Here are the regions of concern:

CRAFF\_9: This had many BLAST hits with capsid morphogenesis protein, which is not on the approved list, so we entered NKF.

CRAFF\_20: This had many BLAST hits with queuine tRNA ribosyltransferase, which is not on the approved list. Since this enzyme is a glycosyltransferase, that’s what we entered.

CRAFF\_39: SSC: 36244-36612 (FWD). There is another possible start at 36238. We picked this start because most of the 1:1 matches are with the shorter ORF.

CRAFF\_61: This had BLAST hits to genes with a number of different names, including replicative helicase, DNA helicase, DNA primase/helicase, DNA primase/polymerase. We entered DNA primase/polymerase based on the top HHpred hits, but there were also good hits for DNA helicase.

CRAFF\_65: SSC: 55543-55695(REV). There are two adjacent starts. We called the second one because we were told that the evidence suggests the second start. However, the first start will give a 4 bp overlap.

CRAFF\_66: SSC: 55695-56132 (REV). We called an ORF with a 32 bp gap so as not to truncate conserved coding sequence. The starterator suggested start at 56051 is called 81% of the time, but mostly in draft sequences. The start we called is called 12.9% of the time, but not in any draft sequences.

CRAFF\_68: SSC: 56580-56789 (FWD). This is an ORF at that was not called by Glimmer or GeneMark.

CRAFF\_69: SSC: 56888-57742 (FWD). This gene start was called by Glimmer at 57011, but we included an additional 123 bp upstream so as not to truncate the ORF. The start we called is annotated only 10.9% of the time, but all in non-draft sequences.

CRAFF\_77: SSC: 61072-61401 (FWD). This gene start was called by Glimmer and Starterator at 61165. But we called the start at 61072, which is called only 2.8% of the time (all in non-draft sequences), so as not to truncate the ORF. This is also the start called in the closely related phage Orion. However, this start gives a 35 bp overlap, which could be problematic..

CRAFF\_78: SSC: 61683-61477 (REV): There are two adjacent starts. We called the second start.

CRAFF\_89: SSC: 64764-64543 (REV): There are two adjacent starts. We called the second start, this start also gives a 4 bp overlap with the previous gene.

Thanks!