iScience, Volume *24*

Supplemental information

A benchmarking study of SARS-CoV-2 whole-genome sequencing protocols using COVID-19 patient samples

Tiantian Liu, Zhong Chen, Wanqiu Chen, Xin Chen, Maryam Hosseini, Zhaowei Yang, Jing Li, Diana Ho, David Turay, Ciprian P. Gheorghe, Wendell Jones, and Charles Wang

Supplementary Figures

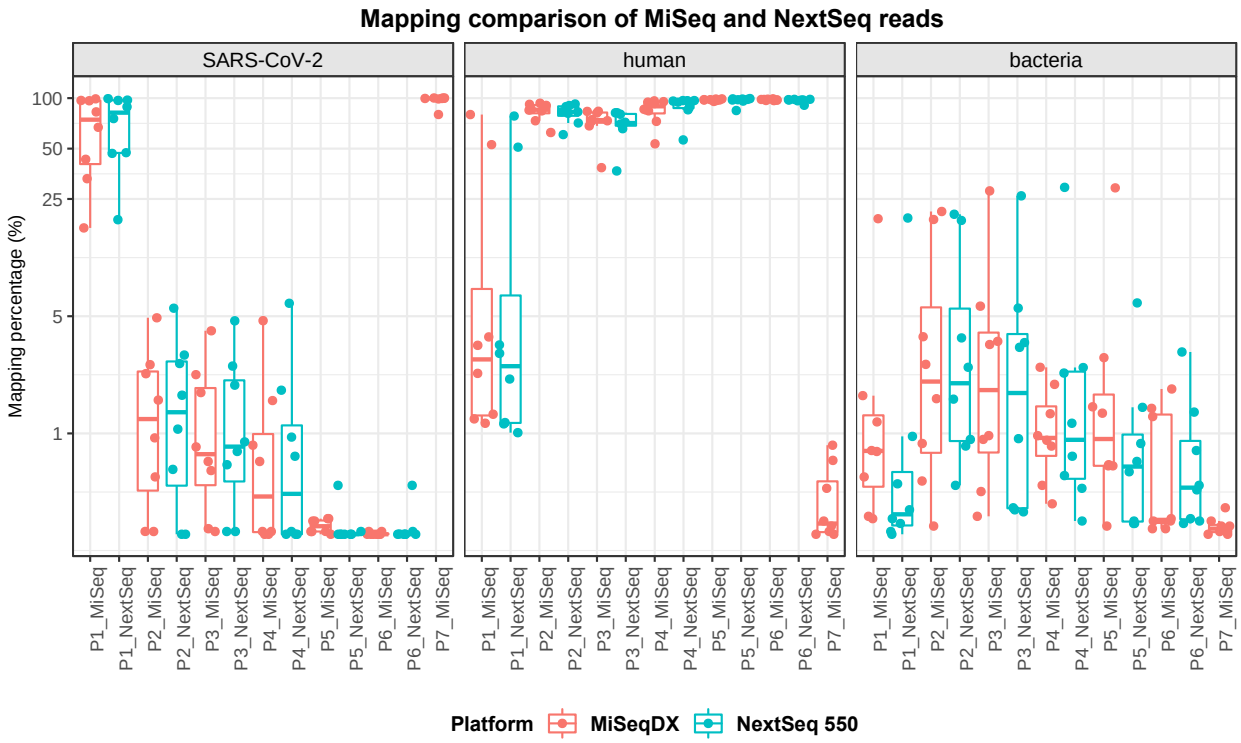


Figure S1. Comparison of mappability of sequencing data between MiSeqDx and NextSeq 550, related to Figure 1 and 2. The SARS-CoV-2 whole-genome sequencing (WGS) libraries constructed from seven protocols were sequenced on an Illumina MiSeqDx at 300x2 bp or 150x2 bp, pair-end (PE), or a NextSeq 550 at 150x2 bp, PE. Mapping percentage of sequencing reads to the SARS-CoV-2, human, and bacterial genomes are presented. Each dot represents one library. Libraries constructed from P7 protocol were only sequenced on MiSeqDx, 300x2 bp, PE.

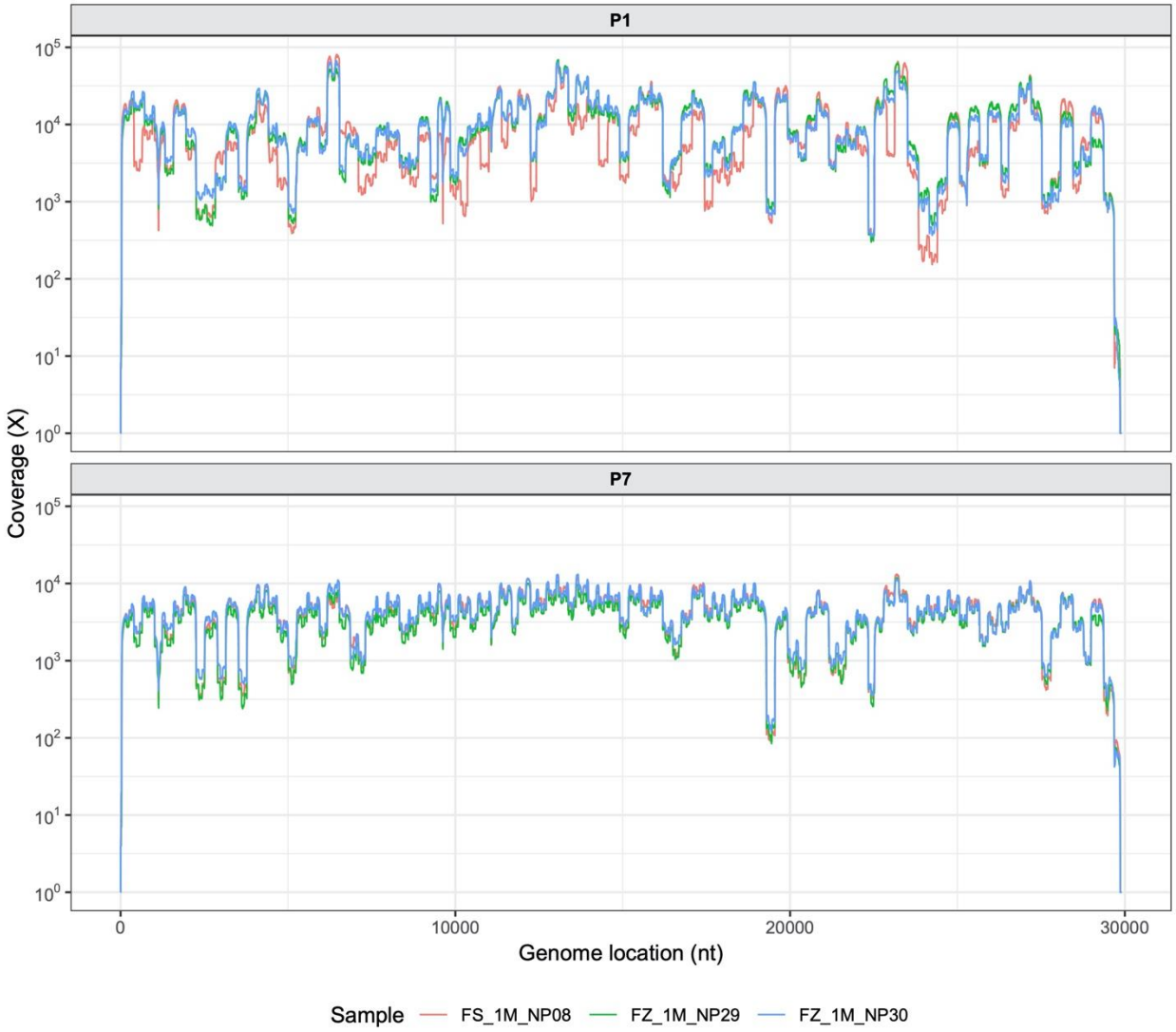


Figure S2. Improved genome coverage in P7 compared to P1 at high viral input, related to Figure 4. Genome coverage across SARS-CoV-2 genome for three clinical samples using 1 million (1M) viral input at 1M sequencing depth was plotted. The upper panel shows the uneven genome coverage of the Qiagen Primer Panel V1 ARTIC V3-based target whole-genome amplification technology (P1). The lower panel shows the improved uniformity on genome coverage of the modified Qiagen Primer Panel Primer Panel V2 ARTIC V3-based target whole-genome amplification technology (P7). FS (fresh): RNA isolated from fresh samples; FZ (frozen): RNA isolated from frozen samples. FS_1M_NP08: fresh RNA sample with 1M viral input for sample NP08; FZ_1M_NP29: RNA from frozen sample with 1M viral input for sample NP29; FZ_1M_NP30, RNA from frozen sample with 1M viral input for sample NP30.

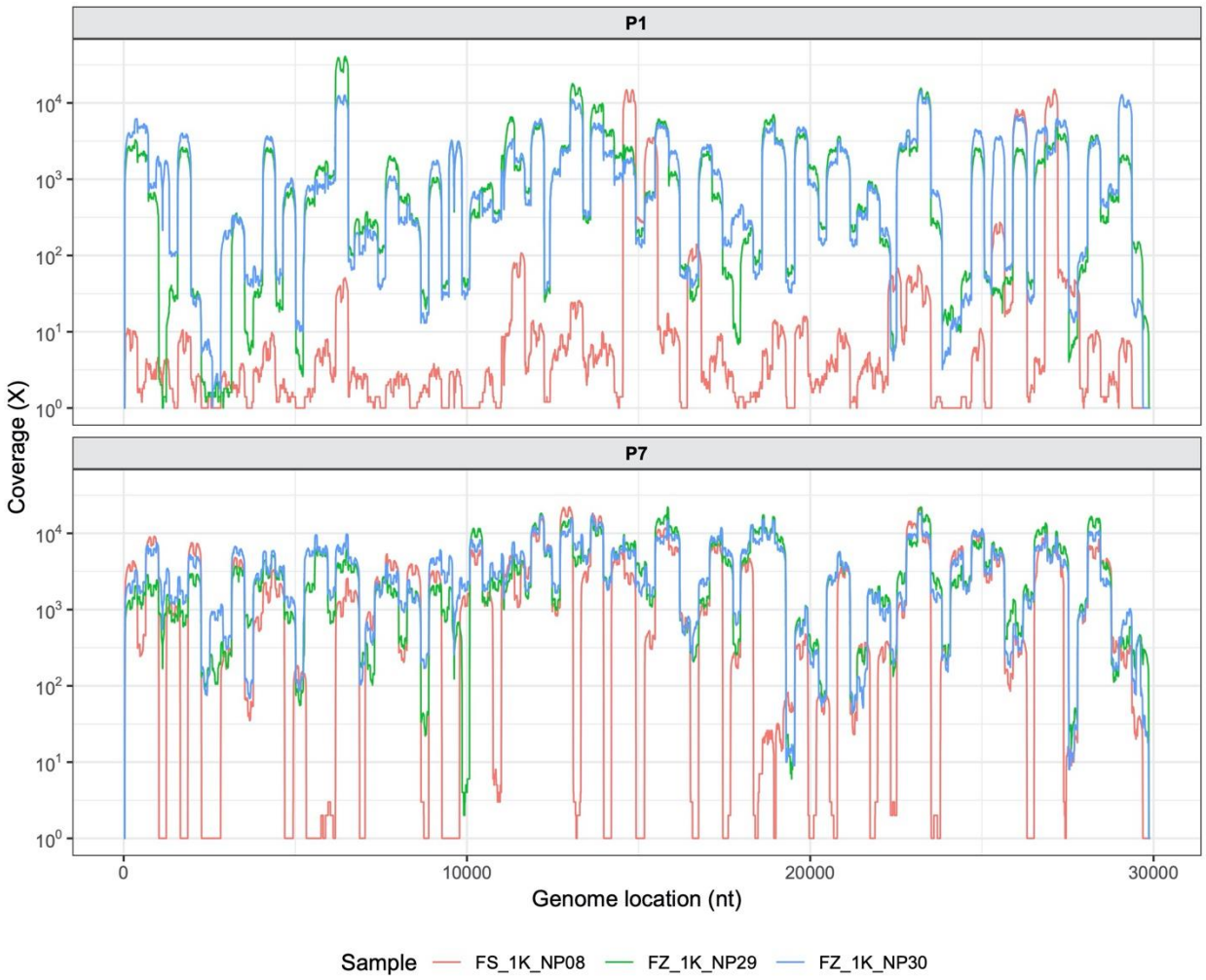


Figure S3. Improved genome coverage in P7 compared to P1 at low viral input, related to Figure 4. Genome coverage across SARS-CoV-2 genome for three clinical samples using 1000 (1K) viral input at 1M sequencing depth was plotted. The upper panel shows the uneven genome coverage of the Qiagen Primer Panel V1 ARTIC V3-based target whole-genome amplification technology (P1). The lower panel shows the improved uniformity on genome coverage of the modified Qiagen Primer Panel Primer Panel V2 ARTIC V3-based target whole-genome amplification technology (P7). FS (fresh): RNA isolated from fresh samples; FZ (frozen): RNA isolated from frozen samples.

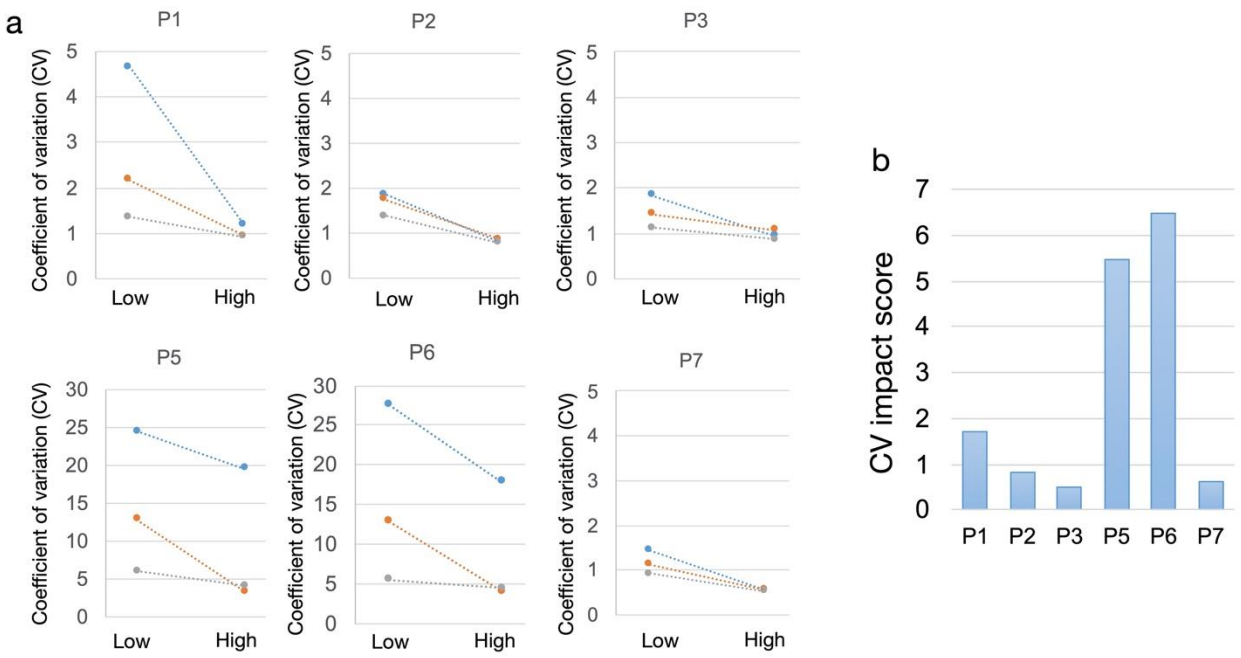


Figure S4. The impact of viral input on coverage evenness of SARS-CoV-2 genome, related to Figure 4. (a) The coefficients of variation (CV) of SARS-CoV-2 genome coverage of three clinical samples (NP08: blue; NP29: orange; and NP30: gray) were shown at low (1000 viral copies) and high (1 million viral copies) viral input for each protocol, respectively. CV metric was computed using the standard deviation (SD) and mean (M) of the coverage at each reference genome position. In each protocol, a trendline was drawn between the CV values of a sample's low and high viral inputs. The absolute value of the trendline' slope, i.e. the magnitude of change from low to high input, was used to evaluate the impact of viral input on CV in each sample, respectively. **(b)** CV impact score, calculated based on average of the absolute value of the slopes from each protocol in panel (**a**) to examine the overall input impact on CV for each protocol. Higher score indicates that a protocol is more sensitive to viral input in term of CV.

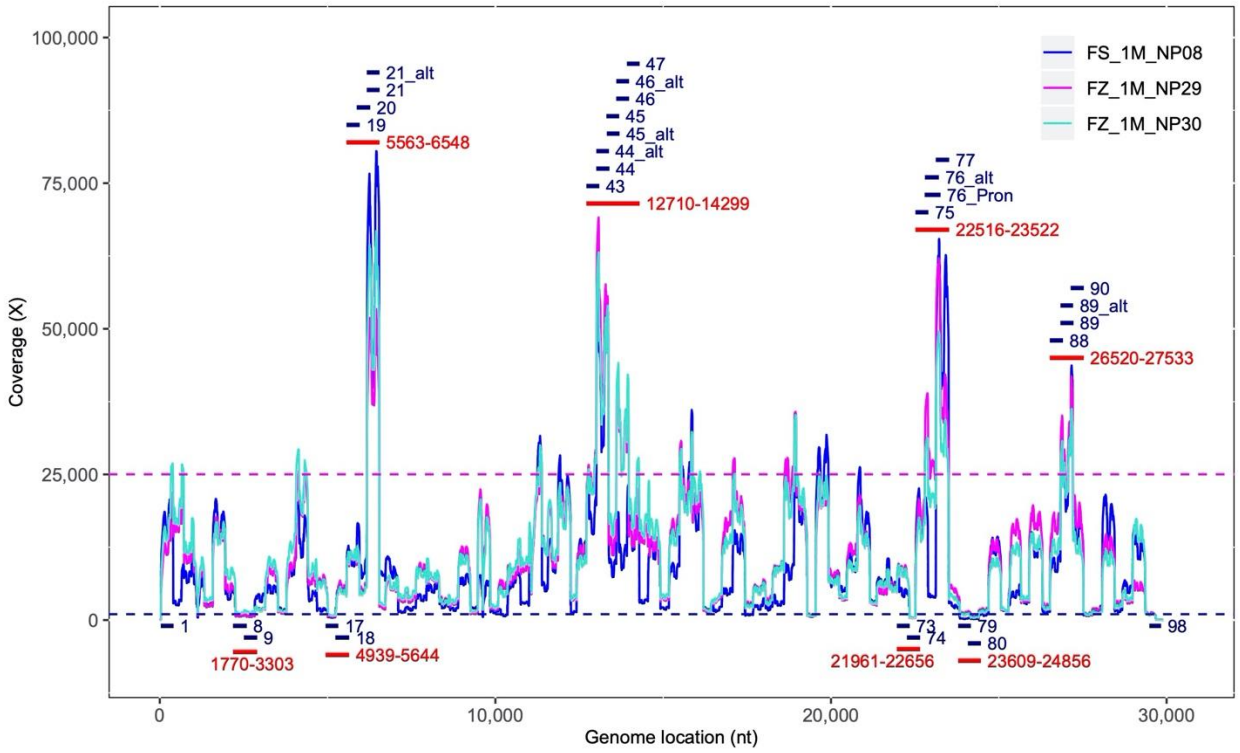


Figure S5. Uneven spiking coverage observed in Qiagen primer panel V1 ARTIC V3-based protocol P1, related to Figure 4. Genome coverage of three clinical samples using 1million viral copies input at 5M read depth were overlapped to show the unevenly covered regions. The ARTIC V3 primers associated with these spiking regions are shown by specific locations. The magenta color dash line indicates 25,000X coverage level, while the blue dash line indicates 1,000X coverage level. The blue segment line shows the position of each amplicon drawn on the real size scale relative to the viral genome size, and the corresponding primer pairs are shown on the right. The red segment line shows the overall region covered by the amplicons listed above. FS: RNA isolated from fresh samples; FZ: RNA isolated from frozen samples.

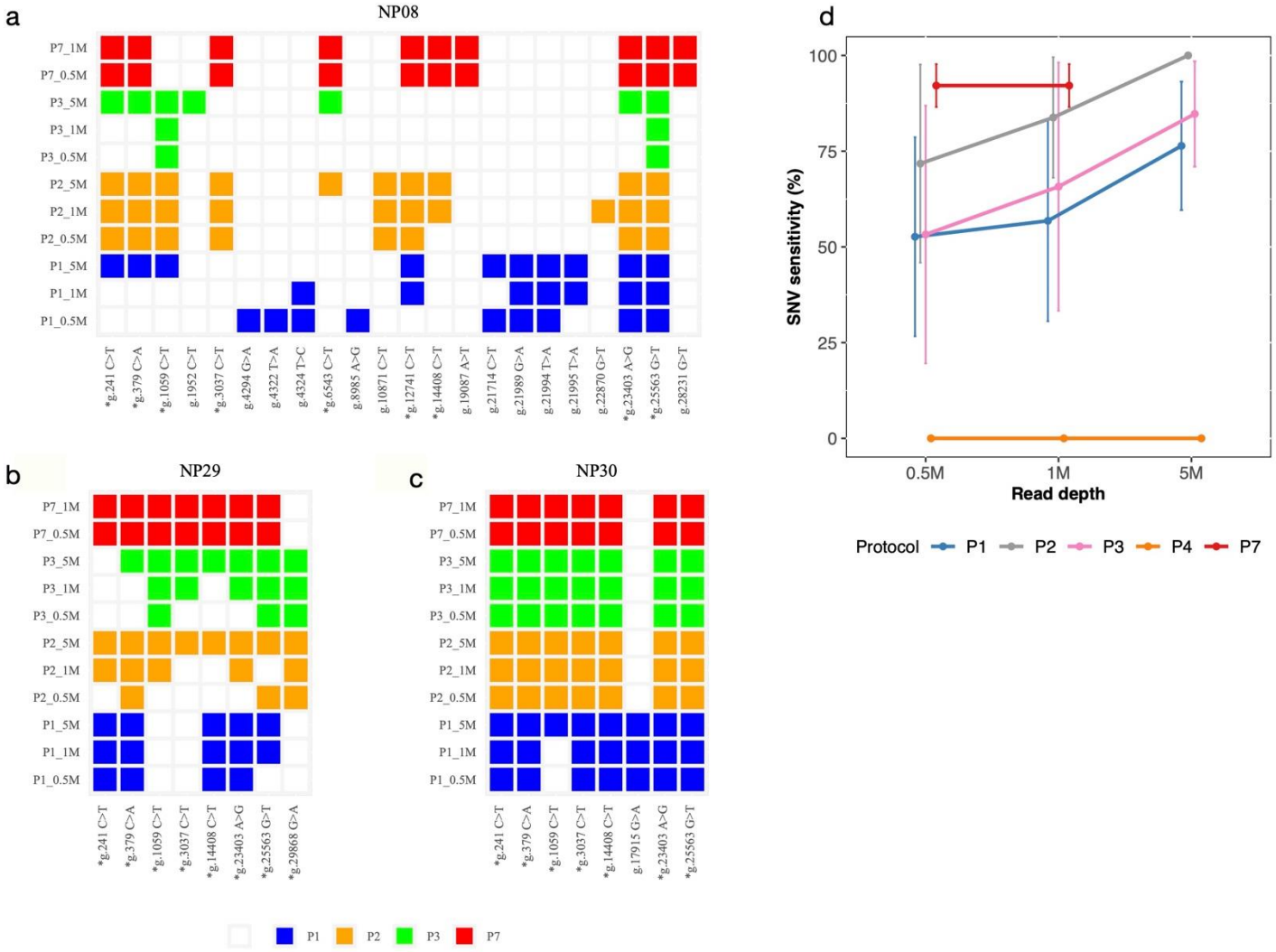
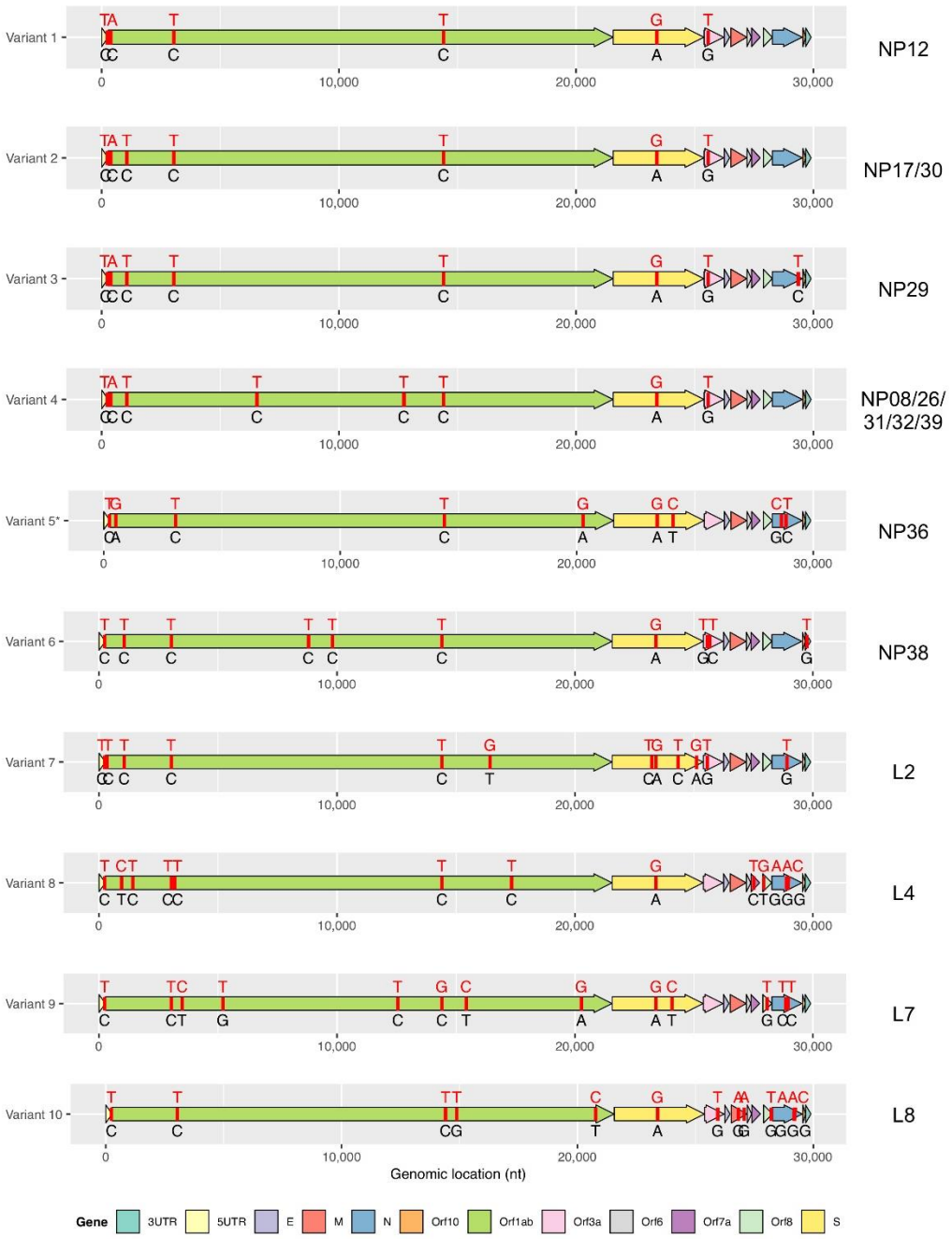


Figure S6. SARS-CoV-2 SNV calling across P1, P2, P3 and P7 protocols at low viral input, related to Figure 5. The SNVs were called at low viral input (1K) at different sequencing read depths in P1, P2, P3, and P7 in sample NP08 (**a**); sample NP29 (**b**); and sample NP30 (**c**). All SNVs with frequency higher or equal to 80% were deemed as true SNVs; and showed in rectangle plot. X-axis indicates the SNV site; Y-axis indicates the protocol at various sequencing depth. (**d**) Influence of sequencing depth on SNV calling sensitivity at low viral input. Sensitivity was measured by sample using the consensus SNVs at low viral input. SNVs were called by five protocols (P1, P2, P3, P4 and P7) at low (1K) viral input at three different sequencing read depths (0.5M, 1M and 5M). For P4 protocol, libraries with low viral input failed to call any SNV due to low genome coverage.



* There is a 9-nucleotides deletion at position 500-508.

Figure S7. SARS-CoV-2 genome variants identified, related to Figure 5. Ten virus genome subtypes were identified in this study. As a note, there were five clinical samples which did not yield enough genome coverage for SNV calling. The true consensus SNVs detected in at least two protocols are depicted on the NC_045512.2 reference genome. The samples are listed from top to bottom by the order of number of SNVs detected. Black letters (nucleotides A, C, G) represent reference sequences at the indicated locations, red letters (nucleotides T, A, G) represent SNVs detected at the genomic locations.

Figure S8. Relationship of SARS-CoV-2 variants identified in this study versus other strains in the world and the United States, related to Figure 5. (a) Phylogenetic tree analysis of ten SARS-CoV-2 genome variants identified in Loma Linda University (LLU_Var1-10) in this study (NP08, NP12, NP17, NP29, NP30, NP36, NP38, L2, L4, L7, and L8 samples) vs. 417 public SARS-CoV-2 genome sequence data from different countries; Color indicates the sample geographic location. Sampling date is shown below the tree. * The variants 1, 2, and 4 were not shown in the phylogenetic tree due to a too similarity with other variants. (**b**) Phylogenetic relationship of designated virus clades for ten SARS-CoV-2 variants obtained in this study vs. 417 publicly available SARS-CoV-2 genome variants.

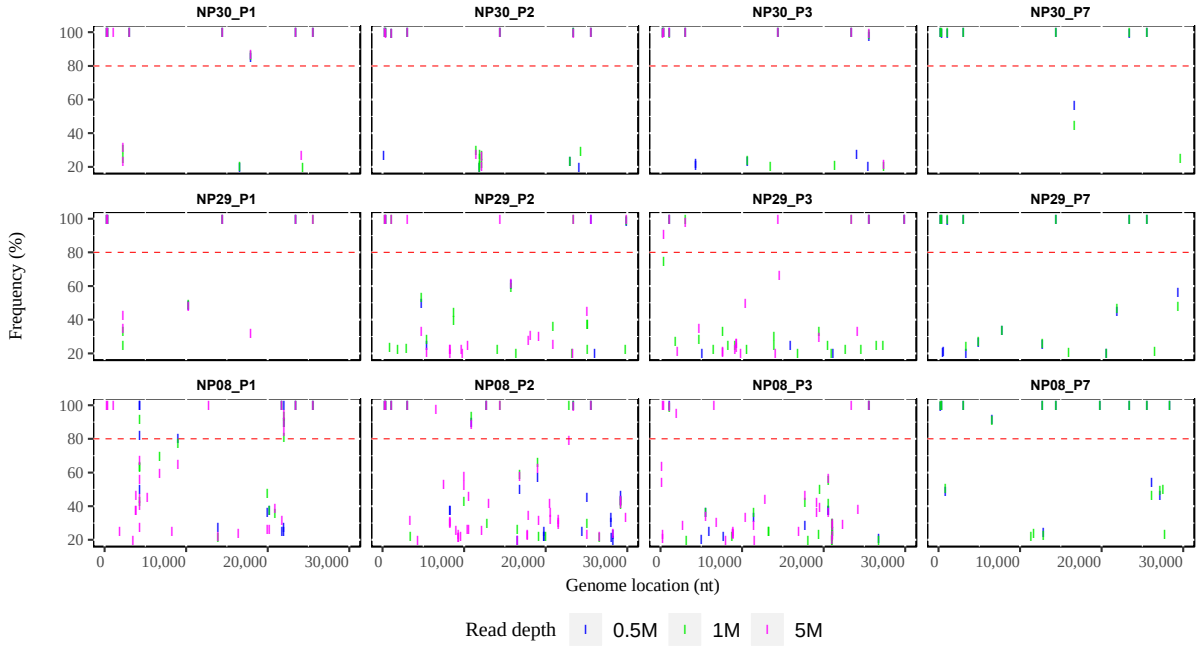


Figure S9. Low viral input had limited capability to make SNV calling, related to Figure 5. The SNVs were called from P1, P2, P3, P4 and P7 protocols from 1000 viral input at 0.5M, 1M, and 5M read depths, respectively. A minimum coverage of 10X and 20% frequency were used as the calling thresholds. SNVs with frequency higher or equal to 80% (red dashed line) were defined as true SNVs. No SNV was called from all three samples prepared with P4 protocol using 1K viral input due to low genome coverage. The SARS-CoV-2 genome NC_045512.2 was used as reference.

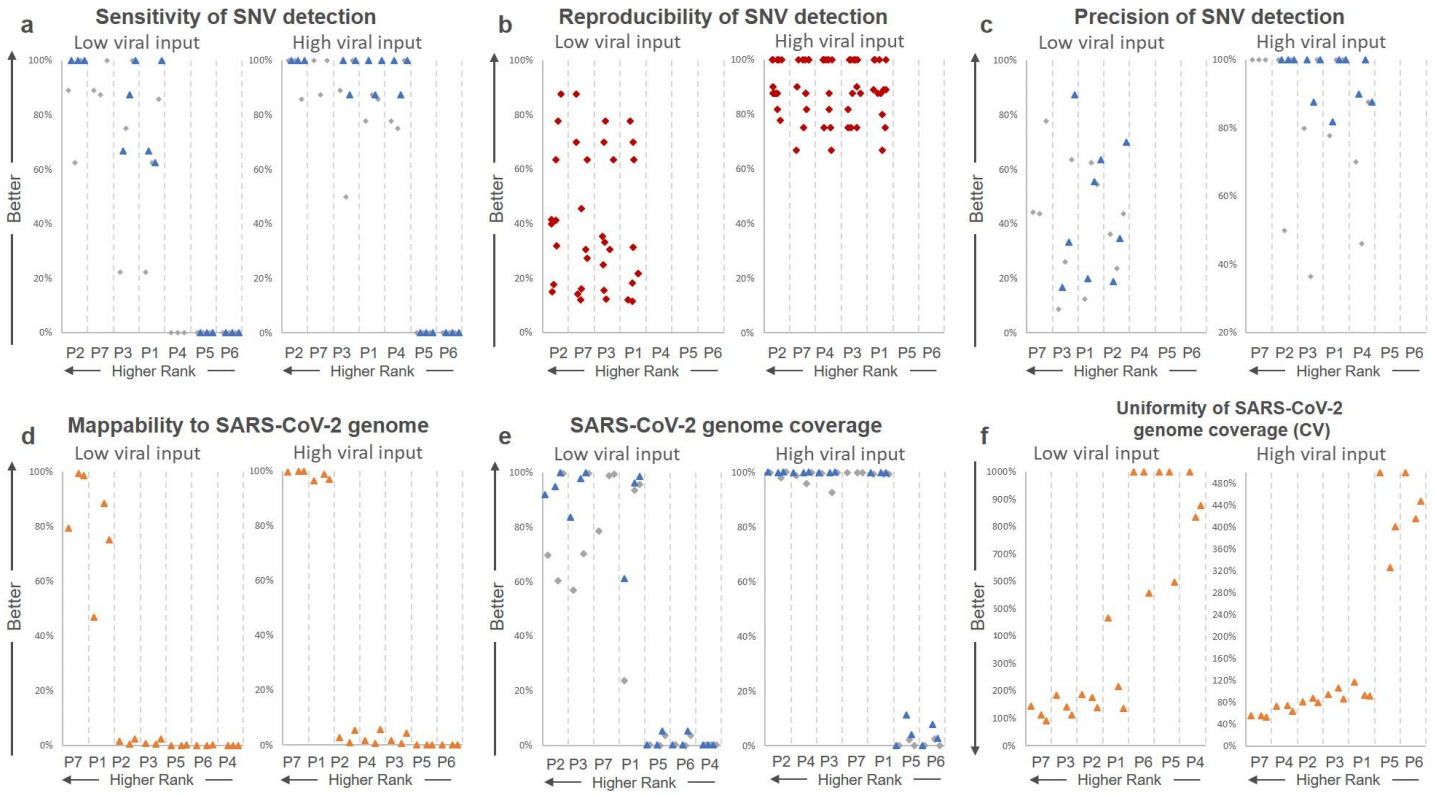


Figure S10. Performance observations of SARS-CoV-2 whole-genome sequencing protocols, related to Figure 8. (**a)** Sensitivity was measured by sample using the consensus SNVs at low and high input levels. Grey diamonds are results at 1M PE reads and blue triangles are results at 5M PE reads. (**b)** Reproducibility was measured between protocols at low and high input levels indicating whether the SNV calls using a VAF threshold of 80% were reproduced in a different protocol. There was one value per sample per protocol pair (P4, P5, and P6 were not used for the low viral input results; P5 & P6 were not used for the high viral input results). All results were based on sample output from 5M PE reads except for P7 which did not have samples with sequencing data at 5M depth and so the same sample results with 1M PE read depth were used instead. (**c)** Precision of SNV detection was measured using a VAF detection threshold of 20% or more and using the consensus SNVs as ground truth. Grey diamonds are results at 1M PE reads and blue triangles are results at 5M PE reads. (**d)** Mappability measured the percentage of reads that aligned to the target (SARS-CoV-2) genome and was estimated overall, and not with respect to a particular read depth. (**e)** Genome coverage was the percentage of the SARS-CoV-2 genome at ≥ 10x. Grey diamonds are results at 1M PE reads and blue triangles are results at 5M PE reads. (**f)** Uniformity of coverage used the coefficient of variation (CV) in read depth. Thus, lower values imply more uniform and efficient, and thus, better coverage. CV was estimated generally using the 5M PE reads for the primary samples except for P7 where the results are based on 1.5M PE reads. Values at the top edge of the scale for each graph are censored values. Graphs (**a-c**) use samples NP08, NP29, and NP30 only. Graphs (**d-f**) generally only use samples NP08, NP29, and NP30 except for P4 at low input, where samples NP16, NP26 and NP27 were used.

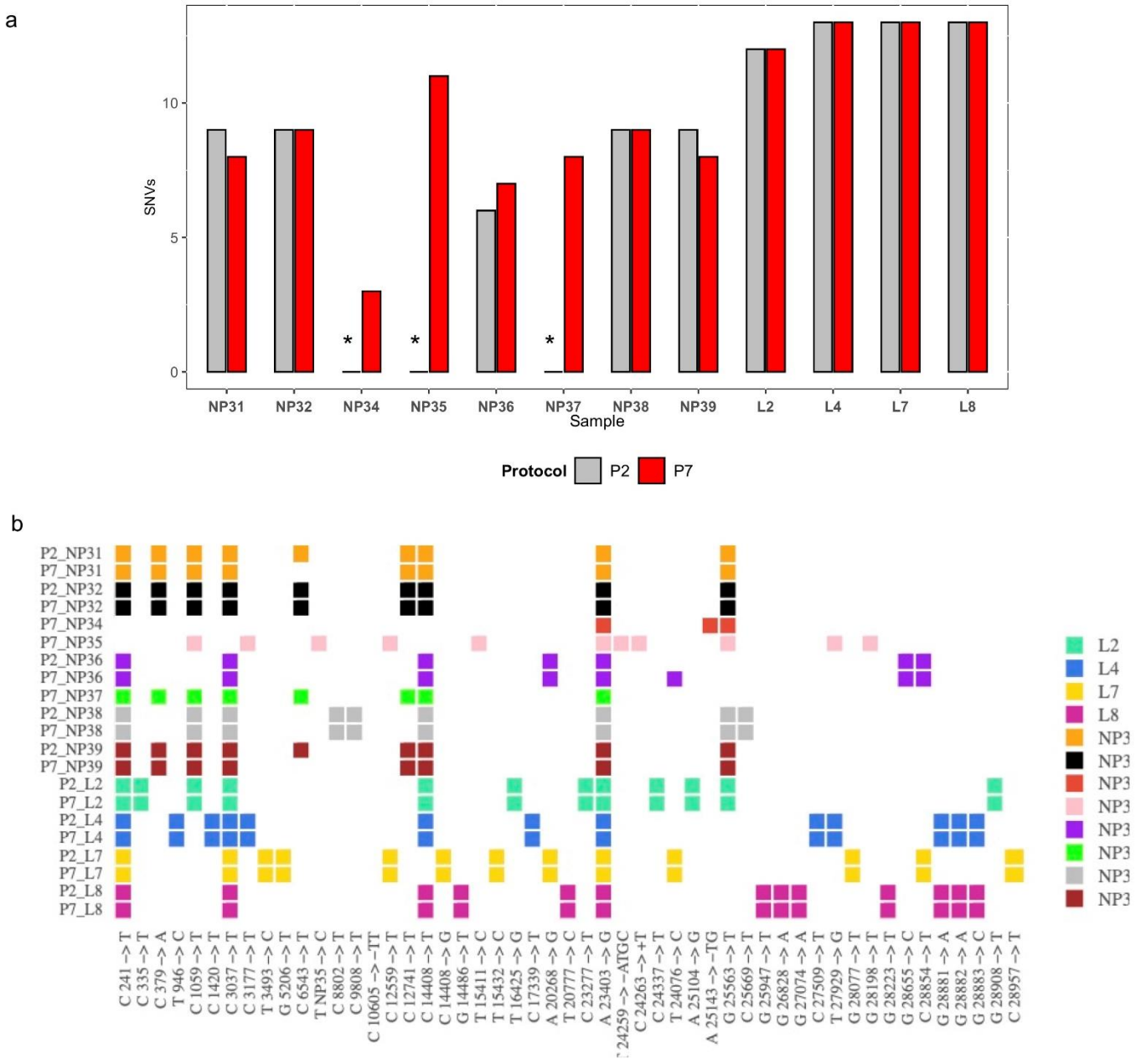


Figure S11. Consistency of SARS-CoV-2 SNV calling between P2 and P7 protocols using additional clinical samples, related to Figure 5. All SNVs with frequency higher or equal to 80% were deemed as true SNVs and were reported here. The viral input was same for each given sample in P2 and P7 protocols, but was different among samples. (**a**) Number of SNVs called in each sample by P2 and P7 protocols, respectively. Asterisk indicates failed SNV calling due to a low genome coverage. (**b**) Depiction of SNVs and their nucleotide transitions on genomic position.