

# Complete Genome Sequence of two *Gordonia rubripertincta* cluster CS4 Bacteriophages, Vitaenoi and Philon9

Joshua T. McMenemy<sup>1\*</sup>, Amanda I. Acra<sup>1\*</sup>, Elisabeth A. Alexander<sup>1\*</sup>, Vanessa K. Cordovez<sup>1\*</sup>, Morgan M. Fitzpatrick<sup>1\*</sup>, Dayanara Mejia<sup>1\*</sup>, Chaeli H. Smith<sup>1\*</sup>, Marie P. Fogarty<sup>2§</sup>

<sup>1</sup>Durham Technical Community College, Durham, North Carolina, United States

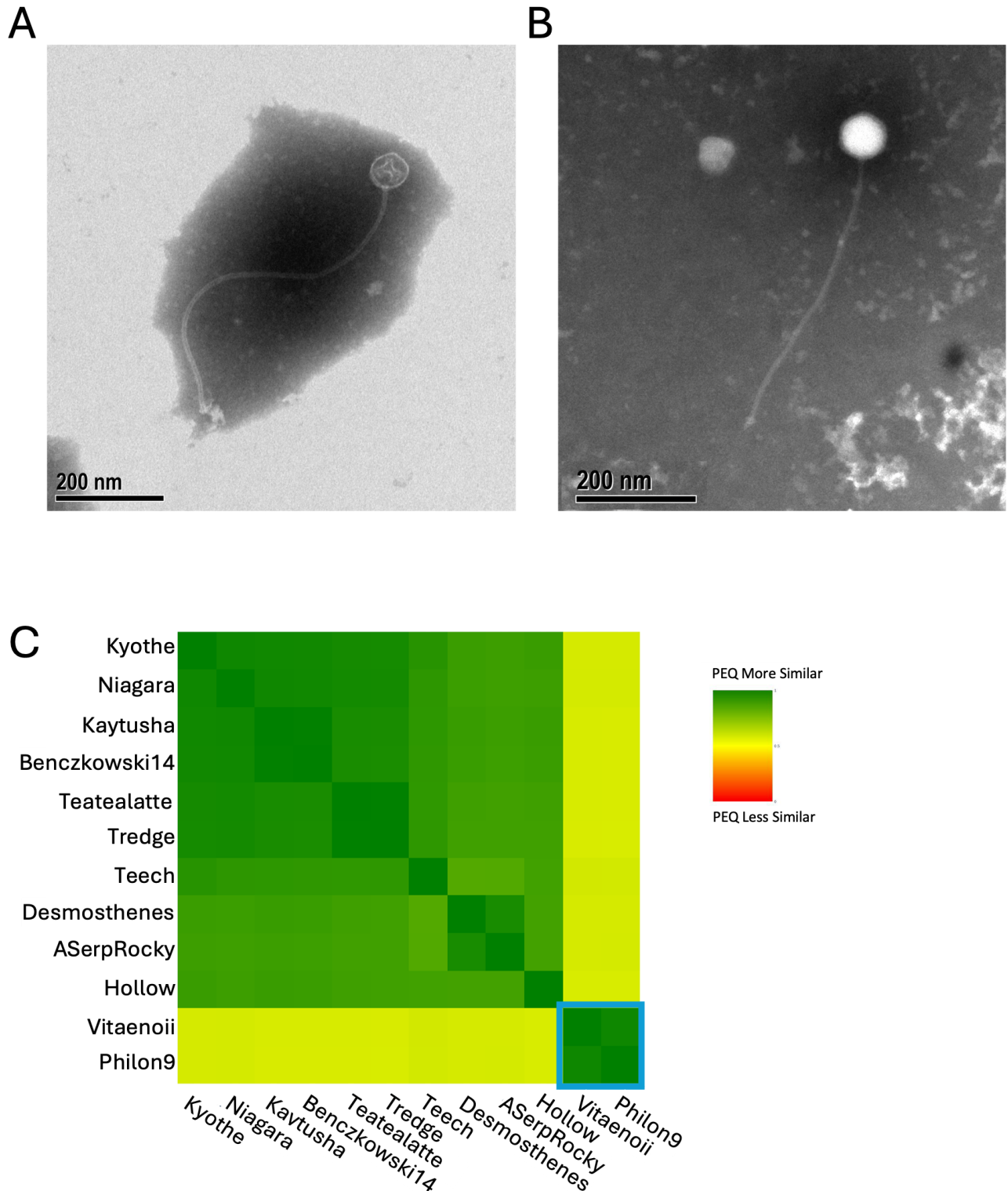
<sup>2</sup>Science and Math, Durham Technical Community College, Durham, North Carolina, United States

§To whom correspondence should be addressed: fogartym@durhamtech.edu

\*These authors contributed equally.

## Abstract

We report the discovery and genome sequences of bacteriophages Vitaenoi and Philon9, isolated from soil samples collected in Durham, North Carolina using the host *Gordonia rubripertincta* NRRL B-16540. The phages have siphoviral morphology and are assigned to bacteriophage subcluster CS4.



**Figure 1. Transmission electron microscopy and comparative genomics analysis of Vitaenoi and Philon9:**

Negative stain (2 % phosphotungstic acid, PTA) TEM images of Vitaenoi (A) and Philon9 (B) show siphoviral morphology, with long and flexible tails. Scale bar = 200 nm. C: Heatmap showing the structure of proteomic equivalence quotient (PEQ) among CS4 subcluster phage. Vitaenoi and Philon9 cluster together within the CS4 subcluster (blue box). The PhamClust workflow was used to calculate PEQ, defined as the extent of amino acid sequence similarity among shared genes within the CS4 subcluster, and to generate the resulting heatmap.

## Description

*Gordonia* are ubiquitous Gram-positive bacteria belonging to the phylum Actinobacteria. *Gordonia* species are found across diverse environments and hold potential for environmental and industrial biotechnological application (Arenskötter et al., 2004, Pope et al., 2017). Exploring the genetic diversity of bacteriophages that infect *Gordonia* species can advance our knowledge of these bacteria (Arenskötter et al., 2004). To contribute to our understanding of *Gordonia* bacteriophage diversity and evolution, we describe the isolation and characterization of two lytic bacteriophages that infect *Gordonia rubripertincta* strain NRRL B-16540, Vitaenoi and Philon9.

Vitaenoi and Philon9 were extracted from soil samples collected in Durham, NC (Table 1). Standard enrichment isolation procedures were used (Zorawik et al., 2024). Briefly, the soil samples were washed with peptone-yeast extract-calcium (PYCa) liquid medium to extract bacteriophages, and the resulting washes were filtered through 0.22  $\mu\text{m}$  filters. The filtrates were inoculated with *G. rubripertincta* NRRL B-16540 and incubated with shaking at 30°C for 2-5 days. Aliquots of the resulting cultures were filtered, spotted on PYCa top agar containing *G. rubripertincta* and incubated at 30°C for 3 – 5 days. Vitaenoi and Philon9 produced clear plaques with well-defined borders between 0.25 – 0.5 mm in size ( $n = 4$ ). Three rounds of individual plaque plating were carried out before lysates were prepared. All plaque assays were incubated for 2–5 days at 30°C. Lysates were negatively stained using 2 % phosphotungstic acid (PTA) and imaged by transmission electron microscopy (TEM) to reveal siphovirus morphologies with very long tails measuring over 400 nm in length (Table 1, Figure 1).

Genomic DNA was isolated from lysates using phenol-chloroform-isoamyl alcohol extraction (Sigma-Aldrich, P2069). Genome sequencing was performed by the Pittsburgh Bacteriophage Institute, using the NEB Ultra II FS kit for library preparation. For genome sequencing, Illumina NextSeq 1000 (with XLEAP-SBS P1 Kit) was used for Vitaenoi to generate 2,804,678 100 base single-end reads and Illumina MiSeq was used for Philon9, to generate 462,218 150 base single-end reads. For Vitaenoi, raw reads were trimmed with cutadapt 4.7 (using the option: `–nextseq-trim 30`) and filtered with skewer 0.2.2 (using the options: `–q 20 –Q 30 –n –l 50`) prior to assembly. Sequence reads were assembled using Newbler v2.9 and genomes were checked for accuracy and completion using Consed v.29 (Silva et al., 2013, Gordon and Green, 2013, Russell, 2018). Both phages have direct terminal repeat genome ends. The terminal repeat lengths are reported in Table 1, along with number of reads, average genome coverage and genome length, and GC content for each phage. Vitaenoi and Philon9 show 99% nucleotide identity as compared to 88% nucleotide identity with Niagara, the next most closely related phage outside the pair.

Gene annotation for each genome was performed using DNA Master v5.23.6 (Pope and Jacobs-Sera, 2018) and PECAAN (Reinhart et al., 2016), platforms integrating multiple bioinformatic gene prediction tools and genomic databases. The genomes were automatically annotated using Glimmer v3.02b (Delcher et al., 2007), and GeneMark v2.5p (Besemer & Borodovsky, 2005) to identify open reading frames and assess coding potential and potential gene start sites. Gene start site refinement and putative gene function assignments were determined using Starterator v1.2 (<http://phages.wustl.edu/starterator/>), Phamerator (Actino\_draft database v578 (Cresawn et al., 2011)), BLASTP searches against NCBI non-redundant and Actinobacteriophage databases (Altschul et al., 1990), and HHpred searches against the PDB\_mmCIF70, Pfam-A\_v36, UniProt-SwissProt-viral70\_3, NCBI\_Conserved\_Domains(CD)\_v3.19 databases (Zimmermann et al., 2018). Deep TMHMM v1.0.24 was used to detect putative transmembrane domains (Hallgren et al., 2022), and Aragorn v1.2.41 (Laslett and Canback, 2004) and tRNAscan-SE v2.0 (Lowe and Chan, 2016) were used for tRNA prediction. Default settings and parameters were used for all software, unless otherwise specified. The GC content of the Vitaenoi genome is 58.8%, and the Philon9 genome is 58.7%. Based on gene content similarity (GCS) of at least 35% to phages in the Actinobacteriophages database, phagesDB (Russell and Hatfull, 2017, Pope et al., 2017), Vitaenoi and Philon9 were assigned to bacteriophage cluster CS, and subcluster CS4. When compared to the additional 10 members of the CS4 subcluster, Vitaenoi and Philon9 show notable similarity to each other based on nucleotide sequence similarity, GCS and PhamClust analysis (Figure 1C and Table 1). PhamClust is a computational approach that efficiently calculates a proteomic equivalence quotient (PEQ) value for each bacteriophage pair based on amino acid sequence identity of shared genes, enabling accurate clustering and subclustering. The resulting heatmap generated allows visualization of cluster structure, subclusters and potential outliers (Gauthier & Hatfull, 2023). As with other cluster CS4 bacteriophage, genes involved in structure and assembly are in the first third of the left arm of the genome and encoded on the forward strand. Genes involved in DNA metabolism (including CobT-like cobalamin biosynthesis protein, Cas4 family exonuclease, nucleoside deoxyribosyltransferase, and ASCE-ATPase) and replication (including DNA primase, DNA helicase, DNA polymerase III sliding clamp (beta) and DnaQ-like DNA Polymerase III subunit) are scattered among the rest of the genome and encoded on the reverse strand. As with previously characterized cluster CS bacteriophages, no integrase or immunity repressor functions could be identified, suggesting they are unlikely to establish lysogeny and consistent with our inability to raise lysogens using standard procedures.

### Nucleotide sequence accession numbers

Vitaenoi is available at GenBank with Accession No. [PV915899](https://www.ncbi.nlm.nih.gov/nuclseq/PRJNA15899) and Sequence Read Archive No. [SRX28943186](https://www.ncbi.nlm.nih.gov/sra/PRJNA15899). Philon9 is available GenBank with Accession No. [PX089653](https://www.ncbi.nlm.nih.gov/nuclseq/PX089653) and Sequence Read Archive No. [SRX28943174](https://www.ncbi.nlm.nih.gov/sra/PRJNA15899).

Table 1: Isolation and sequencing parameters, and phage characteristics

Bacteriophage	Vitaenoi	Philon9
GPS coordinates	36.075457 N, 79.007587 W (Eno River State Park Durham, NC 27705)	36.061757 N, 78.91479 W (Yard of home in Durham, NC 27704)
Plaque size (mm)	0.25-0.5mm (n = 4)	0.25-0.5mm (n = 4)
Capsid size (nm)	65.9± 5.9nm (n=7)	69.1± 7.9nm (n=4)
Tail length (nm)	489.0 ± 74nm (n = 7)	479.3 ± 5nm (n = 4)
Number of Reads	2,804,678	462,218
Average Fold Coverage	2804	73
Genome Length (bp)	74721	74736
Genome End Direct Terminal Repeat	1206 bp	1205 bp
GC content %	58.8%	58.7%
Nucleotide similarity	99%	99%
Gene Content Similarity (GCS)	100 %	100 %
Number of genes with predicted function / Total predicted genes	39/96	39/96

**Acknowledgements:** We thank Daniel Russel and Rebecca Garlena for sequencing and assembling the genomes and the HHMI SEA-PHAGES program for support. We thank the SEA Community for review of the manuscript. We thank Dr. Alex Broussard for discovery course co-instruction and Matthew Trn, Kathryn Shriver, Jasmin Perez Lopez and Izzy Alexander for lab support. We thank Kristen White and Jillian Madren at UNC-CH microscopy services laboratory for taking TEM micrographs. Vitaenoi was discovered by Dex Aucion and X Peterson and Philon9 was discovered by Jennifer Daniels and Alasha Roth.

## References

- Arenskötter M, Bröker D, Steinbüchel A. 2004. Biology of the metabolically diverse genus *Gordonia*. *Appl Environ Microbiol* 70(6): 3195-204. PubMed ID: [15184112](#)
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215(3): 403-10. PubMed ID: [2231712](#)
- Besemer J, Borodovsky M. 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res* 33(Web Server issue): W451-4. PubMed ID: [15980510](#)
- Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF. 2011. Phamerator: a bioinformatic tool for comparative bacteriophage genomics. *BMC Bioinformatics* 12: 395. PubMed ID: [21991981](#)

Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23(6): 673-9. PubMed ID: [17237039](#)

Gauthier CH, Hatfull GF. 2023. PhamClust: a phage genome clustering tool using proteomic equivalence. *mSystems* 8(5): e0044323. PubMed ID: [37791778](#)

Gordon D, Green P. 2013. Consed: a graphical editor for next-generation sequencing. *Bioinformatics* 29(22): 2936-7. PubMed ID: [23995391](#)

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* 10.1101/2022.04.08.487609. DOI: [10.1101/2022.04.08.487609](#)

Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 32(1): 11-6. PubMed ID: [14704338](#)

Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res* 44(W1): W54-7. PubMed ID: [27174935](#)

Pope WH, Jacobs-Sera D. 2018. Annotation of Bacteriophage Genome Sequences Using DNA Master: An Overview. *Methods Mol Biol* 1681: 217-229. PubMed ID: [29134598](#)

Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, et al., Hatfull GF. 2017. Bacteriophages of *Gordonia* spp. Display a Spectrum of Diversity and Genetic Relationships. *mBio* 8(4): 10.1128/mBio.01069-17. PubMed ID: [28811342](#)

Rinehart CA, Gaffney BL, Smith JR, Wood JD. 2016. PECAAN: phage evidence collection and annotation network user guide. *Western Kentucky University Bioinformatics and Information Science Center, Bowling Green, KY*.

Russell DA. 2018. Sequencing, Assembling, and Finishing Complete Bacteriophage Genomes. *Methods Mol Biol* 1681: 109-125. PubMed ID: [29134591](#)

Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. *Bioinformatics* 33(5): 784-786. PubMed ID: [28365761](#)

Silva GG, Dutilh BE, Matthews TD, Elkins K, Schmieder R, Dinsdale EA, Edwards RA. 2013. Combining de novo and reference-guided assembly with scaffold\_builder. *Source Code Biol Med* 8(1): 23. PubMed ID: [24267787](#)

Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, et al., Alva V. 2018. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J Mol Biol* 430(15): 2237-2243. PubMed ID: [29258817](#)

Zorawik M, Jacobs-Sera D, Freise AC, SEA-PHAGES, Reddi K. 2024. Isolation of Bacteriophages on Actinobacteria Hosts. *Methods Mol Biol* 2793: 273-298. PubMed ID: [38526736](#)

**Funding:** This work was funded by Durham Technical Community College Department of Science and Math.

**Conflicts of Interest:** The authors declare that there are no conflicts of interest present.

**Author Contributions:** Joshua T. McMenemy: formal analysis, investigation, writing - original draft, writing - review editing. Amanda I. Acra: investigation, writing - original draft, writing - review editing. Elisabeth A. Alexander: investigation, writing - original draft, writing - review editing. Vanessa K. Cordovez: investigation, writing - original draft, writing - review editing. Morgan M. Fitzpatrick: investigation, writing - original draft, writing - review editing. Dayanara Mejia: investigation, writing - original draft, writing - review editing. Chaeli H. Smith: investigation, writing - original draft, writing - review editing. Marie P. Fogarty: validation, writing - original draft, writing - review editing, supervision, investigation.

**Reviewed By:** Dustin Edwards, Maria Diane Gainey Dr.

**History:** Received November 25, 2025 **Revision Received** April 2, 2026 **Accepted** March 31, 2026 **Published Online** April 7, 2026 **Indexed** April 21, 2026

**Copyright:** © 2026 by the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Citation:** McMenemy JT, Acra AI, Alexander EA, Cordovez VK, Fitzpatrick MM, Mejia D, Smith CH, Fogarty MP. 2026. Complete Genome Sequence of two *Gordonia rubripertincta* cluster CS4 Bacteriophages, Vitaenoi and Philon9. *microPublication Biology*. [10.17912/micropub.biology.001957](#)