CONSIDER FOR TALK

10th Annual SEA Symposium Abstract

Universidad Autónoma de Nuevo León

San Nicolás de los Garza

Corresponding Faculty Member: Patricio Zapata (dr.patriciozm@gmail.com)



Nancy E Rodríguez-Garza

Genomic Annotation of Mycobacteriophage Tubs

Elva T Aréchiga-Carvajal, José I González-Rojas, Patricio A Zapata-Morín, Nancy E Rodríguez-Garza, Francisco R Rodríguez-Recio , Mónica D Rodríguez-Nava

Bacteriophages have been applied by scientists as tools to understand fundamental molecular biology, as vectors of horizontal gene transfer, as sources of diagnostic and genetic tools and as novel therapeutic agents. Therefore isolation, identification, and annotation of their genome is of great interest. For this reason, we were assigned for the genomic annotation of the Mycobacteriophage Tubs, which was isolated from Foxworth, MS USA. This phage belongs to the Siphoviridae family whose particular characteristics are that have long, noncontractile and thin tails as well as dsDNA genome, and is classified in subcluster A9. The sequence of Tubs was entered into the DNA Master program to carry out an autoannotation that needed to be refined because the algorithms Glimmer and GeneMark can only make an accurate assumption of 90% of gene positions. For this, we analyzed the start of each gene and determined if these were the optimal ones. To do this, the results of GenMark, Starterator, BLAST and the RBS scores were considered. Analyzing the results obtained with these tools, the start of the genes 10, 23, 27, 32, 33, 55, 56, 58 & 74 were modified in order to have a conserved start between the pham, match 1 to 1 with the genes of other phages and better RBS scores. In addition, it was determined that some genes were not autoannotated. Genes that weren't called by either Glimmer or GeneMark (genes 1, 4, 31.1 & 56.1) were determined by analyzing the gene synteny with related phages of the Cluster (Pheader, PackMan, Catalina, HortumSL17, Myxus and Scherzo). After this we observed the presence of large gaps (>130 bp) which indicated the possible presence of a gene. The coding potential, RBS scores and BLASTp alignments (match with related phages) in these unannotated genes were strong evidence that supports the presence of a gene. Once all the genes were established, we proceeded to identify its possible function. For this, we used BLAST, where homologous genes were identified in other phages with an assigned function, and HHPred to assign possible function through the finding of conserved domains within the protein structure that can be linked to certain functions. These tools helped us to identify structural genes, DNA replication genes, genes involved in life cycle regulation, etc. Besides a synteny analysis against other cluster members was made to compare the structure, finding that the synteny between the genomes is conserved and the differences in some regions are minimal, especially in the region were the forward genes are. The differences we observed are in the regions of the reverse genes, especially in Packman, were there is a considerable deletion (~1900 bp) that includes 3 missed genes (2 without reported function and a DNA methylase). Also, Scherzo has 3 gene deletions, and 1 gene insertion in comparison with Tubs. And Catalina, Hortum, Myxus and Scherzo have an apparent gene insertion within Tubs and Phaeder gene 61.