European Review for Medical and Pharmacological Sciences 2023; 27: 20-32

Linking pathogenic and likely pathogenic gene variants to long-COVID symptoms

C. MICHELETTI[1], M.C. MEDORI[1], K. DHULI[1], P.E. MALTESE[1], S. CECCHIN[1], G. BONETTI[1,2], F. FIORETTI[3], L. ASSONI[3], A. CALZONI[3], A. PRADERIO[3], M.G. DE ANGELIS[3], K. DONATO[4,5], G. ARABIA[6], L. LORUSSO[7], P. MANGANOTTI[8], E. CAPELLI[9], G. MARCEDDU[4], M. BERTELLI[1,4,5], S. NODARI[3]

[1]MAGI'S LAB, Rovereto, Trento, Italy
[2]Department of Pharmaceutical Sciences, University of Perugia, Perugia, Italy
[3]Department of Medical and Surgical Specialities, Radiological Sciences and Public Health, University and Spedali Civili of Brescia, Brescia, Italy
[4]MAGI EUREGIO, Bolzano, Italy
[5]MAGISNAT, Atlanta Tech Park, Peachtree Corners, GA, USA
[6]Cardiology Unit, Department of Medical and Surgical Specialties, Radiological Sciences and Public Health, University of Brescia, Brescia, Italy
[7]UOC Neurology and Stroke Unit, ASST Lecco, Merate, Italy
[8]Clinical Unit of Neurology, Department of Medicine, Surgery and Health Sciences, Cattinara University Hospital ASUGI, University of Trieste, Trieste, Italy
[9]Department of Earth and Environmental Sciences and Centre for Health Technologies, University of Pavia, Pavia, Italy

C. Micheletti and M.C. Medori contributed equally to this work

Abstract. – OBJECTIVE: Long-COVID is a clinical syndrome characterized by the presence of symptoms related to SARS-CoV-2 infection that persist for at least four weeks after recovery from COVID-19. Genetics have been proposed to play an important role in long-COVID syndrome onset. This study aimed to identify genetic pathogenetic and likely pathogenetic causative variants of Mendelian genetic diseases in patients with Long-COVID syndrome. Additionally, we aimed to establish an association between these genetic variants and the clinical symptoms manifested during long-COVID syndrome.

PATIENTS AND METHODS: 95 patients affected by long-COVID syndrome were analyzed with a Next-Generation Sequencing (NGS) panel comprising 494 genes. The analyzed genes and the symptoms of the patients collected with an *ad-hoc* questionnaire were divided into four groups (cardiological, respiratory, immunological, and neurological). Finally, a statistical analysis comprising descriptive statistics, classification based on reported symptoms, and comparative analysis against a control group of healthy individuals was conducted.

RESULTS: 12 patients resulted positive for genetic testing with an autosomal dominance (8) or autosomal recessive (4) inheritance, showing a higher prevalence of cardiovascular genetic diseases (9) in the analyzed cohort compared to the normal population. Moreover, the onset of the long-COVID syndrome and its cardiovascular manifestations was compliant with the onset reported in the literature for the identified genetic diseases, suggesting that COVID-19 could manifest late-onset genetic diseases associated with their appearance. Apart from the 12 positive patients, 57 were healthy carriers of genetic diseases. Analyzing the whole cohort, a statistical correlation between prevalent symptomatology and the gene class was established, suggesting an association between the genetic susceptibility of an individual and the possibility of developing specific long-COVID syndrome symptoms, especially cardiovascular symptoms. Furthermore, 17 genetic variants were identified in CFTR. Finally, we identified genetic variants in IFNAR2 and POLG, supporting their respective involvement in inflammation and mitochondria mechanisms, correlated with long-COVID syndrome according to literature data.

CONCLUSIONS: This study proposed COVID-19 to act as a manifest of underlying late-onset genetic diseases Mendelian associated with carrier status. Moreover, according to our results, mutations in cardiological genes are more present in patients who show cardiological symptoms during the syndrome. This underscores the necessity for cardiological investigation and genetic screening in long-COVID patients to address existing or potential clinical implications.

Corresponding Author: Cecilia Micheletti, MSc; e-mail: cecilia.micheletti@assomagi.org

Key Words:
Long-COVID syndrome, Next-Generation Sequencing, NGS, Mendelian disease, Respiratory disease, Cardiac disease, Immunological disease, Neurological disease.

Introduction

The appearance of long-COVID syndrome and its impact on those recovering from COVID-19 creates serious concerns in the worldwide health community. Long-COVID syndrome is a collection of symptoms that continue after the acute phase of a SARS-CoV-2 infection has passed[1]. These lasting symptoms range from incapacitating weariness and anosmia to chronic shortness of breath and cognitive deficits[2]. Chronic fatigue, on the other hand, is largely defined by intense and chronic exhaustion that severely limits everyday activity. Though it can be triggered by a variety of factors, it has lately received significant attention due to its frequent relationship with long-COVID syndrome: understanding the complex interplay between these two conditions is crucial[3]. It is critical to understand the genetic association of long-COVID syndrome by delving into the complexity of both multifactorial and Mendelian genetics[4]. Our study investigates how neurological, cardiological, respiratory, and immunological symptoms manifested during long-COVID syndrome may be due to the presence of genetic variants in genes associated with Mendelian genetic diseases. We specifically look at the emergence of symptoms connected to long-COVID syndrome and the possible impact of genetic variations in genes linked to Mendelian diseases[5]. Our study focuses on genes connected to pathways for respiratory, cardiac, immune, and neurological diseases, all of which have been linked to the wide range of symptoms seen in long-COVID patients. To demonstrate the correlation between clinical symptoms and the genetic variants found, a statistical analysis was conducted. We seek to gain crucial knowledge about the biological foundations of long-COVID syndrome and open the door to more specialized therapy approaches by examining these genetic relationships.

Patients and Methods

Patient Recruitment

Between April 2021 and October 2022, 95 adult patients (51 women and 44 men; average age 52; youngest patient 28; oldest patient 81), who exhibited clinical symptoms and severe long-term COVID syndrome, were enrolled in the study. These symptoms covered a wide range, including exhaustion, fever, sleeplessness, mental disorientation, memory loss, anxiety, pain, blood pressure fluctuations, dizziness, and loss of balance. Patients were diagnosed at least six months after SARS-CoV-2 recovery. Blood samples were collected from each patient.

Exclusion criteria included: significant mental retardation, severe disabilities, or in any case subjects unable to provide informed consent; subjects with an oropharyngeal swab still testing positive for COVID-19 at the time of screening; subjects presenting negative serological from COVID-19; minors (subjects under 18 years of age); subjects suffering from pathologies that can alter the normal functioning of nervous system (e.g., serious heart disease, hereditary or acquired neuropathies, eating disorders, psychiatric disorders).

Informed Consent and Ethical Considerations

A written informed consent form was obtained from all patients at the time of enrollment, and each of them was assigned a unique alphanumeric code to protect their anonymity. All research processes were conducted according to the ethical guidelines of the Declaration of Helsinki and its latest amendments and were approved by the Ethical Committee of Azienda Sanitaria dell'Alto Adige, Italy (Approval No. 132-2020).

Data Collection and Symptoms Categorization

Patients were asked to rate their symptoms on a 0-10 scale to reflect the symptom intensity reported during Covid-19. With 0 denoting the lack of symptoms and 10 denoting the presence of severe symptoms. Four distinct domains of symptoms were evaluated, including those related to the cardiovascular system (e.g., arrhythmias, tachycardia), the respiratory system (e.g., rhinorrhea, dyspnea), the nervous system (e.g., tremors, sleep disorders), and the immune system (e.g., fever, diarrhea, skin manifestations). This classification made it possible to thoroughly and methodically examine the wide range of symptoms that our patient group reported. We also noted each patient's comorbidities. These comorbidities enabled us to investigate potential relationships between pre-existing medical illnesses and the long-COVID syndrome presentation.

Next-Generation Sequencing

We developed a Next-Generation Sequencing (NGS) panel comprising 494 genes potentially associated with chronic fatigue syndrome and related fatigue syndromes (**Supplementary Table I**). The selection of genes for this panel was informed by data from reputable sources, including the Human Gene Mutation Database (HGMD Professional), Online Mendelian Inheritance in Man (OMIM), Orphanet, GeneReviews, and PubMed.

DNA Probe Creation

To create the custom DNA probes for this panel, we utilized Twist Bioscience Technology (South San Francisco, CA, USA). Subsequently, Illumina NGS sequencing services were conducted by IntegraGen Genomic Service (Évry, Évry-Courcouronnes, Francia).

Targeted Sequencing and Variant Interpretation

The NGS panel targeted coding exons and adjacent regions within each gene, with a cumulative target length of 839,308 base pairs. The interpretation of identified variants was carried out using the already published classification algorithm[6-8]. Variants were interpreted according to the American College of Medical Genetics guidelines[9].

Data Analysis and Visualization

i. Bar graph for survey responses: we utilized a bar graph to visualize the distribution of survey responses across symptom categories.
ii. ANOVA test: an analysis of variance (ANOVA) was conducted to assess the relationship between gene class and minor allele frequency (MAF).
iii. Regression analysis: to evaluate the relationship between inheritance types and MAF, a regression analysis was performed.
iv. Violin graph for MAF: we created a violin graph to depict the distribution of MAF based on the verdict.
v. Chi-Square test for long-COVID syndrome: to investigate the connection between P and LP mutations in our population and the occurrence of long-COVID syndrome, a Chi-square test was employed.

These analytical tools and methodologies enabled us to conduct a comprehensive examination of the dataset, yielding valuable insights into the relationships among various variables of interest.

Statistical Analysis

For our analysis, we employed a combination of data preparation techniques. Initially, data organization and formatting were carried out in Microsoft Excel to prepare the dataset. Subsequently, we used R Studio, an advanced statistical analysis environment, to import and analyze the dataset. In our statistical analysis, we set a significance threshold at a p-value of less than 0.05.

Results

A comprehensive analysis of 494 genes was conducted in a cohort of 95 patients, utilizing a designed and validated panel developed during this study. Within this analysis, 75 variants were identified and subsequently classified as either pathogenic or likely pathogenic, employing the published interpretation algorithm[6-8]. These notable variants were observed in a total of 57 patients. Significantly, these variants were dispersed across 45 distinct genes, further elucidating the intricate genetic landscape associated with long-COVID syndrome. The genes harboring these variants include *ABCA3, ABCB1, ABCC9, APOB, AKR1C2, AKR1C4, ATP1A2, ATR, CD36, CFTR, CHRNE, CHRNB1, DSG2, DSP, DUOX2, F5, FBN1, FCGR2A, FYCO1, GPR18, GAL3ST2, GDPD1, GLA, GPBAR1, GRIN2D, IFNAR2, JSRP1, KCNA5, KCNN3, LDLR, LRBA, LRP6, MIB1, MYH6, NAT8, ORAI1, POLG, PON3, PROC, PYGM, SCN1B, SCN5A, SDHA, SERPINA1, ST6GALNAC5, TLR4, TRPV1, VAMP1, ZFHX3, TTN, VARS2*. The complete results are reported in **Supplementary Table II**.

Multifactorial Analysis

Figure 1, a bar chart was crafted to present the questionnaire responses about symptom categories. Neurological deficits emerge as the most prevalent, being documented in 87 cases, whereas cardiological deficits are comparatively less common, with 62 reported cases. Respiratory deficits were reported in 74 patients and immunological deficits in 73 patients. The figures also draw attention to the fact that the cardiological category registers the highest number of absent responses, followed sequentially by the immunological, respiratory, and neurological categories.

To determine whether there is a significant association between the presence of P and LP mutations in our population and the occurrence of long-COVID syndrome, we conducted the

chi-square test as a better analytical tool because it can be used with categorical data and can assess the presence of an association between variables.

The Chi-square test (Table I) was applied to our dataset and yielded significant and profound results.

- The Chi-square statistic was 260. This statistic measures the strength of the association between the presence of P and LP mutations and the onset of long-COVID syndrome.
- The test was conducted with 25 degrees of freedom, representing the number of independent comparisons made in the analysis.
- The *p*-value was approximately 5.70e-41, a very small number. This *p*-value represents the probability of obtaining a Chi-square statistic as extreme as that obtained if there was no association between the variables.

To better understand the complex relationship between genetics and allelic variation, we tapped into the ANOVA analytic tool (**Supplementary Table III**). Our main goal was to determine how different gene classes might affect the variation in 'Minor Allele Frequency' (MAF) values. The variance attributed to "Gene Class" unveils an astonishingly high *F*-value of 2.273251e+30, accompanied by an exceedingly minute *p*-value of 6.798358e-136.

In an attempt to understand the interplay between 'Inheritance' and 'Minor Allelic Frequency' (MAF) Patterns, we employed a linear regression analysis (**Supplementary Table IV**). We aimed to understand whether the specific types of "inheritance", classified as "AR" (Autosomal Recessive) and "AD" (Autosomal Dominant), had a discernible impact on the observed variations in allele frequency.

Analyzing the scatter diagram (**Supplementary Figure 1**), the x-axis labels the inheritance types 'AD' and 'AR'. A surprising pattern emerges: all instances of 'AD' have distinct MAF values, different from those of 'AR'.

The violin diagram (Figure 2) attempts to illuminate the potential connections between genetic verdicts and the distribution of allele frequencies within our dataset.

Mendelian Analysis

To illustrate the results of the Mendelian analysis, we have drafted 3 tables. Table II shows subjects with P or LP mutations are likely to be positive in genes with recessive or dominant inheritance. Subjects with two variants in recessive genes or one variant in dominant genes were therefore considered. The table also shows the clinical signs evident for each subject collected employing a questionnaire.

Table I. Conducting a Chi-Square Test Analysis to assess the correlation between the presence of P and LP mutations in our population and the incidence of long-COVID syndrome.

Statistical Analysis	Value
Chi-Squared Statistics	260
Degrees of Freedom	25
p-value	5.697e-41

Supplementary Table V shows 57 carrier subjects for Mendelian genetic disorders with autosomal recessive and X-linked inheritance. There are, therefore, subjects with a variant in heterozygosis for recessive genes and two female subjects with an X-linked variant. The table also shows the clinical signs evident for each subject.

Supplementary Table VI shows the frequencies for each gene in which we found P or LP variants. Three frequencies were then calculated for each gene: the population frequency according to our case history, the carrier frequency based on the population of our case history, and finally, the carrier frequency in the general population. The population frequency, according to our case history, was calculated as the number of subjects with P or LP variant of the gene/total number of subjects in our case history. The population-based carrier frequency of our case series was calculated according to the Hardy-Weinberg formula. The carrier frequency in the general population was calculated according to literature data. Where 0.1/100 is indicated, an estimate based on the definition of a rare disease is given.

Discussion

Multifactorial Analysis

Table I presents compelling evidence from the Chi-square test, conclusively demonstrating a profound and non-random correlation between the presence of P and LP mutations and the onset of long-COVID syndrome. This statistical validation forms a critical foundation for advancing future research initiatives, emphasizing the necessity to explore both the clinical and biological implications of long-COVID syndrome. With a remarkably low *p*-value of 5.70e-41, significantly below the conventional significance threshold of 0.05, we can hypnotize that there is an association between the presence of P and LP mutation and the development of long-COVID syndrome. This robust rejection of the null hypothesis firmly

Table II. Genes with AD or AR inheritance and P or LP Verdict. The patients with two variants for AR genes and one variant for AD genes are considered likely positive for Mendelian respiratory, cardiac, immunological, and neurological diseases.

ID	Gene	Gene Class and gene OMIM	Inheri-tance	Verdict	Respiratory deficit	Cardiological deficit	Neurological deficit	Immunological deficit	Comorbidity
1	ABCC9	Cardiology (OMIM: 601439)	AD	Likely Pathogenic	X		X	X	Cardiovascular deficit, dyslipidemia
1	ABCC9	Cardiology (OMIM: 601439)	AD	Likely Pathogenic	X		X	X	Cardiovascular deficit, dyslipidemia
2	CFTR	Respiratory (OMIM: 602421)	AR	Likely Pathogenic			X		Dyslipidemia
2	CFTR	Respiratory (OMIM: 602421)	AR	Pathogenic			X		Dyslipidemia
13	POLG	Neurology (OMIM: 174763)	AR	Likely Pathogenic		X	X	X	Dyslipidemia
13	POLG	Neurology (OMIM: 174763)	AR	Likely Pathogenic		X	X	X	Dyslipidemia
21	DSP	Cardiology (OMIM: 125647)	AR	Likely Pathogenic	X	X	X	X	
21	DSP	Cardiology (OMIM: 125647)	AR	Likely Pathogenic	X	X	X	X	
26	SCN1B	Cardiology (OMIM: 600235)	AD	Likely Pathogenic	X	X	X		Hypertension arteriosa
33	TTN	Cardiology (OMIM: 188840)	AD	Likely Pathogenic	X	X	X	X	Hypertension
37	KCNA5	Cardiology (OMIM: 176267)	AD	Likely Pathogenic	X	X	X	X	
40	MYH6	Cardiology (OMIM: 160710)	AD	Likely Pathogenic	X	X	X	X	Cardiac insufficiency, dyslipidemia, obesity
42	LRP6	Cardiology (OMIM: 603507)	AD	Likely Pathogenic	X		X	X	Cardiac insufficiency, hypertension, dyslipidemia
46	POLG	Neurology (OMIM: 174763)	AR	Pathogenic	X	X	X	X	
46	POLG	Neurology (OMIM: 174763)	AR	Pathogenic	X	X	X	X	
50	DSG2	Cardiology (OMIM: 125671)	AD	Likely Pathogenic			X		
53	DES	Cardiology (OMIM: 125660)	AD	Likely Pathogenic			X	X	Cardiac insufficiency

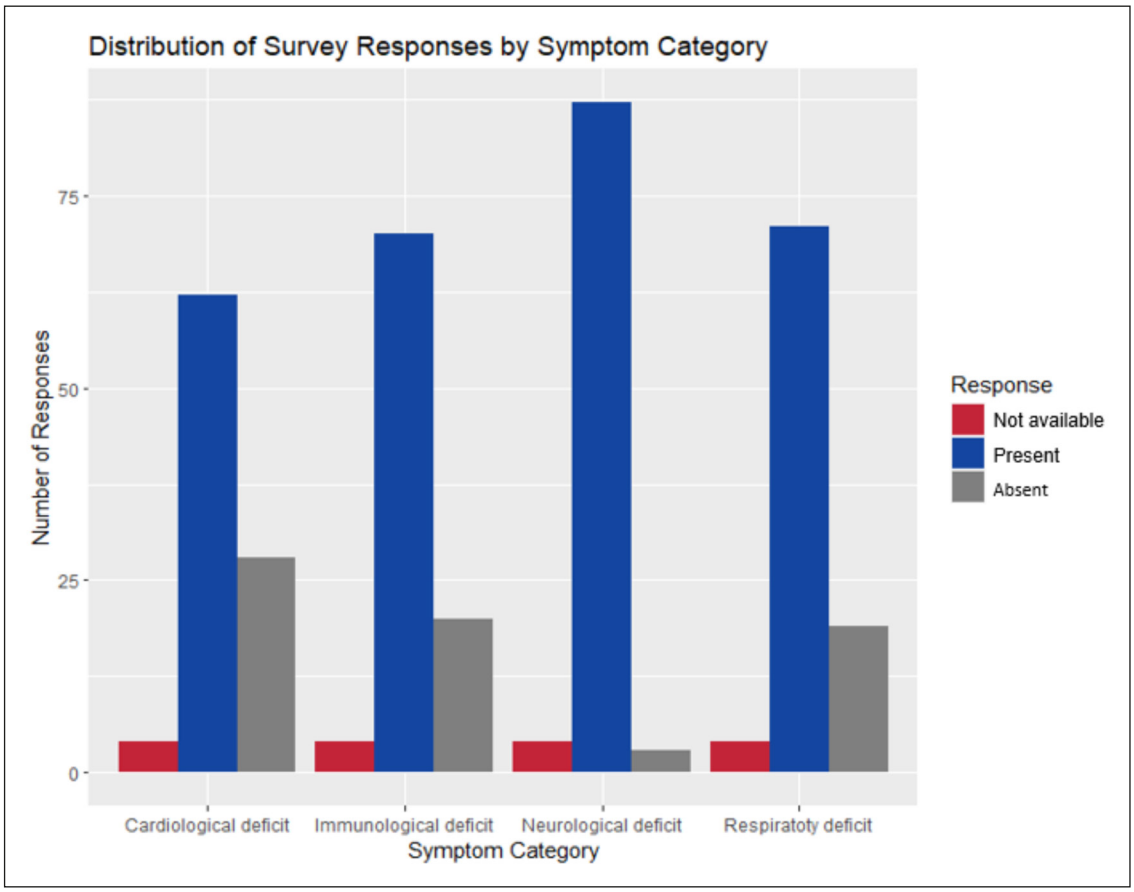

Figure 1. Creating a Bar Graph to illustrate the distribution of survey responses categorized by symptoms.

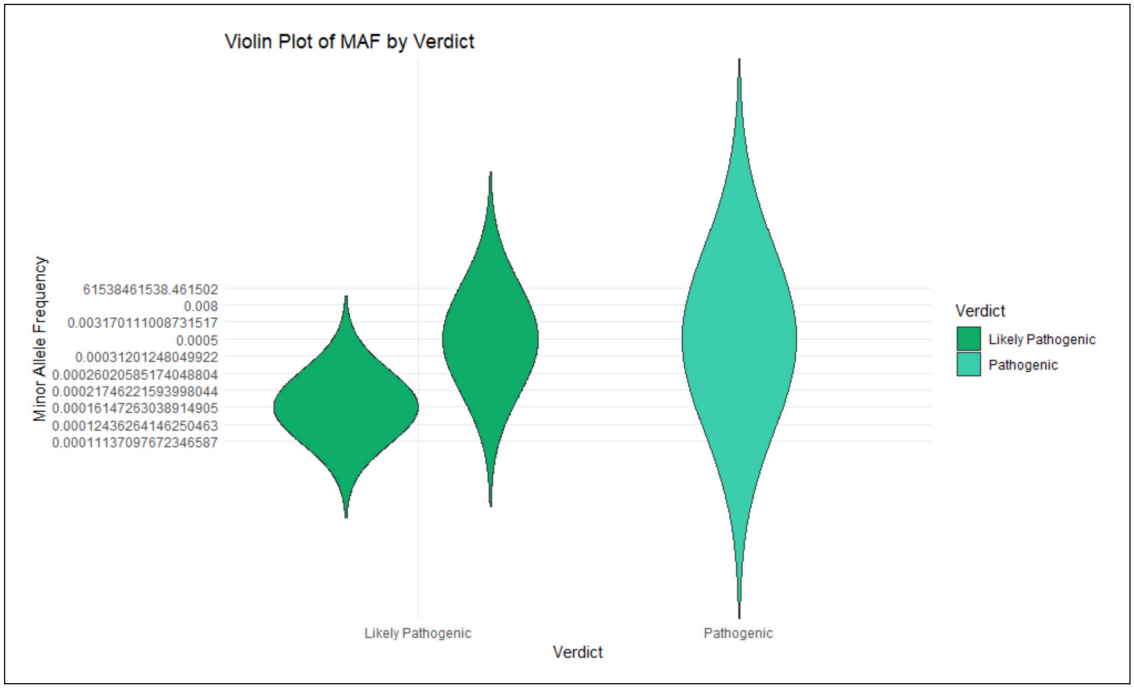

Figure 2. Utilizing a Violin Plot to elucidate potential links between genetic findings and the distribution of allele frequencies in our dataset.

establishes a non-random relationship between these variables. **Supplementary Table III** illustrates how different gene classes may influence the variation in Minor Allele Frequency (MAF) values. The insights are succinctly summarized in the ANOVA results: comparison of MAF across Gene Classes' table weaves a compelling narrative. The variance attributed to 'Gene Class' reveals a high *F*-value of 2.273251e+30, accompanied by an exceedingly small *p*-value of 6.798358e-136. These remarkable statistics collectively underscore a profound distinction in MAF values across diverse gene classes. The negligible residuals, characterized by minuscule sums of squares and mean squares, signal the substantial contribution of gene classes to the observed allelic variance. The remarkably low *p*-value serves as a testament to the pivotal role of gene classes in orchestrating MAF variation, suggesting unique genetic attributes that contribute to the observed heterogeneity in allele frequencies. Furthermore, the substantial *F*-value reinforces the notion that these observed disparities are not transient anomalies but rather indicative of substantial genetic influences. As indicated in Table I, these findings collectively underscore a profound distinction in Minor Allele Frequency (MAF) values across different gene classes. The presence of negligible residuals, characterized by small sums of squares and mean squares, highlights the substantial influence of gene classes on the observed allelic variance. The notably low *p*-value serves as a testament to the pivotal role played by gene classes in shaping MAF variation, suggesting the presence of unique genetic attributes that contribute to the observed heterogeneity in allele frequencies. Moreover, the substantial *F*-value reinforces the notion that these observed disparities are not transient anomalies but rather indicative of significant genetic influences.

Supplementary Table IV, titled 'Results of the Regression Analysis' presents intriguing findings. About the intercept term, the coefficient suggests a minor shift in the Minor Allele Frequency (MAF), represented as 1.915e-04, albeit accompanied by a substantial standard error of 4.986e+09. Notably, the corresponding *t*-value and *p*-value both underscore the insignificance of the intercept's contribution to the model explanation. In statistical terms, this implies that the intercept does not substantially contribute to explaining the model's variability. Turning our attention to 'InheritanceAR' the coefficient estimate stands at 1.758e+09, with a standard error of 5.263e+09. The associated *t*-value of 0.334 indicates a moderate deviation from the null hypothesis. However, the *p*-value of 0.74 accentuates the lack of statistical significance, suggesting that the observed variation in 'InheritanceAR' is not a random occurrence.

As illustrated in **Supplementary Figure 1**, the 'Inheritance' type profoundly impacts the landscape of allele frequencies. These distinct trends unequivocally underscore the pivotal role of genetic inheritance patterns in shaping the landscape of allelic diversity.

Upon close examination of the violin graph (Figure 2), noteworthy patterns come to the forefront. In the LP category, we observe two distinct violins. The first spans a wide MAF range, ranging from 0.000111 to 0.008, indicating significant variability in allele frequencies. In contrast, the second, distinctly higher violin also extends over the same MAF range but reaches a substantially higher peak of 61538461538.461502. This contrast underscores a substantial variability of allele frequencies within the LP category. Conversely, the 'Pathogen' category is represented by a single, dominant violin. This violin spans from the lowest MAF value to the highest point, indicating a more concentrated distribution and suggesting a potentially more coherent pattern of allele frequencies. These findings underscore the intricate relationship between genetic verdicts and allele frequencies. The differing shapes of the violins within the 'Likely Pathogenic' category suggest varying genetic dynamics that may influence the prediction of pathogenicity. Furthermore, the prominence of the single high violin in the 'Pathogenic' category hints at a distinct genetic signature, warranting further in-depth investigation.

Mendelian Analysis

In 57 patients, the genetic analysis unveiled 75 Pathogenic or Likely Pathogenic variants in genes linked to Mendelian genetic illnesses (**Supplementary Table II**). We divided the reported symptoms of the patients and the genes into four different categories – respiratory, cardiovascular, neurological, and immunological deficit – to thoroughly assess the effects of these genetic variations.

Table II lists a group of 12 patients with one P or LP variant in a gene with autosomal dominance inheritance or two variants in a gene with autosomal recessive inheritance who are most likely to have latent Mendelian genetic disorders that manifested as clinical symptoms after COVID-19 infection. In the context of long-COVID syndro-

me, it is possible that COVID-19 infection played a role in the development and clinical manifestation of several Mendelian genetic disorders. These genetic disorders may have been latent or had mild signs before the infection. COVID-19 may have catalyzed the discovery of these underlying genetic predispositions and their eventual manifestation as clinical symptoms because of its complex effects on the human body.

Supplementary Table VII presents patients identified as carriers of genes associated with Mendelian genetic diseases characterized by recessive inheritance patterns. These individuals carry a variant in genes that exhibit recessive inheritance, which means they are not afflicted by a Mendelian genetic disease. Notably, two patients were identified as carriers of variants in a gene with X-linked (XL) inheritance, as clinical data indicated that they were women. It is well-documented in the literature that women with variants in XL genes typically exhibit very mild symptoms or, in some cases, remain asymptomatic. Therefore, further clinical verification is warranted for these cases.

We also focused on investigating patients presenting cardiological symptoms. Predisposition to cardiological diseases carries significant clinical and practical implications, impacting both patient outcomes and management in clinical practice. Our analysis revealed the presence of Pathogenic (P) and Likely Pathogenic (LP) variants in 16 genes related to the cardiological pathway among 18 patients. Except for five patients, the remaining individuals reported cardiovascular deficits as their predominant symptom. In some cases, these cardiological symptoms were pre-existing conditions. Notably, 9 patients were classified as 'probably positive' as they possessed P or LP variants in genes with autosomal dominant transmission (*ABCC9, SCN1B, TTN, KCNA5, MYH6, LRP6, DSG2, DES*), or they exhibited two P or LP variants in genes with autosomal recessive transmission (*DSP*). In summary, among our total cohort of 95 patients, 25% presented P or LP variants in cardiological genes. Within this subgroup, 8% exhibited cardiological symptoms, some of which were pre-existing, and received a likely positive diagnosis for genetic diseases within the cardiological field, as depicted in **Supplementary Table II**.

For genes involved in cardiologic pathways, age at enrollment, symptoms manifested during long-COVID, and comorbidities were considered for each patient. Gene-related pathology and age of disease onset were also considered. The following are the considerations of our interest. In patient RX1037.2022 (ID 1), two dominant variants of the *ABCC9* gene emerged. This gene is associated with dilated cardiomyopathy and atrial fibrillation; for both diseases, transmission is autosomal dominant. Frameshift mutations in the *ABCC9* gene have been described in the literature in patients with dilated cardiomyopathy[10]. The age of onset of the disease is between 40 and 60 years (OMIM). The subject was 57 years old at the time of enrollment. Furthermore, he did not report cardiovascular deficit during the infection, but as a comorbidity.

In subject RX1081.2022 (ID 26), a dominant variant in the *SCN1B* gene emerged. The gene is associated in OMIM with atrial fibrillation with an onset age between 25 and 45 years. The subject was 60 years old at the time of enrolment and manifested cardiological deficits during COVID-19. The subject also reported hypertension as a comorbidity[11].

In the case of subject RX1103.2022 (ID 37), a variant with autosomal dominant inheritance was identified in the *KCNA5* gene. This gene is linked to atrial fibrillation, typically manifesting between the ages of 25 and 45. However, it is important to note that the subject was 75 years old at the time of enrollment and reported cardiovascular deficits[12].

A variant with autosomal dominant inheritance in the *TTN* gene was found in subject RX1092.2022 (ID 33). The gene is associated with dilated cardiomyopathy, which occurs up to the age of 50. The subject was 53 years old at the time of enrolment. The patient also reported cardiovascular deficits, such as hypertension as a comorbidity[13].

A variant with autosomal dominant inheritance in the *MYH6* gene was found in subject RX1133.2022 (ID 40). The gene is associated with dilated cardiomyopathy, which occurs up to the age of 50. The subject was 53 years old at the time of enrolment. The patient also reported a cardiovascular deficit. In the medical record, heart failure, dyslipidemia, and obesity were reported as comorbidities[14].

A variant with autosomal dominant inheritance in the *LRP6* gene was found in subject RX1135.2022 (ID 42). The gene is associated with coronary artery disease, which occurs over the age of 45. The subject was 49 years old at the time of enrolment. The patient reported no cardiovascular deficit. In the medical record, heart failure, hypertension, and obesity were reported as comorbidities[11].

A variant with autosomal dominant inheritance in the *DSG2* gene was found in subject RX1150.2022 (ID 50). The gene is associated

with arrhythmogenic right ventricular dysplasia. The subject was 73 years old at the time of enrolment. The patient reported no cardiovascular deficit and no comorbidities[15].

In subject RX1164.2022 (ID 54) a variant with autosomal dominant inheritance was found in the DES gene. The gene is associated with dilated cardiomyopathy, which occurs up to the age of 50. The subject was 55 years old at the time of enrolment. The patient also did not report cardiovascular deficit during infection but as a comorbidity[14]. Considering each subject's age, the age of onset of symptoms, and the reported comorbidities, COVID-19 may have triggered a cardiological disease from which those subjects were already suffering. It is also advisable to investigate the patients' course through genetic testing.

Below are some considerations made for particularly relevant genes and of interest to us. In the patient RX1037.2022 (ID 1), a single-copy variation in the *IFNAR2* gene that is linked to recessive inheritance was found. One of the two chains of the receptor responsible for identifying interferon alpha and beta is formed by a type I membrane protein that is encoded by the *IFNAR2* gene. Even though it has generally been linked to immunodeficiency, recent research shows its potential relevance in the context of long-COVID syndrome. According to previous research[11,12], those who have *IFNAR2* gene variations may experience more severe symptoms after contracting COVID-19, which would add to the complexity of long-COVID syndrome presentations. *IFNAR2*'s importance stems from its function in the innate immune system, namely in the response to viral infections. This receptor's activation sets off a cascade of immunological and antiviral responses. Genetic variations in *IFNAR2*, such as those linked to COVID-19 severity, may influence the expression and function of the receptor, impacting the body's capacity to fight viral infections like SARS-CoV-2. Additionally, the investigation goes beyond just genetic variations. It explores soluble *IFNAR2* (*sIFNAR2*) plasma levels, an important factor in controlling the interferon response. The observed decreased levels of *sIFNAR2* in those with severe COVID-19 may indicate a dysregulated interferon response, which would relate the genetic and plasma-level characteristics of *IFNAR2* to the severity of long-COVID syndrome[11,12].

Two autosomal recessive variants within the *CFTR* gene were identified in patient RX1038.2022 (ID 2). The *CFTR* gene is responsible for cystic fibrosis and follows an autosomal recessive inheritance pattern. Interestingly, despite carrying these variants, the subject exhibited no respiratory deficits. It is important to note that the Minor Allele Frequency (MAF) values for both variants are 0.009 and 0.007, respectively. These frequencies are notably high, which could potentially indicate a false positive result. This finding aligns with the absence of respiratory deficits and comorbidity reported by the subject[13].

In our case series, 11 dominant and recessive variants in the *POLG* gene were found in 9 subjects. Subject RX1060.2022 (ID 14) has two autosomal recessive variants in the *POLG* gene, as does subject RX1141.2022 (ID 47).

The *POLG* gene in OMIM is associated with Mitochondrial DNA Depletion syndrome 4A, Mitochondrial DNA Depletion syndrome 4B, Mitochondrial Recessive Ataxia syndrome, and Progressive External Ophthalmoplegia: each syndrome has an autosomal recessive inheritance pattern. In addition, the gene is associated with Progressive External Ophthalmoplegia with autosomal dominant inheritance[16]. While there is limited evidence linking the *POLG* gene to long-COVID syndrome, recent insights suggest that mitochondria may play a pivotal role in the host's response to viral infections and immune reactions. Notably, a recent study[17] has detected the presence of the SARS-CoV-2 genome within the host cell's mitochondria. The mitochondrion assumes a central role in orchestrating cellular responses during stress, including its role in inducing the formation of vesicles that facilitate communication between the mitochondrion and the endoplasmic reticulum (ER). In the context of viral infections, the replication of coronaviruses involves the formation of double-membrane vesicles derived from the ER. These vesicles serve as sites for viral replication while aiding in the evasion of the host's cellular defenses[17].

This intricate interplay between mitochondria and the ER highlights the multifaceted nature of host-virus interactions and may shed light on the mechanisms underlying long-COVID syndrome, even in genes like *POLG* with limited direct evidence of association.

Subject RX1060.2022 (ID 14) reported a cardiological deficit, neurological deficit, and immunological deficit. He also reported dyslipidemia as a comorbidity. Subject RX1141.2022 (ID 47) reported respiratory deficit, cardiovascular deficit, neurological deficit, and immunological deficit as a comorbidity.

Since mitochondrial diseases usually show immunological and neurological deficits, this could be compatible with what our patients reported. The symptoms of progressive external ophthalmoplegia have an age of onset in adulthood. Since patient RX1060.2022 (ID 13) is 63 years old, this could not be compatible with developing symptoms. On the other hand, patient RX1141.2022 (ID 46), being 43, could be compatible with the manifestation of the symptoms of external ophthalmoplegia.

In the case of subject RX 1071.2022 (ID 22), two recessive variants interpreted as Likely Pathogenic (LP) were identified within the *DSP* gene. The *DSP* gene is associated in OMIM with various conditions, including arrhythmogenic right ventricular dysplasia (AD), dilated cardiomyopathy with woolly hairs and keratoderma (AR), dilated cardiomyopathy with woolly hairs, keratoderma, and dental agenesis (AD), lethal acantholytic bullous epidermolysis (AR), and striated palmoplantar keratosis (AD). Typically, these diseases manifest symptoms until the age of 37. However, it is noteworthy that the subject was 43 years old at the time of enrollment and reported cardiological and immunological deficits, including brittle nails and hair loss. Given the patient's age, age of symptom onset, and the reported symptoms, it is likely that the COVID-19 event triggered the manifestation of the disease[15].

Two subjects (RX1084.2022, ID 9 and RX 1145.2022, ID 49) with a variant of the *GLA* gene with X-linked inheritance were also included. The subjects were found to be female and were therefore considered to be unaffected carriers[11]. A single-copy variant in the *IFNAR2* gene with recessive inheritance was detected in subject RX1037.2022 (ID 1). The protein encoded by this gene is a type I membrane protein that forms one of the two chains of an interferon alpha and beta receptor. In OMIM, it is associated with immunodeficiency, with autosomal recessive inheritance. Little evidence has emerged in the literature concerning the association between *IFNAR2* and post-COVID syndrome. However, this evidence shows that if an individual has variants in the *IFNAR2* gene, he or she manifests more severe symptoms in post-COVID syndrome[12,13]. In the subject RX1070.2022 (ID 21), a variant in the *ABCA3* gene and a variant in the *POLG* gene were found, respectively: both variants, present in a single copy, were interpreted as LPs, and the genes show recessive inheritance. The ABCA3 gene is associated with a lung disease involving dysfunction in surfactant metabolism. The subject reported respiratory deficit as a clinical sign, and mutations in the *ABCA3* gene have been associated in the literature with dysregulation mechanisms in alveolar epithelial cells.

In this particular case, the presence of variants in both the *POLG* and *ABCA3* genes in a patient experiencing respiratory stress highlights the potential additive effect of multiple genetic variations on an individual's health. Mutational burden underscores the complex interplay between genetic factors and disease susceptibility. It suggests that the cumulative impact of multiple genetic variants, even if individually considered benign or with limited clinical significance, can collectively contribute to the manifestation of clinical symptoms or disease17. In subject RX1100.2022 (ID 37), a single-copy variant was found in the *VARS2* gene, which shows recessive inheritance. The patient reported cardiological and neurological deficits and heart failure as a comorbidity. There is a publication in the literature that correlates the *VARS2* gene with cardiomyopathy and pulmonary hypertension[18]. The same variant associated with the *VARS2* gene was also found in subject RX1096.2022 (ID 36). The subject reported cardiac deficit as a clinical sign and obesity as a comorbidity. It subsequently emerged that subject RX1096.2022 (ID 36) is the mother of RX1100.2022 (ID 37). This could be compatible with the inheritance of the gene. This is the only case of familiarity in our case history whereby two variants were found in common, although it is still unclear whether the two variants have a pathogenetic effect and whether they may influence each other.

Furthermore, a single-copy variant in the *SDHA* gene, associated with mitochondrial deficiency and cardiomyopathy, was found in both subjects. There are articles in the literature confirming this association. The simultaneous presence of *SDHA* and *VARS2* in RX1096.2022 (ID 36) and RX1100.2022 (ID 37) could confirm the involvement of the genes in the mitochondrial and cardiological pathways. Subject RX1092.2022 also has a variant in the *VARS2* gene and reported cardiological, neurological, respiratory, and cardiovascular deficits as clinical signs[19,20].

Supplementary Table VI presents the frequencies of carriers bearing variants linked to Mendelian genetic diseases within our case series exceeding those observed in the general population. This provides evidence that long-COVID syndrome is significantly associated with carriers of Mendelian genetic illnesses. When comparing the frequency of these genetic variations in our com-

munity with that anticipated in a cohort without a family history of genetic illnesses, our analysis indicates a striking disparity. The higher percentages we found in our group may provide clinicians managing long-COVID patients with important information. Healthcare professionals can better understand the genetic characteristics that might predispose people to long-COVID symptoms that are more severe by identifying these elevated frequencies of genetic variations linked to cardiological, pulmonary, immunological, and neurological pathways. With this information at hand, clinicians can adjust how they approach diagnosis, treatment, and care, ultimately resulting in more individualized and effective management strategies for patients with long-COVID syndrome.

In addition, the *CFTR* and *POLG* genes, respectively have a higher frequency in our cohort than in the general population:

- *CFTR* in our population has a frequency of 17/95, corresponding to a carrier incidence of 49/100 *vs.* 1/25 in the general population;
- *POLG* in our population has a frequency of 11/95, which corresponds to a carrier incidence in our population of 45/100 against the literature figure of 1/10020[21];
- *DUOX2* in our population has a frequency of 5/95, corresponding to a carrier incidence of 34/100 *vs.* 1.8/100 in the general population[22].

Limitations

Despite the significant insights our study has provided, several limits must be acknowledged and considered. The genetic disorders found in our patient necessitate a further evaluation using specialized diagnostic techniques.

Conclusions

Long-COVID syndrome presents a complex and multifaceted medical challenge characterized by a range of persistent symptoms following SARS-CoV-2 infection. This syndrome has far-reaching implications for the affected individuals and healthcare systems worldwide.

The purpose of our project was to enroll patients with long-COVID syndrome for NGS analysis to investigate the association between genetic variants and the development of symptoms linked with the condition. We have uncovered associations between genes and cardiological, respiratory, immunological, and neurological symptoms.

Our analysis revealed the presence of Pathogenic and Likely Pathogenic variants in 16 genes related to the cardiological pathway among 18 patients. 9 patients were classified as 'probably positive' as they possessed P or LP variants in genes with autosomal dominant transmission or they exhibited two P or LP variants in genes with autosomal recessive transmission.

We observed a higher presence of mutations than we expected from the general population in *CFTR*, *POLG* and INFT2 genes, which underlines the importance of further exploring the link between mutations and symptom manifestation.

In conclusion, we identified a significant association between symptoms and genetic mutations in long-COVID patients, suggesting that the disease has an important genetic basis and therefore that patients presenting long-COVID symptoms have a genetic predisposition due to variants in genes associated with Mendelian diseases. This suggests that long-COVID syndrome is involved in the manifestation of underdiagnosed or undiagnosed Mendelian diseases in apparently healthy patients before infection.

Availability of Data and Materials

The data are within the test or in the supplementary materials document.

Funding

This research was funded by the Provincia Autonoma di Bolzano in the framework of LP 14/2006.

Authors' Contributions

Project administration M.B. and S.N.; Conceptualization: M.B.; Methodology: K.Dh., P.E.M., S.C., G.B., K.Do.; Investigation: F.F., L.A., A.C., A.P., M.G.D.A., G.A., and S.N.; Writing-original draft preparation: C.M. and M.C.M.; Writing, review and editing: K.Dh., P.E.M., S.C., G.B., F.F., L.A., A.C., A.P., M.G.D.A., G.A., S.N., K.Do., L.L., P.M., E.C., G.M. All authors have read and agreed to the published version of the manuscript.

Ethics Approval

The study was conducted in compliance with the 1975 Declaration of Helsinki and its latest amendments. The study was approved on the 12th of January 2021 by the Ethics Committee of the University of Brescia (Italy) Prot. No. NP4588.

Conflicts of Interest

The authors declare no conflict of interest.

References

1) Niemi MEK, Daly MJ, Ganna A. The human genetic epidemiology of COVID-19. Nat Rev Genet 2022: 533-546.
2) Cappadona C, Rimoldi V, Paraboschi EM, Asselta R. Genetic susceptibility to severe COVID-19. Infect Genet Evol 2023; 100: 105.
3) Kedor C, Freitag H, Meyer-Arndt L, Wittke K, Hanitsch LG, Zoller T, Steinbeis F, Haffke M, Rudolf G, Heidecker B, Bobbert T, Spranger J, Volk HD, Skurk C, Konietschke F, Paul F, Behrends U, Bellmann-Strobl J, Scheibenbogen C. A prospective observational study of Long-COVID-19 chronic fatigue syndrome following the first pandemic wave in Germany and biomarkers associated with symptom severity. Nat Commun 2022; 13: 5104. Erratum in: Nat Commun 2022; 13: 6009.
4) Lv Y, Zhang T, Cai J, Huang C, Zhan S, Liu J. Bioinformatics and systems biology approach to identify the pathogenetic link of Long-COVID and Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. Front Immunol 2022; 13: 952987.
5) Upadhyai P, Shenoy PU, Banjan B, Albeshr MF, Mahboob S, Manzoor I, Das R. Exome-Wide Association Study Reveals Host Genetic Variants Likely Associated with the Severity of COVID-19 in Patients of European Ancestry. Life (Basel) 2022; 12: 1300.
6) Cristofoli F, Sorrentino E, Guerri G, Miotto R, Romanelli R, Zulian A, Cecchin S, Paolacci S, Miertus J, Bertelli M, Maltese PE, Chiurazzi P, Stuppia L, Castori M, Marceddu G. Variant Selection and Interpretation: An Example of Modified VarSome Classifier of ACMG Guidelines in the Diagnostic Setting. Genes (Basel) 2021; 12: 1885.
7) Sorrentino E, Cristofoli F, Modena C, Paolacci S, Bertelli M, Marceddu G. Integration of VarSome API in an existing bioinformatic pipeline for automated ACMG interpretation of clinical variants. Eur Rev Med Pharmacol Sci 2021; 25: 1-6.
8) Cristofoli F, Daja M, Maltese PE, Guerri G, Tanzi B, Miotto R, Bonetti G, Miertus J, Chiurazzi P, Stuppia L, Gatta V, Cecchin S, Bertelli M, Marceddu G. MAGI-ACMG: Algorithm for the Classification of Variants According to ACMG and ACGS Recommendations. Genes (Basel) 2023; 14: 1600.
9) Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee. 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. Genet Med 2015; 17: 405-424.
10) Smeland MF, McClenaghan C, Roessler HI, Savelberg S, Hansen GÅM, Hjellnes H, Arntzen KA, Müller KI, Dybesland AR, Harter T, Sala-Rabanal M, Emfinger CH, Huang Y, Singareddy SS, Gunn J, Wozniak DF, Kovacs A, Massink M, Tessadori F, Kamel SM, Bakkers J, Remedi MS, Van Ghelue M, Nichols CG, van Haaften G. ABCC9-related Intellectual disability Myopathy Syndrome is a KATP channelopathy with loss-of-function mutations in ABCC9. Nat Commun 2019; 10: 4457.
11) Fricke-Galindo I, Martínez-Morales A, Chávez-Galán L, Ocaña-Guzmán R, Buendía-Roldán I, Pérez-Rubio G, Hernández-Zenteno RJ, Verónica-Aguilar A, Alarcón-Dionet A, Aguilar-Duran H, Gutiérrez-Pérez IA, Zaragoza-García O, Alanis-Ponce J, Camarena A, Bautista-Becerril B, Nava-Quiroz KJ, Mejía M, Guzmán-Guzmán IP, Falfán-Valencia R. IFNAR2 relevance in the clinical outcome of individuals with severe COVID-19. Front Immunol 2022; 13: 949413.
12) Yaugel-Novoa M, Bourlet T, Longet S, Botelho-Nevers E, Paul S. Association of IFNAR1 and IFNAR2 with COVID-19 severity. Lancet Microbe 2023; 4: e487.
13) Zhang Q, Bastard P; COVID Human Genetic Effort; Cobat A, Casanova JL. Human genetic and immunological determinants of critical COVID-19 pneumonia. Nature 2022; 603: 587-598.
14) Singh KK, Chaubey G, Chen JY, Suravajhala P. Decoding SARS-CoV-2 hijacking of host mitochondria in COVID-19 pathogenesis. Am J Physiol Cell Physiol 2020; 319: C258-C267.
15) Santos-Rebouças CB, Piergiorge RM, Dos Santos Ferreira C, Seixas Zeitel R, Gerber AL, Rodrigues MCF, Guimarães APC, Silva RM, Fonseca AR, Souza RC, de Souza ATAM, Rossi ÁD, Porto LCMS, Cardoso CC, de Vasconcelos ATR. Host genetic susceptibility underlying SARS-CoV-2-associated Multisystem Inflammatory Syndrome in Brazilian Children. Mol Med 2022; 28: 153.
16) Cabrera-Borrego E, Montero-Vilchez T, Bermúdez-Jiménez FJ, Tercedor-Sánchez J, Tercedor-Sánchez L, Sánchez-Díaz M, Macías-Ruiz R, Molina-Jiménez M, Cañizares-García FJ, Fernández-Segura E, Fernandez-Flores A, Arias-Santiago S, Jiménez-Jáimez J. Heterozygous Arrhythmogenic Cardiomyopathy-desmoplakin Mutation Carriers Exhibit a Subclinical Cutaneous Phenotype with Cell Membrane Disruption and Lack of Intercellular Adhesion. J Clin Med 2021; 10: 4608.
17) Herr DJ, Baarine M, Aune SE, Li X, Ball LE, Lemasters JJ, Beeson CC, Chou JC, Menick DR. HDAC1 localizes to the mitochondria of cardiac myocytes and contributes to early cardiac reperfusion injury. J Mol Cell Cardiol 2018; 114: 309-319.
18) Franklin AD, Chaudhari BP, Koboldt DC, Machut KZ. Polymerase Gamma Mitochondrial DNA Depletion Syndrome Initially Presenting as Disproportionate Respiratory Distress in a Moderately Premature Neonate: A Case Report. Front Genet 2021; 12: 664278.
19) Kušíková K, Feichtinger RG, Csillag B, Kalev OK, Weis S, Duba H-C, Mayr JA, Weis D. Case Report and Review of the Literature: A New and a Recurrent Variant in the VARS2 Gene Are Associated With Isolated Lethal Hypertrophic Cardiomyopathy, Hyperlactatemia, and Pulmonary Hypertension in Early Infancy. Front Pediatr 2021; 9: 660076.

20) Courage C, Jackson CB, Hahn D, Euro L, Nuoffer JM, Gallati S, Schaller A. SDHA mutation with dominant transmission results in complex II deficiency with ocular, cardiac, and neurologic involvement. Am J Med Genet A 2017; 173: 225-230.

21) Nurminen A, Farnum GA, Kaguni LS. Pathogenicity in POLG syndromes: DNA polymerase gamma pathogenicity prediction server and database. BBA Clin 2017; 7: 147-156.

22) Park KS. Analysis of Worldwide Carrier Frequency and Predicted Genetic Prevalence of Autosomal Recessive Congenital Hypothyroidism Based on a General Population Database. Genes (Basel) 2021; 12: 863.