The Quest for Novelty: A Tale of Four Hosts

Sierra C. Oates, Shelby E. Ferebee, Nana Y. Amoh, Zshonel Auma, Simeon A. Brown, Maiya C. Callender, Michael J. Crumpton, Shelsy E. Delgado Huamani, Kathryn R. Eckler, Holly A. Funkhouser, Victoria E. Guillen, Makeda N. Jackson, Katerina M. Kuzel, Yiying Liu, Ellie M. Martini, Alexander Mattei, Arianna Marie A. McFarlane, Alexandra G. Robles, Matthew D. Zachary, Yemi Zewdu Yimer, Elizabeth A. Allen, Albert Andrews, Olivia A. Awate, Sarah Belay, Sarah E. Modlin, Anja E. Nilsson, Caroline A. Golino, Andrew D. Halleran, William J. Buchser, Mark H. Forsyth, Margaret S. Saha

Department of Biology, College of William and Mary, Williamsburg, VA

hhmi science education alliance HOWARD HUGHES MEDICAL INSTITUTE

W&M Phage Lab

1. Abstract

In an effort to expand the diversity of phage discovery in the order Actinomycetales, the 2015-16 William & Mary Phage Lab attempted to isolate novel phages that infect *Rhodococcus erythropolis* or *Streptomyces virginiae*. Soil, compost, sludge, manure, and stream samples from approximately 100 different locations around coastal Virginia including swamps, livestock grazing areas, treatment plants, and runoffs, were used for both direct plating and enrichment throughout the extremely warm fall of 2015. Unfortunately, not a single phage was isolated from either host from any of the samples. Positive controls using previously isolated phage from these hosts worked as expected. After five weeks of working with *Rhodococcus erythropolis* and *Streptomyces virginiae*, we returned to *Mycobacterium smegmatis* as our host. Given the large number of Mycobacteriophage already identified, we conducted our infections at lower temperatures, namely 30C and 25C to optimize the possibility of novel isolates. A total of 20 phages were isolated, all following enrichment; three (Qobbit, Marcolius Prime, and Dismal Stressor) were sequenced by the University of Pittsburgh. Qobbit, a Cluster A9 phage isolated from moist soil in a local woods, displays ~90% identity with other A9 phages. Marcolius Prime and Dismal Stressor, both Cluster K2 phages, were isolated from the muddy banks of the Dismal Swamp in eastern North Carolina and differ only by four bases from each other. Both phage display 99% identity to Milly and 93% identity to TM4. In a continuing effort to isolate novel phages, we pursued two additional approaches. The first approach involved "crowd-sourcing" phage discovery via direct plating to the 480 student Introductory Biology lab where 27 mycobacteriophage were isolated this past year. The second approach involved attempting, during the spring semester, yet another alternative host, *Corynebacterium glutamicum*. Three phages from two different locations were isolated and are currently being sequenced.

2. Background

Goal: To expand the diversity of actinobacteriophage by discovering and characterizing unique phage through: (1) vary the conditions to isolate phage from *M. smegmatis mc²155* and (2) to employ a range of alternative hosts within the actinobacteria family: *Rhodococcus erythropolis; Streptomyces virginiae; Corynebacterium glutamicum.*

Research Questions:
- What is the genomic diversity of rhodococcus phage?
- Can novel phage be isolated from *M. smegmatis* using nonstandard conditions?
- Can underutilized hosts such as Corynebacteria serve as good models to expand actinobacteriophage diversity?

Hypotheses:
- Based on empirical and published data, the genomic diversity of Rhodococcus phages may be more limited.
- Based on rarefaction analysis (Pope et al., 2015), additional diversity of novel phage infecting M. smegmatis is likely and non-standard isolation conditions may result in novel phage infecting M. smegmatis.
- Underutilized hosts (Corynebacteria) will yield novel phage and expand actinobacteriophage diversity.

3. Methods

Hosts: *Four hosts were used; culture conditions followed standard Phages DB/ATCC protocols except as noted.*
1. *Rhodococcus erythropolis* (August-September, 2015)
2. *Streptomyces virginiae* (August-September, 2015)
3. *Mycobacterium smegmatis* (October, 2015); all procedures performed at 30°C.
   Experiment #1: Direct plating and enrichment (n=40)
   Experiment #2: Only direct plating (n=510)
4. *Corynebacterium glutamicum* (March-April, 2016).

Locations tested: Stream beds; compost piles; local woods; animal pastures; Dismal Swamp; soil from building excavations

Phage Isolation & Characterization (Standard SEA-PHAGES/PhagesDB protocol except where noted):

- Collected many soil samples from diverse environments (N=160 for *R. erythropolis; N=40 for S. virginiae; N=550 for M. smegmatis*; N=10 for *C. glutamicum)*
- Performed direct plating for ~20% of samples, except for *M. smegmatis*, 90% of the samples were used for direct plating
- Three rounds of plaque purification, Visualized phage with TEM
- Isolated DNA using PC method and analyzed via gel electrophoresis and Nanodrop to assess DNA concentration and purity
- Selected phages with highest concentration and quality and sent to U. Pittsburgh for sequencing
- Analyzed sequence using DNA Master and associated tools, HHPred, Phamerator
- For phages isolated following SEA Phages DNA submission deadline, phage DNA was sequenced in house on Illumina Mi-Seq

4. Results

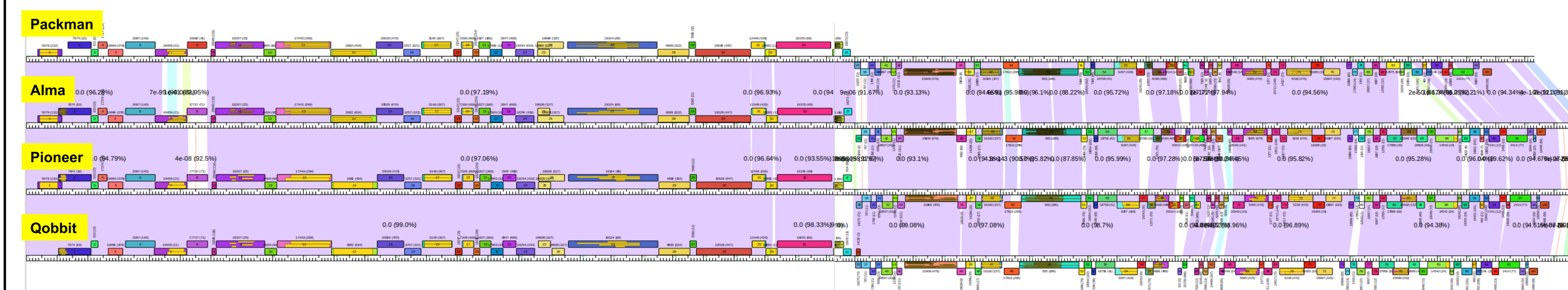
4a. Experiments with *R. erythropolis & S. virginiae* hosts

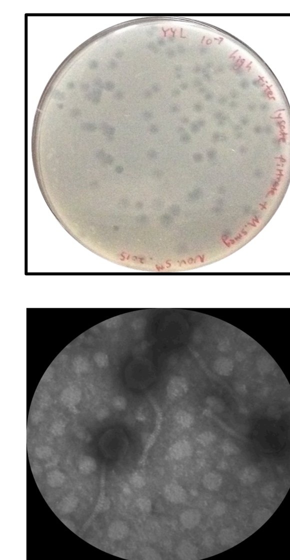
➡ No phage were obtained following repeated attempts by the entire class to isolate phage from diverse locations, either with direct plating or enrichment.

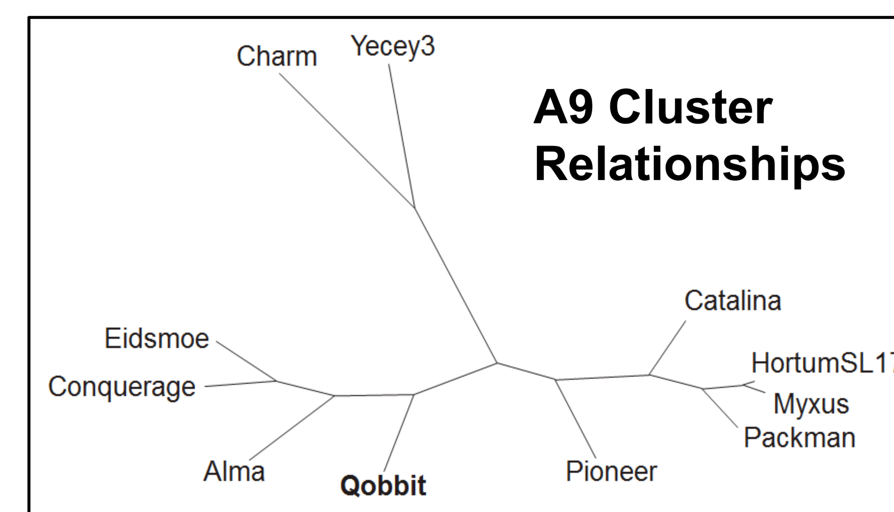
4b. Experiments with *M. smegmatis* as a host

➡ Experiment #1: Most 10/20 students obtained plaques following enrichment on their first or second attempt. Two were sequenced by the University of Pittsburgh: Marcoliusprime and Qobbit.

Qobbit: Cluster A9, 94 genes, 52,911 bp, 62.6% GC **(W&M Campus, VA)**







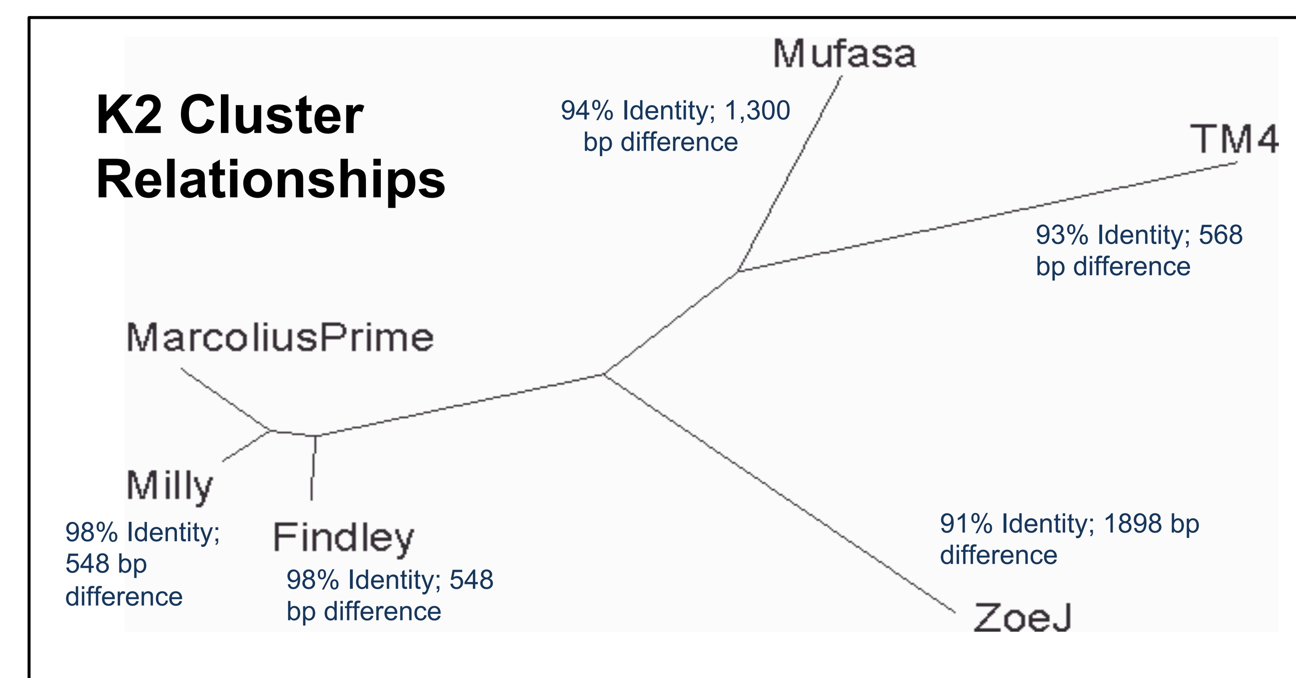
Marcoliusprime: Cluster K2, 92 genes, 58,129 bp, 68.2% GC **(Dismal Swamp, NC)**

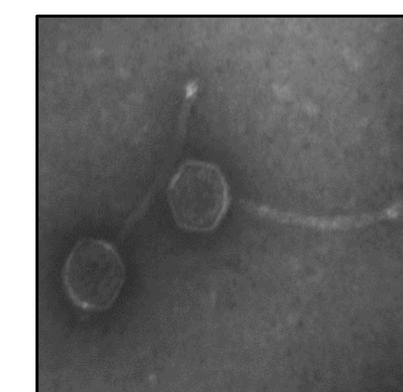
Genes for Infection:
Terminase; Lysin A ; Lysin B; Holin; Terminase small subunit; Integrase

Structural Genes:
Capsid maturation protease; Head-to-tail connector protein; Major capsid protein; Putative Phage Head protein; Portal protein; Major tail protein; Minor tail protein; Tail assembly chaperone; Tape measure protein

Other:
WhiB-like; SprT-like protein; Hydrolase; Putative Peptidase, Putative glucose dehydrogenase; Exonuclease; Putative Transcriptional Regulator; Putative Resolvase transcriptional regulator; YcfA-like protein;Glutaredoxin; tRNA-ribosyltransferase; Queuine tRNA-ribosyltransferase; HTH DNA binding domain; Glutaredoxin; Cro-like protein; Immunity repressor; NrdH-like glutaredoxin

DNA Replication Genes:
DNA binding domain protein; DNA Binding protein; DNA primase/ helicase ; Endonuclease; RusA Nuclease; Putative Nuclease; Transcription factor; Helicase; HNH endonuclease; Scaffolding protein

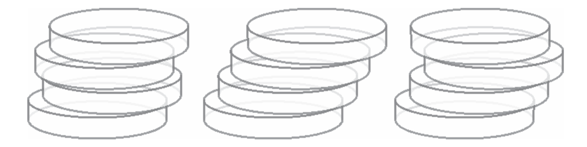




4b. Experiments with *M. smegmatis* as a host *(con't)*

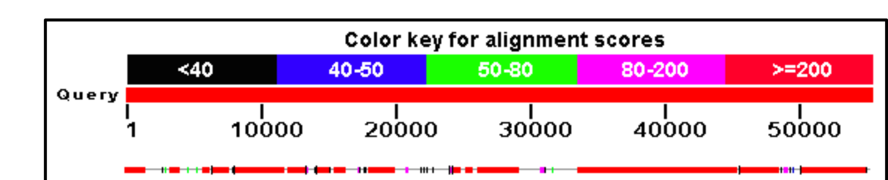
➡ Experiment #2: "Crowdsourcing Phage Discovery" in which 510 Introductory Biology students performed (only) direct plating, also at 30°C; 22/510 students obtained plaques.

Of the 15 phage genomes sequenced on the MiSeq platform, 13 shared 99% identity to:
- Larva – Subcluster K5 (9)
- CrimD – K1 (1)
- ShiLan – F1 (1)
- Redno2 – J (1)
- Wally – C1 (1)



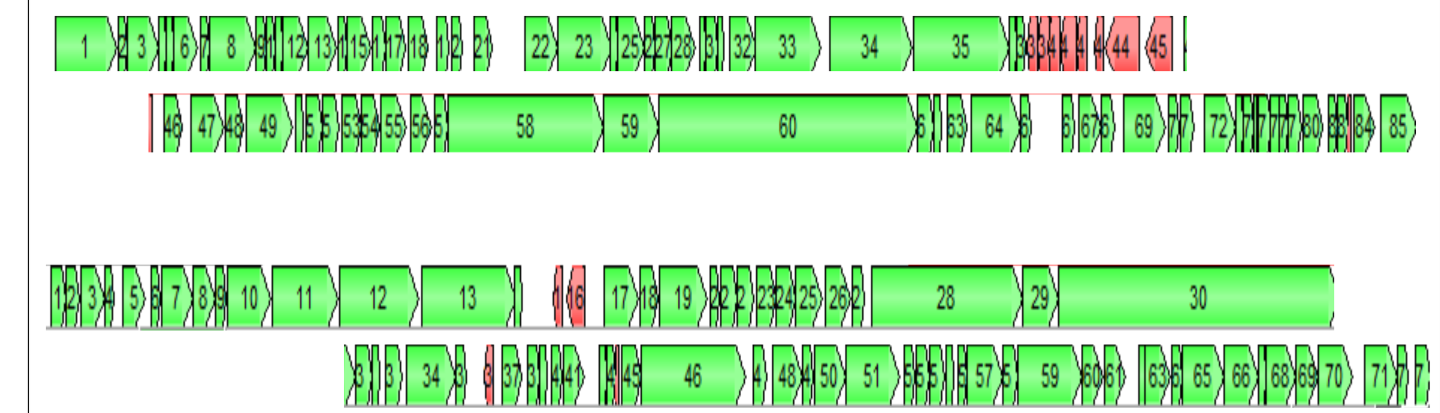
Two phages displayed ~79% identity with Enkosi (subcluster K1) over 75% of the genome..

BLAST of Phage #13 against Enoski



4c. Experiments with *C. glutamicum* as a host

While direct plating did not produce plaques, enrichment yielded three plaques. Illumina sequencing (40,000 reads) revealed two different phages consisting of 55,601 and 52,443 bp genomes.



BLAST searches of entire phage genome did not produce any significant homologies; BLAST searches using individual coding regions revealed only weak homology to Bacillus phages.

5. Conclusions & Future Directions

Conclusions:
- Despite extensive testing, no phages infecting *R. erythropolis* or *S. virginiae* were isolated.
- Use of lower temperature alone for *M. smegmatis* did not result in novel phage.
- Use of direct plating *and* lower temperature resulted in potentially novel phage infecting *M. smegmatis.*
- Preliminary results suggest that *Corynebacterium glutamicum* may serve as a useful host to discover additional actinobacteriophage diversity; two isolates differ from each other and from two published phages.

Future Directions:
- Further analysis of Cluster K phages (W&M continues to be a "hot spot" for Ks)
- Annotation of novel *M. smegmatis* and *C. glutamicum* phages
- Transcriptome analysis of novel phages

6. Acknowledgements

- The Howard Hughes Medical Institute Undergraduate Science Education Program & the SEA-PHAGES Program
- The University of Pittsburgh (Graham Hatfull, Welkin Pope, Debbie Jacobs-Sera, Dan Russell, and entire Pittsburgh team)
- The College of William & Mary and our Instructors, TAs & a huge thanks to Caroline Golino & Andy Halleran for sequencing expertise!