Saudi Journal of Biological Sciences 28 (2021) 6803–6807

Contents lists available at ScienceDirect

Saudi Journal of Biological Sciences

journal homepage: www.sciencedirect.com

Original article

Characterization of altered genomic landscape of SARS-CoV-2 variants isolated in Saudi Arabia in a comparative *in silico* study

Mohammad Fahad Ullah [a,b], Tarig M.S. Alnour [a,b,c], Elmutuz H. Elssaig [a,b], Eltayib H. Ahmed-Abakur [a,b,c,*]

[a] Department of Medical Laboratory Technology (FAMS), University of Tabuk, P.O. Box 741, Tabuk 71411, Saudi Arabia
[b] Prince Fahad Research Chair, University of Tabuk, P.O. Box 741, Tabuk 71411, Saudi Arabia
[c] Faculty of Medical Laboratory Science, Department of Microbiology and Immunology, Alzaiem Alazhari University, Khartoum North 11111, Sudan

ARTICLE INFO

Article history:
Received 10 May 2021
Revised 17 July 2021
Accepted 17 July 2021
Available online 22 July 2021

Keyword:
SARS-CoV-2
Covid 19
BLAST
Mutations
Nucleotide

ABSTRACT

SARS-CoV-2 has become one of the unprecedented global health challenge for human population. Genomic signature studies of SARS-CoV-2 reveals relation between geographical location of the isolates and genetic diversity. The present work is an *in silico*, cross sectional study aimed to determine the genetic heterogeneity of SARS-CoV-2 variants isolated in Saudi Arabia compared to the first isolated strain NC_045512 (reference sequence). Each sequence was aligned against the reference sequence using local alignment search tool (NCBI) Nucleotide-BLAST. A total of 58 SARS-CoV-2 genomes were isolated in KSA and retrieved from NCBI. Our study shows that KSA variants demonstrated homology ranging between 99.96 and 99.98 % compared to the reference strain. There are 89 nucleotide changes that have been identified among the KSA variants; the most common nucleotide change was C: T accounting for 50.6% (45/89). These nucleotides changes resulted in 53.9% (48/89) missense mutations and 42.7% (38/89) silent mutations; while the majority of mutations- 48.3% (43/89) occurred in ORF1ab gene. All structural genes displayed mutations; N gene harbored 16.9% (15/89) mutations, S gene displayed 15.7% (14/89) mutations, M gene exhibited 2.2% (2/89) mutations and E gene showed only 1 mutation which was silent. The most frequently changed nucleotide was C3037T (silent mutation) and A23403G (D614G), each of which occurred in 57 variants out of 58 followed by C14408T (P4715L) and C241T (5′UTR) which were found in 56 and 55 variants respectively. The Phylogenetic trees showed that SARS-CoV-2 variants isolated in Saudi Arabia clustered together closely.

1. Introduction

The global pandemic declared on March 11, 2020 as "COVID-19", the disease caused by the SARS-CoV-2 has become one of the unprecedented global health challenge for human population in the recorded history; disturbing the life of billions of people and claiming more than 2.5 million human lives within a year worldwide (Ashour et al., 2020; Seoane 2021). Coronaviruses (CoVs) are enveloped viruses with a single-stranded (positive-sense) ~30 kb RNA genome, typically accommodating six open reading frames (ORFs). Taxonomically classified into four genera; alpha, beta, gamma, and delta CoVs and as evidence suggest only alpha-CoVs and beta-CoVs have the ability to infect humans (Masters, 2006). Beta-CoV genus in recent years has embodied highly pathogenic human CoVs including SARS-CoV, MERS-CoV and currently the SARS-CoV-2, which are associated with high mortality (Ashour et al., 2020; Fehr and Perlman, 2015).

The viral genome of SARS-CoV-2 varies from 29.8 kb to 29.9 kb and the genome structure shares the specific gene characteristics that are common to known coronaviruses (Wang et al., 2020). The 5′ end accommodates more than two-third of the genome containing 5′UTR (265 nt) and orf1ab (21290 nt) that encodes orf1ab polyproteins (processed into 16 non-structural proteins); while the 3′ end which is nearly one third includes 3′UTR (229 nt) and genes that encode structural proteins such as surface–S (3822 nt), envelope-E (228 nt), membrane-M (669 nt), and nucleocapsid-N (908 nt). The SARS-CoV-2 genome also contains

* Corresponding author at: Department of Medical Laboratory Technology (FAMS), University of Tabuk, P.O. Box 741, Tabuk 71411, Saudi Arabia.
E-mail addresses: m.ullah@ut.edu.sa (M.F. Ullah), telnour@ut.edu.sa (T.M.S. Alnour), eelssaig@ut.edu.sa (E.H. Elssaig), eosman@ut.edu.sa (E.H. Ahmed-Abakur).
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.sjbs.2021.07.054
1319-562X/

six accessory proteins, encoded by ORF3a (828 nt), ORF6 (186 nt), ORF7a (366 nt), ORF7b (132 nt), ORF8 (193 nt) and ORF10 (117 nt) genes (Khailany et al. 2020). A minor change in the sequences might change the function of the constituent proteins of the viruses, therefore studying the types of mutation (missense or silent) is essential for understanding variation during evolution, since the missense mutation alters the amino acid (Das et al. 2021). Determining the genomic variation of SARS-CoV-2 is key to understand the pathogenesis, diagnosis, prevention and treatment (Wang et al., 2020), as it helps in analyzing the genetic diversity and pattern of community spread (Khailany et al., 2020).

The outbreak dynamics of Covid 19 varies among regions and countries; this may be due to several factors such environment, cultural practices and public health system that comprises testing rate and surveillance policies (Bandoy and Weimer, 2021). Recently several studies were conducted regionally and globally to determine the mutations of SARS-CoV-2 (Almubaid and Al-Mubaid, 2021; Al-Qaaneh et al., 2021). Considering the history of MERS-CoV which was first identified in Saudi Arabia in 2012, (WHO, 2021) the population dynamics of Saudi Arabia as center of Islamic world and the great numbers of overseas workers, the determination of the nucleotide changes of SARS-CoV-2 in Saudi Arabia (KSA) may be of great interest. Therefore present study aimed to determine the genetic variations of SARS-CoV-2 variants isolated in Saudi Arabia compared to the first isolated strain (Wuhan strain).

2. Method and materials:

This study is an *in silico*, cross sectional study; it is part of project aimed to study the genetic variations of SARS-CoV-2 according to the geographical location and time of isolation. The SARS-CoV-2 sequences were recovered from severe acute respiratory syndrome coronavirus 2 data hub NCBI (https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/virus?SeqType_s=Nucleotide&VirusLineage_ss=SARS-CoV-2,%20taxid:2697049&Country_s=Saudi%20Arabia). The sequence of Wuhan strain SARS-CoV-2 NC_045512 which represent the first characterized isolate (Lu et al., 2020) was used as the reference sequence. Each sequence was aligned against the reference sequence using local alignment search tool (NCBI) nucleotide-BLAST. The nucleotides variations were determined by studying the pairwise alignment and the corresponding mutations were determined using local alignment search tool (NCBI) protein BLAST.

Evolutionary relationships: The Neighbor-Joining method was used to determine the evolutionary relationship in the form of a phylogenetic tree (Raghav et al., 1987). The associated taxa clustering together in the bootstrap test (500 replicates) have been reported as percentage of replicate trees shown next to the branches (Felsenstein, 1985). The Maximum Composite Likelihood method was used to compute the evolutionary distances (Tamura and Kumar, 2004) and represent units of the number of base substitutions per site. This analysis involved two sets of 29 nucleotide sequences of SARS-2 from Saudi Arabia and the reference strain. All ambiguous positions have been removed for each sequence pair (pairwise deletion option). There were a total of 29,903 positions in the final dataset. Evolutionary studies were conducted in MEGA X software (Barry, 2013).

3. Results

Fifty-eight SARS-CoV-2 genomes were isolated in KSA and retrieved from NCBI; these variants represent all SARS-CoV-2 sequences from KSA published in NCBI until the date 02.05.2021. KSA variants showed homology ranging between 99.96 and 99.98 % and no identical strain is detected compared to the reference strain NC045512. A total of 89 nucleotide changes were identified in the KSA variants; among these, the most changed nucleotide was C: T-50.6% (45/89) followed by G: T- 14.6% (13/89) and A: G- 12.4% (11/89) while the less frequent nucleotide change was G: C 1(1.1%)- Fig. 1.

The nucleotides changes resulted in 53.9% (48/89) missense mutations and 42.7% (38/89) silent mutations. Few mutations including 1.1% (1/89) and 2.2% (2/89) occurred in 5′ UTR and 3′ UTR terminal loops respectively. The majority of mutations 48.3% (43/89) occurred in ORF1ab gene; 51.2 % (22/43) of these mutations were missense. ORF3a displayed 9% (8/89) mutations, most of them 87.5% (7/8) were missense mutations while ORF7a showed 3.4% (3/89) mutations (Table 1). All structural genes showed mutations; N gene harbored 16.9% (15/89) mutations, 60% (9/15) of them were missense mutations; S gene displayed 15.7% (14/89) mutations, half of them 50% (7/14) were missense mutations; M gene exhibited 2.2% (2/89) mutations, one of these 50% (1/2) was missense mutation whereas E gene showed only one mutation which was silent (Table 1).

Fig. 2 showed the landscape of SARS-CoV-2 variants representing amino acid changes (missense mutations). Our analysis showed that the most frequently changed nucleotide were C3037T (silent mutation) and A23403G (D614G) each of which occurred in 57 variants out of 58, followed by C14408T (P4715L) and C241T (5′UTR) which were found in 56 and 55 variants, respectively. The coexistence of C241T (5′UTR), C3037T (ORF1ab), C14408T (ORF1ab) and A23403G (S gene) was observed in 55 variants. Other changes which were frequently identified include G25563T (G57H) and C26735T (silent) each of which occurred in 33 variants, followed by C18877T (silent), G28881A (S202N) and G28883C (G204R) which appeared in 32, 23 and 22 variants respectively.

The evolutionary relationship in the form of a phylogenetic tree (Fig. 3a, Fig. 3b) showed that KSA Covid 19 variants clustered together closely; the accession number of each sequence was cited. Number at nodes shows percent bootstrap value above 50 supported by more than 1000 replicates. The bar shows the Jukes-Cantor evolutionary distance.

4. Discussion

The rapid transmission and world wide spread of SARS-CoV-2 reflect its ability to adapt to different environments and conditions. The nucleotides changes during the course of transmission drive the genome variability and viral evolution in a way that affects the transmission, pathogenesis, diagnosis, disease course, and treatment outcome (Almubaid and Al-Mubaid, 2021; Al-Qaaneh et al., 2021; Islam et al., 2020). Therefore the present study

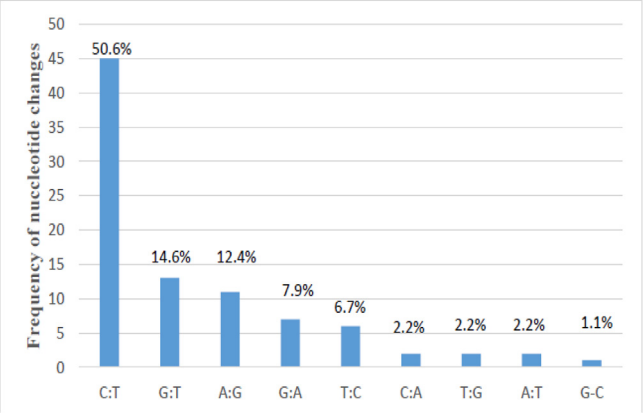

Fig. 1. Types and frequency of nucleotide changes.

Table 1
Type and frequency of mutations according to gene.

Gene	Missense mutations		Silent mutations		Total of the mutation/gene
	Frequency	percentage	Frequency	percentage	
5′UTR					1 (1.1%)
ORF1A/B	22	51.2	21	48.8	43 (48.3%)
S gene	7	50	7	50	14 (15.7%)
ORF3a	7	87.5	1	12.5	8 (9.0%)
E gene	0	0	1	100	1 (1.1%)
M gene	1	50	1	50	2 (2.2%)
ORF7a	2	66.7	1	33.3	3 (3.4%)
N gene	9	60.0	6	40	15 (16.9%)
3′UTR					2 (2.2%)
total	48	100	38		89 (100%)

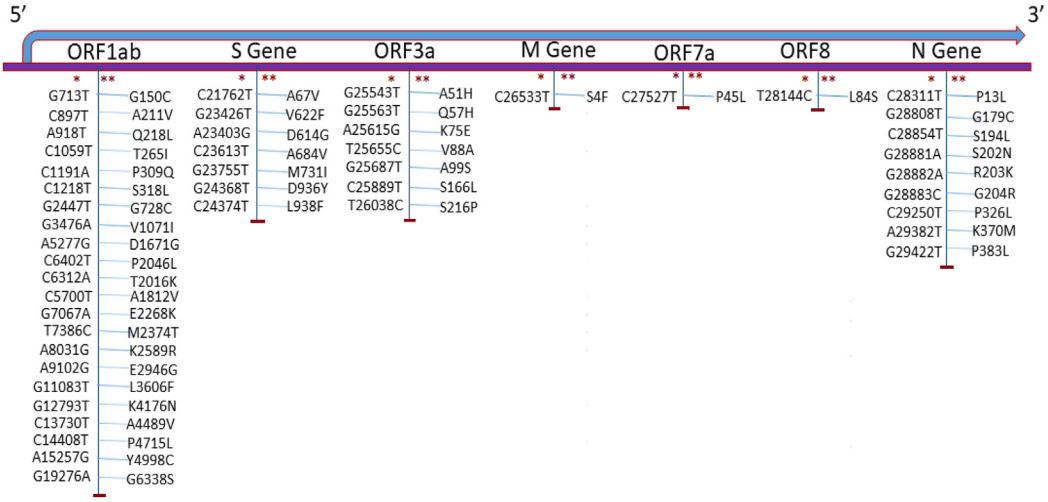

Fig. 2. The genomic landscape (ORF1 a/b, S gene, ORF3a, M gene, orf7a, orf8 and N gene) of KSA variants representing amino acid changes. (*) denotes missense mutations and (**) denotes the corresponding amino acid.

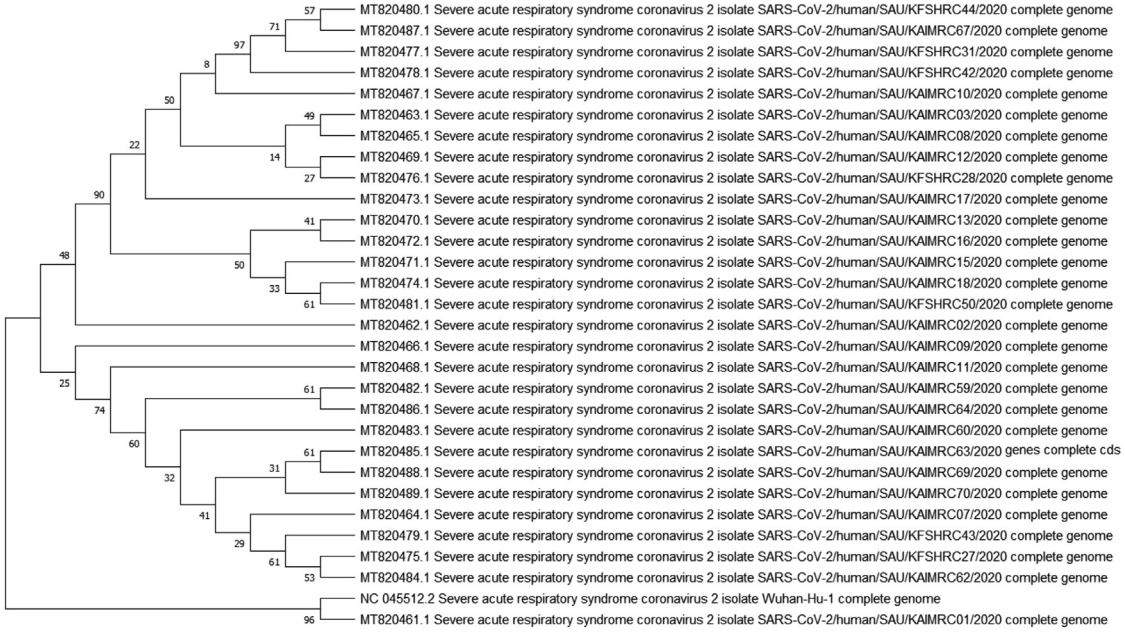

Fig. 3a. Phylogenetic affiliation of KSA variants compared to Wuhan strain.

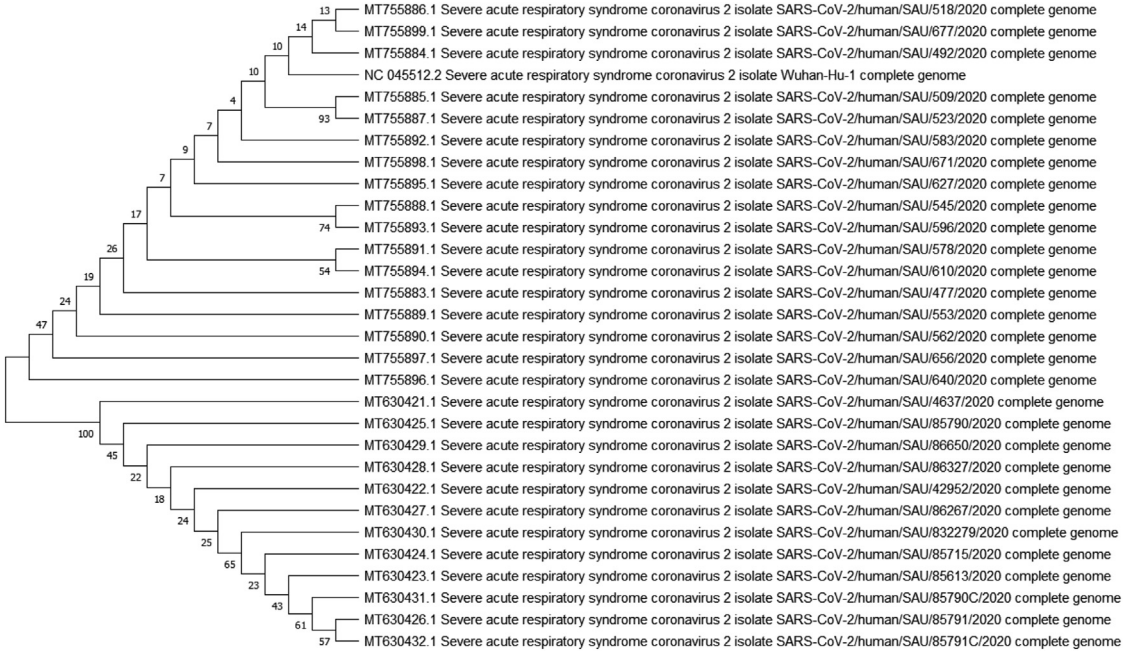

Fig. 3b. Phylogenetic affiliation of KSA variants compared to Wuhan strain.

determined the nucleotides changes in SARS-CoV-2 variants isolated in KSA.

Our study showed high homology of KSA variants compared to the reference strain and pointed 89 mutations, most of them 53.9 % were missense. In alignment to our findings several reports from different regions showed high homology at nucleotide level; Wang et al. (2020) reported similarity between 99.91 and 100%, Lu et al. (2020) displayed high level of conservation more than 99.98% homology, whereas Ceraolo and Giorgi (2020) showed more than 99% identity. At the level of mutations Shishir et al. 2020 detected 209 mutations in 64 variants of SARS-CoV-2 isolated in Bangladesh, Khailany et al. (2020) reported 116 mutations among 95 SARS-CoV-2 genome while among 66 SARS-CoV-2 variants isolated in China Ahmed-Abakur and Alnour (2020) reported 143 mutations. In spite of high homology, each study pointed few different mutations which were either specific or coexist with other mutations. The genetic variation in SARS-CoV-2 genomes could be due to geographical distributions and intra-host viral evolution after infection (Al-Qaaneh et al., 2021; Ceraolo and Giorgi, 2020). Abduljalil and Abduljalil (2020) stated that the genomic signature studies reveal a strong relation between the time of sampling, location of sample and genetic diversity. However the RNA viruses are associated with error-prone RNA dependent RNA polymerases that leads to occurrence of mutations and recombination events rather frequently (Uddin et al., 2020).

Our study showed that the most frequent base changes were C: T followed by G: T; as also reported in several other studies (Almubaid and Al-Mubaid 2021; Khailany et al., 2020; Koyama et al., 2020). The present study showed that the majority of mutations occurred in ORF1ab gene, followed by N gene, S gene, and ORF3a gene. These results are in congruence with Das et al. (2021), who studied 469 SARS-CoV-2 genome sequences, and specified that the susceptible proteins are ORF1ab, spike, nucleocapsid, ORF3a and ORF7a. The ORF1ab is the largest viral segment covering 2/3 of the viral genome (Koyama et al., 2020), contributing important non-structural proteins such as polyprotein, leader protein, 3C-like proteinase, RNA-dependent RNA polymerase, helicase, 3′-to-5′ exonuclease, and endoRNAse, as these proteins influence viral replication and host immune response (Al-Qaaneh et al., 2021). Consequently, many reports stated that the majority of mutations occurred in ORF1ab gene (Khailany et al., 2020; Shishir et al., 2021).

All structural genes in our study showed missense mutations except E gene which showed only one silent mutation; N gene showed highest mutations among the structural genes, the most frequent mutations at positions G28881A (S202N), and G28883C (G204R). The most frequent mutations in spike gene were A23403G (D614G) followed by C22444T (silent). The nucleocapsid protein form the helical structure of nucleocapsid during the assembly and can activate immune responses (Almubaid and Al-Mubaid 2021; Zeng et al., 2020) while the spike protein is a key protein to initiate the attack of the virus on the cellular level (Koyama et al., 2020). It has also been mentioned that persistent mutation in spike protein may indicate that the virus is becoming more effective in human to human transmission (Almubaid and Al-Mubaid 2021), for example, D614G mutation has been reported to enhance infectivity and transmission due to increased interaction with ACE2 receptor present on host cells (Shishir et al., 2021; Tiwari and Mishra, 2020). Several published papers have predicted possible amino acid changes in the viral proteins particularly in the S protein (Islam et al. 2020; Ahmed-Abakur and Alnour 2020; Tiwari and Mishra, 2020).

Our study indicated that the SARS-CoV-2 variants isolated in KSA were characterized with coexistence of C241T (UTR region), C3037T (NSP3 gene), C14408T (NSP12) and A23403G (S gene) nucleotide changes, they occur simultaneously in 55 variants out of 58. Similar pattern of mutations have also been reported in Europe and the United States by Koyama et al. (2020), who investigated 10022 SARS CoV-2 genomes from 68 countries. Our study showed that the mutations at G 25563T, G28881A and G28883C occurred in 33, 23, 22 variants out of 58; according to Shishir et al. (2021), these mutations were found highly prevalent among sequences from the Europe, USA and Bangladesh. This finding might figure out the source of infection in Saudi Arabia. In alignment with our findings several reports showed that C3037T and C241T changes were the two most frequent mutations (Shishir et al., 2021), followed by C14408T, A23403G, and T28144C (L84S,

ORF8) (Koyama et al, 2020; Raghav et al. 2020). Interestingly, C1059T is one of most frequent mutation (Koyama et al., 2020) but did not appear in KSA variants. These results revealed that SARS-CoV-2 variants isolated in KSA are free from E484K, K417N, N501Y mutations which characterized South African variant (Prathiviraj et al., 2020) and from N501Y which characterized UK variant (Al-Qaaneh et al., 2021; Prathiviraj et al., 2020; Al-Qaaneh et al., 2021).

5. Conclusion

The SARS CoV-2 variants isolated in Saudi Arabia showed high homology compared to original strain (Wuhan strain) and revealed characterized pattern of coexisting mutations (C241T, C3037T, C14408T and A23403G).

Funding

This research is NOT funded fully or partially from any Governmental or non-governmental agencies.

Authors contribution

The retrieved data from NCBI was divided equally between the authors, each one of the authors handled different accession numbers. The corresponding author and first author wrote the manuscript, the rest of authors revised it.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abduljalil, J.M., Abduljalil, B.M., 2020. Epidemiology, genome, and clinical features of the pandemic SARS-CoV-2: a recent view. New Microbes New Infect. 35, 100672. https://doi.org/10.1016/j.nmni.2020.100672.

Ahmed-Abakur, E.H., Alnour, T.M.S., 2020. Genetic variations among SARS-CoV-2 strains isolated in China. Gene Rep. 21, 100925. https://doi.org/10.1016/j.genrep.2020.100925.

Almubaid, Z., Al-Mubaid, H., 2021. Analysis and comparison of genetic variants and mutations of the novel coronavirus SARS-CoV-2. Gene Rep. 23, 101064. https://doi.org/10.1016/j.genrep.2021.101064.

Al-Qaaneh, A.M., Alshammari, T., Aldahhan, R., Aldossary, H., Alkhalifah, Z.A., Borgio, J.F., 2021. Genome composition and genetic characterization of SARS-CoV-2. Saudi J. Biol. Sci. 28 (3), 1978–1989. https://doi.org/10.1016/j.sjbs.2020.12.053.

Ashour, H.M., Elkhatib, W.F., Rahman, M.M., Elshabrawy, H.A., 2020. Insights into the Recent 2019 Novel Coronavirus (SARS-CoV-2) in light of past human coronavirus outbreaks. Pathogens. 9 (3), 186. https://doi.org/10.3390/pathogens9030186.

Bandoy, D.J.D.R., Weimer, B.C., 2021. Analysis of SARS-CoV-2 genomic epidemiology reveals disease transmission coupled to variant emergence and allelic variation. Sci. Rep. 11 (1), 7380. https://doi.org/10.1038/s41598-021-86265-4.

Barry, G.H., 2013. Building phylogenetic trees from molecular data with MEGA. Mol. Biol. Evol. 30 (5), 1229–1235. https://doi.org/10.1093/molbev/mst012.

Ceraolo, C., Giorgi, F.M., 2020. Genomic variance of the 2019-nCoV coronavirus. J. Med. Virol. 92 (5), 522–528. https://doi.org/10.1002/jmv.v92.510.1002/jmv.25700.

Das, J.K., Sengupta, A., Choudhury, P.P., Roy, S., 2021. Characterizing genomic variants and mutations in SARS-CoV-2 proteins from Indian isolates. Gene Rep. 19, 101044. doi: 19:101044.

Fehr, A.R., Perlman, S., 2015. Coronaviruses: an overview of their replication and pathogenesis. Methods Mol. Biol. 1282, 1–23. https://doi.org/10.1007/978-1-4939-2438-7_1.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39 (4), 783–791. https://doi.org/10.1111/evo.1985.39.issue-410.1111/j.1558-5646.1985.tb00420.x.

Islam, M.R., Hoque, M.N., Rahman, M.S., Alam, A.S.M.R.U., Akther, M., Puspo, J.A., Akter, S., Sultana, M., Crandall, K.A., Hossain, M.A., 2020. Genome-wide analysis of SARS-CoV-2 virus strains circulating worldwide implicates heterogeneity. Sci. Rep. 10 (1). https://doi.org/10.1038/s41598-020-70812-6.

Khailany, R.A., Safdar, M., Ozaslan, M., 2020. Genomic characterization of a novel SARS-CoV-2. Gene Rep. 19, 100682. https://doi.org/10.1016/j.genrep.2020.100682.

Koyama, T., Platt, D., Parida, L., 2020. Variant analysis of SARS-CoV-2 genomes. Bull. World Health Organ. 98 (7), 495–504. https://doi.org/10.2471/BLT.20.253591.

Lu, R., Zhao, X., Li, J., Niu, P., Yang, B.o., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N. a., Bi, Y., Ma, X., Zhan, F., Wang, L., Hu, T., Zhou, H., Hu, Z., Zhou, W., Zhao, L.i., Chen, J., Meng, Y., Wang, J.i., Lin, Y., Yuan, J., Xie, Z., Ma, J., Liu, W.J., Wang, D., Xu, W., Holmes, E.C., Gao, G.F., Wu, G., Chen, W., Shi, W., Tan, W., 2020. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet 395 (10224), 565–574. https://doi.org/10.1016/S0140-6736(20)30251-8.

Masters, P.S., 2006. Te molecular biology of coronaviruses. Adv. Virus Res. 66, 193–292. https://doi.org/10.1016/S0065-3527(06)66005-3.

Prathiviraj, R., Kiran, G.S., Selvin, J., 2020. Phylogenomic proximity and comparative proteomic analysis of SARS-CoV-2. Gene Rep. 20, 100777. https://doi.org/10.1016/j.genrep.2020.100777.

Raghav, S., Ghosh, A., Turuk, J., Kumar, S., Jha, A., Madhulika, S., Priyadarshini, M., Biswas, V.K., Shyamli, P.S., Singh, B., Singh, N., Singh, D., Datey, A., Avula, K., Smita, S., Sabat, J., Bhattacharya, D., Kshatri, J.S., Vasudevan, D., Suryawanshi, A., Dash, R., Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454.

Raghav, S., Ghosh, A., Turuk, J., Kumar, S., Jha, A., Madhulika, S., Priyadarshini, M., Biswas, V.K., Shyamli, P.S., Singh, B., Singh, N., Singh, D., Datey, A., Avula, K., Smita, S., Sabat, J., Bhattacharya, D., Kshatri, J.S., Vasudevan, D., Suryawanshi, A., Dash, R., Senapati, S., Beuria, T.K., Swain, R., Chattopadhyay, S., Syed, G.H., Dixit, A., Prasad, P., Pati, S., Parida, A., 2020. Analysis of Indian SARS-CoV-2 Genomes Reveals Prevalence of D614G Mutation in Spike Protein Predicting an Increase in Interaction With TMPRSS2 and Virus Infectivity. Front. Microbiol. 11. https://doi.org/10.3389/fmicb.2020.59492810.3389/fmicb.2020.594928.s00110.3389/fmicb.2020.594928.s00210.3389/fmicb.2020.594928.s00310.3389/fmicb.2020.594928.s00410.3389/fmicb.2020.594928.s005.

Seoane, B., 2021. A scaling approach to estimate the age-dependent COVID-19 infection fatality ratio from incomplete data. PLoS ONE 16 (2), e0246831. https://doi.org/10.1371/journal.pone.0246831.

Shishir, T.A., Naser, I.B., Faruque, S.M., 2021. In silico comparative genomics of SARS-CoV-2 to determine the source and diversity of the pathogen in Bangladesh. PLoS One. 16 (1), e0245584. https://doi.org/10.1371/journal.pone.0245584.

Tamura, K., Nei, M., Kumar, S., 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc. Natl. Acad. Sci. U S A. 101 (30), 11030–11035. https://doi.org/10.1073/pnas.0404206101.

Tiwari, M., Mishra, D., 2020. Investigating the genomic landscape of novel coronavirus (2019-nCoV) to identify non-synonymous mutations for use in diagnosis and drug design. J. Clin. Virol. 128, 104441. https://doi.org/10.1016/j.jcv.2020.104441.

Wang, C., Liu, Z., Chen, Z., Huang, X., Xu, M., He, T., Zhang, Z., 2020. The establishment of reference sequence for SARS-CoV-2 and variation analysis. J. Med. Virol. 92 (6), 667–674. https://doi.org/10.1002/jmv.v92.610.1002/jmv.25762.

WHO. Middle East respiratory syndrome coronavirus (MERS-CoV). https://www.who.int/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-(mers-cov) [accessed 5 May 2021].

Uddin, M., Mustafa, F., Rizvi, T.A., Loney, T., Al Suwaidi, H., Al-Marzouqi, A.H.H., Eldin, A.K., Alsabeeha, N., Adrian, T.E., Stefanini, C., Nowotny, N., Alsheikh-Ali, A., Senok, A.C., 2020. SARS-CoV-2/COVID-19: viral genomics, epidemiology, vaccines, and therapeutic interventions. Viruses. 12 (526). https://doi.org/10.3390/v12050526.

Zeng, W., Liu, G., Ma, H., Zhao, D., Yang, Y., Liu, M., Mohammed, A., Zhao, C., Yang, Y., Xie, J., Ding, C., Ma, X., Weng, J., Gao, Y., He, H., Jin, T., 2020. Biochemical characterization of SARS-CoV-2 nucleocapsid protein. Biochem. Biophys. Res. Commun. 527 (3), 618–623. https://doi.org/10.1016/j.bbrc.2020.04.136.