We obtained great coding potential for almost all HunterDalle genes called by GeneMarkS Output. Almost all of our chosen gene start sites were ones that captured all coding potential via GeneMarkS Output.

Most gene start sites agreed with Startorator recommended start sites. If we chose another start site other than the Starorator recommended site it was either because we would not capture all coding potential via GeneMarkS Output and/or the chosen start site had a better SD score and agreed with Glimmer and/or GeneMark.

We did not detect any tRNA or tmRNA genes using both Aragorn and tRNA scan

We annotated genes gp12 & 13 as the tail assembly chaperones with gp13 annotated as a minus one frameshift. Gp12 & 13 genes are next to the tape measure protein (gp14)

Gp15 gives a great hit to a tail protein. Could this be the major tail protein subunit gene?

We also annotated genes gp21 & gp23 as the lys A & B genes respectively. Gene gp22 gives a modest hit (65% probability for a holin gene) via HHPred analysis. Based on its location, we annotate gp22 as the hollin gene. We believe the lysis cassette includes these three genes.

We annotate gp44 as a sole reverse gene flanked by forward genes. Ideally, we would have liked to see a space between gp44 and gp45 (forward gene) for a promoter element. Gp45 captures really good coding potential via GeneMark S Ouput so we want to keep the beginning of the gene as we annotated it. Perhaps gp44 should be deleted from the final annotation but it did give coding potential via GeneMarkS Output.

Finally, we deleted two short potential reverse genes next to gp56 (a forward gene). We expanded the length of gp56 to include all coding potential via GeneMarkS Output. This caused a complete overlap with the two reverse genes that we deleted.