

## Complete genome sequences of *Rhodococcus equi* phages CoffeeBean, Dorin, Francesca, Madraxi, and Tonitrus

Orin R. Daspit<sup>1\*</sup>, Andrew C. Slagter<sup>1\*</sup>, Drew VanWoerkom<sup>1\*</sup>, Derby Effie Addison<sup>1</sup>, Samantha M. Arteaga<sup>1</sup>, Rory E. Barnard<sup>1</sup>, Nick M. Bultje<sup>1</sup>, Junu Chung<sup>1</sup>, Brenna M. Coleman<sup>1</sup>, Cameron D. Figg<sup>1</sup>, Audrey J. Gonzales<sup>1</sup>, Josiah M. Hoekman<sup>1</sup>, Kat H. Irmen<sup>1</sup>, Valerie G. Krause<sup>1</sup>, Natalie L. Luginbill<sup>1</sup>, Rebecca A. Nulty<sup>1</sup>, Alexander W. Scofield<sup>1</sup>, Marguerite C. Sytsema<sup>1</sup>, Charles E. Tavera<sup>1</sup>, Anna G. Thomas<sup>1</sup>, Simon P. Troyer<sup>1</sup>, Hannah Sheppard<sup>1</sup>, Myles D. Radersma<sup>1</sup>, John T. Wertz<sup>1</sup>, Randall J. DeJong<sup>1§</sup>

<sup>1</sup>Calvin University, Grand Rapids, Michigan, United States

<sup>§</sup>To whom correspondence should be addressed: rdejong@calvin.edu

\*These authors contributed equally.

### Abstract

We report genomes of five phages isolated using the actinobacterium *Rhodococcus equi* NRRL B-16538. Based on gene content similarity, one phage is assigned to actinobacteriophage cluster CF, one to cluster CR, two to cluster CG, and one that cannot be assigned to any existing cluster. The latter encodes a five-gene thymine hypermodification system.

Phage Name	CoffeeBean	Dorin	Francesca	Madraxi	Tonitrus
Soil sample collection site	Grand Rapids, MI, 42.931694 N, 85.581889 W	Grand Rapids, MI, 42.928267 N, 85.587137 W	Grand Rapids, MI, 42.93455 N, 85.58755 W	Grandville, MI, 42.897778 N, 85.737222 W	Grand Rapids, MI, 42.93155 N, 85.586913 W
Capsid size (nm) <sup>a</sup>	62-64	75-80	69-88	58-60	56-81
Tail length (nm) <sup>a</sup>	243-250	333-355	353-439	238-242	268-323
No. of 150 bp sequencing reads	2,000,000	2,000,000	2,000,000	856,305	482,745
Average shotgun coverage (fold)	194	104	296	1862	1081
Genome length (bp)	67362	136968	137748	64064	57971
Genome ends	3' single-stranded overhang, 5' AGCCGCGTAC	direct terminal repeat (4720 bp)	direct terminal repeat (4862 bp)	3' single-stranded overhang, 5' CCCGCC	circularly permuted
G + C content (%)	67.3	48.0	48.0	66.2	69.0
No. of ORFs	95	282	281	108	84
Cluster	CR	CG	CG	CF	Singleton
No. of tRNA-coding regions	0	33	33	0	0

**Table 1. Characteristics of five phages isolated from *Rhodococcus equi* NRRL B-16538:**

<sup>a</sup>Minimum 3 phage particles measured

## Description

*Rhodococcus equi* is an Actinobacteria commonly found in soil and known to cause lung infections in livestock and immunocompromised humans (Weinstock & Brown, 2002). Limited phages capable of infecting *Rhodococcus equi* have been documented (Summer et al., 2010; Bonilla et al., 2017; Radersma et al., 2024). Here, we report on five phages isolated in 2023 using *R. equi* NRRL B-16538 and standard methods (Zorawik et al., 2024). Soil samples collected from ~ 5 cm below the soil surface were washed in PYCa broth and filtered (0.22µm pore size). Directly plating the filtrate in PYCa top agar with *R. equi* yielded phages Coffeebean and Madraxi. Incubation of the filtrate with *R. equi* for 2 days at 30°C before refiltration and plating yielded Dorin, Francesca, and Tonitrus. All phages have siphovirus morphologies, as determined by transmission electron microscopy.

Phage DNA was extracted from lysate using a Qiagen DNeasy kit, then prepared for sequencing using the NEBNext Ultra II-FS DNA library prep kit. Sequencing was performed with an Illumina MiSeq (v3 reagents), yielding 150-base single-end reads that were assembled using Newbler v2.9 and Consed v29 with 104 to 1862-fold shotgun coverage (Miller et al., 2010; Gordon & Green, 2013; Russell, 2018). Genomes were then annotated using DNA Master v5.23.6 (Pope & Jacobs-Sera, 2018), PECAAN v20221109 (Rinehart et al., 2016), Glimmer v3.02 (Kelley et al., 2012), Genemark v2.5 (Besemer et al., 2001), Phamerator (Actino\_draft database v578) (Cresawn et al., 2011), BLASTp v2.14.1 (Actinobacteriophage and NCBI non-redundant protein databases) (McGinnis & Madden, 2004), HHPred (PDB, UniProt, Pfam-A v.36, and NCBI v.3.19 databases) (Söding et al., 2005), Aragorn (Laslett & Canback, 2004), tRNAscanSE v2.0 (Lowe & Eddy, 1997), DeepTMHMM v1.0.24

(Hallgren et al., 2022), and TOPCONS v2.0 (Tsirigos et al., 2015), all using default parameters. Sequencing details and genomic characteristics are provided in Table 1.

Phages were assigned to clusters based on gene content similarity (GCS) of at least 35 % (Pope et al., 2017) to phages in the actinobacteriophage database (<https://phagesdb.org/>). Phage Tonitrus did not meet this threshold and is therefore classified as a singleton. The first five genes of Tonitrus may be associated with a thymidine hypermodification (THM) system, including a parB-like nuclease domain, a tet-like J-binding protein, an aGPT-Pplase2 domain protein, a 5-hmU DNA kinase, and a hypothetical protein (Lee et al., 2018, 2022). This system may defend against bacterial restriction endonucleases through chemical modification of thymidine (Lee et al., 2018, 2022; Flodman et al., 2019). Phage CoffeeBean is assigned to cluster CR, which is principally composed of phages isolated using *Gordonia*; to date, only seven of forty-nine cluster CR phages were isolated using *R. equi* (Radersma et al., 2024). Phages Dorin and Francesca share 91% GCS, are assigned to cluster CG, and contain 33 tRNA-coding regions each. Phage Madraxi is assigned to cluster CF and is the only phage presumed to be a temperate phage based on the presence of identifiable serine integrase and immunity repressor functions; these functions could not be identified in the other four phages. Except for Tonitrus, a translational frameshift in the tail assembly chaperone protein was identified for the other phages reported here.

**Data availability.** Annotated genome sequences can be accessed for CoffeeBean, Dorin, Francesca, Madraxi, and Tonitrus at GenBank accession numbers [PP978888](#), [PQ114736](#), [PP978770](#), [PP978820](#), [PP978825](#), respectively. Sequence reads are deposited at NCBI under SRA accession numbers [SRX25029070](#), [SRX25029071](#), [SRX25029053](#), [SRX25029059](#), [SRX25029066](#).

**Acknowledgements:** We thank Lori Keen for her excellent assistance with the Calvin phage research course and the HHMI SEA-PHAGES program for collaborative support.

## References

- Besemer J. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Research* 29: 2607-2618. DOI: [10.1093/nar/29.12.2607](#)
- Bonilla JA, Isern S, Findley AM, Klyczek KK, Michael SF, Saha MS, et al., Hatfull. 2017. Genome Sequences of 19 *Rhodococcus erythropolis* Cluster CA Phages. *Genome Announcements* 5: 10.1128/genomea.01201-17. DOI: [10.1128/genomea.01201-17](#)
- 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: 10.1186/1471-2105-12-395. DOI: [10.1186/1471-2105-12-395](#)
- Flodman K, Tsai R, Xu MY, Corrêa IR, Copelas A, Lee YJ, et al., Xu. 2019. Type II Restriction of Bacteriophage DNA With 5hmU-Derived Base Modifications. *Frontiers in Microbiology* 10: 10.3389/fmicb.2019.00584. DOI: [10.3389/fmicb.2019.00584](#)
- Gordon D, Green P. 2013. *Consed*: a graphical editor for next-generation sequencing. *Bioinformatics* 29: 2936-2937. DOI: [10.1093/bioinformatics/btt515](#)
- 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. : 10.1101/2022.04.08.487609. DOI: [10.1101/2022.04.08.487609](#)
- Kelley DR, Liu B, Delcher AL, Pop M, Salzberg SL. 2011. Gene prediction with Glimmer for metagenomic sequences augmented by classification and clustering. *Nucleic Acids Research* 40: e9-e9. DOI: [10.1093/nar/gkr1067](#)
- Laslett D. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research* 32: 11-16. DOI: [10.1093/nar/gkh152](#)
- Lee YJ, Dai N, Müller SI, Guan C, Parker MJ, Fraser ME, et al., Weigele. 2021. Pathways of thymidine hypermodification. *Nucleic Acids Research* 50: 3001-3017. DOI: [10.1093/nar/gkab781](#)
- Lee YJ, Dai N, Walsh SE, Müller S, Fraser ME, Kauffman KM, et al., Weigele. 2018. Identification and biosynthesis of thymidine hypermodifications in the genomic DNA of widespread bacterial viruses. *Proceedings of the National Academy of Sciences* 115: 10.1073/pnas.1714812115. DOI: [10.1073/pnas.1714812115](#)
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence. *Nucleic Acids Research* 25: 955-964. DOI: [10.1093/nar/25.5.955](#)

- McGinnis S, Madden TL. 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research* 32: W20-W25. DOI: [10.1093/nar/gkh435](https://doi.org/10.1093/nar/gkh435)
- Miller JR, Koren S, Sutton G. 2010. Assembly algorithms for next-generation sequencing data. *Genomics* 95: 315-327. DOI: [10.1016/j.ygeno.2010.03.001](https://doi.org/10.1016/j.ygeno.2010.03.001)
- Pope WH, Jacobs-Sera D. 2017. Annotation of Bacteriophage Genome Sequences Using DNA Master: An Overview. *Methods in Molecular Biology, Bacteriophages* : 217-229. DOI: [10.1007/978-1-4939-7343-9\\_16](https://doi.org/10.1007/978-1-4939-7343-9_16)
- Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, et al., Hatfull. 2017. Bacteriophages of *Gordonia* spp. Display a Spectrum of Diversity and Genetic Relationships. *mBio* 8: 10.1128/mbio.01069-17. DOI: [10.1128/mbio.01069-17](https://doi.org/10.1128/mbio.01069-17)
- Radersma MD, Lathrop G, Moleakunnel KC, Harlow LA, Baker AE, Chen AJ, et al., DeJong. 2024. Complete genome sequences of nine *Rhodococcus equi* phages. *Microbiology Resource Announcements* 13: 10.1128/mra.01088-23. DOI: [10.1128/mra.01088-23](https://doi.org/10.1128/mra.01088-23)
- Rinehart, C. A., Gaffney, B., Wood, J. D., & Smith, J. (2016). *PECAAN, a Phage Evidence Collection And Annotation Network*. <https://discover.kbrinsgd.org/login>
- Russell DA. 2017. Sequencing, Assembling, and Finishing Complete Bacteriophage Genomes. *Methods in Molecular Biology, Bacteriophages* : 109-125. DOI: [10.1007/978-1-4939-7343-9\\_9](https://doi.org/10.1007/978-1-4939-7343-9_9)
- Soding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* 33: W244-W248. DOI: [10.1093/nar/gki408](https://doi.org/10.1093/nar/gki408)
- Summer EJ, Liu M, Gill JJ, Grant M, Chan-Cortes TN, Ferguson L, et al., Young. 2011. Genomic and Functional Analyses of *Rhodococcus equi* Phages ReqiPepy6, ReqiPoco6, ReqiPine5, and ReqiDocB7. *Applied and Environmental Microbiology* 77: 669-683. DOI: [10.1128/AEM.01952-10](https://doi.org/10.1128/AEM.01952-10)
- Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Research* 43: W401-W407. DOI: [10.1093/nar/gkv485](https://doi.org/10.1093/nar/gkv485)
- Weinstock DM, Brown AE. 2002. *Rhodococcus equi*: An Emerging Pathogen. *Clinical Infectious Diseases* 34: 1379-1385. DOI: [10.1086/340259](https://doi.org/10.1086/340259)
- Zorawik M, Jacobs-Sera D, Freise AC, SEA-PHAGES, Reddi K. 2024. Isolation of Bacteriophages on Actinobacteria Hosts. *Methods in Molecular Biology, Phage Engineering and Analysis* : 273-298. DOI: [10.1007/978-1-0716-3798-2\\_17](https://doi.org/10.1007/978-1-0716-3798-2_17)

**Funding:** We thank the Star and Talen families for their generous support of the phage research course at Calvin University.

**Author Contributions:** Orin R. Daspit: investigation, writing - original draft, writing - review editing. Andrew C. Slagter: investigation, writing - original draft, writing - review editing. Drew VanWoerkom: investigation, writing - original draft, writing - review editing. Derby Effie Addison: investigation, writing - original draft. Samantha M. Arteaga: investigation, writing - original draft. Rory E. Barnard: investigation, writing - original draft. Nick M. Bultje: investigation, writing - original draft. Junu Chung: investigation, writing - original draft. Brenna M. Coleman: investigation, writing - original draft. Cameron D. Figg: investigation, writing - original draft. Audrey J. Gonzales: investigation, writing - original draft. Josiah M. Hoekman: investigation, writing - original draft. Kat H. Irmen: investigation, writing - original draft. Valerie G. Krause: investigation, writing - original draft. Natalie L. Luginbill: investigation, writing - original draft. Rebecca A. Nulty: investigation, writing - original draft. Alexander W. Scofield: investigation, writing - original draft. Marguerite C. Sytsema: investigation, writing - original draft. Charles E. Tavera: investigation, writing - original draft. Anna G. Thomas: investigation, writing - original draft. Simon P. Troyer: investigation, writing - original draft. Hannah Sheppard: investigation. Myles D. Radersma: investigation. John T. Wertz: investigation. Randall J. DeJong: investigation, writing - original draft, writing - review editing.

**Reviewed By:** Anonymous

**History:** Received December 31, 2024 **Revision Received** March 14, 2025 **Accepted** March 11, 2025 **Published Online** March 16, 2025 **Indexed** March 30, 2025

**Copyright:** © 2025 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:** Daspit OR, Slagter AC, VanWoerkom D, Addison DE, Arteaga SM, Barnard RE, et al., DeJong RJ. 2025. Complete genome sequences of *Rhodococcus equi* phages CoffeeBean, Dorin, Francesca, Madraxi, and Tonitrus. *microPublication*

---

3/16/2025 - *Open Access*

Biology. [10.17912/micropub.biology.001491](https://doi.org/10.17912/micropub.biology.001491)