water

Supplementary Figures "Genome Sequencing of SARS-CoV-2 Allows Monitoring of Variants of Concern through Wastewater"

Malte Herold, Aymeric Fouquier d'Hérouël, Patrick May, Francesco Delogu, Anke Wienecke-Baldacchino, Jessica Tapp, Cécile Walczak, Paul Wilmes, Henry-Michel Cauchie, Guillaume Fournier and Leslie Ogorzaly

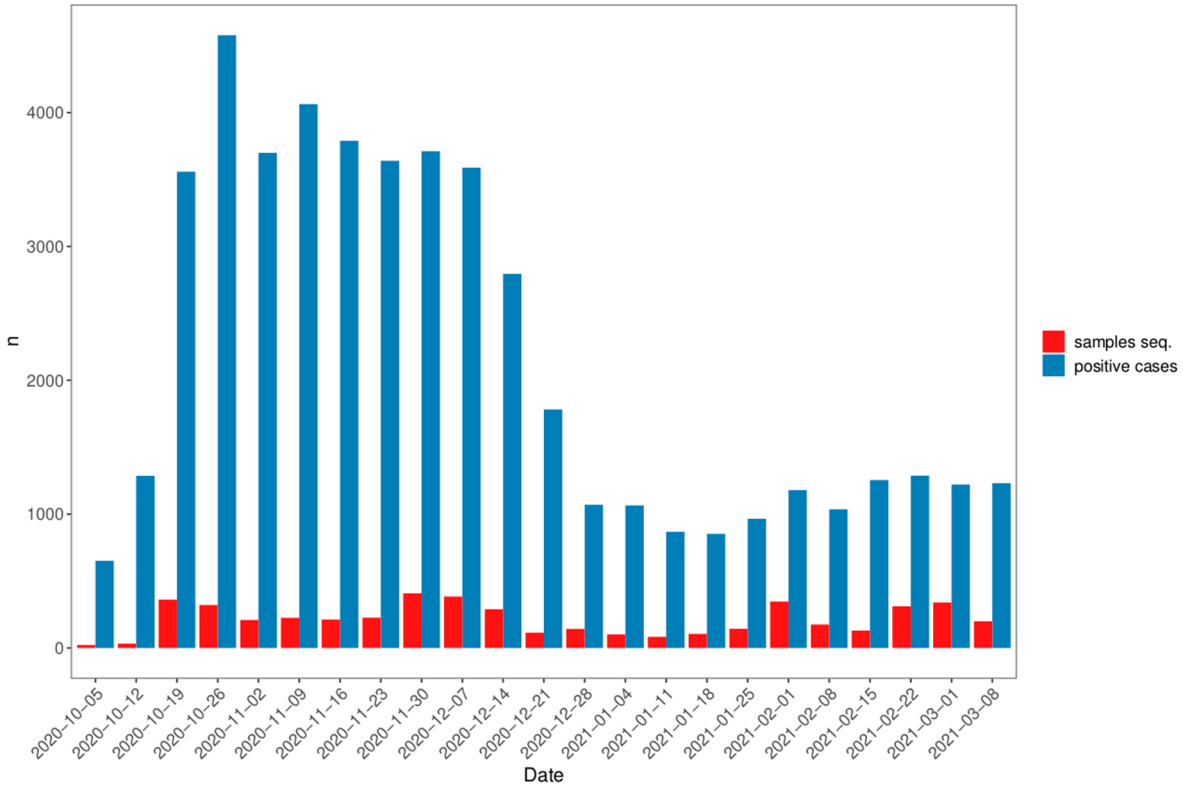

Figure S1. Weekly sequencing coverage.
Weekly sequencing coverage of clinical samples for the time period 1 October 2020 until 8 March 2021. Positive cases per week were retrieved from ECDC (https://www.ecdc.europa.eu/en/publications-data/data-national-14-day-notification-rate-covid-19 accessed: 22 June 2021) while the number of sequenced samples was determined by sequences available in GISAID.

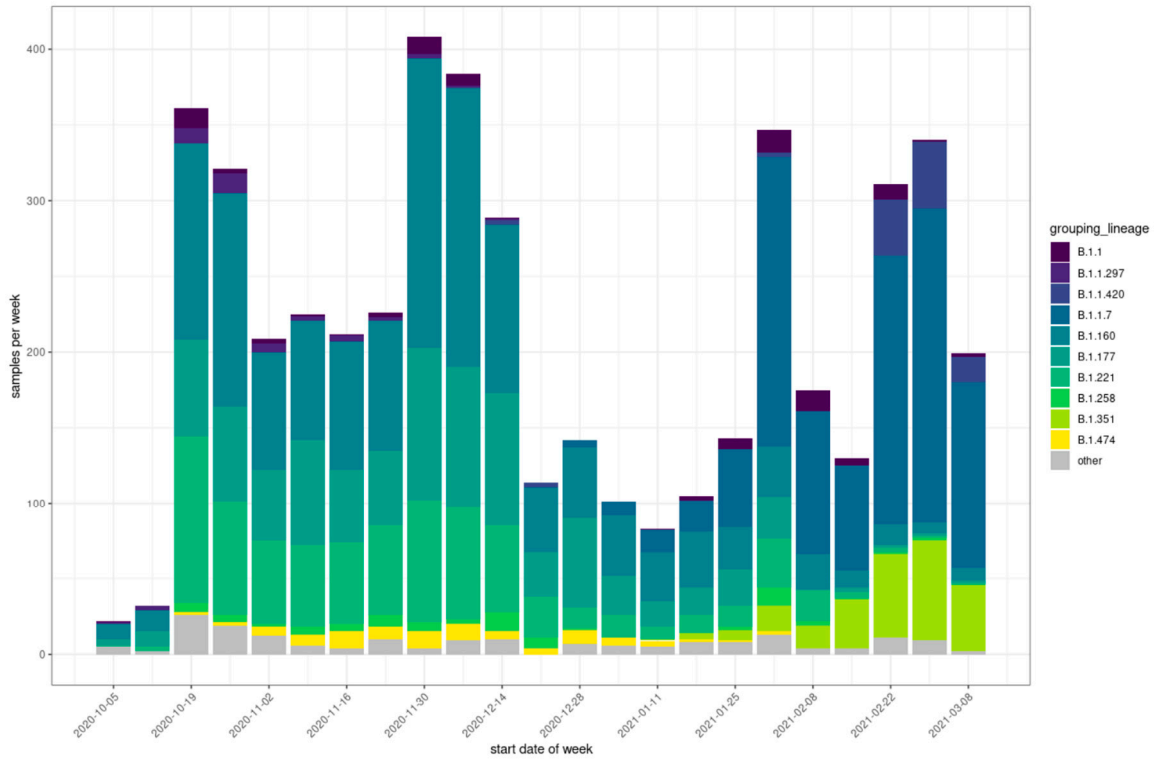

Figure S2. Lineages in clinical consensus sequences.
Number of sequenced clinical samples per week with assigned grouped lineages. Only the 10 most prevalent lineages during the period 1 October 2020 to 9 March 2021 are highlighted by color. The total number of clinical samples shown here is 4714.

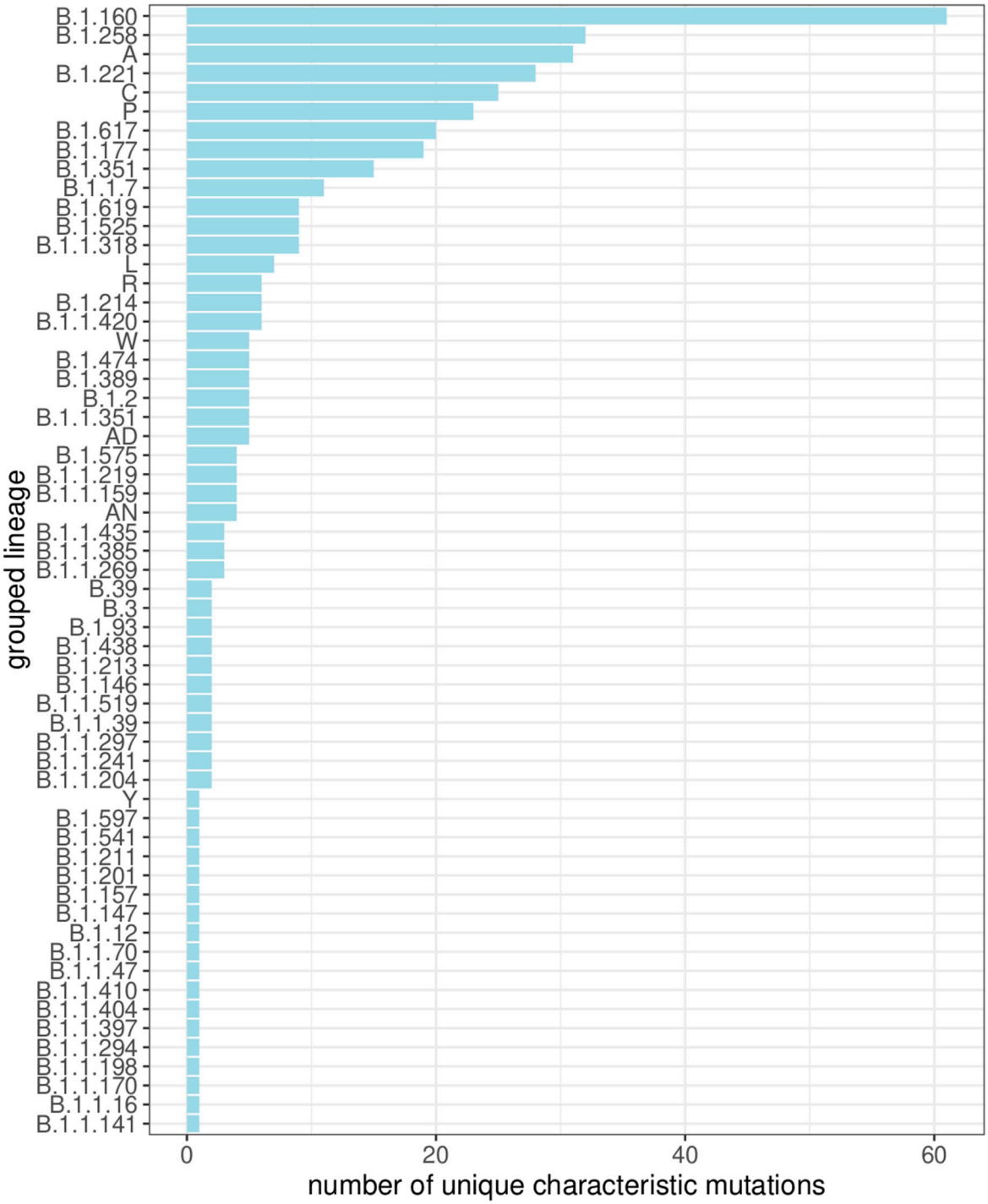

Figure S3. Number of signature mutations for grouped lineages.
Number of unique characteristic mutations for each PANGO-lineage. Lineages were grouped (see methods) and mutations were filtered according to uniquely appearing within a specific group of lineages. Characteristic mutations for each lineage assigned to a clinical sample were downloaded from outbreak.info (https://outbreak.info/ accessed on 15 July 2021).

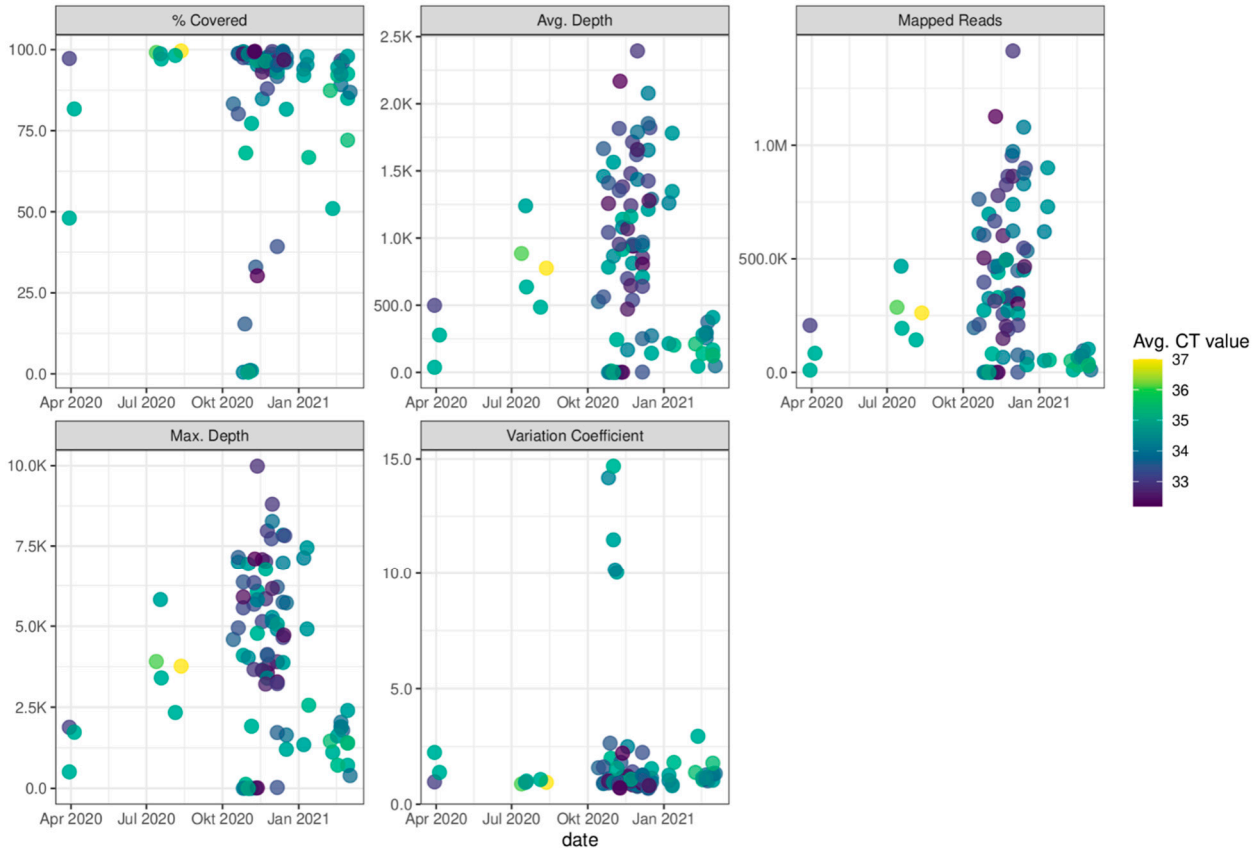

Figure S4. Mapping statistics for wastewater samples.
Mapping statistics for 92 wastewater samples (pre-filtering). The different panels show different measures computed with WeeSAM v.1.6 (https://github.com/centre-for-virus-research/weeSAM). The color indicates the average (across the marker genes) CT value before sequencing.

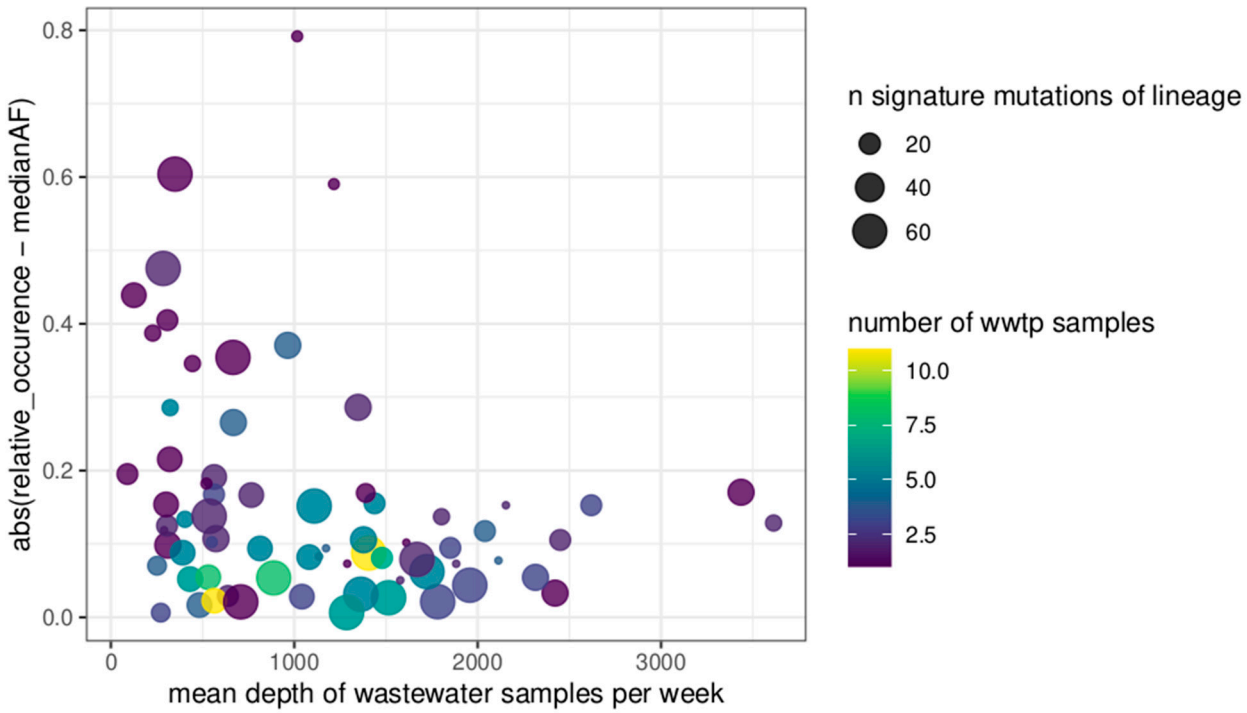

Figure S5. Differences relative occurrence and allele frequencies with sequencing depth and number of signature mutations.

Comparison of median signature mutation allele frequency per lineage (wastewater samples) and clinical samples (relative occurrence per week of that lineage). The absolute difference between the 2 measures (y-axis) is plotted vs. the median sequencing depth of the wastewater samples. Each point represents the comparison of 1 week and 1 lineage. Point size highlights the number of signature mutation of the given lineage while the color indicates the number of wastewater samples that was used to calculate the median allele frequency. The following grouped lineages were selected: "B.1.1.420", "B.1.1.7", "B.1.160", "B.1.177", "B.1.221", "B.1.258", "B.1.351" and "B.1.474".

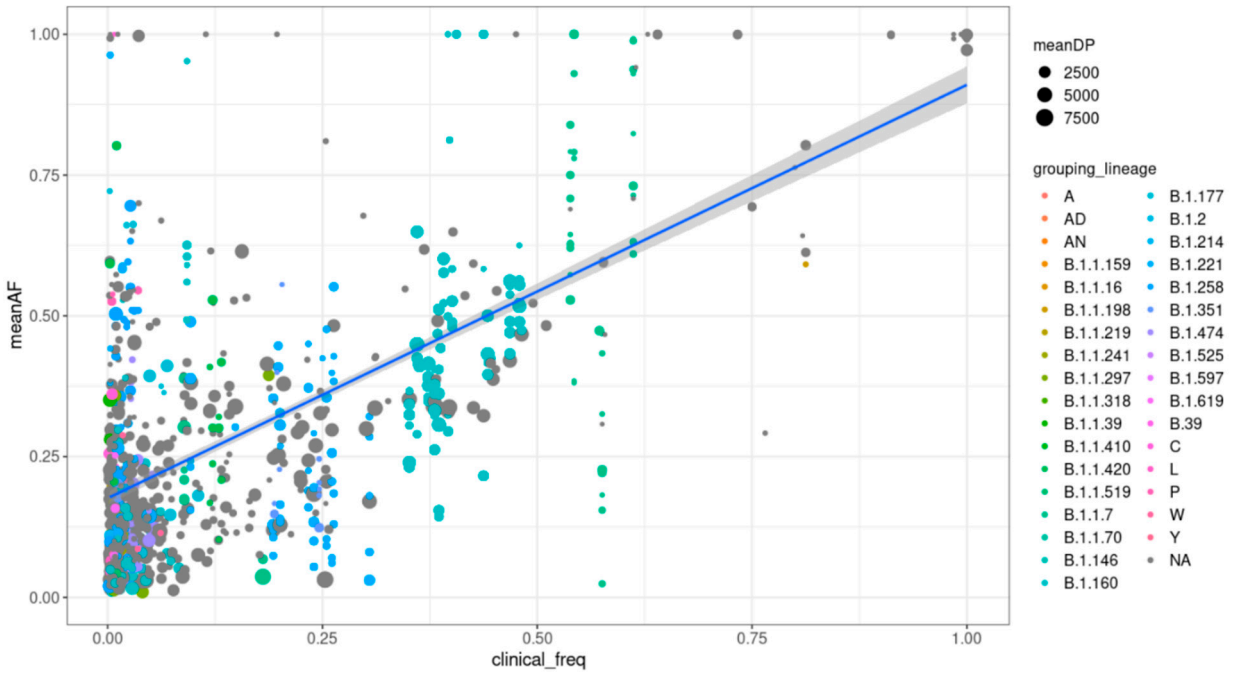

Figure S6. Comparison relative occurrence and allele frequencies grouped by week.
Relative occurrences of amino acid mutations in clinical samples vs allele frequencies of corresponding mutations in wastewater samples grouped by week. Point size reflects mean total depth (meanDP) at the genomic position of the mutations. Colors indicate mutations that uniquely correspond to a grouped lineage. Adjusted R^2: 0.548.

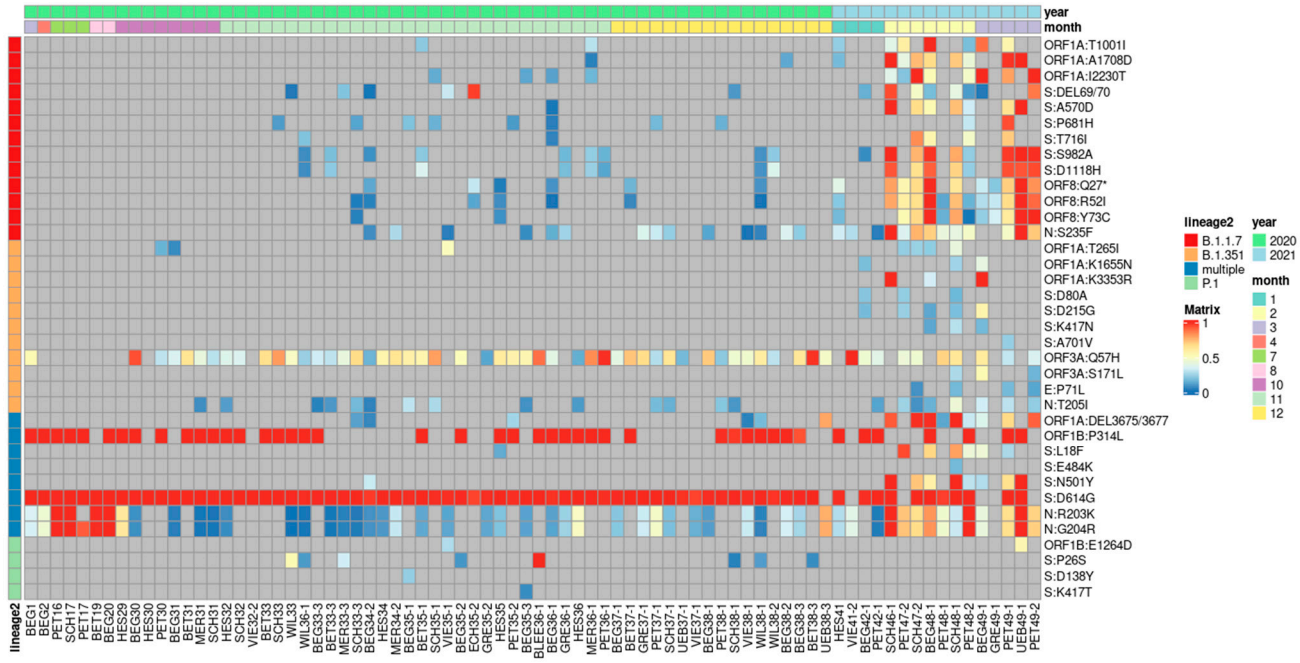

Figure S7. VOC mutations over time.
Sample specific allele frequencies for characteristic mutations derived from literature (outbreak.info) for 3 VOCs: P.1, B.1.351, B.1.1.7. Samples with at least 1 detected mutation are shown arranged by date. Lineages P.1.1 and B.1.351.3 are grouped to P.1 and B.1.351 respectively.