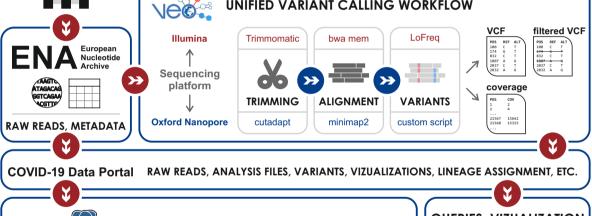
(github.com/enasequence/covid-sequence-analysis-workflow) and features of the COVID-19 Data Portal (https://www.covid19dataportal.org/) are shown that are relevant to the current analysis. More details on both can be found in Rahman et al. (2023). For the CoVEO database, only those tables and fields are displayed that were queried during data processing. Queries, codes, data files and vizualizations are uploaded to the csabaiBio/SARSCoV2-coinf github repository. Postgres, PostgreSQL and the Slonik Logo are trademarks or registered trademarks of the PostgreSQL Community Association of Canada and used with their permission.

**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



QUERIES, VIZUALIZATION COVEO PostgreSQL DATABASE (PYTHON) 1) SAMPLE PRE-FILTERING vcf key metaextension metadata vcf (CoVEO DB sample metadata) O → runid → runid → kev O-O → kev description sample accession runid 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  $\rightarrow$  runid ocountry id **VARIANTS** raw reads from the COVID-19 estimated n content country name host id Data Portal) instrument id host instrument  $\rightarrow$  id0 bi ← Ø SAMPLE METADATA csabaiBio/SARSCoV2-coinf host instrument platform instrument model