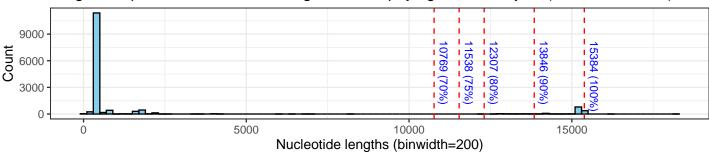
## Nextclade Result Parameter Diagnostic Plots

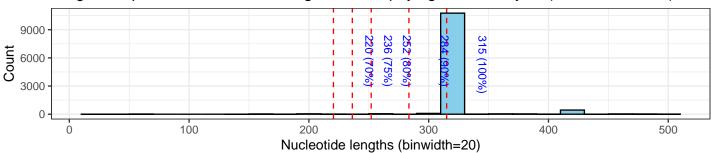
## Full Genome

Diagnostic plot to estimate min–length filter for phylogenetic analysis (all data = 14548)

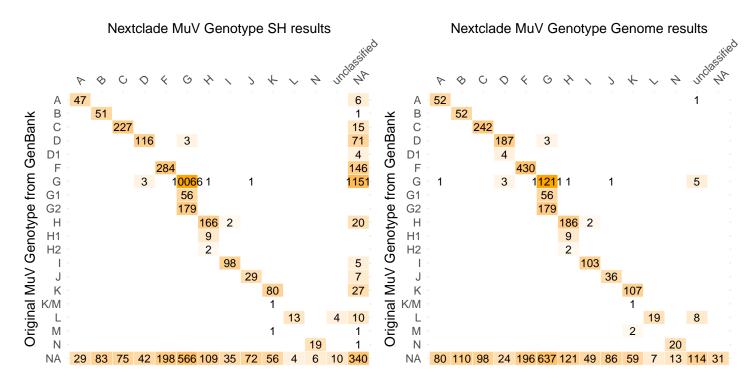


## **SH** Region

Diagnostic plot to estimate min-length filter for phylogenetic analysis (all data = 11657)



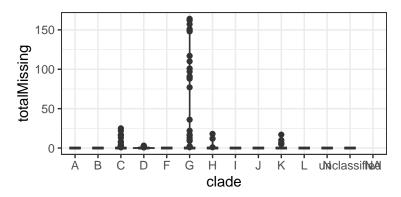
# Nextclade compared with GenBank annotations:



# SH gene Nextclade Dataset

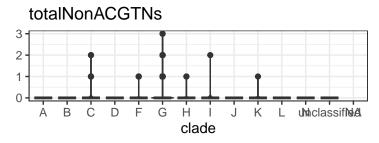
## Missing Data

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



Since this is the SH gene ( 315nt), we do not reach past the 300 minimum. Mixed Sites (M)

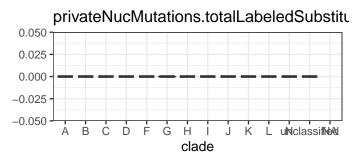
--mixedSitesThreshold 10 # More than this number of mixed sites results in a "bad" score

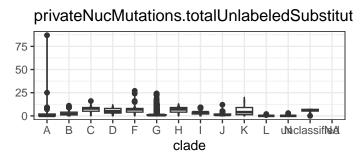


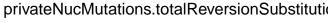
Well below the 10 threshold (but once again this is for a reference that is 315nt long).

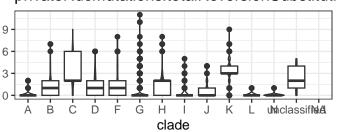
#### **Private Mutations**

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

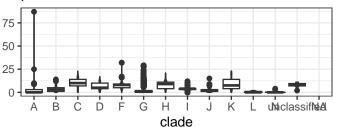








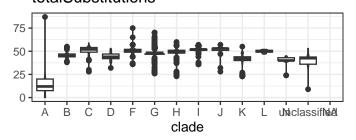


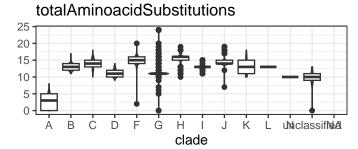


#### Mutation clusters

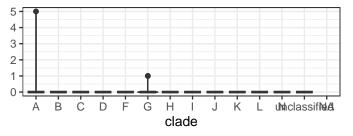
```
"snpClusters": {
   "enabled": true,
   "windowSize": 100,
   "clusterCutOff": 6,
   "scoreWeight": 50
},
```

## totalSubstitutions

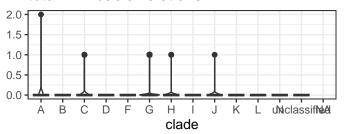




## totalAminoacidInsertions

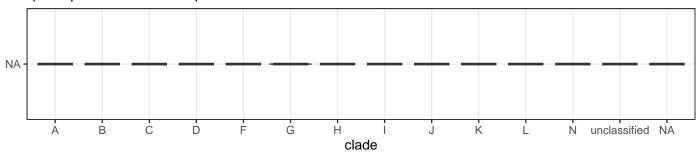


## totalAminoacidDeletions



## **Stop Codons**

# qc.stop Codons.total Stop Codons



## Frameshifts

## totalFrameShifts

