Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Biochemical and Biophysical Research Communications 538 (2021) 88–91

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Evolution patterns of SARS-CoV-2: Snapshot on its genome variants

Marta Giovanetti [a, b, 1], Francesca Benedetti [c, 1], Giovanni Campisi [d], Alessandra Ciccozzi [b], Silvia Fabris [b], Giancarlo Ceccarelli [e], Vittoradolfo Tambone [f], Arnaldo Caruso [d], Silvia Angeletti [g], Davide Zella [c, **], Massimo Ciccozzi [b, *]

[a] Laboratório de Flavivírus, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil
[b] Medical Statistic and Molecular Epidemiology Unit, University of Biomedical Campus, Rome, Italy
[c] Institute of Human Virology, Department of Biochemistry and Molecular Biology, School of Medicine, University of Maryland Baltimore, USA
[d] Department of Molecular and Translational Medicine, Section of Microbiology, Universityof Brescia, Brescia, Italy
[e] Dipartimento di Salute Pubblica e Malattie Infettive, Policlinico Umberto I Universita'Sapienza', Rome, Italy
[f] Anthropology, and Applied Ethics, Campus Bio-Medico University, Rome, Italy
[g] Unit of Clinical Laboratory Science, University Campus Bio-Medico of Rome, Rome, Italy

ARTICLE INFO

Article history:
Received 13 October 2020
Received in revised form
20 October 2020
Accepted 21 October 2020
Available online 6 November 2020

Keywords:
SARS-CoV-2
Pandemic
Mutation
Evolution pattern

ABSTRACT

An acute respiratory syndrome (COVID-19), caused by a novel coronavirus (SARS-CoV-2) with a high rate of morbidity and elevate mortality, has emerged as one of the most important threats to humankind in the last centuries. Rigorous determination of SARS-CoV-2 infectivity is very difficult owing to the continuous evolution of the virus, with its single nucleotide polymorphism (SNP) variants and many lineages. However, it is urgently necessary to study the virus in depth, to understand the mechanism of its pathogenicity and virulence, and to develop effective therapeutic strategies. The present contribution summarizes in a succinct way the current knowledge on the evolutionary and structural features of the virus, with the aim of clarifying its mutational pattern and its possible role in the ongoing pandemic.
© 2020 Elsevier Inc. All rights reserved.

1. Introduction

The respiratory syndrome responsible for the current pandemic initially detected in Wuhan in late December 2019 (Na Z et al., 2020) is an infectious disease caused by a novel coronavirus (SARS-CoV-2). Coronaviruses (CoVs) are a group of enveloped viruses, with a positive single-stranded RNA genome of approximately 30,000 bases with 5′-cap structure and 3′-poly-A tail, belonging to the Coronaviridae family of the order Nidovirales [1]. They cause mainly respiratory and gastrointestinal tract infections and are genetically classified into four major genera: Alphacoronavirus, Betacoronavirus, Gammacoronavirus, and Deltacoronavirus. The Alpha and Beta-CoVs can infect humans, while the Gama and Delta-CoVs infect predominantly birds [2].

The SARS-CoV-2, which is responsible for the current Coronavirus disease (COVID-19), also belongs to the genus Beta-CoV and it is considered the third major coronavirus outbreak in the last 20 years, after Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) [2]. On March 11th, 2020, the World Health Organization (WHO) having established the spread (and severity) of the SARS-CoV-2 infection, declared that the COVID-19 outbreak recorded in the preceding months was a pandemic [3]. It has currently affected >200 countries. As on October 2020, about 35.6 million people have been infected, with more than 1.04 million deaths [4]. More than 24.8 million people have recovered completely, but a large number of the infected people end up in critical condition that require respiratory assistance. The Countries that have been affected most severely are the USA, Brazil, India, Russia, Mexico, South American Countries, most European Countries [4].

Adaptive mutations in the SARS-CoV-2 genome could alter its pathogenic potential, and at the same time would increase the difficulty of drug and vaccine development. This contribution will not deal in detail with the mass of molecular information now available for SARS-CoV-2. It will rapidly summarize the information on its evolutionary and structural features that could be useful for the development of vaccines.

* Corresponding author.
** Corresponding author.
E-mail address: M.ciccozzi@unicampus.it (M. Ciccozzi).
[1] These authors contributed equally to this article.

https://doi.org/10.1016/j.bbrc.2020.10.102
0006-291X/© 2020 Elsevier Inc. All rights reserved.

1. SARS-CoV-2 genome composition, proximal origin and evolution patterns

Genomes of Coronaviruses include a variable number of open reading frames (ORF). The first 5′ ORF (ORF1a/b) corresponds to about 2/3 of genome, and it is translated in the rough endoplasmic reticulum of the host cell into pp1a and pp1ab protein that are cleaved by proteases it yields 16 non-structural proteins (*nsp1-16*) [5]. The 3′ ORF, which corresponds to the remaining third of the genome consists of genes encoding accessory and structural proteins [5]. The four major structural proteins are: the surface Spike (S) protein, which recognizes the receptor of the host cell (the angiotensin converting enzyme 2 (ACE2)), binds to it, and mediates the penetration of the virus into the host cell the envelope E protein, the matrix protein (M), and the nucleocapsid (N) protein that binds the RNA and is fundamental for virion assembly [1,5] (Fig. 1). Additionally, the SARS-CoV-2 contains 6 accessory proteins, encoded by the ORF3a, ORF6, ORF7a, and ORF8, ORF10 genes: their functions are still largely unexplored. Thus, most of the proteins encoded by ORF1a and ORF1ab are essential for virus replication and at least for the adaption of the virus to a new host. In addition, a 5′ untranslated region (UTR) and a 3′-UTR have also been identified in the SARS-CoV-2 genome [1,5]. Some of the *nsp* form the replicase/transcriptase complex: *nsp* 12 is the RNA-dependent RNA polymerase that replicates the RNA of the virus, however, to function perfectly it also needs *nsp* 7 and *nsp* 8 and, possibly, of other non-structural proteins. The genomic organization of the SARS-CoV-2 shares about 89% sequence identity with that of other CoVs. Comparative sequence analysis of the SARS-CoV-2 genome indicates striking similarities to that of the Bat-CoV, suggesting a possible bat origin for that in the affected humans in Wuhan [6]. The possibility of other animal intermediate hosts from bats before the introduction of the virus into humans has been discussed. Initially, the Pangolin origin of COVID-19 in humans had received significant favour, particularly based on the finding that the virus of pangolins could use ACE2 as the receptor in the host cells. However, more recent results of various experimental approaches, e.g., on the very poor affinity of the pangolin virus for the human ACE2 receptor have apparently exempted pangolins from a possible role as intermediate host. The question of the intermediate host in the transmission of SARS-CoV-2 from bats to human is thus not settled, and it appears possible that the transmission occurs directly [7] (Fig. 2).

Generally, the rates of nucleotide substitution of RNA viruses are fast, and this rapid evolution is mainly shaped by natural selection. This high error rate and the consequent rapidly evolving virus populations [8], which could lead to the accumulation of amino acid mutations, might affect the transmissibility of the virus, its cell tropism and pathogenicity. It would unfortunately also present daunting challenges for the design of effective vaccines and diagnostic means. Fortunately, however, until now the observed diversity among SARS-CoV-2 sequences has been low. There has been an exception that has had dire consequences, i.e., the replacement of aspartic acid 614 of the Spike protein with a glycine, which has greatly increased the infectivity of the virus, but, in principle, the possibility of positive natural selection mutations exists. Considering its high transmissibility and the absence of pre-existing immunity in the general population its natural disappearance appears to be unlikely. Furthermore, it is not known whether SARS-CoV-2 is already fully adapted for efficient growth in human cells after its host-jump from bats or from a putative intermediate host [9]. Genomic epidemiology has revealed that the spillover from bats to humans has most likely occurred in late November or early December 2019. and that from that moment the spread of the virus occurred mainly by a human-to-human transmission [6].

2. Recent genomic changes in SARS-CoV-2

Coronaviruses such as SARS-CoV-2 are relatively stable thanks to a proofreading mechanism that operates during replication. Many genomic studies have nevertheless revealed changes in their genomes, including mutations and deletions. The D614G point-mutation in the Spike protein of SARS-CoV-2, which rapidly became the most widespread variant of SARS-CoV-2, has just been mentioned: we were among the first to observe it [10–12], but we had also observed that this mutation clustered with a series of other point mutations, including one in the polymerase gene [10]. A series of other mutations were then identified, allowing the classification of several SARS-CoV-2 lineages [13]. At the same time, additional profound changes in the genome, i.e. deletions, started to be reported. In particular, an extensive deletion in the ORF7a gene [14] and a deletion in the *nsp*2 gene [15]. More recently, analysing a more comprehensive dataset of more than 17.000 sequences obtained from GISAID we identified the emergence of a strain with a deletion of 9 nucleotides in the *nsp1* gene (nucleotides 686–694 corresponding to amino acids 241–243) in patients infected with COVID-19 from different areas of the world. The overall frequency of the genome deletion was 0.44%, but was not distributed homogeneously: for instance, we did not find it in Italy, Germany, and Austria, whereas it was more frequent in Sweden, Israel, and the USA. Structural analysis suggest that this deletion might affect the C-terminal region of the protein that appear to be important for the regulation of viral replication and appear to have negative effect on host's gene expression [16]. Those results were also confirmed by other groups that highlighted that SARS-CoV-2 appear to be undergoing profound genomic changes [17,18].

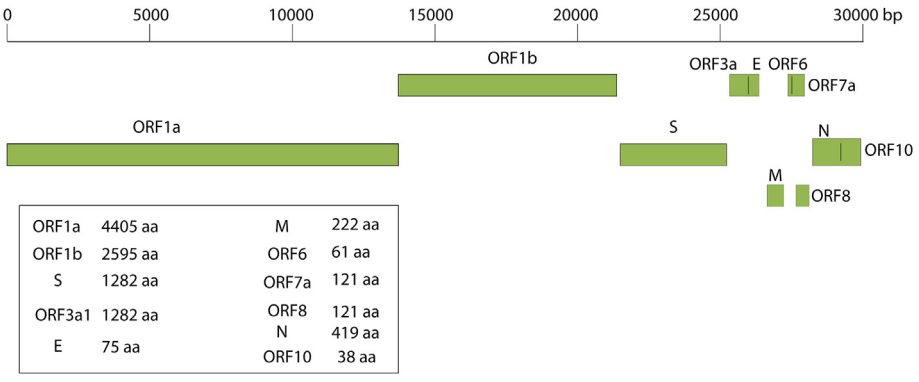

Fig. 1. Genome architecture of SARS-CoV-2.

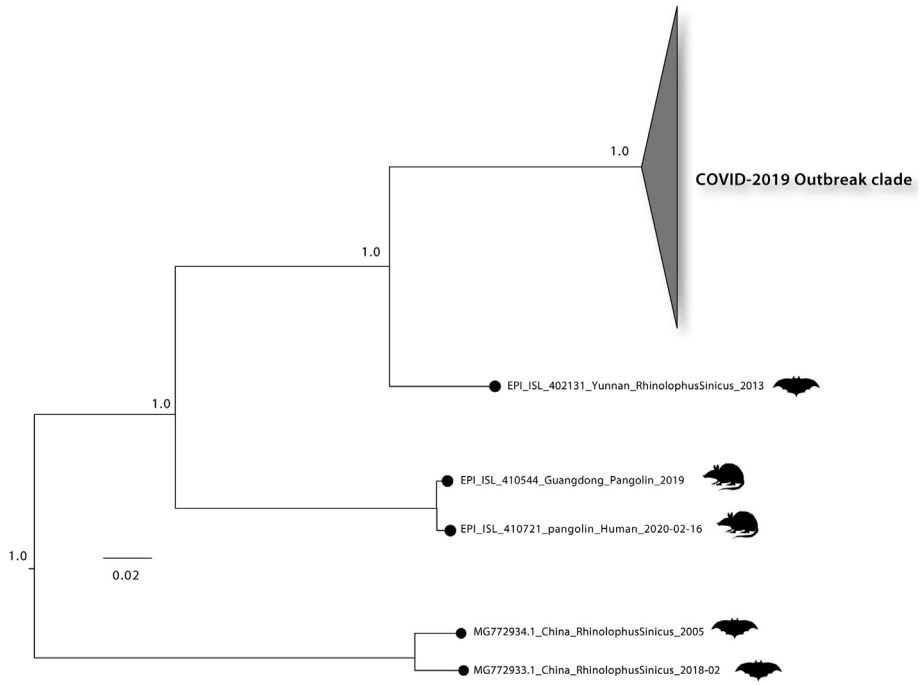

Fig. 2. Maximum likelihood phylogeny was estimated with n = 143 complete genomes sequences from the current (2019–2020) SARS-CoV-2 epidemic plus n = 3 closely related bat strains plus n = 2 pangolin strains retrieved from GISAID. Number along branch represent bootstrap score. Scale bar represents expected substitutions per nucleotide site.

While the D614G mutation confers a selective advantage for SARS-CoV-2 fitness [11,12], the exact biological relevance of the other mutations is still unknown. However, *nsp 1*, which is also known as the leader protein, is central in the inhibition of the antiviral innate immune response, in particular the expression of interferon-alpha [19] and is possibly the most important determinant of the viral pathogenicity. We feel it is appropriate to mention that a viral isolate from an asymptomatic SARS-CoV-2-positive subject had an unprecedented replication ability in VeroE6 cells in the absence of any clear cytopathic effect [9]. Even though it was a single observation, and even if the precise molecular mechanism that explains the absence of cytopathic effect has not yet been identified, we believe that the observation is significant. Such results could indicate the evolution of a possible new viral quasispecie, but further data will be necessary in order to confirm this hypothesis. On this respect we believe that priority in this moment will be the monitoring over time of asymptomatic and paucisymptomatic subjects to confirm the spreading of this particular viral strain with a possible decreased viral pathogenicity.

3. Parameters on viral mutations and spreading and their effect

Environmental factors, e.g., temperature, population density and air pollution, seem to affect viral spreading and mortality rate [20–24]. In addition, interventions aimed at limiting people movement and interactions have been implemented in several areas of the globe to curb the pandemic. The overall effect of these measures has been a reduction of the number of infections and the decrease of death rates. However, while it is intuitively clear that, for instance, full lockdowns by themselves are very effective, they are economically so burdensome to become impractical and unsustainable in the long run. Consequently, it becomes important to implement a series of concomitant and complementary measures to limit viral spread. One important point to be considered concerns the virus itself: at the moment there are no reasons to believe that changes of the virus have occurred in the direction of decreased pathogenicity. The decrease in the seriousness of the infection which is now generally observed is prima facie related to the containment measures: they certainly influence one parameter that is important, i.e. de magnitude of the viral load which is transmitted, thus in a sense they alter, quantitatively if not qualitatively the virus. Whether the measures above also have an effect on the viability of the viral particles has been suggested as a possibility [25,26]. But, as we have discussed above, the chances that the virus itself may change (i.e., mutate) in the direction of decreased pathogenicity could still be considered.

4. Conclusions

The COVID-19 pandemic has stressed our health care systems in an unprecedented way and underlined once more the important role of the molecular evolution succinctly described in this review. Within a few days from the first reported cases of anomalous pneumonia, significant progress was made in the fight against it: the virus was isolated, sequenced, identified and genetically characterized. It was named SARS-CoV-2 because of its phylogenetic relationship with SARS-CoV and bat SARS-like coronaviruses. Based on its genetic features, molecular and serological assays were developed and have been introduced in routine diagnostics. Pharmacological means have been gradually discovered and introduced, and general vaccine strategies have been developed or are in development. Trials are ongoing or are about to start to determine their effectiveness. This contribution has pictured the current research on the molecular evolution of the SARS-CoV-2 after its epidemic outbreak. Phylogenetic analysis and homology modeling have added knowledge to the fine details of the virus, and so have done the studies exploring the genome of the virus and the structure of its proteins. The search for viral variants with decreased or no pathogenic potential would be a significant step.

Funding

MG is supported by Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] F. Wu, S. Zhao, B. Yu, Y.M. Chen, W. Wang, Z.G. Song, Y. Hu, Z.W. Tao, J.H. Tian, Y.Y. Pei, M.L. Yuan, Y.L. Zhang, F.H. Dai, Y. Liu, Q.M. Wang, J.J. Zheng, L. Xu, E.C. Holmes, Y.Z. Zhang, A new coronavirus associated with human respiratory disease in China, Nature 10 (10) (2020) 25–38.
[2] J.F. Chan, K.H. Kok, Z. Zhu, H. Chu, K.K. To, S. Yuan, K.Y. Yuen, Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan, Emerg. Microb. Infect. 28 (2) (2020) 21–236.
[3] Organization WHO, Coronavirus disease (COVID-19) pandemic, 2020. Accessed 2 Jul, 2020.
[4] Organization WHO, Coronavirus Disease (COVID-19) Report, 2020.
[5] R.A. Khailany, M. Safdar, M. Ozaslan, Genomic characterization of a novel SARS- CoV-2, Gene Reports 16 (19) (2020) 100–682.
[6] D. Benvenuto, M. Giovanetti, M. Ciccozzi, S. Spoto, S. Angeletti, M. Ciccozzi, The 2019- new coronavirus epidemic: evidence for virus evolution, J. Med. Virol. 9 (4) (2020) 455–459.
[7] C. Siew Woh, Are pangolins scapegoats of the COVID-19 outbreak-CoV transmission and pathology evidence? J of the Society for conservation Biology (8) (2020) 9–19, 02.
[8] A. Moya, S.F. Elena, A. Bracho, R. Miralles, E. Barrio, The evolution of RNA viruses: a population genetics view, Proc. Natl. Acad. Sci. U.S.A. 97 (2) (2000) 6967–6973.
[9] F. Caccuri, A persistently replicating SARS-CoV-2 variant derived from an asymptomatic individual, J. Transl. Med. 2 (9) (2020) 13–98.
[10] M. Pachetti M, Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant, J. Transl. Med. 2 (18) (2020) 1–79.
[11] B. Korber, Spike mutation pipeline reveals the emergence of a more transmissible form of SARS-CoV-2, bioRxiv (2020), 04.29.069054.
[12] B. Korber, Tracking changes in SARS-CoV-2 Spike: evidence that D614G increases infectivity of the COVID-19 virus, Cell 182 (7) (2020) 812–827.
[13] K.G. Andersen, A. Rambaut, W.I. Lipkin, E.C. Holmes, R.F. Garry, The proximal origin of SARS-CoV-2, Nat. Med. 26 (7) (2020) 450–452.
[14] L.A. Holland, An 81-nucleotide deletion in SARS-CoV-2 ORF7a identified from sentinel surveillance in Arizona (january to March 2020), J. Virol. 20 (9) (2020) 84–94.
[15] A. Bal, Molecular characterization of SARS-CoV-2 in the first COVID-19 cluster in France reveals an amino acid deletion in nsp2 (Asp268del), Clin. Microbiol. Infect. 26 (1) (2020) 960–962.
[16] F. Benedetti, G.A. Snyder, M. Giovanetti, S. Angeletti, R.C. Gallo, M. Ciccozzi, D. Zella, Emerging of a SARS-CoV-2 viral strain with a deletion in nsp1, J. Transl. Med. 18 (11) (2020) 3–29.
[17] T. Phan, Genetic diversity and evolution of SARS-CoV-2, Infect. Genet. Evol. 81 (11) (2020) 104–260.
[18] M.R. Islam, Genome-wide analysis of SARS-CoV-2 virus strains circulating worldwide implicates heterogeneity, Sci. Rep. 10 (12) (2020) 14–40.
[19] A.R. Jauregui, D. Savalia, V.K. Lowry, C.M. Farrell, M.G. Wathelet, Identification of residues of SARS-CoV nsp1 that differentially affect inhibition of gene expression and antiviral signaling, PloS One 29 (4) (2013) 62–416.
[20] M. Sajadi, Temperature and latitude analysis to predict potential spread and seasonality for COVID-19, SSRN Electronic Journal 2 (33) (2020) 10–20.
[21] T. Carleton, K.C. Meng, Causal empirical estimates suggest COVID-19 transmission rates are highly seasonal, medRxiv (2020), 03.26.20044420.
[22] G.F. Ficetola, D. Rubolini, Climate affects global patterns of COVID-19 early outbreak dynamics, MedRxiv (2020), 03.23.20040501.
[23] X. Wu, Exposure to air pollution and COVID-19 mortality in the United States: a nationwide cross-sectional study, medRxiv (2020), 04.05.20054502.
[24] F. Benedetti, M. Pachetti, B. Marini, R. Ippodrino, R.C. Gallo, M. Ciccozzi, D. Zella, Inverse correlation between average monthly high temperatures and COVID-19-related death rates in different geographical areas, J. Transl. Med. 18 (22) (2020) 22–51.
[25] M. Gandhi, G.W. Rutherford, Facial masking for covid-19 — potential for "variolation" as we await a vaccine, N. Engl. J. Med. 9 (1) (2020) 8–10.
[26] Y. Abba, Antiviral activity of resveratrol against human and animal viruses, Adv Virol 18 (12) (2005), 4-4.