**Simpliphy Annotation Cover Sheet, Claire Rinehart, WKU. 6/10/18.**

CDS complement (34879 - 35394)

 /note=Start site 35394 has a -4 gap motif and if start site 35394 were used coding capacity is not used.

CDS complement (35391 - 35573)

 /note=Top NCBI BLAST hits match 1:1.

CDS complement (35570 - 35839)

 /note=The Starterator start shown above is from an old Starterator input file. If you go to the Pham Starterator it actually suggests 35839 for the start.

CDS complement (39551 - 39982)

 /note=Starterator is very split, three ways.

CDS 43303 - 43632

 /note=Starterator suggests 43303, but that does not agree with some close relatives (Tuco, Williez) in Pham maps, excludes coding potential is Genemark, and leaves a large gap with no way to fill it. Conflicted on start site. It does match NoSleep in Pham maps though.

CDS 43629 - 44597

 /note=Distant relative to an exonuclease.

CDS 46493 - 46861

 /note=When there are two starts in a row the second one is used according to Mass Spec data (see SEA PHAGES forum).

CDS 46864 - 47232

 /note=Starterator is fairly split due to many starts at 46864 being drafts. NCBI is fairly split also. 46864 has much better Z and final score. Not sure how I feel selecting against -4 twice.

CDS 47638 - 48051

 /note=Porky is from 2008, but there is also supporting functional evidence in NCBI. Overall functional evidence is somewhat weak.

CDS 51007 - 51267

 /note=Starterator is split evenly between 50992 and 51007, but 51007 looks to be the better choice in NCBI.

CDS 51242 - 51463

 /note=Large overlap looks to be real.

CDS 52526 - 53467

 /note=gap looks conserved.

CDS 53939 - 54037

 /note=Stareterator is fairly split, but based on other evidence 53915 may be a stronger choice. I have gone back and forth on this one, could probably go either way.

CDS 57879 - 58421

 /note=Starterator is split four ways. Not real strong functional evidence.

CDS 60199 - 60978

 /note=Starterator is split a few ways, 60199 has good scores and conserved among closely related phages.

CDS 66378 - 66590

 /note=added gene

CDS 66571 - 66825

 /note=added gene, not a strong start site but matches the close relatives

CDS 66774 - 67235

 /note=66774 has a lot of drafts in Starterator and closlely conserved relatives seem to prefer 66822.

CDS 68008 - 68169

 /note=added gene, based on NCBI could probably go either way, Starterator would be helpful.

CDS complement (68808 - 69197)

 /note=Was originally called as ParB-like dsDNA partitioning protein but only ParB represented in a few Phagesdb hits. Real ParB hits have good HHPred evidence.

CDS complement (72449 - 72796)

 /note=Start site could go two different ways.

CDS complement (73650 - 73892)

 /note=added gene

CDS 74574 - 74795

 /note=Changed start site based on starterator.

CDS 74840 - 75253

 /note=added gene