Complete Genome Sequence of *Gordonia* phage RedRaider

Peyton Abraham[1*], Princess Bola-Lawal[1*], Mary Britten[1*], Randi Childress[1*], Abby Flannagan[1*], Mallory Gramstad[1*], Hayden Groos[1*], Sharon Huffman[1*], Korver Hupke[1*], Corbin Kleis[1*], Jenna Kluxdal[1*], Ayda Maassen[1*], Caden Meyer[1*], Elsa Meyer[1*], Logan Miller[1*], Emma Steinhardt[1*], Travis Sweeney[1*], Alexa Trover[1*], Shayler Van Gelder[1*], Dane Schoenborn[1*], Byron Noordewier[1*], Sara Tolsma[1§*]

[1]Biology, Northwestern College - Iowa, Orange City, Iowa, United States

[§]To whom correspondence should be addressed: stolsma@nwciowa.edu

[*]These authors contributed equally.

Abstract

We isolated and annotated a novel bacteriophage, RedRaider, on *Gordonia terrae* 3612. Annotation revealed genome features typical of cluster CR3 phages with the putative lysin A function encoded in two gene products, a protease C39 domain and a glycosyl hydrolase domain, and lysin B encoded distally in the left arm of the genome.
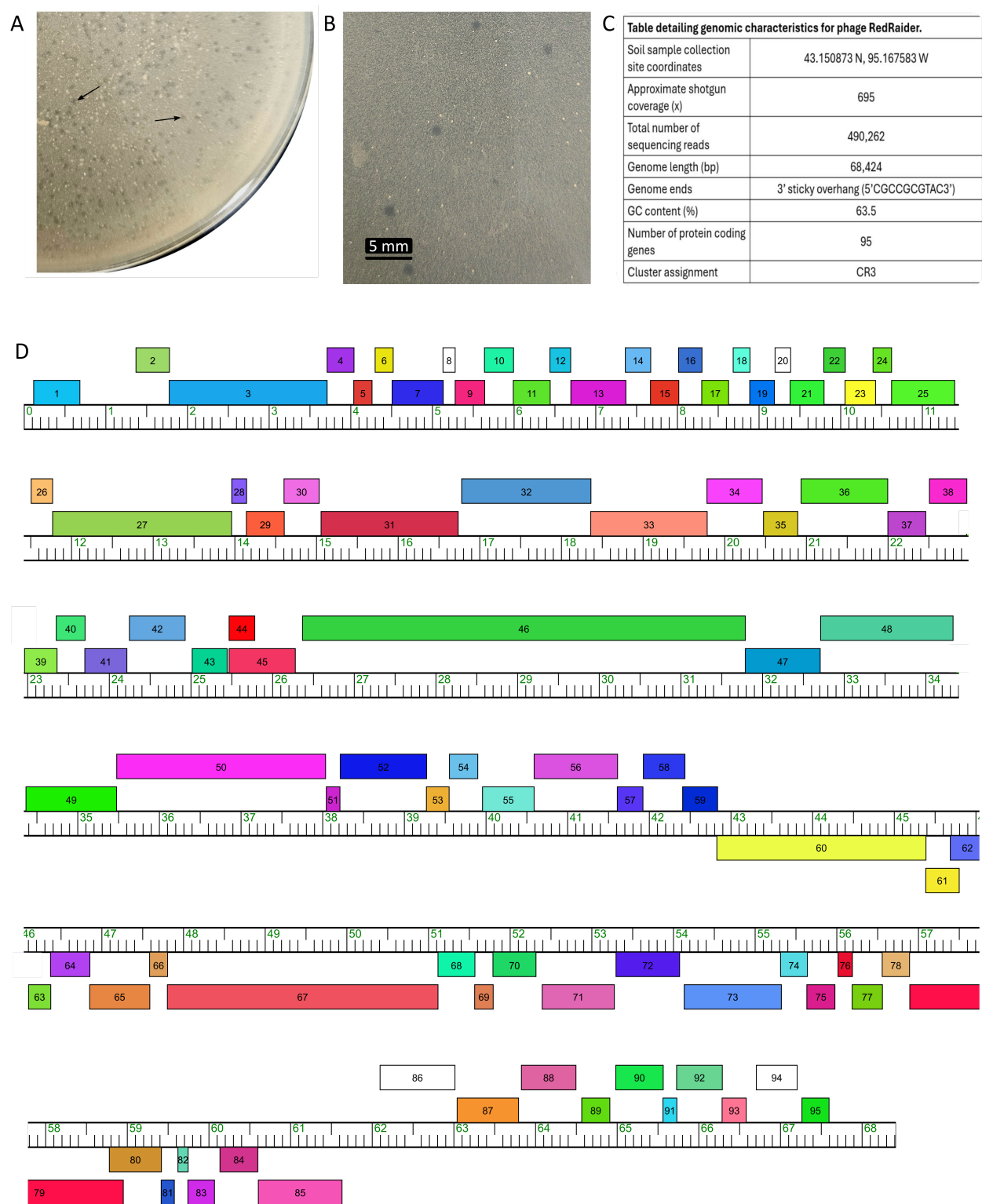
Figure 1. Phage RedRaider:

A. Picture of 1/4 of a 100 mm whole plate plaque assay of phage RedRaider. Arrows identify representative plaques (dark) to distinguish plaques from bacterial clumps in the lawn which appear white. B. RedRaider plaque assay at lower dilution for detailed plaque morphology. Plaques were clear with turbid, irregular edges and ranged from 0.8 to 2.3 mm in diameter (n = 10) with a mean diameter of 1.2 mm. Scale bar = 5 mm. C. Genomic characteristics of phage RedRaider. D. Phamerator map of RedRaider. Colored boxes indicate annotated genes. The ruler indicates the position in the genome. Boxes above the ruler represent genes in the forward direction; boxes below the line represent genes in the reverse direction.

Description

Bacteriophages are a rapidly evolving group of infectious particles (Hendrix 2002). They are remarkably abundant, found in nearly every ecosystem investigated (Rohwer et al., 2014). In this paper, we report the genomic organization of RedRaider, a newly discovered member of this dynamic population.

We isolated RedRaider from 15 g of very dry, surface topsoil that was near a patch of grass clippings using standard methods (see Figure 1C for GPS coordinates) (Poxleitner et al, 2018). Briefly, the soil sample was washed with peptone-yeast extract-calcium (PYCa) liquid medium supplemented with 0.1% dextrose and the suspension was filtered (0.2-µm pore size). The filtrate was inoculated with *Gordonia terrae* 3512 and incubated, with shaking, at 30°C for two days. The resulting culture was filtered (0.2- µm pore size) to remove bacteria and phages in the filtrate were plated in PYCa top agar with *G. terrae*. Plaques of RedRaider were observed after 24 hours and purified with three rounds of plating. RedRaider formed small plaques, 0.8 - 2.3 mm in diameter (n = 10), with slightly turbid, irregular edges after 24-48 hours at 30°C (Figure 1A, B).

We isolated phage DNA from high titer lysate using Wizard DNA Clean-Up System (Promega) then concentrated the DNA with a DNA Clean and Concentrator Kit (ZYMO Research). RedRaider DNA was pooled with DNA isolated from a mycobacteriophage and a microbacterium phage and then prepared for sequencing on an Illumina NextSeq 1000 sequencer using the NEB Ultra II Library kit. The resulting 100 base 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 (Martin 2011; Jiang et al, 2014; Wick et al, 2017; Gordon et al, 1998). Genome ends were determined with Consed (v29.0) using default settings (Gordon and Green 2013; Russell 2018). The DNA sequences from each pooled sample were substantially different, which enabled the assembly of three distinct genomes, one of which was RedRaider (Russell 2018). RedRaider was assigned to cluster CR3 based on gene-content similarity (GCS) of 35% or higher to sequenced bacteriophages present in the Actinobacteriophage database, phagesDB (Russell and Hatfull 2017) using the phagesDB GCS tool and previously described criteria (Hatfull et al., 2010). Sequencing details and genome characteristics are reported in Figure 1C.

RedRaider's genome was autoannotated using DNA Master v5.23.2 (cobamide2.bio.pitt.edu) with embedded Glimmer v3.02 (Delcher et al., 2007) and Genemark v2.5 (Besemer and Borodovsky 2005). The autoannotation was manually refined using these tools along with several others. Start site conservation was analyzed using Starterator (hhttp://phages.wustl.edu/starterator/). BLAST, using the Actinobacteriophage and NCBI non-redundant database (Altschul et al., 1990), Phamerator using Actino draft database v578 (Cresawn et al., 2011), and HHPRED, using the PDB_mmCIF70, Pfam- v.36, NCBI Conserved Domains databases (Söding et al., 2005) were used to make putative functional calls. We used Aragorn v1.1 (Laslett et al., 2004) and tRNAscan-SE v2.0 (Lowe et al., 1997) to scan for tRNA genes and deepTMHMM to identify putative transmembrane proteins (Hallgren et al., 2022). Many of the programs we used embedded in PECAAN (discover.kbrinsgd.org) and we used default parameters for all software programs. Our annotation identified 95 putative protein-coding genes and no tRNA genes (Figure 1C, D).

RedRaider's genome organization is typical of bacteriophages with structural genes located in the left arm of the genome while genes encoding transcription factors and nucleic acid modification enzymes are primarily located in the right arm of the genome (Hatfull and Hendrix 2011). Genes are transcribed in both directions in the genome with 69 genes transcribed in one direction and the remaining 26 transcribed in the opposite direction (Figure 1D). We annotated four orphams, which are genes for which no homologues exist within the Actinobacteriophage database: genes 8, 20, 86, and 94.

Like the other CR3 cluster phages, RedRaider has a non-canonical lysis cassette organization (Pimentel 2014). The putative lysin B gene (gp25) is located in the left arm of the genome, upstream of genes encoding putative structural proteins such as major capsid, tape measure, and minor tail proteins. The putative lysin A gene (gp 55 and gp 56) is located downstream of the putative structural genes, as expected, and is encoded by two gene products, one encoding the protease C39 domain and another the glycosyl hydrolase domain (Figure 1D). We were unable to identify a gene encoding holin function but noted three putative gene products (gp57, gp58, and gp59) with predicted transmembrane domains encoded immediately downstream of the two lysin A genes (Bavda et al., 2021).

RedRaider represents only the sixth cluster CR3 phage from the Actinobacteriophage database discovered making this work an important contribution to our growing understanding of phage diversity, genomic structure, and evolution (Dion et al., 2020).

Data Availability.

RedRaider is available at GenBank with Accession No. PV329801 and Sequence Read Archive (SRA) No. SRX28484027.
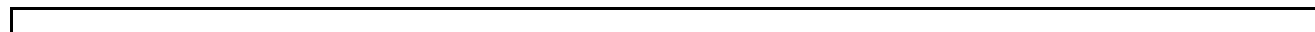
References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. Journal of Molecular Biology 215: 403-410. DOI: 10.1016/S0022-2836(05)80360-2

Bavda VR, Yadav A, Jain V. 2021. Decoding the molecular properties of mycobacteriophage D29 Holin provides insights into Holin engineering. J Virol 95(10): pii: JVI.02173-20. 10.1128/JVI.02173-20. PubMed ID: 33627396

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

Dion MB, Oechslin F, Moineau S. 2020. Phage diversity, genomics and phylogeny. Nat Rev Microbiol 18(3): 125-138. PubMed ID: 32015529

Gordon D, Abajian C, Green P. 1998. Consed: a graphical tool for sequence finishing. Genome Res 8(3): 195-202. PubMed ID: 9521923

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. : 10.1101/2022.04.08.487609. DOI: 10.1101/2022.04.08.487609

Hatfull GF, Hendrix RW. 2011. Bacteriophages and their genomes. Curr Opin Virol 1(4): 298-303. PubMed ID: 22034588

Hatfull GF, Jacobs-Sera D, Lawrence JG, Pope WH, Russell DA, Ko CC, et al., Hendrix. 2010. Comparative Genomic Analysis of 60 Mycobacteriophage Genomes: Genome Clustering, Gene Acquisition, and Gene Size. Journal of Molecular Biology 397: 119-143. DOI: 10.1016/j.jmb.2010.01.011

Hendrix RW. 2002. Bacteriophages: Evolution of the Majority. Theoretical Population Biology 61: 471-480. DOI: 10.1006/tpbi.2002.1590

Jiang H, Lei R, Ding SW, Zhu S. 2014. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics 15: 182. PubMed ID: 24925680

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

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/25.5.0955

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17: 10. DOI: 10.14806/ej.17.1.200

Pimentel M. 2014. Genetics of Phage Lysis. Microbiology Spectrum 2: 10.1128/microbiolspec.mgm2-0017-2013. DOI: 10.1128/microbiolspec.mgm2-0017-2013

Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull GF. 2018. HHMI SEA-PHAGES Phage Discovery Guide https://seaphagesphagediscoveryguide.helpdocsonline.com/home.

Rohwer F, M Youle, H Maughan and N Hisakawa. 2014. Life in Our Phage World. San Diego, CA: Wholon. ISBN 978-0-990-49430-0.

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

Russell DA. 2018. Sequencing, Assembling, and Finishing Complete Bacteriophage Genomes. Methods Mol Biol 1681: 109-125. PubMed ID: 29134591

microPublication
BIOLOGY µP

Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res 33(Web Server issue): W244-8. PubMed ID: 15980461

Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13(6): e1005595. PubMed ID: 28594827

Author Contributions: Peyton Abraham: investigation. Princess Bola-Lawal: investigation. Mary Britten: investigation. Randi Childress: investigation. Abby Flannagan: investigation. Mallory Gramstad: investigation. Hayden Groos: investigation. Sharon Huffman: investigation. Korver Hupke: investigation. Corbin Kleis: investigation. Jenna Kluxdal: investigation. Ayda Maassen: investigation. Caden Meyer: investigation. Elsa Meyer: investigation. Logan Miller: investigation. Emma Steinhardt: investigation. Travis Sweeney: investigation. Alexa Trover: investigation. Shayler Van Gelder: investigation. Dane Schoenborn: investigation. Byron Noordewier: investigation. Sara Tolsma: validation, investigation, visualization, writing - review editing.

Reviewed By: Susan Gurney, Anonymous