# SARS-CoV-2 Guangdong genomic epidemiology

Figure data and plots

Louis du Plessis

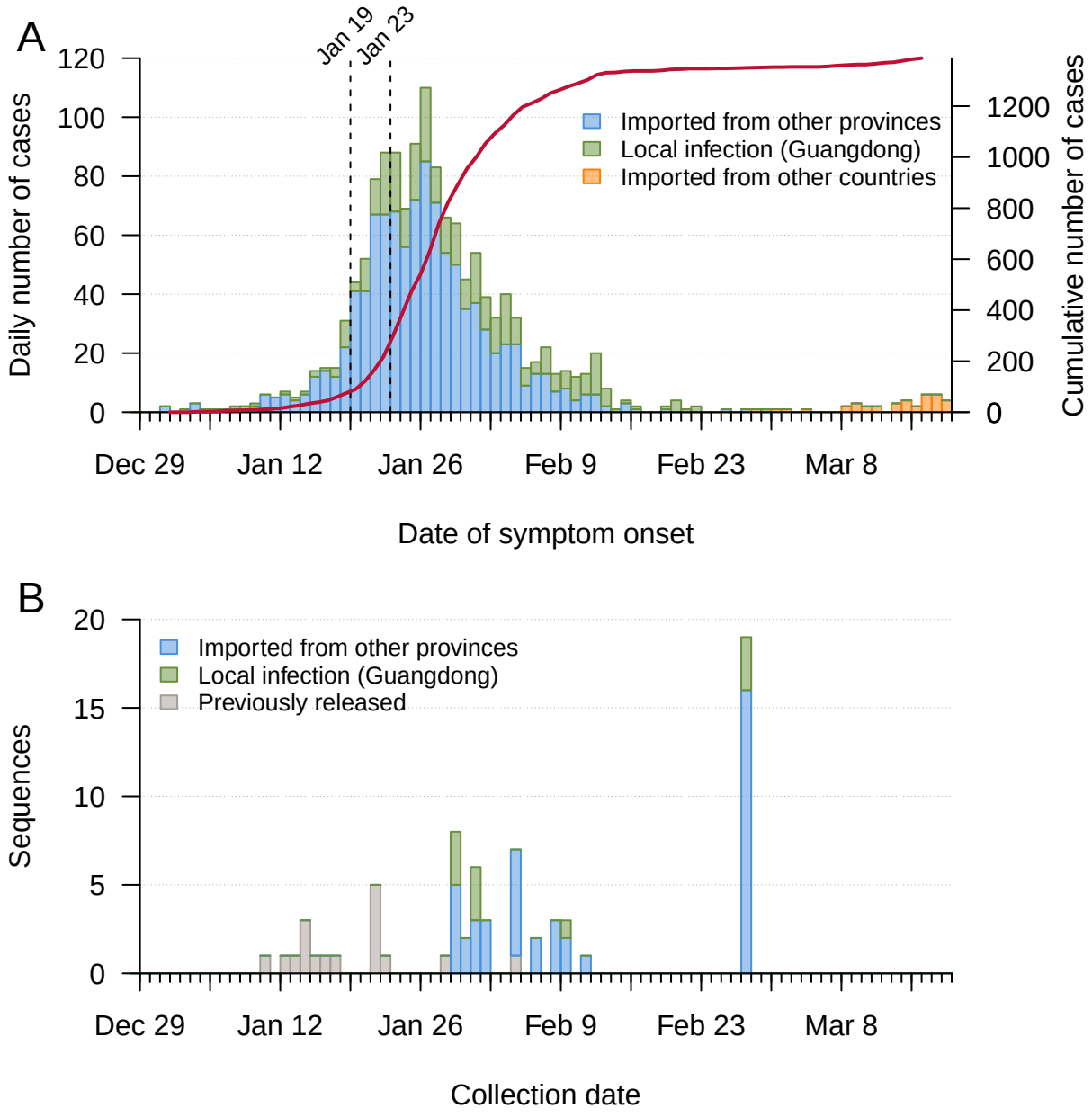
Last modified: 09 Apr 2020

## Process datasets

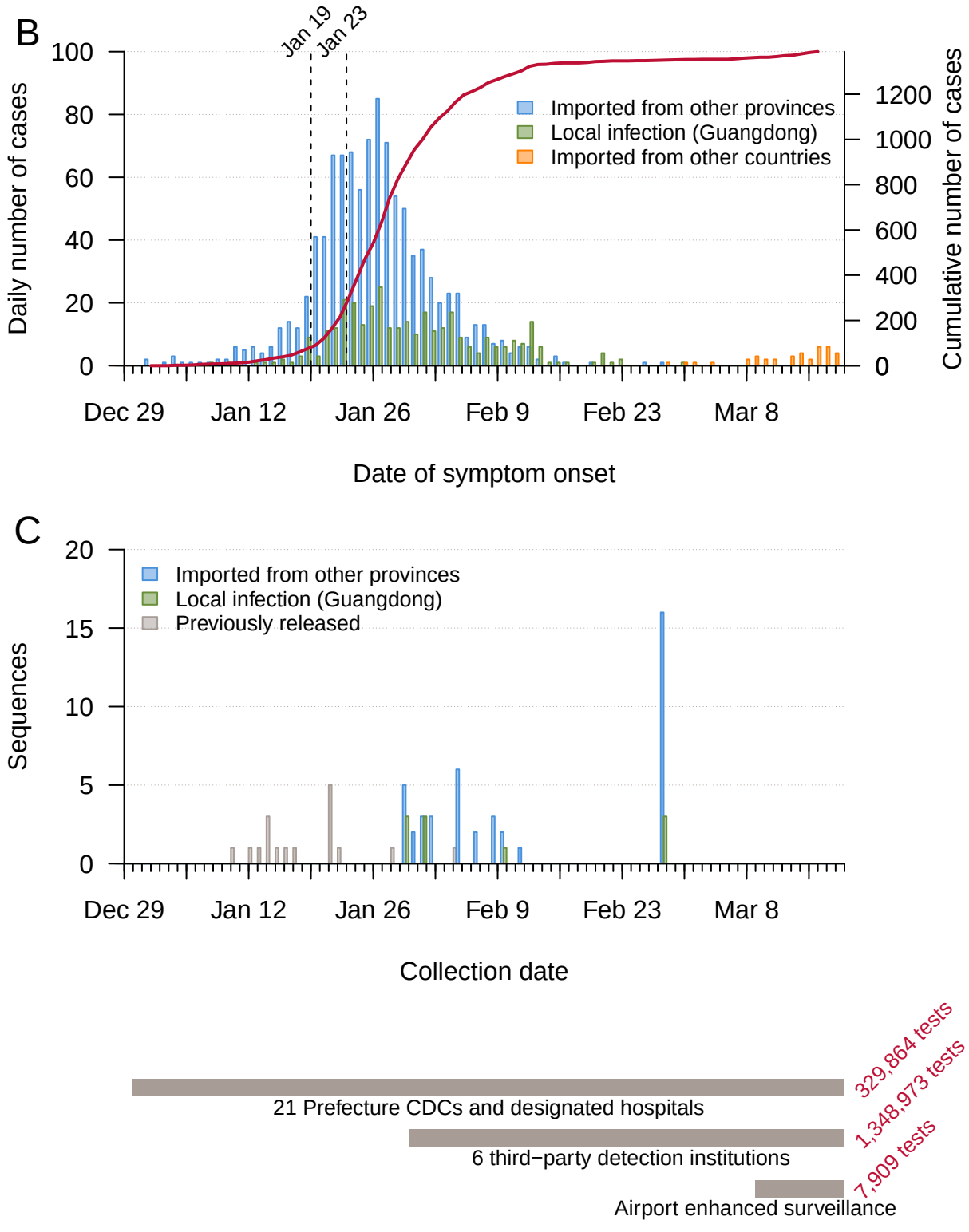
This section does some consistency checks to makes sure different spreadsheets agree with each other (e.g. no samples have different Ct values or collection dates). If no warnings are printed everything is ok.

Checking Ct values and collection dates...

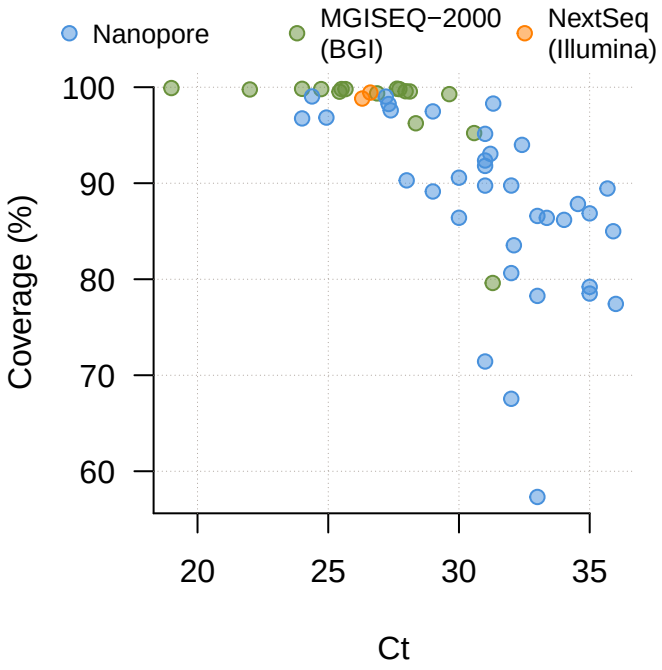
**Figure 1: Cases and Sequences**

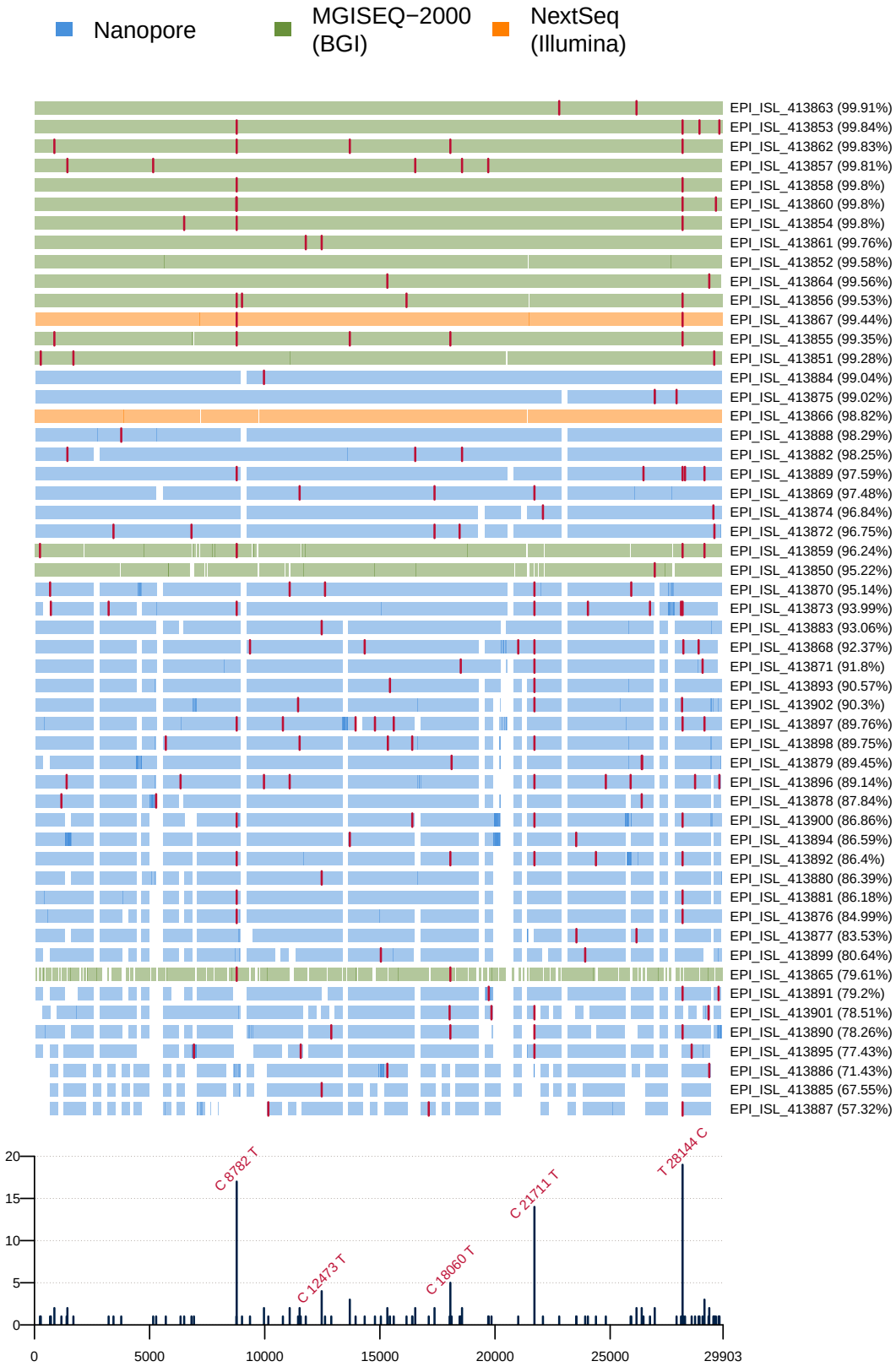


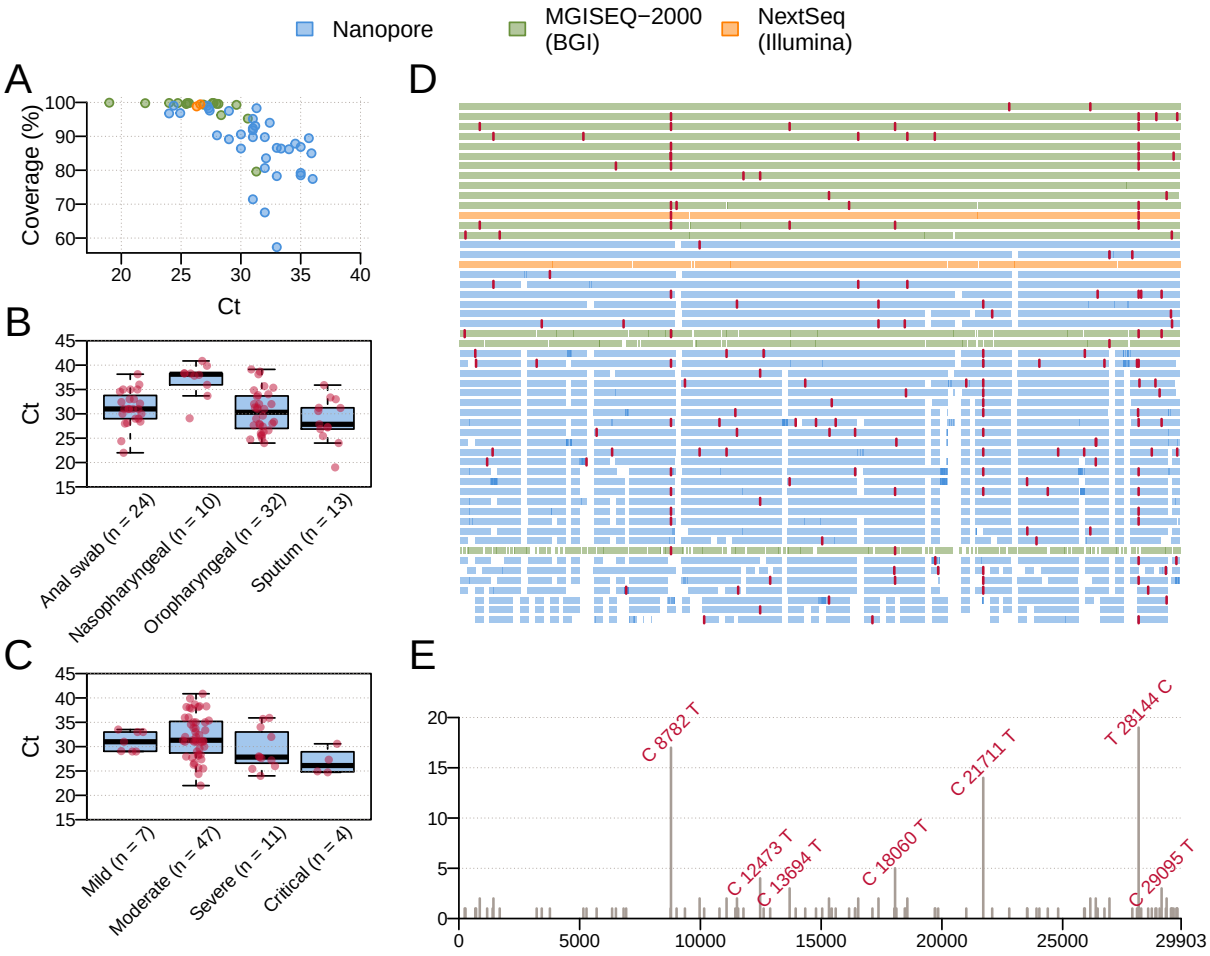
## Fig 1 - Non-stacked



## Figure 2: Sequencing Statistics




Nanopore
MGISEQ−2000 (BGI)
NextSeq (Illumina)
EPI_ISL_413863 (99.91%)
EPI_ISL_413853 (99.84%)
EPI_ISL_413862 (99.83%)
EPI_ISL_413857 (99.81%)
EPI_ISL_413858 (99.8%)
EPI_ISL_413860 (99.8%)
EPI_ISL_413854 (99.8%)
EPI_ISL_413861 (99.76%)
EPI_ISL_413852 (99.58%)
EPI_ISL_413864 (99.56%)
EPI_ISL_413856 (99.53%)
EPI_ISL_413867 (99.44%)
EPI_ISL_413855 (99.35%)
EPI_ISL_413851 (99.28%)
EPI_ISL_413884 (99.04%)
EPI_ISL_413875 (99.02%)
EPI_ISL_413866 (98.82%)
EPI_ISL_413888 (98.29%)
EPI_ISL_413882 (98.25%)
EPI_ISL_413889 (97.59%)
EPI_ISL_413869 (97.48%)
EPI_ISL_413874 (96.84%)
EPI_ISL_413872 (96.75%)
EPI_ISL_413859 (96.24%)
EPI_ISL_413850 (95.22%)
EPI_ISL_413870 (95.14%)
EPI_ISL_413873 (93.99%)
EPI_ISL_413883 (93.06%)
EPI_ISL_413868 (92.37%)
EPI_ISL_413871 (91.8%)
EPI_ISL_413893 (90.57%)
EPI_ISL_413902 (90.3%)
EPI_ISL_413897 (89.76%)
EPI_ISL_413898 (89.75%)
EPI_ISL_413879 (89.45%)
EPI_ISL_413896 (89.14%)
EPI_ISL_413878 (87.84%)
EPI_ISL_413900 (86.86%)
EPI_ISL_413894 (86.59%)
EPI_ISL_413892 (86.4%)
EPI_ISL_413880 (86.39%)
EPI_ISL_413881 (86.18%)
EPI_ISL_413876 (84.99%)
EPI_ISL_413877 (83.53%)
EPI_ISL_413899 (80.64%)
EPI_ISL_413865 (79.61%)
EPI_ISL_413891 (79.2%)
EPI_ISL_413901 (78.51%)
EPI_ISL_413890 (78.26%)
EPI_ISL_413895 (77.43%)
EPI_ISL_413886 (71.43%)
EPI_ISL_413885 (67.55%)
EPI_ISL_413887 (57.32%)
20
15
10
5
0
C 8782 T
C 12473 T
C 18060 T
C 21711 T
T 28144 C
0
5000
10000
15000
20000
25000
29903

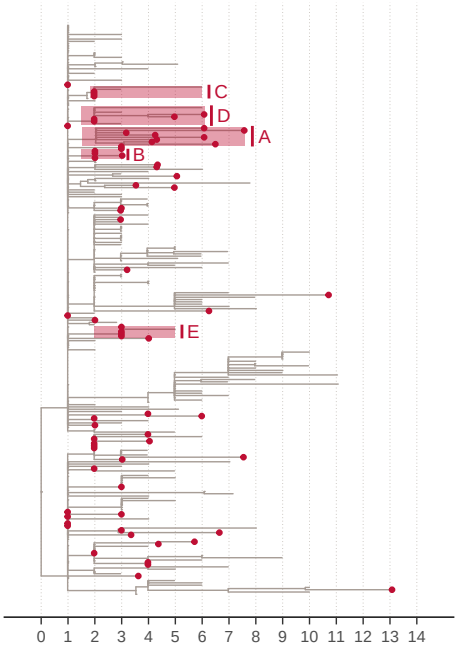



## Kruskal-Wallis H-test

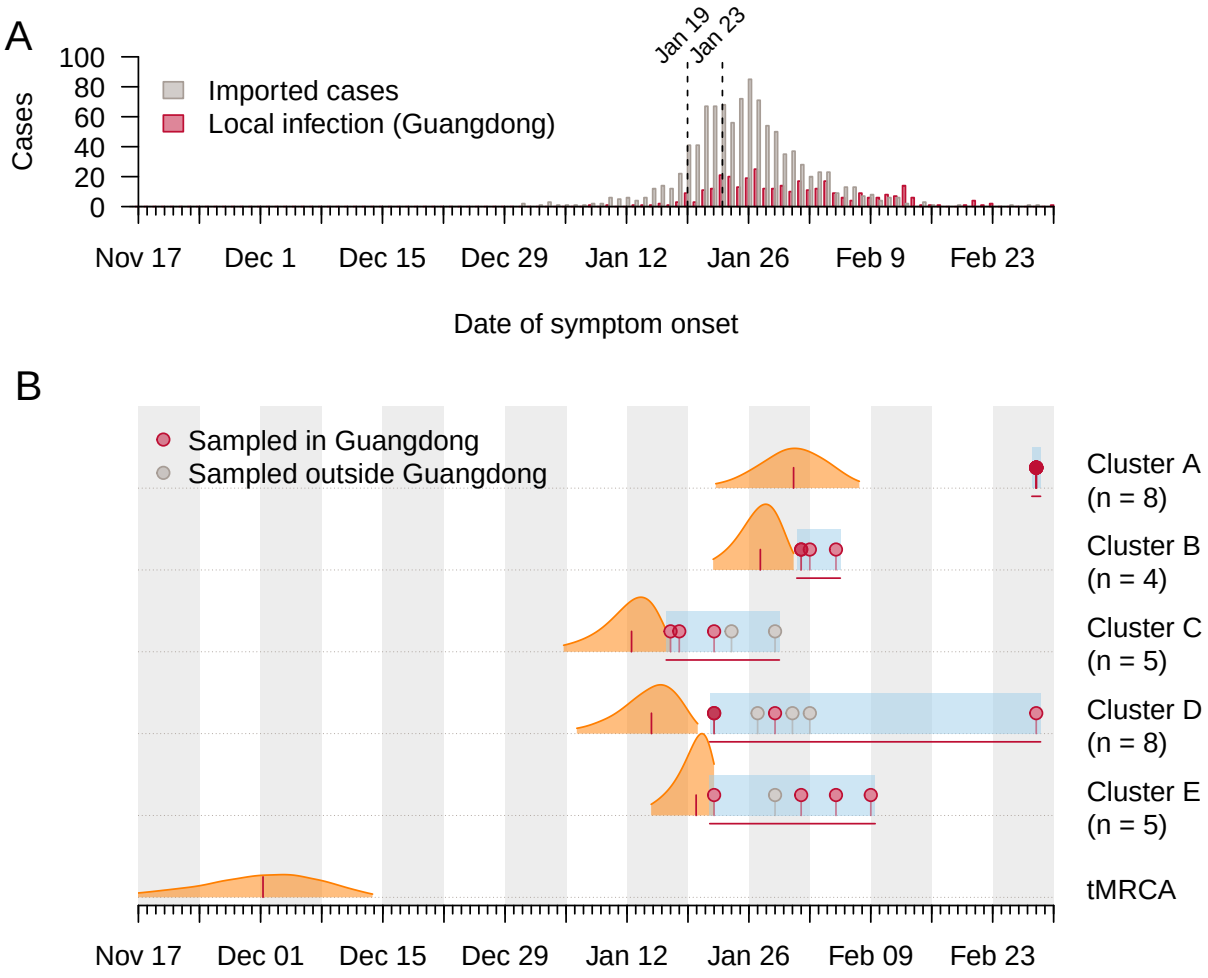
```
##
##  Kruskal-Wallis rank sum test
##
## data:  ct by symptom
## Kruskal-Wallis chi-squared = 8.9434, df = 3, p-value = 0.03005
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  ct by sample_type
## Kruskal-Wallis chi-squared = 18.447, df = 3, p-value = 0.0003557
```

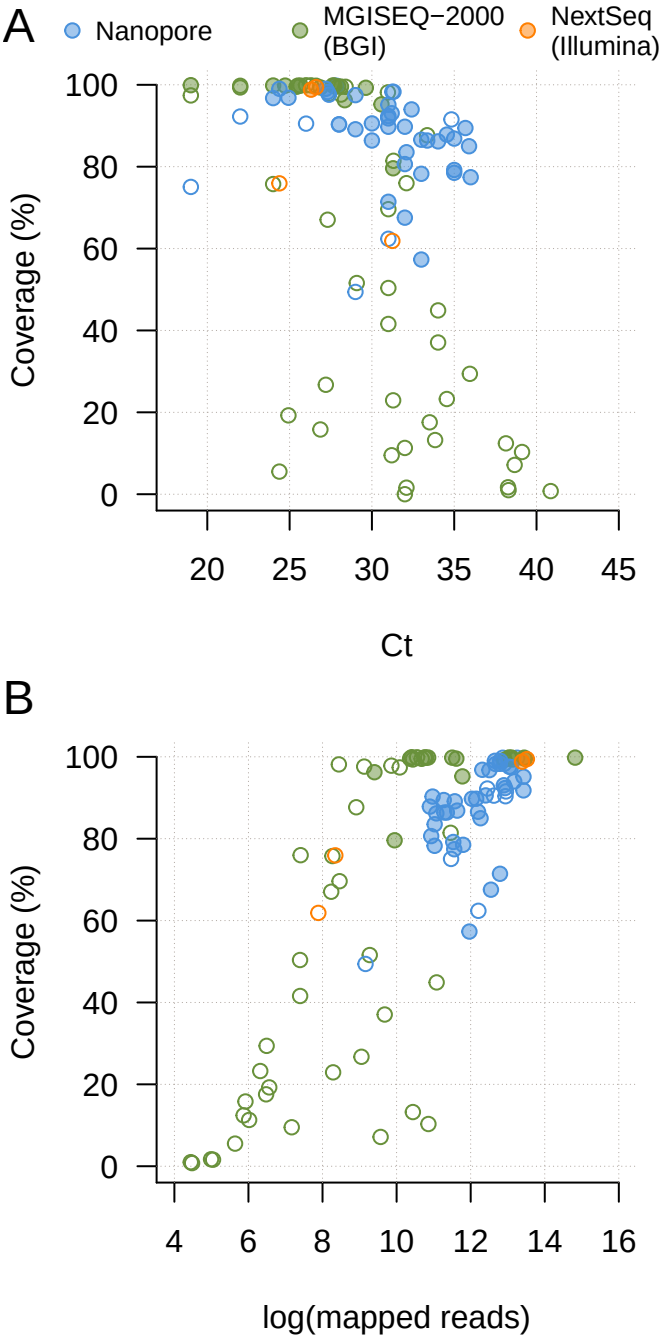
# Fig 3A

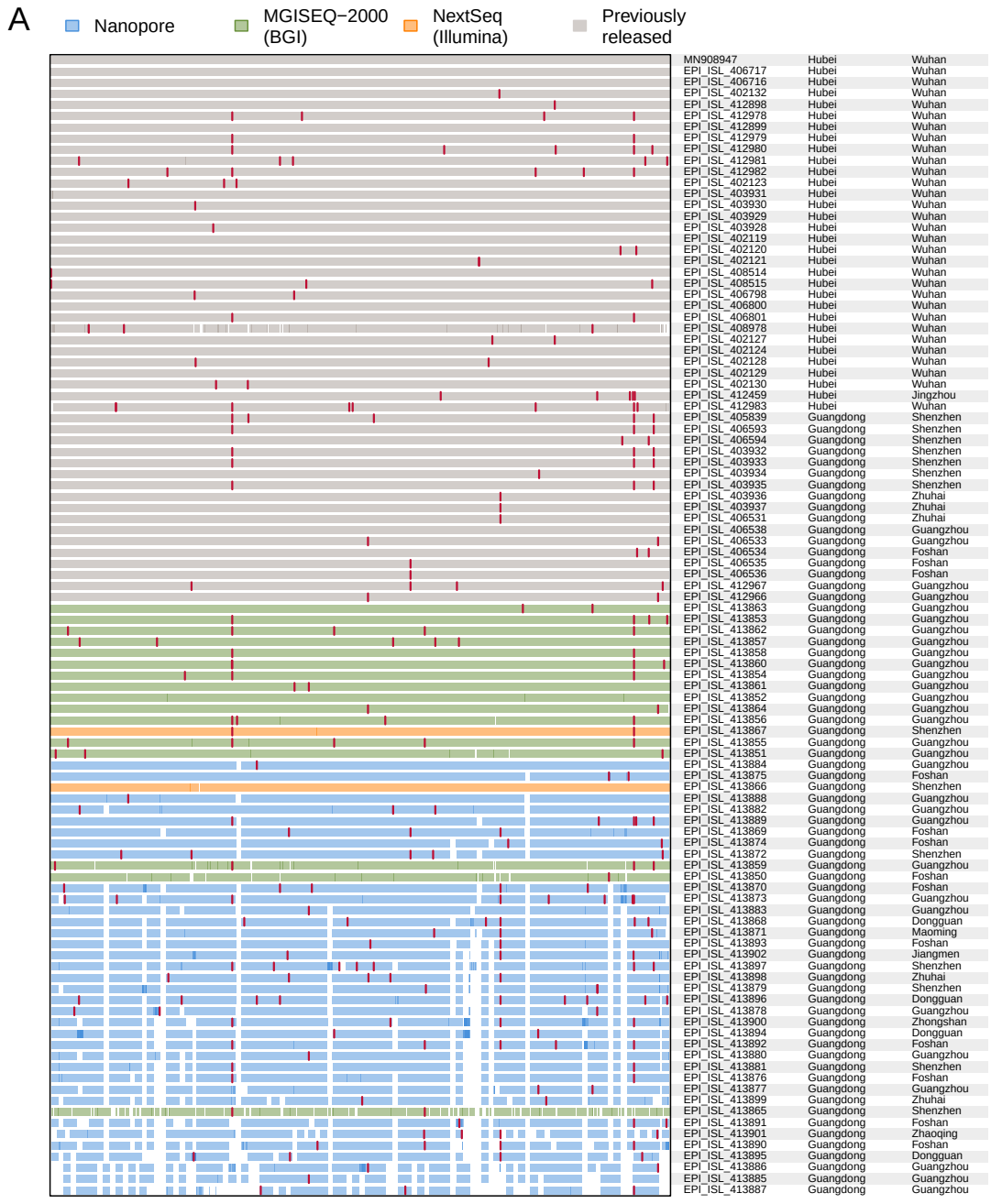


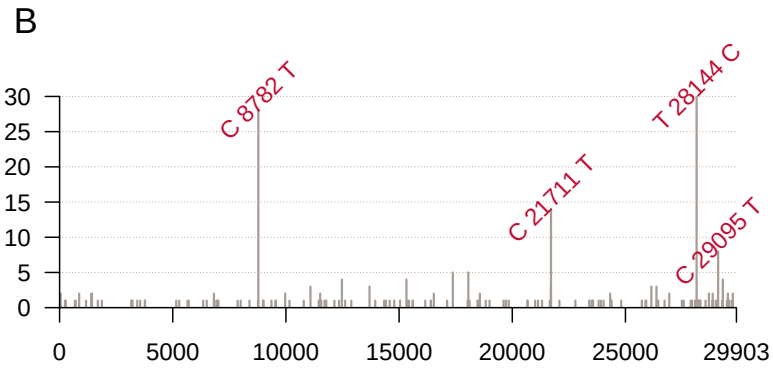
## Figure 4 - TaxonSet Ages

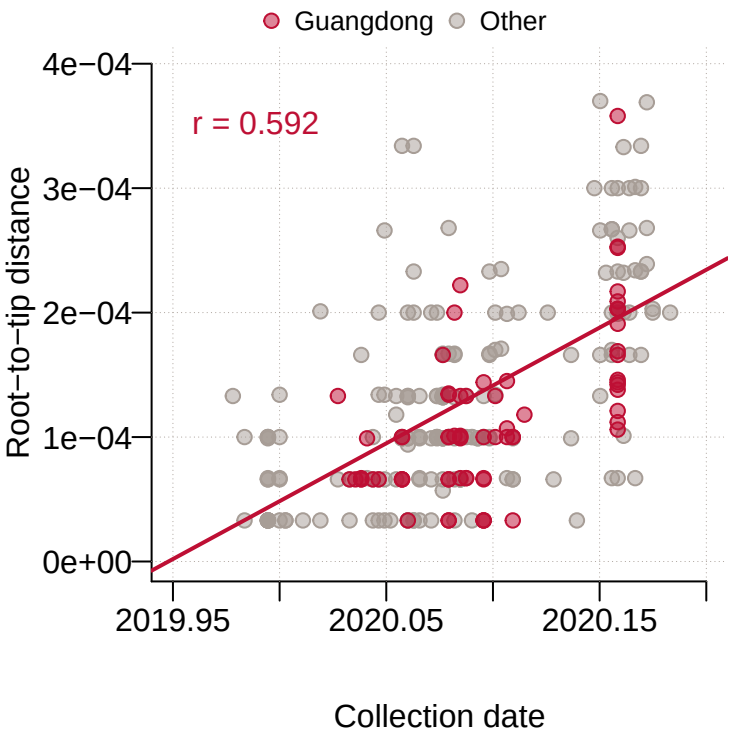


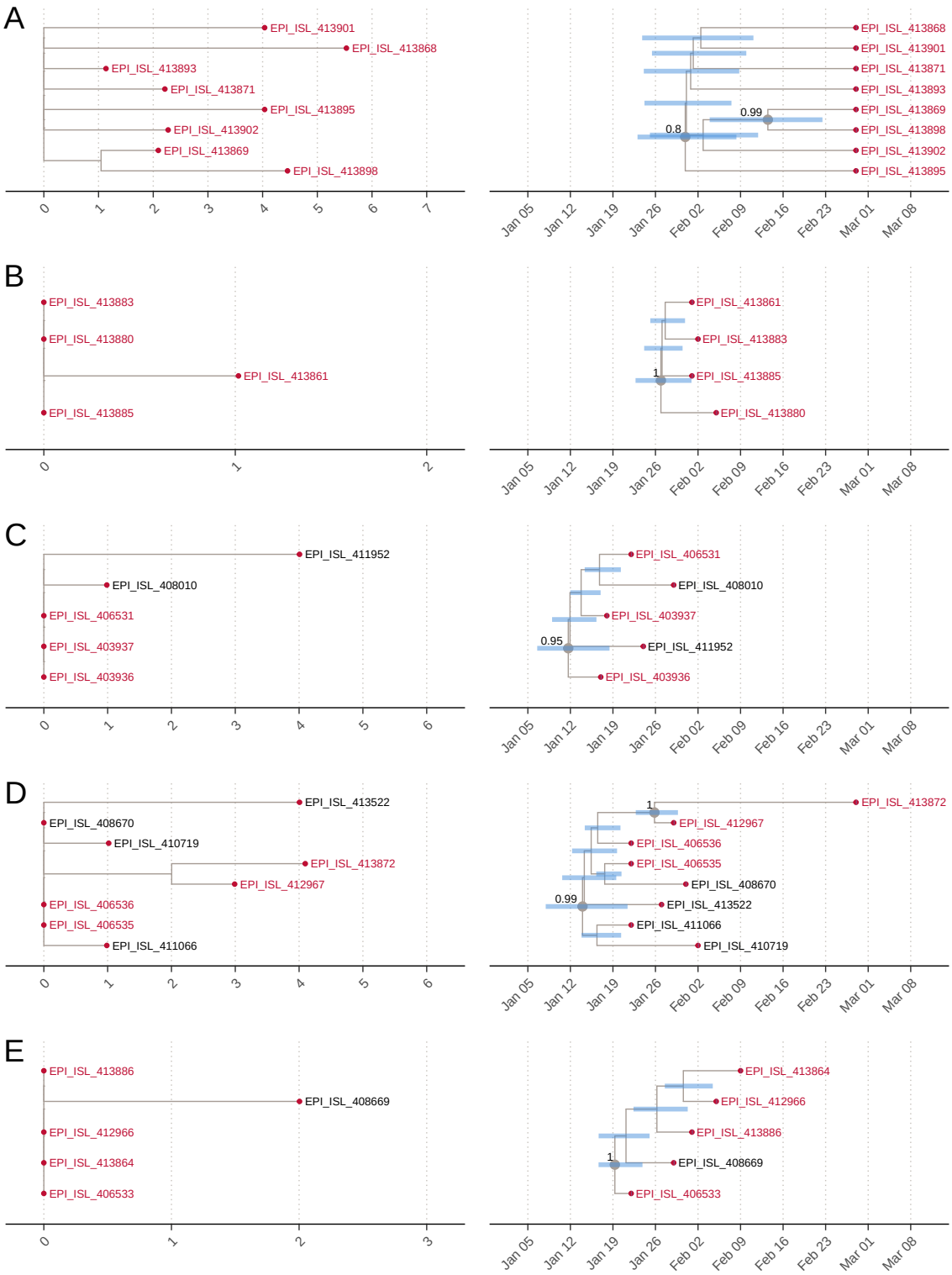
## Supplementary figures (or not included)




A
Nanopore
MGISEQ−2000 (BGI)
NextSeq (Illumina)
Previously released
MN908947 Hubei Wuhan
EPI_ISL_406717 Hubei Wuhan
EPI_ISL_406716 Hubei Wuhan
EPI_ISL_402132 Hubei Wuhan
EPI_ISL_412898 Hubei Wuhan
EPI_ISL_412978 Hubei Wuhan
EPI_ISL_412899 Hubei Wuhan
EPI_ISL_412979 Hubei Wuhan
EPI_ISL_412980 Hubei Wuhan
EPI_ISL_412981 Hubei Wuhan
EPI_ISL_412982 Hubei Wuhan
EPI_ISL_402123 Hubei Wuhan
EPI_ISL_403931 Hubei Wuhan
EPI_ISL_403930 Hubei Wuhan
EPI_ISL_403929 Hubei Wuhan
EPI_ISL_403928 Hubei Wuhan
EPI_ISL_402119 Hubei Wuhan
EPI_ISL_402120 Hubei Wuhan
EPI_ISL_402121 Hubei Wuhan
EPI_ISL_408514 Hubei Wuhan
EPI_ISL_408515 Hubei Wuhan
EPI_ISL_406798 Hubei Wuhan
EPI_ISL_406800 Hubei Wuhan
EPI_ISL_406801 Hubei Wuhan
EPI_ISL_408978 Hubei Wuhan
EPI_ISL_402127 Hubei Wuhan
EPI_ISL_402124 Hubei Wuhan
EPI_ISL_402128 Hubei Wuhan
EPI_ISL_402129 Hubei Wuhan
EPI_ISL_402130 Hubei Wuhan
EPI_ISL_412459 Hubei Jingzhou
EPI_ISL_412983 Hubei Wuhan
EPI_ISL_405839 Guangdong Shenzhen
EPI_ISL_406593 Guangdong Shenzhen
EPI_ISL_406594 Guangdong Shenzhen
EPI_ISL_403932 Guangdong Shenzhen
EPI_ISL_403933 Guangdong Shenzhen
EPI_ISL_403934 Guangdong Shenzhen
EPI_ISL_403935 Guangdong Shenzhen
EPI_ISL_403936 Guangdong Zhuhai
EPI_ISL_403937 Guangdong Zhuhai
EPI_ISL_406531 Guangdong Zhuhai
EPI_ISL_406538 Guangdong Guangzhou
EPI_ISL_406533 Guangdong Guangzhou
EPI_ISL_406534 Guangdong Foshan
EPI_ISL_406535 Guangdong Foshan
EPI_ISL_406536 Guangdong Foshan
EPI_ISL_412967 Guangdong Guangzhou
EPI_ISL_412966 Guangdong Guangzhou
EPI_ISL_413863 Guangdong Guangzhou
EPI_ISL_413853 Guangdong Guangzhou
EPI_ISL_413862 Guangdong Guangzhou
EPI_ISL_413857 Guangdong Guangzhou
EPI_ISL_413858 Guangdong Guangzhou
EPI_ISL_413860 Guangdong Guangzhou
EPI_ISL_413854 Guangdong Guangzhou
EPI_ISL_413861 Guangdong Guangzhou
EPI_ISL_413852 Guangdong Guangzhou
EPI_ISL_413864 Guangdong Guangzhou
EPI_ISL_413856 Guangdong Guangzhou
EPI_ISL_413867 Guangdong Shenzhen
EPI_ISL_413855 Guangdong Guangzhou
EPI_ISL_413851 Guangdong Guangzhou
EPI_ISL_413884 Guangdong Guangzhou
EPI_ISL_413875 Guangdong Foshan
EPI_ISL_413866 Guangdong Shenzhen
EPI_ISL_413888 Guangdong Guangzhou
EPI_ISL_413882 Guangdong Guangzhou
EPI_ISL_413889 Guangdong Guangzhou
EPI_ISL_413869 Guangdong Foshan
EPI_ISL_413874 Guangdong Foshan
EPI_ISL_413872 Guangdong Shenzhen
EPI_ISL_413859 Guangdong Guangzhou
EPI_ISL_413850 Guangdong Foshan
EPI_ISL_413870 Guangdong Foshan
EPI_ISL_413873 Guangdong Guangzhou
EPI_ISL_413883 Guangdong Guangzhou
EPI_ISL_413868 Guangdong Dongguan
EPI_ISL_413871 Guangdong Maoming
EPI_ISL_413893 Guangdong Foshan
EPI_ISL_413902 Guangdong Jiangmen
EPI_ISL_413897 Guangdong Shenzhen
EPI_ISL_413898 Guangdong Zhuhai
EPI_ISL_413879 Guangdong Shenzhen
EPI_ISL_413896 Guangdong Dongguan
EPI_ISL_413878 Guangdong Guangzhou
EPI_ISL_413900 Guangdong Zhongshan
EPI_ISL_413894 Guangdong Dongguan
EPI_ISL_413892 Guangdong Foshan
EPI_ISL_413880 Guangdong Guangzhou
EPI_ISL_413881 Guangdong Shenzhen
EPI_ISL_413876 Guangdong Foshan
EPI_ISL_413877 Guangdong Guangzhou
EPI_ISL_413899 Guangdong Zhuhai
EPI_ISL_413865 Guangdong Shenzhen
EPI_ISL_413891 Guangdong Foshan
EPI_ISL_413901 Guangdong Zhaoqing
EPI_ISL_413890 Guangdong Foshan
EPI_ISL_413895 Guangdong Dongguan
EPI_ISL_413886 Guangdong Guangzhou
EPI_ISL_413885 Guangdong Guangzhou
EPI_ISL_413887 Guangdong Guangzhou


B
C 8782 T
T 28144 C
C 21711 T
C 29095 T
30
25
20
15
10
5
0
0
5000
10000
15000
20000
25000
29903


Guangdong
Other
r = 0.592
Root−to−tip distance
4e−04
3e−04
2e−04
1e−04
0e+00
2019.95
2020.05
2020.15
Collection date


A
EPI_ISL_413901
EPI_ISL_413868
EPI_ISL_413893
EPI_ISL_413871
EPI_ISL_413895
EPI_ISL_413902
EPI_ISL_413869
EPI_ISL_413898
0 1 2 3 4 5 6 7
EPI_ISL_413868
EPI_ISL_413901
EPI_ISL_413871
EPI_ISL_413893
EPI_ISL_413869
EPI_ISL_413898
EPI_ISL_413902
EPI_ISL_413895
0.99
0.8
Jan 05 Jan 12 Jan 19 Jan 26 Feb 02 Feb 09 Feb 16 Feb 23 Mar 01 Mar 08
B
EPI_ISL_413883
EPI_ISL_413880
EPI_ISL_413861
EPI_ISL_413885
0 1 2
EPI_ISL_413861
EPI_ISL_413883
EPI_ISL_413885
EPI_ISL_413880
1
Jan 05 Jan 12 Jan 19 Jan 26 Feb 02 Feb 09 Feb 16 Feb 23 Mar 01 Mar 08
C
EPI_ISL_411952
EPI_ISL_408010
EPI_ISL_406531
EPI_ISL_403937
EPI_ISL_403936
0 1 2 3 4 5 6
EPI_ISL_406531
EPI_ISL_408010
EPI_ISL_403937
EPI_ISL_411952
EPI_ISL_403936
0.95
Jan 05 Jan 12 Jan 19 Jan 26 Feb 02 Feb 09 Feb 16 Feb 23 Mar 01 Mar 08
D
EPI_ISL_413522
EPI_ISL_408670
EPI_ISL_410719
EPI_ISL_413872
EPI_ISL_412967
EPI_ISL_406536
EPI_ISL_406535
EPI_ISL_411066
0 1 2 3 4 5 6
EPI_ISL_413872
1
EPI_ISL_412967
EPI_ISL_406536
EPI_ISL_406535
EPI_ISL_408670
EPI_ISL_413522
EPI_ISL_411066
EPI_ISL_410719
0.99
Jan 05 Jan 12 Jan 19 Jan 26 Feb 02 Feb 09 Feb 16 Feb 23 Mar 01 Mar 08
E
EPI_ISL_413886
EPI_ISL_408669
EPI_ISL_412966
EPI_ISL_413864
EPI_ISL_406533
0 1 2 3
EPI_ISL_413864
EPI_ISL_412966
EPI_ISL_413886
EPI_ISL_408669
EPI_ISL_406533
1
Jan 05 Jan 12 Jan 19 Jan 26 Feb 02 Feb 09 Feb 16 Feb 23 Mar 01 Mar 08

## Session info

```
## R version 3.5.1 (2018-07-02)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS Sierra 10.12.6
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_GB.UTF-8/en_GB.UTF-8/en_GB.UTF-8/C/en_GB.UTF-8/en_GB.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] lubridate_1.7.4 phytools_0.6-99 maps_3.3.0      ape_5.3
##  [5] ggtree_2.1.1    ggplot2_3.2.1   treeio_1.11.2   beastio_0.2.5
##  [9] coda_0.19-3     gplots_3.0.1.1  cowplot_1.0.0
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_1.0.3              lattice_0.20-38         tidyr_1.0.0
##  [4] gtools_3.8.1            assertthat_0.2.1        zeallot_0.1.0
##  [7] digest_0.6.23           R6_2.4.1                backports_1.1.5
## [10] evaluate_0.14           pillar_1.4.2            rlang_0.4.2
## [13] lazyeval_0.2.2          gdata_2.18.0            phangorn_2.5.5
## [16] combinat_0.0-8          Matrix_1.2-18           rmarkdown_2.0
## [19] labeling_0.3            stringr_1.4.0           igraph_1.2.4.2
## [22] munsell_0.5.0           compiler_3.5.1          numDeriv_2016.8-1.1
## [25] xfun_0.11               pkgconfig_2.0.3         mnormt_1.5-5
## [28] htmltools_0.4.0         tidyselect_0.2.5        expm_0.999-4
## [31] tibble_2.1.3            quadprog_1.5-8          crayon_1.3.4
## [34] dplyr_0.8.3             withr_2.1.2             MASS_7.3-51.4
## [37] bitops_1.0-6            grid_3.5.1              nlme_3.1-143
## [40] jsonlite_1.6            gtable_0.3.0            lifecycle_0.1.0
## [43] magrittr_1.5            scales_1.1.0            KernSmooth_2.23-16
## [46] tidytree_0.3.1          stringi_1.4.3           farver_2.0.1
## [49] scatterplot3d_0.3-41    rvcheck_0.1.7           vctrs_0.2.1
## [52] fastmatch_1.1-0         RColorBrewer_1.1-2      tools_3.5.1
## [55] glue_1.3.1              purrr_0.3.3             plotrix_3.7-7
## [58] parallel_3.5.1          yaml_2.2.0              colorspace_1.4-1
## [61] BiocManager_1.30.10     caTools_1.17.1.3        animation_2.6
## [64] clusterGeneration_1.3.4 knitr_1.26
```