

Genetics and the relationship between the virus and the host organism infected with SARS-COV-2

Schröder Verginica¹, Mitea Gabriela¹, Iancu Irina Mihaela^{1*}, Iancu Valeriu¹, Honcea Adina¹, Stoicescu Ramona Mihaela¹

¹Faculty of Pharmacy, “Ovidius” University of Constanta, 900470, Romania

*Corresponding author: irina.iancu@365.univ-ovidius.ro

Abstract. Human genetic diversity is shaped by both demographic and biological factors and has fundamental implications for understanding the genetic basis of disease. Understanding how some genes influence the manifestation of SARS-CoV-2 viral infection, is a desideratum of modern approaches to personalized therapy. With genome-wide prediction of individual pharmacological response this information can be used to obtain new pharmacological targets, and with them new active substances. All this information reinforces the concept of customized medicine, a concept that was utopian for a while, but with technological advances in molecular analysis, this concept is starting to become a reality.

This study is a short review highlighting the main components of the genetic variations in the relationship between the virus and the host organism. Individual variations of genes and their proteins expressions are also a key component in the process of understanding induced immunity after disease as well as after vaccination. Genetics of human cell membranar systems used for virus integrations host and genetics of human intracellular signalling are short presented.

Keywords. genes, virus adaptations, COVID-19 human genetic variations

1. Human genetic variations and medical’s implications

Genetic variation is the difference in DNA sequences that occurs between individuals in a population. Only variation that occurs in germ cells can be inherited from one individual to another and thus affects population dynamics and ultimately evolution. As humans multiplied and spread across the planet, they encountered thousands of local dietary variations and different diseases that generated countless selection pressures [1].

Many of the protein-coding genes are polymorphic, and it appears that those with high frequency in the population are related to selection induced by environmental factors. The selected allelic variants will generate, in populations, enclaves with considerable significance in the case of some pathologies. Human genetic variability has many causes, including SNPs (single nucleotide polymorphism), INDELs (INDELs small insertions and deletions), CNVs (copy number variation – CNV) and other genetic events that can lead to the alteration of the functions of some proteins.

SNPs represent the most common form of DNA polymorphisms (on average 1 in 1000 bases), but the frequency can be higher in certain regions such as CpG islands (1 in 100) or

lower within some genes (1 in 2500 bases). SNPs located in intergenic or synonymous regions within genes are not subject to natural selection.

Technological advances in the area of genomics, proteomics and metabolomics, have led to the development of means of identifying these variations [2, 3].

There are some important human genetic variations for medical sciences [4] that have been more intensively studied:

- variation in disease resistance, well studied for malaria, tuberculosis, hepatitis and leprosy, but not for more than a thousand other human diseases;
- variation in the ability to metabolize two products of the agricultural revolution: alcohol and milk (lactase persistence);
- variation in drug metabolizing capacity, in particular, mediated by variants in the cytochrome P450 and N-acetyl-transferase gene families;
- HLA genetic variation with implications in the immune response.

Molecular diagnostic tools are current and very valuable tools for determining individual genetic variations that can play an important role in diagnosis, but also in individual pharmacological response, being necessary for predicting pharmacological response before starting therapy.

2. The genetic variations in the relationship between the virus and the host organism

2.1. The viral adaptations

The differences generated by individual genetic variations can contribute to understanding the manifestation and severity of the disease identified at the impact of the host organism with the virus. The clinical symptoms of patients infected with SARS-CoV-2 vary greatly and are usually divided into symptomatic or asymptomatic [5, 6].

When SARS-CoV-2 infection occurs there is a constant battle between the viral load and the host's immune strength. Disease progression or a negative SARS-CoV-2 nucleic acid test would be two completely different outcomes of host-virus interactions.

Changes in serum molecular expression profile represent a framework of biological events induced by COVID-19 and could reflect the damage of cells, tissues and organs by SARS-CoV-2 [7].

Starting from these considerations and to understand the human genetic variability in correlation with the clinical manifestation of the COVID 19 disease, research have followed the relationship between the virus and the host organism.

Viral adaptations were very fast, consisting in the creation of opportunities for the virus to access the host cells, by remodeling the spike protein (protein S). Multiple mutant variants, over 5106 of the spike protein, created by independent amino acid substitutions, were identified. The impact of mutations increases transmissibility and affects the recognition by the antibodies of the host organism (camouflage), during the convalescence period.

In addition to protein S, the virus has other proteins that are more or less known, but which allow viral invasion, coordination of changes in the host cells (changes at the level of the endoplasmic reticulum, changes in ionic homeostasis) and facilitation of viral replication (**Table 1**), [8].

There are a number of non-structural viral proteins that play a role in replication and transcription (nsp 1- nsp16). These protein enzymes necessary for the formation of the viral genome are about three times higher than in other viral forms. The important enzyme is RNA

polymerase, but this is aided by other nonstructural proteins (nsp) to complete multiplication and synchronization of the viral cycle with complex cellular mechanisms of the host [9].

Table 1. Viral proteins and their role in the process of viral transmissibility and damage to the host cell

Viral proteins	Key role and host interactions
Structural proteins	
Spike protein	<ul style="list-style-type: none"> - ensure the attachment of the virus to the receptors on the surface of the host cell. - facilitates viral entry into the host cell. - induces cell fusion between infected and adjacent, uninfected cells generating multinucleated entities or syncytia, a strategy to allow the direct spread of the virus between cells and to undermine the antibodies that neutralize the virus.
Membranare protein	<ul style="list-style-type: none"> - is the most abundant structural protein and defines the shape of the viral coat.
Nucleocapsid protein	<ul style="list-style-type: none"> - involved mostly in processes related to the viral genome. - associated with other aspects of the CoV replication cycle. - associated with the host's cellular response to viral infection.
Envelope protein	<ul style="list-style-type: none"> - is abundantly expressed inside the infected cell during the replication cycle, but only a small fraction is incorporated into the membrane. - participates in CoV assembly and budding. - controls intracellular traffic, namely ER, Golgi and ERGIC (intermediate compartments between ER and G), - forms/permeabilizes Ca²⁺ channel at ER level. - acts as a viroporin after insertion into host membranes to generate a hydrophilic pore that allows ion transport altering cellular homeostasis.
Non-structural proteins (Coronavirus replicase proteins)	
Nsp 1	<ul style="list-style-type: none"> - mRNA degradation from host cell
Nsp 4, NSP 6	<ul style="list-style-type: none"> - organelle formation
Nsp 10	<ul style="list-style-type: none"> - the ribosomal regulation

2.2. Genetics of human cell membranar systems used for virus attachment and entry

SARS-CoV-2 entry into the host cell is mediated by the molecular interaction between virus and human proteins such as metalloproteinase angiotensin receptor (ACE 2). The activity of this peptidase is critical for the virion to have access to the host cell [10, 11].

From a genetic point of view, 1700 variants of the ACE2 gene, located on the X chromosome, are known. Also, 15 unique variants are associated with the quantitative regulation of gene expression. At the gene level, 14 SNPs and 1 insertion/deletion (INDEL) polymorphisms with higher minor allelic frequencies (MAF) were noted in the Asian population than in a European population [11].

Other host enzyme systems are involved in the progression or facilitation of viral entry: cathepsins, serine protease, transmembrane protease (TMPRSS2), TMPRSS4 or trypsin-like protease [12].

The activities of these enzymes are conditioned and nuclear coordinated by the genes that specifically encode these proteins and differ from one individual to another. Proteolytic cleavage is followed by fusion of viral and cell membranes. The more numerous these enzymes are because of gene activation, the faster the penetration of the virus. The cell membrane protease that contains the disintegration domain ADAM17 (a disintegrin and metalloproteinase 17) can cleave the domain of the ACE2 receptor that recognizes the spike protein and once released extracellularly will keep the virus bound/blocked at a distance from the host cell membrane. Three unique variant SNPs inducing strong differences in allelic profiles between European and Asian populations were identified for ADAM17 protein [13]. Consequently, the virus that has entered the body can be detected serologically, but it will not induce symptoms of the disease because it has not also entered the intracellular level.

Inhibiting the activity of proteases required to cleave viral spike proteins: for example, inhibiting the enzymatic activity of ADAM17 and TMPRSS2 could serve as novel therapeutic targets [14]. This could block the viral interaction with the receptor and its entry into cells.

ACE receptors are expressed in almost all tissues, while ACE2 is expressed by alveolar epithelial cells and capillary endothelial cells. ACE2 is highly expressed in the lung and kidney, but also in the gut and brain. Genetic polymorphisms of ACE and ACE2 are associated with hypertension, cardiovascular disease, stroke, and diabetes [12].

2.3. Genetics of human intracellular signaling in SARS-CoV-2 infection

Effects of molecular activation, at the cellular level, can induce various signaling pathways, including serine/threonine kinase, ERK, JNK/MAPK as well as PKC. These considerations offer new visions for the development of targeted therapies, because the signaling pathways triggered by ACE and ACE2 increase cytokines (IL-6, TNF α), increase oxidative damage by ROS and amplify endothelial damage by inhibiting NO synthesis and vasoconstriction. Therefore, inhibition of only one of its targets, eg IL-6, may not provide significant therapeutic benefit in these patients [15, 16].

Individual variations are also a key component in the process of understanding induced immunity after disease as well as after vaccination. Further studies on long-term COVID-19 showed that abnormal patterns of immune cell activation persisted in patients' blood during the recovery period and following the onset illness 170 days after discharge [17, 18].

The adaptive evolution of the dominant immune cell subpopulations associated with the new phenotype imposed by the impact with the virus is an important strategy for the immune response [19].

Table 2. The human genes with regulatory mechanisms at a distance associated with COVID-19

Genes	Proteins	Phenotip	Molecular genetics/Expression
<p>SLC6A20</p> <p>3p21.31</p>	<p>membranar transporter</p>	<p>Na(+) and Cl(-)-coupled transporter proteins mediate transit of structurally related small hydrophilic substances across plasma membranes. Expression in human kidney and small intestine, with no expression detected in other tissues.</p>	<p>Renal glycosuria and hyperglycinuria without increased excretion of other amino acids were the features observed of 3 generations of 1 kindred with probable autosomal dominant inheritance [27]</p>
<p>IFNAR2</p> <p>21q22.11</p> <p>(https://www.genecards.org/cgi-bin/carddisp.pl?gene=IFNAR2)</p>	<p>type I membrane protein that forms one of the two chains of a receptor for interferons alpha and beta</p>	<p>Human IFNAR2 deficiency causes fatal susceptibility to live viral vaccines, revealing a vital but narrow nonredundant role for IFN-α/β in viral immunity. Type I interferon (IFN-α/β) is a fundamental antiviral defense mechanism.</p>	<p>A major susceptibility locus for hepatitis B virus [28].</p> <p>Patient cells showed no IFNAR2 expression and were unresponsive to IFNA1/IFNB1 signaling with increased susceptibility to viral infection. These cellular defects could be rescued by transduction of wildtype IFNAR2, including reinstating the ability to control the replication of IFN-sensitive viruses. The patient developed fatal encephalitis following routine immunization with measles/mumps/rubella (MMR) vaccine.</p>
<p>CCR9</p> <p>3p21.31</p> <p>(https://www.ncbi.nlm.nih.gov/gene/10803)</p>	<p>chemokine, CC motif, receptor 9</p>	<p>Chemokines also play fundamental roles in the development, homeostasis, and function of the immune system, and they have effects on cells of the central nervous system as well as on endothelial cells involved in angiogenesis or angiostasis.</p>	<p>This gene and its exclusive ligand, chemokine 25, are overexpressed in a variety of malignant tumors and are closely associated with tumor proliferation, apoptosis, invasion, migration and drug resistance. CCR9 transcript that was highly expressed in thymus and weakly expressed in lymph nodes and spleen.</p>

			Multiple transcript variants encoding different isoforms have been found for this gene.
TLR3 4q35 https://www.genecards.org/cgi-bin/carddisp.pl?gene=TLR3	transmembrane cell-surface receptors, which have a key role in the innate immune system	It recognizes dsRNA associated with viral infection, and induces the activation of NF-kappaB and the production of type I interferons. It thus plays a role in host defense against multiple viruses.	The various TLRs exhibit different patterns of expression. This receptor is most abundantly expressed in placenta and pancreas, and is restricted to the dendritic subpopulation of the leukocytes
TYK2 https://www.genecards.org/cgi-bin/carddisp.pl?gene=TYK2&keywords=TYK2	protein associates with the cytoplasmic domain	Tyrosine kinase of the non-receptor type involved in numerous cytokines and interferons signaling, which regulates cell growth, development, cell migration, innate and adaptive immunity .	This gene has 35 transcripts (splice variants), 197 orthologues, 32 paralogues and is associated with 4 phenotypes. http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000105397;r=19:10350533-10380608 A mutation in this gene has been associated with Immunodeficiency 35. Ubiquitous expression in spleen (RPKM 21.6), bone marrow (RPKM 20.9) and 25 other tissues [29]

Numerous studies have demonstrated that SARS-CoV-2 infection can cause a variety of potential sequelae, including impairment of lung function or imaging abnormalities, cardiovascular injury, neuroinflammation and nerve injury, gastrointestinal injury, and kidney injury [20-23].

Zhang X. *et al.* (2023) in a complex study used information that includes molecular patterns, in patients with COVID 19, of serum abnormalities and their effects on hundreds of cells signaling pathways, thousands of functions of different biological systems, cardiac, liver and kidney organ toxicity and other biological events on cells, tissues and target organs infected with SARS-CoV-2.

Combining information on the abundance represented by synthesized proteins and protein activity correlated with regulatory factors (genes, RNA, transcription factors) and proteins important in the cellular responses with complex mechanisms (**Table 2**), suggests a new perspective to find drugs that can counteract COVID-19 at the molecular level throughout the course of the disease [24-26].

3. Conclusions

Genetic variability is the basis of understanding many processes associated with pathogenesis and can provide support for a series of therapeutic approach mechanisms. The COVID-19 pandemic marked a crucial moment for humanity that laid the foundations for broad interdisciplinary approaches and the consolidation of scientific interventions.

The way the disease manifests, the individual or population differences identified both in the severity of the disease and in the response to therapy required a specific approach and updated the concept of personalized medicine.

References

- [1] Abi-Rached L, Jobin MJ, Kulkarni S, McWhinnie A, Dalva K, Gragert L, Babrzadeh F, Gharizadeh B, Luo M, Plummer FA, Kimani J, Carrington M, Middleton D, Rajalingam R, Beksac M, Marsh SG, Maiers M, Guethlein LA, Tavoularis S, Little AM, Green RE, Norman PJ, Parham P. The shaping of modern human immune systems by multiregional admixture with archaic humans. *Science*. 2011 Oct 7;334(6052):89-94. doi: 10.1126/science.1209202. Epub 2011 Aug 25. PMID: 21868630; PMCID: PMC3677943
- [2] William B. Coleman, Gregory J. Tsongalis, *Molecular Pathology, The Molecular Basis of Human Disease*, 2009 Academic Press. Elsevier Science, 664 pp.
- [3] Debra G.B. Leonard Editor, *Molecular Pathology in Clinical Practice*, 2016, Second Edition Springer International Publishing Switzerland, XIX,1001 pp.
- [4] Gomez F, Hirbo J, Tishkoff SA. Genetic variation and adaptation in Africa: implications for human evolution and disease. *Cold Spring Harb Perspect Biol*. 2014 Jul 1;6(7):a008524. doi: 10.1101/cshperspect.a008524. PMID: 24984772; PMCID: PMC4067985].
- [5] Mousavizadeh L, Ghasemi S. Genotype and phenotype of COVID-19: Their roles in pathogenesis. *J Microbiol Immunol Infect* [Internet]. 2021;54(2):159–63. Available from: <https://doi.org/10.1016/j.jmii.2020.03.022>
- [6] Gao Z, Xu Y, Sun C, Wang X, Guo Y, Qiu S. et al. A systematic review of asymptomatic infections with COVID-19. *J Microbiol Immunol Infect*. 2021; 54:12–6.
- [7] Zhang X, Sh Y, Dong J, Chen Z, Hong F. The landscape of abnormal pathway activation confers COVID-19 patients' molecular sequelae earlier than clinical phenotype. *Theranostics*. 2023 Jun 12;13(10):3451-3466. doi: 10.7150/thno.83405. PMID: 37351167; PMCID: PMC10283057.
- [8] Harvey, W.T., Carabelli, A.M., Jackson, B. et al. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol* **19**, 409–424 (2021). <https://doi.org/10.1038/s41579-021-00573-0>.
- [9] Malone, B., Urakova, N., Snijder, E.J. et al. Structures and functions of coronavirus replication–transcription complexes and their relevance for SARS-CoV-2 drug design. *Nat Rev Mol Cell Biol* **23**, 21–39 (2022). <https://doi.org/10.1038/s41580-021-00432-z>
- [10] Samavati L and Uhal BD (2020) ACE2, Much More Than Just a Receptor for SARS-CoV-2. *Front. Cell. Infect. Microbiol*. 10:317. doi: 10.3389/fcimb.2020.00317
- [11] Cao Y, Li L, Feng Z, Wan S, Huang P, Sun X, Wen F, Huang X, Ning G, Wang W. Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. *Cell Discov* [Internet]. 2020;6(1):4–7. Available from: <http://dx.doi.org/10.1038/s41421-020-0147-1>.
- [12] Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., et al. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by

- a clinically proven protease inhibitor. *Cell*. 181, 271–280. doi: 10.1016/j.cell.2020.02.052].
- [13] Delanghe, Joris R., and Marijn M. Speeckaert. “Chapter Two - Host Polymorphisms and COVID-19 Infection.” In *Advances in Clinical Chemistry*, edited by Gregory S. Makowski, 107:41–77. Elsevier, 2022. <https://doi.org/10.1016/bs.acc.2021.07.002>].
- [14] Shapira T, Monreal IA, Dion SP, Buchholz DW, Imbiakha B, Olmstead AD, Jager M, Désilets A, Gao G, Martins M, Vandal T, Thompson CAH, Chin A, Rees WD, Steiner T, Nabi IR, Marsault E, Sahler J, Diel DG, Van de Walle GR, August A, Whittaker GR, Boudreault PL, Leduc R, Aguilar HC, Jean F. A TMPRSS2 inhibitor acts as a pan-SARS-CoV-2 prophylactic and therapeutic. *Nature*. 2022 May;605(7909):340-348. doi: 10.1038/s41586-022-04661-w. Epub 2022 Mar 28. PMID: 35344983; PMCID: PMC9095466].
- [15] Brest, Patrick, Sadal Refae, Baharia Mograbi, Paul Hofman, and Gerard Milano. “Host Polymorphisms May Impact SARS-CoV-2 Infectivity.” *Trends in Genetics* 36, no. 11 (November 1, 2020): 813–15. <https://doi.org/10.1016/j.tig.2020.08.003>.
- [16] Liu T, Zhang L, Joo D, Sun SC. NF- κ B signaling in inflammation. *Signal Transduct Target Ther*. 2017;2 (April).
- [17] Burn E, Roel E, Pistillo A, Fernández-Bertolín S, Aragón M, Raventós B. et al. Thrombosis and thrombocytopenia after vaccination against and infection with SARS-CoV-2 in Catalonia, Spain. *Nat Commun*. 2022;13:7169
- [18] Asadi K, Gholami A. Virosome-based nanovaccines; a promising bioinspiration and biomimetic approach for preventing viral diseases: A review. *Int J Biol Macromol* [Internet]. 2021;182:648–58. Available from: <https://doi.org/10.1016/j.ijbiomac.2021.04.005>
- [19] Maucourant C, Filipovic I, Ponzetta A, Aleman S, Cornillet M, Hertwig L, Strunz B, Lentini A, Reinius B, Brownlie D, Cuapio A, Ask EH, Hull RM, Haroun-Izquierdo A, Schaffer M, Klingström J, Folkesson E, Buggert M, Sandberg JK, Eriksson LI, Rooyackers O, Ljunggren HG, Malmberg KJ, Michaëlsson J, Marquardt N, Hammer Q, Strålin K, Björkström NK; Karolinska COVID-19 Study Group. Natural killer cell immunotypes related to COVID-19 disease severity. *Sci Immunol*. 2020 Aug 21;5(50):eabd6832. doi: 10.1126/sciimmunol.abd6832. PMID: 32826343; PMCID: PMC7665314.
- [20] Korompoki E, Gavriatopoulou M, Hicklen RS, Ntanasis-Stathopoulos I, Kastritis E, Fotiou D, Stamatelopoulos K, Terpos E, Kotanidou A, Hagberg CA, Dimopoulos MA, Kontoyiannis DP. Epidemiology and organ specific sequelae of post-acute COVID19: A narrative review. *J Infect*. 2021 Jul;83(1):1-16. doi: 10.1016/j.jinf.2021.05.004. Epub 2021 May 14. PMID: 33992686; PMCID: PMC8118709.
- [21] Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. *Nature*. 2021;594:259–64.
- [22] Al-Aly Z, Bowe B, Xie Y. Long COVID after breakthrough SARS-CoV-2 infection. *Nat Med*. 2022;28:1461–7.
- [23] Varney JA, Dong VS, Tsao T, Sabir MS, Rivera AT, Ghula S, Moriles KE, Cherukuri ML, Fazal R, Azevedo CB, Mohamed RM, Jackson GR, Fleming SE, Rochez DE, Abbas KS, Shah JH, Minh LHN, Osman F, Rafla SM, Huy NT. COVID-19 and arrhythmia: An overview. *J Cardiol* [Internet]. 2022;79(4):468–75. Available from: <https://doi.org/10.1016/j.jcc.2021.11.019>.
- [24] Zhang X, Sh Y, Dong J, Chen Z, Hong F. The landscape of abnormal pathway activation confers COVID-19 patients' molecular sequelae earlier than clinical phenotype.

- Theranostics. 2023 Jun 12;13(10):3451-3466. doi: 10.7150/thno.83405. PMID: 37351167; PMCID: PMC10283057.
- [25] Smemo, S. et al. Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* **507**, 371–375 (2014).
- [26] Kwok, A.J., Mentzer, A. & Knight, J.C. Host genetics and infectious disease: new tools, insights and translational opportunities. *Nat Rev Genet* **22**, 137–153 (2021). <https://doi.org/10.1038/s41576-020-00297-6>
- [27] Kaser, H., Cottier, P., Antener, I. Glucoglycinuria, a new familial syndrome. *J. Pediat.* 61: 386-394, 1962. [PubMed].
- [28] Duncan CJ, Mohamad SM, Young DF, Skelton AJ, Leahy TR, Munday DC, Butler KM, Morfopoulou S, Brown JR, Hubank M, Connell J, Gavin PJ, McMahon C, Dempsey E, Lynch NE, Jacques TS, Valappil M, Cant AJ, Breuer J, Engelhardt KR, Randall RE, Hambleton S. Human IFNAR2 deficiency: Lessons for antiviral immunity. *Sci Transl Med.* 2015 Sep 30;7(307):307ra154. doi: 10.1126/scitranslmed.aac4227. PMID: 26424569; PMCID: PMC4926955.
- [29] Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, Habuka M, Tahmasebpoor S, Danielsson A, Edlund K, Asplund A, Sjöstedt E, Lundberg E, Szigyarto CA, Skogs M, Takanen JO, Berling H, Tegel H, Mulder J, Nilsson P, Schwenk JM, Lindskog C, Danielsson F, Mardinoglu A, Sivertsson A, von Feilitzen K, Forsberg M, Zwahlen M, Olsson I, Navani S, Huss M, Nielsen J, Ponten F, Uhlén M. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics.* 2014 Feb;13(2):397-406. doi: 10.1074/mcp.M113.035600. Epub 2013 Dec 5. PMID: 24309898; PMCID: PMC3916642.