REVIEW

COVID-19: impact on Public Health and hypothesis-driven investigations on genetic susceptibility and severity

Susana David[1,2] · Guillermo Dorado[3] · Elsa L. Duarte[4] · Stephanie David-Bosne[5] · João Trigueiro-Louro[6,7,8] · Helena Rebelo-de-Andrade[6,7]

Abstract
COVID-19 is a new complex multisystem disease caused by the novel coronavirus SARS-CoV-2. In slightly over 2 years, it infected nearly 500 million and killed 6 million human beings worldwide, causing an unprecedented coronavirus pandemic. Currently, the international scientific community is engaged in elucidating the molecular mechanisms of the pathophysiology of SARS-CoV-2 infection as a basis of scientific developments for the future control of COVID-19. Global exome and genome analysis efforts work to define the human genetics of protective immunity to SARS-CoV-2 infection. Here, we review the current knowledge regarding the SARS-CoV-2 infection, the implications of COVID-19 to Public Health and discuss genotype to phenotype association approaches that could be exploited through the selection of candidate genes to identify the genetic determinants of severe COVID-19.

Keywords COVID-19 · Public Health · Genetic susceptibility to infection · Genetic determinants of severe disease · Candidate gene association studies (CGAS) · Genome-wide association studies (GWAS)

Introduction

COVID-19 is a new, complex, multisystem disease caused by the novel coronavirus SARS-CoV-2. In slightly over 2 years, it infected nearly 500 million and killed 6 million

✉ Susana David
david.suzana@gmail.com

[1] Departamento de Genética Humana, Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA,IP), Lisboa, Portugal

[2] Instituto de Investigação do Medicamento (iMed.ULisboa), Faculdade de Farmácia, Universidade de Lisboa, Lisboa, Portugal

[3] Atlántida Centro de Investigación y Desarrollo de Estudios Profesionales (CIDEP), Granada, Spain

[4] MED-Instituto Mediterrâneo para a Agricultura, Ambiente e Desenvolvimento, Escola de Ciências e Tecnologia, Universidade de Évora, Évora, Portugal

[5] Université Claude Bernard Lyon 1, Lyon, France

[6] Departamento de Doenças Infeciosas, INSA, IP, Lisboa, Portugal

[7] Host-Pathogen Interaction Unit, Instituto de Investigação do Medicamento (iMed.ULisboa), Faculdade de Farmácia, Universidade de Lisboa, Lisboa, Portugal

[8] Hospital Egas Moniz, Centro Hospitalar Lisboa Ocidental, Lisboa, Portugal

human beings worldwide, causing an unprecedented coronavirus pandemic. Several vaccines, developed in record time, were approved for public use (Forni and Mantovani 2021). However, their real-world effectiveness in controlling the COVID-19 pandemic is still under debate. Consequently, other scientific approaches to control the disease are also needed. The wide spectrum of interindividual variability in the clinical manifestation of COVID-19, ranging from asymptomatic infection to severe disease and death, has gone unexplained by mutations in SARS-CoV-2. Therefore, notwithstanding underlying comorbidities and risk factors, host genetics appears to be a possible source and major contributor of these observed differences in the host–pathogen relationship. The identification of genetic determinants of COVID-19 susceptibility and severity could represent a valuable tool in the control of COVID-19.

In this review, we begin by summarizing recent investigations regarding the epidemiology and Public Health of COVID-19. We then describe the biology and pathogenesis of the SARS-CoV-2 and the immune response to this virus in humans. Finally, we review sources available for candidate gene selection and hypothesis driven investigations on the genetic susceptibility and severity of COVID-19.

An overview of the Public Health impact of SARS-COV-2

Coronaviruses (CoVs) constitute a worldwide threat to Public Health. CoVs may cause respiratory, enteric, hematological, and neurological disorders, to name a few. They are able to infect many animal species, also ranging from asymptomatic to severe clinical forms. There is great concern to the fact that the newly emerging CoVs have demonstrated the capacity to jump across species barriers with animal-to-human transmission, also acquiring the ability of efficient human-to-human transmission (Li 2016; Phan et al. 2018; Chen et al. 2020).

The previous CoV epidemic outbreaks include the 2002/2003 outbreak of Severe Acute Respiratory Syndrome (SARS)-CoV (also reported as SARS-CoV-1). Like SARS-CoV-2, it also originated in China (Guangdong). (SARS)-CoV spread to nearly 30 countries and infected more than 8000 people, causing about 1000 deaths, with an 11% fatality rate (World Health Organization, WHO 2006). The 2012 outbreak of Middle East Respiratory Syndrome (MERS)-CoV, from Saudi Arabia, also originated cases in nearly 30 other countries and was responsible for over 2000 cases and nearly 1000 deaths, with an approximate 35% fatality rate (Donnelly et al. 2019). The (SARS)-CoV-2 has surpassed by several orders of magnitude these earlier CoV outbreaks in the number of cases and deaths. Compared to the previous outbreaks (more confined in time and space), the SARS-CoV-2 outbreak, that started in 2019 in Wuhan, China, has plundered global health as well as the economic and social sectors, in an unprecedented CoV pandemic.

The spread of the SARS-CoV-2 to all continents occurred at a fast pace. On December 31, 2019, the health authorities of Wuhan reported 27 people diagnosed with a pneumonia of unknown origin. In the following days, reports of cases appeared in Thailand, Japan, and the USA, all of them from people who had recently traveled to Wuhan or had been in contact with an infected person (supporting the ability of sustained human-to-human transmission) (Park 2020). On March 11, 2020, in view of its rapid expansion and high fatality rate, the World Health Organization (WHO) announced the SARS-CoV-2 infection as a pandemic (WHO Mar 11, 2020). In less than 1 year, the virus had already infected over 50 million individuals and resulted in the loss of more than one million lives (Dong et al. 2020; WHO November 17, 2020).

It is worth highlighting that the SARS-CoV-2 viral genome was isolated and sequenced at an unequalled pace and that the results became readily available within 3 days after the first coronavirus cases were reported (Genebank accession MN908947) (Wu et al. 2020a). Three months later, GeneBank had more than five hundred genomic sequences of the virus from samples collected in different parts of the world.

COVID-19 has negatively affected several aspects of Public Health. The diversion of the healthcare workforce to the mitigation of SARS-CoV-2, indirectly led to the neglect of measures regarding both communicable and non-communicable diseases. This has been the case for humanity's main infectious disease threats: tuberculosis (TB); human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS); and malaria. Ending these epidemics by 2030 is among the health targets of the 2030 Sustainable Development Goals. However, in the wake of the SARS-CoV-2 pandemic, the milestones are at risk of a substantial regression. Studies modeling the current situation, suggest a potential setback of five to eight years in the fight against the incidence and deaths from TB, 20 years in the fight against mortality due to malaria and about 12 years in AIDS-related deaths (Cilloni et al. 2020; Jewell et al. 2020; Stop TB Partnership 2020; WHO 2020). These top killer infectious diseases concentrate in current high burden countries mainly in Asia and sub-Saharan Africa, whereas the SARS-CoV-2 has taken its toll on high- and low-income countries alike. Perhaps surprisingly, some high-income economies have been among the hardest hit.

For the SARS-CoV-2, the secondary attack rate was estimated to exceed 10% (Burke 2020; Ghinai et al. 2020). A systematic review and meta-analysis reported a household secondary attack rate for the SARS-CoV-2 of 18%, higher than the upper range of estimates for the 2009 H1N1 pandemic influenza (5–15%) and also higher than that observed for both SARS (5–10%) and MERS (4–5%) (Koh et al. 2020). Moreover, there is the possibility of superspreading events (SSEs), with reports of over 50% and 80% attack rates among confirmed and all cases, respectively (Hamner 2020). Another key aspect is that all the observed forms of COVID-19's clinical presentation, from asymptomatic to symptomatic; appear to be potentially infectious (Bai et al. 2020; Chau et al. 2020; Hoehl et al. 2020; Yu et al. 2020).

The relationship between viral load and infectivity or disease severity remains unclear (Walsh et al. 2020). Increased viral load of index cases appears as a leading driver of SARS-CoV-2 transmission (Marks et al. 2021). However, whereas in some studies high viral load appears to be associated with immune factors, disease severity and mortality (Fajnzylber et al. 2020; Pujadas et al. 2020; Shenoy et al. 2021), others have reported conflicting results, namely in children and asymptomatic patients (Hasanoglu et al. 2021; Aykac et al. 2021). Strain-specific differences have also been reported (Luo et al. 2021; He et al. 2021). Currently, no standard measure of viral load in clinical samples is available (Dahdouh et al. 2021); therefore, the relationship between viral load and infectivity or disease severity needs to be further investigated. Moreover, evidence from the literature does not directly suggest that severity would be only mediated by viral load but also by age and the impact that age has on the immune system (Chou et al. 2022).

"Host-jumps" in CoV appear to result in short-term spill-over infections, with little evidence for sustained onward transmission in the new host species (Leopardi et al. 2018). However, this was not the case in the recent SARS-CoV-2 outbreak in minks, where sustained then back to human infection was observed (ECDC 2020; WHO January 20, 2021). In the earlier CoV outbreaks of SARS and MERS, the early SSEs gave way to sustained transmission in the later stages of the epidemics. With the possibility of SSEs and both symptomatic and asymptomatic transmission of the virus, the 1-year fatality rate from the SARS-CoV-2 pandemic has surpassed the levels observed for malaria, HIV/AIDS, influenza virus and TB, the deadliest pre-COVID-19 infectious disease of humanity in modern times (Table 1). It has been suggested that the SARS-CoV-2 may represent the fifth CoV to be introduced and become permanently established in the human population, along with the prior four seasonal CoVs that currently circulate in humans (Escobedo et al. 2022; Li et al. 2021). Ultimately, this will be mainly determined by the forthcoming interplay between the host (e.g., immunity and susceptibility to the virus), the environment and the viral pathogen (e.g., potential emergence of newly SARS-CoV-2 variants) (Escobedo et al. 2022; Li et al. 2021).

Human-tropic COVS

The CoVs belong to the *Coronaviridae* family and the *Orthocoronavirinae* subfamily. They are the largest of the RNA viruses, carrying a positive-sense RNA strand of approximately 26–32 kilobases. The CoVs are spherical or pleomorphic in shape and approximately 125 nm in diameter with surface club-shaped projections, similar to what would resemble a crown, from hence the name (from the Latin "corona," meaning "crown") (Chen et al. 2020). These enveloped RNA viruses replicate in the host cell cytoplasm and antagonize the interferon and other host proteins of the innate immune response.

CoVs consists of four serologically and genetically differentiated genera, the *Alphacoronaviruses* (*Alpha-CoV*), *Betacoronaviruses* (*Beta-CoV*), *Gammacoronaviruses* (*Gamma-CoV*), and *Deltacoronaviruses* (*Delta-CoV*). CoVs infect over 12 mammalian and avian species and demonstrate the capacity to cross the species barrier. They generally exhibit host genus-specificity, characterized by a narrow host range, infecting one or just a few species dictated by the viral interaction with the corresponding host cell receptor (Leopardi et al. 2018). There are seven known human CoVs. These belong to the *Alpha-CoV* and *Beta-CoV* genus. The

Table 1 Comparative burden on human life of five major infectious diseases (1 year estimates)

Disease	Infectious agent	Estimated N cases	Estimated N deaths	Fatality rate (%)	Global spread
Tuberculosis [1]	*M. tuberculosis*	10 000 000 [1]	1 500 000	15.00	World wide [1]
Malaria [2]	*Plasmodium* sp.	219 000 000 [2]	435 000	0.20	WHO African Region [2]
HIV/AIDS [3]	HIV	37 900 000 [3]	770 000	2.03	Sub-Saharan Africa [3]
COVID-19 [4]	SARS-CoV-2	55 000 000 [4]	1 300 000	> 2.00	World wide [4]
Viral influenza [5]	Influenza virus	3 000 000 – 5 000 000	290 000 – 650 000	< 0.25	World wide [5]

TB tuberculosis, *M. tuberculosis Mycobacterium tuberculosis*, *HIV* human immunodeficiency virus, *AIDS* acquired immunodeficiency syndrome, *SARS* severe acute respiratory syndrome, *CoV* coronaviruses, *N* number of, sp. = Species; fatality rate = N. deaths/ N. cases (%), *EMR* mean influenza-associated respiratory excess mortality rate

[1] In 2018, the 30 high TB burden countries accounted for 87% of new TB cases. Eight countries account for two thirds of the total number of cases, led by India, followed by, China, Indonesia, the Philippines, Pakistan, Nigeria, Bangladesh, and South Africa. Number of cases corresponds to the 2018 incidence rate. Number of deaths reported for 2018. https://www.who.int/news-room/fact-sheets/detail/tuberculosis (accessed on June 11, 2020)

[2] In 2017 approximately 93% of all deaths were in the WHO African Region and almost 80% of all deaths in 2017 occurred in 17 countries in the WHO African Region and India. Number of cases corresponds to the 2018 incidence rate. Number of deaths reported for 2018. https://www.who.int/gho/malaria/en/ (accessed on June 11, 2020)

[3] Sub-Saharan Africa remains the most heavily affected region of the world, accounting for approximately two thirds of all incident and prevalent HIV infections and three quarters of all AIDS deaths. Number of cases corresponds to people living with HIV in 2018. Number of deaths reported for 2018. https://www.who.int/gho/hiv/en/ (accessed on June 11, 2020)

[4] Estimate relative to the first year of the COVID-19 pandemic. Almost a year after the initial report by the health authorities in Wuhan, China, on December 31, 2019, the SARS-CoV-2 pandemic reached 190 countries, surpassing the 2018 mortality from malaria, HIV/AIDS and viral influenza, and rapidly approaching that of the most deadly infectious disease, tuberculosis. https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6 (accessed on November 17, 2020)

[5] In a sample representing 57% of the global population, EMR ranged from 0.1 to 6.4 per 100,000 individuals for people younger than 65 years, 2.9 to 44.0 per 100,000 individuals for people aged between 65 and 74 years, and 17.9 to 223.5 per 100,000 for people older than 75 years (Iuliano et al. 2018). https://www.who.int/en/news-room/fact-sheets/detail/influenza-(seasonal) (accessed on June 11, 2020)

CoV-NL63 and CoV-229E are representative human *Alpha-CoVs*. Human *Beta-CoVs* include SARS-CoV-2, SARS-CoV, MERS-CoV, CoV-HKU1, and CoV-OC43. Three of these species are particularly pathogenic to humans: the SARS-CoV, its closely resembling SARS-CoV-2, and the MERS-CoV. SARS-CoV and MERS-CoV have originated from bats with intermediate hosts identified as civet cat and dromedary camels; respectively (Walls et al. 2020). The origin of SARS-CoV-2 is still controversial. The remaining endemic species HCoV-229E, HCoVNL63, HCoV-OC43 and HCoV-HKU1 are the known cause of common colds.

CoVs are composed of four basic structural proteins in their lipid bilayer envelope, the spike (S), envelope (E), membrane (M), and nucleocapsid (N), encoded by one-third of the CoV genomic RNA. The role of the spike glycoprotein (S) is particularly relevant as the determinant of species specificity in CoV infections (Kuo et al. 2000), and has been explored as the major anti-CoV target (Trigueiro-Louro et al. 2020).

SARS-COV-2 infection

Although SARS-CoV-2 is a new challenge, similarities with other CoVs set the groundwork for dissecting the immunological pathways that come to play in COVID-19 (de Wilde et al. 2018; Li et al. 2020a; Ortiz-Prado et al. 2020; Vabret et al. 2020). SARS-CoV-2 pathophysiology is being extensively investigated in the biological pathways governing the early stages and characterizing the immune responses to the infection. These pathophysiological mechanisms underpin the host genetic determinants of the disease.

The early stages of SARS-CoV-2 infection

In a comparative study carried out by Hui and collaborators, SARS-related CoVs showed a natural tropism for the human respiratory tract, although also targeting extrapulmonary cells, such as the vascular endothelial cells (Hui et al. 2020). The early stages of the infection process for SARS-CoV-2 are similar to those described for SARS-CoV, thus: 1) the viral S1 N-terminal domain of the S protein attaches to the angiotensin-converting enzyme 2 (ACE2), the targeted host cell surface receptor of the virus (Hoffmann et al. 2020; Walls et al. 2020; Zhou et al. 2020); 2) redundant proteolytic pathways ensure S protein activation required for cell entry. The S protein acts as the leading mediator of viral entry, binding the virus to the target cell receptor and allowing fusion of the viral envelope with the virus-target cell membrane (Belouzard et al. 2012; Heald-Sargent and Gallagher 2012; Trigueiro-Louro et al. 2020). CoV S proteins are class I membrane fusion proteins, synthesized as inactive precursors,

needing proteolytic activation (Kielian and Rey 2006). The ectodomain of the S protein consists of two subunits, an S1 N-terminal domain responsible for receptor binding and an S2 C-terminal domain responsible for membrane fusion. The two domains are joined by the protease sensitive S1-S2 junction. Two proteolytic pathways assist the virus in entering the cell: the endosomal pathway, characterized by the proteolytic activity of the endosomal cysteine proteases cathepsins (CTS), such as cathepsin L (CTSL) (Simmons et al. 2005), and/or the cell surface non-endosomal pathway, activating the transmembrane protease serine 2 (TMPRSS2) and the human airway trypsin-like (HAT) proteases (Chan et al. 2015). Moreover, as with the MERS-CoV, the proprotein convertase furin appears to preactivate S protein cleavage (Millet and Whittaker 2014), whereas in SARS-CoV the furin-like cleavage site is absent (Coutard et al. 2020). These factors typically co-express in the target cells contributing to the tissue tropism of the virus.

Increased transmissibility of SARS-CoV-2

Despite the similarities between the infection process of SARS-CoV-2 and SARS-CoV, major differences between the two viruses have been identified. Several factors support the increased transmissibility of SARS-CoV-2 in comparison to SARS-CoV. A greater affinity for the ACE2 receptor binding domain (RBD) of SARS-CoV-2 compared to that of SARS-CoV may account for a more efficient cell entry (Shang et al. 2020a). In SARS-CoV-2 the sequence of the S1-S2 junction shows greater affinity for the host proteases that activate the S protein function, than in SARS-CoV (Meng et al. 2020). Moreover, the activation of the SARS-CoV-2 S protein by furin, abundantly expressed in the lungs, reduces the dependence of the virus on the target cell surface protease, TMPRSS2, and the lysosomal proteases, CTSs (Shang et al. 2020b). Interestingly, the SARS-CoV-2 genotypes more ancestrally related to CoVs in animals are being replaced by still emerging genotypes presenting higher transmissibility in humans (Tang et al. 2020; Karim and Karim 2021).

Challenges in the face of the new SARS-CoV-2 variants

Key mutations within the Spike protein affecting the biological functions of SARS-CoV-2 led to the emergence of the viral genetic variants that have been circulating worldwide and contribute to different phenotypes of COVID-19 upon infection by these variants (Osuchowski et al. 2021; Zhou and Wang 2021). The World Health Organization (WHO) along with the Centers for Disease Control and Prevention (CDC) currently define four classes of SARS-CoV-2 variants, namely: Variants under Monitoring (VUM), Variants of Interest (VOI), Variants of Concern (VOC) and Variants

of High Consequence (VOHC) (CDC 2021; WHO 2022). Currently, there are no variants categorized as VOHC (CDC 2021; WHO 2022). The B.1.525 (Eta); B.1.526 (Iota); B.1.526.1; B.1.617; B.1.617.1 (Kappa); B.1.617.3; B.1.1.28.2/P.2 (Zeta); B.1.1.28.3/P.3; and B.1.427 and B.1.429 (Epsilon) variants are currently categorized as VUM or as formerly monitored variants, since no clear evidence of their severe phenotypic or epidemiological impact exists (WHO 2022). These variants are no longer detected or are circulating at very low levels in the USA and Europe, and as such, do not pose a major significant or imminent risk to Public Health (CDC 2021; WHO 2022). The Lambda (C.37) and Mu (B.1.621) variants are labelled as VOI and have been under evaluation with respect to their phenotypic features (WHO 2022).

The Alpha (B.1.1.7), Beta (B.1.351 and descendent lineages), Gamma (B.1.1.28.1/P.1 and descendent lineages), Delta (B.1.617.2 and descendent lineages), and Omicron (B.1.1.529 lineage along with the BA.1, BA.2 and BA.3 descendent sublineages) variants are currently labelled as VOC and are responsible for different SARS-CoV-2 phenotypes (CDC 2021; Zhou and Wang 2021; WHO 2022).

These variants have acquired a transmission advantage in comparison to the wild-type virus. In addition, the Omicron variant has been demonstrated to have a replication advantage over the remaining VOCs (CDC 2021; McIntosh et al. 2022; WHO 2022). The SARS-CoV-2 VOCs (specially, Gamma and Delta variants) are suggested to be more pathogenic (resulting in a higher risk of hospitalization and death); and it is reported that a reduction in neutralization by post-vaccination sera and a potential reduction in neutralization by the approved monoclonal antibody treatments against COVID-19 occurs upon SARS-CoV-2 VOC infection (mainly for Gamma, Delta, and Omicron variants) (Davies et al. 2020; CDC 2021; Khateeb et al. 2021; Lopez Bernal et al. 2021; Zhou and Wang 2021; McIntosh et al. 2022; WHO 2022). All VOC share the Spike mutation D614G (mutation from aspartic acid to glycine at position 614), which was first identified in March 2020 (Lauring and Hodcroft 2021). Considering that the binding of the Spike receptor-binding domain to the ACE2 is essential for SARS-CoV-2 to enter and infect cells, key mutations within the RBD have the potential to modulate the binding ability (Zhou and Wang 2021; McIntosh et al. 2022). Viruses bearing this mutation have rapidly spread and dominated worldwide; they are suggested to have a more transmissible phenotype with respect to a higher replication capacity and higher levels of viral load in the respiratory tract; and associated to an increased ACE2 binding affinity (Cho and Glenn 2020; McIntosh et al. 2022).

Transmission of Alpha, Beta, and Gamma VOCs has been significantly reduced over time, and these variants are not widely circulating (CDC 2021; McIntosh et al. 2022; WHO 2022). Omicron VOC has surpassed the Delta VOC as the predominant variant responsible for SARS-CoV-2 transmission, worldwide (ECDC 2022; McIntosh et al. 2022; WHO 2022). While the BA.1 sub-lineage of Omicron has been established as dominant, BA.2 reported sequences have recently increased in proportion in comparison to other Omicron sub-lineages (ECDC 2022; WHO 2022). Hence, the current epidemiological context with respect to Omicron VOC should be closely monitored.

There is a permanent concern regarding how emerging mutations in the virus may direct it to escape the developed diagnostics and therapeutics, which may consequently reduce the impact of vaccine and drug development efforts (Cho and Glenn 2020). A global overview of the SARS-CoV-2 VOC is summarized in Table 2.

Induction of inflammatory cytokines

The interferon-mediated immune response has a prominent role in limiting infection. It is triggered by the host's perception of pathogen-associated molecular patterns (PAMPs). PAMPs are recognized by the host pattern recognition receptors (PRR). Toll-like receptors (TLRs) and RIG-I-like receptors (RLRs) constitute distinct families of PRR that sense viral nucleic acids, activating transcription factors that lead to the production of type I interferon (IFN-I) (Akira et al. 2006; Kawai and Akira 2008). There are three main pathways leading to the production of IFN-I. The RIG-I receptor signaling pathway, activated by RNA viruses, the TLR3- and TLR4 TRIF-dependent TLR receptor signaling pathway, and the TLR7/8- and TLR9 MyD88-dependent TLR receptor signaling pathway (Kato et al. 2005; Barreiro et al. 2009). The activation of these receptors triggers transcription factors, such as NF-κB, the interferon regulatory factors IRF3 and IRF7, and the production of IFN-I. In the extracellular environment, IFN-I reacts with its specific receptor (IFNAR) to trigger the JAK-STAT signaling cascade (Janus kinase-signal transducer and activator of transcription), inducing the phosphorylation of STAT1 and STAT2 and the formation of the STAT1-STAT2 heterodimer. The STAT1-STAT2 heterodimer connects to the IRF9 transcription factor, forming the IFN Stimulated Gene Factor 3 (ISGF3) complex. In the nucleus, ISGF3 binds to the promoter Interferon-Stimulated Response Element (ISRE). This leads to the transcription of IFN-stimulated (ISN) genes, whose effector products promote the antiviral response by modulating the signaling pathways or by exerting direct antiviral activity (Platanias 2005; Sadler and Williams 2008; Yoneyama and Fujita 2009; Schoggins et al. 2011). Regulation by these cascades has been identified in CoV infections (Dandekar and Perlman 2005; Channappanavar and Perlman 2017; Ortiz-Prado et al. 2020).

Table 2 Summary description of the characteristics of SARS-CoV-2 variants of concern[1]

WHO nomenclature	PANGO[2] lineage/ nomenclature	Key mutations in Spike protein[3]	Date first detected	Location first detected	Phenotype
Alpha	B.1.1.7	ΔH69/V70 Δ144Y **(E484K)**[4] **(S494P)**[4] **N501Y** A570D D614G P681H	September 2020	United Kingdom	•Transmissibility increased on average by 40-50% •Potential increase in disease severity •Minimal/no impact on neutralization capacity by monoclonal antibody therapies against COVID-19 •Minimal or no impact on neutralization capacity by convalescent and post-vaccination sera
Beta	B.1.351	**K417N** **E484K** **N501Y** D614G	August/ September 2020	South Africa	•Transmissibility increased on average by ~50% •Possibly increase in disease severity and mortality (under investigation) •Significant impact on neutralization capacity by bamlanivimab-etesevimab (>45-fold decrease in susceptibility); minimal impact when considering other monoclonal antibody therapies •Moderate reduction in neutralization by convalescent and post-vaccination sera
Gamma	B.1.1.28.1 / P.1	**K417N/T** **E484K** **N501Y** D614G	Late 2020 / January 2021	Brazil/Japan	•Transmissibility increased by > 50% •Increase in disease severity and mortality (including in individuals with no pre-existing disease) •Significant impact on neutralization capacity by bamlanivimab-etesevimab (>511-fold decrease in susceptibility); minimal impact on other monoclonal antibody therapies •Reduction in neutralization capacity by convalescent and post-vaccination sera
Delta	B.1.617.2	Δ156-157 R158G **L452R** **T478K** D614G P681R D950N	December 2020	India	•Increased transmissibility and pathogenicity compared with the Alpha variant •Potential minimal or without reduction in the neutralization capacity by monoclonal antibody therapies •Potential reduction in neutralization by convalescent and post-vaccination sera (vaccine effectiveness is slightly less for delta variant than for Alpha variant)

Table 2 (continued)

WHO nomenclature	PANGO[2] lineage/ nomenclature	Key mutations in Spike protein[3]		Date first detected	Location first detected	Phenotype
Omicron	B.1.1.529	A67V ΔH69-V70 T95I G142D ΔV143-Y145 ΔN211 L212I R214-aa insertion ("EPE") **G339D** **S371L** **S373P** **S375F** **K417N** **N440K** **G446S** **S477N**	**T478K** **E484A** **Q493R** **G496S** **Q498R** **N501Y** **Y505H** **T547K** D614G H655Y N679K P681H N764K D796Y N856K Q954H N969K L981F	November 2021	Botswana (Multiple other countries)	●Transmissibility increased over Delta variant and other VOCs. ●Impact on pathogenicity is under investigation ●A significant reduction in the neutralization capacity by monoclonal antibody therapies (bamlanivimab-etesevimab and casirivimab-imdevimab) is anticipated Expected reduction in neutralization by convalescent and post-vaccination sera (with the standard regimen)

[1]Adapted from; Zhou and Wang 2021; CDC 2021; NSW government 2021; McIntosh et al. 2022; WHO 2022

[2]Pango (or Pangolin) lineage: Phylogenetic Assignment of Named Global Outbreak

[3]Spike-Receptor-binding domain (RBD) mutations are marked in bold

[4]Detected in a limited number of sequenced viruses

Dissimilarities in IFN-I inhibition between SARS-CoV and SARS-CoV-2 have been reported (Lokugamage et al. 2020). SARS-CoV-2 is more sensitive to IFN-I, through phosphorylation of STAT1 and induced ISG proteins, while SARS-CoV is indifferent to its antiviral effects. The differential host innate immune modulation between the two viruses could originate the observed contrasts in disease pathophysiology and progression to severe disease. Investigations using an approach based on the determination of differentially conserved positions (DCPs) affecting viral protein structure and function evidenced changes in viral proteins, between the two epidemic CoV strains, that may affect IFN-I inhibition. An early and properly localized, IFN-I response can effectively limit CoV infection (Rausell et al. 2010; Bojkova et al. 2020; Trigueiro-Louro et al. 2020). Conversely, failure to elicit an early IFN-I response has been associated with severe CoV disease (Channappanavar et al. 2019; Hadjadj et al. 2020). Moreover, IL-6 and C-reactive protein are significantly upregulated in patients that died, compared to convalescent COVID-19 patients (Lavillegrand et al. 2021). High cytokine levels in patients with severe COVID-19, indicating that a storm of chemokines and cytokines occurred could be associated with the severity of the disease (Qin et al. 2020). Increased levels of pro-inflammatory cytokines were associated with major pulmonary inflammation and extensive lung damage in both SARS and MERS (Wong et al. 2004; Mahallawi et al. 2018). A leading cause of COVID-19 mortality is respiratory failure caused by acute respiratory distress syndrome (ARDS). This excessive inflammatory reaction has been considered an important contributor to ARDS and multiple organ dysfunction syndrome (Wang and Ma 2008; Suri et al. 2021).

Interestingly, studies in human alveolar epithelial cells and macrophages showed SARS-CoV-2 to be a less potent inducer of proinflammatory cytokines than influenza viruses or MERS-CoV. In contrast to influenza virus infection, and as previously observed with MERS-CoV, there appears to be no evidence that the elevated levels of cytokines and chemokines observed in intensive-care unit (ICU) COVID-19 patients, are major contributors to SARS-CoV-2 pathogenesis. It has been suggested that these traits perhaps reflect the more severe lung damage that has taken place (Chan et al. 2013; Hui et al. 2020; Liu et al. 2020).

Furthermore, a suggested consequence of macrophage activation and pro-inflammatory cytokine release could be to trigger a coagulation cascade, leading to immune-mediated arterial thrombosis (Merad and Martin 2020). The involvement of TNF and IL-1 cytokines could also imply the activation of cytokine and transcription factor signaling, such as the MAPK and NF-kappa B signaling pathways.

Adaptive immune response

Most viral infections eventually lead to the death of host cells through different types of regulated cell death (RCD) mechanisms (Tang et al. 2019, 2020). CoVs are known to decrease T cell populations, by inducing the intrinsic and extrinsic pathways of T cell apoptosis, thereby compromising B cell activation and humoral immunity (Chu et al. 2016; Channappanavar et al. 2019).

T cells are one of the major components of the adaptive immune system. They play a vital role in antiviral immunity. Major groups of effector T lymphocyte are differentiated according to the co-receptor (CD8 or CD4) used to bind the human leucocyte antigen (HLA) molecules where viral antigens are loaded. On one hand, to eradicate viruses from the host, the CD8 + cytotoxic T cells kill their target cells primarily by releasing molecules such as perforin, granzymes, and IFN-γ. On the other hand, the role of CD4 + helper T cells includes activating other immune cells, releasing cytokines. The CD4 + helper T cells enhance cytotoxic T cells and B cells, to generate stronger antibody responses and help clear the pathogen.

The role of T cell activation in COVID-19 clinical recovery is incontrovertible. T-cell mediated adaptive immune responses have been shown to be critical in the clearance of CoV infections including SARS-CoV-2 infection in animal models (Zhao et al. 2010; Sun et al. 2020). In humans, T cell responses are required for protection from clinical disease, contributing to the elimination of the SARS-CoV-2 virus, conferring protective immunity and leading to clinical recovery (Li et al. 2020a; Nicoli et al. 2020). SARS-CoV-2-specific CD4 + and CD8 + T cell responses were observed in many of the patients who had recovered from COVID-19 (Peng et al. 2020). Zhao and collaborators concluded that T cells made up one of the three critical axes for viral clearance and disease resolution, along with IFN I and signal transducer and activator of transcription 1 (STAT1) (Zhao et al. 2010). T cells are important in immunopathology as well as cross-immunity to other coronaviruses (Zhao et al. 2016; Weiskopf et al. 2020).

In the early stages of severe COVID-19, patients show a significant reduction in the number of CD4 + T cells, CD8 + T cells, B cells and natural killer (NK) cells, monocytes, eosinophils and basophils (Huang et al. 2020; Qin et al. 2020; Wu J et al. 2020b; Xu et al. 2020). In patients with acute moderate or severe COVID-19, however, the absolute numbers and relative frequencies of CD4 + and CD8 + T cells are unphysiologically low (He et al. 2020; Liu et al. 2020; Sekine et al. 2020). Several investigations suggest that this is due to the functional exhaustion of T cells that could result from increased expression of programmed cell death (Diao et al. 2020; Zheng HY et al. 2020a; Zheng M et al. 2020b). Diao and collaborator observed that the numbers of total T cells, CD4 + T and CD8 + T cells are negatively correlated to levels of TNF-α, IL-6, and IL-10, respectively, suggesting these cytokines may be involved in the decrease of T cells detected in COVID-19 (Diao et al. 2020).

The genetic basis of COVID-19 susceptibility and severity

From a perspective on the genetic basis of covid-19 susceptibility and severity, researchers are interested in identifying the molecular pathways governing the remarkable range in the clinical presentation of COVID-19. Although in recent assessments some of the new SARS-CoV-2 variants appear to increase disease severity and mortality, there is no evidence that viral mutations in SARS-CoV-2 could explain the extent in which clinical phenotype variation has been observed nor the observed differences in the geographic distribution of COVID-19 (Debnath et al. 2020; Pan et al. 2020; Zhang et al. 2020a). Disease outcomes are also affected by age, sex, and stochastic events. Moreover, the immune system ages after each infection and is trained by past infections. Nevertheless, the wide range in the clinical presentations observed suggests that host genetics could also be a major contributor to the differences observed in the host–pathogen relationship.

A wide spectrum of COVID-19 clinical presentations

The spectrum of COVID-19 clinical features ranges from mild disease (none or a mild pneumonia), 81%, severe disease (e.g., with dyspnea, hypoxia, or > 50% lung involvement on imaging within 24 to 48 h), 14%, critical disease (e.g., with respiratory failure, shock, or multiorgan dysfunction), 5%, with an overall case fatality rate of 2.3% (McIntosh et al. 2022). Moreover, there remains uncertainty about the proportion of individuals that remain asymptomatic, with a wide range of reports across studies. Pediatric populations aged under 20 years, appear to be less vulnerable to COVID-19 than adults, representing only 2% of the reported cases, with more than 90% of the cases either being asymptomatic, mild, or moderate. Interestingly, the emerging multisystem inflammatory syndrome in children (MIS-C), or the COVID-19 associated pediatric inflammatory multisystem syndrome temporally associated with SARS-CoV-2 (PIMS-TS), are multisystem inflammatory conditions that overlap with the diagnostic criteria of Kawasaki disease (KD) and KD shock syndrome (Jiang et al. 2020).

Population-based genotype to phenotype association studies

Genetic variants affecting clinical phenotype can be identified through genotype to phenotype association studies. Once identified, genetic determinants of disease susceptibility and severity can contribute to disease control as a tool in early diagnosis for risk evaluation, allowing patients at greater risk to be prioritized for protective prophylactic vaccination or to be referenced to special medical care and surveillance. Moreover, genetic association studies have been central to the development of novel therapeutic schemes aimed at surpassing deficiencies in gene products with causal implications in severe disease (Holland 2001; Alangari et al. 2011; Casanova et al. 2014). This is an ongoing strategy in COVID-19 investigations (Bastard et al. 2020; Zhang Q et al. 2020b).

The genome-wide association study (GWAS), also known as the whole genome association study (WGAS), and the candidate gene association study (CGAS) are population-based approaches in genotype to phenotype association. A GWAS is an untargeted hypothesis-free genome-wide screening strategy. It is the usual approach for common variants. The candidate gene approach is a rational approach where the targeted gene/s and variant/s are selected based on a priori hypothesis about their implication in disease causation (Fig. 1) (David 2021). It is more adaptable to identify lower minor allele frequency variants than the GWAS and may also be considered in validation strategies for genome-wide studies. The availability of cost effective next-generation sequencing (NGS) resources has been key to the gained momentum in the CGAS approach.

In simplified terms, the GWAS can be viewed as hypothesis-generating and the CGAS as hypothesis-testing approaches (Fig. 2).

Genome-wide and candidate gene association studies applied to COVID-19

Of interest to genome-wide and candidate gene association studies, the current genetic theory of infectious diseases, suggests a monogenic (single gene) basis of most infections, characterized by incomplete penetrance (non-Mendelian) and genetic heterogeneity (multiple genes and alleles) but physiological homogeneity (a molecular basis pointing to a cellular and immunological mechanism) (Casanova and Abel 2020, 2021). Applied to today's pandemic, most humans would, in fact, be "immunodeficient" with respect to SARS-CoV-2 and a physiological homogeneity in the molecular pathways governing severe COVID-19 is to be expected.

Whole-genome studies and whole-exome sequencing (WES) have taken the lead in the efforts to identify the human genetic determinants of COVID-19 susceptibility and severity. Several consortiums are now contending with this task including the COVID Human Genetic Effort (https://www.Covidhge.com/) (Casanova et al. 2020), deCODE Genetics (https://www.decode.com/) (Hakonarson et al. 2003), the COVID-19-GR (http://www.gsrt.gr/central.aspx?sId=119I428I1089I646I488772) and the Harvard's Wyss Institute and the Personal Genome Project at Harvard University (https://wyss.harvard.edu). In particular, the COVID-19 host genetics initiative provides an environment to foster the sharing of resources to facilitate COVID-19 host genetics research (e.g., protocols, questionnaires), organizes analytical activities across the studies, and provides a platform to share results to the benefit of the broader scientific community (COVID-19 Host Genetics Initiative 2020).

From a different perspective, in a new disease like COVID-19, for which the pathophysiological mechanisms of the disease remain largely unknown, the rational approach of a CGAS poses an obvious challenge. Hypothesis generating for candidate gene studies of COVID-19 relies on our knowledge of the prepandemic CoV infections, our current understanding of the molecular mechanisms governing the pathophysiology of SARS-CoV-2 infection, and recent findings from genome wide studies. Since the beginning of the pandemic, several reviews have been published on the host genetics of the immune response associated with COVID-19 susceptibility, severity, and outcome (Table 3) (Anastassopoulou et al. 2020; Carter-Timofte et al. 2020; Elhabyan et al. 2020; Godri Pollitt et al. 2020; LoPresti et al. 2020; Oladejo et al. 2020; Ovsyannikova et al. 2020; Zunec 2020; Colona et al. 2021a; SeyedAlinaghi et al. 2021; Suh et al. 2022). Moreover, in line with this renewed interest, in a recent systematic review of genes related to viral susceptibility reported in human genetic studies, the CGAS approach was considered as a valid approach (Elhabyan et al. 2020).

Considerations on the definition of clinical phenotype

In genetic association studies, a clear definition of the clinical phenotype is key, for it must be kept in mind that different definitions of phenotype typically lead to different results (Fig. 3). The need for respiratory support or death due to COVID-19 are major criteria for a severe phenotype classification of laboratory confirmed SARS-CoV-2 infection (Table 4). The main strategy undertaken by the consortium COVID Human Genetic Effort is to analyze previously healthy young patients with severe and unexplained forms of COVID-19 (Casanova et al. 2020). A second strategy adopted by this group intends to explain why some individuals, despite heavy exposure to SARS-CoV-2, remain healthy and seronegative, particularly those individuals that also test negative for T cell responses to SARS-CoV-2, that is, they appear to be naturally resistant to SARS-CoV-2 infection.

Genetic association studies
What is the causal variant?

Study design
→ Case-control / cohort

Phenotype definition
→ Enrich for the risk allele in the group of cases

Candidate gene and candidate variant selection
→ Literature review and bioinformatic tools

Haplotype analysis and linkage disequilibrium
→ Use of tag SNPs

Sample size estimation and statistical significance
→ sample size required, given the significance level, effect size and statistical power

Exploratory data analysis and quality control of genotyping data
→ Identify trends as well as possible shortcomings of the data

Testing statistical association
→ Is the risk allele more frequent, than would be expected by chance alone, in the group of cases as compared to the controls?
→ Results are reported with the effect measure and confidence interval

What is the function?

Funcional analysis
→ Validation of the causal hypothesis

Fig. 1 Critical steps to CGAS: scientific protocol, study design, phenotype definition, candidate gene and variant selection, haplotype analysis and linkage disequilibrium (LD), sample size estimation, statistical significance and *p*-value, exploratory data analysis (EDA) and quality control of genotyping data, testing statistical association and functional analysis (adapted from, David 2021)

A detailed list of criteria for phenotype definition of severe COVID-19 is given in the COVID-19 Host Genetics Initiative website (COVID-19 Host Genetics Initiative 2020). It includes demographic data, risk factors, symptoms at admission Moreover, care and hospital related attributes and possible comorbidities of the immune, cardiovascular, neurological, and respiratory systems, as well as cancer, taken into account, also contribute to improve the reproducibility of genetic association studies. Male sex was identified by a global COVID-19 meta-analysis as being a risk factor for death and ITU admission (Peckham et al. 2020). This observation may be due to sex differences in both the innate and adaptive immune systems or sex differences in X chromosome linked ACE2 receptor expression (Peckham et al. 2020). Moreover, significant positive genetic correlations were detected for several adiposity phenotypes (Pairo-Castineira et al. 2021). Vitamin D status has also been recently recognized risk factor for COVID-19, with deficiency conferring greater risk (Grant et al. 2020). Other possible cofactors of disease severity include secondary infections and quantitative traits that have been associated with disease severity such as mannose-binding lectin (MBL), ferritin, D-dimer, C-reactive protein, and increased levels of proinflammatory cytokines, absolute neutrophil counts, and the neutrophil to lymphocyte ratio (NLR) (Ortiz-Prado et al. 2020; Ovsyannikova et al. 2020).

Causality in genotype to phenotype associations

To support a causal association of the alleles identified, case–control genetic association studies need to be validated: Validated results are core in tracking other pathway related targets with increased advantages, such as druggability (Baillie 2014; King et al. 2019). In the case of COVID-19, the selection of possibly more druggable targets is strategic (Parkinson et al. 2020; Novelli et al. 2021). The validation
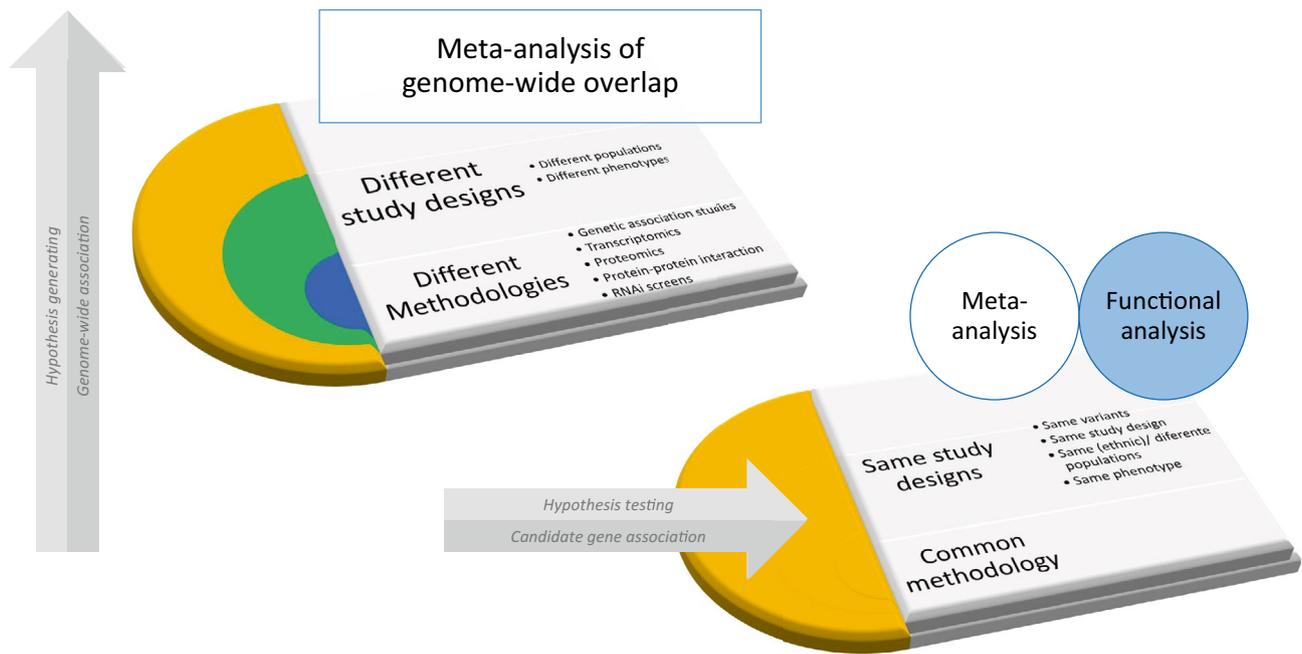
Fig. 2 To support a causal association, case–control genetic association studies can be validated through replication. Candidate gene association studies in targeted biological pathways are useful for the replication of genome-wide results. They are also more adaptable to identify lower minor allele frequency variants and thus more prone to a functional analysis follow-up. Meta-analysis is a statistical method for analyzing large collections of results from independent association studies including candidate gene and genome-wide association studies. Meta-analysis can be used to compile results from methodologically uniform studies (conventional application) (Skol et al. 2006; Zondervan and Cardon 2007) or from various methodological approaches (Li et al. 2020c; Pairo-Castineira et al. 2021; Parkinson et al. 2020). Meta-analysis of genome-wide overlap (MAGO), or Meta-analysis of overlapping SNVs from genome-wide approaches, provides a relevant means of aggregating evidence for causal association. There is an increased need for replication strategies by which to infer causality for candidate gene selection and validation, pointing to the possible role of other candidate pathway-related targets

process is elaborate and begins as early in the process as the initial design of the scientific protocol, from the careful and correct definition of phenotype to in vitro and in vivo validation through functional anylisis (Fig. 1) (David 2021). Other points for CGAS and WGAS are discussed below.

Replication

One means of validation genotype to phenotype associations is through replication (Fig. 2). There is an increased need for replication strategies by which to infer causality for candidate gene selection and validation. Conventionally, a replication study involves the analysis of the same variants in the same direction of effect and in the same (ethnic) population, considering the same phenotype as the original study (Skol et al. 2006; Zondervan and Cardon 2007). However, the findings can also be replicated in different populations, using different study designs, or applying various methodological approaches. Conformingly, the use of meta-analysis (MA), as an element of a systematic review, is a statistical method of analyzing large collection of results from independent CGAS or genome-wide association studies. Data from several independent CGAS

can be grouped by meta-analysis, for power and precision increment in the association results. Furthermore, meta-analysis of overlapping single nucleotide variants (SNVs) from genome-wide approaches, or Meta-analysis of genome-wide overlap (MAGO), provides a relevant means of aggregating evidence for causal association within and across genome-wide methodological approaches.

Criticalities of GWAS

Although the contribution of GWAS to the understanding of the genetic determinants of disease is widely accepted, several caveats to the validation of GWAS results exist (Colona et al. 2021a, 2021b). These include the facts that (1) the effect sizes of the identified alleles, namely in terms of odds ratio (OR), are generally too low (OR < 2) for these to be considered as predictive genomic markers (Table 5); (2) most of these studies do not take into account the contribution of nongenetic and environmental factors implicated in the disease phenotype; (3) many of the biobanks currently available lack essential complementary patient information and are unrepresentative of mixed and complex ethnic groups; (4) computer algorithms can lead to interpretive errors if they are

Table 3 Review articles identifying candidate genes through COVID-19 host-related factors

Description	Reference
Mini-review on the genetic determinants of COVID-19	Godri Pollitt et al. 2020
Proposal of three potentially important genetic gateways governing COVID-19 infection and severity: 1) ACE2 pathway; 2) Human Leukocyte Antigen (HLA) locus; and 3) Toll-like receptor and complement pathways and induced inflammatory pathways	Debnath et al. 2020
Review of the host genetic factors related to CoVs that affect different species, including humans	LoPresti et al. 2020
Review of the prepandemic and COVID-19 association studies	Ovsyannikova et al. 2020
Review of the clinical features and the host immune response to suggest candidate immune pathways for genetic association studies of severe COVID-19	Carter-Timofte et al. 2020
Systematic review on the genetics determinants of SARS Virus disease [1]	Elhabyan et al. 2020
Review of HLA and COVID-19 genetic association studies	Zunec 2020
Systematic review on the genetics determinants of COVID-19 [2]	Anastassopoulou et al. 2020
Updated summary of ultra-rare, rare, and common human variants, haplotypes, and susceptibility gene polymorphisms detected in several studies, through various approaches	Colona et al. 2021a
Systematic review of current evidence to investigate the genetic susceptibility of COVID19 in Scopus, PubMed, Web of Science and Science Direct databases [3]	SeyedAlinaghi et al.2021
Systematic review on the genetics determinants of COVID-19 severity [4]	Suh et al. 2022

[1] Systematic review inquiry: "Genetic Predisposition to Disease"[Majr] AND "Virus Diseases"[Majr]; "SARS Virus"[Mesh] AND Genetic [Title/Abstract]

[2] Systematic review inquiry: "SARS-CoV-2," "2019- nCoV," AND "COVID-19"; "polymorphisms," "allelic variation," "genetic predisposition," "genotype," "clinical outcome"; names of individual genes in which relevant polymorphisms were found

[3] Systematic review inquiry: diferent combinations of keywords in the following orders, 1. "COVID-19" or "SARS-CoV-2" or "Novel Coronavirus" or "2019-nCoV" or "Coronavirus" (title/abstract); 2. "Genetic susceptibility" or "Genetic vulnerability" or "Genetic probability" (title/abstract); 3. (A) and (B)

[4] Systematic review inquiry: ((SARS-CoV-2[tiab] OR COVID-19[tiab] OR "Coronavirus disease"[tiab] OR "Severe acute respiratory syndrome coronavirus 2"[tiab] OR coronavirus[tiab]) OR (SARS-CoV2[Mesh] OR "Spike Glycoprotein, Coronavirus"[Mesh] OR COVID-19[Mesh] OR Betacoronavirus[Mesh] OR "Coronavirus Infections"[Mesh])) AND (("Variant gene"[tiab] OR "whole-exome sequencing"[tiab] OR "allele frequency"[tiab] OR mutations[tiab] OR "protein–protein interaction"[tiab] OR "Significant linkage disequilibrium"[tiab] OR LD[tiab] OR PPI[tiab] OR Variants[tiab] OR Coding[tiab] OR Missense[tiab] OR "epigenetic modifcation"[tiab] OR polymorphism[tiab]) OR ("Molecular Docking Simulation"[Mesh] OR "Protein Interaction Domains and Motifs"[Mesh] OR "Virus Internalization"[Mesh] OR "High-Troughput Nucleotide Sequencing"[Mesh] OR "Polymorphism, Single Nucleotide*"[Mesh] OR "Real-Time Polymerase Chain Reaction"[Mesh]))

based on representative alleles from international reference databases and certain predictive functional bioinformatics characteristics; (5) studies based on sampling hospitalized patients and the general population as controls for which no information is available about exposure to the virus might lead to an overstatement of the results and collider bias.

Novel approaches to circumvent these shortcomings play a major part in the current efforts of the scientific community (Colona et al. 2021b; Moon et al. 2021). Largely the need for more holistic approaches is recognized. These may call upon phenome-wide association study (PheWAS) approaches as well as the use of functional genomic databases (Colona et al. 2021b). The phenome-wide association study (PheWAS) approach is a complementary approach that can use genetic pleiotropy to validate known comorbidities and provide insight into biological and molecular mechanisms (Hebbring 2014; Moon et al. 2021). Moreover, functional genomic databases such as the Genotype-Tissue Expression (GTEx) database, grouping genotyping data linked to genome-wide gene expression patterns across a wide range of tissue types, allow for analysis at tissue-specific resolution (https://www.gtexportal.org) (Lonsdale et al. 2013).

Candidate genes and pathways associated with COVID-19

Recently, several genotype to phenotype association approaches have been deployed, contributing to the identification of candidate genes governing COVID-19 susceptibility and severity. These include in silico approaches; candidate gene approaches based on knowledge of the pathophysiological mechanisms of common human viral infections, COVID-19 and of other prepandemic CoV infections; and genome wide approaches such as GWAS, transcriptomics and CRISPR-Cas9 knockout screens.

Association results | Phenotype definitions

COVID-19 susceptibility
- **Controls** General population without knowledge on the exposure status to SARS-CoV-2
- **Cases** Individuals with laboratory confirmed SARS-CoV-2 infection symptomatic or severely symptomatic for COVID-19

COVID-19 Severity (respiratory failure)
- **Controls** Individuals with laboratory confirmed SARS-CoV-2 infection symptomatic for COVID-19 but not requiring hospitalization
- **Cases** Hospitalized severely symptomatic individuals requiring respiratory support (with or without consideration for age)

Resistance to COVID-19
- **Controls** Individuals exposed or heavily exposed to SARS-CoV-2 but remaining asymptomatic or seronegative
- **Cases** Individuals with laboratory confirmed SARS-CoV-2 infection symptomatic or severely symptomatic for COVID-19

Fig. 3 Illustration of different definitions of phenotype leading up to different results in case–control genetic association studies

In silico approaches

The human leukocyte antigen (HLA) system

The human leukocyte antigen (HLA) system is intrinsically implicated in an effective T cell mediated immune response. A central part of anti-viral immunity is antigen presentation to CD4 + and CD8 + T cells by the HLA system. HLA class I and HLA class II are cell-surface glycoproteins that regulate the immune system by presenting foreign peptides to T cells. In the case of viral epitopes, these are presented by HLA class I molecules to CD8 + T lymphocytes and HLA class II molecules to CD4 + T lymphocytes which can assist antibody-producing B-cells.

The HLA genes have come into focus as a potential marker for COVID-19 susceptibility and severity (Li et al. 2020a; Zunec 2020). Allele and haplotype frequencies in class I and class II HLA molecules have been associated with either protective or deleterious effects to MERS and SARS-CoV. Although, it has been difficult to detect a clear pattern of association due to

Table 4 Phenotypic criteria for genetic association studies of COVID-19

Clinical phenotypes absent or mild	Severe clinical phenotypes
■ General population without knowledge on the exposure status to SARS-CoV-2	■ Hospitalized severely symptomatic individuals requiring respiratory support (with or without consideration for age)
■ Individuals exposed or heavily exposed to SARS-CoV-2 but remaining asymptomatic or seronegative	
	■ Death due to COVID-19
■ Individuals with laboratory confirmed SARS-CoV-2 infection symptomatic for COVID19 but not requiring hospitalization	

Adapted from **COVID-19 Host Genetics Initiative** 2020 (https://www.covid19hg.org/about/) and **COVID Human Genetic Effort** 2020 (https://www.Covidhge.com/)

Table 5 MAIC score rank of some candidate genes from cited references and respective biological pathways as included in KEGG 2019 (Human)

Gene	Chr. location	MAIC [1]	Risk estimate [2]	Biological Pathways	References
IFNAR2 [3]	21q22.11	56	9	Immune System, Infectious Disease: Viral (Influenza A, Herpes Simplex Virus 1 Infection, Cancer, Coronavirus Disease–COVID-19), Signaling Molecules and Interaction, Signal Transduction	(Bastard et al. 2020; Pairo-Castineira et al. 2021; Zhang et al. 2020b; COVID-19 Host Genetics Initiative 2021)
CCHCR1	6p21.33	61	≤ 2	No Hits	(Pairo-Castineira et al. 2021)
TBK1 [3]	12q14.2	67	9	Immune System, Infectious Disease: Viral (Hepatitis B, Hepatitis C, Influenza A, Herpes Simplex Virus 1 Infection, Influenza A, Coronavirus Disease–COVID-19), Signal Transduction	(Bastard et al. 2020; Zhang et al. 2020b)
TLR3 [3]	4q35.1	73	9	Cell Growth and Death, Immune System, Infectious Disease: Viral (Hepatitis B, Hepatitis C, Influenza A, Herpes Simplex Virus 1 Infection, Coronavirus Disease–COVID-19)	(Bastard et al. 2020; Chen et al. 2018; Ciancanelli et al. 2015; Zhang et al. 2020b)
RPLP0	12q24.23	79	NA	Coronavirus Disease–COVID-19	(Amati et al. 2020)
ACTB	7p22.1	97	NA	Cardiovascular Disease, Cell Growth and Death, Cellular Processes, Immune System, Infectious Disease: Bacteria (Pathogenic Escherichia coli Infection, Shigellosis, Salmonella Infection, Yersinia Infection, Bacterial Invasion of Epithelial Cells), Signal Transduction	(Amati et al. 2020)
PCSK3/ FURIN	15q26.1	235	NA	No Hits	(Millet and Whittaker 2014; Coutard et al. 2020; Shang et al. 2020b Amati et al. 2020; Latini et al. 2020)
FYCO1	3p21.31	293	≤ 2	Infectious Disease: Bacterial (Salmonella Infection)	(Pairo-Castineira et al. 2021; Ellinghaus et al. 2020)
IL10RB	21q22.11	398	≤ 2	Infectious Disease: bacterial (Tuberculosis), Infectious Disease: parasitic (Toxoplasmosis), Infectious Disease: Viral (Human cytomegalovirus infection), Signaling molecules and interaction, Signal transduction	(Kousathanas et al. 2022)
HLA-B	6p21.33	479	NA	Cancer, Celular Processes, Immune System, Infectious Disease: Viral (Herpes Simplex Virus 1, Human Immunodeficiency Virus 1 Infection), Signaling Molecules and Interaction	(Nguyen et al. 2020)
IRF7 [3,4]	11p15.5	607	9 (ADM) > 50 (ARM)	Cardiovascular Disease, Immune System, Infectious Disease: Viral (Hepatitis B, Hepatitis C, Influenza A, Herpes Simplex Virus 1 Infection,, Influenza A)	(Bastard et al. 2020; Zhang et al. 2020b)
UNC93B1 [3]	11q13.2	608	9	No Hits	(Bastard et al. 2020; Zhang et al. 2020b, c)

Table 5 (continued)

Gene	Chr. location	MAIC [1]	Risk estimate [2]	Biological Pathways	References
ACE2	Xp22.2	610	NA	Coronavirus Disease – COVID-19	(Hussain et al. 2020; Benetti et al. 2020; Li Q et al. 2020; Hofmann et al. 2005; Lonsdale et al. 2013; Cao et al. 2020; Amati et al. 2020; Suryamohan et al. 2021; Suh et al. 2022)
IRF3 [3]	19q13.33	618	9	Cancer, Immune System, Infectious Disease: Viral (Hepatitis B, Hepatitis C, Human Immunodeficiency Virus 1 Infection, Herpes Simplex Virus 1 Infection, Influenza A, Coronavirus Disease – COVID-19)	(Bastard et al. 2020; Zhang et al. 2020b)
IFNAR1 [3,4]	21q22.11	641	9 (ADM) > 50 (ARM)	Cancer, Immune System, Infectious Disease: Viral (Hepatitis B, Hepatitis C, Human Immunodeficiency Virus 1 Infection,, Herpes Simplex Virus 1 Infection, Coronavirus Disease – COVID-19), Signaling Molecules and Interaction, Signal Transduction	(Bastard et al. 2020; Zhang et al. 2020b)
LZTFL1	3p21.31	669	≤ 2	No Hits	(Pairo-Castineira et al. 2021; Ellinghaus et al. 2020; COVID-19 Host Genetics Initiative 2021)
SLC6A20	3p21.31	708	≤ 2	No Hits	(Ellinghaus et al. 2020; COVID-19 Host Genetics Initiative 2021)
TICAM1 [3]	19p13.3	790	9	Cancer, Cardiovascular Disease, Cell Growth and Death, Immune System, Infectious Disease: Viral (Hepatitis B, Hepatitis C, Influenza A, Herpes Simplex Virus 1 Infection, Influenza A), Signal Transduction	(Bastard et al. 2020; Zhang et al. 2020b)
MAT2B	5q34	881	NA	Metabolic Pathways	(Pairo-Castineira et al. 2021)
CFL1	11q13.1	926	NA	Celular Processes, Immune System, Infectious disease: bacterial (Pertussis), Infectious Disease: Viral (Human Immunodeficiency Virus 1 Infection)	(Godri Pollitt et al. 2020)
PLSCR1	3q24	929	≤ 2	No Hits	(Kousathanas et al. 2022)
ApoE	19q13.32	995	2.3–2.4	Digestive System (Cholesterol Metabolism)	(Kuo et al. 2020; Lu et al. 2020)
TNFSF15	9q32	1230	NA	Signaling Molecules and Interaction	(Pairo-Castineira et al. 2021)
IKBKG	Xq28	1294	NA	Cancer, Cardiovascular Disease, Immune System, Infectious Disease: Viral (Human cytomegalovirus infection, Hepatitis B, Hepatitis C, Measles, Human T-cell leukemia virus 1 infection, Kaposi sarcoma-associated herpesvirus infection, Human Immunodeficiency Virus 1 Infection, Herpes Simplex Virus 1 Infection, Epstein-Barr virus infection), Signal Transduction	(Bastard et al. 2020; Zhang et al. 2020b)

Table 5 (continued)

Gene	Chr. location	MAIC [1]	Risk estimate [2]	Biological Pathways	References
DPP4	2q24.2	1358	NA	Digestive System	(Amati et al. 2020; Latini et al. 2020)
TYK2	19p13.2	1435	≤ 2	Cell growth and death, Immune System, Infectious disease: parasitic (Toxoplasmosis), Infectious Disease: Viral (Hepatitis B, Hepatitis C, Influenza A, Measles, Human papillomavirus infection, Kaposi sarcoma-associated herpesvirus infection, Epstein-Barr virus infection, Herpes Simplex Virus 1 Infection, Coronavirus Disease – COVID-19), Signal transduction	(Pairo-Castineira et al. 2021; COVID-19 Host Genetics Initiative 2021)
BSG	19p13.3	1446	NA	No Hits	(Latini et al. 2020)
ABO	9q34.2	1517	≤ 2	Cancer, Metabolic Pathways Including Digestive System (Cholesterol Metabolism)	(Ellinghaus et al. 2020; COVID-19 Host Genetics Initiative 2021)
CCR2	3p21.31	1818	NA	Immune System, Signaling Molecules and Interaction	(Pairo-Castineira et al. 2021)
HLA-C	6p21.33	NR	3.5	Cancer, Cardiovascular disease, Celular Processes, Endocrine and metabolic disease, Immune disease, Immune System, Infectious Disease: Viral (Epstein-Barr virus infection, Herpes Simplex Virus 1, Human cytomegalovirus infection, Human Immunodeficiency Virus 1 Infection, Human papillomavirus infection, Human T-cell leukemia virus 1 infection, Kaposi sarcoma-associated herpesvirus infection), Signaling Molecules and Interaction	(Weiner 3rd et al. 2021)
TMPRSS2	21q22.3	NR	≤ 2	Cancer, Infectious Disease: Viral (Influenza A, Coronavirus Disease–COVID-19), Transport and Catabolism	(Amati et al. 2020; Latini et al. 2020; Torre-Fuentes et al. 2021; Suh et al. 2022)
IFITM3	11p15.5	NR	≤ 2	No Hits	(Zhang et al. 2020c; Goméz et al. 2021; Suh et al. 2022)
EMILIN1	2p23.3	NR	NA	No Hits	(Amati et al. 2020)
EMILIN2	18p11.32-p11.31	NR	NA	No Hits	(Amati et al. 2020)
MMRN1	4q22.1	NR	NA	No Hits	(Amati et al. 2020)
MMRN2	10q23.2	NR	NA	No Hits	(Amati et al. 2020)
GAPDH	12p13.31	NR	NA	Infectious Disease: Bacteria (Pathogenic Escherichia coli infection, Salmonella infection), Metabolic Pathways	(Amati et al. 2020)
NEDD4	15q21.3	NR	NA	Cellular Processes, Genetic Information Processing, Infectious Disease: Viral (Epstein-Barr virus infection), Organismal Systems	(Novelli et al. 2021)
WWP1	8q21.3	NR	NA	Cellular Processes, Genetic Information Processing,	(Novelli et al. 2021)
TNFRSF13C	22q13.2	NR	12.3	Immune disease, Immune system, Infectious disease: viral (Human T-cell leukemia virus 1 infection), Signaling molecules and interaction, Signal transduction	(Russo et al. 2021)

Table 5 (continued)

Gene	Chr. location	MAIC [1]	Risk estimate [2]	Biological Pathways	References
BCL11A	2p16.1	NR	≤ 2	No Hits	(Kousathanas et al. 2022)
FUT2	19q13.33	NR	≤ 2	Metabolism	(Kousathanas et al. 2022)

Chr. loc. chromosome location; *NA* not available; *NR* not reported in top 2000 MAIC score rankings; *OR* odds ratio

[1]MAIC score rank, within a range of 2000, version April 14, 2021 (https://baillielab.net/maic/covid19, accessed June 18, 2021)

[2]Risk estimates–Odds ratio (OR)

[3]TLR3, UNC93B1, TICAM1, TBK1, IRF3, IRF7, IFNAR1, IFNAR2 (autosomal-dominant model–ADM) with risk estimates (odds ratios) of 9 (Zhang et al. 2020b)

[4]IRF7, IFNAR1 (autosomal-recessive model–ARM) with risk estimates (odds ratios) > 50 (Zhang Q et al. 2020b)

the extreme polymorphic nature and extraordinary variation in the geographic distribution of the alleles, in silico approaches can contribute to the overall strategy of identifying target HLA polymorphisms for CGAS. As a best option for CGAS, *HLA* variants should be analyzed at the allelic level (several linked SNVs) and not at the SNV level. Nguyen and collaborators, in a comprehensive in silico analysis of the binding affinity of SARS-CoV-2 peptides across multiple HLA genotypes, identified HLA-B*46:01 as having the fewest predicted binding peptides for SARS-CoV-2. Based on this finding, the authors suggested its association with higher vulnerability to COVID19. On the contrary, HLA-B*15:03, having the greatest capacity to present highly conserved SARS-CoV-2 peptides that are shared among common human coronaviruses, could suggest a potential cross-protective T cell immunity (Nguyen et al. 2020). Recognizably, large, and reproducible studies are needed to circumvent the polymorphic nature of these genes. To surpass these shortcomings, Weiner 3rd and collaborators using *HLA* sequencing through a combined a multicentric approach with individuals from several countries, included covariates aimed at excluding potential confounding effects, then tested the association results by meta-analysing data from prior genome-wide association studies (GWAS) and combined the analysis with results of in silico epitope modeling (Weiner 3rd et al. 2021). These authors identified HLA-C*04:01 as a risk factor for severe COVID-19 with an OR of 3.5 (Weiner 3rd et al. 2021).

The ACE2 SARS-CoV-2 receptor

Resistance to infections has been associated with the lack of a receptor for the pathogen (Casanova and Abel 2020, 2021). The human ACE2 protein is the recognized receptor for SARS-CoV-2 in human cells. *ACE2* SNVs, have been suggested as potential evidence for a genetic basis underlying variable geographic distributions of COVID19 in human populations**.** Using a structural modeling approach, Hussain and collaborators investigated how the binding of the SARS-CoV-2 spike protein with proteins encoded by different human *ACE2* allelic variants could contribute to

the susceptibility and/ or resistance against the viral infection (Hussain et al. 2020). The authors identified two *ACE2* alleles that may provide potential resistance to SARS-CoV-2 infection, rs73635825 (S19P) and rs143936283 (E329G). In another study, based on the analysis of a large WES and molecular dynamic analysis simulations of protein structural changes caused by a small number of selected missense variants, Benetti and colleagues concluded that ACE2 is one of the main molecules whose genetic heterogeneity can modulate infection and disease progression (Benetti et al. 2020). Other studies also identified *ACE2* variants affecting the binding affinities for the S protein, on the premise that these could influence COVID-19 susceptibility (Li et al. 2020b; Suryamohan et al. 2021). Interestingly, in a recent review on the SNVs affecting COVID-19 severity, *ACE2*, along with *TMPRSS2*, a protease coding gene, and *IFITM3*, an interferon induced antiviral protein gene, were the genes most mentioned as related to SARS-CoV-2 infection (Suh et al. 2022).

Candidate gene approaches

Relying on our current knowledge of the immune response to SARS-COV-2 and examples offered by other human viral infections can guide researchers towards hypothesis driven research on the genetic determinants of COVID-19.

The genes that encode the proteins involved in SARS-CoV-2 entry into the host cells are major targets. The prominent role of proteases during the initial stages of SARS-CoV-2 cell entry, points to the endosomal and non-endosomal proteolytic pathways. Based on our current understanding, candidate genes encoding proteases include the *CTSL*, *TMPRSS2*, *PCSK3/ FURIN*, and *HAT* genes. Considering the critical function of the respiratory tract in the initial contact with the virus, Novelli and collaborators performed a pilot-study aimed to verify, in nasopharyngeal and oropharyngeal samples, the in situ expression of predicted SARS-CoV-2 host invasion genes (*ACE2, TMPRSS2, PCSK3/ FURIN, EMILIN1, EMILIN2, MMRN1, MMRN2,*

DPP4, GAPDH, RPLP0, ACTB) (Amati et al 2020). These genes were subsequently sequenced; identifying genetic variants associated with COVID-19 in the *TMPRSS2 and PCSK3/FURIN* genes (Latini et al. 2020). In a similar sequential approach joining in situ expression of candidate genes, sequencing and functional studies, but directed to a later phase of RNA viruses (cell egression), this group identified rare germline activating variants associated with severe COVID-19 cases in the *NEDD4* and *WWP1* genes of the HECT-E3 ligase family, (Novelli et al. 2021).

In the initial stages of viral infection, hyperactivation of cofilin and inefficient actin polymerization is known to occur (Godri Pollitt et al. 2020). The possible implications in the clinical phenotype of COVID-19 has been suggested for variants in the *DSTN* gene, which encodes the actin-depolymerizing factor (ADF), and in the genes encoding cofilin, *CFL1* and *CFL2*.

The IFN-I response is pivotal to the infectious process of the influenza virus as well as it appears to be for the SARS-CoV-2 infection. The signaling cascades involved, from the recognition of the viral particles by host PRRs to the transcription of ISGs, contain an abundance of candidate genes. Notably the intracellular Toll-like receptors (TLRs), such as TLR3, TLR7, TLR8, and TLR9, are particularly relevant for viral recognition as supported by clinical genetics data and the evidence of strong selective pressure exerted on these genes (Barreiro et al. 2009). Moreover, acting on the premise that influenza virus infection is a possible co-factor of severe COVID-19, and that a deficient IFN-I response appears associated with severe disease in both of these viral infections (Ciancanelli et al. 2015; Chen et al. 2018), it appears relevant to observe the similarities in the immune response to the viruses and explore the possibility of genetic determinants in the signaling cross-talk.

In one study, rare putative loss-of-function variants, recently identified in the *TLR7*, associated with impaired type I and II IFN responses of 4 young male patients with severe COVID19 (Van Der Made et al. 2020).

The most compelling understanding for the successful application of a CGAS approach to the task of identifying reproducible disease gene associations in COVID-19, however, resulted from a recent study by Zhang and collaborators on the IFN pathway genes (Zhang et al. 2020b). A NGS technology was used to investigate 13 candidate genes implicated in life-threatening influenza and other viral illnesses (*TLR3*, *IRF7*, and *IRF9*) and core genes conferring influenza susceptibility (*TICAM1/ TRIF*, *UNC93B1*, *TRAF3*, *TBK1*, *IRF3*, and *NEMO/ IKBKG* from the TLR3-dependent type I IFN induction pathway, and *IFNAR1*, *IFNAR2*, *STAT1*, and *STAT2* from the IRF7- and IRF9-dependent type I IFN amplification pathway) (Zhang et al. 2020b). From this study it appeared that rare coding variants of 8 of these candidate genes (*TLR3*, *UNC93B1*, *TICAM1*, *TBK1*, *IRF3*, *IRF7*, *IFNAR1*, *IFNAR2*)

accounted for about 3% of critical COVID-19 cases, with risk estimates (odds ratios) of 9 (Table 5). These results evidenced the essential role of type I IFNs in the protective immune response against SARS-CoV-2 infection. They pointed to the possible role of the > 400 other candidate type I IFN-related genes. Finally, they incited investigations hypothesizing that neutralizing auto-antibodies (auto-Abs) against type I IFNs might also underlie life-threatening COVID-19 pneumonia. A subsequent study led to the demonstration that auto-Abs against type I IFNs account for at least 10% of critical cases (Bastard et al. 2020).

Downstream from the recognition of viral particles by the PRR, innate response cytokines determine the clinical course of the disease. An exacerbated pro-inflammatory cytokine response may lead to the development of ARDS. These observations and many others provide evidence of the importance of the PRRs, the specific cytokines and their pathways including their receptors and effectors, relating to a myriad of potential candidate genes for genetic association studies.

In COVID-19, both the intrinsic and extrinsic pathway of cell death by apoptosis are activated during later stages of the disease. These are highly dependent on the activity of cysteine proteases of the caspase group. The investigation of these proteases, as well as the CTSs, would also be warranted due to the roles that these proteins play in regulating cell death mechanisms (Chwieralski et al. 2006; Minarowska et al. 2007; Schrader et al. 2010; Repnik et al. 2012; De Castro et al. 2016).

Genome-wide approaches

Genome-wide approaches are relevant to the identification of common variants associated with disease. Although the identified variants may be common in the population and their allelic variability confers only a small but detectable effect on the disease phenotype, they can underline the key molecular pathways implicated in severe disease. Stoeger and Amaral, revising candidate genes in COVID-19 publications, flag that research into human protein-coding genes is disproportionately skewed towards a comparably small set of genes, and that genes that are identified by genome-wide datasets, and hence likely to have biological significance in the context of COVID-19, are at a risk of remaining ignored by researchers (Stoeger and Amaral 2020).

GWAS approaches as a source of candidate genes

In collaboration with the Covid-19 Host Genetics Initiative, the Severe Covid-19 GWAS Group, in a large GWAS, identified, in Spanish and Italian populations, two genomic *loci* associated with severe Covid-19 (Ellinghaus et al. 2020). In the study, severe Covid-19 was defined as hospitalization

with respiratory failure and a confirmed SARS-CoV-2 viral RNA polymerase-chain-reaction (PCR) test. The control group was mainly comprised of persons with unknown SARS-CoV-2 infection status, whereas a small number of participants had evidence of the development of anti–SARS-CoV-2 antibodies. Resulting from the choice of the clinical phenotypes, this investigation provided targets associated with COVID-19 respiratory failure as well as COVID-19 susceptibility per se. Candidate genes were identified from the *ABO* blood group locus at 9q34.2 and from a multigene locus at 3p21.31. A blood-group–specific analysis showed a higher risk in blood group A than in other blood groups and a protective effect in blood group O. This association has also been reported in other studies (Golinelli et al. 2020; Zhao J et al. 2021). Additionally, at locus 3p21.31, six candidate genes were identified (*SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6*, and *XCR1*). These findings grounded new directions, and represented a major leap, in identifying the molecular mechanisms governing the pathophysiology of COVID-19 and the determinants of severe disease. The identified association with the chromosomal region 3p21.31 was replicated in association to critical COVID-19 disease in the GWAS performed by Pairo-Castineira and collaborators (Pairo-Castineira et al. 2021).

A GWAS approach also identified novel candidate genes and risk variants by examining deaths in participants in UK Biobank testing positive for the SARS-CoV-2. These genes include *ALOXE3, TMEM181, BRF2, ERAP2*, and *ApoE* (Kuo et al. 2020; Lu et al. 2020).

A targeted analysis of whole-exome sequencing (WES) data evidenced a mutational enrichment of common variable immunodeficiency (CVID) associated genes in severe COVID-19 patients, in the *TNFRSF13C* gene, encoding the B cell-activating factor receptor (OR = 12.3) (Russo et al. 2021).

Genome-wide CRISPR-Cas9 knockout screens as a source of candidate genes

Genome-wide CRISPR-Cas9 knockout screens contribute to the elucidation of genotype to phenotype relationships by ablating gene expression on a genome-wide scale and studying the resulting phenotypic alterations. Several studies have used CRISPR loss-of-function screens to generate functional catalogues of host factors required for infection by SARS-CoV-2 (Baggen et al. 2021; Daniloski et al. 2021; Schneider et al. 2021; Wang et al. 2021; Wei et al. 2021; Zhu et al. 2021). These studies have highlighted the critical pathways and the importance of multiple genes within each pathway. Notable candidate host factors include genes involved in phosphatidylinositol biosynthesis, cholesterol homeostasis, regulation of intracellular transport, regulation of autophagy, and endosomal trafficking.

Transcriptomics

Transcriptomics and expression quantitative trait loci (eQTLs), that associate genomic and transcriptomic data sets are observations that present additional opportunity for candidate gene discovery for associations (Nica and Dermitzakis 2013; Pairo-Castineira et al. 2021).

The variation in *ACE2* expression level may be associated to susceptibility and clinical phenotype of SARS-CoV infection (Hofmann et al. 2005). The finding however remains controversial (Butler-Laporte et al. 2021). Based on this premise Cao and collaborators, using the GTEx database for eQTL analysis, revealed variable *ACE2* allele frequencies across different populations (Cao et al. 2020). The effect of ACE2 polymorphisms and expression levels on increased morbidity and lethality of COVID-19 in males has also been speculatively associated with its location on the X chromosome (Liao et al. 2005; Gadi et al. 2020; Devaux et al. 2020; Van Der Made et al. 2020). In a genome-wide eQTL analysis, based on high-throughput RNA sequencing, Ansari and collaborators discovered that polymorphisms in the interferon lambda (IFNL) region are associated with *ACE2* expression, suggesting the negative correlation between the interferon response and *ACE2* expression (Ansari et al. 2020). This result could also suggest a possible epistatic relationship between the genetic variants in these genes that also warrants further exploration.

TMPRSS2 levels and genetic variants proved to be possible candidate disease modulators, contributing to the observed epidemiological data among Italian patients (Asselta et al. 2020).[149] A haplotype characterized by 3 SNVs (rs2070788, rs9974589, rs7364083), whose MAF is significantly increased in Europeans in respect to East Asians, was predicted to be associated with higher *TMPRSS2* expression. The rs2070788 SNV has also been associated with susceptibility to influenza (Cheng et al. 2015). Recently a common *TMPRSS2* SNV (rs12329760) was associated with a reduced likelihood of developing severe COVID-19 was identified (OR 0.87) (David et al. 2022).

Meta-analysis of genome-wide overlap

Meta-analysis of genome-wide association studies are resulting from the plethora of data generated since the onset of the pandemic. In particular, two of the recent publications are central to MAGO applications in the validation of candidate genes in COVID-19 (Parkinson et al. 2020; Pairo-Castineira et al. 2021). These include an application of the meta-analysis by information content (MAIC) algorithm (Parkinson et al. 2020), and a GWAS followed by multiple genome-wide replication approaches (Pairo-Castineira et al. 2021).

1) Meta-analysis by information content (MAIC)

The MAIC algorithm was originally developed to identify host genes necessary for Influenza A virus (IAV) replication (Li et al. 2020c). Parkinson and collaborators applied MAIC to identify potential therapeutic targets in COVID-19 (Parkinson et al. 2020). In this approach, it is assumed that (1) there exists a set of true positives: host genes involved in COVID-19 pathogenesis; (2) a gene is more likely to be a true positive if it is found in multiple experiments; (3) a gene is more likely to be a true positive if it occurs in a list containing a higher proportion of genes with supporting evidence from multiple sources; (4) due to experimental biases, the evidence that a gene is a true positive is further increased if it is found across experimental types. With these assumptions, MAIC allows the quantification of the information content in a gene list by comparing that list to the results from other experiments, both within and between methodologies, producing a weighting factor for each experiment used to calculate a score for each gene.

For the COVID-19 MAIC application a systematic literature search of PubMed was carried out for: ("Coronavirus" OR "Severe Acute Respiratory Syndrome" OR "Middle East Respiratory Syndrome" OR "Sars-CoV-2" OR "COVID-19") AND ((gene*.[Title/Abstract]) OR (genom*[Title/Abstract]) OR (transcript*[Title/Abstract]) OR (protein*[Title/Abstract]) OR ("Susceptibility"[Title/Abstract]) OR (siRNA[All Fields])). The methodological types included were CRISPR screens, RNAi screens, Protein–protein interaction e.g. yeast-2-hybrid screens, Host proteins incorporated into virion or virus like particle, Genetic Association Studies, Transcriptomics e.g. RNA-Seq, scRNA-Seq, Proteomic studies e.g. mass-spectrometry, and Selected gene set screens. Candidate-gene human genetic studies were excluded. A final ranked list of genes based on the MAIC score was provided and updated results are made available at https://baillielab.net/maic/covid19 (accessed June 18, 2021). MAIC score distributions of top enriched pathways from KEGG 2019 (Human) Cytokine-cytokine receptor interaction, Influenza A, Hepatitis C, IL-17 signaling pathway, Toll-like receptor signaling pathway, T-cell receptor signaling pathway, MAPK signaling pathway, Apoptosis, Protein processing in endoplasmic reticulum, Antigen processing and presentation. From WikiPathways 2019 (Human), these included Senescence and autophagy in cancer, Toll-like receptor signaling pathway, T-cell antigen receptor signaling pathway, Human complement system, VEGFA-VEGFE2 signaling pathway, Glycolysis and gluconeogenesis, TSLP signaling pathway, IL-3 signaling pathway, Fas ligand pathway/HSP regulation, MAPK signaling pathway.

The MAIC score rank of some candidate genes from references cited in this review are given in Table 5. Some of these genes are currently recognized as belonging to the Coronavirus disease—COVID-19 viral infectious disease pathway included in the KEGG PATHWAY Database (Kanehisa and Goto 2000). (https://www.genome.jp/kegg/pathway.html, accessed July 1, 2021), such as *IFNAR2, TBK1, TLR3, ACE2, IRF3, IFNAR1, TYK2*. Several of these genes are also implicated in the pathways of other viral infectious diseases such as Herpes simplex virus 1 infection, Hepatitis B, Hepatitis C, the Human immunodeficiency virus 1, and Influenza A. A large number of these genes code for proteins that act upon the immune system and some are implicated in signaling and interaction, and signal transduction.

The MAIC score rankings of a gene appear to vary considerably from one observation to another of the dataset. There is yet no indication on the possibility of a stabilization of the gene rankings as the number of integrated studies increases. This advocates the need for further development of this or other meta-analytical algorithms. Moreover, a failing of the MAIC is that it misses important functional studies (Novelli et al. 2021).

2) GWAS followed by multiple genome-wide replication approaches.

In the study presented by Pairo-Castineira and collaborators, MAGO was applied to results replicated from GWAS, Mendelian randomization of gene expression, Transcriptome-wide association study (TWAS) and MAIC analysis (Pairo-Castineira et al. 2021). The GWAS was performed in cases of European descent from GenOMICC and a control population from the UK Biobank. Results from the primary analysis were replicated, in cases and controls, in a meta-analysis of data from other data sources (the COVID-19 Host Genetics Initiative 23andMe Inc). Meta-analysis identified overlapping SNVs from six genomic loci: *LZTFL1, CCHCR1, OAS1–OAS3, DPP9, TYK2, IFNAR2*. Mendelian randomization of gene expression evidenced a causal relationship on the odds of patients developing severe COVID-19: 1) of genes selected a priori as potential therapeutic targets for COVID-19 and; 2) transcriptome-wide. Results were significant for two genes *TYK2, IFNAR2*. TWAS was used to link GWAS results to tissue-specific gene expression data by inferring gene expression from known eQTL in whole blood and lung samples from the GTEx v.8 data base (https://gtexportal.org/home/). A meta-TWAS analysis was performed (Barbeira et al. 2018). The genes identified with significant differences in predicted expression compared to control individuals were *CXCR6, OAS3, CCR3, MAT2B, CCR2, TNFSF15, ICAM5, FYCO1*. The MAIC algorithm was used to analyze these results according to

the data-driven gene list weightings to produce a comprehensive ranked list of host genes associated with COVID-19 (Parkinson et al. 2020). TWAS revealed the greater overlap of results as compared to that of other MAIC data sources of host genes implicated in COVID-19.

A GenOMICC study was again performed in a more recent whole genome sequencing approach. Critical COVID-19-associated variants were identified, including in genes involved in interferon signalling (*IL10RB*, *PLSCR1*), leucocyte differentiation (*BCL11A*), and blood type antigen secretor status (*FUT2*) (Kousathanas et al. 2022). Moreover, in the subsequent TWAS study, using gene expression data for disease-relevant tissues, lung and whole blood, the authors also found evidence implicating multiple genes in critical COVID-19.

3) Meta-analysis of GWAS.

The COVID-19 Host Genetics Initiative network recently published the results of three genome-wide association meta-analyses comprising up to 49,562 COVID-19 patients from 46 studies across 19 countries worldwide (COVID-19 Host Genetics Initiative 2021). Here, conventional case–control meta-analyses was performed in three main categories of COVID-19 disease phenotype according to predefined and partially overlapping phenotypic criteria: (1) critically ill COVID-19 cases defined as those who required in hospital respiratory support or who were deceased due to the disease, (2) cases with moderate or severe COVID-19 defined as those hospitalized due to symptoms associated with the infection, and (3) all cases with reported SARS-CoV-2 infection with or without symptoms of any severity. Controls for all three analyses were selected as genetic ancestry matched samples generally without known SARS-CoV-2 infection. Thirteen genome-wide significant loci associated with SARS-CoV-2 infection or severe COVID-19 were reported. Four loci were associated to susceptibility to SARS-CoV-2 infection, but not with the progression to more severe COVID-19 phenotypes. These included the previously reported *ABO* locus. In contrast, the remaining 9 loci were associated with increased risk of severe symptoms, many of which did not ranked within the MAIC top 2000 scores (https://baillielab.net/maic/covid19/, version April 14, 2021). Smoking and body mass index were identified as causal risk factors for severe COVID-19.

Interactome studies

Along with GWAS, candidate gene and transcriptome approaches, interactome studies of physical interactions among molecules such as protein–protein interactions (PPIs)

between SARS-CoV-2 and human cells can reveal underlying mechanisms of pathogenesis and assist in development of antiviral therapies. In his article, Lee describes the construct of protein sequence-based multi-label classifiers to predict virus-human PPIs with different evidence or confidence levels for SARS-CoV-2 infection (Lee 2021). The model also predicts evidence levels for all PPI pairs between the SARS-CoV-2 and human proteomes to provide a draft virus-host interactome landscape of SARS-CoV-2 infection in humans in silico. Finally, it prioritizes those virus-human PPIs or sub-networks of high evidence for biological relevance and therapeutic opportunities for COVID-19 in future studies. Preliminary observations concluded that most human proteins appear to interact with SARS-CoV-2 Nsp7, Nsp1, and ORF14, with significant enrichment in the top 2 pathways of vascular smooth muscle contraction (CALD1, NPR2, CALML3) and Myc targets (CBX3, PES1). It is also suggested that histone H2A components are targeted by multiple SARS-CoV-2 proteins (Lee 2021).

Conclusion

The overwhelming impact that COVID-19 has had on Public Health worldwide acknowledges the importance of pursuing multidisciplinary research efforts needed to understand the pathogen and control the disease.

Slightly over 2 years after the first cases of infection were reported in Wuhan, China, the new SARS-CoV-2 coronavirus, cause of an entirely new disease COVID-19, had infected over 500 million people, and directly resulted in the loss of 6 million human lives. The COVID-19 fatality rate exceeds that of malaria, HIV/AIDS, and influenza, and within the first year of the pandemic, the yearly fatality rate of rapidly approached that of the deadliest human infectious disease, TB. Thus, the SARS-CoV-2 pandemic has had an indirect negative impact on other areas of Public Health, causing a global setback in the 2030 Sustainable Development Goals, and although several vaccines have been approved for public use, it is apparent that the struggle to control the COVID-19 pandemic will not subside for some time to come.

Throughout human history, deadly infections have shaped our immune system. However, it is still naïve, at least partially, in relation to the new virus SARS-CoV-2. SARS-CoV-2 has gradually acquired several mutations that determined the emergence of new variants characterized by different phenotypes (e.g., with changes in transmissibility, pathogenicity and antigenicity or immune escape ability). To date, however, no specific mutations have been found in SARS-CoV-2 that could explain the extraordinary range in the clinical forms of COVID-19. It is therefore likely that, in human beings, a genetic basis could govern severe COVID-19. Nevertheless, it remains of utmost importance

to routinely monitor SARS-CoV-2 mutations and investigate the relationship between viral genetic variants and the clinical features/ disease phenotype through sequence-based surveillance, epidemiological studies, and clinical and research laboratory investigations (CDC 2021; Bakhshandeh et al. 2021; Duarte et al. 2021; Zhou and Wang 2021).

Currently, the scientific community is making important efforts to elucidate the molecular mechanisms of the pathophysiology of SARS-CoV-2 infection and as more results from genome-wide approaches are available through the ongoing global collaborative efforts; these will undeniably identify target genes, genomic regions and molecular pathways that will contribute to the current understanding of the genetic determinants of COVID-19. A greater awareness of the importance of candidate genes identified by genome-wide datasets is viewed as an imperative for the future success of these efforts. In the current theory of human infection genetics, a homogeneous physiological basis in the molecular pathways that govern the genetics of serious forms of COVID-19 is to be expected. The identification of genes associated with disease severity can flag the main molecular pathways involved, providing an abundance of information from which to select candidate genes for future studies and therapeutic developments. However, the moderate effect sizes of associated common variants from genome-wide approaches make it difficult to establish a functional basis. Therefore, there is an increased need for replication strategies by which to infer causality. The strategic use of meta-analysis of genome-wide overlap, such as the MAIC, or other meta-analytical algorithms still need further development as tools for candidate gene selection and validation.

Except for the interferon circuit genes, there is currently insufficient data, to identify risk and/or resistance profiles for screening with practical implications on Public Health. The combined knowledge of these efforts is at the basis of scientific developments for the future control of COVID-19. Ultimately, such knowledge has the potential to prompt the identification of promising viral- or host-directed targets and to establish an integrated COVID-19 therapeutic discovery and development program (Stevens 2004; Hingorani et al. 2019; Seyhan 2019; Trigueiro-Louro et al. 2019, 2020; Parkinson et al. 2020; CDC 2021).

References

Akira S, Uematsu S, Takeuchi O (2006) Pathogen recognition and innate immunity. Cell 124(4):783–801

Alangari AA, Al-Zamil F, Al-Mazrou A et al (2011) Treatment of disseminated mycobacterial infection with high-dose IFN-γ in a patient with IL-12Rβ1 deficiency. Clin Dev Immunol 2011:691956.

Amati F, Vancheri C, Latini A, Colona VL, Grelli S, D'Apice MR et al (2020) Expression profiles of the SARS-CoV-2 host invasion genes in nasopharyngeal and oropharyngeal swabs of COVID-19 patients. Heliyon 6(10):e05143

Anastassopoulou C, Gkizarioti Z, Patrinos GP, Tsakris A (2020) Human genetic factors associated with susceptibility to SARS-CoV-2 infection and COVID-19 disease severity. Hum Genomics 14(1):1–8

Ansari MA, Marchi E, Ramamurthy N et al (2020) Negative regulation of ACE2 by interferons in vivo and its genetic control medRxiv. https://doi.org/10.1101/2020.04.26.20080408

Asselta R, Paraboschi EM, 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

Aykac K, Cura Yayla BC, Ozsurekci Y et al (2021) The association of viral load and disease severity in children with COVID-19. J Med Virol 93(5):3077–3083

Bai Y, Yao L, Wei T et al (2020) Presumed asymptomatic carrier transmission of COVID-19. JAMA 323(14):1406–1407

Baillie JK (2014) Targeting the host immune response to fight infection. Science 344:807–808

Baggen J, Persoons L, Vanstreels E et al (2021) Genome-wide CRISPR screening identifies TMEM106B as a proviral host factor for SARS-CoV-2. Nat Genet 53:435–444

Bakhshandeh B, Jahanafrooz Z, Abbasi A et al (2021) Mutations in SARS-CoV-2; Consequences in structure function and pathogenicity of the virus. Microb Pathog 154:104831

Barbeira AN, Dickinson SP, Bonazzola R et al (2018) Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat Commun 9(1):1–20

Barreiro LB, Ben-Ali M, Quach H et al (2009) Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense. PLoS Genet 5(7):e1000562

Bastard P, Rosen LB, Zhang Q et al (2020) Autoantibodies against type I IFNs in patients with life-threatening COVID-19. Science 370(6515):eabd4585

Belouzard S, Millet JK, Licitra BN, Whittaker GR (2012) Mechanisms of coronavirus cell entry mediated by the viral spike protein. Viruses 4(6):1011–1033

Benetti E, Tita R, Spiga O et al (2020) ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population. Eur J Hum Genet 28(11):1602–1614

Bojkova D, McGreig JE et al (2020) SARS-CoV-2 and SARS-CoV differ in their cell tropism and drug sensitivity profiles. BioRxiv. https://doi.org/10.1101/2020.04.03.024257

Butler-Laporte G, Nakanishi T, Mooser V et al (2021) The effect of angiotensin-converting enzyme levels on Covid-19 susceptibility and severity: a Mendelian randomization study. Int J Epidemiol 50(1):75–86

Burke RM, Midgley CM, Dratch A et al (2020) Active monitoring of persons exposed to patients with confirmed COVID-19—United States, January–February 2020. Morb Mortal Wkly Rep 69(9):245

Cao Y, Li L, Feng Z et al (2020) Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. Cell Disc 6(1):1–4

Carter-Timofte ME, Jørgensen SE, Freytag MR et al (2020) Deciphering the role of host genetics in susceptibility to severe COVID-19. Front Immunol 11:1606

Casanova JL, Conley ME, Seligman SJ et al (2014) Guidelines for genetic studies in single patients: lessons from primary immunodeficiencies. J Exp Med 211(11):2137–2149

Casanova JL, Abel L (2020) The human genetic determinism of life-threatening infectious diseases: genetic heterogeneity and physiological homogeneity? Hum Genet 139:681–694

Casanova JL, Su HC, Abel L et al (2020) A global effort to define the human genetics of protective immunity to SARS-CoV-2 infection. Cell 181(6):1194–1199

Casanova JL, Abel L (2021) Lethal infectious diseases as inborn errors of immunity: toward a synthesis of the germ and genetic theories. Ann Rev Pathol 16:23–50

Centers for Disease Control and Prevention, CDC (2021). SARS-CoV-2 variant classifications and definitions. https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html#Interest. Accessed 19 Nov 2021

Ciancanelli MJ, Huang SX, Luthra P et al (2015) Life-threatening influenza and impaired interferon amplification in human IRF7 deficiency. Science 348(6233):448–453

Chan RW, Chan MC, Agnihothram S et al (2013) Tropism of and innate immune responses to the novel human betacoronavirus lineage C virus in human ex vivo respiratory organ cultures. J Virol 87(12):6604–6614

Chan JF, Lau SK, To KK, Cheng VC et al (2015) Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease. Clin Microbiol Rev 28:465–522

Channappanavar R, Perlman S (2017) Pathogenic human coronavirus infections: causes and consequences of cytokine storm and immunopathology. Semin Immunopathol 39(5):529–539

Channappanavar R, Fehr AR, Zheng J et al (2019) IFN-I response timing relative to virus replication determines MERS coronavirus infection outcomes. J Clin Invest 129:3625–3639

Chau NVV, Thanh Lam V, Thanh Dung N et al (2020) The natural history and transmission potential of asymptomatic SARS-CoV-2 infection. Clin Infect Dis 71(10):2679–2687

Chen X, Liu S, Goraya MU et al (2018) Host immune response to influenza A virus infection. Front Immunol 9:320

Chen Y, Liu Q, Guo D (2020) Emerging coronaviruses: genome structure, replication, and pathogenesis. J Med Virol 92(4):418–423

Cheng Z, Zhou J, To KKW et al (2015) Identification of TMPRSS2 as a susceptibility gene for severe 2009 pandemic A (H1N1) influenza and A (H7N9) influenza. J Infect Dis 212(8):1214–1221

Cilloni L, Fu H, Vesga JF et al (2020) The potential impact of the COVID-19 pandemic on the tuberculosis epidemic a modelling analysis. E Clin Med 28:100603

Cho NJ, Glenn JS (2020) Materials science approaches in the development of broad-spectrum antiviral therapies. Nat Mater 19(8):813–816

Chou J, Thomas PG, Randolph AG (2022) Immunology of SARS-CoV-2 infection in children. Nat Immunol 1–9

Chu H, Zhou J, Wong BHY et al (2016) Middle East respiratory syndrome coronavirus efficiently infects human primary T lymphocytes and activates the extrinsic and intrinsic apoptosis pathways. J Infect Dis 213(6):904–914

Chwieralski CE, Welte T, Bühling F (2006) Cathepsin-regulated apoptosis. Apoptosis 11(2):143–149

Coutard B, Valle C, de Lamballerie X et al (2020) The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. Antivir Res 176:104742

Colona VL, Vasiliou V, Watt J, Novelli G, Reichardt JK et al (2021a) Update on human genetic susceptibility to COVID-19: susceptibility to virus and response. Hum Genomics 15(1):1–8

Colona VL, Biancolella M, Novelli A, Novelli G (2021b) Will GWAS eventually allow the identification of genomic biomarkers for COVID-19 severity and mortality? J Clin Invest 131(23)

COVID-19 Host Genetics Initiative (2021) Mapping the human genetic architecture of COVID-19 by worldwide meta-analysis MedRxiv. https://doi.org/10.1101/2021.03.10.21252820

COVID-19 Host Genetics Initiative (2020) 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(6):715

Dahdouh E, Lázaro-Perona F, Romero-Gómez MP et al (2021) Ct values from SARS-CoV-2 diagnostic PCR assays should not be used as direct estimates of viral load. J Infect 82(3):414–451

Dandekar AA, Perlman S (2005) Immunopathogenesis of coronavirus infections: implications for SARS. Nat Rev Immunol 5(12):917–927

Daniloski Z, Jordan TX, Wessels HH et al (2021) Identification of required host factors for SARS-CoV-2 infection in human cells. Cell 184(1):92–105

David S (2021) A current guide to candidate gene association studies. Trends Genet 37(12):1056–1059

David A, Parkinson N, Peacock TP, Pairo-Castineira E, Khanna T, Cobat A et al (2022) A common TMPRSS2 variant has a protective effect against severe COVID-19. Current research in translational medicine 70(2):103333

Davies NG, Abbott S, Barnard RC et al (2020) Estimated transmissibility and severity of novel SARS-CoV-2 Variant of Concern 202012/01 in England. Science 372(6538)

Debnath M, Banerjee M, Berk M (2020) Genetic gateways to COVID-19 infection: Implications for risk severity and outcomes. FASEB J 34(7):8787–8795

De Castro MAG, Bunt G, Wouters FS (2016) Cathepsin B launches an apoptotic exit effort upon cell death-associated disruption of lysosomes. Cell Death Discov 2(1):1–8

De Wilde AH, Snijder EJ, Kikkert M, van Hemert MJ (2018) Host factors in coronavirus replication. Curr Top Microbiol Immunol 419:1–42

Devaux CA, Rolain JM, Raoult D (2020) ACE2 receptor polymorphism: Susceptibility to SARS-CoV-2 hypertension multi-organ failure and COVID-19 disease outcome. J Microbiol Immunol Infect 53(3):425–435

Diao B, Wang C, Tan Y et al (2020) Reduction and functional exhaustion of T cells in patients with coronavirus disease 2019 (COVID-19). Front Immunol 11:827

Dong E, Du H, Gardner L (2020) An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 20(5):533–534

Donnelly CA, Malik MR, Elkholy A et al (2019) Worldwide reduction in MERS cases and deaths since 2016. Emerg Infect Dis 25(9):1758

Duarte CM, Jamil T, Gojobori T, Alam I (2021) Detection of SARS-CoV-2 variants requires urgent global coordination Int J Infect Dis 109:50–53

Escobedo A, Kaushal D, Singh K (2022) Insights into the changing landscape of coronavirus disease 2019. Front Cell Infect Microbiol 11

European Centre for Disease Prevention and Control, ECDC (2020) Detection of new SARS-CoV-2 variants related to mink. https://www.ecdc.europa.eu/sites/default/files/documents/RRA-SARS-CoV-2-in-mink-12-nov-2020.pdf

Elhabyan A, El Yaacoub S, Sanad et al (2020) The role of host genetics in susceptibility to severe viral infections in humans and insights into host genetics of severe COVID-19: A systematic review. Virus Res 289:198163

Ellinghaus D, Degenhardt F, Bujanda L et al (2020) Severe Covid-19 GWAS Group Genomewide association study of severe Covid-19 with respiratory failure. N Engl J Med 383(16):1522–1534

European Centers for Disease Control and Prevention, ECDC (2022) SARS-CoV-2 variants of concern as of 17 February 2022. https://www.ecdc.europa.eu/en/covid-19/variants-concern. Accessed 26 Feb 2022

Fajnzylber J, Regan J, Coxen K et al (2020) SARS-CoV-2 viral load is associated with increased disease severity and mortality. Nat Commun 11(1):1–9

Forni G, Mantovani A (2021) COVID-19 vaccines: where we stand and challenges ahead. Cell Death Differ 28(2):626–639

Gadi N, Wu SC, Spihlman AP, Moulton VR (2020) What's sex got to do with COVID-19? Gender-based differences in the host immune response to coronaviruses. Front Immunol 11:2147

Ghinai I, Woods S, Ritger KA et al (2020) Community transmission of SARS-CoV-2 at two family gatherings—Chicago, Illinois, February–March 2020. Morb Mortal Wkly Rep 69(15):446

Golinelli D, Boetto E, Maietti E, Fantini MP (2020) The association between ABO blood group and SARS-CoV-2 infection: A meta-analysis. PLoS One 15(9):e0239508

Godri Pollitt KJ, Peccia J, Ko AI et al (2020) COVID-19 vulnerability: the potential impact of genetic susceptibility and airborne transmission. Hum Genomics 14:1–7

Gómez J, Albaiceta GM, Cuesta-Llavona E, García-Clemente M, López-Larrea C, Amado-Rodríguez L et al (2021) The Interferon-induced transmembrane protein 3 gene (IFITM3) rs12252 C variant is associated with COVID-19. Cytokine 137:155354

Grant WB, Lahore H, McDonnell SL et al (2020) Evidence that vitamin D supplementation could reduce risk of influenza and COVID-19 infections and deaths. Nutrients 12(4):988

Hadjadj J, Yatim N, Barnabei L et al (2020) Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science 369(6504):718–724

Hakonarson H, Gulcher JR, Stefansson K (2003) deCODE genetics Inc. Pharmacogenomics 4(2):209–215

Hamner L (2020) High SARS-CoV-2 attack rate following exposure at a choir practice-Skagit County, Washington, March 2020. MMWR Morbidity and mortality weekly report 69

Hasanoglu I, Korukluoglu G, Asilturk D et al (2021) Higher viral loads in asymptomatic COVID-19 patients might be the invisible part of the iceberg. Infection 49(1):117–126

He X, Hong W, Pan X et al (2021) SARS-CoV-2 Omicron variant: characteristics and prevention. MedComm 2:838–845

He R, Lu Z, Zhang L et al (2020) The clinical course and its correlated immune status in COVID-19 pneumonia. J Clin Virol 127:104361

Hebbring SJ (2014) The challenges, advantages and future of phenomewide association studies. Immunology 141(2):157–165

Heald-Sargent T, Gallagher T (2012) Ready, set, fuse! The coronavirus spike protein and acquisition of fusion competence. Viruses 4(4):557–580

Hingorani AD, Kuan V, Finan C et al (2019) Improving the odds of drug development success through human genomics: modelling study. Sci Rep 9(1):1–25

Hoehl S, Rabenau H, Berger A et al (2020) Evidence of SARS-CoV-2 infection in returning travelers from Wuhan. China N Engl J Med 382(13):1278–1280

Hofmann H, Pyrc K, Van Der Hoek L et al (2005) Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. Proc Natl Acad Sci U S A 102(22):7988–7993

Hoffmann M, Kleine-Weber H, Schroeder S et al (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

Holland SM (2001) Immunotherapy of mycobacterial infections. Semin Respir Infect 16(1):47–59

Hussain M, Jabeen N, Raza F et al (2020) Structural variations in human ACE2 may influence its binding with SARS-CoV-2 spike protein. J Med Virol 92(9):1580–1586

Huang C, Wang Y, Li X et al (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan. China Lancet 395(10223):497–506

Hui KP, Cheung MC, Perera RA et al (2020) Tropism, replication competence, and innate immune responses of the coronavirus SARS-CoV-2 in human respiratory tract and conjunctiva: an analysis in ex-vivo and in-vitro cultures. Lancet Resp Med 8(7):687–695

Jewell BL, Mudimu E, Stover J et al (2020) Potential effects of disruption to HIV programmes in sub-Saharan Africa caused by COVID-19: results from multiple mathematical models. Lancet HIV 7(9):e629–e640

Jiang L, Tang K, Levin M et al (2020) COVID-19 and multisystem inflammatory syndrome in children and adolescents. Lancet Infect Dis 20(11):e276–e288

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28(1):27–30

Karim SSA, Karim QA (2021) Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. Lancet 398(10317):2126–2128

Kato H, Sato S, Yoneyama M et al (2005) Cell type-specific involvement of RIG-I in antiviral response. Immunity 23(1):19–28

Kawai T, Akira S (2008) Toll-like receptor and RIG-1-like receptor signaling. Ann N Y Acad Sci 1143:1–20

King EA, Davis JW, Degner JF (2019) Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. PLoS Genet 15(12):e1008489

Kielian M, Rey FA (2006) Virus membrane-fusion proteins: more than one way to make a hairpin. Nature Rev Microbiol 4(1):67–76

Khateeb J, Li Y, Zhang H (2021) Emerging SARS-CoV-2 variants of concern and potential intervention approaches. Crit Care 25(1):1–8

Koh WC, Naing L, Chaw L et al (2020) What do we know about SARS-CoV-2 transmission? A systematic review and meta-analysis of the secondary attack rate and associated risk factors. PloS one 15(10):e0240205

Kousathanas A, Pairo-Castineira E, Rawlik K, Stuckey A, Odhams CA, Walker S et al (2022) Whole genome sequencing reveals host factors underlying critical Covid-19. Nature 1–10

Kuo CL, Pilling LC, Atkins JL et al (2020) APOE e4 genotype predicts severe COVID-19 in the UK Biobank community cohort. J Gerontol A Biol Sci Med Sci 75(11):2231–2232

Kuo L, Godeke GJ, Raamsman MJ et al (2000) Retargeting of coronavirus by substitution of the spike glycoprotein ectodomain: crossing the host cell species barrier. J Virol 74(3):1393–1406

Latini A, Agolini E, Novelli A, Borgiani P, Giannini R, Gravina P et al (2020) COVID-19 and genetic variants of protein involved in the SARS-CoV-2 entry into the host cells. Genes 11(9):1010

Lauring AS, Hodcroft EB (2021) Genetic variants of SARS-CoV-2—what do they mean? JAMA 325(6):529–531

Lavillegrand JR, Garnier M, Spaeth A et al (2021) Elevated plasma IL-6 and CRP levels are associated with adverse clinical outcomes and death in critically ill SARS-CoV-2 patients: inflammatory response of SARS-CoV-2 patients. Ann Intensive Care 11(1):1–10

Lee HJ (2021) An interactome landscape of SARS-CoV-2 virus-human protein-protein interactions by protein sequence-based multi-label classifiers. bioRxiv

Leopardi S, Holmes EC, Gastaldelli M et al (2018) Interplay between co-divergence and cross-species transmission in the evolutionary history of bat coronaviruses. Infect Genet Evol 58:279–289

Li F (2016) Structure function and evolution of coronavirus spike proteins. Annu Rev Virol 3:237–261

Li X, Geng M, Peng Y et al (2020a) Molecular immune pathogenesis and diagnosis of COVID-19. J Pharm Anal 10(2):102–108

Li Q, Cao Z, Rahman P (2020b) Genetic variability of human angiotensin-converting enzyme 2 (hACE2) among various ethnic populations. Mol Genet Genomic Med 8(8):e1344

Li B, Clohisey SM, Chia BS et al (2020c) Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection. Nat Commun 11(1):1–18

Li Z, Jiang J, Ruan X, Tong Y, Xu S, Han L, Jianguo XuJ (2021) The zoonotic and natural foci characteristics of SARS-CoV-2. Journal of Biosafety and Biosecurity 3(1):51–55

Liao QJ, Ye LB, Timani KA et al (2005) Activation of NF-κB by the full-length nucleocapsid protein of the SARS coronavirus. Acta Biochim Biophys Sin 37(9):607–612

Liu J, Li S, Liu J et al (2020) Longitudinal characteristics of lymphocyte responses and cytokine profiles in the peripheral blood of SARS-CoV-2 infected patients. EBioMedicine 55:102763

Lokugamage KG, Hage A, Schindewolf C et al (2020) Type I interferon susceptibility distinguishes SARS-CoV-2 from SARS-CoV. J Virol 94:e01410-e1420

Lonsdale J, Thomas J, Salvatore M et al (2013) The genotype-tissue expression (GTEx) project. Nat Genet 45(6):580–585

Lopez Bernal J, Andrews N, Gower C et al (2021) Effectiveness of Covid-19 vaccines against the B 1617 2 (Delta) variant. N Engl J Med 385:585–594

LoPresti M, Beck DB, Duggal P et al (2020) The role of host genetic factors in coronavirus susceptibility: review of animal and systematic review of human literature. Am J Hum Genet 107(3):381–402

Luo CH, Morris CP, Sachithanandham J et al (2021) Infection with the SARS-CoV-2 delta variant is associated with higher infectious virus loads compared to the alpha variant in both unvaccinated and vaccinated individuals. medRxiv. https://doi.org/10.1101/2021.08.15.21262077

Luliano AD, Roguski KM, Chang HH et al (2018) Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. Lancet 391(10127):1285–1300

Lu C, Gam R, Pandurangan AP, Gough J (2020) Genetic risk factors for death with SARS-CoV-2 from the UK Biobank. MedRxiv. https://doi.org/10.1101/2020.07.01.20144592

Mahallawi WH, Khabour OF, Zhang Q et al (2018) MERS-CoV infection in humans is associated with a pro-inflammatory Th1 and Th17 cytokine profile. Cytokine 104:8–13

Marks M, Millat-Martinez P, Ouchi D et al (2021) Transmission of COVID-19 in 282 clusters in Catalonia Spain: a cohort study. Lancet Infect Dis 21:629–636

McIntosh K, Hirsch MS, Bloom A (2022) Coronavirus disease 2019 (COVID-19): Epidemiology, virology, and prevention. Lancet Infect Dis 1:2019–20. https://www.uptodate.com/contents/covid-19-epidemiology-virology-and-prevention. Accessed 26 Feb 2022

Meng T, Cao H, Zhang H et al (2020) The insert sequence in SARS-CoV-2 enhances spike protein cleavage by TMPRSS. BioRxiv. https://doi.org/10.1101/2020.02.08.926006

Merad M, Martin JC (2020) Pathological inflammation in patients with COVID-19: a key role for monocytes and macrophages. Nat Rev Immunol 20(6):355–362

Millet JK, Whittaker GR (2014) Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. Proc Natl Acad Sci U S A 111:15214–15219

Minarowska A, Minarowski L, Karwowska A, Gacko M (2007) Regulatory role of cathepsin D in apoptosis. Folia Histochem Cytobiol 45(3):159–163

Moon CY, Schilder, BM, Raj T, Huang KL (2021) Phenome-wide and expression quantitative trait locus associations of coronavirus disease 2019 genetic risk loci. Iscience 24(6):102550

Nica AC, Dermitzakis ET (2013) Expression quantitative trait loci: present and future. Philos Trans R Soc Lond B Biol Sci 368(1620):20120362

Nicoli F, Solis-Soto MT, Paudel D et al (2020) Age-related decline of de novo T cell responsiveness as a cause of COVID-19 severity. Geroscience 42(4):1015–1019

Nguyen A, David JK, Maden SK et al (2020) Human leukocyte antigen susceptibility map for severe acute respiratory syndrome coronavirus 2. J Virol 94(13):e00510-e520

Novelli G, Liu J, Biancolella M, Alonzi T, Novelli A, Patten JJ et al (2021) Inhibition of HECT E3 ligases as potential therapy for COVID-19. Cell Death Dis 12(4):1–18

NSW (2021) Government. Living Evidence - SARS-CoV-2 variants. [WWW Document]. https://aci.health.nsw.gov.au/covid-19/critical-intelligence-unit/sars-cov-2-variants. Accessed 24 Jun 2021

Oladejo BO, Adeboboye CF, Adebolu TT (2020) Understanding the genetic determinant of severity in viral diseases: a case of SARS-Cov-2 infection. Egypt J Med Hum Genet 21(1):1–11

Ortiz-Prado E, Simbaña-Rivera K, Gómez-Barreno L et al (2020) Clinical, molecular and epidemiological characterization of the SARS-CoV-2 virus and the Coronavirus Disease 2019 (COVID-19), a comprehensive literature review. Diagn Microbiol Infect Dis 98(1):115094

Osuchowski MF, Winkler MS, Skirecki T et al (2021) The COVID-19 puzzle: deciphering pathophysiology and phenotypes of a new disease entity. Lancet Resp Med 9:622–642

Ovsyannikova IG, Haralambieva IH, Crooke SN et al (2020) The role of host genetics in the immune response to SARS-CoV-2 and COVID-19 susceptibility and severity. Immunol Rev 296(1):205–219

Pairo-Castineira E, Clohisey S, Klaric L et al (2021) Genetic mechanisms of critical illness in Covid-19. Nature 591(7848):92–98

Pan D, Sze S, Minhas JS et al (2020) The impact of ethnicity on clinical outcomes in COVID-19: a systematic review. E Clin Med 23:100404

Park SE (2020) Epidemiology, virology, and clinical features of severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2; Coronavirus Disease-19). Clin Exp Pediatr 63(4):119

Parkinson N, Rodgers N, Fourman MH et al (2020) Systematic review and meta-analysis identifies potential host therapeutic targets in COVID-19. Sci Rep 10:22303

Peckham H, de Gruijter NM, Raine C et al (2020) Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. Nat Commun 11(1):1–10

Peng Y, Mentzer AJ, Liu G et al (2020) Broad and strong memory CD4+ and CD8+ T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. Nat Immunol 21(11):1336–1345

Phan MV, Ngo Tri T, Hong Anh P et al (2018) Identification and characterization of Coronaviridae genomes from Vietnamese bats and rats based on conserved protein domains. Virus Evol 4(2) vey035

Platanias LC (2005) Mechanisms of type-I-and type-II-interferon-mediated signaling. Nat Rev Immunol 5(5):375–386

Pujadas E, Chaudhry F, McBride R et al (2020) SARS-CoV-2 viral load predicts COVID-19 mortality. Lancet Resp Med 8(9):e70

Qin C, Zhou L, Hu Z et al (2020) Dysregulation of immune response in patients with COVID-19 in Wuhan. China Clin Infect Dis 71(15):762–768

Rausell A, Juan D, Pazos F, Valencia A (2010) Protein interactions and ligand binding: from protein subfamilies to functional specificity. Proc Natl Acad Sci U S A 107(5):1995–2000

Repnik U, Stoka V, Turk V (1824) Turk B (2012) Lysosomes and lysosomal cathepsins in cell death. Biochim Biophys Acta Proteins Proteom 1:22–33

Russo R, Andolfo I, Lasorsa VA, Cantalupo S, Marra R, Frisso G et al (2021) The TNFRSF13C H159Y variant is associated with severe COVID-19: a retrospective study of 500 patients from Southern Italy. Genes 12(6):881

Sadler AJ, Williams BR (2008) Interferon-inducible antiviral effectors. Nat Rev Immunol 8(7):559–568

Schoggins JW, Wilson SJ, Panis M et al (2011) A diverse range of gene products are effectors of the type I interferon antiviral response. Nature 472:481–485

SeyedAlinaghi S, Mehrtak M, MohsseniPour M, Mirzapour P, Barzegary A, Habibi P et al (2021) Genetic susceptibility of COVID-19: a systematic review of current evidence. Eur J Med Res 26(1):1–12

Shang J, Ye G, Shi K et al (2020a) Structural basis of receptor recognition by SARS-CoV-2. Nature 581(7807):221–224

Shang J, Wan Y, Luo C et al (2020b) Cell entry mechanisms of SARS-CoV-2. Proc Natl Acad Sci U S A 117:11727–11734

Shenoy S (2021) SARS-CoV-2 (COVID-19), viral load and clinical outcomes; lessons learned one year into the pandemic: A systematic review. World J Crit Care Med 10(4):132

Simmons G, Gosalia DN, Rennekamp AJ et al (2005) Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. Proc Natl Acad Sci U S A 102:11876–11881

Skol AD, Scott LJ, Abecasis GR, Boehnke M (2006) Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet 38(2):209–213

Sekine T, Perez-Potti A, Rivera-Ballesteros O et al (2020) Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. Cell 183(1):158–168

Suh S, Lee S, Gym H, Yoon S, Park S, Cha J et al (2022) A systematic review on papers that study on Single Nucleotide Polymorphism that affects coronavirus 2019 severity. BMC Infect Dis 22(1):1–11

Stop TB Partnership (2020) The potential impact of the covid-19 response on tuberculosis in high-burden countries, a modelling analysis. https://www.stoptb.org/assets/documents/news/Modeling%20Report_1%20May%202020_FINAL.pdf

Sun J, Zhuang Z, Zheng J et al (2020) Generation of a broadly useful model for COVID-19 pathogenesis vaccination and treatment. Cell 182(3):734–743

Suri JS, Agarwal S, Gupta SK et al (2021) A narrative review on characterization of acute respiratory distress syndrome in COVID-19-infected lungs using artificial intelligence. Comput Biol Med 130:104210

Seyhan AA (2019) Lost in translation: the valley of death across preclinical and clinical divide–identification of problems and overcoming obstacles. Transl Med Commun 4(1):1–19

Schneider WM, Luna JM, Hoffmann HH et al (2021) Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. Cell 184(1):120–132

Schrader K, Huai J, Jöckel L et al (2010) Non-caspase proteases: triggers or amplifiers of apoptosis? Cell Mol Life Sci 67(10):1607–1618

Stoeger T, Amaral LAN (2020) Meta-Research: COVID-19 research risks ignoring important host genes due to pre-established research patterns. Elife 9:e61981

Suryamohan K, Diwanji D, Stawiski EW et al (2021) Human ACE2 receptor polymorphisms and altered susceptibility to SARS-CoV-2. Commun Biol 4:475

Stevens RC (2004) Long live structural biology. Nat Struct Mol Biol 11(4):293–295

Tang D, Kang R, Berghe TV et al (2019) The molecular machinery of regulated cell death. Cell Res 29(5):347–364

Tang D, Comish P, Kang R (2020) The hallmarks of COVID-19 disease. PLoS Pathog 16(5):e1008536

Torre-Fuentes L, Matías-Guiu J, Hernández-Lorenzo L, Montero-Escribano P, Pytel V, Porta-Etessam J et al (2021) ACE2 TMPRSS2 and Furin variants and SARS-CoV-2 infection in Madrid Spain. J Med Virol 93(2):863–869

Trigueiro-Louro JM, Correia V, Santos LA et al (2019) To hit or not to hit: large-scale sequence analysis and structure characterization

of influenza A NS1 unlocks new antiviral target potential. Virology 535:297–307

Trigueiro-Louro J, Correia V, Figueiredo-Nunes I et al (2020) Unlocking COVID therapeutic targets: A structure-based rationale against SARS-CoV-2, SARS-CoV and MERS-CoV Spike. Comput Struct Biotechnol J 18:2117–2131

Vabret N, Britton GJ, Gruber C et al (2020) Immunology of COVID-19: current state of the science. Immunity 52(6):910–941

Van Der Made CI, Simons A, Schuurs-Hoeijmakers J et al (2020) Presence of genetic variants among young men with severe COVID-19. JAMA 324(7):663–673

Walsh KA, Jordan K, Clyne B et al (2020) SARS-CoV-2 detection, viral load and infectivity over the course of an infection: SARS-CoV-2 detection, viral load and infectivity. J Infect 81(3):357–371

Walls AC, Park YJ, Tortorici MA et al (2020) Structure function and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell 181(2):281–292

Wang H, Ma S (2008) The cytokine storm and factors determining the sequence and severity of organ dysfunction in multiple organ dysfunction syndrome. Am J Emerg Med 26(6):711–715

Wang R, Simoneau CR, Kulsuptrakul J et al (2021) Genetic screens identify host factors for SARS-CoV-2 and common cold coronaviruses. Cell 184(1):106–119

Weiner 3rd J, Suwalski P, Holtgrewe M, Rakitko A, Thibeault C, Müller M, Heidecker B (2021) Increased risk of severe clinical course of COVID-19 in carriers of HLA-C* 04: 01. EClinicalMedicine 40:101099

Wei J, Alfajaro MM, DeWeirdt PC et al (2021) Genome-wide CRISPR screens reveal host factors critical for SARS-CoV-2 infection. Cell 184(1):76–91

Weiskopf D, Schmitz KS, Raadsen MP et al (2020) Phenotype and kinetics of SARS-CoV-2–specific T cells in COVID-19 patients with acute respiratory distress syndrome. Sci Immunol 5(48):eabd2071

Wong CK, Lam CWK, Wu AKL et al (2004) Plasma inflammatory cytokines and chemokines in severe acute respiratory syndrome. Clin Exp Immunol 136(1):95–103

World Health Organization (2006) SARS: how a global epidemic was stopped. Manila: WHO Regional Office for the Western Pacific. https://www.file:///D:/DS/suzana.david/Downloads/9290612134_eng%20(1).pdf

World Health Organization, WHO March 11 (2020) WHO Director-General's opening remarks at the media briefing on COVID-19—11 March 2020. https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19-11-march-2020. Accessed 20 Dec 2021. https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencies-coronavirus-press-conference-full-and-final-11mar2020.pdf

World Health organization, WHO November 17 (2020) Weekly epidemiological update - 17 November 2020. https://www.file:///D:/DS/suzana.david/Downloads/weekly-epi-update-14.pdf

World Health Organization, WHO January 20 (2021) SARS-CoV-2 in animals used for fur farming: GLEWS+ risk assessment 20 January 2021 (No WHO/2019-nCoV/fur_farming/risk_assessment/20211). World Health Organization. http://www.fao.org/3/cb3368en/cb3368en.pdf

World Health Organization, WHO (2022) Tracking SARS-CoV-2 variants. https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/. Accessed 26 Feb 2022

World Health Organization, WHO (2020) The potential impact of health service disruptions on the burden of malaria, a modelling analysis for countries in sub-Saharan Africa. https://www.apps.who.int/iris/bitstream/handle/10665/331845/9789240004641-eng.pdf

Wu F, Zhao S, Yu B et al (2020a) A new coronavirus associated with human respiratory disease in China. Nature 579(7798):265–269

Wu J, Yuan X, Wang B et al (2020b) Severe acute respiratory syndrome coronavirus 2: from gene structure to pathogenic mechanisms and potential therapy. Front Microbiol 11:1576

Xu Z, Shi L, Wang Y et al (2020) Pathological findings of COVID-19 associated with acute respiratory distress syndrome. Lancet Resp Med 8(4):420–422

Yoneyama M, Fujita T (2009) RNA recognition and signal transduction by RIG-I-like receptors. Immunol Rev 227(1):54–65

Yu P, Zhu J, Zhang Z, Han Y (2020) A familial cluster of infection associated with the 2019 novel coronavirus indicating possible person-to-person transmission during the incubation period. J Infect Dis 221(11):1757–1761

Zhao J, Zhao J, Perlman S (2010) T cell responses are required for protection from clinical disease and for virus clearance in severe acute respiratory syndrome coronavirus-infected mice. J Virol 84(18):9318–9325

Zhao J, Zhao J, Mangalam AK et al (2016) Airway memory CD4+ T cells mediate protective immunity against emerging respiratory coronaviruses. Immunity 44(6):1379–1391

Zhao J, Yang Y, Huang H et al (2021) Relationship between the ABO blood group and the coronavirus disease 2019 (COVID-19) susceptibility. Clin Infect Dis 73(2):328–331

Zhang X, Tan Y, Ling Y et al (2020a) Viral and host factors related to the clinical outcome of COVID-19. Nature 583:437–440

Zhang Q, Bastard P, Liu Z et al (2020b) Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. Science 370(6515):eabd4570

Zhang Y, Qin L, Zhao Y, Zhang P, Xu B, Li K et al (2020c) Interferon-induced transmembrane protein 3 genetic variant rs12252-C associated with disease severity in coronavirus disease 2019. J Infect Dis 222(1):34–37

Zheng M, Gao Y, Wang G et al (2020b) Functional exhaustion of antiviral lymphocytes in COVID-19 patients. Cell Mol Immunol 17(5):533–535

Zheng HY, Zhang M, Yang CX et al (2020a) Elevated exhaustion levels and reduced functional diversity of T cells in peripheral blood may predict severe progression in COVID-19 patients. Cell Mol Immunol 17(5):541–543

Zhou P, Yang XL, Wang XG et al (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579(7798):270–273

Zhou W, Wang W (2021) Fast-spreading SARS-CoV-2 variants: challenges to and new design strategies of COVID-19 vaccines. Signal Transduct Target Ther 6:1–6

Zhu Y, Feng F, Hu G et al (2021) A genome-wide CRISPR screen identifies host factors that regulate SARS-CoV-2 entry. Nat Commun 12(1):1–11

Zondervan KT, Cardon LR (2007) Designing candidate gene and genome-wide case–control association studies. Nat Protoc 2(10):2492–2501

Zunec R (2020) A review of HLA and COVID-19 association studies. Mol Exp Biol Med 3(2):25–30