1. Introduction

Coronavirus disease-2019 (COVID-19) was declared a pandemic by the World Health Organization on March 11, 2020, affecting more than 200 countries and 7 continents worldwide. The disease is caused by novel coronaviruses, which are zoonotic viruses, thus can transmit the infection from animals to humans. The International Committee on Taxonomy of Viruses (ICTV) has named the virus “Severe Acute Respiratory Syndrome Coronavirus-2” (SARS-CoV-2) [1]. Coronaviruses, scientifically called Orthocoronavirinae or Coronavirinae, are a large family of enveloped viruses that belong to the subfamily Coronavirinae and order Nidovirales [2]. The Coronavirinae has four genera: Alphacoronavirus, Betacoronavirus, Gammacoronavirus, and Deltacoronavirus. To date, there are seven strains of human coronaviruses (HCoVs) discovered, including HCoV-229E, HCoV-NL63, HCoV-OC43, HCoV-HKU1, Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV), Middle East Respiratory Syndrome coronavirus (MERS-CoV), and SARS-CoV-2 [3, 4]. SARS-CoV and MERS-CoV are coronaviruses that caused severe infections in the human population [5-8]. SARS-CoV-2, formerly known as 2019-nCoV, first emerged in 2019 at Wuhan City. According to worldometer reports, the novel coronavirus strain has killed 969,287 people worldwide as of September 21, 2020 [9].

SARS-CoVs-2 have enveloped positive sense and SSRNA viruses, which attach to human cells via spike (S) proteins, thus confirming their human to human transmission [10]. S1 subunit of S proteins has receptor binding domain responsible for attachment to its functional receptor angiotensin-converting enzyme 2 (ACE2) in the human lungs and causing bronchial epithelial cells and type II pneumocytes infection [11-13]. COVID-19 causes mild symptoms like dry cough, sore throat, fever, body ache, fatigue, loss of taste and smell in some patients whereas, severe clinical impacts like seizures, stroke, and pneumonia in others [14]; in contrast, a significant population infected with COVID-19 does not show any symptoms (asymptomatic). Additionally, in many COVID-19 patients, morbidity and mortality are seen to be related to age, sex, and genetic makeup [15]. Therefore, the ability of SARS-CoV-2 infection raises many questions about its susceptibility and pathogenesis in individuals due to these factors. The main objective of this review is based on recent studies showing a link between novel coronavirus strain and host genetics [16]. Thus, the relationship between human genetics and COVID-19 needs to be uncovered to understand SARS-CoV-2 better to find its clinical cure. activity against human leukemia cell lines [9].

2. A brief history of coronaviruses and COVID-19

2.1 Origin of human coronaviruses

The history of human coronaviruses (HCoVs) dates back to the 1960s when a virus named “B814” was screened from the respiratory tract of an adult with a common cold by Tyrrell and Bynoe [17]. The virus was successfully cultured by Hamre and Procknow and was named as ‘229E’ by Hamre [18]. Both ‘B814’ and 229E were found to have lipid covering to be virulent and ether-sensitive, unrelated to known paramyxoviruses [19]. Further, McIntosh et al.
discovered many ether-sensitive virus strains from human lungs and named them as ‘OC’ as they were organ cultured [20]. Electron microscopy done by Almeida and Tyrrell reported the similar morphology (pleomorphic, membrane coated with club-shaped projections and 80-150 nm-sized particle) for ‘229E’ and ‘OC’ viruses, and these were also morphologically indistinguishable to the agents responsible for bronchitis infection in chickens [19-21]. Thus, a new group of viruses, including infectious bronchitis virus, mouse hepatitis virus, and gastroenteritis virus of swine, are categorized as a new genus of viruses. It was named ‘Coronavirus’ because of the crown-like appearance of its surface proteins [22-24].

2.2 Coronavirus outbreaks in the past

Previously coronaviruses were considered to infect animals until severe acute respiratory syndrome (SARS) outbreak caused by SARS-CoV in 2002 in Guangdong, China causing a cough, dyspnea, and pneumonia [25-26]. Later in 2012, an epidemic MERS showed similar clinical symptoms to SARS but was less transmissible with a mortality rate of ~34% compared to SARS, which has ~10% mortality rate [27-28]. Functional receptors for SARS-CoV and MERS-CoV are angiotensin-converting enzyme-2 (ACE2) and dipeptidyl peptidase 4 (DPP4), respectively [28-29]. SARS-CoV, MERS-CoV, and SARS-CoV-2 belong to subgroup beta coronaviruses but are all distinct species of this genus [4, 28]. Presently, the whole world is suffering from COVID-19, and scientists are trying hard to defeat the novel SARS-CoV-2.

2.3 Case Reports: COVID-19 and Human Genetics

Herein we have underlined case studies of genetic factors directly linked to the survival and death of COVID-19 patients. Severe infection of SARS-CoV-2 is rare in healthy younger people. On the contrary, data from 14 May 2020 by the National Institute of Public Health in the Netherlands reported 3.5% of patients hospitalized were <35 years old, and 6 out of 7 patients died were men [30]. A case study in Iran where three brothers aged 54-66 years died from COVID-19 despite having any underlying medical complications led the scientists to hypothesize genetic predisposition to the SARS-CoV-2 infection in some individuals [31]. Additionally, another case report published the effect of X-linked agammaglobulinemia (XLA) on two patients aged 34 and 26 years in Italy who developed severe pneumonia after being infected with SARS-CoV-2 but recovered [32]. Thus, suggesting that the patients born with XLA defects have higher chances of developing acute pneumonia as COVID-19 manifestation. A study of 4 young men with a loss-of-function mutation in the TLR7 gene present on the X chromosome affects type I and type II INFs. This immunological defect resulted in severe COVID-19 in all these patients requiring clinical ventilation and leading to the death of one [33].

A case study of Iran reported a death rate of 10% out of 9.5% of hospitalized COVID-19 patients with diabetes [34]. According to statistics published in a case report, 322,984 patients from the UK with variants of Alzheimer’s disease were more susceptible to SARS-CoV-2 infection and the risk of progression to severe illness [35]. Data collected from 6 hospitals of Vancouver showed 87 percent COVID 19 patients with blood group A or AB required mechanical ventilation than patients with O/B blood types [36]. The location of TMPRSS2 on the 21q22.3 chromosome confirms its potential to cause SARS-CoV-2 infection in a population with Down syndrome [37].

3. Link between human genetics and COVID-19

SARS-CoV-2 is strangely selective in terms of morbidity and mortality in individuals. A large number of infected people did not get sick. Individuals suffering from medical complications like cardiovascular diseases, hypertension, diabetes mellitus, cancer, and the older population reported severe COVID-19 symptoms. However, reported deaths due to COVID-19 also included the young and previously healthy people. Therefore, host genetics analysis can dive deeper to solve these mysteries behind individuals showing diverse clinical outcomes when encountered with the novel coronavirus.

3.1 Genome, transmission and pathogenesis of SARS-CoV-2

After the COVID-19 outbreak, the biological community seems curious to unlock the genome characteristics of SARS-CoV2. Various bioinformatics and biotechnological tools are being employed to accomplish this task. One such study suggests that the viral genome ranges from 26 to 32 kilobases [38]. SARS-CoV2 contains six ORFs in its genome [39]. Among the six ORFs, the first ORF (orf1ab), as shown in figure 1, comprises a significant portion of the genome (67%), is located at 5’end, and encodes 16 non-structural proteins (orf1a, polyproteins) [40]. Papain-like protease (nsp3) and chymotrypsin-like protease (nsp5).

![Figure 1: SARS-CoV2 genome showing non-structural proteins on 5’ UTR of ORF1ab. nsp3 symbolizes papain-like protease, nsp5 represents chymotrypsin-like protease, nsp 13 represents helicase, and nsp12 represents RNA dependent RNA polymerase. In contrast, 3’UTR encoding structural proteins are Surface (S), envelope (E), membrane (M), and nucleocapsid (N) proteins.](https://doi.org/10.52679/tabcj.2020.0007)
helicase (nsp13), and RNA dependent RNA polymerase (nsp12) are involved in its replication and transcription. 3'ORF encodes structural proteins such as surface (S), envelope (E), membrane (M), nucleocapsid (N) proteins [41]. Apart from this, SARS-CoV2 contains two UTRs at 5' end and 3'end and are of an approximate size of 265 and 358 nucleotides, respectively [40]. Among the viral structural proteins, S protein is the principal protein involved in attachment with the host and viral entry in the host [41].

Compared to subtypes SARS and MERS, SARS-CoV-2 is causing more acute clinical symptoms and is more transmissible with high mutational rates [6, 42]. There are various modes of SARS-CoV2 transmission, including contact, droplet, animal to human, etc. [43]. Also, it is a common notion about all the types of SARS viruses that they have a high potential to shift to different host species bringing new diseases to the world. The emergence of SARS-CoV2 is also an example of a host-shifting mechanism. The alterations in Spike protein, particularly in its S1A domain, binds to glycoproteins such as CD147, which further binds to ACE2 determines infection in hosts [44]. Genetic polymorphism and allelic variants of the ACE2 gene decide on its binding and invasion in the host [45]. Moreover, the transmembrane protease serine 2 (TMPRSS2) is also essential for its entry [46]. Further, macrophages act as its reservoir inside the host cell, and other immune cells, including dendritic cells, act as a transporter and responsible for pathogenesis [47].

3.2 Influence of human genetics on SARS-CoV-2 infection

Host genomic variants are essential to understand host-pathogen interactions. Approaches used to study these gene variants include genome-wide association studies (GWAS) and multi-omics based methods [48]. Thus, ABO blood types, HLA haplotypes, Polygenic risk scores (PRS), and the connection of the available biobanks to Electronic health records (EHRS) are used to understand SARS-CoV-2 infection. For instance, the linkage of gene cluster 3p21.31 is known to be responsible for causing COVID-19 and respiratory failure [49-51]. GWAS studies also unfold the mystery of males being more susceptible to COVID-19 due to variation in the TLR7 gene, which is essential in innate immunity and is located on the X chromosome [52]. A recent study concluded that patients with inborn defects in type I IFNs were found to be infected with severe COVID-19 infection irrespective of their age, sex, and nationality [53]. Human Leucocyte Antigen (HLA) systems are essential in deciding the severity of COVID-19 infection as they can critically affect the response of immunoglobulins (Igs) in hosts during the viral life cycle [54-55].

3.2.1 Role of ABO blood type in SARS-CoV-2 susceptibility

Literature involving large numbers of genetic data of SARS-CoV-2 infected and non-infected people revealed a link between human genome association to lower or higher COVID-19 vulnerability and mortality. For instance, the people with blood group O have lower susceptibility towards novel coronavirus [56]. The reason for these linkages can be formulated in the light of binding affinity reactions between the SARS-CoV-2 spike proteins and blood group antigens [41, 57]. The antigens on red blood cells can be expressed in the airway or alveolar epithelial cells and can serve as viable receptors for many infectious microorganisms [58-59]. SARS-CoV and SARS-CoV-2 structural studies showed similarity in the nucleic acid sequence and ACE2 receptor recognition pattern [57]. Antigens present on ABO blood groups lead to the formation of sialoside grouping, which further increases the binding of N-terminal domain (NTD) and RBD domains to CD147 and ACE2 respectively of SARS-CoV-2 virus [41, 60]. Therefore, the absence of blood group antigens from the erythrocyte surface of O group individuals makes them less susceptible to COVID-19 infection [61].

3.2.2 Role of ACE2 polymorphism in COVID-19 susceptibility

ACE2 receptors are found on apical surfaces of respiratory epithelial cells serve a critical role in SARS-CoV 2 spread and pathogenesis; therefore, the expression of these receptors can affect the susceptibility of SARS-CoV2 infection in individuals [62]. ACE2 is a crucial player in the Renin-Angiotensin-Aldosterone system (RAAS) as it hydrolyzes a potent vasoconstrictor, angiotensin II (Ang II), into lung-protective Ang-(1-7), which further binds to type 1 angiotensin II (AT1) receptor causing a cascade of inflammation, vasoconstriction, fibrosis, and thrombosis reactions. Ang-(1-7) binds to type 2 angiotensin II (AT2) receptor driving increased vasodilation and reduced inflammation, fibrosis, and thrombosis [62-63]. Therefore, ACE and ACE2 have opposite functions, and their gene polymorphisms are related to severe hypertension and cardiovascular illness, which are directly associated with severe clinical COVID-19 complications [62,64,65]. The male population could be more affected due to the presence of the ACE2 gene on the X chromosome. According to studies, single nucleotide polymorphisms (SNPs) in ACE2 receptors are associated with cardiovascular illnesses and low expression of ACE2 male carriers, directly linked to severe SARS-CoV-2 infection [66-68].

3.2.3 Role of SARS-CoV2 mutants on COVID-19 susceptibility

Mutations in the novel coronavirus and variations in human genetics can impact the clinical manifestations of COVID-19 to a greater extend [69]. The alterations in the viral genome can affect viral entry into the host cell, viral life cycle, transmission, and the onset of the host’s immune response. However, the mutation rate is much lesser in novel coronavirus strains than older coronaviruses, confirming the high conservation of S protein in SARS-CoV-2 [70]. But a variant of SARS-CoV-2 with a mutation in S protein (D614G) significantly increased the transmissibility of COVID-19 infection, which has spread throughout the United States and Europe since its emergence in Wuhan [71]. Further, mutation R4081 in RBD from India leads to a weaker affinity between SARS-CoV-2 and ACE2, which could be an obstacle in vaccine designing for COVID-19 [70]. A research study conducted
in the United States concluded the higher mutation rates in the SARS-CoV-2 genome in females compared to the male population [72].

<table>
<thead>
<tr>
<th>Human Genes</th>
<th>COVID-19 susceptibility</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ABO Blood Type</strong></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>Increased</td>
</tr>
<tr>
<td>O</td>
<td>Decreased</td>
</tr>
<tr>
<td><strong>Genetic Sex</strong></td>
<td></td>
</tr>
<tr>
<td>Males (XY)</td>
<td>Increased</td>
</tr>
<tr>
<td><strong>HLA Variant</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Increased</td>
</tr>
<tr>
<td><strong>Receptor polymorphism</strong></td>
<td></td>
</tr>
<tr>
<td>ACE2, TMPRSS2, CD14</td>
<td>Increased</td>
</tr>
</tbody>
</table>

4. Conclusion

Studies of SARS-CoV-2, as well as human genetics with the help of tools like genome-wide association studies (GWAS), provide insights into the host genetic, allelic polymorphisms including viral SNPs that are critical for understanding pathophysiology, the severity of clinical outcomes in patients and designing vaccination for COVID-19 pandemic [69]. The effect of some of the essential genetic factors on susceptibility to COVID 19 infection is highlighted in table 1. From recent studies on human genetics, it is concluded that inborn errors in genes can make some people more susceptible to severe forms of SARS-CoV-2 infection. Clinical trial results published by Mike Bogetofte predicted fewer chances of contracting SARS-CoV-2 in individuals with blood group O compared with non-O individuals [61]. The absence of stochastic X chromosome inactivation in males due to one copy of the X chromosome is held responsible for a higher rate of SARS-CoV-2 infection [73-74]. Thus, different clinical outcomes observed with COVID-19 patients could be associated with factors like older age, sex of a person, associated medical complications, ABO type, genetic variations, host-virus binding interactions, prior exposure to similar viruses, and much more. However, the genetic and biochemical aspects of the host and novel COVID-19 virus is yet to be uncovered. Additionally, SNPs in host receptors ACE2, TMPRSS2, CD14, along with human immune factors like HLA, cytokines are reported to have a significant influence on transmission and pathogenesis of SARS-CoV-2 infection [69]. Moreover, the role of several protein complexes genes such as vacuolar ATPases, retromer, P13K, ARP2/3, etc., is also studied, which highlighted the importance of multiple genes in SARS-CoV-2 pathogenesis [75]. A changed viral genome can lead to the elicitation of an entirely different set of cytokines rush, which could be potentially more lethal to the host than wild viral forms. Thus, more novel approaches for in-silico and wet-lab studies are needed to understand SARS-CoV-2 better and its respective response in humans to deal with the COVID-19 pandemic.

Declarations

**Author Contribution** Preetinder Kaur: concept, structuring, writing, figure, and editing. Pawanjot Kaur: writing, editing, and figure.

**Funding** Not applicable

**Acknowledgments** We acknowledge the efforts of all research scholars, scientists, and doctors for their contribution to the better understanding of the COVID-19 pandemic.

**Conflict of Interest** The authors declare no conflict of interest.

**References**


