The Iraqi Journal of Veterinary Medicine, 44 (2): 56–70 **2020**

The evolution of Genetic Molecular Map and phylogenetic tree of Coronavirus (COVID-19): Review Article

Hayder A H AL-Mutar

Department of Surgery and Obstetrics, College of Veterinary Medicine, University of Baghdad, Iraq

ABSTRACT

Coronavirus (COVID-19) is suspected to originate from an animal host (zoonotic) followed by a human to human transmission, The purpose of the present study is to determine the genetic affinity of a phylogenetic tree and conformation of protein between human and bat. Through study of genetic sequencing, as shown in the tree design of strains and genetic variants, the main cause of COVID-19 is the Bat Coronavirus RaTG13 (SARS) virus in ID: MN996532.1 and ID: MG772933.1 (24-Jul-2013 and Feb-2017) shows that the evolution of the Corona virus from 2003, 2013 to 2020 which has become the most deadly peak in humans, the virus evolved from the bat effect on Humans. Protein analysis show 98 change of amino acid form RaTG13 (SARS) virus to COVID-19 Homo sapiens. Consequently, this study increased our understanding of the genetic variety of the COVID-19 carried by bats. So we conclude that a protein conformation drawing shows high identity compatibility between a bat and a human.

Keywords: SARS; COVID-19; molecular evolution; *Rhinolophus affinis*.

Introduction

The first cases of SARS were identified in 2002 in China (Guangdong Province). While in 2003 an epidemic has spread across the world that affected more than three thousand people ensuing in more than a hundred deaths (**1**). SARS-CoV is an element of the Corona viridae family of enveloped positive-Stranded RNA viruses which have a great host range. Some corona virus infect the humans, Domestic birds and ruminant might cause respiratory illness,whereas the other corona virus infected in rodents, feline, sow, and bovine lead to enteric disease. The 27–32 kb genomes of corona viruses are the largest of RNA viruses encode 23 putative proteins, including four major structural proteins with N:nucleocapsid, S:spike, M: membrane and small E:envelope.

The spike protein of the glycoprotein projection on the viral surface is critical for viral attachment and entry into the host cells. In addition, variations of S protein among strains of corona virus are responsible for host range sensitivity and tissue tropism activity (**2**). An ongoing outbreak of pneumonia related with a novel corona virus that called a severe acute respiratory syndrome corona virus 2 (SARS-CoV-2) in the persons was report in china-Wuhan, Hubei province, China in December-2019 (**3**). In the following weeks, the infections spread all countries around the world (**4-5**).

On February 2020, world health organization named the disease caused by the new corona virus as COVID-19 (**6**). Corona virus is an envelope and single-stranded ribonucleic acid named for its solar corona like appearance due to 9-12 nm-long surface spikes (**7**). There are four structural proteins encoded important to envelop the corona viruses genome, one of which is the S protein that binds to angiotensin-converting enzyme2 receptor and mediates later merger between the envelope and host plasma membranes to help viral in gress into the host cell (**8**). The aim of the present study was to determine the genetic affinity of a phylogenetic tree and conformation of protein between human and bat.

---

*Correspondence: almutar.haydar@gmail.com; al_mutar.haydar @covm.uobaghdad.edu.iq
Department of Surgery and Obstetrics, College of Veterinary Medicine, University of Baghdad, Baghdad, Iraq.
Received: 16 May 2020, Accepted: 16 June 2020, Published: 28 December 2020.

Materials and Methods

In this study data were collected from a date March\2003 to Feb\2020 were genome sequences have been submitted to GenBank. For all sequences, the accession numbers are listed in Supplementary Table (1;2;3). The sequence data have been deposited into the NCBI sequence reads archive below the accession number. Using BLAST-NCBI, Routine sequence alignments were performed (9). MEGA6.0 (Phoenix, AZ, USA) was used to align the nucleotide (nt) and amino acid (aa) sequences using phylogenetic analyses.

Results and Discussion

Phylogenetic tree structuring

The phylogenetic tree diagrammatic was done by molecular evolutionary genetics analysis (MEGA) software version 6.0. Sequences (**10**).

Construction of the Neighbor-joining tree showed that the identity (80% - 100%), through comparison between isolated Rhinolophus sinicus and Homo sapiens . hierarchical cluster analysis determine the following clusters: large Clusters divided into several root: first root the SARS coronavirus Rhinolophus sinicus in Homo sapiens from 2003-2016 in China, Hong Kong, Singapore, USA, Canada, Italy and Taiwan showed compatibility and the highest identity at >99%. Second root COVID-19 in Homo sapiens from 2019-2020 in China, Taiwan, USA, Australia, Japan, South Korea, Finland and Nepal showed compatibility and the highest identity at >99%. Third root SARS coronavirus in Rhinolophus affinis (2013; 2017 in China ID: MN996532.1, MG772933.1) showed compatibility with COVID-19 in humans and the highest identity at >97% (Figure 1)

In this study, the investigation of the patterns of molecular divergence between SARS-CoV-2 and other related COVID-19. Although the genomic analyses recommended that SARS-CoV-2 was closest to RaTG13, the virus was evolved from the bat might have an effect on the humans (**11**). It may be that the virus is transmitted to humans from mice through an experiment that conducted the cultivation of SARS virus in mice which led to the development of the virus due to the relationship between mice and humans (**12**). structural, protein (M: membrane, N:nucleocapsid, S:spike and E:envelop and several non-structural) Because of interacting and eating rats in China, this leads to transmission to humans and development (**13**).

The sequencing of amplified product of Amino acid polymorphism of Coronavirus

The result of the sequence analysis was analyzed by blast in the National Center Biotechnology Information (NCBI) online and BioEdit program (**14**) to detect polymorphism after alignment of product amplification in Corona virus, showed that 98 change of amino acid form RaTG13 (SARS) virus *Rhinolophus affinis* to COVID-19 Homo sapiens (Table 4 and Figure 2). The analysis of molecular development and population genetics show some protein changes that might be favored by natural selection during the evolution of SARS-CoV-2 and other related viruses. or f1ab, which encodes replicate/ transcriptase is required for viral genome replication and might also be significant for viral pathogenesis (**15**). COVID-19 is spherical or polymorphous laminated particles that contain the associated mono-stranded RNA (positive sensation) with a nuclear protein inside a capsule consisting of a matrix protein. The envelope contains club-shaped glycoprotein expectations. Some corona viruses also have hem Agglutinin-esterase protein (HE) (**16**, **17**) (Figure 3).

The result showed the similarity of the conformation of protein for COVID-19 shown in figure 4, between the Bat Coronavirus and human corona virus. Identification of other COV in bat species (*R. sinicus, R. ferrumequinum* (**18**,**19**).

Human and animal relations may pose an exacting public health threat in rural communities where frequent contact with animals occurs and where infection prevention measures are less developed (**20**). Bats (order Chiroptera) are reservoir of a great number of zoonotic viruses, including COVID-19 that have caused disease outbreak in human and livestock populations (**21–22**),

Severe Acute Respiratory Syndrome COVID-19, are the causative agent of the SARS outbreak affecting world countries (**23**, **24**).

Mechanism of action for COVID-19

The genome of coronavirus is comprised of three thousands nucleotides. It encodes four proteins of (Figure 3) (**25**, **26**). **Conformation of protein for COVID-19** The COVID-19 replicate gene-encoded two poly proteins, poly protein 1a with MW 450 KD and poly protein 1ab with MW 750 KD, these poly proteins are required for viral replication and transcription (**27**, **28**). Several crystal structure of COVID-19 (Figure 4) in complex with different inhibitor are deposited in the protein by using X-ray diffraction technique at a resolution between a human and a bat (**29**, **30**).

Table 1. Neighbor-joining tree of Coronavirus from animals

Accession	host	isolation source	collection date	Country	Compatibility
ID: AY304488.1 ID: AY304486.1	Animals\ Himalayan palm civets	nasal turbinate	MAY-2003	Hong Kong	14871/18249(81%)
ID: JQ316196.1	Animals\rat	Vero cells	Apr-2003	United Kingdom	14866/18249(81%)
ID: JX163928.1 ID: JX163924.1	Animals\ ferrets	nasal turbinate	Feb-2010	USA	14866/18249(81%)

Table 2. Neighbor-joining tree of Coronavirus from bats

Accession	host	isolation source	collection date	Country	Compatibility
ID: KY417146.1	*Rhinolophus sinicus*	fecal swab	Apr-2013	China	14879/18255(82%)
ID: MN996532.1	*Rhinolophus affinis*	fecal swab	Jul-2013	China	20546/21290(97%)
ID: KJ473816.1	*Rhinolophus sinicus*	fecal swab	2013	China	14872/18254(81%
ID: MG772934.1	*Rhinolophus sinicus*	fecal swab	Jul-2015	China	16148/18231(89%)
ID: MK211376.1	*Rhinolophus affinis*	fecal swab	Sep-2016	China	14875/18254(81%)
ID: MG772933.1	*Rhinolophus sinicus*	fecal swab	Feb-2017	China	18972/21300(89%)

Table 3. Neighbor-joining tree of Coronavirus from human

Accession	host	isolation source	collection date	Country	Compatibility
ID: AY427439.1	Homo sapiens	sputum	March 2003	Italy	14867/18249(81%)
ID: AY278554.2	Homo sapiens	sputum	APR-2003	China: Hong Kong	14869/18249(81%)
ID: AY283796.1	Homo sapiens	sputum	APR-2003	Singapore	14867/18249(81%)
ID: AY274119.3	Homo sapiens	sputum	APR-2003	Canada: Toronto	14867/18249(81%
ID: AY291451.1 ID: AY502928.1 ID: AY502923.1	Homo sapiens	sputum	MAY-2003	Taiwan	14867/18249(81%)
ID: AY323977.2	Homo sapiens	sputum	JUL-2003	Italy	14867/18249(81%)
ID: AY390556.1	Homo sapiens	sputum	SEP-2003	China: Guangzhou	14870/18249(81%)
ID: AY395003.1 ID: AY394996.1	Homo sapiens	sputum	SEP-2003	China	14873/18249(82%)
ID: AY559083.1	Homo sapiens	sputum	2003	Singapore	14868/18249(81%)
ID: AY559096.1 ID: AY559095.1 ID: AY559086.1 ID: AY559085.1	Homo sapiens	sputum	FEB-2004	Singapore	14867/18249(81%)
ID: AY714217.1	Homo sapiens	sputum	AUG-2004	USA	14867/18249(81%)
ID: FJ882963.1	Homo sapiens	sputum	Aug-2004	USA	14867/18249(81%)
ID: DQ898174.1	Homo sapiens	sputum	AUG-2006	Canada	14867/18249(81%)

ID: JX163927.1 ID: JX163926.1 ID: JX163923.1	Homo sapiens	nasal turbinate	Feb-2010	USA	14867/18249(81%)
ID: JX163925.1	Homo sapiens	lung	Feb-2010	USA	14866/18249(81%)
ID: JX163928.1, ID: JX163924.1	Homo sapiens	lung	Feb-2010	USA	14866/18249(81%)
ID: MT019529.1	Homo sapiens	bronchoalveolar lavage fluid	Dec-2019	China: Hubei, Wuhan	21287/21290(99%)
ID: MT019532.1	Homo sapiens	bronchoalveolar lavage fluid	Dec-2019	China: Hubei, Wuhan	21290/21290(100%)
ID: MT019530.1	Homo sapiens	bronchoalveolar lavage fluid	Dec-2019	China: Hubei, Wuhan	21290/21290(100%)
ID: MN996530.1	Homo sapiens	bronchoalveolar lavage fluid	Dec-2019	China: Wuhan	21290/21290(100%)
ID: MN996528.1	Homo sapiens	bronchoalveolar lavage fluid	Dec-2019	China: Wuhan	21290/21290(100%)
ID: NC_045512.2	Homo sapiens	bronchoalveolar lavage fluid	Dec-2019	China	21290/21290(100%)
ID: MT019531.1	Homo sapiens	bronchoalveolar lavage fluid	Dec-2019	China: Hubei, Wuhan	21289/21290(99%)
ID: MN996527.1	Homo sapiens	sputum	Dec-2019	China: Wuhan	21289/21290(99%)
ID: MN996531.1 ID: MN996529.1	Homo sapiens	sputum	Dec-2019	China: Wuhan	21289/21290(99%)
ID: LR757996.1	Homo sapiens	sputum	2019/2020	Wuhan	21290/21290(100%)
ID: LR757995.1	Homo sapiens	sputum	2019/2020	Wuhan	21288/21290(99%)
ID: LC522975.1	Homo sapiens	sputum	2020-01	Japan	21288/21290(99%)
ID: LC522974.1	Homo sapiens	sputum	2020-01	Japan	21288/21290(99%)
ID:LC522972.1 ID: LC522973.1	Homo sapiens	sputum	2020-01	Japan	21288/21290(99%)
ID: MT039873.1	Homo sapiens	sputum	Jan-2020	China: Hangzhou	21290/21290(100%)
ID: MT019533.1	Homo sapiens	sputum	Jan-2020	China: Hubei, Wuhan	21289/21290(99%)
ID: MN988668.1	Homo sapiens	sputum	Jan-2020	China	21290/21290(100%)
ID: MT027064.1	Homo sapiens	sputum	Jan-2020	USA: CA	21289/21290(99%)
ID: MT066175.1	Homo sapiens	sputum	Jan-2020	Taiwan	21289/21290(99%)
ID: MT007544.1	Homo sapiens	sputum	Jan-2020	Australia: Victoria	21289/21290(99%)
ID: MN994468.1	Homo sapiens	nasopharyngeal swab	Jan-2020	USA: CA	21289/21290(99%)
ID: MN938384.1	Homo sapiens	nasopharyngeal swab	Jan-2020	China: Shenzhen	21289/21290(99%)
ID: MN988713.1	Homo sapiens	sputum	Jan-2020	USA: Illinois	21287/21290(99%)
ID: MT027062.1	Homo sapiens	nasopharyngeal swab	Jan-2020	USA: CA	21288/21290(99%)
ID: MN997409.1	Homo sapiens	buccal swab	Jan-2020	USA: AZ	21288/21290(99%)
ID: MN994467.1	Homo sapiens	nasopharyngeal swab	Jan-2020	USA: CA	21288/21290(99%)
ID: MN985325.1 ID: MT039888.1 ID: MT039887.1	Homo sapiens	oropharyngeal swab	Jan-2020	USA	21288/21290(99%)
ID: MT044257.1	Homo sapiens	sputum	Jan-2020	USA: IL	21287/21290(99%)
ID: MN975262.1	Homo sapiens	sputum	Jan-2020	China	21287/21290(99%)

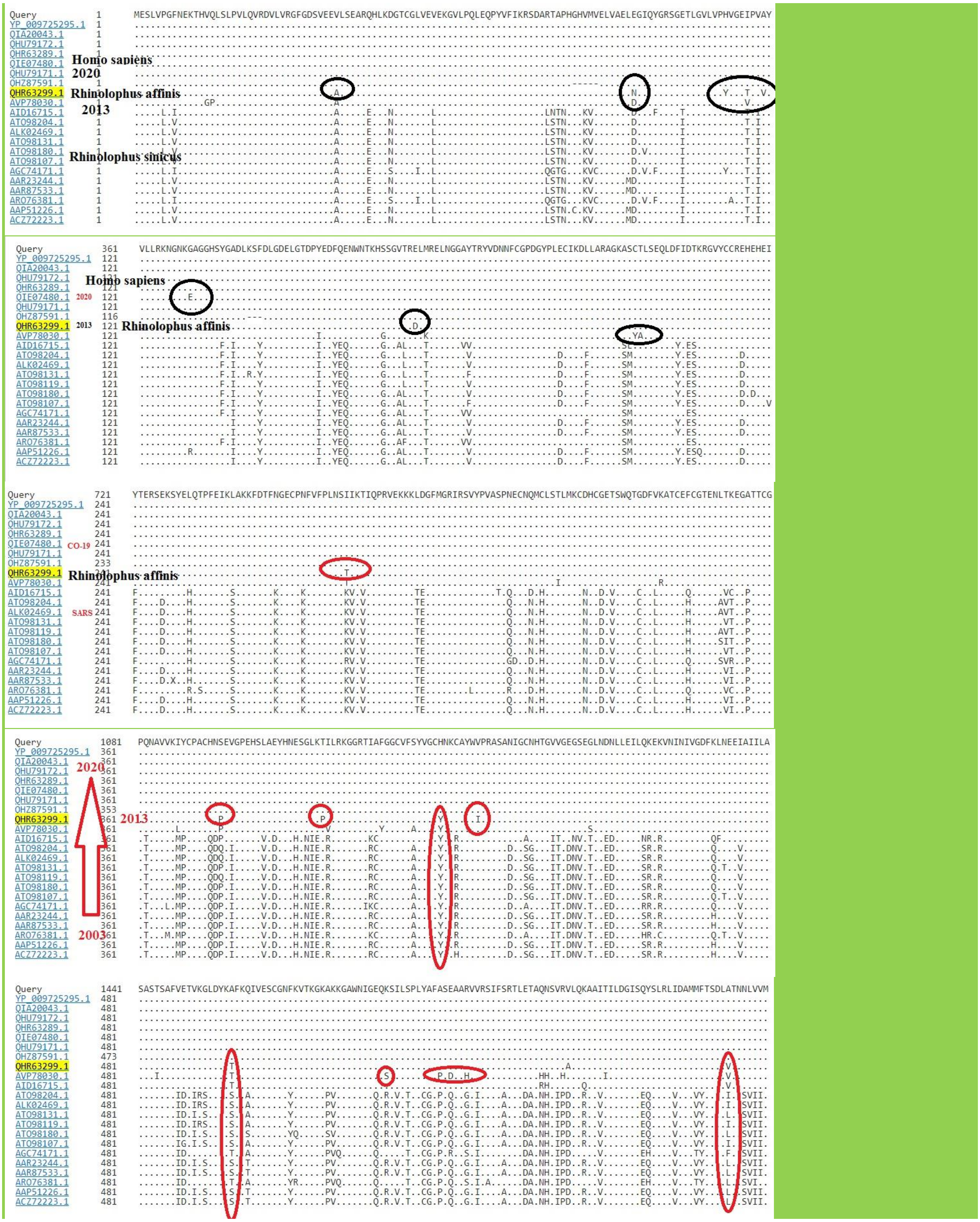
ID: MT039890.1	Homo sapiens	oropharyngeal swab	Jan-2020	South Korea	21285/21290(99%)
ID: MT044258.1	Homo sapiens	oropharyngeal swab	Jan-2020	USA: CA	21266/21290(99%)
ID: LC521925.1	Homo sapiens	oropharyngeal swab	Jan-2020	Japan	21264/21290(99%)
ID: MT020781.1	Homo sapiens	oropharyngeal swab	Jan-2020	Finland	21257/21290(99%)
ID: MT049951.1	Homo sapiens	sputum	Jan-2020	China: Yunnan	21288/21290(99%)
ID: MT093631.1	Homo sapiens	throat swab	Jan-2020	China	21290/21290(100%)
ID: MT072688.1	Homo sapiens	oropharyngeal swab	Jan-2020	Nepal	21290/21290(100%)
ID: MT066176.1	Homo sapiens	oropharyngeal swab	Feb-2020	Taiwan	21288/21290(99%)
ID: MT106053.1	Homo sapiens	nasopharyngeal swab	Feb-2020	USA: CA	21290/21290(100%

Table 3. Represent type of Amino acid polymorphism of Coronavirus isolate between Rhinolophus affinis and Homo sapiens (the Predicted effect is Missense).

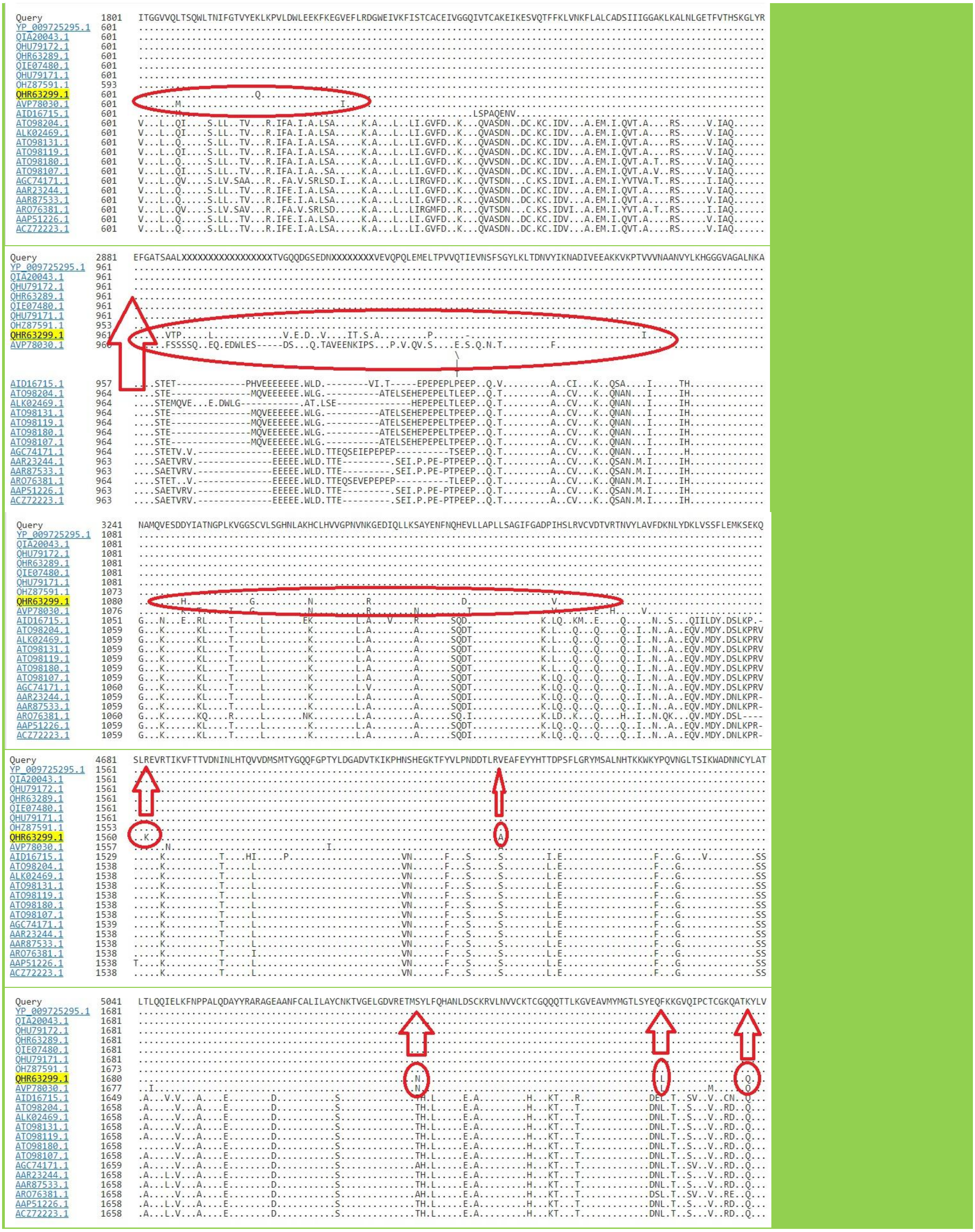
NO.	**Amino acid change between *Rhinolophus affinis\Homo sapiens* ID: MN996532.1 with COVID-19**	**Location**	**NO.**	**Amino acid change *Rhinolophus affinis\Homo sapiens* ID: MN996532.1 with COVID-19**	**Location**
1.	Alanine\ Valine	38	50.	Threonine\ Alanine	1313
2.	Aspartic acid+Asparagine\ Glutamic acid	93	51.	Arginine\ Lysine	1342
3.	Tyrosine \ Histidine	110	52.	Methionine\ Valine	1392
4.	Threonine\ Isoleucine	114	53.	Isoleucine\ Threonine	1541
5.	Valine\ Alanine	117	54.	Isoleucine\ Threonine	1557
6.	Aspartic acid \ Glutamic acid	172	55.	Lysine\ Arginine	1562
7.	Threonine\ Isoleucine	280	56.	Alanine\ Valine	1628
8.	Proline\ Serine	376	57.	Asparagine\ Serine	1732
9.	Proline\ Threonine	395	58.	Glutamine\ Lysine	1794
10.	Tyrosine \Histidine	417	59.	Isoleucine\ Valine	1865
11.	Isoleucine\ Valine	424	60.	Isoleucine\ Threonine	1873
12.	Threonine\ Alanine	498	61.	Isoleucine\ Threonine	1880
13.	Alanine\ Valine	561	62.	Leucine\ Phenylalanine	1936
14.	Valine\ Alanine	591	63.	Phenylalanine \ Leucine	1954
15.	Glutamine\ Lysine	623	64.	Asparagine\ Serine	2014
16.	Proline\ Serine	723	65.	Threonine\ Alanine	2032
17.	Lysine\ Arginine	724	66.	Isoleucine\ Valine	2060
18.	Threonine\ Alanine	859	67.	Aspartic acid \ Asparagine	2081
19.	Isoleucine\ Threonine	891	68.	Glycine \ Serine	2082
20.	Glycine \ Serine	901	69.	Valine\ Alanine	2122
21.	Aspartic acid \ Glutamic acid	940	70.	Phenylalanine \ Serine	2223
22.	Proline\ Serine	944	71.	Threonine\ Lysine	2224
23.	Serine\ Proline	959	72.	Threonine\ Isoleucine	2258
24.	Valine\ Alanine	967	73.	Isoleucine \Threonine	2260

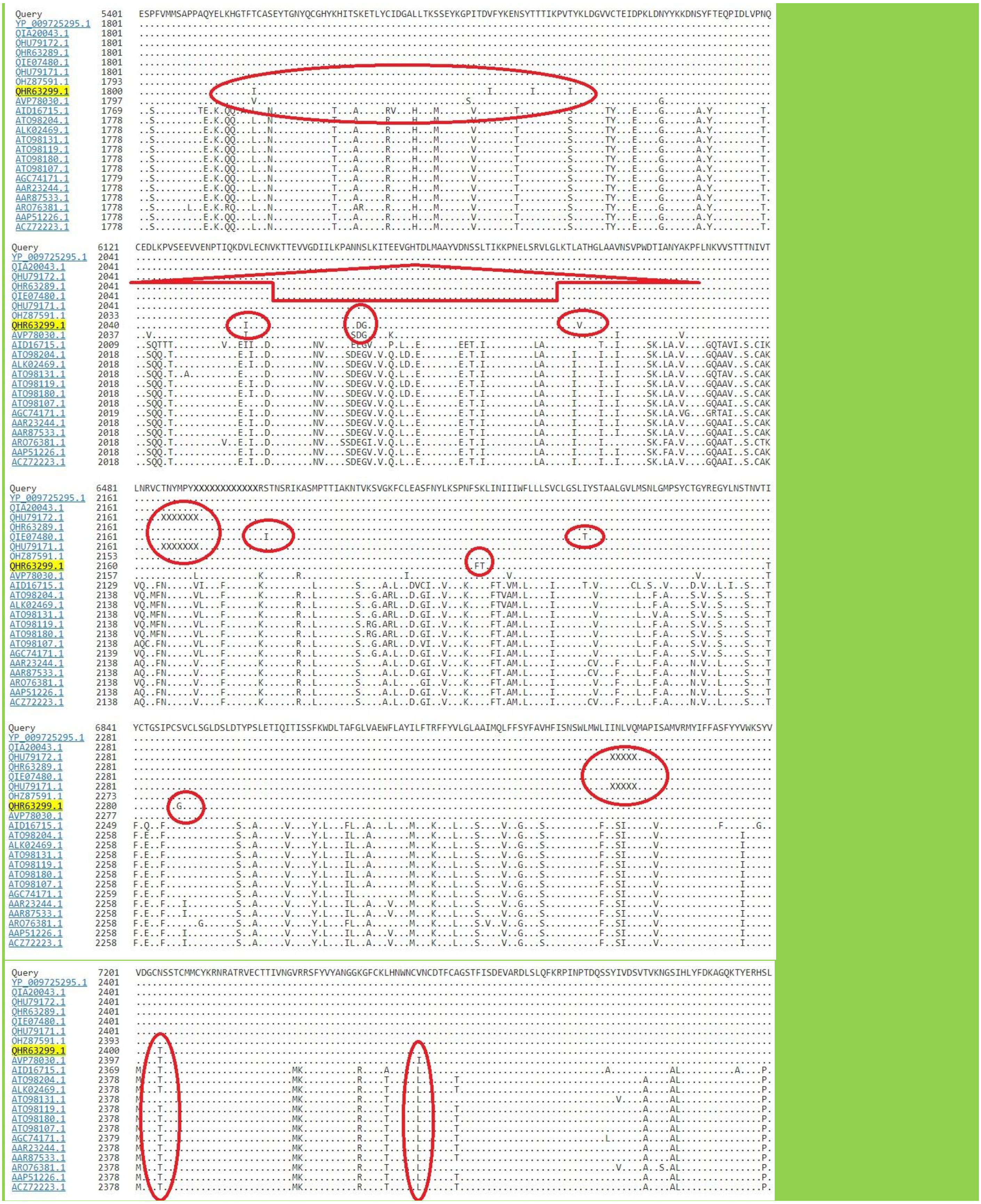
25.	Threonine \ Alanine	968	74.	Glycine \ Serine	2288
26.	Proline\ Leucine	969	75.	Threonine\ Asparagine	2404
27.	Leucine\ Glutamine	975	76.	Serine\ Asparagine	2529
28.	Valine\ Glycine	989	77.	Threonine\ Isoleucine	2583
29.	Glutamic acid\ Glutamine	991	78.	Valine\ Alanine	3142
30.	Aspartic acid \ Glycine	993	79.	Valine\ Isoleucine	3144
31.	Valine\ Aspartic acid	996	80.	Threonine\ Asparagine	3442
32.	Isoleucine\ Threonine	1001	81.	Valine\ Leucine	3605
33.	Threonine\ Isoleucine	1002	82.	Arginine\ Lysine	4534
34.	Serine\ Threonine	1004	83.	Aspartic acid \ Asparagine	4589
35.	Alanine\ Valine	1006	84.	Isoleucine\ Valine	4624
36.	Proline\ Leucine	1016	85.	Histidine\ Tyrosine	5110
37.	Isoleucine\ Valine	1055	86.	Threonine\ Alanine	5828
38.	Histidine\ Aspartic acid	1088	87.	Valine\ Alanine	6004
39.	Glycine \ Serine	1101	88.	Alanine\ Valine	6025
40.	Asparagine\ Histidine	1112	89.	Lysine\ Arginine	6137
41.	Arginine\ Lysine	1123	90.	Arginine\ Leucine	6183
42.	Aspartic acid \ Glutamic acid	1141	91.	Asparagine\ Threonine	6565
43.	Threonine\ Isoleucine	1202	92.	Asparagine\ Aspartic acid	6580
44.	Serine\ Proline	1212	93.	Leucine\ Valine	6623
45.	Serine\ Phenylalanine	1213	94.	Asparagine\ Aspartic acid	6634
46.	Leucine\ Proline	1219	95.	Lysine\ Glutamine	6652
47.	Glutamine\ Lysine	1225	96.	Arginine\ Serine	6694
48.	Valine\ Aspartic acid	1227	97.	Serine\ Phenylalanine	6709
49.	Asparagine\ Serine	1271	98.	Leucine \ Phenylalanine	6714

Figure 1. Neighbor-joining tree of Coronavirus



```
Query  1     MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQHLKDGTCGLVEVEKGVLPQLEQPYVFIKRSDARTAPHGHVMVELVAELEGIQYGRSGETLGVLVPHVGEIPVAY
Query  361   VLLRKNGNKGAGGHSYGADLKSFDLGDELGTDPYEDFQENWNTKHSSGVTRELMRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQLDFIDTKRGVYCCREHEHEI
Query  721   YTERSEKSYELQTPFEIKLAKKFDTFNGECPNFVFPLNSIIKTIQPRVEKKKLDGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSWQTGDFVKATCEFCGTENLTKEGATTCG
Query  1081  PQNAVVKIYCPACHNSEVGPEHSLAEYHNESGLKTILRKGGRTIAFGGCVFSYVGCHNKCAYWVPRASANIGCNHTGVVGEGSEGLNDNLLEILQKEKVNINIVGDFKLNEEIAIILA
Query  1441  SASTSAFVETVKGLDYKAFKQIVESCGNFKVTKGKAKKGAWNIGEQKSILSPLYAFASEAARVVRSIFSRTLETAQNSVRVLQKAAITILDGISQYSLRLIDAMMFTSDLATNNLVVM
```




Query
YP_009725295.1
QIA20043.1
QHU79172.1
QHR63289.1
QIE07480.1
QHU79171.1
QHZ87591.1
QHR63299.1
AVP78030.1
AID16715.1
ATO98204.1
ALK02469.1
ATO98131.1
ATO98119.1
ATO98180.1
ATO98107.1
AGC74171.1
AAR23244.1
AAR87533.1
ARO76381.1
AAP51226.1
ACZ72223.1

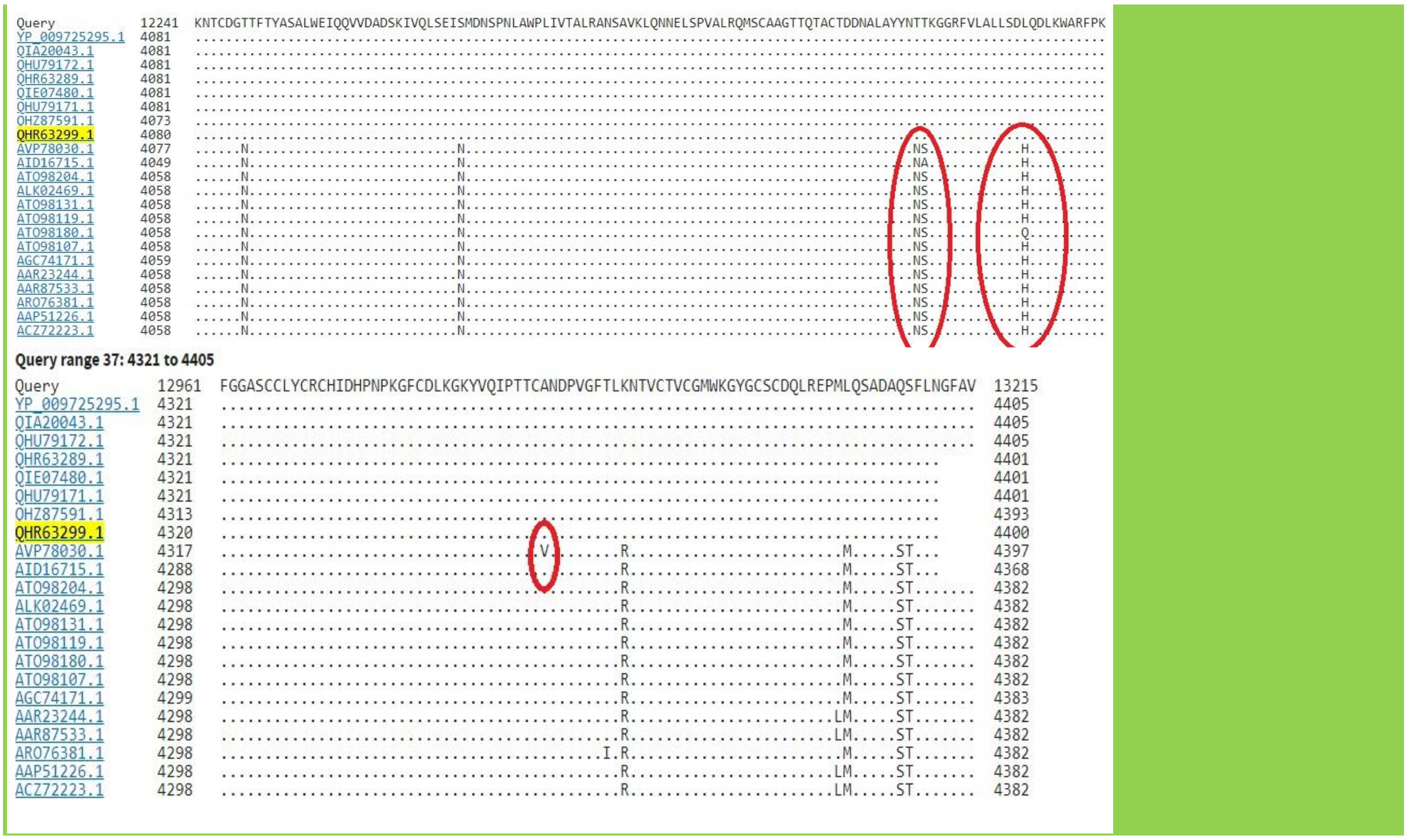
```
Query           7561  FVNLDNLRANNTKGSLPINVIVFDGXXXXXXXXXXXXXVYYSQLMCQPILLLDQALVSDVGDSAEVAVKMFDAYVNTFSSTFNVPMEKLKTLVATAEAELAKNVSLDNVLSTFISAAR
YP_009725295.1  2521
QIA20043.1      2521
QHU79172.1      2521
QHR63289.1      2521
QIE07480.1      2521
QHU79171.1      2521
QHZ87591.1      2513
QHR63299.1      2520
AVP78030.1      2517
AID16715.1      2489
ATO98204.1      2498
ALK02469.1      2498
ATO98131.1      2498
ATO98119.1      2498
ATO98180.1      2498
ATO98107.1      2498
AGC74171.1      2499
AAR23244.1      2498
AAR87533.1      2498
ARO76381.1      2498
AAP51226.1      2498
ACZ72223.1      2498

Query           9001  SGRWVLNNDYYRSLPGVFCGVDAVNLLTNMFTPLIQPIGALDXXXXXXXXXXXXXXXTCLAYYFMRFRRAFGEYSHVVAFNTLLFLMSFTVLCLTPVYSFLPGVYSVIYLYLTFYLTN
YP_009725295.1  3001
QIA20043.1      3001
QHU79172.1      3001
QHR63289.1      3001
QIE07480.1      3001
QHU79171.1      3001
QHZ87591.1      2993
QHR63299.1      3000
AVP78030.1      2997
AID16715.1      2969
ATO98204.1      2978
ALK02469.1      2978
ATO98131.1      2978
ATO98119.1      2978
ATO98180.1      2978
ATO98107.1      2978
AGC74171.1      2979
AAR23244.1      2978
AAR87533.1      2978
ARO76381.1      2978
AAP51226.1      2978
ACZ72223.1      2978

Query           9361  SFLAHIQWMVMFTPLVPFWITIAYIICISTKHFYWFFSNYLKRRVVFNGVSFSTFEEAALCTFLLNKEMYLKLRSDVLLPLTQYNRYLALYNKYKYFSGAMDTTSYREAACCHLAKAL
YP_009725295.1  3121
QIA20043.1      3121
QHU79172.1      3121
QHR63289.1      3121
QIE07480.1      3121
QHU79171.1      3121
QHZ87591.1      3113
QHR63299.1      3120
AVP78030.1      3117
AID16715.1      3089
ATO98204.1      3098
ALK02469.1      3098
ATO98131.1      3098
ATO98119.1      3098
ATO98180.1      3098
ATO98107.1      3098
AGC74171.1      3099
AAR23244.1      3098
AAR87533.1      3098
ARO76381.1      3098
AAP51226.1      3098
ACZ72223.1      3098

Query           9721  FSNSGSDVLYQPPQTSITSAVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVLKLKVDTAN
YP_009725295.1  3241
QIA20043.1      3241
QHU79172.1      3241
QHR63289.1      3241
QIE07480.1      3241
QHU79171.1      3241
QHZ87591.1      3233
QHR63299.1      3240
AVP78030.1      3237
AID16715.1      3209
ATO98204.1      3218
ALK02469.1      3218
ATO98131.1      3218
ATO98119.1      3218
ATO98180.1      3218
ATO98107.1      3218
AGC74171.1      3219
AAR23244.1      3218
AAR87533.1      3218
ARO76381.1      3218
AAP51226.1      3218
ACZ72223.1      3218

Query           10081 TPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNVLAWLYAAVING
YP_009725295.1  3361
QIA20043.1      3361
QHU79172.1      3361
QHR63289.1      3361
QIE07480.1      3361
QHU79171.1      3361
QHZ87591.1      3353
QHR63299.1      3360
AVP78030.1      3357
AID16715.1      3329
ATO98204.1      3338
ALK02469.1      3338
ATO98131.1      3338
ATO98119.1      3338
ATO98180.1      3338
ATO98107.1      3338
AGC74171.1      3339
AAR23244.1      3338
AAR87533.1      3338
ARO76381.1      3338
AAP51226.1      3338
ACZ72223.1      3338
```



```
Query           12241  KNTCDGTTFTYASALWEIQQVVDADSKIVQLSEISMDNSPNLAWPLIVTALRANSAVKLQNNELSPVALRQMSCAAGTTQTACTDDNALAYYNTTKGGRFVLALLSDLQDLKWARFPK
YP_009725295.1  4081   ......................................................................................................................
QIA20043.1      4081   ......................................................................................................................
QHU79172.1      4081   ......................................................................................................................
QHR63289.1      4081   ......................................................................................................................
QIE07480.1      4081   ......................................................................................................................
QHU79171.1      4081   ......................................................................................................................
QHZ87591.1      4073   ......................................................................................................................
QHR63299.1      4080   ......................................................................................................................
AVP78030.1      4077   ......N.............................N.......................................................NS...........H............
AID16715.1      4049   ......N.............................N.......................................................NA...........H............
ATO98204.1      4058   ......N.............................N.......................................................NS...........H............
ALK02469.1      4058   ......N.............................N.......................................................NS...........H............
ATO98131.1      4058   ......N.............................N.......................................................NS...........H............
ATO98119.1      4058   ......N.............................N.......................................................NS...........H............
ATO98180.1      4058   ......N.............................N.......................................................NS...........Q............
ATO98107.1      4058   ......N.............................N.......................................................NS...........H............
AGC74171.1      4059   ......N.............................N.......................................................NS...........H............
AAR23244.1      4058   ......N.............................N.......................................................NS...........H............
AAR87533.1      4058   ......N.............................N.......................................................NS...........H............
ARO76381.1      4058   ......N.............................N.......................................................NS...........H............
AAP51226.1      4058   ......N.............................N.......................................................NS...........H............
ACZ72223.1      4058   ......N.............................N.......................................................NS...........H............
```

Query range 37: 4321 to 4405

```
Query           12961  FGGASCCLYCRCHIDHPNPKGFCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCGMWKGYGCSCDQLREPMLQSADAQSFLNGFAV  13215
YP_009725295.1  4321   .....................................................................................  4405
QIA20043.1      4321   .....................................................................................  4405
QHU79172.1      4321   .....................................................................................  4405
QHR63289.1      4321   .................................................................................      4401
QIE07480.1      4321   .................................................................................      4401
QHU79171.1      4321   .................................................................................      4401
QHZ87591.1      4313   .................................................................................      4393
QHR63299.1      4320   .................................................................................      4400
AVP78030.1      4317   ....................................V........R........................M.....ST...      4397
AID16715.1      4288   .............................................R........................M.....ST...      4368
ATO98204.1      4298   .............................................R........................M.....ST.......  4382
ALK02469.1      4298   .............................................R........................M.....ST.......  4382
ATO98131.1      4298   .............................................R........................M.....ST.......  4382
ATO98119.1      4298   .............................................R........................M.....ST.......  4382
ATO98180.1      4298   .............................................R........................M.....ST.......  4382
ATO98107.1      4298   .............................................R........................M.....ST.......  4382
AGC74171.1      4299   .............................................R........................M.....ST.......  4383
AAR23244.1      4298   .............................................R.......................LM.....ST.......  4382
AAR87533.1      4298   .............................................R.......................LM.....ST.......  4382
ARO76381.1      4298   ...........................................I.R........................M.....ST.......  4382
AAP51226.1      4298   .............................................R.......................LM.....ST.......  4382
ACZ72223.1      4298   .............................................R.......................LM.....ST.......  4382
```

Figure 2. Sequence amino acid analysis of Coronavirus isolate by NCBI.

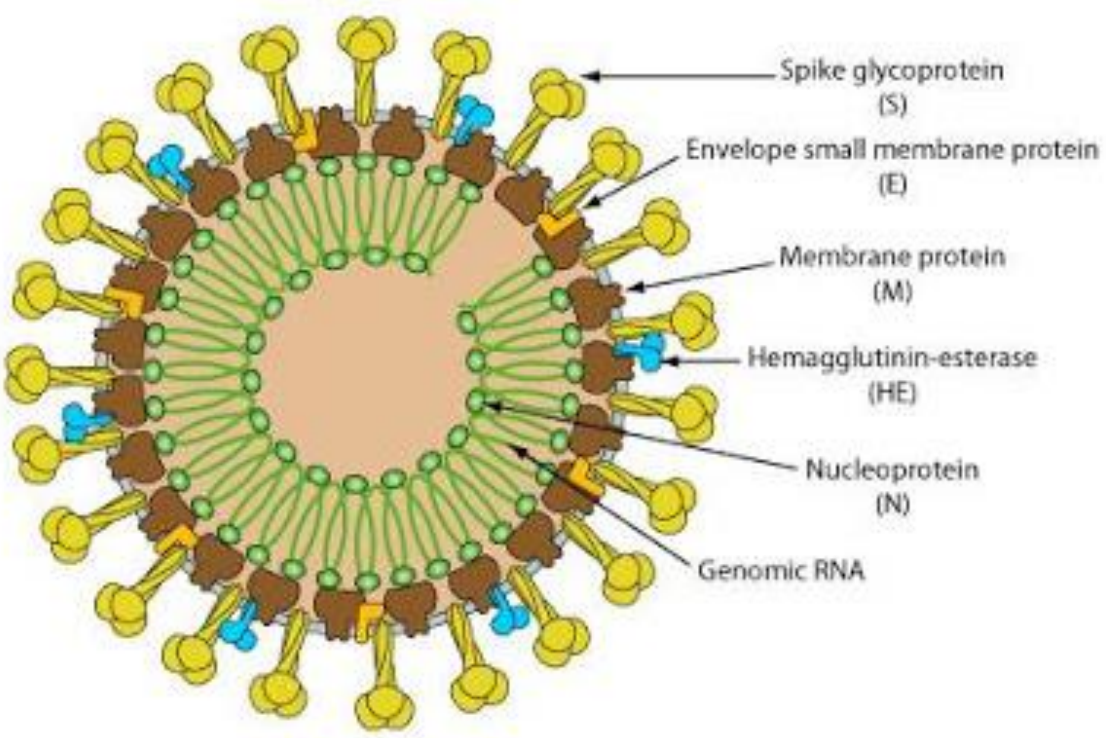


Figure 3. Diagram of a coronavirus structure.

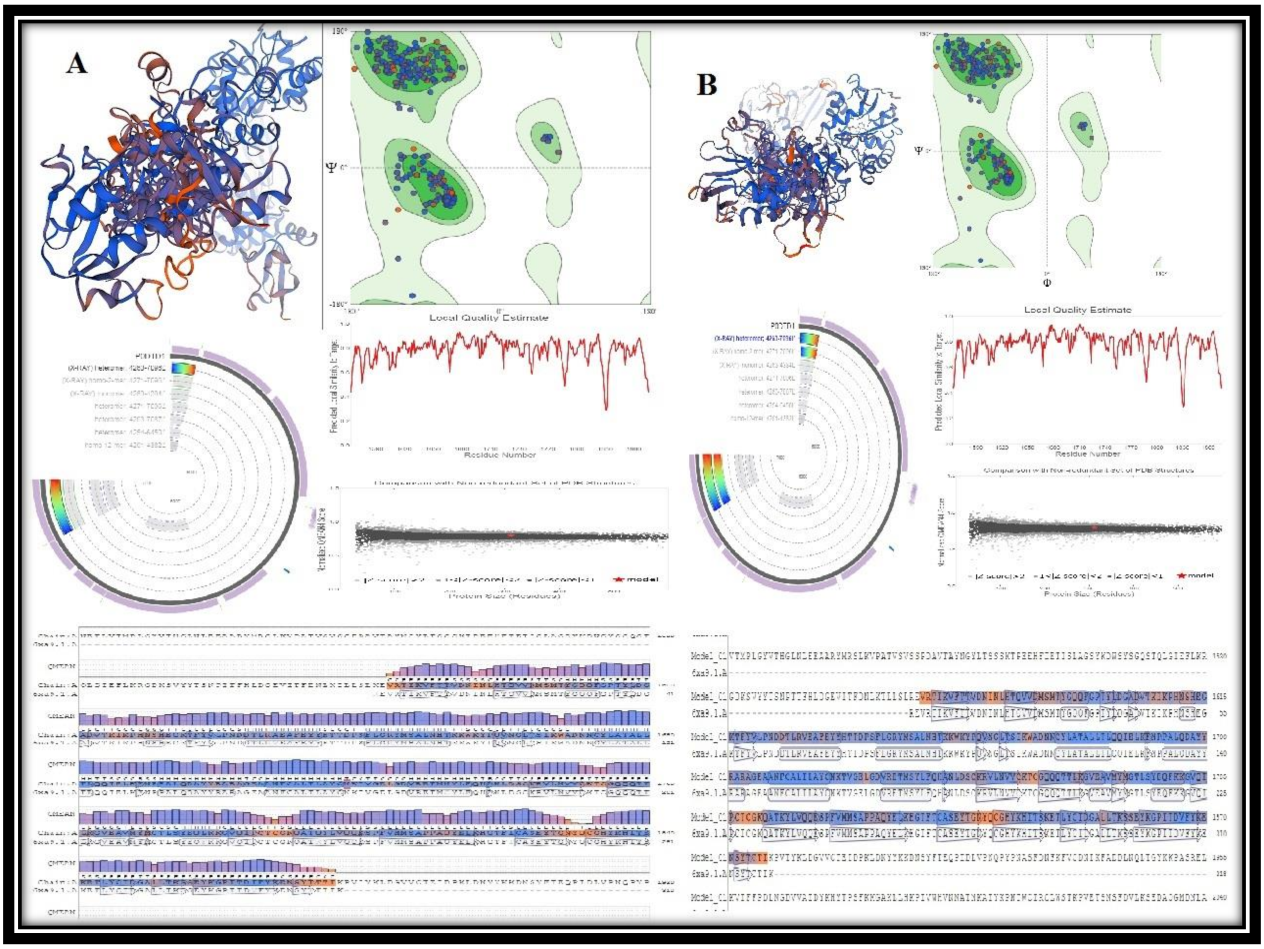


Figure 4. Conformation of protein from COVID-19, (A: Bat Coronavirus and B:human coronavirus).

Acknowledgment

The author is very grateful to the University of Baghdad /College of Veterinary Medicine for their provided facilities, which helped to improve of this study.

.

References

1. China SARS. Severe Acute Respiratory Syndrome (SARS) Statistics [Internet]. Mainland China;2003 May 17. Available from http: // www . chinacdc . net . cn / default.asp.
2. Kuo L, Godeke GJ, Raamsman MJ, Masters PS, Rottier PJ. Retargeting of coronavirus by substitution of the spike glycoprotein ectodomain: crossing the host cell species barrier. J Virol. 2000; 74**:** 1393–406.
3. Huang C, Wang Y, Li X. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, Elsevier. 2020; 24:1-11
4. Phan LT, Nguyen TV, Luong QC. Importation and human-to-human transmission of a novel coronavirus in vietnam. N Engl J Med (NEJM). 2020; 27 (9):872-874
5. Giovanetti M, Benvenuto D, Angeletti S, Ciccozzi M. The first two cases of 2019-nCoV in Italy. J Med Virol. 2020;92(5):518-521.
6. Raquel, s, Julio T, Gloria M. Chinese Society of Radiology. Radiological findings for diagnosis of SARS-CoV-2 pneumonia (COVID-19) Medicina Clínica j.2020; 155 (1): 36-40.
7. Wang Q, Wang YH, Ma JC. Description of the first strain of 2019-nCoV, C-Tan-nCoV Wuhan Strain. Chinese Center for Disease Control and Prevention. 2020;2(6)81-82.
8. Kirchdoerfer RN, Cottrell CA, Wang N. Pre-fusion structure of a human coronavirus spike protein. Nature 2016; 531(7592):118–121.
9. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J. Mol. Biol. 1990; 215: 403-410.

10. Koichiro T, Glen S, Daniel P, Alan F, Sudhir K. Molecular evolutionary Genetics Analysis version 6.0 (MEGA6). Mol. Biol. Evol. 2013; 30(12):2725–2729.

11. Hu B, Ge X, Wang LF, Shi Z. Bat origin of human coronaviruses. Virol. J. 2015; 12 (221):2-10.

12. Ksiazek TG. A novel coronavirus associated with severe acute respiratory syndrome. N. Engl. J. Med. 2003; (348)1953–1966.

13. Dan H, Changqiang Z, Lele A, Ting H, Yi W, Fuqiang Y, Lu Y, Chenxi D. Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats. Emerging Microbes and Infections. 2018; 7(1):154.

14. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 1999; 41: 95-98.

15. Graham RL, Sparks JS, Eckerle LD, Sims AC, Denison MR. SARS coronavirus replicase proteins in pathogenesis. Virus Res. 2008;133(1):88-100.

16. De Haan CM, Kuo L, Masters PS, Vennema H, Rottier PM. Coronavirus particle assembly: primary structure requirements of the membrane protein. J Virol 1998; 72 (8): 6838-6850.

17. Tian X, Li C, Huang A, Xia S, Lu S, Shi Z. Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus-specific human monoclonal antibody..Emer. Micr. and Infec.. 2020; 9(1):382-385.

18. Simmons NB, Wilson DE, Reeder DM. Mammal species of the world. Baltimore: J. Hopkins. 2005; 312–529.

19. Woo PC, Lau SK, Li KS, Poon RW, Wong BH, Tsoi HW. Molecular diversity of coronaviruses in bats. Virology. 2006; 35 (180): 7-8.

20. Wang M, Hu Z. Bats as animal reservoirs for the SARS coronavirus. *Virol Sin* 2013; 28**:** 15–17.

21. Zhou H, Fan T, Lan XL, Yang WF. Shi W, Zhang Y. Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin. Nature 2018; (556): 255-274.

22. Hu LP, Zeng XL, Yang XY, Ge W, Zhang B, Li JZ. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLoS Pathog. 2017; (13): 1-27.

23. Lam Wk, Zhong Ns, Tan Wc. Overview on SARS in Asia and the World. Respirology. 2003; 8: 2–5

24. Karim S, Ali A, Abdulrazak SH, Usama AA. Seropositivity rate of middle east respiratory syndrome coronavirus among Iraqi dromedary Arabic camels. Iraqi J. Vet. Med.2017; 41(2): 163-167.

25. Walls AC, Park YJ, Tortorici M, Wall A, McGuire A. Structure, function, and antigenicity of the SARS-CoV-2 Spike glycoprotein. Cell. 2020; 181(2): 212-218.

26. VanHemert MJ, VanDen W, Knoops K, Mommaas AM, Gorbalenya AE, Snijder E J. SARS-coronavirus replication/ transcription complexes are membrane -protected and need a host factor for activity in vitro. PLoS Pathogens 2008; 133(2): 136–148

27. Wu F, Zhao S, Yu B, Chen YM, Wang W. A new coronavirus associated with human respiratory disease in China. Nature 2020; 579 (7798):1-8.

28. Zhou P, Yang XL, Wang XG, Hu B. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature, 2020; 579(7798), 270–273.

29. Verdia-Baguena, C, Nieto-Torres, JL, Alcaraz A, DeDiego ML. Coronavirus E protein forms ion channels with functionally and structurally-involved membrane lipids. Virology, 2012; 432(2), 485–494.

30. Tian X, Li C, Huang A, Xia S, Lu S, Shi Z. Potent binding of 2019 novel coronavirus spike protein by a SARS coronavirus- specific human monoclonal antibody. Emer. Micr. and Infec. 2020; 9(1): 382–385.

تطور الخريطة الجزيئية الجينية وشجرة الوراثية لفيروس كورونا (كوفيد -19) : المقالة المراجِعية

حيدر عبد الكريم حسن المطر

فرع الجراحة والتوليد، كلية الطب البيطري، جامعة بغداد، بغداد، العراق

الخلاصة

يُشتبه في أن فيروس كورونا (كوفيد -19) نشأ من مضيف حيواني (مشترك) متبوعاً بانتقال من إنسان إلى إنسان، والهدف من مقالة المرجعية هو معرفة الشجرة الوراثية وشكل البروتين للفيروس التاجي .من خلال دراسة التسلسل الجيني، كما هو موضح في تصميم شجرة التطورية والتغايرات الجينية، فإن السبب الرئيسي لـ COVID-19 هو فيروس Bat Coronavirus RaTG13 (SARS) في ID: MN996532.1 و ID: MG772933.1 (24\تموز (يوليو) \2013 وشباط (فبراير)\2017)، حيث يظهر تطور الفيروس التاجي من 2003-2013 إلى 2020، والذي أصبح الذروة الأكثر فتكًا في البشر، إذ تطور الفيروس من تأثير الخفافيش على البشر. إذ أظهرت نتائج تحليل البروتين هنالك 98 تغيير في الأحماض ألأمينيه من RaTG13 (SARS) إلى COVID-19 Homo sapiens. وهكذا، فقد زادت هذه الدراسة من فهمنا للتنوع الجيني لـ COVID-19 التي تحملها الخفافيش، كذلك نستنتج أن تركيب وشكل البروتين متطابق بين الخفافيش والإنسان.

الكلمات المفتاحية: **السارس، كوفيد -19، التطور الجزيئي، الخفاش.**