REVIEW ARTICLE
Volume - 9 | Issue - 12 | December - 2020 | PRINT ISSN No. 2277 - 8179 | DOI : 10.36106/ijsr
INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

GENOMICS OF SARS-COV-2: A STUDY

Healthcare

Rithwik Rajesh B.Tech (Genetic Engineering), SRM Institute of Science and Technology, Kattankulathur.

ABSTRACT

SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) is a type of novel coronavirus responsible for the COVID-19 outbreak, which attained pandemic proportions in 2020. SARS-CoV-2 belongs to that group of coronaviruses which infect the lower respiratory tract. The complete genome was found to be 29.9kb in size. The viral genome encodes 16 non-structural proteins (Nsps), essential for pathogenesis and replication. Structural proteins- envelope (E), membrane (M), nucleocapsid (N), and spike (S) glycoprotein are also encoded, important for virus subtyping and response to vaccines.

L and S type of SARS-CoV-2 were identified in Wuhan initially. Later on, as the pandemic spread across the globe, A and C subtypes were identified to me the common type of strains across the Americas and Europe respectively. In Asia, B subtype was found to be common. ACE2 (angiotensin-converting enzyme 2), a receptor is utilised by SARS-CoV-2 for entry into host cell. People of European descent have a higher frequency of alternate allele of rs763395248 SNP in T92I risk variant compared to African and Asian population. In Asian populations, SNPs like rs758278442 and rs759134032 in the region of protective variants (K31R and Y83H) of ACE2 gene show higher frequency of mutant alleles than American and European populations. According to a structural modelling study, certain ACE2 variants may provide potential resistance to SARS-CoV-2 infection.

KEYWORDS

ACE2, SARS-CoV-2, Cell membrane receptors, Corona virus, Genomic structure, Viral variants

INTRODUCTION

COVID-19 is a disease that took the world by storm. Even more than the fast transmission rate and deadliness, what perplexed healthcare workers, researchers and laymen was the fact that not much was known about the virus SARS-CoV-2 during the initial stages of the pandemic.

COVID-19 is unique for the fact that after a very long time, a disease took the world by surprise. For many decades, human beings were convinced that most pathogens in this world were already known and that most diseases, were containable and even if it wasn't, it could be prevented or cured after a short span of research. The novel coronavirus, now known as SARS-CoV-2 is reported to have its origin in Wuhan, China.

There is significant research into the pathogenesis of the disease and novel corona virus is happening all over the world, with over one billion publications till date, with over 95 million about the genomics of SARS-CoV-2 virus alone. A comprehensive understanding of the genome is critical in understanding the virus better and devising treatment plans to tackle the disease.

This article provides a focused review on the genome aspect of the virus, with special attention to the replication, epidemiology and transmission as well as on the ACE2 (angiotensin-converting enzyme 2), which is an important receptor that the virus utilises to gain entry into the host cell. A narrative review of 33 recent articles on SARS-Cov-2 virus were selected and compared in aspects of genomics, replication, epidemiology, transmission, ACE 2 receptor and its clinical implications were done.

Genomics of SARS-CoV-2

Seven virus species are known as of now, within which, four are known to cause diseases that show symptoms milder in nature because they infect the upper respiratory tract. The other three, which includes the SARS-CoV-2 cause more severe diseases as they infect the lower respiratory tract [1].

SARS-CoV-2 was found to be very similar in genome structure to SARS-CoV and MERS-CoV viruses. Compared with the known SARS-CoV and MERS-CoV genome, SARS-CoV-2 is closer to the SARS-like bat CoVs in terms of the whole genome sequence. Most genomic encoded proteins of SARS-CoV-2 are similar to SARS-CoVs, but also exhibit certain differences. The complete genome of Wuhan-Hu-1 coronavirus (WHCV), one strain of SARS-CoV-2, was found to be 29.9kb [2], compared to SARS-CoV and MERS-CoV, which have positive-sense RNA genomes of 27.9kb and 30.1kb, respectively [3]. A variable number (6–11) of open reading frames (ORFs) were found in the genome of CoVs[4].

16 non-structural proteins (Nsps) are coded by the viral genome. These are essential for pathogenesis and replication. The genome also encodes for four structural proteins- envelope (E), membrane (M), nucleocapsid (N), and spike (S) glycoprotein. These are important for virus subtyping and response to vaccines, and nine other accessory factors [5,6]

Bats have been identified as the primary reservoir of SARS-like coronaviruses after the phylogenetic analyses of SARS-CoV-2 genomes were conducted [7]. High sequence similarity was identified between BatCoV and SARS-CoV-2 genomes (96.2%) [8]. SARS-CoV-2 also showed a 91.02% identity in genome sequences when compared to the whole genome of Pangolin-CoV, which was sequenced from a dead Malayan Pangolin (Manis javanica) [9].

Since SARS-CoV-2 has been reported from many different parts of the world, differences in viral strains have been identified. During the initial months, two major subtypes- L and S had been reported from Wuhan, the origin of the outbreak, 70% of all the reported strains were found to be the L subtype. L was also identified to be more deadly and aggressive than its counterpart, the S type [10]. Using a phylogenetic network analysis approach on 160 full-length genomes, a recent study has shown that the virus seems to be evolving into three distinct clusters, with A being the ancestral type closest to the bat genome and found mostly in Americas and Europe along with the C type, while B being the most common type in east Asia [11].

Replication

S glycoprotein has S1 and S2 subunits [11]. Virus-host range and cellular tropism with the key function domain – RBD (receptor-binding domain) is determined by S1. S2 mediates virus-cell membrane fusion by two tandem domains, heptad repeats 1 (HR1) [13] and HR2 [14]. After the membrane is fused, the viral genome RNA is released into the cytoplasm. Then, two polyproteins, pp1a and pp1ab are translated by the uncoated RNA [15]. They encode non-structural proteins, and is responsible for the formation of replication-transcription complex (RTC) in double-membrane vesicle [16]. A nested set of subgenomic RNAs are then synthesised after the continuous replication of RTC [17], which encode accessory proteins and structural proteins. Mediating endoplasmic reticulum (ER) and Golgi [18], newly formed genomic RNA, nucleocapsid proteins and envelope glycoproteins assemble and form viral particle buds. The virion-containing vesicles fuse with the plasma membrane to release the virus.[19]

Systematic detection of β-CoV receptors showed that human cells expressing ACE2, but not human Dipeptidyl peptidase-4 (DPP4) or APN (Aminopeptidase N), were enhanced entry of SARS-CoV-2 [20]. While, another study showed that S protein and ACE2 binding efficiency is 10- to 20- fold higher than that of SARS-CoV, evidenced by Cryo-EM Structure of the SARS-CoV-2 Spike in the prefusion

conformation [21]. di Viti et al suggests that the SARS-CoV-2 may be less virulent than the SARS-CoV and MERS-CoV, with the currently analyzed mortality of COVID-19 is 3.4%, lower than death rate of SARS (9.6%) and MERS (around 35%), respectively [22].

Role of ACE2 gene

SARS-CoV-2 uses angiotensin-converting enzyme 2 (ACE2), which serves as a receptor for SARS-CoV for viral entry into the host cell [23,24]. The ACE2 gene is located in X chromosome; potential functional variants of ACE2 gene are shown to alter its transcriptional activity [25].

Varying allele frequencies are exhibited by some single nucleotide polymorphisms (SNPs) located in the coding regions of ACE2 gene [26]. Compared to a global average of American and European population, SNPs like rs758278442 and rs759134032 in the region of protective variants (K31R and Y83H) of ACE2 gene show relatively higher frequency of mutant alleles in Asian population. However, compared to African and Asian population, people of European descent have a higher frequency of alternate allele of rs763395248 SNP in T92I risk variant. Certain ACE2 variants might provide potential resistance to SARS-CoV-2 infection according to a structural modelling study [27]. Additionally, the expression quantitative trait loci (eQTL) analysis of ACE2 variants revealed association of some eQTLs with higher expression of ACE2 in tissues and few eQTLs had higher allele frequencies in East Asian Populations than the European populations [26].

Significant allele frequency differences of the missense mutations (K26R, 1468V, N720D, and N638S) of ACE2 gene were detected in various populations. Within these, 1468V mutated more frequently in Asians, whereas K26R mutated more frequently among Caucasians. Variable ACE2 gene expression was reported among different populations, but the magnitude was found to be very minute [28]. However, Chen et al noted that people of Asian descent exhibit similar ACE2 expression like other groups and that no significant differences in gene expressions were reported [29].

Clinical implications of ACE2 Gene

The most common comorbid conditions in COVID-19, Hypertension and Diabetes mellitus (DM) are both are modulated by ACE2 [30]. ACE2 plays a major role in inflammatory processes [31]. Genetic deficiency of ACE2 upregulates the expression of cytokines and induces vascular inflammation in ApoE knockout mouse model [32].Multiple immune signatures such as markers of T cells, B cells, NK cells, and interferon response across various human tissues were found to be associated with ACE2 expression [33]. Therefore, these findings can imply that ACE2 is also involved in mediating the post-infection downstream processes including inflammatory responses, rather than just acting as a receptor for SARS-CoV-2 [25].

CONCLUSION

Genome structure of SARS-CoV-2 is very similar to that of SARS-CoV and MERS-CoV viruses. L and S strains were isolated in Wuhan. Later A, B and C types were isolated, A and C being most common in the Americas and Europe and B being most common in Asia. This data may have helped in the production of country specific or ethnicity specific vaccines.

SNPs like rs758278442 and rs759134032 in the region of protective variants (K31R and Y83H) of ACE2 gene show relatively higher frequency of mutant alleles in Asian population than American and European populations, while people of European descent have a higher frequency of alternate allele of rs763395248 SNP in T92I risk variant compared to African and Asian population. According to a structural modelling study, certain ACE2 variants may provide potential resistance to SARS-CoV-2 infection.

REFERENCES

1. Lu, G.; Wang, Q.; Gao, G.F. Bat-to-human: Spike features determining 'host jump' of coronaviruses SARS-CoV, MERS-CoV, and beyond. Trends Microbiol. 2015, 23, 468–478.
2. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020. https:// doi.org/10.1038/s41586-020-2008-3 [Epub ahead of print].
3. de Wit E, van Doremalen N, Falzarano D, Munster VJ. SARS and MERS: recent insights into emerging coronaviruses. Nat Rev Microbiol. 2016;14(8): 523–34.
4. Song Z, Xu Y, Bao L, Zhang L, Yu P, Qu Y, et al. From SARS to MERS, thrusting coronaviruses into the spotlight. Viruses. 2019;11(1):E59. https:// doi.org/10.3390/v11010059.
5. Kim, D.; Lee, J.Y.; Yang, J.S.; Kim, J.W.; Kim, V.N.; Chang, H. The Architecture of SARS-CoV-2 Transcriptome. Cell 2020, 181, 1–8.
6. Gordon, D.E.; Jang, G.M.; Bouhaddou, M.; Xu, J.; Obernier, K.; White, K.M.; O'Meara, M.J.; Rezelj, V.V.; Guo, J.Z.; Swaney, D.L.; et al. A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug-Repurposing. Nature 2020, in press.
7. Li, W.; Shi, Z.; Yu, M.; Ren, W.; Smith, C.; Epstein, J.H.; Wang, H.; Crameri, G.; Hu, Z.; Zhang, H.; et al. Bats are natural reservoirs of SARS-like coronaviruses. Science 2005, 310, 676–679.
8. Ceraolo, C.; Giorgi, F.M. Genomic variance of the 2019-nCoV coronavirus. J. Med. Virol. 2020, 92, 522–528.
9. Tang, T.; Wu, C.; Li, X.; Song, Y.; Yao, X.; Wu, X.; Duan, Y.; Zhang, H.; Wang, Y.; Qian, Z.; et al. On the origin and continuing evolution of SARS-CoV-2. Natl. Sci. Rev. 2020, in press.
10. Liu, P.; Chen, W.; Chen, J.P. Viral Metagenomics Revealed Sendai Virus and Coronavirus Infection of Malayan Pangolins (Manis javanica). Viruses 2019, 11, 979.
11. Forster, P.; Forster, L.; Renfrew, C.; Forster, M. Phylogenetic network analysis of SARS-CoV-2 genomes. Proc. Natl. Acad. Sci. USA 2020, 117, 9241–9243.
12. Zhang N, Jiang S, Du L. Current advancements and potential strategies in the development of MERS-CoV vaccines. Expert Rev Vaccines. 2014;13(6):761–74.
13. Xia S, Zhu Y, Liu M, Lan Q, Xu W, Wu Y, et al. Fusion mechanism of 2019- nCoV and fusion inhibitors targeting HR1 domain in spike protein. Cell Mol Immunol. 2020. https://doi.org/10.1038/s41423-020-0374-2.
14. Yu F, Du L, Ojcius DM, Pan C, Jiang S. Measures for diagnosing and treating infections by a novel coronavirus responsible for a pneumonia outbreak originating in Wuhan, China. Microbes Infect. 2020. https://doi.org/10.1016/j. micinf.2020.01.003 [Epub ahead of print].
15. de Wilde AH, Snijder EJ, Kikkert M, van Hemert MJ. Host factors in coronavirus replication. Curr Top Microbiol Immunol. 2018;419:1–42.
16. Sawicki SG, Sawicki DL. Coronavirus transcription: a perspective. Curr Top Microbiol Immunol. 2005;287:31–55.
17. Hussain S, Pan J, Chen Y, Yang Y, Xu J, Peng Y, et al. Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus. J Virol. 2005;79(9):5288–95.
18. Perrier A, Bonnin A, Desmarets L, Danneels A, Goffard A, Rouille Y, et al. The C-terminal domain of the MERS coronavirus M protein contains a transGolgi network localization signal. J Biol Chem. 2019;294(39):14406–21.
19. Guo, YR., Cao, QD., Hong, ZS. *et al.* The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak – an update on the status. *Military Med Res* **7**, 11 (2020)
20. Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. Nat Microbiol. 2020. https://doi.org/10.1038/s41564-020-0688-y.
21. Song W, Gui M, Wang X, Xiang Y. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. PLoS Pathog. 2018;14(8):e1007236.
22. de Wit E, van Doremalen N, Falzarano D, Munster VJ. SARS and MERS: recent insights into emerging coronaviruses. Nat Rev Microbiol. 2016;14(8): 523–34
23. Li W, Moore MJ, Vasilieva N, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature. 2003;426(6965):450-454.
24. Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell. 2020;181(2):271-280. e8.
25. Debnath, M, Banerjee, M, Berk, M. Genetic gateways to COVID‒19 infection: Implications for risk, severity, and outcomes. *The FASEB Journal*. 2020; 34: 8787–8795. https://doi.org/10.1096/fj.202001115R
26. Cao Y, Li L, Feng Z, et al. Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. Cell discovery. 2020;6:11.
27. Hussain EW, Diwanji D, Suryamohan K, et al. Structural variations in human ACE2 may influence its binding with SARS-CoV-2 spike protein. J Med Virol. 2020. doi:10.1002/jmv.25832
28. Li Q, Cao Z, Rahman P. Genetic variability of human angiotensin-converting enzyme 2 (hACE2) among various ethnic populations. BioRxiv. 2020. https://doi.org/10.1101/2020.04.14.041434
29. Chen Y, Shan K, Qian W. Asians and other races express similar levels of and share the same genetic polymorphisms of the SARSCoV-2 cell-entry receptor. Preprints. 2020:2020020258. https:// doi.org/10.20944/preprints202002.0258.v1
30. Patel VB, Bodiga S, Basu R, et al. Loss of angiotensin-converting enzyme-2 exacerbates diabetic cardiovascular complications and leads to systolic and vascular dysfunction: a critical role of the angiotensin II/AT1 receptor axis. Circ Res. 2012;110:1322-1335.
31. Gaddam RR, Chambers S, Bhatia M. ACE and ACE2 in inflammation: a tale of two enzymes. Inflamm Allergy Drug Targets. 2014;13:224-234.
32. Thomas MC, Pickering RJ, Tsorotes D, et al. Genetic Ace2 deficiency accentuates vascular inflammation and atherosclerosis in the ApoE knockout mouse. Circ Res. 2010;107:888-897
33. Li M-Y, Li L, Zhang Y, Wang X-S. Expression of the SARS-CoV-2 cell receptor gene ACE2 in a wide variety of human tissues. Infect Dis Poverty. 2020;9:45