Review

An overview of human proteins and genes involved in SARS-CoV-2 infection

Zohreh Jahanafrooz [a,*], Zhishan Chen [b], Jiandong Bao [c], Hongzhi Li [d], Loren Lipworth [b], Xingyi Guo [b,e,*]

[a] Department of Biology, Faculty of Sciences, University of Maragheh, Maragheh, Iran
[b] Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN 37203, USA
[c] College of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou, Fujian 350002, China
[d] Department of Molecular Medicine, City of Hope National Medical Center, Duarte, CA 91010, USA
[e] Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, TN 37203, USA

ARTICLE INFO

ABSTRACT

As of July 2021, the outbreak of coronavirus disease 2019 (COVID-19), caused by SARS-CoV-2, has led to more than 200 million infections and more than 4.2 million deaths globally. Complications of severe COVID-19 include acute kidney injury, liver dysfunction, cardiomyopathy, and coagulation dysfunction. Thus, there is an urgent need to identify proteins and genetic factors associated with COVID-19 susceptibility and outcome. We comprehensively reviewed recent findings of host-SARS-CoV-2 interactome analyses. To identify genetic variants associated with COVID-19, we focused on the findings from genome and transcriptome wide association studies (GWAS and TWAS) and bioinformatics analysis. We described established human proteins including ACE2, TMPRSS2, 40S ribosomal subunit, ApoA1, TOM70, HLA-A, and PALS1 interacting with SARS-CoV-2 based on cryo–electron microscopy results. Furthermore, we described approximately 1000 human proteins showing evidence of interaction with SARS-CoV-2 and highlighted host cellular processes such as innate immune pathways affected by infection. We summarized the evidence on more than 20 identified candidate genes in COVID-19

severity. Predicted deleterious and disruptive genetic variants with possible effects on COVID-19 infectivity have been also summarized. These findings provide novel insights into SARS-CoV-2 biology and infection as well as potential strategies for development of novel COVID therapeutic targets and drug repurposing.

1. Epidemiology and clinical presentation of COVID-19

Coronavirus disease 2019 (COVID-19) is caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Zhu et al., 2020). The outbreak of COVID-19 has been characterized as a global pandemic, and as of July 2021, there were more than 200 million people infected worldwide and more than 4.2 million confirmed deaths (https://coronavirus.1point3acres.com/) (Huang et al., 2020; Liang et al., 2020). The overall mortality rate is approximately 2%, and 23% of COVID-19 patients have severe disease, with 11% requiring intensive care and 7% requiring mechanical ventilation (Li et al., 2021b; Wiersinga et al., 2020). Among hospitalized COVID-19 patients, 75% have pneumonia requiring supplemental oxygen and 15% have acute respiratory distress syndrome with hypoxic respiratory failure being the primary reason for intensive care unit (ICU) admission and the primary cause of death. Case fatality rises to 40% among patients hospitalized in the ICU. The strongest independent risk factors for critical COVID-19 illness are older age, immunosuppression, and comorbidities such as hypertension, diabetes, cardiovascular disease, and obesity (Docherty et al., 2020; Petrilli et al., 2020).

Early in infection, SARS-CoV-2 targets nasal and bronchial epithelial cells; as viral replication accelerates in later stages of infection, the virus directly infects the endothelial cells in multiple organs, including pulmonary capillary endothelial cells, leading to endothelial barrier disruption and diffuse endothelial cell injury (Ackermann et al., 2020; Levi et al., 2020; Varga et al., 2020; Wiersinga et al., 2020). Clinical manifestations of severe COVID-19 include low leukocyte and lymphocyte count; an accentuated inflammatory response including large rapid release of inflammatory signaling molecules, such as C-reactive protein, tumor necrosis factor-α, interleukin-1 (IL-1), and IL-6; and abnormal coagulation parameters, such as thrombocytopenia and elevated D-dimer, suggesting the presence of a hypercoagulable state (Levi et al., 2020). Approximately 10–15% of critically ill COVID-19 patients develop a venous or arterial thromboembolic event, with that percentage increasing to >30% among ICU patients. Other common complications of severe COVID-19 include acute kidney injury, liver dysfunction, ventricular arrhythmia, myocarditis, neurological manifestations (including acute cerebrovascular disease), shock (AlGhatrif et al., 2020), and ophthalmic manifestations (including conjunctivitis, conjunctival hyperemia and chemosis) (Jevnikar et al., 2021). Thus, identification of influencing host proteins and genetic factors is essential for understanding COVID-19 susceptibility and outcomes and development of targeted COVID-19 prevention and therapeutic strategies. In this review, we comprehensively summarized recent progress and findings related to host proteins that interact with SARS-CoV-2 and genetic variants associated with COVID-19 susceptibility and severity.

2. Key human proteins for COVID-19 infectivity

The genome of SARS-CoV-2 has been established with approximately 30 kb including 14 open-reading frames (ORFs), which encode 29 proteins (Gordon et al., 2020b; Wu et al., 2020). These include 16 non-structural proteins (NSP1–NSP16) that form the replicase-transcriptase complex, four typical structural proteins, spike (S), envelope (E), membrane (M) and nucleocapsid (N), and nine accessory proteins from a complement of 3′ ORFs (ORF3a, 3b, 6, 7a, 7b, 8, 9b, 9c, and 10). Of these proteins, the surface spike glycoprotein (S protein) is responsible for SARS-CoV-2 cellular entry through recognition of the peptidase domain of the angiotensin converting enzyme (ACE2) in humans (Li et al., 2005; Yan et al., 2020). The process of cellular entry into host cells requires i)

the receptor binding domain (RBD) of the surface unit (S1) of S protein to bind to ACE2, which facilitates viral attachment to the surface of host cells (Gallagher and Buchmeier, 2001), and ii) the S2 subunit of S protein to drive S protein priming by cellular proteases, which facilitates the fusion of the viral and cellular membranes (Li et al., 2003). Following the virus attachment to ACE2, the S1/S2 junction in S protein is cleaved by the cellular transmembrane protease serine 2 (TMPRSS2). This causes the internalization and fusion process of the virus to the host cell (Hoffmann et al., 2020).

Remarkable progress though structural analysis by cryo-electron microscopy (CryoEM) or X-ray crystallography has revealed the structure of six viral proteins, including S protein, NSP1, ORF3a, ORF9b, E and N proteins in complex with human proteins. SARS-CoV-2 NSP1 binding to human 40S ribosomal subunit can lead to shutdown of host mRNAs and immune evasion (Schubert et al., 2020; Thoms et al., 2020). ORF3a protein is one putative ion channel in SARS-CoV-2 (Kern et al., 2020). In SARS-CoV-1, ORF3a protein is implicated in viral release, inflammasome activation, and cell death; its deletion reduces viral titer and morbidity in animal models (Kern et al., 2020). SARS-CoV-2 ORF3a can bind to human apolipoprotein A1 (ApoA1) protein in *in vitro* lipoprotein nanodiscs (Schubert et al., 2020; Thoms et al., 2020). ApoA1 is a component of high-density lipoprotein (HDL), a cholesterol and phospholipid blood transporter to the liver (Kern et al., 2020). ORF9b protein is localized to mitochondria in infected cells via interaction with mitochondrial import receptor subunit TOM70, which is one major import receptor in the TOM complex (Gordon et al., 2020a; Jiang et al., 2020). The human translocase of the outer membrane (TOM) is a multi-subunit protein complex located in the outer mitochondrial membrane of the mitochondria, which is involved in the recognition, unfolding, and translocation of preproteins into the mitochondria. As with SARS-CoV-1, the C-terminal domain of protein E forms a novel PDZ-binding motif that binds to the PDZ domain of human PALS1 and causes its translocation from tight junction to endoplasmic reticulum (ER)-Golgi intermediate compartment (ERGIC). As a consequence, it dissociates the intercellular tight junction or seal formation between adjacent epithelial cells in epithelia of various organs resulting in alteration of cellular polarity, epithelial morphogenesis, cytokine storm initiation, and virus dissemination (Chai et al., 2021). PALS1 has a different cellular function and is a main component of CRB complex in the intercellular tight junction. The hijacking of human PALS1 by viral protein E may play a critical role in the disruption of the lung epithelium in SARS patients (Chai et al., 2021; Teoh et al., 2010). Moreover, similar to SARS-CoV-1, the N-terminal transmembrane domain of E protein forming the cation channel in ERGIC membrane can cause an increase in cytosolic concentration of Ca2 + and possible activation of host inflammasome (Mandala et al., 2020). There are some immunogenic N protein-derived epitopes that bind to the human leukocyte antigens (HLA)-A of all nucleated cells presented to CD8 + T cells and stimulate them. A study showed that HLA-A02:01 individuals have limited numbers of stable N-derived epitopes-HLA-A complexes, which may help us to understand the CD8 + T cell response toward SARS-CoV-2 infection and determine if a CD8 + T-cell-mediated vaccine can provide broad and long-lasting protection (Szeto et al., 2021). Lower SARS-CoV-2 antigen presentation capacity of HLA-A02:01 compared to other frequent HLA-A polymorphisms (i.e., HLA-A11:01 and HLA-A24:02) was also reported in an in silico analysis which can be a reason for inter-individual variation in COVID-19 severity (Tomita et al., 2020).

3. Characterized human proteins with evidence of interaction partners with SARS-CoV-2

Numerous studies have characterized human proteins from different cellular pathways with evidence of interaction partners of SARS-CoV-2. We reviewed previous studies using experimental protocols for protein-protein-interactions (i.e., affinity purification mass spectrometry [AP-MS] and AP-MS combined with liquid chromatography [AP-LC-MS/MS]) or viral protein and human protein interactomes. More than 1000 proteins have interaction partners with viral proteins (Table S). By interacting with host cell proteins, SARS-CoV-2 disrupts normal cellular function and employs cellular equipment to replicate itself. Here, some recognized host protein-SARS-CoV-2 interactions are summarized according to their relevant disturbed cellular pathways:

3.1. Host translation

Switching translation from host cell mRNAs toward viral mRNAs primarily occurs in viral infected cells, and NSP1 is a proposed viral protein responsible for this phenomenon in SARS-CoV-2 infection by interacting with h18 of the 18S rRNA and uS3 protein in the head and uS5 and eS30 proteins in the body of 40S ribosomal subunit (Schubert et al., 2020; Thoms et al., 2020). Binding of the C-terminal domain of NSP1 to the 40S subunit in ribosomal complexes causes the mRNA entry channel to the ribosome to close, which thereby decreases the available number of active ribosomes and reduces the global cellular translation. However, mRNAs with viral 5′UTR are translated more efficiently compared to mRNAs with host 5′UTR in a cap-dependent translation manner. Under these circumstances, expression of one important host-antiviral mRNA "type I interferon (IFN)" is also decreased leading to enhanced viral replication (Thoms et al., 2020). NSP2 targets eIF4E2 (one of the initiation factors in translation) to employ host translation machinery, and Nsp8 interacts with three components of the signal recognition particle (SRP) to translocate viral proteins synthesizing the ribosome to the ER surface (Gordon et al., 2020b).

3.2. Autophagy

ORF3a of SARS-CoV-2 has been known to form ion channels in host cell membranes (Kern et al., 2020). In related SARS-CoV-1, ORF3a forms channels in late endosomal, lysosomal, and *trans*-Golgi-network membranes and is implicated in viral release, inflammasome activation, non-inflammatory apoptotic cell death, and necrotic cell death (Yue et al., 2018). Its deletion reduces viral titer and morbidity in animal models. ORF3a from SARS-CoV-2 is thought to be an important inhibitor of autophagy and SARS-CoV-2 rescue because ORF3a binds to VPS39 of the HOPS complex on the lysosome and late endosome membranes, which negatively impacts the fusion complex between autophagosome-lysosome (Miao et al., 2021).

3.3. Cell death

Cell death is manipulated by SARS-CoV-2 at the early infection stage to create a bio-factory for its own replication. Apoptosis and necrosis are two types of cell death following virus infection. Mitochondria play a major role in triggering the intrinsic apoptosis in cellular stress and abnormal conditions like virus infection (Atkin-Smith et al., 2018). Interaction between ORF9b and the substrate binding site of TOM70 can suppress IFN responses (Jiang et al., 2020) and may also influence mitochondrial import efficiency and result in modulation of apoptotic signaling (Gordon et al., 2020a). A decrease in the expression level of TOM70 during SARS-CoV-2 infection has been reported (Gordon et al., 2020a). Permeability changes in the mitochondrial membrane can lead to a release of cytochrome *C*, then apoptosome formation in the cytosol in turn followed by caspase-9 activation and cell death (Suhaili et al., 2017). Global analysis of protein-protein interaction has identified

interaction between NSP12 (RNA-dependent RNA polymerase) and receptor-interacting serine/threonine protein kinase 1 (RIPK1) (Gordon et al., 2020b). Death receptors such as tumor necrosis factor receptor 1 (TNFR1) can result in activation of RIPK1 and possible phosphorylation of mixed lineage kinase domain-like protein (MLKL), which then disturbs cell membranes and cause a specific type of cell death called necroptosis. By interaction with FADD-caspase-8 axis, RIPK1 can result in apoptosis (Eng et al., 2020). In addition to RIPK1-dependent cell death, there is also RIPK1-dependent inflammatory signaling and pro-survival signaling via NF-κB- and MAPK-signaling; and RIPK1 dysregulation has shown association with hyper-inflammation (discussed in immune response section) (Eng et al., 2020). Remdesivir, the first Food and Drug Administration (FDA) approved drug for the treatment of COVID-19, inhibits NSP12 (Gordon et al., 2020b).

3.4. Cell cycle and cell growth

Dysregulated cell cycle is a hallmark of a virus infected cell, believed to result from interaction of viral proteins with cell cycle players. DNA polymerase alpha complex (a primary DNA synthase in S phase) was found to interact with NSP1, while centrosome complex (a major player in mitosis phase and mitotic spindle formation) interacts with NSP13 (Gordon et al., 2020a; Gordon et al., 2020b). A study has systematically identified 332 human proteins showing high-confidence protein-protein interactions with SARS-CoV-2 proteins (Gordon et al., 2020b). In other work from the same group, the phosphorylation landscape has been characterized based on a quantitative MS-based phosphoproteomics investigation of SARS-CoV-2 infection in Vero E6 cells (Bouhaddou et al., 2020). It has been suggested that SARS-CoV-2 infection could rewire activity regulation of 97 kinases, including promotion of several members of the p38/MAPK cascade (leading to cytokine production), downregulation of cell growth-related kinases, and inhibition of mitotic kinases (leading to S/G2 cell cycle arrest). Some of the dysregulated kinases directly interact with viral proteins. For instance, microtubule affinity regulating kinase 2 (MARK2) interacts with ORF9b or protein kinase A catalytic subunit alpha (PRKACA) interacts with NSP13. MARK2 phosphorylates a large number of substrates involved in cell growth, cell polarity, cell cycle progression; and PRKACA also contributes to different cellular processes, such as cell cycle and glucose metabolism (Tutuncuoglu et al., 2020). In addition, 40 proteins of the 332 identified SARS-CoV-2 interacting proteins are significantly differentially phosphorylated upon infection (Bouhaddou et al., 2020). By imposing global change in host kinases signaling, SARS-CoV-2 provides phosphorylated or dephosphorylated interactors for itself,; for example, detected interactors for NSP12 are hypo-phosphorylated while for NSP8 are mostly hyper-phosphorylated (Bouhaddou et al., 2020).

3.5. Immune responses

As a pathogen, SARS-CoV-2 infection causes significant changes in both innate and adaptive immune response. Often upon virus infection, production of IFN by most cells induces a downstream cascade that blocks viral replication and leads to cytokine production. In the case of SARS-CoV-1, NSP1, NSP3, ORF3b, ORF6 are detected as IFN antagonists and exert this effect through some mechanisms in the IFN pathway, for instance, blocking the entrance of STAT to nucleus upon IFN or IL-6 stimulation and blocking phosphorylation of IRF3, both of which are transcription factors to control gene expression. In contrast, because of major differences in amino acid sequences between the SARS-CoV-1 protein and SARS-CoV-2, IFN antagonist function is weaker in SARS-CoV-2 and this novel virus is more sensitive to IFN (Lokugamage et al., 2020). Interestingly, association of decreased expression of the cell surface receptor for IFN (IFNAR2) with hospitalization for COVID-19 has been reported (Pairo-Castineira et al., 2021). In one study, it was shown that the N protein of SARS-CoV-2 suppresses phosphorylation/activation and nuclear translocation of STAT1 and STAT2 and

subsequently shuts down the antiviral IFN pathway. SARS-CoV-2 N protein can interact directly with STAT1 and STAT2 and interferes with their interactions with related receptor-associated kinases JAK1 and TYK2, respectively; therefore, N protein is a potent antagonist of IFN signaling in SARS-CoV-2 (Mu et al., 2020). Viral proteins control host gene expression by interaction with both transcription factors and epigenetic regulators; Nsp5 targets histone deacetylase 2 (HDAC2) and cleaves it, which blocks HDAC2 translocation to the nucleus and modulates IFN and immune response related gene expression (Gordon et al., 2020b). Based on CryoEM, AP-MS, and coronavirus-host interactomes analysis, Gordon et al. identified 73 host genetic factors that can induce significant changes in SARS-CoV-2 replication (Gordon et al., 2020a). Of these, they highlighted the physical interaction between ORF8 and IL-17 receptor A (IL-17RA). They further showed that knock-down of IL-17RA leads to a decrease in SARS-CoV-2 replication in A549-ACE2 cells, and that decreased levels of soluble IL-17RA are associated with severe COVID-19. Another important viral-host interaction was reported between prostaglandin E synthase 2 (PGES-2)-Nsp7, which is inhibited by the FDA-approved prescription nonsteroidal anti-inflammatory drug (NSAID) indomethacin (Gordon et al., 2020a). Genome-wide proteomic screening for viral-host protein interaction predicted 286 human proteins which may have interactions with SARS-CoV-2 proteins (Li et al., 2021a). The host-viral protein interactions for TBK1-TANK-TRAF2 ternary complex-ORF6 and ORF3a-KIT were also reported in the study (Li et al., 2021a). TBK1-TANK-TRAF2 ternary complex plays critical roles in IFN and NF-kB pathways upon viral infection, and KIT as a tyrosine kinase facilitates the activation of STATs (Li et al., 2021a). NF-κB-repressing factor (NKRF)-Nsp10 interaction was highlighted with additional evidence that it facilitates IL-8 induction, resulting in subsequent chemotaxis of neutrophils and triggering an inflammatory response. Interaction between Nsp13 and transducin-like enhancer protein 1 (TLE1), TLE3, and TLE5 (which are transcriptional co-repressors of NF-kB) also influences the NF-κB pathway (Kumar et al., 2020). As mentioned in the cell death section, RIPK1 is among the interactors of NSP12; its activation can lead to various outcomes in cells, either cell death or cell growth, dependent on cellular context and infection context. RIPK1 is a downstream kinase for different activated receptors, including TNFR1, toll-like receptors (TLRs), and other immune pattern recognition receptors; RIPK1 activation can trigger NF-κB activation and pro-inflammatory cytokine production, MAPK activation and anti-apoptotic proteins production, and IRF3 activation and IFN production (Eng et al., 2020).

3.6. Other cellular process

A comparative genomic analysis strategy analyzing the IntAct database, Gene Ontology, and Reactome pathways prediction reported that ORF1ab interacts with several host factors, including TLE1, NKRF (a potential regulator of IL-8), and U4/U6-U5 tri-snRNP complex, U4/U6 small nuclear ribonucleoprotein Prp3 (PRPF3), SRSF protein kinase 1 (SRPK1), and PUF60 (components of spliceosome). U4/U6-U5 tri-snRNP complex, PRPF3, SRPK1, and PUF60 play a vital role in pre-mRNA splicing which emphasizes that SARS-CoV-2 manipulates spliceosome machinery, its accuracy, and alternative splicing. Furthermore, an interaction was shown between Nsp8 and POGLUT2, POGLUT3, and POFUT1, which modulate the Notch signaling pathway by fucosylation of Notch1 and also regulate transport of Notch1 and Notch3 to the plasma membrane (Kumar et al., 2020). Notch signaling has a wide range of regulatory functions through transcriptional induction of Notch target genes in different cell types; for example, differentiation of precursor cells to absorptive cells in gut, common lymphoid precursor differentiation to T and B lymphocyte in bone marrow, and maintaining of stem-cell state in adult stem cells (Shang et al., 2016). Nsp2 can interact with ERLIN1/2 and prohibitin complexes, which play a regulatory role in mitochondrial function and Ca2 + flux at ER. They also showed that Nsp4 can interact with N-linked glycosylation

machinery in ER membranes, unfolded protein response (UPR) of ER associated proteins, and antiviral innate immune signaling factors (Davies et al., 2020). Indeed, SARS-CoV-2 employs ER and Golgi for its benefit. Furthermore, Nsp6-SIGMAR1 and ORF9C-SIGMAR2 interactions have been reported which influence lipid composition of ER membranes and impose alteration in SIGMARs function. SIGMARs, as ER stress "gatekeepers," are overexpressed in proliferating tumor cells and activate the PI3K/AKT pathway (a cell growth pathway) and cell motility (Tesei et al., 2018). Drug modulation of SIGMARs showed antiviral activity (Gordon et al., 2020b). E protein is the other viral interactor with BRD2 and BRD4 in ER and Golgi membranes; BRD2 and BRD4 belong to the bromodomain and extraterminal domain (BET) family, which recognizes acetylated histones and recruits other transcription factors to regulate immune gene expression. Noticeably, BRD4 regulates the NF-κB pathway, which is activated upon pathogen binding to endosomal or cell surface TLRs or cytokine binding to cell surface cytokine receptors (Wang et al., 2021).

Interactome and proteome studies have characterized other host cellular processes, including vesicle trafficking, cytoskeleton arrangement, ubiquitin ligation, extracellular matrix, and nuclear transport machinery, and IL-1, IL-6, and chemokine signaling involved in SARS-CoV-2 infectivity (Bouhaddou et al., 2020; Gordon et al., 2020a; Gordon et al., 2020b). Most recently, a proteomic differential analysis of 144 autopsy samples (from lung, spleen, liver, heart, kidney, testis, and thyroid) of 19 COVID-19 patients compared to normal controls has identified thousands of proteins involved in immune system, hypoxia, angiogenesis, blood coagulation, fibrosis in multiple organs, glucose and fatty acid metabolism, and testicular injuries leading to multi-organ injuries (Nie et al., 2021). Together these findings provide strong evidence that thousands of human proteins significantly contribute to complexes and biological processes, including key host cellular processes, kinases, epigenetic and gene-expression regulators, and innate immune pathways underlying SARS-CoV-2 biology and infectivity.

4. Host susceptibility variants and genes identified through genome-wide association studies (GWAS)/transcriptome-wide association studies (TWAS)/integrative genomic analyses

Although remarkable progresses has been made in identifying viral-human proteins from protein-protein interaction and virus-host interactome analysis, whether their related genetic variants affect SARS-CoV-2 recognition and infection could not be directly inferred. To identify common genetic susceptibility variants and genes, both GWAS and TWAS approaches have been applied to evaluate associations between genetic variants and COVID-19 susceptibility and severity. A pioneering GWAS including 835 patients with severe COVID-19 (defined as respiratory failure) and 1255 control participants from Italy, together with 775 patients and 950 control participants from Spain, was conducted by the Severe COVID-19 GWAS Group (Ellinghaus et al., 2020). A risk locus at 3p21.31 was reported to be associated with COVID-19 respiratory failure at GWAS significance level. Candidate susceptibility genes including SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6, and XCR1 in the locus have been suggested. Most recently, a study from the GenOMICC (Genetics Of Mortality In Critical Care), using genetic data from 2244 critically ill COVID-19 patients from 208 UK ICUs, has identified genome-wide significant associations with severe COVID-19, on chr12q24.13 (nearby genes, OAS1, OAS2, OAS3), on chr19p13.2 (TYK2), on chr19p13.3 (DPP9), and on chr21q22.1 (IFNAR2). Based on Mendelian randomization and TWAS approaches, this study has further reported genetically predicted expression of IFNAR2, TYK2, and CCR2 associated with severe COVID-19 (Pairo-Castineira et al., 2021). Wu et al. identified candidate genes for severe COVID-19 through a cross methylome omnibus test and S-PrediXcan analyses in blood tissue, and their integrative multiomics analysis confirmed eight of the candidate genes (XCR1, CCR2, SACM1L, OAS3, IFNAR2, NSF, WNT3, and NAPSA) as putative causal genes for COVID-19 severity in lung tissue (Wu et al.,

2021). Moreover, Pathak et al integrated mRNA/splicing/protein transcriptome-wide association studies (TWAS/spTWAS/PWAS) with GWAS data and identified 27 genes involved in inflammation and coagulation pathways (Table 1). They reported nine genes (including LZTFL1, DPP9, IL10RB, IFNAR2, OAS3, FYCO1, ABO, OAS1, and XCR1) significantly associated with COVID-19 severity identified by both TWAS and spTWAS; ABO and OAS1 genes were also implicated by PWAS. In addition, they highlighted a functional role for the 27 genes through phenome- and laboratory-wide association scanning in Vanderbilt Biobank (n = 85,460) (Pathak et al., 2021).

Pathogenic coding variants may have greater potential to affect disease phenotype than GWAS-identified non-coding variants. Based on whole exome sequencing (WES) analysis on a small cohort of 131 SARS-CoV-2 infected (COVID +) cases and 258 COVID test-negative (COVID-) controls from Italy, gene-based analysis of the collapsing coding variants in ACE2 showed a significant risk association with $P < 0.03$ (Benetti et al., 2020). Specifically, several highlighted missense variants (V749V, N720D, K26R, and G211R) were reported in the ACE2 gene with a potential effect on the interaction between S protein and receptor (Benetti et al., 2020). N720D was also reported as a variant associated with infection rates of SARS-CoV-2 in a Spanish population (Stawiski et al., 2020). In other work though WES analysis in four young patients with severe COVID-19 from two unrelated families of European Ancestry, loss-of-function variants (i.e. c.2129_2132del and V795F) in TLR7 were identified (van der Made et al., 2020). It has been suggested that the loss-of-function variants in TLR7 caused a delay in IFN responses induced by TLRs (van der Made et al., 2020). Zhang et al. conducted WES in genomic DNA samples from 659 patients with life-threatening COVID-19 pneumonia and 534 asymptomatic or benign infected subjects with various ancestries. They focused on loss-of-function and other potential deleterious coding variants at 13 human loci that govern immunity to influenza virus, including TLR3, IRF7, IRF9, TICAM1/TRIF, UNC93B1, TRAF3, TBK1, IRF3, NEMO/IKBKG, IFNAR1, IFNAR2, STAT1, and STAT2. Their findings showed that 23 patients (3.5%) carried 24 deleterious variants in eight genes, including TLR3, IRF7, UNC93B1, TICAM1, TBK1, IRF3, IFNAR1, and IFNAR2 variants (Zhang et al., 2020). In Chinese populations, whole genome sequencing (WGS) (greater than45X) has been conducted to identify coding variants in 332 COVID-19 patients categorized by five levels of disease severity, including asymptomatic, mild, moderate, severe, and critically ill patients (Wang et al., 2020a). A GWAS-identified variant rs6020298 in TMEM189-UBE2V1 that is involved in the IL-1 signaling pathway and a missense variant (V197M) in TMPRSS2 have been suggested to be associated with COVID-19 severity (Wang et al., 2020a). Genotyping of the SNPs in the gene TMPRSS2 showed significant association of CC genotype of rs383510 in the intronic sequence of TMPRSS2 with increased SARS-CoV-2 infection risk in a German case-control study (239 cases and 253 controls). However, TT or TC genotype of rs383510 had been shown to be significantly associated with increased H7N9 influenza A virus infection (Cheng et al., 2015; Schönfelder et al., 2021). Effect of rs383510 polymorphism on TMPRSS2 expression level has not been clearly identified and it may cause up-regulation or down-regulation of TMPRSS2 depending on type of infection (Schönfelder et al., 2021).

Together, the reported susceptibility variants and genes identified through GWAS, TWAS and sequencing in COVID-19 patients are presented in Table 1. In addition to genes involved in SARS-CoV-2 entrance to host cells (including ACE2, TMPRSS2, Furin, DPP7, and DPP9) most susceptibility variants belong to immunity and antiviral genes such as IFNAR2, TYK2, and OAS1, which have been highlighted in relation to critical presentation of COVID-19. It is also evident that IFN binding to IFNR1 and IFNR2 induced a cross-link of these receptors and activation of Janus kinases family like TYK2 and JAK1, which cross-phosphorylate each other on tyrosine and subsequently activate the JAK-STAT signaling pathway (Yan et al., 1996). As one of the downstream induced proteins of INF signaling pathway, OAS1 decreases viral

Table 1
Findings from genetic susceptibility studies in COVID-19 patients.

Locus/noncoding or coding variant	Chr	Gene(s)	Study note
GWAS/TWAS, including 835 (Italian) + 775 (Spanish) patients and 1255 (Italian) + 950 controls (Spanish) (Ellinghaus et al., 2020)			
3p21.31	3	SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6, and XCR1	Significant association of locus 3p21.31 with respiratory failure in COVID-19 patients Association of rs11385942 as the lead variant at 3p21.31 with low expression of CXCR6 and high expression of SLC6A20
rs657152	9	ABO	Confirmation of blood group A as a risk factor
GWAS/TWAS, including 2415 (UK ancestry) + 1128 (EUR ancestry) patients and 477,741(UK ancestry) + 679,531 (EUR ancestry) controls (Pairo-Castineira et al., 2021)			
12q24.13/ rs10735079	12	OAS1, OAS2, and OAS3	Higher expression association of OAS3 with sever COVID-19
19p13.2/rs74956615	19	TYK2	Significant association of high expression of TYK2 with sever COVID-19
19p13.3/rs2109069	19	DPP9	
21q22.1/rs2236757	21	IFNAR2	Association of low expression of IFNAR2 with sever COVID-19
3p21.31	3	CCR2, CCR3, and CXCR6	Association of high expression of CCR2 with severe COVID-19
A GWAS, including 1778 COVID-19 cases (Hu et al., 2021)			
Super-variant chr6_148	6	STXBP5/STXBP5-AS1	Significant association between decreased patients survival with 2 SNPs within chr6_148 through mechanisms related to endothelial exocytosis
Super-variant chr8_99	8	CPQ	Location of all the 7 identified SNPs of chr8_99 in the intron of CPQ
Super-variant chr16_4	16	CLUAP1	Possible relation of single SNP of chr16_4 to cilia dysfunctions
Super-variant	17	WSB1	Association of variation within chr17_26 to decreased survival probability Possible relation of 2 SNPs within chr17_26 to the function of IL-21 receptor and innate immunity
Super-variant chr2_197	2	DNAH7/SLC39A10	Association of variation within ch2_197 to decreased survival probability Possible relation of 3 SNPs in DNAH7 with cilia dysfunctions and 3 SNPs in SLC39A10 with B cell development
Super-variant chr2_221	2	DES/SPEG	Possible association of single SNP located in the downstream of DES gene and the upstream of SPEG gene with

Table 1 (*continued*)

Locus/noncoding or coding variant	Chr	Gene(s)	Study note
			cardiomyopathy in COVID-19 patients
Super-variant chr7_23	7	TOMM7	Possible relation of SNPs within chr7_23 to mitochondrial dysfunctions
Super-variant chr10_57	10	PCDH15	Location of all the 11 SNPs of chr10_57 in the intron of PCDH15 gene
An integrative multiomics analysis, including 9986 patients and 1,877,672 controls (Wu et al., 2021)			
3p21.31	3	XCR1, CCR2, and SACM1L	Inferring causality for COVID-19 severity
12q24.13/ rs10735079	12	OAS3	Significant association of SNPs in OAS3 gene with sever COVID-19
17q21.3	17	NSF and WNT3	
19q13.3	19	NAPSA	
21q22.1	21	IFNAR2	
Integrative genomic analyses, including 7885 patients and 961,804 controls (Pathak et al., 2021)			
3p21	3	SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6, CCR3, CCR2, CCR1, CCR5, CCRL2, LINC02009, XCR1, and GNL3	Identification of ABO and OAS1 by all three TWAS approaches Significant involvement of cytokine receptor signaling and JAK-STAT signaling pathways in immune response to COVID-19 Co-localizing of GNL3 eQTL with pulmonary function
6p21.1	6	FOXP4	
9q34.2	9	ABO and LCN1P1	
12q24.13	12	OAS1 and OAS3	
16p12.3	16	XYLT1 and DNAH3	
17q21.31	17	WNT3	
19p13	19	DPP9, PGLS, and KEAP1	
21q22.1	21	IFNAR2 and IL10RB	
10q23.31	10	IFIT3	
A sequencing study, including 284 patients and 966 controls (Wang et al., 2020a)			
rs143359233	12	GOLGA3	
rs11391519	9	DPP7	Possible effect on asymptomatic disease demonstration Association with low expression of DPP7 in multiple tissues
rs12329760 (V197M)	21	TMPRSS2	Association of V197M with sever COVID-19
6p22.1	6	HLA-A	Significant association of HLA-A11:01, B51:01, and C14:02 alleles with sever COVID-19
6p21.33	6	HLA-B and HLA-C	
rs6020298	20	Intron of a read-through transcript TMEM189–UBE2V1	Significant association of rs6020298 with mild and sever COVID-19
rs657152	9	ABO	Confirmation of blood group A as a risk factor
A sequencing study, including 131 patients and 258 controls (Benetti et al., 2020)			
rs41303171 (N720D) rs4646116 (K26R) rs148771870 (G211R) rs775181355 (V506A) rs767462182 (G377E)	X	ACE2	Significant higher allelic variability in controls compared with patients Destabilizing effect of V506A, V209G, and G377E on S protein and ACE2 interaction

Table 1 (*continued*)

Locus/noncoding or coding variant	Chr	Gene(s)	Study note
rs1327824339 (V209I)			Possible influence of N720D, K26R, and G211R on the interaction between S protein and ACE2
A sequencing study, including 659 life-threatening COVID-19 patients and 534 asymptomatic infected subjects (Zhang et al., 2020)			
S339fs rs121434431 (P554S) rs769624456 (W769*) M870V	4	TLR3	Existence of inborn errors of TLR3- and IRF7-dependent type I IFN immunity at eight loci in 3.5% of patients at various ages and ancestries
P364fs M371V D117N R7fs Q185* P246fs R369Q F95S	11	IRF7	
E96*	11	UNC93B1	
T42I rs201782115 (S60C) Q392K	19	TICAM1	
F24S rs1284582102 (R308*)	12	TBK1	
E49delrs199550479 (N146K)	19	IRF3	
rs181939581 (W73C) rs746291558 (S422R) P335del	21	IFNAR1	
E140fs	21	IFNAR2	
A sequencing study, including Four young (age <35 years) men with sever COVID-19 from 2 unrelated families (van der Made et al., 2020)			
Q710fs rs200553089 (V795F)	X	TLR7	Significant association of rare putative loss-of-function variants of TLR7 with impaired type I and II IFN responses in sever COVID-19
A sequencing study, including 82 patients (Wang et al., 2020b)			
6p22.1	6	HLA-A	Association of HLA-C07:29 and HLA-B15:27 with susceptibility to COVID-19
6p21.33	6	HLA-B	
A sequencing study, including 204 patients and 536 controls (Gómez et al., 2020)			
rs4646994 (I/D polymorphism)	17	ACE	Significant association of D allele with sever COVID-19
A sequencing study, including 25 patients with thromboembolic complications and 43 patients without thromboembolic complications (Calabrese et al., 2020)			
rs1799752 (I/D polymorphism)	17	ACE	Significant association of D/D homozygous polymorphism with thrombo-inflammatory in COVID-19 patients

replication by degradation of the RNA genome of coronavirus (Klaassen et al., 2020). On the other hand, loss-of-function variants in TLR7 caused a delay in IFN responses induced by TLRs and have been reported in young severe COVID-19 patients (van der Made et al., 2020). TLR7 has a known function of inducing innate immunity and endosomal innate immune sensor or receptor for single stranded RNA and subsequently triggers antiviral responses, including IFN production (Karaderi et al., 2020).

5. Host susceptibility variants and genes suggested by bioinformatics and structure-basis analyses

In addition to GWAS, directly evaluating deleterious missense variants in flexible regions of the established human proteins (i.e. ACE2 and TMPRSS2) may also identify putative pathogenic variants, which affect its function and structure, and consequently alter viral recognition. Extensive bioinformatics analysis strategies have been conducted including i) characterization of missense and/or disruptive variants (i.e., with stop gain/loss, splice-site, and frameshift/non-frameshift) in the established proteins from public coding variant databases such as the gnomAD; ii) functional prediction using multiple bioinformatics tools, including PolyPhen2 (Adzhubei et al., 2013), Sorting Intolerant From Tolerant (SIFT) (Ng and Henikoff, 2002), ActiveDriver (Reimand and Bader, 2013), MutationAssessor (Reva et al., 2011), and Combined Annotation Dependent Depletion (CADD) (Rentzsch et al., 2019); and iii) structure-based analysis (i.e., the LiAn program) to provide evidence that missense variants could potentially disrupt the protein structures and/or affect binding efficiencies to SARS-CoV-2 proteins. Using these strategies, previous studies, including our own, aimed to identify possible susceptibility variants in the established human proteins through bioinformatics and structure-basis analysis (Benetti et al., 2020; Guo et al., 2020; Hou et al., 2020). We have characterized potentially susceptible coding variants in ACE2, TMPRSS2, and other established genes reported by previous studies (Table 2).

5.1. ACE2 studies

According to bioinformatics analyses, some residues of ACE2 map to its interface site with S protein, such as 19, 35, 38, 206, 211, 219, 341, 353, 378, 468, and 547 residues (Benetti et al., 2020; Lippi et al., 2020). Missense variants in other residues may also influence binding affinity or the cleavage site of ACE2 and subsequently affect SARS-CoV-2 internalization (Lippi et al., 2020). It has been shown that some ACE2 variants, namely L351V and P389H, directly impact the binding affinities of S protein and ACE2 (Benetti et al., 2020). Variations located at the proteolytic cleavage site of ACE2 can influence virus infection through creating soluble ACE2 as a decoy receptor for the virus (Benetti et al., 2020). According to our analysis, ACE2 variants have population specific frequencies. For instance, L731F and S19P have higher frequencies in African populations and I468V is only present in Asian populations (Guo et al., 2020). Notably, the predicted effect of certain variants was contradictory in some studies; for example, the H378R variant of ACE2 was reported as both reducing and increasing susceptibility to virus binging (Guo et al., 2020; Lippi et al., 2020). In addition to coding variants, non-coding variants in regulatory and noncoding regions such as intronic and promoter regions of ACE2 can play a role in inter-individual differences in expression levels of ACE2 (Lippi et al., 2020). For instance, in one study, 6 SNPs (rs4240157, rs6632680, rs1548474, rs4830965, rs1476524, and rs2048683) out of 61 evaluated in the intronic sequence of ACE2 showed significant association with the expression level of ACE2 and hospitalization in COVID-19 patients (Wooster et al., 2020). CC or CT genotype of rs2106809 showed correlation with up-regulation of soluble ACE2 in the serum; a possible suggested mechanism is translation modulation by microRNA effect on this variant of mRNA in endothelial cells (Ciaglia et al., 2020). SARS-CoV-2 infection leads to down-regulation of ACE2 (Ciulla, 2020). Evidence of the relation between up-regulation or down-regulation of ACE2 and susceptibility to COVID-19 is scarce, and there are inconsistent reports; nevertheless, it is clear that high levels of circulating ACE2 or soluble ACE2 in healthy individuals can act as the decoy receptor for SARS-CoV-2 and decrease coronavirus infection. It is also worth mentioning that in some conditions, for instance type 2 diabetes, up-regulation of ACE2 disturbs the renin-angiotensin-aldosterone system (RAAS) pathway which in combination with the underlying condition can result in enhanced immune response and tissue injury effects seen in life-

Table 2
Predicted genetic variants involved in COVID-19 pathogenesis using in silico studies.

Risk variant	Amino acid change	Evaluated gene	Data source	Ref.
rs73635825	S19P	**ACE2**	**GnomAD (N = 141456)**	(Guo et al., 2020)
rs142984500	H378R			
rs148771870	G211R			
rs142443432	D206G			
rs373025684	S547C			
rs138390800	K341R			
rs372272603	R219C			
rs759590772	R219H			
rs191860450	I468V			
rs147311723	L731F			
rs751603885	R697G			
rs149039346	S692P			
rs146676783	E37K	**ACE2**	**GnomAD (N = 141456)**	(Gibson et al., 2020)
rs781255386	T27A			
rs143936283	K329G			
rs1299103394	K26E			
rs41303171	N720D			
rs1447927937	S43R			
rs759579097	G326E			
rs766996587	M82I			
rs4646116	K26R			
rs781255386	T27A	**ACE2**	**GnomAD (N = 141456)**	(MacGowan and Barton, 2020)
rs759579097	G326E			
rs146676783	E37K			
rs370610075	G352V			
rs961360700	D355N			
rs73635825	S19P	**ACE2**	**GnomAD (N >290000)**	(Stawiski et al., 2020)
rs778030746	I21V			
rs756231991	E23K			
rs4646116	K26R			
rs781255386	T27A			
rs1199100713	N64K			
rs763395248	T92I			
rs1395878099	Q102P			
rs142984500	H378R			
rs758278442	K31R			
rs1348114695	E35K			
rs146676783	E37K			
rs1192192618	Y50F			
rs1569243690	N51S			
rs1325542104	M62V			
rs755691167	K68E			
rs1256007252	F72V			
rs759134032	Y83H			
rs759579097	G326E			
rs370610075	G352V			
rs961360700	D355N			
rs751572714	Q388L			
	N33I			
	H34R			
	D38V			
	D509Y			
rs143936283	E329G	**ACE2**	**GnomAD (N = Not Available)**	(Hussain et al., 2020)
rs73635825	S19P			
rs1352194082	R514G	**ACE2**	**GnomAD (N = ~ 81000)**	(Hou et al., 2020)
rs12329760	V160M	**TMPRSS2**		
rs867186402	D435Y			
rs12329760	V160M	**TMPRSS2**	**Data from 520 healthy individuals and Data from non-Indian ethnicities (N = Not Available)**	(Vishnubhotla et al., 2020)
rs75603675	G8V	**TMPRSS2**	**131 patients and gnomAD (N = Not Available)**	**Data source: (** Latini et al., 2020)
rs114363287	G111R			
rs769208985	R298Q	**Furin**		
rs1236237792	I636V			
rs35074065	..	**TMPRSS2**	**GnomAD (N = 141456)**	(Bhattacharyya et al., 2020)
rs12329760	V160M			
HLA-B46:01	..	**HLA**		

(continued on next page)

Table 2 (*continued*)

Risk variant	Amino acid change	Evaluted gene	Data source	Ref.
HLA-B15:03	..		**Allele Frequency Net Database (N = 3382)**	(Nguyen et al., 2020)
rs117888248	..	**DPP4**	**Neandertal genomes (N = Not Available)**	(Zeberg and Pääbo, 2020)
rs201551785	G146S	**Furin**	**143 Serbian individuals and 1000 Genomes project (N = Not Available)**	(Klaassen et al., 2020)
rs4252187	R261H	**PLG**		
rs4252128	A494V			
rs148440491	N54K	*Trypsin-1*		
rs5030737	R52C	**MBL2**		
rs1800450	G54D			
rs1800451	G57E			
rs751350524	R49Q	**OAS1**		
rs753837415	I99V			
rs1021340095	R130H			

threatening COVID-19 (Banu et al., 2020). On the other hand, it is speculated that down-regulation of ACE2 increases pathogenicity of SARS-CoV-2 because of higher amounts of angiotensin II, the main substrate of ACE2. This, in turn, can cause progression of inflammatory reaction and hyper-coagulation processes and subsequent worse prognosis of COVID-19 (Babadaei et al., 2020; Verdecchia et al., 2020).

5.2. TMPRSS2 studies

Molecular docking studies suggested that K835, Y837, D839, C840, L841, D843, I844, R847, R848 in cup-like architecture of TMPRSS2 are the key interacting residues with S2 domain of S protein (Vishnubhotla et al., 2020). Both coding and noncoding variants of TMPRSS2 could influence COVID-19 severity (Table 2). A single nucleotide deletion (delC) at a variant site (rs35074065) in a *cis*-eQTL of TMPRSS2 is common in Europeans and North Americans which leads to up-regulation of TMPRSS2 and cleavage of a new site in 614G (D614G mutation in S protein) subtype of SARS-CoV-2 (Bhattacharyya et al., 2020). Interestingly, the delC allele results in overexpression of MX1 gene in lung tissue, a nearby gene of TMPRSS2; upregulated MX1 causes neutrophil infiltration, which is one of the key players in innate immunity relevant to viral infection (Bhattacharyya et al., 2020). One suggested reason for higher lethality of COVID-19 in Italy was the lower frequency of deleterious missense mutations in TMPRSS2 when compared to other European countries (Asselta et al., 2020). By in silico approaches, it was confirmed that the V160M mutation decreases S protein priming and viral infection which has lower frequency in Italian populations compared to others (Vishnubhotla et al., 2020).

Although these studies have characterized the putative functional and structurally related variants in ACE2 and TMPRSS2 with top allele frequencies in various populations, studies to date have had limited ability to identify susceptibility variants associated with the phenotype of COVID-19. However, these findings, together with other genetic studies, can prioritize the promising variants in these established genes for further fast-track genotyping in blood samples from COVID-19 patients, which could provide a valuable opportunity to identify susceptibility variants related to symptoms of COVID-19 patients and/or markers of severity of disease. In particular, accumulating evidence suggests that both ACE2 and TMPRSS2 could be potential biomarkers for COVID-19 severity (AlGhatrif et al., 2020; Gue et al., 2020; Strope and Chau, 2020).

6. Concluding remarks and future prospective

The pathophysiology and severity of COVID-19 illness varies among patients and depends in part on underlying risk factors and chronic diseases. Remarkable progress from protein structural and multi-omics

data analyses have characterized human proteins with evidence of interaction with SARS-CoV-2, providing novel insights into underlying SARS-CoV-2 biology and infectivity. Current genetic studies using GWAS and TWAS designs have successfully identified susceptibility genetic variants for COVID-19. WES and WGS designs in COVID-19 patients have also identified putative pathogenic variants, especially in established ACE2, TMPRSS2, 40S ribosomal subunit, and immunity and antiviral signaling pathways, for host susceptibility to COVID-19 and subsequent occurrence of clinical symptoms. Assembling a large study population with well-phenotyped clinical data and banked DNA for etiologic and prognostic research on COVID-19 is urgently needed to comprehensively characterize large effects of pathogenic variants and susceptibility genes with significant clinical implications for disease prevention and control, treatment, and intensive care of severe COVID-19.

Remarkable successes in Moderna and Pfizer vaccines have greatly reduced the likelihood of SARS-CoV-2 infection and, to an even greater extent, the risk of severe COVID-19. As both vaccines are designed to target the SARS-CoV-2 S protein, genetic findings may additionally improve vaccine effectiveness by elucidating potential biological mechanisms of S protein binding efficiency and by identifying populations particularly vulnerable to SARS-CoV-2 infection due to ACE2 or other genetic predisposition. On the other hand, vaccine effectiveness may fail for populations with predisposition to immune deficiency, as demonstrated by a substantial proportion of severe COVID-19 cases likely carrying loss-of-function variants in known loci that impaired type I and II IFN responses. The identification of such pathogenic variants could play an important role in guiding current vaccine development and in the discovery of therapeutic targets needed even after the introduction of successful vaccines. Moreover, new emerging variants of SARS-CoV-2 can have a dramatic effect on viral-host protein interactions, behavior of virus infection and transmission, and short- and long-term outcomes of COVID-19; therefore, the currently authorized mRNA vaccines may need to be updated against those new versions of SARS-CoV-2. Furthermore, a significant proportion of patients have developed post-acute sequelae of SARS-CoV-2 infection (PASC), which can affect multiple organ systems and range from severe sequelae such as stroke, seizures, and anxiety/depression to symptom-based conditions, such as shortness of breath, fatigue, loss of taste or smell, and cognitive impairment. PASC have been documented as long as nine months after illness among those with moderate to severe COVID-19 illness requiring hospitalization but also among those who suffered mild COVID-19 illness. Thus, the ability to identify susceptibility genes related to PASC could provide new insight into biological mechanisms underlying the development of adverse long-term consequences of SARS-CoV-2 infection and could have also high translational potential for the identification of vulnerable populations for PASC surveillance and prevention.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2021.145963.

References

Ackermann, M., Verleden, S.E., Kuehnel, M., Haverich, A., Welte, T., Laenger, F., Vanstapel, A., Werlein, C., Stark, H., Tzankov, A., Li, W.W., 2020. Pulmonary vascular endothelialitis, thrombosis, and angiogenesis in Covid-19. New England Journal of Medicine 383 (2), 120–128.

Adzhubei, I., Jordan, D.M., Sunyaev, S.R., 2013. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet. Chapter 7 (Unit7), 20.

AlGhatrif, M., Cingolani, O., Lakatta, E.G., 2020. The dilemma of coronavirus disease 2019, aging, and cardiovascular disease: insights from cardiovascular aging science. JAMA cardiology 5 (7), 747–748.

Asselta, R., Paraboschi, E.M., Mantovani, A., Duga, S., 2020. ACE2 and TMPRSS2 variants and expression as candidates to sex and country differences in COVID-19 severity in Italy. Aging (Albany NY) 12 (11), 10087.

Atkin-Smith, G.K., Duan, M., Chen, W., Poon, I.K., 2018. The induction and consequences of Influenza A virus-induced cell death. Cell death & disease 9 (10), 1–11.

Babadaei, M.M.N., Hasan, A., Bloukh, S.H., Edis, Z., Sharifi, M., Kachooei, E., Falahati, M., 2020. The expression level of angiotensin-converting enzyme 2 determines the severity of COVID-19: Lung and heart tissue as targets. Journal of Biomolecular Structure and Dynamics 1–7.

Banu, N., Panikar, S.S., Leal, L.R., Leal, A.R., 2020. Protective role of ACE2 and its downregulation in SARS-CoV-2 infection leading to macrophage activation syndrome: therapeutic implications. Life sciences 256, 117905.

Benetti, E., Tita, R., Spiga, O., Ciolfi, A., Birolo, G., Bruselles, A., Doddato, G., Giliberti, A., Marconi, C., Musacchia, F., Pippucci, T., 2020. ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population. European Journal of Human Genetics 28 (11), 1602–1614.

Bhattacharyya, C., Das, C., Ghosh, A., Singh, A.K., Mukherjee, S., Majumder, P.P., Basu, A., Biswas, N.K., 2020. Global spread of SARS-CoV-2 subtype with spike protein mutation D614G is shaped by human genomic variations that regulate expression of TMPRSS2 and MX1 genes. BioRxiv.

Bouhaddou, M., Memon, D., Meyer, B., White, K.M., Rezelj, V.V., Marrero, M.C., Polacco, B.J., Melnyk, J.E., Ulferts, S., Kaake, R.M., Batra, J., 2020. The global phosphorylation landscape of SARS-CoV-2 infection. Cell 182 (3), 685–712.

Calabrese, C., Annunziata, A., Coppola, A., Pafundi, P.C., Guarino, S., Di Spirito, V., Maddaloni, V., Pepe, N., Fiorentino, G., 2020. ACE gene I/D polymorphism and acute pulmonary embolism in Covid 19 pneumonia: a potential predisposing role. Frontiers in medicine 7, 1136.

Chai, J., Cai, Y., Pang, C., Wang, L., McSweeney, S., Shanklin, J., Liu, Q., 2021. Structural basis for SARS-CoV-2 envelope protein recognition of human cell junction protein PALS1. Nature Communications 12 (1), 1–6.

Cheng, Z., Zhou, J., To, K.K.W., Chu, H., Li, C., Wang, D., Yang, D., Zheng, S., Hao, K., Bossé, Y., Obeidat, M.E., 2015. Identification of TMPRSS2 as a susceptibility gene for severe 2009 pandemic A (H1N1) influenza and A (H7N9) influenza. The Journal of infectious diseases 212 (8), 1214–1221.

Ciaglia, E., Vecchione, C., Puca, A.A., 2020. COVID-19 infection and circulating ACE2 levels: protective role in women and children. Frontiers in pediatrics 8, 206.

Ciulla, M.M., 2020. SARS-CoV-2 downregulation of ACE2 and pleiotropic effects of ACEIs/ARBs. Hypertension Research 43 (9), 985–986.

Davies, J.P., Almasy, K.M., McDonald, E.F., Plate, L., 2020. Comparative multiplexed interactomics of SARS-CoV-2 and homologous coronavirus nonstructural proteins identifies unique and shared host-cell dependencies. ACS infectious diseases 6 (12), 3174–3189.

Docherty, A.B., Harrison, E.M., Green, C.A., Hardwick, H.E., Pius, R., Norman, L., Holden, K.A., Read, J.M., Dondelinger, F., Carson, G., Merson, L., 2020. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. bmj 369.

Ellinghaus, D., Degenhardt, F., Bujanda, L., Buti, M., Albillos, A., Invernizzi, P., Fernandez, J., Prati, D., Baselli, G., Asselta, R., Grimsrud, M.M., 2020. Genomewide association study of severe covid-19 with respiratory failure. N Engl J Med 383, 1522–1534.

Eng, V.V., Wemyss, M.A., Pearson, J.S., 2020. August. The diverse roles of RIP kinases in host-pathogen interactions. Academic Press.

Gallagher, T.M., Buchmeier, M.J., 2001. Coronavirus spike proteins in viral entry and pathogenesis. Virology 279 (2), 371–374.

Gibson, W.T., Evans, D.M., An, J. and Jones, S.J. 2020. ACE 2 coding variants: a potential X-linked risk factor for COVID-19 disease. bioRxiv.

Gómez, J., Albaiceta, G.M., García-Clemente, M., López-Larrea, C., Amado-Rodríguez, L., Lopez-Alonso, I., Hermida, T., Enriquez, A.I., Herrero, P., Melón, S., Alvarez-Argüelles, M.E., 2020. Angiotensin-converting enzymes (ACE, ACE2) gene variants and COVID-19 outcome. Gene 762, 145102.

Gordon, D.E., Gordon, D.E., Hiatt, J., Bouhaddou, M., Rezelj, V.V., Ulferts, S., Braberg, H., Jureka, A.S., Obernier, K., Guo, J.Z., Batra, J., Kaake, R.M., 2020a. Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. Science 370 (6521).

Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O'Meara, M. J., Rezelj, V.V., Guo, J.Z., Swaney, D.L., Tummino, T.A., 2020b. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature 583 (7816), 459–468.

Gue, Y.X., Kanji, R., Markides, V., Gorog, D.A., 2020. Angiotensin converting enzyme 2 may mediate disease severity in COVID-19. American Journal of Cardiology 130, 161–162.

Guo, X., Chen, Z., Xia, Y., Lin, W., Li, H., 2020. Investigation of the genetic variation in ACE2 on the structural recognition by the novel coronavirus (SARS-CoV-2). Journal of Translational Medicine 18 (1), 1–9.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A. and Müller, M.A. 2020. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell. 181(2). 271-280.

Hou, Y., Zhao, J., Martin, W., Kallianpur, A., Chung, M.K., Jehi, L., Sharifi, N., Erzurum, S., Eng, C., Cheng, F., 2020. New insights into genetic susceptibility of COVID-19: an ACE2 and TMPRSS2 polymorphism analysis. BMC medicine 18 (1), 1–8.

Hu, J., Li, C., Wang, S., Li, T., Zhang, H., 2021. Genetic variants are identified to increase risk of COVID-19 related mortality from UK Biobank data. Human genomics 15 (1), 1–10.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan. China. The lancet 395 (10223), 497–506.

Hussain, M., Jabeen, N., Raza, F., Shabbir, S., Baig, A.A., Amanullah, A., Aziz, B., 2020. Structural variations in human ACE2 may influence its binding with SARS-CoV-2 spike protein. Journal of medical virology 92 (9), 1580–1586.

Jevnikar, K., Jaki Mekjavic, P., Vidovic Valentincic, N., Petrovski, G., Globocnik Petrovic, M., 2021. An Update on COVID-19 Related Ophthalmic Manifestations. Ocular Immunology and Inflammation 1–6.

Jiang, H.W., Zhang, H.N., Meng, Q.F., Xie, J., Li, Y., Chen, H., Zheng, Y.X., Wang, X.N., Qi, H., Zhang, J., Wang, P.H., 2020. SARS-CoV-2 Orf9b suppresses type I interferon responses by targeting TOM70. Cellular & molecular immunology 17 (9), 998–1000.

Karaderi, T., Bareke, H., Kunter, I., Seytanoglu, A., Cagnan, I., Balci, D., Barin, B., Hocaoglu, M.B., Rahmioglu, N., Asilmaz, E., Taneri, B., 2020. Host genetics at the intersection of autoimmunity and COVID-19: A potential key for heterogeneous COVID-19 severity. Frontiers in Immunology 11, 3314.

Kern, D.M., Sorum, B., Mali, S.S., Hoel, C.M., Sridharan, S., Remis, J.P., Toso, D.B., Kotecha, A., Bautista, D.M., Brohawn, S.G., 2020. Cryo-EM structure of the SARS-CoV-2 3a ion channel in lipid nanodiscs. BioRxiv.

Klaassen, K., Stankovic, B., Zukic, B., Kotur, N., Gasic, V., Pavlovic, S., Stojiljkovic, M., 2020. Functional prediction and comparative population analysis of variants in genes for proteases and innate immunity related to SARS-CoV-2 infection. Infection, Genetics and Evolution 84, 104498.

Kumar, R., Verma, H., Singhvi, N., Sood, U., Gupta, V., Singh, M., Kumari, R., Hira, P., Nagar, S., Talwar, C., Nayyar, N., 2020. Comparative genomic analysis of rapidly evolving SARS-CoV-2 reveals mosaic pattern of phylogeographical distribution. Msystems 5 (4), e00505–20.

Latini, A., Agolini, E., Novelli, A., Borgiani, P., Giannini, R., Gravina, P., Smarrazzo, A., Dauri, M., Andreoni, M., Rogliani, P., Bernardini, S., 2020. COVID-19 and genetic variants of protein involved in the SARS-CoV-2 entry into the host cells. Genes 11 (9), 1010.

Levi, M., Thachil, J., Iba, T., Levy, J.H., 2020. Coagulation abnormalities and thrombosis in patients with COVID-19. The Lancet Haematology 7 (6), e438–e440.

Li, F., Li, W., Farzan, M., Harrison, S.C., 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science 309 (5742), 1864–1868.

Li, J., Guo, M., Tian, X., Wang, X., Yang, X., Wu, P., Liu, C., Xiao, Z., Qu, Y., Yin, Y., Wang, C., 2021a. Virus-host interactome and proteomic survey reveal potential virulence factors influencing SARS-CoV-2 pathogenesis. Med 2 (1), 99–112.

Li, J., Huang, D.Q., Zou, B., Yang, H., Hui, W.Z., Rui, F., Yee, N.T.S., Liu, C., Nerurkar, S. N., Kai, J.C.Y., Teng, M.L.P., 2021b. Epidemiology of COVID-19: A systematic review and meta-analysis of clinical characteristics, risk factors, and outcomes. Journal of medical virology 93 (3), 1449–1458.

Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greenough, T.C., Choe, H., 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 426 (6965), 450–454.

Liang, W., Guan, W., Chen, R., Wang, W., Li, J., Xu, K., Li, C., Ai, Q., Lu, W., Liang, H., Li, S., 2020. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. The lancet oncology 21 (3), 335–337.

Lippi, G., Lavie, C.J., Henry, B.M., Sanchis-Gomar, F., 2020. Do genetic polymorphisms in angiotensin converting enzyme 2 (ACE2) gene play a role in coronavirus disease 2019 (COVID-19)? Clinical Chemistry and Laboratory Medicine (CCLM) 58 (9), 1415–1422.

Lokugamage, K.G., Hage, A., de Vries, M., Valero-Jimenez, A.M., Schindewolf, C., Dittmann, M., Rajsbaum, R., Menachery, V.D., 2020. Type I interferon susceptibility distinguishes SARS-CoV-2 from SARS-CoV. Journal of virology 94 (23), e01410–20.

MacGowan, S., Barton, G.J., 2020. Missense variants in ACE2 are predicted to encourage and inhibit interaction with SARS-CoV-2 Spike and contribute to genetic risk in COVID-19. BioRxiv.

Mandala, V.S., McKay, M.J., Shcherbakov, A.A., Dregni, A.J., Kolocouris, A., Hong, M., 2020. Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. Nature structural & molecular biology 27 (12), 1202–1208.

Miao, G., Zhao, H., Li, Y., Ji, M., Chen, Y., Shi, Y., Bi, Y., Wang, P., Zhang, H., 2021. ORF3a of the COVID-19 virus SARS-CoV-2 blocks HOPS complex-mediated assembly of the SNARE complex required for autolysosome formation. Developmental cell 56 (4), 427–442.

Mu, J., Fang, Y., Yang, Q., Shu, T., Wang, A., Huang, M., Jin, L., Deng, F., Qiu, Y., Zhou, X., 2020. SARS-CoV-2 N protein antagonizes type I interferon signaling by suppressing phosphorylation and nuclear translocation of STAT1 and STAT2. Cell discovery 6 (1), 1–4.

Ng, P.C., Henikoff, S., 2002. Accounting for human polymorphisms predicted to affect protein function. Genome research 12 (3), 436–446.

Nguyen, A., David, J.K., Maden, S.K., Wood, M.A., Weeder, B.R., Nellore, A., Thompson, R.F., 2020. Human leukocyte antigen susceptibility map for severe acute respiratory syndrome coronavirus 2. Journal of virology 94 (13), e00510–20.

Nie, X., Qian, L., Sun, R., Huang, B., Dong, X., Xiao, Q., Zhang, Q., Lu, T., Yue, L., Chen, S., Li, X., 2021. Multi-organ proteomic landscape of COVID-19 autopsies. Cell 184 (3), 775–791.

Pairo-Castineira, E., Clohisey, S., Klaric, L., Bretherick, A.D., Rawlik, K., Pasko, D., Walker, S., Parkinson, N., Fourman, M.H., Russell, C.D., Furniss, J., 2021. Genetic mechanisms of critical illness in Covid-19. Nature 591 (7848), 92–98.

Pathak, G.A., Singh, K., Miller-Fleming, T.W., Wendt, F.R., Ehsan, N., Hou, K., Johnson, R., Lu, Z., Gopalan, S., Yengo, L., Mohammadi, P., 2021. Integrative genomic analyses identify susceptibility genes underlying COVID-19 hospitalization. Nature Communications. 12, 1–11.

Petrilli, C.M., Jones, S.A., Yang, J., Rajagopalan, H., O'Donnell, L., Chernyak, Y., Tobin, K.A., Cerfolio, R.J., Francois, F., Horwitz, L.I., 2020. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. Bmj 369.

Reimand, J., Bader, G.D., 2013. Systematic analysis of somatic mutations in phosphorylation signaling predicts novel cancer drivers. Molecular systems biology 9 (1), 637.

Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J., Kircher, M., 2019. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic acids research 47 (D1), D886–D894.

Reva, B., Antipin, Y. and Sander, C. 2011. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic acids research. 39(17). e118-e118.

Schönfelder, K., Breuckmann, K., Elsner, C., Dittmer, U., Fistera, D., Herbstreit, F., Risse, J., Schmidt, K., Sutharsan, S., Taube, C., Jöckel, K.H., 2021. Transmembrane serine protease 2 Polymorphisms and Susceptibility to Severe Acute Respiratory Syndrome Coronavirus Type 2 Infection: A German Case-Control Study. Frontiers in genetics 12.

Schubert, K., Karousis, E.D., Jomaa, A., Scaiola, A., Echeverria, B., Gurzeler, L.A., Leibundgut, M., Thiel, V., Mühlemann, O., Ban, N., 2020. SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. Nature structural & molecular biology 27 (10), 959–966.

Shang, Y., Smith, S., Hu, X., 2016. Role of Notch signaling in regulating innate immunity and inflammation in health and disease. Protein & cell 7 (3), 159–174.

Stawiski, E.W., Diwanji, D., Suryamohan, K., Gupta, R., Fellouse, F.A., Sathirapongsasuti, J.F., Liu, J., Jiang, Y.P., Ratan, A., Mis, M., Santhosh, D., 2020. Human ACE2 receptor polymorphisms predict SARS-CoV-2 susceptibility. BioRxiv.

Strope, J.D., Chau, C.H., 2020. TMPRSS2: Potential biomarker for COVID-19 outcomes. Journal of clinical pharmacology.

Suhaili, S.H., Karimian, H., Stellato, M., Lee, T.H., Aguilar, M.I., 2017. Mitochondrial outer membrane permeabilization: a focus on the role of mitochondrial membrane structural organization. Biophysical reviews 9 (4), 443–457.

Szeto, C., Chatzileontiadou, D.S., Nguyen, A.T., Sloane, H., Lobos, C.A., Jayasinghe, D., Halim, H., Smith, C., Riboldi-Tunnicliffe, A., Grant, E.J., Gras, S., 2021. The presentation of SARS-CoV-2 peptides by the common HLA-A* 02: 01 molecule. Iscience 24 (2), 102096.

Teoh, K.T., Siu, Y.L., Chan, W.L., Schlüter, M.A., Liu, C.J., Peiris, J.M., Bruzzone, R., Margolis, B., Nal, B., 2010. The SARS coronavirus E protein interacts with PALS1 and alters tight junction formation and epithelial morphogenesis. Molecular biology of the cell 21 (22), 3838–3852.

Tesei, A., Cortesi, M., Zamagni, A., Arienti, C., Pignatta, S., Zanoni, M., Paolillo, M., Curti, D., Rui, M., Rossi, D., Collina, S., 2018. Sigma receptors as endoplasmic reticulum stress "gatekeepers" and their modulators as emerging new weapons in the fight against cancer. Frontiers in pharmacology 9, 711.

Thoms, M., Buschauer, R., Ameismeier, M., Koepke, L., Denk, T., Hirschenberger, M., Kratzat, H., Hayn, M., Mackens-Kiani, T., Cheng, J., Straub, J.H., 2020. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. Science 369 (6508), 1249–1255.

Tomita, Y., Ikeda, T., Sato, R., Sakagami, T., 2020. Association between HLA gene polymorphisms and mortality of COVID-19: An in silico analysis. Immunity, inflammation and disease 8 (4), 684–694.

Tutuncuoglu, B., Cakir, M., Batra, J., Bouhaddou, M., Eckhardt, M., Gordon, D.E., Krogan, N.J., 2020. The landscape of human cancer proteins targeted by SARS-CoV-2. Cancer discovery 10 (7), 916–921.

van der Made, C.I., Van Der Made, C.I., Simons, A., Schuurs-Hoeijmakers, J., Van Den Heuvel, G., Mantere, T., Kersten, S., Van Deuren, R.C., Steehouwer, M., Van Reijmersdal, S.V., Jaeger, M., Hofste, T., 2020. Presence of genetic variants among young men with severe COVID-19. Jama 324 (7), 663–673.

Varga, Z., Flammer, A.J., Steiger, P., Haberecker, M., Andermatt, R., Zinkernagel, A.S., Mehra, M.R., Schuepbach, R.A., Ruschitzka, F., Moch, H., 2020. Endothelial cell infection and endotheliitis in COVID-19. The Lancet 395 (10234), 1417–1418.

Verdecchia, P., Cavallini, C., Spanevello, A., Angeli, F., 2020. The pivotal link between ACE2 deficiency and SARS-CoV-2 infection. European journal of internal medicine 76, 14–20.

Vishnubhotla, R., Vankadari, N., Ketavarapu, V., Amanchy, R., Avanthi, S., Bale, G., Reddy, D.N., Sasikala, M., 2020. Genetic variants in TMPRSS2 and Structure of SARS-CoV-2 spike glycoprotein and TMPRSS2 complex. BioRxiv.

Wang, F., Huang, S., Gao, R., Zhou, Y., Lai, C., Li, Z., Xian, W., Qian, X., Li, Z., Huang, Y., Tang, Q., 2020a. Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility. Cell discovery 6 (1), 1–16.

Wang, N., Wu, R., Tang, D., Kang, R., 2021. The BET family in immunity and disease. Signal transduction and targeted therapy 6 (1), 1–22.

Wang, W., Zhang, W., Zhang, J., He, J., Zhu, F., 2020b. Distribution of HLA allele frequencies in 82 Chinese individuals with coronavirus disease-2019 (COVID-19). Hla 96 (2), 194–196.

Wiersinga, W.J., Rhodes, A., Cheng, A.C., Peacock, S.J., Prescott, H.C., 2020. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): A review. Jama 324 (8), 782–793.

Wooster, L., Nicholson, C.J., Sigurslid, H.H., Cardenas, C.L.L. and Malhotra, R., 2020. Polymorphisms in the ACE2 locus associate with severity of COVID-19 infection. medRxiv.

Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., Meng, J., Zhu, Z., Zhang, Z., Wang, J., Sheng, J., 2020. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. Cell host & microbe 27 (3), 325–328.

Wu, L., Zhu, J., Liu, D., Sun, Y., Wu, C., 2021. An integrative multiomics analysis identifies putative causal genes for COVID-19 severity. Genetics in Medicine 1–11.

Yan, H., Krishnan, K., Lim, J.T., Contillo, L.G., Krolewski, J.J., 1996. Molecular characterization of an alpha interferon receptor 1 subunit (IFNaR1) domain required for TYK2 binding and signal transduction. Molecular and Cellular Biology 16 (5), 2074–2082.

Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., Zhou, Q., 2020. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science 367 (6485), 1444–1448.

Yue, Y., Nabar, N.R., Shi, C.S., Kamenyeva, O., Xiao, X., Hwang, I.Y., Wang, M., Kehrl, J.H., 2018. SARS-Coronavirus Open Reading Frame-3a drives multimodal necrotic cell death. Cell death & disease 9 (9), 1–15.

Zeberg, H. and Pääbo, S. 2020. The MERS-CoV receptor gene is among COVID-19 risk factors inherited from Neandertals. bioRxiv.

Zhang, Q., Bastard, P., Liu, Z., Le Pen, J., Moncada-Velez, M., Chen, J., Ogishi, M., Sabli, I.K., Hodeib, S., Korol, C., Rosain, J., 2020. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. Science 370 (6515).

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R. and Niu, P. 2020. A novel coronavirus from patients with pneumonia in China, 2019. New England journal of medicine.