

# SARS-CoV-2 State Introductions

## Importation Summary for states

Last modified: 27 Jun 2021

## Contents

<b>1</b>	<b>Summary</b>	<b>1</b>
1.1	Data and Method . . . . .	1
<b>2</b>	<b>Germany Sequenced Genomes</b>	<b>2</b>
<b>3</b>	<b>Lineage importation distribution (shifted TMRCA distribution)</b>	<b>3</b>
<b>4</b>	<b>Session info</b>	<b>11</b>

## 1 Summary

This notebook plots figures about the dataset, applies the importation lag model to the Germany transmission lineage TMRCA and plots figures with lineage importations.

### 1.1 Data and Method

- GISAID tree until ??? as initial tree.
- The tree contains ??? Germany sequences.
- The tree is time-calibrated by TreeTime.
- Sankoff algorithm is used to assign location (Germany and non-Germany) to inner vertices of the tree.

## 2 Germany Sequenced Genomes

## #BE0F34 #B0B0B0

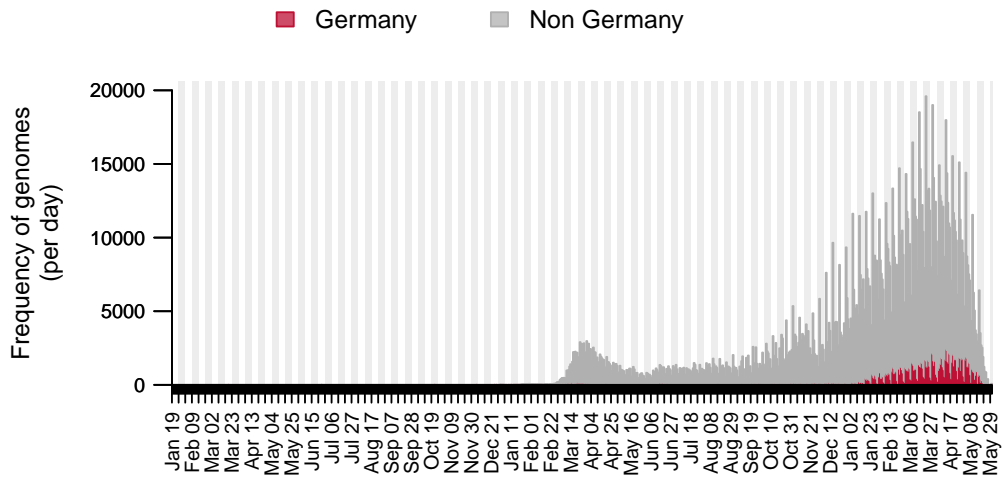


Figure 1: Collection dates of the `nrow(metadata)` genomes analysed here (left-hand axis). Genomes are coloured by sampling location.

## #BE0F34

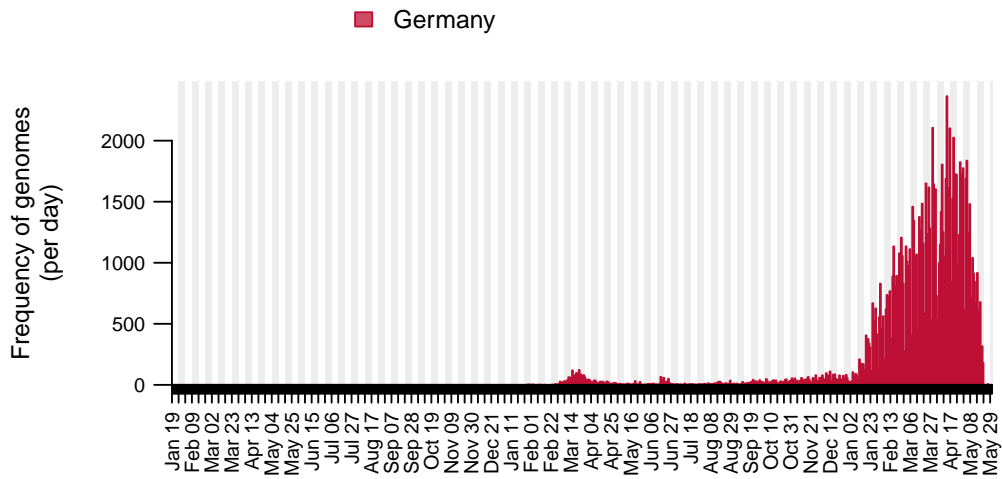


Figure 2: Collection dates of the `nrow(metadata)` genomes analysed here (left-hand axis). Genomes are coloured by sampling location.

### 3 Lineage importation distribution (shifted TMRCA distribution)

- GISAID tree until 2021-06-24 as initial tree.
- The tree is time-collibrated by TreeTime.
- Sankoff algorithm is used to assign location (Germany and non-Germany) to inner vertices of the tree, for each state separately.
- Dataset contains Germany : 248 transmission lineages (2 or more sequences), comprising Germany : 32802 sequences from the Germany, as well as a further Germany : 88005 singletons.
- Mean and SD of the importation (shifted TMRCA) distribution: Germany : NA  $\pm$  NA days (singletons excluded).
- Median and interquartile range of the importation (shifted TMRCA) distribution Germany : NA [Germany : 2020-09-21, Germany : 2021-01-03] (singletons excluded).
- 80% of importations fall in [Germany : 2020-08-15, Germany : 2021-02-11].

Table 1: Estimated importation lags for multistate transmission lineages of different sizes. Importation lag is the waiting time between importation date and the TMRCA of the sampled genomes in the transmission lineage. Detection lag is the waiting time from the importation date to the sampling time of the oldest (first) sampled genome in the transmission lineage.

Lineages of size	No. of lineages of Germany	Importation lag (mean $\pm$ SD) Germany	Importation lag (median and IQR) Germany	Detection lag (mean $\pm$ SD) Germany	Detection lag (median and IQR) Germany
All	248	3.75 $\pm$ 4.2	1.79 [0.99-4.33]	NA $\pm$ NA	NA [27-105]
2 to 10	80	8.65 $\pm$ 4.27	7.95 [4.72-10.36]	NA $\pm$ NA	NA [21.25-56.5]
11 to 100	102	1.78 $\pm$ 0.59	1.64 [1.29-2.24]	66.4 $\pm$ 49.18	55 [30-92]
101 to 1000	60	0.86 $\pm$ 0.08	0.85 [0.8-0.94]	96.28 $\pm$ 51.68	93 [60.75-130]
Bigger than 1000	6	0.73 $\pm$ 0	0.74 [0.73-0.74]	159.5 $\pm$ 15.33	160.5 [148-170.75]

Table 2: 3. Estimated importation and detection lags for Germany transmission lineages ordered by importation date and aggregated by epi-week. Importation lag is the waiting time between importation date and the TMRCA of the sampled genomes in the transmission lineage. Detection lag is the waiting time from the importation date to the sampling time of the oldest (first) sampled genome in the transmission lineage. All statistics show means and standard deviations computed from the MCC trees.

Week start-ing	Epi-week	Estimated no. of importations of Germany	Lineage sizes (median and IQR) Germany	Importation lag (mean $\pm$ SD) Germany	Detection lag (mean $\pm$ SD) Germany
19 Jan 19	3	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Jan 26	4	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Feb 02	5	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Feb 09	6	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Feb 16	7	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Feb 23	8	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Mar 02	9	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Mar 09	10	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Mar 16	11	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Mar 23	12	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Mar 30	13	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Apr 06	14	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Apr 13	15	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Apr 20	16	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Apr 27	17	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 May 04	18	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 May 11	19	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 May 18	20	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 May 25	21	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Jun 01	22	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Jun 08	23	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA

Week start- ing	Epi- week	Estimated no. of importations of Germany	Lineage sizes (median and IQR) Germany	Importation lag (mean $\pm$ SD) Germany	Detection lag (mean $\pm$ SD) Germany
19 Jun 15	24	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Jun 22	25	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Jun 29	26	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Jul 06	27	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Jul 13	28	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Jul 20	29	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Jul 27	30	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Aug 03	31	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Aug 10	32	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Aug 17	33	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Aug 24	34	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Aug 31	35	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Sep 07	36	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Sep 14	37	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Sep 21	38	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Sep 28	39	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Oct 05	40	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Oct 12	41	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Oct 19	42	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Oct 26	43	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Nov 02	44	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Nov 09	45	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Nov 16	46	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Nov 23	47	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Nov 30	48	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA

Week start- ing	Epi- week	Estimated no. of importations of Germany	Lineage sizes (median and IQR) Germany	Importation lag (mean $\pm$ SD) Germany	Detection lag (mean $\pm$ SD) Germany
19 Dec 07	49	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Dec 14	50	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Dec 21	51	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
19 Dec 28	0	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Jan 04	1	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Jan 11	2	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Jan 18	3	3	NA [144-144]	NA $\pm$ NA	NA $\pm$ NA
20 Jan 25	4	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Feb 01	5	3	NA [229-229]	NA $\pm$ NA	NA $\pm$ NA
20 Feb 08	6	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Feb 15	7	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Feb 22	8	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Feb 29	9	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Mar 07	10	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Mar 14	11	3	NA [3-3]	NA $\pm$ NA	NA $\pm$ NA
20 Mar 21	12	3	NA [8-8]	NA $\pm$ NA	NA $\pm$ NA
20 Mar 28	13	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Apr 04	14	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Apr 11	15	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Apr 18	16	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Apr 25	17	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 May 02	18	3	NA [7-7]	NA $\pm$ NA	NA $\pm$ NA
20 May 09	19	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 May 16	20	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 May 23	21	3	NA [10-10]	NA $\pm$ NA	NA $\pm$ NA

Week start- ing	Epi- week	Estimated no. of importations of Germany	Lineage sizes (median and IQR) Germany	Importation lag (mean $\pm$ SD) Germany	Detection lag (mean $\pm$ SD) Germany
20 May 30	22	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Jun 06	23	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Jun 13	24	3	NA [153-153]	NA $\pm$ NA	NA $\pm$ NA
20 Jun 20	25	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Jun 27	26	3	NA [7-7]	NA $\pm$ NA	NA $\pm$ NA
20 Jul 04	27	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
20 Jul 11	28	4	NA [178-290]	NA $\pm$ NA	NA $\pm$ NA
20 Jul 18	29	4	NA [2.5-3.5]	NA $\pm$ NA	NA $\pm$ NA
20 Jul 25	30	6	NA [81.5-198]	NA $\pm$ NA	NA $\pm$ NA
20 Aug 01	31	8	NA [4-209.25]	NA $\pm$ NA	NA $\pm$ NA
20 Aug 08	32	4	NA [728.25-2012.75]	NA $\pm$ NA	NA $\pm$ NA
20 Aug 15	33	7	NA [2-378]	NA $\pm$ NA	NA $\pm$ NA
20 Aug 22	34	11	NA [36-363]	NA $\pm$ NA	NA $\pm$ NA
20 Aug 29	35	8	NA [20.25-199]	NA $\pm$ NA	NA $\pm$ NA
20 Sep 05	36	9	NA [92-968.5]	NA $\pm$ NA	NA $\pm$ NA
20 Sep 12	37	11	NA [11-45]	NA $\pm$ NA	NA $\pm$ NA
20 Sep 19	38	6	NA [2.75-18]	NA $\pm$ NA	NA $\pm$ NA
20 Sep 26	39	14	NA [12.25-41.75]	NA $\pm$ NA	NA $\pm$ NA
20 Oct 03	40	11	NA [7-216]	NA $\pm$ NA	NA $\pm$ NA
20 Oct 10	41	14	NA [8.25-144.25]	NA $\pm$ NA	NA $\pm$ NA
20 Oct 17	42	13	NA [21.5-448.5]	NA $\pm$ NA	NA $\pm$ NA
20 Oct 24	43	8	NA [20.25-163.5]	NA $\pm$ NA	NA $\pm$ NA
20 Oct 31	44	10	NA [16-81.25]	NA $\pm$ NA	NA $\pm$ NA
20 Nov 07	45	12	NA [9.5-79.5]	NA $\pm$ NA	NA $\pm$ NA
20 Nov 14	46	10	NA [5-127]	NA $\pm$ NA	NA $\pm$ NA

Week start- ing	Epi- week	Estimated no. of importations of Germany	Lineage sizes (median and IQR) Germany	Importation lag (mean $\pm$ SD) Germany	Detection lag (mean $\pm$ SD) Germany
20 Nov 21	47	11	NA [3-17]	NA $\pm$ NA	NA $\pm$ NA
20 Nov 28	48	10	NA [5.5-239.75]	NA $\pm$ NA	NA $\pm$ NA
20 Dec 05	49	11	NA [14-75]	NA $\pm$ NA	NA $\pm$ NA
20 Dec 12	50	12	NA [26.5-219.75]	NA $\pm$ NA	NA $\pm$ NA
20 Dec 19	51	6	NA [33.75-67]	NA $\pm$ NA	NA $\pm$ NA
20 Dec 26	0	4	NA [9.5-12.5]	NA $\pm$ NA	NA $\pm$ NA
21 Jan 02	1	14	NA [6.25-55.75]	NA $\pm$ NA	NA $\pm$ NA
21 Jan 09	2	7	NA [8-51]	NA $\pm$ NA	NA $\pm$ NA
21 Jan 16	3	11	NA [15-50]	NA $\pm$ NA	NA $\pm$ NA
21 Jan 23	4	7	NA [17-31]	NA $\pm$ NA	NA $\pm$ NA
21 Jan 30	5	7	NA [39-122]	NA $\pm$ NA	NA $\pm$ NA
21 Feb 06	6	5	NA [19.5-32]	NA $\pm$ NA	NA $\pm$ NA
21 Feb 13	7	6	NA [5.25-29]	NA $\pm$ NA	NA $\pm$ NA
21 Feb 20	8	9	NA [10-31.5]	NA $\pm$ NA	NA $\pm$ NA
21 Feb 27	9	4	NA [17.25-47.75]	NA $\pm$ NA	NA $\pm$ NA
21 Mar 06	10	4	NA [11.25-17.75]	NA $\pm$ NA	NA $\pm$ NA
21 Mar 13	11	6	NA [10.5-42]	NA $\pm$ NA	NA $\pm$ NA
21 Mar 20	12	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
21 Mar 27	13	3	NA [9-9]	NA $\pm$ NA	NA $\pm$ NA
21 Apr 03	14	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
21 Apr 10	15	4	NA [9.25-17.75]	NA $\pm$ NA	NA $\pm$ NA
21 Apr 17	16	4	NA [17.5-22.5]	NA $\pm$ NA	NA $\pm$ NA
21 Apr 24	17	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
21 May 01	18	3	NA [2-2]	NA $\pm$ NA	NA $\pm$ NA
21 May 08	19	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA



Week start- ing	Epi- week	Estimated no. of importations of Germany	Lineage sizes (median and IQR) Germany	Importation lag (mean $\pm$ SD) Germany	Detection lag (mean $\pm$ SD) Germany
21 May 15	20	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
21 May 22	21	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
21 May 29	22	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
21 Jun 05	23	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA
21 Jun 12	24	2	NA [NA-NA]	NA $\pm$ NA	NA $\pm$ NA

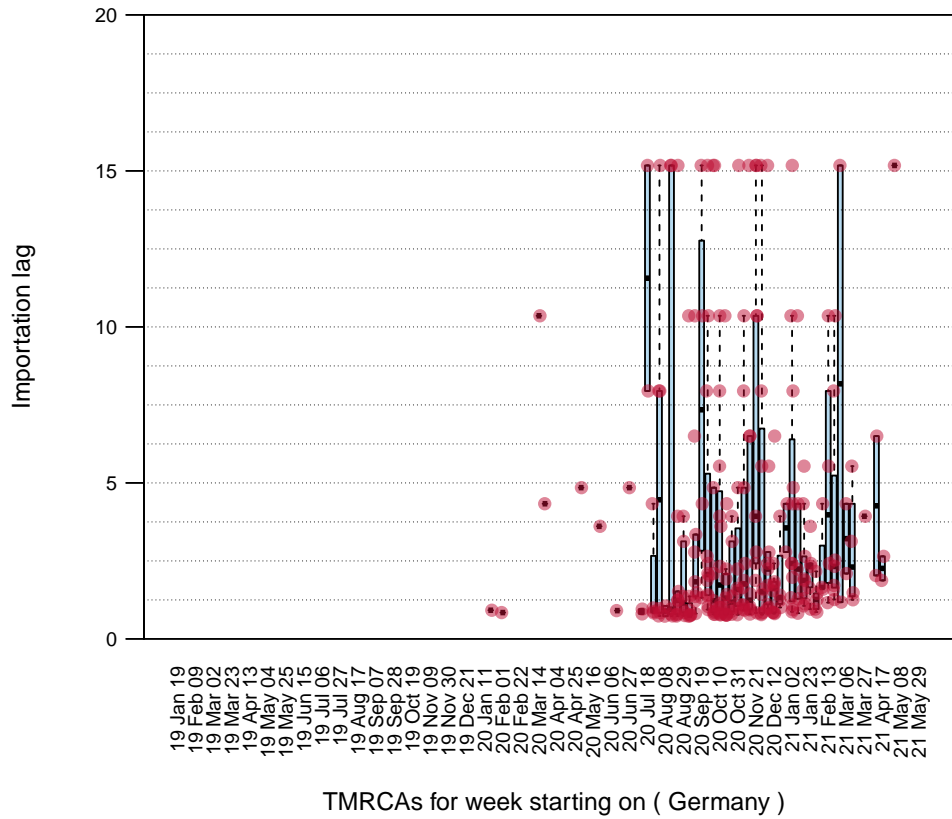


Figure 3: Boxplots of the estimated importation lags for `r states` transmission lineages ordered by importation date and aggregated by epi-week.

```
## Error in seq.int(0, to0 - from, by): 'to' must be a finite number
## Error in seq.int(0, to0 - from, by): 'to' must be a finite number
## Error in seq.int(0, to0 - from, by): 'to' must be a finite number
```

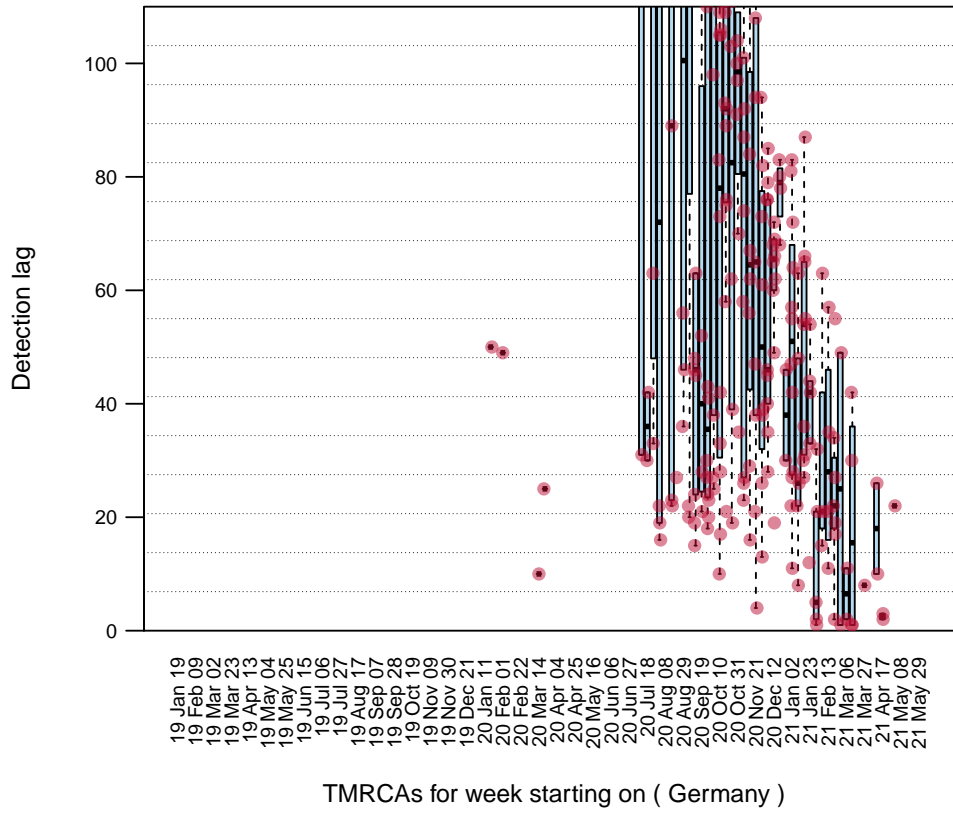


Figure 4: Boxplots of the estimated detection lags for **r states** transmission lineages ordered by importation date and aggregated by epi-week.

## 4 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] phytools_0.7-70  maps_3.3.0      ape_5.5         ggtree_3.0.1
##  [5] ggsci_2.9        ggplot2_3.3.3   stringr_1.4.0   tictoc_1.0.1
##  [9] knitr_1.33       beastio_0.3.3   gplots_3.1.1   plyr_1.8.6
## [13] lubridate_1.7.10
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_1.0.6          lattice_0.20-44    tidyr_1.1.3
##  [4] gtools_3.8.2        digest_0.6.27      utf8_1.2.1
##  [7] R6_2.5.0            evaluate_0.14      coda_0.19-4
## [10] highr_0.9           pillar_1.6.1       rlang_0.4.11
## [13] lazyeval_0.2.2      phangorn_2.7.0     Matrix_1.3-4
## [16] combinat_0.0-8      rmarkdown_2.8      igraph_1.2.6
## [19] munsell_0.5.0       numDeriv_2016.8-1.1 compiler_4.1.0
## [22] xfun_0.23           pkgconfig_2.0.3    mnormt_2.0.2
## [25] tmvnsim_1.0-2       htmltools_0.5.1.1  tidyselect_1.1.1
## [28] tibble_3.1.2        expm_0.999-6       codetools_0.2-18
## [31] quadprog_1.5-8      fansi_0.5.0        crayon_1.4.1
## [34] dplyr_1.0.6         withr_2.4.2        MASS_7.3-54
## [37] bitops_1.0-7        grid_4.1.0         nlme_3.1-152
## [40] jsonlite_1.7.2      gtable_0.3.0       lifecycle_1.0.0
## [43] magrittr_2.0.1      scales_1.1.1       KernSmooth_2.23-20
## [46] tidytree_0.3.4      stringi_1.6.2      scatterplot3d_0.3-41
## [49] ellipsis_0.3.2      rvcheck_0.1.8      generics_0.1.0
## [52] vctr_0.3.8          fastmatch_1.1-0    RColorBrewer_1.1-2
## [55] tools_4.1.0         treeio_1.16.1      glue_1.4.2
## [58] purrr_0.3.4         plotrix_3.8-1      parallel_4.1.0
## [61] yaml_2.2.1          colorspace_2.0-1   BiocManager_1.30.15
## [64] caTools_1.18.2      aplot_0.0.6        clusterGeneration_1.3.7
## [67] patchwork_1.1.1
```