Lehigh University would like to submit the annotated genome of Streptomyces phage Sushi23 for QC. Since this is only the second BE1 genome annotate we have many regions we would like double checked by the QC team. We used GeneMarkS to read coding potential for this genome. Also of note, many starterator calls do not match the chosen start choice because they are based largely on draft files. Many gene products have no BLAST data because these gp are new in BE1 Streptomyces phages, they are found in many of the unannotated genomes.

We deleted many genes because there was no coding potential, no hhpred data, no homology, very small gene products and/or they were called in the same place as genes with these features. The following features were deleted gp10 (5903-5965), gp 12 (5958-6035), gp 18 (7556-7747), gp19 (7698-7997), gp100 (73472-73591R), gp135 (92675-92818), gp136 (92815-93048 R), gp160 (103864-104112R), gp198 (118462-118578R), gp204 (120074-120193), gp222(128756-128890), gp221(128746-128808R), gp229 (130399-130590), and gp230 (130541-130840).

We also added a few genes based on homology with Jay2jay or coding potential in the region of interest. The genes we added were gp101 (73427-73588) and gp249 (123371-123562). The region (123371-123562) is a very large gap as the reading frame goes from positive to negative. This region was investigated for other potential gene calls and none were found. There was very little coding potential, all the gene products would have been very small and no homology was found with Jay2jay.

Additionally, 5 tRNA molecules were added because they were called by both web based Aragorn and tRNA-scan. These also had infernal scores that were above 20. Feature 130 (86244-86315), feature 134 (86883-86958), feature 167 (95294-95365), feature 172 (96164-96238), feature 205 (106726-106809).

We have updated a few of the functional calls as well based on HHPred data. Gp68(52765-53043) in Jay2jay does not have a functional call but strong homology to HNH endonuclease, this function was called. Gp87 (65352-66125) was called as an exonuclease due to strong HHpred data. Gp105(74695-76509) is called as a putative ftsk DNA translocase, since there is high probability and coverage in HHpred we removed the putative. 2 gene products also have homology but we did not call the functions. Gp89 (66965-67423) and gp90 (67416-67580) has HHpred homology to structural genes and gp91(67589-68218) has homology to Thymidylatesynthetase.

Jay2jay has 2 gene products prior to the 1st gp in sushi23 no coding potential in this region, also there is significant sequence divergence between this region in jay2jay and sushi23. We have chosen not to call genes in this region. Both gene products that are present are very small with low homology to even Jay2jay.

The region between gp9-11 (4764-6047) is unusual region. There is one forward gene called in a string of reverse gene. This gene has strong coding potential. All other genes that could be added have no homology or coding potential.