

The Annotation of Genes 17 through 60 of the bacteriophage KADY with the Programs DNA Master, Phamerator, and Starterator

Payal M. Patel, Victoria C. Torres, Kathrine E. Yacoo, Jacob D. Kagey, and Joshua J. Thomson
University of Detroit Mercy, Department of Biology and Chemistry

Abstract

Bacteriophage are viruses that infect and kill bacteria by inserting their genetic material into the genomes of bacteria. They use the bacterium's reproductive mechanisms to replicate. When there is a large accumulation of new phage the overwhelming pressure will cause the cell to burst, releasing the new bacteriophage to spread. They can be used in what is called "phage therapy," the use of bacteriophage to attack and eliminate harmful bacteria that infect humans and livestock. The overall goal of annotating mycobacteriophage KADY, was to analyze its abilities and similarities with other bacteriophages to suggest further host bacterium. Previously, KADY was isolated from an enriched soil sample from Warren, MI through a series of experimental methods. Mycobacterium smegmatis mc²155 was the isolation host. The isolated DNA from KADY was sequencing using the Illumina Sequencing method at the Pittsburgh Bacteriophage Institute. In this study, we annotated the 50898 bp DNA sequence from KADY with the programs DNA Master, Phamerator, and Starterator. We describe here the complete DNA sequence of the KADY genome. In total we changed two genes from the start codons that DNA Master had predicted. Many of the annotated genes do not have predicted function. However, the integrase protein was identified as Gene 34. Genes 23 and 24 were classified as the frame shift, which occur in every bacteriophage genome.

Introduction

- KADY was sequenced at the Pittsburgh Bacteriophage Institute using the Illumina Sequencing method
- Genome length is 50898 bp
- Section B: Genes 17 through 32 are on the forward strand while 33 is on the reverse strand
- Section C: Gene 34 and genes 36 through 60 are on the reverse strand while gene 35 is on the forward strand
- The criteria for annotation included: Start and Stop codons, Coding Potential, Shine-Dalgarno (Kibler6/Karlin Medium), Glimmer and GeneMark start codons, Longest Opening Reading Frame, Gap or overlap size between two genes and BLAST Alignment
- DNA Master auto annotates the genome then the final placement of genes is declared based on the criteria



Figure 1: Plaque Morphology

Plaques are clear zones formed in a lawn of cells due to lysis by phage. The plaques are 3 mm in size and lytic.

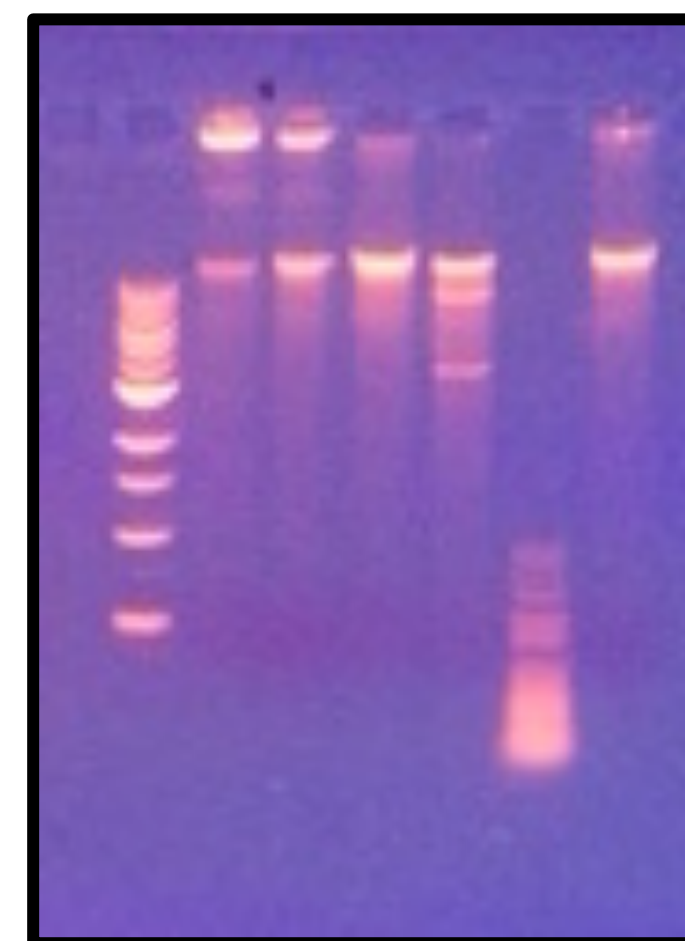


Figure 2: Gel Electrophoresis

A restriction enzyme digest with was run with the following enzymes (in the following order) BamHI, ClaI, EcoRI, HaeIII, HindIII.

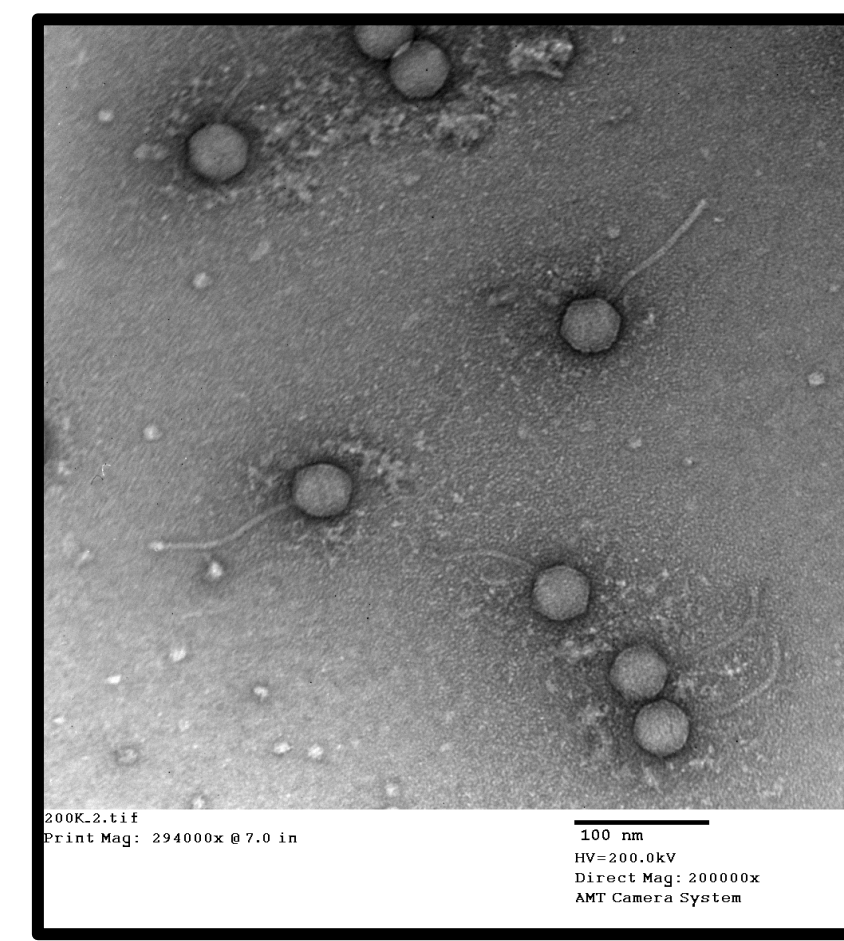


Figure 3: Transmission Electron Microscopy Image of KADY

An image is formed from the interaction of electrons transmitted through the specimen. KADY was classified in the Siphoviridae family.

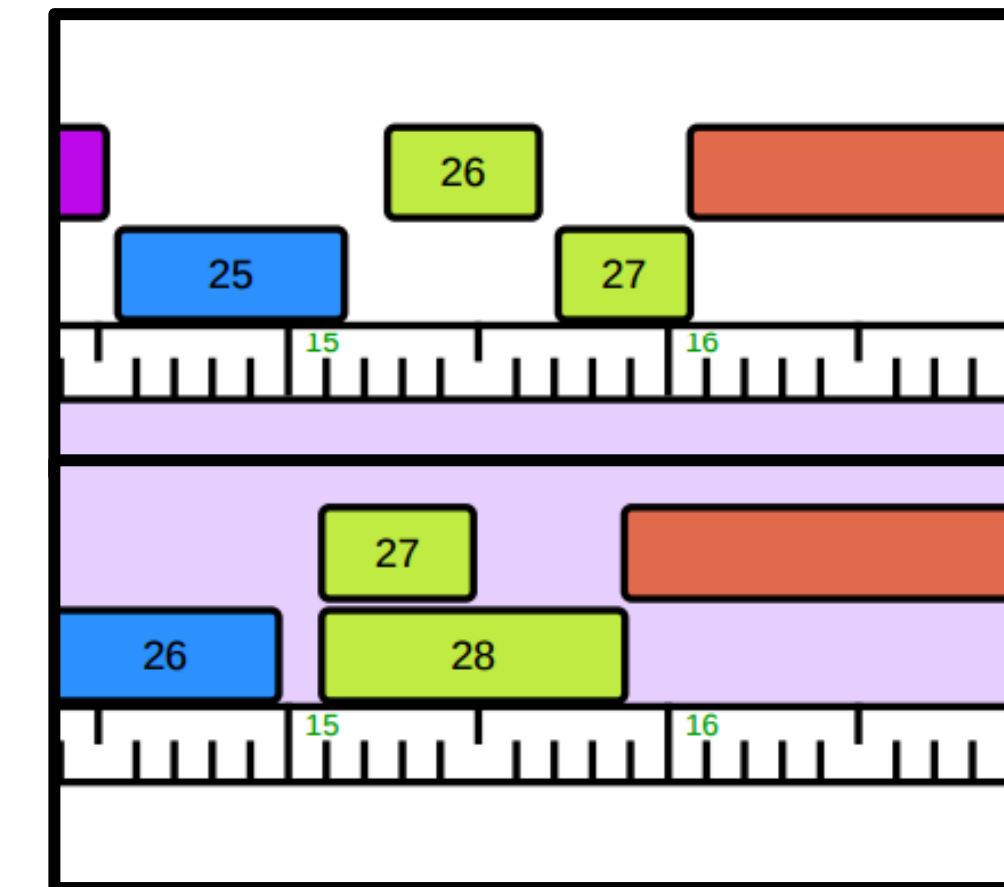


Figure 4A: Phamerator Image of KADY (above) and Wooldri (below)

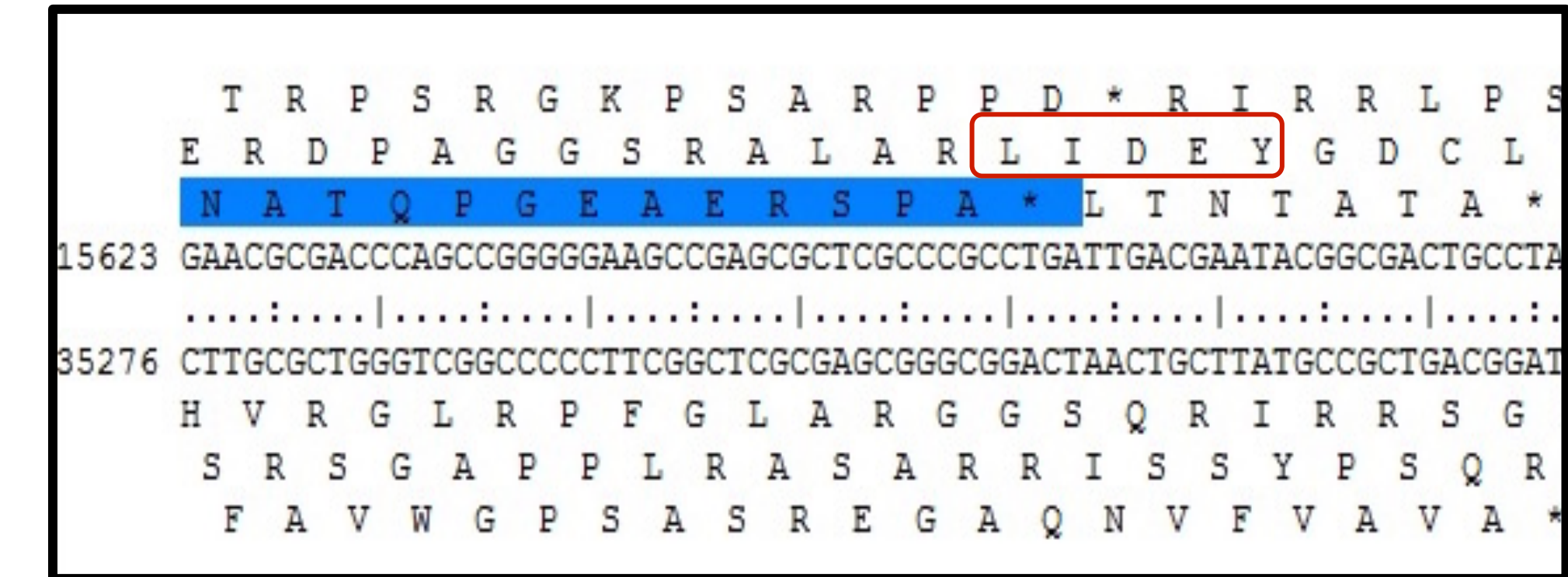


Figure 4B: Six Frame Translation

Frameshift

In Figure 4A, KADY (above) and Wooldri (below) are shown in phamerator where the frameshift occurred. Due to the three tRNA in between genes 7 and 8, genes 23 and 24 of KADY in DNA Master are equivalent to genes 26 and 27 on phamerator. Gp23 of KADY aligned with the tail assembly chaperone protein of Wooldri 28. A frameshift mutation is a genetic mutation caused by an insertion or deletion. In Figure 4B, it is shown that the insertion is an extra Cytosine at the six frame translation LIDEY.

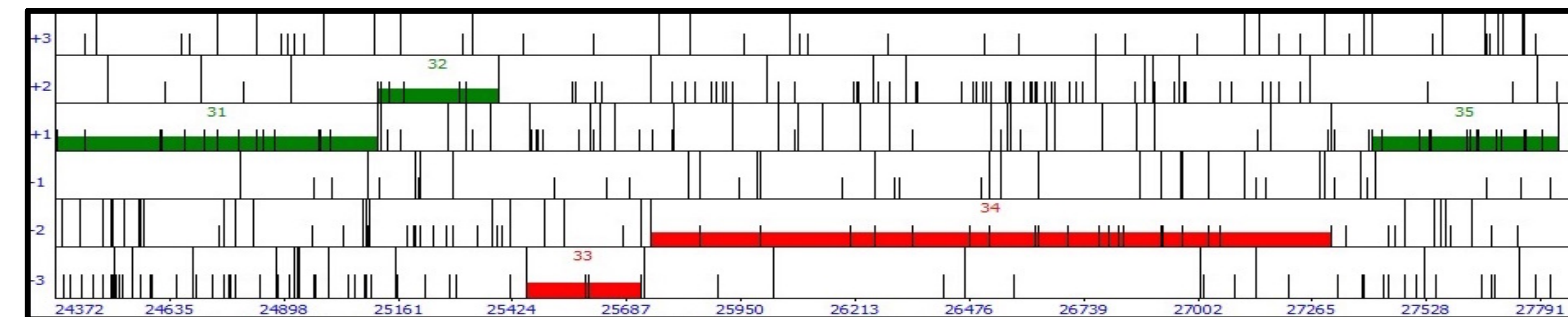


Figure 5: ORF Analysis for KADY

Integrase

The integrase enzyme is a site specific recombinase that moderates the DNA integration, with reference to which bacterial strains are adsorbed. Often, the integrase is found in a gene that is on the reverse strand and is followed by a gene in the forward strand.

Results

- The complete DNA sequence of KADY's genome was found for genes 17 through 60
- The frameshift occurs in 15267-15662
- The integrase protein was defined as gene 34
- In total two genes were changed: genes 38 and 45. Gene 38 was found to have a start at 28574 on a reverse strand. Gene 45 was found to start at 31884 on a reverse strand
- The closest relatives to KADY are Wooldri and Bxz2
- Most of the genes have unknown functions

This work was funded by Howard Hughes Medical Institute (HHMI) Science Education Alliance and the NHI ReBuild Program