

Annotation Review Process

1. Randomly select sequences with and without errors from the pipeline output OR list of differences between annotation output and Virus Variation Resource database

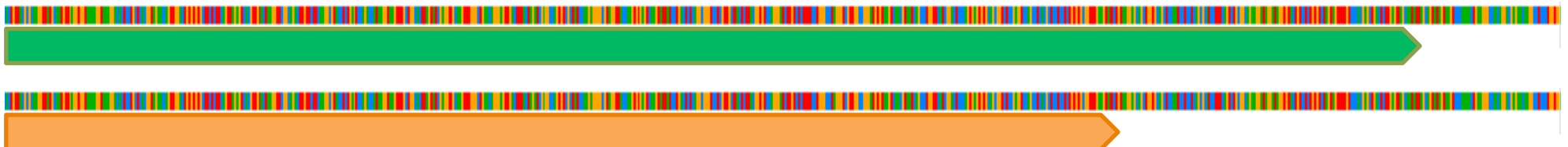
```
160 HM488171 16 MP#16 trc in-frame stop codon exists 5' of stop position predicted by homology to reference [homolog
161 HM488171 19 CDS(MP)#3 trc in-frame stop codon exists 5' of stop position predicted by homology to reference [homolog
162 HQ671730 15 MP#15 ext first in-frame stop codon exists 3' of stop position predicted by homology to reference [h
163 HQ671730 18 CDS(MP)#2 stp predicted CDS stop position is not end of valid stop codon (TAG|TAA|TGA) [CAG ending at p
164 HQ671730 18 CDS(MP)#2 ext first in-frame stop codon exists 3' of stop position predicted by homology to reference [h
165 JX503098 14 MP#14 trc in-frame stop codon exists 5' of stop position predicted by homology to reference [homolog
```

OR

```
1 FJ461303: NS2B 4108-4500 (file) None-None (DB)
2 FJ461303: NS3 4501-6357 (file) None-None (DB)
3 KF955440: NS5 7552-9873 (file) None-None (DB)
4 HQ166035: NS3 4465-4572 (file) 4464-6320 (DB)
5 KF955446: NS4A 6341-6720 (file) 6341-6721 (DB)
6 JQ922552: NS4B 6799-7539 (file) 6799-7542 (DB)
```

2. Use alignments of the RefSeq(s) and the selected sequences to verify or correct each error, and identify errors that may have been missed.

NC_009942 MP#16 (NS1')



HM488171 MP#16 trc in-frame stop codon exists 5' of stop position predicted by homology to reference [homology search predicted 3553..3681 revised to 3553..3603 (stop shifted 78 nt)]

3. Rodney's group provides Alejandro and Eric with a list suggesting improvements that should be made in the code.