

REVIEW

Open Access



# Human genetic factors associated with susceptibility to SARS-CoV-2 infection and COVID-19 disease severity

Cleo Anastassopoulou<sup>1\*</sup> , Zoi Gkizarioti<sup>2</sup>, George P. Patrinos<sup>2,3,4</sup>  and Athanasios Tsakris<sup>1</sup> 

**Abstract: Background:** The emergence of the novel coronavirus in Wuhan, Hubei Province, China, in December 2019 marked the synchronization of the world to a peculiar clock that is counting infected cases and deaths instead of hours and minutes. The pandemic, highly transmissible severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has indeed caused considerable morbidity and mortality and drastically changed our everyday lives. As we continue to become acquainted with the seventh coronavirus known to infect our species, a number of its characteristics keep surprising us. Among those is the wide spectrum of clinical manifestations of the resulting coronavirus disease 2019 (COVID-19), which ranges from asymptomatic or mildly symptomatic infections to severe pneumonia, respiratory failure, and death.

**Main body:** Data, now from patient populations, are beginning to accumulate on human genetic factors that may contribute to the observed diversified disease severity. Therefore, we deemed it prudent to review the associations between specific human genetic variants and clinical disease severity or susceptibility to infection that have been reported in the literature to date (at the time of writing this article in early August 2020 with updates in mid-September). With this work, we hope (i) to assist the fast-paced biomedical research efforts to combat the virus by critically summarizing current knowledge on the potential role of host genetics, and (ii) to help guide current genetics and genomics research towards candidate gene variants that warrant further investigation in larger studies. We found that determinants of differing severity of COVID-19 predominantly include components of the immune response to the virus, while determinants of differing susceptibility to SARS-CoV-2 mostly entail genes related to the initial stages of infection (i.e., binding of the cell surface receptor and entry).

**Conclusion:** Elucidating the genetic determinants of COVID-19 severity and susceptibility to SARS-CoV-2 infection would allow for the stratification of individuals according to risk so that those at high risk would be prioritized for immunization, for example, if or when safe and effective vaccines are developed. Our enhanced understanding of the underlying biological mechanisms could also guide personalized therapeutics. Such knowledge is already beginning to provide clues that help explain, at least in part, current epidemiologic observations regarding the typically more severe or benign disease course in older males and children, respectively.

**Keywords:** Polymorphisms, Allelic variation, Genetic predisposition, Genotype, Clinical outcome, SARS-CoV-2, COVID-19, Coronavirus, X chromosome, TLR7

\* Correspondence: [cleo@med.uoa.gr](mailto:cleo@med.uoa.gr)

<sup>1</sup>Laboratory of Microbiology, Medical School, National and Kapodistrian University of Athens, 75 Mikras Asias Street, 11527 Athens, Greece  
Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Ever since the emergence of the novel coronavirus in Wuhan, Hubei Province, China, in December 2019, the world has been synchronized to a peculiar clock that is not counting hours and minutes, but infected cases and deaths. The pandemic for the past 4 months severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has indeed caused considerable morbidity, mortality, and economic damage and drastically changed our everyday lives. As we continue to learn more about the seventh coronavirus known to infect our species, a number of baffling factors about it keep surprising us. Among those is the wide spectrum of clinical manifestations that characterizes the resulting coronavirus disease 2019 (COVID-19), which ranges from lack of symptoms (as perceived by individuals because measurable signs of infection such as viral RNA may be detectable from various anatomic sites, including the naso- or oro-pharynx and the gastrointestinal tract, even of asymptomatic persons), or mild symptoms of the upper respiratory tract to severe pneumonia with acute respiratory distress syndrome (ARDS) and death [1, 2]. Advanced age, male gender, and comorbidities, such as cardiovascular disease or hypertension, and also diabetes and obesity have been identified as risk factors for more severe COVID-19 [3, 4]. Still, which and to what extent specific genetic factors may account for the predisposition of individuals to develop severe disease or to contract the infection remains unclear.

Herein, we summarize and comment on the associations between specific human genetic variants and clinical disease severity or susceptibility to infection that have been reported in the literature to date (at the time of writing this article in early August 2020 with updates as of mid-September 2020). Our main objectives were twofold: (i) to assist current biomedical research efforts to combat the virus by presenting in a comprehensive manner what is known on the subject, and (ii) to help guide current genetics and genomics research towards candidate gene variants that warrant further investigation in larger studies. In turn, this knowledge could help stratify individuals according to risk, thus allowing for the prioritization of those at greater risk for protection, for instance, if or when safe and effective prophylactic vaccines or therapeutics are developed. Furthermore, we aspire that this accumulated knowledge could ultimately help open new avenues, ideally for innovative personalized treatments. Preliminary information of possible genetic determinants of COVID-19 was highlighted in a mini-review by Godri Pollitt et al. that was published in May [5]. To the best of our knowledge, this is the first comprehensive review on associations between specific genetic loci or genes, which may well differ based on their geographical distribution, and COVID-19. While

revising this article, an exhaustive review of the literature was published by LoPresti et al. on host genetic factors related to coronaviruses that affect different species, including humans [6].

## Main text

### Literature search

We searched PubMed/MEDLINE for all English-language original articles or reviews reporting on potential associations between human genetic factors and susceptibility to SARS-CoV-2 infection or COVID-19 severity, up to August 12, 2020 (with updating as of September 11, 2020, during the revision of the manuscript). Articles on preprint servers (i.e., BioRxiv and MedRxiv) were included in our search. The search was performed using all combinations of terms related to the novel coronavirus and the disease (e.g., “SARS-CoV-2,” “2019-nCoV,” and “COVID-19”) on the one hand, and terms concerning susceptibility to infection or disease severity (e.g., “polymorphisms,” “allelic variation,” “genetic predisposition,” “genotype,” “clinical outcome”) as well as the names of individual genes in which relevant polymorphisms were found (e.g., “TLR7,” “ACE2”), on the other. All study types and countries of origin, irrespective of the size of patients’ cohorts or whether positive or negative results were reported, were included in our analysis. Reference was not made to social and economic inequalities that tend to disproportionately affect populations of specific ancestries or ethnic backgrounds and increase risk mainly due to limited access to health care.

## Results

A summary of reported associations between human genes and differing severity of COVID-19 or differing susceptibility to SARS-CoV-2 that are discussed in detail below is presented in Table 1.

### Determinants of differing severity of COVID-19

#### *Locus 3p21.31 spanning genes SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6, and XCR1*

A genome-wide association study (GWAS) analyzed 8,582,968 single nucleotide polymorphisms (SNPs) from 1980 patients severely afflicted by COVID-19 from the Italian and Spanish epicenters of the pandemic in Europe and conducted a meta-analysis of the two case-control panels that included 835 patients and 1255 control participants from Italy and 775 patients and 950 control participants from Spain, respectively [7]. The study did not detect significant associations of severe disease defined as respiratory failure with a single causative gene, but rather with a multi-gene cluster on chromosome 3. Among those were the *LZTFL1* that is strongly expressed in human lung cells and the sodium-imino acid (proline) transporter 1 (SIT1)-encoding

**Table 1** Summary of reported associations between human genes and COVID-19

Gene(s)	Polymorphism(s)	Chromosome location	Reported COVID-19 associations	Reference(s)
<i>ABO</i>	rs657152	9q34.2	Higher risk of infection for blood group A vs. non-A (OR 1.45, 95% CI 1.20–1.75, $P = 1.48 \times 10^{-4}$ ) and lower risk of infection for blood group O vs. non-O (OR 0.65, 95% CI 0.53–0.79, $P = 1.06 \times 10^{-5}$ )	[7, 38, 39]
<i>ACE2</i>	p.Arg514-Gly	Xp22.2	Cardiovascular and pulmonary conditions in the African/African-American population by altering AGT-ACE2 pathway	[49]
<i>ApoE</i>	rs429358-C-C (e4e4)	19q13.32	Severe disease independently of pre-existing dementia, cardiovascular disease, and type 2 diabetes	[32]
<i>HLA</i>	B*46:01 and B*15:03	6p21.33	Vulnerable to disease for <i>HLA-B*46:01</i> and cross-protective T cell-based immunity for <i>HLA-B*15:03</i>	[15]
<i>IFITM3</i>	rs12252-C/C	11p15.5	Mild-to-moderate disease requiring hospitalization	[35]
<i>SLC6A20</i> , <i>LZTFL1</i> , <i>CCR9</i> , <i>FYCO1</i> , <i>CXCR6</i> , <i>XCR1</i>	rs11385942-GA	3p21.31	Severe disease (respiratory failure) (OR 1.77, 95% CI 1.48–2.11, $P = 1.15 \times 10^{-10}$ )	[7]
<i>TLR7</i>	g.12905756_12905759del and g.12906010G>T	Xp22.2	Severe disease	[19]
<i>TMEM189-UBE2V1</i>	rs6020298-A	20q13.13	Severe disease	[17]
<i>TMPRSS2</i>	p.Val160Met (rs12329760)	21q22.3	Increased susceptibility to disease and for risk factors, e.g., cancer	[49]

*SLC6A20* that functionally interacts with the cell-surface receptor of the novel coronavirus, angiotensin-converting enzyme 2 (*ACE2*) [8, 9]. The remaining genes of the locus encode chemokine receptors of the C-C and CXC families which are defined based on the position of the two conserved cysteine residues in the N-termini of these members of the superfamily of G protein-coupled receptors (GPCRs); chemokines control cell migration associated with immune surveillance by trafficking effector cells to sites of infection and inflammation [10]. *CXCR6*, for instance, regulates the specific location of lung-resident memory CD8<sup>+</sup> T cells throughout the sustained immune response to airway pathogens, including influenza viruses, while flanking genes (e.g., *CCR1* and *CCR2*) also have relevant functions [11, 12]. The risk allele GA of the lead SNP rs11385942 was associated with reduced expression of *CXCR6* and increased expression of *SLC6A20* [7].

Associations within the same locus of ~ 50 kb at chromosome 3p21.31 that may have been inherited from Neandertals [13] were also reported by preliminary results from the COVID-19 Host Genetics Consortium that included both mildly and severely affected patients [14]. Ideally, the comparison of such a GWAS analysis that highlights associations rather than demonstrating causality for a specific genetic variant would be between hospitalized patients with severe disease and infected mildly symptomatic or asymptomatic individuals. Instead, population-based controls were included of people from the same geographic region who were not hospitalized, and, importantly, whose exposure to the virus is unknown. However, given the unprecedented conditions

of the pandemic under which this research was conducted—and in such a short period of time—any potential limitations in design cannot diminish its significance.

Thus, the ~ 1.5 times significantly higher frequency of the risk allele among hospitalized patients who received mechanical ventilation than among those who received supplemental oxygen only, as well as the younger age of patients who were homozygous for the risk allele than patients who were heterozygous or homozygous for the non-risk allele, support the involvement of the chromosome 3p21.31 locus in COVID-19 severity and suggest that some of the protective benefits of younger age are trumped by a double dose [7]. Further studies are needed to explore the underlying biomechanistic details of these associations.

#### ***HLA-B\*46:01 and HLA-B\*15:03***

Potential associations between the genetic variability in major histocompatibility complex (MHC) class I genes (human leukocyte antigen [HLA] A, B, and C) and the susceptibility to SARS-CoV-2 and severity of COVID-19 were investigated by Nguyen et al. who performed a comprehensive in silico analysis of viral peptide-MHC class I binding affinity across 145 HLA-A, HLA-B, and HLA-C genotypes for all SARS-CoV-2 peptides [15]. The fewest predicted binding peptides for SARS-CoV-2 were found for *HLA-B\*46:01*, suggesting that individuals harboring this allele may be particularly vulnerable to COVID-19, as previously shown for SARS-CoV-1 [16]. Conversely, the greatest capacity to present highly conserved SARS-CoV-2 peptides that are shared among common human coronaviruses were detected for *HLA-*

*B\*15:03*, suggesting that this allelic variant could enable cross-protective T cell-based immunity [15]. Meanwhile, preliminary results from a study from China indicated that the *HLA-A\*11:01*, *-B\*51:01*, and *-C\*14:02* alleles significantly predispose patients for the worst clinical outcome [17]. Another recently published report from Italy added three more *HLA* alleles, *HLA-DRB1\*15:01*, *-DQB1\*06:02*, and *-B\*27:07*, to the list that may predispose for a less favorable outcome by analyzing a group of 99 patients affected by a severe or extremely severe form of COVID-19 compared to a reference group of 1017 individuals [18]. Further studies are needed to clarify the role of single *HLA* alleles in COVID-19 severity.

#### The X-chromosomal *TLR7*

Rare, putative loss-of-function variants of the X-chromosomal *Toll-Like Receptor 7 (TLR7)* gene that were associated with impaired type I and II IFN responses were identified by rapid clinical whole-exome sequencing of the patients and segregation in available family members [19]. The study involved a case series of four young males (two pairs of brothers younger than 35 years from two unrelated families) with severe COVID-19 from the Netherlands [19]. A maternally inherited 4-nucleotide deletion (c.2129\_2132del; p.[Gln710Argfs\*18]) was identified in members of the first family that included a brother who died of the infection; the affected members of the second family carried a missense variant (c.2383G>T; p.[Val795Phe]) in *TLR7*. Downstream type I interferon (IFN) signaling was transcriptionally downregulated in primary peripheral blood mononuclear cells (PBMCs) from the patients compared with family members and controls, as measured by the significantly decreased mRNA expression of *IRF7*, *IFNB1*, and *ISG15* on stimulation with the *TLR7* agonist imiquimod. The production of the type II IFN- $\gamma$  was also decreased in patients in response to stimulation with imiquimod. Thus, *TLR7*, which is subject to selective evolutionary constraints against predicted loss of function [20], seems to be an essential component of the innate immunity against coronaviruses, including SARS-CoV-2 [21–23]. Nevertheless, SARS-CoV-2 presumably induces a lower antiviral transcriptional response, marked by low type I IFN levels and elevated chemokine expression, compared to other respiratory viruses [24].

Apart from discovering a new genetic link that could open a novel avenue for the exploration of potential treatments, this study possibly also provides an explanation for the observed trend of higher fatalities from COVID-19 in men than in women. Several immune-related genes and regulatory elements, such as non-coding micro RNA (miRNA), that are extensively involved in both the innate and adaptive immune responses are found in the X chromosome, the

complexity of statistical analysis of which poses an obstacle to its inclusion in genome-wide and candidate association studies of infectious diseases [25]. However, apart from the influence of sex hormones and socioeconomic and behavioral factors, the X chromosome, X-linked genes, and X chromosome inactivation mechanisms contribute to sexual dimorphism. In other words, males are haploid for the X chromosome that they inherit from their mothers and, therefore, any abnormality in genes on the X chromosome is more likely to be expressed phenotypically and have more pronounced immunological consequences. In contrast, the fact that females carry both a maternal and a paternal X chromosome, and are thus functional mosaics for X-linked genes mainly due to X chromosome inactivation, could contribute to an immunological advantage for females in many infections, possibly including COVID-19. Females tend to have higher antibody responses, but also more adverse reactions in response to a number of vaccines and they are more prone to developing autoimmunity, including systemic lupus erythematosus [26–28]. Testosterone acts as an immune suppressor through the upregulation of anti-inflammatory cytokines (interleukin-10, IL-10), while estrogen enhances the immune system by upregulating pro-inflammatory cytokines (tumor necrosis factor alpha (TNF $\alpha$ )) [27, 29].

#### *ApoE*

Pre-existing dementia was also identified as a major risk factor (odds ratio [OR] = 3.07, 95% CI 1.71 to 5.50) for COVID-19 severity in older adults in the UK Biobank [30]. The *Apolipoprotein E (APOE)* gene has three major isoforms, APOE2, APOE3, and APOE4, that are encoded by e2, e3, and e4 alleles, respectively, which in turn are haplotypes of the SNPs rs429358 and rs7412 on chromosome 19 (T-T, C-T, and C-C, correspondingly) [31]. The *ApoE* e4e4 homozygous genotype was found to increase the risk of severe COVID-19, independently of pre-existing dementia, cardiovascular disease, and type 2 diabetes [32]. In addition to affecting lipo-protein function and subsequent cardio-metabolic diseases, the *ApoE* e4 allele moderates macrophage pro-/anti-inflammatory phenotypes [33]. *ApoE* is one of the highly co-expressed genes in type II alveolar cells in the lungs, where the ACE2 receptor that SARS-CoV-2 uses for cell entry is highly expressed [32]. Further studies could help unravel the biological mechanisms linking *ApoE* genotypes to COVID-19 severity.

#### *IFITM3*

The SNP rs12252-C/C in the interferon-induced transmembrane protein 3-encoding gene *IFITM3* that is linked to severe influenza [34] was detected in a patient

from Wuhan, China, with mild-to-moderate COVID-19 that required hospitalization but recovered [35]. This was a finding from a single case report; however, as the prevalence of this SNP is 26.5% in the Chinese population (the 1000 Genomes Project) [36], further investigation of the *IFITM3*-rs12252-C/C allele in larger cohorts of people with COVID-19 may be worth pursuing.

#### ***TMEM189-UBE2V1***

A preprint of the first host genetic study in China analyzed a total of 22.2 million genetic variants by deeply sequencing 332 COVID-19 patients from the Shenzhen Third People's Hospital who had been categorized by varying levels of severity to asymptomatic, mild, moderate, severe, or critical ill, after the correction of potential confounding factors. The most significant gene loci associated with disease severity were located in *TMEM189-UBE2V1* that is involved in the IL-1 signaling pathway [17].

#### **Determinants of differing susceptibility to SARS-CoV-2**

##### ***ABO blood groups***

As with SARS-CoV-1 [37], ABO blood groups have been implicated in susceptibility to SARS-CoV-2 infection as well. In particular, blood groups A and O have been associated with minimally increased and decreased risk of acquiring COVID-19 than non-A and non-O groups, respectively [7, 38, 39]. The biologic mechanisms underlying the potentially differing susceptibility to infection by ABO blood group may stem directly from the ABO blood group and the development of neutralizing antibodies against protein-linked N-glycans, for example, [40], or, indirectly, from other mediated effects that could include the stabilization of von Willebrand factor [41–45]. Zietz and Tatonetti [46] also found evidence for association between ABO and Rh blood groups, with depletion of O and enrichment of B blood groups among SARS-CoV-2-positive patients; moreover, Rh(D)-positive blood types were found to be associated with SARS-CoV-2 infection and death following infection, without confounding due to demographics or other known risk factors [46]. Nonetheless, the observed associations with blood types were not corroborated by the COVID-19 Host Genetics Consortium [14], suggesting they may be circumstantial, stemming from factors unrelated to COVID-19.

##### ***ACE2 and TMPRSS2***

SARS-CoV-2 infection of susceptible cells is dependent on at least two host cell factors: the receptor ACE2 for cell entry, as already mentioned, and the transmembrane serine protease (TMPRSS2) for priming of the spike (S) protein of the virus [47]. Single-cell RNA sequencing has been used recently to

investigate the distribution of *ACE2* expression, which together with *TMPRSS2* are thought to dictate viral tissue tropism [48]. A recently published comparative genetic analysis of approximately 81,000 human genomes across different populations suggested possible associations of coding region variants of *ACE2* and *TMPRSS2* with COVID-19 susceptibility, severity, and clinical outcomes [49].

*ACE2* polymorphisms were more likely to be associated with cardiovascular and pulmonary conditions by altering the angiotensinogen (AGT)-*ACE2* interactions, such as p.Arg514-Gly in the African/African-American population [49]. Apart from the potential effects of differential polymorphisms on susceptibility and outcome in different ethnic populations, the localization of *ACE2* to the X chromosome may help explain the increased risk for males compared to females. As for the X-chromosomal *TLR7* gene, the mono-allelic versus the bi-allelic presence of this gene too may have an adverse impact on the natural history and prognosis of COVID-19 in males.

*TMPRSS2* is a key gene in prostate cancer, which, as an associated translocation, drives ETS-family oncogene expression in a large proportion of tumors [50]. The differential genetic susceptibility to COVID-19 as well as for risk factors, including those with cancer, may be explained at least in part by unique but prevalent polymorphisms, including p.Val160Met (rs12329760), an expression quantitative trait locus (eQTL) in *TMPRSS2*. The likely adverse effect of the rs12329760 polymorphism in the coding region of the *TMPRSS2* gene was corroborated by a recent study that used data of the 1000 Genome Project and web-based tools [51]. This study further identified 17 polymorphisms in the non-coding regions of *TMPRSS2* as well as an additional 31 such polymorphisms with possible functional consequences by providing binding sites for transcription factors and miRNAs, in the non-coding sequences of the following genes: three in *ACE2*, ten in *TMPRSS11A*, twelve in *neutrophil elastase (ELANE)*, and six in *cathepsin L (CTSL)* [51]. The oncogenic roles of *TMPRSS2* may be linked to poor outcomes with COVID-19 as well [52], while the localization of the gene on 21q22.3 could place individuals with Down syndrome at high risk for COVID-19 infection [49]. Interestingly, the developmental regulation of *TMPRSS2*, as suggested by the finding, by Schuler et al. [53], of the highest expression in ciliated cells and type I alveolar epithelial cells (AT1) that increased with aging in humans and mice, may provide the link of the relative protection of children from severe COVID-19. Hence, it may be worth pursuing further studies of age-related polymorphisms for *TMPRSS2*, using such cohorts as the Genetic Epidemiology Research on Adult Health and Aging (GERA) [54].

A study from Italy, the second epicenter of the pandemic after Wuhan, that examined exome and SNP-

array data from a large Italian cohort in comparison to other Europeans and East Asians, found no evidence that *ACE2* is associated with disease severity or sex bias; on the other hand, *TMPRSS2* levels and genetic variants proved to be possible candidate disease modulators in that study, prompting for rapid experimental validations on large patient cohorts [55]. Another recently published study from Italy that mined whole-exome sequencing data of 6930 individuals from five different centers participating in the Network of Italian Genomes (NIG) identified a number of *ACE2* variants with a potential impact on protein stability [56]. Among those were three common missense changes predicted to interfere with protein structure and stabilization, p.(Asn720Asp), p.(Lys26Arg), and p.(Gly211Arg), as well as the rare variants p.(Leu351Val) and p.(Pro389His), which possibly interfere with the binding of the spike of SARS-CoV-2 or with the internalization process [57]. Exome sequencing and analysis of a smaller cohort of 131 COVID-19 patients from Italy for genetic variants of *TMPRSS2*, *PCSK3*, *DPP4*, and *BSG* genes identified 17 variants [58]. The statistically more frequent alleles found in patients compared to the EUR GnomAD reference population included a missense variant (c.893G>A) in *PCSK3* and c.331G>A, c.23G>T, and c.589G>A variant alleles in *TMPRSS2* [58]. Genetic variants of the *ACE2* gene from the same cohort of 131 Italian patients were also examined in comparison to 1000 individuals (500 males and 500 females) [57]. One intronic c.439+4G>A and two missense c.1888G>C p.(Asp630His) and c.2158A>G p.(Asn720Asp) with a similar frequency in male and female were found. However, only the c.1888G>C p.(Asp630His) variant, which was nonetheless detected in just one heterozygous COVID-19 patient, showed a statistically different frequency compared to the ethnically matched population [57]. The initial study of Chinese patients that is still in preprint, showed a decreasing allele frequency of the p.Val197Met missense variant which affects the stability of the *TMPRSS2* protein, among the severely infected compared to the mildly infected patients and the general population [17].

Finally, a recent survey of the critical loci for host-viral interaction and entry into host cells (*ACE2* and *TMPRSS2*) or viral genomic RNA sensing (i.e., *TLR3/7/8*) showed that they are highly conserved in all populations, with very few non-synonymous variants [59]. The authors conclude that non-genetic factors, such as age, underlying medical conditions, and environmental risk factors like air pollution, are likely to contribute to differing susceptibility to SARS-CoV-2 infection [59]. Further studies are needed to confirm this conclusion or to offer support to detected potential associations with host genetics.

## Conclusions

The increasing availability of data from COVID-19 patient populations is allowing for potential associations to be made between specific gene loci and disease severity or susceptibility to infection. Components of the immune response to the virus seem to be principally related to the observed inter-individual variation in disease severity, while genes related to the binding of the *ACE2* cell surface receptor and entry at the initial stages of infection, largely determine the differing susceptibility to SARS-CoV-2. Such evidence-based risk assessment could lead to personalized preventive measures and therapeutic options. Perhaps the time has come for precision medicine strategies that could help us survive, if not win, this tug-of-war with the novel coronavirus.

## Abbreviations

*ACE2*: Angiotensin-converting enzyme 2; *AGT*: Angiotensinogen; *APOE*: Apolipoprotein E; *ARDS*: Acute respiratory distress syndrome; *AT1*: Type I alveolar epithelial cells; *CI*: Confidence interval; *CTSL*: Cathepsin L; *COVID-19*: Coronavirus disease 2019; *ELANE*: Neutrophil elastase; *eQTL*: Expression quantitative trait locus; *GERA*: Genetic Epidemiology Research on Adult Health and Aging; *GPCRs*: G protein-coupled receptors; *GWAS*: Genome-Wide Association Study; *HLA*: Human leukocyte antigen; *IFITM3*: Interferon-induced transmembrane protein 3; *IFN*: Interferon; *IL-10*: Interleukin-10; *MHC*: Major histocompatibility complex; *miRNA*: Micro RNA; *OR*: Odds ratio; *PBMCs*: Peripheral blood mononuclear cells; *SARS-CoV-2*: Severe acute respiratory syndrome coronavirus 2; *SNP*: Single nucleotide polymorphism; *SIT1*: Sodium-imino acid (proline) transporter 1; *S*: Spike; *TLR7*: Toll-like receptor 7; *TMPRSS2*: Transmembrane serine protease; *TNFA*: Tumor necrosis factor alpha

## Acknowledgements

Not applicable.

## Authors' contributions

CA designed and wrote the manuscript. ZG, GPP, and AT also contributed to the writing of the manuscript and made critical revisions. All authors read and approved the final manuscript.

## Funding

No specific funding was obtained for this study.

## Availability of data and materials

Not applicable.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>Laboratory of Microbiology, Medical School, National and Kapodistrian University of Athens, 75 Mikras Asias Street, 11527 Athens, Greece.

<sup>2</sup>Laboratory of Pharmacogenomics and Individualized Therapy, Department of Pharmacy, School of Health Sciences, University of Patras, Patras, Greece.

<sup>3</sup>Zayed Center of Health Sciences, United Arab Emirates University, Al-Ain, United Arab Emirates.

<sup>4</sup>College of Medicine and Health Sciences, Department of Pathology, United Arab Emirates University, Al-Ain, United Arab Emirates.

Received: 21 August 2020 Accepted: 13 October 2020  
Published online: 22 October 2020

## References

- Richardson S, Hirsch JS, Narasimhan M, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the new York City area. *JAMA*. 2020;323(20):2052–9. <https://doi.org/10.1001/jama.2020.6775>.
- Grasselli G, Zangrillo A, Zanella A, et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy region, Italy. *JAMA*. 2020;323(16):1574–81. <https://doi.org/10.1001/jama.2020.5394>.
- Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan China: a retrospective cohort study. *Lancet*. 2020;395(10229):1054–1062. [doi.org/10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3).
- Nakeshbandi M, Maini R, Daniel P, et al. The impact of obesity on COVID-19 complications: a retrospective cohort study. *Int J Obes*. 2020;44(9):1832–7. [doi.org/10.1038/s41366-020-0648-x](https://doi.org/10.1038/s41366-020-0648-x).
- Godri Pollitt KJ, Peccia J, Ko AI, et al. COVID-19 vulnerability: the potential impact of genetic susceptibility and airborne transmission. *Hum Genomics*. 2020;14(1):17. [doi.org/10.1186/s40246-020-00267-3](https://doi.org/10.1186/s40246-020-00267-3).
- LoPresti M, Beck DB, Duggal P, Cummings DAT, Solomon BD. The role of host genetic factors in coronavirus susceptibility: review of animal and systematic review of human literature. *Am J Hum Genet*. 2020;107(3):381–402. [doi.org/10.1016/j.ajhg.2020.08.007](https://doi.org/10.1016/j.ajhg.2020.08.007).
- Ellinghaus D, Degenhardt F, Bujanda L, et al. Genomewide association study of severe COVID-19 with respiratory failure. *N Engl J Med*. 2020; NEJMoa2020283. [doi.org/10.1056/NEJMoa2020283](https://doi.org/10.1056/NEJMoa2020283).
- Vuille-dit-Bille RN, Camargo SM, Emmenegger L, et al. Human intestine luminal ACE2 and amino acid transporter expression increased by ACE-inhibitors. *Amino Acids*. 2015;47(4):693–705. [doi.org/10.1007/s00726-014-1889-6](https://doi.org/10.1007/s00726-014-1889-6).
- Kuba K, Imai Y, Ohto-Nakanishi T, Penninger JM. Trilogy of ACE2: a peptidase in the renin-angiotensin system, a SARS receptor, and a partner for amino acid transporters. *Pharmacol Ther*. 2010;128(1):119–28. [doi.org/10.1016/j.pharmthera.2010.06.003](https://doi.org/10.1016/j.pharmthera.2010.06.003).
- Allen SJ, Crown SE, Handel TM. Chemokine: receptor structure, interactions, and antagonism. *Annu Rev Immunol*. 2007;25:787–820. [doi.org/10.1146/annurev.immunol.24.021605.090529](https://doi.org/10.1146/annurev.immunol.24.021605.090529).
- Wein AN, McMaster SR, Takamura S, et al. CXCR6 regulates localization of tissue-resident memory CD8 T cells to the airways. *J Exp Med*. 2019;216(12):2748–62. [doi.org/10.1084/jem.20181308](https://doi.org/10.1084/jem.20181308).
- Hickey MJ, Held KS, Baum E, Gao JL, Murphy PM, Lane TE. CCR1 deficiency increases susceptibility to fatal coronavirus infection of the central nervous system. *Viral Immunol*. 2007;20(4):599–608. [doi.org/10.1089/vim.2007.0056](https://doi.org/10.1089/vim.2007.0056).
- Zeberg H, Pääbo S. The major genetic risk factor for severe COVID-19 is inherited from Neanderthals. *Nature*. 2020. [doi.org/10.1038/s41586-020-2818-3](https://doi.org/10.1038/s41586-020-2818-3).
- COVID-19 Host Genetics Initiative. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *Eur J Hum Genet*. 2020; 28(6):715–8. [doi.org/10.1038/s41431-020-0636-6](https://doi.org/10.1038/s41431-020-0636-6).
- Nguyen A, David JK, Maden SK, et al. Human leukocyte antigen susceptibility map for severe acute respiratory syndrome coronavirus 2. *J Virol*. 2020;94(13):e00510–20. [doi.org/10.1128/JVI.00510-20](https://doi.org/10.1128/JVI.00510-20).
- Lin M, Tseng H, Trejaut JA, et al. Association of HLA class I with severe acute respiratory syndrome coronavirus infection. *BMC Med Genet*. 2003;4:9. [doi.org/10.1186/1471-2350-4-9](https://doi.org/10.1186/1471-2350-4-9).
- Wang F, Huang S, Gao H, et al. Initial whole genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility. *medRxiv*. 2020. [doi.org/10.1101/2020.06.09.20126607](https://doi.org/10.1101/2020.06.09.20126607).
- Novelli A, Andreani M, Biancolella M, et al. HLA allele frequencies and susceptibility to COVID-19 in a group of 99 Italian patients. *HLA*. 2020. [doi.org/10.1111/tan.14047](https://doi.org/10.1111/tan.14047).
- van der Made CI, Simons A, Schuurs-Hoeijmakers J, et al. Presence of genetic variants among young men with severe COVID-19. *JAMA*. 2020; 324(7):1–11. [doi.org/10.1001/jama.2020.13719](https://doi.org/10.1001/jama.2020.13719).
- Casanova J-L, Abel L, Quintana-Murci L. Human TLRs and IL-1Rs in host defense: natural insights from evolutionary, epidemiological, and clinical genetics. *Annu Rev Immunol*. 2011;29:447–91. [doi.org/10.1146/annurev-immunol-030409-101335](https://doi.org/10.1146/annurev-immunol-030409-101335).
- Cervantes-Barragan L, Züst R, Weber F, et al. Control of coronavirus infection through plasmacytoid dendritic-cell-derived type I interferon. *Blood*. 2007; 109(3):1131–7. [doi.org/10.1182/blood-2006-05-023770](https://doi.org/10.1182/blood-2006-05-023770).
- Moreno-Eutimio MA, López-Macías C, Pastelin-Palacios R. Bioinformatic analysis and identification of single-stranded RNA sequences recognized by TLR7/8 in the SARS-CoV-2, SARS-CoV, and MERS-CoV genomes. *Microbes Infect*. 2020;22(4-5):226–9. [doi.org/10.1016/j.micinf.2020.04.009](https://doi.org/10.1016/j.micinf.2020.04.009).
- Channappanavar R, Fehr AR, Zheng J, et al. IFN-I response timing relative to virus replication determines MERS coronavirus infection outcomes. *J Clin Invest*. 2019;129(9):3625–39. [doi.org/10.1172/JCI126363](https://doi.org/10.1172/JCI126363).
- Blanco-Melo D, Nilsson-Payant BE, Liu WC, et al. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell*. 2020;181(5):1036–1045.e9. [doi.org/10.1016/j.cell.2020.04.026](https://doi.org/10.1016/j.cell.2020.04.026).
- Schurz H, Salie M, Tromp G, Hoal EG, Kinnear CJ, Möller M. The X chromosome and sex-specific effects in infectious disease susceptibility. *Hum Genomics*. 2019;13(1):2. [doi.org/10.1186/s40246-018-0185-z](https://doi.org/10.1186/s40246-018-0185-z).
- Jaillon S, Berthenet K, Garlanda C. Sexual dimorphism in innate immunity. *Clin Rev Allergy Immunol*. 2019;56(3):308–21. [doi.org/10.1007/s12016-017-8648-x](https://doi.org/10.1007/s12016-017-8648-x).
- Klein SL, Marriott I, Fish EN. Sex-based differences in immune function and responses to vaccination. *Trans R Soc Trop Med Hyg*. 2015;109(1):9–15. [doi.org/10.1093/trstmh/tru167](https://doi.org/10.1093/trstmh/tru167).
- Souyris M, Mejia JE, Chaumeil J, Guéry J-C. Female predisposition to TLR7-driven autoimmunity: gene dosage and the escape from X chromosome inactivation. *Semin Immunopathol*. 2019;41(2):153–64. [doi.org/10.1007/s00281-018-0712-y](https://doi.org/10.1007/s00281-018-0712-y).
- Cutolo M, Capellino S, Sulli A, et al. Estrogens and autoimmune diseases. *Ann N Y Acad Sci*. 2006;1089:538–47. [doi.org/10.1196/annals.1386.043](https://doi.org/10.1196/annals.1386.043).
- Atkins JL, Masoli JAH, Delgado J, Et al. Preexisting comorbidities predicting COVID-19 and mortality in the UK Biobank Community Cohort. *J Gerontol A Biol Sci Med Sci*. 2020;glaa183. [doi.org/10.1093/gerona/glaa183](https://doi.org/10.1093/gerona/glaa183).
- Kuo CL, Pilling LC, Atkins JL, Kuchel GA, Melzer D. *ApoE* e2 and aging-related outcomes in 379,000 UK Biobank participants. *Aging (Albany NY)*. 2020;12(12):12222–33. [doi.org/10.18632/aging.103405](https://doi.org/10.18632/aging.103405).
- Kuo CL, Pilling LC, Atkins JL, Et al. *APOE* e4 genotype predicts severe COVID-19 in the UK Biobank community cohort. *J Gerontol A Biol Sci Med Sci*. 2020; glaa131. [doi.org/10.1093/gerona/glaa131](https://doi.org/10.1093/gerona/glaa131).
- Tudorache IF, Trusca VG, Gafencu AV. Apolipoprotein E - a multifunctional protein with implications in various pathologies as a result of its structural features. *Comput Struct Biotechnol J*. 2017;15:359–65. [doi.org/10.1016/j.csbj.2017.05.003](https://doi.org/10.1016/j.csbj.2017.05.003).
- Everitt AR, Clare S, Pertel T, et al. IFITM3 restricts the morbidity and mortality associated with influenza. *Nature*. 2012;484(7395):519–23. [doi.org/10.1038/nature10921](https://doi.org/10.1038/nature10921).
- Thevarajan I, Nguyen THO, Koutsakos M, et al. Breadth of concomitant immune responses prior to patient recovery: a case report of non-severe COVID-19. *Nat Med*. 2020;26(4):453–5. [doi.org/10.1038/s41591-020-0819-2](https://doi.org/10.1038/s41591-020-0819-2).
- Wang Z, Zhang A, Wan Y, et al. Early hypercytokinemia is associated with interferon-induced transmembrane protein-3 dysfunction and predictive of fatal H7N9 infection. *PNAS*. 2014;111(2):769–74. [doi.org/10.1073/pnas.1321748111](https://doi.org/10.1073/pnas.1321748111).
- Cheng Y, Cheng G, Chui CH, et al. ABO blood group and susceptibility to severe acute respiratory syndrome. *JAMA*. 2005;293(12):1450–1. [doi.org/10.1001/jama.293.12.1450-c](https://doi.org/10.1001/jama.293.12.1450-c).
- Zhao J, Yang Y, Huang H, Et al. Relationship between the ABO blood group and the COVID-19 susceptibility. *Clin Infect Dis*. 2020;ciaa1150. [doi.org/10.1093/cid/ciaa1150](https://doi.org/10.1093/cid/ciaa1150).
- Wu Y, Feng Z, Li P, Yu Q. Relationship between ABO blood group distribution and clinical characteristics in patients with COVID-19. *Clin Chim Acta*. 2020;509:220–3. [doi.org/10.1016/j.cca.2020.06.026](https://doi.org/10.1016/j.cca.2020.06.026).
- Breiman A, Ruvén-Clouet N, Le Pendu J. Harnessing the natural anti-glycan immune response to limit the transmission of enveloped viruses such as

- SARS-CoV-2. *PLoS Pathog.* 2020;16(5):e1008556. doi.org. <https://doi.org/10.1371/journal.ppat.1008556>.
41. Comuzzie AG, Cole SA, Laston SL, et al. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. *PLoS One.* 2012;7(12):e51954. doi.org. <https://doi.org/10.1371/journal.pone.0051954>.
  42. Aziz M, Fatima R, Assaly R. Elevated interleukin-6 and severe COVID-19: a meta-analysis. *J Med Virol.* 2020. doi.org. <https://doi.org/10.1002/jmv.25948>.
  43. Naitza S, Porcu E, Steri M, et al. A genome-wide association scan on the levels of markers of inflammation in Sardinians reveals associations that underpin its complex regulation. *PLoS Genet.* 2012;8(1):e1002480. doi.org. <https://doi.org/10.1371/journal.pgen.1002480>.
  44. Franchini M, Crestani S, Frattini F, Sissa C, Bonfanti C. ABO blood group and von Willebrand factor: biological implications. *Clin Chem Lab Med.* 2014; 52(9):1273–6. doi.org. <https://doi.org/10.1515/cclm-2014-0564>.
  45. Murray GP, Post SR, Post GR. ABO blood group is a determinant of von Willebrand factor protein levels in human pulmonary endothelial cells. *J Clin Pathol.* 2020;73(6):347–9. doi.org. <https://doi.org/10.1136/jclinpath-2019-206182>.
  46. Zietz M, Tatonetti NP. Testing the association between blood type and COVID-19 infection, intubation, and death. medRxiv. 2020;2020.04.08. 20058073. doi.org. <https://doi.org/10.1101/2020.04.08.20058073>.
  47. Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell.* 2020;181(2):271–280.e8. doi.org. <https://doi.org/10.1016/j.cell.2020.02.052>.
  48. Zou X, Chen K, Zou J, Han P, Hao J, Han Z. Single-cell RNA-seq data analysis on the receptor ACE2 expression reveals the potential risk of different human organs vulnerable to 2019-nCoV infection. *Front Med.* 2020;14(2): 185–2. doi.org. <https://doi.org/10.1007/s11684-020-0754-0>.
  49. Hou Y, Zhao J, Martin W, et al. New insights into genetic susceptibility of COVID-19: an ACE2 and TMPRSS2 polymorphism analysis. *BMC Med.* 2020; 18(1):216. doi.org. <https://doi.org/10.1186/s12916-020-01673-z>.
  50. Stopsack KH, Mucci LA, Antonarakis ES, Nelson PS, Kantoff PW. TMPRSS2 and COVID-19: serendipity or opportunity for intervention? *Cancer Discov.* 2020; 10(6):779–82. doi.org. <https://doi.org/10.1158/2159-8290.CD-20-0451>.
  51. Vargas-Alarcón G, Posadas-Sánchez R, Ramírez-Bello J. Variability in genes related to SARS-CoV-2 entry into host cells (*ACE2*, *TMPRSS2*, *TMPRSS11A*, *ELANE*, and *CTSL*) and its potential use in association studies. *Life Sci.* 2020; 260:118313. doi.org. <https://doi.org/10.1016/j.lfs.2020.118313>.
  52. Yu J, Ouyang W, Chua MLK, Xie C. SARS-CoV-2 transmission in patients with cancer at a tertiary care hospital in Wuhan, China. *JAMA Oncol.* 2020;6(7): 1108–10. doi.org. <https://doi.org/10.1001/jamaoncol.2020.0980>.
  53. Schuler A, Habermann C, Plosa J, et al. Age-related expression of SARS-CoV-2 priming protease TMPRSS2 in the developing lung. bioRxiv. 2020; 2020. 05.22.111187. doi.org. <https://doi.org/10.1101/2020.05.22.111187>.
  54. Mostafavi H, Berisa T, Day FR, Perry JRB, Przeworski M, Pickrell JK. Identifying genetic variants that affect viability in large cohorts. *PLoS Biol.* 2017;15(9): e2002458. doi.org. <https://doi.org/10.1371/journal.pbio.2002458>.
  55. Asselta R, Paraboschi EM, Mantovani A, Duga S. *ACE2* and *TMPRSS2* variants and expression as candidates to sex and country differences in COVID-19 severity in Italy. *Aging (Albany NY).* 2020;12(11):10087–98. doi.org. <https://doi.org/10.18632/aging.103415>.
  56. Benetti E, Tita R, Spiga O, et al. ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population. *Eur J Hum Genet.* 2020;1–13. doi.org. <https://doi.org/10.1038/s41431-020-0691-z>.
  57. Novelli A, Biancolella M, Borgiani P, et al. Analysis of ACE2 genetic variants in 131 Italian SARS-CoV-2-positive patients. *Hum Genomics.* 2020;14(1):29. doi.org. <https://doi.org/10.1186/s40246-020-00279-z>.
  58. Latini A, Agolini E, Novelli A, et al. COVID-19 and genetic variants of protein involved in the SARS-CoV-2 entry into the host cells. *Genes (Basel).* 2020; 11(9):E1010. doi.org. <https://doi.org/10.3390/genes11091010>.
  59. Lee IH, Lee JW, Kong SW. A survey of genetic variants in SARS-CoV-2 interacting domains of ACE2, TMPRSS2 and TLR3/7/8 across populations. *Infect Genet Evol.* 2020;85:104507. doi.org. <https://doi.org/10.1016/j.meegid.2020.104507>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

