2020 SCIENCELINE
Journal of Life Science and Biomedicine
J Life Sci Biomed, 10 (2): 17-20, 2020

ISSN 2251-9939

DOI: https://dx.doi.org/10.29252/scil.2020.jlsb3

Alignment of SARS-CoV-2 in comparison with other coronaviruses

Mohamed Samy ABOUSENNA

PhD of Virology, Central Laboratory for Evaluation of Veterinary Biologics (CLEVB), Agricultural Research Center (ARC), Cairo, Egypt

Corresponding author's E-Mail: mohamedsamy2020@hotmail.com; ORCID: 0000-0003-2202-9544

ABSTRACT

Introduction. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) is currently declared as pandemic according to the WHO. It was initially detected in China and then rapidly transmitted to most world territories. The SARS-CoV-2 has an ambiguous origin, with unique properties, pathogenesis and transmission rate, thus making its prevention and control a difficult task. Aim. In the present study, we investigated the origin hypotheses through conducting multiple alignments and phylogenetic analysis for surface glycoprotein and complete genome of SARS-CoV-2 in comparison with other coronaviruses of different species. All the data used in this study were obtained from NCBI online database and analyzed using Blast tool. The alignment and phylogenetic analysis of SARS-CoV-2 surface glycoprotein in comparison with spike glycoprotein of Bat coronavirus RaTG13, Pangolin coronavirus, Bat SARS-like coronavirus, SARS-CoV, BCoV, IBV, ECoV, MHV-JHM, MERS-CoV, CCoV, HCoV-229E and FCOV indicated close identical matching to spike protein for Bat coronavirus RaTG13 and Pangolin coronavirus isolate MP789. The similarity was 97.41% and 96.67%, respectively. Also, multiple alignments of complete genome for SARS-CoV-2 and Bat coronavirus RaTG13 showed a significant similarity of 96.11%. Recommendation. Therefore, these relevant results strongly recommend the origin hypothesis of SARS-CoV-2 from Bat coronavirus RaTG13. The nature of evolution is considered to be natural selection.

Original Article
PII: S225199392000003-10

Rec. 05 March 2020
Rev. 20 March 2020
Pub. 25 March 2020

Keywords
**Severe Acute Respiratory Syndrome,
SARS-CoV-2,
COVID-19,
Coronavirus Disease 2019,
Alignment,
Phylogenetic analysis**

INTRODUCTION

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) was recently detected and isolated in Wuhan Province, China. The virus causes severe acute respiratory illness which is called COVID-19 (Coronavirus Disease, 2019) [1]. SARS-CoV-2 is the seventh coronavirus which infects humans; SARS-CoV, MERS-CoV, SARS-CoV-2, HKU1, NL63, OC43, and 229E, while the recorded data till now demonstrate the severe cases of SARS-CoV-2 not more than 10% and the mortality rate not more than 5 % [2, 3].

SARS-CoV-2 belongs to the family *Coronaviridae* in the order *Nidovirales*. It could be classified into four genera: Alphacoronavirus, Betacoronavirus, Gammacoronavirus, and Deltacoronavirus, whereas alpha- and betacoronaviruses infect mammals (human coronavirus NL63 (HCoV-NL63), porcine transmissible gastroenteritis coronavirus (TGEV), PEDV, porcine respiratory coronavirus (PRCV), SARS-CoV, MERS-CoV, bat coronavirus HKU4, mouse hepatitis coronavirus (MHV), bovine coronavirus (BCoV), and human coronavirus OC43). Gammacoronaviruses infect avian species e.g.; avian infectious bronchitis coronavirus (IBV) and deltacoronaviruses infect both mammalian and avian species (porcine deltacoronavirus (PdCV).

Coronaviruses are large, enveloped, positive-stranded RNA viruses. The genome is surrounded by the nucleocapsid protein (N) and further surrounded by an envelope. The viral envelope is associated with three structural proteins: the envelope protein (E) and the membrane protein (M) are involved in virus assembly, whereas the spike protein (S) binds the host cell receptors to mediate virus entry. Furthermore, the spike protein is a critical determinant of viral host range, tissue tropism and a major inducer of host immune responses, which is significant for developing vaccines in many species [4]. It was recently confirmed that SARS-CoV-2 uses the Spike protein (S) to bind human cell receptors to process the cell entry [5] like other related coronaviruses.

The present study aimed to investigate the origin hypothesis of SARS-CoV-2 through conducting spike protein alignment and phylogenetic analysis for different coronaviruses in comparison with SARS-CoV-2.

MATERIAL AND METHODS

Multiple alignment and phylogenetic analysis had been conducted for surface glycoprotein [Severe acute respiratory syndrome coronavirus 2] (accession: QIJ96523.1) in comparison with spike glycoprotein [Bat coronavirus RaTG13] (accession: QHR63300.2), Spike glycoprotein [Pangolin coronavirus isolate MP789] (accession: MT084071.1), spike protein [Bat SARS-like coronavirus] (accession: ATO98132.1), E2 glycoprotein precursor [Severe acute respiratory syndrome-related coronavirus] (accession: AAP41037.1), spike protein [Bovine respiratory coronavirus AH187] (accession: ACT11007.1), spike protein [Infectious bronchitis virus] (accession: AAP92675.1), spike protein [Equine coronavirus] (accession: BAS18866.1), spike glycoprotein (S) [Murine hepatitis virus strain JHM] (accession: AAU06356.1), S protein [Middle East respiratory syndrome-related coronavirus] (accession: ANF29250.1), spike protein [Canine coronavirus] (accession: AKA65829.1), surface glycoprotein [Human coronavirus 229E] (accession: QEO75985.1) and spike protein [Feline coronavirus] (accession: ASU62499.1). SARS-CoV-2 isolate 2019-nCoV/USA-CruiseA-26/2020, complete genome (accession; MT184913.1) had been aligned with Bat coronavirus RaTG13, complete genome (accession: MN996532.1) and Pangolin coronavirus isolate MP789 complete genomic sequence (accession: MT084071.1).

All the data used in this study had been obtained from GenBank (NCBI Database), the alignment and phylogenetic analysis was conducted using Blast tool at NCBI.

RESULTS

The alignment and phylogenetic analysis of SARS-CoV-2surface glycoprotein in comparison with spike glycoprotein of Bat coronavirus RaTG13, Pangolin coronavirus, Bat SARS-like coronavirus, SARS-CoV, Bovine respiratory coronavirus AH187 (BCoV), Infectious bronchitis virus (IBV), Equine coronavirus (ECoV), Murine hepatitis virus strain JHM (MHV-JHM), Middle East respiratory syndrome-related coronavirus (MERS-CoV), Canine coronavirus (CCoV), Human coronavirus 229E (HCoV-229E) and Feline coronavirus (FCoV) indicated close identical matching to spike protein for Bat coronavirus RaTG13 and Pangolin coronavirus isolate MP789 as shown in table 1 and figure 1. Furthermore, alignment of complete genome of Pangolin coronavirus isolate MP789 in comparison with SARS-CoV-2 isolate 2019-nCoV/USA-CruiseA-26/2020 and Bat coronavirus RaTG13 indicated identity matching as shown in table 2, while the alignment of SARS-CoV-2isolate 2019-nCoV/USA-CruiseA-26/2020 in comparison with Bat coronavirus RaTG13 indicated closely identical matching as shown in table 3.

Table 1. Alignment of surface glycoprotein SARS-CoV-2 (accession: QIJ96523.1) in comparison with other coronavirus spike proteins of different species.

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
Spike glycoprotein [Bat coronavirus ratg13]	2565	2565	100%	0.0	97.41%	QHR63300.2
Spike glycoprotein [Pangolin coronavirus isolate MP789]	1013	2047	87%	0.0	96.67%	MT084071.1
Spike protein [Bat SARS-like coronavirus]	2049	2049	99%	0.0	77.23%	ATO98132.1
E2 glycoprotein precursor [Severe acute respiratory syndrome-related coronavirus]	2038	2038	100%	0.0	75.96%	AAP41037.1
Spike protein [Bovine respiratory coronavirus AH187]	483	632	78%	3e-152	37.55%	ACT11007.1
Spike protein [Infectious bronchitis virus]	351	415	50%	5e-106	37.43%	AAP92675.1
Spike protein [Equine coronavirus]	466	580	71%	6e-146	37.14%	BAS18866.1
Spike glycoprotein (S) [Murine hepatitis virus strain JHM]	459	616	70%	6e-144	36.45%	AAU06356.1
S protein [Middle East respiratory syndrome-related coronavirus]	551	632	85%	6e-178	35.10%	ANF29250.1
Spike protein [Canine coronavirus]	340	433	69%	9e-101	32.07%	AKA65829.1
Surface glycoprotein [Human coronavirus 229E]	364	484	65%	1e-110	31.07%	QEO75985.1
Spike protein [Feline coronavirus]	340	510	70%	1e-101	30.75%	ASU62499.1

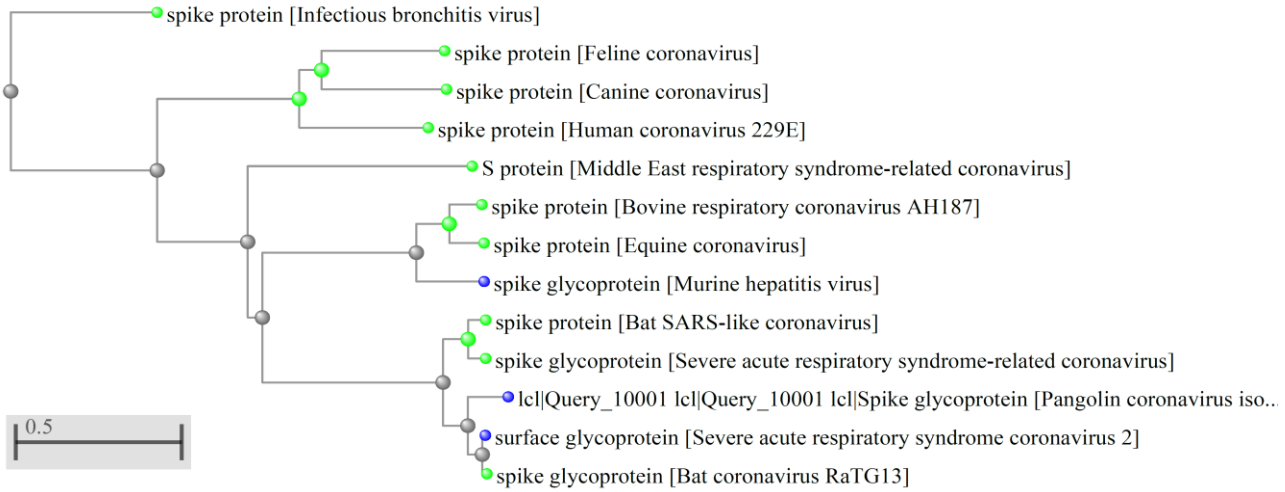

Figure 1. Phylogenetic analysis for surface glycoprotein SARS-CoV-2 (accession: QIJ96523.1) in comparison with other coronavirus spike proteins of different species.

Table 2. Alignment of Pangolin coronavirus isolate MP789 genomic sequence MT084071.1 in comparison with SARS-CoV-2 isolate and Bat coronavirus RaTG13complete genomes.

Description	Mzax Score	Total Score	Query Cover	E value	Per. Ident	Accession
Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CruiseA-26/2020, complete genome	492	489	87%	0.0	91.63%	MT184913.1
Bat coronavirus RaTG13, complete genome	9389	1505	87%	0.0	91.38%	MN996532.1

Table 3. Alignment of SARS-CoV-2 isolate 2019-nCoV/USA-CruiseA-26/2020, complete genome MT184913.1 in comparison with Bat coronavirus RaTG13, complete genome.

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
Bat coronavirus RaTG13, complete genome	48700	48700	99%	0.0	96.11%	MN996532.1

DISCUSSION

Coronaviruses had been a life-threatening impact on the human for almost 18 years, SARS and Middle East respiratory syndrome (MERS) [6, 7], recently the SARS-CoV-2 has been isolated in Wuhan 1, It had been thought the virus could be originally from bats or other animal species or probably passage in intermediate host before human infection, actually, there are different hypothesizes and conflicts.

In the current study, the origin hypothesis of SARS-CoV-2 through conducting spike protein alignment and phylogenetic analysis for different coronaviruses in comparison with SARS-CoV-2 as well as the complete genome multiple alignment for the viruses which indicated significant similarity of spike protein with SARS-CoV-2, has been investigated.

The relevant analysis results indicated significant alignments of the surface glycoprotein SARS-CoV-2 (QIJ96523.1) which was closely related to spike glycoprotein Bat coronavirus RaTG13 (QHR63300.2) and spike glycoprotein Pangolin coronavirus (MT084071.1) showed similarity 97.41% and 96.67% respectively, while the alignment of surface glycoprotein SARS-CoV-2 (QIJ96523.1) in comparison with spike glycoprotein of Bat SARS-like coronavirus, SARS-related coronavirus, Bovine respiratory coronavirus AH187, Infectious bronchitis virus, Equine coronavirus, Murine hepatitis virus strain JHM, Middle East respiratory syndrome-related coronavirus, Canine coronavirus, Human coronavirus 229E and Feline coronavirus showed similarity 77.23%, 75.96%, 37.55%, 37.43%, 37.14%, 36.45%, 35.10%, 32.07%, 31.07%, and 30.75%, respectively (Figure 1 and table 1) ,the phylogenetic analysis suggested the SARS-CoV-2 is belonged to betacoronaviruses [8] and relative closely to Bat coronavirus RaTG13 and Pangolin coronavirus [9, 10], as indicated too in multiple alignments of complete genome for SARS-

CoV-2 and Bat coronavirus RaTG13 which showed a significant similarity 96.11% (Table 3), while the multiple a alignments of complete genome for Pangolin coronavirus isolate MP789 compared to SARS-CoV-2 and Bat coronavirus RaTG13 showed similarity 91.63% and 91.38%, thus strongly recommend the hypothesis that the SARS-CoV-2 is originated of Bat coronavirus RaTG13 regardless other coronaviruses and it is natural selection.

CONCLUSION

It is found that the hypothesis origin of SARS-CoV-2 is almost certainly natural origin, that impetus us for future investigations for the susceptibility of other species to SARS-CoV-2 infection and the impact of immune system of other species on SARS-CoV-2.

Competing interests

The author declares that he has no conflict of interests.

REFERENCES

1. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, Yuan ML. A new coronavirus associated with human respiratory disease in China. Nature. 2020; 579(7798):265-9. doi: https://doi.org/10.1038/s41586-020-2008-3 Google Scholar
2. CDC. Human Coronavirus types 2019. https://www.cdc.gov/coronavirus/types.html
3. WHO. Coronavirus disease 2019 (COVID-19) Situation Report – 64 https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200324-sitrep-64-covid-19.pdf?sfvrsn=703b2c40_2
4. Li F. Structure, function, and evolution of coronavirus spike proteins. Annual review of virology. 2016; 3:237-61. DOI: https://doi.org/10.1146/annurev-virology-110615-042301 Google Scholar
5. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell. 2020; 181(2):271-280. doi: https://doi.org/10.1016/j.cell.2020.02.052 Google Scholar
6. Drosten C, Günther S, Preiser W, Van Der Werf S, Brodt HR, Becker S, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. New England journal of medicine. 2003; 348(20):1967-76. doi: https://doi.org/10.1056/NEJMoa03074 Google Scholar
7. 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. doi: https://doi.org/10.1056/NEJMoa1211721 Google Scholar
8. Jaimes JA, André NM, Chappie JS, Millet JK, Whittaker GR. Phylogenetic analysis and structural modeling of SARS-CoV-2 spike protein reveals an evolutionary distinct and proteolytically sensitive activation loop. Journal of molecular biology. 2020; 432(10):3309-25. doi: https://doi.org/10.1016/j.jmb.2020.04.009 Google Scholar
9. Lam TT, Jia N, Zhang YW, Shum MH, Jiang JF, Zhu HC, et al. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. Nature. 2020; 583(7815):282-5. doi: https://doi.org/10.1038/s41586-020-2169-0 Google Scholar
10. Lv L, Li G, Chen J, Liang X, Li Y. Comparative genomic analysis revealed specific mutation pattern between human coronavirus SARS-CoV-2 and Bat-SARSr-CoV RaTG13. BioRxiv. 2020 Jan 1, doi: https://doi.org/10.1101/2020.02.27.969006 Google Scholar