Phylogenetic Analysis of Severe Associated Respiratory Syndrome (SARS) Related Coronavirus

Saru Maharjan

University of Nottingham, UK

Abstract:

Corona viruses are enveloped viruses with a positive-sense single-stranded RNA genome. It is roughly spherical in shape having a diameter of around 125 nm. The genome size of coronaviruses ranges from 26.4 to 31.7 kilobases and is counted as one of the largest RNA viruses. The first human coronavirus was discovered in 1960s, since then coronavirus family gained notoriety with the emergence of severe acute respiratory syndrome (SARS). In 2002/2003, SARS-CoV was identified that lasted for 8 months and resulted in 8,098 confirmed human cases worldwide, of which 774 were lethal. In 2013, MERS-CoV (Middle East Respiratory Syndrome) appeared in kingdom of Saudi Arabia that continued to cause outbreak resulting in 2,260 affirmed cases in 27 countries of which 803 were fatal. In 2019, a global pandemic of SARS-CoV-2 (Severe Associated Respiratory Syndrome Coronavirus-2) was identified which resulted in 29,466,633 confirmed cases in more than 250 countries and 933,150 have been fatal. SARS-CoV-2 symptoms are milder, yet it is highly transmissible infectious diseases compared to SARS and MERS outbreaks, according to morbidity and mortality rates. Genome sequencing and phylogenetic analysis provide powerful insights into the origins, transmission, and evolution of emerging viral infections such as SARS-CoV-2.

The development of a phylogenetic tree is widely utilized in numerous areas of study where evolutionary divergence can be examined and illustrated. All evolutionary tree provides the basic information about a historical pattern of ancestry, divergence, and descent by illustrating a series of branches that merge at points indicating common ancestors which are connected through distant ancestors. To understand the evolutionary origin of SARS-CoV-2 and its relationship with different SARS related coronaviruses, the sequences of coronavirus family were downloaded from the Genbank and the phylogenetic tree of Nucleocapsid (N) protein and Spike (S) protein of coronavirus was constructed using MEGA7 (Molecular Evolutionary Genetics Analysis Version 7) that revealed bats as the natural reservoir. SARS-CoV-2 showed close relation with bat coronavirus RaTG13 illustrating its inception from bats. Similarly, SARS-CoV shared a common ancestor with bat SARS coronavirus and MERS-CoV with bat coronavirus HKU5 illustrating the origin of SARS coronavirus from bat and the zoonotic transmission of coronavirus from bat to other animals led to a human infection which acted as an intermediate host.

Coronaviruses are notable for their high recurrence of recombination and mutation rates, which allow them to adapt to new hosts and ecological niches. The global diversity of coronavirus is highly associated with the diversity of bat species exhibiting high recombination rates. However, the recombination rate was sparse in human coronavirus compared to that of bats. Bats have been speculated to harbor a wide variety of viruses and many of them cause's disease in humans. Bats live among many different animal species that gives the opportunity to coronavirus transmission. Interestingly, it has been found that coronavirus in bats requires an intermediate host before spilling over into humans. In SARS-CoV-1, palm civet (Paguma larvata) acted as an intermediate host allowing subsequent animal to human transmission. There is no evidence of direct transmission of coronavirus from bats to humans. Likewise, the phylogenetic tree illustrated close relation of SARS-CoV-2 Wuhan isolates with Pangolin coronavirus isolate MP789 and it was demonstrated that both SARS-CoV-2 and Pangolin coronavirus shared a common ancestor with bat coronavirus.

A novel coronavirus SARS-CoV-2 has diversified in several phylogenetic lineages while it spread topographically across the world. To better comprehend the transmission events of SARS-CoV-2, the coronavirus outbreaks were followed, the genome of CoV-sequences (Coronavirus-sequences) were downloaded from the GISAID (Global Initiative on Sharing All Influenza Data), complete genomes were used as only a handful of genomic variants can be seen in novel SARS coronavirus because of its new introduction in human populace. The high coverage (sequences where >95% of the genome was accessible) Nottingham SARS-CoV-2 sequences were analyzed along with hospital sequences and reference sequences (sequences from highly infected countries) that indicated various lineages circulating in Nottingham and based on lineages and phylogenetic tree the transmission events of COVID-19 was determined.

For the evaluation of lineages, Pangolin (Phylogenetic Assignment of Named Global Outbreak LINeages) COVID-19 lineage Assigner was used where nearly 20 different lineages was found to be circulating in Nottingham and majority of

them belonged to lineage B (B.1, B.1.1, B.1.1.1, B.1.1.9, B.1.36, B.1.5, B.15, B.1.34, B.9, B.10, B.3, B.2.2, B.2.6, B.2.1, B.2, B.2.5, B.6), one lineage A (A.2). The hospital sequences revealed lineages B.1 and B.3 indicating multiple introductions of the virus. Through analysis, it was found that lineage B.1 was initially tracked in continents such as North America, Europe, and Australia. There was no lineage transmission from China and South Korea. Moreover, the phylogenetic tree illustrated a close relation of Nottingham isolates to the sequences of the United States of America, Europe, and Australia with shared lineages inferable from inbound international travels that happened to be directly associated with virus transmission.

This study provides valuable insights into the evolution and transmission of the virus, which is important in identifying newly discovered pathogens, epidemiological investigations, and tracking the natural reservoirs of zoonotic diseases. The importance of understanding the origin and transmission networks of the viral infection guides in the implementation of control measures. There were disease outbreaks in the past controlled by good surveillance such as Ebola, Influenza which helped to minimize the spread of viral infection. Strategies like enabling rapid diagnosis, isolation, better management of suspected cases can plunge the transmission events of COVID-19. Moreover, this helps in the rapid detection of viral infection and provides a platform to predict the future trends of upcoming viral outbreaks.