**TheloniousMonk\_AnnotationCoverSheet**

Gene 25 Additional Notes: This gene contains a -1 translational frameshift at bp 17191; the protein sequence used to annotate this gene is NOT the sequence originally called by Glimmer, but the sequence that includes the amino acids called in the first reading frame between bp 17191 and 17634.

Gene 31 Additional Notes: Starterator and Glimmer/GeneMark call the start at 25058 (Z-value = 1.62) but moving the start back to 25043 (Z-value=1.89) has the advantage of a 4 bp overlap instead of a 12 bp gap as well as a better Z-value. That is why the start is at 25043.

Gene 32 Additional Notes: Set the start to match GeneMark since Starterator call it and it gives a 4 bp overlap.

Gene 34 Additional Notes: The starterator call overlaps the previous gene almost 200 bp. Used the Glimmer/GeneMark start.

Gene 35 Additional Notes: Used the Starterator suggested start site at 29176 (z-value=1.936) instead of Glimmer/GeneMark call at 28170 (z-value=1.266). Two codons of coding capacity are gained but there is about 30 bp of coding capacity that is not captured by either start site.

Gene 37 Additional Notes: I am torn whether to go with the structural protein function, which is very vague for a function, or the carboxypeptidase function which is specific but has less support strength. If I consider Structural protein to be just a specific hypothetical protein, then I would not normally annotate a function that way. Therefore, I will go with the carboxypeptidase.

Gene 39 Additional Notes: Starterator had stronger start location than the called by Glimmer/GeneMark. I went with the Starterator start site.

Gene 45 Additional notes: The Glimmer called start at 34272 (Z-value = 2.018) does not capture all of the coding capacity. The start at the suggested Starterator site at 34368 (Z-value = 2.119) captures all of the coding capacity but overlaps the upstream gene by 23 bp. The intermediate start site at 34326 (Z-value = 2.234) captures all of the coding capacity and has a 20 bp gap to the next upstream gene. I choose the start site called by Starterator because it matches Aeneas\_46 and Nepal\_46 at 99% and 1:1 and because the intermediate site does not match other phage start sites 1:1.

Gene 56 Additional Notes: Added this gene to match Museum gp 57, cover coding capacity and to fill in gap. Perhaps uses a frameshift with adjacent primase gene.

Gene 68 Additional Notes: I added this gene to cover coding capacity and fill a gap. The gene overlaps the previous gene by either 1 or 4 bps, I elected to use the 4 bp start. The gene matches Nerujay\_69.

Gene 74 Additional Notes: Original Glimmer start at 46693 (Z-value= 2.278) was not chosen because a 4 bp overlap start site at 46699 (Z-value=1.893) was available.

Gene 83 Additional Notes: I moved the start site from the Glimmer, GeneMark and Starterator suggest start site at 49436 to 49460 which covered the gap between genes and gave a 4 bp overlap with the previous gene.

Gene 86 Additional Notes: Start moved from Glimmer start at 50890 to GeneMark and Stoperator suggested site at 50899. This move also gave a 4 bp overlap.

Gene 87 Additional Notes: I used start 51102 instead of 511099 because it gave a 4 bp overlap instead of a 1 bp overlap with exactly the same Z-value.

Gene 88 Additional Notes: Start moved from Glimmer/GeneMark start at 51242 to 51287 so that there is a 4 bp overlap vs a large gap. Starterator is NI, non-informative.