ARC Journal of Forensic Science
Volume 7, Issue 1, 2023, PP 1-15
ISSN No. (Online) 2456-0049
https://doi.org/10.20431/2456-0049.0701001
www.arcjournals.org

In *Silico* Functional Analysis of Variants in Genes Associated with Sudden Cardiac Death in Cases of the National Institute of Legal Medicine of Colombia

Joseph Alape Ariza[1,2,4*], Andrea Pinzón Reyes [2], Arbey Hernán Medina Rocha[3], Rodrigo Cabrera Perez[4.5], Clara Isabel Bermúdez-Santana[6,7]

[1]Forensic Sciences Research Group, National Institute of Legal Medicine and Forensic Sciences, D.C. Colombia.
[2]Forensic Genetics Research Group, National Institute of Legal Medicine and Forensic Sciences, D.C. Colombia.
[3]Forensic Pathology Group, Bogotá Regional Directorate, National Institute of Legal Medicine and Forensic Sciences, D.C. Colombia.
[4]InstituteofCardiology, Fundación Cardio Infantil, D.C. Colombia
[5]Institute of Translational Medicine (IMT), Rosario University, Colombia
[6]Research Group in Rnomics, Department of Biology, Faculty of Sciences, National University of Colombia, Bogotá, D.C. Colombia.
[7]Center of Excellence in Scientific Computing, National University of Colombia, Colombia

Corresponding Author:** ***Joseph Alape Ariza, *Forensic Sciences Research Group, National Institute of Legal Medicine and Forensic Sciences, D.C. Colombia. https://orcid.org/0000-0002-5764-4873*

Abstract: *Diseases of cardiovascular origin such as cardiomiopathies and channelopathíes are the main causes of sudden cardiac death and in some cases are difficult to diagnose during autoposy. We present an in silico analysis using bioinformatic tools, for the analysis of possibly pathogenic variants in genes associated with sudden cardiac death. Algorithms were used to predict pathogenicity, predicting the impact of SNPs on proteins, genomic specificity andprotein-protein interaction. We found that variants in the KCNH2, ANK3, TTN, CAV3 and DSP genes cause structural alterations, molecular, cellular and interstitial changes in the heart that can trigger sudden death.*

Keywords: *NGS, sudden cardiac death, MutPred2, ShinyGO, protein analysis*

1. INTRODUCTION

Cardiovascular diseases are one of the leading causes of sudden cardiac death (SCD) [1]. Cardiomyopathies and channelopathies are the most frequent, with a hereditary component and difficulty in their detection during medico-legal autopsies. Hereditary cardiomyopathies such as hypertrophic cardiomyopathy, dilated cardiomyopathy and arrhythmogenic cardiomyopathy are difficult to diagnose as they present minimal structural changes in the heart [2,3]. Channelopathies are not associated with anatomical changes, but affect heart rhythm and cardiac electrical conduction triggering sudden cardiac arrest, which also hinder their postmortem diagnosis [2,4].

In recent years with the development of massive parallel sequencing or next-generation sequencing (NGS), which allows sequencing the genome and the complete exome of a patient, to analyze the totality of the genes of an individual. This methodology has made it possible to identify a large number of genetic variants, however, the validation of all of them through electrophysiological functional studies can be costly and time-consuming. One possible strategy to overcome this challenge is to evaluate each potentially pathogenic variant using the bioinformatics approach [5,6].

Different prediction algorithms and statistical models have been developed to explore the association between genetic variants and diseases, considering protein stability, sequence conservation, physical and chemical properties, and structural variations [7,8]. Structural analysis helps to develop hypotheses about the possible impacts of substitutions and their possible links to disease states. Amino acid changes can affect both the normal

function of a protein by changing the hydrogen bond network, pH dependence, disturbing ligand binding, and conformational dynamics, transductional modification [9,10].

Different studies have considered structural variations of proteins and free energy changes when evaluating deregulatory effects [11,12,13],which have demonstrated the effectiveness of predicting the possible influence of amino acid changes on proteins [14].

Considering a group of genes from an NGS study of cases of indeterminate deaths, possibly pathogenic variants were recovered and using structural and functional analysis in silico aimed to describe possibly pathogenic variants associated with sudden cardiac death, to improve predictions of pathogenicity.

2. METHODOLOGY

We selected 68 cases of deaths to be determined, without macro and microscopic lesions during autopsy, with negative toxicology and virology.

Whole blood DNA extraction was performed using the QIAamp®DNA Blood Midi/Maxi kit, following the manufacturer's recomme-ndations. The quantification of DNA and libraries was performed using the Quantitating dsDNA kit using the Quantus™ Fluorometer instrument and following the manufacturer's recommendations. 2x150 bp paired-end NGS sequencing with MiSeq equipment (Illumina Inc., San Diego, CA. USA) using MiSeq Reagent V3 (150 clycles), according to the protocol of the commercial house.

2.1. Data Analysis

The alinators Bowtie2, BWA-MEM, and NovoAlign was aligned to the reference human genome sequence (GRCh37) were used, to identify the variaantes genomic analysis tools HaplotypeCaller (GATK-HC), Samtools mpileup and Freebayes were used. The variants were annotated by SnpEff and ANNOVAR, including predictive algorithms *in silico* SIFT, PolyPhen-2, MutationTaster, LRT, Mutation Assesor, FATHMM (Functi- onal Analysis Through Hidden Markov Models), MetaSVM, RadialSVM, LR, CADD, GERP++, phyloP, and SiPhy, to evaluate pathogenicity. Variant filtering by determining minor allele frequency (MAF) and selecting variants for in silico analysis was performed considering the recommendations of the American College of Medical Genetics and Genomics (ACMG) [15,16,17,18].

2.2. Predicting SNP Impact on Protein Stability

MutPred2: Machine learning-based method that integrates genetic and molecular data to probabilistically infer the pathogenicity of amino acid substitutions and their molecular mechanisms, by general pathogenicity prediction and a ranked list of specific molecular alterations that may affect the phenotype [19].

2.3. Genomic Specificity Analysis by ShinyGO

ShinyGO analyzes the genomic specificity of genes compared to the whole genome. Four aspects are compared: number of exons, number of transcription isoforms per gene, genome extension and length of 3′-UTR (untranslated region) [20]. Novel features of ShinyGO include graphical display of enrichment results and gene characteristics, and access to the application program interface (API) to KEGG and STRING for retrieval of pathway diagrams and protein-protein interaction networks. Shiny GO is a graphical and intuitive web application that can help researchers gain actionable insights from gene lists. Availability: http://ge-lab.org/go/. Gene Ontology (GO) is a unique database that describes the characteristics and cellular localization of each gene [21]. KEGG is a database containing a large number of known metabolic pathways of genes [22].

2.4. Analysis of Protein-protein Interaction and Functional Networks

Additionally, an analysis of protein-protein interaction and functional networks was performed to investigate the direct physical and functional relationships between identified genes, with the database (STRING) (http:// string.embl.de) [23]. The STRING database provided a score for each gene-gene interaction, calculated as the joint probability of the probabilities of the different evidence channels (protein interaction, fusion, co-expression, etc.), an approximate functional network was constructed on the basis of the expression profile of the proteins identified in the present study.

3. Results and Discussion

With the pathogenicity prediction algorithms of the 72 variants detected, in 41.66% it was not possible to obtain pathogenicity prediction results by any of the prediction algorithms used. Analysis with *SIFT* (26) variants was found to be harmful; with Polyphen2 (28) harmful and (8) probably harmful; for Mutation Taster (28) harmful. With reference to the other prediction algorithms there is a discrepancy in the prediction of pathogenicity. For sequence conservation prediction algorithms, it was found that for *GEPP++RS* with cut-off point greater than 4.4, it was found that in 14 of the variants there is no sequence conservation, for CADD with a cut-off point of 20, 12 harmful variants were found.

Relevant protein variants that had a defined crystallography structure were selected for further *in silico* studies that would allow us to understand the molecular mechanism and better classify whether the variant may actually become pathogenic.

3.1. Predicting the Impact of SNPs on Protein Stability

Protein Stability Prediction Analysis with MUPred2

MutPred2, allows to know molecular mechanisms that can be altered by a variant in a protein, Table 1.

Table1. *Prediction of protein stability with MutPred2 for variants of KCNH2, ANK3, TTN, CTDSP2,CAV3 and DSP genes.*

ID	**Substitution**	**MutPred2 score**	**Molecular mechanisms with P-values <= 0.05**	**Probability**	**P-value**	**Affected PROSITE and ELM Motifs**
NP_742053.1 _KCNH2	R92P	0.812	Loss of Helix	0.30	5.3e-03	ELME000002, ELME000102, ELME000233, PS00007
			Gain of Strand	0.30	2.7E-03	
			Loss of SUMOylation at K93	0.21	0.03	
ANK3	P1489S	0.594	Altered Ordered interface	0.25	0.02	ELME000052, ELME000053, ELME000063, ELME000106, ELME000173, ELME000336
			Gain of O-linked glycosylation at T1491	0.24	9.6E-03	
			Altered Transmembrane protein	0.15	0.01	
			Gain of Proteolytic cleavage at R1486	0.14	0.02	
			Loss of Sulfation at Y1494	0.01	0.05	
TTN	P16475Q	0.755	Altered Transmembrane protein	0.30	1.4e-04	ELME000052, ELME000062, ELME000117, ELME000136, ELME000159, ELME000202
			Altered Ordered interface	0.28	0.04	
			Gain of Strand	0.26	0.04	
			Loss of ADP-ribosylation at R16472	0.23	0.02	
TTN	P1698L	0.845	Altered Transmembrane protein	0.20	5.5E-03	None
			Altered Metal binding	0.19	0.02	

TTN	P16475Q	0.755	Altered Stability	0.28	6.6e-03	ELME000052, ELME000063, ELME000070, ELME000182, PS00001
			Altered Transmembrane protein	0.16	0.01	
			Loss of Disulfide linkage at C24664	0.12	0.04	
			Gain of N-linked glycosylation at N24658	0.06	0.02	
TTN	P1698L	0.845	Altered Transmembrane protein	0.20	5.5E-03	None
			Altered Metal binding	0.19	0.02	
TTN	I24660T	0.707	Altered Stability	0.28	6.6e-03	ELME000052, ELME000063, ELME000070, ELME000182, PS00001
			Altered Transmembrane protein	0.16	0.01	
			Loss of Disulfide linkage at C24664	0.12	0.04	
			Gain of N-linked glycosylation at N24658	0.06	0.02	
TTN	P22367L	0.723	Gain of Phosphorylation at Y22368	0.26	0.03	ELME000080
CTDSP2	I106T	0.823	Altered Stability	0.63	8.1e-04	ELME000220, ELME000333, PS00006
			Altered Metal binding	0.29	3.4e-03	
			Loss of Catalytic site at E110	0.19	0.01	
CAV3	Y62C	0.702	Altered Ordered interface	0.35	2.5E-03	ELME000052, ELME000053, ELME000063, ELME000080, ELME000120, ELME000182
			Gain of Helix	0.28	0.02	
			Altered Transmembrane protein	0.20	5.3e-03	
DSP	N4K	0.586	Altered Ordered interface	0.16	0.05	ELME000285, PS00005
			Loss of N-terminal acetylation of M1	0.02	6.2e-03	

Cardiovascular remodeling is defined as a set of molecular, cellular and interstitial changes that occur in the heart and vessels due to different injuries. Changes in size, geometry and function are the key events that occur in the heart. Pathophysiology includes cell death, changes in energy metabolism, inflammation, oxidative stress, alteration in the extracellular matrix, neurohormonal activation, and changes in ion transport [24].

Since the heart undergoes different physiological stimuli, its pathological adaptation can lead to cardiomyopathies, cardiac dysfunction and ultimately heart

failure [25,26]. SUMOylation modifies protein-protein interactions, enzyme activity, or chromatin binding in a multitude of key cellular processes, acting as a highly dynamic molecular switch [27,28,29]. In the present study was found loss of SUMOilación in the protein gene KCNH2, which suggests that this protein, along with loss of the helix and gain in the strand lose its functional stability having a high probability of incidence in the presence of LQTS2.

Mutations in the KCNH2 gene cause LQTS2, is the second most common cause of congenital LQTS and responsible for 35% to 45% of all genotyped LQTS [30], and that 25% of cases of LQTS are not diagnosed, which makes identification in family groups difficult and the determinants of variability in disease severity are still largely unknown, because some members despite having the mutation are asymptomatic [31].

In the heart,O-linked glycosylation is recognized as an important mechanism involved in the regulation of many cellular processes, including cellular metabolism, mitochondrial function, quality control and protein turnover, autophagy, and calcium management [13,32]. Glycation can alter the function and stability of proteins and induce the synthesis of pathogenic molecules that favor the appearance and progression of different diseases, including cardiovascular diseases [33,34].

In the present study it was found that the variant in the ANK3 protein produces an O-linked glycosylation gain, alters the transmembrane protein, produces loss of sulfation in tyrosine and alters the interphase order, which has a high probability of having harmful effects on the regulation of the heart.

The role of O-linked glycosylation in the regulation of cardiovascular function is complex and that, like most studies, focuses on its role in cardiovascular pathophysiology. On the other hand, as we begin to understand more about the cellular functions regulated by O-linked glycosylation protein, it is becoming increasingly clear that a more accurate concept would be that O-linked glycosylation modification of cardiovascular proteins is a dynamic process that is critical to maintaining normal cardiomyocyte function [35,36].

Likewise, in the TTN, CDSP2 and CAV3 proteins, it is also established loss and or changes of the molecular mechanisms, which practically remodel the functioning of the heart, which implies that these variants cause loss in the stability of the protein, which can be the cause of the presence of arrhythmogenic diseases.

N-terminal acetylation is a post-translational modification carried out by N-terminal acetyltransferases (NAT) in nascent protein chains during translation, involved in the maturation of post-translational proteins, physicochemically affecting the N-terminal limb of most proteins [37]. The loss of N-terminal acetylation, and its previously unanticipated role in protein biogenesis, globally remodels the proteome to create a unique phenotype [38]. Therefore, pathogenic variants in DSP with loss of NAT could lead to abnormal contractile function of cardiac cells, additional because this protein is responsible for binding the cardiac desmosome with intermediate filaments. Alterations in this biosynthesis could trigger arrhythmogenic cardiomyopathy that is estimated to affect 0.02% to 0.1% of the population with an increased risk of sudden cardiac death and heart failure [39,40].

3.2. GO Enrichment Analysis

Analysis of GO enrichment of AKAP9, ANK2, ANK3, ANKRD1, CACNA1C, CACNB2, CASQ2, CAV3, CTNNA3, DSG2, DSP, KCND3, KCNE3, KCNH2, KCNJ5, KCNQ1, MYBPC3, NOS1AP, PRDM16, RYR2, SCN10A, SCN4B, SCN5A, SLMAP, SNTA1, TMEM43, TNNT2, TPM1, TRDN, TRPMP4 and TTN, using ShinyGO.

It has been reported that the higher the hierarchical level of GO terms in the tree structure, the more explicit the demonstrated biological function (Jain S, Bader GD.2010). Therefore, only bio-enriched GO terms with a false discovery rate (FDR) value less than 0.05 remained as remarkably enriched terms. Pathways with an FDR value less than 0.05 and containing at least 5 genes were considered significantly enriched.

The pathways with greater biological enrichment highlighted in Table 2, nine pathways were found related to cardiac conduction, action potential and membrane repolarization, caused by defects in the genes encoding ion channels such as potassium, sodium, and calcium or associated proteins that alter the generation and transmission of the action potential that predispose to fatal arrhythmias and sudden cardiacdeath , such as the DSG2, KCNH2, DSP, KCNJ5, AKAP9,

ANK2, CACNB2, KCND3, KCNE3, SCN4B, CTNNA3, and SCN5A genes involved in cardiac channelopathies [41,42].

Table2. *ShinyGO biological enrichment*

Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathway	Genes
1256,943587	5	11	3.70065E+14	AV node cell action potential	CACNB2 SCN4B SCN5A SCN10A RYR2
1256,943587	5	11	3.70065E+14	AV node cell to bundle of His cell signaling	CACNB2 SCN4B SCN5A SCN10A RYR2
2.04292E-13	11	25	3.58223E+14	Ventricular cardiac muscle cell membrane repolarization	KCNH2 SNTA1 KCNJ5 AKAP9 ANK2 KCND3 KCNE3 SCN4B CAV3 SCN5A NOS1AP
1.5183E-06	9	21	3.48918E+14	Reg. of ventricular cardiac muscle cell membrane repolarization	KCNH2 SNTA1 AKAP9 ANK2 KCNE3 SCN4B CAV3 SCN5A NOS1AP
2.83816E-15	13	34	3.1129E+14	Ventricular cardiac muscle cell action potential	DSG2 KCNH2 DSP SNTA1 KCNJ5 ANK2 KCND3 KCNE3 CAV3 CTNNA3 SCN5A RYR2 NOS1AP
4969,229302	5	14	2.90765E+14	Membrane repolarization during ventricular cardiac muscle cell action potential	KCNH2 KCNJ5 KCND3 KCNE3 NOS1AP
7357,844876	5	15	2.71381E+14	Bundle of His cell to Purkinje myocyte communication	DSG2 DSP CTNNA3 SCN5A SCN10A
0,612051969	7	22	2.59045E+14	Membrane depolarization during cardiac muscle cell action potential	ANK2 ANK3 SLMAP CACNB2 SCN4B CAV3 SCN5A

131619,8334	4	13	2.50505E+14	Reg. of ventricular cardiac muscle cell action potential	DSG2 DSP CTNNA3 RYR2
0,008218532	8	27	2.41228E+14	Reg. of cardiac muscle cell action potential	DSG2 DSP AKAP9 ANK2 CAV3 CTNNA3 RYR2 NOS1AP
5.78813E-07	10	34	2.39454E+14	Reg. of membrane repolarization	KCNH2 SNTA1 CASQ2 AKAP9 ANK2 KCNE3 SCN4B CAV3 SCN5A NOS1AP
2.56466E-11	12	41	2.38286E+14	Reg. of heart rate by cardiac conduction	DSG2 KCNH2 DSP KCNJ5 AKAP9 ANK2 CACNB2 KCND3 KCNE3 SCN4B CTNNA3 SCN5A
6.60817E-18	15	52	2.34849E+14	Cardiac muscle cell action potential involved in contraction	DSG2 KCNH2 DSP SNTA1 KCNJ5 ANK2 CACNB2 KCND3 KCNE3 SCN4B CAV3 CTNNA3 SCN5A RYR2 NOS1AP
190236387,4	3	11	2.22039E+14	SA node cell to atrial cardiac muscle cell communication	ANK2 SCN5A RYR2
8.94148E-11	12	45	2.17105E+14	Membrane repolarization	KCNH2 SNTA1 CASQ2 KCNJ5 AKAP9 ANK2 KCND3 KCNE3 SCN4B CAV3 SCN5A NOS1AP

26186,2998	5	19	2.14248E+14	Atrial cardiac muscle cell action potential	KCNJ5 ANK2 CACNB2 SCN5A RYR2
26186,2998	5	19	2.14248E+14	Atrial cardiac muscle cell to AV node cell signaling	KCNJ5 ANK2 CACNB2 SCN5A RYR2
1.45932E-25	19	74	2.09037E+13	Cardiac muscle cell action potential	DSG2 KCNH2 DSP SNTA1 KCNJ5 AKAP9 ANK2 ANK3 SLMAP CACNB2 KCND3 KCNE3 SCN4B CAV3 CTNNA3 SCN5A SCN10A RYR2 NOS1AP
2.54198E-07	10	39	2.08755E+14	Reg. of actin filament-based movement	DSG2 DSP TNNT2 AKAP9 MYBPC3 ANK2 CAV3 CTNNA3 SCN5A RYR2
3.54913E-05	9	36	2.03536E+14	Membrane depolarization during action potential	KCNH2 ANK2 ANK3 SLMAP CACNB2 SCN4B CAV3 SCN5A SCN10A

Table3. *ShinyGOMolecular Enrichment*

Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathway	Genes
2595314,67	4	12	2.71381E+14	Voltage-gated potassium channel activity involved in ventricular cardiac muscle	KCNH2 KCNJ5 KCND3 KCNE3
3914323,131	4	15	2.17105E+14	Nitric-oxide synthase binding	SNTA1 CAV3 SCN5A NOS1AP
3.63898E+13	2	11	1.48026E+14	Protein binding involved in heterotypic cell-cell adhesion	DSG2 DSP
5.02076E+13	2	13	1.25253E+14	Titin binding	MYBPC3 ANKRD1
2611015095	3	24	1.01768E+14	Voltage-gated sodium channel activity	SCN4B SCN5A SCN10A

1.30897E+14	2	22	7.4013E+13	Inward rectifier potassium channel activity	KCNH2 KCNJ5
1.30897E+14	2	22	7.4013E+13	Cytoskeletal anchor activity	ANK2 ANK3
2.92133E+14	1	11	7.4013E+13	High voltage-gated calcium channel activity	CACNB2
2.92133E+14	1	11	7.4013E+13	Structural molecule activity conferring elasticity	TTN
0,000865976	12	140	6.97837E+14	Transmembrane transporter binding	SNTA1 AKAP9 ANK2 ANK3 KCND3 KCNE3 SCN4B CAV3 SCN5A SCN10A TRDN RYR2
8709309595	3	38	6.42744E+14	Sodium channel regulator activity	SNTA1 SCN4B CAV3
1.71884E+14	2	26	6.26264E+14	Protein kinase A regulatory subunit binding	AKAP9 RYR2
3.36429E+14	1	13	6.26264E+14	Outward rectifier potassium channel activity	KCND3
1.95452E+14	2	28	5.81531E+14	Spectrin binding	ANK2 ANK3
3.49314E+14	1	14	5.81531E+14	Myosin heavy chain binding	MYBPC3
3.49314E+14	1	14	5.81531E+14	Protein kinase A catalytic subunit binding	RYR2
3.49314E+14	1	14	5.81531E+14	C3HC4-type RING finger domain binding	KCNH2
3.65446E+14	1	15	5.42762E+14	Histone methyltransferase activity (H3-K9 specific)	PRDM16
3.65446E+14	1	15	5.42762E+14	Muscle alpha-actinin binding	TTN
1.2821E+13	3	46	5.30963E+14	Sodium channel activity	SCN4B SCN5A SCN10A

In Table 3 molecular enrichment, genes are involved in the formation and regulation of different voltage-dependent ion channels, so any alteration in any of the pathways in which they interact will be pathogenic.

In Figure 1, using the chi-square test, the number of exons showed a p-value (0.0023) when comparing SDRs with other genes in the genome. In addition, the number of transcription isoforms per gene were significantly different from the expected value with a value of p = 0.051. These results indicated that our identified SDRs may have strong transcription characteristics with other genes, which may be involved in heart disease. For genome extension analysis, we observed an extremely low p-value (0.00096), while for the 3′-UTR length comparison, we observed a p-value (0.52) and for 5'-UTR p (0.76). There is a strong association between these genes and therefore the presence of possibly pathogenic variants presents an increased risk of death due to cardiac arrhythmias, cardiomyopathies and coronary heart disease. For example, about 50% of ARVC cases are associated with mutations in genes encoding desmosome and cell

adhesion proteins necessary for mechano-electrical coupling in the heart [43].

In Figure 2, you can see the great correlation that exists between the genes studied, that all of them are not only interacting, but participate in many of the biological processes that have to do with the heart. It has been previously reported that the higher the hierarchical level of GO terms in the tree structure, the more explicit the biological function [44].

The networks for the biological, cellular and molecular components were constructed Figure 3. The strong interaction between the genes in each of the components is observed.

For the cellular component it is observed that there are two clusters of interaction, one corresponding to the cellular processes of contact between sarcolemma, contraction fibers, it could be said that everything that has with the mechanism of cellular interaction at the physical level and in the other cluster the genes that have to do with transport through ion channels and their associated proteins. In the molecular component (c), it is observed in a large cluster where all interactions are included at the level ce ion channels during the action potential.

KEGG pathmaps are molecular interaction/ reaction network diagrams represented in terms of the KEGG orthology groups, Figure 4. It shows the enriched pathway diagrams and how the genes studied are interacting with the other molecules. Therefore, any variant in genes can cause the presence of the same or different phenotype.

In this study, it was found enriched ARVC, presents several signaling pathways that are involved in metabolism and cellular behavior presented an enrichment, including calcium signaling pathway, oxytocin signaling pathway, cyclic GMP signaling pathway (cGMP)-protein kinase G (pkg), cyclic AMP (cAMP) signaling pathway, HIF-1 (hypoxic stress response) signaling pathway, mitogen-activated protein kinase (MAPK) signaling pathway, which is consistent with previous studies suggesting that these pathways have important roles in SCD [45,46].

3.3. Protein-Protein Interaction Analysis

The STRING functional networks of the proteins TRDN, TTN, CAV3, DSP, DSG2, TRPM4, AKAP9, ANK3, ANK2, CACNA1C, MYBPC3, KCND3 and KCNH2, resulted in a functional network. What their proteins have more interactions with each other than would be expected from a random set of proteins of the same size and degree of distribution extracted from the genome. Such enrichment indicates that the proteins are at least partially biologically connected, as a group. Figure 5.

The network of these genes when grouped in this way indicates that their proteins have more interactions with each other than would be expected for a random set of proteins of the same size and grade distribution extracted from the genome. Such enrichment indicates that the proteins are at least partially biologically connected, as a group.

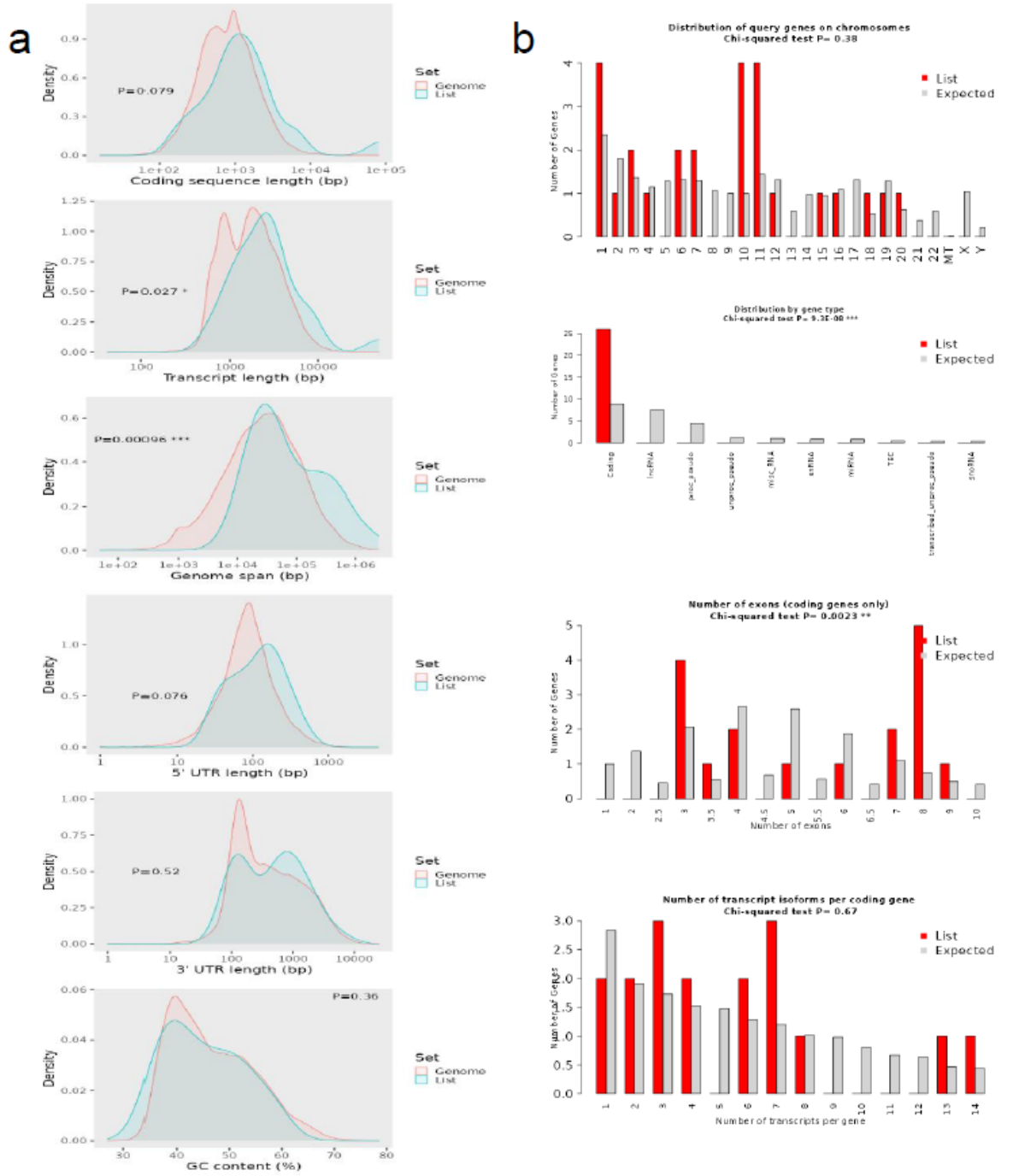


Figure1. *Comparison of AKAP9, ANK2, ANK3, ANKRD1, CACNA1C, CACNB2, CASQ2, CAV3, CTNNA3, DSG2, DSP, KCND3, KCNE3, KCNH2, KCNJ5, KCNQ1, MYBPC3, NOS1AP, PRDM16, RYR2, SCN10A, SCN4B, SCN5A, SLMAP, SNTA1, TMEM43, TNNT2, TPM1, TRDN, TRPMP4 and TTN genes with the rest of the genes in the genome. Chi-square and Student's t tests were run to see if the genes have special characteristics compared to all other genes: a) distribution of UTR lengths, b) distribution of the genes studied, compared to the genome.*

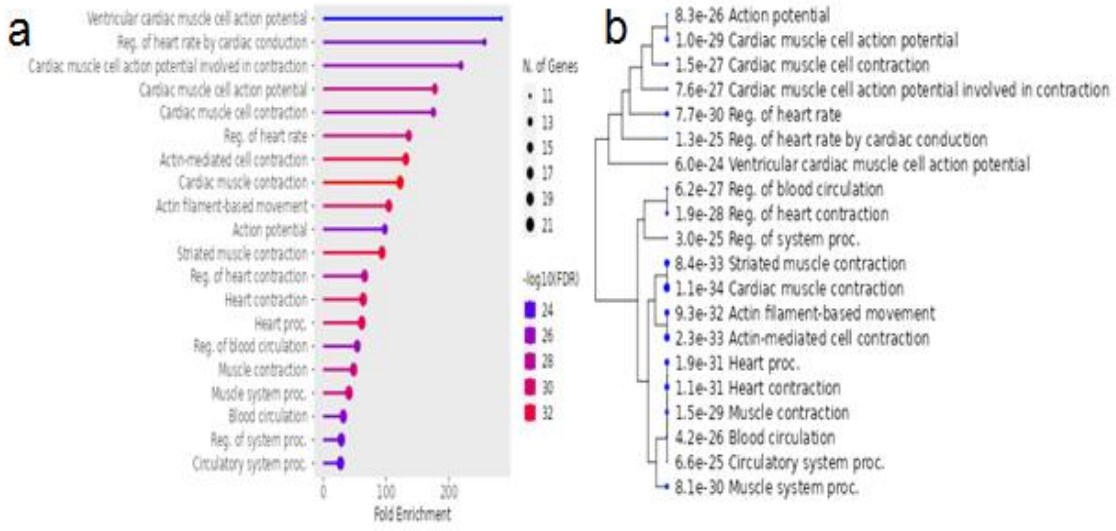
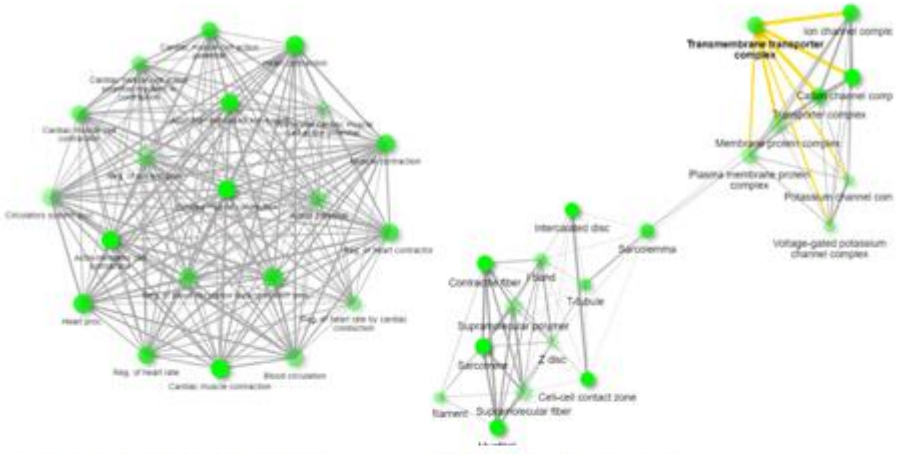


Figure2. *Enrichment of biological processes in terms of GO for AKAP9, ANK2, ANK3, ANKRD1, CACNA1C, CACNB2, CASQ2, CAV3, CTNNA3, DSG2, DSP, KCND3, KCNE3, KCNH2, KCNJ5, KCNQ1, MYBPC3, NOS1AP, PRDM16, RYR2, SCN10A, SCN4B, SCN5A, SLMAP, SNTA1, TMEM43, TNNT2, TPM1, TRDN, TRPMP4 and TTN ($p \leq 0.05$), a) the color gradient represents the adjusted values and the differences in the size of the bubbles correlate with the enrichment factor; b) hierarchical clustering tree, related GO terms are grouped according to the number of genes they share. Larger dots indicate more significant p-values.*



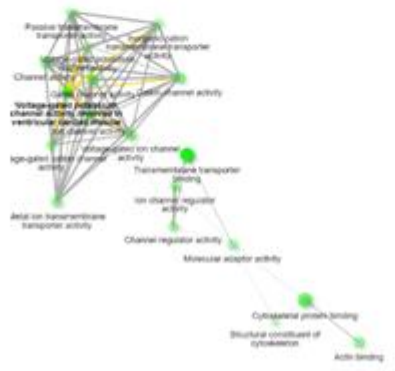


Figure3. *Interaction networks of AKAP9, ANK2, ANK3, ANKRD1, CACNA1C, CACNB2, CASQ2, CAV3, CTNNA3, DSG2, DSP, KCND3, KCNE3, KCNH2, KCNJ5, KCNQ1, MYBPC3, NOS1AP, PRDM16, RYR2, SCN10A, SCN4B, SCN5A, SLMAP, SNTA1, TMEM43, TNNT2, TPM1, TRDN, TRPMP4 and TTN.*

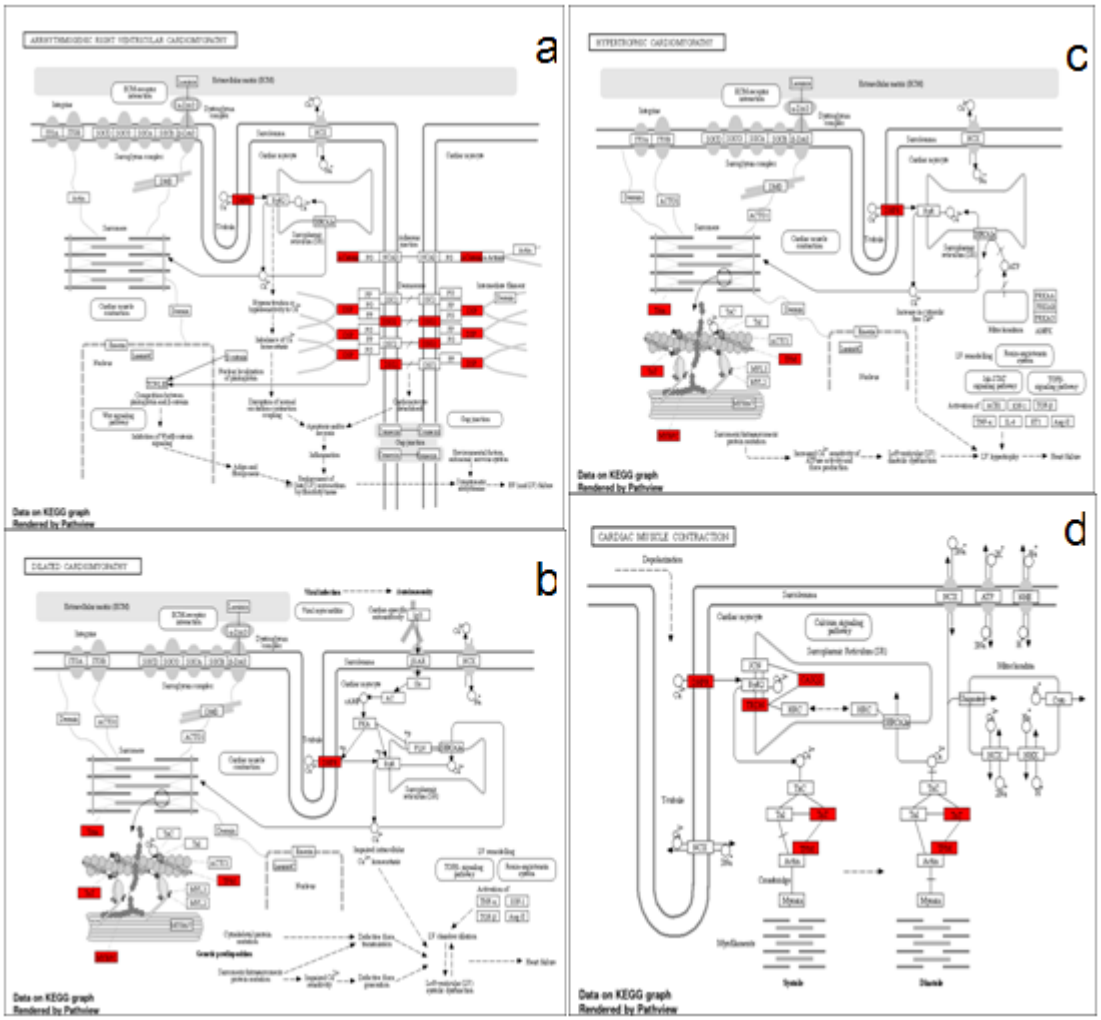


Figure4. *Significant KEGG pathway for a) arrhythmogenic right ventricular cardiomyopathy (ARVC), b) dilated cardiomyopathy (DM), c) hypertrophic cardiomyopathy (HCM) and d) cardiac muscle contraction. The genes of interest are highlighted in red.*

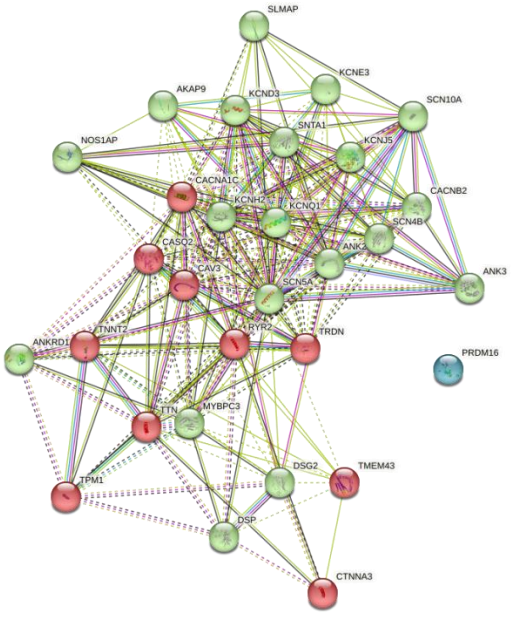


Figure5. *Functional network of TRDN, TTN, CAV3, DSP, DSG2, TRPM4, AKAP9, ANK3, ANK2, CACNA1C, MYBPC3, KCND3 and KCNH2 genes with three clusters: cluster 1 (red), cluster 2 (green) and cluster 3 (blue).*

4. CONCLUSIONS

Both pathogenicity prediction analysis and biological function enrichment analysis allow us to determine that the variants of the group of genes analyzed participate in several biological processes and that the dysfunction of these genes can lead to the appearance of cardiovascular diseases that can lead to sudden cardiac death.

It is of utmost importance when pathogenic variants are present, probably pathogenic and of uncertain significance, to use in *silico* functional analyzes, to have a broader vision that allows a more accurate diagnosis, given the genetic heterogeneity and phenotypic variability that occurs in this type of diseases.

The analysis of cosegregation with relatives in cases of sudden death, is very difficult, this gap can be filled with functional analysis *in silico*, pathogenicity prediction analysis, GO enrichment analysis and the construction of protein-protein interaction networks, since these allow to understand not only the contribution of genetic factors, but also the molecular mechanisms underlying sudden cardiac death.

5. LIMITATIONS OF THE STUDY

The inability to functionally test all candidate variants, with cosegregation studies with relatives, to corroborate the pathogenicity results performed *in silico*.

6. ETHICAL CONSIDERATIONS

It's the first research carried out in the country and that involved the use of postmortem samples for molecular analysis. The research required the use of bodily fluids which means that informed consent had to be given by family members, this was supplied first, through the approval of the ethics committee of the Faculty of Sciences of the National University of Colombia and secondly through Decree 0786 of 1990, with special emphasis on Chapter VI. Of the Viscerotomies and articles 18, 19, and 20.

7. DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

8. ACKNOWLEDGMENT

Thanks to the Faculty of Sciences at National University Of Colombia, Bogota Campus for supporting partialy this research under the project Hermes 49924 (2020).

REFERENCES

[1] Adabag AS, Luepker RV, Roger VL, Gersh BJ. (2010). Sudden cardiac death: epidemiology and risk factors. Nat Rev Cardiol 7: 216–225.

[2] Chugh SS, Reinier K, Teodorescu C, et al. (2008). Epidemiology of sudden cardiac death: Clinical and research implications. Prog Cardiovasc Dis. Vol. 51: 213-28.

[3] Kaltman JR, Thompson PD, Lantos J, et al., 2011. Screening for sudden cardiac death in the young: Report from a national heart, lung, and blood institute working group. Circulation. 123:1911-8.

[4] Tester David J., Argelia Medeiros-Domingo, Melissa L. Will, Carla M. Haglund, and Michael J. Ackerman, (2012). Cardiac Channel Molecular Autopsy: Insights From 173 Consecutive Cases of Autopsy-Negative Sudden Unexplained Death Referred for Postmortem Genetic Testing. Mayo Clin. Proc. 87(6): 524-539.

[5] Wei, Q., Wang, L., Wang, Q., Kruger, W. D. & Dunbrack, R. L., Jr Testing computational prediction of missense mutation phenotypes: functional characterization of 204 mutations of human cystathionine beta synthase. Proteins 78, 2058–74 (2010).

[6] Lincoln, S.E., Hambuch, T., Zook, J.M. et al. One in seven pathogenic variants can be challenging to detect by NGS: an analysis of 450,000 patients with implications for clinical sensitivity and genetic test implementation. Genet Med 23, 1673–1680 (2021). https://doi.org/ 10.1038/s41436-021-01187-w

[7] Schwartz, C. E. & Chen, C. F. Progress in detecting genetic alterations and their association with human disease. J Mol Biol 425, 3914–8 (2013).

[8] Ali MZ, Farid A, Ahmad S, Muzammal M, Mohaini MA, Alsalman AJ, Al Hawaj MA, Alhashem YN, Alsaleh AA, Almusalami EM, Maryam M, Khan MA. In Silico Analysis Identified Putative Pathogenic Missense nsSNPs in Human SLITRK1 Gene. Genes. 2022; 13(4):672. https://doi.org/10.3390/genes13040672

[9] Stefl, S., Nishi, H., Petukh, M., Panchenko, A. R. &Alexov, E. Molecular mechanisms of disease-causing missense mutations. J Mol Biol 425, 3919–36 (2013).

[10] Vymětal, J, Mertová, K, Boušová, K, Šulc, J, Tripsianes, K, Vondrasek, J. Fusion of two unrelated protein domains in a chimera protein

and its 3D prediction: Justification of the x-ray reference structures as a prediction benchmark. Proteins. 2022; 90(12): 2067- 20 79. doi:10.1002/prot.26398

[11] Zhang, Z., Teng, S., Wang, L., Schwartz, C. E. & Alexov, E. Computational analysis of missense mutations causing Snyder-Robinson syndrome. Hum Mutat 31, 1043–9 (2010).

[12] Zhang, Z. et al. A Y328C missense mutation in spermine synthase causes a mild form of Snyder-Robinson syndrome. Hum Mol Genet 22, 3789–97 (2013).

[13] Boccuto, L. et al. A mutation in a ganglioside biosynthetic enzyme, ST3GAL5, results in salt & pepper syndrome, a neurocutaneous disorder with altered glycolipid and glycoprotein glycosylation. Hum Mol Genet (2013).

[14] Ahsan, T., Sajib, A.A. Missense variants in the TNFA epitopes and their effects on interaction with therapeutic antibodies—in silico analysis. J Genet Eng Biotechnol 20, 7 (2022). https://doi.org/ 10.1186/s43141-021-00288-y

[15] Siepel A, Pollard K, Haussler D (2006) New methods for detecting lineage-specific selection. In: Proc. 10th Int'l Conf. on Research in Computational Molecular Biology. Berlin: Springer-Verlag, pp.190–205.

[16] Garber M, Guttman M, Clamp M, Zody MC, Friedman N, Xie X. (2009). Identifying novel constrained elements by exploiting biased substitution patterns. Bioinformatics. Jun 15;25(12):i54-62. doi:10.1093/bioinformati cs/btp190. PMID: 19478016; PMCID: PMC 2687944.

[17] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL (2015). ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. May;17(5):405-24. doi: 10.1038/ gim.2015.30. Epub 2015 Mar 5. PMID: 257 41868; PMCID: PMC4544753.

[18] Li Q, Wang K. (2017). InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. Am J Hum Genet. Feb 2;100(2):267-280. doi: 10.1016/j.ajhg.20 17.01.004. Epub 2017 Jan 26. PMID: 28132688; PMCID: PMC5294755.

[19] Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, Nam H, Mort M, Cooper DN, Sebat J, Iakoucheva LM, Mooney SD, Radivojac P. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nat. Commun. 11, 5918 (2020)

[20] Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics. 2020 Apr 15;36(8): 2628-2629. doi: 10.1093/bioinformatics/btz 931. PMID: 31882993; PMCID:PMC7178 415.

[21] Ashburner M. et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25, 25–29.

[22] Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017 Jan 4;45(D1):D353-D361. doi: 10.1093/nar/gkw1092. Epub 2016 Nov 28. PMID: 27899662; PMCID: PMC 5210567.

[23] Mering V. (2005). STRING: known and predicted protein-protein associations, integrated and transferred across organisms. Nucleic Acids Res. 33(Suppl. 1): D433-D7.

[24] Azevedo, P.S.; Polegato, B.F.; Minicucci, M.F.; Paiva, S.A.; Zornoff, L.A. Cardiac Remodeling: Concepts, Clinical Impact, Pathophysiological Mechanisms and Pharmacologic Treatment. Arq. Bras Cardiol. 2016, 106, 62–69.

[25] Christopher J. Oldfield, Todd A. Duhamel, and Naranjan S. Dhalla. 2019. Mechanisms for the transition from physiological to pathological cardiac hypertrophy. Canadian Journal of Physiology and Pharmacol ogy. 98(2): 74-84. https://doi.org/10.1139/ cjpp-2019-0566

[26] He, X., Du, T., Long, T. et al. Signaling cascades in the failing heart and emerging therapeutic strategies. Sig Transduct Target Ther 7, 134 (2022). https://doi.org/10.1038/ s41392-022-00972-6

[27] Heusch G. (2020). Myocardial Ischaemia-Reperfusion Injury and Cardioprotection in Perspective. Nat. Rev. Cardiol. 17, 773–789. 10.1038/s41569-020-0403-y

[28] Chen Y., Xu T., Li M., Li C., Ma Y., Chen G., et al. (2021). Inhibition of SENP2-Mediated Akt deSUMOylation Promotes Cardiac Regeneration via Activating Akt Pathway. Clin. Sci. (Lond) 135, 811–828. 10.1042/CS20201408

[29] Cai Z., Wang Z., Yuan R., Cui M., Lao Y., Wang Y., et al. (2021). Redox-sensitive Enzyme SENP3 Mediates Vascular Remodeling via De-SUMOylation of β-catenin and Regulation of its Stability. EBioMedicine 67, 103386.10.1016/ j.ebiom. 2021.103386.

[30] Duncan G, Firth K, George V, Hoang MD, Staniforth A, Smith G, Denning C. Drug-Mediated Shortening of Action Potentials in LQTS2 Human Induced Pluripotent Stem

Cell-Derived Cardiomyocytes. Stem Cells Dev. 2017 Dec 1;26(23):1695-1705. doi: 10.1089/scd.2017.0172. Epub 2017 Oct 9. PMID: 28992755; PMCID: PMC5706629.

[31] Wilde AAM, Amin AS, Postema PG Diagnosis, management and therapeutic strategies for congenital long QT syndrome Heart 2022;108:332-338.

[32] Erickson, J.R.; Pereira, L.; Wang, L.; Han, G.; Ferguson, A.; Dao, K.; Copeland, R.J.; Despa, F.; Hart, G.W.; Ripplinger, C.M.; et al. Diabetic hyperglycaemia activates CaMKII and arrhythmias by O-linked glycosylation. Nature 2013, 502, 372–376.

[33] Cho, S.J.; Roman, G.; Yeboah, F.; Konishi, Y. The road to advanced glycation end products: A mechanistic perspective. Curr. Med. Chem. 2007, 14, 1653–1671.

[34] Dozio, E.; Di Gaetano, N.; Findeisen, P.; Corsi Romanelli, M.M. Glycated albumin: From biochemistry and laboratory medicine to clinical practice. Endocrine 2017, 55, 682–690

[35] Wende AR. Post-translational modifications of the cardiac proteome in diabetes and heart failure. Proteomics Clinical Applications. 2016;10(1):25–38.

[36] Wang X, et al. O-GlcNAcase deficiency suppresses skeletal myogenesis and insulin sensitivity in mice through the modulation of mitochondrial homeostasis. Diabetologia. 2016;59(6):1287–1296

[37] Aksnes, H., Ree, R. y Arnesen, T. (2019). Functions cotraduccionales, post-translational and non-catalytic of N-terminal acetyltransferases. mole. Cell 73, 1097–1114. two: 10.1016/j.molcel.2019.02.007

[38] Nguyen KT, Lee CS, Mun SH, Truong NT, Park SK, Hwang CS. N-terminal acetylation and the N-end rule pathway control degradation of the lipid droplet protein PLIN2. J Biol Chem. 2019 Jan 4;294(1):379-388. doi: 10.1074/jbc. RA118.005556. Epub 2018 Nov 13. PMID: 30425097; PMCID: PMC6322874.

[39] Corrado D, Basso C, Judge DP. Arrhythmogenic Cardiomyopathy. Circ Res. 2017 Sep 15;121(7):784-802. doi: 10.1161/ CIRCRESAHA.117.309345.PMID:28912183

[40] Gerull B., Klaassen S., Brodehl A. (2019). "The genetic landscape of cardiomyopathies," in *Genetic Causes of Cardiac Disease*, eds Erdmann J., Moretti A. (Cham: Springer;), 45–91. 10.1007/978-3-030-27371-2_2.

[41] Liu C, Zhao Q, Su T, et al. Postmortem molecular analysis of KCNQ1, KCNH2, KCNE1 and KCNE2 genes in sudden unexplained nocturnal death syndrome in the Chinese Han population. Forensic Sci Int 2013;231:82-7.

[42] Martínez-Barrios, E.; Cesar, S.; Cruzalegui, J.; Hernandez, C.; Arbelo, E.; Fiol, V.; Brugada, J.; Brugada, R.; Campuzano, O.; Sarquella-Brugada, G. Clinical Genetics of Inherited Arrhythmogenic Disease in the Pediatric Population. Biomedicines 2022, 10, 106. https://doi.org/10.3390/bio medicines 10010106

[43] Towbin JA. Inherited cardiomyopathies. Circ J. 2014;78(10):2347-56. doi: 10.1253/circj.cj-14-0893. Epub 2014 Sep 2. PMID: 25186923; PMCID: PMC4467885.

[44] Jain S, Bader GD. An improved method for scoring protein-protein interactions using semantic similarity within the gene ontology. BMC Bioinformatics 2010;11:562.

[45] Yeung CY, Lam KS, Li SW, et al. Sudden cardiac death after myocardial infarction in type 2 diabetic patients with no residual myocardial ischemia. Diabetes Care 2012;35:2564-9.

[46] Ackerman M, Atkins DL, Triedman JK. Sudden cardiac death in the young. Circulation 2016;133:1006-26.

Citation: *Joseph Alape Ariza et al. "In Silico Functional Analysis of Variants in Genes Associated with Sudden Cardiac Death in Cases of the National Institute of Legal Medicine of Colombia" ARC Journal of Forensic Science, vol 7, no.1, 2023, pp. 1-15. DOI: https://doi.org/10.20431/2456-0049.0701001*