Article

Discovery Medicine 2024; 36(186): 1513–1526
https://doi.org/10.24976/Discov.Med.202436186.140

Whole-Exome Sequencing Detecting a Recurrent Pathogenic Mutation, *HFE* p.His63Asp (H63D) in COVID-19 Patients and Its Effect on Mortality

Rashid Mir[1,*], Imadeldin Elfaki[2], Mohammad A Alanazi[1], Naseh A. Algehainy[1], Faisal H Altemani[1], Badr A Alsayed[3], Elsiddig Idriss Mohamed[4], Syed Khalid Mustafa[5], Mamdoh S Moawadh[1], Faris J Tayeb[1], Jaber Alfaifi[6], Sael M Alatawi[1], Mohammad Muzaffar Mir[7], Mohammad Fahad Ullah[1]

[1]Department of Medical Lab Technology, Prince Fahad Bin Sultan Chair for Biomedical Research, Faculty of Applied Medical Sciences, University of Tabuk, 71491 Tabuk, Saudi Arabia
[2]Department of Biochemistry, Faculty of Science, University of Tabuk, 71491 Tabuk, Saudi Arabia
[3]Department of Internal Medicine, Faculty of Medicine, University of Tabuk, 71491 Tabuk, Saudi Arabia
[4]Department of Statistics, University of Tabuk, 71491 Tabuk, Saudi Arabia
[5]Department of Chemistry, Faculty of Science, University of Tabuk, 71491 Tabuk, Saudia Arabia
[6]Department of Child Health, College of Medicine, University of Bisha, 61922 Bisha, Saudi Arabia
[7]Department of Clinical Biochemistry, College of Medicine, University of Bisha, 61922 Bisha, Saudi Arabia
*Correspondence: rashidmirut@gmail.com (Rashid Mir)
Published: 20 July 2024

Background: In recent years, various coronaviruses have caused severe respiratory illnesses worldwide. For example the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) infections of COVID-19 outbreak in 2019 in Wuhan, China. Genome-wide association studies (GWAS) have significantly expanded our comprehension of how specific genetic variations are linked to diseases. Research has demonstrated the existence of genetic factors influencing susceptibility to coronaviruses. The objective of this study was to examine the association of certain loci with the COVID-19 in Saudi population.

Methods: In the present study we have examined the link between the COVID-19 disease and certain genetic variants in hospitalized COVID-19 patients (n = 16) in Tabuk and Bisha, Kingdom Saudi Arabia. We used the genome Analysis Toolkit (GATK) and Comprehensive variant annotation was performed different databases and tools such as Search Tool for the Retrieval of Interacting Genes (STRING), PanelApp and PolyPhen-2.

Results: The study showed that the genetic variants associated with genes such as Homeostatic Iron Regulator (*HFE*) (found in 7 patients, representing 44%), complement factor H (*CFH*) (6 patients, 38%), cadherin 23 (*CDH23*) (4 patients, 25%), cytotoxic T-lymphocyte associated protein 4 (*CTLA-4*) (3 patients, 19%), Transforming Growth Factor Beta 1 (*TGFB1*) (3 patients, 19%), CREB-binding protein (*CREBBP*) (2 patients, 13%), E1A Binding Protein P300 (*EP300*) (2 patients, 13%), hemoglobin subunit beta (*HBB*) (2 patients, 13%), interferon regulatory factor 7 (*IRF7*) (2 patients, 13%), and unc-119 lipid binding chaperone (*UNC119*) (2 patients, 13%) might be associated with susceptibility to coronavirus. We also identified mutations in the COVID-19 patient that are pathogenic or likely pathogenic.

Conclusion: A recurrent pathogenic mutation, *HFE* p.His63Asp (H63D), was identified in 7 patients, suggesting its potential contribution to disease severity. Additionally, a likely pathogenic variant, *HBB* p.Glu7Val (E7V), was present in 2 patients, highlighting its potential role in disease susceptibility. Our results shed light on the key genetic mechanisms of COVID-19 pathogenesis and help to identify and stratify the individuals or populations that are at risk to corona virus infection. The identification of susceptible individuals or populations assist in prevention and/or in treatment programs.

Keywords: coronavirus – COVID-19 – whole exome sequencing - *HFE* gene; GWASs

Introduction

The coronavirus belongs to the viral family Coronaviridae which comprises the subfamilies Coronavirinae and Torovirinae [1]. The members of the Coronaviridae contain an envelope. This envelope contains three proteins, namely the envelope protein (E), membrane protein (M) and spike glycoprotein (S) [1,2]. There are four genera in Coronaviridae family that include Alphacoronavirus, Betacoronavirus, Deltacoronavirus, and Gammacoronavirus [3].

COVID-19 is from the B strain of the genus Betacoronavirus. The Betacoronavirus infects only mammalian species [3]. The coronavirus is positive-sense single-stranded RNA virus with a genome of about 30 thousand nucleotide in length [4]. The severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) is a novel betacoronavirus responsible for the human coronavirus disease COVID-19 which spread in a pandemic proportion globally after initially starting in later months of the year 2019 in Wuhan, China [4]. The COVID-19 infection could rather easily transmit from person to person and rapidly spread among the population worldwide by March 2020 [2]. Symptoms of COVID-19 include headache, hyperthermia, cough, breath shortness, generalized weakness, sputum, and chest pain [5,6]. In addition to some neurological symptoms, dermatological manifestations, loss of appetite, muscles pain, sneezing, sore throat, rhinitis and diarrhea have been extensively reported in patients [5].

Several studies have suggested the role of human host gene variations in the infection of COVID-19 because of the differences in the clinical manifestations and rate of mortality between individuals, populations and ethnicities [3,7]. These studies may provide evidences of the modulation of host genetics on the pathogenesis of COVID-19 virus [7]. Coronaviruses also caused human respiratory infections in China in 2002 - SARS-CoV [8] followed by Middle East respiratory syndrome corona virus (MERS-CoV) in Saudi Arabia in 2012, [8]. It is known that certain risk loci are linked to susceptibility to diseases. Specific genetic variations can increase the susceptibility or severity of individuals or populations to various pathological conditions such as diabetes, cancers, cardiovascular diseases (CVD), and infections [9,10].

The Genome-wide association studies (GWAS) are employed successfully to uncover the gene variations or loci associated with diseases, such as diabetes mellitus, cardiovascular disease, cancers, COVID-19 and others [11–19]. Some of these have reported the linkage between the COVID-19 infection and genetic variants [18]. For example, the loci 3q21.31 (leucine zipper transcription factor like 1 (*LZTFL1*) and chemokine receptor genes) was reported to be linked with the severity of COVID-19 disease in patients, while 9q34.2 (*ABO*) was found to be linked to the susceptibility to COVID-19 infection [18]. Moreover, the genes that are expressed in pulmonary tissues and have been linked to the progression of emphysema, obstruction of airway and surface tension within the respiratory system and the genes that are involved in secretion of inflammatory cytokines from the T lymphocytes were reported to be associated with COVID-19 infection [11].

In the current study, we examined the gene variations associated with the susceptibility to COVID-19 in hospitalized cases from the Saudi Arabian population to explore the risk loci and the likely pathogenic variants associated with this disease. The results will help in identification of the underlying genetic mechanisms of COVID-19 pathogenicity and aid in the detection of susceptible individuals that will help in prevention and/or in treatment of this viral infection.

Materials and Methods

Study Population

The present study included 16 COVID-19 patients from two cities in Saudi Arabia; Tabuk and Bisha. The group included 16 COVID-19 patients admitted to two tertiary care hospitals. Over the study period, 16 consecutive admissions from King Fahad Hospital, Tabuk city and King Abdullah Hospital, Bisha city respectively were included. A positive result on nasal and oropharyngeal swab samples tested with the RealStar® SARS-CoV-2 Real-time polymerase chain reaction (RT-PCR) Kit 1.0 (Altona Diagnostics GmbH, Hamburg, Germany) was considered a laboratory confirmation of COVID-19. We excluded cases with negative results of the nasal and oropharyngeal swab samples.

Sample Collection

A peripheral blood sample of around 2 mL was collected from the 16 subjects in Ethylenediaminetetraacetic acidV (EDTA) tube for the purpose of Whole-Exome Sequencing (WES) and kept in storage right away at –20 °C till analysis. Using structured questionnaire, all included subjects were interviewed with regard to epidemiological and demographic information, history of coronary artery disease (CAD), type 2 diabetes mellitus (T2DM), chronic kidney disease (CKD) and addictions, particularly smoking. Also, family history of any other important illnesses was documented.

DNA Extraction and Sequencing

High-quality human genomic DNA was utilized as the starting specimen for this study. The DNA was extracted with Qiagen DNA extraction kit (Catalog No. 51106), Hilden, Germany. Following the protocols outlined in the twist Exome 2.0, 96 Reactions, Kit (Cat#104136), Twistbioscience, South San Francisco, CA, USA, an optimized library was generated. Subsequently, sequencing was performed on the Illumina NovaSeq 6000 platform (Illumina Way, San Diego, CA, USA), according to the instruc-

tions. To assess the quality of the sequencing reads, FastQC v0.11.9 (Java software) was employed. Subsequently, adapter sequences and low-quality bases were eliminated from the raw reads using TrimGalore v0.6.6. The resulting high-quality reads were mapped onto the hg38 human reference genome assembly.

Variant Calling and Annotation

The genome Analysis Toolkit (GATK) v4.2.4.1 (Broad Institute of MIT and Harvard, Cambridge, MA, USA, https://www.broadinstitute.org/scientific-community/software/genome-analysis-toolkit-gatk) best practice pipeline, incorporating the Haplotype Caller for small-variant calling (single nucleotide variants (SNVs) and short insertions/deletions (InDels)), was implemented for variant discovery. Comprehensive variant annotation was performed leveraging diverse databases and tools. The RefSeq database [20] was utilized to identify and characterize genes harboring potential variants. To elucidate potential disease associations of identified variants, public databases like OMIM (Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD), {December 2023}. World Wide Web URL: https://omim.org/) and ClinVar [21] were extensively employed. Furthermore, population frequency information was compiled from resources such as the 1000 Genomes Project [22], ExAC, GnomAD [23] (exome and genome), and ESP to effectively distinguish variants from common polymorphisms. Variants were filtered according to a minor allele frequency (MAF) <0.01 in population databases. Additionally, synonymous, intronic, and non-coding variants were excluded from the analysis. The PolyPhen-2 (version 2.2.3, release 405c) [24] and SIFT (6.2.1) [25] prediction tools were utilized to assess the potential functional impact of non-synonymous coding SNVs on protein structure and function. Additionally, *in silico* variant effect prediction was performed for all variants using a multitude of independent prediction tools. Finally, all variants were meticulously interpreted in accordance with the American College of Medical Genetics and Genomics (ACMG) guidelines [26]. Pathogenic, likely pathogenic, and variants of uncertain significance were subsequently reported.

Gene Interaction Network Construction

A curated list of genes harboring recurrent pathogenic variants was subjected to network analysis within the Search Tool for the Retrieval of Interacting Genes (STRING) database v12.0 [27]. Utilizing the database's robust predictions of protein-protein interactions, a downloadable tabular file (.tsv) was acquired through the Exports tab. This data served as the foundation for visualizing the network using Cytoscape version 3.10.1 [28], a software specifically designed for biomolecular interaction analysis.

Functional Prioritization and Enrichment Analysis

To prioritize key genes within the network, an additional filtering step was implemented. Genes were cross-referenced with the COVID-19 severity and susceptibility panel available in PanelApp [29]. This resource, populated by expert consensus, enabled the identification of genes possessing high review ratings, suggesting their potential relevance to the disease context. These prioritized genes were subsequently fed into the g: profiler webserver (e110_eg57_p18) [30] for enrichment analysis. This comprehensive platform assessed gene sets for overrepresentation in Gene Ontology (GO) terms, highlighting unique terms with a statistical significance threshold of –log10p score >=3.

Statistical Analysis

Statistical analyses were conducted utilizing SAS Institute Inc.'s statistical software version 9.4 (Cary, NC, USA), Stata Corp.'s Stata statistical software Release 13 (College Station, TX, USA), Med-Calc software version 20.027 (https://www.medcalc.org/), and SPSS Inc.'s SPSS version 16.0 (IBM Corp., Chicago, IL, USA).

Variant Annotation and Prioritization

The Ensembl Variant Effect Predictor (VEP) is a powerful tool that provides comprehensive annotations for SNVs and insertions/deletions (indels) [31]. To ensure accurate results, we utilized default VEP annotations, as well as several plugins including dbNSFP, Combined Annotation-Dependent Depletion (CADD), DisGeNET, Gene Ontology, Geno2MP, LOEUF, mutfunc, and Phenotypes. To predict the functional consequences of mutations, we leveraged tools from dbNSFP such as PROVEAN, REVEL, ClinPred, Aloft, and DANN. For missense mutations, we also incorporated SIFT, Polyphen, and CADD scores for annotation. To determine the functional impact of indels, we used SIFT-indel [32]. Finally, we annotated allele frequencies using gnomAD (exomes) allele frequencies and only retained canonical transcript-dependent consequences per variant in the VEP-annotated file. Following the ACMG/AMP guidelines for the "Benign standalone" (BA1) criterion, we eliminated all alleles with a frequency of 5% or higher in gnomAD [23] or gnomAD Middle East [33]. The remaining variants were then compared to the Clinvar database to ascertain their classification as benign (B), likely benign (LB), likely pathogenic (LP), or pathogenic (P). Variants that were not found in the Clinvar database were deemed "deleterious" based on the following criteria: (i) Variants with a VEP reported impact of "low" were removed. Only variants with a reported impact of "high", "moderate" or "modifier" were considered. (ii) Variants were only considered deleterious if predicted to be so by SIFT [25], Polyphen [24], or PROVEAN [34]. (iii) A CADD [35] score greater than 20 was required. iv. For

DANN or REVEL scores [36,37], a score higher than 0.5 was necessary. V. For indels, SIFT-INDEL [32] must classify them as damaging.

Variant Classification

The aforementioned criteria were used to identify any deleterious variations, which were then categorized as either variants of uncertain significance (VUS) or LP. If three or more tools concurred on the classification, the variations were designated as LP. Conversely, if there were discrepancies among tool results, they were classified as VUS. All the identified LP genes were scrutinized for their connections with immune and respiratory functions or traits using GO process and Human Phenotype Ontology (HPO).

COVID-19 Phenotype Association with Variants

The 40001-U071-COVID-19-virus-identified phenotype was sourced from the AstraZeneca Phewas portal [38], a comprehensive database of gene-phenotype associations. This database is built from electronic health records, questionnaire data, and continuous traits computed on exomes shared by UK Biobank. All the variants that are present in this cohort were thoroughly analyzed for their association with the COVID-19 phenotype, and any overlapping variants between the two datasets were reported.

Results

Patient Characteristics

This study investigated the clinical characteristics of a cohort of 16 patients admitted to the hospital with COVID-19 symptoms. The group comprised 9 female (56.3%) and 7 male (43.8%) individuals. Eight patients (50%) achieved full recovery and were discharged, while eight (50%) succumbed to the severe complications of the disease. The median duration of hospital stay was 12 days (interquartile range: 9–16), with a median intensive care unit (ICU) stay of 8 days (5–12). The median age of the patients was 59 years (50–69). The most prevalent respiratory symptom was cough, reported by 88% (14/16) of patients, followed by fever, present in 63% (10/16). Among pre-existing comorbidities, hypertension was the most common, diagnosed in 75% (12/16) of patients. Cancer and diabetes were also prevalent, observed in 56% (9/16) and 31% (5/16) of patients, respectively. Chronic kidney disease was present in one patient (6.3%). Notably, all patients (100%) received treatment with immunosuppressive steroids and/or antiviral medications. Main clinical characteristics of this cohort are summarized in Table 1.

Table 1. Clinical characteristics of the COVID-19 cases.

Characteristic	COVID-19 Patients (n %)
	N = 16
Gender	
Female	9 (56%)
Male	7 (44%)
Age[1]	59 (50–69)
Hypertension	12 (75%)
Cancer	9 (56%)
Diabetes	
No	2 (13%)
Unknown	9 (56%)
Yes	5 (31%)
Smoking	
No	6 (38%)
Unknown	9 (56%)
Yes	1 (6%)
Fever	
No	3 (19%)
Unknown	3 (19%)
Yes	10 (63%)
Cough	
Unknown	2 (13%)
Yes	14 (88%)
Chronic kidney disease	
No	7 (44%)
Unknown	8 (50%)
Yes	1 (6%)
Ischemic heart disease	8 (50%)
Immunosuppressive medications	
No	2 (13%)
Unknown	3 (19%)
Yes	11 (69%)
Hospital duration (Days)[1]	12 (9–16)
ICU duration (Days)[1]	8 (5–12)
Outcome	
Death	8 (50%)
Improved	8 (50%)

[1]Median (IQR); ICU, intensive care unit.

Detection and Classification of Candidate Variants

Whole-exome sequencing identified 99 variants of interest in 54 genes within the cohort of 16 patients. Among these, 71 variants were unique to individual patients, while nine were recurrent, being present in at least two patients. Notably, 94 of the identified variants were single nucleotide polymorphisms (SNPs), with the remaining 5 categorized as insertions/deletions (indels). Further characterization revealed the predominant variant class to be missense mutations (n = 71), followed by silent (n = 19), intronic (n = 2), frameshift (n = 4), nonsense (n = 2), and a single in-frame indel. Applying the ACMG criteria for variant interpretation, 16 variants across 10 patients were classified as either pathogenic or likely pathogenic. The majority of variants (n = 50) fell into the category of VUS, while 33 were classified as likely benign (Fig. 1).

Focusing on non-synonymous variants present in two or more patients revealed a specific set of potentially influential genes. These included *HFE* (found in 7 patients, rep-

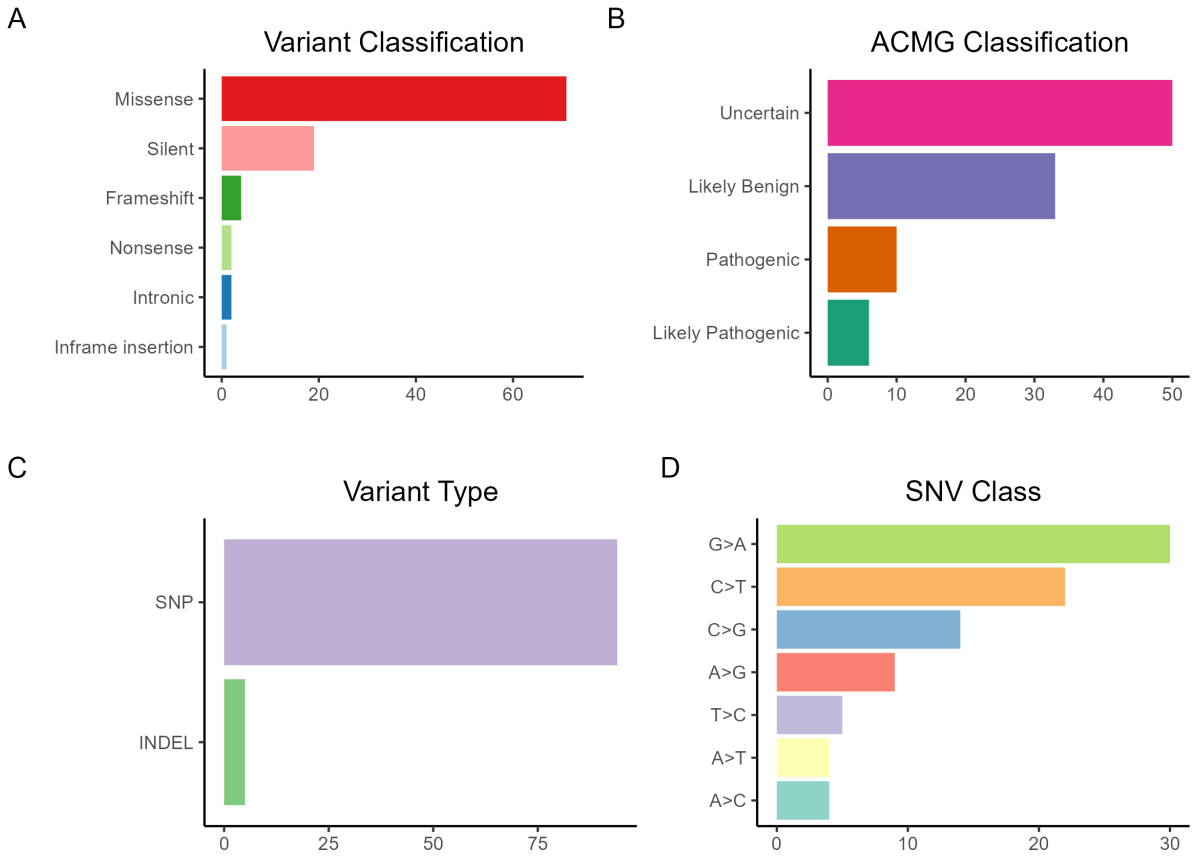

Fig. 1. Barplot showing the count of different categories of gene variants identified in COVID-19 patients shown in the x-axis. (A) Variant classification. (B) American College of Medical Genetics and Genomics (ACMG) Classification. (C) Variant type and (D) single nucleotide variat (SNV) class. SNP, single nucleotide ploymorphism; INDEL, insertion and deletion mutations.

resenting 44%), complement factor H (*CFH*) (6 patients, 38%), cadherin 23 (*CDH23*) (4 patients, 25%), cytotoxic T-lymphocyte associated protein 4 (*CTLA-4*) (3 patients, 19%), Transforming Growth Factor Beta 1 (*TGFB1*) (3 patients, 19%), CREB-binding protein (*CREBBP*) (2 patients, 13%), E1A Binding Protein P300 (*EP300*) (2 patients, 13%), hemoglobin subunit beta (*HBB*) (2 patients, 13%), interferon regulatory factor 7 (*IRF7*) (2 patients, 13%), and unc-119 lipid binding chaperone (*UNC119*) (2 patients, 13%) (Fig. 2).

Identification of Potentially Pathogenic Variants

Detailed analysis of the identified variants revealed 8 genes harboring a total of 16 variants which were classified as pathogenic or likely pathogenic, distributed across 10 patients within the cohort. Notably, among these 10 patients, 8 were female and 2 were male. A recurrent pathogenic mutation, *HFE* p.His63Asp (H63D), was identified in 7 patients, suggesting its potential contribution to disease severity. Additionally, a likely pathogenic variant, *HBB* p.Glu7Val (E7V), was present in 2 patients, highlighting its potential role in disease susceptibility. Table 2 lists details regarding all identified pathogenic and likely pathogenic variants.

Characterization of HFE p.His63Asp Variants and Association with Patient Outcomes

Our analysis revealed distinct patterns in *HFE* p.His63Asp (H63D) variants and their association with patient outcomes. Notably, both male patients presented with homozygous mutations (H63D/H63D), while five females harbored heterozygous mutations (H63D/wild type (WT)). A critical observation was the differential impact of genotype on patient survival. Unfortunately, none of the patients with the homozygous mutation (H63D/H63D) survived the illness, while four out of five (80%) patients with the heterozygous mutation (H63D/WT) demonstrated clinical improvement. This suggests a potential dose-dependent effect of the H63D mutation on disease severity. Furthermore, comorbidity analysis revealed that both patients with the homozygous p.His63Asp/p.His63Asp genotype harbored pre-existing conditions of hypertension and cancer. There was significant association emerged between *HFE* genotype and patient age at the time of infection and hospitalization. Patients bearing *HFE* mutant alleles (p.His63Asp/p.His63Asp or p.His63Asp/WT) presented with COVID-19 symptoms and required hospitalization at a significantly younger age compared to patients with the WT/WT genotype (p =

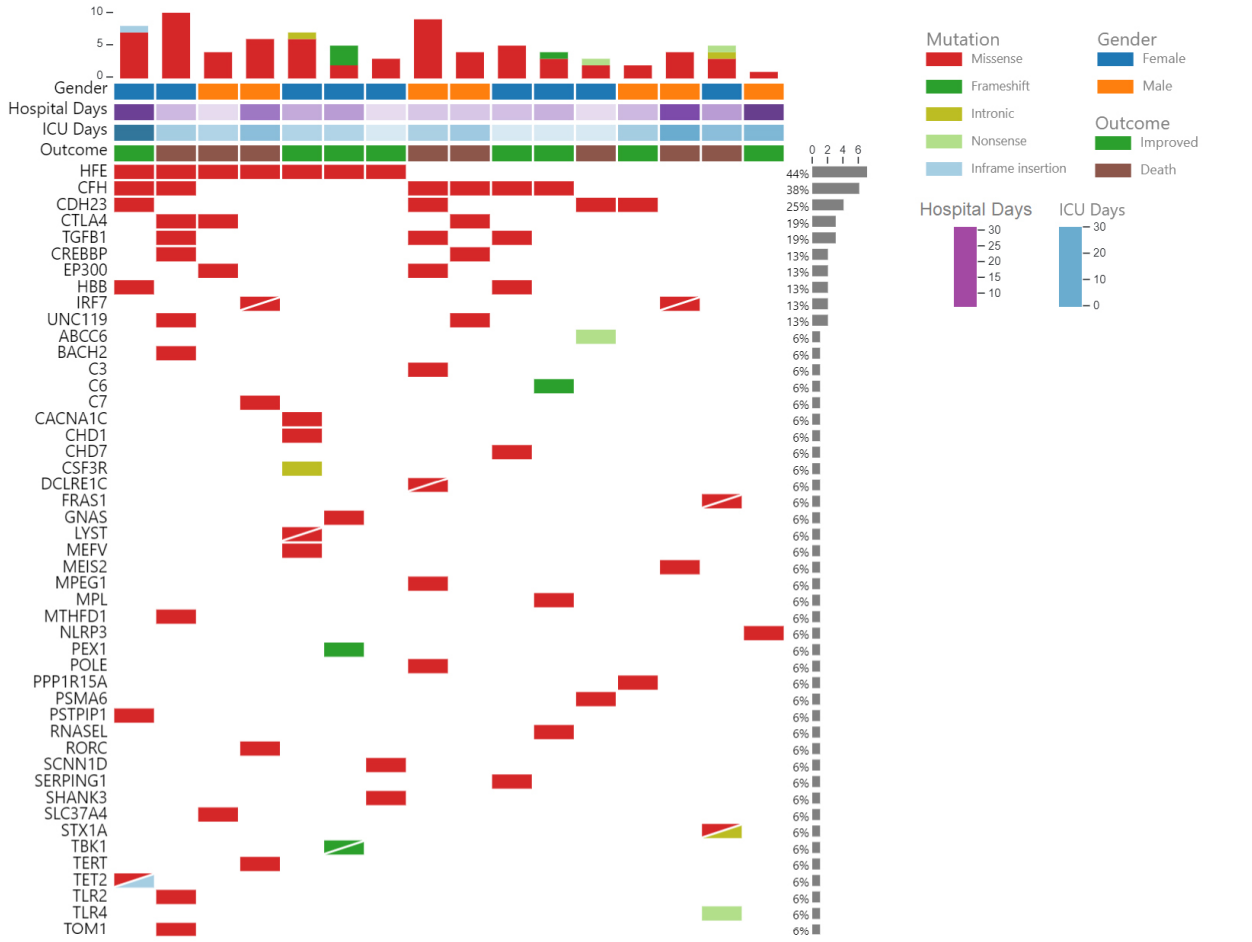

Fig. 2. Comut plot showing the recurrent variants and clinical characteristics of the COVID-19 patients. Each column represents an individual patient.

0.0093). Table 3 lists the clinical characteristics difference between the patients bearing *HFE* mutant alleles vs *HFE* wild types.

Network Analysis and Functional Characterization of HFE in COVID-19

To uncover potential functional partners of the *HFE* gene in the context of COVID-19 severity, a network analysis was conducted. This analysis revealed interactions between *HFE* and 10 other genes (Fig. 3). Subsequently, these 10 genes, along with *HFE*, were cross-referenced with the PanelApp COVID gene list for their association with COVID-19 severity and susceptibility. Interestingly, three genes - *HFE*, *B2M*, and *TFRC* - emerged with a green review rating, indicating strong evidence from expert reviewers that variations in these genes could contribute to disease severity.

Further insights were gained through GO enrichment analysis of these three key genes (Fig. 4). Beyond the expected involvement in the *HFE* transferrin complex, they were found to be significantly enriched in pathways related to regulating adaptive immunity ($p = 1.1 \times 10^{-4}$) and Major Histocompatabilty Complex (MHC) class I antigen presentation ($p = 8.2 \times 10^{-4}$) (Fig. 4). This suggests a potential connection between *HFE* and immune response pathways relevant to COVID-19.

Discussion

Understanding the risk loci will help to identify and stratify the susceptible individuals or populations, which will assist through the genetic testing to prevent or delay the incidence of these diseases [39,40]. In the present study, we examined gene variations in hospitalized patient of COVID-19. Our results showed that there were pathogenic and likely pathogenic variants in patients (Table 2). These gene variations include *HFE* (44%), *CFH* (38%), *CDH23* (25%), *CTLA-4* (19%), *TGFB1* (19%), *CREBBP* (13%), *EP300* (13%), *HBB* (13%), *IRF7* (13%), and *UNC119* (13%) (Fig. 2). The *HFE* gene (H = high; FE = iron) was detected in 44% from the ceases.

The *HFE* gene consists of seven exons expanded in 12-kilobase. *HFE* gene is composed of 9609 base pair of deoxyribonucleic acid and located on chromosome 6p

Table 2. Identified pathogenic (P) and likely pathogenic (LP) variants in patients.

Patient	Gender	Age	Outcome	Hugo symbol	HGVSc	HGVSp	ACMG classification
CCC3_S69	Female	50	Improved	TBK1	NM_013254.4:c.897dupT	p.Ala300fs	LP
				TBK1	NM_013254.4:c.1178delA	p.Ile393fs	LP
				PEX1	NM_000466.3:c.1108delA	p.Ile370fs	P
				HFE	NM_000410.4:c.187C>G	p.His63Asp	P
CCC8_S63	Female	64	Death	ABCC6	NM_001079528.4:c.281G>A	p.Trp94*	LP
CCC9_S56	Female	59	Improved	C6	NM_000065.5:c.1879delG	p.Asp627fs	P
				MPL	NM_005373.3:c.317C>T	p.Pro106Leu	P
CCC19_S71	Female	59	Improved	*HFE*	NM_000410.4:c.187C>G	p.His63Asp	P
CCC30_S58	Male	49	Death	*HFE*	NM_000410.4:c.187C>G	p.His63Asp	P
CCC38_S66	Female	52	Improved	*HFE*	NM_000410.4:c.187C>G	p.His63Asp	P
				MEFV	NM_000243.3:c.2230G>T	p.Ala744Ser	LP
CCC51_S59	Male	36	Death	*HFE*	NM_000410.4:c.187C>G	p.His63Asp	P
CCC52_S64	Female	50	Death	*HFE*	NM_000410.4:c.187C>G	p.His63Asp	P
CCC67_S60	Female	68	Improved	*HBB*	NM_000518.5:c.20A>T	p.Glu7Val	LP
CCC74_S65	Female	49	Improved	*HFE*	NM_000410.4:c.187C>G	p.His63Asp	P
				HBB	NM_000518.5:c.20A>T	p.Glu7Val	LP

p.Trp94* mutation, '*' represents the premature translational stop signal. It is a nonsense type of mutation. Abbreviations: HGVS, Human Genome Variation Society.

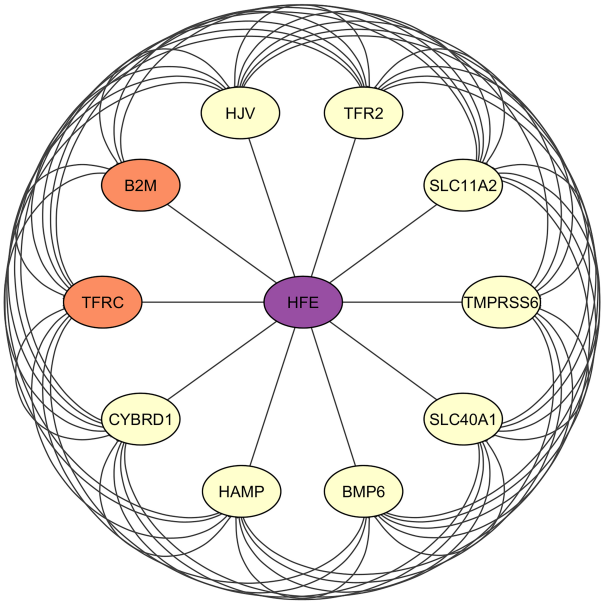

Fig. 3. The *HFE* interacting gene network with highlighted COVID-19 susceptibility genes.

within the extended human leukocytes antigen region tail [41,42]. The exon one encodes the signal peptide, and exons two to four encode alpha one, alpha two and alpha three domains, respectively. While exon 5 encodes for the transmembrane domain, and exon 6 encodes for the cytoplasmic tail [41]. The transmembrane protein *HFE* resembles MHC class I-type proteins and binds beta-2 microglobulin and transferrin receptor. The *HFE* plays a role in synthesis of hepcidin by the hepatocytes [43]. The hepcidin is a peptide hormone that regulates the iron metabolism [43]. It regulates the transport of iron to the blood plasma from the intestine, macrophage recycling, aging or damaged Red Blood Cells (RBCs) and from liver [44]. *HFE* affect the absorption of iron by regulating the hepcidin expression [41,45]. Iron has vital roles in different physiological processes. The iron sulfur proteins are found in all cells and are important for oxidation-reduction reactions producing energy [46], and it is required for oxygen transport by respiratory system to all organs of the body [47]. The loss of iron from bleeding or dermal desquamation are compensated by increased duodenal iron absorption [48]. However, when duodenal iron absorption exceeds the physiological limits iron overload is induced (hereditary hemochromatosis) [48]. The HH results in oxidative damage in parenchymal organs such as hepatic, myocardium, and pancreatic tissues [48–50]. Our result showed that the in *HFE* p.His63Asp is associated with patient outcomes. Both male patients with homozygous mutations (p.His63Asp/p.His63Asp), while five females harbored heterozygous mutations (p.His63Asp/WT). A critical observation was the differential impact of p.His63Asp on patient survival. Unfortunately, no patients with the homozygous mutation (p.His63Asp/ p.His63Asp) survived. While 80% of the cases with the heterozygous mutation (p.His63Asp/WT) improved with time. This indicates a possible dose-dependent effect of the p.His63Asp mutation on COVID-19 severity. The p.His63Asp results in exchange of histidine, a positively charged amino acid to aspartic amino acid, a negatively charged amino acid that may influence the *HFE* protein structure and function. It is possible that the p.His63Asp mutation affect the functions of the *HFE* protein rendering the carrier of this mutation susceptible to COVID-19. This result is consistent with a re-

Table 3. *HFE* p.His63Asp based clinical characteristics of the patients.

Characteristic	COVID-19 Patients (n %)		
	H63D/H63D, N = 2	H63D/WT, N = 5	WT/WT, N = 9
Gender			
Female	0 (0%)	5 (100%)	4 (44%)
Male	2 (100%)	0 (0%)	5 (56%)
Age[1]	43 (39–46)	50 (50–52)	68 (64–73)
Hypertension	2 (100%)	4 (80%)	6 (67%)
Cancer	2 (100%)	3 (60%)	4 (44%)
Diabetes			
No	0 (0%)	1 (20%)	1 (11%)
Unknown	2 (100%)	3 (60%)	4 (44%)
Yes	0 (0%)	1 (20%)	4 (44%)
Smoking			
No	0 (0%)	2 (40%)	4 (44%)
Unknown	2 (100%)	3 (60%)	4 (44%)
Yes	0 (0%)	0 (0%)	1 (11%)
Fever			
No	0 (0%)	0 (0%)	3 (33%)
Unknown	1 (50%)	0 (0%)	2 (22%)
Yes	1 (50%)	5 (100%)	4 (44%)
Cough			
Unknown	1 (50%)	0 (0%)	1 (11%)
Yes	1 (50%)	5 (100%)	8 (89%)
Chronic kidney disease			
No	0 (0%)	2 (40%)	5 (56%)
Unknown	2 (100%)	3 (60%)	3 (33%)
Yes	0 (0%)	0 (0%)	1 (11%)
Ischemic heart disease	2 (100%)	3 (60%)	3 (33%)
Immunosuppressive medications			
No	0 (0%)	1 (20%)	1 (11%)
Unknown	2 (100%)	0 (0%)	1 (11%)
Yes	0 (0%)	4 (80%)	7 (78%)
Hospital duration (Days)[1]	13 (10–17)	13 (11–15)	11 (9–15)
ICU duration (Days)[1]	9 (8–11)	6 (6–8)	8 (0–12)
Outcome			
Death	2 (100%)	1 (20%)	5 (56%)
Improved	0 (0%)	4 (80%)	4 (44%)

[1]Median (IQR). WT, wild Type; H63D, p.His63Asp.

port that the *HFE* mutation is associated with susceptibility to COVID-19 in the Czech and Croatian populations [51, 52]. Elevated iron concentration (caused by p.His63Asp mutation) results in oxidative stress, (through Haber-Weiss reaction), and defective immune response [52]. Thus increased iron concentration can enhance the viral infection and increase viral disease morbidity and mortality rates [52]. This is in agreement with studies which reported that iron overload results in metabolic and immune disturbances and that hereditary hemochromatosis (HH) is a risk factor for viral infection [53,54]. The co-morbidity analysis revealed that both patients with the homozygous p.His63Asp/p.His63Asp genotype harbored pre-existing conditions of hypertension and cancer. This raises intriguing questions about the potential interplay between *HFE* mutations and comorbidities in modulating COVID-19 susceptibility and outcomes. Interestingly, a significant association emerged between *HFE* genotype and patient age at the time of infection and hospitalization. Patients bearing *HFE* mutant alleles (p.His63Asp/p.His63Asp or p.His63Asp/WT) presented with COVID-19 symptoms and required hospitalization at a significantly younger age compared to patients with the WT/WT genotype ($p = 0.0093$). This finding warrants further investigation to elucidate the underlying mechanisms driving this age-dependent association.

Our result showed that the *HFE* gene interact with B2M, and TFRC. This result is in agreement with a study reporting that the B2M is elevated in cases with COVID-

GO Term	Term ID	p_{adj}	-log10p_{adj}
HFE-transferrin receptor complex	GO:1990712	5.7×10^{-9}	8
response to iron ion	GO:0010039	1.5×10^{-6}	6
iron ion transport	GO:0006826	1.0×10^{-5}	5
positive regulation of lymphocyte mediated immunity	GO:0002708	9.9×10^{-5}	4
positive regulation of adaptive immune response	GO:0002821	1.1×10^{-4}	4
antigen processing and presentation of peptide antigen via MHC class Ib	GO:0002428	8.2×10^{-4}	3

Fig. 4. Gene Ontology (GO) enrichment of the COVID-19 susceptibility genes.

19 and it is associated with bad prognosis [55]. The TFRC transferrin 1 receptor was also reported to be increased in COVID-19 infections [56,57]. Beyond the expected involvement in the *HFE* transferrin complex, they were found to be significantly enriched in pathways related to regulating adaptive immunity ($p = 1.1 \times 10^{-4}$) and MHC class I antigen presentation ($p = 8.2 \times 10^{-4}$). This suggests a potential connection between *HFE* and immune response pathways relevant to COVID-19. The *HFE*, hepcidin, and iron are very critical for adaptive immune response [58]. It has been reported that patients with iron overload or HH have T lymphocytes with defective functions, proliferation and differentiation [59], and that these patients are susceptible to infections [59]. Considering the reported loss-of-function nature of the *HFE* p.His63Asp variant [60], these findings support the hypothesis that patients carrying this mutation may have compromised adaptive immune function. This could partially explain the observed earlier age of infection and hospitalization in *HFE* mutant carriers. Nevertheless, additional experimental validation is imperative to reinforce this connection and clarify the underlying mechanisms.

Next, we found mutation in *CFH* gene. This gene is found in on chromosome 1q32 [61]. This gene encode for the factor H which is 155-kDa plasma glycoprotein involved in the regulation of the alternative pathway of complement [61]. The *CFH* polymorphisms were linked to sepsis associated with *Pseudomonas aeruginosa* in Chinese Han population [62], protection against dengue virus infection [63], conditions of complement dysfunction, for instance, typical and atypical hemolytic uremic syndrome, age-related macular degeneration, and dense deposit disease [61,64,65]. Factor H is an important complement soluble inhibitor [66]. It prevents the activation of complement and amplification on surface of host cells [61]. Moreover, tumor cells and pathogens can hijack the factor H to escape the immune system [61]. The COVID-19 spike proteins suppress the binding between factor H and heparin which may result in dysregulation of complement and promote the COVID-19 infection [67]. Our result showed that *CFH* gene mutation might be associated with COVID-19 infection risk. This result is in agreement with a study reported that the dysregulation of complement system is a risk factor for COVID-19 [67].

We also found the mutation in the gene *CDH23* - Cadherin-23. This result is consistent with a study that reported *CDH23* gene to be risk locus for COVID-19 [68]. Cadherin 23 is from the superfamily of cadherin. This family consists of transmembrane proteins that mediate calcium-dependent intracellular adhesion [69].

Our result also showed mutation in the gene Cytotoxic T-lymphocyte antigen 4 (*CTLA-4*). *CTLA-4* is an immune checkpoint and a negative regulator of the immune response and important for the function of the regulatory T-lymphocytes [70,71]. *CTLA-4* mutation results in immune dysfunction syndrome in human [72] such as the rheumatoid arthritis, autosomal dominant immune dysregulation, juvenile idiopathic arthritis and Addison's disease [71]. This mutation was linked with reduced number of T-lymphocytes, B-lymphocytes and natural killer cells [72]. This result comes in line with a study reported that pathogenic variants of *CTLA-4* is associated with prolonged COVID-19 [73]. Nevertheless, Mirsharif *et al.* [74], reported that *CTLA-4* gene variation is not associated with severity and mortality of COVID-19 infections in Iranian population. The effect of *CTLA-4* gene variations on COVID-19 remains to be investigated in further studies.

In addition, we found a mutation in the *TGFB1*. The *TGFB1* gene is found on chromosome 19 and encode for the TGF-β1 cytokine [75]. TGF-β regulates immune responses against microbial infection such as COVID-19 [76,77]. It also has important roles in immune homeostasis and tolerance [76]. The selection of T cells in thymus is regulated by TGF-β, and the homeostasis of the pool of the naïve T-cell is maintained by The TGF-β [76]. It also inhibits types of T cells such as cytotoxic T helpers 1 and 2 cells,

and enhancing the generation of peripheral T regulatory and T helpers 17 and 9 [76]. Regarding the B lymphocytes, it regulates their proliferation, activation, and differentiation [76]. Moreover, TGF-β also promotes the development of natural killer cells, macrophages, dendritic cells, and neutrophils [76]. Inhibitors of the TGF-β were suggested for the prevention and therapy of microvascular thrombosis and immune dysfunction associated with COVID-19 disease [78].

We also found mutation in *CREBBP,* this gene encodes for the CREB-binding protein *(CREBBP).* This conserved protein is a transcription coactivator and has crucial role in angiogenesis [79]. The CREB binding protein and *EP300* (mutation was also detected in *EP300* gene) are histone acetyl-transferases (HATs) [80]. The disruption of the *CREBBP* gene was reported to cause immune dysregulation [81]. Recently it has been reported that *EP300* deficient cells exhibit significant impaired activation of innate immune system [82], and that *EP300* mutations result in chronic stress of DNA replication, instability of genomic and defective innate immune system [82].

Result also showed that there a mutation in *HBB*, NM_000518.5:c.20A>T, p.Glu7Val. The *HBB* gene is found in chromosome 11 p15.5. The Mutations in the *HBB* gene causes inherited hemoglobinopathies, for example, the sickle cell disease (defective hemoglobin structure) and beta-thalassemia (inadequate amount of hemoglobin [83,84]. Patients with NM_000518.5:c.20A>T, p.Glu7Val mutation saw improvement. The beta-globin chain mutations lead to elevated sickling of RBCs in stress conditions, for instance infection dehydration, and hypoxia. This leads to hemolysis and vasooclusion [85]. The proactive testing and low threshold treatment can be opted to improve the conditions in patients with COVID-19 and sickle cell disease [85].

Moreover a mutation in the gene Interferon regulatory factor 7 *(IRF7*) has also been reported in this study. The *IRF7* (*IRF7* gene is located on human chromosome 11p15.5) is a member of the interferon regulatory factors (IRFs) family and an important regulator of type I interferons (IFNs) preventing the bacterial and viral infections and suppressing the tumor progression [86,87]. The infections stimulate *IRF7 via* activating signaling cascades from the pathogen recognition receptors (PRRs) that recognize the nucleic acids of the pathogens [86]. *IRF7* can enhance the inflammation and the development of inflammatory diseases [87]. Innate antiviral signaling stimulates the signaling of IRF3, *IRF7* and nuclear factor kappa B (NF-κB) to induce interferon (IFN) and other pro-inflammatory factors to suppress the replication of the virus [87]. It has been reported that children with mutated *IRF7* suffer from severe influenza since they have reduced amount of IFNs from the WBCs and dendritic, dermal fibroblasts and stem cells [88]. The IFNs exhibit antiviral effects for instance cytotoxic T-cell responses activation, suppression of the translation of viral messenger RNA, the viral RNA damage, editing of the RNA and modulating the production of nitric Oxide [89]. Moreover, the individuals with impaired *IRF7* gene are at risk of severe respiratory viral infections [90].

The study also showed that there is a mutation in the Uncoordinated-119 (*UNC119*) gene. This gene encodes the UNC119 proteins which are multifunctional proteins and part of *UNC119* supergene family [91]. The *UNC119* is important for the T-lymphocytes activity through the T cell receptor [92]. A mutation in the *UNC119* gene causes idiopathic CD4 lymphopenia that is heterogeneous syndrome characterized by reproducible decreased count of the CD4 T-lymphocyte with no immunodeficiency causes [92]. The cells carrying this mutation have exhibited impaired response to stimulation of T cell receptor, with defect in the lymphocyte-specific kinase (Lck) leading to reduced T cell proliferation [92].

Next we also found mutation in the *TANK-binding kinase 1* (*TBK1*) which has important roles in innate immunity against viral infections [93]. *TBK1* is involved in activation of interferon regulatory factor (*IRF*) 3 and antiviral immune response [93]. The TBK1p.Ala300fs deletion mutation is likely pathogenic according to the ACMG Classification. The COVID-19 helicase non-structural protein 13 (NSP13) has an important role in immune evasion by binding to the TBK1 [94]. The NSP13 suppresses the production of type I IFN by targeting the degradation of the TBK1 [95]. The TBK1p.Ala300fs mutation may reduce the affinity of the COVID-19 NSP13 to TBK1 and thus the production of the type I IFN is not affected rendering the carriers of this mutation more resistant to COVID-19. Nevertheless, the effect of this mutation on the COVID-19 patient remain to be investigated in future research. Our results also showed that there a mutation (NM_000466.3:c.1108delA, p.Ile370fs) in the gene *PEX1*. This gene located in on chromosome 7q21.2 [96]. *PEX1* gene encodes a protein with a 143-kDa molecular weight and belongs to the AAA ATPase protein family. These ATPases are involved in different cellular processes [97]. The mutations in *PEX1* gene were associated with recessive inherited disorders for example the Zellweger syndrome spectrum or cerebrohepatorenal syndrome [96]. Patients with *PEX1* gene mutations can be susceptible to infections and respiratory diseases [97].

We also found mutation in the gene *ABCC6* located in chromosome 16p13.11. This gene encodes an ATP-binding cassette (ABC) transmembrane protein expressed primarily in hepatic and renal tissues [98,99]. This protein functions as transporter in HDL metabolism [98,99]. Mutations in this gene were reported to cause Pseudoxanthoma elasticum, an autosomal recessive disorder characterized by ectopic mineralization of skin, eye and arteries and increased risk to CAD [98–101].

We also encountered mutation (NM_000065.5:c.1879delG, p.Asp627fs) in a *C6* gene. *C6* gene encodes a protein that is a part of the membrane attack

complex (MAC) that disturbs the cell membrane leading to cell lysis. Gene variation in the *C6* gene was associated with recurrent bacterial and viral infections [102].

Furthermore, we found a mutation (NM_005373.3:c.317C>T, p.Pro106Leu) in the *MPL* gene. The human c-mpl gene (*MPL*) is from the hematopoietic receptor superfamily. This gene encodes the CD110 protein with 635-amino acid residues [103]. The mutations in *MPL* gene were previously associated with hematopoietic diseases [103]. The p.Pro106Leu mutation in the *MPL* gene causes familial thrombocytosis [104]. The platelets have important roles in the inflammatory response and profibrotic mechanisms [105]. In COVID-19, thrombocytosis (without neoplasm) is associated with good prognosis, but shows increased risk of venous thromboembolism [105].

In addition we found a mutation (NM_000243.3:c.2230G>T, p.Ala744Ser) in the *MEFV* gene. This gene codes for pyrin protein. Mutations in *MEFV* were associated with Familial Mediterranean Fever (FMF) [106]. The FMF is an auto-inflammatory disease characterized by recurrent episodes of hyperthermia, peritonitis, inflammation of the peritoneum, pleura and synovium, and complications of amyloidosis [106]. It has been reported that the FMF is manifested with multisystemic disorders with diverse clinical symptoms for example severe atopic disorder and recurrent infections of the respiratory system [107]. However, it has been reported that FMF patients perhaps are not susceptible to bad outcomes of COVID-19 [108,109]. Our result will support research on personalized medicine for COVID-19 and might help healthcare providers to anticipate the long-term COVID-19 sequelae in patients at risk due to their genetic makeup. Limitations of this study include the limited sample size further future large scale studies in different populations are warranted.

Conclusion

In summary, we have conducted a screen on the genome of the COVID-19 hospitalized cases. Result showed that these genes *HFE*, *CFH*, *CDH23*, *CTLA-4*, *TGFB1*, *CREBBP*, *EP300*, *HBB*, *IRF7*, and *UNC119* might be associated with susceptibility to coronavirus infection. Further large-scale studies in different populations are recommended to verify these findings. After being confirmed these results can be used in genetic testing for the identification and stratification of susceptible population for the programs of prevention and treatment of coronavirus infections.

Availability of Data and Materials

All data supporting the reported results can be found at the Prince Fahad Bin Sultan Chair for Biomedical Research, Faculty of Applied Medical Sciences, University of Tabuk, Tabuk, Saudi Arabia.

Author Contributions

Conceptualization: RM, IE, MAA, NAA, FHA, BAA, EIM, SKM, MSM, FJT, JA, SMA, MMM and MFU; methodology: RM, IE, MAA, NAA, FHA and BAA; formal analysis: RM, IE, EIM, SKM, MSM, FJT, JA, SMA, MMM and MFU; validation: RM, IE, MAA, NAA, FHA, BAA, EIM, SKM, MSM, FJT, JA, SMA and MFU; writing original draft of the manuscript: IE and RM; writing and editing of the manuscript: IE, RM, and MFU. All authors contributed significantly to editorial changes of important content. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

This study was conducted in accordance with international and national regulations and in accordance with the Declaration of Helsinki. The study was approved by the institutional ethics committees at the University of Bisha (Ref. no. UBCOM/H-06-BH-087(05/25)) and University of Tabuk. All the subjects provided written informed consent before their participation in the study.

Acknowledgment

We thank staff from the King Fahad Hospital, Tabuk city and King Abdullah Hospital, Bisha city for assistance in sample collection.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Pal M, Berhanu G, Desalegn C, Kandi V. Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update. Cureus. 2020; 12: e7423.

[2] Elrashdy F, Redwan EM, Uversky VN. Why COVID-19 Transmission Is More Efficient and Aggressive Than Viral Transmission in Previous Coronavirus Epidemics? Biomolecules. 2020; 10: 1312.

[3] da Silva Torres MK, Bichara CDA, de Almeida MDNDS, Vallinoto MC, Queiroz MAF, Vallinoto IMVC, *et al*. The Complexity of SARS-CoV-2 Infection and the COVID-19 Pandemic. Frontiers in Microbiology. 2022; 13: 789882.

[4] Alexandersen S, Chamings A, Bhatta TR. SARS-CoV-2 genomic and subgenomic RNAs in diagnostic samples are not an indicator of active replication. Nature Communications. 2020; 11: 6059.

[5] da Rosa Mesquita R, Francelino Silva Junior LC, Santos Santana FM, Farias de Oliveira T, Campos Alcântara R, Monteiro Arnozo G, *et al*. Clinical manifestations of COVID-19 in the general population: systematic review. Wiener Klinische Wochenschrift. 2021; 133: 377–382.

[6] Baj J, Karakuła-Juchnowicz H, Teresiński G, Buszewicz G, Ciesielka M, Sitarz R, *et al*. COVID-19: Specific and Non-Specific Clinical Manifestations and Symptoms: The Current State of Knowledge. Journal of Clinical Medicine. 2020; 9: 1753.

[7] Hu J, Li C, Wang S, Li T, Zhang H. Genetic variants are identified to increase risk of COVID-19 related mortality from UK Biobank data. Human Genomics. 2021; 15: 10.

[8] Perveen S, Orfali R, Azam MSU, Aati HY, Bukhari K, Bukhari SI, *et al*. Coronavirus nCOVID-19: A pandemic disease and the Saudi precautions. Saudi Pharmaceutical Journal: SPJ: the Official Publication of the Saudi Pharmaceutical Society. 2020; 28: 888–897.

[9] Anastassopoulou C, Gkizarioti Z, Patrinos GP, Tsakris A. Human genetic factors associated with susceptibility to SARS-CoV-2 infection and COVID-19 disease severity. Human Genomics. 2020; 14: 40.

[10] DeForest N, Majithia AR. Genetics of Type 2 Diabetes: Implications from Large-Scale Studies. Current Diabetes Reports. 2022; 22: 227–235.

[11] Mousa M, Vurivi H, Kannout H, Uddin M, Alkaabi N, Mahboub B, *et al*. Genome-wide association study of hospitalized COVID-19 patients in the United Arab Emirates. EBioMedicine. 2021; 74: 103695.

[12] Tcheandjieu C, Zhu X, Hilliard AT, Clarke SL, Napolioni V, Ma S, *et al*. Large-scale genome-wide association study of coronary artery disease in genetically diverse populations. Nature Medicine. 2022; 28: 1679–1692.

[13] Alzahrani OR, Mir R, Alatwi HE, Hawsawi YM, Alharbi AA, Alessa AH, *et al*. Potential Impact of PI3K-AKT Signaling Pathway Genes, KLF-14, MDM4, miRNAs 27a, miRNA-196a Genetic Alterations in the Predisposition and Progression of Breast Cancer Patients. Cancers. 2023; 15: 1281.

[14] Elfaki I, Mir R, Almutairi FM, Duhier FMA. Cytochrome P450: Polymorphisms and Roles in Cancer, Diabetes and Atherosclerosis. Asian Pacific Journal of Cancer Prevention: APJCP. 2018; 19: 2057–2070.

[15] Jha CK, Mir R, Elfaki I, Javid J, Babakr AT, Banu S, *et al*. Evaluation of the Association of Omentin 1 rs2274907 A>T and rs2274908 G>A Gene Polymorphisms with Coronary Artery Disease in Indian Population: A Case Control Study. Journal of Personalized Medicine. 2019; 9: 30.

[16] Almassabi RF, Mir R, Javid J, AbuDuhier FM, Almotairi R, Alhelali MH, *et al*. Differential Expression of Serum Proinflammatory Cytokine TNF-α and Genetic Determinants of TNF-α, CYP2C19*17, miR-423 Genes and Their Effect on Coronary Artery Disease Predisposition and Progression. Life (Basel, Switzerland). 2023; 13: 2142.

[17] Alzahrani OR, Alatwi HE, Alharbi AA, Alessa AH, Al-Amer OM, Alanazi AFR, *et al*. Identification and Characterization of Novel Mutations in Chronic Kidney Disease (CKD) and Autosomal Dominant Polycystic Kidney Disease (ADPKD) in Saudi Subjects by Whole-Exome Sequencing. Medicina (Kaunas, Lithuania). 2022; 58: 1657.

[18] Ferreira LC, Gomes CEM, Rodrigues-Neto JF, Jeronimo SMB. Genome-wide association studies of COVID-19: Connecting the dots. Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases. 2022; 106: 105379.

[19] Elangeeb ME, Elfaki I, Eleragi AMS, Ahmed EM, Mir R, Alzahrani SM, *et al*. Molecular Dynamics Simulation of Kir6.2 Variants Reveals Potential Association with Diabetes Mellitus. Molecules (Basel, Switzerland). 2024; 29: 1904.

[20] O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, *et al*. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Research. 2016; 44: D733–D745.

[21] Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, *et al*. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Research. 2014; 42: D980–D985.

[22] Fairley S, Lowy-Gallego E, Perry E, Flicek P. The International Genome Sample Resource (IGSR) collection of open human genomic variation resources. Nucleic Acids Research. 2020; 48: D941–D947.

[23] Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, *et al*. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020; 581: 434–443.

[24] Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, *et al*. A method and server for predicting damaging missense mutations. Nature Methods. 2010; 7: 248–249.

[25] Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Research. 2001; 11: 863–874.

[26] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, *et al*. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine: Official Journal of the American College of Medical Genetics. 2015; 17: 405–424.

[27] Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, *et al*. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Research. 2011; 39: D561–D568.

[28] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research. 2003; 13: 2498–2504.

[29] Martin AR, Williams E, Foulger RE, Leigh S, Daugherty LC, Niblock O, *et al*. PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. Nature Genetics. 2019; 51: 1560–1565.

[30] Kolberg L, Raudvere U, Kuzmin I, Adler P, Vilo J, Peterson H. g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Research. 2023; 51: W207–W212.

[31] McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, *et al*. The Ensembl Variant Effect Predictor. Genome Biology. 2016; 17: 122.

[32] Hu J, Ng PC. SIFT Indel: predictions for the functional effects of amino acid insertions/deletions in proteins. PloS One. 2013; 8: e77940.

[33] Ghosh R, Harrison SM, Rehm HL, Plon SE, Biesecker LG, ClinGen Sequence Variant Interpretation Working Group. Updated recommendation for the benign stand-alone ACMG/AMP criterion. Human Mutation. 2018; 39: 1525–1530.

[34] Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics (Oxford, England). 2015; 31: 2745–2747.

[35] Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Research. 2019; 47: D886–D894.

[36] Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics (Oxford, England). 2015; 31: 761–763.

[37] Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, *et al*. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. American Journal of Human Genetics. 2016; 99: 877–885.

[38] Wang Q, Dhindsa RS, Carss K, Harper AR, Nag A, Tachmazidou I, *et al*. Rare variant contribution to human disease in 281,104 UK Biobank exomes. Nature. 2021; 597: 527–532.

[39] Ragia G, Manolopoulos VG. Assessing COVID-19 susceptibility through analysis of the genetic and epigenetic diversity of

ACE2-mediated SARS-CoV-2 entry. Pharmacogenomics. 2020; 21: 1311–1329.
[40] Roberts R, Chang CC, Hadley T. Genetic Risk Stratification: A Paradigm Shift in Prevention of Coronary Artery Disease. JACC. Basic to Translational Science. 2021; 6: 287–304.
[41] Barton JC, Edwards CQ, Acton RT. HFE gene: Structure, function, mutations, and associated iron abnormalities. Gene. 2015; 574: 179–192.
[42] Campos WN, Massaro JD, Martinelli ALC, Halliwell JA, Marsh SGE, Mendes-Junior CT, *et al*. HFE gene polymorphism defined by sequence-based typing of the Brazilian population and a standardized nomenclature for HFE allele sequences. HLA. 2017; 90: 238–242.
[43] Kowdley KV, Gochanour EM, Sundaram V, Shah RA, Handa P. Hepcidin Signaling in Health and Disease: Ironing Out the Details. Hepatology Communications. 2021; 5: 723–735.
[44] Ganz T. Hepcidin. [Rinsho Ketsueki] the Japanese Journal of Clinical Hematology. 2016; 57: 1913–1917.
[45] Sangkhae V, Nemeth E. Regulation of the Iron Homeostatic Hormone Hepcidin. Advances in Nutrition (Bethesda, Md.). 2017; 8: 126–136.
[46] Kisgeropoulos EC, Artz JH, Blahut M, Peters JW, King PW, Mulder DW. Properties of the iron-sulfur cluster electron transfer relay in an [FeFe]-hydrogenase that is tuned for H_2 oxidation catalysis. The Journal of Biological Chemistry. 2024; 300: 107292.
[47] Hirota K. An intimate crosstalk between iron homeostasis and oxygen metabolism regulated by the hypoxia-inducible factors (HIFs). Free Radical Biology & Medicine. 2019; 133: 118–129.
[48] Hollerer I, Bachmann A, Muckenthaler MU. Pathophysiological consequences and benefits of *HFE* mutations: 20 years of research. Haematologica. 2017; 102: 809–817.
[49] Aronow WS. Management of cardiac hemochromatosis. Archives of Medical Science: AMS. 2018; 14: 560–568.
[50] Cornelissen A, Guo L, Sakamoto A, Virmani R, Finn AV. New insights into the role of iron in inflammation and atherosclerosis. EBioMedicine. 2019; 47: 598–606.
[51] Hubacek JA, Philipp T, Adamkova V, Majek O, Dusek L. A haemochromatosis-causing HFE mutation is associated with SARS-CoV-2 susceptibility in the Czech population. Clinica Chimica Acta; International Journal of Clinical Chemistry. 2023; 538: 211–215.
[52] Ristić S, Milić S, Tatalović T, Bilobrk M, Rončević D, Ćurko-Cofek B, *et al*. The Influence of Hemochromatosis Gene *(HFE)* Mutations on SARS-CoV-2 Susceptibility and COVID-19 Severity. Balkan Medical Journal. 2023; 40: 229–231.
[53] Reuben A, Chung JW, Lapointe R, Santos MM. The hemochromatosis protein HFE 20 years later: An emerging role in antigen presentation and in the immune system. Immunity, Inflammation and Disease. 2017; 5: 218–232.
[54] Riley MJ, Hicks SR, Irvine S, Blanchard TJ, Britton E, Shawki H, *et al*. Hereditary haemochromatosis, haemophagocytic lymphohistiocytosis and COVID-19. Clinical Infection in Practice. 2020; 7: 100052.
[55] Gong S, Ma R, Zhu T, Ge X, Xie R, Tao Q, *et al*. Elevated serum beta-2 microglobulin level predicts short-term poor prognosis of patients with *de novo* acute omicron variant COVID-19 infection. Frontiers in Cellular and Infection Microbiology. 2023; 13: 1204326.
[56] Nersisyan SA. Induction of Hypoxic Response in Caco-2 Cells Promote the Expression of Genes Involved in SARS-CoV-2 Endocytosis and Transcytosis. Doklady. Biochemistry and Biophysics. 2022; 506: 206–209.
[57] Carapito R, Li R, Helms J, Carapito C, Gujja S, Rolli V, *et al*. Identification of driver genes for critical forms of COVID-19 in a deeply phenotyped young patient cohort. Science Translational Medicine. 2022; 14: eabj7521.
[58] Preston AE, Drakesmith H, Frost JN. Adaptive immunity and vaccination - iron in the spotlight. Immunotherapy Advances. 2021; 1: ltab007.
[59] Cronin SJF, Woolf CJ, Weiss G, Penninger JM. The Role of Iron Regulation in Immunometabolism and Immune-Related Disease. Frontiers in Molecular Biosciences. 2019; 6: 116.
[60] Tomatsu S, Orii KO, Fleming RE, Holden CC, Waheed A, Britton RS, *et al*. Contribution of the H63D mutation in HFE to murine hereditary hemochromatosis. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100: 15788–15793.
[61] Parente R, Clark SJ, Inforzato A, Day AJ. Complement factor H in host defense and immune evasion. Cellular and Molecular Life Sciences: CMLS. 2017; 74: 1605–1624.
[62] Li J, Long D, Wu S, Wu X, Wei B, Chen D, *et al*. Association of CFH polymorphism with susceptibility to sepsis caused by Pseudomonas aeruginosa in Chinese Han populations: A multicenter study. Gene. 2020; 722: 144127.
[63] Pastor AF, Rodrigues Moura L, Neto JWD, Nascimento EJM, Calzavara-Silva CE, Gomes ALV, *et al*. Complement factor H gene (CFH) polymorphisms C-257T, G257A and haplotypes are associated with protection against severe dengue phenotype, possible related with high CFH expression. Human Immunology. 2013; 74: 1225–1230.
[64] Kopp A, Hebecker M, Svobodová E, Józsi M. Factor h: a complement regulator in health and disease, and a mediator of cellular interactions. Biomolecules. 2012; 2: 46–75.
[65] Bottazzi B, Doni A, Garlanda C, Mantovani A. An integrated view of humoral innate immunity: pentraxins as a paradigm. Annual Review of Immunology. 2010; 28: 157–183.
[66] Moore SR, Menon SS, Cortes C, Ferreira VP. Hijacking Factor H for Complement Immune Evasion. Frontiers in Immunology. 2021; 12: 602277.
[67] Yu J, Gerber GF, Chen H, Yuan X, Chaturvedi S, Braunstein EM, *et al*. Complement dysregulation is associated with severe COVID-19 illness. Haematologica. 2022; 107: 1095–1105.
[68] Shcherbak SG, Changalidi AI, Barbitoff YA, Anisenkova AY, Mosenko SV, Asaulenko ZP, *et al*. Identification of Genetic Risk Factors of Severe COVID-19 Using Extensive Phenotypic Data: A Proof-of-Concept Study in a Cohort of Russian Patients. Genes. 2022; 13: 534.
[69] Ramzan K, Al-Numair NS, Al-Ageel S, Elbaik L, Sakati N, Al-Hazzaa SAF, *et al*. Identification of Novel *CDH23* Variants Causing Moderate to Profound Progressive Nonsyndromic Hearing Loss. Genes. 2020; 11: 1474.
[70] Lin TW, Hu YC, Yang YH, Chien YH, Lee NC, Yu HH, *et al*. CTLA-4 gene mutation and multiple sclerosis: A case report and literature review. Journal of Microbiology, Immunology, and Infection. 2022; 55: 545–548.
[71] Irfan M, Iqbal T, Hashmi S, Ghani U, Bhatti A. Insilico prediction and functional analysis of nonsynonymous SNPs in human CTLA4 gene. Scientific Reports. 2022; 12: 20441.
[72] Schwab C, Gabrysch A, Olbrich P, Patiño V, Warnatz K, Wolff D, *et al*. Phenotype, penetrance, and treatment of 133 cytotoxic T-lymphocyte antigen 4-insufficient subjects. The Journal of Allergy and Clinical Immunology. 2018; 142: 1932–1946.
[73] Hoffman TW, Leavis HL, Smits BM, van der Veken LT, van Kessel DA. Prolonged Disease Course of COVID-19 in a Patient with CTLA-4 Haploinsufficiency. Case Reports in Immunology. 2023; 2023: 3977739.
[74] Mirsharif ES, Rostamian A, Salehi M, Askari N, Ghazanfari T. Cytotoxic T-lymphocyte-associated antigen 4 *(CTLA-4)* +49A>G (rs231775) gene polymorphism is not associated with COVID-19 severity and mortality in an Iranian population. Heliyon. 2023; 10: e23308.
[75] Wodziński D, Wosiak A, Pietrzak J, Świechowski R, Kordek R, Balcerczak E. Assessment of the TGFB1 gene expression and

methylation status of the promoter region in patients with colorectal cancer. Scientific Reports. 2022; 12: 11488.
[76] Sanjabi S, Oh SA, Li MO. Regulation of the Immune Response by TGF-β: From Conception to Autoimmunity and Infection. Cold Spring Harbor Perspectives in Biology. 2017; 9: a022236.
[77] Ferreira-Gomes M, Kruglov A, Durek P, Heinrich F, Tizian C, Heinz GA, *et al*. SARS-CoV-2 in severe COVID-19 induces a TGF-β-dominated chronic immune response that does not target itself. Nature Communications. 2021; 12: 1961.
[78] Arguinchona LM, Zagona-Prizio C, Joyce ME, Chan ED, Maloney JP. Microvascular significance of TGF-β axis activation in COVID-19. Frontiers in Cardiovascular Medicine. 2023; 9: 1054690.
[79] Sadeghi H, Esmkhani S, Pirjani R, Amin-Beidokhti M, Gholami M, Azizi Tabesh G, *et al*. CREB-binding protein (CREBBP) and preeclampsia: a new promising target gene. Molecular Biology Reports. 2021; 48: 2117–2122.
[80] Hashwah H, Schmid CA, Kasser S, Bertram K, Stelling A, Manz MG, *et al*. Inactivation of CREBBP expands the germinal center B cell compartment, down-regulates MHCII expression and promotes DLBCL growth. Proceedings of the National Academy of Sciences of the United States of America. 2017; 114: 9701–9706.
[81] Torres LC, Kulikowski LD, Ramos PL, Sugayama SMM, Moreira-Filho CA, Carneiro-Sampaio M. Disruption of the CREBBP gene and decreased expression of CREB, NFκB p65, c-JUN, c-FOS, BCL2 and c-MYC suggest immune dysregulation. Human Immunology. 2013; 74: 911–915.
[82] Barreto-Galvez A, Niljikar M, Gagliardi J, Zhang R, Kumar V, Juruwala A, *et al*. Acetyl transferase EP300 deficiency leads to chronic replication stress mediated by defective fork protection at stalled replication forks. bioRxiv. 2023. (preprint)
[83] Carlice-Dos-Reis T, Viana J, Moreira FC, Cardoso GDL, Guerreiro J, Santos S, *et al*. Investigation of mutations in the HBB gene using the 1,000 genomes database. PloS One. 2017; 12: e0174637.
[84] Saad HKM, Taib WRW, Ab Ghani AS, Ismail I, Al-Rawashde FA, Almajali B, *et al*. HBB Gene Mutations and Their Pathological Impacts on HbE/β-Thalassaemia in Kuala Terengganu, Malaysia. Diagnostics (Basel, Switzerland). 2023; 13: 1247.
[85] Umar Z, Ilyas U, Nso N. Sickle Cell Disease and COVID-19 Infection: Importance of COVID-19 Testing and Approach to Management. Cureus. 2022; 14: e23604.
[86] Ning S, Pagano JS, Barber GN. IRF7: activation, regulation, modification and function. Genes and Immunity. 2011; 12: 399–414.
[87] Qing F, Liu Z. Interferon regulatory factor 7 in inflammation, cancer and infection. Frontiers in Immunology. 2023; 14: 1190841.
[88] Ciancanelli MJ, Huang SXL, Luthra P, Garner H, Itan Y, Volpi S, *et al*. Infectious disease. Life-threatening influenza and impaired interferon amplification in human IRF7 deficiency. Science (New York, N.Y.). 2015; 348: 448–453.
[89] Sodeifian F, Nikfarjam M, Kian N, Mohamed K, Rezaei N. The role of type I interferon in the treatment of COVID-19. Journal of Medical Virology. 2022; 94: 63–81.
[90] Campbell TM, Liu Z, Zhang Q, Moncada-Velez M, Covill LE, Zhang P, *et al*. Respiratory viral infections in otherwise healthy humans with inherited IRF7 deficiency. The Journal of Experimental Medicine. 2022; 219: e20220202.
[91] Constantine R, Zhang H, Gerstner CD, Frederick JM, Baehr W. Uncoordinated (UNC)119: coordinating the trafficking of myristoylated proteins. Vision Research. 2012; 75: 26–32.
[92] Gorska MM, Alam R. A mutation in the human Uncoordinated 119 gene impairs TCR signaling and is associated with CD4 lymphopenia. Blood. 2012; 119: 1399–1406.
[93] Zhao W. Negative regulation of TBK1-mediated antiviral immunity. FEBS Letters. 2013; 587: 542–548.
[94] Rashid F, Suleman M, Shah A, Dzakah EE, Chen S, Wang H, *et al*. Structural Analysis on the Severe Acute Respiratory Syndrome Coronavirus 2 Non-structural Protein 13 Mutants Revealed Altered Bonding Network With TANK Binding Kinase 1 to Evade Host Immune System. Frontiers in Microbiology. 2021; 12: 789062.
[95] Sui C, Xiao T, Zhang S, Zeng H, Zheng Y, Liu B, *et al*. SARS-CoV-2 NSP13 Inhibits Type I IFN Production by Degradation of TBK1 via p62-Dependent Selective Autophagy. Journal of Immunology (Baltimore, Md.: 1950). 2022; 208: 753–761.
[96] Bousfiha A, Bakhchane A, Charoute H, Riahi Z, Snoussi K, Rouba H, *et al*. A novel *PEX1* mutation in a Moroccan family with Zellweger spectrum disorders. Human Genome Variation. 2017; 4: 17009.
[97] Waterham HR, Ebberink MS. Genetics and molecular basis of human peroxisome biogenesis disorders. Biochimica et Biophysica Acta. 2012; 1822: 1430–1441.
[98] Moitra K, Garcia S, Jaldin M, Etoundi C, Cooper D, Roland A, *et al*. ABCC6 and Pseudoxanthoma Elasticum: The Face of a Rare Disease from Genetics to Advocacy. International Journal of Molecular Sciences. 2017; 18: 1488.
[99] Contrò G, Tallerico R, Dattilo V, Fabiani F, Enzo MV, Hladnik U, *et al*. A novel *ABCC6* variant causative of pseudoxanthoma elasticum. Human Genome Variation. 2019; 6: 30.
[100] Ibold B, Tiemann J, Faust I, Ceglarek U, Dittrich J, Gorgels TGMF, *et al*. Genetic deletion of Abcc6 disturbs cholesterol homeostasis in mice. Scientific Reports. 2021; 11: 2137.
[101] Le Saux O, Martin L, Aherrahrou Z, Leftheriotis G, Váradi A, Brampton CN. The molecular and physiological roles of ABCC6: more than meets the eye. Frontiers in Genetics. 2012; 3: 289.
[102] Westra D, Kurvers RAJ, van den Heuvel LP, Würzner R, Hoppenreijs EPAH, van der Flier M, *et al*. Compound heterozygous mutations in the C6 gene of a child with recurrent infections. Molecular Immunology. 2014; 58: 201–205.
[103] He X, Chen Z, Jiang Y, Qiu X, Zhao X. Different mutations of the human c-mpl gene indicate distinct haematopoietic diseases. Journal of Hematology & Oncology. 2013; 6: 11.
[104] El-Harith EHA, Roesl C, Ballmaier M, Germeshausen M, Frye-Boukhriss H, von Neuhoff N, *et al*. Familial thrombocytosis caused by the novel germ-line mutation p.Pro106Leu in the MPL gene. British Journal of Haematology. 2009; 144: 185–194.
[105] Lucijanic M, Krecak I, Soric E, Sedinic M, Sabljic A, Derek L, *et al*. Thrombocytosis in COVID-19 patients without myeloproliferative neoplasms is associated with better prognosis but higher rate of venous thromboembolism. Blood Cancer Journal. 2021; 11: 189.
[106] Moradian MM, Babikyan D, Banoian D, Hayrapetyan H, Manvelyan H, Avanesian N, *et al*. Comprehensive analysis of mutations in the MEFV gene reveal that the location and not the substitution type determines symptom severity in FMF. Molecular Genetics & Genomic Medicine. 2017; 5: 742–750.
[107] Celiksoy MH, Dogan C, Erturk B, Keskin E, Ada BS. The MEFV gene and its association with familial Mediterranean fever, severe atopy, and recurrent respiratory tract infections. Allergologia et Immunopathologia. 2020; 48: 430–440.
[108] Günendi Z, Yurdakul FG, Bodur H, Cengiz AK, Uçar Ü, Çay HF, *et al*. The impact of COVID-19 on familial Mediterranean fever: a nationwide study. Rheumatology International. 2021; 41: 1447–1455.
[109] Salehzadeh F, Pourfarzi F, Molatefi R, Davarnia B, Shahbazfar E, Ahmadabadi F. Immunogenic Potential of the Mediterranean Fever Gene in Patients with Coronavirus Disease: A Cross-Sectional Study. Iranian Journal of Medical Sciences. 2023; 48: 43–48.