INFECTIONAL DISEASES

UDC 616.9+575.1+578.5:578.7:578.834.1

Current knowledge of genetics of COVID-19

D. A. Vologzhanin[1,2], A. S. Golota[1], T. A. Kamilova[1], S. V. Makarenko[1], S. G. Sherbak[1,2]

[1] City Hospital no. 40,
9, ul. Borisova, Sestroretsk, St Petersburg, 197706, Russian Federation
[2] St Petersburg State University,
7–9, Universitetskaya nab., St Petersburg, 199034, Russian Federation

For citation: Vologzhanin D. A., Golota A. S., Kamilova T. A., Makarenko S. V., Sherbak S. G. Current knowledge of genetics of COVID-19. *Vestnik of Saint Petersburg University. Medicine*, 2021, vol. 16, issue 4, pp. 255–272. https://doi.org/10.21638/spbu11.2021.404

The ongoing COVID-19 pandemic, caused by coronavirus SARS-CoV-2, is responsible for a reported 456,797,217 cases of COVID-19, and 6,043,094 deaths worldwide as of 12.03.2022. Following infection with SARS-CoV-2, COVID-19 clinical presentation ranges from asymptomatic or mild (~80 % of infections), to severe disease that typically requires hospitalization and assisted respiration. Innate immune responses to viral infection are also a critical determinant of disease outcome. Genetic risk factors for COVID-19 are in the early stages of study. A number of mutations and polymorphisms have been identified that affect the structure and stability of proteins — factors of susceptibility to SARS-COV-2 infection, a predisposition to the development of respiratory failure, and the need for intensive care. Most of the identified genetic factors are related to the function of the immune system. On the other hand, the genetic polymorphism of the virus itself affects the spread and severity of the course of COVID-19. The genome of the virus accumulates mutations and evolves towards increasing contagiousness, replicative ability, and evasion from the host's immune system. Genetic determinants of infection are potential therapeutic targets. Studying them will provide information for the development of drugs and vaccines to combat the pandemic.

Keywords: COVID-19, coronavirus, SARS-COV-2, genetic predisposition factors, mutation, polymorphism.

Introduction

On March 11, 2020, the World Health Organization declared COVID-19 a pandemic. During the COVID-19 pandemic, the SARS-CoV-2 coronavirus infected 456,797,217 people worldwide (as of 12 March 2022) with a registered death rate of 6,043,094 people [1].

In most patients infected with the SARS-CoV-2 coronavirus, the disease is mild or asymptomatic, while 5 % of patients with COVID-19 develop pneumonia, acute respiratory distress syndrome, septic shock, and often fatal multiple organ failure [2; 3].

Human genetics

Severe COVID-19 is a spectrum of hyperinflammatory, often fatal conditions. Susceptibility to life-threatening infections and immune-mediated diseases has a genetic component. In particular, susceptibility to respiratory viruses such as influenza is inherited and associated with specific genetic variants [4]. Revealing the molecular genetic mechanisms of this variability is of primary biological and medical importance [5]. The determinants of the severity of COVID-19 are almost entirely dependent on host factors, not the virus [6].

D. Ellinghaus and other members of the Severe COVID-19 GWAS Group from Germany, Sweden, Norway, Italy, Spain, Australia, and Lithuania performed a meta-analysis of genome-wide association studies (GWAS) in cohorts of hospitalized patients with severe COVID-19 (defined as respiratory failure) in 7 hospitals in the Italian and Spanish epicenters of the local peak of the epidemic, which received oxygen therapy or mechanical ventilation, and compared the data of these patients with data from healthy blood donors from the same regions [7]. The final case-control data sets comprised 835 patients and 1255 control participants from Italy and 775 patients and 950 control participants from Spain. A total of 8,965,091 single nucleotide polymorphisms (SNP) were included in the Italian cohort and 9140716 SNPs in the Spanish cohort. The study revealed associations of the severity of SARS-CoV-2 infection with polymorphism of polygenic loci 3p21.31 and 9q34.2. Cross-replicating associations were found with the variants rs11385942 (GA insertion/deletion) at the 3p21.31 locus and rs657152 (CA SNP) at the 9q34.2 locus. At the 3p21.31 locus, the association encompassed the genes *SLC6A20*, *LZTFL1*, *CCR9*, *FYCO1*, *CXCR6*, and *XCR1*, and at the 9q34.2 locus, the association signal coincided with the ABO blood group locus — an increased risk in blood group A ($p = 1.5 \times 10^{-4}$) and a protective effect in blood group O compared with other blood groups. The analyses that were corrected for age and sex corroborated the observations at both rs11385942 and rs657152. The biological mechanism underlying the effect of the rs657152 polymorphism at the ABO locus is presumably related to the production of neutralizing antibodies against viral proteins [7]. Meta-analysis showed that it is associated with susceptibility to COVID-19, but not with the severity of the disease [8].

Among the six candidate genes at the 3p21.31 locus, the most convincing is the *LZTFL1* gene with the rs11385942 variant, which is expressed at a high level in human lung cells and encodes a protein involved in the transport of proteins to primary cilia, which are subcellular organelles from microtubules that act as antennas-mechanosensors for extracellular signals. The frequency of the risk allele G at site rs11385942 is higher in ventilated patients than in those receiving supplemental oxygen only, in the main meta-analysis and gender- and age-adjusted meta-analysis. In addition, patients homozygous for the risk allele were younger than those heterozygous or homozygous for the A allele (mean age 59 years vs 66 years; $p = 0.005$) [7].

In T-lymphocytes, the LZTFL1 protein participates in the immunological synapse with antigen-presenting cells. The 3p21.31 locus contains the *SLC6A20* gene, which en-

codes a high intestinal expression transporter protein regulated by the ACE2 receptor, and genes encoding chemokine receptors, including *CXCR6*, which regulates T cell migration and the localization of lung-resident memory CD8 T cells. The *CCR9*, *XCR1*, and *FYCO1* genes are also involved in the function of dendritic and T cells [6]. Thus, the 3p21.31 gene cluster has been identified as a locus of genetic predisposition to the most severe forms of COVID-19.

Researchers from 86 clinics and laboratories in America, Europe, Asia, and Australia sequenced the exome or genome of 659 patients with severe COVID-19-associated pneumonia and 534 patients with asymptomatic or mild infection and found a significant increase in the number of loss-of-function mutations at 13 candidate loci in patients with life-threatening pneumonia compared with patients with asymptomatic or mild infection. In 3.5 % of patients aged 17 to 77 years, 24 pathogenic variants were identified underlying autosomal-recessive deficiencies in the *IRF7* (interferon regulatory factor 7) and *IFNAR1* (interferon-alpha/beta receptor alpha chain) genes and autosomal-dominant deficiencies in the *TLR3*, *UNC93B1*, *TICAM1*, *TBK1*, *IRF3*, *IRF7*, *IFNAR1*, and *IFNAR2* genes involved in TLR3- and IRF7-dependent induction and amplification of type I IFNs. The *IFNAR1* and *IFNAR2* genes are part of a cluster of immunologically important genes and encode subunits 1 and 2 of the IFN-α receptor and IFN-β, respectively, implicated in the pathophysiology of severe COVID-19. Plasmacytoid dendritic cells of patients with IRF7 deficiency do not produce type I IFN when infected with SARS-CoV-2. Fibroblasts from patients with the $TLR3^{-/-}$, $TLR3^{+/-}$, $IRF7^{-/-}$ and $IFNAR1^{-/-}$ phenotypes are susceptible to SARS-CoV-2 infection in vitro. This discovery reveals the essential role of both the double-stranded RNA sensor TLR3 and type I IFN cell-intrinsic immunity in the control of SARS-CoV-2 infection in the lungs. The exogenous type I IFN administration may have a therapeutic effect in patients with COVID-19 — carriers of a certain genotype [5].

Understanding the role of circulating proteins in infectious diseases is challenging because the infection itself often significantly alters the expression of the circulating protein and it might appear that increased levels of circulating proteins, such as cytokines, are associated with a poorer outcome, when in fact it may be a host response to infection and helps mitigate this outcome. Therefore, it is important to know the genetic determinants of protein levels, which reflect the extent to which a person is protected from severe COVID-19. A large-scale randomized study [9] conducted in the USA, Canada, Japan, Sweden, Germany, and England to search for circulating proteins that influence the susceptibility and severity of COVID-19, identified the OAS1 (oligoadenylate synthetase 1) gene associated with reduced COVID-19 susceptibility (14,134 patients/1,284,876 controls), COVID-19 hospitalizations (6,406 patients/902,088 controls) and COVID-19 death or ventilation (4,336 patients / 623,902 controls). By measuring the expression of circulating proteins, the authors demonstrated that this protective effect on the outcome of COVID-19 is provided by increased levels of the p46 OAS1 isoform and total OAS1 protein, which is consistent with the data of H. Zeberg et al. [10].

OAS proteins are part of the innate immune response against RNA viruses. They activate latent RNase L, which cleaves double-stranded RNA, an intermediate in coronavirus replication, leading to direct destruction of viral RNA. SARS-CoV-2 and other beta coronaviruses produce viral proteins that destroy OAS enzymes and counteract RNase L, which degrades viral RNA. This viral activity allows SARS-CoV-2 to evade the host's

immune response. Inhibitors of viral phosphodiesterase-12, which degrades OAS enzymes, enhance the antiviral activity of OAS. Protective isoforms of OAS1, OAS2, and OAS3 proteins increase the expression of the *IRF3* and *IRF7* genes, which are part of the interferon-inducible gene signature [9].

GWAS in 2244 critically ill COVID-19 patients with deep hypoxemic respiratory failure from 208 British hospitals confirmed significant associations of disease severity with several polymorphisms related to key mechanisms of host antiviral defense and mediators of inflammatory organ damage in COVID-19: rs10735079, rs2109069, rs2236757, rs74956615. The rs10735079 is located in the OAS gene cluster (locus 12q24.13), which encodes interferon-inducible activators of enzymes for antiviral protection OAS1, OAS2, OAS3. Transcriptome analysis of lung tissue revealed an association of COVID-19 with OAS3 expression. High levels of OAS3 in lungs and whole blood are associated with worse outcomes in critically ill COVID-19 patients, which is the opposite directional effect compared to OAS1 [8].

To study a haplotype that is protective against the severe form of COVID-19 in the already mentioned OAS gene cluster on chromosome 12, H. Zeberg et al. used the Genetics of Mortality in Critical Care and COVID-19 Host Genetics Initiative databases. This haplotype contains the variants rs2660, rs1859330, rs1859329, rs2285932, rs1293767 [11]. In addition, protective alleles rs4767027-T and rs10774671-G were found in the *OAS1* gene. Alternative splicing of *OAS1*, regulated by the rs10774671-G allele, increases the expression of the p46 isoform, which has a higher antiviral activity than the p42 isoform. Host genetic variants associated with extremely severe diseases help identify therapeutic targets. Molecules are already known that can increase the activity of *OAS1*. Interferon β-1b activates the cytokine cascade leading to an increase in the expression of the *OAS1* gene and the level of OAS1 in the blood. IFN-β1b inhalation therapy may have different effects in populations of different origins due to the presence of different genetic variants, in particular, it is more effective in populations with a higher expression of the p46 isoform [9].

The rs2109069 variant in the *DPP9* gene (dipeptidyl peptidase 9, locus 19p13.3) is associated with idiopathic pulmonary fibrosis. Serine protease DPP9 plays an important role in antigenic presentation and activation of inflammation. The *IFNAR2* gene (locus 21q22.1), which contains the rs2236757 variant, encodes an interferon receptor involved in type 1 IFN signaling. The rs74956615 variant is located near the *TYK2* (tyrosine kinase 2) gene on chromosome 19, the expression of which is associated with the more severe form of COVID-19. *TYK2* is one of the target genes for inhibitors of the JAK/STAT signaling pathway, such as baricitinib [8].

Some of the genetic associations with severe COVID-19 relate to the immune-mediated phase of the disease associated with respiratory failure, requiring invasive mechanical ventilation. Critical illness in COVID-19 is associated with at least two biological mechanisms: innate antiviral protection, which is especially important at the early stage of the disease (*IFNAR2* and *OAS* genes), and inflammatory lung disease, a key mechanism of the late phase of COVID-19 (*DPP9*, *TYK2,* and *CCR2* genes). Interferons are canonical host antiviral signaling mediators and stimulate the release of components of an early response to viral infection. Consistent with the protective role of type I interferons, increased expression of the IFNAR2 interferon receptor subunit reduced the odds of severe COVID-19. Loss-of-function mutations in the IFNAR2 gene are associated with severe COVID-19 and other viral infections. Administration of interferon can reduce the prob-

ability of critical illness in COVID-19, but at what point in the illness the treatment will be effective has not been determined. Exogenous interferon treatment did not lead to a decrease in the mortality of hospitalized patients in large-scale clinical trials [12], it is possible that this genetic effect acts at an early stage of the disease, when the viral load is high [8].

The extrapulmonary effects of COVID-19 can be mediated by an IFN-controlled increase in ACE2 expression on both endothelial and parenchymal cells, which leads to endotheliitis [13] and liver damage in 60 % of severely ill patients [14]. Although a deficiency in type I IFN immunity is associated with life-threatening pneumonia COVID-19 [5], induction of type I IFNs and IFN-stimulated genes are found in some critically ill patients. Profiling immune signatures in bronchoalveolar lavage fluid of eight patients with COVID-19 showed that the expression of 83 pro-inflammatory genes, especially those encoding cytokines (IL1RN and ILB) and chemokines (CXCL17, CXCL8, and CCL2), as well as receptor for the chemokines CXCR2 and CXCL2, markedly increased in COVID-19 cases compared to patients with community-acquired pneumonia and healthy controls, indicating hypercytokinemia in COVID-19 patients caused by the expression of multiple IFN-stimulated genes (ISGs). These ISGs exhibit immunopathogenic potential, with an overrepresentation of genes involved in inflammation. The transcriptome data was also used to estimate immune cell populations, revealing increases in activated dendritic cells and neutrophils. Activation of genes *IL1RN* and *SOCS3*, which encode antagonists of cytokine signaling, suggests that SARS-CoV-2 infection involves a negative feedback loop. The expression of genes involved in the morphogenesis and migration of immune cells (*NCKAP1L*, *DOCK2*, *SPN*, and *DOCK10*) was found to be lower than in healthy controls. The functional analysis revealed a state of high sensitivity to noxious stimuli in COVID-19 cases, characterized by powerful defensive reactions and hyperactive ribosome biogenesis. The study of the dynamics of cytokine expression showed that the expression levels of cytokine-related genes seem to decrease over time. The patient who eventually deceased was the outlier. These observations showed that severe inflammation in COVID-19 gradually resolves, and unresolved inflammation can lead to fatal consequences [15].

Research results suggest that type I IFNs play a bivalent role in the pathobiology of COVID-19, which requires tight regulation, and lead to the hypothesis that JAK/STAT inhibitors are useful early in the disease by reducing IFN-I-induced ACE2 expression. Noteworthy are important qualitative differences between the response of liver spheroids, where IFNs induced ACE2 and increased infectivity, and lung organelles, where IFN signaling did not affect ACE2 expression and viral load. Vascular endothelial cells express high levels of ACE2 [16] and are very sensitive to IFN signaling [17]. Taken together, these data suggest that the effects of baricitinib may differ in different organ systems and that anti-inflammatory effects may be most beneficial in those tissues in which ACE2 gene expression is a response to IFN, including the liver [18].

Genotyping of 322,948 biological samples from the UKB biobank for the *APOE* gene (apolipoprotein E) found that homozygotes *APOE* e4e4 (n = 9,022; 3 %) are more likely to test positive for COVID-19 (OR = 2.31; $p = 1.19 \times 10^{-6}$) compared to e3e3 homozygotes (the most common genotype, n = 223,457; 69 %). This association persisted after excluding patients with diseases associated with the severity of COVID-19 (hypertension, coronary heart disease, myocardial infarction, angina pectoris, diabetes, dementia) from the analy-

sis. Therefore, it can be confidently asserted that the e4 allele of the *APOE* gene increases the risk of severe COVID-19 infection regardless of other risk factors. *APOE* is one of the highly expressed genes in type II alveolar lung cells. The *APOE* e4 variant affects not only the function of lipoproteins and the development of cardiometabolic diseases but also the pro-/antiinflammatory phenotypes of macrophages. Further research is needed to understand the biological mechanisms linking the ApoE genotypes to the severity of COVID-19 [19].

The fact that men are more at risk of severe COVID-19 is due in part to the localization of the *ACE2* gene on the X chromosome [20]. In the region spanning the entire *ACE2* gene and 50,000 base pairs around it, SNPs have been found that carry alleles inherited from Neanderthals. Neanderthal haplotypes in the *DPP4* gene (*DPP9* homolog) are associated with an ~80 % increased risk of hospitalization after infection with SARS-CoV-2. The S-protein of SARS-CoV-2 binds to the membrane-bound DPP4 receptor (known as CD26) [21]. DPP4 is involved in several physiological processes, including the regulation of glucose metabolism. DPP4 inhibitors are used to treat diabetes and are thought to influence COVID-19 outcomes [22]. However, the Neanderthal variant of the *DPP4* gene doubles the risk of severe COVID-19 disease [23]. SNP rs117888248 has the strongest association with severe COVID-19. Neanderthal haplotypes in the *DPP4* gene and on chromosome 3 increase the risk of severe COVID-19 with respiratory failure and the need for mechanical ventilation by 100 % each. Both risk haplotypes in the *ACE2* and *DPP4* genes have stronger effects than the protective Neanderthal haplotype on chromosome 12, which reduces the risk of severe disease by ~23 % [11].

The Neanderthal variant of the *DPP4* gene is present in 1 % of Europeans, 2.5 % of South Asians, 4 % of East Asians, and 0,7 % of Americans. Three available today genomes of Neanderthals from Europe and southern Siberia, whose age varies from 50 to 120 thousand years, are homozygous for risk variants. This means that if a Neanderthal were alive today, he would have a 4–16 times higher risk of severe illness if infected with the SARS-CoV-2 [23].

Advances in proteomics, combined with human genetic data, are helping to identify therapeutic targets and drug development against COVID-19. The finding of a causal relationship between circulating proteins and susceptibility to SARS-CoV-2 infection or the severity of COVID-19 is a promising direction of the development of pharmacotherapy for this disease, in which exposure to SARS-CoV-2 causes profound changes in the levels of circulating proteins. Several genetic associations lead to potential therapeutic approaches for enhancing interferon signaling, counteracting the activation and infiltration of leukocytes into the lungs, or specifically targeting inflammatory pathways [8].

A network model of SARS-CoV-2 antiviral effectors

The innate immune response to viral infection is a critical factor in determining the outcome of the disease. To identify the cellular antiviral response to SARS-CoV-2 infection, L. Martin-Sancho and colleagues screened 399 human interferon-stimulated genes (ISG) [24]. For example, loss of function mutations in the TLR7 immune sensor and suppression of the type I interferon (IFN) response is associated with severe COVID-19. These data highlight the role of IFN-mediated cellular responses in controlling SARS-CoV-2 infection and the severity of COVID-19.

The binding of type I IFN to the IFNAR receptor promotes the activation of transcription of hundreds of ISGs, many of which have antiviral activity. Coordinated regulation and expression of receptors, effector signaling molecules, and transcription factors are essential for a successful antiviral response. Viruses have developed various strategies to interfere with and evade these antivirus programs. RNA sequencing of samples from COVID-19 patients revealed ISG activation. In addition, SARS-CoV-2 is sensitive to IFN treatment [25] and to the activity of several ISGs, including the *LY6E* gene, which inhibits SARS-CoV-2 replication, as well as the *AXIN2*, *CH25H*, *EPSTI1*, *GBP5*, *IFIH1*, *IFITM2,* and *IFITM3* genes.

The limitation of SARS-CoV-2 is mediated by the ISG group, which are regulators of protein degradation in the endoplasmic reticulum (ER), the lipid composition of cell membranes, and vesicular transport. Among these ISGs is the *BST2* gene (bone marrow stromal antigen 2, known as CD317), which inhibits the replication of SARS-CoV-2. The BST2 protein is localized on the plasma membrane and in endosomes, inhibits the release of viruses in the envelope, which bud either at the plasma membrane or the membranes of the ER and the Golgi network. Cells deficient in BST2 released significantly more viral particles. Expression of the viral protein Orf7a partially eliminates BST2-mediated inhibition of SARS-CoV-2 release, although it does not reduce BST2 expression on the cell surface. BST2 and Orf7a are colocalized in the perinuclear region and, according to in silico molecular docking, physically interact. Thus, the viral protein Orf7a makes it possible to evade immunity through antagonism with BST2 [24].

L. Martin-Sancho and colleagues [24] identified 139 ISGs activated in the tracheobronchial epithelium, 121 ISGs activated in alveolar epithelial cells, and 152 ISGs activated in both types of cells. This dataset encompasses ISGs with previously characterized broad-spectrum antiviral activity, including *MX1*, *OAS1*, *OASL,* and *IFI6*. ISGs are a group of genes with functions of inflammatory signaling, intracellular and nuclear transport, and energy metabolism. The list of 65 ISGs that inhibit SARS-CoV-2 replication includes the *MYD88* signal adapter, *STAT1,* and *STAT2* signal transducers, *DDX60* helicase, and *TRIM21* ubiquitin ligase, and signaling pathway effectors including *BST2*, *IFITM2*, and *IFITM3*, which are likely to have direct antiviral activity. Among the ISGs that cause the greatest reduction in SARS-CoV-2 replication is the gene *ELF1* encoding a transcription factor, which drives an antiviral program of more than 300 genes.

To better understand the biochemical and functional context in which these 65 ISGs exhibit antiviral activity, a study was carried out in which closely related clusters of proteins in cellular biological processes were identified. A strong connection was found with IFN signaling pathways, including regulators of phosphorylation (activation) STAT (signal transducer and activator of transcription, a key molecule of IFN-dependent and many other signaling pathways), cell death, transport in the Golgi network or ER, metabolism of nucleotides and sphingolipids. Additional resident ER/Golgi factors identified as potent replication inhibitors of SARS-CoV-2 are apolipoprotein APOL2 (lipid metabolism) and RSAD2/Viperin (lipid synthesis, control of ER membrane curvature). This means that regulation of membrane composition at viral replication or traffic sites may be an important strategy for controlling SARS-CoV-2 replication. This network model illuminates the diversity of cellular activities that function to generate an antiviral response to SARS-CoV-2 replication.

Six ISGs (*LY6E*, *CLEC4D*, *UBD*, *ELF1*, *FAM46C,* and *REC8*) suppress SARS-CoV-2 penetration by 50 %. *ELF1* and *CLEC4D* act as indirect negative regulators of viral infections through a secondary antiviral transcriptional cascade. *LY6E*, *UBD,* and *FAM46C* directly inhibit viral penetration. For example, *LY6E* restricts the entry of live SARS-CoV-2 by inhibiting viral S-protein fusion with the membrane.

ISGs *IFIT3*, *SPATS2L*, *DNAJC6*, *RGSS2*, *LOC152225*, *ZBP1,* and *B4GALT5* suppress SARS-CoV-2 RNA replication by more than 50 % after SARS-CoV-2 enters the cytosol. ISGs that inhibit SARS-CoV-2 replication include genes that encode proteins that bind viral nucleic acid (*ZBP1* and *IFIT1* genes), and ER–Golgi resident proteins (ERG) NAPA, APOL2, and ERLIN1. Five members of the IFIT family (genes *IFIT1*, *IFIT1B*, *IFIT2*, *IFIT3,* and *IFIT5*) prevent viral RNA replication by sequestering single-stranded unmethylated RNA. The *SPAT2SL* gene encodes a protein that reduces viral RNA levels through direct binding. Following stress stimuli, SPAT2SL is recruited into cytoplasmic stress granules, where viral RNA can be isolated. DNAJC6, a member of the heat shock protein 40 (HSP40) family of heat shock proteins, also affects the replicative stage of SARS-CoV-2.

ERG resident proteins inhibit the late stage of SARS-CoV-2 replication. The assembly of the virus after transcription and translation of the subgenomic mRNAs of SARS-CoV-2 and the generation of viral proteins occurs in the ERG. The structural proteins of the virus bind to the viral RNA to form virions, which bud from the ERG and are released through exocytosis. Several ISGs restrict late stages of viral replication, including regulators of ER-associated degradation of the ERLIN1, RETREG1, and FNDC4 proteins, and act as inhibitors of SARS-CoV-2 infection.

SARS-CoV-2 replication negative control proteins, including heat shock protein HSPA8 and phosphodiesterase CNP, are found in a protein complex with NAPA, another identified SARS-CoV-2 limiting factor and a member of the complex that functions to dock and fusion of vesicles to target membranes. The GTPase RAB27A controls exocytic transport through the fusion of multivesicular endosomes with the plasma membrane and also interferes with late stages of replication. These data highlight the role of factors involved in vesicular trafficking as negative regulators of SARS-CoV-2 replication.

Taken together, analyzes of ISGs that interfere with SARS-CoV-2 infection have shown that the IFN response to SARS-CoV-2 infection depends on a limited set of ISGs that govern a variety of cellular functions, including endocytosis, nucleotide biosynthesis, ER-associated protein degradation, and lipid metabolism. Further study of host-pathogen interactions will allow us to understand the role of molecular genetic determinants of innate immune control of SARS-CoV-2 replication in the clinical outcome of the disease [24]. These factors represent potential targets for therapeutic intervention.

Coronavirus SARS-CoV-2 genetics

Genomes sequenced of 5,085 SARS-CoV-2 strains (1,026 strains belonging to the earliest confirmed cases of COVID-19, and 4,059 strains recovered during the massive second wave of the pandemic) in a large metropolis (Houston, USA), an ethnically diverse region with 7 million residents. Analysis of the SARS-CoV-2 strains that caused the disease in the first wave (05.03–11.05.2020) revealed a wide variety of viral genomes, which together represent the main monophylogenetic groups identified in the world today, although not all “branches” of the evolutionary tree SARS-CoV-2 is represented

in this data. The phylogenetic distribution of strains with multiple substitutions of different amino acids at the same site showed their independent origin. Almost all strains (4,054) of the second wave have a substitution of the amino acid asparagine-614 in the receptor-binding domain (RBD) of the S-protein for glycine, which is associated with increased transmissibility and infectivity. Strains with the Gly614 variant in the S-protein accounted for 71 % of SARS-CoV-2 strains at the beginning of the 1st wave and 99.9 % in the 2nd wave ($p < 0.0001$). Patients infected with Gly614 strains had significantly higher upper respiratory viral load at initial diagnosis than patients infected with Asp614 strains. At the same time, the association of the severity of the disease with concomitant diseases and human genetics remains. The presence of the Gly614 variant did not correlate with the disease outcome. Some regions of the S protein — the primary target of global vaccine efforts — are rife with amino acid substitutions, perhaps indicating the action of selection. In RBD S-protein amino acid substitutions are relatively rare compared to other regions of the protein, but some of them reduce the recognition by the neutralizing monoclonal antibody CR3022. This is consistent with the functional role of RBD in interacting with ACE2 and the assumption that new variants of the virus arise due to pressure from the host's immune system [26]. Virus strains with the 614Gly variant show significantly increased replication in human lung epithelial cells in vitro and increased titers in nasal and tracheal washes obtained from experimentally infected hamsters. Thus, the 614Gly variant bestows increased virus fitness in the upper respiratory tract [27].

The SARS-CoV-2 genome encodes an RNA-dependent RNA polymerase (RdRp; also called Nsp12), which is involved in viral replication. Each of the two amino acid substitutions (Phe479Leu and Val556Leu) in the gene encoding RdRp confers to the infection significant resistance to remdesivir, an analog of adenosine [28; 29].

Untranslated flanking regions (5'- and 3'-untranslated region, 5'- and 3'-UTR) of the SARS-CoV-2 genome encode exclusively conserved secondary RNA structures with gene regulatory functions in viral replication and transcription. UTRs can interact with several human and viral protein factors and provide RNA — RNA or RNA — protein interactions through the circularization of the genome. To investigate the genomic stability of SARS-CoV-2, nucleotide variants in isolates collected from the ongoing pandemic were analyzed. 87 SNPs with frequencies > 0.5 % (occurring in at least 93 genomes) were identified. In an extended analysis of 18,599 SARS-CoV-2 genomes, the 241C > T variant was detected with a frequency of 70.2 %. In addition, 6 variants were identified in the 3'-UTR (29700A > G, 29711G > T, 29734G > C, 29742G > T, 29742G > A, 29870C > A) and three in the 5'-UTR (36C > T, 187A > G, 241C > T), which were detected with a frequency of 0.62–1.05 % [30]. A. Mishra et al. identified two positions corresponding to two substitutions found in this analysis — 241C > T in 5'-UTR and 29742G > A/T in 3'-UTR [31]. If an SNP occurs randomly, the probability that it leads to missense, synonymous, and nonsense mutations is 73, 22, and 5 %, respectively, in all 26 viral genes encoding proteins. Analysis of observed amino acid substitutions in 769 SNPs with a variant frequency of 0.05 % or higher revealed fewer than expected missense and nonsense mutations across all genes, except ORF8. The deviations of observed proportions from expected values varied widely across genes. In ORF8, for example, the frequencies of missense, synonymous, and nonsense mutations were 77, 15, and 8 %, respectively, which are close to expected. In contrast, for the processed peptide nsp9 (non-structural

protein 9), the putative function of which is to dimerize and bind RNA, the corresponding proportions were 18, 82, and 0 %, respectively. Selection and evolutionary pressures are likely to differ in individual SARS-CoV-2 genes. Thus, the characterization of SARS-CoV-2 variants suggests a non-random selection pressure that points to hidden driving forces of the evolution of the viral genome associated with a functional or regulatory role of genes [30].

Linkage disequilibrium analysis of SNPs in 18,599 genomes identified a total of 34 groups of coevolving variants (CEV) with a frequency of ≥ 0.1 %. The two groups of CEVs included UTR and other gene features that may indicate functional dependencies or interactions of genomic elements carrying SNP. The first group of CEV (CEVg1) was found in 69.5 % of SARS-CoV-2 genomes and consisted of four variants located in the 5'-UTR (g.241C > T), nsp3 (g.3037C > T, synonymous), RNA-dependent RNA polymerase (g.14408C > T, p.P323L) and S-protein (g.23403A > G, p.D614G). The incidence of CEVg1 increased sharply from 12.2 % to 93.4 % over the three months from February to May 2020 both globally and for each region by continent. The D614G mutation in CEVg1 enhances the transmission of the virus. Another group of CEVs (CEVg5) associated with the 3'-UTR was found in 0.9 % of genomes and included six variants in the leader protein nsp1 (g.490T > A, p.D75E), nsp3 (g.3177C > T, p.P153L), exonuclease (g.18736T > C, p.F233L), S-protein (g.24034C > T, synonymous), membrane protein (g.26729T > C, synonymous) and the 3'-UTR (g.29700A > G). The CEVg5 group remained a minor group in March and April 2020, amounting to 1.2 and 0.53 %, respectively [30]. The nsp3 protein of coronaviruses can block the innate immune response of the host, and other non-structural proteins play a role in evading recognition by the immune system [32]. Overall, a review of variants in 18599 SARS-CoV-2 genomes collected in May 2020 indicates that coevolving and single variants with a likely functional effect on the replicative capacity or pathogenicity of the virus have been identified in both the UTR and functional elements throughout the genome [30].

More than 86,450 SARS-CoV-2 genomes became available in October 2020. This allowed analysis of coevolving variants of 86,450 genomes, which is more than 4 times the size of the first dataset of 18,599 genomes. Comparison of the frequencies of the CEV groups between May and October 2020 datasets provided new insight into the evolution of SARS-CoV-2. First, it confirmed the global dominance of CEVg1 with the D614G mutation in the S-protein, which increased from 69.5 % to 85 % between May and October 2020. Second, the CEVg3 and CEVg4 groups gradually disappeared. Third, two new groups of emerging coevolving mutations have been identified (CEVg6 and CEVg8). CEVg6 appeared in Oceania, its frequency increased from 0 % in April to 96 % in July 2020, while CEVg8 appeared in Europe and its frequency increased from 0 % in June to 36 % in September 2020. The CEVg6 and CEVg8 groups carry new mutations in the S-protein, S477N, and A222V, respectively [30].

Human microRNAs (miRNAs) are evolutionarily conserved noncoding RNAs that can posttranscriptionally suppress gene expression through hybridization of partially homologous sequences, primarily with the 3'-UTR of mRNA. Human miRNAs can target viral RNAs and positively or negatively modulate various stages of viral replication and the viral life cycle [33]. To gain insight into the possible interaction of SARS-CoV UTRs with host microRNAs in modulating the pathogenesis of infection, a search was performed for sequence homology of human miRNAs with

SARS-CoV-2 UTR sequences. MicroRNAs were identified from the miRBase database, including sense and antisense sequences matching the 3'- and 5'-UTRs, respectively. Three miRNAs (hsa-miR-1307-3p, hsa-miR-1304-3p, and hsa-miR-15b-5p) are expressed in all 23 tissues, including lung, heart, liver, kidney, and small intestine, which are severely affected during SARS-CoV-2 infections. Sequences homologous to human hsa-miR-1307-3p and hsa-miR-1304-3p are located in S2m, a conserved genetic element of the virus with unknown function. Based on in silico computer modeling of the interaction between the viral 3'-UTR and human hsa-miR-1307-3p, a possible mechanism of viral survival is presented, according to which a mutation in the 3'-UTR of SARS-CoV-2 weakens the host immune response. M. A. K. Khan et al. [34] identified a target of miR-1307-3p in the 3'-UTR, which mediates antiviral responses and inhibits viral replication. A study by L. Bavagnoli et al. demonstrated the functional role of human miR-1307 in the regulation of influenza A H1N1 virus replication [35] and predicted the complementarity of miR-1307 to the NS1 protein of the H1N1 virus, which limits interferon and pro-inflammatory responses, allowing the virus to evade the innate and adaptive host immunity and replicate efficiently in infected cells. Previously, hsa-miR-1307-3p was associated with lung function [36], as well as with the progression of certain types of cancer in patients with COVID-19 [37]. The C112A mutation, which allows the virus to escape miR-1307, is associated with ARDS. It is noteworthy that in the SARS-CoV-2 genome, the site of interruption of hybridization with miR-1307-3p coincides with the localization of the C112A mutation in the H1N1 genome. Apparently, SARS-CoV-2 shares a mechanism of protection against host immunity with H1N1, if SARS-CoV-2 carries an allele that weakens the function of miR-1307. In support of this hypothesis, analysis of SARS-CoV-2 variations revealed two nearby mutations at positions 29,742 and 29,734, which correspond to positions 7 and 15 of miR-1307, respectively. Mutations at these two sites can disrupt the hybridization of SARS-CoV-2 RNA with miR-1307 to escape from binding and inhibiting infection. As of October 2020, these mutations were detected with a frequency of < 1.2 %. Their relationship with the severity of clinical symptoms is currently unknown and requires further study [30].

Thus, an integrated approach to the analysis of genome variations in circulating strains of SARS-CoV-2 during the current pandemic identified possible interactions of human miR-1307-3p microRNA with the 3'-UTR of the SARS-CoV-2 genome [30], which is confirmed by other researchers [38]. N. Balmeh et al. identified hsa-miR-1307-3p as the best miRNA out of 1872 miRNAs with the highest affinity for the SARS-CoV-2 genome and its related cellular signaling pathways. The results of their study showed that this miRNA is involved in the prevention of the production of the GRP78 protein (glucose regulating protein 78), the cellular co-receptor of the SARS-CoV-2 virus, preventing the virus penetration, endocytosis, proliferation, and development. This creates potential targets for antiviral interventions [38].

Currently, several variants of the Spike protein of the SARS-CoV-2 virus are known, which appeared as a result of mutations. It is unclear whether these variants may have a specific effect on the affinity for the ACE2 receptor, which in turn is characterized by multiple alleles in the human population. Among the 295,000 sequenced SARS-CoV-2 genomes isolated from different patients, several mutations in the Spike protein have been identified that affect interactions with ACE2: S477N, N439K, N501Y, Y453F, E484K, K417N, S477I, and G476S. In particular, the N501Y mutation is one of the events charac-

terizing the SARS-CoV-2 B. 1.1.7 strain with increased infectivity, the frequency of which has recently increased in Europe [39].

A case of chronic infection with SARS-CoV-2 with reduced sensitivity to neutralizing antibodies in an immunosuppressed individual who received convalescent plasma treatment, which generates changes in the viral genome sequence, is described. The analysis covered 23-time points over 101 days. Small changes were observed in the general structure of the viral population after two courses of remdesivir during the first 57 days. However, after plasma treatment, large dynamic changes in the virus population were found with the emergence of a dominant viral strain carrying mutations D796H in the S2 subunit and ΔH69/ΔV70 in the S1 subunit of the Spike protein. The D796H mutation was found to be the main factor in reducing the sensitivity of the virus to plasma antibodies but caused a defect in infectivity. The second mutation, the ΔH69 / ΔV70 deletion, doubled the infectivity compared with the wild type and compensated for the decrease in infectivity resulting from the first D796H mutation. The double mutant carrying the ΔH69 / ΔV70 deletion and the D796H substitution had a moderately reduced sensitivity to antibodies of convalescent plasma in vitro while retaining wild-type infectivity. These data testify to a strong selection of SARS-CoV-2 during convalescent plasma therapy, associated with the emergence of viral variants with reduced sensitivity to neutralizing antibodies [40].

The SARS-CoV-2 strain with a 382-nucleotide deletion (Δ382) in the ORF8 gene appeared in Wuhan at the beginning of the pandemic. The Δ382 deletion truncates the open reading frame and interrupts transcription. The Δ382 variant causes clinically significant disease, including pneumonia, but with a milder course compared to infections caused by wild-type virus. None of the 29 patients (0 %) infected with this variant had hypoxia requiring supplemental oxygen (an indicator of severe COVID-19, primary study endpoint), unlike patients infected with wild-type SARS-CoV-2 virus (28 %). The clinical effect of deletions in the ORF8 region is manifested by the less systemic release of pro-inflammatory cytokines, less systemic inflammation, and a more effective immune response to SARS-CoV-2. Stronger production of IFN-γ at an early stage of infection, which was observed in patients infected with the Δ382 variant, supports T-cell effector functions and a rapid and efficient humoral response to SARS-CoV-2 [41].

The high transmissibility of the SARS-CoV-2 coronavirus by airborne droplets and contact routes has led to the COVID-19 pandemic, which continues to spread throughout the world, despite strict control measures. Moreover, following the relaxation of social distancing policies, there is a resurgence of COVID-19 in many regions. One of the key questions of COVID-19 is whether a real re-infection is occurring. Although neutralizing antibodies develop rapidly after infection, antibody titers begin to decline as early as 1–2 months after an acute infection. Patients tested negative for SARS-CoV-2 RNA and discharged from hospitals sometimes have positive retesting results. These reported cases have caused disagreement among experts about the hypothesis of persistent virus shedding and reinfection.

Studying the viral genome, in particular, sequencing its sequence, is useful not only for tracking its variability and distribution but also for clarifying the question of the possibility of re-infection. The first case report of reinfection was published in August 2020 in Hong Kong. A 33-year-old man recovered from COVID-19 in April and was discharged from the hospital after two negative PCR tests for the presence of SARS-CoV-2 in swabs taken

from the nasopharynx and throat at 24-hour interval tested positive after 4 months test for RNA SARS-CoV-2 in saliva. During the second (asymptomatic) episode of COVID-19, the patient remained in good physical shape, and the blood test results were normal or nearly normal. Serial chest radiographs showed no abnormalities. The patient has not received antiviral treatment. The viral genomes from the first and second episodes belong to different strains of SARS-CoV-2. The first viral genome has a stop codon in the ORF8 gene, leading to a truncation of 58 amino acids, and is phylogenetically closely related to strains collected in March — April 2020, while the genome of the second virus is closely related to strains collected in July — August 2020. Another 23 nucleotide and 13 amino acid differences, located in 9 different proteins, were found between the viruses from the first and second episodes. Epidemiologic, clinical, serologic, and genomic analyzes confirmed that the patient had re-infection instead of persistent viral shedding from the first infection. These data indicate that SARS-CoV-2 may continue to circulate among the human populations despite herd immunity resulting from natural infection or vaccination [42]. Later, the possibility of reinfection was confirmed by other reports.

A 25-year-old man living in the United States was infected with SARS-CoV-2 twice — in April and June 2020. The second infection was symptomatically more severe than the first. The genetic mismatch of SARS-CoV-2 samples in the two episodes of infection was greater than can be explained by the short-term in vivo evolution in the patient's body. These data indicate that the patient was infected with SARS-CoV-2 on two different occasions with genetically different strains of the virus. Thus, the previous exposure to SARS-CoV-2 does not guarantee the emergence of immunity against its new strains [43].

A report from Brazil described a series of 33 reinfection cases, of which 30 were health workers. Sequencing of the viral genome revealed reinfection with a phylogenetically different isolate in each of these patients. The detection of phylogenetically distinct genomic sequences in the first and second episodes argues for de-novo reinfection. Reinfection was due to a decreased humoral response during the first episode of the disease and proves the need for constant vigilance without the assumption of the development of immunity in convalescents [44].

All authors of reports of cases of reinfection insist that patients who have recovered from COVID-19 must comply with epidemiological control measures.

Coronaviruses acquire genetic changes more slowly than other RNA viruses due to correcting RNA-dependent RNA polymerase (RdRp). Repeated deletions in the S-protein gene that alter the amino acid sequence can stimulate and, apparently, accelerate the adaptive evolution of SARS-CoV-2. Deletion variants arise in different genetic and geographic backgrounds, transmit efficiently, and are present in new strains, including those that cause the current global problem. Regions of the genome with recurrent deletion regions (RDRs) map to defined antibody epitopes. Deletions in the RDR confer resistance to neutralizing antibodies. For example, repeated deletions that alter amino acids at positions 144–145 and 243–244 of the S-protein disrupt the binding of the 4A8 antibody, which recognizes an immunodominant epitope in the N terminal domain (NTD) of the S-protein. The antigenic novelty of the virus allows re-infection of previously immunized individuals. During long-term infections in immunocompromised patients, the virus acquires deletions in the S-protein NTD. This phenomenon is called the evolutionary pattern, defined by deletions that change defined antibody

epitopes. Deletions and substitutions in the major epitopes NTD and RBD are likely to continue to contribute to this process. Unlike nucleotide substitutions, deletions cannot be corrected by RdRp polymerase proofreading, and this accelerates the adaptive evolution of SARS-CoV-2. Deletions represent a mechanism through which S glycoprotein rapidly acquires genetic and antigenic novelty of SARS-CoV-2. Since deletions are a product of replication, they will occur at a certain rate, and these variants will likely appear in healthy populations [45].

Some new variants of SARS-CoV-2, emerging in the winter of 2020–2021, such as B.1.351 and B.1.1.17, show higher infectivity and virulence than earlier variants. So far, the analysis of new SARS-CoV-2 variants has focused on point nucleotide substitutions and short deletions that are readily identifiable by comparison to reference genomes. Insertions have largely escaped the attention of researchers, although the furin site insert (PRRA tetrapeptide) in the S-protein is a determinant of SARS-CoV-2 virulence. This insertion is unique to SARS-CoV-2 and introduces a furin cleavage site into the S-protein, enhancing its binding to the receptor. S. K. Garushyants et al. [46] identified 141 unique insertions of different lengths in the SARS-CoV-2 genomes, which reflect the actual virus variance rather than sequencing errors. Two principal mechanisms explain the appearance of insertions in the genomes of SARS-CoV-2. Short insertions in SARS-CoV-2 result from RNA template sliding (polymerase stuttering), while long insertions are due to the formation of non-canonical subgenomic mRNA (sgRNA) due to so-called template switching or duplication of adjacent sequences. Hotspots of template switching are characterized by polymerase "jumping" from one location on the genome to another. In infected cells, beta-coronaviruses produce 5-8 sgRNAs.

Insertions in the S-glycoprotein gene can affect its antigenic properties. At least two insertions in the immediate vicinity of the antibody binding site in the NTD S-protein domain (ins246DSWG and ins15ATLRI) are predicted to elude neutralizing antibodies, while other insertions located in the NTD loops are predicted to elude antibodies and/or T-cell immunity.

Insertions are non-uniformly distributed along the SARS-CoV-2 genome. In particular, 7 out of 25 long insertions are located in the S-protein, suggesting that these insertions persist due to their adaptive value to the virus. The increased frequency of insertions in the S-protein gene indicates that their distribution is driven by positive selection to enhance the interaction of SARS-CoV-2 with human cells, which is given by insertions. To evaluate the functional effects of the insertions in the S-protein NTD, they were plotted on the tertiary structure of the protein. All these insertions were located on the surface of the protein, and ins15ATLRI and ins246DSWG are localized in an epitope recognized by antibodies from convalescent plasma, that is, these insertions may be associated with the escape of SARS-CoV-2 from antibodies. The revealing of multiple insertions at the same site suggests an important role for the NTD domain in SARS-CoV-2 infection, especially since COVID-19 spread-related substitutions and deletions have been reported on this site. Taking into account that NTD is important for the virus to interact with CD4+T lymphocytes, this may be due to evasion of T cell immunity.

Structural variation is an important driving force behind the evolution of betacoronaviruses, and insertions in the S and N genes appear to contribute to the pathogenicity of betacoronavirus. The fact that beta-coronaviruses produce transcripts longer than their genomes suggests that insertions are a natural part of the life cycle of these viruses.

The excess of insertions in the S-protein is compatible with the function of this protein, which is the main area of adaptation of the virus. However, the location of most of the insertions in NTD, as opposed to the RBD, is unexpected. Considering that all detected insertions appeared at a relatively late stage of the pandemic, it seems likely that at the beginning of the pandemic, the RBD structure was already largely optimized for binding to cellular receptors, and insertions in NTD allow the virus to escape immunity without disrupting the interaction with receptors [46].

HLA class, I antigens play a crucial role in the development of a specific immune response to viral infections. Shkurnikov et al. [47] developed a risk scale associated with the ability of HLA class I molecules to present peptides of the SARS-CoV-2 coronavirus. The scores on this scale are significantly higher in the group of adult patients (age ≤ 60) who died from COVID-19 compared to elderly patients (age > 60). In particular, the presence of the HLA-A*01:01 allele is associated with a high risk of death, while HLA-A*02:01 and HLA-A*03:01 are associated with low risk. Analysis of homozygous patients showed that homozygosity for the HLA-A*01:01 allele is associated with the early death of patients with COVID-19. The SR score in an independent cohort of Spanish patients was also associated with disease severity. The results obtained indicate the important role of the presentation of viral peptides by HLA class I molecules in the development of a specific immune response to COVID-19. This conclusion is consistent with the data of Italian researchers that the occurrence of the HLA-A*01:01 and HLA-A*02:01 alleles are associated with the mortality rate of patients with COVID-19 in different regions of Italy [48]. To identify possible associations with clinical information, it is necessary to analyze the whole HLA class I genotype in patients with COVID-19.

Conclusion

In this review, we aimed to illuminate the available information on the genetic determinants of susceptibility to SARS-CoV-2 infection and the severity of COVID-19. The development of new drugs to treat this disease requires knowledge of its molecular pathways and critical target molecules. Blocking viral pathways, including receptors and enzymes, and controlling immune responses are promising strategies for reducing multiple organ dysfunction in patients with COVID-19.

References

1. WHO Coronavirus Disease (COVID-19) Dashboard. Available at: https://covid19.who.int/ (accessed: 10.08.2021).
2. Ahmadian E., Khatibi S. M. H., Soofiyani S. R. COVID-19 and kidney injury: pathophysiology and molecular mechanisms. *Rev. Med. Virol.,* 2021, vol. 31, no. 3, pp. e2176.
3. Sakurai A., Sasaki T., Kato S. Natural history of asymptomatic SARS-CoV-2 infection. *N. Engl. J. Med*., 2020, vol. 383, no. 3, pp. 885–886.
4. Clohisey S., Baillie J. K. Host susceptibility to severe influenza A virus infection. *Critical Care,* 2019, vol. 23, no. 1, p. 303.
5. Zhang Q., Bastard P., Liu Z. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science,* 2020, vol. 370, no. 6515, pp. eabd4570.
6. Kaser A. Genetic risk of severe Covid-19. *N. Engl. J. Med.,* 2020, vol. 383, no. 16, pp. 1590–1591.
7. Ellinghaus D., Degenhardt F., Bujanda L., Buti M., Albillos A., Invernizzi P., Fernández J., Prati D., Baselli G., Asselta R., Grimsrud M. M., Milani C., Aziz F., Kässens J., May S., Wendorff M., Wienbrandt L.,

Uellendahl-Werth F., Zheng T., Yi X., de Pablo R., Chercoles A. G., Palom A., Garcia-Fernandez A. E., Rodriguez-Frias F., Zanella A., Bandera A., Protti A., Aghemo A., Severe COVID-19 GWAS Group. Genomewide association study of severe COVID-19 with respiratory failure. *N. Engl. J. Med.,* 2020, vol. 383, no. 16, pp. 1522–1534.
8. Pairo-Castineira E., Clohisey S., Klaric L. Genetic mechanisms of critical illness in Covid-19. *Nature,* 2021, vol. 591, no. 7848, pp. 92–98.
9. Zhou S., Butler-Laporte G., Nakanishi T. A Neanderthal OAS1 isoform protects against COVID-19 susceptibility and severity: results from mendelian randomization and case-control studies. *Nat. Med.,* 2021, vol. 27, no. 4, pp. 659–667.
10. Zeberg H., Pääbo S. The major genetic risk factor for severe COVID-19 is inherited from Neanderthals. *Nature,* 2020, vol. 587, no. 7835, pp. 610–612.
11. Zeberg H., Pääbo S. A genomic region associated with protection against severe COVID-19 is inherited from Neandertals. *Proc. Natl. Acad. Sci USA,* 2021, vol. 118, no. 9, pp. e2026309118.
12. Pan H., Peto R., Henao-Restrepo A. M. Repurposed antiviral drugs for Covid-19 — Interim WHO Solidarity Trial Results. *N. Engl. J. Med.,* 2021, vol. 384, no. 6, pp. 497–511.
13. Varga Z., Flammer A. J., Steiger P. Endothelial cell infection and endotheliitis in COVID-19. *Lancet,* 2020, vol. 395, no. 10234, pp. 1417–1418.
14. Zhang C., Shi L., Wang F. S. Liver injury in COVID-19: management and challenges. *Lancet Gastroenterol. Hepatol.*, 2020, vol. 5, no. 5, pp. 428–430.
15. Zhou Z., Ren L., Zhang L. Heightened innate immune responses in the respiratory tract of COVID-19 patients. *Cell. Host. Microbe,* 2020, vol. 27, no. 6, pp. 883–890.
16. Hamming I., Timens W., M. Bulthuis L. C. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. The first step in understanding SARS pathogenesis. *J. Pathol.,* 2004, vol. 203, no. 2, pp. 631–637.
17. Jia H., Thelwell C., Dilger P. Endothelial cell functions impaired by interferon in vitro: Insights into the molecular mechanism of thrombotic microangiopathy associated with interferon therapy. *Thromb. Res*., 2018, no. 163, pp. 105–116.
18. Stebbing J., Sánchez Nievas G., Falcone M., Youhanna S., Richardson P., Ottaviani S., Shen J. X., Sommerauer C., Tiseo G., Ghiadoni L., Virdis A., Monzani F., Rizos L. R., Forfori., Céspedes A. A., De Marco S., Carrozzi L., Lena F., Sánchez-Jurado P. M., Lacerenza L. G., Cesira N., Bernardo D. C., Perrella A., Niccoli L., Méndez L. S, Matarrese D., Goletti D., Tan Y. J., Monteil V., Dranitsaris G., Cantini F., Farcomeni A., Dutta S., Burley S. K., Zhang H., Pistello M., Li W., Romero M. M., Pretel F. A., Simón-Talero R. S., García-Molina R., Kutter C., Felce J. H., Nizami Z. F. 9, Miklosi A. G., Penninger J. M., Menichetti F., Mirazimi A., Abizanda P., Lauschke V. M. JAK inhibition reduces SARS-CoV-2 liver infectivity and modulates inflammatory responses to reduce morbidity and mortality. *Sci. Adv.,* 2021, vol. 7, no. 1, p. eabe4724.
19. Kuo C. L., Pilling L. C., Atkins J. L. APOE e4 genotype predicts severe COVID-19 in the UK Biobank community cohort. *J. Gerontol. A. Biol. Sci. Med. Sci.,* 2020, vol. 75, no. 11, pp. 2231–2232.
20. Gemmati D., Bramanti B., Serino M. L. COVID-19, and individual genetic susceptibility/receptivity: role of ACE1/ACE2 genes, immunity, inflammation, and coagulation. Might the double X-chromosome in females be protective against SARS-CoV-2 compared to the single X-chromosome in males? *Int. J. Mol. Sci.,* 2020, vol. 21, no. 10, pp. 3474.
21. Li Y., Zhang Z., Yang L., Lian X., Xie Y., Li S., Xin S., Cao P., Lu J. The MERS-CoV Receptor DPP4 as a Candidate Binding Target of the SARS-CoV-2 Spike. *iScience*, 2020, vol. 23, no. 8, p. 101400.
22. Lim S., Bae J. H., Kwon H. S., Nauck M. A. COVID-19 and diabetes mellitus: from pathophysiology to clinical management. *Nat. Rev. Endocrinol*., 2021, vol. 17, no. 1, pp. 11–30.
23. Lee J. W., Lee I. H., Sato T., Kong S. W., Iimura T. Genetic variation analyses indicate conserved SARS-CoV-2-host interaction and varied genetic adaptation in immune response factors in modern human evolution. *Dev. Growth Differ.,* 2021, vol. 63, no. 3, pp. 219–227.
24. Martin-Sancho L., Lewinski M. K., Pache L., Stoneham C. A., Yin X., Pratt D., Churas C., Rosenthal S. B., Liu S., De Jesus P. D., O'Neill A. M., Gounder A. P., Nguyen C., Pu Y., Oom A. L., Miorin L., Rodriguez-Frandsen A., Urbanowski M., Shaw M. L., Chang M. W., Benner C., Frieman M. B., García-Sastre A., Ideker T., Hultquist J. F., Guatelli J., Chanda S. K. Functional landscape of SARS-CoV-2 cellular restriction. *bioRxiv,* 2020, Sep 30. https://doi.org/10.1101/2020.09.29.319566. Preprint
25. Lokugamage K. G., Hage A., de Vries M, Valero-Jimenez A. M., Schindewolf C., Dittmann M., Rajsbaum R., Menachery V. D. Type I interferon susceptibility distinguishes SARS-CoV-2 from SARS-CoV. *J. Virol.,* 2020, vol. 94, no. 23, pp. e01410–e01420.

26. Long S. W., Olsen R. J., Christensen P. A. Molecular architecture of early dissemination and massive second wave of the SARS-CoV-2 virus in a major metropolitan area. *mBio*., 2020, vol. 11, no. 6, pp. e02707-20.
27. Plante J. A., Liu Y., Liu J., Xia H., Johnson B. A., Lokugamage K. G., Zhang X., Muruato A. E., Zou J., Fontes-Garfias C. R., Mirchandani D., Scharton D., Bilello J. P., Ku Z., An Z., Kalveram B., Freiberg A. N., Menachery V. D., Xie X., Plante K. S., Weaver S. C., Shi P. Y. Spike mutation D614G alters SARS-CoV-2 fitness and neutralization susceptibility. *bioRxiv,* 2020, Sep 2. https://doi.org/10.1101/2020.09.01.278689
28. Shannon A., Le N. T., Selisko B., Eydoux C., Alvarez K., Guillemot J. C., Decroly E., Peersen O., Ferron F., Canard B. Remdesivir and SARS-CoV-2: structural requirements at both nsp12 RdRp and nsp14 exonuclease active-sites. *Antiviral Res.,* 2020, no. 178, pp. 104793.
29. Gordon C. J., Tchesnokov E. P., Woolner E. Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. *J. Biol. Chem*., 2020, vol. 295, no. 20, pp. 6785–6797.
30. Chan AP, Choi Y, Schork N. J. Conserved genomic terminals of SARS-CoV-2 as coevolving functional elements and potential therapeutic targets. *mSphere*, 2020, vol. 5, no. 6, pp. e00754-20.
31. Mishra A., Pandey A. K., Gupta P., Pradhan P., Dhamija1S., Gomes J., Kundu B., Vivekanandan P., Menon M. B. Mutation landscape of SARS-CoV-2 reveals five mutually exclusive clusters of leading and trailing single nucleotide substitutions. *bioRxiv,* 2020. https://doi.org/10.1101/2020.05.07.082768. Preprint
32. Meini S., Zanichelli A., Sbrojavacca R., Iuri F., Roberts A. T., Suffritti C., Tascini C. Understanding the pathophysiology of COVID-19: could the contact system Be the key? *Front. Immunol.,* 2020, no. 11, p. 2014.
33. Girardi E., López P., Pfeffer S. On the importance of host microRNAs during viral infection. *Front. Genet.,* 2018, no. 9, p. 439.
34. Khan M. A. K., Sany M. R. U., Islam M. S., Islam A. Epigenetic regulator miRNA pattern differences among SARS-CoV, SARS-CoV-2, and SARS-CoV-2 worldwide isolates delineated the mystery behind the epic pathogenicity and distinct clinical characteristics of pandemic COVID-19. *Front. Genet.,* 2020, no. 11, p. 765.
35. Bavagnoli L., Campanini G., Forte M., Ceccotti G., Percivalle E., Bione S., Lisa A., Baldanti F., Maga G. Identification of a novel antiviral micro-RNA targeting the NS1 protein of the H1N1 pandemic human influenza virus and a corresponding viral escape mutation. *Antiviral. Res.,* 2019, no. 171, p. 104593.
36. Herrera-Rivero M., Zhang R., Heilmann-Heimbach S. Circulating microRNAs are associated with pulmonary hypertension and development of chronic lung disease in congenital diaphragmatic hernia. *Sci. Rep.,* 2018, vol. 8, no. 1, pp. 10735.
37. Qiu X., Dou Y. miR-1307 promotes the proliferation of prostate cancer by targeting FOXO3A. *Biomed Pharmacother*., 2017, no. 88, pp. 430–435.
38. Balmeh N., Mahmoudi S., Mohammadi N., Karabedianhajiabadi A. Predicted therapeutic targets for COVID-19 disease by inhibiting SARS-CoV-2 and its related receptors. *Inform. Med. Unlocked*., 2020, no. 20, p. 100407.
39. Ortuso F., Mercatelli D., Guzzi P. H., Giorgi F. Structural genetics of circulating variants affecting the SARS-CoV-2 spike/human ACE2 complex. *J. Biomol. Struct. Dyn.,* 2021, pp. 1–11. https://doi.org/10.1080/07391102.2021.1886175. Preprint
40. Kemp S. A., Collier D. A., Datir R. P. et al. SARS-CoV-2 evolution during treatment of chronic infection. *Nature,* 2021, vol. 592, no. 7853, pp. 277–282.
41. Young B. E., Fong S. W., Chan Y. H. Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study. *Lancet*, 2020, vol. 396, no. 10251, pp. 603–611.
42. To K. K., Hung I. F., Ip J. D., Ip J. D., Chu A. W., Chan W. M., Tam A. R. COVID-19 re-infection by a phylogenetically distinct SARS-coronavirus-2 strain confirmed by whole-genome sequencing. *Clin. Infect. Dis.,* 2021, vol. 73, no. 9, pp. e2946–e2951.
43. Tillett R. L., Sevinsky J. R., Hartley P. D. Genomic evidence for reinfection with SARS-CoV-2: a case study. *Lancet Infect. Dis.,* 2021, vol. 21, no. 1, pp. 52–58.
44. Dos Santos L. A., de Góis F. P. G., Fantini S. A. M. Recurrent COVID-19 including evidence of reinfection and enhanced severity in thirty Brazilian healthcare workers. *J. Infect.,* 2021, vol. 82, no. 3, pp. 399–406.
45. McCarthy K. R., Rennick L. J., Nambulli S. Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science,* 2021, vol. 371, no. 6534, pp. 1139–1142.

46. Garushyants S. K., Rogozin I. B., Koonin E. V. Insertions in SARS-CoV-2 genome caused by template switch and duplications give rise to new variants of potential concern. *bioRxiv,* 2021, Apr 26. https://doi.org/10.1101/2021.04.23.441209. Preprint
47. Shkurnikov M., Nersisyan S., Jankevic T., Galatenko A., Gordeev I., Vechorko V. Association of HLA class I genotypes with severity of Coronavirus Disease-19. *Front. Immunol.,* 2021, Feb 23. https://doi.org/10.3389/fimmu.2021.641900. Preprint
48. Pisanti S., Deelen J., Gallina A. M. Correlation of the two most frequent HLA haplotypes in the Italian population to the differential regional incidence of Covid-19. *J. Transl. Med.*, 2020, vol. 18, no. 1, p. 352.

Received: August 10, 2021
Accepted: December 15, 2021

Authors' information:

Aleksandr S. Golota — PhD in Medicine, Associate Professor; golotaa@yahoo.com
Dmitry A. Vologzhanin — Dr. Sci. in Medicine; volog@bk.ru
Tatyana A. Kamilova — PhD in Biology; kamilovaspb@mail.ru
Stanislav V. Makarenko — Assistant Lecturer; st.makarenko@gmail.com
Sergey G. Scherbak — Dr. Sci. in Medicine, Professor