

DATA NOTE

Open Access



Complete genome of a novel mycobacteriophage WXIN isolated in Wuhan, China

Haoming Wu^{1*}, Wenxin Li¹, Chi Zeng¹, Jiaxin Li¹ and Huan Wu²

Abstract

Objectives The rising of antibiotic resistance has sparked a renewed interest in mycobacteriophage as alternative therapeutic strategies against mycobacterial infections. So far, the vast majority of mycobacteriophages have been isolated using the model species *Mycobacterium smegmatis*, implying an overwhelming majority of mycobacteriophages in the environment remain uncultured, unclassified, and their specific hosts and infection strategies are still unknown. This study was undertaken to isolate and characterize novel mycobacteriophages targeting *Mycobacterium septicum*.

Data description Here a novel mycobacteriophage WXIN against *M. septicum* was isolated from soil samples in Wuhan, China. Whole genome analysis indicates that the phage genome consists of 115,158 bp with a GC content of 61.9%. Of the 260 putative open reading frames, 46 may be associated with phage packaging, structure, lysis, lysogeny, genome modification/replication, and other functional roles. The limited genome-wide similarity, along with phylogenetic trees constructed based on viral proteome and orthologous genes show that phage WXIN represents a novel cluster distantly related to cluster J mycobacteriophages (genus *Omegavirus*). Overall, these results provide novel insights into the genomic properties of mycobacteriophages, highlighting the great genetic diversity of mycobacteriophages in relation to their hosts.

Keywords Mycobacteriophage, *Mycobacterium septicum*, Complete genome

Objective

The emergence and rapid spread of drug resistance, especially multidrug-resistant (MDR) and extensively drug-resistant (XDR), poses a great challenge to the clinical control for pathogenic mycobacteria, particularly *Mycobacterium tuberculosis* (*M. tuberculosis*) infection [1]. In recent years, phage treatment has undergone a revitalization as a promising strategy against antimicrobial resistance. Phages, formerly bacteriophages, are natural enemies of bacteria and are believed to be the most abundant organisms on the planet [2]. Unlike antibiotics, phages are characterized by self-replication, high host specificity, biofilm degradation, and low toxicity

*Correspondence:

Haoming Wu
wuhm0701@163.com

¹Pilot Base of Food Microbial Resources Utilization of Hubei Province, School of Life Science and Technology, Wuhan Polytechnic University, Wuhan 430023, China

²Department of Laboratory Medicine, Wuhan Children's Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430019, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

to humans [3, 4]. Several successful clinical trials have been reported to treat patients with disseminated drug-resistant *M. chelonae* and *M. abscessus* infections using naturally occurring phages and/or their genetically engineered derivatives [3, 5, 6].

Since the first mycobacteriophage was identified in 1947 from soil [7], more than 12,000 mycobacteriophages have been isolated from different sources, of which approximately 2,000 have their complete genomes sequenced (<https://phagesdb.org>). These mycobacteriophages are categorized into 30 different types of clusters and at least 10 singletons. Due to the non-pathogenicity, rapid growth, and similarities with other mycobacteria, *M. smegmatis* has been used as a suitable model for studying the pathogenesis of mycobacteria and for initial screening for mycobacteriophages [8]. Hence, almost all known mycobacteriophages were isolated using the *M. smegmatis* mc²155 as host. Further isolation and characterization of mycobacteriophages against different mycobacterial species in different environments will not only enrich our knowledge of phage genetics, ecology, and evolution, but also provide rich phage resources for genetic engineering. *M. septicum* is a rapidly growing non-tuberculosis mycobacteria associated with *Mycobacterium fortuitum* group (MFG) [9]. In this study, a novel mycobacteriophage WXIN against *M. septicum* was isolated from soil samples in Wuhan, China. The results provide some new insights into genomic characteristics of mycobacteriophage.

Data description

The mycobacteriophage WXIN was originally isolated from soils collected in Wuhan, China, using *M. septicum* as host. Phage genomic DNA was extracted using the λ Phage genomic DNA extraction kit (Beijing Baiabolabo Biotechnology Co., Ltd, Beijing, China). Sequencing libraries were prepared using the NEBNext Ultra DNA library prep kit (New England Biolabs, USA) and sequencing was performed on the Illumina NovaSeq 6000 platform to generate 4,191,323 paired-end reads (2×150 bp). The quality of the sequencing reads was evaluated using FastQC v0.12.1 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and was assembled de novo using Megahit v1.2.9 [10]. Genome termini were identified by PCR using primers (WT: 5'-CACATGTCGGCGTGACGT-3' and WW: 5'-CGTCATTAATGTCCCCCTCG-3') facing off the contig ends to produce an approximately 400 bp product, and the PCR product was submitted for Sanger sequencing by Sangon Biotech (Shanghai, China). In total, the complete genome of mycobacteriophage WXIN was 115,158 bp with a GC content of 61.9%. Genome was annotated using the RAST server (<https://rast.nmpdr.org/>). The predicted proteins were assigned putative functions by HHpred [11] and

BLASTp-searching against the NCBI non-redundant (nr) database (<https://www.ncbi.nlm.nih.gov>). A total of 260 putative open reading frames (ORFs) were identified in the genome, with 184 on the sense strand and 76 on the antisense (Data file 1) [12]. Among them, 46 ORFs could be functionally annotated and classified into modules involving in the packaging, structure, virion release, lysogeny, genome modification and replication (Data file 2) [12]. Two tRNA genes were identified by tRNAscan SE 1.21 (<http://lowelab.ucsc.edu/tRNAscan-SE/>). No virulence or drug resistance-related gene was found using the Virulence Factor Database (VFDB, <http://www.mgc.ac.cn/VFs/>) and the Comprehensive Antibiotic Resistance Database (CARD, <https://card.mcmaster.ca/>).

Proteomic trees generated by ViPTree [13] indicates that WXIN is mostly related to cluster J mycobacteriophages (genus *Omegavirus*) and likely represents a novel mycobacteriophage cluster (Data file 3) [12]. Genome-wide comparisons by VIRIDIC (<https://rhea.icbm.uni-oldenburg.de/viridic/>) reveals extremely low sequence similarity between WXIN and cluster J mycobacteriophages (Data file 4, 21.2–24.0% nt identity) [12]. In addition, 15 single-copy orthologous genes are shared by WXIN and the related cluster J, X and E mycobacteriophages, as indicated by Orthofinder v2.2.7 through all-to-all BLASTp analysis [14], including the terminase large subunit (ORF7), tape measure protein (ORF29), minor tail proteins (ORF31 and ORF33) and lysine protein A & B (ORF43 and ORF45) and several hypothetical proteins (ORF53, ORF82, ORF128, ORF130, ORF146, ORF175, ORF184, ORF193 and ORF194). The average amino acid identity (AAI) of the 15 orthologous genes range from 40.7 to 64.1% between WXIN and other related mycobacteriophages (Data file 5) [12]. Phylogenetic analysis based on the concatenated protein sequences of the 15 orthologous genes also revealed similar tree topology with that of ViPTree analysis using IQ-TREE v1.6.5 (Data file 6) [12, 15].

Limitations

This data note was limited to the description of mycobacteriophage WXIN. A larger collection is needed to help us better understand the genetic characteristics of mycobacteriophages for *Mycobacterium septicum*.

Table 1 Overview of data files/data sets

Label	Name of data file/data set	File types (file extension)	Data repository and identifier (DOI or accession number)
Data file 1	Circular genome annotation of mycobacteriophage WXIN	Portable Document Format file (.pdf)	Science data bank (https://doi.org/10.57760/sciencedb.16539) [12]
Data file 2	Predicted genes of mycobacteriophage WXIN	MS Excel file (.xlsx)	Science data bank (https://doi.org/10.57760/sciencedb.16539) [12]
Data file 3	Proteomic tree by ViPTree	Portable Document Format file (.pdf)	Science data bank (https://doi.org/10.57760/sciencedb.16539) [12]
Data file 4	Intergenomic similarities between WXIN and 10 Cluster J mycobacteriophages	Portable Document Format file (.pdf)	Science data bank (https://doi.org/10.57760/sciencedb.16539) [12]
Data file 5	AAI values of 15 orthologous genes	Portable Document Format file (.pdf)	Science data bank (https://doi.org/10.57760/sciencedb.16539) [12]
Data file 6	ML phylogenetic Inference of 15 orthologous genes	Portable Document Format file (.pdf)	Science data bank (https://doi.org/10.57760/sciencedb.16539) [12]
Data set 1	Mycobacteriophage WXIN, complete genome	Genbank file (.gbk)	NCBI Nucleotide, https://identifiers.org/nucleotide:OR813930 [16]
Data set 2	Raw sequencing reads of Mycobacteriophage WXIN	SRA file (.sra)	Sequence Read Archive (https://identifiers.org/ncbi/insdc.sra:SRP492847) [17]

Acknowledgements

We would thank Dr. Haiyan Zeng for her advice and comments in this study.

Author contributions

HM.W, C.Z and H.W designed and supervised the research. WX.L, HM.W and JX.L performed the experiments, data collection and analysis. The manuscript was written by HM.W and WX.L. All authors read and approved the final manuscript.

Funding

This research was supported by grant from the Scientific Research Staring Foundation of Wuhan Polytechnic University (2022RZ052).

Data availability

The genomic sequence described in this Data note can be freely and openly accessed on NCBI GenBank under accession number OR813930. Raw sequence reads are available from NCBI under BioProject PRJNA1082475.

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 1 March 2024 / Accepted: 12 June 2024

Published online: 18 June 2024

References

- Dookie N, Ngema SL, Perumal R, Naicker N, Padayatchi N, Naidoo K. The changing paradigm of drug-resistant tuberculosis treatment: successes, pitfalls, and future perspectives. *Clin Microbiol Rev.* 2022;35(4):e0018019. <https://doi.org/10.1128/cmr.00180-19>.
- Suttle CA. Viruses in the sea. *Nature.* 2005;437(7057):356–61. <https://doi.org/10.1038/nature04160>.
- Little JS, Dedrick RM, Freeman KG, Cristinziano M, Smith BE, Benson CA, Jhaveri TA, Baden LR, Solomon DA, Hatfull GF. Bacteriophage treatment of disseminated cutaneous *Mycobacterium chelonae* infection. *Nat Commun* 2022, 13(1):2313. <https://doi.org/10.1038/s41467-022-29689-4>.
- Kortright KE, Chan BK, Koff JL, Turner PE. Phage therapy: a renewed Approach to combat antibiotic-resistant Bacteria. *Cell Host Microbe.* 2019;25(2):219–32. <https://doi.org/10.1016/j.chom.2019.01.014>.
- Dedrick RM, Guerrero-Bustamante CA, Garlena RA, Russell DA, Ford K, Harris K, Gilmour KC, Soothill J, Jacobs-Sera D, Schooley RT, et al. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*. *Nat Med.* 2019;25(5):730–3. <https://doi.org/10.1038/s41591-019-0437-z>.
- Nick JA, Dedrick RM, Gray AL, Vldar EK, Smith BE, Freeman KG, Malcolm KC, Epperson LE, Hasan NA, Hendrix J et al. Host and pathogen response to bacteriophage engineered against *Mycobacterium abscessus* lung infection. *Cell* 2022, 185(11):1860–e18741812. <https://doi.org/10.1016/j.cell.2022.04.024>.
- McNerney R. TB: the return of the phage. A review of fifty years of mycobacteriophage research. *Int J Tuberc Lung Dis.* 1999;3(3):179–84.
- Sparks IL, Derbyshire KM, Jacobs WR Jr., Morita YS. *Mycobacterium smegmatis*: the Vanguard of Mycobacterial Research. *J Bacteriol.* 2023;205(1):e0033722. <https://doi.org/10.1128/jb.00337-22>.
- Go JR, Wengenack NL, Abu Saleh OM, Corsini Campioli C, Deml SM, Wilson JW. *Mycobacterium septicum*: a 6-Year clinical experience from a Tertiary Hospital and Reference Laboratory. *J Clin Microbiol* 2020, 58(12). <https://doi.org/10.1128/JCM.02091-20>.
- Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de bruijn graph. *Bioinformatics.* 2015;31(10):1674–6. <https://doi.org/10.1093/bioinformatics/btv033>.
- Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 2005. <https://doi.org/10.1093/nar/gki408>. 33(Web Server issue):W244–248.
- Wu H, Li W, Zeng C, Li J, Wu H. Complete genome of a novel mycobacteriophage WXIN isolated in Wuhan, China. *Sci Data Bank.* 2024. <https://doi.org/10.57760/sciencedb.16539>.
- Nishimura Y, Yoshida T, Kuronishi M, Uehara H, Ogata H, Goto S. ViPTree: the viral proteomic tree server. *Bioinf* 2017, 33(15):2379–80. <https://doi.org/10.1093/bioinformatics/btx157>.
- Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 2019;20(1):238. <https://doi.org/10.1186/s13059-019-1832-y>.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol.* 2020;37(5):1530–4. <https://doi.org/10.1093/molbev/msaa015>.

16. Li W, Wu H, Li J, Zeng C, Li J, Wu H. Mycobacterium phage WXIN, complete genome. NCBI. 2023. <https://identifiers.org/nucleotide:OR813930>.
17. Wu H, Li W, Zeng C, Li J, Wu H. Genomic sequence of mycobacteriophage WXIN. NCBI. 2024. <https://identifiers.org/ncbi/insdc.sra:SRP492847>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.