The Emerging of The 2019 Novel Coronavirus 2019-nCoV

Hassan Naji

Abstract—The world is witnessing a major pandemic caused by the 2019 novel coronavirus (2019 nCoV). Efforts are undergoing in order to identify the phylogeny of this elusive virus. It turned out that 2019 nCoV is closely related to two viruses found in bats. Similar to their distant cousins, SARS CoV and MERS CoV, 2019-CoV evolved from bats due to their large genome and greater genetic diversity. Such genetic diversity enables coronaviruses to mutate and form new species that frequently surface among the human population. Developing vaccine and predicting mutation trends in coronaviruses along with surveillance systems are proven pivotal in controlling current and future outbreaks.

Index Terms—2019-nCoV, Bats, Coronavirus, Coronavirus Vaccine, Coronaviridae, Genomic Sequence, MERS CoV, Nidovirales, Orthocoronavirinae, Orthocoronavirinae, SARS CoV.

I. INTRODUCTION

On 31st December 2019, China alerted the World Health Organization (WHO) about a series of pneumonia-like cases in Wuhan. WHO declared the virus a public health emergency of international concern, and health departments around the world started scrambling to control the virus. The causative agent was identified as a novel coronavirus and was provisionally named 2019-novel coronavirus. Novel coronaviruses are now regularly identified. Coronaviruses are zoonotic and have been identified in various mammals including bats, dogs, mice, cats, civets, and in several avian hosts. The epidemiological association of patients with viral pneumonia in December 2019, was found to be with Huanan seafood market in Wuhan, China. In addition, the virus has been identified in several other countries. Infection has been reported in family clusters and medical workers. Also, human to human transmission has been reported (Lu et al., 2020).

2019-nCoV infections cause illnesses ranging from people with little to no symptoms to people being severely ill and dying. Common signs of infection are respiratory symptoms, cough, shortness of breath, fever, and difficulty in breathing. In severe cases, infection can lead to severe acute respiratory syndrome, pneumonia, kidney failure and even death (WHO, 2020).

The purpose of this paper is to shed the light on the genomic sequences of 2019-nCoV, and how it is related to two viruses that emerged in the past and caused pandemics, namely SARS-CoV and MERS-CoV. A special attention is given to historic perspective and the origin of coronaviruses.

II. GENOMIC SEQUENCE

Coronavirus belongs to the family Coronaviridae, subfamily Orthocoronavirinae, and the order Nidovirales. It is composed of a single-stranded, positive-sense RNA genome. Its genome contains 26 to 32 kilobases in length. Among all RNA viruses, the genome of coronavirus is the second largest and non-segmented, reaching up to 30kb in length. This large genome provides greater genetic diversity and increased chances of cross-species transmission. Presently, there are four genera of coronavirus, namely: alpha-, beta-, gamma-, and delta-coronavirus. These genera consist of 38 different species in the subfamily Orthocoronavirinae. During viral replication, numerous sub-genomic RNAs are generated, that is constantly leading to an increase in the number of coronavirus species. Therefore, the taxonomy of coronavirus is constantly changing (International Committee on Taxonomy of Viruses -ICTV, 2018).

Six coronaviruses are known to infect humans, including: OC43, SARS-CoV (Severe Acute Respiratory Syndrome Coronavirus), HKU1, 229E, NL63, and MERS-CoV (Middle East Respiratory Syndrome Coronavirus). Various genetic recombination among coronaviruses, result in novel viruses that are now regularly identified. These novel viruses are often associated with human outbreaks (Su et al., 2016).

The virus genome consists of 6 major open reading frames (ORFs) that are common to coronaviruses and other accessory genes. The genomic RNA encodes for two large overlapping polyproteins, ORF1a and ORF1b, four structural proteins: spike (S), envelope (E), membrane (M), and nucleocapsid (N), other nonstructural and accessory proteins. Open reading frame 1b gene (ORF1b) is used for the classification of viruses, however, the regions of viruses outside ORF1b show great diversity among viruses of the same species, confounding the designation (Woo, Huang, Lau, & Yuen, 2010). The complete genomic sequence of 2019-nCoV is most closely related to bat-SL-CoVZC45 and another beta-coronavirus of bat origin, bat-SL-CoVZXC21. Bat CoVs make up almost 60% of all alpha and beta-coronavirus species (Fan, Zhao, Shi, & Zhou, 2019). Out of the five gene regions: M, 14, E, 7, and N, the highest sequence identity of 98.7% is in the E gene while the lowest sequence identity of 75% is in the S gene with bat-SL-CoVZXC21 and bat-SL-CoVZC45. The 2019-nCoV strains are genetically less similar to MERS-CoV (about 50%) and SARS-CoV (about 79%). The most divergent region of the genome is the region that encodes the N terminus of the spike proteins. The spike protein is usually responsible for receptor binding and membrane fusion and is important for
determining transmission capacity and host tropism. 2019-nCoV encodes a longer spike protein compared with the bat SARS-like coronaviruses, MERS-CoV, SARS-CoV (Fig 1) (Lu et al., 2020). Functionally, the spike protein of coronaviruses is divided into 2 domains namely, S1 and S2 domain. The S1 domain is responsible for receptor binding while the S2 domain is concerned with cell membrane fusion.

Phylogenetic analysis of the 2019-nCov indicates that it is a novel beta-coronavirus, belonging to the subgenus Sarbecovirus. 2019-nCov is distinct from SARS-CoV in a phylogeny of the complete RNA dependent RNA polymerase (RdRp) gene. Sarbecoviruses are classified in 3 clades:
- Clade 1 is formed by two SARS-CoV-related strains from Rhinolophus species from Bulgaria and Kenya.
- Clade 2 is formed by two bat-derived SARS-like strains from Zhoushan in eastern China and the ten 2019-nCov from Wuhan
- Clade 3 is formed by SARS-CoV strains from humans and many genetically similar SARS-like coronaviruses from bats collected from southwestern China (Fig 2) (Lu et al., 2020).

MERS-CoV=Middle East respiratory syndrome coronavirus. 2019-nCoV=2019 novel coronavirus. SARS-CoV=severe acute respiratory syndrome coronavirus.

2019-nCoV is closer to bat-SL-CoVZXC21 and bat-SL-CoVZC45 at the whole genome level, while the receptor-binding domain of 2019-nCoV falls within lineage B and is closer to that of SARS-CoV. The receptor-binding domain of beta-coronaviruses is commonly located in the C-terminal domain of S1, as in MERS-CoV and BatCoV HKU4 for lineage C, and SARS-CoV29 for lineage B (Li, Li, Farzan, & Harrison, 2005). The receptor-binding domain of 2019-nCoV, like other beta-coronaviruses, is composed of an external subdomain and a core. The external subdomain of the 2019-nCoV receptor-binding domain is more similar to that of SARS-CoV, suggesting that 2019-nCoV might also use angiotensin-converting enzyme 2 (ACE2) as a cell receptor. 2019-nCoV does not use other coronavirus receptors like dipeptidyl peptidase 4 and aminopeptidase N (Zhou et al., 2020).

III. EPIDEMIOLOGY OF CORONAVIRUSES

Coronaviruses have been known to infect humans since 1960’s. The first case of human coronavirus (HCoV) infection was reported in 1960 in a patient with common cold. Since then, coronaviruses have been simultaneously maintained in nature and caused fatal human respiratory diseases worldwide. Coronaviruses infect various domestic and wild animals, as well as humans and cause diseases in them. The natural reservoir host of coronaviruses is bat. The alpha- and beta-coronaviruses mainly infect mammals, while gamma- and delta-coronaviruses mainly infect birds. Since 2002, beta-coronaviruses, MERS-CoV and SARS-CoV have caused pandemics in human (Zaki, Van Boheemen, Bestebroer, Osterhaus, & Fouchier, 2012). The 2019-nCov is similar to SARS-CoV, which is closely linked to bat viruses and have a potential to infect humans. The epidemic of 2019-nCov started in Wuhan, China, from December 12th, 2019, from a local fresh seafood market. The epidemic has caused 2010 fatalities by February 19th, 2020 and resulted in more than 75,000 confirmed laboratory cases of infection worldwide. In addition, 2019-nCov has now been reported in many countries around the world such as Japan, Thailand, Singapore, Malaysia, South Korea, Europe, and the USA.

Bats account for about a fifth of all mammalian species and constitute the second largest order of mammals. Bats have the capability of powered flight, which allows them to migrate to a longer range as compared to other land mammals. This migratory capability of bats is relevant with disease transmission and flight enabled viruses to coexist with bats. Viruses belonging to most of the viral families are found in bats. Bats are divided into two large suborders: the Yangochiroptera, consisting of a total of thirteen microbat families, and the Yinpterochiroptera, comprising of one Pteropodidae (megabat) and five Rhinolophoidea (microbat) families (Teeling et al., 2005). Bats have been recognized to harbor the viruses of SARS-CoV, MERS-CoV, and now the 2019-nCoV. Majority of the coronaviruses can be found in China because most of these bat hosts live near humans.

Fig. 1. Genomic organization and sequence comparison of 2019-nCoV. Adopted from Lu et al., 2020.

Open reading frames of more than 100 nucleotides are shown. 2019-nCoV=2019 novel coronavirus. MERS-CoV=Middle East respiratory syndrome coronavirus. SARS-CoV=severe acute respiratory syndrome coronavirus.

Fig. 2. Phylogenetic analysis of full-length genomes of 2019-nCoV and representative viruses of the genus Betacoronavirus. Adopted from Lu et al., 2020.
which potentially transmit viruses to livestock and humans. Several facts suggest that the 2019-nCoV is transmitted to humans via an intermediate host, between bats and humans. And the humans act as terminal hosts. It is because when the epidemic was first reported in Wuhan in December 2019, most of the bat species were hibernating and no bats were being sold or found at the Huanan seafood market; however, various non-aquatic animals (including mammals) were being sold. It is suggested that snakes might be the wildlife animal reservoir, which acted as intermediate host for the 2019-nCoV. However, it is not yet confirmed that which wildlife animal is exactly responsible for transmitting 2019-nCoV to humans, and further investigations are still required (Lu et al., 2020).

Coronaviruses are present in human history from a long time and are associated with certain outbreaks in this century including the SARS (Severe Acute Respiratory Syndrome), MERS (Middle East Respiratory Syndrome), and the SADS (Swine Acute Diarrhea Syndrome). All these outbreaks are related to coronaviruses transmitted through bats via an intermediate host to humans. the SARS-CoV was identified in 2003. The outbreak started in China, in November 2002 and resulted in 8000 cases affecting 37 countries in 2003. The epidemic was at peak in the 1st week of February 2003 and resulted in 774 deaths during 2002–03 (Xu et al., 2004). Since then, few cases have occurred due to laboratory accidents or through animal to human transmission. It was identified that the animal reservoirs of SARS-CoV was bats and through other animals like civet cats it spread to humans. SARS-CoV transmitted primarily from human to human. The Guangdong province of southern China was considered a potential zone of re-emergence of this virus (Wilder-Smith, 2005).

Like many other coronaviruses, the MERS-CoV originated in bats and caused multiple outbreaks. The virus was isolated from a patient with severe acute respiratory syndrome, in June 2012, in Saudi Arabia and it was later named as Middle East Respiratory Syndrome Coronavirus (MERS-CoV). The MERS-CoV of bat origin infected dromedary camels and the infection spread to humans via close dromedary camel contacts (Milne-Price, Miazgowicz, & Munster, 2014). MERS-CoV caused 858 fatalities since September 2012, and 2494 laboratory-confirmed cases of infection in 27 countries.

Also, CoVs are associated with pandemics in domestic and wild animals. SADS-CoV (Swine Acute Diarrhea Syndrome Coronavirus) was recently identified as the causative agent for a large-scale outbreak of fatal disease in pigs in China that resulted in the death of more than 20,000 piglets (Zhou et al., 2018). SADS-related CoV was identified from anal swabs collected from bats in Guangdong province during 2013-2016, majorly in horseshoe bats (Rhinolophus species). Besides these viruses, NL63, beta-coronavirus OC43, HKU1, and alpha coronavirus 229E can also cause respiratory diseases in humans (Zhang & Yoo, 2016). Transmissible gastroenteritis virus (TGEV) and porcine epidemic diarrhea virus (PEDV) are also significant emerging and re-emerging viruses in pigs that are linked with significant economic threat to the swine industry. Moreover, avian infectious bronchitis virus (IBV, gamma-CoV) is the cause of a contagious disease affecting poultry production worldwide. Coronaviruses have also been associated with whale deaths and catarhal gastroenteritis in mink (mink coronavirus) (Vlasova et al., 2011).

For decades, coronaviruses pose huge threat to humans and scientists have been aware of these viruses. In recent years, zoonotic viruses are linked to most of the outbreaks that emerged from viruses and scientists have been making efforts to predict these emerging outbreaks and to warn people. Bats scientists are well aware of the fact that bats carry the highest proportion of viruses that can infect any planet of mammalian species. It is very tricky and hard to predict the effect and spread of new coronaviruses. USAID’s PREDICT program was an initiative that was supposed to predict the emergence of new viruses. The scientists of the PREDICT program identified almost 50-SARS like coronaviruses (Holder, Tunseth, Vodzak, Norton, & Murray, 2016).

In March 2019, scientists working at the Wuhan Institute of Virology, warned about a future outbreak of SARS- and MERS like coronavirus illness. They predicted that it would emerge form bats inhabiting this region. Scientists in China have been raising the flag on these viruses for 15 years. These viruses are constantly infecting human population through bats in China, in Southeast Asia.

IV. VACCINE DEVELOPMENT EFFORTS GOING ON

There is currently no vaccine and no anti-viral treatment available for 2019-nCoV. However, efforts are being speeded up for the development of an effective vaccine. At Inovio’s laboratory in San Diego, scientists are using a new type of DNA technology for the development of a potential vaccine. The vaccine is currently called “INO-4800”. There are plans to enter the vaccine in human trials by early summer. The biotech company in Beijing will collaborate with Inovio to help in the management of human trials. The vaccine is being made using DNA sequences from the virus that targets specific parts of the pathogen. The body’s own natural response mechanisms mount a response to the pathogen and act as a factory for vaccine. Coalition for Epidemic Preparedness Innovations (Cepi) is funding for vaccine development in the lab. Two other programs are also being running for the development of vaccine for 2019-nCoV. Efforts are being made for developing “molecular clamp” which will enable rapid production of vaccine against various viral pathogens. (Chi, n.d.).

Other companies are also being working hard for developing vaccine against the 2019-nCoV. These companies include GlaxoSmithKline and Johnson & Johnson. The development of the vaccine is really challenging because of the unavailability of approved vaccines against similar viruses, like SARS and MERS coronaviruses, which had recently caused outbreaks in China and Saudi Arabia, respectively. NuGenerex Immuno-Oncology is attempting to develop a vaccine using the company’s own peptide-based technology. All these circumstances have urged the Food and Drug Administration to think about new technologies for vaccines like RNA interference and molecular clamps. These technologies will speed up the development of vaccine against 2019-nCoV. Remdesivir, was developed for Ebola
virus by Gilead Sciences in 2017. This drug is injectable and has a broad-spectrum action against zoonotic and epidemic coronaviruses in mouse models. It has been reported that administration of remdesivir improved signs of infection in infected individuals. However, to establish the safety and effectiveness of remdesivir, there is need of a randomized and controlled trial (National Institute of Health – NIH, 2020).

Currently, the best way to prevent infection is by decreasing exposure to the virus. CDC recommends some preventive measures to avoid the spread of respiratory viruses. The recommendations include avoiding close contact with infected persons, washing hands properly and wearing facemasks.

V. IMPORTANCE OF SURVEILLANCE SYSTEM IN PREDICTING OUTBREAKS

Over the past 20 years, coronaviruses have caused two pandemics, severe acute respiratory syndrome (SARS), and Middle East respiratory syndrome (MERS). In recent years, global surveillance programs were initiated to mitigate these pandemic threats. On the basis of past experiences, it is expected that coronaviruses can cause future outbreaks like the 2019-nCoV. Owing to a large genome amongst all RNA viruses, there is a high chance of endless combinations for coronaviruses. Considering the high genetic diversity of these viruses, surveillance on virus transmission ability and viral mutations is very important and urgently needed. Several of these surveillance programs have started and include USAID’s PREDICT program and the Global Virome Project (GVP). These surveillance programs use approaches like NGS (Next Generation Sequencing) to ensure an unbiased analysis of the evolution of bat-CoVs. The implementation of NGS methodology for unbiased viral surveillance is costly. So, the efficiency of NGS has been improved by enrichment of viral nucleic acid using a capture-based system (Li et al., 2020).

USAID’s Emerging Pandemic Threats (EPT) program, the PREDICT was initiated in 2009 for detection and discovery of zoonotic viruses which can cause pandemics. PREDICT is working with partners in 31 countries and is strengthening global surveillance by identifying viruses that can be shared between people and animals. PREDICT is using PCR with high data sequencing. PREDICT has developed detection and identification protocols in laboratories of all 31 countries, ensuring the capacity to detect pandemic risks. PREDICT provides data related to pandemics through viral testing and modeling a scenario of outbreak. Results are then provided to host governments for examination and for approval for public health release. According to the PREDICT, there are almost 1.6 million unknown species of viruses found in avian and mammalian populations, out of which 700,000 viral species have the potential to infect humans (PREDICT, 2014).

The Global Virome Project (GVP, 2020) characterizes the geographic scope and reservoirs and transmission hosts of viruses. It identifies behaviors and monitors the movement of viruses across different regions and hosts. The GVP targets the large pool of future viral threats by developing a genetic and ecologic database of all known naturally occurring viruses. This database enables us to prevent spillover in high risk regions of animal to human transmission. Unknown viruses can be isolated, and their genetic sequences can be shared in a publicly shared database, with the help of GVP. The GVP is concerned to give benefits by developing new diagnostics, risk mitigation strategies and vaccine technologies against emerging viral diseases.

These surveillance systems provide certain benefits in controlling and predicting future viral threats and leads to several advances in global health and science.

VI. CONCLUSION

The novel coronavirus 2019-nCoV is closer to two coronaviruses found in bat (bat-SL-CoVZXC21 and bat-SL-CoVZC45) than SARS-CoV and MERS-CoV. This fact stresses the idea that as humans come closer to wild animal such as bat, the rate of mutation in virus, especially coronavirus is going to stay on the rise. Coronaviruses are peculiar because they possess great genetic diversity due to their numerous sub-genomic RNA that are generated during viral replication leading to an increase in the number of coronavirus species. More research is needed to predict future viral mutation and implement surveillance systems that monitor these viruses as they are diverted into the human population.

REFERENCES

Fan, Y., Zhao, K., Shi, Z. L., & Zhou, P. (2019). Bat coronaviruses in China. Viruses, 11(3), 27–32. https://doi.org/10.3390/v11030210

Holder, K., Tunstel, D., Vodzak, M., Norton, B., & Murray, S. (2016). One health worldwide: Emerging infectious diseases, global health veterinarians, and zoonotic surveillance. American Journal of Tropical Medicine and Hygiene. https://doi.org/http://dx.doi.org/10.4269/ajtmh.abstract2016

International Committee on Taxonomy of Viruses (ICTV). (2018). Virus Taxonomy: 2018 Release ICTV.

Li, F., Li, W., Farzan, M., & Harrison, S. C. (2005). Structural biology: Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science. https://doi.org/10.1126/science.1116480

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., … Tan, W. (2020). Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet (London, England), 6756(20), 1–10. https://doi.org/10.1016/S0140-6736(20)30251-8

Milne-Price, S., Miazgowicz, K. L., & Munster, V. J. (2014). The emergence of the Middle East Respiratory Syndrome coronavirus. Pathogens and Disease. https://doi.org/10.1111/2049-632X.12166

PREDICT. (2014). Predict: Reducing Pandemic Risk, Promoting Global Health. Retrieved from http://data.predictglobal

National Institute of Health (NIH). (2020). Remdesivir prevents MERS Coronavirus Disease in Monkeys. Retrieved from https://www.nih.gov/news-events/news-releases/remdesivir-prevents-mers-coronavirus-disease-monkeys

Li, B., Shi, H., Zhu, Y., … Zhou, P. (2020). Discovery of Bat Coronaviruses through Surveillance and Probe Capture-Based Next-Generation Sequencing. Creative Commons, 5(1), 1–10.

Su, S., Wong, G., Shi, W., Liu, J., Lai, A. C. K., Zhou, J., … Gao, G. F. (2016). Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. Trends in Microbiology. https://doi.org/10.1016/j.tim.2016.03.003

Teeling, E. C., Springer, M. S., Madsen, O., Bates, P., O’Brien, S. J., & Murphy, W. J. (2005). A molecular phylogeny for bats illuminates apathy and the fossil record. Science. https://doi.org/10.1126/science.1105113

Vlasova, A. N., Halpin, R., Wang, S., Ghedin, E., Spiro, D. J., & Saif, L. J. (2011). Molecular characterization of a new species in the genus Alphacoronavirus associated with mink epizootic catarrhal gastroenteritis. Journal of General Virology. https://doi.org/10.1099/vir.0.025353-0
Wilder-Smith, A. (2005). The Severe Acute Respiratory Syndrome: Impact on Travel and Tourism. Travel Medicine and Infectious Disease. 4(2), 53-60. DOI: 10.1016/j.tmaid.2005.04.004

Woo, P. C. Y., Huang, Y., Lau, S. K. P., & Yuen, K. Y. (2010). Coronavirus genomics and bioinformatics analysis. Viruses. https://doi.org/10.3390/v2081803

Xu, R. H., He, J. F., Evans, M. R., Peng, G. W., Field, H. E., Yu, D. W., ... Schnur, A. (2004). Epidemiologic clues to SARS origin in China. Emerging Infectious Diseases. https://doi.org/10.3201/eid1006.030852

Zaki, A. M., Van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. M. E., & Fouchier, R. A. M. (2012). Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. New England Journal of Medicine. https://doi.org/10.1056/NEJMoa1211721

Zhang, Q., & Yoo, D. (2016). Immune evasion of porcine enteric coronaviruses and viral modulation of antiviral innate signaling. Virus Research. https://doi.org/10.1016/j.virusres.2016.05.015

Zhou, P., Fan, H., Lan, T., Yang, X. Lou, Shi, W. F., Zhang, W., ... Ma, J. Y. (2018). Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin. Nature. https://doi.org/10.1038/s41586-018-0010-9

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., ... Shi, Z.-L. (2020). Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. BioRxiv. https://doi.org/10.1101/2020.01.22.914952