Check for updates

| Bacteriophages | Announcement

Genome sequence and annotation of *Arthrobacter globiformis* phage Vulpecula (AS1) isolated from soil in Dahlonega, Georgia

Lily Chuhran,[1] Chase Whitlow,[1] Carson Teems,[1] Alison Kanak,[1] Shane A. Webb[1]

AUTHOR AFFILIATION See affiliation list on p. 3.

ABSTRACT Vulpecula, a temperate bacteriophage collected from soil in Dahlonega, Georgia using host *Arthrobacter globiformis*, is an AS1 subcluster virus of 37,766 bp (67.7% GC). Genome annotation suggests 64 open reading frames, no predicted tRNA genes, and ~98% sequence similarity to AS1 phages Ruchi (from GA) and Jamun (New Hampshire).

KEYWORDS viral bioinformatics, genome annotation

Bacteriophages represent a potential for weaponization in the fight against multidrug resistance (bacteriophage therapy) (1–3). We therefore must understand phage diversity and evolution. Here we contribute to this endeavor with the annotated genome of Vulpecula, an AS1 subcluster bacteriophage.

Vulpecula was isolated in 2022 from enriched, Vickery House Garden soil at UNG in Dahlonega, Georgia (34.53N, 83.98W) following the SEA-PHAGES protocol (4, 5). The soil was mixed with LB and incubated for 1 hour at 30°C. The supernatant was then collected and sterilized by 0.22 µm filtration. Phage presence was confirmed and purified by standard plaque assay using *A. globiformis* B-2979 and then amplified to high titer *via* web plate flooding. A Wizard DNA extraction kit (Promega) was used to produce purified genomic DNA from the lysate. A sequencing library was constructed with an NEBNext Ultra II FS kit (vers3). WGS sequencing (Illumina MiSeq) produced ~2,101× coverage from >630 k 150-base single-end reads. Overlapping terminal reads confirmed genome completeness. Newbler 2.9 (Roche) assembled the genome, and accuracy and completeness were evaluated with Consed 29 (6). The genome was found to have 37,766 bp, 67.7% GC content, and a 3′-overhang of GAGTTGCCGGGA.

Genome annotation depended on phagesdb (7) and software including DNA Master 5.23.6 (8), Glimmer 3.02 (9), GeneMark 3.26 (10), BLAST (11, 12), HHPred 2.08 (13) executed by the MPI Bioinformatics Toolkit (14), Phamerator 505 (Actino_draft) (15), tRNAscan SE 2.0 (16), Aragorn (17), and DeepTMHMM 1.0.24 (18) (default settings for all). ORFs, gaps, and potential ribosomal binding sites were first assigned with DNA Master, Glimmer, and GeneMark. Initial assignments were refined through homology assessment using Phamerator, BLAST, and HHPred. An expect (e) value $<10^{-4}$ was used for function assignments. DeepTMHMM assessed ORFs for trans-membrane domains.

The Vulpecula genome is predicted to contain 64 open reading frames [37 with ascribed function (58%)] and no predicted tRNA genes. Genes 1–24 and 35–64 are forward oriented, and genes 25–34 are reverse oriented. Among the annotated genes are three nucleases (ORFs 29,53,64), endolysin (ORF 23), an immunity repressor (ORF 34) adjacent to tyrosine integrase (ORF 33), and an excise protein (ORF 36) as well as RusA-like resolvase (ORF 41). ORFs 14 and 15 encode overlapping tail assembly chaperones (117 and 241 aa, respectively) with ORF 14 terminated by a −1 frameshift (position 10412). ORF 28 may code for a HIP116 Rad5p N-terminal (HIRAN) domain-containing protein, which appears in numerous *Arthrobacter* species, but appears absent
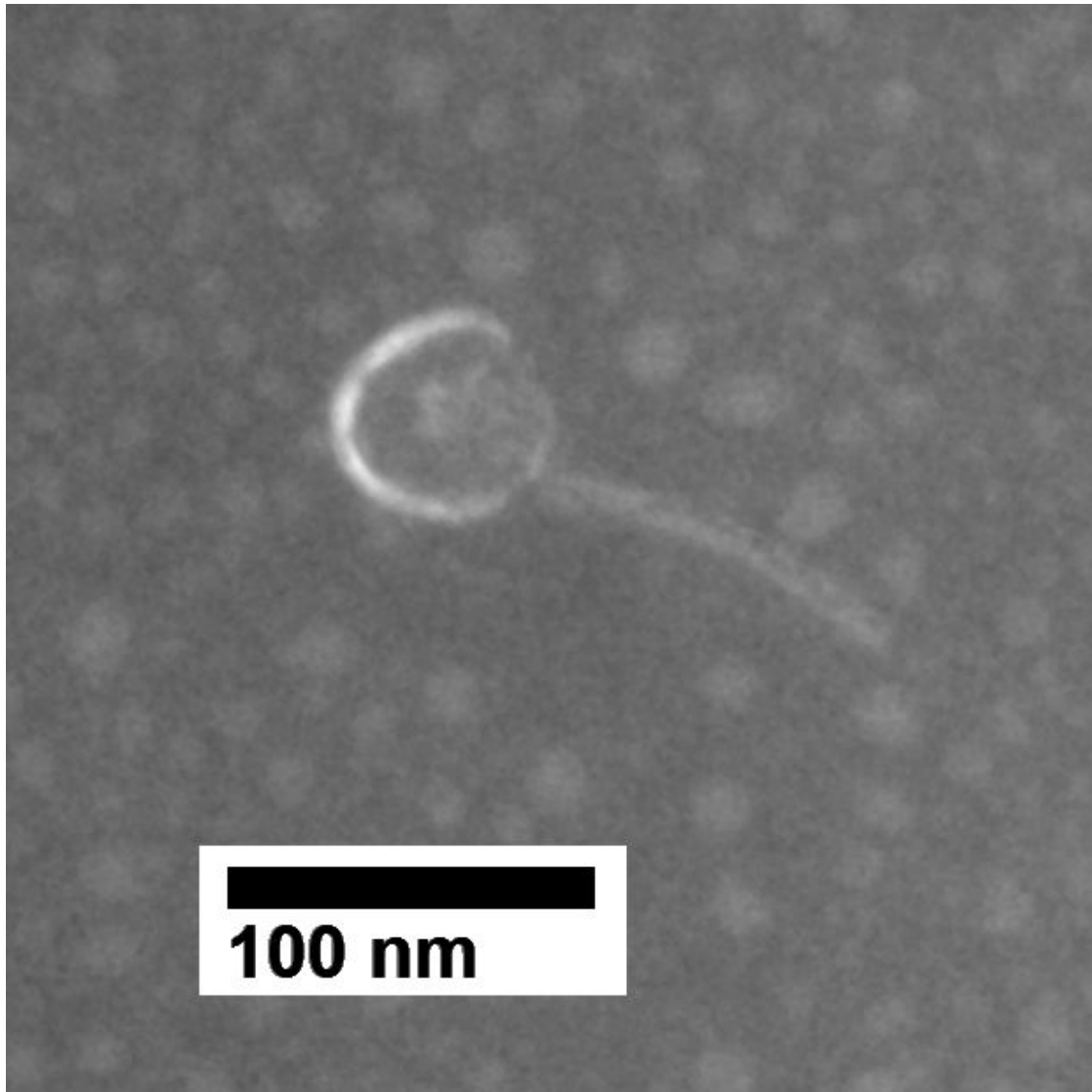
FIG 1 TEM image of phage Vulpecula (head diameter ~57 nm and tail length ~106 nm). The image was captured at the University of Georgia Electron Microscope facility with a JEM-1011 TEM (JOEL, Tokyo, Japan). Phosphotungstic acid was used to stain the lysate.

from other AS1 bacteriophages. Three ORFs with assigned function (ORFs 416,24) and five with unassigned function (ORFs 1,21,22,39,47) likely have transmembrane domains.

Genome BLAST revealed Vulpecula shares the highest nucleotide sequence similarity with *Arthrobacter* phages Ruchi (98.0% identity; OR434022.1) and Jamun (97.4% identity; OP297550.1), which were isolated in 2022 from Lumpkin County, GA (6.6 km from the Vulpecula locality) and in 2021 from Bedford, NH, respectively. Collectively, Vulpecula exhibits siphovirid morphology based on viral particle anatomy (Fig. 1) and homology assessment. Plaque morphology (7 mm circular plaques with hazy peripheries and 2 mm clear centers) and the presence of tyrosine integrase (ORF 33) support that it is temperate.

AUTHOR AFFILIATION

[1]Department of Biology, University of North Georgia, Dahlonega, Georgia, USA

AUTHOR ORCIDs

Alison Kanak http://orcid.org/0000-0001-9053-250X
Shane A. Webb http://orcid.org/0009-0006-5669-404X

AUTHOR CONTRIBUTIONS

Lily Chuhran, Formal analysis, Investigation, Writing – original draft | Chase Whitlow, Formal analysis, Investigation, Writing – original draft | Carson Teems, Formal analysis, Investigation, Writing – original draft | Alison Kanak, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing – review and editing | Shane A. Webb, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing – review and editing

DATA AVAILABILITY

The Vulpecula genome annotation can be accessed through NCBI (GenBank OR475258) and sequencing reads can be obtained from the SRA (SRX22366557).

REFERENCES

1. Vandamme EJ, Mortelmans K. 2019. A century of bacteriophage research and applications: impacts on biotechnology, health, ecology and the economy! J Chem Technol Biotechnol 94:323–342. https://doi.org/10.1002/jctb.5810
2. Marongiu L, Burkard M, Lauer UM, Hoelzle LE, Venturelli S. 2022. Reassessment of historical clinical trials supports the effectiveness of phage therapy. Clin Microbiol Rev 35:e0006222. https://doi.org/10.1128/cmr.00062-22
3. Samir S. 2023. Basic guidelines for bacteriophage isolation and characterization. Recent Pat Biotechnol 17:312–331. https://doi.org/10.2174/1872208317666221017094715
4. Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SCR, et al. 2014. A broadly implementable research course in phage discovery and genomics for first-year undergraduate students. mBio 5:e01051-13. https://doi.org/10.1128/mBio.01051-13
5. Pope W, Jacobs-Sera D, Sussel D, Cresawn S, Hatfull G. 2017. SEA-phages: DNA master annotation guide & bioinformatics guide
6. Russell DA. 2018. Sequencing, assembling, and finishing complete bacteriophage genomes. Methods Mol Biol 1681:109–125. https://doi.org/10.1007/978-1-4939-7343-9_9
7. Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. Bioinformatics 33:784–786. https://doi.org/10.1093/bioinformatics/btw711
8. Lawrence J. 2021. DNA master version 5.23.6 (build 2705; 24 Oct 2021)
9. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23:673–679. https://doi.org/10.1093/bioinformatics/btm009
10. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res 29:2607–2618. https://doi.org/10.1093/nar/29.12.2607
11. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S, Madej T, Marchler-Bauer A, Lanczycki C, Lathrop S, Lu Z, Thibaud-Nissen F, Murphy T, Phan L, Skripchenko Y, Tse T, Wang J, Williams R, Trawick BW, Pruitt KD, Sherry ST. 2022. Database resources of the national center for biotechnology information. Nucleic Acids Res 50:D20–D26. https://doi.org/10.1093/nar/gkab1112
12. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2
13. Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res 33:W244–W248. https://doi.org/10.1093/nar/gki408
14. Gabler F, Nam SZ, Till S, Mirdita M, Steinegger M, Söding J, Lupas AN, Alva V. 2020. Protein sequence analysis using the MPI bioinformatics toolkit. Curr Protoc Bioinform 72:e108. https://doi.org/10.1002/cpbi.108
15. Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF. 2011. Phamerator: a bioinformatic tool for comparative bacteriophage genomics. BMC Bioinform 12:9077–9096. https://doi.org/10.1186/1471-2105-12-395
16. Chan PP, Lin BY, Mak AJ, Lowe TM. 2021. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res 49:9077–9096. https://doi.org/10.1093/nar/gkab688
17. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 32:11–16. https://doi.org/10.1093/nar/gkh152
18. Hallgren J, Tsirigos KD, Pedersen MD, Almagro Armenteros JJ, Marcatili P, Nielsen H, Krogh A, Winther O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. bioRxiv. https://doi.org/10.1101/2022.04.08.487609