CONSIDER FOR TALK

2022 SEA Symposium Abstract

Merrimack College

North Andover MA

Corresponding Faculty Member: Janine LeBlanc-Straceski (janine.leblancstraceski@merrimack.edu)



Sophia Beland



Skylar DiMartino



Erin Lincoln

A Bold Hypothesis for Shygu2

Sophia Beland, Skylar DiMartino, Erin Lincoln, Janine M LeBlanc-Straceski

Shygu2 is a mycobacteriophage captured from Mendel Pond on the Merrimack College Campus in North Andover MA on a 60°F, September 2021 day. Isolated on the *Mycobacterium smegmatis* mc²155 host, its 51410 bp genome contains 87 genes in a typical first half forward (32 genes), second half reverse (55 genes) orientation for cluster A phage. Shygu2’s genetic composition reveals that it is a temperate phage with a siphoviridae morphotype, belonging to the A4 subcluster. Its closest relatives are Morpher26, Katalie136, Bumblebee11, CiCi and AbbysRanger. Some typical A4 subcluster features include 1) the immunity repressor located far down the right arm (at gene 69) instead of adjacent to the serine integrase (gene 33). The large space upstream of the start of the gene (3’ for this reverse gene) has been reported to be a promoter and an operator sequence in other A4 cluster phage. 2) As in Backyardigan, the Shygu2 tapemeasure (gene 25) has a large overlap, -95 bp, with the previous gene, which contains all typical and atypical coding potential. 3) Shygu2 also has 3 minor tail proteins at the left end of the genome between the terminase (gene 2) and lysin A (gene 8) characteristic of Cluster A4 phage like Morpher26 and Katalie136. 4) There is a 5’ most HNH endonuclease gene that needed to be manually added. It is interesting that the gene prediction algorithms consistently miss this 5’ most gene even though it has both typical and atypical coding potential identified by GeneMark. 5) Finally, the two consecutive primase genes, 55 and 56, have the typical lengthy overlap seen in this subcluster. The Starterator suggested start site of gene 55 (a reverse gene) has a 74 bp overlap with gene 56. To date no lab bench evidence of ribosomal slippage during translation has been found. Gene 55 shares domain homology with bacterial DnaG primase. Gene 56 has a zinc finger DNA binding domain homologous to bacterial primases. Heterodimer primases containing two polypeptides, each having distinct functional domains, are found in archaea and eukaryotes (AEP, archaea-eukaryotic primase). It is interesting to hypothesize that these two genes may produce two distinct polypeptides that interact to form a multi-subunit primase in these mycobacteriophage, although why there is a large overlap remains unclear.