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 It Scaly Anteater or Bat A Real Origin of The 2019-Novel CoV: A Probable Hypothesis?

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

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

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

Salvatore Dimonte
Bio Molecular Laboratory
Barletta (BT) 76121,
Italy
salvatore.dimonte@gmail.com

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

Article Info

Special Issue on Coronavirus (COVID-19)

DOI:
10.24017/covid.1

Article history:

Received: 16 April 2020
Accepted: 18 April 2020

Keywords:

Sulaimani city,
COVID-19,
SARS-CoV-2,
Bat-like- coronavirus,
Pangolin.

ABSTRACT

Today, the emerging of the new coronavirus 2019nCoV possesses a global health problem and little is known about its origin. In the current investigation, an evolutionary and molecular epidemiological analysis have provided of this new emerged virus. The phylogenetic trees for animal coronaviruses with the novel coronavirus-2019 have been created using a number of available complete protein sequences of envelope (E), membrane (M), nucleocapsid (N) and spike (S) proteins. The phylogenetic trees analysis illustrated that 2019nCoV in all four proteins are very closely related with coronaviruses isolated from Pangolin (scaly anteater) and Bat-SARS-like-coronavirses because all of them are clustered in the same clade. Whereas, the 2019nCoV is less closely related to coronavirses isolated from Rousettus bat (fruit bat) and MERS coronaviruses isolated from camel because they are gathered in the same clade only in two of the four studied proteins, nucleocapsid (N) and spike (S). In the conclusion, the new 2019nCoV is more likely to be originated from Bat-SARS-like-coronaviruses or/and coronavirus isolated from Pangolin after adaptation and evolution in the human hosts. Because of the

number of infected cases to date indicates a very quick human-to-human transmission. Thus, necessitates a very rapid active surveillance using accurate method to find the original host where the 2019nCoV emerged. This will help in further understanding and creating a better approach to control the spread of SARS-CoV-2 outbreak.

1. INTRODUCTION

In the end of December 2019, a respiratory disease (COVID-19) cases caused by a newly emerging coronavirus occurred in Wuhan, China. At the beginning, this coronavirus was named by World Health Organization (WHO) as the 2019- novel coronavirus (2019-nCoV) on 12 January 2020. WHO officially identified the disease as coronavirus disease 2019 (COVID-19) and the International Committee proposed the name to the virus as SARS-CoV-2, both issued on 11 February 2020. Coronaviridae family is a large group of viruses containing single, positive stranded RNA which is present in a number of host species. Previously, it was known as a causative agent for common cold and sometimes diarrhoea in humans [1,2]. A new severe acute respiratory syndrome coronavirus [SARS CoV] has been discovered and it was known as a cause of the SARS outbreak [1,2]. Recently, a number of severe respiratory infection in humans has been identified in Wuhan city of China, this was due to the emerge a novel coronavirus (2019-nCoV) as a causative agent for severe pneumonia in humans. From December 2019, many infected cases of pneumonia of unknown origin associated with permanence at the Wuhan market in China have been revealed [3,4]. 2019-nCoV belongs to the Orthocoronavirinae subfamily and it is different from MERS-CoV and SARS-CoV. [5]. Till now, a total of 1773,084 people is recorded to be infected with pneumonia globally with 111,652 mortality [6,7]. Animal to human transmission of the 2019-nCoV is contributed the epidemic outbreak, In China. A large number of patients found that they have had visited local fresh fish and wild animal markets at Wuhan in November. Recently after that, evident information has been collected for the zoonotic and human to human transmission of the virus [7,8]. It is shown previously that SARS virus transmission was really related with air droplet of infected patients among intensive care staff [9]. In case of COVID-19 in China, the increased mortality risk is probably due to the failure in the health care system, which indicates the public health interventions enhancement, involving social distance and restrictions of movement [10].

It has been shown that COVID-19 infects all age groups in various geographical areas [11] and the transmission can be occurred through direct and indirect way [12]. The patients have a different response to the infection, which starts from asymptomatic to severe pneumonia and similar condition was seen in MERS outbreaks in the Middle East [13].

According to the recent study through a phylogenetic network analysis of 160 complete genome sequence of the 2019-nCoV which is done by England and German researchers, there are three central variants of the 2019-nCoV are identified until today. The phylogenetic tree was relied on amino acid sequences of the virus and different mutations and changes of amino acids have been discovered. According to these mutations, the variants are named as A, B, and C. In which, the researchers believed that variants B and C are originated from variant A. These variants are noticed to be found in different countries. In Europe, East Asia and America, variants A and C are significantly spread. While, type B causes more outbreaks in East Asia [14]. Therefore, it is important to build a phylogenetic tree of the 2019nCoV in all regions to indicate the types of the variants.

As the early stage of infection control, the proper diagnosis technique and isolation of the infected people are considered as the hallmarks for the control and management of the disease. The current research aims to identify the phylogenetic estimation of the 2019nCoV and its evolutionary rate to more understand the real origin of the virus. Also, phylogenetic analysis is an important outcome to understand the progression of infection, antiviral discovery, and

development of vaccine. Therefore, in this investigation, the phylogenetic trees for animal coronaviruses with the novel coronavirus-2019 have been created using a number of available complete protein sequences of envelope (E), membrane (M), nucleocapsid (N) and spike (S) proteins.

2. METHODS AND MATERIALS

The number of available complete protein sequences of coronaviruses, envelope (E) (56 amino acid sequences, membrane (M) (55 amino acid sequences), and 54 protein sequences for each nucleocapsid (N) and spike (S) proteins of the novel coronavirus and animal coronaviruses isolated from different animal hosts were studied. All the sequences have been downloaded from and GenBank (http://www.ncbi.nlm.nih.gov/genbank/). A basic local alignment search tool has been used for similarity percentage identification (https://blast.ncbi.nlm.nih.gov/Blast.cgi); the datasets were excluded from eventual duplicated and very similar amino acid sequences. BioEdit program v7.0.5. has been used for editing the dataset of the above complete protein sequences [15]. Maximum likelihood (ML) methods were employed for the analyses because they allow for testing different phylogenetic hypotheses by calculating the probability of a given model of evolution generating the observed data and by comparing the probabilities of nested models by the likelihood ratio test. Generalized time reversible plus gamma distribution and invariant sites (+G+I) were used for ML trees reconstruction using MEGAX [16,17]. All the accession numbers of isolates are presented in the figures. For further confirmation and obtaining the final trees the protein domain sequences were aligned with MUSCLE, the phylogenetic tree was created with PhyML and TreeDyn (www.phylogeny.fr). Also, the branch support values were considered.

3. RESULTS

To understand the origin of the 2019nCoV, the 56 available complete protein sequences of animals and human coronaviruses were analysed through ML phylogenetic tree. As its shown in (Figures 1-4), coronaviruses isolated from animals and all the 2019-nCoV are located in three different clusters A, B and C.

Spike Protein

All animal and human coronaviruses used in this study are placed into three different group of clusters (A, B and C). In group A, all coronaviruses belong to different species of bats are seen clustered together. Group B has a larger group of different animals and bats. In this group, coronavirus; avian Infectious Bronchitis Viruses (IBV), infecting poultry are located together into one clade. Porcine coronaviruses isolated in different countries are gathered into one clade next to the feline coronaviruses. Both porcine and feline coronaviruses are neighbored by coronaviruses isolated from different species of bat (Figure 1). In addition, all feline and porcine coronaviruses clade is located next to the clade of sea mammals, including whale and bottlenose dolphin.

Interestingly, most of coronaviruses isolated from bat such as Bat_SARS_like_coronaviruses (MG772934.1_Bat_SARS-like_coronavirus_isolate_bat-SL-CoVZXC21/CHN, DQ022305.2 Bat_SARS_coronavirus_HKU3-1/HK/CHN/2005 and KY417146.1 Bat_SARS-like coronavirus-isolate-Rs4231/CHN/2016) and Pangolin coronaviruses (MT040333.1 Pangolin/China/2020) are located in cluster C (Figure 1). In the same cluster, coronavirus of *Rousettus* bat (NC_030886.1 Rousettus bat/Beijing/CHN/2018, KU762337.1 Rousettus_bat//CHN/2016 and MG762674.1 Rousettus_bat/_HKU9/Jinghong/CHN/2009) and MERS isolated from camel (KJ713299.1 MERS-CAMEL/ Saudi Arabia/ 2014) can be observed which is located in subclades of the same cluster.

While all avian Infectious Bronchitis Viruses (IBV) (MK644086.1, MF508703.1, JF732903.1, KX272465.1, MG233398.1, MH779856.1, GU393337.1, KR902510.1, AJ311317.1, NC_001451.1, FJ888351.1 and KT203557.1) and coronaviruses isolated from pigs

(LC063846.1, LC063847.1, KX580958.1, KR078300.1, MK409659.1, KT021232.1 and KF761675.1) are located in cluster B in two different clades which is located far from the 2019nCoV in the phylogenetic tree (Figure 1).

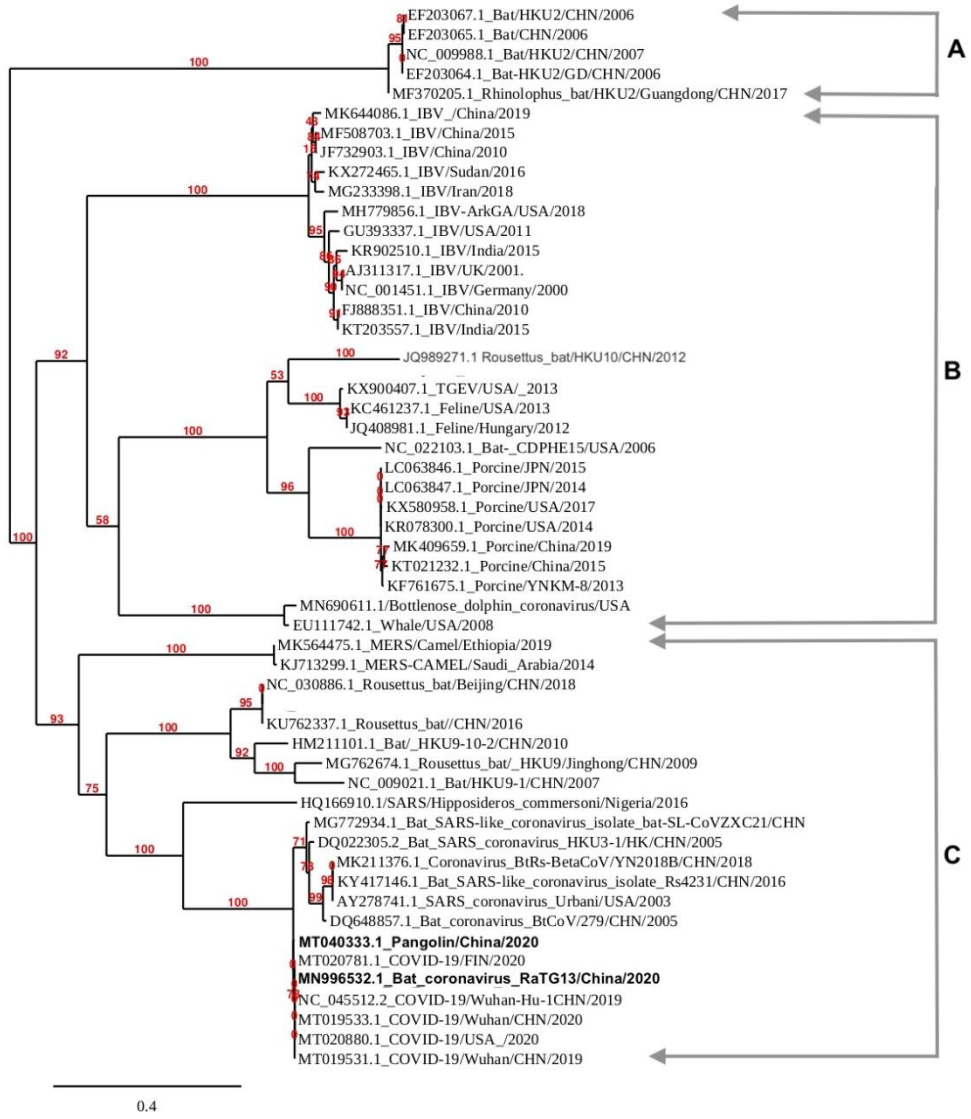

Figure 1: Spike protein phylogenetic tree analysis of the 2019-nCoV with common animal coronaviruses. The maximum likelihood method was used for making phylogenetic tree. Protein domain sequences were aligned with MUSCLE, the phylogenetic tree was created with PhyML and TreeDyn (www.phylogeny.fr). Branch support values are presented in red.

Nucleocapsid Protein

All coronaviruses isolated from different hosts are classified into three different group of clusters (A, B and C). In group A, a large group of different coronaviruses isolated from different bat species and human hosts including the 2019-nCoV are grouped together in this cluster. Coronaviruses isolated from poultry and sea mammal are in group B. In this group, sea-mammal coronaviruses are allocated together into one clade. Furthermore, the clade containing sea mammals, including whale and bottlenose dolphin is placed next to the clade of avian Infectious Bronchitis Viruses (IBV). In group C, all porcine coronaviruses are gathered together

in a single clade, which are accompanied by two different clades collecting coronaviruses isolated from bats and feline hosts (Figure 2).

As it is observed in the (Figure 2), all the 2019nCoV (MT019531.1, NC_045512.2, MT019533.1, MT020880.1 and MT020781.1) are allocated in clusters A, which are divided on some clades. Most of the 2019nCoV in cluster A are allocated next to Bat-SARS-like-coronaviruses(MG772934.1_Bat_SARS-like_coronavirus_isolate_bat-SL-CoVZXC21/CHN, DQ022305.2 Bat_SARS_coronavirus_HKU3-1/HK/CHN/2005 and KY417146.1 Bat_SARS-like coronavirus-isolate-Rs4231/CHN/2016) and Pangolin coronaviruses (MT040333.1 Pangolin/China/2020); whereas, MERS coronavirus isolated from camel (KJ713299.1 MERS-CAMEL/ Saudi_Arabia/ 2014) is located into another clade next to the clade where the 2019nCoV is allocated. Furthermore, all coronaviruses isolated from fruit bats (Rousettus bat) (NC_030886.1, KU762337., NC_009021.1, MG762674.1 and HM211101.1) are allocated in the near subclades.

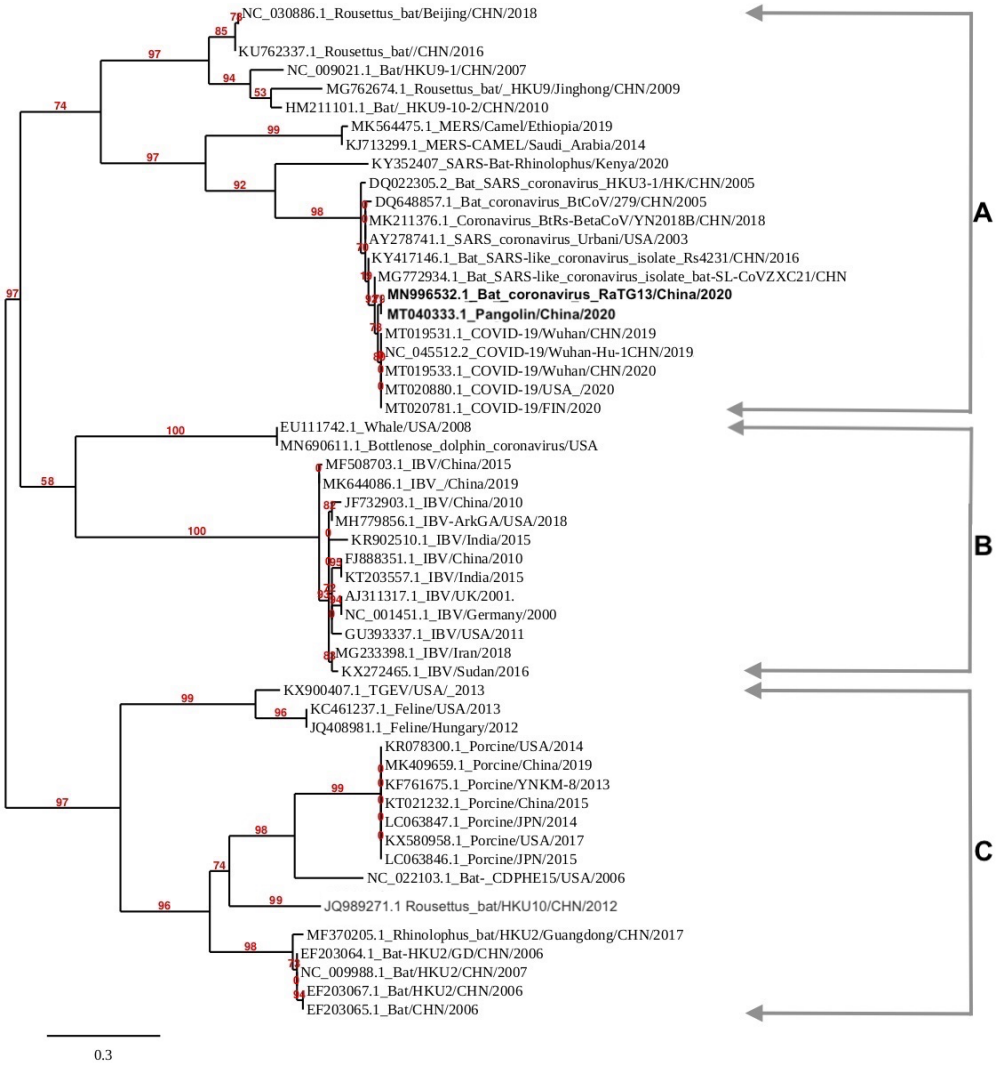

Figure 2: Phylogenetic tree analysis of Nucleocapsid proteins of the 2019-nCoV with common animal coronaviruses. The maximum likelihood method was used for making phylogenetic tree. Protein domain sequences were aligned with MUSCLE, the phylogenetic tree was created with PhyML and TreeDyn (www.phylogeny.fr). Branch support values are presented in red.

Envelope Protein

All amino acid sequences of coronaviruses Envelope protein downloaded from NCBI which are isolated from different hosts (human, animals, bats and sea mammals) in different countries are divided into three different group of clusters (A, B and C).

In group A, all coronaviruses isolated from feline are observed clustered together. Group B contains a large group of different coronaviruses isolated from porcine, bats and human hosts including the 2019-nCoV. Porcine coronaviruses are neighbored by coronaviruses isolated from different species of bat (Figure 3). Furthermore, the 2019-nCoV are in the same group which are far from porcine coronaviruses in different clades. Group C has a large group of different poultry and sea mammal coronaviruses. In this group, coronavirus; avian Infectious Bronchitis Viruses (IBV), infecting poultry are located together into one clade. In addition, all avian Infectious Bronchitis Viruses (IBV) clade is placed next to the clade of sea animals, including whale and bottlenose dolphin.

As it is shown in the (Figure 3), the phylogenetic tree placed the all 2019nCoV isolated from human host (MT019531.1, NC_045512.2, MT019533.1, MT020880.1 and MT020781.1) are in cluster B. Most of the 2019nCoV in cluster B are allocated next to Bat-SARS-like-coronavirus (MG772934.1_Bat_SARS-like_coronavirus_isolate_bat-SL-CoVZXC21/CHN, DQ022305.2 Bat_SARS_coronavirus_HKU3-1/HK/CHN/2005 and KY417146.1 Bat_SARS-like coronavirus-isolate-Rs4231/CHN/2016) and Pangolin coronaviruses (MT040333.1 Pangolin/China/2020); while, MERS coronavirus isolated from camel (KJ713299.1 MERS-CAMEL/ Saudi_Arabia/ 2014) is located into another clade next to the clade where the 2019nCoV is allocated. In addition, all coronaviruses isolated from fruit bats (Rousettus bat) are also observed in the subclade of the same cluster.

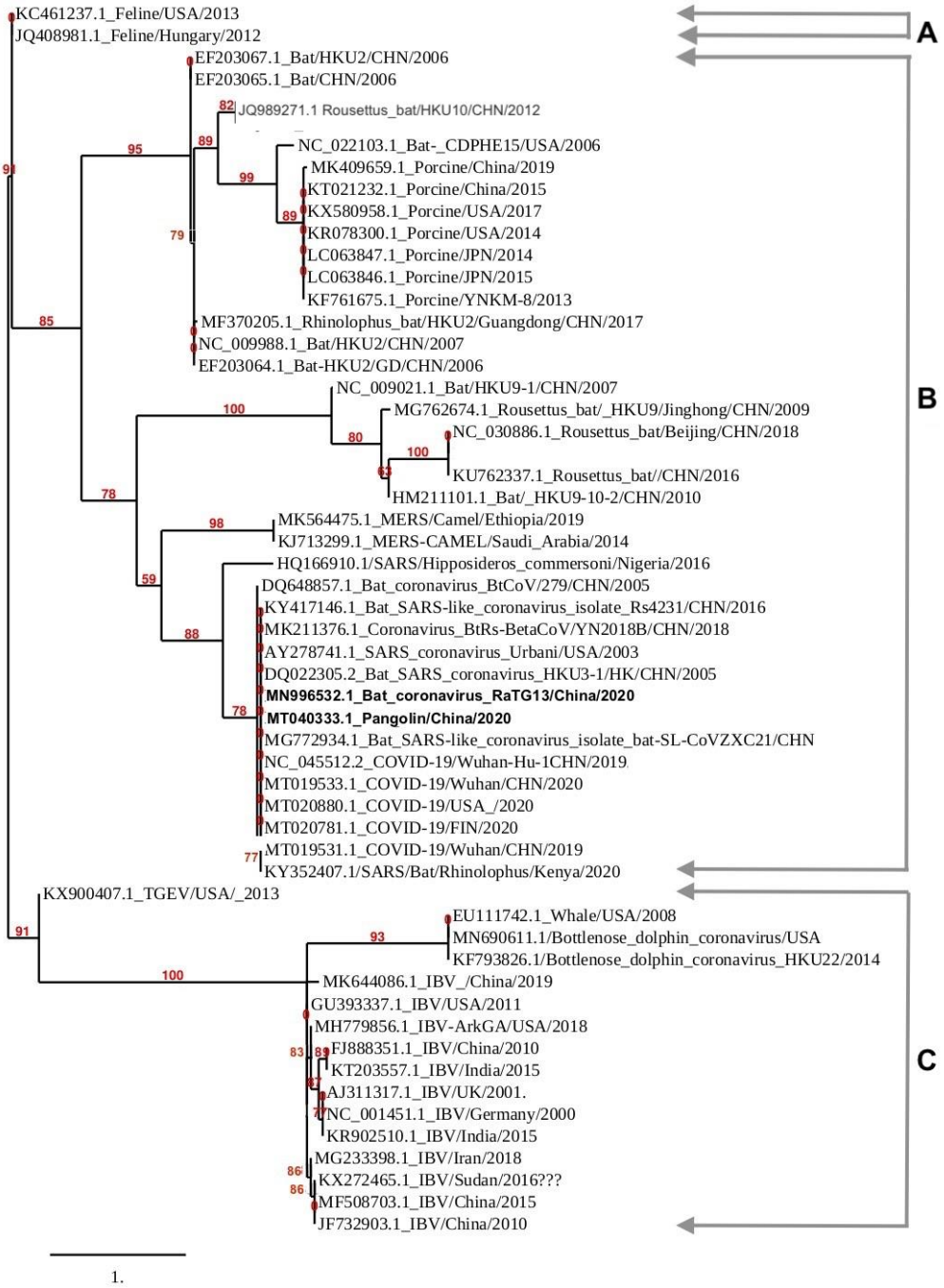

Figure 3: Phylogenetic tree analysis of Envelope proteins of the 2019-nCoV with common animal coronaviruses. The maximum likelihood method was used for making phylogenetic tree. Protein domain sequences were aligned with MUSCLE, the phylogenetic tree was created with PhyML and TreeDyn (www.phylogeny.fr). Branch support values are presented in red.

Membrane Protein

Three groups (A, B and C) of coronaviruses of different hosts are built by phylogenetic tree. Phylogenetic analysis allocated all human coronaviruses including the 2019-nCoV and Bat-SARS-like-coronavirus together in a certain clade of the cluster A. In this protein analysis, the

clades of different species of bat and camel coronaviruses are collected together in group B. In this cluster, sea animal coronaviruses, avian Infectious Bronchitis Viruses (IBV) are observed into one clade. While, all porcine and bat coronaviruses are in group C (Figure 4).

As it is observed in the (Figure 4), all the 2019nCoV (MT019531.1, NC_045512.2, MT019533.1, MT020880.1 and MT020781.1) are allocated in clusters A, which are divided on some clades. all of the 2019nCoV in cluster A are closely related to Bat-SARS-like- coronavirus (MG772934.1_Bat_SARS-like_coronavirus_isolate_bat-SL-CoVZXC21/CHN, DQ022305.2 Bat_SARS_coronavirus_HKU3-1/HK/CHN/2005 and KY417146.1 Bat_SARS-like coronavirus-isolate-Rs4231/CHN/2016) and Pangolin coronaviruses (MT040333.1 Pangolin/China/2020) because all of them are gathered next to each other. In contrast to envelope (E), nucleocapsid (N) and spike (S) proteins; membrane protein of coronaviruses isolated from avian Infectious Bronchitis Viruses (IBV) (MK644086.1, MF508703.1, JF732903.1, KX272465.1, MG233398.1, MH779856.1, GU393337.1, KR902510.1, AJ311317.1, NC_001451.1, FJ888351.1 and KT203557.1), sea animals including dolphin (MN690611.1/Bottlenose_dolphin_coronavirus/USA) and whale (EU111742.1 Whale/USA/ 2008) are allocated in the clade next to the clade where the 2019nCoV is allocated (Figure 4 group A). In addition, all coronaviruses isolated from fruit bats (Rousettus bat) and African (MK564475.1_MERS/Camel/Ethiopia/2019) and Arabic (KJ713299.1_MERS-CAMEL/Saudi_Arabia/2014) camels are allocated in the certain clades of the cluster B. Also, the pig, feline, and some bat coronaviruses are clustered together the cluster group C of the Figure 4.

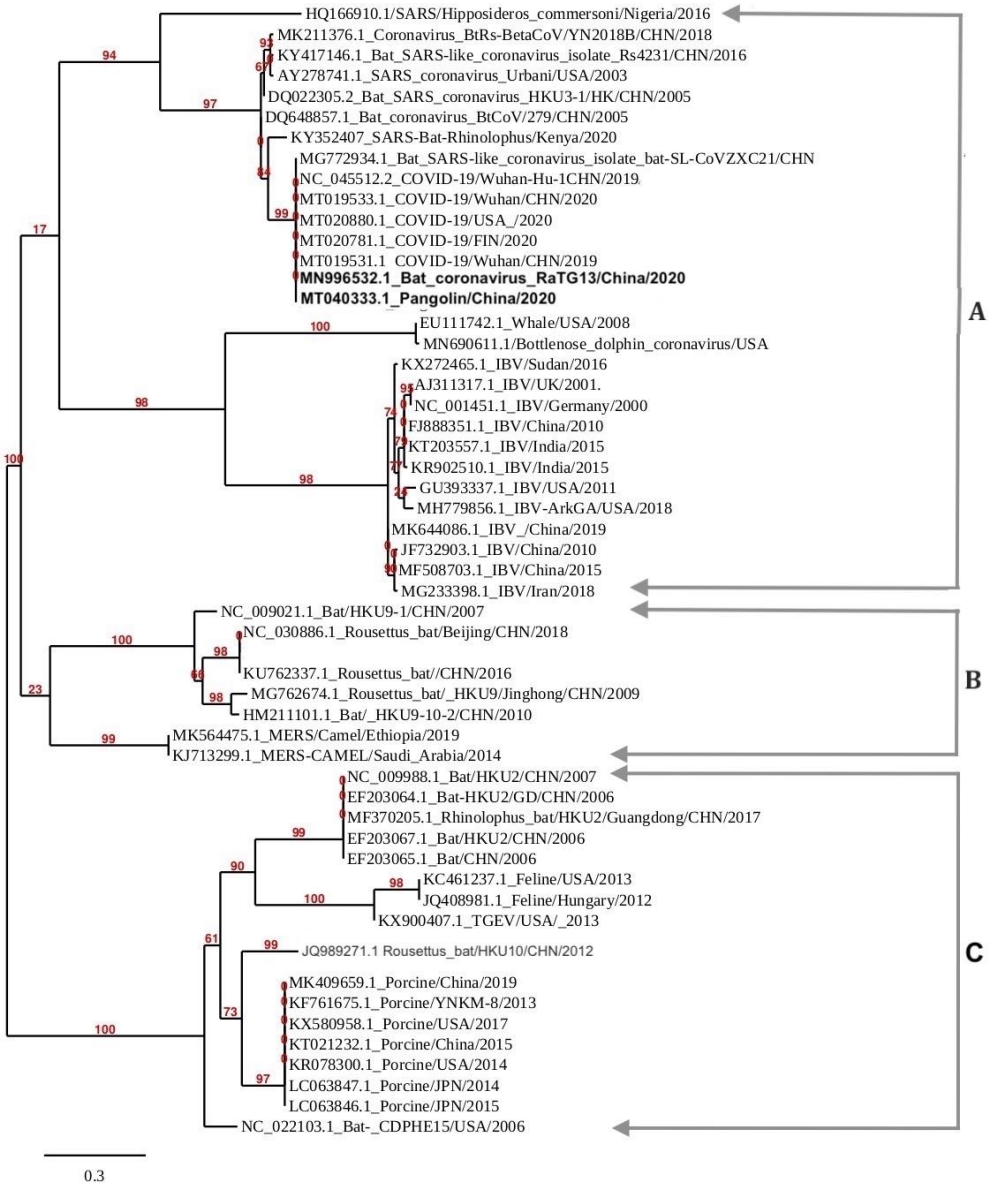

Figure 4: Phylogenetic tree analysis of Membrane proteins of the 2019-nCoV with common animal coronaviruses. The maximum likelihood method was used for making phylogenetic tree. Protein domain sequences were aligned with MUSCLE, the phylogenetic tree was created with PhyML and TreeDyn (www.phylogeny.fr). Branch support values are presented in red.

4. DISCUSSION

Due to the recent pandemic outbreak of deadly COVID-19, it is necessary to understand the epidemiology of the 2019nCoV. So, this study investigated the possible origin of this virus through the phylogenetic analysis that exists in both humans (2019nCoV) and animals. To discover the link between human and animal coronaviruses, the complete protein sequences of envelope (E), membrane (M), nucleocapsid (N) and spike (S) proteins coronaviruses in different animals was analysed and compared to the 2019nCoV through phylogenetic analysis (Figures 1,2,3, and 4).

The phylogenetic tree of the Spike protein (Figures 1) reveals that most of 2019-nCoV are allocated in cluster B with very closely related animal coronaviruses, Bat-SARS-like-coronaviruses (MG772934.1_Bat_SARS-like_coronavirus_isolate_bat-SL-CoVZXC21/CHN, DQ022305.2 Bat_SARS_coronavirus_HKU3-1/HK/CHN/2005 and KY417146.1 Bat_SARS-like coronavirus-isolate-Rs4231/CHN/2016) [18,19,20,21] and Pangolin coronaviruses (MT040333.1 Pangolin/China/2020) [22]. This means that the viruses of this cluster are genetically identical and it is closely related to that coronavirus isolated from in different areas such as China especially during the year 2020. This may conclude that 2019nCoV was mutated from passing through different animal hosts until reached and adapted itself in human host. Periodically in around every five years, the virus evolved from SARS to the new generation of coronavirus in human.

Data analysis of Envelope protein (Figure 3); all coronaviruses are generally classified into two major groups; A and B. All the 2019nCoV are clustered in group A in a similar manner of Spike protein phylogenetic tree, which supports the hypothesis that Spike and Envelope proteins are originated from Bat-SARS-like-coronaviruses and Pangolin coronaviruses. In addition, coronaviruses isolated from pigs are allocated in the same group, but in a very far clade. This concludes that is less possible to be an origin of 2019-nCoV.

According to Nucleocapsid protein phylogenetic tree (Figure 2), all coronaviruses are clustered in two different groups. Group A and B clustered all viruses that cause diseases in pig, poultry, sea mammals, bat and human. Interestingly, all the 2019nCoV coronaviruses are allocated in group A with Bat-SARS-like-coronaviruses (MG772934.1_Bat_SARS-like_coronavirus_isolate_bat-SL-CoVZXC21/CHN, DQ022305.2 Bat_SARS_coronavirus_HKU3-1/HK/CHN/2005 and KY417146.1 Bat_SARS-like coronavirus-isolate-Rs4231/CHN/2016) [18,19,20,21] and Pangolin coronaviruses (MT040333.1 Pangolin/China/2020) [22]. In which, all 2019nCoV are clustered in subclades with coronavirus isolated in *Rousettus* Bat (fruit bat) in 2016 in China (KU762337.1 Rousettus bat//CHN/2016) [23], which are less closely related. This may give a strong evidence that 2019nCoV is originated from Bat_SARS-like coronavirus especially Bat_SARS-coronavirus-RaTG13 (MN996532.1/CHN/2020) [24]. In addition, coronaviruses isolated from avian and sea animals are clustered in the same group, but in in differed clades. The 2019nCoV isolated human hosts is not closely related to coronavirus in Bottlenose Dolphin in the USA (MN690611.1 Bottlenose_Dolphin_coronavirus/USA) [25]. This is for the first time is revealed that there is a genetic homology between the 2019nCoV and marine coronaviruses, but it cannot be a strong evidence to be an origin of the 2019nCoV. Interestingly, a similar result was obtained with Membrane protein (Figure 4), which strengths the postulation that the 2019-nCoV is more possible to come from coronavirus of Bat_SARS-like and Pangolin coronaviruses rather than avian Infectious Bronchitis Viruses (IBV) of poultry. Our findings are comparable to many international studies outcomes [18,21,23,24].

5. CONCLUSION

The result of our study along with analysis of other international researches, enforcing Bat_SARS-like coronavirus especially Bat_SARS-coronavirus-RaTG13 and/or Pangolin 2020 isolates are the probable origin of infection and this may give a clarification of the 2019-nCoV transmission dynamics. However, the postulate of originating 2019nCoV from *Rousettus* bat is weaker because of its less closely related to the 2019-nCoV. This outcome will aid to understanding of the possible origin of the virus and infection control policy. Therefore, to find out the close relationship between the 2019-nCoV and other coronaviruses, amino acids sequence alignment for all structural proteins of the virus and phylogenetic tree of complete genome obtained in every country is required.

ACKNOWLEDGEMENT

We acknowledge the presidency of Sulaimani Polytechnic University and especially Assistant Professor Dr. Alan Faraydoon Ali for his great help and academic support to performing this research. Also, we would like to appreciate Mam Humanitarian Foundation for their financial and logistic support.

REFERENCE

[1] C. Drosten, S. Günther, W. Preiser, et al. Identification of a novel coronavirus associated with severe acute respiratory syndrome. N Engl J Med.; 348:1967-1976, 2003.
[2] Y. Chen, Q. Liu, D. Guo. Emerging coronaviruses: genome structure, replication, and pathogenesis. J Med Virol., 2020. https://doi.org/10. 1002/jmv.25681
[3] JF. Chan, S. Yuan, KH. Kok, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to- person transmission: a study of a family cluster. Lancet., 2020. https:// doi.org/10.1016/S0140-6736(20)30154-9
[4] C. Huang, Y. Wang, X. Li, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet., 2020. https://doi. org/10.1016/S0140-6736(20)30183-5
[5] N. Zhu, D. Zhang, W. Wang, et al. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med., 2020. https://doi.org/ 10.1056/NEJMoa2001017
[6] Coronavirus disease 2019 (COVID-19) Situation Report – 74, World Health Organization (WHO). https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200403-sitrep-74-covid-19-mp.pdf?sfvrsn=4e043d03_12
[7] H. Lu, CW. Stratton, YW. Tang. Outbreak of pneumonia of unknown etiology in Wuhan China: the mystery and the miracle. J Med Virol., 2020. https://doi.org/10.1002/jmv.25678
[8] W. Ji, W. Wang, X. Zhao, J. Zai, X. Li. Homologous recombination within the spike glycoprotein of the newly identified coronavirus may boost cross-species transmission from snake to human. J Med Virol., 2020. https://doi.org/10.1002/jmv.25682
[9] DC. Scales, K. Green, AK. Chan, SM. Poutanen, D. Foster, K. Nowak, et al. Illness in intensive-care staff after brief exposure to severe acute respiratory syndrome. Emerging Infect Dis.; 9:1205–1210, 2003. https://doi.org/10.3201/eid0910.030525
[10] K. Mizumoto, G. Chowell, Estimating Risk for Death from 2019 Novel Coronavirus Disease, China, January–February 2020. Retrieved from Volume 26, Number 6—June 2020. Center for Disease Control and Prevention –CDC. https://doi.org/10.3201/eid2606.200233
[11] Y. Chen, Q. Liu, D. Guo, Emerging coronaviruses: Genome structure, replication, and pathogenesis. J. Med. Virol., 2020. https://doi.org/10.1002/jmv.25681
[12] J. Cai, W. Sun, J. Huang, M. Gamber, J. Wu, G. He, Early Release - Indirect Virus Transmission in Cluster of COVID-19 Cases, Wenzhou, China. Emerging Infect. Dis. 26(6), 2020. https://doi.org/10.3201/eid2606.200412
[13] K. Marie, B. Holly, M. Claire, G. Susan, W. John. Middle East Respiratory Syndrome Coronavirus Transmission. Emerging Infect. Dis. www.cdc.gov/eid • Vol. 26, No. 2, 2020. https://doi.org/10.3201/eid2602.190697
[14] P. Forster, L. Forster, C. Renfrew, M. Forster. Phylogenetic network analysis of SARS-CoV-2 genomes. PNAS, www.pnas.org/cgi/doi/10.1073/pnas.2004999117, 2020.
[15] TA. Hall. BioEditA user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser.; 41:95-98, 1999. https://doi.org/10.14601/Phytopathol_Mediterr-14998u1.29
[16] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol. Biol. Evol. 2018;35(6):1547-1549. https://doi.org/10.1093/molbev/msy096.
[17] A. Dereeper, et al., Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 2008 Jul 1; 36 (Web Server issue): W465-9.
[18] P. Zhou, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature, 2020. https://doi.org/10.1038/s41586-020-2012-7.
[19] D. Hu, C. Zhu, L. Ai L, et al. Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats. Emerg Microbes Infect 2018; 7: 154.
[20] S. Lau, K. Woo, Y. Li, H. Huang, B. Tosi, S. Wong, S. Wong, K. Leung, H. Chan, K. Yuen. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc. Natl. Acad. Sci. USA 102:14040–14045, 2005.
[21] B. Hu, L. Zeng, X. Yang, et al. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLoS Pathog, 13, Article e1006698, 2017.
[22] T. Zhang, Q. Wu, Z. Zhang. Pangolin homology associated with2019-nCoV. BioRxiv, 2020. https://doi.org/10.1101/2020.02.19.950253.
[23] C. Huang, W. Liu W. Xu, T. Jin, Y. Zhao, J. Song, Y. Shi, W. Ji, H. Jia, Y. Zhou, et al. A bat-derived putative

cross-family recombinant coronavirus with a Reovirus gene. PLoS Pathog, 12, p. e1005883, 2016. https://doi.org/10.1371/journal.ppat.1005883.

[24] K. Andersen, A. Rambaut, W. Lipkin, E. Holmes, R. Garry, The Proximal Origin of SARS-CoV-2. ARTIC Network, 2020. https://doi.org/10.1038/s41591-020-0820-9

[25] P. Woo, S. Lau, C. Lam, A. Tsang, S. Hui, R. Fan, P. Martelli, K. Yuen. Discovery of a novel bottlenose dolphin coronavirus reveals a distinct species of marine mammal coronavirus in Gammacoro- navirus. J Virol 88:1318 –1331, 2014. http://dx.doi.org/10.1128/JVI.02351-13.