

Editorial

# The Importance of Research on the Origin of SARS-CoV-2

Kenneth Lundstrom <sup>1,\*</sup>, Murat Seyran <sup>2,3</sup>, Damiano Pizzol <sup>4</sup> , Parise Adadi <sup>5</sup> , Tarek Mohamed Abd El-Aziz <sup>6,7</sup> , Sk. Sarif Hassan <sup>8</sup> , Antonio Soares <sup>6</sup>, Ramesh Kandimalla <sup>9,10</sup>, Murtaza M. Tambuwala <sup>11</sup> , Alaa A. A. Aljabali <sup>12</sup> , Gajendra Kumar Azad <sup>13</sup> , Pabitra Pal Choudhury <sup>14</sup>, Vladimir N. Uversky <sup>15</sup> , Samendra P. Sherchan <sup>16</sup>, Bruce D. Uhal <sup>17</sup> , Nima Rezaei <sup>18,19</sup>  and Adam M. Brufsky <sup>20</sup> 

<sup>1</sup> PanTherapeutics, CH1095 Lutry, Switzerland

<sup>2</sup> Doctoral Studies in Natural Sciences and Technology SPL44, University of Vienna, 1010 Vienna, Austria; muratseyran@gmail.com

<sup>3</sup> Network of Immunity in Infection, Malignancy and Autoimmunity (NIIMA), Universal Scientific Education and Research Network (USERN), Vienna, 1010 Vienna, Austria

<sup>4</sup> Department of Global Health, Italian Agency for Development Cooperation—Khartoum, Al Amarat 111111, Sudan; damianopizzol8@gmail.com

<sup>5</sup> Department of Food Science, University of Otago, Dunedin 9054, New Zealand; pariseadadi@gmail.com

<sup>6</sup> Department of Cellular and Integrative Physiology, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229-3900, USA; mohamedt1@uthscsa.edu (T.M.A.E.-A.); soaresa@uthscsa.edu (A.S.)

<sup>7</sup> Zoology Department, Faculty of Science, Minia University, El-Minia 61519, Egypt

<sup>8</sup> Department of Mathematics, Pingla Thana Mahavidyalaya, Maligram, Paschim Medinipur, West Bengal 721140, India; sarimif@gmail.com

<sup>9</sup> Department of Applied Biology, CSIR-Indian Institute of Chemical Technology, Tarnaka, Hyderabad 500007, India; ramesh.kandimalla@gmail.com

<sup>10</sup> Department of Biochemistry, Kakatiya Medical College/MGM-Hospital, Hyderabad 500007, India

<sup>11</sup> School of Pharmacy and Pharmaceutical Science, Ulster University, Coleraine BT52 1SA, Northern Ireland, UK; m.tambuwala@ulster.ac.uk

<sup>12</sup> Department of Pharmaceutical Sciences, Faculty of Pharmacy, Yarmouk University, Irbid 21163, Jordan; alaaaj@yu.edu.jo

<sup>13</sup> Department of Zoology, Patna University, Patna, Bihar 800005, India; gkazad@patnauniversity.ac.in

<sup>14</sup> Applied Statistics Unit, Indian Statistical Institute, Kolkata, West Bengal 700108, India; pabitrapalchoudhury@gmail.com

<sup>15</sup> Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA; vuffersky@usf.edu

<sup>16</sup> Department of Environmental Health Sciences, Tulane University, New Orleans, LA 70112, USA; sshercha@tulane.edu

<sup>17</sup> Department of Physiology, Michigan State University, East Lansing, MI 48824, USA; bduhal@gmail.com

<sup>18</sup> Center for Immunodeficiencies, Pediatrics Center of Excellence, Children's Medical Center, Tehran University of Medical Sciences, Tehran 1419733151, Iran; rezaei\_nima@yahoo.com

<sup>19</sup> Network of Immunity in Infection, Malignancy and Autoimmunity (NIIMA), Universal Scientific Education and Research Network (USERN), Tehran 1419733151, Iran

<sup>20</sup> UPMC Hillman Cancer Center, Department of Medicine, Division of Hematology/Oncology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213, USA; brufskyam@upmc.edu

\* Correspondence: lundstromkenneth@gmail.com

Received: 7 October 2020; Accepted: 20 October 2020; Published: 22 October 2020



**Abstract:** The origin of the severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2) virus causing the COVID-19 pandemic has not yet been fully determined. Despite the consensus about the SARS-CoV-2 origin from bat CoV RaTG13, discrepancy to host tropism to other human Coronaviruses exist. SARS-CoV-2 also possesses some differences in its S protein receptor-binding domain, glycan-binding N-terminal domain and the surface of the sialic acid-binding domain.

Despite similarities based on cryo-EM and biochemical studies, the SARS-CoV-2 shows higher stability and binding affinity to the ACE2 receptor. The SARS-CoV-2 does not appear to present a mutational “hot spot” as only the D614G mutation has been identified from clinical isolates. As laboratory manipulation is highly unlikely for the origin of SARS-CoV-2, the current possibilities comprise either natural selection in animal host before zoonotic transfer or natural selection in humans following zoonotic transfer. In the former case, despite SARS-CoV-2 and bat RaTG13 showing 96% identity some pangolin Coronaviruses exhibit very high similarity to particularly the receptor-binding domain of SARS-CoV-2. In the latter case, it can be hypothesized that the SARS-CoV-2 genome has adapted during human-to-human transmission and based on available data, the isolated SARS-CoV-2 genomes derive from a common origin. Before the origin of SARS-CoV-2 can be confirmed additional research is required

**Keywords:** coronavirus; COVID-19 pandemic; origin of SARS-CoV-2; receptor binding domain; genome homology; natural selection; zoonotic transfer

---

The COVID-19 pandemic has seriously touched the whole world with over 40 million infections and claiming more than 1 million lives as of today (21 October 2020), also causing social and economic havoc globally. There is currently a tremendous amount of both competitive and collaborative efforts in search of novel antiviral drugs and vaccines reaching almost desperate proportions. Among all this, the question of the origin of the “culprit”, the severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2), is being addressed. Numerous politically motivated conspiracy theories have surfaced, and hypotheses have arisen, including the unintentional or intentional escape/release of the virus from a high-security laboratory facility in Wuhan, China. For instance, it was reported [1] (now withdrawn) that SARS-CoV-2 contained four inserts in its spike (S) glycoprotein, critical for virus entry, which were either identical or similar to motifs found in the Env and Gag proteins of HIV-1. It was speculated that fragments from the HIV-1 genome had been intentionally introduced into the SARS-CoV-2 genome. A thorough bioinformatics analysis and sequence examination of SARS-CoV-2, other Coronaviruses, and HIV-1 from the GenBank database demonstrate that there is currently no compelling evidence that HIV-1 specific inserts in SARS-CoV-2 exist [2]. Additional analysis initially suggests that the probability is low that SARS-CoV-2 is a laboratory construct or intentionally engineered [3].

The self-assembled COVID consortium, consisting of international experts in bioinformatics, structural biology, molecular biology, immunology, and virology, has just published a Letter in the Journal of Medical Virology [4] in response to publications on the natural origin of SARS-CoV-2. It stated that despite the consensus of SARS-CoV-2 originating from bat CoV RaTG13, SARS-CoV-2 had demonstrated significant discrepancies to other human Coronaviruses related to host tropism. Moreover, bat and rodent Coronaviruses have seen some specific changes in the S protein receptor-binding domain (RBD) and the glycan-binding N-terminal domain (NTD) in host tropism [5,6]. However, SARS-CoV-2 sequences do not contain these changes, indicating a very recent origin of RBD and NTD subdomains. For instance, the hidden glycan-binding domains located in cavities in the S protein NTD domain, limiting their access to antibodies and immune cells, are not present in SARS-CoV-2 [6]. The surface of the sialic acid-binding domain of the SARS-CoV-2 S protein is flat and non-sunken, unlike other Coronaviruses, influenza viruses, rhinoviruses, and Meningo viruses showing “canyons”, depression zones, or cavities in accordance with the “Canyon hypothesis” [7].

Although previous cryo-EM structural and biochemical studies on furin-cleaved and native SARS-CoV-2 S protein and bat CoV RaGT13 S protein have indicated strong similarity, the native human S protein showed higher stability and a 1000-fold higher binding affinity to the human ACE2 receptor. It suggests that furin cleavage decreased the overall S protein stability and facilitated the open conformation required for viral particle binding to the ACE2 receptor [8]. Furthermore, it has been demonstrated that the D614G mutation in the SARS-CoV-2 S protein reduced S1 shedding and

increased infectivity [9]. In contrast to bat RaTG13, SARS-CoV-2 recombination presumably occurs between the S1 and S2 domains in the S protein enabling the utilization of furin protease. Despite the analysis of numerous clinical isolates of the SARS-CoV-2 S protein, no alternative recombination seems to occur, suggesting that the furin S1/S2 cleavage site is unique for recombination. Moreover, the four amino acid insertion, which creates a novel furin cleavage site, supports it. Although human Coronaviruses frequently contain “hot spots” for non-synonymous amino acid replacements affecting host tropism/adaptation, resistance to neutralizing antibodies and immune evasion [10], only a single high-frequency non-synonymous mutation (D614G) has been identified from clinical SARS-CoV-2 isolates [11]. Based on these findings, the SARS-CoV-2 S protein does not occur as a mutational “hot spot” in contrast to other human Coronaviruses.

So, the million-dollar question today is the origin of SARS-CoV-2. Why is this important? Perhaps a citation of Theodore Roosevelt is in place: “The more you know about the past the better you are prepared for the future”. Indeed, not only of scientific curiosity or finding someone to blame, but of practical awareness and preparation for potential emerging new outbreaks, the origin of SARS-CoV-2 is of utmost importance. As the current consensus within the scientific community strongly indicates, it is improbable (though not zero) that the SARS-CoV-2 emerged through laboratory manipulations. Two alternative hypotheses for the origin of SARS-CoV-2 have been presented [3]: natural selection in an animal host before zoonotic transfer or natural selection in humans following zoonotic transfer. In the earlier case, despite the high similarity between SARS-CoV-2 and bat Coronaviruses, such as bat Coronavirus RaTG13 (96% identical), there are significant discrepancies as we have pointed out. Moreover, some pangolin Coronaviruses exhibit high similarity to the RBD region in the SARS-CoV-2 S protein, including six RBD key residues [3,12]. In this latter case, the SARS-CoV-2 genome has been postulated to have adapted during undetected human-to-human transmission, which allowed the pandemic to accelerate [13]. It seems that sequence data accumulated from SARS-CoV-2 genomes so far indicate that the isolated SARS-CoV-2 genomes derived from a common ancestor. Moreover, the identification of a very similar RBD sequence in the pangolin Coronavirus S protein to the one found in SARS-CoV-2 S supports the susceptibility of transmission to humans.

In any case, additional research is required before we can confirm the origin of SARS-CoV-2. In addition to our analysis on the SARS-CoV-2 S protein, we have also targeted the SARS-CoV-2 ORF8 and ORF10 proteins in comparison to other coronaviruses to further shed some light on the origin of SARS-CoV-2, which will soon be shared with the scientific community.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Pradhan, P.; Pandey, A.K.; Mishra, A.; Gupta, P.; Tripathi, P.K.; Menon, M.B.; Gomes, J.; Vivekanandan, P.; Kundu, B. Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag. *bioRxiv* **2020**. [[CrossRef](#)]
2. Xiao, C.; Li, X.; Liu, S.; Sang, Y.; Gao, S.-J.; Gao, F. HIV-1 did not contribute to the 2019-nCoV genome. *Emerg. Microb. Inf.* **2020**, *9*, 378–381. [[CrossRef](#)] [[PubMed](#)]
3. Andersen, K.G.; Rambaut, A.; Lipkin, W.I.; Holmes, E.C.; Garry, R.F. The proximal origin of SARS-CoV-2. *Nat. Med.* **2020**, *26*, 450–452. [[CrossRef](#)] [[PubMed](#)]
4. Seyran, M.; Pizzol, D.; Adadi, P.; Mohamed Abd El-Aziz, T.; Hassan, S.S.; Soares, A.; Kandimalla, R.; Lundstrom, K.; Tambuwala, M.; Aljabali, A.A.A.; et al. Questions concerning the proximal origin of SARS-CoV-2. *J. Med. Virol.* **2020**. [[CrossRef](#)] [[PubMed](#)]
5. Hulswit, R.J.; de Haan, C.A.; Bosch, B.J. Coronavirus Spike Protein and Tropism Changes. *Adv. Virus Res.* **2016**, *96*, 29–57. [[CrossRef](#)] [[PubMed](#)]
6. Li, F. Receptor recognition mechanisms of coronaviruses: A decade of structural studies. *J. Virol.* **2015**, *89*, 1954–1964. [[CrossRef](#)] [[PubMed](#)]

7. Rossmann, M.G. The canyon hypothesis. Hiding the host cell receptor attachment site on a viral surface from immune surveillance. *J. Biol. Chem.* **1989**, *264*, 14587–14590. [[CrossRef](#)] [[PubMed](#)]
8. Wrobel, A.G.; Benton, D.J.; Xu, P.; Roustan, C.; Martin, S.R.; Rosenthal, P.B.; Skehel, J.J.; Gambin, S.J. SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nat. Struct. Mol. Biol.* **2020**, *27*, 763–767. [[CrossRef](#)] [[PubMed](#)]
9. Zhang, L.; Jackson, C.B.; Mou, H.; Ojha, A.; Rangarajan, E.S.; Izard, T.; Farzan, M.; Choe, H. The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. *bioRxiv* **2020**. [[CrossRef](#)]
10. Malaiyan, J.; Arumugam, S.; Mohan, K.; Radhakrishnan, G.G. An update on the origin of SARS-CoV-2: Despite closest identity, bat (RaTG13) and pangolin derived coronaviruses varied in the critical binding site and O-linked glycan residues. *J. Med. Virol.* **2020**. [[CrossRef](#)] [[PubMed](#)]
11. Brufsky, A. Distinct viral clades of SARS-CoV-2: Implications for modeling of viral spread. *J. Med. Virol.* **2020**. [[CrossRef](#)] [[PubMed](#)]
12. Zhang, T.; Wu, Q.; Zhang, Z. Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak. *Curr. Biol.* **2020**, *30*, 1346–1351. [[CrossRef](#)] [[PubMed](#)]
13. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.M.; Wang, W.; Song, Z.G.; Hu, Y.; Tao, Z.W.; Tian, J.H.; Pei, Y.Y.; et al. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *579*, 265–271. [[CrossRef](#)] [[PubMed](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).