Genetics and Molecular Biology, 44, 1(suppl 1), e20210036 (2021)

DOI: https://doi.org/10.1590/1678-4685-GMB-2021-0036

Review
COVID-19 – Special Issue

An immunogenetic view of COVID-19

Vitor R. C. Aguiar[1*], Danillo G. Augusto[2,3*], Erick C. Castelli[4*], Jill A. Hollenbach[2*], Diogo Meyer[1*], Kelly Nunes[1*] and Maria Luiza Petzl-Erler[3*]

[1]*Universidade de São Paulo, Departamento de Genética e Biologia Evolutiva, São Paulo, SP, Brazil.*
[2]*University of California, UCSF Weill Institute for Neurosciences, Department of Neurology, San Francisco, CA, USA.*
[3]*Universidade Federal do Paraná, Departamento de Genética, Curitiba, PR, Brazil.*
[4]*Universidade Estadual Paulista, Faculdade de Medicina de Botucatu, Departamento de Patologia, Botucatu, SP, Brazil.*

Abstract

Meeting the challenges brought by the COVID-19 pandemic requires an interdisciplinary approach. In this context, integrating knowledge of immune function with an understanding of how genetic variation influences the nature of immunity is a key challenge. Immunogenetics can help explain the heterogeneity of susceptibility and protection to the viral infection and disease progression. Here, we review the knowledge developed so far, discussing fundamental genes for triggering the innate and adaptive immune responses associated with a viral infection, especially with the SARS-CoV-2 mechanisms. We emphasize the role of the HLA and KIR genes, discussing what has been uncovered about their role in COVID-19 and addressing methodological challenges of studying these genes. Finally, we comment on questions that arise when studying admixed populations, highlighting the case of Brazil. We argue that the interplay between immunology and an understanding of genetic associations can provide an important contribution to our knowledge of COVID-19.

Keywords: Immunogenetics, COVID-19, SARS-CoV-2, HLA, KIR.

Received: January 26, 2021; Accepted: June 12, 2021.

Introduction

The coronavirus disease 19 (COVID-19) caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is imposing severe humanitarian, social, and economic consequences. The ongoing pandemic has affected the lives of millions of people around the world. As of April 2021, Brazil has recorded the third-highest number of COVID-19 cases worldwide, with nearly 15 million infected individuals and the second-highest number of about 400,000 deaths by COVID-19 (WHO COVID-19 | Brazil, 2021).

Two genera of coronaviruses cause human disease: alphacoronaviruses HCoV-229E and HCoV-NL63, and betacoronaviruses HCoV-HKU1, HCoV-OC43, SARS-CoV (presently named SARS-CoV-1), MERS-CoV (Middle East respiratory syndrome), and SARS-CoV-2 (Ogimi *et al.,* 2020). The four HCoV-* viruses cause mild self-limiting respiratory infections, but MERS-CoV, SARS-CoV-1, and the new SARS-CoV-2 may cause significant morbidity and mortality (Song *et al.,* 2019; Ogimi *et al.,* 2020). The most likely natural reservoir of these three viruses are bats, and the possible intermediate hosts are the palm civet for SARS-CoV-1 and the dromedary camel for MERS-CoV (Song *et al.,* 2019; Ogimi *et al.,* 2020). Whether SARS-CoV-2 was transmitted directly from bats to humans or through an intermediate host is still an open question (Lam *et al.,* 2020).

Send correspondence to Maria Luiza Petzl-Erler. Universidade Federal do Paraná, Departamento de Genética, 81531-980, Curitiba, PR, Brazil. E-mail: perler@ufpr.br.

*All authors contributed to the manuscript equally.

The SARS-CoV-2 is phylogenetically close to SARS-CoV-1, which emerged in 2002 in China and caused more than 8,000 cases in 29 countries over eight months, with a case mortality rate of around 10% (Song *et al.,* 2019). The total number of cases reported for MERS was 2,254 from 2012 through January 2020, with 35% mortality (Song *et al.,* 2019; Rabaan *et al.,* 2020). Comparing data from different diseases and sources is tricky, yet the case fatality rate of COVID-19 is definitely much lower, estimated at 2.2% worldwide (WHO COVID-19 Dashboard, April 13 2021). However, the socioeconomic impact of this disease widely surpasses SARS and MERS, because of the high infectivity and rapid spread of the SARS-CoV-2, and the extreme burden placed on healthcare systems due to the need for hospitalization and artificial ventilation for severe cases.

The symptoms after infection by SARS-CoV-2 range from asymptomatic to severe disease and death (McIntosh *et al.,* 2020; Wu and McGoogan, 2020). The most common clinical symptoms are fever, dry cough, dyspnea, fatigue, dysgeusia, and anosmia (taste and smell disorders). Other common symptoms are myalgia, rhinorrhea, sore throat, diarrhea, nausea and/or vomiting, and headache. About 15% of patients develop severe disease with exuberant inflammatory response, lymphopenia, thromboembolic complications, and hypoxemia that eventually leads to acute respiratory distress syndrome (ARDS) and multiple organ dysfunction syndromes (MODS) (McIntosh *et al.,* 2020; Wu and McGoogan, 2020). Patients may also experience arrhythmias, acute cardiac injury, kidney injury, liver dysfunction, or neurologic manifestations. Possible neurological damage after SARS-CoV-2 infection,

even among recovered patients, increases its impact on the healthcare system (Ellul *et al.,* 2020; Kotfis *et al.,* 2020; Moriguchi *et al.,* 2020; Varatharaj *et al.,* 2020; Zanin *et al.,* 2020). Some features of severe COVID-19 and ARDS are presented in Figure 1.

The risk factors for severe illness are older age, male sex, and medical comorbidities such as diabetes mellitus, cardiovascular disease, hypertension, chronic kidney disease, cancer, obesity, and smoking (Espinosa *et al.,* 2020; McIntosh *et al.,* 2020; Wu and McGoogan, 2020). However, these factors do not explain all the variation, as exemplified by reports of young individuals without comorbidities, including children, that developed severe forms of the disease (e.g., van der Made *et al.,* 2020) and several anecdotal reports of elderly patients with other illnesses that fully recovered from COVID-19. Undoubtedly, host genetics plays a pivotal role in influencing the human response to infection, and rare as well as polymorphic genetic variants are expected to underlie the consequences of SARS-CoV-2 infection. Understanding the role of immune-related genes in the response to SARS-CoV-2, as well as the distribution of associated genetic variants in populations across the world, is a critical element in understanding the disease and in seeking strategies to respond to it.

Human populations have been locked in a coevolutionary arms race with pathogens for thousands of years, with natural selection favoring protective genetic variants against pathogenic viruses, bacteria, and other microorganisms. The findings that as many as 30% of all adaptive amino acid changes in the human proteome are related to selective pressures imposed by viruses underscores the critical participation of viruses in this process (Enard *et al.,* 2016).

Pathogens also evolve in response to host adaptive changes, increasing their capacity to evade the immune response, and to invade, multiply, and be transmitted to other hosts. As a result, the host-pathogen arms race is a dynamic process that leaves signatures in the human genome (reviewed in Karlsson *et al.,* 2014). Therefore, identifying these traces allows discovering genes involved in adaptive response against pathogens (Fan *et al.,* 2016).

Scans for natural selection signatures have identified a conspicuous enrichment of genes related to the immune function in the human genome (Hedrick, 2002; Sabeti *et al.,* 2006; Andrés *et al.,* 2009; DeGiorgio *et al.,* 2014; Bitarello *et al.,* 2018). The group of immune-related genes identified to be under selection is broad, encompassing components of both adaptive and innate immune responses (Akey *et al.,*

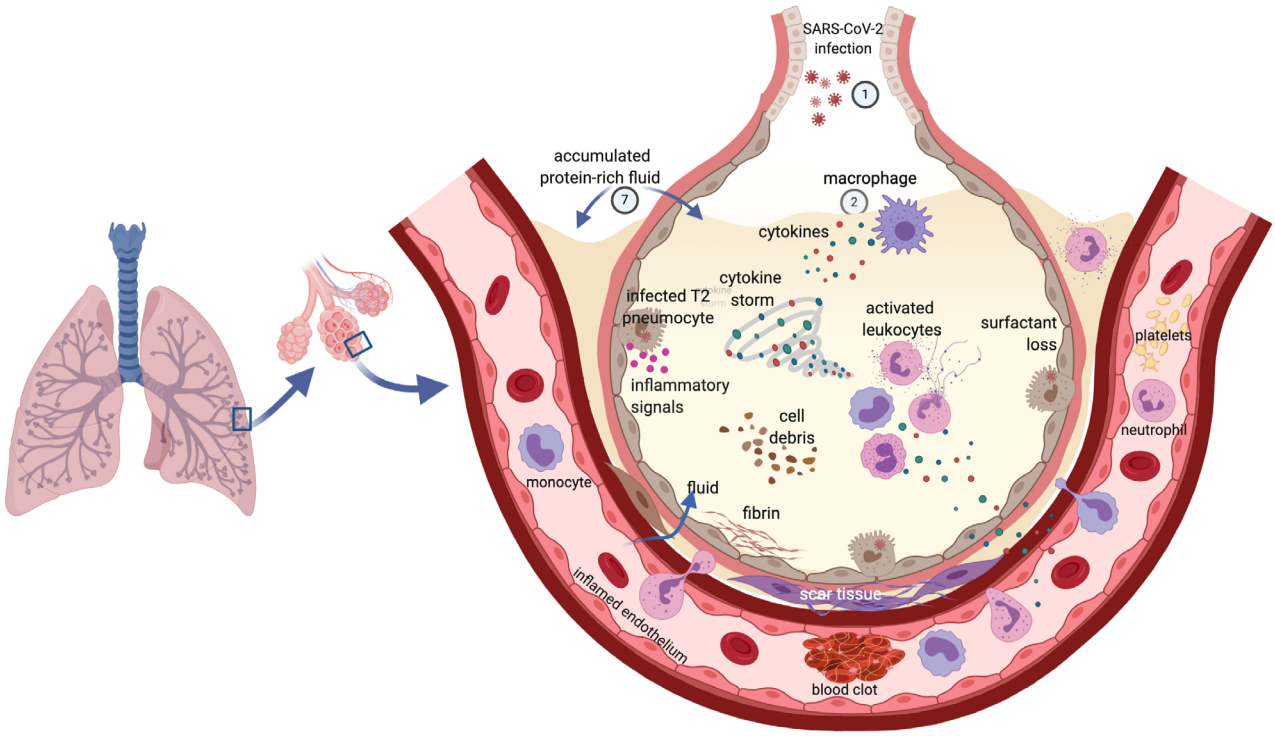

Figure 1 – Features of severe COVID-19 and acute respiratory distress syndrome (ARDS) in the lung. SARS-CoV-2 enters the body by the airways and infects lung cells **(1)**. Immune cells, including macrophages **(2)**, and the infected cells **(3)** react to the virus and produce cytokines, interferons, and additional inflammatory signals **(4)**, which attract other leukocytes. These also produce cytokines **(5)** that may lead to hyper inflammation and the "cytokine storm" **(6)**. The inflamed capillaries allow fluid to sweep into the alveoli and fill the lung cavities **(7)**. Damage to the lung occurs through several processes, including the consequences of surfactant loss **(8)**, the accumulated liquid, and the formation of fibrin and scar tissue **(9)**. With the participation of complement components, coagulation factors, neutrophils, and platelets, blood clots are formed in the inflamed blood vessel **(10)**. The association between thromboembolic events and higher levels of von Willebrand's factor (vWF) and factor VIII with non-group O (see main text) may underlie the association of blood group A with increased susceptibility to severe COVID-19. The deregulated immune response and disturbed coagulation spark inflammation throughout the body damaging other organs and fuels the respiratory failure responsible for most deaths caused by COVID-19. Moreover, the virus may evade the immune response by blocking the effect of immune cells and soluble as well as membrane-bound mediators of immunity and complete its cycle in infected cells, spreading throughout the body. Figure created with Biorender.

2004; Fumagalli *et al.,* 2011). These results show that immune responses against pathogens have been critical for local adaptation of human populations since ancient times. While scans for selection only identify genomic signatures shaped by selection in the past, it is much more challenging to identify evidence of ongoing natural selection.

Many studies have surveyed the association between host genetic variation and susceptibility to infection or disease outcomes. The associations between genetic variants and response to pathogens are the contemporary counterpart of the selective process that generates the evolutionary signatures. Not surprisingly, genes related to immune responses systematically appear in association studies of infectious diseases, with specific variants associated with increased susceptibility, while others are associated with the protection from different diseases (Hill, 2012).

In this review, we discuss the importance of association studies focusing on variation in genes related to immune function, emphasizing HLA (*human leukocyte antigen*) and KIR (*killer-cell immunoglobulin-like receptor)* gene families when searching for host determinants of the response to SARS-CoV-2, without ignoring the paramount importance of other genes involved in fighting viral infections and diseases. The prominent role of the HLA and KIR genes in the viral immune responses, and of HLA in vaccine development, make them natural candidates in the search for COVID-19 disease-relevant variants. We draw attention to the peculiar characteristics of these gene families (e.g., their high polymorphism, distinct allele frequencies in worldwide populations, and their interaction with each other), which impose technical challenges especially for genomewide studies and therefore demand targeted strategies. We list and discuss these challenges and alternative approaches to minimize errors in association studies. Finally, we introduce some research strategies and how the scientific community has been engaged to combine resources, share data, and accelerate the knowledge about the immunogenetics of COVID-19.

HLA in viral infections

The HLA molecules were initially investigated due to their determinant role in allogeneic transplantation outcomes (Dausset, 1958; Klein, 1986; Thorsby, 2009), but their major functions are immunomodulation and the triggering of adaptive immune responses (Bjorkman *et al.,* 1987; Brown *et al.,* 1993; Jones *et al.,* 2006). The classical HLA class I molecules (class Ia), HLA-A, HLA-B, and HLA-C, are expressed in all nucleated cells and present peptides of cytosolic origin to $CD8^+$ T lymphocytes, and together with accessory signals stimulate a cytotoxic response against target cells. The non-classical HLA class I molecules (class Ib) are expressed in specific tissues and their primary function is immunomodulation (Donadi *et al.,* 2011; Kraemer *et al.,* 2014; Persson *et al.,* 2020). In contrast, the classical HLA class II molecules (HLA-DR, HLA-DQ, and HLA-DP) are expressed by antigen-presenting cells and present exogenous antigens to $CD4^+$ T lymphocytes, which in the context of costimulatory signals trigger adaptive immune responses.

The HLA genes are located within the human major histocompatibility complex (MHC), in chromosome region 6p21.3 (Klein and Sato, 2000), and are extraordinarily polymorphic (Robinson *et al.,* 2020) with thousands of alleles in some loci (IPD-IMGT/HLA | Statistics, 2020). The HLA genotypes directly influence the range of antigens presented by a given individual to the T lymphocytes and how the HLA molecules interact with other receptors on NK cells. Consequently, there is variation among individuals regarding the set of viral peptides that they can present and eventually create an efficient response to.

Infectious diseases are one of the leading causes of human mortality (Burgner *et al.,* 2006) and a major selective pressure for human survival (Sabeti *et al.,* 2002; Frodsham and Hill, 2004; Walsh *et al.,* 2006). Among the plethora of genes involved in human immune responses, HLA variants are among those with the strongest reported associations with infection risk and progression (Cooke and Hill, 2001; Tian *et al.,* 2016).

A well-documented example of how HLA alleles influence viral infections is HIV (*human immunodeficiency virus*) infection outcome (Fellay *et al.,* 2007; Kawashima *et al.,* 2009; International HIV Controllers Study *et al.,* 2010). While some HLA-B molecules can accommodate specific HIV antigens and trigger an immune response, others cannot. The homozygosity of class Ia genes was also strongly associated with rapid progression of HIV infection (Carrington *et al.,* 1999; Tang *et al.,* 1999) while differential HLA expression levels were associated with HIV viral load control (Apps *et al.,* 2013; Kulkarni *et al.,* 2013; Ramsuran *et al.,* 2018). Moreover, the HLA variation has been also associated with hepatitis B, hepatitis C, and several other infectious diseases (Blackwell *et al.,* 2009).

During the SARS-CoV-1 (2002-2003) and MERS-CoV (2012) epidemics, several studies demonstrated that HLA alleles were associated with differential susceptibility, but these results were not consistent among studies (Sanchez-Mazas, 2020). The discrepancies are probably explained by the fact that SARS-CoV-1 and MERS-CoV transmission was in specific geographic regions (Asia and the Middle East) and the consequently limited pool of HLA alleles identified in those populations. Due to the pandemic nature of SARS-CoV-2, case-control studies are required to assess a broader range of HLA alleles, potentially revealing new sets of associated alleles differing among geographic regions and populations. Different populations may present distinct alleles associated with susceptibility, depending on the pool of HLA alleles present in each population. Besides differences in study design and statistic power issues, this may be a reason why thus far no conclusive associations between COVID-19 and HLA have been reported (Ellinghaus *et al.,* 2020; Novelli *et al.,* 2020; Wang *et al.,* 2020a,c; Amoroso *et al.,* 2021; Anzurez *et al.,* 2021; Leite *et al.,* 2021; Lorente *et al.,* 2021; Shkurnikov *et al.,* 2021; Yung *et al.,* 2021).

Mapping the potential response to SARS-CoV-2 mediated by HLA peptide presentation

Understanding the repertoire of viral epitopes that specific HLA allotypes can bind provides a mechanistic basis for interpreting genetic associations and contributes to developing vaccines and identifying viral escape epitopes (reviewed in Gfeller and Bassani-Sternberg, 2018).

Despite significant methodological advances in the experimental screening of peptide repertoires presented by HLA (e.g., mass spectrometry and *in vitro* binding assays), only a few HLA allotypes have been studied (Bassani-Sternberg *et al.,* 2015; Caron *et al.,* 2015; Gfeller and Bassani-Sternberg, 2018). Experimental data, together with genetic sequences from pathogens in combination with HLA alleles, make up the reference databases for predictive computational methods (e.g., machine learning, neural network). The success of the computational prediction depends on the availability of large-scale training datasets (Abelin *et al.,* 2017; Dhanda *et al.,* 2019) and the accuracy of the prediction models (Rivino *et al.,* 2013).

SARS-CoV-1 and MERS-CoV experimentally-determined epitopes can be found in several public databases (e.g., the Virus Pathogen Database and Analysis Resource, ViPR (Pickett *et al.,* 2012); The Immune Epitope Database, IEDB (Vita *et al.,* 2019)). Due to their genetic similarity, several studies have used the information obtained experimentally for SARS-CoV-1 to make predictions for SARS-CoV-2 (Ahmed *et al.,* 2020; Grifoni *et al.,* 2020; Lee and Koohy, 2020). However, Ahmed *et al.* (2020) showed that only 23% of known SARS-CoV-1 and SARS-CoV-2 T-cell putative epitopes are identical. Even though part of the SARS-CoV-2 epitope information may not be captured in these comparisons, regions that are identical in SARS-CoV-1 and SARS-CoV-2 are possibly those with a low substitution rate. Consequently, vaccination strategies designed to target the immune response toward these conserved epitope regions could generate immunity that is cross-protective to SARS-CoV-2 and also to other coronaviruses (Grifoni *et al.,* 2020).

HLA affinity prediction for SARS-CoV-1 has been studied based on a limited number of alleles (e.g., *HLA-A*02:01, HLA-A*11:01, HLA-A*24:02*), identifying potential affinities with epitopes from spike (S) and nucleocapsid (N) proteins (Tsao *et al.,* 2006; Rivino *et al.,* 2013). A later study applied an immunization prime-boost strategy to increase the number of memory $CD8^+$ T-cells in the respiratory tract, finding that structural proteins S and N are highly immunogenic and induce longer-lasting neutralizing antibodies than other coronavirus proteins (Channappanavar *et al.,* 2014). Nowadays, many studies of SARS-CoV-2 focus on antigens from these viral structural proteins (Kiyotani *et al.,* 2020; La Porta and Zapperi, 2020; Sanami *et al.,* 2020), as well as on non-structural proteins (Marchan, 2020).

At first, many studies of SARS-CoV-2 were limited to presenting a predicted list of potential candidate epitopes with high affinity to certain HLA allotypes (Grifoni *et al.,* 2020; Kiyotani *et al.,* 2020; Lucchese, 2020; Vashi *et al.,* 2020). However, studies have recently taken a step forward and started evaluating other characteristics of HLA+epitope complexes, such as antigenicity, toxicity, and population coverage (Joshi *et al.,* 2020; Mukherjee *et al.,* 2020; Yarmarkovich *et al.,* 2020). *HLA-A*68:01, B*15:03,* and *DRB1*07:01* are recurrent among the lists of HLA allotypes showing a stronger binding with SARS-CoV-2 predicted peptides (Barquera *et al.,* 2020; Joshi *et al.,* 2020; Iturrieta-Zuazo *et al.,* 2020; La Porta and Zapperi, 2020). Conversely, *HLA-B*14:02*, *B*35:03,* and *B*46:01* have a low predicted binding for SARS-CoV-2 peptides, raising the hypothesis that individuals expressing this molecule may be more vulnerable to COVID-19 (Nguyen *et al.,* 2020). The frequency of each of these HLA alleles varies among populations. For instance, the predicted strong binder *HLA-B*15:03* is frequent in most African populations, particularly in Guinea-Bissau and Uganda, and very rare in Europe and Asia (Figure 2, upper right panel). The predicted weak binder *HLA-B*14:02* is frequent in Europe, Africa, and America, particularly in Brazil (Figure 2, lower right panel). Other examples include the weak binders *HLA-B*35:03,* highly frequent in India and Pakistan, and *B*46:01*, highly frequent and mostly detected in East Asia. This underscores the potential for the genetic basis of response to COVID-19 differing among populations.

Many other HLA alleles were included in some studies but not in others (Barquera *et al.,* 2020; Joshi *et al.,* 2020; Mukherjee *et al.,* 2020; Yarmarkovich *et al.,* 2020; Leite *et al.,* 2021; Shkurnikov *et al.,* 2021). This heterogeneity highlights the methodological differences among studies and the differences in the sets of the selected HLA alleles, which may be restricted to a geographic region in some cases (Kiyotani *et al.,* 2020) (Figure 2, lower left panel), or cosmopolitan (Figure 2, upper left panel) in others (Barquera *et al.,* 2020).

For successful vaccination strategies, it is critical to identify epitopes that can be recognized not only by one but multiple HLA allotypes and, consequently, cover a wide diversity of populations (Ahmed *et al.,* 2020; Barquera *et al.,* 2020). Studies with SARS-CoV-1 drew attention to the binding affinity of viral epitopes based on the functional classification of HLA supertypes (groups of molecules sharing chemical properties in the B and F pockets of the peptide binding region). As expected, allotypes belonging to the same supertype had an affinity to similar viral peptides. In contrast, those belonging to different supertypes had little overlap in the repertoire of viral peptide sets (Sylvester-Hvid *et al.,* 2004). The A3 supertype (*HLA-A*03:01* and *HLA-A*11:01*) has an affinity to the greatest range of SARS-CoV-1 epitopes (Sylvester-Hvid *et al.,* 2004; Blicher *et al.,* 2005).

Certain HLA supertypes are geographically widely distributed. For example, supertype A3 (alleles from the A*03, A*11, A*30, A*31, A*33, A*66, A*68, and A*74 allele groups) is present in at least 44% of the world population (Sette and Sidney, 1999). Moreover, the frequency distribution of supertypes is relatively conserved worldwide (Dos Santos Francisco *et al.,* 2015), which allows the distribution and frequencies of supertypes to also be taken into account in the development of vaccines. In studies for SARS-CoV-2 vaccines, there has been an active effort to identify candidate peptides for which the widely distributed supertypes A3 and B7 have a strong affinity (Kalita *et al.,* 2020), however, these studies are still in the early stages of development.

A major effort of vaccine development is to induce $CD8^+$ cytolytic T lymphocytes (CTL) and $CD4^+$ T-helper immune responses (Ahlers and Belyakov, 2010), and the HLA+peptide complex plays a crucial role in this process. Vaccine development requires a detailed investigation of how SARS-CoV-2 antigens interact with the immune system. However, experimental approaches require long periods of study, which represents a challenge due to the urgency required for the development of effective COVID-19 vaccines.

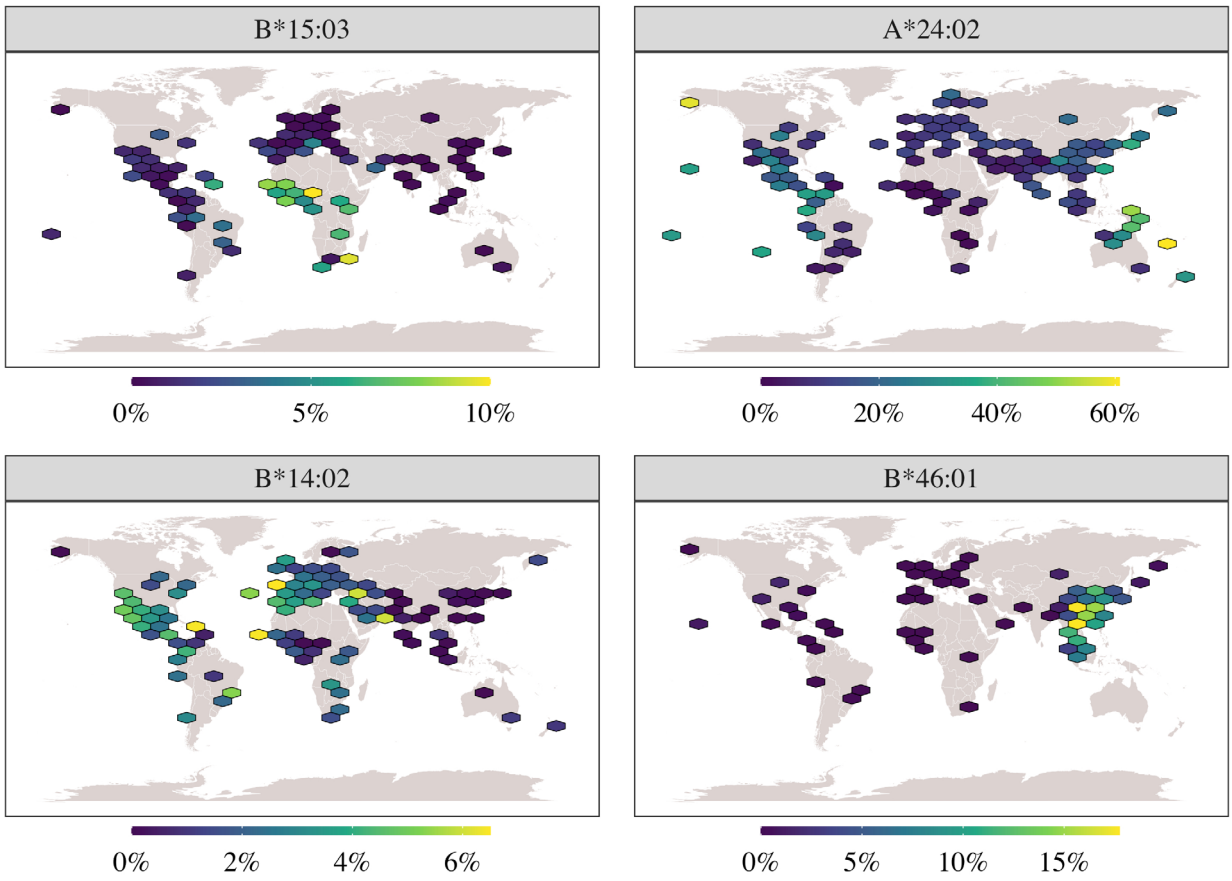

Figure 2 – Global distribution of HLA alleles that are either strong or weak binders of SARS-CoV-2 epitopes. **Upper left panel**: The frequency of the predicted SARS-CoV-2 strong binder *HLA-B*15:03* is high in African populations (usually > 8%) and rare among Europeans, Asians, and Americans. **Lower left panel**: The frequency of the predicted SARS-CoV-2 weak binder *HLA-B*14:02* is high in American populations, particularly Brazil, and also among European and Africans. **Upper right panel**: The frequency of the cosmopolitan allele *A*24:02*. **Lower right panel**: The frequency of the Asian-restricted *B*46:01* allele. Frequency data were obtained from the Brazilian and 1000 Genomes high-coverage sequencing data processed with specific HLA bioinformatics workflow (Naslavsky *et al.,* 2020), and from the allelefrequencies.net website (Gonzalez-Galarza *et al.,* 2020).

Reverse vaccinology assesses the pathogen genome using bioinformatic tools to predict promising target epitopes. Combining it with HLA binding predictions may be an interesting path for vaccine discovery (Enayatkhani *et al.,* 2020; Ong *et al.,* 2020; Tahir Ul Qamar *et al.,* 2020). However, it selects a reduced set of antigens that can better meet a vaccine's requirements: activation of the immune response and effectiveness for most individuals in the population. This strategy does not rule out the vaccine development and testing validation steps required by regulatory agencies to prove the safety and effectiveness of vaccines.

HLA expression in the context of infectious diseases

The differential expression of HLA is also associated with susceptibility to viral infections. For example, higher HLA-C expression leads to a better control of HIV-1 infection (Thomas *et al.,* 2009; Kulkarni *et al.,* 2011; Apps *et al.,* 2013; Bachtel *et al.,* 2018; Parolini *et al.,* 2018); HLA-DP expression has been associated with HBV clearance (Thomas *et al.,* 2012; Ou *et al.,* 2019); HLA-DR levels were shown to correlate with susceptibility to infection by bat Influenza A viruses in human cell lines (Karakus *et al.,* 2019). In a transcriptome-wide association study, Kachuri *et al.* (2020) identified a predominance of associations in HLA class II genes between expression levels and antibody response to multiple prevalent viruses (such as EBV, Herpes, and polyomavirus 2).

Understanding how HLA expression varies among individuals, alleles, and tissues can illuminate the role of HLA genes in SARS-CoV-2 infection. Many studies have profiled HLA alleles with respect to their ability to present SARS-CoV-2 peptides and have identified strong and weak binders. The integration of such data with expression levels could lead to the identification of HLA alleles which are both strong binders and have sufficient expression levels to efficiently elicit an immune response to the virus.

The study of HLA expression is an area of active research and can be undertaken using a wide array of methods. Developments include qPCR-based (Ramsuran *et al.,* 2015) and antibody-based (Apps *et al.,* 2013) approaches to estimate mRNA and surface protein levels, respectively, as well as next generation sequencing (NGS) assays (RNA-seq) that aim to estimate HLA expression at the levels of isoforms or HLA alleles (Cole *et al.,* 2020), and bioinformatics pipelines to extract accurate HLA information from standard RNA-seq data for whole transcriptomes (Boegel *et al.,* 2012; Lee *et al.,* 2018; Aguiar *et al.,* 2019; Orenbuch *et al.,* 2020).

However, there is still scarce knowledge on HLA expression in COVID-19. Previous studies of SARS-CoV-1 and

MERS-CoV indicate that coronaviruses induce transcription changes in infected tissues, including the modulation of HLA genes (Josset *et al.,* 2013; Menachery *et al.,* 2018). The few studies so far on HLA expression in SARS-CoV-2 infection suggest a down-regulation of HLA expression at the mRNA level (Vastrad *et al.,* 2020; Wilk *et al.,* 2020) and at the protein level (Zhang *et al.,* 2020c). These findings indicate that obtaining expression estimates at different levels (mRNA, protein) may help understand HLA regulation in COVID-19.

Whether and how HLA expression affects SARS-CoV-2 infection will also require the disentanglement of different effects to pinpoint which are causal. For example, the protective effects of some HLA alleles may result from the overall gene expression levels which they mark (Thomas *et al.,* 2009; Thomas *et al.,* 2012; Apps *et al.,* 2013; Wissemann *et al.,* 2013); some HLA GWAS SNPs may be non-independent of HLA eQTLs, and HLA eQTLs may be linked to specific HLA lineages (Aguiar *et al.,* 2019).

It may also be necessary to investigate factors located elsewhere in the genome. For example, HLA-C surface levels were associated with a 3'-UTR microRNA binding site, and variation in the microRNA expression itself influences HLA-C levels (Kaur *et al.,* 2017). For HLA class II genes, CIITA is an important transactivator that affects HLA expression regardless of the HLA allele (Carey *et al.,* 2019). Although CIITA plays additional roles in antiviral responses which are not mediated by HLA (Forlani *et al.,* 2016; Bruchez *et al.,* 2020), previous studies reported a downregulation of both CIITA and HLA in MERS-CoV (Josset *et al.,* 2013; Menachery *et al.,* 2018), thus indicating a putative role of the CIITA-HLA interaction in coronavirus diseases.

Killer-Cell Immunoglobulin-Like Receptor (KIR) Molecules Bind HLA and Are Essential for Innate Immunity

Although the primary focus of HLA research is related to antigen presentation to T cells, certain HLA molecules evolved and specialized as ligands for natural killer (NK) cell receptors, activation of the immune response or modulation of its effectiveness (Kiessling *et al.,* 1975; Herberman and Ortaldo, 1981) (Figure 3). Their cytotoxicity against target cells is mediated by surface receptors that recognize abnormal patterns that are characteristics of infected and neoplastic cells (Parham, 2004; Bottino *et al.,* 2005).

Among the variety of NK cell receptors, the killer-cell immunoglobulin-like receptor (KIR) family stands out as the most polymorphic and most explored in the context of diseases. KIR molecules control NK cells' activating and inhibitory signals toward target cells and are regulated by interactions with HLA class I molecules. Infection by several pathogens or neoplastic transformation frequently results in abnormal expression of HLA class I on the cell surface. Abnormal HLA expression ultimately affects the balance of activating and inhibitory signals transduced by NK cell receptors, which triggers the cytotoxic response (Ljunggren and Kärre, 1990; Yawata *et al.,* 2008).

The KIR complex is located at the chromosome region 19q13.42 (Wilson *et al.,* 1997; Wende *et al.,* 1999) and consists of 13 genes and two pseudogenes that exhibit an uncommon structural variation of presence and absence of genes and a high allelic polymorphism. There is growing evidence that KIR and HLA are coevolving as a unique and complex system and are important for human survival (Augusto and Petzl-Erler, 2015). Among all HLA class Ia molecules, HLA-C originated more recently and evolved to primarily bind KIR and regulate NK cell response, while HLA-A and HLA-B retained their primary function as T cell ligands (Older Aguilar *et al.,* 2010). Combinations of KIR-HLA have been associated with diseases (Williams *et al.,* 2005; Kulkarni *et al.,* 2008; Augusto, 2016; Boudreau and Hsu, 2018), including infection (Martin *et al.,* 2007; Podhorzer *et al.,* 2017; Alves *et al.,* 2019; Auer *et al.,* 2020), cancer (Middleton *et al.,* 2007; Al Omar *et al.,* 2010; Kim *et al.,* 2014), and autoimmunity (Suzuki *et al.,* 2004; Augusto *et al.,* 2012; Hollenbach *et al.,* 2016; Anderson *et al.,* 2020).

Variation in genes encoding receptors for natural killer (NK) cells deserves special attention in host genetics affecting the susceptibility to any viral infection, including SARS-CoV-2. For SARS patients, low $CD3^+$, $CD4^+$, $CD8^+$, and NK cell counts might be prognostic indicators to predict admission to intensive care unit (ICU) for SARS patients (Chan *et al.,* 2004). The number of NK cells increased with recovery from SARS, but did not return to normal levels even at the 5th week after the disease onset (Dong *et al.,* 2004). Not only were NK cell counts significantly reduced in SARS patients, but so was the proportion of NK cells expressing the receptor KIR2DL2/3 (CD158b) (National Research Project for SARS, Beijing Group 2004). Moreover, the number of NK and also $KIR2DL2/3^+$ NK cells correlated with disease severity and anti-SARS coronavirus-specific antibodies. More recently, as previously observed for SARS-CoV-1 infection, severe COVID-19 cases exhibited lower counts of NK cells (Jiang *et al.,* 2020; Qin *et al.,* 2020; Sun *et al.,* 2020; Wang *et al.,* 2020b), especially COVID-19 patients admitted to ICU in comparison to no-ICU patients (Bordoni *et al.,* 2020). Combined with the high infiltration of NK cells in the lung from mice infected with SARS-CoV-1 (Yao *et al.,* 2020), these observations provide compelling evidence that NK cells and their receptors may be critical players for immune responses to coronaviruses.

The presence of KIR2DL2/3, whose expression on the NK cell surface was previously implicated with SARS (National Research Project for SARS, Beijing Group 2004), and of their HLA-C ligands were associated with hepatitis B (Gao *et al.,* 2010; Di Bona *et al.,* 2017; Auer *et al.,* 2020). The extensive variation of the KIR genes and the associations with other viral diseases certainly make the KIR family a critical candidate for NK-cell-related COVID-19 studies. A recent meta-analysis (Leite *et al.,* 2021) found no association between the presence/absence of KIR genes and COVID-19 case fatality rate. However, few studies addressed KIR in COVID-19 and much additional work is needed to ascertain if KIR haplotypes, genotypes, and alleles, as well as KIR/HLA compound genotypes are involved in the risk of SARS-CoV-2 infection and COVID-19 outcomes.

The NKG2A (NK group 2 member A) in another NK cell receptor that may be involved in the SARS-CoV-2 response. This receptor recognizes HLA-E as a ligand, suppressing NK cell cytokine secretion and cytotoxicity (Borrego *et al.,* 1998; Braud *et al.,* 1998; Lee *et al.,* 1998).

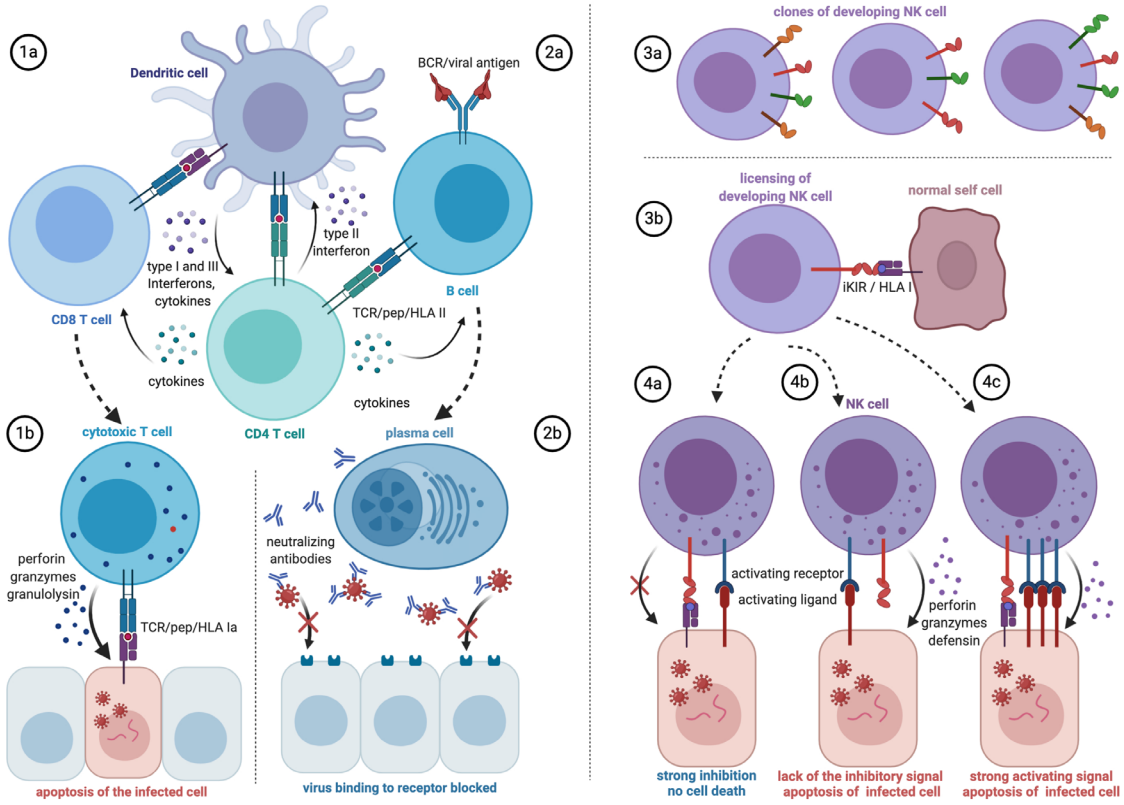

Figure 3 – Predicted HLA involvement in T cell, B cell, and NK cell responses to SARS-CoV-2 infection. **Left**: T and B cells are central to adaptive immunity, whose effectivity is influenced by the individual's HLA genotype. **(1a)** Dendritic cells (DC) present viral peptides bound to HLA Ia and HLA II to, respectively, CD8+ and CD4+ T cell clones displaying specific TCR. (**1b**) The effector cytotoxic CD8+ T cell recognizes the infected target cell by interaction of its specific receptor (TCR) with HLA Ia/peptide on the target and lyses the infected cell. **(2a)** B cells bind viral antigens through specific membrane immunoglobulin receptors (BCR). The internalized antigens go through the class II pathway and HLA II plus viral peptide displayed at the B cell membrane are recognized by the previously primed helper CD4+ T cell to activate the B cells. **(2b)** The activated B cells differentiate in antibody-secreting plasma cells. The neutralizing specific antibodies bind to the virus's antigen, blocking the viral entry into the cell. Other types of antibodies (not shown) may bind to viral antigens at the surface of infected cells to recruit NK cells or trigger the complement cascade. Besides, all the cell-cell interactions indicated in this schematic view depend on signals by accessory membrane molecules (not shown) and soluble factors such as interferons and specific sets of cytokines. **Right**: The natural killer cells (NKc) are important players in innate immunity. NKc repertoires differ among individuals due to the high polymorphism of *KIR* and *HLA* class I, resulting in differential susceptibility to infection and disease. **(3a)** Each individual has numerous NK cell clones that differ for the number and types of inhibitory and activating KIR receptors. **(3b)** During NKc development, the high-affinity binding of inhibitory KIRs (iKIR) with HLA class I (HLA I) enhances the functions of NKc through a process known as licensing. The strength of the interactions depends on the individual's *HLA* and *KIR* genotype and dictates the efficiency of mature NK effector function. **(4a)** The KIR/HLA I interaction inhibits apoptosis of the infected cells even in the presence of activating interactions, especially of NKc licensed by strong interactions; **(4b)** signaling by the activating receptor/ligand leads to apoptosis of the infected target cell when HLA I is absent; **(4c)** strong activating signals may overcome the iKIR/HLA I interaction especially if this interaction is weak, resulting in apoptosis of the target. Moreover, in severe COVID-19, NKc often are reduced in number and dysregulated. Figure created with Biorender.

Moreover, NKG2A expression induces NK and CD8$^+$ T cells to functional exhaustion in viral infections (Li *et al.,* 2013; André *et al.,* 2018). COVID-19 patients exhibited increased expression of NKG2A in comparison to controls, as well as characteristics of functional exhaustion of cytotoxic lymphocytes (Zheng *et al.,* 2020). Like KIR, NKG2A is critical for the education of NK cytotoxicity in early developmental stages (Fauriat *et al.,* 2010; Boudreau and Hsu, 2018; Zhang *et al.,* 2019), and variation in KIR, NKG2A, and their HLA ligands could be responsible for differential immune responses against SARS-CoV-2.

GWAS or Candidate Gene Approaches?

The position of HLA and KIR at the interface between hosts and pathogens, and the extensive list of associations between these loci and diseases, make them obvious candidates when looking for genetic variation associated with COVID-19. However, in the contemporary era of genomic studies, when the whole genome can be routinely queried to identify genetic variants contributing to a phenotype of interest, what could be the rationale for carrying out methods that focus specifically on a subset of loci (i.e., a candidate gene approach)?

The ability to analyze hundreds of thousands or millions of SNPs in microarray-based or NGS-based genome-wide association studies (GWAS) allows the discovery of multiple susceptibility loci in a single study. This exploratory approach is not based on hypotheses, thus permitting the identification of associations with variants that would not even be suspected to be involved in the disease. In fact, GWAS have already identified loci associated with phenotypes of SARS-CoV-2 infection (Ellinghaus *et al.,* 2020; Pairo-Castineira *et al.,* 2020).

However, here we argue that GWAS as implemented leave out relevant variation at HLA, KIR, and other immune-related loci (also discussed in Kwok *et al.,* 2020). These loci

warrant the development of specific approaches to generate data for candidate gene studies, or that can be integrated into a GWAS, allowing to appropriately analyze their role in the host's responses to infection by SARS-CoV-2 and the development of COVID-19.

Even though microarray-based methods can be cost-effective and provide genome-wide genotypic data, they can offer only a coarse map of associations. The SNP microarrays include a limited number of variants in comparison to the total existing variation, and a substantial portion of genetic variation is not captured. This may prove critical in analyses of HLA or KIR, where tag-SNPs in the microarray capture only a fraction of their variation.

Another limitation that affects both SNP microarray-based GWAS and whole genome sequencing-based approaches is the existence of technical hurdles that preclude genotyping coverage for several immune-related regions. Many genes involved in immune responses are extraordinarily polymorphic and were originated by duplications or recombination events. For example, the uncommon structural variation of KIR and high homology among genes result in a lack of suitable reference alignments to include KIR specific SNPs in GWAS arrays. The extensive gene-content variation in KIR haplotypes is incompatible with pre-analysis quality control thresholds typically used in GWAS. Even the ImmunoChip, which was explicitly enriched for this region, only identifies non-coding variants of a single common KIR haplotype (Cortes and Brown, 2011). For example, associations between hepatitis B and C viral infections or diseases and KIR have been observed in candidate gene studies (Gao *et al.,* 2010; De Re *et al.,* 2015; Di Bona *et al.,* 2017; Shan *et al.,* 2018; Auer *et al.,* 2020), but were missed in GWAS (Li *et al.,* 2016; Vergara *et al.,* 2019).

Consequently, the direct association of disease risk with KIR variation has only been detectable by targeted approaches. Other immune system genes, such as immunoglobulin genes, LILR, among others, are also poorly covered in GWAS for similar reasons (Horton *et al.,* 2006; Calonga-Solís *et al.,* 2019; Guselnikov and Taranin, 2019; Lefranc and Lefranc, 2020). HLA genes, on the other hand, may be imputed from SNP microarray data. However, as discussed below, this approach has limitations that may prevent the discovery of relevant associations. Although the MHC region is frequently observed in GWAS (Lenz *et al.,* 2016), the associated HLA alleles or haplotypes are usually missed. Thus, for example, known associations of HLA and malaria were not found in GWAS (Kwok *et al.,* 2020).

Taken together, these points suggest that generating data for KIR, HLA and other genes of the immune system using specific methods may be essential to obtain reliable data for these loci.

Strategies for obtaining HLA and KIR data

The analysis of HLA and KIR loci and their effects on a complex phenotype such as COVID-19 can be carried out in three main frameworks. First, extracting SNP data from microarray-based genotyping methods (SNP microarrays). Secondly, when whole-genome sequencing (WGS) or whole-exome sequencing (WES) data are available, the sequence reads that align to the loci of interest can be selected and processed with appropriate bioinformatics tools to generate SNP and allele calls. Finally, it is possible to obtain HLA data from targeted sequencing through either amplicon or probe-based capture approaches, as many commercial HLA-genotyping kits also do.

When using microarray-based genotyping data, it is possible to infer genotypes of different HLA and KIR genes using SNP data from surrounding regions, an approach known as imputation. The advantage of extracting HLA and KIR data from SNP microarrays is the possibility of analyzing these genes jointly with the genome-wide information, boosted with the relatively reduced costs of the SNP microarrays. Although powerful, HLA imputation is hampered in cases where there is a sparsity of informative markers, which varies among different SNP microarrays and platforms. Imputation methods have delivered HLA genotyping at two-field resolution (protein-level) with accuracy ranging from 89% to 98%, depending on the locus and the population (Zheng *et al.,* 2014; Pappas *et al.,* 2018; Geffard *et al.,* 2020; Chen *et al.,* 2021). Therefore, high-resolution genotypes may be missed, and the inaccuracies could eventually lead to false discoveries. For KIR, imputation methods from SNP microarrays are limited to the assessment of the presence and absence of specific genes (Vukcevic *et al.,* 2015; Chen *et al.,* 2021) because the platforms usually include very few SNPs within the KIR region.

In addition, a major challenge to imputation is the informativeness of the reference panel (i.e., a large subset of samples with both the SNP microarray data and also HLA and KIR alleles genotyped by other methods) for the target sample. Despite efforts (Levin *et al.,* 2014; Okada *et al.,* 2015; Tian *et al.,* 2016; Naslavsky *et al.,* 2020), most reference panels are underrepresented for non-European and/or populations of mixed ancestry (Zheng *et al.,* 2014; Neville *et al.,* 2017), compromising the accuracy of HLA allele imputation for these groups (reviewed in Meyer and Nunes, 2017). For these reasons, there are worldwide initiatives to build better reference panels that include samples of under-represented populations, such as the Brazilian (Naslavsky *et al.,* 2020; Vince *et al.,* 2020).

In this context, WGS and WES data are more attractive – though costly – source of HLA and KIR data. However, the paralogy and extreme polymorphism make it challenging to obtain accurate HLA genotypes from WGS. Mapping short sequencing reads to reference genomes leads to mapping bias and loss of information (Brandt *et al.,* 2015), potentially resulting in erroneous genotyping. Mapping bias is related to two different issues. First, reads carrying many nucleotide differences compared to the reference often fail to align, leading to an overestimation of reference alleles. Second, the cross-mappings between very similar genes reinforce the previous problem and result in the detection of false-positive variants. Bioinformatic approaches have been developed to overcome these difficulties for HLA genes: hla-mapper (Castelli *et al.,* 2018) and MHC-PRG (Dilthey *et al.,* 2015) improve mapping at the HLA region and accuracy of SNP calls; SNP2HLA (Jia *et al.,* 2013), HLA*PRG (Dilthey *et al.,* 2016), HLA-VBSeq (Nariai *et al.,* 2015), Kourami (Lee and Kingsford, 2018), HLAminer (Warren *et al.,* 2012), and OptiType (Szolek *et al.,* 2014) infer HLA alleles directly from short-read sequencing data. There are fewer methods available for KIR genes.

KIR allele genotypes can be obtained from NGS data using the software PING (Norman *et al.,* 2016; Marin *et al.,* 2021), however the cost and computational efforts of analyzing WGS data is a limitation.

Each method has pros and cons, depending on the data available, the number of samples, and the level of resolution needed. For instance, OptiType is highly accurate to predict two and three-field resolution alleles in a single sample basis. However, its database is outdated, and it fails to discriminate alleles differing only in introns and regulatory sequences. HLA-VBSeq also predicts HLA alleles for individual samples, but its accuracy is lower than other methods. Hla-mapper optimizes read alignment in HLA genes, allowing accurate detection of SNP-level genotypes in exons, introns, and regulatory sequences, at the cost of large sample sizes to obtain correct haplotypes. Another essential issue is the bias potentially introduced by hybridization-based capture panels not specifically designed for HLA and KIR genes (such as WES). The polymorphic nature of HLA and KIR may jeopardize the capturing, losing some exonic regions, or only capturing the segments that resemble the reference genome, thus leading to an overestimation of reference alleles. Although some of the methods presented above are compatible with WES, the use of exome data to determine HLA and KIR alleles should be carried out with caution, since the outcome might not represent the correct genotype distribution. Thus, accuracy of HLA and KIR allele calls may be lower for WES when compared to WGS unless specific panels designed for HLA and KIR genes are used.

Finally, targeted sequencing can provide the most precise high-quality data for HLA and KIR. However, this strategy lacks information that allows the analysis of HLA and KIR in combination with other regions of the genome. An interesting alternative is integrating targeted sequencing with a genome-wide approach. For example, Ellinghaus *et al.* (2020) used microarrays to perform a GWAS and applied targeted sequencing to genotype HLA for the association analysis with respiratory failure in COVID-19 patients.

None of these approaches for generating HLA and KIR data is universally preferable. The strategy to be used will depend on the research aims, available funding, synergy with other projects, and the intended speed of delivering results. Regardless of the approach, we emphasize that the complexity of HLA and KIR requires customized methods, or at least specific bioinformatic tools to process sequencing data.

Analytical strategies for HLA and KIR in association studies

Our understanding of how HLA influences susceptibility and resistance to autoimmune and infectious diseases and how HLA and KIR genes have evolved, helps in planning the statistical analyses used in association studies. First, in addition to interrogating SNPs, it is possible to test individual HLA and KIR alleles or HLA supertypes (Ellinghaus *et al.,* 2020), since these are more informative about the functional effects of the HLA molecules. Besides, it is feasible using HLA and KIR alleles as covariates to identify SNP associations that are independent of the HLA alleles or haplotypes (Kachuri *et al.,* 2020).

Another strategy is to code individuals with respect to their heterozygosity over all classical HLA loci, the premise being that individuals that carry a larger number of distinct alleles are more likely to mount an effective response to viral epitopes. This last approach can be further refined by quantifying how much the alleles of an individual differ from each other at the amino-acid level, a measure which Arora *et al.* (2020) found to be correlated to HIV replication. Ellinghaus *et al.* (2020), in their association study of severe COVID-19 with respiratory failure, tested for both the multilocus heterozygosity and the amino acid divergence, but neither was significantly different between cases and controls in their study.

A final challenge for studies on host genetics of COVID-19 refers to the study design itself. While association studies of HLA or KIR and disease phenotypes are most frequently developed in a case-control format, it is by no means clear that this is appropriate for COVID-19. In many countries, including those most affected by the pandemic, the first set of GWAS were carried out at a time when less than 10% of the population had been infected. Thus, random population samples of unaffected individuals comprised a mixture of genetic backgrounds, ranging from possibly susceptible to resistant. This problem likely affected previous studies of SARS and MERS. However, it may be overcome for COVID-19 due to the substantially larger number of infected individuals, and the wide range of disease phenotypes. Accordingly, most studies have focused on comparing "extreme phenotypes", a strategy that may increase the power to detect genetic effects. In this context, it is of utmost importance to carry on studies comparing the genetic background of individuals that have been exposed to the infection without developing the disease (e.g., individuals living in the same house as infected patients) and compare them to those who have contracted COVID-19.

Host Genetics Beyond HLA and KIR

Apart from HLA and KIR, we will now focus on the genes and genetic systems involved in antiviral immune responses that may also be strong candidates for hypothesis-driven COVID-19 studies. Some of the proteins whose genetic variants may be involved in the infection by SARS-CoV-2 or COVID-19 are presented in Figure 4, in the context of an effective antiviral response. The examples we present do not intend to cover the full scope of the genetics of viral infections. Instead, they reveal a small portion of the complexity of the immunogenetics of infectious diseases.

Cytokines and chemokines and their receptors

The cytokine release syndrome (CRS) or "cytokine storm" results from the over-production of soluble mediators of inflammation, which, in turn, sustain an uncontrolled systemic inflammatory response. The CRS characterizes a broad spectrum of non-infectious and infectious diseases, including SARS and MERS, and is common in patients with severe COVID-19 (Coperchini *et al.,* 2020). CSR contributes to the disease's pathophysiology, including hyper inflammation, thrombosis, hypotension, and pulmonary dysfunction in acute respiratory distress syndrome (ARDS) (Moore and June, 2020; Coperchini *et al.,* 2020).

The levels of multiple cytokines and chemokines were significantly higher in SARS-CoV-2 infected patients and

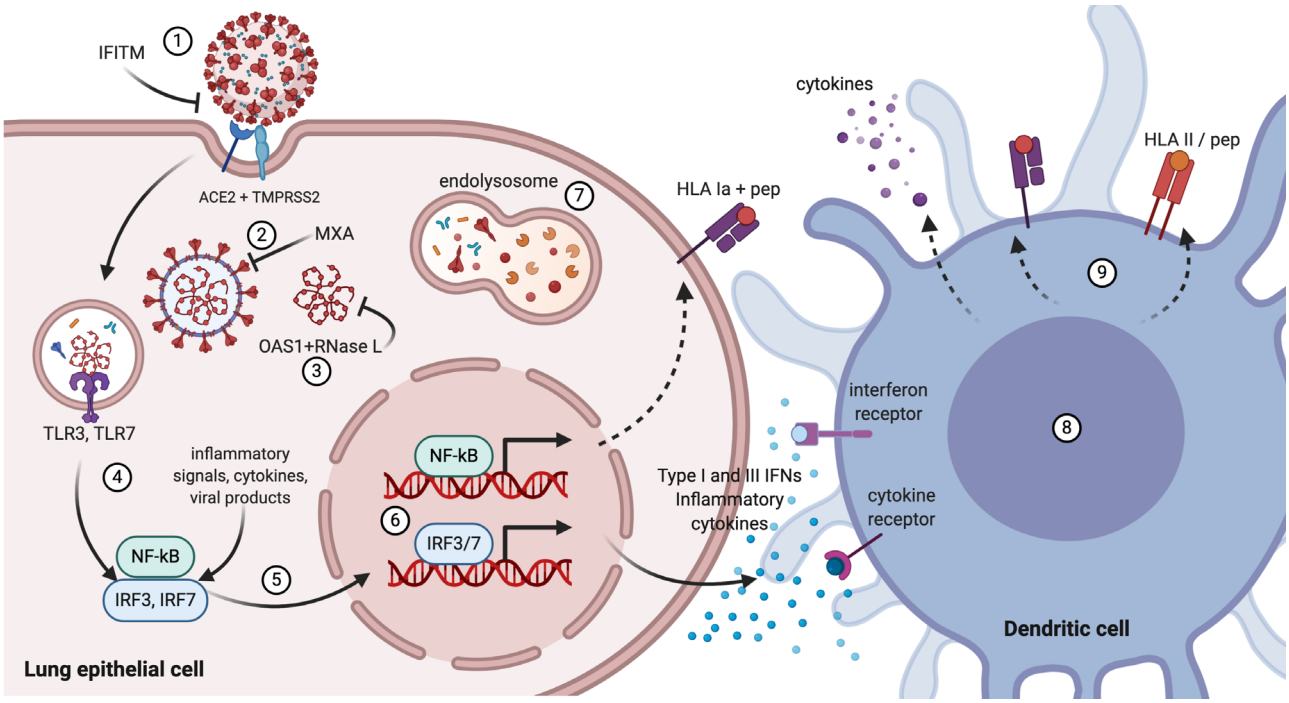

Figure 4 – Some critical molecules involved in antiviral innate immune response, whose loss-of-function mutations and polymorphisms may result in severe COVID-19. SARS-CoV-2 entry into the cell is mediated by TMPRSS2 (transmembrane protease, serine 2) and the receptor ACE2 (angiotensin-converting enzyme 2). **(1)** The interferon-induced transmembrane proteins (**IFITM**) may inhibit the entry of viruses to the host cell cytoplasm. **(2)** The interferon-induced GTP-binding protein **MXA** may block endocytic traffic of incoming virus particles. **(3)** The 2'-5'-oligoadenylate synthase 1 (**OAS1**) binds ribonuclease L (**RNase L**) leading to its activation with subsequent degradation of the viral and cellular RNA, thus terminating viral replication. **(4)** Toll-like receptors 3 and 7 (**TLR3**, **TLR7**) recognize viral RNA leading to activation and **(5)** nuclear translocation of transcription factors NF-kappa-B (NF-kB) and IFN regulatory factors 3 and 7 (IFR3, IRF7). **(6) IFR3** and **IRF7** regulate the transcription of type I IFN genes (IFN-alpha and IFN-beta) and IFN-stimulated genes (ISG); NF-kB is a pleiotropic transcription factor crucial for regulation of numerous genes involved in immunity and other biological processes, such as apoptosis. **(7)** The lysosomal enzymes digest viral components in the phagolysosome. **(8)** Type I and III (IFN-lambda) interferons, NF-kB, and cytokines promote expression of numerous genes involved in innate and adaptive immune responses in different cells, including **(9)** up-regulation of **HLA** gene expression in antigen-presenting cells. Figure created with Biorender.

were associated with the severity of COVID-19 (Chi *et al.,* 2020; Huang *et al.,* 2020). Dysregulated activation of the mononuclear phagocyte compartment may contribute to the COVID-19-associated hyper inflammation (Merad and Martin, 2020). Activation of monocyte-derived macrophages by factors released from SARS-CoV-2 infected alveolar epithelial cells, activated T cells, and others, release massive amounts of IL-6 and other proinflammatory cytokines that initiate downstream signaling pathways. Also, the interaction of circulating activated monocytes and activated or damaged endothelial cells may trigger the extrinsic coagulation pathway, leading to intravascular blood clotting. This process can be amplified by the recruitment of neutrophils by the activated endothelial cells, which triggers the intrinsic / contact coagulation pathway (Merad and Martin, 2020).

Associations with variants in the genes encoding inflammatory cytokines, chemokines, and their receptors, have been reported for many diseases. They include anti-inflammatory cytokines or those that exert both pro- and anti-inflammatory effects, like TGF-b, IL-4, IL-10, and IL-27. Examples of viral diseases are the respiratory syncytial virus (RSV) infection (Gentile *et al.,* 2003) and hepatitis B virus infections (Gusatti *et al.,* 2016). No doubt, a comprehensive study of cytokine and chemokine genetic variation and expression levels will shed light on COVID-19 and its complications.

A strong association between COVID-19 and respiratory failure with a genomic region at 3p21.31 was reported in a GWAS of Italian and Spanish patients with severe disease (Ellinghaus *et al.,* 2020). The association signal revealed a cluster of several protein-coding and lncRNA genes, although the data could not reliably implicate a causal gene. However, three of the six protein-coding genes encode chemokine receptors, including the C-C motif chemokine receptor 9 (CCR9), the C-X-C motif chemokine receptor 6 (CXCR6), and the X-C motif chemokine receptor 1 (XCR1) (Ellinghaus *et al.,* 2020). Thereafter, results from two GWAS confirmed the association with this region at chromosome 3p21.31 (COVID-19 Host Genetics Initiative 2020) | GWAS meta-analyses round 5 2021; Shelton *et al.,* 2021). This genomic region has a large effect on COVID-19 morbidity and mortality. The effects are similar to or larger than the ones of most established risk factors and are age-dependent, such that the risk is higher in younger individuals (Nakanishi *et al.,* 2021).

The associated genetic variants in 3p21.31 are all in strong linkage disequilibrium - LD ($r^2 > 0.98$) and span almost 50 kb. This haplotype occurs in South Asia at a frequency of 30%, in Europe at 8%, and at lower frequencies in East Asia, but is absent in Africans. Therefore, the analysis of Africans could help to narrow-down the COVID-19 causal variant(s). Interestingly, the same extended haplotype was found in 50,000 to 120,000 old Neanderthal genomes, leading

to the conclusion that it is inherited from Neanderthals (Zeberg and Pääbo, 2020). The unusually contrasting haplotype frequencies between South and East Asia indicate the effect of natural selection, possibly because of the exposure of these populations to different pathogens in the past. Concerning the SARS-CoV-2 pandemic, the chromosome 3 haplotype is now under negative selection with dramatic consequences (Zeberg and Pääbo, 2020).

The complement system

The complement system comprises a network of dozens of soluble and cell membrane proteins that work in a coordinated manner towards the activation of three pathways – the classical, alternative, and lectin pathways (reviewed in Beltrame *et al.* (2015) and Reis *et al.* (2019)). These pathways are crucial in innate immunity and crosstalk with the adaptive response components, influencing disease outcomes.

Both deficient and excessive activation of the complement molecules may be harmful to the host. Complement system deficiencies are frequently associated with increased susceptibility to infections and adverse manifestations in other diseases, especially in immunocompromised patients (reviewed in Beltrame *et al.,* 2015; Reis *et al.,* 2019). Conversely, abnormally increased levels of complement proteins, which are partially explained by genetic variation, may contribute to hyperinflammatory responses, observed also in severe COVID-19. Further, the complement system is also involved in other biological processes, including coagulation, which has been implicated in various COVID-19 complications (Bumiller-Bini *et al.,* 2021).

Some features of severe COVID-19 suggest that complement activation is possibly playing a critical role in the pathogenesis of this disease, particularly during exaggerated immune responses (Java *et al.,* 2020). Genetic variants of several complement regulatory proteins and factors, including CD55(DAF), CFH, C3, C4BPA, and COLEC11, have been associated with adverse COVID-19 clinical outcomes (Ramlall *et al.,* 2020). The comparison of postmortem lung biopsies of COVID-19 and H1N1 patients and control patients who died of causes not involving lung lesions showed higher expression of FCN3 (ficolin 3) in both diseases compared to the control group. The MBL2 (mannose-binding lectin) level was increased only in the COVID-19 group (Malaquias *et al.,* 2020). Instead, low MBL2 serum levels and a variant of the gene *MBL2* were pointed out as risk factors for SARS-CoV-1 infection that causes the COVID-19-related disease SARS (Zhang *et al.,* 2005).

Toll-like receptors and genes involved in type I interferon regulated immunity

Cells detect pathogen-associated molecular patterns (PAMP) through pattern-recognition receptors (PRR), which allows semi-specific recognition of pathogenic microorganisms and viruses and influences innate and adaptive immune responses. The toll-like receptors (TLR) constitute one of the PRR classes (reviewed by Frazão *et al.,* 2013). The surface receptors TLR1, TLR2, TLR4, TLR5, TLR6, and TLR10 are mainly responsible for detecting components from extracellular bacteria and fungi. However, these receptors also detect viral capsid proteins, including the SARS-CoV-2 S-protein (Choudhury and Mukherjee, 2020). Conversely, the intracellular TLR3, TLR7, TLR8, and TLR9 primarily recognize nucleic acids from viruses and bacteria. The TLRs upregulate anti-viral and pro-inflammatory mediators and, therefore, modify the infection's outcome positively, limiting the viral load, or negatively, by triggering the exacerbated inflammation associated with the CRS.

In the Netherlands, whole-exome sequencing was performed for four young men hospitalized with severe COVID-19, all without a history of major chronic diseases. The two brother pairs were aged 32 and 29 (family 1, Dutch ancestry), 23 and 21 (family 2, African ancestry). One of the patients died. The study identified two different rare novel *TLR7* loss-of-function variants (van der Made *et al.,* 2020). The analysis of men with severe COVID-19 aged less than 60 years found *TLR7* deleterious missense variants in 2.1% of the patients and in none of the asymptomatic participants (Fallerini *et al.,* 2021). In both studies, the *TLR7* variants were associated with impaired type I and II IFN responses. Recessive or incompletely dominant loss-of-function mutations of *TLR7* and other X-chromosomal genes such as *ACE2* (that encodes angiotensin-converting enzyme 2, the cellular receptor for SARS-CoV-2) and *NEMO* (encodes NF-kappa-B essential modulator, a regulatory protein involved in antiviral response) may be partly responsible for the higher risk of severe COVID-19 and higher death rates in men compared to women (Espinosa *et al.,* 2020; Patil *et al.,* 2020). Moreover, less detrimental common and rare variants of TLRs and other PRRs could contribute to the polygenic component of the COVID-19 susceptibility.

Rare variants at 13 genes known to govern TLR3- and IRF7-dependent type I interferon (IFN) immunity to viruses were searched in 659 patients with severe COVID-19 pneumonia and 534 individuals with asymptomatic infection or mild disease. The study unveiled 24 loss-of-function variants underlying autosomal recessive or dominant deficiencies in 23 (3.5%) of critically ill patients, aged 17 to 77 (Zhang *et al.,* 2020a). The 24 mutations were found in 8 of the 13 genes: *TLR3*, *UNC93B1* (protein unc-93 homolog B1 that regulates nucleotide-sensing TLR signaling), *TICAM1* (TIR domain-containing adapter molecule 1, a cytoplasmic viral sensor which participates in activation of transcription factors and induction of proinflammatory cytokines), *TBK1* (a multifunctional serine/threonine-protein kinase that plays an essential role in the TLR3- and IFN-dependent control of viral infections), *IRF3* and *IRF7* (IFN regulatory factors 3 and 7, key transcriptional regulators of type I IFN-dependent immune responses against DNA and RNA viruses), and *IFNAR1* and *IFNAR2* (IFN alpha/beta receptors 1 and 2, which associate to form the type I IFN receptor) (Zhang *et al.,* 2020a). Polymorphisms of *IFNAR2* and *TYK2* (involved in the initiation of type I IFN signaling) were also associated with critical COVID-19 in a recent GWAS of individuals of mostly European descent (Pairo-Castineira *et al.,* 2020). Moreover, evidence in support of a causal link between low expression of *IFNAR2* and high expression of *TYK2* with life-threatening COVID-19 was reported (Pairo-Castineira *et al.,* 2020). Remarkably, life-threatening COVID-19 pneumonia can also result from

auto-immune phenocopies of these inborn errors of type I IFN immunity. At least 10.2% (2.6% of women and 12.5% of men) of 987 critically ill COVID-19 patients had neutralizing IgG autoantibodies against type I IFNs; they were aged 25 to 87 years and 95 were men. None of the 663 subjects with asymptomatic SARS-CoV-2 infection or mild COVID-19 had these autoantibodies, which were present in only 0.0033% of 1,227 healthy individuals (Bastard *et al.,* 2020).

Proteins that antagonize viral entry into the host cell and replication

The 2'-5'oligoadenylate synthetase (OAS) protein family consists of the OAS1, OAS2, OAS3, and OAS-like (OASL) proteins. Type I and II IFNs induce synthesis of the OAS proteins that recognize exogenous nucleic acid to initiate antiviral pathways (Hovanessian and Justesen, 2007). The OAS1 is a tetrameric interferon-induced dsRNA-activated antiviral enzyme. It leads to dimerization and activation of ribonuclease L (RNase L), culminating in cellular and viral RNA degradation, thus inhibiting protein synthesis and viral replication. Alternatively, the antiviral effect can also be mediated via a pathway independent of RNase L (Uniprot | P00973, 2020). The paralogous genes *OAS1-3* are closely linked at the chromosomal position 12q24.13 and are involved in the same general function of inhibiting the early viral replication. The fourth member of the family, *OASL,* is located at 12q24.31 and has anti- and pro-viral dual functions, which depend on various mechanisms and the phase of the infection (Choi *et al.,* 2015).

The SNP rs2660 G>A in *OAS1* was associated with SARS-CoV-1 infection in Han Chinese from Beijing. The allele rs2660*G conferred a dominant protective effect on SARS infection (He *et al.,* 2006). Other viral infections are influenced by *OAS* gene variants or expression levels as well. *OAS1-OAS3-OAS2* haplotypes were associated with clinical outcomes of dengue virus infection in India (Alagarasu *et al.,* 2013). The Zika virus (ZIKV) infection of A549 cells induces *OAS2* expression that inhibits ZIKV replication through enhanced IFNβ expression, which leads to the induction of the Jak/STAT signaling pathway (Liao *et al.,* 2020). Association of critical COVID-19 with rs10735079 in the *OAS1-3* gene cluster was reported for a sample of mostly European ancestry in a recent GWAS (Pairo-Castineira *et al.,* 2020) and increased levels of OAS1 decreases the susceptibility to COVID-19 and, in particular, severe COVID-19 (Hernández-Cordero *et al.,* 2021). The splice-site variant rs10774671*G in the gene *OAS1* is associated with greater OAS1 expression and has strong alternative splicing quantitative trait loci (asQTL) effect on *OAS1* (Sams *et al.,* 2016) that results in higher levels of the p46 isoform and reduced COVID-19 susceptibility and severity (Zhou *et al.,* 2021). Interestingly, rs10774671 presents strong linkage disequilibrium (LD; $r^2 > 0.8$) with other ~130 SNPs in non-Africans (including rs2669 and rs10735079 (cited above). In Africans, however, only weak LD is observed between rs10735079 and all other SNPs ($r^2 \leq 0.5$), according to LDlink (Machiela and Chanock, 2015). The reason for such distinct LD patterns among continental populations was explored and discussed in detail by Sams *et al.* (2016). The *OAS* genomic region exhibits an elevated frequency of Neandertal-derived alleles in non-African populations, despite the known purifying selection against Neandertal ancestry observed in humans (Sankararaman *et al.,* 2014; Gallego Llorente *et al.,* 2015). Therefore, the OAS Neandertal-introgressed haplotype was subjected to positive selection in human populations, possibly because it reintroduced the ancestral splice-site variant rs10774671*G (Sams *et al.,* 2016). This is a magnificent example of the relevance of studying population genetics and evolution of immune-related genes and highlights why differences between African and non-African populations must be considered in future studies of *OAS* variation and SARS-CoV-2 infection.

The members of the interferon-induced transmembrane (IFITM) family are antiviral cell-intrinsic restriction factors that inhibit the viral entry into the host cells by restricting the membrane fusion. The IFITM proteins are active against SARS-CoV-1, influenza A virus, Ebola virus, dengue virus, HIV-1, among others (Bailey *et al.,* 2014), and therefore are candidates for genetic studies in SARS-CoV-2 infection. The three paralogous *IFITM1-3* genes are located at chromosome region 11p15.5. In a preliminary study of COVID-19, homozygosity for the *C* allele of rs12252 in the gene *IFITM3* was associated with more severe outcomes in an age-dependent manner (Zhang *et al.,* 2020b). The splice-site variant rs12252 is also a 5'UTR and synonymous variant of *IFITM3,* as well as a long non-coding RNA (lncRNA) variant (Ensembl | rs12252, 2020; GTEx Portal | rs12252, 2020). Future studies should explore multiple tag SNPs along the gene-dense genomic region that hosts the three IFITM genes (Ensembl | IFITM genes region, 2020), besides the *NLRP6* (NLR family pyrin domain containing 6) gene that encodes the sensor component of the NLRP6 inflammasome and is involved in innate immunity and inflammation, and *IRF7*, already shown to be implicated in severe COVID-19 (see above).

The human myxovirus resistance protein MxA also plays an important role in the outcome of human viral infections. The intracellular MxA protein is induced by type I interferons and has broad activity against diverse RNA viruses and a few DNA viruses. *MX2* produces nuclear and cytoplasmic MxB forms and has potent activity against human immunodeficiency virus type 1 (HIV-1) and herpesviruses (Staeheli and Haller, 2018). The *MX1*, *MX2* and *TMPRSS2* (transmembrane protease serine 2, which is critical for SARS-CoV-2 host cell entry (Hoffmann *et al.,* 2020), are closely linked at the chromosomal region 21q22.3 and linkage disequilibrium in that region is strong ($r^2 > 0.8$) in non-African populations. The minor (less frequent) alleles of five SNPs correlated with high of *MX1* expression in blood and a reduced risk of developing severe COVID-19 in Europeans (Andolfo *et al.,* 2021). Previously, polymorphisms in the *MX1* promoter associated with increased transcription *in vitro* were associated with decreased susceptibility to SARS in the Chinese Hong Kong population (Ching *et al.,* 2010).

The blood groups

The biological role of ABO (alternatively ABH) blood groups and their effects on infections, immunity, thrombosis, cardiovascular disease, and metabolism have been thoroughly reviewed (Nordgren and Svensson, 2019; Stowell and Stowell, 2019a,b). The synthesis of these histo-blood group antigens

is mediated by fucosyl- and glycosyltransferases under the genetic control of *FUT2* (secretor), *FUT3* (Lewis), and *ABO*(*ABH*) genes. The variants of these genes are associated with susceptibility to infection and the severity of many human pathogens, including *Helicobacter pylori* and *Vibrio cholerae* (Stowell and Stowell, 2019a) and Norovirus (Nordgren and Svensson, 2019). The highly diverse Noroviruses (or Norwalk-like viruses) are the most common etiological agent of acute gastroenteritis worldwide, with the viral genotype GII.4 mostly implicated in human disease (Nordgren and Svensson, 2019). Disease susceptibility is dependent on the Norovirus genotype and is mediated by the presence or absence of the A, B, and/or H antigens on gut epithelial surfaces, the so-called secretor phenotype (Nordgren and Svensson, 2019). Non-secretor individuals, those having a nonsense mutation in *FUT2* that causes the expression of an inactive FUT2 enzyme, do not express these blood antigens in mucosal tissues and are also resistant to GII.4 and several other Norovirus genotypes. The *FUT2* protective allele is fully penetrant against infection as none of the non-secretor individuals develop Norovirus infection (Nordgren and Svensson, 2019).

A study of SARS demonstrated that patients from blood group O exhibited a lower risk of infection by SARS-CoV-1 when compared with non-O participants (Cheng *et al.,* 2005). As to COVID-19, significantly decreased and increased risk was reported for individuals from blood groups O and A, respectively, in 2,173 hospitalized patients with a confirmed infection by SARS-CoV-2 in Wuhan and Shenzhen, China, compared to the frequencies in populations from the same regions (Zhao *et al.,* 2020). Numerous studies in different populations followed. The consensus that emerged is that the association of COVID-19 with the ABO locus is highly significant and the risk of SARS-CoV-2 infection and possibly also COVID-19 severity is slightly lower for group O than for non-O groups (Ellinghaus *et al.,* 2020; Liu *et al.,* 2020; Pendu *et al.,* 2021).

Besides, associations between arterial and venous thromboembolic events with non-group O have been consistently observed in the literature. Interestingly, von Willebrand's factor (vWF) and factor VIII levels are significantly higher in non-group O individuals, which might predispose them to pathologic clot formation (Stowell and Stowell, 2019b). In addition, there is evidence that ABO blood groups may affect vascular biology independently of, or in conjunction with, alterations in hemostasis (Stowell and Stowell, 2019b). For example, ABO blood group status may influence outcomes following acute vascular injury. Critically ill patients who presented a major trauma or severe sepsis were evaluated for ABO blood group status in addition to the development of acute respiratory distress syndrome (ARDS) or acute kidney injury (AKI). Among patients of European ancestry, but not African, blood group A individuals were more likely to develop ARDS or AKI (Stowell and Stowell, 2019b). In COVID-19, ARDS, AKI, and hemostasis alteration have also been reported (Wu *et al.,* 2020, McIntosh, 2020). The analysis by a multi-omics approach suggested that the increased ABO protein level is a causal risk factor for COVID-19 susceptibility and severity (Hernández-Cordero *et al.,* 2021). Clinical COVID-19 phenotypes, especially circulatory system complications, including thrombotic and coagulation-related phenotypes, were associated with genetically predicted increased ABO expression levels (Pathak *et al.,* 2020).

Genes involved in mucociliary clearance

Eight potential genomic regions ("super-variants") associated with mortality by COVID-19 were identified in a GWAS with more than 18,600,000 SNPs (Hu *et al.,* 2021). The white British patients' sample included 1,096 SARS-CoV-2 infected cases, of which 292 were deaths and 804 were survivors. The disruption of DNAH7 (dynein heavy chain 7, axonemal) function may cause ciliary dysmotility and weakened mucociliary clearance capability, and CLUAP1 (clusterin associated protein 1) is required for ciliogenesis. This finding evidences the importance of respiratory cilia functioning properly in COVID-19 patients. Interestingly, *DNAH7* is the most downregulated gene after *in vitro* infection of human bronchial epithelial cells with SARS-CoV-2 (Nunnari *et al.,* 2020). The protein WSB1 (WD repeat and SOCS box-containing protein 1) is involved in the innate immune response and antigen processing and presentation by HLA class I molecules and may enhance maturation of the IL-21 receptor, which is involved in NK and T cell functions. The other genomic regions contain variants related to thromboembolic disease, mitochondrial dysfunction, and cardiovascular disease (Hu *et al.,* 2021).

Proteins of the mucin family are components of the mucociliary clearance system, and therefore are crucial for innate defense. The mucin 5B, oligomeric mucus/gel-forming protein (MUC5B), is secreted in the lung, saliva, and cervix (Uniprot | Q9HC84, 2020). A polymorphism in the *MUC5B* gene is strongly associated with protein expression levels and susceptibility to some diseases. Patients with idiopathic pulmonary fibrosis (IPF) had significantly higher levels of MUC5B in the lung than controls, and the allele rs35705950*T was strongly associated with risk, especially in homozygosis (Seibold *et al.,* 2011). This SNP is located upstream of the *MUC5B* transcription start site and is predicted to affect the binding affinity of different transcription factors and to be a splice-site variant of the long non-coding RNA *AC061979.1* gene that overlaps the *MUC5B* gene (Ensembl | *MUC5B* gene region, 2020). The allele rs35705950*T, a risk factor for IPF, was associated with protection against the development of severe COVID-19 in older adults (Fadista *et al.,* 2021). The observed association with rs35705950 could be due to a protective effect of high mucin production in the airways, but the authors could not rule out a patient selection bias associated with the rs35705950 SNP (van Moorsel *et al.,* 2020) Other mucin family members might also be candidates for studies of COVID-19 complications in the gut and airways.

Consolidating Information

The importance of investigating genetic associations in diverse countries has recently become more evident (Hindorff *et al.,* 2018), especially because the genetic basis of many complex phenotypes differs significantly among geographic regions (Martin *et al.,* 2017). These differences may be explained by both geographic-related pathogen variability and variation in host genetics, which might influence the function of specific

genes (reviewed in Domínguez-Andrés and Netea, 2019). Differential allele frequencies and linkage disequilibrium are the major factors responsible for discordant associations in different populations. For example, LD in regions 3p21.31 and 12q24.13 is strong and extended in Europeans, but absent in Africans and admixed Latin-American populations from the 1000G project (LDlink April 19, 2021). These regions harbor, respectively, a multigene cluster and the OAS gene cluster, both associated with severe COVID-19 in Europeans (see Section "Host genetics beyond HLA and KIR"). A clear example of discordant results due to contrasting allele frequencies between populations comes from HLA in endemic pemphigus foliaceus. Among other differences, the *HLA-DRB1* alleles associated with the highest risk in the Brazilian population of predominantly European ancestry are *DRB1*01:02* and **04:04*, but in Native Americans, highest risk is associated with *DRB1*04:04* only, because *DRB1*01:02* is not present (Petzl-Erler, 2020). Thus, the analysis of non-Europeans and admixed populations is urgently needed to identify the COVID-19 causal variants and to evaluate their impact on COVID-19 in worldwide populations. To this end, several collaborative initiatives and consortia have been launched and the first results are emerging (for example, Castelli *et al.* (2021); Castro *et al.* (2021); and Rede Genômica IPEC).

The extent of the admixture in Brazil poses challenges in study designing since ancestry heterogeneity is known to underlie spurious associations (Seldin *et al.,* 2011; Thornton and Bermejo, 2014). The ancestry varies among Brazilian geographic regions, with a higher proportion of Native American contribution in the North, larger proportion of African in the Northeast, and a predominance of European background in the South and Southeast (reviewed by Souza *et al.* (2019). In addition, the ancestry of traditional communities such as *Ribeirinhos* and *Quilombolas* differs significantly from other populations within the same region (Kimura *et al.,* 2013; Gontijo *et al.,* 2018).

Some reports suggest that ancestry may be associated with differential risk for COVID-19 (Sze *et al.,* 2020; Andrasfay and Goldman, 2021; Shelton *et al.,* 2021). However, it is challenging to disentangle the effects of ancestry from other confounding factors. For example, in countries such as the USA, UK, and Brazil, the non-European ancestry is associated with a lower socioeconomic status, which by its turn is also associated with higher incidence of some comorbidities (Franceschini *et al.,* 2013; Musemwa and Gadegbeku, 2017) and with conditions that increase transmissibility and vulnerability. Among these conditions are overcrowded housing, occupation, access to healthcare, stress, unbalanced diets, use of public transport, and others (Hawkins, 2020; Sze *et al.,* 2020). Several studies indicate that COVID-19 is more severe and deadly in groups with lower socioeconomic status (Demenech *et al.,* 2020; Figueiredo *et al.,* 2020; Hawkins *et al.,* 2020; Clouston *et al.,* 2021). The conditions associated with socioeconomic deprivation may interfere in the immunological response to COVID-19, especially in an exacerbated pro-inflammatory response and senescent phenotypes (Holuka *et al.,* 2020).

Notwithstanding, other studies have found that non-European ancestries are associated with an increased risk of death by COVID-19, even after controlling for comorbidities and socioeconomic status (Harrison *et al.,* 2020; Shelton *et al.,* 2021). Contrasting frequencies of genetic variants that are implicated in differential susceptibility among populations is a plausible explanation for these results. However, it is still unclear if there are additional reasons for ancestry-related differential risk to COVID-19.

Investigating ancestry-associated risk may be even more challenging in studies of stratified admixed populations, in which it is necessary to adjust for population structure. This adjustment may eventually overcorrect and lead to a substantial loss in power to detect the effects of alleles that are geographically restricted and/or specific to certain ancestries (Martin *et al.,* 2018). Consequently, studies of stratified populations require appropriate approaches that assume differential association according to genetic ancestry, as exemplified by admixture mapping and association mapping (Shriner, 2017; Skotte *et al.,* 2019).

Research Strategies and Community Engagement

In only a few months, an impressive abundance of HLA-focused COVID-19 studies have been initiated or are in advanced stages. While most of them focus on specific questions regarding the relationship between genetic variation and COVID-19 outcomes, there are a myriad of different approaches. Large biobanks, donor registries, and others are leveraging pre-existing data resources to analyze genetic variants involved in immune responses, genotyped either as targeted candidate genes or in genome-wide methods. While many HLA laboratories intend to perform high-resolution genotyping in COVID-19 cohorts, multiple strategies are being employed in collecting these cohorts. It is expected that several research groups will perform genotyping of cohorts from their local clinic or hospital. Others will examine their previously genotyped HLA or KIR data of transplant donors and recipients, recontacting those individuals or accessing their medical records to ascertain COVID-19 status. This strategy is also being employed by donor registries with pre-existing genotype data. Likewise, large biobanks with either high-density SNP genotyping data suitable for imputation of HLA, or WGS or WES data from which HLA genotypes can be extracted, are being examined in light of either medical records or patient self-reported COVID-19 disease, outcomes, or symptoms.

An important endeavor to coordinate research activities is the *COVID-19 HLA & Immunogenetics Consortium* (CHIC). This effort unites the global community of HLA and immunogenetics experts and thought leaders to focus efforts, combine resources and share data to accelerate discovery in understanding the role of HLA and other immunogenetic loci in disease outcomes. This consortium currently numbers over 100 members, including HLA and immunogenetics-focused academic research scientists, HLA clinical laboratory directors, leaders of the major international immunogenetics scientific societies, Editors-in-Chief of immunogenetics-focused journals, and commercial vendors with HLA laboratory and bioinformatics products. A website describing these efforts and including information and resources for the research community can be found at hlacovid19.org. The intention of the web resources is to both centralize access to COVID-19-related HLA data through a dedicated database (HLA | COVID-19 Database portal) that includes HLA data-management and analysis tools,

as well as to connect COVID-19 researchers and clinicians seeking HLA genotyping services with the immunogenetics laboratories that can provide them.

The database is being funded under an emergency United States National Institutes of Health (NIH) COVID-19 supplement to an existing grant. Under the parent grant, tools and services have been developed for the standardized analysis, collection, exchange and storage of immunogenomic data. Under the supplemental funding, efforts have begun to apply these tools and resources for HLA data management into this public database to promote data-sharing and accelerate discovery of the relationship between HLA variation and COVID-19. Because the complexity and extreme polymorphism of the HLA region make consolidation, equivalency, analysis, and biological interpretation of HLA data challenging, a centralized resource that aggregates data from disparate sources and platforms and provides well-curated bioinformatics and analytical tools will serve to accelerate discovery. This platform is intended to centralize access to COVID-19-related HLA data and HLA data-management and analysis tools. It will serve both as a knowledge base and technical resource for HLA and immunogenetics research on the COVID-19 pandemic. Many consortium members have already committed to depositing their data in the database.

Aside from the groups centered on HLA and KIR, numerous other international, national, and regional collaborative efforts have been established to discover the host genetic determinants of COVID-19 susceptibility, severity, and outcomes, and contribute to the knowledge of the biology of SARS-CoV-2 infection and disease. The most associated variants localize in genes involved in the innate and adaptive immune responses, as exemplified by the initial results of *The COVID-19 Host Genetics Initiative* commented above in the topic *Host genetics beyond HLA and KIR*.

Concluding Remarks

The prime importance of the immune responses in susceptibility to viral infections urges deep investigation to understand the impact of host immunogenetics in COVID-19 pathogenesis.

The influence of HLA diversity on viral infections is well-documented. The involvement of HLA may occur along different paths, which includes the innate response through interaction with KIR and other less well-defined receptors of innate lymphoid cells (ILC). Additionally, HLA is directly involved in the adaptive response, playing a crucial role in antigen-presentation to $CD4^+$ and $CD8^+$ T-cells, with implications also for vaccine development. The great similarity between the paralogous genes and their astonishing polymorphism render the identification of genotypes at allele-level resolution a challenging endeavor for HLA and, especially, KIR, making targeted approaches and dedicated pipelines preferred over microarray-based or NGS-based GWAS.

Apart from HLA and their receptors, a vast collection of host molecules certainly also plays crucial roles in the immune responses to SARS-CoV-2 and COVID-19 pathogenesis. Included are viral restriction elements, toll-like receptors, cytokines and chemokines and their receptors, the complement system, among others. The genes of many (if not most) of these molecules present functional polymorphism. The GWAS as well as the screening of candidate genomic regions with a dense set of tag SNPs will identify genetic variants that modulate the host immune response to infection and COVID-19 manifestations and severity. Moreover, to uncover rare highly penetrant variants, i.e., genetic variants that follow a Mendelian or oligogenic quasi-Mendelian inheritance, segregation and linkage studies in families with unusual SARS-CoV-2 infection outcomes and the comparison of patients with contrasting extreme disease phenotypes are desirable.

For HLA and their receptors as well as for other genes that participate in immune responses, differences of allele and haplotype frequencies among populations, along with genetic variations among SARS-CoV-2 strains implicate the need for carefully studying populations of diverse geographic regions. To maximize the likelihood of discovering the disease-relevant genetic variants, patient samples covering the whole spectrum of SARS-CoV-2 infection outcomes, from asymptomatic to critical COVID-19 and characterized for disease complications, should be investigated in local and worldwide collaborations.

Integrative host immunogenetic studies will likely deliver ground-breaking insight into the pathogenesis of COVID-19, with considerable biological and clinical implications.

Acknowledgments

The authors acknowledge funds from the United States National Institutes of Health (NIH R01 GM075091; VRCA, KN, DM); Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 308783/2019-0; MLPE); United States National Institutes of Health (NIH grant 1R01NS102153; JH).

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

All authors contributed to the manuscript equally.

References

Abelin JG, Keskin DB, Sarkizova S, Hartigan CR, Zhang W, Sidney J, Stevens J, Lane W, Zhang GL, Eisenhaure TM *et al.* (2017) Mass spectrometry profiling of HLA-Associated peptidomes in mono-allelic cells enables more accurate epitope prediction. Immunity 46:315–326.

Aguiar VRC, César J, Delaneau O, Dermitzakis ET and Meyer D (2019) Expression estimation and eQTL mapping for HLA genes with a personalized pipeline. PLoS Genet 15:e1008091.

Ahlers JD and Belyakov IM (2010) Memories that last forever: Strategies for optimizing vaccine T-cell memory. Blood 115:1678–1689.

Ahmed SF, Quadeer AA and McKay MR (2020) Preliminary identification of potential vaccine targets for the COVID-19 coronavirus (SARS-CoV-2) based on SARS-CoV immunological studies. Viruses 12:254.

Akey JM, Eberle MA, Rieder MJ, Carlson CS, Shriver MD, Nickerson DA and Kruglyak L (2004) Population history and natural selection shape patterns of genetic variation in 132 genes. PLoS Biol 2:e286.

Alagarasu K, Honap T, Damle IM, Mulay AP, Shah PS and Cecilia D (2013) Polymorphisms in the oligoadenylate synthetase gene cluster and its association with clinical outcomes of dengue virus infection. Infect Genet Evol 14:390–395.

Al Omar S, Middleton D, Marshall E, Porter D, Xinarianos G, Raji O, Field JK and Christmas SE (2010) Associations between genes for killer immunoglobulin-like receptors and their ligands in patients with solid tumors. Hum Immunol 71:976–981.

Alves HV, de Moraes AG, Pepineli AC, Tiyo BT, de Lima Neto QA, Santos TS, Teixeira JJV, Ambrosio-Albuquerque EP, Sell AM and Visentainer JEL (2019) The impact of KIR/HLA genes on the risk of developing multibacillary leprosy. PLoS Negl Trop Dis 13:e0007696.

Amoroso A, Magistroni P, Vespasiano F, Bella A, Bellino S, Puoti F, Alizzi S, Vaisitti T, Boros S, Grossi PA *et al.* (2021) HLA and AB0 polymorphisms may influence SARS-CoV-2 infection and COVID-19 severity. Transplantation 105:193–200.

Anderson KM, Augusto DG, Dandekar R, Shams H, Zhao C, Yusufali T, Montero-Martín G, Marin WM, Nemat-Gorgani N, Creary LE *et al.* (2020) Killer cell immunoglobulin-like receptor variants are associated with protection from symptoms associated with more severe course in Parkinson disease. J Immunol 205:1323–1330.

Andolfo I, Russo R, Lasorsa VA, Cantalupo S, Rosato BE, Bonfiglio F, Frisso G, Abete P, Cassese GM, Servillo G *et al.* (2021) Common variants at 21q22.3 locus influence and gene expression and susceptibility to severe COVID-19. iScience 24:102322.

Andrasfay T and Goldman N (2021) Reductions in 2020 US life expectancy due to COVID-19 and the disproportionate impact on the Black and Latino populations. Proc Natl Acad Sci U S A 118:e2014746118.

André P, Denis C, Soulas C, Bourbon-Caillet C, Lopez J, Arnoux T, Bléry M, Bonnafous C, Gauthier L, Morel A *et al.* (2018) Anti-NKG2A mAb is a checkpoint inhibitor that promotes anti-tumor immunity by unleashing both T and NK Cells. Cell 175:1731–1743.e13.

Andrés AM, Hubisz MJ, Indap A, Torgerson DG, Degenhardt JD, Boyko AR, Gutenkunst RN, White TJ, Green ED, Bustamante CD *et al.* (2009) Targets of balancing selection in the human genome. Mol Biol Evol 26:2755–2764.

Anzurez A, Naka I, Miki S, Nakayama-Hosoya K, Isshiki M, Watanabe Y, Nakamura-Hoshi M, Seki S, Matsumura T, Takano T *et al.* (2021) Association of HLA-DRB1*09:01 with severe COVID-19. HLA 98:37-42.

Apps R, Qi Y, Carlson JM, Chen H, Gao X, Thomas R, Yuki Y, Del Prete GQ, Goulder P, Brumme ZL *et al.* (2013) Influence of HLA-C expression level on HIV control. Science 340:87–91.

Arora J, Pierini F, McLaren PJ, Carrington M, Fellay J and Lenz TL (2020) HLA heterozygote advantage against HIV-1 is driven by quantitative and qualitative differences in HLA allele-specific peptide presentation. Mol Biol Evol 37:639–650.

Auer ED, Van Tong H, Amorim LM, Malheiros D, Hoan NX, Issler HC, Petzl-Erler ML, Beltrame MH, Boldt ABW, Toan NL *et al.* (2020) Natural killer cell receptor variants and chronic hepatitis B virus infection in the Vietnamese population. Int J Infect Dis 96:541–547.

Augusto DG (2016) The impact of KIR polymorphism on the risk of developing cancer: Not as strong as imagined? Front Genet 7:121.

Augusto DG, Lobo-Alves SC, Melo MF, Pereira NF and Petzl-Erler ML (2012) Activating KIR and HLA Bw4 ligands are associated to decreased susceptibility to pemphigus foliaceus, an autoimmune blistering skin disease. PLoS One 7:e39991.

Augusto DG and Petzl-Erler ML (2015) KIR and HLA under pressure: Evidences of coevolution across worldwide populations. Hum Genet 134:929–940.

Bachtel ND, Umviligihozo G, Pickering S, Mota TM, Liang H, Del Prete GQ, Chatterjee P, Lee GQ, Thomas R, Brockman MA *et al.* (2018) HLA-C downregulation by HIV-1 adapts to host HLA genotype. PLoS Pathog 14:e1007257.

Bailey CC, Zhong G, Huang I-C and Farzan M (2014) IFITM-Family proteins: The cell's first line of antiviral defense. Annu Rev Virol 1:261–283.

Barquera R, Collen E, Di D, Buhler S, Teixeira J, Llamas B, Nunes JM and Sanchez-Mazas A (2020) Binding affinities of 438 HLA proteins to complete proteomes of seven pandemic viruses and distributions of strongest and weakest HLA peptide binders in populations worldwide. HLA 96:277-298.

Bassani-Sternberg M, Pletscher-Frankild S, Jensen LJ and Mann M (2015) Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. Mol Cell Proteomics 14:658–673.

Bastard P, Rosen LB, Zhang Q, Michailidis E, Hoffmann H-H, Zhang Y, Dorgham K, Philippot Q, Rosain J, Béziat V *et al.* (2020) Auto-antibodies against type I IFNs in patients with life-threatening COVID-19. Science 370:eabd4585.

Beltrame MH, Boldt ABW, Catarino SJ, Mendes HC, Boschmann SE, Goeldner I and Messias-Reason I (2015) MBL-associated serine proteases (MASPs) and infectious diseases. Mol Immunol 67:85–100.

Bitarello BD, de Filippo C, Teixeira JC, Schmidt JM, Kleinert P, Meyer D and Andrés AM (2018) Signatures of long-term balancing selection in human genomes. Genome Biol Evol 10:939–955.

Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL and Wiley DC (1987) Structure of the human class I histocompatibility antigen, HLA-A2. Nature 329:506–512.

Blackwell JM, Jamieson SE and Burgner D (2009) HLA and infectious diseases. Clin Microbiol Rev 22:370–385.

Blicher T, Kastrup JS, Buus S and Gajhede M (2005) High-resolution structure of HLA-A*1101 in complex with SARS nucleocapsid peptide. Acta Crystallogr D Biol Crystallogr 61:1031–1040.

Boegel S, Löwer M, Schäfer M, Bukur T, de Graaf J, Boisguérin V, Türeci O, Diken M, Castle JC and Sahin U (2012) HLA typing from RNA-Seq sequence reads. Genome Med 4:102.

Bordoni V, Sacchi A, Cimini E, Notari S, Grassi G, Tartaglia E, Casetti R, Giancola L, Bevilacqua N, Maeurer M *et al.* (2020) An inflammatory profile correlates with decreased frequency of cytotoxic cells in COVID-19. Clin Infect Dis 71:2272-2275.

Borrego F, Ulbrecht M, Weiss EH, Coligan JE and Brooks AG (1998) Recognition of human histocompatibility leukocyte antigen (HLA)-E complexed with HLA class I signal sequence-derived peptides by CD94/NKG2 confers protection from natural killer cell-mediated lysis. J Exp Med 187:813–818.

Bottino C, Castriconi R, Moretta L and Moretta A (2005) Cellular ligands of activating NK receptors. Trends Immunol 26:221–226.

Boudreau JE and Hsu KC (2018) Natural killer cell education in human health and disease. Curr Opin Immunol 50:102–111.

Brandt DYC, Aguiar VRC, Bitarello BD, Nunes K, Goudet J and Meyer D (2015) Mapping bias overestimates reference allele frequencies at the HLA genes in the 1000 genomes project phase I data. G3 (Bethesda) 5:931–941.

Braud VM, Allan DS, O'Callaghan CA, Söderström K, D'Andrea A, Ogg GS, Lazetic S, Young NT, Bell JI, Phillips JH *et al.* (1998) HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. Nature 391:795–799.

Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL and Wiley DC (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. Nature 364:33–39.

Bruchez A, Sha K, Johnson J, Chen L, Stefani C, McConnell H, Gaucherand L, Prins R, Matreyek KA, Hume AJ *et al.* (2020) MHC class II transactivator CIITA induces cell resistance to Ebola virus and SARS-like coronaviruses. Science 370:241–247.

Bumiller-Bini V, Oliveira-Toré CF, Carvalho TM, Kretzschmar GC, Gonçalves LB, Alencar NM, Gasparetto Filho MA, Beltrame MH and Boldt ABW (2021) MASPs at the crossroad between the complement and the coagulation cascades - the case for COVID-19. Genet Mol Biol 44:e20200199.

Burgner D, Jamieson SE and Blackwell JM (2006) Genetic susceptibility to infectious diseases: Big is beautiful, but will bigger be even better? Lancet Infect Dis 6:653–663.

Calonga-Solís V, Malheiros D, Beltrame MH, Vargas LB, Dourado RM, Issler HC, Wassem R, Petzl-Erler ML and Augusto DG (2019) Unveiling the diversity of immunoglobulin heavy constant gamma (IGHG) gene segments in Brazilian populations reveals 28 novel alleles and evidence of gene conversion and natural selection. Front Immunol 10:1161.

Carey BS, Poulton KV and Poles A (2019) Factors affecting HLA expression: A review. Int J Immunogenet 46:307–320.

Caron E, Kowalewski DJ, Chiek Koh C, Sturm T, Schuster H and Aebersold R (2015) Analysis of major histocompatibility complex (MHC) immunopeptidomes using mass spectrometry. Mol Cell Proteomics 14:3105–3117.

Carrington M, Nelson GW, Martin MP, Kissner T, Vlahov D, Goedert JJ, Kaslow R, Buchbinder S, Hoots K and O'Brien SJ (1999) HLA and HIV-1: Heterozygote advantage and B*35-Cw*04 disadvantage. Science 283:1748–1752.

Castelli EC, de Castro MV, Naslavsky MS, Scliar MO, Silva NSB, Andrade HS, Souza AS, Pereira RN, Castro CFB, Mendes-Junior CT *et al.* (2021) Immunogenetics of resistance to SARS-CoV-2 infection in discordant couples. medRxiv. DOI: 10.1101/2021.04.21.21255872

Castelli EC, Paz MA, Souza AS, Ramalho J and Mendes-Junior CT (2018) Hla-mapper: An application to optimize the mapping of HLA sequences produced by massively parallel sequencing procedures. Hum Immunol 79:678–684.

Castro MV de, Santos KS, Apostolico JS, Fernandes ER, Almeida RR, Levin G, Magawa JY, S. Nunes JP, Bruni M, Yamamoto MM *et al.* (2021) Monozygotic twins discordant for severe clinical recurrence of COVID-19 show drastically distinct T cell responses to SARS-Cov-2. bioRxiv. DOI: 10.1101/2021.03.26.21253645

Chan MHM, Wong VWS, Wong CK, Chan PKS, Chu CM, Hui DSC, Suen MWM, Sung JJY, Chung SSC and Lam CWK (2004) Serum LD1 isoenzyme and blood lymphocyte subsets as prognostic indicators for severe acute respiratory syndrome. J Intern Med 255:512–518.

Channappanavar R, Fett C, Zhao J, Meyerholz DK and Perlman S (2014) Virus-specific memory CD8 T cells provide substantial protection from lethal severe acute respiratory syndrome coronavirus infection. J Virol 88:11034–11044.

Cheng Y, Cheng G, Chui CH, Lau FY, Chan PKS, Ng MHL, Sung JJY and Wong RSM (2005) ABO blood group and susceptibility to severe acute respiratory syndrome. JAMA 293:1450–1451.

Chen J, Madireddi S, Nagarkar D, Migdal M, Heiden JV, Chang D, Mukhyala K, Selvaraj S, Kadel EE, Brauer MJ *et al.* (2021) In silico tools for accurate HLA and KIR inference from clinical sequencing data empower immunogenetics on individual-patient and population scales. Brief Bioinform 22:bbaa223.

Ching JC-Y, Chan KYK, Lee EHL, Xu M-S, Ting CKP, So TMK, Sham PC, Leung GM, Peiris JSM and Khoo U-S (2010) Significance of the myxovirus resistance A (MxA) gene -123C>A single-nucleotide polymorphism in suppressed interferon beta induction of severe acute respiratory syndrome coronavirus infection. J Infect Dis 201:1899–1908.

Chi Y, Ge Y, Wu B, Zhang W, Wu T, Wen T, Liu J, Guo X, Huang C, Jiao Y *et al.* (2020) Serum cytokine and chemokine profile in relation to the severity of Coronavirus disease 2019 (COVID-19) in China. J Infect Dis 222:746-754.

Choi UY, Kang J-S, Hwang YS and Kim Y-J (2015) Oligoadenylate synthase-like (OASL) proteins: Dual functions and associations with diseases. Exp Mol Med 47:e144.

Choudhury A and Mukherjee S (2020) In silico studies on the comparative characterization of the interactions of SARS-CoV-2 spike glycoprotein with ACE-2 receptor homologs and human TLRs. J Med Virol 92:2105-2113.

Clouston SAP, Natale G and Link BG (2021) Socioeconomic inequalities in the spread of coronavirus-19 in the United States: A examination of the emergence of social inequalities. Soc Sci Med 268:113554.

Cole C, Byrne A, Adams M, Volden R and Vollmers C (2020) Complete characterization of the human immune cell transcriptome using accurate full-length cDNA sequencing. Genome Res 30:589–601.

Cooke GS and Hill AV (2001) Genetics of susceptibility to human infectious disease. Nat Rev Genet 2:967–977.

Coperchini F, Chiovato L, Croce L, Magri F and Rotondi M (2020) The cytokine storm in COVID-19: An overview of the involvement of the chemokine/chemokine-receptor system. Cytokine Growth Factor Rev 53:25–32.

Cortes A and Brown MA (2011) Promise and pitfalls of the Immunochip. Arthritis Res Ther 13:101.

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 28:715–718.

Dausset J (1958) Iso-leuko-antibodies. Acta Haematol 20:156–166.

DeGiorgio M, Lohmueller KE and Nielsen R (2014) A model-based approach for identifying signatures of ancient balancing selection in genetic data. PLoS Genet 10:e1004561.

Demenech LM, Dumith S de C, Vieira MECD and Neiva-Silva L (2020) Income inequality and risk of infection and death by COVID-19 in Brazil. Rev Bras Epidemiol 23:e200095.

De Re V, Caggiari L, De Zorzi M, Repetto O, Zignego AL, Izzo F, Tornesello ML, Buonaguro FM, Mangia A, Sansonno D *et al.* (2015) Genetic diversity of the KIR/HLA system and susceptibility to hepatitis C virus-related diseases. PLoS One 10:e0117420.

Dhanda SK, Mahajan S, Paul S, Yan Z, Kim H, Jespersen MC, Jurtz V, Andreatta M, Greenbaum JA, Marcatili P *et al.* (2019) IEDB-AR: Immune epitope database-analysis resource in 2019. Nucleic Acids Res 47:W502–W506.

Di Bona D, Aiello A, Colomba C, Bilancia M, Accardi G, Rubino R, Giannitrapani L, Tuttolomondo A, Cascio A, Caiaffa MF *et al.* (2017) KIR2DL3 and the KIR ligand groups HLA-A-Bw4 and HLA-C2 predict the outcome of hepatitis B virus infection. J Viral Hepat 24:768–775.

Dilthey A, Cox C, Iqbal Z, Nelson MR and McVean G (2015) Improved genome inference in the MHC using a population reference graph. Nat Genet 47:682–688.

Dilthey AT, Gourraud P-A, Mentzer AJ, Cereb N, Iqbal Z and McVean G (2016) High-accuracy HLA type inference from whole-genome sequencing data using population reference graphs. PLoS Comput Biol 12:e1005151.

Domínguez-Andrés J and Netea MG (2019) Impact of historic migrations and evolutionary processes on human immunity. Trends Immunol 40:1105–1119.

Donadi EA, Castelli EC, Arnaiz-Villena A, Roger M, Rey D and Moreau P (2011) Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association. Cell Mol Life Sci 68:369–395.

Dong Q-M, He Z-P, Zhuang H, Song S-J, Dai W-S, Zhang S-P, Chen Z-H and Sun J-Y (2004) [Dynamics of peripheral blood B lymphocytes and natural killer cells in patients with severe acute respiratory syndrome]. Zhonghua Liu Xing Bing Xue Za Zhi 25:695–697.

Dos Santos Francisco R, Buhler S, Nunes JM, Bitarello BD, França GS, Meyer D and Sanchez-Mazas A (2015) HLA supertype variation across populations: new insights into the role of natural selection in the evolution of HLA-A and HLA-B polymorphisms. Immunogenetics 67:651–663.

Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, Invernizzi P, Fernández J, Prati D, Baselli G, Asselta R *et al.* (2020) Genomewide association study of severe Covid-19 with respiratory failure. N Engl J Med 383:1522-1534.

Ellul MA, Benjamin L, Singh B, Lant S, Michael BD, Easton A, Kneen R, Defres S, Sejvar J and Solomon T (2020) Neurological associations of COVID-19. Lancet Neurol 19:767–783.

Enard D, Cai L, Gwennap C and Petrov DA (2016) Viruses are a dominant driver of protein adaptation in mammals. Elife 5:e12469.

Enayatkhani M, Hasaniazad M, Faezi S, Gouklani H, Davoodian P, Ahmadi N, Einakian MA, Karmostaji A and Ahmadi K (2020) Reverse vaccinology approach to design a novel multi-epitope vaccine candidate against COVID-19: An in silico study. J Biomol Struct Dyn 39:2857-2872.

Espinosa OA, Zanetti ADS, Antunes EF, Longhi FG, Matos TA de and Battaglini PF (2020) Prevalence of comorbidities in patients and mortality cases affected by SARS-CoV2: A systematic review and meta-analysis. Rev Inst Med Trop São Paulo 62:e43.

Fadista J, Kraven LM, Karjalainen J, Andrews SJ, Geller F, COVID-19 Host Genetics Initiative, Baillie JK, Wain LV, Jenkins RG and Feenstra B (2021) Shared genetic etiology between idiopathic pulmonary fibrosis and COVID-19 severity. EBioMedicine 65:103277.

Fan S, Hansen MEB, Lo Y and Tishkoff SA (2016) Going global by adapting local: A review of recent human adaptation. Science 354:54–59.

Fauriat C, Ivarsson MA, Ljunggren H-G, Malmberg K-J and Michaëlsson J (2010) Education of human natural killer cells by activating killer cell immunoglobulin-like receptors. Blood 115:1166–1174.

Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A *et al.* (2007) A whole-genome association study of major determinants for host control of HIV-1. Science 317:944–947.

Figueiredo AM de, Figueiredo DCMM de, Gomes LB, Massuda A, Gil-García E, Vianna RP de T and Daponte A (2020) Social determinants of health and COVID-19 infection in Brazil: an analysis of the pandemic. Rev Bras Enferm 73:e20200673.

Forlani G, Turrini F, Ghezzi S, Tedeschi A, Poli G, Accolla RS and Tosi G (2016) The MHC-II transactivator CIITA inhibits Tat function and HIV-1 replication in human myeloid cells. J Transl Med 14:94.

Franceschini N, Fox E, Zhang Z, Edwards TL, Nalls MA, Sung YJ, Tayo BO, Sun YV, Gottesman O, Adeyemo A *et al.* (2013) Genome-wide association analysis of blood-pressure traits in African-ancestry individuals reveals common associated genes in African and non-African populations. Am J Hum Genet 93:545–554.

Frazão JB, Errante PR and Condino-Neto A (2013) Toll-like receptors' pathway disturbances are associated with increased susceptibility to infections in humans. Arch Immunol Ther Exp 61:427–443.

Frodsham AJ and Hill AVS (2004) Genetics of infectious diseases. Hum Mol Genet 2:R187–94.

Fumagalli M, Sironi M, Pozzoli U, Ferrer-Admetlla A, Pattini L and Nielsen R (2011) Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. PLoS Genet 7:e1002355.

Gallego Llorente M, Jones ER, Eriksson A, Siska V, Arthur KW, Arthur JW, Curtis MC, Stock JT, Coltorti M, Pieruccini P *et al.* (2015) Ancient Ethiopian genome reveals extensive Eurasian admixture throughout the African continent. Science 350:820–822.

Gao X, Jiao Y, Wang L, Liu X, Sun W, Cui B, Chen Z and Zhao Y (2010) Inhibitory KIR and specific HLA-C gene combinations confer susceptibility to or protection against chronic hepatitis B. Clin Immunol 137:139–146.

Geffard E, Limou S, Walencik A, Daya M, Watson H, Torgerson D, Barnes KC, CAAPA, Cesbron Gautier A, Gourraud P-A *et al.* (2020) Easy-HLA: A validated web application suite to reveal the full details of HLA typing. Bioinformatics 36:2157–2164.

Gentile DA, Doyle WJ, Zeevi A, Howe-Adams J, Kapadia S, Trecki J and Skoner DP (2003) Cytokine gene polymorphisms moderate illness severity in infants with respiratory syncytial virus infection. Hum Immunol 64:338–344.

Gfeller D and Bassani-Sternberg M (2018) Predicting antigen presentation-what could we learn from a million peptides? Front Immunol 9:1716.

Gontijo CC, Mendes FM, Santos CA, Klautau-Guimarães M de N, Lareu MV, Carracedo Á, Phillips C and Oliveira SF (2018) Ancestry analysis in rural Brazilian populations of African descent. Forensic Sci Int Genet 36:160–166.

Gonzalez-Galarza FF, McCabe A, Santos EJMD, Jones J, Takeshita L, Ortega-Rivera ND, Cid-Pavon GMD, Ramsbottom K, Ghattaoraya G, Alfirevic A *et al.* (2020) Allele frequency net database (AFND) 2020 update: Gold-standard data classification, open access genotype data and new query tools. Nucleic Acids Res 48:D783–D788.

Grifoni A, Sidney J, Zhang Y, Scheuermann RH, Peters B and Sette A (2020) A sequence homology and bioinformatic approach can predict candidate targets for immune responses to SARS-CoV-2. Cell Host Microbe 27:671–680.e2.

Gusatti C de S, Costi C, de Medeiros RM, Halon ML, Grandi T, Medeiros AFR, da Silva CMD, Rodenbusch R, Silva MSN, Niel C *et al.* (2016) Association between cytokine gene polymorphisms and outcome of hepatitis B virus infection in southern Brazil. J Med Virol 88:1759–1766.

Guselnikov SV and Taranin AV (2019) Unraveling the LRC evolution in mammals: IGSF1 and A1BG provide the keys. Genome Biol Evol 11:1586–1601.

Harrison EM, Docherty AB, Barr B, Buchan I, Carson G, Drake TM, Dunning J, Fairfield CJ, Gamble C, Green CA *et al.* (2020) Ethnicity and Outcomes from COVID-19: The ISARIC CCP-UK prospective Observational Cohort Study of Hospitalised Patients. Lancet. DOI: 10.2139/ssrn.3618215

Hawkins D (2020) Differential occupational risk for COVID-19 and other infection exposure according to race and ethnicity. Am J Ind Med 63:817–820.

Hawkins RB, Charles EJ and Mehaffey JH (2020) Socio-economic status and COVID-19-related cases and fatalities. Public Health 189:129–134.

Hedrick PW (2002) Pathogen resistance and genetic variation at MHC loci. Evolution 56:1902–1908.

He J, Feng D, de Vlas SJ, Wang H, Fontanet A, Zhang P, Plancoulaine S, Tang F, Zhan L, Yang H *et al.* (2006) Association of SARS susceptibility with single nucleic acid polymorphisms of OAS1 and MxA genes: a case-control study. BMC Infect Dis 6:106.

Herberman RB and Ortaldo JR (1981) Natural killer cells: Their roles in defenses against disease. Science 214:24–30.

Hernández-Cordero AI, Li X, Milne S, Yang CX, Bossé Y, Joubert P, Timens W, van den Berge M, Nickle D, Hao K *et al.* (2021) Multi-omics highlights ABO plasma protein as a causal risk factor for COVID-19. Hum Genet 140:969-979.

Hill AVS (2012) Evolution, revolution and heresy in the genetics of infectious disease susceptibility. Philos Trans R Soc Lond B Biol Sci 367:840–849.

Hindorff LA, Bonham VL, Brody LC, Ginoza MEC, Hutter CM, Manolio TA and Green ED (2018) Prioritizing diversity in human genomics research. Nat Rev Genet 19:175–185.

Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu N-H, Nitsche A *et al.* (2020) SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell 181:271–280.e8.

Hollenbach JA, Pando MJ, Caillier SJ, Gourraud P-A and Oksenberg JR (2016) The killer immunoglobulin-like receptor KIR3DL1 in combination with HLA-Bw4 is protective against multiple sclerosis in African Americans. Genes Immun 17:199–202. doi: 10.1038/gene.2016.5

Holuka C, Merz MP, Fernandes SB, Charalambous EG, Seal SV, Grova N and Turner JD (2020) The COVID-19 pandemic: Does our early life environment, life trajectory and socioeconomic status determine disease susceptibility and severity? Int J Mol Sci 21:5094.

Horton R, Coggill P, Miretti MM, Sambrook JG, Traherne JA, Ward R, Sims S, Palmer S, Sehra H, Harrow J *et al.* (2006) The LRC haplotype project: A resource for killer immunoglobulin-like receptor-linked association studies. Tissue Antigens 68:450–452.

Hovanessian AG and Justesen J (2007) The human 2'-5'oligoadenylate synthetase family: Unique interferon-inducible enzymes catalyzing 2'-5' instead of 3'-5' phosphodiester bond formation. Biochimie 89:779–788.

Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X *et al.* (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395:497–506.

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

International HIV Controllers Study, Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PIW, Walker BD, Ripke S, Brumme CJ, Pulit SL *et al.* (2010) The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science 330:1551–1557.

Iturrieta-Zuazo I, Rita CG, García-Soidán A, de Malet Pintos-Fonseca A, Alonso-Alarcón N, Pariente-Rodríguez R, Tejeda-Velarde A, Serrano-Villar S, Castañer-Alabau JL and Nieto-Gañán I (2020) Possible role of HLA class-I genotype in SARS-CoV-2 infection and progression: A pilot study in a cohort of Covid-19 Spanish patients. Clin Immunol 219:108572.

Java A, Apicelli AJ, Liszewski MK, Coler-Reilly A, Atkinson JP, Kim AH and Kulkarni HS (2020) The complement system in COVID-19: Friend and foe? JCI Insight 5:e140711.

Jiang M, Guo Y, Luo Q, Huang Z, Zhao R, Liu S, Le A, Li J and Wan L (2020) T-Cell subset counts in peripheral blood can be used as discriminatory biomarkers for diagnosis and severity prediction of Coronavirus disease 2019. J Infect Dis 222:198–202.

Jia X, Han B, Onengut-Gumuscu S, Chen W-M, Concannon PJ, Rich SS, Raychaudhuri S and de Bakker PIW (2013) Imputing amino acid polymorphisms in human leukocyte antigens. PLoS One 8:e64683.

Jones EY, Fugger L, Strominger JL and Siebold C (2006) MHC class II proteins and disease: A structural perspective. Nat Rev Immunol 6:271–282.

Joshi A, Joshi BC, Mannan MA-U and Kaushik V (2020) Epitope based vaccine prediction for SARS-COV-2 by deploying immuno-informatics approach. Inform Med Unlocked 19:100338.

Josset L, Menachery VD, Gralinski LE, Agnihothram S, Sova P, Carter VS, Yount BL, Graham RL, Baric RS and Katze MG (2013) Cell host response to infection with novel human coronavirus EMC predicts potential antivirals and important differences with SARS coronavirus. MBio 4:e00165–13.

Kachuri L, Francis SS, Morrison ML, Wendt GA, Bossé Y, Cavazos TB, Rashkin SR, Ziv E and Witte JS (2020) The landscape of host genetic factors involved in immune response to common viral infections. Genome Med 12:93.

Kalita P, Padhi AK, Zhang KYJ and Tripathi T (2020) Design of a peptide-based subunit vaccine against novel coronavirus SARS-CoV-2. Microb Pathog 145:104236.

Karakus U, Thamamongood T, Ciminski K, Ran W, Günther SC, Pohl MO, Eletto D, Jeney C, Hoffmann D, Reiche S *et al.* (2019) MHC class II proteins mediate cross-species entry of bat influenza viruses. Nature 567:109–112.

Karlsson EK, Kwiatkowski DP and Sabeti PC (2014) Natural selection and infectious disease in human populations. Nat Rev Genet 15:379–393.

Kaur G, Gras S, Mobbs JI, Vivian JP, Cortes A, Barber T, Kuttikkatte SB, Jensen LT, Attfield KE, Dendrou CA *et al.* (2017) Structural and regulatory diversity shape HLA-C protein expression levels. Nat Commun 8:15924.

Kawashima Y, Pfafferott K, Frater J, Matthews P, Payne R, Addo M, Gatanaga H, Fujiwara M, Hachiya A, Koizumi H *et al.* (2009) Adaptation of HIV-1 to human leukocyte antigen class I. Nature 458:641–645.

Kiessling R, Klein E, Pross H and Wigzell H (1975) "Natural" killer cells in the mouse II cytotoxic cells with specificity for mouse moloney leukemia cells characteristics of the killer cell. Eur J Immunol 5:117–121.

Kim H-J, Choi H-B, Jang J-P, Baek I-C, Choi E-J, Park M, Kim T-G and Oh S-T (2014) HLA-Cw polypmorphism and killer cell immunoglobulin-like receptor (KIR) gene analysis in Korean colorectal cancer patients. Int J Surg 12:815–820.

Kimura L, Ribeiro-Rodrigues EM, De Mello Auricchio MTB, Vicente JP, Batista Santos SE and Mingroni-Netto RC (2013) Genomic ancestry of rural African-derived populations from Southeastern Brazil. Am J Hum Biol 25:35–41.

Kiyotani K, Toyoshima Y, Nemoto K and Nakamura Y (2020) Bioinformatic prediction of potential T cell epitopes for SARS-Cov-2. J Hum Genet 65:569–575.

Klein J (1986) Natural history of the major histocompatibility complex. John Wiley and Sons, New York, USA, 775 p.

Klein J and Sato A (2000) The HLA system first of two parts. N Engl J Med 343:702–709.

Kotfis K, Williams Roberson S, Wilson JE, Dabrowski W, Pun BT and Ely EW (2020) COVID-19: ICU delirium management during SARS-CoV-2 pandemic. Crit Care 24:176.

Kraemer T, Blasczyk R and Bade-Doeding C (2014) HLA-E: A novel player for histocompatibility. J Immunol Res 2014:352160.

Kulkarni S, Martin MP and Carrington M (2008) The yin and yang of HLA and KIR in human disease. Semin Immunol 20:343–352.

Kulkarni S, Qi Y, O'hUigin C, Pereyra F, Ramsuran V, McLaren P, Fellay J, Nelson G, Chen H, Liao W *et al.* (2013) Genetic interplay between HLA-C and MIR148A in HIV control and Crohn disease. Proc Natl Acad Sci U S A 110:20705–20710.

Kulkarni S, Savan R, Qi Y, Gao X, Yuki Y, Bass SE, Martin MP, Hunt P, Deeks SG, Telenti A *et al.* (2011) Differential microRNA regulation of HLA-C expression and its association with HIV control. Nature 472:495–498.

Kwok AJ, Mentzer A and Knight JC (2020) Host genetics and infectious disease: New tools, insights and translational opportunities. Nat Rev Genet 22:137-153.

Lam SD, Bordin N, Waman VP, Scholes HM, Ashford P, Sen N, van Dorp L, Rauer C, Dawson NL, Pang CSM *et al.* (2020) SARS-CoV-2 spike protein predicted to form complexes with host receptor protein orthologues from a broad range of mammals. Sci Rep 10:16471.

La Porta CAM and Zapperi S (2020) Estimating the binding of Sars-CoV-2 peptides to HLA class I in human subpopulations using artificial neural networks. Cell Syst 11:412-417.e2.

Lee CH and Koohy H (2020) In silico identification of vaccine targets for 2019-nCoV. F1000Res 9:145.

Lee H and Kingsford C (2018) Kourami: graph-guided assembly for novel human leukocyte antigen allele discovery. Genome Biol 19:16.

Lee N, Llano M, Carretero M, Ishitani A, Navarro F, López-Botet M and Geraghty DE (1998) HLA-E is a major ligand for the natural killer inhibitory receptor CD94/NKG2A. Proc Natl Acad Sci U S A 95:5199–5204.

Lee W, Plant K, Humburg P and Knight JC (2018) AltHapAlignR: Improved accuracy of RNA-seq analyses through the use of alternative haplotypes. Bioinformatics 34:2401–2408.

Lefranc M-P and Lefranc G (2020) Immunoglobulins or antibodies: IMGT bridging genes, structures and functions. Biomedicines 8:319.

Leite M de M, Gonzalez-Galarza FF, Silva BCC da, Middleton D and Santos EJMD (2021) Predictive immunogenetic markers in COVID-19. Hum Immunol 82:247–254.

Lenz TL, Spirin V, Jordan DM and Sunyaev SR (2016) Excess of deleterious mutations around HLA genes reveals evolutionary cost of balancing selection. Mol Biol Evol 33:2555–2564.

Levin AM, Adrianto I, Datta I, Iannuzzi MC, Trudeau S, McKeigue P, Montgomery CG and Rybicki BA (2014) Performance of HLA allele prediction methods in African Americans for class II genes HLA-DRB1, -DQB1, and -DPB1. BMC Genet 15:72.

Liao X, Xie H, Li S, Ye H, Li S, Ren K, Li Y, Xu M, Lin W, Duan X *et al.* (2020) 2', 5'-Oligoadenylate Synthetase 2 (OAS2) Inhibits Zika Virus replication through activation of type I IFN signaling pathway. Viruses 12:418.

Li F, Wei H, Wei H, Gao Y, Xu L, Yin W, Sun R and Tian Z (2013) Blocking the natural killer cell inhibitory receptor NKG2A increases activity of human natural killer cells and clears hepatitis B virus infection in mice. Gastroenterology 144:392–401.

Liu N, Zhang T, Ma L, Zhang H, Wang H, Wei W, Pei H and Li H (2020) The impact of ABO blood group on COVID-19 infection risk and mortality: A systematic review and meta-analysis. Blood Rev 48:100785.

Li Y, Si L, Zhai Y, Hu Y, Hu Z, Bei J-X, Xie B, Ren Q, Cao P, Yang F *et al.* (2016) Genome-wide association study identifies 8p21.3 associated with persistent hepatitis B virus infection among Chinese. Nat Commun 7:11664.

Ljunggren H-G and Kärre K (1990) In search of the "missing self": MHC molecules and NK cell recognition. Immunol Today 11:237–244.

Lorente L, Martín MM, Franco A, Barrios Y, Cáceres JJ, Solé-Violán J, Perez A, Marcos Y Ramos JA, Ramos-Gómez L, Ojeda N *et al.* (2021) [HLA genetic polymorphisms and prognosis of patients with COVID-19]. Med Intensiva 45:96–103.

Machiela MJ and Chanock SJ (2015) LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics 31:3555–3557.

Malaquias MAS, Gadotti AC, Motta-Junior J da S, Martins APC, Azevedo MLV, Benevides APK, Cézar-Neto P, Panini do Carmo LA, Zeni RC, Raboni SM *et al.* (2020) The role of the lectin pathway of the complement system in SARS-CoV-2 lung injury. Transl Res 231:55-63.

Marchan J (2020) Conserved HLA binding peptides from five non-structural proteins of SARS-CoV-2 – An *in silico glance*. Hum Immunol 81:588–595.

Marin WM, Dandekar R, Augusto DG, Yusufali T, Heyn B, Hofmann J, Lange V, Sauter J, Norman PJ and Hollenbach JA (2021) High-throughput interpretation of killer-cell immunoglobulin-like receptor short-read sequencing data with PING. bioRxiv. DOI: 10.1101/2021.03.24.436770

Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, Daly MJ, Bustamante CD and Kenny EE (2017) Human demographic history impacts genetic risk prediction across diverse populations. Am J Hum Genet 100:635–649.

Martin ER, Tunc I, Liu Z, Slifer SH, Beecham AH and Beecham GW (2018) Properties of global- and local-ancestry adjustments in genetic association tests in admixed populations. Genet Epidemiol 42:214–229.

Martin MP, Qi Y, Gao X, Yamada E, Martin JN, Pereyra F, Colombo S, Brown EE, Shupert WL, Phair J *et al.* (2007) Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. Nat Genet 39:733–740.

McIntosh K, Hirsch MS and Bloom A (2020) Coronavirus disease 2019 (COVID-19): Epidemiology, virology, and prevention, https://www.uptodate.com/contents/coronavirus-disease-2019-covid-19-epidemiology-virology-and-prevention (accessed 7 July 2020).

Menachery VD, Schäfer A, Burnum-Johnson KE, Mitchell HD, Eisfeld AJ, Walters KB, Nicora CD, Purvine SO, Casey CP, Monroe ME *et al.* (2018) MERS-CoV and H5N1 influenza virus antagonize antigen presentation by altering the epigenetic landscape. Proc Natl Acad Sci U S A 115:E1012–E1021.

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

Meyer D and Nunes K (2017) HLA imputation, what is it good for? Hum Immunol 78:239–241.

Middleton D, Vilchez JR, Cabrera T, Meenagh A, Williams F, Halfpenny I, Maleno I, Ruiz-Cabello F, Lopez-Nevot MA and Garrido F (2007) Analysis of KIR gene frequencies in HLA class I characterised bladder, colorectal and laryngeal tumours. Tissue Antigens 69:220–226.

Moore JB and June CH (2020) Cytokine release syndrome in severe COVID-19. Science 368:473–474.

Moriguchi T, Harii N, Goto J, Harada D, Sugawara H, Takamino J, Ueno M, Sakata H, Kondo K, Myose N *et al.* (2020) A first case of meningitis/encephalitis associated with SARS-Coronavirus-2. Int J Infect Dis 94:55–58.

Mukherjee S, Tworowski D, Detroja R, Mukherjee SB and Frenkel-Morgenstern M (2020) Immunoinformatics and structural analysis for identification of immunodominant epitopes in SARS-CoV-2 as potential vaccine targets. Vaccines (Basel) 8:290.

Musemwa N and Gadegbeku CA (2017) Hypertension in African Americans. Curr Cardiol Rep 19:129.

Nakanishi T, Pigazzini S, Degenhardt F, Cordioli M, Butler-Laporte G, Maya-Miles D, Nafría-Jiménez B, Bouysran Y, Niemi M, Palom A *et al.* (2021) Age-dependent impact of the major common genetic risk factor for COVID-19 on severity and mortality. medRxiv. DOI: 10.1101/2021.03.07.21252875

Nariai N, Kojima K, Saito S, Mimori T, Sato Y, Kawai Y, Yamaguchi-Kabata Y, Yasuda J and Nagasaki M (2015) HLA-VBSeq: Accurate HLA typing at full resolution from whole-genome sequencing data. BMC Genomics 16:S7.

Naslavsky MS, Scliar MO, Yamamoto GL, Wang JYT, Zverinova S, Karp T, Nunes K, Ceroni JRM, de Carvalho DL, da Silva Simões CE *et al.* (2020) Whole-genome sequencing of 1,171 elderly admixed individuals from the largest Latin American metropolis (São Paulo, Brazil). bioRxiv. DOI: 10.1101/2020.09.15.298026

National Research Project for SARS, Beijing Group (2004) The involvement of natural killer cells in the pathogenesis of severe acute respiratory syndrome. Am J Clin Pathol 121:507–511.

Neville MJ, Lee W, Humburg P, Wong D, Barnardo M, Karpe F and Knight JC (2017) High resolution HLA haplotyping by imputation for a British population bioresource. Hum Immunol 78:242–251.

Nguyen A, David JK, Maden SK, Wood MA, Weeder BR, Nellore A and Thompson RF (2020) Human leukocyte antigen susceptibility map for severe acute respiratory syndrome Coronavirus 2. J Virol 94:e00510-20.

Nordgren J and Svensson L (2019) Genetic susceptibility to human norovirus infection: An update. Viruses 11:226.

Norman PJ, Hollenbach JA, Nemat-Gorgani N, Marin WM, Norberg SJ, Ashouri E, Jayaraman J, Wroblewski EE, Trowsdale J, Rajalingam R *et al.* (2016) Defining KIR and HLA class I genotypes at highest resolution via high-throughput sequencing. Am J Hum Genet 99:375–391.

Novelli A, Andreani M, Biancolella M, Liberatoscioli L, Passarelli C, Colona VL, Rogliani P, Leonardis F, Campana A, Carsetti R *et al.* (2020) HLA allele frequencies and susceptibility to COVID-19 in a group of 99 Italian patients. Hladnikia 96:610–614.

Nunnari G, Sanfilippo C, Castrogiovanni P, Imbesi R, Li Volti G, Barbagallo I, Musumeci G and Di Rosa M (2020) Network perturbation analysis in human bronchial epithelial cells following SARS-CoV2 infection. Exp Cell Res 395:112204.

Ogimi C, Kim YJ, Martin ET, Huh HJ, Chiu C-H and Englund JA (2020) What's new with the old Coronaviruses? J Pediatric Infect Dis Soc 9:210–217.

Okada Y, Momozawa Y, Ashikawa K, Kanai M, Matsuda K, Kamatani Y, Takahashi A and Kubo M (2015) Construction of a population-specific HLA imputation reference panel and its application to Graves' disease risk in Japanese. Nat Genet 47:798–802.

Older Aguilar AM, Guethlein LA, Adams EJ, Abi-Rached L, Moesta AK and Parham P (2010) Coevolution of killer cell Ig-like receptors with HLA-C to become the major variable regulators of human NK cells. J Immunol 185:4238–4251.

Ong E, Wong MU, Huffman A and He Y (2020) COVID-19 coronavirus vaccine design using reverse vaccinology and machine learning. bioRxiv. DOI: 10.1101/2020.03.20.000141

Orenbuch R, Filip I, Comito D, Shaman J, Pe'er I and Rabadan R (2020) ArcasHLA: High-resolution HLA typing from RNAseq. Bioinformatics 36:33–40.

Ou G, Liu X, Yang L, Yu H, Ji X, Liu F, Xu H, Qian L, Wang J and Liu Z (2019) Relationship between HLA-DPA1 mRNA expression and susceptibility to hepatitis B. J Viral Hepat 26:155–161.

Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, Walker S, Parkinson N, Fourman MH, Russell CD *et al.* (2020) Genetic mechanisms of critical illness in Covid-19. Nature 591:92-98.

Pappas DJ, Lizee A, Paunic V, Beutner KR, Motyer A, Vukcevic D, Leslie S, Biesiada J, Meller J, Taylor KD *et al.* (2018) Significant variation between SNP-based HLA imputations in diverse populations: The last mile is the hardest. Pharmacogenomics J 18:367–376.

Parham P (2004) Killer cell immunoglobulin-like receptor diversity: Balancing signals in the natural killer cell response. Immunol Lett 92:11–13.

Parolini F, Biswas P, Serena M, Sironi F, Muraro V, Guizzardi E, Cazzoletti L, Scupoli MT, Gibellini D, Ugolotti E *et al.* (2018) Stability and expression levels of HLA-C on the cell membrane modulate HIV-1 infectivity. J Virol 92:e01711-17.

Pathak GA, Singh K, Miller-Fleming TW, Wendt F, Ehsan N, Hou K, Johnson R, Lu Z, Gopalan S, Dimbou LY *et al.* (2020) Integrative analyses identify susceptibility genes underlying COVID-19 hospitalization. medRxiv. DOI: 10.1101/2020.12.07.20245308

Patil A, Tripathy JP, Deshmukh V, Sontakke B and Tripathi SC (2020) SeXX and COVID-19: Tussle between the two. Monaldi Arch Chest Dis 90.

Pendu JL, Breiman A, Rocher J, Dion M and Ruvoën-Clouet N (2021) ABO blood types and COVID-19: Spurious, anecdotal, or truly important relationships? A reasoned review of available data. Viruses 13:160.

Persson G, Jørgensen N, Nilsson LL, Andersen LHJ and Hviid TVF (2020) A role for both HLA-F and HLA-G in reproduction and during pregnancy? Hum Immunol 81:127–133.

Petzl-Erler ML (2020) Beyond the HLA polymorphism: A complex pattern of genetic susceptibility to pemphigus. Genet Mol Biol 43:e20190369.

Pickett BE, Sadat EL, Zhang Y, Noronha JM, Squires RB, Hunt V, Liu M, Kumar S, Zaremba S, Gu Z *et al.* (2012) ViPR: an open bioinformatics database and analysis resource for virology research. Nucleic Acids Res 40:D593–8.

Podhorzer A, Dirchwolf M, Machicote A, Belen S, Montal S, Paz S, Fainboim H, Podestá LG and Fainboim L (2017) The clinical features of patients with chronic hepatitis C virus infections are associated with killer cell immunoglobulin-like receptor genes and their expression on the surface of natural killer cells. Front Immunol 8:1912.

Qin C, Zhou L, Hu Z, Zhang S, Yang S, Tao Y, Xie C, Ma K, Shang K, Wang W *et al.* (2020) Dysregulation of immune response in patients with COVID-19 in Wuhan, China. Clin Infect Dis 71:762-768.

Rabaan AA, Al-Ahmed SH, Haque S, Sah R, Tiwari R, Malik YS, Dhama K, Yatoo MI, Bonilla-Aldana DK and Rodriguez-Morales AJ (2020) SARS-CoV-2, SARS-CoV, and MERS-COV: A comparative overview. Infez Med 28:174–184.

Ramlall V, Thangaraj PM, Meydan C, Foox J, Butler D, Kim J, May B, De Freitas JK, Glicksberg BS, Mason CE *et al.* (2020) Immune complement and coagulation dysfunction in adverse outcomes of SARS-CoV-2 infection. Nat Med 26:1609–1615.

Ramsuran V, Kulkarni S, O'huigin C, Yuki Y, Augusto DG, Gao X and Carrington M (2015) Epigenetic regulation of differential HLA-A allelic expression levels. Hum Mol Genet 24:4268–4275.

Ramsuran V, Naranbhai V, Horowitz A, Qi Y, Martin MP, Yuki Y, Gao X, Walker-Sperling V, Del Prete GQ, Schneider DK *et al.* (2018) Elevated expression impairs HIV control through inhibition of NKG2A-expressing cells. Science 359:86–90.

Reis ES, Mastellos DC, Hajishengallis G and Lambris JD (2019) New insights into the immune functions of complement. Nat Rev Immunol 19:503–516.

Rivino L, Tan AT, Chia A, Kumaran EAP, Grotenbreg GM, MacAry PA and Bertoletti A (2013) Defining CD8+ T cell determinants during human viral infection in populations of Asian ethnicity. J Immunol 191:4010–4019.

Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P and Marsh SGE (2020) IPD-IMGT/HLA Database. Nucleic Acids Res 48:D948–D955.

Sabeti PC, Reich DE, Higgins JM, Levine HZP, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ *et al.* (2002) Detecting recent positive selection in the human genome from haplotype structure. Nature 419:832–837.

Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O, Palma A, Mikkelsen TS, Altshuler D and Lander ES (2006) Positive natural selection in the human lineage. Science 312:1614–1620.

Sams AJ, Dumaine A, Nédélec Y, Yotova V, Alfieri C, Tanner JE, Messer PW and Barreiro LB (2016) Adaptively introgressed neandertal haplotype at the OAS locus functionally impacts innate immune responses in humans. Genome Biol 17:246.

Sanami S, Zandi M, Pourhossein B, Mobini G-R, Safaei M, Abed A, Arvejeh PM, Chermahini FA and Alizadeh M (2020) Design of a multi-epitope vaccine against SARS-CoV-2 using immunoinformatics approach. Int J Biol Macromol 164:871–883.

Sanchez-Mazas A (2020) HLA studies in the context of coronavirus outbreaks. Swiss Med Wkly 150:w20248.

Sankararaman S, Mallick S, Dannemann M, Prüfer K, Kelso J, Pääbo S, Patterson N and Reich D (2014) The genomic landscape of neanderthal ancestry in present-day humans. Nature 507:354–357.

Seibold MA, Wise AL, Speer MC, Steele MP, Brown KK, Loyd JE, Fingerlin TE, Zhang W, Gudmundsson G, Groshong SD *et al.* (2011) A common MUC5B promoter polymorphism and pulmonary fibrosis. N Engl J Med 364:1503–1512.

Seldin MF, Pasaniuc B and Price AL (2011) New approaches to disease mapping in admixed populations. Nat Rev Genet 12:523–528.

Sette A and Sidney J (1999) Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. Immunogenetics 50:201–212.

Shan Z, Huang J, Liao Q, Huang K, Wang M, Xu R, Tang X, Zhang W, Nelson K, Fu Y *et al.* (2018) Association of killer cell immunoglobulin-like receptors with spontaneous clearance of hepatitis C virus in the Chinese population. Transfusion 58:1028–1035.

Shelton JF, Shastri AJ, Ye C, Weldon CH, Filshtein-Sonmez T, Coker D, Symons A, Esparza-Gordillo J, Aslibekyan S and Auton A (2021) Trans-ancestry analysis reveals genetic and nongenetic associations with COVID-19 susceptibility and severity. Nat Genet 53:801-808.

Shkurnikov M, Nersisyan S, Jankevic T, Galatenko A, Gordeev I, Vechorko V and Tonevitsky A (2021) Association of HLA class I genotypes with severity of Coronavirus disease-19. Front Immunol 12:641900.

Shriner D (2017) Overview of Admixture Mapping. Curr Protoc Hum Genet 94:1.23.1–1.23.8.

Skotte L, Jørsboe E, Korneliussen TS, Moltke I and Albrechtsen A (2019) Ancestry-specific association mapping in admixed populations. Genet Epidemiol 43:506–521.

Song Z, Xu Y, Bao L, Zhang L, Yu P, Qu Y, Zhu H, Zhao W, Han Y and Qin C (2019) From SARS to MERS, thrusting Coronaviruses into the spotlight. Viruses 11:59.

Souza AM de, Resende SS, Sousa TN de and Brito CFA de (2019) A systematic scoping review of the genetic ancestry of the Brazilian population. Genet Mol Biol 42:495–508.

Staeheli P and Haller O (2018) Human MX2/MxB: A potent interferon-induced postentry inhibitor of Herpesviruses and HIV-1. J Virol 92:e00709-18.

Stowell CP and Stowell SR (2019a) Biologic roles of the ABH and Lewis histo-blood group antigens part I: Infection and immunity. Vox Sang 114:426–442.

Stowell SR and Stowell CP (2019b) Biologic roles of the ABH and Lewis histo-blood group antigens part II: thrombosis, cardiovascular disease and metabolism. Vox Sang 114:535–552.

Sun D-W, Zhang D, Tian R-H, Li Y, Wang Y-S, Cao J, Tang Y, Zhang N, Zan T, Gao L *et al.* (2020) The underlying changes and predicting role of peripheral blood inflammatory cells in severe COVID-19 patients: A sentinel? Clin Chim Acta 508:122–129.

Suzuki Y, Hamamoto Y, Ogasawara Y, Ishikawa K, Yoshikawa Y, Sasazuki T and Muto M (2004) Genetic polymorphisms of killer cell immunoglobulin-like receptors are associated with susceptibility to psoriasis vulgaris. J Invest Dermatol 122:1133–1136.

Sylvester-Hvid C, Nielsen M, Lamberth K, Røder G, Justesen S, Lundegaard C, Worning P, Thomadsen H, Lund O, Brunak S *et al.* (2004) SARS CTL vaccine candidates; HLA supertype-, genome-wide scanning and biochemical validation. Tissue Antigens 63:395–400.

Sze S, Pan D, Nevill CR, Gray LJ, Martin CA, Nazareth J, Minhas JS, Divall P, Khunti K, Abrams KR *et al.* (2020) Ethnicity and clinical outcomes in COVID-19: A systematic review and meta-analysis. EClinicalMedicine 29:100630.

Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M and Kohlbacher O (2014) OptiType: precision HLA typing from next-generation sequencing data. Bioinformatics 30:3310–3316.

Tahir Ul Qamar M, Shahid F, Aslam S, Ashfaq UA, Aslam S, Fatima I, Fareed MM, Zohaib A and Chen L-L (2020) Reverse vaccinology assisted designing of multiepitope-based subunit vaccine against SARS-CoV-2. Infect Dis Poverty 9:132.

Tang J, Costello C, Keet IP, Rivers C, Leblanc S, Karita E, Allen S and Kaslow RA (1999) HLA class I homozygosity accelerates disease progression in human immunodeficiency virus type 1 infection. AIDS Res Hum Retroviruses 15:317–324.

Thomas R, Apps R, Qi Y, Gao X, Male V, O'hUigin C, O'Connor G, Ge D, Fellay J, Martin JN *et al.* (2009) HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. Nat Genet 41:1290–1294.

Thomas R, Thio CL, Apps R, Qi Y, Gao X, Marti D, Stein JL, Soderberg KA, Moody MA, Goedert JJ *et al.* (2012) A novel variant marking HLA-DP expression levels predicts recovery from hepatitis B virus infection. J Virol 86:6979–6985.

Thornton TA and Bermejo JL (2014) Local and global ancestry inference and applications to genetic association analysis for admixed populations. Genet Epidemiol 38:S5–S12.

Thorsby E (2009) A short history of HLA. Tissue Antigens 74:101–116.

Tian C, Hromatka BS, Kiefer AK, Eriksson N, Tung JY and Hinds DA (2016) Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. Nat Commun 8:599.

Tsao Y-P, Lin J-Y, Jan J-T, Leng C-H, Chu C-C, Yang Y-C and Chen S-L (2006) HLA-A*0201 T-cell epitopes in severe acute respiratory syndrome (SARS) coronavirus nucleocapsid and spike proteins. Biochem Biophys Res Commun 344:63–71.

van der Made CI, Simons A, Schuurs-Hoeijmakers J, van den Heuvel G, Mantere T, Kersten S, van Deuren RC, Steehouwer M, van Reijmersdal SV, Jaeger M *et al.* (2020) Presence of genetic variants among young men with severe COVID-19. JAMA 324:663-673.

van Moorsel CHM, van der Vis JJ, Benschop C, Ruven HJT, Quanjel M and Grutters JC (2020) The MUC5B promotor polymorphismassociates with severe Covid-19. medRxiv. DOI: 10.1101/2020.05.12.20099333

Varatharaj A, Thomas N, Ellul MA, Davies NWS, Pollak TA, Tenorio EL, Sultan M, Easton A, Breen G, Zandi M *et al.* (2020) Neurological and neuropsychiatric complications of COVID-19 in 153 patients: a UK-wide surveillance study. The Lancet Psychiatry 7:875–882.

Vastrad B, Vastrad C and Tengli A (2020) Identification of potential mRNA panels for severe acute respiratory syndrome coronavirus 2 (COVID-19) diagnosis and treatment using microarray dataset and bioinformatics methods. 3 Biotech 10:422.

Vergara C, Thio CL, Johnson E, Kral AH, O'Brien TR, Goedert JJ, Mangia A, Piazzolla V, Mehta SH, Kirk GD *et al.* (2019) Multi-ancestry genome-wide association study of spontaneous clearance of hepatitis C virus. Gastroenterology 156:1496–1507.e7.

Vince N, Douillard V, Geffard E, Meyer D, Castelli EC, Mack SJ, Limou S and Gourraud P-A (2020) SNP-HLA reference consortium (SHLARC): HLA and SNP data sharing for promoting MHC-centric analyses in genomics. Genet Epidemiol 44:733-740.

Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, Wheeler DK, Sette A and Peters B (2019) The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res 47:D339–D343.

Vukcevic D, Traherne JA, Næss S, Ellinghaus E, Kamatani Y, Dilthey A, Lathrop M, Karlsen TH, Franke A, Moffatt M *et al.* (2015) Imputation of KIR types from SNP variation data. Am J Hum Genet 97:593–607.

Walsh EC, Sabeti P, Hutcheson HB, Fry B, Schaffner SF, de Bakker PIW, Varilly P, Palma AA, Roy J, Cooper R *et al.* (2006) Searching for signals of evolutionary selection in 168 genes related to immune function. Hum Genet 119:92–102.

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

Wang F, Nie J, Wang H, Zhao Q, Xiong Y, Deng L, Song S, Ma Z, Mo P and Zhang Y (2020b) Characteristics of peripheral lymphocyte subset alteration in COVID-19 Pneumonia. J Infect Dis 221:1762–1769.

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

Warren RL, Choe G, Freeman DJ, Castellarin M, Munro S, Moore R and Holt RA (2012) Derivation of HLA types from shotgun sequence datasets. Genome Med 4:95.

Wende H, Colonna M, Ziegler A and Volz A (1999) Organization of the leukocyte receptor cluster (LRC) on human chromosome 19q13.4. Mamm Genome 10:154–160.

Wilk AJ, Rustagi A, Zhao NQ, Roque J, Martínez-Colón GJ, McKechnie JL, Ivison GT, Ranganath T, Vergara R, Hollis T *et al.* (2020) A single-cell atlas of the peripheral immune response in patients with severe COVID-19. Nat Med 26: 1070-1076.

Williams AP, Bateman AR and Khakoo SI (2005) Hanging in the balance KIR and their role in disease. Mol Interv 5:226–240.

Wilson MJ, Torkar M and Trowsdale J (1997) Genomic organization of a human killer cell inhibitory receptor gene. Tissue Antigens 49:574–579.

Wissemann WT, Hill-Burns EM, Zabetian CP, Factor SA, Patsopoulos N, Hoglund B, Holcomb C, Donahue RJ, Thomson G, Erlich H *et al.* (2013) Association of Parkinson disease with structural and regulatory variants in the HLA region. Am J Hum Genet 93:984–993.

Wu Z and McGoogan JM (2020) Characteristics of and important lessons from the Coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72 314 cases from the chinese center for disease control and prevention. JAMA 323:1239-1242.

Yao Z, Zheng Z, Wu K and Junhua Z (2020) Immune environment modulation in pneumonia patients caused by coronavirus: SARS-CoV, MERS-CoV and SARS-CoV-2. Aging 12:7639–7651.

Yarmarkovich M, Warrington JM, Farrel A and Maris JM (2020) Identification of SARS-CoV-2 vaccine epitopes predicted to induce long-term population-scale immunity. Cell Reports Medicine 1:100036.

Yawata M, Yawata N, Draghi M, Partheniou F, Little A-M and Parham P (2008) MHC class I-specific inhibitory receptors and their ligands structure diverse human NK-cell repertoires toward a balance of missing self-response. Blood 112:2369–2380.

Yung Y-L, Cheng C-K, Chan H-Y, Xia JT, Lau K-M, Wong RSM, Wu AKL, Chu RW, Wong ACC, Chow EYD *et al.* (2021) Association of HLA-B22 serotype with SARS-CoV-2 susceptibility in Hong Kong Chinese patients. Hladnikia 97:127–132.

Zanin L, Saraceno G, Panciani PP, Renisi G, Signorini L, Migliorati K and Fontanella MM (2020) SARS-CoV-2 can induce brain and spine demyelinating lesions. Acta Neurochir 162:1491–1494.

Zeberg H and Pääbo S (2020) The major genetic risk factor for severe COVID-19 is inherited from neanderthals. Nature 587:610-612.

Zhang H, Zhou G, Zhi L, Yang H, Zhai Y, Dong X, Zhang X, Gao X, Zhu Y and He F (2005) Association between mannose-binding lectin gene polymorphisms and susceptibility to severe acute respiratory syndrome coronavirus infection. J Infect Dis 192:1355–1361.

Zhang Q, Bastard P, Liu Z, Le Pen J, Moncada-Velez M, Chen J, Ogishi M, Sabli IKD, Hodeib S, Korol C *et al.* (2020a) Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. Science 370:eadbd4570.

Zhang X, Feng J, Chen S, Yang H and Dong Z (2019) Synergized regulation of NK cell education by NKG2A and specific Ly49 family members. Nat Commun 10:5010.

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

Zhang Y, Zhang J, Chen Y, Luo B, Yuan Y, Huang F, Yang T, Yu F, Liu J, Liu B *et al.* (2020c) The ORF8 protein of SARS-CoV-2 mediates immune evasion through potently downregulating MHC-I. bioRxiv. DOI: 10.1101/2020.05.24.111823

Zheng M, Gao Y, Wang G, Song G, Liu S, Sun D, Xu Y and Tian Z (2020) Functional exhaustion of antiviral lymphocytes in COVID-19 patients. Cell Mol Immunol 17:533–535.

Zheng X, Shen J, Cox C, Wakefield JC, Ehm MG, Nelson MR and Weir BS (2014) HIBAG--HLA genotype imputation with attribute bagging. Pharmacogenomics J 14:192–200.

Zhou S, Butler-Laporte G, Nakanishi T, Morrison DR, Afilalo J, Afilalo M, Laurent L, Pietzner M, Kerrison N, Zhao K *et al.* (2021) A Neanderthal OAS1 isoform protects individuals of European ancestry against COVID-19 susceptibility and severity. Nat Med 27:659–667.

Internet Resources

COVID-19 DataSharing/BR FAPESP COVID-19 DataSharing/BR, http://repositoriodatasharingfapesp.uspdigital.usp.br (accessed 23 December 2020).

COVID-19 Host Genetics Initiative (2021) GWAS meta-analyses round 5, http://www.covid19hg.org/results (accessed 18 January 2021).

Ensembl IFITM genes region, http://www.ensembl.org/Homo_sapiens/Location/View?db=core;r=11:320722-320822;source=dbSNP;v=rs12252;vdb=variation;vf=83319956;time=1591125630 (accessed 2 October 2020).

Ensembl MUC5B region, http://www.ensembl.org/Homo_sapiens/Location/View?db=core;r=11:1219941-1220041;source=dbSNP;v=rs35705950;vdb=variation;vf=87816932;time=1591119800 (accessed 2 October 2020).

Ensembl MX1 and MX2 region, http://www.ensembl.org/Homo_sapiens/Location/View?db=core;g=ENSG00000157601;r=21:41419526-41459992 (accessed 2 October 2020).

Ensembl rs12252, http://www.ensembl.org/Homo_sapiens/Variation/Mappings?db=core;r=11:320272-321272;v=rs12252;vdb=variation;vf=83319956 (accessed 15 October 2020).

EpiCovid-19 http://www.epicovid19brasil.org (accessed 23 December 2020).

GTEx Portal rs12252, https://gtexportal.org/home/snp/rs12252 (accessed 15 October 2020).

HLA | COVID-19 Database portal, https://www.database.covid19.org (accessed 15 August 2020).

IPD-IMGT/HLA Statistics, http://www.ebi.ac.uk/ipd/imgt/hla/stats.html (accessed 15 October 2020).

LDlink LDproxy Tool, https://analysistools.cancer.gov/LDlink/?var=rs10774671&pop=MXL%2BPUR%2BCLM%2BPEL&r2_d=r2&window=500000&tab=ldproxy (accessed 19 April 2020).

Rede Genômica IPEC (2021) Rede Genômica IPEC/Guarapuava, https://ipec.org.br/rede-genomica-ipec-guarapuava (accessed 3 May 2021).

UniProtKB MUC5B, Q9HC84, http://www.uniprot.org/uniprot/Q9HC84 (accessed 30 September 2020).

UniProtKB MX1, P20591, http://www.uniprot.org/uniprot/P20591 (accessed 30 September 2020).

UniProtKB MX2, P20592, http://www.uniprot.org/uniprot/P20592 (accessed 30 September 2020).

UniProtKB RNAse L, P00973, http://www.uniprot.org/uniprot/P00973 (accessed 30 September 2020).

WHO COVID-19 | Brazil, http://covid19.who.int/region/amro/country/br (accessed 19 December 2020).

WHO COVID-19 Dashboard, https://covid19.who.int/ (accessed 13 April 2021).

Associate Editor: Carlos F.M. Menck