

SARS-CoV-2 State introductions

Transmission Lineage Summary for Hamburg

Last modified: 27 Jun 2021

Contents

1	Summary	1
2	Transmission lineage statistics	1
2.1	Input data	1
3	Transmission lineage TMRCA distribution	3
4	Size and duration vs. TMRCA	4
5	Time since lineage last sampled	6
6	Session info	7

1 Summary

This notebook plots summary statistics and figures of the German transmission lineages extracted from the GISAID tree and Sankoff algorithm analyses.

2 Transmission lineage statistics

2.1 Input data

- Transmission lineages and singletons, on the dataset, $n = 178672$ sequences ($n = \text{Germany} : 111283$, Germany : 62% from the Germany‘).

```
## Warning in if (class(try(col2rgb(palette), silent = TRUE)) == "try-error")
## stop("color palette is not correct"): the condition has length > 1 and only the
## first element will be used
```

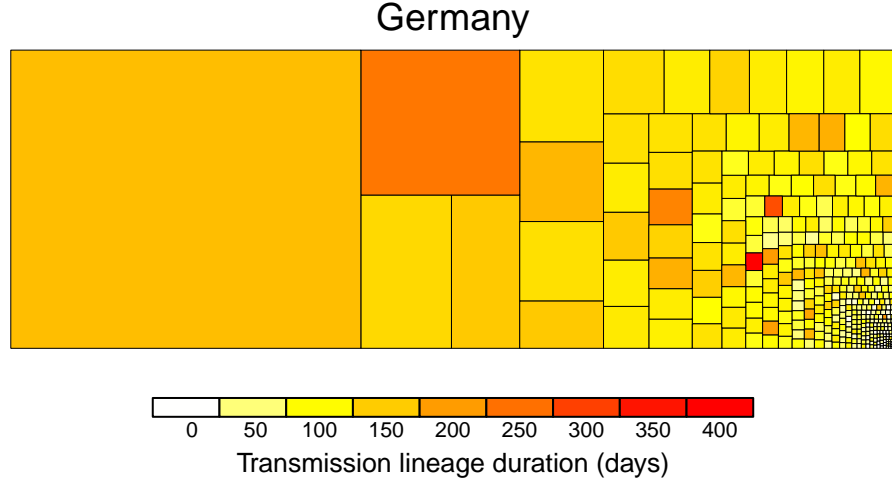


Figure 1: Partition of `r nrow(metadata)` `r states` genomes into `r states` transmission lineages and singletons, coloured by duration of lineage detection (time between the lineage's oldest and most recent genomes). (Transmission lineages from the MCC trees).

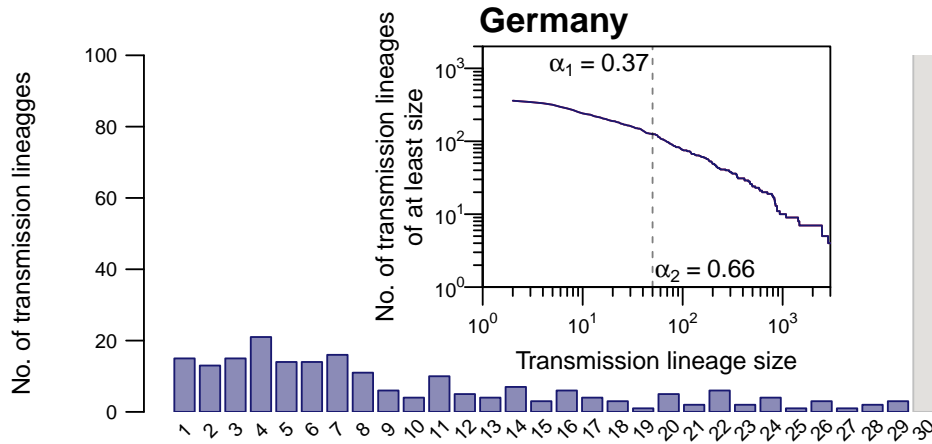


Figure 2: Distribution of `r states` transmission lineage sizes (MCC trees). Blue bars show the number of transmission lineages of each size (red bars=95% HPD of these sizes across the posterior tree distribution). Inset: the corresponding complementary cumulative frequency distribution of lineage size (blue line), on double logarithmic axes (red shading=95% HPD of this distribution across the posterior tree distribution). Values either side of vertical dashed line show coefficients of power-law distributions ($P[X \geq x] \sim x^{\alpha}$) fitted to lineages containing ≤ 50 (α_1) and > 50 (α_2) virus genomes, respectively.

3 Transmission lineage TMRCAs distribution

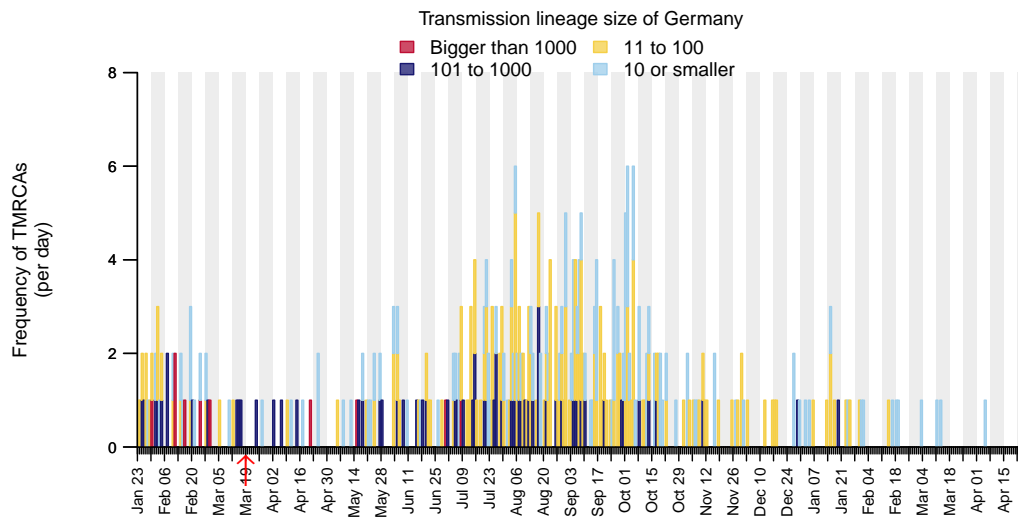


Figure 3: Histogram of lineage TMRCAs, coloured by lineage size. Inset: expanded view of the days prior to Germany lockdown. Left-hand arrow = collection date of the Germany's first laboratory-confirmed case; right-hand arrow = collection date of the earliest Germany virus genome in our dataset.

4 Size and duration vs. TMRCA

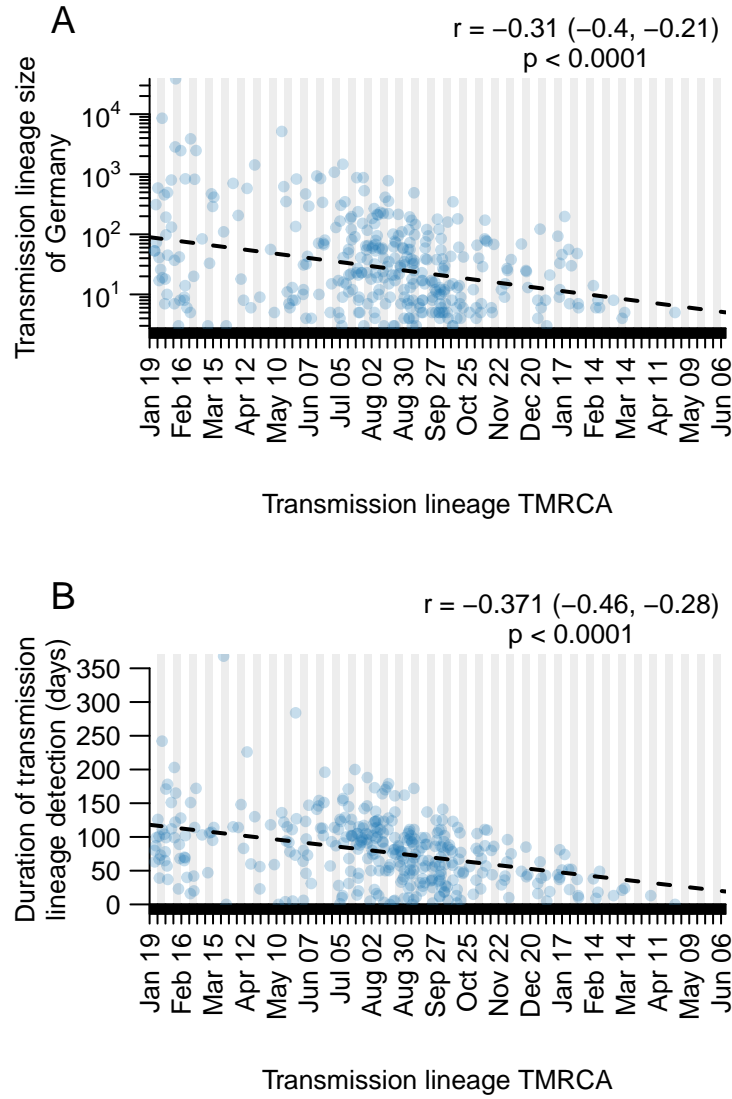


Figure 4: Scatterplots showing the relationship between (A) **r states** transmission lineage size and lineage TMRCA and between (B) **r states** transmission lineage sampling duration and lineage TMRCA. Pearson correlation coefficients, 95% CIs and p-values are shown in the top-right corners.

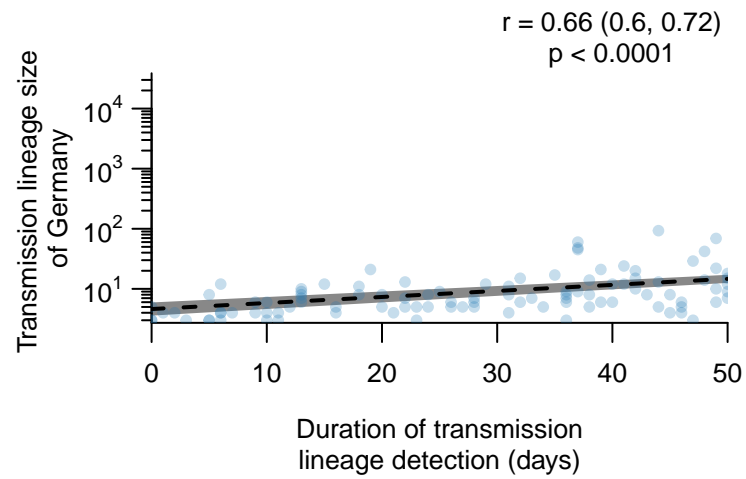


Figure 5: Scatterplot showing the strong relationship between **r states** transmission lineage size and sampling duration. The Pearson correlation coefficient, 95% CI and p-value are shown.

5 Time since lineage last sampled

2.013 sec elapsed

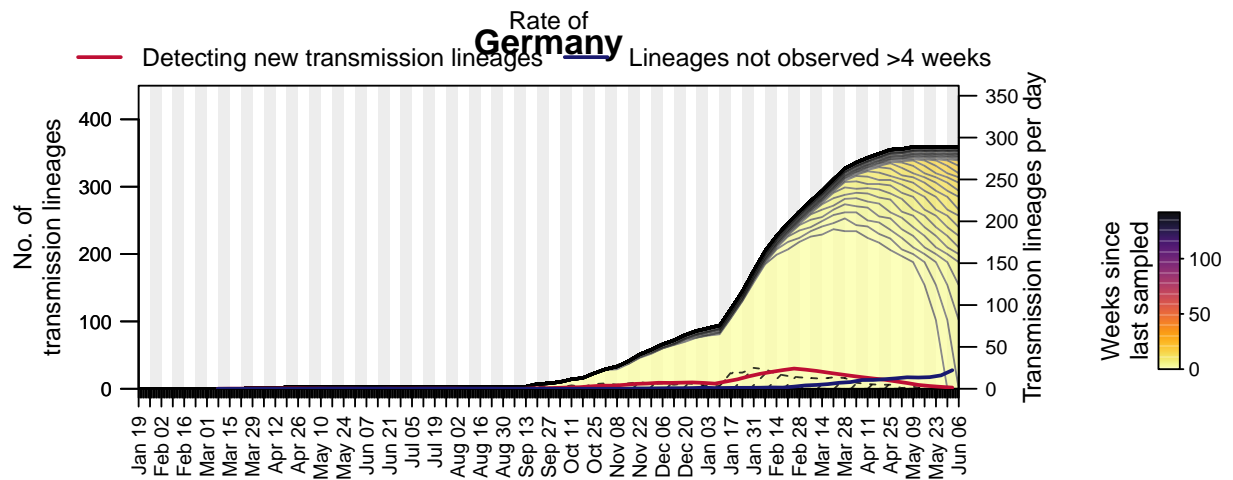


Figure 6: Trends through time in the detection of r states transmission lineages (absolute values). For each day, all lineages detected up to that day are coloured by the time since the transmission lineage was last sampled. Isoclines correspond to weeks. Shaded area=transmission lineages that were first sampled <1 week ago.

6 Session info

```
## R version 4.1.0 (2021-05-18)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 20.04.2 LTS
##
## Matrix products: default
## BLAS: /usr/lib/x86_64-linux-gnu/atlas/libblas.so.3.10.3
## LAPACK: /usr/lib/x86_64-linux-gnu/atlas/liblapack.so.3.10.3
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C
##  [3] LC_TIME=en_US.UTF-8      LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=en_US.UTF-8  LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8     LC_NAME=C
##  [9] LC_ADDRESS=C             LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods    base
##
## other attached packages:
## [1] tictoc_1.0.1      beastio_0.3.3      viridis_0.6.1      viridisLite_0.4.0
## [5] coda_0.19-4       treemap_2.4-2      gplots_3.1.1       lubridate_1.7.10
## [9] stringr_1.4.0
##
## loaded via a namespace (and not attached):
##  [1] gtools_3.8.2      tidyselect_1.1.1   xfun_0.23          purrr_0.3.4
##  [5] lattice_0.20-44   colorspace_2.0-1   generics_0.1.0     vctrs_0.3.8
##  [9] htmltools_0.5.1.1 yaml_2.2.1         utf8_1.2.1         rlang_0.4.11
## [13] later_1.2.0       pillar_1.6.1       glue_1.4.2         RColorBrewer_1.1-2
## [17] lifecycle_1.0.0   munsell_0.5.0      gtable_0.3.0       caTools_1.18.2
## [21] codetools_0.2-18  evaluate_0.14      knitr_1.33         fastmap_1.1.0
## [25] httpuv_1.6.1      fansi_0.5.0        highr_0.9          Rcpp_1.0.6
## [29] KernSmooth_2.23-20 xtable_1.8-4       promises_1.2.0.1   scales_1.1.1
## [33] mime_0.10         gridExtra_2.3      ggplot2_3.3.3      digest_0.6.27
## [37] stringi_1.6.2     dplyr_1.0.6        shiny_1.6.0        grid_4.1.0
## [41] tools_4.1.0       bitops_1.0-7       magrittr_2.0.1     tibble_3.1.2
## [45] crayon_1.4.1      pkgconfig_2.0.3    ellipsis_0.3.2     gridBase_0.4-7
## [49] data.table_1.14.0 rmarkdown_2.8      R6_2.5.0           igraph_1.2.6
## [53] compiler_4.1.0
```