Consequences of Mutations in Severe Acute Respiratory Syndrome Coronavirus 2 (Sars-Cov-2) Genome in Comparison to Other Pathogenic Coronaviruses

Yusuf Amuda Agabi, Ismaila Shittu, Kennedy Iliya Amagon, James Garba Damen, Richard Joseph Kutshik, Ukpe Ajima, Bukata Bayero Bukar, Kakjing Dadul Falang, Ishaya Yohanna Longdet, Simji Samuel Gomerep, Stephen Daniel Davou, Jacob Adegboyega Kolawole, and Noel Nen’man Wannang

Abstract

Background: Late December 2019, an unknown incidence of Pneumonia was observed among some residents of Wuhan city, China. The disease named coronavirus disease 2019 (COVID-19) and declared as a pandemic by the WHO on the March 11th, 2020 by the World Health Organization (WHO) has resulted to the death of million people across the globe. Prior to the current COVID-19 pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), two other outbreaks of coronaviruses namely severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV) have been experienced within the last few decades. This review looks at the unique characteristics of SARS-CoV-2 to the other coronaviruses (SARS-CoV and MERS-CoV) and its significance(s) in the control strategies including diagnostics.

Materials and Methods: Using the keywords “coronavirus mutation”, “nucleotide substitution”, “coronavirus evolution”, “SARS-CoV-2”, “COVID-19” published literatures on coronaviruses and SARS-CoV-2 were retrieved from MEDLINE and reviewed for gaps and current knowledge as it relates to evolution of SARS-CoV-2.

Results: In comparison with seasonal flu, investigations revealed that SARS-CoV-2 mutates less rapidly which provides an edge in the possible development of a long-lasting vaccine to combat the spread of the virus. Though, several mutations in the genome of the virus with dire consequences on the diagnostics have been identified.

Conclusions: The unique importance of mutation as a mechanism of survival for viruses cannot be overemphasized. Several mutations have been observed in SARS-CoV-2 genome whose implications as regards diagnostics and control measures have been highlighted herein.

Keywords: Coronaviruses, SARS-CoV-2, Genetic mutations, Control strategies, Diagnostic.
I. INTRODUCTION

Late December 2019, an unknown incidence of Pneumonia was observed among some residents of Wuhan city, China. Investigation revealed the aetiological agent as a 2019 novel Coronavirus (2019-nCoV) which was later changed to Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1], [2]. Subsequently, the virus spread to several countries across the globe and was declared a pandemic by the World Health Organization (WHO) on the March 11th, 2020 [3]. As of 17th November 2020, approximately 55 million cases of Coronavirus disease 19 (COVID-19) have been reported to WHO across the world with the Americas leading with over 23 million [4]. As the virus spreads from continent to continent, the case fatality rate reported in different countries differs considerably [5]. In sub-Saharan Africa, the fatality rate was lower than projected [6]. In the last few decades, similar outbreaks of respiratory infections caused by coronaviruses namely severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV) have been reported [7]. These outbreaks have resulted to loss of millions of lives as well as creating public health concerns across the globe. Aside the death and public health concerns caused by the current COVID-19 outbreaks, livelihood and economies of countries have been affected due to movement restrictions and lockdowns introduced as a way of curtailing the spread of the disease.

Coronaviruses are enveloped, positive-sense, single-strand RNA viruses with wide host range which include but not limited to avian, mammalian and human. Bats are the natural reservoir of most coronaviruses, there is no substantial evidence so far that SARS-CoV-2 originated in or was transmitted from a seafood market [8]. Based on phylogenetic analysis, SARS-CoV-2 is classified as a member of the genus Sarbecovirus, subfamily Orthocoronavirinae, family Coronaviridae and order Nidovirales [9]. Most members of this family are zoonotic viruses transmitted to humans through contact with infected animals. RNA viruses are characterized by a high mutation rate, up to a million times higher than that of their hosts. Mutation rate drives viral evolution and genome variability, thereby enabling viruses to escape host immunity and to develop drug resistance [10]. A recent report identified specific mutations in the receptor binding domain (RBD) of spike protein which is the most variable part in coronavirus genome [11].

There is increasing epidemiological and clinical evidence that suggests the SARS-CoV-2 has a stronger transmissibility than SARS-CoV and lower pathogenicity. However, the mechanism of high transmission of SARS-CoV-2 is unclear. DNA sequence comparisons using single nucleotide polymorphisms (SNPs) are often used for evolutionary studies and can be especially beneficial in recognizing the mutated coronavirus genomes, where high mutations can occur due to an error prone RNA-dependent RNA polymerase in genome replication. To understand the virus evolution of SARS-CoV-2 from the genome mutation context, this study establishes the SNP genotyping method and investigate the genotype changes during the transmission of SARS-CoV-2. The results show that the genotypes of the virus are not uniformly distributed among the complete genomes of SARS-CoV-2. This genotyping study discovers a few highly frequent mutations in the SARS-CoV-2 genomes. The highly frequent SNP mutations might be associated with the changes in transmissibility and virulence of the virus. The mutations are located in the spike (S) protein, RNA polymerase, RNA, primase, and nucleoprotein, which are fundamental proteins for vaccine efficacy. Therefore, the high-frequency SNP mutations are important factors when developing vaccines for preventing the infection of SARS-CoV-2 [12].

Being the first pandemic associated with coronavirus, close monitoring of the different mutations is of paramount importance to track the evolution of the virus and devise appropriate control measures to mitigate its spread and impact on public health. This review looks at the different mutations identified in SARS-COV-2 and its implication in the disease control including diagnostics.

II. MATERIALS AND METHOD

In writing the review, information was sourced from the MEDLINE using the keywords “coronavirus mutation”, “nucleotide substitution”, “coronavirus evolution”, “SARS-CoV-2”, “COVID-19”.

III. REPLICAION OF SARS-COVT AND OTHER CORONAVIRUSES

Different researchers in China have reported that SARS-CoV-2, just like SARS-CoV, requires the angiotensin converting enzyme 2 (ACE2) [13] as a receptor to enter cells [14]. The binding of the virus with host cell receptors is a significant determinant for the pathogenesis of infection. The possible origin of SARS-CoV can be traced back to bats. Invariably, the virus adapted to non-bat ACE2 variants as it crossed the species barrier to infect humans [15]. Dipeptidyl peptidase 4 (DPP4, also known as CD26) was identified as a functional receptor for MERS-CoV, because the receptor-binding S1 domain of the MERS-CoV spike protein was copurified with DPP4 specifically from lysates of susceptible Huh-7 cells [16]. MERS-CoV can bind DPP4 from multiple species, which promotes the transmission to humans, and other species, and infection of cells from a large number of species [17]. A better understanding of the relative effects of receptor binding and protease action will help predict whether specific zoonotic coronaviruses infect humans and the possibility of adaptation [15]. Importantly, the sequence of the receptor-binding domain of SARS-CoV-2 spikes is similar to that of SARS-CoV. Data from previous studies strongly suggests that entry into the host cells is most likely via the ACE2 receptor [18], [19].

IV. STRUCTURE OF SARS-COVT AND OTHER CORONAVIRUSES

Structural analysis shows that SARS-CoV-2 probably derives from a bat SARS-like coronavirus, which has mutated in the spike glycoprotein (protein S) and nucleocapsid N protein. The positive-sense RNA genomes of SARS-CoV-2 differ from SARS-CoV and MERS-CoV, being approximately 29.9 kb, 27.9 kb and 30.1 kb, respectively.
The SARS-CoV-2 complete genome was annotated to possess 14 open reading frames (ORFs) encoding 27 proteins. Sequence analysis revealed that SARS-CoV-2 has greater than 80% identity with SARS-CoV and 50% with MERS-CoV, which originated in bats. In addition, the spherical external S protein displays a characteristic crown shape with electron microscopy [8]. Genes are involved in proteolysis, replication and transcription of virions, thus, serve as potential drug targets. Of note amongst these gene in SARS-CoV-2 and other coronaviruses include main protease or 3C-like protease (Mpro or 3CLpro), papain-like protease (PLpro), non-structural protein 12 (nsp12) and RNA-dependent RNA polymerase (RdRp) [20]. In general, the envelope (E), membrane (M), spike (S), and the nucleocapsid protein (N) (Fig. 1) are responsible for the biosynthesis of coronavirus polymerases and structural genes.

A number of genes encoded by SARS-CoV such as ORF3a, -3b, -6, -7a, -7b, -8a, -8b, and -9b [21] function as accessory genes which are found in the 3' loci of the viral genome (Fig. 1). However, SARS-CoV-2 encodes additional accessory genes namely: ORF1a, ORF1b, ORF6, ORF7a, ORF7b, ORF8, Nucleocapsid (N), and ORF10 [20], [22]. Studies have shown that there is a high degree of structural similarity between closely related coronaviruses such as SARS-CoV-2, SARS-CoV and MERS-CoV (Fig 1). However, the spike glycoprotein of SARS-CoV and SARS-CoV-2 differs remarkably from each other resulting in the structural difference which impacts on the variable affinity of the RNA-binding domain (RBD) with ACE2, the receptor binding domain in the host cell. As a result antibodies and therapeutic molecules known to be previously potent against SARS-CoV are not effective against SARS-CoV-2 [23], [20]. Therefore, a gap in the understanding of the structure and function of SARS-CoV-2 RBD must be addressed in order to develop novel antiviral drugs to curb transmission of SARS-CoV-2. RdRP is notably the core enzyme of polyprotein replicase-transcriptase complex (RTC) required for transcription as well as replication of coronaviruses. Therefore, antiviral drug discovery explores the inhibition of RdRp and Helicase by blocking or destroying viral proteases and consequently stopping replication and transcription of the viral ribonucleic acid [24].

Considering the structural genes of SARS-CoV-2 and other coronaviruses, the S glycoprotein is primarily involved in pathogenesis as it begins the process of recognition, attachment, and binding to host cells via the RBD. The RBD of S protein is the most variable part of the SARS-CoV-2 and comprises of subunits 1, 2 and 3'. Subunit 1, S1 serves as an attachment spike to the host cell through the ACE2 receptor which facilitates the beginning of pathogenesis while S2 serve as fusion protein that fuses the virion with the mammalian cell membrane [25]. Subsequently, the cleavage of S2 domain (S2’ site) takes place in order to release the fusion peptide which in turn triggers the activation of the membrane fusion mechanism [22], [26], [27].

The E protein is made up of approximately 76 to 109 amino acids being the smallest of the structural proteins and helps with virus assembly and release of new viral particles [28], [22]. The envelope protein covers the genetic material of the virus. The E protein of SARS-CoV-2 may serve as good target for drugs and therapeutic peptides. The envelope protein of BAT-CoV, SARS-CoV, and SARS-CoV-2 is structurally similar but the envelope of MERS-CoV shows some sequence dissimilarity [22]. The M protein consists of approximately 222 amino acids with three transmembrane domains and reported as the most common glycoprotein. The M protein helps with the intracellular processes involving synthesis of virus particles such as RNA packaging, giving a unique morphology, the crown shape to coronaviruses. Additionally, the M protein is highly conserved among coronaviruses [29], [30]. The N protein on the other hand is involved in the replication and transcription of viral particles including packaging of viral genome into ribonucleocapsid [22], [31]. The N protein is suggested as an important site for antiviral drug activity. The N protein shows a high sequence similarity among coronaviruses, thus, antibodies against N proteins of SARS-CoV may be effective against N proteins of SARS-CoV-2 [22]. Another important structural protein in coronaviruses is Replicase multiprotein which helps to transcribe and replicate viral RNA using the host cell RNA as template in addition to many other roles it plays to facilitate the virulence and infectivity of viruses [22]. The replicase gene is involved in the coding of RdRp and a number of non-structural proteins (nsps) [24] all of which play a role in replication and transcription.

**fig. 1:** Genome organization of Coronaviruses: a) severe acute respiratory syndrome coronavirus (sar-cov); b) severe acute respiratory syndrome coronavirus 2 (sar-cov-2); c) middle east respiratory syndrome coronavirus (mers-cov).
immune responses, gene to gene activation and regulation. Sixteen nsps (nspl-nsp16) have been reported to be encoded by most of the coronaviruses except Gammacoronaviruses which lack nspl [22], [30].

Taken together, the understanding of the structure and function of the different proteins in coronaviruses would help in the design and development of drugs and vaccines towards the prevention and management of COVID-19 pandemic.

V. HIGHLIGHTS OF DIFFERENT MUTATIONS IN SARS-COVID-2 AND THEIR IMPLICATIONS ON DIAGNOSTIC AND DISEASE CONTROL

Due to selective pressure, viruses tend to mutate to adapt to its host. A mutation, when it occurs, is either neutral or harmful to the virus. In nature, viral evolution occurs as a way of survival for the virus and nucleotide substitution accounts for one of the mechanisms deployed by the virus. The hallmark of RNA viruses is their ability to mutate rapidly while replicating in their hosts. This phenomenon is due to the infidelity of the RdRp involved in making copies of the virus during replication. Influenza A viruses mutate more rapidly than coronaviruses. This is due to the presence of proofreading machinery in the latter [32]. These mutations have been responsible for the evolutionary and genetic variability observed in the viruses over the years. Invariably, this has impacted on the pathogenicity, tropism, and transmissibility of the virus. Since the onset of the current COVID-19 pandemic, several researches have been conducted on SARS-CoV-2 virus with variations observed across the genome. Notably, [33] identified 13 mutations in the open reading frame (ORF)1a, 1b, 3a, 8, S, N and M gene (Fig. 1). Also, 8 repeated mutations were observed at different positions in SARS-CoV-2 genomes from sequences recovered from patients in Europe and North America [10]. In a study conducted on over 31,000 SARS-CoV-2 genomes retrieved from the genome repositories, it was revealed that the diagnostic protocols used for COVID-19 have undergone mutations mostly at the N-gene primers and probe region [34]. This has a dire consequence on the diagnosis of the disease and calls for constant monitoring of the primers and probes used for COVID-19 diagnosis or search for alternative method of diagnosis. Recent studies have shown the prevalence of SARS-CoV-2 with mutation (D614G) at the S protein [35]. This mutation which was observed in the early part of the pandemic has been reported to be widespread in recent SARS-CoV-2. Viruses with the mutation (D614G) have been reported to have higher transmissibility and higher viral loads in infected patients [35], [36].

VI. CONCLUSION

By their very nature mutations are a random event, and so predicting when and where they will occur and the impact this might have is a very difficult task. However, monitoring the occurrence of mutations is not just important for understanding how changes in the virus could alter its lethality or transmissibility, but also how they could affect the diagnostic tools we are using in response. Therefore, whether SARS-CoV-2 will evolve to become less lethal over time remains to be seen. Currently, the number of complete genomes of the SARS-CoV-2 in the genetic data repositories is still on the increase. Genetic information available from Africa on SARS-CoV-2 is still low and this needs to be improved upon in order to have a better understanding of the dynamics of the virus on the Africa continent. Fortunately, the Africa Centre for Disease Control recently identified three centres of excellence on genomic for the continent. The establishment of these centres is expected to help provide additional genomic data on the virus in Africa. As more genomic sequences of SARS-CoV-2 are made available there is possibility of identification of novel mutations hitherto unreported as recently observed in Turkey [11]. To keep abreast of the virus, analysis of new sequences for possible deleterious mutations needs to be monitored.

VII. CONFLICT OF INTEREST

The authors declare no conflict or competing interest on the manuscript.

REFERENCES

[1] D.S. Hui, E.J.A. Madani, T.A. Ntoumi, F. Kock, R.O. Dar. The continuing 2019- nCoV epidemic threat of novel coronaviruses to global health – the latest 2019 novel coronavirus outbreak in Wuhan. China. International Journal of Infectious Diseases; 91:264–6, 2020. doi.org/10.1016/j.ijid.2020.01.009 12.

[2] R. Lu, X. Zhao, J. Li, P., Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. Lancet. 395:1022(4):565-574, 2020. doi: 10.1016/S0140-6736(20)30251-8.

[3] Archived: WHO Timeline - COVID-19. https://www.who.int/newsroom/detail/27-04-2020-who-timeline—covid-19.

[4] Coronavirus disease (COVID-19) pandemic. https://www.who.int/emergencies/diseases/novel-coronavirus-2019.

[5] J.B. Dowd, L. Andriano, D.M. Brazel, V. Rotondi, P. Block, X. Ding, Y. Liu, and M.C. Mills. Demographic science aids in understanding the spread and fatality rates of COVID-19. Proceedings of the National Academy of Sciences; 117(18): 9696-9698, 2020. doi: 10.1073/pnas.2004911117.

[6] S. Ghisolfi, I. Almás, J.C. Sandefur, T. von Carnap, J. Heitner, and T. Bold. Predicted COVID-19 fatality rates based on age, sex, comorbidities and health system capacity. BMJ Global Health, 2020;5:e003094. doi:10.1136/bmjgh-2020-003094.

[7] J. Cui, F. Li and Z.L. Shi. Origin and evolution of pathogenic coronaviruses. Nature Review Microbiology; 17(3):181-192, 2019. doi: 10.1038/s41579-018-0118-9. PMID: 30531947.

[8] A. A. Dawood. Mutated COVID-19 may foretell a great risk for mankind in the future. New Microbes and New Infections. 35: 100673, 2020. doi: 10.1016/j.jmn.2020.100673.

[9] Coronavirusidae.https://talk.ictvonline.org/taxonomy/p/taxonomyhistory?taxnode_id=201906129.

[10] R. Lu, X. Zhao, J. Li, P., Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. Lancet. 395(10224):565-574, 2020. doi: 10.1016/S0140-6736(20)30251-8.

[11] Y. Yin, and R.G. Wunderink. MERS, SARS and other coronaviruses as causes of pneumonia. Respiratory. 23(2):130-137, 2018. doi: 10.1111/res.13196.

[12] P. Zhou, X.L. Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, H.R. Si, Y. Zhu, B. Li, C.L. Huang, H.D. Chen, J. Chen, Y. Luo, H. Guo, R.D. Jiang, M.Q. Liu, Y. Chen, X. R. Shen, X. Wang, Zheng, X.S., K. Zhao, Q.F. Chen, F. Deng, L. L. Liu, B. Yan, F.X. Zhan, Y.Y. Wang, G.F. Xiao, and Z.L. Shi. A pneumonia outbreak associated with a new...
coronavirus of probable bat origin. *Nature*, 579(7798): 270-273, 2020. doi:10.1038/s41586-020-0212-7.

[14] W. Li, M.J. Moore, and N. Vasiliyeva. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature*, 426(6965):450-454, 2003. doi: 10.1038/nature02145.

[15] H. Li, L. Geng, Y. Peng, L. Meng, S. Lu. Molecular Immune Pathogenesis and Diagnostics of COVID-19. *Journal of Pharmaceutical Analysis*; 10(2):102-108, 2020. doi: 10.1016/j.jpha.2020.03.001.

[16] V.S. Raj, H. Mou, and S.L. Smits. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature*, 495: 251-254, 2013. doi:10.1038/nature12005.

[17] A. Barlan, J. Zhao, M.K. Sarkar, K. Li, P.B. McCray, S. Perlman, and T. Gallagher. Receptor variation and susceptibility to Middle East respiratory syndrome coronavirus infection. *Journal of Virology*; 88(9):4953-61, 2014. doi:10.1128/JVI.00161-14.

[18] S. Belouzard, V.C. Chu, and G.R. Whittaker. Activation of the SARS coronavirus spike protein by sequential proteolytic cleavage at two distinct sites. *Proceedings National Academy Science*; 106: 5871-5876, 2009. doi:10.1073/pnas.0809524106.

[19] H.A. Rothan, and S.N. Byrareddy. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *Journal of Autoimmunity*; 109:102433, 2020.

[20] A.A.T. Naqvi, K. Fatima, T. Mohammad, U. Fatima, I.K. Singh, A. Singh, S.M. Atif, G. Hariprasad, G.M. Hasan, M.I. Hassan. Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach. *Biochimica et Biophysica Acta -Molecular Basis of Disease*; 1866(10):165878, 2020. doi:10.1016/j.bbadis.2020.165878.

[21] S.R. Schaecher, and A. Pekosz. SARS coronavirus accessory gene expression and function. In Lal S. *Coronavirus*, Springer Verlag, Berlin, Germany; 2010.

[22] S. Boopathi, A.B. Poma, and P. Kolandavel. Novel 2019 coronavirus structure, mechanism of action, antiviral drug promises and rule out against its treatment. *Journal of Biomolecular Structure and Dynamics*, 2020. doi:10.1080/07391102.2020.1758788.

[23] A. Zumla, J.F. Chan, E.I. Azhar, D.S. Hui, K.Y. Yuen. Coronaviruses - drug discovery and therapeutic options. *Nature Reviews Drug Discovery*; 15(5):327-47, 2016. doi:10.1038/nrd.2015.37.

[24] A. Gaurav, and M. Al-Nema. Polymerases of Coronaviruses: Structure, Function, and Inhibitors in: *Viral Polymerases*; pp:271-300, 2019. doi: https://doi.org/10.1016/B978-0-12-815422-9.00010-3.

[25] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Kruger, T. Herrler, S. Erichsen, T.S. Schiergens, G. Herrler, N.H. Wu, A. Nitsche, M.A. Muller, C. Drosten, S. Pohlmann. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*; 181(2): 271-280, 2020. doi:10.1016/j.cell.2020.02.052.

[26] J.F. Chan, C.C. Yip, K.K. To, T.H. Tang, S.C. Wong, K.H. Leung, A.Y. Fung, A.C. Ng, Z. Zou, H.W. Tsui, G.K. Choi, A.R. Tam, V.C. Cheng, K.H. Chan, O.T. Tsang, and K.Y. Yuen. Improved molecular diagnosis of COVID-19 by the novel, highly sensitive and specific COVID-19-RdRp/Hel real-time reverse transcription-polymerase chain reaction assay validated in vitro and with clinical specimens. *Journal of Clinical Microbiology*; 58(5): e00310-20, 2020. doi:10.1128/JCM.00310-20.

[27] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.L. Hsieh, O. Abona, B.S. Graham, J.S. McLellan. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*; 367: 1260–1263, 2020.

[28] D. Schoeman, and B.C. Fielding. Coronavirus envelope protein: current knowledge. *Virolology Journal*; 16 (69). doi:10.1186/s12985-019-1182-0, 2019.

[29] L. Moussavizadeh, and S. Ghaseemi. Genotype and phenotype of COVID-19: Their roles in pathogenesis. *Journal of Microbiology, Immunology and Infection*. doi:10.1016/j.jmi.2020.03.022. PMID: 32265180, 2020.

[30] T. Tang, M. Bidon, J.A. Jaimes, G.R. Whittaker, S. Daniel. Coronavirus membrane fusion mechanism offers a potential target for antiviral development. *Antiviral Research*; 104792. doi:10.1016/j.antiviral.2020.104792, 2020.

[31] B.W. Neuman, and M.J. Buchmeier. Supramolecular architecture of the coronavirus particle. *Advances in Virus Research*; 96: 1–27, 2016. doi:10.1016/bs.avir.2016.08.005.

[32] E.C. Smith, H. Blanc, M. Vignuzzi, and M.R. Denison. Coronaviruses Lacking Exoribonuclease Activity Are Susceptible to Lethal Mutagenesis: Evidence for Proofreading and Potential Therapeutics. *PLoS Pathogen* 9(8): e1003565, 2013. https://doi.org/10.1371/journal.ppat.1003565, 2013.

[33] C. Wang, Z. Liu, Z. Chen, X. Huang, M. Xu, T. He, Z. Zhang. The establishment of reference sequence for SARS-CoV-2 and variation analysis. *Journal of Medical Virology*, https://doi.org/10.1002/jmv.25762, 2020.