Lehigh would like to submit Streptomyces phage Warpy for QC.

Many tRNAs were added because they were called by Aragorn and tRNA-scan. We added feature 117 (77861-77950), feature 140 (87245-87329), feature 173 ( 94943-95017), feature 174 (95103-95176), feature 178 (95841-95916), feature 211 (106130-106213).

One tRNA was deleted (100367-100506) because it was not called by Aragorn and tRNA-can called it as an undetermined tRNA.

Many gene products were deleted because they had no homology with Jay2jay, low or no coding potential, and poor SD scores. The following genes were deleted gp1(95-214R), gp2 (563-829), gp11 (5473-6039), gp17 (7041-7538), gp21(7932-8396), gp244 (132991-133108), gp151 (97647-97784), gp177 (106908-107189), gp214 (121503-121646), gp215 (121633-121752R), gp225 (127011-127577), gp231 (128579-129076), and gp235(129470-129934).

A few genes were added because they had strong coding potential and genes called in Jay2jay. The genes added were gp1(8-220), gp253 (121546-121758).

We would like the following regions investigated further.

Gp72 (53028-53324) should be called an HNH endonuclease based on HHpred data. We have called this function even though it was not called in Jay2jay.

Gp75 (54400-54579) was not called the same as Jay2jay because of the following reasons: the sd score is higher, the coding potential matched better, and it is a longer ORF with the start site we chose.

Gp125 (81580-81768) did not call Q1:S1 with Jay2jay, the coding potential matches this start better, the sd score is higher, and the gap is smaller.

Gp150 (97215-97412) we have chosen to delete this gene based on homology with Jay2jay. There is no coding potential however a large gap exists without it.

Gp216 (107155-107349) has a changed start site so it no longer is 1:1 with Jay2jay. This was done because it aligned with coding potential better, it has a better SD score.

Gp226 (113064-113399) was not called to match Jay2jay Q1:S1. There is high homology but the ORF we called has better sd score, and has coding potential. This does have a large overlap however.

gp249 (120037-120504) was not Q1:S1 with Jay2jay. The SD score is much higher, the coding potential is a better fit for the gene we called.

The region between 121009-122456 should be investigated further for any genes missing. This area is highly homologous with Jay2jay however some genes that Jay2jay calls are not in Warpy. We called all genes in this region based on coding potential.