JMolMed

ORIGINAL ARTICLE

Rare variants in cardiomyopathy genes predispose to cardiac injury in severe COVID-19 patients of African or Hispanic ancestry

Hui-Qi Qu[1] · Matthew S. Delfiner[2] · Chethan Gangireddy[2] · Anjali Vaidya[2] · Kenny Nguyen[1] · Isaac R. Whitman[2] · JuFang Wang[2] · Jianliang Song[2] · Michael R. Bristow[3] · Charles F. McTiernan[4] · Glenn S. Gerhard[5] · Hakon Hakonarson[1,6,7] · Arthur M. Feldman[2]

Abstract

In one of the earliest reports from China during COVID-19, it was noted that over 20% of patients hospitalized with the disease had significant elevations of troponin, a marker of myocardial tissue damage, that put them at a higher risk. In a hypothesis-independent whole exome sequencing (WES) study in hospitalized COVID-19 patients of diverse ancestry, we observed putative enrichment in pathogenic variants in genes known to be involved in the pathogenesis of cardiomyopathy. This observation led us to hypothesize that the observed high morbidity and mortality in these patients might be due to the presence of rare genetic factors that had previously been silent but became relevant as a consequence of the severe stress inflicted by an infection with SARS-CoV-2. To test this hypothesis, we analyzed our WES data generated from a cohort of 325 patients sequentially admitted for COVID-19 infection. In this predominantly minority population (53.9% African ancestry and 37.9% Hispanic/Latin ancestry), our initial analysis screen identified 263 variants that were identified as highly deleterious (HD) from a total of 26,661 variants of interest that represented 215 genes. Of those, we identified 46 genes (in 58 patients) harboring rare HD coding variants that were previously implicated in dilated cardiomyopathy and were considered as disease initiators for the severe COVID-19 in this study. These findings offer valuable insights into the molecular mechanisms and genetic susceptibility to heart injury in severe COVID-19.

Key messages

- COVID-19 may cause cardiac damage in some affected patients without a plausible biological explanation.
- Our study reveals an enrichment of highly deleterious variants linked to cardiomyopathy in severe COVID-19 patients.
- Genetic profiling unveils the molecular basis of severe COVID-19-related heart injury, potentially aiding in patient stratification.

Introduction

The severe acute respiratory syndrome virus-2 (SARS-CoV-2) has infected over 62 million people in the USA and the resulting disease, COVID-19, has contributed to the deaths of millions worldwide [1]. Most COVID-19 patients

Hui-Qi Qu and Matthew S. Delfiner contributed equally to this work.

Extended author information available on the last page of the article

have had relatively minor symptoms, but typical of other respiratory viruses, a subgroup presented with acute respiratory distress syndrome (ARDS) requiring hospitalization and an admission to an intensive care unit [2]. Consistent with previous studies with other respiratory viruses, mortality was worse in patients who were older and had a history of prior pulmonary disease [3]. However, early reports from China noted that 19.7% of hospitalized COVID-19 patients also had increased levels of high sensitivity troponin I (hsTnI), suggesting that these patients had sustained myocardial damage [4, 5]. Although many of the patients had typical risk

factors for cardiovascular disease, the group also included young patients without known risk factors. Increased levels of C-reactive protein (CRP), consistent with systemic inflammation, and high levels of N-terminal pro-B-type natriuretic peptide (NT-proBNP), suggesting myocardial congestion, were also present. As COVID-19 spread to the USA, a study of 2736 hospitalized COVID-19 patients in New York found that 36% had elevated cardiac troponin I (cTnI) concentrations and that even a small elevation was associated with an increased risk of death [6]. However, there was no plausible biological explanation for why only a small number of people developed cardiac damage after an infection with SARS-CoV-2 [7].

Recognizing the complexity of the cardiac inflammasome and the relatively modest number of COVID-19 patients who were hospitalized with cardiac damage, we hypothesized that the group of COVID-19 patients who develop cardiac damage had a pre-existing risk that made them susceptible to the cardiac stress of a severe viral infection. We also based our hypothesis on our knowledge of another enigmatic disease: post-partum cardiomyopathy (PPCM). In 1990, we reported for the first time that 14 of 18 women with PPCM had myocarditis on an initial endomyocardial biopsy (EMB) [8, 9]; however, most had no serious abnormalities on repeat biopsy six weeks later. Those patients whose repeat biopsies were normal had the eventual restoration of normal LV function whereas those who failed to clear their inflammation had persistent HF symptoms requiring transplantation. In a similar patient population to that of PPCM, Ware et al. found that 15% of the women had truncating variants of putative functional relevance, with the most frequent variants being found in the well-established cardiomyopathy gene, *Titin* (*TNN*) [10], a finding confirmed in a subsequent study [11].

As the results of these prior studies would suggest that similar underlying genetic variances might be a cause for the myocardial injury in a subset of patients with COVID-19, we elected to search for disease causing variants in our cohort. By contrast with earlier studies in myocarditis that focused on an established group of genes, the unique features of infections with COVID-19 led us to look more broadly for potentially causative variants. We therefore used WES to investigate the presence of genetic variants in inpatients with COVID-19 who were admitted to a large urban hospital that cares for a largely minority population. Not only is this among the first studies to systematically decipher the role of genetics in severe COVID-19 but it is also the first genomics study to focus a genetic analysis on minority populations that have been severely under-represented in genetic analyses to this date.

Materials and methods

Patients

Patients admitted consecutively between March 15, 2020 and July 26, 2020 were included in this analysis. The COVID-19 Cardiac Outcomes study (COVID CARDOS) was a retrospective registry of hospitalized COVID-19 patients that had an a priori goal of collecting and curating hospital records from 500 subjects. All hospitalized patients were 18 years of age or older and diagnosed with COVID-19 by polymerase chain reaction (PCR). Clinical datapoints included demographic information, medical history, medications on admission, laboratory values at admission, inpatient treatments including supplemental oxygen, patient disposition, and laboratory values throughout the hospitalization. Subjects were noted to have myocardial injury if their initial cTnI level was greater than the 99th percentile of the normal value. Patients were excluded from the registry if key demographic or clinical datapoints were missing from the medical record. All clinical data were obtained from the electronic medical record, anonymized by an independent broker, and entered into a secure database with REDCap. This study was performed in line with the principles of the Declaration of Helsinki. The Temple University Institutional Review Board (IRB) approved the protocol. All the patients were recruited before the availability of COVID-19 vaccines.

DNA collection and processing

Blood was collected as part of a separate IRB approved biobank from COVID-19 inpatients. In brief, EDTA collection tubes from which a small sample had been removed for automated blood analyses were collected daily by research personnel from the CLIA certified genomics laboratory. The patient identifier on each tube was removed and a replacement identifier affixed. The prior identifier was entered into the secure database, which was then transmitted to the independent broker, who then entered the information into a secure REDCap database. In this way, only the independent broker was able to link a blood sample to clinical data and neither representatives from the clinical registry nor the biobank had the ability to cross from one data set to the other. The collection tubes were stored at − 70 °C in the Human Genetics Biobank until such time as they were sequenced. DNA was isolated using the QIAGEN QIAamp Blood Mini protocol according to the manufacturer's instructions. After thawing, blood was

incubated with QIAGEN Protease or proteinase K at 56 °C for 10 min, 96–100% ethanol was added for loading on to a QIAamp Mini spin column followed by centrifugation at 6000 × g for 1 min. One wash buffer was applied followed by centrifugation at 6000 × g for 1 min with a second wash buffer at 20,000 × g for 3 min followed by a second spin for 1 min. DNA was eluted from the column using QIA-GEN elution buffer after incubation at room temperature (15–25 °C) for 5 min and centrifugation at 6000 × g for 1 min. Eluted DNA was stored frozen at – 20 °C. A third IRB approval allowed for the independent broker to provide the requisite data for the analyses performed as part of this study by coupling genetic data with clinical outcome data from the same patient with all personal identifiers removed.

Exome sequencing

Exome sequencing was done by the Helix Exome + ® v2 Assay (San Mateo, CA). Illumina NovaSeq was used for paired end sequencing (San Diego, CA). Data were processed using Sentieon DNASeq Variant Calling Workflow (https://www.sentieon.com/) and aligned to GRCh38. The WES callings were annotated using the ANNOVAR software [12]. Genetic variants which have minor allele frequency (MAF) less than 0.001 in the African ancestry group (the largest racial group), based on the Exome Aggregation Consortium (ExAC) database, were identified [13]. Functional candidate variants were selected by the prediction results with at least 1 of a number of genetic variant prediction software, i.e., SIFT_pred = "D" or Polyphen2_HDIV_pred = "D" or Polyphen2_HDIV_pred = "P" or Polyphen2_HVAR_pred = "D" or Polyphen2_HVAR_pred = "P" or LRT_pred = "D" or MutationTaster_pred = "A" or MutationTaster_pred = "D" or MutationAssessor_pred = "H" or MutationAssessor_pred = "M" or FATHMM_pred = "D" or PROVEAN_pred = "D" or MetaSVM_pred = "D" or MetaLR_pred = "D", based on the annotation with the ANNO-VAR software, as we have previously described [14, 15].

Data analysis

To determine the classification of pathogenic (P) or likely-pathogenic (LP) variants the rare coding variants in the sample were annotated with the InterVar software [16] for automating functional classification of sequence variants by the ACMG criteria, as well as the most recent data release of the Human Gene Mutation Database (HGMD_Pro_2022.1_hg38) [17]. Pathogenic (P) or likely pathogenic (LP) variants were also identified using ClinVar annotations, as described in detail, previously [18, 19]. Coding variants classified as pathogenic (P) or likely pathogenic (LP) were accepted only when identified by both InterVar/HGMD and ClinVar.

To identify the burden of rare coding variants, supplementary to the approach for Highly Deleterious (HD) variants, burden of functional rare coding variants was tested in the cases against the frequencies in the ExAC African ancestry population, and the frequencies in the ExAC Non-Finish European (NFE) population, respectively. Functional rare coding variants identified based on the ANNOVAR software were counted in the cases with the Test Rare vAriants with Public Data (TRAPD) software [20], and compared with the ExAC African ancestry and NFE controls under the dominant inheritance model by one tailed Fisher exact test. We have optimized the TRAPD algorithm with normalized genome coverage to capture causal variants with effects in the same directions [14, 15]. We corrected significance level for multiple testing by Bonferroni correction. The genome-wide significance of the gene-based burden test was defined as $\alpha = 0.05/21,306 = 2.35E\text{-}06$, assuming 21,306 protein-coding genes in the human genome [21]. Due to the racial heterogeneity of the research subjects, burden of rare coding variants in cases were tested against both African ancestry and NFE populations. Significance for a gene was required for both tests against African ancestry and NFE populations. Over-representation analysis (ORA) was done using the WebGestalt (WEB-based Gene SeT AnaLysis Toolkit) web tool by the human phenotype ontology (HPO) approach [22]. Statistics analysis for the multinomial logistic regression was done using the IBM SPSS Statistics Version 23 software.

Results

Study subjects

Of the 500 patients included in a cardiac clinical registry established to study COVID-19 patients, 325 had biological samples available for genetic analysis. Ages ranged from 18 to 90 years old, and there were 175 (53.9%) patients identified as African American (AA), 23 (7.1%) as White, 2 (0.6%) as Asian, 123 (37.9%) as Hispanic/Latin, and 2 (0.6%) were undetermined. Of the total, 139 (42.8%) were Female. In the cohort of 325 patients evaluated in this study, 175 (53.9%) were of African ancestry of whom 74 out of 164 measured (45.1%) had positive plasma troponin at admission; 89 of 168 (53.0%) had positive plasma troponin at some point in their hospitalization (i.e., 4 additional patients had troponin levels measured during the hospital stay but not on admission). Among non-African American (non-AA) patients, 48 out of 140 (34.3%) measured had positive plasma troponin at admission, while 61 out of 143 (42.7%) had positive plasma troponin at some point during their hospitalization.

As shown in Table 1, the baseline characteristics of the AA population were comparable to those of the Hispanic ancestry group except for a higher percentage of patients

Table 1 Clinical characteristics, laboratory values and clinical outcomes

	African American (n = 197)	Non-African American (n = 173)	p value
a. Baseline characteristics and presenting laboratory values			
Age (years)	61.5 (14.3)	61.0 (15.1)	0.75
Female gender	94 (48)	67 (39)	0.90
Hypertension	164 (84)	108 (62)	< 0.001
Chronic kidney disease	52 (26)	28 (38)	0.02
Coronary artery disease	28 (14)	34 (20)	0.16
Diabetes mellitus	89 (45)	77 (45)	0.56
History of stroke	27 (14)	26 (15)	
Heart failure with reduced Ejection fraction	21 (11)	15 (9)	0.53
Heart failure with preserved ejection fraction	22 (11)	22 (13)	0.61
Body mass index (kg/m^2)	32.1 (9.2)	31.1 (7.4)	0.27
Troponin (ng/mL)*	0.69 (2.3)	1.6 (15)	0.37
Elevated Troponin*	82 (42)	57 (33)	0.08
B-type natriuretic peptide (pg/mL)*	244 (636)	336 (707)	0.24
Elevated B-type natriuretic peptide*	55 (27)	54 (31)	0.88
C reactive protein (mg/dL)#	13.5 (43.2)	13.9 (28.2)	0.92
b. Clinical outcomes			
Lactic acid (mmol/L)#	1.8 (1.5)	1.9 (1.3)	0.76
Inpatient mortality	38 (19)	28 (16)	0.450
Intensive care	72 (37)	64 (37)	0.900
Intubation	40 (20)	25 (14)	0.146
Supplemental oxygen	143 (73)	130 (75)	0.398
Non-invasive positive pressure ventilation	35 (18)	41 (24)	0.152
High flow oxygen	62 (31)	61 (35)	0.420
c. WES	n = 175	n = 150	

Continuous variables noted as mean (standard deviation). Categorical variables noted as number (percent). *Elevated troponin determined by values > 0.012 ng/mL. Elevated B-type natriuretic peptide (BNP) determine by values > 100 pg/mL. Troponin and BNP are peak levels during hospital stay.# C reactive protein and lactic acid are levels at admission

of African ancestry had hypertension and chronic kidney disease when compared with the non-AA group. Similarly, clinical outcomes were not different between the two groups. Interestingly, 40.8% of the COVID-19 patient population had myocardial injury as evidenced by an elevated troponin at admission. That troponin was a useful tool for identifying patients with myocardial damage as evidenced by the finding that twice (27.3%) the number of individuals who had serum troponin detected at the time of admission died in the hospital compared to 12.4% deaths amongst those who did not have troponin leaks at the time of admission ($P = 0.021$ by binary logistic regression, corrected by age and BMI).

To conduct meaningful genetic analyses, ensuring homogeneity within cohorts is essential for the accuracy and reliability of our findings. For this reason, we stratified participants into two groups: AA and non-AA. This distinction recognizes the ancestral diversity of human populations, with African populations representing the root lineage for all modern humans based on evolutionary and genetic evidence. Dividing the cohort in this manner allows us to investigate

potential genetic and prognostic variations between these groups, which may influence health outcomes in the context of COVID-19. By integrating evolutionary biology with socioeconomic considerations related to race, our approach aims to illuminate the complex interplay between genetic predispositions and environmental or systemic factors contributing to health disparities. For example, hypertension, a major risk factor for severe COVID-19, is more prevalent in African American populations (Table 1). Additionally, genetic analyses revealed differences in the frequencies of HD variants in cardiomyopathy genes in AA and non-AA participants. This genetic heterogeneity further highlights the importance of stratified analyses by population groups to ensure robust and meaningful insights.

Sequencing

The sequencing was done in both the AA and non-AA cohorts. In this cohort, we identified 26,661 deleterious rare coding variants in 10,522 genes with allele frequencies in the

African ancestry population of <0.1%. Allele frequencies of functional variants can vary significantly across human populations due to evolutionary pressures and genetic drift [23]. Given the large representation of African American participants in this study, we opted to use allele frequencies from the African ancestry population to filter out common variants that are less likely to have a significant functional impact. Frequencies derived from other populations, such as European or Asian populations, may be influenced by positive or purifying selection related to their significant effects, which could lead to the misclassification of certain variants as rare or common. Variants were considered HD if, upon data aggregation, they were identified as deleterious by two of the following: ClinVar, HGMD classification (HGMD-Pro-2022.1-hg38), or InterVar software. (Supplementary Data 1).

Variants in cardiomyopathy genes

From the 26,661 variants of interest, we identified 263 variants from 215 genes as HD (Fig. 1). The 263 HD variants included 200 nonsynonymous variants, 46 stop/gain variants, and 17 splicing variants (Fig. 1). HD variants were seen in 88 of 175 patients of African ancestry (50.3%: Supplementary Data 2) and 115 of the 150 non-African ancestry patients (76.6%). (Supplementary Data 2).

Using the WebGestalt (WEB-based Gene SeT AnaLysis Toolkit) web tool, we performed over-representation analysis (ORA) using the Human Phenotype Ontology (HPO) pre-specified phenotype terms [24], for the 215 genes identified as HD variants in the COVID-19 patients. Among 10 gene sets found to be enriched for statistical significance (Supplementary Table 1), the most highly enriched gene set was cardiomyopathy (HP:0001638) which had a false discovery rate (FDR) of 7.08×10^{-05}. Included in this gene set were 46 genes that potentially play roles in cardiomyocytes, all of which

harbor HD variants as determined by the three programs used in this study (Fig. 2, Supplementary Table 2). The HD variants in these 46 genes were detected in 58 patients. Twenty-two of these patients were of African ancestry (37.9%) and were positive for at least one of these variants. Interestingly, in the patients harboring HD variants in these cardiomyopathy genes, we observed a significantly higher incidence of elevated troponin levels during their hospitalization [16 out of 21 patients with HD variants (76.2%) vs. 73 out of 147 patients without HD variants (49.7%) in African Americans, $p = 0.020$]. Covariate analysis showed that age was a risk factor for a positive troponin ($p = 0.014$), but BMI ($p = 0.597$) or grades of obesity ($p = 0.487$) were not risk factors.

Mutation burden analysis of rare coding variants

In parallel with the analysis we describe above, we performed mutation burden analysis of all of the 26,661 rare functional coding variants, to prioritize candidate genes. We identified 206 genes with significant p values ($p < 2.35 \times 10^{-06}$) when compared to both African ancestry and the ExAC Non-Finish European (NFE) reference population (Supplementary Data 3). Among these 206 genes, 16 genes overlapped with the list of 215 genes with HD variants. In addition, 46 cardiomyopathy genes identified as harboring HD variants showed varied significance of burden with coding variants (Table 2). Two mutations, i.e., caveolin 3 (*CAV3*) C72W and cadherin related 23 (*CDH23*) R1060W, were identified in patients of African and non-African ancestry. Furthermore, 2 of the 46 cardiomyopathy genes, *CDH23* and myosin binding protein C (*MYBPC3*), were significantly enriched for rare functional coding variants, as revealed by mutation burden analysis (Supplementary Data 3). Our further experimental research on these two genes in heart failure is described in the supplementary methods and results.

Fig. 1 Algorithm of genetic variant analysis in this study. By whole exome sequencing (WES), 26,661 variants of interest (VOI) were identified. Among these VOIs, 263 variants in 215 genes were identified as highly deleterious (HD). Over-representation analysis by Human Phenotype Ontology (HPO) identified 51 HD variants from 46 genes in the HPO gene set Cardiomyopathy (HP:0001638). In parallel, mutation burden analysis of the 26,661 coding variants identified 206 genes with significant *p* values ($p < 2.347 \times 10^{-6}$)
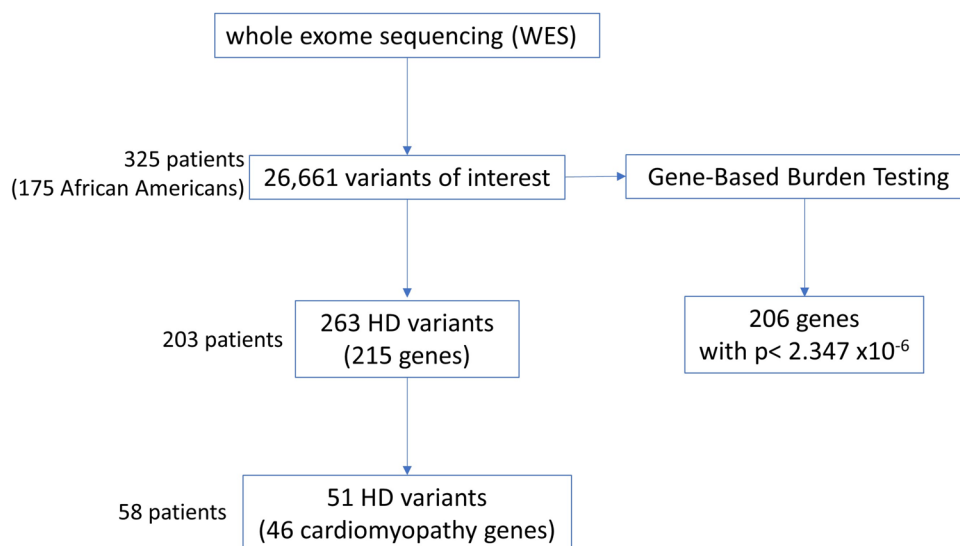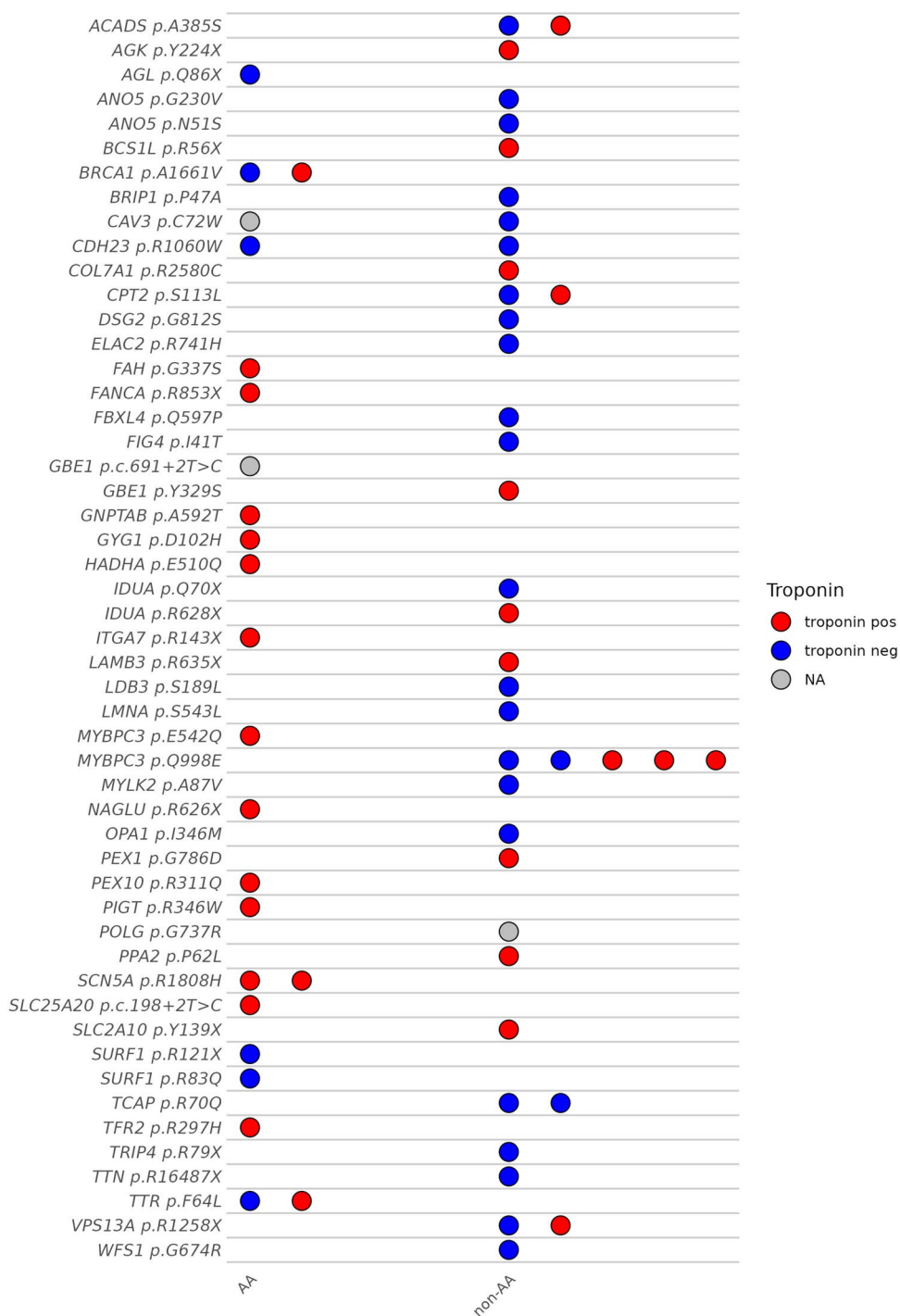
Fig. 2 Distribution of HD variants in each group. Each circle represents an individual with a mutation, with the corresponding troponin status indicated by color

Discussion

This analysis of severe COVID-19 patients represents a groundbreaking discovery, as it unveils a genetic predisposition that is notably enriched for cardiomyopathy-associated genes, all identified through hypothesis-independent analysis. Building upon this compelling observation of heightened cardiomyopathy gene enrichment, we conducted a targeted investigation into "well-defined" cardiomyopathy genes using a hypothesis-driven approach. In this regard, we hypothesized that patients who developed myocardial injury during an infection with the SARS-CoV-2 virus carried underlying variants in cardiac genes that limited the heart's ability to support normal cardiac output during the stress of a severe viral infection. A corollary to this underlying hypothesis was that the abundance of genetic variation in the human genome accounts for a wide range of responses to physiologic stress—in this case to the stress

Table 2 Cardiomyopathy genes identified of HD variants and burden with coding variants

Gene Symbol	Variant	Count_Afr	Count_NonAfr	subtotal	*Heterogeneity P* (2 tailed Fisher's)	ExAC_AFR	ExAC_Freq (all)
ACADS	A385S	0	2	2	0.212	0	4.33E-05
AGK	Y224X	0	1	1	0.462	0	2.49E-05
AGL	Q86X	1	0	1	1.000	3.00E-04	2.48E-05
ANO5	N51S	0	1	1	0.462	7.00E-04	1.90E-03
	G230V	0	1	1	0.462	2.00E-04	1.10E-03
BCS1L	R56X	0	1	1	0.462	9.61E-05	2.00E-04
BRCA1	A1661V	2	0	2	0.501	4.00E-04	3.32E-05
BRIP1	P47A	0	1	1	0.462	0	2.00E-04
CAV3	C72W	1	1	2	1.000	6.00E-04	1.10E-03
CDH23	R1060W	1	1	2	1.000	2.00E-04	4.00E-04
COL7A1	R2580C	0	1	1	0.462	0	3.96E-05
CPT2	S113L	0	2	2	0.212	2.00E-04	1.30E-03
DSG2	G812S	0	1	1	0.462	0	1.66E-05
ELAC2	R741H	0	1	1	0.462	4.00E-04	5.00E-04
FAH	G337S	1	0	1	1.000	3.00E-04	1.00E-04
FANCA	R853X	1	0	1	1.000	9.69E-05	8.26E-06
FBXL4	Q597P	0	1	1	0.462	0	1.65E-05
FIG4	I41T	0	1	1	0.462	9.91E-05	1.00E-03
GBE1	Y329S	0	1	1	0.462	0	3.00E-04
	c.691 + 2 T > C	1	0	1	1.000	0	1.10E-03
GNPTAB	A592T	1	0	1	1.000	0	1.00E-04
GYG1	D102H	1	0	1	1.000	2.00E-04	9.00E-04
HADHA	E510Q	1	0	1	1.000	4.00E-04	1.20E-03
IDUA	R628X	0	1	1	0.462	0	3.35E-05
	Q70X	0	1	1	0.462	1.00E-04	7.00E-04
ITGA7	R143X	1	0	1	1.000	2.00E-04	4.12E-05
LAMB3	R635X	0	1	1	0.462	2.00E-04	8.00E-04
LDB3	S189L	0	1	1	0.462	0	6.00E-04
LMNA	S543L	0	1	1	0.462	0	1.00E-04
MYBPC3	Q998E	0	5	5	0.020*	4.00E-04	5.20E-03
	E542Q	1	0	1	1.000	1.00E-04	2.49E-05
MYLK2	A87V	0	1	1	0.462	3.00E-04	1.00E-04
NAGLU	R626X	1	0	1	1.000	2.00E-04	2.74E-05
OPA1	I346M	0	1	1	0.462	4.00E-04	6.00E-04
PEX1	G786D	0	1	1	0.462	9.71E-05	3.00E-04
PEX10	R311Q	1	0	1	1.000	5.00E-04	3.77E-05
PIGT	R346W	1	0	1	1.000	0	1.65E-05
POLG	G737R	0	1	1	0.462	3.00E-04	7.00E-04
PPA2	P62L	0	1	1	0.462	9.89E-05	2.00E-04
SCN5A	R1808H	2	0	2	0.501	1.00E-04	3.31E-05
SLC25A20	c.198 + 2 T > C	1	0	1	1.000	6.00E-04	4.94E-05
SLC2A10	Y139X	0	1	1	0.462	0	8.24E-06
SURF1	R83Q	1	0	1	1.000	0	1.80E-05
	R121X	1	0	1	1.000	9.63E-05	2.00E-04
TCAP	R70Q	0	2	2	0.212	0	2.09E-05
TFR2	R297H	1	0	1	1.000	9.83E-05	7.00E-04
TRIP4	R79X	0	1	1	0.462	0	2.47E-05
TTN	R16487X	0	1	1	0.462	0	8.34E-06
TTR	F64L	2	0	2	0.501	6.00E-04	5.77E-05

Table 2 (continued)

Gene Symbol	Variant	Count_Afr	Count_NonAfr	subtotal	*Heterogeneity P* (2 tailed Fisher's)	ExAC_AFR	ExAC_Freq (all)
VPS13A	R1258X	0	2	2	0.212	0	8.24E-06
WFS1	G674R	0	1	1	0.462	0	3.00E-04

*Among the 51 variants from the 46 genes, 28 are currently classified as P/LP, 16 with conflicting interpretations of pathogenicity (CIP), 6 as uncertain significance (VUS), and 1 as Benign/Likely benign. All the 16 CIP variants were previously classified as P/LP by ClinVar according to ANNOVAR, while 4 were predicted as LP by InterVar, and the left were classified as DM by HGMD. All the 6 VUS were previously classified as P/LP by ClinVar according to ANNOVAR, while 1 was predicted as LP by InterVar, and the left were classified as DM by HGMD. The 1 B/LB variant was previously classified as P/LP by ClinVar according to ANNOVAR, and classified as DM by HGMD. In particular, the controversial B/LB variant in *MYBPC3* was only seen in non-African individuals, and was not seen in African Americans that we demonstrated of genetic correlation with elevated troponin. Burden test of rare VOIs in *MYBPC3* has genome-wide significance as well. More information is shown in Supplementary Data 2. The 46 genes are included in all these HPO genesets: Abnormality of cardiovascular system morphology (HP:0030680); Abnormal heart morphology (HP:0001627); Abnormal myocardium morphology (HP:0001637); Cardiomyopathy (HP:0001638); Abnormality of the cardiovascular system (HP:0001626)

of a severe viral infection [25]. While the number of HD variants in this patient cohort is high, it is important to note that these hospitalized for COVID-19 patients were all very sick and there were no COVID-19 specific treatments at the time of the study conduct. On a single day at the mid-point of the data collection there were over 500 new cases in the Philadelphia County averaging about 40 deaths per day, and Temple University carried the highest census as all elective procedures had been discontinued and a 100 bed unit within the hospital was established only to treat COVID-19. Therefore, it would not be unexpected to see a significant number of maladaptive genetic variants in a population enriched for SARS-CoV-2 infected patients, including severe pulmonary and cardiac disease.

Among the characteristic features of patients who were hospitalized for COVID-19 was that they demonstrated an elevation of cTnI. The choice of troponin as a biomarker was supported by a recent study performed at four hospitals in New York that enrolled 2736 COVID-19 patients and used measures of cTnI as a marker of myocardial injury [6]. They found that small elevations of troponin representing limited myocardial injury were associated with death while greater amounts of troponin were associated with a higher risk of death. Their results were consistent with earlier studies [26] as well as with the present study, in which we found that an increase in cTnI levels was associated with a worse outcome.

The present study is the first to identify a spectrum of rare HD variants in the landscape of COVID-19 myocardial injury. Among these, 10 genes, *CAV3*, *DSG2*, *LDB3*, *LMNA*, *MYBPC3*, *MYLK2*, *SCN5A*, *TCAP*, *TTN*, and *TTR*, have been defined as the associated genes of cardiomyopathy by the 2023 European Society of Cardiology (ESC) Guidelines [27]. The remaining 36 genes, though not explicitly recognized by the ESC Guidelines, are implicated in diverse mechanisms that maintain cardiac structure and function. These mechanisms include contributions to the stability of sarcomeric and extracellular matrix proteins, which preserve the structural integrity of cardiomyocytes and prevent damage caused by mechanical stress [28–30]. Dysfunctions in autophagy, a cellular process essential for the removal of damaged organelles and proteins, disrupt cardiomyocyte homeostasis and lead to the accumulation of toxic cellular debris [31, 32]. Impaired mitochondrial maintenance reduces ATP production, increases oxidative stress, and heightens the heart's vulnerability to energy deficits [33–37]. DNA repair mechanisms protect cardiomyocytes from genotoxic stress and ensuring genomic stability, with deficiencies potentially leading to apoptosis or senescence [38, 39]. Furthermore, disruptions in lipid metabolism [40–42], glycogen storage [43–45], and cellular signaling pathways[29, 46, 47] can drive hypertrophic and dilated cardiomyopathies by altering energy homeostasis and metabolic efficiency. The potential roles of these 46 genes in non-ischemic heart failure, as previously reported, are referenced in Supplementary Table 2.

The major limitation of our study is the small size owing to the fact that we collected samples and data from only a single institution and were constrained by the timing of the introduction of antibodies, vaccines and other treatment strategies. We also acknowledge that focusing solely on hospitalized patients may introduce selection bias, as this group represents a subset of the COVID-19 population with more severe disease manifestations. Follow-up studies are warranted to identify whether the differences in genetic variants seen in this first study of COVID-19 patients and prior studies in non-infectious DCM are attributable to differences in the genetic landscape or to fundamental biological differences in genomic variants across different ancestral groups. Tools like WES are indeed critical for studying rare variants, particularly in addressing minority populations and health disparities. This work provides a starting point for those studies, which may present a new perspective of precision

medicine in both cardiology and infectious diseases. Also, given the study's limited sample size, it is important to acknowledge its lack of statistical power to identify a specific gene as a novel contributor to viral heart injury. It did not include a comparable hospitalized control population without COVID. Additionally, the patients exhibited heterozygous phenotypes. Functional cardiac data, such as ejection fraction (EF) measured by transthoracic echocardiography (TTE), was not available for analysis, rendering comparison of cardiac function between subjects with and without potentially pathogenic variants more difficult. We envision that further research on available samples from existing SARS-CoV-2 studies may be able to address some of these limitations and confirm our findings. Expanding analyses to larger and more diverse cohorts will enhance the generalizability of our findings, helping to confirm they are not limited to specific populations. However, the success of vaccines and other treatment strategies necessitates larger sample sizes to effectively study cardiac injury in COVID-19 patients.

Declarations

Ethics approval and consent to participate The Temple University Institutional Review Board (IRB) approved the protocol. Informed consent was obtained from all subjects.

Competing interests Dr. Feldman received research support from, was a consultant to, and held equity in Renovacor Inc. (now Rocket Pharmaceuticals); however, no Renovacor funds were used to support the studies described herein. The investigative group have no other conflicts to describe.

References

1. Yang J, Zhang Q, Cao Z, Gao J, Pfeiffer D, Zhong L, Zeng DD (2021) The impact of non-pharmaceutical interventions on the prevention and control of COVID-19 in New York City<? A3B2 show [editpick]?>. Chaos: An Interdisciplinary Journal of Nonlinear Science 31:021101

2. Gibson PG, Qin L, Puah SH (2020) COVID-19 acute respiratory distress syndrome (ARDS): clinical features and differences from typical pre-COVID-19 ARDS. Med J Aust 213(54–56):e51

3. Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, Huang H, Zhang L, Zhou X, Du C et al (2020) Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. JAMA Intern Med 180:934–943. https://doi.org/10.1001/jamainternmed.2020.0994

4. Shi S, Qin M, Yang B (2020) Coronavirus disease 2019 (COVID-19) and cardiac injury-reply. JAMA Cardiol 5:1199–1200. https://doi.org/10.1001/jamacardio.2020.2456

5. Shi S, Qin M, Shen B, Cai Y, Liu T, Yang F, Gong W, Liu X, Liang J, Zhao Q et al (2020) Association of cardiac injury with mortality in hospitalized patients with COVID-19 in Wuhan, China. JAMA Cardiol 5:802–810. https://doi.org/10.1001/jamacardio.2020.0950

6. Lala A, Johnson KW, Januzzi JL, Russak AJ, Paranjpe I, Richter F, Zhao S, Somani S, Van Vleck T, Vaid A et al (2020) Prevalence and impact of myocardial injury in patients hospitalized with COVID-19 infection. J Am Coll Cardiol 76:533–546. https://doi.org/10.1016/j.jacc.2020.06.007

7. Topol EJ (2020) COVID-19 can affect the heart. Science (New York, NY) 370:408–409. https://doi.org/10.1126/science.abe2813

8. Felker GM, Jaeger CJ, Klodas E, Thiemann DR, Hare JM, Hruban RH, Kasper EK, Baughman KL (2000) Myocarditis and long-term survival in peripartum cardiomyopathy. Am Heart J 140:785–791. https://doi.org/10.1067/mhj.2000.110091

9. Midei MG, DeMent SH, Feldman AM, Hutchins G, Baughman K (1990) Peripartum myocarditis and cardiomyopathy. Circulation 81:922–928

10. Ware JS, Li J, Mazaika E, Yasso CM, DeSouza T, Cappola TP, Tsai EJ, Hilfiker-Kleiner D, Kamiya CA, Mazzarotto F et al (2016) Shared genetic predisposition in peripartum and dilated cardiomyopathies. N Engl J Med 374:233–241. https://doi.org/10.1056/NEJMoa1505517

11. Goli R, Li J, Brandimarto J, Levine LD, Riis V, McAfee Q, DePalma S, Haghighi A, Seidman J, Seidman CE (2021) Genetic and phenotypic landscape of peripartum cardiomyopathy. Circulation 143:1852–1862

12. Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38:e164. https://doi.org/10.1093/nar/gkq603

13. Karczewski KJ, Weisburd B, Thomas B, Solomonson M, Ruderfer DM, Kavanagh D, Hamamsy T, Lek M, Samocha KE, Cummings BB et al (2017) The ExAC browser: displaying reference data information from over 60 000 exomes. Nucleic Acids Res 45:D840–D845. https://doi.org/10.1093/nar/gkw971

14. Qu HQ, Qu J, Vaccaro C, Chang X, Mentch F, Li J, Mafra F, Nguyen K, Gonzalez M, March M et al (2022) Genetic analysis for type 1 diabetes genes in juvenile dermatomyositis unveils genetic disease overlap. Rheumatology (Oxford) 61:3497–3501. https://doi.org/10.1093/rheumatology/keac100

15. Liu Y, Qu HQ, Qu J, Chang X, Mentch FD, Nguyen K, Tian L, Glessner J, Sleiman PMA, Hakonarson H (2022) Burden of rare coding variants reveals genetic heterogeneity between obese and non-obese asthma patients in the African American population. Respir Res 23:116. https://doi.org/10.1186/s12931-022-02039-0

16. Li Q, Wang K (2017) InterVar: clinical interpretation of genetic variants by the 2015 ACMG-AMP guidelines. Am J Hum Genet 100:267–280. https://doi.org/10.1016/j.ajhg.2017.01.004

17. Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NS, Abeysinghe S, Krawczak M, Cooper DN (2003) Human Gene Mutation Database (HGMD): 2003 update. Hum Mutat 21:577–581. https://doi.org/10.1002/humu.10212

18. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J (2016) ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res 44:D862–D868

19. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Jang W et al (2018) ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res 46:D1062–D1067. https://doi.org/10.1093/nar/gkx1153

20. Guo MH, Plummer L, Chan YM, Hirschhorn JN, Lippincott MF (2018) Burden testing of rare variants identified through exome sequencing via publicly available control data. Am J Hum Genet 103:522–534. https://doi.org/10.1016/j.ajhg.2018.08.016

21. Pertea M, Shumate A, Pertea G, Varabyou A, Chang Y-C, Madugundu AK, Pandey A, Salzberg SL (2018) Thousands of large-scale RNA sequencing experiments yield a comprehensive new human gene list and reveal extensive transcriptional noise. bioRxiv 332825. https://doi.org/10.1101/332825

22. Kohler S, Vasilevsky NA, Engelstad M, Foster E, McMurry J, Ayme S, Baynam G, Bello SM, Boerkoel CF, Boycott KM et al (2017) The human phenotype ontology in 2017. Nucleic Acids Res 45:D865–D876. https://doi.org/10.1093/nar/gkw1039

23. Hancock AM, Alkorta-Aranburu G, Witonsky DB, Di Rienzo A (2010) Adaptations to new environments in humans: the role of subtle allele frequency shifts. Philosophical Transactions of the Royal Society B: Biological Sciences 365:2459–2468

24. Köhler S, Gargano M, Matentzoglu N, Carmody LC, Lewis-Smith D, Vasilevsky NA, Danis D, Balagura G, Baynam G, Brower AM (2021) The human phenotype ontology in 2021. Nucleic Acids Res 49:D1207–D1217

25. Kanneganti T-D (2010) Central roles of NLRs and inflammasomes in viral infection. Nat Rev Immunol 10:688–698

26. Blankenberg S, Salomaa V, Makarova N, Ojeda F, Wild P, Lackner KJ, Jorgensen T, Thorand B, Peters A, Nauck M et al (2016) Troponin I and cardiovascular risk prediction in the general population: the BiomarCaRE consortium. Eur Heart J 37:2428–2437. https://doi.org/10.1093/eurheartj/ehw172

27. Arbelo E, Protonotarios A, Gimeno JR, Arbustini E, Barriales-Villa R, Basso C, Bezzina CR, Biagini E, Blom NA, de Boer RA et al (2023) 2023 ESC Guidelines for the management of cardiomyopathies. Eur Heart J 44:3503–3626. https://doi.org/10.1093/eurheartj/ehad194

28. Sidwell RU, Yates R, Atherton D (2000) Dilated cardiomyopathy in dystrophic epidermolysis bullosa. Arch Dis Child 83:59–63. https://doi.org/10.1136/adc.83.1.59

29. Bugiardini E, Nunes AM, Oliveira-Santos A, Dagda M, Fontelonga TM, Barraza-Flores P, Pittman AM, Morrow JM, Parton M, Houlden H (2022) Integrin α7 mutations are associated with adult-onset cardiac dysfunction in humans and mice. J Am Heart Assoc 11:e026494

30. Will LM, Reichrath J, Vogt T (2021) Epidermolysis bullosa dystrophica pretibialis—clinical snapshot and management of a rare orphan disease. JDDG: Journal der Deutschen Dermatologischen Gesellschaft 19: 983–986. https://doi.org/10.1111/ddg.14446

31. Sabouny R, Wong R, Lee-Glover L, Greenway SC, Sinasac DS, Canada CR, Khan A, Shutt TE (2019) Characterization of the C584R variant in the mtDNA depletion syndrome gene FBXL4, reveals a novel role for FBXL4 as a regulator of mitochondrial fusion. Biochimica et Biophysica Acta (BBA)-Mol Basis Dis 1865:165536

32. Deneubourg C, Ramm M, Smith LJ, Baron O, Singh K, Byrne SC, Duchen MR, Gautel M, Eskelinen E-L, Fanto M (2022) The spectrum of neurodevelopmental, neuromuscular and neurodegenerative disorders due to defective autophagy. Autophagy 18:496–517

33. Al-Owain M, Colak D, Albakheet A, Al-Younes B, Al-Humaidi Z, Al-Sayed M, Al-Hindi H, Al-Sugair A, Al-Muhaideb A, Rahbeeni Z (2013) Clinical and biochemical features associated with BCS1L mutation. J Inherit Metab Dis 36:813–820

34. Ojala T, Nupponen I, Saloranta C, Sarkola T, Sekar P, Breilin A, Tyni T (2015) Fetal left ventricular noncompaction cardiomyopathy and fatal outcome due to complete deficiency of mitochondrial trifunctional protein. Eur J Pediatr 174:1689–1692

35. Chen L, Liu T, Tran A, Lu X, Tomilov AA, Davies V, Cortopassi G, Chiamvimonvat N, Bers DM, Votruba M (2012) OPA 1 Mutation and late-onset cardiomyopathy: mitochondrial dysfunction and mtDNA instability. J Am Heart Assoc 1:e003012

36. Spracklen TF, Kasher PR, Kraus S, Botha TL, Page DJ, Kamuli S, Booi Z, Chin A, Laing N, Keavney BD (2021) Identification of a POLG variant in a family with arrhythmogenic cardiomyopathy and left ventricular fibrosis. Circulation: Genomic Precis Med 14:e003138

37. Wedatilake Y, Brown RM, McFarland R, Yaplito-Lee J, Morris AA, Champion M, Jardine PE, Clarke A, Thorburn DR, Taylor RW et al (2013) SURF1 deficiency: a multi-centre natural history study. Orphanet J Rare Dis 8:96. https://doi.org/10.1186/1750-1172-8-96

38. Shukla PC, Singh KK, Quan A, Al-Omran M, Teoh H, Lovren F, Cao L, Rovira II, Pan Y, Brezden-Masley C et al (2011) BRCA1 is an essential regulator of heart function and survival following myocardial infarction. Nat Commun 2:593. https://doi.org/10.1038/ncomms1601

39. Cantor SB, Bell DW, Ganesan S, Kass EM, Drapkin R, Grossman S, Wahrer DC, Sgroi DC, Lane WS, Haber DA et al (2001) BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. Cell 105:149–160. https://doi.org/10.1016/s0092-8674(01)00304-x

40. Dessein A-F, Fontaine M, Joncquel-Chevalier Curt M, Briand G, Sechter C, Mention-Mulliez K, Dobbelaere D, Douillard C, Lacour A, Redonnet-Vernhet I et al (2017) Fluxomic evidence for impaired contribution of short-chain acyl-CoA dehydrogenase to

mitochondrial palmitate β-oxidation in symptomatic patients with ACADS gene susceptibility variants. Clin Chim Acta 471:101–106. https://doi.org/10.1016/j.cca.2017.05.026

41. Longo N, Amat di San Filippo C, Pasquali M (2006) Disorders of carnitine transport and the carnitine cycleAmerican Journal of Medical Genetics Part C: Seminars in Medical Genetics Wiley Online Library, pp. 77–85.

42. Iacobazzi V, Invernizzi F, Baratta S, Pons R, Chung W, Garavaglia B, Dionisi-Vici C, Ribes A, Parini R, Huertas MD (2004) Molecular and functional analysis of SLC25A20 mutations causing carnitine-acylcarnitine translocase deficiency. Hum Mutat 24:312–320

43. Lucchiari S, Fogh I, Prelle A, Parini R, Bresolin N, Melis D, Fiori L, Scarlato G, Comi G (2002) Clinical and genetic variability of glycogen storage disease type IIIa: seven novel AGL gene mutations in the Mediterranean area. Am J Med Genet 109:183–190

44. Fernandez C, Halbert C, De Paula AM, Lacroze V, Froissart R, Figarella-Branger D, Chabrol B, Pellissier JF (2010) Non-lethal neonatal neuromuscular variant of glycogenosis type IV with novel GBE1 mutations. Muscle Nerve 41:269–271

45. Hedberg-Oldfors C, Glamuzina E, Ruygrok P, Anderson LJ, Elliott P, Watkinson O, Occleshaw C, Abernathy M, Turner C, Kingston N (2017) Cardiomyopathy as presenting sign of glycogenin-1 deficiency—report of three cases and review of the literature. J Inherit Metab Dis 40:139–149

46. Kvarnung M, Nilsson D, Lindstrand A, Korenke GC, Chiang SC, Blennow E, Bergmann M, Stödberg T, Mäkitie O, Anderlid B-M (2013) A novel intellectual disability syndrome caused by GPI anchor deficiency due to homozygous mutations in PIGT. J Med Genet 50:521–528

47. Johannsen J, Heinemeyer J, Hempel M, Lessel D, Mueller-Stoever S, Strom TM, Matschke J, Denecke J (2017) A novel nonsense mutation in TRIP4 gene causes severe muscular weakness with respiratory failure and cardiomyopathy but without skin, joint, and/or bone abnormalities. Neuropediatrics 48:PP07. https://doi.org/10.1055/s-0037-1602929

Authors and Affiliations

Hui-Qi Qu[1] · Matthew S. Delfiner[2] · Chethan Gangireddy[2] · Anjali Vaidya[2] · Kenny Nguyen[1] · Isaac R. Whitman[2] · JuFang Wang[2] · Jianliang Song[2] · Michael R. Bristow[3] · Charles F. McTiernan[4] · Glenn S. Gerhard[5] · Hakon Hakonarson[1,6,7] · Arthur M. Feldman[2]

✉ Hakon Hakonarson
hakonarson@chop.edu

✉ Arthur M. Feldman
arthur.feldman@tuhs.temple.edu

[1] The Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, USA

[2] Department of Medicine, Division of Cardiology, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA

[3] Division of Cardiology, University of Colorado Anschutz Medical Campus, Aurora, CO, USA

[4] Division of Cardiology, Heart and Vascular Institute, University of Pittsburgh, Pittsburgh, PA, USA

[5] Department of Human Genetics and Molecular Biochemistry, Lewis Katz School of Medicine at Temple University, Philadelphia, PA, USA

[6] Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA, USA

[7] Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA