ILeeKay Cover Letter

This is the first annotation for Northwestern College so we are eager to get feedback from the SMART team. We annotated an adopted phage. Next Fall semester we will annotate two phages discovered by students at our institution. What follows is a brief summary of our annotation.

**Genes added**: DNAM5\_6 (5b) and DNAM5\_80 (80b)

**Genes deleted**: Originally listed as genes 44 and 61

**Start site changes**: DNAM5\_31, DNAM5\_39, DNAM5\_58, DNAM5\_63, DNAM5\_67, DNAM5\_71, DNAM5\_74, DNAM\_81, DNAM5\_86, DNAM5\_87, DNAM5\_88, DNAM5\_89 (we also resolved Glimmer/GeneMark disagreements)

**Gaps**: We have three gaps that we tried to fill but could not. Note that the frames on our GeneMark coding potential graphs for reverse genes did not align with the frames in DNA Master. In other words, Frame 4 in DNA Master corresponded to Frame 5 in GeneMark, Frame 5 in DNA Master to Frame 6 in GeneMark, and Frame 6 in DNA Master to Frame 4 in GeneMark.

* Between DNAM5\_58 and DNAM5\_59 there is a 339 bp gap. We searched the gap for potential genes. The only possible ORF was one in frame 5 (DNA Master) that had a single, weak BLAST hit to Nepal\_57. There is coding potential in Frame 5 of GeneMark (Frame 4 in DNA Master) but any ORF in Frame 4 would overlap DNAM5\_58 dramatically.
* Between DNAM5\_83 and DNAM5\_84 there is a 268 bp gap. Extending the start site for DNAM5\_83 was not possible and there are not reasonable ORFs in the gap in other frames.
* Between DNAM5\_84 and DNAM5\_85 there is a 207 bp gap. Extending the start site for DNAM\_84 moved it from the most conserved and did not include any additional coding potential. We looked for potential genes/ORFs in the gap but did not find any.

**tRNA genes**: We looked for but did not find any tRNA genes in ILeeKay’s genome. We used Aragorn and tRNAscan-SE.

**Programmed translational frameshift**: We annotated a -1 programmed translational frameshift in genes DNAM5\_24 and DNAM5\_25 (tail assembly chaperone).

**Additional notes**: We called gene DNAM5\_66 to agree with the Glimmer call. The original BLAST hit genes from a few related phages. When we re-BLASTed this gene after our annotation was complete we were not able to get any matches in DNA Master. If we take the product and BLAST it in NCBI we find several hits but the e values are not lower than 10-18. We acknowledge that this is a small gene but we decided to leave it in because several non-draft phages that we found were similar to ILeeKay decided to include it.

Thank you!

Sara Tolsma and the NWC SEA-PHAGES Annotation Team

[stolsma@nwciowa.edu](mailto:stolsma@nwciowa.edu)