SUPPLEMENTARY MATERIAL

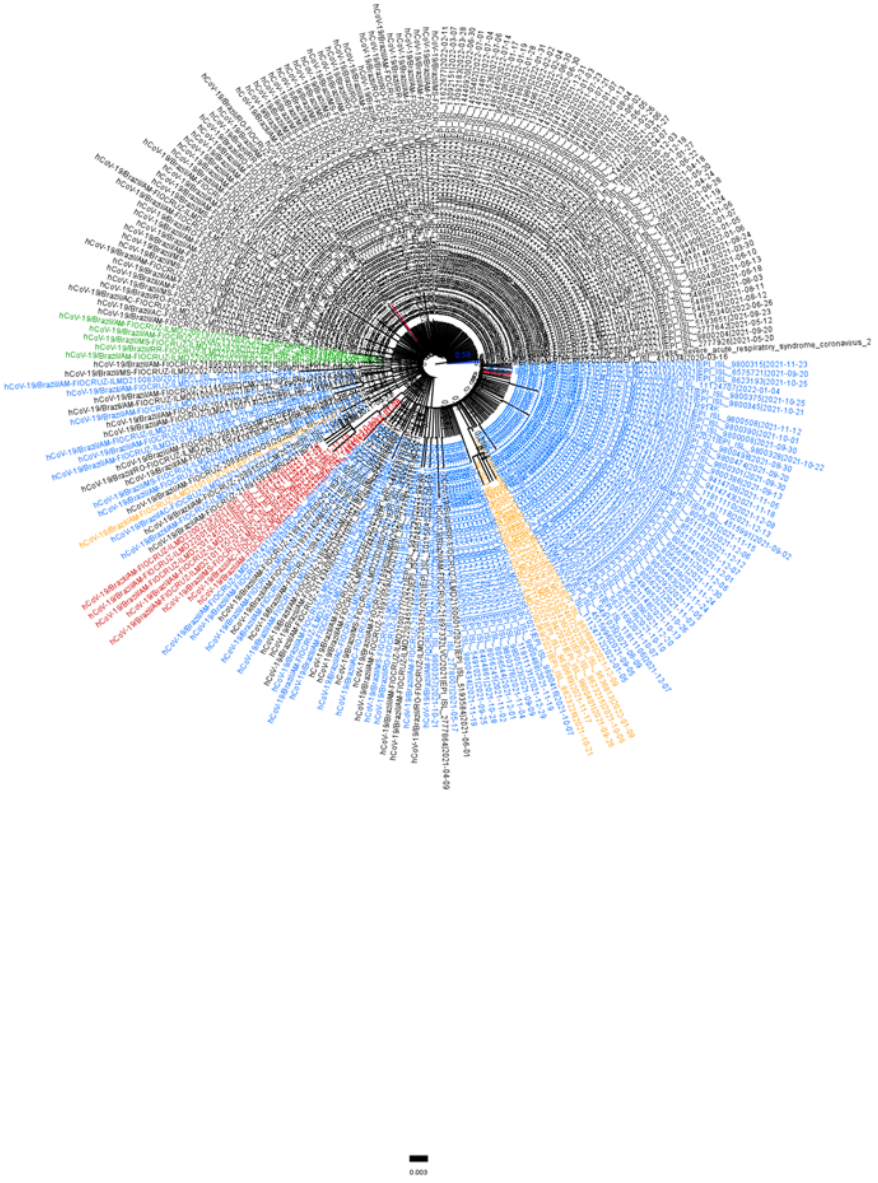

Figure S1. The phylogenetic tree was inferred by using the maximum likelihood method and the JTT matrix-based model (Jones et al. 1992) and the tree with the highest log likelihood (-2469.92) is shown. The percentage of trees in which the associated taxa clustered together is shown below the branches. Initial tree(s) for the heuristic search were obtained automatically by applying neighbor-joining and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with a superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 2.2595)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 236 amino acid sequences. There was a total of 177 positions in the final dataset. Analyses were conducted in MEGA11 (Tamura et al. 2021). Red branches represent the singletons M796D and R914K described in Table I. Blue branches represent the Wuhan isolate and a strain which shares a proline at position,323 and a single strain with the P323F mutation. Samples named in blue presented the mutation G671S or G671C. Samples named in orange presented the mutations L838I and L838M. Samples named in red presented the mutations P94L and P94S. Samples named in green presented the mutations A529T, A529V and A529S (Figures S1 and S2).

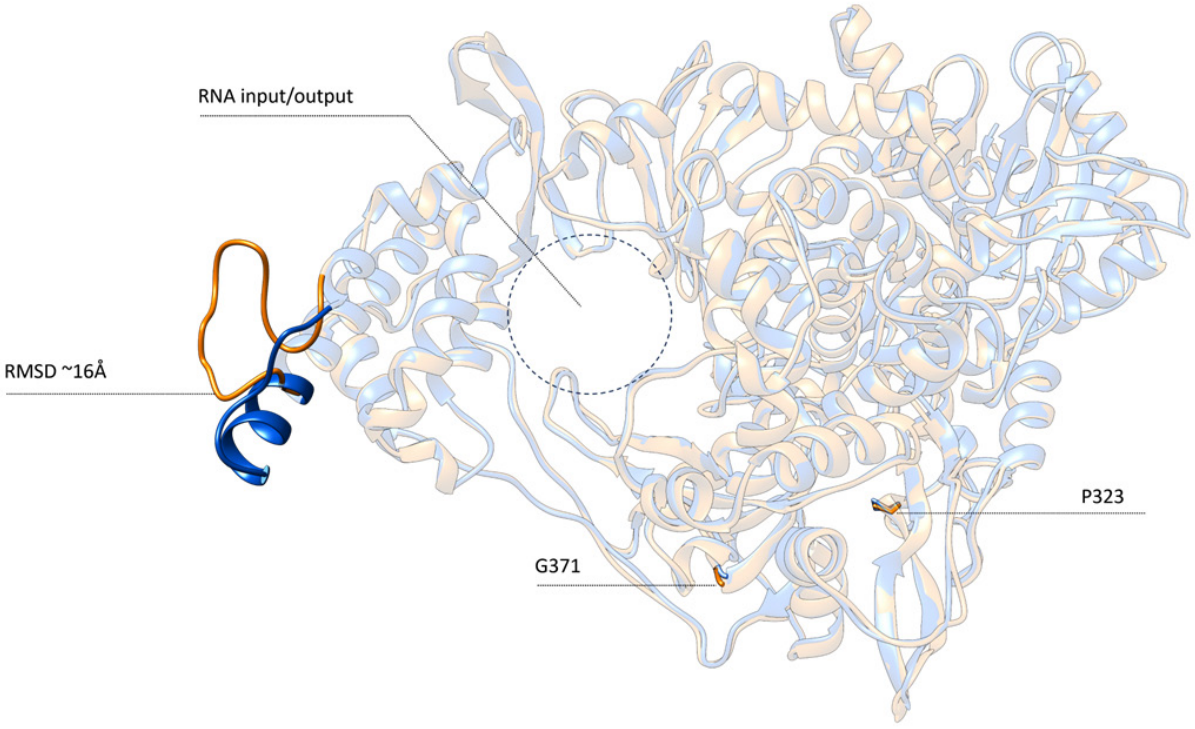

Figure S2. Overlap of the NSP12 structures of SARS-CoV (transparent blue) and SARS-CoV-2 (transparent orange) with the identification of the RNA entry and exit site. The total RMSD was 0.583 Å. Considering only the highlighted region, the RMSD showed a deviation of approximately 16 Å. We also identified the positions in which there were more mutations found in patients from Manaus, Amazonas (amino acids G371 and P323).