UTILIZING A GENOMIC DATABASE FOR IDENTIFYING GENOMIC VARIATION FOR SARS-CoV-2 RECEPTORS

by Universitas Ahmad Dahlan Yogyakarta 30

Submission date: 11-Jan-2024 01:02PM (UTC+0700)
Submission ID: 2269244554
File name: 7.pdf (918.12K)
Word count: 5253
Character count: 28924

Medical Sains : Jurnal Ilmiah Kefarmasian
Vol. 8 No. 2, April - June 2023
https://ojs.stfmuhammadiyahcirebon.ac.id/index.php/iojs

UTILIZING A GENOMIC DATABASE FOR IDENTIFYING GENOMIC VARIATION FOR SARS-CoV-2 RECEPTORS

Melodia Rezadhini[1], Lalu Muhammad Irham[2*], Anisa Nova Puspitaningrum[2], Arief Rahman Afief[2], Wirawan Adikusuma[3], Dyah Aryani Perwitasari[2]

[1]*Agricultural Microbiology, Faculty of Agriculture, Universitas Gadjah Mada Bulaksumur, Yogyakarta, Indonesia*

[2]*Faculty of Pharmacy, Universitas Ahmad Dahlan, Indonesia*
Jl. Prof. Dr. Soepomo, S.H, Warungboto, Umbulharjo, Yogyakarta, Indonesia

[3]*Faculty of Medical Science, Universitas Muhammadiyah Mataram, Indonesia*
Jl. KH. Ahmad Dahlan No.1, Pagesangan, Mataram, Nusa Tenggara Barat, Indonesia

Email Corresponding: lalu.irham@pharm.uad.ac.id

Submitted: 1 April 2023 ***Revised: 6 May 2023*** ***Accepted: 6 May 2023***

ABSTRACT

Coronavirus disease 2019 (COVID-19) is a severe acute respiratory syndrome caused by the coronavirus SARS-CoV-2. This disease progresses through different stages, and its common symptoms include fever, malaise, dry cough, dyspnea, and pneumonia. Since its discovery in Wuhan, Hubei, China, in December 2019, SARS-CoV-2 transmission has rapidly spread. Unfortunately, the number of cases is increasing beyond expectations, making the spread of confirmed COVID-19 cases easier. To detect specific symptoms in a large population, the relationship between gene variants and patients with COVID-19 can be determined. The data of COVID-19 variations can be obtained from the Ensembl Genome Browser. Herein, the gene associated with 19 variants was selected according to the Haploreg 4.1 version. Additionally, protein expression of missense genes variants using the GTEx portal to determine 2 selected variants: rs200553089 encode the *TLR7* gene and rs1061622 encode the *TNFRSF1B* gene, respectively. The highest protein expression of the *TLR7* gene is found in lymphocyte cells, while *TNFRSF1B* shows the highest expression in the whole blood system. According to the data obtained from the Ensembl Genome Browser, the two most prevalent populations for SNP rs19085059, which is associated with the *GPHN* gene, are in Africa and East Asia, while SNP rs74956615, linked to the *RAVER1* gene, is most prevalent in East Asia. The allele frequency for the *TLR7* gene cannot be found in five countries, except in America.

Keywords: COVID-19, SARS-CoV-2, Single nucleotide polymorphism, Variants.

INTRODUCTION

Coronavirus disease 2019 (COVID-19) is an illness that infects the human body. Its first discovery was in Wuhan, Hubei, China, in December 2019 (Singhal, 2020), and it has since spread worldwide. The rapid spread of the virus is caused by its ease of transmission from animals to humans and between humans. Symptoms of this disease include fever, malaise, dry cough, dyspnea, and a diagnosis of viral pneumonia (Liu *et al*., 2020). Ultimately, the virus was identified through whole-genome sequencing as a novel coronavirus, the seventh member of the coronavirus family that can infect humans (Liu *et al*., 2020). Subsequently, the World Health Organization (WHO) announced that the virus is the cause of infectious diseases, namely COVID-19.

The transmission of SARS-CoV-2 has spread rapidly and caused a sudden increase in the incidence of the disease in a population beyond normal expectations. The number of cases in population density levels has reached 22 people/km2 (Sy *et al*., 2021). Population density is an important factor in the spread of COVID-19 (Ganasegeran *et al*., 2021), as a high number of people lead to easy infection spread and an increase in disease incidence. As of February 6,

2020, there were 565 confirmed cases of COVID-19, with 25 deaths reported in 25 countries (Wu *et al.*, 2020). As of August 21 - September 3, 2020, cases increased in 50 countries, and the highest number of cases was found in the United States (Noh & Danuser, 2021). Regarding a gender comparison, a higher severity of the disease was found in men than in women. The mortality rate in men infected with COVID-19 was 2.4 times higher than in women (Jin *et al.*, 2020).

The detection of specific symptoms and the level of spread in the highest population of COVID-19 can be attempted by examining Deoxyribonucleic acid (DNA). These gene variations are linked to the progression and pathogenesis of a disease. The search for these gene variations is obtained by searching for Single Nucleotide Polymorphism (SNP) in positive SARS-CoV-2 samples as genetic markers in predicting the spread and transmission of COVID-19 (Harper *et al.*, 2021). Collecting gene variations or SNPs obtained from the Ensembl Genome Browser is very useful for identifying SNP types that play a role in COVID-19. This study used various databases to determine the highest gene expression of *TLR7* and *TNFRSF1B* in human tissue and the highest population spread of specific SNP variants in various regions around the world (Huang *et al.*, 2014). These functional applications in the future can be applied to the handling and risk of SARS-CoV-2 infection spread to COVID-19 in multiple continents.

METHODS

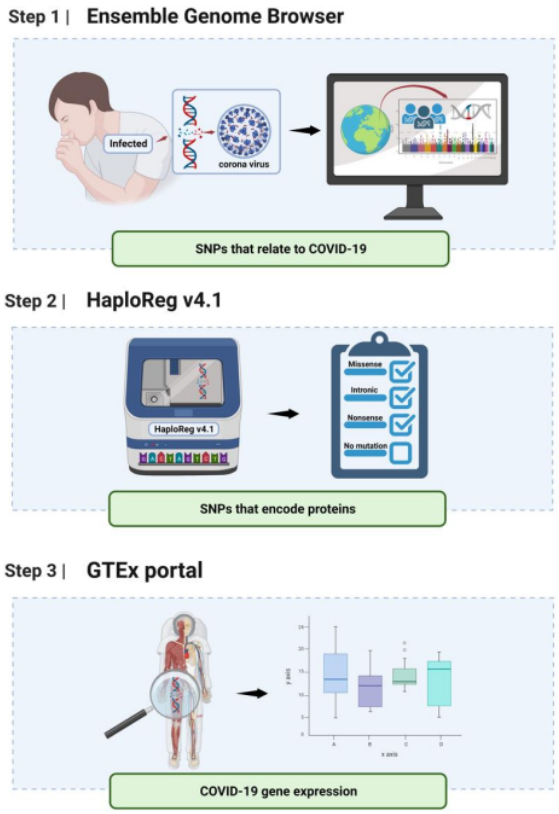

Figure 1. Genomic Database Workflow for Identifying Potential SNP Biomarkers for COVID-19 [Copyright Licence; XI24ODRG2W]

Variants of SNP in COVID-19 can be identified by searching SNP databases through the Ensembl Genome Browser (https://www.ensembl.org/index.html). This search resulted in data on 18 different SNP variants. Access to other aspects of each of the 18 variants is necessary, such as position, allele or nitrogen base changes, and the number of allele or nitrogen base changes in various populations around the world, including Europe, Africa, America, East Asia, Southeast Asia, Han China, and Japan. This information can also be accessed from the same database, the Ensembl Genome Browser. To determine the gene and its location for each COVID-19 SNP variant, a search was conducted using the HaploReg v4.1 database (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) to identify variants encoded the gene. The gene and location data obtained from the 18 COVID-19 variants were used to identify gene expression in missense regions, which affect changes in single amino acids (Petrosino *et al.*, 2021). Among the 18 variants, two variants with genes that experienced missense were identified, namely rs200553089 encoded the *TLR7* gene and rs1061622 encoded the *TNFRSF1B* gene. Furthermore, the *TLR7* and *TNFRSF1B* gene expressions were identified in various human tissue types by searching the GTEx portal database (http://www.gtexportal.org/home/).

RESULT AND DISCUSSION

We retrieved 19 SNPs from the search results of COVID-19 in the Ensembl Genome Browser to determine the allele frequencies of SNP variants in different populations. Among these 19 SNPs, two missense variants were chosen for further analysis: rs200553089 with the *TLR7* gene and rs1061622 with the *TNFRSF1B* gene. Subsequently, we investigated the expression of the *TLR7* and *TNFRSF1B* genes in the human body.

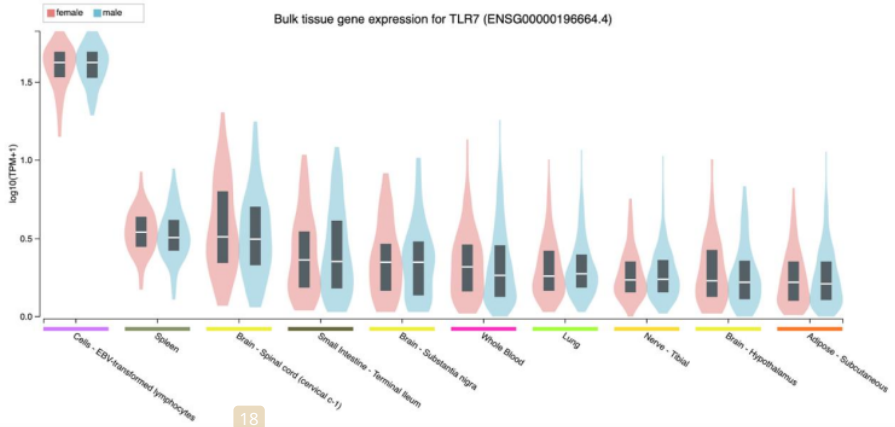

Figure 2. Expression of TLR7 Gene in Human Body Tissues.

The expression of the *TLR7* gene in human tissues is highest in lymphocytes, spleen, bone marrow, small intestine or ileum, midbrain, the entire circulatory system, lungs, the nervous system, hypothalamus, and subcutaneous adipose tissue. This high expression corresponds to its role in recognizing and responding to ssRNA viruses, including SARS-CoV-2, which infects human cells. Based on gender, *TLR7* gene expression is higher in males than in females, which is consistent with the higher number of SARS-CoV-2 infections in males compared to females.

Toll-Like Receptor 7 (TLR7) is a Pattern Recognition Receptor (PRR) located on the X chromosome locus Xp22.2, responsible for recognizing ssRNA viruses like SARS-CoV-2 (Celhar *et al.*, 2019). The missense gene *TLR7* plays a crucial role in immunity and the regulation of the interferon pathway. *TLR7* is expressed on the endosome membrane of plasmacytoid dendritic cells (pDCs), the main producers of type I interferon (IFN-I) (Asano

et al., 2021), as well as on B cells in monocytes/macrophages (Sindhu *et al.,* 2015). In macrophages, *TLR7* can stimulate the production of pro-inflammatory cytokines. *TLR7* also recognizes ssRNA from viruses and bacteria in humans and responds to RNA-associated autoantigens, making it a potential factor in recognizing coronaviruses.

Regarding ssRNA recognition, *TLR7* activates the MyD88-dependent pathway, which triggers the activation of nuclear factor kappa Beta (NFκB) and the type 1 interferon (IFN-I) pathways in pDCs (Spiering & Vries, 2021). Additionally, the *TLR7* pathway induces a response of type 2 interferon (IFN-2) in COVID-19 (van der Made *et al.,* 2020). This process is associated with the CXorf21 gene product, which is a gene on the X chromosome that encodes the *TLR* associated with the endolysosomal *SLC15A4* (TASL) protein. This process can also stimulate the migration of interferon regulatory factor 5 (*IRF5*) and interferon regulatory factor 7 (*IRF7*) (Spiering & Vries, 2021).

TLR7 gene expression has been observed in various human tissues, including the kidney (Asano *et al.,* 2021), adipose tissue (Sindhu *et al.,* 2015), peripheral blood leukocytes (PBL) with the highest expression, spleen tissue, intestine, brain, liver, ovary, prostate, lung, small intestine, pancreas, placenta, testis, and thymus (Zarember & Godowski, 2002).

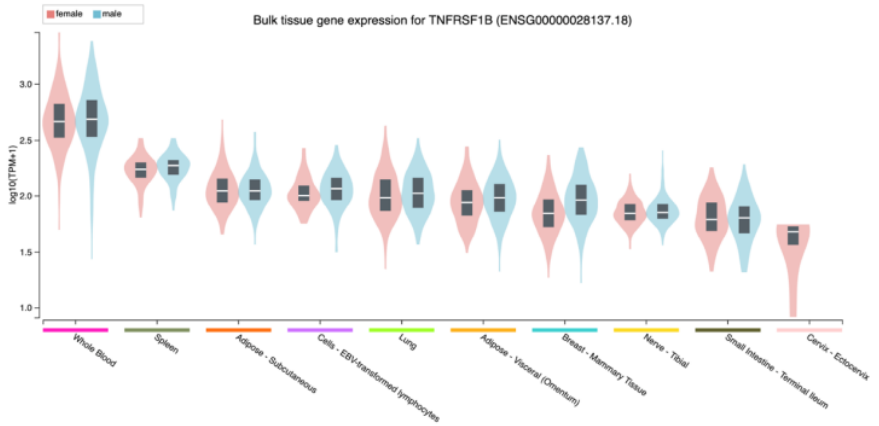

Figure 3. Expression of the TNFRSF1B Gene in Human Body Tissues.

The *TNFRSF1B* gene shows its highest expression in various human body tissues, including the entire circulatory system, spleen, subcutaneous adipose tissue, lymphocytes, lungs, visceral adipose tissue, breasts, nerves, ileum of the small intestine, and cervix. This high expression is closely associated with SARS-CoV-2 infection within the entire circulatory system. Moreover, the expression of the *TNFRSF1B* gene is found to be higher in men compared to women, which aligns with the higher incidence of COVID-19 observed in men.

The *TNFRSF1B* gene is a receptor gene that interacts with the TNF gene variant, which has been identified in COVID-19 patient plasma levels (Fricke-Galindo *et al.,* 2022). Located on chromosome 1 at the locus 1p36.22, *TNFRSF1B* exhibits a high affinity for the TNFSF2/TNF-alpha receptor (NCBI, 2022). In COVID-19 patients, the gene demonstrates high expression in TNF receptor-associated factor 3 (TRAF3), which plays a role in controlling the production of type I interferon (IFN-I) (Zhou *et al.,* 2021). *TNFRSF1B* interacts with TNF-receptor 1, forming a hetero-complex with two anti-apoptotic proteins, c-IAP1 and c-IAP2, which activate ubiquitin ligase (NCBI, 2022). Protein expression of TNFRSF1B is detected in various tissues, such as the appendix, spleen (NCBI, 2022), and blood cells (Fricke-Galindo *et al.,* 2022).

Table I. SNP Variants, Genes, Locations, Alleles, and Population Allele Frequencies Found.

SNP	Position	Gene	Location	Allele		Allele Frequency				
				Ref	Eff*	AFR	AMR	EAS	EUR	SAS
rs200553089	Chromosome x : 12906010	*TLR7*	missense	G	T	Not available				
rs71325088	Chromosome 3 : 45862952	*LZTFL1*	none	T	C	T: 0.996	T: 0.957	T: 0.995	T: 0.918	T: 0.704
rs13050728	Chromosome 21: 34615210	*IFNAR2*	intronic	T	C	T: 0.194	T: 0.451	T: 0.576	T: 0.333	T: 0.519
rs361525	Chromosome 6 : 31543101	*TNF*	none	A	G	A: 0.038	A: 0.082	A: 0.031	A: 0.064	A: 0.105
rs12564811	not found	not found		G	A	G: 0.818	G: 0.95	G: 0.823	G: 0.923	G: 0.915
rs73064425	Chromosome 3 : 45901089	*LZTFL1*	none	C	T	C: 0.995	C: 0.957	C: 0.995	C: 0.92	C: 0.706
rs6489867	Chromosome 12: 113363550	*OAS1*	none	C	T	C: 0.269	C: 0.264	C: 0.253	C: 0.367	C: 0.286
rs2109069	Chromosome 19 : 4719431	*DPP9*	intronic	G	A	G: 0.804	G: 0.781	G: 0.86	G: 0.679	G: 0.814
rs3397	Chromosome 1 : 12267292	*TNFRSF 1B*	3'-UTR	C	T	C: 0.835	C: 0.308	C: 0.711	C: 0.379	C: 0.555
rs1061622	Chromosome 1 : 12252955	*TNFRSF 1B*	missense	T	G	T: 0.813	T: 0.873	T: 0.85	T: 0.782	T: 0.724
rs11385942	Chromosome 3 : 45876459	*LZTFL1*	Intronic	AA	AAA	AA: 0.947	AA: 0.954	AA: 0.995	AA: 0.919	AA: 0.704
rs767455	Chromosome 12 : 6450945	*TNFRSF 1A*	synonymous	T	C	T: 0.637	T: 0.699	T: 0.878	T: 0.571	T: 0.728
rs143334143	Chromosome 6 : 31121426	*CCHCR 1*	Intronic	A	G	A: 0.057	A: 0.137	A: 0.035	A: 0.112	A: 0.077
rs190850598	Chromosome 14 : 67112421	*GPHN*	Intronic	A	G	A: 1	A: 0.983	A: 1	A: 0.983	A: 0.996
rs1800629	Chromosome 6 : 31543031	*TNF*	none	A	G	A: 0.12	A: 0.069	A: 0.059	A: 0.134	A: 0.053
rs657152	Chromosome 9 : 136139265	*ABO*	Intronic	A	C/T	A: 0.436	A: 0.305	A: 0.363		A: 0.461

rs1800693	Chromosome 12 : 6440009	*TNFRSF1A*	Intronic	T	C	T: 0.639	T: 0.719	T: 0.877	T: 0.575	T: 0.717
rs74956615	Chromosome 19 : 10427721	*RAVER1*	3'-UTR	T	A/C	T: 0.998	T: 0.974	T: 1	T: 0.97	T: 0.98
rs11085727	Chromosome 19 : 10466123	*TYK2*	Intronic	C	T	C: 0.906	C: 0.742	C: 0.431	C: 0.73	C: 0.7

Genetic variants associated with COVID-19 in the *TLR7* and *TNFRSF1B* genes were identified using available databases. Genotype-Tissue Expression (GTEx) database is used to identify genotypes and expression in human and mammalian tissues. GTEx has various tools for searching gene, variant, and tissue-related databases with molecular genomic and phenotypic datasets (Stanfill & Cao, 2021). In addition, the Ensembl Genome Browser is a single-scale genomic data entry with a graphical user interface that involves OMIM, dbSNP, and the NHGRI GWAS catalog (Spudich & Fernandez-Suarez, 2012). The Ensembl Genome Browser also provides comprehensive human genome annotation data sources with biomedical analysis and their relationships with human genetic diseases. Genome annotation, analysis, and dissemination are also provided on the Ensembl Genome Browser with reference sequences, gene models, transcription data, polymorphisms, and comparative analysis (Kersey *et al.,* 2012). The HaploReg database identifies Genome-Wide Association Study (GWAS) associations with putative haplotype relationships, cell type predictions, target gene predictions with systematic comparison, epigenomic and regulatory annotation, functional non-coding sequence prediction, and epigenomic information identification (Ward & Kellis, 2012). HaploReg can also be used to identify SNP associations with diseases based on expression quantitative trait locus (eQTL) with specific network and target genes from the GTEx portal (Ward & Kellis, 2016).

By identifying SNP variants in COVID-19 through Ensembl Genome Browser database searching, 19 SNP variants were found. The 19 SNP variants are rs200553089, rs71325088, rs13050728, rs361525, rs12564811, rs73064425, rs6489867, rs2109069, rs3397, rs1061622, rs11385942, rs767455, rs143334143, rs190850598, rs1800629, rs657152, rs1800693, rs74956615, and rs11085727. From the 19 SNP variants, genetic variants were identified using HaploReg, resulting in SNP variants encoding missense mutations in the *TLR7* and *TNFRSF1B* genes.

Among the 19 SNP variants identified using HaploReg, rs200553089 encodes the *TLR7* gene with a missense mutation. The *TLR7* gene encodes a toll-like receptor that can recognize SARS-CoV-2 RNA and trigger antiviral responses (Deng *et al.,* 2021). The missense SNP variant rs200553089 identifies a loss-of-function (LOF) in the TLR7 gene related to immune deficiency responses to type I and II interferons (Deng *et al.,* 2021). Moreover, this SNP variant identifies a high consideration of COVID-19 cases in males compared to females, with 2% of cases (Deng *et al.,* 2021). In previous sequencing studies, four males under 35 were infected with SARS-CoV-2 and contracted COVID-19 with the rs200553089 SNP variant and the *TLR7* gene encoding missense mutations (van der Made *et al.,* 2020).

Another missense-encoding variant is rs1061622 in the *TNFRSF1B* gene. In a previous study, blood plasma from patients with COVID-19 was identified, and the rs1061622 variant was found to encode the *TNFRSR1B* gene. It was associated with chronic diseases, such as cystic fibrosis and lung disease (Fricke-Galindo *et al.,* 2022). In addition, this SNP variant was also found to be associated with back and chest pain (Xing-rong *et al.,* 2018). In a previous study, it was also found that rs1061622 affects the level and function of *TNFR* p75

and has an effect on the response to anti-TNF treatment (Xing-rong *et al.,* 2018). The SNP variant rs1061622 is associated with sTNFR2 in the periphery (Pillai *et al.,* 2021).

Genetic variants associated with COVID-19 in the *TLR7* and *TNFRSF1B* genes were identified using available databases. The Genotype-Tissue Expression (GTEx) database is used to identify genotypes and expression in human and mammalian tissues. GTEx offers various tools for searching gene, variant, and tissue-related databases with molecular genomic and phenotypic datasets (Stanfill & Cao, 2021). Additionally, the Ensembl Genome Browser is a single-scale genomic data entry with a graphical user interface that incorporates OMIM, dbSNP, and the NHGRI GWAS catalog (Spudich & Fernandez-Suarez, 2012). The Ensembl Genome Browser provides comprehensive human genome annotation data from various sources, facilitating biomedical analysis and understanding of their relationships with human genetic diseases. It includes reference sequences, gene models, transcription data, polymorphisms, and comparative analysis (Kersey *et al.,* 2012). The HaploReg database identifies Genome-Wide Association Study (GWAS) associations, predicts haplotype relationships, cell types, and target genes, and provides epigenomic and regulatory annotations, functional non-coding sequence predictions, and epigenomic information (Ward & Kellis, 2016). HaploReg can also be used to identify SNP associations with diseases based on expression quantitative trait locus (eQTL) using specific network and target genes from the GTEx portal (Ward & Kellis, 2012).

By searching the Ensembl Genome Browser database, 19 SNP variants associated with COVID-19 were found. These variants include rs200553089, rs71325088, rs13050728, rs361525, rs12564811, rs73064425, rs6489867, rs2109069, rs3397, rs1061622, rs11385942, rs767455, rs143334143, rs190850598, rs1800629, rs657152, rs1800693, rs74956615, and rs11085727. Using HaploReg, genetic variants were identified among these 19 SNP variants, resulting in SNP variants that encode missense mutations in the *TLR7* and *TNFRSF1B* genes.

Among the 19 SNP variants identified using HaploReg, rs200553089 encodes the TLR7 gene with a missense mutation. The TLR7 gene encodes a toll-like receptor that recognizes SARS-CoV-2 RNA and triggers antiviral responses (Deng *et al.,* 2021). The missense SNP variant rs200553089 identifies a loss-of-function (LOF) mutation in the TLR7 gene, which is associated with immune deficiency responses to type I and II interferons (Deng *et al.,* 2021). Moreover, this SNP variant indicates a higher risk of COVID-19 cases in males compared to females, with a prevalence of 2% (Deng *et al.,* 2021). Previous sequencing studies have found the rs200553089 SNP variant and TLR7 gene encoding missense mutations in four males under 35 who were infected with SARS-CoV-2 and developed COVID-19 (van der Made *et al.,* 2020).

Another missense-encoding SNP variant is rs1061622 in the *TNFRSF1B* gene. In a previous study, the rs1061622 SNP variant was identified in the blood plasma of patients with COVID-19 and was found to be associated with the *TNFRSF1B* gene (Fricke-Galindo *et al.,* 2022). This SNP variant has also been linked to chronic diseases such as cystic fibrosis and lung disease (Fricke-Galindo *et al.,* 2022). Additionally, rs1061622 has been associated with back and chest pain in previous research (Xing-rong *et al.,* 2018). Another study revealed that rs1061622 affects the level and function of *TNFR* p75 and has an impact on the response to anti-TNF treatment (Xing-rong *et al.,* 2018). Furthermore, the SNP variant rs1061622 is associated with *TNFR2* in the periphery, as demonstrated in a study by (Pillai *et al.,* 2021).

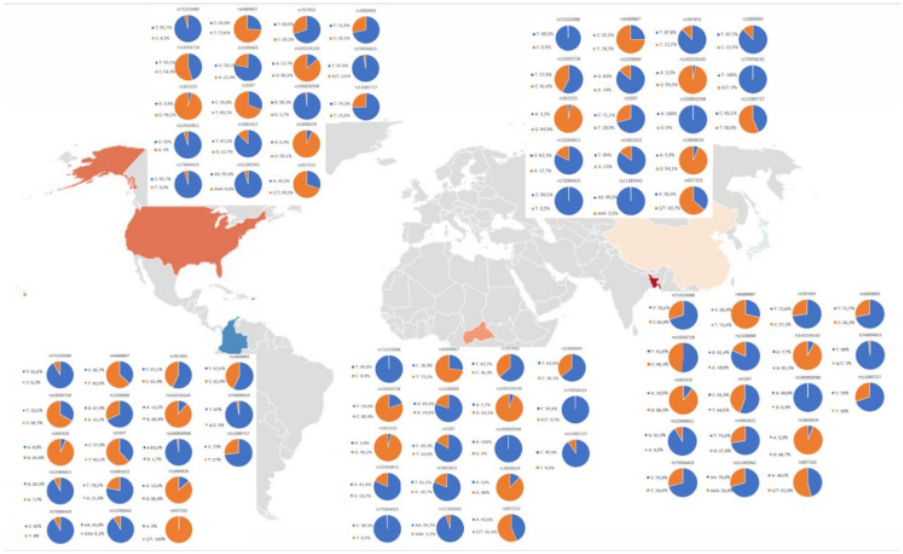

Figure 4. Allele frequencies in populations from Africa, America, East Asia, Europe, and South Asia.

The Ensembl Genome Browser was used to identify the densest populations, and it revealed the highest allele frequency (100%) in Africa, America, East Asia, Europe, and South Asia. Specifically, the SNP variant rs190850598, which codes for the GPHN gene, exhibited the highest allele frequency (100%) in both Africa and East Asia. Similarly, the SNP variant rs74956615, encoding the *RAVER1* gene, was also found in East Asia with the highest allele frequency (100%). In the case of the missense mutation-coding SNP variant rs1061622, which codes for the *TNFRSF1B* gene, the highest allele frequency (87.3%) was observed in America. However, no allele frequency was found for the SNP variant rs200553089, which codes for the *TLR7* gene, in any of the five countries.

The highest allele frequency (100%) in the SNP variant rs190850598 was found in Africa. According to the World Health Organization (WHO) report as of August 21, 2020, Africa had recorded 22,536,278 confirmed cases of COVID-19, with 789,197 deaths and a COVID-19 mortality rate of 1.7%. This mortality rate is slightly lower than the global COVID-19 mortality rate of 3.3% (Altable & de la Serna, 2021). Africa experienced the spread of COVID-19 relatively later compared to other regions, but its population is susceptible to viral infections such as SARS-CoV-2. Previous studies have identified higher prevalence rates in three African countries: Mombasa (9.3%), Nairobi (8.5%), and Kisumu (6.5%) (Altable & de la Serna, 2021). Furthermore, studies have also reported dense populations in Africa and Asia with INF-y SNPs associated with SARS-CoV-2 (Vakil *et al*., 2022). The African population's immunity is often compromised due to a high prevalence of underlying diseases like cardiovascular, cancer, and chronic respiratory diseases (Lone & Ahmad, 2020). This may be a contributing factor to weakened immunity. Additionally, the effectiveness of vaccines administered to the local population is considered limited due to non-specific drug derivatives targeting disease genes, leading to polymorphism (Altable & de la Serna, 2021). Moreover, transmission rates tend to be higher during the winter season compared to the summer season. Previous studies have also identified a variant that codes for the *OAS1* gene in the African population, which can produce degradative enzymes against pathogens, including SARS-CoV-2 (Biancolella *et al*., 2022).

In East Asia, the SNP variant rs190850598, which encodes the *GPHN* gene, also showed the highest allele frequency of 100%. A previous study in East Asia identified another

specific SNP variant associated with lung cancer and interstitial lung disease, which is related to the expression of the *FOXP4* gene (Niemi *et al*., 2022). The *FOXP4* gene is secreted by alveolar cells in the lungs and plays a role in lung function (Niemi *et al*., 2022). Furthermore, a high population in East Asia was also found to have the SNP variant rs1061622, which encodes the *TNFRSF1B* gene. This variant has been associated with increased response to SARS-CoV-2, leading to the production of pro-inflammatory cytokines and dysregulation of the immune response (Celik *et al*., 2020).

The rs1061622 SNP variant, which encodes the *TNFRSF1B* missense gene, exhibits the highest allele variation trend in the American population. This trend correlates with the increasing number of COVID-19 cases in the country, which rose from 11.3% to 24.5% in 2020 (Pei *et al*., 2021). In a previous study, it was found that adults over the age of 65 in the American population had the highest risk of contracting COVID-19, with a rate of 43% compared to 28% among individuals aged 18-24 years (Raifman & Raifman, 2020).

CONCLUSION

Our identification results of SNP variants influencing COVID-19 disease reveal the presence of 19 SNP variants. Among these variants, the TLR7 gene exhibited the highest expression levels in various tissues, including lymphocytes, spleen, bone marrow, small intestine or ileum, midbrain, entire circulatory system, lungs, nerves, hypothalamus, and subcutaneous adipose tissue. The TNFRSF1B gene expression was detected in the circulatory system, spleen, subcutaneous adipose tissue, lymphocytes, lungs, visceral adipose tissue, breasts, nerves, small intestine ileum, and cervix. The SNP variants, specifically rs200553089 and rs1061622, affected these genes and were observed at varying frequencies across different populations and genders. Therefore, the allele frequency of each variant becomes an important consideration in predicting the susceptibility of a population to COVID-19. The findings of this study provide insights for further research in determining the association between SNP variants, gene expression, and the severity and susceptibility levels of COVID-19 in dominant human populations worldwide. In the future, this study's findings can be applied to guide risk assessment and management strategies related to the identified SNP variants and their associated gene expression in COVID-19.

REFERENCES

Asano, T., Boisson, B., Onodi, F., Matuozzo, D., Moncada-Velez, M., Renkilaraj, M. R. L. et al.,, 2021, X-linked recessive TLR7 deficiency in `1% of men under 60 years old with life-threatening COVID-19, *Science Immunology*. 6.

Altable, M. and de la Serna, J. M, 2021, Protection against COVID-19 in african population: immunology, genetics, and malaria clues for therapeutic targets, *Virus Research*. 299(198347).

Biancolella, M., Colona, V. L., Mehrian-Shai, R., Watt, J. L., Luzzatto, L., Novelli, G., and Reichardt, J. K. V, 2022, Covid-19 2022 update: transition of the pandemic to the endemic phase, *Human Genomics*. 16(19).

Celhar, T., Lu, H. K., Benso, L., Rakhilina, L., Lee, H. Y., Tripathi, S., Zharkova, O., Ong, W. Y., Yasuga, H., Au, B., Marlier, D., Lim, L. H. K., Thamboo, T. P., Mudgett, J. S., Mackey, M. F., Zaller, D. M., Connolly, J. E., and Fairhurst, A, 2019, TLR7 protein expression in mild and severe lupus-prone models is regulated in a leukocyte, genetic, and IRAK4 dependent manner, *Frontiers in Immunology*. 10(1546).

Celik, S. K., Genc, G. C., and Dursun, A, 2020, A bioinformatic approach to investigating cytokine genes and their receptor variants in relation to COVID-19 progression, *International Journal of Immunogenetics*. 48(2): 211-218.

Deng, H., Yan, X. and Yuan, L, 2021, Human genetic basis of coronavirus disease 2019, *Signal Transduction and Targeted Therapy*. 6(344).

Fricke-Galindo, I., Buendia-Roldan, I., Ruiz, A., Palacios, Y., Perez-Rubio, G., Hernandez-Zenteno, R. J., et al., 2022, *TNFRSF1B* and *TNF* variants are associated with differences

in levels of soluble tumor necrosis factor receptors in patients with severe COVID-19, *The Journal of Infectius Diseases*. 226(5): 778-787.

Ganasegeran, K., Jamil, M. F. A., Ch'ng, A. S. H., Looi, I., and Peariasamy, K. M, 2021, Influence of population density for COVID-19 spread in Malaysia: an ecological study, *International Journal of Environmental Research and Public Health*. 18(18): 9866.

Harper, H., Burridge, A., Winfield, M., Finn, A., Davidson, A., Matthews, D., Hutchings, S., Vipond, B., Jain, N., The Covid-19 Genomics UK (COG-UK) Consortium, Edwards, K., and Barker, G, 2021, Detecting SARS-CoV-2 variants with SNP genotyping, *PLoS ONE*. 16(2): 1-12.

Huang, Y. VanderWeele, T. J., and Lin, X. 2014. Joint analysis of SNP and gene expression data in genetic association studies of complex diseases, *The Annals of Applied Statistics,* 8(1): 352.

Jin, J., Bai, P., He, W., Wu, F., Liu, X., Han, D., Liu, S., and Yang, J, 2020, Gender differences in patients with COVID-19: focus on severity and mortality, *Frontier Public Health*. 8:152.

Kersey, P. J., Staines, D. M., Lawson, D., Kulesha, E., Derwent, P., Humphrey, J. C., et al., 2012, Ensembl genomes: an integrative resource for genome-scale data from non-vertebrate species, *Nucleic Acids Research*. 40(1): 91-97.

Liu, Y-C., Kuo, R-L., and Shih, S-R, 2020, COVID-19: The first documented coronavirus pandemic in history, *Biomedical Journal*. 43(4): 328-333.

Lone, S. A. and Ahmad, A, 2020, COVID-19 pandemic-an african perspective, *Emerg Microbes Infect*. 9(1): 1300-1308.

NCBI, 2022, TNFRSF1B TNF receptor superfamily member 1B [*Homo sapiens* (human)], National Center for Biotechnology Information.

Niemi, M. E. K., Daly, M. J., and Ganna, A, 2022, The human genetic epidemiology of COVID-19, *Nature Reviews Genetics*. 23(9): 533-546.

Noh, J. and Danuser, G, 2021, Estimation of the fraction of COVID-19 infected people in U.S. states and countries worldwide, *PLoS ONE*. 16(2): 1-10.

Pei, S., Yamana, T. K., Kandula, S., Galanti, M., and Shaman, J, 2021, Burden and characteristics of COVID-19 in the United States during 2020, *Nature*. 598: 338-341.

Petrosino, M., Novak, L., Pasquo, A., Chiaraluce, R., Turina, P., Capriotti, E., and Consalvi, V, 2021, Analysis and interpretation of the impact of missense variants in cancer, *International Journal of Molecular Sciences*. 22(11): 5416.

Pillai, J. A., Bebek, G., Khrestian, M., Bena, J., Bergmann, C. C., Bush, W. S., Leverenz, J. B., and Bekris, L. M, 2021, *TNFRSF1B* gene variants and related soluble TNFR2 levels impact resilience in alzheimer's disease, *Frontiers Aging Neuroscience*. 13.

Raifman, M. A. and Raifman, J. R, 2020, Disparities in the population at risk of severe illness from COVID-19 by rare/ethnicity and income, *American Journal of Preventive Medicine*. 59(1): 137-139.

Sindhu, S., Wilson, A., Akhter, N., Shenouda, S., Kochumon, S., Behbehani, K., and Ahmad, R, 2015, Increased adipose tissue expression of Toll-Like Receptor (TLR)-7 in obese individuals: significance in metabolic disease. *Journal of Glycomics & Lipidomics*. 5(4): 1-6.

Singhal, T, 2020, A review of coronavirus disease-2019 (COVID-19), *Indian Journal of Pediatrics*. 87(4): 281-286.

Spiering, A. E. and Vriess, T. J, 2021, Why females do better: the x chromosomal TLR7 gene-dose effect in COVID-19, *Frontiers in Immunology*. 12(756262).

Spudich, G. M. and Fernandez-Suarez, X. M, 2012, Disease and phenotype data at Ensembl. *Current Protocols in Human Genetics*. 69(1): 1-34.

Stanfill, A. G. and Cao, X, 2021, Enhancing research through the use of the genotype-tissue expression (GTEx) database, *Biological Research for Nursing*. 23(3): 553-540.

Sy, K. T. L., White, L. F., and Nichols, B. E, 2021, Population density and basic reproductive number of COVID-19 across United States counties, *PLoS ONE*. 16(4): 1-11.

Vakil, M. K., Mansoori, Y., Al-Awsi, G. R. L., Hosseinipour, A., Ahsant, A., Ahmadi, S., Ekrahi, M., et al., 2022, Individual genetic variability mainly of proinflammatory cytokines, cytokine receptors, and toll-like receptors dictates pathophysiology of COVID-19 disease, *Journal of Medical Virology*. 94(9): 4088-4096.

Van der Made, C. I., Simons, A., Schuurs-Hoeijmakers, J., van den Heuvel, G., Mantere, T., Kersten, S., van Deuren, R. C., et al. 2020, Presence of genetic variants among young men with severe COVID-19, *The Journal of the American Medical Association*. 324(7): 663-673.

Ward, L. D. and Kellis, M, 2012, HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants, *Nucleic Acids Research*. 40: 930-934.

Ward, L. D. and Kellis, M, 2016, HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease, *Nucleic Acids Research*. 44(4): 887-881.

Wu, Y., Chen, C., and Chan, Y, 2020, The outbreak of COVID-19: an overview, *Journal of the Chinese Medical Association*. 83(3): 217-220.

Xing-rong, W., Sheng-qian, X., Wen, L., Shan, Q., Fa-ming, P., and Jian-hua, X, 2018, Role of *TNFRSF1A* and *TNFRSF1B* polymorphisms in susceptibility, severity, and therapeutic efficacy of etanercept in human leukocyte antigen-B27-positive Chinese Han patients with ankylosing spondylitis, *Medicine (Baltimore)*. 97(31):11677.

Zarember, K. A. and Godowski, P. J, 2002, Tissue expression of human toll-like receptors and differential regulation of toll-like receptor mRNAs in leukocytes in response to microbes, their products, and cytokines, *The Journal of Immunology*. 168(2): 554-561.

Zhou, Y., Xu, J., Hou, Y., Leverenz, J. B., Kallianpur, A., Mehra, R., Liu, Y., Yu, H., Pieper, A. A., Jehi, L., and Cheng, F, 2021, Network medicine links SARS-CoV-2/COVID-19 infection to brain microvascular injury and neuroinflammation in dementia-like cognitive impairment, *Alzheimer's Research & Therapy.* 13(110).

UTILIZING A GENOMIC DATABASE FOR IDENTIFYING GENOMIC VARIATION FOR SARS-CoV-2 RECEPTORS

ORIGINALITY REPORT

16%	12%	11%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Liverpool John Moores University Student Paper	1%
2	www.genominfo.org Internet Source	1%
3	egonw.github.io Internet Source	1%
4	rupress.org Internet Source	1%
5	www.researchsquare.com Internet Source	1%
6	Sevim Karakas Celik, Gunes Cakmak Genc, Ahmet Dursun. "A bioinformatic approach to investigating cytokine genes and their receptor variants in relation to COVID-19 progression", International Journal of Immunogenetics, 2020 Publication	<1%
7	www.ncbi.nlm.nih.gov	

	Internet Source	<1%
8	Manuel Martinez. "FROM PLANT GENOMES TO PROTEIN FAMILIES: COMPUTATIONAL TOOLS", Computational and Structural Biotechnology Journal, 2013 Publication	<1%
9	www.ideals.illinois.edu Internet Source	<1%
10	Submitted to Oklahoma State University Student Paper	<1%
11	www.journal.uad.ac.id Internet Source	<1%
12	www.news-medical.net Internet Source	<1%
13	www.researchwithnj.com Internet Source	<1%
14	operamedphys.org Internet Source	<1%
15	journals.lww.com Internet Source	<1%
16	uwspace.uwaterloo.ca Internet Source	<1%
17	Madeleine J. Bunders, Marcus Altfeld. "Implications of Sex Differences in Immunity	<1%

for SARS-CoV-2 Pathogenesis and Design of Therapeutic Interventions", Immunity, 2020
Publication

18	Oliveira Adrhyan Araújo da Silva. "ACE2 and TMPRSS2 polymorphisms and the development of COVID-19: a review of the literature", International Journal of Clinical Virology, 2022 Publication	<1 %
19	www.referencecitationanalysis.com Internet Source	<1 %
20	Malik Ibrahim, Anisah, Wita Ratna Puspita, Rody Satriawan, Lalu Jaswandi, Aminullah, Johri Sabaryati, Syaharuddin. "Maximum turning point and final spread of COVID-19 in Indonesia: An analysis of trends and data patterns", AIP Publishing, 2022 Publication	<1 %
21	Wang Xing-rong, Xu Sheng-qian, Liu Wen, Qi Shan, Pan Fa-ming, Xu Jian-hua. "Role of TNFRSF1A and TNFRSF1B polymorphisms in susceptibility, severity, and therapeutic efficacy of etanercept in human leukocyte antigen-B27-positive Chinese Han patients with ankylosing spondylitis", Medicine, 2018 Publication	<1 %
22	Marcos Altable, Juan Moisés de la Serna. "Protection against COVID-19 in African	<1 %

population: Immunology, genetics, and malaria clues for therapeutic targets", Qeios Ltd, 2021
Publication

23	downloads.hindawi.com Internet Source	<1 %
24	Jungsik Noh, Gaudenz Danuser. "Estimation of the fraction of COVID-19 infected people in U.S. states and countries worldwide", PLOS ONE, 2021 Publication	<1 %
25	Yue Jiang, Yanlin Li, Aijun Yao, Xiangzhi Gao, Hui Li. "Soil Deformation Investigation of a Piled-Raft Foundation Pit Under-Crossed by a Super-Large Diameter Shield Tunnel", Applied Sciences, 2023 Publication	<1 %
26	covid.kemri-wellcome.org Internet Source	<1 %
27	ojs.stikesylpp.ac.id Internet Source	<1 %
28	repo.lib.tokushima-u.ac.jp Internet Source	<1 %
29	www.journal.ugm.ac.id Internet Source	<1 %
30	www2.mdpi.com Internet Source	

		<1 %
31	Ingrid Fricke-Galindo, Ivette Buendia-Roldan, Leslie Chavez-Galan, Gloria Pérez-Rubio et al. "SERPINE1 rs6092 Variant Is Related to Plasma Coagulation Proteins in Patients with Severe COVID-19 from a Tertiary Care Hospital", Biology, 2022 Publication	<1 %
32	immunodeficiency.ca Internet Source	<1 %
33	repositorio.cientifica.edu.pe Internet Source	<1 %
34	scholarworks.iupui.edu Internet Source	<1 %
35	univendspace.univen.ac.za Internet Source	<1 %
36	www.koreascience.or.kr Internet Source	<1 %
37	www.techscience.com Internet Source	<1 %
38	Levente Zsichla, Viktor Müller. "Risk Factors of Severe COVID-19: A Review of Host, Viral and Environmental Factors", Viruses, 2023 Publication	<1 %

39	ddd.uab.cat Internet Source	<1 %
40	escholarship.org Internet Source	<1 %
41	journals.plos.org Internet Source	<1 %
42	libweb.kpfu.ru Internet Source	<1 %
43	livrepository.liverpool.ac.uk Internet Source	<1 %
44	medical.science.journal-asia.education Internet Source	<1 %
45	periodicos.uninove.br Internet Source	<1 %
46	repository.upnvj.ac.id Internet Source	<1 %
47	ueaeprints.uea.ac.uk Internet Source	<1 %
48	Lawrence Omo-Aghoja, Emuesiri Goodies Moke, Kenneth Kelechi Anachuna, Adrian Itivere Omogbiya et al. "COVID-19 pandemic: the implications of the natural history, challenges of diagnosis and management for care in sub-Saharan Africa", Beni-Suef	<1 %

University Journal of Basic and Applied Sciences, 2021
Publication

---

49 Moritz Anft, Krystallenia Paniskaki, Arturo Blazquez-Navarro, Adrian Doevelaar et al. "COVID-19-Induced ARDS Is Associated with Decreased Frequency of Activated Memory/Effector T Cells Expressing CD11a++", Molecular Therapy, 2020 <1%
Publication

---

50 Ville Salo, Juhani Määttä, Eeva Sliz, Ene Reimann et al. "Genome-wide meta-analysis conducted in three large biobanks expands the genetic landscape of lumbar disc herniations", Cold Spring Harbor Laboratory, 2023 <1%
Publication

---

51 Augusta Ortolan, Giacomo Cozzi, Mariagrazia Lorenzin, Paola Galozzi, Andrea Doria, Roberta Ramonda. "The Genetic Contribution to Drug Response in Spondyloarthritis: A Systematic Literature Review", Frontiers in Genetics, 2021 <1%
Publication

---

52 Petter Brodin. "SARS-CoV-2 infections in children: Understanding diverse outcomes", Immunity, 2022 <1%
Publication

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		