|  |  |
| --- | --- |
|  | **Omnicron**  Florida Gulf Coast University  May 13, 2014 |

We are fairly confident with Omnicron’s annotation. Since it’s a K5 phage, with only 5 other subcluster members, 4 of them still in draft form, functional calls were not as straight forward as when annotating Power, an A2 phage with many subcluster members. In summary: (1) synteny was followed with the structural genes, (2) a t-RNA was discovered, (3) there is a frameshift in the tail assembly chaperone, (4) 3 putative novel genes of unknown function were discovered, (5) an operon of interest was found with a putative toxin/antitoxin system, and (6) several large gaps were identified

**t-RNA**: 789-865 bp. A tryptophan t-RNA (anticodon CCA) was discovered. Ends were trimmed to Aragorn call. COVE=80.15.

**Frameshift**: 10271-10759 bp and 10870-11151 bp for tail assembly chaperone. The slippery sequence involved was GGGAAAA. This orf bore a high similarity to Larva. The frameshift was annotated by reading the first A in the slippery sequence counted twice, which allowed the gene to shift up one reading frame (from +2 to +3) at 10723 bp changing from GKSR to GKIA, etc.

**Putative novel genes:**

(1) Feature 34727-34936 bp orf, gp47. This orf has strong coding potential in the reverse direction and is conserved in its location in K5 phage, except for Larva (published). An autoannotation of the Larva fasta file also revealed this feature. Initially, it appears that this feature might be a miscall, and our intitial inclination was to delete it (as was done with Larva) since it is between two orfs that are oriented in the forward direction, and creates a large overlap (234 bp). The pham for the orf in Omnicron is unique. The other K5 draft genomes have two additional phams (two members in each pham). Thus, the sequence itself does not appear to be conserved, just the location. We erred on the side of caution and optimism hoping that this reverse feature, in this unique location might serve a function in K5 phage, so we opted to leave it in.

(2) Feature 38414-39298 bp, gp54. This is a relatively large orf (885 bp) of unknown function, with strong coding potential in the GeneMark graphical output with *M. tuberculosis* preferences. This orf is absent in the other K5 phage and is only found in one unrelated phage (Mutaform13) with a low (39%) identity with 96% query cover. New gene?!

(3) Feature 49348-49560 bp orf, gp71. This feature yielded no good matches to other phage in PhagesDB or NCBI (poor E-values). The coding potential in the GeneMark graphical output was not as strong as for other orfs. We left it in the annotation, but we’re not sure if this could be a novel gene or a miscall.

**Operon of interest**: 58318-58644 bp, gp 90, and 58646-59038, gp91. Though we functionally listed these features as “NKF”, based on strong HHpred matches, we found that gp90 resembles the toxin and gp91 the anti-toxin of a putative bacterial toxin/antitoxin system. We think it’s worth exploring.

**Gaps**: Large gaps were identified: between gp96 and gp1 (181 bp), t-RNA and gp 5 (292 bp), gp11 and gp12 (241 bp), gp 38 and gp39 (273 bp), gp 39 and gp 40 (216 bp), gp42 and gp43 (119 bp), gp43 and gp44 (136 bp), gp63 and gp64 (109 bp), gp81 and gp82 (159 bp), gp88 and gp89 (113 bp), gp91 and gp92 (145 bp), gp92 and gp93 (280 bp), gp93 and gp94 (252 bp), and gp94 and gp95 (210 bp). We did not find good coding potential or BLAST matches in these regions.