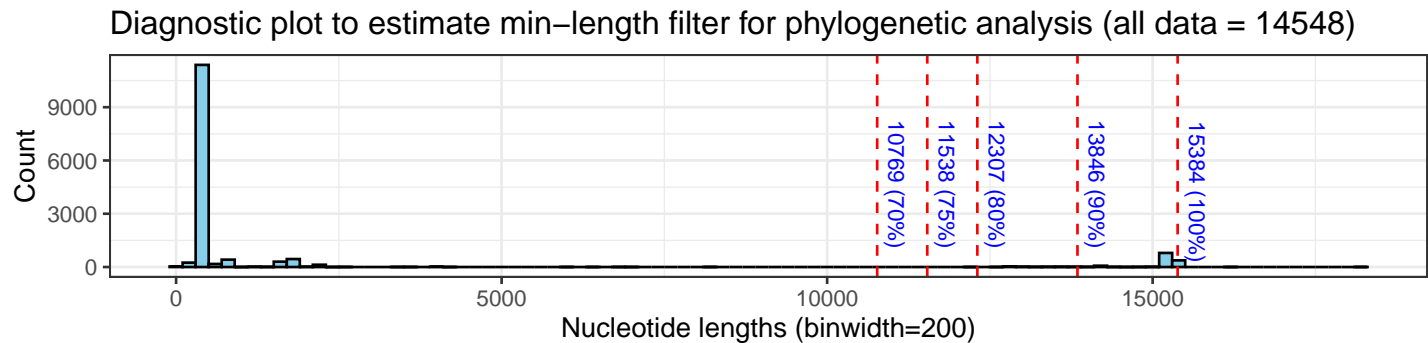
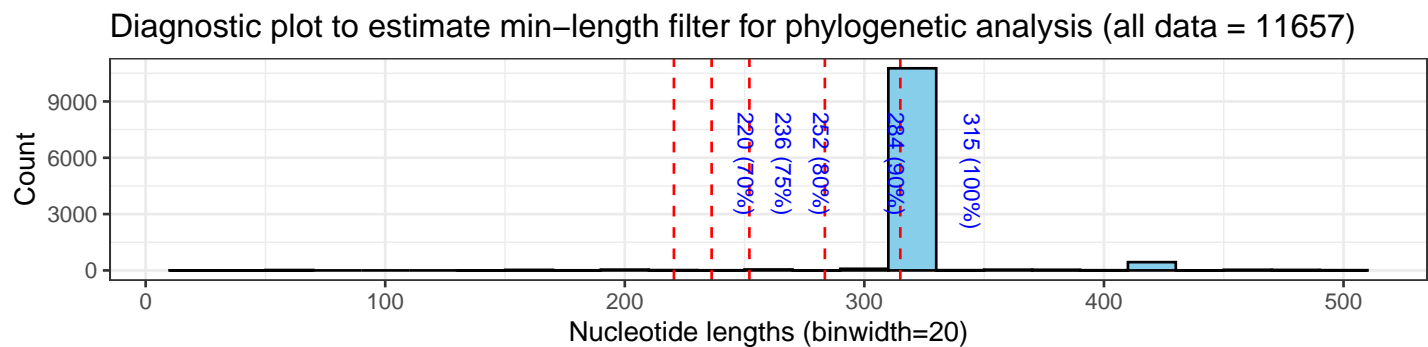


Nextclade Result Parameter Diagnostic Plots

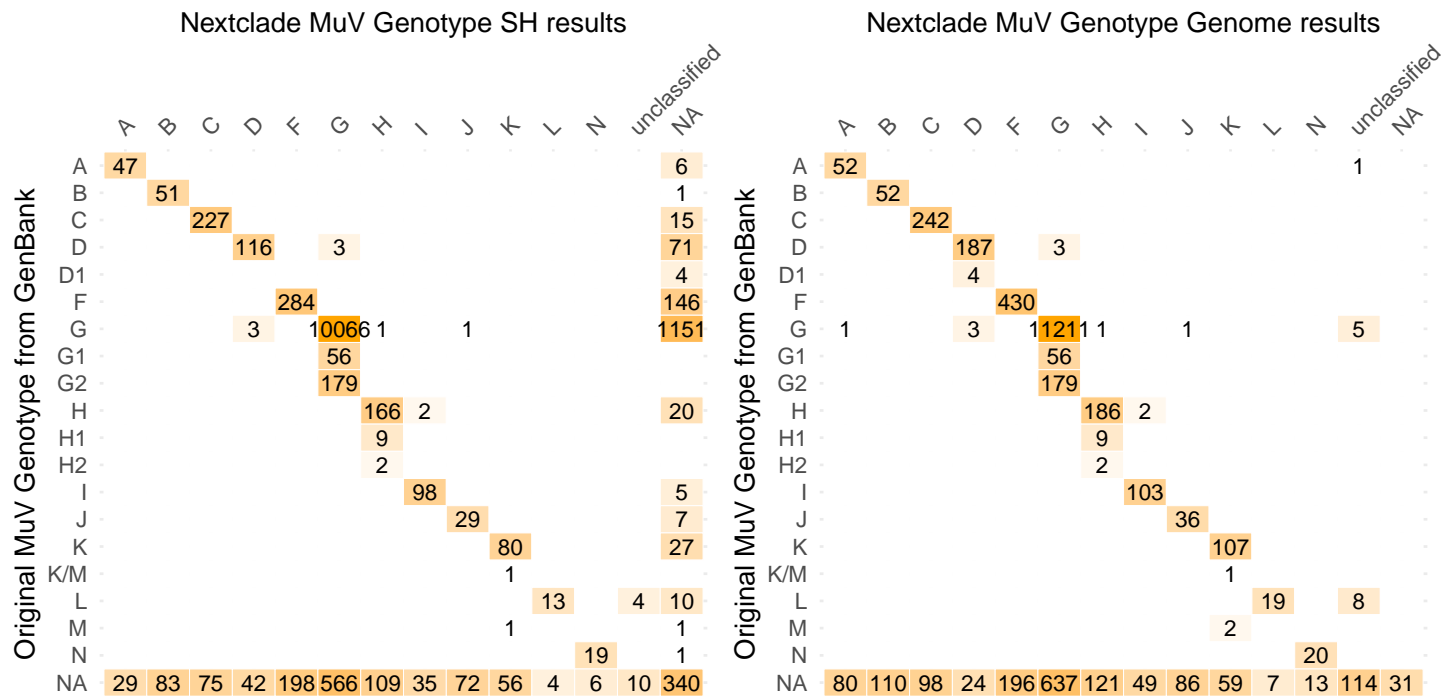
Full Genome



SH Region



Nextclade compared with GenBank annotations:



SH gene Nextclade Dataset

Alignment check

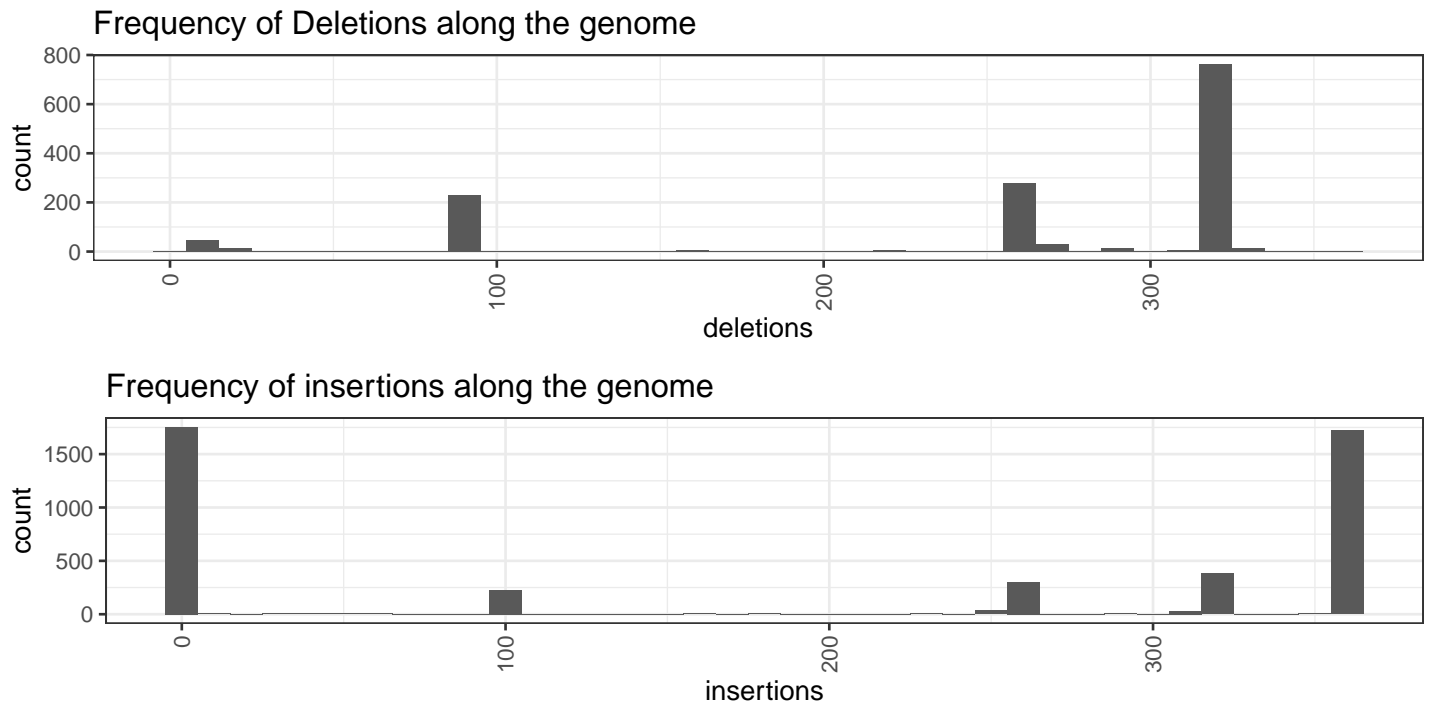


Table 1: Top Frameshifts

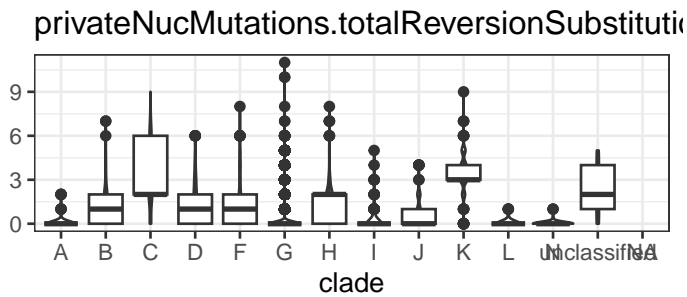
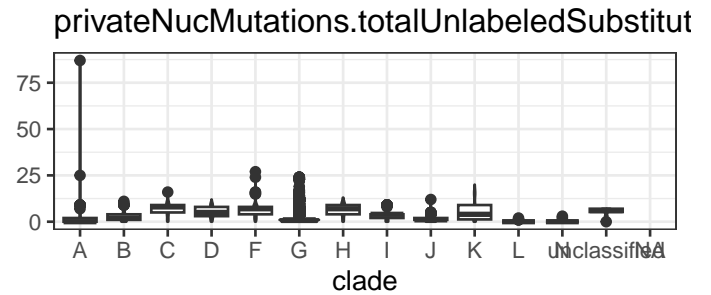
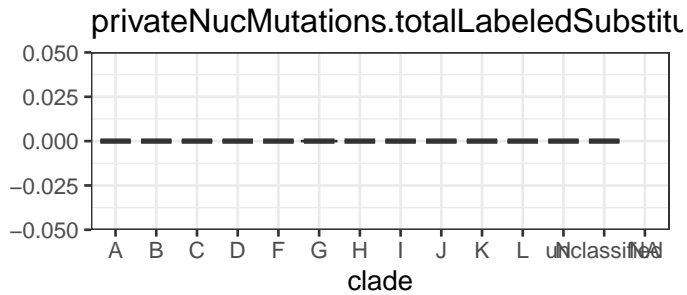
frameShifts	n
SH:7	227
SH:empty range	32
SH:49	4
SH:27-28	3
SH:27-58	2
SH:28	2
SH:31-33	2
SH:15-58	1
SH:18-26	1
SH:2-3	1

Missing Data

```
--scoreBias 300          # First 300 not penalized
--missingDataThreshold 3000 # Top threshold, scale the next set of missing data from 0 to 100
```

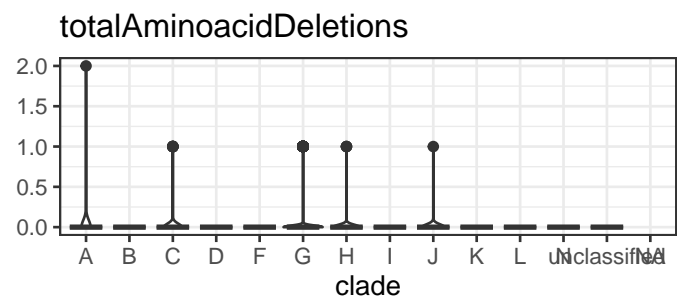
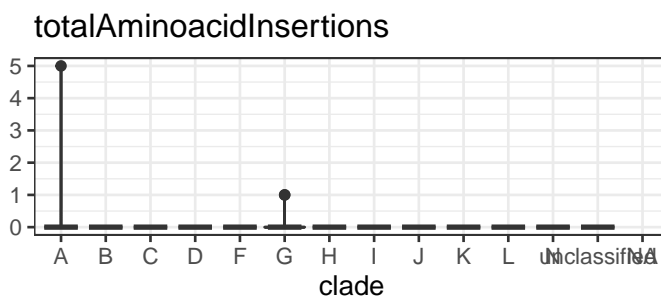
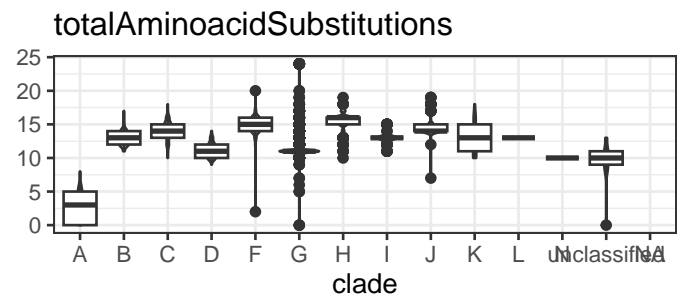
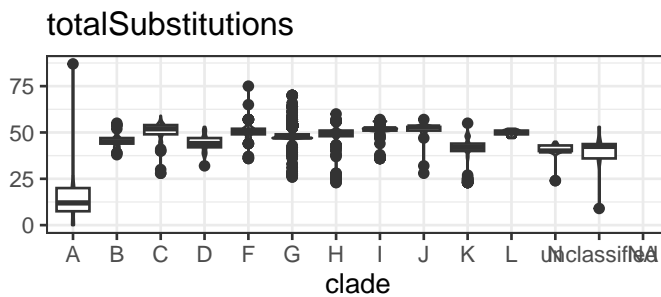

Private Mutations

```
"privateMutations": {
  "enabled": true,
  "typical": 8,
  "cutoff": 24,
  "weightLabeledSubstitutions": 4,
  "weightReversionSubstitutions": 6,
  "weightUnlabeledSubstitutions": 1
},
```



Mutation clusters

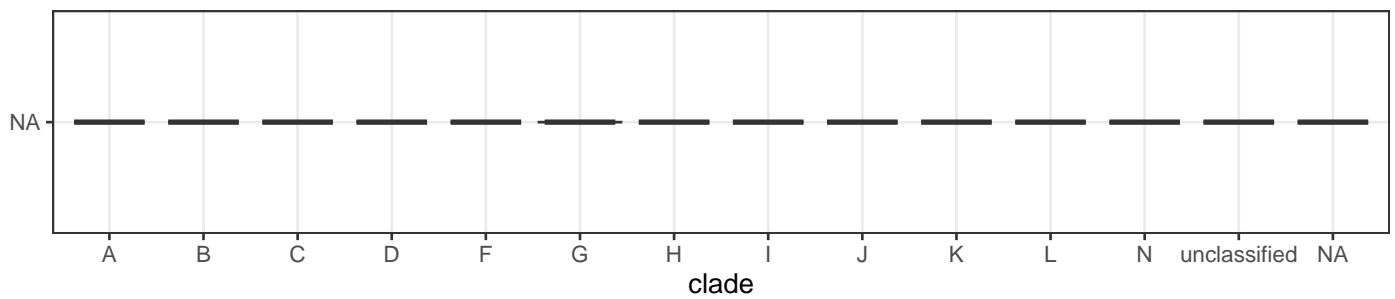
```
"snpClusters": {
  "enabled": true,
  "windowSize": 100,
  "clusterCutOff": 6,
  "scoreWeight": 50
},
```



Stop Codons

```
"stopCodons": {  
  "enabled": true,  
  "ignoredStopCodons": [  
    {  
      "geneName": "ORF8",  
      "codon": 26  
    },  
    {  
      "geneName": "ORF8",  
      "codon": 67  
    }  
  ]  
}
```

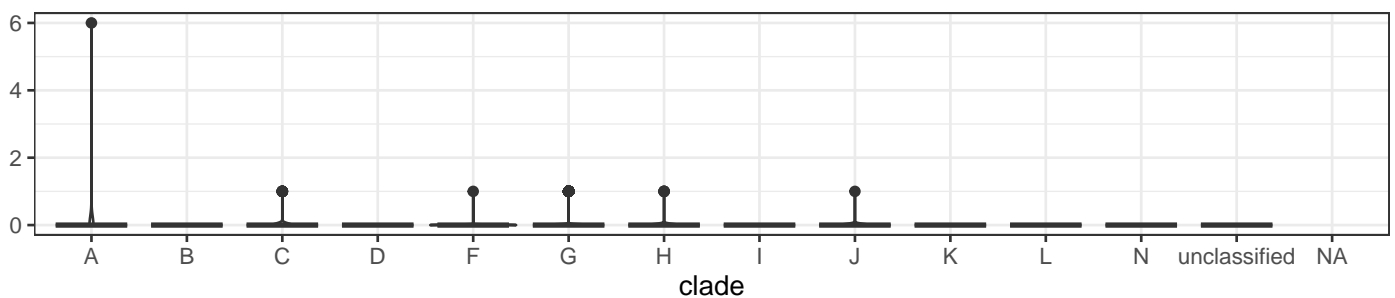
qc.stopCodons.totalStopCodons



Frameshifts

```
"frameShifts": {  
  "enabled": true,  
  "ignoredFrameShifts": [  
    {  
      "geneName": "ORF3a",  
      "codonRange": {  
        "begin": 256,  
        "end": 276  
      }  
    },  
    {  
      "geneName": "ORF3a",  
      "codonRange": {  
        "begin": 258,  
        "end": 276  
      }  
    }  
  ]  
},
```

totalFrameShifts



Genome Nextclade Dataset

Alignment check

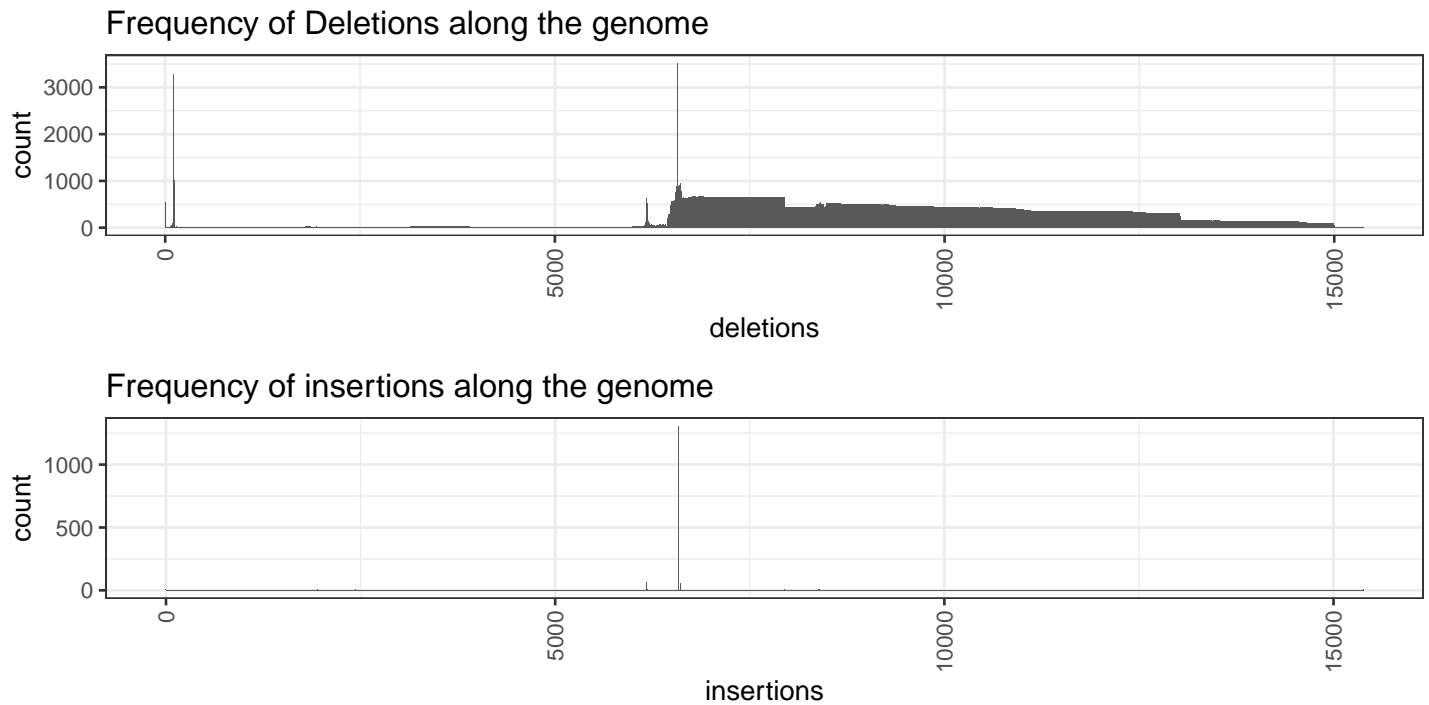
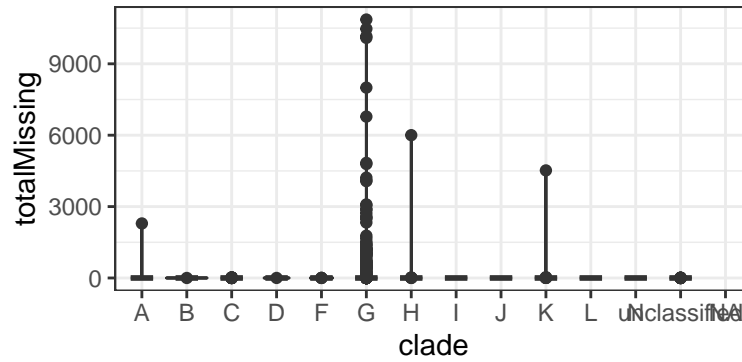


Table 2: Top Frameshifts

frameShifts	n
HN:445-446	7
P:156-391	6
SH:55-58	4
L:1385-2262	3
NP:114-550	2
SH:27-58	2
F:483-488	1
F:495-503	1
F:503-510	1
F:517-520	1

Missing Data

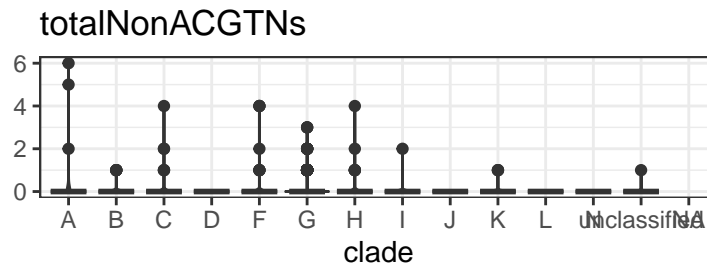
```
--scoreBias 300          # First 300 not penalized
--missingDataThreshold 3000 # Top threshold, scale the next set of missing data from 0 to 100
```



The full genome does reach the 300 minimum, especially for samples that are only reporting one gene (SH region or 315 out of 14k genome length)

Mixed Sites (M)

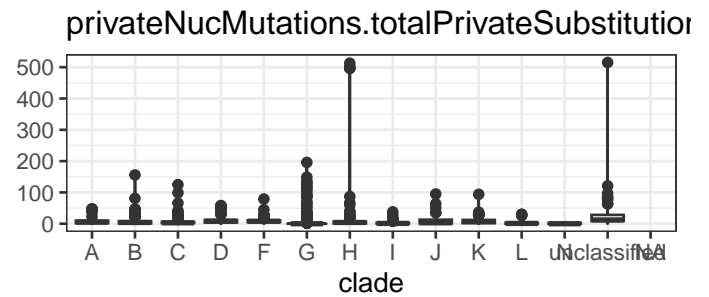
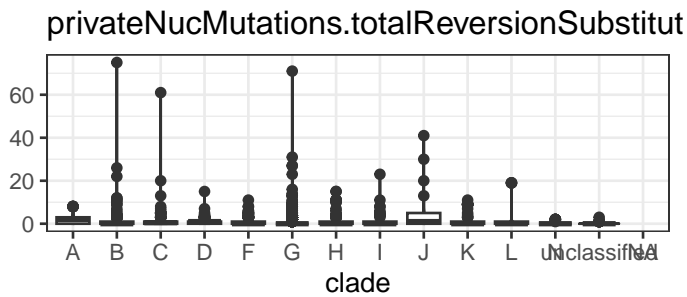
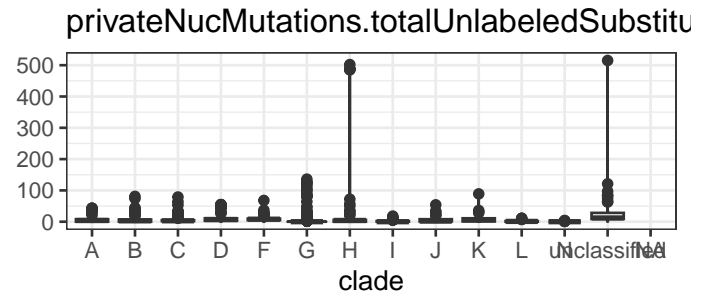
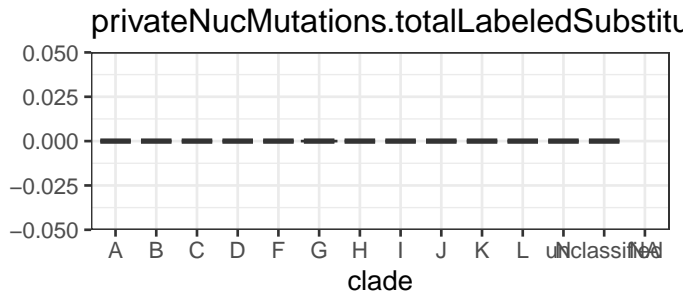
```
--mixedSitesThreshold 10 # More than this number of mixed sites results in a "bad" score
```



This is still well below the 10 threshold.

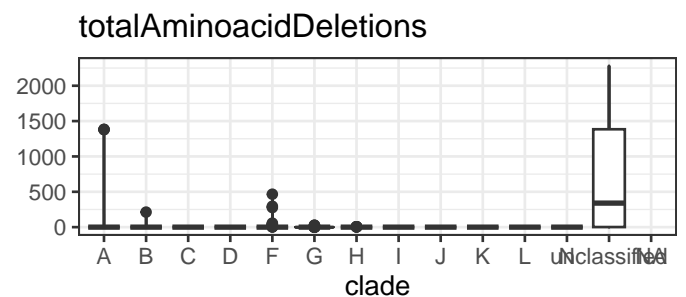
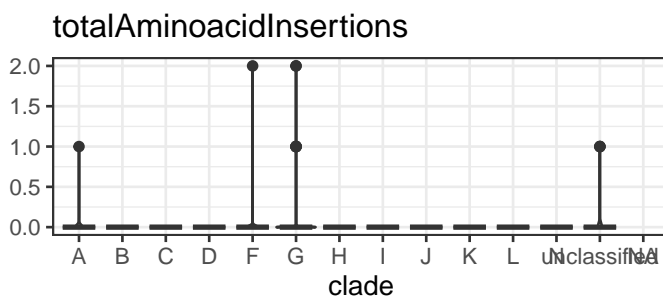
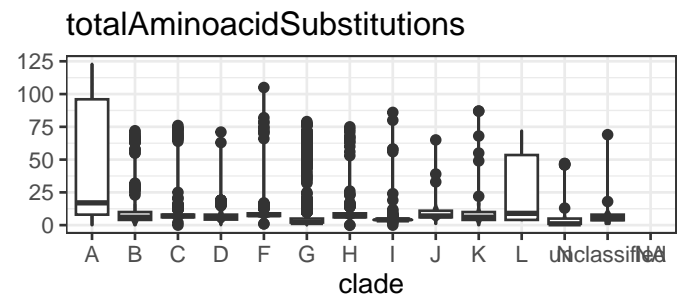
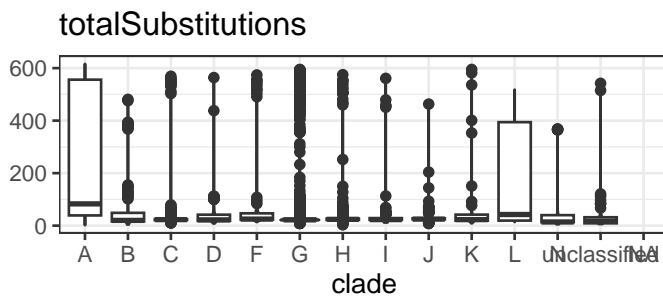
Private Mutations

```
"privateMutations": {
  "enabled": true,
  "typical": 8,
  "cutoff": 24,
  "weightLabeledSubstitutions": 4,
  "weightReversionSubstitutions": 6,
  "weightUnlabeledSubstitutions": 1
},
```



Mutation clusters

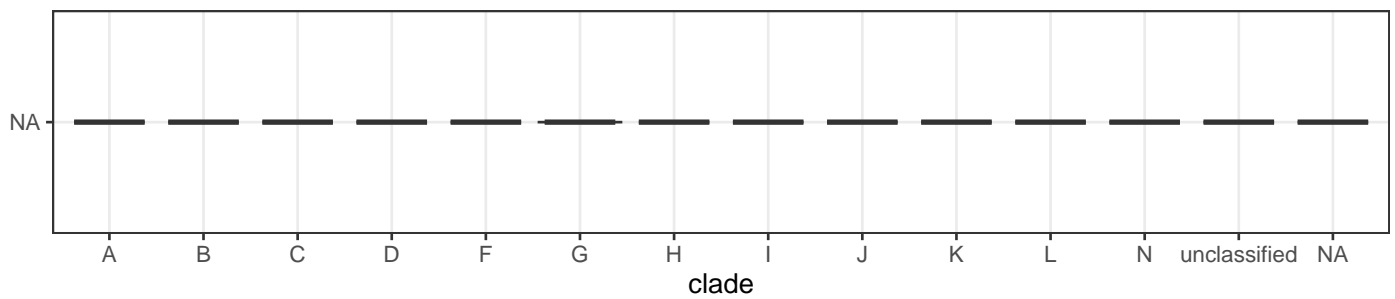
```
"snpClusters": {
  "enabled": true,
  "windowSize": 100,
  "clusterCutoff": 6,
  "scoreWeight": 50
},
```



Stop Codons

```
"stopCodons": {
  "enabled": true,
  "ignoredStopCodons": [
    {
      "geneName": "ORF8",
      "codon": 26
    },
    {
      "geneName": "ORF8",
      "codon": 67
    }
  ]
}
```

qc.stopCodons.totalStopCodons



Frameshifts

```
"frameShifts": {
  "enabled": true,
  "ignoredFrameShifts": [
    {
      "geneName": "ORF3a",
      "codonRange": {
        "begin": 256,
        "end": 276
      }
    },
    {
      "geneName": "ORF3a",
      "codonRange": {
        "begin": 258,
        "end": 276
      }
    }
  ]
}
```

totalFrameShifts

