

and other SARSr-CoVs (FIG. 2). Using sequences of five conserved replicative domains in ppfab (3C-like protease (3CLpro), nidovirus RNA-dependent RNA polymerase (RdRp)-associated nucleotidyltransferase (NiRAN), RdRp, zinc-binding domain (ZBD) and HEL1), the *Coronaviridae* Study Group of the International Committee on Taxonomy of Viruses estimated the pairwise patristic distances between SARS-CoV-2 and known coronaviruses, and assigned SARS-CoV-2 to the existing species SARSr-CoV¹⁷. Although phylogenetically related, SARS-CoV-2 is distinct from all other coronaviruses from bats and pangolins in this species.

The SARS-CoV-2 S protein has a full size of 1,273 amino acids, longer than that of SARS-CoV (1,255 amino acids) and known that SARS-CoV's S protein (1,269 amino acids). It is distinct from the S proteins of most members of the subgenus Sarbecovirus, sharing amino acid sequence similarities of 76.7%–77.0% with SARS-CoVs from civets and humans,