



Annotation of Two Soil Mycobacteriophages: **JacoRen57** and **DrLupo**

Ali Almail, Kelsey Dorhout, Jackson Johnson, Hannah Jorgensen, Byron Noordewier, and Sara S. Tolsma
Northwestern College, Department of Biology, Orange City, IA 51041

Abstract

We discovered two novel Mycobacteriophages from soil samples. **JacoRen57** is a singleton, most closely related to cluster AB phages. **DrLupo** is a member of the rare H2 sub-cluster containing only one other member. Both phages exhibit a siphoviridae morphotype, double-stranded DNA genome, and a non-contractile tail. The genome of **JacoRen57** is composed of 52,411 base pairs with a 56.7% GC content and includes genes and plaque morphology that suggest a lytic life cycle. The genome is organized according to the following order from the left to right arm: 33 forward genes followed by 16 reverse genes and 24 forward genes. We identified functions for 33 of the 73 protein coding genes using bioinformatics software. The 70,030 base-pair genome of **DrLupo** has a 57.5% GC content and contains genes that support a lytic life cycle. We identified 110 forward genes and assigned functions to 25 of them.

Introduction

Bacteriophages are a genetically diverse group of infectious particles (Hatfull, 2015). The advent of economical and rapid sequencing technology has made it possible to compare and analyze the genomes of various phages. The analysis reveals the mosaic genome composition between phages within a cluster due to horizontal gene transfer (Hatfull, 2015). Through the SEA-PHAGES program, students have the opportunity to annotate phage genomes, uncovering the diversity and evolutionary history of bacteriophages.

Northwestern College students discovered **JacoRen57** and **DrLupo**, which are two mycobacteriophages that infect *Mycobacterium smegmatis* mc²155, during the spring semester of 2018. Their genomes were sequenced and then annotated during the fall semester of 2018 by the students in the Genetics and Genomics course.

Materials and Methods

JacoRen57 was discovered by Tanner Rensink and **DrLupo** was discovered by Connor Tupper through the direct or enriched isolation protocol using the host organism *Mycobacterium smegmatis* mc²155 (Discovery Guide Protocol 5.2, 5.5).

The phages were amplified, purified, and characterized according to the Discovery Guide protocols. Phage DNA was isolated from a high titer phage solution and characterized using a restriction digest. Viruses were sequenced using Illumina Sequencing Technology at the Pittsburgh Bacteriophage Institute.

DNA Master was the primary platform for the annotation of **JacoRen57**. Auto-annotation was performed using DNA Master software, Starterator, Phamerator (www.phamerator.org), NCBI BLAST (<https://www.ncbi.nlm.nih.gov>), GeneMark, Glimmer, Phagesdb (<http://phagesdb.org/>), and HHPred (<https://toolkit.tuebingen.mpg.de/hhpred>) were used to identify start, stop, transcriptional direction, and function of genes. **DrLupo** was annotated using online software PECAAN (<https://discover.kbrinsgd.org>), which streamlined the process by directly retrieving data from various bioinformatics software such as HHPred, NCBI BLAST, and Phamerator.

References

Hatfull, G. F. (2015). Dark Matter of the Biosphere: the Amazing World of Bacteriophage Diversity. *Journal of Virology*, 89(16), 8107–8110. <https://doi.org/10.1128/JVI.01340-15>
Almail, A., Dorhout, K.E., Johnson, J., Jorgensen, H.J., Tolsma, S., Garlena, R.A., Russell, D.A., Pope, W.H., Jacobs-Sera, D., and Hatfull, G.F. (2019) Mycobacterium phage DrLupo. Complete Genome. Accession number MK279909.1. <https://www.ncbi.nlm.nih.gov/genbank/MK279909>
Almail, A., Dorhout, K.E., Johnson, J., Jorgensen, H.J., Tolsma, S., Garlena, R.A., Russell, D.A., Pope, W.H., Jacobs-Sera, D., and Hatfull, G.F. (2019) Mycobacterium phage JacoRen57. Complete Genome. Accession number MK279840.1. <https://www.ncbi.nlm.nih.gov/genbank/MK279840>

Support for this project was provided by the Howard Hughes Medical Institute, the Hatfull Lab at the University of Pittsburgh, SEA-PHAGES, and Northwestern College.

Results: **JacoRen57**

JacoRen57 is a singleton most closely related to cluster AB phages. The genome is composed of forward and reverse genes, with structural genes on the left arm, and genes involved in DNA replication and expression on the right arm. We resolved start site disagreements, removed, and added genes if we disagreed with the auto-annotation. We assigned functions to 33 of its 73 genes, and some of the functions, such as Lysin A, Lysin B, and Holin, suggest a lytic life cycle.

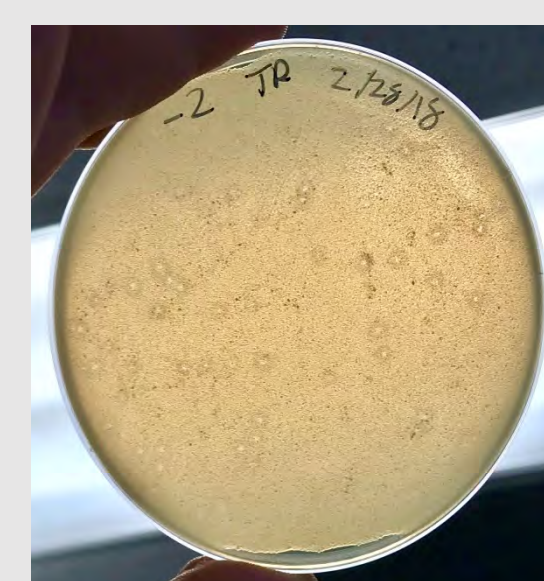


Figure 1A: Plaque picture of **JacoRen57**. The plaque morphology is small with light interior and turbid exterior.

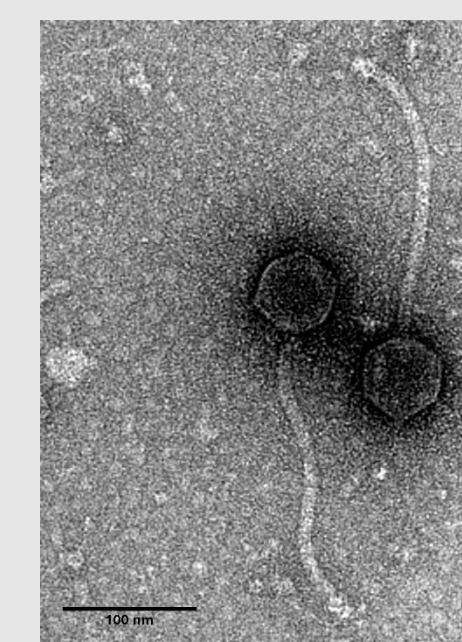


Figure 1B: Transmission electron micrograph of **JacoRen57** showing its siphoviridae morphology.

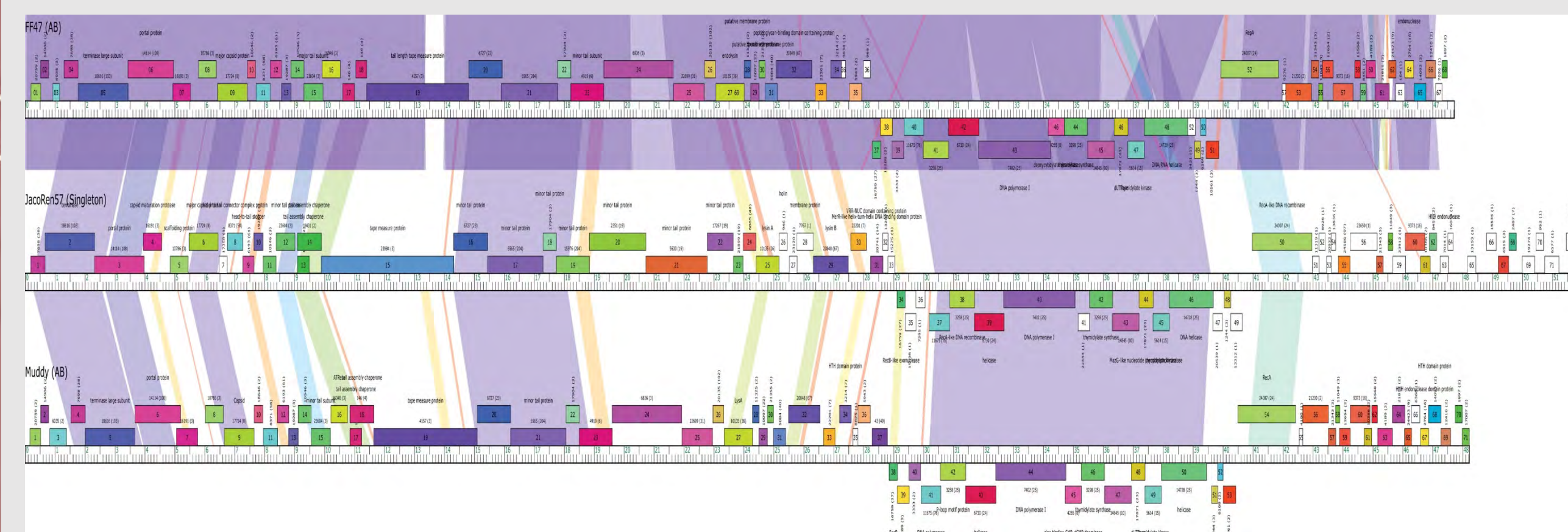


Figure 1C: Phamerator map comparing sequence similarities between **JacoRen57** and the two cluster AB phages: FF47 and Muddy.

Gene(s)	Function
2	Terminase
3	Portal Protein
4	Capsid Maturation Protease
5	Scaffolding Protein
6	Major Capsid Protein
8	Head-to-tail Connector Complex Protein
9	Head-to-tail Stopper
13, 14*	Tail Assembly Chaperone
15	Tape Measure Protein
12, 16, 17, 18, 19, 20, 21, 22	Minor Tail Protein
25	Lysin A
26	Holin
28	Membrane Protein
29	Lysin B
31	MerR-like Helix-turn-helix DNA Binding Domain Protein
34	RecB-like Exonuclease
38	RecA-like DNA Recombinase
39	Helicase
40	DNA Polymerase I
42	Thymidylate Synthase
44	MazG-like Nucleotide Pyrophosphohydrolase
45	Thymidylate Kinase
46	DNA Helicase
50	RecA-like DNA Recombinase
63	HNH Endonuclease

Table 1A: Genes for which we were able to call functions in **JacoRen57**. Genes 13/14* represent a single protein product due to a -1 programmed translational frameshift.

Sequencing Information	Results
Approximate Shotgun Coverage	693
Genome Length (bp)	52411
Character of Genome Ends	3' Sticky Overhang
Overhang Length (bp)	11

Table 1B: **JacoRen57** genome details.

Results: **DrLupo**

DrLupo belongs to the rare H2 sub-cluster and its genome contains 110 forward genes. We assigned functions to 25 of the 110 genes, and some of the functions, such as Holin, Lysin A, Lysin B, are indicative of a lytic life cycle. We resolved start site disagreements, added, and removed genes if we disagreed with the auto-annotation. We did not annotate a programmed translational frameshift in the tail assembly chaperone due to its non-canonical nature that is yet to be supported by evidence from the laboratory bench.

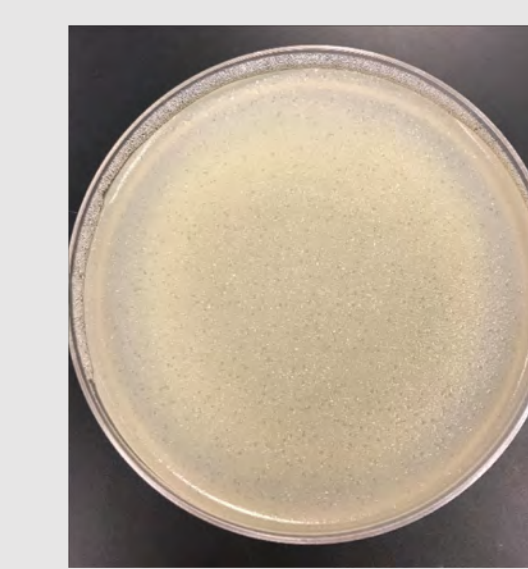


Figure 2A: Plaque picture of **DrLupo**. The plaque morphology is small with light interior and turbid exterior.

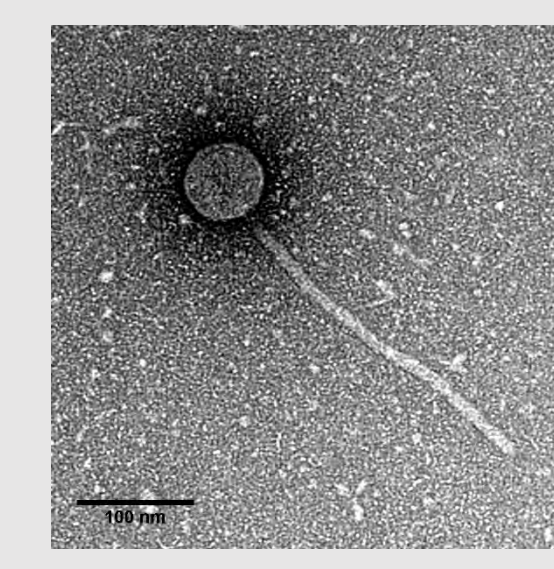


Figure 2B: Transmission electron micrograph of **DrLupo** showing its siphoviridae morphology.



Figure 2C: Phamerator map comparing sequence similarities between the sub-cluster H2 phages **DrLupo** and **Barnyard**.

Gene(s)	Function
6	Terminase
7	Portal Protein
18	Capsid Maturation Protease
19	Protease
20	MuF-like Minor Capsid Protein
23	Major Capsid Protein
29	Tail Terminator
32	Major Tail Protein
33,34	Tail Assembly Chaperone
35	Tape Measure Protein
36,37,38,39	Minor Tail Protein
41	Lysin A
42	Lysin B
43	Holin
67	Helicase
75	Exonuclease
79	RecA-like DNA Recombinase
85	DnaE-like DNA Polymerase III
91	RuvC-like Resolvase
97	Hydrolase
109	DNA Primase/ Polymerase

Table 2A: Genes for which we were able to call functions in **DrLupo**.

Sequencing Information	Results
Approximate Shotgun Coverage	195
Genome Length (bp)	70030
Character of Genome Ends	Circularly Permuted

Table 2B: **DrLupo** genome details.

Discussion

The genomes of **JacoRen57** and **DrLupo** exhibit mosaicism which is hypothesized to be due to horizontal gene transfer. **JacoRen57** is a singleton, and is most genetically similar to **Muddy** and **FF47**, both from the AB cluster. **DrLupo** exhibits sequence similarity to **Barnyard**. These two phages are the only known members of the rare H2 sub-cluster. Our work contributed to the growing understanding of phages. Our annotation results have passed peer review and were published in GenBank. The GenBank Accession Numbers for **JacoRen57** and **DrLupo** are MK279840 and MK279909 respectively.