

Annotating Phages Using Bioinformatics to Further Understanding of Phage Biology

N.L. Arocho, W. Earley, B.M. Mazen

Department of Biology • Coastal Carolina University • Conway South Carolina



Background Information

A phage is a type of virus known to infect bacterial cells for reproduction. Phages can play a key role in the fight against antibiotic resistance. The understanding of these phages will lead to a better understanding of phage biology and underlying mechanisms. CCU students have discovered phage Phayeta, which attacks bacteria *Mycobacterium smegmatis*, a bacterium closely related to the causative agent for Tuberculosis. This research looks mainly at Phayeta's minor tail proteins which gives us a good insight on how phages mutate and evolve. This, in turn, also helps us understand phage biology.

Methods

Use of bioinformatics:

- Annotation
- Blast Results
- Comparative analysis

Findings

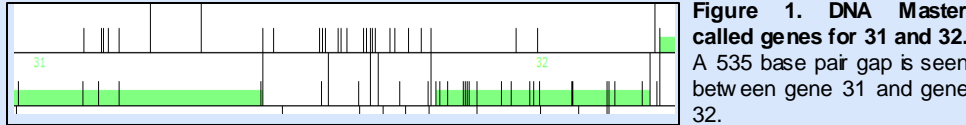


Figure 1. DNA Master called genes for 31 and 32. A 535 base pair gap is seen between gene 31 and gene 32.

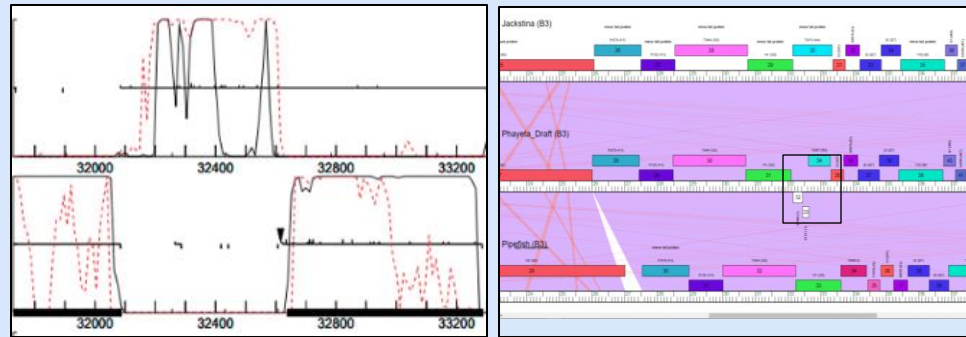


Figure 2. Coding potential for gene 31 and 32. Coding potential was found between gene 31 and 32 on a separate track. The coding potential was not called as a gene in this program.

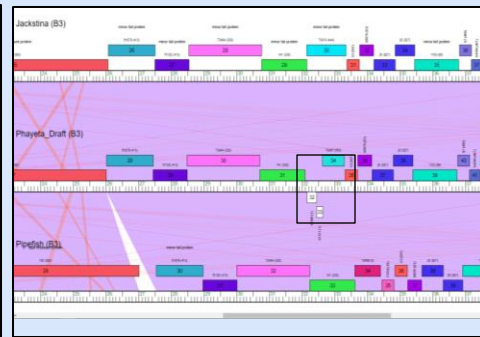


Figure 3. Phage gene comparison. Phayeta compared to other phages annotated within the same domain. Coding potential was recognized and called as genes 32 and 33. The comparison to other phages shows a potential mutation within the Phayeta's genome.

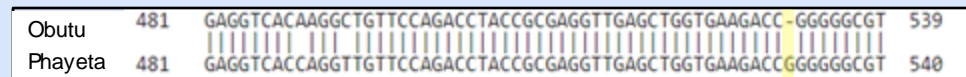


Figure 4. BLAST results of Phayeta with a phage previously annotated. BLAST result of gene 32 on Phayeta when compared with a gene of similar function on phage Obutu. A single nucleotide insertion occurred within gene 32 of the Phayeta genome, causing the minor tail protein gene to be cut and to become unfunctional.

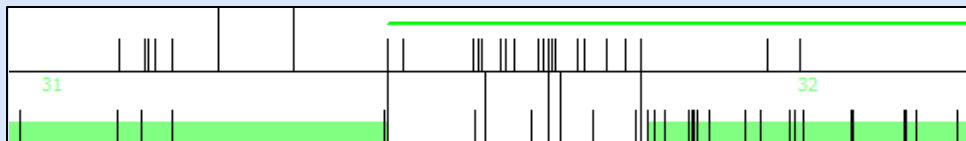


Figure 5. Frames view of the new start of gene 32. Discovery of the nucleotide insertion prompted the removal of the previously predicted gene 32 (bottom right green line) and the replacement with the gene start that contains all coding potential for the minor tail protein (upper right green line).

Conclusions

By annotating Phayeta's genome, evolution in action was observed within a conserved gene through a single nucleotide insertion. The discovery of this insertion led to the removal of the original called gene 32 and the addition of a new gene that contains the full coding potential.

Future Directions

- Continue discovery and annotation of phages to expand the knowledge of phage biology.
- Identify the function of genes with unknown functions using applications like Alpha Fold.
- Further the understanding of phage medical applications such as phage therapy.

Acknowledgements

This research was conducted for BIOL303-Q1 at Coastal Carolina University as part of the SEAPHAGES program supported by HHMI. Mentored by Dr. Daniel Williams