

Cell Host & Microbe, Volume 29

Supplemental information

One year of SARS-CoV-2 evolution

Aiping Wu, Lulan Wang, Hang-Yu Zhou, Cheng-Yang Ji, Shang Zhou Xia, Yang Cao, Jing Meng, Xiao Ding, Sarah Gold, Taijiao Jiang, and Genhong Cheng

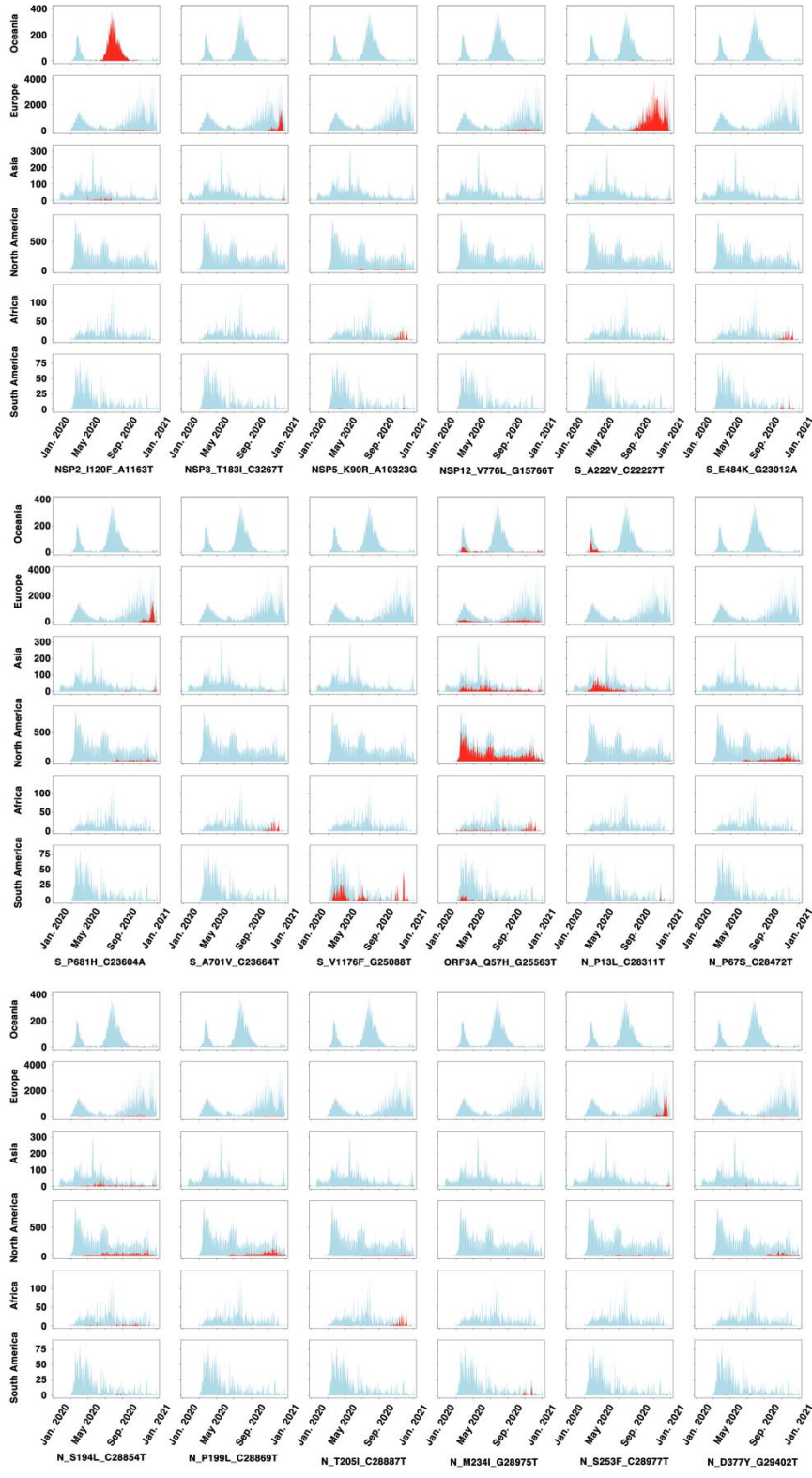
1 **Supplemental Data**

2 Table S1. (A) All detected amino acid mutations across the whole genomes in 3823 rep-
3 resentative SARS-CoV-2 strains. The cases indicated the mutation allele frequency (MAF)
4 of these mutations in the >320,000 sequences collected until 11st January 2021. (B) All
5 of the antibodies used for analysis that related to Figure 2D.

8 Figure S1. Phylogenetic tree with fixed nucleotide mutations (A) and amino acid (B) for
9 representative SARS-CoV-2 strains selected by Nextstrain until 11th January, 2021. Fig-
10 ure S1 is related to Figure 1A.

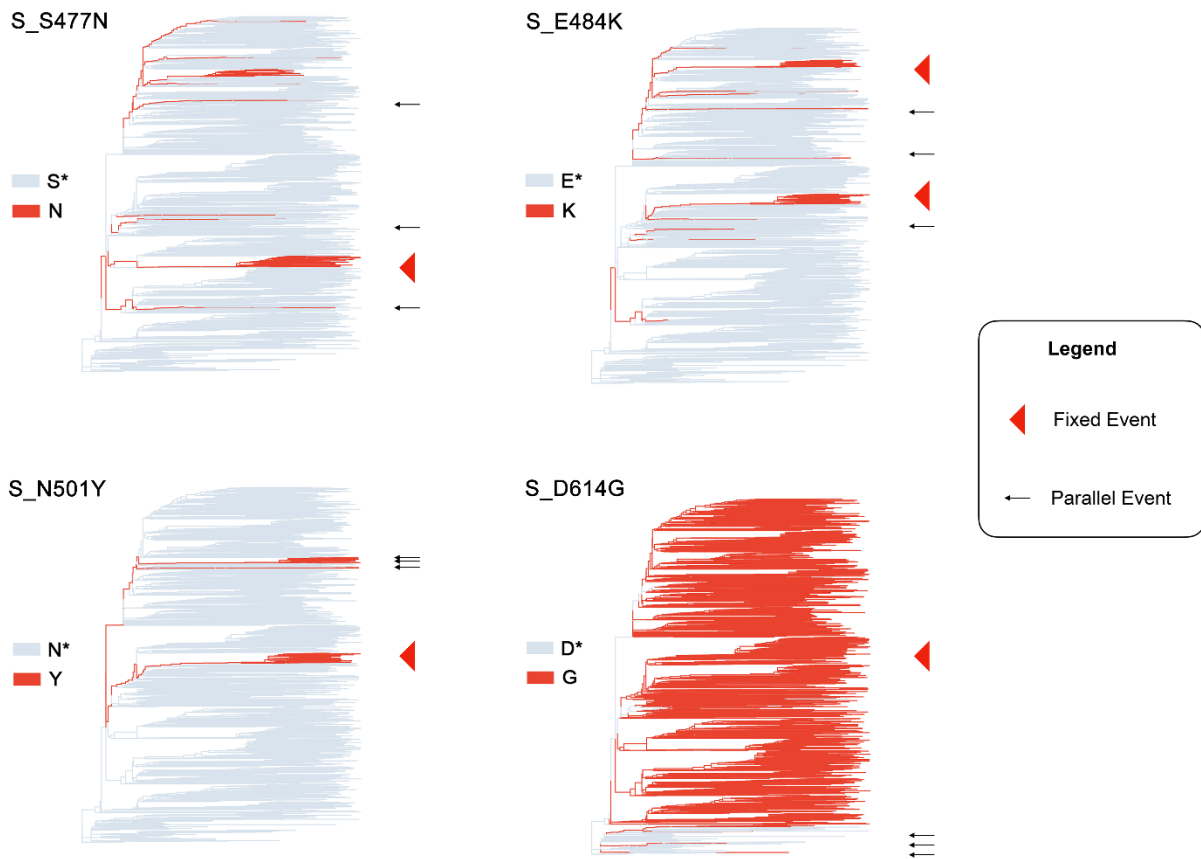
11

12 Supplemental Figure 2



14 Figure S2. The spatiotemporal distribution of collected genomes in GISAID with 18 mu-
15 tations in the NSP2, NSP5, NSP12, S, ORF3A and N proteins. In the histogram for each
16 continent/mutation pairing, the X-axis represents the collection date of the sequenced
17 viruses and the Y-axis represents the number of sequences with the indicated mutation.
18 The number of mutated sequences is shown in red, with the total number of all strains
19 shown in light blue for comparison. Figure S2 is related to Figure 1C.

20 **Supplemental Figure 3**



21

22 Figure S3. Phylogenetic pattern of four amino acid mutations in the RBD region of the S
23 protein. The red triangle indicates the event of a fixed mutation within a sub-clade. The
24 black arrow indicates a recurrent mutation among multiple and independent sub-clades.

25 *Most of the viral strains (>99.8%) while not all of them have the labeled amino acid.

26 Figure S3 is related to Figure 2B.