Comparative Analysis of SARS-Cov-2 Haut-de-France Viral Sequenced Samples

Vivian Lynn Ly

11 May, 2021

Background and Overview

This is a report on SARS-CoV-2, including some variant analysis (Koyama *et al.*, 2020).

Methods

See the set of tutorials on the vcfR package website.

You may also want to use any of a range of different COVID data packages and data sources:

- https://kjhealy.github.io/covdata/
- https://github.com/como-ph/oxcovid19
- https://ropensci.org/blog/2020/10/20/searching-medrxivr-and-biorxiv-preprint-data/
- https://covidtracking.com/data/api
    - `readr::read_csv("https://api.covidtracking.com/v1/states/daily.csv")`

Subsections are ok too

Results and Discussion

```
## Error in dimnames(x) <- dn: length of 'dimnames' [2] not equal to array extent
```

```
## Error in is.data.frame(stacked_vcf): object 'stacked_vcfs' not found
```

Figures

```
## Error in eval(lhs, parent, parent): object 'vcf_with_metadata' not found
```

Figure 1: N and S genes have more unique SNPs in the set of samples analyzed.

Tables

Gene Name	Start	End	Length
S	21563	25384	3821
ORF3a	25393	26220	827
E	26245	26472	227
M	26523	27191	668
ORF6	27202	27387	185
ORF7a	27394	27759	365
ORF7b	27756	27887	131
ORF8	27894	28259	365
N	28274	29533	1259
ORF10	29558	29674	116

Table 1: Gene names, locations, and lengths in the SARS-CoV-2 genome. Higher SNP counts in the S and N genes may be related to the larger size of these genes.

Sources Cited

Koyama,T. *et al.* (2020) Variant analysis of sars-cov-2 genomes. *Bulletin of the World Health Organization*, **98**, 495.