

216 countries and regions from all six continents had reported more than 20 million cases of COVID-19, and more than 733,000 patients had died²².¹ High mortality occurred especially when health-care resources were overwhelmed. The USA is the country with the largest number of cases so far.

Although genetic evidence suggests that SARS-CoV-2 is a natural virus that likely originated in animals, there is no conclusion yet about when and where the virus first entered humans. As some of the first reported cases in Wuhan had no epidemiological link to the seafood market²², it has been suggested that the market may not be the initial source of human infection with SARS-CoV-2. One study from France detected SARS-CoV-2 by PCR in a stored sample from a patient who had pneumonia at the end of 2019, suggesting SARS-CoV-2 might have spread there much earlier than the generally known starting time of the outbreak in France²³. However, this individual early report cannot give a solid answer to the origin of SARS-CoV-2 and contamination, and thus a false positive result cannot be excluded. To address this highly controversial issue, further retrospective investigations involving a larger number of banked samples from patients, animals and environments need to be conducted worldwide with well-validated assays.

Genomics, phylogeny and taxonomy

As a novel betacoronavirus, SARS-CoV-2 shares 79% genome sequence identity with SARS-CoV and 50% with MERS-CoV²⁴.¹ Its genome organization is shared with other betacoronaviruses. The six functional open reading frames (ORFs) are arranged in order from 5' to 3' as follows: replicase (ORF1ab), spike (S), envelope (E), membrane (M) and nucleocapsid (N). In addition, seven putative ORFs encoding accessory proteins are interspersed between the structural genes²⁵. Most of the proteins encoded by SARS-CoV-2 have a similar