Virology and Epidemiology of the 2019 Coronavirus Disease (Covid - 19)

Abba Musa Abdullahi

Neurology/Neuroscience, Hasiya Bayero Paediatric Hospital, Kano, NGA

Corresponding author: Abdullahi AM, MBBS, MPH, Neurology/Neuroscience, Hasiya Bayero Paediatric Hospital, Kano, NGA

Received date: 20 August 2020; Accepted date: 28 August 2020; Published date: 03 September 2020


Copyright: © 2020 Abdullahi AM. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract
Coronavirus disease of the 2019 (COVID-19) is a contagious disease with a varying range of severity ranging from mild to moderate, severe, and critical. It has devastating consequences on the healthcare system, economy, and political system making it a public health emergency. The viral aetiological agent was named by the International Committee on Taxonomy of Virus (ICTV) as SARS-CoV-2. It was first discovered in the Wuhan city of Hubei province, China in mid-December from where it spread to every continent on earth. The viral envelope is nearly spherical with crown-like surface projections known as spikes which aid in viral attachment to the hosts’ tissues. It has a single-stranded RNA genome with many nucleotides bases encoding thousands of amino acids with two adjoining untranslated regions (UTRs) and one long open reading frame (ORF) encoding polyproteins. The main route of transmission is a human-to-human transmission via direct contact or via respiratory droplets spread by coughing, sneezing, or talking from an infected individual with an average incubation period of 4-6 days, however, 14 days incubation period was suggested by experts especially during the quarantine. As of June 2, 2020, the global confirmed cases were reported by the European Centre for Disease Control and Prevention (ECDC) as 6,245,352 cases with global death was recorded as 376,427 deaths.

Keywords: coronavirus, COVID-19, SARS-CoV-2, virology, genome, virion, spikes, epidemiology

1. Introduction

The term coronavirus (CoV) was derived from the Latin name ‘corona’ which means the crown-like spikes or halo around the viral envelop viewed under the electron microscope and belongs to the family coronaviridae [1]. Coronaviruses have wide range of natural hosts and are believed to be the largest group of positive-sense RNA viruses causing an increasing threat to public health worldwide with devastating economic impacts [2]. The virus was first isolated from chickens by Beaudette and Hudson in 1937 which was termed as infectious bronchitis virus (IBV) [3], and in 1946 a transmissible gastroenteritis in swine was identified which was then followed by recovery of murine hepatitis viruses (MHV) in 1949 by Cheever and colleagues [4]. The human coronaviruses (HCoVs) were first discovered in 1965 by Tyrrell and Bynoe from nasal washings of a male child, which led to the proper categorization of the coronavirus genus [5]. These viruses are commonly found in bats, however, intermediate hosts like livestock and household pets are identified to be responsible for emergence and transmission of these viruses from bats to humans [6], and knowledge of bat origin of human coronaviruses could have a tremendous impact in predicting and prevention of possible occurrence of the coronavirus-related pandemic [7].

2. Virology

2.1 Overview

Generally, most coronaviruses are similar in structures which are enveloped but nearly spherical with a diameter of about 80-220nm. They are marked by prominent crown-like surface projections (spikes), 20nm in length, made up of heavily glycosylated type I glycoprotein, called spike protein (S) [10]. They are positive-sense single-stranded RNA viruses with a nucleocapsid of helical symmetry [6]. The virion envelope is made of inner and outer shells separated by a translucent space with a tongue-like inner membrane. Additionally, the virion has an internal ribonucleoprotein (RNP) measuring 1-2nm in length which is condensed into coiled structures [10]. The genome is composed of a single-stranded RNA linear molecule which is capped, polyadenylated and infectious harboring about 15000 to 20000 nucleotides with many genes encoding the polymerase and structural proteins as well as several open reading frames (ORFs) interspersed between these genes which encode non-structural proteins [11].

This genome is organized into 6-7 regions separated by junction sequences which contain signals for transcriptions of several mRNAs and each region contains one or more ORFs. There are three major structural proteins in the coronavirus virion: a nucleocapsid protein (N), a small integral membrane glycoprotein (M) and a large spike glycoprotein (S), however, additional proteins are found in some coronaviruses [12]. The N-protein play an important role in the encapsidating of genomic RNA, its subsequent incorporation into the virion through the formation of RNP and...
replication of RNA; M-protein facilitates virus maturation and determine the site for viral particles assembly; S-protein play a role in receptor recognition and membrane fusion which facilitate entry into the host cells [13].

The whole genome sequence of Covid-19 showed that it is a member of β-CoVs of group 2B with 79.0 % and 51.8 % similarity with SARS-CoV and MERS - CoV, respectively [14]. The SARS-CoV is a positive-sense single-stranded ribonucleic acid (RNA) virus consisting of thousands of nucleosides bases which phylogenetically related to other human coronaviruses especially HCoV-OC43 and HCoV-229E [15]. The genome of the SARS-CoV is made up of basically five main open reading frames (ORFs) which encode the replicase polyprotein; the spike (S), envelope (E), and membrane (M) glycoproteins; and the nucleocapsid protein (N) that are structurally and functionally similar to proteins of other coronaviruses. It also has a transcription regulating sequence (TRS) located near the beginning of each ORF and at the 3’end which discontinue RNA transcription and therefore facilitate syntheses of subgenomic negative-sense RNA species which may be translated in infected cells [16] (Figure 1).

![Figure 1: structure of β-Coronavirus genome](Image)

**2.2Source and Genome Characterization**

Genome sequencing of samples from bronchoalveolar lavage fluid and cultured isolates in late December 2019 from nine patients presented to Wuhan hospitals with symptoms of viral pneumonia who were tested negative for common causes of pneumonia revealed similarity among the samples with over 99.9 % sequence identity, closely resembling SARS-CoV (about 79 %) and MERS-CoV(about 50 %) but not identical [17], resulting in the viral identification by WHO as 2019-αCoV [9]. However, due to its close but a relatively distant relationship with SARS-CoV and MERS-CoV, the Coronavirus Study Group (CSG) categorized the virus pathogen as SARS-CoV-2 and the disease entity it caused was named by WHO as COVID-19 [18].

The origin of SARS-CoV-2 is debatable. Many attributed the source to bats due to 96 % similarity of the SARS-CoV-2 genome at the whole-genome level with bat-derived coronavirus genome [19]. In a study by Zhang and colleagues, Pangolin-CoV at the whole-genome level is found to be 91.02 % identical to SARS-CoV-2 indicating the probable Pangolin origin of SARS-CoV-2 and this is supported by the consistency of the five main amino acids in the RBD between Pangolin-CoV and SARS-CoV-2 [20]. However, a unique RRAR motif in the spike protein of human SARS-CoV-2 that is not present in coronaviruses isolated from pangolins indicates that SARS-CoV-2 may not come directly from pangolins [21]. Additionally, some strains of SARS-CoV-2 were isolated a few months earlier before COVID-19 was officially reported, thus disputing bats as a potential origin [22]. Some pieces of evidence from phyloepidemiological approaches are suggesting the possibility of importation of SARS-CoV-2 from other places to the Huanan Seafood Market [23]. Similarly, the virus was shown to be a recombinant virus between the bat coronavirus and another coronavirus with unknown origin with a snake as the most likely animal reservoir [9].

SARS-CoV-2 was found to be single-stranded RNA genome with 29891 nucleotides bases encoding 9860 amino acids with two adjoining untranslated regions (UTRs) and one long ORF encoding polyproteins [20]. The genome has six main ORFs with many more accessory genes with a 29,891-base-pair CoV genome [19]. As with other human coronaviruses, the SARS-CoV-2 genome has a transcription regulatory sequences and two terminal untranslated regions which facilitate the synthesis of subgenomic mRNA expressed ORFs [24]. Additionally, the genome has 5’ and 3’ terminal sequences (265 nt at the 5’ terminal and 229 nt at the 3’ terminal region) resembling other human coronaviruses with 5’ replicate open reading frame (ORF) 1ab-S-envelope(E)- membrane(M)-N-3’ gene order (Jin et al., 2020). The genome has a diameter of about 60-100nm with a length of about 29.9 kb [25]. It is unstable and can easy be deactivated by UV or when heated at 56°C for at least 30 min, and sensitive to most disinfectants such as diethyl ether, 75 % ethanol, chlorine, peracetic acid, and chloroform [26]. Three deletions were found in isolates from Japan, USA and Australia with ninety-three
mutations found over the entire genomes. Additionally, eight mutations were found in the spike surface glycoprotein, particularly three mutations (D354, Y364, and F367) found in the spike surface glycoprotein a receptor-binding domain which may indicate an important revolutionary trend of SARS-CoV-2 in the future [21] (Figure 2).

Figure 2: SARS-CoV-2 Structure (adapted from Astuti and Ysrafil (2020): Severe Acute Respiratory Syndrome Coronavirus 2 (SARSCoV-2): An Overview of Viral Structure and Host Response).

2.3 Virion Proteins and Receptor Interaction

The virion has four main structural proteins: The Spike (S), Envelope (E), Membrane (M) and Nucleocapsid (N) proteins as in other human coronaviruses but differentially deficient in the hemagglutinin-esterase gene [27]. The spike (S) protein is made up of two subunits, S1 and S2. The former consists of a signal peptide with two domains: an N-terminal domain (NTD) and a receptor-binding domain (RBD); whereas the latter consists of conserved fusion peptide (FP), heptad repeat (HR) 1 and 2, transmembrane domain (TM), and a cytoplasmic domain (CP) [28]. The similarity between S1 and S2 subunits with both the two bats SARS-like CoVs (SL-CoV ZXC21 and ZC45) and human SARS-CoV is found to be 70% and 99% respectively making the S2 subunit as an important target for both preventive and therapeutic approach against SARS-CoV-2 [29,30]. The receptor-binding domain (RBD) of S1 subunit is made up of two subdomains: external and internal subdomains. The former harbors most of the amino acid differences of RBD and is the facilitator for the direct viral interaction with the host receptor, human angiotensin-converting enzyme 2 (ACE2); while the latter is highly conserved [29]. (Figure 3).

Figure 3: Receptor interaction (Adopted with modification from Walls et al., 2020: structure, function and antigenicity of the SARS-CoV-2 glycoproteins).

The Orf3b and Orf8 are accessory proteins that are found in human coronaviruses that have also been identified in SARS-CoV-2 with some peculiarities. The Orf3b the protein of SARS-CoV-2 encodes a short protein with a putative binding site for Karyopherin (K) proteins which impair the interferon signaling network and the host innate defense by binding to the K-proteins, a function resembling SARS-CoV Orf6 protein function [24]. Contrary to the Orf8 of other coronaviruses, the SARS-CoV-2 Orf8 does not contain any functional domain due to lack of VLVVL (amino acid 75–79) motif found in SARS-CoV Orf8b in which it stimulates intracellular stress pathways with activation of NLRP3 inflammasomes [29]. With additional study, SARS-CoV-2 was demonstrated to use the same cell entry receptor, ACE2, as SARS-CoV, not CD26 as MERS-CoV. Also, structural analysis by Cryo-EM have shown that n The SARS-CoV-2 protein binds ACE2 with 10-20 folds higher affinity than SARS-CoVs indicating that, SARS-CoV-2 may be more infectious to humans than SARS-CoV [21].
2.4 Genomic Comparison
The whole level sequence analysis of SARS-CoV-2 revealed that it shares 79% sequence identity to SARS-CoV and around 50% to MERS-CoV2. However, the seven conserved replicate domains in ORFlab (used for CoV species classification) of SARS-CoV-2 shows 94.6% similarity to SARS-CoV, meaning that the two belong to same species. The receptor-binding protein spike (S) gene of SARS-CoV-2 is highly divergent to all previously described SARS-CoVs with less than 75% sequence identity to except a 93.1% identity to RaTG13. Homology modeling revealed SARS-CoV-2 had a similar receptor-binding domain structure to that of SARS-CoV [21]. Early sequencing the analysis revealed more than 99.9% similarity among the samples collected, however, further studies indicate that SARS-
CoV-2 has undergone mutation and substitutions at some codons evolving into two major types: L and S types with the L type exhibiting more aggressive behavior and likely to spread more quickly than the S type [26].

3. Epidemiology

3.1 Disease Source
The main sources of infection are the COVID19 infected patients and the more severe the disease, the more contagious it is [31]. Additionally, individuals without clinical signs and symptoms of COVID19 (asymptomatics) were reported to be potential sources of infection as they were proven to shed the infectious virus [32]. Similarly, patients recovered from COVID19 were reported to have shown positive RT-PCR test again, however, it is clear whether they are still infectious and can transmit the disease [33].

3.2 Incubation Period
The average incubation period is between 4-6days [34]. In a Wuhan study of 425 confirmed cases, the mean incubation period was found to be 5.2days [35]. In another study by Linton and colleagues, the mean incubation period was estimated to be 5days [36]. In the Chinese studies conducted outside Wuhan among 88 confirmed cases, a mean incubation of 6.4days was reported [37]. Lauer and colleagues estimated that 2.5% of the patients will develop symptoms within 2.2 days and 97.5% of patients will develop symptoms within 11.5 days [38]. However, an incubation period of 1-19days was reported in a familial cluster of five patients indicating that the COVID-19 incubation period is almost like MERS and SARS but slightly longer than influenza [39]. The 19days incubation period is a rare occurrence; therefore, specialists and professionals suggest the maximum incubation period as 14days for quarantine [40].

3.3 Disease Pattern
As most other viral respiratory diseases, COVID19 is also a self-limiting disease with the majority of the cases occurring as mild diseases that can recover within 1-2weeks [41]. The disease can occur in five different patterns: asymptomatic without any clinical sign or symptom, mild with mild upper respiratory tract symptoms, moderate with symptoms of pneumonia, severe with severe symptoms of pneumonia requiring oxygen and critical with acute respiratory distress syndrome (ARDS) or respiratory failure, shock, or multiple organ dysfunction [39]. In a report published by the Chinese Center for Disease Control and Prevention, 1.2% cases were asymptomatic, 81% were mild cases, 14% were severe cases, 5% were critical cases with a case fatality rate of 2.3% which is remarkably high in elderly and individuals with associated co-morbid conditions [42]. Healthcare workers are one of the most susceptible hosts of COVID19 due to their close contact with infected individuals. About 20% and 19% of healthcare workers in Italy and the United States were infected with COVID19 with the majority of the cases reported as mild cases but with significant mortalities [43].

3.4 Route of Transmission and Susceptible Host
The main route of transmission is by a human-to-human transmission via direct contact or via respiratory droplets spread by coughing, sneezing, or talking from an infected individual resembling the mode of spread of influenza [44]. Currently, there some evidence of the virus detected in blood and stool specimen, thus serving as potential routes of transmissions [21]. Based on the available evidence, anybody, irrespective of his age, sex, or race that has close contact with an individual infected with COVID19 is susceptible to contract the disease [34]. In a study by Wu and colleagues, the sex distribution was 1.06:1 indicating that the disease affects both females and males almost equally. However, the age distribution was remarkable with most of the cases (87%) affecting individuals within the age range of 30-79yrs almost sparing children (1%) and teenagers (1%) [41]. Currently, no evidence to support the vertical transmission of COVID19, however, adverse health outcomes were reported in some newborns born to COVID19 infected mothers [45]. The risks for adverse outcomes in COVID19 infected patients include increasing age and associated co-morbidities like cardiovascular disease, diabetes mellitus, chronic respiratory disease, hypertension, and cancer [43].

3.5 Geographical Distribution
Following the early report of COVID-19 identification from mainland China, the disease rapidly spread across the continent and the number of cases increased exponentially worldwide with the first imported case reported from Thailand on early January [43]. As of May 31, 2020, the COVID-19 pandemic has affected six geographical regions: Africa, Americas, Eastern Mediterranean region, Europe, South-East Asia, and Western Pacific [46]. By June 2, 2020, the European Centre for Disease Prevention and Control (ECDC) reported the highest incidence of COVID-19 in Americas region (n = 2,956,532) with United States having the highest incidence worldwide (n = 1,811,277), followed by Europe (n = 1,975,341), Asia (n = 1,151,637) and then Africa (n = 152,485). Similarly, the number of new cases were highest is Americas region, followed by Europe, Asia and then
Africa. However, the overall death rate was highest in Europe \((n = 175,572)\), followed by Americas \((n = 165,262)\), Asia \((n = 31,110)\) and then Africa \((4,344)\) with the least death recorded in Oceania \((ECDC COVID-19 Situation Update Worldwide, June 2, 2020)\) (Figure 4).

**Figure 4:** geographical distribution of Covid19 Worldwide, as of June 2, 2020. Adapted from ECDC Situation Update Worldwide, June 2, 2020.

### 3.6 Disease Chronology

Many cases of pneumonia-like cases of unknown causes were reported in the mid-December 2019 from the central Chinese city of Wuhan [44]. Five patients presented with fever, cough, and dyspnea accompanied by complications of acute respiratory distress syndrome were hospitalized from December 18 to December 29, 2019 [47]. On December 31, 2019, China alerted WHO about a cluster of 41 patients with new cases of unusual pneumonia in Wuhan city comprising of about 11 million people located in the central Hubei province [48]. Around January 2, 2019, 41 hospitalized patients were reported to be laboratory-confirmed cases of the new virus, and most of them were males some of whom were exposed to the Huanan seafood market [49]. The World Health Organization (WHO) on January 7, 2020, identified the virus as a novel coronavirus and officially but tentatively named it as 2019-nCoV, the new coronavirus in 2019 [9].

Around January 13, 16 and 21, confirmed imported cases in three Asian countries comprising of Thailand, Japan, and Korea were respectively reported marking the beginning of the global pandemic [2]. By Jan 22, 2020, a total of 571 cases of the 2019-new coronavirus \((2019-nCoV)\) have been reported in 25 provinces (districts and cities) in China [50]. On Jan 24, 2020, with reports of thirteen exportation events, the cumulative incidence of the new coronavirus \((2019-nCoV)\) in China was estimated to be 5502 cases [50]. By Jan 25, 2020, a total of 1975 cases of 2019-nCoV infection were confirmed in mainland China with a total of 56 deaths that have occurred [52]. As of January 30, 2020, the confirmed cases in China were estimated to be 7734 cases and concurrently 90 more cases were also confirmed in many countries including Taiwan, Thailand, Vietnam, Malaysia, Nepal, Sri Lanka, Cambodia, Japan, Singapore, Republic of Korea, United Arab Emirates, United States, The Philippines, India, Australia, Canada, Finland, France, and Germany with a case fatality rate estimated to be 2.2% [53]. On January 31, the WHO declared the disease as the Public Health Emergency of International Concern (PHEIC), which indicate that the disease can pose threats to many countries and therefore requires an active coordinated international response [1]. By 31 Jan 2020, the total number of confirmed cases was 9,826 cases globally affecting 19 countries with 213 mortalities [54].

On February 11, both ICTV and WHO comes up with new but different names for the virus. The former named as SARS-CoV2 while the later named it as COVID19 (WHO COVID-19 Situation Report-22, 2020). As of February 29, 2020, the total number of confirmed cases was 85,403 cases globally involving 53 countries with 2,924 global deaths [55]. On March 11, the WHO declared the disease as a “pandemic public health menace” [1] (Figure 5). By March 31, 2020, the global confirmed cases reached 750,890 cases with 36,405 global deaths [56]. A global confirmed case was reported to be 3,090,445 cases on April 30, 2020 with 217,769 total deaths [57]. As of May 31, 2020, the total confirmed cases globally were 5,934,936 cases with 367,166 global deaths [46]. As of June 2, 2020 (the last day of preparing this manuscript), the global confirmed cases were reported by ECDC as 6,245,352 cases with global deaths recorded as 376,427 deaths \((ECDC COVID-19 Situation Update Worldwide, June 2, 2020)\) (Table 1).
4. Conclusion

SARS-CoV-2 is an emerging disease of public health importance. The majority of the cases are mild self-limiting conditions but few cases present as serious and critical conditions requiring high input of clinical management. It is highly contagious with devastating consequences especially in the elderly and those with associated medical conditions like diabetes, hypertension, respiratory diseases, or cardiovascular conditions. The primary route of transmission is through contact directly with an infected person or via respiratory droplets with an average incubation period of 4–6 days. The structure of the virus resembles SARS-CoV and MERS-CoV in many aspects but appeared to have some peculiarities with bats as the most probable source of the disease. It has a rapid geographical distribution making it a global pandemic and serious infectious disease requiring urgent and active measures to curtail its spread. It is clear that the Wuhan Seafood market was not the only origin of this disease indicating other potential origins as was identified in this paper. Effective control measures, appropriate preventive strategies, and active case surveillance are the best approaches to curtail the spread of this virus and to win the battle against it.

Conflict of Interest: None to disclose

Funding: None

References


Table 1: Shows the Global Incidence of COVID-19 across Countries with the Case Fatality Rate from December to May, 2020

<table>
<thead>
<tr>
<th>Date</th>
<th>Number of Confirmed Cases</th>
<th>Deaths</th>
<th>Case Fatality Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>December 31, 2019</td>
<td>41</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>January 31, 2020</td>
<td>9826</td>
<td>213</td>
<td>2.2</td>
</tr>
<tr>
<td>February 29, 2020</td>
<td>85,403</td>
<td>2924</td>
<td>3.4</td>
</tr>
<tr>
<td>March 31, 2020</td>
<td>750,890</td>
<td>36,405</td>
<td>4.8</td>
</tr>
<tr>
<td>April 30, 2020</td>
<td>5,090,445</td>
<td>217,769</td>
<td>7.0</td>
</tr>
<tr>
<td>May 31, 2020</td>
<td>5,934,936</td>
<td>367,166</td>
<td>5.2</td>
</tr>
</tbody>
</table>


