Stormbreaker8 and A3Wally Bacteriophage Genome Annotations

Riley Chartier, Abigail Clarke, Lindsey Groen, Marshall Kleinhesselink, Cassandra Koel, Madison Korthals, Jordyn Kramer, Garrett Raymon, Krista Riensche, Sadie Sandbulte, Annika Stecker, Daniel Strand, Kaytlin Wojciechowski, Ali Almail, Byron Noordewier, and Sara S. Tolsma
Northwestern College, Department of Biology, Orange City, IA 51041

Abstract

Stormbreaker8 and A3Wally are two novel bacteriophages isolated and purified on Microbacterium foliorum NRRL B-24224 by students in the Fall 2020 Discovery course. Stormbreaker8, an EA1 cluster lytic phage, was isolated from soil collected in Orange City, IA. Its circular permuted genome contains 41,751 base-pairs with 63.4% GC content. A3Wally, a GD cluster phage, was isolated from soil collected in Sioux Center, IA. Its genome is 60.1% GC, contains 194,724 base-pairs, and its ends are direct terminal repeats. Spring 2021 Genetics students annotated the genomes using bioinformatics software.

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 a large number of phages. Sequencing analysis reveals the mosaic genome composition between phages within and between clusters due to horizontal gene transfer (Hatfull, 2015). Upon sequencing, phages that are genetically similar can be grouped into clusters and the linear sequence can be analyzed using bioinformatics software for the location of genes nad assigning of likely gene functions.

As part of the SEA-PHAGES program, students at Northwestern College annotate phage genomes, uncovering the diversity and evolutionary history of bacteriophages. We analyzed and annotated the genomes of two novel Actinobacteriophages, Stormbreaker8 and A3Wally.

Discussion

The genomes of Stormbreaker8 and A2Wally exhibit mosaicism which is hypothesized to be due to horizontal gene transfer. Our work contributes to the growing understanding of phages. Our annotations are nearly ready for submission to GenBank for publication.

Future work: We are especially curious about the large number of apparent gene duplications in A3Wally and wonder what selective advantage these duplications give a phage that needs to maintain such a large genome.

Materials and Methods

Stormbreaker8 was discovered by Annika Stecker and A3Wally was discovered by Abigail Clarke through the direct or enriched isolation protocol using the host organism *Microbacterium foliorum* NRRL B-24224 (*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 used for Illumina Sequencing Technology at the Pittsburgh Bacteriophage Institute.

DNA Master was the primary platform for the annotation of Stormbreaker. 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. A3Wally 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 Starterator.

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/nuccore/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/nuccore/MK279840.
D Jacobs-Sera, LA Abad, RM Alvey, KR Anders, HG Aull, SS Bhalla, ... (2020) Genomic diversity of bacteriophages infecting Microbacterium spp, Plos one 15 (6), e0234636

Results: Stormbreaker8

Stormbreaker8 was discovered in 2020 from soil collected in Orange City, IA. It infects the host *Microbacterium foliorum* NRRL B-24224. It is a Cluster EA1 phage. Its circularly permuted genome contains 41,751 bp with a GC content of 63.4%. Stormbreaker8 appears to be lytic by plaque morphology and we did not find an integrase gene when we annotated its genome further supporting the conclusion that this phage is lytic. Its genome contains 62 genes of which we were able to assign functions to 25.

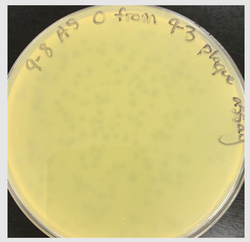


Figure 1. *Plaque picture of phage Stormbreaker8. Plaque morphology is medium-sized with a slightly cloudy appearance.*

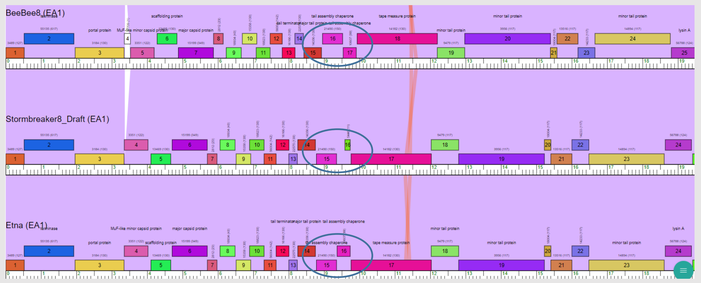


Figure 2. *Stormbreaker8 is most similar to Etna, Balsa, and BeeBee8, all of which are also cluster EA1 phages. Like other EA1 cluster phages, Stormbreaker8 has no clear, canonical slippery sequence in the genes for the tail assembly chaperone so we have not annotated a programmed translational frameshift.*

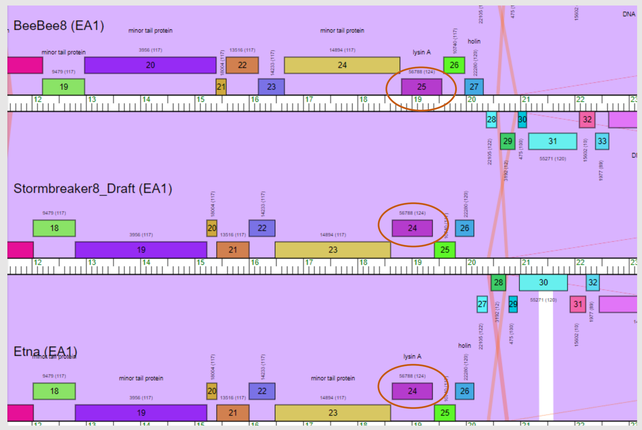


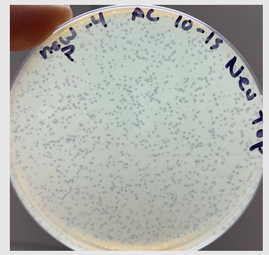
Figure 3. *Stormbreaker8 has no identifiable lysin B gene so the gene similar to lysin A is called as an endolysin. It is close to the holin as in typical phage synteny.*

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: A3Wally

A3Wally was discovered in 2020 in a direct isolation from soil collected in Sioux Center, IA. It infects the host *Microbacterium foliorum* NRRL B-24224. Its genome consists of a whopping 194,724 bp. Its ends are 15,503 bp direct terminal repeats and it has a relatively low GC content of 60.1%. It is a cluster GD phage. A3Wally appears to be lytic by plaque morphology and, when annotating, we did not find an integrase gene, further supporting this conclusion. A3Wally's genome contains 351 protein-coding genes and 44 tRNA genes, making it the largest phage ever discovered and annotated at Northwestern College. A3Wally is a cluster GD phage. To date only four GD cluster phages have been discovered, and only one of those, PauloDiabli, has a published annotation, making our annotation work a bit more difficult than other phages. Of A3Wally's 351 protein-coding genes, we were able to call functions for 66 of them.

Figure 4. *Plaque picture of A2Wally. The plaque morphology is small and largely clear.*



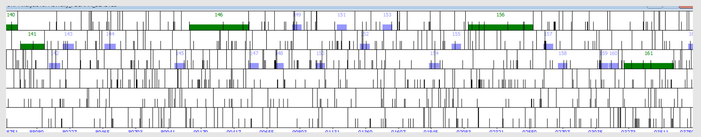


Figure 5. *The 44 tRNA genes are located in a large tRNA island near the center of the genome with only a few protein-coding genes scattered within the region. A portion of this large region is shown here. This is one of the regions of the genome that shows the most sequence divergence as compared to PauloDiabli, the only other annotated Cluster GD phage.*

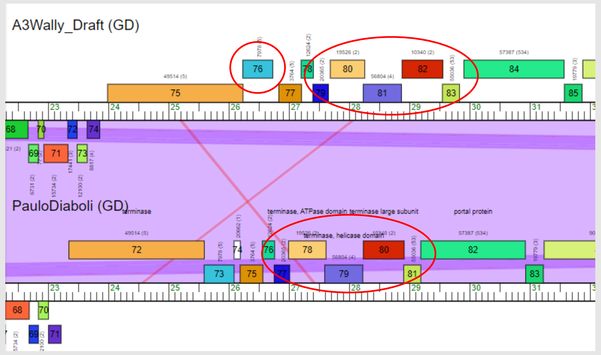


Figure 5. *A3Wally has two intact copies of the large subunit of the terminase, plus two more genes that encode large subunit of the terminase domains. PauloDiaboli also has this interesting duplication of terminase genes. We noticed additional genes in A2Wally that BLASTed to other A2Wally genes suggesting that having duplicated genes may be a characteristic of this very large phage and something we are interested in investigating further.*