Abstract

The outbreak of coronavirus disease 2019 (COVID-19) in December 2019, reportedly caused by novel coronavirus known as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), thereby the concerns to combat the pandemic has risen worldwide. The pandemic began in Wuhan, China and soon spread at an unprecedented rate that cost hundreds of thousands of lives due to high transmissibility and pathogenicity. The disease severity has been the subject of various factors including host immune response, gender, genetic factors, pre-existing medical conditions, and the virus’s evolutionary pattern. Although the causes of evolution are yet to be known, the research studies demonstrate the changes in the structural genome of the virus, particularly the genes that encode the receptor-binding proteins (RBD). The purpose of this review is to provide the current literature and findings on virology of SARS-CV-2 structure and genetic factors of the host and the virus that may shed lights to future studies of diagnosis and therapies.

Keywords: SARS-CoV-2; COVID-19; Coronavirus; Receptor binding proteins; Pandemic

Introduction

Background and Origin

Coronaviruses are a diverse group of viruses that cause infection in animals and humans and they comprise a positive strand genomic RNA. The members of this family have distinctive morphological features and cause mild to severe respiratory infections in the host body. However, the twenty-first century began with the emergence of two pathogenic coronaviruses known as severe acute respiratory syndrome (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS/CoV) and recent emergence of novel SARS-CoV-2 that were detected to have had zoonotic origin (that are transmitted from animals to humans) [1-3]. SARS-CoV-2, later known as coronavirus 2019 (COVID-19) was first detected in the city of Wuhan, China [4] that cause mild to severe viral pneumonia followed by fever, severe breathing problems (dyspnea), chest discomfort and lung infiltration [4]. In comparison with MERS-CoV and SRAS-CoV, SARS-CoV-2 spreads faster and wider and so far has caused a remarkable threat to human race worldwide [5]. The outbreak was reported to have originated from seafood market in Wuhan, China that sells live animals and poultry. Although the first cases of COVID-19 are dated back to late December 2019, the first reported case was reported on 8th of December 2019 [6].

RNA sequencing analysis of the isolated virus from infected patients detected novel SARS-CoV-2 to be belonged to β-coronavirus subgenus of the Coronoviridae family which is part of virus order known as nidovirales [7-11] with zoonotic nature of transmission. Further nosocomial infections at hospitals and family members identified the transmission to have occurred from human to human [12-14].

In the statement made by the world health organization on January 31, 2020, Public Health Emergency of International Concern (PHEIC) listed COVID-19 as a posing a threat for to many countries and called for international precaution. The most
common features of COVID-19 in patients are respiratory tract infection which are detected with computer tomography (CT) and chest X-ray in the acute period [15]. In a mild case of COVID-19, patients may experience mild respiratory infections which can turn to become acute with the severe case of pneumonia. Based on WHO announcement, Approximately 80% of patients were asymptomatic or developed mild symptoms, while 20% are detected with serious clinical conditions which are particularly observed with older people or the individuals with hypertension, cardiovascular diseases, chronic obstructive pulmonary disease, diabetes, obesity, and any malignancy. Genetic analysis of viral isolates from infected patients showed SARS-CoV-2 nucleotides to have nearly 88% similarity with bat-like SARS coronavirus indicating a zoonotic nature of the virus. Bats were found to be the reservoir hosts for SARS-CoV-2 [16]. The virus is thought to be transmitted from person to person by droplets that can spread from the infected people with talking, sneezing/coughing, touching the mouth, nose, or eye mucosa after the contact with infected areas. The virus has a mean incubation period of 6.4 days [17]. Clinical manifestations in COVID-19 patients, including fever, cough, and tiredness are similar to that of SARS-CoV and MERS-CoV cases. Shortness of breath, chills, sore throat, loss of taste or smell and chest pain are listed among other disease symptoms [15]. The severity of the disease varies with the immune system response as well as pre-existing clinical conditions from person to person. As the current treatments have been insufficiently effective on the patients, the scientists and researchers have been trying to develop the efficient therapy with minimal short- and long-term side effects and to reduce the mortality rate worldwide.

Phylogenetic analysis of viral isolates from the patients reveals various evolutionary processes of the virus. Epidemiological studies have shown that clinical outcomes are greatly influenced by the individual’s age, gender, complications, and other possible unknown clinical parameters [18]. All societies should take the necessary precautions tackle the current pandemic.

The origin and structural features of SARS-CoV-2

Coronaviridae virus families constitute the Nidovirales Arteriviridae, Roniviridae, and Mesoniviridae (Figure 1). SARS-CoV-2, along with SARS-CoV and MERS-CoV, is a member of the subfamily Coronavirinae called Beta-coronavirus [19]. Four other types of disease-causing CoVs are endemic in humans: 229E, OC43, NL63 and HKU and they cause common cold symptoms. SARS-CoV causes serious respiratory distress and death in infected cases. The first SARS-CoV case was identified in 2003. The Middle East respiratory syndrome coronavirus (MERS-CoV) was described highly pathogenic in 2012. SARS-CoV-2 is the 7th known type of coronavirus that cause severe symptoms after identification of MERS and SARS, while HKU1, NL63, OC43, and 229E have been associated with milder symptoms [20].

Coronaviruses are the largest known RNA viruses with no segments, single-stranded, positive chain and genome' length ranges from 26 to 32kb [21,22]. The typical coronavirus genome organization is 5'-1a-1b-S-E-M-N-3', and this sequence is common among coronaviruses. The partially overlapping open frames 1a and 1b makeup two-thirds of the viral genome and encode all the elements necessary for viral RNA synthesis. This region is defined as the replicase gene [23]. Two large polyprotein (pp1a and pp1b) are encoded from coronavirus replicase genes (ORF 1a and ORF 1b). Structural proteins of the coronavirus are encoded by almost a third of the viral genome [23].

SARS-CoV-2 includes nonsegmented, single-stranded, positive-sense genomic RNA, with 5′ cap structure and 3′ poly-A tail similar
to other coronaviruses [24]. The genome sequence of the SARS-CoV-2 BatCoV has 96.2% similarity with RaTG13 coronavirus in bats and 79.5% similarity with SARS-CoV [25]. Thus, bat is still considered the most likely species for the 2019 new coronavirus (SARS-CoV-2) source.

Generally, coronaviruses encode four structural proteins in their genomes [26]. Genomic RNA of the coronaviruses is packaged by the viral nucleocapsid (N) protein. Three other structural proteins of coronavirus are Spike (S), Membrane (M), and Envelope (E) proteins that form the envelope of virions. These four structural proteins are common among coronaviruses; (N, M, S, E). The S glycoprotein provides the coronaviruses' entry into the host cell, a class I virus fusion protein [27]. Coronavirus membrane (M) protein is the most common structural protein and plays a vital role in forming virus particles that interacts with N protein in many coronavirus species. Coronavirus nucleocapsid (N) protein is required for the efficient packaging of viral genomic RNA into the helical ribonucleoprotein (RNP) complex during the formation of virions [28]. Coronavirus envelope (E) protein is the main component of the viral envelope. It also plays an important role in releasing virions from the cell and the regulation of other cellular functions [29]. The S genome of the SARS-CoV-2 shows 93.1% high similarity with the RaTG13 S genome, while it has a similarity rate of less than 75% with SARS-CoV [25]. It is stated that the highest similarity between SARS-CoV-2 and SARS-CoV is in the N terminus [30,31].

Coronaviruses are introduced into the cell by the S protein of the virus. While the mouse coronavirus enters the MHV cell via the CEACAM1 receptor, SARS-CoV can enter the cell by binding S protein to angiotensin-converting enzyme 2 (ACE2). When the S1 subunit of the S protein is attached to the host cell receptor, the S protein undergoes a series of structural (conformational) changes to ensure fusion with the host cell membrane [21]. The SARS-CoV-2-S protein binds ACE2 with higher affinity than SARS-CoV Receptor Binding Domain (RBD) [32]. After S protein binds with the ACE2 receptor on the host cell, the S protein is cleaved by host cell proteases and viral-host membrane fusion occurs. Depending on the type of virus, the fusion may occur in the host cell's plasma membrane or acidified endosomes after receptor-mediated endocytosis. The binding of the virus with host cell receptors is a critical factor for the virus's pathogenesis [33,34]. A better understanding of receptor binding and protease effect relative characteristics will also help predict how the coronaviruses infect humans and adapt to the host cell.

**Diagnosis of COVID-19 and current treatment**

Epidemiological history, clinical findings, nucleic acid detection, and CT scan are mainly used to diagnose COVID-19 (Figure 2). Additionally, the IgM/IgG Fluorescence Immunochromatographic Assay method (POCT) produces rapid early diagnosis results. In cases with strong epidemiological relationship with COVID-19 infection, the application of serological tests in serum samples taken in the acute or convalescent phase may support the diagnosis. For this purpose, enzyme-linked immunosorbent assay (ELISA) or rapid antibody tests that detect IgM/IgG help investigate the ongoing outbreak and provide a retrospective evaluation of the attack rate and the severity of the epidemic. In addition, recent studies on viral isolates collected from stool samples show SARS-CoV-2 replication in a stool sample [35] that could be considered as a potential source of viral isolation.

The clinical symptoms of patients infected with SARS-CoV-2 are different from each other, that is, atypical. In cases, fever, shortness of breath and radiological findings are compatible with bilateral pulmonary infiltrate. Although different protocols were published to target N, E and S genes for diagnosis of COVID-19, it is more efficient to adopt more accurate technique such as RT-qPCR can identify SARS-CoV-2 virus based on N gene and RdRP, ORF1a and reduce the errors in the test result [36].

Among drugs that are candidates for treating COVID-19, re-evaluation of old medications for antiviral therapy is a usual strategy because their safety profile, side effects, posology, and drug interactions are well known [37,38]. Apart from antiviral treatment, vaccine development studies, efficacy studies of biological agents (MABs), and different antiparasitic drugs are being tested. In the absence of the vaccine, an international investigation “Solidarity” was launched by WHO to overcome the pandemic. The drugs included in this study are listed as lopinavir and ritonavir, lopinavir and ritonavir interferon beta, chloroquine and remdesivir [39].

Interleukin-6 (IL-6) pathway inhibitors are another alternative form of treatment. Tocilizumab was developed against recombinant human IL-6 receptor monoclonal antibody. In COVID-19 patients, high IL-6, IL-1 and TNF-α levels were observed with severe clinical manifestations that result in "cytokine storm". Tocilizumab is an immunosuppressive drug against the IL-6 receptor in COVID-19 patients who developed severe pulmonary symptoms. Although the use of Tocilizumab against IL-6 receptor in the cytokine storm is supported, there is no published clinical data with proven effectiveness [40].

Kinases are another alternative way of therapy. p21-activated protein kinases (PAKs) are small (p21) GTPases that contain members of the Cdc42 and Rac families. It has been shown that PAK1 has an important role in the entry, replication and spread of many important viruses, including influenza and HIV [41,42]. Coronaviruses use macropinocytosis to enter the host cell, and this process has been shown to depend on PAK1 activity. Thus, this suggests that PAK1 inhibitors can be valuable in the treatment of COVID-19 [43].
Pathogenesis and Immunity of SARS-COV-2

As it is known, ARDS (Acute Respiratory Distress Syndrome) is pulmonary edema caused by increased permeability of the capillaries in the lungs. It causes the fluid transfer to the pulmonary region as a result of acute inflammatory lung injury. Here, the causes damaging the alveolocapillary membrane may be due to the direct airway, such as pneumonia, aspiration, or indirect hematogenous causes such as sepsis and uremia. In the study by Xu et al., [53] ARDS was shown to be one of the main causes of death in severe COVID-19.

Cytokine storm is one of the primary mechanisms for ARDS formation. It is an over-developing systemic inflammatory response from uncontrolled release of pro-inflammatory cytokines (IFN-α, IFN-γ, IL-1, IL-6, IL-12, IL-18, IL-33, TNF-α) and chemokines (CCL2, CCL3, CCL5, CXCL8, CXCL9, CXCL10) by immune effector cells in SARS-CoV infection. This uncontrolled immune response was stated to be closely related to the determination of COVID-19 severity. Also, increased interleukin-6 (IL-6) levels with “cytokine storm” have been observed in COVID-19 patients with severe clinical findings [54] (Figure 3). In the cytokine storm seen in severe COVID-19 patients, an immunosuppressive drug against IL-6 pathway was recommended against the IL-6 receptor, targeting the IL-6 receptor [55].

MAS (macrophage activation syndrome) may also develop in COVID-19 disease, but there is no evidence on the frequency of occurrence and definitive treatment. It is believed that 10% of COVID-19 patients may be in the critical condition and cytokine storm caused by MAS may worsen the prognosis. For this reason, the correct and timely development of effective treatment against the cytokine storming can efficiently reduce the mortality rate. The case series and observations reported to date indicate that the MAS findings in COVID-19 patients generally show similarity to MAS tables in hereditary HLH (hemophagocytic lymphohistiocytosis) rheumatic diseases [56]. However, MAS/HLH findings may not develop in all patients depending on the disease’s course. The scores or criteria used in the diagnosis of other conditions may not always help decide the best treatment method [56].

The presence of signs such as persistent fever, an elevated level of C-reactive protein (CRP) and ferritin values, a high D-dimer, lymphopenia and thrombocytopenia, liver deterioration function tests, hypofibrinogenemia or increased triglyceride values may indicate that the MAS deteriorate the patients’ conditions. MAS is a complication that requires close follow-up and early treatment. When the patient diagnosed, it can become much more difficult or impossible to suppress the developing cytokine storm within hours [56].

Possible impact of frequent mutations in the novel SARS-CoV-2 genome on pathogenicity and transmission ability

The increasing epidemiological and clinical evidence implicates that the SARS-CoV-2 has stronger transmission power and lower pathogenicity than SARS-CoV [57]. However, the high transmission mechanism of SARS-CoV-2 has not been clarified yet. Researches have found that the virus is rapidly mutating. As a result of mutation, they have discovered three different “variants” of COVID-19 consisting of closely related strains called “A”, “B” and “C”. Based on SARS-CoV-2 evolutionary tree, and it was stated what the mutations occurred during the viral spread from Wuhan to Europe and North America. Thus, the scientists have been investigating to determine the factors causing the evolution of SARS-CoV-2 to type-A, type-B and type-C with limited genome data [58].

The coronavirus subtype "A" variant is most closely associated with the virus in both bats and pangolins. Researchers have described as “the source of the epidemic”. The type “B” is derived from "A" by two mutations. Later, a type C variant formed which derived from B type coronavirus with a uniform mutation. Type A has been found in Americans who are living in Wuhan and patients diagnosed in the United States and Australia. While type B coronavirus is common throughout East Asia, the C variant is predominantly in Europe; it has been expressed as the primary type of virus found in patients from France, Italy, Sweden and the UK [59].
Detection of mutations occurring in the SARS-CoV-2 genome is important to understand the evolution of the virus and reveal genotype changes that occur during the transmission of SARS-CoV-2. Genotyping studies have demonstrated a few common mutations in SARS-CoV-2 genomes. It has been stated that the mutations might be related to the virus permeability and the extent of pathogenicity. Detected mutations are found mostly in S protein, RNA polymerase, RNA primase, and nucleoprotein, which are essential for the vaccine efficacy [58]. For this reason, mutation analysis might be the results that support the literature ineffective treatment studies to prevent SARS-CoV-2 coronavirus infection.

Discussion

A recent outbreak of COVID-19 has led to global and panic and cost more than a million lives worldwide. Genome analysis of SARS-CoV-2 remarked on the significance of spike (S) glycoprotein that maintains viral invasion by binding to angiotensin-converting Enzyme 2 (ACE2) [60,61]. SARS-CoV-2, similar to other RNA viruses, display high mutation rate that may affect the virus’s pathogenicity [62] and consider the importance of S protein in SARS-CoV-2 virulence, a mutation on the spike region is likely to enhance the virus’s pathogenicity [63]. S protein was shown to undergo mutation owing to a receptor-binding domain (RBD) of S protein, the most variable genomic component of SARS-CoV-2 [64]. Recent studies detected that infected patients with the SARS-CoV-2 with substitution mutation of Asp to Gly in 614 position (D614G) in the spike glycoprotein show higher viral loads in the upper respiratory tract [65-67] and this could suggest the impact of the mutation on the transmission rate. Moreover, the D614G strain also manifests stronger binding affinity to angiotensin-converting (ACE2) [68,69] that could require further investigation of the patients’ pathogenicity and treatment. The ex vivo and in vivo studies of D614G SARS-CoV-2 strain on different cell lines detected higher replication rate of the strain in upper respiratory tract cell lines with larger numbers of human ACE2 (hACE2) receptors [70,71]. However, further analysis of hamster did not signify a higher pathogenicity rate with the studied strain [4,70].

Further studies suggest a higher prevalence rate of D614G SARS-CoV-2 strain in younger ages which could relate to the higher replication rate [72,73]. However, the follow-up examination and observation could not confirm the effect of the mutation on virulence or pathogenicity [72]. Similar to other viruses’ features, investigation on different types of SARS-CoV-2 shown to have a higher transmission rate was followed by lower pathogenicity degree [72-74]. Genomic analysis of SARS-CoV-2 on patients in China detected several single nucleotide variants (SNVs) on S protein that each manifested noticeable changes in the virus’s pathogenicity [75]. Furthermore, mutations on nonstructural proteins (nsps) of SARS-CoV-2 was described to be the cause virus’s high contagion rate compared to that of SARS-CoV [76,77]. Some other studies on different mutation positions affect SARS-CoV-2 pathogenicity and decrease virulence. Study on full-length SARS-CoV-2 discovered 81 base pair deletion on AZ-AU2923 genome, an open reading frame (ORF)7a gene that synthesizes the accessory proteins replication, infection, and proliferation of the virus inside host body [66,78]. Detected substitution of Serine by Arginine (S247R) in S1 subunit of S protein was in the SARS-CoV-2 isolates of some patients [35]. Although the position is within the N-terminal domain of S1 subunit that is not directly part of S subunit binding to ACE2 receptors, N-terminal is adjacent to C-terminal that binds to ACE2 receptors. Mutation in this position could affect the binding affinity of the virus. Confirmation of such possibilities needs further investigation. Indicated substitution of Asparagine by Arginine in 439th position (N439R) in RBD increase SARS-CoV-2 binding affinity to ACE2 receptors and enhance viral transmission from human to human [79,80]. However, detected various mutation points on SARS-CoV-2 RBD, the mutation rate on the virus reported to be remarkably low compared to other members of coronavirus families. On the contrary, the higher viral loads of SARS-CoV-2 in humans are likely to result in mutation at the infection period. Even though almost all reported mutations did not display increased pathogenicity of the virus, more than one mutation in RBD is likely to enhance virulence. However, further in vitro and in vivo investigations are required to confirm a change in viral pathogenicity that can lead to drug resistance or further treatments. Investigating these mutations in the SARS-CoV-2 genome whether it can affect the function of the structural genes and the pathogenicity of SARS-CoV-2, revealed that the mutation analysis is critical for the drug and vaccine development against COVID-19 in the stage of the pandemic.

Conclusion

Coronaviruses have been known as the common cause of death for many years, and with the emergence of the new type of SARS-CoV-2 in early 2020, they pose a more severe threat to human health, spread to the whole world in a short time, causing thousands of cases and hundreds of deaths. Within the scope of many molecular studies that are being conducted on different populations on a global basis, it is thought that COVID-19 outbreak analysis will also contribute to the control of the pandemic in terms of providing epidemiological data, investigating viral resistance development mechanisms, directing treatment and clinical monitoring of the disease.

Authors Contributions

Şeyda Demirkol drafted the paper and prepared the images and figures. Nasim Kherad collected the recent articles and information on COVID-19 and drafted the paper. Elif Aslan helped in data analysis and contributed in data collection. Saadet Busra Aksoyer Sezgin, Sema Ketenci, and Ebru Nur Ay helped in writing the paper and arranging the citation. Meryem Alagoz and İlhan Yaylım did the study design, checked the information and data accuracy and finalized the draft.
References


