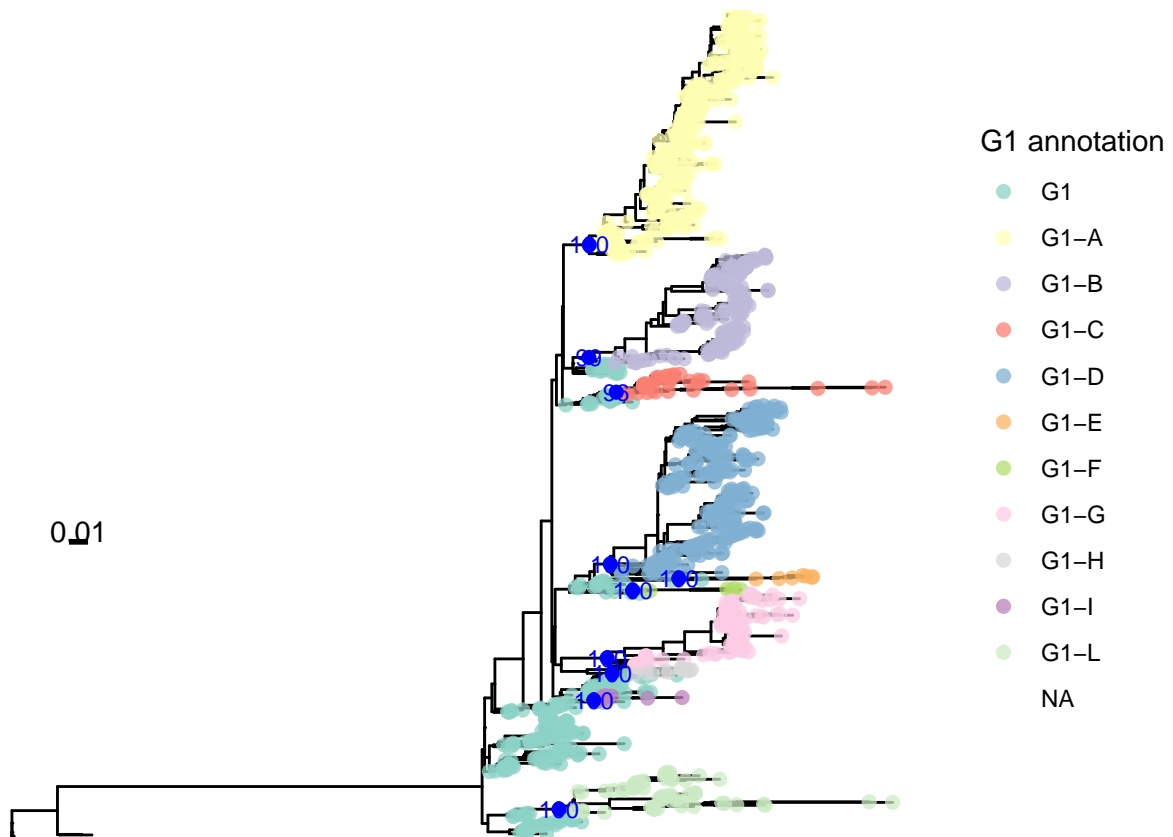


H9_G1_KOR_pairwise_nodedistance

Jiani Chen

9/9/2021

G1 clades defined from annotation files

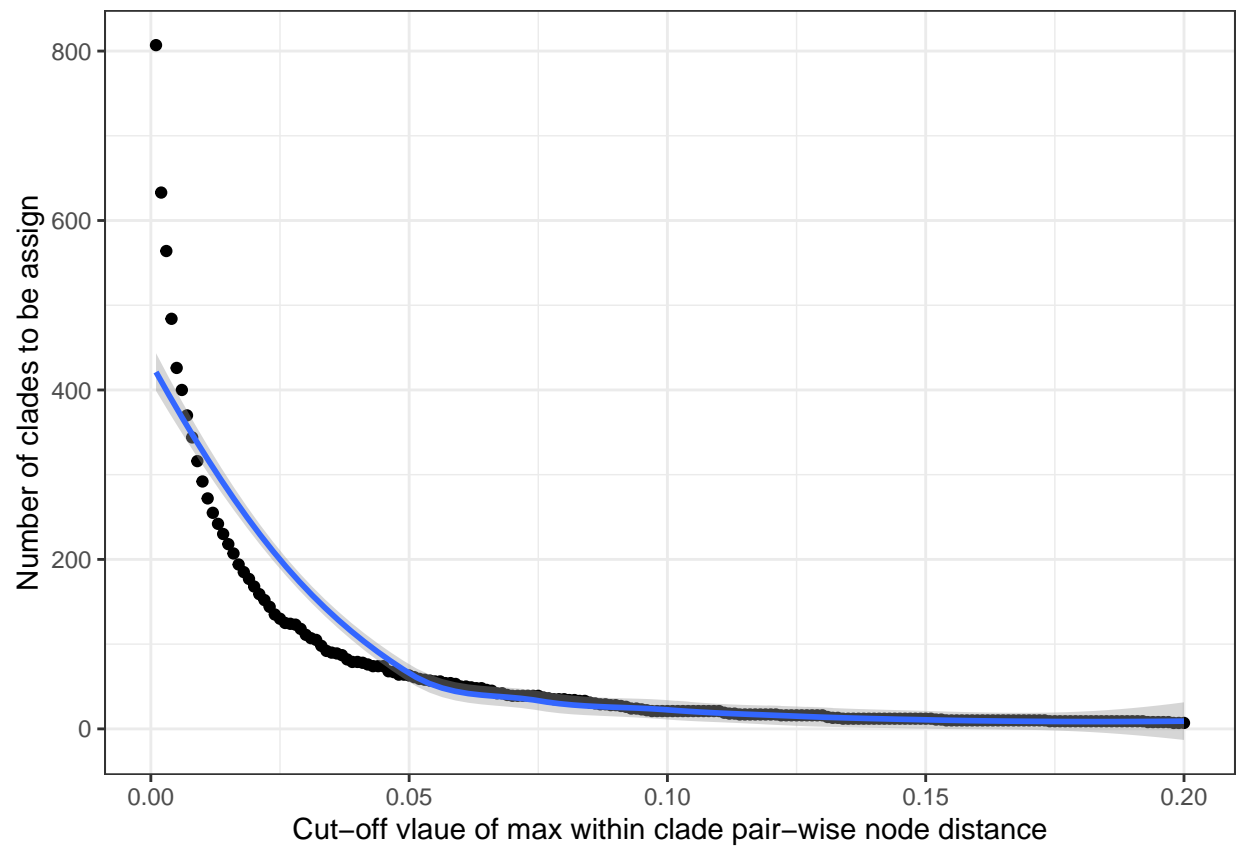


Calculate the maximal pair-wise node distance within the assigned clades

```
##      clade max.within.clade.pair.nodedist
## 1    G1-A          0.15301150
## 2    G1-B          0.13781677
## 3    G1-C          0.13020160
## 4    G1-D          0.19205041
## 5    G1-E          0.05739515
```

| | | |
|-------|------|------------|
| ## 6 | G1-F | 0.04719568 |
| ## 7 | G1-G | 0.10645165 |
| ## 8 | G1-H | 0.02182253 |
| ## 9 | G1-I | 0.02052572 |
| ## 10 | G1-L | 0.13336139 |

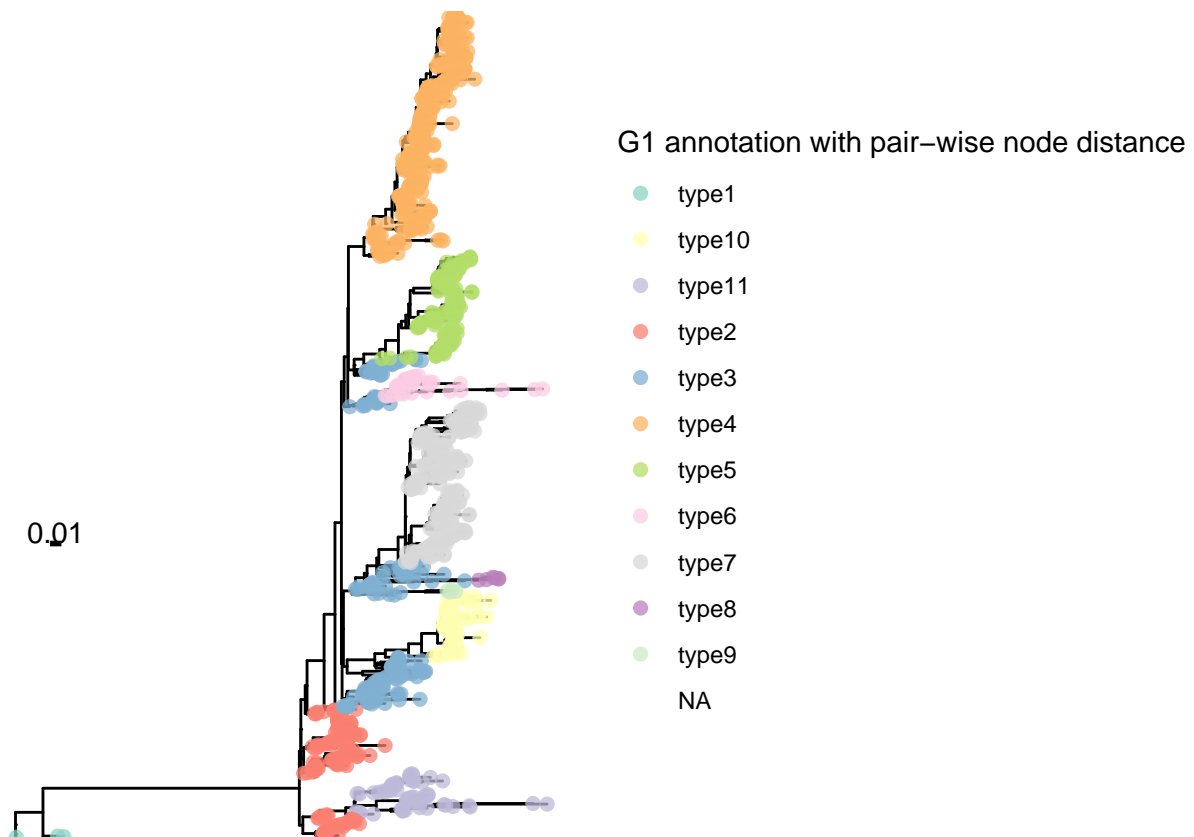
Simulation for G1-clade



I re-assign the clades with maximal pair-wise node distance within clade as 0.16 subs/site

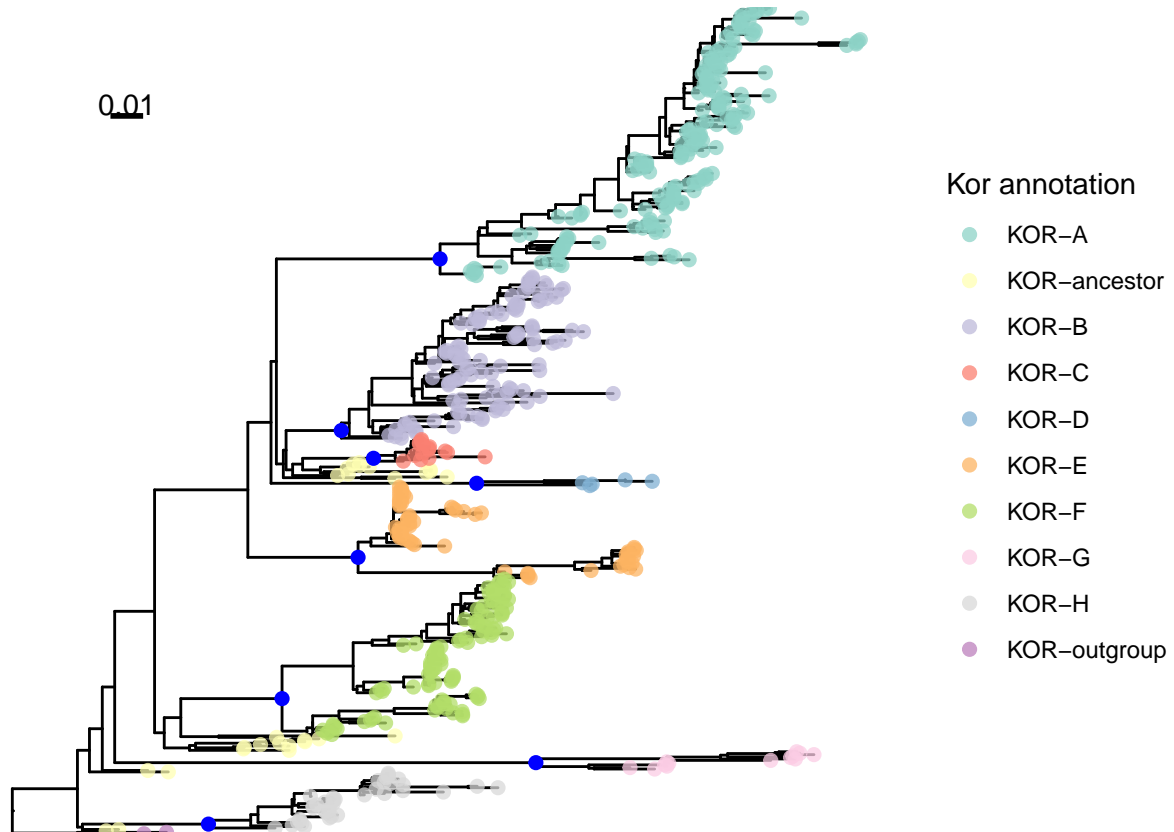
Similar results from the clades defined in annotation file

G1-D clade, small propotion of sequences was unable to classified with this approach (type 8, green)



Analysis for KOR clade

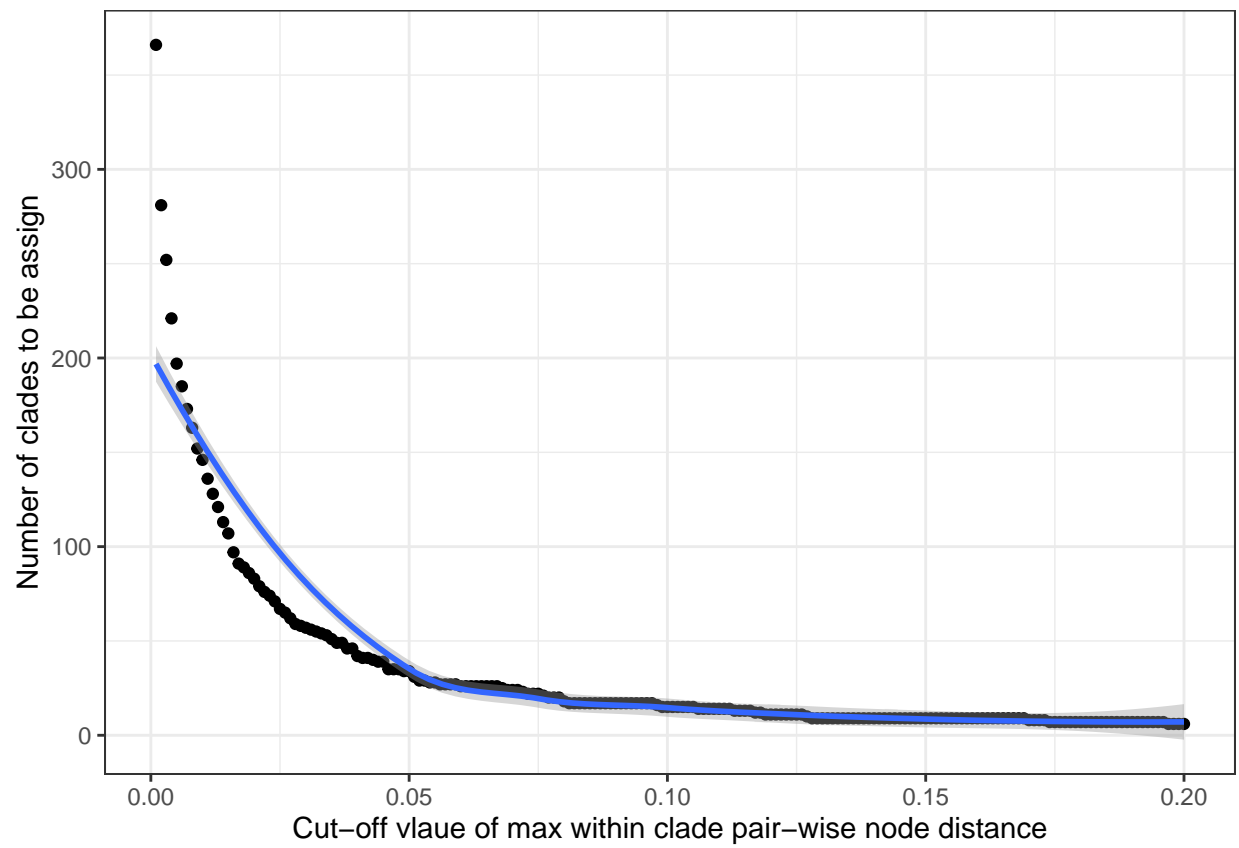
Kor clades defined from annotation files



Calculate the maximal pair-wise node distance within the assigned clades

```
##   clade max.within.clade.pair.nodedist
## 1 KOR-A 0.173678
## 2 KOR-B 0.097446
## 3 KOR-C 0.016257
## 4 KOR-D 0.079447
## 5 KOR-E 0.118378
## 6 KOR-F 0.126374
## 7 KOR-G 0.116707
## 8 KOR-H 0.079360
```

simulation with different maximal within clade pair-wise node distances



I re-assign the clades with maximal pair-wise node distance within clade as 0.16 subs/site

Similar results from the clades defined in annotation files, only differences in KOR-A

KOR-A was divided into two propotion with pair-wise node distance criteria (type1 and type 3)

