



Microbes and Infectious Diseases

Journal homepage: <https://mid.journals.ekb.eg/>

Mini-review article

Structure and genome of severe acute respiratory syndrome coronavirus 2

Ibrahim Mohammed Hussaini ^{*1}, Musa Bishir ², Suleiman Gide ³, Sheriff Wakil ⁴, Anas Gide ⁵

1- Department of Microbiology, Faculty of Life Sciences, Ahmadu Bello University, Zaria Nigeria.

2- Department of Microbiology, Ahmadu Bello University Zaria, Nigeria.

3- Desert Research Centre, Yobe State University, Damaturu, Nigeria.

4- Department of Microbiology, Yobe State University, Damaturu, Nigeria.

5- Nigerian Institute for Trypanosomiasis Research, Sokoto State, Nigeria.

ARTICLE INFO

Article history:

Received 13 April 2022

Received in revised form 12 May 2022

Accepted 15 May 2022

Keywords:

SARS-CoV-2

COVID-19

Genome

Structure

Coronavirus

ABSTRACT

SARS-CoV-2 is a novel coronavirus responsible for the current COVID-19 pandemic. It is the seventh coronavirus known to infect humans (the previous human coronaviruses are HCoV-OC43, HCoV-229E, HCoV-HKU1, HCoV-NL63, SARS-CoV and MERS-CoV) and the third human coronavirus known to cause severe illness in human after SARS-CoV and MERS-CoV. These three coronaviruses have caused three different severe respiratory diseases outbreaks within the last two decades: SARS in 2002-2003, MERS in 2012 and COVID-19 in 2020. The aim of this review was to summarize information on the genome and structure of SARS-CoV-2. SARS-CoV-2 is an enveloped positive-sense single-stranded RNA virus with a crown-like appearance due to the presence of spike glycoprotein on the envelope. The nonsegmented genome of SARS-CoV-2 of approximately 30kb encodes two large polyproteins, four main structural proteins namely spike, membrane, envelope and nucleocapsid proteins as well as several accessory proteins. Analysis of SARS-CoV-2 genome shows that it is highly related to coronavirus from the bat (96%), pangolin (91%) and SARS-CoV (80%). Variants of SARS-CoV-2 have evolved continuously as a result of genetic mutations and are circulating worldwide. These variants have varying degrees of transmissibility, disease severity, susceptibility to therapeutics and detection by diagnostic tools. Understanding the structure and genome of SARS-CoV-2 is important in the control, management, diagnosis and treatment of COVID-19 as well as vaccine development.

Introduction

Coronaviruses belong to the family Coronaviridae, subfamily Coronavirinae and order Nidovirales. There are four genera in the Coronavirinae subfamily namely Alphacoronavirus, Betacoronavirus, Gammacoronavirus, and Deltacoronavirus [1].

Coronaviruses are spherical, enveloped, positive-sense single-stranded RNA viruses with helical nucleocapsid surrounded nonsegmented genome [2]. They can infect the respiratory,

gastrointestinal, hepatic and central nervous system of both humans and animals [1, 3]. Coronaviruses of medical importance were first isolated from respiratory secretion of a common cold patient in the 1960s [2].

In December 2019, a severe respiratory disease outbreak caused by a novel coronavirus was reported in Wuhan city, Hubei province of China [4]. On 11th of February 2020, this severe respiratory disease was named coronavirus disease

2019 (COVID-19) by World Health Organization [5] and the novel coronavirus was named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) by the International Committee on Taxonomy of Viruses (ICTV) [6]. COVID-19 was listed as Public Health Emergency of International Concern (PHEIC) on the 31st of January, 2020 [7] and declared as pandemic on the 11th on March 2020 by WHO [8].

SARS-CoV-2 is the seventh coronavirus known to cause respiratory infections in humans. Two of these human coronaviruses HCoV-229E (named after sample code of the student 229E) and HCoV-NL63 (Netherland 63) are alphacoronaviruses (α -CoVs) while five namely; HCoV-OC43 (organ culture 43), HCoV-HKU1 (Hong Kong University 1), SARS-CoV (Severe Acute Respiratory Syndrome Coronavirus), MERS-CoV (Middle East Respiratory Syndrome Coronavirus) and SARS-CoV-2 are betacoronaviruses (β -CoVs) [9]. Mild infections are caused by HCoV-OC43, HCoV-229E, HCoV-HKU1 and HCoV-NL63 while severe respiratory diseases are caused by SARS-CoV, MERS-CoV and SARS-CoV-2 [3]. Within the last two decades, severe respiratory disease outbreaks caused by these three coronaviruses were recorded. Severe Acute Respiratory Syndrome (SARS) outbreak caused by SARS-CoV was reported in 2002 in China and spreads to other countries [10]. A decade later (2012), Middle East respiratory syndrome (MERS) caused by MERS-CoV was reported in Saudi Arabia [11] and late 2019; COVID-19 caused by SARS-CoV-2 was reported in China.

Structure of SARS-CoV-2

SARS-CoV-2 is a spherical or moderately pleomorphic (~60 to 140 nm), positive-sense single-stranded enveloped RNA (+ssRNA) virus [12]. The spike glycoprotein on the membrane of the virus forms peplomers and gives it a crown-like appearance hence the name corona. A ring structure is formed by the membrane glycoprotein and the envelope protein while the helical nucleocapsid lies within the virus. The viral nucleocapsid is made up of nucleocapsid protein complexed with the +ssRNA [13].

The spike glycoprotein consists of: a large ecto domain, a single-pass transmembrane anchor, and a short C-terminal intracellular tail. The ecto domain is made up of a receptor binding unit S1 and a membrane-fusion unit S2 [14].

The receptor-binding domain (RBD) found in the S1 subunit of spike proteins of SARS-CoV-2 mediates entry in the host cell by binding to the angiotensin-converting enzyme 2 (ACE2) receptor on the epithelial cells of the lungs and small intestines [3].

Genome of SARS-CoV-2

The genome of SARS-CoV-2 was isolated and sequenced for the first in January 2020. This genome was named Wuhan-Hu-1 with NCBI reference sequence of NC_045512 [12, 13].

Genes responsible for replication and pathogenesis of SARS-CoV-2 are contained in the 5'-end of the genome while genes that code for nucleocapsid and membrane proteins are found in the 3'-end of the genome [15, 16].

SARS-CoV-2 poses a positive-sense single-stranded RNA genome of approximately 30 kb comprising a 5' terminal cap structure and a 3' poly A tail [3]. The viral genome contains 10 - 14 open reading frames (ORFs). The first ORFs (ORF1a/b) found in the 5'-end of the viral genome are the largest and make up about two-third of SARS-CoV-2 genome. ORF1a/b encode two large polyproteins (pp1a/pp1b) that contain 15 non-structural proteins (nsps): nsp1-nsp10 and nsp12-nsp16. The remaining ORFs found in the 3'-end of the genome encode majorly four structural proteins (spike, envelope, membrane and nucleocapsid proteins) and accessory proteins (3a, 3b, 6, 7a, 7b, 8 and 9b) as illustrated in **figure (2)** [17].

The nsps are assembled into the replicase-transcriptase complex which generates anti-sense (-) genomic that serves as a template for positive-sense genome (+) and sub-genomic RNAs which serves as a template for sub-genomic mRNA synthesis [9]. Most of the nsps play important roles in replication and pathogenesis of SARS-CoV-2 such as Cellular mRNA degradation, blocking of host immune response, polypeptides cleaving, dimerization and RNA binding [1].

The structural proteins are essential in assembly and infection of SARS-CoV-2. The spike protein is essential in binding to receptors on host cells while the membrane protein is essential in shaping the virion and binds to the nucleocapsid. Assembly and release of the virion is mediated by the envelope protein while the two domains of the nucleocapsid protein bind the viral genome [1].

Understanding the structure and genome of SARS-CoV-2 is vital as it is important in the

development of vaccine and other therapeutics. Since the spike glycoprotein mediates the entry of the virus into the host cells, it can be exploited to prevent viral entry. The spike protein of SARS-CoV-2 has also served as target for vaccines such as the DNA vaccine [18]. Design and development of direct acting antiviral therapy against SARS-CoV-2 requires understanding of its genome. Hence an insight into the structure and genome of SARS-CoV-2 is vital in the control, management and treatment of COVID-19.

Nucleotide identity analysis revealed that the genome of SARS-CoV-2 is similar to the genome of other coronaviruses. SARS-CoV-2 genome shows similarity of approximately 96% to the bat coronavirus BatCoV RaTH13 [3], 91% to PangolinCoV [19], 80% to SARS-CoV [12] and 50% with MERS-CoV [14, 20]. SARSCoV-2 and the other members of Betacoronavirus (MERS-CoV and SARS-CoV) mutate at high rate hence they are genetically diverse and can adapt to infect wide host range [21].

Variants of SARS-CoV-2

Changes that occur in the genetic code of SARS-CoV-2 during replication of its genome (i.e. genetic mutations) result in the continuous evolution of SARS-CoV-2 [22]. Some of these changes may affect the transmissibility of the virus, severity of clinical disease, susceptibility of the virus to vaccines, Emergency Use Authorization (EUA) monoclonal and therapeutic or performance of diagnostic tools [23].

Some of these genetic mutations confer competitive advantage to the variants by altering various aspects of the virus biology such as viral antigenicity, infectivity or pathogenicity. Mutations that affect antigenic phenotype of the virus will enable the variant evade immunity conferred naturally or through vaccination as well as affect the degree of immune recognition [24].

Mutation that occurs in areas critical for primer or antibody binding increases the rate of false negative results for RT-PCR and immunoassays [25]. Data from the global tracking indicate that

variant with G614 mutation spread faster and is more infectious than variant with D614 mutation [26].

Variants of SARS-CoV-2 (with each variant containing one or more mutations that differentiate it from other variants) have emerged and are circulating globally [22]. Four classes of SARS-CoV-2 variants are defined by the SARS-CoV-2 Interagency Group (SIG) as follows [22]:

a. Variant Being Monitored (VBM)

Alpha, Beta, Gamma, Epsilon, Eta, Iota, Kappa, Mu and Zeta.

This class include variants whose data indicates a potential or clear impact on approved or authorized therapeutics or variants associated with severe disease or those with increased transmissibility but are no longer detected, or are circulating at very low levels [22].

b. Variant of Interest (VOI)

There is no currently circulating VOI as at 6th April, 2022.

Included in this class are variants with specific genetic markers that are associated with increased transmissibility, increased disease severity, reduced neutralization by antibodies, reduced treatment and diagnostic efficacy [22].

c. Variant of Concern (VOC) Delta and Omicron.

Variants with evidence of: increased transmissibility, more severe disease, significant reduction in antibody neutralization, reduced treatment and vaccine effectiveness or increased diagnostic detection failure are included in this class [22].

d. Variant of High Consequence (VOHC)

There is no currently circulating VOHC as at 6th April, 2022

These include variants that have clear evidence of significantly reduced efficacy of prevention measures or medical countermeasures (MCMs) relative to previously circulating variants. They have significantly reduced susceptibility to multiple EUA and approved therapeutics, More severe disease outcome and increased hospitalizations [22].

Figure 1. Structure of SARS-CoV-2 virion. Source: **Ortiz-Prado et al.** [3].

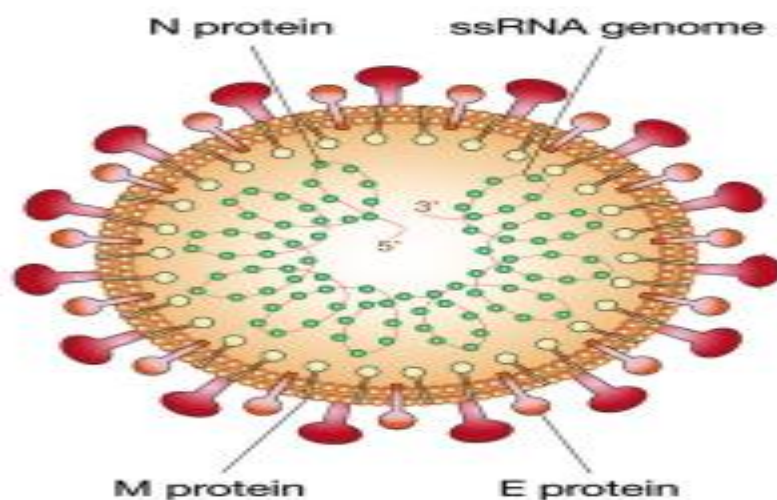


Figure 2. Structural organization of SARS-CoV-2 genome.



References

- 1-Chen Y, Liu Q, Guo D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. *Journal Medical Virology* 2020; 92:418-423.
- 2-Saif LJ, Wang Q, Vlasova AN, Jung K, Xiao S. *Coronaviruses*. In: Zimmerman JJ, Karriker LA, Ramirez A, Schwartz KJ, Stevenson GW, Zhang J, editors. *Diseases in swine*. 7th ed. Chichester: John Wiley & Sons 2019; Ltd; p. 488–523.
- 3-Ortiz-Prado E, Simbaña-Rivera K, Gómez-Barreno L, Rubio-Neira M, Guaman LP, Kyriakidis NC, et al. Clinical, molecular, and epidemiological characterization of the SARS-CoV-2 virus and the Coronavirus Disease 2019 (COVID-19), a comprehensive literature review. *Diagnostic Microbiology and Infectious Disease* 2020; 98:115094.
- 4-Lorusso A, Calistria P, Mercantea MT, Monaco F, Portantia O, Marcaccia M, et al. “One-Health” approach for diagnosis and molecular characterization of SARS-CoV-2 in Italy *One Health* 2020; 10:100135.
- 5-World Health Organization. Naming the coronavirus disease (COVID-19) and the virus that causes it. World Health Organization 2020a. Accessed August 6th, 2020).
- 6-World Health Organization. Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases. 2020b. Accessed August 6th, 2020.

- 7-Li X, Geng M, Peng Y, Meng L, Lu S.** Molecular immune pathogenesis and diagnosis of COVID-19. *Journal of Pharmaceutical Analysis* 2020; 10:102-108.
- 8-World Health Organization.** Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). WHO, Geneva, Switzerland. 2020c Accessed August 6th , 2020).
- 9-Neerukonda SN, Katneni U.** A Review on SARS-CoV-2 Virology, Pathophysiology, Animal Models, and Anti-Viral Interventions. *Pathogens* 2020; 9:426.
- 10-Fung TS, Liu DX.** Human Coronavirus: Host-Pathogen Interaction. *Annual Review of Microbiology* 2019; 73:529–557.
- 11-de Groot RJ, Baker SC, Baric RS, Brown CS, Drosten C, Enjuanes L, et al.** Middle East respiratory syndrome coronavirus (MERS-CoV): announcement of the Coronavirus Study Group. *Journal of Virology* 2013; 87(14):7790-7792.
- 12-Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q.** Structural basis for the recognition of the SARS-CoV-2 by full-length human ACE2. *Science* 2020; 367(6485):1444-1448.
- 13-Gralinski LE, Menachery VD.** Return of the coronavirus: 2019-nCoV. *Viruses* 2020; 12:135.
- 14-Su YCF, Anderson DE, Young BE, Linster M, Zhu F, Jayakumar J, et al.** Discovery of a 382-nt deletion during the early evolution of SARS-CoV-2. *mBio* 2020;11(4)
- 15-Yoshimoto FK.** The proteins of severe acute respiratory syndrome coronavirus-2 (SARS CoV-2 or n-COV19), the cause of COVID-19. *Protein Journal* 2020; 39:198–216.
- 16-Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, et al.** Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* 2020; 27:325-328.
- 17-Yang Y, Xiao Z, Ye K, He X, Sun B, Qin Z, et al.** SARS-CoV-2: characteristics and current advances in research. *Virology Journal* 2020; 17:117.
- 18-Yu J, Tostanoski LH, Peter L, Mercado NB, McMahan K, Mahrokhian SH, et al.,** DNA vaccine protection against SAR-CoV-2 in rhesus macaques. *Science* 2020; 369(6505): 806-811.
- 19-Zhang T, Wu Q, Zhang Z.** Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Current Biology* 2020; 30:1346-1351.
- 20-Wu C, Liu Y, Yang Y, Zhang P, Zhong W, Wang Y, et al.** Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharmaceutica Sinica B* 2020; 10(5):766-788.
- 21-Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D.** Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* 2020; 181(2):281-292.
- 22-Centers for Disease Control and Prevention (2021).** SARS-CoV-2 Variant Classifications and Definitions. Available at: www.cdc/coronavirus/2019-ncov/variants/variant-classifications.html
- 23-World Health Organization.** Tracking SARS-CoV-2 variants. 2020. Available at: <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>
- 24-Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al.** SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews Microbiology* 2021; 19: 409-424

25-Ascoli CA. Could mutations of SARS-Cov-2 suppress diagnostic detection? *Nature Biotechnology* 2021; 39: 274-275

26-Akkiz H. Implications of the Novel mutations in the SARS-CoV-2 Genome for Transmission, Disease severity, and the vaccine Development. *Frontier in Medicine* 2021; 7(8): 636532.

Hussaini IM, Bishir M, Gide S, Wakil S, Gide A. Structure and genome of severe acute respiratory syndrome coronavirus 2. *Microbes Infect Dis* 2022; 3(3): 516-521.