

SARS-CoV-19 State introductions

Transmission Lineage Breakdown for Hamburg

Last modified: 28 Aug 2021

Contents

1	Summary	1
1.1	Data and Method	1
2	Sample breakdown (daily)	2
3	Sample breakdown into lineages (weekly)	4
4	Individual transmission lineage plots	13
4.1	Biggest transmission lineages	14
4.2	Earliest transmission lineages	18
4.3	Newest transmission lineages	20
4.4	Longest unobserved period before reactivating	24
5	Session info	26

1 Summary

This notebook plots breakdowns of Germany transmission lineages over time (using only the assignment on the MCC trees).

1.1 Data and Method

- GISAID tree until 2021-06-03 as initial tree.
- The tree contains 541,000 ??? sequences.
- The small branches are collapsed.
- 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 Sample breakdown (daily)

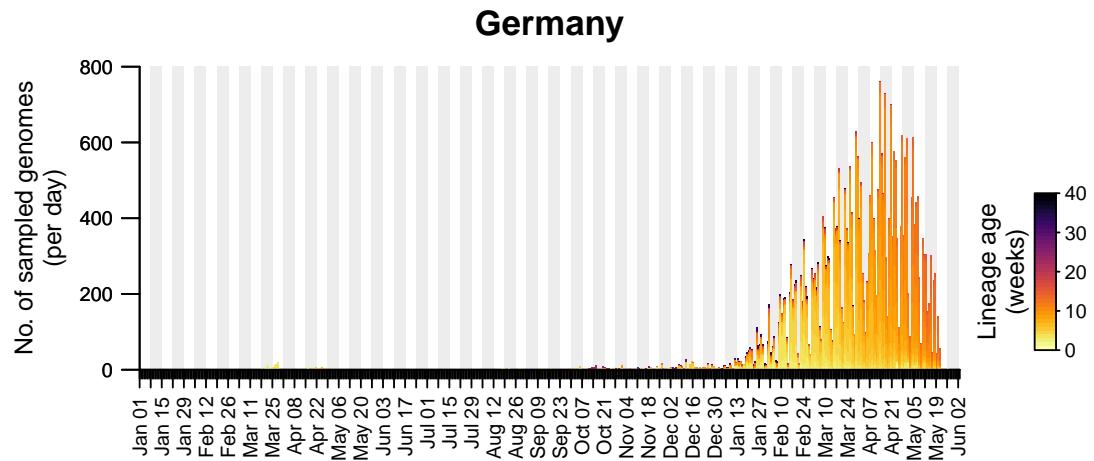


Figure 1: Number of genomes collected each day, coloured by the time since the age of the transmission lineage when the genome was collected (time from the oldest sampled genome in the lineage to the sampling time of the genome). Note that **only** genomes in transmission lineages are shown (no singletons).

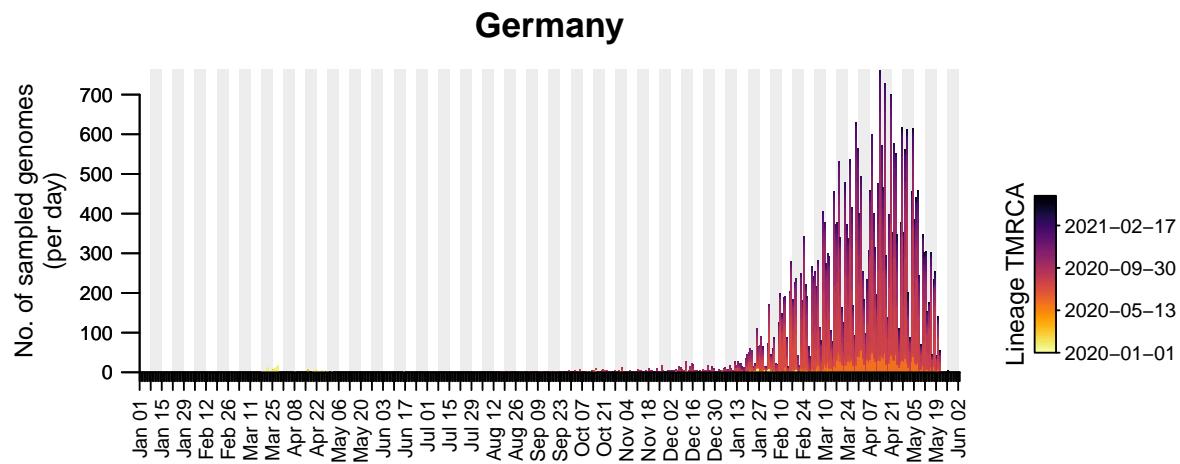


Figure 2: Number of genomes collected each day, coloured by the TMRCA of the transmission lineage. Note that **only** genomes in transmission lineages are shown (no singletons).

3 Sample breakdown into lineages (weekly)

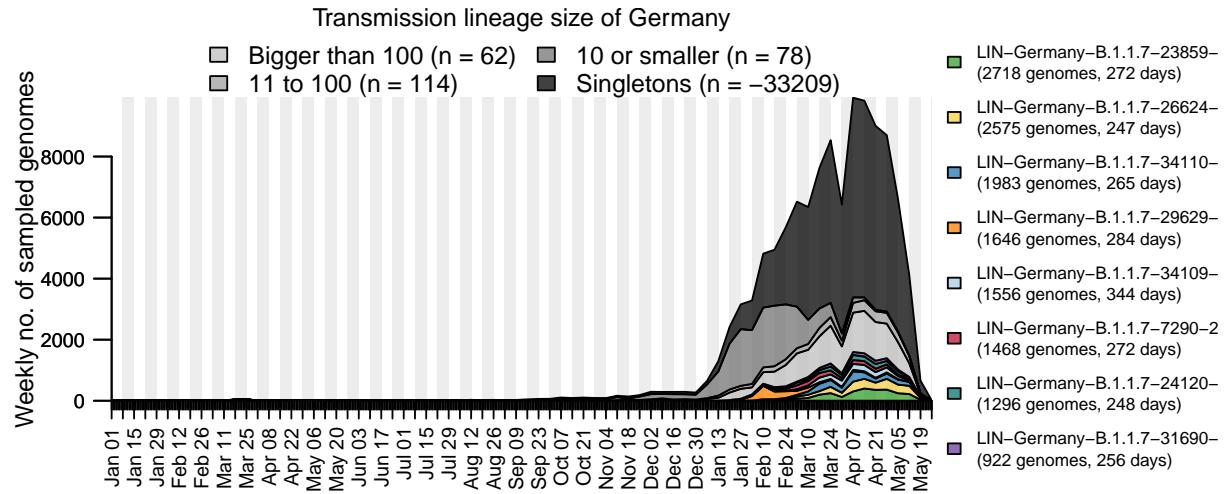
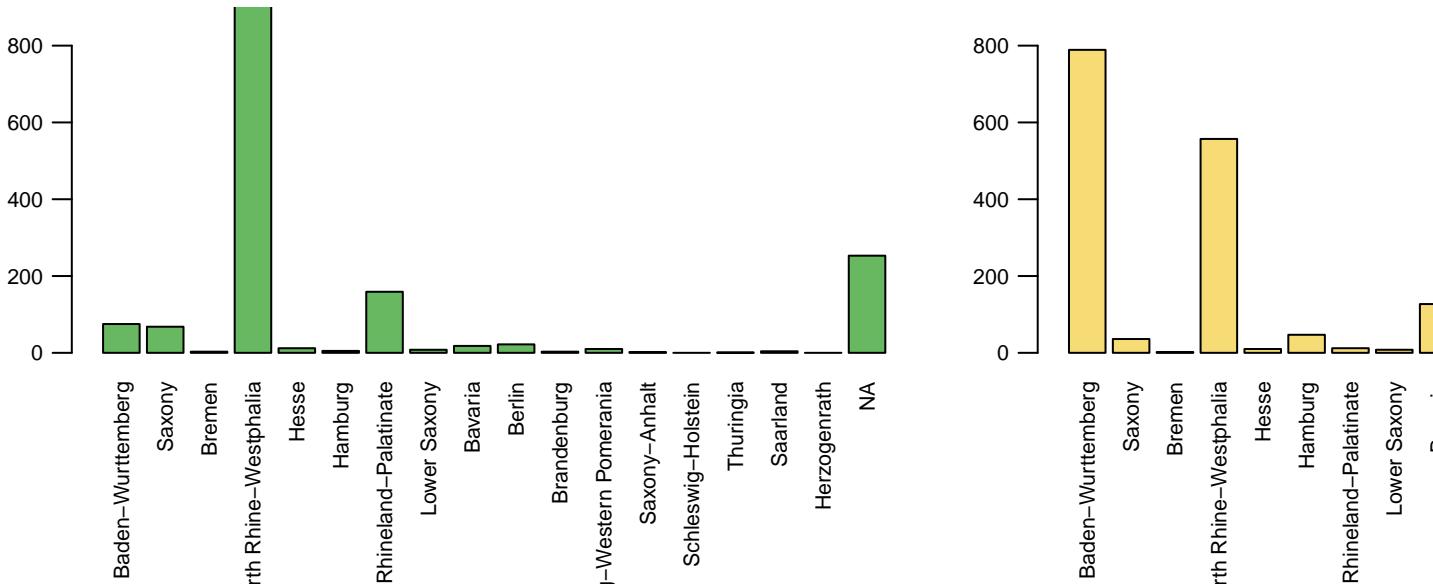


Figure 3: Lineage size breakdown of Germany genomes collected each week. The 8 largest lineages are coloured.



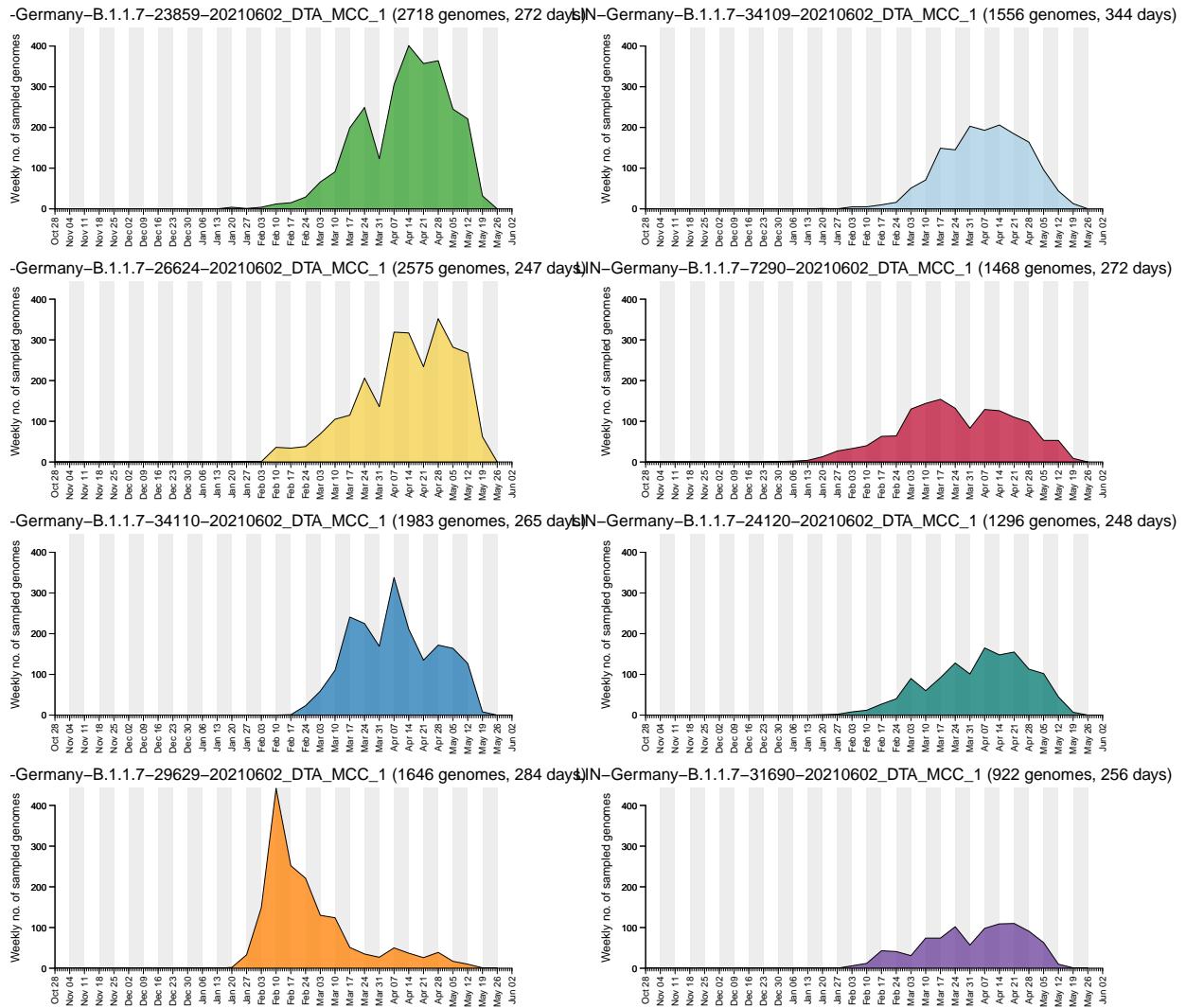


Figure 4: The weekly sampling frequency of the 8 largest r state transmission lineages.

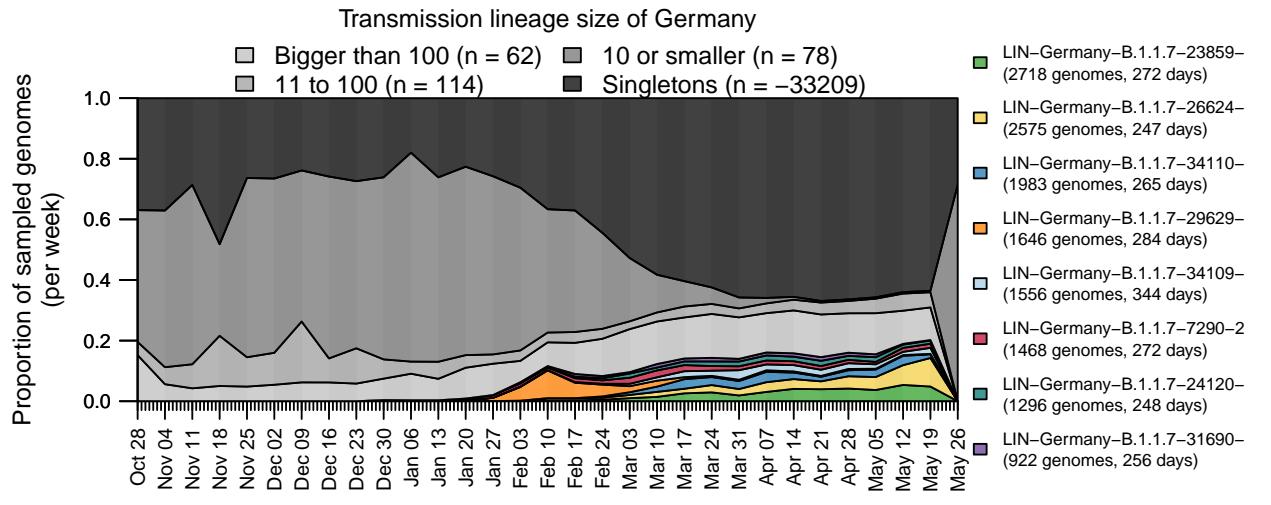
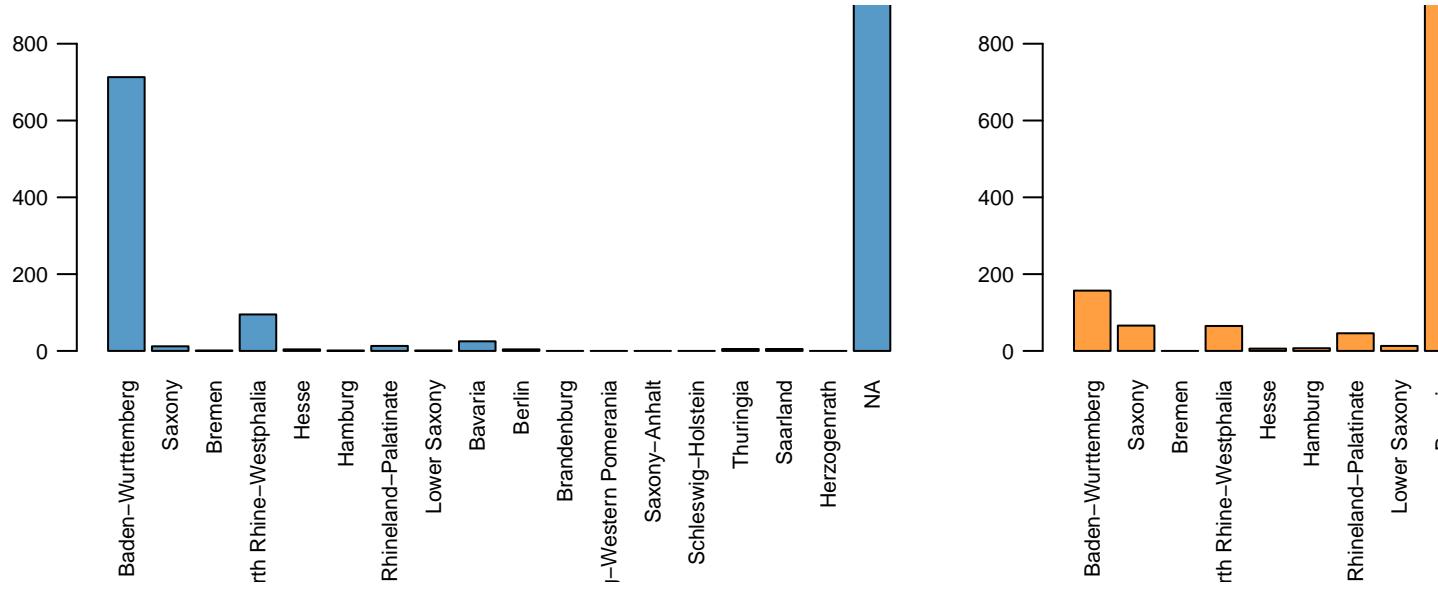
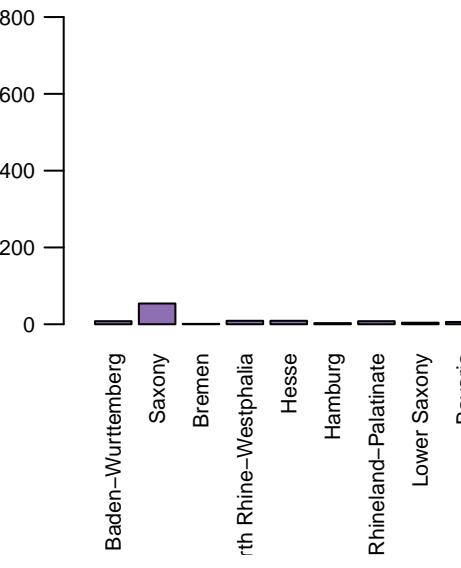
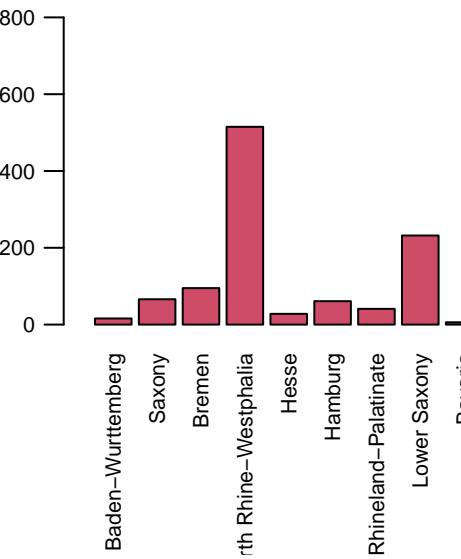
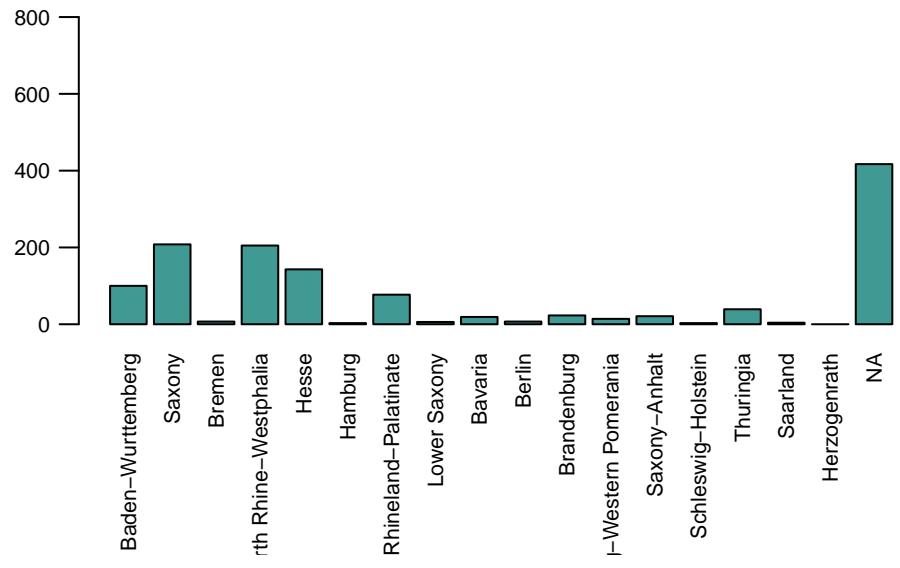
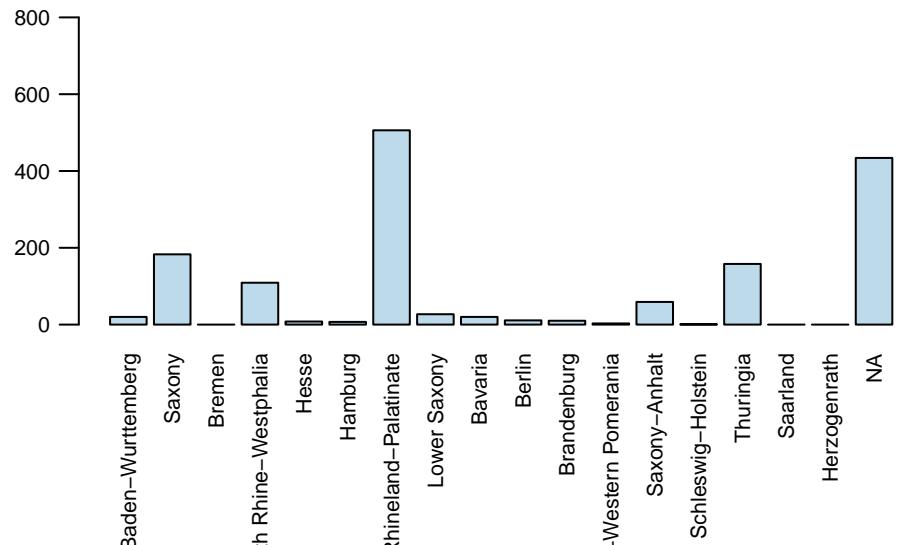
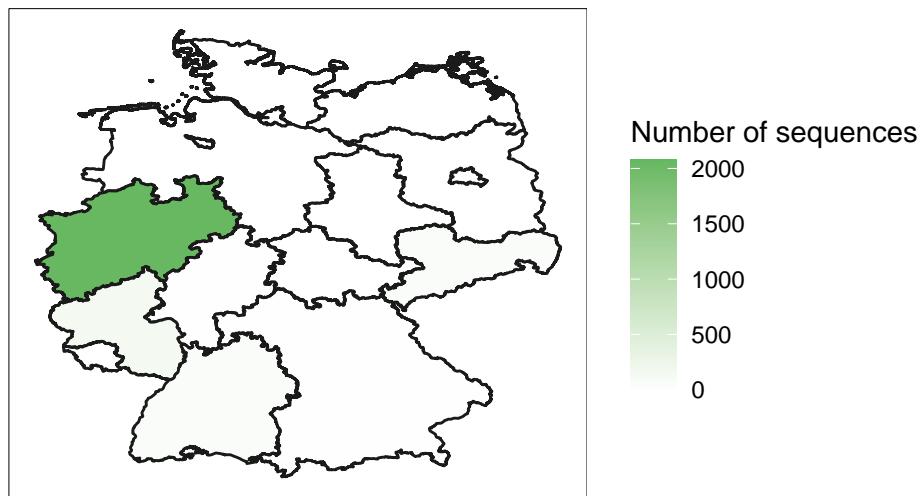


Figure 5: Lineage size breakdown of Germany genomes collected each week. The 8 largest lineages are coloured.

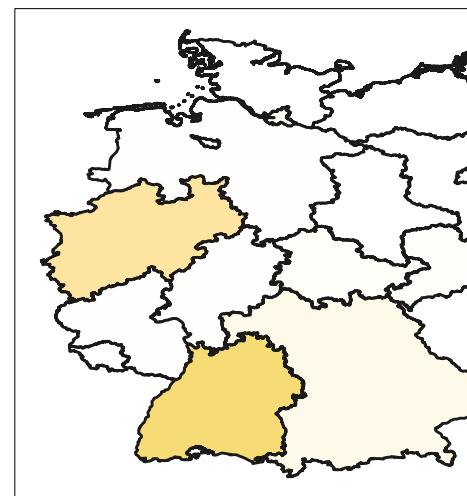




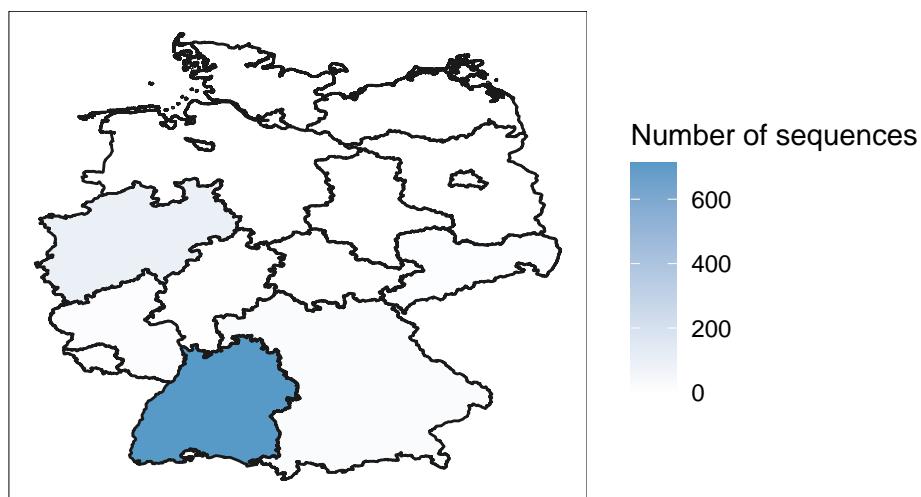
LIN-Germany-B.1.1.7-23859-20210602_DTA_MCC_1



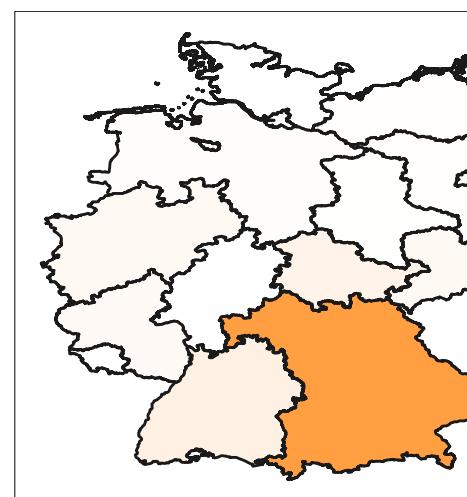
LIN-Germany-B.1.1.7-26624



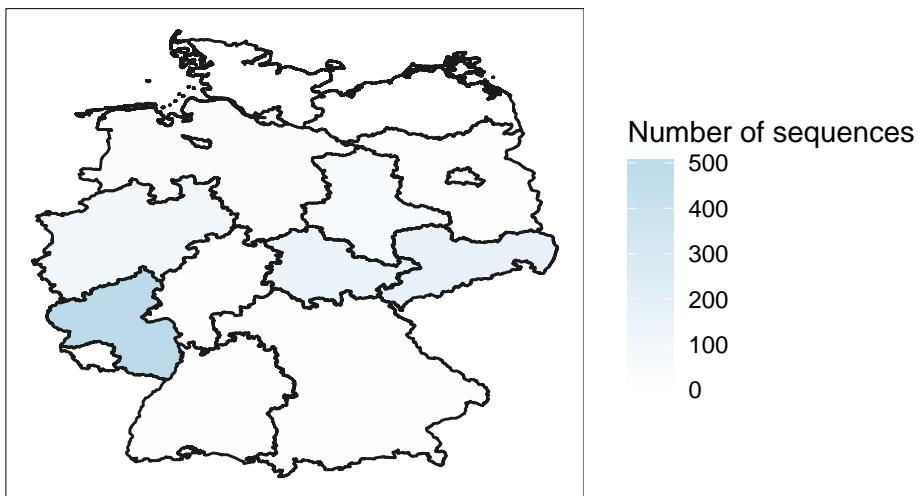
LIN-Germany-B.1.1.7-34110-20210602_DTA_MCC_1



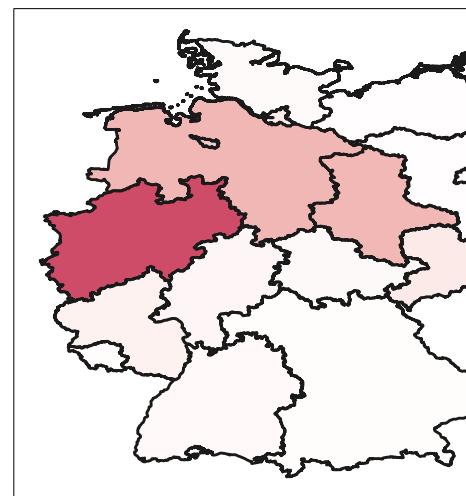
LIN-Germany-B.1.1.7-29629



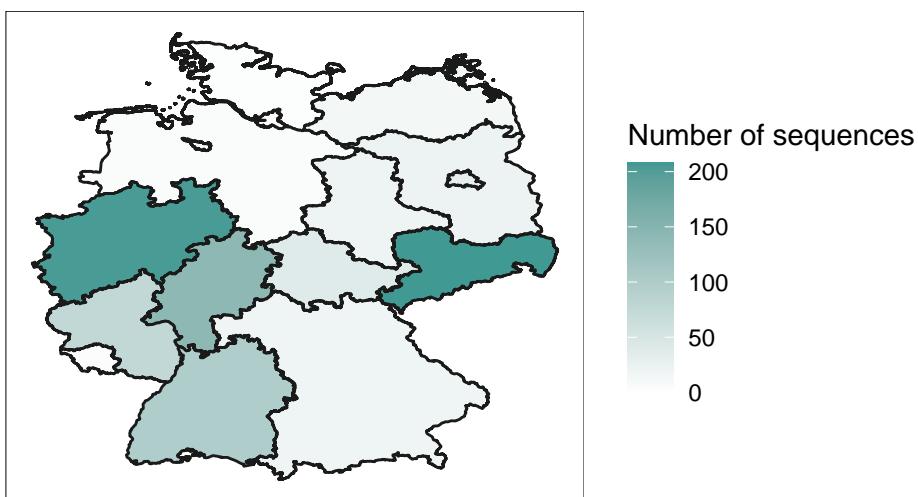
LIN-Germany-B.1.1.7-34109-20210602_DTA_MCC_1



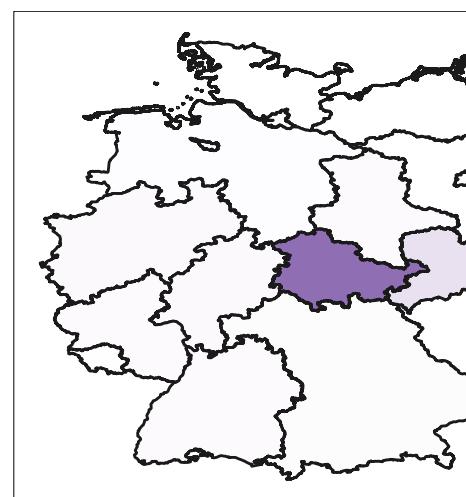
LIN-Germany-B.1.1.7-7290-



LIN-Germany-B.1.1.7-24120-20210602_DTA_MCC_1



LIN-Germany-B.1.1.7-31690-



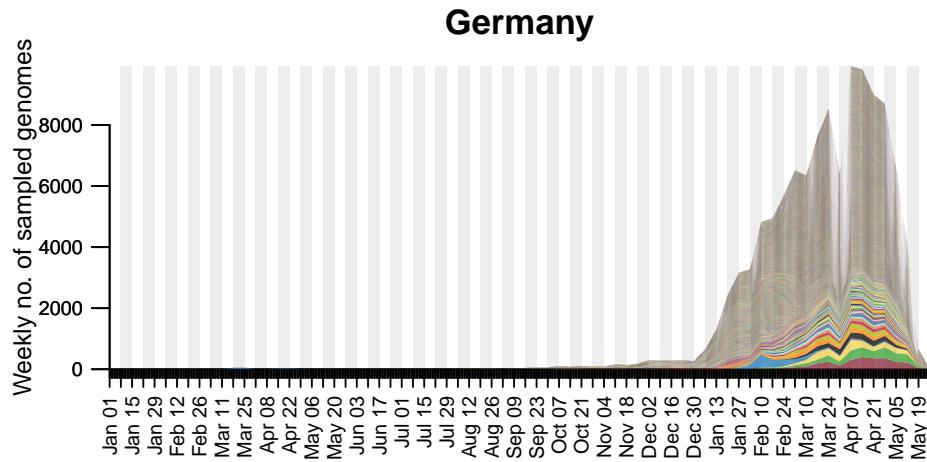


Figure 6: Lineage size breakdown of Germany genomes collected each week.

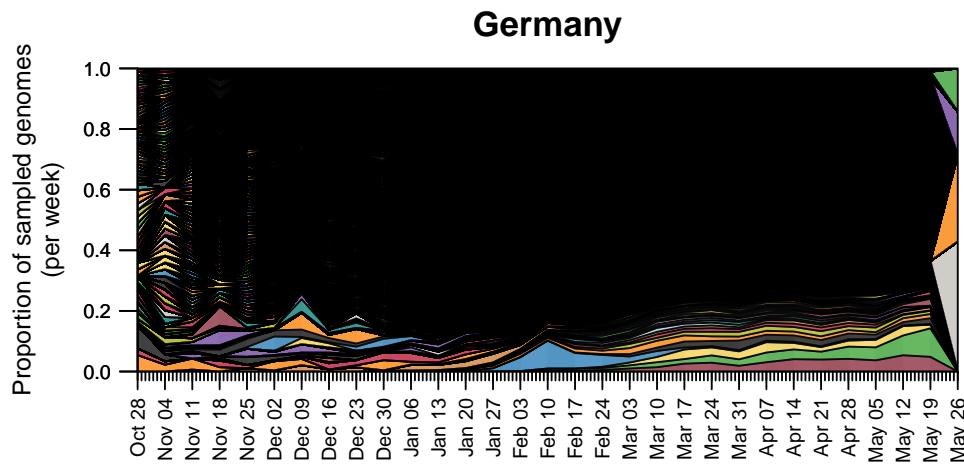


Figure 7: Lineage size breakdown of Germany genomes collected each week.

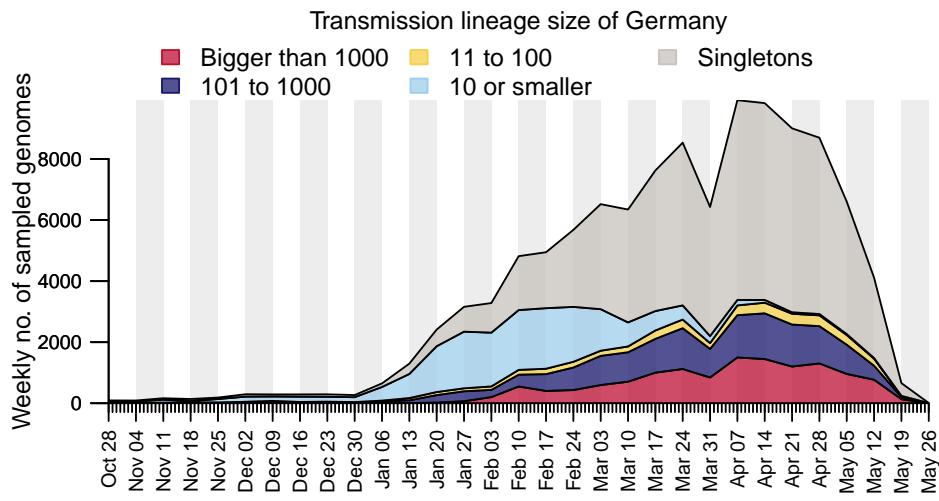
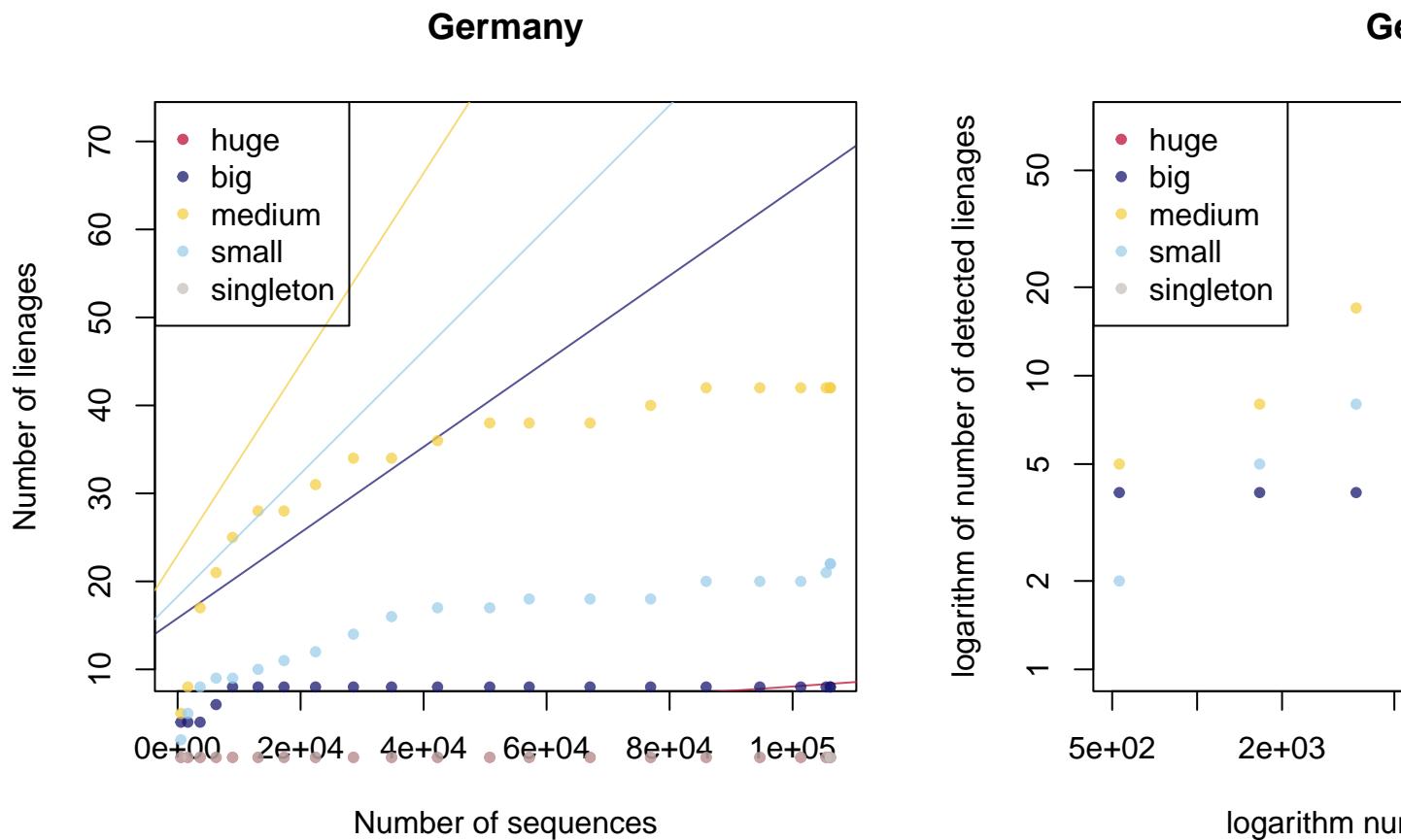


Figure 8: Lineage size breakdown of Germany genomes collected each week.

```
## Warning in xy.coords(x, y, xlabel, ylabel, log): 21 y values <= 0 omitted from
## logarithmic plot
```



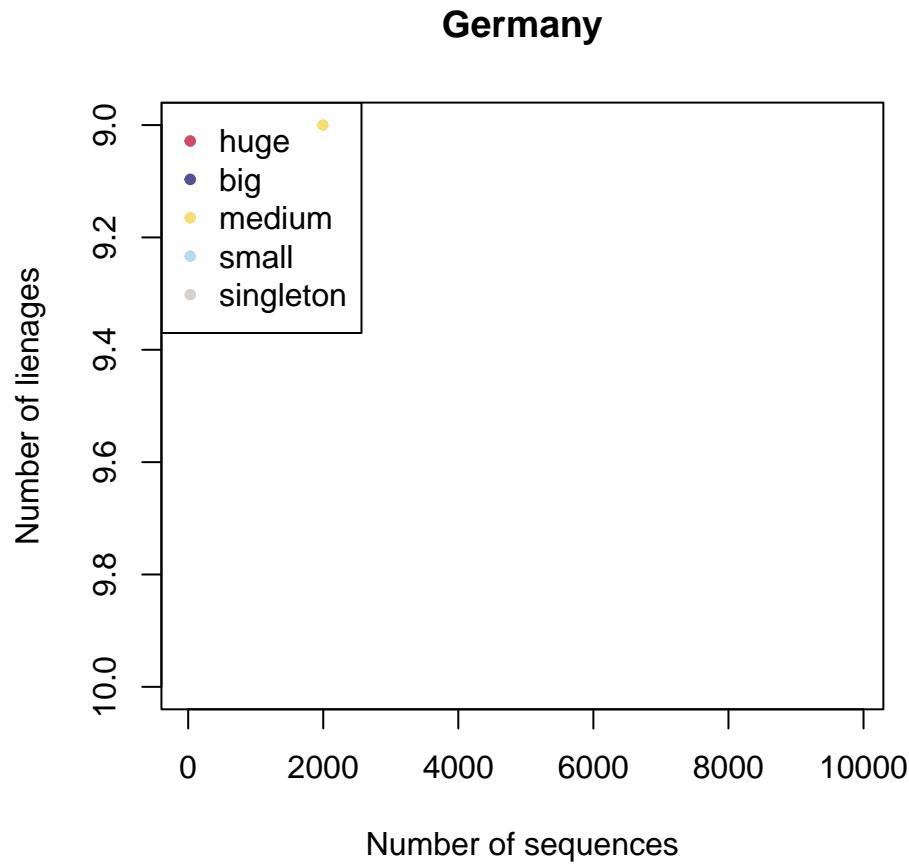


Figure 9: Lineage size breakdown of Germany genomes collected each week.

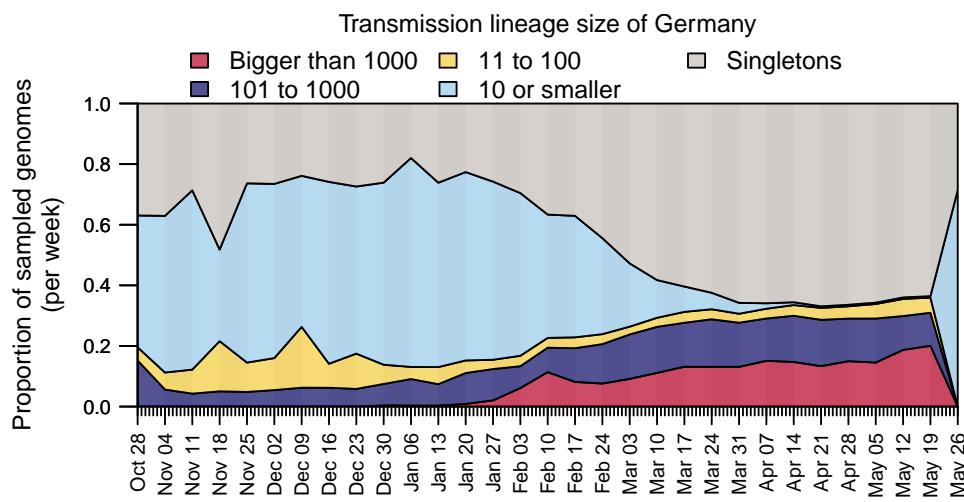


Figure 10: Lineage size breakdown of Germany genomes collected each week.

4 Individual transmission lineage plots

4.1 Biggest transmission lineages

cluster	Pango.lineage	sample_date	lineage_seqs	state
LIN-Germany-B.1.1.7-23859-20210602_DTA_MCC_1	B.1.1.7	2021-01-22	2718	North Rhine-Westphalia
LIN-Germany-B.1.1.7-23859-20210602_DTA_MCC_1	B.1.1.7	2021-01-22	2718	North Rhine-Westphalia
LIN-Germany-B.1.1.7-26624-20210602_DTA_MCC_1	B.1.1.7	2021-01-31	2575	Saxony
LIN-Germany-B.1.1.7-34110-20210602_DTA_MCC_1	B.1.1.7	2021-02-18	1983	Baden-Wurttemberg
LIN-Germany-B.1.1.7-29629-20210602_DTA_MCC_1	B.1.1.7	2021-01-25	1646	Thuringia
LIN-Germany-B.1.1.7-34109-20210602_DTA_MCC_1	B.1.1.7	2021-01-22	1556	Thuringia
LIN-Germany-B.1.1.7-7290-20210602_DTA_MCC_1	B.1.1.7	2021-01-02	1468	North Rhine-Westphalia
LIN-Germany-B.1.1.7-24120-20210602_DTA_MCC_1	B.1.1.7	2021-01-25	1296	Rhineland-Palatinate
LIN-Germany-B.1.1.7-31690-20210602_DTA_MCC_1	B.1.1.7	2021-02-05	922	Thuringia
LIN-Germany-B.1.1.7-31690-20210602_DTA_MCC_1	B.1.1.7	2021-02-05	922	Thuringia
LIN-Germany-B.1.1.7-31690-20210602_DTA_MCC_1	B.1.1.7	2021-02-05	922	Thuringia
LIN-Germany-B.1.1.7-31690-20210602_DTA_MCC_1	B.1.1.7	2021-02-05	922	Thuringia
LIN-Germany-B.1.1.7-55470-20210602_DTA_MCC_1	B.1.1.7	2021-01-22	767	North Rhine-Westphalia
LIN-Germany-B.1.1.7-33946-20210602_DTA_MCC_1	B.1.1.7	2021-02-08	635	Bavaria
LIN-Germany-B.1.1.7-35161-20210602_DTA_MCC_1	B.1.1.7	2021-01-19	608	Saxony
LIN-Germany-B.1.1.7-25980-20210602_DTA_MCC_1	B.1.1.7	2021-02-09	577	Saxony
LIN-Germany-B.1.1.7-63701-20210602_DTA_MCC_1	B.1.1.7	2021-02-11	520	Berlin
LIN-Germany-B.1.1.7-9569-20210602_DTA_MCC_1	B.1.1.7	2021-02-09	518	North Rhine-Westphalia
LIN-Germany-B.1.1.7-27729-20210602_DTA_MCC_1	B.1.1.7	2021-01-27	517	Thuringia
LIN-Germany-B.1.1.7-9106-20210602_DTA_MCC_1	B.1.1.7	2021-01-28	478	Baden-Wurttemberg
LIN-Germany-B.1.1.7-38045-20210602_DTA_MCC_1	B.1.1.7	2021-01-27	438	Rhineland-Palatinate
LIN-Germany-B.1.1.7-38114-20210602_DTA_MCC_1	B.1.1.7	2021-01-14	432	Baden-Wurttemberg
LIN-Germany-B.1.1.7-35231-20210602_DTA_MCC_1	B.1.1.7	2021-02-09	430	
LIN-Germany-B.1.1.7-39671-20210602_DTA_MCC_1	B.1.1.7	2021-02-08	426	North Rhine-Westphalia
LIN-Germany-B.1.1.7-38575-20210602_DTA_MCC_1	B.1.1.7	2021-02-17	415	Berlin

cluster	Pango.lineage	sample_date	lineage_seqs	state
LIN-Germany-B.1.160-20210602_DTA_MCC_1029	B.1.160	2020-11-13	378	Saarland
LIN-Germany-B.1.1.7-54961-20210602_DTA_MCC_1	B.1.1.7	2021-02-09	336	Bremen
LIN-Germany-B.1.1.7-54961-20210602_DTA_MCC_1	B.1.1.7	2021-02-09	336	Bremen
LIN-Germany-B.1.1.7-32894-20210602_DTA_MCC_1	B.1.1.7	2021-02-24	314	Thuringia
LIN-Germany-B.1.1.7-32894-20210602_DTA_MCC_1	B.1.1.7	2021-02-24	314	Thuringia
LIN-Germany-B.1.1.7-32894-20210602_DTA_MCC_1	B.1.1.7	2021-02-24	314	Thuringia
LIN-Germany-B.1.1.7-33722-20210602_DTA_MCC_1	B.1.1.7	2021-02-12	313	Saxony-Anhalt
LIN-Germany-B.1.1.7-33722-20210602_DTA_MCC_1	B.1.1.7	2021-02-12	313	Saxony-Anhalt
LIN-Germany-B.1.1.7-59513-20210602_DTA_MCC_1	B.1.1.7	2021-02-13	280	Bavaria
LIN-Germany-B.1.1.317-20210602_DTA_MCC_38	B.1.1.317	2021-01-07	274	Bavaria
LIN-Germany-B.1.1.7-30660-20210602_DTA_MCC_1	B.1.1.7	2021-02-26	274	Rhineland-Palatinate
LIN-Germany-B.1.1.7-30660-20210602_DTA_MCC_1	B.1.1.7	2021-02-26	274	Saxony
LIN-Germany-B.1.1.7-18644-20210602_DTA_MCC_1	B.1.1.7	2021-02-18	259	Lower Saxony
LIN-Germany-B.1.1.7-27635-20210602_DTA_MCC_1	B.1.1.7	2021-02-22	243	Saxony-Anhalt
LIN-Germany-B.1.1.7-27635-20210602_DTA_MCC_1	B.1.1.7	2021-02-22	243	
LIN-Germany-B.1.1.7-23525-20210602_DTA_MCC_1	B.1.1.7	2021-01-29	231	Bavaria
LIN-Germany-B.1.1.7-23525-20210602_DTA_MCC_1	B.1.1.7	2021-01-29	231	Bavaria
LIN-Germany-B.1.1.7-23525-20210602_DTA_MCC_1	B.1.1.7	2021-01-29	231	Bavaria
LIN-Germany-B.1.221-20210602_DTA_MCC_4	B.1.221.2	2020-03-21	229	Saarland
LIN-Germany-B.1.1.7-64316-20210602_DTA_MCC_1	B.1.1.7	2021-02-16	223	North Rhine-Westphalia
LIN-Germany-B.1.221-20210602_DTA_MCC_306	B.1.221	2020-10-17	222	Saarland
LIN-Germany-B.1.221-20210602_DTA_MCC_306	B.1.221	2020-10-17	222	Saarland
LIN-Germany-B.1.1.7-40790-20210602_DTA_MCC_1	B.1.1.7	2021-02-13	210	Bremen
LIN-Germany-B.1.1.7-40790-20210602_DTA_MCC_1	B.1.1.7	2021-02-13	210	Bremen
LIN-Germany-B.1.1.7-40790-20210602_DTA_MCC_1	B.1.1.7	2021-02-13	210	Bremen
LIN-Germany-B.1.351-20210602_DTA_MCC_1624	B.1.351	2021-02-04	205	Saarland

cluster	Pango.lineage	sample_date	lineage_seqs	state
LIN-Germany-B.1.351-20210602_DTA_MCC_1624	B.1.351	2021-02-04	205	Saarland
LIN-Germany-B.1.1.7-54346-20210602_DTA_MCC_1	B.1.1.7	2021-03-08	191	Baden-Wurttemberg
LIN-Germany-B.1.1.7-60099-20210602_DTA_MCC_1	B.1.1.7	2021-02-25	188	Thuringia
LIN-Germany-B.1.1.7-17101-20210602_DTA_MCC_1	B.1.1.7	2021-02-19	186	Hamburg
LIN-Germany-B.1.1.7-17101-20210602_DTA_MCC_1	B.1.1.7	2021-02-19	186	Hamburg
LIN-Germany-B.1.1.7-53110-20210602_DTA_MCC_1	B.1.1.7	2021-03-05	183	North Rhine-Westphalia
LIN-Germany-B.1.1.7-53110-20210602_DTA_MCC_1	B.1.1.7	2021-03-05	183	North Rhine-Westphalia
LIN-Germany-B.1.1.7-53110-20210602_DTA_MCC_1	B.1.1.7	2021-03-05	183	North Rhine-Westphalia
LIN-Germany-B.1.1.7-62408-20210602_DTA_MCC_1	B.1.1.7	2021-01-21	173	Hesse
LIN-Germany-B.1.1.7-26692-20210602_DTA_MCC_1	B.1.1.7	2021-02-27	170	Baden-Wurttemberg
LIN-Germany-A-20210602_DTA_MCC_68	A	2020-03-14	144	Hamburg
LIN-Germany-B.1.1.317-20210602_DTA_MCC_63	B.1.1.317	2021-01-22	142	Bavaria
LIN-Germany-B.1.1.7-4662-20210602_DTA_MCC_1	B.1.1.7	2021-02-19	142	Thuringia
LIN-Germany-B.1.1.7-36198-20210602_DTA_MCC_1	B.1.1.7	2021-03-02	139	Baden-Wurttemberg
LIN-Germany-B.1.221-20210602_DTA_MCC_920	B.1.221	2020-10-01	139	Saarland
LIN-Germany-B.1.1.7-2546-20210602_DTA_MCC_1	B.1.1.7	2021-03-26	136	North Rhine-Westphalia
LIN-Germany-B.1.1.7-8475-20210602_DTA_MCC_1	B.1.1.7	2021-03-04	131	North Rhine-Westphalia
LIN-Germany-B.1.221-20210602_DTA_MCC_90	B.1.221	2020-10-19	131	Saarland

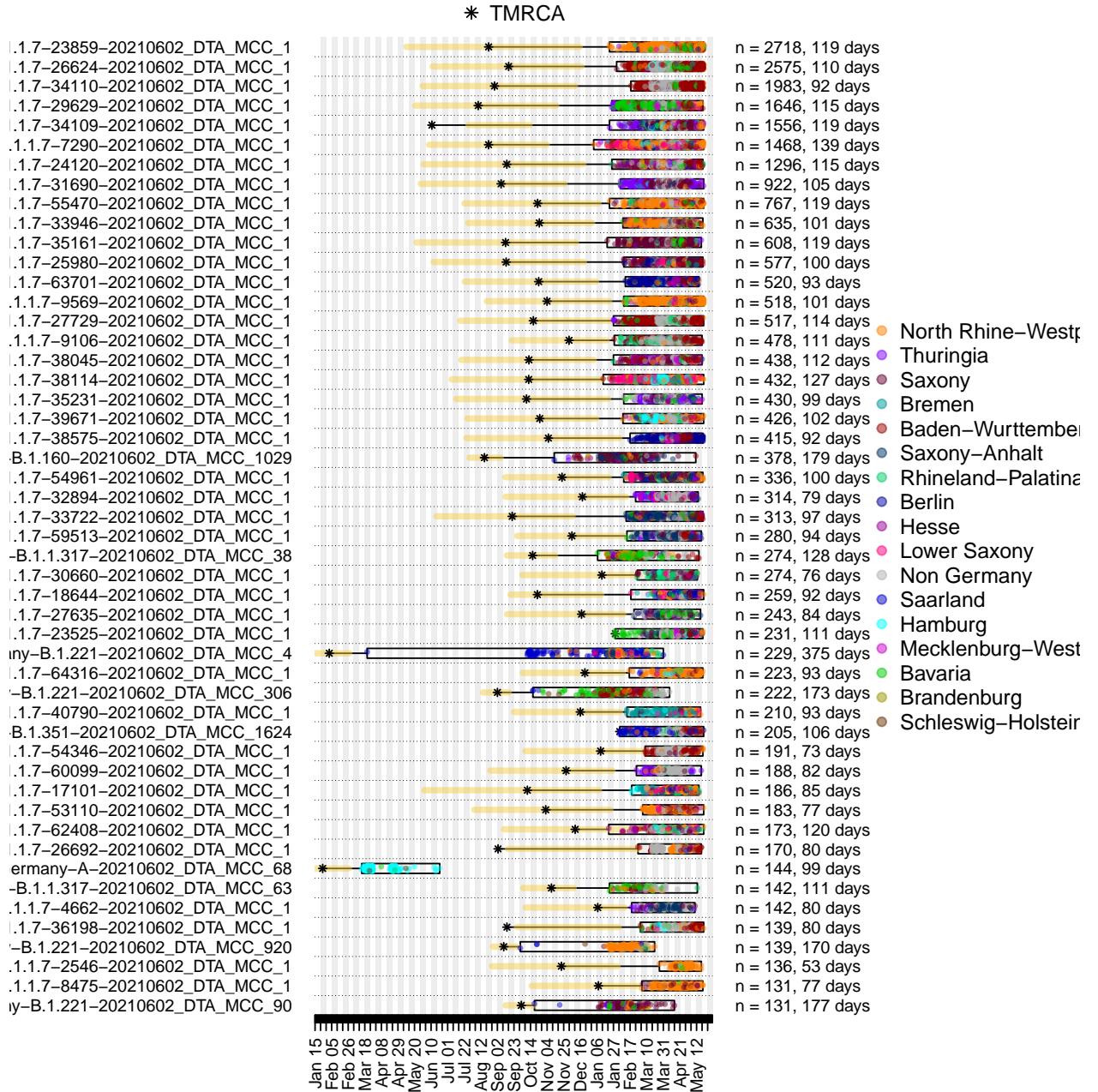


Figure 11: Illustration of the time course of the 50 largest Düsseldorf transmission lineages in our dataset. Each row is a transmission lineage. Dots are genome sampling times and boxes show the range of sampling times for each transmission lineage (sampling duration). Asterisks show the TMRCA of each lineage. On the right, n indicates the number of Düsseldorf genomes in the lineage and the duration of lineage detection (time between the lineage's oldest and most recent genomes). Sampling times of the first 500 SARS-CoV-2 genomes collected in the Düsseldorf have been obscured.

4.2 Earliest transmission lineages

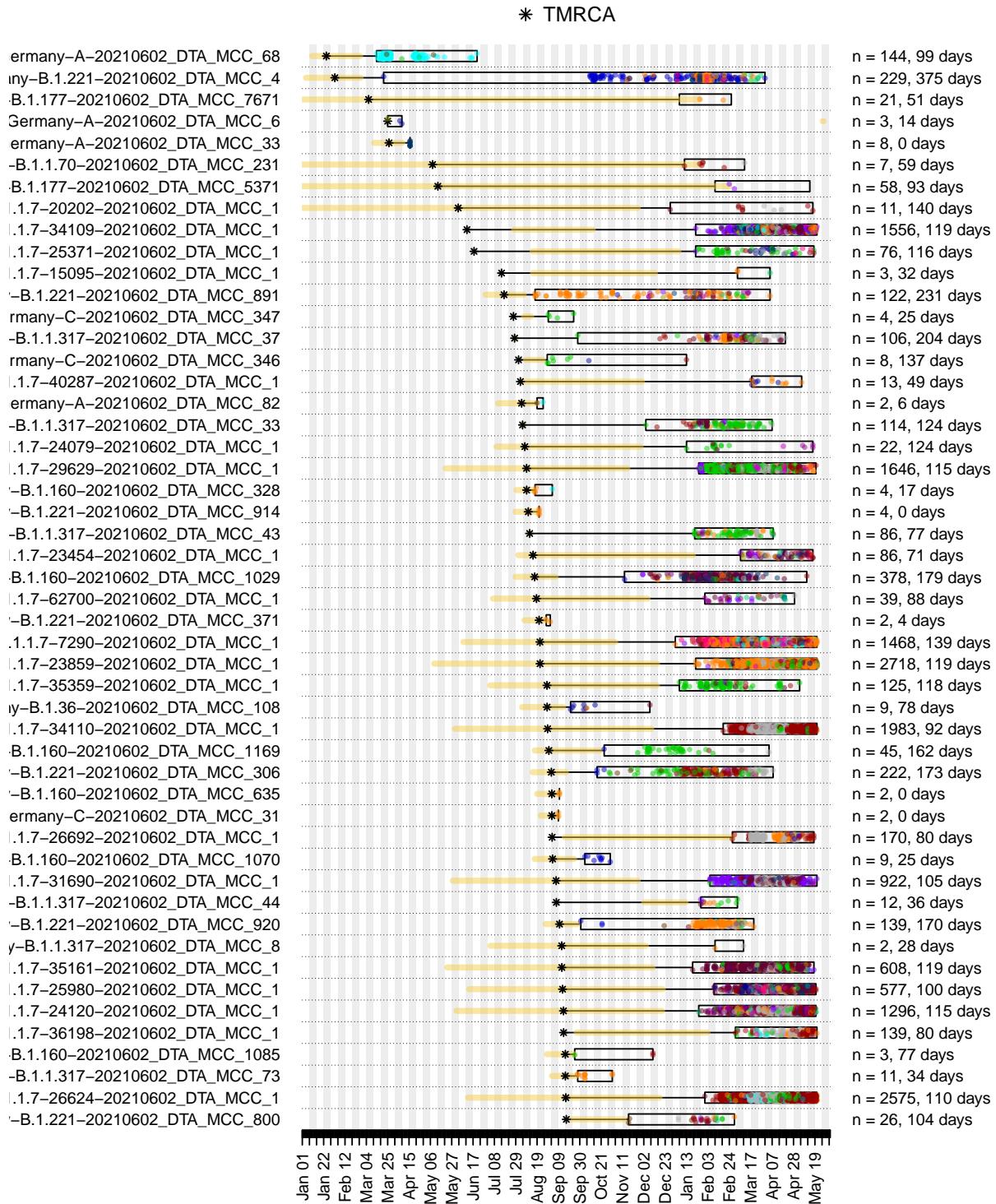
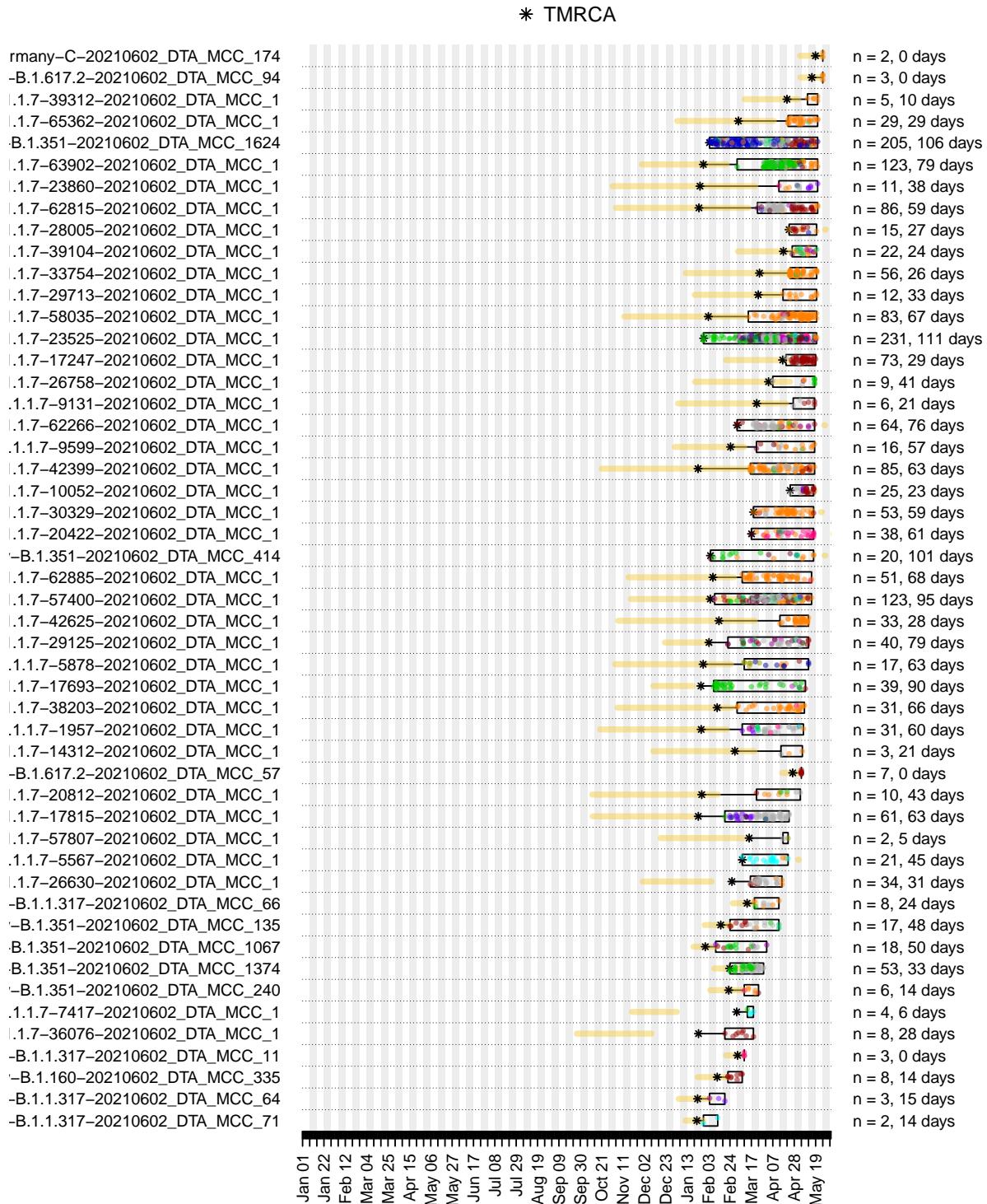
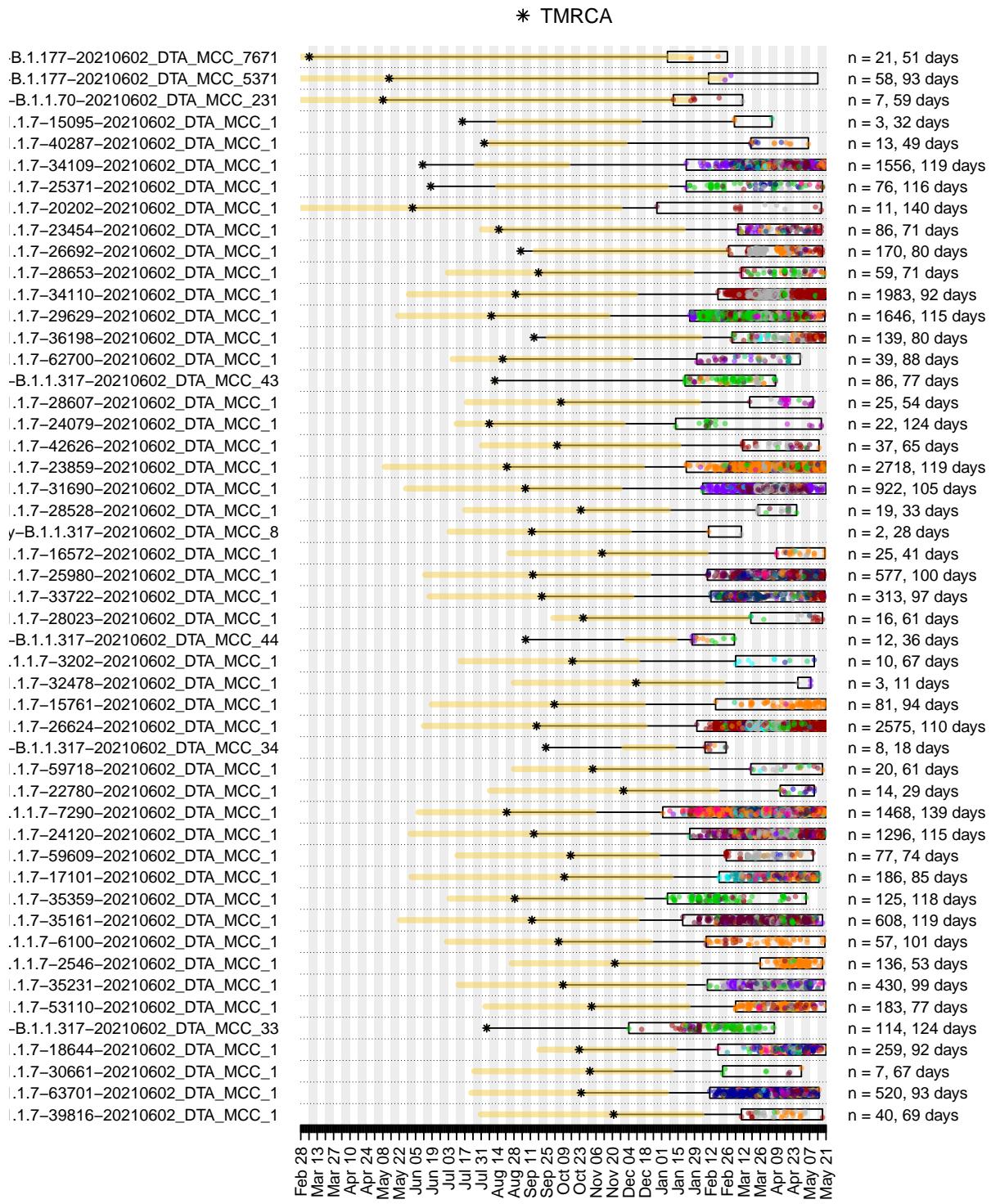


Figure 12: Illustration of the time course of the 50 earliest Dusseldorf transmission lineages in our dataset. Each row is a transmission lineage. Dots are genome sampling times and boxes show the range of sampling times for each transmission lineage (sampling duration). Asterisks show the TMRCA of each lineage. On the right, n indicates the number of Dusseldorf genomes in the lineage and the duration of lineage detection (time between the lineage's oldest and most recent genomes). Sampling times of the first 500 SARS-CoV-2 genomes collected in the Dusseldorf have been obscured.

4.3 Newest transmission lineages



Longest periods of cryptic circulation



Longest sampling period

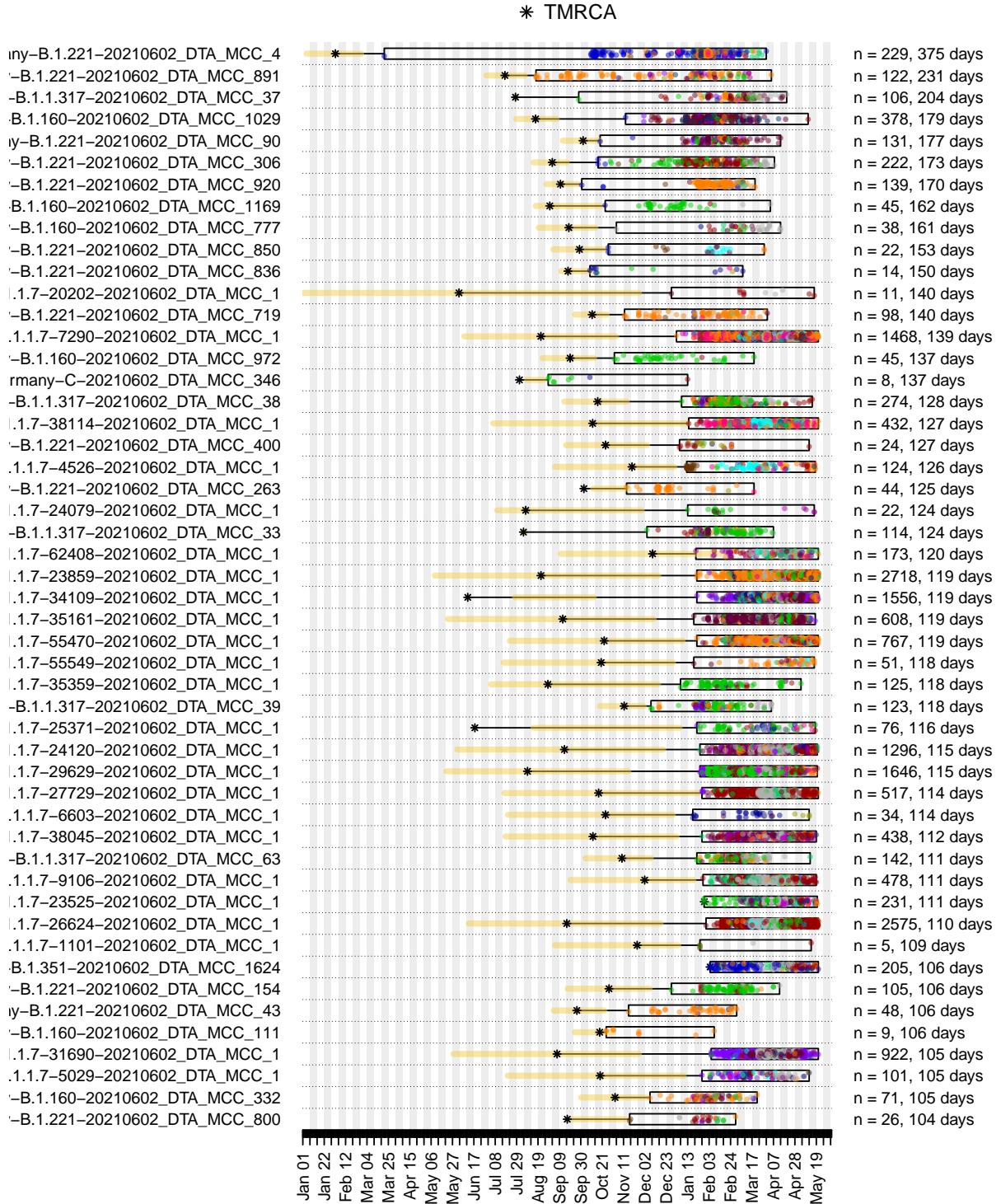


Figure 13: Illustration of the time course of the 50 Dusseldorf transmission lineages in our dataset with the longest sampling duration (from earliest to most recently collected genome). Each row is a transmission lineage. Dots are genome sampling times and boxes show the range of sampling times for each transmission lineage (sampling duration). Asterisks show the TMRCA of each lineage. On the right, n indicates the number of Dusseldorf genomes in the lineage and the duration of lineage detection (time between the lineage's oldest and most recent genomes). Sampling times of the first 500 SARS-CoV-2 genomes collected in the Dusseldorf have been obscured.²³

4.4 Longest unobserved period before reactivating

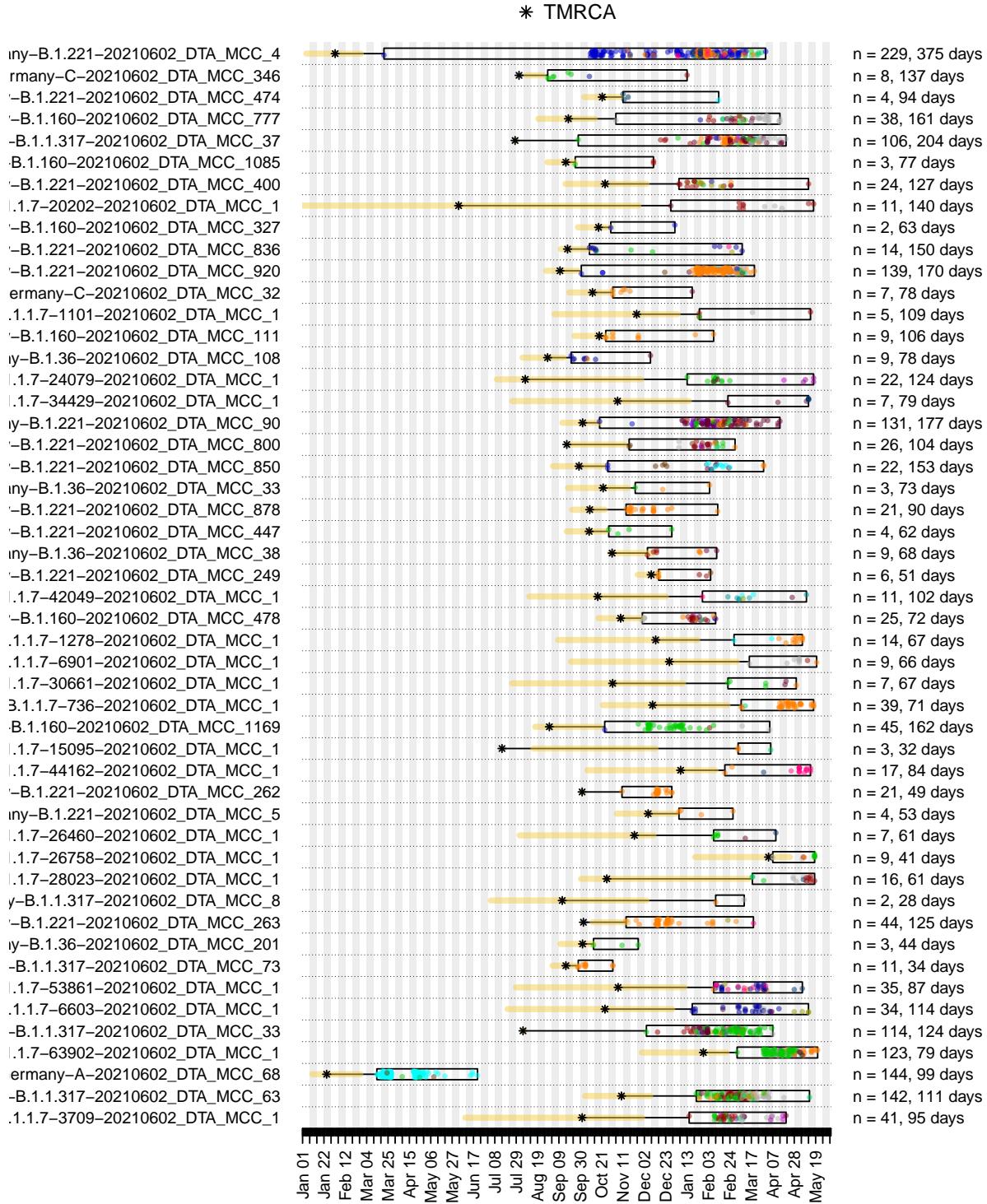


Figure 14: Illustration of the time course of the 50 Dusseldorf transmission lineages in our dataset with the longest unobserved period before reemerging. Each row is a transmission lineage. Dots are genome sampling times and boxes show the range of sampling times for each transmission lineage (sampling duration). Asterisks show the TMRCA of each lineage. On the right, n indicates the number of Dusseldorf genomes in the lineage and the duration of lineage detection (time between the lineage's oldest and most recent genomes). Sampling times of the first 500 SARS-CoV-2 genomes collected in the Dusseldorf have been obscured.
25

5 Session info

```
## R version 4.1.1 (2021-08-10)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 20.04.3 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] maptools_1.1-1    colorspace_2.0-1  ggplot2_3.3.3    raster_3.4-10
## [5] sp_1.4-5        knitr_1.33     dplyr_1.0.6      stringr_1.4.0
## [9] tictoc_1.0.1     viridis_0.6.1   viridisLite_0.4.0 gplots_3.1.1
## [13] lubridate_1.7.10
##
## loaded via a namespace (and not attached):
## [1] Rcpp_1.0.6       highr_0.9       RColorBrewer_1.1-2 compiler_4.1.1
## [5] pillar_1.6.1     bitops_1.0-7     tools_4.1.1       digest_0.6.27
## [9] lattice_0.20-44  evaluate_0.14   lifecycle_1.0.0   tibble_3.1.2
## [13] gridExtra_2.3    gtable_0.3.0    rlang_0.4.11     yaml_2.2.1
## [17] vctrs_0.3.8     xfun_0.23      withr_2.4.2      rgeos_0.5-5
## [21] generics_0.1.0   glue_1.4.2      gtools_3.8.2     caTools_1.18.2
## [25] tidyselect_1.1.1  grid_4.1.1      glue_1.4.2      R6_2.5.0
## [29] fansi_0.5.0     foreign_0.8-81  rmarkdown_2.8    farver_2.1.0
## [33] purrrr_0.3.4    magrittr_2.0.1   codetools_0.2-18 scales_1.1.1
## [37] htmtools_0.5.1.1 ellipsis_0.3.2  labeling_0.4.2   KernSmooth_2.23-20
## [41] utf8_1.2.1      stringi_1.6.2   munsell_0.5.0   crayon_1.4.1
```