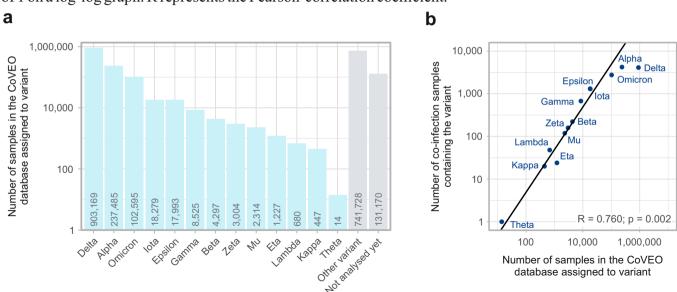
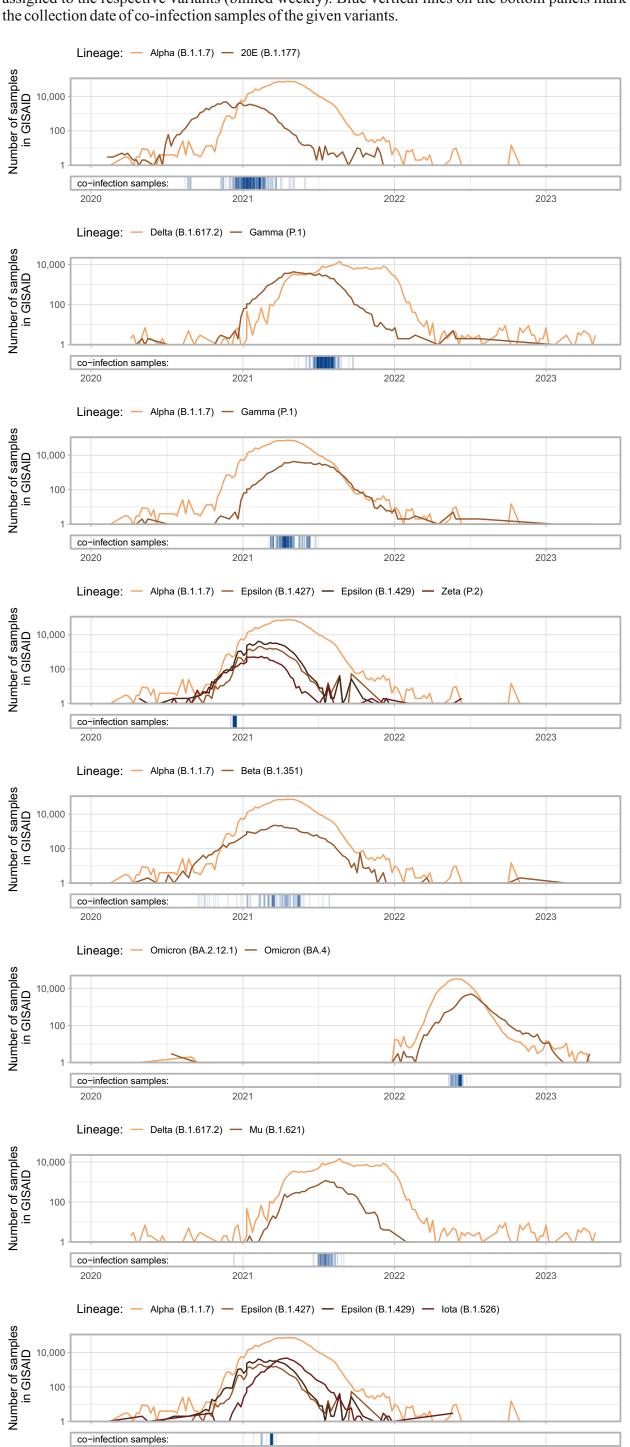
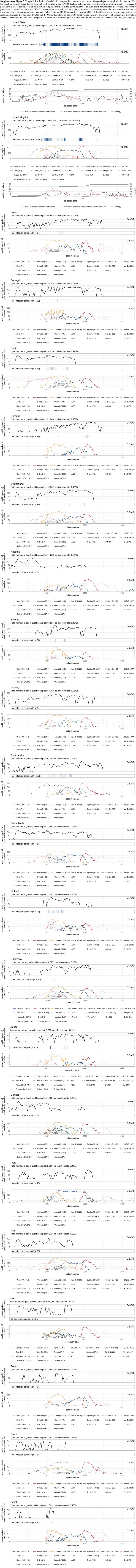
SARS-CoV-2 samples with a human host assigned to different variants in the CoVEO database. **b.** The relationship between the total number of samples assigned to a specific variant and the number of coinfection samples containing the variant. The straight black line represents a linear dependence with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

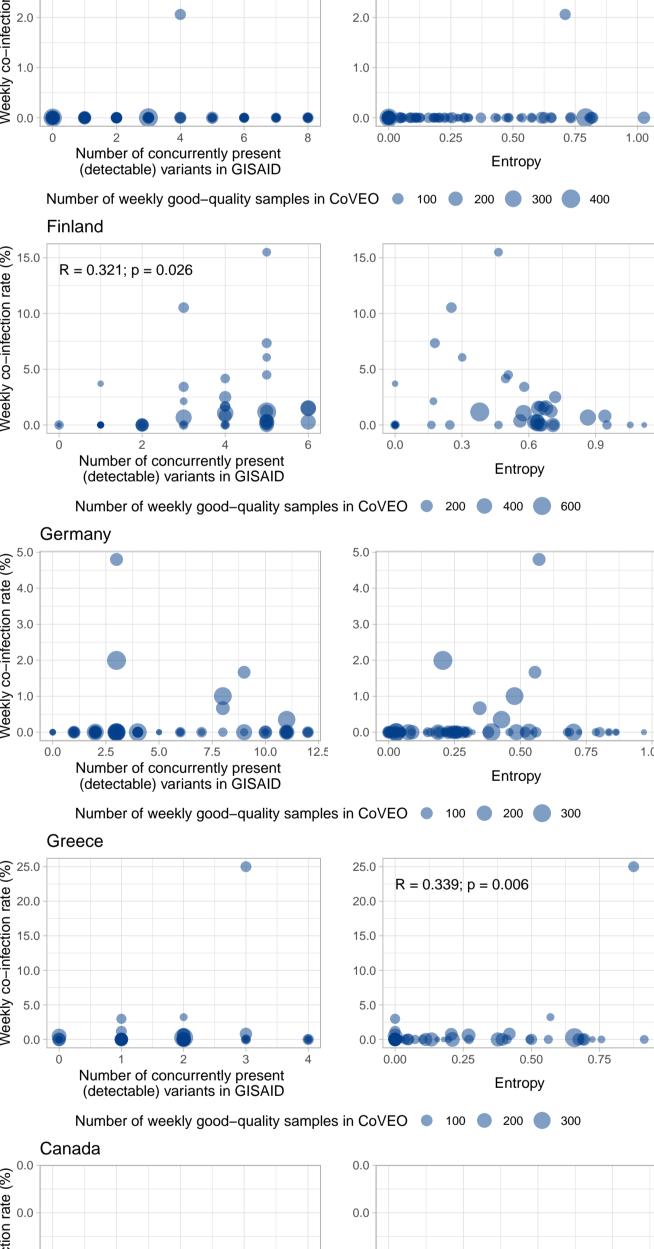
Supplementary Figure 1. The number of samples in the CoVEO database. a. The number of good-quality

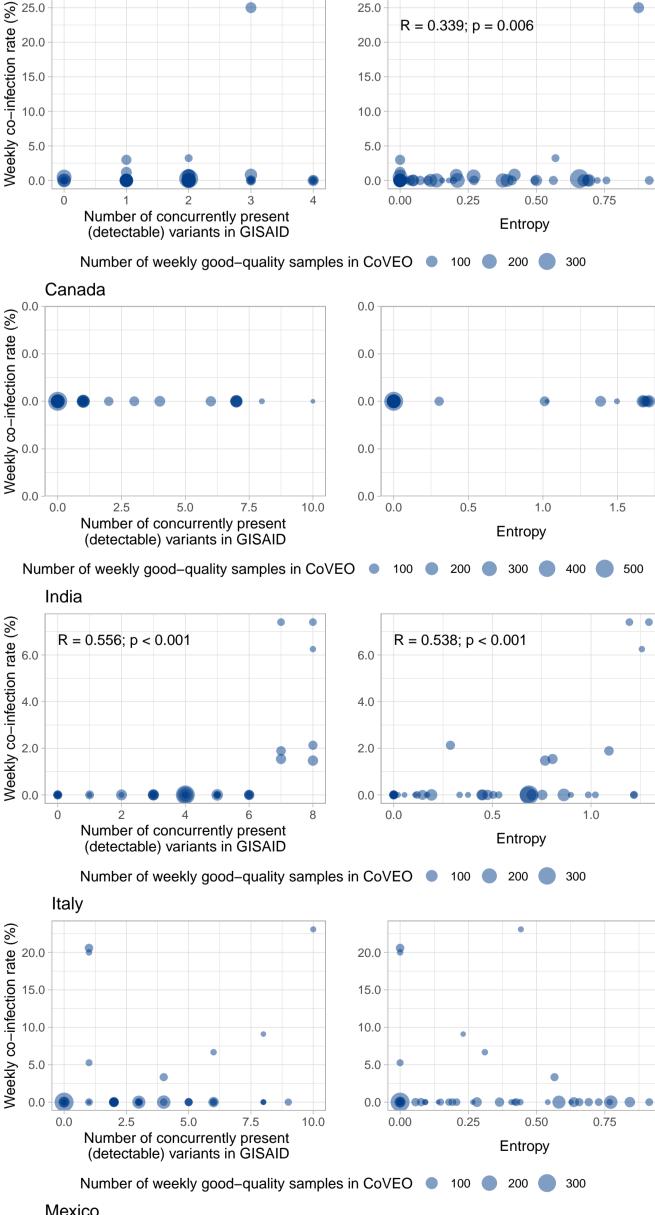


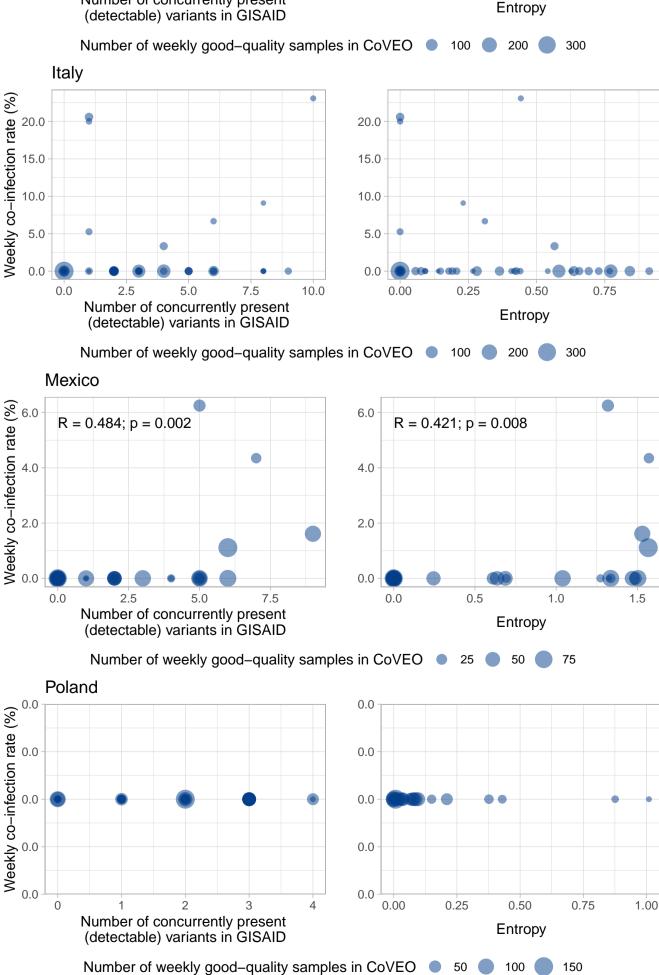
Supplementary Figure 2. Temporal distribution of co-infection samples for variant combinations with more than 50 samples. (The same figures for the top 4 most abundant combinations are shown in Figure 1C of the main manuscript.) Prevalence curves indicate the number of GISAID samples assigned to the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











0.0

0.0

0.0

0.0

0.0

4.0

3.0

2.0

1.0

0.0

0.0

0.1

50

0.2

100

Entropy

0.3

0.4

0.0

0.2

30

0.4

Entropy

0.6

8.0

Brazil

Weekly co-infection rate (%)

0.0

0.0

0.0

0.0

4.0

3.0

2.0

0.0

0

Weekly co-infection rate (%)

Ó

Israel

Number of concurrently present

(detectable) variants in GISAID

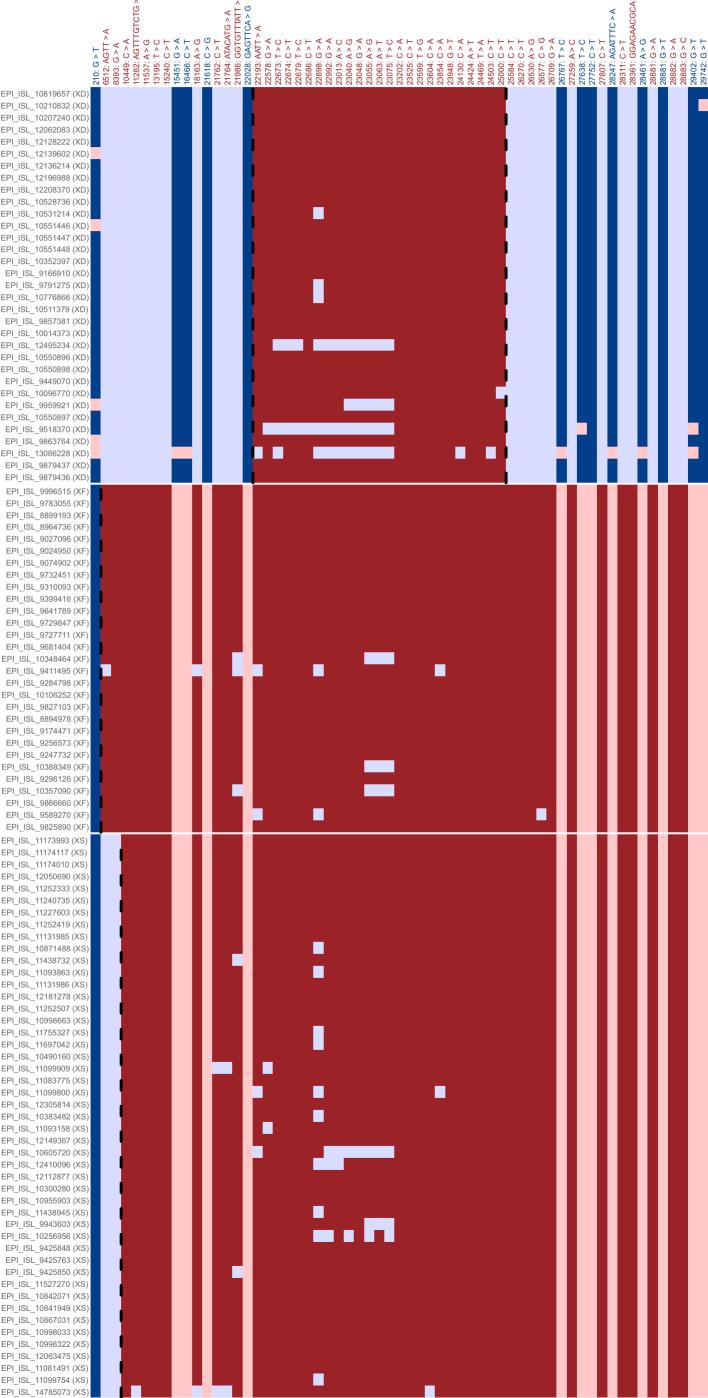
Number of concurrently present

(detectable) variants in GISAID

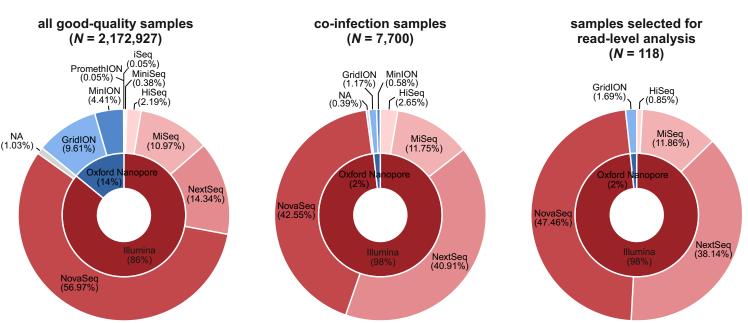
Number of weekly good–quality samples in CoVEO •

Number of weekly good-quality samples in CoVEO

Supplementary Figure 5. Disjunct defining mutations in GISAID samples assigned to the XD, XF and XS 'Deltacron' (recombinant Delta and Omicron) lineages. Samples (rows) are listed by their GISAID ID, with their Pango lineage indicated in brackets. Defining mutations (columns) are colored by their respective variant (blue for Delta and red for Omicron). Heatmap colors show whether the given sample contains the given mutation or not (dark blue: present Delta-mutation; light blue: missing Omicron-mutation; dark red: present Omicronmutation; light red: missing Delta-mutation). Dashed vertical lines mark the recombination breakpoint range(s) for each lineage. (The same genomic regions are indicated with red shaded areas in Figure 4 of the main manuscript.) 10449: C > A 11822: AGTITGTCTG > A 11822: AGTITGTCTG > A 11957: A > G 11957: A > G 11646: C > T 16461: G > A 11646: C > T 16461: C > T 1764: C > T 1764: ATTACATG > A 21986: GGTGTTTATT > G 22028: GAGTTCA > G 22193: AATT > A 22193: AATT > A 22673: T > C 22674: C > T 22674: C > T 22674: C > T 22674: C > T 22674: C > A 22673: T > C 22674: C > A 22675: T > C 23065: A > T 23065: A > T 23065: A > T 23075: T > C 23069: C > A 23650: C > T 23650: C > A 23650: C > T 23650: C > T 25600: C > T 25600: C > T 26500: C > T 26500: C > T 26500: C > T 2650: C > T 2759: A > C 2759: A > C 2750: C > T EPI_ISL 10819657 (XD) EPI_ISL_10210832 (XD) EPI ISL 10207240 (XD) EPI_ISL_12062083 (XD) EPI ISL_12128222 (XD) EPI_ISL_12139602 (XD) EPI_ISL_12136214 (XD) EPI_ISL_12196988 (XD) EPI_ISL_12208370 (XD) EPI_ISL_10528736 (XD) EPI ISL_10531214 (XD) EPI_ISL_10551446 (XD) EPI_ISL 10551447 (XD)



Supplementary Figure 6. Distribution of sequencing platforms and instruments within all good-quality samples included in the study, for co-infection samples and for samples selected for downstream read-level analysis.

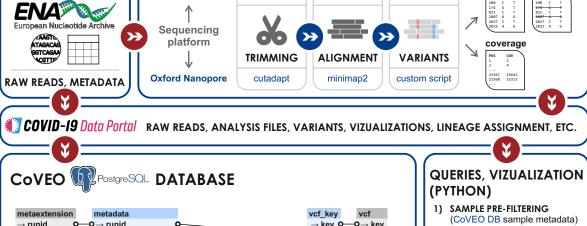


Data Portal (https://www.covid19dataportal.org/) are shown that are relevant to the current analysis. More details on both can be found in Rahman et al. (2023). For the CoVEO database, only those tables and fields are displayed that were queried during data processing. Queries, codes, data files and vizualizations are uploaded to the csabaiBio/SARSCoV2-coinf github repository.

UNIFIED VARIANT CALLING WORKFLOW

| UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CALLING WORKFLOW | UNIFIED VARIANT CAL

Supplementary Figure 7. Schematic diagram of the workflow used to produce the data analysed in the study. Only those steps of the VEO variant calling pipeline (github.com/enasequence/covid-sequence-analysis-workflow) and features of the COVID-19



→ runid O → runid → key O-O → kev description sample accession Orunid dp 2) CO-INFECTION SAMPLES first created experiment accession pos (CoVEO DB mutations) base_count study_accession ref collection date INTRA-HOST RECOMBINATION country n content collection date valid (CoVEO DB mutations (af) & \rightarrow id $b \rightarrow \text{runid}$ ocountry id **VARIANTS** raw reads from the COVID-19 country name estimated n content host id Data Portal) instrument id host instrument → id0 bi ← Ø csabaiBio/SARSCoV2-coinf host SAMPLE METADATA instrument platform instrument model