Kurdistan Journal of Applied Research (KJAR)
Print-ISSN: 2411-7684 | Electronic-ISSN: 2411-7706

Website: Kjar.spu.edu.iq | Email: kjar@spu.edu.iq

Is The SARS-CoV2 Evolved in Human Being: A prospective Genetic Analysis

Salvatore Dimonte
BioMolecular Lab
Barletta (BT) 76121
Italy
salvatore.dimonte@gmail.com

Paywast Jamal Jalal
Biology Department
College of Science
University of Sulaimani
Sulaimani, Iraq
paywast.jalal@univsul.edu.iq

Hiwa Abdulrahman Ahmad
Health Research Center
Faculty of Sciences
University of Koya
Koya, Iraq
hiwa.abdulrahman@koyauniversity.org

Safa Bakr Karim
Anesthesia Department
Technical College of Health
Sulaimani Polytechnic University
Sulaimani, Iraq
kamal.rashid@spu.edu.iq

Taib Ahmed Hama Soor
Medical Laboratory Department
Technical College of Health
Sulaimani Polytechnic University
Sulaimani, Iraq
taib.hama-soor@spu.edu.iq

Kamal Jalal Rashid
Anesthesia Department
Technical College of Health
Sulaimani Polytechnic University
Sulaimani, Iraq
kamal.rashid@spu.edu.iq

Salar Ibrahim Ali
Nursing Department
Technical College of Health
Sulaimani Polytechnic University
Sulaimani, Iraq
salar.ali@spu.edu.iq

Muhammad Babakir-Mina
Medical Laboratory Department
Technical College of Health
Sulaimani Polytechnic University
Sulaimani, Iraq
muhammed.babakir-mina@spu.edu.iq

Francesco Greco
Department of Intensive Care Medicine and Neonatology
University Children's Hospital
Zurich, Switzerland
grecofnc@hotmail.it

Article Info

Volume 5 – Special Issue: 4th International Conference on the Health and Medical Science : Medical Researches Improve Life Quality (ICHMS 2020)

DOI: 10.24017/science.2020.ICHMS2020.19

Article history:

ABSTRACT

COVID-19 is the deadly respiratory disease of the century caused by new type unknown origin Coronavirus. The recent effort of the word researchers is toward finding the origin of the virus. The current study investigated the extent of molecular similarity and divergence between SARS-CoV2 and other related Coronavirus. An attempt has been made to investigate the epidemiological study of this new contagious virus using molecular

Received: 21 September 2020
Accepted: 26 September 2020

Keywords:

SARS-CoV2, SARS-CoV, MERS-CoV, COVID-19 and 2019-nCoV

biology techniques. The phylogenetic trees for all human coronaviruses with the novel Coronavirus have been built using a several complete amino acid sequences of the four known structural proteins, S (spike), E (envelope), M (membrane), and N (nucleocapsid). The result of the study revealed that the SARS-CoV2 is related to human SARS-CoV isolated from different countries very cloely, especially those strains recovered from China in recent times, 2020. The evolutionary changes observed in the inserted 23 amino acids in the RNA binding domain (RBD) of the coronvirus spike glycoprotein which cannot be detected in any other human coronavirus. Moreover, the 2019-nCoV is not closely related to other alpha, beta and gamma human Coronavirus, including MERS-CoV. The current study concluded that 2019-nCoV is more likely believed to originated from SARS-CoV. The probability is more vital to be originated from the strain isolated in China in 2020, which is coincident with the spraed of COVID-19 in the same country. The phyloepidemiologic analyses suggested that the coronaviruses are circulating in human hosts evolving gradually by times in response to the different environment stimuli facing the virus inside the host in different geographical areas. Furthermore, the analysis showed the flow of transmission, and evolutionary changes of SARS-CoV2 which may be directed from the transmission of SARS-CoV from human to Bat and Pangolin then jumped to human again in the crowded market Wuhan city in China.

1. INTRODUCTION

The enveloped Betacoronavirus possesses a single-stranded RNA with a positive polarity, and they are hosted in the mammals [1]. Before the COVID-19 outbreaks, only four strains of Betacoronavirus, including SARS-CoV, MERS-CoV, OC43 and HKU1 had been reported to infect human and cause severe pneumonia [1]. Since December 2019, another strain of Betacoronavirus named SARS-CoV2 detected to cause severe human pneumonia named as Coronavirus Infectious disease-2019 (COVID-19) [2, 3]. The new Coronavirus has been reported in a wet market in Wuhan city which is located in Hubei province, Central China [4]. Since its appearance, it causes a lockdown with the infection of over 4800000 people in all over the World and recording mortality of more than 300000 people till april 2020 and unfortunately this record increase on the daily bases.

According to the genomic analysis of some published paper, it is suggested that Bat is the origin of SARS-CoV2, and Pangolin is an intermediate mammal's host [5-7]. According to the medical information, the initially infected cases were related to the seafood market in one of China city,Wuhan c [8, 9] transmitted from animals to human [10]. Later, the human to human transmission was reported even in peoples whom no previous contact with animals or even the market was documented [11-13]. However, other researcher argued that the wet market in China was not the only source of SARS-CoV2 transmission to humans [14, 15]. Because the market closed in the very beginning of raising the COVID-19, so this makes the identification

of the intermediate host difficult. Due to this reason, SARS-CoV2 origin and transmission are still debatable.

Four structural proteins encoded by the genome of SARS-CoV2 which participating in different viral processes, for instance, formation of the particle. The immunogenicity rate to T-Lymphocyte cell is higher in the structural proteins than non-structural proteins [16]. All the coronaviruses have these structural proteins, S (spike), E (envelope), M (membrane), and N (nucleocapsid). [17, 18]. The spike glycoprotein composed of two domains S1 and S2, which both enhance the efficient binding to the ACE2 (angiotensin-converting enzyme 2 receptor) in the host cell membrane [19-21]. After the internalization of SARS-CoV2 in the endosome, a sequence of conformational change in the spike domains lead to activation of the membrane fusion within the endosome [21-24]. The small and integral envelop (E) protein believed to have an essential role in the viral life cycle of the virus from assembly to envelope generation, and pathogenesis. The E protein act as an ion-channel viroporin and can interact with the host cell proteins and other coronavirus proteins [25]. The Membrane (M) protein composed of three transmembrane domain that shapes of the virion, enhances membrane curvature, and reachjes to the nucleocapsid for binding [26]. During virion assembly, the positive-strand RNA of the virus packaged by the Nucleoprotein (ORF9a) into a helical ribonucleocapsid (RNP) by binding of the viral genome to the membrane protein [27]. Moreover, the nucleoprotein enhance viral replication and transcription.

In the current study, complete structural proteins of the SARS-CoV2 were retrieved from NCBI Data Base for a reason for the analysis of the evolution and transmission of this novel Coronavirus since it is appearance. We aimed to explore the genetic variation in the structural proteins, infer the evolutionary flow in comparison to the previous and different isolates of the SARS-CoV2 in different countries, and lastly, indicate the history of transmission to estimate to find the hypothesis or real origin of the virus.

2. METHODS AND MATERIALS

The multiple sequence alignment was performed based on the amino acid using MultAlin software [28]to show the variation in the structural proteins of SARS-CoV2 in comparison to SARS-CoV. The start codon of each sequence of the structural protein was gained from NCBI graphics of the following selected sequence, SARS-CoV2 [NC045512.2 (Wuhan-Hu-1CHN/2019/First Isolate), MT152824.1 (USA/2020), LC528233.1 (Japan/2020), MT240479.1 (Pakistan/2020) and MN996531.1 (China/2020)]; SARS-CoV [KF514397.1 (USA/2009), AB257344.1 (Germany/2006), MK062184.1 (USA/2018), MK062180.1 (USA/2018), AY463060.1 (China/2004), NC004718.3 (Canada/2018) and KF514422.1 (USA/2014).

Different strains of human Coronavirus were downloaded from the NCBI database (http://www.ncbi.nlm.nih.gov) to characterize the evolution of each structural protein. The accession numbers of the sequences were shown in figure 1, 2, 3 and 4 respectively for the spike protein (68 sequences), Nucleocapsid protein (61 sequences), Membrane protein (61 sequences) and Envelop protein (62 protein). A typical local alignment search engine was used to check the percentage of similarity (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The duplication sequences were excluded to avoid the mistake of calculating the rate of similar amino acid sequences variations. Maximum likelihood (ML) methods were used to create a phylogenetic tree. Generalized time-reversible plus gamma distribution and invariant sites (+G+I) has been utilized for ML trees rebuilding using MEGAX for all protein sequences that their accession numbers available in the figures. The MUSCLE software was used to further confirm and building the last trees of the protein sequences were aligned with, the phylogenetic tree was designed using TreeDyn and PhyML (www.phylogeny.fr).

3. RESULTS

1) Most SARS-CoV2 variation located in the Spike region

The multiple sequence analysis [supplementary data] of the selected sequences showed that

most variations in the SARS-CoV2 located in the receptor-binding site in the spike region, followed by the nucleocapsid proteins when compared to the SARS-CoV strains [Table-1]. Out of 1277 amino acids in the spike region, there were 225 amino acids substituted and interestingly 23 amino acid which is approximately 1.8% of the total amino acids in the spike inserted in the receptor-binding site. The SARS-CoV2 envelop consist of 76 amino acid which only three amino acids substituted and one deleted in comparison to SARS-CoV. The Membrane (M) protein is followed by the spike protein in the high percentage of substituted proteins with 7.2% substitution and insertion of one amino acid. The nucleocapsid of SARS-CoV2 composed of 422 amino acid which 25 amino acid substituted and three amino acid is deleted. The lowest number of amino acid deletion and substitution observed in envelop of the SARS-CoV2.

Table 1: The percentage of the substituted, inserted and deleted amino acids in the structural proteins of SARS-CoV2 in comparison to SARS-CoV.

Structural protein	No. of Amino acid	Deletion (%)	Substitution (%)	Insertion (%)
Spike	1277	4 (0.3%)	225 (17.6%)	23 (1.8%)
ENVELOP	76	1 (1.3%)	3 (3.9%)	None
MEMBRANE	222	None	16 (7.2%)	1 (0.45%)
NUCLEOCAPSID	422	3 (0.7%)	25 (5.9%)	None

2) SARS-CoV2 is in close relation to SARS-CoV based on their structural proteins

The phyloepidemiologic analysis showed that most strains found in the NCBI database when compared to different isolates of human Coronavirus are in a high percentage of similarity to human SARS-CoV. The four structural proteins of human Coronavirus were compared separately. In each tree, all the isolates divided into three cluster A, B and C.

The phylogenetic tree of the spike protein was analyzed 1277 amino acids of 36 different isolate of SARS-CoV2 with 8 isolates of SARS-CoV, 9 isolates of MERS-CoV, 9 isolates of OC43, 4 isolates of NL63, 1 isolate of each HKU1 and 229E. All the strains are isolated in different countries and years. As it is apparent from the tree [Figure 1], the strains divided into three clusters A, B and C. The alpha coronaviruses (NL63 and 229E) are grouped in cluster A. Group B is a relatively more extensive group of different beta-human Coronaviruses and MERS-CoV. The group B is divided into two branches; one is including the MERS-CoV and the other one gathering OC43 and HKU1. Group C consists the largest group of the tree, including the novel SARS-CoV2 and previously isolated SARS-CoV. The human Coronavirus in group C has a high similarity in their sequence of the spike region.

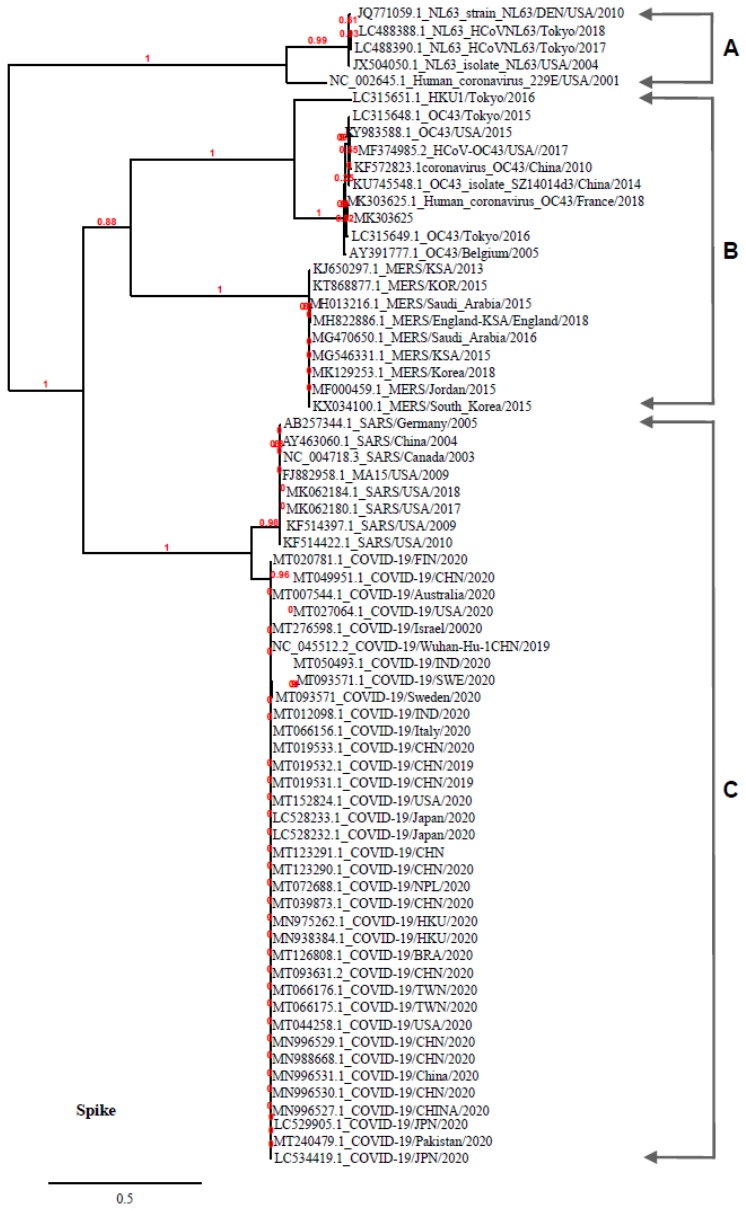

Figure 1: Phylogenetic tree of the SARS-CoV2 (Spike protein) with other human coronaviruses. The phylogenetic tree was buit using the maximum likelihood method . The sequences of the spike were aligned with MUSCLE, and PhyML and TreeDyn were used to design The phylogenetic tree (www.phylogeny.fr). The values of different support Branches are indicated in red color.

The nucleocapsid tree analyses 422 amino acids in different selected human Coronavirus. As showed in figure 2, the tree divided into two main clades. One including group A with both HKU1 and OCA43 human coronaviruses with the similarity of 66%. The other clade also divided into two branches, including both group B and C. Group B gathering some strains of OCA43 with NL63 and 229E human Coronavirus. Interestingly, the close similarity between the SARS-CoV2, SARS-CoV and MERS-CoV could be observed in group C with the similarity of 96%. The novel coronaviruses are less closly related to MERS-CoV than SARS-CoV based on the analysis of the nucleocapsid amino acids.

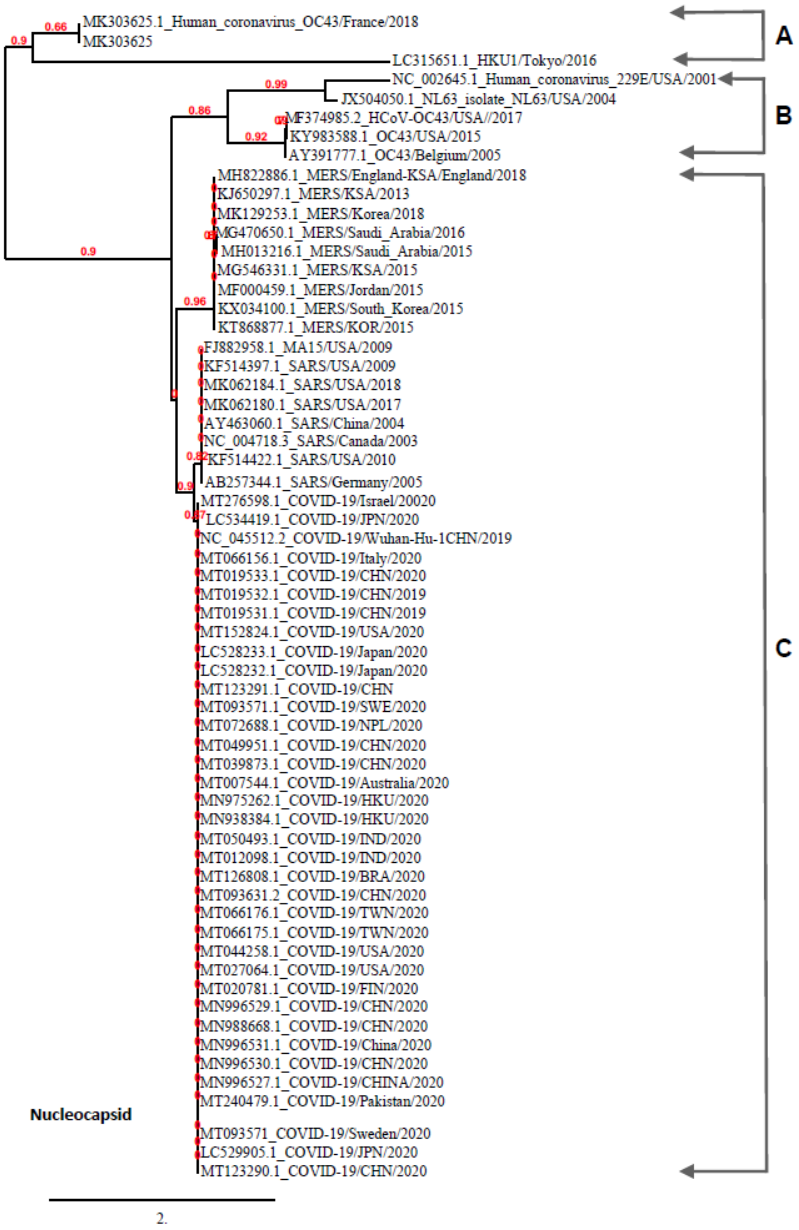

Figure 2: Phylogenetic tree analysis of the SARS-CoV2 (Nucleocapsid proteins) with other human coronaviruses. The phylogenetic tree was buit using the maximum likelihood method . The sequences of the spike were aligned with MUSCLE, and PhyML and TreeDyn were used to design The phylogenetic tree (www.phylogeny.fr). The values of different support Branches are indicated in red color.

The comparison of the 222 amino acid of the Membrane (M) showed the two major clades among different strains of the human coronaviruses. One of the clades including group A of both 229E and NL63 human coronaviruses. The other clade, including both group B and group C. Group B including the OCA43 human coronaviruses strains. All the human beta coronaviruses are gathered in group C including the novel coronavirus SARS-CoV2, SARS-CoV and MERS-CoV. As the comparison of spike and nucleocapsid, the amino acid of the membrane of the SARS-CoV2 strains is with a close similarity to SARS-CoV.

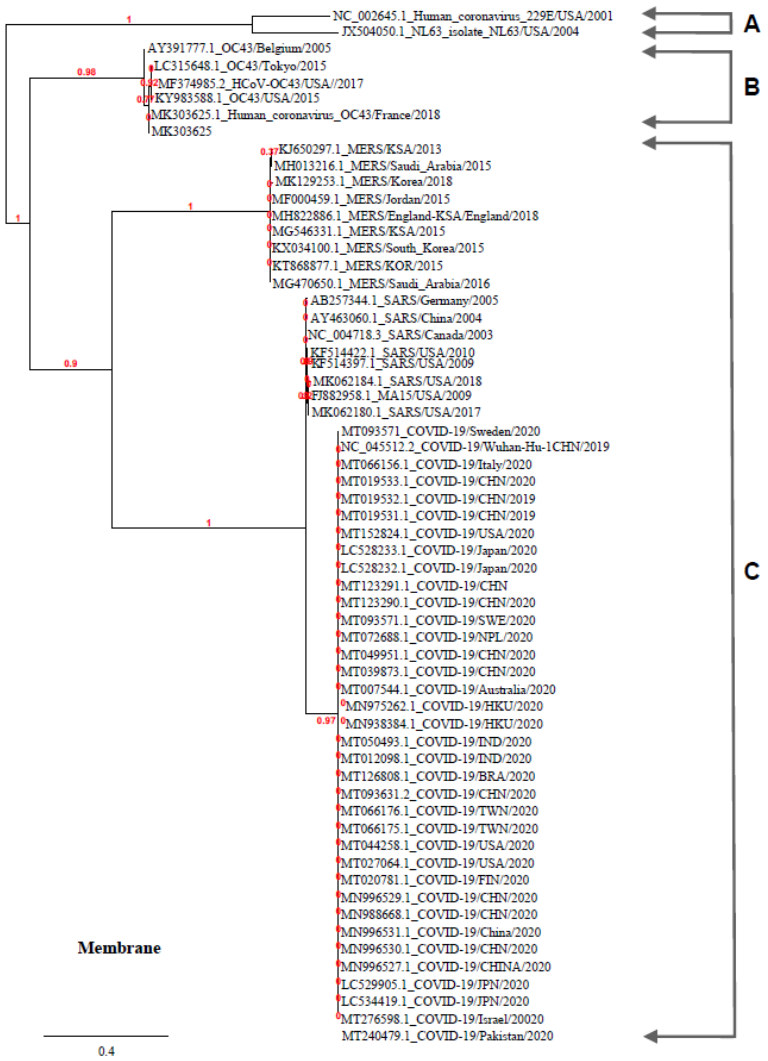

Figure 3: Phylogenetic tree of the SARS-CoV2 (Membrane protein) with other human coronaviruses. The phylogenetic tree was buit using the maximum likelihood method . The sequences of the spike were aligned with MUSCLE, and PhyML and TreeDyn were used to design The phylogenetic tree (www.phylogeny.fr). The values of different support Branches are indicated in red color.

The envelop consists of 76 amino acids and the comparison between SARS-CoV2 with the other SARS-CoV showed the lowest variation among the novel coronaviruses. However, the phylogenetic analysis revealed that SARS-CoV2 strains are located in one of the major clades and belonging to group A. This group consist of both SARS-CoV2 and SARS-CoV with a high degree of similarity (93%) in the envelope region. Moreover, this clade includes group B too, which gathering MERS-CoV, HKU1 and OCA43 with a similarity of 77%. The other clade includes group C, which both NL63 and 229E human Coronavirus.

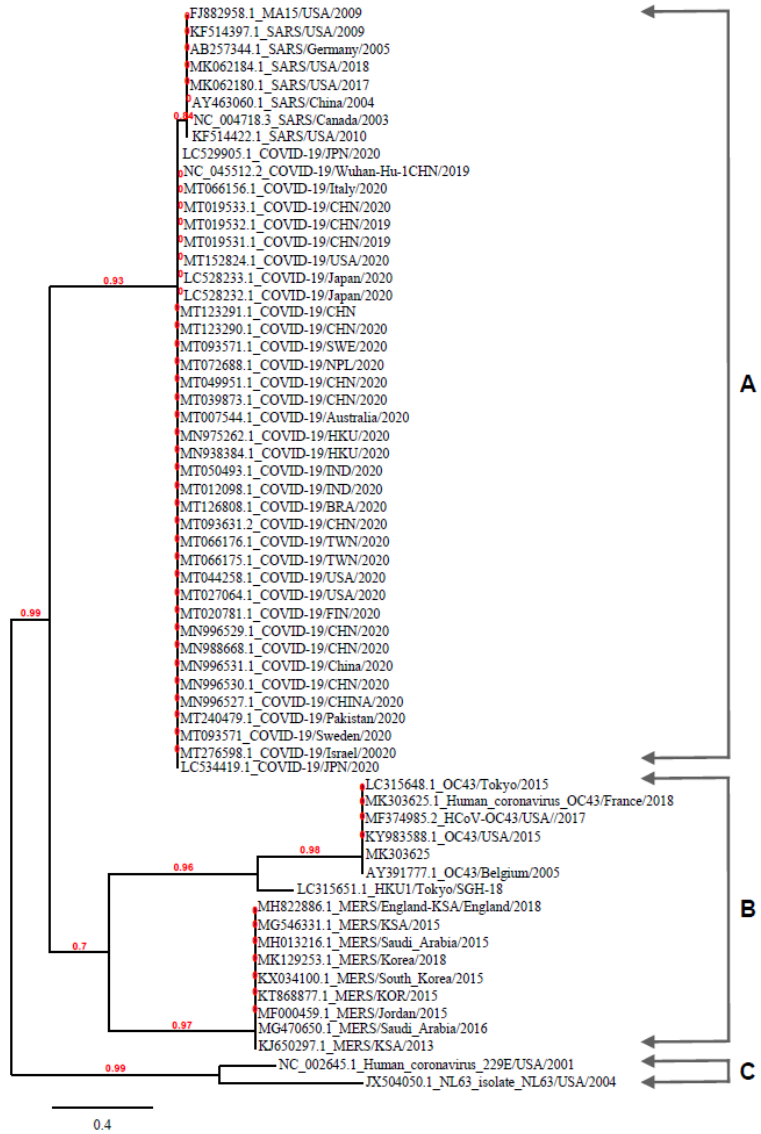

Figure 4: Phylogenetic tree of the SARS-CoV2 (Envelop protein) with other human coronaviruses. The phylogenetic tree was buit using the maximum likelihood method . The sequences of the spike were aligned with MUSCLE, and PhyML and TreeDyn were used to design The phylogenetic tree (www.phylogeny.fr). The values of different support Branches are indicated in red color.

4. DISCUSSION

COVID-19 is the pandemic severe acute respiratory disease, and it causes high morbidity and mortality rates. The origin of this novel Coronavirus is still unrecognized, so it is vital to understand the epidemiology of the 2019-nCoV. The bioinformatics tools have been used successfully as a powerful tool to analyze the infectious disease outbreak, transmission and understanding the pathogen [15]. In this report, we aimed to use this tool to understand the novel coronavirus real origin and to show the most frequent variation in the structural proteins. Many strains of human Coronavirus used for the reason of the phyloepidemiological analysis based on the full amino acids of the structural proteins; S (spike), E (envelope), M (membrane), and N (nucleocapsid). of human coronaviruses in various countries and times. Our analysis showed that most variations located in the spike region of the SARS-CoV2 when

compared to SARS-CoV. Moreover, the phylogenetic tree revealed that all the selected strains of the novel coronavirus are gathered together in one clade, and they are in close relation to the SARS-CoV. According to the bioinformatics analysis, the origin of the SARS-CoV2 is human SARS-CoV.

As it is evident that SARS-CoV2 is an RNA virus which makes its replication machinery inside the host cell [26] and it originates from Bat and Pangolin [7, 29]. The previous human SARS-CoV (2003) and MERS-CoV (2012) were also originated from the Bat [30-32]. The spillover of these viruses from Bat to human is believed to be related to the high mutation rate in the RNA viruses which make an error in the new progeny of the viruses due to the RNA dependent RNA polymerase that does not have the proofreading activity [26, 33].

The high similarity between SARS-CoV2 and SRAS-CoV in their structural regions give support to our hypothesis that this novel Coronavirus is originated from the SARS-CoV. We believed that the scenario of the transmission of the SARS-CoV2 is that a sequential of antigenic shift and drift occurred in the SARS-CoV RNA in the Bat then transmitted to Pangolin to give some other antigenic capability for binding to the host receptors. Then indirect transmission occurred to human due to viral contamination of subject or asymptomatic infected person [12].

Our analysis sequence data revealed that the high mutation rate occurred in the spike region of SARS-CoV2 in comparison to SARS-CoV. Any evolution and mutation happen in this protein may change the ability of the virus to attack different hosts and even with different strengths. The virus uses the spike region for binding to the specific receptor in the host cell [34]. Therefore, the high mutation rate in this region is to enhance the entry and escaping from the immunity of the host [35].

5. CONCLUSION

The genetic evolution of the collected data demonstrated that SARS-CoV2 most likely originated from SARS-CoV infected Bat through sequential antigenic shift and drift in the RNA genome. This may give a clue that human coronaviruses can evolve to new strains of the virus, such as the 2019-nCoV by passing through different hosts and bypassing times. However, the postulate of originating 2019-nCoV from other human coronaviruses (alpha, beta and gamma Coronavirus) is weaker because they are less closely related to the 2019-nCoV. This fining will help to find the possible origin of the virus, which is essential for infection control policy.

REFERENCE

[1] J. Cui, F. Li, and Z. L. Shi, "Origin and evolution of pathogenic coronaviruses," *Nat Rev Microbiol,* vol. 17, pp. 181-192, Mar 2019.

[2] WHO, "Coronavirus disease (COVID-2019) situation reports. https://www.who.int/emergencies/diseases/novel-coronavirus," *2019/situation-reports/ [accessed 12 February 2020]*. 2020.

[3] D. S. Hui, I. A. E, T. A. Madani, F. Ntoumi, R. Kock, O. Dar, *et al.*, "The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China," *Int J Infect Dis,* vol. 91, pp. 264-266, Feb 2020.

[4] S. Zhao, S. S. Musa, Q. Lin, J. Ran, G. Yang, W. Wang, *et al.*, "Estimating the Unreported Number of Novel Coronavirus (2019-nCoV) Cases in China in the First Half of January 2020: A Data-Driven Modelling Analysis of the Early Outbreak," *J Clin Med,* vol. 9, Feb 1 2020.

[5] F. Greco, S. Dimonte, M. Babakir-Mina, T. Hama-Soor, and S. Ali, "Is It Scaly Anteater or Bat A Real Origin of The 2019-Novel CoV: A Probable Hypothesis?," *Kurdistan Journal of Applied Research,* vol. 5, pp. 1-12. DOI https://doi.org/10.24017/covid.1, 2020.

[6] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, *et al.*, "A pneumonia outbreak associated with a new coronavirus of probable bat origin," *Nature,* vol. 579, pp. 270-273. https://doi.org/10.1038/s41586-020-2012-7, 2020.

[7] T. Zhang, Q. Wu, and Z. Zhang, "Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak," *Current Biology,* vol. 30 pp. 1346-1351. DOI: 10.1016/j.cub.2020.03.022, Mar 19 2020.

[8] W. Ji, W. Wang, X. Zhao, J. Zai, and X. Li, "Cross-species transmission of the newly identified coronavirus 2019-nCoV," *J Med Virol,* vol. 92, pp. 433-440. https://doi.org/10.1002/jmv.25682, Apr 2020.

[9] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, *et al.*, "Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia," *N Engl J Med,* vol. 382, pp. 1199-1207. DOI: 10.1056/NEJMoa2001316, Mar 26 2020.

[10] M. A. Shereen, S. Khan, A. Kazmi, N. Bashir, and R. Siddique, "COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses," *J Adv Res,* vol. 24, pp. 91-98, Jul 2020.

[11] J. F.-W. Chan, S. Yuan, K.-H. Kok, K. K.-W. To, H. Chu, J. Yang, *et al.*, "A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster," *The Lancet,* vol. 395, pp. 514-523. DOI:https://doi.org/10.1016/S0140-6736(20)30154-9, 2020.

[12] J. Cai, W. Sun, J. Huang, M. Gamber, J. Wu, and G. He, "Indirect Virus Transmission in Cluster of COVID-19 Cases, Wenzhou, China, 2020," *Emerging Infectious Diseases journal,* vol. 26, p. DOI: 10.3201/eid2606.200412, 2020.

[13] P. Li, J. B. Fu, K. F. Li, Y. Chen, H. L. Wang, L. J. Liu, *et al.*, "Transmission of COVID-19 in the terminal stage of incubation period: a familial cluster," *Int J Infect Dis,* Mar 16 2020.

[14] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, *et al.*, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *The Lancet,* vol. 395, pp. 497-506. DOI:https://doi.org/10.1016/S0140-6736(20)30183-5, 2020.

[15] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, *et al.*, "A Novel Coronavirus from Patients with Pneumonia in China, 2019," *N Engl J Med,* vol. 382, pp. 727-733, Feb 20 2020.

[16] C. K.-f. Li, H. Wu, H. Yan, S. Ma, L. Wang, M. Zhang, *et al.*, "T cell responses to whole SARS coronavirus in humans," *The Journal of Immunology,* vol. 181, pp. 5490-5500, 2008.

[17] E. G. Brown and J. A. Tetro, "Comparative analysis of the SARS coronavirus genome: a good start to a long journey," *The Lancet,* vol. 361, pp. 1756-1757, 2003.

[18] J. S. Kahn and K. McIntosh, "History and recent advances in coronavirus discovery," *Pediatr Infect Dis J,* vol. 24, pp. S223-7, discussion S226, Nov 2005.

[19] K. Kuba, Y. Imai, S. Rao, H. Gao, F. Guo, B. Guan, *et al.*, "A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury," *Nat Med,* vol. 11, pp. 875-9, Aug 2005.

[20] B. J. Bosch, R. van der Zee, C. A. de Haan, and P. J. Rottier, "The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex," *J Virol,* vol. 77, pp. 8801-8811. DOI: 10.1128/JVI.77.16., Aug 2003.

[21] A. C. Walls, Y. J. Park, M. A. Tortorici, A. Wall, A. T. McGuire, and D. Veesler, "Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein," *Cell,* vol. https://doi.org/10.1016/j.cell.2020.02.058 Mar 6 2020.

[22] J. K. Millet and G. R. Whittaker, "Physiological and molecular triggers for SARS-CoV membrane fusion and entry into host cells," *Virology,* vol. 517, pp. 3-8, Apr 2018.

[23] A. Hasan, B. A. Paray, A. Hussain, F. A. Qadir, F. Attar, F. M. Aziz, *et al.*, "A review on the cleavage priming of the spike protein on coronavirus by angiotensin-converting enzyme-2 and furin," *J Biomol Struct Dyn,* pp. 1-13. DOI: 10.1080/07391102.2020.1754293, Apr 10 2020.

[24] B. Coutard, C. Valle, X. de Lamballerie, B. Canard, N. G. Seidah, and E. Decroly, "The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade," *Antiviral Res,* vol. 176, p. 104742. DOI: 10.1016/j.antiviral.2020.104742, Apr 2020.

[25] D. Schoeman and B. C. Fielding, "Coronavirus envelope protein: current knowledge," *Virol J,* vol. 16, p. 69, May 27 2019.

[26] Y. Chen, Q. Liu, and D. Guo, "Emerging coronaviruses: Genome structure, replication, and pathogenesis," *J Med Virol,* vol. 92, pp. 418-423, Apr 2020.

[27] L. Mousavizadeh and S. Ghasemi, "Genotype and phenotype of COVID-19: Their roles in pathogenesis," *J Microbiol Immunol Infect,* vol. S1684-1182:30082- 30087, p. http://doi.org/10.1016/j.jmii.2020.3.022, Mar 31 2020.

[28] F. Corpet, "Multiple sequence alignment with hierarchical clustering," *Nucleic Acids Research,* vol. 16, p. DOI: 10.1093/nar/16.22.10881, 1988.

[29] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, *et al.*, "Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin," *BioRxiv. https://doi.org/10.1101/2020.01.22.914952,* 2020.

[30] S. R. Weiss and S. Navas-Martin, "Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus," *Microbiol Mol Biol Rev,* vol. 69, pp. 635-664. doi:10.1128/MMBR.69.4, Dec 2005.

[31] Y. Fan, K. Zhao, Z. L. Shi, and P. Zhou, "Bat Coronaviruses in China," *Viruses,* vol. 11, Mar 2 2019.

[32] Z. Song, Y. Xu, L. Bao, L. Zhang, P. Yu, Y. Qu, *et al.*, "From SARS to MERS, Thrusting Coronaviruses into the Spotlight," *Viruses,* vol. 11, Jan 14 2019.

[33] N. S. Ogando, F. Ferron, E. Decroly, B. Canard, C. C. Posthuma, and E. J. Snijder, "The Curious Case of the Nidovirus Exoribonuclease: Its Role in RNA Synthesis and Replication Fidelity," *Front Microbiol,* vol. 10, p. 1813, 2019.

[34] W. Tai, L. He, X. Zhang, J. Pu, D. Voronin, S. Jiang, *et al.*, "Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine," *Cell Mol Immunol,* Mar 19 2020.

[35] X. Ou, Y. Liu, X. Lei, P. Li, D. Mi, L. Ren, *et al.*, "Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV," *Nat Commun,* vol. 11, pp. 1620. https://doi.org/10.1038/s41467-020-15562-9, Mar 27 2020.