**Annotation Notes: Phage Correa**

We annotated the phage Correa genome for coding potential using GeneMarkS Output.

Most of our gene start calls agree with Staterator recommended start sites. The exceptions are gp27, gp29, gp47 & gp59. Both gp 27 & 47 are homologs to Circum genes and have been annotated at the same positions as the Circum homologs. The start site chosen for gp29 gives the longest ORF and best SD score.

Genes gp90, 92 & 93 were not called by Glimmer or GeneMark but fill genomic gaps and have coding potential.

Genes of Special Interest:

Gp2 is a lone reverse gene but has good coding potential and has a homolog in the related AM Cluster Circum genome annotated as sole reverse gene.

Gp12 appears to be a prohead protease/major capsid fusion gene. The left part of the gene gives good hits to prohead protease genes and the right part of the gene gives good hits to major capsid genes. **We note that the closely related AM cluster Circum genome homolog has been annotated as only encoding the prohead protease protein.**

Interestingly, both gp15 & gp17 give good hits as encoding a major tail subunit protein. **We not that the closely related AM cluster Circum genome homologs have been annotated as capsid genes. We feel these two related homolog are in fact MCP genes.**

Gp96 gives really good blast hits to homologs of tail proteins of phages that can infect different types of bacteria hosts. There is precedence in the S cluster mycobacteriophages of tail genes located at the right hand side of phage genomes.

Lastly, using the programs Aragorn and tRNA scan we were not able to identify any tRNA or tmRNA genes within the phage Correa genome.