

Figure S1: Profiles of SARS-CoV-2 sequences. **A)** Age and gender demographics of the sequenced individuals. **B)** A representative coverage plot for amplicon-based sequencing (top) and direct RNA sequencing (bottom). X-axis represents the 29.9 Kb genome and Y axis represents the depth in log scale. **C)** Average coverage obtained for each of the 98 amplicons in the ARTIC protocol. Except for the 98th amplicon (the last violin), all amplicons showed consistent amplification with many samples showing more than 10000x coverage.

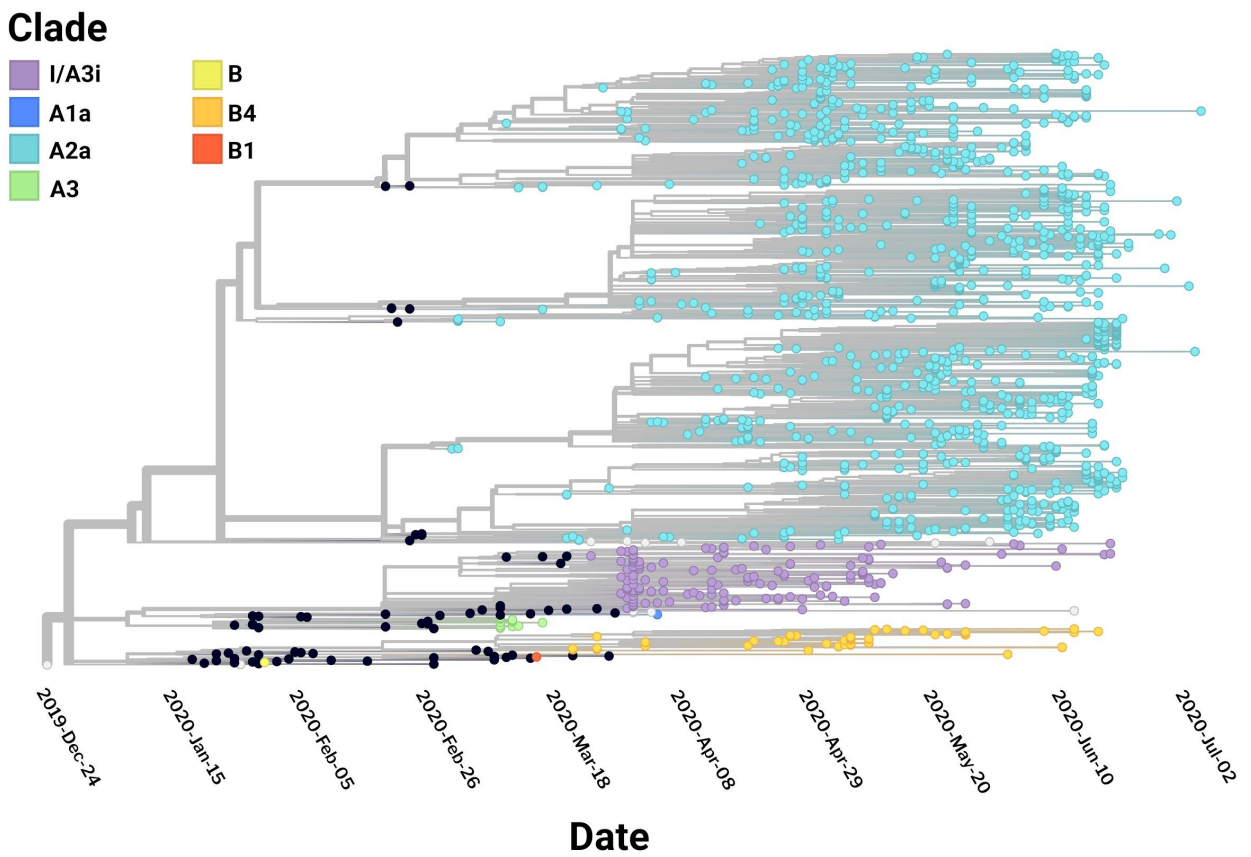


Figure S2: Indian SARS-CoV-2 genomes in the context of high quality global genomes. 10 high quality genomes from other countries belonging to each of the different clades were chosen and incorporated in the phylogenetic tree. Genomes sampled from other countries are colored black, whereas I/A3i clade is in purple.

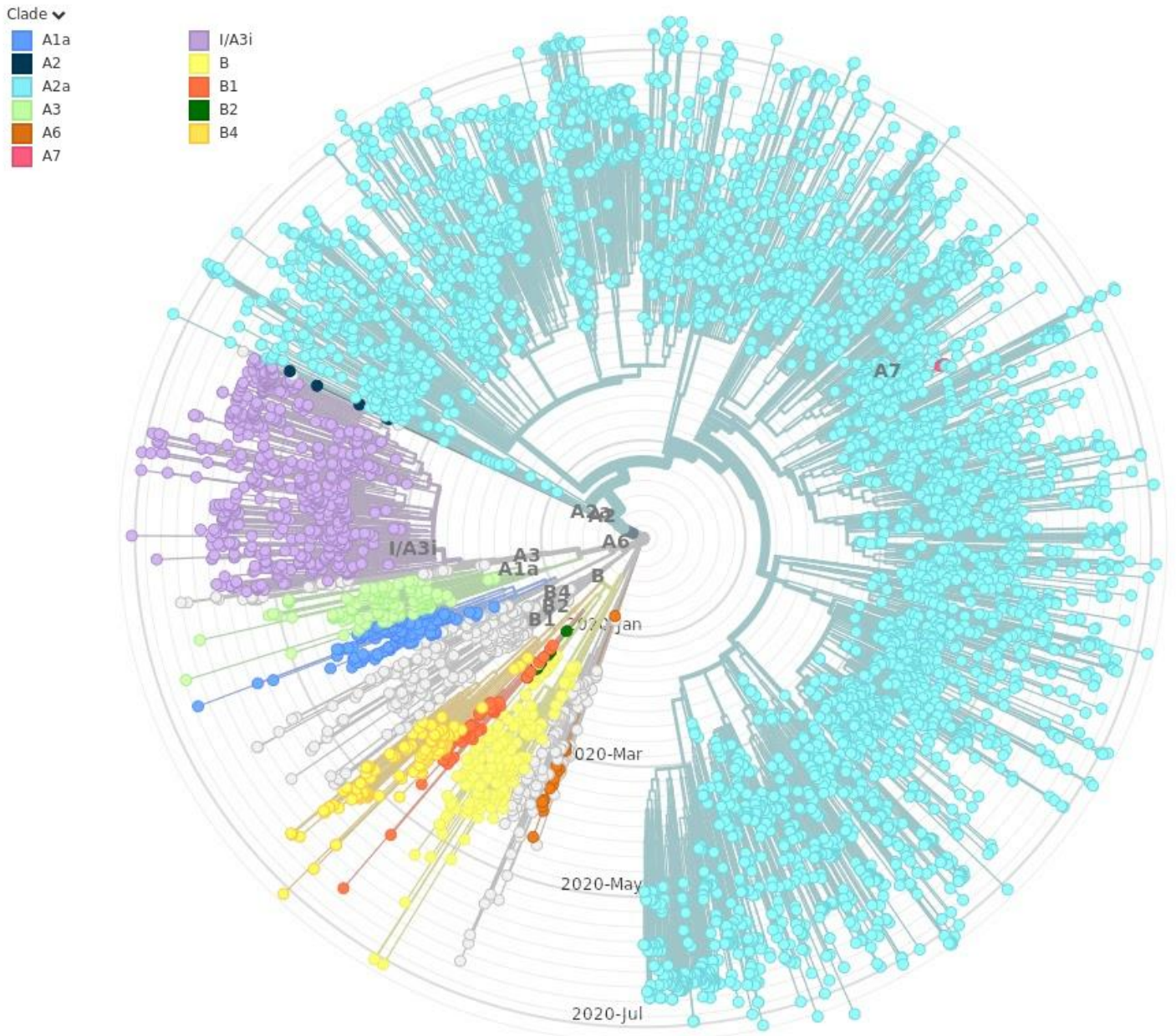


Figure S3: Phylogenetic tree of global SARS-CoV-2 genomes. 5571 high quality genomes were chosen based on the criteria established by the Nexstrain group, which include 1749 Indian genomes and 3882 genomes from other countries. Phylogenetic tree was constructed using IQTREE, and visualized using Auspice. I/A3i clade is marked in purple.