The Genetic Variations that Increase the Risk of Heart Failure of Pakistani Descendants

By

[Name]

[Student Number]

[Name of Supervisor]

Declaration

Acknowledgements

Table of Contents

Title **P.No.**

Abstract

Introduction: Heart failure is a pathophysiologic condition that occurs secondary to cardiovascular diseases and in which the requirements of cardiac tissues are not adequately met. Although the causes for heart failure are not clear, several studies have suggested a strong genetic basis for its development. **Aim:** This study sought to identify those genetic variants that might specifically lead to heart failure in the Pakistani population. **Methods:** The entire list of genes implicated in heart failure was acquired from the OMIM database and the Pakistani data set was acquired from the 1000 Genomes database. The list of variants in each of these genes in the Pakistani population was acquired and their pathogenecities were predicted. **Results:** A total of 11 genes and 15 variants were found to be deleterious in the Pakistani population. **Discussion:** The possibly lethal variants in cardiac-related genes were identified in the Pakistani population. These need to be checked further through gene sequencing in susceptible individuals in order to identify their role in causing heart failure.

Chapter 1: Introduction

1.1 Heart Failure

Heart failure is a pathophysiologic condition in which blood and nutrient delivery is insufficient for cardiac tissue requirements. It usually occurs secondary to cardiovascular diseases such as cardiomyopathy, atherosclerosis, myocarditis, valvular disease, and congential conditions. Although the etiology of heart failure is unclear, it is largely attributed to lifestyle factors such as weight management, physical activity, diet, smoking, and drinking, and to a smaller extent to genetic causes (Morita et al., 2005). In a majority of cases, it is believed that a multifactorial interaction between lifestyle and genetic factors lead to the development of heart diseases (Skrzynia et al., 2015).

The development of heart failure is a multifactorial and multi-step process beginning with initial damage to the myocardium, extrinsic forces acting on the heart, and an immense functional load on the left ventricle. A combination of two or more of these factors lead to the development of chronic congestive heart failure over months or even years. Myocardial damage is most commonly caused due to coronary artery disease that can lead to ischemia and myocardial infarction in patients. Other possible causes of myocardial damage include autoimmune conditions, metabolic injuries, and infections (Figueroa and Peters, 2006). The initial damage to myocardium is most commonly followed by ventricular remodeling and further damage to the normal areas of the myocardial tissues. As a result, several changes in cell morphology and metabolism take place, eventually leading to fibrosis and hypertrophy. All these changes predispose the heart to ventricular remodeling wherein the efficiency of cardiac function decreases, consequentially leading to heart failure (Braunwald, 2013).

Heart failure is a common condition in all countries around the world, affecting around 5 million individuals in the United States of America (USA) and 2.8 million individuals in Pakistan (Sheikh, 2006). It is widely prevalent in the elderly population with around 80% of the cases occuring in people aged 65 and above (Nadar et al., 2005). The mortality rate of heart failure is estimated to be around 2 years in about 45 to 50% of all cases (Khand et al., 2000). The financial burden of heart failure is also huge

for a nation, accounting for around 27.9 billion dollars in the United States annually which constitutes around 3% of the entire federal budget (Nadar et al., 2005). In Pakistan alone, around 8.2% of the elderly population is afflicted with heart failure, and this number is projected to increase to around 20% by 2050. Around 50 heart patients in Pakistan suffer a heart attack every hour leading to the death of around 1,100 people per day due to heart failure. The most common risk factors for heart failure in Pakistan are found to be obesity, diabetes, hypertension, unhealthy lifestyle habits, and lack of sufficient check-ups in people susceptible to heart failure (Pillai and Ganapathi, 2013).

1.2 Genetic Basis of Heart Failure

The genome of every human being contains several million genetic variants either in the form of Single Nucleotide Polymorphisms (SNPs) or large copy number (around 1000 nucleotides) variants. Several of these genetic variants are common in over 90% of the human population whereas a few are specific to a given family or clan reflecting the birth of novel mutations with each generation. A vast majority of these variations are benign and do not confer a major phenotypic change to the individual; however, a few of these variations are connected to specific phenotypes or disorders in humans. A comprehensive study of these few specific genetic variants can provide deep insights into the genetic basis and linkage for heart failure, and enable us to understand the percentage of the population that is at risk for developing heart disease (Skrzynia et al., 2015).

It is estimated that around 18% of people who develop heart disease are geneticaly predisposed to this condition. Analysis of the Framingham database cohort has shown that if one of the parents has heart disease, their child has a 70% higher risk of acquiring heart disease as compared to the general population. Additionally, the Framingham Offspring Study found that heart failure in parents was highly associated with left ventricular dysfunction and an increased risk of heart failure in the offspring (Lee et al., 2006). Another study carried out in Sweden further confirmed genetic predisposition to be an important risk factor for heart failure. This study found that presence of genetic factors was associated with early-onset heart failure and the risk for acquiring heart failure was not only tied to parents but also to siblings suffering from this

condition (Lindgren et al., 2016). Although, statistically, the contribution of genetic variants in the development of heart failure is minor, its role cannot be ignored, and studying its effects can help us understand the complex genetic mechanisms that contribute to heart disease (Morita et al., 2005).

Specifically in the case of heart failure, the genetic etiology is contributed by several variants in multiple genes and therefore, heart failure is considered to be a multifactorial condition. The affected genes and their protein products may be a part of several diverse pathways such as those that are involved in the initial myocardial damage, progression of the heart condition, or response of an individual to therapy. Studies that aim to identify possible genetic causes for heart failure employ a candidate gene approach wherein the gene sequences of proteins that are involved in, for instance, adrenergic and renin-angiotensin-aldosterone systems pathways are scanned for possibly lethal variants (Skrzynia et al., 2015).

1.3 Literature Review

As heart failure is considered multifactorial, it has been identified that a number of genetic variations with minor effects are involved in its etiology rather than one major mutation. Genes involved in adrenergic and renin-angiotensin-aldosterone system pathways have been extensively linked to heart failure (Cappola and Dorn, 2011). In particular, polymorphisms in the β-1 adrenergic receptor genes (Dorn, 2010), G-protein coupled receptor kinase genes (Liggett et al., 2008), angiotensin converting enzyme (ACE) genes (Danser et al., 1995), ubiquitin-specific protease genes, LRIG3 gene, and CMTM7 gene (Smith et al., 2010) have been shown to be associated with heart failure in candidate gene studies. Although these genetic variants are not exclusive in the development of heart failure, a study of their prevalence in different groups of populations gives several insights into the mechanisms that contribute to the development of cardiovascular disease and heart failure in different populations.

Two of the most common SNPs that are found to be associated with heart failure are p.Arg83Gly in the CLCNKA gene (Tavira et al., 2013) and p.Ser38Gly in the KCNE gene (Fatini et al., 2010). Genome-Wide Association Studies (GWAS) performed by

Larson et al. (2007) analysed 70,987 SNPs at once, and they found a single SNP, rs740363, to be associated with heart failure. In another study carried out by the Cohorts for Heart and Ageing Research in Genomic Epidemiology (CHARGE), 2,478,304 SNPs were evaluated in 24,000 participants with heart failure belonging to 4 different cohorts. This study identified 2 chromosomal locations, rs10519210 at 15q22 and rs11172782 at 12q14, which are located close to the genes USP3, CA12, and LRIG3 (Smith et al., 2010). However, the challenge here is that these three genes have not yet been implicated in cardiovascular function, and therefore, the mechanisms with which they may contribute to heart failure remain poorly understood (Pastorekova et al., 2008).

A study conducted by Cappola et al. (2010) identified two more SNPs, rs1739843 and rs6787362, both of which have been replicated in other studies. The SNP, rs1739843, has been associated with both ischemic heart failure as well as non-ischemic heart failure and therefore, it is a strong candidate for providing a genetic basis for heart failure. This SNP is located within the HSPB7 gene which codes for a heat shock protein B7 expressed exclusively in the skeletal and cardiac muscles (Kampinga et al., 2009). Variants in another protein HSPB5 belonging to the same family as the previous gene have been found to be associated with familial cardiomyopathy (Krief et al., 1999). Furthermore, the gene HSPB7 is found close to another gene CLCNKA which has also demonstrated the presence of SNPs in other GWAS studies for heart failure (Barlassina et al., 2007). On the other hand, the SNP rs6787362 is present within the FRMD4B gene which is a participant in the PI-3 kinase signaling pathway that is widely implicated in the cardiovascular system (Klarlund et al., 2001).

The challenge with GWAS studies is that it becomes very difficult to understand the biological relationship between the variant and its contribution to heart failure. Therefore, another approach that has been widely adopted by researchers is candidate gene analysis and several studies have sought to identify possibly lethal variations in the ADRA1A, ADRB2, PLN, HSPB7, and HSP7 genes. Twelve variants identified in the HSPB7 gene were found to be associated with systolic heart failure; however, none of

them were found to bring about any alteration in the amino acid sequence of the gene product (Matkovich et al., 2010).

Apart from heart failure, several studies have been conducted to identify genetic variants that could possibly contribute to the etiology and heredity of other cardiac diseases. One of the most significant causes of heart failure is dilated cardiomyopathy (DCM) accounting for about 40% of all heart failure cases (Karkkainen and Peuhkurinen, 2007). One of the SNPs mentioned above, rs1739843, is found to be a common variant for both DCM and heart failure, thereby enhancing its significance as a potential genetic factor in the etiology of heart failure (Watkins et al., 2011).

Several studies have focused on genetic variations involved in causing different types of cardiovascular diseases in Pakistani subjects that ultimately lead to heart failure. Shakeel et al (2018) analyzed publicly available genomic datasets of Pakistani subjects to identify pathogenic genetic variants linked to cardiovascular disease. They identified 115 deleterious and 44 likely deleterious genetic variants linked to cardiovascular disease that were found in much higher proportion in Pakistani subjects (Shakeel et al., 2018). Another study conducted by Shakeel et al (2018) aimed to analyze one of the leading causes of heart failure, dilated cardiomyopathy (DCM). This study identified 5 specific mutations that contributed to DCM in the affected subjects. These included loss-of-function mutations in MYOM3, TMED4, and C2orf40; a frameshift mutation in RTKN2; and a deletion mutation in SLC6A6. These mutations were found to have a higher prevalence in affected Pakistani subjects as compared to other population groups indicating that these were unique to the Pakistani cohort. However, this study analyzed the genomes of only 5 patients making it a small sample size. Analysis of a larger population would have given stronger and deeper genetic insights into the causes of heart failure (Shakeel et al., 2018).

Other studies have concentrated on mutations in specific genes for cardiovascular diseases in the Pakistani population. Rehman et al (2019) have studied the low density lipoprotein receptor gene and the contribution of its specific variants to the development of coronary artery disease (CAD). In this study too, the sample size was small (40 patients) and a larger sample size could have given further insights into

the role of this gene in CAD (Rehman et al., 2019). Cheema et al (2015) studied genetic variants of the APOE gene and their role in coronary stenosis in Pakistani subjects. A genome analysis of 695 subjects found the presence of APOE polymorphisms in more than 70% of samples indicating a strong association of this gene with coronary stenosis (Cheema et al., 2015).

Despite studies that have focused on different causes of heart failure, no comprehensive study has been performed on Pakistani subjects that has analyzed the role of genetic variants in heart failure. Although several genes have been identified to contribute to the etiology of heart failure, the prevalence of genetic variations in these genes has not been studied in the Pakistani population. Knowledge of genetic variants that are directly linked to heart failure in the Pakistani cohort can help in the design of testing and therapeutic strategies to reduce the burden of cardiovascular diseases in Pakistan.

1.4 Rationale of the Study

Pakistan is the second largest country in South Asia and the sixth largest country in the world (The World Bank, 2016). In the Pakistani population alone, around 2.8 million people are affected with congestive heart failure (Sheikh, 2006). This accounts for around 26.9% of the total Pakistani population with equal prevalence in both men and women (Jafar et al., 2005). Out of the affected population, more than 30% of people with cardiovascular disease and heart failure belong to the age group of over 45 years. As per reports from the World Health Organization, heart failure accounts for the largest number of deaths annually due to non-communicable diseases (Shahid et al., 2018).

Identification of genetic variants that lead to heart failure is of great clinical significance in predicting risk factors and response to therapeutic interventions. These types of studies are important for the development of comprehensive genetic tests that can identify people at risk for developing heart disease and treat them effectively for better prognosis. It can help identify predisposition to heart failure before visible symtpoms develop and make use of preventative interventions to prevent or delay the development of cardiovascular disease (Czepluch et al., 2018).

Currently, due to the rapid development in gene sequencing techniques and the overload of genetic data from hundreds of studies around the world, we have at our disposal in-depth genetic and proteomic information of different cohorts. This technological advancement has resulted in the development of huge DNA sequence datasets at a population level that has been accomplished by large-scale projects such as the 1000 Genomes project (Auton et al., 2015) and the Exome Aggregation Consortium (ExAC) (Lek et al., 2016). The data available through these projects have been widely used for analysing frequencies of different variants at individual loci in specific population groups and to discover cohort-specific genes and their variations in relation to disease susceptibility (Lettre, 2014).

1.5 Research Hypotheses

This research study has been designed on the basis of the following hypotheses:

1. There are several genetic variants, specific to the Pakistani population, which are associated with the risk of heart failure.
2. The identified genetic variants are deleterious to the Pakistani cohort in that they may result in either the structural and/or functional disruption in protein function thereby contributing to the pathology of heart failure.

1.6 Aims and Objectives

This research study aims to analyze the specific genetic variants in the Pakistani population that increase the risk of heart failure in this cohort. The objectives of this study are three-fold:

- To identify genetic variants that are linked to the risk of heart failure in Pakistani subjects
- To identify the prevalence of these variants in the Pakistani cohort in order to understand how strongly they are associated with the risk of heart failure
- To study the pathogenecity of these variants in the human genome in order to understand their deleterious effects in the Pakistani cohort

Chapter 2: Methods

2.1 Preparation of Genes List

The two databases, Online Mendelian Inheritance in Man (OMIM) and ClinVar, were used to acquire a list of genes that were associated with heart failure. As this study intended to identify genetic variants associated only with heart failure and not any other cardiovascular condition, the terms ‘heart failure’ and ‘cardiac failure’ alone were used to search for the gene list in all these databases. The gene list obtained was also compared with the entries in the Human Phenotype Ontology and International Classification of Diseases (ICD-10) databases. Once the gene lists from all these databases were obtained, they were compiled into a single list by eliminating duplicate entries and grouping them based on their position on the chromosome.

2.2 Preparation of Data Set

In order to identify the genetic variants that were possibly associated with heart failure in the Pakistani population, the Pakistani data set was acquired from the 1000 Genomes Project browser on the NCBI website. This provided an exhaustive list of variants found in the Pakistani population throughout the genome.

2.3 Data Analysis

A mix of manual and computational analysis was carried out in order to determine possibly lethal variations that are associated with heart failure in the Pakistani population. The genomic coordinates of each of the selected genes were noted using the NCBI website. In order to account for promoter regions and other possible regulatory regions, 2000 nucleotides were added both before the start position of the gene and after the end position of the gene by subtracting 2000 from the first number and adding 2000 to the second number of the gene position. Once the genomic coordinates of the selected genes were calculated, all variants within the gene position were acquired.

As we intended to identify those variants that could possibly be involved in the pathogenesis of heart failure in Pakistani individuals, we decided to use only variants

that resulted in frameshift or nonsense mutations in the amino acid sequence. Additionally, we decided to analyse variants having low-allele frequencies in the population as this type of variants is predicted to be associated with complex conditions in humans. For this study, we decided to use variants with a frequency of 0.05 or lower in the population. In order to determine the possible effects of variants on the protein structure and function, we used Ensembl Variation Viewer to predict the pathogenecity of each variant. This website uses several prediction tools and algorithms such as Sorting Intolerant From Tolerant (SIFT), PolyPhen, Rare Exome Variant Ensembl Learner (REVEL), MetaIR, and MutationAssessor to predict if a given variant is benign, possibly lethal, or deleterious.

Chapter 3: Results

3.1 Preparation of Genes List

The OMIM and ClinVar databases were used to acquire a complete list of genes that were associated with heart failure in humans. After comparing the entries acquired from both these databases and eliminating duplicates, all genes were categorized based on their chromosomal location and a final list of 336 genes were obtained. In order to understand the physiological roles of these genes in the human body, a few representative genes were checked in the UniProt Gene Ontology Annotation database. The findings showed that most of these genes were involved in several structural and functional processes of the heart such as physiological regulation, formation of anatomical structures, genesis and organization of cellular compartments, metabolic processes, developmental processes, catalysis, binding, and molecular transduction. This indicates that these genes may be involved in the structural and functional processes relating to heart failure.

3.2 Acquisition of Data Set

The Pakistani data set was obtained from the 1000 Genomes browser on the NCBI website. For each variant, the allele frequencies were accessed as this indicated if a given variant was present in the Pakistani population and if it was present, the frequency at which it was present was noted.

3.3 Preparation of Variants List

Each gene in the list was checked in the 1000 Genomes browser of the NCBI database and the list of all variants in the specific gene was acquired for the Pakistani population. While compiling the list, it was noted that several variants present in other population groups of the world were absent in the Pakistani population. As we were only concerned with variants present in the Pakistani population, these variants were eliminated from the list.

Once the entire variant list was acquired for each gene, the variants with an allele frequency of 0.05 or higher were eliminated from the list. As we wanted to identify the

variants that were associated with a risk of heart failure, we decided to narrow down our search to those variants that had a low minor allele frequency as these are predicted to have more deleterious effects in the human genome. Additionally, we also eliminated intron variants and synonymous variants as these have a lower chance of contributing to a disruption in protein structure and function. Therefore, the final list of variants included 5' UTR variants, 3' UTR variants, missense variants, stop gained variants, initiator codon variants, and splice donor variants bringing the final number to 149 variants with a low minor allele frequency and predicted to bring about a structural and/or functional disruption to the protein product.

3.4 Pathogenecity Prediction of Variants

Using the Ensembl Variation Viewer, each of the variants was checked for its predicted deleterious effects on protein function. The pathogenecity of the variants were predicted using five different algorithms and the variants that were considered to be deleterious were chosen on the basis that they were predicted to be pathogenic by at least two of the five algorithms.

SIFT uses sequence homology and similarities in chemical composition and hydophobicity between the two alternate amino acids to predict the possibility of a variant affecting protein function. Variants with scores less than 0.05 were considered deleterious and those with scores greater than 0.05 were considered to be benign or tolerated. PolyPhen uses sequence homology and 3D protein structures to predict the pathogenecity of a variant. In contrast to SIFT, higher scores in PolyPhen indicate pathogenecity whereas lower scores indicate that the variant is benign. In our study, a value greater than 0.908 was considered 'probably damaging', a value between 0.446 and 0.908 was considered 'possibly damaging', and a value less than 0.446 was considered benign.

Three other algorithms that were used for pathogenecity prediction of variants were REVEL, MetalR, and MutationAssessor. REVEL uses predictions from more than 10 different algorithms including SIFT and PolyPhen to predict if a variant is disease causing or benign. Scores greater than 0.5 classified the variant as 'likely disease

causing', and scores lower than 0.5 classified the variant as 'likely benign'. MetaIR uses allele frequencies and logistic regression to classify variants as damaging or tolerated. Similar to REVEL, score higher than 0.5 are considered damaging and scores lower than 0.5 are considered tolerated. MutationAssessor uses the evolutionary conservation status of the specific amino acid affected by the variant to predict if the presence of the variant is likely to disrupt protein function. This algorithm classifies variants as low, medium, high, and neutral based on its effect on protein function where higher scores represent high likelihood of pathogenecity and lower scores represent low likelihood of pathogenecity.

Apart from these five tools, two other tools were also used to understand the biological effects of a given variant. One of them was Combined Annotation Dependent Depletion (CADD) and the other was Genomic Evolutionary Rate Profiling (GERP). CADD is a tool that is used to predict the deleterious effects of SNPs and insertion/deletion variants by combining data from evolutionary conservation of the amino acid and functional effects on the protein. CADD scores above 30 are considered to be deleterious and lower than 30 are considered to be benign. GERP is a tool that evaluates evolutionary conservation of the specific amino acid by performing multiple sequence alignment with gene sequences of several species. If a positive score is obtained, it means that the amino acid is highly conserved among species and if the score is negative, it means that the amino acid is not conserved across different species.

After checking each variant in the Ensembl Variation Viewer and assessing the pathogenecity predictions provided by different algorithms, a total of 55 variants from the entire list of 149 were found to be deleterious or possibly lethal to the structure and/or function of the protein product. Each of the genes that hosted these variants is described in detail in the subsequent section.

3.5 Genes Found to Have Variants with Potentially Deleterious Effects

An entire list of genes along with their variants that were predicted to be disease causing is provided in Table 3.1.

S.No.	Gene Name	Variant(s)
1	SELENON	rs546041571
2	BIN1	rs571987587
3	IDUA	rs370442463; rs574025652
4	IL6ST	rs551130374; rs2228044
5	PRDM6	rs550841106
6	SGCD	rs45559835
7	DSP	rs17604693
8	HFE	rs1800730
9	CD36	rs548507859; rs149985988
10	ZFPM2	rs121908601
11	FXN	rs541981554
12	FKTN	rs41277797
13	PHYH	rs538686726; rs62619919
14	CTNNA3	rs543012210; rs570159272
15	LDB3	rs565557622; rs45618633
16	DCHS1	rs140245002
17	GRK2	rs528538949
18	DDX11	rs201612562
19	GNPTAB	rs555336070
20	SGCG	rs527562042
21	PDPK1	rs370069297
22	TNFSF12	rs540997935
23	ERBB2	rs564064363
24	JUP	rs200327969
25	CDK5RAP3	rs544149741
26	SGCA	rs565069721
27	ACE	rs551801825; rs565263717
28	GAA	rs573556709; rs1800307
29	PIEZO2	rs536250412
30	DSG2	rs546455032; rs576404380
31	EPG5	rs186446511; rs72918350; rs78339727; rs527293119
32	ATP8B1	rs535998516; rs571233811
33	RTTN	rs34717557; rs577535023
34	XRCC1	rs2307177; rs2228487; rs572056306
35	RBCK1	rs544518686
36	JAG1	rs574205422
37	GLA	rs28935490
38	FHL1	rs764929890
39	FLNA	rs782269425

Table 3.1: List of variants that were found to be deleterious in the Pakistani population in genes implicated in cardiac function

The structural and functional aspects of each of these genes along with their expression patterns as given in the NCBI website are discussed below.

3.5.1 SELENON

Selenoprotein-N is a glycoprotein that plays a role in protection of the cell against oxidative stress and in the regulation of calcium homeostasis in the cell. It is present in

the endoplasmic reticulum where it is found to execute most of its functions. Mutations in this protein have been found to be associated with muscle disorders or myopathies such as muscular dystrophy. It is expressed in 27 tissues in the body including the heart.

3.5.2 BIN1

Bridging Integrator 1 is a nucleocytoplasmic adaptor protein with tumor suppression properties. There are several isoforms of this protein expressed in different tissues such as the central nervous system, muscle, brain, kidney, and heart. Preliminary studies in mouse have suggested that this protein may have a significant role to play in the development of cardiac muscles.

3.5.3 IDUA

Alpha L-iduronidase, as the name suggests, is a protein that hydrolyzes the alpha L-iduronic acid residues of heparan sulphate and dermatan sulphate in their lysosomal degradation process. Mutations in this gene have been linked to Mucopolysaccharidosis Type I (MPS I). This protein is expressed in 27 different tissues in the body including the heart, although with a lower expression profile as compared to other tissues.

3.5.4 IL6ST

Interleukin 6 Signal Transducer functions as a signal transducer protein in a cytokine receptor complex for many different cytokines such as interleukin, neurotrophic factor, oncostatin, and leukemia inhibitory factor. Studies have shown that this gene may have an important role to play in apoptosis in myocytes. This protein is expressed in heart at a lower concentration as compared to other tissues.

3.5.5 PRDM6

PR/SET Domain 6 is a transcriptional repressor protein and mutations in this gene have been linked to Patent Ductus Arteriosus 3 (PDA 3). This protein is involved in

regulating contractile proteins in vascular smooth muscle cells. It is expressed in the heart at a very low concentration as compared to other tissues.

3.5.6 SGCD

Sarcoglycan Delta is one of the four components of the sarcoglycan complex which forms a link between the extracellular matrix and the actin cytoskeleton. This gene is heavily expressed in cardiac and skeletal muscles, and mutations in this gene have been linked to dilated cardiomyopathy and autosomal recessive limb-girdle muscular dystrophy.

3.5.7 DSP

Desmoplakin is used as an anchor for intermediate proteins within functional desmosomes. Although it is expressed at a low concentration in the heart, mutations in this gene are linked to different types of cardiomyopathies and also, keratodermas.

3.5.8 HFE

Homeostatic Iron Regulator associates with beta-2 microglobulin in the regulation of iron absorption. Mutations in this gene are associated with hemochromatosis, which is an iron storage disorder. This gene has a medium-level expression in the heart tissue.

3.5.9 CD36

The CD36 protein is a surface glycoprotein on platelets and it acts as a receptor for thrombospondin involved in various cell adhesion processes. This gene is well-expressed in the heart tissue and mutations in this gene are shown to be linked to platelet glycoprotein deficiency.

3.5.10 ZFPM2

ZFPM2 is a zinc finger protein belonging to the FOG family of transcription factors and it regulates the activity of GATA proteins. It shows medium-level expression

in the heart and it has been widely implicated in the processes of cardiogenesis and hematopoiesis.

3.5.11 FXN

Frataxin is a mitochondrial protein that regulates respiration and iron transport in the mitochondria. It is expressed in the heart along with 26 other tissues in the body and mutations in this gene have been linked to Friedreich ataxia.

3.5.12 FKTN

Fukutin is a transmembrane protein that localizes to the Golgi apparatus where it functions as a glycosyltransferase for alpha-dystroglycan in the skeletal muscle. This protein is shown to play an important role in the development processes of the brain and mutations in this gene are found to be associated with dilated cardiomyopathy and muscular dystrophy.

3.5.13 PHYH

Phytanoyl-CoA 2-hydroxylase is a peroxisomal protein that is involved in the process of fatty acid oxidation. Mutations in this gene have been associated with Refsum disease, Zellweger syndrome, and rhizomelic chondrodysplasia punctata. This gene shows a medium level expression in the heart tissues.

3.5.14 CTNNA3

Catenin alpha-3 is found mostly in muscle cells where it functions in cell-cell adhesion processes. It shows a very high expression level in the heart and mutations in this gene have been linked to arrhythmogenic right ventricular dysplasia.

3.5.15 LDB3

The Lim Domain Binding 3 protein is involved in cytoskeletal assembly, clustering, and targeting of membrane proteins. It is highly expressed in the heart, and mutations in this gene have been linked to dilated cardiomyopathy and myofibrillar myopathy.

3.5.16 DCHS1

Dachsous Cadherin-related 1 protein functions in cell-cell adhesion processes in a calcium-dependent manner. It is expressed in the fibroblasts where it is shown to promote wound healing. It shows a medium-level expression in the heart along with 23 other tissues in the body.

3.5.17 GRK2

G-protein coupled Receptor Kinase 2 phosphorylates several important receptor proteins and transcription factors involved in various regulatory processes in the body. This protein has been shown to play a role in heart function, embryonic development, and metabolism. High levels of this protein have been found in patients who have had heart failure.

3.5.18 DDX11

DEAD/H-box Helicase 11 is an RNA helicase that is involved in altering the secondary structure of RNA to initiate translation. Apart from that, it is also involved in nuclear splicing and spliceosome assembly. This protein has been implicated in spermatogenesis, embryogenesis, cellular growth, and cell division. It shows a low to medium-level expression in the heart.

3.5.19 GNPTAB

This gene encodes the alpha and beta subunits of N-acetylglucosamine-1-phosphate Transferase. This protein is present in the Golgi apparatus where it plays a role in certain synthesis processes. It shows a medium level expression in the heart and mutations in this gene have been linked to mucolipidosis.

3.5.20 SGCG & SGCA

These genes encode the gamma and alpha subunits of sarcoglycan respectively, which is a sarcolemmal transmembrane glycoprotein involved in the dystrophin-glycoprotein complex (DGC). This gene shows high expression in the heart and mutations in this gene have been linked to muscular dystrophy.

3.5.21 PDPK1

3-Phosphoinositide Dependent Protein Kinase 1 is a mitochondrial protein which is involved in the oxidative decarboxylation of pyruvate and homeostasis regulation of carbohydrates. It shows a high level of expression in the heart along with several other tissues in the body.

3.5.22 TNFSF12

TNF Superfamily member 12 is a cytokine which is involved in signaling pathways for apoptosis and cell death. Apart from that, it regulates angiogenesis by promoting the proliferation and translocation of endothelial cells. It is expressed at a medium level in the heart.

3.5.23 ERBB2

Erb-b2 Receptor Tyrosine Kinase 2 is a receptor for epidermal growth factor (EGF) and mutations in this gene are shown to play a role in breast and ovarian cancers. It shows a medium-level expression in the heart.

3.5.24 JUP

Junction Plakoglobin is a cytoplasmic protein which forms an important component of plaques in both intermediate junctions and desmosomes. It shows a medium-level expression in the heart and mutations in this gene have been associated with Naxos disease.

3.5.25 CDK5RAP3

CDK5 Regulatory subunit Associated Protein 3 plays an important role in cellular processes of cell cycle progression and transcriptional regulation. It shows a low-level expression in the heart and is involved in tumorigenesis and metastasis.

3.5.26 ACE

Angiotensin I Converting Enzyme is a protein that plays a role in the regulation of blood pressure and electrolyte balance. It is involved in the conversion of angiotensin I

into angiotensin II, which is a vasopressor and stimulator of aldosterone. Although it shows a low-level expression in the heart, mutations in the gene are implicated in cardiovascular conditions, renal disease, psoriasis, stroke, and Alzheimer's disease.

3.5.27 GAA

Alpha-Glucosidase is a lysosomal enzyme which is involved in the conversion of glycogen into glucose. As a result, mutations in this gene are implicated in Pompe's disease or glycogen storage disease. It shows a medium level expression in the heart.

3.5.28 PIEZO2

Piezo type mechanosensitive ion channel component 2 is involved in connecting mechanical forces to physiological signals by means of mechanically-activated cation channels. It shows minimal expression in the heart and mutations in this gene are linked to distal arthrogryposis.

3.5.29 DSG2

Desmoglein 2 is a transmembrane glycoprotein present on the desmosomes and is involved in cell-cell adhesion especially in the myocardial and epithelial cells. It shows a medium level expression in the heart and mutations in this gene have been shown to be associated with arrhythmogenic right ventricular dysplasia.

3.5.30 EPG5

Ectopic P-Granules autophagy protein 5 homolog plays an important role in autophagy when the cell is under starvation conditions. It shows a medium level expression in the heart and mutations in this gene are associated with Vici syndrome.

3.5.31 ATP8B1

ATPase Phospholipid Transporting 8B1 is involved in the transport of phosphatidylethanolamine and phosphatidylserine across the lipid bilayer membrane. This gene has a very low expression profile in the heart and mutations in this gene have been associated with intrahepatic cholestasis.

3.5.32 RTTN

Rotatin is a protein whose function is not clearly known yet. Mutations in this gene have been shown to be associated with seizures and polymicrogyria. Also, it is found to localize to ciliary basal bodies in the cell and shows a medium level expression in the heart.

3.5.33 XRCC1

X-ray Repair Cross Complementing 1 is involved in DNA repair of single-stranded breaks via the base excision repair pathway. It shows a medium level expression profile in the heart and mutations in this gene have been implicated in certain cancers.

3.5.34 RBCK1

RANBP2-type and C3HC4-type zinc finger containing 1 protein interacts with the UIP28 protein in cells. It shows a medium level expression in the heart.

3.5.35 JAG1

Jagged Canonical Notch Ligand 1 is involved in various signaling processes and hematopoiesis. It shows a medium level expression in the heart and mutations in this gene are linked to Alagille syndrome.

3.5.36 GLA

Galactosidase Alpha is involved in the hydrolysis of terminal galactose units from glycoproteins and glycolipids. It has a low level expression in the heart and mutations in this gene are shown to result in a lysosomal storage disorder known as Fabry disease.

3.5.37 FHL1

Four and a Half Lim Domains 1 comprises of zinc finger domains and is involved in the regulation of several cellular processes. It shows a very high level of expression in the heart and mutations in this gene have been linked to Emery-Dreifuss muscular dystrophy.

3.5.38 FLNA

Filamin A binds actin elements and crosslinks them to membrane glycoproteins. As a result, it is involved in cytoskeleton remodeling, alteration of cell shape, and migration of the cell. It shows a low level expression in the heart and mutations in this gene are associated with heterotopia, otopalatodigital syndrome, frontometaphyseal dysplasia, and idiopathic intestinal pseudoobstruction.

3.6 Genes with Deleterious Variants Involved in Heart Function

Based on the above discussion, a few of the genes that were found to possess potentially lethal variants in the Pakistani population and were directly implicated in heart function include BIN1, SGCD, DSP, ZFPM2, FKTN, CTNNA3, LDB3, GRK2, ACE, DSG2, and FHL1. The specific variants of these genes along with their structural change at the protein level and pathogenecity prediction scores are given in Table 3.2.

Chr	Position	Variant ID	Reference Allele	Alternative Allele	MAF	Gene	SIFT	PolyPhen	REVEL	MetaIR	Mutation Assessor
2	127816616	rs571987587	C	T	0.0052	BIN1	0.03	0.991	0.168	0.316	0.403
5	155935708	rs45559835	G	A	0.0052	SGCD	0	0.308	0.422	0.336	0.48
6	7565727	rs17604693	A	T	0.0208	DSP	0.01	0.836	0.593	0.86	0.696
8	106431420	rs121908601	A	G	0.0052	ZFPM2	0.03	0.537	0.329	0.041	0.295
9	108363426	rs41277797	C	T	0.0052	FKTN	0.01	0.543	0.412	0.549	0.518
10	67748542	rs543012210	G	C	0.0052	CTNNA3	0.02	0.228	0.283	0.227	0.54
10	69299367	rs570159272	C	T	0.0052	CTNNA3	0	0.973	0.34	0.354	0.518
10	88439873	rs565557622	C	A	0.0052	LDB3	0	0.922	0.245	0.362	0.835
10	88478529	rs45618633	G	A	0.0104	LDB3	0	0.975	0.605	0.717	0.603
11	67049412	rs528538949	C	G	0.0052	GRK2	0	1	0.659	0.384	0.994
17	61561215	rs551801825	T	G	0.0052	ACE	0	1	0.575	0.417	0.96
17	61571795	rs565263717	C	A	0.0052	ACE	0	0.915	0.641	0.41	0.949
18	29104457	rs546455032	G	A	0.0052	DSG2	0.01	0.949	0.46	0.449	0.649
18	29115328	rs576404380	A	G	0.0052	DSG2	0	0.964	0.404	0.367	0.901
X	135289290	rs764929890	A	G	0.0069	FHL1	0.03	0.793	0.55	0.689	0.448

Table 3.2: List of variants in genes that are involved in heart function along with their pathogenecity prediction

Chapter 4: Discussion

To our knowledge, till date, no study has ever reported the possibly lethal genetic variants and/or genes that are associated with heart failure in the Pakistani population. Using publicly available databases and bioinformatics tools, we finalised a list of 336 genes which have been found to be associated with heart failure. The possible variants in these genes were identified and checked for their presence and allele frequencies in the Pakistani data set. After eliminating intron variants, synonymous variants, and those variants that were predicted to be benign in their effects on protein structure and function, we were left with a list of 55 deleterious or possibly deleterious variants in 39 genes. Each of these genes was further analysed for its expression level in the heart, its involvement in cardiac function, and its contribution to cardiac diseases and a final list of 11 genes was obtained.

BIN1 is a membrane scaffolding protein present in cardiomyocytes and is responsible for the regulation of t-tubule function and calcium signaling in these cells. This protein has 10 different isoforms with one of the isoforms expressed in the heart. The cardiac BIN1 protein has been shown to be reduced in patients with heart failure, thereby increasing the risk of ventricular arrhythmias and cardiomyopathies (Zhou and Hong, 2017). A recent study by Jiang et al. (2019) found that reduction in the levels of BIN1 in the heart can lead to abnormalities in cardiac contraction increasing the risk of arrhythmias and ultimately leading to heart failure. The study also suggested that BIN1 can be used as an important predictor of heart failure in the future in susceptible patients (Jiang et al., 2019).

GRK2 has also been consistently demonstrated to play a role in heart failure. This biomarker is involved in the desensitization of GPCRs and beta-adrenergic receptors, and is found to be upregulated in patients with heart failure. It has been seen that high levels of this protein can disrupt the functions of beta-adrenergic receptors and lowering the levels of this protein can efficiently reverse these effects (Woodall et al., 2015). Studies conducted in different species have shown that inhibiting GRK2 in cardiac cells can prevent and/or reverse the effects of heart failure. Research groups

around the world have also seriously considered using GRK2 as a gene therapy target to reverse the symptoms of heart failure in patients (Reinkober et al., 2012).

The enzymatic product of ACE, angiotensin II, is responsible for maintaining cardiovascular homeostasis, and monitoring blood pressure and systemic volume. It is involved in several signaling processes for arterial vasoconstriction, vascular smooth muscle contraction, cell proliferation, and oxidative stress. Increased levels of angiotensin II in the heart can lead to the development of myocardial fibrosis and hypertrophy, both of which are characteristic features of heart failure. Use of ACE inhibitors such as lisinopril and enalapril for the management of heart failure has been shown to reduce mortality in these patients (Shah et al., 2017). Therefore, the role of ACE in the development of heart failure has been well studied.

DSG2 is another gene that has been linked to arrhythmogenic right ventricular cardiomyopathy (ARVC) as a hereditary component (Syrris et al., 2007). In another study, DSG2 has shown a high degree of association with sustained ventricular arrhythmia, heart failure, and death due to cardiac conditions and/or heart transplantation (Hermida et al., 2019). Thus, mutations in DSG2 have been highly implicated in the etiology of heart failure.

DSP is a desmosomal gene that has been linked to arrhythmogenic right ventricular cardiomyopathy, left ventricular predominant disease, ventricular tachyarrhythmia, cardiocutaneous syndrome, and it may also be associated with dilated cardiomyopathy (Lakdawala et al., 2013). A study by Smith et al. (2020) has revealed that mutations in the DSP gene are linked to arrhythmogenic cardiomyopathy, episodic myocardial injury, left ventricular fibrosis, systolic dysfunction, and ventricular arrhythmias. No functional studies have been performed to identify the role of this gene in heart failure; however, genome-wide association studies (GWAS) and candidate gene studies have found significant links between mutations in DSP and heart failure (Tayal et al., 2017).

CTNNA3 has been shown to be important for proper regulation and functioning of cardiomyocytes with a particularly high expression in these cells (Ameri et al., 2020).

Studies have identified links between mutations in this gene and the development of dliated cardiomyopathy, which is a precursor of heart failure, arrhythmias, and arrhythmogenic cardiomyopathy (Janssens et al., 2003; Norland et al., 2019; Sommariva et al., 2017). However, all these studies are genetic studies and no functional study has been undertaken to uncover links between variants in CTNNA3 and the development of heart disease. FHL-1 mutations have been shown to be linked to left ventricular hypertrophy, ventricular tachycardia, and hypertrophic cardiomyopathy leading to sudden death in patients (Scott Binder et al., 2020). Apart from that, it is also associated with X-linked myopathy which leads to the development of cardiac abnormalities such as spongious hypertrophic cardiomyopathy characterised by reduced systolic and diastolic functions (Binder et al., 2012). However, no study has directly linked FHL-1 mutations to heart failure.

Other genes such as the SGCD have been implicated in heart function without much evidence about its role in heart failure. For instance, people with mutation in this gene are known to develop fatal cardiomyopathies and myocardial damage. This has been confirmed by using Sgcd-null mice and visualizing the progression of cardiomyopathy in these mice along with symptoms such as decrease in myocardial contractility, increase in preload and decrease in afterload, and a high cardiac output (Bauer et al., 2008). However, its role in the development of heart failure has not yet been studied.

A study conducted by Qian et al. (2017) has reported candidate gene analysis of a few genes and the results show significant associations between mutations in ZFPM2 and congenital heart disorder. However, no other studies have reported the role of this gene in heart function, or the development of heart failure or other heart conditions. Similarly, genetic studies have described links between mutations in the FKTN gene and the development of cardiac disorders related to muscular dystrophies and systolic dysfunction (Finsterer and Stollberger, 2016). No other studies have been performed to analyse a possible role of FKTN in cardiac diseases not related to muscular dystrophies. Mutations in the gene LDB3 have also been associated with idiopathic

dilated cardiomyopathy (Wang et al., 2019) without any clear implications for heart failure.

Based on our analysis, it was apparent that literature in the field of genetic variants related to heart failure is lacking. Several studies have reported the genetic and physiological basis of cardiomyopathies and other heart conditions. While most of these are precursors to the development of heart failure, there are very few studies that have specifically demonstrated the role of genes and their variants in the development of heart failure. Also, population-level studies for specific genetic variants and their allele frequencies within the population are very few. This limits our knowledge of genetic variants related to heart failure in specific populations. This study aimed to bridge this gap by identifying those pathogenic genetic variants that were present in the Pakistani population and could be contributing to heart failure. More studies of this type will help us understand hereditary mechanisms of life-threatening conditions and focus on targeted therapies for management of these conditions.

Limitations of the Study

This study has several limitations. The first one is that it has only considered Single Nucleotide Polymorphisms (SNPs) in the genome of the Pakistani cohort and not copy number variants or, on a much larger scale, chromosomal aberrations. However, the latter two categories may also be involved in increasing the susceptibility of the Pakistani population to heart failure; however, they have not been addressed in this study.

The second limitation of this study is that heart failure is a multi-etiological and multi-factorial condition. However, the variants of genes that have been indirectly associated with heart failure have not been comprehensively assessed. Rather, only variants of genes that have a direct correlation with heart development, heart function, and pathophysiology of heart disorders have been considered to play a role in contributing to heart failure. This does not intend to neglect the role of those genes that have not yet been directly linked to heart function, and there is always the possibility that these genes may directly and strongly contribute to the etiology of heart failure.

The third limitation of this study is that it has used the publicly available Pakistani data set from the 1000 Genomes project as sample rather than whole genome sequence data of actual Pakistani subjects. As spontaneous mutations can take place at any time in a person's life, some of the more recent variants in this cohort may have been overlooked that may be contributing to heart failure in the Pakistani population.

Implications for Future Studies

Our study has identified deleterious variants in 11 genes that may be contributing to heart failure in the Pakistani population. Future studies need to recruit patients with heart failure and analyse the sequences of these 11 genes in their and their family's DNA samples. This will enable us to see if the variants found in our study match with the variants identified in the sequences of actual Pakistani patients with heart failure. Apart from that, functional studies also need to be carried out to understand the structural and functional consequences of each of these variants on the protein product. These studies will help in understanding and uncovering the exact physiological mechanisms by which these variants may be contributing to heart failure.

Final Conclusion

This study aimed to identify those genetic variants that were implicated in the development of heart failure in the Pakistani population. Using bioinformatics tools and publicly available databases, the genes implicated in heart failure were identified and the specific variants of these genes present in the Pakistani cohort were acquired. Those variants that had high allele frequencies and were found to be benign were eliminated from the final list. At the end of the process, we were left with 11 genes with deleterious variants in the Pakistani cohort that could possibly be associated with heart failure risk in the population. Functional studies on these variants can further confirm their physiological consequences in the heart tissues.

References

Ameri, P., Schiattarella, G. G., Crotti, L., Torchio, M., Bertero, E., Rodolico, D., . . . Perrino, C. (2020). Novel Basic Science Insights to Improve the Management of Heart Failure: Review of the Working Group on Cellular and Molecular Biology of the Heart of the Italian Society of Cardiology. *International Journal of Molecular Sciences, 21*(4), 1192. doi:10.3390/ijms21041192

Barlassina, C., Fiume, C. D., Lanzani, C., Manunta, P., Guffanti, G., Ruello, A., . . . Cusi, D. (2007). Common genetic variants and haplotypes in renal CLCNKA gene are associated to salt-sensitive hypertension. *Human Molecular Genetics, 16*(13), 1630-1638. doi:10.1093/hmg/ddm112

Bauer, R., Macgowan, G., Blain, A., Bushby, K., & Straub, V. (2008). Steroid treatment causes deterioration of myocardial function in the -sarcoglycan-deficient mouse model for dilated cardiomyopathy. *Cardiovascular Research, 79*(4), 652-661. doi:10.1093/cvr/cvn131

Binder, J. S., Weidemann, F., Schoser, B., Niemann, M., Machann, W., Beer, M., . . . Windpassinger, C. (2012). Spongious Hypertrophic Cardiomyopathy in Patients With Mutations in the Four-and-a-Half LIM Domain 1 Gene. *Circulation: Cardiovascular Genetics, 5*(5), 490-502. doi:10.1161/circgenetics.111.962332

Braunwald, E. (2013). Heart Failure. *JACC: Heart Failure, 1*(1), 1-20. doi:10.1016/j.jchf.2012.10.002

Cappola, T. P., & Dorn, G. W. (2011). Clinical Considerations of Heritable Factors in Common Heart Failure. *Circulation: Cardiovascular Genetics, 4*(6), 701-709. doi:10.1161/circgenetics.110.959379

Cappola, T. P., Li, M., He, J., Ky, B., Gilmore, J., Qu, L., . . . Dorn, G. W. (2010). Common Variants in HSPB7 and FRMD4B Associated With Advanced Heart Failure. *Circulation: Cardiovascular Genetics, 3*(2), 147-154. doi:10.1161/circgenetics.109.898395

Cheema, A. N., Bhatti, A., Wang, X., Ali, J., Bamne, M. N., Demirci, F. Y., & Kamboh, M. I. (2015). APOEGene Polymorphism and Risk of Coronary Stenosis in Pakistani Population. *BioMed Research International, 2015*, 1-5. doi:10.1155/2015/587465

Czepluch, F. S., Wollnik, B., & Hasenfuß, G. (2018). Genetic determinants of heart failure: Facts and numbers. *ESC Heart Failure, 5*(3), 211-217. doi:10.1002/ehf2.12267

Danser, A. H., Schalekamp, M. A., Bax, W. A., Van den Brink, A. M., Saxena, P. R., Riegger, G. A., & Schunkert, H. (1995). Angiotensin-converting enzyme in the

human heart. Effect of the deletion/insertion polymorphism. *Circulation, 92*, 1387-1388.

Dorn, G. W. (2010). Adrenergic Signaling Polymorphisms and Their Impact on Cardiovascular Disease. *Physiological Reviews, 90*(3), 1013-1062. doi:10.1152/physrev.00001.2010

Fatini, C., Sticchi, E., Marcucci, R., Verdiani, V., Nozzoli, C., Vassallo, C., . . . Gensini, G. F. (2010). S38G single-nucleotide polymorphism at the KCNE1 locus is associated with heart failure. *Heart Rhythm, 7*(3), 363-367. doi:10.1016/j.hrthm.2009.11.032

Figueroa, M. S., & Peters, J. I. (2006). Congestive Heart Failure: Diagnosis, Pathophysiology, Therapy, and Implications for Respiratory Care. *Respiratory Care, 51*(4), 403-412.

Finsterer, J., & Stöllberger, C. (2016). Heart Disease in Disorders of Muscle, Neuromuscular Transmission, and the Nerves. *Korean Circulation Journal, 46*(2), 117. doi:10.4070/kcj.2016.46.2.117

Hermida, A., Fressart, V., Hidden-Lucet, F., Donal, E., Probst, V., Deharo, J., . . . Gandjbakhch, E. (2019). High risk of heart failure associated with desmoglein-2 mutations compared to plakophilin-2 mutations in arrhythmogenic right ventricular cardiomyopathy/dysplasia. *European Journal of Heart Failure, 21*(6), 792-800. doi:10.1002/ejhf.1423

Jafar, T. H., Jafary, F. H., Jessani, S., & Chaturvedi, N. (2005). Heart disease epidemic in Pakistan: Women and men at equal risk. *American Heart Journal, 150*(2), 221-226. doi:10.1016/j.ahj.2004.09.025

Janssens, B., Mohapatra, B., Vatta, M., Goossens, S., Vanpoucke, G., Kools, P., . . . Towbin, J. A. (2003). Assessment of the CTNNA3 gene encoding human αT-catenin regarding its involvement in dilated cardiomyopathy. *Human Genetics, 112*(3), 227-236. doi:10.1007/s00439-002-0857-5

Jiang, X., Zhu, Y., Liu, H., Chen, S., & Zhang, D. (2019). Effect of BIN1 on cardiac dysfunction and malignant arrhythmias. *Acta Physiologica, 228*(3). doi:10.1111/apha.13429

Kampinga, H. H., Hageman, J., Vos, M. J., Kubota, H., Tanguay, R. M., Bruford, E. A., . . . Hightower, L. E. (2009). Guidelines for the nomenclature of the human heat shock proteins. *Cell Stress Chaperones, 14*(1), 105-111.

Kärkkäinen, S., & Peuhkurinen, K. (2007). Genetics of dilated cardiomyopathy. *Annals of Medicine, 39*(2), 91-107. doi:10.1080/07853890601145821

Khand, A., Gemmell, I., Clark, A. L., & Cleland, J. G. (2000). Is the prognosis of heart failure improving? *J Am Coll Cardiol, 36*(7), 2284-2286. doi:10.1016/s1388-9842(99)00032-x

Klarlund, J. K., Holik, J., Chawla, A., Park, J. G., Buxton, J., & Czech, M. P. (2001). Signaling Complexes of the FERM Domain-containing Protein GRSP1 Bound to ARF Exchange Factor GRP1. *Journal of Biological Chemistry, 276*(43), 40065-40070. doi:10.1074/jbc.m105260200

Krief, S., Faivre, J. F., Robert, P., Le Douarin, B., Brument-Larignon, N., Lefrere, I., . . . Bril, A. (1999). Identification and characterization of cvHsp. A novel human small stress protein selectively expressed in cardiovascular and insulin-sensitive tissues. *J Biol Chem, 274*(51), 36592-36600.

Lakdawala, N. K., Winterfield, J. R., & Funke, B. H. (2013). Dilated Cardiomyopathy. *Circulation: Arrhythmia and Electrophysiology, 6*(1), 228-237. doi:10.1161/circep.111.962050

Larson, M. G., Atwood, L. D., Benjamin, E. J., Cupples, L. A., D'agostino, R. B., Fox, C. S., . . . Levy, D. (2007). Framingham Heart Study 100K project: Genome-wide associations for cardiovascular disease outcomes. *BMC Medical Genetics, 8*(S1). doi:10.1186/1471-2350-8-s1-s5

Lee, D. S., Pencina, M. J., Benjamin, E. J., Wang, T. J., Levy, D., O'donnell, C. J., . . . Vasan, R. S. (2006). Association of Parental Heart Failure with Risk of Heart Failure in Offspring. *New England Journal of Medicine, 355*(2), 138-147. doi:10.1056/nejmoa052948

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., & Fennell, T. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature, 536*(7616), 285-291.

Lettre, G. (2014). Rare and low-frequency variants in human common diseases and other complex traits. *Journal of Medical Genetics, 51*(11), 705-714. doi:10.1136/jmedgenet-2014-102437

Liggett, S. B., Cresci, S., Kelly, R. J., Syed, F. M., Matkovich, S. J., Hahn, H. S., . . . Ii, G. W. (2008). A GRK5 polymorphism that inhibits β-adrenergic receptor signaling is protective in heart failure. *Nature Medicine, 14*(5), 510-517. doi:10.1038/nm1750

Lindgren, M. P., Smith, J. G., Li, X., Sundquist, J., Sundquist, K., & Zöller, B. (2016). Sibling risk of hospitalization for heart failure – A nationwide study. *International Journal of Cardiology, 223*, 379-384. doi:10.1016/j.ijcard.2016.08.067

Matkovich, S. J., Booven, D. J., Hindes, A., Kang, M. Y., Druley, T. E., Vallania, F. L., . . . Dorn, G. W. (2010). Cardiac signaling genes exhibit unexpected sequence

diversity in sporadic cardiomyopathy, revealing HSPB7 polymorphisms associated with disease. *Journal of Clinical Investigation, 120*(1), 280-289. doi:10.1172/jci39085

Morita, H., Seidman, J., & Seidman, C. E. (2005). Genetic causes of human heart failure. *The Journal of Clinical Investigation, 115*(3), 518-526. doi:10.1172/JCI200524351

Nadar, S., Prasad, N., Taylor, R. S., & Lip, G. Y. (2005). Positive pressure ventilation in the management of acute and chronic cardiac failure: A systematic review and meta-analysis. *International Journal of Cardiology, 99*(2), 171-185. doi:10.1016/j.ijcard.2004.03.047

Norland, K., Sveinbjornsson, G., Thorolfsdottir, R. B., Davidsson, O. B., Tragante, V., Rajamani, S., . . . Stefansson, K. (2019). Sequence variants with large effects on cardiac electrophysiology and disease. *Nature Communications, 10*(1). doi:10.1038/s41467-019-12682-9

Pastorek, J., Pastorekova, S., & Zatovicova, M. (2008). Cancer-Associated Carbonic Anhydrases and Their Inhibition. *Current Pharmaceutical Design, 14*(7), 685-698. doi:10.2174/138161208783877893

Pillai, H. S., & Ganapathi, S. (2013). Heart Failure in South Asia. *Current Cardiology Reviews, 9*(2), 102-111. doi:10.2174/1573403x11309020003

Qian, Y., Xiao, D., Guo, X., Chen, H., Hao, L., Ma, X., . . . Wang, H. (2017). Multiple gene variations contributed to congenital heart disease via GATA family transcriptional regulation. *Journal of Translational Medicine, 15*(1). doi:10.1186/s12967-017-1173-0

Rehman, S., Ahmad, T. M., Hayat, A., & Tahir, S. (2019). Low-density lipoprotein receptor gene mutation at Exon 2 and 4 in premature coronary artery disease in our population. *Pakistan Journal of Medical Sciences, 35*(4). doi:10.12669/pjms.35.4.1308

Reinkober, J., Tscheschner, H., Pleger, S. T., Most, P., Katus, H. A., Koch, W. J., & Raake, P. W. (2012). Targeting GRK2 by gene therapy for heart failure: Benefits above β-blockade. *Gene Therapy, 19*(6), 686-693. doi:10.1038/gt.2012.9

Scott Binder, M., Brown, E., Aversano, T., Wagner, K. R., Calkins, H., & Barth, A. S. (2020). Novel FHL1 Mutation Associated With Hypertrophic Cardiomyopathy, Sudden Cardiac Death, and Myopathy. *JACC: Case Reports, 2*(3), 372-377. doi:10.1016/j.jaccas.2019.11.075

Shah, A., Gandhi, D., Srivastava, S., Shah, K. J., & Mansukhani, R. (2017). Heart Failure: A Class Review of Pharmacotherapy. *P&T, 42*(7), 464-472.

Shahid, S. U., Shabana, N. A., Rehman, A., & Humphries, S. (2018). GWAS implicated risk variants in different genes contribute additively to increase the risk of coronary artery disease (CAD) in the Pakistani subjects. *Lipids in Health and Disease, 17*(1). doi:10.1186/s12944-018-0736-2

Shakeel, M., Irfan, M., & Khan, I. A. (2018). Estimating the mutational load for cardiovascular diseases in Pakistani population. *Plos One, 13*(2). doi:10.1371/journal.pone.0192446

Shakeel, M., Irfan, M., & Khan, I. A. (2018). Rare genetic mutations in Pakistani patients with dilated cardiomyopathy. *Gene, 673*, 134-139. doi:10.1016/j.gene.2018.06.019

Sheikh, S. A. (2006). HEART FAILURE IN PAKISTAN: A Demographic Survey. *Journal of Cardiac Failure, 12*(8). doi:10.1016/j.cardfail.2006.08.069

Skrzynia, C., Berg, J., Willis, M., & Jensen, B. (2014). Genetics and Heart Failure: A Concise Guide for the Clinician. *Current Cardiology Reviews, 11*(1), 10-17. doi:10.2174/1573403x09666131117170446

Smith, E. D., Lakdawala, N. K., Papoutsidakis, N., Aubert, G., Mazzanti, A., Mccanta, A. C., . . . Helms, A. S. (2020). Desmoplakin Cardiomyopathy, a Fibrotic and Inflammatory Form of Cardiomyopathy Distinct From Typical Dilated or Arrhythmogenic Right Ventricular Cardiomyopathy. *Circulation, 141*(23), 1872-1884. doi:10.1161/circulationaha.119.044934

Smith, N. L., Felix, J. F., Morrison, A. C., Demissie, S., Glazer, N. L., Loehr, L. R., . . . Vasan, R. S. (2010). Association of Genome-Wide Variation With the Risk of Incident Heart Failure in Adults of European and African Ancestry. *Circulation: Cardiovascular Genetics, 3*(3), 256-266. doi:10.1161/circgenetics.109.895763

Sommariva, E., Stadiotti, I., Perrucci, G. L., Tondo, C., & Pompilio, G. (2017). Cell models of arrhythmogenic cardiomyopathy: Advances and opportunities. *Disease Models & Mechanisms, 10*(7), 823-835. doi:10.1242/dmm.029363

Syrris, P., Ward, D., Asimaki, A., Evans, A., Sen-Chowdhry, S., Hughes, S. E., & Mckenna, W. J. (2006). Desmoglein-2 mutations in arrhythmogenic right ventricular cardiomyopathy: A genotype-phenotype characterization of familial disease. *European Heart Journal, 28*(5), 581-588. doi:10.1093/eurheartj/ehl380

Tavira, B., Gómez, J., Ortega, F., Tranche, S., Díaz-Corte, C., Alvarez, F., . . . Coto, E. (2013). A CLCNKA polymorphism (rs10927887; p.Arg83Gly) previously linked to heart failure is associated with the estimated glomerular filtration rate in the RENASTUR cohort. *Gene, 527*(2), 670-672. doi:10.1016/j.gene.2013.06.055

Tayal, U., Prasad, S., & Cook, S. A. (2017). Genetics and genomics of dilated cardiomyopathy and systolic heart failure. *Genome Medicine, 9*(1). doi:10.1186/s13073-017-0410-8

The World Bank. (2016). The World Bank data for world populations. Accessed from http://databank.worldbank.org/data/ download/POP.pdf

Wang, D., Lyu, J., Fang, J., Chen, J., Chen, W., Huang, J., . . . Guo, X. (2019). Impact of LDB3 gene polymorphisms on clinical presentation and implantable cardioverter defibrillator (ICD) implantation in Chinese patients with idiopathic dilated cardiomyopathy. *Journal of Zhejiang University-SCIENCE B, 20*(9), 766-775. doi:10.1631/jzus.b1900017

Watkins, H., Ashrafian, H., & Redwood, C. (2011). Inherited cardiomyopathies. *N Engl J Med, 364*(17), 1643-1656.

Woodall, M. C., Ciccarelli, M., Woodall, B. P., & Koch, W. J. (2014). G Protein–Coupled Receptor Kinase 2. *Circulation Research, 114*(10), 1661-1670. doi:10.1161/circresaha.114.300513

Zhou, K., & Hong, T. (2016). Cardiac BIN1 (cBIN1) is a regulator of cardiac contractile function and an emerging biomarker of heart muscle health. *Science China Life Sciences, 60*(3), 257-263. doi:10.1007/s11427-016-0249-x