

Supplemental information for: Longitudinal sequencing and variant detection of SARS-CoV-2
across Southern California wastewater from April 2020 – August 2021.

Jason A. Rothman^{a#}, Andrew Saghir^a, Amity G. Zimmer-Faust^b, Kylie Langlois^b, Joshua A.
Steele^b, John F. Griffith^b, Katrine L. Whiteson^{a#}

^a Department of Molecular Biology and Biochemistry, University of California, Irvine, Irvine,
CA, USA.

^b Southern California Coastal Water Research Project, Costa Mesa, CA, USA

Co-corresponding authors: Jason A. Rothman, University of California, Irvine, Irvine, CA,
92697, (949) 824-3509, rothmanj@uci.edu. Katrine L. Whiteson, University of California,
Irvine, Irvine, CA, 92697, (949) 824-9032, katrine@uci.edu.

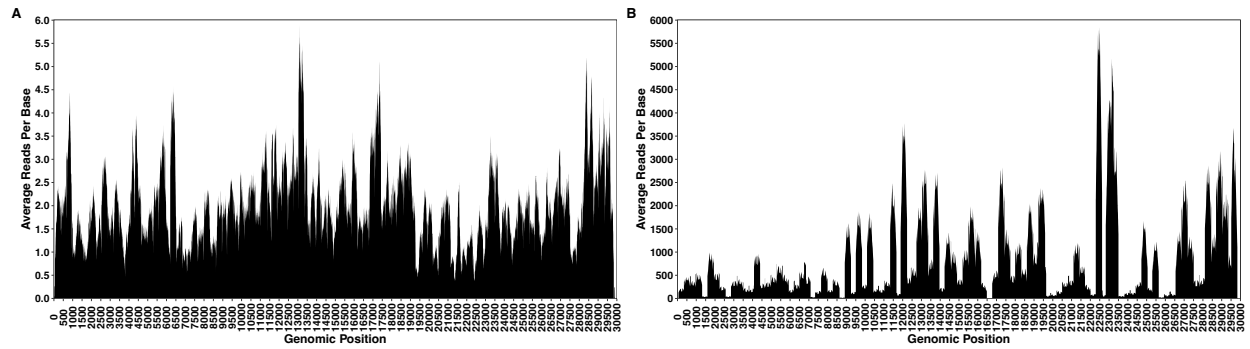
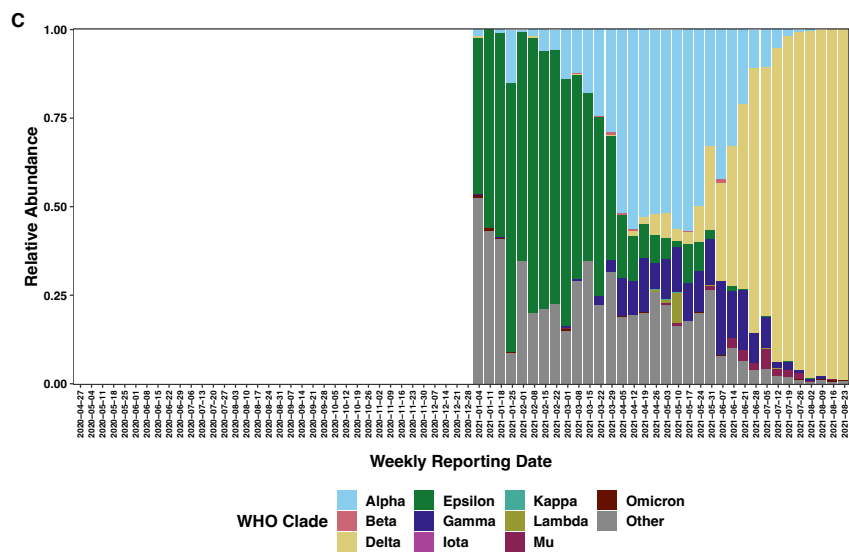
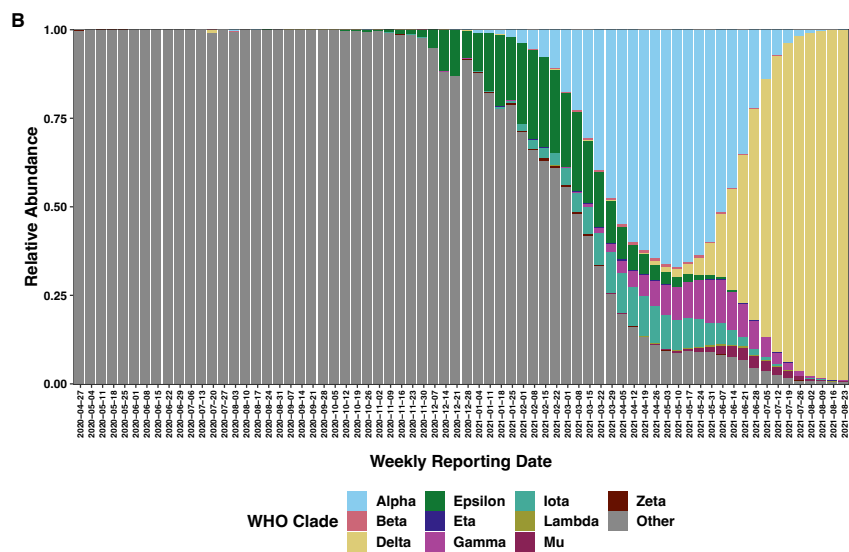
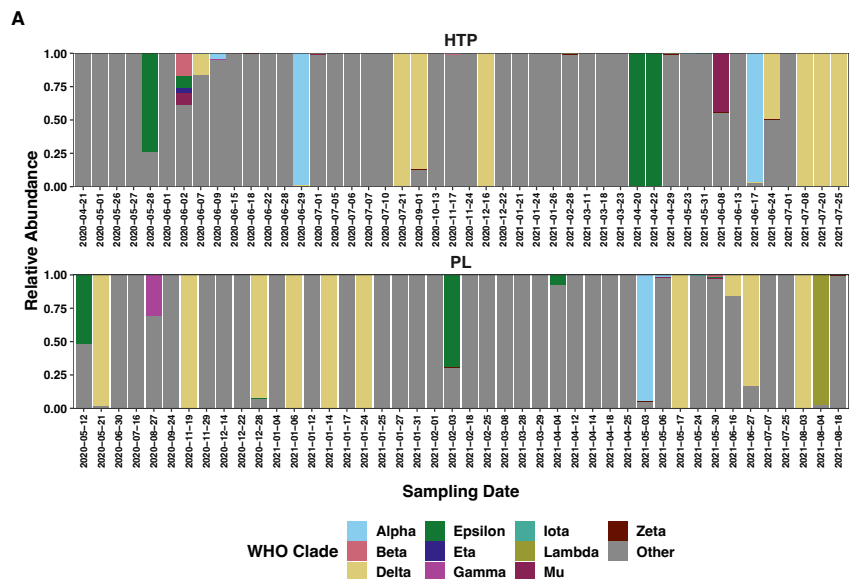


Figure S1: Area plot of the average reads per base mapped to the SARS-CoV-2 genome across all samples. Panel A represents libraries prepared with the Illumina Respiratory Virus Enrichment Panel (IRV) and Panel B represents libraries prepared for tiled amplicon sequencing.



21 Figure S2: The relative proportional abundance of the ten most abundant SARS-CoV-2 lineages
22 plus others in A) tiled-amplicon libraries, B) weekly GISAID-reported data for clinical samples
23 from the United States, and C) CHSS-reported clinical data. Panel A is faceted by WTP and
24 labeled with sampling date and panel B is labeled by the aggregate of GISAID-reported data.
25 Note that one sample date from the North City Water Reclamation Plant is not shown, and
26 variant data from CHSS are not available before 1/1/21 so those dates are left blank.
27

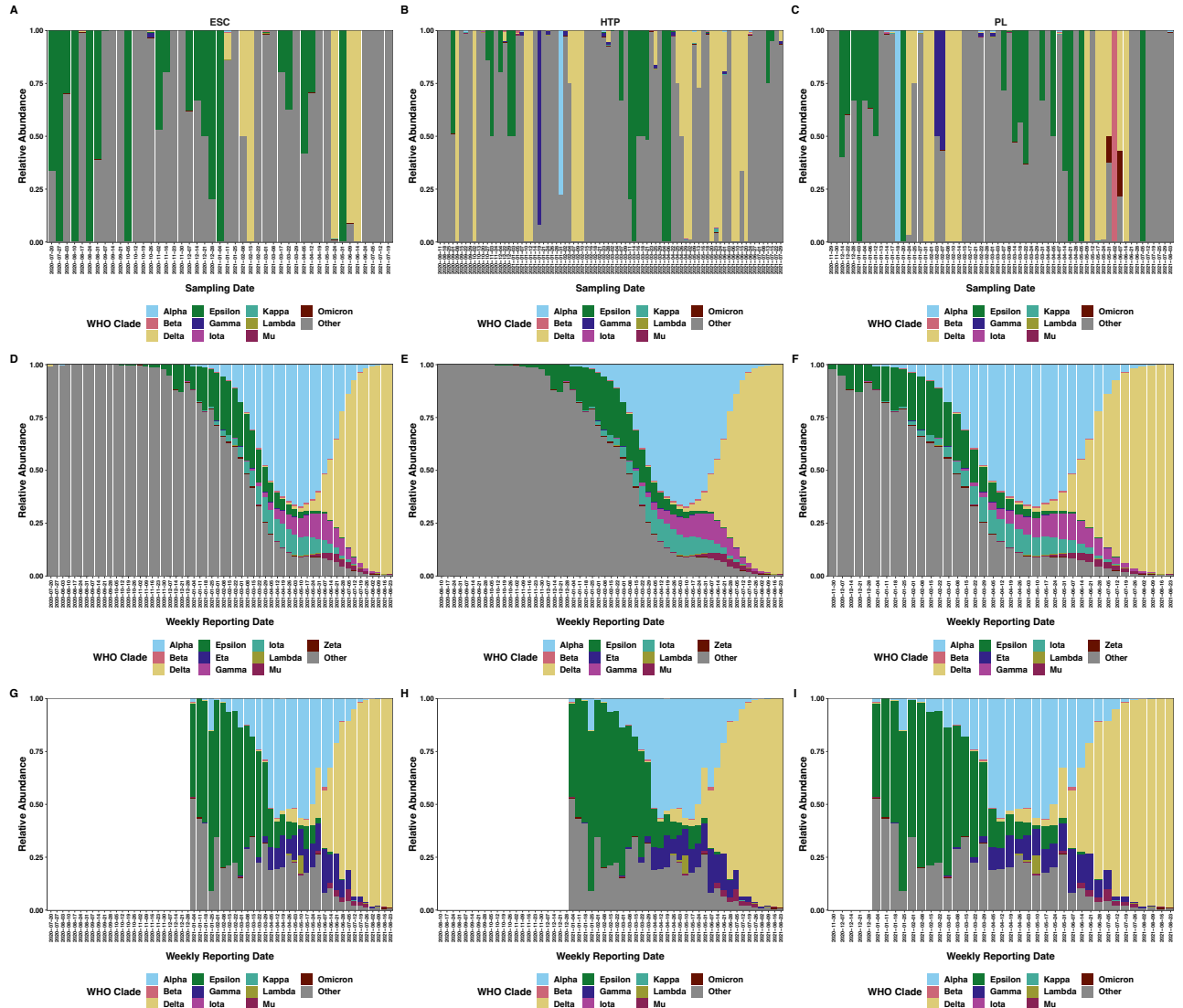


Figure S3: The relative proportional abundance of the ten most abundant SARS-CoV-2 lineages plus others in Illumina Respiratory Virus Panel enriched libraries for A) Escondido, B) Hyperion, and C) Point Loma water treatment facilities. Panels D, E, and F are weekly GISAID-reported data for clinical samples from the United States corresponding to the approximate time periods of panels A, B, and C, respectively. Panels G, H, and I are CHSS-reported clinical data corresponding to the approximate time periods of panels A, B, and C, respectively. Variant data from CHSS are not available before 1/1/21 so those dates are left blank.

37 Supplemental File SF1: This file contains several worksheets of information: A README sheet
38 with detailed information, sample metadata, single nucleotide variants for IRV and tiled
39 amplicon samples, and PANGO sublineages as output by Freyja for IRV and tiled amplicon
40 samples.