SARS-CoV-2 Molecular and Phylogenetic Analysis in COVID-19 Patients: A Preliminary Report from Iran

Alireza Tabibzadeh[1], Farhad Zamani[1], Azadeh Laali[1], Maryam Esghaei[1], Fahimeh Safarnezhad Tameshkel[1], Hossein Keyvani[1], Mahin Jamshidi Makiani[1], Mahshid Panahi[1], Nima Motamed[2], Dhayaneethie Perumal[3], Mahmoodreza Khoonsari[1], Hossein Ajdarkosh[1], Masodreza Sohrabi[1], Behrooz Ghanbari[1], Shokoufeh Savaj[1], Alireza Mosavi-Jarrahi[4], and Mohammad Hadi Karbalaie Niya[1]

[1]Iran University of Medical Sciences
[2]Zanjan University of Medical Sciences
[3]Kingston University
[4]Shahid Beheshti University of Medical Sciences

May 6, 2020

Abstract

Background: The aim of the current study was to investigate and track the SARS-CoV-2 in Iranian Coronavirus Disease 2019 (COVID-19) patients using molecular and phylogenetic methods. Methods: We enrolled seven confirmed cases of COVID-19 patients for the phylogenetic assessment of the SARS-CoV-2 in Iran. The nsp-2, nsp-12, and S genes were amplified using one-step RT-PCR and sequenced using Sanger sequencing method. Popular bioinformatics software were used for sequences alignment and analysis as well as phylogenetic construction. Results: The mean age of the patients in the present study was 60.42 ± 9.94 years and 57.1% (4/7) were male. The results indicated high similarity between Iranian and Chinese strains. We could not find any particular polymorphisms in the assessed regions of the three genes. Phylogenetic trees by neighbor-joining and maximum likelihood method of nsp-2, nsp-12, and S genes showed that there are not any differences between Iranian isolates and those of other countries. Conclusion: As a preliminary phylogenetic study in Iranian SARS-CoV-2 isolates, we found that these isolates are closely related to the Chinese and reference sequences. Also, no sensible differences were observed between Iranian isolates and those of other countries. Further investigations are recommended using more comprehensive methods and larger sample sizes.

1- Introduction

Coronaviruses are responsible for respiratory tract infections (RTIs). In the December of 2019, cases of pneumonia with unknown origin were reported in Wuhan, China [1-6]. The disease started spreading in the China first and then around world. Until March 20, 2020, this virus resulted in 234'073 confirmed cases and 9'840 deaths worldwide; the statistics are 18'407 confirmed cases and 1'284 deaths in Iran [7]. This is while the number of studies on this novel coronavirus is growing all around the world[8-11].

Coronaviruses can infect the respiratory and intestinal epithelial cells and generates a wide range of the clinical presentations in humans[12]. Coronaviruses are classified in *Nidovirales*order and *Cornidoviridae* sub-order. *Orthocoronavirinae*subfamily in *Coronaviridae* family is divided into four genera. Severe acute respiratory syndrome-related coronavirus type 2 (SARS-CoV-2), and other SARS-like coronaviruses are classified in*Sarbecovirs* subgenus in *Betacoronaviruses* genera[10,13,14]. Endemic coronaviruses, such as HCoV-NL63, HCoV-229E, HCoV-OC43, and HKU1 can induce upper respiratory tract infections (UTRI) with mild respiratory manifestation in immune competent adults while infection with the epidemic coronaviruses, like

1

SARS-CoV and Middle East respiratory syndrome-related coronavirus (MERS-CoV), could lead to the lower respiratory tract infections (LRTI) as well as serious respiratory manifestations [15,16]. In endemic areas for coronaviruses, it is suggested that the HCoV-229E and HCoV-OC43 are responsible for 15-29% of the common cold cases, while during the epidemic coronaviruses, there is another scenario at work [17]. The SARS-CoV epidemic started from Guangdong, China, in 2002-2003 and resulted in 8'089 infected cases and a 9.6% mortality rate [7]. Also, the MERS-CoV in 2012 firstly started in Middle East and infected 2'465 cases demonstrating a high mortality [18,19]. All the mentioned pandemics as well as the current pandemic condition with SARS-CoV-2 reflect the importance of following coronaviruses in human population. The aim of the current study was to investigate and track the SARS-CoV-2 in Iranian Coronavirus Disease 2019 (COVID-19) patients using molecular and phylogenetic methods.

2- Materials and Methods

2-1- Cases

We enrolled seven cases between February-March, 2020, with the clinical presentations including fever, cough, and radiologic features of the COVID-19 who referred to Firoozgar Hospital, Iran University of Medical Science, Tehran, Iran, after obtaining the written consent from all of the enrolled cases. The study was approved by the Ethical committee of the Iran University of Medical Science, Tehran, Iran (ethical code no. IR.IUMS.REC.1398.1344). The laboratory records including biochemical and hematological parameters, blood gases analyses, and cardiac markers were recorded. Also, all virological findings were obtained and used for further investigations.

2-2- SARS-CoV-2 pathogen detection

Respiratory samples, including the nasopharyngeal and throat swabs, were obtained from the enrolled patients. The SARS-CoV-2 primary detections were performed using the standard real-time PCR method recommended by CDC using Rotor-Gene-Q 6000 thermocycler (Corbett, Australia) according to the recommended protocol [20]. After the SARS-CoV-2 was confirmed in cases, the samples were stored in -70ºC for further analysis.

2-3- RT-PCR and sequencing

The RNA was extracted from the clinical samples using the Extraction of viral DNA/RNA kit (FAVORGEN Biotech Corporation, Taiwan) based on the manufactures' instructions. All the extracted RNAs were evaluated for the quality of the extraction via spectrophotometry method using NanoDrop ND-1000® (Thermo Fisher Scientific Inc., Waltham, MA, USA). PrimeScript One Step RT-PCR Kit (Takara Bio, Japan) was used for the synthesis of complementary DNA (cDNA) and PCR according to the manufactures' instructions.

We used the specific primers for amplifying the RdRp (RNA dependent RNA polymerase), the S protein, and nsp-2 (non-structural protein 2)[21,22]. Conventional PCR method was used for amplification of 344 bp region of RNA-dependent RNA polymerase (RdRp) gene via forward 5'-CAA GTG GGG TAA GGC TAG ACT TT-3' and reverse 5'-ACT TAG GAT AAT CCC AAC CCA T-3' primers; and for the 158 bp region of Spike (S) protein forward 5'-CCT ACT AAA TTA AAT GAT CTC TGC TTT ACT-3' and reverse 5'- CAA GCT ATA ACG CAG CCT GTA-3' primers; and a nested-PCR for amplification of 346 and 322 bp region of nsp-2 gene by first round forward 5'-CTC GAA CTG CAC CTC ATG G-3' and reverse 5'- CAG AAG TTG TTA TCG ACA TAG C-3' primers and second round 5'-ACC TCA TGG TCA TGT TAT GG-3' and 5'-GAC ATA GCG AGT GTA TGC C-3' primers.

A total volume of 50 ul Master Mix was performed using 25ul 2x EnzymeMix (Takara Bio, Japan) for one-step RT PCR of RdRp, S and the first round of nsp-2 PCR, and 2x Super PCR Mastermix (Yekta Tajhiz Azma Co., Iran) for the second round of PCR for nsp-2 gene amplification, 1ul (10 pMol concentration) of each primers, 4 µl (0.2–0.5 µM concentration) of extracted specimens, and distilled RNase/DNase free water added to reach the rest of total volume.

A Bio-Rad (T100 Thermal Cycler) was performed for heating program. Heating protocol, for RdRp, S, and

the first round of nsp-2 PCR, was 1 step at 50°C 40min, 94°C 5min, 40 cycles at 94°C 30s, 58°C 30s, 72°C 30s, and 1 step at 72°C 5 min. The second round of PCR for nsp-2 gene included 94°C 3 min, 94°C 20s, 58°C 20s, 72°C 20s, and 72°C 3 min. Visualization of PCR products was performed using 1.5% concentration of agarose gel into an electrophoresis system and stained using safe stain (YTA, Yekta Tajhiz Azma Co., Iran) behind UV translaminator.

2-4- Phylogenetic assessment

PCR products underwent purification using High Pure PCR Product Purification Kit (Roche Diagnostic GmbH, Mannheim, Germany) corresponding to the manufacturer's instructions. Also, an ABI 3730 XL sequencer was used for bidirectional sequencing. Raw data were trimmed and analyzed using popular bio-informatics software CLC workbench 5, and Basic Local Alignment Search Tool (BLAST). All sequences were submitted into GeneBank database (https://www.ncbi.nlm.nih.gov/genbank/) and released accession numbers were used to draw phylogenetic tree. SARS-CoV-2 reference sequence (NC_045512.2) and other coronavirus sequences were obtained from NCBI (https://blast.ncbi.nlm.nih.gov) database. Moreover, MEGA X software was used for multiple sequence alignment (MSA) and the phylogenetic trees were drown using 1000 replicate bootstrap method. In addition, the neighbor joining and maximum likelihood methods were carried out for similarity assessment and evolution analysis.

3- Results

The mean age ± SD of the patients in the present study was 60.42±9.94 years and 57.1% of patients were male. The mean of the white blood cell count (WBC) and platelets count were 6.97±4.01 X1000 cells/mm³ and 181.4±28.87 X1000 Platelets/mm³, respectively. Furthermore, the differential WBC indicated that the polymorph nuclear leukocytes (PMN) and lymphocytes were 75±11.8% and 18±8.9%, respectively. The investigations of the inflammatory markers were assessed by qualitative C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR). The CRP was positive in all patients and the ESR was 33.2±20.4 millimeters per hour (mm/hr). Also, the sequence analyses were performed on 3 out of 7 cases due to insufficient quality of the rest of the samples.

Obtained sequence was used for multiple sequence alignment (MSA) as illustrated in Figure 1. The MSA analysis could not show any particular polymorphism in the assessed regions of all three genes (only nsp-2 is illustrated in Figure 1). Furthermore, the neighbor joining phylogenetic analysis showed the highest similarity between Iranian sequences for nsp-2 and the obtained sequence from Wuhan of Hubei province, China, as illustrated in Figure 2. Also, the neighbor joining phylogenetic analysis showed the most similarity for RdRp between Iranian sequences and those from China and USA (Figure 3). Also, the same results were observed in the assessment of the S gene (Figure 4). The phylogenetic assessment indicated that the bat coronaviruses and other SARS-related bat coronaviruses are highly similar to each other regarding their ancestors (Figure 5).

4- Discussion

The COVID-19 statistics show a high distribution of this disease in Iran and all around the word [7]. The SARS-CoV-2 shows a great potential for the transmission due to its highly contagious nature[23]. Considering the importance of the public knowledge about the SARS-CoV-2 epidemiologic and its phylogenetic analysis, in the present preliminary study in Iran, we aimed to investigate and track the SARS-CoV-2 in Iranian COVID-19 patients using molecular and phylogenetic methods.

The neighbor joining phylogenetic analysis demonstrated the highest similarity between Iranian sequences for nsp-2 and the obtained sequence from Wuhan of Hubei province, China (NC_045512) as well as a familial cluster of SARS-CoV-2 in Hong Kong- Shenzhen Hospital, Shenzhen, Guangdong, China, whose direct relationship with the seafood market of Wuhan of Hubei province, China (MN975262) [21] is confirmed, as illustrated in Figure 2. Also, the neighbor joining phylogenetic analysis showed the most similarity for RdRp with obtained sequence from MN938384 [21] and MN985325, which were the first obtained sequences from the USA (Figure 3). Analysis of the S sequence showed a high similarity with NC_045512 and MN938384

3

sequences (Figure 4). The maximum likelihood phylogenetic indicates that the Bat coronavirus RaTG13 (MN996532), Bat coronavirus Cp/Yunnan2011 (JX993988), and other SARS-related bat coronaviruses have the highest similarity revealing a close relationship among their ancestors considering the RdRp (Figure 5).

As a major consideration, coronaviruses encode six main and variable numbers of accessory open reading frames and in case of the SARS-CoV-2, from 5' to 3' ORFs including ORF1a/b, spike protein or surface glycoprotein, membrane , envelope and nucleocapsid main ORFs, and 3a, 6, 7a, 7b, 8, and 10 accessories [24-26]. The S glycoprotein is responsible for the attachment of the virus into the host cell. The S glycoprotein is cleaved to S1 and S2 in SARS-CoV and SARS-CoV-2 by host TMPRSS2 serine protease [27,28]. The S1 domain is vital for the virus attachment as receptor in host cell and the S2 domain is useful for the membrane fusion [27,28]. The SARS-CoV and -CoV-2 use the Angiotensin Converting Enzyme 2 for the attachment [29]. The conducted studies reported that the SARS-CoV-2 S protein shows 80% and 76% similarity with bat-SL-CoVZXC21 and SARS-CoV, respectively [30]. Regardless of S protein, the ORF 1a/b of the virus encodes 16 nonstructural proteins, which is known as nsp (nonstructural protein). The nsp-12 of the SARS-CoV-2 plays a role as dependent polymerase, while the nsp-2 function is not clear yet [30]. Our current phylogenetic study showed that the Iranian isolates have the most similarity with the Chinese isolates. The study conducted by Eden *et al.* [31] showed that the isolated sequences from the Australian and New Zealand patients who travelled from Iran showed a clade different from those of other circulating strains. They also suggested that three major substitutions in the coding and non-coding regions of the nsp-2 include G1397A, T28688C, and G29742T in this particular clade. The results of the current study did not show any differences or clades in the assessed sequences between Iranian strains and those obtained from other countries in the phylogenetic study of the RdRp, nsp-2, and S proteins. However, it should be noted that the Eden *et al.* [31]phylogenetic study was conducted using the whole genome of the virus in contrast to our study. The MSA assessment of the nsp-2 for the three suggested substitutions by Eden *et al.* could not be assessed due to differences between our sequencing regions. Furthermore, Bal *et al.* [9] could find three nucleotide variations in some of the assessed samples in France. They suggested that these variations might be associated with intra host evolution of the virus. In another study, conducted by Sardar *et al* [32], it is suggested that some of the mutations in isolated viruses could be unique. For instance, a substation in S protein C24351T is unique in Indian isolates.

A major limitation of the current study was the limited region targeted for the nucleotide sequencing. Using high throughput methods and next generation sequencing (NGS) could result in a great advancement in the assessment of these variations [33]. These comprehensive methods can also further our knowledge on Iranian SARS-CoV-2 isolates.

In conclusion, the results of the present study indicated that the SARS-CoV-2 sequences from Iran have great similarity with Chinese and reference sequences although limited regions were sequenced and no variations were observed among them. To sum up, we found no identified differences between Iranian isolates and those from other regions. The current investigation was conducted as a preliminary study in Iran; further studies are to be performed to increase our knowledge regarding SARS-CoV-2 isolates.

5- Conflict of interest

There is no conflict of interests between authors in the present study.

6- Acknowledgement

7- Contributions

Study design: MHKN, AT; Laboratory testing: FST, MP, ME, HK; Data collection: AL, MJM, FZ, MS; Data analysis: NM, AMJ, SS; Writing: DP, MHKN, BG; Revision and finalize the manuscript: AMJ, DP; providing patients to a study: MK, AL, MJM, HA.

8- References

1. Hui D, Madani T, Ntoumi F, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health-The latest 2019 novel coronavirus outbreak in Wuhan, China. *International journal of infectious diseases: IJID: official publication of the International Society for Infectious Diseases.* 2020;91:264.

2. Zaki AM, Van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *New England Journal of Medicine.* 2012;367(19):1814-1820.

3. Woo PC, Lau SK, Chu C-m, et al. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *Journal of virology.* 2005;79(2):884-895.

4. van der Hoek L, Pyrc K, Jebbink MF, et al. Identification of a new human coronavirus. *Nature medicine.* 2004;10(4):368-373.

5. van der Hoek L. Human coronaviruses: what do they cause?*Antiviral therapy.* 2007;12(4 Pt B):651.

6. Bradburne A, Bynoe M, Tyrrell D. Effects of a" new" human respiratory virus in volunteers. *British medical journal.* 1967;3(5568):767.

7. Zhong N, Zheng B, Li Y, et al. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. *The Lancet.* 2003;362(9393):1353-1358.

8. Biscayart C, Angeleri P, Lloveras S, Chaves TDSS, Schlagenhauf P, Rodríguez-Morales AJ. The next big threat to global health? 2019 novel coronavirus (2019-nCoV): What advice can we give to travellers? – Interim recommendations January 2020, from the Latin-American society for Travel Medicine (SLAMVI). *Travel Medicine and Infectious Disease.* 2020.

9. Bal A, Destras G, Gaymard A, et al. Molecular characterization of SARS-CoV-2 in the first COVID-19 cluster in France reveals an amino-acid deletion in nsp2 (Asp268Del). *bioRxiv.*2020:2020.2003.2019.998179.

10. Lu H, Stratton CW, Tang YW. Outbreak of Pneumonia of Unknown Etiology in Wuhan China: the Mystery and the Miracle. *Journal of Medical Virology.*

11. Bogoch II, Watts A, Thomas-Bachli A, Huber C, Kraemer MU, Khan K. Pneumonia of Unknown Etiology in Wuhan, China: Potential for International Spread Via Commercial Air Travel. *Journal of Travel Medicine.* 2020.

12. Wevers BA, van der Hoek L. Recently discovered human coronaviruses.*Clinics in laboratory medicine.* 2009;29(4):715-724.

13. Luk HK, Li X, Fung J, Lau SK, Woo PC. Molecular epidemiology, evolution and phylogeny of SARS coronavirus. *Infection, Genetics and Evolution.* 2019.

14. van Boheemen S, de Graaf M, Lauber C, et al. Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. *MBio.*2012;3(6):e00473-00412.

15. Forni D, Cagliani R, Clerici M, Sironi M. Molecular evolution of human coronavirus genomes. *Trends in microbiology.*2017;25(1):35-48.

16. Su S, Wong G, Shi W, et al. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends in microbiology.*2016;24(6):490-502.

17. Monto AS. Medical reviews. Coronaviruses. *The Yale journal of biology and medicine.* 1974;47(4):234.

18. Azhar EI, Hui DSC, Memish ZA, Drosten C, Zumla A. The Middle East Respiratory Syndrome (MERS). *Infectious Disease Clinics of North America.* 2019;33(4):891-905.

19. Reusken CB, Haagmans BL, Müller MA, et al. Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. *The Lancet infectious diseases.*2013;13(10):859-866.

20. Division of Viral Diseases CfDCaP, 2020. https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html.

21. Chan JF-W, Yuan S, Kok K-H, 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.*2020;395(10223):514-523.

22. Nao N, Shirato K, Katano H, Matsuyama S, Takeda M. Detection of second case of 2019-nCoV infection in Japan (corrected version). In:2020.

23. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine.*2020.

24. Wu A, Peng Y, Huang B, et al. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host & Microbe.* 2020.

25. Chen Y, Liu Q, Guo D. Coronaviruses: genome structure, replication, and pathogenesis. *J Med Virol.* 2020.

26. Hussain S, Chen Y, Yang Y, et al. Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus. *Journal of virology.*2005;79(9):5288-5295.

27. Glowacka I, Bertram S, Müller MA, et al. Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. *Journal of virology.* 2011;85(9):4122-4134.

28. 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.

29. Liu Z, Xiao X, Wei X, et al. Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. *J Med Virol.* 2020.

30. Chan JF-W, Kok K-H, Zhu Z, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerging Microbes & Infections.* 2020;9(1):221-236.

31. Eden J-S, Rockett R, Carter I, et al. An emergent clade of SARS-CoV-2 linked to returned travellers from Iran. *bioRxiv.*2020.

32. Sardar R, Satish D, Birla S, Gupta D. Comparative analyses of SAR-CoV2 genomes from different geographical locations and other coronavirus family genomes reveals unique features potentially consequential to host-virus interaction and pathogenesis.*bioRxiv.* 2020:2020.2003.2021.001586.

33. Xiao M, Liu X, Ji J, et al. Multiple approaches for massively parallel sequencing of HCoV-19 genomes directly from clinical samples.*bioRxiv.* 2020.

Tables and figures
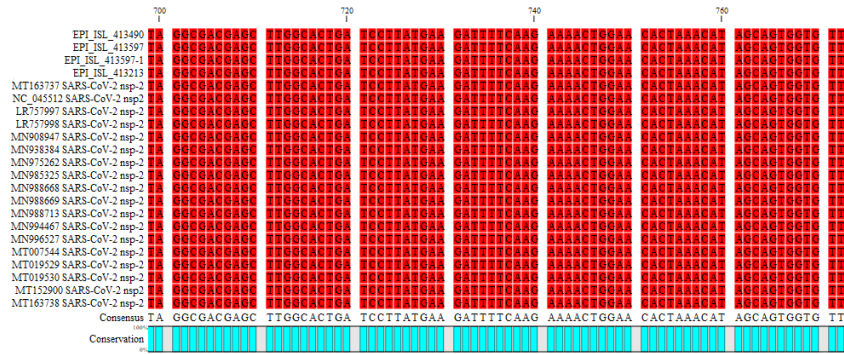
Figure 1. MSA investigation for SNPs in nsp-2 did not show any polymorphisms, the red nucleotides show conservation between this study isolates and prototype or other countries isolated sequences

MN988713.1 2019-nCoV/USA-IL1/2020
MT123293.1 SARS-CoV-2/IQTC03/human/2020/CHN
MN997409.1 2019-nCoV/USA-AZ1/2020
MN996530.1 WIV06
LR757996.1 Wuhan seafood market pneumonia virus
MT019530.1 BetaCoV/Wuhan/IPBCAMS-WH-02/2019
MT020781.1 nCoV-FIN-29-Jan-2020
LC522974.1 2019-nCoV/Japan/TY/WK-501/2020
MT039887.1 2019-nCoV/USA-WI1/2020
MT066175.1 SARS-CoV-2/NTU01/2020/TWN
MT093631.1 SARS-CoV-2/WH-09/human/2020/CHN
MT118835.1 2019-nCoV/USA-CA9/2020
LC528232.1 SARS-CoV-2/Hu/DP/Kng/19-020
MT123292.1 SARS-CoV-2/IQTC04/human/2020/CHN
MT106054.1 2019-nCoV/USA-TX1/2020
MT093571.1 SARS-CoV-2/01/human/2020/SWE
MT049951.1 SARS-CoV-2/Yunnan-01/human/2020/CHN
MT039890.1 SNU01
LC522973.1 2019-nCoV/Japan/TY/WK-012/2020
MT019533.1 BetaCoV/Wuhan/IPBCAMS-WH-05/2020
MT019529.1 BetaCoV/Wuhan/IPBCAMS-WH-01/2019
LR757995.1 Wuhan seafood market pneumonia virus
MN996529.1 WIV05
MN994468.1 2019-nCoV/USA-CA2/2020
MN985325.1 2019-nCoV/USA-WA1/2020
MT027062.1 2019-nCoV/USA-CA3/2020
LC528233.1 SARS-CoV-2/Hu/DP/Kng/19-027
99 MT123290.1 SARS-CoV-2/IQTC01/human/2020/CHN
MT106052.1 2019-nCoV/USA-CA7/2020
MT066176.1 SARS-CoV-2/NTU02/2020/TWN
MT039888.1 2019-nCoV/USA-MA1/2020
LC522975.1 2019-nCoV/Japan/TY/WK-521/2020
MT027064.1 2019-nCoV/USA-CA5/2020
MT019531.1 BetaCoV/Wuhan/IPBCAMS-WH-03/2019
LR757997.1 Wuhan seafood market pneumonia virus
MN996531.1 WIV07
MN996527.1 WIV02
MN988668.1 2019-nCoV WHU01
MN938384.1 2019-nCoV_HKU-SZ-002a_2020
MT044258.1 2019-nCoV/USA-CA6/2020
MT123291.1 SARS-CoV-2/IQTC02/human/2020/CHN
MT106053.1 2019-nCoV/USA-CA8/2020
MT072688.1 SARS0CoV-2/61-TW/human/2020/NPL
MT044257.1 2019-nCoV/USA-IL2/2020
MT039873.1 HZ-1
LC522972.1 2019-nCoV/Japan/KY/V-029/2020
MT019532.1 BetaCoV/Wuhan/IPBCAMS-WH-04/2019
LR757998.1 Wuhan seafood market pneumonia virus
MT007544.1 Australia/VIC01/2020
MN996528.1 WIV04
MN994467.1 2019-nCoV/USA-CA1/2020
MN975262.1 2019-nCoV_HKU-SZ-005b_2020
▲ NC_045512.2 Wuhan-Hu-1
▲ MT152900.1
▲ MT163737.1
▲ MT163738.1
83
MN996532.1 Bat coronavirus RaTG13
MG772934.1 Bat SARS-like coronavirus isolate bat-SL-CoVZXC21
99 MG772933.1 Bat SARS-like coronavirus isolate bat-SL-CoVZC45
MT084071.1 Pangolin coronavirus isolate MP789

0.01

Figure 2. Phylogenetic tree for 322 nucleotides of the nsp-2 gene of the SARS-CoV-2 using neighbor joining method and 1000 bootstrap, the red triangle shows the reference sequence for the SARS-CoV-2 and the blue triangle indicates Iranian isolates.

8

MN938385.1 2019-nCoV_HKU-SZ-001_2020
MT072668.1 SARS-CoV-2/GHB-03021/human/2020/BEL
▲ NC_045512.2 Wuhan-Hu-1
MN975262.1 2019-nCoV_HKU-SZ-005b_2020
MN988713.1 2019-nCoV/USA-IL1/2020
MN994467.1 2019-nCoV/USA-CA1/2020
MN997409.1 2019-nCoV/USA-AZ1/2020
MN996528.1 WIV04
MN996530.1 WIV06
MT007544.1 Australia/VIC01/2020
LR757996.1 Wuhan seafood market pneumonia virus
MT019529.1 BetaCoV/Wuhan/IPBCAMS-WH-01/2019
MT019531.1 BetaCoV/Wuhan/IPBCAMS-WH-03/2019
MT019533.1 BetaCoV/Wuhan/IPBCAMS-WH-05/2020
MT027062.1 2019-nCoV/USA-CA3/2020
LC522972.1 2019-nCoV/Japan/KY/V-029/2020
LC522974.1 2019-nCoV/Japan/TY/WK-501/2020
MT039873.1 HZ-1
MT039887.1 2019-nCoV/USA-WI1/2020
MT044257.1 2019-nCoV/USA-IL2/2020
MT049951.1 SARS-CoV-2/Yunnan-01/human/2020/CHN
MT066176.1 SARS-CoV-2/NTU02/2020/TWN
MT093571.1 SARS-CoV-2/01/human/2020/SWE
MT106052.1 2019-nCoV/USA-CA7/2020
MT106054.1 2019-nCoV/USA-TX1/2020
MT123290.1 SARS-CoV-2/IQTC01/human/2020/CHN
MT123292.1 SARS-CoV-2/IQTC04/human/2020/CHN
99   LC528232.1 SARS-CoV-2/Hu/DP/Kng/19-020
LC528233.1 SARS-CoV-2/Hu/DP/Kng/19-027
MT123293.1 SARS-CoV-2/IQTC03/human/2020/CHN
MT123291.1 SARS-CoV-2/IQTC02/human/2020/CHN
MT118835.1 2019-nCoV/USA-CA9/2020
MT106053.1 2019-nCoV/USA-CA8/2020
MT093631.1 SARS-CoV-2/WH-09/human/2020/CHN
MT072688.1 SARS0CoV-2/61-TW/human/2020/NPL
MT066175.1 SARS-CoV-2/NTU01/2020/TWN
MT044258.1 2019-nCoV/USA-CA6/2020
MT039888.1 2019-nCoV/USA-MA1/2020
MT039890.1 SNU01
LC522975.1 2019-nCoV/Japan/TY/WK-521/2020
LC522973.1 2019-nCoV/Japan/TY/WK-012/2020
MT027064.1 2019-nCoV/USA-CA5/2020
MT020781.1 nCoV-FIN-29-Jan-2020
MT019532.1 BetaCoV/Wuhan/IPBCAMS-WH-04/2019
MT019530.1 BetaCoV/Wuhan/IPBCAMS-WH-02/2019
LR757998.1 Wuhan seafood market pneumonia virus
LR757995.1 Wuhan seafood market pneumonia virus
MN996531.1 WIV07
MN996529.1 WIV05
MN996527.1 WIV02
MN994468.1 2019-nCoV/USA-CA2/2020
MN988668.1 2019-nCoV WHU01
MN985325.1 2019-nCoV/USA-WA1/2020
MN938384.1 2019-nCoV_HKU-SZ-002a_2020
▲ MT232869
▲ MT232870
99
MN996532.1 Bat coronavirus RaTG13
JX993988.1 Bat coronavirus Cp/Yunnan2011
KJ473815.1 BtRs-BetaCoV/GX2013
KY417145.1 Bat SARS-like coronavirus isolate Rf4092
KF569996.1 Rhinolophus affinis coronavirus isolate LYRa11
KF294439.1 SARS-related bat coronavirus isolate Anlong-29
KF294455.1 SARS-related bat coronavirus isolate Anlong-111
KF294440.1 SARS-related bat coronavirus isolate Anlong-97
KY417146.1 Bat SARS-like coronavirus isolate Rs4231
KY770859.1 Bat coronavirus isolate Anlong-112
KY770858.1 Bat coronavirus isolate Anlong-103
KT444582.1 SARS-like coronavirus WIV16
KY417150.1 Bat SARS-like coronavirus isolate Rs4874
KY417143.1 Bat SARS-like coronavirus isolate Rs4081
MK211377.1 BtRs-BetaCoV/YN2018C
MK211378.1 BtRs-BetaCoV/YN2018D
89
KY417142.1 Bat SARS-like coronavirus isolate As6526
KJ473816.1 BtRs-BetaCoV/YN2013
KY417151.1 Bat SARS-like coronavirus isolate Rs7327
KY417152.1 Bat SARS-like coronavirus isolate Rs9401
MK211376.1 BtRs-BetaCoV/YN2018B
KY417149.1 Bat SARS-like coronavirus isolate Rs4255
KF367457.1 Bat SARS-like coronavirus WIV1
KC881006.1 Bat SARS-like coronavirus Rs3367
KC881005.1 Bat SARS-like coronavirus RsSHC014
KY417144.1 Bat SARS-like coronavirus isolate Rs4084
KY417147.1 Bat SARS-like coronavirus isolate Rs4237
KY417148.1 Bat SARS-like coronavirus isolate Rs4247
MK211375.1 BtRs-BetaCoV/YN2018A
DQ071615.1 Bat SARS coronavirus Rp3
AY559097.1 SARS coronavirus Sin3408L
MK062184.1 SARS coronavirus Urbani isolate icSARS-C7-MA
MK062183.1 SARS coronavirus Urbani isolate icSARS-C7
MK062182.1 SARS coronavirus Urbani isolate icSARS-C3-MA
MK062181.1 SARS coronavirus Urbani isolate icSARS-C3
MK062180.1 SARS coronavirus Urbani isolate icSARS-MA
MK062179.1 SARS coronavirus Urbani isolate icSARS
AH013709.2 SARS coronavirus Sin_WNV
AH012999.2 SARS coronavirus Shanghai LY polyprotein
KP886809.1 Bat SARS-like coronavirus YNLF_34C
KP886808.1 Bat SARS-like coronavirus YNLF_31C
KF514407.1 SARS/VeroE6_lab/USA/ExoN1_c5.7P20/2010
JX163928.1 SARS coronavirus isolate Tor2/FP1-10895
JX163927.1 SARS coronavirus isolate Tor2/FP1-10851
JX163926.1 SARS coronavirus isolate Tor2/FP1-10912
JX163925.1 SARS coronavirus isolate Tor2/FP1-10895

0.005

Figure 3. Phylogenetic tree for the 338 nucleotides of RdRp gene of the SARS-CoV-2 using neighbor joining method and 1000 bootstrap, the red triangle shows the reference sequence for the SARS-CoV-2 and the blue triangle indicates Iranian isolates.

LC528233.1 SARS-CoV-2/Hu/DP/Kng/19-027
LC528232.1SARS-CoV-2/Hu/DP/Kng/19-020
MT123293.1 SARS-CoV-2/IQTC03/human/2020/CHN
MT123292.1 SARS-CoV-2/IQTC04/human/2020/CHN
MT123291.1 SARS-CoV-2/IQTC02/human/2020/CHN
MT123290.1 SARS-CoV-2/IQTC01/human/2020/CHN
MT118835.1 2019-nCoV/USA-CA9/2020
MT106054.1 2019-nCoV/USA-TX1/2020
MT106053.1 2019-nCoV/USA-CA8/2020
MT106052.1 2019-nCoV/USA-CA7/2020
MT093631.1 SARS-CoV-2/WH-09/human/2020/CHN
MT093571.1 SARS-CoV-2/01/human/2020/SWE
MT072688.1 SARS0CoV-2/61-TW/human/2020/NPL
MT066176.1 SARS-CoV-2/NTU02/2020/TWN
MT066175.1 SARS-CoV-2/NTU01/2020/TWN
MT049951.1 SARS-CoV-2/Yunnan-01/human/2020/CHN
MT044258.1 2019-nCoV/USA-CA6/2020
MT044257.1 2019-nCoV/USA-IL2/2020
MT039888.1 2019-nCoV/USA-MA1/2020
MT039887.1 2019-nCoV/USA-WI1/2020
MT039890.1 SNU01
MT039873.1 HZ-1
LC522975.1 2019-nCoV/Japan/TY/WK-521/2020
LC522974.1 2019-nCoV/Japan/TY/WK-501/2020
LC522973.1 2019-nCoV/Japan/TY/WK-012/2020
LC522972.1 2019-nCoV/Japan/KY/V-029/2020
MT027064.1 2019-nCoV/USA-CA5/2020
MT027062.1 2019-nCoV/USA-CA3/2020
MT020781.1 nCoV-FIN-29-Jan-2020
MT019533.1 BetaCoV/Wuhan/IPBCAMS-WH-05/2020
MT019532.1 BetaCoV/Wuhan/IPBCAMS-WH-04/2019
MT019531.1 BetaCoV/Wuhan/IPBCAMS-WH-03/2019
MT019530.1 BetaCoV/Wuhan/IPBCAMS-WH-02/2019
MT019529.1 BetaCoV/Wuhan/IPBCAMS-WH-01/2019
LR757998.1 Wuhan seafood market pneumonia virus
LR757996.1 Wuhan seafood market pneumonia virus
LR757995.1 Wuhan seafood market pneumonia virus
MT007544.1 Australia/VIC01/2020
MN996531.1 WIV07
MN996530.1 WIV06
MN996529.1 WIV05
MN996528.1 WIV04
MN996527.1 WIV02
MN997409.1 2019-nCoV/USA-AZ1/2020
MN994468.1 2019-nCoV/USA-CA2/2020
MN994467.1 2019-nCoV/USA-CA1/2020
MN988668.1 2019-nCoV WHU01
MN988713.1 2019-nCoV/USA-IL1/2020
MN985325.1 2019-nCoV/USA-WA1/2020
MN975262.1 2019-nCoV_HKU-SZ-005b_2020
MN938384.1 2019-nCoV_HKU-SZ-002a_2020
NC_045512.2 Wuhan-Hu-1
MT232871
MT232872
MN938387.1 2019-nCoV_HKU-SZ-001_2020
MN996532.1 Bat coronavirus RaTG13
LC469301.1 Bat coronavirus Rc-CoV-3

80

0.02

Figure 4. Phylogenetic tree for 157 nucleotides of the S gene of the SARS-CoV-2 using neighbor joining method and 1000 bootstrap, the red triangle shows the reference sequence for the SARS-CoV-2 and the blue triangle indicates Iranian isolates.

10

LC528233.1 SARS-CoV-2/Hu/DP/Kng/19-027
LC528232.1 SARS-CoV-2/Hu/DP/Kng/19-020
MT123293.1 SARS-CoV-2/IQTC03/human/2020/CHN
MT123292.1 SARS-CoV-2/IQTC04/human/2020/CHN
MT123291.1 SARS-CoV-2/IQTC02/human/2020/CHN
MT123290.1 SARS-CoV-2/IQTC01/human/2020/CHN
MT118835.1 2019-nCoV/USA-CA9/2020
MT106054.1 2019-nCoV/USA-TX1/2020
MT106053.1 2019-nCoV/USA-CA8/2020
MT106052.1 2019-nCoV/USA-CA7/2020
MT093631.1 SARS-CoV-2/WH-09/human/2020/CHN
MT093571.1 SARS-CoV-2/01/human/2020/SWE
MT072688.1 SARS0CoV-2/61-TW/human/2020/NPL
MT066176.1 SARS-CoV-2/NTU02/2020/TWN
MT066175.1 SARS-CoV-2/NTU01/2020/TWN
MT049951.1 SARS-CoV-2/Yunnan-01/human/2020/CHN
MT044258.1 2019-nCoV/USA-CA6/2020
MT044257.1 2019-nCoV/USA-IL2/2020
MT039888.1 2019-nCoV/USA-MA1/2020
MT039887.1 2019-nCoV/USA-WI1/2020
MT039890.1 SNU01
MT039873.1 HZ-1
LC522975.1 2019-nCoV/Japan/TY/WK-521/2020
LC522974.1 2019-nCoV/Japan/TY/WK-501/2020
LC522973.1 2019-nCoV/Japan/TY/WK-012/2020
LC522972.1 2019-nCoV/Japan/KY/V-029/2020
MT027064.1 2019-nCoV/USA-CA5/2020
MT027062.1 2019-nCoV/USA-CA3/2020
MT020781.1 nCoV-FIN-29-Jan-2020
MT019533.1 BetaCoV/Wuhan/IPBCAMS-WH-05/2020
MT019532.1 BetaCoV/Wuhan/IPBCAMS-WH-04/2019
MT019531.1 BetaCoV/Wuhan/IPBCAMS-WH-03/2019
MT019530.1 BetaCoV/Wuhan/IPBCAMS-WH-02/2019
MT019529.1 BetaCoV/Wuhan/IPBCAMS-WH-01/2019
LR757998.1 Wuhan seafood market pneumonia virus
LR757996.1 Wuhan seafood market pneumonia virus
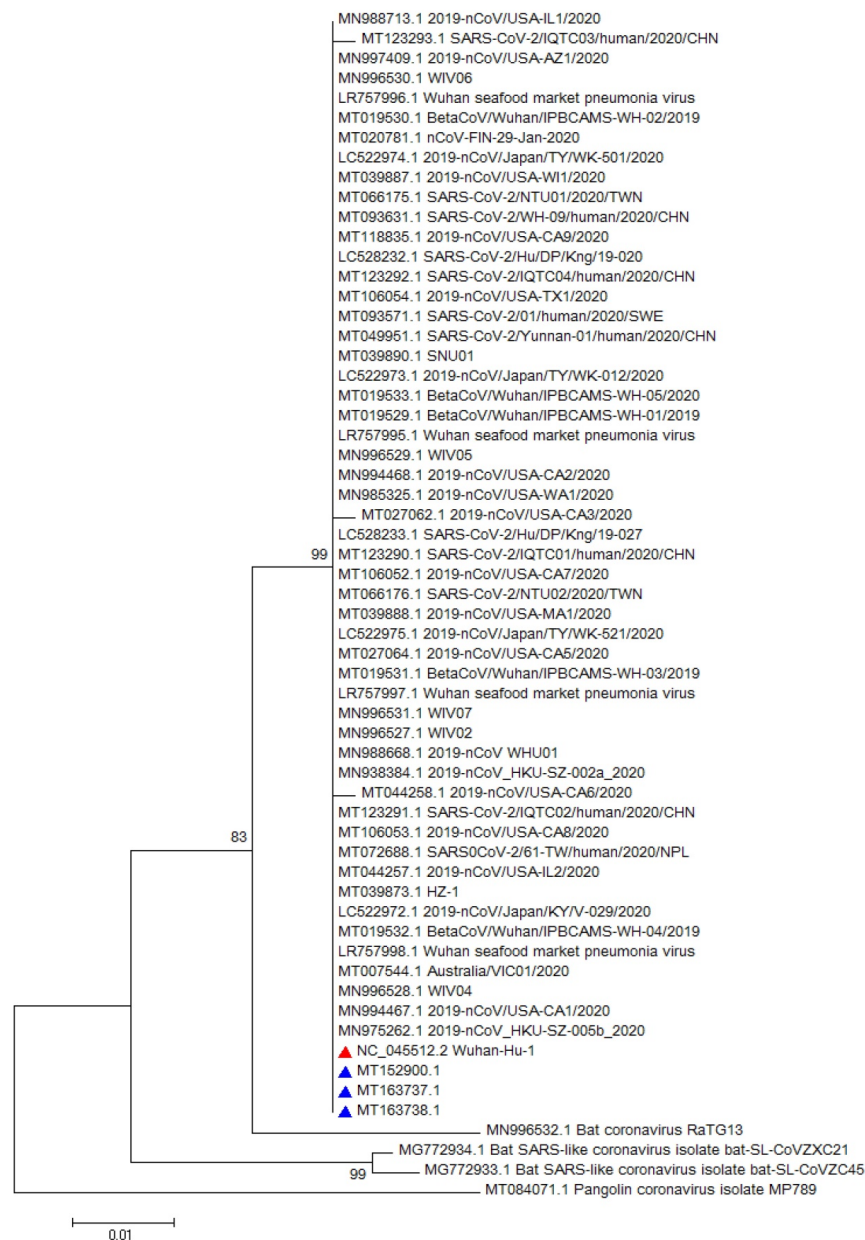LR757995.1 Wuhan seafood market pneumonia virus
MT007544.1 Australia/VIC01/2020
MN996531.1 WIV07
MN996530.1 WIV06
MN996529.1 WIV05
MN996528.1 WIV04
MN996527.1 WIV02
MN997409.1 2019-nCoV/USA-AZ1/2020
MN994468.1 2019-nCoV/USA-CA2/2020
MN994467.1 2019-nCoV/USA-CA1/2020
MN988668.1 2019-nCoV WHU01
MN988713.1 2019-nCoV/USA-IL1/2020
MN985325.1 2019-nCoV/USA-WA1/2020
MN975262.1 2019-nCoV_HKU-SZ-005b_2020
MN938384.1 2019-nCoV_HKU-SZ-002a_2020
NC_045512.2 Wuhan-Hu-1
MN938385.1 2019-nCoV_HKU-SZ-001_2020
MT232869
MT232870
MT072668.1 SARS-CoV-2/GHB-03021/human/2020/BEL

98

99

MN996532.1 Bat coronavirus RaTG13
JX993988.1 Bat coronavirus Cp/Yunnan2011
KJ473815.1 BtRs-BetaCoV/GX2013
KF294439.1 SARS-related bat coronavirus isolate Anlong-29
KF294455.1 SARS-related bat coronavirus isolate Anlong-111
KF294440.1 SARS-related bat coronavirus isolate Anlong-97
KY417145.1 Bat SARS-like coronavirus isolate Rf4092
KF569996.1 Rhinolophus affinis coronavirus isolate LYRa11
KT444582.1 SARS-like coronavirus WIV16
KY417150.1 Bat SARS-like coronavirus isolate Rs4874
KY417143.1 Bat SARS-like coronavirus isolate Rs4081
MK211377.1 BtRs-BetaCoV/YN2018C
MK211378.1 BtRs-BetaCoV/YN2018D
KY417146.1 Bat SARS-like coronavirus isolate Rs4231
KY770859.1 Bat coronavirus isolate Anlong-112
KY770858.1 Bat coronavirus isolate Anlong-103
KF367457.1 Bat SARS-like coronavirus WIV1
KC881006.1 Bat SARS-like coronavirus Rs3367
KY417149.1 Bat SARS-like coronavirus isolate Rs4255
KC881005.1 Bat SARS-like coronavirus RsSHC014
KY417144.1 Bat SARS-like coronavirus isolate Rs4084
KY417147.1 Bat SARS-like coronavirus isolate Rs4237
KY417148.1 Bat SARS-like coronavirus isolate Rs4247
MK211375.1 BtRs-BetaCoV/YN2018A
MK211376.1 BtRs-BetaCoV/YN2018B
KY417152.1 Bat SARS-like coronavirus isolate Rs9401
KY417151.1 Bat SARS-like coronavirus isolate Rs7327
KY417142.1 Bat SARS-like coronavirus isolate As6526
KJ473816.1 BtRs-BetaCoV/YN2013
DQ071615.1 Bat SARS coronavirus Rp3
KY559097.1 SARS coronavirus Sin3408L
MK062184.1 SARS coronavirus Urbani isolate icSARS-C7-MA
MK062183.1 SARS coronavirus Urbani isolate icSARS-C7
MK062182.1 SARS coronavirus Urbani isolate icSARS-C3-MA
MK062181.1 SARS coronavirus Urbani isolate icSARS-C3
MK062180.1 SARS coronavirus Urbani isolate icSARS-MA
MK062179.1 SARS coronavirus Urbani isolate icSARS
AH013709.2 SARS coronavirus Sin_WNV
AH012999.2 SARS coronavirus Shanhgai LY polyprotein
KP886809.1 Bat SARS-like coronavirus YNLF_34C
KP886808.1 Bat SARS-like coronavirus YNLF_31C

88

11

Figure 5. Phylogenetic tree for 338 nucleotide of the RdRp gene of the SARS-CoV-2 using maximum likelihood method and 1000 bootstrap, the red triangle shows the reference sequence for the SARS-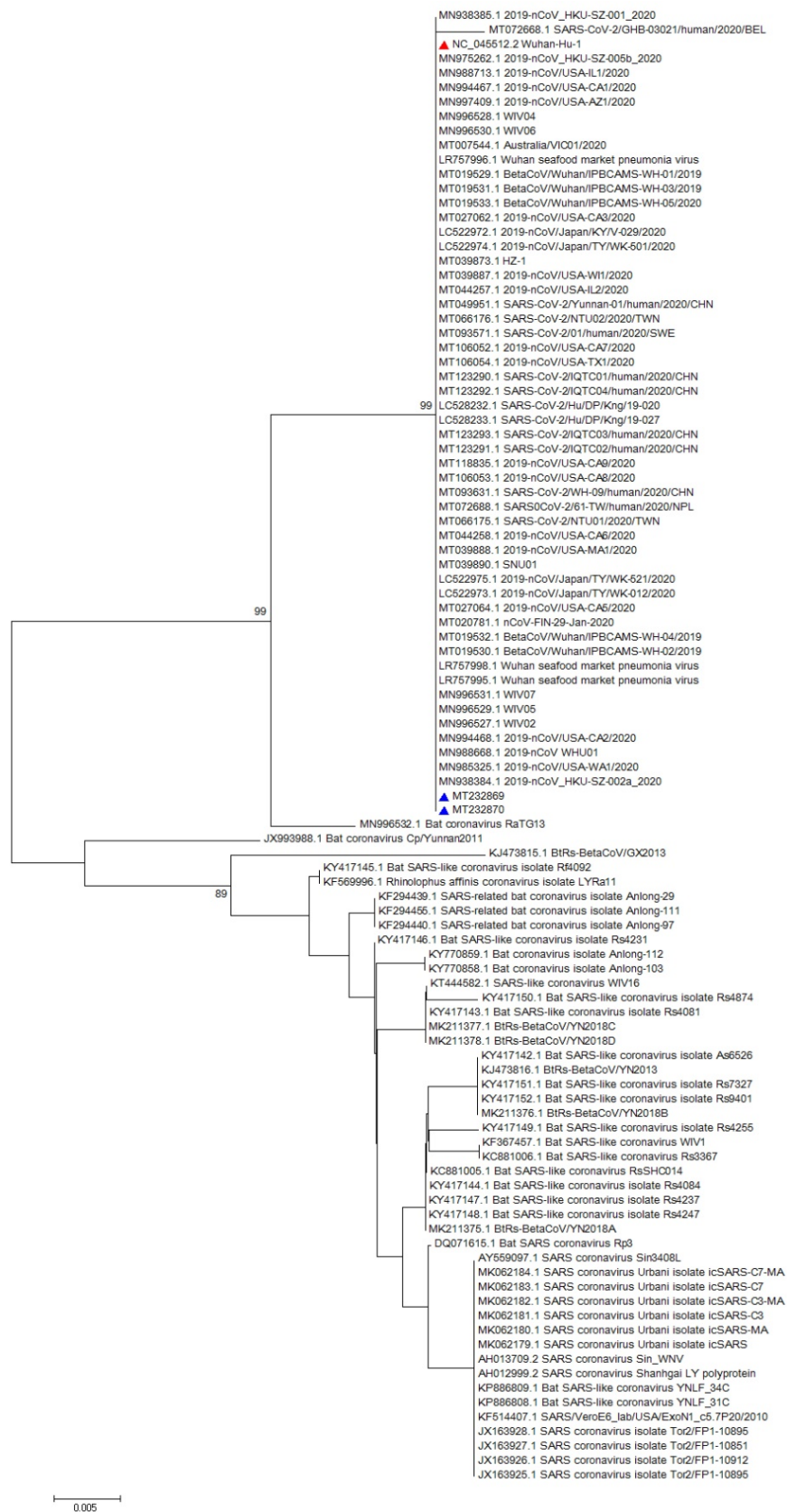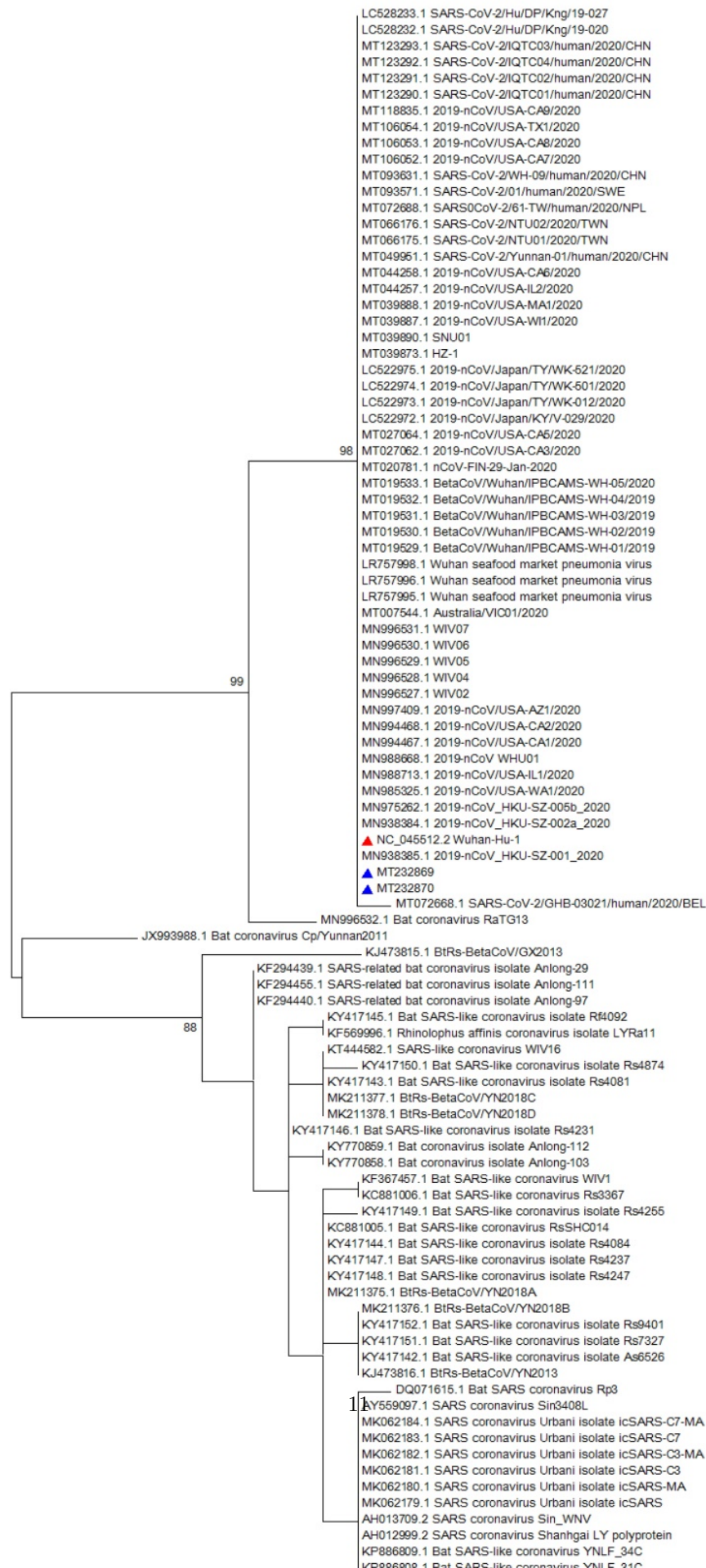CoV-2 and the blue triangle indicates Iranian isolates.