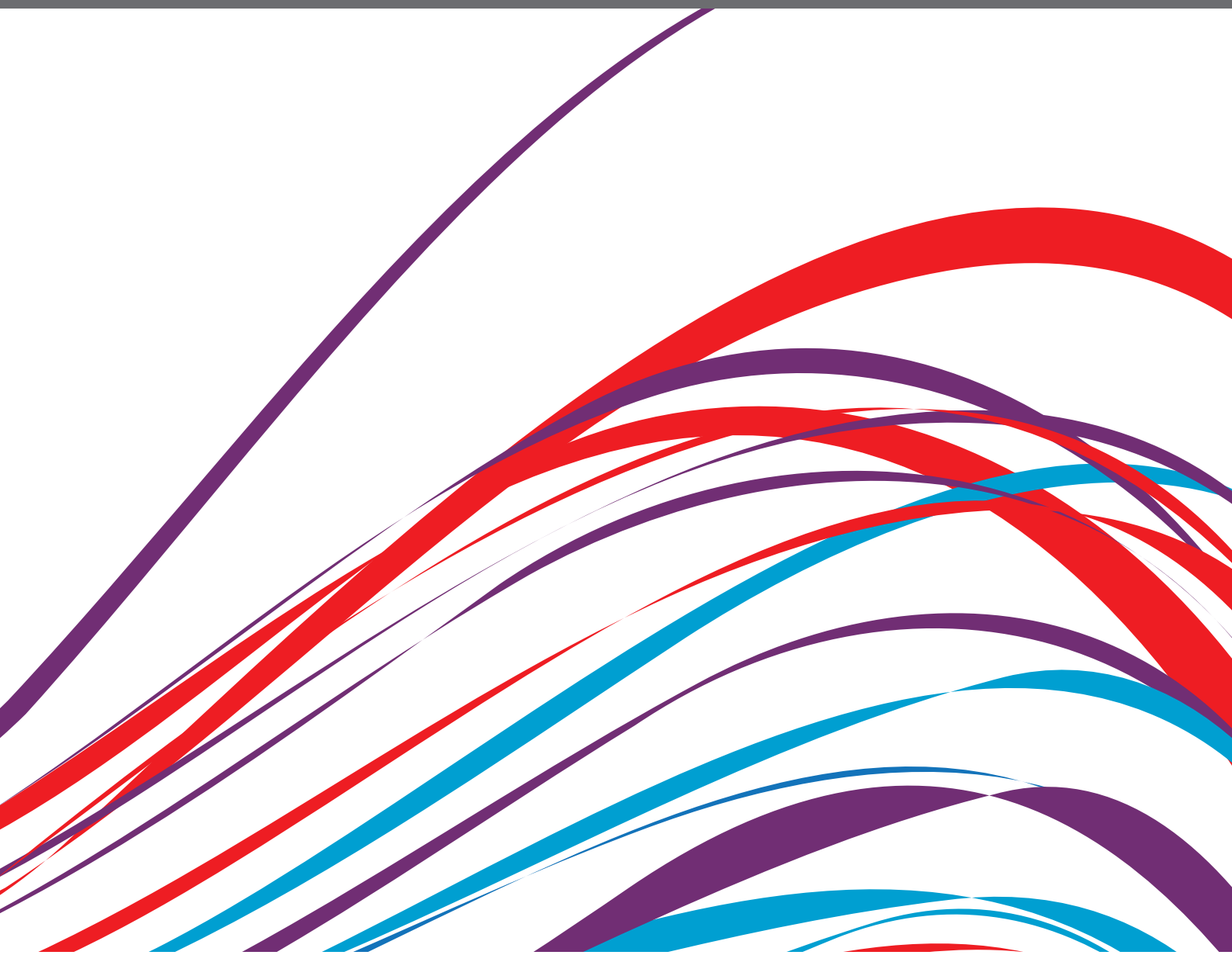
GENOMIC ARCHITECTURE OF CARDIOVASCULAR DISEASES

EDITED BY: Seitaro Nomura and Kaoru Ito
PUBLISHED IN: Frontiers in Cardiovascular Medicine



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-393-1
DOI 10.3389/978-2-88976-393-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

GENOMIC ARCHITECTURE OF CARDIOVASCULAR DISEASES

Topic Editors:
Seitaro Nomura, The University of Tokyo, Japan
Kaoru Ito, RIKEN Yokohama, Japan

Citation: Nomura, S., Ito, K., eds. (2022). Genomic Architecture of Cardiovascular Diseases. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-393-1

Table of Contents

frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 28 October 2020
doi: 10.3389/fcvm.2020.580573

Comprehensive Identification of Key Genes Involved in Development of Diabetes Mellitus-Related Atherogenesis Using Weighted Gene Correlation Network Analysis

Qi Huang[1,2†], *Guoxiong Deng*[1,2†], *Rongguo Wei*[3,4†], *Qiaoye Wang*[5,6], *Donghua Zou*[5,6*] and *Jinru Wei*[1,2*]

[1] Department of Cardiology, The Fifth Affiliated Hospital of Guangxi Medical University, Nanning, China, [2] Department of Cardiology, The First People's Hospital of Nanning, Nanning, China, [3] Department of Clinical Laboratory, The Fifth Affiliated Hospital of Guangxi Medical University, Nanning, China, [4] Department of Clinical Laboratory, The First People's Hospital of Nanning, Nanning, China, [5] Department of Neurology, The Fifth Affiliated Hospital of Guangxi Medical University, Nanning, China, [6] Department of Neurology, The First People's Hospital of Nanning, Nanning, China

OPEN ACCESS

Edited by:
Seitaro Nomura,
The University of Tokyo, Japan

Reviewed by:
Ville-Petteri Makinen,
South Australian Health and Medical Research Institute (SAHMRI), Australia
Paras Kumar Mishra,
University of Nebraska Medical Center, United States

***Correspondence:**
Donghua Zou
danvor0922@hotmail.com
Jinru Wei
weijinru@stu.gxmu.edu.cn

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: 06 July 2020
Accepted: 15 September 2020
Published: 28 October 2020

Citation:
Huang Q, Deng G, Wei R, Wang Q, Zou D and Wei J (2020) Comprehensive Identification of Key Genes Involved in Development of Diabetes Mellitus-Related Atherogenesis Using Weighted Gene Correlation Network Analysis. Front. Cardiovasc. Med. 7:580573. doi: 10.3389/fcvm.2020.580573

Coronary heart disease (CHD) is common in patients with diabetes mellitus (DM), however, the relevant mechanism remains elusive. The whole blood gene expression profiles of healthy control, patients with DM, patients with DM and CHD (DMCHD) were used to performed weight gene correlation network analysis (WGCNA) to identify the gene modules associated with DM-related atherogenesis. The candidate module was significantly involved in immune- and T cell activity-related biological process. GSEA results suggested that lysosome and apoptosis were enriched in DM and DMCHD samples. The protein-protein-KEGG pathway network may reveal the potential transcriptional regulatory network involving in DM-related atherosclerosis. Nineteen genes (RTKN, DCP1B, PDZD4, CACNA2D2, TSEN54, PVRIG, PLEKHF1, NKG7, ZAP70, NUDCD3, SLAMF6, CCDC107, NAG6, ZDHHC14, EOMES, VIL2, WDR54, DMAP1, and PMPCA) were considered as DM-related atherogenesis genes (DRAGs). The Gene Set Variation Analysis (GSVA) score of the DRAG set gradually increased in the control, DM and DMCHD. ROC curve analysis showed that ZAP70, TSEN54, and PLEKHF1 may be potential blood circulation biomarkers for DMCHD in patients with DM. In conclusion, we identified nineteen hallmark genes involving in DM-related atherogenesis and constructed a potential transcriptional regulatory network involving in DM-related atherosclerosis.

Keywords: diabetes mellitus, coronary heart disease, hallmark gene set, WGCNA, atherogenesis

INTRODUCTION

Diabetes mellitus (DM) is a metabolic disorder of the endocrine system, clinically characterized by hyperglycemia as well as alterations in lipids, carbohydrates, and protein metabolism (1, 2). DM can be divided into type 1 and type 2, of which more than 90% of cases are type 2 (3, 4). Thus, only type 2 DM was focused in our present study. The increasing prevalence of type 2 DM means that 430 million people are expected to suffer the condition by 2030 (5, 6). DM renders patients more

susceptible to arterial atherosclerosis, which causes cardiovascular disease, especially coronary heart disease (CHD). CHD is the most serious complication of DM and a frequent cause of death among DM patients (7). Thus, elucidating why DM increases risk of CHD is of great importance.

Several studies have suggested that the presence of hypertension, dyslipidemia, and/or obesity in most patients with DM may help explain the increased incidence of CHD, since these factors also increase CHD risk in individuals without DM (8–14). Given that DM is multifactorial and quite heterogeneous (15), a comprehensive examination is required to understand key mechanisms driving development of CHD in patients.

To address this issue, we attempted to identify hub genes or gene sets involved in DM-associated CHD (DMCHD). Using weighted correlation network analysis (WGCNA), we identified 19 DM-related atherogenesis genes (DRAGs). Based on protein-protein interaction and functional enrichment analysis, we constructed a protein-protein pathway network to uncover the potential transcriptional regulatory network involved in DM-related atherosclerosis.

MATERIALS AND METHODS

Data Processing

We downloaded gene expression profiles from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/) (16) to analyze genetic differences between DM and DMCHD. We extracted the data set from the human whole blood gene expression profile of *GSE*34198 based on GPL6947 (17), which included 33 healthy control samples, 34 CHD samples, 15 DM samples and 15 DMCHD samples. In GSE34198, CHD patients refers to patients who eventually develop acute myocardial infarction and DM refers to type 2 diabetes. The clinical information including sex, age, body mass index and smoking status was provided in the **Supplementary Table 1**. GSE90074 (18) based on GPL6480 was as a validation set and contained human whole blood mononuclear cells gene expression profiles from 33 healthy control samples, 55 CHD samples, 17 DM samples, and 38 DMCHD samples. The *normalize Between Arrays* function in the *limma* package (19) was used to normalize the gene expression profiles. Probes that corresponded to multiple genes were removed. If a gene was detected using multiple probes, the expression level of the gene was calculated by averaging the expression from all the probes.

Screening Differently Expressed Genes (DEGs) in DM and DMCHD Compared to Healthy Control

To identify the genes involving in DM-related atherogenesis, the DEGs in DM and DMCHD compared to healthy control were screened using limma package. A gene with P adjusted by false discovery rate <0.05 was considered significantly differently expressed.

WGCNA

It is well-known that DM increases risk of CHD, thus, healthy control, DM, and DMCHD were considered as different stages of DM-related atherosclerosis. According to WGCNA official website (https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/), this method is suitable for identifying gene modules with the certain phenotype (DM-related atherosclerosis). DM increases the risk of CHD is considered the result of multiple molecules working together. WGCNA was used to identify the related gene modules (each module contain multiple molecules). Thus, WGCNA (20) was selected in the present study. The DEGs were subjected to perform a weighted gene expression network in GSE34198 using the WGCNA package (20) in R version 3.5 (https://www.r-project.org/). Candidate power (1 to 20) was used to test the average connectivity degrees of different modules and degrees of independence. If the degree of independence was >0.85, the power value was selected. The *hclust* function was used to cluster samples and check for outliers. Subsequently, a heat map was constructed to correlate gene modules with phenotype and statistical significance. High correlation meant that the genes of the corresponding module tended to be highly correlated with development of CHD in DM patients. In the present study, the module with the most significant positive correlation to DM-related atherogenesis was considered the candidate module.

Functional Enrichment Analysis of Gene Modules

In order to explore the signaling pathways and biological characteristics of related genes in phenotype-related modules, we conducted Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. *The clusterProfiler* package (21) in R were applied to these functional enrichment analysis. In addition, a biological processes network was constructed using the *ClueGO* plug-in (22) in *Cytoscape* (23). GSEA software (https://www.gsea-msigdb.org/gsea/downloads.jsp) was used to analyze gene set enrichment (24) based on the reference gene sets "c5.bp.v7.0.symbols.gmt" and "c2.cp.kegg.v7.0.symbols.gmt" (25). $P < 0.05$ was considered significant.

Construction of Protein-Protein KEGG Pathway Networks

The STRING database (https://string-db.org/) (26) was used to extract protein-protein interactions between candidate module genes. *Cytoscape* was used to construct a protein-protein network based on KEGG pathways involved in DM-related atherogenesis.

Identification of DM-Related Atherogenesis Genes (DRAGs) and Assessment of Predictive Potential

In the candidate module based on WGCNA, genes with significance (GS) > 0.2 and module membership (MM) > 0.7 were considered as DM-related atherogenesis genes (DRAGs). The *GSVA* package (27) was used to perform gene set variation analysis (GSVA) to score individual samples against the DRAG

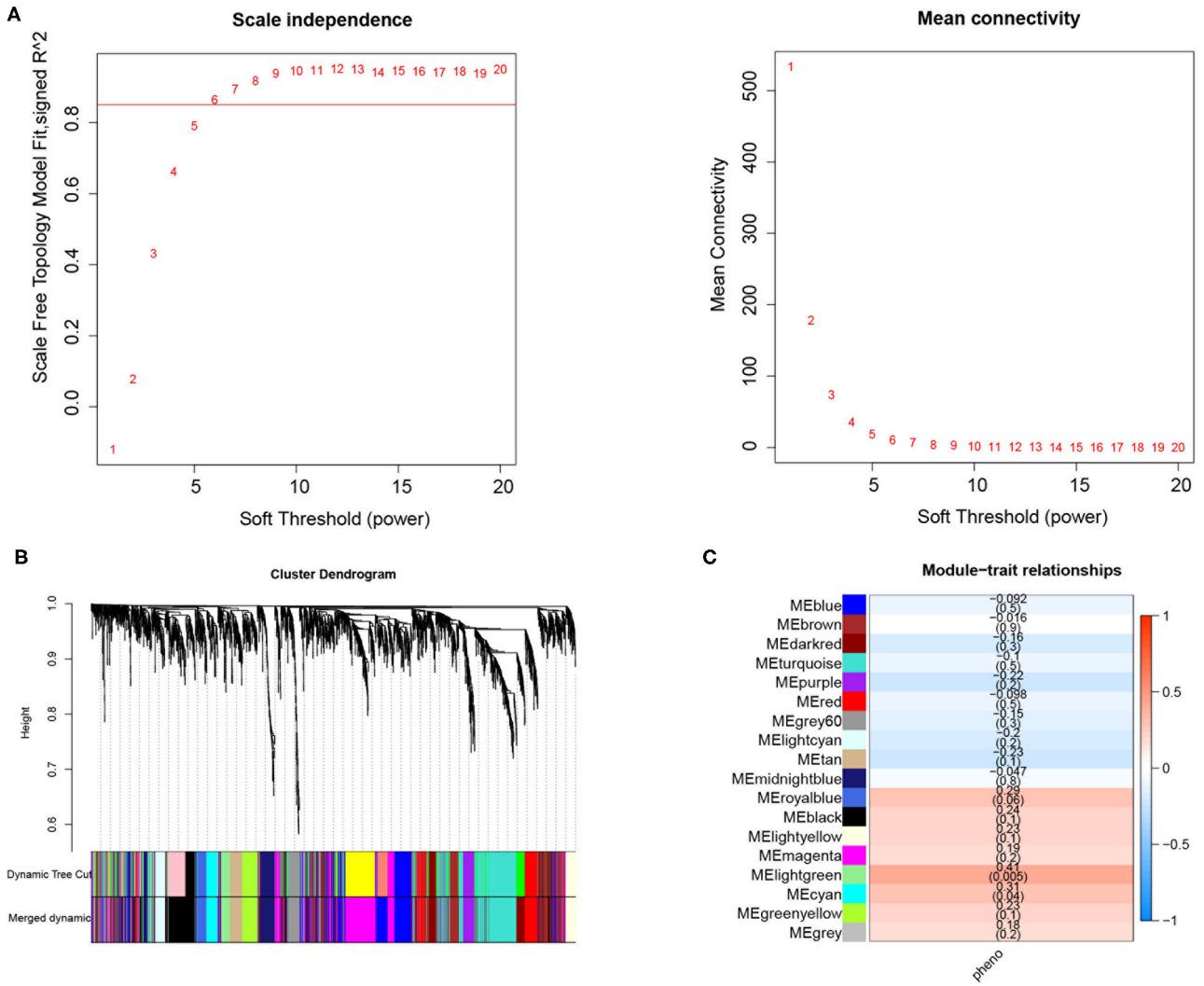


FIGURE 1 | Weighted correlation network analysis to obtained the module associated with the development of diabetes mellitus-related atherogenesis in 33 healthy control samples, 15 DM samples and 15 DMCHD samples. **(A)** Graphical representation showing the selection of soft-thresholding powers. **(B)** Hierarchical clustering dendogram. **(C)** Histogram showing module-trait correlations. Light green module was the modules which had the strongest correlation with diabetes mellitus-related atherogenesis. DM, diabetes mellitus. DMCHD, diabetes mellitus-induced coronary heart disease.

set, and the DM-related atherogenesis GSVA score was obtained for each sample. The ability of the DM-related atherogenesis GSVA score to predict DMCHD was evaluated using receiver operating characteristic (ROC) curve analysis with the *pROC* package (28). The aberrant expression of DRAGs were validated using the GSE90074 dataset. DRAGs with high predictive value were considered as biomarkers of DMCHD.

Cell Subpopulation Analysis

The host's response to DM and DMCHD may also be reflected in blood cell subsets. The webtool xCell (29) (https://xcell.ucsf.edu/) was used to perform cell type enrichment analysis from gene expression data. Use analysis of variance to compare the cell subpopulations abundance in control, DM, and DMCHD.

RESULTS

Determining DRAG Modules

In the current study, a total of 2,442 DEGs were screened with 1,029 up-regulated and 1,413 down-regulated in in DM and DMCHD compared to healthy control. The expression profiles of the 2,442 DEGs were subjected to WGCNA. In the WGCNA, according to the results of the scale-free fit index and average connectivity, the soft-thresholding was determined as six (**Figure 1A**). Based on hierarchical clustering, the genes were divided into 18 modules (**Figure 1B**). The light green module was involved in immune- and T cell activity-related biological processes, and displayed the most significant positive correlation between modules and DM-related atherogenesis ($r = 0.41$, $p = 0.005$) (**Figure 1C**).

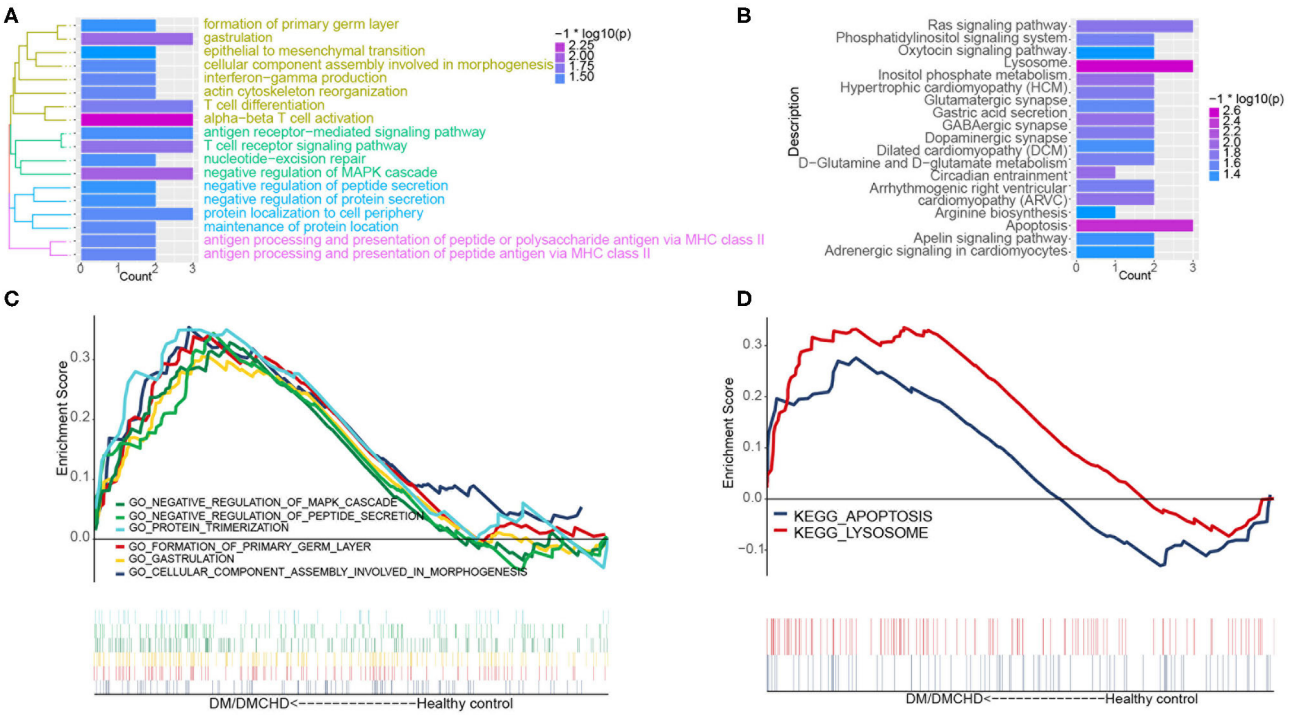


FIGURE 2 | The exploration of DM and DMCHD-related biological process and pathways using functional enrichment analysis. **(A)** Significantly enriched biological processes corresponding to genes in the light green module. $P < 0.05$ was considered significant. **(B)** Significant enrichment of KEGG pathways for genes in the light green module. $P < 0.05$ was considered significant. **(C)** Biological processes identified by gene set enrichment analysis and found to be significantly enriched in samples from patients with diabetes mellitus or both diabetes mellitus (DM) and coronary heart disease (DMCHD). KEGG, Kyoto Encyclopedia of Genes and Genomes. **(D)** Biological processes determined from gene clusters. DM, diabetes mellitus; DMCHD, diabetes mellitus and coronary heart disease.

It was therefore considered as the candidate module in subsequent analyses.

Biological Functions Involving the Candidate Module

The light green module contained genes significantly involved in biological processes related to cytoskeletal anchoring at plasma membrane, protein localization to cell cortex and protein deneddylation (**Figure 2A**). Genes within the module were associated with autophagy, phosphatidylinositol biosynthetic processes and G protein-coupled receptor signaling pathways (**Supplementary Figure 1**). Moreover, the light green module was significantly involved in pathways involving lysosomes, apoptosis, Ras signaling pathways and inositol phosphate metabolism (**Figure 2B**). DM and DMCHD samples showed enrichment in gastrulation and protein trimerization (**Figure 2C**), as well as lysosome and apoptosis pathways (**Figure 2D**).

Interaction Between Proteins Encoded by DM and CHD

A protein-protein KEGG pathway network was constructed to provide a potential mechanism explaining how DM increases risk of CHD. The network included 51 module-related genes and 18 KEGG pathways (**Figure 3A**). Previous studies proposed that arginine biosynthesis, GABAergic synapse and gastric acid secretion pathways overlap between DM and CHD (30–32). In this study, we elucidated the following pathways potentially involved in DMCHD: adrenergic signaling in cardiomyocytes, apelin signaling, circadian entrainment, oxytocin signaling pathway, apoptosis and Ras signaling (**Figure 3B**). The arenergic signaling in cardiomyocytes was activated by Ca^{2+} signaling, cell phosphorylation was mediated by CAMKII and CACNA2D2. In addition, we found that CNGT2 could affect cell proliferation. The apelin signaling pathway might cause vasodilatation and cardiovascular development. The G protein-coupled receptor APJ and its cognate ligand what apelin were widely expressed throughout the human body. They implicated in different key physiological processes such as angiogenesis, cardiovascular functions and energy metabolism regulation. Moreover, they also involved in the development and progression of different pathologies including diabetes, obesity and cardiovascular disease (33). In the circadian entrainment, presynaptic RGC neuron secreted L-glutamate and PACAP. Furthermore, PAC1 could activate c-Fos by CNGT2. In the oxytocin signaling pathway, CACNA2D2 and CALM3 could indirectly affect cardiovascular system and renal epithelial cell. Besides, the oxytocin signaling pathway was also activated by Ca2+ signaling. In the apoptosis, sphingosine and lysosomotropic detergents

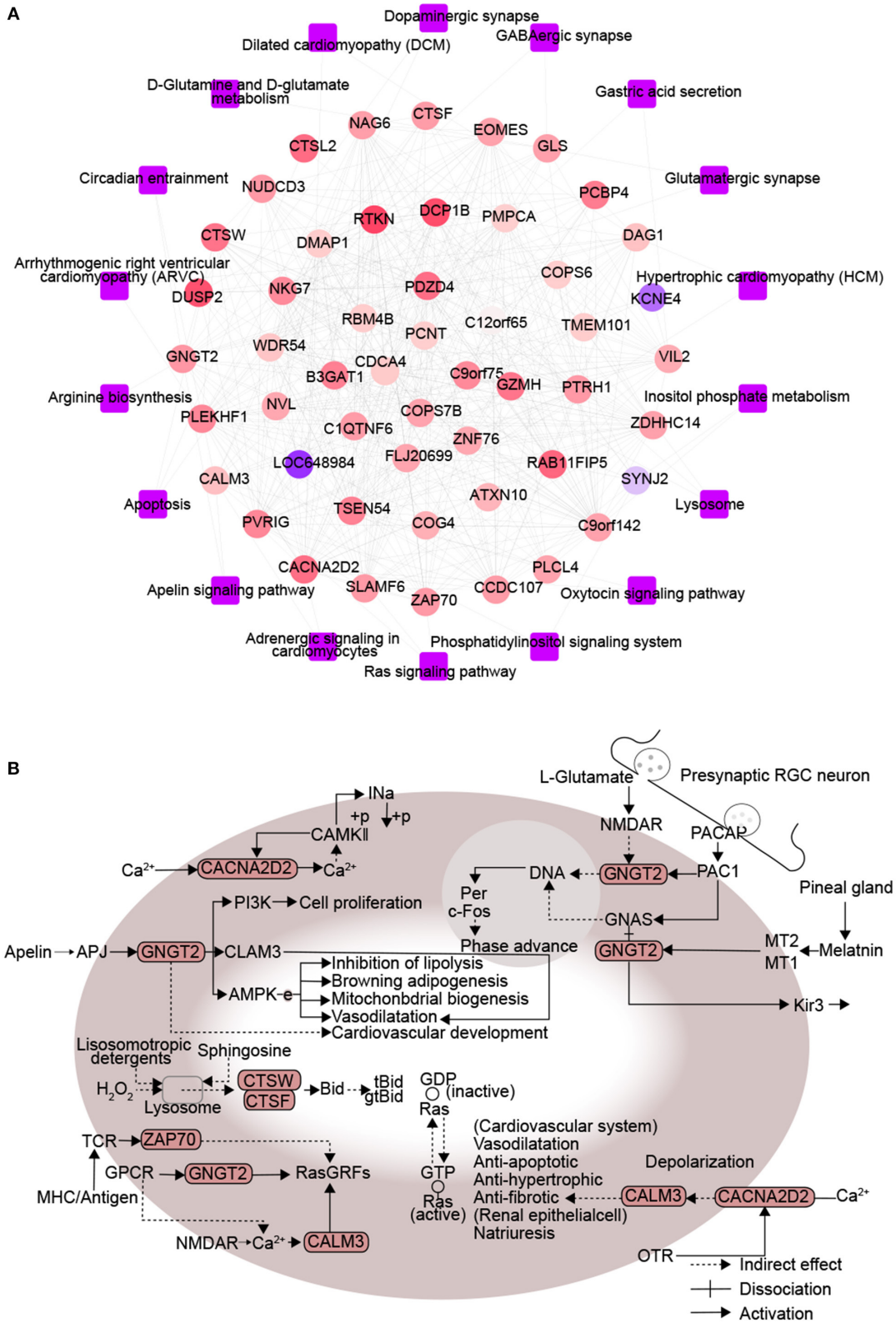


FIGURE 3 | Protein interactions associated with signaling pathways enriched in diabetes mellitus-associated coronary heart disease. **(A)** Network diagram of protein-protein-KEGG pathway related to DMCAD. The ellipse for the module gene. The round rectangle for the KEGG pathway. The color of gene nodes represents MM, and the color at the edge of gene nodes represents modules. **(B)** CALM3, ZAP70, GNGT2, CACNA2D2, CTSF, and CTSW-related KEGG pathways. DMCHD, diabetes mellitus-induced coronary heart disease. DRAG, DM-related atherogenesis gene. MM, module membership.

and H2O2 could indirectly affect lysosome. CTSW and CTSF could be indirectly affected by lysosome, thereby affecting cell apoptosis. In the Ras signaling pathway, CNGT2 and CALM3 could activate GDP-Ras inactive to GTP-Ras active process by RasGRFs. Moreover, ZAP70 could also indirectly affect the process by RasGRFs.

The Expression of DRAGs and ROC Analysis of DM-Related Atherogenesis GSVA Score

We identified the following genes from the light green module as significant DRAG candidates: *RTKN*, *DCP1B*, *PDZD4*, *CACNA2D2*, *TSEN54*, *PVRIG*, *PLEKHF1*, *NKG7*, *ZAP70*, *NUDCD3*, *SLAMF6*, *CCDC107*, *NAG6*, *ZDHHC14*, *EOMES*, *VIL2*, *WDR54*, *DMAP1*, and *PMPCA*. Their expression progressively increased in healthy control, DM and DMCHD (**Figure 4A**), the differential expression of these genes in three different groups healthy control, DM, DMCHD might be the reason which lead to DM patients with coronary heart disease, but whether this is the cause of this result needs further exploration. ROC curves showed that the DM-related atherogenesis GSVA score can predict DMCHD, with an area under the ROC curve of 0.743 (**Figure 4B**).

Validating Aberrant Expression of DRAGs

To validate the DRAGs we identified, we examined their expression levels in the GSE90074 dataset. We confirmed that *ZAP70*, *PLEKHF1,* and *TSEN54* expression significantly progressively increased in healthy control, DM and DMCHD (**Figures 5A–D**), these genes were considered to be the progression genes of DM complicated with CHD. ROC curve analysis suggested that all three genes may be potential biomarkers of CHD development in DM (**Figure 5E**).

Various Cell Types Vary in Control, DM and DMCHD

After estimating and comparing the cell subpopulation abundance, we found the various cell types vary in control, DM and DMCHD. B cells and Macrophages (M2) showed an increasing trend, while eosinophils, common lymphoid progenitor cell (CLP), and granulocyte macrophage progenitor (GMP) showed a downward trend. Osteoblast, Gamma delta T cells (Tgd cells), and endothelial cells are also different in control, DM and DMCHD (**Figure 6**).

DISCUSSION

DM is a syndrome characterized by chronic hyperglycemia accompanied by disorder of lipid and protein metabolism. The pathogenesis of diabetes includes insulin resistance and insulin secretion disorders of islet beta cells. CHD is a kind of heart disease caused by atherosclerotic lesion of coronary artery, which leads to stenosis or obstruction of vascular lumen and myocardial ischemia, hypoxia or necrosis. DM and CHD are common chronic diseases (34). People with DM have a much higher incidence of CHD because of dyslipidemia and metabolic disorders, particularly in the formation of glycated proteins, causing inflammation and artery stiffening as well as dysregulating hormones, immune responses and dendritic cell activity (35). However, the genetic changes that underlie development of CHD in patients with DM is not fully understood. In the present study, we performed WGCNA in a whole-blood gene expression profile to screen for gene modules involved in the pathogenesis of DM-related atherogenesis. We found that a light green module, a cluster of immune system genes, had the strongest correlation with emergence of CHD in DM. Gene sets associated with apoptosis and lysosomes were particularly enriched in DM and DMCHD samples. These findings are consistent with a known link between apoptosis and CHD (36) and between lysosome activity and DMCHD (37).

A protein-protein KEGG pathway network was constructed to help reveal the molecular mechanism of DMCHD. Among the six pathways identified, previous work has already implicated apelin signaling (38), apoptosis (39), and Ras signaling (40) in DMCHD. In addition, our study suggests that adrenergic signaling in cardiomyocytes, circadian entrainment and oxytocin signaling pathway are also related to DMCHD. *CACNA2D2* and *GNGT2* play important roles in regulating these pathways. Expression of *CACNA2D2* in the adrenergic signaling pathway elevates intracellular Ca^{2+} in cardiomyocytes, enhancing myocardial contractility (41), and more generally changes in gene expression have been associated with Ca^{2+} mobilization in CHD (42). Regulation of *GNGT2* affects circadian entrainment, in which c-Fos may be activated to regulate apoptosis, although this is controversial (43, 44). The activation of c-Fos in gestational diabetes suggests a link to glucose metabolism, which may therefore be linked to DMCHD, which requires further study. Notably, oxytocin promotes glucose metabolism in cultured cardiomyocytes from newborn and adult rats (45). Furthermore, it can directly improve pancreatic functions (46), presumably by increasing insulin early (47). Oxytocin signaling pathway is also activated by Ca^{2+} signaling in a *CACNA2D2*-dependent manner. These data support targeting oxytocin in DM patients in order to prevent CHD.

DM is a complicated, polygenic disease that exerts pleotropic effects that may increase risk of other diseases, such as CHD. Therefore, it is necessary to identify hallmark genes that explain potential mechanism of DM-related atherogenesis. We identified 19 genes in the light green module as DRAGs (*RTKN*, *DCP1B*, *PDZD4*, *CACNA2D2*, *TSEN54*, *PVRIG*, *PLEKHF1*, *NKG7*, *ZAP70*, *NUDCD3*, *SLAMF6*, *CCDC107*, *NAG6*, *ZDHHC14*, *EOMES*, *VIL2*, *WDR54*, *DMAP1,* and *PMPCA*), the differential expression of these genes in three different groups healthy control, DM, DMCHD might be the reason which lead to DM patients with coronary heart disease, but whether this is the cause of this result needs further exploration. We found that the DM-related atherogenesis GSVA score progressively increased in healthy control, DM and DMCHD samples. ROC analysis showed that DM-related atherogenesis GSVA score may have predictive potential for identifying genes responsible for DMCHD development. Our work suggests that the DRAGs *ZAP70*, *PLEKHF1,* and *TSEN54* may be biomarkers for CHD

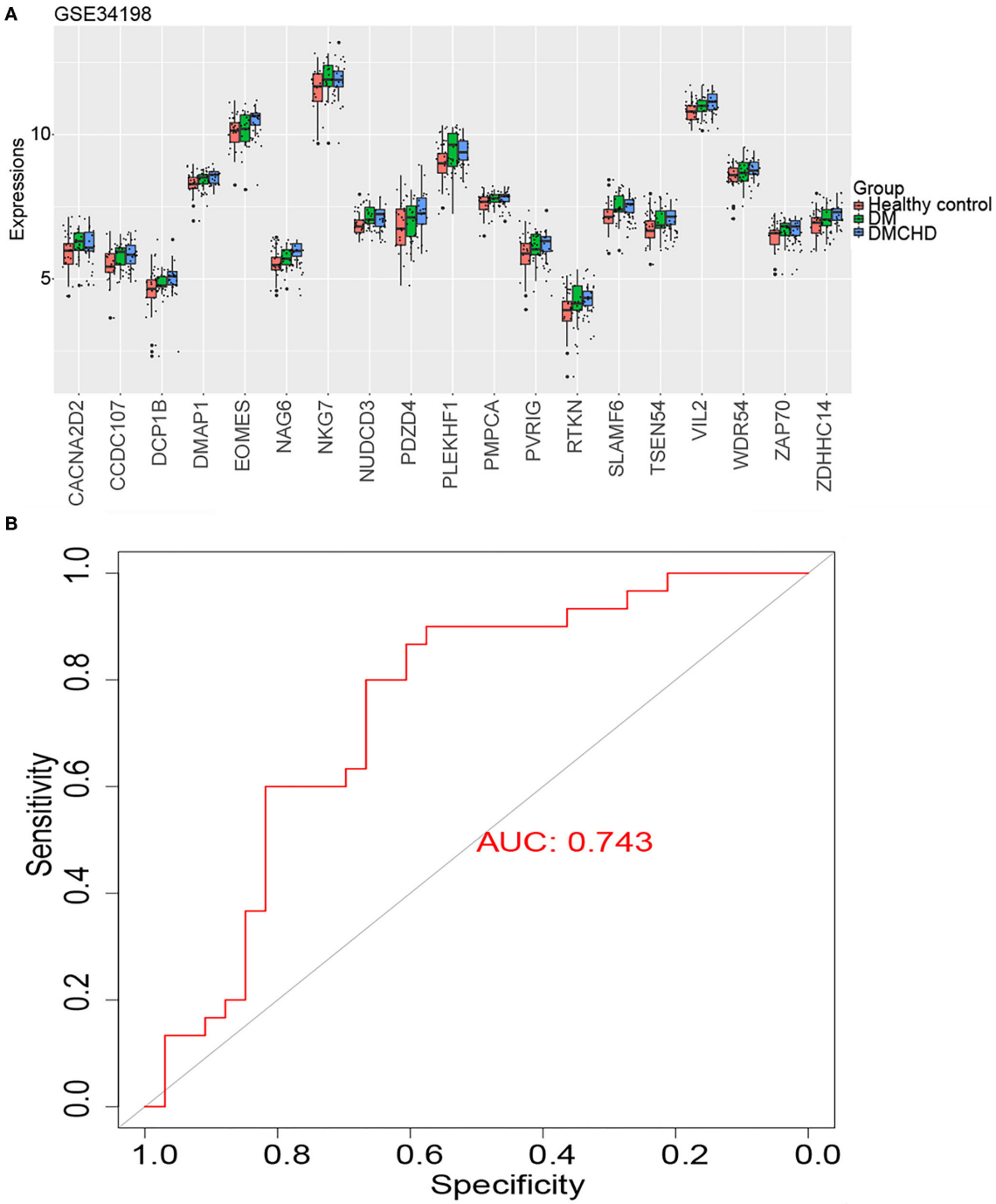


FIGURE 4 | Expression of diabetes mellitus-related atherogenesis genes (DRAGs) and their ability to predict diabetes mellitus-associated coronary heart disease (DMCHD). **(A)** The expression level of 19 DRAGs (genes with significance (GS) > 0.2 and module membership (MM) > 0.7 were considered as DRAGs) in tissue from healthy individuals and from patients with diabetes mellitus (DM) or DMCHD. **(B)** Receiver operating characteristic curve analysis of DM-related atherogenesis GSVA score for predicting DMCHD.

in DM patients. In addition, we also found some types of cells in whole blood showed upward or downward trend in control, DM and DMCHD. However, the abundance of the cell subsets will be affected by many factors, and whether it can be used as a specific marker of DM or DMCHD still needs further exploration.

There are several notable limitations in our study. First, we took a bioinformatics approach to mine potential molecular

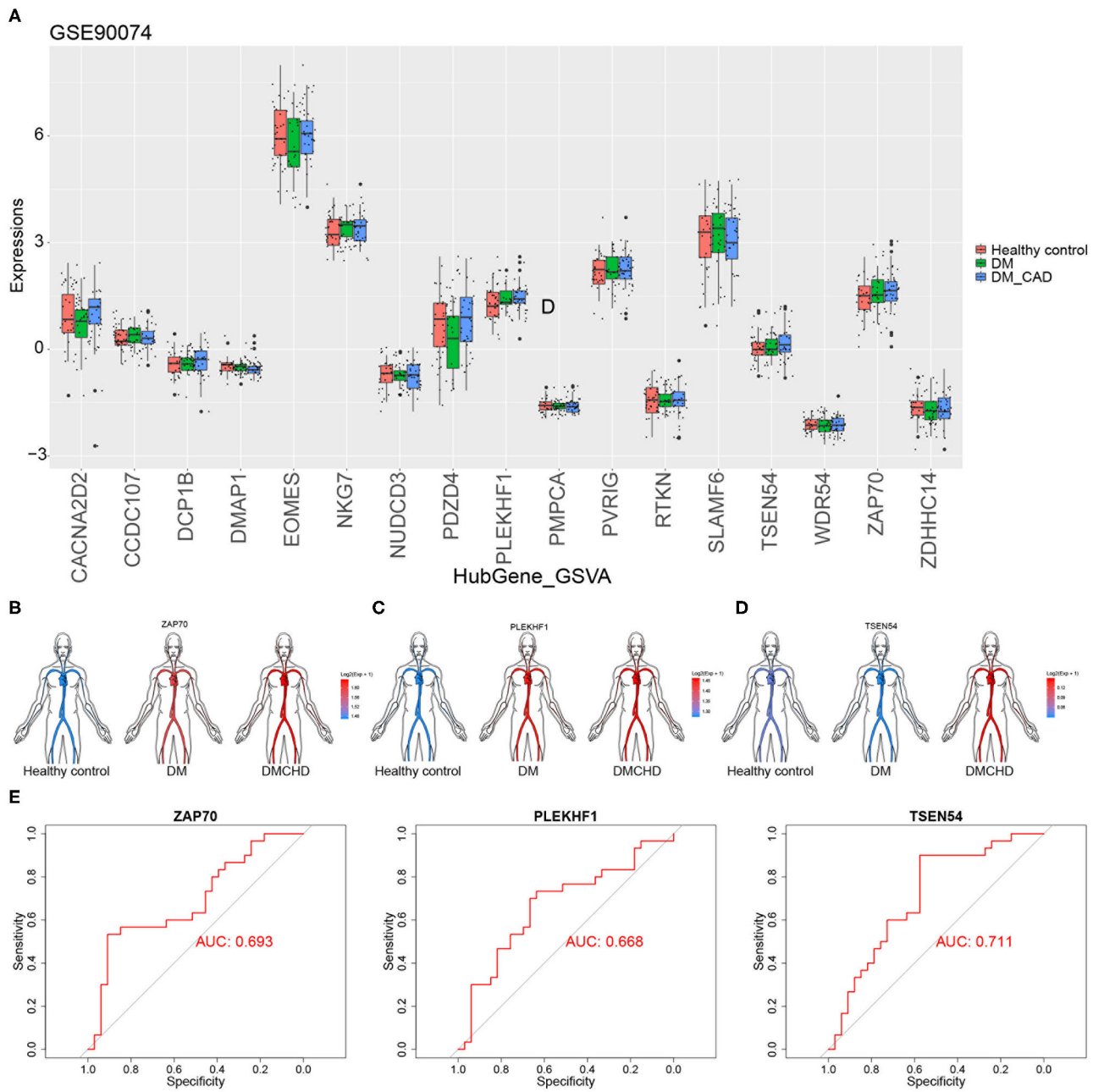


FIGURE 5 | Validation of diabetes mellitus-related atherogenesis genes (DRAGs). **(A)** The expression of DRAGs in tissue samples from healthy controls or patients with diabetes mellitus (DM) or both diabetes mellitus and coronary heart disease (DMCHD) in a validation data set (GSE90074). Expression levels in GSE34198 of **(B)** *ZAP70*, **(C)** *PLEKHF1*, and **(D)** *TSEN54* in human circulation (https://github.com/jespermaag/gganatogram). **(E)** Receiver operating characteristic curves of *ZAP70*, *PLEKHF1* and *TSEN54* as potential biomarkers for predicting coronary heart disease in patients with diabetes mellitus.

mechanisms causing DMCHD. Our findings will need to be followed up with mechanistic studies. Second, using GSVA to assess presence of genetic risk factors in individual patients is cost-prohibitive and therefore not applicable in clinics yet. Third, only one dataset was used for validation. Whether *ZAP70*, *PLEKHF1*, and *TSEN54* are true biomarkers for DMCHD requires further examination in additional, larger data sets. In addition, it is worth noting that the whole blood gene expression profiles were used in the present study. We also found various cell types vary in control, DM and DMCHD, however, we can not elaborate on the interactions between cells in the DM-related atherogenesis in the present study.

In conclusion, we identified 19 hallmark genes of DM-related atherogenesis, which might be the biomarkers for

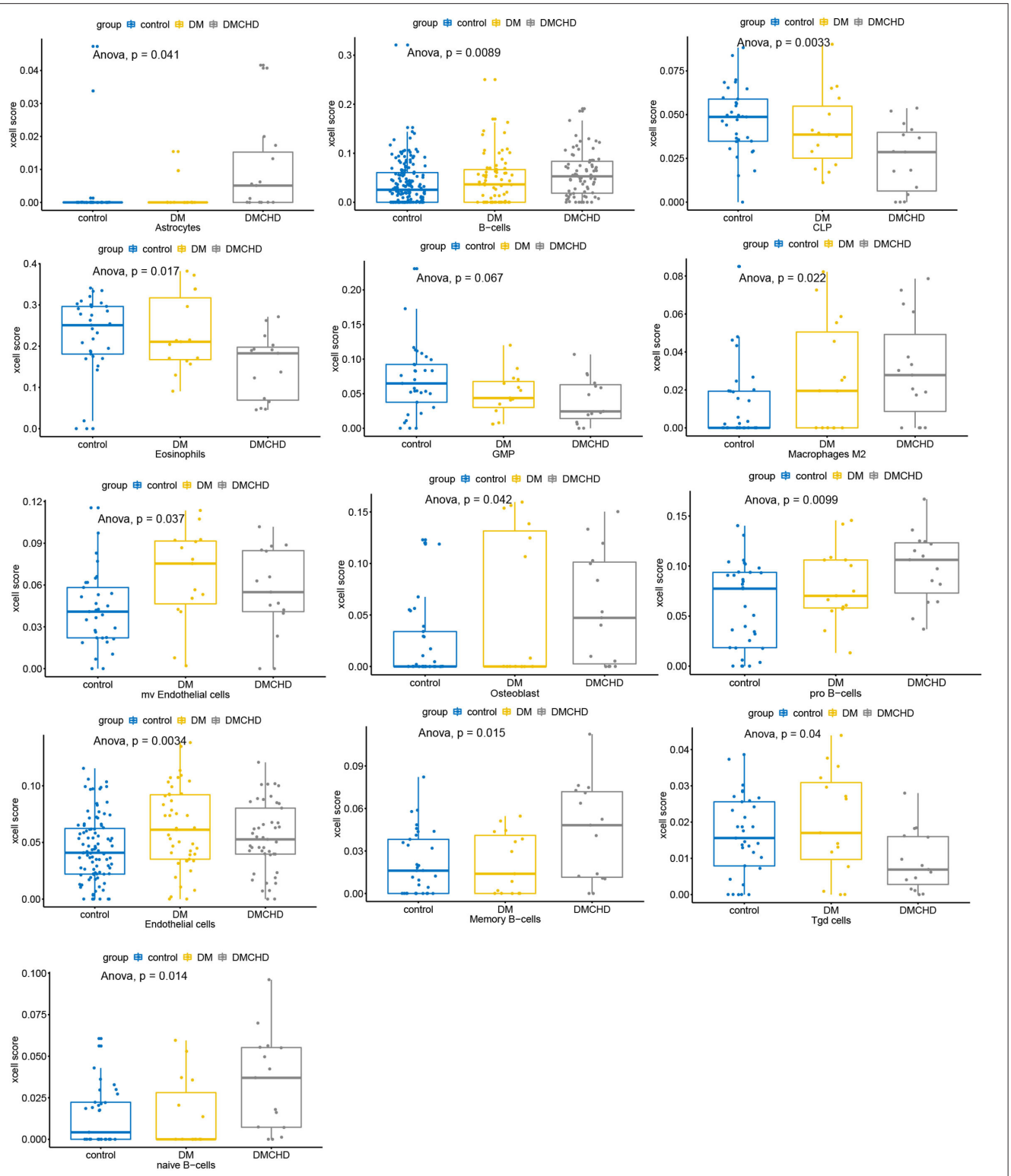


FIGURE 6 | Various cell types vary in control, DM and DMCHD. DM, diabetes mellitus, DMCHD, diabetes mellitus and coronary heart disease, CLP, common lymphoid progenitor cell; GMP, granulocyte macrophage progenitor; Tgd cells, gamma delta T cells; mv endothelial cells, mitral valve endothelial cells. Anova, analysis of variance.

identifying DM and CHD. The DM-related atherogenesis GSVA score may predict susceptibility of DM patients to developing CHD. Our protein-protein KEGG pathway network may reveal molecular mechanisms underlying DMCHD development. These results provide numerous leads for future studies of CHD risk factors in DM, and may provide a theoretical basis for the treatment of DM and CHD patients.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

All authors participated in the design, interpretation of the studies, analysis of the data, and review of the manuscript. QH and GD conducted the experiments. RW supplied critical reagents. DZ wrote the manuscript. JW contributed to critically revise the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Project of Nanning Scientific Research and Technology Development Plan (20193252), the Nanning Qingxiu District Science and Technology Plan Project (2018029), the High-Level Medical Expert Training Program of Guangxi 139 Plan Funding (G201903049), and Guangxi medical and health key discipline construction project. This study was also sponsored by the Nanning Excellent Young Scientist Program and the Guangxi Beibu Gulf Economic Zone Major Talent Program (RC20190103).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2020.580573/full#supplementary-material

Supplementary Figure 1 | Biological process enrichment analysis using *ClueGO* plugin for genes in the light green module.

REFERENCES

1. Ramachandran S, Rajasekaran A, Manisenthilkumar KT. Investigation of hypoglycemic, hypolipidemic and antioxidant activities of aqueous extract of Terminalia paniculata bark in diabetic rats. *Asian Pac J Trop Biomed.* (2012) 2:262–8. doi: 10.1016/S2221-1691(12)60020-3
2. Rashidi AA, Mirhashemi SM, Taghizadeh M, Sarkhail P. Iranian medicinal plants for diabetes mellitus: a systematic review. *Pak J Biol Sci.* (2013) 16:401–11. doi: 10.3923/pjbs.2013.401.411
3. Yang SH, Dou KF, Song WJ. Prevalence of diabetes among men and women in China. *N Engl J Med.* (2010) 362:2425–6. doi: 10.1056/NEJMoa0908292
4. Patil P, Mandal S, Tomar SK, Anand S. Food protein-derived bioactive peptides in management of type 2 diabetes. *Eur J Nutr.* (2015) 54:863–80. doi: 10.1007/s00394-015-0974-2
5. Chang MH, Valdez R, Ned RM, Liu T, Yang Q, Yesupriya A, et al. Influence of familial risk on diabetes risk-reducing behaviors among U.S. adults without diabetes. *Diabetes Care.* (2011) 34:2393–9. doi: 10.2337/dc11-0876
6. Mustafa SB, Mehmood Z, Akhter N, Kauser A, Hussain I, Rashid A, et al. Review-medicinal plants and management of diabetes mellitus: a review. *Pak J Pharm Sci.* (2016) 29:1885–91.
7. Ullah F, Afridi AK, Rahim F, Ur Rahman S, Ashfaq M. Efficacy of 5 Mg and 10 Mg rosuvastatin in Type-2 diabetes mellitus with hypercholesteroalemia. *J Ayub Med Coll Abbottabad.* (2015) 27:564–8.
8. Defronzo RA. Insulin resistance, hyperinsulinemia, and coronary artery disease: a complex metabolic web. *J Cardiovasc Pharmacol.* (1992) 20(Suppl. 11):S1–16. doi: 10.1097/00005344-199200111-00002
9. Solymoss BC, Marcil M, Chaour M, Gilfix BM, Poitras AM, Campeau L. Fasting hyperinsulinism, insulin resistance syndrome, and coronary artery disease in men and women. *Am J Cardiol.* (1995) 76:1152–6. doi: 10.1016/S0002-9149(99)80326-9
10. Despres JP, Lamarche B, Mauriege P, Cantin B, Dagenais GR, Moorjani S, et al. Hyperinsulinemia as an independent risk factor for ischemic heart disease. *N Engl J Med.* (1996) 334:952–7. doi: 10.1056/NEJM199604113341504
11. Grossman E, Messerli FH. Diabetic and hypertensive heart disease. *Ann Intern Med.* (1996) 125:304–10. doi: 10.7326/0003-4819-125-4-199608150-00009
12. Roh E, Ko SH, Kwon HS, Kim NH, Kim JH, Kim CS, et al. Prevalence and management of dyslipidemia in Korea: Korea national health and nutrition examination survey during 1998 to 2010. *Diabetes Metab J.* (2013) 37:433–49. doi: 10.4093/dmj.2013.37.6.433
13. Kim CS, Ko SH, Kwon HS, Kim NH, Kim JH, Lim S, et al. Prevalence, awareness, and management of obesity in Korea: data from the Korea national health and nutrition examination survey (1998-2011). *Diabetes Metab J.* (2014) 38:35–43. doi: 10.4093/dmj.2014.38.1.35
14. Ko SH, Kwon HS, Kim DJ, Kim JH, Kim NH, Kim CS, et al. Higher prevalence and awareness, but lower control rate of hypertension in patients with diabetes than general population: the fifth korean national health and nutrition examination survey in 2011. *Diabetes Metab J.* (2014) 38:51–7. doi: 10.4093/dmj.2014.38.1.51
15. Zimmet P, Alberti KG, Shaw J. Global and societal implications of the diabetes epidemic. *Nature.* (2001) 414:782–7. doi: 10.1038/414782a
16. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res.* (2013) 41:D991–5. doi: 10.1093/nar/gks1193
17. Jiang YX, Jing LD, Jia YH. Clinical characteristics and risk factors of left ventricular thrombus after acute myocardial infarction: a matched case-control study. *Chin Med J.* (2015) 128:2415–9. doi: 10.4103/0366-6999.164869
18. Ravi S, Schuck RN, Hilliard E, Lee CR, Dai X, Lenhart K, et al. Clinical evidence supports a protective role for CXCL5 in coronary artery disease. *Am J Pathol.* (2017) 187:2895–911. doi: 10.1016/j.ajpath.2017.08.006
19. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* (2015) 43:e47. doi: 10.1093/nar/gkv007
20. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* (2008) 9:559. doi: 10.1186/1471-2105-9-559
21. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* (2012) 16:284–7. doi: 10.1089/omi.2011.0118
22. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics.* (2009) 25:1091–3. doi: 10.1093/bioinformatics/btp101

23. Kohl M, Wiese S, Warscheid B. Cytoscape: software for visualization and analysis of biological networks. *Methods Mol Biol.* (2011) 696:291–303. doi: 10.1007/978-1-60761-987-1_18
24. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102
25. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst.* (2015) 1:417–25. doi: 10.1016/j.cels.2015.12.004
26. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* (2017) 45:D362–8. doi: 10.1093/nar/gkw937
27. Herranz Amo F, Rivero Sanchez E, Verdu Tartajo F, Hernandez Fernandez C, Diez Cordero JM. [Cysts of the seminal vesicles: apropos of a case]. *Actas Urol Esp.* (1987) 11:210–3
28. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* (2011) 12:77. doi: 10.1186/1471-2105-12-77
29. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* (2017) 18:220. doi: 10.1186/s13059-017-1349-1
30. Feldman M, Corbett DB, Ramsey EJ, Walsh JH, Richardson CT. Abnormal gastric function in longstanding, insulin-dependent diabetic patients. *Gastroenterology.* (1979) 77:12–7. doi: 10.1016/S0016-5085(79)80003-7
31. Ramirez-Zamora S, Mendez-Rodriguez ML, Olguin-Martinez M, Sanchez-Sevilla L, Quintana-Quintana M, Garcia-Garcia N, et al. Increased erythrocytes by-products of arginine catabolism are associated with hyperglycemia and could be involved in the pathogenesis of type 2 diabetes mellitus. *PLoS ONE.* (2013) 8:e66823. doi: 10.1371/journal.pone.0066823
32. Wake H, Ortiz FC, Woo DH, Lee PR, Angulo MC, Fields RD. Nonsynaptic junctions on myelinating glia promote preferential myelination of electrically active axons. *Nat Commun.* (2015) 6:7844. doi: 10.1038/ncomms8844
33. Chaves-Almagro C, Castan-Laurell I, Dray C, Knauf C, Valet P, Masri B. Apelin receptors: from signaling to antidiabetic strategy. *Eur J Pharmacol.* (2015) 763:149–59. doi: 10.1016/j.ejphar.2015.05.017
34. Zhou T, Guan H, Yao J, Xiong X, Ma A. The quality of life in Chinese population with chronic non-communicable diseases according to EQ-5D-3L: a systematic review. *Qual Life Res.* (2018) 27:2799–814. doi: 10.1007/s11136-018-1928-y
35. Price CL, Knight SC. Advanced glycation: a novel outlook on atherosclerosis. *Curr Pharm Des.* (2007) 13:3681–7. doi: 10.2174/138161207783018608
36. Dong Y, Chen H, Gao J, Liu Y, Li J, Wang J. Molecular machinery and interplay of apoptosis and autophagy in coronary heart disease. *J Mol Cell Cardiol.* (2019) 136:27–41. doi: 10.1016/j.yjmcc.2019.09.001
37. Reiner Z, Guardamagna O, Nair D, Soran H, Hovingh K, Bertolini S, et al. Lysosomal acid lipase deficiency–an under-recognized cause of dyslipidaemia and liver dysfunction. *Atherosclerosis.* (2014) 235:21–30. doi: 10.1016/j.atherosclerosis.2014.04.003
38. Kidoya H, Takakura N. Biology of the apelin-APJ axis in vascular formation. *J Biochem.* (2012) 152:125–31. doi: 10.1093/jb/mvs071
39. Wang F, Fisher SA, Zhong J, Wu Y, Yang P. Superoxide dismutase 1 *in vivo* ameliorates maternal diabetes mellitus-induced apoptosis and heart defects through restoration of impaired wnt signaling. *Circ Cardiovasc Genet.* (2015) 8:665–76. doi: 10.1161/CIRCGENETICS.115.001138
40. Zhang Z, Zhang HY, Zhang Y, Li H. Inactivation of the Ras/MAPK/PPARgamma signaling axis alleviates diabetic mellitus-induced erectile dysfunction through suppression of corpus cavernosal endothelial cell apoptosis by inhibiting HMGCS2 expression. *Endocrine.* (2019) 63:615–31. doi: 10.1007/s12020-018-1810-2
41. Medzikovic L, Schumacher CA, Verkerk AO, Van Deel ED, Wolswinkel R, Van Der Made I, et al. Orphan nuclear receptor Nur77 affects cardiomyocyte calcium homeostasis and adverse cardiac remodelling. *Sci Rep.* (2015) 5:15404. doi: 10.1038/srep15404
42. Wan X, Belanger K, Widen SG, Kuyumcu-Martinez MN, Garg NJ. Genes of the cGMP-PKG-Ca(2+) signaling pathway are alternatively spliced in cardiomyopathy: role of RBFOX2. *Biochim Biophys Acta Mol Basis Dis.* (2020) 1866:165620. doi: 10.1016/j.bbadis.2019.165620
43. Gajate C, Alonso MT, Schimmang T, Mollinedo F. C-Fos is not essential for apoptosis. *Biochem Biophys Res Commun.* (1996) 218:267–72. doi: 10.1006/bbrc.1996.0047
44. Kermani ES, Nazari Z, Mehdizadeh M, Shahbazi M, Golalipour MJ. Gestational diabetes influences the expression of hypertrophic genes in left ventricle of rat's offspring. *Iran J Basic Med Sci.* (2018) 21:525–8. doi: 10.22038/IJBMS.2018.25116.6233
45. Florian M, Jankowski M, Gutkowska J. Oxytocin increases glucose uptake in neonatal rat cardiomyocytes. *Endocrinology.* (2010) 151:482–91. doi: 10.1210/en.2009-0624
46. Watanabe S, Wei FY, Matsunaga T, Matsunaga N, Kaitsuka T, Tomizawa K. Oxytocin protects against stress-induced cell death in murine pancreatic beta-cells. *Sci Rep.* (2016) 6:25185. doi: 10.1038/srep25185
47. Klement J, Ott V, Rapp K, Brede S, Piccinini F, Cobelli C, et al. Oxytocin improves beta-cell responsivity and glucose tolerance in healthy men. *Diabetes.* (2017) 66:264–71. doi: 10.2337/db16-0569

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 20 January 2021
doi: 10.3389/fcvm.2020.617561

Genetics and Clinical Features of Noncompaction Cardiomyopathy in the Fetal Population

Hairui Sun [1,2,3†], Xiaoyan Hao [1,2†], Xin Wang [1,2†], Xiaoxue Zhou [1,2], Ye Zhang [1,2], Xiaowei Liu [1,2], Jiancheng Han [1,2], Xiaoyan Gu [1,2], Lin Sun [1,2], Ying Zhao [1,2], Tong Yi [1,2], Hongjia Zhang [1,3] and Yihua He [1,2,3*]*

[1] Beijing Anzhen Hospital, Capital Medical University, Beijing, China, [2] Beijing Key Laboratory of Maternal-Fetal Medicine and Fetal Heart Disease, Beijing Anzhen Hospital, Capital Medical University, Beijing, China, [3] Key Laboratory of Medical Engineering for Cardiovascular Disease, Ministry of Education, Beijing, China

OPEN ACCESS

Edited by:
Seitaro Nomura,
The University of Tokyo, Japan

Reviewed by:
Keiichi Hirono,
University of Toyama, Japan
Steven Clive Greenway,
University of Calgary, Canada

***Correspondence:**
Hongjia Zhang
zhanghongjia722@hotmail.com
Yihua He
heyihuaecho@hotmail.com

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: 15 October 2020
Accepted: 16 December 2020
Published: 20 January 2021

Citation:
Sun H, Hao X, Wang X, Zhou X, Zhang Y, Liu X, Han J, Gu X, Sun L, Zhao Y, Yi T, Zhang H and He Y (2021) Genetics and Clinical Features of Noncompaction Cardiomyopathy in the Fetal Population. Front. Cardiovasc. Med. 7:617561. doi: 10.3389/fcvm.2020.617561

Objectives: Noncompaction Cardiomyopathy (NCCM) has been classified as primary genetic cardiomyopathy and has gained increasing clinical awareness; however, little is known about NCCM in the fetal population. We aimed to investigate the clinical characteristics and genetic spectrum of a fetal population with NCCM.

Methods: We retrospectively reviewed all fetuses with a prenatal diagnosis of NCCM at a single center between October 2010 and December 2019. These cases were investigated for gestational age at diagnosis, gender, left or biventricular involvement, associated cardiac phenotypes, outcomes, and genetic testing data.

Results: We identified 37 fetuses with NCCM out of 49,898 fetuses, indicating that the incidence of NCCM in the fetal population was 0.07%. Of the 37 fetuses, 26 were male, ten were female and one was of unknown gender. NCCM involvement biventricle is the most common (n = 16, 43%), followed by confined to the left ventricle (n = 14, 38%). Nineteen (51%) had additional congenital heart defects, with right-sided lesions being the most common (n = 14, 74%), followed by ventricular septal defects (n = 10, 53%). Hydrops fetalis was present in 12 cases (32%), of which four were atypical (pericardial effusion only). Sequencing analysis was performed at autopsy (n = 19) or postnatally (n = 1) on 20 fetuses. Of the 20 fetuses undergoing copy number variation sequencing and whole-exome sequencing, nine (47%) had positive genetic results, including one with a pathogenic copy number variant and eight with pathogenic/likely pathogenic variants. Non-sarcomere gene mutations accounted for the vast majority (n = 7). In contrast, sarcomere gene mutations occurred in only one case (TPM1), and no mutations were identified in the three most common sarcomere genes (MYH7, TTN, and MYBPC3) of pediatric and adult patients. Pathogenic/likely pathogenic variants were significantly more frequent in fetuses with congenital heart defects than those without congenital heart defects.

Conclusions: Our data demonstrate that fetal NCCM is a unique entity. Compared with pediatric and adult NCCM, fetal NCCM is more prone to biventricle involvement, more likely to be complicated with congenital heart defects, and has a distinct genetic spectrum.

Keywords: fetal cardiovascular abnormality, genetics, left ventricular noncompaction, noncompaction cardiomyopathy, prenatal diagnosis, whole-exome sequencing

INTRODUCTION

Noncompaction cardiomyopathy (NCCM), which is also known as left ventricular noncompaction, is an increasingly recognized type of cardiomyopathy characterized by prominent left ventricular trabeculations, deep intertrabecular recesses communicating with the ventricular cavity, and a thin and compacted epicardial layer. While NCCM is considered genetic cardiomyopathy by the American Heart Association (1), the European Society of Cardiology categorizes it as unclassified cardiomyopathy (2).

Genetics plays an essential role in NCCM because a genetic cause can be identified in 17–40% of cases in differently sized studies of pediatric or adult patients, with sarcomere genes being the most common mutant genes (3–12). However, these studies applied various criteria for the interpretation of sequence variants, as well as various sequencing strategies, mostly Sanger or gene panel sequencing containing a limited and varied number of genes. Therefore, the exact genetic spectrum of NCCM remains uncertain. Furthermore, these studies were in cohorts that were primarily adults and children, rarely involving fetuses. Consequently, Little is known about the genetic spectrum of NCCM in fetuses, and whether the insights obtained from pediatric and adult patients can be extended to fetal NCCM is uncertain. In addition, since NCCM has seldom been described prenatally, the incidence and clinical characteristics of NCCM in the fetus are also mostly unknown.

The present retrospective study aimed to identify the genetic spectrum of fetal NCCM using whole-exome sequencing (WES) and copy number variation sequencing (CNV-seq), to assess the incidence of NCCM in the fetal population, and to characterize the largest cohort of fetal NCCM clinically. The objective was also to identify new potential candidate genes for NCCM. These findings provide novel insights into the clinical and genetic characteristics of NCCM in the fetal population, a segment of the human population not well studied. These insights may also help to improve the prenatal and postnatal management of NCCM.

METHODS

Ethics Statement

This retrospective study was approved by the institutional review board of the Medical Ethics Committee of Beijing AnZhen Hospital. All parents agreed to participate in this study and provided signed informed consent.

Data Collection

We retrospectively reviewed our experience with fetuses diagnosed with NCCM at Beijing Anzhen Hospital, Capital Medical University, from October 2010 to December 2019. Clinical data were collected regarding gestational age at diagnosis, gender, family history, indications for fetal echocardiography, left or biventricular involvement, associated cardiac and extracardiac abnormalities, and pregnancy outcome. Available postnatal echocardiograms and autopsy studies were reviewed to confirm the diagnosis of noncompaction.

Echocardiographic Data

Due to the lack of fetal criteria, we used postnatal echo findings of NCCM in this prenatal population. According to the Jenni criteria, the diagnosis of NCCM was based on a consensus of re-evaluated echocardiography by two dedicated participating cardiologists (13). The Jenni criteria were also applied in right ventricular noncompaction. Considering the retrospective nature of this study, we first screened fetuses with a diagnosis of NCCM in the report and then reviewed the echocardiographic images to ensure that the patients met the inclusion criteria. Fetuses without complete, high-quality images were excluded from further analysis.

Copy Number Variation Sequencing and Data Analyzing

Both CNV-seq and WES were done in the setting of a purely research-based protocol. Next generation sequencing–based CNV analysis is increasingly used in clinical testing (14), and CNV-seq, a viable alternative to arrays for detection of CNV (15), was performed using methods as described previously (16, 17). Briefly, RNA-free genomic DNA was isolated from the umbilical cord using the DNeasy Blood and Tissue Kit (Qiagen GmbH, Hilden, Germany) following the manufacturer's protocol. The quality and concentration of DNA from the samples were assessed using the Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA). A total amount of 1 μg DNA per sample was used as input material for the DNA library preparations. The sequencing library was generated using Truseq Nano DNA HT Sample Prep Kit (Illumina USA) following the manufacturer's recommendations and index codes were added to each sample. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using Hiseq PE Cluster Kit (Illumina) according to the manufacturer's instructions. After cluster generation, the DNA libraries were sequenced on Illumina Hiseq 4000 or Illumina

Novaseq (Illumina, Inc., San Diego, CA, USA) and 150 bp paired-end reads were generated. Then, raw image files were processed by the Bcl To Fastq (Illumina) for base calling and generating the raw data. Reads with adaptors and low quality reads (quality score of <20) were removed. Finally, about 5 million sequencing reads per sample were mapped to the NCBI human reference genome (hg19/GRCh37) using the Burrows-Wheeler Aligner and the allocated to 20-kb sequencing bins with 5-kb sliding to achieve a higher resolution for CNV detection. CNV-seq profiles of each chromosome were represented as log2 of the mean CNV of each sequencing bin along the length of the chromosome.

Detected CNVs were evaluated based on literature reviews and the following public databases: DECIPHER (https://decipher.sanger.ac.uk/), DGV (http://dgv.tcag.ca/), the 1000 Genomes Project (http://www.internationalgenome.org/), OMIM (http://omim.org/), ClinVar (http://www.ncbi.nlm.nih.gov/clinvar), ClinGen (https://www.clinicalgenome.org/) and ISCA CNV (https://www.iscaconsortium.org). According to the American College of Medical Genetics (ACMG) standards and guidelines for interpretation of CNVs, CNVs were classified into five categories: pathogenic, likely pathogenic, uncertain significance, likely benign or benign (14).

Whole-Exome Sequencing, Variant Annotation, Filtering, and Classification

Genomic DNA was extracted from the umbilical cord and parental blood using a Qiagen DNA Blood Midi/Mini kit (Qiagen GmbH, Hilden, Germany). DNA libraries were prepared using an Agilent liquid capture system (Agilent SureSelect Human All Exon V6; Agilent Technologies, Santa Clara, CA, USA) according to the manufacturer's protocol. The size distribution and concentration of the libraries were determined by Agilent 2100 Bioanalyzer (Agilent Technologies) and quantified using real-time PCR. The DNA library was sequenced on Illumina Hiseq 4000 or Illumina Novaseq (Illumina, Inc., San Diego, CA, USA) for paired-end 150 bp reads according to the manufacturer's protocol. Raw image files were processed using Bcl To Fastq (Illumina Inc.) for base calling and generating raw data. Low-quality sequencing reads were filtered out using a quality score of ≥20. The reads were aligned to the NCBI human reference genome (hg19/GRCh37) using the Burrows-Wheeler Aligner. BAM files were subjected to single nucleotide polymorphism (SNP) analysis, duplication marking, indel realignment and recalibration using GATK and Picard.

After variant detection, ANNOVAR was used for annotation (http://wannovar.wglab.org/). Variant frequencies were determined in dbSNP150 (https://www.ncbi.nlm.nih.gov/SNP/), the 1000 Genomes Project, Exome Variant Server (http://evs.gs.washington.edu/EVS) and gnomAD (http://gnomad-old.broadinstitute.org/) to remove common SNP (minor allele frequency > 0.05%). Then, nonsynonymous, splicing, frameshift and non-frameshift variants, as well as variants located in splice sites within 10 base pairs of an exon were prioritized for study. SIFT (http://sift.jcvi.org), PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2), MutationTaster (http://www.mutationtaster.org) and CADD (http://cadd.gs.washington.edu) were used for predicting the pathogenicity of missense variants, while Human Splicing Finder (http://www.umd.be/HSF.html) and MaxEntScan (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html) were used to evaluate the effects on splicing. Moreover, databases such as OMIM, ClinVar and Human Gene Mutation Database(http://www.hgmd.org) were used to determine variant harmfulness and pathogenicity where appropriate. Pathogenicity of variants was determined according to current ACMG guidelines that recommend classifying variants into five categories: pathogenic, likely pathogenic, uncertain significance, likely benign or benign (18).

About the TTN gene, according to a recent publication on TTN mutations in cardiomyopathies, we only evaluated the pathogenicity of loss-of-function (LoF) variants (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single or multiexon deletion) and excluded other variants (19). Variants classified as pathogenic or likely pathogenic were considered positive genetic results. Sanger sequencing was used to validate the presence of positive genetic results.

The novel candidate genes are referred to as genes that have not previously been implicated in NCCM or for which the published data to support the NCCM association may not yet be definitive. For variants in novel candidate genes, we relied on model organism data, tolerance of the gene to sequence variation, co-segregation analysis and *in-silico* prediction where possible.

Statistical Analysis

Categorical variables are presented as frequencies (percentage) and were compared using the Pearson $\chi 2$ test or Fisher's exact test. Statistical analysis was performed using SPSS version 23 (SPSS, Chicago, IL). Two-sided *P*-values < 0.01 were considered significant.

RESULTS

From October 2010 to December 2019, 40 patients with a diagnosis of NCCM in the report were reviewed, and 37 patients fit our inclusion criteria, including five previously published patients with NONO mutations (20). Three cases were excluded for incomplete imaging or lack of evidence for NCCM. The total number of fetuses undergoing ultrasound examinations in our center was 49,898, indicating an overall incidence of 0.07%.

This cohort was composed of 26 males, ten females and one of unknown gender. The median maternal age was 29 (range 20–40) years, and the fetuses were assessed at a median gestational age of 25 (range 20–33) weeks. The indications for referral for fetal echocardiography were suspected ventricular noncompaction detected by ultrasound screening ($n = 15$), congenital heart defects (CHD; $n = 12$), multisystem malformations ($n = 1$), fetal arrhythmia ($n = 1$), cardiac enlargement ($n = 1$), thickened ventricular wall ($n = 1$) and family history of NCCM ($n = 5$) or CHD ($n = 1$) (**Supplementary Table 1**).

Pregnancy Outcome

Thirty-three fetuses (89%) were electively terminated after detailed counseling, with confirmation of NCCM at autopsy in 14 cases. The remaining four (11%) cases were born. Of these

TABLE 1 | General information of the entire cohort.

Gender		
	Male/Female ratio	2.6:1
	Male	26 (72%)
	Female	10 (28%)
	Unknown	1
Maternal Age (years)		
	20–40, median: 29	
Gestational Age (weeks)		
	20–33, median: 25	
Pregnancy Outcome		
	Terminated	33 (89%)
	Born	4 (11%)
Noncompaction Involvement		
	Biventricle	16 (43%)
	Left ventricle	14 (38%)
	Right ventricle	7 (19%)
Concomitant Cardiac Defects		19 (51%)
	Pulmonic stenosis	11
	Ventricular septal defect	10
	Persistent left superior vena cava	5
	Ebstein's anomaly	3
	Pulmonary atresia	3
	Hypoplasia of right ventricle	3
	Tricuspid stenosis	2
	Right aortic arch with mirror image branching	2
	Aortic valve stenosis	2
	Others*	
Rhythm Disturbances		3 (8%)
Hydrops Fetalis		12 (32%)
Genetic Testing		20 (54%)

**Represents congenital heart defect that occurs only once, including: vascular ring, tetralogy of Fallot, atrial septal defect, aberrant right subclavian artery, dysplastic tricuspid valve, ectopic ductus arteriosus, right ventricular diverticulum, right ductus arteriosus, abnormal branching pattern of the aortic arch, interrupted aortic arch type A, coarctation of aorta and anomalous origin of left pulmonary artery from ascending aorta.*

four born patients, one was lost to follow-up after birth; one with moderate heart failure was lost to follow-up after 7 months old; one survived without clinical symptoms at 2 years old; one survived but had decreased left ventricular function and a dilated left ventricle, and was admitted to hospital for medical treatment when she was 7 months old.

Clinical Features

Of the 37 fetal NCCM, 16 (43%) were biventricular involvement, and 14 (38%) were confined to the left ventricle. In the remaining seven cases (19%), NCCM was confined to the right ventricle.

Hydrops fetalis was present in 12 cases (32%), of which four were atypical (pericardial effusion only). NCCM was associated with atrioventricular block II and sinus bradycardia in one case, sinus bradycardia in one case, and atrial bigeminy in one case.

Of the 37 fetuses with NCCM, 19 had additional CHD, the most common being the right-sided lesions ($n = 14$; 74%), followed by ventricular septal defects. This included pulmonary stenosis ($n = 11$), ventricular septal defects ($n = 10$), Ebstein's anomaly ($n = 3$), pulmonary atresia ($n = 3$) and right ventricular hypoplasia ($n = 3$) (**Table 1** and **Supplementary Table 1**).

Genetic Results

Sequencing analysis was performed at autopsy ($n = 19$) or postnatally ($n = 1$) on 20 fetuses. Of the 20 cases undergoing CNV-seq and WES, nine cases (47%) had a positive genetic result, including one with a chromosomal abnormality and eight with pathogenic/likely pathogenic variants. The genotype and phenotype information of the 9 cases are shown in **Table 2**.

Likely-pathogenic and pathogenic variants were identified in genes reported previously in association with NCCM, including NONO ($n = 5$), KCNH2 ($n = 1$), TPM1 ($n = 1$), and PRKAG2 ($n = 1$). Interestingly, no mutations were identified in TTN, MYH7, and MYBPC3; mutations in these three genes together account for more than 50% of the genetic causes in pediatric and adult patients (3, 4, 21).

In the present cohort, fetal NCCM with CHD were more likely to have a positive result than those without CHD (67 vs. 25%; Fisher exact probability test, $p = 0.17$). However, because of the small sample size, this difference must be confirmed in larger cohorts. When considering the known inheritance patterns of eight cases with pathogenic/likely pathogenic variants, five (62.5%) cases with NONO mutations are X-linked recessive and three (27.5%) cases with mutations in KCNH2, TPM1, and PRKAG2 are autosomal dominant inheritance.

Besides the findings presented above, we also detected variants in novel candidate genes (LHX9, CACNA1A, UBQLN4, CACNA1G, and MYH11) for NCCM in five cases (**Supplementary Table 2**). All these variants predicted to be deleterious in potential candidate genes met the following conditions: (1) the allele frequency in the gnomAD database $<$ 0.00001, (2) were LoF variants, and (3) occurred in genes that have a probability of LoF intolerance score $>$ 0.9, supporting deleterious effect for predicted LoF pathogenic variants (22). Two of the five mutations were *de novo*, and the other three were inherited from one parent. Further research is required to evaluate any of the suggested candidate genes. Finally, we also identified four variants of uncertain significance in cardiomyopathy-associated genes (TNNT2, PRDM16, LDB3, and SDHA) in three fetuses (**Supplementary Table 3**). All these variants were paternal inherited, rare and putatively damaging variants.

DISCUSSION

We present the largest fetal series to date, describing the incidence, clinical characteristics, and genetic spectrum of NCCM. The incidence of NCCM in the fetal population was estimated to be 0.07% from our single-center study. We showed that about 43% of fetuses diagnosed with NCCM were biventricular involvement, which was in contrast to the predominant left ventricular involvement seen in adults

TABLE 2 | Genotype and phenotype of patients with positive genetic results.

ID	NCCM Pattern	CHD	Gene	Variant	Zygosity	Parental origin	Pathogenicity
PATHOGENIC/LIKELY PATHOGENIC VARIANTS							
2	BV	EA; PS; VSD	NONO	NM_001145408.1: c.246_249del, p.P83fs7*	Hemizygous	Maternal	Pathogenic
3	LV	PA; VSD; PLSVC; RAAWMIB; RDA	NONO	NM_001145408.1: c.246_249del, p.P83fs7*	Hemizygous	Maternal	Pathogenic
4	LV	PS; VSD; RVD; PLSVC	NONO	NM_001145408.1: c.246_249del, p.P83fs7*	Hemizygous	Maternal	Pathogenic
6	LV	EA; ASD; PS; ABPOTAA	NONO	NM_001145408.1: c.471del, p.Q157fs18*	Hemizygous	Maternal	Pathogenic
9	BV	NA	PRKAG2	NM_016203.3: c.G1592A, p.R531Q	Heterozygous	De novo	Pathogenic
10	BV	PS	KCNH2	NM_000238.4: c.A1847G, p.Y616C	Heterozygous	De novo	Pathogenic
17	BV	EA; PS; VSD	NONO	NM_001145408.1: c.471del, p.Q157fs18*	Hemizygous	Maternal	Pathogenic
36	LV	NA	TPM1	NM_001018005.2: c.G398A, p.R133Q	Heterozygous	De novo	Likely pathogenic
ID	**NCCM**	**CHD**		**Copy Number Variant**	**Size (Mbp)**	**Parental Origin**	**Pathogenicity**
PATHOGENIC/LIKELY PATHOGENIC COPY NUMBER VARIANTS							
13	LV	VSD		seq[hg19]del(1)(q42.3q44) chr1:g.236262500_249214999del*	12.952	De novo	Pathogenic

*ABPOTAA, abnormal branching pattern of the aortic arch; BV, biventricle; CHD, congenital heart defect; EA, Ebstein's anomaly; LV, left ventricle; NCCM, noncompaction cardiomyopathy; PA, Pulmonary atresia; PLSVC, Persistent left superior vena cava; PS, Pulmonic stenosis; RAAWMIB, Right aortic arch with mirror image branching; RDA, Right ductus arteriosus; RVD, Right ventricular diverticulum; VSD, ventricular septal defect. *, This CNV contains 116 genes, among which fourteen genes FH, MTR, HNRNPU, AKT3, RYR2, EDARADD, ZBTB18, GREM2, COX20, CHRM3, NLRP3, FMN2, ACTN2, and SDCCAG8 are associated with diseases in OMIM database. RYR2 gene is known to be associated with cardiomyopathy (arrhythmogenic right ventricular dysplasia-2).*

and children. In this cohort, 51% of cases had significant congenital heart disease, with right-sided lesions being the most common (78%). Of the 20 fetuses undergoing genetic testing, 50% had a positive genetic result. Fetal diagnosis may be associated with more obvious and potentially more severe diseases, which may have lead resulted in different gene mutations being found. For example, in our study, mutations in sarcomere genes are rare. These findings provide novel insights into the clinical and genetic characteristics of NCCM in the fetal population, a segment of the human population not well studied.

Our study's significant strength is that we performed CNV-seq and WES rather than targeted panel sequencing, which only screened genes associated with cardiomyopathies, such as sarcomere genes. This sequencing strategy led to a sweeping landscape of genetic variation of NCCM because of the ability to assess chromosomal abnormalities and identify novel candidate genes, rather than just sequencing the known NCCM genes. This advantage is well illustrated in cases with LoF variants in the NONO gene (**Table 2**) (20). At the time of initial analysis, NONO was not considered as a cardiomyopathy related gene in any published literature. Recently, LoF variants in NONO have been reported in several males with NCCM and CHD (23–25), and the results from the mouse model suggested that NONO deficiency was associated with developing heart defects (26). Furthermore, the cardiac phenotype of the five fetuses with NONO mutations is highly consistent (NCCM, pulmonic stenosis/atresia and septal defects). Taken together, we considered NONO a new gene associated with NCCM and reclassified the variants presented in these fetuses as pathogenic during reassessment at a multidisciplinary team meeting.

NCCM is considered rare, and there is no data on its incidence in the fetal population in previous literature. The estimated incidence of fetal NCCM from our single-center study was 0.07%, similar to that in the adult population. Interestingly, there was an increasing trend in the annual incidence of fetal NCCM in our center during the study (data not shown), probably because of increased awareness, improved imaging technology, and increased popularity of prenatal ultrasound screening in mainland China. This phenomenon indicates that the actual prevalence of NCCM may be higher than previously recognized.

There appears to be a distinct spectrum of gene variants in fetuses with NCCM. Variants in non-sarcomere genes constituted the majority genetic etiology of NCCM, whereas the prevalence of sarcomere gene variants was unexpectedly low (**Table 2** and **Figure 1**). Furthermore, an X-linked recessive inheritance (NONO mutation) occurs more commonly in this fetal cohort. This genetic spectrum is quite different from previous studies in pediatric and adult patients with NCCM. In adult patients with NCCM, mutations in sarcomere genes, especially TTN, MYH7, and MYBPC3, are most commonly detected and account for the majority of genetic etiology of NCCM (3, 4, 11, 12, 21), among which TTN is the most common (8, 21, 27, 28). In contrast, sarcomere gene mutations are less common in children than in adults, especially in TTN, defects of which are rarely observed in children (4–9), whereas an X-linked or mitochondrial inherited defect or chromosomal anomalies are more common in children than in adults (3). Taken together, these data suggest significant differences in the genetic spectrum of NCCM among fetal, pediatric, and adult patients. These differences should be considered in the clinic and when performing molecular diagnostics. Most importantly,

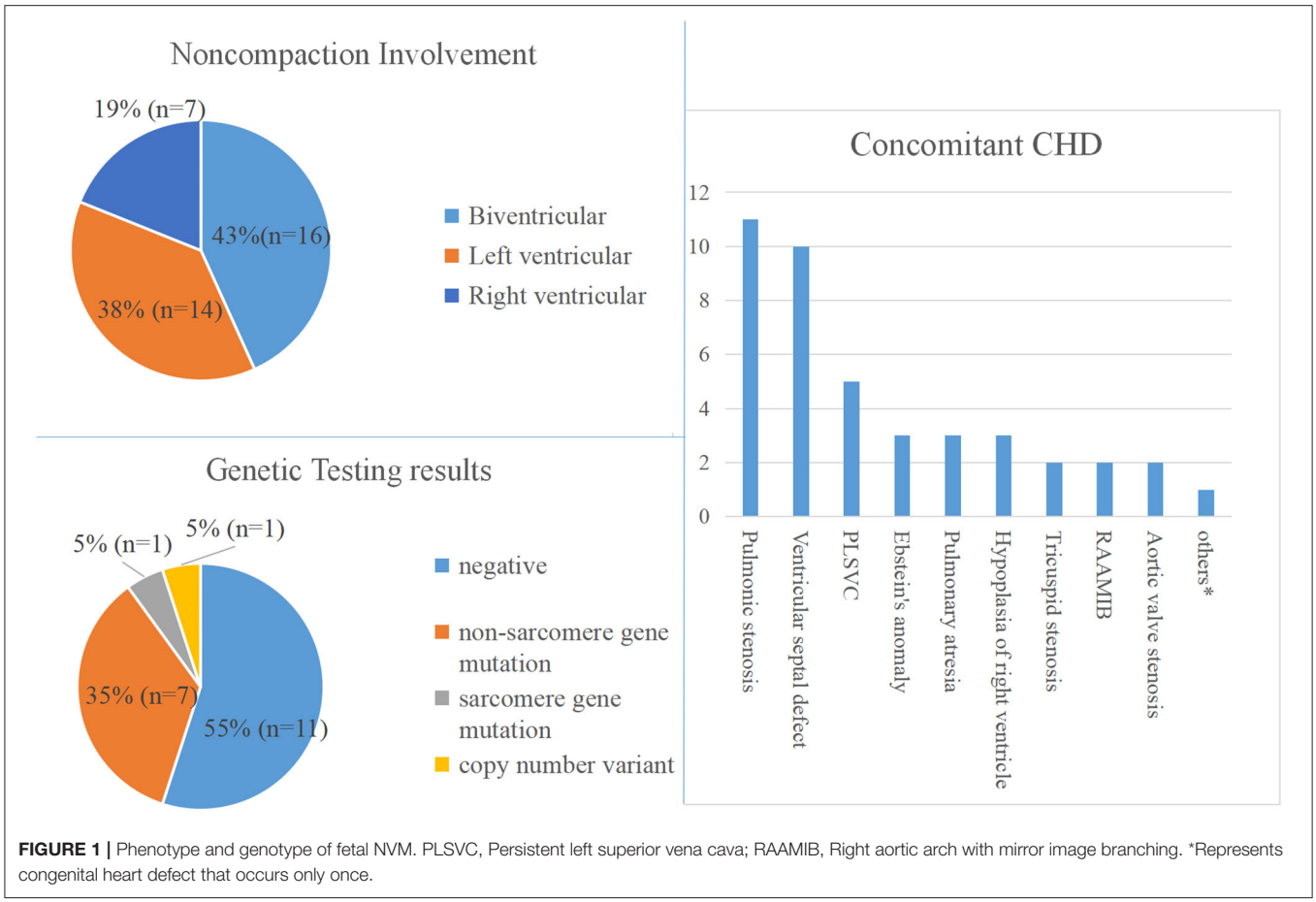


FIGURE 1 | Phenotype and genotype of fetal NVM. PLSVC, Persistent left superior vena cava; RAAMIB, Right aortic arch with mirror image branching. *Represents congenital heart defect that occurs only once.

these differences evoke the question of whether the insights obtained from children and adults with NCCM can be extended to fetal patients, especially when determining the gene range for a panel sequencing, simultaneously highlighting that the genetic spectrum of fetal NCCM must be evaluated in larger studies.

Concomitant CHD was more prevalent in our fetal cohort than reported in children and adults. Of the 37 fetuses with NCCM, 51% had significant CHD. Other studies in the literature on fetal NCCM support the high incidence of concomitant CHD (29–32). In contrast, CHD's prevalence in adult NCCM is much lower, with a reported frequency of 5–12% (4, 21, 33). There is some suggestion that children with NCCM have an increased risk for CHD compared with adults, with a frequency of 30–27% (3, 4, 9). Altogether, these data indicate that the prevalence of concomitant CHD is highest among fetuses with NCCM, followed by children, and lowest among adults. Part of the reason for this phenomenon may be that associated additional CHD may confer additional risk of poor prognosis for NCCM, and NCCM fetuses with significant CHD may suffer fetal demise or have died before adulthood (32, 34, 35).

The distribution of CHD subtypes varied among studies. In a recent study of fetal NCCM involving 22 cases with significant CHD, double outlet right ventricle was the most common (64%) (32). Of the 19 cases with CHD in our cohort, right-sided lesions were the most common (74%), particularly pulmonic stenosis (58%) (**Table 1**). Of the 24 adult NCCM with CHD reported by Stähli et al., left ventricular outflow tract abnormalities were the most frequent (46%) (33). These wide variations in the distribution of CHD subtypes may be partly due to the small number of affected patients and the variable characteristics of the cohorts (e.g., isolated LVNC or not, diagnostic criteria, age at diagnosis).

In contrast to the predominant left ventricular involvement of NCCM seen in adults, one observation emerging from this study is the high incidence of biventricular involvement (43%) in fetal patients. The high incidence of biventricular involvement in fetal NCVM has been described in previous fetal NCCM cases and small series and may be related to the right ventricular dominance of fetal circulation (30, 36–39). The difference between the fetus and the adult in the affected ventricles supports the view that the predominance of circulation may play an important role in noncompaction, and NCCM may reflect an impaired adaptation to special hemodynamics rather than an arrest during myocardial embryogenesis (40).

STUDY LIMITATIONS

This study has some limitations. The sample size was relatively small. Eighty-nine percentage of fetuses had the pregnancy

terminated, so the evolution of the phenotypic findings observed in these fetuses is unknown. Besides, subtle dysmorphic features cannot be determined using fetal ultrasound, and some phenotypes, neurodevelopmental disorders in particular, are impossible to determine in the prenatal setting. Since normally the right ventricle is much more trabeculated than the left, the application of left ventricular noncompaction diagnostic criteria to the right ventricle is not rigorous. However, no recognized diagnostic criteria are available for right ventricular noncompaction. To solve the issue, we chose a compromise method. While applying the left ventricular criteria to the right ventricular, we also listed the right ventricular noncompaction/compaction ratio of fetuses diagnosed with right ventricular or biventricular noncompaction in **Supplementary Table 1**.

CONCLUSION

By clinical phenotyping and next generation sequencing of the largest cohort of fetuses with NCCM, we find that fetal NCCM is a unique entity compared to pediatric and adult NCCM: 1. Concomitant congenital heart defects are more prevalent; 2. Noncompaction involves biventricle is much more common; 3. Mutations in sarcomere genes are rarer. These differences evoke the question of whether the insights obtained from children and adults with NCCM can be extended to fetal patients, especially when determining the gene range for a panel sequencing.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://bigd.big.ac.cn/gsa-human/browse/HRA000495.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the institutional review board of the Medical Ethics Committee of Beijing Anzhen Hospital. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

HZ, HS, and YH designed the study. XH, XW, XZ, YZhan, XL, JH, XG, LS, YZhao, and TY collected the study materials or samples. HS and TY completed the sequencing experiments. HS, XH, and XW collected and aggregated the data. HS, XH, and XW analyzed and interpreted the data. HS and XH wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by National Key R&D Program of China (2018YFC1002300), Beijing Lab for Cardiovascular Precision Medicine, Beijing, China, code: PXM2020_014226_000017_00377132_FCG, and Beijing Municipal Science & Technology Commission (Z181100001918008).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2020.617561/full#supplementary-material

REFERENCES

1. Maron BJ, Towbin JA, Thiene G, Antzelevitch C, Corrado D, Arnett D, et al. Contemporary definitions and classification of the cardiomyopathies. *Circulation.* (2006) 113:1807–16. doi: 10.1161/CIRCULATIONAHA.106.174287
2. Elliott P, Andersson B, Arbustini E, Bilinska Z, Cecchi F, Charron P, et al. Classification of the cardiomyopathies: a position statement from the european society of cardiology working group on myocardial and pericardial diseases. *Eur Heart J.* (2007) 29:270–6. doi: 10.1093/eurheartj/ehm342
3. van Waning JI, Moesker J, Heijsman D, Boersma E, Majoor Krakauer D. Systematic review of genotype-phenotype correlations in noncompaction cardiomyopathy. *J Am Heart Assoc.* (2019) 8:e12993. doi: 10.1161/JAHA.119.012993
4. van Waning JI, Caliskan K, Hoedemaekers YM, van Spaendonck-Zwarts KY, Baas AF, Boekholdt SM, et al. Genetics, clinical features, and long-term outcome of noncompaction cardiomyopathy. *J Am Coll Cardiol.* (2018) 71:711–22. doi: 10.1016/j.jacc.2017.12.019
5. Li S, Zhang C, Liu N, Bai H, Hou C, Wang J, et al. Genotype-positive status is associated with poor prognoses in patients with left ventricular noncompaction cardiomyopathy. *J Am Heart Assoc.* (2018) 7:e9910. doi: 10.1161/JAHA.118.009910
6. Finsterer J, Stöllberger C, Towbin JA. Left ventricular noncompaction cardiomyopathy: cardiac, neuromuscular, and genetic factors. *Nat Rev Cardiol.* (2017) 14:224–37. doi: 10.1038/nrcardio.2016.207
7. Wang C, Hata Y, Hirono K, Takasaki A, Ozawa SW, Nakaoka H, et al. A wide and specific spectrum of genetic variants and genotype-phenotype correlations revealed by next-generation sequencing in patients with left ventricular noncompaction. *J Am Heart Assoc.* (2017) 6:e006210. doi: 10.1161/JAHA.117.006210
8. Sedaghat-Hamedani F, Haas J, Zhu F, Geier C, Kayvanpour E, Liss M, et al. Clinical genetics and outcome of left ventricular non-compaction cardiomyopathy. *Eur Heart J.* (2017) 38:3449–60. doi: 10.1093/eurheartj/ehx502.P1338
9. Miller EM, Hinton RB, Czosek R, Lorts A, Parrott A, Shikany AR, et al. Genetic testing in pediatric left ventricular noncompaction. *Circ Cardiovasc Genet.* (2017) 10:e001735. doi: 10.1161/CIRCGENETICS.117.001735
10. Arbustini E, Favalli V, Narula N, Serio A, Grasso M. Left ventricular noncompaction. *J Am Coll Cardiol.* (2016) 68:949–66. doi: 10.1016/j.jacc.2016.05.096
11. van Waning JI, Caliskan K, Michels M, Schinkel AFL, Hirsch A, Dalinghaus M, et al. Cardiac phenotypes, genetics, and risks in familial noncompaction cardiomyopathy. *J Am Coll Cardiol.* (2019) 73:1601–11. doi: 10.1016/j.jacc.2018.12.085

12. Klaassen S, Probst S, Oechslin E, Gerull B, Krings G, Schuler P, et al. Mutations in sarcomere protein genes in left ventricular noncompaction. *Circulation.* (2008) 117:2893–901. doi: 10.1161/CIRCULATIONAHA.107.746164
13. Jenni R. Echocardiographic and pathoanatomical characteristics of isolated left ventricular non-compaction: a step towards classification as a distinct cardiomyopathy. *Heart.* (2001) 86:666–71. doi: 10.1136/heart.86.6.666
14. Riggs ER, Andersen EF, Cherry AM, Kantarci S, Kearney H, Patel A, et al. Technical standards for the interpretation and reporting ofconstitutional copy-number variants: a joint consensus recommendation of the AmericanCollege of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet Med.* (2020) 22:245–57. doi: 10.1038/s41436-019-0686-8
15. Liang D, Peng Y, Lv W, Deng L, Zhang Y, Li H, et al. Copy number variation sequencing for comprehensive diagnosis of chromosome disease syndromes. *J Mol Diagn.* (2014) 16:519–26. doi: 10.1016/j.jmoldx.2014.05.002
16. Liu S, Song L, Cram DS, Xiong L, Wang K, Wu R, et al. Traditional karyotyping vs copy number variation sequencing for detection of chromosomal abnormalities associated with spontaneous miscarriage. *Ultrasound Obst Gyn.* (2015) 46:472–7. doi: 10.1002/uog.14849
17. Sun H, Yi T, Hao X, Yan H, Wang J, Li Q, et al. Contribution of single-gene defects to congenital cardiac left-sided lesions in the prenatal setting. *Ultrasound Obst Gyn.* (2020) 56:225–32. doi: 10.1002/uog.21883
18. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* (2015) 17:405–23. doi: 10.1038/gim.2015.30
19. Ware JS, Cook SA. Role of titin in cardiomyopathy: from DNA variants to patient stratification. *NAT Rev Cardiol.* (2018) 15:241–52. doi: 10.1038/nrcardio.2017.190
20. Sun H, Zhou X, Hao X, Zhang Y, Zhang H, He Y. Characteristics of cardiac phenotype in prenatal familial cases with NONO mutations. *Circ Genom Precis Med.* (2020) 13:e2847. doi: 10.1161/CIRCGEN.119.002847
21. Kayvanpour E, Sedaghat-Hamedani F, Gi W, Tugrul OF, Amr A, Haas J, et al. Clinical and genetic insights into non-compaction: a meta-analysis and systematic review on 7598 individuals. *Clin Res Cardiol.* (2019) 108:1297–308. doi: 10.1007/s00392-019-01465-3
22. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* (2016) 536:285–91. doi: 10.1038/nature19057
23. Carlston CM, Bleyl SB, Andrews A, Meyers L, Brown S, Bayrak-Toydemir P, et al. Expanding the genetic and clinical spectrum of the NONO-associated X-linked intellectual disability syndrome. *Am J Med Genet A.* (2019) 179:792–6. doi: 10.1002/ajmg.a.61091
24. Scott DA, Hernandez-Garcia A, Azamian MS, Jordan VK, Kim BJ, Starkovich M, et al. Congenital heart defects and left ventricular non-compaction in males with loss-of-function variants in NONO. *J Med Genet.* (2016) 54:47–53. doi: 10.1136/jmedgenet-2016-104039
25. Reinstein E, Tzur S, Cohen R, Bormans C, Behar DM. Intellectual disability and non-compaction cardiomyopathy with a *de novo* NONO mutation identified by exome sequencing. *Eur J Hum Genet.* (2016) 24:1635–8. doi: 10.1038/ejhg.2016.72
26. Xu X, Jiang H, Lu Y, Zhang M, Cheng C, Xue F, et al. Deficiency of NONO is associated with impaired cardiac function and fibrosis in mice. *J Mol Cell Cardiol.* (2019) 137:46–58. doi: 10.1016/j.yjmcc.2019.10.004
27. Richard P, Ader F, Roux M, Donal E, Eicher JC, Aoutil N, et al. Targeted panel sequencing in adult patients with left ventricular non-compaction reveals a large genetic heterogeneity. *Clin Genet.* (2018) 95:356–67. doi: 10.1111/cge.13484
28. Miszalski-Jamka K, Jefferies JL, Mazur W, Głowacki J, Hu J, Lazar M, et al. Novel genetic triggers and genotype-phenotype correlations in patients with left ventricular noncompaction. *Circ Cardiovasc Genet.* (2017) 10:e001763. doi: 10.1161/CIRCGENETICS.117.001763
29. Stöllberger C, Wegner C, Finsterer J. Fetal ventricular hypertrabeculation/noncompaction: clinical presentation, genetics, associated cardiac and extracardiac abnormalities and outcome. *Pediatr Cardiol.* (2015) 36:1319–26. doi: 10.1007/s00246-015-1200-y
30. Menon SC, O'Leary PW, Wright GB, Rios R, MacLellan-Tobert SG, Cabalka AK. Fetal and neonatal presentation of noncompacted ventricular myocardium: expanding the clinical spectrum. *J Am Soc Echocardiog.* (2007) 20:1344–50. doi: 10.1016/j.echo.2007.04.028
31. Ursell PC. Noncompaction in the fetus and neonate: an autopsy study. *Am J Med Genet C Semin Med Genet.* (2013) 163:169–77. doi: 10.1002/ajmg.c.31367
32. Arunamata A, Punn R, Cuneo B, Bharati S, Silverman NH. Echocardiographic diagnosis and prognosis of fetal left ventricular noncompaction. *J Am Soc Echocardiog.* (2012) 25:112–20. doi: 10.1016/j.echo.2011.09.019
33. Stähli BE, Gebhard C, Biaggi P, Klaassen S, Valsangiacomo Buechel E, Attenhofer Jost CH, et al. Left ventricular non-compaction: prevalence in congenital heart disease. *Int J Cardiol.* (2013) 167:2477–81. doi: 10.1016/j.ijcard.2012.05.095
34. Aras D, Tufekcioglu O, Ergun K, Ozeke O, Yildiz A, Topaloglu S, et al. Clinical features of isolated ventricular noncompaction in adults long-term clinical course, echocardiographic properties, and predictors of left ventricular failure. *J Card Fail.* (2006) 12:726–33. doi: 10.1016/j.cardfail.2006.08.002
35. Ramachandran P, Woo JG, Ryan TD, Bryant R, Heydarian HC, Jefferies JL, et al. The impact of concomitant left ventricular non-compaction with congenital heart disease on perioperative outcomes. *Pediatr Cardiol.* (2016) 37:1307–12. doi: 10.1007/s00246-016-1435-2
36. Sleurs E, De Catte L, Benatar A. Prenatal diagnosis of isolated ventricular noncompaction of the myocardium. *J Ultras Med.* (2005) 24:1325–9. doi: 10.7863/jum.2005.24.9.1325
37. Moura C, Hillion Y, Daikha-Dahmane F, Eydoux P, Fallet C, Oury JF, et al. Isolated non-compaction of the myocardium diagnosed in the fetus: two sporadic and two familial cases. *Cardiol Young.* (2002) 12:278–83. doi: 10.1017/S1047951102000598
38. Tian L, Zhou Q, Zhou J, Zeng S, Cao D, Zhang M. Ventricular non-compaction cardiomyopathy: prenatal diagnosis and pathology. *Prenatal Diag.* (2015) 35:221–7. doi: 10.1002/pd.4523
39. Stöllberger C, Finsterer J. Trabeculation and left ventricular hypertrabeculation/noncompaction. *J Am Soc Echocardiog.* (2004) 17:1120–1. doi: 10.1016/j.echo.2004.06.009
40. Vidaeff AC, Ramin SM, Glaser AM, Gupta-Malhotra M. Differences in fetal and postnatal presentations of isolated noncompaction of the ventricular myocardium. *J Ultras Med.* (2011) 30:293–5. doi: 10.7863/jum.2011.30.3.293

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 09 March 2021
doi: 10.3389/fcvm.2021.650667

Loss-of-Function Variants in the *SYNPO2L* Gene Are Associated With Atrial Fibrillation

Alexander Guldmann Clausen[1], *Oliver Bundgaard Vad*[1], *Julie Husted Andersen*[1] and *Morten Salling Olesen*[1,2*]

[1] Department of Biomedical Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, [2] Laboratory for Molecular Cardiology, Department of Cardiology, The Heart Centre, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark

Multiple genome-wide association studies (GWAS) have identified numerous loci associated with atrial fibrillation (AF). However, the genes driving these associations and how they contribute to the AF pathogenesis remains poorly understood. To identify genes likely to be driving the observed association, we searched the FinnGen study consisting of 12,859 AF cases and 73,341 controls for rare genetic variants predicted to cause loss-of-function. A specific splice site variant was found in the *SYNPO2L* gene, located in an AF associated locus on chromosome 10. This variant was associated with an increased risk of AF with a relatively high odds ratio of 3.5 ($p = 9.9 \times 10^{-8}$). *SYNPO2L* is an important gene involved in the structural development and function of the cardiac myocyte and our findings thus support the recent suggestions that AF can present as atrial cardiomyopathy.

Keywords: atrial fibrillation, genetics, arrhythmia, splice site variant, cardiomyopathy, cardiology

OPEN ACCESS

Edited by:
Seitaro Nomura,
The University of Tokyo, Japan

Reviewed by:
Tom Robert Webb,
University of Leicester, United Kingdom
Douglas A. Marchuk,
Duke University, United States

***Correspondence:**
Morten Salling Olesen
morten.salling.olesen@gmail.com

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: *07 January 2021*
Accepted: *18 February 2021*
Published: *09 March 2021*

Citation:
Clausen AG, Vad OB, Andersen JH and Olesen MS (2021) Loss-of-Function Variants in the SYNPO2L Gene Are Associated With Atrial Fibrillation. Front. Cardiovasc. Med. 8:650667. doi: 10.3389/fcvm.2021.650667

INTRODUCTION

Atrial Fibrillation (AF) is the most common cardiac arrhythmia with more than 33 million people affected worldwide (1). The disease has been associated with an increased risk of stroke, heart failure as well as increased all-cause mortality (2). While AF is a complex disease associated with multiple risk factors including age, sex, lifestyle and existing heart disease, a significant genetic component of the disease has also been established (3). In 2018, a meta-analysis of genome wide association studies (GWAS) of AF identified a total of 134 loci associated with AF (4). Many of these loci have been linked to genes related to ion channels, structural proteins of the myocardium and transcription factors (5). However, a large number of the identified loci are located in non-coding regions and in the vicinity of multiple genes. More work is needed to identify the causal genes mediating the observed associations. The purpose of this paper is to investigate potentially causal genes at known GWAS loci in order to further elucidate the mechanisms behind AF pathogenesis. This was done by searching genes in a publicly available dataset containing phenotypes linked with imputed genotypes for AF-associated variants in loci previously associated with AF in GWAS.

MATERIALS AND METHODS

To identify genes likely to be driving the association between the observed GWAS loci and AF, we searched the FinnGen database (6) for variants significantly enriched in AF cases. Only variants located in one of the 134 GWAS loci previously identified in our GWAS meta-analysis (4) were

selected for further investigation. To increase the likelihood of the identified variants contributing to the pathogenesis of AF development, we confined our analysis of the FinnGen data to only include predicted loss-of-functions variants. These were defined as variants leading to a pre-mature stop codon, deletions or insertions leading to a frameshift or variants involving splice sites. To further evaluate the identified candidate genes, we analyzed mRNA expression in right atrial appendage tissue integrated with genotype data obtained through the Genotype-Tissue Expression (GTEx) project. This was done by searching the publicly available GTEx release v8 database (https://gtexportal.org/home/gene/SYNPO2L). A brief description of the used data from the FinnGen database and the GTEx project is given below.

FinnGen Database

The publicly available FinnGen database contains genotype data from Finnish biobanks linked with corresponding national health registry data from 12,859 AF cases and 73,341 controls. Data from the FinnGen database was retrieved from freeze number 3. FinnGen participants were genotyped by the FinnGen investigators using various Illumina and Affymetrix GWAS arrays. This was followed by a genotype imputation by the Beagle v.4.1 software using 3,775 high coverage whole genome sequences of Finnish origin as a population specific reference panel (6). We used the FinnGen disease endpoint "Atrial Fibrillation and Flutter." The SAIGE software was used for the genetic association tests of the FinnGen data. SAIGE is a mixed model logistic regression R/C++ package that is able to adjust for relatedness of the analyzed samples while also controlling for unbalanced case-control ratios that can otherwise lead to inflated type 1 error rates (7). SAIGE first generates relatedness estimates and a null model. Next, the null model is fitted with the respective SNP, and the genetic kinship matrix and the model itself is then adjusted for the following covariates: sex, age, 10 principal components and genotyping batch. For more information on the included biobanks, disease endpoint definitions and genotype data in the FinnGen database, please visit https://finngen.gitbook.io/documentation/v/r3/.

GTEx Project

The GTEx Project database contains analyses of mRNA levels in 47 different tissues, including atrial appendage tissue obtained from 429 deceased donors with available genotype data. The data was obtained by massive parallel sequencing of RNA strands using Illumina TrueSeq RNA sequencing followed by qPCR (8).

COLOCALIZATION ANALYSIS

In order to further explore the genes responsible for the identified genome-wide associated risk locus, we performed a colocalization analysis. Using the LocusFocus tool (9), we tested whether the GWAS signals are colocalized with eQTL signals, and therefore are likely to share a causal variant. The LocusFocus tool utilizes the *Simple Sum* colocalization method based on a frequentist framework. To perform the colocalization analysis, we used the publicly available FinnGen AF GWAS summary statistics and integrated them with cis-eQTL data from atrial appendage tissue from the GTEx project v8. For an identified locus, we used GWAS summary statistics from all SNPs within a 100 kb window centered around the identified lead SNP. The performance of the Simple Sum colocalization method compared with other colocalization methods has been described by Gong et al. (10).

PATHWAY ANALYSIS

We performed an *in-sillico* pathway analysis of interactions between the protein product of *SYNPO2L* and other proteins in AF, using the tool String (v11) (11). String is a large database of protein-protein interactions, derived from multiple sources, including both physical and functional associations. We conducted a search of proteins thought to be important in structural function of cardiomyocytes, as defined in a recent review of AF genetics by Andersen et al. (12). Afterwards a search of all proteins was done, limited to the 10 interactions predicted with highest confidence. Evaluation of confidence in interactions have been described in detail by Szklarczyk et al. (11).

EVALUATION OF VARIANT DELETERIOUSNESS

Variants in candidate genes were evaluated using the *in silico* tool Combined Annotation Dependent Depletion Score (CADD), in order to predict the effects and deleteriousness of the concerned variants (13). This method uses a machine learning model to incorporate multiple different forms of variant annotations into one integrated score, in order to better identify causal variants. Variants with a CADD-score >20 are predicted to be among the 1% most deleterious variants in the human genome. Those with a CADD-score >30 are predicted to be among the 0.1% most deleterious.

RESULTS

FinnGen Database

A total of 119 variants were associated with AF in the FinnGen study. Of these, 13 variants were located in genes in loci previously identified in our GWAS meta-analysis. The 13 variants are presented in **Table 1**. Of the 13 variants, only one variant, rs766868752, was predicted to be a loss-of-function variant. This variant was located in the *SYNPO2L* gene. The variant was imputed with an INFO-score of 0.972, indicating a very high confidence imputation.

The rs766868752 variant in *SYNPO2L* is characterized as a splice site donor variant and is located on position 73,655,817 on the first base of intron 1. In carriers of the variant, the C base of the *non-template strand* is replaced with an A base. Thus, the normal splice donor site of 5′ GT 3′ in the template strand is changed to a 5′ TT 3 ′. In the 12,859 AF cases in the FinnGen study, the variant was found with a minor allele frequency (MAF) of 0.002585. In the 73,341 control participants, the variant was found with a MAF of 0.001022. We found an increased Odds Ratio (OR) of developing AF of 3.5 in persons carrying the c.105+1G>T variant in *SYNPO2L* ($p = 9.9 \times 10^{-8}$). According

TABLE 1 | Variants associated with AF in the GWAS meta-analysis by Roselli et al. that were located near genes also associated with AF in the FinnGen study.

Meta-analysis SNP	Chr	Nearest gene	RR (95 % CI)	*p*-value	Effect/ref allele	Effect allele frequency	SNP placement relative to nearest gene	FinnGen variant rsid	FinnGen variant OR (95% CI)	*p*-value	Effect/ref allele	Effect allele frequency	FinnGen variant placement relative to nearest gene
rs880315	1	*CASZ1*	1.04 (1.03–1.06)	$p = 5.0 \times 10^{-9}$	C/T	0.37	Intronic	rs34071855	1.1 (1.07–1.15)	$p = 3.6 \times 10^{-8}$	G/C	0.41	Intronic
rs12044963	1	*KCND3*	1.08 (1.06–1.11)	$p = 1.6 \times 10^{-12}$	T/G	0.18	Intronic	rs34119223	0.88 (0.84–0.92)	$p = 1.3 \times 10^{-7}$	T/TCC	0.19	Intronic
rs35504893	2	*TTN*	1.09 (1.08–1.11)	$p = 6.9 \times 10^{-25}$	T/C	0.25	Intronic	rs16866400	1.1 (1.09–1.19)	$p = 4.4 \times 10^{-9}$	A/G	0.22	Intronic
rs6810325	3	*CAND2*	1.08 (1.06–1.09)	$p = 5.2 \times 10^{-23}$	C/T	0.63	Intronic	rs7650482	1.1 (1.06–1.14)	$p = 1.3 \times 10^{-6}$	G/A	0.61	Intronic
rs7632427	3	*EPHA3*	1.04 (1.03–1.06)	$p = 1.10 \times 10^{-8}$	T/C	0.61	Downstream	rs7633500	0.92 (0.88–0.95)	$p = 4.2 \times 10^{-6}$	A/G	0.45	Intronic
rs73366713	6	*ATXN1*	1.11 (1.09–1.14)	$p = 5.8 \times 10^{-21}$	G/A	0.86	Intronic	rs112583508	0.86 (0.81–0.92)	$p = 4.7 \times 10^{-6}$	A/G	0.098	Intronic
rs3176326	6	*CDKN1A*	1.06 (1.04–1.08)	$p = 8.0 \times 10^{-11}$	G/A	0.80	Intronic	rs3176323	0.9 (0.86–0.94)	$p = 3.0 \times 10^{-7}$	C/T	0.29	Intronic
rs17079881	6	*SLC35F1*	1.09 (1.07–1.11)	$p = 4.2 \times 10^{-16}$	G/A	0.14	Intronic	rs111862156	1.5 (1.28–1.80)	$p = 1.7 \times 10^{-6}$	T/C	0.012	Intronic
rs11773845	7	*CAV1*	1.12 (1.11–1.14)	$p = 4.6 \times 10^{-58}$	A/C	0.59	Intronic	rs3807990	0.89 (0.86–0.93)	$p = 5.5 \times 10^{-8}$	T/C	0.28	Intronic
rs60212594	10	*SYNPO2L*	1.12 (1.09–1.14)	$p = 6.5 \times 10^{-27}$	G/C	0.85	Intronic	rs766868752	3.5 (2.20–5.48)	$p = 9.9 \times 10^{-8}$	A/C	$1.3{\cdot}10^{-3}$	Splice site donor
rs113819537	12	*SSPN*	1.05 (1.03–1.07)	$p = 2.2 \times 10^{-09}$	C/G	0.74	Upstream	rs113819537	0.9 (0.87–0.94)	$p = 4.3 \times 10^{-7}$	G/C	0.31	5′ UTR
rs883079	12	*TBX5*	1.13 (1.11–1.14)	$p = 1.3 \times 10^{-51}$	T/C	0.69	3′ UTR	rs7312625	1.1 (1.07–1.16)	$p = 1.9 \times 10^{-7}$	A/G	0.7	Intronic
rs2359171	16	*ZFHX3*	1.21 (1.19–1.23)	$p = 2.9 \times 10^{-100}$	A/T	0.19	Intronic	rs2359171	1.2 (1.14–1.24)	$p = 3.8 \times 10^{-15}$	A/T	0.23	Intronic

All FinnGen variants had an imputation info score >0.95. RR, Relative Risk; OR, Odds Ratio; CI, Confidence Interval.

to the gnomAD browser, a database containing more than 76,000 exome sequences, the specific variant was almost exclusively found in people of Finnish descent (14). Of the 41 carriers of the mutation in the gnomAD database, 40 was of Finnish descent and only one carrier was of non-Finnish European descent. Only one other splice site donor variant has been identified in the gnomAD v3 database, suggesting that splice site mutations in *SYNPO2L* are rare and likely deleterious.

Two isoforms of the *SYNPO2L* genes are expressed *in-vivo*. A long isoform with the ID ENST00000394810.2 as well as a short isoform with the ID ENST00000372873.8 (14). The splice site mutation found in the FinnGen study is located in the long transcript as seen in **Figure 1**. As the only gene positioned in a locus identified in our GWAS meta-analysis that also was found to have predicted loss-of-function variants enriched in the FinnGen study population, the *SYNPO2L* gene was selected for further investigation.

Colocalization Analysis

To determine whether the GWAS signals at the *SYNPO2L* locus are likely to be driven by changes in *SYNPO2L* expression in atrial appendage tissue, we performed a colocalization analysis using the LocusFocus tool. Our analysis supports a strong colocalization of GWAS signals with eQTLs of *SYNPO2L* in atrial appendage tissue with a *Simple Sum p*-value of 2.1×10^{-6}. The colocalization analysis has been visualized in **Figure 2**.

GTEx Project Results

To further assess the involvement of the *SYNPO2L* gene in AF, we examined whether the expression of the *SYNPO2L* gene was affected in carriers of the SNP rs60212594. This common SNP is located at the *SYNPO2L* locus and was associated with an increased risk of AF in our GWAS meta-analysis (4). This was done by a quantitative trait locus (cis-eQTL) analysis using data from the GTEx project. The rs60212594 SNP identified in the GWAS was shown to be associated with a minor but significant change in *SYNPO2L* expression (15). Carriers of the C risk allele had a 0.08-fold decrease in *SYNPO2L* expression with a *p*-value of 2.9×10^{-8}, which is illustrated in **Figure 3**.

Pathway Analysis

Using the tool String, we examined interactions between the protein product of *SYNPO2L* and those of other structural genes previously associated with AF. The protein product of *SYNPOL2* as well as *TTN* was found to be co-expressed in numerous species as illustrated in **Figures 4A–E**, indicating a functional association between the protein products of these genes. When searching for interactions between *SYNPO2L* and all genes, we found the *SYNPO2L* to interact with and be co-expressed with the protein products of several other genes. These genes are associated with the structure of sarcomeres, calcium signaling and cardiomyopathies. The proteins interacting with the *SYNPO2L* product are summarized in **Supplementary Table 1**.

Variant Deleteriousness Prediction

Using *in silico* analyses we investigated the potential deleteriousness of the rs766868752 variant in the *SYNPO2L* gene. The variant was found to have a CADD score of 35, indicating that it is among the 0.1% most deleterious variants in the human genome.

DISCUSSION

In a previously performed meta-analysis of GWAS, 134 loci associated with AF were identified in the human genome (4). However, these are all common variants with small effect sizes. In this study we report on the first loss-of-function variant in the FinnGen study significantly associated with atrial fibrillation. This rare variant has a substantial effect size and is found in the *SYNPO2L* gene located on chromosome 10q22. This region has previously been associated with AF as the intronic SNP rs60212594 in *SYNPO2L* has been identified in our GWAS meta-analysis. This common variant was found with an associated OR of 1.12 ($p = 6{,}48 \times 10^{-27}$). Interestingly, a study by Brugada et al. (16) from 1997 identified a family of 26 with AF being inherited in an autosomal dominant pattern. The responsible genetic locus was localized to chromosome 10q22-10q24. Combined with GWAS, this strongly implicate genes in chromosome 10q22 as candidate genes with a potential large significant effect on the risk of AF.

Different studies have suggested both *MYOZ1* (17, 18) and *SYNPO2L* (19, 20) both located on chromosome 10q22, as the likely causal genes driving the association. Studies suggesting *MYOZ1* as the causal gene, have done so primarily based on eQTL studies showing an altered expression of the *MYOZ1* gene in carriers of AF risk alleles identified through GWAS studies. However, expression of the *MYOZ1* gene has been shown to be primarily restricted to skeletal muscle, with very low *MYOZ1* mRNA levels in cardiac muscle tissue (21, 22). Additionally, no variants in the *MYOZ1* gene were found to be significantly enriched in the FinnGen study. These findings suggest *MYOZ1* as an unlikely candidate gene in the locus. Conversely, the findings of this study showing a predicted loss-of-function variant in the *SYNPO2L* gene with an associated OR of 3.5, strongly suggests *SYNPO2L* as the likely causal gene in the locus. This is supported by the colocalization of the AF GWAS signals with the eQTLs of *SYNPO2L* as seen in **Figure 2**.

Functions of the SYNPO2L Gene Product

The functions of the *SYNPO2L* gene product were first studied by Beqqali et al. (23), naming the protein product Cytoskeletal Heart-enriched Actin-associated Protein (CHAP). Beqqali et al. identified two different isoforms both expressed in cardiac tissue. The long isoform was named CHAPa (=978 aa) coded by the ENST00000394810.2 transcript, and the short isoform was named CHAPb (=749 aa) coded by the ENST00000372873.4 transcript. The authors furthermore found that the long isoform CHAPa was predominately expressed in adult cardiac tissue, whereas the short isoform, CHAPb, was primarily expressed in embryonic cardiac tissue. Expression of *SYNPO2L* in various tissues can be seen in **Supplementary Figure 1** (24). The splice-site variant identified in the FinnGen database is only present in the long isoform coding the CHAPa. Using *in situ* hybrdization in mice, Beqqali et al. determined the subcellular localization of the proteins to the Z-disc of both skeletal and cardiac muscle tissue by co-localization with a-actinin (23). Both the CHAPa

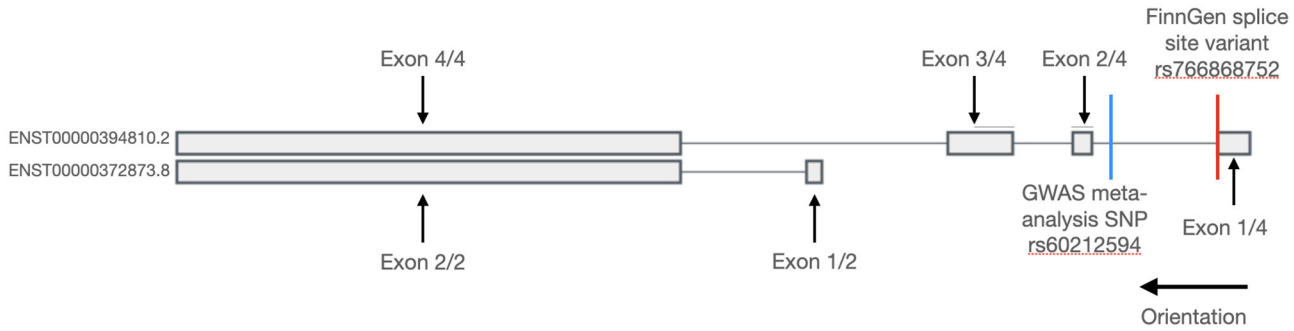


FIGURE 1 | Transcripts of the *SYNPO2L* gene. Exons are shown with thick bars and introns are shown with thin lines. The *SYNPO2L* gene is expressed in a short and a long isoform. The splice site variant found in the FinnGen database is shown with a red line and is located in the first intron of the long isoform. The position of the common SNP, rs60212594, identified in the GWAS meta-analysis is shown with a blue line.

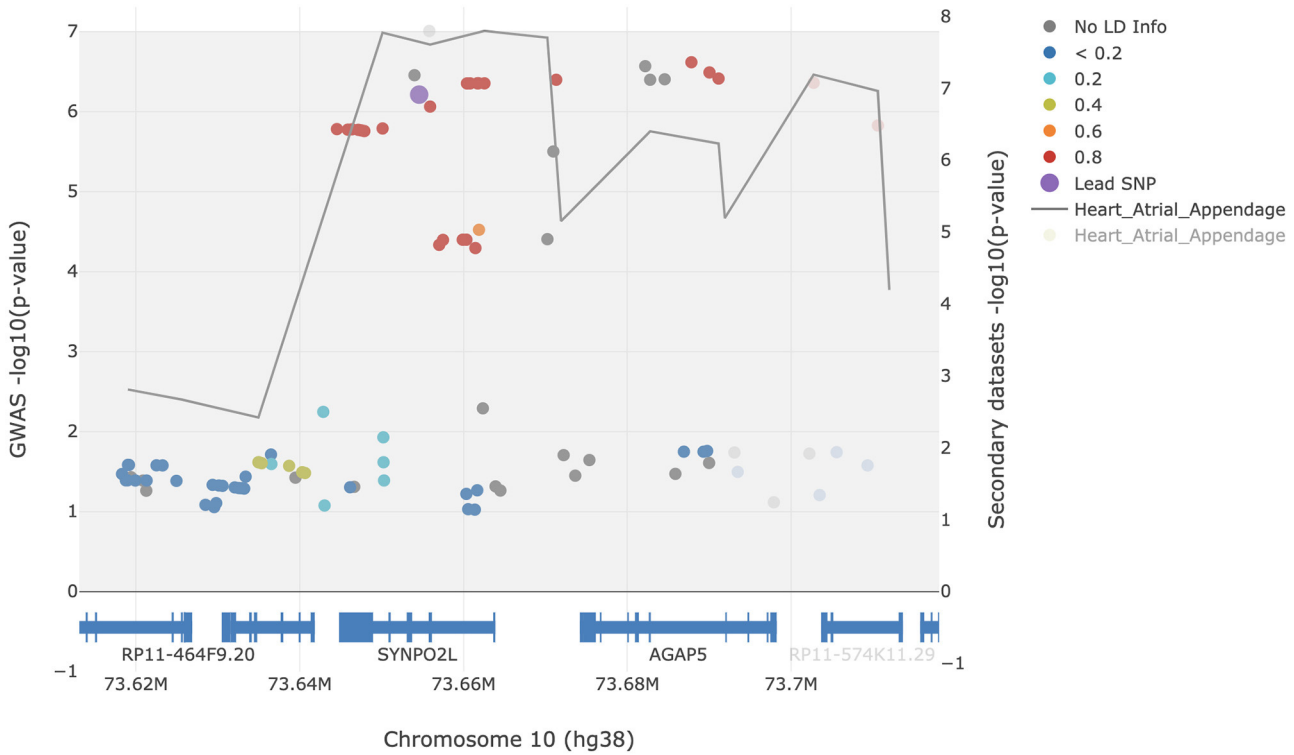


FIGURE 2 | Colocalization analysis of AF GWAS signals with eQTL signals of *SYNPO2L* in atrial appendage tissue. Filled circles represent AF GWAS -$\log_{10}$(p-values) (left y-axis). The gray line represents the eQTL signals and traces the lowest p-value [right y-axis, showing -$\log_{10}$(*p*-values)] per window of 22.5 bp. The rs60212594 SNP was defined as the lead SNP and is presented in purple. The European 1000 Genomes Project was used for the LD calculations in reference to the lead SNP. The figure was generated using the LocusFocus tool.

and CHAPb isoforms contain a nuclear localization signal (NLS) which was found to be highly conserved in both the mouse and zebrafish orthologs. Accordingly, CHAPb, the embryonic form, was also localized to the nucleus in embryonic mice cardiac tissue.

To further investigate the function of the *SYNPO2L* product, Beqqali et al. knocked down the CHAPa and CHAPb isoforms of the *SYNPO2L* gene product in a zebrafish model using antisense morpholino oligonucleotides (23). The zebrafish kncokdown-models showed a phenotype with highly disorganized sarcomeres and Z-disks, as well as decreased cardiac contractility. α-actinin-2, which was shown to interact with the CHAP proteins, was also displaced resulting in disorganized myofibrils. This suggests that the *SYNPO2L* gene product is highly involved in the structural development and function of the sarcomeres. Due to its nuclear and Z-disc localization, this role could be carried

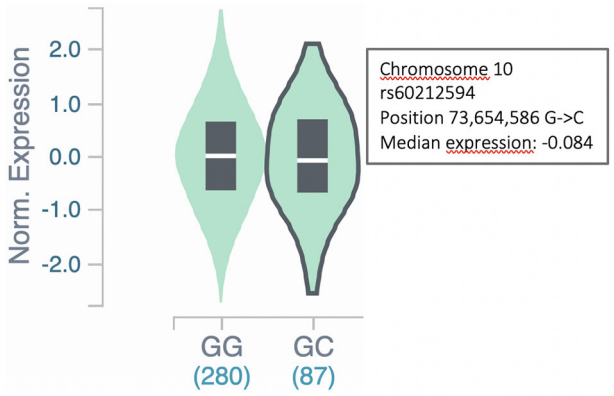


FIGURE 3 | *SYNPO2L* expression in carriers of the rs60212594 SNP. Expression of the *SYNPO2L* gene is decreased ($p = 2.9 \times 10^{-8}$) in heterozygous carriers of the C allele of the rs60212594 SNP associated with AF. The data and figure were obtained from the GTEx project eQTL analysis (https://gtexportal.org/home/snp/rs60212594).

out by the protein acting as both a structural component and also being involved in cellular signaling processes. This is in line with previous findings suggesting other Z-disc proteins to be associated with a-actinin. One example is MLP, which can translocate from the cytoplasm to the nucleus and interact with varius transcription factors (25).

The consequences of splice site mutations on the protein product depend on multiple different factors. These are complex to predict, but can include exon skipping, intron inclusion or activation of different cryptic splice sites (26). If the splice site mutation in intron 1 of *SYNPO2L* leads to a change in the remaining reading frame, the mutation is likely to lead to a pre-mature stop codon followed by non-sense mediated decay (NMD) (27). Therefore, the splice site mutation could dispose to AF either by a lack of protein product or by an altered dysfunctional protein product. The pathogenicity of the splice variant identified in this study is underlined by *in sillico* analysis, predicting the variant to be among the 0.1% variants in the genome most likely to be deleterious. Future analysis of mRNA extracted from cell lines or patients with the splice site variant could help determine how the mutation affects splicing and protein product. Carriers of the common AF risk variant rs60212594 show a decreased expression of *SYNPO2L,* as shown in the eQTL analysis (**Figure 3**). Expression of *SYNPO2L* in carriers of this common variant was examined as the variant is associated with an increased risk of AF in our GWAS meta-analysis (4). The decreased expression is consistent with the present finding that the rs766868752 splice site variant, possibly leading to mRNA NMD, was strongly associated with AF. Both findings suggest that a decreased amount of *SYNPO2L* gene product pre-disposes to AF.

Genetic variants in other structural cardiac proteins e.g., titin, have previously been shown to be strongly associated with AF (28). *TTN* encodes a giant sarcome protein (titin) expressed in all chambers of the human heart. Ahlberg et al. (28) found sarcomere defects in mutant zebrafish carrying a truncated variant in ttn, the homologs *TTN* gene in zebrafish. This was also the case in the *SYNPO2L* knockdown zebrafish model studied by Beqqali et al. (23). Furthermore, a prolonged PR interval as well as an increased amount of fibrosis was shown in zebra fish models carrying the truncated titin variant (28). Another interesting gene, *MYL4*, is responsible for contractile, electrical and structural integrity of the atrium. Peng et al. found a loss-of function variant to cause early atrial fibrosis leading to atrial arrhytmia and atrial cardiomyopathy (29). In line with this, a mouse model with an overexpression of the *SYNPO2L* gene studied by Van Eldik et al. (30), also displayed both increased interstitial fibrosis and electrical abnormalities. Cardiac fibrosis has been shown to be increased in AF patients (31). The findings of increased cardiac fibrosis and atrial cardiomyopathy, resulting from abnormal sarcomere organization, propose a possible mechanism of how variants in the *SYNPO2L* gene and other structural proteins of the cardiomyocyte pre-dispose to AF.

However, this study also has several limitations. While the association seems biologically plausible as the genetic locus is associated with AF and the variant has been predicted to be deleterious, the splice variant, rs766868752, is found with a very low allele frequency and is almost exclusively found in persons of Finnish heritage. Furthermore, the allele frequency of the variant has been imputed. Although the variant was imputed with a very high confidence in imputation, one must consider the possible risk of inaccuracies in imputation. Finally, the common variant analyzed in this study, rs60212594, while significantly associated with AF, has a quite modest effect on disease risk, and may not be of major clinical significance for individuals carrying this variant. We therefore urge our readers to interpret our results with caution.

CONCLUSION AND SIGNIFICANCE

For many years, AF has been considered an electrical disease, but our understanding of the pathophysiology of AF has improved and the field continues to grow as we explore new associations in AF genetics. Our findings suggest that mutations in the *SYNPO2L* gene encoding the structural protein CHAP pre-dispose to AF. This is primarily based on the loss-of-function splice site variant in the *SYNPO2L* gene that was associated with a significantly increased risk of AF in a Finnish population ($p = 9.9 \times 10^{-8}$). The loss-of-function variant in the *SYNPO2L* gene is too rare to be of any major clinical significance alone. However, the findings that mutations in another gene involved in the structural arrangement of the sarcomeres pre-dispose to AF, supports the theory that AF is not only an electrophysiological disease. Future studies are needed to investigate if *SYNPO2L* loss-of-function variants in general pre-dispose for AF with a similar effect size. The discovery that atrial fibrosis plays a significant role in AF might explain why existing treatment options of AF are sometimes inadequate and may help identify new directions of AF treatments and risk stratifications in the future. It also underlines the complexity of the disease and

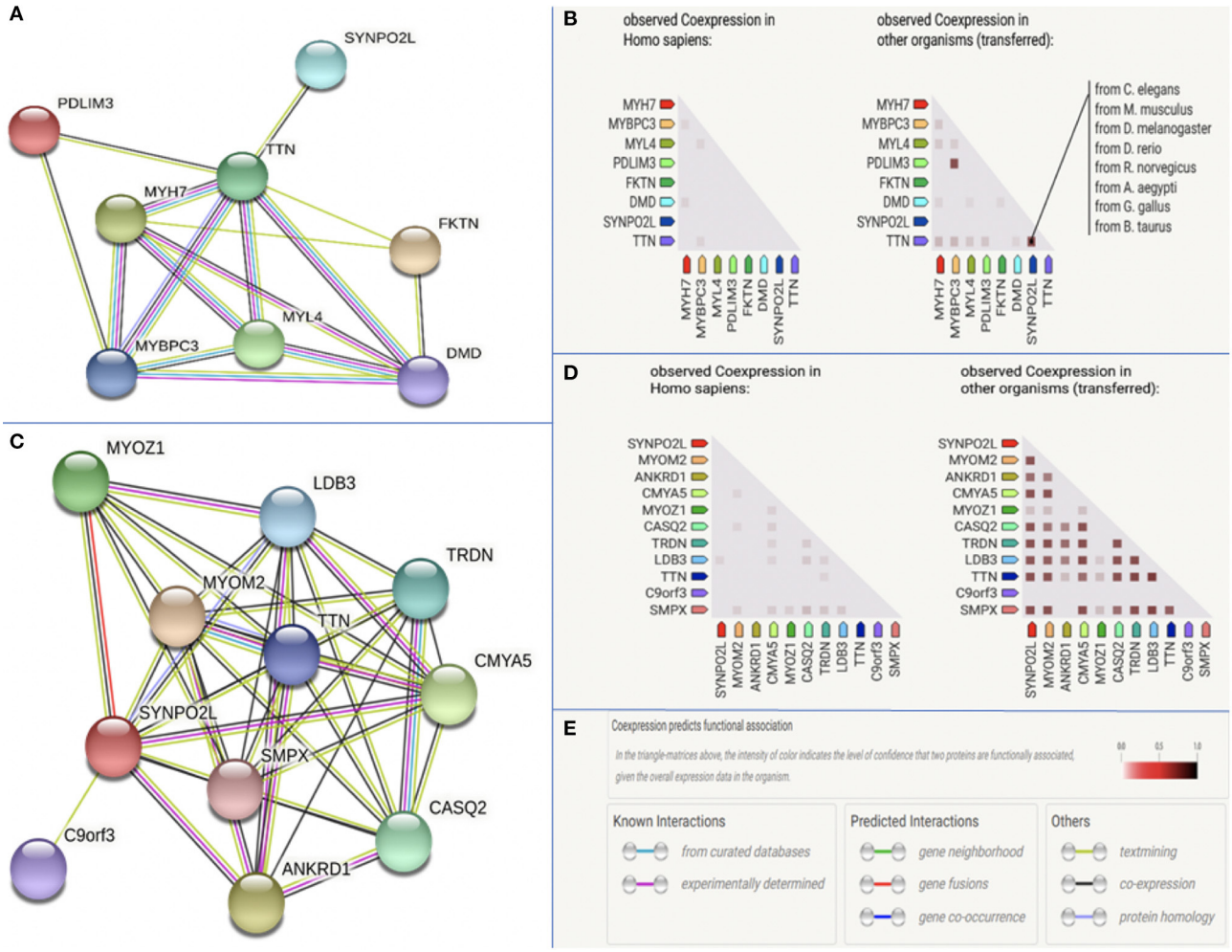


FIGURE 4 | Protein-protein interaction analysis. The figure is modified from data provided by the tool String v11. **(A)** Protein interactions between the product of *SYNPO2L* and seven other structural genes previously associated with AF. **(B)** Co-expression data across Homo Sapiens and other species for the product of *SYNPO2L* and seven other structural genes previously associated with AF. Species in which co-expression has been observed has been highlighted for *SYNPO2L* and *TTN*. **(C)** Protein interactions between the product of *SYNPO2L* and the products of the 10 genes with the highest confidence in interaction prediction. **(D)** Co-expression data across Homo Sapiens and other species for the product of *SYNPO2L* and the products of the 10 genes with the highest confidence in interaction prediction. **(E)** Legend of color codes used in this figures **(A–D)**.

suggests that different genes in different geographic regions and populations may pre-dispose for the disease through diverse pathophysiological mechanisms.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.finngen.fi/en/access_results.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

AC and MO: conceived the idea. AC, MO, and OV: performed the analyses. AC: prepared the original draft. AC, MO, OV, and JA: helped write, edit and review the manuscript for key intellectual content. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was supported by The John and Birthe Meyer Foundation, The Research Foundation at Rigshospitalet, Villadsen Family Foundation, The Arvid Nilsson Foundation, and The Hallas-Møller Emerging Investigator Novo Nordisk (NNF17OC0031204). The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS.

ACKNOWLEDGMENTS

We want to acknowledge the participants and investigators of FinnGen study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2021.650667/full#supplementary-material

REFERENCES

1. Chugh SS, Havmoeller R, Narayanan K, Singh D, Rienstra M, Benjamin EJ, et al. Worldwide epidemiology of atrial fibrillation. *Circulation*. (2014) 129:837–47. doi: 10.1161/CIRCULATIONAHA.113.005119
2. Ruddox V, Sandven I, Munkhaugen J, Skattebu J, Edvardsen T, Otterstad JE. Atrial fibrillation and the risk for myocardial infarction, all-cause mortality and heart failure: a systematic review and meta-analysis. *Eur J Prev Cardiol.* (2017) 24:1555–66. doi: 10.1177/2047487317715769
3. Fox CS, Parise H, D'Agostino Ralph BS, Lloyd-Jones DM, Vasan RS, Wang TJ, et al. Parental atrial fibrillation as a risk factor for atrial fibrillation in offspring. *JAMA*. (2004) 291:2851–5. doi: 10.1001/jama.291.23.2851
4. Roselli C, Chaffin MD, Weng L-C, Aeschbacher S, Ahlberg G, Albert CM, et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet.* (2018) 50:1225–33. doi: 10.1038/s41588-018-0133-9
5. Gutierrez A, Chung MK. Genomics of atrial fibrillation. *Curr Cardiol Rep.* (2016) 18:6. doi: 10.1007/s11886-016-0735-8
6. FinnGen. *FinnGen Documentation of R3 Release*. (2020). Available online at: https://finngen.gitbook.io/documentation/ (accessed November 22, 2020).
7. Zhou W, Nielsen JB, Fritsche LG, Dey R, Gabrielsen ME, Wolford BN, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet.* (2018) 50:1335–41. doi: 10.1038/s41588-018-0184-y
8. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The genotype-tissue expression (GTEx) project. *Nat Genet.* (2013) 45:580–5. doi: 10.1038/ng.2653
9. Panjwani N, Wang F, Mastromatteo S, Bao A, Wang C, He G, et al. LocusFocus: web-based colocalization for the annotation and functional follow-up of GWAS. *PLoS Comput Biol.* (2020) 16:e1008336. doi: 10.1371/journal.pcbi.1008336
10. Gong J, Wang F, Xiao B, Panjwani N, Lin F, Keenan K, et al. Genetic association and transcriptome integration identify contributing genes and tissues at cystic fibrosis modifier loci. *PLoS Genet.* (2019) 15:e1008007. doi: 10.1371/journal.pgen.1008007
11. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* (2019) 47:D607–13. doi: 10.1093/nar/gky1131
12. Andersen JH, Andreasen L, Olesen MS. Atrial fibrillation—a complex polygenetic disease. *Eur J Hum Genet.* (2020). doi: 10.1038/s41431-020-00784-8. [Epub ahead of print].
13. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* (2019) 47:D886–94. doi: 10.1093/nar/gky1016
14. GnomAD. *GnomAD browser v3.1*. Available online at: https://gnomad.broadinstitute.org/gene/ENSG00000166317?dataset=gnomad_r2_1 (accessed November 7, 2020).
15. *GTEx Project, Single Tissue eQTL Browser*. Available online at: https://gtexportal.org/home/snp/rs60212594 (accessed November 27, 2020).
16. Brugada R, Tapscott T. Identification of a genetic locus for familial atrial fibrillation. *N Engl J Med.* (1997) 336:905–11.
17. Lin H, Dolmatova EV, Morley MP, Lunetta KL, McManus DD, Magnani JW, et al. Gene expression and genetic variation in human atria. *Hear Rhythm.* (2014) 11:266–71. doi: 10.1016/j.hrthm.2013.10.051
18. Martin RIR, Babaei MS, Choy MK, Owens WA, Chico TJA, Keenan D, et al. Genetic variants associated with risk of atrial fibrillation regulate expression of PITX2, CAV1, MYOZ1, C9orf3 and FANCC. *J Mol Cell Cardiol.* (2015) 85:207–14. doi: 10.1016/j.yjmcc.2015.06.005
19. Nielsen JB, Fritsche LG, Zhou W, Teslovich TM, Holmen OL, Gustafsson S, et al. Genome-wide study of atrial fibrillation identifies seven risk loci and highlights biological pathways and regulatory elements involved in cardiac development. *Am J Hum Genet.* (2018) 102:103–15. doi: 10.1016/j.ajhg.2017.12.003
20. Lubitz SA, Brody JA, Bihlmeyer NA, Roselli C, Weng LC, Christophersen IE, et al. Whole exome sequencing in atrial fibrillation. *PLoS Genet.* (2016) 12:1–12. doi: 10.1371/journal.pgen.1006284
21. Gontier Y, Taivainen A, Fontao L, Sonnenberg A, van der Flier A, Carpen O, et al. The Z-disc proteins myotilin and FATZ-1 interact with each other and are connected to the sarcolemma via muscle-specific filamins. *J Cell Sci.* (2005) 118:3739–49. doi: 10.1242/jcs.02484
22. Posch MG, Perrot A, Dietz R, Ozcelik C, Pankuweit S, Ruppert V, et al. Mutations in MYOZ1 as well as MYOZ2 encoding the calsarcins are not associated with idiopathic and familial dilated cardiomyopathy. *Mol Genet Metab.* (2007) 91:207–8. doi: 10.1016/j.ymgme.2007.02.014
23. Beqqali A, Monshouwer-Kloots J, Monteiro R, Welling M, Bakkers J, Ehler E, et al. CHAP is a newly identified Z-disc protein essential for heart and skeletal muscle function. *J Cell Sci.* (2010) 123:1141–50. doi: 10.1242/jcs.063859
24. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-based map of the human proteome. *Science.* (2015) 347:1260419. doi: 10.1126/science.1260419
25. Ecarnot-Laubriet A, De Luca K, Vandroux D, Moisant M, Bernard C, Assem M, et al. Downregulation and nuclear relocation of MLP during the progression of right ventricular hypertrophy induced by chronic pressure overload. *J Mol Cell Cardiol.* (2000) 32:2385–95. doi: 10.1006/jmcc.2000.1269
26. Anna A, Monika G. Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J Appl Genet.* (2018) 59:253–68. doi: 10.1007/s13353-018-0444-7
27. Baker KE, Parker R. Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Curr Opin Cell Biol.* (2004) 16:293–9. doi: 10.1016/j.ceb.2004.03.003
28. Ahlberg G, Refsgaard L, Lundegaard PR, Andreasen L, Ranthe MF, Linscheid N, et al. Rare truncating variants in the sarcomeric protein titin associate with familial and early-onset atrial fibrillation. *Nat Commun.* (2018) 9:4316. doi: 10.1038/s41467-018-06618-y
29. Peng W, Li M, Li H, Tang K, Zhuang J, Zhang J, et al. Dysfunction of myosin light-chain 4 (MYL4) leads to heritable atrial cardiomyopathy with electrical, contractile, and structural components: evidence from genetically-engineered rats. *J Am Heart Assoc.* (2017) 6:e007030. doi: 10.1161/JAHA.117.007030
30. van Eldik W, den Adel B, Monshouwer-Kloots J, Salvatori D, Maas S, van der Made I, et al. Z-disc protein CHAPb induces cardiomyopathy and contractile dysfunction in the postnatal heart. *PLoS ONE.* (2017) 12:e0189139. doi: 10.1371/journal.pone.0189139

31. Dzeshka MS, Lip GYH, Snezhitskiy V, Shantsila E. Cardiac fibrosis in patients with atrial fibrillation. *J Am Coll Cardiol.* (2015) 66:943–59. doi: 10.1016/j.jacc.2015.06.1313

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cardiovascular Medicine

REVIEW
published: 16 March 2021
doi: 10.3389/fcvm.2021.653244

Genetic and Cellular Interaction During Cardiovascular Development Implicated in Congenital Heart Diseases

Kazuki Kodo†, Keiko Uchida† and Hiroyuki Yamagishi†*

Division of Pediatric Cardiology, Department of Pediatrics, Keio University School of Medicine, Tokyo, Japan

Congenital heart disease (CHD) is the most common life-threatening congenital anomaly. CHD occurs due to defects in cardiovascular development, and the majority of CHDs are caused by a multifactorial inheritance mechanism, which refers to the interaction between genetic and environmental factors. During embryogenesis, the cardiovascular system is derived from at least four distinct cell lineages: the first heart field, second heart field, cardiac neural crest, and proepicardial organ. Understanding the genes involved in each lineage is essential to uncover the genomic architecture of CHD. Therefore, we provide an overview of recent research progress using animal models and mutation analyses to better understand the molecular mechanisms and pathways linking cardiovascular development and CHD. For example, we highlight our recent work on genes encoding three isoforms of inositol 1,4,5-trisphosphate receptors (IP_3R1, 2, and 3) that regulate various vital and developmental processes, which have genetic redundancy during cardiovascular development. Specifically, IP_3R1 and 2 have redundant roles in the atrioventricular cushion derived from the first heart field lineage, whereas IP_3R1 and 3 exhibit redundancy in the right ventricle and the outflow tract derived from the second heart field lineage, respectively. Moreover, 22q11.2 deletion syndrome (22q11DS) is highly associated with CHD involving the outflow tract, characterized by defects of the cardiac neural crest lineage. However, our studies have shown that *TBX1*, a major genetic determinant of 22q11DS, was not expressed in the cardiac neural crest but rather in the second heart field, suggesting the importance of the cellular interaction between the cardiac neural crest and the second heart field. Comprehensive genetic analysis using the Japanese genome bank of CHD and mouse models revealed that a molecular regulatory network involving GATA6, FOXC1/2, TBX1, SEMA3C, and FGF8 was essential for reciprocal signaling between the cardiac neural crest and the second heart field during cardiovascular development. Elucidation of the genomic architecture of CHD using induced pluripotent stem cells and next-generation sequencing technology, in addition to genetically modified animal models and human mutation analyses, would facilitate the development of regenerative medicine and/or preventive medicine for CHD in the near future.

Keywords: heart field, neural crest, outflow tract, inositol trisphosphate receptor, TBX1, 22q11.2 deletion

OPEN ACCESS

Edited by:
Seitaro Nomura,
The University of Tokyo, Japan

Reviewed by:
Robert Kelly,
UMR7288 Institut de Biologie du Développement de Marseille (IBDM), France
Patrick G. Burgon,
Qatar University, Qatar

***Correspondence:**
Hiroyuki Yamagishi
hyamag@keio.jp

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: *14 January 2021*
Accepted: *22 February 2021*
Published: *16 March 2021*

Citation:
Kodo K, Uchida K and Yamagishi H (2021) Genetic and Cellular Interaction During Cardiovascular Development Implicated in Congenital Heart Diseases. Front. Cardiovasc. Med. 8:653244. doi: 10.3389/fcvm.2021.653244

INTRODUCTION

Congenital heart disease (CHD) is the most common life-threatening congenital anomaly that occurs in ∼1% of live births. With advances in pediatric cardiology and cardiac surgery, most patients with CHD survive to adulthood; therefore, understanding the inheritance of CHD has become an increasingly critical clinical issue. Although insight gleaned from molecular genetics combined with developmental biology approaches has helped to uncover the detailed mechanisms of cardiovascular development, the genomic architecture of CHD remains largely unknown. We provide an overview of the progress that research has made till date in understanding the molecular mechanisms contributing to cardiovascular development, which in turn can provide new directions for research to uncover the inheritance of CHD and key susceptibility genes. We first provide general background into the etiology of CHD and the nature of cardiac development, highlighting our work on the role of inositol 1,4,5-trisphosphate receptors (IP_3Rs) in this process. In addition, we focus on recent research demonstrating a mechanistic link of the T-box-containing transcription factor (TBX1) with CHD in the context of 22q11.2 deletion syndrome (22q11DS). Finally, we highlight progress to date in understanding the general genetic architecture associated with CHD and the underlying regulatory mechanisms.

ETIOLOGY OF CHD

CHD is considered to occur due to defects in cardiovascular development during the first 6 weeks of gestation. At this stage, the heart and vessels develop from a simple primitive tube structure into a four-chambered heart with two great vessels. Genetic factors, including chromosomal abnormalities, are estimated to account for approximately 8% of CHD cases, with single-gene mutations accounting for about 5% of cases, and environmental factors, including maternal infections, systemic diseases, and administration of drugs, accounting for about 2% of CHD cases. However, the etiology of the remaining ∼85% of CHDs is generally unknown, and is therefore attributed to so-called "multifactorial inheritance," which refers to the interaction between certain genetic and environmental factors (1, 2). Recently, more genetic factors associated with CHD have been reported, including chromosomal abnormalities for 12% of cases, *de novo* copy number variants such as chromosomal microdeletion accounting for 15% of cases and *de novo* gene mutation affecting protein function in 10% of cases, and inherited gene mutations in 1.3% of cases (**Tables 1**, **2**) (3–5). As shown in **Tables 1**, **2**, candidate monogenic factors include many transcription factors and signal molecules that are essential for development of the heart and are responsible for multiple types of CHD. Genetic alterations of these factors are considered to disrupt the spatiotemporal regulation of complex three-dimensional heart structure. However, the interaction of multiple genetic and environmental factors is still considered as the primary etiology of the remaining majority of CHDs.

DEVELOPMENTAL ORIGINS OF THE CARDIOVASCULAR SYSTEM

Current knowledge in molecular embryology suggests that the cardiovascular system is derived from at least four distinct cell lineages, namely, the first heart field (FHF), second heart field (SHF), cardiac neural crest (CNC), and proepicardial organ (PEO) (**Figure 1**) (6–9). The FHF stands for the crescent shaped heart primordium that is derived from the anterior lateral plate mesoderm. The FHF cells (shown in red in **Figure 1**) form a primitive straight heart tube, consisting of an interior endocardial layer and an exterior myocardial layer along with cardiac jelly (extracellular matrix) layer in between. In addition to the FHF, the SHF (shown in blue in **Figure 1**) develops medially to the cardiac crescent from the splanchnic mesoderm and lies along the pharyngeal region dorsal to the primitive heart tube derived from the FHF (10–12). Eventually, the heart tube provides a scaffold and cardiac progenitor cells derived from the SHF migrate into both anterior and posterior ends of it. The heart tube proceeds looping rightward, the cells originally from the FHF finally form exclusively the left ventricle and part of the atria, whereas cells from the SHF migrated into the anterior portions of the heart tube form a large portion of the outflow tract of and the right ventricle. In addition, cells from the SHF cross the pharyngeal mesoderm into the posterior end of the heart tube contribute to a part of the atria. Meanwhile, CNC cells (shown in yellow in **Figure 1**), specifically developed in the dorsal region of the neural tube between the mid-otic placode and the third somite, migrate to the outflow tract where they give rise to the outflow tract septum to separate the truncus arteriosus into the aorta and pulmonary artery (13–15). CNC cells also migrate to pharyngeal arch arteries 3, 4, and 6, where they differentiate into smooth muscle cells of the great vessels. The neural crest cells from the preotic region of the neural tube contribute to the development of coronary arteries (16). The PEO (shown in green in **Figure 1**) is derived from the coelomic mesothelium that overlays the liver bud and gives rise to the epicardial layer over the heart (12). Some epicardial cells invade the subepicardial space through a process of epithelial–mesenchymal transformation, and contribute to the development of the coronary vessels and connective tissues (17, 18). To further uncover the genetic architecture of CHD, it is essential to adopt an approach for identifying the specific genes involved in each of these progenitor cell lineages, and to determine how their interaction regulates cardiovascular development.

IP_3Rs IN CARDIOVASCULAR DEVELOPMENT

We have investigated the roles of three isoforms of IP_3R (IP_3R1, 2, and 3) in cardiovascular development, demonstrating their genetic redundancy (**Figure 2**). In particular, IP_3R1 and 2 have redundant roles in the FHF-derived lineage, whereas IP_3R1 and 3 exhibit redundancy in SHF-derived lineages. IP_3Rs are intracellular Ca^{2+}-release channels, which are opened by

TABLE 1 | Genetic causes of non-syndromic congenital heart diseases.

	Gene	Cardiovascular malformation	Gene MIM
Transcription factors	*CITED2*	ASD, VSD, AS, PS, SIT, Dextrocardia, TGA, TOF, RVOTO, TAPVR	602937
	GATA4	Dextrocardia, AVSD, DORV,TOF, BAV, CoA, AR, PAPVR,PDA, PS, ASD, VSD	600576
	GATA5	AVSD, DORV, LVNC, BAV, CoA	611496
	GATA6	AVSD, TOF, PDA, PTA, PS, ASD, VSD	601656
	MED13L	TGA	608771
	NR2F2	AVSD, AS, CoA, VSD, HLHS, TOF, DORV	107773
	NKX2–5	ASD, AVSD, BAV, CoA, Dextrocardia, DORV, Ebstein's anomaly, HTX, HLHS, IAA, LVNC, Mitral valve anomalies, PA, PAPVR, PDA, PS, SVAS, TA, TAPVR, TGA, TOF, PTA, VSD	600584
	NKX2–6	PTA	611770
	TBX1	DORV, TOF, IAA, PTA, VSD,	602054
	TBX2	ASD, VSD, RVOTO	600747
	TBX5	AVSD, TOF, BAV, CoA, ASD, VSD	601620
	TBX20	ASD, VSD, MS, DCM, LVNC	606061
	MEF2C	DORV	600602
	ZFPM2/FOG2	AVSD, DORV, TOF, VSD	603693
	FOXH1	TOF, TGA, HTX, VSD	603621
	FOXO1	TOF	136533
	FOXP1	AVSD, HLHS	605515
	HAND1	AVSD, DORV, HLHS, HLV, HRV, ASD, VSD	602406
	HAND2	TOF, LVNC, VSD	602407
	MSX1	BAV, CoA	142983
	NFATC1	TOF, LVNC, BAV, CoA, TA, VSD	600489
	ETS1	DORV, HLHS, ASD, VSD	164720
	JARID2	Left-sided lesions	601594
	NR1D2	AVSD	602304
	RBPJ	HLHS	147183
	RFX3	PTA	601337
Cell signaling and adhesion proteins	*ACVR1/ALK2*	HTX, AVSD, DORV, TGA, Left-sided lesions, ASD	102576
	ACVR2B	HTX, Dextrocardia, AVSD, DORV, TGA, HLHS, PS, Venous anomaly	602730
	BMPR1A	AVSD	601299
	BMPR2	AVSD, PDA, PAPVR, ASD, VSD	600799
	GDF1	HTX, AVSD, DORV, TGA, TOF	602880
	SMAD6	HLHS, AS, BAV, CoA	602931
	CRELD1	ASD, AVSD	607170
	GJA1	HLHS, VSD, PA	121014
	JAG1	Aortic dextroposition, TOF, BAV, CoA, PS, VSD	601920
	NOTCH1	HTX, AVSD, TOF, HLHS, LVNC, BAV, CoA, AS, MS, VSD	190198
	NOTCH2	AVSD, TOF, BAV, CoA, PS,	600275
	PDGFRA	TAPVR	173490
	TAB2	BAV, AS, TOF	605101
	ADAM17	AVSD	603639
	HES1	TGA	139605
	HEY2	AVSD	604674
	APC	BAV, CoA	611731
	DCHS1	LVNC, MVP	603057
	DVL1	LVNC, PDA	601365
	EDN1	TOF	131240
	PCDHA9	HLHS	606315
	VEGFA	TOF, PDA, PTA, AS, BAV, CoA, IAA, VSD	192240
Structural proteins	*ACTC1*	ASD, HCM, DCM, LVNC	102540
	DCHS1	MVP	603057

(Continued)

TABLE 1 | Continued

	Gene	Cardiovascular malformation	Gene MIM
	ELN	SVAS	130160
	MYH6	ASD, HCM, DCM	160710
	MYH7	Ebstein's anomaly, LVNC, HCM, DCM	160760
	MYH11	PDA, TAA	160745

AS, aortic stenosis; ASD, atrial septal defect; AVSD, atrioventricular septal defect; BAV, bicuspid aortic valve; CoA, coarctation of the aorta; DCM, dilated cardiomyopathy; DORV, double outlet right ventricle; HCM, hypertrophic cardiomyopathy; HLHS, hypoplastic left heart syndrome; HLV, hypoplastic left ventricle; HRV, hypoplastic right ventricle; HTX, heterotaxy; IAA, interrupted aortic arch; LVNC, left ventricular noncompaction; MS, mitral stenosis; MVP, mitral valve prolapse; PA, pulmonary atresia; PAPVR, partial anomalous pulmonary venous return; PDA, patent ductus arteriosus; PS, pulmonary valve stenosis; PTA, persistent truncus arteriosus; RVOTO, right ventricular outflow tract obstruction; SIT, situs inversus totalis; SVAS, supravalvular aortic stenosis; TA, tricuspid atresia; TAA, thoracic aortic aneurysm; TAPVR, total anomalous pulmonary venous return; TGA, transposition of the great arteries; TOF, tetralogy of Fallot; VSD, ventricular septal defect.

TABLE 2 | CNVs associated with CHD.

Locus	Size (kb)	Mode	CNV	Copy number	Candidate genes for CHD	Type of CHD
1q21.1	418-3,981	*de novo*, inherited, n/a	Gain, Loss	3–45	*PRKAB2, FM05, CHD1L, BCL9, ACP6, GJA5, CD160, PDZK1, NBPF11, JM05, GJA8*	TOF, AS, COA, PA, VSD
3p25.1	175-12,380	*de novo*, inherited	Gain	2	*RAFJ, TMEM40*	TOF
3q22.1-3q26.1	680-32134	inherited, n/a	Gain, Loss	0–300	*FOXL2, NPHP3, FAM62C, CEP70, FAIM, PIK3CB, FOXL2, BPESC1*	DORV, TAPVR, AVSD
4q22.1	45	*de novo*	Gain	1	*PPM1K*	TOF
5q14.1-q14.3	4937-5454	Inherited, *de novo*	Gain	41,103	*EDIL3, VCAN, SSBP2, TMEM167A*	TOF
5q11.1	0.6	*de novo*	Gain	1	*ISL1*	TOF
5q35.3	264-1777	*de novo*, n/a	Gain	19–38	*CNOT6, GFPT2, FLT4, ZNF879, ZNF 345C, ADAMTS2, NSD1*	TOF
7q11.23	330-348	n/a	Gain	5–8	*FKBP6*	HLHS, Ebstein
8p23.1	67-12,000	n/a	Gain, Loss	4	*GATA4,NEIL2, FDFT1, CSTB, SOX7*	AVSD, VSD, TOF, ASD, BAV
9q34.3	190-263	*de novo*	Loss	2–9	*NOTCH1, EHMT1*	TOF, COA, HLHS
9q34.3	1.7	*de novo*	Gain	1	*NOTCH1*	TOF
11p15.5	256-271	n/a	Gain	13	*HRAS*	SV, AS
13q14.11	55-1430	n/a, *de novo*	Gain	7	*TNFSF11*	TOF, TAPVR, VSD, BAV
15q11.2	238-2,285	n/a	Loss	4	*TUBGCP5, CYFIP1, NIPA2, NIPA1*	COA, ASD, VSD, TAPVR
16p13.11	1414-2903	n/a	Gain	11–14	*MYH11*	HLHS
18q11.1-2	308-6118	n/a	Gain	1–28	*GATA6*	VSD
19p13.3	52-805	n/a, *de novo*	Gain, Loss	1–34	*MIER2, CNN2, FSTL3, PTBP1, WDR18, GNA11, S1PR4*	TOF
22q11.21	0.7-13	*de novo*	Gain	1	*PRODH*	TOF
Xp22.2	509-615	n/a	Gain	2–4	*MID1*	TOF, AVSD

CNV, copy number variation; CHD, congenital heart disease; ASD, atrial septal defect; VSD, ventricular septal defect; PDA, patent ductus arteriosus; TOF, tetralogy of Fallot; COA, coarctation of the aorta; TAPVR, total anomalous pulmonary venous return; AVSD, atrioventricular septal defect; PA, pulmonary atresia; DORV, double outlet right ventricle; BAV, bicuspid aortic valve; HLHS, hypoplastic left heart syndrome; AS, aortic stenosis; SV, single ventricle.

IP_3 binding to regulate various vital processes for diverse cell functions (19). As the modifications distinguishing the isoforms vary, such as phosphorylation sites, splicing sites, and associated molecules, each IP_3R may play a distinct role as a signaling hub offering different trajectories of cell signaling (20). In cardiovascular development, expression of IP_3R1 was detectably higher in the atrial than in the ventricular myocardium, IP_3R2 was mainly expressed in the trabecular layer of the ventricular myocardium, and IP_3R3 was uniformly expressed in the atrial and ventricular myocardia from embryonic day 9.5. These dynamic and complementary expression patterns of each subtype of IP_3R suggest their specific and/or redundant functions during the development of the heart. Although single subtype-knockout mice showed no developmental disorders and could survive after birth, IP_3R1-IP_3R2 double-knockout mice died *in utero* with developmental defects of the ventricular myocardium and atrioventricular canal of the heart, along with impaired Ca^{2+}-dependent calcineurin/NFATc signaling by embryonic day 11.5 (21). Moreover, IP_3R1-IP_3R3 double-knockout embryos showed hypoplasia of the outflow tract and the right ventricle,

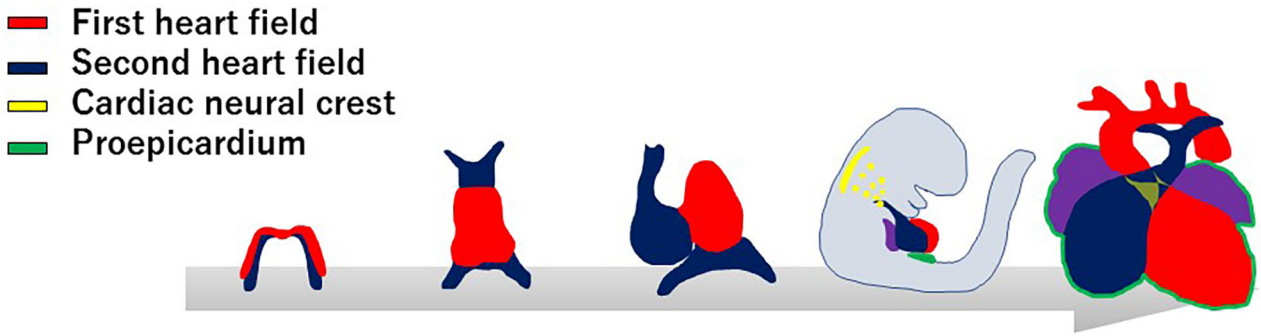


FIGURE 1 | Developmental origins of the cardiovascular system. The progenitor cells of the first heart field (red) form a cardiac crescent under the anterior part of the embryo, then move ventrally to the midline and form a linear heart tube. The second heart field (blue) is situated in the medial splanchnic pharyngeal mesoderm, and migrates to the anterior and posterior parts of the linear heart tube, providing the source of the right ventricle, outflow tract, and atrial cardiomyocytes. After looping of the heart tube, cardiac neural crest cells (yellow) migrate from the dorsal neural tube to pharyngeal arch arteries III, IV, and VI, and contribute to vascular smooth muscle cells of the aortic arch and the cardiac outflow tract. At the same time, the progenitor cells from proepicardial organ (green) contact the surface of the developing heart, giving rise to the epicardium.

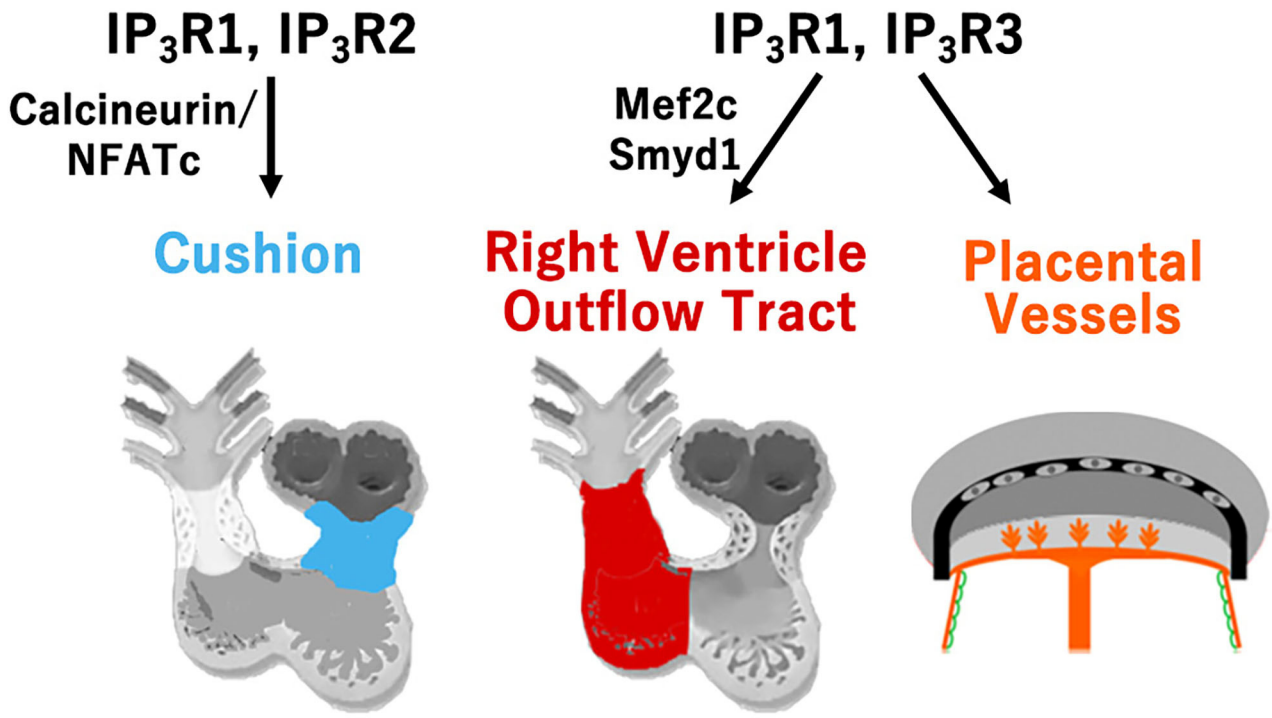


FIGURE 2 | Type-specific roles of inositol trisphosphate receptors (IP_3R) for cardiovascular development. Redundant roles of IP_3R1 and IP_3R2 in the development of atrioventricular cushion via calcineurin/NFATc signaling, and redundant roles of IP_3R1 and IP_3R3 in the development of the right ventricle and outflow tract via the Mef2c-Smyd1 pathway, and in the development of the extra-embryonic vessels of the placenta, allantois, and yolk sac are shown.

reduced expression of specific molecular markers, and enhanced apoptosis of mesodermal cells in the SHF with reduced activity of the Mef2C-Smyd1 pathway, a transcriptional cascade essential for the SHF (22). In addition, IP_3R1 and IP_3R3 were found to be required for extra-embryonic vascularization in the placenta, allantois, and yolk sac at the embryonic-maternal interface (23).

GENOMIC ARCHITECTURE OF CHD IMPLICATED WITH 22Q11.2 DELETION SYNDROME

22q11DS is the most common chromosomal microdeletion syndrome and is also known as DiGeorge syndrome or Takao syndrome (24, 25). 22q11DS is highly associated with

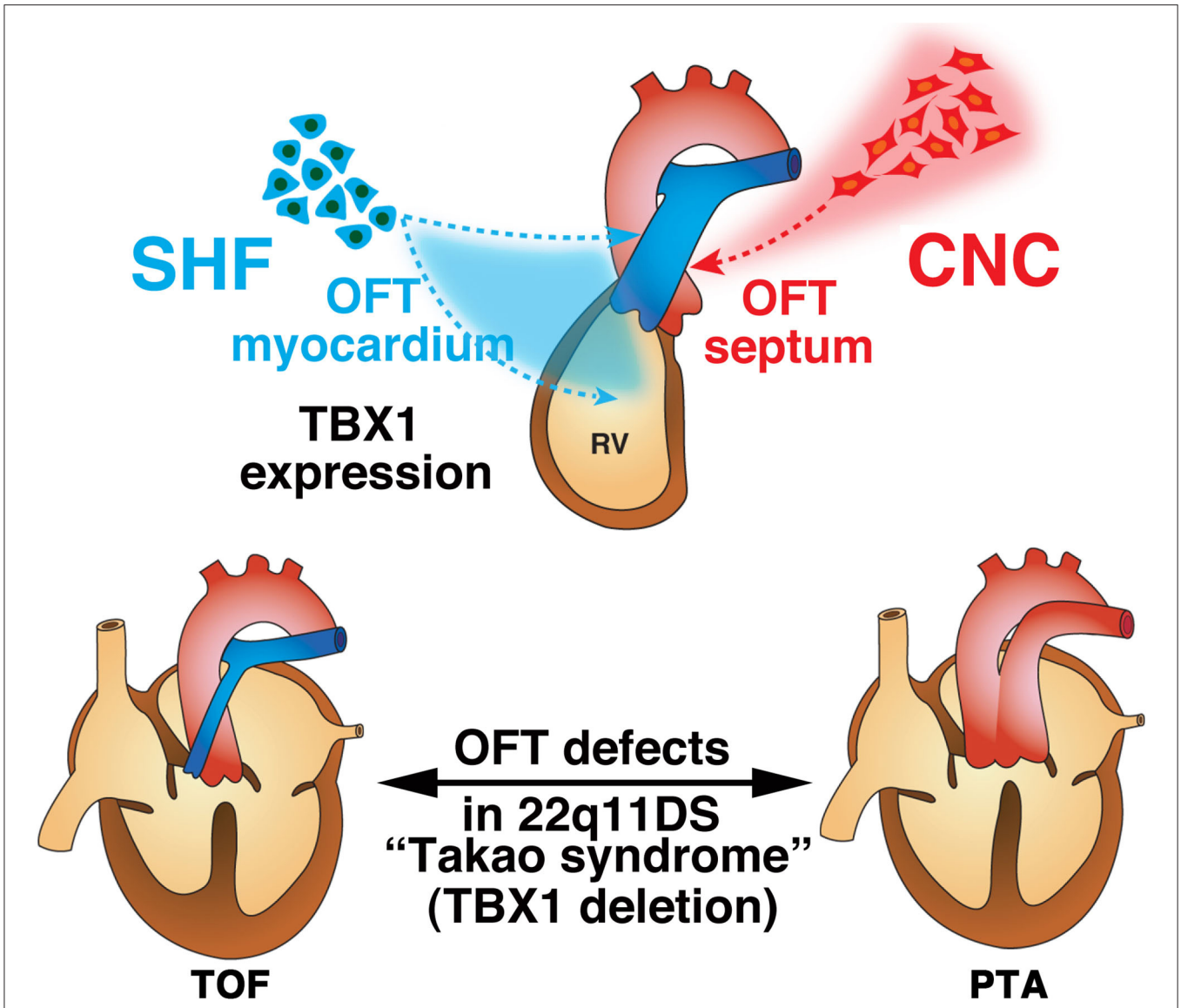


FIGURE 3 | Developmental architecture underlying 22q11.2 (*TBX1*) deletion. Second heart field (SHF) cells give rise to the cardiac outflow tract (OFT) myocardium, while cardiac neural crest (CNC) cells give rise to the OFT septum during normal development. The anatomical defects in tetralogy of Fallot (TOF) and persistent truncus arteriosus (PTA) are believed to result from malrotation of the OFT and aplasia of the OFT septum, respectively. The contribution of CNC cells is thought to be essential for proper rotation and septation of the OFT. Alternatively, developmental defects of the SHF may cause hypoplasia of the pulmonary infundibulum, resulting in TOF, and a more severe decrease in the number or absence of this subset of cells may affect the development and/or migration of CNC cells, resulting in PTA. This is consistent with the notion that OFT defects ranging from TOF to PTA are highly associated with 22q11.2 deletion syndrome (22q11DS).

CHD, involving the outflow tract, including persistent truncus arteriosus (PTA) and tetralogy of Fallot (TOF). Based on observations from experimental ablation of the CNC in chicken embryos, the outflow tract defects implicated in 22q11DS were thought to be the primary defect of the CNC development that leads to the outflow tract septum of the heart. At the beginning of the twenty-first century, the transcription factor TBX1 was identified to be the major etiology of outflow tract defects in this syndrome using new genetic engineering methods to model 22q11DS in mice (26–28). Mice with null or hypomorphic mutations for *Tbx1* demonstrate PTA (28, 29). Delineation of the expression pattern of TBX1 shed further light on the molecular and cellular basis of normal and abnormal development of the outflow tract. We and other groups surprisingly revealed that TBX1 was not expressed in the CNC, but was robustly expressed in the core region of pharyngeal mesoderm in the pharyngeal arch as well as in the SHF, pharyngeal endoderm, and head mesenchyme (30–32). Moreover, we showed that TBX1-expressing descendants that represent a subset of cells originated from the SHF, predominantly contribute to the right ventricular

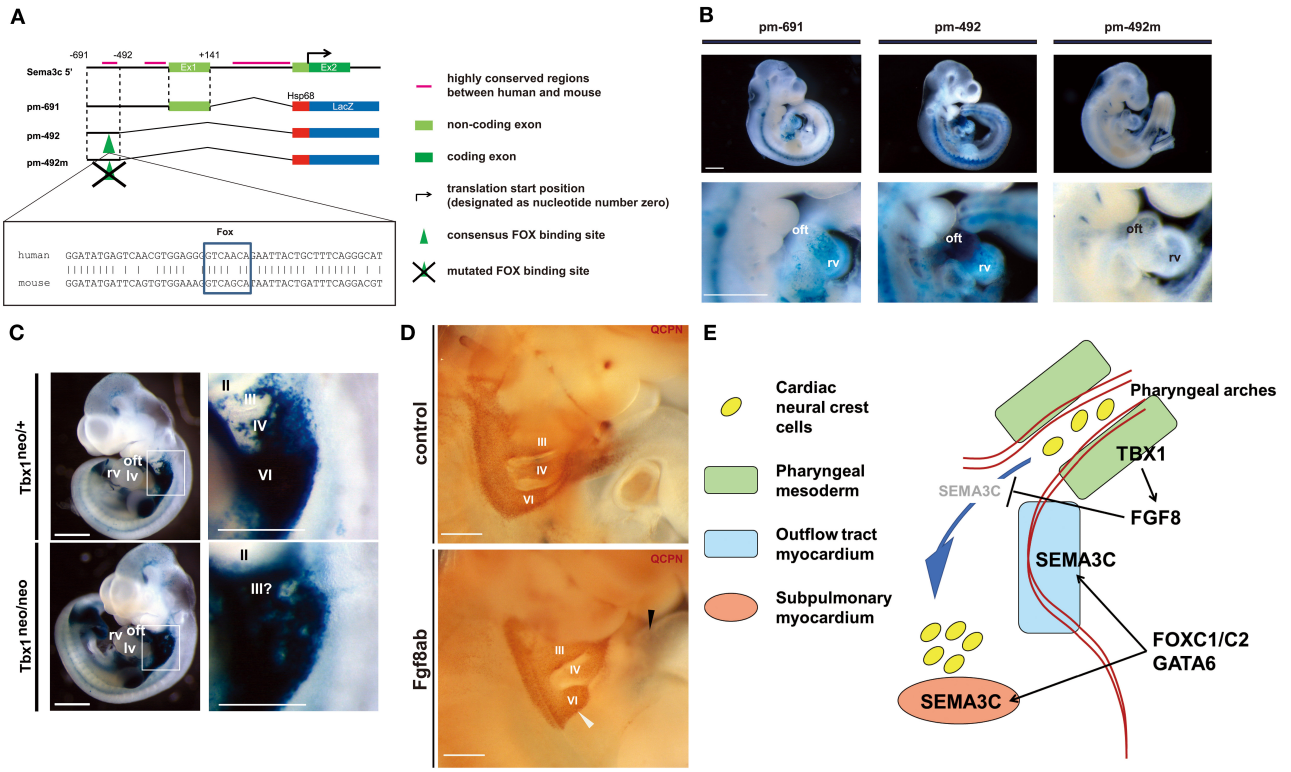


FIGURE 4 | The regulatory mechanism of *SEMA3C* expression during cardiogenesis by *FOXC1, FOXC2, GATA6,* and *TBX1-FGF8* signaling. **(A)** Genomic organization of the 5′ mouse *Sema3c* locus and flanking region. Green boxes indicate exons (Ex), black bars indicate highly conserved regions between human and mouse, and the translation start site (arrow) is designated as nucleotide number zero. Each construct number is indicated on the left. The consensus Fox-binding site is indicated by a green arrowhead. The construct pm-492m has a mutated Fox site in the context of pm-492. Conserved alignments of genomic sequences including the Fox site (box) between human (upper) and mouse (lower) are shown. **(B)** Right lateral views of the hearts of representative embryos obtained with each construct. Lower lane pictures show higher-magnification views of the heart. **(C)** *LacZ* expression in *Sema3c* int1/3′*-lacZ:Tbx1neo*/+ (control: upper, left) and *Tbx1neo/neo* (Tbx1 hypomorphic: lower, left) double-transgenic (tg) embryos at embryonic day (E)10.5, reminiscent of endogenous *Sema3c* mRNA expression. High-magnification views of the pharyngeal arch region (white box) are shown in the right lane. Enhanced *lacZ* expression in the pharyngeal arch region of *Sema3c* int1/3′*-lacZ* tg:*Tbx1neo/neo* embryos (lower, right) compared to *Sema3c* int1/3′*-lacZ* tg:*Tbx1neo*/+ embryos (upper, right). **(D)** Pharyngeal arch regions of chick-quail chimeras with bilateral transplantation of cardiac neural crest cells (cNCCs). Injection of anti-Fgf8 antibody (Fgf8ab) causes less migration of cNCCs from the 6th pharyngeal arch region (white arrowhead) to the outflow tract (black arrowhead) compared to that of the control. **(E)** A working model for the regulation of SEMA3C in developing cNCCs and second heart field (SHF) progenitor cells. During normal development, TBX1 may restrict the expression of SEMA3C in the SHF along the pharyngeal arch region by blocking ectopic SEMA3C expression in cNCCs via FGF8 signaling. FOXC1, FOXC2, and GATA6 may then synergistically activate the expression of SEMA3C in the SHF during migration into the outflow tract myocardium, leading to differentiation. Panels **(A–D)** are under a Creative Commons Attribution 4.0 license.

outflow tract and pulmonary trunk (33). These findings are very intriguing because they suggest that deletion of TBX1 in 22q11DS may result in defects of CNC-derived tissues in a non-cell-autonomous fashion through the cellular interaction between CNC and the SHF. It is believed that TOF results from malalignment of the outflow tract septum, leading to an overriding aorta with malaligned ventricular septal defect (34, 35). The developmental defects of CNC is considered to cause malalignment of outflow tract septum, thus leading to TOF. Alternatively, developmental defects of the SHF may cause hypoplasia of the right ventricular outflow tract that may also result in pulmonary stenosis and malalignment of the outflow tract septum with overriding aorta (34, 35). Our data about TBX1 in the SHF provides a new insight into the developmental mechanisms underlying TOF where cellular and molecular interaction of CNC and SHF are essential (33). As for PTA in 22q11DS or TBX1 deletion, it is considered that the TBX1-expressing descendants are severely decreased in number, affecting the development and/or migration of CNC cells, thus result in complete absence of the outflow tract septum. Indeed, we recently showed that PTA in mice hypomorphic for *Tbx1* might result from agenesis of the pulmonary trunk using IP_3R2-LacZ mice, in which a *LacZ* gene was genetically inserted in-frame at the translation initiation site of IP_3R2 locus on the mouse genome as a molecular marker (36). This developmental model is consistent of the observation that the outflow tract defects ranging from TOF to PTA are highly associated with 22q11DS (**Figure 3**).

EXPLORING THE GENOMIC ARCHITECTURE OF CHD AND THE REGULATORY MECHANISM UNDERLYING THE INTERACTION OF CARDIAC PROGENITOR LINEAGES

To further elucidate the genomic architecture of CHD, we performed mutation analysis using the genome bank of Japanese patients with non-syndromic CHD, and identified *GATA6* as the genetic cause of PTA (37). Mutations in *GATA6* disturb the transcriptional regulation of downstream target genes that play an essential role in cardiac development, including semaphorin 3C (*SEMA3C*) and plexin A2 (*PLXNA2*). SEMA3C is a neurovascular guiding molecule that functions as a ligand for PLXNA2 and an attractant for CNC cells (38). Mutation of *GATA6* eliminates activation of *SEMA3C* and *PLXNA2*. Mutation of the GATA sites on the enhancer elements of *SEMA3C* and *PLXNA2* abolished these transactivation activities in the outflow tract myocardium and the CNC derivatives in the outflow tract. Further analysis of the regulatory mechanism of *SEMA3C* revealed that a molecular network involving *GATA6, FOXC1/2, TBX1, SEMA3C,* and *FGF8* plays an important role in the interaction between SHF and CNC cells (39). Moreover, we found that *TBX1* restricts the expression of *SEMA3C* to the SHF in the pharyngeal arch region by inhibiting ectopic *SEMA3C* expression in CNC cells during migration via FGF8 signaling, whereas *GATA6, FOXC1*, and *FOXC2* activate the expression of *SEMA3C* in the SHF in the outflow tract myocardium at the same time. A recent report also showed the positive regulation of *SEMA3C* expression in the proximal outflow tract by TBX1 (40). This spatial and temporal regulation of *SEMA3C* expression is essential for proper homing of CNC cells from the pharyngeal region to the outflow tract. With loss of *TBX1*, downregulation of TBX1-FGF8 signaling in the pharyngeal region may lead to misexpression of *SEMA3C* in the migrating CNC cells, resulting in the failure of their migration with ectopic aggregation, ultimately causing outflow tract defects (**Figure 4**) (39). Although many other genes are also associated with the regulation of CNC cell migration, our results regarding the SEMA3C regulatory mechanism provide important evidence of interactions between CNC and the SHF for the developmental basis of CHD.

CONCLUDING REMARKS

In recent decades, detailed molecular biological analyses using genetically modified animals and accumulation of solid evidence from human mutation studies have dramatically advanced the understanding of cardiovascular development. In addition, with the recent development of stem cell science, including induced pluripotent stem cells and comprehensive expression analysis procedures using next-generation sequencing, elucidating the more detailed temporal and spatial gene regulatory mechanisms underlying cardiovascular development has become possible with evaluations at the single-cell level (9, 41–43). As a future direction for clinical application, detailed elucidation of the genomic architecture of CHD implicated in the mechanism regulating interactions between cells of multiple different origins would facilitate the development of regenerative medicine and/or preventive medicine for complex heart diseases such as CHD.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This study was supported by the Japan Agency for Medical Research and Development (AMED) under grant number 20ek0109487h0001.

REFERENCES

1. Nora JJ. Congenital heart disease. In: Nora JJ, Berg K, Nora AH, editors. *Genetics in Cardiovascular Diseases: Genetics, Epidemiology and Prevention*. New York, NY: Oxford University Press. (1991). p. 53–80.
2. Yamagishi H. Life-long managements for congenital heart diseases: from embryology to adult congenital heart diseases influencing the next generation. *Pediatr Cardiol Cardiovasc Surg.* (2010) 26:29–32.
3. Akhirome E, Walton NA, Nogee JM, Jay PY. The complex genetic basis of congenital heart defects. *Circ J.* (2017) 81:629–34. doi: 10.1253/circj.CJ-16-1343
4. Pierpont ME, Brueckner M, Chung WK, Garg V, Lacro RV, McGuire AL, et al. Genetic basis for congenital heart disease: Revisited: A scientific statement from the American Heart Association. *Circulation.* (2018) 138:e653–e711. doi: 10.1161/CIR.0000000000000606
5. Williams K, Carson J, Lo C. Genetics of congenital heart disease. *Biomolecules.* (2019) 9:879. doi: 10.3390/biom9120879
6. Srivastava D. Making or breaking the heart: From lineage determination to morphogenesis. *Cell.* (2006) 126:1037–48. doi: 10.1016/j.cell.2006.09.003
7. Buckingham M, Meilhac S, Zaffran S. Building the mammalian heart from two sources of myocardial cells. *Nat Rev Genet.* (2005) 6:826–35. doi: 10.1038/nrg1710
8. Kodo K, Yamagishi H. A decade of advances in the molecular embryology and genetics underlying congenital heart defects. *Circ J.* (2011) 75:2296–304. doi: 10.1253/circj.cj-11-0636
9. Kodo K, Yamagishi H. Current insights into genetics of congenital heart diseases: GATA and T-box cardiac transcription factors as the hotspot pathogenesis. *J Pediatr Cardiol Card Surg.* (2017) 1:18–27. doi: 10.24509/JPCCS.170105
10. Kelly RG, Brown NA, Buckingham ME. The arterial pole of the mouse heart forms from FGF10-expressing cells in pharyngeal mesoderm. *Dev Cell.* (2001) 1:435–40. doi: 10.1016/s1534-5807(01)00040-5
11. Mjaatvedt CH, Nakaoka T, Moreno-Rodriguez R, Norris RA, Kern MJ, Eisenberg CA, et al. The outflow tract of the heart is recruited from a novel heart-forming field. *Dev Biol.* (2001) 238:97–109. doi: 10.1006/dbio.2001.0409
12. Waldo KL, Kumiski DH, Wallis KT, Stadt HA, Hutson MR, Platt DH, et al. Conotruncal myocardium arises from a secondary heart field. *Development.* (2001) 128:3179–88.
13. Kirby ML, Gale TF, Stewart DE. Neural crest cells contribute to normal aorticopulmonary septation. *Science.* (1983) 220:1059–61. doi: 10.1126/science.6844926
14. Hutson MR, Kirby ML. Model systems for the study of heart development and disease. Cardiac neural crest and conotruncal malformations. *Semin Cell Dev Biol.* (2007) 18:101–10. doi: 10.1016/j.semcdb.2006.12.004

15. Yamagishi H. Cardiac neural crest. *Cold Spring Harb Perspect Biol.* (2021) 13:a036715. doi: 10.1101/cshperspect.a036715
16. Arima Y, Miyagawa-Tomita S, Maeda K, Asai R, Seya D, Minoux M, et al. Preotic neural crest cells contribute to coronary artery smooth muscle involving endothelin signalling. *Nat Commun.* (2012) 3:1267. doi: 10.1038/ncomms2258
17. Mikawa T, Gourdie RG. Pericardial mesoderm generates a population of coronary smooth muscle cells migrating into the heart along with ingrowth of the epicardial organ. *Dev Biol.* (1996) 174:221–32. doi: 10.1006/dbio.1996.0068
18. Olivey HE, Svensson EC. Epicardial-myocardial signaling directing coronary vasculogenesis. *Circ Res.* (2010) 106:818–32. doi: 10.1161/CIRCRESAHA.109.209197
19. Berridge MJ. The inositol trisphosphate/calcium signaling pathway in health and disease. *Physiol Rev.* (2016) 96:1261–96. doi: 10.1152/physrev.00006.2016
20. Mikoshiba K. Role of IP3 receptor signaling in cell functions and diseases. *Adv Biol Regul.* (2015) 57:217–27. doi: 10.1016/j.jbior.2014.10.001
21. Uchida K, Aramaki M, Nakazawa M, Yamagishi C, Makino S, Fukuda K, et al. Gene knock-outs of inositol 1,4,5-trisphosphate receptors types 1 and 2 result in perturbation of cardiogenesis. *PLoS ONE.* (2010) 5:e12500. doi: 10.1371/journal.pone.0012500
22. Nakazawa M, Uchida K, Aramaki M, Kodo K, Yamagishi C, Takahashi T, et al. Inositol 1,4,5-trisphosphate receptors are essential for the development of the second heart field. *J Mol Cell Cardiol.* (2011) 51:58–66. doi: 10.1016/j.yjmcc.2011.02.014
23. Uchida K, Nakazawa M, Yamagishi C, Mikoshiba K, Yamagishi H. Type 1 and 3 inositol trisphosphate receptors are required for extra-embryonic vascular development. *Dev Biol.* (2016) 418:89–97. doi: 10.1016/j.ydbio.2016.08.007
24. Yamagishi H. The 22q11.2 deletion syndrome. *Keio J Med.* (2002) 51:77–88. doi: 10.2302/kjm.51.77
25. Yamagishi H, Srivastava D. Unraveling the genetic and developmental mysteries of 22q11 deletion syndrome. *Trends Mol Med.* (2003) 9:383–9. doi: 10.1016/s1471-4914(03)00141-2
26. Lindsay EA, Vitelli F, Su H, Morishima M, Huynh T, Pramparo T, et al. Tbx1 haploinsufficieny in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature.* (2001) 410:97–101. doi: 10.1038/35065105
27. Merscher S, Funke B, Epstein JA, Heyer J, Puech A, Lu MM, et al. TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. *Cell.* (2001) 104:619–29. doi: 10.1016/s0092-8674(01)00247-1
28. Jerome LA, Papaioannou VE. DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. *Nat Genet.* (2001) 27:286–91. doi: 10.1038/85845
29. Hu T, Yamagishi H, Maeda J, McAnally J, Yamagishi C, Srivastava D. Tbx1 regulates fibroblast growth factors in the anterior heart field through a reinforcing autoregulatory loop involving forkhead transcription factors. *Development.* (2004) 131:5491–502. doi: 10.1242/dev.01399
30. Garg V, Yamagishi C, Hu T, Kathiriya IS, Yamagishi H, Srivastava D. Tbx1, a DiGeorge syndrome candidate gene, is regulated by sonic hedgehog du ring pharyngeal arch development. *Dev Biol.* (2001) 235:62–73. doi: 10.1006/dbio.2001.0283
31. Yamagishi H, Maeda J, Hu T, McAnally J, Conway SJ, Kume T, et al. Tbx1 is regulated by tissue-specific forkhead proteins through a common Sonic hedgehog-responsive enhancer. *Genes Dev.* (2003) 17:269–81. doi: 10.1101/gad.1048903
32. Xu H, Morishima M, Wylie JN, Schwartz RJ, Bruneau BG, Lindsay EA, et al. Tbx1 has a dual role in the morphogenesis of the cardiac outflow tract. *Development.* (2004) 131:3217–27. doi: 10.1242/dev.01174
33. Maeda J, Yamagishi H, McAnally J, Yamagishi C, Srivastava D. Tbx1 is regulated by forkhead proteins in the secondary heart field. *Dev Dynam.* (2006) 235:701–10. doi: 10.1002/dvdy.20686
34. Thom T, Haase N, Rosamond W, Howard VJ, Rumsfeld J, Manolio T, et al. Heart disease and stroke statistics—2006 update: a report from the american heart association statistics committee and stroke statistics subcommittee. *Circulation.* (2006) 113:e85–e151. doi: 10.1161/CIRCULATIONAHA.105.171600
35. Siwik ES, Patel CR, Zahka KG. Tetralogy of fallot. In: Allen HD, Gutgesell HP, Clark EB, Driscoll DJ, editors. *Moss and Adams' Heart Disease in Infants, Children, and Adolescents Including the Fetus and Young Adult*. Philadelphia: Lippincott Williams & Wilkins (2001). p. 880–902.
36. Ishizaki-Asami R, Uchida K, Tsuchihashi T, Shibata A, Kodo K, Emoto K, et al. Inositol 1,4,5-trisphosphate receptor 2 as a novel marker of vasculature to delineate processes of cardiopulmonary development. *Dev Biol.* (2020) 458:237–45. doi: 10.1016/j.ydbio.2019.11.011
37. Kodo K, Nishizawa T, Furutani M, Arai S, Yamamura E, Joo K, et al. GATA6 mutations cause human cardiac outflow tract defects by disrupting semaphorin-plexin signaling. *Proc Natl. Acad Sci USA.* (2009) 106:13933–8. doi: 10.1073/pnas.0904744106
38. Toyofuku T, Yoshida J, Sugimoto T, Yamamoto M, Makino N, Takamatsu H, et al. Repulsive and attractive semaphorins cooperate to direct the navigation of cardiac neural crest cells. *Dev Biol.* (2008) 321:251–62. doi: 10.1016/j.ydbio.2008.06.028
39. Kodo K, Shibata S, Miyagawa-Tomita S, Ong SG, Takahashi H, Kume T, et al. Regulation of Sema3c and the interaction between cardiac neural crest and second heart field during outflow tract development. *Sci Rep.* (2017) 7:6771. doi: 10.1038/s41598-017-06964-9
40. Théveniau-Ruissy M, Dandonneau M, Mesbah K, Ghez O, Mattei MG, Miquerol L, et al. The del22q11.2 candidate gene Tbx1 controls regional outflow tract identity and coronary artery patterning. *Circ Res.* (2008) 103:142–8. doi: 10.1161/circresaha.108.172189
41. Kodo K, Ong SG, Jahanbani F, Termglinchan V, Hirono K, InanlooRahatloo K, et al. iPSC-derived cardiomyocytes reveal abnormal TGF-β signalling in left ventricular non-compaction cardiomyopathy. *Nat Cell Biol.* (2016) 18:1031–42. doi: 10.1038/ncb3411
42. Zhang JZ, Termglinchan V, Shao NY, Itzhaki I, Liu C, Ma N, et al. A human iPSC double-reporter system enables purification of cardiac lineage subpopulations with distinct function and drug response profiles. *Cell Stem Cell.* (2019) 24:802–11.e5. doi: 10.1016/j.stem.2019.02.015
43. de Soysa TY, Ranade SS, Okawa S, Ravichandran S, Huang Y, Salunga HT, et al. Single-cell analysis of cardiogenesis reveals basis for organ-level developmental defects. *Nature.* (2019) 572:120–4. doi: 10.1038/s41586-019-1414-x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 09 April 2021
doi: 10.3389/fcvm.2021.631374

Prenatal Genetic Diagnosis in Three Fetuses With Left Heart Hypoplasia (LHH) From Three Unrelated Families

Sukun Luo[1†], Luyi Chen[2†], Weizhong Wei[3†], Li Tan[1], Meng Zhang[3], Zhengrong Duan[2], Jiangxia Cao[2], Yan Zhou[2], Aifen Zhou[2*] and Xuelian He[1*]*

[1] Precision Medical Center, Tongji Medical College, Wuhan Children's Hospital (Wuhan Maternal and Child Healthcare Hospital), Huazhong University of Science and Technology, Wuhan, China, [2] Prenatal Diagnosis Center, Tongji Medical College, Wuhan Children's Hospital (Wuhan Maternal and Child Healthcare Hospital), Huazhong University of Science and Technology, Wuhan, China, [3] Ultrasonic Diagnosis Department, Tongji Medical College, Wuhan Children's Hospital (Wuhan Maternal and Child Healthcare Hospital), Huazhong University of Science and Technology, Wuhan, China

OPEN ACCESS

Edited by:
Seitaro Nomura,
The University of Tokyo, Japan

Reviewed by:
Olivier M. Vanakker,
Ghent University, Belgium
Douglas A. Marchuk,
Duke University, United States

***Correspondence:**
Xuelian He
hexuelian2013@hotmail.com
Aifen Zhou
937577332@qq.com
Yan Zhou
zhouyan030118@126.com

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: 19 November 2020
Accepted: 04 March 2021
Published: 09 April 2021

Citation:
Luo S, Chen L, Wei W, Tan L, Zhang M, Duan Z, Cao J, Zhou Y, Zhou A and He X (2021) Prenatal Genetic Diagnosis in Three Fetuses With Left Heart Hypoplasia (LHH) From Three Unrelated Families. Front. Cardiovasc. Med. 8:631374. doi: 10.3389/fcvm.2021.631374

Background: Congenital heart defects (CHDs) are the most common birth defects, and left heart hypoplasia (LHH) is a severe form of CHD and responsible for more than 20% cardiac deaths during the first week of life, however, its genetic causes remain largely elusive.

Methods: Three families with fetal LHH were recruited. Genomic DNA from amniotic fluid or peripheral blood, and trio whole exome sequencing (trio-WES) and copy number variation sequencing (CNV-seq) were performed.

Results: All the three couples had no family history, and mid-gestation ultrasound revealed LHH and other variable cardiovascular defects in the fetuses. Trio-WES revealed *de novo* pathogenic variations in *KMT2D* (p.Gly3465Aspfs*37) (NM_003482) and *WDFY3* (p.Ser117Xfs*) (NM_014991), and CNV-seq identified a deletion of 150 kb encompassing *NOTCH1*. *KMT2D* and *NOTCH1* previously have been reported to be associated with CHDs, however, *WDFY3* is reported for the first time to be possibly related to CHD in human.

Conclusion: Our study suggested that genetic component is an important risk factor for the development of LHH, and next generation sequencing is a powerful tool for genetic diagnosis in fetuses with CHDs and genetic counseling, however, more studies and data are need to establish the correlation of fetal phenotypes and genotypes.

Keywords: prenatal diagnosis, congenital heart defects, *KMT2D*, *NOTCH1*, *WDFY3*

INTRODUCTION

Congenital heart defects (CHDs) are a spectrum of diseases involving structural abnormalities of the heart and great vessels that affect about 1% of liveborn infants and occur in 10% of aborted fetuses (1). Left heart hypoplasia (LHH), a severe and complex phenotype of CHDs, refers to the underdeveloped left heart structures including left ventricle, aorta and mitral valve (2). Hypoplastic left heart syndrome (HLHS) is used to describe the most severe condition of LHH, and cause 23% of cardiac deaths in the first week of life without early surgery (3). Although a three-stage surgical palliation has significantly increased the 5-year survival rate (50–70%) (4), the extent of morbidity

is considerable and life-expectance remains reduced, and some of them will still develop heart failure and need cardiac transplantation. To achieve substantive progress in the development of therapeutic intervention, the understanding of the formation of the 4-chambered vertebrate heart and the pathogenic mechanism causing LHH is needed.

Several lines of evidence support a genetic contribution to HLHS. A number of reports suggest HLHS is highly associated with chromosomal abnormalities (Turner, trisomy 13, 18, and 21) (5–7), as well as copy number variations (CNV) (DiGeorge, and Jacobsen syndrome) (8, 9). Genetic studies have also identified variants in several candidate genes such as *NKX2.5*, *GJA1*, *NOTCH1*, *HAND1*, *FOXC2*, *FOXL1*, *IRX4*, *MYH6*, and *RBFOX2*, in patients with HLHS (10–15), and these genes are very important for cardiogenesis.

The aim of this study is to identify the pathogenic variants in 3 fetuses with LHH by using next generation sequencing, and we found a *de novo* loss-of-function variant in *KMT2D*, and *WDFY3* (WD40 repeat and FYVE domain containing 3) gene, respectively, and a *de novo* heterozygous microdeletion variant in chromosome 9q34.3, encompassing *NOTCH1*, which aid prenatal genetic counseling and understanding of the pathogenesis of LHH.

METHODS

Patient Enrollment

This study was approved by the Wuhan Children's Hospital (Wuhan Maternal and Child Healthcare Hospital) ethics committee and informed consent was obtained from the parents of the fetuses. We recruited 3 fetuses with LHH detected by ultrasound during middle gestation. The parents were healthy and non-consanguineous, and denied any adverse history. The maternal age range at diagnosis of LHH was 29–30 years, and the gestation age range at diagnosis was 23–25 weeks. Fetal samples were from amniotic fluid during the invasive prenatal diagnosis. A total of 2 ml of peripheral blood was collected in EDTA-containing tubes from the couples. Multidisciplinary consultation including genetic counseling regarding the risk of LHH and the chance of surgery as well as the benefits and limitations of whole exome sequencing (WES) and CNV sequencing (CNV-seq) were introduced to the couples.

Sample Preparation

Genomic DNA was extracted from 20 ml of amniotic fluid or 2 ml whole blood according to the protocol of the Omega DNA Mini Kit. A NanoDropTM spectrophotometer was used for quality control (QC) of DNA purity and concentration.

Whole Exome Sequencing and Data Analysis

Genomic DNA was sheared by sonication and exome sequences were enriched by IDT xGen® Exome Researcher Panel v1.0, according to the manufacturer's protocol. DNA libraries were sequenced on Illumina Hiseq XTen with a method of pair-end 150 bp reads. Raw image files were processed by the Bcl To Fastq (Illumina) for base calling and raw data production. Low quality variants were filtered out using the quality score $\geq$ 20 (Q20) standard. The sequencing reads were then mapped to the human reference genome (hg19/GRCh37) with BWA. Single nucleotide variations (SNVs) and Indels were scored by GATK 3.8. All short variants were annotated with databases including 1,000 genomes, dbSNP, Exac, GnomAD, ClinVar, HGMD, and OMIM. Impacts of variants on protein functions were predicted using software packages Provean, SIFT, Polyphen2_HDIV, Polyphen2_HVAR, Mutationtaster, M-CAP, and REVEL. The pathogenicity of candidate variants was evaluated on the base of mutated frequency, conservation of amino acid change and structural or function domain of the protein as well as inheritance pattern. The pathogenicity of identified variants was assessed in accordance with the American College of Medical Genetics (ACMG) guidelines and Sherloc (16, 17).

Generally, the variants detected by Trio-WES were analyzed according to ACMG standards and guidelines, and classified into pathogenic variants, likely pathogenic variants, uncertain significance variants, likely benign variants and benign variants. The former three types were further annotated in OMIM diseases database and filtered through family co-segregation according to genetic rules. Potential pathogenic genes were identified combining disease correlation and clinical observations. In addition, the incidental findings were disposed according to the document of ACMG about prenatal diagnosis (18). Pathogenic and likely pathogenic variants causing neurodevelopmental disorders, intellectual disability or metabolic diseases were reported.

CNV Analysis

A total amount of more than 1.2 μg purified genomic DNA was broken into 200–300 bp fragments by sonication, followed by fragment DNA being sequentially repaired, tailed, ligated with an adaptor and amplified for library construction. The samples were then subjected to Illumina NovaSeq 6000 (Illumina, San Diego, USA). Raw data were processed by the basecall analysis software fastq v0.18.1. The clean data were then mapped to human reference genome (hg19) using BWA. The candidate CNVs were first filtered through normal frequency databases DGV, and, subsequently, were annotated based on scientific literature review and public databases including Decipher, ClinVar, ClinGen, ISCA, and dbVar. According to the ACMG, the candidate CNVs were classified into five categories: pathogenic, likely pathogenic, uncertain clinical significance, likely benign and benign (19).

Sanger Sequencing

The identified variants in the *KMT2D* and *WDFY3* genes were confirmed by Sanger sequencing. Primers were designed to amplify the exon 36 of *KMT2D* and the exon 6 of *WDFY3*. PCR products were subjected to ABI 3730XL and analyzed by DNAStar software.

Quantitative PCR Assay

Semi-quantitative qPCR was performed to confirm the presence of the microdeletion identified in case 3 by detecting the copies of DNA fragments within the region on both the fetus and the parents. Primers were designed to quantify the relative copies of

TABLE 1 | Prenatal phenotype and genotype information for the cohort.

	Case 1	Case 2	Case 3
Maternal age	30 Y	29 Y	30 Y
Prenatal Gestational age	23 W	24 W + 3	25 W
Congenital heart defect	HLH, double-outlet right ventricle, perpetuate left superior vena cava, cardiac arrhythmias, single umbilical artery	Small left heart, coaractation of aortic arch, ascending arota, small VSD	HLH, atresia of mitral valves, stenosis of aortic valves, VSD
Extracardiac malformation	Horseshoe kidney	Congenital cystic adenomatoid malformation of the lung(type II), polydactylism	/
NT	3.5 mm	1.8 mm	1.2 mm
Karyotype	Normal	NA	Normal
CMA	Normal	del9q34.3	Normal
Pathogenic gene	*KMT2D*	*NOTCH1*, *INPP5E*, *PMPCA*	*WDFY3*
Variant	c.10394delG (exon36) NM_003482	/	c.349_c.350delAG (exon6) NM_014991
Amino acid change	p.Gly3465Aspfs*37	/	p.Ser117Xfs*1
AF in gnomAD	Absence	Absence	Absence
HGMD	Included	Not included	Not included
ClinVar	Not included	Not included	Not included
Zygosity	Heterozygous	Heterozygous	Heterozygous
Mutation type	Frameshift	Microdeletion	Frameshift
Parental origin	*De novo*	*De novo*	*De novo*
Disease	Kabuki syndrome 1	AOS 5/Aortic valve disease 1	Microcephaly 18
Inheritance Model	AD	AD	AD
Reference (PMID)	21671394	26820064	Novel
Pathogenicity	Pathogenic	Pathogenic	Pathogenic

**HLH, Hypoplasia left heart; VSD, ventricular septal defect; AOS, Adams-Oliver syndrome; AF, allele frequency; AD, autosomal dominant.*

PMPCA exon 2 and 13, *INPP5E* exon 10 as well as *NOTCH1* exon 1 and exon 34. The primers for these DNA fragments are listed as following: *PMPCA* exon 2(Forward: 5′-GTG CCT ATC CCA ACA TCC-3′ and Reverse: 5′-GCG AAG CCC ATT ATC CAA-3′), *PMPCA* exon 13(Forward: 5′-CCG TTC CCG TGC GTG TTA-3′ and Reverse: 5′-TTC CTG GCT TGC GGT GGT-3′), *INPP5E* exon 10(Forward: 5′-GCA CCA TCT GCT CCG TTT C-3′ and Reverse: 5′-TTC CTT CCT GGG ACG CTG-3′), *NOTCH1* exon 1(Forward: 5′-GAG CGC AGC GAA GGA ACG AG-3′ and Reverse: 5′-CCT CTC TTC CCC GGC TGG CT-3′) and *NOTCH1* exon 34 (Forward: 5′-GCA CAG GAG CGC ATG CAT CA-3′ and Reverse: 5′-GCT GGG CTT GCG GAC CTT CT-3′). The 36B4 gene was used as reference gene. q-PCR was performed on a ABI 7500 apparatus using SYBR Green Real-Time PCR Master Mix (Takara) according to the following conditions: 95°C for 2 min, followed by 40 cycles of 95°C for 15 s, 53°C for 15 s and 72°C for 34 s.

RESULTS

After performing genetic analysis, variants of two genes, *KMT2D* (NM_003482) and *WDFY3*(NM_014991), and a microdeletion involving *NOTCH1* were detected in three fetuses, respectively. All the three variants were classified into pathogenic variants because they were LOF and *de novo* mutations, and absent in control population. The detailed clinical features and genetic information were shown in **Table 1**. No other pathogenic variants in known pathogenic genes, which were in line with family separation, were identified in the three cases.

Case 1 was identified to be a *de novo* heterozygous variant (c.10394delG, p.Gly3465Aspfs*37) in the exon 36 of *KMT2D* gene (**Figure 1A**), which was reported previously and included in the HGMD (Accession number: CD114940) (20). At 23-week of gestation (wg), the ultrasound, shown in **Figures 2A,B**, revealed complex congenital heart defects (hypoplastic left heart, ventricular septal defect, double-outlet right ventricle, and perpetuate left superior vena cava), horseshoe kidney, and single umbilical artery, and the fetus presented arrhythmia. Mid-pregnancy ultrasound showed the fetus had increased nuchal translucency of 3.5 mm (normal value < 2.5 mm) at 12 wg, and, at 18 wg, the tests of karyotyping and CMA analyses by amniocentesis were negative.

Case 2 was found to carry a *de novo* 150 kb microdeletion at 9q34.3 (chr9:139, 315, 643-139, 465, 759), resulting in deletion of full-length of *NOTCH1, INPP5E, and SEC16A* as well as part of *PMPCA* (**Figure 1C**). The ultrasound at 24 wg detected severe congenital heart defects, including small left heart (**Figure 2C**), coarctation of ascending aorta and aortic ache, and a small ventricular septal defect. Polydactyly and cystic adenomatoid of lung were also observed (**Table 1**). The nuchal

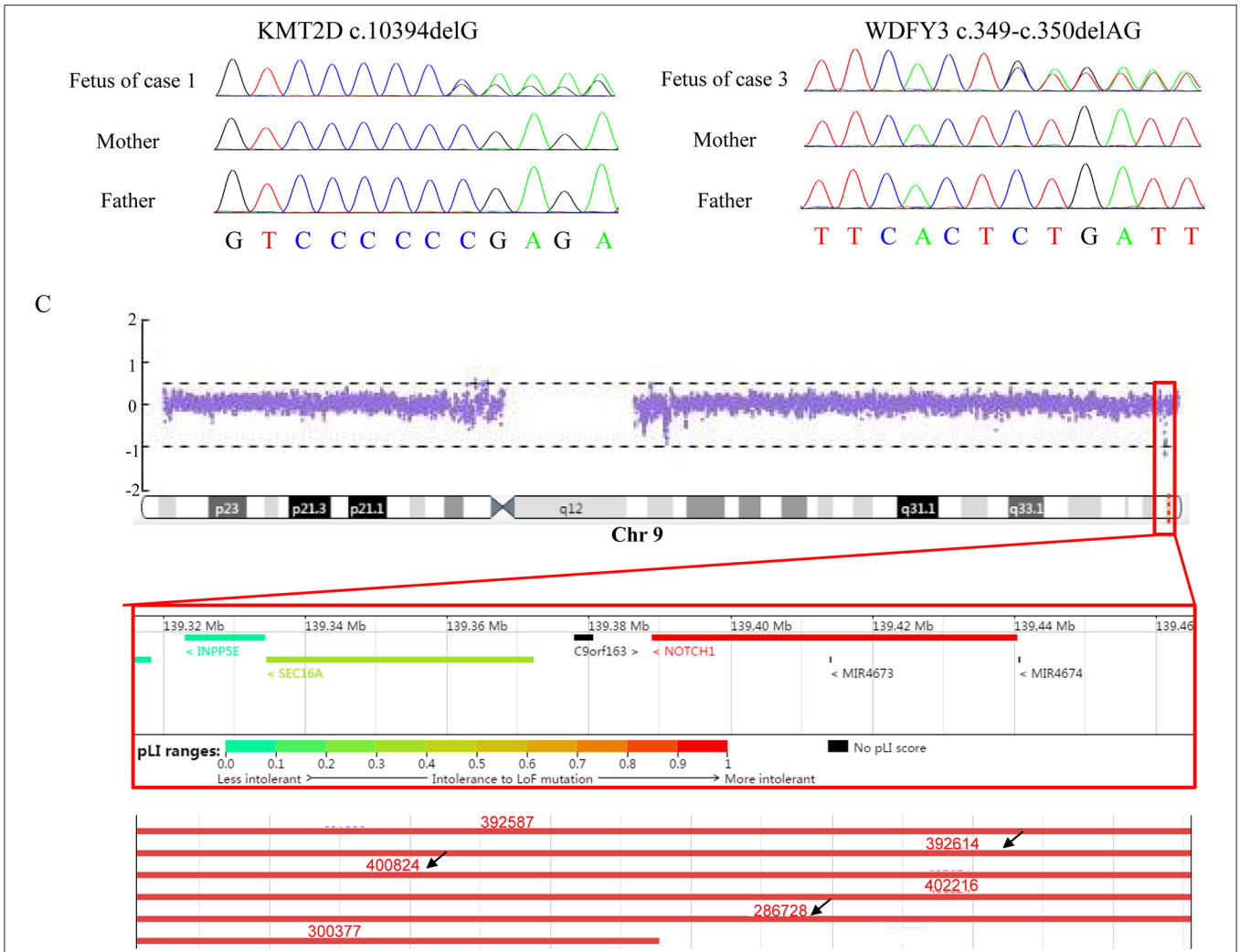


FIGURE 1 | Results of Sanger sequencing and CNV-seq. **(A)** The mutation, c.10394delG, identified in *KMT2D* in the fetus (case 1) but not in the parents. **(B)** The mutation, c.394-c.350delAG, found in *WDFY3* in the fetus (case 3) but not in the parents. **(C)** A 150 kb microdeletion at 9q34.3 resulting in deletion of full-length of *NOTCH1* detected in case 2 and part of microdeletion variants encompassing the region in Decipher. Variants pointed by black arrow displayed heart abnormalities.

translucency measurement and noninvasive prenatal testing were unremarkable.

Case 3 was identified to have a *de novo* heterozygous frameshift variant (c.349-c.350delAG, p.Ser117Xfs*1) in the exon 6 of *WDFY3* gene (**Figure 1B**), which was a novel pathogenic variant. Ultrasound of this fetus at 25 wg showed left heart dysplasia with mitral atresia and aortic valve stenosis (**Figures 2D,E**), and ventricular septum defect. No other abnormality was found during mid-pregnancy examination. Although, *WDFY3* pathogenic variants were reported to cause microcephaly or macrocephaly (21, 22), the described fetus in this case had normal head circumference.

DISCUSSION

Although increasing evidence supports a genetic etiology (23–25), the precise pathogenesis of LHH or HLHS is unknown. In this study, we reported 2 *de novo* frameshift variations in the *KMT2D* and *WDFY3* genes, respectively, and a 9q34.3 microdeletion of 150 kb encompassing *NOTCH1*, in three fetuses with LHH and other structural malformations.

WES identified a *de novo* heterozygous frameshift variant (c.10394delG) in the *KMT2D* gene, resulting in a truncated protein, in Case 1. The *KMT2D* encodes a histone H3K4 methytransferases that is associated with transcriptionally active genes. Mutations in *KMT2D* are most frequently, up to 70%, found in patients with Kabuki syndrome (KS, MIM: #147920), which is a rare, autosomal dominant multiple congenital anomaly syndrome characterized by recognizable facial features, global developmental delay, intellectual disability, short stature, and musculoskeletal abnormalities (26). Approximately 55% of patients with *KMT2D* mutations have been reported to present CHDs, including both septal defects and aortic coarctation (27). Studies in animal models have indicated *KMT2D* is essential for

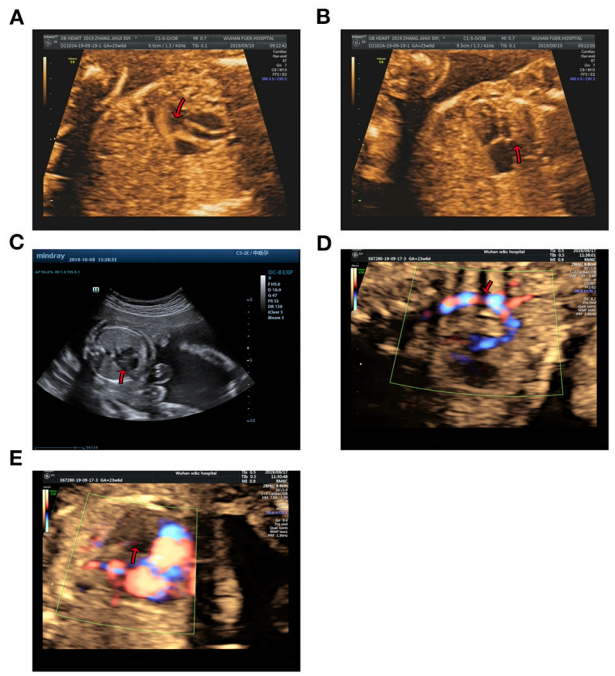


FIGURE 2 | Echocardiographic images of fetal heart structures. **(A,B)** show fetal ventricular septal defect and hypoplasia of left heart in case 1. **(C)** displays a smaller left heart in case 2. **(D,E)** show coarctation of the aorta and mitral atresia in the fetal heart of case 3.

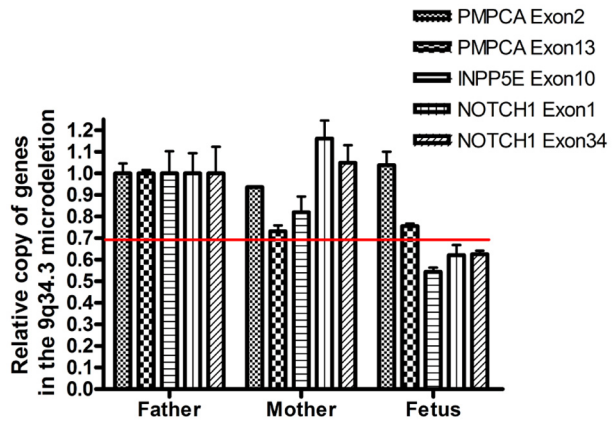


FIGURE 3 | Confirmation of the presence of the deletion by semi-quantitative q-PCR. The copy number of *PCPMA*, *INPP5E*, *NOTCH1,* and *36B4* was evaluated by semi-quantitative PCR. The relative copies of *PCPMA* (exon 2 and 13), *INPP5E* (exon 10) and *NOTCH1* (exon 1 and 34) were normalized using, a valid single copy gene.

regulating cardiac gene expression during heart development and is required for the establishment of the primary and secondary heart fields (28, 29). Recently, a study including 80 fetuses with congenital cardiac left-sided lesions showed that *KMT2D* was the most frequently mutated gene (10.6%) followed by *NOTCH1* (6.1%) (30). *NOTCH1*, encoding one of core receptors in Notch signaling pathway, is a well-known CHD-related gene, and loss-of-function variants in this gene confer a higher risk for and segregates with left-sided-CHD (31, 32). Deletion of the whole *NOTCH1* gene was previously reported to cause non-syndromic Tetralogy of Fallot and HLHS (32, 33). Our Case 2 carried a 9q34.3 microdeletion surrounding full-length *NOTCH1*, and presented underdeveloped left heart and ventricular septal defect. In addition to *NOTCH1*, the microdeletion contained three other protein-coding genes, *PMPCA*, *INPP5E,* and *SEC16A* according to human hg19/GRCh37. However, our q-PCR assay confirmed that the microdeletion might not destroy the *PMPCA* gene which contained 13 exons, since the relative copies of *PMPCA* exon2 and exon 13 in the fetus was comparable to that in its parents (**Figure 3**). *INPP5E* were associated with autosomal recessive diseases, while *SEC16A* had not been reported to be related to a known disease. Deletion of the whole *NOTCH1* gene was the most likely cause for LHH of the fetus in case 2.

Although a clear prenatal genetic diagnosis can be made only when definitely pathogenic genes and variants are identified, prenatal WES is invaluable for the diagnosis of causal genes but also for the discovery of candidate genes. In case 3, we did not detect any pathogenic variants in well-established CHD pathogenic genes, but identified a frameshift (c.349-c.350delAG, p.Ser117Xfs*1) variant in *WDFY3*, thus, we propose novel candidate gene for early cardiac development supported by functional and phenotyping evidence in animal model (34). *WDFY3* is located on chromosome 4q21.23, encompassing 74 exons and encoding a protein with 3,526 amino acids. The encoded protein is an adaptor protein involved in selective degradation of protein aggregates by autophagy, and is critical for brain development (35, 36). It was included in OMIM and associated with microcephaly 18, an autosomal dominant disorder. It is extremely intolerant to loss-of-function (LOF) mutation with a pLI score of 1. Pathogenic *WDFY3* variants caused mild to moderate neurodevelopmental delay and intellectual disability, most of which (8/13) were loss-of-function variant and putatively lead to haploinsufficiency (21).

Homozygous knockout of this gene in mice model would lead to variable cardiac anomalies, such as ventricular septum defect, double outlet right ventricle, aortic overriding, thinning of ventricular wall, ventricular dilation and disorganized trabeculation, and abnormal NOTCH1 signaling in the heart of mice (34), suggesting that *WDFY3* is critical for cardiac development. However, this report did not mention whether heterozygous *WDFY3* knockout mice presented visible heart abnormalities (34). Unfortunately, we did not obtain the heart tissue of the aborted fetus to examine the impacts of *WDFY3* variant on signaling pathways, such as NOTCH1 and Wnt pathways related to early cardiac development. This is the first report of a human phenotype caused by a frameshift variant in *WDFY3*, more cases and further functional studies are

needed before we make a conclusion that *WDFY3* is pathogenic candidate gene for LHH. In this study, the pregnancy was terminated due to the poor prognosis of LHH and the risk of neurodevelopmental delay. The pregnant woman and her family were informed that the recurrence risk is very low except that the variant is moasic in germline cells.

In conclusion, next generation sequencing is a powerful tool for prenatal genetic diagnosis, in particular, for fetuses with LHH. Apart from two known genes, *KMT2D* and *NOTCH1*, we reported a potential pathogenic gene *WDFY3*, and this is the first case of a fetal CHD with a frameshift variant in *WDFY3*. Before this genetic variant is used to genetic counseling, more sequencing data and functional studies are needed to determine the contribution of this genetic variant to CHDs.

BULLETED STATEMENTS

Genetic variants in *KMT2D* and *NOTCH1* previously have been reported in patients with CHDs, however, *WDFY3* is reported for the first time to be possibly related to CHD in human. Genetic factor is an important factor for the pathogenesis of LHH.

DATA AVAILABILITY STATEMENT

The data of KMT2D p.Gly3465Aspfs* and WDFY3 p.Ser117Xfs*1 presented in the study are deposited in the ClinVar database, and the accession numbers were SCV001449147 and SCV001449148, respectively.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Wuhan Children's Hospital (Wuhan Maternal and Child Healthcare Hospital) ethics committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

HX and ZA conceived the study. CL, WW, ZM, DZ, CJ, and ZY clinically analyzed the patients. LS and TL conducted genetic data interpretation. LS and HX performed literature review and prepared the manuscript. ZA, ZY, and HX revised the manuscript. All authors reviewed and approved the manuscript.

FUNDING

This work was supported by the grants of Wuhan Municipal Health Commission (No. WX19C19, WX14A06, and WX16C17); Youth Program of National Natural Science Foundation of China (No. 81700302); Natural Science Foundation of Hubei Province (2017CFB322); Hubei Health and Family Planning Commission (WJ2015MB247).

ACKNOWLEDGMENTS

We gratefully acknowledge our patients in this study and their families for their consent and participation in this study.

REFERENCES

1. van der Linde D, Konings EE, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJ, et al. Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis. *J Am Coll Cardiol.* (2011) 58:2241–7. doi: 10.1016/j.jacc.2011.08.025
2. Hickey EJ, Caldarone CA, McCrindle BW. Left ventricular hypoplasia: a spectrum of disease involving the left ventricular outflow tract, aortic valve, and aorta. *J Am Coll Cardiol.* (2012) 59 (1 Suppl.):S43–54. doi: 10.1016/j.jacc.2011.04.046
3. Khairy P, Fernandes SM, Mayer JE Jr, Triedman JK, Walsh EP, Lock JE, et al. Long-term survival, modes of death, and predictors of mortality in patients with Fontan surgery. *Circulation.* (2008) 117:85–92. doi: 10.1161/CIRCULATIONAHA.107.738559
4. Feinstein JA, Benson DW, Dubin AM, Cohen MS, Maxey DM, Mahle WT, et al. Hypoplastic left heart syndrome: current considerations and expectations. *J Am Coll Cardiol.* (2012) 59 (1 Suppl.):S1–42. doi: 10.1016/j.jacc.2011.09.022
5. Allen RH, Benson CB, Haug LW. Pregnancy outcome of fetuses with a diagnosis of hypoplastic left ventricle on prenatal sonography. *J Ultrasound Med.* (2005) 24:1199–203. doi: 10.7863/jum.2005.24.9.1199
6. Madriago E, Nguyen T, McFerson M, Larson EV, Airhart N, Moller JH, et al. Frequency and outcomes of cardiac operations and catheter interventions in Turner syndrome. *Am J Cardiol.* (2012) 110:580–5. doi: 10.1016/j.amjcard.2012.04.036
7. Zakaria D, Tang X, Bhakta R, ElHassan NO, Prodhan P. Chromosomal abnormalities affect the surgical outcome in infants with hypoplastic left heart syndrome: a large cohort analysis. *Pediatr Cardiol.* (2018) 39:11–8. doi: 10.1007/s00246-017-1717-3
8. Mattina T, Perrotta CS, Grossfeld P. Jacobsen syndrome. *Orphanet J Rare Dis.* (2009) 4:9. doi: 10.1186/1750-1172-4-9
9. Warburton D, Ronemus M, Kline J, Jobanputra V, Williams I, Anyane-Yeboa K, et al. The contribution of de novo and rare inherited copy number changes to congenital heart disease in an unselected sample of children with conotruncal defects or hypoplastic left heart disease. *Hum Genet.* (2014) 133:11–27. doi: 10.1007/s00439-013-1353-9
10. Iascone M, Ciccone R, Galletti L, Marchetti D, Seddio F, Lincesso AR, et al. Identification of de novo mutations and rare variants in hypoplastic left heart syndrome. *Clin Genet.* (2012) 81:542–54. doi: 10.1111/j.1399-0004.2011.01674.x
11. Elliott DA, Kirk EP, Yeoh T, Chandar S, McKenzie F, Taylor P, et al. Cardiac homeobox gene NKX2-5 mutations and congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome. *J Am Coll Cardiol.* (2003) 41:2072–6. doi: 10.1016/S0735-1097(03)00420-0
12. Dasgupta C, Martinez AM, Zuppan CW, Shah MM, Bailey LL, Fletcher WH. Identification of connexin43 (alpha1) gap junction gene mutations in patients with hypoplastic left heart syndrome by denaturing gradient gel electrophoresis (DGGE). *Mutat Res.* (2001) 479:173–86. doi: 10.1016/S0027-5107(01)00160-9
13. Reamon-Buettner SM, Ciribilli Y, Inga A, Borlak J. A loss-of-function mutation in the binding domain of HAND1 predicts hypoplasia of the human hearts. *Hum Mol Genet.* (2008) 17:1397–405. doi: 10.1093/hmg/ddn027
14. Homsy J, Zaidi S, Shen Y, Ware JS, Samocha KE, Karczewski KJ, et al. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science.* (2015) 350:1262–6. doi: 10.1126/science.aac9396
15. Tomita-Mitchell A, Stamm KD, Mahnke DK, Kim MS, Hidestrand PM, Liang HL, et al. Impact of MYH6 variants in hypoplastic left heart syndrome. *Physiol Genomics.* (2016) 48:912–21. doi: 10.1152/physiolgenomics.00091.2016
16. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus

recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* (2015) 17:405–24. doi: 10.1038/gim.2015.30
17. Nykamp K, Anderson M, Powers M, Garcia J, Herrera B, Ho YY, et al. Correction: Sherloc: a comprehensive refinement of the ACMG-AMP variant classification criteria. *Genet Med.* (2020) 22:240. doi: 10.1038/s41436-019-0624-9
18. Monaghan KG, Leach NT, Pekarek D, Prasad P, Rose NC, on behalf of the APP, Guidelines Committee. The use of fetal exome sequencing in prenatal diagnosis: a points to consider document of the American College of Medical Genetics and Genomics (ACMG). *Genet Med.* (2020) 22:675–80. doi: 10.1038/s41436-019-0731-7
19. Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med.* (2011) 13:680–5. doi: 10.1097/GIM.0b013e3182217a3a
20. Hannibal MC, Buckingham KJ, Ng SB, Ming JE, Beck AE, McMillin MJ, et al. Spectrum of MLL2 (ALR) mutations in 110 cases of Kabuki syndrome. *Am J Med Genet A.* (2011) 155a:1511–6. doi: 10.1002/ajmg.a.34074
21. Le Duc D, Giulivi C, Hiatt SM, Napoli E, Panoutsopoulos A, Harlan De Crescenzo A, et al. Pathogenic WDFY3 variants cause neurodevelopmental disorders and opposing effects on brain size. *Brain.* (2019) 142:2617–30. doi: 10.1093/brain/awz198
22. Kadir R, Harel T, Markus B, Perez Y, Bakhrat A, Cohen I, et al. ALFY-controlled DVL3 autophagy regulates Wnt signaling, determining human brain size. *PLoS Genet.* (2016) 12:e1005919. doi: 10.1371/journal.pgen.1005919
23. Grossfeld PD. Hypoplastic left heart syndrome: it is all in the genes. *J Am Coll Cardiol.* (2007) 50:1596–7. doi: 10.1016/j.jacc.2007.06.045
24. Hinton RB Jr, Martin LJ, Tabangin ME, Mazwi ML, Cripe LH, Benson DW. Hypoplastic left heart syndrome is heritable. *J Am Coll Cardiol.* (2007) 50:1590–5. doi: 10.1016/j.jacc.2007.07.021
25. McBride KL, Zender GA, Fitzgerald-Butt SM, Koehler D, Menesses-Diaz A, Fernbach S, et al. Linkage analysis of left ventricular outflow tract malformations (aortic valve stenosis, coarctation of the aorta, and hypoplastic left heart syndrome). *Eur J Hum Genet.* (2009) 17:811–9. doi: 10.1038/ejhg.2008.255
26. Bögershausen N, Wollnik B. Unmasking Kabuki syndrome. *Clin Genet.* (2013) 83:201–11. doi: 10.1111/cge.12051
27. Armstrong L, Abd El Moneim A, Aleck K, Aughton DJ, Baumann C, Braddock SR, et al. Further delineation of Kabuki syndrome in 48 well-defined new individuals. *Am J Med Genet A.* (2005) 132A:265–72. doi: 10.1002/ajmg.a.30340
28. Schwenty-Lara J, Nürnberger A, Borchers A. Loss of function of Kmt2d, a gene mutated in Kabuki syndrome, affects heart development in Xenopus laevis. *Dev Dyn.* (2019) 248:465–76. doi: 10.1002/dvdy.39
29. Ang SY, Uebersohn A, Spencer CI, Huang Y, Lee JE, Ge K, et al. KMT2D regulates specific programs in heart development via histone H3 lysine 4 di-methylation. *Development.* (2016) 143:810–21. doi: 10.1242/dev.132688
30. Sun H, Yi T, Hao X, Yan H, Wang J, Li Q, et al. Contribution of single-gene defects to congenital cardiac left-sided lesions in the prenatal setting. *Ultrasound Obstet Gynecol.* (2020) 56:225–32. doi: 10.1002/uog.21883
31. Helle E, Córdova-Palomera A, Ojala T, Saha P, Potiny P, Gustafsson S, et al. Loss of function, missense, and intronic variants in NOTCH1 confer different risks for left ventricular outflow tract obstructive heart defects in two European cohorts. *Genet Epidemiol.* (2019) 43:215–26. doi: 10.1002/gepi.22176
32. Kerstjens-Frederikse WS, van de Laar IM, Vos YJ, Verhagen JM, Berger RM, Lichtenbelt KD, et al. Cardiovascular malformations caused by NOTCH1 mutations do not keep left: data on 428 probands with left-sided CHD and their families. *Genet Med.* (2016) 18:914–23. doi: 10.1038/gim.2015.193
33. Greenway SC, Pereira AC, Lin JC, DePalma SR, Israel SJ, Mesquita SM, et al. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet.* (2009) 41:931–5. doi: 10.1038/ng.415
34. Zhang S, Song Z, An L, Liu X, Hu XW, Naz A, et al. WD40 repeat and FYVE domain containing 3 is essential for cardiac development. *Cardiovasc Res.* (2019) 115:1320–31. doi: 10.1093/cvr/cvy285
35. Isakson P, Holland P, Simonsen A. The role of ALFY in selective autophagy. *Cell Death Differ.* (2013) 20:12–20. doi: 10.1038/cdd.2012.66
36. Dragich JM, Kuwajima T, Hirose-Ikeda M, Yoon MS, Eenjes E, Bosco JR, et al. Autophagy linked FYVE (Alfy/WDFY3) is required for establishing neuronal connectivity in the mammalian brain. *Elife.* (2016) 5:e14810. doi: 10.7554/eLife.14810.026

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cardiovascular Medicine

MINI REVIEW
published: 13 April 2021
doi: 10.3389/fcvm.2021.646816

The Evolving Story in the Genetic Analysis for Heart Failure

*Kazuo Miyazawa and Kaoru Ito**

Laboratory for Cardiovascular Genomics and Informatics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan

Genomic studies of cardiovascular diseases have achieved great success, not only in Mendelian genetic diseases such as hereditary arrhythmias and cardiomyopathies, but also in common diseases such as ischemic heart disease and atrial fibrillation. However, only limited success has been achieved in heart failure due to the complexity of its disease background. In this paper, we will review the genetic research for heart failure to date and discuss how we can discover new aspects of heart failure from the viewpoint of genomic perspective.

Keywords: heart failure, cardiomyopathy, genetic architecture, genome-wide association study, big data

OPEN ACCESS

Edited by:
Christian Schulte,
University Heart and Vascular Center Hamburg (UHZ), Germany

Reviewed by:
Dipender Gill,
St George's, University of London, United Kingdom
Thorsten Kessler,
Technical University Munich, Germany
Jan Haas,
Heidelberg University Hospital, Germany

****Correspondence:***
Kaoru Ito
kaoru.ito@riken.jp

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: *28 December 2020*
Accepted: *19 March 2021*
Published: *13 April 2021*

Citation:
Miyazawa K and Ito K (2021) The Evolving Story in the Genetic Analysis for Heart Failure. Front. Cardiovasc. Med. 8:646816. doi: 10.3389/fcvm.2021.646816

INTRODUCTION

In recent years, human genomic research has made great strides, riding the wave of big data. In the field of cardiology, the analyses of hereditary arrhythmias and cardiomyopathies have been very active, and many disease-related genes and causative gene mutations have been identified to date. In particular, cardiomyopathy research could be useful in identifying common mechanisms that cause abnormalities in cardiac function, since the main locus of the abnormality is the myocardium itself (1). But whether it can help us to understand the "common" heart failure (HF) encountered in our daily clinical practice is controversial. On the other hand, with the advent of genome-wide association studies (GWAS), >200 disease-susceptibility loci have been identified thus far in common cardiovascular diseases such as atrial fibrillation (AF) (2) and ischemic heart disease (IHD) (3), which are caused by the accumulation of weak-effect genetic polymorphisms that differ from monogenic diseases such as hereditary cardiomyopathies (Mendelian diseases). Although these advances in technology and methods are gradually revealing the genetic background of many cardiovascular diseases, not many studies have been able to approach the genetic background of "common" HF. This article will briefly review the genetic studies for HF and then suggest how those genetic studies should be viewed in light of the commonality of genetic background between cardiomyopathy and HF.

RECENT ADVANCES IN GENOMIC RESEARCH AND CARDIOVASCULAR DISEASE

Since the Human Genome Project was completed in 2003 and the outline of the human genome sequence was determined, the researches in this field have developed at an extraordinary pace. In addition to the completion of the reference map of the human genome, the involvement of technologies such as genotyping arrays and next-generation sequencers that can determine genetic variants inexpensively and accurately has made this possible. At the same time, data processing environments, such as computer and data storage, have been improved in the last few years to cope with the growing volume of genomic data.

In recent years, large databases of genetic variants including very rare ones have been developed, such as gnomAD (https://gnomad.broadinstitute.org/) (4). Also databases including the information of neighboring omics layers, such as epigenome and transcriptome data that are important to understand genomic information, are now available; ENCODE (https://www.encodeproject.org/), Roadmap Epigenetics (http://www.roadmapepigenomics.org/), GTEx (https://www.gtexportal.org/home/), and other international projects have enriched the data for each organ and cell. Such multi-omics data allow us to infer the function of identified genetic variants without additional wet experiments.

Cardiovascular genomics has also made great strides along these lines. The implementation of targeted resequencing, as well as whole-exome sequencing s and whole-genome sequencing, has dramatically increased the detection rate of mutations that cause hereditary arrhythmias and cardiomyopathies. However, new problems, such as "Variant of Unknown Significance," have arisen due to the availability of a large amount of genomic data. This problem also seems to be gradually being solved with the advent of organ-specific RNA sequencing (5) and superior *in-silico* prediction algorithms (6, 7). On the other hand, "common" cardiovascular diseases such as IHD and AF are being addressed by large international consortia and mega-biobanks such as the UK Biobank, which have made samples of hundreds of thousands of people available for analysis. The resulting dramatic increase in the statistical power of studies has enabled the identification of many disease-susceptible gene regions (2, 3) and accelerated our understanding of disease mechanisms. And now, there is a growing momentum for clinical applications of genomic information, such as studies using polygenic risk scores, which are genetic risk scores that combine the information of a large number of genetic variants in order to aim achieving precision medicine (8).

GENOMIC RESEARCH FOR HF UNDER DIFFICULT CONDITIONS

While human genomic research is making remarkable progress, genomic research on so-called "HF" seems to have had only limited success. The reason for this is that HF is not a single disease concept but a syndrome that can be caused by a variety of factors, including IHD, valvular heart disease, infectious diseases, congenital heart diseases, and hereditary cardiomyopathies, all of which share a common endpoint. In other words, although the final pathology may be the same, the processes leading to it are different, which makes them incompatible with genomic research that defines the starting point of the disease. Under these circumstances, several HF studies have been conducted, but the full picture of the genetic makeup of HF remains unclear.

Since HF is the end result of many different diseases, it seems like a good idea to consider each disease separately and conduct each genomic research. The less variability there is in the disease, the more statistically advantageous it is and the easier it is to reach a conclusion. However, there is an attempt to identify a common pathway for HF arriving from such a variety of causes. This type of HF is called "all-cause HF (HF)," meaning that no specific cause is assumed. There is also another definition "non-ischemic HF" that excludes only HF caused by IHD. Although the analysis of these disease concepts is inefficient because of the variability in causes, it is expected to discover common pathways for HF independent of the cause. Such studies may lead to the discovery of mechanisms to distinguish between cases that progress to HF and those that do not, in the presence of various types of cardiac stressors.

WHAT ALL-CAUSE HF OR NON-ISCHEMIC HF GENOMIC STUDIES SUGGEST

Let's take a look at the status of genomic studies on all-cause HF and non-ischemic HF: among the major GWAS studies (9, 10) conducted until the end of 2018, only three disease susceptibility regions have achieved genome-wide significance levels ($P < 5 \times 10^{-8}$; **Table 1**). This is a very small number compared to other "common" cardiovascular diseases such as IHD and AF. However, a careful reading of these can shed light on the nature of HF. The first one, reported by Smith et al. (9), is the disease susceptibility locus (lead variant: rs10519210) located in the intergenic region between *CA12* and *USP3*. Such disease-susceptibility regions identified by GWAS are often not in the protein-coding regions of genes, but in enhancer regions that regulate a gene expression (which may be remote from the target gene). Since an enhancer and the target gene are in close proximity in three dimensions, Hi-C database (11) can indicate the physical proximity to the enhancer region. According to the Hi-C data, the enhancer may act on both *RAB8B* and *FBXL22* genes. Based on the previously reported GWAS (https://www.ebi.ac.uk/gwas/) and the UK Biobank GWAS catalog (https://www.nealelab.is/uk-biobank), *RAB8B* affects AF, systolic blood pressure (SBP), heart rate (HR), and body mass index (BMI), while *FBXL22* affects AF, HR, and BMI. Next, HF susceptibility locus (lead variant: rs1906609) reported by Li et al. was located upstream of the *PITX2* gene, one of the most famous AF-related genes (12), suggesting a genetic relationship between HF and AF. In other words, these two regions reconfirmed clinically known HF risk factors such as AF, SBP, HR, and obesity.

Another non-ischemic HF-related locus (lead variant: rs2234962) reported by Aragam et al. (10) was located on the *BAG3* gene, which is a well-known causative gene for idiopathic dilated cardiomyopathy (DCM). *BAG3* plays a variety of roles in maintaining cardiac function, including linking adrenergic receptors to L-type Ca^{2+} channels, providing structural support for sarcomeres, and inhibiting apoptosis (13). In addition, the *BAG3* gene has also been reported to be associated with other diseases such as Takotsubo cardiomyopathy, HIV-associated cardiomyopathy, and virus-associated myocarditis, in which the myocardium develops cardiomyopathy or myocarditis when exposed to stresses or infection. This suggests that HF and these diseases may share some common genetic background, which means that they may share common disease mechanisms by having the same disease-related genes, and that HF genomics can help us to understand these diseases and vice versa.

TABLE 1 | Genetic loci identified by GWAS for all-cause or non-ischemic HF.

Study	Year	The number of samples	Chromosome	Lead variant	Nearest gene(s)	*P*-value	Suggested Associated Traits
(9)	2010	Case 2,526 Control 20,926	15	rs10519210	*USP3*	1.4×10^{-8}	AF, SBP, HR, BMI
(10)	2018	Case 7,382	4	rs1906609	*PITX2*	9.1×10^{-10}	AF
		Control 480,628	10	rs121429633	*BAG3*	2.3×10^{-9}	DCM
(14)	2020	Case 47,309	1	rs660240	*CELSR2*	3.3×10^{-10}	LDL-C, IHD
		Control 930,014	4	rs17042102	*PITX2/FAM241A*	5.7×10^{-20}	AF
			5	rs11745324	*KLHL3*	2.4×10^{-8}	AF
			6	rs4135240	*CDKN1A*	6.8×10^{-9}	Reduced LV function
			6	rs55730499	*LPA*	1.8×10^{-11}	LDL-C, IHD
			6	rs140570886	*LPA*	7.7×10^{-11}	LDL-C, IHD
			9	rs1556516	9p21/*CDKN2B-AS1*	1.6×10^{-15}	IHD
			9	rs600038	*ABO/SURF1*	3.7×10^{-9}	LDL-C, IHD, T2D
			10	rs4746140	*SYNPO2L/AGAP5*	1.1×10^{-9}	AF
			10	rs17617337	*BAG3*	3.7×10^{-9}	DCM, reduced LV function
			12	rs4766578	*ATXN2*	4.9×10^{-8}	SBP, DBP
			16	rs56094641	*FTO*	1.2×10^{-8}	BMI, T2D

AF, atrial fibrillation; SBP, systolic Blood Pressure; HR, heart rate; DCM, dilated cardiomyopathy; LDL-C, low density lipoprotein cholesterol; IHD, ischemic heart disease; LV, left ventricular; T2D, type 2 diabetes mellitus; DBP, diastolic blood pressure; BMI, body mass index.

FINDINGS FROM THE LARGEST HF GWAS IN THE WORLD

Implications from genomic studies of HF have been useful but, as already mentioned, extremely limited. To overcome this situation, the largest HF GWAS study as of 2020 has been conducted, which is a meta-analysis of global HF studies. In the study presented by Shah et al. (14), 47,309 HF patients and 930,014 control samples were analyzed. As a result, they identified 11 novel disease-susceptibility loci (**Table 1**). These suggested the associations of HF with IHD and AF from the viewpoint of genetics. They also reported loci involved in myocardial development (*MYOZ1*, *SYNPO2L*) and cellular senescence (*CDKN1*). In particular, Frey et al. (15) have already suggested in a mouse model that *MYOZ1* regulates exercise performance *in vivo* via modulation of calcineurin/NFAT activity and concomitant changes in myofiber type composition. They also utilized the Mendelian randomization method to show that other factors besides IHD, such as AF, hypertension, BMI, hypertension, and triglycerides, are important in the development of severe HF. Even with this unprecedentedly large sample size, the number of new disease susceptibility loci identified was not so many, and new findings were limited, again highlighting the difficulties of genomic research for HF.

MEASUREMENTS OF CARDIAC FUNCTION, CARDIAC MORPHOLOGY, BIOMARKER FOR HF, AND GENETIC FACTORS

The genetic analysis of all-cause HF was very challenging. On the other hand, genetic analysis using individual cardiac parameters as an outcome variable not only gives more statistical power (i.e., dichotomous vs. continuous variables), but also provides various suggestions regarding cardiac function. Newly, Aung et al. (16) in 2019, Pirruccello et al. (17) in 2020 performed GWAS using parameters obtained from cardiac MRI, which are closely related to cardiac function. They identified 57 new susceptibility loci, 45 of which were novel and had not been reported in previous genomic analyses. In addition, familial cardiomyopathy genes such as *BAG3*, *FLNC*, *TTN*, *GATA4*, *MYH6*, *MYH7*, *NKX2-5*, *PLN*, *RBM20*, and *RYR2* were found to be significantly enriched in the susceptibility loci, suggesting an association between genetic factors that define cardiac function in healthy individuals and cardiomyopathy-related genes. They also showed that polygenic risk scores (PRS) derived from the cardiac MRI GWAS were significantly correlated with HF and DCM (it should be noted that the PRS derived from the GWAS of left ventricular diameter in systole (LVDs) was the most correlated). They also showed that these genetic loci also influence cardiac function in patients with cardiomyopathy with low penetrance. Collectively, they demonstrated the association of myocardial measurements with cardiomyopathy genes.

Another structure specific to myocardium that can be observed on MRI is the cardiac trabeculae. This is a complex network of muscular strands thought to be a remnant of embryonic development. Meyer et al. (18) quantified the complexity of myocardial trabeculae using a deep learning algorithm from cardiac MRI images and performed a GWAS. They not only identified loci related to hemodynamic phenotypes and regulation of cytoskeletal arborization, but also showed that trabecular morphology was an important determinant of cardiac performance. Furthermore, they performed a Mendelian randomization study and showed a causal relationship between trabecular morphology and cardiovascular diseases. Their

findings were very interesting in that they clarified the significance of myocardial trabeculae on cardiac function and cardiovascular diseases, since the significance of which had not been well-understood before.

In addition to measurements from cardiac imaging, the circulating level of cardiac peptides such as BNP and NT-proBNP, is a useful biomarker used in clinical practice to reflect the state of HF. BNP is also known to have a cardioprotective effect, and its levels increase in worsening HF. Musani et al. (19) identified two loci associated with BNP levels. One of them was the region containing the *NPPB* gene encoding BNP (lead variant: rs198389) as expected, while the other was a missense variant located in *KLKB1* locus (rs3733402). The *NPPB* locus was related to SBP, and the *KLKB1* locus to the aldosterone/renin ratio, both of which were associated with left ventricular mass, suggesting the existence of a shared genetic architecture between these traits. This genomic analysis using cardiac measurements and biomarkers, which does not rely on the conventional case-control analysis, is very meaningful because it shows a new path of genomic analysis for HF and suggests an clear association with hereditary cardiomyopathy.

GENETIC ANALYSIS FOR DISEASES LEADING TO HF

Other conditions that can cause HF include IHD, AF, and valvular heart disease. IHD and AF have been reported to have more than 100 disease susceptibility loci as already described (2, 3), where IHD has been suggested to be strongly associated with lipids, hypertension, diabetes, inflammation, and other atherosclerosis-related traits, while AF has been suggested to have mechanisms strongly related to HF, such as cardiac development and myocardial contractility, apart from electrophysiological traits. As for valvular disease, Chen et al. (20) have recently identified nine disease-susceptibility loci in aortic stenosis, suggesting the involvement of lipids and the immune system. In addition, there are a few reports on mitral valve prolapse syndrome (21, 22) in which mitral valve repair and cardiac development are reported to be involved in the pathogenesis. These reports suggest that HF-related diseases have both their own unique genetic factors and the shared genetic background, which must be useful in understanding the genetic basis of HF.

STRUCTURAL VARIATIONS, NON-PROTEIN-CODING REGIONS, EPIGENOME, AND HF

Although the studies described so far are mainly based on DNA polymorphisms, we should also consider the influence of alterations in non-protein-coding regions through DNA structural variations, and the influence of the epigenomic changes on cardiac function. Haas et al. (23) conducted a genome-wide evaluation of the association between DNA structural variations (SVs) and gene expression in the human heart and detected SVs affecting non-coding RNAs, protein-coding transcripts and regulatory genomic regions (e.g., enhancers, transcription factor binding sites), demonstrating the importance of SVs on gene transcription in HF. Regarding the effect of epigenetic alterations on HF, Meder et al. (24) used arrays to generate genome-wide DNA methylation profiles of patients with DCM and compared them to controls, and identified 59 differentially methylated CpG loci. Of these, about half each were hypomethylated or hypermethylated, indicating a variety of directions of DNA methylation associated with HF. In addition, the authors identified 517 epigenetic loci associated with DCM and cardiac gene expression. This study also reported overlapping methylation patterns between myocardium and peripheral blood, including demethylated *NPPA* and *NPPB* loci, which encodes an important cardiac stress marker ANP and BNP, respectively. Thus, epigenomic information seems to be more dynamic than expected and may be responsible for various pathological conditions in HF.

TTN GENE AND HF

The *TTN* gene is one of genes that encode proteins that make up the sarcomere, the contractile unit of the heart muscle. The sarcomere genes such as *MYH7* and *MYBPC3* are known to be causative for hereditary cardiomyopathies. Among them, the *TTN* gene is so large that it was not comprehensively studied in the days when only Sanger sequencing was available. When high-throughput sequencers started to become available, targeted resequencing of the *TTN* gene was performed. Then, it was reported that about 25% of DCM was caused by *TTN* protein truncating variants (25). Subsequently, it was reported that about 15% of DCM was caused by *TTN* gene abnormalities (26). Although the percentage decreased in the later report, it remained the most common of all responsible genes for DCM in both adults and children. However, there was no report of DCM caused by *TTN* mutation in infants, suggesting that *TTN* truncating variants themselves do not immediately cause DCM, but rather have characteristics that develop as a result of a certain amount of stress or after a certain amount of time has passed. It has also been reported that *TTN* truncating variants are relatively common even in healthy individuals, where as much as 1% of them have *TTN* mutations (27). Whether or not *TTN* truncating variants cause cardiomyopathy has been proven, to some extent, to depend on the region of the TTN protein where the mutation is located. The A- or I-band distal ends of titin proteins, which have high exon utilization, have high odds ratios for the disease development due to protein truncation mutations. These effects of TTNtv are thought to be due to "poison peptide" or haploinsufficiency. In addition, the accumulation of the effects of common variants affects the phenotype of TTNtv (17).

According to a report by Schafer et al. (28) in 2017, a comparison of cardiac MRI images of healthy subjects with and without TTNtv revealed that those with TTNtv had a small but significant change in left ventricular morphology. They then showed in rats that such TTNtv can cause cardiac dysfunction under stress. The results also suggest that TTNtv may cause haploinsufficiency due to nonsense-medicated decay regardless of its position (even on the I-band). Collectively,

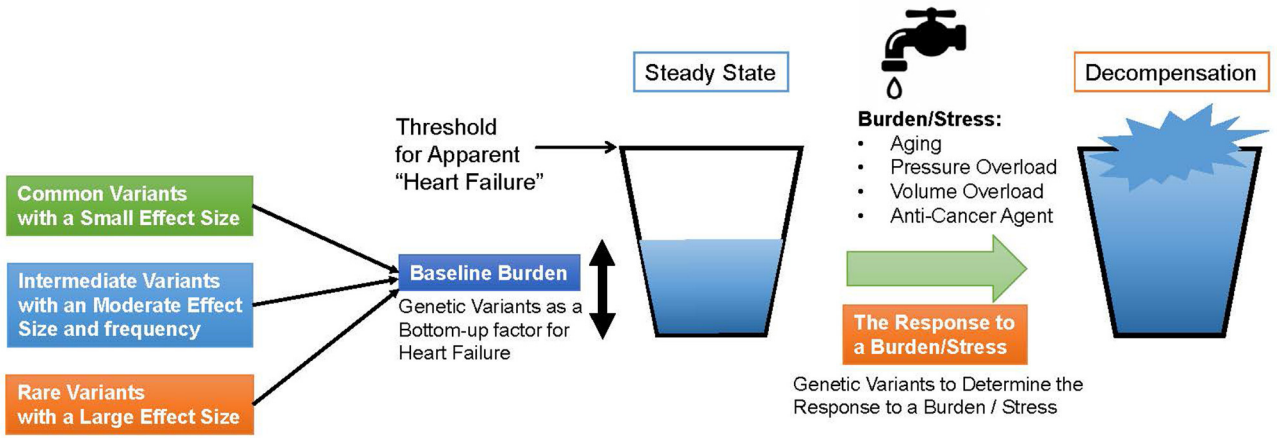


FIGURE 1 | A model for heart failure development. The horizontal line denotes a threshold for apparent heart failure. First, genetic Variants determine the baseline burden of the heart and serve a bottom-up factor for heart failure. If a patient has a cardiomyopathy-causing variant, the baseline burden itself may exceed the threshold. When the heart is subjected to a burden or a stress such as aging, pressure or volume overload and anti-cancer agent, the amount of the burden grows up and exceeds the threshold. In the process, genetic variants determine the response to the burden or the stress.

they demonstrated that TTNtv, which are normally considered to be unproblematic, can define the stress response and lead to HF under stress. Ware et al. (29) analyzed perinatal cardiomyopathy, a condition that causes HF in pregnancy, and reported that two-thirds of the causative gene mutations were *TTN* truncating variants. Another report showed that *TTN* truncating mutations were identified in patients with adriamycin cardiomyopathy who developed HF due to the administration of the anticancer drug (30). In perinatal cardiomyopathy and adriamycin cardiomyopathy, patients who originally had no abnormal cardiac function develop HF upon stress loading such as pregnancy or administration of anticancer drugs. In other words, this is consistent with the characteristic of *TTN* mutations, which does not cause any particular problem under steady state conditions, but develops HF when stressed. Thus, the *TTN* mutation may determine the stress response of the myocardium.

Additionally, *TTN* is a well-known splicing target of the protein encoded by *RBM20*, which contains zinc finger domains, an RNA recognition motif (RRM), a serine- and arginine-rich region (RS region), a leucine-rich region, and a glutamate-rich region. These regions, especially the RS region, are highly conserved among species, and mutations in these regions can lead to loss of RBM20 function and the development of DCM (31). It has also been reported that other splicing targets of RBM20 including important DCM-related genes, *CACNA1C* and *CAMK2D*, may be involved in the pathogenesis of cardiomyopathy caused by *RBM20* mutations (32). Specifically, during transcription, TTN pre-mRNA and RBM20 form clusters that spatially attract other RBM20 targets on different chromosomes such as *CACNA1C* and *CAMK2D* loci. Furthermore, this colocalization leads to increased splicing activity of RBM20. Conversely, the loss of TTN pre-mRNA causes splicing changes in *CACNA1C* and *CAMK2D* similar to the loss of RBM20 itself. This mechanism promotes alternative splicing of RBM20-dependent transcripts, suggesting the existence of a cardiac-specific trans-interacting chromatin domain that functions as a "splicing factory."

DISCUSSION

Genomic studies of HF not only traced known clinical risk factors, but also revealed a common genetic basis with various cardiomyopathies and myocarditis, because of the shared causative genes such as *BAG3* and *TTN*. These similarities lead to the following findings. These commonalities suggest that (i) HF has a common pathway, (ii) some cardiomyopathies, which share some genetic basis with HF, develop under load and may define baseline load and stress response of the heart, and (iii) it may define the baseline morphology and function of the myocardium. From these properties, we should further try to prove the hypothesis that genetic factors, identified in the genomic studies of all-cause HF and non-ischemic HF, play a shared role in the development of various types of HF, as shown in **Figure 1**. In this model, under steady-state conditions, HF-associated genetic variants act as "bottom-up" factors in HF and define the baseline burden. In the steady state, the threshold for HF is not exceeded, and HF symptoms do not develop. HF-associated genetic variants include common variants [minor allele frequency (MAF) >5%], which have been identified through GWAS and generally have a small effect size but have a high frequency of being associated with disease risk; intermediate effect variants (MAF 1–5%), which have a moderate effect size and frequency between common variants and rare disease variants; rare (MAF <1%) and highly effective variants that cause Mendelian disease. As mentioned earlier (17), factors such as common variants and intermediate effect variants have also been

shown to affect the penetrance of individuals with Mendelian disorders by pushing the genetic burden toward the threshold of the disease. When the myocardium is further stressed by aging, pregnancy, pressure overload, volume overload, anticancer drugs, etc., the threshold is exceeded and HF develops. Here, HF-related gene mutations define the stress response. In other words, through genomic studies of HF we can learn about the genetic underpinnings that define the basic characteristics of cardiac function and determine susceptibility to the development of HF.

Thus, genomic studies can provide interesting insights into our understanding of human HF. However, even with the world's largest HF GWAS study, the number of novel disease susceptibility loci was not so many, and the mechanisms of disease susceptibility identified in this study were limited to tracing previous discoveries. Therefore, it is important to promote HF genomic research in a more multifaceted manner by integrating other omics data and using methods such as machine learning in addition to conventional frequency-based statistical methods.

AUTHOR CONTRIBUTIONS

KM and KI wrote the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

KM and KI is funded by Japan Agency for Medical Research and Development, AMED, under Grant Numbers JP16ek0109070h0003, JP18kk0205008h0003, JP18kk0205001s0703, JP20km0405209, and JP20ek0109487.

REFERENCES

1. Garfinkel AC, Seidman JG, Seidman CE. Genetic pathogenesis of hypertrophic and dilated cardiomyopathy. *Heart Fail Clin.* (2018) 14:139–46. doi: 10.1016/j.hfc.2017.12.004
2. Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM, et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet.* (2018) 50:1225–33. doi: 10.1038/s41588-018-0133-9
3. Koyama S, Ito K, Terao C, Akiyama M, Horikoshi M, Momozawa Y, et al. Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease. *Nat Genet.* (2020) 52:1169–77. doi: 10.1038/s41588-020-0705-3
4. Koch L. Exploring human genomic diversity with gnomAD. *Nat Rev Genet.* (2020) 21:448. doi: 10.1038/s41576-020-0255-7
5. Marco-Puche G, Lois S, Benítez J, Trivino JC. RNA-Seq perspectives to improve clinical diagnosis. *Front Genet.* (2019) 10:1152. doi: 10.3389/fgene.2019.01152
6. Kim SS, Dey KK, Weissbrod O, Márquez-Luna C, Gazal S, Price AL. Improving the informativeness of Mendelian disease-derived pathogenicity scores for common disease. *Nat Commun.* (2020) 11:6258. doi: 10.1038/s41467-020-20087-2
7. Grujic O, Phung TN, Kwon SB, Arneson A, Lee Y, Lohmueller KE, et al. Identification and characterization of constrained non-exonic bases lacking predictive epigenomic and transcription factor binding annotations. *Nat Commun.* (2020) 11:6168. doi: 10.1038/s41467-020-19962-9
8. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet.* (2018) 50:1219–24. doi: 10.1038/s41588-018-0183-z
9. Smith NL, Felix JF, Morrison AC, Demissie S, Glazer NL, Loehr LR, et al. Association of genome-wide variation with the risk of incident heart failure in adults of European and African ancestry: a prospective meta-analysis from the cohorts for heart and aging Research in Genomic Epidemiology (CHARGE) Consortium. *Circ Cardiovasc Genet.* (2010) 3:256–66. doi: 10.1161/CIRCGENETICS.109.895763
10. Aragam KG, Chaffin M, Levinson RT, McDermott G, Choi SH, Shoemaker MB, et al. Phenotypic refinement of heart failure in a national biobank facilitates genetic discovery. *Circulation.* (2019) 139:489–501. doi: 10.1161/CIRCULATIONAHA.118.035774
11. Javierre BM, Burren OS, Wilder SP, Kreuzhuber R, Hill SM, Sewitz S, et al. Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell.* (2016) 167:1369–84. doi: 10.1016/j.cell.2016.09.037
12. Li N, Dobrev D, Wehrens XH. PITX2: a master regulator of cardiac channelopathy in atrial fibrillation? *Cardiovasc Res.* (2016) 109:345–7. doi: 10.1093/cvr/cvw008
13. Myers VD, McClung JM, Wang J, Tahrir FG, Gupta MK, Gordon J, et al. The multifunctional protein BAG3: a novel therapeutic target in cardiovascular disease. *JACC Basic Transl Sci.* (2018) 3:122–31. doi: 10.1016/j.jacbts.2017.09.009
14. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. *Nat Commun.* (2020) 11:163. doi: 10.1038/s41467-019-13690-5
15. Frey N, Frank D, Lippl S, Kuhn C, Kögler H, Barrientos T, et al. Calsarcin-2 deficiency increases exercise capacity in mice through calcineurin/NFAT activation. *J Clin Invest.* (2008) 118:3598–608. doi: 10.1172/JCI36277
16. Aung N, Vargas JD, Yang C, Cabrera CP, Warren HR, Fung K, et al. Genome-wide analysis of left ventricular image-derived phenotypes identifies fourteen loci associated with cardiac morphogenesis and heart failure development circulation. (2019) 140:1318–30. doi: 10.1161/CIRCULATIONAHA.119.041161
17. Pirruccello JP, Bick A, Wang M, Chaffin M, Friedman S, Yao J, et al. Analysis of cardiac magnetic resonance imaging in 36,000 individuals yields genetic insights into dilated cardiomyopathy. *Nat Commun.* (2020) 11:2254. doi: 10.1038/s41467-020-15823-7
18. Meyer HV, Dawes TJW, Serrani M, Bai W, Tokarczuk P, Cai J, et al. Genetic and functional insights into the fractal structure of the heart. *Nature.* (2020) 584:589–94. doi: 10.1038/s41586-020-2635-8
19. Musani SK, Fox ER, Kraja A, Bidulescu A, Lieb W, Lin H, et al. Genome-wide association analysis of plasma B-type natriuretic peptide in blacks: the Jackson Heart Study. *Circ Cardiovasc Genet.* (2015) 8:122–30. doi: 10.1161/CIRCGENETICS.114.000900
20. Chen HY, Cairns BJ, Small AM, Burr HA, Ambikkumar A, Martinsson A, et al. Association of FADS1/2 locus variants and polyunsaturated fatty acids with aortic stenosis. *JAMA Cardiol.* (2021) 5:694–702. doi: 10.1001/jamacardio.2020.0246
21. Dina C, Bouatia-Naji N, Tucker N, Delling FN, Toomer K, Durst R, et al. Genetic association analyses highlight biological pathways underlying mitral valve prolapse. *Nat Genet.* (2015) 47:1206–11. doi: 10.1038/ng.3383
22. Yu M, Georges A, Tucker NR, Kyryachenko S, Toomer K, Schott JJ, et al. Genome-wide association study-driven gene-set analyses, genetic, and functional follow-up suggest GLIS1 as a susceptibility gene for mitral valve prolapse. *Circ Genom Precis Med.* (2019) 12:e002497. doi: 10.1161/CIRCGEN.119.002497
23. Haas J, Mester S, Lai A, Frese KS, Sedaghat-Hamedani F, Kayvanpour E, et al. Genomic structural variations lead to dysregulation of important coding and

non-coding RNA species in dilated cardiomyopathy. *EMBO Mol Med.* (2018) 10:107–20. doi: 10.15252/emmm.201707838
24. Meder B, Haas J, Sedaghat-Hamedani F, Kayvanpour E, Frese K, Lai A, et al. Epigenome-wide association study identifies cardiac gene patterning and a novel class of biomarkers for heart failure. *Circulation*. (2017) 17:136:1528–44. doi: 10.1161/CIRCULATIONAHA.117.027355
25. Herman DS, Lam L, Taylor MR, Wang L, Teekakirikul P, Christodoulou D, et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med.* (2012) 366:619–28. doi: 10.1056/NEJMoa1110186
26. Pugh TJ, Kelly MA, Gowrisankar S, Hynes E, Seidman MA, Baxter SM, et al. The landscape of genetic variation in dilated cardiomyopathy as surveyed by clinical DNA sequencing. *Genet Med.* (2014) 16:601–8. doi: 10.1038/gim.2013.204
27. Akinrinade O, Koskenvuo JW, Alastalo TP. Prevalence of titin truncating variants in general population. *PLoS One*. (2015) 10:e0145284. doi: 10.1371/journal.pone.0145284
28. Schafer S, de Marvao A, Adami E, Fiedler LR, Ng B, Khin E, et al. Titin-truncating variants affect heart function in disease cohorts and the general population. *Nat Genet.* (2017) 49:46–53. doi: 10.1038/ng.3719
29. Ware JS, Li J, Mazaika E, Yasso CM, DeSouza T, Cappola TP, et al. Shared genetic predisposition in peripartum and dilated cardiomyopathies. *N Engl J Med.* (2016) 374:233–41. doi: 10.1056/NEJMoa1505517
30. Garcia-Pavia P, Kim Y, Restrepo-Cordoba MA, Lunde IG, Wakimoto H, Smith AM, et al. Genetic variants associated with cancer therapy-induced cardiomyopathy. *Circulation.* (2019) 140:31–41. doi: 10.1161/CIRCULATIONAHA.118.037934
31. Lennermann D, Backs J, van den Hoogenhof MMG. New Insights in RBM20 cardiomyopathy. *Curr Heart Fail Rep*. (2020) 17:234–46. doi: 10.1007/s11897-020-00475-x
32. Bertero A, Fields PA, Ramani V, Bonora G, Yardimci GG, Reinecke H, et al. Dynamics of genome reorganization during human cardiogenesis reveal an RBM20-dependent splicing factory. *Nat Commun*. (2019) 10:1538. doi: 10.1038/s41467-019-09483-5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 13 April 2021
doi: 10.3389/fcvm.2021.625852

Prevalence and Impact of Apolipoprotein E7 on LDL Cholesterol Among Patients With Familial Hypercholesterolemia

Hayato Tada, Kan Yamagami, Nobuko Kojima, Junichi Shibayama, Tetsuo Nishikawa, Hirofumi Okada, Akihiro Nomura, Soichiro Usui, Kenji Sakata, Masayuki Takamura and Masa-aki Kawashiri*

Department of Cardiovascular Medicine, Kanazawa University Graduate School of Medical Sciences, Kanazawa, Japan

Background: It has been suggested that a rare mutant apolipoprotein E7, APOE7 (p.Glu262Lys, p.Glu263Lys), has been identified to be associated with hyperlipoproteinemia in the general population. Moreover, its prevalence has been shown to be 0.005–0.06%. However, there are no prior data regarding its prevalence and impact on serum lipids in patients with familial hypercholesterolemia (FH).

Methods: We recruited 1,138 patients with clinically diagnosed FH [mean age = 48, men = 512, median low-density lipoprotein (LDL) cholesterol = 231 mg/dl]. The coding regions of three FH genes (*LDLR*, *APOB*, and *PCSK9*) and apolipoprotein E (*APOE*) gene were sequenced. We investigated the prevalence and impact of APOE7 mutant on serum lipid levels in patients with FH.

Results: We identified 29 patients (2.5 %) with a mutant APOE7 (heterozygote), which is apparently much higher than that of the general population. Moreover, when we focus on those without FH mutation ($n = 540$), we identified 21 patients (3.9 %) with a mutant APOE7. Patients with a mutant APOE7 exhibited significantly higher median LDL cholesterol and triglyceride levels compared with those without this rare mutant (249 vs. 218 mg/dl, $p < 0.05$, 216 vs. 164 mg/dl, $p < 0.05$, respectively). Moreover, LDL cholesterol levels in the APOE7-oligogenic FH individuals, with a pathogenic mutation in FH genes and APOE7 mutant, were significantly higher than that in monogenic FH patients (265 vs. 245 mg/dl, $p < 0.05$).

Conclusion: We identified more patients with a mutant APOE7 than expected among those diagnosed with FH clinically, especially among those without FH-causing mutation. This implies a mutant APOE7 may be one of the causes FH, especially among those without FH mutations.

Keywords: APOE, familial hypercholesterolemia, PCSK9, LDLR, LDL-C

OPEN ACCESS

Edited by:
Seitaro Nomura,
The University of Tokyo, Japan

Reviewed by:
Ville-Petteri Makinen,
South Australian Health and Medical Research Institute (SAHMRI), Australia
Judith Sluimer,
Maastricht University, Netherlands

****Correspondence:***
Hayato Tada
ht240z@sa3.so-net.ne.jp

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: *04 November 2020*
Accepted: *11 March 2021*
Published: *13 April 2021*

Citation:
Tada H, Yamagami K, Kojima N, Shibayama J, Nishikawa T, Okada H, Nomura A, Usui S, Sakata K, Takamura M and Kawashiri M-a (2021) Prevalence and Impact of Apolipoprotein E7 on LDL Cholesterol Among Patients With Familial Hypercholesterolemia. Front. Cardiovasc. Med. 8:625852. doi: 10.3389/fcvm.2021.625852

INTRODUCTION

Patients with familial hypercholesterolemia (FH), caused by genetic mutations in low-density lipoprotein (LDL)-associated genes (FH genes), are typically exhibiting extreme hyper-LDL cholesterolemia, tendinous xanthomas, and premature atherosclerotic cardiovascular disease (ASCVD) (1, 2). Even though it is a genetic disorder, at least a part of clinically diagnosed FH

patients do not have any causative mutations in FH genes even with the comprehensive genetic analyses (3). It is possible that novel genes are contributing to this disease, or accumulation of LDL-associated common genetic variations may have an impact on this disease. In contrast, apolipoprotein E (APOE) has been shown to interact with LDL receptor (LDLR) (4). Moreover, the *APOE* is highly polymorphic: APOE-ε2 (cys112/cys158), APOE-ε3 (cys112/arg158), and APOE-ε4 (arg112/arg158) (5, 6). These common polymorphisms have been associated with LDL cholesterol levels (7). In addition to these common polymorphisms, a rare mutant of APOE7 (p.Glu262Lys, p.Glu263Lys) has been identified to be associated with dyslipidemia in the general population. It has been shown that additional lysine residues of apoE7 were associated with reduced binding affinity to LDLR and increased affinity to heparin; both of those features appear to lead to their atherogenic lipoprotein profile (8). Moreover, its prevalence has been shown to be 0.005–0.06% (9–12). According to gnomAD browser, the allele frequency of the e7 variant is 0.002698 in East Asians, and it is not found among other ethnicities (13). Another study showed that the LDLR binding activity of APOE7 has been reduced to 23% *in vitro* (14). Moreover, interesting cases with a mutant APOE7 exhibit phenocopy of FH, including extremely high LDL cholesterol level and Achilles tendon thickness (15).

Recently, we showed that loss-of function mutations in ATP-binding cassette subfamily G member 5 (*ABCG5*) or ATP-binding cassette subfamily G member 8 (*ABCG8*) considered as causes of recessive disorder contribute to mimicking or exacerbating the phenotype of FH (16, 17). Considering these backgrounds, we tried to investigate whether a rare mutant of APOE7 is also contributing to mimic and/or exacerbate of the phenotype of FH.

METHODS

Study Population

In this study, we included 1,262 clinically diagnosed FH patients (18) at our institution from 2014 to 2020. We used the clinical diagnostic criteria of FH determined by Japan Atherosclerosis Society. A total of 121 subjects were removed due to missing data. Moreover, three individuals with double mutations were excluded. Accordingly, 1,138 subjects [male, 45%; female, 55%; mean age = 48 years; ASCVD = 211 (19%)] were finally included in this study.

Genetic Analysis

The coding region of 21 genes (*ABCA1*, *ABCG5*, *ABCG8*, *ANGPTL3*, *APOA1*, *APOB*, *APOC2*, *APOC3*, *APOA5*, *APOE*, *CETP*, *GPIHBP1*, *LCAT*, *LDLR*, *LDLRAP1*, *LIPG*, *LMF1*, *LPL*, *MTTP*, *PCSK9*, and *SAR1B*) associated with Mendelian lipid diseases, including three FH genes (*LDLR*, *APOB*, and *PCSK9*) and *APOE* gene were sequenced. In this study, we defined pathogenic mutations as those raising LDL cholesterol. The pathogenicity of the variants was assessed through the allele frequency in the Asian population of The Exome Aggregation Consortium (19), computational pathogenicity prediction tools, and ClinVar. The variants with allele frequency <5% are considered as rare. In addition, we determined variants as pathogenic using the standard American5College of Medical Genetics (ACMG) criteria. The subjects were divided based on the presence of FH mutation and APOE7 mutant (heterozygote; **Figure 1**).

Ethical Considerations

The Institutional Review Board of Kanazawa University approved this study. All procedures were in accordance with the ethical standards of the committee on human experimentation and Declaration of Helsinki of 1975, revised in 2008. Informed consent was given to us from all of the study participants in this study.

Biochemical Test

All blood tests were performed at fasting state. The serum lipids, including total cholesterol, triglycerides, and high-density lipoprotein cholesterol, were assessed enzymatically (20). LDL cholesterol levels were calculated using the Friedewald formula if the triglyceride levels were <400 mg/dl; otherwise, they were determined directly. All of the data presented in this study were assessed at the point without any lipid-lowering therapies.

Clinical Assessments

We defined hypertension as systolic blood pressure 140 mmHg or greater, diastolic blood pressure of 90 mmHg or greater, or under the antihypertensive treatment. We used the definition of diabetes determined by the Japan Diabetes Society (21). Coronary artery disease (CAD) was defined as any incident associated with coronary artery stenosis assessed through angiogram and/or computed tomography (22).

Statistical Analysis

Fisher's exact test or chi-square test was used for categorical variables. Continuous variables with a normal distribution were presented as mean ± standard deviation. Median values and interquartile ranges (IQRs) were presented for those exhibiting non-normal distribution. Student's *t*-test was used for continuous variables. Median values were compared using nonparametric Wilcoxon Mann–Whitney rank sum test or chi-square test with Fisher's *post-hoc* test. Multivariable linear analyses, including age and sex, were used to assess the association between genetic variations and LDL cholesterol. Kruskal–Wallis one-way analysis of variance followed by Tukey *post-hoc* test were used to test the differences among the groups. All of the statistical analyses were conducted with R statistical software, and $p < 0.05$ were considered significant statistically.

RESULTS

Characteristics

The characteristics are presented in **Table 1**. The mean age was 48 years, and 211 (19%) of the patients had CAD. There were 590 patients with monogenic FH with a pathogenic mutation in one of the FH genes without APOE7 mutant (52%, monogenic FH, group 3), whereas 519 patients did not have FH mutations or APOE7 mutant (46%, mutation negative, group 1). Twenty-one patients had APOE7 mutant but not FH genes (2%, APOE7 mutant carriers, group 2). We identified eight patients with

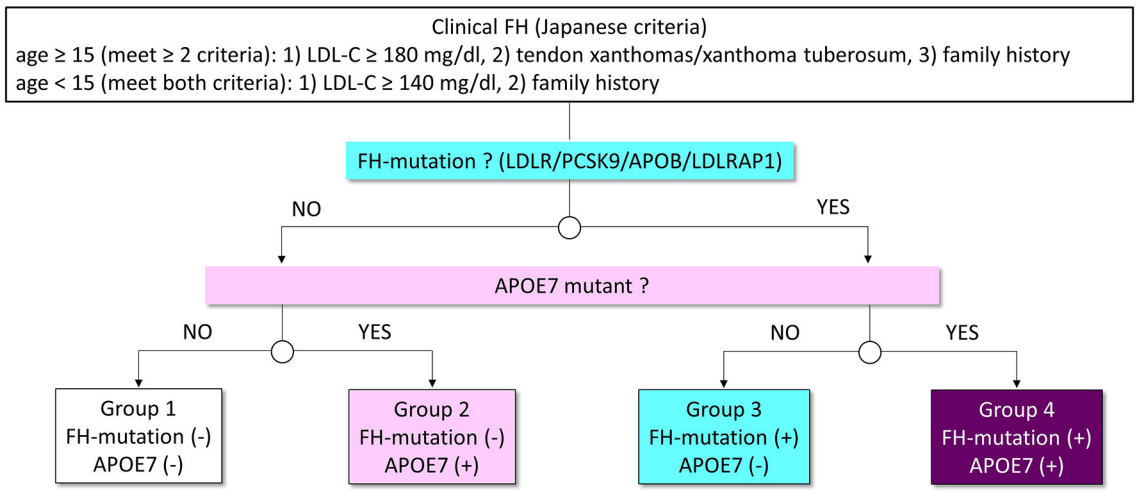


FIGURE 1 | Analytic scheme in this study. Patients with FH were classified into four groups based on the status of FH mutation and APOE7 mutant (Groups 1–4). White indicates patients without mutations. Pink indicates patients with APOE7 mutant. Light blue indicates patients with a mutation in FH genes. Purple indicates patients with a mutation in FH genes and APOE7 mutant.

TABLE 1 | Characteristics of the patients.

Variable	All (n = 1,138)	Classification				p-value
		Group 1 (n = 519) No mutation	Group 2 (n = 21) APOE7 mutant	Group 3 (n = 590) Monogenic FH	Group 4 (n = 8) Oligogenic FH (monogenic FH + APOE7 mutant)	
Age (years)	48 ± 20	53 ± 18	44 ± 21	40 ± 19	42 ± 21	$<1 \times 10^{-16}$
Male	512 (45%)	236 (45%)	11 (52%)	260 (44%)	5 (63%)	0.18
Hypertension	213 (19%)	117 (23%)	5 (24%)	91 (15%)	0 (0%)	0.012
Diabetes	69 (6%)	33 (6%)	5 (24%)	31 (5%)	0 (0%)	6.9×10^{-4}
Smoking	221 (19%)	124 (24%)	7 (33%)	89 (15%)	1 (13%)	0.14
Total cholesterol (mg/dl)	306 [279–366]	303 [271–327]	343 [277–389]	327 [264–391]	355 [298–398]	4.6×10^{-5}
Triglyceride (mg/dl)	169 [88–204]	164 [81–182]	216 [94–256]	145 [73–199]	188 [87-314]	8.9×10^{-4}
HDL cholesterol (mg/dl)	52 [42–64]	52 [43–65]	51 [44–61]	52 [44–63]	51 [46–62]	0.23
LDL cholesterol (mg/dl)	231 [191–271]	218 [189–243]	249 [196–277]	245 [201–281]	265 [216–319]	7.2×10^{-6}
RLP cholesterol (mg/dl)	7.4 [2.5–9.8]	7.3 [2.4–9.6]	12.3 [5.8–21.5]	7.2 [2.6–8.8]	11.4 [2.5–19.0]	1.6×10^{-4}
CAD	211 (19%)	102 (20%)	2 (10%)	105 (18%)	2 (25%)	0.026

FH, familial hypercholesterolemia; CAD, coronary artery disease; RLP, remnant-like particle.

pathogenic mutations in FH genes as well as in APOE7 mutant (0.7%, APOE7-oligogenic FH, group 4).

Impact of APOE7 (Heterozygote) Mutant on LDL Cholesterol

LDL cholesterol levels in patients with APOE7-oligogenic FH, who had pathogenic mutation in FH genes as well as APOE7 mutant, were significantly higher than those in patients with monogenic FH with a single pathogenic mutation in FH gene (265 vs. 245 mg/dl; **Table 1** and **Figure 2**). Moreover, LDL cholesterol level in patients with monogenic FH was significantly higher than those without any mutations (group 1) (245 vs. 218 mg/dl; **Table 1** and **Figure 2**). Under these conditions, LDL cholesterol level of APOE7 mutant carriers (group 2) was significantly higher than that of the patients without any mutations (group 1) (249 vs. 218 mg/dl) (**Table 1** and **Figure 2**).

Impact of APOE7 Mutant on Triglycerides and Remnant-Like Particle Cholesterol

In addition to LDL cholesterol, triglyceride levels in the APOE7 mutant group (group 2) were significantly higher than those without any mutations (group 1) (216 vs. 164 mg/dl; **Table 1** and **Figure 3A**). Additionally, remnant-like particle cholesterol levels in the APOE7 mutant group (group 2) were significantly higher than those without any mutations (group 1) (12.3 vs. 7.3 mg/dl; **Table 1** and **Figure 3B**).

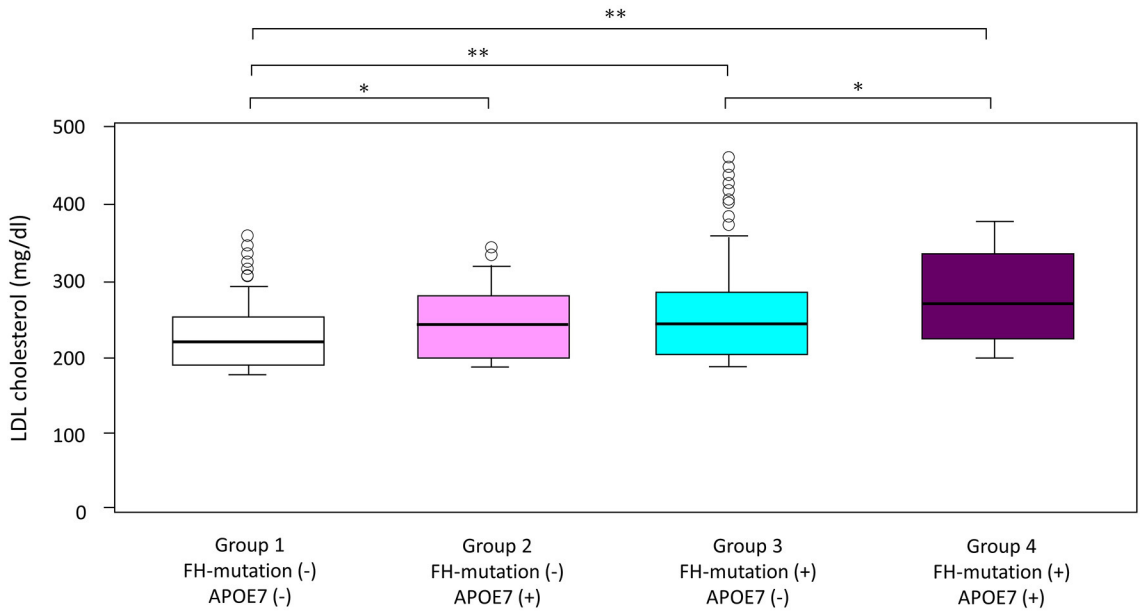


FIGURE 2 | Impact of FH gene and APOE7 on LDL cholesterol. Boxplots are showing LDL cholesterol. *$p < 0.05$ and **$p < 0.001$. Analyses of variance followed by Tukey *post-hoc* test were used to test the differences among the groups.

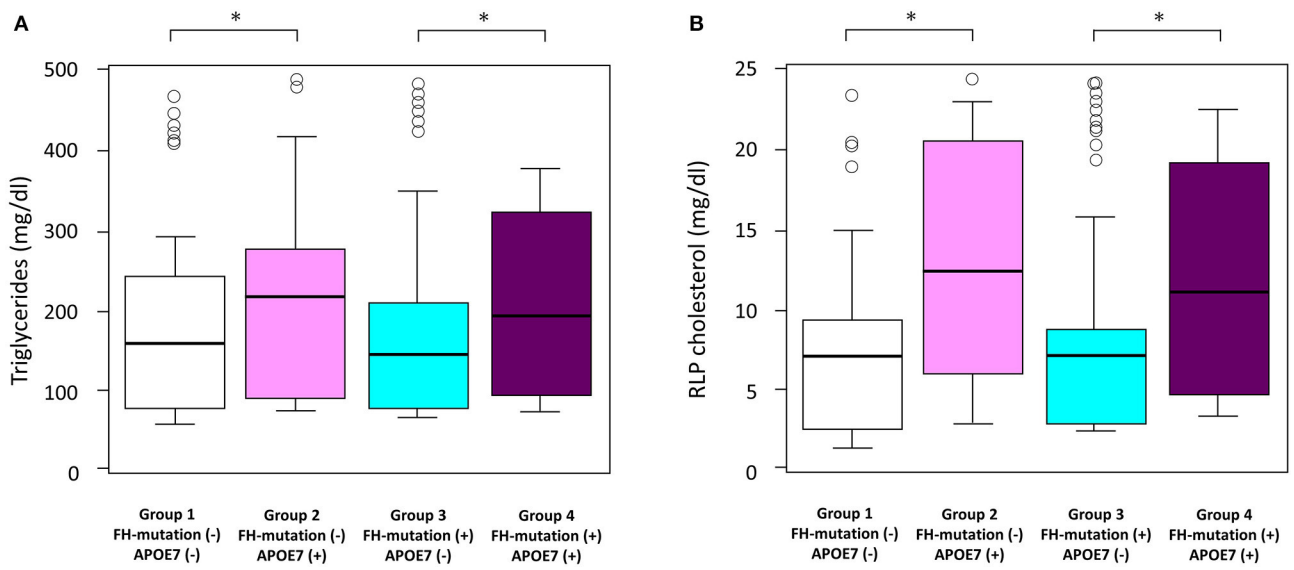


FIGURE 3 | Impact of FH gene and APOE7 on **(A)** triglyceride and **(B)** RLP cholesterol levels. Boxplots are showing triglyceride and RLP cholesterol. *$p < 0.05$. Analyses of variance followed by Tukey *post-hoc* test were used to test the differences among the groups.

Associations Between LDL-Associated Genetic Mutations and LDL Cholesterol

Multivariate linear analyses revealed that one mutation in the FH genes increased LDL cholesterol level by 46 mg/dl (95% CI, 21–51; $p = 2.3 \times 10^{-12}$; **Table 2**), while an APOE7 mutant elevated LDL cholesterol by 31 mg/dl (95% CI, 6–45; $p = 7.4 \times 10^{-4}$; **Table 2**). Moreover, we found that one protein-truncating mutation had much larger effect on LDL cholesterol than one missense mutation in *LDLR* or *PCSK9* (**Table 2**).

DISCUSSION

We aimed to assess the impact of a rare mutant APOE7 on serum lipids in clinically diagnosed FH patients. Among patients without any FH mutations, 21 (4%) had APOE7 mutant (heterozygote). We also identified eight individuals with a FH mutation and APOE7 mutant (heterozygote) exhibiting significantly elevated LDL cholesterol levels.

TABLE 2 | Mutation type and its association with LDL cholesterol.

Variable	Effect (mg/dl)	95% CI	*p*-value
Mutation (all FH genes) [$n = 598$]	46	21–51	2.3×10^{-12}
Protein truncating mutation (*LDLR*) [$n = 305$]	50	29–71	$<1 \times 10^{-16}$
Missense mutation (*LDLR*) [$n = 214$]	38	23–44	4.4×10^{-7}
Missense mutation (*PCSK9*) [$n = 79$]	28	10–39	3.9×10^{-5}
APOE7 mutant [$n = 29$]	31	6–45	7.4×10^{-4}

FH, familial hypercholesterolemia; LDLR, LDL receptor; PCSK9, proprotein convertase subtilisin/kexin type 9; CI, confidence interval.
Multivariable linear analyses, including age and sex, were used to assess the association between genetic variations and LDL cholesterol.

According to gnomAD browser, the allele frequency of APOE7 mutant is 0.002698 in East Asians, and it is not found among other ethnicities. The unexpectedly high prevalence of this variant among the phenotype of FH strongly suggests the association between this variant and LDL cholesterol.

Despite the rigorous effort of uncovering the genetic backgrounds of FH, approximately 30–40% of patients with clinical FH do not exhibit pathogenic genetic mutation(s) in established FH genes. There are several potential explanations for this situation, including so-called polygenic FH caused by the accumulation of common genetic variations elevating their LDL cholesterol. Recently, we have shown that loss-of function mutations in *ABCG5* or *ABCG8* typically regarded as the cause of recessive disorder are contributing to mimic and/or exacerbate the FH phenotype (16, 17). In this study, we extended this notion to *APOE*, at least, a rare mutant of APOE7. The clinical impact of an APOE7 mutant appears to be smaller than that of other FH mutations in *LDLR*, or *PCSK9*. However, a proportion of individuals with a mutant of APOE7 can be considered to have typical FH who are showing dominant inheritance pattern.

APOE plays an important role as the ligand, whereas the transportation of lipids mediates interaction between lipoproteins and lipoprotein receptors, including LDLR and remnant receptor. Moreover, dysfunction caused by homozygous APOE2 has been shown to cause dyslipidemia, called type III hyperlipoproteinemia (23, 24). This situation has been shown to be associated with premature ASCVD. In addition to a rather "common" situation, rare genetic variations, including a mutant APOE7, have been identified as the cause of inherited dyslipidemia. However, its prevalence and clinical impact have long been unclear. Recent advances in comprehensive genetic analytic scheme enabled us to assess this issue. We believe that the increase in triglyceride and remnant cholesterol levels is an additional risk factor for ASCVD beyond increase in LDL cholesterol level.

Study Limitations

This study declares some limitations. First, it was an observational study assessed cross-sectionally conducted in a retrospective manner. Moreover, relatively small sample size without replication may lead to a false conclusion. Second, some patients were excluded from this analysis due to missing data, which may cause a bias. Third, validations through Sanger sequencing were not performed. Fourth, we essentially used Friedewald formula, which is closer to LDL cholesterol plus intermediate-density lipoprotein (IDL) cholesterol rather than LDL cholesterol alone. In addition, we used LDL cholesterol determined directly when triglycerides $\geq$400 mg/dl. The mixture of the determination may lead to biased observation. Fourth, the number of individuals with oligogenic FH was small; thus, the statistical analyses relating this group may be underpowered. Lastly, we could not account for common genetic variations associated with LDL cholesterol level in this study. Other single-nucleotide polymorphisms (SNPs) are more present and have been causally related to FH; hence, the impact of this low prevalent mutation may be small.

CONCLUSIONS

We identified more patients with a APOE7 mutant than expected in clinically diagnosed FH patients, especially among those without FH-causing mutation. This implies that a proportion of FH cases may be caused by this particular genetic mutation.

DATA AVAILABILITY STATEMENT

Requests to access the datasets should be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by IRB of the Kanazawa University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

HT contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript. KY, NK, JS, TN, AN, SU, KS, MT, and M-aK contributed to recruiting the patients and to the writing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work has been supported by scientific research grants from the Ministry of Education, Science and Culture of Japan (No. 19K08575); Health, Labor, and Welfare Sciences Research Grant for Research on Rare and Intractable Diseases; and Japanese Circulation Society.

ACKNOWLEDGMENTS

We would like to express special thanks to Mr. Sachio Yamamoto, Ms. Kazuko Honda, and Yoko Iwauchi (Kanazawa University) for their technical assistance.

REFERENCES

1. Goldstein JL, Hobbs HH, Brown MS. Familial hypercholesterolemia. In: Scriver CR, Beaudet AL, Sly WS, Valle D, editors. *The Metabolic and Molecular Bases of Inherited Disease*. New York, NY: McGraw-Hill (2001). 2863e913 p.
2. Mabuchi H. Half a century tales of familial hypercholesterolemia (FH) in Japan. *J Atheroscler Thromb.* (2017) 24:189–207. doi: 10.5551/jat.RV16008
3. Sturm AC, Knowles JW, Gidding SS, Ahmad ZS, Ahmed CD, Ballantyne CM, et al. Clinical genetic testing for familial hypercholesterolemia: JACC scientific expert panel. *J Am Coll Cardiol.* (2018) 72:662–80. doi: 10.1016/j.jacc.2018.05.044
4. Soutar AK, Myant NB, Thompson GR. The metabolism of very low density and intermediate density lipoproteins in patients with familial hypercholesterolaemia. *Atherosclerosis.* (1982) 43:217–31. doi: 10.1016/0021-9150(82)90024-7
5. Zannis VI, Breslow JL. Apolipoprotein E. *Mol Cell Biochem.* (1982) 42:3–20. doi: 10.1007/BF00223534
6. Utermann G. Apolipoprotein E (role in lipoprotein metabolism and pathophysiology of hyperlipoproteinemia type III). *Ric Clin Lab.* (1982) 12:23–30.
7. Bennet AM, Di Angelantonio E, Ye Z, Wensley F, Dahlin A, Ahlbom A, et al. Association of apolipoprotein E genotypes with lipid levels and coronary risk. *JAMA*. (2007) 298:1300–11. doi: 10.1001/jama.298.11.1300
8. Yamamura T, Dong LM, Yamamoto A. Characterization of apolipoprotein E7 (Glu(244)–>Lys, Glu(245)—>Lys), a mutant apolipoprotein E associated with hyperlipidemia and atherosclerosis. *J Lipid Res.* (1999) 40:253–9. doi: 10.1016/S0022-2275(20)33364-2
9. Kitahara M, Shinomiya M, Shirai K, Saito Y, Yoshida S. Frequency and role of apo E phenotype in familial hypercholesterolemia and non-familial hyperlipidemia in the Japanese. *Atherosclerosis*. (1990) 82:197–204. doi: 10.1016/0021-9150(90)90041-G
10. Yamanouchi Y, Arinami T, Tsuchiya S, Miyazaki R, Takaki H, Takano T, et al. Apolipoprotein E5 and E7 in apparently healthy Japanese males: frequencies and relation to plasma lipid levels. *Jpn J Hum Genet.* (1994) 39:315–24. doi: 10.1007/BF01874050
11. Matsunaga A, Sasaki J, Moriyama K, Arakawa F, Takada Y, Nishi K, et al. Population frequency of apolipoprotein E5 (Glu3–>Lys) and E7 (Glu244–>Lys, Glu245–>Lys) variants in western Japan. *Clin Genet.* (1995) 48:93–9. doi: 10.1111/j.1399-0004.1995.tb04063.x
12. Yanagi K, Yamashita S, Hiraoka H, Ishigami M, Kihara S, Hirano K, et al. Increased serum remnant lipoproteins in patients with apolipoprotein E7 (apo E Suita). *Atherosclerosis.* (1997) 131:49–58. doi: 10.1016/S0021-9150(96)06068-6
13. gnomAD browser. Available online at: https://gnomad.broadinstitute.org/ (accessed November 29, 2020).
14. Yamamura T, Dong LM, Yamamoto A. Apolipoprotein E polymorphism and coronary heart disease. *Chin Med J.* (1992) 105:738–41.
15. Ueyama Y, Nozaki S, Yanagi K, Hiraoka H, Nakagawa T, Sakai N, et al. Familial hypercholesterolaemia-like syndrome with apolipoprotein E-7 associated with marked Achilles tendon xanthomas and coronary artery disease: a report of two cases. *J Intern Med.* (1994) 235:169–74. doi: 10.1111/j.1365-2796.1994.tb01051.x
16. Tada H, Kawashiri MA, Nomura A, Teramoto R, Hosomichi K, Nohara A, et al. Oligogenic familial hypercholesterolemia, LDL cholesterol, and coronary artery disease. *J Clin Lipidol.* (2018) 12:1436–44. doi: 10.1016/j.jacl.2018.08.006
17. Tada H, Okada H, Nomura A, Yashiro S, Nohara A, Ishigaki Y, et al. Rare and deleterious mutations in ABCG5/ABCG8 genes contribute to mimicking and worsening of familial hypercholesterolemia phenotype. *Circ J.* (2019) 83:1917–24. doi: 10.1253/circj.CJ-19-0317
18. Harada-Shiba M, Arai H, Ishigaki Y, Ishibashi S, Okamura T, Ogura M, et al. Guidelines for diagnosis and treatment of familial hypercholesterolemia 2017. *J Atheroscler Thromb.* (2018) 25:751–70. doi: 10.5551/jat.CR003
19. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* (2016) 536:285–91. doi: 10.1038/nature19057
20. Tada H, Kawashiri MA, Yoshida T, Teramoto R, Nohara A, Konno T, et al. Lipoprotein(a) in familial hypercholesterolemia with proprotein convertase subtilisin/kexin type 9 (PCSK9) gain-of-function mutations. *Circ J.* (2016) 80:512–8. doi: 10.1253/circj.CJ-15-0999
21. Committee of the Japan Diabetes Society on the Diagnostic Criteria of Diabetes Mellitus, Seino Y, Nanjo K, Tajima N, Kadowaki T, Kashiwagi A, et al. Report of the committee on the classification and diagnostic criteria of diabetes mellitus. *J Diabetes Investig.* (2010) 1:212–28. doi: 10.1111/j.2040-1124.2010.00074.x
22. Tada H, Kawashiri MA, Nohara A, Inazu A, Mabuchi H, Yamagishi M. Impact of clinical signs and genetic diagnosis of familial hypercholesterolaemia on the prevalence of coronary artery disease in patients with severe hypercholesterolaemia. *Eur Heart J.* (2017) 38:1573–9. doi: 10.1093/eurheartj/ehx004
23. Phillips MC. Apolipoprotein E isoforms and lipoprotein metabolism. *IUBMB Life.* (2014) 66:616–23. doi: 10.1002/iub.1314
24. Hopkins PN, Brinton EA, Nanjee MN. Hyperlipoproteinemia type 3: the forgotten phenotype. *Curr Atheroscler Rep.* (2014) 16:440. doi: 10.1007/s11883-014-0440-2

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 20 April 2021
doi: 10.3389/fcvm.2021.658915

Common Variants Associated With *OSMR* Expression Contribute to Carotid Plaque Vulnerability, but Not to Cardiovascular Disease in Humans

Danielle van Keulen[1,2,3,4], Ian D. van Koeverden[1], Arjan Boltjes[2], Hans M. G. Princen[4], Alain J. van Gool[5,6], Gert J. de Borst[7], Folkert W. Asselbergs[8,9,10], Dennie Tempel[2,3,11], Gerard Pasterkamp[2†] and Sander W. van der Laan[2*†]*

[1] Laboratory of Experimental Cardiology, University Medical Center Utrecht, University of Utrecht, Utrecht, Netherlands, [2] Central Diagnostics Laboratory, University Medical Center Utrecht, University of Utrecht, Utrecht, Netherlands, [3] Quorics B.V., Rotterdam, Netherlands, [4] TNO-Metabolic Health Research, Gaubius Laboratory, Leiden, Netherlands, [5] Translational Metabolic Laboratory, Radboudumc, Nijmegen, Netherlands, [6] TNO- Microbiology & Systems Biology, Zeist, Netherlands, [7] Department of Vascular Surgery, University Medical Center Utrecht, University of Utrecht, Utrecht, Netherlands, [8] Department of Cardiology, Division Heart & Lungs, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands, [9] Faculty of Population Health Sciences, Institute of Cardiovascular Science, University College London, London, United Kingdom, [10] Health Data Research UK and Institute of Health Informatics, University College London, London, United Kingdom, [11] SkylineDx B.V., Rotterdam, Netherlands

OPEN ACCESS

Edited by:
Seitaro Nomura,
The University of Tokyo, Japan

Reviewed by:
Christoph Sinning,
University Heart and Vascular Center Hamburg (UHZ), Germany
Lasse Folkersen,
Danish National Genome Center, Denmark

***Correspondence:**
Gerard Pasterkamp
g.pasterkamp@umcutrecht.nl
Sander W. van der Laan
s.w.vanderlaan-2@umcutrecht.nl

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: 26 January 2021
Accepted: 09 March 2021
Published: 20 April 2021

Citation:
van Keulen D, van Koeverden ID, Boltjes A, Princen HMG, van Gool AJ, de Borst GJ, Asselbergs FW, Tempel D, Pasterkamp G and van der Laan SW (2021) Common Variants Associated With OSMR Expression Contribute to Carotid Plaque Vulnerability, but Not to Cardiovascular Disease in Humans. Front. Cardiovasc. Med. 8:658915. doi: 10.3389/fcvm.2021.658915

Background and Aims: Oncostatin M (OSM) signaling is implicated in atherosclerosis, however the mechanism remains unclear. We investigated the impact of common genetic variants in *OSM* and its receptors, *OSMR* and *LIFR,* on overall plaque vulnerability, plaque phenotype, intraplaque *OSMR* and *LIFR* expression, coronary artery calcification burden and cardiovascular disease susceptibility.

Methods and Results: We queried Genotype-Tissue Expression data and found that rs13168867 (C allele) was associated with decreased *OSMR* expression and that rs10491509 (A allele) was associated with increased *LIFR* expression in arterial tissues. No variant was significantly associated with *OSM* expression.

We associated these two variants with plaque characteristics from 1,443 genotyped carotid endarterectomy patients in the Athero-Express Biobank Study. After correction for multiple testing, rs13168867 was significantly associated with an increased overall plaque vulnerability ($\beta = 0.118 \pm$ s.e. $= 0.040$, $p = 3.00 \times 10^{-3}$, C allele). Looking at individual plaque characteristics, rs13168867 showed strongest associations with intraplaque fat ($\beta = 0.248 \pm$ s.e. $= 0.088$, $p = 4.66 \times 10^{-3}$, C allele) and collagen content ($\beta = -0.259 \pm$ s.e. $= 0.095$, $p = 6.22 \times 10^{-3}$, C allele), but these associations were not significant after correction for multiple testing. rs13168867 was not associated with intraplaque *OSMR* expression. Neither was intraplaque *OSMR* expression associated with plaque vulnerability and no known *OSMR* eQTLs were associated with coronary artery calcification burden, or cardiovascular disease susceptibility. No associations were found for rs10491509 in the *LIFR* locus.

Conclusions: Our study suggests that rs1316887 in the OSMR locus is associated with increased plaque vulnerability, but not with coronary calcification or cardiovascular disease risk. It remains unclear through which precise biological mechanisms OSM signaling exerts its effects on plaque morphology. However, the OSM-OSMR/LIFR pathway is unlikely to be causally involved in lifetime cardiovascular disease susceptibility.

Keywords: cardiovascular disease, atherosclerosis, plaque, genetics, OSM, OSMR, LIFR

INTRODUCTION

The prevalence of cardiovascular disease (CVD) is high, poses a significant global burden and is expected to rise (1). Arterial inflammation, leading to asymmetric focal arterial thickening and atherosclerotic plaque formation and progression, is the primary mechanism underlying CVD (2). Inflammatory cytokines contribute to arterial inflammation and subsequent atherosclerotic plaque formation (3). One cytokine, for which there is mounting evidence suggesting a role in atherosclerosis development is OSM (4, 5). It has been shown that OSM is present in both murine and human atherosclerotic plaques (6). Moreover, murine studies showed that OSM receptor $(OSMR)^{-/-}ApoE^{-/-}$ mice have reduced plaque size and improved plaque stability compared to their *OSMR*-expressing littermates (7), indicating that OSM drives atherosclerosis development. These observations are in line with our previous work, in which we showed that simultaneous signaling of OSM through OSMR and leukemia inhibitory factor receptor (LIFR), induces activation in human endothelial cells, suggestive of a role in atherosclerosis development (8). In contrast, chronic OSM administration to APOE*3Leiden.CETP mice reduces the atherosclerotic lesion size and severity, and high circulating OSM levels correlate with increased post-incident coronary heart disease survival probability in humans (9).

Although all these studies implicate that OSM is involved in atherosclerosis, little is known about the effects of OSM on plaque composition in humans. Grouped in the interleukin 6 subfamily of cytokines, OSM is released by activated immune cells (10–12), and exerts pleiotropic effects on cell proliferation, inflammation, hematopoiesis, tissue remodeling, and development (13). Its signals are transduced through binding to either OSMR or LIFR, which form a heterodimer with glycoprotein 130 (8, 14), that in turn activates multiple pathways (14). It is suggested that the ratio of the two receptor types expressed on the cell membrane is a potential regulatory mechanism for the multiple, and sometimes opposing, effects that are exerted by OSM (15).

Thus, given its pleiotropic function, it is difficult to predict how OSM contributes to atherosclerotic plaque formation. Cell and murine studies have shown that OSM promotes angiogenesis (4), endothelial activation (8), vessel permeability (16), and osteoblastic differentiation (17). Therefore, increased OSM levels hypothetically results in a higher intraplaque microvessel density, intraplaque hemorrhages and plaque calcification, thereby contributing to plaque destabilization (18, 19). In other cell and murine studies, OSM promotes fibroblast proliferation (20), collagen formation (20), smooth muscle cell proliferation (6), and M2 macrophage polarization (21). These processes hypothetically lead to enhanced fibrosis, and attenuated inflammation, thereby contributing to plaque stabilization (22).

Large-scale studies have shown that *cis*-acting genetic variants associated to gene expression [expression quantitative trait loci (eQTLs) (23)] are key to disease susceptibility (24). This means that gene expression in a given tissue differs between individuals carrying different genotypes which ultimately results in differential disease susceptibility. Thus, on the premise that alleles are randomly distributed at conception and are invariant throughout a lifetime, meaning that genetics is not influenced by disease or risk factors, eQTLs can be used as proxies of gene expression to examine the effect on plaque morphology (25). We hypothesized that if circulating *OSM*, or arterial *OSMR* or *LIFR* expression has an effect on plaque morphology, these phenotypic differences will be observed among genotype groups of the eQTL. We aimed to investigate the double-edged sword of OSM signaling on the composition of human atherosclerotic plaques using known eQTLs of circulating *OSM*, and arterial *OSMR* and *LIFR*.

MATERIALS AND METHODS

Sample Collection

The Athero-Express Biobank Study (https://www.atheroexpress.nl) is an ongoing prospective study, containing biomaterial of patients elected for endarterectomy at two Dutch tertiary referral centers. Details of the study design were described before (26). Briefly, blood subfractions are obtained before and arterial plaque material during endarterectomy. Each plaque is dissected into segments of 0.5 cm. The culprit lesion is reserved for histological assessment (see below), while surrounding segments are immediately snap frozen in liquid nitrogen and stored at −80°C for later use, e.g., in order to perform RNA-seq (see below). Only carotid endarterectomy (CEA) patients were included in the present study. All research was conducted according to the principles of the Declaration of Helsinki and its later amendments, all patients provided informed consent and the study was approved by the medical ethics committees.

Athero-Express Genotyping, Quality Control, and Imputation

Details of genotyping were previously described (26). Briefly, DNA was extracted from EDTA blood or (when no blood was available) plaque samples of 1,858 consecutive patients from the Athero-Express Biobank Study and genotyped in 2 batches. For

the Athero-Express Genomics Study 1 (AEGS1), 836 patients, included between 2002 and 2007, were genotyped using the Affymetrix Genome-Wide Human SNP Array 5.0 (SNP5) chip (Affymetrix Inc., Santa Clara, CA, USA). For the Athero-Express Genomics Study 2 (AEGS2), 1,022 patients, included between 2002 and 2013 and not overlapping AEGS1, were genotyped using the Affymetrix Axiom® GW CEU 1 Array (AxM).

Both studies were carried out according to OECD standards. After genotype calling, we adhered to community standard quality control and assurance (QA/QA) procedures of the genotype data from AEGS1 and AEGS2. Samples with low average genotype calling and sex discrepancies (compared to the clinical data available) were excluded. The data was further filtered on (1) individual (sample) call rate >97%, (2) SNP call rate >97%, (3) minor allele frequencies (MAF) >3%, (4) average heterozygosity rate ± 3.0 s.d., (5) relatedness (pi-hat >0.20), (6) Hardy–Weinberg Equilibrium (HWE $p < 1.0 \times 10^{-6}$), and (7) population stratification (based on HapMap 2, release 22, b36) by excluding samples deviating more than 6 standard deviations from the average in five iterations during principal component analysis (PCA) and by visual inspection as previously described (26). After QA/QA, 657 samples and 403,789 SNPs in AEGS1, and 869 samples and 535,983 SNPs in AEGS2 remained. To correct for genetic ancestry and population stratification we performed PCA in each cleaned dataset to obtain principal components for downstream analyses as described before (26).

We used SHAPEIT2 (27) for phasing and finally the data was imputed with 1000G phase 3 (28) and GoNL 5 (29) as a reference on genome build 37. Note that we only selected the CEA patients in these datasets, leaving 1,443 samples for our further analyses.

Variant Selection

We queried data from the Genotype-Tissue Expression (GTEx) Portal (https://gtexportal.org) (23) for *cis*-acting variants [defined as variants within 1Mb of a given gene (30)] that alter *OSM* expression in whole blood, and *OSMR* or *LIFR* expression in non-diseased arterial tissue. We selected common variants with a MAF > 3%, which yielded two variants in total: rs13168867 for *OSMR* in tibial arterial tissue and rs10491509 for *LIFR* in aortic arterial tissue. We found no eQTL for circulating *OSM* expression, i.e., in whole blood. We harmonized the effect alleles and effect sizes from these eQTLs to match the allele orientation in the Athero-Express Biobank Study data.

Plaque Phenotyping

The (immuno)histochemical analysis of plaques have been described previously (26, 31, 32). Briefly, per plaque, the culprit lesion was identified directly after dissection, fixed in 4% formaldehyde, embedded in paraffin and cut in 5 μm sections on a microtome for (immuno)histochemical analysis by pathology experts. Calcification (hematoxylin & eosin, H&E) and collagen content (picrosirius red) were semi-quantitatively scored and defined as no/minor or moderate/heavy. Atheroma size (H&E and picrosirius red) was defined as <10% or ≥10% fat content. Macrophages (CD68) and smooth muscle cells (ACTA2) were quantitatively scored and classified as percentage of plaque area. Intraplaque hemorrhage (H&E) was defined as absent or present, and vessel density was classified as the number of intraplaque vessels (CD34) per 3–4 hotspots.

Plaque Vulnerability

Assessment of overall plaque vulnerability was performed as previously described by Verhoeven et al. (25). Briefly, macrophages and smooth muscle cells were semi-quantitatively defined as no/minor or moderate/heavy. Each plaque characteristic that defines a stable plaque (i.e., no/minor macrophages, moderate/heavy collagen, moderate/heavy smooth muscle cells and <10% fat) was given a score of 0, while each plaque characteristic that defines a vulnerable plaque (i.e., moderate/heavy macrophages, no/minor collagen, no/minor smooth muscle cells and ≥10% fat) was given a score of 1. The score of each plaque characteristic was summed resulting in a final plaque score ranging from 0 (most stable plaque) to 4 (most vulnerable plaque). Intraobserver and interobserver variability were examined previously and showed good concordance (κ = 0.6–0.9) (33).

Plaque Expression

Detailed information on the RNA sequencing (RNAseq) experiment is described in the **Supplemental Material**. In short, to assess the global expression profile, plaque segments were thawed, cut up, and homogenized using ceramic beads and tissue homogenizer (Precellys, Bertin instruments, Montigny-le-Bretonneux, France), in the presence of TriPure (Sigma Aldrich), and RNA was isolated according to TriPure manufacturer's protocol.

Library Preparation

was performed, adapting the CEL-Seq2 protocol for library preparation (34, 35), as described before (36). The primer used for initial reverse-transcription reaction was designed as follows: an anchored polyT, a unique 6bp barcode, a unique molecular identifier (UMI) of 6bp, the 5' Illumina adapter and a T7 promoter, as described (36). Complementary DNA (cDNA) was then used in the *in vitro* transcription (IVT) reaction (AM1334; Thermo-Fisher). Amplified RNA (aRNA) was fragmented, and cleaned, and RNA yield and quality in the suspension were checked by Bioanalyzer (Agilent).

cDNA library construction was initiated according to the manufacturer's protocol, adding randomhexRT primer as random primer. PCR amplification was done with Phusion High-Fidelity PCR Master Mix with HF buffer (NEB, MA, USA) and a unique indexed RNA PCR primer (Illumina) per reaction. Library cDNA yield and quality were checked by Qubit fluorometric quantification (Thermo-Fisher) and Bioanalyzer (Agilent), respectively. Libraries were sequenced on the Illumina Nextseq500 platform; paired end, 2 × 75 bp.

Upon sequencing, retrieved fastq files were de-barcoded, split into forward and reverse reads. Subsequently, these were mapped making use of Burrows-Wheel aligner [BWA (37)] version 0.7.17-r1188 and a cDNA reference (assembly hg19, Ensembl release 84). Read counts and UMI counts were derived from SAM files using custom perl code, and then gathered into count matrices. Genes were annotated with Ensembl ID's,

and basic quality control was performed, encompassing filtering out samples with low gene numbers (<10,000 genes), and read numbers (<18,000 reads). These steps resulted in 641 samples with up to 60,674 genes (Ensembl ID's), and median of 178,626 reads per sample.

Data Analysis

Plaque vulnerability scores, and genotypes for rs10491509 and rs13168867, were added to metadata, upon which this was combined with counts and annotation in a SummarizedExperiment object (38). Counts were normalized and transformed making use of the variance stabilization transformation function (vst()) in DESeq2 (39). This results in transformed data on a $\log_2$-scale, normalized for library size, for visualization and ordination purposes. Differential expression analysis between plaque vulnerability scores or genotypes, used as "condition variables" was performed using the DESeq2-function DESeq() on the raw counts. In short, three steps are performed: 1. estimation of size factors, controlling for sequencing depth; 2. estimation of dispersion values, that capture variation around expected values. These expected values take into account sequencing depth and differences caused by variables in the design formula argument, i.e., "design = ~ condition" where condition is a variable that specifies which group samples belong to; and 3. fitting a generalized linear model using the above-mentioned size factors and dispersion values, estimating log fold changes. This results in a results table, showing estimated log2 fold changes and p-values comparing between two levels of the condition variable. Complete details for statistical procedures used by the DESeq function are described elsewhere (39).

Genetic Analyses

Quantitatively scored characteristics (macrophages, smooth muscle cells, and the vessel density) were Box-Cox transformed (40) to obtain a normal distribution. For genetic analyses we used GWASToolKit (https://swvanderlaan.github.io/GWASToolKit/) which is a wrap-around collection of scripts for SNPTEST (41). Continuous and categorical variables were tested using linear and logistic regression models, respectively. Models for genetic analyses were corrected for age, sex, genotyping chip, and genetic ancestry using principal components 1 through 4. Thus, the models were of the form

$$\text{phenotype} \sim \text{SNP} + \text{age} + \text{sex} + \text{genotypingchip} + \text{PC 1} + \text{PC 2} + \text{PC 3} + \text{PC 4}.$$

Multiple Testing and Power

Correction for multiple testing resulted in a corrected p-value of $p = 0.05/[(7 \text{ plaque phenotypes} + \text{plaque vulnerability}) \times 2 \text{ common variants}] = 3.13 \times 10^{-3}$. The power of the study was estimated at ±75% based on a sample size of 1,443, a minor allele frequency (MAF) of 0.409 and a relative risk of 1.28 (http://csg.sph.umich.edu/abecasis/cats/gas_power_calculator/).

TABLE 1 | Baseline characteristics of genotyped CEA patients from the Athero-Express Biobank Study.

Patient characteristics		Missing data (%)
Sex, male, *n* (%)	976 (64.0)	5.7
Age in years, mean (SD)	68.84 (9.33)	5.7
History		
Cerebrovascular disease, *n* (%)	478 (33.2)	5.7
Coronary artery disease, *n* (%)	430 (29.9)	5.8
Peripheral artery disease, *n* (%)	297 (20.7)	5.8
Risk factors		
Type 2 diabetes mellitus, *n* (%)	332 (23.1)	5.7
Hypertension, *n* (%)	1,017 (73.0)	8.7
Current smoker, *n* (%)	492 (34.9)	7.5
BMI, median (IQR)	26.0 (24.0–28.4)	11.5
eGFR, median (IQR)	72.3 (58.7–85.4)	8.1
Total cholesterol in mmol/L, median (IQR)	4.38 (3.60–5.25)	22.8
LDL in mmol/L, median (IQR)	2.40 (1.81–3.13)	27.8
HDL in mmol/L, median (IQR)	1.06 (0.87–1.30)	25.0
Triglycerides in mmol/L, median (IQR)	1.50 (1.08–2.04)	24.6
Medication		
Anti-hypertensives, *n* (%)	1,110 (77.2)	5.8
Lipid lowering drugs, *n* (%)	1,112 (77.4)	5.8
Anti-thrombotics, *n* (%)	1,272 (88.6)	6.0
Symptoms		
Asymptomatics, *n* (%)	195 (13.6)	6.0
Ocular, *n* (%)	221 (15.4)	6.0
TIA, *n* (%)	634 (44.2)	6.0
Stroke, *n* (%)	384 (26.8)	6.0

Cerebrovascular disease history is defined by ischemic stroke prior to surgery. Coronary artery disease history includes diagnosed coronary artery disease, myocardial infarction, percutaneous coronary intervention, and coronary artery bypass grafting. Peripheral disease history includes diagnosed peripheral arterial occlusive disease, femoral artery interventions, and ankle-brachial index <70. Type 2 diabetes mellitus includes all individuals with diagnosed type 2 diabetes mellitus and those on appropriate medication. Hypertension includes all individuals with self-reported hypertension. Current smokers include all individuals smoking up to 6 months until the surgery date. BMI, kg/m^2. eGFR rate was based on the Modification of Diet in Renal Disease formula, mL/min/1.73 m^2. Anti-hypertensives include all anti-hypertension medication. Anti-thrombotics include clopidogrel, dipyridamole, acenocoumarol, ascal, and other anti-platelet drugs. Missing data shows the percentage of the patients of which we lack information on the specific patient characteristic.

Data and Scripts

Data is available upon request. Scripts are posted at GitHub https://github.com/swvanderlaan/2019_vankeulen_d_osmr.

RESULTS

Common Variants Altering *OSM*, *OSMR*, and *LIFR* Expression

We included and genotyped 1,443 carotid endarterectomy patients in this study. We combined these groups (**Table 1**) for overall plaque vulnerability and phenotype analyses, as we previously showed that the baseline characteristics between the two genotyping groups (AEGS1 and AEGS2) are comparable (26).

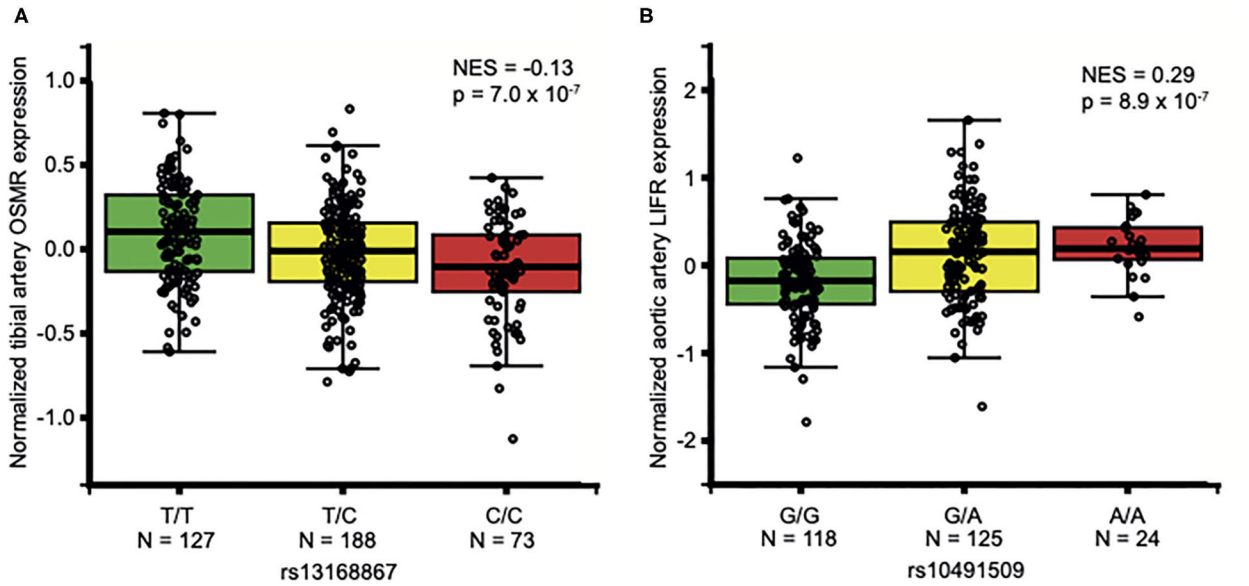


FIGURE 1 | Association of *OSMR* and *LIFR* variants in non-diseased arterial tissues. Per variant, the normalized expression of *OSMR* **(A)** and *LIFR* **(B)** is given in non-diseased arterial tissue. Data from GTEx Portal (www.gtexportal.org). NES: Normalized effect size. In aortic arterial tissue, rs13168867 had a NES of −0.123 and in tibial arterial tissue, rs10491509 had a NES of 0.0881 (**Supplementary Figures 1**, **2**).

OSM is secreted by, among others, neutrophils (12), monocytes (11), macrophages (11) and T-cells (10), and acts through binding to OSMR and LIFR (14, 42, 43) in the arterial wall (7, 44). Thus, we queried data from the Genotype-Tissue Expression project (GTEx) (23) for SNPs that alter *OSM* expression in whole blood and *LIFR* and *OSMR* expression in arterial tissue. There were no significant eQTLs for *OSM*, but there were two eQTLs associated with altered *OSMR* (rs13168867) or *LIFR* (rs10491509) expression in arterial tissue. The C allele of rs13168867 is associated with decreased *OSMR* expression in the tibial artery (**Figure 1A**), and the A allele of rs10491509 is associated with increased *LIFR* expression in the aortic artery (**Figure 1B**). Cross-tissue meta-analysis showed that these variants have m-values >0.9 in both tibial and aortic artery tissue, indicating a high posterior probability that they are single *cis*-eQTLs in both tissues (**Supplementary Figures 1**, **2**).

Genetic Association With Plaque Vulnerability

To determine the effect of OSM signaling on the overall plaque vulnerability, we correlated the rs13168867 and rs10491509 genotypes to the overall plaque vulnerability, which was given a score ranging from 0 (least vulnerable plaque) to 4 (most vulnerable plaque). The effect allele of variant rs13168867 in the *OSMR* locus was significantly correlated with an increased overall plaque vulnerability ($\beta = 0.118 \pm$ s.e. $= 0.040$ (C allele), $p = 3.00 \times 10^{-3}$, **Figure 2**). No association was observed with rs10491509 and overall plaque vulnerability.

Genetic Association With Plaque Phenotypes

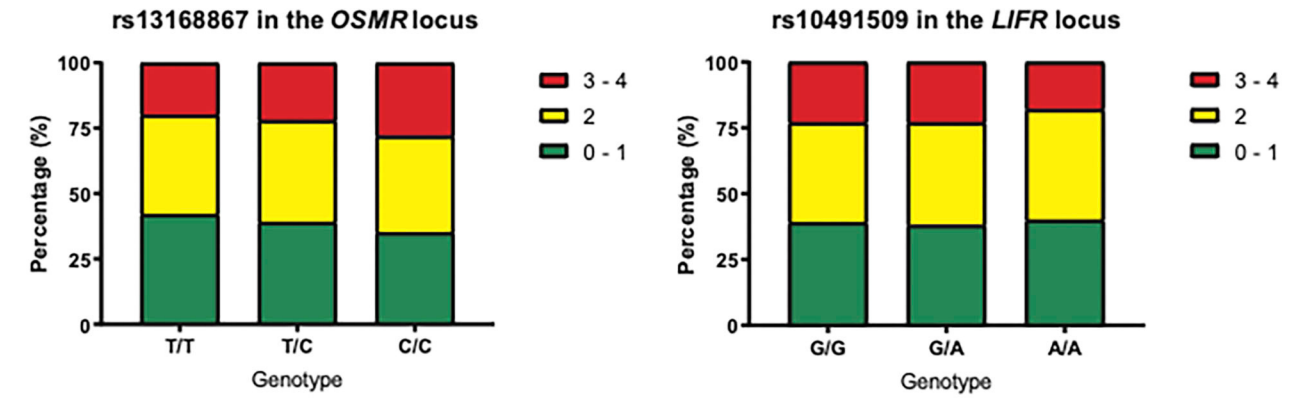
To determine the effect of OSM signaling on the individual plaque characteristics comprising the overall plaque vulnerability, we assessed the association between rs13168867 and rs10491509 and seven plaque phenotypes in the Athero-Express Biobank Study. Although not significant after correction for multiple testing, the strongest associations were observed between the effect allele of variant rs13168867 in the *OSMR* locus and intraplaque fat ($\beta = 0.248 \pm$ s.e. $= 0.088$ (C allele), $p = 4.66 \times 10^{-3}$), and collagen content ($\beta = -0.259 \pm$ s.e. $= 0.095$ (C allele), $p = 6.22 \times 10^{-3}$, **Table 2**). No associations were observed between rs10491509 and any of the plaque phenotypes.

Known eQTLs of *OSMR* and *LIFR* Expression in Non-diseased Arterial Tissue Are Not Associated With Expression in Atherosclerotic Plaques

Atherosclerotic disease progression changes the artery-specific transcriptional dynamics, and may therefore abolish the effects of known *OSMR* and *LIFR* eQTLs in non-atherosclerotic arterial tissues. Thus, we tested whether these eQTLs were associated with *OSMR* and *LIFR* expression in carotid atherosclerotic plaques. Neither variant showed associations with expression of *OSM*, *OSMR*, and *LIFR* (**Supplementary Table 1**).

Intraplaque *OSMR* Expression Is Not Associated With Plaque Vulnerability

As rs13168867 was associated with an increased overall plaque vulnerability, we next investigated if intraplaque *OSMR*



Gene	Variant	NES	Alleles	EAF	Info	β (s.e.)	*p*
OSMR	rs13168867	-0.13	C/T	0.393	0.999	0.118 (0.040)	**3.00×10^{-3}**
LIFR	rs10491509	0.29	A/G	0.351	0.989	0.001 (0.043)	0.981

FIGURE 2 | Association of *OSMR* and *LIFR* variants with overall plaque vulnerability. The plaques were given a vulnerability score ranging from 0 (least vulnerable) to 4 (most vulnerable). The bars represent the proportion of each plaque score per genotype for rs13168867 in the *OSMR* locus **(A)** and rs10491509 in the *LIFR* locus **(B)**. The table shows the results from GTEx Portal (23) where *Gene* is the gene of interest; *Variant* is the eQTL as found in GTEx Portal, and *NES* refers to the normalized effect size on expression from GTEx Portal. The *Alleles* refer to the effect allele and the other allele in both GTEx Portal and the Athero-Express Biobank. The remaining columns in the table are referring to the analysis of these eQTLs with respect to plaque vulnerability in the Athero-Express Biobank Study. Here *EAF* represents effect allele frequency; *Info* refers to the estimated imputation score. The effect size (β) and standard error (s.e.) are relative to the effect allele; *p* indicates the *p*-value of association with plaque vulnerability of the given eQTL; Bold *p*: *p*-value of association surpasses significance threshold ($p < 3.13 \times 10^{-3}$).

expression levels associated with overall plaque vulnerability. Differential expression analyses, comparing the reference score (0) with each increasing vulnerability score (1, 2, 3, or 4) showed no associations between *OSMR* plaque expression levels and plaque vulnerability (**Supplementary Table 2**). Neither did intraplaque *OSM* or *LIFR* expression associate with plaque severity.

Known eQTLs of *OSMR* and *LIFR* Do Not Associate With Cardiovascular Diseases

The Athero-Express comprises patients with advanced stage atherosclerotic plaques. Therefore, we assessed the effect of known *OSMR* and *LIFR* eQTLs on coronary calcification (CAC) as intermediate phenotype of atherosclerotic burden, and primary cardiovascular outcomes as clinical manifestation. We queried summary statistics from GWAS on CAC ($n = 2{,}674$) (45), coronary artery disease (CAD, $n = 336{,}755$) (45, 46), and ischemic stroke subtypes (sample sizes 242,573–522,258) (45–47). Neither eQTL associated with increased CAC burden, or cardiovascular disease susceptibility (**Supplementary Table 3**).

DISCUSSION

We investigated whether common variants associated to gene expression, i.e., eQTLs, near *OSM, OSMR* and *LIFR* affect overall plaque vulnerability and phenotype. We showed that one *cis*-acting eQTL (rs13168867, C/T), of which the C allele associates with reduced *OSMR* expression in non-diseased arterial tissue, significantly associates with increased plaque vulnerability after correction for multiple testing. This suggests that a decrease in *OSMR* expression and therefore possibly a decrease in OSM signaling, increases the chance of developing a vulnerable plaque.

To gain further insight into the role of genetically decreased *OSMR* expression on developing a vulnerable plaque, we examined the effect of rs13168867 on individual plaque characteristics in more detail. The strongest associations were found for rs13168867 with increased intraplaque fat and decreased collagen content, suggesting that reduced OSM signaling results in a larger lipid core and less fibrosis - in line with a more vulnerable plaque phenotype. We previously showed that OSM enhances intercellular adhesion molecule (ICAM)-1 expression on human endothelial cells (8). Reduced *OSMR* expression, which hypothetically results in reduced OSM signaling, may therefore result in reduced ICAM-1 expression. ICAM-1 depletion leads to M1 macrophage polarization (48), which is the pro-inflammatory macrophage subtype that promotes an unstable plaque phenotype (49). Reduced OSM signaling could also explain the decreased collagen content as OSM enhances *in vitro* fibroblast proliferation and collagen formation (20). Moreover, it was previously shown that OSM

TABLE 2 | *OSMR* and *LIFR* variants and their association with plaque phenotypes.

Gene	Variant	NES	Alleles	EAF	Info	Phenotype	β (s.e.)	*p*
OSMR	rs13168867	−0.13	C/T	0.393	0.999	Calcification	0.036 (0.077)	0.637
						Collagen	−0.259 (0.095)	6.22×10^{-3}
						Fat content	0.248 (0.088)	4.66×10^{-3}
						Intraplaque hemorrhage	−0.014 (0.080)	0.862
						Smooth muscle cells	0.001 (0.011)	0.913
						Vessel density	-1.06×10^{-4} (0.004)	0.976
						Macrophages	0.004 (0.015)	0.809
LIFR	rs10491509	0.29	A/G	0.351	0.989	Calcification	0.046 (0.082)	0.577
						Collagen	0.134 (0.104)	0.194
						Fat content	0.086 (0.094)	0.363
						Intraplaque hemorrhage	0.071 (0.086)	0.414
						Smooth muscle cells	−0.003 (0.012)	0.840
						Vessel density	0.002 (0.004)	0.577
						Macrophages	0.015 (0.016)	0.354

For each variant, the association with plaque phenotypes is given. NES: the normalized effect size on expression; Alleles: the effect allele and the other allele, respectively; EAF: effect allele frequency in the Athero-Express Biobank; Info: estimated imputation score in the Athero-Express Biobank; β, effect size; s.e., standard error; p, p-value of association; Adj. p, Bonferroni adjusted p-value. Calcification and collagen were classified as no/minor vs. moderate/heavy, fat content as 10 vs. >10% fat of plaque area, intraplaque hemorrhage was classified as absent vs. present. Smooth muscle cells and macrophages were classified as Box-Cox transformed percentage of plaque area and vessel density as Box-Cox transformed number of vessels/hotspot. None of the p-values surpassed significance threshold (3.13×10^{-3}).

enhances liver fibrosis in mice (50) and that OSM is upregulated in patients with pulmonary fibrosis (51). A reduction in OSM signaling caused by decreased OSMR expression may therefore result in decreased collagen content. Further studies are needed to investigate these hypotheses.

A possible explanation for the lack of associations for the variant (rs10491509) in the *LIFR* locus could be that an increase in *LIFR* expression would not affect OSM signaling as, hypothetically, there might already be a *LIFR* surplus and therefore, an increase in *LIFR* expression will not affect OSM signaling.

Although rs13168867 did associate with plaque vulnerability, no associations were found between rs13168867 and intraplaque *OSMR* expression, intraplaque *OSMR* expression and plaque vulnerability, nor did rs13168867 associate with cardiovascular disease outcomes. Possibly, OSM signaling mainly affects atherogenesis and atherosclerosis development in the initial phases of the disease. Arterial *OSMR* expression is reduced in human atherosclerotic plaques when compared to normal arteries9 and may therefore have bigger effects in the initial phase, when *OSMR* expression is still high. Another possible explanation is that OSM signaling may be overruled by for example, other cytokines in later stages of the disease. Lastly, although coronary thrombosis, and therefore cardiovascular disease, is most often caused by plaque rupture, which is most likely to happen in vulnerable plaques, thrombosis can also be triggered by other processes, including plaque erosion and atrial fibrillation (52, 53). Xie et al. showed that OSM is associated with thrombosis in patients with atrial fibrillation and suggested that OSM exerts thrombogenic effects by increasing tissue factor expression and decreasing the expression of tissue factor pathway inhibitors (53). So, OSM could potentially increase the risk of cardiovascular disease through its thrombogenic effects and at the same time decrease the risk of cardiovascular disease by its atheroprotective effects. Potentially, the seemingly atheroprotective effect of OSM that we described in our current study may be neutralized by the thrombogenic or potential other cardiovascular disease driving effects of OSM.

A limitation of association studies like ours, is that it is challenging to uncover the biological meaning of the discovered associations. It is likely that a reduction in OSMR expression on the arterial wall reduces OSM signaling, but this is difficult to verify. Firstly, OSM signaling is not only dependent on OSMR, but also on the blood OSM levels. If there is no or little OSM present in the blood, there might have been a surplus of OSMR and in this case, there will be no change in OSM signaling. Another possibility is that there is not only a decrease of OSMR expression on the arterial wall, but also a decrease in circulating OSMR levels, which can also bind to OSM and acts as a neutralizer (54), also resulting in no net difference in OSM signaling. Moreover, this study cannot make a distinction between the timing and the duration of OSM signaling, which may differentially affect atherosclerosis development as previous studies have shown that OSM, like IL-6, can act differently in the acute phase than in the chronic phase (8, 9, 55, 56). Finally, we focused on only three genes (*OSM, OSMR* and *LIFR*), while atherosclerosis is a multifactorial disease. Although studies like ours can be very insightful to better understand the disease, single variants seldomly show big correlations with disease outcome.

Compared to genome-wide association studies that include thousands of individuals, the Athero-Express Biobank Study is relatively small (n=1,443), and, given its design, finite in size. However, it is well suited to examine the effect of common disease-associated genetic variation on plaque morphology and characteristics. Indeed, we estimated the power at ±75% given a MAF=0.40 (approximately the frequency of rs13168867) and relative risk = 1.28 (http://csg.sph.umich.edu/abecasis/gas_power_calculator/).

Recent developments in single-cell expression analyses might extend on the present study by investigating which cell types, that are present in the plaque, most abundantly express *OSM, OSMR* and *LIFR*. Furthermore, it would be interesting to investigate if the *OSMR/LIFR* expression ratio correlates with plaque vulnerability and if this ratio might be a predictor of plaque vulnerability.

CONCLUSION

Based on this work we conclude that the variant rs13168867 in the *OSMR* locus is associated with increased plaque vulnerability, but not with coronary calcification or cardiovascular disease susceptibility. Given the multiple testing burden for individual plaque characteristics, it remains unclear through which precise biological mechanisms OSM signaling exerts its effects on plaque morphology, although our data point toward lipid metabolism and extracellular matrix remodeling. However, the OSM-OSMR/LIFR pathway is unlikely to be causally involved in lifetime cardiovascular disease susceptibility as none of the investigated eQTLs associated with cardiovascular diseases.

DATA AVAILABILITY STATEMENT

Data can be accessed at https://dataverse.nl/ using accession number 0RB5IZ.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Medisch Ethische Toetsingscommissie (METC) Utrecht. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

DvK: conceptualization, formal analysis, and writing - original draft. IvK: data curation. AB: conceptualization, formal analysis, and writing – review & editing. HP and AvG: writing – review & editing. GdB and DT: conceptualization. DT and GP: conceptualization and writing – review & editing. SvL: conceptualization, formal analysis, writing – original draft and review & editing. All authors contributed to the article and approved the submitted version.

FUNDING

SvL was funded through grants from the Netherlands CardioVascular Research Initiative of the Netherlands Heart Foundation [CVON 2011/B019 and CVON 2017-20: Generating the best evidence-based pharmaceutical targets for atherosclerosis (GENIUS I&II)]. This work was supported by ERA-CVD, grant number: 01KL1802. FA was supported by UCL Hospitals NIHR Biomedical Research Center. DvK, HP, and DT were funded through the FP7 EU project CarTarDis (FP7/2007-2013) under grant agreement 602936. AB was funded through the Taxinomisis grant, part of the European Union's Horizon 2020 research and innovation program (No 755320). HP received funding from the TNO research program Preventive Health Technologies. The funding sources were not involved in study design, collection, analysis and interpretation of data, writing of the report and in the decision to submit the article for publication.

ACKNOWLEDGMENTS

We would like to thank Dr. Jessica van Setten and acknowledge her for imputing our datasets using an in-house developed imputation pipeline. Evelyn Velema and Petra Homoet-Van der Kraak are acknowledged for the immunohistochemical stainings. We also acknowledge the support from the Netherlands CardioVascular Research Initiative from the Dutch Heart Foundation, Dutch Federation of University Medical Centers, the Netherlands Organization for Health Research and Development and the Royal Netherlands Academy of Sciences (GENIUS I & II, CVON2011-19) and the TNO research program Preventive Health Technologies.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2021.658915/full#supplementary-material

REFERENCES

1. Moodie DS. The global burden of cardiovascular disease. *Congenit Heart Dis.* (2016) 11:213. doi: 10.1111/chd.12383
2. Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med.* (2005) 352:1685–95. doi: 10.1056/NEJMra043430
3. Sprague AH, Khalil RA. Inflammatory cytokines in vascular dysfunction and vascular disease. *Biochem Pharmacol.* (2009) 78:539–52. doi: 10.1016/j.bcp.2009.04.029
4. Vasse M, Pourtau J, Trochon V, Muraine M, Vannier J-P, Lu H, et al. Oncostatin M induces angiogenesis *in vitro* and *in vivo*. *Arterioscler Thromb Vasc Biol.* (1999) 19:1835–42. doi: 10.1161/01.ATV.19.8.1835
5. Nagata T, Kai H, Shibata R, Koga M, Yoshimura A, Imaizumi T. Oncostatin M, an interleukin-6 family cytokine, upregulates matrix metalloproteinase-9 through the mitogen-activated protein kinase kinase-extracellular signal-regulated kinase pathway in cultured smooth muscle cells. *Arterioscler Thromb Vasc Biol.* (2003) 23:588–93. doi: 10.1161/01.ATV.0000060891.31516.24
6. Albasanz-Puig A, Murray J, Preusch M, Coan D, Namekata M, Patel Y, et al. Oncostatin M is expressed in atherosclerotic lesions: a role for Oncostatin M in the pathogenesis of atherosclerosis. *Atherosclerosis.* (2011) 216:292–8. doi: 10.1016/j.atherosclerosis.2011.02.003
7. Zhang X, Li J, Qin J-J, Cheng W-L, Zhu X, Gong F-H, et al. Oncostatin M receptor β deficiency attenuates atherogenesis by inhibiting

JAK2/STAT3 signaling in macrophages. *J Lipid Res.* (2017) 58:895–906. doi: 10.1194/jlr.M074112
8. Van Keulen D, Pouwer MG, Pasterkamp G, Van Gool AJ, Sollewijn Gelpke MD, Princen HMG, et al. Inflammatory cytokine oncostatin M induces endothelial activation in macro-and microvascular endothelial cells and in APOE*3Leiden.CETP mice. *PLoS ONE.* (2018) 13:e0204911. doi: 10.1016/j.atherosclerosissup.2018.04.056
9. van Keulen D, Pouwer MG, Emilsson V, Matic LP, Pieterman EJ, Hedin U, et al. Oncostatin M reduces atherosclerosis development in APOE3Leiden.CETP mice and is associated with increased survival probability in humans. *PLoS One.* (2019) 14:e0221477. doi: 10.1371/journal.pone.0221477
10. Brown TJ, Lioubin MN, Marquardt H. Purification and characterization of cytostatic lymphokines produced by activated human T lymphocytes. Synergistic antiproliferative activity of transforming growth factor beta 1, interferon-gamma, and oncostatin M for human melanoma cells. *J Immunol.* (1987) 139:2977–83.
11. Kastl SP, Speidl WS, Kaun C, Katsaros KM, Rega G, Afonyushkin T, et al. In human macrophages the complement component C5a induces the expression of oncostatin M via AP-1 activation. *Arterioscler Thromb Vasc Biol.* (2008) 28:498–503. doi: 10.1161/ATVBAHA.107.160580
12. Grenier A, Dehoux M, Boutten A, Arce-Vicioso M, Durand G, Gougerot-Pocidalo M-A, et al. Oncostatin M production and regulation by human polymorphonuclear neutrophils. *Blood.* (1999) 93:1413–21. doi: 10.1182/blood.V93.4.1413
13. Tanaka M, Miyajima A. Oncostatin M, a multifunctional cytokine. *Rev Physiol Biochem Pharmacol.* (2003) 149:39–52. doi: 10.1007/s10254-003-0013-1
14. Dey G, Radhakrishnan A, Syed N, Thomas JK, Nadig A, Srikumar K, et al. Signaling network of Oncostatin M pathway. *J Cell Commun Signal.* (2013) 7:103–8. doi: 10.1007/s12079-012-0186-y
15. Mosley B, De Imus C, Friend D, Boiani N, Thoma B, Park LS, et al. Dual oncostatin M (OSM) receptors. Cloning and characterization of an alternative signaling subunit conferring OSM-specific receptor activation. *J Biol Chem.* (1996) 271:32635–43. doi: 10.1074/jbc.271.51.32635
16. Takata F, Sumi N, Nishioku T, Harada E, Wakigawa T, Shuto H, et al. Oncostatin M induces functional and structural impairment of blood–brain barriers comprised of rat brain capillary endothelial cells. *Neurosci Lett.* (2008) 441:163–6. doi: 10.1016/j.neulet.2008.06.030
17. Guihard P, Danger Y, Brounais B, David E, Brion R, Delecrin J, et al. Induction of osteogenesis in mesenchymal stem cells by activated monocytes/macrophages depends on oncostatin M signaling. *Stem Cells.* (2012) 30:762–72. doi: 10.1002/stem.1040
18. Virmani R, Kolodgie FD, Burke AP, Finn A V, Gold HK, Tulenko TN, et al. Atherosclerotic plaque progression and vulnerability to rupture angiogenesis as a source of intraplaque hemorrhage plaque rupture is the dominant cause of acute coronary thrombosis. *Arterioscler Thromb Vasc Biol.* (2005) 25:2054–61. doi: 10.1161/01.ATV.0000178991.71605.18
19. Hutcheson JD, Maldonado N, Aikawa E. Small entities with large impact: microcalcifications and atherosclerotic plaque vulnerability. *Curr Opin Lipidol.* (2014) 25:327–32. doi: 10.1097/MOL.0000000000000105
20. Scaffidi AK, Mutsaers SE, Moodley YP, McAnulty RJ, Laurent GJ, Thompson PJ, et al. Oncostatin M stimulates proliferation, induces collagen production and inhibits apoptosis of human lung fibroblasts. *Br J Pharmacol.* (2002) 136:793–801. doi: 10.1038/sj.bjp.0704769
21. Shrivastava R, Asif M, Singh V, Dubey P, Ahmad Malik S, Lone M-U-D, et al. M2 polarization of macrophages by Oncostatin M in hypoxic tumor microenvironment is mediated by mTORC2 and promotes tumor growth and metastasis. *Cytokine.* (2018) 118:130–43. doi: 10.1016/j.cyto.2018.03.032
22. Van Der Wal AC, Becker AE. Atherosclerotic plaque rupture - pathologic basis of plaque stability and instability. *Cardiovasc Res.* (1999) 41:334–44. doi: 10.1016/S0008-6363(98)00276-4
23. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The genotype-tissue expression (GTEx) project. *Nat Genet.* (2013) 45:580–5. doi: 10.1038/ng.2653
24. Grundberg E, Small KS, Hedman ÅK, Nica AC, Buil A, Keildson S, et al. Mapping cis-and trans-regulatory effects across multiple tissues in twins. *Nat Genet.* (2012) 44:1084–9. doi: 10.1038/ng.2394
25. Verhoeven B, Hellings WE, Moll FL, De Vries JP, De Kleijn DPV, De Bruin P, et al. Carotid atherosclerotic plaques in patients with transient ischemic attacks and stroke have unstable characteristics compared with plaques in asymptomatic and amaurosis fugax patients. *J Vasc Surg.* (2005) 42:1075–81. doi: 10.1016/j.jvs.2005.08.009
26. Van Der Laan SW, Foroughi Asl H, van den Borne P, van Setten J, van der Perk MEM, van de Weg SM, et al. Variants in ALOX5, ALOX5AP and LTA4H are not associated with atherosclerotic plaque phenotypes: The Athero-Express Genomics Study. *Atherosclerosis.* (2015) 239:528–38. doi: 10.1016/j.atherosclerosis.2015.01.018
27. O'Connell J, Gurdasani D, Delaneau O, Pirastu N, Ulivi S, Cocca M, et al. A general approach for haplotype phasing across the full spectrum of relatedness. *PLoS Genet.* (2014) 10:1004234. doi: 10.1371/journal.pgen.1004234
28. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, et al. A global reference for human genetic variation. *Nature.* (2015) 526:68–74. doi: 10.1038/nature15393
29. Francioli LC, Menelaou A, Pulit SL, Van Dijk F, Palamara PF, Elbers CC, et al. Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nat Genet.* (2014) 46:818–25. doi: 10.1038/ng.3021
30. Aguet F, Barbeira AN, Bonazzola R, Brown A, Castel SE, Jo B, et al. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science (80-).* (2020) 369:1318–30. doi: 10.1126/science.aaz1776
31. Verhoeven BAN, Velema E, Schoneveld AH, De Vries JPPM, De Bruin P, Seldenrijk CA, et al. Athero-express: differential atherosclerotic plaque expression of mRNA and protein in relation to cardiovascular events and patient characteristics. Rationale and design. *Eur J Epidemiol.* (2004) 19:1127–33. doi: 10.1007/s10564-004-2304-6
32. Van Lammeren GW, Den Ruijter HM, Vrijenhoek JEP, Van Der Laan SW, Velema E, De Vries JPPM, et al. Time-dependent changes in atherosclerotic plaque composition in patients undergoing carotid surgery. *Circulation.* (2014) 129:2269–76. doi: 10.1161/CIRCULATIONAHA.113.007603
33. Hellings WE, Pasterkamp G, Vollebregt A, Seldenrijk CA, De Vries JPPM, Velema E, et al. Intraobserver and interobserver variability and spatial differences in histologic examination of carotid endarterectomy specimens. *J Vasc Surg.* (2007) 46:1147–54. doi: 10.1016/j.jvs.2007.08.018
34. Hashimshony T, Wagner F, Sher N, Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep.* (2012) 2:666–73. doi: 10.1016/j.celrep.2012.08.003
35. Hashimshony T, Senderovich N, Avital G, Klochendler A, de Leeuw Y, Anavy L, et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol.* (2016) 17:77. doi: 10.1186/s13059-016-0938-8
36. Ferraz MAMM, Rho HS, Hemerich D, Henning HHW, van Tol HTA, Hölker M, et al. An oviduct-on-a-chip provides an enhanced in vitro environment for zygote genome reprogramming. *Nat Commun.* (2018) 9:1–14. doi: 10.1038/s41467-018-07119-8
37. Li H, Durbin R. Fast and accurate short read alignment with burrows–wheeler transform. *Bioinformatics.* (2009) 25:1754–60. doi: 10.1093/bioinformatics/btp324
38. Morgan M, Obenchain V, Hester J, Pagès H. *Summarized Experiment: Summarized Experiment container*. R package version 1.20.0 (2020). Available online at: https://bioconductor.org/packages/SummarizedExperiment
39. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* (2014) 15:550. doi: 10.1186/s13059-014-0550-8
40. Box GEP, Cox DR. An analysis of transformations. *J. R. Stat. Soc.* (1964) 26:211–43. doi: 10.1111/j.2517-6161.1964.tb00553.x
41. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet.* (2007) 39:906–13. doi: 10.1038/ng2088
42. Thoma B, Bird TA, Friend DJ, Gearing DP, Dower SK. Oncostatin M and leukemia inhibitory factor trigger overlapping and different signals through partially shared receptor complexes. *J Biol Chem.* (1994) 269:6215–22. doi: 10.1016/S0021-9258(17)37590-7
43. Hermanns H, Radtke S, Haan c, Schmitz-Van de Leur H, Tavernier J, Heinrich P, et al. Contributions of leukemia inhibitory factor receptor and oncostatin M

receptor to signal transduction in heterodimeric complexes with glycoprotein 130 - PubMed. *J Immunol.* (1999) 163:6651–8.
44. Rolfe B, Stamatiou S, World C, Brown L, Thomas A, Bingley J, et al. Leukaemia inhibitory factor retards the progression of atherosclerosis. *Cardiovasc Res.* (2003) 58:222–30. doi: 10.1016/S0008-6363(02)00832-5
45. van Setten J, Isgum I, Smolonska J, Ripke S, de Jong PA, Oudkerk M, et al. Genome-wide association study of coronary and aortic calcification implicates risk loci for coronary artery disease and myocardial infarction. *Atherosclerosis.* (2013) 228:400–5. doi: 10.1016/j.atherosclerosis.2013.02.039
46. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat Genet.* (2017) 49:1385–91. doi: 10.1038/ng.3913
47. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Publisher correction: multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet.* (2019) 51:1192–3.doi: 10.1038/s41588-018-0058-3
48. Gu W, Yao L, Li L, Zhang J, Place AT, Minshall RD, et al. ICAM-1 regulates macrophage polarization by suppressing MCP- 1 expression via miR-124 upregulation. *Oncotarget.* (2017) 8:111882–901. doi: 10.18632/oncotarget.22948
49. Barrett TJ. Macrophages in atherosclerosis regression. *Arterioscler Thromb Vasc Biol.* (2020) 40:20–33. doi: 10.1161/ATVBAHA.119.312802
50. Matsuda M, Tsurusaki S, Miyata N, Saijou E, Okochi H, Miyajima A, et al. Oncostatin M causes liver fibrosis by regulating cooperation between hepatic stellate cells and macrophages in mice. *Hepatology.* (2018) 67:296–312. doi: 10.1002/hep.29421
51. Mozaffarian A, Brewer AW, Trueblood ES, Luzina IG, Todd NW, Atamas SP, et al. Mechanisms of Oncostatin M-induced pulmonary inflammation and fibrosis. *J Immunol.* (2008) 181:7243–53. doi: 10.4049/jimmunol.181.10.7243
52. Bentzon JF, Otsuka F, Virmani R, Falk E. Mechanisms of plaque formation and rupture. *Circ Res.* (2014) 114:1852–66. doi: 10.1161/CIRCRESAHA.114.302721
53. Xie J, Zhu S, Dai Q, Lu J, Chen J, Li G, et al. Oncostatin M was associated with thrombosis in patients with atrial fibrillation. *Medicine.* (2017) 96:e6806. doi: 10.1097/MD.0000000000006806
54. Diveu C, Venereau E, Froger J, Ravon E, Grimaud L, Rousseau F, et al. Molecular and functional characterization of a soluble form of Oncostatin M/interleukin-31 shared receptor. *J Biol Chem.* (2006) 281:36673–382. doi: 10.1074/jbc.M607005200
55. Low ASL, Symmons DPM, Lunt M, Mercer LK, Gale CP, Watson KD, et al. Relationship between exposure to tumour necrosis factor inhibitor therapy and incidence and severity of myocardial infarction in patients with rheumatoid arthritis. *Ann Rheum Dis.* (2017) 76:654–60. doi: 10.1136/annrheumdis-2016-209784
56. Watson C, Whittaker S, Smith N, Vora AJ, Dumonde DC, Brown KA. IL-6 acts on endothelial cells to preferentially increase their adherence for lymphocytes. *Clin Exp Immunol.* (1996) 105:112–9. doi: 10.1046/j.1365-2249.1996.d01-717.x

Conflict of Interest: DvK is employed by Quorics B.V., and DT is employed by SkylineDx B.V and Quorics B.V. Quorics B.V. and SkylineDx B.V. had no part whatsoever in the conception, design, or execution of this study, nor the preparation and contents of this manuscript.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 23 June 2021
doi: 10.3389/fcvm.2021.639148

Cardiomyocyte Proliferative Capacity Is Restricted in Mice With *Lmna* Mutation

Kenji Onoue[1,2], Hiroko Wakimoto[2], Jiangming Jiang[2], Michael Parfenov[2], Steven DePalma[2], David Conner[2], Joshua Gorham[2], David McKean[2], Jonathan G. Seidman[2], Christine E. Seidman[2,3] and Yoshihiko Saito[1]*

[1] Department of Cardiovascular Medicine, Nara Medical University, Kashihara, Japan, [2] Department of Genetics, Harvard Medical School, Boston, MA, United States, [3] Division of Cardiovascular Medicine, Howard Hughes Medical Institute, Brigham and Women's Hospital, Boston, MA, United States

OPEN ACCESS

Edited by:
Kaoru Ito,
RIKEN Yokohama, Japan

Reviewed by:
Toru Kubo,
Kochi University, Japan
Blanca Morales Rodriguez,
Université Paris-Sorbonne, France

***Correspondence:**
Kenji Onoue
konoue@naramed-u.ac.jp

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: 08 December 2020
Accepted: 21 May 2021
Published: 23 June 2021

Citation:
Onoue K, Wakimoto H, Jiang J, Parfenov M, DePalma S, Conner D, Gorham J, McKean D, Seidman JG, Seidman CE and Saito Y (2021) Cardiomyocyte Proliferative Capacity Is Restricted in Mice With Lmna Mutation. Front. Cardiovasc. Med. 8:639148. doi: 10.3389/fcvm.2021.639148

LMNA is one of the leading causative genes of genetically inherited dilated cardiomyopathy (DCM). Unlike most DCM-causative genes, which encode sarcomeric or sarcomere-related proteins, *LMNA* encodes nuclear envelope proteins, lamin A and C, and does not directly associate with contractile function. However, a mutation in this gene could lead to the development of DCM. The molecular mechanism of how *LMNA* mutation contributes to DCM development remains largely unclear and yet to be elucidated. The objective of this study was to clarify the mechanism of developing DCM caused by *LMNA* mutation.

Methods and Results: We assessed cardiomyocyte phenotypes and characteristics focusing on cell cycle activity in mice with *Lmna* mutation. Both cell number and cell size were reduced, cardiomyocytes were immature, and cell cycle activity was retarded in *Lmna* mutant mice at both 5 weeks and 2 years of age. RNA-sequencing and pathway analysis revealed "proliferation of cells" had the most substantial impact on *Lmna* mutant mice. *Cdkn1a,* which encodes the cell cycle regulating protein p21, was strongly upregulated in *Lmna* mutants, and upregulation of p21 was confirmed by Western blot and immunostaining. DNA damage, which is known to upregulate *Cdkn1a*, was more abundantly detected in *Lmna* mutant mice. To assess the proliferative capacity of cardiomyocytes, the apex of the neonate mouse heart was resected, and recovery from the insult was observed. A restricted cardiomyocyte proliferating capacity after resecting the apex of the heart was observed in *Lmna* mutant mice.

Conclusions: Our results strongly suggest that loss of lamin function contributes to impaired cell proliferation through cell cycle defects. The inadequate inborn or responsive cell proliferation capacity plays an essential role in developing DCM with *LMNA* mutation.

Keywords: dilated cardiomyopathy, lamin A/C, cell cycle, p21, repressed proliferating capacity

INTRODUCTION

Dilated cardiomyopathy (DCM), a significant cause of heart failure, is genetically inherited in 30 to 50% of cases (1, 2). *LMNA* is one of the leading causative genes of genetically inherited DCM as well as *TTN* or *MYH7* (2–4). Most DCM-causative genes (e.g., *TTN, MYH7,* and *TNNT2*) encode sarcomeric proteins or sarcomere-related proteins and are directly involved in the generation or transmission of the contractile force of the cardiomyocyte. Unlike these sarcomere-related genes, *LMNA* encodes nuclear envelope proteins, lamin A and C, and does not directly associate with contractile function. However, a mutation in this gene could lead to the development of DCM (4, 5), which in many cases is also accompanied by defects in the conduction system (2, 6) and poor prognosis (7). Moreover, mutations in *LMNA* are also known to cause a range of diseases, including myopathies and neuropathies such as limb-girdle muscular dystrophy (8), Emery-Dreifuss muscular dystrophy (9), Charcot-Marie-Tooth neuropathy (10), and Hutchinson-Gilford progeria (11).

Lamin proteins are structural proteins of the inner nuclear membrane, and an *LMNA* mutation is reported as the cause of morphological changes in nuclei such as flattening and bleb formation (12–15). Nikolova et al. reported the cardiac phenotypes resulting from this mutation using *Lmna* homozygous knockout ($Lmna^{-/-}$) mice, which presented left ventricular (LV) dilation, reduced systolic function, and died around 6 weeks. $Lmna^{-/-}$ myocytes showed altered nuclear shapes, decreased size, impaired contractility, and diminished Ca^{2+} binding affinity to myofilament (13). Morales Rodriguez et al. reported the altered calcium cycling could be related to LV dysfunction (16). Wolf et al. reported the impaired contractility of myocytes accompanied by atrioventricular conduction defects in *Lmna* heterozygous knockout ($Lmna^{+/-}$) mice, mimicking human patients' phenotypes (17). Macquart et al. proved that altered distribution of a major gap junction protein Cx43 contributes to conduction defects using *Lmna* mutant mice (18). Thus, the molecular mechanism of how *LMNA* mutation contributes to DCM development is gradually elucidated, however, there remain unclear mechanisms to be elucidated.

Recently, cell cycle activity is reported to be related to the phenotype of the mammalian heart (19, 20). The development of the mammalian heart is characterized by cardiomyocyte proliferation and hypertrophy (21). In most mammalian hearts, cardiomyocytes proliferate before birth, followed by exiting the cell cycle by the postnatal change of nutrients, increased hemodynamic stress, and increased oxygen concentration (22, 23). Before exiting the cell cycle, murine neonatal cardiomyocytes undergo karyokinesis without cytokinesis from around 4 days after birth (21, 24), making the number of nuclei double per cell. By 12 days after birth, approximately 90% of the cardiomyocytes are binuclear, which is the physiological termination of the cell cycle. The heart grows mainly with hypertrophy of the cardiomyocyte after this period. When this physiological cell cycle activity is impaired by some reasons such as preterm birth, cardiac morphological change is observed (25, 26), and sometimes cardiomyopathy develops (27). On the contrary, under the condition in which the cell cycle is activated, the heart proliferates too much and sometimes develops hypertrophic phenotype as previously reported in the mouse with *MYBPC3* mutation (28) or hypoxic condition (19). Thus, the cell cycle activity is essential in developing the heart and could influence cardiac morphology. This cell cycle activity is reported to be retarded in lamin A/C-depleted cells (29) and skeletal muscle in $Lmna^{-/-}$ mice (30). However, the cell cycle activity in the heart under the physiological condition, stress, and its relation to the development of DCM in *Lmna* mutant mice has not been elucidated to date.

We, therefore, studied the cell cycle activity in the mouse model with *Lmna* mutation and pursued the mechanism of developing DCM caused by *LMNA* mutation.

MATERIALS AND METHODS

All experimental protocols for the animal models were approved by the Animal Care and Use Committee of Harvard Medical School (#2530) and Nara Medical University (#11251, #11355).

Mice

Lmna mutant mice were generated by the deletion of exon 8 to 11 of *Lmna* as described before (12). In brief, the targeting vector, removing exon 8 to 11 of *Lmna*, was electroporated into W9.5 embryonic stem cells. Two clones were injected into C57BL/6 blastocysts, and chimeras were bred to produce germline offspring. This line was backcrossed to 129S6/SvEvTac strain, obtained from Taconic Biosciences Inc. (Rensselaer, NY, USA), for more than 10 generations. Mouse genotypes were determined by PCR amplification of tail genomic DNA.

Echocardiographic Studies

Mice were anesthetized with 1–1.5% inhalational isoflurane. Each limb was placed on the ECG leads on a Vevo Mouse Handling Table (FujiFilm VisualSonics Inc., Toronto, ON, Canada), maintaining the body temperature at 37°C during the study. Transthoracic echocardiography was performed using a Vevo 770 High-Resolution *In vivo* Micro-Imaging System and RMV 707B scan-head (FujiFilm VisualSonics Inc.) with heart rate at 450–550 beats per minute. The images were acquired as 2-dimensional mode (left parasternal long and short axes) and M-mode (left parasternal short axis). Measurements averaged from 3 consecutive heartbeats of M-mode tracings were used for LV wall thickness, LV end-diastolic, and end-systolic diameters. All echocardiographic measurements were done blinded to mouse genotype.

EdU Cell Proliferation Assay

The thymidine analog 5-ethynyl-2′-deoxyuridine (EdU) cell proliferation assay was performed using Click-iT® EdU Alexa Fluor® 594 Imaging Kit (Life Technologies, Carlsbad, CA, USA). EdU is incorporated into DNA during DNA synthesis, the same as 5-bromo-2′-deoxyuridine (BrdU). EdU (5 μg/g body weight) was injected intraperitoneally for 5 consecutive days starting from the day of birth. Mice were sacrificed at 3 weeks of age and processed for EdU detection following the company's protocol.

Apical Resection

LV apical resection procedure was performed as described before (31). In brief, mice at the age of 2 days after birth were anesthetized on ice to induce sedation. The thoracic cavity was opened at the fourth intercostal space, and the exposed LV apex was then resected. Following surgery, mice were warmed up to 37°C and were monitored for viability. Sham-operated mice underwent identical procedures without LV apical resection. Five μg/g body weight of EdU was administered subcutaneously after the surgery for 10 days every 2 days. Mice were euthanized at 3 weeks of age under 4–5% continuous inhalational isoflurane by exsanguination to assess cardiomyocyte proliferation.

Histological Examination

To quantitate the number of nuclei per myocyte, isolated myocytes were labeled with 4′,6-Diamidine-2′-phenylindole dihydrochloride (DAPI, Sigma-Aldrich, St. Louis, MO, USA) and counterstained with anti-troponin I antibody (Ab) (1:200 dilution, Abcam, Cambridge, UK). To quantify cell cycle activity, isolated myocytes were labeled with anti-phospho-histone H_3 Ab (Ser10, 1:100 dilution, Millipore, Billerica, MA, USA). Four percentage PFA-fixed paraffin-embedded hearts were cross-sectioned in 4 μm thickness and stained with hematoxylin and eosin and Masson's trichrome staining. The tissue sections were stained with Alexa Fluor® 488 or 594 conjugate of wheat germ agglutinin (WGA, Life technologies) for cell number counting and cell size measurement. To quantify the number of myocytes in tissue sections, cell counts were obtained from 10 different layers of the heart. We utilized Image J (NIH software) for counting cell numbers and measurement. Immunostaining for p21 was performed with anti-p21 Ab (ab2961, 1:40 dilution, Abcam). We applied anti-phospho-histone $H_2A.X$ Ab (pH_2AX; Ser139, 1:480 dilution, Cell Signaling Technology, Danvers, MA, USA) to detect double-strand DNA break. EdU detection was performed in either 4% PFA-fixed frozen tissues embedded in optimal cutting temperature (OCT) compound (Sakura Finetek, Tokyo, Japan) or 10% Formalin-fixed paraffin-embedded tissue. Counterstaining was done with DAPI. Immunostaining positive cells or the number of nuclei were counted independently by two authors blinded to specimens' background.

Protein Analysis

Western blot analyses were performed as in the conventional method. Briefly, total protein was extracted from frozen tissue, separated by SDS-PAGE, and hybridized with primary anti-p53 Ab (sc-100, 1:200 dilution, Santa Cruz Biotechnology, Dallas, TX, USA), anti-p21 Ab (sc6246, 1:200 dilution, Santa Cruz Biotechnology), or anti-β actin Ab (ab8227, 1:1000 dilution, Abcam). Hybridized signals were quantified by Image J and normalized to β actin.

RNA-Sequence and Data Analysis

Total RNA was extracted from flash-frozen heart specimens with Trizol (Life Technologies), followed by mRNA purification by poly-A selection and cDNA synthesis using standard protocols to construct RNA-seq libraries as reported previously (32). The index was added to cDNA to distinguish samples. 20 nmol of each library was analyzed using the next-generation sequencing platform, Illumina HiSeq 2000 (Illumina, San Diego, CA, USA). Paired-end, 50 bp reads were aligned to mouse genome mm9 using TopHat ver1.4 (http://tophat.cbcb.umd.edu/). Gene expression profiles were generated as described before using a Bayesian *P*-value (33). The number of reads was normalized to total aligned reads on gene loci per 1 million reads. RNA-seq results were analyzed by IPA (http://www.ingenuity.com). Up- or down-regulated gene was defined as having more than or less than 1.5-fold transcriptional level compared to age-matched wild type (wt). Genes with expression values <1 read per 1 million reads (as evaluated in the wt sample) were excluded from the analysis.

Statistics

Continuous data are expressed as mean ± SD. The significance of differences between two groups was determined using the Student's T-test, and that between more than 3 groups was determined with 1-way ANOVA. *Post-hoc* pairwise comparisons were performed with the Tukey-Kramer test. $P < 0.05$ were considered statistically significant. The *P-value* associated with functional analysis in IPA was calculated with the right-tailed Fisher's exact test.

RESULTS

Small Heart Phenotype in *Lmna*$^{-/-}$ Mouse Is Caused by the Reduction of Both the Number and the Size of Cardiomyocytes

As reported previously (34), 2-year old *Lmna*$^{+/-}$ mice were observed to have dilated LV in both systolic and diastolic phases and reduced LV contraction, whereas 5-week and 1-year-old *Lmna*$^{+/-}$ mice did not present a DCM phenotype yet (**Table 1**). Overall, *Lmna*$^{-/-}$ mice were smaller in body size than their age-matched wt and *Lmna*$^{+/-}$ littermates. LV chamber size was enlarged in both systolic and diastolic phases after body surface area correction, and LV contractile function was reduced in *Lmna*$^{-/-}$ mice compared to their controls. Masson's trichrome staining showed massive fibrosis in *Lmna*$^{-/-}$ mice at 5 weeks of age (**Supplementary Figure 1**). Interestingly, those cardiac phenotypes such as enlarged LV chamber, reduced contractile function, and severe fibrosis were comparable to those of *Lmna*$^{+/-}$ mice observed at 2 years of age (**Table 1**, **Supplementary Figure 1**), which suggested that the cardiac phenotypes developed more rapidly in *Lmna*$^{-/-}$ mice than in *Lmna*$^{+/-}$ mice, sharing similar pathology at different ages. Thus, the *Lmna*$^{-/-}$ mice offer a model for studying DCM that rapidly develops within the lifespan of the mice. Identifying the molecular determinants of DCM and morphological characteristics in these mice allows for opportunities to understand and potentially pharmacologically treat the pathology in both *Lmna*$^{-/-}$ and *Lmna*$^{+/-}$ mice, with extensions to understanding how DCM develops in humans with *LMNA* mutations, which typically appear as autosomal dominant heterozygous mutations.

The total number of myocytes was counted in 3-week-old wt and *Lmna*$^{-/-}$ mice. Myocyte counts from tissue sections of

TABLE 1 | Physical and echocardiographic characteristics of the mice used in this study.

Age group	0 year (5 weeks old)			1 year (43–53 weeks old)		2 years (86–100 weeks old)	
Genotype	wt ($n = 6$)	*Lmna*$^{+/-}$ ($n = 8$)	*Lmna*$^{-/-}$ ($n = 5$)	wt ($n = 5$)	*Lmna*$^{+/-}$ ($n = 9$)	wt ($n = 6$)	*Lmna*$^{+/-}$ ($n = 10$)
BW (g)	18.1 ± 2.5	19.7 ± 2.7	8.5 ± 1.7†	34.5 ± 2.0	33.5 ± 3.3	36.5 ± 7.1	34.4 ± 7.8
HW/BW (mg/g)	5.8 ± 0.4	5.6 ± 0.6	6.5 ± 1.1	5.0 ± 0.7	4.6 ± 0.4	5.7 ± 1.0	8.9 ± 2.9†
IVS (mm)	0.64 ± 0.03	0.63 ± 0.05	0.45 ± 0.08†	0.76 ± 0.05	0.73 ± 0.05	0.82 ± 0.14	0.74 ± 0.08
LVDd (mm)	3.03 ± 0.28	3.25 ± 0.22	2.62 ± 0.27*	3.60 ± 0.32	3.54 ± 0.33	4.10 ± 0.26	4.81 ± 0.71
LVDd/BSA (mm/m^2)	924 ± 99	934 ± 66	1,331 ± 163†	709 ± 60	713 ± 77	778 ± 115	977 ± 234
LVDs (mm)	1.82 ± 0.18	1.82 ± 0.36	1.81 ± 0.43	2.33 ± 0.38	2.09 ± 0.25	2.78 ± 0.46	3.99 ± 0.80†
LVDs/BSA (mm/m^2)	555 ± 67	520 ± 81	914 ± 182†	458 ± 61	421 ± 62	524 ± 89	817 ± 251†
PW (mm)	0.62 ± 0.06	0.64 ± 0.06	0.48 ± 0.03†	0.88 ± 0.05	0.79 ± 0.10	1.03 ± 0.11	0.77 ± 0.11†
FS (%)	40.0 ± 3.8	44.1 ± 9.1	31.6 ± 9.0	35.2 ± 9.1	40.9 ± 7.0	32.3 ± 9.7	17.5 ± 7.0†

*BSA, body surface area; BW, body weight; FS, fractional shortening; HW, heart weight; IVS, interventricular septum; LVDd, diastolic left ventricular diameter; LVDs, systolic left ventricular diameter; PW, posterior wall; FS was calculated by [(LVDd-LVDs)/LVDd] × 100. LV mass was calculated by (IVS + PW + LVDd)^3-LVDd^3. *$p < 0.05$ vs. wt, †$p < 0.01$ vs. wt, Student's 2-tailed T-test.*

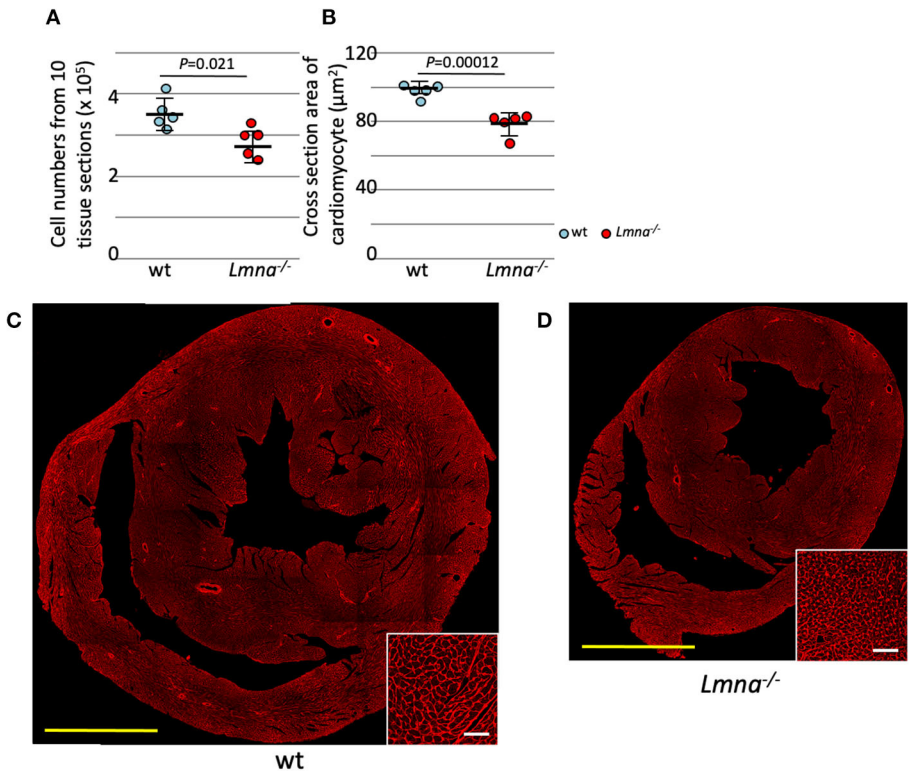


FIGURE 1 | Cardiomyocyte cell number and size in wild type (wt) and *Lmna*$^{-/-}$ mice. Both the cell number and size of cardiomyocytes were significantly reduced in *Lmna*$^{-/-}$ mice (3 weeks old). **(A)** Total counts of cardiomyocytes in 10 tissue layers of the heart ($n = 5$ in each group). **(B)** Cross-section area of cardiomyocytes measured in a tissue section ($n = 5$ in each group). **(C,D)** Tissue sections stained with wheat germ agglutinin in wt **(C)** and *Lmna*$^{-/-}$ **(D)** mice. The open circle represents wt, and the solid circle represents *Lmna*$^{-/-}$ mice. The significance of differences between two groups was determined using the Student's 2 tailed *T*-test. Scale bar: 1 mm in lower magnification and 50 μm in higher magnification in inset.

10 levels from apex to the base sequentially with the distance of 100 μm between each level (2.85 ± 0.37 × 10^5 in *Lmna*$^{-/-}$ and 3.54 ± 0.39 × 10^5 in wt, $P = 0.021$; **Figure 1A**) showed a significant decrease in the number of *Lmna*$^{-/-}$ myocytes. Moreover, myocyte cell size in tissue sections was significantly reduced in *Lmna*$^{-/-}$ hearts (75.1 ± 6.5 μm^2 in *Lmna*$^{-/-}$ and

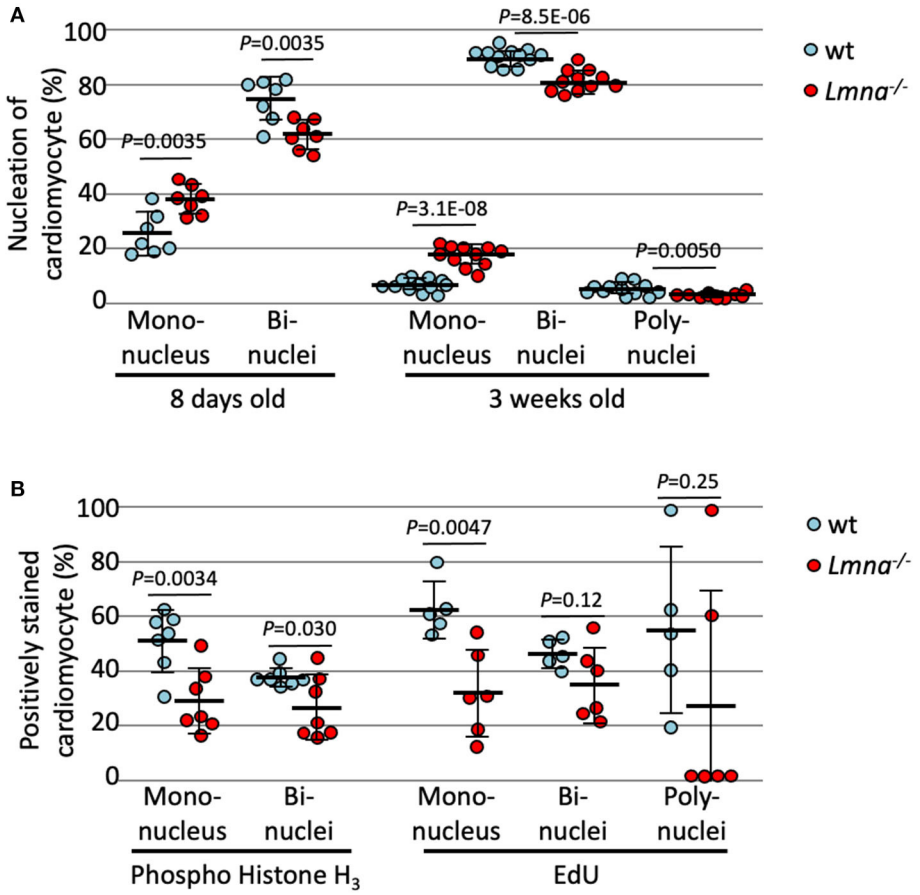


FIGURE 2 | Nucleation and cell cycle markers of cardiomyocytes in wild type (wt) and *Lmna*$^{-/-}$ mice. **(A)** Nucleation of cardiomyocytes in 8 days old ($n = 7$ in each group) and 3 weeks old ($n = 12$ in wt and 11 in *Lmna*$^{-/-}$) showed more mononuclear and less bi- or poly-nuclear cardiomyocytes in *Lmna*$^{-/-}$ mice. **(B)** Percentage of phospho-histone H_3 positive cardiomyocytes in 8 days old ($n = 7$ in each group) and of EdU positive cells in 3 weeks old ($n = 5$ in wt and 6 in *Lmna*$^{-/-}$) showed retarded cell cycle in *Lmna*$^{-/-}$ mice. Dots represent individual mouse data. The significance of differences between two groups was determined using the Student's 2 tailed *T*-test.

$100.9 \pm 5.2\ \mu m^2$ in wt, $P = 0.00012$, **Figures 1B–D** insets). These results indicated a defect in both proliferation and growth of cardiomyocytes in *Lmna*$^{-/-}$ mice.

Retardation of Nucleation and Cell Cycle Activity in *Lmna*$^{-/-}$ Mice

As previously mentioned, murine cardiomyocytes transition to binuclear from mononuclear ∼4 days after birth, and by 12 days, more than 90% of myocytes become binuclear. We hypothesized that binuclear or multinuclear cardiomyocytes are mature cells, while mononuclear cells are immature, and assessed the maturity of cardiomyocytes by counting the number of nuclei per cell. At 8 days after birth, *Lmna*$^{-/-}$ had a significantly higher number of mononuclear myocytes compared to wt (37.9 ± 5.4% in *Lmna*$^{-/-}$, 24.7 ± 8.0% in wt, $P = 0.0035$, $n = 7$ in each group, approximately 250 myocytes were counted in each mouse, **Figure 2A**), in consequence, binuclear cell counts were reduced in *Lmna*$^{-/-}$ compared to wt. At 3 weeks after birth, mononuclear myocytes were observed 16.6 ± 3.9% in *Lmna*$^{-/-}$ but only 5.6 ± 2.2% in wt ($P = 3.1 \times 10^{-8}$, $n = 11$, and 12, respectively, **Figure 2A**), while binuclear cells were 81.7 ± 4.1% in *Lmna*$^{-/-}$ but 90.4 ± 3.0% in wt ($P = 8.5 \times 10^{-6}$, **Figure 2A**). Furthermore, 4.0 ± 2.2% of myocytes in wt presented more than 3 nuclei, whereas only 1.7 ± 0.9% of myocytes were multinuclear in *Lmna*$^{-/-}$ ($P = 0.0050$, **Figure 2A**). These results suggest that the lack of *LMNA* retards the cardiomyocyte maturity.

To exclude the possibility that increased mononuclear cells in *Lmna*$^{-/-}$ myocytes are the results of the completion of cell division rather than immaturity (28), we evaluated the cell cycle status of each cardiomyocyte with immunocytochemistry of phospho-histone H_3 (pH_3), which is a mitosis phase marker, and EdU cell proliferation assay, which is a synthesis phase marker. As shown in **Figure 2B**, reduced pH_3 positive cardiomyocytes at 8 days after birth were observed in *Lmna*$^{-/-}$ compared to wt in both mononuclear and binuclear myocyte populations (28.3 ± 12.1% vs. 51.0 ± 11.3%, *Lmna*$^{-/-}$ and wt respectively, in mononuclear myocyte, $P = 0.0034$, 25.7 ± 11.9% vs. 37.2 ± 3.4%, *Lmna*$^{-/-}$ and wt, in binuclear myocyte, $P = 0.030$, $n = 7$ in each

group). EdU was injected intraperitoneally for 5 consecutive days from the day of birth and assessed at 21 days after birth. In the mononuclear myocyte population, EdU incorporated myocytes were also reduced in $Lmna^{-/-}$ compared to wt (31.2 ± 16.3% vs. 62.7 ± 10.3%, $P = 0.0047$). EdU incorporation in both bi and polynuclear $Lmna^{-/-}$ myocytes compared to wt was reduced but not significant. These results implicate that cardiomyocyte maturation is retarded because of the delayed cell cycle in $Lmna^{-/-}$ mice.

Reduced Cell Cycle Activity in Older *Lmna*$^{+/-}$ Mice

We also counted the number of nuclei per cell in 2-year-old $Lmna^{+/-}$, which developed DCM phenotype, and wt mice. At 3 weeks of age, the nucleation was not different between wild type and $Lmna^{+/-}$ mice (**Figure 3A**). In 2 years of age, however, $Lmna^{+/-}$ mice had a significantly higher number of mononuclear myocytes compared to wt as same as observed in the younger generation of $Lmna^{-/-}$ mice (19.8 ± 2.1% in $Lmna^{+/-}$, 10.6 ± 1.5% in wt, $P = 0.00040$, $n = 4$ in each group, ∼250 myocytes were counted in each mouse, **Figure 3B**), and binuclear cell counts were reduced in $Lmna^{+/-}$ compared to wt (77.6 ± 2.2% in $Lmna^{+/-}$, 87.3 ± 1.6% in wt, $P = 0.00039$, **Figure 3B**). The percentage of myocytes with more than 3 nuclei was not different between $Lmna^{+/-}$ and wt in 2 years old ($P = 0.35$). These results suggest that insufficiency of *LMNA* reduces cell cycle activity and retards cardiomyocyte maturity throughout the lifetime.

RNA-Sequencing of *Lmna* Mutant Mice

To elucidate the mechanism of reduced cell number, reduced cell size, immaturity, and cell cycle defects in *Lmna* mutant mice, we performed RNA-seq with RNA extracted from LV in $Lmna^{-/-}$, $Lmna^{+/-}$, and wt mice. One μg of RNA per mouse was pooled from 3 male mice and processed to construct a library as described in methods. Briefly, a total of 7 groups of mice were used: 5 weeks old wt, $Lmna^{+/-}$, and $Lmna^{-/-}$ mice, as well as wt and $Lmna^{+/-}$ mice at 1 and 2 years of age. The numbers of reads processed after sequencing and alignment to the reference genome mm9 are shown in **Supplementary Table 1**. The number of reads aligned to each gene was normalized by the total number of aligned reads. We then compared the expression profile between *Lmna* mutant and age-matched wt mice in 4 groups; namely $Lmna^{-/-}$/wt in 5 weeks, $Lmna^{+/-}$/wt in 5 weeks, $Lmna^{+/-}$/wt in 1 year, and $Lmna^{+/-}$/wt in 2 years. Among 30,387 genes listed in mm9, more than 2,000 genes had altered expression profiles in $Lmna^{-/-}$/wt in 5 weeks and $Lmna^{+/-}$/wt in 2 years groups. In contrast, only <250 genes had altered expression profiles in $Lmna^{+/-}$/wt in 5 weeks and $Lmna^{+/-}$/wt in 1-year groups, which corresponds to the echocardiographic results before developing overt phenotype (the numbers of up- or down-regulated genes are shown in **Supplementary Table 2**). There were no up- or down-regulated genes overlapping across all 4 groups. We analyzed these data using Ingenuity Pathway Analysis (IPA), a pathway analysis software, to understand the biological functions of the genes with altered expression. The results are shown in **Table 2** and **Supplementary Table 3**. We

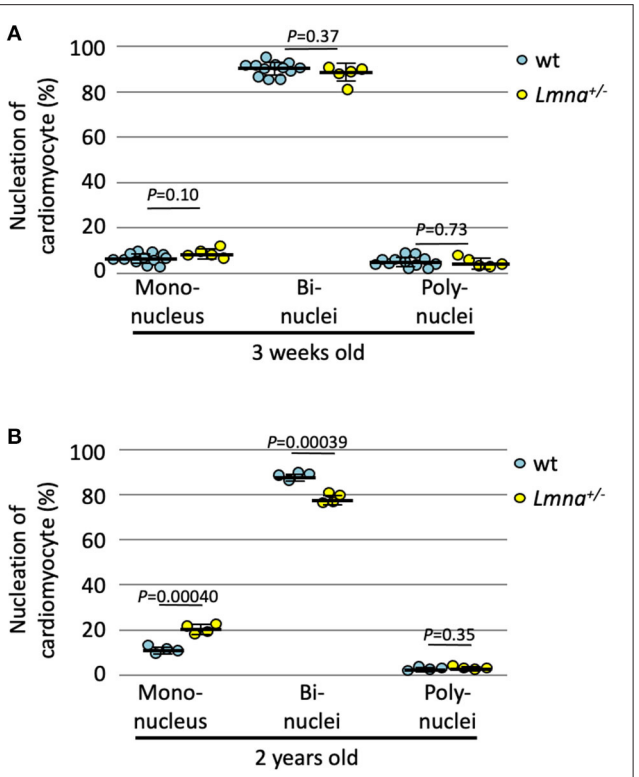


FIGURE 3 | Nucleation of cardiomyocytes in young and old wild type (wt) and $Lmna^{+/-}$ mice. **(A)** Nucleation of cardiomyocytes in 3 weeks old ($n = 12$ in wt and 5 in $Lmna^{+/-}$) was similar between wt and $Lmna^{+/-}$ mice. **(B)** Nucleation of cardiomyocytes in 2 years old ($n = 4$ in each group) showed more mononuclear and less binuclear cardiomyocytes in $Lmna^{+/-}$ mice. Dots represent individual mouse data. The significance of differences between two groups was determined using the Student's 2 tailed *T*-test.

focused on the results of $Lmna^{-/-}$ mice in 5 weeks old and of $Lmna^{+/-}$ mice in 2 years old since those two groups develop similar cardiac phenotypes with different onset. We found a significant number of standard biological functions that have been impacted by this *Lmna* mutation. The "proliferation of cells" was identified as one of the categories that had a strong impact both on $Lmna^{-/-}$/wt in 5 weeks and $Lmna^{+/-}$/wt in 2 years groups. **Supplementary Table 4** shows the genes in this category with a higher evidence level, associated with more than 50 papers reporting effects of these genes in cell proliferation. Among these genes, *Cdkn1a* had the most significant fold changes in both $Lmna^{-/-}$/wt in 5 weeks and $Lmna^{+/-}$/wt in 2 years groups, which encodes cell cycle regulating protein p21. Next, we analyzed which transcription regulators are mainly involved in *Lmna* mutation by IPA. As shown in **Supplementary Table 5**, TP53, which encodes p53 protein, was the most significantly affected transcriptional regulator in both $Lmna^{-/-}$/wt in 5 weeks and $Lmna^{+/-}$/wt in 2 years groups. As these genes exert significant effects in cell cycle regulation and proliferation, we inferred that upregulations of p53 and p21 contribute to the reduced cell proliferation and delayed cell cycle in *Lmna* mutant mice.

TABLE 2 | Top biological functions related to *Lmna* mutant mice in RNA-seq analysis.

Lmna$^{-/-}$/wt in 5 weeks			*Lmna*$^{+/-}$/wt in 2 years		
Biological function	**Number of genes/total genes in the category**	***P*-value**	**Biological function**	**Number of genes/total genes in the category**	***P*-value**
Proliferation of cells	686/5,802	1.48E-33	Cell movement	429/3,105	8.54E-48
Cell death	652/5,975	3.7E-31	Migration of cells	394/2,784	2.24E-46
Apoptosis	550/4,663	4.16E-31	Proliferation of cells	620/5,802	1.54E-38
Necrosis	508/4,635	6.94E-27	Leukocyte migration	218/1,449	1.77E-32
Migration of cells	382/2,784	9.68E-26	Organization of cytoskeleton	269/1,685	3.35E-32
Cell movement	413/3,105	6.20E-25	Development of blood vessel	196/1,176	2.17E-30
Leukocyte migration	199/1,449	1.50E-17	Apoptosis	486/4,663	5.71E-30
Development of blood vessel	185/1,176	2.42E-17	Organization of cytoplasm	278/1,816	1.77E-29
Vascular disease	178/1,280	3.92E-17	Cell death	569/5,975	1.88E-28
Development of cardiovascular system	225/1,460	5.38E-17	Necrosis	460/4,635	2.48E-28

Right-tailed Fisher's exact test.

p53 and p21 Are Highly Expressed; DNA Damage Is More Observed in *Lmna*$^{-/-}$ Mice

We performed protein analysis to evaluate whether those proteins are also upregulated by *Lmna* mutation. As shown in **Figure 4A**, p21 was highly expressed in *Lmna*$^{-/-}$ mice, especially in the nuclei of cardiomyocytes. Western blotting of protein extracted from LV revealed that not only p21 but p53 protein levels were also upregulated in *Lmna*$^{-/-}$ mice with 4.7-fold for p21 ($P = 0.012$) and 3.7-fold for p53 ($P = 0.030$) compared to wt (**Figures 4B,C**).

As DNA damage is known to induce both p21 and p53 (35), we performed immunohistochemistry of phospho-histone H_2AX (pH_2AX) to detect DNA double-strand breaks (**Supplementary Figure 2**). We found a significantly higher percentage of pH_2AX positive myocytes in *Lmna*$^{-/-}$ compared to *Lmna*$^{+/-}$ or wt mice (0.0043 ± 0.0011%, 0.069 ± 0.015%, and 0.49 ± 0.18%, wt, *Lmna*$^{+/-}$ and *Lmna*$^{-/-}$, respectively, **Figure 4D**). This result suggested that mice with *Lmna* mutation have more DNA damage than wt, possibly responsible for the increased p53 and p21 activities.

Lmna Mutant Mice had a Repressed Cardiomyocyte Proliferative Capacity After Resecting the Apex of the Heart

Next, we performed LV apical resection to assess the ability of cardiomyocyte proliferation. The experiment was performed on wt, *Lmna*$^{+/-}$, and *Lmna*$^{-/-}$ mice at the age of 2 days as described in Materials and Methods. EdU incorporated myocytes distributed around the resected area and remote area diffusely in the heart as previously reported (36). Eight high power fields per each cross-sectioned sample were chosen so that each field was evenly distributed across the section, and the percentage of EdU positive cardiomyocyte nuclei among total cardiomyocyte nuclei was counted. In the sham operation, the percentage of EdU positive cardiomyocyte nuclei was not different among 3 genotypes (17.5 ± 2.9%, 15.7 ± 1.6% and 15.3 ± 2.3%, wt, *Lmna*$^{+/-}$, and *Lmna*$^{-/-}$, respectively, ANOVA $P = 0.0961$, **Figure 5**). However, after apical resection, EdU incorporation into the nucleus was significantly accelerated, especially in wt from 17.5 ± 2.9% to 24.3 ± 2.2% ($P = 0.00017$). In *Lmna*$^{+/-}$ mice, a similar effect was observed from 15.7 ± 1.6% to 19.1 ± 2.4% ($P = 0.0014$), although the change was milder than that in wt. On the other hand, in *Lmna*$^{-/-}$ mice, the percentage of EdU positive cardiomyocyte nuclei did not increase from sham operation (15.3 ± 2.3%) even after apical resection (15.8 ± 1.6%, $P = 0.58$). The difference between these 3 groups after apical resection was significant ($P < 0.0001$). These data suggested that *Lmna* mutant mice have defective cardiomyocyte proliferative capacity compared to wt after stimulation, which normally leads to cell cycle activation.

DISCUSSION

We report here that the *Lmna* mutation retards cardiomyocyte proliferation and maturation processes and prevents compensatory proliferative response induced by apical resection, which could be one of the mechanisms of DCM development in patients with *LMNA* mutation.

It was reported that *Lmna* mutations are associated with DNA damage (12, 14, 37). As a structural protein of the nuclear envelope, the abnormality of lamin leads to deformity of the nuclear membrane and results in defective disassembly and reassembly processes of the nuclear envelope during the mitosis phase of the cell cycle. Since lamin also works as an anchor protein of chromatin, the abnormality of lamin further leads to defective chromatin arrangement (38, 39). These problems can influence DNA replication and thus be the cause for DNA damage (12, 14, 37). Once DNA is damaged, p53 and p21 are activated as previously reported (35, 40), which consequently retards the cell cycle and restricts cardiomyocyte proliferation as well as cardiomyocyte growth. In this study, DNA double-strand break was detected with pH_2AX immunostaining, which can then activate the p53-p21 signaling axis. Actually, RNA-seq elucidated that p21 was most strongly activated in

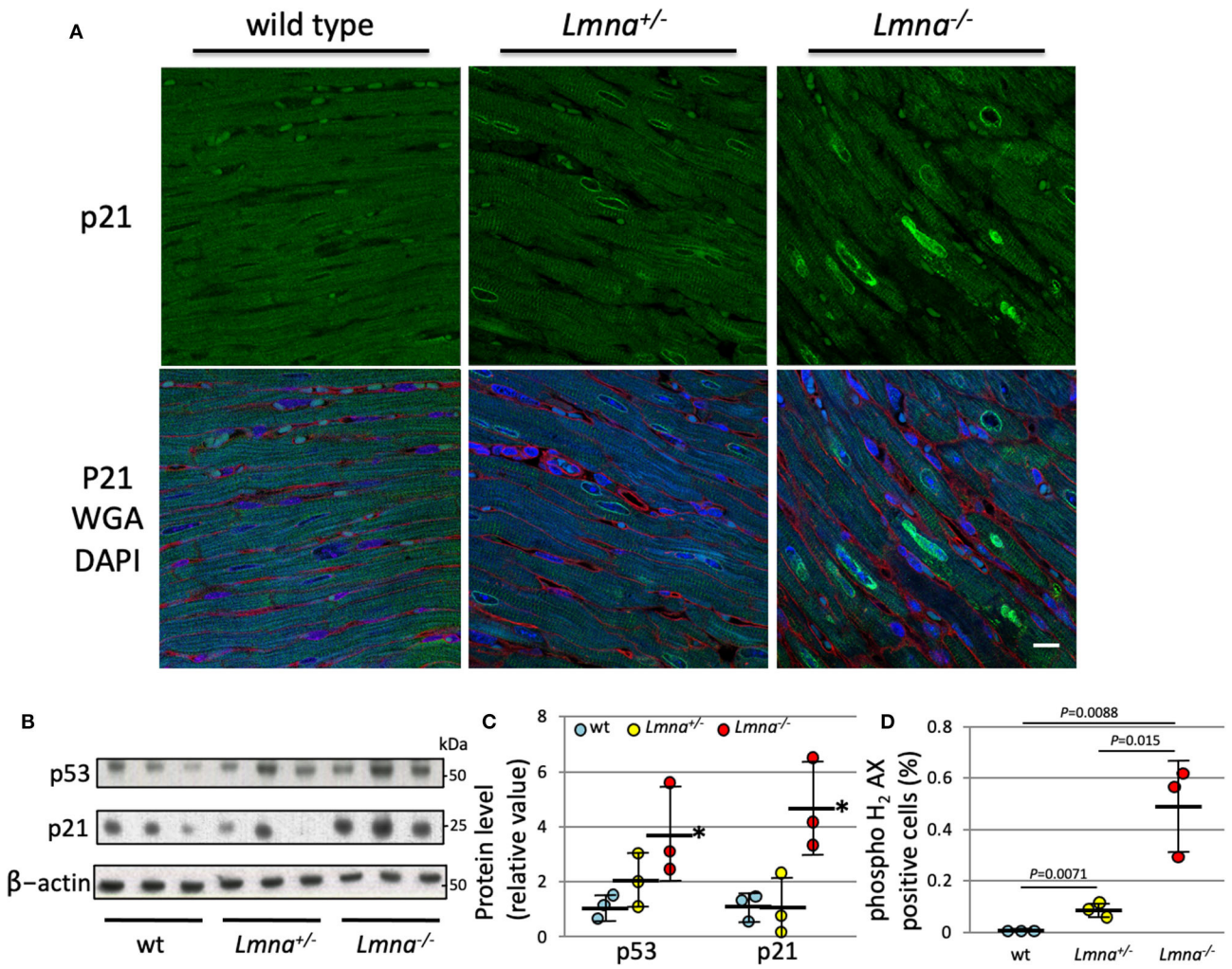


FIGURE 4 | Protein expression of p21 and p53, and DNA damage of cardiomyocyte in 5 weeks. **(A)** Immunostaining against anti-p21 antibody in heart tissue section of wt, $Lmna^{+/-}$, and $Lmna^{-/-}$. The top panel shows p21 staining, and the bottom panel shows merged figures of anti-p21 antibody (green), WGA (red), and DAPI (blue). Scale bar: 10 μm. **(B)** Western blot analysis of LV tissue hybridized with p53, p21, and β-actin Abs. The data shown are representative of three independent experiments. **(C)** Protein expression level of p53 and p21 normalized by β-actin. *$P < 0.05$ vs. respective wt. **(D)** Percentage of phospho histone H_2 AX staining counted in cardiomyocytes (**Supplementary Figure 2**) ($n = 3$ in each group). Dots represent individual mouse data. The significance of differences between 3 groups was determined with 1-way ANOVA. *Post-hoc* pairwise comparisons were performed with the Tukey–Kramer test.

Lmna mutant mice, which was also confirmed by the protein expression. These results indicate that DNA damage, which arises from nuclear membrane malformation, structural and functional abnormalities as reported before, disturbs cell maturation and proliferation through the p53-p21 axis. Moiseeva et al. reported that depletion of *Lmna* leads to cell cycle arrest via p21 up-regulation by using primary human fibroblasts (29). In this study, we observed this cell cycle defects also *in vivo* heart as decreased cell numbers counted by tissue section.

Not only decreased cell numbers but also diminished cardiomyocyte binucleation was observed in *Lmna* mutant mice, suggesting inadequate maturation caused by cell cycle defects. As previously reported, mammalian cardiomyocytes exit the cell cycle after karyokinesis without cytokinesis shortly after birth (21, 24), making cardiomyocytes binucleated or multinucleated. The percentage of cardiomyocytes with more than 1 nucleus in adult varies depending on species, e.g., approximately 90% in mouse (21, 24), 70% in rat (41) and sheep (42), and 30% in human (43). In a murine heart, the higher the percentage of mononucleated cardiomyocyte exists, the more immaturity the cardiomyocyte has, because the ratio of a mononucleated cardiomyocyte is reported to be reduced as the heart develops (26). This decreased percentage of mononucleated cardiomyocyte was also observed in this study when we compared 8 days and 3 weeks old mice. The percentage of mononucleated cardiomyocytes was higher in $Lmna^{-/-}$ than in wt mice in both ages. This fact also indicates cell cycle is retarded before the mitosis phase in *Lmna* mutant mice.

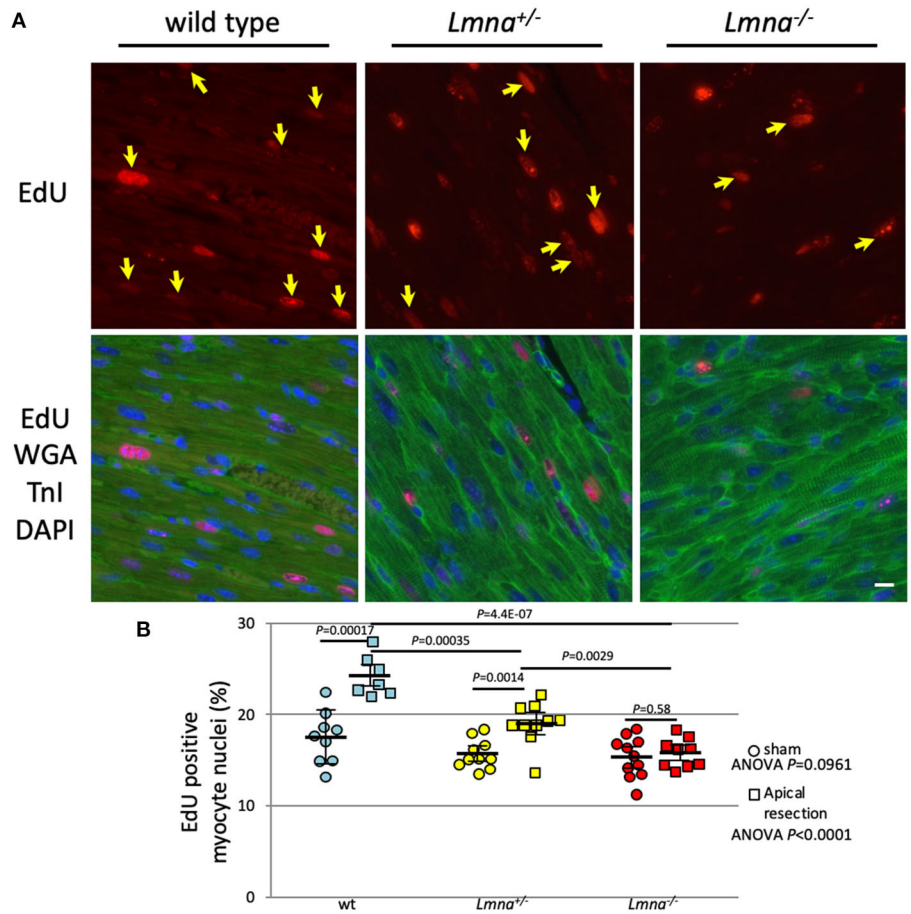


FIGURE 5 | EdU incorporation and the percentage of EdU positive myocyte nuclei after apical resection. **(A)** Representative immunostaining for EdU in cardiac tissue sections of wt, $Lmna^{+/-}$, and $Lmna^{-/-}$. The top panel shows EdU staining, and the bottom panel shows merged figures of immunostaining for EdU (red), WGA, and troponin I (green) and DAPI (blue). Yellow arrows indicate EdU positive nucleus only in cardiomyocytes. Scale bar: 10 μm. **(B)** Percentage of EdU incorporated myocyte nuclei ($n = 7$ to 11 in each group). A circle indicates data from mice with the sham operation, square with apical resection. Dots represent individual mouse data. The significance of differences between 3 groups was determined with 1-way ANOVA. *Post-hoc* pairwise comparisons were performed with Tukey–Kramer test.

We propose that the vicious cascade of cell cycle defects in patients with *LMNA* mutations not only diminishes cell growth and proliferation in the physiological setting but alters the ability of the cell to respond against many kinds of stresses, including cardiomyocyte injury or damage, leading to the development of DCM. As shown in **Figure 5**, cardiomyocyte proliferative capacity after apical resection was diminished in $Lmna^{-/-}$ mice. As Porrello et al. reported previously, once some signal concerning acceleration of cell proliferation is turned on, cardiomyocytes of mice can replicate in response to that if mice are not older than a week of age (31). Even after entering adulthood, cardiomyocytes can proliferate under stress, such as in ischemic heart disease (19). This compensatory myocyte proliferation was actually activated by apical resection in this study. However, the extent of DNA synthesis differed between wt and *Lmna* mutant mice. Furthermore, cardiomyocyte is reported to continuously but very slowly regenerate in human throughout life, 1% annually in young adult and 0.45% at the age of 75. And ∼50% of cardiomyocytes are replaced during an average life span in humans (44). We, therefore, speculate that both the physiological regeneration process of cardiomyocytes during lifetime and the appropriate regenerative response under stress are retarded in patients with *LMNA* mutation, as suggested from the results of inappropriate cardiomyocyte maturation observed in 2-year-old *Lmna* mice and defective proliferative capacity after apical resection in *Lmna* mutant mice at 2 days after birth. And those defects could lead to the development of DCM, as observed in the reduced heart function at the age of 2 years in $Lmna^{+/-}$ mice. **Figure 6** shows our model of disease development for patients with *LMNA* mutations.

As we reported recently, cardiomyocytes in *Mybpc3* mutant mice also have fewer nuclei numbers (28). We concluded this is because cardiomyocytes undergo an additional round of cell division within 10 days after birth since MYBPC has inhibitory functions during postnatal cardiomyocyte cytokinesis and cell cycle progression. In *Mybpc3* mutant mice, DNA

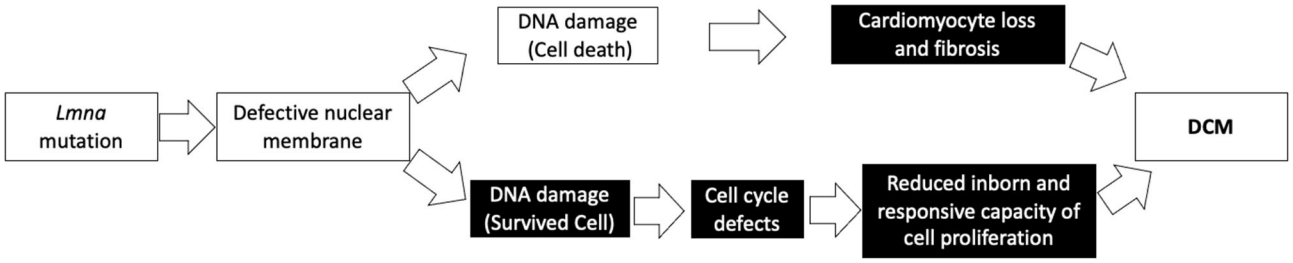


FIGURE 6 | Our model depicting a possible mechanism in the development of DCM with a *Lmna* mutation. Black boxes denote findings from this study.

synthesis in cardiomyocytes and mitosis markers were observed in higher levels than wt mice, which is entirely opposite in *Lmna* mutant mice in this study. In *Mybpc3* mutant mice, the cell cycle progresses an additional round resulting in a higher percentage of mononuclear cardiomyocytes and a higher number of cardiomyocytes, whereas in *Lmna* mutant mice, cell cycle delays compared to wt resulting in a higher percentage of mononuclear cardiomyocytes but a fewer number of cardiomyocytes. This difference comes from a difference in the cell cycle, accelerated in *Mybpc3* and delayed in *Lmna* mutant mice, and presumably contributes to the phenotypic divergence of developing hypertrophic cardiomyopathy in patients with *MYBPC3* mutation while developing DCM in those with *LMNA* mutation.

We performed RNA-seq in this study, which resulted in robust transcriptional profiles (33, 45). We have successfully identified some biological functions that correlate with *Lmna* mutation. Among these, "cell proliferation" had the most prominent relationship to the *Lmna* mutation by the IPA pathway analysis. Indeed, we observed the cell number difference in physiological conditions as well as the repressed response after apical resection in *Lmna* mutant mice. Besides that, the biological functions related to "cell death" also significantly influenced *Lmna* mutant mice. It was difficult to identify the signal of cell death as either apoptosis or necrosis in this study, however, interstitial fibrosis in the heart was more frequently and massively observed in *Lmna* mutant mice, which suggested that injured cardiomyocytes were more abundant in *Lmna* mutant mice and were replaced by fibrosis. This progression of fibrosis could also be involved in DCM development with *Lmna* mutation (**Supplementary Figure 1**).

In this study, like many other studies, we found the body size and heart size difference between wt and $Lmna^{-/-}$ mice, although their size is comparable when they are born. Lamin B is another structural protein in the inner nuclear membrane coded by *LMNB1*. $Lmnb1^{-/-}$ mice are known to be embryonic lethal (46), while $Lmna^{-/-}$ mice are not. In the embryonic stage, lamin B may have vital roles, and a class switch system from lamin B to lamins A and C at the perinatal stage may occur. Actually, when we compared the expression profiles in LV between embryonic day 14.5 (E14.5) and postnatal day 0, the expression of *Lmna* increases to 174.8%, while *Lmnb1* decreases to 34.8% of E14.5 mice at birth, respectively. This class switch of lamins might contribute to the cardiac growth defects after birth in *Lmna* mutant mice. Elucidation of the mechanism of the class switch system may also further reveal strategies to treat patients with *LMNA* mutation by either the inhibition of this class switch or by promoting the activation of the *LMNB1* gene.

Besides having a role in reinforcing the structure of the nucleus, lamin proteins are reported to be significant players in signal transduction or chromatin regulation (47, 48). Muchir et al. reported that mitogen-activated protein kinase (MAPK) signaling is activated in *Lmna* H222P knock-in mouse ($Lmna^{H222P/H222P}$) hearts from abnormal activation of extracellular signal-regulated kinase (ERK) and Jun amino-terminal kinase (JNK), and this activation leads to DCM phenotype through actin depolymerization (49, 50). They also succeeded in proving that the inhibition of these pathways can prevent the mice from developing DCM (51). Another group showed only up-regulation of ERK and not of JNK after pressure overload induced by transverse aortic constriction (34). Our study did not observe any significant differences in these genes at any time point, which may be due to the different mouse lines between those previous studies and this study.

As the previous papers reported, conduction disorder and ventricular arrhythmia are also the characteristics of DCM with *LMNA* mutation (5, 17, 18). In most cases, the conduction disorder preceded LV dysfunction and ventricular arrhythmia (52, 53). In this study, although we didn't perform the precise evaluation on the arrhythmic status by electrocardiogram or telemetric studies as the previous study examined these electrophysiological properties using the same mutant mice line (17), we speculate the massive fibrosis more observed in *Lmna* mutant mice as shown in **Supplementary Figure 1** could be the focus of ventricular arrhythmia. The reason why the conduction disorder precedes LV dysfunction or ventricular arrhythmia should be elucidated in a future study.

After the development, our model mice generated by the deletion of exon 8 to 11 of *Lmna* had been thought to be a "null" mutation (12). Recently Jahn et al. reported this mouse line has a truncated lamin A protein (54). Many researchers have used this *Lmna* mutant mouse line, and all studies are reported that this truncated protein does not work as a dominant-negative manner but as a loss-of-function manner (13, 17, 30, 55). Hence, we described this mouse line as *Lmna* knockout instead of $Lmna^{\Delta 8-11}$ in this study. Although careful interpretation is

needed to use this mouse line, we believe this truncated protein works as a loss-of-function manner because the heterozygous mice live as long as wt under normal conditions, although they develop heart dysfunction in the last stage of life.

In conclusion, we found that the cell cycle alterations, including activation of the p53-p21 axis and inadequate responses against stresses in *Lmna* mutant mice. We speculate these phenomena could play essential roles in the development of DCM caused by *LMNA* mutations. Although further studies are needed to fully understand the mechanism, modulating the cell cycle activity could be an efficient treatment strategy for DCM patients with *LMNA* mutation.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the NCBI BioProject repository, accession number PRJNA732812. This data can be found online at: http://www.ncbi.nlm.nih.gov/bioproject/732812.

ETHICS STATEMENT

The animal study was reviewed and approved by Harvard Medical School and Nara Medical University.

AUTHOR CONTRIBUTIONS

KO designed the study, conducted and analyzed the experiments, performed the statistical analysis, and wrote the manuscript. HW contributed to study design, edited the manuscript, and conducted the animal experiments with JJ. MP and SD analyzed RNA-seq data. DC designed and conducted the animal study. JG contributed to extract RNA and build RNA-seq libraries. DM analyzed RNA-seq data, performed the statistical analysis, and edited the manuscript. JS, CS, and YS contributed to study design, provided scientific input, and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported in part by Howard Hughes Medical Institute (to CS, JJ, and SD), National Heart, Lung, and Blood Institute, NIH (5R01HL084553 and 5R01HL080494 to HW, KO, DM, JS, and CS), Fondation Leducq (to CS and JS), Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (26461135 to KO and 25670393 to YS) and Japan Heart Foundation Research Grant on Dilated Cardiomyopathy (to KO).

ACKNOWLEDGMENTS

The authors would like to thank Colin L. Stewart for allowing us to use *Lmna* mutant mice generated by him and their colleagues at Harvard Medical School and Nara Medical University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2021.639148/full#supplementary-material

REFERENCES

1. Watkins H, Ashrafian H, Redwood C. Inherited cardiomyopathies. *N Engl J Med.* (2011) 364:1643–56. doi: 10.1056/NEJMra0902923
2. van Tintelen JP, Hofstra RM, Katerberg H, Rossenbacker T, Wiesfeld AC, du Marchie Sarvaas GJ, et al. High yield of LMNA mutations in patients with dilated cardiomyopathy and/or conduction disease referred to cardiogenetics outpatient clinics. *Am Heart J.* (2007) 154:1130–39. doi: 10.1016/j.ahj.2007.07.038
3. Herman DS, Lam L, Taylor MR, Wang L, Teekakirikul P, Christodoulou D, et al. Truncations of titin causing dilated cardiomyopathy. *N Engl J Med.* (2012) 366:619–28. doi: 10.1056/NEJMoa1110186
4. Kayvanpour E, Sedaghat-Hamedani F, Amr A, Lai A, Haas J, Holzer DB, et al. Genotype-phenotype associations in dilated cardiomyopathy: meta-analysis on more than 8000 individuals. *Clin Res Cardiol.* (2017) 106:127–39. doi: 10.1007/s00392-016-1033-6
5. Fatkin D, MacRae C, Sasaki T, Wolff MR, Porcu M, Frenneaux M, et al. Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N Engl J Med.* (1999) 341:1715–24. doi: 10.1056/NEJM199912023412302
6. Arbustini E, Pilotto A, Repetto A, Grasso M, Negri A, Diegoli M, et al. Autosomal dominant dilated cardiomyopathy with atrioventricular block: a lamin A/C defect-related disease. *J Am Coll Cardiol.* (2002) 39:981–90. doi: 10.1016/S0735-1097(02)01724-2
7. Tobita T, Nomura S, Fujita T, Morita H, Asano Y, Onoue K, et al. Genetic basis of cardiomyopathy and the genotypes involved in prognosis and left ventricular reverse remodeling. *Sci Rep.* (2018) 8:1998. doi: 10.1038/s41598-018-20114-9
8. Muchir A, Bonne G, van der Kooi AJ, van Meegen M, Baas F, Bolhuis PA, et al. Identification of mutations in the gene encoding lamins A/C in autosomal dominant limb girdle muscular dystrophy with atrioventricular conduction disturbances (LGMD1B). *Hum Mol Genet.* (2000) 9:1453–9. doi: 10.1093/hmg/9.9.1453
9. Bonne G, Di Barletta MR, Varnous S, Becane HM, Hammouda EH, Merlini L, et al. Mutations in the gene encoding lamin A/C cause autosomal dominant emery-dreifuss muscular dystrophy. *Nat Genet.* (1999) 21:285–8. doi: 10.1038/6799
10. De Sandre-Giovannoli A, Chaouch M, Kozlov S, Vallat JM, Tazir M, Kassouri N, et al. Homozygous defects in LMNA, encoding lamin A/C nuclear-envelope proteins, cause autosomal recessive axonal neuropathy in human (charcot-marie-tooth disorder type 2) and mouse. *Am J Hum Genet.* (2002) 70:726–36. doi: 10.1086/339274
11. Eriksson M, Brown WT, Gordon LB, Glynn MW, Singer J, Scott L, et al. Recurrent *de novo* point mutations in lamin A cause hutchinson-gilford progeria syndrome. *Nature.* (2003) 423:293–8. doi: 10.1038/nature01629
12. Sullivan T, Escalante-Alcalde D, Bhatt H, Anver M, Bhat N, Nagashima K, et al. Loss of A-type lamin expression compromises nuclear envelope integrity leading to muscular dystrophy. *J Cell Biol.* (1999) 147:913–20. doi: 10.1083/jcb.147.5.913
13. Nikolova V, Leimena C, McMahon AC, Tan JC, Chandar S, Jogia D, et al. Defects in nuclear structure and function promote dilated cardiomyopathy in lamin A/C-deficient mice. *J Clin Invest.* (2004) 113:357–69. doi: 10.1172/JCI200419448

14. Butin-Israeli V, Adam SA, Goldman AE, Goldman RD. Nuclear lamin functions and disease. *Trends Genet.* (2012) 28:464–71. doi: 10.1016/j.tig.2012.06.001
15. Bhattacharjee P, Dasgupta D, Sengupta K. DCM associated LMNA mutations cause distortions in lamina structure and assembly. *Biochim Biophys Acta Gen Subj.* (2017) 1861 (11 Pt. A):2598–608. doi: 10.1016/j.bbagen.2017.08.016
16. Morales Rodriguez B, Dominguez-Rodriguez A, Benitah JP, Lefebvre F, Marais T, Mougenot N, et al. Activation of sarcolipin expression and altered calcium cycling in LMNA cardiomyopathy. *Biochem Biophys Rep.* (2020) 22:100767. doi: 10.1016/j.bbrep.2020.100767
17. Wolf CM, Wang L, Alcalai R, Pizard A, Burgon PG, Ahmad F, et al. Lamin A/C haploinsufficiency causes dilated cardiomyopathy and apoptosis-triggered cardiac conduction system disease. *J Mol Cell Cardiol.* (2008) 44:293–303. doi: 10.1016/j.yjmcc.2007.11.008
18. Macquart C, Juttner R, Morales Rodriguez B, Le Dour C, Lefebvre F, Chatzifrangkeskou M, et al. Microtubule cytoskeleton regulates connexin 43 localization and cardiac conduction in cardiomyopathy caused by mutation in A-type lamins gene. *Hum Mol Genet.* (2019) 28:4043–52. doi: 10.1093/hmg/ddy227
19. Nakada Y, Canseco DC, Thet S, Abdisalaam S, Asaithamby A, Santos CX, et al. Hypoxia induces heart regeneration in adult mice. *Nature.* (2017) 541:222–7. doi: 10.1038/nature20173
20. Borden A, Kurian J, Nickoloff E, Yang Y, Troupes CD, Ibetti J, et al. Transient Introduction of miR-294 in the heart promotes cardiomyocyte cell cycle reentry after injury. *Circ Res.* (2019) 125:14–25. doi: 10.1161/CIRCRESAHA.118.314223
21. Li F, Wang X, Capasso JM, Gerdes AM. Rapid transition of cardiac myocytes from hyperplasia to hypertrophy during postnatal development. *J Mol Cell Cardiol.* (1996) 28:1737–46. doi: 10.1006/jmcc.1996.0163
22. Siggens L, Figg N, Bennett M, Foo R. Nutrient deprivation regulates DNA damage repair in cardiomyocytes via loss of the base-excision repair enzyme OGG1. *FASEB J.* (2012) 26:2117–24. doi: 10.1096/fj.11-197525
23. Puente BN, Kimura W, Muralidhar SA, Moon J, Amatruda JF, Phelps KL, et al. The oxygen-rich postnatal environment induces cardiomyocyte cell-cycle arrest through DNA damage response. *Cell.* (2014) 157:565–79. doi: 10.1016/j.cell.2014.03.032
24. Walsh S, Ponten A, Fleischmann BK, Jovinge S. Cardiomyocyte cell cycle control and growth estimation *in vivo*–an analysis based on cardiomyocyte nuclei. *Cardiovasc Res.* (2010) 86:365–73. doi: 10.1093/cvr/cvq005
25. Lewandowski AJ, Augustine D, Lamata P, Davis EF, Lazdam M, Francis J, et al. Preterm heart in adult life: cardiovascular magnetic resonance reveals distinct differences in left ventricular mass, geometry, and function. *Circulation.* (2013) 127:197–206. doi: 10.1161/CIRCULATIONAHA.112.126920
26. Jonker SS, Louey S, Giraud GD, Thornburg KL, Faber JJ. Timing of cardiomyocyte growth, maturation, and attrition in perinatal sheep. *FASEB J.* (2015) 29:4346–57. doi: 10.1096/fj.15-272013
27. Hille S, Dierck F, Kuhl C, Sosna J, Adam-Klages S, Adam D, et al. Dyrk1a regulates the cardiomyocyte cell cycle via D-cyclin-dependent Rb/E2f-signalling. *Cardiovasc Res.* (2016) 110:381–94. doi: 10.1093/cvr/cvw074
28. Jiang J, Burgon PG, Wakimoto H, Onoue K, Gorham JM, O'Meara CC, et al. Cardiac myosin binding protein C regulates postnatal myocyte cytokinesis. *Proc Natl Acad Sci USA.* (2015) 112:9046–51. doi: 10.1073/pnas.1511004112
29. Moiseeva O, Bourdeau V, Vernier M, Dabauvalle MC, Ferbeyre G. Retinoblastoma-independent regulation of cell proliferation and senescence by the p53-p21 axis in lamin A /C-depleted cells. *Aging Cell.* (2011) 10:789–97. doi: 10.1111/j.1474-9726.2011.00719.x
30. Cohen TV, Gnocchi VF, Cohen JE, Phadke A, Liu H, Ellis JA, et al. Defective skeletal muscle growth in lamin A/C-deficient mice is rescued by loss of Lap2alpha. *Hum Mol Genet.* (2013) 22:2852–69. doi: 10.1093/hmg/ddt135
31. Porrello ER, Mahmoud AI, Simpson E, Hill JA, Richardson JA, Olson EN, et al. Transient regenerative potential of the neonatal mouse heart. *Science.* (2011) 331:1078–80. doi: 10.1126/science.1200708
32. Christodoulou DC, Gorham JM, Herman DS, Seidman JG. Construction of normalized RNA-seq libraries for next-generation sequencing using the crab duplex-specific nuclease. *Curr Protoc Mol Biol.* (2011) Chapter 4:Unit4 12. doi: 10.1002/0471142727.mb0412s94
33. Christodoulou DC, Gorham JM, Kawana M, DePalma SR, Herman DS, Wakimoto H. Quantification of gene transcripts with deep sequencing analysis of gene expression (DSAGE) using 1 to 2 microg total RNA. *Curr Protoc Mol Biol.* (2011) Chapter 25:Unit25B 9. doi: 10.1002/0471142727.mb25b09s93
34. Chandar S, Yeo LS, Leimena C, Tan JC, Xiao XH, Nikolova-Krstevski V, et al. Effects of mechanical stress and carvedilol in lamin A/C-deficient dilated cardiomyopathy. *Circ Res.* (2010) 106:573–82. doi: 10.1161/CIRCRESAHA.109.204388
35. Li R, Waga S, Hannon GJ, Beach D, Stillman B. Differential effects by the p21 CDK inhibitor on PCNA-dependent DNA replication and repair. *Nature.* (1994) 371:534–7. doi: 10.1038/371534a0
36. Porrello ER, Olson EN. A neonatal blueprint for cardiac regeneration. *Stem Cell Res.* (2014) 13 (3 Pt. B):556–70. doi: 10.1016/j.scr.2014.06.003
37. Worman HJ, Fong LG, Muchir A, Young SG. Laminopathies and the long strange trip from basic cell biology to therapy. *J Clin Invest.* (2009) 119:1825–36. doi: 10.1172/JCI37679
38. Dechat T, Adam SA, Goldman RD. Nuclear lamins and chromatin: when structure meets function. *Adv Enzyme Regul.* (2009) 49:157–66. doi: 10.1016/j.advenzreg.2008.12.003
39. Khadija SG, Chen F, Hadden T, Commissaris RL, Kowluru A. Nuclear Lamins: biology and roles in cell function and dysregulation. *Recent Pat Endocr Metab Immune Drug Discov.* (2015) 9:111–20. doi: 10.2174/1872214809666151009120402
40. Herbert KE, Mistry Y, Hastings R, Poolman T, Niklason L, Williams B. Angiotensin II-mediated oxidative DNA damage accelerates cellular senescence in cultured human vascular smooth muscle cells via telomere-dependent and independent pathways. *Circ Res.* (2008) 102:201–8. doi: 10.1161/CIRCRESAHA.107.158626
41. Lushnikova EL, Tolstikova TG, Nepomnyashchikh LM, Klinnikova MG, Molodykh OP, Sviridov EA, et al. Cardiomyocyte count in rat myocardium under the effect of antitumor agents cyclophosphamide and triterpenoids. *Bull Exp Biol Med.* (2007) 144:355–61. doi: 10.1007/s10517-007-0332-3
42. Barbera A, Giraud GD, Reller MD, Maylie J, Morton MJ, Thornburg KL. Right ventricular systolic pressure load alters myocyte maturation in fetal sheep. *Am J Physiol Regul Integr Comp Physiol.* (2000) 279:R1157–64. doi: 10.1152/ajpregu.2000.279.4.R1157
43. Olivetti G, Cigola E, Maestri R, Corradi D, Lagrasta C, Gambert SR, et al. Aging, cardiac hypertrophy and ischemic cardiomyopathy do not affect the proportion of mononucleated and multinucleated myocytes in the human heart. *J Mol Cell Cardiol.* (1996) 28:1463–77. doi: 10.1006/jmcc.1996.0137
44. Bergmann O, Bhardwaj RD, Bernard S, Zdunek S, Barnabe-Heider F, Walsh S, et al. Evidence for cardiomyocyte renewal in humans. *Science.* (2009) 324:98–102. doi: 10.1126/science.1164680
45. Christodoulou DC, Wakimoto H, Onoue K, Eminaga S, Gorham JM, DePalma SR, et al. 5'RNA-Seq identifies Fhl1 as a genetic modifier in cardiomyopathy. *J Clin Invest.* (2014) 124:1364–70. doi: 10.1172/JCI70108
46. Vergnes L, Peterfy M, Bergo MO, Young SG, Reue K. Lamin B1 is required for mouse development and nuclear integrity. *Proc Natl Acad Sci USA.* (2004) 101:10428–33. doi: 10.1073/pnas.0401424101
47. Andres V, Gonzalez JM. Role of A-type lamins in signaling, transcription, chromatin organization. *J Cell Biol.* (2009) 187:945–57. doi: 10.1083/jcb.200904124
48. Solovei I, Wang AS, Thanisch K, Schmidt CS, Krebs S, Zwerger M, et al. LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell.* (2013) 152:584–98. doi: 10.1016/j.cell.2013.01.009
49. Muchir A, Pavlidis P, Decostre V, Herron AJ, Arimura T, Bonne G, et al. Activation of MAPK pathways links LMNA mutations to cardiomyopathy in emery-dreifuss muscular dystrophy. *J Clin Invest.* (2007) 117:1282–93. doi: 10.1172/JCI29042
50. Chatzifrangkeskou M, Yadin D, Marais T, Chardonnet S, Cohen-Tannoudji M, Mougenot N, et al. Cofilin-1 phosphorylation catalyzed by ERK1/2 alters cardiac actin dynamics in dilated cardiomyopathy caused by lamin A/C gene mutation. *Hum Mol Genet.* (2018) 27:3060–78. doi: 10.1093/hmg/ddy215
51. Wu W, Muchir A, Shan J, Bonne G, Worman HJ. Mitogen-activated protein kinase inhibitors improve heart function and prevent fibrosis in

cardiomyopathy caused by mutation in lamin A/C gene. *Circulation.* (2011) 123:53–61. doi: 10.1161/CIRCULATIONAHA.110.970673
52. Arimura T, Onoue K, Takahashi-Tanaka Y, Ishikawa T, Kuwahara M, Setou M, et al. Nuclear accumulation of androgen receptor in gender difference of dilated cardiomyopathy due to lamin A/C mutations. *Cardiovasc Res.* (2013) 99:382–94. doi: 10.1093/cvr/cvt106
53. van Rijsingen IA, Nannenberg EA, Arbustini E, Elliott PM, Mogensen J, Hermans-van Ast JF, et al. Gender-specific differences in major cardiac events and mortality in lamin A/C mutation carriers. *Eur J Heart Fail.* (2013) 15:376–84. doi: 10.1093/eurjhf/hfs191
54. Jahn D, Schramm S, Schnolzer M, Heilmann CJ, de Koster CG, Schutz W, et al. A truncated lamin A in the Lmna -/- mouse line: implications for the understanding of laminopathies. *Nucleus.* (2012) 3:463–74. doi: 10.4161/nucl.21676
55. Nikolova-Krstevski V, Leimena C, Xiao XH, Kesteven S, Tan JC, Yeo LS, et al. Nesprin-1 and actin contribute to nuclear and cytoskeletal defects in lamin A/C-deficient cardiomyopathy. *J Mol Cell Cardiol.* (2011) 50:479–86. doi: 10.1016/j.yjmcc.2010.12.001

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 12 August 2021
doi: 10.3389/fcvm.2021.713068

Effect of *SYTL3-SLC22A3* Variants, Their Haplotypes, and G × E Interactions on Serum Lipid Levels and the Risk of Coronary Artery Disease and Ischaemic Stroke

Peng-Fei Zheng[1], *Rui-Xing Yin*[1]*, *Xiao-Li Cao*[2], *Wu-Xian Chen*[1], *Jin-Zhen Wu*[1] and *Feng Huang*[1]

[1] Department of Cardiology, Institute of Cardiovascular Diseases, The First Affiliated Hospital, Guangxi Medical University, Nanning, China, [2] Department of Neurology, The First Affiliated Hospital, Guangxi Medical University, Nanning, China

OPEN ACCESS

Edited by:
Seitaro Nomura,
The University of Tokyo, Japan

Reviewed by:
Liu Miao,
Liuzhou People's Hospital, China
Raj Sewduth,
VIB KU Leuven Center for Cancer Biology, Belgium

***Correspondence:**
Rui-Xing Yin
yinruixing@163.com
orcid.org/0000-0001-7883-4310

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: 21 May 2021
Accepted: 21 July 2021
Published: 12 August 2021

Citation:
Zheng P-F, Yin R-X, Cao X-L, Chen W-X, Wu J-Z and Huang F (2021) Effect of SYTL3-SLC22A3 Variants, Their Haplotypes, and G × E Interactions on Serum Lipid Levels and the Risk of Coronary Artery Disease and Ischaemic Stroke. Front. Cardiovasc. Med. 8:713068. doi: 10.3389/fcvm.2021.713068

Background: The current study aimed to investigate the effects of synaptotagmin-like 3 (*SYTL3*) and solute carrier family 22 member 3 (*SLC22A3*) single nucleotide polymorphisms (SNPs) and gene-environment (G × E) interactions on blood lipid levels as well as the risk of coronary artery disease (CAD) and ischaemic stroke (IS) in the Southern Chinese Han population.

Methods: The genetic makeup of 6 *SYTL3-SLC22A3* SNPs in 2269 unrelated participants (controls, 755; CAD, 758 and IS, 756) of Chinese Han ethnicity was detected by the next-generation sequencing techniques.

Results: The allele and genotype frequencies of the *SYTL3* rs2129209 and *SLC22A3* rs539298 SNPs were significantly different between the case and control groups. The *SLC22A3* rs539298 SNP was correlated with total cholesterol (TC) levels in controls, the rs539298G allele carriers maintained lower TC levels than the rs539298G allele non-carriers. At the same time, the *SLC22A3* rs539298 SNP interacted with alcohol consumption reduced the risk of CAD and IS. The *SYTL3-SLC22A3* A-C-A-A-A-A, G-T-C-G-C-A and A-T-A-A-C-A haplotypes increased and the A-C-A-A-C-G haplotype reduced the risk of CAD, whereas the *SYTL3-SLC22A3* A-C-A-A-A-A, G-T-C-G-A-G and A-T-A-A-C-A haplotypes increased and the A-C-A-A-A-G and A-C-A-A-C-G haplotypes reduced the risk of IS. In addition, several SNPs interacted with alcohol consumption, body mass index ≥ 24 kg/m^2 and cigarette smoking to affect serum lipid parameters such as triglyceride, high-density lipoprotein cholesterol, TC, and apolipoprotein A1 levels.

Conclusions: Several *SYTL3-SLC22A3* variants, especially the rs539298 SNP, several haplotypes, and G × E interactions, were related to blood lipid parameters and the risk of CAD and IS in the Southern Chinese Han population.

Keywords: coronary artery disease, ischaemic stroke, *SYTL3*, *SLC22A3*, single nucleotide polymorphism, haplotype, lipids

INTRODUCTION

Ischaemic cardiovascular and cerebrovascular diseases, including coronary artery disease (CAD) and ischaemic stroke (IS), are the primary contributors to global disability and death despite the vast therapeutic and diagnostic methods innovated in the last 10 years. Several recent studies have shown that as a multifactorial and complex disorder, CAD or IS occurs due to numerous factors, such as unhealthy lifestyles, alterations of serum lipid levels, genomic background and environmental factors and interactions among these factors (1–3). As the common pathological basis of CAD (4) and IS (5), atherosclerosis is the result of chronic inflammation (6) and abnormal lipid metabolism, such as increased levels of triglyceride (TG) (7), apolipoprotein (Apo) B (8), low-density lipoprotein cholesterol (LDL-C) (9), and total cholesterol (TC) (10), along with reduced levels of ApoA1 (8) and high-density lipoprotein cholesterol (HDL-C) (11) in serum.

Many genes and genetic loci associated with CAD (12) or IS (13) have been identified and reported in previous genome-wide association studies (GWASes). Similarly, a large number of genes or loci associated with lipid metabolism have also been associated with CAD and/or IS (1–3). Currently, several compelling genes closely associated with blood lipid parameters, including synaptotagmin-like 3 (*SYTL3*) and solute carrier family 22-member 3 (*SLC22A3*), have been identified in the European population by GWASes (14). The *SYTL3* (also named *SLP3*, gene ID: 94120, OMIM: 608441, HGNC:15587) is located on chromosome 6q25.3 (exon count: 22), and its encoded protein plays a key role in vesicular trafficking. Several compelling studies have suggested that the transport of lipids and proteins between eukaryotic cells via secretion and endocytosis is primarily facilitated by vesicular trafficking (15, 16). Muse et al. (17) reported that the *SYTL3* was associated with the incidence of acute myocardial infarction (AMI). However, its mechanism remains unclear.

The *SLC22A3* (also known as *OCT3*; *EMTH*; *EMT*, gene ID: 6581, OMIM: 604842, HGNC:10967) is located on chromosome 6q25.3 (exon count: 15), and it encodes an organic cation transporter 3 protein, which is a member of the peripheral membrane protein family and plays a key role in biogenic histamine deactivation and synthesis (18, 19). As an effective proinflammatory mediator, histamine can enhance the deposition of LDL-C in vascular endothelial cells by inducing inflammatory mediators such as adhesion molecules, cytokines, chemokines and others. Inflammatory mediators further increase the permeability of vascular endothelial cells, so histamine can create a favorable environment for the formation of atherosclerotic plaques (20, 21). Previous studies showed that the *SLC22A3* silencing could effectively inhibit the synthesis of histamine and proinflammatory mediators (*MCP-1*, *IL-8* and *IL-6*) and reduce the infiltration of monocytes and the adhesion between leukocytes and the endothelium (22). The *SLC22A3* expression is widespread, and it can be found in skeletal muscle, heart, brain, and placenta. The *SLC22A3* is highly expressed in the human heart, with the strongest *SLC22A3* immunoreactivity found in vascular endothelial cells (23, 24). A series of studies have confirmed that the *SLC22A3* rs1810126, rs2048327 and rs3088442 SNPs reduce the risk of CAD by downregulating *SLC22A3* transcription and protein levels (22, 25). Chen et al. (26) found that the *SLC22A3* could be regarded as a gene vector for the association between lipid metabolism and CAD. Nevertheless, the association between the *SYTL3* rs9364496, rs6455600, rs2129209, rs9456350, *SLC22A3* rs446809 and rs539298 SNPs and the risk of CAD and IS is still unclear and not reported in the Chinese Han population. Thus, this research aimed to understand the relationship between the six selected SNPs and the risk of CAD and IS in the Southern Chinese Han population.

Abbreviations: ANCOVA, covariance analysis; Apo, apolipoprotein; BMI, body mass index; CAD, coronary artery disease; CI, confidence interval; DNA, deoxyribonucleic acid; GWAS, genome-wide association study; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; OR, odds ratio; *SYTL3*, synaptotagmin like 3; *SLC22A3*, solute carrier family 22 member 3; SNP, single nucleotide polymorphisms; T2DM, type 2 diabetes mellitus; TC, total cholesterol; TG, triglyceride; AMI, acute myocardial infarction; MRI, magnetic resonance imaging; MI, myocardial infarction; LD, linkage disequilibrium; NGS, next-generation sequencing; ASW, Americans of African Ancestry in the Southwestern USA; GIH, Gujarati Indians in Houston; LWK, Luhya in Webuye, Kenya; MEX, Mexican ancestry in Los Angeles, California; MEX, Maasai in Kinyawa, Kenya; CHB, Han Chinese in Beijing.

METHODS

Subjects

A case-control study was conducted among the Chinese Han participants in Guangxi Zhuang Autonomous Region, a province in Southern China. A total of 2,269 samples were collected from the First Affiliated Hospital of Guangxi Medical University, including controls, 755; CAD patients, 758; and IS cases, 756. CAD was defined as significant coronary artery stenosis ($\geq$ 50%) in at least one of the three major coronary arteries or their major branches (branch diameter $\geq$ 2 mm) (27). All patients suffering from IS received detailed and rigorous neurological examination as well as brain magnetic resonance imaging (MRI) scans. The diagnostic criteria for IS were based on the International Classification of Diseases (9th Revision). Subjects with a history of neoplastic, autoimmune, type 1 diabetes mellitus, haematologic, thyroid, renal or liver diseases were excluded. Patients suffering from IS had no history of CAD, and patients suffering from CAD also had no history of IS.

After adjusting for sex and age, all of the healthy control participants were randomly recruited from the Physical Examination Center of the First Affiliated Hospital of Guangxi Medical University during the same period. All participants were healthy, and none of them had a history of myocardial infarction (MI), type 2 diabetes mellitus (T2DM), CAD or IS as judged by a critical clinical examination and medical history collection. All subjects in the current research signed written informed consent forms. The research proposal was approved by the Ethics Committee of the First Affiliated Hospital, Guangxi Medical University (No: Lunshen-2014-KY-Guoji-001; Mar. 7, 2014).

Biochemical Index Detection

After fasting for at least 12 h, a sample of 5 ml of peripheral venous blood was drawn from each participant. A portion of the blood sample (2 ml) was used to measure serum lipid levels, and another portion of the blood sample (3 ml) was used to extract genomic DNA. The methods for measuring serum ApoB, LDL-C, ApoA1, TC, HDL-C, and TG were described in detail in our previous study (28). All determinations were performed using an autoanalyser (Type 7170A; Hitachi Ltd., Tokyo, Japan) in the Clinical Science Experiment Center of the First Affiliated Hospital, Guangxi Medical University (29).

Diagnostic Criteria

The definition for normal values of serum ApoB (0.80–1.05 g/L), TG (0.56–1.70 mmol/L), ApoA1 (1.20–1.60 g/L), LDL-C (2.70–3.10 mmol/L), TC (3.10–5.17 mmol/L), HDL-C (1.16–1.42 mmol/L), and the ApoA1/ApoB ratio (1.00–2.50) was based on our previous study (30). The diagnostic criteria for hyperlipidaemia (31), hypertension (32, 33), obesity, overweight, and normal weight (34, 35) were described in detail in previous studies (31–35). Participants with a previous diagnosis of diabetes or a fasting plasma glucose $\geq$ 7.0 mmol/L or 2 h postprandial plasma glucose $\geq$ 11.1 mmol/L were defined as diabetic patients (36).

SNP Selection and Genotyping

Six SNPs located on *SYTL3* and *SLC22A3* were chosen based on predetermined criteria: (1) The *SYTL3-SLC22A3* cluster was selected from previous GWASes associated with blood lipid levels. (2) Tagging SNPs were identified *via* Haploview (Broad Institute of MIT and Harvard, USA, version 4.2), and potential lipid metabolism-associated functional SNPs were predicted using the latest version of the 1000 Genome Project Database. (3) More complete details of the selected SNPs were collected from NCBI dbSNP Build 132. (4) Regarding SNP selection, we also referenced a previous study by Ober et al. (14), and the minor allele frequency (MAF) of all SNPs was more than 1% and associated with blood lipid profiles in a previous study. (5) Six SNPs of *SYTL3* (rs9364496, rs6455600, rs2129209 and rs9456350) and *SLC22A3* (rs446809 and rs539298) were selected using the block-based method, which involves marking the association of linkage disequilibrium (LD) amongst the chosen SNPs ($r^2 > 0.8$). White blood cells were used for genomic DNA extraction with phenol-chloroform (37). All obtained DNA samples were numbered and maintained at −20°C until further studies were carried out. Next-generation sequencing technology (NGS) was used to analyse the genotypes of the 6 selected SNPs at the Center for Human Genetics Research, Shanghai Genesky Bio-Tech Co. Ltd., China (38). **Supplementary Table 1** depicts the relevant primer sequences. Detailed steps for multiplex PCR and high throughput sequencing are also shown in **Supplementary Material**.

Statistical Analyses

SPSS (Version 22.0) software was used to analyse all data collected from the current research. An independent sample *t*-test was used to evaluate the continuous data (mean ± SD) that were normally distributed between control and patient groups. TG levels that were not normally distributed were expressed using median and quartile ranges and were evaluated using the Wilcoxon-Mann-Whitney test. Data such as the genotype distribution, sex ratio, and the number of smokers and drinkers

TABLE 1 | Comparison of demographic, lifestyle characteristics, and serum lipid levels of the participants.

Characteristic	Control (n = 755)	Case CAD (n = 758)	IS (n = 756)	P Control vs. CAD	P Control vs. IS
Male/female	575/180	553/205	602/154	0.153	0.104
Age (years)	61.86 ± 9.96	62.63 ± 9.46	62.36 ± 10.97	0.270	0.354
BMI (kg/m²)	22.57 ± 2.94	23.60 ± 2.96	23.61 ± 3.48	0.000	0.000
Smoking, *n* %	322 (42.6)	346 (45.6)	354 (46.8)	0.240	0.103
Alcohol, *n* %)	349 (46.2)	205 (27.0)	212 (28.0)	0.000	0.000
SBP (mmHg)	128.69 ± 19.29	134.27 ± 23.05	147.86 ± 21.79	0.000	0.000
DBP (mmHg)	80.84 ± 11.56	79.89 ± 13.85	82.00 ± 12.60	0.144	0.064
PP (mmHg)	47.84 ± 14.48	54.39 ± 18.27	65.86 ± 18.07	0.000	0.000
Glu (mmol/L)	6.00 ± 1.57	6.32 ± 1.54	6.19 ± 1.43	0.000	0.014
TC (mmol/L)	4.89 ± 0.96	4.49 ± 0.99	4.54 ± 1.03	0.000	0.002
TG (mmol/L)	1.03 (0.73)	1.30 (0.93)	1.41 (0.97)	0.000	0.000
HDL-C (mmol/L)	1.91 ± 0.51	1.13 ± 0.34	1.22 ± 0.43	0.000	0.000
LDL-C (mmol/L)	2.74 ± 0.78	2.71 ± 0.97	2.77 ± 0.91	0.480	0.577
ApoA1 (g/L)	1.41 ± 0.27	1.03 ± 0.36	1.02 ± 0.22	0.000	0.000
ApoB (g/L)	0.89 ± 0.21	0.89 ± 0.26	0.91 ± 0.24	0.739	0.078
ApoA1/ApoB	1.68 ± 0.64	1.27 ± 0.79	1.21 ± 0.48	0.000	0.000

SBP, systolic blood pressure; DBP, diastolic blood pressure; PP, pulse pressure; Glu, glucose; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; Apo, apolipoprotein; TC, total cholesterol; TG, triglyceride. The value of triglyceride was presented as median (interquartile range), the difference between the control and CAD/IS groups was determined by the Wilcoxon-Mann-Whitney test.

were analyzed by the chi-square test. The standard goodness-of-fit test was used to test the Hardy-Weinberg equilibrium (HWE). The correlation between genotypes and blood lipid levels was tested by the covariance analysis (ANCOVA). Any variants related to the blood lipid parameter at a value of $P \leq 0.008$ were considered statistically significant after Bonferroni correction. Age, alcohol consumption, sex, cigarette smoking and body mass index (BMI) were adjusted for the statistical analysis. Unconditional logistic regression analysis was used to calculate the 95% confidence interval (CI) and the odds ratio (OR). The interactions of the six selected SNPs with cigarette smoking, sex, age, alcohol consumption and BMI $\geq$ 24 kg/m^2 on blood lipid levels and the risk of CAD and IS were performed by using a factorial regression analysis after

TABLE 2 | Genotype and allele frequencies of the 6 *SYTL3- SLCC2A3* SNPs in cases and controls [*n*(%)].

Genotype	Control (*n* 755)	CAD (*n* = 758)	IS (*n* = 756)	Allele	Control (*n* = 1,510)	CAD (*n* = 1,516)	IS (*n* = 1,512)
rs9364496							
AA	481 (63.7)	448 (59.1)	486 (64.3)				
AG	237 (31.4)	276 (36.4)	234 (30.9)	A	1199 (79.4)	1172 (77.3)	1206 (79.8)
GG	37 (4.9)	34 (4.5)	36 (4.8)	G	311 (21.6)	344 (23.7)	306 (20.2)
x^2		4.258	0.058			1.958	0.060
P		0.119	0.971			0.162	0.807
P_{HWE}	0.268	0.298	0.257				
rs6455600							
CC	365 (48.3)	364 (48.0)	356 (47.1)				
CT	332 (44.0)	332 (43.8)	327 (43.2)	C	1062 (70.3)	1060 (69.9)	1039 (68.7)
TT	58 (7.7)	62 (8.2)	73 (9.7)	T	448 (29.7)	456 (30.1)	473 (31.3)
x^2		0.129	1.867			0.061	0.929
P		0.938	0.393			0.805	0.335
P_{HWE}	0.140	0.256	0.868				
rs2129209							
AA	375 (49.7)	427 (56.3)	423 (56.0)				
AC	328 (43.4)	282 (37.2)	294 (38.9)	A	1078 (71.4)	1136 (74.9)	1140 (75.4)
CC	52 (6.9)	49 (6.5)	39 (5.2)	C	432 (28.6)	380 (25.1)	372 (24.6)
x^2		6.924	6.602			4.838	21.577
P		0.031	0.037			0.028	0.001
P_{HWE}	0.081	0.791	0.185				
rs9456350							
AA	572 (75.8)	595 (78.5)	568 (75.1)				
AG	173 (22.9)	153 (20.2)	178 (23.5)	A	1317 (87.2)	1343 (88.6)	1314 (86.9)
GG	10 (1.3)	10 (1.3)	10 (1.3)	G	193 (12.8)	173 (10.4)	198 (12.1)
x^2		1.674	0.085			1.335	0.066
P		0.433	0.959			0.248	0.797
P_{HWE}	0.446	0.963	0.343				
rs446809							
CC	378 (50.1)	370 (48.8)	360 (47.6)				
CA	315 (41.7)	318 (42.0)	331 (43.8)	C	1071 (70.9)	1058 (69.8)	1051 (69.5)
AA	62 (8.2)	70 (9.2)	65 (8.6)	A	439 (28.1)	458 (29.2)	461 (30.5)
x^2		0.579	0.906			1.728	1.454
P		0.749	0.636			0.189	0.228
P_{HWE}	0.749	0.888	0.365				
rs539298							
AA	383 (50.7)	439 (57.9)	452 (59.8)				
AG	317 (42.0)	279 (36.8)	265 (35.0)	A	1083 (71.7)	1157 (76.3)	1169 (77.3)
GG	55 (7.3)	40 (5.3)	39 (5.2)	G	427 (28.3)	359 (23.7)	343 (22.7)
x^2		8.600	13.242			8.316	12.447
P		0.014	0.001			0.004	0.000
P_{HWE}	0.335	0.614	0.984				

CAD, coronary artery disease; IS, ischemic stroke.

controlling for potential confounders. A $P \leq 0.00125$ was considered statistically significant after Bonferroni correction. Haploview (Broad Institute of MIT and Harvard, USA, version 4.2) software was used to calculate the haplotype frequencies and pair wise LD among the six detected SNPs. The heatmap of the inter-locus models was measured by using R software (version 4.1.0).

RESULTS

Common and Biochemical Characteristics

As presented in **Table 1**, there were no differences in the proportion of smokers, diastolic blood pressure, age, the sex ratio, or serum ApoB and LDL-C levels between the controls and cases. The glucose, pulse pressure, systolic blood pressure,

TABLE 3 | Genotypes of the six *SYTL3-SLC22A3* SNPs and the risk of CAD and IS.

SNP/model	Ref. genotype	Effect genotype	CAD (OR 95% CI)	P_{CAD}	IS (OR 95% CI)	P_{IS}
rs9364496						
Codominant	A/A	A/G	1.26 (1.01–1.59)	0.12	0.99 (0.79–1.25)	0.99
		G/G	1.01 (0.61–1.67)		0.96 (0.59–1.58)	
Dominant	A/A	A/G+G/G	1.23 (0.99–1.53)	0.064	0.99 (0.79–1.23)	0.93
Recessive	A/A+A/G	G/G	0.92 (0.56–1.52)	0.76	0.96 (0.59–1.58)	0.89
Overdominant	A/A+G/G	A/G	1.26 (1.01–1.58)	0.041	1.00 (0.79–1.25)	0.98
Log-additive			1.14 (0.95–1.37)	0.15	0.99 (0.82–1.18)	0.9
rs6455600						
Codominant	C/C	C/T	1.09 (0.87–1.36)	0.74	1.01 (0.81–1.27)	0.61
		T/T	1.10 (0.73–1.64)		1.22 (0.82–1.81)	
Dominant	C/C	C/T+T/T	1.09 (0.88–1.35)	0.44	1.05 (0.85–1.29)	0.68
Recessive	C/C+C/T	T/T	1.06 (0.71–1.57)	0.78	1.21 (0.83–1.77)	0.32
Overdominant	C/C+T/T	C/T	1.07 (0.86–1.33)	0.53	0.98 (0.79–1.22)	0.89
Log-additive			1.06 (0.90–1.26)	0.47	1.07 (0.90–1.26)	0.45
rs2129209						
Codominant	A/A	A/C	0.82 (0.65–1.02)	0.18	0.81 (0.65–1.02)	0.063
		C/C	0.82 (0.53–1.27)		0.65 (0.41–1.03)	
Dominant	A/A	A/C+C/C	0.82 (0.66–1.01)	0.062	0.79 (0.64–0.98)	0.031
Recessive	A/A+A/C	C/C	0.90 (0.59–1.37)	0.62	0.71 (0.46–1.12)	0.14
Overdominant	A/A+C/C	A/C	0.83 (0.67–1.04)	0.1	0.85 (0.69–1.05)	0.14
Log-additive			0.86 (0.73–1.02)	0.086	0.81 (0.68–0.97)	0.019
rs9456350						
Codominant	A/A	A/G	0.89 (0.69–1.16)	0.67	0.99 (0.77–1.28)	0.96
		G/G	0.85 (0.34–2.12)		1.13 (0.46–2.81)	
Dominant	A/A	A/G+G/G	0.89 (0.69–1.15)	0.37	1.00 (0.78–1.28)	0.99
Recessive	A/A+A/G	G/G	0.87 (0.35–2.17)	0.76	1.13 (0.46–2.81)	0.79
Overdominant	A/A+G/G	A/G	0.90 (0.69–1.16)	0.41	0.99 (0.77–1.28)	0.95
Log-additive			0.90 (0.72–1.13)	0.37	1.01 (0.80–1.27)	0.94
rs446809						
Codominant	C/C	C/A	1.15 (0.92–1.45)	0.064	1.20 (0.96–1.51)	0.044
		A/A	1.58 (1.06–2.36)		1.65 (1.08–2.53)	
Dominant	C/C	C/A+A/A	1.22 (0.98–1.52)	0.08	1.25 (1.00–1.57)	0.046
Recessive	C/C+C/A	A/A	1.47 (1.01–2.15)	0.046	1.48 (0.99–2.22)	0.054
Overdominant	C/C+A/A	C/A	1.07 (0.86–1.33)	0.55	1.11 (0.89–1.37)	0.37
Log-additive			1.22 (1.03–1.44)	0.024	1.25 (1.05–1.49)	0.014
rs539298						
Codominant	A/A	A/G	0.77 (0.62–0.96)	0.0082	0.67 (0.53–0.84)	0.0511
		G/G	0.56 (0.36–0.88)		0.66 (0.42–1.04)	
Dominant	A/A	A/G+G/G	0.74 (0.60–0.91)	0.0053	0.67 (0.54–0.83)	2E-04
Recessive	A/A+A/G	G/G	0.63 (0.40–0.98)	0.038	0.79 (0.51–1.22)	0.28
Overdominant	A/A+G/G	A/G	0.82 (0.66–1.02)	0.07	0.70 (0.56–0.87)	0.0012
Log-additive			0.76 (0.64–0.91)	0.002	0.74 (0.62–0.88)	6E-04

SNP, single nucleotide polymorphism; CAD, coronary artery disease; IS, ischemic stroke. Adjusted for sex, age, smoking, drinking, BMI, diabetes, hypertension.

BMI, serum TG levels were significantly lower, and the serum TC, ApoA1, and HDL-C levels; the proportion of drinkers; and the ApoA1/ApoB ratio were significantly higher in controls than in CAD and IS patients.

Genotypic and Allelic Frequencies

As shown in **Table 2**, all detected SNPs in the CAD, IS and control groups conformed to HWE ($P > 0.05$). The frequencies of the rs2129209AC/CC genotypes and the rs2129209C allele were lower in CAD (AC/CC, 43.7%; C, 25.1%) and IS (AC/CC, 44.1%; C, 24.6%) patients than in control participants (AC/CC, 50.3%; C, 28.6%, $P < 0.05$-0.01 for all). The frequencies of the rs539298AG/GG genotypes and rs539298G allele were also lower in CAD (AG/GG, 42.1%; G, 23.7%) and IS (AG/GG, 40.2%; G, 22.7%) patients than in control subjects (AG/GG, 49.3%; G, 28.3%, $P < 0.05$-0.001 for all). The specific and detailed genotype distribution of the six selected SNPs in the control, CAD, and IS groups is shown in **Supplementary Material**.

Genotypes and the Risk of Diseases

As described in **Table 3**, only the rs539298 SNP decreased the risk of CAD and IS ($P < 0.008$ was considered statistically significant after Bonferroni correction). The dominant model (AG/GG *vs.* AA) for CAD group showed OR = 0.74, 95% CI = 0.60–0.91, $P = 0.0053$; and the log-additive model (A vs. G) illustrated OR = 0.76, 95% CI = 0.64–0.91, $P = 0.002$. The dominant model (AG/GG *vs.* AA) for IS group displayed OR = 0.67, 95% CI = 0.54–0.83, $P = 2E\text{-}04$; and the log-additive model (A vs. G) revealed OR = 0.74, 95% CI = 0.62–0.88, $P = 6E\text{-}04$.

Interaction Between the rs539298 SNP and Environmental Factors on the Risk of Diseases

As shown in **Figure 1**, the interaction between the rs539298AG/GG genotypes and alcohol consumption reduced the risk of CAD (OR = 0.53, 95% CI = 0.37–0.77, $P < 0.01$) and IS (OR = 0.44, 95% CI = 0.30–0.65, $P < 0.01$).

Haplotypes and the Risk of Diseases

As presented in **Figure 2**, strong LD was found among the *SYTL3* rs9364496, rs6455600, rs2129209 and rs9456350 SNPs as well as the *SLC22A3* rs446809 and rs539298 SNPs in the CAD (**Figure 2A**) and IS (**Figure 2B**) groups ($D' = 0.64$–0.95).

Haplotype analysis among the six selected SNPs showed that the main haplotype was the *SYTL3-SLC22A3* A-C-A-A-C-A (**Table 4**). The *SYTL3-SLC22A3* A-C-A-A-A-A (adjusted OR = 2.19, 95% CI = 1.49–3.21, $P = 1E\text{-}04$), G-T-C-G-C-A (adjusted OR = 1.82, 95% CI = 1.07–3.10, $P = 0.026$) and A-T-A-A-C-A (adjusted OR = 2.04, 95% CI = 1.14–3.66, $P = 0.016$) haplotypes increased the risk of CAD, while the *SYTL3-SLC22A3* A-C-A-A-C-G (adjusted OR = 0.52, 95% CI = 0.34–0.78, $P = 0.0016$) haplotype reduced the risk of CAD. Meanwhile, the *SYTL3-SLC22A3* A-C-A-A-A-A (adjusted OR = 1.98, 95% CI = 1.39–2.84, $P = 2E\text{-}04$), G-T-C-G-A-G (adjusted OR = 2.79, 95% CI = 1.87–4.15, $P < 0.0001$) and A-T-A-A-C-A (adjusted OR = 3.24, 95% CI = 1.92–5.48, $P < 0.0001$) haplotypes increased the risk of IS, whereas the *SYTL3-SLC22A3* A-C-A-A-A-G (adjusted OR = 0.63, 95% CI = 0.42–0.95, $P = 0.027$) and A-C-A-A-C-G (adjusted OR = 0.59, 95% CI = 0.40–0.88, $P = 0.0099$) haplotypes reduced the risk of IS.

Interactions Between the Haplotypes and Environmental Factors on the Risk of Diseases

The interactions between the *SYTL3-SLC22A3* haplotypes and several environmental factors on the risk of CAD and IS are shown in **Figure 3**. The interactions of the *SYTL3-SLC22A3* A-C-A-A-A-A-smoking (adjusted OR = 3.70, 95% CI = 2.29–5.96, $P < 0.01$), *SYTL3-SLC22A3* A-C-A-A-A-A-age > 75 (adjusted OR = 1.91, 95% CI = 1.28–2.84, $P < 0.05$), *SYTL3-SLC22A3* G-T-C-G-C-A-smoking (adjusted OR = 2.25, 95% CI = 1.43–3.54, $P < 0.01$), *SYTL3-SLC22A3* G-T-C-G-C-A-BMI $\geq 24\ kg/m^2$ (adjusted OR = 2.80, 95% CI = 1.78–4.40, $P < 0.01$), *SYTL3-SLC22A3* A-T-A-A-C-A-age > 75 (adjusted OR = 2.72, 95% CI = 1.62–4.56, $P < 0.01$), and *SYTL3-SLC22A3* A-T-A-A-C-A-smoking (adjusted OR = 2.38, 95% CI = 1.27–4.43, $P < 0.01$) increased the risk of CAD,

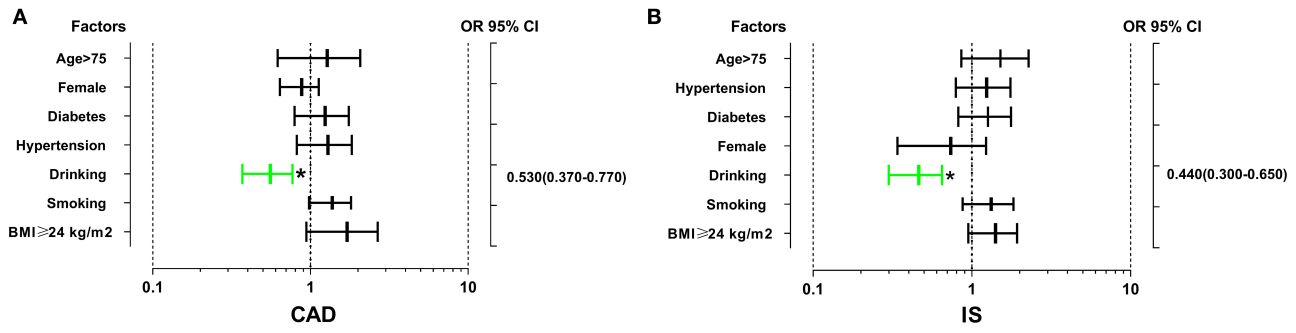


FIGURE 1 | The interactions of the *SLC22A3* rs539298 SNP and diabetes, drinking, hypertension, smoking, age, sex, and BMI on the risk of CAD and IS. The rs539298AG/GG genotypes interacted with alcohol consumption to decrease the risk of CAD **(A)** (OR = 0.53, 95% CI = 0.37–0.77) and IS **(B)** (OR = 0.44, 95% CI = 0.30–0.65). *$P < 0.01$.

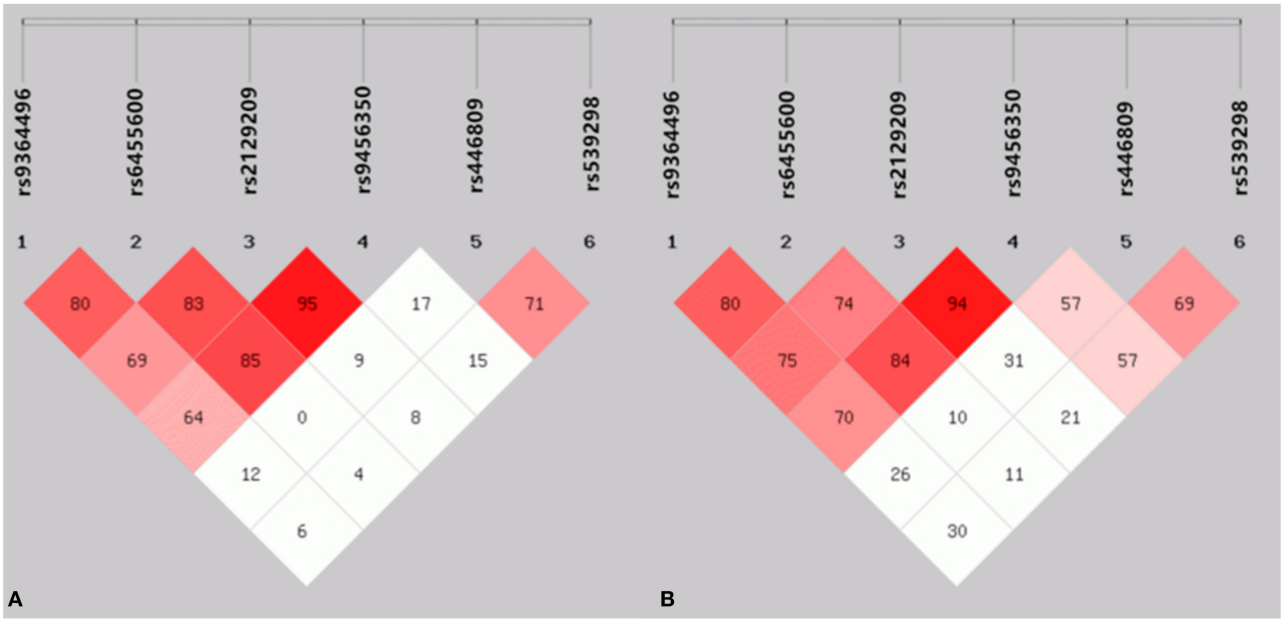


FIGURE 2 | The linkage disequilibrium (LD) represents pair wise *D'* × 100 in the CAD **(A)** and IS **(B)** groups.

TABLE 4 | Haplotype frequencies of the six *SYTL3-SLC22A3* SNPs and the risk of CAD and IS.

Haplotype	Control frequency	CAD			IS		
		Frequency	OR (95% CI)	*P*	Frequency	OR (95% CI)	*P*
A-C-A-A-C-A	0.4458	0.4446	1.00	–	0.4813	1.00	–
A-C-A-A-A-G	0.1052	0.1242	1.32 (0.94–1.84)	0.11	0.0528	0.63 (0.42–0.95)	0.027
A-C-A-A-A-A	0.0362	0.0645	2.19 (1.49–3.21)	1E-04	0.0641	1.98 (1.39–2.84)	2E-04
G-T-C-A-C-A	0.0398	0.0573	1.18 (0.78–1.79)	0.43	0.0357	0.94 (0.58–1.54)	0.820
G-T-C-G-A-G	0.0448	0.0403	0.97 (0.61–1.56)	0.91	0.1175	2.79 (1.87–4.15)	<0.0001
A-C-A-A-C-G	0.0486	0.0236	0.52 (0.34–0.78)	0.0016	0.0291	0.59 (0.40–0.88)	0.0099
G-T-C-G-C-A	0.0264	0.0472	1.82 (1.07–3.10)	0.026	0	–	–
A-T-C-A-C-A	0.0330	0.0366	1.13 (0.69–1.85)	0.64	0.0341	0.75 (0.46–1.23)	0.25
A-T-A-A-C-A	0.0195	0.0359	2.04 (1.14–3.66)	0.016	0.0753	3.24 (1.92–5.48)	<0.0001

CAD, coronary artery disease; IS, ischemic stroke. The haplotypes consist of six alleles in the order of rs9364496, rs6455600, rs2129209, rs9456350, rs446809, and rs539298 SNPs. Adjusted for sex, age, smoking, drinking, BMI, diabetes, hypertension.

whereas the interaction of the *SYTL3-SLC22A3* A-C-A-A-C-G-drinking (adjusted OR = 0.31, 95% CI = 0.14–0.67, $P < 0.01$) decreased the risk of CAD.

The interactions of the *SYTL3-SLC22A3* A-C-A-A-A-A-hypertension (adjusted OR = 2.27, 95% CI = 1.32–3.90, $P < 0.01$), *SYTL3-SLC22A3* A-C-A-A-A-A-smoking (adjusted OR = 2.19, 95% CI = 1.49–3.21, $P < 0.01$), *SYTL3-SLC22A3* A-C-A-A-C-G-hypertension (adjusted OR = 1.18, 95% CI = 1.23–2.67, $P < 0.05$), *SYTL3-SLC22A3* A-T-A-A-C-A-age > 75 (adjusted OR = 3.22, 95% CI = 1.74–5.98, $P < 0.01$), *SYTL3-SLC22A3* A-T-A-A-C-A-smoking (adjusted OR = 3.67, 95% CI = 1.99–6.67, $P < 0.01$) increased the risk of IS, whereas the interactions of the *SYTL3-SLC22A3* A-C-A-A-A-G-female (adjusted OR = 0.45, 95% CI = 0.24–0.82, $P < 0.05$), *SYTL3-SLC22A3* A-C-A-A-A-G-drinking (adjusted OR = 0.26, 95% CI = 0.14–0.49, $P < 0.01$), and *SYTL3-SLC22A3* A-C-A-A-C-G-drinking (adjusted OR = 0.37, 95% CI = 0.22–0.65, $P < 0.01$) decreased the risk of IS. However, there was no significant interaction between the *SYTL3-SLC22A3* G-T-C-G-A-G haplotype and several environmental factors including gender, age, diabetes, hypertension, BMI, smoking and drinking.

Relationship Between Genotypes and Serum Lipid Parameters

The correlation between the *SYTL3-SLC22A3* SNPs and serum lipid parameters in the control group is shown in **Figure 4**. There were significant differences in TC levels between the rs539298AA and rs539298AG/GG genotypes ($P < 0.008$ was considered statistically significant after Bonferroni correction), the subjects with rs539298AG/GG genotypes had lower serum

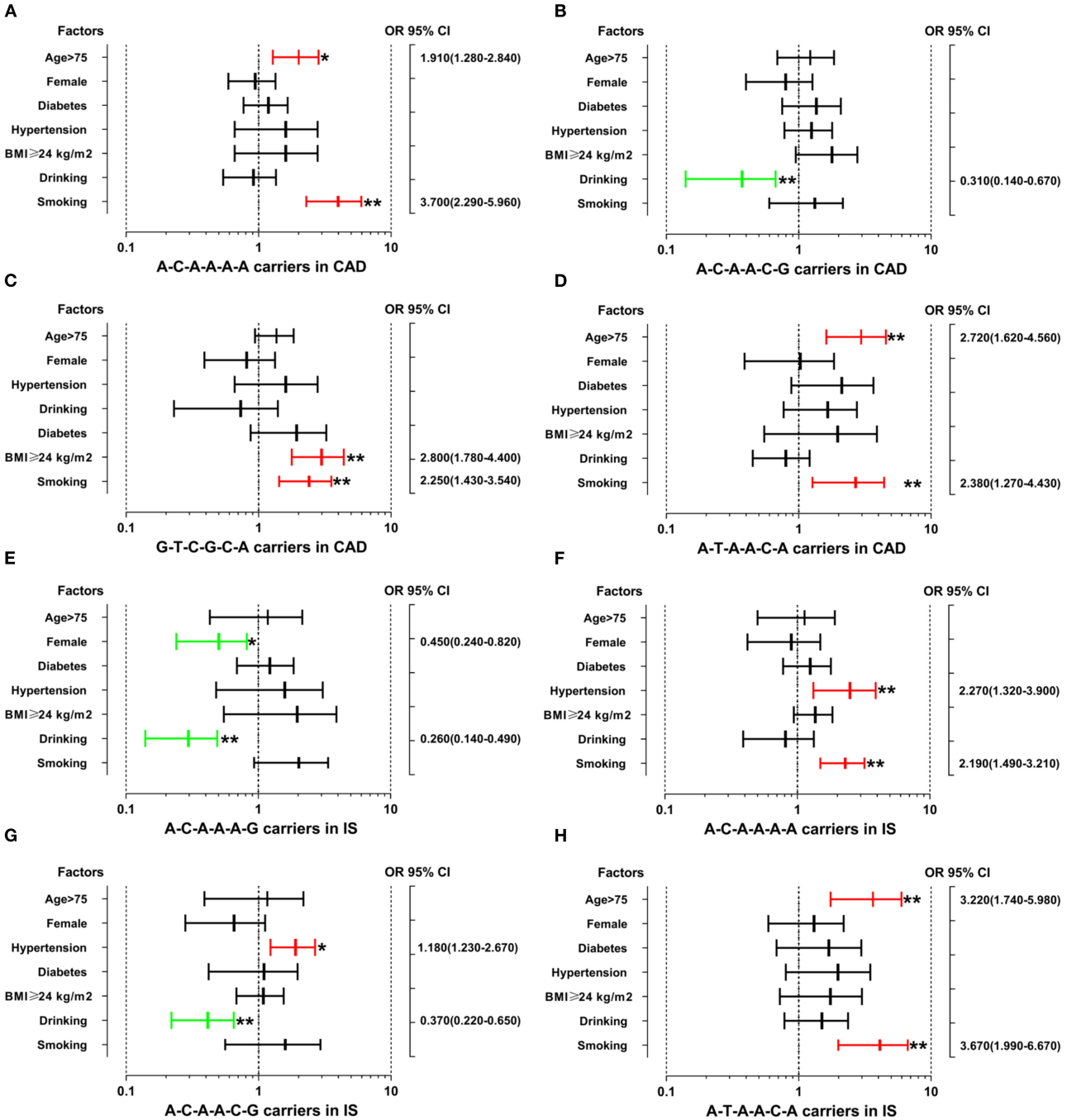


FIGURE 3 | The interactions between the *SYTL3-SLC22A3* haplotypes and the risk of CAD and IS. CAD, coronary artery disease; IS, ischaemic stroke. **(A)** The *SYTL3-SLC22A3* A-C-A-A-A-A haplotype interacted with age > 75 and smoking to increase the risk of CAD. **(B)** The interaction between the *SYTL3-SLC22A3* A-C-A-A-C-G haplotype and alcohol consumption decreased the risk of CAD. **(C)** The interactions between *SYTL3-SLC22A3* G-T-C-G-C-A and cigarette smoking and BMI ≥ 24 kg/m^2 increased the risk of CAD. **(D)** The *SYTL3-SLC22A3* A-T-A-A-C-A haplotype interacted with age > 75 and smoking to increase the risk of CAD. **(E)** The interactions between the *SYTL3-SLC22A3* A-C-A-A-A-G haplotype and female and alcohol consumption decreased the risk of IS. **(F)** The *SYTL3-SLC22A3* A-C-A-A-A-A haplotype interacted with hypertension and smoking to increase the risk of IS. **(G)** The interactions between the *SYTL3-SLC22A3* A-C-A-A-C-G haplotype and hypertension increased the risk of IS, and alcohol consumption decreased the risk of IS. **(H)** The interactions between the *SYTL3-SLC22A3* A-T-A-A-C-A haplotype and age > 75 and smoking increased the risk of IS. *P < 0.05 and **P < 0.01.

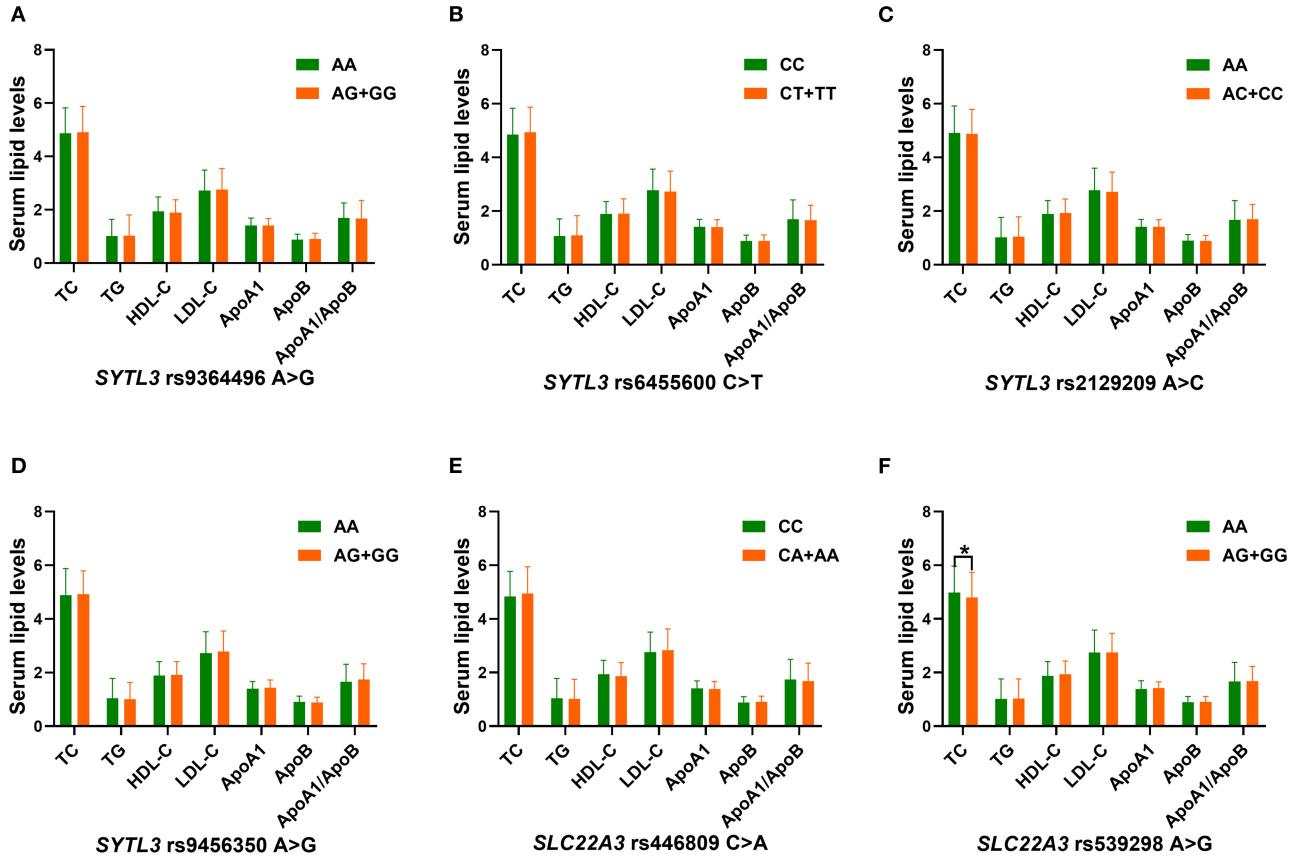


FIGURE 4 | The association between the genotypes of six *SYTL3-SLC22A3* SNPs and blood lipid levels in the control group. HDL-C, high-density lipoprotein cholesterol; TC, total cholesterol; ApoA1, apolipoprotein A1; TG, triglyceride; ApoB, apolipoprotein B; LDL-C, low-density lipoprotein cholesterol; ApoA1/ApoB, the ratio of apolipoprotein A1 to apolipoprotein B. A $^*P < 0.008$ (after adjusting for 6 independent tests by the Bonferroni correction) was considered statistically significant. **(A)**, rs9364496; **(B)**, rs6455600; **(C)**, rs2129209; **(D)**, rs9456350; **(E)**, rs446809; and **(F)**, rs539298 SNPs.

TC levels than those with rs539298AA genotype. However, there was no significant correlation between the remaining 5 SNPs and serum lipid levels in the control group ($P \geq 0.008$ for all).

Interactions Between the *SYTL3-SLC22A3* SNPs and Several Environmental Factors on Serum Lipid Parameters and the Risk of CAD/IS

As listed in **Table 5**, several interactions between the *SYTL3-SLC22A3* SNPs and some environmental factors on serum lipid parameters were detected in the control group. The interactions of the *SYTL3* rs6455600-smoking, *SYTL3* rs6455600-BMI, and *SYTL3* rs9456350-BMI affected TG levels. The interaction of the *SYTL3* rs2129209- and *SLC22A3* rs539298-BMI influenced TC levels; and the interaction of the *SYTL3* rs9364496-smoking or *SYTL3* rs9364496-BMI influenced TG or TC levels. The interaction of the *SYTL3* rs2129209-, *SYTL3* rs9456350- and *SLC22A3* rs446809-alcohol consumption influenced HDL-C levels; and the interaction of the *SLC22A3* rs539298-alcohol consumption affected HDL-C and ApoA1 levels.

As presented in **Figure 5**, the interactions of the *SYTL3* rs6455600CT/TT-BMI $\geq$ 24 kg/m^2 or smoking, and *SYTL3* rs9456350AG/GG-BMI $\geq$ 24 kg/m^2 increased TG levels. The interactions of the *SLC22A3* rs539298AG/GG-BMI $<$ 24 kg/m^2, and *SYTL3* rs2129209AC/CC-BMI $<$ 24 kg/m^2 reduced TC levels; and the interaction of *SYTL3* rs9364496AG/GG-BMI $\geq$ 24 kg/m^2 or smoking increased TC or TG levels, respectively. The interactions of the *SYTL3* rs9456350AG/GG-, *SLC22A3* rs446809CA/AA-, and *SYTL3* rs2129209AC/CC-alcohol consumption increased HDL-C levels; and the interaction of the *SLC22A3* rs539298AG/GG-alcohol consumption increased HDL-C and ApoA1 levels.

Relative Factors for Serum Lipid Parameters

As shown in **Figure 6**, Pearson correlation analysis revealed that there were significant correlations between the 6 SNPs and several environmental factors, including alcohol consumption, BMI, cigarette smoking, blood glucose, age, blood pressure, sex and serum lipid profiles, in the CAD (**Figure 6A**) and IS groups (**Figure 6B**).

TABLE 5 | The *P*-values for the interactions of genotypes and gender, age, drinking, smoking, and BMI on serum lipid levels and the risk of CAD and IS.

SNP/factor	Lipid							CAD	IS
	TC	TG	HDL-C	LDL-C	ApoA1	ApoB	ApoA1/ApoB		
rs9364496									
Gender	0.622	0.028	0.004	0.022	0.022	0.003	0.031	0.052	0.089
Age	0.793	0.153	0.645	0.478	0.478	0.275	0.444	0.550	0.430
Smoking	0.004	0.000	0.003	0.002	0.002	0.877	0.136	0.174	0.156
Drinking	0.399	0.020	0.956	0.040	0.023	0.278	0.401	0.022	0.017
BMI	0.000	0.003	0.092	0.006	0.006	0.610	0.083	0.019	0.005
rs6455600									
Gender	0.655	0.004	0.004	0.022	0.002	0.071	0.388	0.286	0.366
Age	0.684	0.056	0.828	0.681	0.228	0.992	0.374	0.216	0.164
Smoking	0.473	0.000	0.089	0.002	0.424	0.293	0.013	0.251	0.363
Drinking	0.359	0.030	0.042	0.003	0.056	0.906	0.374	0.221	0.265
BMI	0.006	0.001	0.304	0.010	0.144	0.101	0.004	0.010	0.131
rs2129209									
Gender	0.364	0.063	0.012	0.008	0.235	0.021	0.659	0.605	0.005
Age	0.900	0.428	0.688	0.497	0.198	0.479	0.266	0.004	0.407
Smoking	0.849	0.005	0.105	0.011	0.798	0.296	0.061	0.880	0.654
Drinking	0.373	0.002	0.000	0.023	0.102	0.500	0.115	0.020	0.189
BMI	0.001	0.013	0.207	0.006	0.235	0.082	0.006	0.016	0.031
rs9456350									
Gender	0.281	0.098	0.021	0.017	0.013	0.026	0.285	0.809	0.162
Age	0.916	0.561	0.084	0.637	0.083	0.679	0.230	0.039	0.281
Smoking	0.933	0.027	0.010	0.023	0.126	0.293	0.027	0.494	0.037
Drinking	0.108	0.010	0.000	0.020	0.017	0.412	0.047	0.008	0.030
BMI	0.011	0.001	0.032	0.006	0.106	0.072	0.005	0.130	0.071
rs446809									
Gender	0.314	0.090	0.009	0.005	0.004	0.011	0.013	0.890	0.759
Age	0.215	0.266	0.286	0.009	0.170	0.191	0.009	0.768	0.298
Smoking	0.294	0.032	0.054	0.007	0.463	0.018	0.024	0.281	0.381
Drinking	0.057	0.040	0.001	0.012	0.183	0.049	0.003	0.113	0.167
BMI	0.004	0.013	0.112	0.021	0.225	0.010	0.004	0.005	0.076
rs539298									
Gender	0.678	0.002	0.004	0.032	0.002	0.075	0.630	0.781	0.738
Age	0.385	0.131	0.052	0.839	0.033	0.928	0.343	0.006	0.180
Smoking	0.066	0.689	0.030	0.011	0.193	0.374	0.089	0.216	0.143
Drinking	0.097	0.447	0.000	0.003	0.000	0.890	0.103	0.000	0.000
BMI	0.000	0.002	0.045	0.010	0.058	0.003	0.007	0.334	0.200

SNP, single nucleotide polymorphism; TC, total cholesterol; TG, triglyceride; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; ApoA1, apolipoprotein A1; ApoB, apolipoprotein B; CAD, coronary artery disease; IS, ischemic stroke; BMI, body mass index. $P \leq 0.00125$ was considered statistically significant after Bonferroni correction.

DISCUSSION

The main findings of the current research included the following aspects: (1) The allelic and genotypic frequencies of the *SYTL3* rs2129209 and *SLC22A3* rs539298 SNPs were different between controls and CAD/IS patients. The rs2129209AC/CC genotype and rs2129209C allele frequencies as well as the rs539298AG/GG genotype and rs539298G allele frequencies were lower in CAD/IS patients than in controls, respectively. (2) The *SLC22A3* rs539298 SNP was correlated with TC levels in the control group, the rs539298G allele carriers had lower TC levels than the rs539298G allele non-carriers. At the same time, the *SLC22A3* rs539298 SNP interacted with alcohol consumption would reduce the risk of CAD and IS. (3) Several haplotypes were associated with an increased or decreased risk of CAD and IS, and the haplotypes could explain more changes in the risk of CAD/IS than any single SNP alone. In addition, the interactions between several haplotypes and different environmental factors on the onset of CAD and IS were also observed. (4) The interactions between the *SYTL3-SLC22A3* SNPs and several environmental

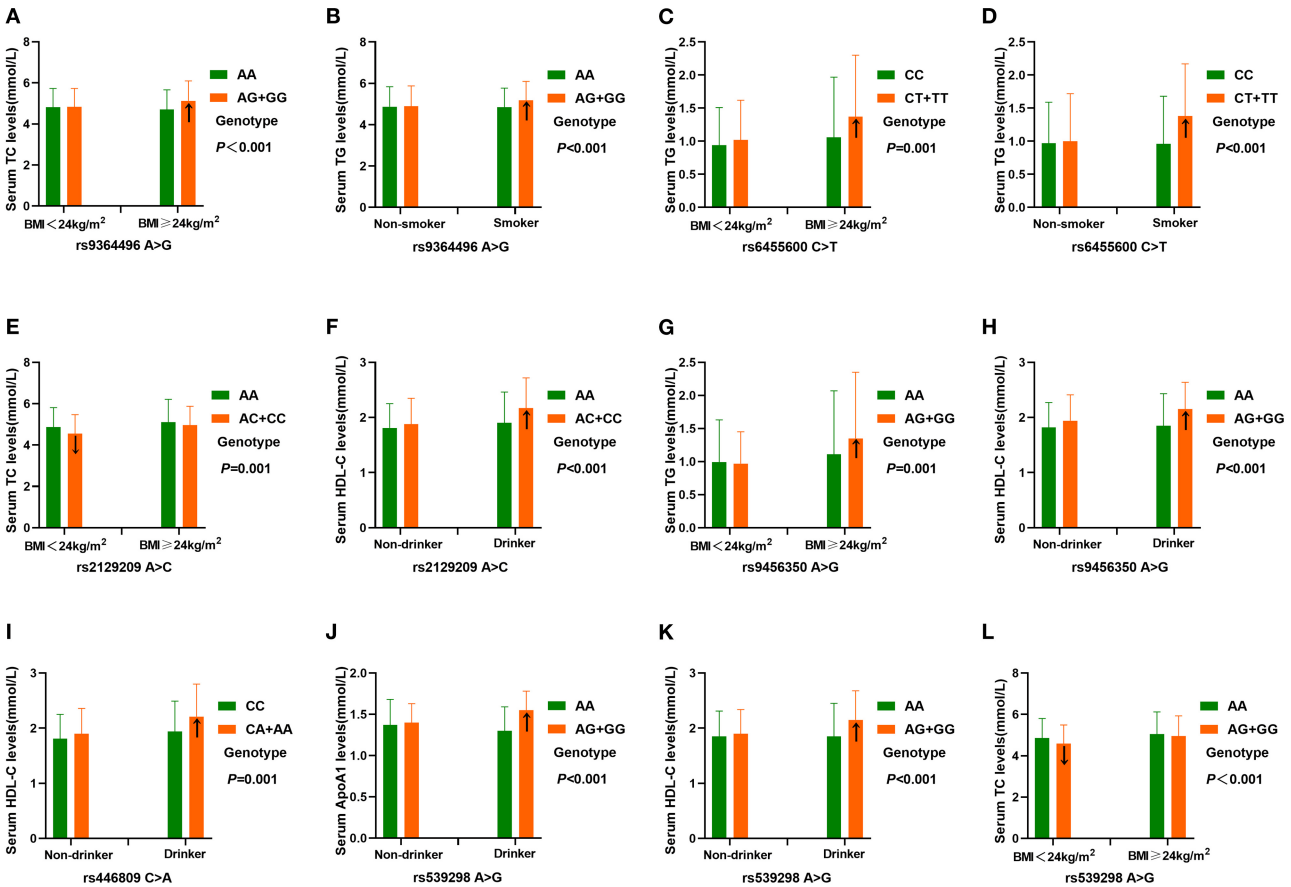


FIGURE 5 | The interactions of the *SYTL3-SLC22A3* SNPs and drinking, BMI and smoking on serum lipid levels. TC, total cholesterol; TG, triglyceride; BMI, body mass index; ApoA1, apolipoprotein A1; HDL-C, high-density lipoprotein cholesterol. A $P \leq 0.00125$ (corresponding to $P < 0.05$ after adjusting for five environmental exposures multiplied by eight outcomes by the Bonferroni correction) was considered statistically significant. **(A)**, rs9364496-BMI on TC; **(B)**, rs9364496-smoking on TG; **(C)**, rs6455600-BMI on TG; **(D)**, rs6455600-smoking on TG; **(E)**, rs2129209-BMI on TC; **(F)**, rs2129209-drinking on HDL-C; **(G)**, rs9456350-BMI on TG; **(H)**, rs9456350-drinking on HDL-C; **(I)**, rs446809-drinking on HDL-C; **(J)**, rs539298-drinking on ApoA1; **(K)**, rs539298-drinking on HDL-C; **(L)**, rs539298-BMI on TC.

factors, including smoking, alcohol consumption and BMI, could affect serum lipid parameters such as TC, TG, HDL-C and ApoA1 levels.

The genotypic and allelic frequencies of the *SLC22A3* rs539298 SNP in different racial/ethnic groups are not well-known. Based on the data derived from the International 1000 Genomes database (https://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/), the frequencies of the rs539298G allele and the AG, GG genotypes were, respectively, 52.21, 49.56, and 27.43% in Europeans; 30.23, 41.86 and 9.30% in Japanese; 59.82, 53.57, and 33.04% in Sub-Saharan Africans; 52.04, 59.18, and 22.45% in Americans of African Ancestry in the Southwestern USA (ASW); 39.20, 48.86, and 14.77% in Gujarati Indians in Houston (GIH); 51.67, 45.56, and 28.89% in Luhya in Webuye, Kenya (LWK); 49.00, 46.00, and 25.00% in those with Mexican ancestry in Los Angeles, California (MEX); 58.10, 50.00, and 33.10% in Maasai in Kinyawa, Kenya (MKK); and 26.83, 48.84, and 4.65% in Han Chinese in Beijing (CHB). In the current study, we noticed that the allele and genotype frequencies of the *SLC22A3* rs539298 SNP were significantly different between the controls and cases. The frequencies of the G allele and AG and GG genotypes in our study populations were 28.28, 41.99, and 7.28% in controls; 23.68, 36.81, and 5.28% in CAD patients; and 22.69, 35.05, and 5.16% in IS cases ($P < 0.05$–0.001, respectively). These findings suggest that the *SLC22A3* rs539298 SNP may have racial/ethnic and population specificity. However, these findings need to be confirmed in other populations or ethnic groups with larger sample size.

Hyperlipidaemia is closely related to the occurrence and development of atherosclerotic cardiovascular and cerebrovascular diseases (39, 40). Previous studies showed that the *SLC22A3* rs2048327 SNP was significantly associated with cardiovascular disease events in familial hypercholesterolemia subjects (41). A genome-wide haplotype association (GWHA) research revealed that the *SLC22A3-LPAL2-LPA* cluster was a strong susceptibility locus for CAD (42). Furthermore, Wang, et al. suggested that the rs3088442G allele might inhibit miR-147a binding to the 3' UTR region of *SLC22A3*, resulting in

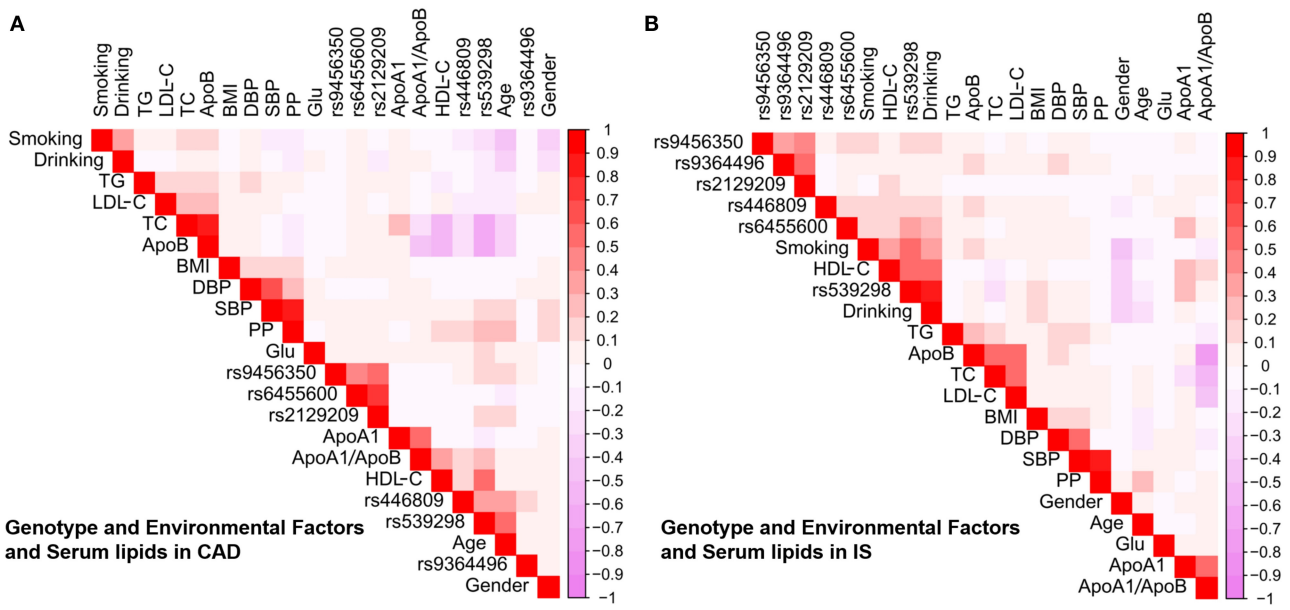


FIGURE 6 | Association between environmental exposures as well as the candidate loci and serum lipid parameters in CAD **(A)** or IS **(B)**. CAD, coronary artery disease; IS, ischaemic stroke. SBP, systolic blood pressure; TC, total cholesterol; BMI, body mass index; TG, triglyceride; ApoA1/B, the ratio of apolipoprotein A1 to apolipoprotein B; HDL-C, high-density lipoprotein cholesterol; DBP, diastolic blood pressure; LDL-C, low-density lipoprotein cholesterol; ApoA1, apolipoprotein A1; PP, pulse pressure; ApoB, apolipoprotein B; Glu, glucose.

increased the expression levels of *SLC22A3*, and ultimately lead to increased risk of CAD (43). In addition, another case control study noticed that four SNPs (rs2048327, rs3127599, rs7767084 and rs10755578) in the *SLC22A3-LPAL2-LPA* cluster were not significantly associated with the risk of CAD (44). However, no study has yet well-elucidated the potential association between the six selected SNPs in the *SYTL3-SLC22A3* cluster and serum lipid levels and the risk of CAD and IS. To meet this need, the potential correlation between the six selected SNPs in the *SYTL3-SLC22A3* cluster and serum lipid levels in humans has been well-elaborated in our previous research (45). We noticed that the rs6455600, rs2129209 and rs539298 SNPs were associated with TC levels; and the rs446809 SNP was associated with TG and LDL-C levels in the Chinese Han population. Furthermore, we also observed that the rs539298G allele carriers had lower serum TC levels than the rs539298G allele non-carriers, and the dominant model of the rs539298 SNP reduced the morbidity of hyperlipidaemia in the Chinese Han population. In the current research, we obtained the similar findings to our previous study. We noticed that serum TC levels in the control group were significantly different between the rs539298AA and rs539298AG/GG genotypes, and participants with the rs539298AG/GG genotypes had lower serum TC levels than those with the rs539298AA genotype. In addition, we also found that the rs539298G allele carriers had lower risk of CAD (adjusted OR = 0.74, 95% CI: 0.60–0.91) and IS (adjusted OR = 0.67, 95% CI: 0.54–0.83) than the rs539298G allele non-carriers. These findings suggested that the *SLC22A3* rs539298A allele may be a genetic risk factor for ischaemic cardiovascular and cerebrovascular diseases, and the *SLC22A3* rs539298G allele may reduce the risk of CAD and IS by affecting serum TC levels. However, the correlation between the *SLC22A3* rs539298 SNP and serum TC levels needs to be confirmed by proteomics studies in future research.

Haplotype analysis showed that the *SYTL3-SLC22A3* A-C-A-A-A-A, G-T-C-G-C-A and A-T-A-A-C-A haplotypes increased the risk of CAD, while the *SYTL3-SLC22A3* A-C-A-A-C-G haplotype reduced the risk of CAD. Meanwhile, the *SYTL3-SLC22A3* A-C-A-A-A-A, G-T-C-G-A-G and A-T-A-A-C-A haplotypes increased the risk of IS, whereas the *SYTL3-SLC22A3* A-C-A-A-A-G and A-C-A-A-C-G haplotypes reduced the risk of IS. These results suggest that haplotypes could explain more changes in the risk of CAD/IS than any single SNP alone.

A large number of studies have suggested that as a multifactorial and complex disorder, CAD or IS occurs due to numerous pathogenic factors, including genetic background, environmental exposures and their interactions (46–48). However, the potential mutual effect between the *SLC22A3* rs539298 SNP and environmental factors on blood lipid levels and the risk of CAD and IS is still unknown. In the current research, we firstly noticed that the interaction of the *SLC22A3* rs539298-alcohol consumption increased serum ApoA1 as well as HDL-C levels; and *SLC22A3* rs539298-BMI < 24 kg/m^2 reduced serum TC levels. These findings could partly account for a decreased risk of CAD and IS in rs539298G allele carriers. Several other potential interactions between haplotypes and environmental factors on the risk of CAD and IS were also

observed in this study. The interactions between the *SYTL3-SLC22A3* A-C-A-A-A-A and A-T-A-A-C-A haplotypes and age > 75 and smoking increased the risk of CAD. The interactions between the *SYTL3-SLC22A3* G-T-C-G-C-A haplotype and smoking and BMI $\geq$ 24 kg/m^2 increased the risk of CAD. The interaction between the *SYTL3-SLC22A3* A-C-A-A-C-G haplotype and alcohol consumption reduced the risk of CAD. In addition, the interaction between the *SYTL3-SLC22A3* A-C-A-A-A-G haplotype and female sex and alcohol consumption reduced the risk of IS. The interactions between the *SYTL3-SLC22A3* A-C-A-A-A-A and smoking and hypertension increased the risk of IS. The interactions between the *SYTL3-SLC22A3* A-T-A-A-C-A and smoking and age > 75 increased the risk of IS. The interactions between the *SYTL3-SLC22A3* A-C-A-A-C-G haplotype and hypertension increased the risk of IS, while its interaction with alcohol consumption reduced the risk of IS. However, these findings have yet to be confirmed by additional research.

The current study may have several limitations. First, compared to other studies, the number of controls and patients was relatively small. Larger samples are necessary to validate our results in future research. Second, most of the patients with CAD or IS received some secondary prevention drugs. All of these drugs may have a certain effect on blood lipid levels. Third, to improve the accuracy of the statistical analysis, we adjusted for several environmental factors, such as BMI, age, smoking, sex and alcohol consumption, but the potential effects of the above factors on blood lipid levels and the risk of CAD and IS could not be completely eliminated. Finally, *in vitro* and *in vivo* studies are necessary to strengthen the significance of our results.

CONCLUSIONS

In conclusion, the current research showed that the allelic and genotypic frequencies of the *SLC22A3* rs539298 SNP were significantly different between controls and CAD/IS patients. The *SLC22A3* rs539298G carriers in the control group had lower serum TC levels than the G allele non-carriers. The interaction between the *SLC22A3* rs539298 SNP and alcohol consumption reduced the risk of CAD and IS. In addition, haplotypes could explain more changes in the risk of CAD/IS than any single SNP alone. Several potential interactions between the haplotypes and some environmental factors affected the risk of CAD and IS. These findings revealed that the *SLC22A3* rs539298A allele may be a new genetic marker for CAD and IS in our study populations. The correlation between the *SLC22A3* rs539298 SNP and CAD/IS may be partly explained by its association with decreased serum TC levels in this study.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the First Affiliated hospital, Guangxi Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

P-FZ conceived the study, participated in the design, performed the statistical analyses, and drafted the manuscript. R-XY conceived the study, participated in the design, carried out the epidemiological survey, collected the samples, and helped to draft the manuscript. X-LC, W-XC, J-ZW, and FH carried out the epidemiological survey and collected the samples. All authors read and approved the final manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (No. 81460169) and the Youthful Science Foundation of Guangxi Province (No. 2017GXNSFBA198067). There was no role of the funding body in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2021.713068/full#supplementary-material

REFERENCES

1. Zhang QH, Yin RX, Chen WX, Cao XL, Wu JZ. TRIB1 and TRPS1 variants, G × G and G × E interactions on serum lipid levels, the risk of coronary heart disease and ischemic stroke. *Sci Rep.* (2019) 9:2376. doi: 10.1038/s41598-019-38765-7
2. Zheng PF, Yin RX, Deng GX, Guan YZ, Wei BL, Liu CX. Association between the XKR6 rs7819412 SNP and serum lipid levels and the risk of coronary artery disease and ischemic stroke. *BMC Cardiovasc Disord.* (2019) 19:202–2. doi: 10.1186/s12872-019-1179-z
3. Li KG, Yin RX, Huang F, Chen WX, Wu JZ, Cao XL. XKR6 rs7014968 SNP increases serum total cholesterol levels and the risk of coronary heart disease and ischemic stroke. *Clin Appl Thromb-Hem.* (2020) 26:1076029620902844. doi: 10.1177/1076029620902844
4. Libby P, Theroux P. Pathophysiology of coronary artery disease. *Circulation.* (2005) 111 3481–8. doi: 10.1161/CIRCULATIONAHA.105.537878
5. Holmstedt CA, Turan TN, Chimowitz MI. Atherosclerotic intracranial arterial stenosis: risk factors, diagnosis, and treatment. *Lancet Neurol.* (2013) 12:1106–14. doi: 10.1016/S1474-4422(13)70195-9
6. Li B, Li W, Li X, Zhou H. Inflammation: a novel therapeutic target/direction in atherosclerosis. *Curr Pharm Des.* (2017) 23:1216–27. doi: 10.2174/1381612822666161230142931

7. Criqui MH, Heiss G, Cohn R, Cowan LD, Suchindran CM, Bangdiwala S, et al., Plasma triglyceride level and mortality from coronary heart disease. *N Engl J Med.* (1993) 328:1220–5. doi: 10.1056/NEJM199304293281702
8. Kwiterovich PO Jr, Coresh J, Smith HH, Bachorik PS, Derby CA, Pearson TA. Comparison of the plasma levels of apolipoproteins B and A-1, and other risk factors in men and women with premature coronary artery disease. *Am J Cardiol.* (1992) 69:1015–21. doi: 10.1016/0002-9149(92)90856-T
9. Crouse JR, Parks JS, Schey HM, Kahl FR. Studies of low density lipoprotein molecular weight in human beings with coronary artery disease. *J Lipid Res.* (1985) 26:566–74. doi: 10.1016/S0022-2275(20)34343-1
10. Shekelle RB, Shryock AM, Paul O, Lepper M, Stamler J, Liu S, et al. Diet, serum cholesterol, and death from coronary heart disease. The Western Electric study. *N Engl J Med.* (1981) 304:65–70. doi: 10.1056/NEJM198101083040201
11. Silbernagel G, Schottker B, Appelbaum S, Scharnagl H, Kleber ME, Grammer TB, et al. High-density lipoprotein cholesterol, coronary artery disease, and cardiovascular mortality. *Eur Heart J.* (2013) 34:3563–71. doi: 10.1093/eurheartj/eht343
12. Sabater-Lleal M, Huang J, Chasman D, Naitza S, Dehghan A, Johnson AD, et al. Multiethnic meta-analysis of genome-wide association studies in >100 000 subjects identifies 23 fibrinogen-associated Loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease. *Circulation.* (2013) 128:1310–24. doi: 10.1161/CIRCULATIONAHA.113.002251
13. Bevan S, Traylor M, Adib-Samii P, Malik R, Paul NL, Jackson C, et al. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. *Stroke.* (2012) 43:3161–7. doi: 10.1161/STROKEAHA.112.665760
14. Ober C, Nord AS, Thompson EE, Pan L, Tan Z, Cusanovich D, et al. Genome-wide association study of plasma lipoprotein(a) levels identifies multiple genes on chromosome 6q. *J Lipid Res.* (2009) 50:798–806. doi: 10.1194/jlr.M800515-JLR200
15. Agmon E, Stockwell BR. Lipid homeostasis and regulated cell death. *Curr Opin Chem Biol.* (2017) 39:83–9. doi: 10.1016/j.cbpa.2017.06.002
16. Stefan CJ, Trimble WS, Grinstein S, Drin G, Reinisch K, De Camilli P, et al. Membrane dynamics and organelle biogenesis-lipid pipelines and vesicular carriers. *BMC Biol.* (2017) 15:102. doi: 10.1186/s12915-017-0432-0
17. Muse ED, Kramer ER, Wang H, Barrett P, Parviz F, Novotny MA, et al. A whole blood molecular signature for acute myocardial infarction. *Sci Rep.* (2017) 7:12268. doi: 10.1038/s41598-017-12166-0
18. Verhaagh S, Schweifer N, Barlow DP, Zwart R. Cloning of the mouse and human solute carrier 22a3 (Slc22a3/SLC22A3) identifies a conserved cluster of three organic cation transporters on mouse chromosome 17 and human 6q26-q27. *Genomics.* (1999) 55:209–18. doi: 10.1006/geno.1998.5639
19. Schneider E, Machavoine F, Pléau JM, Bertron AF, Thurmond RL, Ohtsu H, et al. Organic cation transporter 3 modulates murine basophil functions by controlling intracellular histamine levels. *J Exp Med.* (2005) 202:387–93. doi: 10.1084/jem.20050195
20. Rozenberg I, Sluka SH, Rohrer L, Hofmann J, Becher B, Akhmedov A, et al. Histamine H1 receptor promotes atherosclerotic lesion formation by increasing vascular permeability for low-density lipoproteins. *Arterioscler Thromb Vasc Biol.* (2010) 30:923–30. doi: 10.1161/ATVBAHA.109.201079
21. Kimura S, Wang KY, Tanimoto A, Murata Y, Nakashima Y, Sasaguri Y. Acute inflammatory reactions caused by histamine via monocytes/macrophages chronically participate in the initiation and progression of atherosclerosis. *Pathol Int.* (2004) 54:465–74. doi: 10.1111/j.1440-1827.2004.01653.x
22. Li L, He M, Zhou L, Miao X, Wu F, Huang S, et al. A solute carrier family 22 member 3 variant rs3088442 G→ A associated with coronary heart disease inhibits lipopolysaccharide-induced inflammatory response. *J Biol Chem.* (2015) 290:5328–40. doi: 10.1074/jbc.M114.584953
23. Chen L, Pawlikowski B, Schlessinger A, More SS, Stryke D, Johns SJ, et al. Role of organic cation transporter 3 (SLC22A3) and its missense variants in the pharmacologic action of metformin. *Pharmacogenet Genomics.* (2010) 20:687–99. doi: 10.1097/FPC.0b013e32833fe789
24. Solbach TF, Grube M, Fromm MF, Zolk O. Organic cation transporter 3: expression in failing and nonfailing human heart and functional characterization. *J Cardiovasc Pharmacol.* (2011) 58:409–17. doi: 10.1097/FJC.0b013e3182270783
25. Zhao Q, Wei H, Liu D, Shi B, Li L, Yan M, et al. PHACTR1 and SLC22A3 gene polymorphisms are associated with reduced coronary artery disease risk in the male Chinese Han population. *Oncotarget.* (2017) 8:658–63. doi: 10.18632/oncotarget.13506
26. Chen L, Yao Y, Jin C, Wu S, Liu Q, Li J, et al. Integrative genomic analysis identified common regulatory networks underlying the correlation between coronary artery disease and plasma lipid levels. *BMC Cardiovasc Disord.* (2019) 19:310. doi: 10.1186/s12872-019-01271-9
27. Franceschini N, Carty C, Buzkova P, Reiner AP, Garrett T, Lin Y, et al. Association of genetic variants and incident coronary heart disease in multiethnic cohorts: the PAGE study. *Circ Cardiovasc Genet.* (2011) 4:661–72. doi: 10.1161/CIRCGENETICS.111.960096
28. Sun JQ, Yin RX, Shi GY, Shen SW, Chen X, Bin Y, et al. Association of the ARL15 rs6450176 SNP and serum lipid levels in the Jing and Han populations. *Int J Clin Exp Pathol.* (2015) 8:12977–94.
29. Guo T, Yin RX, Nie RJ, Chen X, Bin Y, Lin WX. Suppressor of cytokine signaling 3 A+930–>G (rs4969168) polymorphism is associated with apolipoprotein A1 and low-density lipoprotein cholesterol. *Int J Clin Exp Pathol.* (2015) 8:7305–17.
30. Li WJ, Yin RX, Huang JH, Bin Y, Chen WX, Cao XL. Association between the PPP1R3B polymorphisms and serum lipid traits, the risk of coronary artery disease and ischemic stroke in a southern Chinese Han population. *Nutr Metab.* (2018) 15:27. doi: 10.1186/s12986-018-0266-y
31. Zheng PF, Yin RX, Liu CX, Deng GX, Guan YZ, Wei BL. SYNE1-QK1 SNPs, G × G and G × E interactions on the risk of hyperlipidaemia. *J Cell Mol Med.* (2020) 24:5772–85. doi: 10.1111/jcmm.15239
32. Yin RX, Aung LH, Long XJ, Yan TT, Cao XL, Huang F, et al. Interactions of several genetic polymorphisms and alcohol consumption on blood pressure levels. *Biofactors.* (2015) 41:339–51. doi: 10.1002/biof.1234
33. Yin RX, Wu DF, Wu JZ, Cao XL, Aung LHH, Miao L, et al. Interactions of several lipid-related gene polymorphisms and cigarette smoking on blood pressure levels. *Int J Biol Sci.* (2012) 8:685–96. doi: 10.7150/ijbs.4401
34. Zhou BF, Cooperative Meta-Analysis Group of the Working Group on Obesity in China. Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults—study on optimal cut-off points of body mass index and waist circumference in Chinese adults. *Biomed Environ Sci.* (2002) 15:83–96.
35. Wildman RP, Gu D, Reynolds K, Duan X, He J. Appropriate body mass index and waist circumference cutoffs for categorization of overweight and central adiposity among Chinese adults. *Am J Clin Nutr.* (2004) 80:1129–36. doi: 10.1093/ajcn/80.5.1129
36. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med.* (1998) 15: 539–53.
37. Yin RX, Wang Y, Chen GQ, Lin WX, Yang DZ, Pan SL. Lipoprotein lipase gene polymorphism at the PvuII locus and serum lipid levels in Guangxi Hei Yi Zhuang and Han populations. *Clin Chem Lab Med.* (2006) 44:1416–21. doi: 10.1515/CCLM.2006.273
38. Onda Y, Takahagi K, Shimizu M, Inoue K, Mochida K. Multiplex PCR targeted amplicon sequencing (MTA-Seq): simple, flexible, and versatile SNP genotyping by highly multiplexed PCR amplicon sequencing. *Front Plant Sci.* (2018) 9:201. doi: 10.3389/fpls.2018.00201
39. Yeramaneni S, Kleindorfer DO, Sucharew H, Alwell K, Moomaw CJ, Flaherty ML, et al. Hyperlipidemia is associated with lower risk of poststroke mortality independent of statin use: a population-based study. *Int J Stroke.* (2017) 12:152–60. doi: 10.1177/1747493016670175
40. Zhang X, Patel A, Horibe H, Wu Z, Barzi F, Rodgers A, et al. Cholesterol, coronary heart disease, and stroke in the Asia Pacific region. *Int J Epidemiol.* (2003) 32:563–72. doi: 10.1093/ije/dyg106
41. Paquette M, Bernard S, Baass A. SLC22A3 is associated with lipoprotein (a) concentration and cardiovascular disease in familial hypercholesterolemia. *Clin Biochem.* (2019) 66:44–8. doi: 10.1016/j.clinbiochem.2019.02.008
42. Trégouët DA, König IR, Erdmann J, Munteanu A, Braund PS, Hall AS, et al. Genome-wide haplotype association study identifies the SLC22A3-LPAL; 2-LPA gene cluster as a risk locus for coronary artery disease. *Nat Genet.* (2009) 41:283–5. doi: 10.1038/ng.314

43. Wang L, Chen J, Zeng Y, Wei J, Jing J, Li G, et al. Functional variant in the SLC22A3-LPAL2-LPA gene cluster contributes to the severity of coronary artery disease. *Arterioscler Thromb Vasc Biol.* (2016) 36:1989–96. doi: 10.1161/ATVBAHA.116.307311
44. Lv X, Zhang Y, Rao S, Liu F, Zuo X, Su D, et al. Lack of association between four SNPs in the SLC22A3-LPAL2-LPA gene cluster and coronary artery disease in a Chinese Han population: a case control study. *Lipids Health Dis.* (2012) 11:128. doi: 10.1186/1476-511X-11-128
45. Zheng PF, Yin RX, Guan YZ, Wei BL, Liu CX, Deng GX. SYTL3-SLC22A3 SNPs, G × G and G × E interactions on the risk of hyperlipidaemia. *Front Genet*. (2021) 12:679027. doi: 10.3389/fgene.2021.679027
46. Yang Q, Yin RX, Cao XL, Huang F, Zhou YJ, Chen WX. ANGPTL4 variants and their haplotypes are associated with serum lipid levels, the risk of coronary artery disease and ischemic stroke and atorvastatin cholesterol-lowering responses. *Nutr Metab (Lond).* (2018) 15:70. doi: 10.1186/s12986-018-0308-5
47. Li WJ, Yin RX, Cao XL, Chen WX, Huang F, Wu JZ. DOCK7-ANGPTL3 SNPs and their haplotypes with serum lipid levels and the risk of coronary artery disease and ischemic stroke. *Lipids Health Dis.* (2018) 17:30. doi: 10.1186/s12944-018-0677-9
48. Zhang QH, Yin RX, Chen WX, Cao XL, Chen YM. Association between the TIMD4-HAVCR1 variants and serum lipid levels, coronary heart disease and ischemic stroke risk and atorvastatin lipid-lowering efficacy. *Biosci Rep*. (2018) 38:BSR20171058. doi: 10.1042/BSR20171058

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Cardiovascular Medicine

CASE REPORT
published: 13 September 2021
doi: 10.3389/fcvm.2021.711694

Case Report: *BMPR2*-Targeted MinION Sequencing as a Tool for Genetic Analysis in Patients With Pulmonary Arterial Hypertension

Tomoya Takashima[1], *Sophie Brisset*[1,2,3], *Asuka Furukawa*[1], *Hirohisa Taniguchi*[1,4], *Rika Takeyasu*[1], *Akio Kawamura*[4] *and Yuichi Tamura*[1,4]*

[1] Pulmonary Hypertension Center, International University of Health and Welfare, Mita Hospital, Tokyo, Japan, [2] Faculty of Medicine, Université Paris-Saclay, Le Kremlin-Bicêtre, France, [3] Service d'Histologie, Embryologie et Cytogénétique, Assistance Publique Hôpitaux de Paris (AP-HP), Hôpital Antoine Béclère, Clamart, France, [4] Department of Cardiology, International University of Health and Welfare School of Medicine, Narita, Japan

OPEN ACCESS

Edited by:
Seitaro Nomura,
The University of Tokyo, Japan

Reviewed by:
Douglas A. Marchuk,
Duke University, United States
Crescenzio Francesco Minervini,
University of Bari Aldo Moro, Italy

***Correspondence:**
Yuichi Tamura
tamura.u1@gmail.com

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: 19 May 2021
Accepted: 11 August 2021
Published: 13 September 2021

Citation:
Takashima T, Brisset S, Furukawa A, Taniguchi H, Takeyasu R, Kawamura A and Tamura Y (2021) Case Report: BMPR2-Targeted MinION Sequencing as a Tool for Genetic Analysis in Patients With Pulmonary Arterial Hypertension. Front. Cardiovasc. Med. 8:711694. doi: 10.3389/fcvm.2021.711694

Background: Mutations in the bone morphogenetic protein receptor type 2 gene (*BMPR2*) represent a major genetic cause of pulmonary arterial hypertension (PAH). Identification of *BMPR2* mutations is crucial for the genetic diagnosis of PAH. MinION nanopore sequencer is a portable third-generation technology that enables long-read sequencing at a low-cost. This nanopore technology-based device has not been used previously for PAH diagnosis. This study aimed to determine the feasibility of using MinION nanopore sequencing for the genetic analysis of PAH patients, focused on *BMPR2*.

Methods: We developed a protocol for the custom bioinformatics pipeline analysis of long reads generated by long-PCR. To evaluate the potential of using MinION sequencing in PAH, we analyzed five samples, including those of two idiopathic PAH patients and a family of three members with one affected patient. Sanger sequencing analysis was performed to validate the variants.

Results: The median read length was around 3.4 kb and a good mean quality score of approximately 19 was obtained. The total number of reads generated was uniform among the cases and ranged from 2,268,263 to 3,126,719. The coverage was consistent across flow cells in which the average number of reads per base ranged from 80,375 to 135,603. We identified two polymorphic variants and three mutations in four out of five patients. Certain indel variant calling-related errors were observed, mostly outside coding sequences.

Conclusion: We have shown the ability of this portable nanopore sequencer to detect *BMPR2* mutations in patients with PAH. The MinION nanopore sequencer is a promising tool for screening *BMPR2* mutations, especially in small laboratories and research groups.

Keywords: nanopore sequencing, MinION, BMPR2, PAH, next generation sequencing

INTRODUCTION

Pulmonary arterial hypertension (PAH) is a rare, severe, and progressive disorder associated with vascular remodeling and the narrowing of small pulmonary arteries. The increase in pulmonary vascular resistance can result in right heart failure and eventual death (1–3). Pulmonary arterial hypertension may be inherited from an affected relative (heritable PAH) or induced by a drug, toxin, or associated with another disease (congenital heart disease, connective tissue disease, HIV). However, in most cases of PAH, there is no identifiable cause or family history of disorder (idiopathic PAH) (4).

Developments in genetics and technological advances have contributed to our knowledge of this lethal disease. Most heritable PAH cases are caused by a mutation in the bone morphogenetic protein receptor type 2 gene (*BMPR2*), a receptor for the transforming growth factor-beta (TGF-β) superfamily (5, 6). *BMPR2* mutations are detected in approximately 80% of individuals with family history and in 11–40% of patients with sporadic PAH (7). The rate of occurrence of *BMPR2* mutations in Japanese patients is the same as that reported in other countries. Mutations have been identified less frequently in other genes, including *ACVRL1* (also known as *ALK1*), *ENG, TBX4, KCNK3, CAV1*, and *SMAD9*, among others (8, 9).

The identification of *BMPR2* mutations is a crucial step in the genetic diagnosis of PAH. It has implications for personalized therapy and genetic counseling (7–11). The usual approach is to perform Sanger sequencing, but the process is time consuming and labor-intensive. The complexity, efficiency, and rapidity of sequencing methods continues to increase. Next-generation sequencing (NGS) technology transformed the manner in which sequencing was performed (12). Next-generation sequencing exhibited a higher sensitivity in the detection of novel and rare variants (9). The past few years have seen the emergence of a nanopore-based third-generation sequencing technology, known as nanopore DNA sequencing (13). The low-cost and rapid MinION-based nanopore sequencing method has emerged as a competitive and portable technology that has provided accurate long sequence reads (14). Here, we developed a rapid method for genetic analysis that employed MinION technology through a custom analysis pipeline. Because of the high incidence of *BMPR2* mutations, we particularly focused on this gene. To the best of our knowledge, this article is the first to report about the use of MinION for PAH genetic analysis.

MATERIALS AND METHODS

Patients and Genomic DNA

Five samples, including those of two patients with idiopathic PAH (Cases #1 and #2) and a family of three (Cases #3, #4, and #5) were investigated in this study. In the family, only the father (#3) was diagnosed with PAH; the siblings, a daughter (#4) and a son (#5), were unaffected. This study was approved by a local institutional ethics committee at the International University of Health and Welfare (approval number 5-16-25); informed consent was obtained from each patient. Whole blood samples were collected in EDTA-containing tubes and sent to a commercial laboratory (SRL, Tokyo, Japan) for DNA extraction, according to standard procedures. After genomic DNA extraction, DNA samples were stored in the laboratory at 4°C. The integrity of the DNA was checked using a NanoDrop One Spectrophotometer and the DNA was diluted to a concentration of 20 ng/μl.

Long-PCR Target Enrichment

Before sequencing, DNA was amplified by targeted long-PCR. Primers were designed using Primer-BLAST to amplify the *BMPR2* exonic sequences (including UTR regions) and surrounding intronic sequences (adjacent, upstream, and/or downstream exonic sequences). The primer locations and corresponding amplicon sizes are presented in **Figure 1A**. Details regarding the primer sequences and conditions are described in **Supplementary Table 1**.

We used a long-PCR based strategy with *BMPR2*-specific primers tailed with universal sequences, as recommended by Oxford Nanopore. PCR was performed using the KAPA HiFi HotStart ReadyMix PCR Kit (KAPA BIOSYSTEMS) on a T100 Thermal Cycler (BIO-RAD).

The amplification quality was checked via agarose gel electrophoresis and PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN), according to the manufacturer's protocol.

Library Preparation and Nanopore Sequencing

Library preparation was performed using the protocol 1D amplicon/cDNA, by ligation (SQK-LSK109) (Oxford Nanopore). Sequencing was performed using the MinION device, along with a FLO-MIN106 (Pore type R9.4) Flowcell. The MinION device was controlled by the MinKNOW software during the run time of 24 h. During sequencing, base calling was performed using the MinKNOW software (v19.06.8).

Data Analysis

A pipeline was specifically developed for analyzing the data produced by MinION sequencing (**Figure 1B**). Raw electric signals classified as "pass" signals were converted to raw reads in the FASTQ format using MinKNOW. Average read quality was determined using FASTQC (v0.11.6). The filtering of parameters was performed based on quality and read length, and trimming was performed using Nanofilt (v2.6.0). A cutoff value of 10 was set as an acceptance quality score. The 50 nucleotide long reads were trimmed from the 5′ end and filtered for average quality (quality score >10) and size (>100 nt). We mapped the quality-controlled reads to a human *BMPR2* region (GRCh38/hg38, chr2: 202371789-202568449) using the Minimap2 program (v2.13-r850). These unique, mapped reads were extracted and converted to another format for mutation analysis using Samtools (v1.9). Single nucleotide variations (SNVs) and small insertions and deletions (Indels) were identified using Varscan (v2.4.3). Variant

Abbreviations: *BMPR2*, bone morphogenetic protein receptor type 2; NGS, Next generation sequencing; PAH, pulmonary arterial hypertension; PCR, polymerase chain reaction; SNV, Single nucleotide variant.

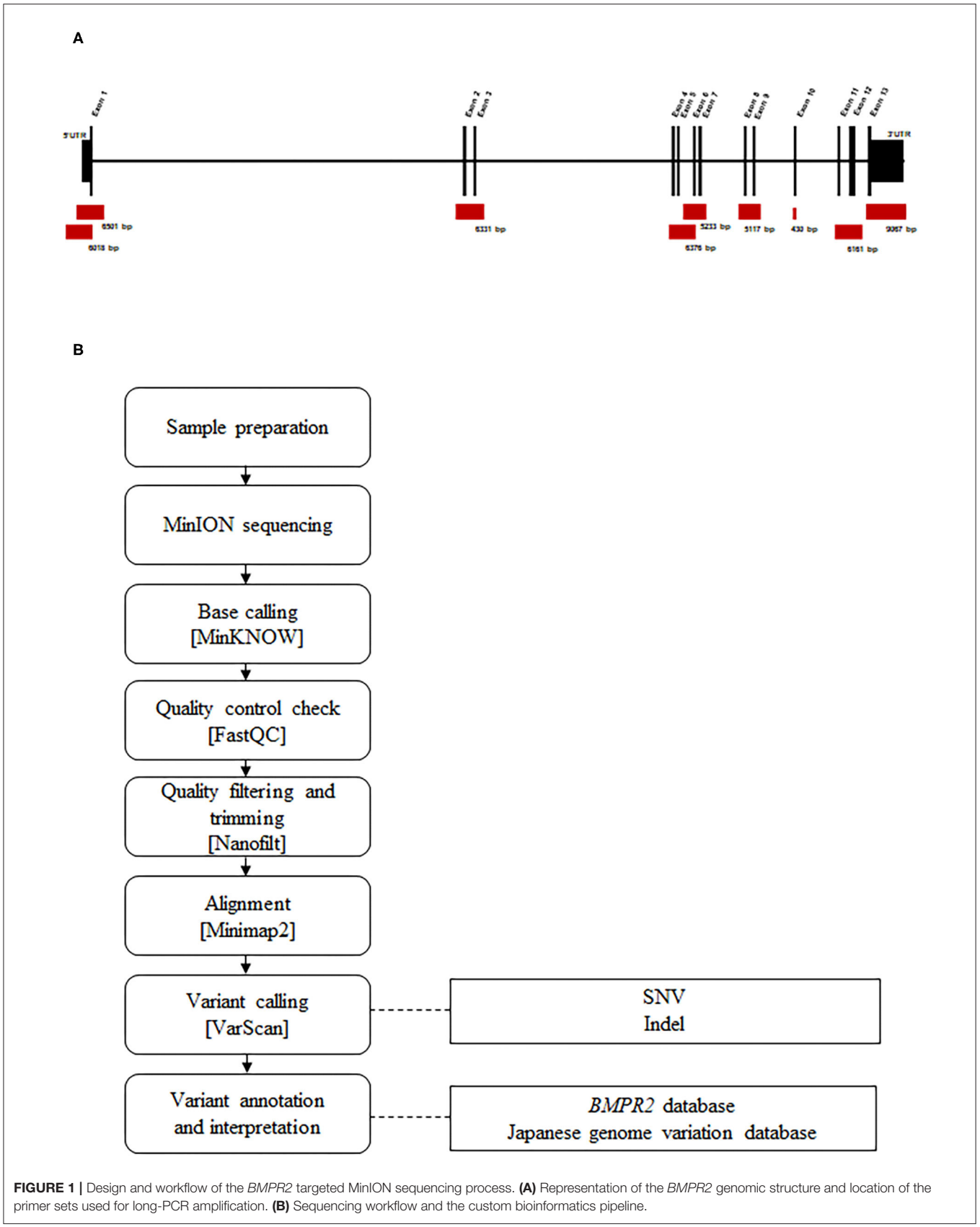


FIGURE 1 | Design and workflow of the *BMPR2* targeted MinION sequencing process. **(A)** Representation of the *BMPR2* genomic structure and location of the primer sets used for long-PCR amplification. **(B)** Sequencing workflow and the custom bioinformatics pipeline.

TABLE 1 | Read statistics derived from MinION sequencing data for each case.

Case	#1	#2	#3	#4	#5
Median length (bp)	3,524	3,454	3,244	3,120	3,465
1st Quartile length (bp)	1,653	1,450	1,297	1,225	1,462
3rd Quartile length (bp)	5,681	5,133	4,884	4,923	5,127
Mean quality score	20.26	19.26	18.80	18.09	18.93
Base called reads *n*	2,638,348	3,017,449	2,297,139	3,126,718	2,268,263
Reads after filtering *n* (%)	2,219,645 (84%)	2,453,722 (81%)	1,802,236 (78%)	2,283,628 (73%)	1,746,233 (77%)
Reads after alignment *n* (%)	1,550,471 (59%)	1,683,653 (56%)	975,403 (42%)	1,520,099 (56%)	1,061,661 (47%)
Average sequencing depth	135,603	133,736	80,375	119,464	106,229

Reads were base-called using MinKNOW and filtered and trimmed using Nanofilt before alignment. n, number of reads; %, percentage of reads after quality filtering/trimming or after alignment is indicated in brackets. The average sequencing depth corresponds to the average number of reads per base.

calls were filtered using the following parameters: a minimum coverage of 2,000 reads, an allele frequency of at least 10%, p-value < 0.05, and a fraction of reads from each strand (plus/minus) with a range of 0.1–10.

Then, the filtered variants were manually compared to a population database of genomic variations detected by the whole genome sequencing of Japanese individuals (https://jmorp.megabank.tohoku.ac.jp) and a database of known *BMPR2* variants (https://arup.utah.edu/database/BMPR2/BMPR2_welcome.php).

Statistic Evaluation and Sequencing Data Visualization

Average read lengths were calculated using awk commands. The distribution of quality and length (using a Kernel density estimation) was performed using the ggplot2 (v3.2.1) package of R software.

Sanger Sequencing

All five samples were analyzed for *BMPR2* mutation by Sanger sequencing. Primers were designed using Primer BLAST. Genomic DNA was amplified by PCR using a T100 Thermal Cycler (BIO-RAD) and TaKaRa Ex Taq DNA Polymerase (TAKARA BIO INC). Primer sequences and PCR conditions are available upon request. PCR products were subjected to nucleotide sequencing by a commercial sequencing service (FASMAC, Japan). Nucleotide sequences were compared with human reference genome sequences. Electropherograms were analyzed by visual inspection and the ATGC Software (GENETYX).

RESULTS

MinION sequencing runs were performed on PCR products generated by long-PCR. Five runs were carried out on five different flow cells. A summary of the number and length of reads, the quality score and the average per-base coverage for each experiment is detailed in **Table 1**.

The mean quality score of approximately 19 was considered a good score (**Table 1**). The plotting of the quality score distribution for the five runs showed that the quality was consistent from one flow cell to another (**Figure 2A**).

The median read length was approximately 3.4 kb (**Table 1**). Upon plotting the distribution of read lengths obtained for the five runs, we could observe that the range of read lengths was in agreement with the physical size of the sequencing library, and the amplicon size ranged from 430 bp to 9 kb (**Figure 2B**).

The total number of base-called reads was significantly uniform between cases and ranged from 2,268,263 to 3,126,718 reads (**Table 1**). After quality filtering and trimming (quality score >10, sequences shorter than 100 nucleotides removed, first 50 nucleotides of all reads trimmed), this number was reduced, and ranged from 73 to 84% of base-called reads. Finally, we mapped an average of 51% of the base-called reads (42–59%) to the reference genome after alignment (**Table 1**).

The average coverage was considerably uniform, with a similar number of per-base reads across flow cells, ranging from 80,375 to 135,603 reads per base (**Table 1**). Among aligned reads, the ratio of reads assigned to each exon (from exon 1 to exon 13) was consistent across cases (**Figure 2C**).

Following variant calling with VarScan's filtering parameters, we detected mostly SNVs and fewer Indels. All the variants identified for each case are listed in **Table 2**. These variants were interpreted based on their genomic location (exonic, splicing variant, or intronic), functional effect (synonymous or non-synonymous), and data obtained from a Japanese population database and a *BMPR2* variant database. Most variants were identified in intronic regions. No splicing variants and Indels were located within coding sequences. Sanger sequencing could not confirm the presence of Indel variants and resulted in false positive results. We checked the genomic sequences flanking these called Indels and observed nucleotide sequence repeats. These indel errors were probably related to these homopolymeric regions.

Finally, MinION sequencing revealed three mutations and two polymorphisms in four of the five cases (Cases #1, #2, #3, and #5). All these variants were validated by Sanger sequencing. Only

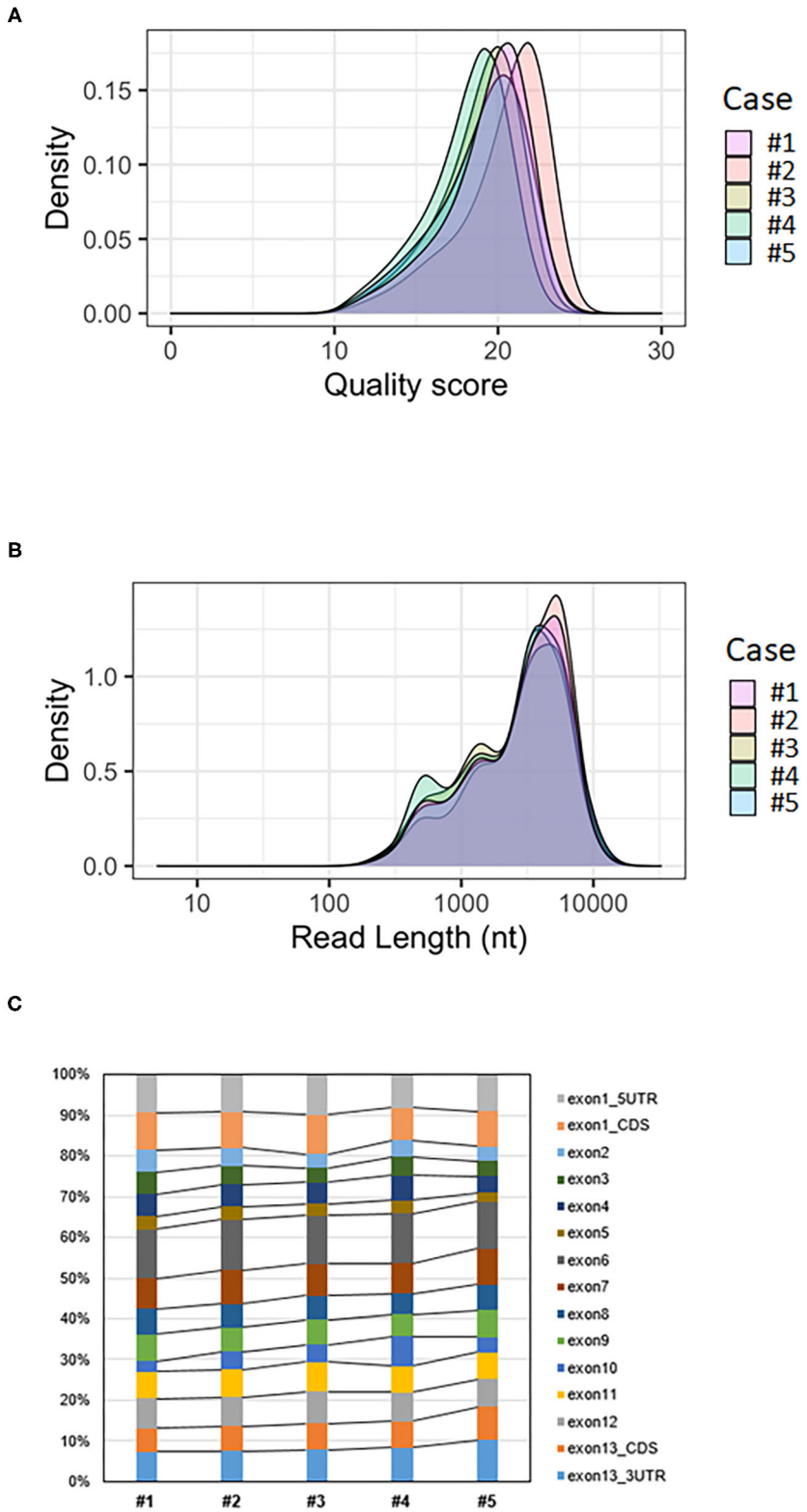


FIGURE 2 | Analysis of *BMPR2* targeted MinION sequencing data. **(A)** Quality score distribution of reads. **(B)** Read length distribution of reads; the X-axis is log10 transformed; nt = nucleotides. **(C)** Ratio of aligned reads per exon.

TABLE 2 | Number and type of all *BMPR2* variants detected using MinION sequencing.

Case	Patient background	SNVs					SNVs identified in Japanese database
		Total	CDS	UTR	Splice site	Intron	
#1	PAH	32	2	3	0	27	18
#2	PAH	30	1	1	0	28	10
#3	PAH	23	2	0	0	21	9
#4	Unaffected	25	0	1	0	24	4
#5	Unaffected	28	2	0	0	26	8
Case	**Patient background**	**Indels**					**Indels identified in Japanese database**
		Total	**CDS**	**UTR**	**Splice site**	**Intron**	
#1	PAH	0	0	0	0	0	0
#2	PAH	3	0	0	0	3	0
#3	PAH	0	0	0	0	0	0
#4	Unaffected	2	0	1	0	1	0
#5	Unaffected	1	0	0	0	1	0

Variants were manually checked for redundancy with respect to variants in an available Japanese database providing genomic variation data of whole genome sequencing of Japanese individuals (https://jmorp.megabank.tohoku.ac.jp). SNV, Single nucleotide variant; CDS, coding sequence; PAH, pulmonary arterial hypertension; UTR, untranslated region; Indel, short insertions or deletions.

cases #1, #2, and #3 were diagnosed with PAH. A final summary of the mutations and polymorphisms detected by MinION and Sanger sequencing has been shown in **Table 3**.

Case #1 presents two variants in the coding sequence of exon 12, c.2379A>C (p.Thr793=) and c.2811G>A (p.Arg937=). These two synonymous variants are considered as benign polymorphisms and are registered in the NCBI dbSNP database (rs3731697 and rs1061157, respectively). No pathogenic mutation was identified for this case.

Case #2 was heterozygous for a nonsense mutation c.2695C>T (p.Arg899Ter) in exon 12. This variant was considered to be pathogenic in the disease database.

Case #3 is heterozygous for a nonsense mutation c.2617 C>T (p.Arg273Ter) in exon 12 (**Figure 3A**; **Supplementary Figure 1**). The same mutation was identified in his unaffected son (#5), who exhibited no clinical features of the disease (**Figure 3B**; **Supplementary Figure 1**). His healthy daughter (#4) did not carry this mutation (**Figure 3C**; **Supplementary Figure 1**). This variant is considered pathogenic in the disease database. The father (#3) and his son (#5) also presented with a missense variant c.276 A>C (p.Gln92His) located in exon 3. The significance of this variant is considered to be uncertain.

DISCUSSION

Here, we developed a protocol and performed a pipeline analysis via the targeted nanopore sequencing of *BMPR2*. Nanopore sequencing has been successfully used to sequence genomes of many species (viruses, bacteria, animals, plants, and humans), but limited data are available regarding the use of this technology in clinical settings (15–22). To the best of our knowledge, the MinION platform has not been applied for PAH diagnosis.

Sequencing is essential for determining the genetic basis of pathogenicity and for genetic counseling. Current Sanger technology is time consuming and labor intensive. New sequencing technologies have revolutionized DNA sequencing. However, though NGS generally requires a massive investment in capital equipment, its use could be limited in small laboratories, owing to its high cost and high input material requirements. Another limitation of second-generation platforms is their short-read lengths, which range from 150 to 300 bp (12, 23). The nanopore sequencing platform is a third-generation platform that generates long read lengths; its sequencing process occurs in real-time (13, 14). The MinION nanopore sequencer is a small, easy to manipulate, and affordable device, capable of sequencing many gigabases (Gb) in a single experiment (24). This portable device allows sequencing to be performed in a small laboratory or research unit, where the large-scale installation of sequencing equipment is unfeasible.

We are the first to report about the use of MinION sequencing for the rapid genetic analysis of PAH patients. *BMPR2* was focused upon because its frequency was the highest. Heterozygous germline mutations in *BMPR2* represent the central susceptibility factor in the precipitation and progression of PAH (9). Our laboratory has developed a pipeline for rapidly sequencing human samples and validating them via their comparison with standard Sanger sequencing results.

To evaluate the potential of MinION sequencing for genetic analysis, we analyzed a small set of five samples. We have discussed the ability of MinION to detect *BMPR2* variants below.

We identified two polymorphisms and three mutations, comprising two nonsense mutations considered to be pathogenic, and one missense mutation (see below). Both MinION sequencing and Sanger sequencing have detected these

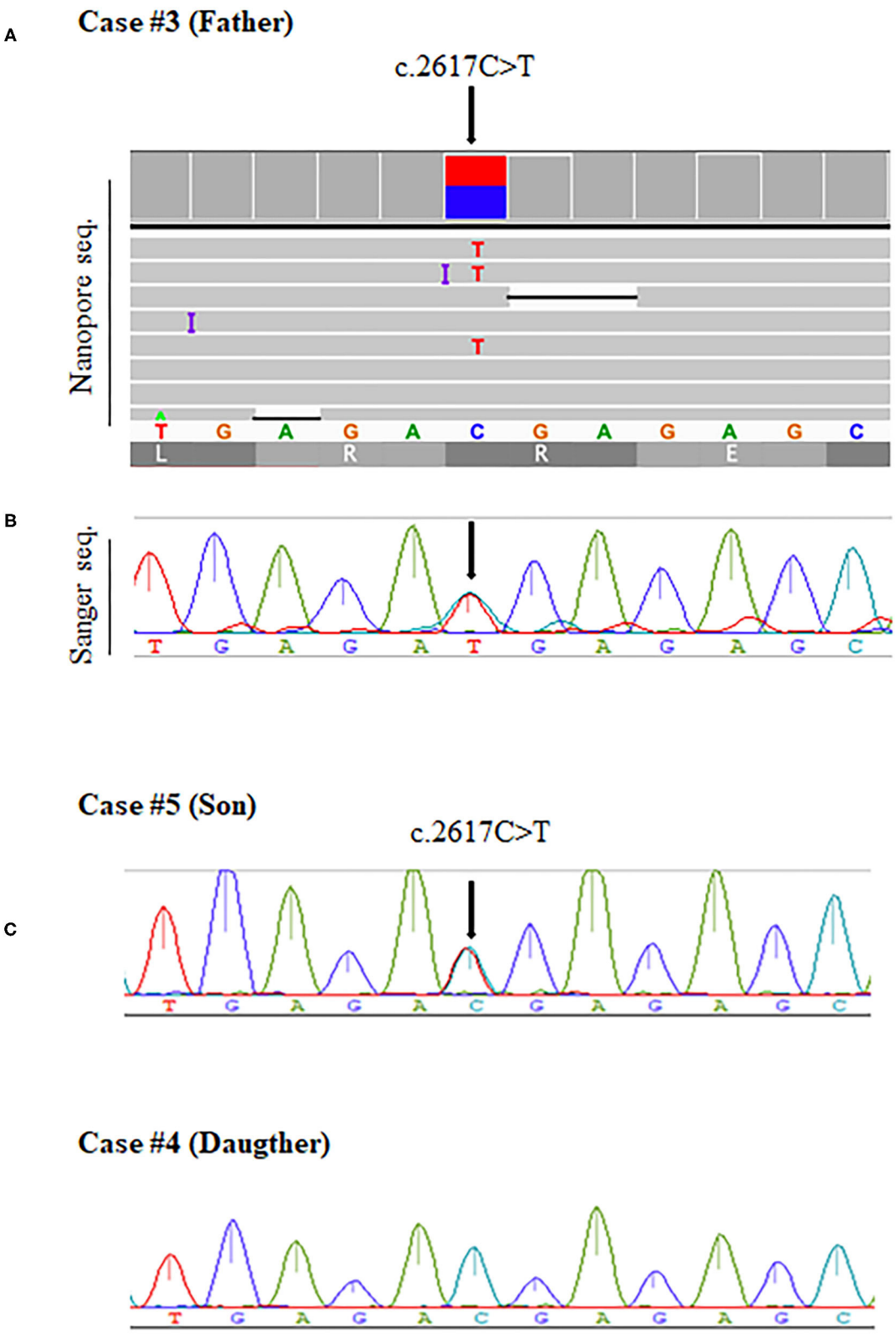


FIGURE 3 | The *BMPR2* mutation c.2617C>T was detected in two members of a family (Cases #3 and #5) and validated by Sanger sequencing. Arrows indicate mutations in the heterozygous state. **(A)** Integrative genomics viewer (IGV) screenshot and Sanger validation of case #3 with PAH. **(B)** Sanger sequencing results show the same mutation being detected for case #5, the unaffected son. **(C)** Sanger sequencing results for case #4, the healthy daughter, show no mutations.

TABLE 3 | Description and annotation of *BMPR2* mutations and polymorphisms detected using MinION sequencing for each case.

Case	Location	Nucleotide change	Variant type	Protein change	Mutation type	Clinical significance*	Reference SNP report**
#1	Exon 12	c.2379 A>C	SNV	p.Thr793=	–	Benign	rs3731697
#1	Exon 12	c.2811 G>A	SNV	p.Arg937=	–	Benign	rs1061157
#2	Exon 12	c.2695 C>T	SNV	p.Arg899XTer	Nonsense	Pathogenic	–
#3	Exon 3	c.276 A>C	SNV	p.Gln92His	Missense	Benign	–
#3	Exon 12	c.2617 C>T	SNV	p.Arg873XTer	Nonsense	Pathogenic	–
#5	Exon 3	c.276 A>C	SNV	p.Gln92His	Missense	Benign	–
#5	Exon 12	c.2617 C>T	SNV	p.Arg873XTer	Nonsense	Pathogenic	–

All variants were validated by Sanger sequencing. SNV, Single nucleotide variant; SNP, single nucleotide polymorphism.
** in the database of known BMPR2 variants (https://arup.utah.edu/database/BMPR2/BMPR2_welcome.php).*
*** in the Japanese database (https://jmorp.megabank.tohoku.ac.jp).*

BMPR2 variants, which were located in exonic regions. There was no discrepancy.

Two polymorphic variants in exon 12 (c.2379 A>C and c.2811 G>A), observed in Case #1, caused non-synonymous changes in codons and did not alter amino acid sequences (p.Thr793= and p.Arg937=). These polymorphic variants were reportedly observed in unaffected individuals (25, 26) and were not expected to be clinically significant.

The nonsense mutation c.2695C>T (p.Arg899Ter) in exon 12 (Case #2) has been previously reported in individuals with PAH (6, 25–27). In the family mentioned above, we could identify a nonsense mutation c.2617 C>T (p.Arg273Ter) in the affected father (Case #3) and his son (Case #5), who did not exhibit symptoms yet. This mutation was previously reported to be pathogenic and showed reduced penetration among carriers (5, 26). The disease onset varied widely between individuals who harbored the same gene defect within a family, suggesting that other factors, genetic and/or environmental, were important for disease progression (10, 26). Furthermore, it is essential to identify this mutation for clinical management; the necessity to do so extends beyond the needs of the affected patient, as it enables earlier disease diagnosis and adoption of preventive measures in carriers (28).

In the same family, we also identified a c.276 A>C (p.Gln92His) missense variant in the father and son. This mutation was classified as a benign mutation in our *BMPR2* database (http://arup.utah.edu/database/BMPR2/BMPR2_welcome.php, last update January 2013). However, this missense variant is listed in the ClinVar database (http://www.ncbi.nlm.nih.gov/clinvar, last evaluated March 2016) and interpreted as a variant with uncertain significance, as reported previously (29, 30). It should be noted that several years have elapsed since variant interpretation was performed in our *BMPR2* database. Changes in variant classification might have occurred over time due to the advent of NGS and the detection and interpretation of numerous variants (31). Hence, special care is warranted while determining their possible pathogenicity. Nevertheless, in this case, the change in variant classification did not lead to a clinically significant change.

Here, we demonstrated the feasibility of using MinION sequencing as a tool for genetic analysis in patients with PAH. However, there are some limitations, including Indel variant calling-related errors. It is well-known that the process of nanopore sequencing of homopolymeric regions is difficult and error prone (19, 20, 32). Improvements are essential to enable accurate indel calling. The sample size of our study is small; however, given the satisfactory results obtained, further studies should be performed to analyze more cases.

One advantage is that there is no instrument cost with a cheap starter kit (including MinION device, two flow cells and library preparation kit) costing approximately US$ 1,000. Then single extra flow cell costs US$ 500–900 (depending on the number purchased) and reagents cost approximately US$ 100 per sample. To increase cost-effectiveness, sample multiplexing would be necessary. Moreover, ongoing developments in the flow cell design have contributed to cost-efficient sequencing with the introduction of Flongle (Flow cell Dongle). Flongle is an adapter for MinION that enables DNA sequencing on smaller, lower cost (US$ 90 per flow cell), single-use flow cells. Thus, the lower cost per sample makes Flongle more suitable for targeted sequencing and smaller tests.

Finally, we have focused on *BMPR2* alone, but the analysis of a panel of genes implicated in PAH is a challenge, and should be attempted in future studies.

CONCLUSION

MinION sequencing can be used to detect *BMPR2* mutations in patients with PAH. The MinION device is easy to use and provides long sequence reads. Thus, it can be concluded that the MinION nanopore sequencer is a promising tool for screening *BMPR2* mutations.

DATA AVAILABILITY STATEMENT

The sequencing data presented in this study are deposited in DNA Data Bank of Japan, accession number: DRA012598.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by institutional ethics committee at the International University of Health and Welfare (approval number 5-16-25). The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the relevant individual(s), and/or minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

YT takes responsibility for the content of the manuscript, including the data, and analysis. YT and TT conceived and designed the study. TT performed nanopore sequencing, bioinformatics analysis, and primary data analysis. SB and YT participated in data analysis and interpretation. AF, HT, RT, and AK contributed to sample collection. SB prepared the first draft of the manuscript, along with TT. YT supervised the manuscript preparation process. All authors reviewed the content of the draft, approved the final manuscript, and agreed to be accountable for all aspects of the work.

FUNDING

This work was supported by MEXT KAKENHI, via grant number JP18K08183, and the Actelion Academia Prize and GSK Japan Research Grant 2018.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2021.711694/full#supplementary-material

REFERENCES

1. Rubin LJ. Diagnosis and management of pulmonary arterial hypertension: ACCP evidence-based clinical practice guidelines. *Chest*. (2004) 126(1 Suppl):7S–10S. doi: 10.1378/chest.126.1_suppl.7S
2. Montani D, Günther S, Dorfmüller P, Perros F, Girerd B, Garcia G, et al. Pulmonary arterial hypertension. *Orphanet J Rare Dis*. (2013) 8:97. doi: 10.1186/1750-1172-8-97
3. Galiè N, Humbert M, Vachiery JL, Gibbs S, Lang I, Torbicki A, et al. 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The joint task force for the diagnosis and treatment of pulmonary hypertension of the European society of cardiology (ESC) and the European respiratory society (ERS): Endorsed by: Association for European paediatric and congenital cardiology (AEPC), International society for heart and lung Transplantation (ISHLT). *Eur Heart J*. (2016) 37:67–119. doi: 10.1093/eurheartj/ehv317
4. Simonneau G, Gatzoulis MA, Adatia I, Celermajer D, Denton C, Ghofrani A, et al. Updated clinical classification of pulmonary hypertension. *J Am Coll Cardiol*. (2013) 62(25 Suppl):D34–41. doi: 10.1016/j.jacc.2013.10.029
5. Deng Z, Morse JH, Slager SL, Cuervo N, Moore KJ, Venetos G, et al. Familial primary pulmonary hypertension (gene PPH1) is caused by mutations in the bone morphogenetic protein receptor-II gene. *Am J Hum Genet*. (2000) 67:737–44. doi: 10.1086/303059
6. Lane K, Machado R, Pauciulo M, Thomson JR, Phillips JA III, Loyd JE, et al. Heterozygous germline mutations in *BMPR2*, encoding a TGF-β receptor, cause familial primary pulmonary hypertension. *Nat Genet*. (2000) 26:81–4. doi: 10.1038/79226
7. Evans JD, Girerd B, Montani D, Wang XJ, Galiè N, Austin ED, et al. BMPR2 mutations and survival in pulmonary arterial hypertension: an individual participant data meta-analysis. *Lancet Respir Med*. (2016) 4:129–37. doi: 10.1016/S2213-2600(15)00544-5
8. Morrell NW, Aldred MA, Chung WK, Elliott CG, Nichols WC, Soubrier F, et al. Genetics and genomics of pulmonary arterial hypertension. *Eur Respir J*. (2019) 53:1801899. doi: 10.1183/13993003.01899-2018
9. Southgate L, Machado RD, Gräf S, Morrell NW. Molecular genetic framework underlying pulmonary arterial hypertension. *Nat Rev Cardiol*. (2020) 17:85–95. doi: 10.1038/s41569-019-0242-x
10. Austin ED, Loyd JE. The genetics of pulmonary arterial hypertension. *Circ Res*. (2014) 115:189–202. doi: 10.1055/s-0033-1355443
11. Condon DF, Nickel NP, Anderson R, Mirza S, de Jesus Perez VA. The 6th world symposium on pulmonary hypertension: what's old is new. *F1000Res*. (2019) 8(F1000 Faculty Rev):888. doi: 10.12688/f1000research.18811.1
12. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet*. (2016) 17:333–51. doi: 10.1038/nrg.2016.49
13. van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C. The third revolution in sequencing technology. *Trends Genet*. (2018) 34:666–81. doi: 10.1016/j.tig.2018.05.008
14. Lu H, Giordano F, Ning Z. Oxford nanopore MinION sequencing and genome assembly. *Genom Proteom Bioinf*. (2016) 14:265–79. doi: 10.1016/j.gpb.2016.05.004
15. Loman N. J., Quick J., Simpson J. T. (2015). A complete bacterial genome assembled *de novo* using only nanopore sequencing data. *Nat Methods*. (2015) 12:733–5. doi: 10.1038/nmeth.3444
16. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature*. (2016) 530:228–32. doi: 10.1038/nature16996
17. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat Protoc*. (2017) 12:1261–76. doi: 10.1038/nprot.2017.066
18. Minervini CF, Cumbo C, Orsini P, Brunetti C, Anelli L, Zagaria A, et al. TP53 gene mutation analysis in chronic lymphocytic leukemia by nanopore MinION sequencing. *Diagn Pathol*. (2016) 11:96. doi: 10.1186/s13000-016-0550-y
19. Jain, M., Koren, S., Miga, K, Quick J, Rand AC, Sasani TA, et al. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat Biotechnol*. (2018) 36:338–45. doi: 10.1038/nbt.4060
20. Orsini P, Minervini CF, Cumbo C, Anelli L, Zagaria A, Minervini A, et al. Design and MinION testing of a nanopore targeted gene sequencing panel for chronic lymphocytic leukemia. *Sci Rep*. (2018) 8:11798. doi: 10.1038/s41598-018-30330-y
21. Midha MK, Wu M, Chiu KP. Long-read sequencing in deciphering human genetics to a greater depth. *Hum Genet*. (2019) 138:1201–15. doi: 10.1007/s00439-019-02064-y
22. Scheunert A, Dorfner M, Lingl T, Oberprieler C. Can we use it? On the utility of de novo and reference-based assembly of Nanopore data for plant plastome sequencing. *PLoS ONE*. (2020) 15:e0226234. doi: 10.1371/journal.pone.0226234
23. Alekseyev YO, Fazeli R, Yang S, Basran R, Maher T, Miller NS, et al. A next-generation sequencing primer-how does it work and what can it do? *Acad Pathol*. (2018) 5:2374289518766521. doi: 10.1177/2374289518766521
24. Laver T, Harrison J, O'Neill PA, Moore K, Farbos A, Paszkiewicz K, et al. Assessing the performance of the Oxford nanopore technologies MinION. *Biomol Detect Quantif*. (2015) 3:1–8. doi: 10.1016/j.bdq.2015.02.001

25. Morisaki H, Nakanishi N, Kyotani S, Takashima A, Tomoike H, Morisaki T. *BMPR2* mutations found in Japanese patients with familial and sporadic primary pulmonary hypertension. *Hum Mutat*. (2004) 23:632. doi: 10.1002/humu.9251
26. Machado RD, Aldred MA, James V, Harrison RE, Patel B, Schwalbe EC, et al. Mutations of the TGF-beta type II receptor BMPR2 in pulmonary arterial hypertension. *Hum Mutat*. (2006) 27:121–32. doi: 10.1002/humu.20285
27. Machado RD, Eickelberg O, Elliott CG, Geraci MW, Hanaoka M, Loyd JE, et al. Genetics and genomics of pulmonary arterial hypertension. *J Am Coll Cardiol*. (2009) 54(1 Suppl):S32–42. doi: 10.1016/j.jacc.2009.04.015
28. Girerd B, Montani D, Jaïs X, Eyries M, Yaici A, Sztrymf B, et al. Genetic counselling in a national referral centre for pulmonary hypertension. *Eur Respir J*. (2016) 47:541–52. doi: 10.1183/13993003.00717-2015
29. Kabata H, Satoh T, Kataoka M, Tamura Y, Ono T, Yamamoto M, et al. Bone morphogenetic protein receptor type 2 mutations, clinical phenotypes and outcomes of Japanese patients with sporadic or familial pulmonary hypertension. *Respirology*. (2013) 18:1076–82. doi: 10.1111/resp.12117
30. Machado RD, Southgate L, Eichstaedt CA, Aldred MA, Austin ED, Best DH, et al. Pulmonary arterial hypertension: a current perspective on established and emerging molecular genetic defects. *Hum Mutat*. (2015) 36:1113–27. doi: 10.1002/humu.22904
31. Chisholm C, Daoud H, Ghani M, Mettler G, McGowan-Jordan J, Sinclair-Bourque L, et al. Reinterpretation of sequence variants: one diagnostic laboratory's experience, and the need for standard guidelines. *Genet Med*. (2018) 20:365–8. doi: 10.1038/gim.2017.191
32. Cumbo C, Minervini CF, Orsini P, Anelli L, Zagaria A, Minervini A, et al. Nanopore targeted sequencing for rapid gene mutations detection in acute myeloid leukemia. *Genes (Basel)*. (2019) 10:1026. doi: 10.3390/genes10121026

Conflict of Interest: This work was supported by MEXT KAKENHI, via grant number JP18K08183, and the Actelion Academia Prize and GSK Japan Research Grant 2018. Both Actelion Academia Prize and GSK Japan Research Grant are independent from industrial activities and totally funded on scientific activities. These funds contributed only our experimental procedures and the preparation of manuscripts.

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Cardiovascular Medicine

REVIEW
published: 16 December 2021
doi: 10.3389/fcvm.2021.783398

Recent Progress in Cardiovascular Research Involving Single-Cell Omics Approaches

*Zhehao Dai and Seitaro Nomura**

Department of Cardiovascular Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Cardiovascular diseases are among the leading causes of morbidity and mortality worldwide. Although the spectrum of the heart from development to disease has long been studied, it remains largely enigmatic. The emergence of single-cell omics technologies has provided a powerful toolbox for defining cell heterogeneity, unraveling previously unknown pathways, and revealing intercellular communications, thereby boosting biomedical research and obtaining numerous novel findings over the last 7 years. Not only cell atlases of normal and developing hearts that provided substantial research resources, but also some important findings regarding cell-type-specific disease gene program, could never have been established without single-cell omics technologies. Herein, we briefly describe the latest technological advances in single-cell omics and summarize the major findings achieved by such approaches, with a focus on development and homeostasis of the heart, myocardial infarction, and heart failure.

Keywords: single-cell omics, cardiovascular research, heart development, cardiac homeostasis, myocardial infarction, heart failure

OPEN ACCESS

Edited by:
Christoph Dieterich, Heidelberg University, Germany

Reviewed by:
Robert Kelly, UMR7288 Institut de Biologie du Développement de Marseille (IBDM), France
Monika Gladka, Academic Medical Center, Netherlands

***Correspondence:**
Seitaro Nomura
senomura-cib@umin.ac.jp

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: *26 September 2021*
Accepted: *22 November 2021*
Published: *16 December 2021*

Citation:
Dai Z and Nomura S (2021) Recent Progress in Cardiovascular Research Involving Single-Cell Omics Approaches. Front. Cardiovasc. Med. 8:783398. doi: 10.3389/fcvm.2021.783398

INTRODUCTION

Cardiovascular diseases are among the major causes of morbidity and mortality worldwide. Yet, despite relentless research efforts in recent decades, the complex structure of the heart, its incessantly beating nature, and its wide range of abnormalities and disorders made us still struggle to understand the molecular mechanisms underlying the normal development, homeostasis, and diseases of this organ.

Recent technological advances have enabled researchers to understand biology at the single-cell resolution. Single-cell RNA sequencing (scRNA-seq) and single-cell assay for transposase-accessible chromatin sequencing (scATAC-seq) allow an unbiased assessment of transcriptomics and epigenomics in heterogeneous tissues and have identified not only previously unknown cell populations, but also dynamic cellular transitions and intercellular interactions in various tissues (1, 2). These cutting-edge technologies have been applied not only to healthy humans and mice to create reference maps, namely, the Human Cell Atlas and the Tabula Muris (3, 4), but also to normal development and various diseases. Recent advances in spatial transcriptomics, the molecular profiling method that compensates for the lack of spatial information in scRNA-seq, have furthered our understanding (5). The use of single-cell omics technologies in cardiovascular research has been increasing exponentially since 2015 (**Figure 1**) and, as expected, is providing more insight into the workings of both healthy and pathological hearts.

In this review, we briefly introduce the latest technological advances in single-cell omics and summarize recent progress in cardiovascular research that has applied these technologies, with

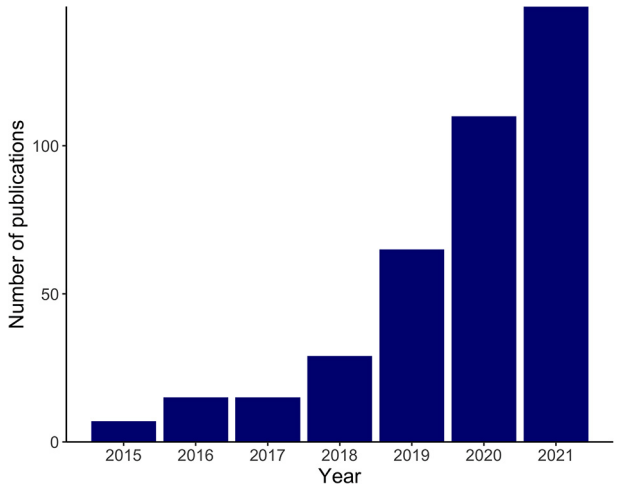


FIGURE 1 | Exponential increase in cardiovascular research implementing single-cell omics. The bar plot shows the numbers of search results in PubMed in recent years with the following search term and the search conducted on November 2, 2021: ("RNA-Seq"[Mesh] OR "Transcriptome"[Mesh] OR "Chromatin Immunoprecipitation Sequencing"[Mesh] OR multi-omics[Title/Abstract] OR multiomics[Title/Abstract] OR omics[Title/Abstract] OR sequencing[Title/Abstract] OR seq[Title/Abstract]) AND ("Single-Cell Analysis"[Mesh] OR "single cell"[Title/Abstract] OR "single nucleus"[Title/Abstract] OR "single nuclei"[Title/Abstract]) AND ("heart"[Mesh] OR heart[Title/Abstract] OR cardiac[Title/Abstract] OR cardiomyocyte[Title/Abstract] OR myocardial[Title/Abstract]).

special focus on development and homeostasis of the heart, myocardial infarction (MI) and ischemia, and heart failure (HF). We also discuss how single-cell omics can, and should support future cardiovascular research and contribute to precision cardiovascular medicine.

TECHNOLOGICAL ADVANCES

The experimental pipelines of single-cell omics, represented by scRNA-seq, typically begin with single-cell isolation and capture, followed by sequencing and downstream analyses (6). In the early days, integrated fluidic circuit (IFC) systems, such as Fluidigm C1, were used to isolate cells into individual chambers, where they were then lysed and reverse-transcribed (plate-based method). The IFC system has been largely replaced by fluorescence-activated cell sorting (FACS), which can not only isolate cells, but also distinguish live cells through viability stains. A major drawback of plate-based systems, however, is the limited number of cells that can be analyzed. To overcome this, droplet-based methods were developed, in which single cells are encapsulated in droplets and then undergo barcoding and reverse transcription. This technology has greatly improved throughput, increasing the number of cells analyzed to up to 10,000 cells, albeit with a sacrifice in read depth.

Both IFC systems used in plate-based methods and microfluidic devices used in droplet-based methods limit the size of the cells that can be processed to $< \sim 40\,\mu m$, which are applicable to non-cardiomyocytes, neonate cardiomyocytes (CMs), and pluripotent stem cell-derived CMs. However, adult CMs ($>100\,\mu m$ in length) cannot be processed by these methods. Moreover, commercially available FACS nozzles (70–130 μm in diameter) might induce terminal damage to live CMs and so are not suitable either (6, 7). Although successful isolation of viable single CMs using large-particle FACS with a nozzle of 500 μm in diameter has recently been reported (8), most published studies applying scRNA-seq in CMs have relied on either manual pick-up of CMs combined with plate-based methods (9) or the iCELL8 system, a large nozzle dispenser with a 5,184-nanowell plate and an imaging system to differentiate live cells from dead ones (10). An alternative solution is to isolate and sequence the single nuclei of CMs (11), which are smaller and can pass through FACS and droplet-based systems. Single-nucleus RNA sequencing (snRNA-seq), although is applicable in archived frozen specimens and can minimize the alteration of gene expression that may be caused by dissociation (12), should be used and its results interpreted with awareness of how it differs from scRNA-seq. A comparison of snRNA-seq and scRNA-seq showed that they both demonstrated similar cell clusters in the process of induced pluripotent stem cell (iPSC) differentiation to CMs. However, scRNA-seq captures more genes than snRNA-seq, represented by mitochondrial and ribosomal genes, while snRNA-seq detects more non-coding RNA and more reads that are mapped to intronic regions (11, 13). Importantly, the multinucleation and polyploidy of CMs also contribute to the divergence of snRNA-seq from scRNA-seq when applied to CMs, thus should be taken into consideration (14).

Besides the advances made in single-cell transcriptomics, its integration with single-cell epigenomics, namely, scATAC-seq, provides more information on both cell clustering and interactions between gene expression and open chromatin states (15, 16). Furthermore, spatial transcriptomics (i.e., technologies that recover transcriptomic information from tissue while preserving its spatial localization) is undergoing exponential development. Spatial transcriptomics uses either multiplexed RNA imaging [e.g., sequential fluorescence *in situ* hybridization (seqFISH)] or spatial barcoding (e.g., Visium, slide-seq) (5). Because the structure of the heart is complex and its tissues are highly heterogeneous, the addition of spatial information substantially improves our understanding of tissue architectures. It is also better able to uncover intercellular communications, thereby boosting the exploration of the underlying molecular pathways in the developing, normal, and diseased heart (5, 17).

In the following sections of this review, we focus on recent cardiovascular research that has applied single-cell omics technologies.

DEVELOPMENT OF THE HEART

Early scRNA-seq studies of the heart mainly focused on its development because fetal and neonatal cardiac tissues are easier to digest and their cells, including CMs, are smaller and so can pass through various single-cell platforms. Several such studies are highlighted in **Table 1**.

TABLE 1 | Summary of studies of heart development.

References	Species	Developmental stage	Target tissues/cells	Modalities	Major findings
Chan et al. (18)	Mouse	Mouse ESC-derived embryoid bodies day 4	Dissociated cells from embryoid bodies	scRNA-seq	*Mesp1*$^+$ mesoderm cells are highly heterogeneous
Lescroart et al. (19)	Mouse	Embryonic day 6.25 and 7.5	*Mesp1*$^+$ or *Mesp1* KO CPCs	scRNA-seq	Mesp1 was essential for cell exit from the pluripotent state and the induction of the cardiac gene expression program
DeLaughter et al. (20)	Mouse	Embryonic day 9.5, 11.5, 14.5, and 18.5; postnatal day 0, 3, and 21	Microdissected embryonic heart tissues of each chamber	scRNA-seq	Single cells were classified into CMs, ECs, and fibroblast-enriched cells. Markers of temporal- and chamber-specific developmental programs were identified. *Nkx2-5* haploinsufficiency delayed developmental programs in both CMs and ECs
Li et al. (21)	Mouse	Embryonic day 8.5, 9.5, and 10.5	Microdissected embryonic heart tissues of each chamber	scRNA-seq	Using a random forest algorithm, the origins of single CMs were successfully reconstructed. Loss of *Nks2-5* led to failed differentiation into ventricular CMs
Jia et al. (15)	Mouse	Embryonic day 7.5, 8.5, and 9.5	*Isl1*$^+$ or *Nkx2-5*$^+$ CPCs	scRNA-seq and scATAC-seq	*Isl1*$^+$ CPCs passed through an attractor state before segregating into different developmental branches. Continued expression of *Nkx2-5* established a unidirectional cardiomyocyte fate for CPCs
Xiong et al. (22)	Mouse	Embryonic day 7.75, 8.25, and 9.75	*Isl1*$^+$ or *Nkx2-5*$^+$ CPCs	scRNA-seq, CHIP-seq	CPCs in first and second heart fields exhibited different differentiation kinetics toward CMs. FHF cardiomyocytes guided the migration of SHF cells through the MIF-CXCR2/CXCR4 chemotaxis
Ivanovitch et al. (23)	Mouse	Embryonic day 6.9 (mid to late streak), day 7.1 (no-bud to early-bud), and day 7.3 (early headfold),	Tamoxifen treated $T^{nEGFP-CreERT2/+}$;$R26R^{tdTomato/tdTomato}$ embryos	scRNA-seq	scRNA-seq demonstrated that the primitive streak cells contributing to the ventricles had a distinct molecular signature from those forming the outflow tract and atria, suggesting that cardiac progenitors were prepatterned in the primitive streak before migration.
de Soysa et al. (24)	Mouse	Embryonic day 7.75, 8.25, and 9.25	CPCs from control and *Hand2*-null embryos	scRNA-seq	*Hand2* was a specifier of outflow tract cells. *Hand2*-null mouse embryos failed to develop the right ventricle, which was due to failure of outflow tract myocardium specification
Xiao et al. (25)	Mouse	Embryonic day 13.5 and 14.5	Cardiac tissue from control or *Lats1/2* conditional (epicardium) KO mice	scRNA-seq	*Lats1/2* in EPDCs were essential for their differentiation into cardiac fibroblasts and for coronary vessel remodeling
Quijada et al. (26)	Mouse	Embryonic day 12.5 and 16.5; embryonic day 14.5	EPDCs from control at embryonic day 12.5 and 16.5; from conditional KO of myocardin-related transcription factors in the epicardium, embryonic day 14.5	scRNA-seq	*Slit2*$^+$ EPDCs emerged following EMT. Genetic disruption of EMT altered the expression of vascular guidance cues such as *Slit2* and disturbed EC maturation and localization in the coronary vasculature
Sereti et al. (27)	Mouse	Embryonic day 9.5 and 12.5 and postnatal day 1	αMHC$^+$ cardiomyocytes	scRNA-seq	Expression of cell cycle genes decreased to a minimal level postnatally, along with CM maturation
Su et al. (28)	Mouse	Embryonic day 12.5 and 14.5	Sinus venosus-derived ECs from control at embryonic day 12.5; from control and *Nr2f2* overexpression hearts at embryonic day 14.5	scRNA-seq	Vein cells underwent an early cell fate switch to pre-artery ECs. NR2F2 inhibited the pre-artery population *via* the induction of cell cycle genes
Li et al. (29)	Mouse	Embryonic day 10.5	Cardiac cells from dissected heart chambers and *Isl1*$^+$ cardiac cells	scRNA-seq	CMs in G2/M phase downregulated sarcomeric and cytoskeletal markers
Wunnemann et al. (30)	Mouse	Embryonic day 14.5	Wild-type and *Adamts19*-null heart cells	scRNA-seq	*Adamts19* was a marker gene of valvular interstitial cells. Gene regulatory network analysis positioned *Adamts19* in a network downstream of the WNT signaling pathway

(Continued)

TABLE 1 | Continued

References	Species	Developmental stage	Target tissues/cells	Modalities	Major findings
Hulin et al. (31)	Mouse	Postnatal day 7 and 30	Cells from aortic and mitral valve leaflets	scRNA-seq	Subpopulations of endothelial and immune cells were comparable between the two time points, whereas interstitial cells were more diverse at postnatal day 30, expressing complement factors, ECM proteins, and osteogenic genes
Wang et al. (32)	Mouse	Postnatal day 1, 4, 7, 14 and 56	CMs and non-CMs from left ventricles	scRNA-seq	Genes involved in ECM organization, BMP signaling pathway and CM differentiation were enriched in mature fibroblasts when compared with immature ones. Maturation of CMs were promoted when co-cultured with adult fibroblasts, but not with neonatal fibroblast.
Goodyer et al. (33)	Mouse	Embryonic day 16.5	Cells from three zones of microdissected hearts: sinoatrial node region, atrioventricular node/His region, and bundle branch/Purkinje fiber region	scRNA-seq	The study revealed clusters of sinoatrial node cells, atrioventricular node/His cells, and Purkinje fiber and transitional Purkinje fiber cells, as well as their cell subtypes, and thus delineated the first transcriptional landscape of the developing cardiac conduction system
Suryawanshi et al. (34)	Human	19–22 weeks gestation	Cells from healthy or autoimmune-associated CHB fetal hearts	scRNA-seq	Fetal heart with autoimmune-associated CHB exhibited varying degrees of increased interferon responses in all cell types. Matrisome transcripts were highly enriched in the fibroblasts and smooth muscle cells in the CHB sample
Sahara et al. (35)	Human	Human ESC-derived cardiac lineages day 3, 6, and 15; human embryonic/fetal hearts 4.5–10 weeks gestation	Human ESC-derived cardiac lineages; human embryonic/fetal cardiac cells	scRNA-seq	Cono-ventricular progenitors were marked by *LGR5*, which is human-specific. *LGR5*$^{+}$ progenitors promoted cardiogenesis *via* expansion of committed cardiac intermediates
Cui et al. (36)	Human	5–25 weeks gestation	Anatomically informed cardiac cells from human embryos	scRNA-seq	The study identified major cell types and demonstrated the similarities in and differences between humans and mice
Lahm et al. (37)	Human/mouse	4.5–10 weeks gestation and adults	Human embryonic/fetal cardiac cells; adult human atrial and ventricular cells	scRNA-seq	The four detected genes—*MACROD2*, *GOSR2*, *WNT3*, and *MSX1*—were highly expressed throughout development, and *MACROD2* remained expressed in adult heart
Asp et al. (38)	Human	4.5–5, 6.5, and 9 weeks post-conception	Human embryonic/fetal cardiac cells	scRNA-seq, spatial barcoding, and *in situ* sequencing	The study generated the first spatiotemporal cell atlas for the developing human heart
Tyser et al. (39)	Mouse	Embryonic day 7.75–8.25 (divided into six stages)	Microdissected anterior cardiac region	scRNA-seq, multiplexed RNA imaging	A previously unknown CPC pool marked by *Mab21l2* was identified, which contributed to not only CMs, but also the epicardium
Mantri et al. (40)	Chicken	Embryonic day 4, 7, 10, and 14 (corresponding to HH21–24, HH30–31, HH35–36, and ∼HH40)	Enzymatically digested cardiac ventricular tissues	scRNA-seq, spatial barcoding, and multiplexed RNA imaging	Spatially restricted genes during development were identified. The study also discovered a *TMSB4X*-expressing subpopulation that overlapped coronary vascular cells throughout coronary development

CHB, congenital heart block; ChIP-seq, chromatin immunoprecipitation with parallel sequencing; CM, cardiomyocyte; CPC, cardiac progenitor cell; EC, endothelial cell; EMT, epithelial-to-mesenchymal transition; ESC, embryonic stem cell; FHF, first heart field; HH, Hamburger-Hamilton; KO, knockout; scATAC-seq, single-cell assay for transposase-accessible chromatin sequencing; scRNA-seq, single-cell RNA sequencing; SHF, second heart field.

It had been unclear whether *Mesp1*$^+$ mesoderm cells were intrinsically heterogeneous or were homogeneous but capable of multiple lineage decisions. Taking advantage of scRNA-seq, Chan et al. unraveled the heterogeneity of *Mesp1*$^+$ mesoderm cells from mouse embryonic stem cell (ESC)-derived embryoid bodies (18). Lescroart et al., by examining *Mesp1*$^+$ cardiac progenitor cells (CPCs) as well as *Mesp1*-expressing cells in the *Mesp1* knockout context, found that *Mesp1* was essential for exit from the pluripotent state and for induction of the cardiac gene expression program (19).

Efforts to establish a transcriptomic atlas of the developing heart have been undertaken by multiple teams. By investigating >1,200 cells from murine heart chambers of different developmental stages, DeLaughter et al. not only successfully clustered CMs, endothelial cells (ECs), and fibroblast-enriched cells, but also identified markers of temporal- and chamber-specific developmental programs (e.g., genes involved in cell cycle at embryonic day 9.5–11.5 and genes involved in cellular growth and CM differentiation between embryonic day 14.5 and postnatal day 3) (20). Li et al. used a similar approach together with a random forest algorithm to reconstruct the spatial origins of single CMs (21). Both studies demonstrated that *Nkx2-5* deficiency led to failure of CM maturation (20, 21). Jia et al. integrated scRNA-seq and scATAC-seq in *Nkx2-5*$^+$ and *Isl1*$^+$ CPCs, enabling developmental trajectories to be reconstructed and revealing that *Isl1*$^+$ CPCs passed through an attractor state before segregating into different developmental branches, whereas continued expression of *Nkx2-5* established a unidirectional CM fate for CPCs (15). A similar analysis of *Nkx2-5*$^+$ and *Isl1*$^+$ CPCs found interactions between first and second heart field CPCs, with second heart field CPCs attracted to the first heart field-populated heart tube region through MIF (macrophage migration inhibitory factor)-CXCR2/CXCR4 chemotaxis, contributing to the growth of the outflow tract and right ventricle (22). Furthermore, using genetic lineage tracing and live imaging, Ivanovitch et al. discovered the temporal order in which different cardiac lineages arose within the primitive streak: the progenitors of the left ventricle at the mid-streak stage, those of the right ventricle at the late-streak stage, and those of the outflow tract and atria at the no-bud to early-bud stage. More importantly, scRNA-seq demonstrated distinct molecular signatures of these subpopulations of primitive streak cells, indicating their prepatterned fate in the primitive streak before migration (23). Using a Boolean network-based lineage-specifier prediction method in the downstream analysis of scRNA-seq of mouse embryos as an alternative approach, de Soysa et al. identified *Hand2* as a specifier of outflow tract cells. *Hand2*-null mouse embryos failed to specify outflow tract myocardium, whereas the right ventricle myocardium was specified but not properly differentiated and migrated, which was associated with disrupted anterior–posterior patterning of CPCs (24).

With a focus on epicardium-derived cells (EPDCs), Xiao et al. demonstrated that epicardium conditional knockout of *Lats1/2*, encoding LATS1 and LATS2 that are negative regulators in the Hippo signaling pathway, caused arrested differentiation into fibroblasts and defective coronary vasculature remodeling (25). Quijada et al. also investigated epicardial contributions to the growing coronary plexus by applying scRNA-seq to EPDCs and uncovered a *Slit2*$^+$ subpopulation that emerged following epithelial-to-mesenchymal transition (EMT). Genetic disruption of EMT altered the expression of vascular guidance cues such as *Slit2* and disturbed EC maturation and localization in the coronary vasculature (26). Su et al. analyzed sinus venosus-derived ECs, discovered that the vein cells underwent an early cell fate switch to pre-artery ECs, and identified the transcription factor NR2F2 as a vein specifier, whose overexpression prevented differentiation to the pre-artery population *via* the induction of cell cycle genes (28). By performing a genome-wide association study (GWAS) in humans, Wunnemann et al. detected mutations in *ADAMTS19* that are responsible for valvular diseases. They then applied scRNA-seq to wild-type and *Adamts19*-null murine heart at embryonic day 14.5 and identified *Admats19* as a marker of valve interstitial cells. Gene regulatory network analysis helped to identify the Wnt-Adamts19-Klf2 axis as being essential for valve development (30). Another study of murine postnatal valve leaflets demonstrated that, whereas subpopulations of ECs and immune cells were comparable between day 7 and 30, interstitial cells were more diverse at postnatal day 30, expressing complement factors, extracellular matrix proteins, and osteogenic genes (31). Wang et al. analyzed the dynamic change of subtypes of non-CMs along with postnatal CM maturation in murine hearts. scRNA-seq of CMs and non-CMs of murine hearts at postnatal day 1, 4, 7, 14, and 56 showed that fibroblasts increasingly expressed maturation promoting genes such as *Dcn*. Co-culture experiments demonstrated that maturation of CMs was promoted when co-cultured with adult fibroblasts, but not with neonatal fibroblasts. Their results indicated fibroblasts subtype switching as a crucial microenvironmental event regulating postnatal CM maturation (32).

Development of the conduction system has also attracted researchers' interest. Goodyer et al. performed scRNA-seq on cells from murine embryonic heart regions including the conduction system. The study revealed clusters of sinoatrial node cells, atrioventricular node/His cells, and Purkinje fiber and transitional Purkinje fiber cells, thus delineated the first transcriptional landscape of the developing cardiac conduction system (33). Suryawanshi et al. investigated single-cell transcriptomes from healthy human fetal hearts and a fetal heart with autoimmune-associated congenital heart block and showed, through a comparison with normal fetal hearts, that the fetal heart with autoimmune-associated congenital heart block exhibited increased interferon responses of varying degrees in all cell types. Additionally, matrisome transcripts were highly enriched in the fibroblasts and smooth muscle cells in the congenital heart block sample compared with healthy hearts, indicating their contribution to fibrosis (34).

Other studies of human heart development, although rare, have also been conducted to unravel the human-specific features. Cui et al. identified cell types from human fetal hearts and compared their gene expression profiles with those of mice (36). A study of human ESC-derived cardiac derivates and human embryonic/fetal hearts revealed a human-specific cono-ventricular progenitor population marked by *LGR5*. LGR5

(leucine rich repeat containing G protein-coupled receptor 5) is a receptor for R-spondins, which is involved in the WNT signaling pathway. *LGR5*$^{+}$ progenitors were shown to promote cardiogenesis *via* expansion of committed cardiac intermediates (35). Lahm et al. identified four congenital heart disease-associated genes—*MACROD2*, *GOSR2*, *WNT3*, and *MSX1*—by GWAS and explored their expression in embryonic and adult human heart by scRNA-seq. They found that *MACROD2*, *GOSR2*, *WNT3*, and *MSX1* were highly expressed throughout development and that *MACROD2* expression persisted in adult heart (37).

More recent research implemented multiple advanced technologies along with scRNA-seq. Asp et al. generated the first spatiotemporal cell atlas for human developing heart by integrating scRNA-seq, spatial barcoding, and multiplexed RNA imaging (38). Tyser et al., by combining scRNA-seq with genetic lineage labeling and abundant imaging modalities including multiplexed RNA *in situ* hybridization in murine embryonic hearts, identified and located the CPC subpopulations. *Via* this spatial transcriptomic approach, they unraveled a previously unknown progenitor subpopulation marked by *Mab21l2* that located at the rostral border of the cardiac crescent and contributed to both CMs and epicardium. Mab21l2 (Mab-21 like 2) was previously known to be required for neural and eye development, was here implied to represent the earliest progenitors of the epicardium (39). By integrating scRNA-seq and spatial transcriptomics in developing chicken hearts, Mantri et al. identified spatially restricted genes in cardiac tissue during development. They found that *TBX5* was enriched in the left ventricle compared with the right but decreased from embryonic day 7 to 14 and that *CHGB* expression was restricted to the right ventricle from day 7 onward. In addition, they discovered a *TMSB4X*-expressing subpopulation that overlapped with coronary vascular cells throughout coronary development. TMSB4 (thymosin β4) participates in regulating actin polymerization, thus is involved in cell proliferation, migration, and differentiation. Its role in heart development had long been under significant debate, given the paradoxical results of several previous *Tmsbx* knockout and knockdown experiments in mouse models, and was provided deeper insights into by this study (40). Undoubtedly, the latest technological advances, represented by spatial transcriptomics, are furthering our understanding of the molecular mechanisms of cardiac development.

HOMEOSTASIS OF THE ADULT HEART

Single-cell omics approaches have also been implemented in studies of adult mice and humans in homeostasis (**Table 2**), with more studies focusing on non-CMs than on CMs. Yekelchyk et al. performed scRNA-seq on murine CMs using the iCELL8 platform, finding that gene expression was homogeneous between mono- and multi-nucleated CMs in heathy ventricles (41). By performing scRNA-seq on non-CMs of murine heart, Skelly et al. detected major cell types and transcriptional heterogeneity, revealed communications between cell types, and showed cell type-specific sexual dimorphism of cardiac gene expression, such as male upregulation of *Irf8* (interferon regulatory factor 8) and female upregulation of *Tsc22d3* (TSC22 domain family member 3, a transcription factor implicated in anti-inflammatory functions) in macrophages. Their results provided molecular cues for the sexual difference of cardiac responses to the environmental insults (42). Chakarov et al. identified two distinct interstitial macrophage populations occupying different niches, which were conserved across lung, heart, fat, and dermis tissues: Lyve1^{low}MHCII^{high}CX3CR1^{high} and Lyve1^{high}MHCII^{low}CX3CR1^{low} monocyte-derived interstitial macrophages. They also demonstrated that depletion of Lyve1^{high}MHCII^{low}CX3CR1^{low} interstitial macrophages exacerbated lung and heart fibrosis (43), which was highly consistent with the results obtained using a mouse MI model (55). Macrophages also facilitate conduction in the atrioventricular node. Genes involved in cardiac conduction are enriched in macrophages from the atrioventricular node when compared with those from other tissues. Atrioventricular node macrophages highly express *Cx43*, contributing to the gap junctions that link them to neighboring CMs (44). In regard to the conduction system, quantitative proteomics of the murine sinus node demonstrated a remarkable abundance of ion channels responsible for the pacemaking process (e.g., HCN4), which were indicated by snRNA-seq to be predominantly expressed by sinus node myocytes (45). Liang et al. then verified the existence of a cell cluster in the sinus node marked by *Hcn1* and *Hcn4*, where *Vsnl1* was also highly expressed. VSNL1 (visinin-like 1) is a member of visinin/recoverin subfamily of neuronal calcium sensor proteins, and is involved in calcium signaling pathways. Adeno-associated virus-mediated knockdown of *Vsnl1* in mice resulted in a reduced heart rate and decreased expression of *Hcn4*, *Cacna1d*, *Cacna1i*, and *Serca2a*, but not that of *Ryr2* or of the genes involved in potassium channels (46).

Cardiac ECs are another research interest. Paik et al. re-explored the scRNA-seq data from tissue-specific ECs in the Tabula Muris (4) and demonstrated that cardiac ECs could be classified as a separate cluster, even though they had considerable transcriptomic overlap with ECs from other tissues (47). Yucel et al. unexpectedly observed that adult murine cardiac ECs actively expressed cardiac myofibril genes such as *Tnnt2* and *Myh6*, which was consistent with the analytical results of multiple datasets; this result was validated by RNA *in situ* hybridization. They also showed that cardiac ECs had open chromatins at cardiac myofibril gene promoters, suggesting that the results were not due to technical contamination or paracrine transfer of mRNA. Interestingly, the accessibility of cardiac myofibril genes was no longer maintained in cardiac ECs upon culture. Although the significance of the expression of cardiac myofibril genes in cardiac ECs is still unclear, the authors hypothesized that those genes played a role in maintaining cardiac EC maturity and specificity, which needs further investigation (48). Hu et al. used scRNA-seq to analyze cells from human aorta, pulmonary artery, and coronary arteries collected from patients who underwent heart transplantation. Artery-specific cell subpopulations with distinct transcriptional profiles were

TABLE 2 | Summary of studies of homeostasis of the adult heart.

References	Species	Target tissues/cells	Modalities	Major findings
Yekelchyk et al. (41)	Mouse	CMs from both healthy and hypertrophic ventricles	scRNA-seq	Gene expression was homogeneous between mono- and multi-nucleated CMs in homeostasis. Heterogeneity was introduced by TAC
Skelly et al. (42)	Mouse	Non-CMs from the heart	scRNA-seq	Major cell types and transcriptional heterogeneity were detected, as well as communications among cell types. The study also showed cell type-specific sexual dimorphism of cardiac gene expression
Chakarov et al. (43)	Mouse	Lung interstitial macrophages (validated in the heart, fat, and skin by flow cytometry)	scRNA-seq	Two distinct IM populations conserved across the lung, heart, fat, and dermis tissues were identified. Depletion of Lyve1^{high}MHCII^{low}CX3CR1^{low} IMs exacerbated lung and heart fibrosis
Hulsmans et al. (44)	Mouse	Macrophages in the AV nodes	scRNA-seq	Genes involved in cardiac conduction were enriched in macrophages from the AV node compared with those from other tissues. AV node macrophages highly expressed *Cx43*, forming the gap junctions linking them to neighboring CMs
Linscheid et al. (45)	Mouse	Cells from sinus node tissue and the adjacent atrium	snRNA-seq, bulk proteomics	Quantitative proteomics of murine sinus node demonstrated significant abundancy of ion channels responsible for the pacemaking process (e.g., HCN4), which were predominantly expressed by sinus node myocytes, as revealed by snRNA-seq
Liang et al. (46)	Mouse, rabbit, monkey	Cells from microdissected sinus node tissues and those from atrial and ventricular CMs	scRNA-seq	A cell cluster expressing *Hcn1*, *Hcn4*, and *Vsnl1* was found. Knockdown of *Vsnl1* in mice resulted in reduced heart rate and decreased *Hcn4*, *Cacna1d*, *Cacna1i*, and *Serca2a* expression, but not that of *Ryr2* or those involved in potassium channels
Paik et al. (47)	Mouse	scRNA-seq data from tissue-specific ECs from the Tabula Muris	scRNA-seq	Cardiac ECs could be classified as a separate cluster but had considerable transcriptomic overlap with ECs from other tissues. Markers of tissue-specific ECs identified in mice were also enriched in their corresponding human tissue-specific ECs
Yucel et al. (48)	Mouse	ECs and non-ECs from murine heart	Bulk RNA-seq, snRNA-seq, bulk ATAC-seq, multiplexed RNA imaging	Cardiac ECs actively expressed cardiac myofibril genes such as *Tnnt2* and *Myh6* and had open chromatin at cardiac myofibril gene promoters
Hu et al. (49)	Human	Cells from human aorta, pulmonary artery, and coronary arteries collected from patients undergoing heart transplantation	scRNA-seq	Artery-specific cell subpopulations with distinct transcriptional profiles were identified in VSMCs and fibroblasts, but not in ECs. Intercellular communication between macrophages and ECs was predicted to increase in atherosclerosis
Wolfien et al. (50)	Mouse	Nuclei isolated from entire murine hearts	snRNA-seq	Distinct cell clusters were identified, including immune cells and cells of neuronal origin. RNA velocity enabled interrogation of transcriptional kinetics
Vidal et al. (51)	Mouse	Nuclei isolated from entire hearts of 12-week-old and 18-month-old mice	snRNA-seq	Angiogenesis-related extracellular protein-encoding genes including *Serpine1* and *Serpine2* were enriched in fibroblasts derived from aged hearts, indicating an affected paracrine crosstalk between fibroblasts and ECs during aging. Conditioned medium derived from aged fibroblasts had higher levels of serpins and showed a reduced angiogenic property, which was mediated by an impairment of EC function
Tucker et al. (52)	Human	Tissue samples taken from the lateral aspect of the four cardiac chambers from potential transplant donors	snRNA-seq	CMs were the most heterogeneous of various cell types. When combined with GWAS data, genes at the loci associated with heart rhythm were enriched in CMs, whereas those associated with CAD were enriched in pericytes
Litvinukova et al. (53)	Human	Full-thickness myocardial biopsies from the left and right atria, left and right ventricles, and interventricular septum and apex from deceased transplant organ donors	scRNA-seq, snRNA-seq, and multiplexed RNA imaging	CMs, pericytes, and fibroblasts were the most heterogeneous. Predicted intercellular communications among CMs, fibroblasts, and immune cells differed between atria and ventricles
Wang et al. (54)	Mouse	Murine heart non-myocytes	scRNA-seq, scATAC-seq	Differential accessibilities of the *cis*-regulatory elements among different subpopulations in each cell type that potentially regulate the expression of marker genes

ATAC-seq, assay for transposase-accessible chromatin sequencing; AV, atrioventricular; CAD, coronary artery disease; CM, cardiomyocyte; GWAS, genome-wide association study; IM, interstitial macrophage; EC, endothelial cell; scATAC-seq, single-cell assay for transposase-accessible chromatin sequencing; scRNA-seq, single-cell RNA sequencing; snRNA-seq, single-nucleus RNA sequencing; TAC, transverse aorta constriction; VSMC, vascular smooth muscle cell.

identified in vascular smooth muscle cells and fibroblasts, but not in ECs (49).

Recently, efforts have focused on depicting the full picture of the cell atlas of the whole heart. snRNA-seq of murine hearts successfully clustered distinct cell subpopulations, including immune cells, cells of neuronal origin, and cardiac glial cells (50). With snRNA-seq, Vidal et al. investigated the transcriptomes of murine hearts of 12-week-old and 18-month-old, and found upregulated angiogenesis-related extracellular protein-encoding genes in aged hearts, indicating an affected paracrine crosstalk between fibroblasts and ECs during aging. *In vitro* experiments confirmed that conditioned medium derived from aged fibroblasts showed a reduced angiogenic property, which was mediated by an impairment of EC function (51). In humans, samples of myocardium from cardiac chambers, rather than of the whole heart, have been studied. By applying snRNA-seq to biopsied human cardiac chambers, Tucker et al. showed that CMs were the most spatially heterogeneous of the distinct cell types. Furthermore, by combining the snRNA-seq results with GWAS data for cardiometabolic traits, they revealed that genes at the loci associated with heart rhythm were enriched in CMs, whereas those at the loci associated with coronary artery disease were enriched in pericytes (52). Litvinukova et al. investigated full-thickness biopsy samples from deceased transplant donors and were able to provide more in-depth information, predicting different intercellular communications among immune cells, CMs, and fibroblasts between ventricles and atria (53). Wang et al. integrated scRNA-seq and scATAC-seq for non-CMs in murine hearts and revealed the heterogeneity of non-CM populations and their subpopulations at transcriptomic and epigenomic levels. Cardiac fibroblasts were classified into three distinct functional states related to response to stimuli (marked by *Hsd11b1*; 11β-hydroxysteroid dehydrogenase type 1), the cytoskeleton (marked by *Gfpt2*; glutaminefructose-6-phosphate transaminase 2, involved in the hexosamine biosynthesis pathway), and immune regulation (marked by *C1qa*), respectively. Importantly, they also identified differential accessibilities of the *cis*-regulatory elements among different subpopulations of each cell type that potentially regulate the expression of marker genes (54). The success of these studies indicates that single-cell omics represents an indispensable toolbox for broadening our insights into the homeostasis of the adult heart, particularly the heterogeneity of cells and their functions, as well as intercellular communications.

MYOCARDIAL INFARCTION/ISCHEMIA-REPERFUSION INJURY

In addition to development and adult homeostasis, single-cell omics approaches have played a crucial role in research into diseased hearts (**Table 3**). MI, as a consequence of atherosclerotic disease, may have a marked effect on cardiac function and quality of life. The toolbox from single-cell omics technologies has enabled detailed investigation of not only the cells involved in the process of plaque formation and rupture, but also the ischemic myocardium that is susceptible to remodeling (71).

MI has long been known to involve both pro- and anti-inflammatory cascades. King et al. were the first to apply scRNA-seq to leukocytes isolated from infarcted and non-infarcted murine hearts, finding that MI induced an IRF3-dependent type I interferon response in a distinct subpopulation of cardiac macrophages and that interruption of IRF3-dependent signaling decreased the cardiac expression of inflammatory cytokines and chemokines and improved cardiac function (56). Cardiac-infiltrating neutrophils were demonstrated to locally acquire a SiglecF^{high} ICAM^{high}CXCR4^{high} state from day 3 after MI, indicating local neutrophil activation and aging. Similar neutrophil subpopulation was also found in atherosclerotic aortas, suggesting their roles both in acute and chronic cardiovascular inflammation (57). Regulatory T cells (Tregs) accumulated in the injured myocardium after MI or ischemia-reperfusion injury. The majority of these heart Tregs were derived from the circulating Treg pool; a considerable fraction of them experienced clonal expansion and displayed a unique T-cell receptor repertoire. Importantly, extracellular matrix-associated genes, including *Sparc*, were enriched in heart Tregs compared with other Tregs. Further experiments revealed that heart Tregs contributed to increased collagen content and prevented rupture and that *Sparc* was critical in this process (58). After MI, cardiac B cells also accumulated rapidly *via* the CXCL13-CXCR5 axis but were largely polyclonal and contributed to local TGFβ production, indicating their involvement in fibroblast activation (59).

Fibroblasts are directly involved in post-infarct scar formation. Gladka et al. performed scRNA-seq on cells sorted by FACS with a 130-μm nozzle from infarcted and control murine hearts and were able to identify distinct clusters of CMs, fibroblasts, ECs, and macrophages. A subset of activated fibroblasts was revealed to be specific to injured hearts and to strongly express *Ckap4.* CKAP4 (cytoskeleton associated protein 4) is a transmembrane receptor whose function in cardiac fibroblasts had remained unknown. *In vitro* inhibition of *Ckap4* resulted in increased expression of genes related to activated fibroblasts after TGFβ induction, suggesting a modulating function for CKAP4 in fibroblast activation (60). Molenaar et al. focused on the ligands upregulated in CMs 1 day after ischemia, whose cognate receptors were expressed by another cell type, and detected *B2m* (β2 microglobulin) in CMs and its receptors in fibroblasts. *In vitro* experiments demonstrated that fibroblasts expressed myofibroblast markers when cultured with B2M (61). Farbehi et al. discovered a novel fibroblast subpopulation with upregulated *Wif1* expression, in sham hearts, which persisted after MI, had an anti-WNT/CTGF/TGFβ signature, and interacted with ECs. Immunofluorescence revealed WIF1^{+} fibroblasts in close proximity with ECs in the border zone at post-MI day 3, but not in sham or post-MI day 1 or 7 hearts, indicating the injury-dependent and post-transcriptional regulation of WIF expression, and its function of inhibiting fibrosis and angiogenesis in post-MI repair (62). By investigating the murine heart 14 days after MI, Kretzschmar et al. identified a subpopulation of proliferative fibroblasts resembling neonatal

TABLE 3 | Summary of studies of myocardial infarction or ischemia-reperfusion injury.

References	Species	Ischemia model	Target tissues/cells	Modalities	Major findings
King et al. (56)	Mouse	Permanent ligation of left coronary artery	Leukocytes isolated from wild-type and *Irf3*-null heart at day 4 post-MI or sham	scRNA-seq	MI induced an IRF3-dependent type I interferon response in a distinct subpopulation of cardiac macrophages. Interruption of IRF3-dependent signaling decreased the cardiac expression of inflammatory cytokines and chemokines and improved cardiac function and survival
Vafadarnejad et al. (57)	Mouse	Permanent ligation of LAD	Neutrophils isolated from infarcted hearts and blood at day 1, 3, and 5 post-MI or sham	scRNA-seq, CITE-seq	Post-MI cardiac neutrophils were temporally heterogeneous. Infiltrating neutrophils were demonstrated to locally acquire a SiglecF^{high} state at day 3 onward after MI, in which transcripts associated with neutrophil aging and activation were also enriched
Xia et al. (58)	Mouse	Permanent ligation of LAD; 60-min ligation of LAD	Regulatory and conventional T cells from heart, spleen, non-draining LNs, and mediastinal LNs at day 7 post-MI or sham	Bulk RNA-seq, scRNA-seq, scTCR-seq	Tregs were recruited to the injured myocardium after MI or I/R injury from the circulating Treg pool. A considerable fraction of heart Tregs was clonally expanded. ECM-associated genes, including Sparc, were enriched in heart Tregs. Heart Tregs led to increased collagen content and prevented rupture, with Sparc playing a critical role in this process
Heinrichs et al. (59)	Mouse	Permanent ligation of LAD	B cells from heart and mediastinal LNs at day 5 post-MI or sham	scRNA-seq, scBCR-seq	Cardiac B cells accumulated rapidly after MI *via* the CXCL13-CXCR5 axis. These cells were largely polyclonal and contributed to local TGFβ production. CXCR5-deficiency reduced B cell infiltration and local *Tgfb1* expression but did not change post-MI contractile function or myocardial morphology
Gladka et al. (60)	Mouse	75-min ligation of left coronary artery	Cells from the infarct and border zone region from infarcted heart at day 3 post-MI or sham	scRNA-seq	A subcluster of fibroblasts was revealed to be specific to injured hearts, highly expressing *Ckap4*, along with well-known myofibroblast markers such as *Postn* and *Cthrc1*. *In vitro* inhibition of *Ckap4* resulted in increased expression of myofibroblasts after TGFβ induction
Molenaar et al. (61)	Mouse	60-min ligation of LAD	Cells from non-infarct regions from the heart at day 1, 3, and 14 post-injury or sham	scRNA-seq	*Mfge8, Calr* and *B2m* were upregulated in CMs at day 1 post-injury, with their cognate receptors expressed in other cell types (dominantly in fibroblasts). B2M induced the expression of myofibroblast markers in fibroblasts *in vitro*
Farbehi et al. (62)	Mouse	Permanent ligation of LAD	All non-CMs and enriched (*Pdgfra*-GFP^{+}) fibroblast lineage cells from ventricles at day 3 and 7 post-MI or sham	scRNA-seq	A novel fibroblast subpopulation expressing *Wif1* was discovered in both sham and post-MI hearts, which presented an anti-WNT/CTGF/TGFβ signature, and interacted with ECs. IF revealed WIF1^{+} fibroblasts in the border zone at post-MI day 3, but not in sham or post-MI day 1 or 7 hearts, indicating the post-transcription and injury-dependent regulation of WIF expression, and its function of inhibiting fibrosis and angiogenesis
Kretzschmar et al. (63)	Mouse	Permanent ligation of LAD; 60-min ligation of LAD	Cells from neonatal heart or adult heart 14 days after MI, I/R, or sham	scRNA-seq	By investigating murine heart 14 days after MI with genetic lineage tracing using *Ki67* knockin models, no cycling CMs were found in infarcted hearts. A subpopulation of proliferative fibroblasts expressing *Fstl1* was identified, which resembled neonatal cardiac fibroblasts. Conditional KO of *Fstl1* in fibroblasts resulted in more rupture
Forte et al. (64)	Mouse	Permanent ligation of LAD	Cardiac interstitial cells at post-MI day 1, 3, 5, 7, 14, and 28 or sham	scRNA-seq	Epicardial-derived injury-response fibroblasts arose immediately after MI and were replaced by myofibroblasts, followed by matrifibrocytes and late-response fibroblasts. Different strains of mice exhibited different post-MI dynamics of fibroblasts, which was related to cardiac rupture
Yokota et al. (65)	Mouse	Permanent ligation of LAD	Non-CMs from heart at post-MI day 7 in wild-type and *Col5a* conditional KO (in fibroblasts) mice	scRNA-seq	Deficiency in type V collagen resulted in a paradoxical increase in post-MI scar tissue with altered mechanical properties of scars and myofibroblast induction
Li et al. (66)	Mouse	Permanent ligation of LAD	Cardiac ECs from post-MI day 7 or sham heart in EC-specific multispectral lineage tracing mice	scRNA-seq	Clonal proliferation of resident ECs occurred in the infarct border zone. Subpopulations of ECs expressing *Plvap* increased after MI. *In vitro* inhibition of Plvap reduced EC proliferation
Tombor et al. (67)	Mouse	Permanent ligation of LAD	Cardiac non-CMs from post-MI day 1, 3, 5, 7, 14, and 28 or sham	scRNA-seq	ECs underwent transient mesenchymal transition on days 3–7 after MI

(Continued)

TABLE 3 | Continued

References	Species	Ischemia model	Target tissues/cells	Modalities	Major findings
Gladka et al. (68)	Mouse	Permanent ligation of LAD	Cells from the infarct and border zone region from infarcted heart at day 3 post-MI or sham	scRNA-seq	scRNA-seq revealed upregulated *Zeb2* expression in injured CMs, which facilitated release of TMSB4 and PTMA. Overexpression of *Zeb2*, *Tmsb4x*, and *Ptma* in CMs enhanced EC migration and proliferation *in vitro*
Kuppe et al. (69)	Human	-	Heart tissues and isolated cells from patients with acute/chronic MI or non-transplanted donor	snRNA-seq, snATAC-seq, spatial barcoding	Spatially distinct CM, fibroblast, and EC subclusters were identified. Trajectory analysis from snRNA-seq and snATAC-seq data revealed an increase in both the expression and transcription factor motif accessibility of *RUNX1*, along with myofibroblast differentiation
Zhang et al. (70)	Mouse	Permanent ligation of LAD	Isolated nuclei from ventricles at post MI day 5 or sham, from triple-transgenic linage tracing mice (*αMHC-MerCreMer*; *RFP^{fl}/GFP*; *Myh6-H2BBFP6xHis*)	snRNA-seq	snRNA-seq detected an increased BFP^{-} proportion in post-MI CMs, within which *Mki67*, *Runx1* and cell cycle genes were expressed in a higher level, pathways for cardiac contraction and rhythm were downregulated, and survival/proliferation-related pathways and pathways for extracellular matrix receptor interaction were upregulated. This population was thought to represent the dedifferentiated CMs post-MI

ATAC-seq, assay for transposase-accessible chromatin sequencing; CITE-seq, cellular indexing of transcriptomes and epitopes by sequencing; CM, cardiomyocyte; EC, endothelial cell; ECM, extracellular matrix; IF, immunofluorescence; I/R, ischemia/reperfusion; KO, knockout; LAD, left anterior descending artery; LN, lymph node; MI, myocardial infarction; RNA-seq, RNA sequencing; scBCR-seq, single-cell B cell receptor sequencing; scRNA-seq, single-cell RNA sequencing; scTCR-seq, single-cell T cell receptor sequencing; Treg, regulatory T cell.

cardiac fibroblasts, which presented with upregulated *Fstl1* (follistatin-like 1, which had been known to be expressed in adult cardiac fibroblasts). Conditional knockout of *Fstl1* in fibroblasts induced less proliferative cells at the infarct and border zones and lead to more cardiac ruptures, indicating the importance of autocrine FSTL1 signaling in preventing rupture (63). Longitudinal scRNA-seq analysis of murine heart interstitial cells from post-MI day 1 to 28 uncovered a dynamic shift in the subpopulations. In particular, epicardial-derived injury-response fibroblasts arose immediately after MI and were replaced by myofibroblasts, followed by matrifibrocytes and late-response fibroblasts. Different strains of mice exhibited different post-MI dynamics of fibroblasts, which was related to the risk of cardiac rupture (64). Yokota et al. demonstrated that a deficiency in type V collagen resulted in a paradoxical increase of scar tissue, where mechanical properties of scars were altered, and myofibroblasts were induced (65).

ECs are involved in angiogenesis after MI and have thus been another target of single-cell studies. Li et al. discovered that clonal proliferation of resident ECs occurred in the infarct border zone, in which *Plvap* was enriched. PLVAP (plasmalemma vesicle associated protein) had been known as an EC-specific membrane protein involved in the formation of transendothelial channels and microvascular permeability. *In vitro* inhibition of *Plvap* reduced EC proliferation, indicating its involvement in neoangiogenesis (66). Li et al. also observed transient endothelial mesenchymal activation, but the extent to which was comparable between post-MI day 7 and sham. By contrast, Tombor et al. demonstrated that ECs underwent transient mesenchymal activation on days 1 to 7 after MI, which returned to homeostasis afterwards, indicating its involvement in facilitating regeneration of the vascular network *via* EC migration and clonal expansion (67). A crosstalk between injured CMs and ECs has also been identified: transcription factor *Zeb2* expression was upregulated in injured CMs, which facilitated the release of TMSB4 and PTMA (prothymosin α). Overexpression and knockdown experiments of *Zeb2*, *Tmsb4x*, and *Ptma* in iPSC-derived CMs clearly demonstrated their involvement in enhancing EC proliferation and migration (68).

Kuppe et al. integrated snRNA-seq, single-nucleus assay for transposase-accessible chromatin sequencing (snATAC-seq), and spatial transcriptomics in human heart tissues from patients with and without a history of MI. Trajectory analysis from snRNA-seq and snATAC-seq data revealed an increase in both the expression and the transcription factor motif accessibility of *RUNX1* in fibroblasts, as well as in their differentiation to myofibroblasts. *In vitro* experiments then validated that *RUNX1* amplified TGFβ signaling and thus mediated myofibroblast differentiation (69). Zhang et al. developed a triple-transgenic mouse model (*αMHC-MerCreMer*; *RFP^{fl}/GFP*; *Myh6-H2BBFP6xHis*, tamoxifen treated) for fate mapping of CMs after MI. They discovered an increase in the rare GFP^{+}BFP^{low} CMs in post-MI hearts compared to sham hearts that presented with a high BrdU^{+} incorporation rate, which were thought to represent dedifferentiated CMs. snRNA-seq also detected an increased BFP^{-} proportion in post-MI CMs, within which *Mki67*, *Runx1*, and cell cycle genes were

expressed in a higher level. They further showed that pathways for cardiac contraction and rhythm were downregulated, and pathways for survival/proliferation and extracellular matrix receptor interaction were upregulated (70).

MI is not a local event but rather a spatiotemporally heterogeneous state of disease involving the entire heart. The conventional approach for increasing the spatial resolution in earlier studies was dissecting infarct hearts into ischemic zone, border zone, and remote zone. To this end, spatial transcriptomics adds a continuous spatial resolution to the transcriptomic information, which is being used in ongoing research into MI.

HEART FAILURE AND MISCELLANEOUS

HF presents with various underlying causes and culminates in cardiac dysfunction and pump insufficiency, with the most often used experimental animal model of HF being transverse aorta constriction (TAC)-induced pressure overload in mice (**Table 4**). See et al. investigated single nuclei from CMs in the mouse pressure overload model, in which CMs accurately segregated into clusters specific to sham or TAC groups. Weighted gene co-expression network analysis (WGCNA) revealed a disease module in which the long non-coding RNAs (lncRNAs) *Gas5* and *Sghrt* were identified as key nodal regulators and were correlated with cell cycle genes and the fetal reprogramming marker *Nppa*, providing early evidence for lncRNA-regulated state alteration in stress-response CMs (72). Nomura et al. investigated CMs at various time points after TAC and revealed that activated p53 signaling in the late stage of hypertrophy facilitated the HF gene program (73). They also analyzed single-cardiomyocyte transcriptomes from patients with HF and confirmed the conservation of pathological gene programs. Ren et al. used the iCELL8 platform for scRNA-seq of both CMs and non-CMs in murine hearts after TAC and revealed that fibroblasts underwent a subtype switching away from protective features at the initial stage of hypertrophy. The study also showed that the macrophages switched toward a pro-inflammatory state and that their considerable interaction with CMs was associated with deterioration of cardiac function (74). With the iCELL8 platform, Wang et al. constructed a comprehensive resource of single-cell transcriptomes of both CMs and non-CMs from normal, failed, and recovered human heart biopsy samples. In addition to revealing the inter- and intra-compartmental CM heterogeneity, a major finding of this study was that ACKR1$^+$ ECs, which had a protective function and possessed the highest counts of interactions with other cell types, decreased in failed hearts. This observation was validated by *in vivo* experiments where injection of ACKR1$^+$ ECs rescued cardiac function after MI by increasing vessel density in both infarct and border regions in mice (10). Zaman et al. demonstrated that cardiac macrophages in the acute and chronic phase after angiotensin II-induced hypertensive stress were enriched in *Igf1*-containing reparative pathways related to wound healing and angiogenesis and that conditional knockout of *Igf1* in tissue-resident macrophages inhibited adaptive CM hypertrophy and led to cardiac dysfunction after long-term hypertensive stress (75). Besides the interaction with CMs, cardiac macrophages have also been reported to activate fibroblasts through paracrine signaling in HF. Ramanujam et al. applied ligand–receptor pair analysis to scRNA-seq data from non-CMs in pressure-overloaded murine hearts by analyzing the transcripts of secreted protein ligands in one cell type and their receptors in another cell type, and predicted that M1 like macrophages interacted with activated fibroblasts *via* secreted proteins such as TGFβ and fibronectin. These interactions were demonstrated to be mediated by macrophage miR-21, depletion of which rescued hypertrophic phenotype after TAC (76). Martini et al., by sequencing cardiac CD45$^+$ cells, showed increases in *Osm*$^+$ pro-inflammatory macrophages, as well as *Pdcd1*$^+$ Tregs, in pressure-overloaded murine hearts when compared with healthy hearts (77).

Single-cell omics approaches have also played important roles in recent studies investigating SARS-CoV-2–associated cardiac injury. ACE2 (the cellular receptor for virus spike protein) is expressed in CMs and pericytes and is upregulated at both mRNA and protein levels in failing hearts, indicating a vulnerability of these cell states to infection (78, 79). By combining phosphoproteomics and snRNA-seq, Mills et al. reported that the inflammatory response in human cardiac organoids was mediated by STAT1 and epigenetic activation, including BRD4 (bromodomain containing 4, a member of bromodomain and extraterminal protein family). They then validated the results in the hearts of SARS-CoV-2–infected K18-hACE2 mice. Bromodomain and extraterminal family inhibitors were identified as drug candidates that could prevent COVID-19–mediated cardiac damage (80).

FUTURE PERSPECTIVES

Single-cell omics technologies are advancing rapidly and are being applied to a wide range of research subjects. Meanwhile, enabled by progress in bioinformatic analysis, the integration of multi-omics data, as well as spatial transcriptomics, is becoming more common and is providing us with more comprehensive information. These technologies offer deeper insights into the molecular pathways involved in heart development, homeostasis, and specifically in cardiovascular diseases and are thus helping to identify therapeutic candidates.

Besides elucidating pathophysiology, single-cell omics can also contribute to precision medicine. The response to treatment in cardiovascular diseases is non-uniform, as in other organ systems, and this might be associated with interindividual differences in not only genetic backgrounds but also cell populations and microenvironment. Single-cell omics enables assessment of cell populations, functional states, and intercellular communications and thus can enhance patient-specific understanding of diseases alongside the use of existing diagnostic tools. Kim et al. exemplified how single-cell analysis can facilitate a personalized therapeutic approach in a patient with drug-induced hypersensitivity syndrome/drug reaction with eosinophilia and systemic symptoms (81). Such an application to cardiovascular diseases is expected in the near future.

TABLE 4 | Summary of studies of heart failure.

References	Species	Model/disease	Target tissues/cells	Modalities	Major findings
See et al. (72)	Mouse/human	TAC/DCM	Nuclei of CMs isolated from LVs of mice 8 weeks after TAC or sham surgery/end-stage DCM patients or control	snRNA-seq	WGCNA revealed a disease module in which lncRNAs *Gas5* and *Sghrt* were identified as core genes and were correlated with *Nppa*, *Nppb*, and *Ccng1*. *NPPA* and *NPPB* were also upregulated in human DCM CMs
Nomura et al. (73)	Mouse/human	TAC/DCM	CMs isolated from LVs of mice after sham surgery or 3 days and 1, 2, 4, and 8 weeks after TAC/DCM patients or normal control	scRNA-seq, ChIP-seq	Early hypertrophy was associated with ERK1/2 and NRF1/2 transcriptional networks independent from p53. Activated p53 signaling in late hypertrophy facilitated the heart failure gene program, which was conserved between humans and mice
Ren et al. (74)	Mouse/human	TAC/DCM or HCM	CMs and non-CMs isolated from LVs of mice after sham or 2, 5, 8, and 11 weeks after TAC/end-stage DCM, HCM patients, and control	scRNA-seq	The dynamics of various cell types in the spectrum of heart failure were revealed. Macrophage switching toward a pro-inflammatory state and their considerable interaction with CMs were associated with deterioration of cardiac function, which could be ameliorated by dapagliflozin
Wang et al. (10)	Human	Ischemic HF or DCM	CMs and non-CMs from biopsy samples of LA and LVs of normal, failed and recovered adult human hearts	scRNA-seq	A comprehensive resource of single-cell transcriptomes of both CMs and non-CMs from normal, failed, and recovered human heart was constructed. ACKR1$^+$ ECs with a protective function decreased in failed hearts. Injection of ACKR1$^+$ ECs rescued cardiac function after MI by increasing vessel density in both infarct and border region in mice
Zaman et al. (75)	Mouse	Angiotensin II	Sorted cardiac macrophages from mice after 4 or 28 days of angiotensin II infusion or sham	scRNA-seq	Cardiac macrophages exposed to hypertensive stress were enriched in reparative pathways encompassing *Igf1*. Either depletion of tissue-resident macrophages or conditional knockout of *Igf1* in tissue-resident macrophages inhibited adaptive CM hypertrophy and thus led to cardiac dysfunction
Ramanujam et al. (76)	Mouse	TAC	Non-CMs from pressure-overloaded hearts 6 days after TAC in wild-type and macrophage-specific miR-21–deficient mice	scRNA-seq	Macrophage-specific deficiency of miR-21 resulted in less fibrosis and attenuated cardiac dysfunction in pressure-overloaded murine hearts. Cardiac macrophages were primary paracrine inducers of fibroblast activation possibly *via* secreted proteins such as TGFβ and fibronectin, which was mediated by macrophage miR-21
Martini et al. (77)	Mouse	TAC	CD45$^+$ cells from LVs of mice 1 or 4 weeks after TAC/sham	scRNA-seq	Upon pressure overload, immune activation occurred across the entire range of immune cell types, leading to upregulation of key subset-specific molecules, such as *Osm* in pro-inflammatory macrophages and *Pdcd1* in regulatory T cells

ChIP-seq, chromatin immunoprecipitation with parallel sequencing; CM, cardiomyocyte; DCM, dilated cardiomyopathy; HCM, hypertrophic cardiomyopathy; HF, heart failure; lncRNA, long non-coding RNA; LA, left atria; LV, left ventricle; scRNA-seq, single-cell RNA sequencing; snRNA-seq, single-nucleus RNA sequencing; TAC, transverse aorta constriction; WGCNA, weighted gene co-expression network analysis.

AUTHOR CONTRIBUTIONS

ZD: writing—original draft preparation. SN: writing—review and editing, supervision, and funding acquisition. Both authors have approved the submitted version and agreed to be accountable for the content of the work.

FUNDING

This work was supported by grants from a Grant-in-Aid for Young Scientists (to SN), the Japan Foundation for Applied Enzymology (to SN), the SENSHIN Medical Research Foundation (to SN), the Kanae Foundation for the Promotion of Medical Science (to SN), MSD Life Science Foundation (to SN), the Tokyo Biomedical Research Foundation (to SN), Astellas Foundation for Research on Metabolic Disorders (to SN), the Novartis Foundation (Japan) for the Promotion of Science (to SN), the Japanese Circulation Society (to SN), the Takeda Science Foundation (to SN), the Cell Science Research Foundation (to SN), a Grant-in-Aid for Scientific Research (B) (to SN), and AMED (JP21ek0210152, JP21gm6210010, JP20bm0704026, JP21ek0210141, JP21ek0109440, JP21ek0109487, JP21gm0810013, JP21km0405209, JP21ek0210118, JP21ek0109406, JP21ek0109543, JP21ek0109569, and JP21tm0724601).

REFERENCES

1. Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature.* (2015) 523:486–90. doi: 10.1038/nature14590
2. Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, et al. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods.* (2014) 11:163–6. doi: 10.1038/nmeth.2772
3. Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, et al. The human cell atlas. *Elife.* (2017) 6:e27041. doi: 10.7554/eLife.27041
4. Tabula Muris Consortium. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature.* (2018) 562:367–72. doi: 10.1038/s41586-018-0590-4
5. Longo SK, Guo MG, Ji AL, Khavari PA. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat Rev Genet.* (2021) 22:627–44. doi: 10.1038/s41576-021-00370-8
6. Paik DT, Cho S, Tian L, Chang HY, Wu JC. Single-cell RNA sequencing in cardiovascular development, disease and medicine. *Nat Rev Cardiol.* (2020) 17:457–73. doi: 10.1038/s41569-020-0359-y
7. Yamada S, Nomura S. Review of single-cell RNA sequencing in the heart. *Int J Mol Sci.* (2020) 21:218345. doi: 10.3390/ijms21218345
8. Kannan S, Miyamoto M, Lin BL, Zhu R, Murphy S, Kass DA, et al. Large particle fluorescence-activated cell sorting enables high-quality single-cell RNA sequencing and functional analysis of adult cardiomyocytes. *Circ Res.* (2019) 125:567–9. doi: 10.1161/CIRCRESAHA.119.315493
9. Katoh M, Nomura S, Yamada S, Aburatani H, Komuro I. Single-cardiomyocyte RNA sequencing to dissect the molecular pathophysiology of the heart. *Methods Mol Biol.* (2021) 2320:183–92. doi: 10.1007/978-1-0716-1484-6_18
10. Wang L, Yu P, Zhou B, Song J, Li Z, Zhang M, et al. Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. *Nat Cell Biol.* (2020) 22:108–19. doi: 10.1038/s41556-019-0446-7
11. Selewa A, Dohn R, Eckart H, Lozano S, Xie B, Gauchat E, et al. Systematic comparison of high-throughput single-cell and single-nucleus transcriptomes during cardiomyocyte differentiation. *Sci Rep.* (2020) 10:1535. doi: 10.1038/s41598-020-58327-6
12. Lacar B, Linker SB, Jaeger BN, Krishnaswami SR, Barron JJ, Kelder MJE, et al. Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat Commun.* (2016) 7:11022. doi: 10.1038/ncomms11022
13. Bakken TE, Hodge RD, Miller JA, Yao Z, Nguyen TN, Aevermann B, et al. Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS ONE.* (2018) 13:e0209648. doi: 10.1371/journal.pone.0209648
14. Bergmann O, Zdunek S, Felker A, Salehpour M, Alkass K, Bernard S, et al. Dynamics of cell generation and turnover in the human heart. *Cell.* (2015) 161:1566–75. doi: 10.1016/j.cell.2015.05.026
15. Jia G, Preussner J, Chen X, Guenther S, Yuan X, Yekelchyk M, et al. Single cell RNA-seq and ATAC-seq analysis of cardiac progenitor cell transition states and lineage settlement. *Nat Commun.* (2018) 9:4877. doi: 10.1038/s41467-018-07307-6
16. Wang Z, Cui M, Shah AM, Tan W, Liu N, Bassel-Duby R, et al. Cell-type-specific gene regulatory networks underlying murine neonatal heart regeneration at single-cell resolution. *Cell Rep.* (2020) 33:108472. doi: 10.1016/j.celrep.2020.108472
17. Rao A, Barkley D, Franca GS, Yanai I. Exploring tissue architecture using spatial transcriptomics. *Nature.* (2021) 596:211–20. doi: 10.1038/s41586-021-03634-9
18. Chan SS, Chan HHW, Kyba M. Heterogeneity of Mesp1+ mesoderm revealed by single-cell RNA-seq. *Biochem Biophys Res Commun.* (2016) 474:469–75. doi: 10.1016/j.bbrc.2016.04.139
19. Lescroart F, Wang X, Lin X, Swedlund B, Gargouri S, Sanchez-Danes A, et al. Defining the earliest step of cardiovascular lineage segregation by single-cell RNA-seq. *Science.* (2018) 359:1177–81. doi: 10.1126/science.aao4174
20. DeLaughter DM, Bick AG, Wakimoto H, McKean D, Gorham JM, Kathiriya IS, et al. Single-cell resolution of temporal gene expression during heart development. *Dev Cell.* (2016) 39:480–90. doi: 10.1016/j.devcel.2016.10.001
21. Li G, Xu A, Sim S, Priest JR, Tian X, Khan T, et al. Transcriptomic profiling maps anatomically patterned subpopulations among single embryonic cardiac cells. *Dev Cell.* (2016) 39:491–507. doi: 10.1016/j.devcel.2016.10.014
22. Xiong H, Luo Y, Yue Y, Zhang J, Ai S, Li X, et al. Single-cell transcriptomics reveals chemotaxis-mediated intraorgan crosstalk during cardiogenesis. *Circ Res.* (2019) 125:398–410. doi: 10.1161/CIRCRESAHA.119.315243
23. Ivanovitch K, Soro-Barrio P, Chakravarty P, Jones RA, Bell DM, Mousavy Gharavy SN, et al. Ventricular, atrial, and outflow tract heart progenitors arise from spatially and molecularly distinct regions of the primitive streak. *PLoS Biol.* (2021) 19:e3001200. doi: 10.1371/journal.pbio.3001200
24. de Soysa TY, Ranade SS, Okawa S, Ravichandran S, Huang Y, Salunga HT, et al. Single-cell analysis of cardiogenesis reveals basis for organ-level developmental defects. *Nature.* (2019) 572:120–4. doi: 10.1038/s41586-019-1414-x
25. Xiao Y, Hill MC, Zhang M, Martin TJ, Morikawa Y, Wang S, et al. Hippo signaling plays an essential role in cell state transitions during cardiac fibroblast development. *Dev Cell.* (2018) 45:153–69 e6. doi: 10.1016/j.devcel.2018.03.019
26. Quijada P, Trembley MA, Misra A, Myers JA, Baker CD, Perez-Hernandez M, et al. Coordination of endothelial cell positioning and fate specification by the epicardium. *Nat Commun.* (2021) 12:4155. doi: 10.1038/s41467-021-24414-z
27. Sereti KI, Nguyen NB, Kamran P, Zhao P, Ranjbarvaziri S, Park S, et al. Analysis of cardiomyocyte clonal expansion during mouse heart development and injury. *Nat Commun.* (2018) 9:754. doi: 10.1038/s41467-018-02891-z
28. Su T, Stanley G, Sinha R, D'Amato G, Das S, Rhee S, et al. Single-cell analysis of early progenitor cells that build coronary arteries. *Nature.* (2018) 559:356–62. doi: 10.1038/s41586-018-0288-7
29. Li G, Tian L, Goodyer W, Kort EJ, Buikema JW, Xu A, et al. Single cell expression analysis reveals anatomical and cell cycle-dependent transcriptional shifts during heart development. *Development.* (2019) 146:173476. doi: 10.1242/dev.173476
30. Wunnemann F, Ta-Shma A, Preuss C, Leclerc S, van Vliet PP, Oneglia A, et al. Loss of ADAMTS19 causes progressive non-syndromic heart valve disease. *Nat Genet.* (2020) 52:40–7. doi: 10.1038/s41588-019-0536-2

31. Hulin A, Hortells L, Gomez-Stallons MV, O'Donnell A, Chetal K, Adam M, et al. Maturation of heart valve cell populations during postnatal remodeling. *Development*. (2019) 146:173047. doi: 10.1242/dev.173047
32. Wang Y, Yao F, Wang L, Li Z, Ren Z, Li D, et al. Single-cell analysis of murine fibroblasts identifies neonatal to adult switching that regulates cardiomyocyte maturation. *Nat Commun.* (2020) 11:2585. doi: 10.1038/s41467-020-16204-w
33. Goodyer WR, Beyersdorf BM, Paik DT, Tian L, Li G, Buikema JW, et al. Transcriptomic profiling of the developing cardiac conduction system at single-cell resolution. *Circ Res.* (2019) 125:379–97. doi: 10.1161/CIRCRESAHA.118.314578
34. Suryawanshi H, Clancy R, Morozov P, Halushka MK, Buyon JP, Tuschl T. Cell atlas of the foetal human heart and implications for autoimmune-mediated congenital heart block. *Cardiovasc Res.* (2020) 116:1446–57. doi: 10.1093/cvr/cvz257
35. Sahara M, Santoro F, Sohlmer J, Zhou C, Witman N, Leung CY, et al. Population and single-cell analysis of human cardiogenesis reveals unique LGR5 ventricular progenitors in embryonic outflow tract. *Dev Cell.* (2019) 48:475–90 e7. doi: 10.1016/j.devcel.2019.01.005
36. Cui Y, Zheng Y, Liu X, Yan L, Fan X, Yong J, et al. Single-cell transcriptome analysis maps the developmental track of the human heart. *Cell Rep.* (2019) 26:1934–50 e5. doi: 10.1016/j.celrep.2019.01.079
37. Lahm H, Jia M, Dressen M, Wirth F, Puluca N, Gilsbach R, et al. Congenital heart disease risk loci identified by genome-wide association study in European patients. *J Clin Invest.* (2021) 131. doi: 10.1172/JCI141837
38. Asp M, Giacomello S, Larsson L, Wu C, Furth D, Qian X, et al. A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. *Cell.* (2019) 179:1647–60 e19. doi: 10.1016/j.cell.2019.11.025
39. Tyser RCV, Ibarra-Soria X, McDole K, Arcot Jayaram S, Godwin J, van den Brand TAH, et al. Characterization of a common progenitor pool of the epicardium and myocardium. *Science.* (2021) 371:abb2986. doi: 10.1126/science.abb2986
40. Mantri M, Scuderi GJ, Abedini-Nassab R, Wang MFZ, McKellar D, Shi H, et al. Spatiotemporal single-cell RNA sequencing of developing chicken hearts identifies interplay between cellular differentiation and morphogenesis. *Nat Commun.* (2021) 12:1771. doi: 10.1038/s41467-021-21892-z
41. Yekelchyk M, Guenther S, Preussner J, Braun T. Mono- and multi-nucleated ventricular cardiomyocytes constitute a transcriptionally homogenous cell population. *Basic Res Cardiol.* (2019) 114:36. doi: 10.1007/s00395-019-0744-z
42. Skelly DA, Squiers GT, McLellan MA, Bolisetty MT, Robson P, Rosenthal NA, et al. Single-cell transcriptional profiling reveals cellular diversity and intercommunication in the mouse heart. *Cell Rep.* (2018) 22:600–10. doi: 10.1016/j.celrep.2017.12.072
43. Chakarov S, Lim HY, Tan L, Lim SY, See P, Lum J, et al. Two distinct interstitial macrophage populations coexist across tissues in specific subtissular niches. *Science.* (2019) 363:aau0964. doi: 10.1126/science.aau0964
44. Hulsmans M, Clauss S, Xiao L, Aguirre AD, King KR, Hanley A, et al. Macrophages facilitate electrical conduction in the heart. *Cell.* (2017) 169:510–22 e20. doi: 10.1016/j.cell.2017.03.050
45. Linscheid N, Logantha S, Poulsen PC, Zhang S, Schrolkamp M, Egerod KL, et al. Quantitative proteomics and single-nucleus transcriptomics of the sinus node elucidates the foundation of cardiac pacemaking. *Nat Commun.* (2019) 10:2889. doi: 10.1038/s41467-019-10709-9
46. Liang D, Xue J, Geng L, Zhou L, Lv B, Zeng Q, et al. Cellular and molecular landscape of mammalian sinoatrial node revealed by single-cell RNA sequencing. *Nat Commun.* (2021) 12:287. doi: 10.1038/s41467-020-20448-x
47. Paik DT, Tian L, Williams IM, Rhee S, Zhang H, Liu C, et al. Single-cell RNA sequencing unveils unique transcriptomic signatures of organ-specific endothelial cells. *Circulation.* (2020) 142:1848–62. doi: 10.1161/CIRCULATIONAHA.119.041433
48. Yucel N, Axsom J, Yang Y, Li L, Rhoades JH, Arany Z. Cardiac endothelial cells maintain open chromatin and expression of cardiomyocyte myofibrillar genes. *Elife.* (2020) 9:e55730. doi: 10.7554/eLife.55730
49. Hu Z, Liu W, Hua X, Chen X, Chang Y, Hu Y, et al. Single-cell transcriptomic atlas of different human cardiac arteries identifies cell types associated with vascular physiology. *Arterioscler Thromb Vasc Biol.* (2021) 41:1408–27. doi: 10.1161/ATVBAHA.120.315373
50. Wolfien M, Galow AM, Muller P, Bartsch M, Brunner RM, Goldammer T, et al. Single-nucleus sequencing of an entire mammalian heart: cell type composition and velocity. *Cells.* (2020) 9:20318. doi: 10.3390/cells 9020318
51. Vidal R, Wagner JUG, Braeuning C, Fischer C, Patrick R, Tombor L, et al. Transcriptional heterogeneity of fibroblasts is a hallmark of the aging heart. *JCI Insight.* (2019) 4:131092. doi: 10.1172/jci.insight.131092
52. Tucker NR, Chaffin M, Fleming SJ, Hall AW, Parsons VA, Bedi KC Jr, et al. Transcriptional and cellular diversity of the human heart. *Circulation.* (2020) 142:466–82. doi: 10.1161/CIRCULATIONAHA.119.045401
53. Litvinukova M, Talavera-Lopez C, Maatz H, Reichart D, Worth CL, Lindberg EL, et al. Cells of the adult human heart. *Nature.* (2020) 588:466–72. doi: 10.1038/s41586-020-2797-4
54. Wang L, Yang Y, Ma H, Xie Y, Xu J, Near D, et al. Single cell dual-omics reveals the transcriptomic and epigenomic diversity of cardiac non-myocytes. *Cardiovasc Res.* (2021). doi: 10.1093/cvr/cvab134
55. Bajpai G, Bredemeyer A, Li W, Zaitsev K, Koenig AL, Lokshina I, et al. Tissue resident CCR2- and CCR2+ cardiac macrophages differentially orchestrate monocyte recruitment and fate specification following myocardial injury. *Circ Res.* (2019) 124:263–78. doi: 10.1161/CIRCRESAHA.118.314028
56. King KR, Aguirre AD, Ye YX, Sun Y, Roh JD, Ng RP Jr, et al. IRF3 and type I interferons fuel a fatal response to myocardial infarction. *Nat Med.* (2017) 23:1481–7. doi: 10.1038/nm.4428
57. Vafadarnejad E, Rizzo G, Krampert L, Arampatzi P, Arias-Loza AP, Nazzal Y, et al. Dynamics of cardiac neutrophil diversity in murine myocardial infarction. *Circ Res.* (2020) 127:e232–49. doi: 10.1161/CIRCRESAHA.120.317200
58. Xia N, Lu Y, Gu M, Li N, Liu M, Jiao J, et al. A unique population of regulatory T cells in heart potentiates cardiac protection from myocardial infarction. *Circulation.* (2020) 142:1956–73. doi: 10.1161/CIRCULATIONAHA.120.046789
59. Heinrichs M, Ashour D, Siegel J, Buchner L, Wedekind G, Heinze KG, et al. The healing myocardium mobilises a distinct B-cell subset through a CXCL13-CXCR5-dependent mechanism. *Cardiovasc Res.* (2021). doi: 10.1093/cvr/cvab181
60. Gladka MM, Molenaar B, de Ruiter H, van der Elst S, Tsui H, Versteeg D, et al. Single-cell sequencing of the healthy and diseased heart reveals cytoskeleton-associated protein 4 as a new modulator of fibroblasts activation. *Circulation.* (2018) 138:166–80. doi: 10.1161/CIRCULATIONAHA.117.030742
61. Molenaar B, Timmer LT, Droog M, Perini I, Versteeg D, Kooijman L, et al. Single-cell transcriptomics following ischemic injury identifies a role for B2M in cardiac repair. *Commun Biol.* (2021) 4:146. doi: 10.1038/s42003-020-01636-3
62. Farbehi N, Patrick R, Dorison A, Xaymardan M, Janbandhu V, Wystub-Lis K, et al. Single-cell expression profiling reveals dynamic flux of cardiac stromal, vascular and immune cells in health and injury. *Elife.* (2019) 8:e43882. doi: 10.7554/eLife.43882
63. Kretzschmar K, Post Y, Bannier-Helaouet M, Mattiotti A, Drost J, Basak O, et al. Profiling proliferative cells and their progeny in damaged murine hearts. *Proc Natl Acad Sci USA.* (2018) 115:E12245–54. doi: 10.1073/pnas.1805 829115
64. Forte E, Skelly DA, Chen M, Daigle S, Morelli KA, Hon O, et al. Dynamic interstitial cell response during myocardial infarction predicts resilience to rupture in genetically diverse mice. *Cell Rep.* (2020) 30:3149–63 e6. doi: 10.1016/j.celrep.2020.02.008
65. Yokota T, McCourt J, Ma F, Ren S, Li S, Kim TH, et al. Type V collagen in scar tissue regulates the size of scar after heart injury. *Cell.* (2020) 182:545–62 e23. doi: 10.1016/j.cell.2020.06.030
66. Li Z, Solomonidis EG, Meloni M, Taylor RS, Duffin R, Dobie R, et al. Single-cell transcriptome analyses reveal novel targets modulating cardiac neovascularization by resident endothelial cells following myocardial infarction. *Eur Heart J.* (2019) 40:2507–20. doi: 10.1093/eurheartj/ehz305
67. Tombor LS, John D, Glaser SF, Luxan G, Forte E, Furtado M, et al. Single cell sequencing reveals endothelial plasticity with transient mesenchymal activation after myocardial infarction. *Nat Commun.* (2021) 12:681. doi: 10.1038/s41467-021-20905-1
68. Gladka MM, Kohela A, Molenaar B, Versteeg D, Kooijman L, Monshouwer-Kloots J, et al. Cardiomyocytes stimulate angiogenesis after ischemic injury in a ZEB2-dependent manner. *Nat Commun.* (2021) 12:84. doi: 10.1038/s41467-020-20361-3

69. Kuppe C, Ramirez Flores RO Li Z, Hannani M, Tanevski J, Halder M, et al. Spatial multi-omic map of human myocardial infarction. *bioRxiv*. (2020). doi: 10.1101/2020.12.08.411686
70. Zhang Y, Gago-Lopez N, Li N, Zhang Z, Alver N, Liu Y, et al. Single-cell imaging and transcriptomic analyses of endogenous cardiomyocyte dedifferentiation and cycling. *Cell Discov*. (2019) 5:30. doi: 10.1038/s41421-019-0095-9
71. van Blokland IV, Groot HE, Franke LH, van der Wijst MGP, van der Harst P. Translational insights from single-cell technologies across the cardiovascular disease continuum. *Trends Cardiovasc Med*. (2021). doi: 10.1016/j.tcm.2021.02.009
72. See K, Tan WLW, Lim EH, Tiang Z, Lee LT, Li PYQ, et al. Single cardiomyocyte nuclear transcriptomes reveal a lincRNA-regulated de-differentiation and cell cycle stress-response *in vivo*. *Nat Commun*. (2017) 8:225. doi: 10.1038/s41467-017-00319-8
73. Nomura S, Satoh M, Fujita T, Higo T, Sumida T, Ko T, et al. Cardiomyocyte gene programs encoding morphological and functional signatures in cardiac hypertrophy and failure. *Nat Commun*. (2018) 9:4435. doi: 10.1038/s41467-018-06639-7
74. Ren Z, Yu P, Li D, Li Z, Liao Y, Wang Y, et al. Single-cell reconstruction of progression trajectory reveals intervention principles in pathological cardiac hypertrophy. *Circulation*. (2020) 141:1704–19. doi: 10.1161/CIRCULATIONAHA.119.043053
75. Zaman R, Hamidzada H, Kantores C, Wong A, Dick SA, Wang Y, et al. Selective loss of resident macrophage-derived insulin-like growth factor-1 abolishes adaptive cardiac growth to stress. *Immunity*. (2021). doi: 10.1016/j.immuni.2021.07.006
76. Ramanujam D, Schon AP, Beck C, Vaccarello P, Felician G, Dueck A, et al. MicroRNA-21-dependent macrophage-to-fibroblast signaling determines the cardiac response to pressure overload. *Circulation*. (2021) 143:1513–25. doi: 10.1161/CIRCULATIONAHA.120.050682
77. Martini E, Kunderfranco P, Peano C, Carullo P, Cremonesi M, Schorn T, et al. Single-cell sequencing of mouse heart immune infiltrate in pressure overload-driven heart failure reveals extent of immune activation. *Circulation*. (2019) 140:2089–107. doi: 10.1161/CIRCULATIONAHA.119.041694
78. Chen L, Li X, Chen M, Feng Y, Xiong C. The ACE2 expression in human heart indicates new potential mechanism of heart injury among patients infected with SARS-CoV-2. *Cardiovasc Res*. (2020) 116:1097–100. doi: 10.1093/cvr/cvaa078
79. Delorey TM, Ziegler CGK, Heimberg G, Normand R, Yang Y, Segerstolpe A, et al. COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets. *Nature*. (2021) 595:107–13. doi: 10.1038/s41586-021-03570-8
80. Mills RJ, Humphrey SJ, Fortuna PRJ, Lor M, Foster SR, Quaife-Ryan GA, et al. BET inhibition blocks inflammation-induced cardiac dysfunction and SARS-CoV-2 infection. *Cell*. (2021) 184:2167–82 e22. doi: 10.1016/j.cell.2021.03.026
81. Kim D, Kobayashi T, Voisin B, Jo JH, Sakamoto K, Jin SP, et al. Targeted therapy guided by single-cell transcriptomic analysis in drug-induced hypersensitivity syndrome: a case report. *Nat Med*. (2020) 26:236–43. doi: 10.1038/s41591-019-0733-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Cardiovascular Medicine

CASE REPORT
published: 06 January 2022
doi: 10.3389/fcvm.2021.798985

Case Report: Whole Exome Sequencing Identifies Compound Heterozygous Variants in *TSFM* Gene Causing Juvenile Hypertrophic Cardiomyopathy

Jamie O. Yang[1], *Hapet Shaybekyan*[2,3], *Yan Zhao*[2,3], *Xuedong Kang*[2,3], *Gregory A. Fishbein*[4], *Negar Khanlou*[4], *Juan C. Alejos*[2], *Nancy Halnon*[2], *Gary Satou*[2], *Reshma Biniwale*[2], *Hane Lee*[4,5†], *Glen Van Arsdell*[2], *Stanley F. Nelson*[2,4,5,6], *Marlin Touma*[1,2,3,6,7,8,9*], *the UCLA Clinical Genomics Center*[1] *and the UCLA Congenital Heart Defects-BioCore Faculty*[1]

[1] David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States, [2] Department of Pediatrics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States, [3] Neonatal/Congenital Heart Laboratory, Cardiovascular Research Laboratory, University of California, Los Angeles, Los Angeles, CA, United States, [4] Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States, [5] Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States, [6] Institute for Precision Health, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States, [7] Department of Pediatrics, David Geffen School of Medicine, Children's Discovery and Innovation Institute, University of California, Los Angeles, Los Angeles, CA, United States, [8] The Molecular Biology Institute, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States, [9] Eli and Edythe Broad Stem Cell Research Center, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States

We report a case of hypertrophic cardiomyopathy and lactic acidosis in a 3-year-old female. Cardiac and skeletal muscles biopsies exhibited mitochondrial hyperplasia with decreased complex IV activity. Whole exome sequencing identified compound heterozygous variants, p.Arg333Trp and p.Val119Leu, in *TSFM*, a nuclear gene that encodes a mitochondrial translation elongation factor, resulting in impaired oxidative phosphorylation and juvenile hypertrophic cardiomyopathy.

Keywords: mitochondrial cardiomyopathy, hypertrophic cardiomyopathy, COXPD3, whole exome sequencing, mitochondrial hyperplasia, TSFM

OPEN ACCESS

Edited by:
Kaoru Ito,
Riken Yokohama, Japan

Reviewed by:
Isabelle Thiffault,
Children's Mercy Hospital,
United States
Atsuhito Takeda,
Hokkaido University Hospital, Japan

***Correspondence:**
Marlin Touma
mtouma@mednet.ucla.edu

†Present address:
Hane Lee,
3billion, Inc., Seoul, South Korea

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: *20 October 2021*
Accepted: *08 November 2021*
Published: *06 January 2022*

Citation:
Yang JO, Shaybekyan H, Zhao Y, Kang X, Fishbein GA, Khanlou N, Alejos JC, Halnon N, Satou G, Biniwale R, Lee H, Van Arsdell G, Nelson SF, Touma M, the UCLA Clinical Genomics Center and the UCLA Congenital Heart Defects-BioCore Faculty (2022) Case Report: Whole Exome Sequencing Identifies Compound Heterozygous Variants in TSFM Gene Causing Juvenile Hypertrophic Cardiomyopathy. Front. Cardiovasc. Med. 8:798985. doi: 10.3389/fcvm.2021.798985

INTRODUCTION

Inherited mitochondrial diseases, or diseases leading to a defect in mitochondrial oxidative phosphorylation, are estimated to occur in 1 in 5,000 live births (1). Mitochondrial diseases are uniquely under dual control by the maternally inherited mitochondrial genome and the Mendelian-inherited nuclear genome. There is a total of 37 genes in the mitochondrial genome coding for rRNAs, tRNAs, and subunits of the respiratory chain. In contrast, there are over 900 nuclear gene products that code for mitochondrial component. In adults, the vast majority of identified mitochondrial respiratory chain defects are associated with mitochondrial DNA (2). In children however, pathogenic variants in mitochondrial DNA are estimated to account for <10% of all mitochondrial disorders, with nuclear DNA defects being the main cause of mitochondrial disease (2). With the advent of next generation sequencing

platforms, including whole exome sequencing (WES), nuclear DNA-encoded genes are increasingly recognized to also cause mitochondrial respiratory chain defects. A growing number of disease-causing variants in nuclear genes are being identified including elongation factors (EF-Ts), mitoribosomal proteins, aminoacyl tRNA synthetases, and release factors.

Mitochondrial diseases present heterogenously affecting multiple organ systems with variable clinical severity, but they particularly affect organs with high metabolic demand such as the heart due to the crucial role of mitochondria in energy production. One of the most common cardiac manifestations is cardiomyopathy, occurring in 20–40% of children with mitochondrial diseases (3). Primary mitochondrial cardiomyopathy is defined as abnormal myocardial structure or function caused by genetic defects involving the mitochondrial respiratory chain without the presence of coronary artery disease, hypertension, valve disease, or congenital heart disease. The clinical phenotype of mitochondrial cardiomyopathy is highly variable. In the pediatric population, inborn errors of metabolism account for <10% of pediatric cardiomyopathy, with a large proportion of these cases being primarily due to mitochondrial oxidative phosphorylation defects (4). The most commonly affected nuclear and mitochondrial encoded genes with pathogenic variants causing pediatric mitochondrial cardiomyopathies are systematically reviewed by Enns et al. (2).

Herein, we report a unique case of primary mitochondrial cardiomyopathy in a female toddler who presented with cardiomegaly and advanced left ventricular failure at 3 years of age. Blood lactate, lactate-to-pyruvate ratio, brain type-natriuretic peptide (BNP) and cardiac troponin levels were elevated. Biopsies from cardiac and skeletal muscles exhibited mitochondrial hyperplasia of abnormal morphology and diminished complex IV activity. Targeted sequencing of mitochondrial encoded genes was negative. Trio (proband and parents) WES revealed novel compound heterozygous variants in *TSFM* [NM_001172696.2; OMIM 604723], a nuclear gene encoding a mitochondrial translation elongation factor (EF-T), that resulted in impaired mitochondrial biogenesis and respiratory function, leading to early-onset hypertrophic cardiomyopathy. Highlighted in this report are the unique pathological and molecular features of this extremely rare mitochondrial cardiomyopathy.

History of Presentation

The proband is a previously healthy 3-year-old female of non-consanguineous Caucasian parents, who presented with a 10-day history of cough, rhinorrhea, fever, and diarrhea and a 2-day history of emesis, found to have severe cardiomegaly on X-ray (**Figure 1A**). She had no prior history of shortness of breath, syncope or chest pain. She was born at term without complications after an uneventful pregnancy. Early growth and development were typical (able to walk by 14 months) and she reported no difficulty with running and keeping up with peers. Her older sister and both parents were healthy. The family history was significant for a paternal grandfather with adult-onset coronary artery disease, but otherwise non-contributory. There was no history of congenital disease or sudden infant deaths in the family. Physical examination was largely unremarkable. Three generation-family pedigree is shown in **Figure 1B**.

Abbreviations: WES, Whole exome sequencing; EFTs, Elongation Factor Ts; EFTu, Elongation Factor Tu; PAS, Periodic acid-Schiff stain; MAF, Minor Allele Frequency; COXPD3, Combined oxidative phosphorylation deficiency 3; DCA, Dichloroacetate.

Investigations

Clinical Workup

On admission, a chest X-ray demonstrated profound cardiomegaly (**Figure 1A**). Her BNP, cardiac troponin, and serum lactate were elevated (**Figure 1C**). She was found to have severe dilated cardiomyopathy with concentric left ventricular thickening and an ejection fraction (EF) of 10-15% per Echocardiography (**Figure 1D** and **Supplementary Video**). Her EKG was profoundly abnormal with huge voltages suggesting a storage disease (**Supplementary Figure 1**). Subsequently, a cardiac catheterization was done, which demonstrated elevated filling pressures in both ventricles and increased systemic vascular resistance, but normal caliber coronary arteries, aorta, and renal arteries, effectively ruling out renal artery stenosis and aortic coarctation as causes of the cardiomyopathy. Infectious workup for myocarditis also came back negative. Metabolic workup revealed metabolic acidosis (PH 7.3), mildly elevated blood lactate level of 2.94 nmol/L (norm: 0.56–1.39) with a lactate/pyruvate ratio of 18.375 (pyruvate: 0.16 nmol/L, norm: 0.03–0.2).

Biopsy Findings

An endomyocardial biopsy taken at the time of the catheterization revealed prominent vacuolar cardiomyocytes with finely granular component in hypertrophied myofibers, likely representing mitochondrial hyperplasia. Periodic Acid Schiff (PAS) staining of the cardiac specimen revealed no glycogen storage abnormalities, and no evidence of inflammation or necrosis was detected (**Figures 2A–C**). At the ultrastructural level, electron microscopic images revealed inter-myofibrillar mitochondrial hyperplasia with polymorphic features including pelaconia and megaconia forms. Relative paucity of internal cristae was observed, and linear structures representing mitochondrial DNA were noted, together, raising suspicion for a primary mitochondrial disorder (**Figures 2D,E**).

Two skeletal muscle biopsies from the thigh (quadriceps muscle) were subsequently obtained to assess for global myopathy. The skeletal muscle showed evidence of lipid and glycogen accumulation suggestive of a combined lipid and glycogen storage myopathy and there was T1 myofiber predominance with preferential type 2 myofiber atrophy (**Figures 2F–H**). A primary carnitine deficiency or glycogen storage disorder was suspected. However, plasma free/total acyl-carnitine profile was normal and analysis of exons 2-20 of the glucosidase alpha acid (*GAA*) gene from her blood revealed no mutation known to be associated with Glycogen Storage Disease II/Pompe disease (**Supplementary Table 1**). No mitochondrial abnormality was detected by immunohistochemistry (**Figure 2I**), and electron microscopy analysis of skeletal myotubes was

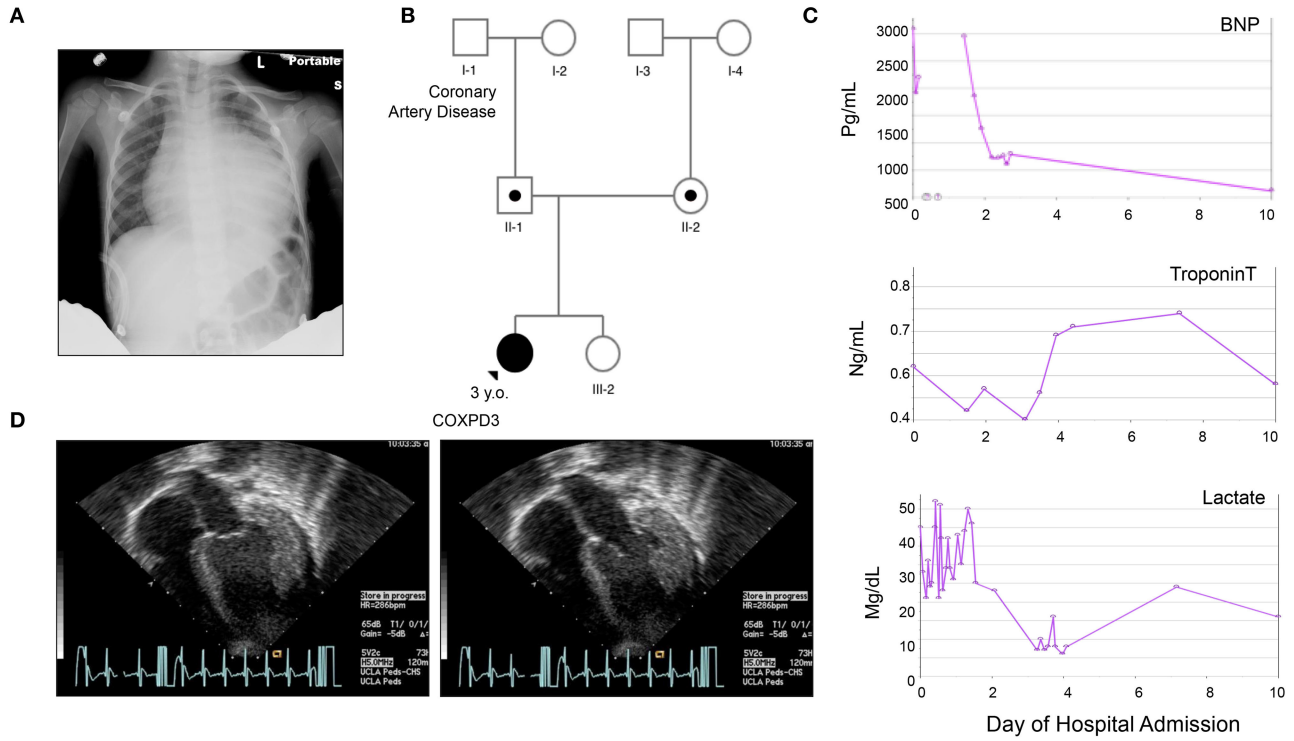


FIGURE 1 | Clinical Presentation of the Proband. **(A)** Chest X ray on admission demonstrating severe cardiomegaly. **(B)** Three-generation family pedigree: III-1, Index case; II-1, father carries the c.355G>C (p.Val119Leu) variant; II-2, mother carries the c.997C>T (p.Arg333Trp) variant. III-2, older healthy sister that has not been tested; No other family history of heart disease besides [I-1], grandfather with coronary heart disease. **(C)** Graphs representing summary of serum brain-natriuretic peptide (BNP), cardiac Troponin T, and Lactate levels during the first 10 h after admission. **(D)** Representative echocardiogram, apical views on admission during systole (a) and diastole (b) demonstrating global left ventricular hypokinesis with estimated EF of 10–15%, significant concentric hypertrophy of the left ventricle, normal coronary arteries take offs, circumferential effusion without chamber compression measuring 14 mm posterior and 8 mm anterior, mild mitral regurgitation. *Please see accompanied **Supplementary Videos** included in **Supplementary Materials**.

essentially normal (**Figures 2J,K**). Hence, biochemical analysis of mitochondrial electron transport complexes was performed on skeletal muscle specimen homogenate and revealed significant reduction in COX activity (16% of the controls) consistent with deficient electron transport complex IV deficiency.

Genetic Workup

The phenotype of mitochondrial IV deficiency can be contributed by disease-causing variants in the mitochondrial genome or in the nuclear mitochondrial genes. To identify the potential genetic defect responsible for the observed cardiac phenotype, a targeted sequence analysis of the mitochondrial DNA coding genes for COX (Complex IV) subunits, *COI*, *COII*, and *COIII* was first performed but no known pathogenic variants were identified. Likewise, the nuclear genes associated with complex IV including *COX10*, *COX15*, *COX6B1*, *SCO1*, *SCO2*, *SURF1*, and *TACO1*, were sequenced, and no significant variations were detected. Sequencing of *FASTKD2*, a nuclear gene that is involved with mitochondrial RNA maturation, was also negative for deleterious variants.

Whole Exome Sequencing

Trio WES was subsequently performed on the proband and both parents. No regions of homozygosity [>5 MB] were identified in the proband's genome suggesting there is no evidence of consanguinity, large deletions or uniparental disomy. Two novel compound heterozygous variants in the *TSFM* gene (NM_001172696.1) were detected in trans: c.997C>T [p.(Arg333Trp)] and c.355G>C [p.(Val119Leu)] (**Figures 3A–D**). Both variants were confirmed in the proband DNA through Sanger sequencing (**Figure 3E**). Pathogenic Variants in *TSFM* are known causal for autosomal recessive combined oxidative phosphorylation deficiency type 3 syndrome (COXPD3) [OMIM: 610505] with heterogenous clinical presentation including cardiomyopathy, encephalopathy, leigh disease, ataxia, or combinations of these disorders. *TSFM* encodes the mitochondrial translation elongation factor Ts. The encoded protein (EF-Ts, 346 aa) is an enzyme that catalyzes the exchange of guanine nucleotides on the translation elongation factor Tu during the elongation step of mitochondrial protein translation, thereby, brings an amino-acylated tRNAs to the ribosomal A site as a ternary complex with guanosine triphosphate (GTP).

The maternally inherited variant c.997C>T is predicted to be a p.(Arg333Trp) change in the EF-Ts subdomain interacting with EF-Tu and has been previously reported as pathogenic (5–7) (dbSNP; rs121909485) and disease-causing for COXPD3

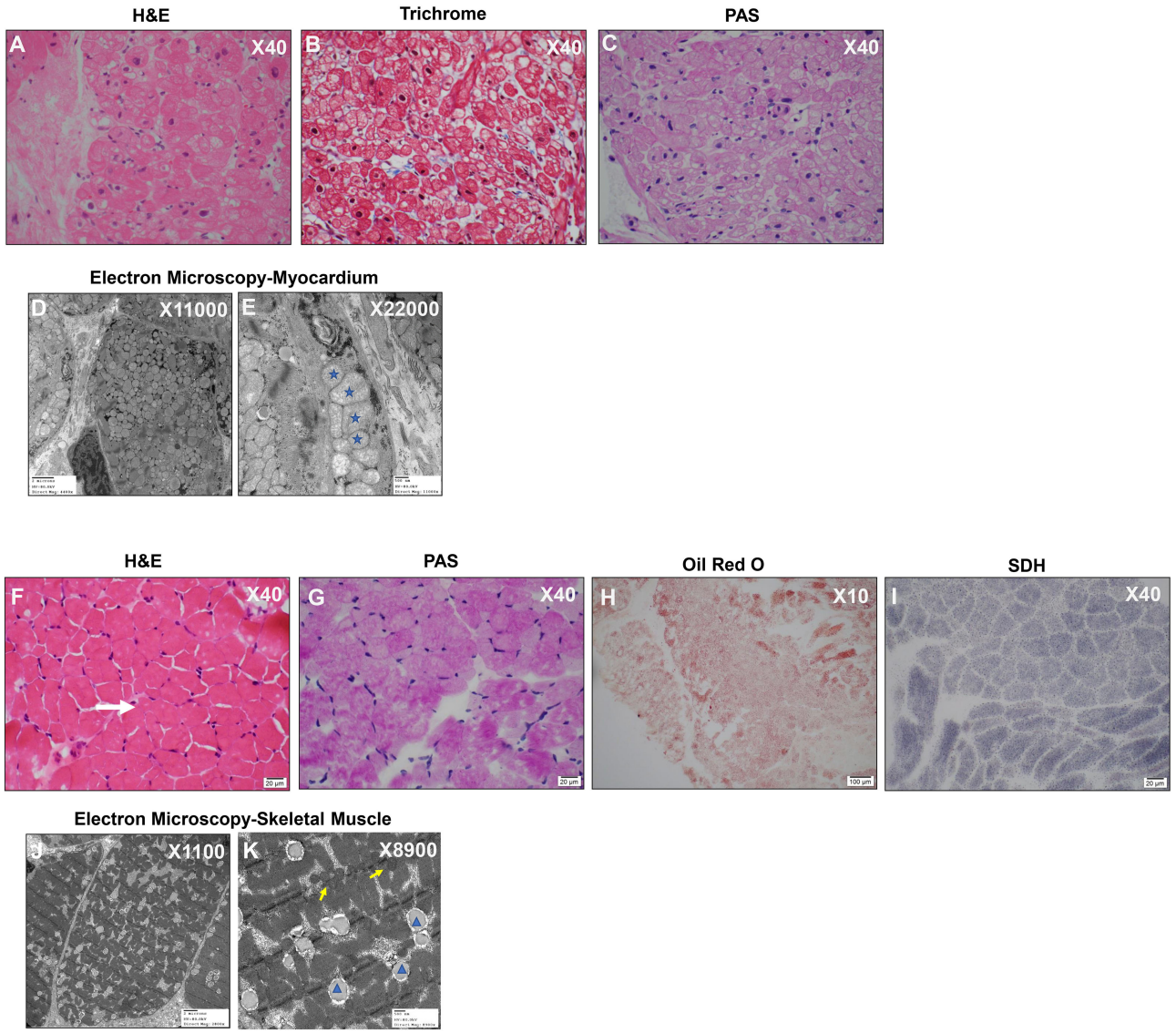


FIGURE 2 | Myocardium and Skeletal Muscle Biopsies Reveal Hyperplastic and Polymorphic Mitochondria and Lipid Accumulation. **(A)** Myocardium: H&E X40 shows diffuse generalized granular vacuolization likely representing mitochondria. **(B)** Myocardium: Mason Trichrome X40 reveals minimal fibrosis. **(C)** Myocardium: Periodic Acid Schiff (PAS) X40 Negative staining indicates absence of abnormal glycogen content. **(D,E)** Myocardium: Electron microscopy– myocardium X11000 **(D)** and X22000 **(E)**: Mitochondrial hyperplasia, pleoconia, and megaconial forms and loss of internal cristae; Stars indicate mitochondria of variable and abnormal size. **(F)** Skeletal Muscle: H&E X40 Moderate fiber size variation with occasional atrophic fibers (white arrow). No evidence of degeneration/regeneration/inflammation. **(G)** Skeletal Muscle: PAS X40 Mild glycogen accumulation. **(H)** Skeletal Muscle: Oil Red O X10 Increased cytoplasmic lipid. **(I)** Skeletal Muscle: SDH X40 complex II mitochondrial respiratory chain enzyme-normal appearance. **(J,K)** Electron microscopy-skeletal muscle direct magnification of X11000 **(J)** and X8900 **(K)**: Mitochondria are normal in size and shape (yellow arrows); there is slight increase in cytoplasmic Glycogen (orange star) lipid (blue arrowhead).

(ClinVar ID: 5379). The paternally inherited c.355G>C is predicted to be a p.(Val119Leu) change. This variant was not previously reported in literature or Clin Var, and was annotated as a variant of uncertain significance (VUS) in Varsome https://varsome.com/variant/hg38/rs863224936?annotation-mode=germline. This variant is not observed in the normal population database, gnomAD, and is highly evolutionally conserved. Substitution of Valine with Leucine is predicted to be deleterious/probably damaging by three *in silico* prediction algorithms (SIFT, PolyPhen2, Condel). Being in trans with a known pathogenic variant in a known autosomal recessive disease that fits the phenotype of the patient well, the variant is classified as likely pathogenic according to published American College of Medical Genetics and Genomics (ACMGG) standards and guidelines and this compound heterozygous genotype is considered diagnostic.

A

Table: Rare and Deleterious Variants in *TSFM* Gene Detected in Trio WES

Gene	Genomic Position	DNA Change	Transcript	Protein Change	Inheritance	Zygosity	MAF	gmomAD Alt Allele Freq (AF)
TSFM	12:58190322	c.997C>T	NM_001172696.2	p.Arg333Trp	Maternal	Heterozygous	0	0
TSFM	12:58180069	c.355G>C	NM_001172696.2	p.Val119Leu	Paternal	Heterozygous	0	0

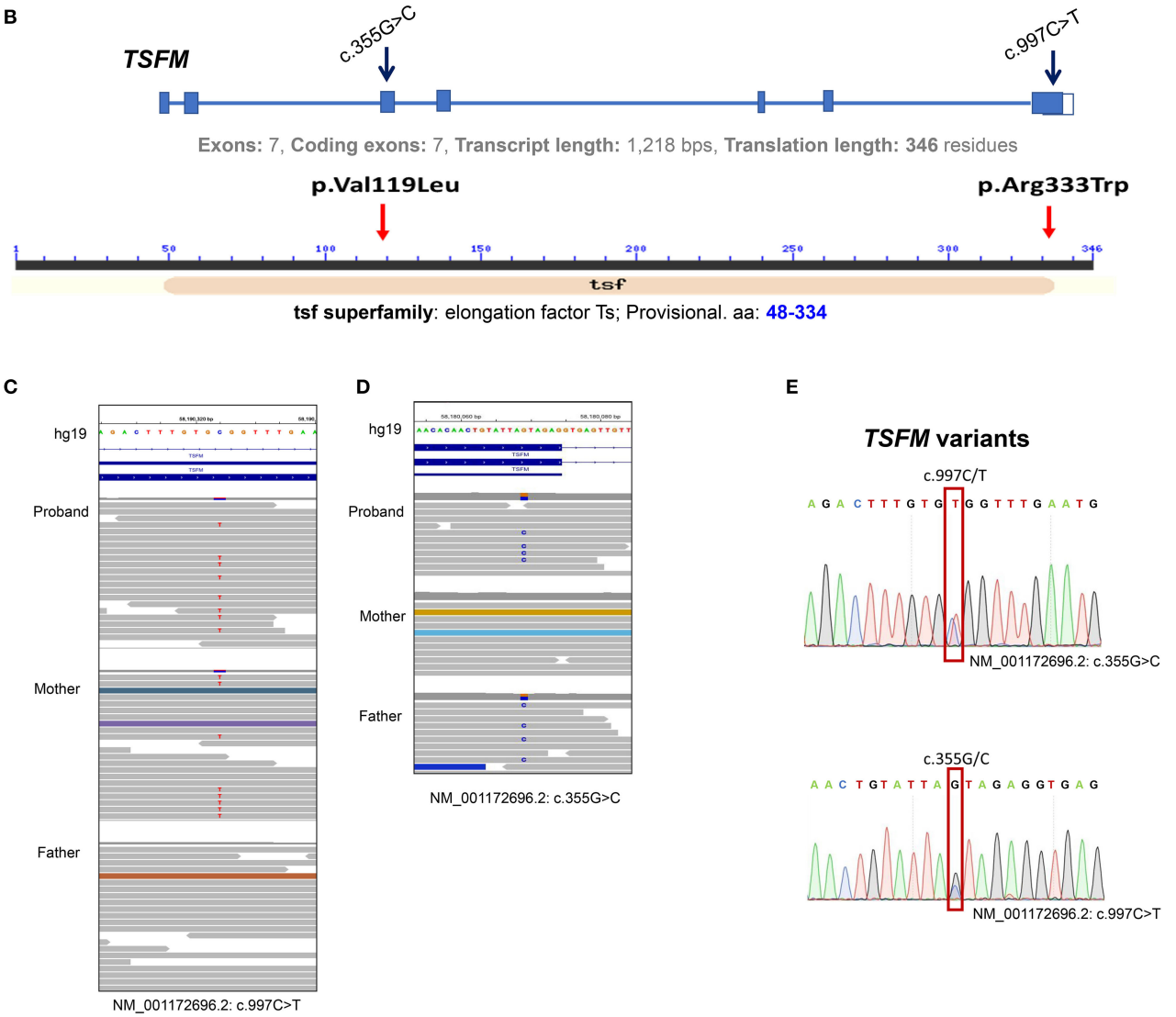


FIGURE 3 | Compound Heterozygous Variants of *TSFM* Revealed by Trio (Proband/Parents) Whole Exome Sequencing (WES). **(A)** Table summary depicts rare compound heterozygous variants in *TSFM* genes detected by WES. **(B)** Schematic representation of *TSFM* gene (upper) and TSFM protein (lower) and known structural motifs. Black arrows indicate exons affected by deleterious mutations. Red arrows indicate the positions of AA substitutions in TSFM protein sequence caused by *TSFM* variants. **(C,D)** Integrated genomic viewer windows depict tow heterozygous variants in *TSFM* in the proband genome, the maternally inherited c.997C>T **(C)** and the paternally inherited c.355G>C **(D)**. **(E)** Sanger sequencing confirms that both *TSFM* variants are in trans configuration in the proband DNA.

Management

On admission, she was treated with continuous intravenous (IV) infusion of milrinone (0.3-1.0 mcg/kg/min) for ionotropic support and afterload reduction, IV furosemide for diuresis, and IV heparin for anticoagulation. The patient was also initially given continuous IV infusion of epinephrine (0.01-0.04 mcg/kg/min) for 3 days, which was subsequently discontinued after lactate downtrend, and later avoided due to concern for its effect in increasing afterload. With remarkable improvement of ejection fraction from 10 to 30% and alleviation of heart failure symptoms, she was discharged with oral therapy of lisinopril, furosemide, chlorothiazide, spironolactone, and

aspirin. A summary of clinical time-course and cardiac function, represented by percentage left ventricle ejection fraction (LVEF%), is shown in **Figure 4**.

At the time of this publication, the patient is 15 years old with mild symptoms of heart failure [New York Heart Association (NYHA) class II-III] and left ventricular ejection fraction >35%, but progressively worsening exercise tolerance with a maximum rate of oxygen (VO2max) of 11 ml/kg. She has a low anaerobic threshold and is not physically active though attends school and participates in usual activities of daily living. She has been hospitalized four times for acute heart failure exacerbations since her initial presentation, which typically resolve within a few days with aggressive diuresis. She was evaluated for cardiac transplantation at the time of presentation. However, she did not qualify because her heart failure symptoms stabilized on oral medications. Her chronic cardiomyopathy therapy includes aspirin, furosemide, losartan, and spironolactone. Her parents and older sister remain asymptomatic.

To date, it appears that the primary clinical manifestations and pathological findings of COXPD3 are restricted to the heart. The skeletal muscles, while exhibiting biochemical features of the disorder, seem to have been spared with no other notable organ involvement. The patient has not manifested signs of muscle weakness, though anaerobic threshold is low. Her neurological, neurosensory and executive functions remain normal.

DISCUSSIONS

In this report, we describe an isolated concentric hypertrophic cardiomyopathy in a 3-year-old proband carrying compound heterozygous variants in *TSFM*, the c.997C>T [p.(Arg333Trp)] variant and the c.355G>C [p.(Val119Leu)] variant, both in highly conserved regions of the TSFM protein. We present the unique histopathological characteristics and the genetic findings of this rare early-onset mitochondrial cardiomyopathy.

TSFM encodes a mitochondrial translation elongation factor Ts (EF-Ts). The encoded protein (346 aa) is a guanine nucleotide exchange factor that plays an essential function during the elongation step of mitochondrial protein translation (8). Pathogenic variants in *TSFM* are known causal for COXPD3 with variable presentations ranging from fatal neonatal-onset to moderate disease courses.

Disease causing variants in *TSFM* are extremely rare. To date, disorders caused by variants in the *TSFM* gene have been reported in only 17 patients and published in 11 reports. These reports are summarized in **Supplementary Table 2**. In terms of cardiomyopathy, variable clinical phenotypes have been correlated with *TSFM* pathogenic variants, including hypertrophic cardiomyopathy, dilated cardiomyopathy, arrhythmogenic cardiomyopathy with fibro-adipose replacement and biventricular failure. These cases were isolated, or associated with varying levels of multisystem involvement (**Supplementary Table 2**).

The p.(Arg333Trp) variant identified in our proband, has been previously detected in seven patients, transmitted in an autosomal recessive manner, manifesting with heterogeneous phenotypes and are all associated with death during infancy. Smeitink et al. reported two unrelated patients with this variant, one with mitochondrial encephalomyopathy, and another patient with concentric hypertrophic cardiomyopathy and muscular hypotonia, both of whom died at 7 weeks (5). Vedrenne et al. reported the same *TSFM* variant in two patients with infantile liver failure (6). Two other patients with this homozygous variant had cardioencephalomyopathy and COXPD3 (7).

The p.(Val119Leu) variant is a likely pathogenic heterozygous variant occurring at a highly conserved residue within the protein sequence, in the setting of a compound *TSFM* variant. Furthermore, this variant is associated with a pathogenic variant in a known autosomal recessive disease, in which the phenotypic, the histopathological findings, and the impaired mitochondrial complex IV activity are consistent with mitochondrial disease caused by pathogenic variants in the *TSFM* gene. Remarkably, no cases with homozygous *TSFM* [p.(Val119Leu)] genotype have been reported to date.

There are several important points to be highlighted in this report. Firstly, our patient's phenotype supports that EF-Ts and mitochondrial translation are critical for cardiac function in early childhood. It is estimated that about 20–30% of children with mitochondrial disease will develop hypertrophic cardiomyopathy (3). Patients with mitochondrial translation defects often present with cardiomyopathy at a young age (3); however, as seen in our patient, if patients can survive during their initial presentation, their cardiomyopathy may potentially remain fairly stable without needing a heart transplant into their adolescent years. This highlights the importance of early diagnosis and aggressive treatment to support cardiac function in the early stages, and the value of developing clinical algorithms for mitochondrial disease-related cardiomyopathy. Understanding the natural history and its correlation with genotypic information may inform treatment decisions and prioritization for heart transplantation.

Secondly, the phenotypes previously described in patients with homozygous genotype for the recurrent p.(Arg333Trp) variant caused infantile deaths, whereas our patient with a compound heterozygous variant survived infancy with a stable cardiomyopathy. Previously described compound heterozygous variants in *TSFM* also demonstrated improved survival with patients living into their adolescent years (8). This suggests that although the number of reported cases is not sufficient to establish a phenotype-genotype correlation, compound heterozygous mutations in *TSFM* appear less severe, as the severe juvenile cardiac phenotype stabilized later. The mechanism of this stabilization remains to be fully revealed, but potentially due to mitochondrial hyperplasia and residual functional EF-Ts allowing for mitochondrial translation to occur, or secondary to compensatory upregulation in EF-Tu that was demonstrated to rescue EF-Ts deficits.

Thirdly, as seen in the previously published reports of *TSFM* variants, phenotypes can range from mitochondrial encephalopathy to optic atrophy, Leigh syndrome, and cardiomyopathy. The reasons that underlie the high variability of phenotypes are unclear at this time (9). Tissue specific genetic modifiers likely exist, but their identification is a challenge given the rarity of these variants (5). This also leads to difficulties with providing accurate prognosis to patients and their families. From the current published *TSFM* cases, hypertrophic cardiomyopathy is the most common manifestation, occurring in 65% of patients.

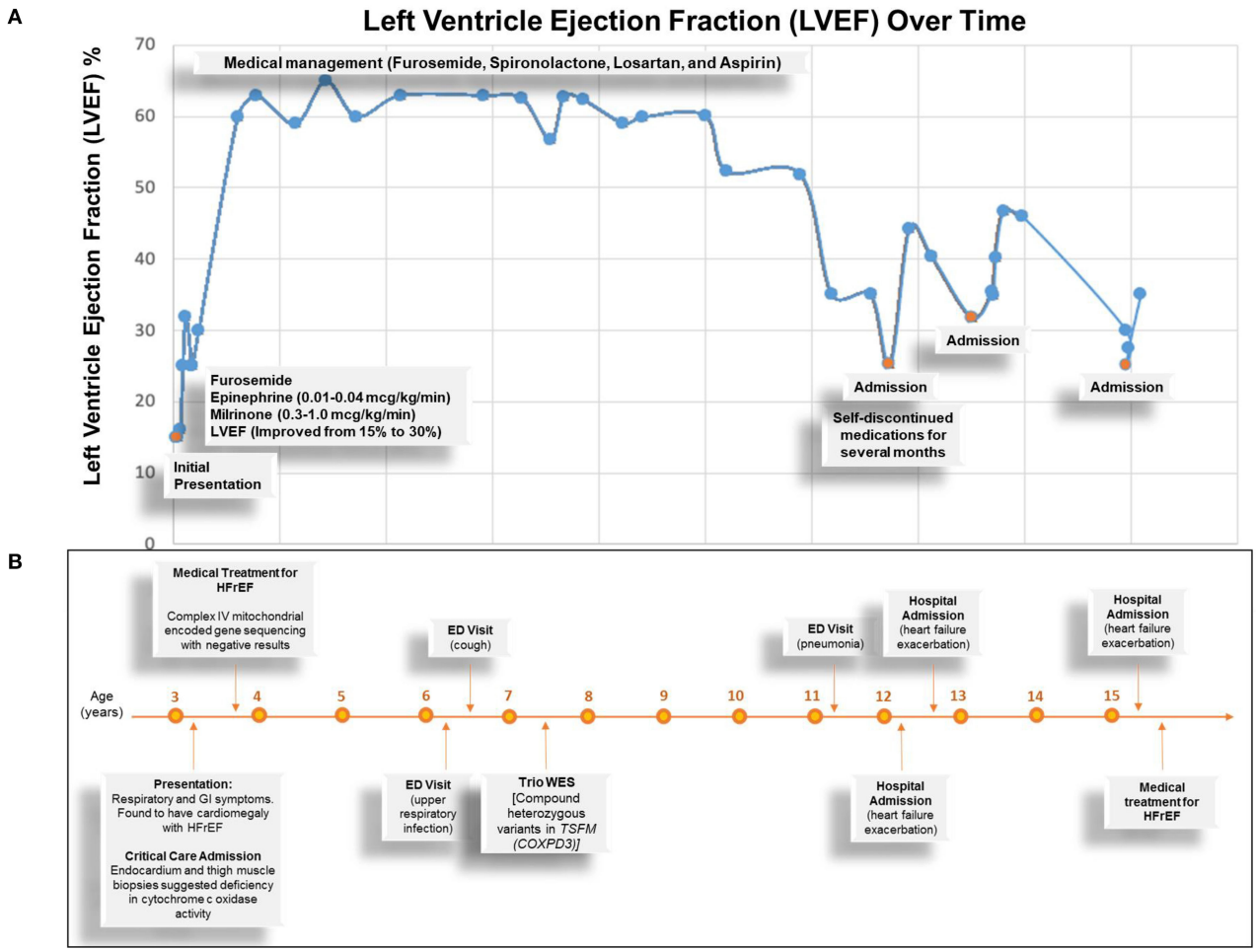


FIGURE 4 | Summary of Clinical Time Course of the Proband. **(A)** Graph depicts left ventricle ejection fraction percentage overtime. Blue dots represent the time of measurements. Orange dots represent the times of admission and related events (texts). **(B)** Summary of clinical time course depicts important clinical events over time. HFrEF, heart failure with reduced ejection fraction; ED, emergency department; COXPD3, combined oxidative phosphorylation deficiency 3.

This is followed by hypotonia (41% of patients) and optic atrophy (29% of patients). Recently, *TSFM* has also been linked to hyperkinetic movement disorders. Continued monitoring of these patients as well as genetic screening and employing WES for early detection of new patients with similar phenotypes will be invaluable in furthering our understanding of mitochondrial oxidation phosphorylation deficiency diseases and informing clinical care.

Lastly, based on the currently reported cases, more than 75% of patients with pathogenic *TSFM* variants have lactic acidosis on presentation. In one *TSFM* case, treatment with sodium dichloroacetate (DCA), an activator of pyruvate dehydrogenase, helped decrease left ventricle mass in a patient with hypertrophic cardiomyopathy (10). Although there is currently no targeted therapy for *TSFM* gene variants, lowering lactate levels with DCA, in combination with L-carnitine and Coenzyme Q supplementation, may potentially lend therapeutic value in the future.

CONCLUSIONS

Pathogenic variants of *TSFM* gene present heterogeneously, and presentation and prognosis can be more severe in homozygous recessive cases. Our report implicates novel compound heterozygous variants, extends the morphologic phenotypes associated with *TSFM* mutations, and confirms that the heart is the main target. EF-Ts and mitochondrial translation apparatus are critical for cardiac function in early childhood, and if the patient survives the critical years, the cardiac status may stabilize. Isolated juvenile cardiomyopathy should prompt consideration of primary mitochondrial dysfunction caused by potential mutations in either mitochondrial or nuclear mitochondrial genes involved in the mitochondrial oxidative phosphorylation process. Whole exome sequencing can improve early detection of causal nuclear mitochondrial genes and facilitate early diagnosis, genetic counseling and discovery of novel treatment strategies.

DATA AVAILABILITY STATEMENT

The data presented in this study must be deposited and made publicly available in an acceptable repository, prior to publication. The study was registered in dbGaP under Novel Gene-Environment Regulatory Circuit in Chamber-Specific Growth of Perinatal Heart, Study ID: 45333. The stable dbGaP accession for this study is phs002725.v1.p1.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of California Los Angeles-Institutional Review Board. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

JY collected the data, participated in data analysis, and drafted the manuscript. HS, XK, and YZ participated in data analysis and manuscript writing. GF and NK performed pathological studies. NH, GS, and JA participated in patient enrollment and provided clinical insights. RB and GVA provided tissue specimens. HL and SN participated in genetic studies. MT conceived and designed the project, analyzed most of the data, obtained funding, and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from NIH/NHLBI 1R01 HL153853-01 (MT), the Department of Defense-Congressionally Directed Medical Research Programs W81XWH-18-1-0164 (MT), and the UCLA Center for Clinical Translational Science Institute Research Grant NIH/UL1TR000124 (MT).

ACKNOWLEDGMENTS

We acknowledge the support of the UCLA Children's Discovery and Innovation Institute, California Center for Rare Disease (CCRD) at the UCLA Institute for Precision Health, the Clinical Genomics Center (Wayne Grody, Hane Lee, Joshua Deignan, Stanley F. Nelson) and the UCLA Congenital Heart Defects-BioCore Faculty (Marlin Touma, Nancy Halnon, Glen Van Arsdell, Mark Sklansky, Juan C. Alejos, Reshma Biniwale, Myke Federman, Leah Reardon, Viviana Fajardo, Meena Garg, John P. Finn, Fabiola Quintero, Wayne Grody, Stanley F. Nelson, and Yibin Wang).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2021.798985/full#supplementary-material

REFERENCES

1. Koenig MK. Presentation and diagnosis of mitochondrial disorders in children. *Pediatr Neurol.* (2008) 38:305–13. doi: 10.1016/j.pediatrneurol.2007.12.001
2. Enns GM. Pediatric mitochondrial diseases and the heart. *Curr Opin Pediatrics.* (2017) 29:541–51. doi: 10.1097/MOP.0000000000000535
3. Holmgren D, Wåhlander H, Eriksson BO, Oldfors A, Holme E, Tulinius M. Cardiomyopathy in children with mitochondrial disease; clinical course and cardiological findings. *Eur Heart J.* (2003). 24:280–8. doi: 10.1016/S0195-668X(02)00387-1
4. Linhart A. *Metabolic Cardiomyopathy.* In: Clinical Cardiogenetics, editors. Cham: Springer International Publishing (2020). p. 151–66. doi: 10.1007/978-3-030-45457-9_9
5. Smeitink JA, Elpeleg O, Antonicka H, Diepstra H, Saada A, Smits P, et al. Distinct clinical phenotypes associated with a mutation in the mitochondrial translation elongation factor EFTs. *Am J Hum Genet.* (2006) 79:869–77. doi: 10.1086/508434
6. Vedrenne V, Galmiche L, Chretien D, de Lonlay P, Munnich A, Rötig A. Mutation in the mitochondrial translation elongation factor EFTs results in severe infantile liver failure. *J Hepatol.* (2012) 56:294–7. doi: 10.1016/j.jhep.2011.06.014
7. Calvo SE, Compton AG, Hershman SG, Lim SC, Lieber DS, Tucker EJ, et al. Molecular diagnosis of infantile mitochondrial disease with targeted next-generation sequencing. *Sci Transl Med.* (2012) 4:118ra10. doi: 10.1126/scitranslmed.3003310
8. Ahola S, Isohanni P, Euro L, Brilhante V, Palotie A, Pihko H, et al. Mitochondrial EFTs defects in juvenile-onset Leigh disease, ataxia, neuropathy, optic atrophy. *Neurology.* (2014) 83:743–51. doi: 10.1212/WNL.0000000000000716
9. Pickett SJ, Grady JP, Ng YS, Gorman GS, Schaefer AM, Wilson IJ, et al. Phenotypic heterogeneity in m.3243A>G mitochondrial disease: the role of nuclear factors. *Ann Clin Transl Neurol.* (2018) 5:333–45. doi: 10.1002/acn3.532
10. Seo GH, Oh A, Kim EN, Lee Y, Park J, Kim T, et al. Identification of extremely rare mitochondrial disorders by whole exome sequencing. *J Hum Genet.* (2019) 64:1117–25. doi: 10.1038/s10038-019-0660-y

Conflict of Interest: HL is now employed by 3billion, Inc. but her contribution to the manuscript was completed while she was employed at UCLA.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 20 January 2022
doi: 10.3389/fcvm.2021.643958

Molecular Mechanisms of Cardiac Injury Associated With Myocardial SARS-CoV-2 Infection

Xianfang Liu[1†], *Longquan Lou*[2†] *and Lei Zhou*[1*]

[1] Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China, [2] Department of General Surgery, The Third People's Hospital of Hangzhou, Hangzhou, China

Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has spread around the world. The development of cardiac injury is a common condition in patients with COVID-19, but the pathogenesis remains unclear. The RNA-Seq dataset (GSE150392) comparing expression profiling of mock human induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) and SARS-CoV-2-infected hiPSC-CMs was obtained from Gene Expression Omnibus (GEO). We identified 1,554 differentially expressed genes (DEGs) based on GSE150392. Gene set enrichment analysis (GSEA), Gene ontology (GO) analysis, and Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis showed that immune-inflammatory responses were activated by SARS-CoV-2, while muscle contraction, cellular respiration, and cell cycle of hiPSC-CMs were inhibited. A total of 15 hub genes were identified according to protein–protein interaction (PPI), among which 11 upregulated genes were mainly involved in cytokine activation related to the excessive inflammatory response. Moreover, we identified potential drugs based on these hub genes. In conclusion, SARS-CoV-2 infection of cardiomyocytes caused a strong defensive response, leading to excessive immune inflammation, cell hypoxia, functional contractility reduction, and apoptosis, ultimately resulting in myocardial injury.

Keywords: COVID-19, cardiac injury, SARS-CoV-2, RNA-Seq, bioinformatics analysis

OPEN ACCESS

Edited by:
Seitaro Nomura,
The University of Tokyo, Japan

Reviewed by:
Nur Izzah Ismail,
The Chinese University of Hong Kong, China
Guangdong Yang,
Laurentian University, Canada

***Correspondence:**
Lei Zhou
zhoulei@njmu.edu.cn

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: *19 December 2020*
Accepted: *29 November 2021*
Published: *20 January 2022*

Citation:
Liu X, Lou L and Zhou L (2022) Molecular Mechanisms of Cardiac Injury Associated With Myocardial SARS-CoV-2 Infection. Front. Cardiovasc. Med. 8:643958. doi: 10.3389/fcvm.2021.643958

INTRODUCTION

Coronavirus disease 2019 (COVID-19) is a viral pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (1). In December 2019, COVID-19 was first reported in Wuhan of China as a severe unknown form of pneumonia. Subsequently, a global pandemic was declared by the WHO in March 2020. At the time of preparing this manuscript, there were more than 5,00,000 fatalities caused by COVID-19.

Acute myocardial damage is the most common described cardiovascular (CV) complication in patients with COVID-19 (2). In multivariable-adjusted models, cardiac injury has been identified as a significant and independent risk factor [hazard ratios (HR) = 4.26] associated with mortality (3). Cytokine storms (caused by acute systemic inflammation) (1, 4), hypoxemia (5), and pathogen-mediated damage (6, 7) were considered as the potential mechanisms responsible for CV complications in COVID-19. However, the exact mechanism of SARS-CoV-2 infection-related myocardial injury remains unclear.

The study of Arun Sharma et al. reported that SARS-CoV-2 directly infected human induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) *in vitro*, inducing contractility depletion, and apoptosis (8). They obtained the expression profiling of three SARS-CoV-2 infected hiPSC-CMs samples and three mock hiPSC-CMs samples by high throughput sequencing and deposit them on the National Center for Biotechnology Information (NCBI, USA) Gene Expression Omnibus (GEO) database (GSE150392). In this study, we performed a detailed bioinformatics analysis of the GSE150392 RNA-seq data to further examine the specific mechanisms of myocardial damage caused by SARS-CoV-2.

MATERIALS AND METHODS

Data Sources

We obtained the RNA-seq data (GSE150392) from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus GEO database. The GSE150392 dataset consists of six samples divided into two groups. The Cov group included three hiPSC-CMs (human-induced pluripotent stem cell-derived cardiomyocytes) samples pretreated with SARS-COV-2 for 72 h (Cov-1-3: GSM4548303-5), and the Mock group included three hiPSC-CMs samples with a mock treatment without virus (Mock1-3: GSM4548306-8).

Identification of Differentially Expressed Genes

We used R (version 4.0.1, https://www.R-project.org/) package "DESeq2" (version 1.28.1) to determine differentially expressed genes (DEGs) (9) (*P*adj <0.01), | (log2FoldChange | > 2) between SARS-CoV-2 infected hiPSC-CMs (Cov group) and Mock hiPSC-CMs (Mock group). The R package "ggplot2" was used to graph the Gene expression values boxplot and minus-vs.-add (MA) plot. *P*adj is the *p*-value using the Benjamini-Hochberg procedure correction, i.e., false discovery rate (FDR).

Gene Set Enrichment Analysis of All Detected Genes

Gene set enrichment analysis (GSEA, a joint project of UC San Diego and Broad Institute, USA) was performed using the GSEA software (version 4.0.3) implementation in our study for identifying potential hallmarks of SARS-CoV-2 infected hiPSC-CMs (10). The annotated gene sets of "h.all.v7.1.symbols.gmt" were adopted from The Molecular Signatures Database (MSigDB, a joint project of UC San Diego and Broad Institute, USA; http://www.broad.mit.edu/gsea/msigdb/index.jsp). The collapse dataset to gene symbols was "False." The permutation type was "gene set." Other parameters were set to default values. Gene sets were considered statistically significant when nominal *p*-value <0.05, FDR was <0.05, and normalized enrichment score (NES) > 1.5.

Enrichment Analyses of DEGs

Gene ontology (GO) analysis (biological processes, molecular function, and cellular component) and Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment were analyzed using R package clusterProfiler (version 3.16) (11). Gene ontology analysis (Immune System Process), KEGG pathways functionally grouped networks, and ClinVar human diseases analysis using the CluGO (version 2.5.7) (12) and CluePedia (version 1.5.7) (13) apps of Cytoscape Software (version 3.8) (14). After inputting DEGs in CluGO, select an analysis method in "Ontologies/Pathways" with the following settings: Statistical Test Used was "Enrichment/Depletion (Two-sided hypergeometric test)," Correction Method Used was "Benjamini-Hochberg," GO Fusion was "false," and GO Group was "true." The one with the smallest *P*adj value was retained among the results of the same GO group. *P*adj <0.05 was considered as the cut-off criteria.

Protein–Protein Interaction Network

The online tool of a search tool for the retrieval of interacting genes (STRING, Swiss Institute of Bioinformatics, Switzerland; https://string-db.org) (15) was applied to establish a Protein–Protein Interactions (PPIs) of DEGs with the score (median confidence) >0.4. Cytoscape and CytoHubba (version 0.1) (16) were used to visualize the PPI network and identify hub genes (Top 15). GeneMANIA online database (https://genemania.org/) was used to analyze the hub genes (17).

Identifying Drug Candidates

We used the Drug Signatures database DSigDB (University of Colorado at Denver and Health Sciences Center, USA) tool of Enrichr (https://maayanlab.cloud/Enrichr/) to identify drug candidates based on hub genes (18). Drug candidates are ranked by a combined score from highest to lowest.

RESULTS

Severe Acute Respiratory Syndrome Coronavirus 2 Infection Caused a Large Number of Gene Expression Changes in hiPSC-CMs

The boxplot of expression values of all transcripts indicated similar whole transcriptome expression in each sample (**Figure 1A**) and sample to sample distance heatmap revealed a relatively clear distinction between samples in the two groups (**Figure 1B**). Based on the criteria of *P*adj <0.01, and |logFC(fold change)| > 2 (**Figure 1C**), a total of 1,554 DEGs was identified from GSE150392, including 726 upregulated genes and 828 downregulated genes. Sharma et al. (8) performed a similar DEGs analysis in their previous study, which was not significantly different from the results of this study.

Gene Set Enrichment Analysis Showed the SARS-CoV-2 Infection Was Mainly Associated With Immune-Inflammatory Response Activity

To obtain insight into the effect of SARS-CoV-2 on the heart, GSEA was used to map into hallmarks between two groups (Cov vs Mock), 14 significant gene sets were

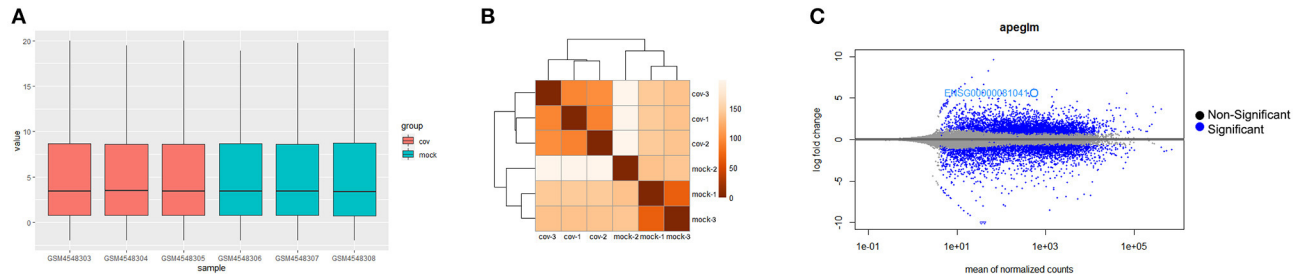


FIGURE 1 | Differential expression of data between two sets of samples. **(A)** Gene expression values boxplot of each sample after normalization using DESeq method; **(B)** Sample to sample distance heatmap of all samples in Cov and Mock group; **(C)** MA plot of genes (The blue points represent genes screened based on |fold change| > 2 and a corrected *P*adj of < 0.01).

TABLE 1 | The most significant gene sets of Cov vs. Mock in gene set enrichment analysis (GSEA).

Gene set follow link to MSigDB	NES	*P*adj
HALLMARK_TNFA_SIGNALING_VIA_NFKB	2.81	0
HALLMARK_INTERFERON_GAMMA_RESPONSE	2.69	0
HALLMARK_INTERFERON_ALPHA_RESPONSE	2.56	0
HALLMARK_INFLAMMATORY_RESPONSE	2.53	0
HALLMARK_IL6_JAK_STAT3_SIGNALING	2.15	0
HALLMARK_IL2_STAT5_SIGNALING	1.85	4.36E-04
HALLMARK_HYPOXIA	1.73	1.02E-02
HALLMARK_P53_PATHWAY	1.73	9.44E-04
HALLMARK_APOPTOSIS	1.57	6.50E-02
HALLMARK_OXIDATIVE_PHOSPHORYLATION	−2.48	0
HALLMARK_E2F_TARGETS	−2.29	0
HALLMARK_G2M_CHECKPOINT	−2.00	0
HALLMARK_MYOGENESIS	−1.70	3.65E-02
HALLMARK_MYC_TARGETS_V1	−1.57	0.01

NES, normalized enrichment score.

shown (**Table 1**; **Figures 2a–o**). Gene set enrichment analysis demonstrated that many inflammation-related gene sets, such as TNFα signaling *via* NF-κB, interferon-γ response, interferon-α response, inflammatory response, IL6-JAK-STAT3 signaling, IL2-STAT5 signaling, hypoxia, P53 pathway, and apoptosis gene sets were enriched by SARS-CoV-2 infection in hiPSC-CMs. Oxidative phosphorylation, E2F targets, G2M checkpoint, myogenesis, and MYC targets v1 gene sets associated with cell cycle, aerobic respiration, and myogenesis were inhibited in hiPSC-CMs with SARS-CoV-2 infection.

Gene Ontology Term Enrichment Analyses and ClinVar Human Diseases Analysis of DEGs

Gene ontology analysis of DEGs was divided into four functional groups, including biological processes (BP), cell composition (CC), molecular function (MF), and Immune System Process (ISP). The top five results of BP, CC, and MF are shown in **Figure 3**.

In the BP group, the upregulated genes were mainly enriched in response to the virus, response to lipopolysaccharide, response to molecule of bacterial origin, leukocyte cell–cell adhesion, and defense response to the virus. The downregulated genes were mainly concentrated in the muscle system process, muscle contraction, oxidative phosphorylation, striated muscle contraction, and respiratory electron transport chain.

In the CC group, the upregulated genes were mainly enriched in receptor complex. The downregulated genes were mainly concentrated in myofibril, contractile fiber, sarcomere, I band, and the inner mitochondrial membrane protein complex.

In the MF group, the upregulated genes were mainly involved in terms of cytokine receptor binding, cytokine activity, DNA-binding transcription activator activity (RNA polymerase II-specific), DNA-binding transcription activator activity, and chemokine activity. The downregulated genes were mainly enriched in actin binding, NADH dehydrogenase activity, NADH dehydrogenase (ubiquinone) activity, NADH dehydrogenase (quinone) activity, and structural constituent of muscle.

The top five ISP results were shown in **Table 2**. Type I interferon signaling pathway, regulation of adaptive immune response, neutrophil migration, regulation of type 2 immune response, and CD4-positive, α-β T cell cytokine production were enriched by upregulated genes. Positive regulation of megakaryocyte differentiation, regulation of megakaryocyte differentiation, and thymus development were enriched by downregulated genes.

ClinVar human diseases analysis indicates that DEGs caused by SARS-CoV-2 infection are significantly involved in terms of myocardial infarction 1, non-syndromic hearing loss and deafness, cardiomyopathy, limb-girdle muscular dystrophy, mitochondrial complex IV deficiency, myofibrillar myopathy, and primary ciliary dyskinesia (**Table 3**).

KEGG Enrichment Analysis of DEGs

The signaling pathways of DEGs were shown in **Figure 4**. The data were imported into Cytoscape to calculate the topological characteristics of the network and determine each node. Cytokine–cytokine receptor interaction, TNF signaling pathway, influenza A, NF-κB signaling pathway, osteoclast differentiation, measles, viral–protein interaction with cytokine and cytokine receptor, IL-17 signaling pathway, Toll-like receptor signaling pathway, and rheumatoid arthritis were significantly enriched by upregulated DEGs (**Figure 4A**). Thermogenesis, Parkinson's disease, oxidative phosphorylation, Huntington disease,

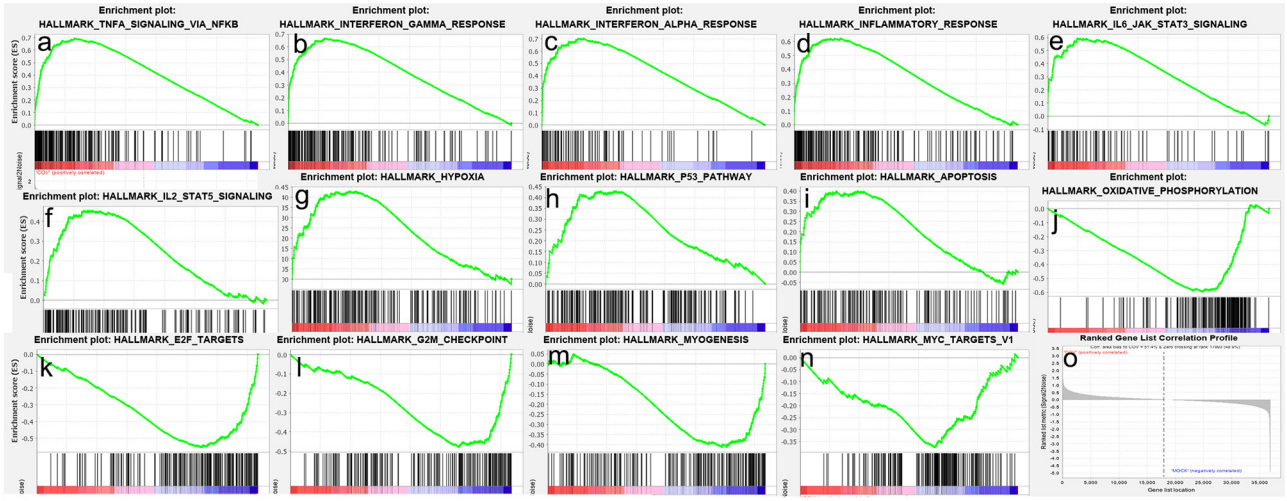


FIGURE 2 | Gene set enrichment analysis (GSEA) of SARS-CoV-2 infection in hiPSC-CMs. **(a–i)** Nine representative gene sets enriched in hiPSC-CMs with SARS-CoV-2 infection; **(j–n)** Five representative gene sets inhibited in hiPSC-CMs with SARS-CoV-2 infection; **(o)** Ranked gene list correlation profile. The red color indicates the "COV" group, the blue color indicates the "MOCK" group, the black line indicates the genes in the GENE SET, and the green line consists of the enrichment fraction of each gene.

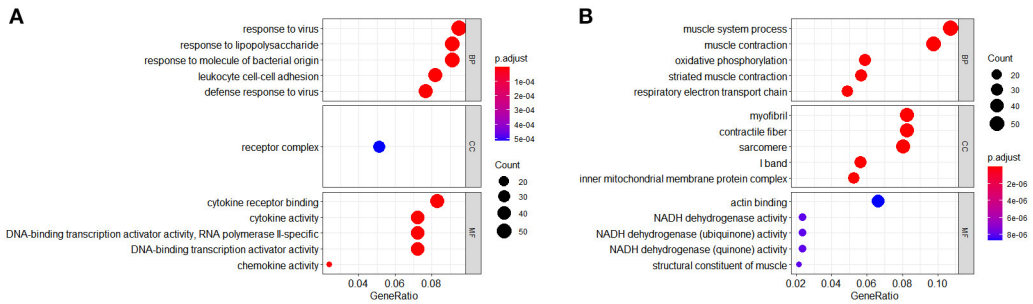


FIGURE 3 | Gene ontology (GO) analysis (BP, CC, and MF) of differentially expressed genes (DEGs). **(A)** GO terms of upregulated DEGs. **(B)** GO terms of downregulated DEGs.

cardiac muscle contraction, non-alcoholic fatty liver disease (NAFLD), retrograde endocannabinoid signaling, adrenergic signaling in cardiomyocytes, dilated cardiomyopathy (DCM), and hypertrophic cardiomyopathy (HCM) were significantly enriched by downregulated DEGs (**Figure 4B**). The top seven *p*-value KEGG pathways with their target genes were shown (**Figure 4C**).

PPI Network Analysis and Hub Genes Recognition

The top 15 genes with the highest interaction degrees were identified, including 11 upregulated genes and four downregulated genes (**Table 4**). Then those genes were used for constructing the PPI network (**Figure 5A**). A total of 11 upregulated hub genes showed the complex PPI network with the co-expression of 66.42%, co-localization of 9.39%, genetic interactions of 8.72%, predicted of 8.6%, shared protein domains of 3.99%, pathway of 1.8%, and physical interactions of 1.09% (**Figure 5C**). Cytokine activity, chemokine activity, inflammatory response, leukocyte chemotaxis, and lipopolysaccharide-mediated signaling pathway were identified as the main function of those genes. Furthermore, four downregulated hub genes showed the complex PPI network with the physical interactions of 67.64%, co-expression of 13.50%, predicted of 6.35%, co-localization of 6.17%, the pathway of 4.35%, genetic interactions of 1.4%, and shared protein domains of 0.59% (**Figure 5B**). Those genes were related to mitosis, nuclear division, metaphase/anaphase transition of the cell cycle, and regulation of ubiquitin–protein ligase activity.

Identification of Candidate Drugs

The study further involved 15 hub genes that were used to screen potential drugs for the treatment of SARS-CoV-2 infection-related myocardial injury. According to the combined

TABLE 2 | Gene ontology (GO) (immune system process) analysis of differentially expressed genes (DEGs).

Term	Description	*P*adj
Upregulated genes		
GO:0060337	Type I interferon signaling pathway	1.02E-08
GO:0002819	Regulation of adaptive immune response	6.53E-05
GO:1990266	Neutrophil migration	2.32E-04
GO:0002828	Regulation of type 2 immune response	2.67E-04
GO:0035743	CD4-positive, alpha-beta T cell cytokine production	8.47E-04
Downregulated genes		
GO:0045654	Positive regulation of megakaryocyte differentiation	1.45E-03
GO:0045652	Regulation of megakaryocyte differentiation	1.88E-03
GO:0048538	Thymus development	4.20E-02

TABLE 3 | ClinVar human diseases analysis of DEGs.

Term	Description	*P*adj
Upregulated genes		
C1832662	Myocardial infarction 1	0.02
CN043648	Non-syndromic hearing loss and deafness	0.03
Downregulated genes		
C0878544	Cardiomyopathy	6.57E-15
C0686353	Limb-girdle muscular dystrophy	2.47E-04
C0268237	Mitochondrial complex IV deficiency	5.69E-04
C2678065	Myofibrillar myopathy	1.06E-03
C0008780	Primary ciliary dyskinesia	3.32E-02

score, the top 10 candidate drugs were generated (**Table 5**). Acetovanillone, antimycin A, proscillaridin, chromium (III) chloride, and hyperforin were identified by 11 upregulated hub genes. Irinotecan hydrochloride, troglitazone, 67526-95-8 (thapsigargin), genistein, and testosterone were identified by four downregulated hub genes.

DISCUSSION

Severe acute respiratory syndrome coronavirus 2, SARS-CoV, and MERS-CoV (Middle East respiratory syndrome coronavirus) are enveloped positive RNA viruses, belong to the Coronaviridae family β genera (19, 20). Severe acute respiratory syndrome coronavirus 2 and SARS-CoV utilize the same receptor angiotensin-converting enzyme 2 (ACE2) for invading human bodies. Those coronaviruses spread widely among people and lead to potentially fatal diseases (21, 22). Coronavirus disease 2019 caused by SARS-CoV-2 has been spreading in 219 countries/areas/territories, with over 50,000,000 confirmed cases (WHO statistics as of November 9, 2020) (23).

The cardiac injury in COVID-19 was generally defined as the elevation of cardiac-specific biomarkers (e.g., troponin concentration above the 99th percentile upper reference limit) (1, 2). It occurs in roughly 8–13% (24, 25) in confirmed cases and as high as 23–44% in severe patients (26–30). Studies have shown that mortality was markedly higher in patients with cardiac injury than in patients without cardiac injury (51.2–59.6 vs. 4.5–8.9%) (3, 31). Besides, patients with underlying cardiovascular disease (CVD) are more likely to develop acute myocardial injury (29, 31). Concurrent occurrence of underlying CVD and myocardial injury causes significantly higher mortality than patients with underlying CVD but without cardiac injury (69.4 vs. 13.3%) (31), suggesting that myocardial injury played a greater role in the fatal outcome of COVID-19 than the presence of underlying CVD itself.

Angiotensin-converting enzyme 2 is an important target for SARS-CoV and SARS-CoV-2 infection, and the heart is one of the organs which express ACE2 (32). During the SARS epidemic, the SARS-CoV viral RNA was detected on autopsied human heart samples (33). Similarly, Guido Tavazzi et al. reported the first case of myocardial localization of SARS-CoV-2 (34). Furthermore, hiPSC-CMs are susceptible to SARS-CoV-2 infection *in vitro* (8). Those studies all suggested that SARS-CoV-2 may directly infect the heart in the human body through ACE2. However, the pathogenesis of SARS-CoV-2 infection-related acute myocardial injury is still unknown. Previous studies suggested that pathogen-mediated direct myocardial injury, acute systemic inflammatory response (1, 4), and low blood oxygen levels (7) may be the most critical causes of myocardial damage in patients with COVID-19. In this study, we described several specific mechanisms of cardiac injury caused by SARS-CoV-2 direct infection.

We used R software and bioinformatics to deeply analyze the RNA-Seq dataset GSE150293, which compared the gene expression between hiPSC-CMs and SARS-CoV-2-infected hiPSC-CMs samples. The results identified 1554 DEGs, including 726 upregulated genes and 828 downregulated genes. In order to reduce the bias, we first performed a GSEA analysis of all detected genes. Gene set enrichment analysis revealed that SARS-CoV-2 infection activated the inflammatory response, and inhibited oxidative phosphorylation and myogenesis, resulting in hypoxia and apoptosis in cardiomyocytes. Subsequently, GO analysis of DEGs revealed specific mechanisms of SARS-CoV-2 impairing myocardial function. Gene ontology analysis of upregulated DEGs suggested several specific mechanisms of inflammatory response activation, including promotion of receptor complexes expression, enhancement of cytokine and chemokine activity, promotion of cytokine binding to receptors, and boosted leukocyte adhesion. Gene ontology analysis of downregulated DEGs showed muscle system process, muscle contraction, oxidative phosphorylation, striated muscle contraction, and respiratory electron transport chain were significantly inhibited by SARS-CoV-2 infection. The expression of myofibril, contractile fiber, sarcomere, I band, and inner mitochondrial membrane protein complex was also suppressed. In the MF group, SARS-CoV-2 inhibits actin binding, NADH dehydrogenase (such as ubiquinone and quinone) activity, and structural constituent of muscle. Kyoto encyclopedia of genes and genomes pathway analysis

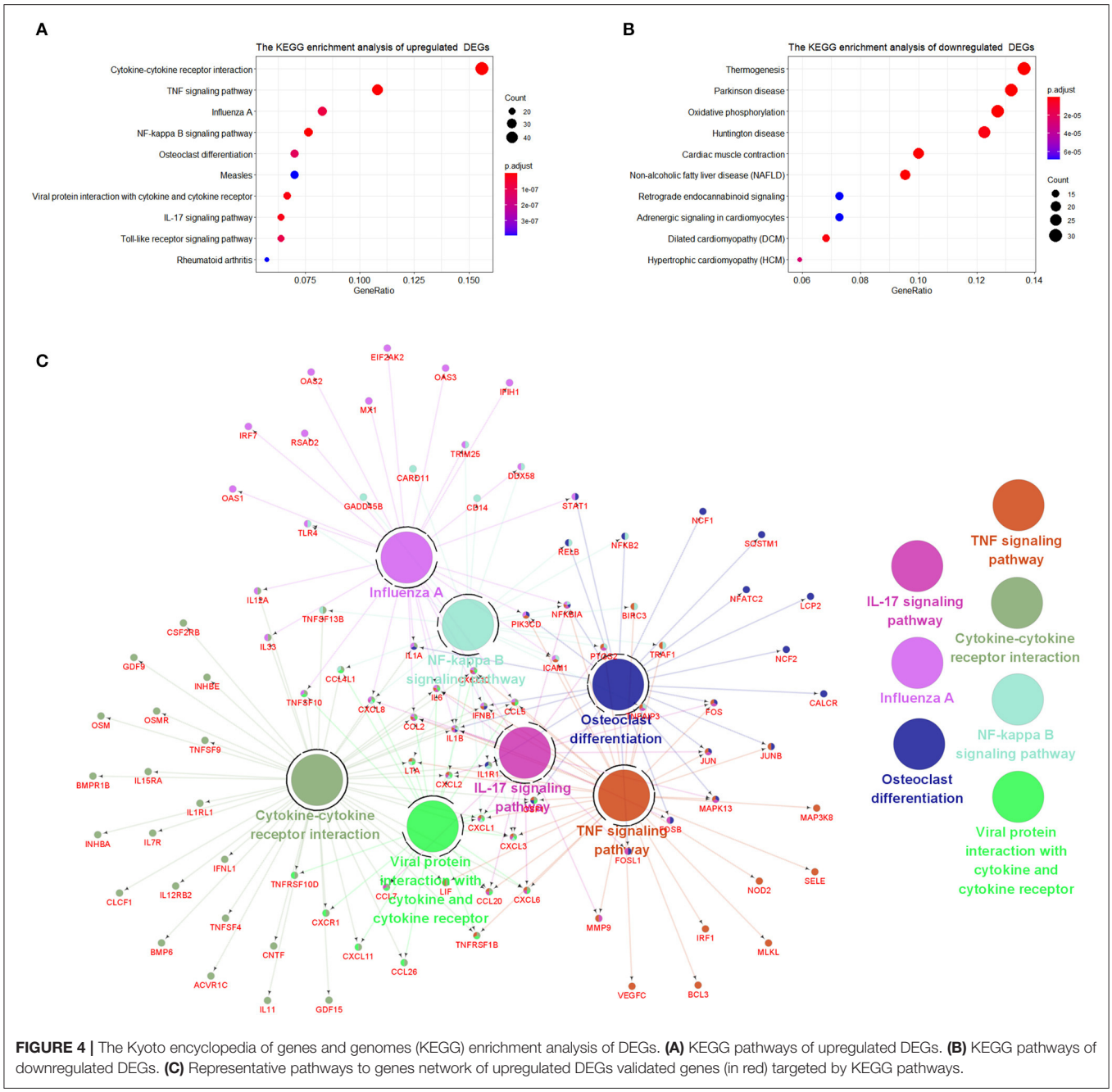


FIGURE 4 | The Kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis of DEGs. **(A)** KEGG pathways of upregulated DEGs. **(B)** KEGG pathways of downregulated DEGs. **(C)** Representative pathways to genes network of upregulated DEGs validated genes (in red) targeted by KEGG pathways.

showed thermogenesis, oxidative phosphorylation, cardiac muscle contraction, retrograde endocannabinoid signaling, and adrenergic signaling in cardiomyocytes were enriched with downregulated DEGs. Those results reveal specific mechanisms by which SARS-CoV-2 infection causes respiration dysfunction and muscle contraction disorders in hiPSC-CMs.

Previous studies have shown that SARS-CoV-2 infection activates multiple immune responses (35, 36), and to clarify the specific immune response processes involved, we performed ISP enrichment analysis. Similar to previous studies, ISP analysis of upregulated DEGs indicated that adaptive and type-2 immune responses were activated by SARS-CoV-2. In addition, neutrophil migration and CD4-positive α-β T cell cytokine production were promoted. Similarly, GSEA and KEGG pathway analysis showed that immune response correlative signal pathways were activated by SARS-CoV-2, including TNFα, IL6-JAK-STAT3, IL2-STAT5, NF-κB, IL17, and Toll-like receptor signaling pathway. Consistent with the inflammatory cytokines detected in the blood of patients with COVID-19 (1, 37), hiPSC-CMs produce and activate various cytokines, including TNFα, chemokines, interleukins

TABLE 4 | Top 15 hub genes with the highest interaction degrees in protein–protein interaction (PPI) network analysis.

Gene ID	Description
Upregulated genes	
IL-6	Interleukin 6
CXCL8	C-X-C motif chemokine ligand 8
TLR4	Toll like receptor 4
STAT1	Signal transducer and activator of transcription 1
IL-1B	Interleukin 1 beta
CXCL10	C-X-C motif chemokine ligand 10
ICAM1	Intercellular adhesion molecule 1
JUN	Jun proto-oncogene, AP-1 transcription factor subunit
CCL5	C-C motif chemokine ligand 5
CCL2	C-C motif chemokine ligand 2
CD44	CD44 molecule
Downregulated genes	
CDK1	Cyclin dependent kinase 1
UBE2C	Ubiquitin conjugating enzyme E2 C
CDC20	Cell division cycle 20
AURKB	Aurora kinase B

TABLE 5 | Drug candidates combined with hub genes.

Drug name	*P*adj	Combined score	Genes
Upregulated hub genes			
Acetovanillone CTD 00002374	6.75E-14	35430.08	CXCL10; JUN; CXCL8; STAT1; CCL5; CCL2; ICAM1
Antimycin A CTD 00005427	3.86E-10	21794.77	JUN; CXCL8; CCL5; CCL2; ICAM1
Proscillaridin CTD 00006639	7.50E-07	16586.61	CXCL8; CCL5; CCL2
Chromium (III) Chloride CTD 00001073	7.50E-07	16586.61	CXCL8; CCL2; ICAM1
Hyperforin CTD 00000051	9.25E-07	14503.67	JUN; CXCL8; ICAM1
Downregulated hub genes			
Irinotecan hydrochloride CTD 00002224	2.16E-04	1109900.66	CDC20; UBE2C; CDK1; AURKB
Troglitazone CTD 00002415	2.33E-04	1061016.69	CDC20; UBE2C; CDK1; AURKB
67526-95-8 CTD 00007263	5.84E-04	928089.07	CDC20; UBE2C; CDK1; AURKB
Genistein CTD 00007324	8.24E-04	837567.36	CDC20; UBE2C; CDK1; AURKB
Testosterone CTD 00006844	8.24E-04	832973.85	CDC20; UBE2C; CDK1; AURKB

(IL), and interferons after SARS-CoV-2 infection. The massive cytokine release reflects an excessive immune defense, which may cause harmful heart damage. These findings complement our understanding of the human immune response mechanism triggered by SARS-CoV-2 (35) and contribute to the treatment of patients with COVID-19. Similar to our study, Omar Pacha et al. suggested IL17 is immunologically plausible as a target to prevent ARDS in COVID-19 (38).

Furthermore, ISP analysis of downregulated DEGs showed positive regulation of megakaryocyte differentiation, and thymus development were repressed by SARS-CoV-2 infection. These mechanisms may be responsible for the significant reduction of platelet count (39) and peripheral blood T cells (37) in patients with severe COVID-19.

After removing confounding results of genetic disorders, ClinVar human diseases analysis showed SARS-CoV-2 infection may be associated with myocardial infarction (40, 41) and cardiomyopathy (42), which are similar to the clinically reported cardiac complications in patients with COVID-19. Cardiac injury often results in chronic heart failure due to the loss of cardiomyocytes (43). Therefore, we speculate that some patients with COVID-19 combined with myocardial damage may have long-term cardiac insufficiency after recovery.

The top genes with the highest degree of interaction in the PPI network are considered to be hub genes. The identification of hub genes may be critical for the diagnosis and treatment of myocardial injury in patients with COVID-19. In this study, functional analysis of 11 upregulated hub genes showed association with cytokine activity, chemokine activity, inflammatory response, leukocyte chemotaxis, and lipopolysaccharide-mediated signaling pathway. These results were highly similar to the GO analysis of all upregulated DEGs, suggesting that these 11 upregulated hub genes play an irreplaceable role in the SARS-CoV-2-induced myocardial injury process.

It is generally accepted that the regeneration of cardiomyocytes was highest in human infants (44, 45) and declines to relative stagnation in adults (46, 47). Restoring the mitotic activity of cardiomyocytes is considered to be the ultimate solution for the treatment of irreversible heart failure (48). The up-regulation of positive cell cycle regulators is one of the endogenous mechanisms for cardiomyocyte proliferation and regeneration (49). Adenoviral transfection of E2F2 expression in mouse cardiomyocytes leads to the proliferation of adult cardiomyocytes (50). Overexpression of G2/M transition promoter cyclin A2 in ischemic porcine hearts promotes recovery after ischemic injury and induced cytokinesis (51). In this study, GSEA showed that cell cycle-related gene sets such as E2F targets, G2/M checkpoint, myogenesis, and MYC targets v1 were inhibited by SARS-CoV-2 infection. The analysis of the PPI network showed that SARS-CoV-2 inhibits the regenerative potency of hiPSC-CMs by inhibiting the metaphase/anaphase transition of the mitotic cycle, suggesting that the downregulated four hub genes (CDK1/UBE2C/CDC20/AURKB) may be important genes for myogenesis. More research is needed to explain the role of SARS-CoV-2 infection-related mitotic pathways modification in cardiomyocytes, which may contribute to cardiac function recovery in patients with COVID-19.

This study identified 10 drug molecules by hub genes, including acetovanillone, antimycin A, proscillaridin chromium

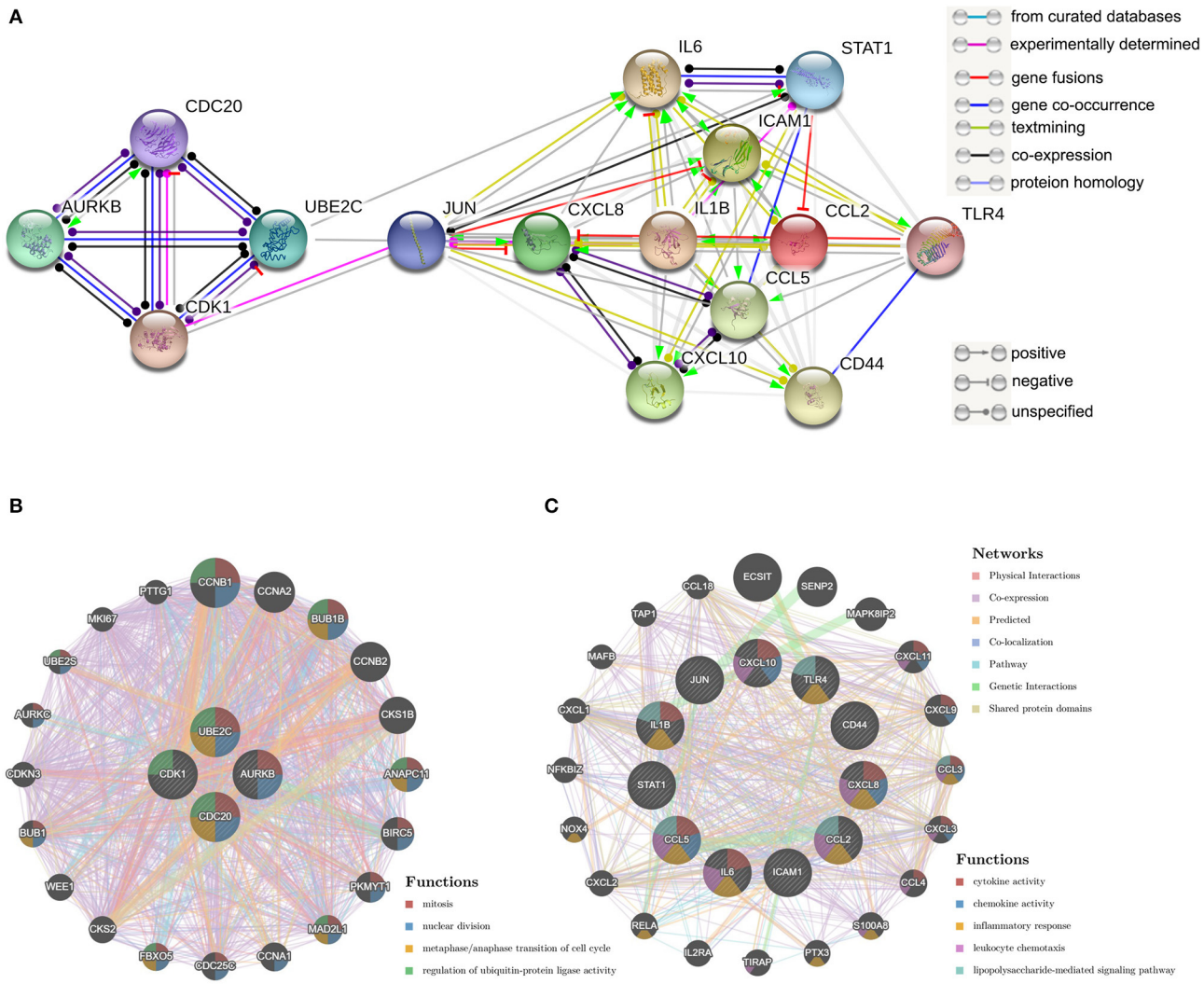


FIGURE 5 | Protein–protein interaction (PPI) network of hub genes. **(A)** The PPI network of the top 15 hub genes created by STRING; **(B)** PP1 networks and function analyses of the four downregulated hub genes; **(C)** PPI networks and function analyses of the 11 upregulated hub genes. Inner circles represent the input genes and outer circles correspond to GeneMANIA proposed hub genes, the size of the circles indicates the correlation with the input genes.

(III) chloride, hyperforin, irinotecan hydrochloride, troglitazone, thapsigargin, genistein, and testosterone. Among these 10 drugs, the cytotoxic drugs chromium (III) chloride and irinotecan hydrochloride should not be considered as candidates. Antimycin A (52), proscillaridin (53, 54), troglitazone (55, 56), thapsigargin (57, 58), genistein (59, 60), and testosterone (61) were considered as potential drugs for COVID-19 in previous studies. Acetovanillone and hyperforin have not been studied for the treatment of COVID-19. The therapeutic effect of these eight candidate drugs on myocardial injury in COVID-19 needs further study.

Overall, our study suggested that SARS-CoV-2 infection can induce a strong immune-inflammatory response, reduce contractility, increase hypoxia, and induce apoptosis in cardiomyocytes. Excessive inflammatory response, respiration dysfunction, and contraction disorders work together to cause cardiac injury. This study supports that direct infection of cardiomyocytes by SARS-CoV-2 is the primary cause of myocardial injury. It should be noted that the occurrence of myocardial injury may be a sign of SARS-CoV-2 entering the circulation, which also explains the acute kidney injury is more common among patients with cardiac injury (3).

The current study has several limitations. The study included only one available data packet of six hiPSC-CMs samples *in vitro*, which could not fully reflect the effect of SARS-CoV-2 on the heart *in vivo*. Only bioinformatics analysis was conducted in this study, further clinical reports or *in vitro* studies are needed to support our conclusions.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE150392.

AUTHOR CONTRIBUTIONS

XL, LL, and LZ designed analyses and drafted the manuscript. XL and LL performed the bioinformatics analysis. LZ reviewed the manuscript and provided funding support. All authors contributed to the final manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (Grant/Award Numbers: 81970723).

REFERENCES

1. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* (2020) 395:497–506. doi: 10.1016/S0140-6736(20)30183-5
2. Bansal M. Cardiovascular disease and COVID-19. *Diabetes Metab Syndr.* (2020) 14:247–50. doi: 10.1016/j.dsx.2020.03.013
3. Shi, S., Qin, M., Shen, B., Cai, Y., Liu, T., Yang, F., et al. (2020). Association of cardiac injury with mortality in hospitalized patients with COVID-19 in Wuhan, China. *JAMA Cardiol.* 5:802–10. doi: 10.1001/jamacardio.2020.0950
4. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet.* (2020) 395:1054–62. doi: 10.1016/S0140-6736(20)30566-3
5. Cameron MJ, Bermejo-Martin JF, Danesh A, Muller MP, Kelvin DJ. Human immunopathogenesis of severe acute respiratory syndrome (SARS). *Virus Res.* (2008) 133:13–9. doi: 10.1016/j.virusres.2007.02.014
6. Li B, Yang J, Zhao F, Zhi L, Wang X, Liu L, et al. Prevalence and impact of cardiovascular metabolic diseases on COVID-19 in China. *Clin Res Cardiol.* (2020) 109:531–8. doi: 10.1007/s00392-020-01626-9
7. Xiong TY, Redwood S, Prendergast B, Chen M. Coronaviruses and the cardiovascular system: acute and long-term implications. *Eur Heart J.* (2020) 41:1798–800. doi: 10.1093/eurheartj/ehaa231
8. Sharma A, Garcia G Jr, Wang Y, Plummer JT, Morizono K, Arumugaswami V, et al. Human iPSC-derived cardiomyocytes are susceptible to SARS-CoV-2 infection. *Cell Rep Med.* (2020) 1:100052. doi: 10.1016/j.xcrm.2020.100052
9. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* (2014) 15:550. doi: 10.1186/s13059-014-0550-8
10. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA.* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102
11. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* (2012) 16:284–7. doi: 10.1089/omi.2011.0118
12. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics.* (2009) 25:1091–3. doi: 10.1093/bioinformatics/btp101
13. Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and *in silico* data. *Bioinformatics.* (2013) 29:661–3. doi: 10.1093/bioinformatics/btt019
14. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* (2003) 13:2498–504. doi: 10.1101/gr.1239303
15. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* (2019) 47:D607–13. doi: 10.1093/nar/gky1131
16. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol.* (2014) 8(Suppl 4):S11. doi: 10.1186/1752-0509-8-S4-S11
17. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* (2010) 38:W214–20. doi: 10.1093/nar/gkq537
18. Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, et al. Gene set knowledge discovery with enrichr. *Curr Protoc.* (2021) 1:e90. doi: 10.1002/cpz1.90
19. Li F. Structure, function, and evolution of coronavirus spike proteins. *Annu Rev Virol.* (2016) 3:237–61. doi: 10.1146/annurev-virology-110615-042301
20. Alsaadi EAJ, Jones IM. Membrane binding proteins of coronaviruses. *Future Virol.* (2019) 14:275–86. doi: 10.2217/fvl-2018-0144
21. Al-Tawfiq JA, Auwaerter PG. Healthcare-associated infections: the hallmark of Middle East respiratory syndrome coronavirus with review of the literature. *J Hosp Infect.* (2019) 101:20–9. doi: 10.1016/j.jhin.2018.05.021
22. Palacios Cruz M, Santos E, Velazquez Cervantes MA, Leon Juarez M. COVID-19, a worldwide public health emergency. *Rev Clin Esp.* (2020). 221:55–61. doi: 10.1016/j.rce.2020.03.001
23. WHO. *WHO Coronavirus Disease (COVID-19) Dashboard.* Geneva: World Health Organization (2020). Available online at: https://covid19.who.int/
24. Lippi G, Plebani M. Laboratory abnormalities in patients with COVID-2019 infection. *Clin Chem Lab Med.* (2020) 58:1131–4. doi: 10.1515/cclm-2020-0198
25. Yang X, Yu Y, Xu J, Shu H, Xia J, Liu H, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med.* (2020) 8:475–81. doi: 10.1016/S2213-2600(20)30079-5
26. Chen T, Wu D, Chen H, Yan W, Yang D, Chen G, et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. *BMJ.* (2020) 368:m1091. doi: 10.1136/bmj.m1091
27. Deng Q, Hu B, Zhang Y, Wang H, Zhou X, Hu W, et al. Suspected myocardial injury in patients with COVID-19: evidence from front-line clinical observation in Wuhan, China. *Int J Cardiol.* (2020) 311:116–21. doi: 10.1016/j.ijcard,0.2020.03.087
28. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med.* (2020) 382:1708–20. doi: 10.1056/NEJMoa2002032
29. Lala A, Johnson KW, Januzzi JL, Russak AJ, Paranjpe I, Richter F, et al. Prevalence and impact of myocardial injury in patients hospitalized with COVID-19 infection. *J Am Coll Cardiol.* (2020). 76:533–46. doi: 10.1016/j.jacc.2020.06.007
30. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA.* (2020) 323:1061–9. doi: 10.1001/jama.2020.1585
31. Guo T, Fan Y, Chen M, Wu X, Zhang L, He T, et al. Cardiovascular implications of fatal outcomes of patients with coronavirus disease 2019 (COVID-19). *JAMA Cardiol.* (2020) 5:811–8. doi: 10.1001/jamacardio.2020.1017
32. South AM, Diz DI, Chappell MC. COVID-19, ACE2, and the cardiovascular consequences. *Am J Physiol Heart Circ Physiol.* (2020) 318:H1084–90. doi: 10.1152/ajpheart.00217.2020

33. Oudit GY, Kassiri Z, Jiang C, Liu PP, Poutanen SM, Penninger JM, et al. SARS-coronavirus modulation of myocardial ACE2 expression and inflammation in patients with SARS. *Eur J Clin Invest.* (2009) 39:618–25. doi: 10.1111/j.1365-2362.2009.02153.x
34. Tavazzi G, Pellegrini C, Maurelli M, Belliato M, Sciutti F, Bottazzi A, et al. Myocardial localization of coronavirus in COVID-19 cardiogenic shock. *Eur J Heart Fail.* (2020) 22:911–5. doi: 10.1002/ejhf.1828
35. Catanzaro M, Fagiani F, Racchi M, Corsini E, Govoni S, Lanni C. Immune response in COVID-19: addressing a pharmacological challenge by targeting pathways triggered by SARS-CoV-2. *Signal Transduct Target Ther.* (2020) 5:84. doi: 10.1038/s41392-020-0191-1
36. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature.* (2020) 579:265–9. doi: 10.1038/s41586-020-2008-3
37. Qin C, Zhou L, Hu Z, Zhang S, Yang S, Tao Y, et al. Dysregulation of immune response in patients with COVID-19 in Wuhan, China. *Clin Infect Dis.* (2020) 71:762–8. doi: 10.1093/cid/ciaa248
38. Pacha O, Sallman MA, Evans SE. COVID-19: a case for inhibiting IL-17? *Nat Rev Immunol.* (2020) 20:345–6. doi: 10.1038/s41577-020-0328-z
39. Lippi G, Plebani M, Henry BM. Thrombocytopenia is associated with severe coronavirus disease 2019 (COVID-19) infections: a meta-analysis. *Clin Chim Acta.* (2020) 506:145–8. doi: 10.1016/j.cca.2020.03.022
40. De Rosa S, Spaccarotella C, Basso C, Calabro MP, Curcio A, Filardi PP, et al. Reduction of hospitalizations for myocardial infarction in Italy in the COVID-19 era. *Eur Heart J.* (2020) 41:2083–8. doi: 10.1093/eurheartj/ehaa409
41. Mahmud E, Dauerman HL, Welt FGP, Messenger JC, Rao SV, Grines C, et al. Management of acute myocardial infarction during the COVID-19 pandemic. *J Am Coll Cardiol.* (2020). 96:336–45. doi: 10.1016/j.jacc.2020.04.039
42. Arentz M, Yim E, Klaff L, Lokhandwala S, Riedo FX, Chong M, et al. Characteristics and outcomes of 21 critically ill patients with COVID-19 in Washington State. *JAMA*. (2020) 323:1612–4. doi: 10.1001/jama.2020.4326
43. Bergmann O, Bhardwaj RD, Bernard S, Zdunek S, Barnabe-Heider F, Walsh S, et al. Evidence for cardiomyocyte renewal in humans. *Science.* (2009) 324:98–102. doi: 10.1126/science.1164680
44. MacMahon HE. Hyperplasia and regeneration of the myocardium in infants and in children. *Am J Pathol.* (1937) 13:845–54.
45. Zerbini C, Weinberg DS, Perez-Atayde AR. DNA ploidy analysis of myocardial hyperplasia. *Hum Pathol.* (1992) 23:1427–30. doi: 10.1016/0046-8177(92)90064-a
46. Mollova M, Bersell K, Walsh S, Savla J, Das LT, Park SY, et al. Cardiomyocyte proliferation contributes to heart growth in young humans. *Proc Natl Acad Sci USA.* (2013) 110:1446–51. doi: 10.1073/pnas.1214608110
47. Bergmann O, Zdunek S, Felker A, Salehpour M, Alkass K, Bernard S, et al. Dynamics of cell generation and turnover in the human heart. *Cell.* (2015) 161:1566–75. doi: 10.1016/j.cell.2015.05.026
48. Laflamme MA, Murry CE. Heart regeneration. *Nature.* (2011) 473:326–35. doi: 10.1038/nature10147
49. Yester JW, Kuhn B. Mechanisms of cardiomyocyte proliferation and differentiation in development and regeneration. *Curr Cardiol Rep.* (2017) 19:13. doi: 10.1007/s11886-017-0826-1
50. Ebelt H, Zhang Y, Kampke A, Xu J, Schlitt A, Buerke M, et al. E2F2 expression induces proliferation of terminally differentiated cardiomyocytes *in vivo*. *Cardiovasc Res.* (2008) 80:219–26. doi: 10.1093/cvr/cvn194
51. Shapiro SD, Ranjan AK, Kawase Y, Cheng RK, Kara, RJ, Bhattacharya R, et al. Cyclin A2 induces cardiac regeneration after myocardial infarction through cytokinesis of adult cardiomyocytes. *Sci Transl Med.* (2014) 6, 224ra227. doi: 10.1126/scitranslmed.3007668
52. Yang CW, Peng TT, Hsu HY, Lee YZ, Wu SH, Lin WH, et al. Repurposing old drugs as antiviral agents for coronaviruses. *Biomed J.* (2020) 43:368–74. doi: 10.1016/j.bj.2020.05.003
53. Aishwarya S, Gunasekaran K, Margret AA. Computational gene expression profiling in the exploration of biomarkers, non-coding functional RNAs and drug perturbagens for COVID-19. *J Biomol Struct Dyn.* (2020) 2020:1–16. doi: 10.1080/07391102.2020.1850360
54. Wang Z, Yang L. Chinese herbal medicine: fighting SARS-CoV-2 infection on all fronts. *J Ethnopharmacol.* (2021) 270:113869. doi: 10.1016/j.jep.2021.113869
55. Chandra A, Gurjar V, Ahmed MZ, Alqahtani AS, Qamar I, Singh N. Exploring potential inhibitor of SARS-CoV2 replicase from FDA approved drugs using insilico drug discovery methods. *J Biomol Struct Dyn.* (2021) 1–8. doi: 10.1080/07391102.2020.1871416
56. Shahjaman M, Rezanur Rahman M, Rabiul Auwul M. A network-based systems biology approach for identification of shared Gene signatures between male and female in COVID-19 datasets. *Inform Med Unlocked.* (2021) 25:100702. doi: 10.1016/j.imu.2021.100702
57. Al-Beltagi S, Preda CA, Goulding LV, James J, Pu J, Skinner P, et al. Thapsigargin is a broad-spectrum inhibitor of major human respiratory viruses: coronavirus, respiratory syncytial virus and influenza A virus. *Viruses*. (2021) 13:234. doi: 10.3390/v13020234
58. Shaban MS, Muller C, Mayr-Buro C, Weiser H, Meier-Soelch J, Albert BV, et al. Multi-level inhibition of coronavirus replication by chemical ER stress. *Nat Commun.* (2021) 12:5536. doi: 10.1038/s41467-021-25551-1
59. Elfiky AA. Natural products may interfere with SARS-CoV-2 attachment to the host cell. *J Biomol Struct Dyn.* (2021) 39:3194–203. doi: 10.1080/07391102.2020.1761881
60. Gour A, Manhas D, Bag S, Gorain B, Nandi U. Flavonoids as potential phytotherapeutics to combat cytokine storm in SARS-CoV-2. *Phytother Res.* (2021) 35:4258–83. doi: 10.1002/ptr.7092
61. Al-Lami RA, Urban RJ, Volpi E, Algburi AMA, Baillargeon J. Sex hormones and novel corona virus infectious disease (COVID-19). *Mayo Clin Proc.* (2020) 95:1710–4. doi: 10.1016/j.mayocp.2020.05.013

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 04 April 2022
doi: 10.3389/fcvm.2022.798033

Deleterious Rare Mutations of *GLI1* Dysregulate Sonic Hedgehog Signaling in Human Congenital Heart Disease

Rui Peng[1,2†], *Binbin Li*[3,4†], *Shuxia Chen*[1,2], *Zhiwen Shi*[1,2], *Liwei Yu*[1,5], *Yunqian Gao*[1,6], *Xueyan Yang*[1,2], *Lei Lu*[1,2*] and *Hongyan Wang*[1,2*]

[1] NHC Key Laboratory of Reproduction Regulation, State Key Laboratory of Genetic Engineering, Obstetrics and Gynecology Hospital, Shanghai Institute of Planned Parenthood Research, Institute of Reproduction and Development, Children's Hospital, Fudan University, Shanghai, China, [2] Shanghai Key Laboratory of Metabolic Remodeling and Health, School of Life Sciences, Institute of Metabolism and Integrative Biology, Fudan University, Shanghai, China, [3] Department of Molecular, Cellular, and Developmental Biology, University of Colorado Boulder, Boulder, CO, United States, [4] Center for Craniofacial Molecular Biology, University of Southern California, Los Angeles, CA, United States, [5] SUNY Downstate Medical Center, Children's Hospital at Downstate, Brooklyn, NY, United States, [6] Department of Laboratory Medicine, Shanghai Children's Medical Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China

OPEN ACCESS

Edited by:
Seitaro Nomura,
The University of Tokyo, Japan

Reviewed by:
Bruno Marino,
Sapienza University of Rome, Italy
Joshi Stephen,
Baylor College of Medicine, United States
Shuchao Pang,
Jining Medical University, China

***Correspondence:**
Lei Lu
lulei@fudan.edu.cn
Hongyan Wang
wanghylab@fudan.edu.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Cardiovascular Genetics and Systems Medicine, a section of the journal Frontiers in Cardiovascular Medicine

Received: 19 October 2021
Accepted: 03 March 2022
Published: 04 April 2022

Citation:
Peng R, Li B, Chen S, Shi Z, Yu L, Gao Y, Yang X, Lu L and Wang H (2022) Deleterious Rare Mutations of GLI1 Dysregulate Sonic Hedgehog Signaling in Human Congenital Heart Disease. Front. Cardiovasc. Med. 9:798033. doi: 10.3389/fcvm.2022.798033

The Glioma-associated oncogene (Gli) family members of zinc finger DNA-binding proteins are core effectors of Sonic hedgehog (SHH) signaling pathway. Studies in model organisms have identified that the *Gli* genes play critical roles during organ development, including the heart, brain, kidneys, *etc*. Deleterious mutations in *GLI* genes have previously been revealed in several human developmental disorders, but few in congenital heart disease (CHD). In this study, the mutations in *GLI1-3* genes were captured by next generation sequencing in human cohorts composed of 412 individuals with CHD and 213 ethnically matched normal controls. A total of 20 patient-specific nonsynonymous rare mutations in coding regions of human *GLI1-3* genes were identified. Functional analyses showed that *GLI1* c.820G>T (p.G274C) is a gain-of-function mutation, while *GLI1* c.878G>A (p.R293H) and c.1442T>A (p.L481X) are loss-of-function mutations. Our findings suggested that deleterious rare mutations in *GLI1* gene broke the balance of the SHH signaling pathway regulation and may constitute a great contribution to human CHD, which shed new light on understanding genetic mechanism of embryo cardiogenesis regulated by SHH signaling.

Keywords: congenital heart disease (CHD), *GLI1*, genetic variant, rare mutation, Sonic hedgehog (SHH) signaling pathway

INTRODUCTION

Congenital heart disease (CHD) is the most common developmental anomaly and the leading non-infectious cause of mortality in newborns (1). In China, the incidence of CHD has been increased from 0.201‰ in 1980–1984 to 4.905‰ in 2015–2019 (2). Ventricular septal defect (VSD) and atrial septal defect (ASD) are the most common congenital heart defects subtypes among the offspring. The etiology of CHD is complicated, and genetic factors play important roles in CHD occurrence

(3–5). However, identification of these factors has been historically slow due to technical limitations and short understanding of signaling pathways regulating embryonic cardiovascular development (5, 6).

In mammals, the Hedgehog (HH) family genes of Sonic hedgehog (SHH), Indian hedgehog (IHH) and Desert hedgehog (DHH) encode evolutionarily conserved ligand proteins initiating pathways crucial for embryogenesis (7). The SHH signaling pathway is transduced by the seven-transmembrane G-protein-coupled receptor (GPCR)-like protein Smoothened (Smo), leading to the activation of Glioma-associated oncogene (Gli) family of transcription factors and downstream target genes transcription (8–10). In vertebrates SHH signaling pathway, *Gli* gene family contains three members of *Gli1, Gli2*, and *Gli3* (11). Gli1 protein acts as a transcriptional activator and provides a positive feedback loop of signaling, whereas Gli2 and Gli3 serve as both transcriptional activators and repressors, depending on specific post-translational modifications and proteolysis processes (12, 13).

SHH signaling pathway has recently been implicated in the specification of early embryonic cardiac progenitor fate. SHH signaling specifies atrial septum from non-septum atrial progenitors (14). Dysregulated SHH signaling pathway involves in numerous human diseases, including birth defects and cancers (15). Various cardiac malformations are observed in $Shh^{-/-}$ mouse embryos (16). The critical roles of *GLI1-3* in embryonic development have been well established (17, 18). *GLI1* participates in differentiation and development of several organs in humans through SHH signaling pathway (19, 20). According to a recent study, *GLI1* nonsynonymous variants were identified among 25 heterotaxy syndrome (HS) patients with CHD (21). Moreover, $Gli2^{-/-}$ and $Gli3^{+/-}$ double mutant mice show a full complement of VACTERL syndrome including cardiac defects (22). These results raise the possibilities that functional rare mutations in *GLI1-3* may also affect human embryo cardiogenesis.

To understand the association between *GL1-3* variants and the risk of CHD in humans, we screened 5′-untranslated region (UTR), 3′-UTR and coding regions of *GLI1-3* genes in a Chinese cohort with 412 cases and 213 matched controls by next generation sequencing. As a result, a total of 20 patient-specific nonsynonymous rare mutations in coding regions of human *GLI1-3* genes were identified. Our *in vitro* and *in vivo* functional analyses showed that *GLI1* c.820G>T (p.G274C) is a gain-of-function mutation, *GLI1* c.878G>A (p.R293H) and c.1442T>A (p.L481X) are loss-of-function mutations. Thus, our data implicate the association between dysregulated SHH signaling pathway and CHD occurrence, and broaden the current knowledge of human embryonic cardiogenesis.

MATERIALS AND METHODS

Human Samples

Sample collection was performed as described previously (23–25). Blood samples from 412 CHD patients (mean age 2.9 ± 2.7 years, 55.6% males) were collected from the Cardiovascular Disease Institute of Jinan Military Command (Jinan, China). The patients were diagnosed by echocardiography, and some were further confirmed surgically. Patients with clinical features of developmental anomalies, positive family history of CHD in a first-degree relative, maternal diabetes mellitus, maternal exposure to known teratogens or any therapeutic drugs during gestation were excluded. The 213 controls (mean age 7.1 ± 3.7 years, 49.8% males) were ethnically and gender-matched, unrelated healthy volunteers recruited from the same geographical area. Sample collection and protocols used in this study were reviewed and approved by the Ethics Committee of the School of Life Sciences, Fudan University and local ethics committees before the start of the present study. All procedures were in accordance with the Declaration of Helsinki. Informed consents were signed by the parents or guardians of the children. Detailed information of the samples was shown in **Table 1**.

TABLE 1 | Information of CHD cases and controls.

Variable	Case (%)	Control (%)
Sequencing group	412	213
Region	Shandong	Shandong
Age: years (mean ± S.D.)	2.9 ± 2.7	7.1 ± 3.7
Male [no. (%)]	229 (55.6)	106 (49.8)
Female [no. (%)]	183 (44.4)	107 (50.2)
CHD classification		
Septation defects [no. (%)]	136 (33.0)	
TOF[a] [no. (%)]	72 (17.5)	
AVSD[b] [no. (%)]	64 (15.5)	
DORV[c] [no. (%)]	39 (9.5)	
PDA[d] [no. (%)]	29 (7.0)	
APVR[e] [no. (%)]	11 (2.7)	
TAPVC[f] [no. (%)]	9 (2.2)	
Others[g] [no. (%)]	52 (12.6)	

[a] *TOF, Tetralogy of Fallot.*
[b] *AVSD, Atrioventricular septal defect.*
[c] *DORV, Double-outlet right ventricle.*
[d] *PDA, Patent ductus arteriosus.*
[e] *APVR, Anomalous pulmonary venous return.*
[f] *TAPVC, Total anomalous pulmonary venous connection.*
[g] *"Others" includes CHDs with other defects such as heterotaxy.*

DNA Sequencing and Data Analysis

Peripheral blood samples were collected. Genomic DNA was extracted and target-sequenced was conducted as described previously (23–25). The genomic structures of human *GLI1-3* genes were determined using NCBI Genebank (mRNA references are NM_001166045, NM_005270, NM_000168, respectively). The 5′ -UTR, 3′ -UTR and coding regions in *GLI1-3* were screened. Identified variants were filtered using the dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP), the 1000 genomes projects (http://www.1000genomes.org/), the Genome Aggregation Database (gnomAD, http://gnomad-sg.org/), and the HuaBiao Database (https://www.biosino.org/wepd) (26). All the patient-specific nonsynonymous mutations were confirmed by Sanger sequencing, the PCR primers were listed in **Supplementary Table S1**.

Plasmids

Human *GLI1-3* ORF without stop codon were amplified by PCR (**Supplementary Table S2**) using cDNA reverse transcripted from human total RNA. *GLI1* and *GLI3* were subcloned into SgfI/MluI restriction enzyme sites. *GLI2* was subcloned into SgfI/NotI restriction enzyme sites of pCMV6-AC-HA (Origene, #PS100004). All plasmid were verified by Sanger sequencing. We then performed the site-directed mutagenesis using QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent, #210518) according to manufacturer's instruction. For *GLI1* c.1442T>A (p.L481X) stop-gain mutation, sequence after premature stop codon (including stop codon) was removed for C-terminal tag fusion. The pGMGLI-Lu firefly luciferase SHH signaling pathway reporter plasmid was obtained from Genomeditech (#GM-021024). The constitutive Renilla luciferase reporter pRL-TK was from Promega (#E6921).

Cell Culture and Transfection

HEK 293T cells (ATCC, #CRL-3216) and HeLa cells (ATCC, #CCL-2) were cultured in high-glucose Dulbecco's Modified Eagle Medium (DMEM, Thermo Fisher Scientific, #11995065) supplemented with 10% FBS (Thermo Fisher Scientific, #A3840001) at 37°C with 5% CO_2. The cells were seeded and maintained overnight to reach ∼80% confluency at the time of transfection by Lipofectamine 2000 Transfection Reagent (Thermo Fisher Scientific, #11668019) and Lipofectamine 3000 Transfection Reagent (Thermo Fisher Scientific, #L3000015) separately according to manufacturer's protocols.

Dual-Luciferase Reporter Assay

Dual-luciferase reporter assay was performed as described previously (24, 27). Briefly, in a 24-well plate well, 300 ng of empty vector or *GLI1* wild-type/mutant constructs, 200 ng of pGMGLI-Lu firefly luciferase reporter plasmid, 10 ng of Renilla luciferase plasmid serving as an internal control were transfected. Reporter assay was performed with the Dual-Luciferase Reporter Assay System (Promega, #E1910) on GloMax Navigator Microplate Luminometer (Promega, #GM2010) 24 h post transfection. At least three independent biological repeats were performed, and data were presented as the mean ± S.D. *P*-values were calculated by Student's *t*-test and considered significant when <0.05.

Western Blotting

C-terminal HA-tagged wild-type or mutant *GLI1* expressing plasmids were transfected into HEK293T cells. Forty-eight hours later, cells were lysed with Cell lysis buffer for Western and IP (Beyotime, #P0013) containing a cocktail of protease inhibitors (ThermoFisher Scientific, #1862209) and heated for 10 min at 100°C. Cell lysates were loaded and separated on a 10% SDS-polyacrylamide gel electrophoresis and transferred onto a PVDF membrane (Merck Millipore, #ISEQ00010). After blocking for 1 h with 5% non-fat milk, the membrane was incubated with mouse HA-tag antibody (Proteintech, #66006-2-Ig) and mouse anti-GAPDH antibody (Proteintech, #66004-1-Ig) at 4°C overnight. Horseradish peroxidase-conjugated anti-mouse IgG was served as secondary antibody (Cell signaling Technology, #7076) for 2 h at room temperature and visualized through the ECL Detection System (Tanon, Shanghai, China) (25). Three independent experiments were performed and representative results were shown.

Electrophoretic Gel Mobility Shift Assay (EMSA)

The probes (F: 5′-AGCTACCTGGGTGGTCTCT-3′, R: 5′-TCGAAGAGACCACCCAGGT-3′) were designed according to the consensus GLI-binding site (5′-TGGGTGGTC-3′) from *PTCH1* promoter region (**Supplementary Table S3**) (28). The DNA-binding affinity of the different GLI proteins was determined by EMSA according to previous description (29). Briefly, nuclear extracts were prepared from HEK 293T cells transfected with empty vector or *GLI1* wild-type/mutation constructs using the NE-PER Nuclear and Cytoplasmic Extraction kit (ThermoFisher Scientific, #78835) and stored at −80°C before use. Protein concentrations were determined by Pierce BCA Protein Assay Kit (ThermoFisher Scientific, #23227). Nuclear proteins were mixed and incubated with indicated probes and subsequently processed using LightShift Chemiluminescent EMSA Kit (ThermoFisher Scientific, #20148) following manufacturer's instruction.

Immunofluorescence

HeLa cells were plated into 35 mm glass bottom dishes and cultured overnight to reach ∼80% confluency. Then cells were transiently transfected with HA-tagged *GLI1* wild-type, p.G274C, p.R293H, or p.L481X constructs. 24 h later, cells were fixed in 4% PFA for 15 min at room temperature, then immunofluorescence was performed using HA antibody (MilliporeSigma, #H6908) as described previously (25), and images were captured on a Zeiss LSM700 confocal microscope under 40x objective lens. Experiment was repeated at least in triplicates and representative result was presented.

Zebrafish Embryo Microinjection

Wild-type AB and Cmlc2-mCherry strain zebrafish (Danio rerio) were maintained under standard conditions (25, 30). All plasmids of the empty vector, *GLI1* wild-type, p.G274C, p.R293H, and p.L481X were extracted by Endo-free Mini Plasmid Kit II (Tiangen Biotech, #DP118), and diluted in nuclease-free water. In this study, each plasmid was injected into >200 zebrafish embryos at 1–2 cell stage. From each plasmid 2–3 nl was injected at a concentration of 40 ng/μl. 72 h post injection, photographs were taken using Leica MZ95 stereo microscope (30) or Olmpus IX83 fluorescence inverted microscope and the percentages of pericardial abnormal embryos were calculated (25). Phenotype distribution differences compared with wild-type group were calculated using $\chi 2$ analysis (25).

RESULTS

Identification of Variants in *GLI1-3* Genes of CHD Patients

A Chinese CHD cohort with ethnically and gender-matched healthy controls was used in this study. The 5′-UTR, 3′-UTR and coding regions of *GLI1-3* genes were sequenced in all patients and controls. CHD is one of severe disorders that impacts mortality and reproductive fitness. A very large negative selection eliminates highly deleterious common mutations from the human population. Thus the existing common variants like single nucleotide polymorphism (SNP) may make less contributions to this disease. Therefore, it is most likely that novel nonsynonymous rare mutations make significant contributions to population prevalence of this defect, especially for sporadic cases. Based on this hypothesis, we sorted out the variants with minor allele frequencies (MAF) < 1% as rare variants.

A total of 94 variants in the *GLI1-3* genes were identified in CHD patients (these data were not shown). 20 nonsynonymous patient-specific rare mutations were identified (**Table 2**, **Figure 1**) and further confirmed by Sanger sequencing (**Supplementary Figure S1**). The effect of rare missense variants on the protein function was predicted *in silico* using SIFT, PolyPhen-2 and CADD (**Table 2**) (31–34). We also evaluated the evolutionary conservativeness of these mutations with Genomic Evolutionary Rate Profiling (GERP, the higher of the GERP score means that the site is more conserved) (**Table 2**) (35), as well as alignment assay across several species (**Supplementary Figure S2**). For *GLI1* gene, we identified one nonsense stop-gain rare mutation and five missense rare mutations. Of which, the stop-gain mutation of c.1442T>A (p.L481X) leads to loss of several crucial domains including the transcriptional activation domain. Mutations of c.820G>T (p.G274C), c.878G>A (p.R293H), c.1776A>T (p.R592S) and c.1924C>T (p.R642S) are predicted as harmful mutations by both SIFT and PolyPhen-2. In addition, mutations of p.G274C and p.R293H locate in the zinc finger domain, which probably interfere with the protein-DNA affinity.

Three *GLI1* Mutations Had Significantly Different Activities Compared to Wild-Type *GLI1*

As GLI1 protein only serves as transcriptional activator whereas GLI2 and GLI3 can be processed into both transcriptional activator and repressor under specific condition, we first put our attentions on studying GLI1 mutations. To explore the effect of these nonsynonymous rare variants of *GLI1* on regulating SHH signaling pathway, we conducted dual-luciferase reporter assay using pGMGLI Lu reporter containing GLI binding sites upstream of firefly luciferase, and Renilla luciferase plasmid as internal control. As expected, wild-type *GLI1* dramatically activates the signaling compared with the empty vector group, whereas the p.G274C mutation shows increased activation, in contrast, p.R293H and p.L481X result in significantly decreased activation of the signaling in HEK293T cells. The result indicates that p.G274C is a gain-of-function mutation, whereas p.R293H and p.L481X are loss-of-function mutations. As predicted, the nonsense mutation p.L481X generating a truncated protein (**Supplementary Figure S3**) almost completely loses its ability in SHH signaling pathway activation (**Figure 2**).

p.G274C Mutation *GLI1* Protein Enhance the Protein Binding Affinity With DNA

As there is no doubt that the nonsense mutation of p.L481X disrupts protein function due to the loss of large portion of several critical domains (**Figure 1**), we focused on the other two harmful missense mutations of p.G274C and p.R293H as indicated above. Immunoblot analysis revealed no significant protein stability change made by two mutations (**Supplementary Figure S3**). Both mutations are mapped into the zinc finger domain of GLI1 protein (**Figure 1**), raising the possibility that they probably affect the protein binding ability with DNA. To this end, we performed electrophoretic mobility shift assay (EMSA) using nuclear extract from *GLI1* wild-type/mutation constructs transfected cells and synthesized probes containing GLI protein bind motif from *PTCH1* promoter. The result demonstrates that, the mutation of p.G274C results in significantly higher binding ability with GLI-Probe than wild-type (**Figure 3**), which probably accounts for its enhanced SHH signaling activation. Slightly higher binding ablility with GLI-Probe was also found in p.R293H (**Figure 3**).

The Mutation p.L481X Affects GLI1 Protein Subcellular Localization

GLI1 protein has two putative nuclear localization signals (NLS), including a monopartite signal NLS1 (aa 79-84) and a bipartite signal NLS2 (aa 383-401) (36). In addition, GLI1 possesses a leucine-rich nuclear export signal (NES) (aa 496-504), which is fully conserved in other vertebrate species (36, 37). In order to evaluate the functional consequences of the mutations, HeLa cells were transfected using Lipofectamine 3000 with various *GLI1* plasmid constructs. No significant differences in level of expression were detected between these GLI1 variants. Both proteins were readily detectable in the nucleus of transfected cells (**Figure 4**). However, compared to wild-type GLI1 protein, p.L481X leads to more accumulation in the nuclei, whereas p.G274C and p.R293H don't affect protein subcellular localization obviously.

Assessment of Human *GLI1* Wild-Type or Mutant Overexpression on Zebrafish Cardiogenesis

The Sonic hedgehog (SHH) signaling pathway is evolutionarily conserved and plays critical roles in organogenesis, particularly of the embryonic heart. Both elevated and repressed core factors in signal transduction will cause failure of heart development as precise spacial and temporal signaling regulation is crucial (38, 39). Hence, we performed plasmid microinjection into one to two cell stage fertilized zebrafish embryos to evaluate the dominant negative teratogenic effect of human *GLI1* overexpression on zebrafish cardiogenesis. Indeed, we observed

TABLE 2 | Bioinformatics analysis of patient-specific nonsynonymous rare mutations identified in *GLI1-3*.

Case	Age	Sex	Phenotype	Gene	Position	NA variant*	AA variant	Type of variant	GERP score	MAF in gnomAD**	MAF in HuaBiao	SIFT[a]	PP2[b]	CADD[c]
#1	0.6	F	VSD,ASD	*GLI1*	chr12:57860080	c.820G>T	p.G274C	Missense	4.93	0.00003184	NA	D	PrD	P
#2	8	F	TOF	*GLI1*	chr12:57860138	c.878G>A	p.R293H	Missense	4.93	NA	0.0001	D	PrD	P
#3	6	F	ASD, PDA, PH	*GLI1*	chr12:57863347	c.1442T>A	p.L481X	Stop-gain	4.66	NA	NA	NA	NA	P
#4	9.25	F	ASD, PS	*GLI1*	chr12:57864299	c.1776A>T	p.R592S	Missense	−0.05	NA	NA	D	PrD	P
#5	7.25	F	VSD, ASD, PDA	*GLI1*	chr12:57864374	c.1851G>A	p.M617I	Missense	−0.43	0.00001195	NA	T	B	B
#6	1.58	F	ASD, PDA	*GLI1*	chr12:57864447	c.1924C>T	p.P642S	Missense	4.44	0.000004010	NA	D	PrD	P
#7	1.17	M	AVCD, TAPVC, DORV	*GLI2*	chr2:121685021	c.233A>C	p.H78P	Missense	1.56	NA	NA	T	PrD	P
#8	1.17	M	VSD, PS, DORV	*GLI2*	chr2:121729629	c.1172C>T	p.A391V	Missense	−0.31	0.00001066	NA	T	B	B
#9	3.33	F	AVCD, PH	*GLI2*	chr2:121743938	c.2041G>A	p.V681M	Missense	−2.72	0.00007115	0.0001	D	B	P
#10	6.75	F	VSD, ASD, PDA	*GLI2*	chr2:121744044	c.2147G>T	p.G716V	Missense	0.81	0.000003988	NA	T	PrD	P
#11	1.75	F	ASD	*GLI2*	chr2:121744105	c.2208G>C	p.K736N	Missense	1.97	0.00005313	0.0004	D	PrD	P
#12	5	F	PDA	*GLI2*	chr2:121744161	c.2264A>G	p.N755S	Missense	−3.54	NA	NA	T	B	B
#13	3	M	VSD	*GLI2*	chr2:121746044	c.2554G>A	p.A852T	Missense	2.91	0.00003428	0.0004	T	PrD	P
#14	0.25	F	TAPVC	*GLI2*	chr2:121746828	c.3338C>T	p.A1113V	Missense	1.94	NA	NA	T	B	B
#15	7	M	DORV, PS	*GLI2*	chr2:121747359	c.3869C>T	p.P1290L	Missense	2.76	NA	NA	T	B	B
#16	11	M	VSD	*GLI2*	chr2:121747842	c.4352T>G	p.I1451S	Missense	4.22	NA	NA	T	PrD	P
#17	1.92	M	CoA	*GLI3*	chr7:42187969	c.223C>G	p.P75A	Missense	5.07	0.00004245	0.0001	T	PrD	P
#18	0.25	F	AVCD	*GLI3*	chr7:42088263	c.506C>T	p.P169L	Missense	4.91	0.00001593	0.0001	D	B	P
#19	0.92	F	AVCD	*GLI3*	chr7:42079808	c.857C>T	p.A286V	Missense	4.13	0.000007072	NA	T	B	P
#20	3	M	VSD	*GLI3*	chr7:42005382	c.3289G>T	p.V1097L	Missense	4.58	NA	NA	T	PrD	P

VSD, Ventricular septal defect; ASD, Atrial septal defect; PH, Pulmonary hypertension; PS, Pulmonary stenosis; AVCD, Atrioventricular canal defect; CoA, Coarctation of aorta.

**All the variants are heterozygous.*

***All of the variants we identified are heterozygous in gnomAD.*

[a]SIFT predictions: D represents damaging, T represents tolerated.

[b]PolyPhen-2 (PP2) predictions: B represents benign, PrD represents probably damaging.

[c]CADD predictions: P represents pathogenic, B represents benign.

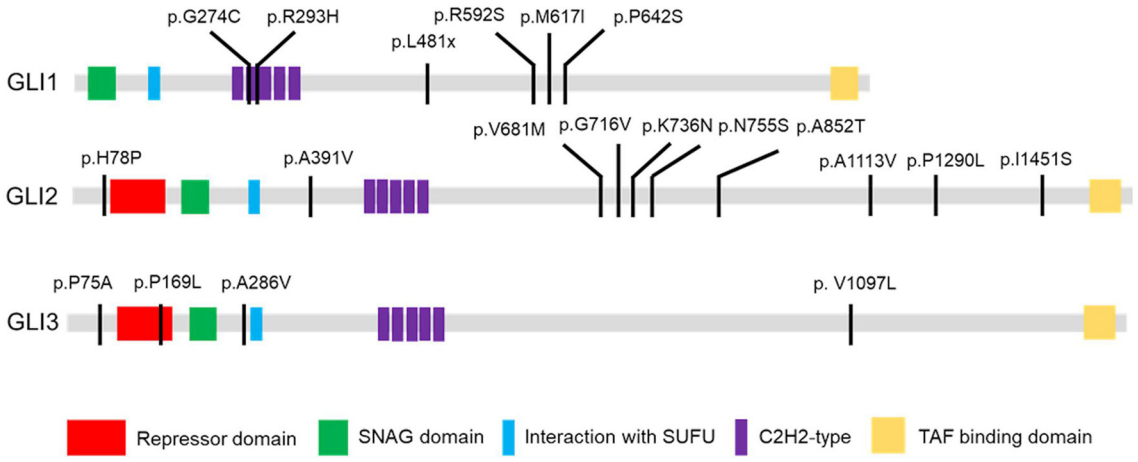


FIGURE 1 | GLI1-3 mutations were identified in CHD patients. Schematic diagram of case-specific mutations in GLI1, GLI2 and GLI3 protein.

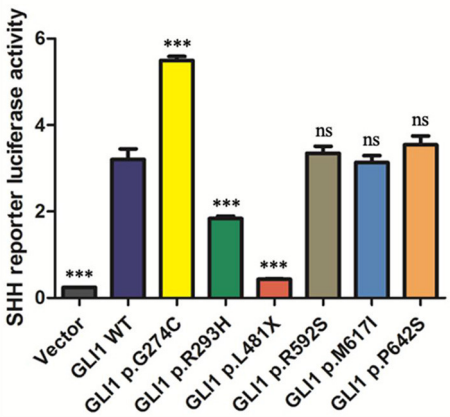


FIGURE 2 | The statistical analyses of luciferase relative activity for HEK293T cells co-transfected with expressing plasmids as indicated and Sonic hedgehog (SHH) signaling pathway reporter plasmids. Activation of SHH signaling pathway pGMGLI-Lu firefly luciferase reporter by human GLI1 wild-type or mutants 24 h post-transfection in HEK 293T cells. Constitutively expressed Renilla Reniformis Luciferase served as an internal control (n ≥ 3, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, ns, not significant, compared with wild-type group).

multiple cardiovascular abnormalities as a result of wild-type human *GLI1* gene overexpression in zebrasish embryos, including malformations on atrium, ventricle, atrioventricular septal and heart blood flow and beat rate (**Figure 5C**, **Supplementary Videos S1–S5**). We then calculated the enlarged pericardium (**Figure 5A**) which is a common phenotype of zebrafish heart malformations (25, 40–42) to evaluate GLI1 function alteration by mutations. As shown in **Figure 5B**, compared to the uninjected and empty vector groups, p.G274C resulted in the most heart malformation occurence, indicating its enhanced protein function over the wild-type form. Mutations p.R293H and p.L481X generated less abnormal embryos than wild-type GLI1 protein, showed the protein function loss by those two mutations. This result further confirmed that the function of *GLI1* gene was seriously affected by those three mutations.

DISCUSSION

SHH signaling exists in a variety of animals and plays a fundamental role in regulating accurate organization of the body plan, and its abnormal activation occurs in a variety of tumor cells (17). The GLI proteins are the main effectors of SHH signal and characterized as DNA-binding transcription factors (11, 43). The critical roles of *GLI1-3* genes in embryonic development has been well established and is manifested by the clinical features of patients with mutations in these genes (17, 18). However, the contribution of these genes on human cardiovascular system formation remains less studied. Herein, we conducted genetic screening in CHD patients and identified several case specific nonsynonymous rare mutations.

Previous studies have shown that perturbations of Hedgehog pathway can lead to developmental errors presenting partially overlapping clinical manifestations and atrioventricular canal defects (AVCD) as a common denominator. Moreover, Shh pathway also involved in cardiac outflow tract and neural crest development and, therefore, $Shh^{-/-}$ embryos display conotruncal and pharyngeal arch artery defects (44, 45). Some variants of *GLI1-3* genes have been found in several diseases including polydactyly, holoprosencephaly, Pallister-Hall syndrome and others (18, 46–51). However, our CHD patients are sporadic and non-syndromic cases in which ones showing additional symptoms such as polydactyly or holoprosencephaly should be excluded already. All the variants of *GLI1-3* genes we identified are heterozygous, and showed much stronger connection with AVSD and PDA than other subtypes of CHD. While this phenotype-genotype correlation deserves further validation by expanding the sample size in rare mutation study.

Dual-luciferase reporter assay revealed p.G274C mutation enhanced transcription activation, p.R293H decreased

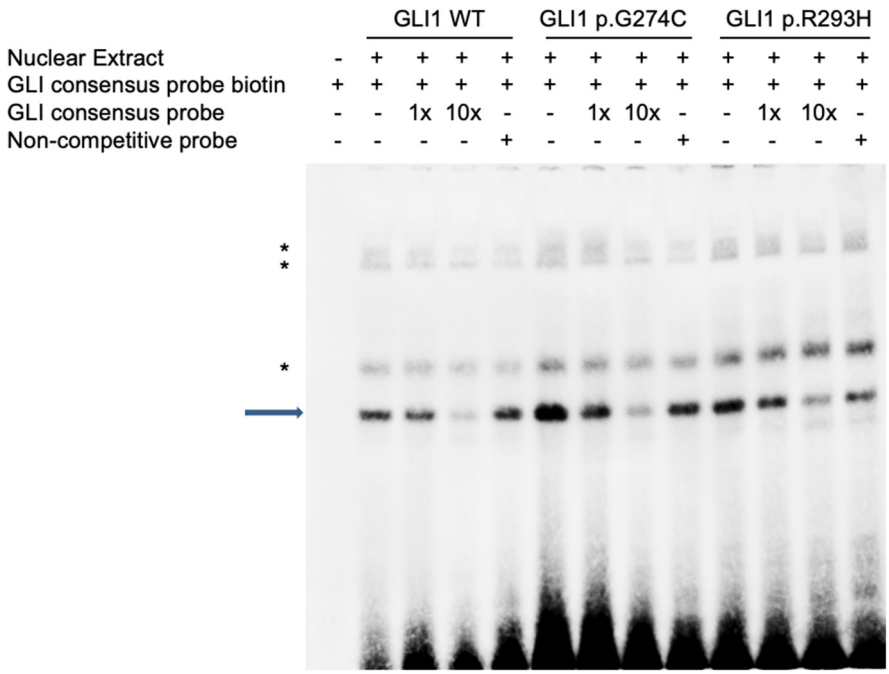


FIGURE 3 | GLI1 mutations affect protein binding affinity with DNA in Electrophoretic Mobility Shift Assay (EMSA). HEK293T cells were transfected with wild-type or mutant GLI1 expression constructs. Nuclear extract was prepared and incubated with specific DNA oligos as indicated. Arrow highlights specific GLI1-DNA binding signal. Asterisks represent non-specific signals.

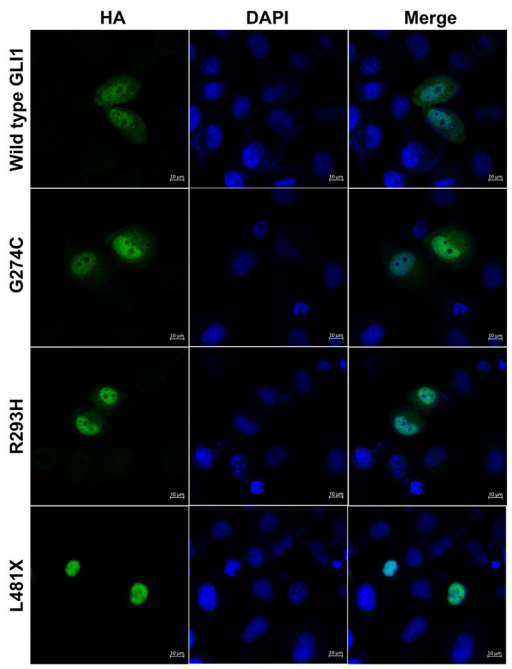


FIGURE 4 | GLI1 p.L481X affect protein subcellular localization. HeLa cells were transfected with HA-tagged wild-type or mutant *GLI1* expression constructs. Immunofluorescence was performed using anti-HA antibody (Green). Nuclei were counterstained with Hoechst (Blue). Scale bars are 10 μm.

transcription activation, whereas p.L481X almost completely loses transcription activation of SHH signaling pathway. EMSA showed p.G274C increased DNA-binding affinity and so did the mutant p.R293H. However, the luciferase assay showed that mutant p.G274C gains stronger function and mutant p.R293H decreases activation of the signaling compared to the function of wild GLI1 protein. Then we defined mutant p.G274C as gain-of-function and mutant p.R293H as loss-of-function. The fact that p.R293H shows loss-of-function does not necessarily need to be associated with decreased DNA-binding ability. The mutation specific characteristics may determine that its impact on the reduction of function is much greater than its impact on the ability to bind DNA. Therefore, the comprehensive performance of mutant p.G293H is that the function is reduced and the DNA-binding ability is reversely increased. Alternatively, the loss-of-function gained by mutant p.G293H might be too prominent to be balanced by the reverse effect of increased DNA-binding ability.

Usually, stop-gain mutations with aberrant translation termination leads to the degradation of the mRNA via activation of nonsense-mediated mRNA decay (NMD) and fail to produce a stable truncated protein (52). In response to NMD of mutated allele, wild-type allele always increasingly expressed to avoid haploinsufficiency. However, our data have demonstrated that the L481X truncated protein can exist stably and the truncated protein is largely produced *in vivo* (**Supplementary Figure S3**). The truncated protein not only failed to activate SHH signaling, but also placed additional inhibitory effects on SHH target

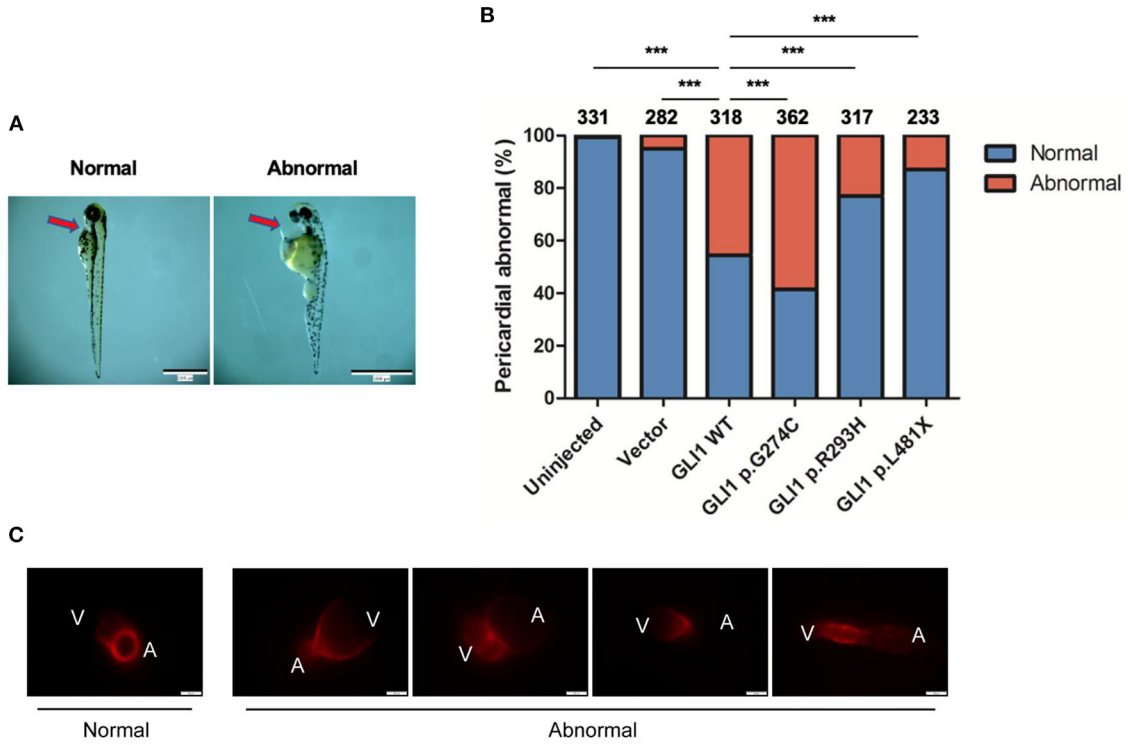


FIGURE 5 | Effect of overexpressed human *GLI1* wild-type or mutant on zebrafish heart development. Empty vector or *GLI1* wild-type/mutant expression construct were injected into 1-2 cell stage fertilized zebrafish embryos, and the teratogenic effect in zebrafish cardiogenesis was calculated 72 h post injection. **(A)** Red arrows indicate the pericardial cavity of zebrafish embryos. **(B)** The frequencies of pericardial abnormalities. ($n \geq 200$, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, compared with wild-type group). **(C)** Phenotype of Cmlc2-mCherry zebrafish after overexpressing wild-type human *GLI1*. *GLI1* wild-type expression construct were injected into 1-2 cell stage fertilized Cmlc2-mCherry zebrafish embryos. Images were captured with Olmpus IX83 72 hours post injection. V, ventricle; A, atrium.

genes (**Figures 2**, **4**). Therefore, mutation p.L481X resulting haploinsufficiency was supported by the evidence that the heterozygous mutation p.L481X caused reverse phenotypes when the wild-type allele expression was strictly limited in half by the existed truncated allele. Because the result that heterozygous mutation p.L481X caused reverse phenotypes was partially attributed to the truncated allele, it's also hard to conclude that what we observed was merely caused by the half-expressed wild-type allele.

The alterations of critical genes in the early stage of cardiac development, such as *de novo* mutations, copy number variants, common variants, noncoding mutations and epigenetic modification, have contributed to the occurrence of CHD (5, 53–55). The *in vitro* function assays indicated that compared to the normal function of wild-type GLI1 protein, mutation p.R293H shows loss of function effect and mutation p.L481X also causes severely loss-of-function effect due to the stop-gain, while, mutation p.R274C shows gain of function. When we used overexpression system to test the cardiac morphology in zebrafish, as expected, the overexpressed wild-type *GLI1* gene would lead to observed cardiac anomalies because of the dosage imbalance in addition to the existed endogenous *GLI1* working normally. Similarly, compared to the wild-type GLI1 protein, the overexpressed p.R293H mutated GLI1 or p.L481X mutated GLI1 protein would result in less reverse effect *in vivo* because they are partially or maximum lose the normal function as wild-type GLI1 protein. Supposed mutation p.R274C gains of the original function, the most severe reverse effect on zebrafish phenotypes would be imaginable. In other words, the *in vitro* and *in vivo* function assays kept consistent although the outcomes looked a sign of contradiction.

Our functional study demonstrated some of those mutations significantly interfere with the protein function, thus lead to dysregulated SHH signaling pathway. Our study suggests the strong association of *GLI1* gene mutations with the CHD occurrence. Interestingly, both gain and loss-of-function nonsynonymous rare mutations in *GLI1* gene were identified in the CHD patients. As GLI1 protein simply serves as a transcriptional activator in SHH signaling pathway, our results suggest that either elevated or repressed GLI1 protein activity caused imbalanced signaling transduction could be a causative factor of CHD. This fits the fact that temporally and spatially tight regulation of key factors in fundamental developmental pathways underlies the normal embryogenesis.

However, since our samples are all sporadic and lack pedigrees, the mutations we found might not be the only causative factor leading to CHD. It is possible that the patients with *GLI1-3* mutations may also carry harmful mutation(s)

in other gene(s). We presume that single deleterous mutation contributes partially to the CHD and multiple such mutations help individual approaches the threshold of CHD occurrence, as our group demonstrated in another congenital malformation recently (56).

In conclusion, in present study, we systematically identified a set of novel rare mutations in *GLI1-3* genes in human CHD that have the potential to be used diagnostically. Our data support our hypothesis that in *GLI1*, both gain and loss-of-function mutations are likely to be associated with CHD. Further understanding of the regulation and balance of SHH signaling pathway is crucial for understanding CHD etiology.

DATA AVAILABILITY STATEMENT

The data in the present study are available from the corresponding authors upon reasonable request. The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the School of Life Sciences, Fudan University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. The animal study was reviewed and approved by Ethics Committee of the School of Life Sciences, Fudan University.

AUTHOR CONTRIBUTIONS

RP, BL, LL, and HW designed the project and prepared and edited the manuscript. RP, BL, and LL performed the experiments and analyzed the data. SC, YG, and XY contributed to the zebrafish embryo injection experiment. ZS and LY contributed to the clinical specimens collection and sequencing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from National Key R&D Program of China (2021YFC2701101, HW) and National Natural Science Foundation of China (81930036 and 82150008, HW) and the Commission for Science and Technology of Shanghai Municipality (20ZR1404800, LL).

ACKNOWLEDGMENTS

We sincerely thank all the individuals who participated in this study and the clinicians who helped in recruiting the volunteers.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcvm.2022.798033/full#supplementary-material

REFERENCES

1. Triedman JK, Newburger JW. Trends in congenital heart disease: the next decade. *Circulation.* (2016) 133:2716–33. doi: 10.1161/CIRCULATIONAHA.116.023544
2. Zhao LJ, Chen LZ, Yang TB, Wang TT, Zhang SM, Chen LT, et al. Birth prevalence of congenital heart disease in China, 1980–2019: a systematic review and meta-analysis of 617 studies. *Eur J Epidemiol.* (2020) 35:631–42. doi: 10.1007/s10654-020-00653-0
3. Botto LD, Mulinare J, Erickson JD. Occurrence of congenital heart defects in relation to maternal multivitamin use. *Am J Epidemiol.* (2000) 151:878–84. doi: 10.1093/oxfordjournals.aje.a010291
4. Botto LD, Mulinare J, Erickson JD. Do multivitamin or folic acid supplements reduce the risk for congenital heart defects? Evidence and gaps. *Am J Med Genet A.* (2003) 121A:95-101. doi: 10.1002/ajmg.a.20132
5. Gelb BD, Chung WK. Complex genetics and the etiology of human congenital heart disease. *Cold Spring Harb Perspect Med.* (2014) 4:a013953. doi: 10.1101/cshperspect.a013953
6. Blue GM, Krik EP, Giannoulatou E, Sholler GF, Dunwoodie SL, Harvey RP, et al. Advances in the genetics of congenital heart disease: a clinician's guide. *J Am Coll Cardiol.* (2017) 69:859–70. doi: 10.1016/j.jacc.2016.11.060
7. Ingham PW, Nakano Y, Seger C. Mechanisms and functions of Hedgehog signalling across the metazoa. *Nat Rev Genet.* (2011) 12:393–406. doi: 10.1038/nrg2984
8. Carpenter D, Stone DM, Brush J, Ryan A, Armanini M, Frantz G, et al. Characterization of two patched receptors for the vertebrate hedgehog protein family. *Proc Natl Acad Sci USA.* (1998) 95:13630–4. doi: 10.1073/pnas.95.23.13630
9. Zaphiropoulos PG, Undén AB, Rahnama F, Hollingsworth RE, Toftgård RE. *PTCH2*, a novel human patched gene, undergoing alternative splicing and up-regulated in basal cell carcinomas. *Cancer Res.* (1999) 59:787–92.
10. Altaba ARI, Sánchez P, Dahmane N. Gli and hedgehog in cancer: tumours, embryos and stem cells. *Nat Rev Cancer.* (2002) 2:361–72. doi: 10.1038/nrc796
11. Altaba ARI. Gli proteins encode context-dependent positive and negative functions: implications for development and disease. *Development.* (1999) 126:3205–16. doi: 10.1007/s004290050265
12. Hynes M, Stone DM, Dowd M, Pitts-Meek S, Goddard A, Gurney A, et al. Control of cell pattern in the neural tube by the zinc finger transcription factor and oncogene *Gli-1*. *Neuron.* (1997) 19:15–26. doi: 10.1016/S0896-6273(00)80344-X
13. Persson M, Despina S, Pascal TW, Elisabet A, Jens BS, Ulrich R, et al. Dorsal-ventral patterning of the spinal cord requires Gli3 transcriptional repressor activity. *Genes Dev.* (2002) 16:2865. doi: 10.1101/gad.243402
14. Hoffmann AD, Peterson MA, Friedland-Little JM, Anderson SA, Moskowitz IP. Sonic hedgehog is required in pulmonary endoderm for atrial septation. *Development.* (2009) 136:1761–70. doi: 10.1242/dev.034157
15. Jiang J, Hui CC. Hedgehog signaling in development and cancer. *Dev Cell.* (2008) 15:801–12. doi: 10.1016/j.devcel.2008.11.010
16. Bijlsma MF, Peppelenbosch MP, Spek CA. Hedgehog morphogen in cardiovascular disease. *Circulation.* (2006) 114:1985–91. doi: 10.1161/CIRCULATIONAHA.106.619213
17. Riddle RD, Johnson RL, Laufer E, Tabin C. Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell.* (1993) 75:1401–16. doi: 10.1016/0092-8674(93)90626-2

18. Palencia-Campos A, Ullah A, Nevado J, Yildirim R, Unal E, Ciorraga M, et al. GLI1 Inactivation is associated with Developmental Phenotypes Overlapping with Ellis-Van Creveld Syndrome. *Hum Mol Genet.* (2017) 26:4556–71. doi: 10.1093/hmg/ddx335
19. Goel HL, Underwood JM, Nickerson JA, Hsieh CC, Languino LR. β1 integrins mediate cell proliferation in three-dimensional cultures by regulating expression of the sonic hedgehog effector protein, GLI1. *J Cell Physiol.* (2010) 224:210–7. doi: 10.1002/jcp.22116
20. Palle K, Mani C, Tripathi K, Athar M. Aberrant GLI1 activation in DNA damage response, carcinogenesis and chemoresistance. *Cancers (Basel).* (2015) 7:2330–51. doi: 10.3390/cancers7040894
21. Liang SZ, Shi X, Yu CX, Shao XL, Zhou HT Li XY, et al. Identification of novel candidate genes in heterotaxy syndrome patients with congenital heart diseases by whole exome sequencing. *Biochim Biophys Acta Mol Basis Dis.* (2020) 1866:165906. doi: 10.1016/j.bbadis.2020.165906
22. Kim PCW, Mo R, Hui CC. Murine models of VACTERL syndrome: role of sonic hedgehog signaling pathway. *J Pediatr Surg.* (2001) 36:381–4. doi: 10.1053/jpsu.2001.20722
23. Qiao XJ, Liu YH Li PQ, Chen ZZ Li HL, Yang XY, et al. Genetic analysis of rare coding mutations of *CELSR1-3* in congenital heart and neural tube defects in Chinese people. *Clin Sci (Lond).* (2016) 130:2329–40. doi: 10.1042/CS20160686
24. Li BB Yu LW, Liu D, Yang XY, Zheng YF, Gui YH, et al. MIB1 mutations reduce Notch signaling activation and contribute to congenital heart disease. *Clin Sci (Lond).* (2018) 132:2483–91. doi: 10.1042/CS20180732
25. Zhu MJ, Ma XY, Ding PC, Tang HF, Peng R, Lu L, et al. Novel mutations of *AXIN2* identified in a Chinese congenital heart disease cohort. *J Hum Genet.* (2019) 64:427–35. doi: 10.1038/s10038-019-0572-x
26. Hao M, Pu WL, Li Y, Wen SQ, Sun C, Ma YY, et al. The huabiao project: whole-exome sequencing of 5,000 han chinese individuals. *J Genet Genomics.* (2021) doi: 10.1016/j.jgg.2021.07.013
27. Li BB, Niswander LA. TMEM132A, a novel wnt signaling pathway regulator through wntless (WLS) interaction. *Front Cell Dev Biol.* (2020) 8:599890. doi: 10.3389/fcell.2020.599890
28. Agren M, Kogerman P, Kleman MI, Wessling M, Toftgård R. Expression of the *PTCH1* tumor suppressor gene is regulated by alternative promoters and a single functional Gli-binding site. *Gene.* (2004) 330:101–14. doi: 10.1016/j.gene.2004.01.010
29. Yu LW, Wang F, Yang XY, Sun SN, Zheng YF Li BB, et al. Mild decrease in *TBX20* promoter activity is a potentially protective factor against congenital heart defects in the Han Chinese population. *Sci Rep.* (2016) 6:23662. doi: 10.1038/srep23662
30. Ye JH, Tong YL, Lv JS, Peng R, Chen SX, Kuang LL, et al. Rare mutations in the autophagy-regulating gene *AMBRA1* contribute to human neural tube defects. *Hum Mutat.* (2020) 41:1383–93. doi: 10.1002/humu.24028
31. Ng PC, Henikoff S, SIFT. Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* (2003) 31:3812–4. doi: 10.1086/175612
32. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet.* (2013). doi: 10.1002/0471142905.hg0720s76
33. Rentzsch P, Witten D, Cooper GM, Shendure J. Kircher. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* (2019) 47:D886–94. doi: 10.1093/nar/gky1016
34. van der Velde KJ, de Boer EN, van Diemen CC, Sikkema-Raddatz B, Abbott KM, Knopperts A, et al. GAVIN: Gene-Aware Variant INterpretation for medical sequencing. *Genome Biol.* (2017) 18:6. doi: 10.1186/s13059-016-1141-7
35. Cooper GM, Stone EA, Asimenos G, Green ED, Batzoglou S, Sidow A. Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res.* (2005) 15:901–13. doi: 10.1101/gr.3577405
36. Barnfield PC, Zhang XY, Thanabalasingham V, Yoshida M, Hui CC. Negative regulation of Gli1 and Gli2 activator function by Suppressor of fused through multiple mechanisms. *Differentiation.* (2005) 73:397–405. doi: 10.1111/j.1432-0436.2005.00042.x
37. Kogerman P, Grimm T, Kogerman L, Krause D, Undén AB, Sandstedt B, et al. Mammalian suppressor-of-fused modulates nuclear-cytoplasmic shuttling of Gli-1. *Nat Cell Biol.* (1999) 1:312–9. doi: 10.1038/13031
38. Olson EN. Gene regulatory networks in the evolution and development of the heart. *Science.* (2006) 313:1922–7. doi: 10.1126/science.1132292
39. Verma MK, Lenka N. Temporal and contextual orchestration of cardiac fate by WNT-BMP synergy and threshold. *J Cell Mol Med.* (2010) 14:2094–108. doi: 10.1111/j.1582-4934.2009.00774.x
40. Nasevicius A, Larson J, Ekker SC. Distinct requirements for zebrafish angiogenesis revealed by a *VEGF-A* morphant. *Yeast.* (2000) 17:294–301. doi: 10.1002/1097-0061(200012)17:4<294::AID-YEA54>3.0.CO;2-5
41. Wang WD, Huang CJ, Lu YF, Hsin JP, Prabhakar VR, Cheng CF, et al. Heart-targeted overexpression of *Nip3a* in zebrafish embryos causes abnormal heart development and cardiac dysfunction. *Biochem Biophys Res Commun.* (2006) 347:979–87. doi: 10.1016/j.bbrc.2006.06.174
42. Gou DZ, Zhou J, Song QX, Wang ZJ, Bai XM, Zhang YD, et al. *Mog1* knockout causes cardiac hypertrophy and heart failure by downregulating *tbx5-cryab-hspb2* signalling in zebrafish. *Acta Physiol (Oxf).* (2021) 231:e13567. doi: 10.1111/apha.13567
43. Hui CC, Angers S. Gli proteins in development and disease. *Annu Rev Cell Dev Biol.* (2011) 27:513–37. doi: 10.1146/annurev-cellbio-092910-154048
44. Digilio MC, Pugnaloni F, Luca AD, Calcagni G, Baban A, Dentici ML, et al. Atrioventricular canal defect and genetic syndromes: The unifying role of sonic hedgehog. *Clin Genet.* (2019) 95:268–76. doi: 10.1111/cge.13375
45. Washington Smoak IW, Byrd NA, Abu-Issa R, Goddeeris MM, Anderson R, Morris J, et al. Sonic hedgehog is required for cardiac outflow tract and neural crest cell development. *Dev Biol.* (2005) 283:357–72. doi: 10.1016/j.ydbio.2005.04.029
46. Ullah A, Umair M, Majeed AI. Abdullah, Jan A, Ahmad W. A novel homozygous sequence variant in GLI1 underlies first case of autosomal recessive pre-axial polydactyly. *Clin Genet.* (2019) 95:540–1. doi: 10.1111/cge.13495
47. Palencia-Campos A, Martínez-Fernández M, Altunoglu U, Soto-Bielicka P, Torres A, Marín P, et al. Heterozygous pathogenic variants in *GLI1* are a common finding in isolated postaxial polydactyly A/B. *Hum Mutat.* (2020) 41:265–76. doi: 10.1002/humu.23921
48. Yousaf M, Ullah A, Azeem Z, Majeed AI, Memon MI, Ghous T, et al. Novel heterozygous sequence variant in the *GLI1* underlies postaxial polydactyly. *Congenit Anom (Kyoto).* (2020) 60:115–9. doi: 10.1111/cga.12361
49. Umair M, Ahmad F, Ahmad S, Alam Q, Rehan M, Alqosaibi AI, et al. A novel homozygous missense mutation in the zinc finger DNA binding domain of GLI1 causes recessive post-axial polydactyly. *Front Genet.* (2021) 12:746949. doi: 10.3389/fgene.2021.746949
50. Arnhold IJP, França MM, Carvalho LR, Mendonca BB, Jorge AAL. Role of *GLI2* in hypopituitarism phenotype. *J Mol Endocrinol.* (2015) 54:R141–50. doi: 10.1530/JME-15-0009
51. Andreu-Cervera A, Catala M, Schneider-Maunoury S. Cilia, ciliopathies and hedgehog-related forebrain developmental disorders. *Neurobiol Dis.* (2021) 150:105236. doi: 10.1016/j.nbd.2020.105236
52. Karousis ED, Mühlemann O. Nonsense-Mediated mRNA Decay Begins Where Translation Ends. *Cold Spring Harb Perspect Biol.* (2019) 11:a032862. doi: 10.1101/cshperspect.a032862
53. Pierpont ME, Basson CT, Benson DW, Gelb BD, Giglia TM, Goldmuntz E, et al. Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation.* (2007) 115:3015–38. doi: 10.1161/CIRCULATIONAHA.106.183056
54. Bentham J. Bhattacharya. Genetic mechanisms controlling cardiovascular development. *Ann N Y Acad Sci.* (2008) 1123:10–9. doi: 10.1196/annals.1420.003
55. Ordovás JM, Smith CE. Epigenetics and cardiovascular disease. *Nat Rev Cardiol.* (2010) 7:510–9. doi: 10.1038/nrcardio.2010.104
56. Chen ZZ, Lei YP, Zheng YF, Aguiar-Pulido V, Ross ME, Peng R, et al. Threshold for neural tube defect risk by accumulated

singleton loss-of-function variants. *Cell Res.* (2018) 28:1039–41. doi: 10.1038/s41422-018-0061-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership