




## REVIEW ARTICLE OPEN

## Human genetic basis of coronavirus disease 2019

Hao Deng<sup>1,2,3,4</sup> , Xue Yan<sup>1,2,3</sup> and Lamei Yuan<sup>1,2,3,4</sup>

Coronavirus disease 2019 (COVID-19) caused by a novel coronavirus, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has resulted in considerable morbidity and mortality worldwide. COVID-19 incidence, severity, and mortality rates differ greatly between populations, genders, ABO blood groups, human leukocyte antigen (HLA) genotypes, ethnic groups, and geographic backgrounds. This highly heterogeneous SARS-CoV-2 infection is multifactorial. Host genetic factors such as variants in the angiotensin-converting enzyme gene (*ACE*), the angiotensin-converting enzyme 2 gene (*ACE2*), the transmembrane protease serine 2 gene (*TMPRSS2*), along with HLA genotype, and ABO blood group help to explain individual susceptibility, severity, and outcomes of COVID-19. This review is focused on COVID-19 clinical and viral characteristics, pathogenesis, and genetic findings, with particular attention on genetic diversity and variants. The human genetic basis could provide scientific bases for disease prediction and targeted therapy to address the COVID-19 scourge.

Signal Transduction and Targeted Therapy (2021)6:344

; <https://doi.org/10.1038/s41392-021-00736-8>

## INTRODUCTION

Coronavirus disease 2019 (COVID-19) is this century's third plague and was declared as the sixth international concerned public health emergency by the World Health Organization (WHO) on 30 January 2020.<sup>1,2</sup> The responsible pathogen is a previously unknown RNA coronavirus.<sup>1,3,4</sup> It was designated as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the International Committee on Taxonomy of Viruses.<sup>5</sup> As of 17 May 2021, the COVID-19 pandemic has resulted in 162,494,817 cases and 3,494,424 deaths worldwide (<https://covid19.who.int/>). COVID-19 is highly heterogeneous and its severity may relate to multiple factors including health care, quarantine effectiveness, governmental policies, societal norms of behavior, economics, cultural practices, climate, pollution, and viral characteristics, as well as host-associated factors.<sup>6–10</sup> In the aspect of host-associated factors, in addition to age (>60 years), initial health status, pre-existing diseases, smoking history, and previous vaccinations, individual genetic basis contributes to individual susceptibility, severity, and outcomes of COVID-19.<sup>7,11,12</sup> Classical twin studies indicated 31% heritability for predicted COVID-19.<sup>13</sup> Human genetic basis may implicate in significant diversities of COVID-19 among populations with different genders, ABO blood groups, human leukocyte antigen (HLA) genotypes, ethnic groups, and geographic backgrounds.<sup>6,14–16</sup> Several gene variants related to gene expression and protein function changes were reported as explaining the individual susceptibility, severity, and outcomes.<sup>12,17</sup>

In this review, clinical and viral characteristics, pathogenesis, and the human genetic basis associated with COVID-19 are investigated. Focus is on the protective and risk effects of variants in related genes such as the angiotensin-converting enzyme gene (*ACE*), the angiotensin-converting enzyme 2 gene (*ACE2*),

the transmembrane protease serine 2 gene (*TMPRSS2*), and ABO blood groups and HLA genotypes (Table 1 and Fig. 1).

## CLINICAL CHARACTERISTICS

The COVID-19 clinical spectrum is heterogeneous (Fig. 2) and ranges from asymptomatic (~5.4–15.0%), mild–moderate (~50.0%), severe (~13.8–16.0%) to critical (~4.0–25.6%) status.<sup>18–22</sup> Typical symptoms in most patients are mild and nonspecific, including fever (~72.4–91.3%), cough (~53.8–68.6%), smell dysfunction (~59.9%), taste dysfunction (~57.5%), fatigue (~25.0–51.0%), dyspnea (~12.3–30.4%), myalgia (~15.3–28.5%), expectoration (~23.0–28.2%), chest discomfort (~14.9–19.3%), anorexia (~17.1%), sore throat or pharyngalgia (~11.1–16.2%), chill (~15.0%), headache (~9.4–14.0%), dizziness or confusion (~7.6–9.2%), rhinorrhoea (~3.5–9.2%), diarrhea (~4.8–8.4%), nausea or vomiting (~3.6–6.5%), abdominal pain (~5.1%), nasal congestion (~1.8–4.9%), and hemoptysis (~2.0%).<sup>2,3,19,21–31</sup> Some patients (usually those with advanced age and comorbidities) may rapidly progress to viral pneumonia, life-threatening acute respiratory distress syndrome, or multiple organ failures.<sup>32,33</sup> In addition to the respiratory tract and lungs being primarily affected, other organs and systems such as the heart, blood vessels, gastrointestinal tract, liver, kidneys, skin, and nervous systems, can be adversely involved.<sup>34,35</sup> Common laboratory abnormalities in COVID-19 patients are decreased albumin (~43.0–60.6%), increased C-reactive protein (~44.3–87.0%), D-dimer (~29.3–48.0%), aspartate aminotransferase (~18.6–47.0%), glucose (~45.0%), procalcitonin (~18.6–36.0%), creatine kinase (~10.8–32.0%), troponin I/troponin T (~29.0%), alanine aminotransferase (~14.2–28.9%), total bilirubin (~10.7–14.3%), and creatinine (~3.1–11.0%), and decreased or

<sup>1</sup>Health Management Center, the Third Xiangya Hospital, Central South University, Changsha, China; <sup>2</sup>Center for Experimental Medicine, the Third Xiangya Hospital, Central South University, Changsha, China; <sup>3</sup>Disease Genome Research Center, Central South University, Changsha, China and <sup>4</sup>Department of Neurology, the Third Xiangya Hospital, Central South University, Changsha, China

Correspondence: Hao Deng (hdeng008@yahoo.com)

These authors contributed equally: Hao Deng, Xue Yan.

Received: 11 April 2021 Revised: 28 July 2021 Accepted: 8 August 2021

Published online: 20 September 2021

**Table 1.** Summary of human genes associated with COVID-19

Locus	Gene(s) or genotype	Variant(s) or allele(s) or haplotype(s)			References	
		Risk	Protective	Uncertain		
2q24.2	<i>IFIH1</i>	/	rs1990760	/	102,103	
3p21.31	<i>SLC6A20</i>	/	/	/	104–108	
		<i>LZTFL1</i>	rs11385942	/		/
		<i>FYCO1</i>	/	/		/
		<i>CXCR6</i>	/	/		/
		<i>XCR1</i>	/	/		/
		<i>CCR9</i>	/	/		/
6p21.3	HLA	HLA-A*11	HLA-A*02:02	HLA-A*02:01	9,15,16,18,97,115,117–127,130	
		HLA-A*25	HLA-A*02:03			
		HLA-A*25:01	HLA-A*02:05			
		HLA-A*25:02	HLA-A*02:06			
		HLA-B*08	HLA-A*02:09			
		HLA-B*15:01	HLA-A*02:11			
		HLA-B*15:27	HLA-A*02:12			
		HLA-B*27:07	HLA-A*02:22			
		HLA-B*44	HLA-A*02:24			
		HLA-B*46:01	HLA-A*02:35			
		HLA-B*51	HLA-A*02:40			
		HLA-B*54:01	HLA-A*11:01			
		HLA-B*55:01	HLA-A*24:02			
		HLA-B*55:07	HLA-B*14			
		HLA-B*55:12	HLA-B*15:03			
		HLA-B*56:01	HLA-B*18			
		HLA-C*01	HLA-B*49			
		HLA-C*01:02	HLA-B*52:01			
		HLA-C*03	HLA-C*12:02			
		HLA-C*04:01	HLA-C*12:03			
		HLA-C*05	HLA- DRB1*12:01			
		HLA-DRB1*01:01	HLA- DPB1*03:01			
		HLA-DRB1*14:04	HLA-A*02:01- B*18:01- C*07:01- DRB1*11:04			
HLA-DRB1*15:01	HLA-A*02:05- B*58:01- DRB1*08:01					
HLA-DQA1*01:01	HLA-A*02:05- B*58:01- C*07:01					
	HLA-DQA1_509					
	HLA-DQB1*04					
	HLA-DQB1*06:02					
	HLA-A*11:01-B*51:01-C*14:02					
	HLA-A*01:01-B*08:01-C*07:01-DRB1*03:01					
9q34.2	<i>ABO</i>	rs495828	/	/	104,135,137–139	
		rs8176746				
		rs657152				
		rs8176746–rs8176740–rs495828–rs12683493				
9q34.3	<i>DPP7</i>	1-bp insertion	/	/	106	
11p15.5	<i>IFITM3</i>	rs12252	/	/	141–146	
		rs34481144				
12q24.33	<i>GOLGA3</i>	rs143359233	/	/	106	

**Table 1.** continued

Locus	Gene(s) or genotype	Variant(s) or allele(s) or haplotype(s)			References
		Risk	Protective	Uncertain	
13q12.3	<i>HMGB1</i>	/	/	/	149,150
15q26.1	<i>FURIN</i>	rs6226 rs8039305	rs4702 rs769208985	/	20,70,151
17q23.3	<i>ACE</i>	D-allele	I-allele	/	80,86,154–158
19q13.32	<i>APOE</i>	ε4ε4 genotype	/	/	160–163
21q22.3	<i>TMPRSS2</i>	rs61299115 rs4303794 rs11088551 rs8134378 rs2070788 rs464397 rs469390 rs383510 rs2070788–rs9974589–rs7364083–rs8134378 rs463727–rs34624090–rs55964536–rs734056–rs4290734– rs34783969–rs11702475–rs35899679–rs35041537	p.Asp435Tyr	rs12329760	106,165–171
Xp22.2	<i>TLR7</i>	p.Ser301Pro rs2042915990 rs200553089 rs189681811 rs147244662	/	/	181,182
Xp22.22	<i>ACE2</i>	rs4646114 rs4646115 rs4646116 rs191860450 p.Arg514Gly rs41303171	rs2285666 rs2106809 rs73635825 rs766996587 rs1448326240 rs143936283 p.Leu351Val rs961360700 rs1396769231 rs762890235 p.Arg708Trp p.Arg710Cys p.Arg710His p.Arg716Cys	/	4,8,17,151,152,165,166,168,171,183,184,191–197
Xq12	<i>AR</i>	Shorter CAG repeat	/	/	101,176,200–203

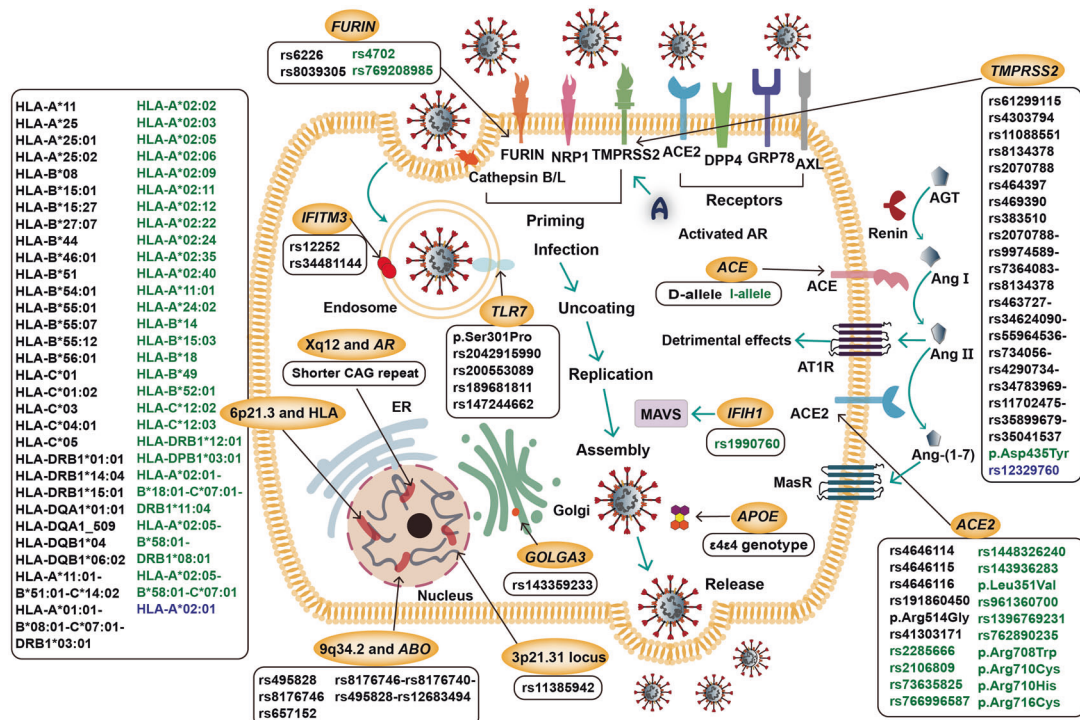
*IFIH1* the interferon induced with helicase C domain 1 gene, *SLC6A20* the solute carrier protein family 6 member 20 gene, *LZTFL1* the leucine zipper transcription factor like 1 gene, *FYCO1* the FYVE and coiled-coil domain autophagy adaptor 1 gene, *CXCR6* the C-X-C motif chemokine receptor 6 gene, *XCR1* the X-C motif chemokine receptor 1 gene, *CCR9* the C-C motif chemokine receptor 9 gene, HLA human leukocyte antigen, *ABO* the ABO, alpha 1–3-N-acetylgalactosaminyltransferase and alpha 1–3-galactosyltransferase gene, *DPP7* the dipeptidyl peptidase 7 gene, *IFITM3* the interferon induced transmembrane protein 3 gene, *GOLGA3* the golgin A3 gene, *HMGB1* the high-mobility group box 1 gene, *FURIN* the furin, paired basic amino acid cleaving enzyme gene, *ACE* the angiotensin-converting enzyme gene, *APOE* the apolipoprotein E gene, *TMPRSS2* the transmembrane protease serine 2 gene, *TLR7* the Toll-like receptor 7 gene, *ACE2* the angiotensin-converting enzyme 2 gene, *AR* the androgen receptor gene.

increased lactate dehydrogenase (~57.0% ↓ vs ~28.3–69.0% ↑), lymphocytes (~57.4–68.0% ↓ vs ~8.2% ↑), neutrophils (~3.6–9.0% ↓ vs ~25.9–31.0% ↑), leukocytes (~20.1–29.4% ↓ vs ~9.8–22.0% ↑), and platelets (~11.4–20.0% ↓ vs ~6.0% ↑).<sup>25,28,29,36</sup> Chest computed tomography (CT) scans revealed ground-glass opacity (~64.6–91.2%), lesions consistent with bilateral (~64.6–73.2%) or unilateral (~21.3–25.3%) pneumonia, vascular changes (~62.9–74.0%), air bronchogram (~39.7–50.5%), bilateral or local patchy shadowing (~43.0%, ~36.5%), halo sign (~27.3%), solid nodules (~5.2–20.7%), septal thickening (~6.5–55.0%), interstitial abnormalities (~14.1%), crazy paving pattern (~15.0–32.0%),

consolidation (~27.7–73.5%), bronchial wall thickening (~19.4–24.0%), fibrous stripes (~25.9–37.2%), spider web design (~22.3%), subpleural lines (~15.0–28.0%), pleural effusion (~3.0–7.8%), intrathoracic lymph node enlargement (~3.0–5.3%), and pericardial effusion (~3.0%).<sup>21,22,25,37–40</sup>

#### THE CAUSATIVE PATHOGEN, SARS-COV-2

SARS-CoV-2 is an enveloped, non-segmented, positive-sense single-stranded RNA virus with 29,903 nucleotides in its genome sequence containing 5' capped and 3' polyadenylated.<sup>41–43</sup>



**Fig. 1** Pathogenesis of SARS-CoV-2 and genetic variants associated with COVID-19. After the recognition of ACE2, DPP4, GRP78, and AXL receptors and the priming by TMPRSS2, Furin, and NRP1, as well as cathepsin B/L, SARS-CoV-2 enters cells and starts the replication process to assemble and release. Activated AR induces *TMPRSS2* expression. ACE/Ang II/AT1R and ACE2/Ang(1-7)/MasR axes regulate RAAS to involve in COVID-19. The risk (black), protective (green), and uncertain (blue) variants or alleles or haplotypes for COVID-19 are highlighted. SARS-CoV-2 severe acute respiratory syndrome coronavirus 2, COVID-19 coronavirus disease 2019, ACE2 angiotensin-converting enzyme 2, DPP4 dipeptidyl peptidase 4, GRP78 glucose-regulated protein-78, AXL anaxelekto, TMPRSS2 transmembrane protease serine 2, Furin furin, paired basic amino acid-cleaving enzyme, NRP1 neuropilin-1, AR androgen receptor, AGT angiotensinogen, Ang angiotensin, ACE angiotensin-converting enzyme, AT1R angiotensin II type 1 receptor, MasR Mas receptor, *TLR7* the Toll-like receptor 7 gene, *IFITM3* the interferon induced transmembrane protein 3 gene, HLA human leukocyte antigen, *GOLGA3* the golgin A3 gene, *ABO* the ABO, alpha 1-3-N-acetylglucosaminyltransferase and alpha 1-3-galactosyltransferase gene, *APOE* the apolipoprotein E gene, *IFIH1* the interferon induced with helicase C domain 1 gene, MAVS mitochondrial antiviral signaling protein, ER endoplasmic reticulum

The 5' two-thirds region of the genome is occupied by two large open reading frames (ORFs), ORF1a and ORF1b, that encodes 15–16 nonstructural proteins.<sup>44</sup> Other functional ORFs encode structural and accessory proteins.<sup>45</sup> The structural proteins include the distinctive spike (S), envelope (E), membrane (M), and nucleocapsid (N) proteins, among which the S, E, and M proteins compose the envelope structure, while the N protein encapsulates the viral genome.<sup>46</sup> SARS-CoV-2 is generally spherical with some pleomorphism and a diameter of ~60–140 nm.<sup>3</sup> The viral-enveloped lipid bilayer consists of cholesterol and phospholipids, which makes the virus susceptible to dry heat, detergents, and organic solvents.<sup>46</sup> This novel coronavirus was assigned to the genus *Betacoronavirus* in the family *Coronaviridae* of the order *Nidovirales* (Fig. 3).<sup>18,42,47</sup> SARS-CoV-2 has a 96.2% identity throughout the genome to RaTG13, a bat-borne coronavirus in *Rhinolophus affinis*.<sup>48</sup> Droplet, aerosol, contact, fecal-oral and transplacental transmissions are documented human-to-human transmission routes.<sup>49–53</sup> Small droplets with SARS-CoV-2 can travel tens of meters in favorable atmospheric conditions and remain viable and infectious from 3 h to days.<sup>32,54,55</sup> Patients mildly affected and asymptomatic carriers constituting the majority of COVID-19 cases are thought to be primarily responsible for the spread of SARS-CoV-2.<sup>10,56</sup>

As of 17 May 2021, the worldwide circulating SARS-CoV-2 variants mainly include the B.1.1.7 (63%), B.1.617.2 (22%), P.1 (6%), B.1.526 (2%), and others (<https://covid19dashboard.regeneron.com>). Lineage B.1.1.7 first detected in the United Kingdom in September 2020 has 21 characteristic mutations and exists in comparative transmission effectivity.<sup>57</sup> Among the S protein

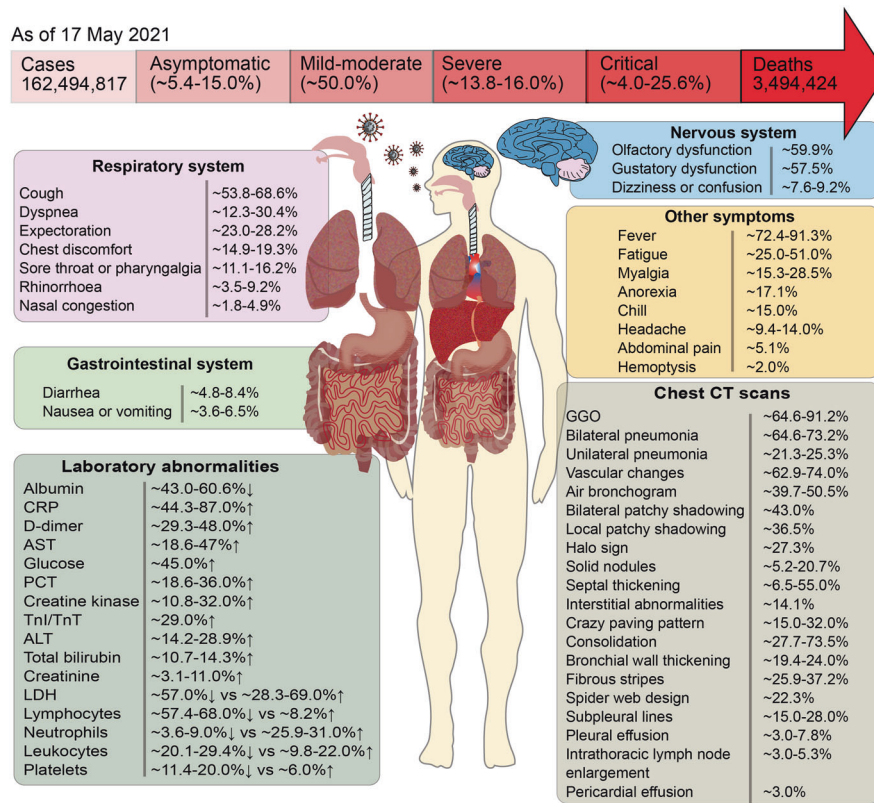
mutations, the N501Y substitution would change the receptor-binding domain conformation and may slightly increase 18% of fatality risk. Mutation del69–70, considered to be responsible for stronger viral transmissibility, may cause S-gene target failure and produce a positive result for other targets at real-time reverse transcription-polymerase chain reaction assays, which could be a proxy for diagnosing B.1.1.7 infections.<sup>58,59</sup> Other two S protein substitutions, E484K in some strains of B.1.1.7 and L452R in B.1.617.2, may lead to much poorer effectivity of specific monoclonal antibody treatment in the corresponding infected cases (<https://www.cdc.gov>).

## PATHOGENESIS OF SARS-COV-2

### Entry and replication of SARS-CoV-2

Cell entry of SARS-CoV-2 depends on two determinants: (1) the viral S protein recognizes ACE2 receptor and (2) TMPRSS2 primes S protein.<sup>60,61</sup> The S1 subunit of the envelope-embedded S glycoprotein attaches to the cellular ACE2 receptor via the polar contacts of hydrophilic residues<sup>60,62</sup> TMPRSS2 may trigger a proteolytic cleavage at the S1/S2 multibasic cleavage site.<sup>63,64</sup> Other than ACE2, some studies reported that the human dipeptidyl peptidase 4 (DPP4), the cell-surface glucose-regulated protein-78, and the receptor tyrosine kinase anaxelekto may also be conducive to viral entry and infection.<sup>65–68</sup> Cathepsin B/L and furin, paired basic amino acid cleaving enzyme (FURIN) may also catalyze S protein proteolytic cleavage.<sup>64,69</sup> As the S protein is cleaved to the S1 and the S2 domain, the cell surface receptor neuropilin-1 (NRP1) may bind to the C-terminal functional





**Fig. 2** Overview of clinical characteristics of COVID-19. CRP C-reactive protein, AST aspartate aminotransferase, PCT procalcitonin, TnI/TnT troponin I/troponin T, ALT alanine aminotransferase, LDH lactate dehydrogenase, CT computed tomography, GGO ground-glass opacity

furin-cleavage sequence of the S1 domain more strongly in some cases and help the S2 isolating from the S1 domain.<sup>70,71</sup> After S1 detachment, the S2 subunit undergoes a conformational change and mediates the fusion between the virus and the host membranes to mediate viral infection.<sup>71,72</sup> The virus enters the cytoplasm and starts the replication process to assemble new viral particles and amplify its viral load.<sup>69,73</sup> During the viral entry and replication process, cyclophilin A is essential for viral replication and its interaction with CD147 may mediate SARS-CoV-2 entering the host cells.<sup>65</sup>

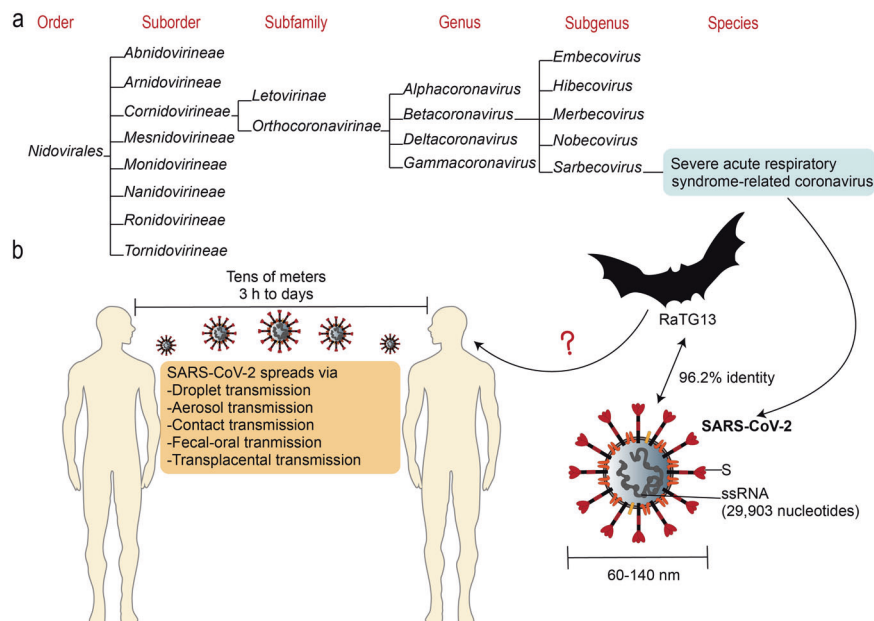
#### Renin-angiotensin-aldosterone system (RAAS)

RAAS is a complex system involved in multiple biological processes that are responsible for inducing a cascade of vasoactive peptides, which regulate vascular and renal functions.<sup>54,74,75</sup> In the blood, the precursor angiotensinogen (AGT) is hydrolyzed to angiotensin I (Ang I) by the active renin.<sup>76,77</sup> In the lungs, ACE removes the C-terminal dipeptide of Ang I to produce a potent vasoconstrictor, Ang II, which promotes detrimental effects by acting on Ang II type 1 receptor (AT1R).<sup>78-80</sup> ACE2 removes a single C-terminal amino acid from Ang I and Ang II to generate Ang-(1-9) and Ang-(1-7), of which the former is in turn converted to Ang-(1-7) by ACE.<sup>81</sup> Ang-(1-7) counters Ang II cellular and molecular effects by binding and activating the G-protein-coupled Mas receptor (MasR).<sup>79,82,83</sup> ACE2 catalytic efficiency with Ang II as a substrate is 400-fold higher than with Ang I.<sup>83,84</sup> Accordingly, ACE2, an ACE homolog, is a key negative regulator antagonizing the activation of the classical RAAS by counterbalancing ACE actions.<sup>82,85</sup> ACE and ACE2 maintain the homeostasis via the "adverse" ACE/Ang II/AT1R axis and the "protective" ACE2/Ang-(1-7)/MasR axis. In general, ACE/ACE2 imbalance contributes to RAAS overactivation and pulmonary shutdown, and a high ACE activity and a reduced ACE2 expression would increase the risk of pulmonary and cardiovascular diseases.<sup>86</sup> ACE2 mainly binds to

cell membranes and rarely exists in a circulating soluble form.<sup>72</sup> Continued viral infection and replication markedly downregulate ACE2 receptors, leading to the loss of the catalytic effect of membrane ACE2, and thus results in unopposed acute Ang II aggregation and local RAAS activation.<sup>74</sup> During ACE2 downregulation, an accentuating RAAS imbalance may further exacerbate pathophysiological alteration in COVID-19.<sup>85,87</sup>

#### Immunopathogenesis

SARS-CoV-2 activates innate and acquired immune response, and further impairs the immune system and causes cytokine storm, which is an uncontrolled inflammatory response with elevations of circulating cytokine levels.<sup>73,88,89</sup> The initial antiviral responses are promoted by pattern recognition receptors (PRRs) detecting pathogen-associated molecular patterns (PAMPs).<sup>90</sup> Retinoic acid-inducible gene I (RIG-I) is an interferon (IFN)-stimulated gene (ISG), and RIG-I-mediated signaling could promote induction of antiviral IFN responses.<sup>90,91</sup> Recognition of virus promotes downstream transduction in nuclear factor- $\kappa$ B, IFN regulatory factor-3, and Janus kinase-signal transducer and activator of transcription signaling pathways.<sup>92</sup> Innate immune cells, such as parenchymal cells, neutrophils, dendritic cells, and macrophages, are stimulated to secrete inflammatory mediators.<sup>93</sup> SARS-CoV-2 disturbs the immune system with its immune evasion strategies, in which viral PAMPs escape from the detection of cytosolic PRRs efficiently.<sup>94</sup> The virus weakens the antiviral effects of ISG products through dysregulating IFN signaling and IFN generation.<sup>94,95</sup> In acquired immunity, SARS-CoV-2 may target the CD147 spike protein of T lymphocytes.<sup>96</sup> Viral peptides are presented by major histocompatibility complex (MHC) Class-I molecules to CD8<sup>+</sup> T cells (cytotoxic T cells) to kill the virus directly, while by MHC Class-II molecules to CD4<sup>+</sup> T cells (helper T cells).<sup>96,97</sup> CD4<sup>+</sup> T cells generate proinflammatory cytokines and mediators to facilitate other immune cells.<sup>88</sup> B lymphocytes are directly stimulated by



**Fig. 3** The causative pathogen, SARS-CoV-2. **a** The taxonomy of SARS-CoV-2 is shown (from <https://talk.ictvonline.org/taxonomy/>). **b** SARS-CoV-2 is ssRNA virus with a diameter of ~60–140 nm and whole viral genome sequence of 29,903 nucleotides, and possesses distinctive S protein; 96.2% identity is shown between SARS-CoV-2 and a bat-borne coronavirus, RaTG13. SARS-CoV-2 spreads from human to human via droplet, aerosol, contact, fecal–oral, and transplacental transmissions. Droplets with SARS-CoV-2 can spread up to tens of meters and remain viable and infectious for 3 h to days. SARS-CoV-2 severe acute respiratory coronavirus 2, S spike protein, ssRNA single-strand RNA

SARS-CoV-2 and interact with CD4<sup>+</sup> T cells to produce substantial immunoglobulin G antibodies, which lead to the disruption of the virus and increased proinflammatory cytokines.<sup>14,96</sup> Complement activation through classical or alternative pathways generates a number of chemotactic/inflammatory mediators.<sup>91</sup> The release of proinflammatory cytokines also could be induced by increased A disintegrin and metalloproteinase 17 activity due to viral invasion.<sup>82</sup> The cytokine storm, which is characterized by a radical rise in the number of inflammatory cytokines/chemokines such as interleukin-2 (IL-2), IL-6, IL-7, IL-10, granulocyte-colony stimulating factor, IFN- $\gamma$ -inducible protein 10, tumor necrosis factor- $\alpha$ , macrophage inflammatory protein-1 $\alpha$ , monocyte chemoattractant protein-1, C-X-C motif ligand 9 (CXCL9), CXCL10, and CXCL11, triggers extensive tissue injury and body dysfunction and is considered as the primary contribution to mortality in COVID-19.<sup>32,33,97,98</sup> Among these inflammatory cytokines/chemokines, IL-6 was reported to play a key role in COVID-19 cytokine storm development.<sup>89</sup> However, a study found that only one cytokine, macrophage migration inhibitory factor, was significantly higher in COVID-19 patients than healthy controls. In addition, elevations of IL-6 were only found in some severe/critical patients and much less than patients with other cytokine storm syndrome-associated diseases.<sup>99</sup> These controversial results pointed that SARS-CoV-2 causes a chemokine storm, not a cytokine storm, providing an interesting insight into COVID-19 immunopathogenesis.<sup>99</sup> Genetically determined individual differences in immunity may relate to both variants in the immune-related genes and the inherent differences in the X- and Y-chromosome gene expressions.<sup>100,101</sup>

**AUTOSOMAL LOCI AND GENES ASSOCIATED WITH COVID-19**  
 2q24.2 and the interferon induced with helicase C domain 1 gene (*IFIH1*)

The *IFIH1* protein is a primary PRR that first senses the coronavirus RNA and then triggers innate immunity and activates mitochondrial antiviral signaling protein.<sup>102</sup> The variant rs1990760 (p.Ala946Thr) of the *IFIH1* gene has been reported to be positively related to increased expression of the viral

resistance gene *IFIH1* and IFN-induced gene.<sup>103</sup> This polymorphic variant in various ethnic populations is correlated with population migration and originated from the European and Asian populations. It is expected that rs1990760 T-allele confers carriers more resistance to COVID-19, e.g., Africans and African-Americans with low-frequency ranging from 0.06 to 0.35 have a more vulnerable risk for COVID-19 than Caucasians and Indians with an overall frequency of 0.56.<sup>102</sup>

**3p21.31**

A genome-wide association study (GWAS) conducted in Italy and Spain revealed that a 3p21.31 gene cluster comprised of the solute carrier protein family 6 member 20 gene (*SLC6A20*), the leucine zipper transcription factor like 1 gene (*LZTFL1*), the FYVE and coiled-coil domain autophagy adaptor 1 gene (*FYCO1*), the C-X-C motif chemokine receptor 6 gene (*CXCR6*), the X-C motif chemokine receptor 1 gene (*XCR1*), and the C-C motif chemokine receptor 9 gene (*CCR9*) is a genetically susceptible locus in severe COVID-19 patients with respiratory failure.<sup>104,105</sup> The *SLC6A20* gene encodes a transporter, signaling threshold regulating transmembrane adaptor, which functionally interacts with ACE2.<sup>105</sup> The *CXCR6* gene and the *CCR9* gene encoding chemokine receptors are implicated in T cell differentiation and recruitment.<sup>95</sup> Rs11385942 in the *LZTFL1* gene associated with increased *SLC6A20* expression and reduced *CXCR6* expression is a risk variant and is common in Europeans, Africans, and South Asians, but almost absent in the East Asians.<sup>106</sup>

Another study suggests that a 49.4 kb haplotype in high linkage disequilibrium (LD) on 3p21.31 is the most highly correlated to severe COVID-19, and this core haplotype is thought to have entered the human population from the Neanderthals, an extinct hominin ~40,000 to 60,000 years ago.<sup>107,108</sup> Neanderthal-derived core haplotype frequency varies significantly among populations that 63% of the Bengalese, ~30% of the South Asians, 8% of the Europeans, 4% of the Americans, a lower frequency of the East Asians, and almost none of the Africans carry this risk haplotype.<sup>109</sup> This could explain that Briton-originated studies showed higher mortality in COVID-19 patients of Bangladeshi

ethnicity (~2 times higher) and of South Asian descent.<sup>108,110</sup> Correspondingly, mortality rates reported on 14 July 2020 in South Africa, Japan, South Korea, and China were substantially lower than the Western countries (in North America and West Europe).<sup>111</sup> However, black individuals were at higher risk compared with white people in England and the United States.<sup>112–114</sup> This paradoxical fact may be explained by the impacts of other genetic and environmental factors.

#### 6p21.3 and HLA genotype

The HLA system containing nearly 27,000 alleles in classes I, II, and III is an exceedingly polymorphic region.<sup>115,116</sup> Genetic variations across the *HLA-A*, *HLA-B*, *HLA-C*, *HLA-DR*, *HLA-DP*, and *HLA-DQ* genes, which encode MHC molecules, might change the process of viral infection by differentially mediating antiviral immunity.<sup>109,117</sup> Several studies suggest that there may be specific risk and protective HLA alleles or haplotypes for COVID-19 incidence and mortality.<sup>9</sup>

HLA-A and HLA-C were reported to have the relatively greatest and least capacities for presenting SARS-CoV-2, respectively, and HLA-B preferentially involves susceptibility to COVID-19.<sup>16,118</sup> HLA-A\*25:01, HLA-A\*25:02, HLA-B\*46:01, HLA-C\*01:02, and HLA-B22 serotype, including HLA-B\*54:01, HLA-B\*55:01, HLA-B\*55:07, HLA-B\*55:12, and HLA-B\*56:01, are weak presenters, and thus individuals with these alleles may be COVID-19 susceptible.<sup>118–120</sup> HLA-A\*02 subtypes such as HLA-A\*02:02, HLA-A\*02:03, HLA-A\*02:05, HLA-A\*02:06, HLA-A\*02:09, HLA-A\*02:11, HLA-A\*02:12, HLA-A\*02:22, HLA-A\*02:24, HLA-A\*02:35, and HLA-A\*02:40, as well as HLA-A\*24:02, HLA-B\*15:03, HLA-B\*52:01, HLA-C\*12:02, and HLA-C\*12:03, are strong presenters for SARS-CoV-2 epitopes and predicted to be protective.<sup>119,121,122</sup> SARS-CoV-2 peptides presented by HLA-B\*15:03 are common among human coronaviruses and enable cross-protective T cell-based immunity.<sup>120</sup> HLA-A\*02:01 has varying capacities for presenting SARS-CoV-2 antigens in different studies.<sup>15,117</sup>

Several studies concluded that HLA-A\*25, HLA-B\*08, HLA-B\*15:01, HLA-B\*15:27, HLA-B\*27:07, HLA-B\*44, HLA-B\*51, HLA-C\*01, HLA-C\*03, HLA-C\*04:01, HLA-DRB1\*15:01, HLA-DQA1\_509, HLA-DQB1\*04, and HLA-DQB1\*06:02 were associated with higher occurrence and mortality, while HLA-B\*14, HLA-B\*18, and HLA-B\*49 showed an inverse log-linear relationship with COVID-19.<sup>123–127</sup> HLA-A\*11 was positively associated with COVID-19 mortality, but another analysis suggested that HLA-A\*11:01 could generate efficient antiviral responses.<sup>115,117</sup> HLA-DRB1\*01:01 (severe 2.2% vs mild 0.5%), HLA-DRB1\*14:04 (severe 2.0% vs mild 0.5%), and HLA-DQA1\*01:01 (severe 2.9% vs mild 0.9%) are risk alleles for severe COVID-19, while HLA-DRB1\*12:01 (severe 2.2% vs mild 3.7%) and HLA-DPB1\*03:01 (severe 0.7% vs mild 4.5%) were protective.<sup>106</sup> HLA-C\*05 is significantly correlated to increased COVID-19 death risk and each increase of 1% in HLA-C\*05 frequency is followed by an increase of 44 deaths/million. Its receptor KIR2DS4fl is located on natural killer (NK) cells and recognizes viral peptides bound to HLA-C\*05 to generate a potent activation signal, leading to NK cell-induced hyperactive antiviral immunity jointly with HLA-C\*05.<sup>9</sup> Several South-East Asian and Oceania regions seem to correspond to higher predicted protective allele frequencies than other global regions based on data from the Allele Frequency Net Database (<http://www.allelefrequencies.net/hla.asp>; Supplementary Fig. S1, 2). HLA-A\*24:02 was found to bind the peptide VYIGDPAQL, which is a virus helicase fragment shared between SARS-CoV-2 and two common cold coronaviruses, human coronavirus OC43 and HKU1. Thus, it was assumed that the anti-VYIGDPAQL T cells primed by previous OC43 or HKU1 infections could be restimulated after SARS-CoV-2 infection.<sup>128</sup> HLA-A\*24:02 allele carried by 25.5–98.0% of Chinese may partly explain the better epidemic prevention effect in China.

HLA is codominant and expresses all the alleles in the high gene density, complex LD, and homology regions.<sup>109,116</sup> This suggests studying complete HLA genotypes for each individual rather than being limited to a few protective or harmful alleles as a wiser course.<sup>129</sup> Haplotype HLA-A\*11:01-B\*51:01-C\*14:02 was more common in severe COVID-19 patients than in mild ones.<sup>106</sup> An Italian study found that haplotype HLA-A\*01:01-B\*08:01-C\*07:01-DRB1\*03:01 contributed to COVID-19 higher occurrence and mortality in northern Italy, while haplotype HLA-A\*02:01-B\*18:01-C\*07:01-DRB1\*11:04 closely linked to lower occurrence and mortality in central-southern Italy.<sup>130</sup> A Sardinian study identified two haplotypes HLA-A\*02:05-B\*58:01-DRB1\*08:01 and HLA-A\*02:05-B\*58:01-C\*07:01 as being protective against severe COVID-19.<sup>123</sup>

Since COVID-19 vaccines may have variable binding affinities with different HLA genotypes in different populations, predicting good binders across certain HLA alleles may contribute to design an efficacious COVID-19 vaccine with corresponding epitope targets.<sup>131</sup>

#### 9q34.2 and the ABO, alpha 1–3-N-acetylgalactosaminyltransferase and alpha 1–3-galactosyltransferase gene (*ABO*)

A, B, and O blood groups possess A-antigen, B-antigen, and the biosynthetic precursor H-antigen, respectively.<sup>132,133</sup> The antigen-encoding gene comprises A, B, and O alleles and is expressed in four genetic phenotypes.<sup>132</sup> SARS-CoV-2 susceptibility and survival following infection may relate to ABO blood groups. Individuals carrying blood group A have a higher COVID-19 risk, while blood group O exerts a relatively protective effect.<sup>132,134</sup> In the blood group A, A-antigen causes more P-selectin and intercellular cell adhesion molecule 1 attached to endothelial cells to increase cardiovascular disease likelihood. Blood group O individuals with ~25% decreased levels of von Willebrand factor might have lower thrombotic disease risk.<sup>134–136</sup> In the blood group B, natural anti-A antibodies might exert a neutralizing activity blocking adhesion between S proteins and ACE2.<sup>133,134</sup>

The GATC haplotype rs8176746–rs8176740–rs495828–rs12683493, of which position is coincident with *ABO* locus, is common in people with non-O blood groups and positively correlated to ACE activity, while blood group O is characterized by intermediate ACE activity.<sup>135,137</sup> Variants account for 15% of ACE activity variance, of which rs8176746 and rs495828 may independently reckon 2.8% and 4.9%, respectively.<sup>137–139</sup> The *ABO* variant rs657152 was considered as a significant signal associating with severe COVID-19 in Italian and Spanish cohorts.<sup>104</sup>

#### 9q34.3 and the dipeptidyl peptidase 7 gene (*DPP7*)

A 1-base pair (bp) insertion in the *DPP7* gene destroying *DPP7* transcription may have a potential monogenic effect for asymptomatic COVID-19 in a Chinese family analysis.<sup>106</sup> *DPP7* known as a survival factor to maintain lymphocytes quiescently may potentially involve in COVID-19 immunopathogenesis.<sup>140</sup> The specific functional effects of the *DPP7* gene in COVID-19 still need further clarification.

#### 11p15.5 and the interferon induced transmembrane protein 3 gene (*IFITM3*)

The rs12252 C-allele homozygosity in the *IFITM3* gene relates to COVID-19 patient disease severity, and CC-homozygote patients have a 6.37 times higher risk of severity after a SARS-CoV-2 infection.<sup>141–143</sup> This association is not thought to stem directly from rs12252, but from a functional variant existing LD with rs12252 of *IFITM3* or a nearby gene.<sup>144</sup> Rs34481144 A-allele (38–56% in Europeans, 2–14% in Africans, and 1–2% in Chinese) might increase COVID-19 susceptibility by triggering methylation of the *IFITM3* promoter to decrease *IFITM3* mRNA expression in CD8<sup>+</sup> T cells and depressing surrounding gene transcription.<sup>145,146</sup>



**12q24.33 and the golgin A3 gene (*GOLGA3*)**

Pedigree analysis of Chinese suggested the splice acceptor variant rs143359233 in the *GOLGA3* gene potentially implicated in critically ill COVID-19 patients as a monogenic factor.<sup>106</sup> The *GOLGA3* gene encodes a Golgi complex-associated protein, which participates in protein transportation, cell apoptosis, Golgi positioning, and spermatogenesis.<sup>147</sup> Its defect was proved to lead to male infertility previously, but the reliable relationship between the *GOLGA3* gene and COVID-19 remains uncertain.<sup>147</sup> *GOLGA3* may implicate COVID-19 severity by influencing the interaction of SARS-CoV-2 to innate immune pathways.<sup>148</sup>

**13q12.3 and the high mobility group box 1 gene (*HMGB1*)**

The *HMGB1* gene encodes a DNA-binding protein, which is a critical damage-associated molecular pattern (DAMP) and probably regulates a proviral gene expression program. *HMGB1* may interact with Toll-like receptor 4 (TLR4) and the advanced glycosylation end-product specific receptor to induce cytokine storm in immune cells and *ACE2* expression in alveolar epithelial cells, further increasing COVID-19 susceptibility.<sup>149,150</sup>

**15q26.1 and the *FURIN* gene**

The *FURIN* gene encodes a ubiquitous membrane-bound pro-protein convertase that cleaves the SARS-CoV-2 S protein into the S1 and S2 subunits. Two highly frequent *FURIN* variants relating to upregulated *FURIN* in Africans, rs6226 (93%) and rs8039305 (81%), are associated with increased hypertension risk and SARS-CoV-2 infection.<sup>151</sup> A common variant, rs4702, may directly reduce SARS-CoV-2 infection. The variant rs769208985 (p.Arg298Gln), representing glutamine residue by replacing arginine in a highly conserved position (R298), might influence *FURIN* recognition of the SARS-CoV-2 S protein.<sup>20,70</sup>

**17q23.3 and the *ACE* gene**

The insertion of an Alu repeat element into *ACE* intron 16 may result in alternative splicing in which the *ACE* I-allele leads to protein shortening and the loss of a catalytically active protein domain, while the *ACE* D-allele still maintains two active protein domains catalyzing Ang I to Ang II.<sup>152,153</sup> Approximately 60% of *ACE* level variability in general populations is likely to be determined by the *ACE* I/D variant.<sup>154,155</sup> The I/D variant is associated with *ACE* circulating and tissue concentrations, which means that *ACE* activity levels in I/I carriers are about half of that of D/D carriers.<sup>87,156</sup> COVID-19 variable recovery and prevalence rates correlate to the ratio of the *ACE* I/D allele frequency and the geographical variations of the *ACE* I/D variant.<sup>157,158</sup>

The racial difference in the *ACE* gene polymorphism is well understood. According to the “thrifty genotype” hypothesis put forward by J.V. Neel, after modern human ancestors expanded out of Africa ~200,000 years ago, genetic variation of the D-allele occurred as the D-allele favoring the retention of salt and water became detrimental.<sup>159</sup> Middle Eastern populations, particularly those in Lebanon with a relatively low I-allele frequency, are believed to be the ancestor of the *ACE* variant.<sup>86</sup> I/I genotype increases westwards and eastwards from the Middle East. The distribution of D-allele is characterized by the highest frequency of D-allele in Africa and Arab regions, medium frequencies in Europe, Australia, and America, and the lowest frequency in East Asia.<sup>159</sup> Therefore, the higher recovery rate in East Asians and disproportionately higher fatality rate in African Americans are unsurprising.<sup>155,158</sup>

**19q13.32 and the apolipoprotein E gene (*APOE*)**

The *APOE* gene has three common alleles,  $\epsilon 2$ ,  $\epsilon 3$ , and  $\epsilon 4$ , which are haplotypes of rs429358 and rs7412.<sup>160,161</sup> Compared to the most common *APOE*  $\epsilon 3\epsilon 3$  genotype, individuals who are homozygous for *APOE*  $\epsilon 4$  have twice the risk of severe COVID-19, although mortalities between *APOE*  $\epsilon 3\epsilon 4$  and  $\epsilon 3\epsilon 3$  COVID-19-positive

subjects have no significant difference.<sup>162,163</sup> The *APOE*  $\epsilon 4\epsilon 4$  homozygous genotype might have a higher risk of severe COVID-19 due to regulating proinflammatory pathways and lipoprotein function being affected.<sup>161,163</sup> An African-American  $\epsilon 4$ -allele frequency of 29.5% compared to a Caucasian rate of 12.1% may explain the diverse mortalities.<sup>162</sup>

**21q22.3 and the *TMPRSS2* gene**

The *TMPRSS2* gene variants may play a significant role in the interindividual differences particularly in the gender-related bias of COVID-19 susceptibility and severity.<sup>164</sup> Rs61299115, rs4303794, and rs11088551 have relatively high frequencies in the general populations (25–36%), but much lower, 2%, in East Asian populations. They potentially enhance *TMPRSS2* transcription, and thus the rarity of these three single-nucleotide variants (SNVs) among the East Asians results in lower *TMPRSS2* expression levels.<sup>165</sup> Rs12329760 (p.Val197Met) located at the exonic splicing enhancer site might considerably increase the *TMPRSS2* faulty expression, weaken *TMPRSS2* stability, and inhibit S protein and *ACE2* interaction, which may contribute to asymptomatic and mild patients in Chinese with higher variant frequency.<sup>166</sup> However, in Italian populations, rs12329760, as well as a haplotype rs2070788–rs9974589–rs7364083–rs8134378, trigger increased *TMPRSS2* expression and may explain the higher mortality rate among the Italians with higher variant frequency.<sup>167</sup> Rs8134378 close to an androgen-responsive enhancer possibly increases the *TMPRSS2* gene expression in males in an androgen-specific manner and is co-regulated with a “European” haplotype rs463727–rs34624090–rs55964536–rs734056–rs4290734–rs34783969–rs11702475–rs35899679–rs35041537.<sup>168</sup> The variants rs2070788, rs464397, rs469390 (p.Val379Ile), and rs383510 could upregulate *TMPRSS2* expression in lung tissue and have lower frequencies in East Asians than Africans, Europeans, and Americans, which might explain the different COVID-19 susceptibilities in different populations.<sup>169,170</sup> Conversely, p.Asp435Tyr only presenting at a low frequency in Europeans leads to the lack of a key residue catalyzing substrate binding.<sup>171</sup>

**X- OR Y-LINKED LOCI AND GENES ASSOCIATED WITH COVID-19**

Consistent with Lyon’s theory, X-chromosome inactivation (XCI), which occurs in females in the late blastocyst stage, is a fundamental event in the epigenetic gene regulation that one of the X chromosomes is stochastically inactivated to equal X-linked gene dosage between genders.<sup>167,172–174</sup> Two noncoding RNAs control this complex inactivation process, which condenses one X chromosome into a compact structure, Barr body, and maintains an active X chromosome simultaneously. Approximately 15–30% of X-linked genes, most are on the short arm (p), can escape from the XCI.<sup>172,175</sup> Interestingly, XCI is cell-specific such that some cells express the maternal copy, while others express the paternal copy, and escape from XCI can be variable between individuals, among cells in a tissue, and during growth and aging.<sup>176,177</sup> The skewed XCI may bypass the deleterious X-linked variants in females, while any abnormal gene variants on the X chromosome of males are more likely to express phenotypically and to cause more pronounced consequences due to hemizygosity.<sup>160,175,178,179</sup> This appears to explain SARS-CoV-2 infection rate gender bias.

**Xp22.2 and the *TLR7* gene**

The *TLR7* gene encodes a Toll-like receptor that could recognize SARS-CoV-2 RNA and trigger the antiviral response.<sup>180</sup> An analysis performed on two young brother pairs with severe COVID-19 identified a maternally inherited variant rs2042915990 (p.Gln710Argfs\*18) and a missense variant rs200553089 (p.Val795Phe) as rare loss-of-function (LOF) variants in the *TLR7* gene, which result in immunodeficiencies in type I and II



interferon responses.<sup>181</sup> Further, a nested case-control study identified the *TLR7* gene variants p.Ser301Pro, rs189681811 (p.Arg920Lys), and rs147244662 (p.Ala1032Thr) as LOF variants, which in young, male, severe COVID-19 patients were considered to account for COVID-19 susceptibility in up to 2% cases.<sup>182</sup>

#### Xp22.22 and the *ACE2* gene

The *ACE2* gene encoding a dipeptidyl carboxypeptidase with 805 amino acids is a putative risk factor for SARS-CoV-2 infection.<sup>72,183</sup> *ACE2* contains a potential N-terminal signal peptide, a peptidase domain, and a C-terminal collectrin-like domain, which ends with the single transmembrane helix. A ferredoxin-like fold “Neck” domain is between the peptidase domain and transmembrane helix. The crucial roles of the peptidase and neck domains (residues 19–726) in the *ACE2* homodimerization allow for positing that variants affecting these amino acid residues may influence viral infection.<sup>184–187</sup>

Up to the date of this writing, no genetically monogenic, naturally resistant *ACE2* mutations which counter S protein binding have been reported.<sup>188</sup> However, a number of *ACE2* variants may influence COVID-19 susceptibility and outcomes via three primary routines: (1) alterations of *ACE2*-binding properties to sirtuin 1, which regulates transcriptional and post-translational modifications of the *ACE2* gene, (2) alteration of the soluble *ACE2* levels in circulation and the affinity and density of *ACE2* for the S protein, and (3) alteration of circulating Ang-(1–7), which causes a greater marked RAAS imbalance and greater disease severity.<sup>80,87,185,189</sup>

Among the most significant variants for *ACE2* activity and levels, the most frequent is the transition rs2285666 (G8790A).<sup>152,190</sup> The A-allele carriers may have higher serum *ACE2* levels than the G-allele carriers, of which A/A genotype had almost 50% higher *ACE2* expression levels than the G/G genotype.<sup>191,192</sup> Rs2285666 located in the intronic-consensus splice site region might theoretically affect the processing of total RNA to mRNA with alternative splicing mechanisms and further the amount of protein.<sup>193</sup> The transition G8790A was predicted to lead to ~9.2% increased strength of the splice site and further elevated *ACE2* serum levels.<sup>192</sup> Accordingly, rs2285666 is suggested to be a protective variant to COVID-19, and lower morbidity and mortality in Indians could be explained by the A-allele of rs2285666.<sup>193</sup> The variant rs2106809 reported in Indians and Saudi Arabians may primarily influence serum *ACE2* levels, and that the C/C or C/T genotype has comparatively higher levels than the T/T genotype.<sup>166</sup>

Rs4646114 and rs4646115, which are more prevalent in African descent populations with frequencies of 5.0–7.2% and 1.4–1.8%, could accelerate viral infection and spread, and thus may associate with higher COVID-19 susceptibility.<sup>165</sup> Rs4646116 (p.Lys26Arg), which is quite frequent in Caucasians, but has not yet been detected in the East-Asian populations, activates *ACE2* and boosts binding to S protein, while rs191860450 (p.Ile468Val), which is more prevalent in Asians, may alter the *ACE2*-S protein interaction characteristics, but the significance of this is unclear.<sup>183,194</sup> The variant p.Arg514Gly, located in the AGT-*ACE2* interaction surface, was predicted to increase COVID-19 risk by altering RAAS function.<sup>171</sup> The higher COVID-19 mortality in Italy may be partly explained by the role of rs41303171 (p.Asn720Asp), which is more prevalently carried by Italians, in promoting *TMPRSS2* cleaving and viral intake.<sup>195</sup>

The variants rs73635825 (p.Ser19Pro) and rs766996587 (p.Met82Ile) exclusively presented in Africans may reduce encoded protein stability and binding affinity to S protein binding sites.<sup>4,8,196</sup> The European-specific variant rs1448326240 (p.Glu239-His) is thought to be an interaction-inhibiting variant and lead to a lower SARS-CoV-2 susceptibility.<sup>184</sup> Rs143936283 (p.Glu329Gly) has a lower binding affinity for the S protein, implying that rs143936283 may confer a lower probability of viral attachment

and some level of resistance against infection.<sup>195</sup> The variants p.Leu351Val and rs762890235 (p.Pro389His), which occur in the *ACE2*-S protein interaction region, are predicted to interfere with the internalization process.<sup>197</sup> Rs961360700 (p.Asp355Asn) and rs1396769231 (p.Met383Thr) were also predicted to adversely affect *ACE2* stability.<sup>196</sup> The four variants located in the *ACE2* dimeric interface, p.Arg708Trp, p.Arg710Cys, p.Arg710His, and p.Arg716Cys, could affect *ACE2* cleavage by *TMPRSS2* and change the dimer formation, which may be responsible for the milder COVID-19 symptoms in Europeans having these four variants.<sup>171</sup>

Heterogeneous *ACE2* expression in different ethnic groups might be a measure of differential population reactions to COVID-19. For example, Asians have a higher *ACE2* expression than African Americans and Caucasians. The expression quantitative loci for upregulating *ACE2* can be up to almost 100% in East Asians, which are over 30% higher than other racial groups.<sup>185,198</sup> The prevalence of *ACE2*-downregulating variants is 54% in non-Finnish Europeans, 39% in Africans/African Americans, and 2–10% in Latinos/admixed Americans, East Asians, Finns, and South Asians, while Amish and Ashkenazi Jewish populations seem to carry none of such variants.<sup>171</sup> Highly penetrant dominant trait presenting in the *ACE2* gene probably affects familial clusters.<sup>10</sup> Approximately 320–365 out of every 100,000 humans possess SNVs decreasing spike binding, while 4–12 of every 100,000 humans possess SNVs increasing spike binding. Specific SNVs affecting S protein binding are more abundant in individuals of a certain ancestry, and this frequency may vary six-fold between different ancestries.<sup>199</sup>

Furthermore, the *ACE2* gene, mapped at the pseudoautosomal regions of X-chromosome, could escape XCI more probably, which likely confers females a double *ACE2* dosage to compensate for the loss of membrane *ACE2* due to SARS-CoV-2.<sup>175,178</sup> One study showed that, in the hemizygous state, >50% of the variants probably influence the binding of the human *ACE2* and the viral S1 protein.<sup>184</sup> *ACE2* interaction-booster and interaction-inhibitor variants can be more significant in males and the former may result in a higher mortality rate in males than females.<sup>100</sup> The fact that the *ACE2* gene expression could be elevated in females due to a skewed XCI, providing a larger *ACE2* pool to maintain the fundamental balance of RAAS-regulatory axis in multiple organs after viral infection, could partly explain the lower frequency of severe COVID-19 in females than males.<sup>17,172</sup>

#### Xq12 and the androgen receptor gene (*AR*)

The 15-bp *AR* binding element is the critical part of the *TMPRSS2* promoter for androgen's binding and its transcription regulation.<sup>200,201</sup> This *AR* element is a polymorphic unit with CAG trinucleotide repeat length variations in the *AR* gene's first exon.<sup>202</sup> Length variation can determine both transcriptional activity and androgen resistance strength that shorter CAG repeat polymorphism may lead to *TMPRSS2* overexpression and link to androgen sensitivity, as well as more severe COVID-19.<sup>176</sup> The variation of greater COVID-19 mortality in males than females or infants of either gender may be explained by androgen-mediated *ACE2* and *TMPRSS2* expressions.<sup>101,200,203</sup> Ethnicity-based vulnerability may be explained by the *AR* gene polymorphisms that African-American males with shorter CAG repeat lengths have a disproportionate mortality rate than non-Hispanic white males.<sup>201,202</sup>

## CONCLUSIONS AND PERSPECTIVES

The world is still suffering from the COVID-19 outbreak and the ultimate outcomes are, so far, unmeasurable, but the global economic, social, and political disruptions caused by this pandemic are poised to worsen in the foreseeable future.<sup>11,204</sup> Therefore, understanding the causal relationships between host genetic basis and COVID-19 is urgently needed to identify

biomarkers for individuals at high risk, which might also provide potential targets for therapy.<sup>6,98</sup> Focusing on these variants, which relate to disease susceptibility and severity through viral trafficking pathways or drug curative effect, could better identify risk subjects and effectively control the disease. Large data consortiums are organizing to produce, share, and analyze data, such as the COVID-19 Host Genetics Initiative and the COVID Human Genetic Effort.<sup>12,205</sup> As more host genetic factors associated with COVID-19 are identified, it should become possible to create tests that would predict the susceptible populations and allow for classifying and safeguarding them.<sup>190</sup> In summary, we extract and review positive results from vast reported papers. However, for a certain variant, large-scale meta-analyses combining data from multiple consistent studies with reliable statistical significance would be helpful to find therapy targets.

Ongoing investigations into COVID-19 and individual genetic makeup are fueling global research to develop vaccines, prioritize individuals for treatment, and discover potential drug target candidates.<sup>95</sup> The antiviral drug Veklury (remdesivir) is the first and the only treatment for COVID-19 approved by the US Food and Drug Administration (<https://www.fda.gov>). Given that remdesivir is a broad-spectrum antiviral drug, better direct-target-based antiviral therapies that intervene in SARS-CoV-2-infected pathways are anticipated.<sup>206</sup> Human genetic basis of COVID-19, which is well known to impact disease susceptibility and severity, may offer novel insights into COVID-19 therapies and controls through identifying particular genes and pathways. Large-scale screening of potential targeted drugs and experimental therapeutic studies would be helpful to develop new drugs or discover repurposing opportunities for existing drugs.<sup>148</sup> In order to end this century nightmare early, when it comes to ethical considerations and/or societal questions, further investigations should pay attention to (1) ensuring the validity and usefulness of the reported studies, (2) not undermining the necessity of solidarity in the public health action, (3) not affecting individuals' action ability or making them be discrimination targets, and (4) perfecting genetic information-related legislation.<sup>207</sup>

## ACKNOWLEDGEMENTS

We acknowledge all the participants of our study. This work was supported by the Special Emergency Project for Novel Coronavirus Pneumonia from the Science and Technology Department of Hunan Province (2020SK3032).

## AUTHOR CONTRIBUTIONS

H.D., X.Y., and L.Y. contributed equally to researching data for the article, discussions of the content, and writing the manuscript. H.D. and L.Y. contributed to reviewing and editing the manuscript before submission. All authors read and approved the final manuscript.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41392-021-00736-8>.

**Competing interests:** The authors declare no competing interests.

## REFERENCES

1. Yoo, J. H. The fight against the 2019-nCoV outbreak: an arduous march has just begun. *J. Korean Med. Sci.* **35**, e56 (2020).
2. Lai, C. C., Shih, T. P., Ko, W. C., Tang, H. J. & Hsueh, P. R. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges. *Int. J. Antimicrob. Agents* **55**, 105924 (2020).
3. Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).

4. Ravaoli, S. et al. ACE2 and TMPRSS2 potential involvement in genetic susceptibility to SARS-CoV-2 in cancer patients. *Cell Transplant.* **29**, 963689720968749 (2020).
5. Gorbalenya, A. E. et al. The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **5**, 536–544 (2020).
6. Hu, J., Li, C., Wang, S., Li, T. & Zhang, H. Genetic variants are identified to increase risk of COVID-19 related mortality from UK biobank data. *Hum. Genomics* **15**, 10 (2021).
7. Zbinden-Foncea, H., Francaux, M., Deldicque, L. & Hawley, J. A. Does high cardiorespiratory fitness confer some protection against proinflammatory responses after infection by SARS-CoV-2? *Obesity* **28**, 1378–1381 (2020).
8. Calcagnile, M. et al. Molecular docking simulation reveals ACE2 polymorphisms that may increase the affinity of ACE2 with the SARS-CoV-2 spike protein. *Biochimie* **180**, 143–148 (2021).
9. Sakuraba, A., Haider, H. & Sato, T. Population difference in allele frequency of HLA-C\*05 and its correlation with COVID-19 mortality. *Viruses* **12**, 1333 (2020).
10. Ikitimur, H. et al. Determining host factors contributing to disease severity in a family cluster of 29 hospitalized SARS-CoV-2 patients: could genetic factors be relevant in the clinical course of COVID-19? *J. Med. Virol.* **93**, 357–365 (2021).
11. Rajarshi, K. et al. Essential functional molecules associated with SARS-CoV-2 infection: potential therapeutic targets for COVID-19. *Gene* **768**, 145313 (2021).
12. Ovsyannikova, I. G., Haralambieva, I. H., Croke, S. N., Poland, G. A. & Kennedy, R. B. The role of host genetics in the immune response to SARS-CoV-2 and COVID-19 susceptibility and severity. *Immunol. Rev.* **296**, 205–219 (2020).
13. Williams, F. M. K. et al. Self-reported symptoms of COVID-19, including symptoms most predictive of SARS-CoV-2 infection, are heritable. *Twin Res. Hum. Genet.* **23**, 316–321 (2020).
14. Agrawal, H. et al. An assessment on impact of COVID-19 infection in a gender specific manner. *Stem Cell Rev. Rep.* **17**, 94–112 (2021).
15. Zhao, J. et al. Relationship between the ABO blood group and the coronavirus disease 2019 (COVID-19) susceptibility. *Clin. Infect. Dis.* **73**, 328–331 (2021).
16. Nguyen, A. et al. Human leukocyte antigen susceptibility map for severe acute respiratory syndrome coronavirus 2. *J. Virol.* **94**, e00510–20 (2020).
17. Gómez, J. et al. Angiotensin-converting enzymes (ACE, ACE2) gene variants and COVID-19 outcome. *Gene* **762**, 145102 (2020).
18. Rahimi, A., Mirzazadeh, A. & Tavakolpour, S. Genetics and genomics of SARS-CoV-2: a review of the literature with the special focus on genetic diversity and SARS-CoV-2 genome detection. *Genomics* **113**, 1221–1232 (2021).
19. Nascimento, V. A. D. et al. Genomic and phylogenetic characterisation of an imported case of SARS-CoV-2 in Amazonas State, Brazil. *Mem. Inst. Oswaldo Cruz* **115**, e200310 (2020).
20. Dobrindt, K. et al. Common genetic variation in humans impacts in vitro susceptibility to SARS-CoV-2 infection. *Stem Cell Rep.* **16**, 505–518 (2021).
21. Jutzeler, C. R. et al. Comorbidities, clinical signs and symptoms, laboratory findings, imaging features, treatment strategies, and outcomes in adult and pediatric patients with COVID-19: a systematic review and meta-analysis. *Travel Med. Infect. Dis.* **37**, 101825 (2020).
22. Pormohammad, A. et al. Clinical characteristics, laboratory findings, radiographic signs and outcomes of 61,742 patients with confirmed COVID-19 infection: a systematic review and meta-analysis. *Microb. Pathog.* **147**, 104390 (2020).
23. Wang, C., Horby, P. W., Hayden, F. G. & Gao, G. F. A novel coronavirus outbreak of global health concern. *Lancet* **395**, 470–473 (2020).
24. Shu, T. et al. Plasma proteomics identify biomarkers and pathogenesis of COVID-19. *Immunity* **53**, 1108–1122.e5 (2020).
25. Fu, L. et al. Clinical characteristics of coronavirus disease 2019 (COVID-19) in China: a systematic review and meta-analysis. *J. Infect.* **80**, 656–665 (2020).
26. Wong, C. K. H., Wong, J. Y. H., Tang, E. H. M., Au, C. H. & Wai, A. K. C. Clinical presentations, laboratory and radiological findings, and treatments for 11,028 COVID-19 patients: a systematic review and meta-analysis. *Sci. Rep.* **10**, 19765 (2020).
27. Yang, J. et al. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *Int. J. Infect. Dis.* **94**, 91–95 (2020).
28. Li, L. Q. et al. COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis. *J. Med. Virol.* **92**, 577–583 (2020).
29. Ghayda, R. A. et al. Correlations of clinical and laboratory characteristics of COVID-19: a systematic review and meta-analysis. *Int. J. Environ. Res. Public Health* **17**, 5026 (2020).
30. Chua, T. H., Xu, Z. & King, N. K. K. Neurological manifestations in COVID-19: a systematic review and meta-analysis. *Brain Inj.* **34**, 1549–1568 (2020).
31. Makhoul, K. & Jankovic, J. Parkinson's disease after COVID-19. *J. Neurol. Sci.* **422**, 117331 (2021).

32. Singhal, T. A review of coronavirus disease-2019 (COVID-19). *Indian J. Pediatr.* **87**, 281–286 (2020).
33. Harrison, A. G., Lin, T. & Wang, P. Mechanisms of SARS-CoV-2 transmission and pathogenesis. *Trends Immunol.* **41**, 1100–1115 (2020).
34. Hossain, M. F. et al. COVID-19 outbreak: pathogenesis, current therapies, and potentials for future management. *Front. Pharm.* **11**, 563478 (2020).
35. Simoneau, C. R. & Ott, M. Modeling multi-organ infection by SARS-CoV-2 using stem cell technology. *Cell Stem Cell* **27**, 859–868 (2020).
36. Hu, J. & Wang, Y. The clinical characteristics and risk factors of severe COVID-19. *Gerontology* **6**, 1–12 (2021).
37. Kronbichler, A. et al. Asymptomatic patients as a source of COVID-19 infections: a systematic review and meta-analysis. *Int. J. Infect. Dis.* **98**, 180–186 (2020).
38. Wan, S. et al. CT manifestations and clinical characteristics of 1115 patients with coronavirus disease 2019 (COVID-19): a systematic review and meta-analysis. *Acad. Radiol.* **27**, 910–921 (2020).
39. Yang, H., Lan, Y., Yao, X., Lin, S. & Xie, B. The chest CT features of coronavirus disease 2019 (COVID-19) in China: a meta-analysis of 19 retrospective studies. *Virology* **17**, 159 (2020).
40. Muhammad, S. Z. et al. Chest computed tomography findings in hospitalized COVID-19 patients: a systematic review and meta-analysis. *Infez. Med.* **28**, 295–301 (2020).
41. Wu, F. et al. A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020).
42. Huang, C. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506 (2020).
43. Khomari, F., Nabi-Afjadi, M., Yarahmadi, S., Eskandari, H. & Bahreini, E. Effects of cell proteostasis network on the survival of SARS-CoV-2. *Biol. Proced. Online* **23**, 8 (2021).
44. V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.* **19**, 155–170 (2021).
45. Bzówka, M. et al. Structural and evolutionary analysis indicate that the SARS-CoV-2 Mpro is a challenging target for small-molecule inhibitor design. *Int. J. Mol. Sci.* **21**, 3099 (2020).
46. Adedokun, K. A., Olarinmoye, A. O., Mustapha, J. O. & Kamorudeen, R. T. A close look at the biology of SARS-CoV-2, and the potential influence of weather conditions and seasons on COVID-19 case spread. *Infect. Dis. Poverty* **9**, 77 (2020).
47. Chen, L. et al. RNA based mNGS approach identifies a novel human coronavirus from two individual pneumonia cases in 2019 Wuhan outbreak. *Emerg. Microbes Infect.* **9**, 313–319 (2020).
48. Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
49. Shereen, M. A., Khan, S., Kazmi, A., Bashir, N. & Siddique, R. COVID-19 infection: origin, transmission, and characteristics of human coronaviruses. *J. Adv. Res.* **24**, 91–98 (2020).
50. Morawska, L. & Milton, D. K. It is time to address airborne transmission of coronavirus disease 2019 (COVID-19). *Clin. Infect. Dis.* **71**, 2311–2313 (2020).
51. Jin, Y. H. et al. A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). *Mil. Med. Res.* **7**, 4 (2020).
52. Wang, Z. et al. Transmission and prevention of SARS-CoV-2. *Biochem. Soc. Trans.* **48**, 2307–2316 (2020).
53. Vivanti, A. J. et al. Transplacental transmission of SARS-CoV-2 infection. *Nat. Commun.* **11**, 3572 (2020).
54. Khalaf, K. et al. SARS-CoV-2: pathogenesis, and advancements in diagnostics and treatment. *Front. Immunol.* **11**, 570927 (2020).
55. Morawska, L. & Cao, J. Airborne transmission of SARS-CoV-2: the world should face the reality. *Environ. Int.* **139**, 105730 (2020).
56. Bai, Y. et al. Presumed asymptomatic carrier transmission of COVID-19. *JAMA* **323**, 1406–1407 (2020).
57. Frampton, D. et al. Genomic characteristics and clinical effect of the emergent SARS-CoV-2 B.1.1.7 lineage in London, UK: a whole-genome sequencing and hospital-based cohort study. *Lancet Infect. Dis.* **21**, 1246–1256 (2021).
58. Zhao, S. et al. Inferring the association between the risk of COVID-19 case fatality and N501Y substitution in SARS-CoV-2. *Viruses* **13**, 638 (2021).
59. Galloway, S. E. et al. Emergence of SARS-CoV-2 B.1.1.7 lineage—United States, December 29, 2020–January 12, 2021. *MMWR Morb. Mortal. Wkly. Rep.* **70**, 95–99 (2021).
60. Wang, Q. et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell* **181**, 894–904.e9 (2020).
61. Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e8 (2020).
62. Li, F., Li, W., Farzan, M. & Harrison, S. C. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **309**, 1864–1868 (2005).
63. Zipeto, D., Palmeira, J. D. F., Argañaraz, G. A. & Argañaraz, E. R. ACE2/ADAM17/TMPRSS2 interplay may be the main risk factor for COVID-19. *Front. Immunol.* **11**, 576745 (2020).
64. Smatti, M. K., Al-Sarraj, Y. A., Albagha, O. & Yassine, H. M. Host genetic variants potentially associated with SARS-CoV-2: a multi-population analysis. *Front. Genet.* **11**, 578523 (2020).
65. Liu, C., von Brunn, A. & Zhu, D. Cyclophilin A and CD147: novel therapeutic targets for the treatment of COVID-19. *Med. Drug Discov.* **7**, 100056 (2020).
66. Wang, S. et al. AXL is a candidate receptor for SARS-CoV-2 that promotes infection of pulmonary and bronchial epithelial cells. *Cell Res.* **31**, 126–140 (2021).
67. Ibrahim, I. M., Abdelmalek, D. H., Elshahat, M. E. & Elfiky, A. A. COVID-19 spike-host cell receptor GRP78 binding site prediction. *J. Infect.* **80**, 554–562 (2020).
68. Vankadari, N. & Wilce, J. A. Emerging COVID-19 coronavirus: glycan shield and structure prediction of spike glycoprotein and its interaction with human CD26. *Emerg. Microbes Infect.* **9**, 601–604 (2020).
69. Daniloski, Z. et al. Identification of required host factors for SARS-CoV-2 infection in human cells. *Cell* **184**, 92–105.e16 (2021).
70. Latini, A. et al. COVID-19 and genetic variants of protein involved in the SARS-CoV-2 entry into the host cells. *Genes* **11**, 1010 (2020).
71. Li, Z. L. & Buck, M. Neuropilin-1 assists SARS-CoV-2 infection by stimulating the separation of spike protein domains S1 and S2. *Biophys. J.* **120**, 2828–2837 (2021).
72. Verdecchia, P., Cavallini, C., Spanevello, A. & Angeli, F. The pivotal link between ACE2 deficiency and SARS-CoV-2 infection. *Eur. J. Intern. Med.* **76**, 14–20 (2020).
73. Rodríguez, C. et al. Pulmonary endothelial dysfunction and thrombotic complications in COVID-19 patients. *Am. J. Respir. Cell Mol. Biol.* **64**, 407–415 (2020).
74. Vaduganathan, M. et al. Renin-angiotensin-aldosterone system inhibitors in patients with COVID-19. *N. Engl. J. Med.* **382**, 1653–1659 (2020).
75. Medina-Enríquez, M. M. et al. ACE2: the molecular doorway to SARS-CoV-2. *Cell Biosci.* **10**, 148 (2020).
76. Warner, F. J., Smith, A. I., Hooper, N. M. & Turner, A. J. Angiotensin-converting enzyme-2: a molecular and cellular perspective. *Cell Mol. Life Sci.* **61**, 2704–2713 (2004).
77. Bader, M. & Ganten, D. Update on tissue renin-angiotensin systems. *J. Mol. Med.* **86**, 615–621 (2008).
78. Donoghue, M. et al. A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. *Circ. Res.* **87**, e1–e9 (2000).
79. Patel, V. B., Zhong, J. C., Grant, M. B. & Oudit, G. Y. Role of the ACE2/angiotensin 1-7 axis of the renin-angiotensin system in heart failure. *Circ. Res.* **118**, 1313–1326 (2016).
80. Sierko, J. et al. COVID-19: the influence of ACE genotype and ACE-I and ARBs on the course of SARS-CoV-2 infection in elderly patients. *Clin. Inter. Aging* **15**, 1231–1240 (2020).
81. Hamming, I. et al. The emerging role of ACE2 in physiology and disease. *J. Pathol.* **212**, 1–11 (2007).
82. Haitao, T. et al. COVID-19 and sex differences: mechanisms and biomarkers. *Mayo Clin. Proc.* **95**, 2189–2203 (2020).
83. Burrell, L. M., Johnston, C. I., Tikellis, C. & Cooper, M. E. ACE2, a new regulator of the renin-angiotensin system. *Trends Endocrinol. Metab.* **15**, 166–169 (2004).
84. Vickers, C. et al. Hydrolysis of biological peptides by human angiotensin-converting enzyme-related carboxypeptidase. *J. Biol. Chem.* **277**, 14838–14843 (2002).
85. Cheng, H., Wang, Y. & Wang, G. Q. Organ-protective effect of angiotensin-converting enzyme 2 and its effect on the prognosis of COVID-19. *J. Med. Virol.* **92**, 726–730 (2020).
86. Yamamoto, N. et al. SARS-CoV-2 infections and COVID-19 mortalities strongly correlate with ACE1 I/D genotype. *Gene* **758**, 144944 (2020).
87. Arnold, R. H. COVID-19—does this disease kill due to imbalance of the renin angiotensin system (RAS) caused by genetic and gender differences in the response to viral ACE2 attack? *Heart Lung Circ.* **29**, 964–972 (2020).
88. Mangalmurti, N. & Hunter, C. A. Cytokine storms: understanding COVID-19. *Immunity* **53**, 19–25 (2020).
89. Yang, L. et al. COVID-19: immunopathogenesis and immunotherapeutics. *Signal Transduct. Target. Ther.* **5**, 128 (2020).
90. Molaie, S., Dadkhah, M., Asgharizadeh, V., Karami, C. & Safarzadeh, E. The immune response and immune evasion characteristics in SARS-CoV, MERS-CoV, and SARS-CoV-2: vaccine design strategies. *Int. Immunopharmacol.* **92**, 107051 (2021).
91. Mohamed Khosroshahi, L., Rokni, M., Mokhtari, T. & Noorbakhsh, F. Immunology, immunopathogenesis and immunotherapeutics of COVID-19; an overview. *Int. Immunopharmacol.* **93**, 107364 (2021).



92. Catanzaro, M. et al. Immune response in COVID-19: addressing a pharmacological challenge by targeting pathways triggered by SARS-CoV-2. *Signal Transduct. Target. Ther.* **5**, 84 (2020).
93. Ebrahimi, N. et al. Recent findings on the coronavirus disease 2019 (COVID-19); immunopathogenesis and immunotherapeutics. *Int. Immunopharmacol.* **89**, 107082 (2020).
94. Kumar, S., Nyodu, R., Maurya, V. K. & Saxena, S. K. in *Coronavirus Disease 2019 (COVID-19): Epidemiology, Pathogenesis, Diagnosis, and Therapeutics* (ed. Saxena, S. K.) Ch. 5 (Springer Singapore, 2020).
95. Schurr, T. G. Host genetic factors and susceptibility to SARS-CoV-2 infection. *Am. J. Hum. Biol.* **32**, e23497 (2020).
96. Azkur, A. K. et al. Immune response to SARS-CoV-2 and mechanisms of immunopathological changes in COVID-19. *Allergy* **75**, 1564–1581 (2020).
97. Dhama, K. et al. An update on SARS-CoV-2/COVID-19 with particular reference to its clinical pathology, pathogenesis, immunopathology and mitigation strategies. *Travel Med. Infect. Dis.* **37**, 101755 (2020).
98. Ramos-Lopez, O. et al. Exploring host genetic polymorphisms involved in SARS-CoV infection outcomes: implications for personalized medicine in COVID-19. *Int. J. Genomics* **2020**, 6901217 (2020).
99. Syed, F. et al. Excessive matrix metalloproteinase-1 and hyperactivation of endothelial cells occurred in COVID-19 patients and were associated with the severity of COVID-19. *J. Infect. Dis.* **224**, 60–69 (2021).
100. von der Thüsen, J. & van der Eerden, M. Histopathology and genetic susceptibility in COVID-19 pneumonia. *Eur. J. Clin. Invest.* **50**, e13259 (2020).
101. Gebhard, C., Regitz-Zagrosek, V., Neuhauser, H. K., Morgan, R. & Klein, S. L. Impact of sex and gender on COVID-19 outcomes in Europe. *Biol. Sex Differ.* **11**, 29 (2020).
102. Maiti, A. K. The African-American population with a low allele frequency of SNP rs1990760 (T allele) in IFIH1 predicts less IFN-beta expression and potential vulnerability to COVID-19 infection. *Immunogenetics* **72**, 387–391 (2020).
103. Molineros, J. E. et al. Admixture mapping in lupus identifies multiple functional variants within IFIH1 associated with apoptosis, inflammation, and autoantibody production. *PLoS Genet.* **9**, e1003222 (2013).
104. Ellinghaus, D. et al. Genomewide association study of severe COVID-19 with respiratory failure. *N. Engl. J. Med.* **383**, 1522–1534 (2020).
105. Ellinghaus, D. et al. The ABO blood group locus and a chromosome 3 gene cluster associate with SARS-CoV-2 respiratory failure in an Italian-Spanish genome-wide association analysis. Preprint at <https://www.medrxiv.org/content/early/2020/06/02/2020.05.31.20114991> (2020).
106. Wang, F. et al. Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility. *Cell Discov.* **6**, 83 (2020).
107. Zeberg, H. & Pääbo, S. The major genetic risk factor for severe COVID-19 is inherited from Neanderthals. *Nature* **587**, 610–612 (2020).
108. Nelson, R. Risk variant for severe COVID-19 inherited from Neanderthals. *Am. J. Med. Genet. A* **182**, 2203–2204 (2020).
109. Kwok, A. J., Mentzer, A. & Knight, J. C. Host genetics and infectious disease: new tools, insights and translational opportunities. *Nat. Rev. Genet.* **22**, 137–153 (2021).
110. Patel, P., Hiam, L., Sowemimo, A., Devakumar, D. & McKee, M. Ethnicity and covid-19. *BMJ* **369**, m2282 (2020).
111. Sornette, D., Mearns, E., Schatz, M., Wu, K. & Darcet, D. Interpreting, analysing and modelling COVID-19 mortality data. *Nonlinear Dyn.* **101**, 1751–1776 (2020).
112. Williamson, E. J. et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature* **584**, 430–436 (2020).
113. Rentsch, C. T. et al. Patterns of COVID-19 testing and mortality by race and ethnicity among United States veterans: a nationwide cohort study. *PLoS Med.* **17**, e1003379 (2020).
114. Webb Hooper, M., Nápoles, A. M. & Pérez-Stable, E. J. COVID-19 and racial/ethnic disparities. *JAMA* **323**, 2466–2467 (2020).
115. Lorente, L. et al. HLA genetic polymorphisms and prognosis of patients with COVID-19. *Med. Intens.* **45**, 96–103 (2021).
116. La Porta, C. A. M. & Zapperi, S. Estimating the binding of SARS-CoV-2 peptides to HLA class I in human subpopulations using artificial neural networks. *Cell Syst.* **11**, 412–417 (2020).
117. Tomita, Y., Ikeda, T., Sato, R. & Sakagami, T. Association between HLA gene polymorphisms and mortality of COVID-19: an in silico analysis. *Immun. Inflamm. Dis.* **8**, 684–694 (2020).
118. Yung, Y. L. et al. Association of HLA-B22 serotype with SARS-CoV-2 susceptibility in Hong Kong Chinese patients. *HLA* **97**, 127–132 (2021).
119. Barquera, R. et al. Binding affinities of 438 HLA proteins to complete proteomes of seven pandemic viruses and distributions of strongest and weakest HLA peptide binders in populations worldwide. *HLA* **96**, 277–298 (2020).
120. Mohammadpour, S., Torshizi Esfahani, A., Halaji, M., Lak, M. & Ranjbar, R. An updated review of the association of host genetic factors with susceptibility and resistance to COVID-19. *J. Cell. Physiol.* **236**, 49–54 (2021).
121. Debnath, M., Banerjee, M. & Berk, M. Genetic gateways to COVID-19 infection: implications for risk, severity, and outcomes. *FASEB J.* **34**, 8787–8795 (2020).
122. Kiyotani, K., Toyoshima, Y., Nemoto, K. & Nakamura, Y. Bioinformatic prediction of potential T cell epitopes for SARS-CoV-2. *J. Hum. Genet.* **65**, 569–575 (2020).
123. Littera, R. et al. Human leukocyte antigen complex and other immunogenetic and clinical factors influence susceptibility or protection to SARS-CoV-2 infection and severity of the disease course. The Sardinian experience. *Front. Immunol.* **11**, 605688 (2020).
124. Kolin, D. A., Kulm, S., Christos, P. J. & Elemento, O. Clinical, regional, and genetic characteristics of Covid-19 patients from UK Biobank. *PLoS One* **15**, e0241264 (2020).
125. Wang, W., Zhang, W., Zhang, J., He, J. & Zhu, F. Distribution of HLA allele frequencies in 82 Chinese individuals with coronavirus disease-2019 (COVID-19). *HLA* **96**, 194–196 (2020).
126. Novelli, A. et al. HLA allele frequencies and susceptibility to COVID-19 in a group of 99 Italian patients. *HLA* **96**, 610–614 (2020).
127. Correale, P. et al. HLA-B\*44 and C\*01 prevalence correlates with COVID-19 spreading across Italy. *Int. J. Mol. Sci.* **21**, 5205 (2020).
128. Dijkstra, J., Frenette, A. & Dixon, B. Most Japanese individuals are genetically predisposed to recognize an immunogenic protein fragment shared between COVID-19 and common cold coronaviruses. *F1000Res.* **10**, 196 (2021).
129. Iturrieta-Zuazo, I. et al. Possible role of HLA class-I genotype in SARS-CoV-2 infection and progression: a pilot study in a cohort of Covid-19 Spanish patients. *Clin. Immunol.* **219**, 108572 (2020).
130. Pisanti, S. et al. Correlation of the two most frequent HLA haplotypes in the Italian population to the differential regional incidence of COVID-19. *J. Transl. Med.* **18**, 352 (2020).
131. Prachar, M. et al. Identification and validation of 174 COVID-19 vaccine candidate epitopes reveals low performance of common epitope prediction tools. *Sci. Rep.* **10**, 20465 (2020).
132. Liu, N. et al. The impact of ABO blood group on COVID-19 infection risk and mortality: a systematic review and meta-analysis. *Blood Rev.* **48**, 100785 (2020).
133. Fan, Q. et al. Association between ABO blood group system and COVID-19 susceptibility in Wuhan. *Front. Cell. Infect. Microbiol.* **10**, 404 (2020).
134. Valenti, L. et al. Association of ABO blood group and secretor phenotype with severe COVID-19. *Transfusion* **60**, 3067–3070 (2020).
135. Dai, X. ABO blood group predisposes to COVID-19 severity and cardiovascular diseases. *Eur. J. Prev. Cardiol.* **27**, 1436–1437 (2020).
136. Franchini, M., Crestani, S., Frattini, F., Sissa, C. & Bonfanti, C. ABO blood group and von Willebrand factor: biological implications. *Clin. Chem. Lab. Med.* **52**, 1273–1276 (2014).
137. Terao, C. et al. Quantitative variation in plasma angiotensin-I converting enzyme activity shows allelic heterogeneity in the ABO blood group locus. *Ann. Hum. Genet.* **77**, 465–471 (2013).
138. Chung, C. M. et al. A genome-wide association study identifies new loci for ACE activity: potential implications for response to ACE inhibitor. *Pharmacogenomics J.* **10**, 537–544 (2010).
139. Yamagata University Genomic Cohort Consortium (YUGCC) Pleiotropic effect of common variants at ABO Glycosyltransferase locus in 9q32 on plasma levels of pancreatic lipase and angiotensin converting enzyme. *PLoS One* **9**, e55903 (2014).
140. Mele, D. A., Bista, P., Baez, D. V. & Huber, B. T. Dipeptidyl peptidase 2 is an essential survival factor in the regulation of cell quiescence. *Cell Cycle* **8**, 2425–2434 (2009).
141. Pati, A., Padhi, S., Suvankar, S. & Panda, A. K. Minor allele of interferon-induced transmembrane protein 3 polymorphism (rs12252) is covered against severe acute respiratory syndrome coronavirus 2 infection and mortality: a worldwide epidemiological investigation. *J. Infect. Dis.* **223**, 175–178 (2021).
142. Zhang, Y. et al. Interferon-induced transmembrane protein 3 genetic variant rs12252-C associated with disease severity in coronavirus disease 2019. *J. Infect. Dis.* **222**, 34–37 (2020).
143. Thevarajan, I. et al. Breadth of concomitant immune responses prior to patient recovery: a case report of non-severe COVID-19. *Nat. Med.* **26**, 453–455 (2020).
144. Gómez, J. et al. The interferon-induced transmembrane protein 3 gene (IFITM3) rs12252 C variant is associated with COVID-19. *Cytokine* **137**, 155354 (2021).
145. Nikoloudis, D., Kountouras, D. & Hiona, A. The frequency of combined IFITM3 haplotype involving the reference alleles of both rs12252 and rs34481144 is in line with COVID-19 standardized mortality ratio of ethnic groups in England. *PeerJ* **8**, e10402 (2020).
146. Allen, E. K. et al. SNP-mediated disruption of CTCF binding at the IFITM3 promoter is associated with risk of severe influenza in humans. *Nat. Med.* **23**, 975–983 (2017).
147. Bentson, L. F. et al. New point mutation in Golga3 causes multiple defects in spermatogenesis. *Andrology* **1**, 440–450 (2013).



148. Gordon, D. E. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).
149. Wei, J. et al. Genome-wide CRISPR screens reveal host factors critical for SARS-CoV-2 infection. *Cell* **184**, 76–91.e13 (2021).
150. Chen, R. et al. HMGB1 as a potential biomarker and therapeutic target for severe COVID-19. *Heliyon* **6**, e05672 (2020).
151. Al-Mulla, F. et al. ACE2 and FURIN variants are potential predictors of SARS-CoV-2 outcome: a time to implement precision medicine against COVID-19. *Heliyon* **7**, e06133 (2021).
152. Gemmati, D. & Tisato, V. Genetic hypothesis and pharmacogenetics side of renin-angiotensin-system in COVID-19. *Genes* **11**, 1044 (2020).
153. Purwaningroom, D. L., Saifurrohman, M., Widodo, Putri, J. F. & Lukitasari, M. Alteration of splicing pattern on angiotensin-converting enzyme gene due to the insertion of Alu elements. *IJCB* **4**, 53–58 (2015).
154. Gemmati, D. et al. COVID-19 and individual genetic susceptibility/receptivity: role of ACE1/ACE2 genes, immunity, inflammation and coagulation. Might the double X-chromosome in females be protective against SARS-CoV-2 compared to the single X-chromosome in males? *Int. J. Mol. Sci.* **21**, 3474 (2020).
155. Zheng, H. & Cao, J. J. Angiotensin-converting enzyme gene polymorphism and severe lung injury in patients with coronavirus disease 2019. *Am. J. Pathol.* **190**, 2013–2017 (2020).
156. Delanghe, J. R., Speeckaert, M. M. & De Buyzere, M. L. COVID-19 infections are also affected by human ACE1 D/I polymorphism. *Clin. Chem. Lab. Med.* **58**, 1125–1126 (2020).
157. Delanghe, J. R., Speeckaert, M. M. & De Buyzere, M. L. The host's angiotensin-converting enzyme polymorphism may explain epidemiological findings in COVID-19 infections. *Clin. Chim. Acta* **505**, 192–193 (2020).
158. Hatami, N. et al. Worldwide ACE (I/D) polymorphism may affect COVID-19 recovery rate: an ecological meta-regression. *Endocrine* **68**, 479–484 (2020).
159. Li, X., Sun, X., Jin, L. & Xue, F. Worldwide spatial genetic structure of angiotensin-converting enzyme gene: a new evolutionary ecological evidence for the thrifty genotype hypothesis. *Eur. J. Hum. Genet.* **19**, 1002–1008 (2011).
160. Anastassopoulou, C., Gkizarioti, Z., Patrinos, G. P. & Tsakris, A. Human genetic factors associated with susceptibility to SARS-CoV-2 infection and COVID-19 disease severity. *Hum. Genomics* **14**, 40 (2020).
161. Kasparian, K., Graykowski, D. & Cudaback, E. Commentary: APOE e4 genotype predicts severe COVID-19 in the UK biobank community cohort. *Front. Immunol.* **11**, 1939 (2020).
162. Inal, J. Biological factors linking APOE e4 variant and severe COVID-19. *Curr. Atheroscler. Rep.* **22**, 70 (2020).
163. Kuo, C. L. et al. APOE e4 genotype predicts severe COVID-19 in the UK biobank community cohort. *J. Gerontol. A Biol. Sci. Med. Sci.* **75**, 2231–2232 (2020).
164. Alshahawey, M., Raslan, M. & Sabri, N. Sex-mediated effects of ACE2 and TMPRSS2 on the incidence and severity of COVID-19; the need for genetic implementation. *Curr. Res. Transl. Med.* **68**, 149–150 (2020).
165. Barash, A., Machluf, Y., Ariel, I. & Dekel, Y. The pursuit of COVID-19 biomarkers: putting the spotlight on ACE2 and TMPRSS2 regulatory sequences. *Front. Med.* **7**, 582793 (2020).
166. Singh, H., Choudhari, R., Nema, V. & Khan, A. ACE2 and TMPRSS2 polymorphisms in various diseases with special reference to its impact on COVID-19 disease. *Microb. Pathogenesis* **150**, 104621 (2021).
167. Yildirim, Z., Sahin, O. S., Yazar, S. & Bozok Cetintas, V. Genetic and epigenetic factors associated with increased severity of Covid-19. *Cell Biol. Int.* **45**, 1158–1174 (2021).
168. Choudhary, S., Sreenivasulu, K., Mitra, P., Misra, S. & Sharma, P. Role of genetic variants and gene expression in the susceptibility and severity of COVID-19. *Ann. Lab. Med.* **41**, 129–138 (2021).
169. Irlam, L. M. et al. Genetic variants that influence SARS-CoV-2 receptor TMPRSS2 expression among population cohorts from multiple continents. *Biochem. Biophys. Res. Commun.* **529**, 263–269 (2020).
170. Piva, F., Sabanovic, B., Cecati, M. & Giulietti, M. Expression and co-expression analyses of TMPRSS2, a key element in COVID-19. *Eur. J. Clin. Microbiol. Infect. Dis.* **40**, 451–455 (2021).
171. Hou, Y. et al. New insights into genetic susceptibility of COVID-19: an ACE2 and TMPRSS2 polymorphism analysis. *BMC Med.* **18**, 216 (2020).
172. Foresta, C. & Rocca, M. S. & Di Nisio, A. Gender susceptibility to COVID-19: a review of the putative role of sex hormones and X chromosome. *J. Endocrinol. Invest.* **16**, 1–6 (2020).
173. Kloc, M., Ghobrial, R. M. & Kubiak, J. Z. The role of genetic sex and mitochondria in response to COVID-19 infection. *Int. Arch. Allergy Immunol.* **181**, 629–634 (2020).
174. Li, Y., Jerkic, M., Slutsky, A. S. & Zhang, H. Molecular mechanisms of sex bias differences in COVID-19 mortality. *Crit. Care* **24**, 405 (2020).
175. Viveiros, A. et al. Sex differences in COVID-19: candidate pathways, genetics of ACE2, and sex hormones. *Am. J. Physiol. Heart Circ. Physiol.* **320**, H296–H304 (2021).
176. Gadi, N., Wu, S. C., Spihlman, A. P. & Moulton, V. R. What's sex got to do with COVID-19? Gender-based differences in the host immune response to coronaviruses. *Front. Immunol.* **11**, 2147 (2020).
177. Wang, J. et al. Unusual maintenance of X chromosome inactivation predisposes female lymphocytes for increased expression from the inactive X. *Proc. Natl. Acad. Sci. USA* **113**, E2029–E2038 (2016).
178. Kelada, M., Anto, A., Dave, K. & Saleh, S. N. The role of sex in the risk of mortality from COVID-19 amongst adult patients: a systematic review. *Cureus* **12**, e10114 (2020).
179. Penna, C., Mercurio, V., Tocchetti, C. G. & Pagliaro, P. Sex-related differences in COVID-19 lethality. *Br. J. Pharmacol.* **177**, 4375–4385 (2020).
180. van de Veerdonk, F. L. & Netea, M. G. Rare variants increase the risk of severe COVID-19. *Elife* **10**, e67860 (2021).
181. van der Made, C. I. et al. Presence of genetic variants among young men with severe COVID-19. *JAMA* **324**, 663–673 (2020).
182. Fallerini, C. et al. Association of Toll-like receptor 7 variants with life-threatening COVID-19 disease in males: findings from a nested case-control study. *Elife* **10**, e67569 (2021).
183. Ghafouri-Fard, S. et al. Effects of host genetic variations on response to, susceptibility and severity of respiratory infections. *Biomed. Pharmacother.* **128**, 110296 (2020).
184. Darbani, B. The expression and polymorphism of entry machinery for COVID-19 in human: juxtaposing population groups, gender, and different tissues. *Int. J. Environ. Res. Public Health* **17**, 3433 (2020).
185. Devaux, C. A., Rolain, J. M. & Raoult, D. ACE2 receptor polymorphism: susceptibility to SARS-CoV-2, hypertension, multi-organ failure, and COVID-19 disease outcome. *J. Microbiol. Immunol. Infect.* **53**, 425–435 (2020).
186. Tipnis, S. R. et al. A human homolog of angiotensin-converting enzyme. Cloning and functional expression as a captopril-insensitive carboxypeptidase. *J. Biol. Chem.* **275**, 33238–33243 (2000).
187. Yan, R. et al. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**, 1444–1448 (2020).
188. Cao, Y. et al. Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. *Cell Discov.* **6**, 11 (2020).
189. Mariappan, V., Rao, S. R. & Balakrishna Pillai, A. Angiotensin-converting enzyme 2: a protective factor in regulating disease virulence of SARS-CoV-2. *IUBMB Life* **72**, 2533–2545 (2020).
190. Iyer, G. R. et al. Infectivity and progression of COVID-19 based on selected host candidate gene variants. *Front. Genet.* **11**, 861 (2020).
191. Wu, Y. H., Li, J. Y., Wang, C., Zhang, L. M. & Qiao, H. The ACE2 G8790A polymorphism: involvement in type 2 diabetes mellitus combined with cerebral stroke. *J. Clin. Lab. Anal.* **31**, e22033 (2017).
192. Asselta, R., Paraboschi, E. M., Mantovani, A. & Duga, S. ACE2 and TMPRSS2 variants and expression as candidates to sex and country differences in COVID-19 severity in Italy. *Aging* **12**, 10087–10098 (2020).
193. Srivastava, A. et al. Genetic association of ACE2 rs2285666 polymorphism with COVID-19 spatial distribution in India. *Front. Genet.* **11**, 564741 (2020).
194. Li, Q., Cao, Z. & Rahman, P. Genetic variability of human angiotensin-converting enzyme 2 (hACE2) among various ethnic populations. *Mol. Genet. Genom. Med.* **8**, e1344 (2020).
195. Lippi, G., Lavie, C. J., Henry, B. M. & Sanchis-Gomar, F. Do genetic polymorphisms in angiotensin converting enzyme 2 (ACE2) gene play a role in coronavirus disease 2019 (COVID-19)? *Clin. Chem. Lab. Med.* **58**, 1415–1422 (2020).
196. Hussain, M. et al. Structural variations in human ACE2 may influence its binding with SARS-CoV-2 spike protein. *J. Med. Virol.* **92**, 1580–1586 (2020).
197. Benetti, E. et al. ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population. *Eur. J. Hum. Genet.* **28**, 1602–1614 (2020).
198. Chen, J. et al. Individual variation of the SARS-CoV-2 receptor ACE2 gene expression and regulation. *Aging Cell* **19**, e13168 (2020).
199. Heinzelman, P. & Romero, P. A. Discovery of human ACE2 variants with altered recognition by the SARS-CoV-2 spike protein. *PLoS One* **16**, e0251585 (2021).
200. McCoy, J. et al. Racial variations in COVID-19 deaths may be due to androgen receptor genetic variants associated with prostate cancer and androgenic alopecia. Are anti-androgens a potential treatment for COVID-19? *J. Cosmet. Dermatol.* **19**, 1542–1543 (2020).
201. Carethers, J. M. Insights into disparities observed with COVID-19. *J. Intern. Med.* **289**, 463–473 (2020).
202. Wambier, C. G. et al. Androgen sensitivity gateway to COVID-19 disease severity. *Drug Dev. Res.* **81**, 771–776 (2020).

203. Mjaess, G., Karam, A., Aoun, F., Albisinni, S. & Roumeguère, T. COVID-19 and the male susceptibility: the role of ACE2, TMPRSS2 and the androgen receptor. *Prog. Urol.* **30**, 484–487 (2020).
204. Sironi, M. et al. SARS-CoV-2 and COVID-19: a genetic, epidemiological, and evolutionary perspective. *Infect. Genet. Evol.* **84**, 104384 (2020).
205. 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.* **28**, 715–718 (2020).
206. Asselah, T., Durantel, D., Pasmant, E., Lau, G. & Schinazi, R. F. COVID-19: discovery, diagnostics and drug development. *J. Hepatol.* **74**, 168–184 (2021).
207. Milne, R. Societal considerations in host genome testing for COVID-19. *Genet. Med.* **22**, 1464–1466 (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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021