

Metadata check

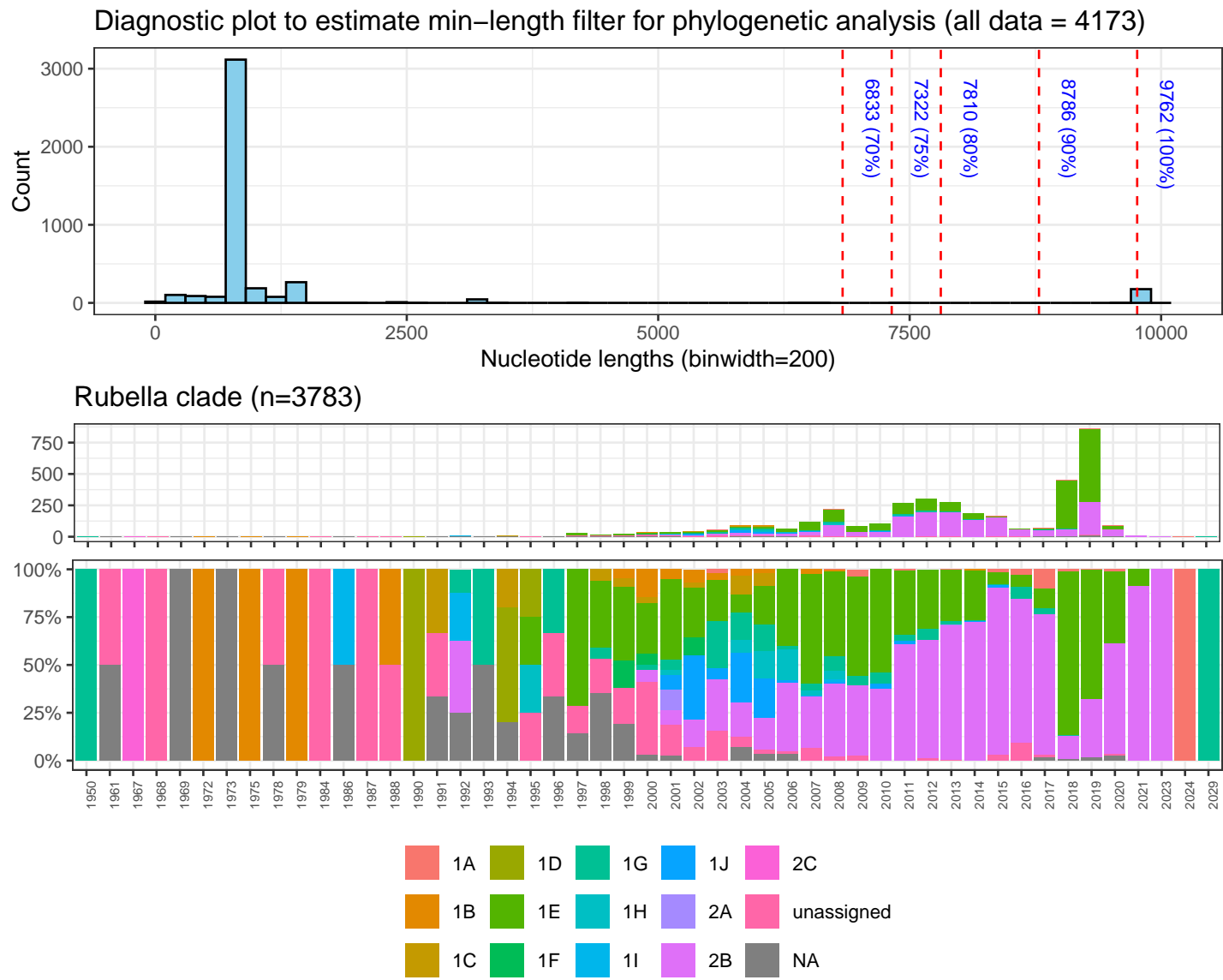


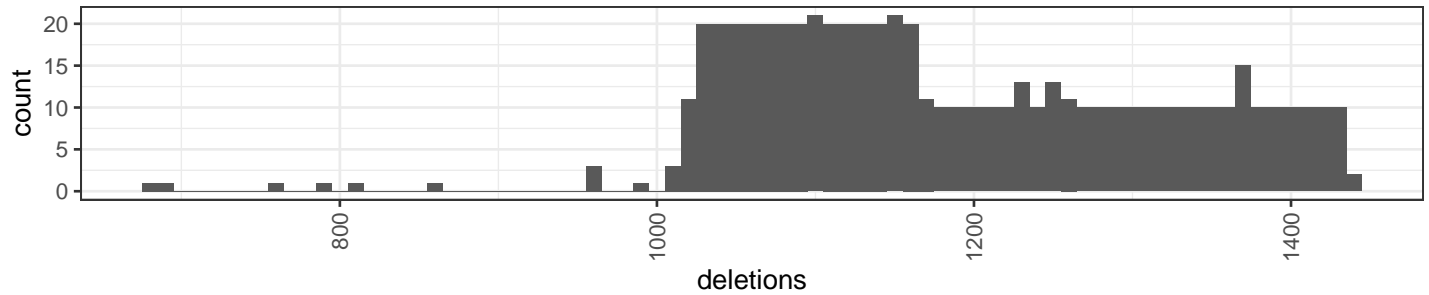
Table 1: Date check

accession	strain	date	date_released	date_updated
EU016200	RVs/Nijmegen.NET/05.29/TT	2029-01-XX	2007-07-30	2016-07-26

Nextclade alignment check

```
"alignmentParams": {
  "minSeedCover": 0.01,
  "terminalBandwidth": 100,
  "excessBandwidth": 18,
  "minMatchLength": 20
},
```

Frequency of deletions along rubella



Frequency of insertions along rubella

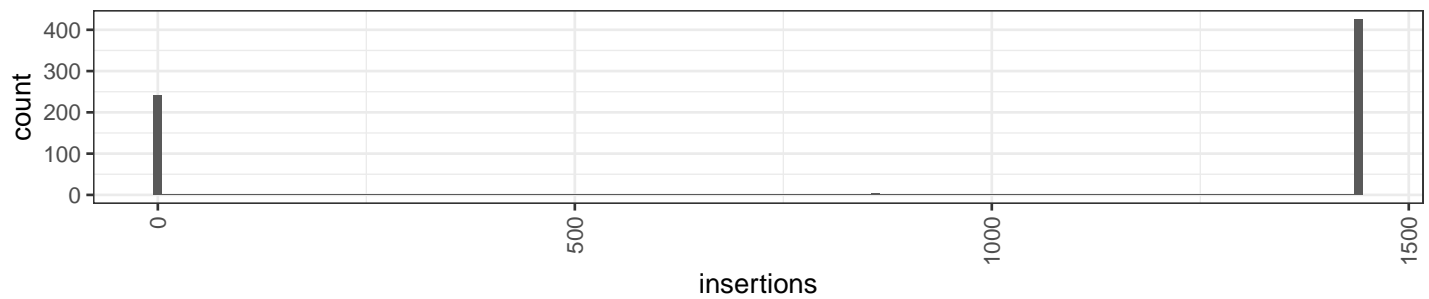


Table 2: Top Frameshifts

frameShifts	n
E1:290-481	4
E1:228-266	1
E1:232-481	1
E1:266-267	1
E1:272-331	1
E1:312-333	1
E1:342-359	1
E1:367-481	1
E1:385-416	1
E1:390-481	1

```
"missingData": {
  "enabled": true,
  "missingDataThreshold": 20, # Top threshold, scale the next set of missing data from 0 to 100
  "scoreBias": 4 # First 4 not penalized
},
```

Mixed Sites (M)

totalNonACGTNs

clade

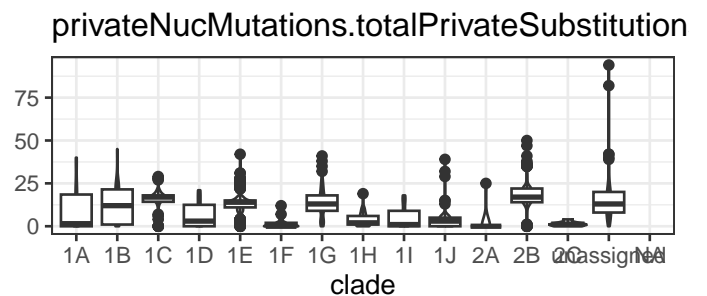
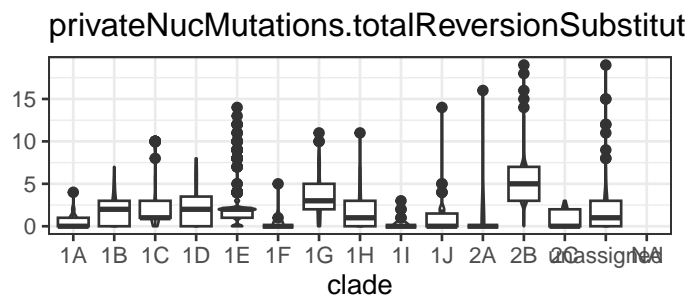
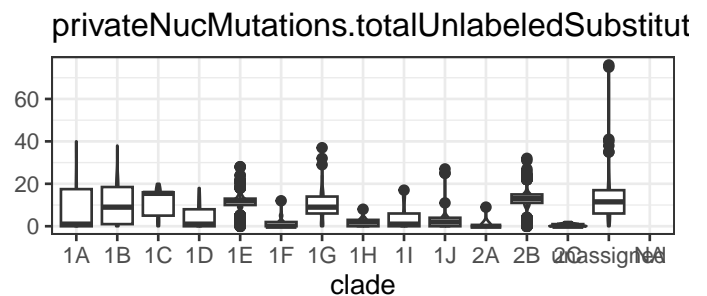
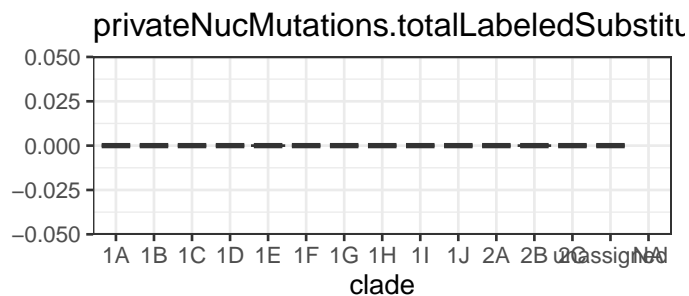
clade	totalNonACGTNs
1A	1
1A	2
1A	3
1A	4
1A	5
1A	6
1A	7
1A	8
1A	9
1A	10
1B	1
1B	2
1B	3
1B	4
1B	5
1C	1
1C	2
1C	3
1C	4
1C	5
1D	1
1D	2
1D	3
1D	4
1D	5
1E	1
1E	2
1E	3
1E	4
1E	5
1F	1
1F	2
1F	3
1F	4
1F	5
1G	1
1G	2
1G	3
1G	4
1G	5
1H	1
1H	2
1H	3
1H	4
1H	5
1I	1
1I	2
1I	3
1I	4
1I	5
1J	1
1J	2
1J	3
1J	4
1J	5
2A	1
2A	2
2A	3
2A	4
2A	5
2B	1
2B	2
2B	3
2B	4
2B	5
2B	6
2B	7
2B	8
2B	9
2B	10
2B	11
2B	12
2B	13
2B	14
2B	15
2B	16
2B	17
2B	18
2B	19
2B	20
2B	21
2B	22
2B	23
2B	24
2C	1
2C	2
2C	3
2C	4
2C	5
2C	6
2C	7
2C	8
2C	9
2C	10
assigned	1
assigned	2
assigned	3
assigned	4
assigned	5
assigned	6
assigned	7
assigned	8
assigned	9
assigned	10
NA	1
NA	2
NA	3
NA	4
NA	5
NA	6
NA	7
NA	8
NA	9
NA	10

Table 4: Records above totalNonACGTNs threshold

seqName	totalNonACGTNs
PV035063	24
PV035062	13
KT369573	10
OQ296620	10
L19420	9
PV035058	9
OM022831	8
PP536671	8
PV035060	8
MK780810	7
PV035061	7
U58299	7
AF039106	5
L16230	5

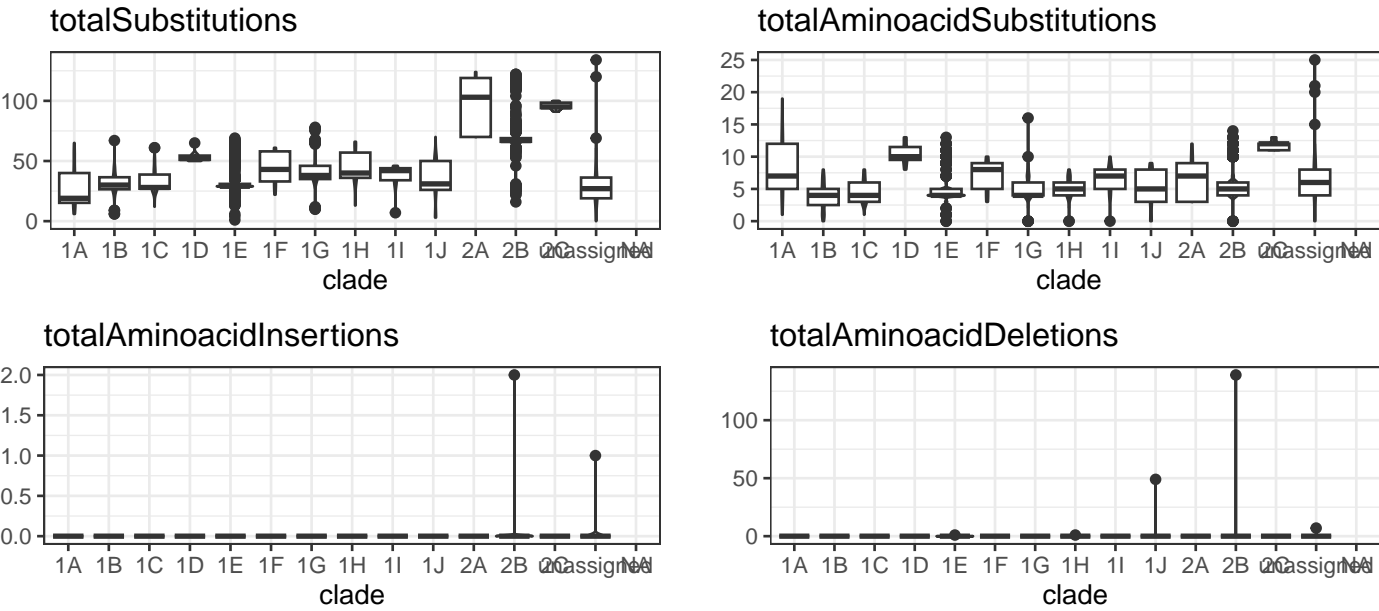
Private Mutations

```
"privateMutations": {
  "enabled": true,
  "cutoff": 45,
  "typical": 4,
  "weightLabeledSubstitutions": 1,
  "weightReversionSubstitutions": 1,
  "weightUnlabeledSubstitutions": 1
},
```



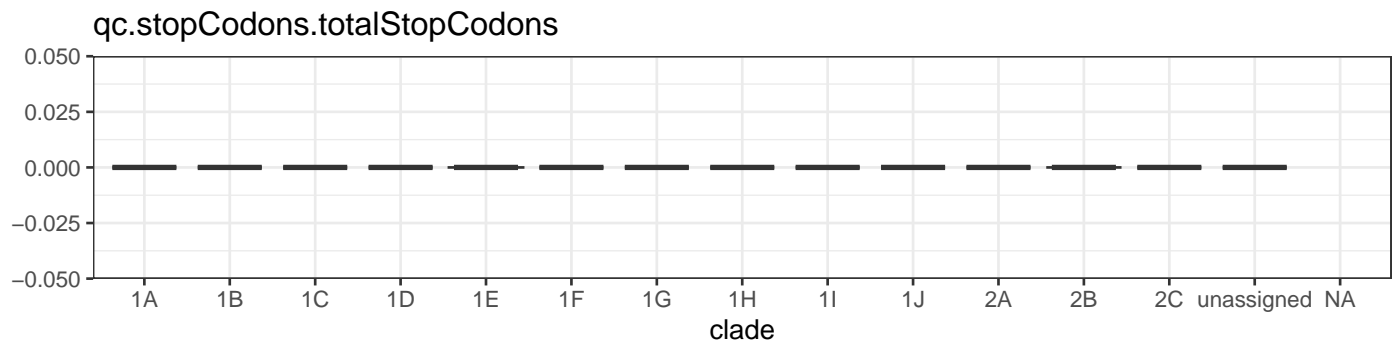
Mutation clusters

```
"snpcClusters": {
  "enabled": true,
  "clusterCutOff": 10,
  "scoreWeight": 50,
  "windowSize": 50
}
```



Stop Codons

```
# Example, not set in rubella
"stopCodons": {
  "enabled": true,
  "ignoredStopCodons": [
    {
      "geneName": "ORF8",
      "codon": 26
    },
    {
      "geneName": "ORF8",
      "codon": 67
    }
  ]
}
```



Frameshifts

```
"frameShifts": {
  "enabled": true,
  # Example known frame shifts, enabled in rubella but no ignored ones set
  "ignoredFrameShifts": [
    {
      "geneName": "ORF3a",
      "codonRange": {
        "begin": 256,
        "end": 276
      }
    },
    {
      "geneName": "ORF3a",
      "codonRange": {
        "begin": 258,
        "end": 276
      }
    }
  ]
},
```

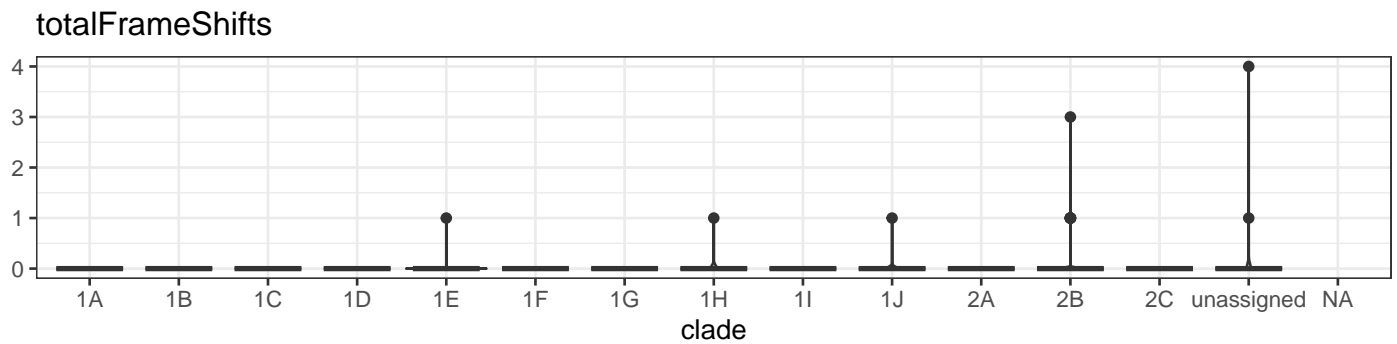


Table 5: Top Frameshifts

frameShifts	n
E1:290-481	4
E1:228-266	1
E1:232-481	1
E1:266-267	1
E1:272-331	1
E1:312-333	1
E1:342-359	1
E1:367-481	1
E1:385-416	1
E1:390-481	1

Table 6: Records with top frameShifts

seqName	frameShifts
FR693375	E1:290-481
FR693376	E1:290-481
FR693377	E1:290-481
FR693378	E1:290-481