Global Scientific JOURNALS

GSJ: Volume 8, Issue 4, April 2020, Online: ISSN 2320-9186

www.globalscientificjournal.com

Covid-19: Biological Origin and Phylogenetic Analysis

Sarah N. Kituyi[1], Joseph A. Gisaina[1], Eric M. Muthanje[1]

[1]University of Embu

Department of Biological Sciences

P. O. Box 6-60100, Embu, Kenya

Email addresses of co-authors: gisainajoseph@gmail.com,ericmurimi9@gmail.com

Corresponding Author: Sarah Naulikha Kituyi, skituyi@gmail.com

Conflicting Interests

The authors declare not conflict of interest

Abstract

Effective management and response to the Covid-19 pandemic requires an ideal understanding of the biology of Corona viruses with borrowing from previous experiences drawn from epidemics by related viruses. Among the human Corana viruses are the OC43, NL63, HKU1 and the 229E viruses. Besides them, are other Corana viruses of a zoonotic origin that have been shown to infect humans such as MERs and SARs-Cov. These viruses have the ability to exhibit inter and intra species transmission and like most other viruses, Corana viruses have also been shown to adapt very fast to environmental and host variability in an effort to foster natural selection and escape available natural immune responses and conventional therapy. For instance, most current epidemics are characterized by a unique selection of hosts with mutations that make it relatively hard to trace the original host organism. The Covid-19 pandemic caused by the SARs-Cov-2 virus is an example of a viral infection where natural selection has empowered the virus to effectively challenge efforts geared at drug design and the rapid identification of the primary host, but there is evidence pointing to a very high sequence similarity to a Corana virus in the *Rhinolophilid* bats hence suggesting a zoonosis. The variable regions as reported in the RBD within the S protein of this virus could have occurred as a result of a recombination event in an intermediate host prior to infection of a human host or as result of a natural selection event in the virus upon infection of a human host. This review will shed light on the biology of the SARS-Cov-2 virus, the history and the myths and misconceptions about Covid-19, the emergence and spread across the globe, disease prevention and diagnosis and phylogenetics

where we provide an analysis of sequences from 257 Covid-19 isolates in the viral database as at the time of writing this paper. The analysis presents evidence that the virus is mutating, and that the USA has sequences that are much more closely related to the Wuhan virus than the rest of the world.

The Biology of the SARs-Cov-2 Virus

Coronaviruses are group of viruses belonging to the order *Nidovirales* which consist of three families:*Arteriviridae, Coronaviridae, andRoniviridae* (Yang and Leibowitz, 2015).Classification via serological tests further subdivided the family *Coronaviridae* into four sub families namely: alpha, beta, gamma and delta coronaviruses (Fehr & Perlman, 2015;Yang and Leibowitz, 2015).Like their counterparts in the *Nidovirales* order, corona viruses are enveloped RNA viruses consisting of linear RNA genomes approximately up to 30 kb(Fehr & Perlman, 2015).The genome consist of 5 open reading frames encoding membrane proteins, replicase enzymes, spike glycoproteins,the envelope protein and nucleocapsid proteins (Han et al., 2019). The enveloped SARs-Cov -2 virion particle has been shown to comprise of the spiked protein also referred to as the (S) protein that has a unique role in dictating entry into a host via association with host cell receptors (Ou et al., 2020). The membrane protein designated as (M) and the envelope protein (E) have unique roles in the assembly of the envelope protein besides regulating the assembly of virion proteins while the nucleocapsid protein designated as N similarly interacts with the M protein to foster the assembly of the nucleocapsid-RNA genome (Schoeman & Fielding 2019). These proteins interact in a manner that heavily sustains the assembly of the viral like particles, the shape of the envelope protein, the stability of the RNA genome and until recently, they were thought to have an equal role in dictating the infection, and viability of the virus but it now emerges that some proteins especially the S and the M proteins might have a greater role (Schoeman & Fielding 2019; Guo et al., 2020).

The genomic RNA is used as a direct template for translation of polyproteins and non-structural proteins which form a replication- transcription complex (Snijder et al., 2006; Guo et al., 2020). This is then followed by synthesis of sub genomic RNAs in a discontinuous transcription process (Chen et al., 2020). The processed sub genomic RNAs possess 5´ leader and 3´terminal sequences and transcription is terminated at regulatory sequences located within the open reading frames (Hussain et al., 2005). Prior to the replication of the virus in the host, Corona viruses require to bind to their respective receptors in the host and the SARs-Cov-2 virus which causes Covid-19 has been shown to uniquely bind onto the human Angiotensin converting enzyme-2 (ACE-2) (Johnson et al., 2018; Kaiser 2020) on the surface of the lungs (Ou et al., 2020; Anderson et al., 2020). This binding is catalyzed by the S protein which has been shown to have a strong affinity for the receptor in its S1 domain (Johnson et al., 2018) and mutations in this protein and in the ACE2 receptor have been associated with a unique host selection pressure and varying disease pathogenicity (Fehr & Perlman, 2015; Johnson et al., 2018).The entry of the virus into the host cell upon binding of the S protein to the host cell receptor is via endocytosis (Wang et al., 2008; Yang et al., 2020) and association with the host receptor is fostered by the

cleavage of the S protein into the S1 and S2 domains which occurs upon recognition of the cleavage site by the host proteases (Hoffmann 2020). Upon successful entry into the host cell, the virus unpacks it's nucleocapsid-RNA genome into the cytoplasm and its then translated in a region believed to be an intermediate between the ER and the golgi (Guo et al., 2020) from where it is not well known howfurther processing of the virion proteins ensues to yield an assembly of viral like particles ready to bind membranes of other host cells to enhance further infectivity and disease progression(Fehr & Perlman, 2015; Guo et al., 2020).

The history, myths and misconceptions about Covid-19

The SARs-Cov-2 virus which causes Covid-19 was first reported in December 2019 in Wuhan, Hubei province in China (Xu et al., 2020).The disease was named and described as a pandemic by WHO in 2020(Lai et al., 2020;WHO, 2020). The virus is presumed to have originated from a large seafood and animal market in Wuhan (Yang & Sheng 2020). Despite the fact that there are numerous conspiracy theories revolving around the manufacture of the virus in the lab (Rabbi et al., 2020), there is sufficient evidence to indicate that the virus is a naturally occurring organism exhibiting high sequence similarity (96%) to a SARs-Cov virus in the *Rhinolophilid* bats (Andersen 2020, 2020;Zhou et al., 2020; Rabbi et el., 2020). However, studies have reportedthat a unique sequence disparity in the S-protein of the SARs-Cov-2 allows it to uniquely bind to the human ACE-2 receptor unlike the counter part in the *Rhinolophilid* bats suggesting that the virus might have undergone a natural selection event to allow for a human host receptor selection (Andersen et al., 2020). Further evidence suggests that the receptor binding domain (RBD) in the S protein of SARs-Cov-2 virus is similar to the RBD in a SARs-Cov virus of the pangolins which would further suggest the role of a pangolin as an intermediate between the *Rhinolophilid* bats and humans (Rabbi et al., 2020) but further analysis of the S-protein in the SARs-Cov-2 points to the presence of a polybasic cleavage site between the S1 and the S2 domains that is lacking in the SARs-Cov virus of the pangolins (Andersen et al., 2020). This cleavage site accessed by human proteases is vital for the binding of the SARs-Cov-2 virus to the human ACE-2 receptor (Hoffmann 2020). The later suggests that the virus might be undergoing selection either prior or upon infection of a human host to further confer human to human transmission (Andersen et al., 2020). Because the backbone of the SARs-Cov-2 virus cannot be matched to any available corona virus, it thus seems impossible for the virus to have been engineered in the lab(Andersen, 2020). Previous studies have also indicated the possibility of the emergence of novel viruses following genetic recombination events between existing viruses and new hosts (Holmes, 2004; Johnson et al., 2018). Gradual evolution of Corona lineages such as SARS-CoV has been associatedwith recombination (Graham and Baric, 2010) and evolutionary studies of SARS-Cov2 suggests that recombination could have occurred in the RBD in the spike gene of the 2019-nCoV ancestor (Andersen et al., 2020).

The emergence and spread across the globe, disease diagnosisand prevention

Since the first reported case in China in December 2019, the virus has spread across many countries of the world owing to human to human transmission through droplets or direct contact with infected persons or surfaces(Li et al., 2020). The disease has high transmission efficiency

catalyzed by the transmission by asymptomatic carriers as well as the convenience of global travel (Biscayart et al., 2020). The incubation period of the SARS-CoV2 virus is between 2 to 11 days (Lauer et al., 2020; Li et al., 2020) and studies in China suggest that the incidence rate of Covid-19 follows an exponential growth (Wu et al., 2020; Zhao et al., 2020). The disease has since grown into a pandemic resulting in small chains of transmission in some countries in Africa and large chains of transmission in other countries like Italy, united states of America, China, Germany, South Korea and even Japan (WHO, 2020). As at 2nd April, there were over 900,000 confirmed cases and over 40,000 deaths. United states of America has the highest number of reported cases followed by Italy, Spain, Germany and China https://www.worldometers.info/coronavirus/.

Covid-19 patients present with common flu- like complications which range from mild fever, cough, running nose and general body weakness while severe cases may entail pneumonia characterized by respiratory distress requiring management via ventilators (Adhikari et al., 2020). Laboratory diagnosis of Covid-19 is via molecular testing for the viral genetic material in respiratory specimens and blood samples using the Polymerase Chain Reaction (RT-PCR) detection as the gold standard (WHO,2020). Radiological examinations might equally support the confirmation of the disease but unfortunately this testing does not provide enough ground for detecting ground glass-opacity (GGO) and thus not considered the first line imaging platform for Covid-19(Zu et al., 2020). Despite their non-specificity for Covid-19, radiological examinations, especially chest CT scans, are recommended for early detection of the disease and as a guide to clinical management (Kanne, 2020). It is expected that advances in testing would yield more ground for serological testing using techniques such as ELISA and antibody-antigen based rapid diagnostic testing(WHO,2020).

There is currently no available treatment or vaccine for Covid-19, however, case management as dictated by disease presentation is currently fostered via approaches such as antiviral therapy, empirical antibiotics, and mechanical ventilators (Chen et al., 2020) with various other treatment options still undergoing testing. Available drugs for other ailments such as Chloroquine and hydroxychloroquine are currently under clinical trials as potential Covid-19 drugs (Lai et al., 2020). Chloroquine has been tested before in treating chronic viral infections via increasing the endosomal pH required for virus-cell fusion and interfering with glycosylation of the SARS-CoV receptors (Savarino et al., 2003; Yan et al., 2013). The drug has shown modest activity against Hepatitis C infections before (Touret & de Lamballerie, 2020). According to the interim trials carried out in China, the drug has shown efficacy against theCorana virus and it may be considered an option for treatment in the future upon further testing(Gao et al., 2020). Other drugs currently undergoing testing includeIvermectin which has shown activity in cell lines and awaits further testing (Caly et al., 2020). There is also evidence that convalescent serum from recovering Covid-19 patients might provide passive therapy in critically ill patients (Guo et al.,2020).

Preventive measures are highly recommended to reduce global spread of the virus especially via the human to human transmission and contact with contaminated surfaces(Lai et al., 2020). The

world health organization recommends control of the source of infection through early diagnosis of the disease, isolation of suspected cases, real time reporting of new cases and supportive treatment (WHO,2020). In line with WHO recommendations,various governments across the globe have implemented measures to ensure social distancing among people which include; banning public gatherings in events and placessuch as churches and mosques, funerals and weddings, and restricting public transport at national level. Contact tracing of individuals, testing and quarantine are the major recommendations by the world health organization in the fight against the spread of the pandemic (WHO,2020). At an individual level, WHO recommends washing hands regularly, wearing personal protective equipment, maintaining social distance of at least one meter apart and keeping rooms well ventilated.

Phylogenetic studies

To further understand the genetic relatedness of the currently spreading Coronaviruses in the world, we performed phylogenetic analysis on a collection of 257 novel Coronavirus sequences deposited in virus pathogen resource repository (VPIR) **(**https://www.viprbrc.org/brc/home.spg?decorator=vipr**)** database from various countries at the time when we were preparing this review. Understanding the genetic relatedness among the reported Corona viruses in the wake of Covid-19 would shed light on the evolutionary trends of the virus and point to novel approaches that can help in circumventing the mechanisms that foster host selection. The alignment of the gene sequences was done using the VIPR workbench. Phylogenetic model generation and tree visualization was done using MEGAX (Hall, 2013), via the Maximum Parsimony method (Hall, 2013; Kumar et al., 2011). The most parsimonious tree with length = 3350 is shown. The MP tree was obtained using the Subtree-Pruning-Regrafting (SPR) algorithm with search level 1 in which the initial trees were obtained by the random addition of sequences (10 replicates). The tree was drawn to scale, with branch lengths calculated using the average pathway method and are in the units of the number of changes over the whole sequence. This analysis involved 257 nucleotide sequences. There were a total of 30572 positions in the final dataset. From our analysis, the sequences of the viruses deposited in this database exhibited a high level of similarity (sequences are highlighted in yellow) by tracing back to the very first strain sequences released from China isolates by Wuhan-Hu-1 (GenBank Accession MN908947) and other 13 first isolates collected in the year 2019 in China.

Table 1: Sequences from China the origin of the COVID-19

Strain Name	Virus Species	GenBank Accession	Sequence Length	Collection Date
Wuhan-Hu-1	Severe acute respiratory syndrome-related coronavirus	MN908947	781	12/2019
BetaCoV/Wuhan/IPBCAMS-WH-01/2019	Severe acute respiratory syndrome-related coronavirus	MT019529	829	12/23/2019
UNKNOWN-LR757995	Severe acute respiratory syndrome-related coronavirus	LR757995	845	12/26/2019
UNKNOWN-LR757998	Severe acute respiratory syndrome-related coronavirus	LR757998	850	12/26/2019
WIV07	Severe acute respiratory syndrome-related coronavirus	MN996531	780	12/30/2019
BetaCoV/Wuhan/IPBCAMS-WH-03/2019	Severe acute respiratory syndrome-related coronavirus	MT019531	819	12/30/2019
WIV02	Severe acute respiratory syndrome-related coronavirus	MN996527	824	12/30/2019
WIV05	Severe acute respiratory syndrome-related coronavirus	MN996529	825	12/30/2019
WIV06	Severe acute respiratory syndrome-related coronavirus	MN996530	826	12/30/2019
BetaCoV/Wuhan/IPBCAMS-WH-04/2019	Severe acute respiratory syndrome-related coronavirus	MT019532	828	12/30/2019
WIV04	Severe acute respiratory syndrome-related coronavirus	MN996528	830	12/30/2019
BetaCoV/Wuhan/IPBCAMS-WH-02/2019	Severe acute respiratory syndrome-related coronavirus	MT019530	842	12/30/2019
UNKNOWN-LR757997	Severe acute respiratory syndrome-related coronavirus	LR757997	854	12/31/2019

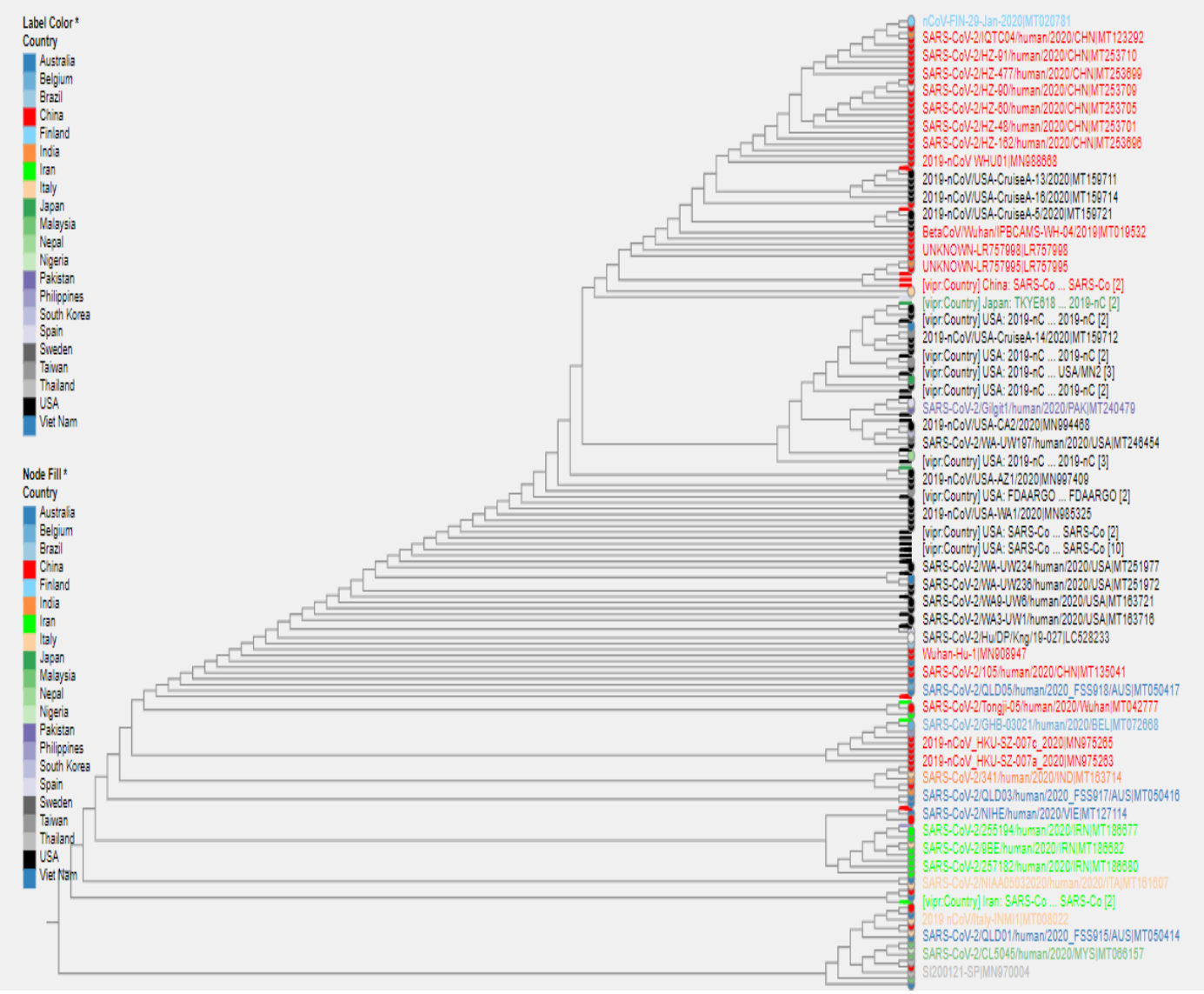

FIG 1: A Phylogentictree showing the China earliest isolates reported and their relationship with the isolates reported from different countries.

The analysis suggests that despite the differences in the collection date and geographical origin, the strains from China were nearly identical and most clustered together as shown in figure 1. Surprisingly majority of them shared a high degree of similarity with the strains reported from USA. We also noted that a large majority of the viruses reported from other parts of the globe are

divergent from the earliest first reported Wuhan 1 strain. These observations may shed light on the observed differences in case presentation and fatalities observed globally. Guo et al.,2020 reports that an overexpression of the human ACE-2 may be a predisposing factor to the entry of the virus into a human host and subsequent development of Covid-19. In addition, Kaiser 2020 reports that the differences in the clinical presentation and exposure to Covid-19 is determined by variations in the gene that codes for the expression of the ACE-2 receptor. The close relatedness of the sequences in the USA to the China isolates and the observed divergence in the isolates in other regions across the globe might equally explain the high case presentation in the USA where the current numbers as at 5th April 2020 far outweigh the reported cases in Chinahttps://www.worldometers.info/coronavirus/. It is likely that the SARs-Cov-2 virus in the US is closely related to the isolates in China while recombination events might be responsible for the divergency in the isolates across other regions. This observation alongside the reports on the variation in case presentation driven by the ACE-2 receptor may suggest that the observed differences in the presentation of Covid -19 are driven by genetic factors in both the host and the virus.

Conclusion

The unique presentation of Covid-19 in different individuals has been associated with age and health status where the elderly and individuals with preexisting conditions such as high blood pressure are at a high risk. From the phylogenetic analysis, it is also likely that the nature of the virus would equally determine the case presentation where cases infected by viruses that are closely related to the first isolated Wuhan virus and the viruses in China would present with more case severity like in the USA and parts of Europe. Divergence from the China isolates would equally suggest differencesin case presentation and it is also likely that differences in the expression of the human ACE-2 receptor would trigger either susceptibility or protection from the virus. This prompts insight into further analysis of the genetics of both the virus and the host receptor as key determinants of the disease.

References

Adhikari S, MengS, Wu Y, Mao Y, Ye1 R, WangQ, SunC, SylviaS, Rozelle S, Hein Raat H & Zhou H. (2020).Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review *Infectious Diseases of Poverty*9:29 https://doi.org/10.1186/s40249-020-00646-x

Andersen k, (2020). The COVID-19 coronavirus epidemic has a natural origin, scientists say. *Scripps Research*

Andersen K, Rambaut A, Ian Lipkin W, Holmes E& Garry R (2020) The proximal origin of SARS-CoV-2. *Nature Medicine*

Biscayart, C., Angeleri, P., Lloveras, S., Chaves, T., Schlagenhauf, P., & Rodríguez-Morales, A. J. (2020). The next big threat to global health? 2019 novel coronavirus (2019-nCoV): What advice can we give to travelers? -Interim recommendations January 2020, from the Latin-American society for Travel Medicine (SLAMVI). *Travel Medicine and Infectious Disease*, 33, 101567.

Caly, L., Druce, J.D., Catton, M.G., Jans, D.A., Wagstaff, K.M. (2020) The FDA approved Drug Ivermectin inhibits the replication of SARS-CoV-2 in vitro, *Antiviral Research*, https:// doi.org/10.1016/j.antiviral.2020.104787

Fehr, A. R., & Perlman, S. (2015). Coronaviruses: an overview of their replication and pathogenesis. In Coronaviruses (pp. 1–23). *Springer*.

Felsenstein J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-791.

Gao, J., Tian, Z., & Yang, X. (2020). Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Bioscience Trends*.

Graham, R. L. and Baric, R. S. (2010) 'Recombination, Reservoirs, and the Modular Spike: Mechanisms of Coronavirus Cross-Species Transmission', *Journal of Virology*, 84(7), pp. 3134–3146.

Guo Y, Cao Q, Hong Z, Tan Y, Chen S, Jin H, Tan K, Wang D, & Yan Y (2020). The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak an update on the status. *Military Medical Research* 7:11.

Hall, B. G. (2013). Building Phylogenetic Trees from Molecular Data with MEGA, 30(5), 1229–1235. https://doi.org/10.1093/molbev/mst012

Han, Y., Du, J., Su, H., Zhang, J., Zhu, G., Zhang, S., … Jin, Q. (2019). Identification of diverse bat alphacoronaviruses and betacoronaviruses in China provides new insights into the evolution and origin of coronavirus-related diseases. Frontiers in Microbiology, 10, 1900.

Hoffmann M, Kleine-Webber H, Schroder S, Muller M, Drostenn C, Pohlmann S, SARsCov-2 entry depends on ACE-2 and TMPRSS2 and it is blocked by a clinically proven protease inhibitor. Cell Press, 781; 1-10.

Holmes, E. C. (2004) 'The phylogeography of human viruses', *Molecular Ecology*, 13(4), pp. 745–756.

Hussain, S., Chen, Y., Yang, Y., Xu, J., Peng, Y., Wu, Y., … Guo, D. (2005). Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus. Journal of Virology, 79(9), 5288–5295.

Johnson B, Graham R, &Menachery V (2018) Viral metagenomics, protein structure, and reverse genetics: Key strategies for investigating coronaviruses. *Journal of Virology* https://doi.org/10.1016/j.virol.2017.

Kaiser J. (2020). How sick will the coronavirus make you? The answer may be in your genes. Science's extensive COVID-19 coverage https://www.sciencemag.org/

Kanne, J. P. (2020). Chest CT findings in 2019 novel coronavirus (2019-nCoV) infections from Wuhan, China: key points for the radiologist. *Radiology*, 200241.

Kumar S., Stecher G., Li M., Knyaz C., and Tamura K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547-1549.

Lai, C.-C., Shih, T.-P., Ko, W.-C., Tang, H.-J., & Hsueh, P.-R. (2020). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virus disease-2019 (COVID-19): the epidemic and the challenges. *International Journal of Antimicrobial Agents*, 105924.

Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., … Lessler, J. (2020). The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Annals of Internal Medicine*.

Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., … Wong, J. Y. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England Journal of Medicine.*

Nei M. and Kumar S. (2000). Molecular Evolution and Phylogenetics. *Oxford University Press*, New York.

Ou X, Liu Y, Lei X, Li P, Mi D, Ren L, Guo L, Guo R, Chen T, Hu J, Xiang Z, Mu Z, Chen X, Chen J, Hu K, Jin Q, Wang J &Qia Z, (2020) Characterization of spike glycoprotein of

SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nature Communications*. https://doi.org/10.1038/s41467-020-15562-9

Rabi F, Al Zoubi M, Kasasbeh G, Salameh D, & Al-Nasser A (2020). SARS-CoV-2 and Coronavirus Disease 2019: What We Know So Far. *Pathogens* 9(3), 231

Savarino, A., Boelaert, J. R., Cassone, A., Majori, G., & Cauda, R. (2003). Effects of chloroquine on viral infections: an old drug against today's diseases. *The Lancet Infectious Diseases*, 3(11), 722–727.

Schoeman D & Fielding B (2019) Coronavirus envelope protein: current knowledge. *Virology Journal* 16:69

Snijder, E. J., Van Der Meer, Y., Zevenhoven-Dobbe, J., Onderwater, J. J. M., van der Meulen, J., Koerten, H. K., &Mommaas, A. M. (2006). Ultrastructure and origin of membrane vesicles associated with the severe acute respiratory syndrome coronavirus replication complex. Journal of Virology, 80(12), 5927–5940.

Touret, F., & de Lamballerie, X. (2020). Of chloroquine and COVID-19. *Antiviral Research*, 104762.

Wang H, Yang P, Liu K, Guo F, Zhang Y, Zhang G, Jiang C (2008) SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway. *Cell Research* 18:290-301.

WHO (2020). Coronavirus disease 2019 (COVID-19) situation report—56.

World Health Organization (2020) Coronavirus disease (COVID-19) technical guidance: Laboratory testing for 2019-nCoV in humans. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/laboratory-guidance

Wu, J. T., Leung, K., & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*, 395(10225), 689–697.

Xu, X. et al. (2020) 'Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission', China *Life Sciences*, 63(3), pp. 457–460.

Yang, D. and Leibowitz, J. L. (2015) 'The structure and functions of coronavirus genomic 3' and 5' ends', *Virus Research. Elsevier* B.V., 206, pp. 120–133.

Zhao, S., Lin, Q., Ran, J., Musa, S. S., Yang, G., Wang, W., … He, D. (2020). Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *International Journal of Infectious Diseases, 92, 214–217.*

Zhou, P. *et al.* (2020) 'A pneumonia outbreak associated with a new coronavirus of probable bat origin', *Nature*.579(7798): 270–273.