents

List of Tables	xii
List of Figures	xiv
List of Abbreviations.....	xviii
CHAPTER 1	1
Introduction and Research Objectives.....	1
CHAPTER 2	8
Review of literature.....	8
CHAPTER 3	28
Isolation and identification of culturable bacteria from hot springs.	28
3.1 Materials:.....	30
3.2 Methods:.....	30
3.2.1 Study Area and Sampling.....	30
3.2.2 Physicochemical analysis of the hot spring water sample	32
3.2.3 Isolation of culturable bacteria from hot spring	32
3.2.4 Filtration and concentration of water sample for virus enrichment	33
3.2.5 Preliminary screening for bacteriophages	34
3.2.6 Morphological and biochemical analysis.....	36
3.2.7 Molecular identification and 16S rRNA gene sequencing.....	36
3.2.8 Phylogenetic analysis	38
3.3 Results and Discussion.....	39
3.3.1 Physicochemical Analysis of the hot springs	39
3.3.2 Isolation of culturable bacteria from hot spring	41
3.3.3 Bacteria for isolation of bacteriophage	46
3.3 Conclusion.....	55

CHAPTER 4	56
Isolation of bacteriophage from hot springs	56
4.1 Materials:.....	58
4.2 Methods	58
4.2.1 Lytic bacteriophage of a thermophilic bacterium <i>Meiothermus</i> sp. JE2.....	58
4.2.2 Isolation of Temperate <i>Pseudomonas</i> phage.....	61
4.3 Results and Discussion.....	63
4.3.1 Lytic bacteriophage of a thermophilic bacterium <i>Meiothermus</i> sp.JE2	63
4.3.2 Isolation of temperate <i>Pseudomonas</i> bacteriophage.....	66
4.4 Conclusion:.....	70
CHAPTER 5	71
Whole genome sequencing of a potential lysogenic bacterium.	71
5.1 Materials:.....	74
5.2 Methods:.....	74
5.2.1 Bacterial strain VB3 sample.....	74
5.2.2 Genomic DNA extraction and 16S rRNA gene sequencing.....	74
5.2.3 De-Novo whole genome sequencing.....	75
5.3 Results and Discussion.....	77
5.3.1 Bacterial selection.....	77
5.3.2 16S rRNA gene sequencing.....	77
5.3.3 De Novo whole genome sequencing and annotation.....	77
5.3.4 Prediction of Prophage and Preliminary induction assay using Mitomycin C	92
5.4 Conclusion:.....	97

CHAPTER 6	99
Characterization of bacteriophage for a specific bacterial strain.	99
6.1 Materials:.....	102
6.2 Methods:	102
6.2.1 Bacterial culture host organism and maintenance:.....	102
6.2.2 Prophage induction and purification of prophages	102
6.2.3 Scanning electron Microscopy (SEM) image of strain VB3.....	103
6.2.4 Determination of host range specificity for bacteriophage after induction	103
6.2.5 Large Scale prophage induction.....	103
6.2.6 Concentration of phage particle using PEG 8000	104
6.2.7 CsCl density gradient ultracentrifugation.	104
6.2.8 Transmission Electron Microscopy.....	104
6.2.9 Phage DNA isolation.....	105
6.2.10 Genome sequencing of temperate phage of <i>Pseudomonas</i> sp. VB3	106
6.3 Results and Discussion.....	108
6.3.1 SEM imaging of <i>Pseudomonas</i> sp. strain VB3	108
6.3.2 Host range specificity for bacteriophage after induction	109
6.3.3 Purification of concentrated Phage VB3 particles using CsCl density gradient ultracentrifugation	110
6.3.4 Transmission electron Microscopy	110
6.3.5 Genome of phage vB_Psp_VB3.	112
6.4 Conclusion:.....	131
CHAPTER 7	132
Metavirome analysis of two hot spring sediments from the West Coast geothermal province (WCGP), Maharashtra, India.	132
7.1 Materials:.....	135
7.2 Methods:	135
7.2.1 Sample collection	135

7.2.2	Physicochemical analysis of hot spring water sample	138
7.2.3	Virus Enrichment from the sediment samples followed by transmission electron microscopy analysis.....	138
7.2.4	Concentration of viral particles from hot spring for metagenomic DNA isolation.....	139
7.2.5	Metagenomic Viral DNA extraction from sediment.....	139
7.2.6	NGS Library preparation and Illumina sequencing.	140
7.2.7	Metagenomic sequence assembly and annotation:.....	141
7.2.8	Taxonomic Annotation of Metagenome	141
7.2.9	Taxonomical classification and abundance.....	141
7.2.10	Prediction of viral sequences.....	142
7.2.11	Viral clustering and classification	142
7.2.12	Viral contig prediction and AMG prediction	142
7.2.13	Identification, annotation, and classification of predicted viral genomes.	143
7.2.14	Genome analysis of the predicted complete circular contig.	143
7.3	Results and discussion.....	144
7.3.1	Physicochemical analysis of Rajwadi and Tural hot spring water.....	144
7.3.2	Virus Enrichment from Sediment Samples Followed by Transmission Electron Microscopy	146
7.3.3	Shotgun metavirome sequencing	148
7.3.4	Microbial Community Composition of Rajwadi and Tural Metagenome.....	149
7.3.5	Viral taxonomy classification and abundance analysis by Kraken Virus database	151
7.3.6	Viral prediction, clustering and curation.....	155
7.3.7	Viral clustering	156
7.3.8	VIBRANT prediction of Auxiliary Metabolic Genes (AMG)	161
7.3.9	VIP tree proteomic tree analysis of predicted viral contigs	165

7.3.10	Quality check and geNomad score prediction.....	168
7.3.11	Sequence similarity with ReSeq viral genomes and Uncultivated viral genomes.....	169
7.4	Conclusion.....	171
SUMMARY AND CONCLUSION.....		172
BIBLIOGRAPHY		175
APPENDIX		
PRESENTATIONS AT CONFERENCES		
PUBLICATIONS		

List of Tables

Chapter 2

Table 2. 1 Thermophilic bacteriophages from hot springs.....	12
--	----

Table 2. 2: Studies in Indian hot springs metagenomics and viral metagenomics.....	27
--	----

Chapter 3

Table 3. 1: Sample collection sites.	31
---	----

Table 3. 2: Primers used for 16S rRNA sequencing.	37
--	----

Table 3. 3: PCR reaction mixture	37
--	----

Table 3. 4: Physicochemical parameters of hot springs from the Himalayan and West Coast geothermal province of India.	40
---	----

Table 3. 5: Host bacterial cultures that were selected for identification based on preliminary screening assay for bacteriophage isolation.....	42
--	----

Table 3. 6: 16S rRNA sequencing and Identification of bacterial hosts, which were positive for preliminary screening of bacteriophages.	44
---	----

Table 3. 7: Morphological and biochemical characteristics of selected bacterial strain JE2	47
---	----

Table 3. 8: Characteristics of <i>Meiothermus</i> sp. strain JE2 compared with other reported <i>Meiothermus</i> species isolated from Indian hot springs.	50
--	----

Table 3. 9: Morphological and biochemical characteristics of the selected isolate VB3.	52
---	----

Chapter 5

Table 5. 1: Basic statistics of the draft genome of <i>Pseudomonas</i> sp. strain VB3.	78
---	----

Table 5. 2: General information of <i>Pseudomonas</i> sp. strain VB3, according to the recommendation of Minimum Information about any (X) sequence (MIxS) standard checklist.	79
---	----

Table 5. 3: The digital DNA-DNA hybridization (dDDH) scores for <i>Pseudomonas</i> sp.	
--	--

strain VB3 genome and other related organisms at the Type Strain genome server...	83
Table 5. 4: Secondary metabolite biosynthetic gene clusters in <i>Pseudomonas</i> sp. strain VB3 as predicted by antiSMASH 6.1.1.....	86
Table 5. 5: Genes annotated by GC type in various databases.....	88
Table 5. 6: CRISPR-Cas Finder analysis of <i>Pseudomonas</i> sp. VB3 genome.	91
Table 5. 7: Putative prophage DNA sequences detected using PHAge Search Tool Enhanced Release (PHASTER).	96
Chapter 6	
Table 6. 1: Host range analysis for temperate <i>Pseudomonas</i> phage.	109
Table 6. 2: Similarity with NCBI-Virus Ref seq complete genome.....	113
Table 6. 3: Genome annotation bacteriophage vB_Psp_VB3.....	115
Chapter 7	
Table 7. 1: Physicochemical parameters of Rajwadi and Tural hot spring, Maharashtra.	145
Table 7. 2: Basic read statistics obtained from Illumina sequencing.	148
Table 7. 3: After trimming and metaSPAdes assembly statistics using metaQUAST.....	149
Table 7. 4: Total metagenome contigs from Rajwadi and Tural sediment samples classified by Kaiju analysis.	151
Table 7. 5: Viral Taxonomical Annotation of viral contigs MVRS1 and MVTS1 using Kraken Virus database.....	153
Table 7. 6: Summary of the number of predicted viral contigs identified using VirSorter	155
Table 7. 7: VIBRANT predicted whole viral circular genomes in MVRS1.	161
Table 7. 8: VIBRANT predicted whole viral circular genomes in MVTS1.....	162
Table 7. 9: CheckV analysis of two complete circular contigs using geNomad on the NMDC-EDGE platform.....	169
Table 7. 10: Sequence similarity of Predicted Viral contig NODE_171 in NCBI virus database.	170

List of Figures

Chapter 1

Figure 1. 1 Geothermal provinces of India.....	6
--	---

Chapter 3

Figure 3. 1: Workflow for the isolation of culturable bacteria from hot springs.	29
Figure 3. 2: Map showing the location of Indian hot springs used in the present study.	31
Figure 3. 3: Water samples collected from Indian Hot springs.	32
Figure 3. 4: Distribution of culturable bacterial isolates obtained from seven different Indian hot springs.	41
Figure 3. 5: Geographical location of Rajwadi Hot Spring and sampling sites.	46
Figure 3. 6: Growth of thermophilic <i>Meiothermus</i> sp. strain JE2 at different temperatures...	48
Figure 3. 7: Purified 16S rRNA PCR product of approximately 1.5 kb	49
Figure 3. 8 : Molecular phylogenetic analysis of <i>Meiothermus</i> sp. strain JE2 by Maximum likelihood method.	49
Figure 3. 9 : Geographical location of Vashisht Hot Spring and sampling sites.....	51
Figure 3. 10: Growth of bacterial strain VB3 at different temperatures.....	53
Figure 3. 11: Growth of bacterial strain VB3 at different pH.	53
Figure 3. 12: Neighbor-joining phylogenetic tree based on 16S rRNA gene sequencing of <i>Pseudomonas</i> sp. strain VB3 (with a Bootstrap value of 1000 replication).....	54

Chapter 4

Figure 4. 1: Workflow for isolation of bacteriophage from hot spring.	57
Figure 4. 2: vB_Msp_JE2 clear plaques detected by the double-layer overlay method using the red-pigmented thermophilic <i>Meiothermus</i> sp. strain JE2 as host.	63
Figure 4. 3: One-step growth curve of phage vB_Msp_JE2.	64
Figure 4. 4 : Thermostability of bacteriophage vB_Msp_JE2.....	65

Figure 4. 5: pH sensitivity of bacteriophage vB_Msp_JE2.....	65
Figure 4. 6: Prophage induction of bacterial strain VB3 with and without Mitomycin-C..	66
Figure 4. 7: Spot test after induction with Mitomycin-C on the lawn of <i>Pseudomonas</i> sp. VB3.....	67
Figure 4. 8: Preliminary enumeration of phage particles after prophage induction with mitomycin C and concentrated through PEG 8000.	68
Figure 4. 9: Preliminary enumeration of phage particles after mitomycin-C treatment and concentrated through PEG 8000.....	69
Chapter 5	
Figure 5. 1: Workflow of the WGS of a potential lysogenic bacterium.....	73
Figure 5. 2: RAST annotation of <i>Pseudomonas</i> sp. VB3.....	80
Figure 5. 3: Circular representation of <i>Pseudomonas</i> sp. strain VB3 contigs using Proksee	81
Figure 5. 4: Phylogenomic Tree based on <i>Pseudomonas</i> sp. strain VB3 genome sequences in TYGS (Type strain Genome server).	84
Figure 5. 5: CARD analysis of <i>Pseudomonas</i> sp. strain VB3	89
Figure 5. 6: COG functional analysis of <i>Pseudomonas</i> sp. strain VB3.....	90
Figure 5. 7: KEGG analysis of <i>Pseudomonas</i> sp. strain VB3.	90
Figure 5. 8 : PHASTER analysis of <i>Pseudomonas</i> sp. Strain VB3.....	92
Figure 5. 9: Prophage region 1 detected by PHASTER web server tool.....	93
Figure 5. 10: Prophage region 2 detected by PHASTER web server tool.....	94
Chapter 6	
Figure 6. 1: Workflow of the characterization of bacteriophage from a specific bacteria <i>Pseudomonas</i> sp. strain VB3.....	101
Figure 6.2: SEM image of <i>Pseudomonas</i> sp. strain VB3	108
Figure 6. 3: CsCl ultracentrifugation concentration of VB3 phage particles after treatment with mitomycin-C.....	110

Figure 6. 4: TEM Images of induced temperate bacteriophage vB_Psp_VB3.	111
Figure 6. 5: Phage VB3 DNA isolated from <i>Pseudomonas</i> sp. VB3	112
Figure 6. 6: Genome of phage vB_Psp_VB3 in PHASTEST linear genome viewer.....	123
Figure 6. 7: Viral Proteomic tree VipTree v1.9.1 of phage vB_Psp_VB3.....	124
Figure 6. 8: VIRFAM analysis of the phage phage vB_Psp_VB3.....	126
Figure 6. 9: VIRDIC analysis heat map showing the intergenomic distance of the phage vB_Psp_VB3 based on the genome sequence.	127
Figure 6. 10: Phylogeny tree based on similarity with NCBI virus complete genomes based on the protein clusters analysis in VirClust (Moraru, 2023). The values above the branch	129
Figure 6. 11: The protein cluster-based intergenomic distance of phage vB_Psp_VB3 based on VirClust analysis.	129
 Chapter 7	
Figure 7.1: Workflow of Metaviromic (viral metagenomic) analysis of two hot springs from the West Coast geothermal province (WCGP) in Maharashtra, India.....	134
Figure 7. 2: Map showing Rajwadi and Tural hot springs belonging to the WCGP.....	136
Figure 7. 3: Sampling sites of Hot springs.	137
Figure 7. 4: TEM images of Rajwadi and Tural sediment sample.....	147
Figure 7. 5: Overview of output from Kaiju analysis.....	150
Figure 7. 6. Viral community composition in MVRs1 (Kraken virus database).....	154
Figure 7. 7: Viral community composition in MVTS1 (Kraken virus database).....	154
Figure 7. 8: Protein clustering network of virus genomes recovered from Rajwadi sediment sample MVRs1 compared to previously described viruses.	157
Figure 7. 9: Protein clustering network of virus genomes recovered from Tural sediment sample MVTS1 compared to previously described viruses.	159
Figure 7. 10: AMG metabolism and Category.	164

Figure 7. 11: Viral proteomic tree using VipTree v. 1.9.1. for MVRs1 complete circular contigs (5)	166
Figure 7. 12: Viral proteomic tree using VipTree v. 1.9.1. for MVTs1 complete circular contigs (10).	167
Figure 7. 13: Graphical comparison of the homology MVRs1 complete circular viral contig.	168

List of Abbreviations

AAI:	average amino acid identity
AMG:	Auxiliary metabolic genes
ANI:	Average nucleotide identity
BLAST	Basic Local Alignment Search Tool
bp:	base pair
BM:	Basal medium
CRISPR:	clustered regularly interspaced short palindromic repeats
CsCl:	caesium chloride or cesium chloride
CTAB:	cetyltrimethylammonium bromide
DNA:	deoxyribonucleic acid
dsDNA:	double-stranded DNA
dsRNA:	double-stranded RNA
DTRs:	direct terminal repeats
EDTA:	ethylenediaminetetraacetic acid
EFM:	epifluorescence microscopy
EGTA:	ethylene glycol-bis(β -aminoethyl)-N,N,N',N'-tetra acetic acid
EtOH:	ethanol
e.g.:	exempli gratia
et al.:	et alia
g:	gram or gramme
g/L:	Gram per litre
Gb:	giga bases
GC content:	guanine-cytosine content
gDNA:	genomic DNA
h:	Hour (s)
HMM:	Hidden Markov Model
HQ:	high quality
ICTV:	International Committee on Taxonomy of Viruses
IMG/VR:	Integrated Microbial Genome/Virus
kb or kbp:	kilo base pair

kDa: Kilodalton
KEGG: Kyoto Encyclopedia of Genes and Genomes
kV: kilovolts
L: litre or liter
LASL: linker amplification shotgun libraries
LSU-rRNA: large subunit rRNA
M: Molar
Mb: mega bases
MDA: multiple displacement amplification
ml: millilitre
mM: millimolar
MWCO: molecular weight cut-off
NASL: non-amplified shotgun library
NCBI: National Centre for Biotechnology Information
ng: nanogram
NGS: next-generation sequencing
nM: nanomolar
nm: nanometre
OD_{600 nm}: Optical density 600 nanometres
ONT: Oxford Nanopore Technologies
PBS: phosphate-buffered saline
PC: polycarbonate or protein cluster
PCR: polymerase chain reaction
PE: paired-end
PEG: polyethylene glycol
PES: polyethersulfone
PFU or pfu: plaque-forming unit
pH: potential of hydrogen or power of hydrogen
Phage: bacteriophage
PVDF: polyvinylidene fluoride or polyvinylidene difluoride
Q30: Phred quality score 30
QC: quality control

RNA: ribonucleic acid
rRNA: ribosomal RNA
SM buffer: sodium chloride-magnesium sulphate buffer
ssDNA: single-stranded DNA
ssRNA: single-stranded RNA
SYBR: Synergy Brands, Inc.
TE buffer: Tris-EDTA buffer
TEM: transmission electron microscopy
TFF: tangential flow filtration
Tris-HCl: Tris-hydrochloride
U: enzyme units
UpViG(s): uncultivated provirus genome(s)
UViG(s): uncultivated virus genome(s)
UV: Ultra violet
VIBRANT: Virus Identification By iteRative ANnoTation
VC: viral cluster
vConTACT: Viral CONTigs Automatic Clustering and Taxonomy
VLP(s): virus-like particle(s)
VOGDB: Virus Orthologous Groups Database
v/v: volume to volume
WGS: whole genome sequencing
w/o: without
μ : Micron(s)
μg : Micrograms
μg/μL: Micrograms/ microlitre
μL : Micro litre
μm : Micro meter
μmol : Micro mole
μmol/L : Micro moles per litre
Φ: phi, representing a bacteriophage or phage
% Percentage
°C Degrees Centigrade

CHAPTER 1

Introduction and Research Objectives

1.1 What are hot springs?

Geothermally heated groundwater emerges from the Earth's surface, forming terrestrial hot springs or geothermal springs. Terrestrial geothermal springs are widely distributed worldwide with different temperatures and pH ranges. From a biological perspective, the most extensively studied, well-known geothermal regions occur in Iceland, New Zealand, Japan, China, Italy, Indonesia, Central America, and the United States. Globally, among the reported hot springs, Yellowstone National Park in Wyoming (USA) is one of the best-studied hot spring (Zablocki et al., 2018).

1.2 Why study viruses?

Viruses are the smallest and most abundant entities in the world, transcending every other form of life on Earth (Bergh et al., 1989; Fuhrman, 1999; Suttle 2005; Graham et al., 2019). The number of viruses on Earth is approximately 10^{31} (Hendrix et al., 1999; Bar-On et al., 2018; Labbé et al., 2020). To conceptualize this staggering number, their current biomass is equivalent to the biomass of 75 million blue whales (approximately 200 million tons), and if their virions are arranged end to end, their total length will span 65 galaxies (Suttle, 2013; Simmonds et al., 2017). They play a significant role in biogeochemical cycles, horizontal gene transfer, and regulating host community diversity in aquatic ecosystems (Fuhrman, 1999; Kimura et al., 2008). Viruses are obligate intracellular parasites that probably infect all life forms (Bacteria, Archaea, and Eukarya) (Mougari et al., 2019). They are found in the world's deepest ocean to the most extreme environments and contribute significantly to the global biogeochemical cycles (Jover et al., 2014). Viral DNA were explored and found in almost all the biosphere, such as oceans, freshwater, soil, animal gut, and even mummies (Biagini et al., 2012; Dion et al., 2020; Duggan et al., 2016; Santiago-Rodriguez and Hollister, 2022) including extreme environments such as hot springs,

(Nagayoshi et al., 2016) hypersaline environments, (Demina et al., 2017) and hydrothermal vents (Liu et al., 2006).

Virologists, like all scientists, tend to view viruses through the lens of their position as disease-causing agents. While there is no doubt that viruses cause serious human, animal, and plant diseases, this narrow perspective overlooks the potential benefits they bring to their hosts.

1.3 Why explore hot spring viruses?

The natural settings of hot springs continue to be fascinating for studying the origin and evolution of life at higher temperatures. Hot springs are the most likely preliminary environment in which life emerged and evolved before mesophilic organisms (Weiss et al., 2016). Additionally, terrestrial geothermal locations are hotspots for biotechnological exploitation, necessitating the isolation and characterization of bacteriophage enzymes. Hot spring viruses are of great interest as they serve as a model system for understanding life's biochemistry and molecular biology at high temperatures. Hot springs are places where hot groundwater comes out from Earth. Hot springs have a unique environment that harbors a consortium of thermophilic bacteria and archaea. Terrestrial geothermal springs have been a prime location for exploring and characterizing archaeal viruses (Häring et al., 2005; Peng et al., 2012; Prangishvili et al., 2017). However, hot spring even comprises viruses infecting thermophilic bacteria (bacteriophages) and eukaryotic viruses (algal viruses). Bacteriophages that infect bacteria or archaea in hot springs are abundant, reaching a concentration of a million viruses per milliliter ($10^6/\text{mL}$) (Breitbart et al., 2004). Further, metagenomic studies have revealed hidden and new viral diversities in hot springs (Gudbergssdóttir et al., 2016; Zablocki et al., 2017; Castelán-Sánchez et al., 2020; Jarett et al., 2020), rekindling interest in the study of virus evolution and diversity.

1.4 Overview of Hot spring viruses worldwide

In the 1980s, Wolfram Zillig's laboratory initiated early studies on the hyperthermophile virus. A systematic screening of surface hot springs in Japan, Iceland, New Zealand, Italy, Russia, and the United States led to the isolation of an unprecedented diversity of new viruses (Rice et al., 2001; Rachel et al., 2002; Häring et al., 2005).

The hot spring viral research and their ecological significance is still in its infancy and least explored due to the difficulty in bacterial host cultivation and the respective viruses that infect them. Over 48 years ago, the first hot spring virus *Thermus* myovirus ϕ YS40 was partially characterized and was isolated from Japan's hot spring (Sakaki and Oshima, 1975). Afterwards, they reported the isolation of phage ϕ NS11, infecting the extreme acidophile *Bacillus acidocaldarius* TA6 (Sakaki and Oshima, 1976). Two later studies reported bacteriophage morphologies observed *in situ* (i.e., without enrichment) in hot springs in New Zealand (Patel et al., 1989) and Yellowstone National Park (Rachel et al., 2002).

A considerable amount of understanding has been gained about the archaeal viruses and their biological function at high temperatures (Rice et al., 2001; Prangishvili et al., 2013), and advances in thermophilic bacteriophages lagged. Limited diversity studies or isolations of these viruses or phages have been reported from lower temperature range hot springs (40–71°C). Amongst the three Yellowstone National Park springs (one neutral and two acidic pools), bacteriophages were detected only in the neutral spring (75–93 °C; pH 6) but absent in acidic springs (85 °C, pH 2), in which only archaeal virus morphologies were detected.

Although bacteriophages are abundant and the most diverse biological entities on Earth (Suttle, 2005), only a few archaeal viruses, thermophilic viruses, and eukaryotic viruses have been studied worldwide (Breitbart et al., 2004; Peng et al., 2012; Prangishvili, 2013; Prangishvili et al., 2017; Schoenfeld et al., 2008). These bacteriophages can influence bacterial communities, abundance, diversity, and evolution in almost every biological niche. Similarly, in geothermal environments, these phages play a significant role in biogeochemical cycles, ecology, and horizontal gene transfer (Nagayoshi et al., 2016).

Thermophilic phages have diverse morphologies (e.g., tailed, filamentous) and unique virion structures (e.g., highly long-tailed *Siphoviruses*). Genome sequences of hot springs viruses are scarcely reported (Zablocki et al., 2018). The Promega Cooperation had undertaken a single, most comprehensive effort to list the thermophages (Yu et al., 2006). Around 115 *Thermus* phages were isolated from seven culturable *Thermus* strains and were categorized based on virion morphology (Yu et al., 2006). A startling finding was that most viruses were categorized as *Tectiviridae* and 45% exhibited a head-tail shape (Order: *Caudovirales*).

1.5 Studying viruses in the era of metagenomics

The recent advance in metagenomic sequencing has revealed an astonishingly large virome in the biosphere. Most of the bacterial hosts from hot springs are uncultivable. Additionally, the lack of conserved regions in the viral genome makes hot spring viral research extremely challenging (Edwards and Rohwer, 2005; Pride and Schoenfeld, 2008). However, with the development of next-generation sequencing (NGS) technology in the last few decades, viral research has significantly expanded novel opportunities for characterizing viral diversity and tapping the richness of viral assembly in extreme environments such as hot springs (Bolduc et al., 2015; Zablocki et al., 2017; Castelán-Sánchez et al., 2020; Jarett et al., 2020; Guajardo-leiva et al., 2021).

Metagenomics has become a powerful tool for understanding viral diversity and variability since the first marine viral metagenomic research in 2002 (Breitbart et al., 2002). It can directly resolve uncultured environmental viral genomes (Angly et al., 2006; Culley et al., 2006; Hurwitz et al., 2013). Unravelling novel viruses from viral dark matter is possible through viral metagenomics (Roux et al., 2015). Investigating microbial communities in extreme environments has revealed hidden viruses that infect uncultivable bacterial hosts, proving that viruses are Earth's most prevalent biological entities (Krupovic et al., 2019). Thus, the extremophiles act as hosts for viral replication, and viruses balance population dynamics in extreme environments (Dávila-Ramos et al., 2019).

In the mid-21st century and continuing until today, the NGS and high-throughput sequencing (HTS) technology have advanced, and the cost of sequencing coming down several orders of magnitude, has tremendously increased viral metagenomic sequencing (Goodwin et al., 2016). Armed with new technology in computational biology, our understanding of viruses has tremendously increased, which led to the growth of large databases of the complete viral genome and viral protein (Paez-Espino et al., 2016, 2017; Roux et al., 2015; Roux et al., 2021). Our understanding of the global virosphere is reshaping with the genomes of uncultivated viruses (UViGs') found in shotgun sequencing datasets rather than being produced by laboratory virus cultivation. Since 2016, the IMG/VR database (genomes of cultivated and uncultivated viruses) has served as a leading resource in the field of viral ecology (Paez-Espino et al., 2017; Roux et al., 2021). The new version IMG/VR v.4 (<https://img.jgi.doe.gov/cgi-bin/vr/main.cgi>) has the most extensive UViGs

collection to date, and will undoubtedly reveal numerous novel insights into the ecological parameters influencing viral diversity, virus evolutionary history, and viral-encoded functional diversity (Camargo et al., 2023).

The International Committee on Taxonomy of Viruses (ICTV), the foremost body, recently opted in 2017 to add virome-derived genomes to its database because the number of viral sequences of uncultivated viruses has surpassed that of the cultured viral sequences (Simmonds et al., 2017). ICTV recently revised the viral taxonomic ranking structure in 2020, increasing the number of hierarchical ranks to 15 to classify novel viruses and allowing the use of genomic sequence data in the absence of Electron microscopy images (ICTV, 2020).

Recently Bacterial Viruses Subcommittee (BVS) of ICTV has moved towards a more coherent and unified system of virus classification based on genome-level relationships. Substantial changes were introduced: the paraphyletic morphological families *Podoviridae*, *Siphoviridae*, and *Myoviridae* as well as the order *Caudovirales* were abolished, and a binomial system of nomenclature for species was established. In addition, one order, 22 families, 30 subfamilies, 321 genera, and 862 species were newly created, promoted, or moved (Turner et al., 2023).

1.6 A scenario of viruses studied in Indian hot springs.

The Geological Survey of India (GSI) has identified 400 hot springs in India (Chandrasekharam, 2000; Chandrasekhar and Chandrasekharam, 2010; Singh et al., 2016). India has seven geothermal provinces: Himalaya, Sohana, Cambay, West Coast, Son-Narmada Tapi (SONATA), Godavari, and Mahanandi (Figure 1.1). In India, geothermal springs are routinely used for bathing and balneotherapy and, occasionally as an energy source for cooking. Hot springs in India are grouped into three categories: moderately hot (37–42 °C), hot (50–70 °C), and very hot (>70 °C). The highest temperature has been recorded for hot springs at Soldhar (Arya et al., 2015) and Manikaran, where the water attained nearly boiling conditions (90–98°C) (Sangwan et al., 2015). Hot springs in India have neutral to alkaline pH (7–7.5), and the microbial community is different from those reported in acidic hot springs environments (Kanokratana et al., 2004; Mathur et al., 2007; Aditiawati et al., 2009).

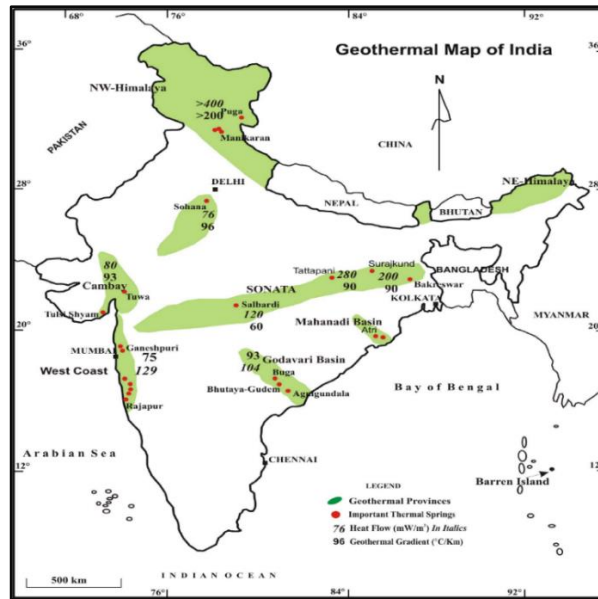


Figure 1. 1 Geothermal provinces of India.

(Image source: Chandrasekhar and Chandrasekharam, 2010)

Despite a wide distribution of hot springs throughout India, limited studies have been conducted on the microbiological analysis of these hot springs, especially viruses. There are minimal reports of viruses from Indian hot springs. Microbial diversity studies in Indian hot springs focus mainly on Himalayan geothermal provinces at higher temperatures and altitudes. There are only two hot springs viral diversity metagenomic studies reported from India belonging to the Himalayan geothermal province (Das et al., 2020; Sharma et al., 2018).

Most published articles report microbiological isolation of thermophilic bacteria or metagenomic bacterial diversity studies that can be exploited for biotechnological potential (Poddar and Das, 2018; Vijayan et al., 2023). Only two metagenomics reports on viral diversity studies have been reported in India from Manikaran (Sharma et al., 2018) and Sikkim hot springs (Das et al., 2020), belonging to the Himalayan geothermal province. There are no reports on the isolation of culturable bacteria which serve as potential hosts for bacteriophage isolation from other Indian hot springs. Also, Maharashtra is one such bio-diverse state of India that harbors 60 diverse alkaline hot springs in the West Coast Geothermal province of India, and it is relatively unexplored for thermophilic bacteria and viruses.

Research Objectives

The following research objectives were proposed based on the above scenario of ongoing research on hot spring viruses being conducted globally.

The proposal aims to isolate culturable bacteria from Indian hot springs and to use and identify potential bacterial hosts for infecting possible viruses from hot springs. The water from the hot springs would be screened for isolating bacteriophage infecting their respective bacterial host. One goal would be to develop a specific virus-host system in the laboratory and explore possible temperate phages. Additionally, to expand our knowledge and overcome the challenges of conventional isolation techniques, we attempted to illustrate the metavirome analysis of two hot spring sediments belonging to the West Coast geothermal province.

The objectives of the study were as follows:

1. Isolation and identification of culturable bacteria from hot springs.
2. Isolation of bacteriophage from hot springs.
3. Whole genome sequencing of a potential lysogenic bacterium.
4. Characterization of bacteriophage for a specific bacterial strain.
5. Metaviromic analysis of two hot spring sediments from the West Coast geothermal province in Maharashtra, India.

CHAPTER 2

Review of literature

This chapter comprises bibliographic information concerning the proposed research objectives. A brief description of the thermophages isolated from hot springs in different parts of the world so far, and information about Indian hot springs. Further, followed by a review of viral metagenomics.

Over the past three years, in the pandemic era, epidemiologists have been working around the clock on how to fight COVID-19 and the healthcare system taking care of the sick due to this infectious disease caused by the SARS COVID-19 virus. The world has come to unprecedented measures to combat the SARS-COVID-19 virus. This is where Yellowstone Hot Springs National Park comes into the picture. We would not have been able to test for COVID-19 without microbes from the hot spring. The thermostable Taq polymerase enzyme isolated from *Thermus aquaticus* YT-1 from Yellowstone National Park is used in Polymerase chain reaction (PCR) reaction.

Early Virology

Van Leeuwenhoek's descriptions of tiny, microscopic organisms astounded people in the seventeenth century, but a significant portion of the microbial world was hidden from his vision. Ivanovski and Beijerinck both asserted in the 1890s that tobacco plants were infected by an organism that was smaller than bacteria (Horzinek, 1997). Twort and d'Herelle later separately reported infectious agents that could destroy bacteria and gave them the name "bacteriophages" in 1915 and 1917, respectively (Keen, 2015). The existence of such filterable infectious agents, as well as whether they were liquid or particulate in nature and whether they were living or not, were hotly contested topics in the years that followed (Van Helvoort, 1994). The need to view infectious virus particles served as one of

the primary drivers for the creation of electron microscopy (EM) in the 1930s (Kruger et al., 2000).

EM's ability to provide thorough descriptions fuelled efforts to classify viruses into groups with similar structures. Eisenstark, for example, provided morphological descriptions, including EM pictures, of the 111 known bacteriophages in 1967; he was followed by Ackermann, who evaluated phages as a group across several decades (Ackermann, 2007; 2009).

EM was utilized in the 1980s to discover that the quantity and diversity of bacteriophages in saltwater samples were exceptionally high (Torrella and Morita, 1979). Following these revelations, researchers found viruses in marine samples from all across the world, and the numbers found, both in quantity and diversity, were orders of magnitude higher than previously anticipated (Bergh et al., 1989). Bacteriophages were the most prevalent, with millions of bacteriophages found in one milliliter of water (Fuhrman, 1999; Suttle, 1993, 2005)

What are viruses?

Viral particles are not only far smaller than most cellular life, but they are also much simpler, consisting of a short piece of DNA or RNA surrounded by a layer of protein known as a capsid and, in some viruses, a lipid membrane known as the viral envelope. Viruses do not have their own metabolism, thus in order to replicate, they must infect a host cell and exploit its metabolic system, followed by the release of new copies of the virus into the environment. Viruses have one of three lifestyles:

(a) **Lytic**: in which host infection is followed directly by replication and the release of new viral particles, killing the host;

(b) **Chronic**: in which viral replication occurs over long periods of time and the release of new virions is less than lethal to the host; or

(c) **Lysogenic**: in which the viral genome is integrated into the host's and remains dormant for a period of time before being activated and lysing the host (Howard-Varona et al., 2017).

Bacteriophages

Bacteriophages, also referred to as phages, are bacteria-infecting viruses. Bacteriophages are thought to be the most abundant entities (Clokier et al., 2011; Keen, 2015; Salmond and Fineran, 2015), with an estimated population of more than 10^{31} (Hendrix et al., 1999; Bar-On et al., 2018; Labbé et al., 2020). They can be identified and isolated from the soil, oceans and lakes, humans and animals, and hot springs and hydrothermal vents, among other places where their host bacteria are found. The temperature at which their host bacteria develop divides phages into three types.

a) Mesophilic: Phages that infect bacteria with an optimal growth temperature of 37°C are known as mesophilic phages.

b) Thermophilic phages: phages that infect bacteria that thrive at temperatures as high as 75°C.

c) Hyperthermophilic phages: phages that infect bacteria that thrive at high temperatures >75°C. Thermophilic and hyperthermophilic phages are usually referred to as thermophages. In general, phages have numerous applications in human, animal, and plant health, as well as food safety (Liu et al., 2023).

Overview of thermophages

Thermophages can be isolated from hot springs, including land-based (Matsushita and Yanase, 2009; Yu et al., 2006) and undersea alkaline freshwater hot springs, as well as coastal springs, deep-sea hydrothermal vents (Liu et al., 2006), compost piles, dairy settings, soil, industrial hot waters, and so on. The selectivity for the bacterial host they infect, genome/particle size, shape, and structure of thermophages change depending on where they are isolated (Table 2. 1).

Some thermophages can infect a large range of thermophilic bacterial species, whilst others can only infect a subset of thermophilic bacteria. Phages P23-45 and P74-26B, for example, can infect *Thermus aquaticus*, *T. flavus*, *T. thermophilus* HB8, and *Thermus* sp. ATCC 31,674 (Yu et al., 2006), although the filamentous phage OH3 appears to infect only the HB8 strain of *T. thermophilus* (and not other strains like HB27, AT62, TMY, etc.) (Nagayoshi et al., 2016).

Some thermophage particles have an icosahedral capsid (oval head without a tail and

are called tail-less phages) (Pawlowski et al., 2014), while others have an oval with a tail (tailed phages) (Cheepudom et al., 2015). Some thermophages, in addition to tailless and tailed thermophages, have a filamentous form (Nagayoshi et al., 2016; Pederson et al., 2001). Their genome length range from 5.7 kb to 152.4 kb (Hjorleifsdottir et al., 2014; Nagayoshi et al., 2016; Tamakoshi et al., 2011), and virion particles range in size from 60-130 nm broad (for oval-shaped thermophages) to 910 nm long (for filamentous thermophages) (Pederson et al., 2001). Their genomes can be linear, circular, single- or double-stranded. GC content of the genome ranges from 32 to 68%. The majority of thermophages infect and multiply in bacteria with optimal temperatures ranging from 55 to 70°C (Table 2. 1).

The small genome size of 6.5 kb (no genome sequence available) of filamentous thermophage (PH75) is the result of evolutionary adaptation to efficiently utilize host resources and replicate within the specific bacteria in a geothermal environment. The inter-subunit interactions within the capsid and genome organization of PH75 contribute to the enhanced thermostability when compared to mesophilic filamentous bacteriophages (Overman et al., 2005). The genome of ϕ OH3 has eight open reading frames in the 5.7 kb circular genome. The genome was highly novel, with most ORFs bearing no homology to other filamentous phages except only for ORF 5 which had 60% amino acid identity with the major coat protein of PH75. Interestingly, this was only observed in thermophilic phages and not in mesophilic ones suggesting that there might be a deep evolutionary connection between hyperthermostability determinants (Tsuboi et al., 2005).

Table 2. 1 Thermophilic bacteriophages from hot springs.

Phage	Phage Genome	Genome Size (kb)	Morphology	Size of Phage	Isolation Origin	Host Bacteria Isolated from	Temperature of Host Bacteria (°C)	pH of Host Bacteria	References
P74–26	dsDNA	83.32	Oval (capsid) head and tail	Head: 824Å Tail: ~800 nm	Kamchatka Peninsula Hot Springs, Russia	<i>Thermus thermophilus</i>	65	6-7	(Minakhin et al., 2008)
φNS11	dsDNA	57	Oval capsid with spikes-like structures	Diameter: 60–75 nm	Beppu Hot Springs, Japan	<i>A. acidocaldarius</i> strain TA6	60-65	3.5	(Yu et al., 2006)
MMP17	dsDNA	33.5 to 39.5	Oval capsid and tail	Head: 42 nm Tail: 120 nm long and 17 nm wide	Eryuan Hot Spring, China	<i>Meiothermus</i> sp.	55-60	6-7	(Lin et al., 2011)
MMP7	dsDNA	32,864 bp	Hexagonal head and tail	134 nm length 14 nm width head (~48 nm in diameter).	Tengchong Hot Spring, China	<i>Meiothermus</i> sp. TC7	55	6-7	(Wang et al., 2020)

Table 2.1: Continued.

Phage	Phage Genome	Genome Size (kb)	Morphology	Size of Phage	Isolation Origin	Host Bacteria Isolated from	Temperature of Host Bacteria (°C)	pH of Host Bacteria	References
P23-77	dsDNA	17.04	Icosahedral capsid with spikes on vertices, and an internal lipid membrane, no tail	Diameter: 78nm	Alkaline hot springs, New Zealand	<i>T. thermophilus</i>	70	6-7	(Jaatinen et al., 2008)
ΦIN93	dsDNA	19.6	Icosahedral	Diameter 130nm	Hot Spring, Japan	<i>T. aquaticus</i>	70	6-7	(Matsushita and Yanase, 2009)
φTMA	dsDNA	151.5	Icosahedral non-contractile tail	Head: 125 nm Tail: 175 nm	Atagawa Hot Spring, Japan	<i>T.thermophilus</i>	65	6-7	(Tamakoshi et al., 2011)
φYS40	dsDNA	152.4	Oval capsid and contractile tail	Head:125nm Tail: 178nm long 27nm wide	Atagawa Hot Spring, Japan	<i>T.thermophilus</i>	65	6-7	(Naryshkina et al., 2006)
φOH3	ssDNA	5.7	Filamentous	830 nm long and 8 nm wide	Obama Hot Spring, Japan	<i>T.thermophilus</i> HB8A.	70	7	(Nagayoshi et al., 2016)
PH75	ssDNA	6.5	Filamentous	Length 910±17 nm	Hot spring New Zealand	<i>T.thermophilus</i>	70	8.3	(Pederson et al., 2001; Yu et al., 2006)

Bacteriophages Isolation studies

Culture-based bacteriophage studies

Phages in nature have been isolated by using plaque assay and liquid medium-based-dilution culturing techniques. This classical approach to studying phage is to isolate them and perform morphological, physiological, and host range experiments.

Plaque based method

The widely used method is the plaque-based method (Adams, 1959; Kropinski et al., 2009), which is based on the direct visualization of clear zones on the lawn of bacteria formed due to cell lysis of lytic phages. This classical approach has allowed us to isolate and study numerous phages from various niches. Phage concentrations from environmental sources, filtered with 0.2 μm pore-sized filters, can either be put on the solidified bacterial lawn plate or added to the molten agar to thoroughly mix with the bacterial host. Numerous lytic phages have been isolated by this technique of plaque assays.

Liquid medium-based isolation technique

The liquid medium-based isolation technique is used for bacteria that do not form colonies on agar solid plates and can only be isolated in a liquid medium (Adams 1959). Compared to the plaque analysis, the phage growth cannot be measured as easily as plaque assay, instead indirectly the bacterial growth is monitored by recording optical densities (OD) or flow cytometry (Zhao et al., 2013). As lytic phages multiply, bacterial cells start to be lysed, cell densities drop, and OD or cell density values are decreased as a result. The majority of their phages have been isolated using the liquid medium-based strategy because slow-growing bacteria in oligotrophic habitats, such as open oceans, pristine lakes, and groundwater, frequently do not form colonies on agar plates. Particularly, dilution-to-extinction cultivation techniques have made it easier to isolate host-phage systems using bench-top flow cytometers and multi-well microtiter plates (Buchholz et al., 2021; Kim et al., 2020).

Prophage studies

Prophage induction with Ultraviolet radiation, poly-aromatic hydrocarbon (PAH), and Mitomycin-C are the most common physical and chemical agents used for prophage induction

In both the "dormant" and "active" stages, prophages can impact both their host and the surrounding microbial ecosystems.

By modifying gene expression patterns, causing DNA transfer or recombination events, and offering virulence features, prophages can cause physiological changes in the host while still in the dormant stage (Banks et al., 2003; Canchaya et al., 2003; Dedrick et al., 2017; Eggers et al., 2016; Wendling et al., 2021). For instance, the presence of integrated prophage sequences is necessary for some strains of *Staphylococcus aureus* to be harmful (Bae et al., 2006).

Phage lysis has a substantial impact on microbial communities during the active stage by releasing vital nutrients, including carbon, nitrogen, and sulfur (Anantharaman et al., 2014; Emerson et al., 2018; Suttle, 2007; Waldbauer et al., 2019). By varying community structures and enhancing niche options, bacterial population lysis also affects entire microbiomes (Gobler et al., 1997; Zimmerman et al., 2020). For instance, the "Kill the Winner" model of virus population growth contends that dominant bacterial populations are more vulnerable to phage predation, allowing less abundant taxa to spread as the dominant populations are lysed. Despite the impact of phage lysis on microbial communities, it is unclear how many phages enter the lytic cycle.

Advances in high-throughput metagenomic sequencing have accelerated the capacity to find a wide variety of lytic and lysogenic phage genomes in addition to conventional methods like phage isolation. VIBRANT (Kieft et al., 2020), VirSorter (Roux et al., 2015), PHASTER (Arndt et al., 2016), and Prophage Hunter (Song et al., 2019) are examples of recently created software tools that have enabled reliable characterization of phages in both isolation and metagenomic assembly genomes. Prophages *in-situ* state of being in the lysogenic or lytic stage of infection cannot be inferred from just identifying their genomic sequences.

These software tools have so far been able to explore the potential consequences of prophages in great detail and estimate the overall diversity of prophages in nature. However,

determining just the prophage genome sequences does not put their *in-situ* infection stage of lysogenic or lytic into context. Such information is crucial because it identifies the prophage or phage populations that are actively causing lysis events in a microbial community.

Enumeration studies of virus

Transmission electron microscopy

Viruses are nano-scale-sized particles and transmission electron microscopy has proven to be a powerful tool for the investigation of bacteriophages. TEM is valuable in the elucidation of bacteriophage attachment and replication studies inside the host bacterial cell. Smith and Krueger (1954) were the first to morphologically characterize a phage isolated from the estuarine environment. TEM involves two types of staining: positive and negative staining using dyes that are specific to the technique of staining (Ackermann, 2009). Hara et al. (1991) studied the abundance of viruses from the coast of Japan and assessed bacteriophages by epifluorescence microscopy and transmission electron microscopy. On comparison between the two techniques, for the enumeration of bacteriophages, it was observed that the epifluorescence microscopic method is simple and inexpensive for enumeration. However, transmission microscopy provides the ultra-structure of bacteriophages to reveal their morphology, thereby proving to be of great value in the identification and classification of marine phages. Electron microscopy has been instrumental in bacteriophage identification, enumeration and classification and has therefore revolutionized the study of viruses. Improvement in electron microscopy by the introduction of newer technologies along with ongoing improvement in technical expertise in viral transmission electron microscopy allows the proper investigation of viruses that have been isolated (Richert-Pöggeler et al., 2018). Bacteriophages are classified as per ICTV guidelines using morphology and the nature of nucleic acid as the basis of viral classification (Ackermann, 2009).

Epifluorescence Microscopy for Virus Enumeration

In a departure from the trend of major innovations in virus research coming from the Western World, a team of Japanese scientists (Hara et al., 1991) were the first to report a different method for the enumeration of viruses in natural waters. Virus particles were stained with DAPI (4',6-diamidino-2-phenylindole), a fluorescent stain that binds strongly to double-stranded DNA. Larger DNA-containing particles – bacteria and plankton were

also stained for comparison. This was done by collecting whole seawater samples on 0.015 µm polycarbonate filters, followed by staining and immediate viewing under a epifluorescence microscope (EFM). Viral counts thus obtained were well correlated with those obtained through TEM. As EFM does not reveal virus structure, but only a quantitative estimate of virus particles, the term virus-like particles (VLPs) was used to denote those satisfying the size criteria of viruses in a given sample analysed through EFM.

Flow Cytometry for Virus Enumeration

The drawbacks of TEM were the time, expense and effort associated with sample preparation as well as analysis. Further, TEM is strictly a laboratory method, and could not be applied to field investigations. EFM, while scoring over TEM in most of the above aspects, still involved tedious sample preparation and strenuous microscopic counting. The technique of flow cytometry (FCM), involving rapid and specific staining of all cells, with nucleic acid-specific stains such as SYBR Green, had been used for the detection and enumeration of phytoplankton and bacteria. Marie and co-workers (Marie et al., 1999) extended this technique to viruses, thus establishing a simple, elegant and speedy method to enumerate viruses in a variety of water samples, especially applicable to on-board studies. With the high throughput capacity of FCM, a large number of samples could be rapidly analysed, generating statistically significant information. The method was first optimized with an algal virus culture, after which it was successfully used with seawater samples from several locations. FCM was further applied to several families of viruses – *Baculoviridae*, *Podoviridae*, *Herpesviridae*, *Myoviridae*, *Siphoviridae*, *Phycodnaviridae*, and others (Brussaard et al., 2000). They found that the output cytogram could differentiate various classes of viruses, with fluorescence signal emitted is proportional to the genome size of the virus. At the lowest limit of detection were small genome-sized RNA viruses. Up to four different groups of viruses (differing according to size and genetic material) could be distinguished by their position in the cytogram. Brussaard provides a comprehensive description of the methodology used to obtain accurate virus counts and eliminate background signals – a major challenge with this method (Brussaard, 2004). Numerous researchers have subsequently applied this method to marine samples

Phage characterization and limitations

Once phage isolates have been acquired, numerous phenotypic and genomic data, including morphology for phage, can be retrieved from taxonomy, whole-genome sequencing (WGS), burst size, and one-step growth curve patterns.

Among these techniques, WGS is particularly beneficial for the analysis and discovery of phage genetic diversity. Due to their tiny size and frequent recombination events with bacterial host genomes, phage genomes are very varied (Hatfull, 2008). Therefore, phage genome analysis frequently results in the identification of bacterial and viral genes that have not been previously described or characterized.

The most notable instance is the discovery of auxiliary metabolic genes (AMGs) from marine cyanobacterial cyanophages. The genes *psbA* and *psbD* that code for the D1 and D2 proteins were discovered during a genome scan of the cyanophage S-PM2, which infects *Synechococcus* species.

Despite the numerous benefits of phage cultivation, there are significant drawbacks because phage isolation requires the identification and cultivation of suitable bacterial host. Overall, microbial community that can be cultivated using existing cultivation techniques is less than 1%. Furthermore, phages lack the globally conserved marker sequences needed for viral taxonomy, unlike prokaryotes. Although terminases, capsid proteins, and tape-measuring proteins are frequently found in phages of the order *Caudovirales* (Low et al., 2019), these genes are only found in certain types of phages. As a result, these genes are ineffective as universal phylogenetic markers. To overcome this challenge and limitations in the culture-based method, metaviromics, a recent approach that does not require viral isolation or amplification of genetic markers, was recently introduced to examine the viral population to get over these restrictions. As a result, recent developments in phage ecology have proven quite significant.

The cultivation-based method is mainly restricted to culturable bacterial strains, which is a major limitation for isolating and exploring viral diversity. The viral classification does not depend upon any conserved universal genetic markers, as in the case of prokaryotes which depend on 16sRNA genetic markers for molecular classification.

Culture-independent technique:

Metaviromics for aquatic virology

Metaviromics, a recent approach that does not require viral isolation or amplification of genetic markers, was recently introduced to examine the viral population to get over these restrictions. As a result, recent developments in phage ecology have proven quite significant.

After samples are pre-treated with cell removal, virus concentration, virus purification, and nucleic acid extraction, virome sequence data can be produced using viral shotgun metagenomic sequencing.

Filtration

The inclusion of prokaryotic sequences in virome data may result in incorrect interpretations of virome since metagenome sequences are ambiguous about the distinction between bacterial and viral genes. To selectively collect viral particles without contaminating prokaryotic cells, the samples for virome analysis should be pre-filtered. 0.2- μm pore size filters are frequently employed to get rid of any cellular organisms because phages are known to have a diameter that varies from 0.045 to 0.185 μm (Dion et al., 2020). Some environmental samples may be filtered via pores as small as 0.1 μm to guarantee that all prokaryotic cells have been removed (Pesant et al., 2015).

However, this practice greatly increases filtration time. The recent discovery of giant viruses with a capsid size of ~ 500 nm (Van Etten, Lane, and Dunigan., 2010) has raised the question of whether current filtration practices lead to the loss of giant phages in the filtrate, suggesting the use of 0.45- μm or larger pore size filters to filtrate viruses more efficiently and comprehensively (Dávila-Ramos et al., 2019). Therefore, the choice of pore size for pre-filtration should be flexibly determined depending on whether the samples must contain giant viruses or minimize prokaryotic cell contamination.

Concentration

Due to the small size of viral genomes, copious amounts of viral particles must be collected from a large volume of samples to obtain sufficient viral nucleic acids for metagenome sequencing. To improve nucleic acid extraction efficiency, phage particles were concentrated using diverse methods such as tangential flow filtration (TFF) and FeCl_3 flocculation. Sample concentration to recover viruses from environmental samples, possess

unique challenges. TFF is the most prominent method used to concentrate viruses from natural water systems, because it reduces filter clogging and allows concentration from large volumes (Wommack et al., 2010). Although this approach uses different TFF cartridge pore size cut-off (30-100kDa), resulting in a significant reduction in the final viral concentrate volume (Rastrojo and Alcamí, 2017). TFF has been used to isolate viral particles diverse environments (Bench et al., 2007; Schoenfeld et al., 2008). TFF concentration method was even used to concentrate the dsDNA viruses in South African hot spring for metagenomics studies (Zablocki et al., 2017).

The particles smaller than the filter pore size are pushed out through the filters (filtrate). A back pressure is used to force the filtrate through the holes. The remaining sample (retentate) is collected into a reservoir basin, repeatedly cycled through the filters (Thurber et al., 2009). The TFF method being the most frequently used viral concentration method, it requires expensive equipment's (hundred to thousand US Dollars) and time consuming and results in highly low virus recoveries (Colombet et al., 2007; John et al., 2011; Wommack et al., 2010). The recovery rate of virus particles can be increased by further concentrating the TFF concentrate using ultrafiltration (Langenfeld et al., 2021). Another popular technique for concentrating viruses is to flocculate the viruses with FeCl₃, and then collect the flocs onto a 0.22-µm filter (John et al., 2011; Thurber, 2009). This technique concentrates viruses more quickly than the TFF method does (Hurwitz et al., 2013).

Purification

Concentrated virome samples are commonly purified through three methods: chloroform treatment, 0.22-µm filtration, and density-gradient centrifugation. Chloroform treatment effectively removes bacterial cells by disrupting lipid cell membranes (Conceição-Neto et al., 2015). However, this approach may also eliminate the lipids of enveloped phages, thus leading to biases in the virome sequencing data. Density-gradient ultracentrifugation using cesium-chloride or sucrose allows for the acquisition of phages with specific density, enabling microbiologists to selectively collect phages of interest. However, this approach neglects the phages with unknown density (Thurber et al., 2009).

Sequencing

Using established procedures for creating viral metagenome libraries, viral nucleic acids, either DNA or RNA, can be easily extracted from phage concentrates. Multiple-displacement amplification (MDA) of DNA can be used (Yoshida et al., 2018), when the required quantity of DNA for the building of metagenome libraries cannot be attained. The MDA amplifies a low amount of viral nucleic acid and this method can be biased towards ssDNA viruses. The overrepresentation of ssDNA sequences in the virome data is due to MDA's known bias towards ssDNA viruses (Kim et al., 2008).

Furthermore, because the amount of nucleic acids required for library synthesis has decreased due to advancements in next-generation sequencing (NGS), MDA is no longer widely employed for the new generation of metagenome libraries. Typically, Thermo-Fisher Ion Torrent or Illumina MiSeq/HiSeq sequencers are frequently used to generate virome data. Illumina HiSeq sequencers have been reported to generate more raw reads than Ion Torrent systems when starting with the same amount of DNA (Solonenko et al., 2013).

Long viral metagenomic reads can now be produced thanks to the recently developed Oxford Nanopore long-read sequencing technology (ONT), which has contributed to the creation of novel viral genomes (Ji et al., 2020). In comparison, to the Illumina sequencing platform, ONT takes a disproportionately large amount of input DNA and has frequent error rates (Guerin and Hill, 2020); as a result, scientists must consider these limitations when choosing a sequencing method. By correcting for the regional sequencing mistakes in ONT and the short virome read sequences produced by the Illumina platform, the combined use of ONT and the Illumina sequencing platform can be a good alternative for producing virome reads with enhanced quality.

RNA virospheres

Due to their established dominance, dsDNA phages have been the focus of the majority of prior virome research. However, freshwater and marine virome investigations have shown that, like dsDNA viruses, RNA viruses are also numerically dominant and preferentially infect archaea or eukaryotes (Djikeng et al., 2009; Miranda et al., 2016). There were incredibly few sequenced and reported RNA phages up until recently (Greninger, 2018; Yang et al., 2016).

This means that, despite efforts to create RNA virome data by creating cDNA from

extracted viral RNA, RNA phages are difficult to classify using public databases. The number of known RNA virus sequences has doubled as a result of a recent marine virome study that focused on RNA viruses, indicating the future possibility of using RNA viromes to find new RNA virospheres (Wolf et al., 2020).

Uncultivated viral genomes (UViGs) assembled from virome sequences

Virome reads are used to build viral contigs, especially for uncultivated viral genomes (UViG), like how bacterial metagenome sequences are used to generate metagenome-assembled genomes. The best quality contigs with long sequences are produced using the most popular virome read assembly tools, SPAdes (Bankevich et al., 2012) and IDBA-UD (Peng et al., 2012).

Based on a few viral genomic traits, viral or non-viral contigs are sorted within the assembled contigs. Based on the presence of crucial phage sequences including genes encoding the capsid protein, tail, and terminase, the VirSorter program searches the assembled virome contigs for suspected viral sequences (Roux, et al., 2015). Through the use of deep learning training and k-mer frequency profiles based on both phage genome databases and accumulated virome sequences, recently developed artificial intelligence technology has improved the detection of viral sequences (Ren et al., 2020). Utilizing network analysis based on k-mer distribution patterns (Jang et al., 2019), sequence similarities based on the Dice coefficient (Meier-Kolthoff and Göker, 2017), and proteomic tree inference (Nishimura et al., 2017), assembled and sorted UViGs can subsequently be appraised and taxonomically categorized.

“crAssphage

Before cultivation, the assembly of virome contigs may result in the identification of new phage sequences. One of the most abundant and widely distributed phage genomes within human gut viromes was discovered in 2014 when Dutilh and coworkers (2014) built a putative phage genome from a human faecal virome. The putative phage genome was given the name "crAssphage," even though this phage has not yet been cultivated, and thorough virome investigations revealed that crAssphage populations displayed spatial speciation.

Although raw virome reads can be assembled into UViGs that closely match viral

genome sequences, the uniqueness of viral sequences and the absence of genetic markers that are generally conserved prevent the majority of virome contigs from being taxonomically categorized. As a result, various methods have been tried to identify these virome contigs. It is believed that phages carry a piece of their host's genome because they are known to be prone to horizontal gene transfer (HGT) when they infect host cells.

Virome sequences were used to identify viral contigs that may infect actinobacterial strains, as suggested by the presence of a *whiB* transcription factor annotated in the *Actinobacteria* acI clade (Ghai et al., 2017). Similar to this, 77 single-copy protein indicators, such as terminase, capsid, and DNA polymerase, that are often found in the order *Caudovirales* were used to screen probable phage genomes predicted to infect bacterial strains belonging to the SAR11 clade (Low et al., 2019). 26 more potential phage genomes have been predicted from marine metagenomes (Zaragoza-Solas et al., 2020) based on sequence similarity and phylogenetic inference of SAR11-infecting phage sequences, such as HTVC010P (Zhao et al., 2013).

The extensive distribution of potential phage sequences that may infect freshwater SAR11 groups was also found by virome research carried out in freshwater lakes (Chen et al., 2019). The International Committee on Taxonomy of Viruses (ICTV) recently agreed to include virome-derived genomes in its database because the number of viral sequences recovered from viromes currently outnumbers the number of cultured sequences (Simmonds et al., 2017).

Virome-derived AMGs

Virome comparative examination of novel phage genomes frequently results in the identification of unexpectedly functional phage genes. Phages may accidentally pick up pieces of bacterial genes during replication; these acquired genes may stay in the phage genome as AMGs and function during host-phage interactions. The existence of host-derived AMGs in the phage genomes has been demonstrated by comparative virome analyses. Phages have been anticipated to exhibit a range of class I and class II AMGs in addition to the photosynthesis-related AMGs outlined above (Hurwitz and U'Ren, 2016). For instance, the analysis of viral contigs obtained from the TARA ocean expedition led to the identification of the gene **amoC**, which codes for the ammonia subunit C of ammonia monooxygenase, demonstrating for the first time that phages can interfere with ammonia

oxidation in host bacteria (Roux et al., 2016).

The *pmoC* gene, a component of the methane monooxygenase involved in methane oxidation, was recently identified through a virome analysis conducted in freshwater lakes containing methanotrophs (Chen et al., 2020). Increased *pmoC* expression was seen in the transcriptional analysis of *pmoC*-containing virome contigs, suggesting that the phage *pmoC* gene may also be expressed in bacterial host cells to support bacterial metabolism.

Bacterial genes that confer antibiotic resistance (ARGs) are hardly found in phage genomes. Due to their rarity in phage genomes and the diversity of gene sequences in virome reads, there has been a discussion on whether or not phage-encoded ARGs are functional (Enault et al., 2017). Recently, novel phage-beta-lactamase genes have been revealed in urban river virome data; these genes have been confirmed to be functional, and when produced in *E. coli*, they exhibit antibiotic resistance (Moon et al., 2020). As a result, phage-encoded ARGs probably operate as AMGs and can spread via transduction between bacterial populations. In addition to enabling the extension of phage gene databases, the finding of additional AMGs, including ARGs inside virome contigs, has also improved our understanding of phage-host dynamics and the ecological significance of phages.

Viruses in the Era of Metagenomics

In the mid-21st century and continuing until today, the NGS and high-throughput sequencing (HTS) technology have advanced and the cost of sequencing coming down several orders of magnitude has tremendously increased viral metagenomic sequencing (Goodwin et al., 2016). Armed with new technology in computational biology, our understanding of viruses has tremendously increased, which led to the growth of large databases of the complete viral genome and viral proteins (Camargo et al., 2023; Paez-Espino et al., 2017; Roux et al., 2021). Our understanding of the global virosphere is reshaping with the genomes of uncultivated viruses (UViGs) found in shotgun sequencing datasets rather than being produced by laboratory virus cultivation. Since 2016, IMG/VR has served as a leading resource in the field of viral ecology (Paez-Espino et al., 2017; Roux et al., 2021), the new version IMG/VR v.4 (<https://img.jgi.doe.gov/cgi-bin/vr/main.cgi>) has the most extensive UViGs collection to date, and will undoubtedly reveal numerous novel insights into the ecological parameters influencing viral variety, virus evolutionary history, and viral-encoded functional diversity (Camargo et al., 2023).

ICTV

The International Committee on Taxonomy of Viruses (ICTV), the foremost body, recently opted in 2017 to add virome-derived genomes to its database because the number of viral sequences of uncultivated viruses has surpassed that of the cultured viral sequences (Simmonds et al., 2017). ICTV recently revised the viral taxonomic ranking structure in 2020, increasing the number of hierarchical ranks to 15 to classify novel viruses and allowing the use of genomic sequence data in the absence of electron microscopy images (ICTV, 2020).

Historical metaviromics and phage culture techniques developments in aquatic viral ecology. By using the plaque assay to isolate phages that infect pathogenic bacteria, culture-dependent methods first isolated a variety of phages that interacted with important bacterial players in nature. Numerous culture-neutral methods that embraced the most recent advancements in NGS have resulted in the compilation of enormous volumes of virome data from various ecosystems.

Massive volumes of virome data from many ecosystems have been accumulated as a result of a few culture-independent methodologies using the recent advancement of NGS. The aquatic viral metagenome effectively became a "black box" after comparing these two methods and discovering that the genome sequences of cultivated phages only make up a small percentage of it. Unknown sequences found in the virome may have been produced by phages that have not yet been cultivated but are infecting known hosts of cultured bacteria or by phages that have not yet been cultured but are infecting uncultured bacteria. As a result, efforts to decipher the aforementioned viral "black box" must concentrate on isolating phages as well as using novel techniques to cultivate the bulk of bacterial hosts that are not currently grown.

In this regard, it is important to draw attention to the recently created novel technique for defining phage-host couples that uses proximity ligation methods to connect viral and rRNA transcripts (Ignacio-Espinoza et al., 2020). In conclusion, research into phage-host dynamics will surely benefit from the separation of both bacterial hosts and their phages as well as the development of innovative bioinformatic or experimental methods.

Despite a wide distribution of hot springs throughout India, limited studies have been conducted on the microbiological analysis of these hot springs, especially viruses. There are minimal reports of viruses from Indian hot springs. Microbial diversity studies in Indian hot springs focus mainly on Himalayan geothermal provinces at higher temperatures and altitudes. There are only two hot springs viral diversity metagenomic studies reports from India belonging to the Himalayan geothermal province (Sharma et al., 2018; Das et al., 2020).

Most published articles report microbiological isolation of thermophilic bacteria or metagenomic bacterial diversity studies that can be exploited for biotechnological potential (Poddar and Das, 2018; Vijayan et al., 2023). Only two metagenomics reports on viral diversity studies have been reported in India from Mannikaran (Sharma et al., 2018) and Sikkim hot springs (Das et al., 2020), belonging to the Himalayan geothermal province (Table 2. 2).

There are no reports on the isolation of culturable bacteria that serve as potential hosts for bacteriophage isolation from Indian hot springs. Also, Maharashtra is one such bio-diverse state of India that harbors 60 diverse alkaline hot springs in the West Coast Geothermal province of India, and it is relatively unexplored for thermophilic bacteria and viruses.

Table 2. 2: Studies in Indian hot springs metagenomics and viral metagenomics.

Study	Hot spring	Reference
Metavirome and its functional diversity analysis through microbiome study of the Sikkim Himalayan hot spring solfataric mud sediments.	Sikkim	(Das et al., 2020)
Bacterial and Archaeal Viruses of Himalayan Hot Springs at Manikaran Modulate Host Genomes	Manikaran	(Sharma et al., 2018)
Draft Genome Sequence of <i>Anoxybacillus mongoliensis</i> Strain MB4, a Sulfur Utilizing Aerobic Thermophile Isolated from a Hot Spring in Tattapani, Central India	Tattapani	(Parul et al., 2017)
Metagenomic Analysis of Hot Springs in Central India Reveals Hydrocarbon Degrading Thermophiles and Pathways Essential for Survival in Extreme Environments	Anhoni and Tattapani	(Saxena et al., 2017)
Microbial diversity of Jakrem hot spring, Meghalaya using metagenomic analysis.	Jakrem	(Panda et al., 2015)
Thermophilic bacteria <i>Aeromonashydrophila</i> isolated from hot Water spring	Vajreshwari Ganeshpuri hot spring,	(Patil et al., 2014)
Thermophilic bacteria isolated from the Tarabalo hot spring with considerable protease activity.	Tarabalo	(Panda et al., 2013)
Thermophilic bacteria from the Bakreshwar hot spring.	Bakreshwar	(Ghati et al., 2013)
Cellulase production by <i>Bacillus</i> sp. isolated from Vajreshwari hot spring	Vajreshwari	(Acharya et al., 2012)
Biodiversity of thermophiles in Manikaran and Vashisht hot spring	Manikaran	(Sharma et al., 2010)

CHAPTER 3

Isolation and identification of culturable bacteria from hot springs.

This chapter describes the isolation and identification of selected hot spring culturable bacteria that can be hosts for bacteriophage infection.

Spring water samples were collected from various Indian hot springs.

The success of cultivating the required host as a pure bacterial isolate is crucial for the isolation and maintenance of bacteriophage in the laboratory. The first goal was to isolate and maintain the potential host bacteria that could be exploited to isolate bacteriophages from hot springs. Therefore, preliminary screening for bacteriophages was used as the basis of the sampling technique to isolate culturable bacteria from few Indian hot springs. Furthermore, the potential bacterial host that was used for bacteriophage screening was subjected to polyphasic characterization and molecular phylogenetic studies.

As mentioned in the ‘Research objectives’ in chapter one, the culturable bacteria isolated from hot springs were primarily used to isolate and characterize bacteriophage from hot springs.

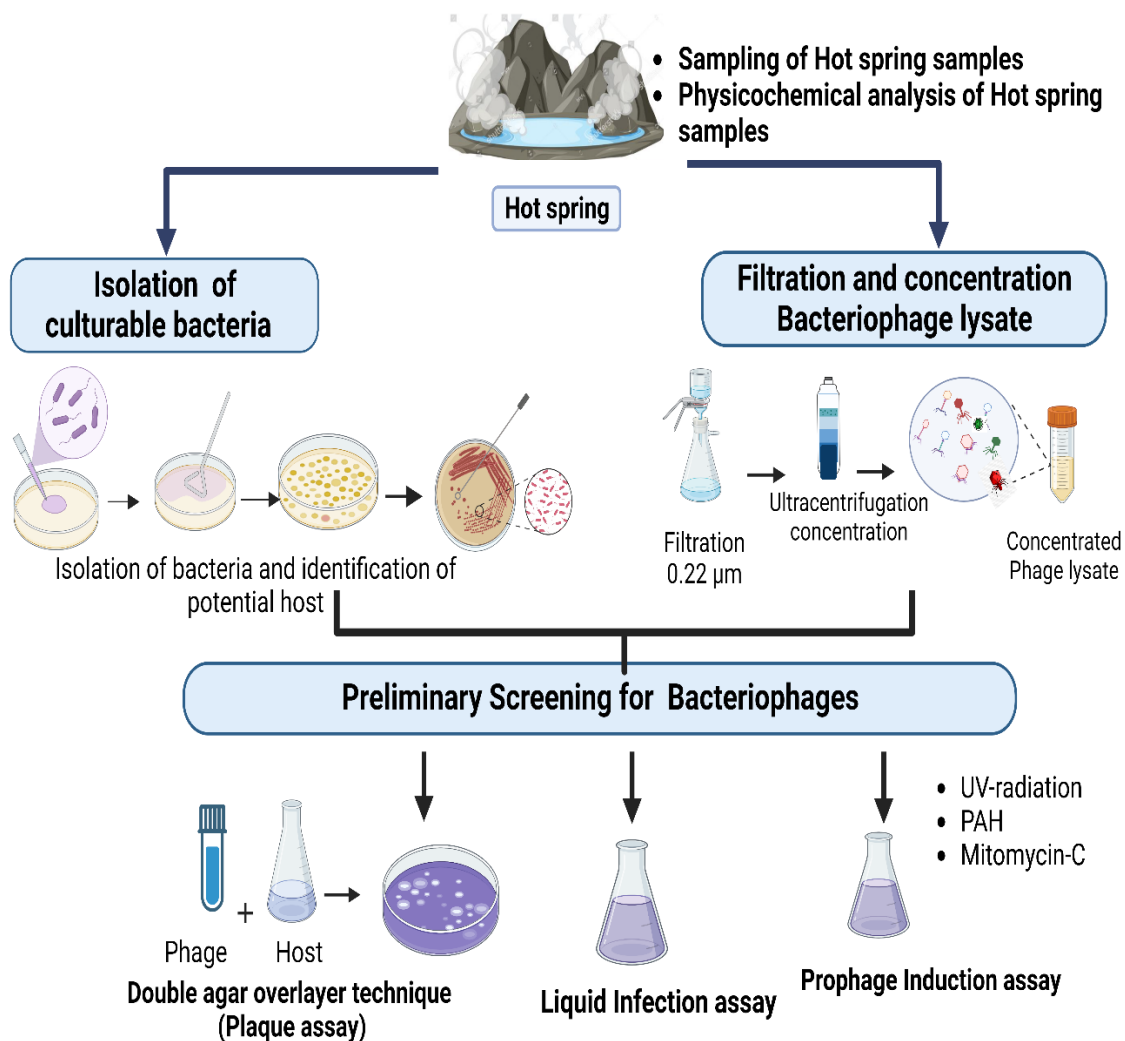
Workflow:

Figure 3. 1: Workflow for the isolation of culturable bacteria from hot springs.

Created with BioRender.com (<https://www.biorender.com/>).

3.1 Materials:

Thermus medium, Nutrient broth, Luria-Bertani broth. Bacterial isolate VB3, Bacterial isolate *JE2*, GeneiPure™ bacterial DNA purification Kit (Cat # 2115900031730) (GeNei, Bangalore, India). Qubit™ dsDNA BR Assay Kit (cat. #. Q32850, Invitrogen, Carlsbad, CA). HiMedia Carbohydrate utilizing kits (KB002 and KB009- Part A, B, and C) and all other reagents/chemicals were AR grade (analytical reagent), and double glass filtered distilled water was used.

3.2 Methods:

3.2.1 Study Area and Sampling

The basic workflow for this chapter is described in Figure 3. 1. Water samples were collected from six hot springs from the West Coast geothermal province (WCGP) and one from the Himalayan Geothermal province at different periods (Table 3. 1). The WCGP in Maharashtra harbors several thermal springs along the 350 km linear stretch. Vajreshwari and Ganeshpuri of WCGP stretch about 7 km in the bed of River Tansa, Thane district, Maharashtra, whereas hot springs Rajwadi, Tural, and Aravali, are located at the N20°W-S20°E fracture zone, 10 kilometers north of Sangameshwar town on the banks of a tributary that joins the Shastri River.

These hot springs in WCGP of Maharashtra are manifested in the basaltic terrain of the Deccan traps of Cretaceous-Paleocene age (~65 Ma). The hot water emerges from the margins of dolerite dykes, intense fractures, and shears (Low et al., 2020; Saxena, 1987).

Several hot springs are in the Mandi and Kullu districts of Himachal Pradesh. 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 (Shirkot et al., 2018).

Water samples were collected from various hot springs in India like Vashisht (Himachal Pradesh, Himalayan province), Aravali, Tural, Rajwadi, Rajapur, Vajreshwari, and Ganeshpuri (Maharashtra) WCGP of India (Figure 3. 2 and Figure 3. 3). The collected samples were transported to the laboratory and analysed within 12h.

Table 3. 1: Sample collection sites.

S No	Sampling Site	Geothermal province	Location
1.	Vajreshwari , Maharashtra	WCGP	19°29'44.9"N 73°02'05.8"E
2.	Ganeshpuri , Maharashtra	WCGP	19°29'22.5"N 73°01'01.7"E
3.	Aravali , Maharashtra	WCGP	17°18'36.9"N 73°31'37.0"E
4.	Tural , Maharashtra	WCGP	17°25'54" N, 73°52'99" E
5.	Vashisht , Himachal Pradesh	Himalayan geothermal province	32°16'00.8"N 77°11'15.0"E
6.	Rajapur , Maharashtra	WCGP	16°38'44.6"N 73°31'52.1"E
7.	Rajwadi , Maharashtra	WCGP	17°14'23.1"N 73°33'42.2"E

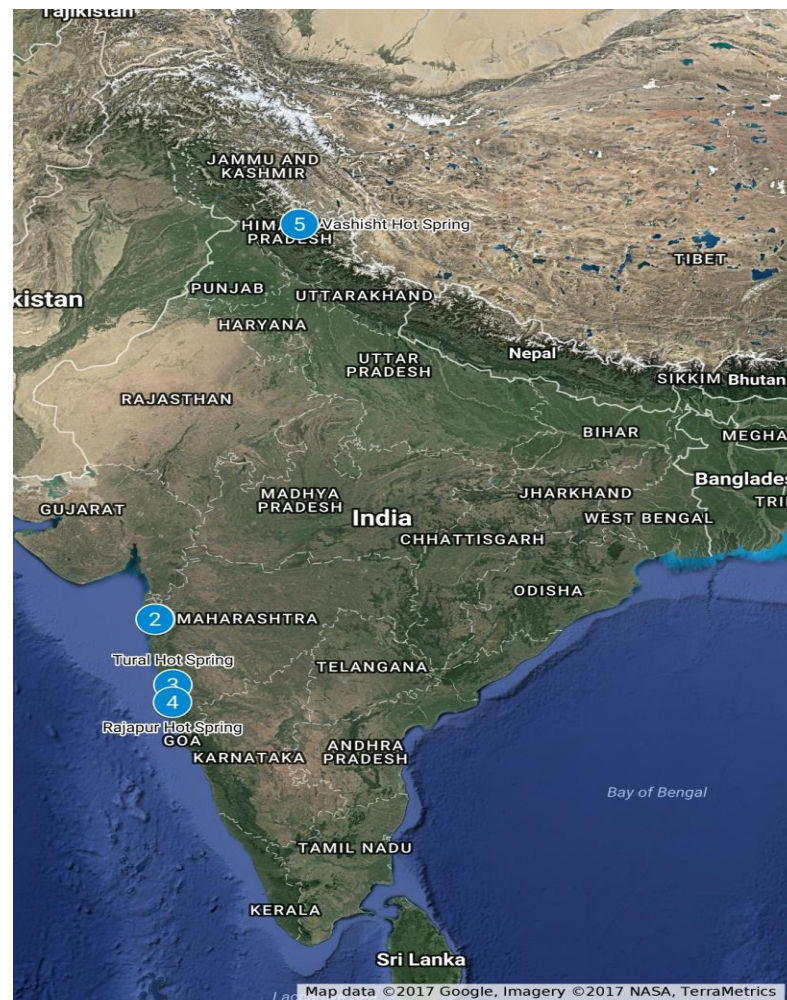


Figure 3. 2: Map showing the location of Indian hot springs used in the present study.



Figure 3. 3: Water samples collected from Indian Hot springs.

A) Vashisht hot spring, B) Tural, (C) Vajeshwari, (D) Rajwadi hot spring Kund, (E) Rajwadi sprouts around the area, (F) Rajapur Hot spring.

3.2.2 Physicochemical analysis of the hot spring water sample

In situ, physicochemical parameters such as temperature, pH, electrical conductivity, and TDS of Vajreshwari, Ganeshpuri, Aravali, Tural, Vashisht, Rajapur and Rajwadi hot spring water were recorded using a multi-parameter (CyberScan PC 650 Multi-Parameter, Eutech Instruments, Thermo Fisher Scientific, India).

3.2.3 Isolation of culturable bacteria from hot spring

The water samples were screened for the growth of bacteria at 40°C, 50°C, and 60°C on different mediums such as *Thermus medium*: [(ATCC medium 697) containing 0.5% NaCl, 0.5% peptone, 0.4% beef extract, 0.2% yeast extract and 0.8% gelrite along with 0.1%

MgSO₄.7H₂O (Gursahani and Gupta, 2011)], Basal Medium (BM) (g/L): [KH₂PO₄: 1.36; (NH₄)₂SO₄: 1.0; MgSO₄.7H₂O: 0.2; FeSO₄: 0.001; NaCl: 2.0; yeast extract, 1.0 and supplemented with 0.2% Sodium succinate), Nutrient broth (NA) and Luria-Bertani medium (Patel et al., 2006).

Briefly, 100 µl of the water sample was spread-plated on solid agar/gelrite plates and incubated at 40°C, 50°C, and 60°C in sealed plastic bags. After incubating on solid agar/gelrite plates for 3-5 days at 40, 50, and 60°C, morphologically different bacteria were picked and purified on a fresh plate. The bacterial isolates were preserved in respective mediums containing 15% glycerol at -80°C. The isolated cultures were purified and maintained in the laboratory. These bacterial isolates were evaluated for being a potential host for bacteriophage infection.

3.2.4 Filtration and concentration of water sample for virus enrichment

Approximately 3L of hot spring water was filtered and concentrated for virus enrichment. The water sample was centrifuged at 3000 x g for 10 min and filtered through 0.22 µm pore-size PVDF filters (Millipore, USA). The filtrate was stored at 4°C until further downstream processing. 10% (w/v) polyethylene glycol 8000 (PEG 8000) and 1M NaCl were added to this filtrate to precipitate the viral particles (Sambrook, 1989). This mixture was incubated at 4°C overnight and centrifuged at 11,000 x g for 20 min. The viral particles precipitated were suspended in 50 ml SM buffer (Saline magnesium buffer: 100 mM NaCl, 8.1 mM MgSO₄, 50 mM Tris-HCl, 0.01% gelatin, and pH 7.5). Virus concentration by ultracentrifugation was carried out in a tabletop ultracentrifuge Optima MAX-XP using an MLS-50 rotor (Beckman Coulter, United States) (Prata et al., 2012). The viral sample was ultracentrifuged at 80,000 x g for 4 h at 4°C under vacuum. The supernatant was discarded, and viral particles were resuspended in 10 ml SM buffer and left overnight.

3.2.5 Preliminary screening for bacteriophages

Bacteriophages were screened against the bacterial isolate by double-layer agar overlay technique (Kropinski et al., 2009) after enriching the concentrated viral sample with 10 X BM broth. Briefly, the 10 ml viral concentrate from the hot spring water was added to 100 ml BM broth inoculated with each bacterial isolate and incubated at 40, 50, and 60°C for 24 h. This enriched culture was centrifuged at 10,000 x g for 15 min, filtered through a 0.22 µm pore-size PVDF filter (Millipore, USA), and stored as phage stock at 4°C.

- a) Plaque assay:** Bacteriophage infecting the host bacterial isolates were screened by plating on a semi-solid gelrite/agar medium using a double-layer agar overlay technique (Kropinski et al., 2009). All mediums were supplemented with 10 mM CaCl₂ for bacteriophage propagation. Briefly, in a 2 mL microcentrifuge tube, 400 µL of filtered enriched phage stock and 800 µL of logarithmic growing bacterial isolate culture were mixed and incubated for 20 min at room temperature for adsorption. This mixture was then added to a 4 mL respective top medium containing 0.4% gelrite or 0.7% agarose and poured onto the medium culture plate containing 0.8% gelrite or 1.5% agar. After 48 h, if clear plaques were observed, they were picked up and suspended in SM buffer and kept at 4°C. This was followed by centrifugation at 10,000 x g for 10 min, and subsequently, the phage supernatant was passed through 0.22 µm pore-size PVDF filters to remove any bacterial debris. Single plaques were excised aseptically in SM buffer and used to re-infect bacterial isolates to obtain plaques of similar morphology. After three rounds of bacterial infection, the phage from one such purified plaque was used for further studies.
- b) Spot assay:** Spot assay was also carried out on each of the lawns of bacterial isolates to check the infectivity of the isolated phage particles (Sambrook, 1989). A fraction of bacteriophage lysate/ filtrate (10µl) was spotted on top agar or agarose plate pre-plated with the respective host culture. The plates were then incubated at respective temperatures for 24-48 h (Kutter, 2009).
- c) Liquid infection assay:** The isolated host bacteria were treated with concentrated phage lysate. Compared to the control, the decrease in OD₆₀₀ of the host bacteria after infection with phage lysate indicates phage infection. This was further confirmed with plaque assay on a solid medium to check the viability of the phage (Sambrook, 1989).

d) Prophage induction

Prophage induction with Ultraviolet radiation

Mid-log phase-grown bacterial isolates were exposed to ultraviolet radiation (40W/cm²) at a distance of 15cm for 5 minutes (intervals of 30 s) inside the laminar airflow. For each interval, 100 µL fractions of respective bacterial cultures were aseptically withdrawn in 2 mL micro-centrifuge tubes and centrifuged at 10,000 g for 10 min at 30 °C. Aseptically, 10 µL of the culture supernatant was spotted on a previously pre-plated bacterial lawn to observe a zone of phage-mediated bacterial lysis indicated by a clearance zone (Jiang and Paul, 1996). A sterile medium was used as a control.

Prophage induction with Poly-aromatic hydrocarbon (PAH)

Mid-log phase grown bacterial strains were exposed to PAH (poly-aromatic hydrocarbon) such as 50 µg/mL pyrene, phenanthrene, and naphthalene to induce lysogen (Jiang and Paul, 1996). For each of the hydrocarbon-treated bacterial cultures, 100 µL fractions were aseptically withdrawn in 2 mL micro-centrifuge tubes and centrifuged at 10,000 g for 10 min at 4 °C. Aseptically, 10 µL of the supernatant was spotted on a medium plate containing previously pre-plated bacterial lawn to observe a zone of phage mediated bacterial lysis indicated by clearance zone. A sterile medium flask with each PAH was used as a control.

Prophage Induction using Mitomycin-C

Each bacterial isolate was grown in the respective medium till the OD₆₀₀ was 0.3, after which Mitomycin-C was added at 1 µg/mL (Chen et al., 2006). The OD₆₀₀ nm was recorded at intervals of two hours for 24h. Subsequently, the OD₆₀₀ was recorded every two hours. Cultures showing a reduction in absorbance were clarified by centrifugation at 5000 ×g for 10 min, followed by filtration through a 0.22 µm filter. 10µl of lysate was spotted on a previously pre-plated bacterial lawn plate (Sambrook, 1989).

The control experiment involves spotting 10µl of the respective bacterial isolate (uninduced by mitomycin C) on the host bacterium's pre-plated lawn.

3.2.6 Morphological and biochemical analysis.

The bacterial isolates that were preliminarily screened for bacteriophages and demonstrated plaque formation, spot assay, or a decrease in OD₆₀₀ after prophage induction were selected for morphological and biochemical analysis as illustrated in Bergey's Manual of Systematic Bacteriology (Bergey, 1994).

The cultures were Gram stained according to Hans Christian Gram (1884). After observing their morphology, the bacteria were also characterized as cocci (spherical cells) or rods (cylindrical cells). The Gram-positive cultures were stained according to Schaeffer-Fulton's method (Schaeffer and Fulton, 1933). The position of the spore was observed and recorded as central, sub-terminal, or terminal. The hanging drop method determined the cultures' motility (Collins and Lyne, 1970). The bacterial culture was grown on agar plates for 24 h for the catalase test. A loopful of the culture was taken on a glass slide, and 3 % hydrogen peroxide (2 drops) was added. The bacterial culture was observed for effervescence production (Vasanthakumari, 2009). The oxidase test was carried out by oxidase discs (HiMedia DD018). Studies on carbohydrate utilization of the bacteria were carried out using KB009 HiCarbo™ kits (HiMedia KB009A/ KB009B1 /KB009C). The kit contains a total of 35 tests for the utilization of carbohydrates. The utilization profile was determined as per the manufacturer's instructions (HiMedia Laboratories Pvt. Ltd, India). Biochemical tests for nitrate reduction, amylase, cellulase, protease, and lipase activity for the potential bacterial isolates were determined as per standard protocols. The bacterial isolates were preserved in respective media containing 15% (v/v) glycerol at -80°C.

3.2.7 Molecular identification and 16S rRNA gene sequencing

Bacterial isolates were inoculated into respective mediums and were incubated at 40, 50, and 60°C for 24 h. The bacterial genomic DNA of isolates was extracted using the GeneiPure™ bacterial DNA purification Kit (GeNei, Bangalore, India). The eluted DNA was stored at -20°C. Later, isolated DNA was quantified using the Qubit dsDNA HS Assay Kit and the Qubit 2.0 Fluorometer (Invitrogen, USA) as per the manufacturer's instruction. One µl of genomic DNA sample was diluted 200-fold in Qubit dsDNA HS assay buffer in clear plastic Qubit assay tubes and measured using a Qubit fluorometer.

The 16S rRNA gene sequencing was determined after PCR amplification of genomic DNA using 27F and 1525R universal primers (Table 3. 2 and Table 3. 3) (Brosius et al., 1978). The amplification was conducted in a 50 μ L reaction mixture containing approximately 100 ng bacterial genomic DNA, 10X polymerase buffer, 5 μ L dNTP mix, 0.25 μ M of each forward and reverse primer, and 1.5 Units of Taq polymerase. The reaction mixture was heated at 95°C for 5 min and immediately subjected to 30 cycles of denaturation at 95 °C for 1 min, annealing at 55 °C for 1 min, extension at 72 °C for 1 min, and final extension step at 72 °C for 10 mins. The amplified PCR product was resolved on 0.7% agarose gel to check for the PCR amplified product. The amplified PCR product was purified using a QIAquick PCR purification kit (Qiagen Inc, Hilden, Germany). The purified product was sequenced by ABI 3730XI electrophoresis system (Applied Biosystems, Foster City, CA, USA).

Table 3. 2: Primers used for 16S rRNA sequencing.

Primers	Sequence (Brosius et al., 1978)	Product size
Forward 27F	5'AGATTGATCMTGGCTCAG-3'	1.5 kb
Reverse 1525R	5'-AAGGAGGTGATCCAGCC-3'	

Table 3. 3: PCR reaction mixture

Sterile MilliQ Water	33.2μl
10X Polymerase Buffer	5.0 μ l
DNTP's mix	5.0 μ l
Template genomic DNA (approximately 100 ng)	5.0 μ l
Forward primer 27 F (0.25 μ m)	0.7 μ l
Reverse primer 1525 R (0.25 μ m)	0.6 μ l
Taq DNA polymerase	0.5 μ l

3.2.8 Phylogenetic analysis

The 16S rRNA PCR amplified product was sequenced. The obtained raw sequences were used to assemble the contigs using BioEdit software, version 7.2.5 (Hall, 1999), and aligned. BLAST analysis of the 16S rRNA sequence of the bacterial isolates with other sequences in the database was performed (<https://blast.ncbi.nlm.nih.gov/>). The closest recognized relatives of the bacterial isolates were identified and downloaded by comparing the 16S rRNA gene sequence of the strain to those available in the EzBioCloud database (<https://www.ezbiocloud.net/>). Multiple sequence alignment of 16S rRNA sequence was carried out using the Clustal W program. A phylogenetic tree was constructed by the Maximum likelihood method using MEGA-X software v.10.1.8 (Molecular Evolutionary Genetics Analysis across computing platforms) (Kumar et al., 2018). Bootstrap analysis was done with 1000 replication using the bootstrap option in MEGA-X software. The 16S rRNA sequence was deposited in NCBI- GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>).

3.3 Results and Discussion

3.3.1 Physicochemical Analysis of the hot springs

Intermittent gas emissions with a sulfurous smell were observed at hot springs' sprouts. The water samples temperature, pH and other physical parameters were recorded onsite (Table 3. 4). The temperature of the seven hot springs ranged between 40 and 60°C. The highest temperature was recorded at Tural (60°C), whereas the lowest was at Rajapur (40°C). The pH of the hot spring sample ranged from 7.2 to 7.7, which is quite similar to other Indian hot springs (Gurav et al., 2016; Poddar and Das, 2018). Hot springs in India are grouped into three categories: moderately hot (37–42 °C), hot (50–70 °C), and very hot (>70 °C). The highest temperature has been recorded for hot springs at Soldhar (Arya et al., 2015) and Manikaran, where the water attained nearly boiling conditions (90–98°C) (Sangwan et al., 2015). Hot springs in India have neutral to alkaline pH (7–7.5), and the microbial community is different from those in acidic hot springs environments (Kanokratana et al., 2004; Mathur et al., 2007; Aditiawati et al., 2009).

Table 3. 4: Physicochemical parameters of hot springs from the Himalayan and West Coast geothermal province of India.

Parameters	Vashisht (Himachal Pradesh)	Vajreshwari (Maharashtra)	Ganeshpuri (Maharashtra)	Aravali (Maharashtra)	Tural (Maharashtra)	Rajwadi (Maharashtra)	Rajapur (Maharashtra)
Temperature (°C)	50	45	50	50	60	55	40
pH	7.4	7.2	7.3	7.4	7.5	7.5	7.7
Conductivity ($\mu\text{S}/\text{cm}$)	162	556	2335	865	1475	1445	545
TDS (mg/L)	780	358	1492	553	945	956	345

3.3.2 Isolation of culturable bacteria from hot spring

A total of **144** bacterial isolates were obtained, which were cultured and maintained in the laboratory until further preliminary screening for bacteriophages was carried out (Figure 3. 4). Using these bacterial strains as hosts, screening for bacteriophages from various hot springs in India like Vashisht (Himachal Pradesh), Aravali, Tural, Rajwadi, Rajapur, Vajreshwari, and Ganeshpuri (Maharashtra) were carried out.

Out of the **144** bacterial isolates, only ten were observed to show either plaque formation, decrease in optical density after treatment with phage lysate, or prophage induction based on preliminary screening for bacteriophages. A total of ten bacterial isolates were selected as potential hosts for bacteriophage infection and hence were identified using the polyphasic approach and molecular phylogenetics. Among the ten bacterial cultures identified, two bacteria, namely, bacterial isolate VB3 isolated at 50°C and JE2 isolated at 60°C served as potential hosts for isolating bacteriophages of VB3 and JE2, respectively. (described in Chapter 4)

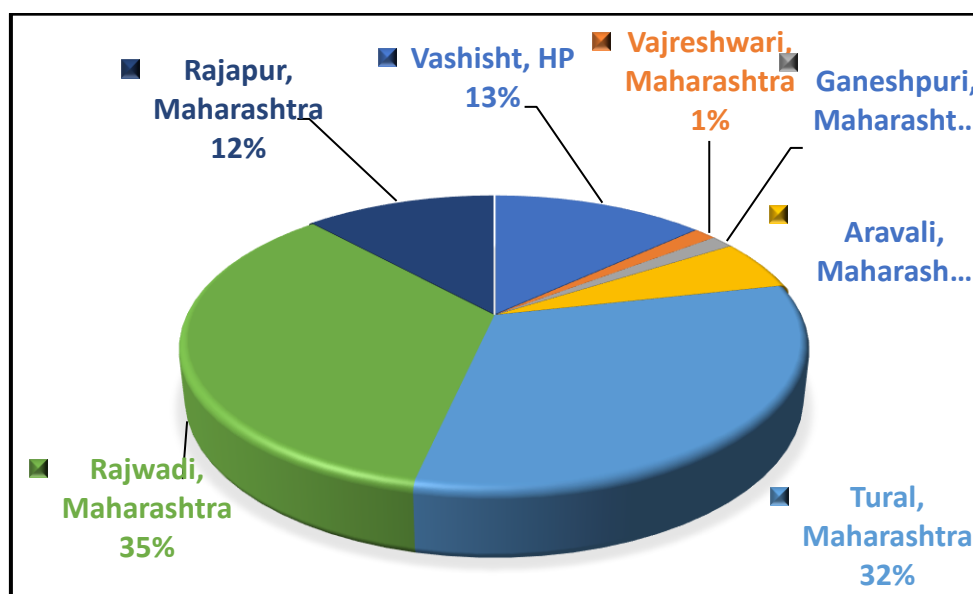


Figure 3. 4: Distribution of culturable bacterial isolates obtained from seven different Indian hot springs.

Table 3. 5: Host bacterial cultures that were selected for identification based on preliminary screening assay for bacteriophage isolation.

Bacterial isolated										
	JE2	VB3	RM2	RM6	RM8	RS3	RS5	TS1	TS2	NATY1
Hot spring	Rajwadi,	Vashisht	Rajwadi,	Rajwadi	Rajwadi,	Rajwadi,	Rajwadi,	Tural	Tural	Tural
Host genus	<i>Meiothermus</i>	<i>Pseudomonas</i>	<i>Anoxybacillus</i>	<i>Anoxybacillus</i>	<i>Parageobacillus</i>	<i>Anoxybacillus</i>	<i>Anoxybacillus</i>	<i>Anoxybacillus</i>	<i>Anoxybacillus</i>	<i>Anoxybacillus</i>
Preliminary Assay	Plaque	Prophage induction Mitomycin C	Liquid infection	Liquid infection	Liquid infection	Liquid infection	Liquid infection	Liquid infection	Liquid infection	Liquid infection
	Plaque formation	Decrease in cell density at OD 600nm and Spot assay	Decrease in cell density at OD 600nm	Decrease in cell density at OD 600nm	Decrease in cell density at OD 600nm	Decrease in cell density at OD 600nm	Decrease in cell density at OD 600nm	Decrease in cell density at OD 600nm	Decrease in cell density at OD 600nm	Decrease in cell density at OD 600nm
Indicative of lytic or lysogeny lifestyle	Lytic	Lysogeny	Lytic	Lytic	Lytic	Lytic	Lytic	Lytic	Lytic	Lytic
Phage Isolation and Propagation	Partially successful	Successful	Failed	Failed	Failed	Failed	Failed	Failed	Failed	Failed

Table 3. 5 depicts the host bacterium cultures selected for identification based on the preliminary screening assay for bacteriophage isolation. Bacterial isolates JE2 and VB3 were selected based on preliminary screening assays, and the plaque and prophage induction assay, respectively. The bacteriophage infecting the culture JE2 was indicative of a lytic lifestyle by showing the presence of plaque formation. Bacteriophage infecting VB3 culture is indicative of a lysogeny lifestyle with a decrease in cell density at OD_{600nm} after prophage induction with Mitomycin C. The bacteriophages infecting other cultures (RM2, RM6, RM8, RS3, RS5, TS1, TS2, and NATY1) were indicative of lytic lifestyles with a decrease in cell density after the liquid infection assay. These cultures were not further studied as the infection was lost after the second round of propagation. Thus, bacterial isolates JE2 and VB3 (as highlighted in Table 3. 6) were selected for identification studies and propagation.

Most microorganisms dwelling in extreme habitats are impossible to cultivate in the laboratory, preventing the study of their associated viruses. While many bacteriophages can be cultivated successfully in laboratory conditions, there are several reasons why they may not grow under certain circumstances. The bacteriophage interacts with their specific bacterial host with complex microbial communities in the natural environment (in our case thermal environment). By isolating and propagation phages in the laboratory cultures, we might have disrupted the natural interactions, potentially altering the phage biology and ecology. Also, even very low concentrations (low phage titer) in monoculture may lead to the rapid emergence of phage-resistant bacterial mutants or bacterial strains with altered phage susceptibility. This phage-host coevolutionary process can influence phage-host interactions and limit the effectiveness of cultured-dependent approaches for studying phage biology and evolution (Piel et al., 2022). Even very low concentrations (low phage titer) in environmental samples, make their isolation and cultivation challenging. Even if phages are present, they may not reach a detectable level in laboratory cultures without enrichment or concentration steps. Some bacteriophages may require additional factors, such as specific ions or cofactors, for successful infection and replication. If these factors are not present or are present in insufficient amounts in the growth medium, phage growth may be impaired (Ács et al., 2020).

Table 3. 6: 16S rRNA sequencing and Identification of bacterial hosts, which were positive for preliminary screening of bacteriophages.

Sno	Name	Accession number Sequence(bp)	Ezbiocloud (Type strains)	Ezbiocloud similarity
1	<i>Pseudomonas</i> sp. VB3	MZ500213 (1386 bp)	<i>Pseudomonas gessardii</i> DSM17152 ^(T)	99.63%
2	<i>Meiothermus</i> sp. JE2	MW186236 (1380bp)	<i>Meiothermus ruber</i> DSM 1279 ^(T)	100%
3	<i>Anoxybacillus</i> sp. RM2	OQ380510 (731bp)	<i>Anoxybacillus gonensis</i> G2 ^(T)	100%
4	<i>Anoxybacillus</i> sp. RM6	OQ380516 (702bp)	<i>Anoxybacillus</i> <i>salavatliensis</i> A343 ^(T)	95.21%
5	<i>Parageobacillus</i> sp. RM8	OQ380518 (1489bp)	<i>Parageobacillus toebii</i> NBRC 107807 ^(T)	95.52%
6	<i>Anoxybacillus</i> sp. RS3	OQ383222 (1469 bp)	<i>Anoxybacillus gonensis</i> G2 ^(T)	95.68%
7	<i>Anoxybacillus</i> sp. RS5	OR178142 (1481bp)	<i>Anoxybacillus</i> <i>pushchinoensis</i> K1 ^(T)	99.73%
8	<i>Anoxybacillus</i> sp. TS1	OQ383223 (1430bp)	<i>Anoxybacillus</i> <i>pushchinoensis</i> K1 ^(T)	97.38%
9	<i>Anoxybacillus</i> sp. NATY1.	OQ372992 (1409 bp)	<i>Anoxybacillus gonensis</i> G2 ^(T)	96.48%
10	<i>Anoxybacillus</i> sp. TS2	OQ380606 (1488bp)	<i>Anoxybacillus</i> <i>pushchinoensis</i> ^T K1 ^(T)	99.73%

Isolating and detecting viruses or bacteriophages from high-temperature hot springs is experimentally challenging, especially since they require specific growth conditions. The virus particles in hot springs have been observed by electron microscopy and propagated using cultivable bacteria and archaea isolated from these extreme environments (Rice et al., 2001). The typical lifecycle of a bacteriophage falls into two categories: lytic and lysogeny. Under certain external conditions (e.g., physical UV radiation or mitomycin C induction), active prophages can be induced during host cell division. Various reports suggest that the viruses in hot springs prefer to be in the lysogeny stage to survive in these extreme environments (Breitbart et al., 2004; Pride and Schoenfeld, 2008; Sharma et al., 2018; Jarett et al., 2020; Guajardo-leiva et al., 2021). Multiple viruses have been isolated from different hot springs around the world. However, there are no reports on isolating bacteriophage from Indian hot springs.

3.3.3 Bacteria for isolation of bacteriophage

A) Bacterial strain JE2

During the preliminary screening of the Rajwadi hot spring, a thermophilic red pigment-producing bacterium was isolated on BM and designated as bacterial strain JE2 (Figure 3. 5A). The bacteria formed smooth, convex, circular colonies with entire margins (Figure 3. 5 B). Strain JE2 was obligate thermophile, Gram-negative, and non-motile rods. Bacterial strain JE2 grew optimally at 60°C and pH 7.5. No growth was observed below 50°C and above 70°C (Figure 3. 6). The bacterial strain JE2 was catalase and oxidase positive. Further, the bacterial strain JE2 utilized various substrates such as D-lactose, D-fructose, D-galactose, D-mannose, maltose, D-cellobiose, D-raffinose, D-sorbitol, D-mannitol, and glycerol (Table 3. 7). Several biochemical characteristics of bacterial strain JE2 were similar to the other type strains of the genus *Meiothermus* (Nobre et al., 1996; Tindall et al., 2010).

Subsequently, a lytic bacteriophage designated as vB_Msp_JE2 was isolated from the Rajwadi hot spring that was propagated using thermophilic *Meiothermus* sp. strain JE2 (See Chapter 4 for detailed study).

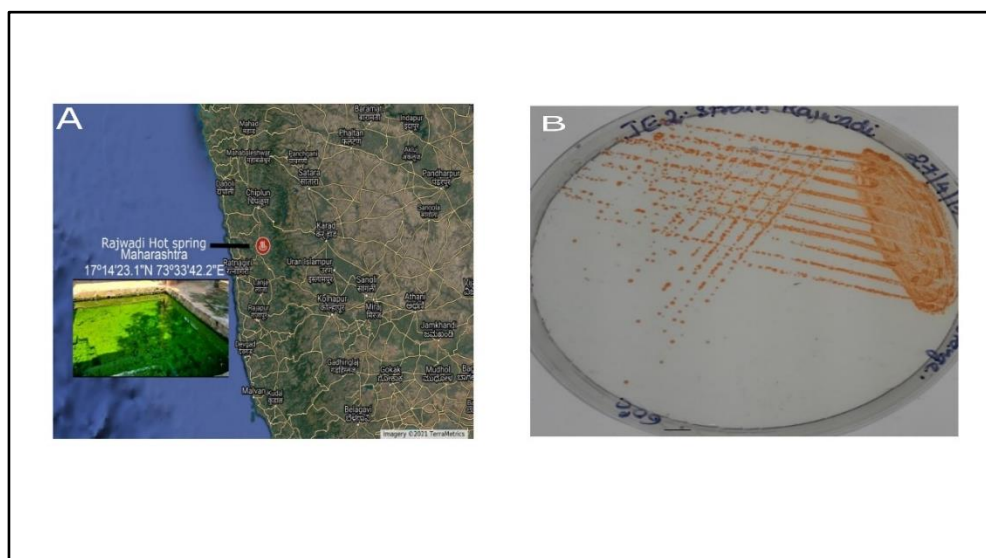


Figure 3. 5: (A) Geographical location of Rajwadi Hot Spring and sampling sites.

The given map was retrieved from Google Maps (2021). (B) Bacterial strain JE2 isolated from Rajwadi hot spring.

Table 3. 7: Morphological and biochemical characteristics of selected bacterial strain JE2

Characteristics	<i>Meiothermus</i> sp. Strain JE2
Gram character	Gram-negative
Pigmentation	Red
Motility	Non-motile
Shape	Rods
Endospore	NA
Reduction of NO ₃ ⁻ to NO ₂ ⁻	-
H ₂ S production	-
Vogues Proskauer	-
Catalase	+
Oxidase	+
Hydrolysis	
Starch	-
Casein	-
Xylan	-
Carbohydrate utilized	Lactose, fructose, galactose, mannose, maltose, cellobiose, raffinose, sorbitol, mannitol, and glycerol

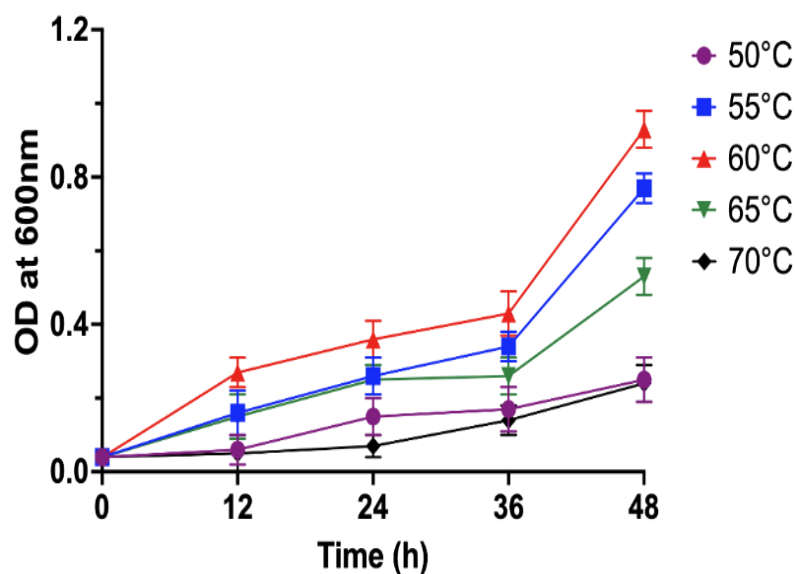


Figure 3. 6: Growth of thermophilic *Meiothermus* sp. strain JE2 at different temperatures.

16s rRNA sequencing and phylogenetic

Figure 3. 7 depicts the PCR amplified product of approximately 1.5 kb. BLAST analysis of strain JE2 16S rRNA sequence (MW186236) with 16S rDNA sequences of type strains from the EzBioCloud demonstrated 99% similarity with *Meiothermus ruber* and 98% similarity with *M. cateniformans*, *M. cerbereus*, and *M. taiwanensis*, respectively. A phylogenetic tree constructed using the sequence of closely related bacterial type strains by Maximum likelihood method indicated that the strain JE2 formed a coherent cluster with the clade comprising the genus *Meiothermus* type strains (Figure 3. 8). Strain JE2 is closely related to *Meiothermus ruber* (ATCC35948) and was identified as *Meiothermus* sp. strain JE2. *Meiothermus* sp. is slightly thermophilic, aerobic bacteria which belongs to phylum *Deinococcus-Thermus*. Species of the genus *Meiothermus* formerly belonged to the genus *Thermus* and are different from the species of *Thermus* by their ability to grow at lower temperatures. The optimum growth temperature of *Meiothermus* species is 50-65°C, and they can form red- or yellow pigment colonies. *Meiothermus* species are widely found in hot springs and artificial thermal environments. *Meiothermus* sp. strain JE2 isolated from the Rajwadi hot spring was compared with two other red-pigment-producing *Meiothermus* species isolated from the Paniphala hot spring, India (Table 3. 8).

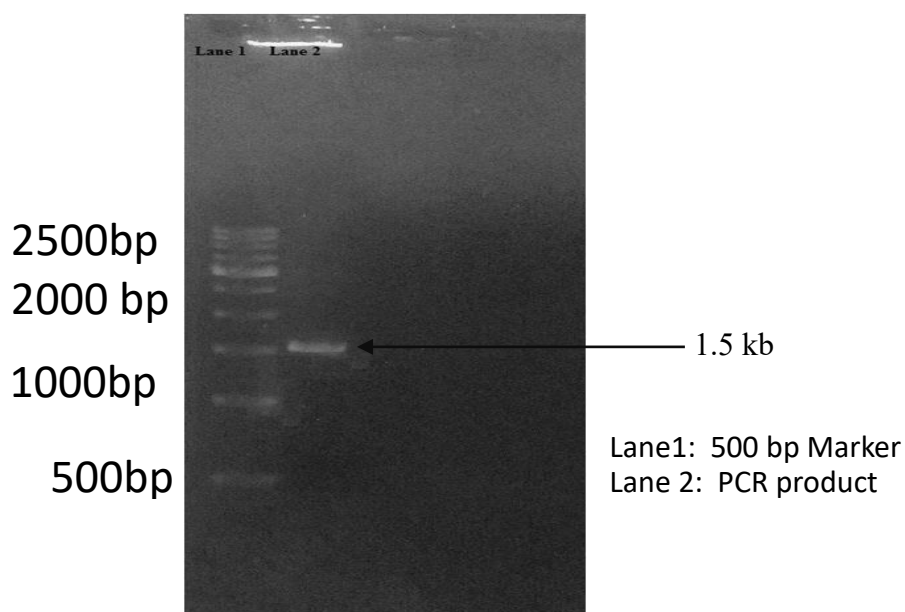


Figure 3. 7: Purified 16S rRNA PCR product of approximately 1.5 kb

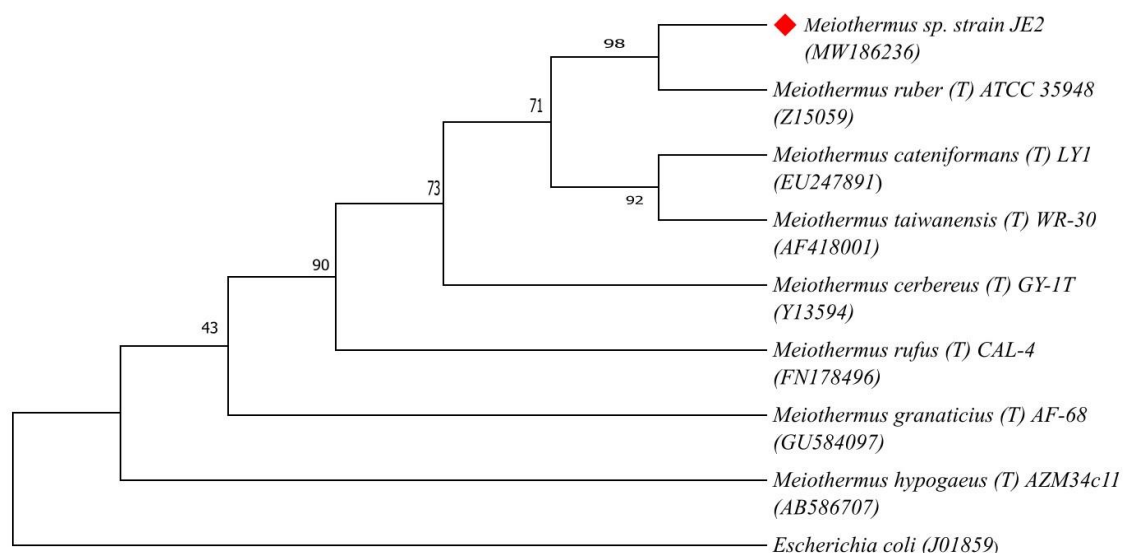


Figure 3. 8 : Molecular phylogenetic analysis of *Meiothermus* sp. strain JE2 by Maximum likelihood method.

Bootstrap analysis was done with 1000 replication using the bootstrap option in MEGA X. *E. coli* was selected as the out-group member.

Table 3. 8: Characteristics of *Meiothermus* sp. strain JE2 compared with other reported *Meiothermus* species isolated from Indian hot springs.

Characteristics	<i>Meiothermus</i> sp.JE2 (This study)	<i>Meiothermus</i> sp.RP (Mukherjee et al., 2017)	<i>Meiothermus</i> sp.TP (Mukherjee et al., 2017)
Growth temperature range (°C)	50-70	55-70	55-70
Optimal growth temperature (°C)	60	60	60
Pigment	Red	Red	Red
Geographic location	Rajwadi, Maharashtra, India	Paniphala Hot Spring, West Bengal, India	Paniphala Hot Spring, West Bengal, India
Casein hydrolysis	-	+	+
Amylase	-	+	+
Xylanase	-	+	+
Cellulase	+	+	+
Lipase	-	-	-

“+” -Present, “-” Negative results.

Several culture-dependent (Gupta et al., 2012; Mukherjee et al., 2016, 2017) and independent studies (Panda et al., 2015, 2016; Saxena et al., 2017) from Indian hot springs also demonstrated genus *Meiothermus* in these extreme geothermal environments. The genus *Meiothermus* was dominant in metagenomic samples with high temperatures (66-67°C) (Panda et al., 2015, 2016; Saxena et al., 2017). These results indicated that temperature is an essential factor for the growth of this organism in these thermal springs.

B) Isolation of bacterial strain VB3

The temperature and pH of the water sample collected from the Vashisht hot spring were recorded as 50°C and 7.4, respectively (Figure 3. 9 A). The conductivity and TDS measured *in situ* were 162 $\mu\text{S cm}^{-1}$ and 780 mg/L, respectively. While isolating culturable bacteria from Vashisht hot springs on BM plates, five morphologically distinct colonies were isolated on the BM plates. During preliminary bacteriophage screening, one of the bacterial isolates showed a reduction in cell density after induction with mitomycin C. The bacterial strain was designated as strain VB3, indicative of a potential lysogenic strain (Figure 3. 9 B). Growth of bacterial strain VB3 was observed from 25-50°C, with optimum temperature for growth observed at 30°C (Figure 3. 10). Growth of strain VB3 was observed between pH 6-10, with optimum pH of 7.5 being optimal for growth (Figure 3. 11). Strain VB3 grew on BM medium containing 0.1-0.2% NaCl. No growth was observed above 3% NaCl. The bacterial strain VB3 was Gram-negative, rods, and motile. Strain VB3 was positive for the catalase and oxidase test. No hydrolysis of casein and gelatin was observed. Nitrate reduction and ONPG tests were negative. The morphological and biochemical characteristics of strain VB3 is described in Table 3. 7.

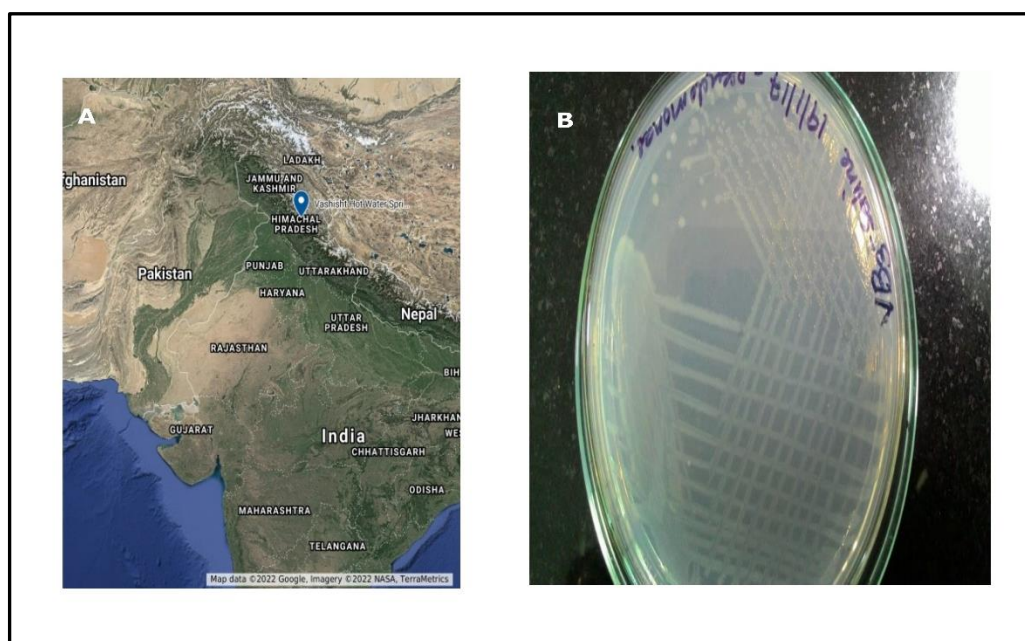


Figure 3. 9 : (A): Geographical location of Vashisht Hot Spring and sampling sites.

The given map was retrieved from Google Maps (2022). (B) Bacterial strain VB3 isolated from Vashisht hot spring.

Table 3. 9: Morphological and biochemical characteristics of the selected isolate VB3.

Property	Term
Gram character	Gram Negative
Cell-shape	Rods
Average size	1.416µm x 0.4 µm
Catalase	+
Oxidase	+
Endospore	Absent
Motility	Motile
Temperature (Optimum)	25-50 °C (30°C)
pH range (Optimum)	6-10 (7)
Salt tolerance	0.1-0.2%
Indole	-
Methyl red	-
Vogues Prosakuers	-
Nitrate reduction	-
Citrate	+
Urease	-
H ₂ S Production	-
Amylase	-
Protease	-
Cellulase	+
Lipase	+
Carbohydrate utilized	
Lactose, Xylose, Maltose, Fructose, Dextrose, Galactose, Melibiose, Sucrose, L-Arabinose, Mannose, Inulin, Sodium gluconate, glycerol, D-Arabinose, inositol, sorbitol, mannitol, cellobiose, citrate	+
Raffinose, Trehalose, Salicin, Dolcitol, adonitol, arabitol, erythriol, α methyl D-gluconate, rhamnose, melizitose, α methyl D-Mannoside, xylitol, α ortho-nitrophenyl β-galactoside (ONPG), esculin hydrolysis, D-arabinose, malonate, sorbose	-

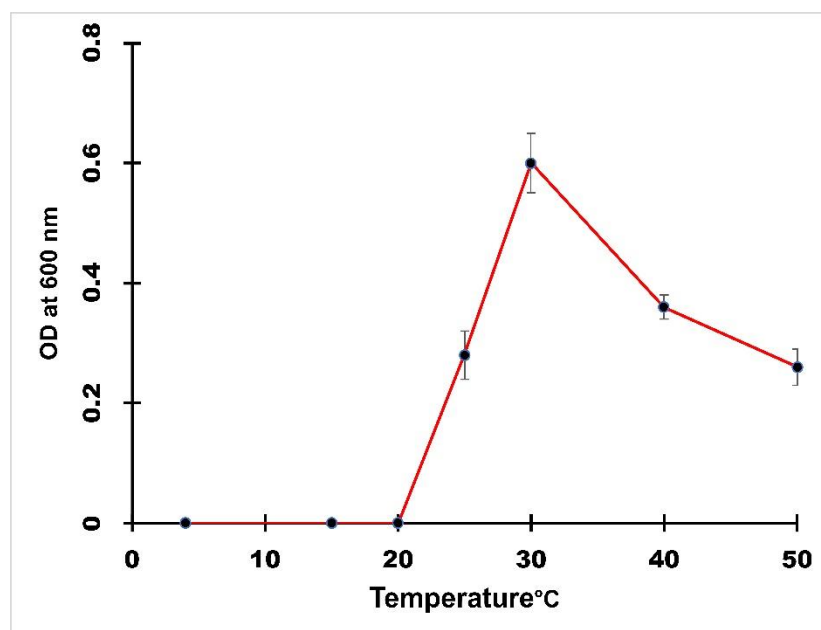


Figure 3. 10: Growth of bacterial strain VB3 at different temperatures.

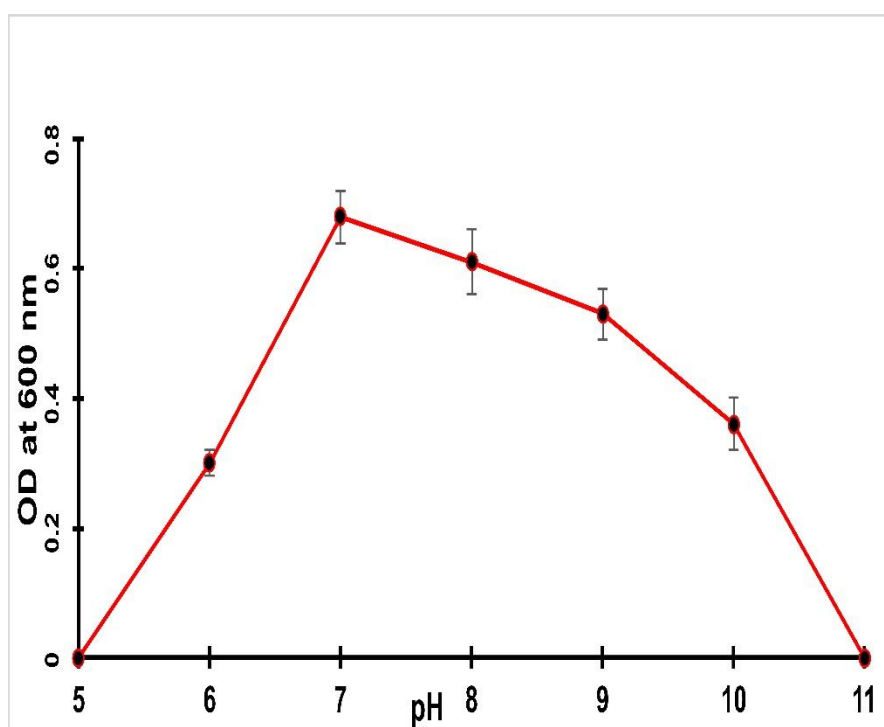


Figure 3. 11: Growth of bacterial strain VB3 at different pH.

16S rRNA sequencing and phylogenetic analysis

The 16S rRNA gene sequence of bacterial strain VB3 (MZ520013) demonstrated pairwise sequence similarity to *Pseudomonas gessardii* strain DSM 17152^T (99.63%), *P. synxantha* strain DSM 18928^T (99.55%), *Pseudomonas libanensis* strain CIP 105460^T (99.48%), *P. mucidolens* strain LMG 2223^T (99.40%), *P. azotoforman* strain DSM 18862^T (99.33%), *P. carnis* strain B4-1^T (99.33%), *P. paralactis* strain DSM 29164^T (99.33%), *P. lactis* strain DSM 29167^T (99.25%), and *P. 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* (Figure 3. 12). This topology relationship was supported by the Maximum-Likelihood tree also. Thus, strain VB3 is closely related to *P. gessardii* and was identified as *Pseudomonas* sp. VB3.

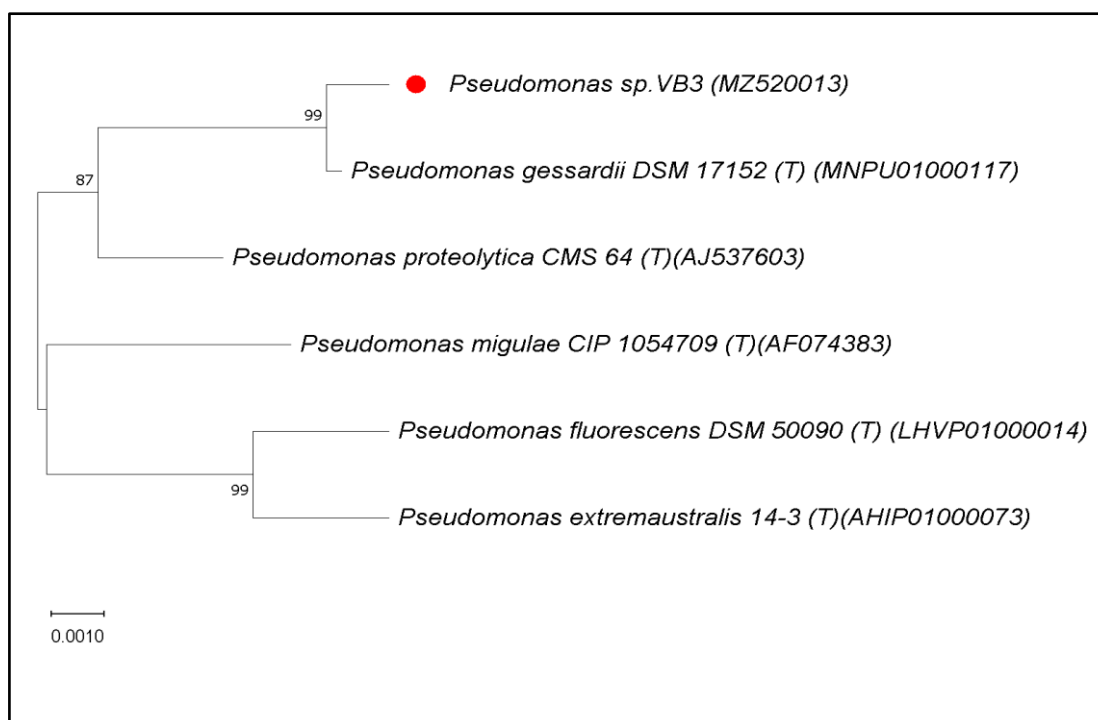


Figure 3. 12: Neighbor-joining phylogenetic tree based on 16S rRNA gene sequencing of *Pseudomonas* sp. strain VB3 (with a Bootstrap value of 1000 replication).

3.3 Conclusion

A total of **144** bacterial strains were obtained, which were cultured and maintained in the laboratory until further bacteriophage screening was conducted. Using these bacterial cultures as hosts, screening for bacteriophages from various hot springs in India like Vashisht (Himachal Pradesh), Aravali, Tural, Rajwadi, Rajapur, Vajreshwari, and Ganeshpuri (Maharashtra) was carried out. A total of ten bacteria were selected as potential hosts for bacteriophage infection and hence were identified using the polyphasic approach and molecular phylogenetics. Among the ten bacterial cultures identified, two bacteria, namely, bacterial strain **VB3** (isolated at 50°C) and **JE2** (isolated at 60°C), served as potential hosts for isolating bacteriophages of VB3 and JE2, respectively. Phylogenetic studies based on 16Sr RNA sequence analysis indicated that bacterial strains VB3 and JE2 formed coherent clusters with species of *Pseudomonas* and *Meiothermus*, respectively.

CHAPTER 4

Isolation of bacteriophage from hot springs

The chapter describes the isolation of bacteriophage from hot springs.

*This chapter describes the isolation of lytic *Meiothermus* phage and temperate *Pseudomonas* phage from hot spring.*

Bacteriophages are present in all the possible niches where bacteria thrive. In extreme environments like terrestrial thermal springs, viruses are one of the main predators of prokaryotes. Isolating and detecting viruses or bacteriophages from high-temperature hot springs has been challenging. The virus particles in hot springs have been observed by electron microscopy and propagated using cultivable bacteria and archaea isolated from these extreme environments. The typical life cycle of a bacteriophage falls into two categories: lytic and lysogeny. During host cell division, prophages can be induced under certain external conditions (e.g., physical UV radiation or mitomycin C induction). Various studies suggest that the hot spring viruses prefer lysogeny over the lytic cycle to survive extreme environments (Breitbart et al., 2004; Guajardo-leiva et al., 2021; Jarett et al., 2020; Pride and Schoenfeld, 2008; Sharma et al., 2018). Multiple viruses have been isolated from different hot springs around the world. But to our knowledge, there are no reports on isolating bacteriophage from Indian hot springs.

The isolation of a lytic phage infecting the thermophilic bacterium *Meiothermus* sp from the Rajwadi hot spring is described in this chapter. However, there are only two lytic thermophilic *Meiothermus* phages, MMP17 and MMP7, that are reported to date from Eryuan and Tengchong hot springs in Yunnan Province of China, respectively (Lin et al., 2011; Wang et al., 2020). The most challenging aspect of isolating thermophilic viruses was to avoid phage viability loss. Thus, the most challenging part of investigating these viruses from extreme environments is the maintenance of viruses-host in the laboratory conditions.

For studying the temperate bacteriophages, various prophage induction assays using mitomycin C, UV-induction and Polyaromatic hydrocarbons (PAH) were used (Chen et al., 2006) (as discussed in Chapter 3). After the prophage induction assay by mitomycin-C, the

presence of prophage was demonstrated by the reduced growth rate of the *Pseudomonas* sp. strain VB3 that correlated with the detection of the clearance zone in the spot assay.

This chapter addresses the isolation of bacteriophages following a preliminary screening (Chapter 3). We selected two bacterial isolates, *Meiothermus* sp. JE2 and *Pseudomonas* strain VB3, to isolate viruses from hot springs. The lytic phage infecting *Meiothermus* sp. JE2 and a temperate phage from *Pseudomonas* sp. strain VB3 were obtained from Rajwadi (Maharashtra) and Vashisht (Himachal Pradesh) hot springs, respectively.

Workflow:

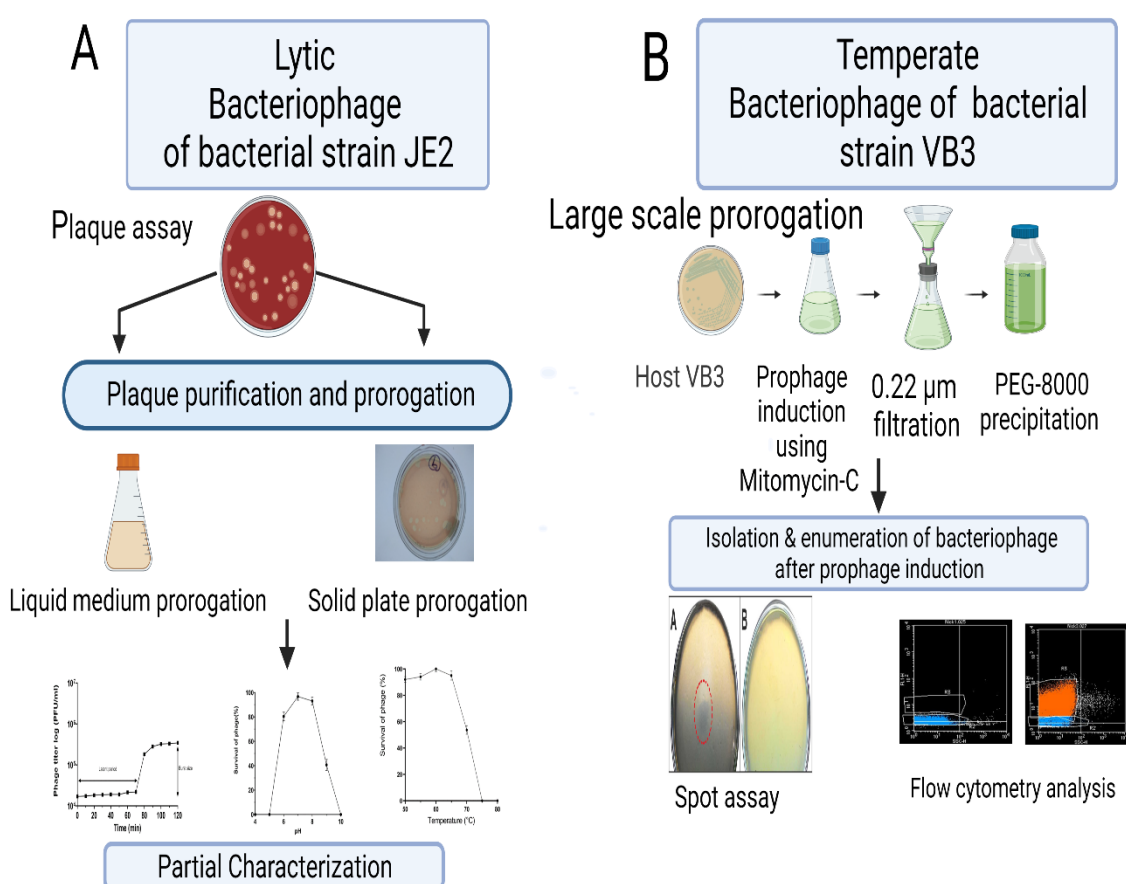


Figure 4. 1: Workflow for isolation of bacteriophage from hot spring. Created with BioRender.com (<https://www.biorender.com/>).

4.1 Materials:

0.22µm pore-size PVDF filter (Millipore, USA), *Meiothermus* sp. JE2, *Pseudomonas* sp. VB3, Gelrite, agar/agarose, Mitomycin-C (Sigma Aldrich), PEG 8000 (Hi-Media), SYBR Green I (Invitrogen, by Thermo Fisher scientific), Tris-EDTA buffer, Saline Magnesium (SM) buffer, BM media.

4.2 Methods

The workflow for isolating bacteriophage from hot springs in this chapter is illustrated in Figure 4. 1.

A) Lytic *Meiothermus* phage

B) Temperate *Pseudomonas* phage

4.2.1 A) Lytic bacteriophage of a thermophilic bacterium *Meiothermus* sp. JE2

4.2.1.1 Filtrations and concentration of water samples

Water sample was collected from Rajwadi hot spring, WCGP, India, and was immediately processed to remove bacteria by filtration through 0.22 µm filters (Middelboe et al., 2010; Steward and Culley, 2010). The viruses from the filtrate were concentrated by PEG 8000 and ultracentrifugation (Colombet et al., 2007) (See Chapter 3, Section 3.2.4). Screening for bacteriophage infection was conducted using the isolated hot springs bacterial cultures as hosts. The lytic infection was detected using a double gelrite overlay technique (Kropinski et al., 2009) and spot assay.

4.2.1.2 Bacteriophage isolation, infection and purification

Bacteriophages were screened against the thermophilic bacterial strain JE2 by double-layer overlay technique (Kropinski et al., 2009) after enriching the concentrated viral sample with 10 X BM broth. Briefly, 10 ml viral concentrate from the hot spring water was added to 100 ml BM broth inoculated with bacterial strain JE2 and incubated at 60°C for 24 h. This enriched culture was centrifuged at 10,000 x g for 15 min, filtered through a 0.22 µm pore-size PVDF filter (Millipore, USA), and stored as phage stock at 4°C.

4.2.1.3 Plaque assay (Double-layer overlay technique)

Bacteriophage infecting thermophilic bacterial strain JE2 was isolated by plating on semi-solid gelrite BM medium using the double-layer overlay technique. The BM medium was supplemented with 10 mM CaCl₂ for bacteriophage propagation. Briefly, in a 2 mL microcentrifuge tube, 400 µL of filtered enriched phage stock and 800 µL of logarithmic growing bacterial strain JE2 culture were mixed and incubated for 20 min at room temperature for adsorption. This mixture was added to a 4 mL top-gelrite medium (0.4%) and poured onto the culture plate with a bottom-gelrite layer (0.8%). After 48 h, clear plaques were observed at 60°C. Single clear plaques were picked up and suspended in SM buffer. This was followed by centrifugation at 10,000 x g for 10 min, and subsequently, the phage supernatant was passed through 0.22 µm pore-size filters to remove any bacterial debris, and the phage was stored at 4°C

4.2.1.4 Spot assay

Spot assay was also carried out on the lawn of bacterial strain JE2 to check the infectivity of the isolated phage particles (Sambrook, 1989). Single plaques were excised aseptically in SM buffer and used to reinfect bacterial strain JE2 to obtain plaques of similar morphology. After three rounds of bacterial infection, the phage from one such purified plaque was designated as vB_Msp_JE2 and used for further studies.

4.2.1.5 One-step growth curve

For a one-step growth curve, phage particles were mixed with 1 ml of fresh *Meiothermus* sp. strain JE2 (1x10⁶ CFU/mL) to obtain a multiplicity of infection (MOI) of 0.01 and incubated for 10 min at 60°C (Kropinski, 2018; Middelboe et al., 2010). After absorption, the mixture was centrifuged at 13000 x g for 2 min. The pellet was suspended in 1ml fresh BM, and this step was repeated thrice. To 50 ml fresh BM, 50 µL of the suspended pellet was added and incubated at 60°C. Phage titer was assayed by double-layer overlay technique at every 10-min interval throughout 2 h to estimate the number of phages as plaque-forming units (PFU). Burst size was calculated as the ratio of the final count of free phage particles to the initial count during the latent period.

4.2.1.6 Phage stability analysis

The thermostability of the phage was determined by pre-incubating the phage in the

BM medium at 50°C- 80°C for 30 min (Liu et al., 2006). Subsequently, the survival of the phage sample was also assayed by the double-layer overlay technique (Kropinski et al., 2009).

To assess the pH sensitivity, the phage (10^6 PFU/ml) was diluted 100-fold with SM buffer of different pH values and incubated for 30 min (Liu et al., 2006). The double-layer overlay technique determined the phage's survival rate (Kropinski et al., 2009).

Note:

Due to the loss in phage viability during long-term storage, the lytic *Meiothermus* phage vB_Msp_JE2 was only partially characterized, and additional studies were hindered. The establishment of host-phage interaction in a laboratory environment that does not resemble the hot spring state is the most challenging part for investigating these viruses from extreme environment.

4.2.2 B) Isolation of Temperate *Pseudomonas* phage

Bacterial growth and prophage induction conditions

Pseudomonas sp. VB3 was grown on fresh basal medium (g/l) [KH₂PO₄: 1.36; (NH₄)₂SO₄: 1.0; MgSO₄·7H₂O: 0.2; FeSO₄: 0.001; NaCl: 2.0; yeast extract, 1.0, pH 7] (Patel et al., 2006) supplemented with 0.2% sodium succinate at 30°C for 24 h.

Prophage Induction with Mitomycin-C

Pseudomonas sp. strain VB3 was grown in BM at 30°C for 24 hrs. The prophage induction with mitomycin-C of bacterial strain VB3 was done as mentioned by Chen et al., 2006., with slight modifications. Briefly, 10 ml of starter culture was transferred to 200 ml fresh basal medium. At 0.3 OD₆₀₀, the culture was divided into two sterile flasks (100 ml each). One flask was treated with mitomycin-C (final concentration 1 µg/ml), whereas the other served as a control. Both control and mitomycin-C treated bacterial cells were incubated, and cell density at OD₆₀₀ nm was monitored every two hours for 24 h. Cell debris was removed by centrifugation at 5000 × g for 20 min, and lysates was then passed through a 0.22 µm PVDF filter.

Spot assay

The presence of phage was evaluated by spot test. 10 µl of lysate was spotted on a previously pre-plated bacterial lawn plate (Sambrook, 1989). The control experiment involves spotting 10 µl of respective bacterial isolate (un-induced by mitomycin C) on the host bacterium's pre-plated lawn.

Large Scale prophage induction

Mitomycin C-mediated prophage inductions were performed on a large scale (1000 ml) to obtain a large quantity of cell lysates. Briefly, this involved the addition of mitomycin C (final concentration 1 µg/ml) to early exponential phase cultures grown in a 1000 ml volume of BM broth, followed by overnight incubation at room temperature. Cell debris was removed by centrifugation at 7560 × g for 20 min, and lysates was passed through a 0.22 µm PVDF filter (Millipore).

Precipitation of phage particles using PEG 8000

The induced phage particles were precipitated by overnight incubation at 4°C with 10% PEG 8000 supplemented with 1M NaCl (Sambrook, 1989). One liter of mitomycin-C induced phage lysate was centrifuged at 12,000 x g for 15 min. The supernatant was filtered through a 0.22µm-pore size filter (type HA; Millipore) to remove bacterial host cells and cellular debris. Phage particles from the filtrate were treated with PEG 8000 (final concentration: 100 g/l) and incubated overnight at 4°C followed by centrifugation at 12,000 x g for 20 min to obtain phage pellet. The phage pellet was re-suspended in 10 ml SM buffer (100 mM NaCl, 50 mM Tris-HCl, and 8.1 mM MgSO₄, pH 7.5) and incubated at 4°C.

Preliminary enumeration of induced phage particles using flow cytometry (FCM) analysis

Following large-scale prophage induction, as described above, released phage particles were detected by Flow Cytometry (FCM) analysis. A standard protocol (Brussaard, 2004) was used for the flow cytometric enumeration of induced phage particles. Briefly, 1 ml of 0.22 µm-filtered water sample (per se or diluted 1:1 with Tris-EDTA buffer) was stained with SYBR Green at a final concentration of 10⁻⁴ of the commercial stock. Tris-EDTA (pH 8) buffer was set as a negative control. The staining process of the mitomycin-C induced phage lysates took place at 80°C in the dark for 10 minutes. The mitomycin-C induced phage lysate and subsequent PEG-precipitated phage lysate were examined on a BD FACS Calibur flow cytometer after a 5-minute cooling time, which was equipped with a 15 mW 488 nm air-cooled argon-ion laser and a standard filter setup. Green fluorescence was set up as the trigger. Flow cytometry acquisition was set on the FL1 channel (threshold value-500), and analysis was carried out for 30 seconds (Flow rate: 60 µL min⁻¹). A combination of FSC (forward scatter) and SSC (side scatter) was used to distinguish virus particles from bacteria if present in the sample. The data were analyzed with the BD CellQuest™ Pro software. The phage quantification (the number of virus particles in a sample was obtained by applying the formula shown below

$$\text{Number of virus particles (per ml)} = \frac{\text{Total events}}{\text{Time} \times \text{flow rate calibration factor}} \times 1000$$

4.3 Results and Discussion

4.3.1 Lytic bacteriophage of a thermophilic bacterium *Meiothermus* sp. JE2

Bacteriophage Isolation, infection, and Purification

After filtration of the water sample through a 0.22 μm pore-size filter and subsequent concentration by ultracentrifugation, the concentrated water sample from Rajwadi was screened for potential bacteriophage that would use *Meiothermus* sp. strain JE2 as a host. A lytic bacteriophage designated as vB_Msp_JE2 was isolated from Rajwadi hot spring. The bacteriophage vB_Msp_JE2 produced a clear plaque of a 1-3 mm diameter after 48 h at 60°C (Figure 4. 2). Currently, there are only two lytic thermophilic *Meiothermus* phages MMP17 and MMP7, that are reported to date from Eryuan and Tengchong hot springs in Yunnan Province of China, respectively (Lin et al., 2011; Wang et al., 2020).



Figure 4. 2: vB_Msp_JE2 clear plaques detected by the double-layer overlay method using the red-pigmented thermophilic *Meiothermus* sp. strain JE2 as host.

One-step growth curve

A one-step growth curve experiment determined the latent period and burst size of phage vB_Msp_JE2. The latent period of the phage was 70 min, and the burst size was 20 (Figure 4. 3). Long latent periods in the 60-80 min range were also reported for other thermophilic *Meiothermus* and *Thermus* phages (Lin et al., 2010, 2011; Naryshkina et al., 2006). Additionally, the burst size of 20 for the phage isolated in the present study was almost similar to that reported for *Meiothermus* phage MMP17 (Lin et al., 2011). In contrast, other thermophilic *Thermus* phages stated a larger burst size of approximately 80 for ϕYS40 (Naryshkina et al., 2006), 109 for ϕOH3 (Nagayoshi et al., 2016) and 200 for TSP4 (Lin et al., 2010).

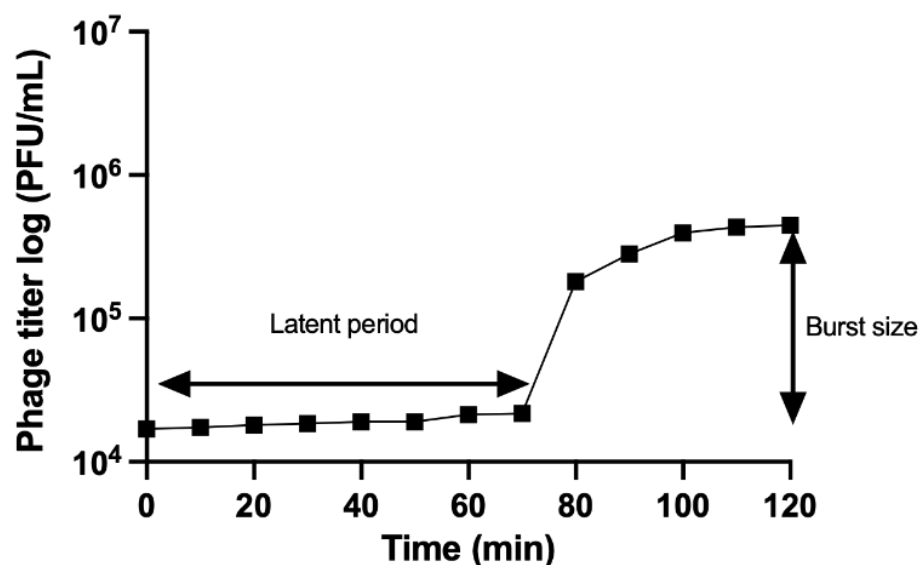


Figure 4. 3: One-step growth curve of phage vB_Msp_JE2.

Thermostability and pH Sensitivity

Phage vB_Msp_JE2 was stable at 60°C (Figure 4. 4). *Meiothermus* phage MMP17 (Lin et al., 2010) was stable between 55 to 60°C whereas *Thermus* phage TSP4 (Lin et al., 2010), and *Geobacillus* phage GVE1 (Liu et al., 2006) were stable at 60°C. In contrast, *Thermus* phages such as φOH3 (Nagayoshi et al., 2016) and φYS40 (Naryshkina et al., 2006) were stable at 70°C.

The thermostability of the phage vB_Msp_JE2 is correlated to the growth of the host bacterium *Meiothermus* sp. JE2 isolated from Rajwadi hot spring water having a temperature of 55°C. Phage vB_Msp_JE2 was most stable at pH 6.0-8.0 and was sensitive to acidic pH (Figure 4. 5). *Meiothermus* phage MMP17 was also described as stable at pH 6-7 (Lin et al., 2011). These results suggest that the geographic origin of the phage influences the temperature and pH stability of the phage.

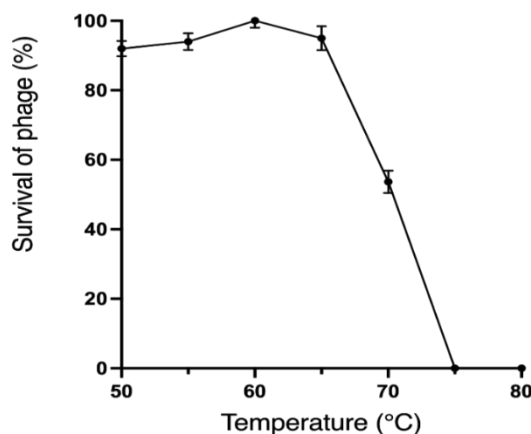


Figure 4. 4 : Thermostability of bacteriophage vB_Msp_JE2

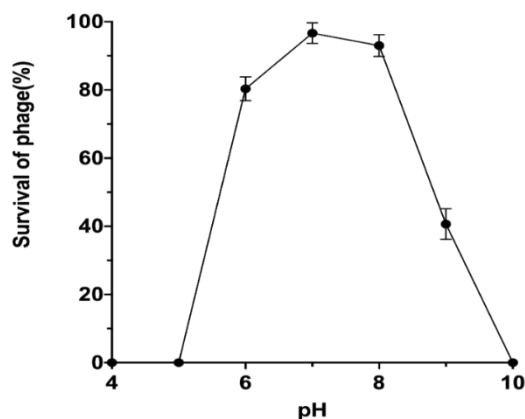


Figure 4. 5: pH sensitivity of bacteriophage vB_Msp_JE2

Meiothermus species have been reported from hot springs and artificial thermal environments like paper mills (Ekman et al., 2007; Kolari et al., 2003) and spent nuclear fuel (Masurat et al., 2005). *Meiothermus* sp. was first described as a paper machine contaminant by Kolari et al., 2003, a common biofouling causing biofilm formation in the paper industry. Compared to the Himalayan hot spring, there needs to be more focus on microbial studies at WCGP. Thus, the present study is the first preliminary study on isolating culturable *Meiothermus* sp. and its lytic phage vB_Msp_JE2 from Rajwadi hot spring, WCGP, India. Due to the loss in phage viability during long-term storage, the lytic *Meiothermus* phage vB_Msp_JE2 was only partially characterized, and additional studies were hindered.

4.3.2 Isolation of temperate *Pseudomonas* bacteriophage.

Prophage Induction with Mitomycin C

After mitomycin-C treatment of *Pseudomonas* sp. strains VB3, cell lysis was detected by a decrease in the optical density of the bacterial culture. Mitomycin-C treatment induced decreased growth curves was indicative of lysogen induction in *Pseudomonas* sp. VB3 (Figure 4. 6). An initial increase followed by a sharp decrease in optical density in the mitomycin-C treatment 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 (Figure 4. 7).

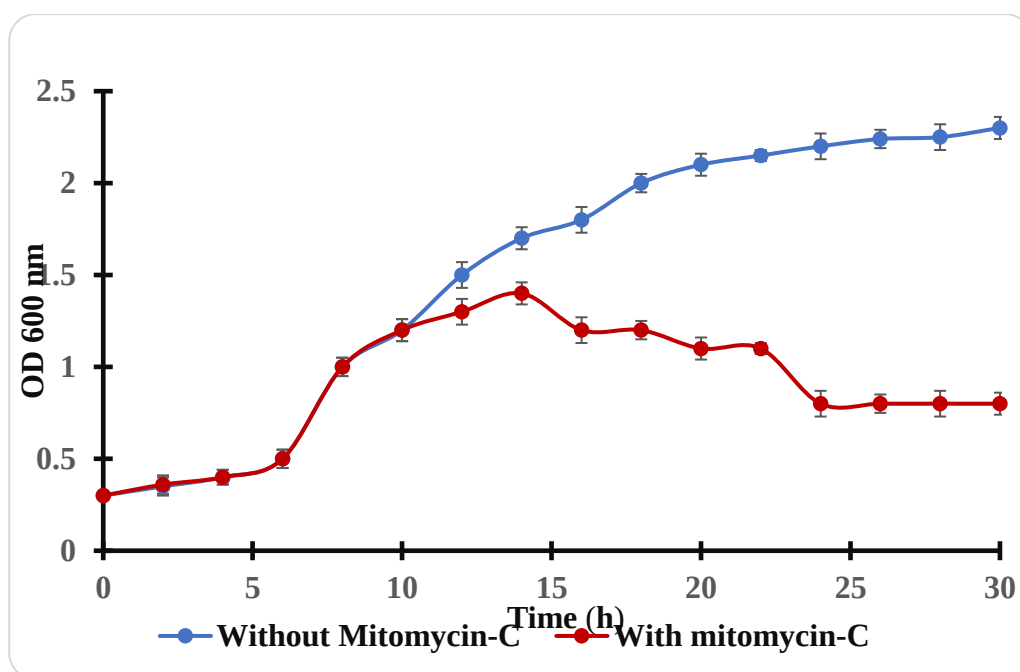


Figure 4. 6: Prophage induction of bacterial strain VB3 with and without Mitomycin-C.

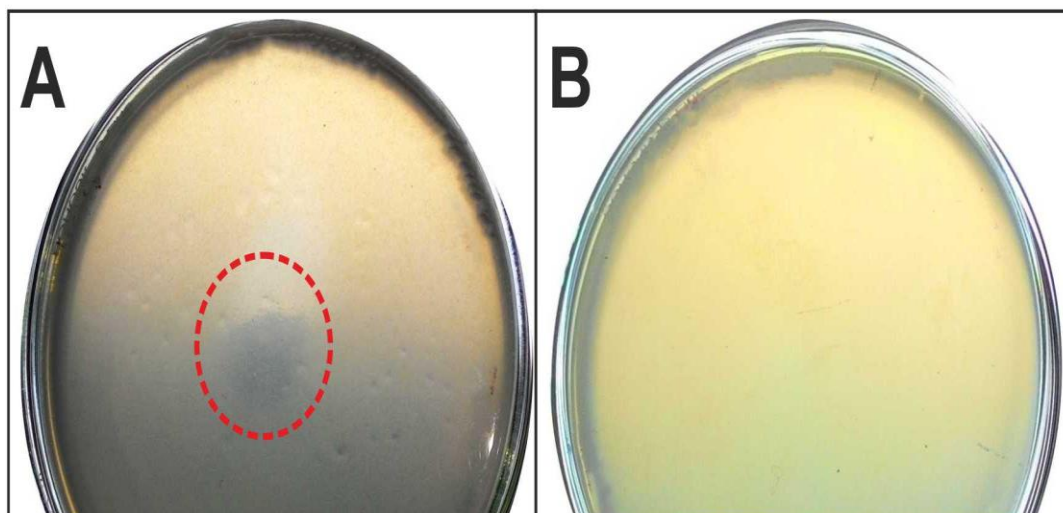


Figure 4. 7: Spot test after induction with Mitomycin-C on the lawn of *Pseudomonas* sp. VB3.

(A) Mitomycin-C treated, clearance zone is circled, (B) Control, without Mitomycin-C.

When tested the potential infectivity of the released phage particles by the double agar layer method, but no obvious plaques appeared on the soft agar. Such phenomena have been reported previously (Cornuault and Moineau, 2021) due to superinfection exclusion. Our indicator bacterium had developed resistance to the released phage due to the defence system. These systems might play a role in preventing plaque formation or there was no infectivity in the phage lysate or else the copy number of temperate phage was low (Vervliet et al., 1975).

Flow cytometry analysis.

A typical plot obtained in the FCM measurement is depicted in Figure 4. 8. This is a plot of fluorescence (proportional to the DNA content of the particle) against side scatter (proportional to particle size). The region designated R3 represents fluorescence and particle size values typical of the bacteriophage, as per the existing literature (Brussaard, 2004). The signal obtained in this region is quantified to arrive at a count of phage particles in the sample.

FCM analysis was used to count phage particles after prophage induction with the mitomycin-C treatment of strain VB3 culture. The results of this investigation have clearly shown that phage particles propagated and concentrated on a large scale using PEG 8000

outperformed non-PEG precipitated VB3 Mito-C samples. The virus count in the VB3 Mito-C PEG precipitated sample was 2.39×10^6 VLP/mL compared to 1.5×10^6 VLP/mL in the VB3 Mito-C non-PEG precipitated lysate (Figure 4. 9). Overall, large-scale concentration by PEG 8000 precipitation indicated the presence of prophage after induction. Further characterization of the induced phage particles using TEM imaging and WGS of bacteria will provide a deeper understanding of the possibility of temperate prophages induced by Mitomycin-C.

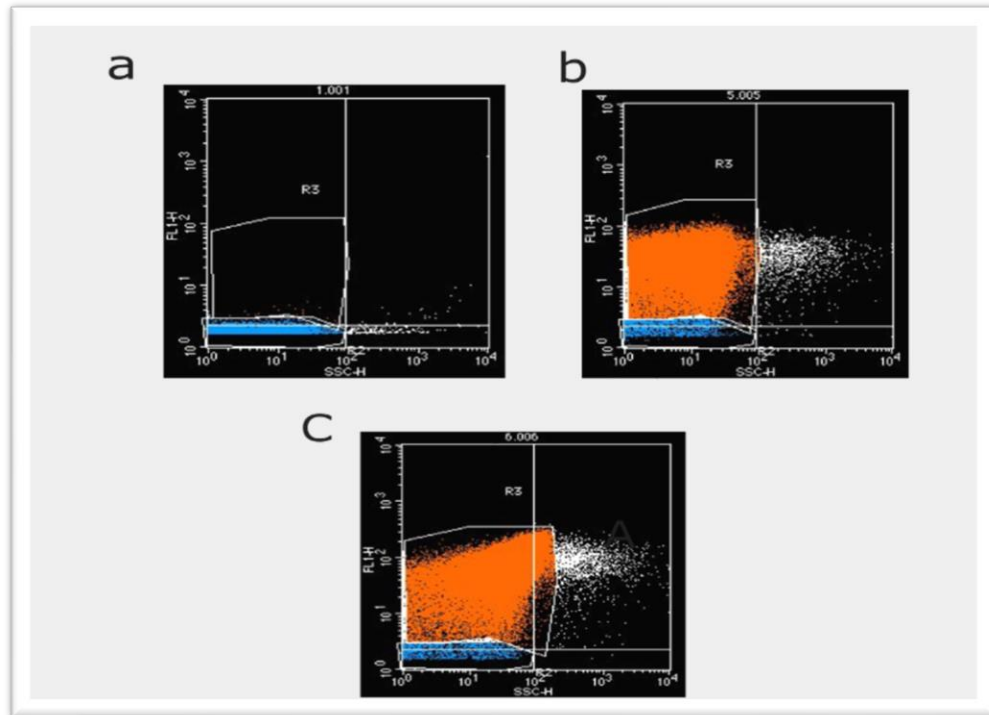


Figure 4. 8: Preliminary enumeration of phage particles after prophage induction with mitomycin C and concentrated through PEG 8000. a) Negative control, b) VB3 Mito-C, c) VB3 Mito-C + PEG 8000.

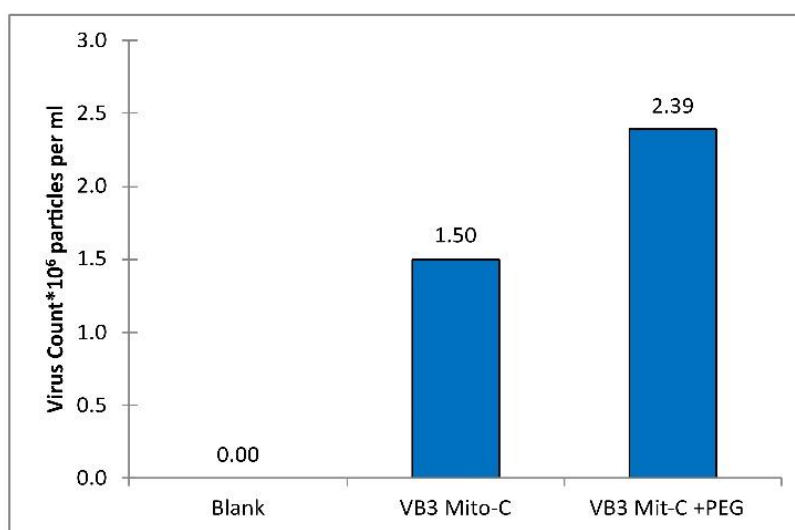


Figure 4. 9: Preliminary enumeration of phage particles after mitomycin-C treatment and concentrated through PEG 8000.

Identifying and enumerating phage particles using FCM was possible, but it was not possible to tell whether such particles represented one or numerous different phages. FCM was a more reliable approach for detecting prophage released than plaque assays when no suitable host was available for plaque assays, as in this case (the induced temperate phage was unable to infect its original *Pseudomonas* lysogen host) or epifluorescence microscopy (EFM) (when the titre of the phage is low). Despite the fact that the double-agar plaque assay is probably the most efficient method for detecting and quantifying intact (i.e., lytic or infectious) phages, it cannot be used when an appropriate host for a released (pro)phage cannot be discovered (Oliveira et al., 2017).

Several methods have been used such as bacterial genome sequencing analysis (Chopin et al., 2001), real-time PCR (Lunde et al., 2003), TEM and EFM analysis have all been used to detect (induced) temperate phages. Lambeth and coworkers established the use of FCM-based assays for dengue virus particle quantifications as a suitable technique for titrating clinical isolates of dengue that frequently do not produce distinct plaques (Lambeth et al., 2005).

4.4 Conclusion:

Isolation of lytic phage infecting *Meiothermus* sp. JE2 from Rajwadi (Maharashtra) was partially characterized. We could not further characterize due to loss in phage viability during long-term storage. Characterizing lytic *Meiothermus* phage vB_Msp_JE2 posed challenges, particularly due to its origin in extreme environments like hot springs. The difficulty lies in replicating the natural conditions of these environments in a laboratory setting. Hot springs harbor unique physical and chemical parameters, such as high temperatures, high mineral content, and pH, which significantly influence microbial life and viral dynamics. Despite preservation methods, some phages may still lose viability over time due to factors like DNA degradation, mutations, or contamination. Also, challenges, particularly due to its origin in extreme environment. This can make it challenging to obtain intact DNA for sequencing, particularly for long-term storage periods.

A lysogenic bacteria *Pseudomonas* train VB3, was isolated from Vashisht hot spring (Himachal Pradesh), India. After the prophage induction assay by mitomycin-C, the presence of prophage was demonstrated by the reduced growth rate of the *Pseudomonas* sp. strain VB3 that correlated with the detection of the clearance zone in the spot assay. The phage particles induced were successfully precipitated and concentrated by PEG 8000 on a large scale. Preliminary enumeration studies by flow cytometry analysis also indicated prophage induction after mitomycin-C treatment. For this next chapter, we will describe the WGS of this potential lysogenic bacterium in order to mine the prophages produced after mitomycin C induction.

CHAPTER 5

Whole genome sequencing of a potential lysogenic bacterium.

The chapter describes the whole-genome sequencing of a lysogenic Pseudomonas sp. strain VB3 isolated from the Vashisht hot spring, Himachal Pradesh, India.

Bacteriophages can sustain two life cycles (lytic and lysogeny). Temperate (lysogenic) phages follow a life cycle that involves the integration of their genome into the host organism's genome to become a prophage. Analysis of bacterial genomes has revealed that prophages are common and have diverse genetic elements. Whole-genome sequencing reveals that prophage DNAs are widespread among bacterial genomes, even up to 20% of the host genome content (Casjens, 2003). This objective was designed to identify the prophage region in the lysogenic bacteria isolated from the Vashisht hot spring. Thus, the potential lysogenic *Pseudomonas* sp. strain VB3 was subjected to NGS Whole-genome sequencing.

The genus *Pseudomonas* was first discovered by Migula in 1894, and around 586 species have been described in the List of Prokaryotic Names with Standing Nomenclature (LPSN) (<http://www.bacterio.net>) (accessed on 27 Jan, 2023) (Parte et al., 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 and Lavigne, 2010).

Two novel *Pseudomonas* species have been reported and studied from Taiwan's hot springs (Liu et al., 2013; Wong and Lee, 2014). Additionally, metagenomic studies have revealed *Pseudomonas* species from Indian hot springs (Saxena et al., 2017; Poddar and Das, 2018; Sharma et al., 2018). Although lysogeny is reported as a favorable lifestyle for bacteriophages in hot springs (Breitbart et al., 2004), viruses from Indian hot springs are

least explored. The previous chapters described the isolation and screening 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.

Because strain VB3 was identified as a probable lysogenic bacterium among the isolated bacterial strains, WGS was selected for whole genome sequencing (WGS) to identify putative prophage regions in the host genome.

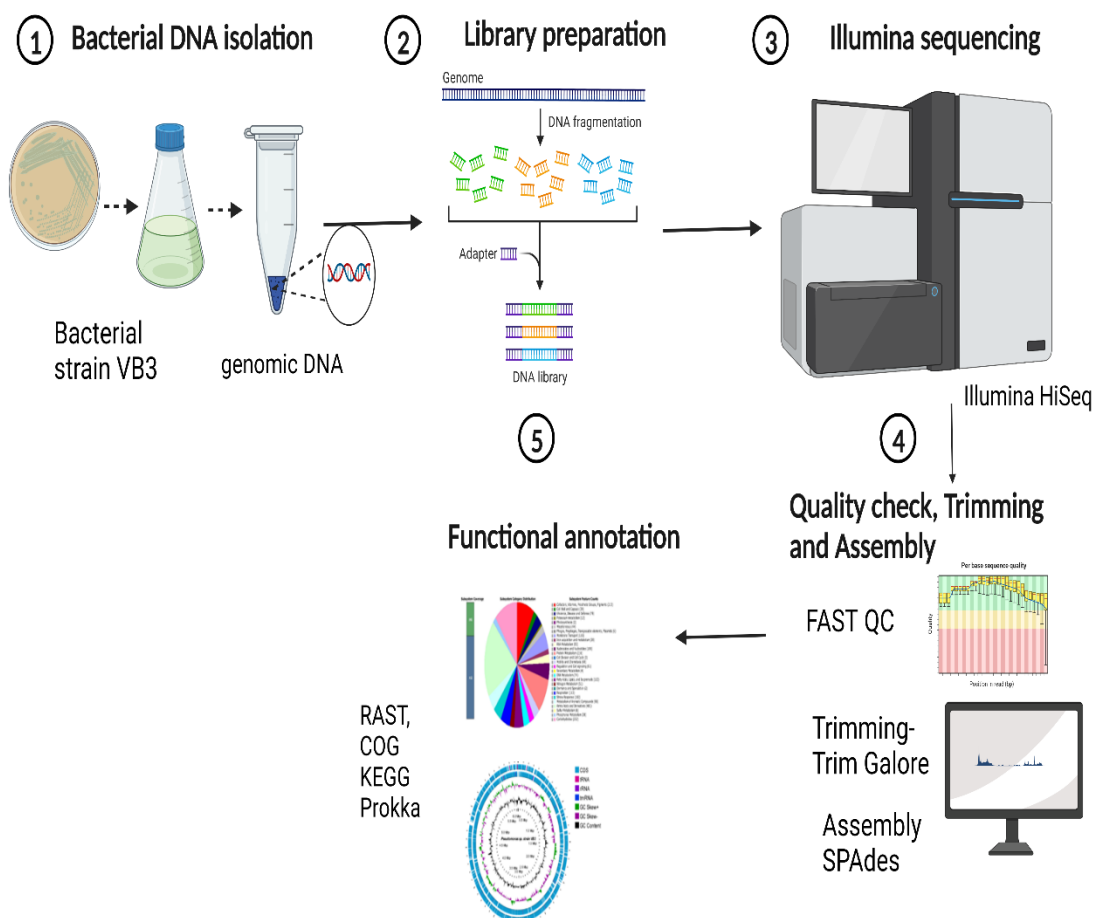
Workflow:

Figure 5. 1: Workflow of the WGS of a potential lysogenic bacterium.

Created with BioRender.com (<https://www.biorender.com/>).

5.1 Materials:

Bacterial strain VB3, GeneiPure™ bacterial DNA purification Kit (Cat # 2115900031730 GeNei, Bangalore, India), Nextera XT Library Prep Kit (Cat #FC-131-1024), Qubit™ dsDNA BR Assay Kit (cat. #. Q32850, Invitrogen, Carlsbad, CA).

5.2 Methods:

5.2.1 Bacterial strain VB3 sample

The lysogenic bacterial strain VB3 was selected for WGS and maintained in Basal medium (BM) agar plates (g/L) containing KH_2PO_4 : 1.36, $(\text{NH}_4)_2\text{SO}_4$:1.0, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$: 0.2, FeSO_4 : 0.001, NaCl : 2.0, yeast extract, 1.0 and supplemented with 0.2% sodium succinate and were incubated at 30°C. Figure 5. 1 describes the workflow of the WGS carried out for the potential lysogenic *Pseudomonas* sp. strain VB3 isolated from the Vashisht hot springs, Himachal Pradesh, India.

5.2.2 Genomic DNA extraction and 16S rRNA gene sequencing

Bacterial strain VB3 was inoculated into BM and was incubated at 30°C for 24 h. The bacterial genomic DNA of strain VB3 was extracted using the GeneiPure™ bacterial DNA purification Kit (GeNei, Bangalore, India). The eluted DNA was stored at -20°C. Later, isolated DNA was quantified using the Qubit dsDNA HS Assay Kit and the Qubit 2.0 Fluorometer (Invitrogen, USA) as per the manufacturer's instruction. One µl of genomic DNA sample was diluted 200-fold in Qubit dsDNA HS assay buffer in clear plastic Qubit assay tubes and measured using a Qubit fluorometer.

To confirm the identity of bacterial strain VB13, 16S rRNA gene sequencing was determined after PCR amplification of genomic DNA using 27F and 1429R universal primers (Frank et al., 2008). The amplification was conducted in a 50 µL reaction mixture containing approximately 100 ng bacterial genomic DNA, 10X polymerase buffer, 5 µL dNTP mix, 0.25 µM of each forward and reverse primer, 1.5 Units of Taq polymerase. The reaction mixture was heated at 95°C for 5 min and immediately subjected to 30 cycles of

denaturation at 95 °C for 1 min, annealing at 55 °C for 1 min, extension at 72 °C for 1 min, and final extension step at 72 °C for 10 mins. The amplified PCR product was resolved on 0.7% agarose gel to check for the PCR amplified product. The amplified PCR product was purified using a QIAquick PCR purification kit (Qiagen Inc, Hilden, Germany). The purified product was sequenced by 3730XI (96 capillary) electrophoresis instruments (Applied Biosystems, Foster City, CA, USA).

5.2.3 De-Novo whole genome sequencing

Library preparation

For next-generation sequencing, the genomic library was prepared from bacterial strain VB3 genomic DNA using the Nextera XT Library Prep Kit as per the manufacturer's instructions (Cat #FC-131-1024, Illumina, San Diego, CA) at Genotypic Technology Pvt Ltd, Bengaluru. One ng of Qubit quantified genomic DNA was tagmented (fragmented and adaptor tagged) using Amplicon Tagment Mix provided in the Nextera XT Kit. The adapter-tagged DNA was subjected to 12 cycles of Indexing-PCR (72°C for 3mins followed by denaturation at 95°C for 30 sec, cycling followed by 95°C for 10 sec, 55°C for 30 sec, 72°C for 30 sec) and 72°C for 5 mins) to enrich the adapter-tagged fragments. The PCR product was purified using magnetic beads JetSeq™ Clean (Bioline, Cat No: BIO-68031). The Illumina-compatible sequencing libraries were quantified by Qubit fluorometer (Thermo Fisher Scientific, MA, and USA), and their fragment size distribution was analysed on Agilent 2200 Tape Station.

Illumina Sequencing quality check

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 base pair reads with a mean fragment length of ~358 bp were obtained. The quality check, trimming, assembly, and annotation for the genome was done on the Galaxy Australia version 20.05 web server platform (<https://usegalaxy.org.au/>) (Jalili et al., 2020). The Illumina pair-ended reads were quality-checked using Fast QC v 0.11.8 (Andrews, 2010). Trim Galore (v.0.6.3) was used to trim the quality reads. (<https://github.com/FelixKrueger/TrimGalore/>).

Assembly

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 k-mer value was set to 77 K-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).

Annotation

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 support 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 circular genomic map was constructed by using the Proksee server (<https://proksee.ca/>) (Stothard et al., 2019). The genome was annotated by the Global catalog of type strain (Shi et al., 2021) (gcType) (<http://gctype.wdcm.org/>), and genes predicted were annotated using several databases, including KEGG, COG, and CARD analysis. The Clustered regularly interspaced short palindromic repeats (CRISPR) and associated proteins (Cas) from the CRISPR-Cas systems were determined using CRISPR-Cas FINDER (<https://crisprcas.i2bc.paris-saclay.fr/>) (Couvin et al., 2018).

Prophage Prediction:

Prophage prediction was made 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.

5.3 Results and Discussion

5.3.1 Bacterial selection

The isolated bacterial strains in the present studies were investigated for possible prophage induction by exposure to UV-light radiation, polycyclic aromatic hydrocarbon (PAH), or mitomycin-C (Chapter 3). The PAH treatment or exposure to UV-light radiation failed to induce any prophages from the bacterial strains. Out of the 144 bacterial strains, only bacterial strain VB3 from Vashisht hot Springs, Himachal Pradesh, India, demonstrated a decrease in cell density after exposure to Mitomycin-C (Chapters 3 and 4). As bacterial strain VB3 was an evidential lysogenic strain, it was selected for WGS to identify the putative prophage regions.

5.3.2 16S rRNA gene sequencing

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%), *P. libanensis* strain CIP 105460^T (99.48%), *P. mucidolens* strain LMG 2223^T (99.40%), *P. azotoforman* strain DSM 18862^T (99.33%), *P. carnis* strain B4-1^T (99.33%), *P. paralactis* strain DSM 29164^T (99.33%), *P. lactis* strain DSM 29167^T (99.25%), and *P. bernneri* strain CFML 97-391^T (99.25%) and lower sequence similarity to other species of the genus *Pseudomonas*.

5.3.3 De Novo whole genome sequencing and annotation

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) on 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% compared 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 5.1). The general features of the

Pseudomonas sp. strain VB3 are summarised in Table 5. 2, ensuring the Minimum information about any (X) sequence as per the (MIxS) standard checklist.

Table 5. 1: Basic statistics of the draft genome of *Pseudomonas* sp. strain VB3.

Attributes	Value
Total number of contigs	256
DNA coding number of bases (bp)	6,065,353
GC Content	60.5%
Longest sequence (bp)	197576
N50 (bp)	56396
Total number of coding sequences	5607
Number of subsystems	388
Number of RNAs	62
Number of rRNA genes	4
Number of tRNA genes	58

Table 5. 2: General information of *Pseudomonas* sp. strain VB3, according to the recommendation of Minimum Information about any (X) sequence (MIxS) standard checklist.

Items	Description
<u>General Features</u>	
Classification	Domain: <i>Bacteria</i> Phylum: <i>Proteobacteria</i> Class: <i>Gammaproteobacteria</i> Order: <i>Pseudomonadales</i> Family: <i>Pseudomonadaceae</i> Genus: <i>Pseudomonas</i>
Gram staining	Negative
Cell shape	Rod-shaped
Motility	Motile
MIxS data	
Submitted to INSDC	JAPQMF000000000 (GenBank)
Investigation type	Bacteria
Project name	<i>Pseudomonas</i> sp. VB3, Whole genome shotgun sequencing project
BioProject	PRJNA899365
BioSample	SAMN31655254
Geographic location	India: Himachal Pradesh
Collection Date	2015-05-12
Environment (biome)	Hot spring
Environmental material	Water
Depth	0-2 m
Relationship with oxygen	Aerobic
Sequencing Method	Illumina HiSeq
Assembly	SPAdes (v 3.12.0)
Finishing strategy	Draft, over 100X coverage, 256 contigs

RAST annotation

The genome sequence of *Pseudomonas* sp. VB3 includes 5607 protein-coding genes, four rRNAs, and 58 tRNAs in 256 contigs. A total of 5607 protein-coding genes were assigned to 388 subsystems using the RAST (Figure 5. 2).

Prokka annotation

The number of protein-coding sequences within the genome is about 5607, with 62 non-coding RNAs including tRNA (58) and rRNA (4) (Figure 5. 3).

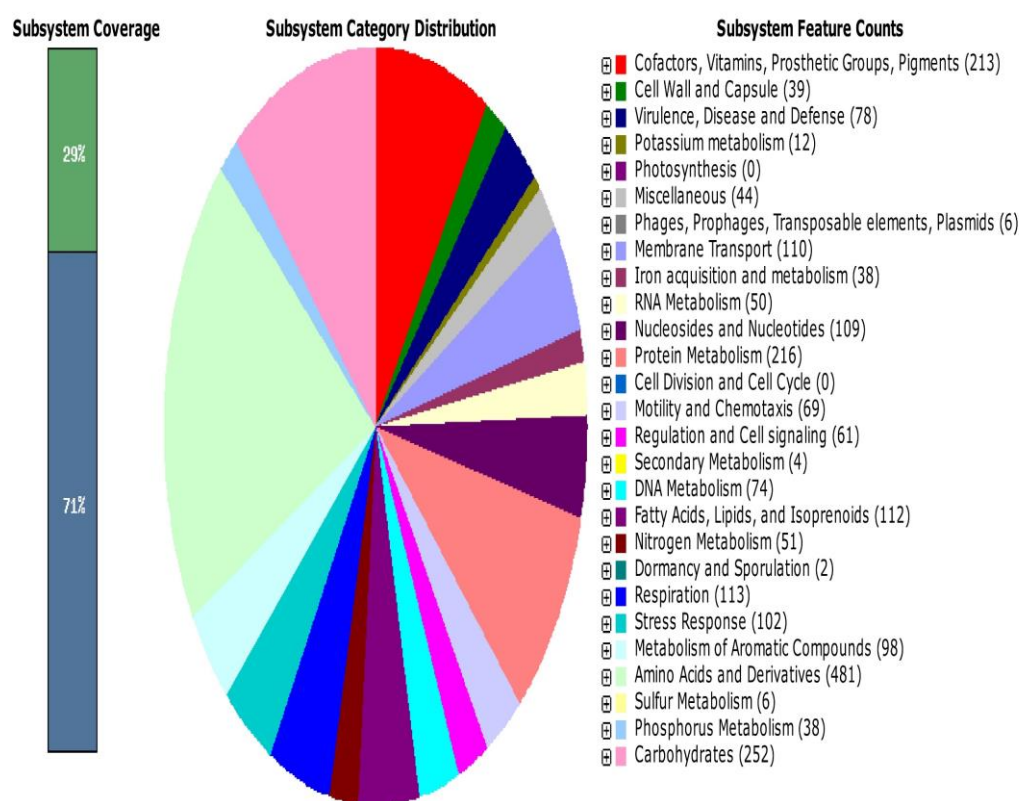


Figure 5. 2: RAST annotation of *Pseudomonas* sp. VB3.

The subsystem category distribution shows the percentage distribution of the genes in different pathways labelled with different colors. The number of genes in each category is mentioned in the bracket.

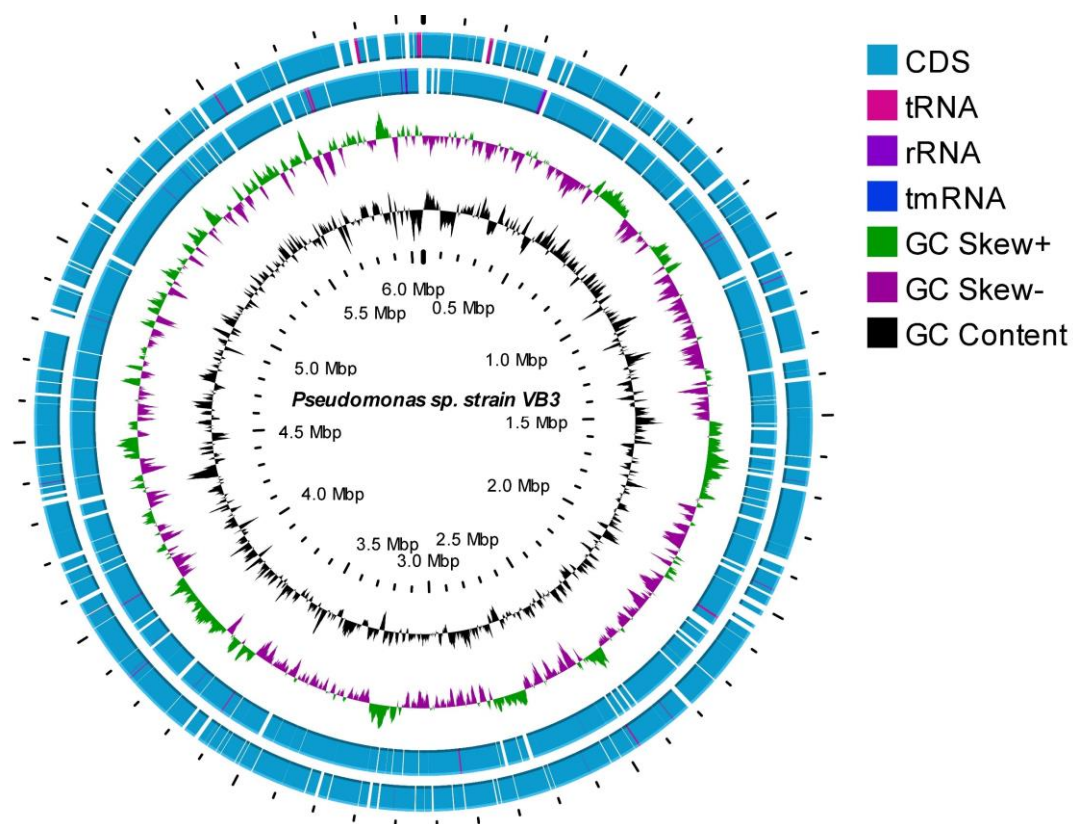


Figure 5. 3: Circular representation of *Pseudomonas* sp. strain VB3 contigs using Proksee (<https://proksee.ca/>).

The scale is shown in megabases on the inside central circle. Moving outward, two outer blue circles show forward and reverse strand CDS, respectively. The tRNAs (magenta), rRNAs (purple), and tmRNA (dark blue) are shown in CDS circles. The next circle shows the GC content (black) followed by the GC Skew (dark green and purple).

TYGS (Type strain Genome server)

The genome's maximum values of average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH) were 90.45 % and 49.5%, respectively (Table 5. 3). The ANI threshold for species delineation is recommended at 95–96% (Richter and Rosselló-Móra, 2009), and ANI between *Pseudomonas* sp. strain VB3, and phylogenetically closest relatives are $\leq 90\%$. The ANI values were significantly lower than the threshold for describing prokaryote species (95-96%). The dDDH values of *Pseudomonas* sp. strain VB3 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 (Figure 5. 4). However, the dDDH analysis obtained from TYGS stated that the strain VB3 is separate from *P. gessardii* DSM17152 as the dDDH values between them were 49.5%, which is less than the recommended threshold (<70%). The dDDH values of other *Pseudomonas* sp. were $\leq 29.1\%$ is much lower than the species threshold of 70% recommended for species differentiation (Meier-Kolthoff et al., 2013) (Table 5. 3).

Table 5. 3: The digital DNA-DNA hybridization (dDDH) scores for *Pseudomonas* sp. strain VB3 genome and other related organisms at the Type Strain genome server.

Subject strains	dDDH (d4, in %)
<i>Pseudomonas gessardii</i> DSM 17152	49.5
<i>Pseudomonas proteolytica</i> LMG 22710	45
<i>Pseudomonas brenneri</i> DSM 15294	43.4
<i>Pseudomonas brenneri</i> JCM 13307	43.4
<i>Pseudomonas mucidolens</i> LMG 2223	30.9
<i>Pseudomonas yamanorum</i> LMG 27247	30.5
<i>Pseudomonas fildesensis</i> KG01	30.2
<i>Pseudomonas aylmerensis</i> S1E40	30
<i>Pseudomonas allii</i> MAFF 301514	29.8
<i>Pseudomonas azotoformans</i> LMG 21611	29.7
<i>Pseudomonas lactis</i> DSM 29167	29.5
<i>Pseudomonas carnis</i> B4-1T	29.4
<i>Pseudomonas libanensis</i> DSM 17149	29.3
<i>Pseudomonas paralactis</i> DSM 29164	29.3
<i>Pseudomonas orientalis</i> DSM 17489	29.3
<i>Pseudomonas paracarnis</i> V5/DAB/2/5	29.1
<i>Pseudomonas synxantha</i> DSM 18928	29.1

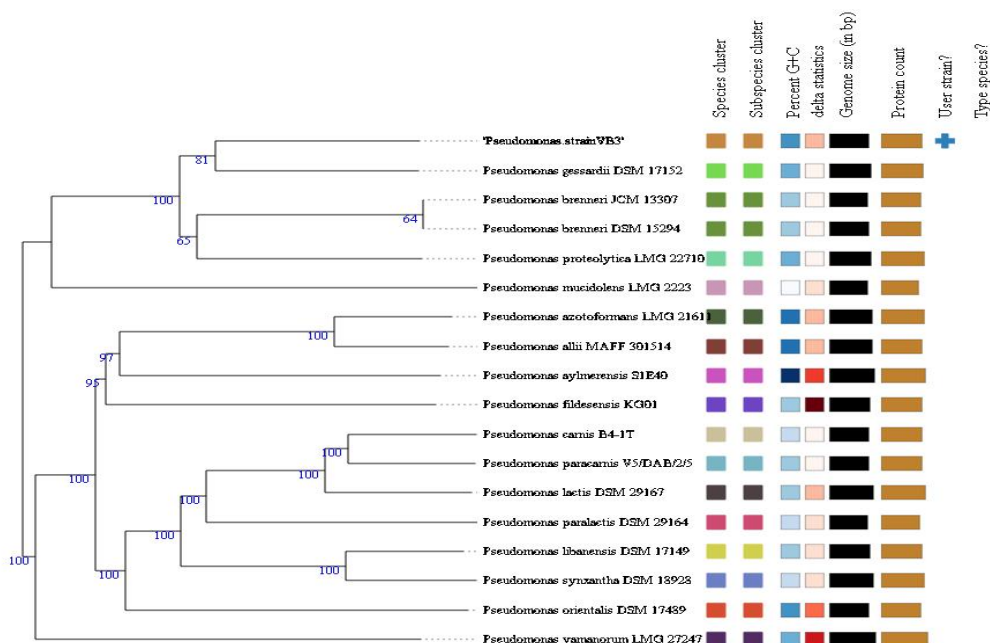


Figure 5. 4: Phylogenomic Tree based on *Pseudomonas* sp. strain VB3 genome sequences in TYGS (Type strain Genome server) (<https://tygs.dsmz.de/>).

Inferred with FastME 2.1.6.1 from Genome BLAST distance Phylogeny (GDPB) approach; distances calculated from the genome sequence. The numbers depicted at the branch point are pseudo-bootstrap support values from 100 replication (values >60% shown)

Secondary metabolite biosynthetic gene clusters

Genome mining revealed that *Pseudomonas* sp. strain VB3 harbors 14 gene clusters coding for the biosynthesis of bioactive secondary metabolites (Table 5. 4). Based on antiSMASH 6.1.1 analysis, the genome of *Pseudomonas* sp. strain VB3 harbours NRPS (Non-ribosomal peptide synthase), aryl polyene, RiPP-like, redox-cofactor, betalactone, 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 might 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) that is 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.

Table 5. 4: Secondary metabolite biosynthetic gene clusters in *Pseudomonas* sp. strain VB3 as predicted by antiSMASH 6.1.1.

Clusters	Type	Most similar know clusters with MIBiG	MIBIG-BGC-ID (Gene location predicted by antiSMASH)
Cluster 1 (Region 5.1)	Aryl polyene	APE Vf biosynthetic gene cluster (35% of gene show similarity) ^b	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	Hserlactone	N.D	17,929 - 38,507 nt.

(Region 47.1) Cluster 10	NAGGN	N.D	17,929 - 38,507 nt.
(Region 70.1) Cluster 11	NRPS	Pyoverdine biosynthetic gene cluster (9% of gene show similarity)	BGC0000413 1 - 28,761 nt.
(Region 73.1) Cluster 12	NRPS-like	N.D	1 - 17,231 nt.
(Region 105.1) Cluster 13	NRPS	Pyoverdine biosynthetic gene cluster (3% of gene show similarity)	BGC0000413 1 - 15,624 nt.
(Region 111.1) Cluster 14	NRPS	N.D	1 - 7,765 nt.
(Region 151.1)			

N.D: no defined known clusters are available; NAGGN, N-acetyl glutaminyl glutamine amide; NRPS, non-ribosomal peptide synthase; RiPP-Like, other unspecified ribosomally synthesized and post-transcriptionally modified peptide product; APE Vf aryl polyene cluster.

^a '-' indicates no match found.

^b The similarities indicate the percentage of genes in BGC similar to other known clusters predicted by antiSMASH.

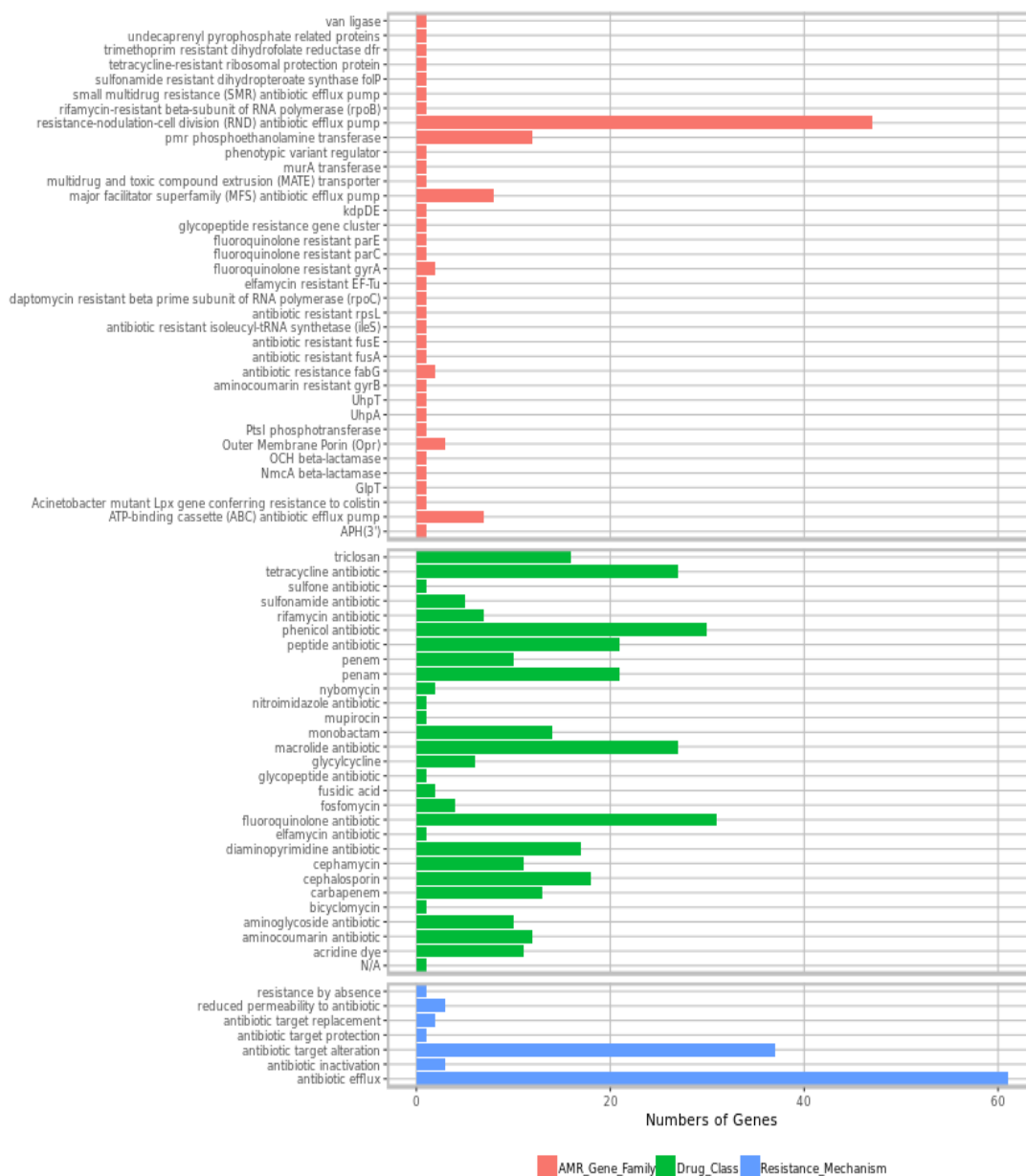
Functional annotation using Global catalogue of type strain (GC type)

Genes were predicted and are aligned with several databases to obtain their corresponding annotation with aligner such as DIAMOND and HMMER. The genome was annotated by the global catalogue of type strain (gcType) (<http://gctype.wdcm.org/>), and the genes predicted were annotated using several databases, including with CARD, MetaCyc NR, CAZy, PHI, VFDB, KEGG, Swiss-Prot, Pfam, KEGG and COG databases (Table 5. 5).

Table 5. 5: Genes annotated by GC type in various databases

Databases	Annotated genes	Ratio
CARD	81	0.01
MetaCyc	1226	0.22
NR	0	0
CAZy	118	0.02
PHI	343	0.06
VFDB	515	0.09
Swiss-Prot	2259	0.4
Pfam	4723	0.84
KEGG	5233	0.93
COG	4216	1

Genes were annotated with CARD protein ID and grouped by drug class, resistance mechanism, and AMR gene family. The number of genes annotated with CARD is as shown in Figure 5. 5. Genes were annotated with the COG protein ID first, and then the genes with the ID were mapped to 26 COG categories and further grouped into 4 clusters. The number of genes in each category is shown in Figure 5. 6. Genes were annotated with KEGG protein IDs first, then the genes with the ID were mapped to KO entry and further mapped to the KEGG pathway. The number of genes in each category is shown in Figure 5. 7.

Figure 5. 5: CARD analysis of *Pseudomonas* sp. strain VB3

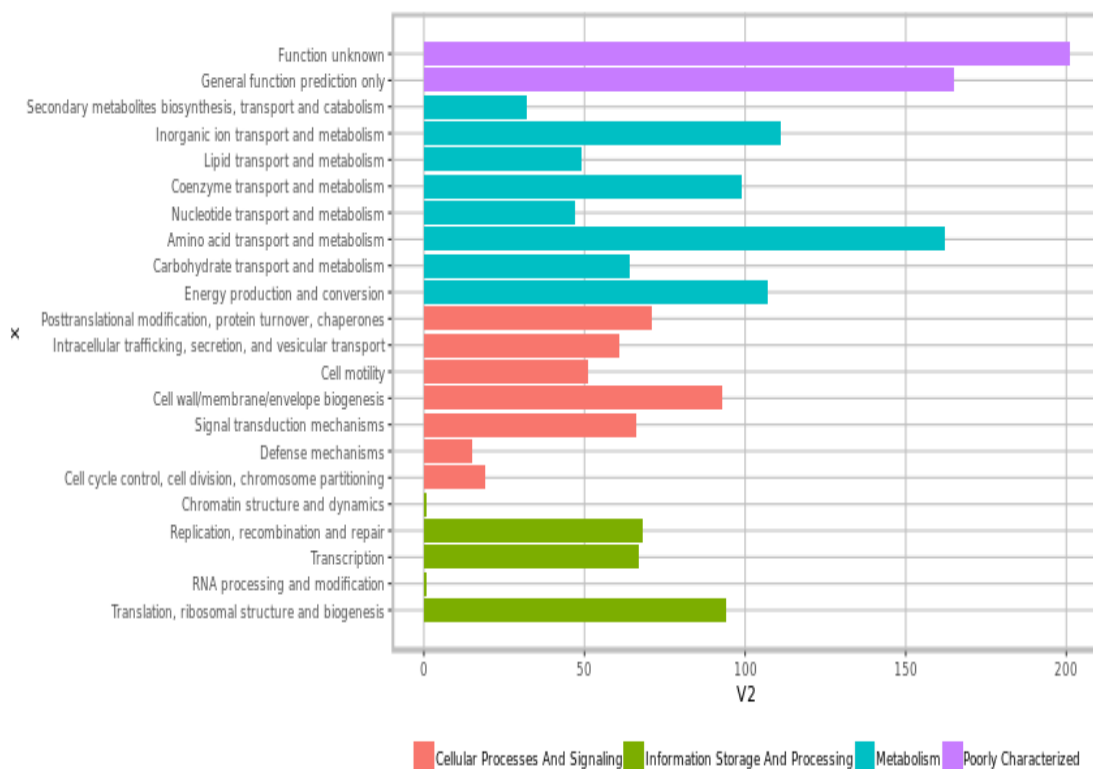


Figure 5. 6: COG functional analysis of *Pseudomonas* sp. strain VB3.

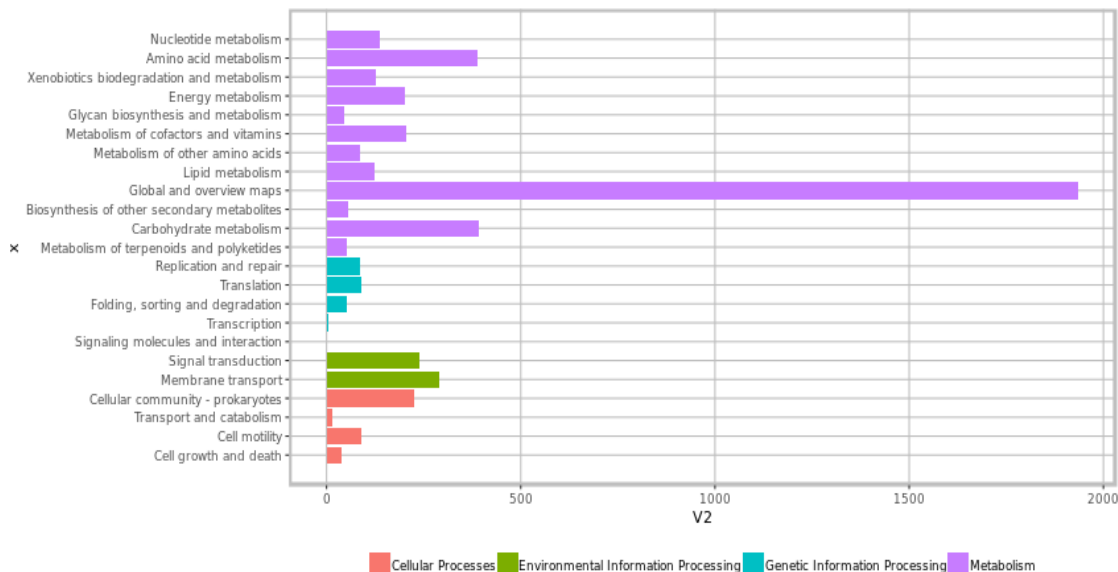


Figure 5. 7: KEGG analysis of *Pseudomonas* sp. strain VB3.

CRISPR-CAS region identification

CRISPR-associated Cas proteins are viewed as a prokaryotic immune system that confers resistance to foreign genetic elements such as phages. We predicted CRISPR in the genome of bacterial strain VB3. Analysis of CRISPR-Cas system using CRISPR-Cas finder revealed the presence of one CRISPR array containing one spacer and no associated Cas genes (Table 5. 6). CRISPR elements provide a bacterial defences mechanism against viral challenges, and research suggests CRISPR elements with fewer spacers and no CAS are more susceptible to prophage integration (Watson et al., 2021).

Table 5. 6: CRISPR-Cas Finder analysis of *Pseudomonas* sp. VB3 genome.

Contig No	CRISPR location	CRISPR length	No Repeats	Repeat length (bp)	Spacer count
228	959-1064	GTTCCGGGCAAGGACACCAAAGTGC	1	25	1

5.3.4 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. PHASTER assigns a completeness score to the Prophage regions based on the proportion of phage genes identified within the region.

The score criteria were:

Intact: A completeness score greater than 90 indicates the prophage region is intact.

Questionable: A completeness score ranging from 70-90 is questionable in terms of completeness.

Incomplete: A completeness score below than 70 indicates the prophage region is incomplete.

Five putative prophages-like elements were identified from the *Pseudomonas* sp. strain VB3 genome (Figure 5. 8). The genome sizes of *Pseudomonas* VB3 prophages ranged from approximately 7.4- 47.6 kb and the GC content varied between 52.49 and 60.53% (Table 5. 7). 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.

PHASTER New Search Genomes Help About My Search

Total: 5 prophage regions have been identified, of which 2 regions are intact, 3 regions are incomplete, and 0 regions are questionable.

Region	Region Length	Completeness	Score	# Total Proteins	Region Position	Most Common Phage	GC %	Details
NODE_9_length_110045_cov_78.539402								
1	20.5Kb	intact	130	29	63689-84222	PHAGE_Vibrio_VP882_NC_009016(14)	60.53%	Show
NODE_11_length_101800_cov_82.326671								
2	47.6Kb	intact	150	45	14568-62244	PHAGE_Iodobacter_phiPLPE_NC_011142(14)	56.88%	Show
3	10.3Kb	incomplete	20	19	53016-63359	PHAGE_Pseudo_PMG1_NC_016765(3)	57.38%	Show
NODE_14_length_97041_cov_86.025762								
4	24.2Kb	incomplete	30	12	19430-43688	PHAGE_Escher_500465_1_NC_049342(6)	58.75%	Show
NODE_132_length_11059_cov_64.387999								
5	7.4Kb	incomplete	30	10	3009-10452	PHAGE_Shigel_Sfil_NC_021857(2)	52.49%	Show

■ Intact (score > 90)
■ Questionable (score 70-90)
■ Incomplete (score < 70)

Region: The number assigned to the region.
Region Length: The length of the sequence of that region (in bp).
Completeness: A prediction of whether the region contains a intact or incomplete prophage based on the above criteria.
Score: The score of the region based on the above criteria.

Figure 5. 8 : PHASTER analysis of *Pseudomonas* sp. Strain VB3

There were two intact prophage regions and three prophage regions incomplete.

Prophage region 1: The prophage region 1 was intact with a score of 130 and the size of the prophage was 20.5kb with GC content 60.53%. The Prophage region 1 proteins coded for protease, tail and plate phage proteins (Figure 5. 9). The total number of ORF present in this region were 29, among which 21 proteins hits were for the phage proteins and 8 were hypothetical proteins. The region is intact with a score 130 and the size was 20.5kb. The major proteins were protease, tail and proteins.

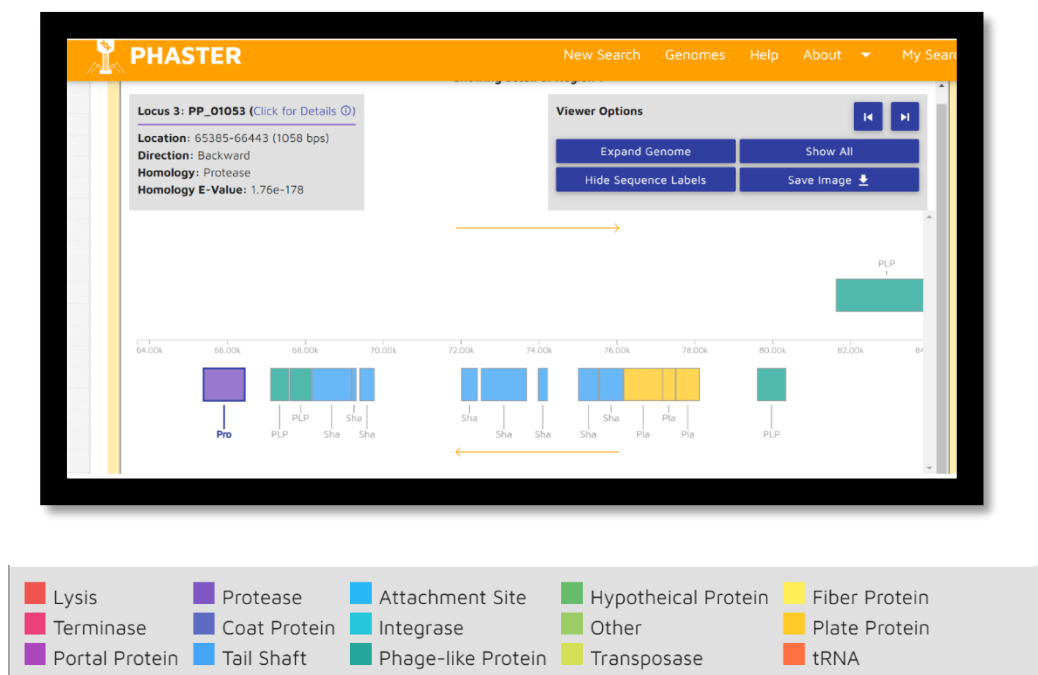


Figure 5. 9: Prophage region 1 detected by PHASTER web server tool.

Prophage region 2

The prophage region 2 was intact with a score of 150 and the size of the prophage was 47.6kb with GC content 56.88%. The prophage region 2 proteins coded for integrase, tail, terminase, plate, capsid, head, and portal proteins. (Figure 5. 10). The total number of ORF present in this region were 45, among which 37 proteins hits were for the phage proteins and 8 were hypothetical proteins. No putative tRNA was recognized in the prophage region 2. The ORFs in prophage region 2 are flanked by **attL** and **attR** sites. Most temperate phages enter the lytic cycle depending on the integrase gene, which functions in chromosomal integration and excision. The integrase gene was identified only in prophage region 2, suggesting that the prophage has the ability to enter a lysogenic replication cycle.

All these finding would suggest that prophage region 2 was excised from the host genome *Pseudomonas* VB3 after mitomycin C induction.

Further to see if the phage excised from the host genome is prophage region 2, phage DNA isolation was carried out after induction with mitomycin C (See Chapter 6).



Figure 5. 10: Prophage region 2 detected by PHASTER web server tool.

The region had proteins especially coding for integrase, terminase, tail, head, capsid, and portal proteins. This prophage had all the modules for perfect temperate phage. Also, the presence of Cro and CI proteins which decide the lysogeny to lytic shift was present.

Prophage regions 3, 4, 5 were not complete prophages. The sizes of these phages were very small. They could be just remnants of phage present in the host to increase the fitness of the bacterium to survive in harsh conditions.

Another interesting fact was prophage region 3 of size 10.3 kb is present on the same contig. BLAST analysis of the phage showed similarities with *Pseudomonas* sp. PHIAH1a (NC_074659) (Accessed on 23 May 2021). These phages could be like a Satellite-helper phage (Dziewit and Radlinska, 2016). Satellite-helper phages often carry genes that interfere with the replication or assembly of the helper phage, ensuring their own propagation. They use the structural proteins supplied by another ‘helper’ phage for assembly of their own virions, and thus, for their propagation and spread. The best studied examples of satellite-helper phages are those between the *Enterobacteria* satellite phage P4 (or the related retronphage ϕ R73) and the fully functional phage P2. Were, P4 is a satellite phage that depends for its propagation on the presence of the helper phage P2 (Lindqvist et

al., 1993). Further, mitomycin C induction experiments and isolation of the temperate phage might answer these findings.

Table 5. 7: Putative prophage DNA sequences detected using PHAge Search Tool Enhanced Release (PHASTER).

Prophage	Score	Region Length	No of proteins	Completeness	Closest relative	Contig No	%GC
Prophage region 1	130	20.5kb	29	Intact	PHAGE_Vibrio_VP88 2_NC_009016 (14)	NODE_9_length_110045_cov_78.5394 02	60.53%
Prophage region 2	150	47.6kb	45	Intact	PHAGE_Iodobacter_phPL PE_NC_0011142(14)	NODE_11_length_101800_cov_82.326 671	56.88%
Prophage region 3	20	10.3kb	19	Incomplete	PHAGE_Pseudo PMG 1_NC_016765(3)	NODE_11_length_101800_cov_82.326 671	57.38%
Prophage region 4	30	24.2kb	12	Incomplete	PHAGE_Escher_5004 65_1_NC_049342 (6)	NODE_14_length_97041_cov_86.0257 62	58.75 %
Prophage region 5	30	7.4kb	10	Incomplete	PHAGE_Shigella_sf11 NC_021857 (2)	NODE_132_length_11059_cov_64 .387999	52.49%

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 induction in *Pseudomonas* sp. VB3 (refer in Chapter 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 and its induction. In spot assay, lysis was observed on the host lawn as clearance on the lawn (refer in Chapter 4).

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 functional genes (Casjens, 2003). Then the bacterial host population could adapt to a specific environment (Zeng et al., 2016). Further characterization of the induced virus particle through TEM imaging and NGS sequencing will better understand which putative temperate prophage from strain VB3 is inducible by Mitomycin-C.

5.4 Conclusion:

In this chapter, we determined the draft genome sequence of *Pseudomonas* sp. VB3 from the Vashisht geothermal spring in the Himalayan geothermal province. 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 gene sequencing assigned the bacterial strain VB3 to the genus *Pseudomonas*. The predicted genes were annotated using several databases, including CARD, MetaCyc NR, CAZy, PHI, VFDB, KEGG, Swiss-Prot, Pfam, KEGG, and COG databases. Genome mining revealed fourteen gene clusters coding for the biosynthesis of bioactive secondary metabolites. The genome of *Pseudomonas* sp. VB3 harbors NRPS, aryl polyene, RiPP-like, redox-cofactor, betalactone, butyrolactone, hserlactone, NAGGN, NRPS-like producing secondary metabolite biosynthetic gene clusters. PHASTER analysis identified two intact and three incomplete putative prophage-related sequence regions in the draft genome of *Pseudomonas* sp. VB3.

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; Poddar and Das, 2018; Sharma et al., 2018). Although lysogeny is a favorable lifestyle for bacteriophages in hot springs (Breitbart et al., 2004), viruses in Indian hot springs are least explored. A recent study using viral metagenomics at Manikaran hot springs (Himachal Pradesh, India) (Sharma et al., 2018) depicted that lysogeny is hot springs dominant viral lifestyle strategy.

Pseudomonas phages might have a temperate lifestyle in hot springs again suggesting that lysogeny is the predominant lifestyle in these extreme environments. Further, analysis can be carried out to induce these prophages using various agents and to elucidate the co-evolution strategies of phage-host in extreme environments such as hot springs. Characterization of inducible phages will better understand putative prophages from the hot springs.

Culture collection deposit number:

The bacterial culture *Pseudomonas* sp. VB3 from the Vashisht hot spring was deposited with culture collection National Centre for Microbial Resources (NCMR), National Centre for Cell Science (NCCS), Pune with the accession no: MCC5368.

Nucleotide sequence accession number:

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.

CHAPTER 6

Characterization of bacteriophage for a specific bacterial strain.

The chapter describes the characterization of the inducible temperate bacteriophage of Pseudomonas sp. VB3 isolated from Vashisht hot spring, India.

Bacteriophages are estimated to have a more extensive repertoire of genetic capabilities and are responsible for horizontal gene transfer throughout microbial communities. Although there are few reports on metagenomic viral diversity studies, scanty reports on lysogenic hot spring bacteria have been described. One of the challenges has been to induce these lysogens and study these hot spring temperate viruses. The genomic sequence of the bacteriophage would give insight into the genome architecture and the level of similarity compared to currently available phages. Prophage induction and characterization of phage induced from the host organism would throw more light on the lysogeny genes involved. Transmission electron microscopy can study the bacteriophage structure, which would let us classify the virus as tailed or non-tailed.

The chapter describes the characterization of an inducible prophage from the potential lysogenic host *Pseudomonas* sp. strain VB3 isolated from the Vashisht hot springs, Himachal Pradesh, India. In the previous chapter 5, the WGS sequencing of the bacteria revealed the presence of 5 prophage regions in *Pseudomonas* sp. VB3. Two of the five prophage regions were intact, and 3 were incomplete genomes based on PHASTER analysis. To better understand prophage induction, the prophages were induced by various agents. The phage genome was obtained after several attempts after large-scale propagation of the bacteriophage lysate. The major challenge was to concentrate these phage particles and characterize them. Mitomycin C was used to induce the potential temperate phages in *Pseudomonas* sp. VB3. Mitomycin C is reported as a standard chemical treatment to induce temperate phages (Oliveira et al., 2017).

According to Olivera et al. (2017), mitomycin C is a common chemical therapy used to produce temperate phages. Taking into account the effectiveness and security of

mitomycin C in prophage induction, we selected this chemical agent for induction.

Members of the genus *Pseudomonas* are among the most significant model organisms in theoretical study and biotechnological applications due to their tremendous genetic flexibility and metabolic diversity. Phage-host interactions can also be studied using *Pseudomonas* sp. and associated phages. Additionally, it has been noted that prophages are frequently discovered in the genomes of *Pseudomonas* and have significant roles in influencing the phenotypic characteristics of their hosts. For instance, the prophage Pf4 is essential for *P. aeruginosa* to produce biofilms and increases its pathogenicity (Rice et al., 2009). *P. putida* KT2440 has a higher resilience to environmental stressors as a result of having all four prophage components removed, which increases the usefulness of this microbe for biotechnological applications (Martínez-García et al., 2015).

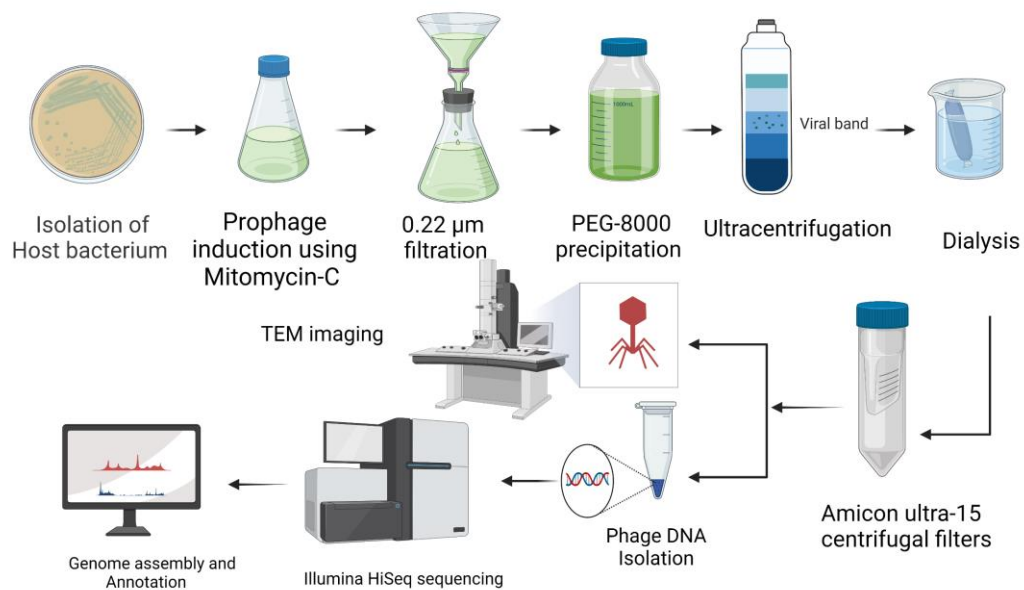
Workflow:

Figure 6. 1: Workflow of the characterization of bacteriophage from a specific bacteria *Pseudomonas* sp. strain VB3.

Created with BioRender.com (<https://www.biorender.com/>).

6.1 Materials:

Bacterial isolate *Pseudomonas* sp. VB3, Mitomycin-C (Sigma Aldrich), PEG 8000 (Hi Media), CsCl (Hi-media), ethanol, isopropanol, TE buffer (pH8), CTAB, 3M sodium acetate (pH 5.2), SM buffer (100 mM NaCl, 50 mM Tris-HCl, 8.1 mM MgSO₄, and 0.01% gelatin), Amicon Ultra-15 (Millipore, USA), phenol: chloroform: isoamyl alcohol (25:24:1), 3M sodium acetate (pH 5.2), Qubit HS dsDNA kit (Life technologies, USA).

6.2 Methods:

6.2.1 Bacterial culture host organism and maintenance:

The bacterial isolate *Pseudomonas* sp. strain VB3 was isolated from Vashisht hot spring, Himalayan Geothermal province, India. The workflow of the characterization of bacteriophage for a specific bacterium, *Pseudomonas* sp. strain VB3 is illustrated in Figure 6. 1. The potential lysogenic bacterial strain VB3 was selected and maintained at 30°C in Basal medium (BM) agar plates (g/L) containing KH₂PO₄:1.36, (NH₄)₂SO₄:1.0, MgSO₄.7H₂O: 0.2, FeSO₄: 0.001, NaCl: 2.0, yeast extract, 1.0 and supplemented with 0.2% sodium succinate.

6.2.2 Prophage induction and purification of prophages

Pseudomonas sp. strain VB3 was grown in BM at 30°C for 24h. The prophage induction from bacterial strain VB3 with mitomycin-C was done as described by Chen et al. (2006) with slight modifications. Briefly, 10 ml of starter culture was transferred to 200 ml fresh basal medium. At OD_{600 nm} of 0.3, the culture was divided into two sterile flasks (100 ml each). One of the flask was treated with mitomycin-C (final concentration 1 µg/ml) whereas the other served as control. Both control and mitomycin-C treated bacterial cells were incubated and cell density at OD_{600 nm} was monitored at every two hours for 24 h. Cell debris was removed by centrifugation at 5000 × g for 20 min, and lysates were then passed through a 0.22 µm PVDF filter. After induction, the bacteria DNA was degraded with DNaseI before concentration by ultra-centrifugation (Sambrook, 1989). The virus sample was fixed with formaldehyde (2%, v/v final), and viruses were counted using a wet-mount

method with SYBR Gold and epifluorescence microscopy (Cunningham et al., 2015).

6.2.3 Scanning electron Microscopy (SEM) image of strain VB3

The Mitomycin-C treated, and untreated bacterial suspension of VB3 was smeared on a 10 mm x 10 mm glass slide and air dried, then fixation with 2% (v/v) glutaraldehyde overnight. The next day it was washed with 0.1M sodium phosphate buffer and dehydrated using increasing acetone concentrations (30%, 50%, 70%, and 90%) for 10 min each and finally with 100% Acetone for 30 min. Slides were air dried and sputter coated with Gold and subjected to SEM.

6.2.4 Determination of host range specificity for bacteriophage after induction

The host range of bacteriophage VB3 was assayed by spot assay. The host range specificity of phage VB3 was determined against six *Pseudomonas* species. *P. syringae* (MTCC No 11950), *P. fluorescens* (MTCC No: 103), *P. putida* (MTCC No:102), *P. stutzeri* (MTCC No: 101), *P. fragi* (MTCC No: 510), and *P. alcaligenes* (MTCC No: 493) were obtained from MTCC, Chandigarh, India. The bacteria were grown on Nutrient Agar media at 25°C for 24 h in the incubator.

Mid-log phase bacterial cells were mixed with molten top agar (0.6%) and poured on 2% bottom agar for the host range specificity analysis. After solidification, 10 µL of VB3 phage lysate was spotted on each plate against the lawn of 6 different strains of *Pseudomonas* sp. and incubated at 25°C for 24 h. A zone of clearance around the spot was examined to determine the host range specificity (Kutter, 2009).

6.2.5 Large Scale prophage induction

In order to obtain large quantity cell lysates, mitomycin C-mediated prophage inductions were performed as described before, but on a larger scale (2000 ml). Briefly, this involved the addition of mitomycin C (final concentration 1µg/ml) to early exponential phase cultures grown in a 2000 ml volume of BM broth, followed by overnight incubation. Cell debris was removed by centrifugation at $7560 \times g$ for 20 min, and the lysate was passed through a 0.22 µm PVDF filter (Millipore).

6.2.6 Concentration of phage particle using PEG 8000

Induced phage particles were precipitated with 10% PEG 8000 containing 1M NaCl after overnight incubation at 4°C (Sambrook, 1989). Two liters of mitomycin- C induced phage lysate was centrifuged at 12,000 x g for 15 min. The supernatant was filtered through a 0.22µm-pore size PVDF filter (Millipore) to remove bacterial host cells and cellular debris. Phage particles from the filtrate were treated with PEG 8000 (final concentration: 100 g/l) and incubated overnight at 4°C followed by centrifugation at 12,000 x g for 20 min to obtain phage pellet. The phage pellet was re-suspended in 10 ml SM buffer (100 mM NaCl, 50 mM Tris, and 8.1 mM MgSO₄, pH 7.5) and incubated at 4°C.

6.2.7 CsCl density gradient ultracentrifugation.

The phage suspension was suspended in Cesium Chloride (1.45, 1.5, and 1.6 g/ml densities) and centrifuged at 45,000 rpm for 24 h in a Beckman ultracentrifuge (MLS-50 rotor) (Sambrook, 1989). After ultracentrifugation, the visible bluish phage band was collected and dialysed against 1L of SM buffer for 1h, 3h, and finally dialysed overnight. The dialyzed fraction was stored at 4°C for further characterization.

6.2.8 Transmission Electron Microscopy

Before TEM analysis, the phage was concentrated using Cesium chloride density gradient centrifugation (1.45, 1.5, and 1.6 g/ml densities) and centrifuged at 45,000 rpm for 24 h in a Beckman ultracentrifuge (MLS-50 rotor) (Sambrook, 1989). After centrifugation, the phages formed bluish white band at the interface between 1.45 and 1.5 g/ml CsCl. To visualise the phage morphology, small fraction of the concentrated phage suspension (CsCl purified and dialysed) was adsorbed onto a formvar carbon-coated copper grid. The grid was carefully placed using forceps, over the drops of phage, with carbon side facing down. After approximately 1min, the grid was later fixed and stained negatively with 2% (w/v) phosphotungstic acid. The excessive stain was removed, and the phage particles were examined under the transmission electron microscope (Jeol/JEM 2100).

6.2.9 Phage DNA isolation

The phage VB3 DNA was isolated after induction of *Pseudomonas* strain VB3 with 1 µg/ml Mitomycin-C (Chen et al., 2006). The 2L phage lysate was filtered through a 0.22 µm membrane filter (Millipore, USA) after induction with Mitomycin C. Phage particles were precipitated with 10% PEG 8000 containing 1M NaCl after overnight incubation at 4°C (Sambrook, 1989). The mixture was centrifuged at 10000 x g for 20 min at 4°C, and the pellet was suspended in SM buffer and kept at 4°C. The virus was pelleted by centrifugation (12 000 x g for 10 min at 4°C), and the supernatant containing viral particles was transferred to a fresh tube. The viral suspension was concentrated using CsCl ultracentrifugation (described above) and Amicon Ultra-15 centrifugal filter units with a molecular cut-off of 100 KDa (Millipore, USA). The concentrated phage preparation was treated with DNase I at a final concentration of 2.5 U/ml to remove any host DNA, and RNase A with a concentration of 1 µg/ml was added to the purified phage suspension. The temperate phage VB3 DNA was extracted using Formamide/CTAB method and followed by phenol: chloroform isoamyl alcohol (25:24:1) extraction method as described previously (Thurber, 2009). Briefly, the nuclease-treated phage preparation was treated with proteinase K (20mg/ml), 10% (w/v) SDS, and 0.5 M EDTA (pH 8.0) and incubated in a water bath at 55°C for 2 h. This was followed by phenol: chloroform: isoamyl alcohol (25:24:1) extraction at 12 000 x g at 4°C for 5 min.

Again, the aqueous phase was purified by chloroform: isoamyl alcohol (24:1) extraction at 12 000 x g at 4°C for 5 min. The supernatant was mixed with 1/10 volume of 3 M sodium acetate (pH 5.2) and 0.7 volume of isopropanol and incubated overnight at 4°C for precipitation. The precipitated phage DNA was centrifuged at 12000 x g for 20 min at 4°C. The pellet was washed with 70% ethanol and air-dried. The phage DNA was resuspended in 50 µl 10 mM Tris-HCl (pH 8.0) and stored at -20°C. The phage DNA was quantified on Qubit 2.0 fluorometer using a Qubit HS dsDNA kit (Life Technologies, USA) and evaluated on a 1% agarose gel. The temperate phage VB3 DNA sample was sequenced on the HiSeq Illumina Platform.

6.2.10 Genome sequencing of temperate phage of *Pseudomonas* sp. VB3

After several attempts, to purify and concentrate the phage DNA, the induced prophage was isolated. A WGS library was prepared with an Illumina-compatible NEXTflex rapid DNA sequencing kit (BIOO Scientific). Qubit-quantified DNA (100 ng) was sheared using a Covaris S220 sonicator to generate specific fragments in a 300–400 bp size range. The fragment size distribution was verified on an Agilent 2200 Tape Station and purified using High-Prep magnetic beads (MagBio, Genomics). Purified fragments were end-repaired, adenylated, and ligated to Illumina multiplex barcode adapters as per the NEXTflex rapid DNA sequencing kit protocol. Illumina universal adapters used in the study are:

5'-AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC T-3' and Index Adapter 5'-GAT CGG AAG AGC ACA CGT CTG AAC TCC AGT CAC [INDEX] ATC TCG TAT GCC GTC TTC TGC TTG-3'.

Adapter-ligated DNA was purified using High-Prep beads. Resultant fragments were amplified for 5 cycles of PCR using Illumina-compatible primers.

The Phage DNA was sequenced on the Illumina Hi-Seq platform. The raw reads were quality controlled, trimmed, and assembled. Briefly, the Illumina pair-ended reads were quality-checked using FAST QC (version 0.11.8) (Andrews, 2010) before quality trimming using Trimmomatic (version.0.39) (Bolger et al., 2014). The high-quality trimmed reads were assembled with SPAdes (version 3.12.0, -only assembler) (Bankevich et al., 2012).

The potential phage contigs were filtered with VirSorter (default settings, Virome decontamination mode) from iVirus (<https://de.iplantcollaborative.org/de/>) (Bolduc, 2021; Bolduc, et al., 2017; Roux et al., 2015) and followed by identification of phage sequences with PHASTER (<https://phaster.ca>) (Arndt et al., 2016; Zhou et al., 2011). The annotation was first performed using RAST's software (<http://rast.nmpdr.org/rast.cgi>, with default parameters) (Aziz et al., 2008; McNair et al., 2018)

Using the Viral Proteomic Tree server (<http://www.genome.jp/viptree>) (ViP-Tree v.3.4) (<https://www.genome.jp/viptree/>) (Nishimura et al., 2017), the entire genome "proteomic tree" was created. This tree is a dendrogram that was generated using tblastx to compare entire genome sequences (Altschul et al., 1990). The phage vB_Psp_VB3 sequence

was also analyzed on VIRFAM (<http://biodev.cea.fr/virfam/>), VIRDIC (<https://rhea.icbm.uni-oldenburg.de/viridic/>) (Moraru et al., 2020) and VirClust analysis (<http://rhea.icbm.uni-oldenburg.de/virclust/>) (Moraru, 2023).

6.3 Results and Discussion

6.3.1 SEM imaging of *Pseudomonas* sp. strain VB3

After Mitomycin- C treatment (1 $\mu\text{g/ml}$), the bacterial cells showed elongation of cells (Figure 6.2). The cells exposed to the DNA-damaging agent mitomycin-C exhibited an elongated morphology. The elongation morphology of the cell is due to the SOS response mechanism of the bacteria. A similar kind of elongation of cells was noticed in a study of *Corynebacterium glutamicum* bacteria, suggesting that cells elongate following DNA damage due to the Rec A and Div S-dependent suppression of cell division suppressor (Ogino et al., 2008).

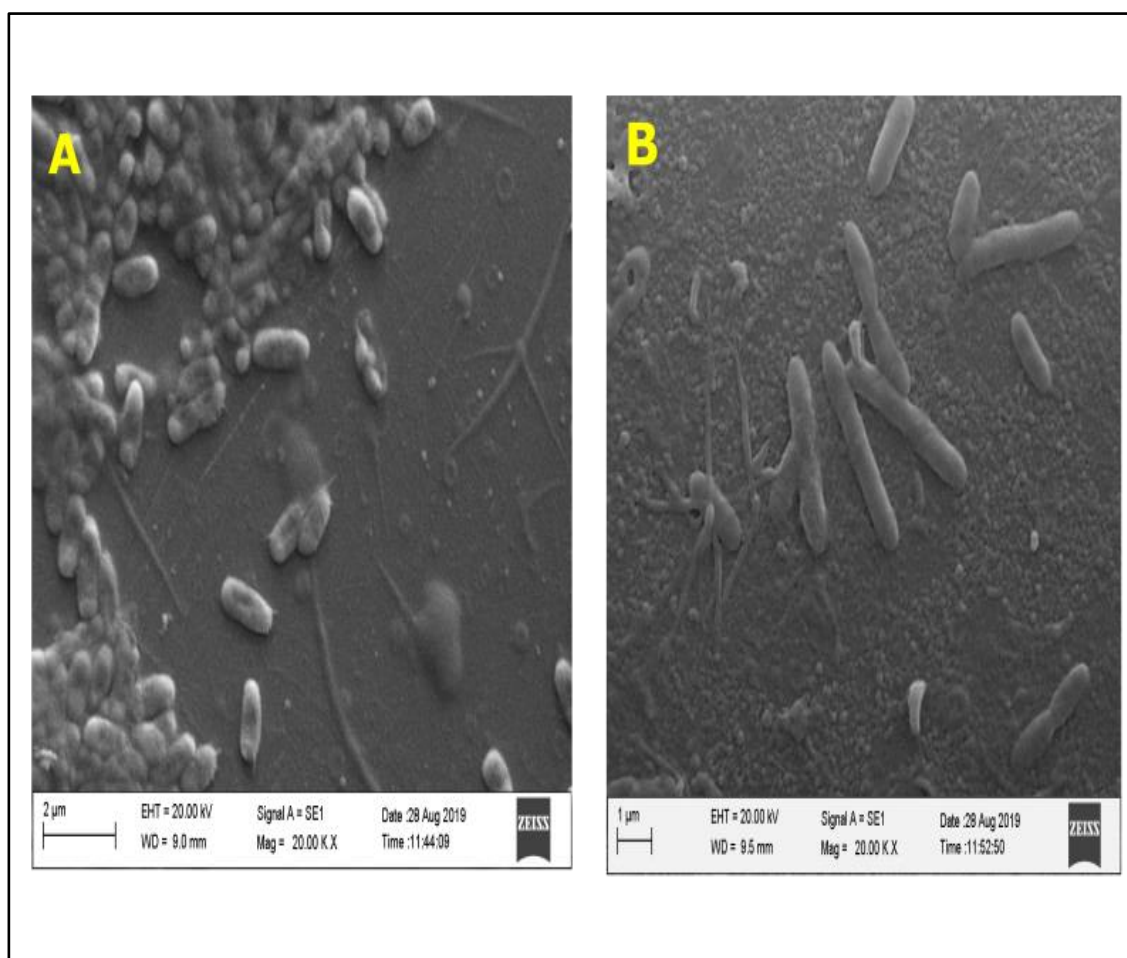


Figure 6.2: SEM image of *Pseudomonas* sp. strain VB3; A) without Mitomycin C, B) after Mitomycin C treatment.

6.3.2 Host range specificity for bacteriophage after induction

The host range specificity for the bacteriophage VB3 was checked against 6 different *Pseudomonas* species. The spot assay method was used to check the infectivity. None of the *Pseudomonas* cultures showed a clearance zone upon infection with phage particles isolated (Table 6. 1). Thus, the bacteriophage obtained during the induction assay had a very limited host range.

Most bacteriophages can only infect a limited number of strains within a single bacterial species. The extremely precise interaction between host cell surface receptors and phage attachment structures determines the potential to infect the host (Weinbauer, 2004).

The isolated phage could not infect any of the *Pseudomonas* species examined in the present investigation, indicating that it probably has a limited host range. However, evaluation with other *Pseudomonas* species would unravel its host range. A certain phage's host range may be affected by both host resistance mechanisms and phage invasion processes. The phage benefits from specialization for a single host over selection for a wide host range (Koskella and Meaden, 2013).

Table 6. 1: Host range analysis for temperate *Pseudomonas* phage.

S.No.	<i>Pseudomonas</i> strains	Infectivity
1.	<i>Pseudomonas stutzeri</i> ^T (MTCC No: 101)	-
2.	<i>Pseudomonas fragi</i> ^T (MTCC No: 510)	-
3.	<i>Pseudomonas alcaligenes</i> ^T (MTCC No: 493)	-
4.	<i>Pseudomonas fluorescens</i> ^T (MTCC No: 103)	-
5.	<i>Pseudomonas syringae</i> ^T (MTCC No 11950)	-
6.	<i>Pseudomonas putida</i> ^T (MTCC No:102)	-
“-“ No clearance zone		

6.3.3 Purification of concentrated Phage VB3 particles using CsCl density gradient ultracentrifugation

A bluish band was visible after CsCl ultracentrifugation for 24 h (Figure 6. 3). The band was collected and dialysed against the SM buffer. The dialyzed fraction of phage VB3 was stored at 4°C until further use for TEM imaging studies and phage DNA isolation.

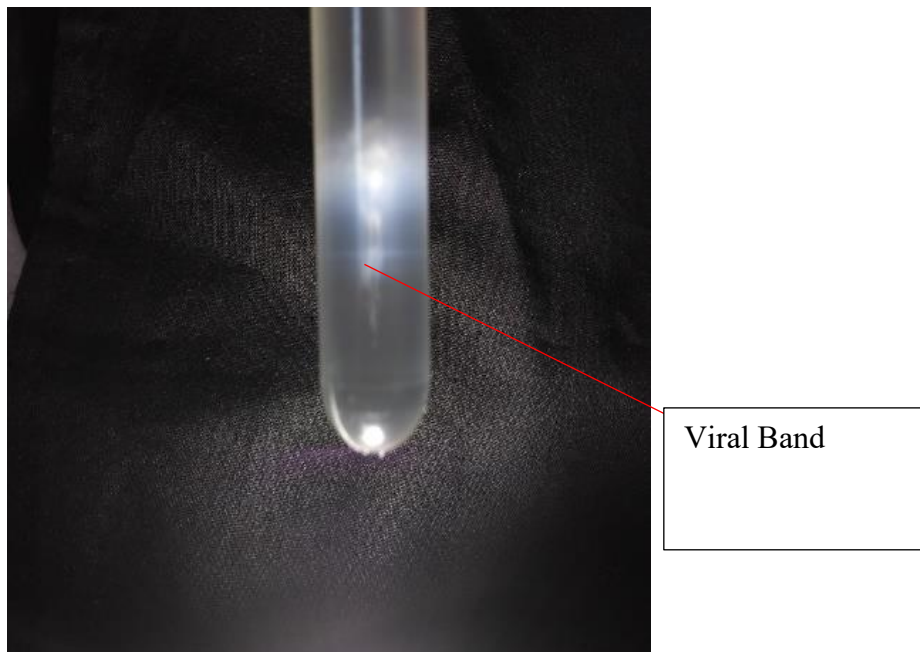


Figure 6. 3: CsCl ultracentrifugation concentration of VB3 phage particles after treatment with mitomycin-C.

6.3.4 Transmission electron Microscopy

The electron microscope image of negatively stained VB3 phage particles revealed a head with a tail morphology structure. The phage head was 30 nm in diameter and 36 nm long. (Figure 6. 4). Bacteriophage characterization by electron microscopy is one of the most widely used characterization techniques in virology (Ackermann, 2009). Transmission electron microscopy showed that phage vB_Psp_VB3 had an icosahedral head and a tail morphology. Based on the latest International Committee on Taxonomy of Viruses (ICTV) classification criteria, phage vB_Psp_VB3 belongs to the class *Caudoviricetes* (Turner et al., 2023).

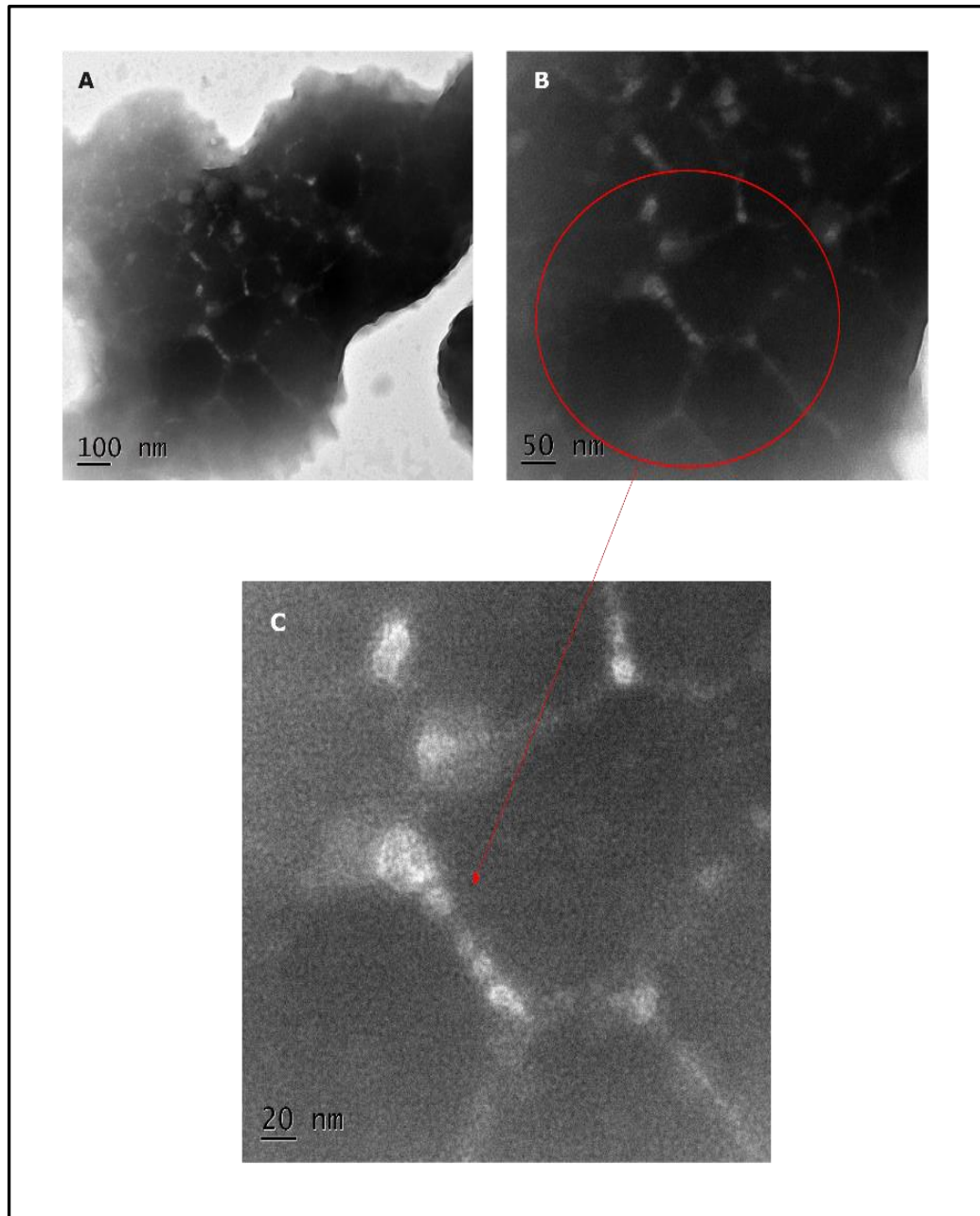


Figure 6. 4: TEM Images of induced temperate bacteriophage vB_Psp_VB3.

100nm scale, B) 50 nm scale and C) 20 nm scale.

6.3.5 Genome of phage vB_Psp_VB3.

The extracted Phage DNA (Figure 6. 5). was sequenced on the Illumina Hi-Seq platform. The genome sequence of vB_Psp_VB3 is composed of doubled-stranded DNA of approximately **48.6 kb** in length with GC content of **56.83%** containing putative **43 ORF** regions. This sequence was similar to the intact prophage region in the host *Pseudomonas* sp. VB3 (See Chapter 5).

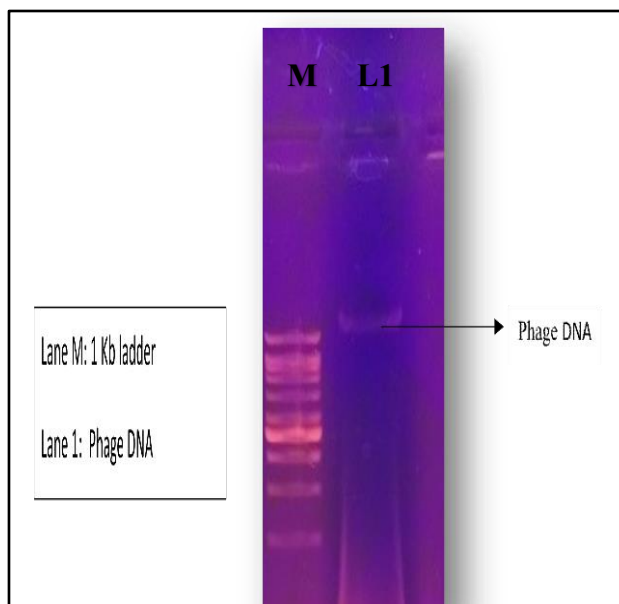


Figure 6. 5: Phage VB3 DNA isolated from *Pseudomonas* sp. VB3

Phage vB_Psp_VB3 had the highest sequence similarity to *Pseudomonas* phage phiPSA1 and *Pseudomonas* phage phiAH14a and a sequence identity of 89% and 92% respectively, when searched against NCBI-Virus database (<https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/>) (Table 6. 2).

Table 6. 2: Similarity with NCBI-Virus Ref seq complete genome.

Accession No	Phage	Genome Length	Host	Identity
NC_074659	<i>Pseudomonas</i> phage phiAH14a	55060	<i>Pseudomonas</i> sp. ANT_H14	92%
NC_024365	<i>Pseudomonas</i> phage phiPSA1	51090	<i>Pseudomonas syringae</i>	89%
NC_070881	<i>Pseudomonas</i> phage UFJF_PfDIW6,	42322	<i>Pseudomonas fluorescens</i> UFV 041	87%
NC_074661	<i>Pseudomonas</i> phage psageK9	51465	<i>Pseudomonas syringae</i>	89%
NC_074660	<i>Pseudomonas</i> phage MR15	51347	<i>Pseudomonas</i>	85%

The G+ C content of the induced phage vB_Psp_VB3 is 56.83%, which is less than the host *Pseudomonas* sp. host GC content. VB3 60.56%. Compared to the host organism's G+C content, the inducible phage's lysogeny lifestyle exhibits a lower G+C content, which is supported by research that correlates similarly (Almpanis et al., 2018).

The inducible phage vB_Psp_VB3 sequence was similar and matched with the intact prophage sequence in the host genome of strain VB3 of 48.6 kb. As phage vB_Psp_VB3 was induced from *Pseudomonas* strain VB3, the functions of viral genes were determined according to the genome annotation of *Pseudomonas* strain VB3. The genome was viewed in PHASTEST linear genome viewer (Wishart et al., 2023) which is the successor of PHAST and PHASTER (Figure 6. 6).

There were 50 viral genes CDS gene predicted in the vB_Psp_VB3 genome, but only 45 genes had functional annotations (Table 6. 3), which were further categorized into four groups: structure and assembly, replication and regulation, lysogeny, and lysis. No tRNA genes were found on the phage vB_Psp_VB3.

The structure and assembly module includes genes encoding the P27 family phage terminase small subunit, terminase large subunit, phage portal protein, phage major capsid protein, phage gp6-like head-tail connector protein, tape measure protein, and HNH endonuclease (Table 6. 3).

An integrase gene that encodes an integrase family protein was discovered next to the att-P site in the phage genome's lysogeny module. A characteristic of temperate phages

is the presence of an integrase gene. Prophage integration is done by this enzyme's catalyzes the recombination between the att-P site in the phage genome and the att-B site in the bacterial genome (Fogg et al., 2014).

In the lysis module, two genes are predicted to be involved in bacterial lysis protein and phage-encoded holin. Most phage-encoded holins' primary function is to create cytoplasmic membrane lesions that allow endolysins to be released (Song et al., 2016). There were genes encoding for regulatory proteins like the helix-turn transcriptional regulatory protein.

With the presence of Cro and CI proteins regulatory systems, the decision of the phage to enter either into a lytic or lysogeny lifestyle is determined by these proteins were also present (Pedersen et al., 2010).

Table 6. 3: Genome annotation bacteriophage vB_Psp_VB3

	Strand		Type	size	Top Homology	E-VALUE:
ORF1	+1	1-183	Regulatory protein	183 bps	helix-turn-helix transcriptional regulator; phage; -; PHAGE_Faecal_FP_Taranis_NC_047914;	3.7e-11
ORF 2	+1	1204..2400	Integrase	1197 bps	phage integrase family site-specific recombinase; phage; PHAGE_Erwini_phiEt88_NC_015295;	4.23e-92
ORF3	-1	2379..2891	Hypothetical	513 bps	PHAGE_Pseudo_PPpW_3_NC_023006	2.06e-27
ORF4	-1	2940..3500	Hypothetical	561 bps	PHAGE_Pseudo_PMG1_NC_016765	2.24e-38
ORF5	-1	3739..4302	Phage-like protein	564 bps	glycoside hydrolase family 19 protein;	6.36e-62

					PHAGE_Pseudo_phiPSA1_N C_024365	
ORF 6	-1	4364..6205	Phage- like_protein	1842 bps	putative acyltransferase; PHAGE_Enterо_Tyrion_NC_ 031077	3.86e-154
ORF 7	-1	6320..8440	Fiber_protein	2121 bps	putative tail fiber protein; PHAGE_Enterо_vB_EcoS_A CG_M12_NC_019404	2.44e-11
ORF8	-1	8517..10127	Terminase	1611 bps	Putative TerL large-subunit; terminasePHAGE_Pectob_ZF 40_NC_019522	2.83e-178
ORF9	-1	10124..10567	Terminase	444 bps	phage terminase, small subunit; phage; -; PHAGE_Erwini_phiEt88_NC _015295	1.29e-39
ORF10	-1	10784..11635	Fiber_protein	852 bps	putative tail-collar fiber protein; phage; -; PHAGE_Bordet_vB_BbrM_P HB04_NC_047861;	2.78e-18

ORF11	-1	11635..12294	Hypothetical protein	660 bps	hypothetical protein; phage; -; PHAGE_Iodoba_phiPLPE_N C_011142;	3.02e-29
ORF12	-1	12291..13550	Plate protein	1260 bps	putative baseplate protein; phage; -; PHAGE_Pectob_ZF40_NC_0 19522;	1.91e-77
ORF13	-1	13550..13915	Hypothetical protein	366 bps	hypothetical protein; phage; -; PHAGE_Iodoba_phiPLPE_N C_011142;	5.54e-25
ORF14	-1	13912..14568	Plate protein	657 bps	putative baseplate protein; phage; -; PHAGE_Vibrio_vB_VchM_1 38_NC_019518;	1.72e-35
ORF15	-1	14565..15515	Hypothetical protein	951 bps	hypothetical protein; phage; -; PHAGE_Vibrio_vB_VchM_1 38_NC_019518	4.07e-32
ORF16	-1	15508..15810	Hypothetical protein	303 bps	hypothetical protein; phage; -;	9.43e-7

					PHAGE_Pectob_ZF40_NC_019522;	
ORF17	-1	15807..16427	Hypothetical protein	621 bps	hypothetical protein; phage; -; PHAGE_Edward_Pei21_NC_021342	2.36e-37
ORF18	-1	16424..17965	Fiber protein	1542 bps	putative tail-fiber/lysozyme protein; phage;PHAGE_Vibrio_CP_T1_NC_019457;	2.41e-24
ORF19	-1	18192..18638	Hypothetical protein	447 bps	hypothetical protein; phage; -; PHAGE_Vibrio_CP_T1_NC_019457;	1.58e-7
ORF20	-1	18635..19039	Hypothetical protein	405 bps	hypothetical protein; phage; -; PHAGE_Klebsi_JD001_NC_020204;	3.2e-27
ORF21	-1	19053..20183	Hypothetical protein	1131 bps	hypothetical protein; phage; -; PHAGE_Edward_Pei21_NC_021342;	3.29e-103

ORF22	-1	20167..20724	Hypothetical protein	558 bps	hypothetical protein; phage; -; PHAGE_Iodoba_phiPLPE_N C_011142;	3.06e-22
ORF23	-1	20717..21136	Hypothetical protein	420 bps	hypothetical protein; phage; -; PHAGE_Iodoba_phiPLPE_N C_011142;	7.9e-16
ORF24	-1	21136..21633	Hypothetical protein	498 bps	hypothetical protein; phage; -; PHAGE_Pectob_ZF40_NC_0 19522	1.99e-28
ORF25	-1	21630..22073	Hypothetical protein	444 bps	hypothetical protein; phage; -; PHAGE_Pectob_ZF40_NC_0 19522	4.54e-28
ORF26	-1	22137..22415	Phage-like protein	279 bps	gp7; phage; -; PHAGE_Burkho_KS14_NC_ 015273	1.96e-11
ORF27	-1	22698..23675	Head protein	978 bps	putative major capsid protein; phage; -; PHAGE_Pectob_ZF40_NC_0 19522	4.51e-42

ORF28	-1	23688..24197	Hypothetical protein	510 bps	hypothetical protein; phage; -; PHAGE_Pectob_ZF40_NC_019522	1.91e-9
ORF29	-1	24197..25291	Head protein	1095 bps	putative head protein; phage; -; PHAGE_Edward_MSW_3_NC_020082	3.32e-76
ORF30	-1	25305..26195	Head protein	891 bps	putative head protein; phage; -; PHAGE_Pectob_ZF40_NC_019522	7.35e-62
ORF31	-1	26161..27573	Portal protein	1413 bps	gp19 putative portal protein; phage; -; PHAGE_Iodobacter_phiPLPE_NC_011142	2.53e-128
ORF32	-1	27802..28068	Hypothetical protein	267 bps	hypothetical protein; phage; -; PHAGE_Pseudo_JBD44_NC_030929	7.81e-21
ORF33	-1	28068..28385	Holin	318 bps	putative holin protein; phage; -	1.11e-8

					; PHAGE_Burkho_BcepC6B_ NC_005887	
ORF34	-1	28871..29536	Hypothetical protein	666 bps	hypothetical protein; phage; -; PHAGE_Pseudo_PS_1_NC_0 29066	1.58e-48
ORF35	-1	29533..31086	Hypothetical protein	1554 bps	hypothetical protein; phage; -; PHAGE_Pseudo_pf16_NC_0 41881	2.21e-81
ORF36	-1	31083..31361	Hypothetical protein	279 bps	hypothetical protein; phage; -; PHAGE_Salmon_9NA_NC_0 25443	4.46e-21
ORF37	-1	31549..32148	Phage-like protein	600 bps	NinG protein; phage; -; PHAGE_Pseudo_PMG1_NC_ 016765	4.91e-90
ORF38	-1	32145..32558	Phage-like protein	414 bps	gp50; phage; -; PHAGE_Burkho_Bcep176_N C_007497	1.77e-35
ORF39	-1	32555..33025	Hypothetical	471 bps	hypothetical protein; phage; -;	0.0000134

			protein		PHAGE_Pseudo_Pq0_NC_02 9100	
ORF40	-1	33459..34250	Hypothetical protein	792 bps	hypothetical protein; phage; -; PHAGE_Pseudo_F116_NC_0 06552	4.23e-56
ORF41	-1	34237..35196	Phage-like protein	960 bps	HTH domain-containing protein; phage; -; PHAGE_Serrat_Parlo_NC_04 8758	3.05e-29
ORF42	-1	35754..35966	Repressor	213 bps	putative repressor Cro; phage; -; PHAGE_Pseudo_PAJU2_NC _011373	1.86e-7
ORF43	1	36088..36756	Repressor	669 bps	putative repressor cI; phage; -; PHAGE_Pseudo_PAJU2_NC _011373	2.33e-89

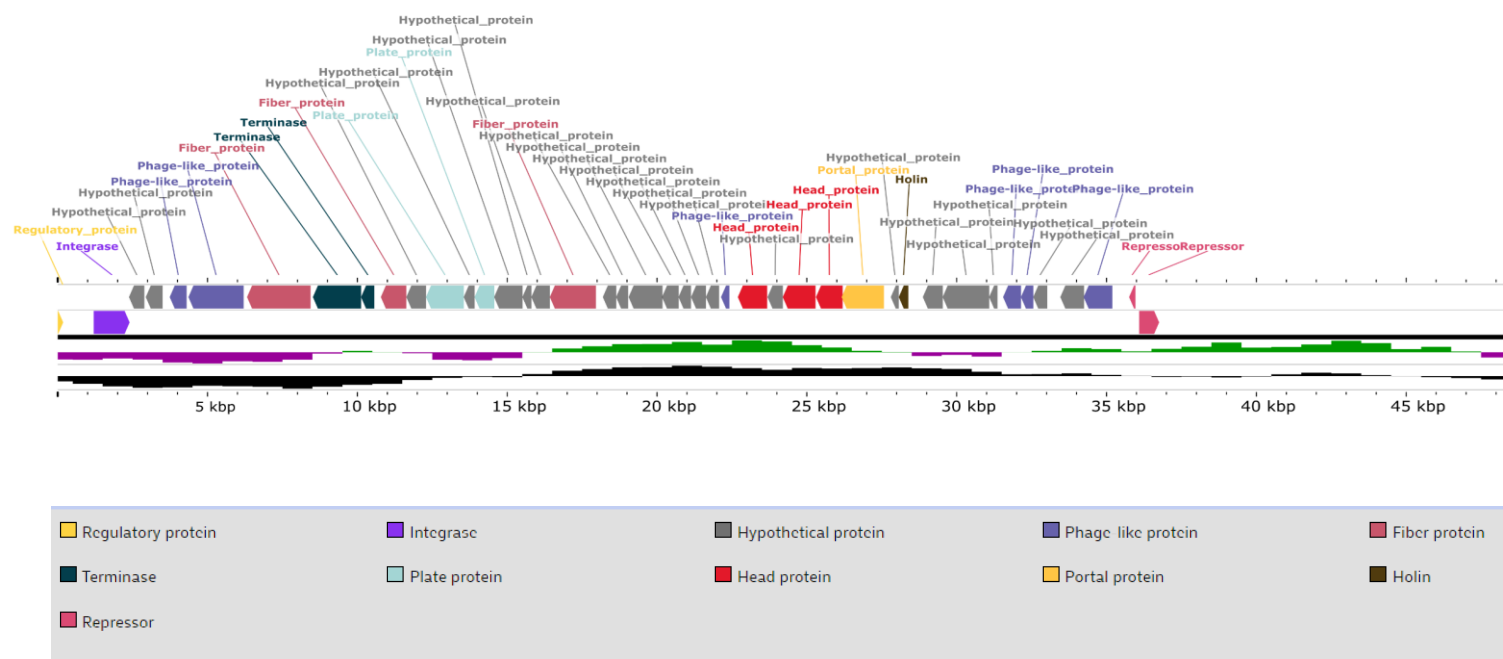


Figure 6. 6: Genome of phage vB_Psp_VB3 in PHASTEST linear genome viewer.

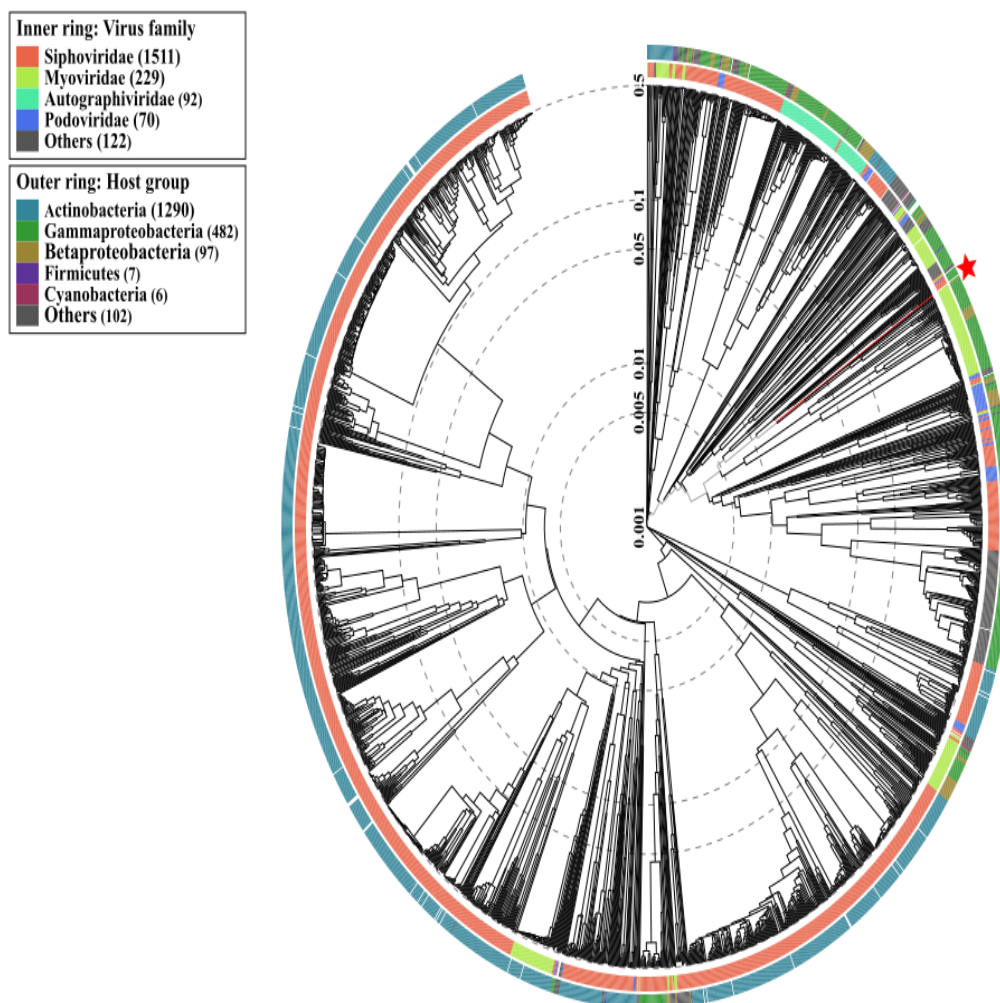


Figure 6. 7: Viral Proteomic tree VipTree v1.9.1 of phage vB_Psp_VB3 represented in circular view.

The coloured rings represent the host groups (outer ring) and virus families (inner ring). The VipTree analysis of phage vB_Psp_VB3, which is marked with a red star.

In order to study the phylogenetic relationship between vB_Psp_VB3 and other isolated phages, a circular proteome tree was constructed using VipTree based on the whole amino acid sequence of vB_Psp_VB3 and other reference phages (Figure 6. 7)

ViTree analysis of the phage vB_Psp_VB3 clustered with the *Myoviridae* family of viruses (Nishimura et al., 2017) , supporting TEM image results reported in our study.

Based on the VIRFAM analysis of the phage vB_Psp_VB3 sequence showed head, neck, and tail proteins, and the phage vB_Psp_VB3 belonged to *Myoviridae* cluster 7 confirming the observation made by TEM imaging of the phage (Figure 6. 8). The vB_Psp_VB3 was clustered with Cluster 7 phages sharing the most similar head-neck tail module. The phages in the cluster 7 strictly is composed of Myophages with small genome size. They infect all the host belonging to *Proteobacteria*. The vB_Psp_VB3 was clustered with Cluster 7 phages Felix 01, B054, Aaphi23, OP2, Bcep781, Bcep 43, Bcep1, BcepNY3, BcepB1A, BcepF1, F8 phages sharing the most similar head-neck tail module in ACLAME which exhibit a type 1 neck and belong to Cluster 7 (Lopes et al., 2014).

Based on the morphological characteristics of phage tail and VIRFAM analysis the phage was evaluated to belong to the class *Caudoviricetes* (Tailed Phages). The genomic analysis also revealed the presence of head and tail protein and major capsid proteins.

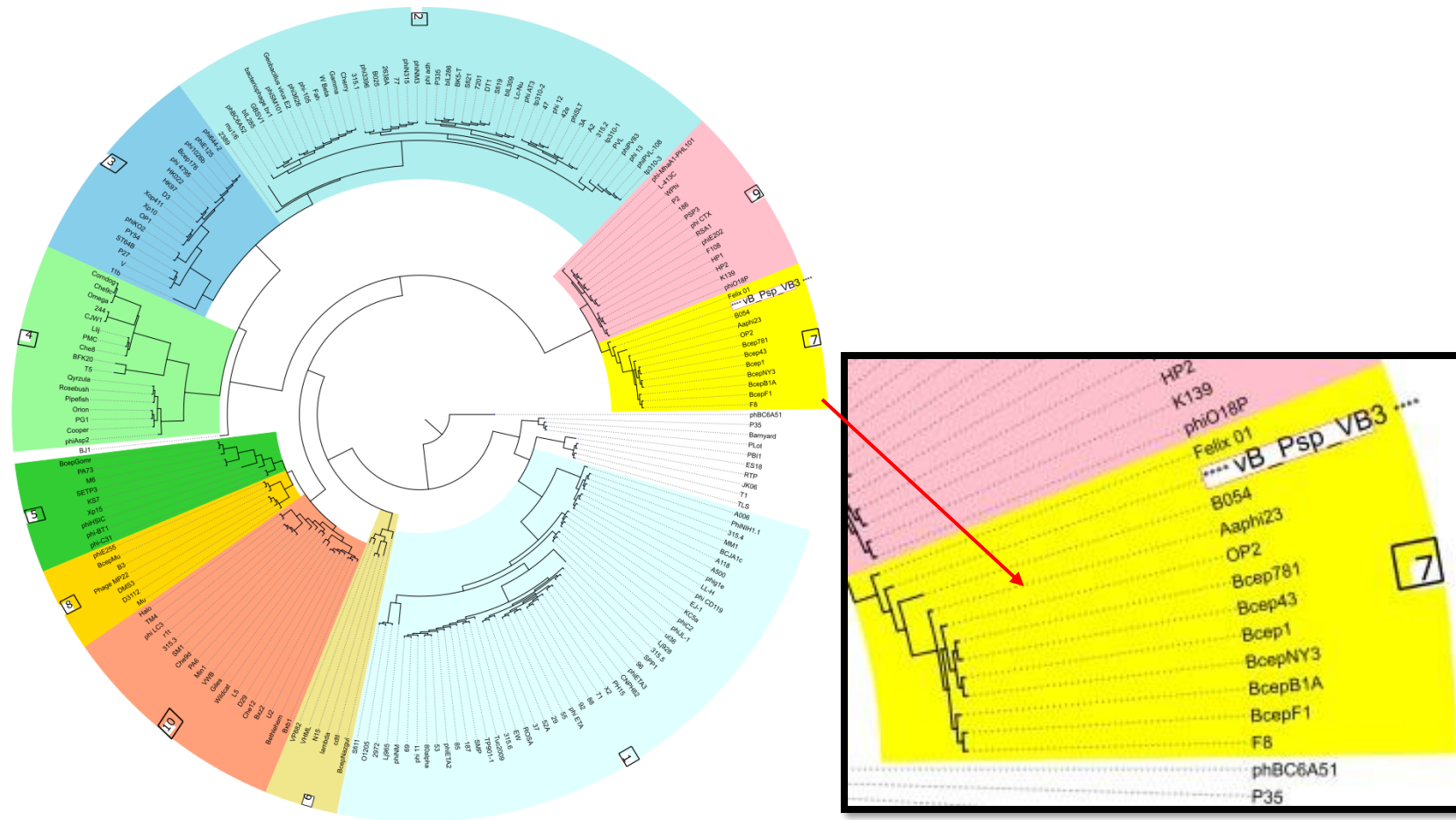


Figure 6. 8: VIRFAM analysis of the phage phage vB_Psp_VB3.
The vB_Psp_VB3 was clustered with Cluster 7 phages

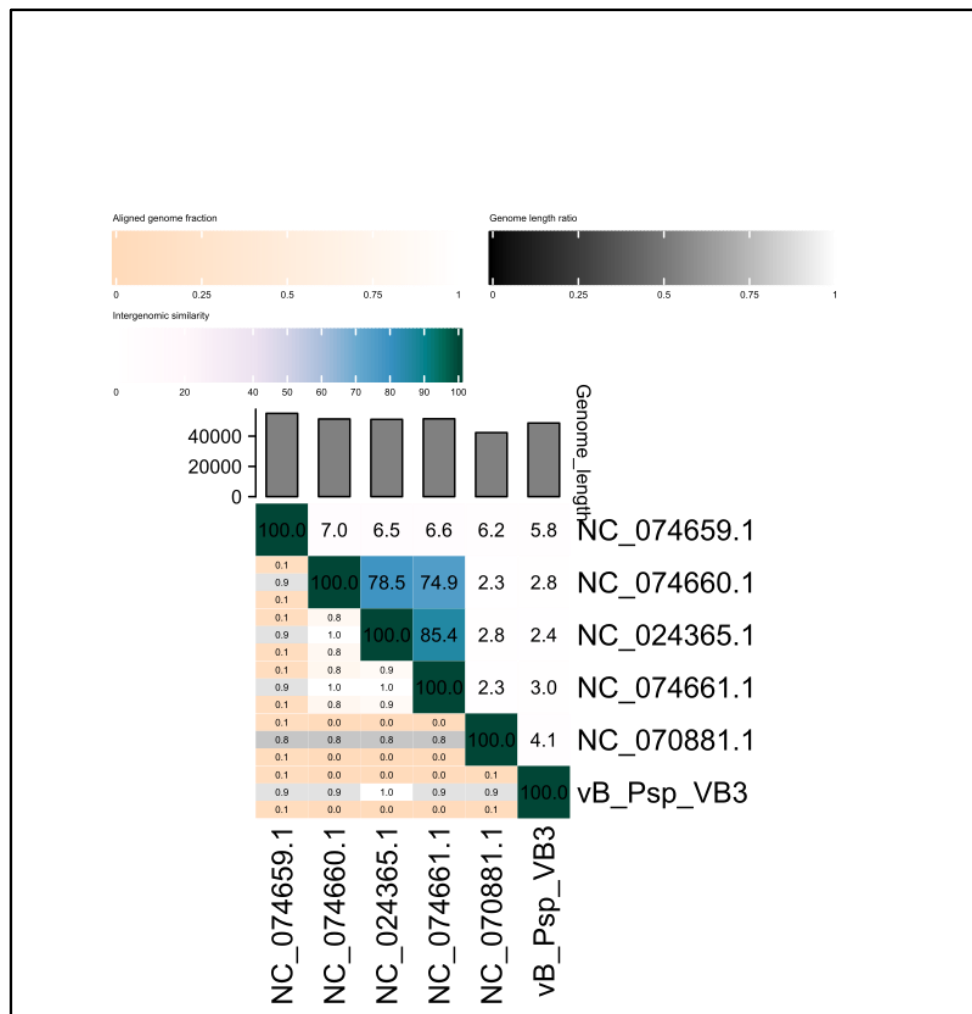


Figure 6. 9: VIRDIC analysis heat map showing the intergenomic distance of the phage vB_Psp_VB3 based on the genome sequence.

The Virus Intergenomic Distance Calculator (VIRDIC) computes pairwise intergenomic distances among the viral genomes (Moraru et al., 2020). Phage vB_Psp_VB3 is *Pseudomonas* phage phiAH14a (NC_074659) with a score of 5.8, which is comparatively less than the distinctly distant based on the intergenomic distance calculated by VIRDIC (Figure 6.9). The heat map showed that the ratio of shared genes between Phage vB_Psp_VB3 and other phages is very low. When the nucleotide sequence homology of bacteriophage is greater than 70%, ICTV regards them as members of same genus (Bin Jang et al., 2019). All phylogenetic analysis showed that the phage vB_Psp_VB3 was notably different from other isolated phages and should therefore be classified as a representative of an unknown virus family.

VirClust analysis (Moraru, 2023) of phage vB_Psp_VB3 based on protein clustering of the phage with other phages in the NCBI Virus database, which much far different from the rest of the phages analyzed (Figure 6. 10 and Figure 6. 11). For phage vB_Psp_VB3, the analysis would involve comparing its protein profiles with those of other phages in the NCBI Virus database. Phage vB_Psp_VB3 was found to be significantly different from the other phages analysed, this could suggest unique evolutionary traits or functional capabilities. Such differences can be due to a variety of factors, including unique host interactions, specific environmental adaptations, or distinct evolutionary pathways.

Recent years have seen significant modifications in the taxonomy and classification of viruses. A viral mega taxonomy based on viral genome characteristics, including common genes (proteins), has replaced the Baltimore categorization system, which was primarily based on the kind of viral nucleic acid (Koonin et al., 2020).

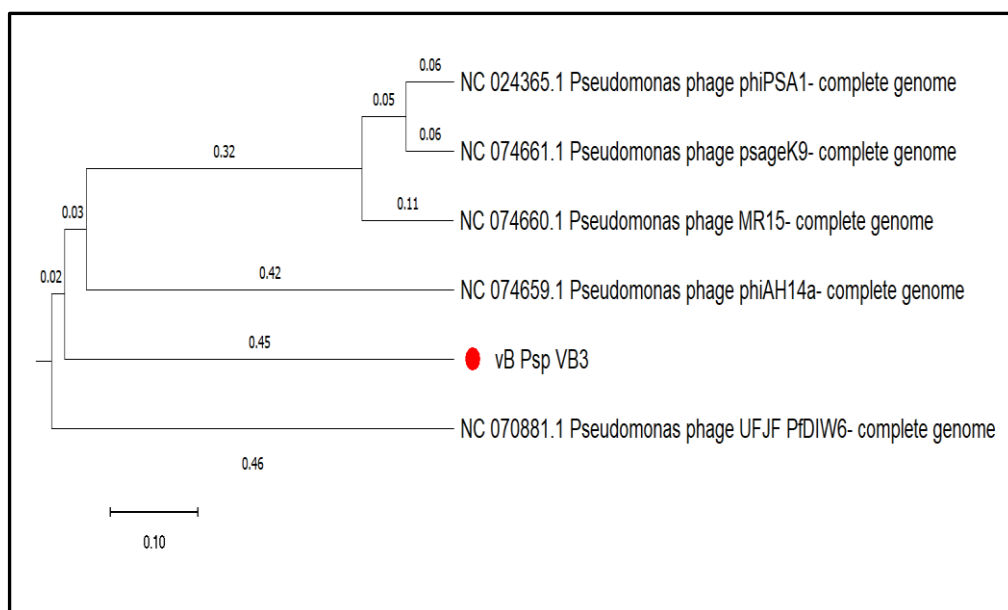


Figure 6. 10: Phylogeny tree based on similarity with NCBI virus complete genomes based on the protein clusters analysis in VirClust (Moraru, 2023) with 1000 Bootstraps.

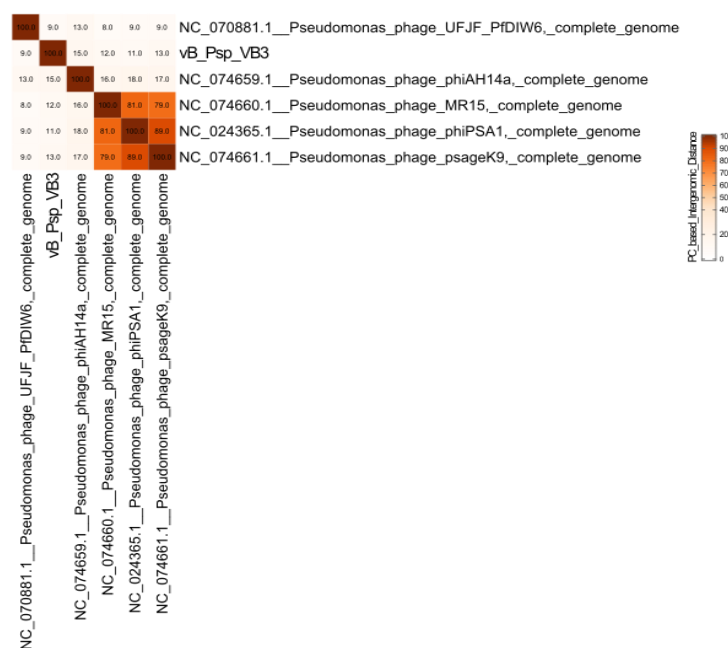


Figure 6. 11: The protein cluster-based intergenomic distance of phage vB_Psp_VB3 based on VirClust analysis.

The isolated phage had a genome supporting a lysogeny lifestyle, a survival strategy used in harsh environments. The genomic analysis of Phage vB_PSp_VB3 confirmed this phage's lysogenic nature, which encodes the typical lysogen-related integrase and Cro and CI regulators on the ends of the genome. The genome analysis of the induced phage was similar to the annotation we observed on prophage region 2. This suggests that the induced phage is similar to prophage region 2 harbored in host genome and hence the induced phage vB_Psp_VB3 was a intact active phage.

There are 1084 complete *Pseudomonas* phage genomes in the NCBI GeneBank database (<https://www.ncbi.nlm.nih.gov>, accessed as on May 16, 2023); more than 750 of these belong to *Pseudomonas aeruginosa* phages, 140 to *Pseudomonas syringae* phages, 60 to *P. fluorescens* phages, and 40 to *Pseudomonas putida* phages. *Pseudomonas agarici*, *Pseudomonas fremontii*, *Pseudomonas chlororaphis*, *Pseudomonas savastanoi*, and *Pseudomonas stutzeri*-specific phage genomes are also presented in the GeneBank.

Two novel *Pseudomonas* species from Taiwan's hot springs have been reported and studied (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; Poddar and Das, 2018; Sharma et al., 2018). Although lysogeny is a favorable lifestyle for bacteriophages in hot springs (Breitbart et al., 2004), particularly in Indian hot springs, they are the least explored. The metavirome study of the Manikaran hot spring in Himachal Pradesh, India (which is closer to the Vashisht hot spring) also reported the presence of temperate *Pseudomonas* phage in their metavirome data sets (Sharma et al., 2018). The phages reported here ranged from sizes 10.1-37.5 kb, which showed temperate and lytic lifestyles with predictions of tailed phages belonging to *Myoviridae*, *Siphoviridae*, or *Podoviridae* families. According to ICTV new taxonomy release (no 37), the majority of the phages in this report belonged to the class *Caudoviricetes* (Turner et al., 2023).

6.4 Conclusion:

- Host range analysis done with limited *Pseudomonas* species appears narrow for phage vB_Psp_VB3.
- The TEM images of the phage samples obtained by PEG precipitation and ultracentrifugation showed structures compatible with phage head and tail structures. The phage belongs to the class *Caudoviricetes* as per the new taxonomy release (No: 37) by ICTV(Turner et al., 2023).
- Sequence analysis of phage vB_Psp_VB3 showed that the phage genome was approximately 48.6 kb in length with GC content of 56.83% containing putative 43 ORF regions.
- Phage vB_Psp_VB3 had the highest sequence similarity to ***Pseudomonas* phage phiPSA1** and ***Pseudomonas* phage phiAH14a** with a sequence identity of 89% and 92%, respectively.
- Phage possesses lysogeny related genes such as integrase, terminase repressor CI, regulatory protein CII, and anti-repressor Cro.
- In these extreme environments, the phage enters the lysogeny state, facilitating long-term co-evolution with the bacterial host and establishing a mutually beneficial interaction for their survival.
- The study suggested that lysogeny may be a viable lifestyle in hot springs.

CHAPTER 7

Metavirome analysis of two hot spring sediments from the West Coast geothermal province (WCGP), Maharashtra, India.

To expand our knowledge and overcome the challenges of conventional isolation techniques, in this chapter, we describe the viral metagenomic analysis of two Indian hot springs, Rajwadi and Tural, belonging to the WCGP in Maharashtra, India.

In the metagenomics era, virus discovery *via* culture-independent techniques that involve sequencing viral genomes directly from the environment is becoming universal. The viral metagenomics (viromes) analysis has produced thousands of viral genomes, delivering unprecedented insights into virus diversity, distribution, and evolution.

Based on the tectonic features and geothermal gradients across India, the Geological Survey of India (GSI) has identified seven geothermal provinces encompassing over 400 thermal springs (Chandrasekharam, 2000; Singh et al., 2016). With temperatures ranging from 42°C at Rajapur to 71°C in Unhavre, the WCGP of Maharashtra, India, contains close to 60 thermal springs clustered around 18 localities. The Rajwadi and Tural hot springs are situated in the southern region of WCGP in the Ratnagiri districts of Maharashtra. The hot springs, a part of the WCGP of India, are found in the N20°W-S20°E fracture zone and are 10 kilometers north of Sangmeshwar on the banks of a tributary that joins the Shastri River. In the past, the geology and geochemistry of these hot springs in WCGP were characterized and reported by GSI (Pitale, et al., 1989; Chatterjee et al., 2016; Gurav et al., 2016; Low et al., 2020). In recent the decade, bacterial life in these extreme environments has been explored for its biotechnological application of thermophiles (Poddar and Das, 2018; Vijayan et al., 2023). The Indian hot springs have a neutral to slightly alkaline pH and therefore harbor a microbial community different from the acidic hot spring environments (Bize et al., 2008; Bolduc et al., 2015; Mathur et al., 2007).

Metagenomics has become a powerful tool for understanding viral diversity and variability since the first marine viral metagenomic research in 2002 (Breitbart et al., 2002). It can directly resolve uncultured environmental viral genomes (Angly et al., 2006; Culley et al., 2006; Hurwitz et al., 2013). Unravelling novel viruses from “**Viral dark**” matter is possible through viral metagenomics (Roux et al., 2015). Investigating microbial communities in extreme environments has revealed hidden viruses that infect uncultivable bacterial hosts, proving that viruses are Earth's most prevalent biological entities (Krupovic et al., 2019). Thus, the extremophiles act as hosts for viral replication, and viruses are responsible for maintaining the balance of population dynamics in extreme environments (Dávila-Ramos et al., 2019).

The recent advance in metagenomic sequencing has revealed an astonishingly large virome in the biosphere. Most of the bacterial hosts from hot springs are uncultivable. The lack of conserved regions in the viral genome makes hot spring viral research extremely challenging (Edwards and Rower, 2005; Pride and Schoenfeld, 2008). However, with the development of next-generation sequencing (NGS) technology in the last few decades, viral research has significantly expanded novel opportunities for characterizing viral diversity and tapping the richness of viral assembly in extreme environments such as hot springs (Bolduc et al., 2015; Zablocki et al., 2017; Castelán-Sánchez et al., 2020; Jarett et al., 2020; Guajardo-leiva et al., 2021).

Despite numerous hot springs in India, there are no reports on the viral diversity of hot springs from the West Coast geothermal province (WCGP), Maharashtra, India. Maharashtra is one such bio-diverse state of India that harbors 60 diverse alkaline hot springs in India's WCGP. Additionally, to expand our knowledge and overcome the challenges of conventional isolation techniques, we assessed the viral metagenomic analysis of two Indian hot springs, Rajwadi and Tural, belonging to the WCGP in Maharashtra, India.

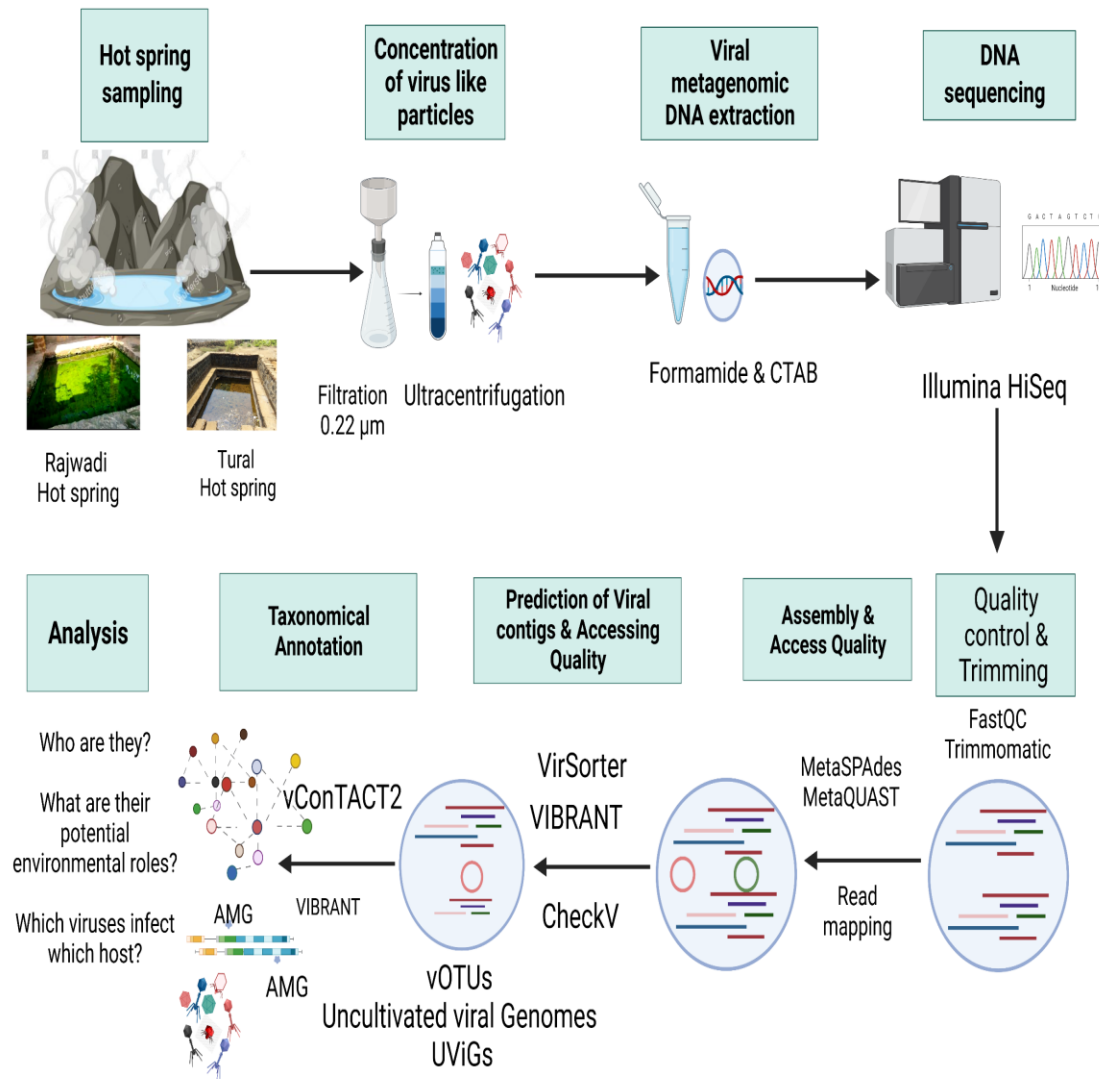
Workflow:

Figure 7.1: Workflow of Metaviromic (viral metagenomic) analysis of two hot springs from the West Coast geothermal province (WCGP) in Maharashtra, India.

Created with BioRender.com (<https://www.biorender.com/>).

7.1 Materials:

Ammonium acetate, PEG 8000, phosphotungstic acid (PTA), CsCl, formamide, chloroform, phenol:chloroform:isoamyl alcohol 25:24:1 mixture, cetyltrimethylammonium bromide (CTAB), DNase I, 0.5M EDTA, Proteinase K, Glycerol, Ethanol, Isopropanol, Gelatin. TE buffer (2M Tris/0.2M EDTA, pH 8.5), Saline Magnesium buffer (pH 7.5).

7.2 Methods:

7.2.1 Sample collection

The basic workflow for generating the viral metagenome is described in Figure 7.1. Water samples and wet sediments were collected from the Rajwadi (17°14'23.1"N 73°33'42.2" E) and Tural (17°14'57.2"N 73°33'16.0" E) hot springs in the Ratnagiri district of Maharashtra belonging to the West coast geothermal province of India in October 2019 (Figure 7.2). The water and wet sediment were collected in sterile bottles from the spring sprout where there were no human anthropologic activities. The collected samples were transported to the laboratory and analysed within six hours.

Water samples collected from the Rajwadi and Tural hot springs in the Ratnagiri district of Maharashtra are constituents of the WCGP of India. The hot springs are located at the N20°W-S20°E fracture zone, 10 kilometers north of Sangameshwar town on the banks of a tributary that joins the Shastri River. These hot springs in WCGP of Maharashtra are manifested in the basaltic terrain of the Deccan traps of Cretaceous-Paleocene age (~65 Ma). The hot water emerges from the margins of dolerite dykes, intense fractures, and shears (Saxena, 1987; Low et al., 2020). Intermittent gas emissions with a sulfurous smell was observed at the sprout of the Rajwadi and Tural hot springs (Figure 7. 3).

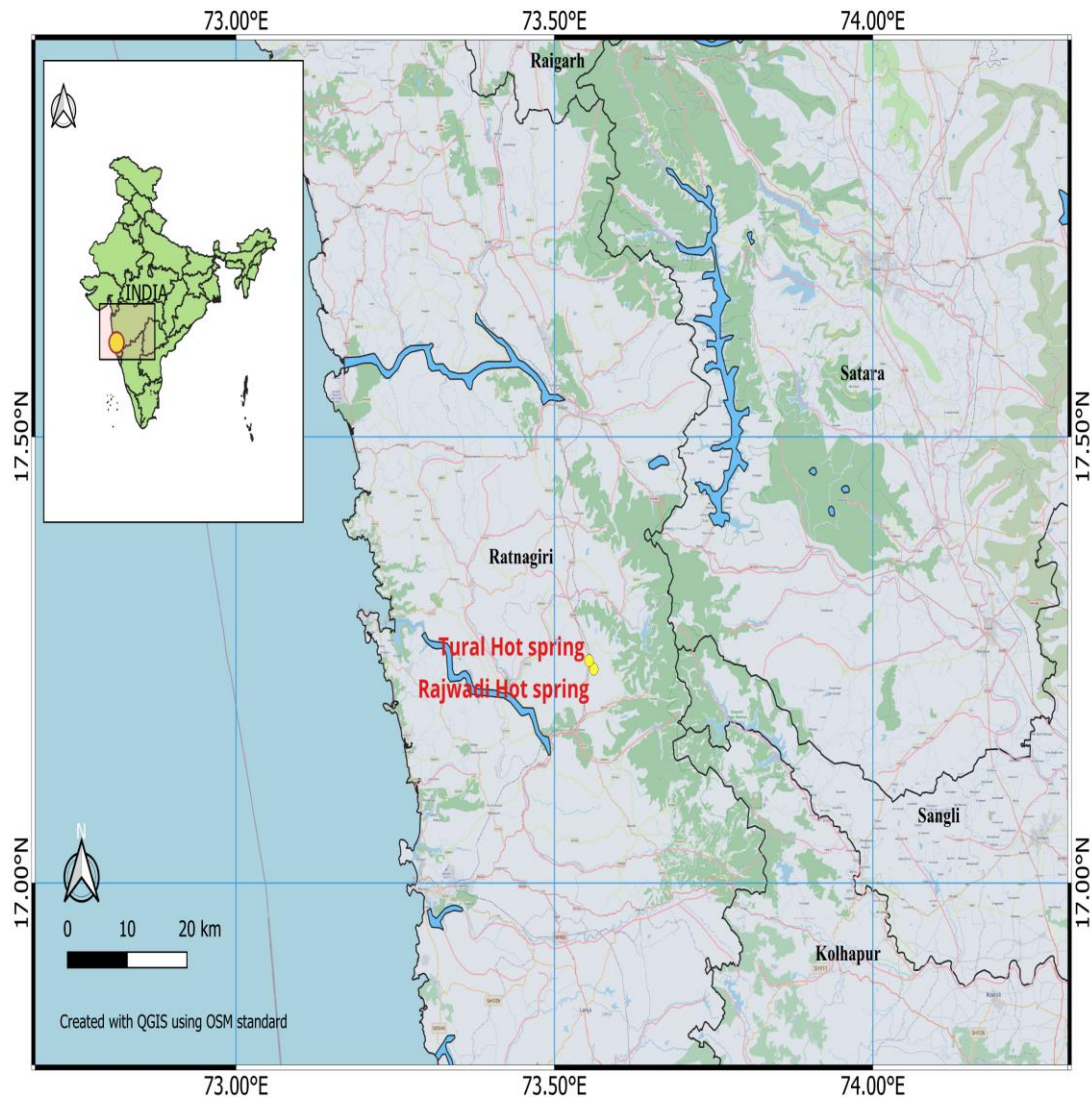


Figure 7. 2: Map showing Rajwadi and Tural hot springs belonging to the WCGP.

The study area map was created using QGIS version 3.28.3. Quantum GIS Development Team. (2017). *Quantum GIS Geographic Information System*. Open Source Geospatial Foundation Project (<https://www.qgis.org>).

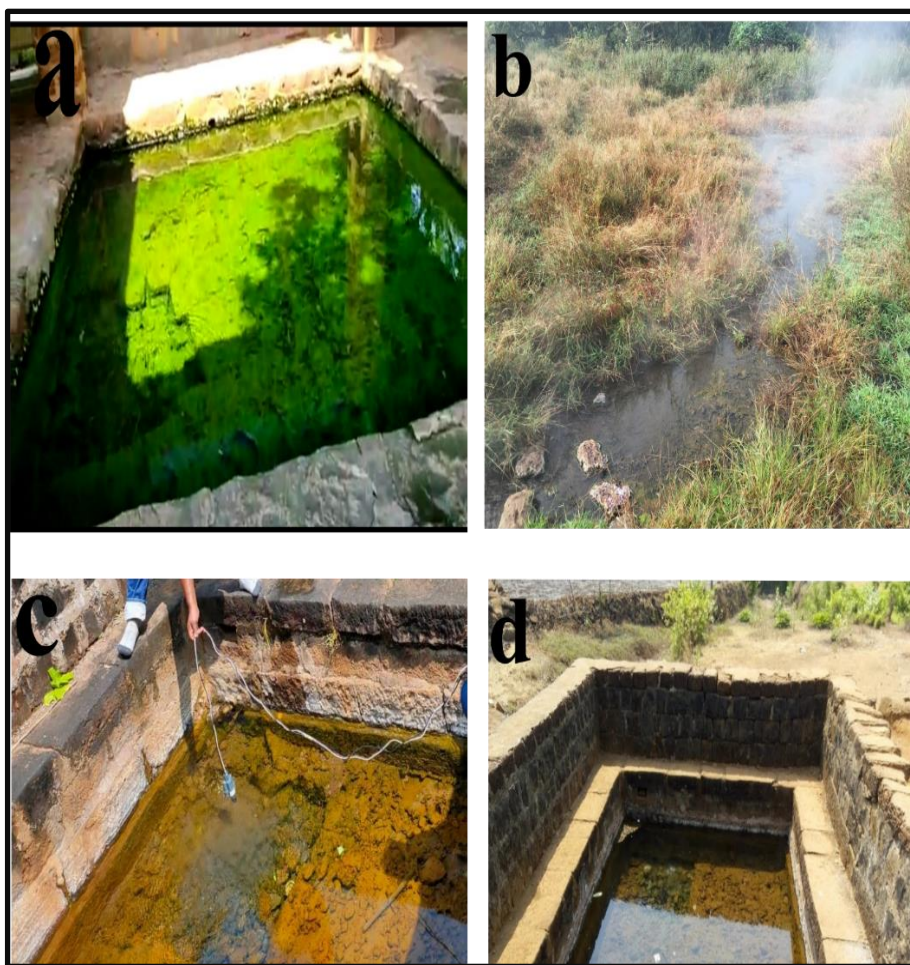


Figure 7. 3: Sampling sites of Hot springs; a and b: Rajwadi hot spring and c & d: Tural hot spring.

7.2.2 Physicochemical analysis of hot spring water sample

In situ, physicochemical parameters such as temperature, pH, electrical conductivity, and TDS of Rajwadi and Tural hot spring water were recorded using a multi-parameter (CyberScan PC 650 Multi-Parameter, Eutech Instruments, Thermo Fisher Scientific, India). The elemental analyses of both hot spring water samples were done using by inductive coupled plasma atomic emission spectrophotometry (ICP-AES, SPECTRO Analytical instrument, GmbH, Germany).

7.2.3 Virus Enrichment from the sediment samples followed by transmission electron microscopy analysis

Five grams of sediment from each of the Rajwadi (MVRS1) and Tural (MVTs1) samples were transferred into a 50 mL tube, and 0.1M Ammonium acetate was added until a total volume of 45 mL was reached (Sharma et al., 2018). The samples were then placed on a shaker for two hours and kept at room temperature in an incubator. The tubes were centrifuged at $10,000 \times g$ for 20 min at 4°C. The supernatant was filtered to remove cells and cell debris using 0.22 μm -pore size syringe filters (Merck Millipore, Germany). The filtrate was stored at 4°C until downstream processing. Virus concentration and purification were carried out with two aliquots (5.5 mL in polyallomer tubes) of each sample in a table-top ultracentrifuge Optima MAX-XP (Beckman Coulter, United States) equipped with a swinging bucket rotor MLS50 (Beckman Coulter, United States). The samples were spun at $80,000 \times g$ at 4°C under vacuum conditions for 4 h. The supernatant was discarded and 50 μL of 0.1 M ammonium acetate (pH 7.0) was added to each centrifuge tube to resuspend the virus pellet and left overnight.

Afterward, 10 μL aliquots of the virus preparations were placed onto a formvar-coated copper grid and left to sit for 1 min to allow the virus to attach to the grid. The excess liquid was blotted off the surface of the grid, and the dried grid was stained with 2% (w/v) phosphotungstic acid (PTA) for 30 seconds. The grid was left to dry in a petri-dish overnight, after which the viral particles were analyzed using High-resolution transmission electron microscopy (HR-TEM) (Tecnai TM).

7.2.4 Concentration of viral particles from hot spring for metagenomic DNA isolation

Nearly 20 L water samples collected from both the hot springs were filtered through 0.22 μm membrane filters (Millipore, USA) and concentrated by Cesium chloride (CsCl) density gradient ultracentrifugation using 1M NaCl and 10% PEG 8000. In the present study, we could not detect sufficient meta viral DNA for sequencing in the water samples. Thus, water samples were excluded from the analyses in the present study.

The viral concentrates for Rajwadi and Tural hot spring sediment samples were prepared with slight modifications as per mentioned protocol (Motlagh et al., 2017). Two fifty grams of the wet sediment samples collected from Rajwadi and Tural hot spring were diluted with three volumes of filter sterilized 1% (w/v) potassium citrate buffer (10 g/l potassium citrate, 1.44 g/l of disodium phosphate, 0.24 g/l of monopotassium phosphate, pH 7). The buffer and the sediment solution were placed on a shaker overnight at 400 rpm at 4°C to suspend the viral particles. The samples were centrifuged for 20 mins at 1500 x g at 4°C to pellet down the bacterial debris and the sediments, and the supernatant was transferred into new tubes for further processing. The retained supernatant was then sequentially filtered through various pore-sized membrane filter units (1.2 μm , 0.8 μm , 0.45 μm , 0.22 μm) (Millipore, USA) to avoid any contamination from microbial cells. The viral particles were concentrated using 1M NaCl and 10% PEG 8000. They were further purified using CsCl density gradient at 1.4–1.6 g/cc density by isopycnic centrifugation at 35,000 rpm in Beckman ultracentrifuge for 4h using MLS-50 rotor (Beckman Ultracentrifuge) at 4°C for 4 h (Angly et al., 2006). The cesium chloride purification was carried out twice to ensure the complete removal of any residual bacterial cell debris and other particles. The purified virus fractions were dialyzed overnight against SM buffer (5.8 g/l sodium chloride, 2 g/l magnesium sulphate, 5 ml/l of 2% (w/v) gelatin, 50 ml/l of 1M Tris-HCl, pH 7.5). Subsequently, the concentrated viral fraction was used to extract metagenomic DNA.

7.2.5 Metagenomic Viral DNA extraction from sediment

The metagenomic viral DNA was extracted from the concentrated viral lysate of Rajwadi (MVRS1) and Tural (MVTS1) sediments samples and was processed using the CTAB method. (Sambrook et al., 1989, Thurber et al., 2009). The viral sample was subjected to DNase treatment with RNase-free DNase I 2.5U/ml for 1 h at 37°C, followed by the

DNase I inactivation at 75°C for 5 min before the lysis process, to remove any residual DNA in the sample. The DNase-treated samples were further concentrated using a centrifugal concentration filter (Amicon Ultra-15 100kDa cutoff, Millipore, USA) before DNA extraction. The metagenomic viral DNA was extracted using formamide and CTAB protocol (Sambrook, 1989; Thurber, 2009) with slight modifications. For every 100 µl of DNase I treated filtrate, 0.1 volume of 2 M TrisCl (pH 8.5)/0.2 M EDTA, 100 µl of 0.5 M EDTA (pH 8.0) per 10 ml sample, 1 volume of formamide and 1 µl of glycogen (20 mg/ml) was added. The mixture was incubated at room temperature (30°C) for 30 min. After incubation for 30 min at room temperature, two volumes of absolute ethanol was added. The DNA was pelleted by centrifugation for 20 min at 8000 x g at 4 °C. The pellet was washed twice with 70% ethanol and air dried for approximately 15 min. The pellets were pooled and resuspended in 567 µl of TE buffer and stored at -20°C. To the resuspended pellet, 30 µl of 10% (w/v) SDS and 3 µl of 20 mg/ml proteinase K were added and mixed thoroughly and incubated for 2 h at 55°C. After lysis, 100 µl of 5 M NaCl was added and mixed thoroughly. 80 µl CTAB/NaCl solution was added to the lysed mixture and incubated for 10 min at 65°C. Metagenomic viral DNA was extracted using phenol: chloroform: isoamyl alcohol (25:24:1) extraction at 12000 x g at 4°C for 5 min. Again, the aqueous phase was extracted by chloroform: isoamyl alcohol (24:1) and centrifuged at 12000 g at 4°C for 5 min. The supernatant was mixed with 1/10 volume of 3M sodium acetate (pH 5.2) followed by the addition of 0.7 volume of isopropanol and incubated overnight at 4°C. The precipitated viral metagenomic DNA was centrifuged at 12000 x g for 20 min at 4°C. The pellet was washed with 70% ethanol and air-dried. The metagenomic viral DNA was resuspended in 50 µl 10 mM Tris-HCl (pH 8.0) and stored at -20°C. The metagenomic viral DNA was quantified on Qubit 2.0 fluorometer using Qubit HS dsDNA kit (Life Technologies, USA) and checked by electrophoresis on a 0.7% agarose gel.

7.2.6 NGS Library preparation and Illumina sequencing.

Next generation sequencing (NGS) library was prepared by Illumina-compatible NEXTflex Rapid DNA sequencing Bundle (BIOO Scientific, Inc. USA.) Briefly, quantified viral DNA was sheared using Covaris S220 sonicator (Covaris, Inc. USA) to generate a specific fragment size range of 200-300 bp. The fragments were purified using HiPrepclean up beads. Purified fragments were end-repaired, adenylated, and ligated to Illumina

multiplex barcode adaptors as per NEXT flex Rapid DNA sequencing bundle kit protocol. The adapter-ligated DNA was purified using High Prep beads. The resultant fragments were PCR amplified for 8 cycles using Illumina-compatible primers provided in the NEXTFlex Rapid DNA sequencing kit. The final PCR product was purified with High Prep cleanup beads and a library quality control check. The sequencing libraries were quantified by Qubit fluorometer (Thermo Fisher Scientific, MA, USA). The library fragment size distribution was analyzed on Agilent 2200 Tape Station (Illumina, USA), which showed a 200–900 bp range for all libraries. Samples were then loaded onto an Illumina HiSeq cartridge (Illumina, CA, USA) and run in 2 x 150 mode on an Illumina HiSeq next-generation sequencer (HiSeq. 2500 sequencer, Illumina, CA, USA) as per the manufacturer's instructions (<https://www.illumina.com>).

7.2.7 Metagenomic sequence assembly and annotation:

The iVirus pipeline (Bolduc, 2021; Bolduc et al., 2021; Bolduc et al., 2017) through the CyVerse platform (Goff et al., 2011) in Discovery Environment (<https://de.cyverse.org/de/>) was used for processing the viral metagenomics data. The raw paired-end reads generated by Illumina HiSeq sequencer for both Rajwadi (MVRS1) and Tural (MVTS1) sediments were first quality checked using FastQC v 0.11.9 (Andrews, 2010) and then trimmed using Trimmomatic v.0.38 (Bolger et al., 2014) with a cutoff score as 30 and a minimum length set at 50. The trimmed sequences were assembled using metaSPAdes version 3.9.0.1 (Nurk et al., 2017) with assembly parameters K-mer values: (21,33,55,77, 99, 121) and default settings. Only contigs longer than 1 kb were retained and designated as MVRS1 and MVTS1 for Rajwadi and Tural datasets, respectively.

7.2.8 Taxonomic Annotation of Metagenome

Taxonomic assignments were performed with the software **Kaiju** v 1.8.2 (Menzel et al., 2016) against the RefSeq complete genome database sourced from the National Center for Biotechnology Information (NCBI) databases using the maximum exact matches and 11 as a minimum match length.

7.2.9 Taxonomical classification and abundance

The putative microbial population reads from both samples were classified using Kraken against the reference database containing all Ref-Seq virus genomes (Wood et al.,

2014). Krona was used to visualize the abundance of viral families in both meta virome datasets.

7.2.10 Prediction of viral sequences

In the present studies, VirSorter v1.0.3 (Roux et al., 2015) and VIBRANT v1.2.0 (Kieft et al., 2020) were used to identify viral contigs in the hot spring metagenome assemblies, followed by manual curation. The assembled sequences of MVRs1 and MVTS1 were examined using VirSorter (Roux et al., 2015) to identify the function and phylogeny of viruses within assembled metagenomes. In the first approach, VirSorter v 1.0.3 (Roux et al., 2015) was used to determine the viral contigs, using the viral hallmark genes annotated as “main capsid protein”, “portal”, “large subunit of the terminase”, “tail”, and “envelope”, among others. The entire contig was considered viral if more than 80% of predicted genes on a contig had a viral signal. The CyVerse Discovery Environment (<https://de.cyverse.org/de/>) was used to run the application VirSorter 1.0.3 (Roux et al., 2015) with the viromes database and decontamination mode. A list of predicted viral contig sequences were obtained and organized by category from 1 to 6. In VirSorter, categories 1 and 4 contigs were assigned to categories with the most confidence, whereas categories 2 and 5 were assigned to likely predictions.

7.2.11 Viral clustering and classification

Viral vConTACT2 v.0.9.19 (Bolduc et al., 2017; Bin Jang et al., 2019) was used to cluster and classify selected viral scaffolds against the Prokaryotic Viral Refseq database, and the resulting clusters were subsequently visualized using Cytoscape v 3.9.1 (<https://cytoscape.org/>) (Shannon et al., 2003). The network was filtered and manually curated by removing duplicated edges and applying the Edge-weighted Spring Embedded Layout.

7.2.12 Viral contig prediction and AMG prediction

Alternatively, a complementary approach was also used. The program VIBRANT v1.2.0 (Kieft et al., 2020) was used to predict viral contigs among all metagenome contigs of Rajwadi and Tural. VIBRANT is a unique and novel tool that identifies viral contigs by integrating a supervised machine learning approach with viral scoring obtained by annotations using Pfam, Virus Orthologous Groups (VOG), and Kyoto Encyclopedia of

Genes and Genomes (KEGG). The circular viral contigs predicted by VIBRANT as for MVRS1 and MVTs1 datasets were uploaded to the Viral Proteomic Tree server (<http://www.genome.jp/viptree>) (VIPTree v.3.4) to generate a proteomics tree for classification of viruses based on genome-wide similarities (Nishimura et al., 2017).

7.2.13 Identification, annotation, and classification of predicted viral genomes.

A recent novel computational framework tool, geNomad (Camargo et al., 2023), that identifies, annotates, and categories viruses and plasmids in sequencing data were used for our UViGs of MVRS1 and MVTs1, which VIBRANT and VIRSorter predicted. The geNomad online web application was run for this purpose on the NMDC-EDGE platform (<https://portal.nersc.gov/genomad>), which also incorporates the CheckV tool (Nayfach et al., 2021) for the predicted viral genomes of MVRS1 and MVTs1. Among all the circular contigs predicted by VIBRANT and VIRSorter, the completeness and quality of the predicted viral contigs were done by CheckV. This novel tool geNomad predicts a score and assigns the identified virus to the taxonomical lineage that follows ICTV taxonomy release. The UViGs, which showed good score and completeness, was taken for further analysis.

7.2.14 Genome analysis of the predicted complete circular contig.

Among all the viral contigs analysed using the geNomad tool, the contig NODE_171 followed the standards of completeness with a good geNomad score above 0.7. The top hits determined the closest relative for the uncultivated viral genome NODE_171, identified in the NCBI virus database [organism: viruses (tax id: 10239)] (<https://www.ncbi.nlm.nih.gov/labs/virus/vssi/>) using BLASTn program with E-value cutoff 10^{-5} . Additionally, the viral contig NODE_171 was searched against the largest uncultivated viral sequence database, IMG/VR v.4 (<https://img.jgi.doe.gov/cgi-bin/vr/main.cgi>) using the BLASTn program with an E-value cutoff of 10^{-5} (Camargo et al., 2023).

7.3 Results and discussion

7.3.1 Physicochemical analysis of Rajwadi and Tural hot spring water

The temperature, pH and other physical parameters of the water samples were recorded onsite (Table 7. 1). The temperature of Rajwadi and Tural hot spring water was recorded as 55°C and 60°C, respectively, whereas the pH was determined as 7.5 and 7.7, respectively. Most hot springs in India report a temperature range between 40 to 98°C (Poddar and Das, 2018). However, the hot springs of WCGP have temperatures ranging from 42°C (Rajapur hot springs) to 71°C (Unhavre hot springs) (Gurav et al., 2016). The Indian hot springs were either neutral, moderately alkaline in nature (Poddar and Das, 2018) and harbored a microbial community different from the acidic hot springs across the world (Mathur et al., 2007). Table 7. 1 depicts the physicochemical parameters of Rajwadi and Tural hot spring water analyzed using ICP-AES. The analysed physicochemical parameters of Rajwadi and Tural hot spring water samples validated previous observations reported elsewhere (Chandrasekhar et al., 2016; Gurav et al., 2016). The Ca concentration in Rajwadi and Tural hot springs were similar to the values reported elsewhere (Chandrasekhar et al., 2016; Gurav et al., 2016). Previous physicochemical studies of WCGP thermal springs revealed that chloride content varies from 78 to 4800 ppm, sodium 76 to 410 ppm, with low magnesium, boron, and fluoride content (Pitale et al., 1989; Sarolkar, 2005; Chatterjee et al., 2016). In general, the spring was indicative of a Na-Ca-Cl or Ca-Na-Cl spring which is the case for other reported WCGP thermal springs (Gurav et al., 2016). Although, elements such as Mn, Cu, Fe, Zn, and Pb were reported in acidic geothermal springs they were almost below the limit of quantification in the water sample collected from the Rajwadi and Tural hot springs. The alkaline hot springs at Yellowstone National Park (63°C), reported elevated levels of sodium, chloride, and fluoride. In contrast, heavy metals, such as Pb, Cr, Cu, Hg, *etc.*, were below the limit of quantification (Bowen De León et al., 2013). Since the WCGP hot springs are near neutral to alkaline in nature, with pH varying from 7.08 to 8.82 from Koknere to Rajapur (Gurav et al., 2016; Low et al., 2020), these elements were below the limit of quantification.

Table 7. 1: Physicochemical parameters of Rajwadi and Tural hot spring, Maharashtra, India.

Hot spring	<i>In situ</i> temperature (°C)	pH	Conductivity (μS/cm)	TDS	Na ⁺	Ca ²⁺	Mg ⁺	Cl ⁻	K ⁺	S	Fe	Cu	Mn	Pb	Zn
Rajwadi	55	7.5	1445	956	230	53.3	0.498	354	7.5	36.25	0.014	*	*	*	*
Tural	60	7.7	1475	945	232	51.1	0.834	415	7.4	33.23	0.01	*	*	*	*

*: (Not detected) < 0.01 ppm

7.3.2 Virus Enrichment from Sediment Samples Followed by Transmission Electron Microscopy

TEM images of viruses isolated from the enriched Rajwadi sediment sample predominantly showed the presence of tailless virus-like particles. These virus particles were the most dominant ones, with a more than 100 nm diameter in the Rajwadi hot spring sample (Figure 7. 4 a). However, the virus-like particles from Tural hot spring had a size of approximately 50-60 nm width (Figure 7. 4 b). Unlike other hot springs, no tailed virus particles or other morphology of virus particles were observed during TEM analyses. These can be attributed to low concentrations of viral particles or the sampling pre-processing. Viruses are lost at every step of purification and concentration. Large viruses might have been possibly removed typically during the filtration step. Some may have been sensitive to chemicals used during purification based on the structure of their capsids. In contrast, the viruses might also have been lost due to the differences in buoyant density, which is critical during density gradient ultracentrifugation (Thurber, 2009). Thus, selective filtration strategies are recommended to recover extreme viruses. Since morphological characterization cannot establish the taxonomic identity of the viruses, genomic analysis was conducted to elucidate the viral diversity of Rajwadi and Tural hot springs of WCGP.

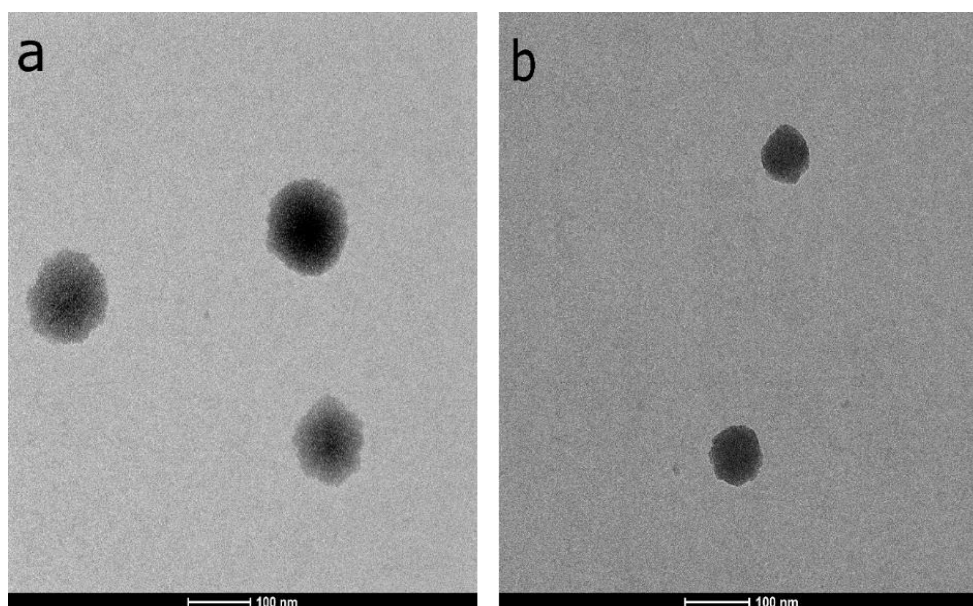


Figure 7. 4: TEM images. a) Rajwadi; b) Tural sediment sample; Scale bar 100 nm.

7.3.3 Shotgun metavirome sequencing

Illumina HiSeq sequencing was used to sequence the NGS library of the metavirome of Rajwadi (MVRS1) and Tural sediment (MVTs1) hot spring sediment samples. The procedure yielded 23,282,842 million reads with an average length of 150 bps and 14,868,246 million reads with an average length of 150 bps (Table 7. 2). We used the iVirus pipeline (Bolduc et al., 2017; Bolduc, 2021) through the CyVerse platform (Goff et al., 2011 in Discovery Environment (<https://de.cyverse.org/de/>) for processing the viral metagenomics. Out of 23,282,842 generated sequencing reads for the MVRS1 sample, 22,399,288 (22.39 million) HQ reads with phred score ≥ 30 were obtained after quality filtering and trimming. Similarly, out of 14,868,246 generated sequencing reads for MVTs1, 14,460,084 (14.46 million) reads with phred score ≥ 30 were obtained after quality filtering and trimming. After trimming and metaSPAdes, assembly statistics were analyzed using metaQUAST (Table 7. 3). The viral metagenomic reads in this study were deposited at the National Center for Biotechnology Information (NCBI) sequence read archive (SRA) (<https://www.ncbi.nlm.nih.gov/sra>) repository under the BioProject **PRJNA906748**.

Table 7. 2: Basic read statistics obtained from Illumina sequencing.

	Rajwadi sediment (MVRS1)	Tural sediment (MVTs1)
Total raw reads (million)	23,282,842	14,868,246
Read of poor quality	0	0
Mean Sequence length (bp)	150	150
Mean GC content (%)	58	58

Table 7. 3: After trimming and metaSPAdes assembly statistics using metaQUAST.

	Rajwadi sediment (MVRS1)	Tural sediment (MVTS1)
After trimming reads	22,399,288	14,460,084
Total no of contigs	214960	470702
Largest contig	1454558	175348
Total length	82740058	124963381
GC (%)	62.36	58.35
N50 value	17975	2530

7.3.4 Microbial Community Composition of Rajwadi and Tural Metagenome

The MVRS1 and MVTS1 hot spring wet sediments represented all three domains of life and viruses. The diversity and abundance of microorganisms in the hot spring samples from Rajwadi and Tural hot springs were determined by shotgun metagenomic sequencing (Figure 7. 5). The microbiome diversity was determined from the metagenomes obtained from these two locations. To this end, the assembled contigs were classified with Kaiju (Menzel et al., 2016). The results at the domain level showed that bacteria in MVRS1 and MVTS1 represent between 97% and 98% of the total microorganisms, followed by Archaea (0.1 and 2%), and viruses (0.1 and 0.04%), respectively (Table 7. 4). The metagenome microbial compositions were consistent with those in other WCGP hot springs, such as Unkeshwar (Mehetre et al., 2016) and Tural hot springs in Maharashtra (Mathan Kumar et al., 2022). Additionally, another metagenomic microbial diversity analysis in Maharashtra, Gujarat, and Telangana hot springs (temperature ranging 38-68°C; pH 7-7.6), suggested bacteria as the dominant group (99.3%) when compared with the archaeal group (0.7%) in these hot springs (Narsing et al., 2021). Similar microbial community composition was also reported with a low abundance of Archaea and Viruses in Chignahuapan, Puebla, Mexico springs (Castelán-Sánchez et al., 2020) that are said to have a neutral pH and a moderate temperature (50°C). Similar to our results (Figure 7. 5), bacteria were the predominant life

forms reported in all the Indian hot springs, with an abundance of phyla *Firmicutes*, *Proteobacteria*, *Actinobacteria*, *Thermi*, *Bacteroidetes*, *Deinococcus-Thermus* and *Chloroflexi* (Poddar and Das, 2018; Vijayan et al., 2023). Archaeal organism abundance was low in the Rajwadi and Tural Hot spring samples and is consistent with observations from other neutral springs in Maharashtra (Mehetre et al., 2016; Rao et al., 2021; Kumar et al., 2022) and circumneutral hot springs in Chignahuapan, Puebla, Mexico (Castelán-Sánchez et al., 2020).

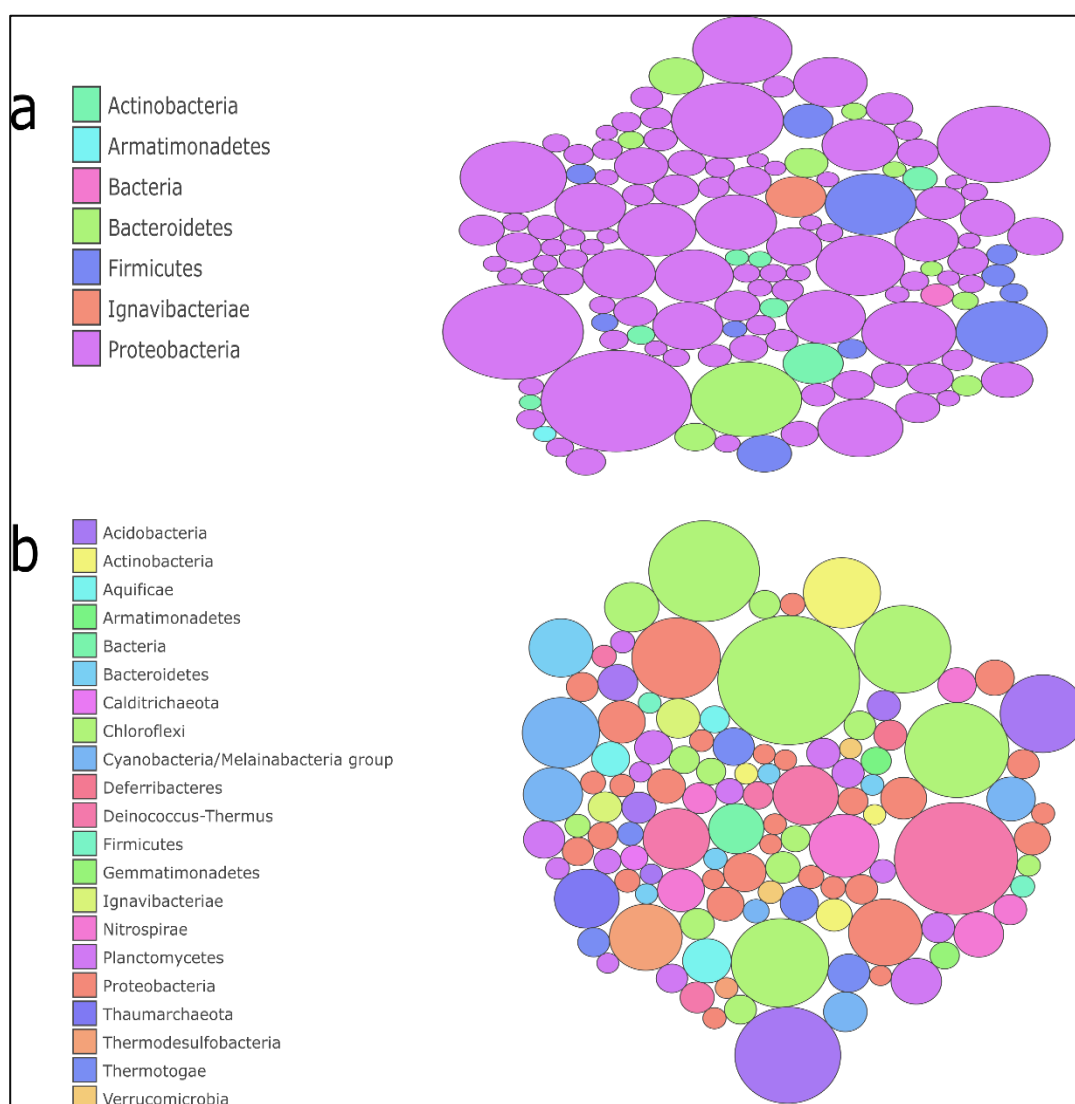


Figure 7. 5: Overview of output from Kaiju analysis

a) MVRs1 and b) MVTS1.

Table 7. 4: Total metagenome contigs from Rajwadi and Tural sediment samples classified by Kaiju analysis.

Sample	No of total contigs	Total Contigs		% Contigs classified within Domain		
		Classified	Unclassified	Bacteria	Archaea	Virus
MVRS1	214960	143745 (67%)	71215 (33%)	139875 (97%)	48 (0.1%)	137 (0.1%)
MVTS1	470702	178392 (38%)	268 (62%)	178044 (98%)	3794 (2%)	78 (0.04%)

7.3.5 Viral taxonomy classification and abundance analysis by Kraken Virus database

The Kraken virus database processed the high-quality assembled reads for taxonomy viral abundance and classification. The analysis revealed the presence of viral families that infected bacterial hosts such as *Siphoviridae*, *Myoviridae*, and *Podoviridae* belonging to order *Caudovirales* (now known as class *Caudoviricetes*) (Figure 7. 6 and Figure 7. 7). Additionally, viruses infecting invertebrates such as *Baculoviridae*, eukaryotic algae *Phycodnaviridae*, and Protozoan *Mimiviridae* were also observed, which was also similarly reported in the metavirome studies in the Sikkim Himalayan hot spring (Das et al., 2020). The analysis indicates the ubiquitous presence of *Myoviridae* and *Siphoviridae* families in the viral metagenomes of MVRS1 and MVTS1 (Table 7. 5).

Interestingly, certain reads in the datasets of MVRS1 and MVTS1 belonged to unclassified giant DNA viruses from the genera *Pandoravirus* containing species such as *Thermus_phage_phi_OH2*, *Thermus_phage_IN93*, *Thermus_phage_P23-77*, *Deep-sea_thermophilic_phage_D6E* and *Geobacillus_virus_E2*. The study results were in concurrence with the metavirome analysis of solfataric mud sediments from Sikkim

Himalayan hot springs that revealed the dominance of dsDNA viruses in class *Caudoviricetes*, (Das et al., 2020) and the presence of giant DNA virus-like genera *Pandoravirus*. On the contrary, the presence of archaeal viruses from the *Fuselloviridae* family and dsDNA bacteriophages in class *Caudoviricetes* were abundant in the Himalayan Manikaran hot spring (Sharma et al., 2018). The presence of archaeal viruses reported by Sharma and coworkers in Manikaran, Himalayan hot spring may be attributed to high temperatures of 98°C. In contrast to the well-researched acidic and higher temperature hot springs around the world, the MVRS1 and MVTs1 metavirome viral analysis and TEM studies, did not demonstrate the presence of archaeal viruses (Rachel et al., 2002; Prangishvili, 2013). *Thermus*-infecting phages have been majorly reported from the various hot springs worldwide. *Thermus myovirus* phiYS40, the first discovered virus from a hot spring, was isolated from Japan almost four decades earlier (Sakaki and Oshima, 1975). Soon after this great discovery, many studies slowly followed the worldwide discovery of viruses or VLPs (Rachel et al., 2002).

Rajwadi and Tural springs have a neutral pH and are moderately hot (55-60°C). The Manikaran sediment sample from the Himalayan hot spring (78 to 98°C) revealed archaeal viruses belonging to the family *Fuselloviridae* with varying sizes as well as tail and tailless bacteriophages from microbial mat samples (Sharma et al., 2018). The temperature and pH play a pivotal role in influencing the microbial community. Thus, the Manikaran Himalayan hot spring sediment samples were observed to be dominated by archaea and microbial mat (50°C) and had a high abundance of the lysogenic bacterial population (Sangwan et al., 2015).

Table 7. 5: Viral Taxonomical Annotation of viral contigs MVRs1 and MVTS1 using Kraken Virus database.

	Rajwadi sediment (MVRs1)	Tural sediment (MVTS1)
Order <i>Caudovirales</i> (total)	53%	44%
Family <i>Siphoviridae</i>	26%	13%
Family <i>Myoviridae</i>	22%	26%
Family <i>Podoviridae</i>	2%	4%
Order <i>Herpesvirales</i>	8%	14%
Family <i>Phycodnaviridae</i>	4%	4%
Family <i>Polydnaviridae</i>	4%	2%
Family <i>Inoviridae</i>	1%	nd
Family <i>Mimiviridae</i>	0.8%	1%
Family <i>Nudiviridae</i>	nd	0.6%
Unclassified viruses	10%	13%

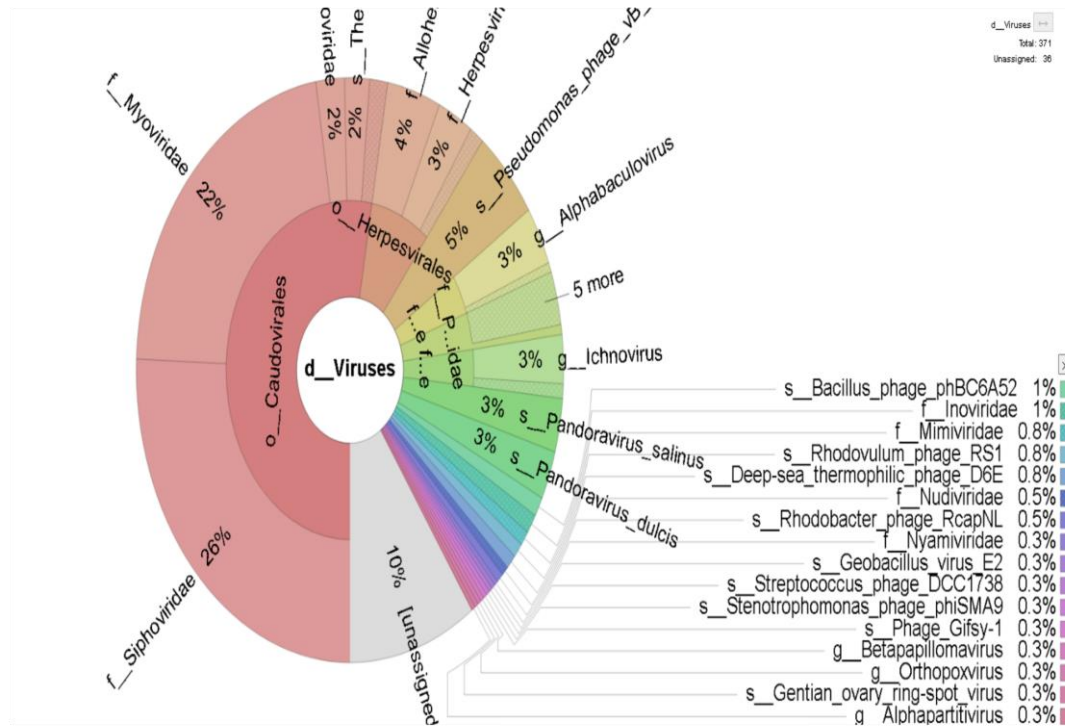


Figure 7. 6. Viral community composition in MVRs1 (Kraken virus database).

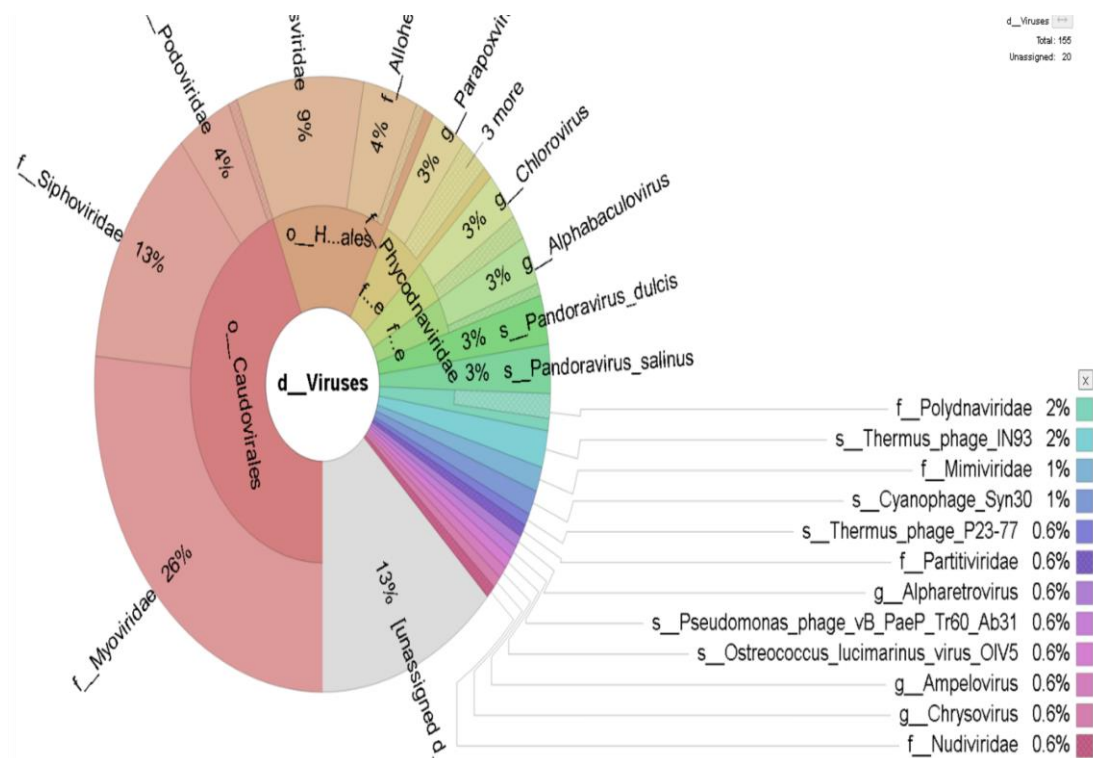


Figure 7. 7: Viral community composition in MVTs1 (Kraken virus database)

7.3.6 Viral prediction, clustering and curation

VirSorter predicted two types of viral-gene group categories: lytic and lysogenic viral contigs, with both lytic and lysogenic genes being further divided into confidence categories (the most confident ones category 1 and category 2 likely prediction and Category 3 possible predictions), while firstly eliminating all bacterial genes. VirSorter predicted 59 and 87 lytic viral contigs from MVRs1 and MVTS1 samples, respectively. The lysogenic (prophage) viral contigs predicted by VirSorter were 2 and 15 for the lysogenic category in MVTS1 and MVRs1 samples, respectively (Table 7. 6). In this approach, VirSorter v 1.0.3 (Roux et al., 2015) was used to determine the viral contigs, using the viral hallmark genes annotated as “main capsid protein”, “portal”, “large subunit of the terminase”, “tail”, and “envelope”, among others (Roux et al., 2015). The entire contig was considered viral if more than 80% of predicted genes on a contig had a viral signal.

Table 7. 6:Summary of the number of predicted viral contigs identified using VirSorter

Sample Name	Virsorter viral Category 1 (Lytic)	Virsorter viral Category 2 (Lytic)	Virsorter prophage Category 1 (Lysogenic)	Virsorter prophage Category 2 (Lysogenic)	Circular
MVRs1	7	52	2	13	2
MVTS1	13	74	0	2	0

Hot springs comprise microbial communities that are dominated by a few species of microorganisms and, as a consequence, have become a model for studying community ecology, the biotic and abiotic interactions that influence their structure. Viruses in hot springs have been proven to be ubiquitous, abundant, and active members of these communities (Breitbart et al., 2004). According to the "kill-the-winner" hypothesis, lytic viruses specifically target dominant or active species in the microbial community and selectively lyse them. This has an impact on the composition of microbial community composition through predator-prey dynamics (Thingstad, 2000). On the contrary, the "piggyback-the-winner" ecological model (Knowles et al., 2016) has predicted that viruses

lysogenize the most metabolically active and occasionally dominant species in the microbial community. A study by Guajardo-Leiva and coworkers suggest that lysogeny is a viable lifestyle in hot springs (Guajardo-leiva et al., 2021). The viruses infecting cyanobacterium *Fischerella* sp., some of the most active and abundant phototropic microbial mats, which physically support the mat and are responsible for nitrogen-carbon fixation. Moreover, a study in hot springs confirms that lysogeny is a broadly distributed state within viral populations that infect chemoheterotrophic bacteria of the phyla *Proteobacteria* and *Firmicutes*. These phyla take advantage of the abundant nutrients derived from primary producers in these thermophilic phototropic microbial mats; therefore, their associated lysogenic viruses better fit the “piggyback-the-winner” model. Similarly, lysogeny-associated viral genomes were found in the metagenome of Himalayan Manikaran hot spring microbial mats, and sediments were associated with *Proteobacteria* and *Firmicutes* (Sharma et al., 2018).

7.3.7 Viral clustering

VIRSorter v 1.0.3 programs were used to retrieve high-quality viral lytic contigs and prophage contigs and processed with vConTACT2 v 0.9.19 (Bolduc et al., 2017, Bin Jang et al., 2019), that compares both the presence and absence of homologous proteins between viral contigs and the sequence identity between homologous proteins and then generates probabilistic clusters of contigs with high similarity. Protein clusters were developed with the MCL algorithm, and viral clusters (VCs) were generated with the cluster ONE algorithm. vConTACT2 v 0.9.19 was used to cluster and classify selected viral contigs against the Prokaryotic Viral RefSeq v94 database. The resulting clusters were subsequently visualized using Cytoscape v 3.9.1 (Figure 7. 8 and Figure 7.9). Nodes depict viral contigs and are coloured according to the source of the contig. Edges connect significantly similar nodes, meaning that the contig pair shares more protein clusters at higher similarity than expected by chance. The viral contig NODE_171 clustered with RefSeq dataset of viruses with *Rhizobium* phages based on vConTACT2 v 0.9.19 analysis. (Figure 7. 8 b).

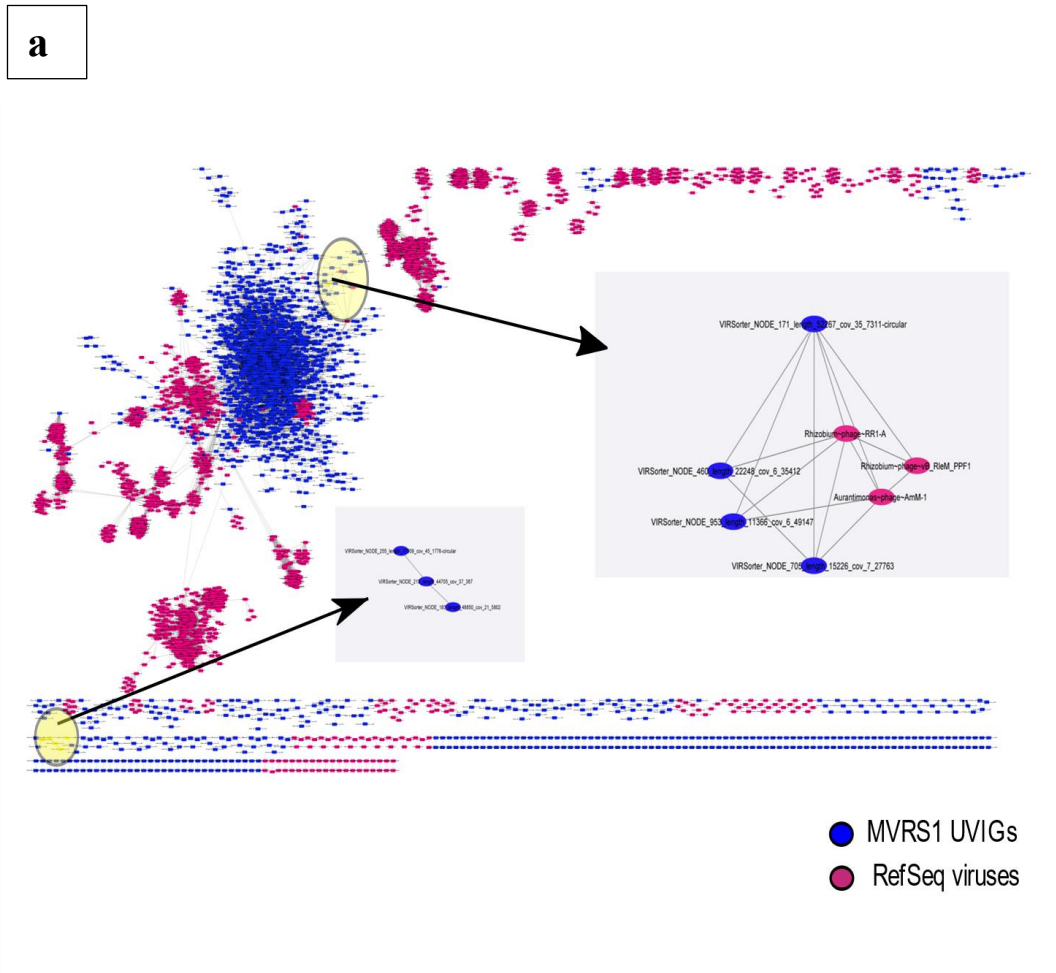


Figure 7. 8: a) Protein clustering network of virus genomes recovered from Rajwadi sediment sample MVRS1 compared to previously described viruses.

A monopartite network analysis was constructed with vContact2 v0.9.19 and Cytoscape v3.9.1 for the taxonomic assignment of MVRS1 UViGs. Nodes represent individual genomes, and edges indicate similarity among genomes within a viral cluster. Blue nodes represent MVRS1 UViGs, and purple nodes represent other viruses from the vConTACT RefSeq dataset.

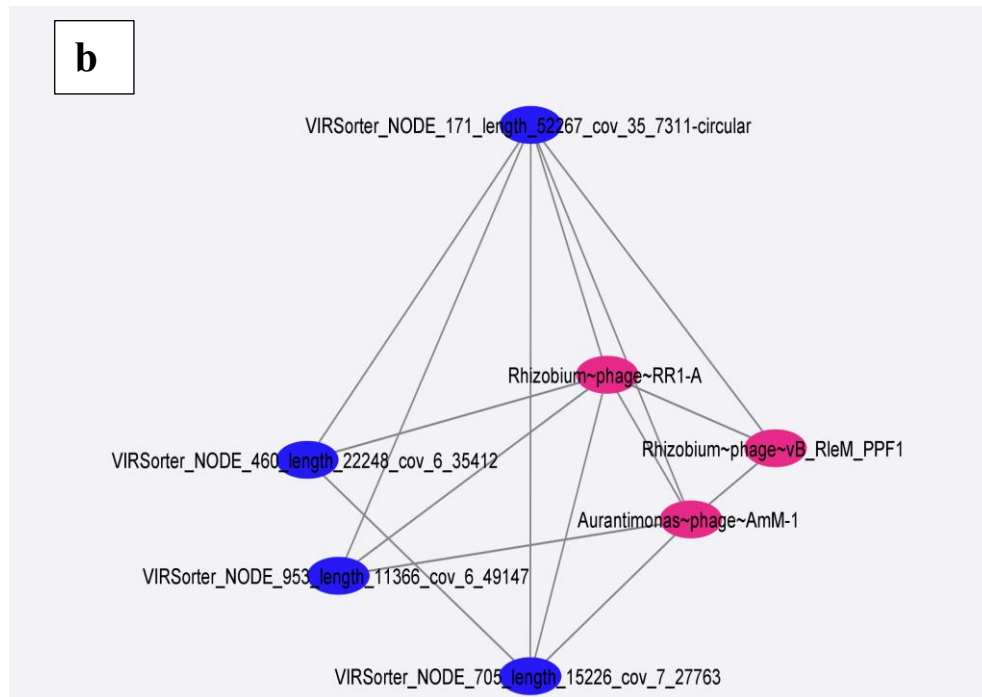


Figure 7.8 b) Viral clusters of MVRs1 the viral contigs clustering within the RefSeq viruses.

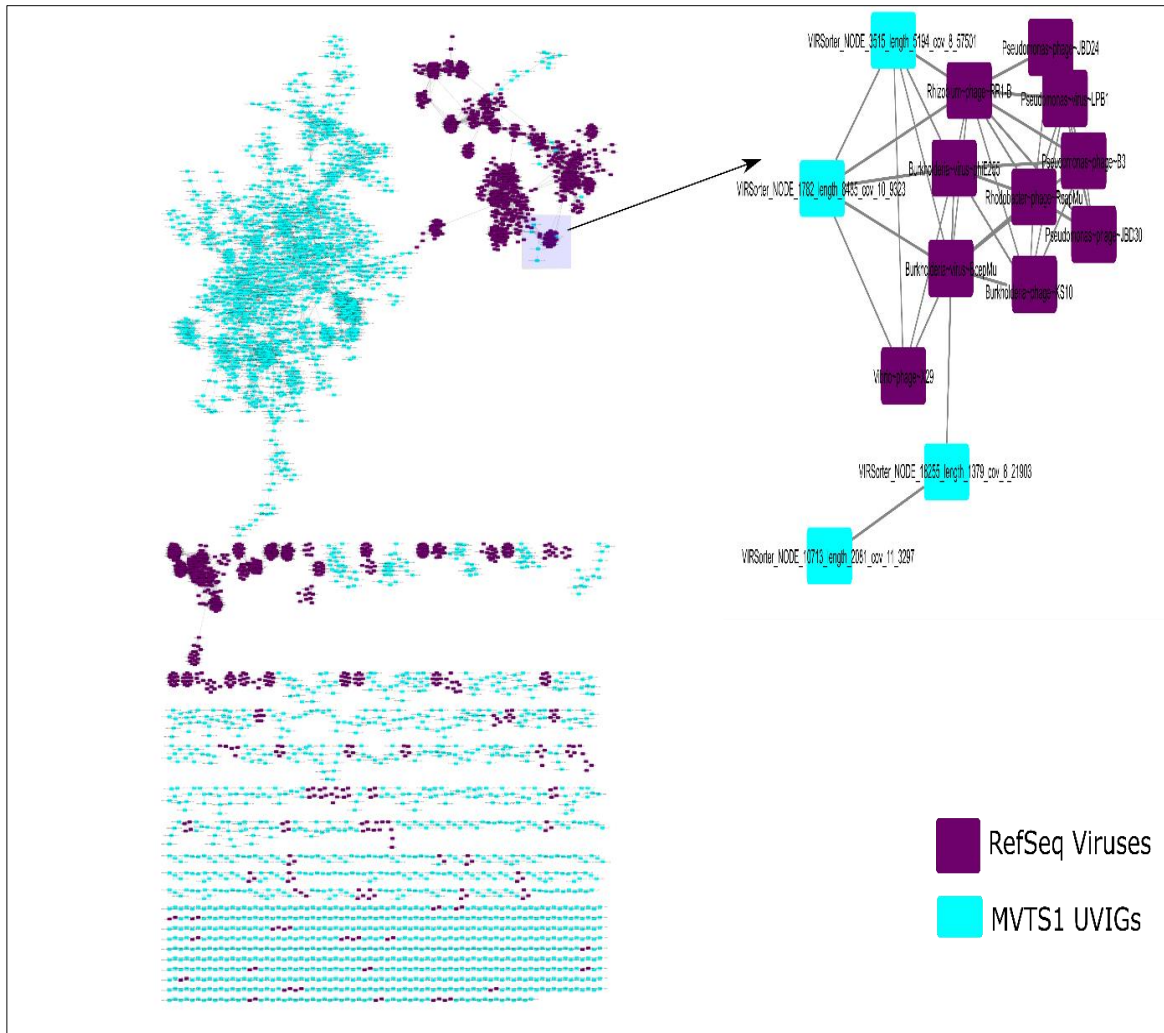


Figure 7. 9: Protein clustering network of virus genomes recovered from Tural sediment sample MVTs1 compared to previously described viruses.

A monopartite network analysis was constructed with vContact2 v0.9.19 and Cytoscape v3.9.1 for the taxonomic assignment of MVTs1 UViGs. Nodes represent individual genomes, and edges indicate similarity among genomes within a viral cluster. Cyan Blue nodes represent MVTs1 UViGs, and purple nodes represent other viruses from the vConTACT RefSeq dataset.

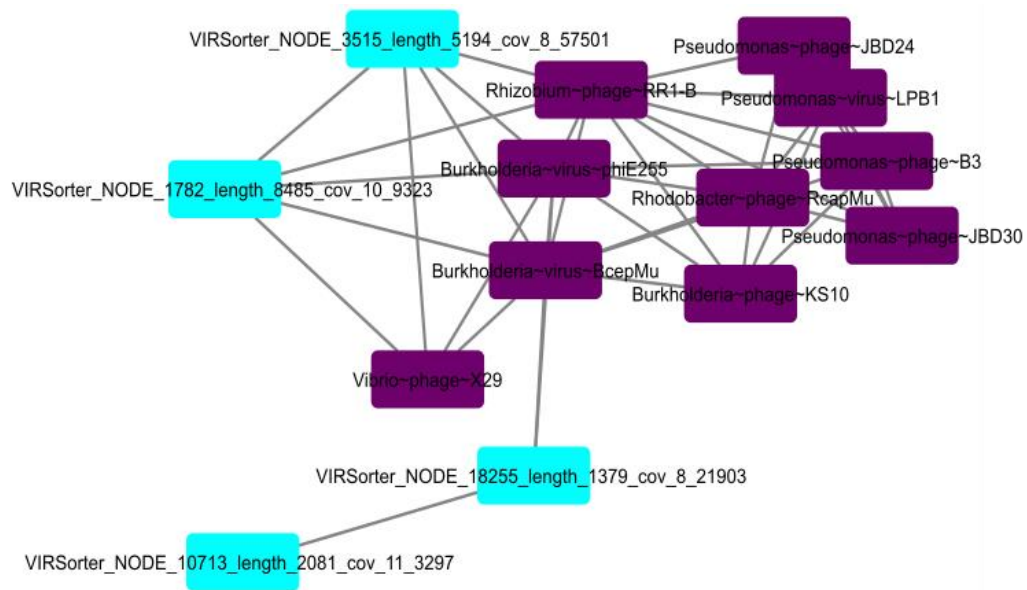


Figure 7.9 b) Viral cluster of MVTs1 showing the clustering within the RefSeq viruses.

7.3.8 VIBRANT prediction of Auxiliary Metabolic Genes (AMG)

A second approach used VIBRANT 1.2.0 programs to retrieve viral contigs. VIBRANT 1.2.0 recovered 510 and 821 viral contigs from Rajwadi and Tural hot spring metavirome assemblies, respectively, and flagged dozens of potential AMGs on viral contigs. VIBRANT prediction of whole viral circular genomes and auxiliary metabolic genes was able to predict 5 contigs and 10 contigs as whole viral circular genomes from MVRS1 and MVTS1 datasets, respectively (Table 7. 7 and Table 7. 8).

Table 7. 7: VIBRANT predicted whole viral circular genomes in MVRS1 (5 contigs).

Viral contigs	Phage nature	Complete Circular viral genomes
NODE_171_length_52267_cov_35.7311	Lysogenic	complete circular
NODE_1059_length_10203_cov_174.356	Lytic	complete circular
NODE_5041_length_1851_cov_38.1102	Lytic	complete circular
NODE_3675_length_2531_cov_20.0856	Lytic	complete circular
NODE_255_length_37909_cov_45.1776	Lytic	complete circular

Table 7. 8: VIBRANT predicted whole viral circular genomes in MVTs1 (10 contigs).

Viral contigs	Phage nature	Complete Circular viral genomes
NODE_6018_length_3336_cov_7.03596	Lytic	complete circular
NODE_9703_length_2249_cov_5.74977	Lytic	complete circular
NODE_10004_length_2197_cov_5.88702	Lytic	complete circular
NODE_2620_length_6459_cov_27.726	Lytic	complete circular
NODE_24405_length_1111_cov_3.82481	Lytic	complete circular
NODE_5761_length_3452_cov_7.50633	Lytic	complete circular
NODE_6097_length_3296_cov_596.379	Lytic	complete circular
NODE_3783_length_4910_cov_7.73615	Lytic	complete circular
NODE_6006_length_3341_cov_8.25289	Lytic	complete circular
NODE_19515_length_1312_cov_30.4885	Lytic	complete circular

7.3.8.1 Auxiliary Metabolic Genes

AMGs in viruses have been attributed to augment fitness and alter/complement the bacterial host's metabolism, enabling their survival under adverse conditions. Various genes related to photosynthesis, carbon fixation, sulphur, and nitrogen biogeochemical cycles have been associated with AMGs. The most abundant AMGs were classified as per the functional categories. They were involved in the metabolism of cofactors and vitamins (MCV), carbohydrates metabolism, amino acids metabolism, metabolism of terpenoids, and polyketides (Figure 7. 10). In the MCV category, the viruses were observed to contribute to the folate biosynthesis pathway (KEGG entries identified: K00287, K00796, K01495, K01737, K09457, K10026, K17745, K06920, K10026). Tetrahydrofolate is a cofactor in all thermophilic bacteria and is essential for the growth; synthesis of formyl methionyl tRNA fMet (White, 1991). Also, photosynthesis-related genes, such as the *cobS* gene among the *psbA* gene, are regarded as core AMGs (Ignacio-Espinoza and Sullivan, 2012). This protein catalyzes the last stage of bacterial cobalamin (vitamin B12) production (Magnúsdóttir et

al., 2015). All these genes were found in our present study suggesting the prevalence of cyanophages. The presence of these AMG genes in our samples suggests that thermophilic cyanophages enhance their fitness by strategically distributing and maintaining specific AMGs in these extreme environments. A previous study on benthic and pelagic viruses from the basaltic sea detected AMGs involved in photosynthesis, 7-cyano-7-deazaguanine synthesis, and cobalamin biosynthesis (Heyerhoff et al., 2022). Many AMGs discovered during the current studies have also been reported in a large-scale examination of metagenomes from various sources, such as marine areas, wastewater, and the human gut. This tends support to the hypothesis of globally conserved core AMGs found across virus genomes regardless of host or environment.

The energy category AMGs were found at low frequency. AMGs involved primarily in photosynthesis, carbon fixation pathways, sulfur metabolism, nitrogen metabolism, and methane metabolism have been widely reported. One of the most important enzymes that we found was **cysH** phosphoadenosine phosphosulfate reductase, that is involved in the reduction in 3-phosphoadenosine-5-phosphosulfate (**PAPS**) to sulfite, an intermediate step of the pathway assimilatory sulfate reduction. These results reveal that viruses contribute to carbohydrate and sulfur metabolism (Figure 7.9 a, and b). Xenobiotics degradation and metabolism category has AMGs related to benzoate, fluorobenzoate, chloro-cyclohexane and chlorobenzene, toluene, nitrotoluene, and caprolactam degradation were also present in both datasets.

Overall, various AMGs suggest that viral infections alter biogeochemical cycles, contribute to host metabolism, and carry out significant ecological tasks in their environments. Viral metagenomics can provide an extensive perspective of viral diversity (Breitbart et al., 2002), and it has uncovered a significant amount of previously unknown diversity and an astonishing richness of viral communities (Edwards and Rohwer, 2005). Viruses are the most abundant entities in the biosphere, with the ability to infect all domains of life, causing disease, influencing biogeochemical cycles, facilitating gene transfer, and influencing the course of microbial evolution. Incorporating AMGs (Breitbart et al., 2018) may modulate cellular activity by expressing enzymes that participate in C, N, P, and S metabolic pathways (Hurwitz and Sullivan, 2013; Castelán-Sánchez et al., 2019). Our results emphasize the need to complement studies of bacteria and archaea communities with the characterization of viral communities to understand microbe-driven elemental cycles in these extreme environments.

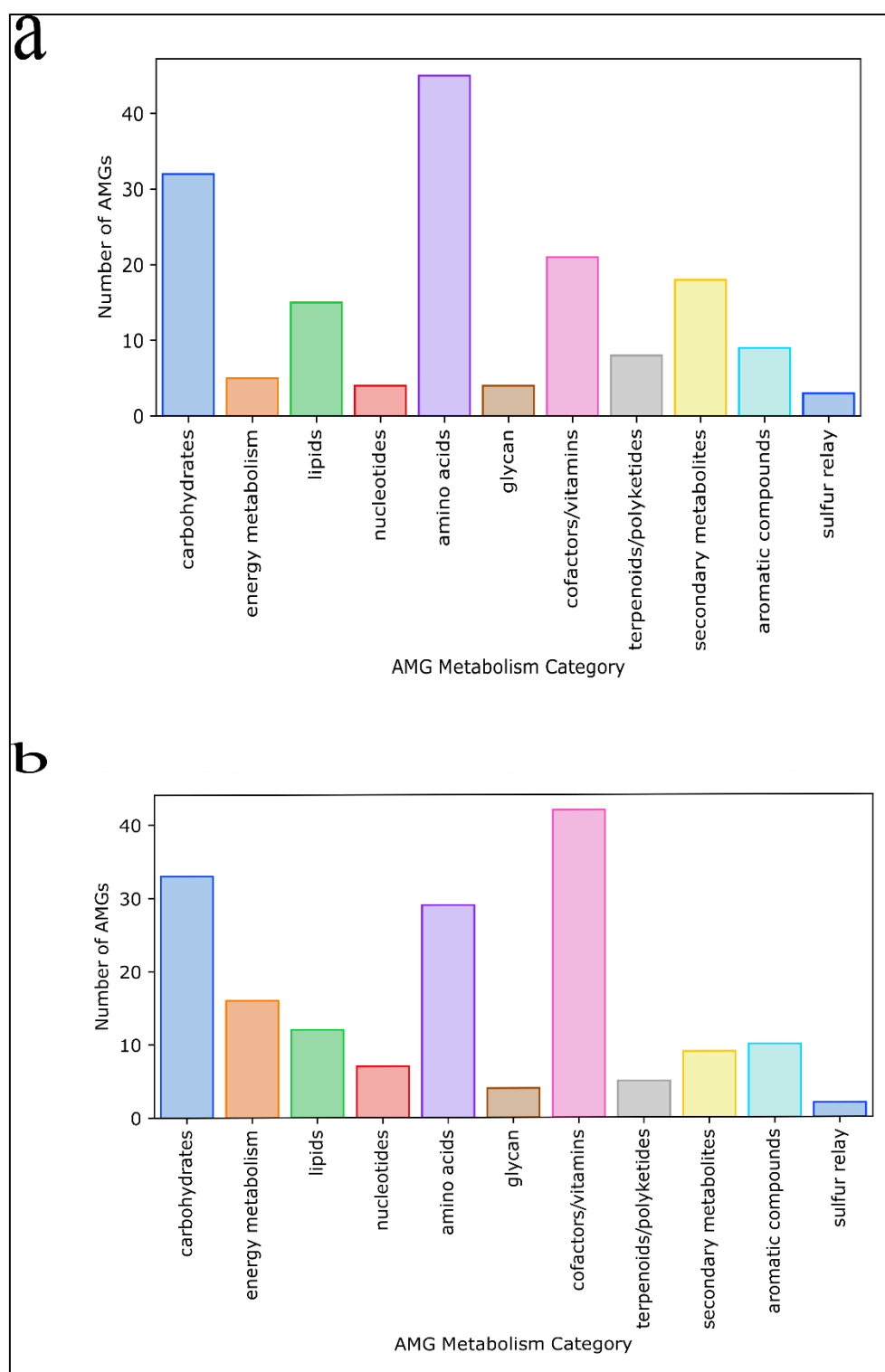


Figure 7. 10: AMG metabolism and Category a) MVRs1 and b) MVTs1.

7.3.9 VIP tree proteomic tree analysis of predicted viral contigs

The viral proteomic tree for whole genome circular contigs from MVRS1 and MVTS1 was visualized using VipTree v. 3.4 (Figure 7. 11 and Figure 7. 12). The VIP tree generated sequence alignment for NODE_ 171, which was complete and circular as predicted by VIRSorter and VIBRANT. The NODE_171 was similar to *Rhizobium* phage RR1-A visualized in the VIPtree web server (Figure 7. 13). This profile was also consistent with the protein network sharing analysis vConTACT 2, where NODE_171 interacts with *Rhizobium* phage RR-1 A and other *Rhizobium* phages (See Figure 7. 8 b).

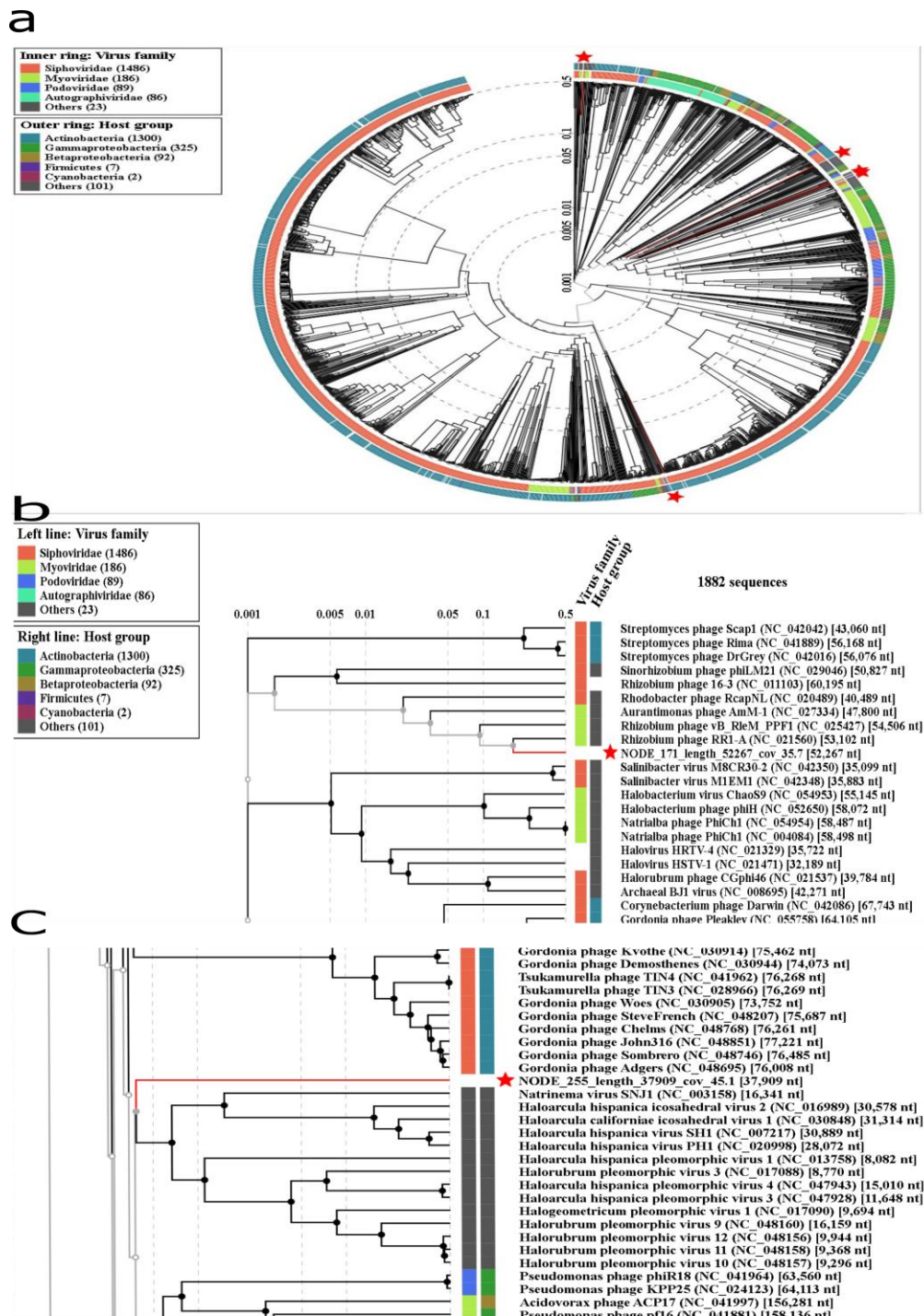


Figure 7. 11: Viral proteomic tree using VipTree v. 1.9.1. for MVRS1 complete circular contigs (5)

a: Circular VIPtree; b and c: expanded rectangular VipTree for two complete circular genome contig (Star) predicted from MVRS1 using VIBRANT and VIRSorter.

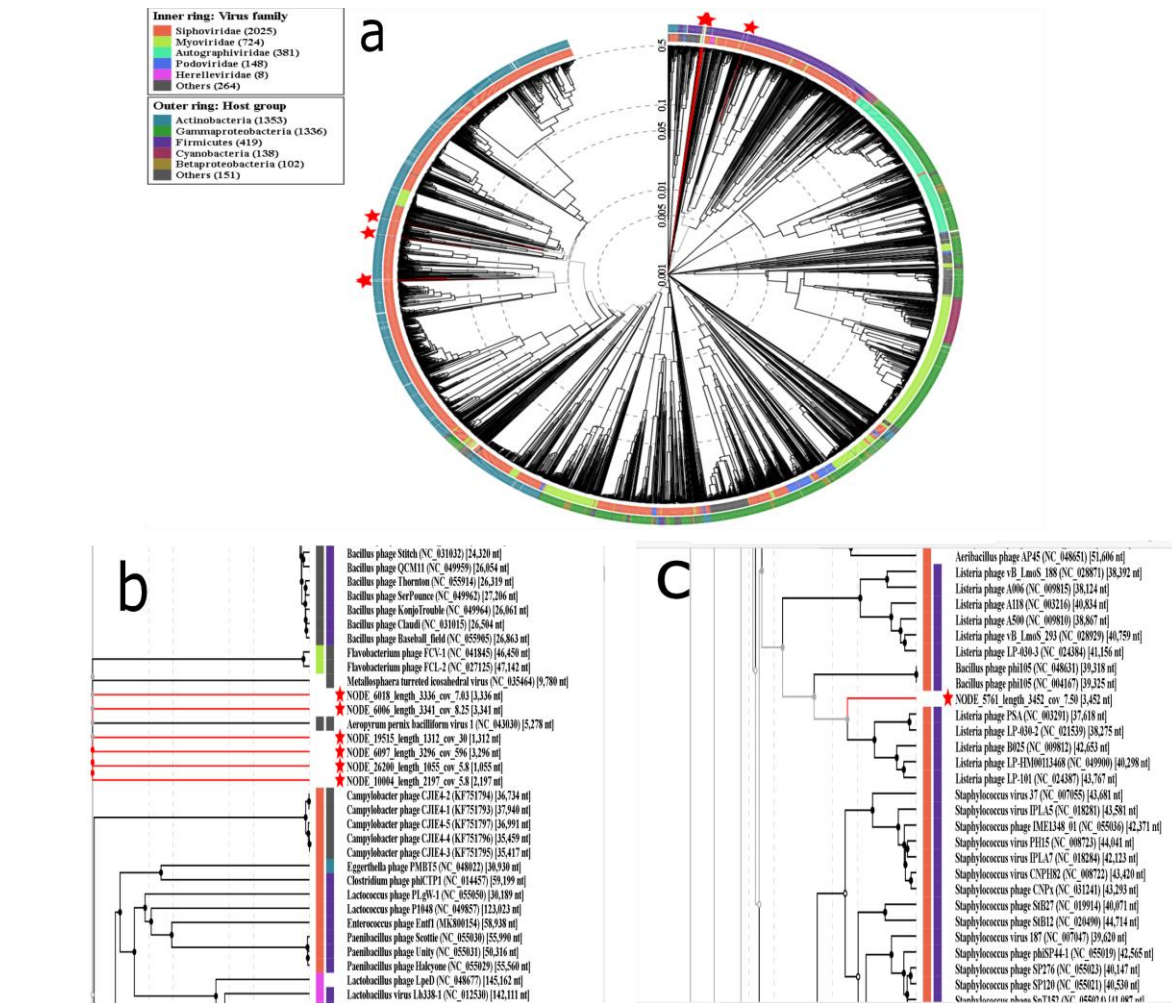


Figure 7. 12: Viral proteomic tree using VipTree v. 1.9.1. for MVTS1 complete circular contigs (10).

a: Circular VIPtree; b and c: expanded rectangular VipTree for ten complete circular genome contig predicted (Star) from MVTS1 using VIBRANT and VIRSorter.

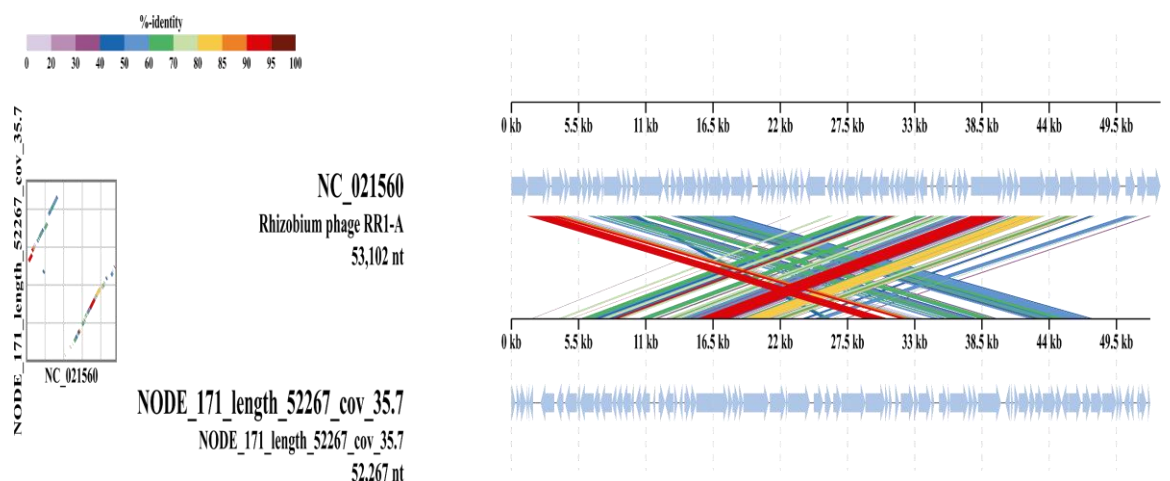


Figure 7. 13: Graphical comparison of the homology MVRS1 complete circular viral contig.

NODE_171_Length 52267 sequences aligned to *Rhizobium* phage RR1-A using VipTree v. 3.4. The schematic representation was conducted with VipTree v3.4 (<https://www.genome.jp/viptree/>), accessed in March 2023)

7.3.10 Quality check and geNomad score prediction

As per the geNomad prediction of viruses and plasmids, the NODE_171 and NODE_255 in the MVRS1 dataset were predicted to be complete by CheckV (Table 7. 9). NODE_171 had an average amino acid identity (AAI)-based completeness of >90% and a geNomad virus score of 0.946, while NODE_255 had an AAI-based completeness of <80% and a geNomad score of 0.8978. The rest of the contigs in both MVRS1 and MVTS1 datasets were of either low or medium quality based on CheckV. The geNomad tool also predicted that these two viral contigs belonged to the taxonomy lineage of tailed viruses class *Caudovirivetes*. Millions of viral genome sequences have been identified from metagenomics data, but the quality and completeness of these sequences vary. Although the vast majority of viral contigs predicted in MVRS1 and MVTS1 were small fragments, and is one of the biggest challenges in the assemblage of viral genomes from short-read genome sequencing. The viral contigs were identified as complete with the DTR (Direct terminal repeats) approach method, the most well-recognized established approached MIUViG standard checklist (Roux et al., 2019).

Table 7. 9: CheckV analysis of two complete circular contigs using geNomad on the NMDC-EDGE platform.

Contig_id	Contig length	Prediction type	Confidence level	Confidence reason
NODE_171_length_52267_cov_35.7311	52267	DTR	High	AAI-based completeness > 90%
NODE_255_length_37909_cov_45.1776	37909	DTR	Low	AAI-based completeness < 80%

7.3.11 Sequence similarity with ReSeq viral genomes and Uncultivated viral genomes

By comparing the UViG NODE_171 viral genome against the nr-NCBI Virus nucleotide database [organism: viruses (tax id: 10239)] (<https://www.ncbi.nlm.nih.gov/labs/virus/vssi/>) using BLASTn program with E-value cutoff 10^{-5} , we found that the viral contig NODE_171 with a genome size of 52.26 kb shared sequence similarity to other known bacteriophages of *Rhizobiaceae* family as host. The closest relative of UViGS NODE_171 was Rhizobium phage RR1-A (Genbank ID: JF974314), infecting the host *Agrobacterium tumefaciens* (*Rhizobium radiobacter*), with 80% nucleotide sequence identity and only 12% query coverage. The UViG NODE_171 was similar to other Rhizobium phage *Sinorhizobium* phage phiLM (GenBank ID: KJ743987), infecting host *Sinorhizobium* sp. LM21, with 79% identity. The *Rhizobium* phage RR1-A of the host *Rhizobium radiobacter*, isolated by prophage induction with mitomycin-C from deep sea subseafloor sediment sample, suggested that an ongoing viral production in 11Ma old sediments (Engelhardt et al., 2013). The abundance of temperate phages in these sediments indicates that there is a continuous viral generation coming from likely natural induction events. Thus, host populations and viral ecosystems in the deep subseafloor appear to be dominated by lysogeny, similar to the hot spring environment (Breitbart et al., 2004). We believe that novel phage development is possible even in slow-growing subterranean microbial communities. Overall, viral studies in sedimentary habitats

are likely the sum of virus preservation, degradation, and ongoing production. Our findings were also in agreement with the Manikarana hot spring in the Himalayan geothermal region where they have predicted the *Rhizobium* phage genome with temperate lifestyle and genome sizes of 20.1kb and 12.6 kb (Sharma et al., 2018). Our study on contrast has predicted a complete circular NODE_171 *Rhizobium* phage with a genome size of 52.2kb and >90% completeness as per the Minimum Information about an Uncultivated Virus Genome (MIUViG) standards (Roux et al., 2019) checklist for uncultivated viral genomes.

Similarly, the UViG NODE_171 was searched against the largest Viral sequence database **IMG/VR v.4** (<https://img.jgi.doe.gov/cgi-bin/vr/main.cgi>) for UViGs using BLASTn program with E-value cutoff of 10^{-5} (Camargo et al., 2023), and the result showed high similarity with the UViGS metagenome vOTUS with (identity 73%-100%) in the global virome database (Table 7. 10). The similarities were with the vOTUs associated with the host of the *Rhizobiaceae* family, which predicted viral lineage to class *Caudovirivetes*-tailed bacteriophages or viruses.

Table 7. 10: Sequence similarity of Predicted Viral contig NODE_171 in NCBI virus database.

Accession	Organism Name	Identity	Species	Family
NC_021560	<i>Rhizobium phage RR1-A</i>	80	<i>Rhizobium</i> phage RR1-A	<i>Myoviridae</i>
NC_025427	<i>Rhizobium phage vB_RleM_PPF1</i>	75	<i>Rhizobium</i> phage vB_RleM_PPF1	<i>Myoviridae</i>
NC_029046	<i>Sinorhizobium phage phiLM21</i>	79	<i>Sinorhizobium</i> phage phiLM21	<i>Siphoviridae</i>
NC_070834	<i>Arthrobacter phage Persistence</i>	70	<i>Arthrobacter</i> phage Persistence	<i>Siphoviridae</i>

7.4 Conclusion

We investigated the viral communities of two WCGP hot springs Rajwadi and Tural from Maharashtra, India, based on the metagenomic data. The present studies on these WCGP hot springs reveal large reservoirs of novel viruses not classified or detected in other virome datasets. These viruses target both eukaryotic and prokaryotic organisms as hosts and modulate their metabolism. Our investigation adds to the unexplored viral diversity studies from the WCGP Indian hot springs. The study highlights the viral diversity and significant roles of viruses in these extreme environments. The slightly alkaline Rajwadi and Tural hot spring water reported 55°C and 60°C, respectively. The transmission electron microscopy of both hot springs revealed the presence of large tailless virus-like particles. The viral metagenomic analysis in the hot springs demonstrated the presence of viruses belonging to the class *Caudoviricetes*. The protein-sharing network analysis of viruses from Rajwadi and Tural hot springs reflects the presence of novel viral clusters unique to particular sites. The two hot spring viromes shared standard metabolic features and encoded auxiliary metabolic genes (AMG) involved in carbohydrate, amino acid, cofactor, and vitamin metabolism. Finally, we predicted an almost complete UVIG (uncultivated viral genome) from the Rajwadi sample, with completeness >90% and a geNomad score of 0.946 and closely related to temperate *Rhizobium* phage RR1-A (80% similarity). Thus, this prediction of UVIG's *Rhizobium* phage suggests that lysogeny is a much more favourable lifestyle for hot spring viruses. A more comprehensive study of virus-host interactions in geothermal hot springs systems will be possible with advancements in sampling, culture-dependent and culture-independent technologies.

Data availability:

The viral metagenomic datasets used in this study are publicly available in the National Center for Biotechnology Information (NCBI) sequence read archive (SRA) (<https://www.ncbi.nlm.nih.gov/sra>) repository under the BioProject PRJNA906748.

SUMMARY AND CONCLUSION

The present study contributes to research in the field of aquatic virology, focusing on hot spring viruses from Indian hot springs that have not previously been explored from a virological perspective. There are two parts to the study:

A) Isolation of virus-host systems from Indian hot springs and to characterize the isolated bacteriophage.

In the first part of the study, a total of 144 bacterial strains were obtained from 7 different Indian hot springs, which were cultured and maintained in the laboratory until further bacteriophage screening was carried out. A total of ten bacteria were selected as potential hosts for bacteriophage infection and they were identified.

Among the ten bacterial cultures identified, two bacteria, namely, bacterial strain **VB3** and **JE2** served as potential hosts for isolating bacteriophages. Phylogenetic studies based on 16Sr RNA sequence analysis indicated that bacterial strains VB3 and JE2 formed coherent clusters with species of *Pseudomonas* and *Meiothermus*, respectively.

Isolation of lytic phage infecting *Meiothermus* sp. JE2 and a temperate phage infecting *Pseudomonas* sp. strain VB3 were obtained from Rajwadi (Maharashtra) and Vashisht (Himachal Pradesh) hot springs, respectively. The lytic bacteriophage vB_Msp_JE2 infecting *Meiothermus* sp. strain JE2 was isolated and partially characterized from the Rajwadi hot spring of the West Coast geothermal province, producing a clear plaque after 48 h at 60°C. It was stable at pH 6.0-8.0 and thermostable at 60°C and completely inactivated above 75°C. We were unable to further characterize due to loss in phage viability during long-term storage.

A lysogenic bacteria *Pseudomonas* strain VB3 was isolated from Vashisht hot spring (Himachal Pradesh), belonging to the Himalayan geothermal province. After the prophage induction assay, the presence of prophage was indicated by the reduction in the growth rate of the bacterium upon mitomycin-C treatment, which was also concurrent with the detection of the clearance zone in the spot assay.

Draft genome sequences of a lysogenic bacteria *Pseudomonas* sp. VB3's comprise approximately 6.06 MB DNA and 5607 predicted genes with a G+C content of 60.56%.

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 of *Pseudomonas* sp. VB3.

The temperate bacteriophage after prophage induction with mitomycin C was isolated and characterized. The inducible bacteriophage has a narrow host specificity, and TEM imaging revealed that it is a tailed phage (Class: *Caudoviricetes*). The phage proteins identified during phage genome assembly and annotation were integrase, tail, terminase, plate, capsid, head, and portal proteins. Sequence analysis of phage vB_Psp_VB3 showed that the phage genome was approximately 48.6 kb in length with GC content of 56.83% containing putative 43 ORF regions.

Phage vB_Psp_VB3 had the highest sequence similarity to ***Pseudomonas* phage phiPSA1** and ***Pseudomonas* phage phiAH14a** with a sequence identity of 89% and 92%, respectively.

Phage vB_Psp_VB3 is the inducible temperate phage of *Pseudomonas* sp. strain VB3 isolated from Vashisht hot spring, Himachal Pradesh, India.

B). Metavirome analysis of two hot springs belonging to the West Coast geothermal province, Maharashtra, India.

In the second part of the study, metavirome analysis of Rajwadi and Tural hot springs belonging to the West Coast geothermal province (WCGP) of Maharashtra, India.

The viral metagenomic analysis in the hot spring sediments demonstrated the presence of viruses belonging to the class *Caudoviricetes*.

The protein-sharing network analysis of viruses from Rajwadi (MVRS1) and Tural (MVTS1) hot springs reflects the presence of novel viral clusters unique to particular sites.

The two hot spring viromes shared standard metabolic features and encoded auxiliary metabolic genes (AMG) that are involved in carbohydrate, amino acid, cofactor, and vitamin metabolism as well as hydrocarbon degradation.

An almost complete UVIG (uncultivated viral genome) was predicted from the Rajwadi sample (MVRS1), with completeness >90% and a geNomad score of 0.946, and closely related to temperate *Rhizobium* phage RR1-A, which suggests that lysogeny is a much more favourable lifestyle for hot spring viruses.

CONCLUSION:

Even though India has numerous hot springs, they have received little attention in terms of virus research. In this context, we isolated and partially characterized lytic phage infecting *Meiothermus* sp. JE2 and a temperate phage infecting *Pseudomonas* sp. strain VB3 were obtained from Rajwadi (Maharashtra) and Vashisht (Himachal Pradesh) hot springs, respectively. The lytic *Meiothermus* phage vB_Msp_JE2 was partially characterized, but the subsequent studies were hindered by phage viability loss during long-term storage. The WGS of the lysogenic bacteria *Pseudomonas* sp. VB3 contained putative prophage-related sequence regions, according to PHASTER analysis. Further, genomic characterization of the mitomycin-C-inducible phage of *Pseudomonas* sp. strain VB3 revealed the presence of genes required for a temperate lifestyle to thrive in hot springs.

Additionally, to expand our knowledge and overcome the challenges of conventional isolation techniques, we assessed the viral metagenomic analysis of two hot springs of WCGP which demonstrated the presence of viruses belonging to the class *Caudoviricetes*. The recovered uncultivated viral contigs contained several AMGs belonging to metabolic pathways especially involved in carbohydrate, amino acid, cofactor, and vitamin metabolism as well as hydrocarbon degradation. We predicted an almost complete UVIG (uncultivated viral genome) from the Rajwadi sample and closely related to temperate *Rhizobium* phage RR1-A, suggesting that lysogeny is a much more favourable lifestyle for hot spring viruses. Altogether, major UViGs remain taxonomically unassigned, indicating the novel nature of the viral community present in these hot springs, necessitating such virome explorations to enrich existing viral databases.

Overall, in these extreme environments, phages enter into a lysogeny state, facilitating long-term co-evolution with the bacterial host, and thereby establishing a mutually beneficial interaction for their survival. Both aspects of the investigations point to lysogeny as the most favourable way of lifestyle in hot springs. Even though metaviromics is a significant tool for the study of viral ecology, this method needs to be coupled with phage-host identification studies to unveil the "**viral dark matter**," which is an ongoing worldwide endeavour to explore the virosphere in different ecosystems.

BIBLIOGRAPHY

- Ackermann, H. W. (2007). 5500 Phages examined in the electron microscope. *Arch Virol*, 152(2), 227–243. <https://doi.org/10.1007/s00705-006-0849-1>
- Ackermann, H.-W. (2009). Phage Classification and Characterization. In M. R. J. Clokie & A. M. Kropinski (Eds.), *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions* (pp. 127–140). Humana Press. https://doi.org/10.1007/978-1-60327-164-6_13
- Ács, N., Gambino, M., & Brøndsted, L. (2020). Bacteriophage Enumeration and Detection Methods. *Frontiers in Microbiology*, 11. <https://doi.org/10.3389/fmicb.2020.594868>
- Adams, M. H., Anderson, E. S., Gots, J. S., Jacob, F., & Wollman, E. L. (1959). Bacteriophages. In *Bacteriophages*.
- Aditiawati, P., Yohandini, H., Madayanti, F., & Akhmaloka. (2009). Microbial diversity of acidic hot spring (Kawah Hujan B) in geothermal field of Kamojang area, west Java-Indonesia. *The Open Microbiology Journal*, 3, 58–66. <https://doi.org/10.2174/1874285800903010058>
- Almpanis, A., Swain, M., Gatherer, D., & McEwan, N. (2018). Correlation between bacterial G+C content, genome size and the G+C content of associated plasmids and bacteriophages. *Microbial Genomics*, 4(4). <https://doi.org/10.1099/mgen.0.000168>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Anantharaman, K., Duhaime, M. B., Breier, J. A., Wendt, K. A., Toner, B. M., & Dick, G. J. (2014). Sulfur oxidation genes in diverse deep-sea viruses. *Science (New York, N.Y.)*, 344(6185), 757–760. <https://doi.org/10.1126/science.1252229>
- Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence*. <https://doi.org/http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Angly, F. E., Felts, B., Breitbart, M., Salamon, P., Edwards, R. A., Carlson, C., Chan, A. M., Haynes, M., Kelley, S., Liu, H., Mahaffy, J. M., Mueller, J. E., Nulton, J.,

- Olson, R., Parsons, R., Rayhawk, S., Suttle, C. A., & Rohwer, F. (2006). The marine viromes of four oceanic regions. *PLoS Biology*, 4(11), e368. <https://doi.org/10.1371/journal.pbio.0040368>
- Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., & Wishart, D. S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, 44(W1), W16-21. <https://doi.org/10.1093/nar/gkw387>
- Arya, M., Joshi, G. K., Gupta, A. K., Kumar, A., & Raturi, A. (2015). Isolation and characterization of thermophilic bacterial strains from Soldhar (Tapovan) hot spring in Central Himalayan Region, India. *Annals of Microbiology*, 65(3), 1457–1464. <https://doi.org/10.1007/s13213-014-0984-y>
- 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., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., ... Zagnitko, O. (2008). The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, 9(1), 75. <https://doi.org/10.1186/1471-2164-9-75>
- Bae, T., Baba, T., Hiramatsu, K., & Schneewind, O. (2006). Prophages of *Staphylococcus aureus* Newman and their contribution to virulence. *Molecular Microbiology*, 62(4), 1035–1047.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, 19(5), 455–477. <https://doi.org/10.1089/cmb.2012.0021>
- Banks, D. J., Lei, B., & Musser, J. M. (2003). Prophage induction and expression of prophage-encoded virulence factors in group A *Streptococcus* serotype M3 strain MGAS315. *Infection and Immunity*, 71(12), 7079–7086. <https://doi.org/10.1128/IAI.71.12.7079-7086.2003>
- Bar-On, Y. M., Phillips, R., & Milo, R. (2018). The biomass distribution on Earth. *Proceedings of the National Academy of Sciences of the United States of America*, 115(25), 6506–6511. <https://doi.org/10.1073/pnas.1711842115>
- Bench, S. R., Hanson, T. E., Williamson, K. E., Ghosh, D., Radosovich, M., Wang, K., &

- Wommack, K. E. (2007). Metagenomic characterization of Chesapeake Bay viroplankton. *Appl Environ Microbiol*, 73(23), 7629–7641.
<https://doi.org/10.1128/AEM.00938-07>
- Bergey, D. H. (1994). *Bergey's manual of determinative bacteriology*. Lippincott Williams & Wilkins.
- Bergh, O., Børsheim, K. Y., Bratbak, G., & Heldal, M. (1989). High abundance of viruses found in aquatic environments. *Nature*, 340. <https://doi.org/10.1038/340467a0>
- Biagini, P., Thèves, C., Balaesque, P., Géraut, A., Cannet, C., Keyser, C., Nikolaeva, D., Gérard, P., Duchesne, S., Orlando, L., Willerslev, E., Alekseev, A. N., de Micco, P., Ludes, B., & Crubézy, E. (2012). Variola virus in a 300-year-old Siberian mummy. *The New England journal of medicine*, 367(21), 2057–2059.
<https://doi.org/10.1056/NEJMc1208124>
- Bin Jang, H., Bolduc, B., Zablocki, O., Kuhn, J. H., Roux, S., Adriaenssens, E. M., Brister, J. R., Kropinski, A. M., Krupovic, M., Lavigne, R., Turner, D., & Sullivan, M. B. (2019). Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nature Biotechnology*, 37(6).
<https://doi.org/10.1038/s41587-019-0100-8>
- Bize, A., Peng, X., Prokofeva, M., MacLellan, K., Lucas, S., Forterre, P., Garrett, R. A., Bonch-Osmolovskaya, E. A., & Prangishvili, D. (2008). Viruses in acidic geothermal environments of the Kamchatka Peninsula. *Res Microbiol*, 159(5), 358–366. <https://doi.org/10.1016/j.resmic.2008.04.009>
- Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., Van Wezel, G. P., Medema, M. H., & Weber, T. (2021). antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Research*, 49(W1), W29–W35.
<https://doi.org/10.1093/nar/gkab335>
- Bolduc, B. (2021). *Processing a Viral Metagenome Using iVirus V.3*.
<https://doi.org/dx.doi.org/10.17504/protocols.io.buisnuee>
- Bolduc, B., Jang, H. Bin, Doulier, G., You, Z. Q., Roux, S., & Sullivan, M. B. (2017). vConTACT: An iVirus tool to classify double-stranded DNA viruses that infect Archaea and Bacteria. *PeerJ*, 2017(5). <https://doi.org/10.7717/peerj.3243>
- Bolduc, B., Wirth, J. F., Mazurie, A., & Young, M. J. (2015). Viral assemblage composition in Yellowstone acidic hot springs assessed by network analysis. *ISME*

- Journal*, 9(10), 2162–2177. <https://doi.org/10.1038/ismej.2015.28>
- Bolduc, B., Youens-Clark, K., Roux, S., Hurwitz, B. L., & Sullivan, M. B. (2017). iVirus: facilitating new insights in viral ecology with software and community data sets imbedded in a cyberinfrastructure. *The ISME journal*, 11(1), 7–14. <https://doi.org/10.1038/ismej.2016.89>
- Bolduc, B., Zablocki, O., Guo, J., Zayed, A. A., Vik, D., Dehal, P., Wood-Charlson, E. M., Arkin, A., Merchant, N., Pett-Ridge, J., Roux, S., Vaughn, M., & Sullivan, M. B. (2021). iVirus 2.0: Cyberinfrastructure-supported tools and data to power DNA virus ecology. *ISME Communications*, 1(1). <https://doi.org/10.1038/s43705-021-00083-3>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bowen De León, K., Gerlach, R., Peyton, B. M., & Fields, M. W. (2013). Archaeal and bacterial communities in three alkaline hot springs in Heart Lake Geyser Basin, Yellowstone National Park. *Frontiers in Microbiology*, 4, 330. <https://doi.org/10.3389/fmicb.2013.00330>
- Breitbart, M., Bonnain, C., Malki, K., & Sawaya, N. A. (2018). Phage puppet masters of the marine microbial realm. *Nature Microbiology*, 3(7), 754–766. <https://doi.org/10.1038/s41564-018-0166-y>
- Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J. M., Segall, A. M., Mead, D., Azam, F., & Rohwer, F. (2002). Genomic analysis of uncultured marine viral communities. *Proceedings of the National Academy of Sciences of the United States of America*, 99(22), 14250–14255. <https://doi.org/10.1073/pnas.202488399>
- Breitbart, M., Wegley, L., Leeds, S., Schoenfeld, T., & Rohwer, F. (2004). Phage Community Dynamics in Hot Springs. *Applied and Environmental Microbiology*, 70(3), 1633–1640. <https://doi.org/10.1128/aem.70.3.1633-1640.2004>
- Brussaard, C. P. D. (2004). Optimization of procedures for counting viruses by flow cytometry. *Applied and Environmental Microbiology*, 70(3), 1506–1513. <https://doi.org/10.1128/AEM.70.3.1506-1513.2004>
- Brussaard, C. P., Marie, D., & Bratbak, G. (2000). Flow cytometric detection of viruses. *Journal of Virological Methods*, 85(1–2), 175–182. <https://doi.org/10.1016/s0166->

- Buchholz, H. H., Michelsen, M. L., Bolaños, L. M., Browne, E., Allen, M. J., & Temperton, B. (2021). Efficient dilution-to-extinction isolation of novel virus-host model systems for fastidious heterotrophic bacteria. *The ISME Journal*, 15(6), 1585–1598. <https://doi.org/10.1038/s41396-020-00872-z>
- Camargo, A. P., Nayfach, S., Chen, I. M. A., Palaniappan, K., Ratner, A., Chu, K., Ritter, S. J., Reddy, T. B. K., Mukherjee, S., Schulz, F., Call, L., Neches, R. Y., Woyke, T., Ivanova, N. N., Elie-Fadrosh, E. A., Kyrpides, N. C., & Roux, S. (2023). IMG/VR v4: an expanded database of uncultivated virus genomes within a framework of extensive functional, taxonomic, and ecological metadata. *Nucleic Acids Research*, 51(D1), D733–D743. <https://doi.org/10.1093/nar/gkac1037>
- Canchaya, C., Proux, C., Fournous, G., Bruttin, A., & Brüssow, H. (2003). Prophage genomics. *Microbiology and Molecular Biology Reviews : MMBR*, 67(2), 238–276, table of contents. <https://doi.org/10.1128/MMBR.67.2.238-276.2003>
- Casjens, S. (2003). Prophages and bacterial genomics: What have we learned so far? *Molecular Microbiology*, 49(2), 277–300. <https://doi.org/10.1046/j.1365-2958.2003.03580.x>
- Castelán-Sánchez, H. G., López-Rosas, I., García-Suastegui, W. A., Peralta, R., Dobson, A. D. W., Batista-García, R. A., & Dávila-Ramos, S. (2019). Extremophile deep-sea viral communities from hydrothermal vents: Structural and functional analysis. *Marine Genomics*, 46, 16–28. <https://doi.org/10.1016/j.margen.2019.03.001>
- Castelán-Sánchez, H. G., Meza-Rodríguez, P. M., Carrillo, E., Ríos-Vázquez, D. I., Liñan-Torres, A., Batista-García, R. A., ... & Dávila-Ramos, S. (2020). The microbial composition in circumneutral thermal springs from Chignahuapan, Puebla, Mexico reveals the presence of particular sulfur-oxidizing bacterial and viral communities. *Microorganisms*, 8(11), 1677. <https://doi.org/10.3390/microorganisms8111677>
- Ceyssens, P.-J., & Lavigne, R. (2010). Bacteriophages of *Pseudomonas*. *Future Microbiology*, 5(7), 1041–1055. <https://doi.org/10.2217/fmb.10.66>
- Chandrasekhar, T., Chandrasekhar, V., & Chandrasekharam, D. (2016). Geothermometry of west coast geothermal province, maharashtra, India. *Transactions - Geothermal Resources Council*, 40, 495–499.

- Chandrasekhar, V., & Chandrasekharam, D. (2010). Energy Independence Through CDM Using Geothermal Resources : Indian Scenario. *Earth, April*, 25–29.
- Chandrasekharam, D. (2000). Geothermal energy resources of India. *IBC Conference “Geothermal Power Asia 2000” Manila, Philippines.*
- Chatterjee, S., Sharma, S., Ansari, Md. A., Deodhar, A. S., Low, U., Sinha, U. K., & Dash, A. (2016). Characterization of subsurface processes estimation of reservoir temperature in Tural Rajwadi geothermal fields, Maharashtra, India. *Geothermics*, 59, 77–89. <https://doi.org/10.1016/j.geothermics.2015.10.011>
- Cheepudom, J., Lee, C.-C., Cai, B., & Meng, M. (2015). Isolation, characterization, and complete genome analysis of P1312, a thermostable bacteriophage that infects *Thermobifida fusca*. *Frontiers in Microbiology*, 6, 959. <https://doi.org/10.3389/fmicb.2015.00959>
- Chen, F., Wang, K., Stewart, J., & Belas, R. (2006). Induction of multiple prophages from a marine bacterium: a genomic approach. *Appl Environ Microbiol*, 72(7), 4995–5001. <https://doi.org/10.1128/AEM.00056-06>
- Chen, L.-X., Méheust, R., Crits-Christoph, A., McMahon, K. D., Nelson, T. C., Slater, G. F., Warren, L. A., & Banfield, J. F. (2020). Large freshwater phages with the potential to augment aerobic methane oxidation. *Nature Microbiology*, 5(12), 1504–1515.
- Chen, L.-X., Zhao, Y., McMahon, K. D., Mori, J. F., Jessen, G. L., Nelson, T. C., Warren, L. A., & Banfield, J. F. (2019). Wide Distribution of Phage That Infect Freshwater SAR11 Bacteria. *MSystems*, 4(5). <https://doi.org/10.1128/mSystems.00410-19>
- Chopin, A., Bolotin, A., Sorokin, A., Ehrlich, S. D., & Chopin, M. (2001). Analysis of six prophages in *Lactococcus lactis* IL1403: different genetic structure of temperate and virulent phage populations. *Nucleic Acids Research*, 29(3), 644–651. <https://doi.org/10.1093/nar/29.3.644>
- Clokier, M. R. J., Millard, A. D., Letarov, A. V., & Heaphy, S. (2011). Phages in nature.pdf. *Bacteriophage*, 1(1), 31–45. <https://doi.org/10.4161/bact.1.1.14942>
- Collins, C. H., & Lyne, P. M. (1970). Microbiological methods. *Microbiological Methods.*, 3rd. Edition.
- Colombet, J., Robin, A., Lavie, L., Bettarel, Y., Cauchie, H. M., & Sime-Ngando, T. (2007). Virioplankton “pegylation”: Use of PEG (polyethylene glycol) to

- concentrate and purify viruses in pelagic ecosystems. *Journal of Microbiological Methods*, 71(3), 212–219. <https://doi.org/10.1016/j.mimet.2007.08.012>
- Conceição-Neto, N., Zeller, M., Lefrère, H., De Bruyn, P., Beller, L., Deboutte, W., Yinda, C. K., Lavigne, R., Maes, P., Van Ranst, M., Heylen, E., & Matthijnssens, J. (2015). Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Scientific Reports*, 5, 16532. <https://doi.org/10.1038/srep16532>
- Cornuault, J. K., & Moineau, S. (2021). Induction and Elimination of Prophages Using CRISPR Interference. *The CRISPR Journal*, 4(4), 549–557. <https://doi.org/10.1089/crispr.2021.0026>
- Couvin, D., Bernheim, A., Toffano-Nioche, C., Touchon, M., Michalik, J., Néron, B., Rocha, E. P. C., Vergnaud, G., Gautheret, D., & Pourcel, C. (2018). CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Research*, 46(W1), W246–W251. <https://doi.org/10.1093/nar/gky425>
- Cunningham, B. R., Brum, J. R., Schwenck, S. M., Sullivan, M. B., & John, S. G. (2015). An inexpensive, accurate, and precise wet-mount method for enumerating aquatic viruses. *Applied and Environmental Microbiology*, 81(9), 2995–3000. <https://doi.org/10.1128/AEM.03642-14>
- Das, S., Kumari, A., Sherpa, M. T., Najar, I. N., & Thakur, N. (2020). Metavirome and its functional diversity analysis through microbiome study of the Sikkim Himalayan hot spring solfataric mud sediments. *Current Research in Microbial Sciences*, 1, 18–29. <https://doi.org/10.1016/j.crmicr.2020.05.002>
- Dávila-Ramos, S., Castelán-Sánchez, H. G., Martínez-Ávila, L., Sánchez-Carbente, M. D. R., Peralta, R., Hernández-Mendoza, A., Dobson, A. D. W., Gonzalez, R. A., Pastor, N., & Batista-García, R. A. (2019). A Review on Viral Metagenomics in Extreme Environments. *Frontiers in microbiology*, 10, 2403. <https://doi.org/10.3389/fmicb.2019.02403>
- Dedrick, R. M., Jacobs-Sera, D., Bustamante, C. A. G., Garlena, R. A., Mavrich, T. N., Pope, W. H., Reyes, J. C. C., Russell, D. A., Adair, T., Alvey, R., Bonilla, J. A., Bricker, J. S., Brown, B. R., Byrnes, D., Cresawn, S. G., Davis, W. B., Dickson, L. A., Edgington, N. P., Findley, A. M., ... Hatfull, G. F. (2017). Prophage-mediated

- defence against viral attack and viral counter-defence. *Nature Microbiology*, 2, 16251. <https://doi.org/10.1038/nmicrobiol.2016.251>
- Demina, T. A., Pietilä, M. K., Svirskaitė, J., Ravantti, J. J., Atanasova, N. S., Bamford, D. H., & Oksanen, H. M. (2017). HCIV-1 and other tailless icosahedral internal membrane-containing viruses of the family Sphaerolipoviridae. *Viruses*, 9(2). <https://doi.org/10.3390/v9020032>
- Dion, M. B., Oechslin, F., & Moineau, S. (2020). Phage diversity, genomics and phylogeny. *Nature Reviews. Microbiology*, 18(3), 125–138. <https://doi.org/10.1038/s41579-019-0311-5>
- Djikeng, A., Kuzmickas, R., Anderson, N. G., & Spiro, D. J. (2009). Metagenomic analysis of RNA viruses in a fresh water lake. *PloS One*, 4(9), e7264. <https://doi.org/10.1371/journal.pone.0007264>
- Duggan, A. T., Perdomo, M. F., Piombino-Mascali, D., Marciniak, S., Poinar, D., Emery, M. V, Buchmann, J. P., Duchêne, S., Jankauskas, R., Humphreys, M., Golding, G. B., Southon, J., Devault, A., Rouillard, J.-M., Sahl, J. W., Dutour, O., Hedman, K., Sajantila, A., Smith, G. L., ... Poinar, H. N. (2016). 17(th) Century Variola Virus Reveals the Recent History of Smallpox. *Current Biology : CB*, 26(24), 3407–3412. <https://doi.org/10.1016/j.cub.2016.10.061>
- Dutilh, B. E., Cassman, N., McNair, K., Sanchez, S. E., Silva, G. G. Z., Boling, L., Barr, J. J., Speth, D. R., Seguritan, V., Aziz, R. K., Felts, B., Dinsdale, E. A., Mokili, J. L., & Edwards, R. A. (2014). A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nature Communications*, 5, 1–11. <https://doi.org/10.1038/ncomms5498>
- Dziewit, L., & Radlinska, M. (2016). Two Inducible Prophages of an Antarctic *Pseudomonas* sp. ANT_H14 Use the Same Capsid for Packaging Their Genomes - Characterization of a Novel Phage Helper-Satellite System. *PloS One*, 11(7), e0158889. <https://doi.org/10.1371/journal.pone.0158889>
- Edwards, R. A., & Rohwer, F. (2005). Viral metagenomics. *Nature Reviews. Microbiology*, 3(6), 504–510. <https://doi.org/10.1038/nrmicro1163>
- Eggers, C. H., Gray, C. M., Preisig, A. M., Glenn, D. M., Pereira, J., Ayers, R. W., Alshahrani, M., Acabbo, C., Becker, M. R., Bruenn, K. N., Cheung, T., Jendras, T. M., Shepley, A. B., & Moeller, J. T. (2016). Phage-mediated horizontal gene

- transfer of both prophage and heterologous DNA by ϕ BB-1, a bacteriophage of *Borrelia burgdorferi*. *Pathogens and Disease*, 74(9).
<https://doi.org/10.1093/femspd/ftw107>
- Ekman, J., Kosonen, M., Jokela, S., Kolari, M., Korhonen, P., & Salkinoja-Salonen, M. (2007). Detection and quantitation of colored deposit-forming *Meiothermus* spp. in paper industry processes and end products. *J Ind Microbiol Biotechnol*, 34, 203–211. <https://doi.org/10.1007/s10295-006-0187-z>
- Emerson, J. B., Roux, S., Brum, J. R., Bolduc, B., Woodcroft, B. J., Jang, H. B., Singleton, C. M., Solden, L. M., Naas, A. E., Boyd, J. A., Hodgkins, S. B., Wilson, R. M., Trubl, G., Li, C., Frolking, S., Pope, P. B., Wrighton, K. C., Crill, P. M., Chanton, J. P., Saleska, S. R., ... Sullivan, M. B. (2018). Host-linked soil viral ecology along a permafrost thaw gradient. *Nature microbiology*, 3(8), 870–880.
<https://doi.org/10.1038/s41564-018-0190-y>
- Enault, F., Briet, A., Bouteille, L., Roux, S., Sullivan, M. B., & Petit, M.-A. (2017). Phages rarely encode antibiotic resistance genes: a cautionary tale for virome analyses. *The ISME Journal*, 11(1), 237–247. <https://doi.org/10.1038/ismej.2016.90>
- Engelhardt, T., Sahlberg, M., Cypionka, H., & Engelen, B. (2013). Biogeography of *Rhizobium radiobacter* and distribution of associated temperate phages in deep subseafloor sediments. *ISME Journal*, 7(1), 199–208.
<https://doi.org/10.1038/ismej.2012.92>
- Fogg, P. C. M., Colloms, S., Rosser, S., Stark, M., & Smith, M. C. M. (2014). New applications for phage integrases. *Journal of Molecular Biology*, 426(15), 2703–2716. <https://doi.org/10.1016/j.jmb.2014.05.014>
- Frank, J. A., Reich, C. I., Sharma, S., Weisbaum, J. S., Wilson, B. A., & Olsen, G. J. (2008). Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Applied and Environmental Microbiology*, 74(8), 2461–2470. <https://doi.org/10.1128/AEM.02272-07>
- Fuhrman, J. A. (1999). Marine viruses and their biogeochemical and ecological effects. *Nature*, 399(6736), 541–548. <https://doi.org/10.1038/21119>
- Ghai, R., Mehrshad, M., Mizuno, C. M., & Rodriguez-Valera, F. (2017). Metagenomic recovery of phage genomes of uncultured freshwater actinobacteria. *The ISME Journal*, 11(1), 304–308. <https://doi.org/10.1038/ismej.2016.110>

- Gobler, C. J., Hutchins, D. A., Fisher, N. S., Cosper, E. M., & Sañudo-Wilhelmy, S. A. (1997). Release and bioavailability of C, N, P Se, and Fe following viral lysis of a marine chrysophyte. *Limnology and Oceanography*, 42(7), 1492–1504.
- Goff, S. A., Vaughn, M., McKay, S., Lyons, E., Stapleton, A. E., Gessler, D., Matasci, N., Wang, L., Hanlon, M., Lenards, A., Muir, A., Merchant, N., Lowry, S., Mock, S., Helmke, M., Kubach, A., Narro, M., Hopkins, N., Micklos, D., ... Stanzione, D. (2011). The iPlant Collaborative: Cyberinfrastructure for Plant Biology. *Frontiers in Plant Science*, 2, 34. <https://doi.org/10.3389/fpls.2011.00034>
- Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews. Genetics*, 17(6), 333–351. <https://doi.org/10.1038/nrg.2016.49>
- Graham, E., Paez-Espino, D., Brislawn, C., Hofmockel, K., Wu, R., Kyrpides, N., Jansson, J., & McDermott, J. (2019). *Untapped viral diversity in global soil metagenomes*. bioRxiv. <https://doi.org/10.1101/583997>
- Greenwald, J., Hoegy, F., Nader, M., Journet, L., Mislin, G. L. A., Graumann, P. L., & Schalk, I. J. (2007). Real time fluorescent resonance energy transfer visualization of ferric pyoverdine uptake in *Pseudomonas aeruginosa*. A role for ferrous iron. *The Journal of biological chemistry*, 282(5), 2987–2995. <https://doi.org/10.1074/jbc.M609238200>
- Greninger, A. L. (2018). A decade of RNA virus metagenomics is (not) enough. *Virus Research*, 244, 218–229. <https://doi.org/10.1016/j.virusres.2017.10.014>
- Guajardo-Leiva, S., Santos, F., Salgado, O., Regeard, C., Quillet, L., & Díez, B. (2021). Unveiling Ecological and Genetic Novelty within Lytic and Lysogenic Viral Communities of Hot Spring Phototrophic Microbial Mats. *Microbiology spectrum*, 9(3), e0069421. <https://doi.org/10.1128/Spectrum.00694-21>
- Gudbergssdóttir, S. R., Menzel, P., Krogh, A., Young, M., & Peng, X. (2016). Novel viral genomes identified from six metagenomes reveal wide distribution of archaeal viruses and high viral diversity in terrestrial hot springs. *Environmental microbiology*, 18(3), 863–874. <https://doi.org/10.1111/1462-2920.13079>
- Guerin, E., & Hill, C. (2020). Shining Light on Human Gut Bacteriophages. *Frontiers in Cellular and Infection Microbiology*, 10, 481. <https://doi.org/10.3389/fcimb.2020.00481>

- Gupta, R., Xu, S. Y., Sharma, P., & Capalash, N. (2012). Characterization of MspNI (G/GWCC) and MspNII (R/GATCY), novel thermostable Type II restriction endonucleases from *Meiothermus* SP., isoschizomers of AvaII and BstYI. *Molecular Biology Reports*, 39(5), 5607–5614. <https://doi.org/10.1007/s11033-011-1365-2>
- Gurav, T., Singh, H. K., & Chandrasekharam, D. (2016). Major and trace element concentrations in the geothermal springs along the west coast of Maharashtra, India. *Arabian Journal of Geosciences*, 9(1), 1–15. <https://doi.org/10.1007/s12517-015-2139-2>
- Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUASt: quality assessment tool for genome assemblies. *Bioinformatics (Oxford, England)*, 29(8), 1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
- Gursahani Y H and Gupta. (2011). Decolourization of textile effluent by a thermophilic bacteria *Anoxybacillus rupiensis*. *Journal of Petroleum & Environmental Biotechnology*, 02(02). <https://doi.org/10.4172/2157-7463.1000111>
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41(41), 95–98.
- Hara, S., Terauchi, K., & Koike, I. (1991). Abundance of viruses in marine waters: assessment by epifluorescence and transmission electron microscopy. *Appl Environ Microbiol*, 57.
- Häring, M., Rachel, R., Peng, X., Garrett, R. A., Prangishvili, D., Haring, M., Rachel, R., Peng, X., Garrett, R. A., & Prangishvili, D. (2005). Viral diversity in hot springs of Pozzuoli, Italy, and characterization of a unique archaeal virus, Acidianus bottle-shaped virus, from a new family, the Ampullaviridae. *J Virol*, 79(15), 9904–9911. <https://doi.org/10.1128/JVI.79.15.9904-9911.2005>
- Hatfull, G. F. (2008). Bacteriophage genomics. *Current Opinion in Microbiology*, 11(5), 447–453. <https://doi.org/10.1016/j.mib.2008.09.004>
- Hendrix, R. W., Smith, M. C., Burns, R. N., Ford, M. E., & Hatfull, G. F. (1999). Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proceedings of the National Academy of Sciences of the United States of America*, 96(5), 2192–2197. <https://doi.org/10.1073/pnas.96.5.2192>

- Heyerhoff, B., Engelen, B., & Bunse, C. (2022). Auxiliary Metabolic Gene Functions in Pelagic and Benthic Viruses of the Baltic Sea. *Frontiers in Microbiology*, 13(July), 1–16. <https://doi.org/10.3389/fmicb.2022.863620>
- Hjorleifsdottir, S., Aevansson, A., Hreggvidsson, G. O., Fridjonsson, O. H., & Kristjansson, J. K. (2014). Isolation, growth and genome of the Rhodothermus RM378 thermophilic bacteriophage. *Extremophiles*, 18(2), 261–270. <https://doi.org/10.1007/s00792-013-0613-x>
- Horzinek, M. C. (1997). The birth of virology. *Antonie van Leeuwenhoek*, 71(1–2), 15–20. <https://doi.org/10.1023/a:1000197505492>
- Howard-Varona, C., Hargreaves, K. R., Abedon, S. T., & Sullivan, M. B. (2017). Lysogeny in nature: Mechanisms, impact and ecology of temperate phages. *ISME Journal*, 11(7), 1511–1520. <https://doi.org/10.1038/ismej.2017.16>
- Hurwitz, B. L., Deng, L., Poulos, B. T., & Sullivan, M. B. (2013). Evaluation of methods to concentrate and purify ocean virus communities through comparative, replicated metagenomics. *Environmental Microbiology*, 15(5), 1428–1440. <https://doi.org/10.1111/j.1462-2920.2012.02836.x>
- Hurwitz, B. L., & Sullivan, M. B. (2013). The Pacific Ocean virome (POV): a marine viral metagenomic dataset and associated protein clusters for quantitative viral ecology. *PloS One*, 8(2), e57355. <https://doi.org/10.1371/journal.pone.0057355>
- Hurwitz, B. L., & U'Ren, J. M. (2016). Viral metabolic reprogramming in marine ecosystems. *Current Opinion in Microbiology*, 31, 161–168. <https://doi.org/10.1016/j.mib.2016.04.002>
- ICTV. (2020). How to write virus, species, and other taxa names. *International Committee on Taxonomy of Viruses*, 1–3.
- Ignacio-Espinoza, J. C., Laperriere, S. M., Yeh, Y.-C., Weissman, J., Hou, S., Long, A. M., & Fuhrman, J. A. (2020). Ribosome-linked mRNA-rRNA chimeras reveal active novel virus host associations. *BioRxiv*, 2010–2020.
- Ignacio-Espinoza, J. C., & Sullivan, M. B. (2012). Phylogenomics of T4 cyanophages: lateral gene transfer in the “core” and origins of host genes. *Environmental Microbiology*, 14(8), 2113–2126. <https://doi.org/10.1111/j.1462-2920.2012.02704.x>
- Jaatinen, S. T., Happonen, L. J., Laurinmäki, P., Butcher, S. J., & Bamford, D. H. (2008).

- Biochemical and structural characterisation of membrane-containing icosahedral dsDNA bacteriophages infecting thermophilic *Thermus thermophilus*. *Virology*, 379(1), 10–19. <https://doi.org/10.1016/j.virol.2008.06.023>
- 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 Research*, 48(W1), W395–W402. <https://doi.org/10.1093/nar/gkaa434>
- Jarett, J. K., Džunková, M., Schulz, F., Roux, S., Paez-Espino, D., Eloë-Fadrosh, E., Jungbluth, S. P., Ivanova, N., Spear, J. R., Carr, S. A., Trivedi, C. B., Corsetti, F. A., Johnson, H. A., Becraft, E., Kyrpides, N., Stepanauskas, R., & Woyke, T. (2020). Insights into the dynamics between viruses and their hosts in a hot spring microbial mat. *ISME Journal*, 14(10), 2527–2541. <https://doi.org/10.1038/s41396-020-0705-4>
- Ji, P., Aw, T. G., Van Bonn, W., & Rose, J. B. (2020). Evaluation of a portable nanopore-based sequencer for detection of viruses in water. *Journal of Virological Methods*, 278, 113805. <https://doi.org/10.1016/j.jviromet.2019.113805>
- Jiang, S. C., & Paul, J. H. (1996). Occurrence of Lysogenic Bacteria in Marine Microbial Communities as determined by prophage induction. *Marine Ecology Progress Series*, 142, 27–38.
- John, S. G., Mendez, C. B., Deng, L., Poulos, B., Kauffman, A. K. M., Kern, S., Brum, J., Polz, M. F., Boyle, E. A., & Sullivan, M. B. (2011). A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environmental Microbiology Reports*, 3(2), 195–202. <https://doi.org/10.1111/j.1758-2229.2010.00208.x>
- Jover, L. F., Effler, T. C., Buchan, A., Wilhelm, S. W., & Weitz, J. S. (2014). The elemental composition of virus particles: implications for marine biogeochemical cycles. *Nature Reviews. Microbiology*, 12(7), 519–528. <https://doi.org/10.1038/nrmicro3289>
- Kanokratana, P., Chanapan, S., Pootanakit, K., & Eurwilaichitr, L. (2004). Diversity and abundance of Bacteria and Archaea in the Bor Khlueng Hot Spring in Thailand. *Journal of Basic Microbiology*, 44(6), 430–444. <https://doi.org/10.1002/jobm.200410388>

- Keen, E. C. (2015). A century of phage research: bacteriophages and the shaping of modern biology. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, 37(1), 6–9. <https://doi.org/10.1002/bies.201400152>
- Kieft, K., Zhou, Z., & Anantharaman, K. (2020). VIBRANT: Automated recovery, annotation and curation of microbial viruses, and evaluation of virome function from genomic sequences. *BioRxiv*, 1–23. <https://doi.org/10.1101/855387>
- Kim, K.-H., Chang, H.-W., Nam, Y.-D., Roh, S. W., Kim, M.-S., Sung, Y., Jeon, C. O., Oh, H.-M., & Bae, J.-W. (2008). Amplification of uncultured single-stranded DNA viruses from rice paddy soil. *Applied and Environmental Microbiology*, 74(19), 5975–5985. <https://doi.org/10.1128/AEM.01275-08>
- Kim, S., Park, M. S., Song, J., Kang, I., & Cho, J.-C. (2020). High-throughput cultivation based on dilution-to-extinction with catalase supplementation and a case study of cultivating acI bacteria from Lake Soyang. *Journal of Microbiology (Seoul, Korea)*, 58(11), 893–905. <https://doi.org/10.1007/s12275-020-0452-2>
- Kimura, M., Jia, Z. J., Nakayama, N., & Asakawa, S. (2008). Ecology of viruses in soils: Past, present and future perspectives. *Soil Science and Plant Nutrition*, 54(1), 1–32. <https://doi.org/10.1111/j.1747-0765.2007.00197.x>
- Knowles, B., Silveira, C. B., Bailey, B. A., Barott, K., Cantu, V. A., Cobián-Güemes, A. G., Coutinho, F. H., Dinsdale, E. A., Felts, B., Furby, K. A., George, E. E., Green, K. T., Gregoracci, G. B., Haas, A. F., Haggerty, J. M., Hester, E. R., Hisakawa, N., Kelly, L. W., Lim, Y. W., ... Rohwer, F. (2016). Lytic to temperate switching of viral communities. *Nature*, 531(7595), 466–470. <https://doi.org/10.1038/nature17193>
- Kolari, M., Nuutinen, J., Rainey, F. A., & Salkinoja-Salonen, M. S. (2003). Colored moderately thermophilic bacteria in paper-machine biofilms. *Journal of Industrial Microbiology and Biotechnology*, 30(4), 225–238. <https://doi.org/10.1007/s10295-003-0047-z>
- Koonin, E. V., Dolja, V. V., Krupovic, M., Varsani, A., Wolf, Y. I., Yutin, N., Zerbini, F. M., & Kuhn, J. H. (2020). Global Organization and Proposed Megataxonomy of the Virus World. *Microbiology and Molecular Biology Reviews : MMBR*, 84(2). <https://doi.org/10.1128/MMBR.00061-19>
- Koskella, B., & Meaden, S. (2013). Understanding bacteriophage specificity in natural

- microbial communities. *Viruses*, 5(3), 806–823. <https://doi.org/10.3390/v5030806>
- Kropinski, A. M. (2018). Practical advice on the one-step growth curve. *Methods in Molecular Biology*, 1681, 41–47. https://doi.org/10.1007/978-1-4939-7343-9_3
- Kropinski, A. M., Mazzocco, A., Waddell, T. E., Lingohr, E., & Johnson, R. P. (2009). Enumeration of bacteriophages by double agar overlay plaque assay. *Bacteriophages: methods and protocols, volume 1: isolation, characterization, and interactions*, 69–76. https://doi.org/10.1007/978-1-60327-164-6_7
- Kruger, D. H., Schneck, P., & Gelderblom, H. R. (2000). Helmut Ruska and the visualisation of viruses. *Lancet (London, England)*, 355(9216), 1713–1717. [https://doi.org/10.1016/s0140-6736\(00\)02250-9](https://doi.org/10.1016/s0140-6736(00)02250-9)
- Krupovic, M., Dolja, V. V., & Koonin, E. V. (2019). Origin of viruses: primordial replicators recruiting capsids from hosts. *Nature Reviews Microbiology*, 17(7), 449–458. <https://doi.org/10.1038/s41579-019-0205-6>
- Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, 35(6), 1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Kutter, E. (2009). Phage host range and efficiency of plating. *Methods in Molecular Biology (Clifton, N.J.)*, 501, 141–149. https://doi.org/10.1007/978-1-60327-164-6_14
- Labbé, M., Girard, C., Vincent, W. F., & Culley, A. I. (2020). Extreme Viral Partitioning in a Marine-Derived High Arctic Lake. *MSphere*, 5(3). <https://doi.org/10.1128/msphere.00334-20>
- Lambeth, C. R., White, L. J., Johnston, R. E., & de Silva, A. M. (2005). Flow cytometry-based assay for titrating dengue virus. *Journal of Clinical Microbiology*, 43(7), 3267–3272. <https://doi.org/10.1128/JCM.43.7.3267-3272.2005>
- Langenfeld, K., Chin, K., Roy, A., Wigginton, K., & Duhaime, M. B. (2021). Comparison of ultrafiltration and iron chloride flocculation in the preparation of aquatic viromes from contrasting sample types. *PeerJ*, 9, e11111. <https://doi.org/10.7717/peerj.11111>
- Lin, L., Han, J., Ji, X., Hong, W., Huang, L., & Wei, Y. (2011). Isolation and characterization of a new bacteriophage MMP17 from *Meiothermus*.

- Extremophiles*, 15(2), 253–258. <https://doi.org/10.1007/s00792-010-0354-z>
- Lin, L., Hong, W., Ji, X., Han, J., Huang, L., & Wei, Y. (2010). Isolation and characterization of an extremely long tail *Thermus* bacteriophage from Tengchong hot springs in China. *Journal of Basic Microbiology*, 50(5), 452–456. <https://doi.org/10.1002/jobm.201000116>
- Lindqvist, B. H., Dehò, G., & Calendar, R. (1993). Mechanisms of genome propagation and helper exploitation by satellite phage P4. *Microbiological Reviews*, 57(3), 683–702. <https://doi.org/10.1128/mr.57.3.683-702.1993>
- Liu, B., Wu, S., Song, Q., Zhang, X., & Xie, L. (2006). Two novel bacteriophages of thermophilic bacteria isolated from deep-sea hydrothermal fields. *Curr Microbiol*, 53(2), 163–166. <https://doi.org/10.1007/s00284-005-0509-9>
- Liu, H., Kheirvari, M., & Tumban, E. (2023). Potential Applications of Thermophilic Bacteriophages in One Health. *International Journal of Molecular Sciences*, 24(9). <https://doi.org/10.3390/ijms24098222>
- Liu, Y. C., Young, L. S., Lin, S. Y., Hameed, A., Hsu, Y. H., Lai, W. A., Shen, F. T., & Young, C. C. (2013). *Pseudomonas guguanensis* sp. nov., a gammaproteobacterium isolated from a hot spring. *International journal of systematic and evolutionary microbiology*, 63(Pt 12), 4591–4598. <https://doi.org/10.1099/ij.s.0.047712-0>
- Lopes, A., Tavares, P., Petit, M. A., Guérois, R., & Zinn-Justin, S. (2014). Automated classification of tailed bacteriophages according to their neck organization. *BMC Genomics*, 15(1), 1–17. <https://doi.org/10.1186/1471-2164-15-1027>
- Low, S. J., Džunková, M., Chaumeil, P. A., Parks, D. H., & Hugenholtz, P. (2019). Evaluation of a concatenated protein phylogeny for classification of tailed double-stranded DNA viruses belonging to the order Caudovirales. *Nature Microbiology*, 4(8), 1306–1315. <https://doi.org/10.1038/s41564-019-0448-z>
- Low, U., Absar, A., Duraiswami, R., & Singh, A. (2020). Geophysical exploration of Tural-Rajwadi group of hot springs, West Coast Geothermal Province, Maharashtra, India and its implications. *Geothermics*, 88, 101874. <https://doi.org/10.1016/j.geothermics.2020.101874>
- Lunde, M., Blatny, J. M., Lillehaug, D., Aastveit, A. H., & Nes, I. F. (2003). Use of real-time quantitative PCR for the analysis of phiLC3 prophage stability in *Lactococci*. *Applied and Environmental Microbiology*, 69(1), 41–48.

<https://doi.org/10.1128/AEM.69.1.41-48.2003>

- Magnúsdóttir, S., Ravcheev, D., de Crécy-Lagard, V., & Thiele, I. (2015). Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. *Frontiers in Genetics*, 6, 148. <https://doi.org/10.3389/fgene.2015.00148>
- Marie, D., Brussaard, C. P. D., Thyraug, R., Bratbak, G., & Vaulot, D. (1999). Enumeration of marine viruses in culture and natural samples by flow cytometry. *Applied and Environmental Microbiology*, 65(1), 45–52. <https://doi.org/10.1128/AEM.65.1.45-52.1999>
- Martínez-García, E., Jatsenko, T., Kivisaar, M., & de Lorenzo, V. (2015). Freeing *Pseudomonas putida* KT2440 of its proviral load strengthens endurance to environmental stresses. *Environmental Microbiology*, 17(1), 76–90. <https://doi.org/10.1111/1462-2920.12492>
- Masurat, P., Fru, E. C., & Pedersen, K. (2005). Identification of *Meiothermus* as the dominant genus in a storage system for spent nuclear fuel. *Journal of Applied Microbiology*, 98(3), 727–740. <https://doi.org/10.1111/j.1365-2672.2004.02519.x>
- Mathan Kumar, R., Jani, K., Parvathi, J. R., Thomas, B. M., Raja, S. S. S., Pandey, A., & Sharma, A. (2022). Bacterial diversity of geochemically distinct hot springs located in Maharashtra, India. *Archives of Microbiology*, 204(1), 1–11. <https://doi.org/10.1007/s00203-021-02728-2>
- Mathur, J., Bizzoco, R. W., Ellis, D. G., Lipson, D. A., Poole, A. W., Levine, R., & Kelley, S. T. (2007). Effects of abiotic factors on the phylogenetic diversity of bacterial communities in acidic thermal springs. *Applied and Environmental Microbiology*, 73(8), 2612–2623. <https://doi.org/10.1128/AEM.02567-06>
- Matsushita, I., & Yanase, H. (2009). The genomic structure of thermus bacteriophage ϕ iN93. *Journal of Biochemistry*, 146(6), 775–785. <https://doi.org/10.1093/jb/mvp125>
- McNair, K., Aziz, R. K., Pusch, G. D., Overbeek, R., Dutilh, B. E., & Edwards, R. (2018). Phage genome annotation using the RAST pipeline. *Methods in Molecular Biology*, 1681, 231–238. https://doi.org/10.1007/978-1-4939-7343-9_17
- Mehetre, G. T., Paranjpe, A. S., Dastager, S. G., & Dharne, M. S. (2016). Complete metagenome sequencing based bacterial diversity and functional insights from basaltic hot spring of Unkeshwar, Maharashtra, India. *Genom Data*, 7, 140–143.

<https://doi.org/10.1016/j.gdata.2015.12.031>

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(D1), D801–D807.

<https://doi.org/10.1093/nar/gkab902>

Meier-Kolthoff, J. P., & Göker, M. (2017). VICTOR: genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics (Oxford, England)*, 33(21), 3396–3404. <https://doi.org/10.1093/bioinformatics/btx440>

Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications*, 7(1), 11257.

<https://doi.org/10.1038/ncomms11257>

Middelboe, M., Chan, A. M., & Bertelsen, S. K. (2010). Isolation and life cycle characterization of lytic viruses infecting heterotrophic bacteria and cyanobacteria. *Manual of Aquatic Viral Ecology*, 118–133.

<https://doi.org/10.4319/mave.2010.978-0-9845591-0-7.118>

Minakhin, L., Goel, M., Berdygulova, Z., Ramanculov, E., Florens, L., Glazko, G., Karamychev, V. N., Slesarev, A. I., Kozyavkin, S. A., Khromov, I., Ackermann, H. W., Washburn, M., Mushegian, A., & Severinov, K. (2008). Genome Comparison and Proteomic Characterization of *Thermus thermophilus* Bacteriophages P23-45 and P74-26: Siphoviruses with Triplex-forming Sequences and the Longest Known Tails. *Journal of Molecular Biology*, 378(2), 468–480.

<https://doi.org/10.1016/j.jmb.2008.02.018>

Miranda, J. A., Culley, A. I., Schvarcz, C. R., & Steward, G. F. (2016). RNA viruses as major contributors to Antarctic viroplankton. *Environmental Microbiology*, 18(11), 3714–3727. <https://doi.org/10.1111/1462-2920.13291>

Moon, K., Kim, S., Kang, I., & Cho, J. C. (2020). Viral metagenomes of Lake Soyang, the largest freshwater lake in South Korea. *Scientific Data*, 7(1), 1–6.

<https://doi.org/10.1038/s41597-020-00695-9>

Moraru, C. (2023). VirClust-A Tool for Hierarchical Clustering, Core Protein Detection and Annotation of (Prokaryotic) Viruses. *Viruses*, 15(4).

<https://doi.org/10.3390/v15041007>

Moraru, C., Varsani, A., & Kropinski, A. M. (2020). VIRIDIC-A Novel Tool to Calculate

- the Intergenomic Similarities of Prokaryote-Infecting Viruses. *Viruses*, 12(11).
<https://doi.org/10.3390/v12111268>
- Motlagh, A. M., Bhattacharjee, A. S., Coutinho, F. H., Dutilh, B. E., Casjens, S. R., & Goel, R. K. (2017). Insights of phage-host interaction in hypersaline ecosystem through metagenomics analyses. *Frontiers in Microbiology*, 8(MAR), 1–15.
<https://doi.org/10.3389/fmicb.2017.00352>
- Mougari, S., Sahmi-Bounsiar, D., Levasseur, A., Colson, P., & Scola, B. La. (2019). Virophages of giant viruses: An update at eleven. *Viruses*, 11(8), 1–28.
<https://doi.org/10.3390/v11080733>
- Mukherjee, T., Bose, S., & Mukhopadhyay, S. K. (2017). Antioxidant properties of the carotenoid extracts of three Deinococcus–Thermus phylum bacteria, Meiothermus sp. strains RP and TP and Thermus sp. strain YY from Paniphala hot spring, India. *Nutrire*, 42(1). <https://doi.org/10.1186/s41110-017-0032-3>
- Mukherjee, T., Bose, S., Sen, U., Roy, C., Rameez, M. J., Ghosh, W., & Mukhopadhyay, S. K. (2016). Genome sequence of the red pigment-forming Meiothermus taiwanensis strain RP isolated from Paniphala hot spring, India. *Genome Announcements*, 4(3). <https://doi.org/10.1128/genomeA.00629-16>
- Nagayoshi, Y., Kumagae, K., Mori, K., Tashiro, K., Nakamura, A., Fujino, Y., Hiromasa, Y., Iwamoto, T., Kuhara, S., Ohshima, T., & Doi, K. (2016). Physiological properties and genome structure of the hyperthermophilic filamentous phage ϕ OH3 which infects Thermus thermophilus HB8. *Frontiers in Microbiology*, 7(FEB), 1–11. <https://doi.org/10.3389/fmicb.2016.00050>
- Narsing Rao, M. P., Dong, Z. Y., Luo, Z. H., Li, M. M., Liu, B. B., Guo, S. X., Hozzein, W. N., Xiao, M., & Li, W. J. (2021). Physicochemical and Microbial Diversity Analyses of Indian Hot Springs. *Frontiers in Microbiology*, 12(March).
<https://doi.org/10.3389/fmicb.2021.627200>
- Naryshkina, T., Liu, J., Florens, L., Swanson, S. K., Pavlov, A. R., Pavlova, N. V, Inman, R., Minakhin, L., Kozyavkin, S. A., Washburn, M., Mushegian, A., & Severinov, K. (2006). Thermus thermophilus bacteriophage phiYS40 genome and proteomic characterization of virions. *Journal of Molecular Biology*, 364(4), 667–677.
<https://doi.org/10.1016/j.jmb.2006.08.087>
- Nayfach, S., Camargo, A. P., Schulz, F., Eloie-Fadrosh, E., Roux, S., & Kyrpides, N. C.

- (2021). CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nature Biotechnology*, 39(5), 578–585.
<https://doi.org/10.1038/s41587-020-00774-7>
- Nishimura, Y., Yoshida, T., Kuronishi, M., Uehara, H., Ogata, H., & Goto, S. (2017). ViPTree: the viral proteomic tree server. *Bioinformatics (Oxford, England)*, 33(15), 2379–2380. <https://doi.org/10.1093/bioinformatics/btx157>
- Nobre, M. F., Trüper, H. G., & Da Costa, M. S. (1996). Transfer of *Thermus ruber* (Loginova et al. 1984), *Thermus silvanus* (Tenreiro et al. 1995), and *Thermus chliarophilus* (Tenreiro et al. 1995) to *Meiothermus* gen. nov. as *Meiothermus ruber* comb. nov., *Meiothermus silvanus* comb. nov., and *Meiothermus chliarop*. *International Journal of Systematic Bacteriology*, 46(2), 604–606.
<https://doi.org/10.1099/00207713-46-2-604>
- Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Research*, 27(5), 824–834.
<https://doi.org/10.1101/gr.213959.116>
- Ogino, H., Teramoto, H., Inui, M., & Yukawa, H. (2008). DivS, a novel SOS-inducible cell-division suppressor in *Corynebacterium glutamicum*. *Molecular Microbiology*, 67(3), 597–608. <https://doi.org/10.1111/j.1365-2958.2007.06069.x>
- Oliveira, J., Mahony, J., Hanemaaijer, L., Kouwen, T. R. H. M., Neve, H., MacSharry, J., & van Sinderen, D. (2017). Detecting *Lactococcus lactis* prophages by mitomycin C-mediated induction coupled to flow cytometry analysis. *Frontiers in Microbiology*, 8, 1–11. <https://doi.org/10.3389/fmicb.2017.01343>
- Oostu, K., Matsumoto, T., Harada, S., & Kishi, T. (1975). Antitumor and immunosuppressive activities of lankacidin-group antibiotics: structure-activity relationships. *Cancer Chemotherapy Reports*, 59(5), 919–928.
- Overman, S. A., Bondre, P., Maiti, N. C., & Thomas, G. J. J. (2005). Structural characterization of the filamentous bacteriophage PH75 from *Thermus thermophilus* by Raman and UV-resonance Raman spectroscopy. *Biochemistry*, 44(8), 3091–3100. <https://doi.org/10.1021/bi048163d>
- Overview of Virus Metagenomic Classification Methods and Their Biological Applications, 9 *Frontiers in Microbiology* (2018).
<https://doi.org/10.3389/fmicb.2018.00749>

- Paez-Espino, D., Chen, I.-M. A., Palaniappan, K., Ratner, A., Chu, K., Szeto, E., Pillay, M., Huang, J., Markowitz, V. M., Nielsen, T., Huntemann, M., K Reddy, T. B., Pavlopoulos, G. A., Sullivan, M. B., Campbell, B. J., Chen, F., McMahon, K., Hallam, S. J., Denef, V., ... Kyrpides, N. C. (2017). IMG/VR: a database of cultured and uncultured DNA Viruses and retroviruses. *Nucleic Acids Research*, 45(D1), D457–D465. <https://doi.org/10.1093/nar/gkw1030>
- Paez-Espino, D., Eloë-Fadrosch, E. A., Pavlopoulos, G. A., Thomas, A. D., Huntemann, M., Mikhailova, N., Rubin, E., Ivanova, N. N., & Kyrpides, N. C. (2016). Uncovering Earth's virome. *Nature*, 536(7617), 425–430. <https://doi.org/10.1038/nature19094>
- Panda, A. K., Bisht, S. S., DeMandal, S., Kumar, N. S., De Mandal, S., & Kumar, N. S. (2016). Bacterial and archeal community composition in hot springs from Indo-Burma region, North-east India. *J*, 6(1), 6–11. <https://doi.org/10.1186/s13568-016-0284-y>
- Panda, A. K., Bisht, S. S., Kumar, N. S., & De Mandal, S. (2015). Investigations on microbial diversity of Jakrem hot spring, Meghalaya, India using cultivation-independent approach. *Genomics Data*. <https://doi.org/10.1016/j.gdata.2015.04.016>
- Parte, A. C., Sardà Carbasse, J., Meier-Kolthoff, J. P., Reimer, L. C., & Göker, M. (2020). List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *International Journal of Systematic and Evolutionary Microbiology*, 70(11), 5607–5612. <https://doi.org/10.1099/ijsem.0.004332>
- Patel, B. K. C., Chalcraft, J. P., Morgan, H. W., & Daniel, R. M. (1989). In Situ Morphologies of Some Bacteria from New Zealand Hot Springs. *Systematic and Applied Microbiology*, 11(2), 187–193. [https://doi.org/10.1016/S0723-2020\(89\)80060-8](https://doi.org/10.1016/S0723-2020(89)80060-8)
- Patel, M. A., Ou, M. S., Harbrucker, R., Aldrich, H. C., Buszko, M. L., Ingram, L. O., & Shanmugam, K. T. (2006). Isolation and characterization of acid-tolerant, thermophilic bacteria for effective fermentation of biomass-derived sugars to lactic acid. *Applied and Environmental Microbiology*, 72(5), 3228–3235. <https://doi.org/10.1128/AEM.72.5.3228-3235.2006>
- Pawlowski, A., Rissanen, I., Bamford, J. K. H. H., Krupovic, M., & Jalasvuori, M. (2014). Gammasphaerolipovirus, a newly proposed bacteriophage genus, unifies viruses of halophilic archaea and thermophilic bacteria within the novel family

- Sphaerolipoviridae. *Archives of Virology*, 159(6), 1541–1554.
<https://doi.org/10.1007/s00705-013-1970-6>
- Pedersen, M., Ligowska, M., & Hammer, K. (2010). Characterization of the CI repressor protein encoded by the temperate Lactococcal phage TP901-1. *Journal of Bacteriology*, 192(8), 2102–2110. <https://doi.org/10.1128/JB.01387-09>
- Pederson, D. M., Welsh, L. C., Marvin, D. A., Sampson, M., Perham, R. N., Yu, M., & Slater, M. R. (2001). The protein capsid of filamentous bacteriophage PH75 from *Thermus thermophilus*. *Journal of Molecular Biology*, 309(2), 401–421.
<https://doi.org/10.1006/jmbi.2001.4685>
- Peix, A., Ramírez-Bahena, M. H., & Velázquez, E. (2009). Historical evolution and current status of the taxonomy of genus *Pseudomonas*. *Infection, Genetics and Evolution*, 9(6), 1132–1147. <https://doi.org/10.1016/J.MEEGID.2009.08.001>
- Peng, X., Garrett, R. A., & She, Q. X. (2012). Archaeal viruses-novel, diverse and enigmatic. *Science China Life Sciences*, 55(5), 422–433.
<https://doi.org/10.1007/s11427-012-4325-8>
- Peng, Y., Leung, H. C. M., Yiu, S. M., & Chin, F. Y. L. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics (Oxford, England)*, 28(11), 1420–1428.
<https://doi.org/10.1093/bioinformatics/bts174>
- Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., Gorsky, G., Iudicone, D., Karsenti, E., Speich, S., Troublé, R., Dimier, C., & Searson, S. (2015). Open science resources for the discovery and analysis of Tara Oceans data. *Scientific Data*, 2, 150023. <https://doi.org/10.1038/sdata.2015.23>
- Piel, D., Bruto, M., Labreuche, Y., Blanquart, F., Goudenège, D., Barcia-Cruz, R., Chenivresse, S., Le Panse, S., James, A., Dubert, J., Petton, B., Lieberman, E., Wegner, K. M., Hussain, F. A., Kauffman, K. M., Polz, M. F., Bikard, D., Gandon, S., Rocha, E. P. C., & Le Roux, F. (2022). Phage-host coevolution in natural populations. *Nature Microbiology*, 7(7), 1075–1086.
<https://doi.org/10.1038/s41564-022-01157-1>
- Pitale, U. L., Dubey, R., Saxena, R. K., Prasad, J. M., Muthuraman, K., & Thussu, J. L. (1989). Review of geothermal studies of west coast hot spring belt, Maharashtra. *Records of the Geological Survey of India*, 115(3), 97-136.

- Poddar, A., & Das, S. K. (2018). Microbiological studies of hot springs in India: a review. *Archives of microbiology*, 200(1), 1–18. <https://doi.org/10.1007/s00203-017-1429-3>
- Prangishvili, D. (2013). The wonderful world of archaeal viruses. *Annual Review of Microbiology*, 67, 565–585. <https://doi.org/10.1146/annurev-micro-092412-155633>
- Prangishvili, D., Bamford, D. H., Forterre, P., Iranzo, J., Koonin, E. V., & Krupovic, M. (2017). The enigmatic archaeal virosphere. *Nature reviews. Microbiology*, 15(12), 724–739. <https://doi.org/10.1038/nrmicro.2017.125>
- Prangishvili, D., Koonin, E. V., & Krupovic, M. (2013). Genomics and biology of Rudiviruses, a model for the study of virus–host interactions in Archaea. *Biochemical Society Transactions*, 41(1), 443–450. <https://doi.org/10.1042/BST20120313>
- Prata, C., Ribeiro, A., Cunha, A., Gomes, N. C., & Almeida, A. (2012). Ultracentrifugation as a direct method to concentrate viruses in environmental waters: virus-like particle enumeration as a new approach to determine the efficiency of recovery. *J Environ Monit*, 14(1), 64–70. <https://doi.org/10.1039/c1em10603a>
- Pride, D. T., & Schoenfeld, T. (2008). Genome signature analysis of thermal virus metagenomes reveals Archaea and thermophilic signatures. *BMC Genomics*, 9, 420. <https://doi.org/10.1186/1471-2164-9-420>
- Rachel, R., Bettstetter, M., Hedlund, B. P., Häring, M., Kessler, A., Stetter, K. O., & Prangishvili, D. (2002). Remarkable morphological diversity of viruses and virus-like particles in hot terrestrial environments. *Archives of Virology*, 147(12), 2419–2429. <https://doi.org/10.1007/s00705-002-0895-2>
- Rastrojo, A., & Alcamí, A. (2017). Aquatic viral metagenomics: Lights and shadows. *Virus Research*, 239, 87–96. <https://doi.org/10.1016/j.virusres.2016.11.021>
- Ren, J., Song, K., Deng, C., Ahlgren, N. A., Fuhrman, J. A., Li, Y., Xie, X., Poplin, R., & Sun, F. (2020). Identifying viruses from metagenomic data using deep learning. *Quantitative Biology (Beijing, China)*, 8(1), 64–77. <https://doi.org/10.1007/s40484-019-0187-4>
- Rice, G., Stedman, K., Snyder, J., Wiedenheft, B., Willits, D., Brumfield, S., McDermott, T., & Young, M. J. (2001). Viruses from extreme thermal environments. *Proceedings of the National Academy of Sciences of the United States of America*,

- 98(23), 13341–13345. <https://doi.org/10.1073/pnas.231170198>
- Rice, S. A., Tan, C. H., Mikkelsen, P. J., Kung, V., Woo, J., Tay, M., Hauser, A., McDougald, D., Webb, J. S., & Kjelleberg, S. (2009). The biofilm life cycle and virulence of *Pseudomonas aeruginosa* are dependent on a filamentous prophage. *The ISME Journal*, 3(3), 271–282. <https://doi.org/10.1038/ismej.2008.109>
- Richert-Pöggeler, K. R., Franzke, K., Hipp, K., & Kleespies, R. G. (2018). Electron Microscopy Methods for Virus Diagnosis and High-Resolution Analysis of Viruses. *Frontiers in Microbiology*, 9, 3255. <https://doi.org/10.3389/fmicb.2018.03255>
- Roux, S., Adriaenssens, E. M., Dutilh, B. E., Koonin, E. V., Kropinski, A. M., Krupovic, M., Kuhn, J. H., Lavigne, R., Brister, J. R., Varsani, A., Amid, C., Aziz, R. K., Bordenstein, S. R., Bork, P., Breitbart, M., Cochrane, G. R., Daly, R. A., Desnues, C., Duhaime, M. B., ... Elie-Fadrosh, E. A. (2019). Minimum information about an uncultivated virus genome (MIUVIG). *Nature Biotechnology*, 37(1), 29–37. <https://doi.org/10.1038/nbt.4306>
- Roux, S., Brum, J. R., Dutilh, B. E., Sunagawa, S., Duhaime, M. B., Loy, A., Poulos, B. T., Solonenko, N., Lara, E., Poulain, J., Pesant, S., Kandels-Lewis, S., Dimier, C., Picheral, M., Searson, S., Cruaud, C., Alberti, A., Duarte, C. M., Gasol, J. M., ... Sullivan, M. B. (2016). Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature*, 537(7622), 689–693. <https://doi.org/10.1038/nature19366>
- Roux, S., Enault, F., Hurwitz, B. L., & Sullivan, M. B. (2015). VirSorter: mining viral signal from microbial genomic data. *PeerJ*, 3, e985. <https://doi.org/10.7717/peerj.985>
- Roux, S., Hallam, S. J., Woyke, T., & Sullivan, M. B. (2015). Viral dark matter and virus–host interactions resolved from publicly available microbial genomes. *ELife*, 4(January), 1–20. <https://doi.org/10.7554/elife.08490>
- Roux, S., Páez-Espino, D., Chen, I. M. A., Palaniappan, K., Ratner, A., Chu, K., ... & Kyrpides, N. C. (2021). IMG/VR v3: An integrated ecological and evolutionary framework for interrogating genomes of uncultivated viruses. *Nucleic acids research*, 49 (D1), D764–D775. <https://doi.org/10.1093/nar/gkaa946>
- Sakaki, Y., and Oshima, T. (1975). Isolation and characterization of a bacteriophage

- infectious to an extreme thermophile, *Thermus thermophilus* HB8. *Journal of Virology*, 15(6), 1449–1453. <https://doi.org/10.1128/JVI.15.6.1449-1453.1975>
- Sakaki, Y., and Oshima, T. (1976). A new lipid-containing phage infecting acidophilic thermophilic bacteria. *Virology*, 75(1), 256–259. [https://doi.org/10.1016/0042-6822\(76\)90024-6](https://doi.org/10.1016/0042-6822(76)90024-6)
- Salmond, G. P., & Fineran, P. C. (2015). A century of the phage: past, present and future. *Nature Reviews Microbiology*, 13(12), 777–786. <https://doi.org/10.1038/nrmicro3564>
- Sambrook, J. (1989). Molecular cloning: a laboratory manual. *Cold Spring Harbor Laboratory*.
- Sangwan, N., Lambert, C., Sharma, A., Gupta, V., Khurana, P., Khurana, J. P., Sockett, R. E., Gilbert, J. A., & Lal, R. (2015). Arsenic rich Himalayan hot spring metagenomics reveal genetically novel predator-prey genotypes. *Environmental Microbiology Reports*, 7(6), 812–823. <https://doi.org/10.1111/1758-2229.12297>
- Santiago-Rodriguez, T. M., & Hollister, E. B. (2022). Unraveling the viral dark matter through viral metagenomics. *Frontiers in Immunology*, 13(September), 1–13. <https://doi.org/10.3389/fimmu.2022.1005107>
- Sarolkar, P. B. (2005). Geochemical characters of hot springs of west coast, Maharashtra state, India. In *Proceedings world geothermal congress, Antalya, Turkey* (24–29).
- Saxena, R., Dhakan, D. B., Mittal, P., Waiker, P., Chowdhury, A., Ghatak, A., Sharma, V. K., Chowdhury, A., Ghatak, A., & Sharma, V. K. (2017). Metagenomic analysis of hot springs in central India reveals hydrocarbon degrading thermophiles and pathways essential for survival in extreme environments. *Frontiers in Microbiology*, 7, 2123. <https://doi.org/10.3389/fmicb.2016.02123>
- Schaeffer, A. B., & Fulton, M. D. (1933). A Simplified method of staining endospores. *Science (New York, N.Y.)*, 77(1990), 194. <https://doi.org/10.1126/science.77.1990.194>
- Schoenfeld, T., Patterson, M., Richardson, P. M., Wommack, K. E., Young, M., & Mead, D. (2008). Assembly of viral metagenomes from Yellowstone hot springs. *Applied and Environmental Microbiology*, 74(13), 4164–4174. <https://doi.org/10.1128/AEM.02598-07>
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*

(Oxford, England), 30(14), 2068–2069.

<https://doi.org/10.1093/bioinformatics/btu153>

- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11), 2498–2504. <https://doi.org/10.1101/gr.1239303>
- Sharma, A., Schmidt, M., Kiesel, B., Mahato, N. K., Cralle, L., Singh, Y., Richnow, H. H., Gilbert, J. A., Arnold, W., & Lal, R. (2018). Bacterial and Archaeal Viruses of Himalayan Hot Springs at Manikaran Modulate Host Genomes. *Frontiers in Microbiology*, 9, 3095. <https://doi.org/10.3389/fmicb.2018.03095>
- Shi, W., Sun, Q., Fan, G., Hideaki, S., Moriya, O., Itoh, T., Zhou, Y., Cai, M., Kim, S.-G., Lee, J.-S., Sedlacek, I., Arahall, D. R., Lucena, T., Kawasaki, H., Evtushenko, L., Weir, B. S., Alexander, S., Dénes, D., Tanasupawat, S., ... Ma, J. (2021). gcType: a high-quality type strain genome database for microbial phylogenetic and functional research. *Nucleic Acids Research*, 49(D1), D694–D705. <https://doi.org/10.1093/nar/gkaa957>
- Shirkot, P., Verma, A., Gupta, M., Manjul, A. S., & Sharma, R. (2018). Assessment of Physicochemical Characteristics and Thermophilic Bacterial Enumeration from Hot Water Springs of Himachal Pradesh. *National Academy Science Letters*, 41(5), 285–287. <https://doi.org/10.1007/s40009-018-0676-4>
- Simmonds, P., Adams, M. J., Benk, M., Breitbart, M., Brister, J. R., Carstens, E. B., Davison, A. J., Delwart, E., Gorbalenya, A. E., Harrach, B., Hull, R., King, A. M. Q., Koonin, E. V., Krupovic, M., Kuhn, J. H., Lefkowitz, E. J., Nibert, M. L., Orton, R., Roossinck, M. J., ... Zerbini, F. M. (2017). Consensus statement: Virus taxonomy in the age of metagenomics. *Nature Reviews Microbiology*, 15(3), 161–168. <https://doi.org/10.1038/nrmicro.2016.177>
- Singh, H. K., Chandrasekharam, D., Trupti, G., Mohite, P., Singh, B., Varun, C., & Sinha, S. K. (2016). Potential Geothermal Energy Resources of India: A Review. *Current Sustainable/Renewable Energy Reports*, 3(3–4), 80–91. <https://doi.org/10.1007/s40518-016-0054-0>
- Smith, L. S., & Krueger, A. P. (1954). Characteristics of a new vibrio-bacteriophage system. *The Journal of General Physiology*, 38(2), 161–168.

<https://doi.org/10.1085/jgp.38.2.161>

- Solonenko, S. A., Ignacio-Espinoza, J. C., Alberti, A., Cruaud, C., Hallam, S., Konstantinidis, K., Tyson, G., Wincker, P., & Sullivan, M. B. (2013). Sequencing platform and library preparation choices impact viral metagenomes. *BMC Genomics*, 14, 320. <https://doi.org/10.1186/1471-2164-14-320>
- Song, J., Xia, F., Jiang, H., Li, X., Hu, L., Gong, P., Lei, L., Feng, X., Sun, C., Gu, J., & Han, W. (2016). Identification and characterization of HolGH15: the holin of *Staphylococcus aureus* bacteriophage GH15. *The Journal of General Virology*, 97(5), 1272–1281. <https://doi.org/10.1099/jgv.0.000428>
- Song, W., Sun, H.-X., Zhang, C., Cheng, L., Peng, Y., Deng, Z., Wang, D., Wang, Y., Hu, M., Liu, W., Yang, H., Shen, Y., Li, J., You, L., & Xiao, M. (2019). Prophage Hunter: an integrative hunting tool for active prophages. *Nucleic Acids Research*, 47(W1), W74–W80. <https://doi.org/10.1093/nar/gkz380>
- Steward, G. F., & Culley, A. I. (2010). Extraction and purification of nucleic acids from viruses. *Manual of aquatic viral ecology*, 16(1), 154-165.
- Stothard, P., Grant, J. R., & Van Domselaar, G. (2019). Visualizing and comparing circular genomes using the CGView family of tools. *Briefings in Bioinformatics*, 20(4), 1576–1582. <https://doi.org/10.1093/bib/bbx081>
- Suttle, C. (1993). Enumeration and isolation of viruses. In P. F. Kemp, J. J. Cole, B. F. Sherr, & E. B. Sherr (Eds.), *Handbook of Methods in Aquatic Microbial Ecology*. CRC Press.
- Suttle, C. A. (2005). Viruses in the sea. *Nature*, 437(7057), 356–361. <https://doi.org/10.1038/nature04160>
- Suttle, C. A. (2007). Marine viruses--major players in the global ecosystem. *Nature Reviews. Microbiology*, 5(10), 801–812. <https://doi.org/10.1038/nrmicro1750>
- Suttle, C. A. (2013). Viruses: unlocking the greatest biodiversity on Earth. *Genome*, 56(10), 542-544. <https://doi.org/10.1139/gen-2013-0152>
- Tamakoshi, M., Murakami, A., Sugisawa, M., Tsuneizumi, K., Takeda, S., Saheki, T., Izumi, T., Akiba, T., Mitsuoka, K., Toh, H., Yamashita, A., Arisaka, F., Hattori, M., Oshima, T., & Yamagishi, A. (2011). Genomic and proteomic characterization of the large Myoviridae bacteriophage ϕ TMA of the extreme thermophile *Thermus thermophilus*. *Bacteriophage*, 1(3), 152–164.

<https://doi.org/10.4161/bact.1.3.16712>

Thingstad, T. F. (2000). Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems.

Limnology and Oceanography, 45(6), 1320–1328.

Thurber, R. V. (2009). Current insights into phage biodiversity and biogeography. *Current Opinion in Microbiology*, 12(5), 582–587.

<https://doi.org/10.1016/j.mib.2009.08.008>

Thurber, R. V., Haynes, M., Breitbart, M., Wegley, L., & Rohwer, F. (2009). Laboratory procedures to generate viral metagenomes. *Nature Protocols*, 4(4), 470–483.

<https://doi.org/10.1038/nprot.2009.10>

Tindall, B. J., Sikorski, J., Lucas, S., Goltsman, E., Copeland, A., del Rio, T. G., Nolan, M., Tice, H., Cheng, J. F., Han, C., Pitluck, S., Liolios, K., Ivanova, N., Mavromatis, K., Ovchinnikova, G., Pati, A., Fährnich, R., Goodwin, L., Chen, A., ... Lapidus, A. (2010). Complete genome sequence of *Meiothermus ruber* type strain (21 T). *Standards in Genomic Sciences*, 3(1), 26–36.

<https://doi.org/10.4056/sigs.1032748>

Torrella, F., & Morita, R. Y. (1979). Evidence by electron micrographs for a high incidence of bacteriophage particles in the waters of Yaquina Bay, Oregon: ecological and taxonomical implications. *Applied and Environmental Microbiology*, 37(4), 774–778.

<https://doi.org/10.1128/aem.37.4.774-778.1979>

Trubl, G., Jang, H. Bin, Roux, S., Emerson, J. B., Solonenko, N., Vik, D. R., Solden, L., Ellenbogen, J., Runyon, A. T., Bolduc, B., Woodcroft, B. J., Saleska, S. R., Tyson, G. W., Wrighton, K. C., Sullivan, M. B., & Rich, V. I. (2018). Soil Viruses Are Underexplored Players in Ecosystem Carbon Processing. *MSystems*, 3(5).

<https://doi.org/10.1128/mSystems.00076-18>

Tsuboi, M., Benevides, J. M., Bondre, P., & Thomas, G. J. J. (2005). Structural details of the thermophilic filamentous bacteriophage PH75 determined by polarized Raman microspectroscopy. *Biochemistry*, 44(12), 4861–4869.

<https://doi.org/10.1021/bi0479306>

Turner, D., Shkoporov, A. N., Lood, C., Millard, A. D., Dutilh, B. E., Alfenas-Zerbini, P., van Zyl, L. J., Aziz, R. K., Oksanen, H. M., Poranen, M. M., Kropinski, A. M., Baryliski, J., Brister, J. R., Chanisvili, N., Edwards, R. A., Enault, F., Gillis, A.,

- Knezevic, P., Krupovic, M., ... Adriaenssens, E. M. (2023). Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. *Archives of Virology*, 168(2), 74. <https://doi.org/10.1007/s00705-022-05694-2>
- Van Etten, J. L., Lane, L. C., & Dunigan, D. D. (2010). DNA viruses: the really big ones (giruses). *Annual Review of Microbiology*, 64, 83–99. <https://doi.org/10.1146/annurev.micro.112408.134338>
- Van Helvoort, T. (1994). History of virus research in the twentieth century: the problem of conceptual continuity. *History of Science*, 32(2), 185–235. <https://doi.org/10.1177/007327539403200204>
- Vasanthakumari, R. (2009). *Practical microbiology BI publications Pvt Ltd*.
- Vervliet, G., Holsters, M., Teuchy, H., Van Montagu, M., & Schell, J. (1975). Characterization of different plaque-forming and defective temperate phages in *Agrobacterium*. *The Journal of General Virology*, 26(1), 33–48. <https://doi.org/10.1099/0022-1317-26-1-33>
- Vijayan, J., Nathan, V. K., Ammini, P., & Ammanamveetil, A. M. H. (2023). Bacterial diversity in the aquatic system in India based on metagenome analysis—a critical review. *Environmental Science and Pollution Research*, 30(11), 28383–28406. <https://doi.org/10.1007/s11356-023-25195-2>
- V.K. Saxena, M. L. G. (1987). Evaluation of reservoir temperatures and local utilization of geothermal waters of the Konkan Coast, India. *Journal of Volcanology and Geothermal Research*, 33, 337–342.
- Waldbauer, J. R., Coleman, M. L., Rizzo, A. I., Campbell, K. L., Lotus, J., & Zhang, L. (2019). Nitrogen sourcing during viral infection of marine cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 116(31), 15590–15595. <https://doi.org/10.1073/pnas.1901856116>
- Wang, F., Xiao, Y., Xiong, Y., Jiao, Y., Zhang, Q., & Lin, L. (2020). Complete genome sequence of MMP7, a novel *Meiothermus* bacteriophage of the family Myoviridae isolated from a hot spring. *Archives of Virology*, 165(3), 753–756. <https://doi.org/10.1007/s00705-019-04462-z>
- Wang, F., Xiong, Y., Xiao, Y., Han, J., Deng, X., & Lin, L. (2020). MMPphg from the thermophilic *Meiothermus* bacteriophage MMP17 as a potential antimicrobial

- agent against both Gram-negative and Gram-positive bacteria. *Virology Journal*, 17(1), 1–10. <https://doi.org/10.1186/s12985-020-01403-0>
- Watson, B. N. J., Steens, J. A., Staals, R. H. J., Westra, E. R., & van Houte, S. (2021). Coevolution between bacterial CRISPR-Cas systems and their bacteriophages. *Cell Host & Microbe*, 29(5), 715–725. <https://doi.org/10.1016/j.chom.2021.03.018>
- Weinbauer, M. G. (2004). Ecology of prokaryotic viruses. *FEMS Microbiol Rev*, 28(2), 127–181. <https://doi.org/10.1016/j.femsre.2003.08.001>
- Weiss, M. C., Sousa, F. L., Mrnjavac, N., Neukirchen, S., Roettger, M., Nelson-Sathi, S., & Martin, W. F. (2016). The physiology and habitat of the last universal common ancestor. *Nature Microbiology*, 1(9), 16116. <https://doi.org/10.1038/nmicrobiol.2016.116>
- Wendling, C. C., Refardt, D., & Hall, A. R. (2021). Fitness benefits to bacteria of carrying prophages and prophage-encoded antibiotic-resistance genes peak in different environments. *Evolution; International Journal of Organic Evolution*, 75(2), 515–528. <https://doi.org/10.1111/evo.14153>
- White, R. H. (1991). Distribution of folates and modified folates in extremely thermophilic bacteria. *Journal of Bacteriology*, 173(6), 1987–1991. <https://doi.org/10.1128/jb.173.6.1987-1991.1991>
- Wick, R. R., Schultz, M. B., Zobel, J., & Holt, K. E. (2015). Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics*, 31(20), 3350–3352. <https://doi.org/10.1093/bioinformatics/btv383>
- Wishart, D. S., Han, S., Saha, S., Oler, E., Peters, H., Grant, J. R., Stothard, P., & Gautam, V. (2023). PHASTEST : faster than PHASTER , better than PHAST. *Nucleic acids research*, 51(W1), W443–W450. <https://doi.org/10.1093/nar/gkad382>
- Wolf, Y. I., Silas, S., Wang, Y., Wu, S., Bocek, M., Kazlauskas, D., Krupovic, M., Fire, A., Dolja, V. V., & Koonin, E. V. (2020). Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome. *Nature Microbiology*, 5(10), 1262–1270. <https://doi.org/10.1038/s41564-020-0755-4>
- Wommack, K. E., Sime-Ngando, T., Winget, D. M., Jamindar, S., & Helton, R. R. (2010). *Manual of Aquatic Viral Ecology (MAVE)*.
- Wong, B. T., & Lee, D. J. (2014). *Pseudomonas yangmingensis* sp. nov., an alkaliphilic denitrifying species isolated from a hot spring. *Journal of Bioscience and*

- Bioengineering*, 117(1), 71–74. <https://doi.org/10.1016/j.jbiosc.2013.06.006>
- Yang, Y., Lu, S., Shen, W., Zhao, X., Shen, M., Tan, Y., Li, G., Li, M., Wang, J., Hu, F., & Le, S. (2016). Characterization of the first double-stranded RNA bacteriophage infecting *Pseudomonas aeruginosa*. *Scientific Reports*, 6, 38795. <https://doi.org/10.1038/srep38795>
- Yoon, S.-H., Ha, S.-M., Kwon, S., Lim, J., Kim, Y., Seo, H., & Chun, J. (2017). Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *International Journal of Systematic and Evolutionary Microbiology*, 67(5), 1613–1617. <https://doi.org/10.1099/ijsem.0.001755>
- Yoshida, M., Mochizuki, T., Urayama, S.-I., Yoshida-Takashima, Y., Nishi, S., Hirai, M., Nomaki, H., Takaki, Y., Nunoura, T., & Takai, K. (2018). Quantitative Viral Community DNA Analysis Reveals the Dominance of Single-Stranded DNA Viruses in Offshore Upper Bathyal Sediment from Tohoku, Japan. *Frontiers in Microbiology*, 9, 75. <https://doi.org/10.3389/fmicb.2018.00075>
- Yu, M. X., Slater, M. R., & Ackermann, H. W. (2006). Isolation and characterization of *Thermus* bacteriophages. *Arch Virol*, 151(4), 663–679. <https://doi.org/10.1007/s00705-005-0667-x>
- Zablocki, O., Kirby, B., Trindade, M., Corinaldesi, C., Tangherlini, M., Dell’anno, A., Zablocki, O., van Zyl, L. J., Kirby, B., & Trindade, M. (2017). Diversity of dsDNA viruses in a South African hot spring assessed by metagenomics and microscopy. *Viruses*, 9(11), 1–12. <https://doi.org/10.1038/s41598-017-08783-4>
- Zablocki, O., van Zyl, L., & Trindade, M. (2018). Biogeography and taxonomic overview of terrestrial hot spring thermophilic phages. *Extremophiles*, 22(6), 827–837. <https://doi.org/10.1007/s00792-018-1052-5>
- Zaragoza-Solas, A., Rodriguez-Valera, F., & López-Pérez, M. (2020). Metagenome Mining Reveals Hidden Genomic Diversity of Pelagimyophages in Aquatic Environments. *MSystems*, 5(1). <https://doi.org/10.1128/mSystems.00905-19>
- 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 Research*, 3(1), 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 Journal*, 10(12), 2787–2800. <https://doi.org/10.1038/ismej.2016.85>
- Zhao, Y., Temperton, B., Thrash, J. C., Schwalbach, M. S., Vergin, K. L., Landry, Z. C., Ellisman, M., Deerinck, T., Sullivan, M. B., & Giovannoni, S. J. (2013). Abundant SAR11 viruses in the ocean. *Nature*, 494(7437), 357–360. <https://doi.org/10.1038/nature11921>
- Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J., & Wishart, D. S. (2011). PHAST: a fast phage search tool. *Nucleic Acids Research*, 39(suppl_2), W347-52. <https://doi.org/10.1093/nar/gkr485>
- Zimmerman, A. E., Howard-Varona, C., Needham, D. M., John, S. G., Worden, A. Z., Sullivan, M. B., ... & Coleman, M. L. (2020). Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. *Nature Reviews Microbiology*, 18(1), 21-34. <https://doi.org/10.1038/s41579-019-0270-x>

APPENDIX

- **Saline magnesium Buffer**

Saline magnesium buffer Mix 100 mM NaCl, 8.1 mM MgSO₄, 50 mM Tris-HCl (pH 7.5) and 0.01% gelatin (w/v). Keep sterile and store at room temperature for a maximum of 1 month.

- **Tris-EDTA (TE) Buffer**

1 M Tris-Cl: Dissolve 121.1 g of Tris base in 800 ml of H₂O. Adjust the pH to the desired value by adding concentrated HCl. Adjust the volume of the solution to 1 L with H₂O. Dispense into aliquots and sterilize by autoclaving.

0.5 M EDTA: Add 186.1 g of disodium EDTA•2H₂O to 800 ml of H₂O. Stir vigorously on a magnetic stirrer. Adjust the pH to 8.0 with NaOH. Dispense into aliquots and sterilize by autoclaving

To prepare 100 ml of TE buffer:

Reagent	Amount to add	Final concentration
Tris-Cl (1 M, pH 8.0)	1 ml	10 mM
EDTA (0.5 M, pH 8.0)	200 µl	1 mM
H ₂ O	98.8 ml	

Store at room temperature.

- **Tris Borate EDTA (TBE) Buffer:**

For a 5X stock solution:

Tris base	54.0 g
Boric acid	27.5 g
0.5 M EDTA (pH 8.0)	20 ml
H ₂ O	to 1 L

Adjust the pH to 8.3. Pass the stock solution through a 0.22 µm-filter and store at roomtemperature.

- **Cetyltrimethylammonium bromide:** Add 1 g of NaCl to 15 ml of water. Slowly add 2.5 g of CTAB. If necessary, heat to 65 °C until dissolved. Purify through a 0.02-mm filter. Keep sterile and store at room temperature for a maximum of 1 month. 2 M Tris/0.2 M EDTA, pH 8.5 (TE) Add 24.228 g of Tris base, 40 ml of 0.5 M EDTA (pH 8.0) and 30 ml of H₂O; adjust the pH to 8.5 and add water to make up a final volume of 100 ml.

PRESENTATIONS AT CONFERENCES

- Paper entitled “*Metavirome analysis of two hot springs in Maharashtra, India.*” Presented at the international conference **Viruses of Microbes 2024** (VoM2024), **Cairns, Australia**, 15-19 July 2024.
- Paper entitled " *Genomic characterization of a lysogenic bacteriophage of Pseudomonas sp. isolated from geothermal spring of India*" presented at the international conference **Viruses of Microbes 2022** (VoM2022), **Guimarães, Portugal**: 18-22 July 2022.
- Paper entitled "*Prophage induction by mitomycin C treatment of hot spring bacteria Pseudomonas sp. strain VB3*" presented at **VIROCON 2016**, The International Conference on "*Global Perspectives In Virus Diseases Management*", Indian Institute of Horticulture Research, Bengaluru, India: 8-10 December 2016.
- Paper entitled "*Biochemical and Molecular Phylogenetic Analysis of Thermotolerant Anoxybacillus sp. from the hot spring of Tural, Maharashtra*" presented at International Seminar on "*New Frontiers in Microbiology and Applied Biology*", Department of Microbiology, St Xavier's College, Mapusa, Goa on 7-8 Jan 2016.

AWARDS

- **Won the Best Poster Award (Second Place)** for this poster entitled "Prophage induction by mitomycin C treatment of hot spring bacteria *Pseudomonas* sp. strain Vb3" at VIROCON 2016, IIHR, Bengaluru, India.
- Selected among the 25 candidates to attend the Second South Asian Thematic School in Aquatic Microbial Ecology (**ASIAME2**) workshop held at NIOT in Port Blair, India, from November 6 to 18, 2017. Funded by CNRS, DBT, IRD, CEFIPRA, MOORE Foundation, and NIOT.
- **Early career scientist-FEMS** grant for the VoM2022, at Guimarães, Portugal.
- **Early career scientist-ISVM** travel grant for the VoM2022, at Guimarães, Portugal.
- **DBT-CTEP travel grant** for VoM2022, Guimarães, Portugal.
- **Travel scholarship for VOM2024**, at Cairns, Australia.

PUBLICATIONS

- **Faria, N., & Ghadi, S. C. (2023).** Draft genome sequence of lysogenic *Pseudomonas* sp. strain VB3 isolated from Vashisht hot spring, India. *Biologia*, 78(11), 3241-3248.
- **Faria, N., & Ghadi, S. C. (2021).** Isolation of a thermophilic *Meiothermus* sp. strain JE2 and its infective phage vB_Msp_JE2 from Rajwadi hot spring, India. *Indian Journal of Biotechnology*, NIScPR-CSIR, India, Vol. 20, 2021, pp 355-363.

Under Communication

- "Isolation and genomic characterization of inducible temperate bacteriophage of *Pseudomonas* strain VB3. from Vashisht hot spring, India. Nicola Faria and Sanjeev C Ghadi, School of Biological Sciences and Biotechnology, Goa University.
- "Metaviromics analysis of two hot springs belonging to the West coast geothermal province in Maharashtra, India". Nicola Faria and Sanjeev C Ghadi, School of Biological Sciences and Biotechnology, Goa University.



Isolation of a thermophilic *Meiothermus* sp. strain JE2 and its infective phage vB_Msp_JE2 from Rajwadi hot spring, India

Nicola Faria and Sanjeev C. Ghadi*

Department of Biotechnology, Goa University, Taleigao Plateau, Goa- 403206, India

Received 23 September 2021; revised & accepted 30 September 2021

A thermophilic, red-pigmented aerobic bacteria designated as strain JE2 was isolated from the water sample of the Rajwadi hot spring. Bacterial strain JE2 was Gram-negative, rod-shaped that grew optimally at pH 7.5 and 60°C. Molecular identification of the 16S rRNA sequence indicated that bacterial strain JE2 formed a coherent cluster with genus *Meiothermus*. Subsequently, a lytic bacteriophage designated as vB_Msp_JE2 was isolated from the Rajwadi hot spring that could be propagated using *Meiothermus* sp. strain JE2. One-step growth curve using *Meiothermus* sp. strain JE2 as host demonstrated a latent period of 70 min and a burst size of 20. Bacteriophage vB_Msp_JE2 was stable at pH 6.0 - 8.0. Additionally, the phage was thermostable at 60°C and was completely inactivated above 75°C. Since, most microbial studies have been reported predominantly from the Himalayan geothermal province, the present study reports the association of a unique culturable pigmented thermophilic bacterial host and bacteriophage from Rajwadi hot spring present in the West coast geothermal province, India.

Keywords: Hot spring, thermophilic bacteria, bacteriophage, *Meiothermus* sp.

Introduction

The Geological Survey of India has reported seven geothermal provinces covering nearly 400 thermal springs based on the tectonic elements and geothermal gradients across India¹⁻². The west coast geothermal province (WCGP) of Maharashtra, India, comprises nearly 60 thermal springs grouped around 18 locations with a temperature ranging from 42°C at Rajapur to 71°C at Unhaver³⁻⁴. The Rajwadi hot spring is located in the Ratnagiri districts of Maharashtra in the southern part of WCGP⁵. The hot spring is confined to the N20°W - S20°E fracture zone and is situated at the banks of the tributary joining Shastri river and located 10 km north of Sangmeshwar belonging to WCGP of India.

The hot spring is a unique extremophilic site that harbours several thermophilic bacteria and bacteriophages. Thermophilic microorganisms and their phages have been isolated from hot springs⁶⁻⁸ hydrothermal vents⁹, compost pile¹⁰ and salted geothermal sites¹¹. Although, several reports on isolation of thermophilic bacteria from hot springs have been reported worldwide¹²⁻¹⁴, isolation of thermophilic phage from hot springs remains

untapped primarily¹⁵. Most of the reports on hot springs phages concern thermophilic archaeal phages¹⁶⁻¹⁷ and a few from the genus of *Thermus* bacteriophages¹⁸⁻²⁰.

Meiothermus sp. are moderately thermophilic, aerobic bacteria belonging to the family *Thermaceae* of the phylum *Deinococcus-Thermus*²¹. *Meiothermus* species are widely found in hot springs, occurring through a wide range of pH and temperatures²². The optimal growth temperature of *Meiothermus* ranges from 50 - 65°C and pH 7 - 8.5. Some of the bacteria present in thermal springs were also dominant species in hot fermenters fed with wastewater²², paper mill machines²³⁻²⁴ and storage systems for spent nuclear fuel²⁵. The optimal growth temperature of *Meiothermus* ranges from 50 - 65°C and pH 7 - 8.5, which is also common in paper machine water systems²⁴. Biofouling can cause complications such as filter clogging, microbial-induced corrosion, and lowered heat conductivity, and the cost of controlling biofouling (biocide, cleaners, dispersants) would be an expensive affair. Currently, only two lytic *Meiothermus* phages are reported from two hot springs in Yunnan Province of China²⁶⁻²⁷. Bacteriophage-based anti-biofouling strategies might possess unique properties and considerable promise in controlling biofilm formation on solid surfaces in an eco-friendly manner.

* Author for correspondence:
saga@unigoa.ac.in
Phone: +91-8669609242

Though several thermophilic bacteria have been reported from thermal hot springs of India²⁸, no infective hot spring phage for any bacterial host has been described. Additionally, although previous studies report the isolation of *Meiothermus* sp. from Indian hot springs²⁹⁻³¹ but present study reports the isolation of *Meiothermus* phage from the Rajwadi hot spring. Most of the microbial diversity studies of Indian hot springs are focused on Himalayan geothermal provinces present at higher temperatures and altitudes²⁸. Maharashtra is one such bio-diverse state of India, which harbors 60 diverse alkaline hot springs in the WCGP of India, which is relatively unexplored for thermophilic bacteria and viruses. Hence, the Rajwadi hot spring was selected for the screening of thermophilic bacteria and their possible potential infective bacteriophage.

Materials and Methods

Site Description and Sample Collection

Water samples were collected from the Rajwadi hot spring (17°14'23.1"N 73°33'42.2"E) in the Ratnagiri district of Maharashtra (Fig. 1). The water samples were collected using sterile bottles from the sprout of the spring where there were no human activities or interference. The samples collected were

transported within 12 h to the laboratory for further studies.

Physicochemical Analysis of Hot Spring Water Sample

In situ physicochemical parameters such as temperature, pH, electrical conductivity and total dissolved salt (TDS) of Rajwadi hot spring water were recorded using multi-parameter (Cyber Scan PC 650, Eutech Instruments, Thermo Fisher Scientific, India). The elemental analyses of the Rajwadi hot spring water sample were performed by inductive coupled plasma atomic emission spectrophotometry (ICP-AES, SPECTRO Analytical instrument, GmbH, Germany). The collected water samples were immediately used for the isolation of thermophilic bacteria and screening of bacteriophages.

Isolation and Biochemical Characterization of Thermophilic Bacteria

For the isolation of thermophilic bacteria, water sample was spread plated on basal media (BM) (g/l) (KH_2PO_4 : 1.36, $(\text{NH}_4)_2\text{SO}_4$:1.0, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$: 0.2, FeSO_4 : 0.001, NaCl : 2.0, yeast extract, 1.0 and supplemented with 0.2% sodium succinate)³². Briefly, 100 μl of the water sample was spread plated on solid gelrite plates and incubated at 60°C in sealed plastic bags. After incubating the gelrite plates for 3 - 5 days at

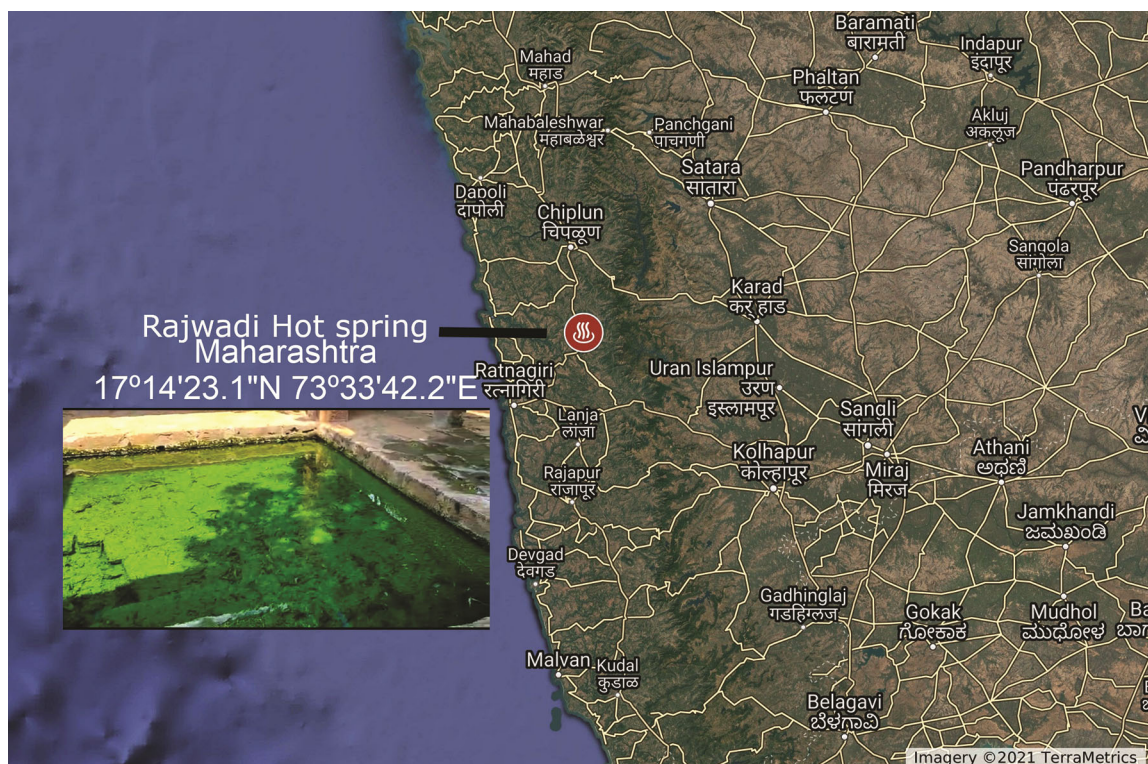


Fig. 1 — Geographical location of Rajwadi hot spring and sampling sites. The given map was retrieved from Google Maps (2021).

60°C, a bacterial strain that was morphologically different were picked and purified on a fresh plate. The bacterial strain was preserved in a basal medium (BM) containing 15% glycerol at -80°C. The colony morphology and the Gram staining of the thermophilic bacterial strain were recorded. The bacterial strain was further characterized by biochemical and growth (temperature and pH tolerance) characteristics. For determining the growth temperature range, the strains were incubated at varying temperatures. The carbon utilization pattern of the bacterial strains was assessed usingTM. The bacterial isolate was also evaluated for various biochemical tests for catalase, oxidase, nitrate reduction and casein hydrolysis. Screening for industrially important enzymes like amylase, cellulase, xylanase, protease and lipase activity as per standard protocols were also carried³³⁻³⁴.

DNA Extraction and 16S rRNA Sequencing

Genomic DNA from bacterial strain JE2 were isolated by GeneiPureTM bacterial DNA purification kit. The eluted DNA was stored at -20°C. Later, isolated DNA was quantified using the Qubit dsDNA HS Assay Kit and the Qubit 2.0 Fluorometer (Invitrogen, USA) as per the manufacturer's instruction. Amplification of 16S rRNA from the bacterial isolate was performed using universal primers 27F and 1525R³⁵. The amplification was conducted in a 50 µL reaction mixture containing approximately 100 ng genomic DNA of bacterial strain JE2, 10X polymerase buffer, 5 µL dNTP mix, 0.25 µM of each forward and reverse primer, 1.5 Units of *Taq* polymerase. The reaction mixtures were heated at 95°C for 5 min and subjected to 30 cycles of denaturation at 95°C for 1 min, annealing at 55°C for 1 min, elongation at 72°C for 1 min and final elongation step at 72°C for 10 min. The amplified PCR product was resolved on 0.8% agarose gel to observe the integrity of the PCR amplified product. The amplified PCR product was purified using a QIA quick PCR purification kit (Qiagen Inc, Hilden, Germany). The amplified PCR product was sequenced by ABI 3130 (48 capillary) or 3730XI (96 capillary) electrophoresis instruments (Applied Biosystems, Foster City, CA, USA). Basic Local Alignment Search Tool (BLAST) analysis of 16S rRNA sequence of bacterial isolates with other sequences in the database was performed (<https://blast.ncbi.nlm.nih.gov/>). Homologous sequences that were similar to bacterial isolate were retrieved

from the ribosomal database project (RDP) (<http://rdp.cme.msu.edu/>). Multiple sequence alignment of 16S rDNA sequence was carried out using the ClustalW program. A phylogenetic tree was constructed by the maximum likelihood method using MEGA-X software v.10.2.6. Bootstrap analysis was done with 1000 replication using the bootstrap option in MEGA-X software. The 16S rRNA sequence was deposited to GenBank and was assigned National Center of Biotechnology Information (NCBI) accession number MW186236.

Filtration and Concentration of Water Sample for Virus Enrichment

Approximately 3 L of hot spring water was filtered and concentrated for virus enrichment. The water sample was centrifuged at 3000 x g for 10 min and filtered through 0.22 µm pore-size polyvinylidene fluoride (PVDF) filters (Millipore, USA). The filtrate was stored at 4°C until further downstream processing. To this filtrate, 10% (w/v) polyethylene glycol-8000 and 1 M NaCl was added to precipitate the viral particles³⁷. This mixture was incubated at 4°C overnight and centrifuged at 11,000 x g for 20 min. The viral particles precipitated were suspended in 50 ml saline magnesium (SM) buffer: 100 mM NaCl, 8.1 mM MgSO₄, 50 mM Tris, 0.01% gelatin and pH 7.5. Virus concentration by ultracentrifugation was carried out in a tabletop ultracentrifuge Optima MAX-XP using an MLS-50 rotor (Beckman Coulter, United States)³⁸. The viral sample was ultracentrifuged at 80,000 x g for 4 h at 4°C under vacuum conditions. The supernatant was discarded and viral particles were re-suspended in 10 ml of SM buffer and left overnight.

Bacteriophage Isolation, Infection and Purification

Bacteriophages were screened against the thermophilic bacterial strain JE2 by double-layer overlay technique³⁹⁻⁴⁰ after enriching the concentrated viral sample with 10X BM broth. Briefly, the 10 ml viral concentrate from the hot spring water was added to 100 ml BM broth inoculated with bacterial strain JE2 and incubated at 60°C for 24 h. This enriched culture was centrifuged at 10,000 x g for 15 min and filtered through a 0.22 µm pore-size PVDF filter (Millipore, USA) and stored as phage stock at 4°C.

Bacteriophage infecting the thermophilic bacterial strain JE2 was isolated by plating on semi-solid gelrite BM medium using double-layer overlay

technique. The BM medium was supplemented with 10 mM CaCl_2 for bacteriophage propagation. Briefly, in a 2 mL micro-centrifuge tube, 400 μL of filtered enriched phage stock and 800 μL of logarithmic growing bacterial strain JE2 culture were mixed and incubated for 20 min at room temperature for adsorption. This mixture was then added to a 4 mL top-gelrite medium (0.4%) and poured onto the culture plate with a bottom-gelrite layer (0.8%). After 48 h, clear plaques were observed at 60°C. Single clear plaques were picked up and suspended in SM buffer and kept at 4°C. This was followed by centrifugation at 10,000 x g for 10 min and subsequently, the phage supernatant was passed through 0.22 μm pore-size filters to remove any bacterial debris. Spot assay was also carried out on the lawn of bacterial strain JE2 to check the infectivity of the isolated phage particles. Single plaques were excised aseptically in SM buffer and used to reinfect bacterial strain JE2 to obtain plaques. After three rounds of bacterial infection, the phage from one such purified plaque was designated as vB_Msp_JE2 and used for further studies.

One-Step Growth Curve

For a one-step growth curve, phage particles were mixed with 1 ml of fresh *Meiothermus* sp. strain JE2 (1×10^6 CFU/mL) to obtain a multiplicity of infection (MOI) of 0.01 and incubated for 10 min at 60°C⁴¹. After absorption, the mixture was centrifuged at 13000 x g for 2 min. The pellet was suspended in 1ml fresh BM and this step was repeated thrice. To 50 ml fresh BM, 50 μL of the suspended pellet was added and incubated at 60°C. Phage titer was assayed by double-layer overlay technique at every 10-min interval throughout 2 h to estimate the number of phages as plaque-forming units (PFU).

Phage Stability Analysis

The thermostability of the phage was determined by pre-incubating the phage in the BM medium at 50°C - 80°C for 30 min. Subsequently, the survival of the phage sample was also assayed by the

double-layer overlay technique. To assess the pH sensitivity, the phage (10^6 PFU/ml) was diluted 100-fold with SM buffer of different pH values and incubated for 30 min. The double-layer overlay technique was used to determine the rate of survival of the phage.

Results and Discussion

Physicochemical Analysis of Rajwadi Hot Spring Water

The hot springs in WCGP of Maharashtra are manifested in the hard rock of Deccan basalt terrain. The hot water emerges from the margins of dykes, intense fractures and shears. Intermittent gas emissions with sulphurous smell were noticed at the sprout of the Rajwadi hot spring. The *in situ* measured temperature and pH of the Rajwadi hot spring water were 55°C and pH 7.5, respectively. Most of the hot springs in India have a temperature ranging between 40°C to 98°C. Hot springs of WCGP have temperatures ranging from 42°C at Rajapur to 71°C at Unhaver. The pH of Indian hot spring water is neutral, moderately alkaline, or alkaline in nature and harbours a microbial community different from the acidic hot springs across the world⁴²⁻⁴³. Table 1 depicts the physicochemical parameters of Rajwadi hot spring water analyzed using ICP-AES. The analysed physicochemical parameters of Rajwadi hot spring water samples validated previous observations reported elsewhere⁴⁴. The Ca concentration in Rajwadi hot spring was similar to the values reported previous study. Previous physicochemical studies of WCGP thermal springs revealed that chloride content varies from 78 to 4800 ppm, sodium 76 to 410 ppm, with low magnesium, boron and fluoride content⁴⁵⁻⁴⁷. In general, the spring was indicative of Na-Ca-Cl or Ca-Na-Cl spring which is the most case for thermal springs of WCGP. Elements such as Mn, Cu, Fe, Zn and Pb were detected in acidic geothermal spring and were almost below the limit of quantification in the water sample collected from the Rajwadi hot spring. In a study of alkaline hot springs in Yellowstone National Park

Table 1 — Physicochemical parameters of hot spring water (Rajwadi, Maharashtra).

<i>In situ</i> temperature (°C)	pH	Conductivity ($\mu\text{S}/\text{cm}$)	TDS	Na^+	Ca^{2+}	Mg^+	Cl^-	K^+	S	Fe	Cu	Mn	Pb	Zn
55	7.5	1445	956	230	53.3	0.498	354	7.5	36.25	0.01	*	*	*	*

*: (Not detected) < 0.01 ppm

(63°C), the hot spring had elevated levels of sodium, chloride, fluoride and heavy metals such as Pb, Cr, Cu, Hg, etc were below the limit of quantification⁴⁸. Since the spring is moderately alkaline in nature, these elements were below the limit of quantification.

Isolation, Biochemical and Molecular Characterization of the Bacterial Isolate

During the preliminary screening of the Rajwadi hot spring, a thermophilic red pigment-producing bacterium was isolated on BM and designated as bacterial strain JE2. The bacteria formed smooth, convex, circular colonies with entire margin (Fig. 2). Strain JE2 was obligate thermophile, Gram-negative and non-motile rods. Bacterial strain JE2 grew optimally at 60°C and pH 7.5. No growth was observed below 50°C and above 70°C (Fig. 3). The bacterial strain JE2 was catalase and oxidase positive. Further, the bacterial strain JE2 utilized various substrates such as D-lactose, D-fructose, D-galactose, D-mannose, maltose, D-cellobiose, D-raffinose, D-sorbitol, D-mannitol, and glycerol (Table 2). Several biochemical characteristics of bacterial strain

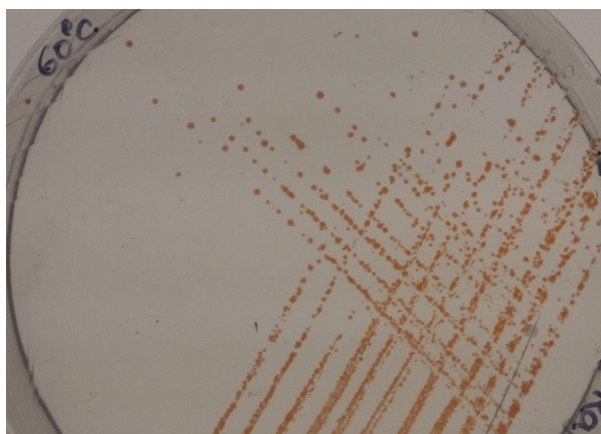


Fig. 2 — Bacterial strain JE2 isolated from Rajwadi hot spring, Maharashtra

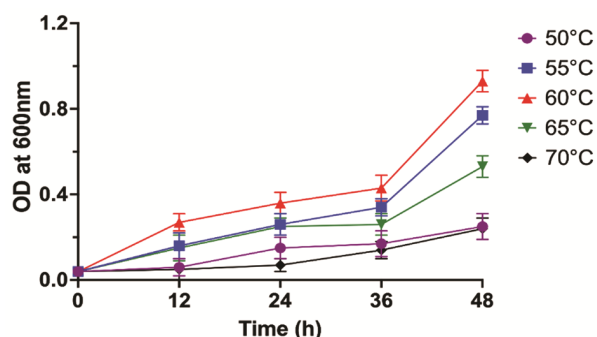


Fig. 3 — Growth of thermophilic *Meiothermus* sp. strain JE2 at different temperatures.

JE2 were similar to the other type strains of the genus *Meiothermus*⁴⁹.

16S rRNA Sequencing

BLAST analysis of strain JE2 16S rRNA sequence (MW186236) with 16S rDNA sequences of type strains from the RDP demonstrated 99% similarity with *Meiothermus ruber* and 98% similarity with

Table 2 — Biochemical characteristics of *Meiothermus* sp. strain JE2

Characteristics	<i>Meiothermus</i> sp. Strain JE2
Pigmentation	Red
Motility	Non-motile
Reduction of NO ₃ ⁻ to NO ₂ ⁻	-
Presence of	
Catalase	+
Oxidase	+
Hydrolysis	
Starch	-
Casein	+
Xylan	-
Utilization of	
Lactose	+
Xylose	+
Maltose	+
Fructose	+
Dextrose	+
Galactose	+
Raffinose	+
Trehalose	+
Melibiose	+
Sucrose	+
L-Arabinose	-
Mannose	+
Inulin	+
Sodium gluconate	+
Glycerol	+
Salicin	-
Dulcitol	-
Inositol	+
Sorbitol	+
Mannitol	+
Adonitol	+
Arabitol	+
Erythritol	+
alpha-Methyl-D-glucoside	+
Rhamnose	-
Cellobiose	+
Melezitose	-
alpha-Methyl-D-Mannoside	-
Xylitol	+
Ortho-nitrophenyl β-galactoside (ONPG)	+
Esculin	+
D-arabinose	-
Citrate	-
Malonate	+
Sorbose	+

Symbols: +, positive, - negative results

M. cateniformans, *M. cerbereus* and *M. taiwanensis*, respectively. Phylogenetic tree constructed using the sequence of closely related bacterial type strains by Maximum likelihood method indicated that the strain JE2 formed a coherent cluster with the clade that comprises the type strains of genus *Meiothermus* (Fig. 4). Strain JE2 is closely related to *Meiothermus ruber* (ATCC35948) and was identified as *Meiothermus* sp. strain JE2. *Meiothermus* sp. are slightly thermophilic, aerobic bacteria which belongs to phylum *Deinococcus-Thermus*. Species of the genus *Meiothermus* formerly belonged to the genus *Thermus* and are different from the species of *Thermus* by their ability to grow at lower temperatures. The optimum growth temperature of *Meiothermus* species is the range of 50-65°C and can form red- or yellow pigment colonies. *Meiothermus* species are widely found in hot springs and artificial thermal environments. *Meiothermus* sp. strain JE2 isolated from Rajwadi hot spring was compared with other two red-pigment producing *Meiothermus* species isolated from Paniphala hot spring, India (Table 3). Several culture-dependent and-independent studies⁵⁰⁻⁵¹

from Indian hot springs also demonstrated genus *Meiothermus* in these extreme geothermal environments. The genus *Meiothermus* was also dominant in metagenomic samples with high temperatures (66-67°C)⁵⁰⁻⁵¹. These results indicated that temperature is an essential factor for the growth of this organism in these thermal springs.

Bacteriophage Isolation, Infection and Purification

After filtration of the water sample through 0.22 µm pore-size filter and subsequent concentration by ultra-centrifugation, the water sample was screened for potential bacteriophage that would use *Meiothermus* sp. strain JE2 as host. A lytic bacteriophage designated as vB_Msp_JE2 was isolated from Rajwadi hot spring. The bacteriophage vB_Msp_JE2 produced a clear plaque of a 1 - 3 mm diameter after 48 h at 60°C (Fig. 5). Currently, there are only two lytic thermophilic *Meiothermus* phages MMP17 and MMP7 that are reported to date from Eryuan and Tengchong hot springs in Yunnan Province of China, respectively²⁶⁻²⁷.

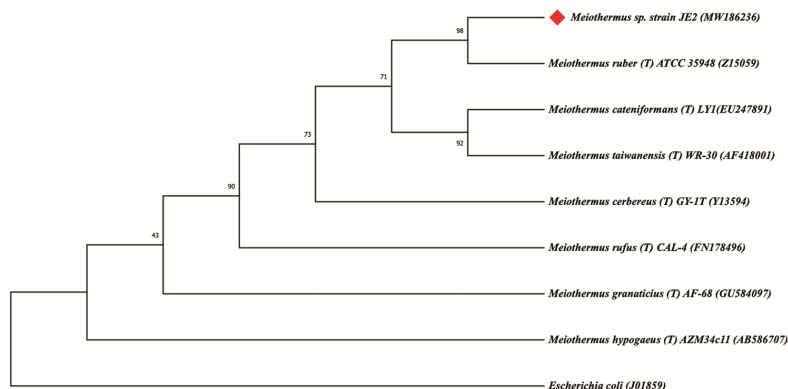


Fig. 4 — Molecular phylogenetic analysis of *Meiothermus* sp. strain JE2 by maximum likelihood method. Bootstrap analysis was done with 1000 replication using the bootstrap option in MEGA X. *Escherichia coli* was selected as the out-group member.

Table 3 — Characteristics of *Meiothermus* sp. strain JE2 in comparison with other reported *Meiothermus* species isolated Indian hot springs.

Characteristics	<i>Meiothermus</i> sp. JE2 (This study)	<i>Meiothermus</i> sp. RP ³¹	<i>Meiothermus</i> sp. TP ³¹
Growth temperature range (°C)	50-70	55-70	55-70
Optimal growth temperature (°C)	60	60	60
Pigment	Red	Red	Red
Geographic location	Rajwadi, Maharashtra, India	Paniphala hot spring, West Bengal, India	Paniphala hot spring, West Bengal, India
Casein hydrolysis	-	+	+
Amylase	-	+	+
Xylanase	-	+	+
Cellulase	+	+	+
Lipase	-	-	-

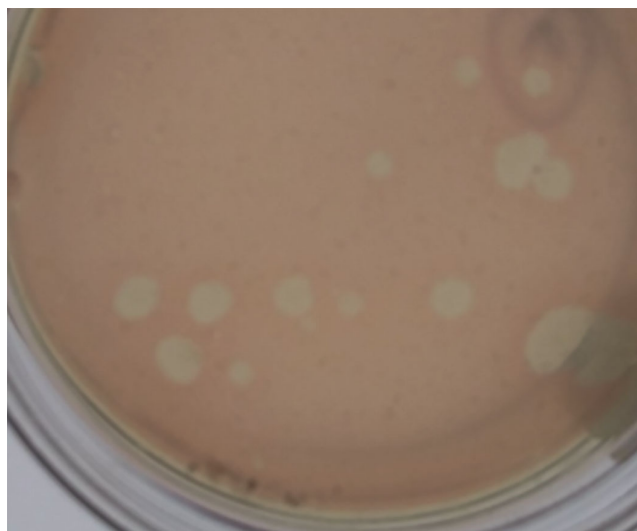


Fig. 5 — Bacteriophage vB_Msp_JE2 clear plaques detected by the double-layer overlay method using the red-pigmented thermophilic *Meiothermus* sp. strain JE2 as host.

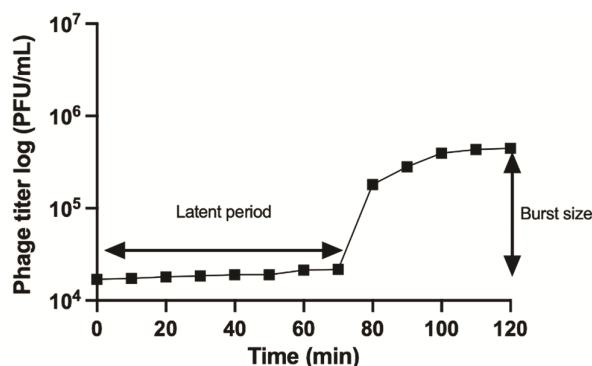


Fig. 6 — One-step growth curve of bacteriophage vB_Msp_JE2.

One-Step Growth Curve

The latent period and burst size of phage vB_Msp_JE2 were determined from a one-step growth curve experiment. The latent period of the phage was 70 min, and the burst size was 20 (Fig. 6). Similar, long latent periods in the range of 60 - 80 min were also reported for other thermophilic *Meiothermus* and *Thermus* phages⁵². Additionally, the burst size of 20 for the phage isolated in the present study was almost similar to that reported for *Meiothermus* phage MMP17²⁶. In contrast, other thermophilic *Thermus* phages stated a larger burst size of approximately 80 for ϕ YS40, 109 for ϕ OH3 and 200 for TSP4⁵³⁻⁵⁵.

Thermostability and pH Sensitivity

Phage vB_Msp_JE2 was stable at 60°C (Fig. 7). *Meiothermus* phage MMP17²⁶ was stable between 55 to 60°C whereas *Thermus* phage TSP4⁵³ and *Geobacillus* phage GVE1⁹ were stable at 60°C. In

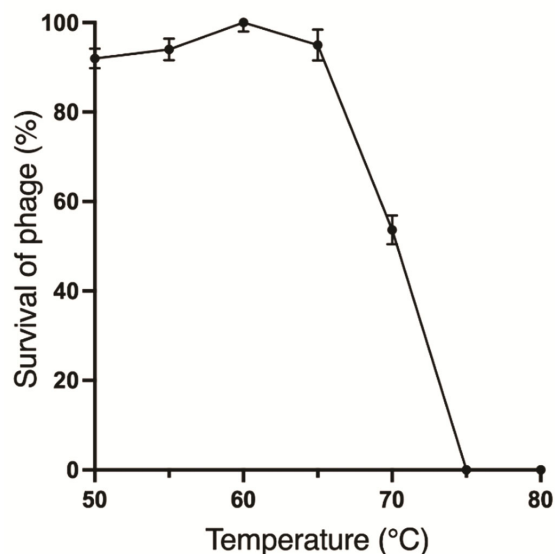


Fig. 7 — Thermostability of bacteriophage vB_Msp_JE2

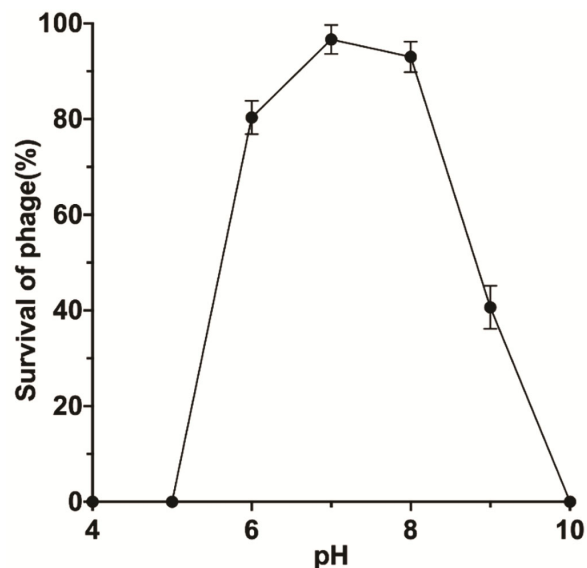


Fig. 8 — pH sensitivity of bacteriophage vB_Msp_JE2

contrast, *Thermus* phages such as ϕ OH3⁵⁵ and ϕ YS40⁵⁴ were stable at 70°C. The thermostability of the phage vB_Msp_JE2 is correlated to the growth of the host bacterium *Meiothermus* sp. JE2 isolated from Rajwadi hot spring water having a temperature of 55°C. Phage vB_Msp_JE2 was most stable at pH 6.0 - 8.0 and was sensitive to acidic pH (Fig. 8). *Meiothermus* phage MMP17 was also described²⁶ to be stable at pH 6 - 7. These results suggest that the temperature and pH stability of the phage are influenced by the geographic origin of the phage.

Meiothermus species have been reported from hot springs and artificial thermal environments like paper

mills²³⁻²⁴ and spent nuclear fuel²⁵. *Meiothermus* sp. was first described as a paper machine contaminant by Kolari *et al* (2003), a common biofouler causing biofilm formation in the paper industry. In our view, this lytic bacteriophage vB_Msp_JE2 isolated from a natural geothermal environment might possess unique properties to control the biofilm formation in an ecofriendly manner.

Further, analysis of this phage-host relationship will yield new data regarding the microbial ecology of the hot spring. In comparison to the Himalayan hot spring, there is limited focus on microbial studies done at WCGP. Thus, the present study is the first preliminary report on the isolation of culturable *Meiothermus* sp. and its associated phage vB_Msp_JE2 from Rajwadi hot spring (WCGP), India.

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).

References

- Chandrasekharam D, Geothermal resources of India. In IBC Conference on *Geothermal Power Asia- 2000*, (Manila, Philippines), Feb 2000.
- Singh H K, Chandrasekharam D, Trupti G, Mohite P, Singh B *et al*, Potential geothermal energy resources of India: A review, *Current Sustainable/Renewable Energy Reports*, 3 (2016) 80-91.
- Minissale A, Vaselli O, Chandrasekharam D, Magro G, Tassi F *et al*, Origin and evolution of 'intracratonic' thermal fluids from Central-Western peninsular India, *Earth Planet Sci Lett*, 181 (2000) 377-394.
- Gurav T, Singh H K & Chandrasekharam D, Major and trace element concentrations in the geothermal springs along the west coast of Maharashtra, India, *Arab J Geosci*, 9 (2016) 1-5.
- Chandrasekharam D, Structure and evolution of the western continental margin of India deduced from gravity, seismic, geomagnetic and geochronological studies, *Phys Earth Planet Inter*, 41 (1985) 186-198.
- Rice G, Tang L, Stedman K, Roberto F, Spuhler J *et al*, The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life, *Proc Natl Acad Sci*, 101 (2004) 7716-7720.
- Rachel R, Bettstetter M, Hedlund B P, Häring M, Kessler A *et al*, Remarkable morphological diversity of viruses and virus-like particles in hot terrestrial environments, *Arch Virol*, 147 (2002) 2419-2429.
- Prangishvili D, Evolutionary insights from studies on viruses of hyperthermophilic archaea, *Res Microbiol*, 154 (2003) 289-294.
- Liu B, Wu S, Song Q, Zhang X & Xie L, Two novel bacteriophages of thermophilic bacteria isolated from deep-sea hydrothermal fields, *Curr Microbiol*, 53 (2006) 163-166.
- Marks T J & Hamilton P T, Characterization of a thermophilic bacteriophage of *Geobacillus kaustophilus*, *Arch Virol*, 159 (2014) 2771-2775.
- Hjorleifsdottir S, Aevansson A, Hreggvidsson G O, Fridjonsson O H & Kristjansson J K, Isolation, growth and genome of the *Rhodothermus* RM378 thermophilic bacteriophage, *Extremophiles*, 18 (2014) 261-270.
- Adiguzel A, Ozkan H, Baris O, Inan K, Gulluce M *et al*, Identification and characterization of thermophilic bacteria isolated from hot springs in Turkey, *J Microbiol Methods*, 79 (2009) 321-328.
- Boomer S M, Noll K L, Geesey G G & Dutton B E, Formation of multilayered photosynthetic biofilms in an alkaline thermal spring in Yellowstone National Park, Wyoming, *Appl Environ Microbiol*, 75 (2009) 2464-2475.
- Derekova A, Mandeva R & Kambourova M, Phylogenetic diversity of thermophilic carbohydrate degrading bacilli from Bulgarian hot springs, *World J Microbiol Biotechnol*, 24 (2008) 1697-1702.
- Zablocki O, van Z L & Trindade M, Biogeography and taxonomic overview of terrestrial hot spring thermophilic phages, *Extremophiles*, 22 (2018) 827-837.
- Prangishvili D, Stedman K & Zillig W, Viruses of the extremely thermophilic archaea on *Sulfolobus*, *Trends Microbiol*, 9 (2001) 39-43.
- Rice G, Stedman K, Snyder J, Wiedenheft B, Willits D *et al*, Viruses from extreme thermal environments, *Proc Natl Acad Sci*, 98 (2001) 13341-13345.
- Sakaki Y & Oshima T, Isolation and characterization of a bacteriophage infectious to an extreme thermophile, *Thermus thermophilus* HB8, *J Virol*, 15 (1975) 1449-1453.
- Yu M X, Slater M R & Ackermann H W, Isolation and characterization of *Thermus* bacteriophages, *Arch Virol*, 151 (2006) 663-679.
- Nagayoshi Y, Kumagai K, Mori K, Tashiro K, Nakamura A *et al*, Physiological properties and genome structure of the hyperthermophilic filamentous phage ϕ OH3 which infects *Thermus thermophilus* HB8, *Front Microbiol*, 7 (2016) 1-11.
- Nobre M F, Truper H G & Da Costa M S, Transfer of *Thermus ruber* (Loginova *et al* 1984), *Thermus silvanus* (Tenreiro *et al* 1995), and *Thermus chliarophilus* (Tenreiro *et al* 1995) to *Meiothermus* gen. nov. as *Meiothermus ruber* comb. Nov., *Meiothermus silvanus* comb. nov., and *Meiothermus chliarophilus* comb. nov., respectively, and emendation of the genus *Thermus*, *Int J Syst Bacteriol*, 46 (1996) 604-606.
- Nobre M F & Da Costa M S, Genus II *Meiothermus*, In *Bergey's Manual of Systematic Bacteriology*, 2nd edn, vol 1, edited by Boone D R, Castenholz R W, Garrity G M (Springer, New York), 2001, 414-420.
- Kolari M, Nuutinen J, Rainey F A & Salkinoja-Salonen MS, Colored moderately thermophilic bacteria in paper-machine biofilms, *J Ind Microbiol Biotechnol*, 30 (2003) 225-238.
- Ekman J, Kosonen M, Jokela S, Kolari M, Korhonen P *et al*, Detection and quantification of colored deposit-forming *Meiothermus* spp. in paper industry processes and end products, *J Ind Microbiol Biotechnol*, 34 (2007) 203-211.

- 25 Masurat P, Fru E C & Pedersen K, Identification of *Meiothermus* as the dominant genus in a storage system for spent nuclear fuel, *J Appl Microbiol*, 98 (2005) 727-740.
- 26 Lin L, Han J, Ji X, Hong W, Huang L *et al*, Isolation and characterization of a new bacteriophage MMP17 from *Meiothermus*, *Extremophiles*, 15 (2011) 253-258.
- 27 Wang F, Xiao Y, Xiong Y, Jiao Y, Zhang Q *et al*, Complete genome sequence of MMP7, a novel *Meiothermus* bacteriophage of the family *Myoviridae* isolated from a hot spring, *Arch Virol*, 165 (2020) 753-756.
- 28 Poddar A & Das S K, Microbiological studies of hot springs in India: A review, *Arch Microbiol*, 200 (2018) 1-18.
- 29 Gupta R, Xu S Y, Sharma P & Capalash N, Characterization of *MspNI* (G/GWCC) and *MspNII* (R/GATCY), novel thermostable Type II restriction endonucleases from *Meiothermus* sp., isoschizomers of *AvaII* and *BstYI*, *Mol Biol Rep*, 39 (2012) 5607-5614.
- 30 Mukherjee T, Bose S, Sen U, Roy C, Rameez M J *et al*, Genome sequence of the red pigment-forming *Meiothermus taiwanensis* strain RP isolated from Paniphala hot spring, India, *Genome Announc*, 4 (2016) 1-2.
- 31 Mukherjee T, Bose S & Mukhopadhyay S K, Antioxidant properties of the carotenoid extracts of three *Deinococcus-Thermus* phylum bacteria, *Meiothermus* sp. strains RP and TP and *Thermus* sp. strain YY from Paniphala hot spring, India, *Nutrire*, 42 (2017) 1-11.
- 32 Patel M A, Ou M S, Harbrucker R, Aldrich H C, Buszko M L *et al*, Isolation and characterization of acid-tolerant, thermophilic bacteria for effective fermentation of biomass-derived sugars to lactic acid, *Appl Environ Microbiol*, 72 (2006) 3228-3235.
- 33 Smbert R M & Krieg N R, *Phenotypic characterization. Methods for general and molecular bacteriology*, (American Society of Microbiology, Washington, DC) 1994, 611-651.
- 34 Krieg N R & Holt J G, *Bergey's Manual of Systematic Bacteriology*, (Williams & Wilkins Co, Baltimore) 1984.
- 35 Brosius J, Palmer M L, Kennedy P J & Noller H F, Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*, *Proc Natl Acad Sci U.S.A*, 75 (1978) 4801-4805.
- 36 Kumar S, Stecher G, Li M, Knyaz C & Tamura K, MEGA X: molecular evolutionary genetics analysis across computing platforms, *Mol Bio Evol*, 35 (2018) 1547-1549.
- 37 Sambrook J, Fritsch E F & Maniatis T, *Molecular cloning: A laboratory manual*, 2nd edn, vol 1, (Cold spring Harbor Laboratory press, New York) 1989.
- 38 Prata C, Ribeiro A, Cunha Â, Gomes N C & Almeida A, Ultracentrifugation as a direct method to concentrate viruses in environmental waters: Virus-like particle enumeration as a new approach to determine the efficiency of recovery, *J Environ Monit*, 14 (2012) 64-70.
- 39 Adams MH, *Bacteriophages*, (Interscience Publishers, Inc, New York) 1959, 450-451.
- 40 Kropinski A M, Mazzocco A, Waddell T E, Lingohr E & Johnson R P, *Bacteriophages* (Humana Press) 2009, 69-76.
- 41 Middelboe M, Chan A M & Bertelsen S K, Isolation and life cycle characterization of lytic viruses infecting heterotrophic bacteria and cyanobacteria, *Manual of Aquatic Viral Ecology*, 13 (2010) 118-133.
- 42 Low U, Absar A, Duraiswami R & Singh A, Geophysical exploration of Tural-Rajwadi group of hot springs, West Coast Geothermal Province, Maharashtra, India and its implications, *Geothermics*, 88 (2020) 1-13.
- 43 Mathur J, Bizzoco R W, Ellis D G, Lipson D A, Poole A W *et al*, Effects of abiotic factors on the phylogenetic diversity of bacterial communities in acidic thermal springs, *Appl Environ Microbiol*, 73 (2007) 2612-2623.
- 44 Chandrasekhar T, Chandrasekhar V & Chandrasekharam D, Geothermometry of West Coast Geothermal Province, Maharashtra, India, *Trans Geotherm Resour Counc*, 40 (2016) 495-499.
- 45 Pitale U L, Dubey R, Saxena R K, Prasad J M & Muthuraman K, Review of geothermal studies of west coast hot spring belt, Maharashtra, *Records of the Geological Survey of India*, 115 (1989) 97-136.
- 46 Sarolkar P B, Geochemical characters of hot springs of west coast, Maharashtra state, India. In *Proc World Geotherm. Congress*, (Antalaya, Turkey), 2005, 24-29.
- 47 Chatterjee S, Sharma S, Ansari M A, Deodhar A S, Low U *et al*, Characterization of subsurface processes estimation of reservoir temperature in Tural Rajwadi geothermal fields, Maharashtra, India, *Geothermics*, 59 (2016) 77-89.
- 48 León K B D, Gerlach R, Peyton B M & Fields M W, Archaeal and bacterial communities in three alkaline hot springs in Heart Lake Geyser Basin, Yellowstone National Park, *Front Microbiol*, 4 (2013) 1-10.
- 49 Tindall B J, Sikorski J, Lucas S, Goltsman E, Copeland A *et al*, Complete genome sequence of *Meiothermus ruber* type strain (21^T), *Stand Genom Sci*, 3 (2010) 26-36.
- 50 Saxena R, Dhakan D B, Mittal P, Waiker P, Chowdhury A *et al*, Metagenomic analysis of hot springs in Central India reveals hydrocarbon degrading thermophiles and pathways essential for survival in extreme environments, *Front Microbiol*, 7 (2017) 1-17.
- 51 Panda A K, Bisht S S, Kumar, N S & De Mandal S, Investigations on microbial diversity of Jakrem hot spring, Meghalaya, India using cultivation-independent approach, *Genom Data*, 4 (2015) 156-157.
- 52 Sakaki Y & Oshima T, Isolation and characterization of a bacteriophage infectious to an extreme thermophile, *Thermus thermophilus* HB8, *J Virol*, 151 (1975) 1449-1453.
- 53 Lin L, Hong W, Jian X, Huang L, Yunlin W *et al*, Isolation and characterization of an extremely long tail *Thermus* bacteriophage from Tengchong hot springs in China, *J Basic Microbiol*, 50 (2010) 452-456.
- 54 Naryshkina T, Liu J, Florens L, Swanson S K, Pavlov A R *et al*, *Thermus thermophilus* bacteriophage phi YS40 genome and proteomic characterization of virions, *J Mol Biol*, 364 (2006) 667-677.
- 55 Nagayoshi Y, Kumagae K, Mori K, Tashiro K, Nakamura A *et al*, Physiological properties and genome structure of the hyperthermophilic filamentous phage ϕ OH3 which infects *Thermus thermophilus* HB8, *Front Microbiol*, 7 (2016) 50.



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

Nicola Faria¹ · Sanjeev C. Ghadi¹

Received: 31 January 2023 / Accepted: 6 June 2023

© The Author(s), under exclusive licence to Plant Science and Biodiversity Centre, Slovak Academy of Sciences (SAS), Institute of Zoology, Slovak Academy of Sciences (SAS), Institute of Molecular Biology, Slovak Academy of Sciences (SAS) 2023

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-acetylglutaminylglutamine 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

Introduction

Geothermal springs from India are grouped into seven geological provinces: Himalaya, Sohana, Cambay, West Coast, Sona-Narmada Tapi (SONATA), Godavari, and Mahanandi

(Chandrasekharam and Chandrasekhar 2000). The Geological Survey of India (GSI) has reported around 400 thermal springs based on tectonic elements and geothermal gradients across India. (Chandrasekhar and 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 et al. 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 and 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.

✉ Sanjeev C. Ghadi
saga@unigoa.ac.in

¹ School of Biological Sciences & Biotechnology, Goa University, Taleigao Plateau, Goa 403206, India

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.

Materials and methods

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, $(\text{NH}_4)_2\text{SO}_4$: 1.0, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$: 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 (150×2 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 77 K-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 OrthoANIm 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 $\mu\text{g/mL}$ (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 $5000\times g$ for 10 min, followed by filtration through a 0.22 μm filter. 10 μL of filtered lysate was spotted on the lawn of *Pseudomonas* sp. VB3 host (Sambrook et al. 1989). Similarly, 10 μL 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 50°C and 7.4 respectively. The conductivity and TDS measured *in situ* were 162 $\mu\text{S cm}^{-1}$ 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 azotoformans* strain DSM 18862^T (99.33%), *Pseudomonas carnis* strain B4-1^T (99.33%), *Pseudomonas paraceticus* 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

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

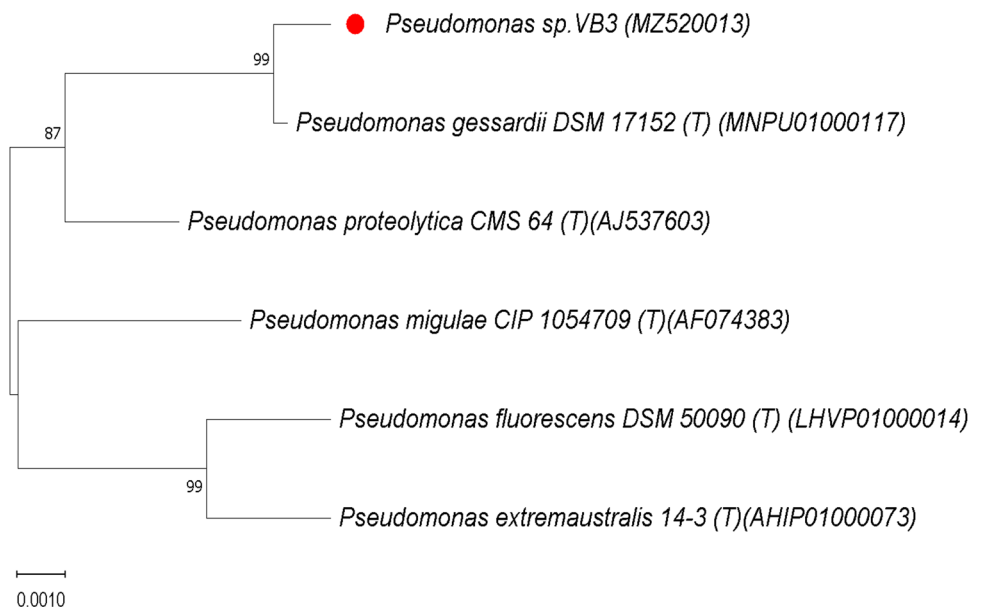
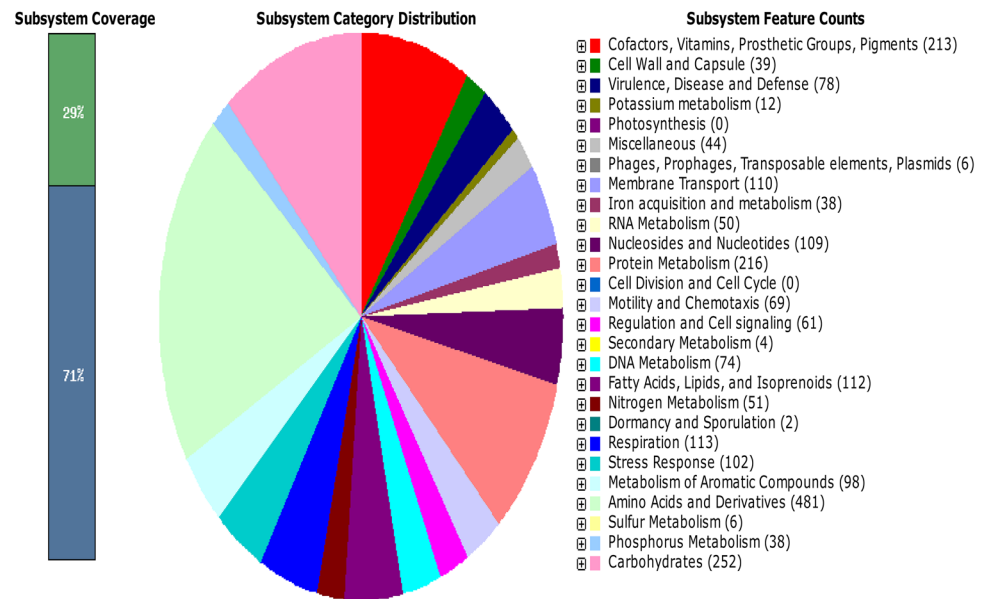


Fig. 2 RAST annotation of *Pseudomonas* sp. VB3: The subsystem category distribution shows the percentage distribution of the genes in different pathways, that are labelled with different colours and the number of genes in each are mentioned in the bracket



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

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)

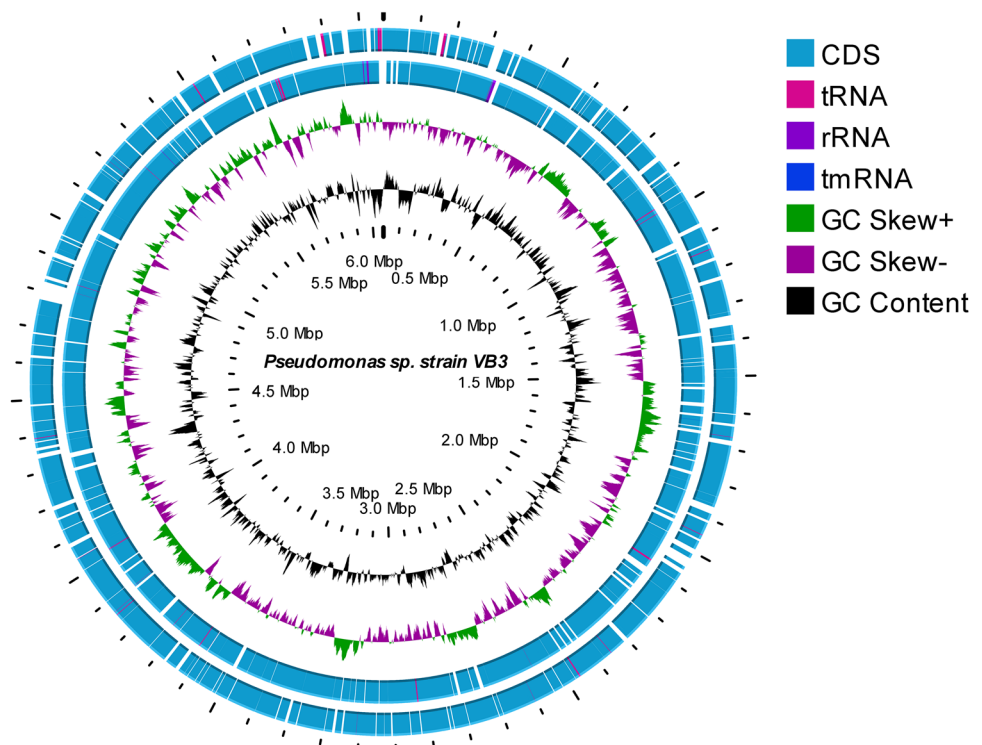


Table 1 Secondary metabolite biosynthetic gene clusters in *Pseudomonas* sp. strain VB3 as predicted by antiSMASH 6.1.1

Clusters	Type	Most similar know cluster with MIBiG	MIBiG-BGC-ID (Gene location predicted by anti- iSMASH)
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) ^b	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-acetylglutaminylglutamine 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

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 $\leq 90\%$. 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).

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	20.5 kb	29	Intact	PHAGE_Vibrio_VP882_NC_009016 (14)	60.53%
Prophage region 2	150	47.6 kb	45	Intact	PHAGE_Iodoba_phPLPE_NC_0011142(14)	56.88%
Prophage region 3	20	10.3 kb	19	Incomplete	PHAGE_Pseudo_PMG1_NC_016765(3)	57.38%
Prophage region 4	30	24.2 kb	12	Incomplete	PHAGE_Escher_500465_1_NC_049342 (6)	58.75%
Prophage region 5	30	7.4 kb	10	Incomplete	PHAGE_Shigel_sf11_NC_021857 (2)	52.49%

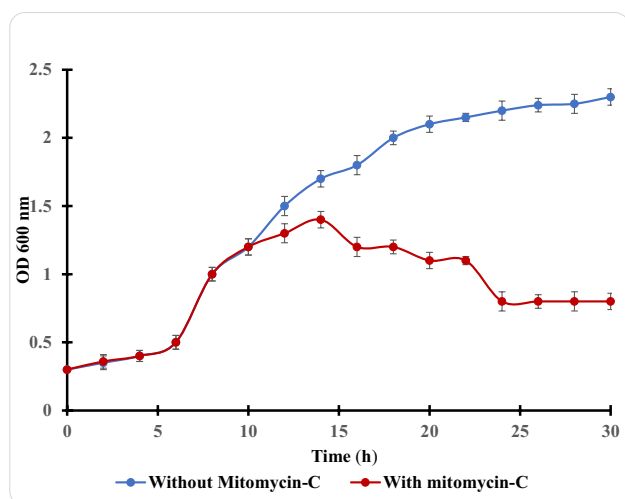


Fig. 4 Prophage induction in bacterial strain VB3 with and without Mitomycin-C

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 $\leq 29.1\%$ 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, betalactone, butyrolactone, hserlactone, NAGGN (N-acetylglutaminylglutamine 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).

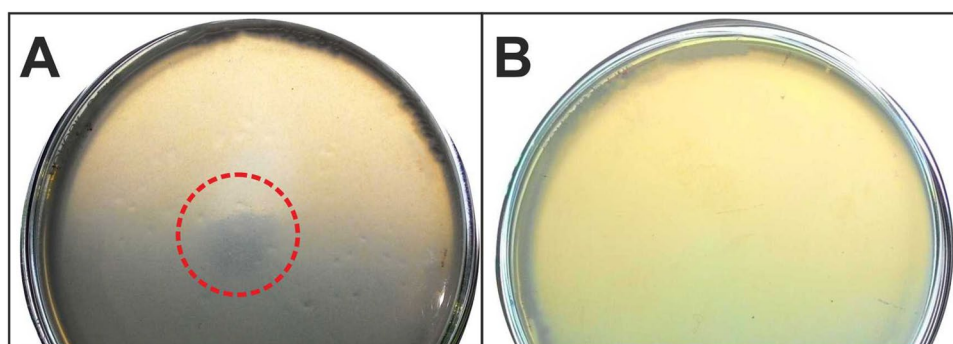
The BGC for lankacidin C is known to have antitumor and immunosuppressive activities on cancer cell lines (Ootsu et al. 1975). Additionally, fengycin has been reported as a biocontrol against fungal disease in crop plants (Zeng et al. 2021). Further characterization of these pathways and secondary metabolites 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

Fig. 5 Spot test after induction with Mitomycin-C on the lawn of *Pseudomonas* sp. VB3. **A** Mitomycin-C treated, Circle depicts clearance zone, **B** Control, without Mitomycin-C



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.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11756-023-01470-y>.

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).

Data availability 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.

Declarations

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

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

References

- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed 15 June 2020
- 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* 19:455–477. <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* 5:1041–1055. <https://doi.org/10.2217/fmb.10.66>
- Chandrasekhar V, Chandrasekhar D (2010) Energy independence through CDM using geothermal resources: Indian scenario. In *Proceedings World Geothermal Congress*
- Chandrasekhar 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):2461–2470. <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* 282:2987–2995. <https://doi.org/10.1074/jbc.M609238200>
- Gurevich A, Saveliev V, Vyahhi 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 SC, Paul JH (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–D807. <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* 59:919–928
- Parte AC, Carbasse JS, Meier-Kolthoff JP, Reimer LC, Göker M (2020) List of Prokaryotic names with Standing in Nomenclature (LPSN)

- p 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* 9:1132–1147. <https://doi.org/10.1016/j.meegid.2009.08.001>
- Poddar A, Das SK (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* 106:19126–19131. <https://doi.org/10.1073/pnas.0906412106>
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold spring harbor laboratory press, New York
- 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* 30:2068–2069. <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* 20:1576–1582. <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* 31:3350–3352. <https://doi.org/10.1093/bioinformatics/btv383>
- Wong BT, Lee DJ (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 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* 10:2787–2800. <https://doi.org/10.1038/ismej.2016.85>
- 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>
- Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
- 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.
- Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.