We adopted singleton phage Bipper from Dr. Welkin Pope’s class. Much appreciated! ☺ We are confident with most areas of Bipper’s annotation. Here is what we found:

1. We annotated 135 features including 134 ORFs and a Met(cat) tRNA [gp41] at position 32891 bp.
2. We called functions in roughly 1/3 of the ORFs whereas the remaining 2/3 were NKF. *{Typical for a singleton?}*
3. 57% of the phams were orphams (77 out of 134).
4. There was a striking abundance of long stretches of overlapping ORFs; as many as 8 consecutive ORFs overlapped by exactly 4 bp. A stretch of 36 ORFs [gp75 to gp110] encompassing 16634 bp only had a single small gap (2 bp). *{Is this unusual?}*
5. 6% of ORFs (8 proteins) had putative transmembrane domains [gp2\_NKF, gp10\_NKF, gp36\_Holin, gp37\_NKF, gp50\_Y-int, gp53\_NKF, gp56\_NKF and gp107\_NKF].
6. ATG (55%) and GTG (44%) were primarily used as start sites. Only one ORF used TTG [gp81\_NKF] as a start codon *{we found it unusual that there was only one TTG}.*
7. Synteny was followed with the structural genes. However, we did find a couple of structural genes near the 3’ end of the genome, outside the typical cassette; gp125\_Minor tail protein and gp130\_Terminase domain. *{Take a look at these}*
8. We annotated a typical tail assembly chaperone [gp20 and gp21] translational frameshift [-1 shift] with slippery sequence GCGGGGG, the 3rd [G] at position 13619 bp was repeated.
9. There was a second putative translational frameshift [gp70\_NKF and gp71\_NKF]. The predicted slippery sequence is: CGCCCC with a [+1 shift]. The 4th nucleotide [C] at 48823 bp was skipped. Since it has not yet been functionally validated, we did not include this frameshift in the annotation. We described it in the notes window of the Complete Notes file. *{take a closer look}*
10. There are several gaps around 100 bp or longer. We did not find good coding potential (with GeneMark M. smegmatis MC2 155 nor TB preferences) or BLAST matches in these regions *{take another look}.* The first major gap is 604 bp between gp135\_HNH endonuclease at the 3’ end of the genome and gp1\_ParB at the 5’ end. Near the **5’ end**, there’s a 182 bp gap between gp12\_Capsid maturation protease and gp13\_Major capsid protein. For the most part, gaps were **clustered** **between gp38 and gp52**: 108 bp between gp38\_NKF and gp39\_HTH DNA binding protein; 201 bp between gp40\_NKF and gp42\_Ribonuclease toxin (the tRNA is found in this gap); 166 bp between gp42\_Ribonuclease toxin and gp43\_NKF; 163 bp between gp43\_NKF and gp44\_NKF; 231 bp between gp44\_NKF and gp45\_NKF; 130 bp between gp45\_NKF and gp46\_NKF; 130 bp between gp47\_NKF and gp48\_Antitoxin; 168 bp between gp48\_Antitoxin and gp49\_NKF; 124 bp between gp49\_NKF and gp50\_Y-int; 272 bp between gp50\_Y-int and gp51\_Immunity repressor; and 224 bp between gp51\_Immunity repressor and gp52\_Recombination directionality factor. Several other gaps were found intermittently nearing the **3’ end** of the genome: 99 bp between gp74\_NKF and gp75\_NKF; 190 bp between gp112\_NKF and gp113\_NKF; 112 bp between gp119\_NKF and gp120\_NKF; and 430 bp between gp123\_NKF and gp124\_NKF.
11. We deleted 3 ORFs called by the autoannotation for lack of coding potential; [FWD:50180-50323 bp, which allowed us to lengthen the downstream ORF (gp75\_NKF) to its full coding potential], [REV:67809-67928], and [REV:72556-72669]. We looked at gp38\_NKF carefully since it was the only reverse ORF in that region, but could not justify deleting it since it had very strong coding potential. *{take another look}*
12. We found several interesting repeat elements (MEME). One motif was particularly long (50 bp) and had 5 occurrences in locations correlating to large gaps in the genome (in blue). [GTCTAGTGAGCTATAACTGACTTACCGGGGCAGGTACCAGACCAGCCCCC]: at 32997 bp [gap between gp40\_NKF and gp42\_Ribonuclease toxin], 33401 bp [gap between gp42\_Ribonuclease toxin and gp43\_NKF], 35254 bp [gap between gp44\_NKF and gp45\_NKF], 36,952 bp [gap between gp48\_Antitoxin and gp49\_NKF], and 71347 bp [gap between gp119\_NKF and gp120\_NKF].