

Supplementary Information for manuscript

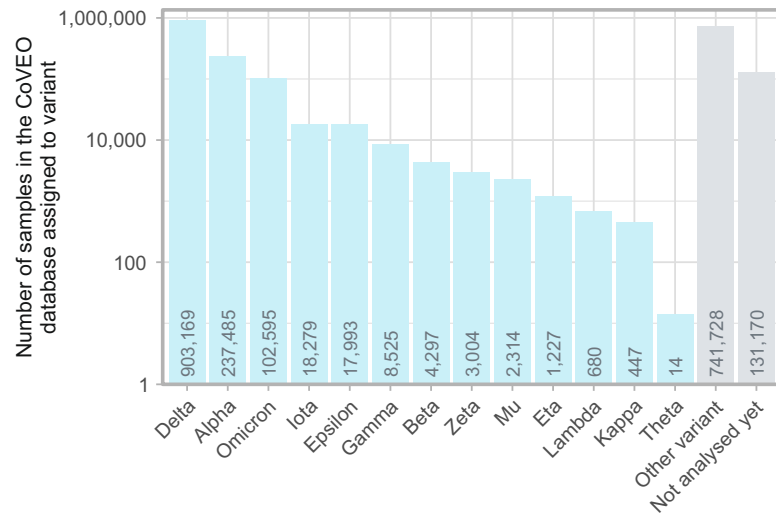
Systematic detection of co-infection and intra-host recombination in more than 2 million global SARS-CoV-2 samples

Orsolya Anna Pipek, Anna Medgyes-Horváth, József Stéger, Krisztián Papp, Dávid Visontai, Marion Koopmans, David Nieuwenhuijse, Bas B. Oude Munnink, VEO Technical Working Group, István Csabai

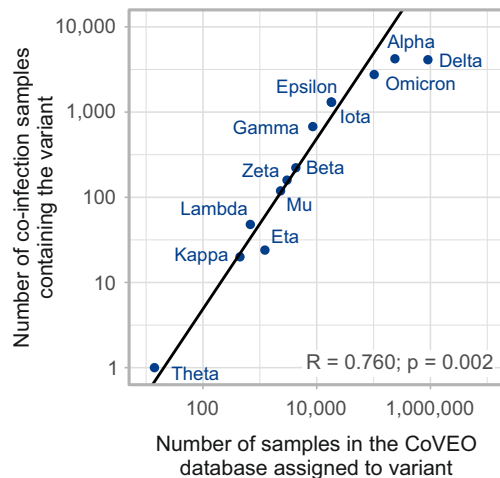
Supplementary Figures 1-3.

Supplementary Figure 1. The number of samples in the CoVEO database. **a.** The number of good-quality SARS-CoV-2 samples with a human host assigned to different variants in the CoVEO database. **b.** The relationship between the total number of samples assigned to a specific variant and the number of co-infection samples containing the variant. The straight black line represents a linear dependence with a slope of 1 on a log-log graph. R represents the Pearson-correlation coefficient.

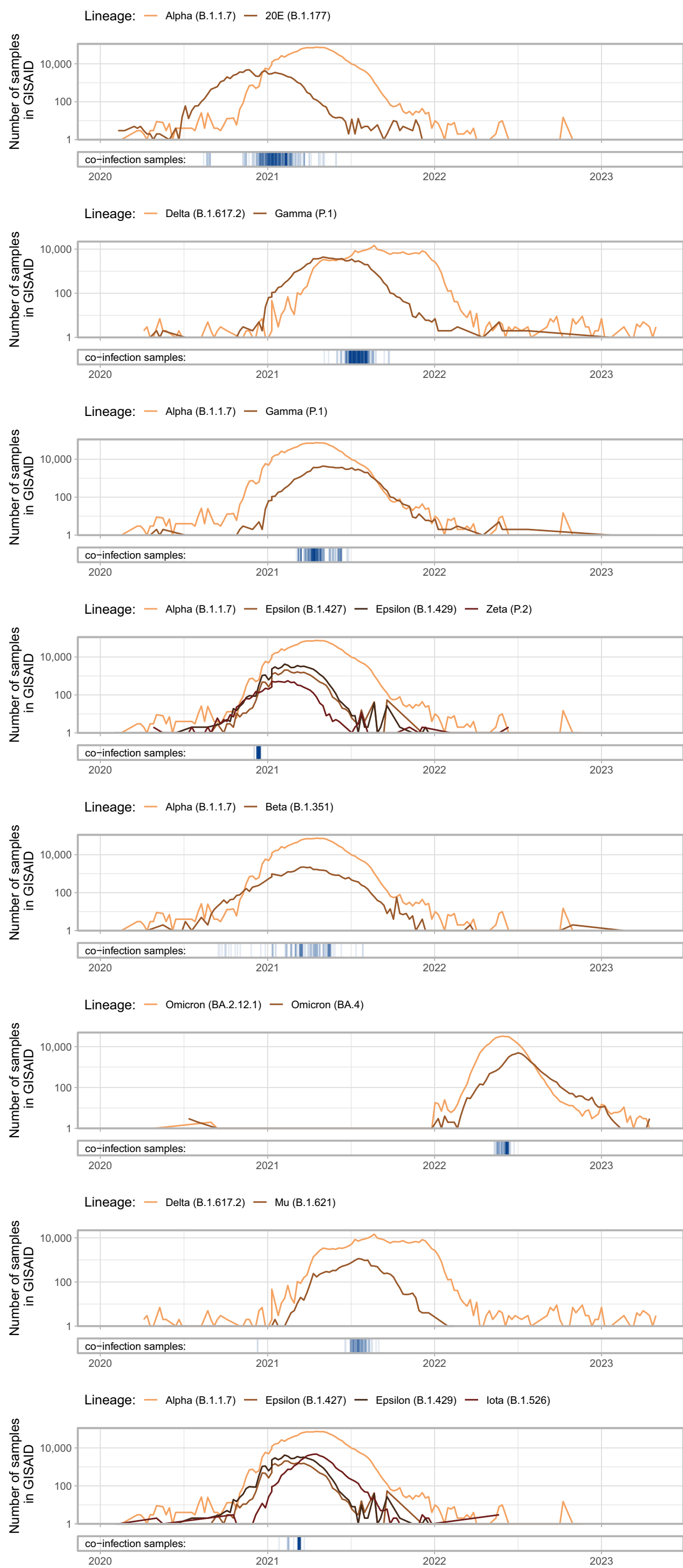
a



b



Supplementary Figure 2. Temporal distribution of co-infection samples for variant combinations with more than 50 samples. (The same figures for the top 4 most abundant combinations are shown in Figure 1C of the main manuscript.) Prevalence curves indicate the number of GISAID samples assigned to the respective variants (binned weekly). Blue vertical lines on the bottom panels mark the collection date of co-infection samples of the given variants.



mutational status