COVID-19: The ethno-geographic perspective of differential immunity.

Usman Abdullah[a], Ned Saleh[b], Peter Shaw[c*], Nasir Jalal[c*]

[a]*Department of Biomedical Sciences, Pak-Austria Fachhochschule, Pakistan.*[b] *Syncal AHLLC. USA.*[c*] *Oujiang Lab, Wenzhou, China.*

Author details and affiliations:

1. Usman Abdullah
Designation: BS student
Pak-Austria Fachhochschule, Haripur, Pakistan
Email: kirito63194@gmail.com; b20f0005bms01@fbse.paf-iast.edu.pk
ORCID: 0000-0003-3920-0563
LinkedIn: www.linkedin.com/in/usman-abdullah-512b33232

2. Ned Saleh (Ph.D.)
Designation: Founder
Synsal-AHLLC, USA
Email: ned.saleh@synsal.com
Linkedin : linkedin.com/in/ned-saleh-phd-3224293

3. *Peter Shaw (Ph.D.)
Co-Corresponding author
Designation: Professor
Oujiang Lab, Wenzhou, China.
Email: petershaw@ojlab.ac.cn
ORCID: 0000-0002-3187-8938
Linkedin : linkedin.com/in/peter-shaw-a775338

4. *Nasir Jalal (Ph.D.)
Corresponding author
Designation: Associate Professor
Oujiang Lab, Wenzhou, China.
Email: nasirjalal@ojlab.ac.cn
ORCID: 0000-0002-6160-0988
LinkedIn: linkedin.com/in/nasir-jalal-member-royal-society-of-biologists-25909023

Second affiliation
Designation: Associate Professor
Biomedical Sciences Department
Pak-Austria Fachhochschule, Haripur, Pakistan.
Email: nasir.jalal@fbse.paf-iast.edu.pk

 doi:10.20944/preprints202212.0582.v1

ABSTRACT

Coronavirus disease 2019 (COVID-19), the emissary behind the worst global pandemic of the 21st century, is primarily a respiratory disease-causing virus called SARS-CoV-2 which is responsible for millions of new cases (incidence) and deaths (mortalities), worldwide. Many factors have played a role in the differential morbidity and mortality experienced by nations and ethnicities against SARS-CoV-2, such as the quality of primary medical health facilities or enabling economies. Nevertheless, the most important variable, i.e., the subsequent ability of individuals to be immunologically sensitive or resistant to the infection, was not properly discussed before. Therefore, an astounding issue arose when some developed countries experienced higher morbidity and mortality, compared with their relatively underdeveloped counterparts, despite having excellent medical health facilities. Hence this investigative review attempts to analyze the issue from an angle of previously undiscussed genetic, epigenetic, and molecular immune resistance mechanisms in correlation with the pathophysiology of SARS-CoV-2 and varied ethnicity-based immunological responses against it. The biological factors discussed here include the overall landscape of human microbiota, endogenous retroviral genes spliced into the human genome, copy number variation, and how they could modulate the innate and adaptive immune systems, which put a particular ethnic genetic architecture at a higher risk of SARS-CoV-2 infection than others. Considering an array of these factors in their entirety may help explain the geographic disparity of disease incidence, severity, and subsequent mortality associated with the disease while at the same time encouraging scientists to design new experimental approaches to investigation.

Keywords: COVID-19, copy number variation (CNV), virome, microbiome, endoretroviral genome (ERV), geographic disparity.

 doi:10.20944/preprints202212.0582.v1

1. INTRODUCTION

The global pandemic, nicknamed coronavirus disease-2019 (COVID-19), is a disease that is caused by the severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2). The virus has caused about 652 million confirmed cases of COVID-19 and 6.5 million confirmed deaths until the filing of this article.[1] Interestingly 64 percent of the morbidities (i.e., 55 million) could be linked to race/ethnicity, while 85 percent of the mortalities (i.e., 750,000) could be linked to race/ethnicity.[2] The world has been fighting this pandemic by developing new treatment methods while also trying to administer vaccines to achieve herd immunity in public. About 10925, 055,390 doses of various vaccines (RNA, protein, and attenuated virus-based) have been manufactured and shipped worldwide, causing the pandemic to slow, but still, it is far from being over.[1] SARS-CoV-2 is a beta corona virus with a positive sense, single-stranded RNA about 30Kb in size.[2] The virus causes infection of the upper respiratory tract in mild infections, but as the disease progresses, pneumonia and hypoxemia may develop. A patient suffering from a severe version of the disease may develop acute respiratory distress syndrome (ARDS), shock, encephalopathy, myocardial injury, and acute kidney failure.[3] Some risk factors already identified for severe and potentially fatal COVID-19 illness include age, sex, and pre-existing medical conditions such as hypertension, diabetes mellitus, cancer, and autoimmune conditions. We have observed disparity in the general severity and mortality rates of COVID-19 across various geographic regions predominantly occupied by specific ethnic groups or populations. Nations, such as Bangladesh and Pakistan, which were previously predicted to be on the worse end of the pandemic, remain seemingly unimpaired, all the while exhibiting low numbers of both mortality and severity of the disease.[4] Even though these developing countries have relatively poor healthcare infrastructures and high population densities, accompanied by poor hygienic conditions among the population, which essentially create a perfect breeding ground for communicable respiratory diseases such as COVID-19, thus the hitherto predicted high viral transmissibility rates (R_0).[4] Nevertheless, these developing countries' actual morbidity, mortality, and transmissibility rates are noticeably low. Whereas if we consider the western, more developed countries in Europe and North America, the severity of the disease and its mortality rates are comparatively much higher, with mortality rates reaching as high as 30 percent in some well-developed countries, such as Belgium, during different phases of the pandemic.[5] With the risk of viral transmission rising in these regions, more infectious strains/variants have appeared, such as omicron.[6] Some studies have even suggested that 37 percent of all mortalities of the pandemic came from the region known as the European Union (EU).[7] In contrast, the EU only accounts for about 9.78 percent of the world's population.[8]

Despite these discouraging statistics, Europe has a much better healthcare infrastructure, hygienic conditions, and better health awareness than developing countries such as Pakistan and Bangladesh. Until now, the biggest argument to justify this disparity was the population's average age in these regions. In the case of Pakistan, about 60 percent of the population was recorded to be below the age of 30 years.[9] Nevertheless, in this review, we argue that there must be some other biological factors at play that may also contribute to this disparity. Thus, the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 December 2022 doi:10.20944/preprints202212.0582.v1

following article is more about proposing a viable hypothesis based on previous studies and the authors' deductions; however, all of the forthcoming hypotheses defend the demographic disparity idea; in other words, the biological mechanisms show variability among populations from different demographic regions. We also briefly discuss the role of these biological phenomena in viral infectability and repercussions for the more comprehensive host immune system and present evidence from existing literature on how they show some variation among members of different populations.

Data Availability Statement

We referred to the online, publicly available COVID-19 data-sharing resources such as WHO and Kaiser Family Foundation (KFF) for most data discussed in this article. Besides these resources, we also consulted peer-reviewed articles (as referenced) that provide insight into the phenomena discussed.

2. Human Microbiota

The human body acts as a host for approximately trillions of pathogens, bacteria, viral particles, viruses, and bacteriophages that help us maintain various internal mechanisms of our body for homeostasis.[10] These organisms that collectively form the microbiota of the human body are, as yet, the unrecognized agents that play a crucial role in the overall working of our body. Some of those processes are:

1. Digestion and ability to harvest nutrients. [11]
2. Modulation of the appetite signal of the body;[11]
3. Production of important vitamins for the body (including vitamin K);[12]
4. Regulation of epithelial growth and development;[13]
5. The metabolization of various xenobiotics in the body; [13]
6. Modulation of the host immune system.[14]

The viral and bacteriophage components of the microbiota that make up the "virome" are found in almost every organ and organ system of the body, with the largest reservoir being the gastrointestinal system.[14] While there have been studies proving the presence of these viral microbiotas detected in the oral cavity, [14] healthy human lungs and respiratory systems also host these viral microbiotas in the air passages.[15] These microbiotas play a role in the human immune system through (Fig. 1) the following main methods.

- The bacteriophages in our body help fight against recurring bacterial infections by first targeting the bacterial pathogen and then through host immunomodulation by developing an adaptive memory of that initial infection.[16]
- The most notable is how it fights off other incoming viruses due to competition in the region while also immune priming the body in preparation for infection by other external pathogenic viruses.[15]
- Viral-induced interferon-α (IFN-α) production helps against viral infections.[17]

The antimicrobial peptides such as the mouse Regenerating islet-derived protein III-gamma (Reg-IIIγ) and its human analog, Pancreatitis-associated protein (HIP/PAP), a subset of the calcium-dependent C-type lectin superfamily, bind their bacterial targets via a peptidoglycan and represent a primitive form of innate immunity. The expression of these antimicrobial proteins require signals from the microbiota.[18]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 December 2022 doi:10.20944/preprints202212.0582.v1

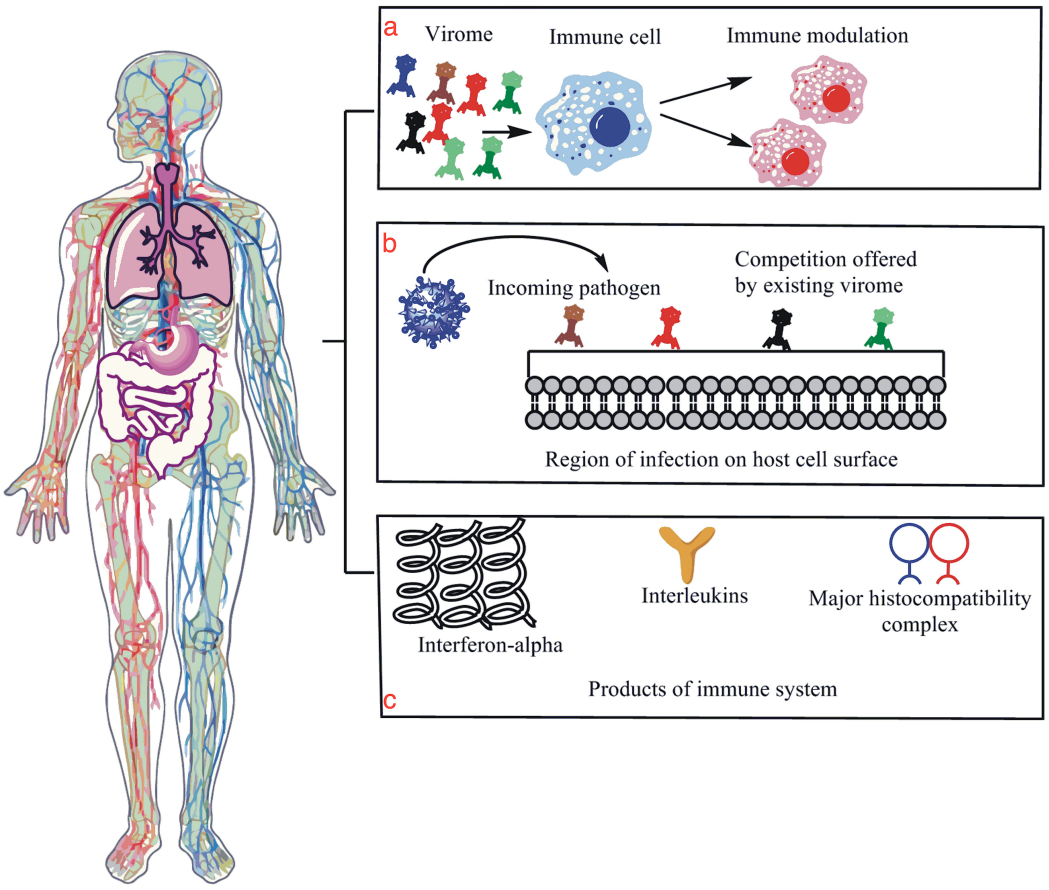

Fig. 1: This figure is an illustration of the human microbiome, specifically the human virome and its few functions related to the immunity against viral infection. (a) is an illustration that shows modulation or the positive reinforcement of the host immune system by virome molecules and particles that enhance the immune system's ability to fight the incoming infection. (b) is a diagram illustrating the competition that is given to the incoming viral particles by the host virome which essentially protects the host cells from infection as the virome particles act as guards of the region. (c) shows the production of IL, Interferon, MHC molecules all of which are essential parts of the host immune system and their production with the help of virome demonstrates their role in the proper functioning of the host's immune system.

Many large-scale studies have illustrated the variation in the metagenomic and meta-communal microbiota, especially the virome data among people of different ethnicities and geographic locations.[19,20] Indeed, other pre-disposition factors exist, such as diet and age, but in one study, geographic factors presented the most significant reason for virome and overall microbiome variability.[20]

Having established a connection between the immune functions of any viral microbiome and the geographic variability of the virome,[21] we present the hypothesis that the geographically variable virome and overall microbiome can play a role in regulating the severity and mortality rates of COVID-19.

As per our hypothesis, people in some regions may have a microbiome and virome that consists of viral particles and phages that offer a more fierce competition against SARS-CoV-2 and pneumococcus (cause of pneumonia and resultant co-morbidity associated with COVID-19), thus lowering the chances of mortality and disease progression into a more severe form; giving these individuals an edge of survival over similar people who were exposed to different viral and microbial environments in other geographical locations.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 December 2022 doi:10.20944/preprints202212.0582.v1

3. Endogenous Retroviruses

Previously accepted and understood as part of junk DNA or part of the "genomic dark matter," endogenous retroviral genomic components account for more than 8 percent of our whole genome.[22] These endogenous retroviruses used to be part of exogenous pathogenic viruses[23] that infected the cells of our ancestors but remained in the lysogenic phases and eventually became a part of the vertical host DNA [24] over time.[25] It was also previously thought that this part of our genome plays no role in the body's normal functioning, but this belief could not have been more wrong as we now know, owing to more extensive research studies) that this part of the human genome has a fundamental role in antiviral immunity, among other functions.[26] Previous retroviral infections can also play a role in modulating the immune response; one such example comes from the human endogenous retroviruses (HERVs), which impart antiviral immunity through the following mechanisms (fig. 2):

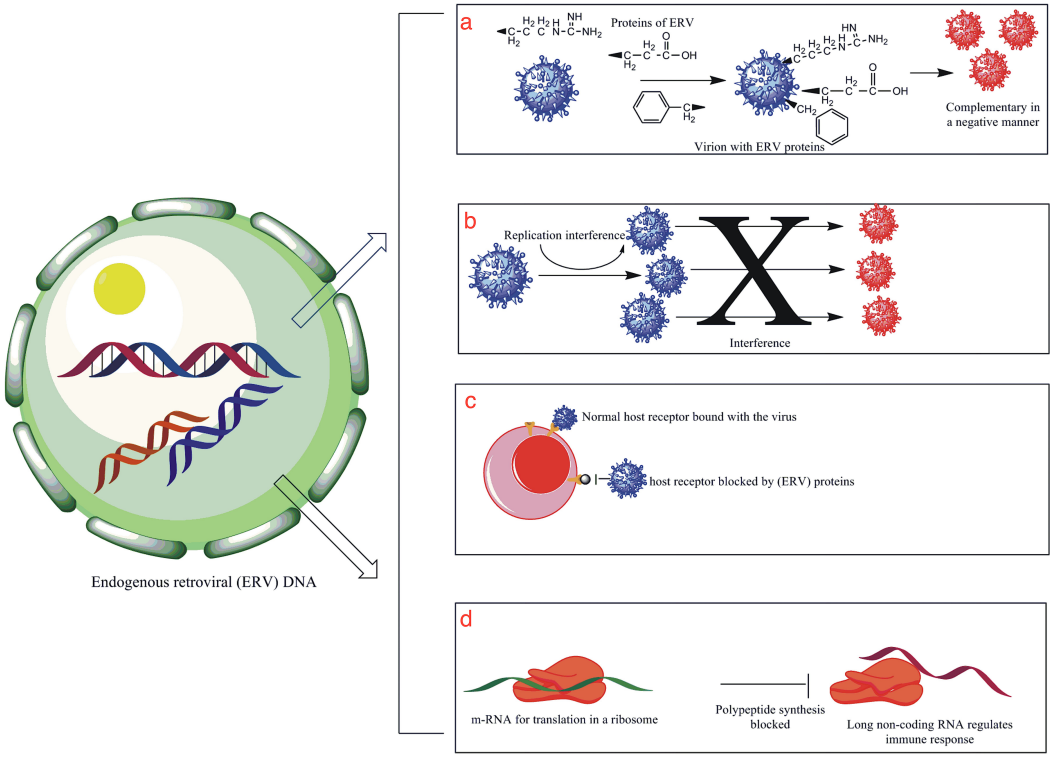

Fig. 2: This is an illustration of some of the functions that HERV genome plays in the protection of the human body against external viral infection. (a) shows production of viral proteins by the HERV genome that are complimentary to the external virus in a negative way thus affecting their success. (b) is an illustration showing the interference produced by HERV proteins that result in inhibiting or blocking the viral replication cycle. (c) is an illustration showing the phenomenon of receptor interference that involved the blockage of the site of entry for the foreign virus by HERV proteins thus resulting in the prevention of viral infection of the cell. (d) is an illustration that shows the production of long non-coding RNA sequences that help in modulation of the host antiviral immune system.

3.1. Activation Of Innate Immune System:

The long-misunderstood "Junk DNA" plays an important role in many functions. One of those many functions is providing an antiviral immune response. Some junk DNA comprises spliced endoretroviral (ERV) genome particles. These pieces of ERV act as a genetic memory and trigger the innate immune system sensing cascades through HERV-synthesized polypeptides in a much shorter time, resulting in a much more efficient immune response against recurring viral infections.[27]

 doi:10.20944/preprints202212.0582.v1

3.2 Production of long non-coding RNAs (lncRNAs):

Regulating the antiviral immune response is achieved by producing long non-coding RNAs (lncRNAs).[28] An example of this would be the production of murine endogenous retrovirus (ERV)-derived lncRNA termed lnc-EPAV that plays a role by boosting the antiviral gene expression. Besides that, it also increases the production of antiviral cytokines such as IFN-β and IL-6. As a positive feedback loop, the ERV pathway of the lnc-EPAV also increases the antiviral immune response. Hence the SFPQ, a proline, glutamine-rich splicing factor that plays a role in splicing RNA molecules related to the immune system, also interacts with multiple HERV-derived RNAs, suggesting a similar immune response in humans. This SFPQ splicing factor participates in diverse molecular functions, including paraspeckle formation, microRNA synthesis, and transcription regulation, and has been shown to regulate the host's innate immune response to viruses. [29]

3.3. Modulation of the immune system by ERV based proteins:

Modulating the immune activation is accomplished by the production of immune cascade-activating proteins.[31] This modulation occurs by the deliberate triggering of plasma membrane-associated toll-like receptors (TLRs) by the retroviral proteins, which cause the release of massive amounts of IL-1β, IL-6, TNF-α, and IL-12p40. The retroviral proteins also cause the release of nitric oxide, which plays an essential role in the innate immune system.[30]

3.4. Receptor interference:

Blocking the virus entry receptors by their envelopes is done through a phenomenon called receptor interference.[31,32] The viral gene products of ERV genomes, called ENVs, are essential components of viral membranes that mediate receptor interference by binding to the host cell receptors before the pathogenic particle can bind. The ERV-derived ENVs (viral gene protein products) might as well bind to the newly synthesized entry receptors, thus preventing their appearance and transport to the cell surface, essentially closing the gateway of the virion entry into the host cell. The ENV proteins conformed with this phenomenon of receptor interference of avian leukosis viruses which were observed to induce resistance to the exogenous viruses of the same subgroup. [32, 33]

Angiotensin-converting enzyme II (ACE2) is the primary receptor of entry for SARS-CoV-2 inside human host cells.[33] Studies have shown that this virus could not enter the ACE2 receptor null host cells.[34] The spike (S) protein of the virus binds with high affinity to the ACE2 receptor, thus modulating the effective entry of the virus into the host cell.[35] Host proteins such as the Transmembrane Serine Protease 2 (TMPRSS2) are known to facilitate the entry mechanism of SARS-CoV-2 into the cells by cleaving the viral S-protein, thus allowing for effective viral fusion with the host cell.[36] Therefore, we propose that if some agonistic proteins facilitate the entry of the virus, then some antagonistic proteins must also act as a blockade or interference to the receptor-based viral entry of SARS-CoV-2. Hence, there might be an ERV-based protein that could explain the low susceptibility of people from a certain population to the virus. We propose that this hypothesis could be tested in a lab, in which case, it may allow us to use the presence of that unknown protein to produce drugs that could upregulate the genetic expression of the proteins responsible for receptor interference or exogenously inject recombinant ERV proteins that mediate ACE2 receptor interference.

 doi:10.20944/preprints202212.0582.v1

3.5. Complementing in a negative manner:
Complementing the external virion in a dominantly negative manner with their proteins can be another factor of differential immunity against SARS-CoV-2.[37,38] An example would be a reduction in the viral infectivity and susceptibility because of the murine FV-4 (a gene that has shown a reduced susceptibility to murine leukemia virus in mice).[38]

3.6. Interference in the replicative cycle:
Creating interference in the replicative cycles of the incoming viral particles was successfully demonstrated in a lab using a mouse gene called Fv1. This gene encodes a restriction fragment that inhibits the entry of murine leukemia virus and seems to have been inserted into the murine genome by an ancient provirus about 45 million years ago. These studies also suggest that retroviruses are adapted to living within their hosts through the vertebrate evolutionary process. [39,40]

3.7. Increase in the diversity and plasticity of immune genes
An increase in the plasticity of the genes that modulate and regulate the immune responses can be observed in the recombination process with the silent ERV genes.[41,42] This homologous recombination may result in the loss of genes or gene fragments in monogenic systems. Nevertheless, let us consider the multigenic families. The risk-to-reward ratio is different, and the recombination or insertion of the ERV sequences may allow the immune system to facilitate the viral pathogen better. One example would be the major histocompatibility complex (MHC) locus in primates. Such involvement and recombination of the ERV genetic segments has resulted in increased diversity of alleles and thus increased the spectrum of viral peptides that can be presented to T cells. This increased diversity in the viral peptides due to the diversity of alleles causes a more robust immune system that is better able to fight against diverse groups of pathogens.[42–47]

3.8. LTR's based enhancement of antiviral genes
Regulating the antiviral gene expression with their promoters and enhancers,[48,49] there are a lot of promoter, inhibitor, and enhancer sites in the ERV's long-term repeats (LTRs) that promote the encoding of antiviral proteins such as INF-α. Thus, the immune system can be boosted or enhanced by the function of the ERV and LTR promoter and enhancer sites.
All the mechanisms of the retroviral genome mentioned above contribute to our defense against pathogenic infections, especially those caused by disease-causing viruses. Furthermore, this biological factor, like the one discussed above, is also variable among the populations of different geographical locations.[50] This phenomenon leads us to speculate that this may add to the disparity observed in the severity and mortality rates of COVID-19 cases among people living in different demographic locations.

4. Previous unrecorded local epidemics
Pakistan is a region with insufficient focus on medical and biological sciences, especially epidemiology. The country faces many flu-based epidemics each winter season, but some are recorded with that much enthusiasm, such as the COVID-19 situation. This hypothesis is based on pure speculation with little evidence due to the lack of public records or pre-existing data in this field. Nevertheless, let us investigate the biology of how a previous epidemic may have been affecting the population of a region concerning COVID-19. Being infected with a similar upper

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 December 2022 doi:10.20944/preprints202212.0582.v1

respiratory tract flu infection that had milder effects on the host may have caused the immune system to adapt to the succeeding SARS-CoV-2 infection so that adaptive immunity clears the infection more quickly.

5. Copy number variation

Copy number variation (CNV) refers to the variation in the number of a specific gene repeated several times or not on the genome. Copy number variations cover about 12 percent of the genome [51]. Copy number variations of genes play a role in genetic variability and epigenetics. It has also been observed that the CNV phenomenon also plays a role in spiking up the innate and adaptive immune systems, thus better enabling them to clear off infections.[52] Copy number variation also seems to have a role in the expression of genes altering the normal physiological functions of organisms. Such as the expression of the alpha-amylase-1 (AMY-1) gene, which is directly proportional to its dosage in an individual's genome. This discovery was seen in an extensive study involving various regions' populations. Their genome analysis was done, and it was seen that people of different regions had different gene dosages and copy numbers of the genome, depending on their lifestyle and food habits. Since the AMY-1 gene is responsible for the production of amylase enzyme, its crucial role in digestion can be observed in the life of individuals separated by distinct geographic locations. [55] Therefore, the gene (AMY-1) variation in populations around the globe leads many to speculate that the phenomenon may be involved in the evolutionary and adaptive change related to dietary variations.[53]

The CNV of a trinucleotide sequence $(CAG)_n$ is also seen to be the causative agent of Hungtinson's disease.[54] In another study, it was reported that the CNV of this sequence refers to the age of the onset of Hungtinson's disease.[55] The evidence mentioned above is probably enough to push the idea that CNVs play a role in genetic as well as infectious diseases by regulating the immune system. The severity of COVID-19, although dependent on many factors, is highly dependent on the cytokine response of the body. A cytokine storm is mainly caused by agents such as IL-2, IL-6 TNF-α, IL-10, etc.[56] Severe forms of cytokine response may also cause the body to go into ARDS and shock resulting in mortality of the patient. A recent article published [60] on the overall long-term effects of COVID-19 also analyzed potential risk factors, including age, gender, acute COVID-19 severity, obesity, and pre-existing allergic diseases. Still, none of these factors could have a strong association with the long-term effects of the disease. The article also points out the lack of race/ethnicity, socioeconomic, education, and employment-based data to extrapolate the analyzed data.

The antibody enhancement mechanism is a phenomenon by which viral entry and growth are enhanced. This increase is due to the presence of specific antibodies. This phenomenon seems to play a role in the severity of infection in people who have received vaccine therapy or are reinfected with the viral disease. This mechanism is also true for the more recent Omicron BA.1 and 20.A variants [54]. It is also important to mention that CNVs, like the above-mentioned biological phenomenon, are also variable and different among people of different geographical populations.[53] Hence, we would like to represent the hypothesis that the severity of the illness of the covid-19 and its mortality may have a direct or indirect relation to the copy number variation of genes associated with the immune system and cytokine response, just to name a few. Alternatively, the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 December 2022 doi:10.20944/preprints202212.0582.v1

CNVs of the genomic components' other involved mechanisms may be over or under-regulated, thus explaining the disparity of the COVID-19 severity and mortality among people of different geographical populations.

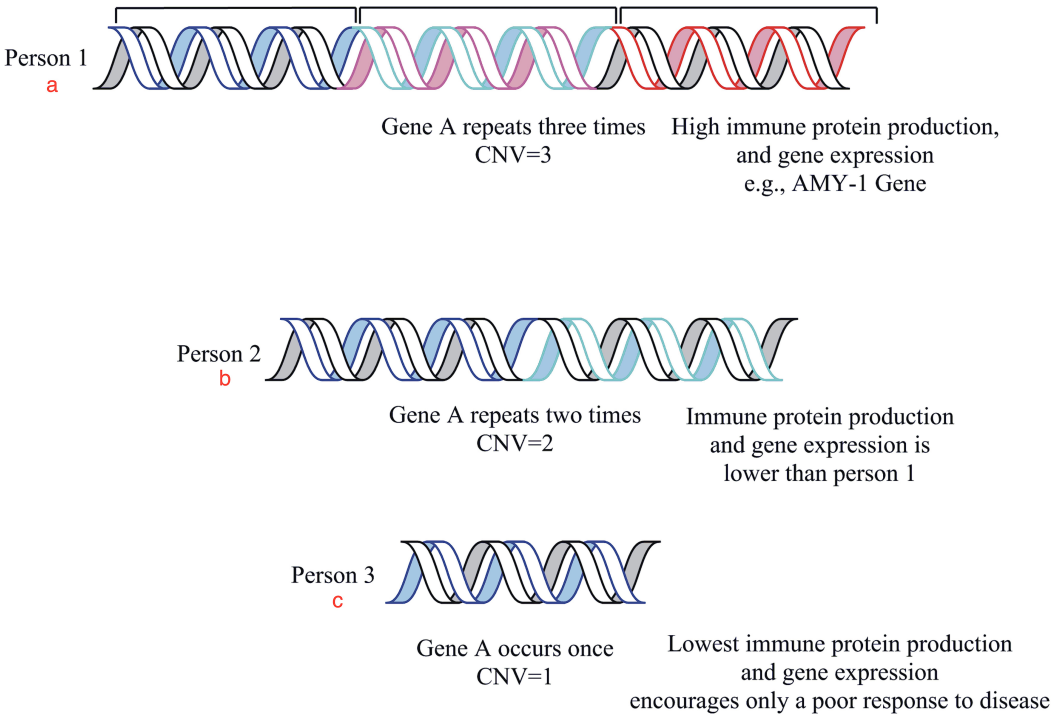

Fig. 3: This is a diagrammatic representation of the relation between gene dosage and gene expression in three different individuals. The diagram is an example of the copy number variation phenomenon showing us the variable copy numbers of a gene in different people. The protein production is observed as (a) high in person 1 (b) moderate in 2 as compared to (a and c) while (c) low in person 3 due to the variable number of gene copies.

6. Potential gene-based immunotherapy, human gene transfer based genetic engineering and gene drive technologies.

Genetic immunotherapy is a therapy that uses immune cells with modified genetics for an improved capability to fight infections or track, trace, and destroy cancer cells. Suppose the theories presented here are correct; certain ethnic population groups (with subsequent differential genetic components) possess improved protection against particular infections and diseases. In that case, we can replicate the same process in less fortunate gene groups using the techniques and principles of genetic immunotherapy, as this strategy could provide viable immediate treatment options in case of new emerging pandemics.[57]

Human gene therapy-based genetic research focuses on transferring a genetic component from one individual to another for therapeutic or beneficial purposes. Therefore if we can identify and clone genes that provide special immunity to some individuals in populations or ethnic groups, then using this technology will also allow us to give better immunological stability to other population groups.[58]

In that regard, clustered, regularly interspaced short palindromic repeats (CRISPR)-Cas-9-based gene-editing technologies are gaining acceptability in the field of research. We can also use this technology to enhance the survival rate of genetically predisposed individuals belonging to certain ethnic groups.[59]

 doi:10.20944/preprints202212.0582.v1

CONCLUSION

The analysis and hypotheses presented in this writing may be speculative in parts. Still, it presents a range of new experiments and analyses that should be done to look at the disease etiology and ecology from a holistic perspective. Such studies based on biological phenomena that we have ignored until now may, in fact, finally lead us to investigate the underlying reasons for the geographic disparity of COVID-19. Investigations into these phenomena may also help obtain a more rigorous understanding of previous, current, or future pandemics.

FUTURE WORK DIRECTIONS

The ideas presented in our paper are, of course, still theories at best, even with supporting evidence from other cases. Therefore, to prove them and get some benefit out of them would require various research projects, and the following steps will be essential to prove these hypotheses:

(1) Collection of strategically randomized samples of DNAs, i.e., setting up a broad spectrum of subjects from major demographic regions of the world. Important things to keep in mind while choosing subjects would be that they should be genetically diverse from each other and not just from different countries (for example, the Eastern Punjab region of Pakistan and the Western Punjab region of India have people with similar gene pools; therefore the results of this sample might not be so prominent). It is important to at least have subjects from major parts of Europe, North and South America, and Asia (South East Asian island nations and south Asian countries included)

(2) For the Virome and microbiome hypothesis, metagenomic analysis and classification and then a comparison of subject's microbiome with each other, finding the potentially helpful groups of bacteria, viruses and bacteriophages may allow us to go towards unleashing the therapeutic arenas of microbiome and virome, which include the modulation of a person's microbiome, phage therapy, pre-, and pro-biotic therapies, microbiome transplantation, etc.

(3) HERVs will require genomic analysis and genome sequences. If we can determine the region of HERV that is potentially boosting the immune system of a subject, we may be able to replicate that and so find its molecular basis and biological causative agents involved; thus allowing us to get a better understanding of the disease host interaction, the role of ERV in our genome all of which in turn may allow us to utilize possibly therapeutic benefits.

(4) To test the ENV hypothesis, an extensive genomic analysis (NextGen) would be required to identify potential genes and find their respective roles in epigenetics and immune modulation, allowing us to understand the mechanism's function better.

(5) A study on the population of Pakistan and Bangladesh for Hypothesis 3 can also be done to investigate whether such a case exists or not, it will require a lot of data analysis, bioinformatics, and big data comparative analysis, but if such a case is indeed found, then we may be able to exploit the same phenomenon to potentiate relative immunity against a future pathogen with the help of the region's previously encountered and less dangerous pathogens.

(6) Lastly, having all of this big data will allow researchers to use Machine Learning (ML) and various novel network search algorithms [61, 62] to predict the effect of future SARS-CoV* strains in a particular region or a geographical location based on their biological and immunological capabilities and not just the genetic backbone, which will give the scientific and medical communities a considerable advantage in refocusing their limited resources.

FUNDING DETAILS
This work is kindly supported by the start-up grant # OJQD2022005 funded by the Oujiang lab (Zhejiang lab for regenerative medicine, vision, and brain health), Wenzhou, Zhejiang province, 325000, China.

CONFLICT OF INTEREST
The authors report that there are no competing interests to declare.

REFERENCES
1. Coronavirus Disease (COVID-19) – World Health Organization Available online: https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (accessed on 19 March 2022).
2. Kim, D.; Lee, J.-Y.; Yang, J.-S.; Kim, J.W.; Kim, V.N.; Chang, H. The Architecture of SARS-CoV-2 Transcriptome. *Cell* **2020**, *181*, 914-921.e10, doi:10.1016/j.cell.2020.04.011.
3. Yuki, K.; Fujiogi, M.; Koutsogiannaki, S. COVID-19 Pathophysiology: A Review. *Clin. Immunol.* **2020**, *215*, 108427, doi:10.1016/j.clim.2020.108427.
4. Shoaib, N.; Noureen, N.; Munir, R.; Shah, F.A.; Ishtiaq, N.; Jamil, N.; Batool, R.; Khalid, M.; Khan, I.; Iqbal, N.; et al. COVID-19 Severity: Studying the Clinical and Demographic Risk Factors for Adverse Outcomes. *PLOS ONE* **2021**, *16*, e0255999, doi:10.1371/journal.pone.0255999.
5. Is Severe COVID-19 Over in Europe? - HealthManagement.Org Available online: https://healthmanagement.org/c/icu/issuearticle/is-the-severe-covid-19-over-in-europe (accessed on 20 March 2022).
6. Statement – Risk Remains High in Eastern Europe and Central Asia with Arrival of COVID-19 Omicron Variant Available online: https://www.euro.who.int/en/media-centre/sections/statements/2022/statement-risk-remains-high-in-eastern-europe-and-central-asia-with-arrival-of-covid-19-omicron-variant (accessed on 20 March 2022).
7. Konstantinoudis, G.; Cameletti, M.; Gómez-Rubio, V.; Gómez, I.L.; Pirani, M.; Baio, G.; Larrauri, A.; Riou, J.; Egger, M.; Vineis, P.; et al. Regional Excess Mortality during the 2020 COVID-19 Pandemic in Five European Countries. *Nat. Commun.* **2022**, *13*, 482, doi:10.1038/s41467-022-28157-3.
8. Population of Europe (2022) - Worldometer Available online: https://www.worldometers.info/world-population/europe-population/ (accessed on 20 March 2022).
9. Unleashing the Potential of a Young Pakistan | Human Development Reports Available online: https://hdr.undp.org/en/content/unleashing-potential-young-pakistan (accessed on 28 April 2022).
10. Sender, R.; Fuchs, S.; Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLOS Biol.* **2016**, *14*, e1002533, doi:10.1371/journal.pbio.1002533.
11. Raoult, D. Obesity Pandemics and the Modification of Digestive Bacterial Flora. *Eur. J. Clin. Microbiol. Infect. Dis.* **2008**, *27*, 631–634, doi:10.1007/s10096-008-0490-x.
12. Sekirov, I.; Russell, S.L.; Antunes, L.C.M.; Finlay, B.B. Gut Microbiota in Health and Disease. *Physiol. Rev.* **2010**, *90*, 859–904, doi:10.1152/physrev.00045.2009.

13. LeBlanc, J. g.; Laiño, J. e.; del Valle, M.J.; Vannini, V.; van Sinderen, D.; Taranto, M. p.; de Valdez, G.F.; de Giori, G.S.; Sesma, F. B-Group Vitamin Production by Lactic Acid Bacteria – Current Knowledge and Potential Applications. *J. Appl. Microbiol.* **2011**, *111*, 1297–1309, doi:10.1111/j.1365-2672.2011.05157.x.
14. Clarke, G.; Sandhu, K.V.; Griffin, B.T.; Dinan, T.G.; Cryan, J.F.; Hyland, N.P. Gut Reactions: Breaking Down Xenobiotic–Microbiome Interactions. *Pharmacol. Rev.* **2019**, *71*, 198–224, doi:10.1124/pr.118.015768.
15. Breitbart, M.; Hewson, I.; Felts, B.; Mahaffy, J.M.; Nulton, J.; Salamon, P.; Rohwer, F. Metagenomic Analyses of an Uncultured Viral Community from Human Feces. *J. Bacteriol.* **2003**, *185*, 6220–6223, doi:10.1128/JB.185.20.6220-6223.2003.
16. Young, J.C.; Chehoud, C.; Bittinger, K.; Bailey, A.; Diamond, J.M.; Cantu, E.; Haas, A.R.; Abbas, A.; Frye, L.; Christie, J.D.; et al. Viral Metagenomics Reveal Blooms of Anelloviruses in the Respiratory Tract of Lung Transplant Recipients. *Am. J. Transplant.* **2015**, *15*, 200–209, doi:10.1111/ajt.13031.
17. Łusiak-Szelachowska, M.; Weber-Dąbrowska, B.; Żaczek, M.; Borysowski, J.; Górski, A. The Presence of Bacteriophages in the Human Body: Good, Bad or Neutral? *Microorganisms* **2020**, *8*, 2012, doi:10.3390/microorganisms8122012.
18. Cash, H.L.; Whitham, C.V.; Behrendt, C.L.; Hooper, L.V. Symbiotic Bacteria Direct Expression of an Intestinal Bactericidal Lectin. *Science* **2006**, *313*, 1126–1130, doi:10.1126/science.1127119.
19. An Enteric Virus Can Replace the Beneficial Function of Commensal Bacteria - PMC Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4257755/ (accessed on 20 April 2022).
20. Rampelli, S.; Turroni, S.; Schnorr, S.L.; Soverini, M.; Quercia, S.; Barone, M.; Castagnetti, A.; Biagi, E.; Gallinella, G.; Brigidi, P.; et al. Characterization of the Human DNA Gut Virome across Populations with Different Subsistence Strategies and Geographical Origin. *Environ. Microbiol.* **2017**, *19*, 4728–4735, doi:10.1111/1462-2920.13938.
21. Zuo, T.; Sun, Y.; Wan, Y.; Yeoh, Y.K.; Zhang, F.; Cheung, C.P.; Chen, N.; Luo, J.; Wang, W.; Sung, J.J.Y.; et al. Human-Gut-DNA Virome Variations across Geography, Ethnicity, and Urbanization. *Cell Host Microbe* **2020**, *28*, 741-751.e4, doi:10.1016/j.chom.2020.08.005.
22. McPherson, J.D.; Marra, M.; Hillier, L.; Waterston, R.H.; Chinwalla, A.; Wallis, J.; Sekhon, M.; Wylie, K.; Mardis, E.R.; Wilson, R.K.; et al. A Physical Map of the Human Genome. *Nature* **2001**, *409*, 934–941, doi:10.1038/35057157.
23. Wang, M.; Qiu, Y.; Liu, H.; Liang, B.; Fan, B.; Zhou, X.; Liu, D. Transcription Profile of Human Endogenous Retroviruses in Response to Dengue Virus Serotype 2 Infection. *Virology* **2020**, *544*, 21–30, doi:10.1016/j.virol.2020.01.014.
24. Vargiu, L.; Rodriguez-Tomé, P.; Sperber, G.O.; Cadeddu, M.; Grandi, N.; Blikstad, V.; Tramontano, E.; Blomberg, J. Classification and Characterization of Human Endogenous Retroviruses; Mosaic Forms Are Common. *Retrovirology* **2016**, *13*, 7, doi:10.1186/s12977-015-0232-y.
25. Esposito, F.; Tramontano, E. Past and Future. Current Drugs Targeting HIV-1 Integrase and Reverse Transcriptase-Associated Ribonuclease H Activity: Single and Dual Active Site Inhibitors. *Antivir. Chem. Chemother.* **2014**, *23*, 129–144, doi:10.3851/IMP2690.

26. Bannert, N.; Kurth, R. The Evolutionary Dynamics of Human Endogenous Retroviral Families. *Annu. Rev. Genomics Hum. Genet.* **2006**, *7*, 149–173, doi:10.1146/annurev.genom.7.080505.115700.
27. Srinivasachar Badarinarayan, S.; Sauter, D. Switching Sides: How Endogenous Retroviruses Protect Us from Viral Infections. *J. Virol. 95*, e02299-20, doi:10.1128/JVI.02299-20.
28. Yoneyama, M.; Kikuchi, M.; Natsukawa, T.; Shinobu, N.; Imaizumi, T.; Miyagishi, M.; Taira, K.; Akira, S.; Fujita, T. The RNA Helicase RIG-I Has an Essential Function in Double-Stranded RNA-Induced Innate Antiviral Responses. *Nat. Immunol.* **2004**, *5*, 730–737, doi:10.1038/ni1087.
29. Zhou, B.; Qi, F.; Wu, F.; Nie, H.; Song, Y.; Shao, L.; Han, J.; Wu, Z.; Saiyin, H.; Wei, G.; et al. Endogenous Retrovirus-Derived Long Noncoding RNA Enhances Innate Immune Responses via Derepressing RELA Expression. *mBio* **2019**, *10*, e00937-19, doi:10.1128/mBio.00937-19.
30. NeST, a Long Noncoding RNA, Controls Microbial Susceptibility and Epigenetic Activation of the Ifng Locus - PMC Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3577098/ (accessed on 20 April 2022).
31. Human Endogenous Retrovirus Type W Envelope Protein Inhibits Oligodendroglial Precursor Cell Differentiation - PubMed Available online: https://pubmed.ncbi.nlm.nih.gov/23836485/ (accessed on 20 March 2022).
32. Vogt, P.K.; Ishizaki, R. Patterns of Viral Interference in the Avian Leukosis and Sarcoma Complex. *Virology* **1966**, *30*, 368–374, doi:10.1016/0042-6822(66)90115-2.
33. Role of Angiotensin-Converting Enzyme 2 (ACE2) in COVID-19 | Critical Care | Full Text Available online: https://ccforum.biomedcentral.com/articles/10.1186/s13054-020-03120-0 (accessed on 28 May 2022).
34. A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin - PubMed Available online: https://pubmed.ncbi.nlm.nih.gov/32015507/ (accessed on 28 May 2022).
35. Li, W.; Moore, M.J.; Vasilieva, N.; Sui, J.; Wong, S.K.; Berne, M.A.; Somasundaran, M.; Sullivan, J.L.; Luzuriaga, K.; Greenough, T.C.; et al. Angiotensin-Converting Enzyme 2 Is a Functional Receptor for the SARS Coronavirus. *Nature* **2003**, *426*, 450–454, doi:10.1038/nature02145.
36. Abbasi, A.Z.; Kiyani, D.A.; Hamid, S.M.; Saalim, M.; Fahim, A.; Jalal, N. Spiking Dependence of SARS-CoV-2 Pathogenicity on TMPRSS2. *J. Med. Virol.* **2021**, *93*, 4205–4218, doi:10.1002/jmv.26911.
37. Steck, F.T.; Rubin, H. The Mechanism of Interference between an Avian Leukosis Virus and Rous Sarcoma Virus. II. Early Steps of Infection by RSV of Cells under Conditions of Interference. *Virology* **1966**, *29*, 642–653, doi:10.1016/0042-6822(66)90288-1.
38. Taylor, G.M.; Gao, Y.; Sanders, D.A. Fv-4: Identification of the Defect in Env and the Mechanism of Resistance to Ecotropic Murine Leukemia Virus. *J. Virol.* **2001**, *75*, 11244–11248, doi:10.1128/JVI.75.22.11244-11248.2001.
39. Brinzevich, D.; Young, G.R.; Sebra, R.; Ayllon, J.; Maio, S.M.; Deikus, G.; Chen, B.K.; Fernandez-Sesma, A.; Simon, V.; Mulder, L.C.F. HIV-1 Interacts with Human Endogenous Retrovirus K (HML-2) Envelopes Derived from Human Primary Lymphocytes. *J. Virol.* **2014**, *88*, 6213–6223, doi:10.1128/JVI.00669-14.
40. Boso, G.; Buckler-White, A.; Kozak, C.A. Ancient Evolutionary Origin and Positive Selection of the Retroviral Restriction Factor Fv1 in Muroid Rodents. *J. Virol.* **2018**, *92*, e00850-18, doi:10.1128/JVI.00850-18.

41. Yan, Y.; Buckler-White, A.; Wollenberg, K.; Kozak, C.A. Origin, Antiviral Function and Evidence for Positive Selection of the Gammaretrovirus Restriction Gene Fv1 in the Genus Mus. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 3259–3263, doi:10.1073/pnas.0900181106.
42. Sun, C.; Skaletsky, H.; Rozen, S.; Gromoll, J.; Nieschlag, E.; Oates, R.; Page, D.C. Deletion of Azoospermia Factor a (AZFa) Region of Human Y Chromosome Caused by Recombination between HERV15 Proviruses. *Hum. Mol. Genet.* **2000**, *9*, 2291–2296, doi:10.1093/oxfordjournals.hmg.a018920.
43. Campbell, I.M.; Gambin, T.; Dittwald, P.; Beck, C.R.; Shuvarikov, A.; Hixson, P.; Patel, A.; Gambin, A.; Shaw, C.A.; Rosenfeld, J.A.; et al. Human Endogenous Retroviral Elements Promote Genome Instability via Non-Allelic Homologous Recombination. *BMC Biol.* **2014**, *12*, 74, doi:10.1186/s12915-014-0074-4.
44. Feschotte, C.; Gilbert, C. Endogenous Viruses: Insights into Viral Evolution and Impact on Host Biology. *Nat. Rev. Genet.* **2012**, *13*, 283–296, doi:10.1038/nrg3199.
45. Hughes, J.F.; Coffin, J.M. Evidence for Genomic Rearrangements Mediated by Human Endogenous Retroviruses during Primate Evolution. *Nat. Genet.* **2001**, *29*, 487–489, doi:10.1038/ng775.
46. Doxiadis, G.G.M.; de Groot, N.; Bontrop, R.E. Impact of Endogenous Intronic Retroviruses on Major Histocompatibility Complex Class II Diversity and Stability. *J. Virol.* **2008**, *82*, 6667–6677, doi:10.1128/JVI.00097-08.
47. Kulski, J.K.; Gaudieri, S.; Martin, A.; Dawkins, R.L. Coevolution of PERB11 (MIC) and HLA Class I Genes with HERV-16 and Retroelements by Extended Genomic Duplication. *J. Mol. Evol.* **1999**, *49*, 84–97, doi:10.1007/pl00006537.
48. Ito, J.; Sugimoto, R.; Nakaoka, H.; Yamada, S.; Kimura, T.; Hayano, T.; Inoue, I. Systematic Identification and Characterization of Regulatory Elements Derived from Human Endogenous Retroviruses. *PLoS Genet.* **2017**, *13*, e1006883, doi:10.1371/journal.pgen.1006883.
49. Sundaram, V.; Cheng, Y.; Ma, Z.; Li, D.; Xing, X.; Edge, P.; Snyder, M.P.; Wang, T. Widespread Contribution of Transposable Elements to the Innovation of Gene Regulatory Networks. *Genome Res.* **2014**, *24*, 1963–1976, doi:10.1101/gr.168872.113.
50. Wildschutte, J.H.; Williams, Z.H.; Montesion, M.; Subramanian, R.P.; Kidd, J.M.; Coffin, J.M. Discovery of Unfixed Endogenous Retrovirus Insertions in Diverse Human Populations. *Proc. Natl. Acad. Sci.* **2016**, *113*, E2326–E2334, doi:10.1073/pnas.1602336113.
51. McCarroll, S.A.; Altshuler, D.M. Copy-Number Variation and Association Studies of Human Disease. *Nat. Genet.* **2007**, *39*, S37–S42, doi:10.1038/ng2080.
52. Machado, L.R.; Ottolini, B. An Evolutionary History of Defensins: A Role for Copy Number Variation in Maximizing Host Innate and Adaptive Immune Responses. *Front. Immunol.* **2015**, *6*.
53. Perry, G.H.; Dominy, N.J.; Claw, K.G.; Lee, A.S.; Fiegler, H.; Redon, R.; Werner, J.; Villanea, F.A.; Mountain, J.L.; Misra, R.; et al. Diet and the Evolution of Human Amylase Gene Copy Number Variation. *Nat. Genet.* **2007**, *39*, 1256–1260, doi:10.1038/ng2123.
54. Macdonald, M. A Novel Gene Containing a Trinucleotide Repeat That Is Expanded and Unstable on Huntington's Disease Chromosomes. *Cell* **1993**, *72*, 971–983, doi:10.1016/0092-8674(93)90585-E.
55. Myers, R.H. Huntington's Disease Genetics. *NeuroRx* **2004**, *1*, 255–262.

56. Ragab, D.; Salah Eldin, H.; Taeimah, M.; Khattab, R.; Salem, R. The COVID-19 Cytokine Storm; What We Know So Far. *Front. Immunol.* **2020**, *11*.
57. Genetic Engineering of T Cells for Immunotherapy | Nature Reviews Genetics Available online: https://www.nature.com/articles/s41576-021-00329-9 (accessed on 12 December 2022).
58. mdabrow1 Gene Transfer Research Available online: https://www.hopkinsmedicine.org/institutional_review_board/guidelines_policies/guidelines/gene_transfer.html (accessed on 12 December 2022).
59. Bier, E. Gene Drives Gaining Speed. *Nat. Rev. Genet.* **2022**, *23*, 5–22, doi:10.1038/s41576-021-00386-0.
60. Pillay J, Rahman S, Guitard S, Wingert A & Hartling L. Risk factors and preventive interventions for post Covid-19 condition: systematic review, Emerging Microbes & Infections. 2022; DOI: 10.1080/22221751.2022.2140612.
61. Abu-Khzam, F. N., Barr, J. R., Fakhereldine, A., and Shaw, P. (2021) "A Greedy Heuristic for Cluster Editing with Vertex Splitting," 2021 4th International Conference on Artificial Intelligence for Industries (AI4I), pp. 38-41, doi: 10.1109/AI4I51902.2021.00017.
62. Grady, Stephen K., et al. (2022) "Domination based classification algorithms for the controllability analysis of biological interaction networks." Scientific Reports 12.1, pp. 1-10.