In the Cis-Element CACCC-box a Master Regulatory Element during Cardiovascular Disease? A Bioinformatics Approach from the Perspective of the Kruppel-like Family of Transcription Factors

Juan Andrés García-Loredo [1,2], Michelle G. Santoyo-Suarez [1], Oscar Rodríguez-Núñez [1], Diego Francisco Benitez Chao [1], Elsa N. Garza-Treviño [1], Patricio A. Zapata-Morin [2], Gerardo R. Padilla-Rivas [1] and Jose Francisco Islas [1,*]

[1] Departamento de Bioquímica y Medicina Molecular, Facultad de Medicina, Universidad Autónoma de Nuevo León, Monterrey, Nuevo León 64460, México. JAG-L andres.garcialrd@uanl.edu.mx; MGS-S s.antoyo@hotmail.es; OR-N Oscar.rodrigueznu@uanl.edu.mx; DFB-C diego.benitezch@uanl.edu.mx; ENG-T elsa.garzatr@uanl.edu.mx; GRP-R Gerardo.padillarv@uanl.edu.mx

[2] Laboratorio de Micología y Fitopatología, Facultad de Ciencias Biológicas, Universidad Autónoma de Nuevo León, San Nicolás de los Garza 66451, México. PAZ-M Patricio.zapatamor@uanl.edu.mx

*   Correspondence: JFI jislas.me0117@uanl.edu.mx

Abstract: The CACCC-box motif emerges as a pivotal cis-regulatory element implicated in diverse developmental processes and diseases, particularly cardiovascular diseases (CVDs). This study centers on the intricate interplay between the CACCC-box and its binding proteins such as: the Krüppel-like family (KLF) of transcription factors as primary effectors in the context of CVDs. Our analysis was through a bioinformatics approach, which reveal significant transcriptional activity among KLF subgroup 2, exhibiting the highest number of interactions focusing on the established roles: pluripotency, cancer, and cardiovascular development and diseases. Our analysis reveals KLF's interactions with GATA4, MEF2C, NKX2.5 and other ~ 90 potential genes that participate in the regulation of the hypertrophic environment (or CVD's Environment). Also, the GO analysis showed genes containing the motif CACCC were enriched for multiple CVD's, in combination with STRING analysis these results pointed to a link between KLFs and these diseases. The analysis further identifies other potential CACCC-box binding factors, such as SP family members, WT1, VEZF1, and -SALL4, which are implicated in cardiac contraction, remodeling, and inflammation processes.

Keywords: Krüppel-like factor; transcription factors; CVD's; CACCC-box; diabetes; SP1

1. Introduction

The Gene is the most basic unit of information; contained within itself are the blueprints required for the development of all the final bioactive products in the cell, whether it may be a miRNA or other small RNA, all the way to a messenger RNA leading to a protein. Besides these products, there is further information pertaining to the gene's own expression. Importantly, regulation is an integral part of expression, a process controlled through elements found in the gene, particularly sequences in the promoter region[1,2].

Gene expression is a process regulated by transcription factors in which there is activation or repression of the transcription activity through their binding to specific DNA motifs mediated by their DNA-binding domains[3]. DNA motifs can be seen as short conserved sequences (ranging from 2 to 20 bp) to which sets of transcription factors (or entire families) bind. One such example of a DNA motif is CACCC-box, which serves as a binding site for several transcription factors, including the

Krüppel-like factor (KLF) family[4], the Specificity Protein (SP) family [5], and the Wilms tumor gene (WT1)[6].

As a regulatory element, the DNA-motif CACCC-box, alongside other Cis-elements such as TATA and CAAT boxes, work together in promoters, such as seen in the cardiac B-type Natriuretic Peptide (BNP) gene, wherein, deletions result in reduced transcriptional activity in postnatal cardiomyocytes; an activity regulated by KLF13[7–10].

Out of the earlier mentioned transcription factor families which bind to the CACCC-box, much research has been devoted to the KLFs for their roles in diverse heart-related processes such as the maturation of cardiac myocytes, and their dysregulation is associated with several CVDs as exemplified by cardiomyopathies, infarctions, left ventricular hypertrophy and diabetic cardiomyopathies[4,11–14].. Particularly, CVDs continue as the leading cause of death worldwide. According to a recent publication by the American College of Cardiology "CVDs are a persistent challenge that led to an enormous number of premature and preventable deaths", specifically ischemic heart disease ranks #1 in mortality leading to 108 deaths per 100,000, overall CVD-related deaths accounted for 19.8 million in 2022 [15]. As stated, KLFs have been shown to play a crucial role in the progression and control of CVDs[4]. Some examples of regulating CVDs by KLFs can be seen in atherosclerosis, as the expression of KLF-5 switches a proliferative phenotype in vascular smooth muscle cells, while the repression of KLF-2 leads to an inflammatory state and vasculature remodeling [16–18]. During myocardial infarction, there is a noted elevation of KLF-4, which induces the secretion of collagen type I and III via the TGF-β1/Smad3 pathway, contributing to myofibroblast differentiation[19–21]. In diabetic cardiomyopathy KLF-5 increase has been shown to upregulate NADPH oxidase 4, a primary cause of cardiomyocyte superoxide accumulation, moreover, KLF-5 further leads to the activation of the serine palmitoyl transferase [SPT] long-chain base subunit 1 and 2, enzymes involved in ceramide synthesis; this process in turn changes the lipid landscape of the heart further contributing to the cardiomyopathy phenotype[11,22]. Interestingly, within diabetic patient's mutations in KLF-15, such as rs9838915 are associated with an increase in the risk of heart failure, as rs9838915 is linked to increase LV mass and thickening of the septal wall [23]. KLF-13 loss-of-function variant in the heart have been associated with double-outlet right ventricle and ventricular septal defects [10]. Hence reducing the activity of GATA-6, GATA-4, and ANP promoters. Additional KLF-13 mutations have also been connected to congenital heart defects, particularly KLF-13 is an activator of VEGF-a and ANP, nonetheless, its mutants are linked to the bicuspid aortic valve, patent ductus arteriosus, and ventricular septal defect [8]. Finally, in stroke there is a reduction in the expression of KLF-2 which directs proinflammatory effects by permitting the expression of the NF-B/p65 pathway[24].

Other examples of regulatory effects structured by the CACCC-box can be seen by the governing effects of SP family[25]. Among SP members, the transcriptional factor Sp1 has been reported to work with KLF-4 and even overlap its binding site. Particularly, during esophagus carcinoma, levels of Sp1 rise disrupting KRT19 regulation by KLF-4 which leads to malignant transformation and metastasis[25].

As mentioned, gene expression is a tightly controlled process facilitated in part by the accessibility of the DNA and regulated by the elements to which transcription factors can bind to. Therefore, we sought to take a bioinformatic approach and investigate gene enrichment, specially targeting diseases and metabolic pathways by establishing associations between genes containing the CACCC-box upstream their promoter, furthermore we investigated their relation to the KLF family as regulators or key players in CVDs to comprehensively understand the interplay between these genetic elements.

2. Materials and Methods

Raw data Collection & Web Scraping

The genome annotation of H. sapiens GRCh38/hg38 (2013) from the Ensembl project series was utilized in this study, employing the Signal Search Analysis Server (SSA) software package

(https://epd.expasy.org/ssa/findm.php) with the promoter prediction tool. The -CACCC- motif sought 150 base pairs upstream of the TATA-Box / Goldberg - Hogness Box functional site, selecting the best options.

The sequences containing the TATA-Box motif were obtained, and an algorithm was used to search for the -CACCC- motif within the first 150 bases. Subsequently, these sequences underwent multiple sequence alignment analyses using BLAST. Given the large number of sequences, an algorithm utilizing Web Scraping and Python code with the Selenium WebDriver was employed to extract the gene names associated with each sequence.

With the Selenium Python library and the web driver functionality, the program automates the process of inputting the search query on the BLAST NCBI website. The web driver interacts with the website's interactive slot, allowing the script to enter the desired query and initiate the search engine. Once the search is performed, the web scraper identifies the specific result page containing the relevant information. The required data from the result page is extracted by analyzing the website's HTML structure and utilizing techniques such as HTML parsing and Text pattern matching.

The extracted data was transformed into a CSV, for further analysis and storage. To be easily processed and utilized for further bioinformatic analysis.

Implementation with Python and Selenium

The Python programming language and the Selenium library (v. 4.9.0) were used to automate web scraping on the BLAST NCBI website. By leveraging the Selenium web driver and the Python framework, a web bot was created to complete the search form and crawl through the links on the main page. Specific parameters were established to narrow down the search to H*omo sapiens* (taxid:9606) and (RefSeqGene[Title]) OR gene[Title] in the Entrez query. The Python code utilizes XPath and EC.presence_of_element_located condition to target and extract the desired elements, ensuring only relevant data is included. The extracted gene information can be printed for further processing.

MEME

A randomized sampling of 40 sequences was conducted using Python from the dataset containing 3044 sequences to observe the filtered motif. The most representative motifs were selected by XSTREME (5.5.5) a motif discovery and enrichment analysis with p-value threshold of 0.05 and motif sequences in the range of 8 to 15 nucleotides. Similar motifs to those searched were compared within the same platform using Tomtom (5.5.5) a motif comparison tool.

STRING

The genes were categorized into two groups, direct and indirect, employing STRING (11.5) as TSV format and visualizing the interactions using Excel. The direct group comprised genes exhibiting various levels of interaction, ranging from weak to strong, with members of the KLF family. Conversely, the remaining genes were assigned to the indirect group.

ShinyGO

We employed ShinyGo (version 0.77), a specialized analysis tool tailored to the target species, *Homo sapiens*, to analyze the gene sets. We set a statistical significance threshold of 0.05 (false discovery rate, FDR), ensuring robust results.

To focus on relevant findings, we applied filtering criteria by considering only pathways with a minimum of 10 genes. This step aimed to reduce noise and present meaningful results. Ultimately, the analysis yielded a total of 30 metabolic pathways that were statistically significant and enriched with genes related to KLFs.

The pathways analysis provided valuable insights into the biological processes, cellular components, molecular functions, and diseases associated with the gene sets. This approach

facilitated the interpretation and analysis of the results, contributing to a comprehensive understanding of the interaction and functional implications of genes with KLF transcription factors.

Using the same dataset, relevant information regarding cardiovascular diseases could also be extracted through the representative KEGG pathway maps.

Enrichr & Appyter

Continuing from the previous point, utilizing the same dataset, we conducted a similar analysis using Enrichr. Specifically, we focused on disease analysis, utilizing the JENSEN database as a resource. Coupled with this analysis, the Appyter platform provided various visualization options for the results. This approach allowed us to identify the genes associated with the presented cardiovascular diseases and facilitated the creation of corresponding tables for further exploration.

3. Results

The CACCC-box motif was initially searched using Signal Search Analysis Server (SSAS) at 150bp upstream and 300bp downstream of transcription start sites (TSS) 5′-TATA(A/T)A(A/T)-3′(CACCC-regions-), Genome Reference Consortium Human Build 38 Organism: *Homo sapiens* (GRCh38) was used as reference. Results obtained a total of 3044 hits for the CACCC-box with 40 diverse matrices regions. Figure 1. Shows the 3 highest representative core matrix regions found encompassing the core CACCC-box. CCCCCACCCCCAC(C/T) sequence was found in 67.5% of the CACCC-regions with a score value of 7.8e-012, next CTCCCCCTCCT was found in 72.5% of the regions and with a score value of 6.9e-009, and finally the sequence CCCCTCC(C/T)(C/T)CCTC found in 62.5% of the regions and with a score of 3.2e-006. Additional sequences can be found in supplemental Table 1. Moreover Figure 2. shows a representative diagram for each chromosome, denoting CACCC-box binding (red marks) for identifiable genes. Out of 3044 potential hits; these included nonsense repeats, as well as non-transcription starting points, 1174 hits were related to identifiable genes.

Motif	Logo MT	Logo MT RC	Score	Sites	Distribution
CCCCCACCCCCACY			7.8e-012	27 (67.5%)	
AGGAGGGGGAG			6.9e-009	29 (72.5%)	
CCCCTCCYYCCTC			3.2e-006	25 (62.5%)	
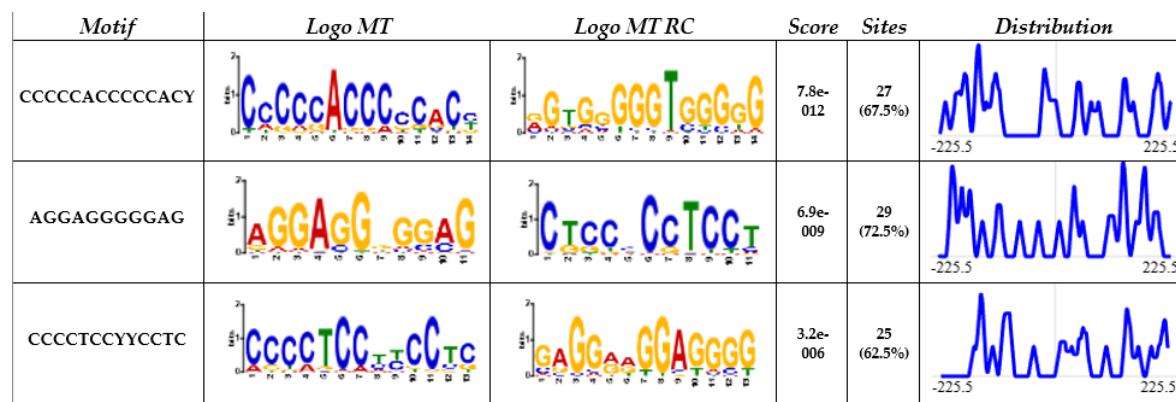
Figure 1. From left to right, the columns present: (1) motif sequence, (2) motif logo, (3) reverse complementary motif logo, (4) p-value, (5) percentage of motif occurrence in the dataset, and (6) distribution of the motive within the sequences.
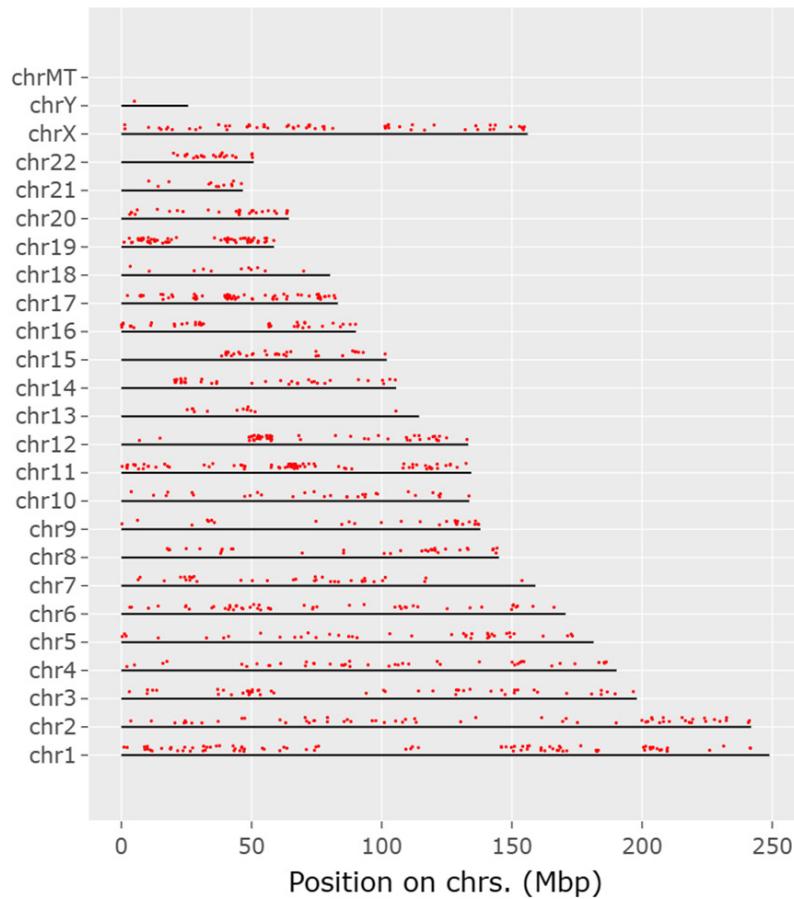
Figure 2. Chromosomic schematic representing CACCC-boxes (red marks) at 150bp from known TATA-boxes.

Our initial objective was to resolve the involvement of the KLF family, after processing our data registered 95 hits confirmed by STRING for KLF interactions. Table 1 shows the interacting genes related to their corresponding KLF members. Surprisingly, KLF-16 did not show any gene interaction, moreover, most confirmed gene interactions are linked to the activator group (Group 2), wherein KLF-4 had the most interactions with nearly 50 hits. Opposingly, Group 1 had the least number of interactions with a total of 8 hits; mostly related to the Homeobox NK family. It is important to note that while the CACCC box is a direct binding motif for KLFs other important factors involved in cardiovascular regulation can potentially bind. Using the same methodology as previously mentioned, we found potential motif sires for WT1, FATZ1, KMT2A, Specific Protein (SP), SALL4, and VEZF1. Supplemental Figure 1. Shows both the confirmed binding motif of these factors and the theoretical motifs found through our analysis. Supplemental Table 2 shows interactions between non-members of the KLF family and their target genes, including several genes which seem to be regulated either in combination KLFs and non KLFs or in alternative modes. Some of these genes include IGF1, MEF2C, MYH11, MYOD, TCF7L2 amongst others.

Table 1. Krüppel-like family of transcription factors gene interactions according to the STRING database. Table shows direct interactions between confirmed members of the KLF family and their target genes. These direct interactions are involved in diverse processes, including cellular development and differentiation.

Family member	KLF	Direct interaction
1	KLF-3	NKX2-8, NKX2-2, LLGL2
1	KLF-8	NKX2-8, NKX2-2

1	KLF-12	TFAP2A, NKX2-2, DNASE1
2	KLF-1	SPI1, HBZ, NKX2-8, TET2, CD44, HBB, HBE1, HBG2, GATA2, AHSP EPB42, UBC
2	KLF-2	GZMB, TNNT2, DUSP1, CCR7, LEF1, ZAP70, CXCR4, TEK, HAS2, CD44, FOXP3, FOS, FBXW7, ICAM1, GATA2, TRPV4, CD8A, TBX3
2	KLF-4	SALL4, COL1A1, SPI1, TNNT2, SOX9, MYOD1, TBX3, EED, EPAS1, EPCAM, LEF1, CDH2, NODAL, GRHL3, T, IGF1, FOS, HAS2, TBX5, TUBB3, NKX2-5, DLK1, MEF2C, GATA2, FN1, PAX9, SOX15, LEFTY2, NES, TCF7L2, FLNA, DKK1, RARG, KLF-6, KDM2B, TET2, ESR1, GFAP, CXCR4, IL6, MYH11, CD24, CD44, TCF4, PAX6, PBX1
2	KLF-5	SOX9, TBX3, TP63, NCOR1, FBXW7, NODAL, UCP3, GATA2, MKX, SUMO1, ESR1
2	KLF-6	PPP1R15A, FOS, DUSP1, KLF-4, ESR1, CYP2E1, NAT1, ANGPT1
2	KLF-7	PROKR2, GABPB1, TBX3, LEF1, NFIA, NKX2-8
3	KLF-9	NKX2-8, THRA, RB1, MAPK10, TFAP2A, ACTG2
3	KLF-10	NKX2-8, SMAD7, NKX2-2, FOS, FOXP3
3	KLF-11	HNF1A, NKX2-2, FOXP3,
3	KLF-13	HBB, HBE1, TBX5
3	KLF-14	HHEX, TCF7L2, FHL2, CAMK1D, JAZF1
3	KLF-16	-
-	KLF-15	NR3C1, NFIA, ARNTL, GATA2, MKX, NKX2-8
-	KLF-17	TG, NODAL, CD44, TBX3, SLC5A5

Taking into consideration all 1174 genes, we ran a ShinyGO analysis focused on diseases using Jessen database. Overall, there were 30 known diseases, with an overrepresentation of genes with the CACCC-motif at the TSS. Of these diseases 4 were directly linked to cancer, and 7 to cardiovascular disease, the focus of this manuscript. Other minor representations included neurological condition, pancreatis, lung, and congenital diseases (Figure 3A). In addition, we went ahead with and ran a second ShinyGO analysis using only the 95 genes confirmed to have KLF interaction. This second analysis showed 5 similar diseases wherein CACCC-box is overrepresented (Figure 3B).

The resulting information was then taken to ENRICHR for gene identification. Table 2A shows the relation between the identified disease and the set of CACCC-box related genes for all 1174 potential genes. Meanwhile, Table 2B denotes the relation between the identified disease and the set of CACCC-box related genes for the 95 confirmed genes related to the KLF family. Interestingly, both tables show similar gene enrichment for holt-oram, hypertension, cerebrovascular disease, and heart conduction disease. Meanwhile, cardiomyopathy and coronary artery disease were only present in Table 2A, while diabetes and diabetic retinopathy were present in Table 2B. Out of all the present genes, the troponin I isoform 3 (TNNI3) was the gene most prevalent as seen in hypertension, cardiomyopathy, cerebrovascular disease, and coronary artery disease. Other highly represented genes were ICAM and IL6, both related to inflammation. Additionally, there was an important presence of the TBX family, particularly 3, 5, and 6, and troponin T isoform 2 (TNNT2).
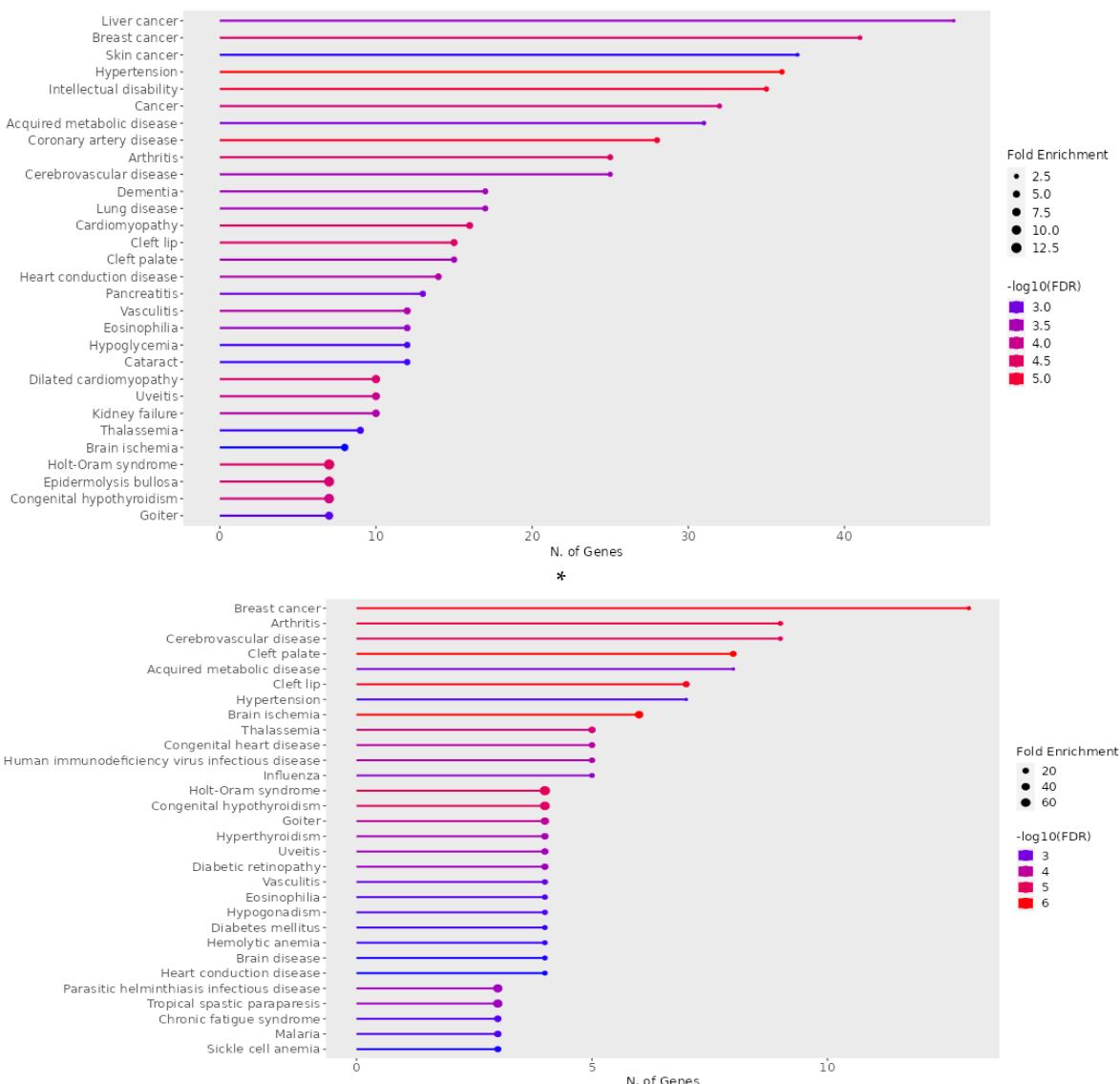
Figure 3. ShinyGo enrichment analysis of CACCC-box involvement in diseases A) (Upper) Complete set B) (Bottom) KLF specific interactions. Horizontal plane correlates to the number of genes involved per disease. Fold enrichment and False discovery rate (FDR) apply to the integrated data to resolve for differentially confirmed gene indicating overrepresentation within the disease. Fold enrichments directly relates to gene target confirmation. FDR relates to the expected number of positive or significant classifications.

Table 2. ENRICHR analysis of genes and cardiac-related diseases directly regulated by CACCC-box (A) all genes, (B) confirmed interactions to the KLF family.

A) Disease	q - value	p–value	Genes
Holt-Oram syndrome	2.91E-06	0.001117367	TFAP2B, SALL4, GJA5, TBX6, TBX5, NKX2-5, TBX3
Hypertension	6.45E-06	0.00148591	ENPEP, NCF1, CUL3, NPR3, KLK4, UMOD, ADM, CACNA1D, DBH, AAAS, NR3C1, ICAM1, ADD2, GJA5, NPY, TNNI3, GGT1, DUOX2, KL, ACE, ANGPT1, CYP4F2, CAV1, ATP6AP2, MTHFR, IGF1, FOS, RHOA, SELP, IL6, TH,

			NEDD4, SCNN1B, CYP11B1, TRPV4, REN, ANG
Dilated cardiomyopathy	8.98E-06	0.001549478	BAG3, LAMA4, TNNC1, TNNT2, ANKRD1, LDB3, TNNI3, DSG2, CRYAB, VCL
Coronary artery disease	9.42E-06	0.001549478	SLC22A4, KLK4, ADM, TNFRSF11B, LDB3, AAAS, MPO, CX3CL1, THBS4, ICAM1, XIRP1, TNNI3, CD36, GGT1, LDLR, SCAI, ACE, ANGPT1, APOA2, MTHFR, APOC3, IGF1, SELP, FABP3, IL6, TNNT1, TNNT2, REN, CD68
Cerebrovascular disease	0.0015946	0.04946049	NTRK2, PRKCH, ACE, ANGPT1, CYP4F2, F11, KLK4, MTHFR, CXCR4, ADM, IGF1, FOS, MPO, RHOA, ICAM1, GFAP, SELP, IL6, TNNT1, TNNT2, REN, TNNI3, CD68, NES
Cardiomyopathy	1.33E-05	0.001700831	TNNC1, LAMA4, PSEN1, TBX5, MTO1, JPH2, BAG3, GJA5, MYL2, TNNT2, DSG2, TNNI3, CTNNA3, FPGT-TNNI3K, CRYAB, VCL
Heart conduction disease	0.000268757	0.013461245	CAV1, IL5RA, DPT, CXCR4, MTHFR, TBX5, SLC4A4, DKK1, ARHGAP24, SFRP2, NFIA, CLDN14, TRPM7, BMPR1B

B) Disease	q - value	p-value	Genes
Holt-Oram syndrome	3.30E-07	2.16E-05	SALL4, TBX5, NKX2-5, TBX3
Hypertension	0.00045705	0.006255711	NR3C1, ICAM1, ANGPT1, IGF1, FOS, IL6, TRPV4
Diabetes mellitus	0.000426195	0.006255711	UCP3, HNF1A, FOXP3
Diabetic retinopathy	3.12E-05	0.000720179	IL6, ANGPT1, IGF1, ICAM1
Cerebrovascular disease	1.25E-06	6.14E-05	IL6, ANGPT1, TNNT2, CXCR4, FOS, IGF1, NES, GFAP, ICAM1
Congenital heart disease	1.14E-05	0.000405807	MEF2C, NKX2-5, TBX5, LEFTY2, NODAL
Heart conduction disease	0.000629384	0.007274938	CXCR4, TBX5, DKK1, NFIA

Further ShinyGo analysis was run for the general 1174 genes which reveal their involvement in major (Cardiac) metabolic pathways such as PI3K-Akt, and MAPK signaling. (Figure 4A). Additionally diabetic and hypertrophic cardiomyopathy-related pathways. Moreover, the 95 gene specific analysis reveals enrichment in MAPK signaling, AGE-RAGE, RAS signaling, and HIF-1 signaling amongst other pathways. (Figure 4B)
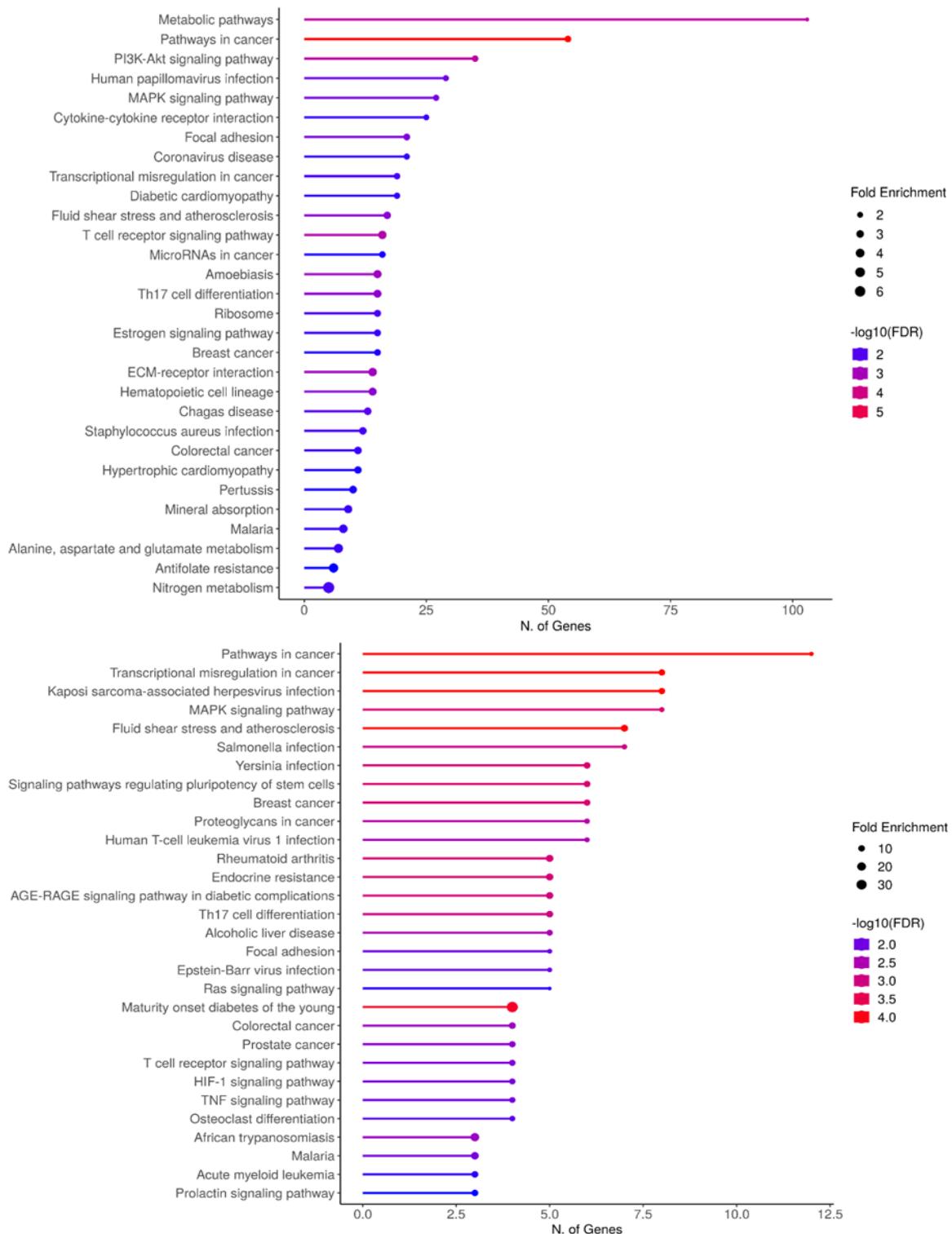
Figure 4. ShinyGo enrichment analysis of CACCC-box involvement in biological processes A) (Upper) Complete set B) (Bottom) KLF specific interactions. Horizontal plane correlates to the number of genes involved per biological process. Fold enrichment and False discovery rate (FDR) apply to the integrated data to resolve for differentially confirmed gene indicating overrepresentation within the biological process. Fold enrichments directly relates to gene target confirmation. FDR relates to the expected number of positive or significant classifications.

ShinyGo analysis directly links to the KEGG database, to directly visualize gene interaction in the cell (cardiomyocyte). As a representation of cardiac disease involvement, we present in Figure 5 a schematic depiction of hypertrophic cardiomyopathy and Figure 6 of diabetic cardiomyopathy.

Blue-colored genes represent KLF-confirmed interacting genes, while red colored represents the general 1174 (CACCC-motif) genes. Furthermore, to show direct protein-protein (KLF) interaction, a STRING (12.0) analysis was run for both diseases (supplemental Figure 2). For hypertrophic cardiomyopathy, the analysis showed a direct interaction, specifically co-expression and text mining, between IL-6 and KLF-2, KLF-4, and KLF-6. Similarly, IL-6 also directly interacts with TNNT2, TNNI3, and ACE, suggesting the potential for indirect interactions mediated by IL-6, KLFs, and genes related to the sarcomere.

For diabetic hypertrophy, KLF-4 exhibits a robust interaction with COL1A1, which, in turn, demonstrates a co-expression interaction with CD36. Similarly, ACE, NCF1, REN, and SLC2A1 also display co-expression interactions with CD36. The presence of genes related to the mitochondrial cell component, such as NDUFB9, NDUFA6, NDUFA12, ATP5MC3, COX7B, NDUFA7, and ATP5P, which specifically function in mitochondrial proton transport, is also observed with a strong correlation.
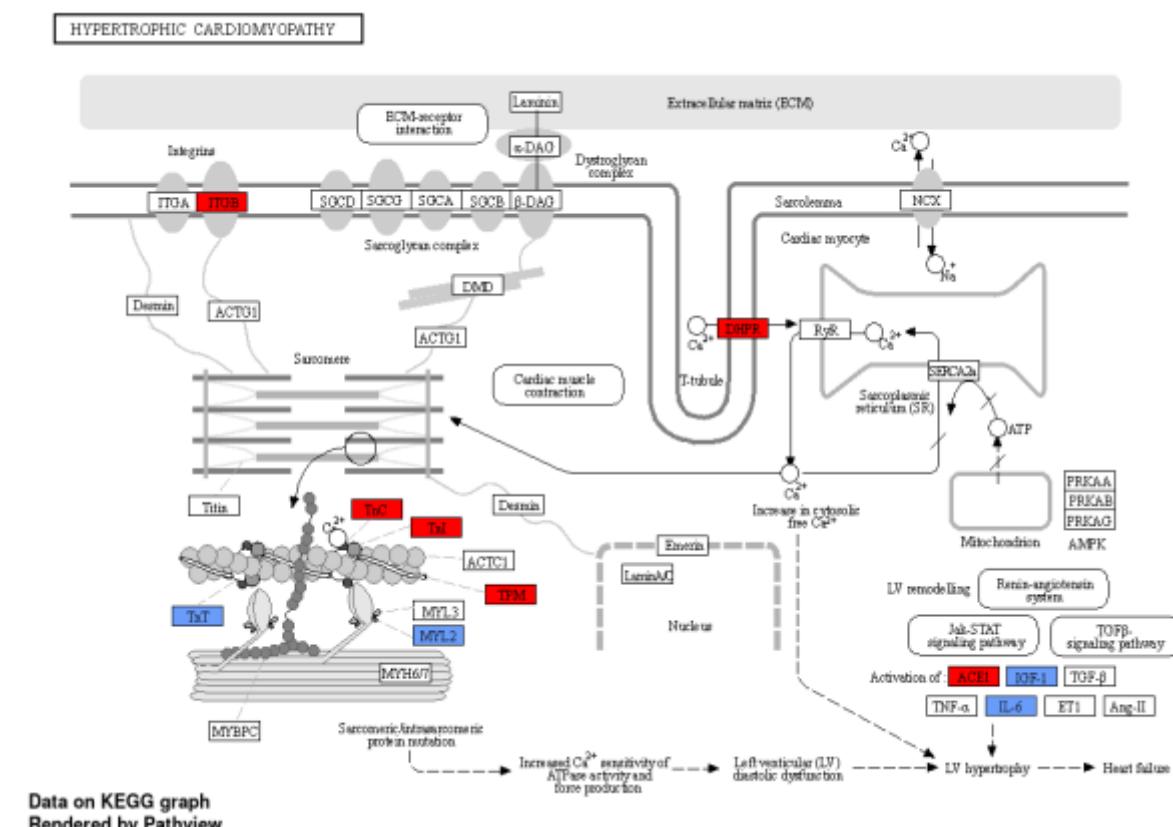
Figure 5. KEGG pathway representative of hypertrophic cardiomyopathy demonstrating biological gene interactions. CACCC-box containing genes (Red) not directly related to KLF interaction (blue) with direct KLF interaction.
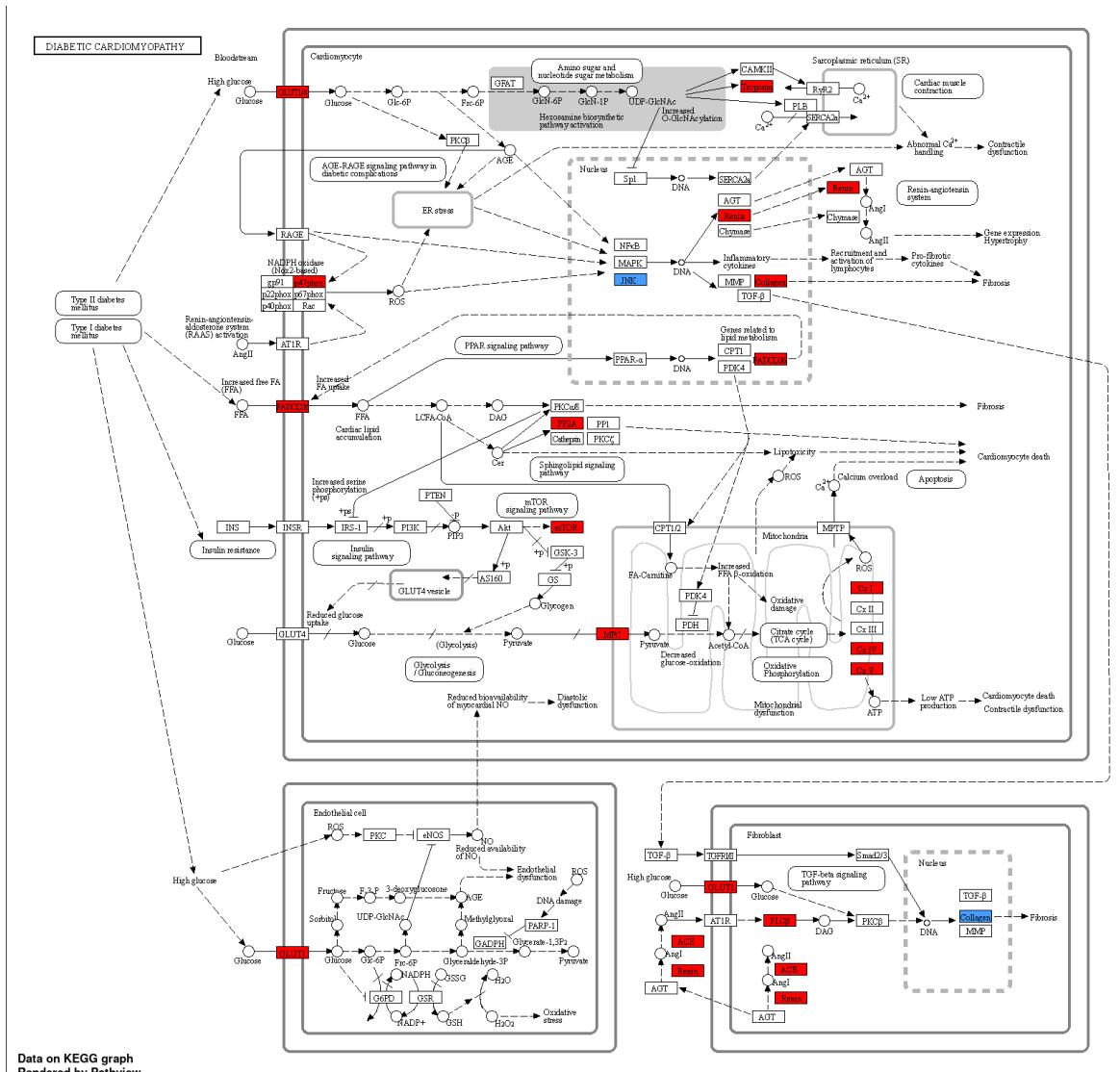
Figure 6. KEGG pathway representative of diabdetic cardiomyopathy demonstrating biological gene interactions. CACCC-box containing genes (Red) not directly related to KLF interaction (blue) with direct KLF interaction.

4. Discussion

The CACCC-box motif, a well-recognized cis-regulatory element, plays a fundamental role in a diverse range of developmental processes and diseases, including CVDs. This research delves specifically into the CACCC-box and its intricate connection to CVDs, focusing on the KLF of transcription factors as the primary effectors.

Our findings illuminate significant activity among various KLFs, particularly those of the KLF subgroup 2 recognized as transcriptional activators (Table 1). Notably, KLF-4 stands out as the KLF with the highest number of interactions, likely because of its well-established roles in pluripotency, cancer, and most importantly, cardiovascular development and disease (Table 1, Table2, Figure 3, Figure 4 and supplemental Figure 2)[4,11,26–28]. This is further confirmed by several crucial KLF-4 interactions in our work, including those with PAX9, PAX6, TBX5, and TBX3 (Table 1 and supplemental Figure 2). Previous research has established a link between KLF-4 and PAX9 (Table 1 and supplemental Figure 2) during cardiovascular development, specifically regarding the crucial WNT signaling pathway. In the pharyngeal endoderm, PAX9 interacts with TBX1 and GBX2, tightly controlling the intricate process of pharyngeal arch morphogenesis [29]. Similarly, PAX6 and NKX2.2 (Table 1) have been shown to orchestrate the differentiation of neural tube ventral progenitors by mediating the sonic hedgehog signaling pathway[30].

Our exploration of the motif landscape yielded intriguing insights into pathological hypertrophy, revealing striking similarities to signaling pathways observed during early cardiac development (Figure 3). This resulted in the identification of KLF-4 interactions with genes like GATA4, MEF2C, TBX5, NKX2.5, and SRF (Table 1, Table 2, and supplemental Table 2, supplemental Figure 2), all of which are known regulators of cardiac hypertrophy and pro-hypertrophic pathways[31–33]. Interestingly, KLF-15 presents a contrasting role, as it was found to repress TCF7L2 expression specifically in the postnatal heart, influencing cardiac growth. TCF7L2 is associated with diabetes and exhibits dependence on KLF-14 (Table 1 and supplemental Table 2), where reduced expression increases pre-adipocyte proliferation but impairs lipogenesis[34]. KLF-15 plays a vital role in regulating cardiac physiology, including circadian oscillations in cardiac cells, and regulating genes involved in cardiac metabolism. Notably, both KLF-15 and ARNTL work in parallel to control cardiac circadian rhythms, even though no evidence of co-regulation with the ARNTL gene was observed (Table 1). KLF-15 plays a key role in binding circadian repressors REV-ERBα (NR1D1) and NCOR (Table 1), ensuring stable expression of a subset of cardiac genes [35,36].

The significant impact of hypertrophic cardiomyopathy and diabetic cardiomyopathy on cardiovascular health cannot be overstated. While hypertrophic cardiomyopathy affects approximately 1 in 500 individuals globally, with a significant proportion of cases remaining undiagnosed, diabetic cardiomyopathy emerges as a serious complication of diabetes mellitus, significantly increasing the risk of heart failure and premature death [37,38]. Both conditions highlight the critical importance of early detection, accurate diagnosis, and implementing effective management strategies to minimize the burden of cardiovascular disease on the population.

As diabetes is a condition linked to increased cardiovascular risk (Table 2 and Figure 3), it was unsurprising that our work revealed gene interaction between KLFs, particularly KLF-14, and genes associated with diabetic risk factors, such as CAMK1D, HHEX, JAZF1, and TCF7L2 (Table 1 and supplemental Table 2). Notably, TCF7L2's ability to modulate PIK3R1 expression suggests its involvement in insulin signaling pathways, potentially mediating diabetes risk through interactions with KLF-14 [39,40].

As illustrated in Figure 4, metabolic pathways like MAP Kinases and PI3K/Akt were overrepresented, with approximately 40 genes associated with these processes, their participation, or regulation. KLF-4 emerged as the regulator of genes such as FN1, IGF1, and COL1A1 (Table 1). Upon binding to the $\alpha5\beta1$ membrane receptor, FN1 activates the PI3K/Akt signaling pathway [29]. Additionally, the PI3K/Akt signaling pathway induces the proliferation of cardiac fibroblasts and collagen accumulation, thus contributing to fibrosis, a process associated with collagen (COL1A1) (Table 1 and supplementary Table 2) [41]. This process is a consequence of various diseases identified in the enrichment analysis (Figure 3 and Table 2), such as hypertension and its vascular remodeling process, which alters the composition of the vascular wall, leading to stiffness and loss of elasticity [42].

Furthermore, our analysis identified other factors, including members of the SP family, WT1, and others (supplemental Table 2 and supplemental Figure 1), that potentially bind to the CACCC-box motif. Notably, earlier studies showed that SP members 1 and 3 are further involved in contraction, as they play critical roles in calcium management by regulating the Ca2+ ATPase pump SERCA2 [43]. Other SP members have been shown to participate in other cardiac-related pathways, such is the case of SP6 which can directly interact with β-catenin and modulate the expression of TCF7L2, a result confirmed by the presence of the CACCC-box in the promoter region of TCF7L2 (Table 1). These findings implicate all these members as potential regulators in electrically related CVDs.

Moreover, our results also showed regulatory effects of SP7 over COL1A1, members of the FGF and TNF families, as well as DKK-1 and IL-6 (supplemental Table 2). This implicates SP7 in potential remodeling and inflammatory processes. Specifically, in the heart, SP7 has been previously shown to play a decremental role in favor of calcification by permitting the activity of Osteoporin and Matrix Gla protein. Downregulation of SP7 in VSMCs leads to a loss of contractile markers and paves the way for the upregulation of RUNX2 [44,45]. RUNX2 is also regulated by another CACCC-box motif

binding protein KLF-5, which regulates the activity of SP7. Overall, the process of calcification further induces remodeling via inflammation, to compensate for the reduction in contractile force[45,46].

Furthermore, our results revealed that the CACCC-box can serve as a complementary reverse binding site for PATZ1 and WIT1, suggesting a broader transcriptional regulation network. In addition, other potential binding proteins found included VEZF1 and SALL4 (supplemental Table 2). The VEZF1 gene encodes a nuclear protein highly conserved among vertebrates, regulating genes related to cardiac muscle contraction and potentially regulating dilated cardiomyopathy[47]. Regarding the interaction between SALL4 and KLFs in the heart, SALL4 interacts directly with members of the T-box family 3 and 5 and regulates the expression of Gap junction alpha-5 (GJA5) and KLF-4. This final regulation of KLF-4 is further seen in stem cells, as SALL4 has important role in proliferation and survival, mediated by the Bmi-1 pathway[48]. Meanwhile, in cardiac development, another KLF, KLF-5, is involved in regulating TBX5, as KLF-5 absence reduces expression of TBX5 (Table 1, Table 2, and supplemental Table 2), contrasting with the interactions described between SALL4 and TBX5 in cardiac patterning[49].

From a clinical perspective, our study holds value at a diagnostic and potentially prognostic level. By continuing to understand the molecular mechanisms that govern gene expression, new potential biomarkers for diverse CVDs can be found and used, giving rise to predicting the development of these pathologies. It is worth noting, that through a deep understanding of the dysregulation of specific genes and pathways, targets such as downstream effectors of KLF4, 5, and 15 who are involved in hypertrophy and fibrosis may be new avenues for treatment in conditions such as diabetic cardiomyopathy and heart failure. In addition, our finding suggested pathways like the PI3k/Akt which, when properly targeted, could attenuate the development of cardiovascular complications in diabetic patients.

Through the discovery and comprehension on interactions between KLFs and other CACCC-box binding factors, regulatory effects can be better explained, additionally the further exploration of adjacent motif and potential binding partners sets-up future studies to have a more developed picture at a particular pathology. Targeting these interactions may offer new opportunities for therapeutic intervention, either through small molecule inhibitors or gene editing approaches aimed at modulating the activity of specific transcription factors involved in cardiovascular gene expression.

From a diagnostic perspective, the identification of gene signatures associated with cardiovascular pathology may help identify novel biomarkers for risk stratification, early detection, and monitoring of disease progression in patients at risk of CVD's. Biomarkers derived from the transcriptional profiles of CACCC box-regulated genes, especially those under the control of KLF family members, may offer improved sensitivity and specificity compared to existing biomarkers, enabling more accurate diagnosis and prognosis of CVD's. By leveraging the understanding of the molecular mechanisms underlying CVD's, clinicians can look into the future to tailor therapeutic interventions targeting specific pathways or gene networks dysregulated in each patient, thereby optimizing treatment efficacy, and minimizing adverse effects.

5. Limitations

The bioinformatic approach employed in this study offers valuable perceptions into the regulatory landscape of cardiovascular gene expression. It is essential to acknowledge its inherent limitations. The stringency parameters applied in the analysis may be a risk of overlooking potentially relevant motifs or genes that do not meet the predefined criteria, which may introduce biases or inaccuracies in the results. However, the genomic data sourced from the Ensembl project series provides a widely used and reliable information. Experimental and future research are necessary to further validate the functional significance of the identified genes and pathways in the context of CVDs.

6. Conclusions

In conclusion, this study sheds light on the intricate network of CACCC-box regulators, particularly KLFs and their interactions with various genes and pathways, predominantly in the

context of cardiovascular diseases. Our findings highlight the crucial role of KLFs, especially KLF-4, in these conditions and pave the way for further exploration of their potential as therapeutic targets. Additionally, the identification of other CACCC-box binding factors like SP family members suggests their potential involvement in various CVDs, warranting further investigation. Overall, this work contributes to our understanding of the complex regulatory mechanisms underlying cardiovascular health and disease.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Supplemental Table 1. Python analysis results of 3044hits from Signal Search Analysis Server resulted in 40 representative sequences. Supplemental Table 2. Transcription factors interactions according to the STRING database. The table shows direct interactions between non-members of the KLF family and their target genes. Supplemental Figure 1 CACCC-box binding partners (Non KLFs) Supplemental Figure 2. KLF and CACCC-box gene interactions (Left) Hypertrophic cardiomyopathy (Right) Diabetic cardiomyopathy.

Author Contributions: Conceptualization, JAG-L., PAZ-M and JFI.; methodology, JAG-L and PAZ-M.; software, PAZ-M; validation, MGS-S, OR-N and DFB-C.; formal analysis JAG-L and GRP-R.; investigation, MGS-S, ENG-T.; writing—original draft preparation, JAG-L, PAZ-M; writing—review and editing, JFI.; "All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by CONAHCYT, grant number CF2009-51208.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Davie, K.; Jacobs, J.; Atkins, M.; Potier, D.; Christiaens, V.; Halder, G.; Aerts, S. Discovery of Transcription Factors and Regulatory Regions Driving in Vivo Tumor Development by ATAC-Seq and FAIRE-Seq Open Chromatin Profiling. PLoS Genet **2015**, 11, e1004994, doi:10.1371/journal.pgen.1004994.

2. Liu, H.; Lei, C.; He, Q.; Pan, Z.; Xiao, D.; Tao, Y. Nuclear Functions of Mammalian MicroRNAs in Gene Regulation, Immunity and Cancer. Mol Cancer **2018**, 17, 1–14, doi:10.1186/s12943-018-0765-5.

3. Frietze, S.; Farnham, P.J. Transcription Factor Effector Domains. In; 2011; pp. 261–277.

4. Santoyo-Suarez, M.G.; Mares-Montemayor, J.D.; Padilla-Rivas, G.R.; Delgado-Gallegos, J.L.; Quiroz-Reyes, A.G.; Roacho-Perez, J.A.; Benitez-Chao, D.F.; Garza-Ocañas, L.; Arevalo-Martinez, G.; Garza-Treviño, E.N.; et al. The Involvement of Krüppel-like Factors in Cardiovascular Diseases. Life **2023**, 13, 420, doi:10.3390/life13020420.

5. Okumura, K.; Sakaguchi, G.; Takagi, S.; Naito, K.; Mimori, T.; Igarashi, H. Sp1 Family Proteins Recognize the U5 Repressive Element of the Long Terminal Repeat of Human T Cell Leukemia Virus Type I through Binding to the CACCC Core Motif. Journal of Biological Chemistry **1996**, 271, 12944–12950, doi:10.1074/jbc.271.22.12944.

6. Fraizer, G.C.; Shimamura, R.; Zhang, X.; Saunders, G.F. PAX 8 Regulates Human WT1 Transcription through a Novel DNA Binding Site. Journal of Biological Chemistry **1997**, 272, 30678–30687, doi:10.1074/jbc.272.49.30678.

7. Lavallée, G.; Andelfinger, G.; Nadeau, M.; Lefebvre, C.; Nemer, G.; Horb, M.E.; Nemer, M. The Kruppel-like Transcription Factor KLF13 Is a Novel Regulator of Heart Development. EMBO Journal **2006**, 25, 5201–5213, doi:10.1038/sj.emboj.7601379.

8. Abhinav, P.; Zhang, G.-F.; Zhao, C.-M.; Xu, Y.-J.; Wang, J.; Yang, Y.-Q. A Novel KLF13 Mutation Underlying Congenital Patent Ductus Arteriosus and Ventricular Septal Defect, as Well as Bicuspid Aortic Valve. Exp Ther Med **2022**, 23, 311, doi:10.3892/etm.2022.11240.

9. Li, W.; Li, B.; Li, T.; Zhang, E.; Wang, Q.; Chen, S.; Sun, K. Identification and Analysis of KLF13 Variants in Patients with Congenital Heart Disease. BMC Med Genet **2020**, 21, 1–8, doi:10.1186/S12881-020-01009-X/FIGURES/3.

10. Wang, S.-S.; Wang, T.-M.; Qiao, X.-H.; Huang, R.-T.; Xue, S.; Dong, B.-B.; Xu, Y.-J.; Liu, X.-Y.; Yang, Y.-Q. KLF13 Loss-of-Function Variation Contributes to Familial Congenital Heart Defects. Eur Rev Med Pharmacol Sci **2020**, 24, 11273–11285, doi:10.26355/eurrev_202011_23617.

11. Kyriazis, I.D.; Hoffman, M.; Gaignebet, L.; Lucchese, A.M.; Markopoulou, E.; Palioura, D.; Wang, C.; Bannister, T.D.; Christofidou-Solomidou, M.; Oka, S.I.; et al. KLF5 Is Induced by FOXO1 and Causes Oxidative Stress and Diabetic Cardiomyopathy; 2021; Vol. 128; ISBN 1215707142.

12. Gao, C.; Qian, H.; Shi, Q.; Zhang, H. MicroRNA-363-3p Serves as a Diagnostic Biomarker of Acute Myocardial Infarction and Regulates Vascular Endothelial Injury by Targeting KLF2. Cardiovasc Diagn Ther **2020**, 10, 421–430, doi:10.21037/cdt-19-700.

13. Dai, Y.; Yan, T.; Gao, Y. Silence of MiR-32-5p Promotes Endothelial Cell Viability by Targeting KLF2 and Serves as a Diagnostic Biomarker of Acute Myocardial Infarction. Diagn Pathol **2020**, 15, 19, doi:10.1186/s13000-020-00942-y.

14. Vinjamur, D.S.; Wade, K.J.; Mohamad, S.F.; Haar, J.L.; Sawyer, S.T.; Lloyd, J.A. Krüppel-like Transcription Factors KLF1 and KLF2 Have Unique and Coordinate Roles in Regulating Embryonic Erythroid Precursor Maturation. Haematologica **2014**, 99, 1565–1573, doi:10.3324/haematol.2014.104943.

15. Walther, O. New Study Reveals Latest Data on Global Burden of Cardiovascular Disease Available online: https://www.acc.org/About-ACC/Press-Releases/2023/12/11/18/48/New-Study-Reveals-Latest-Data-on-Global-Burden-of-Cardiovascular-Disease (accessed on 4 March 2024).

16. Xie, Z.; Chen, J.; Wang, C.; Zhang, J.; Wu, Y.; Yan, X. Current Knowledge of Krüppel-like Factor 5 and Vascular Remodeling: Providing Insights for Therapeutic Strategies. J Mol Cell Biol **2021**, 13, 79–90, doi:10.1093/jmcb/mjaa080.

17. Chang, E.; Nayak, L.; Jain, M.K. Krüppel-like Factors in Endothelial Cell Biology. Curr Opin Hematol **2017**, 24, 224–229, doi:10.1097/MOH.0000000000000337.

18. Xu, Y.; Xu, S.; Liu, P.; Koroleva, M.; Zhang, S.; Si, S.; Jin, Z.G. Suberanilohydroxamic Acid as a Pharmacological Kruppel-Like Factor 2 Activator That Represses Vascular Inflammation and Atherosclerosis. J Am Heart Assoc **2017**, 6, doi:10.1161/JAHA.117.007134.

19. Zhang, Y.; Wang, Y.; Liu, Y.; Wang, N.; Qi, Y.; Du, J. Krüppel-Like Factor 4 Transcriptionally Regulates TGF-B1 and Contributes to Cardiac Myofibroblast Differentiation. PLoS One **2013**, 8, 0–9, doi:10.1371/journal.pone.0063424.

20. Yang, F.; Luo, L.; Zhu, Z. De; Zhou, X.; Wang, Y.; Xue, J.; Zhang, J.; Cai, X.; Chen, Z.L.; Ma, Q.; et al. Chlorogenic Acid Inhibits Liver Fibrosis by Blocking the MiR-21-Regulated TGF-B1/Smad7 Signaling Pathway in Vitro and in Vivo. Front Pharmacol **2017**, 8, 1–13, doi:10.3389/fphar.2017.00929.

21. Li, H.; Wang, Y.; Liu, J.; Chen, X.; Duan, Y.; Wang, X.; Shen, Y.; Kuang, Y.; Zhuang, T.; Tomlinson, B.; et al. Endothelial Klf2-Foxp1-TGFβ Signal Mediates the Inhibitory Effects of Simvastatin on Maladaptive Cardiac Remodeling. Theranostics **2021**, 11, 1609–1625, doi:10.7150/thno.48153.

22. Hoffman, M.; Palioura, D.; Kyriazis, I.D.; Cimini, M.; Badolia, R.; Rajan, S.; Gao, E.; Nikolaidis, N.; Schulze, P.C.; Goldberg, I.J.; et al. Cardiomyocyte Krüppel-Like Factor 5 Promotes De Novo Ceramide Biosynthesis and Contributes to Eccentric Remodeling in Ischemic Cardiomyopathy. Circulation **2021**, 143, 1139–1156, doi:10.1161/CIRCULATIONAHA.120.047420.

23. Patel, S.K.; Wai, B.; Lang, C.C.; Levin, D.; Palmer, C.N.A.; Parry, H.M.; Velkoska, E.; Harrap, S.B.; Srivastava, P.M.; Burrell, L.M. Genetic Variation in Kruppel like Factor 15 Is Associated with Left Ventricular Hypertrophy in Patients with Type 2 Diabetes: Discovery and Replication Cohorts. EBioMedicine **2017**, 18, 171–178, doi:10.1016/j.ebiom.2017.03.036.

24. Sweet, D.R.; Lam, C.; Jain, M.K. Evolutionary Protection of Krüppel-Like Factors 2 and 4 in the Development of the Mature Hemovascular System. Front Cardiovasc Med **2021**, 8, doi:10.3389/fcvm.2021.645719.

25. Kim, C.-K.; He, P.; Bialkowska, A.B.; Yang, V.W. SP and KLF Transcription Factors in Digestive Physiology and Diseases. Gastroenterology **2017**, 152, 1845–1875, doi:10.1053/j.gastro.2017.03.035.

26. Zheng, B.; Han, M.; Wen, J.-K. Role of Krüppel-like Factor 4 in Phenotypic Switching and Proliferation of Vascular Smooth Muscle Cells. IUBMB Life **2010**, NA-NA, doi:10.1002/iub.298.

27. Shankman, L.S.; Gomez, D.; Cherepanova, O.A.; Salmon, M.; Alencar, G.F.; Haskins, R.M.; Swiatlowska, P.; Newman, A.A.C.; Greene, E.S.; Straub, A.C.; et al. KLF4-Dependent Phenotypic Modulation of Smooth Muscle Cells Has a Key Role in Atherosclerotic Plaque Pathogenesis. Nat Med **2015**, 21, 628–637, doi:10.1038/nm.3866.

28. Shindo, T.; Manabe, I.; Fukushima, Y.; Tobe, K.; Aizawa, K.; Miyamoto, S.; Kawai-Kowase, K.; Moriyama, N.; Imai, Y.; Kawakami, H.; et al. Krüppel-like Zinc-Finger Transcription Factor KLF5/BTEB2 Is a Target for Angiotensin II Signaling and an Essential Regulator of Cardiovascular Remodeling. Nat Med **2002**, 8, 856–863, doi:10.1038/nm738.

29. Ji, J.; Chen, L.; Zhuang, Y.; Han, Y.; Tang, W.; Xia, F. Fibronectin 1 Inhibits the Apoptosis of Human Trophoblasts by Activating the PI3K/Akt Signaling Pathway. Int J Mol Med **2020**, doi:10.3892/ijmm.2020.4735.

30. Ericson, J.; Rashbass, P.; Schedl, A.; Brenner-Morton, S.; Kawakami, A.; van Heyningen, V.; Jessell, T.M.; Briscoe, J. Pax6 Controls Progenitor Cell Identity and Neuronal Fate in Response to Graded Shh Signaling. Cell **1997**, 90, 169–180, doi:10.1016/S0092-8674(00)80323-2.

31. Christoforou, N.; Chellappan, M.; Adler, A.F.; Kirkton, R.D.; Wu, T.; Addis, R.C.; Bursac, N.; Leong, K.W. Transcription Factors MYOCD, SRF, Mesp1 and SMARCD3 Enhance the Cardio-Inducing Effect of GATA4, TBX5, and MEF2C during Direct Cellular Reprogramming. PLoS One **2013**, 8, doi:10.1371/journal.pone.0063577.

32. Cornwell, J.D.; McDermott, J.C. MEF2 in Cardiac Hypertrophy in Response to Hypertension. Trends Cardiovasc Med **2022**, doi:10.1016/j.tcm.2022.01.002.

33. Chen, C.Y.; Schwartz, R.J. Recruitment of the Tinman Homolog Nkx-2.5 by Serum Response Factor Activates Cardiac Alpha-Actin Gene Transcription. Mol Cell Biol **1996**, 16, 6372–6384.

34. Dahlgren, A.; Zethelius, B.; Jensevik, K.; Syvänen, A.-C.; Berne, C. Variants of the TCF7L2 Gene Are Associated with Beta Cell Dysfunction and Confer an Increased Risk of Type 2 Diabetes Mellitus in the ULSAM Cohort of Swedish Elderly Men. Diabetologia **2007**, 50, 1852–1857, doi:10.1007/s00125-007-0746-5.

35. Aggarwal, A.; Costa, M.J.; Rivero-Gutiérrez, B.; Ji, L.; Morgan, S.L.; Feldman, B.J. The Circadian Clock Regulates Adipogenesis by a Per3 Crosstalk Pathway to Klf15. Cell Rep **2017**, 21, 2367–2375, doi:10.1016/j.celrep.2017.11.004.

36. Zhang, L.; Prosdocimo, D.A.; Bai, X.; Fu, C.; Zhang, R.; Campbell, F.; Liao, X.; Coller, J.; Jain, M.K. KLF15 Establishes the Landscape of Diurnal Expression in the Heart. Cell Rep **2015**, 13, 2368–2375, doi:10.1016/j.celrep.2015.11.038.

37. Butzner, M.; Leslie, D.L.; Cuffee, Y.; Hollenbeak, C.S.; Sciamanna, C.; Abraham, T. Stable Rates of Obstructive Hypertrophic Cardiomyopathy in a Contemporary Era. Front Cardiovasc Med **2022**, 8, doi:10.3389/fcvm.2021.765876.

38. Dandamudi, S.; Slusser, J.; Mahoney, D.W.; Redfield, M.M.; Rodeheffer, R.J.; Chen, H.H. The Prevalence of Diabetic Cardiomyopathy: A Population-Based Study in Olmsted County, Minnesota. J Card Fail **2014**, 20, 304–309, doi:10.1016/j.cardfail.2014.02.007.

39. Grarup, N.; Andersen, G.; Krarup, N.T.; Albrechtsen, A.; Schmitz, O.; Jørgensen, T.; Borch-Johnsen, K.; Hansen, T.; Pedersen, O. Association Testing of Novel Type 2 Diabetes Risk Alleles in the JAZF1, CDC123/CAMK1D, TSPAN8, THADA, ADAMTS9, and NOTCH2 Loci With Insulin Release, Insulin Sensitivity, and Obesity in a Population-Based Sample of 4,516 Glucose-Tolerant Middle-Aged Danes. Diabetes **2008**, 57, 2534–2540, doi:10.2337/db08-0436.

40. Jonsson, A.; Ladenvall, C.; Ahluwalia, T.S.; Kravic, J.; Krus, U.; Taneera, J.; Isomaa, B.; Tuomi, T.; Renström, E.; Groop, L.; et al. Effects of Common Genetic Variants Associated With Type 2 Diabetes and Glycemic Traits on α- and β-Cell Function and Insulin Action in Humans. Diabetes **2013**, 62, 2978–2983, doi:10.2337/db12-1627.

41. Li, X.; Wang, G.; QiLi, M.; Liang, H.; Li, T.; E, X.; Feng, Y.; Zhang, Y.; Liu, X.; Qian, M.; et al. Aspirin Reduces Cardiac Interstitial Fibrosis by Inhibiting Erk1/2-Serpine2 and P-Akt Signalling Pathways. Cellular Physiology and Biochemistry **2018**, 45, 1955–1965, doi:10.1159/000487972.

42. Intengan, H.D.; Schiffrin, E.L. Vascular Remodeling in Hypertension. Hypertension **2001**, 38, 581–587, doi:10.1161/hy09t1.096249.

43. Brady, M. Sp1 and Sp3 Transcription Factors Are Required for Trans-Activation of the Human SERCA2 Promoter in Cardiomyocytes. Cardiovasc Res **2003**, 60, 347–354, doi:10.1016/S0008-6363(03)00529-7.

44. Nishio, Y.; Dong, Y.; Paris, M.; O'Keefe, R.J.; Schwarz, E.M.; Drissi, H. Runx2-Mediated Regulation of the Zinc Finger Osterix/Sp7 Gene. Gene **2006**, 372, 62–70, doi:10.1016/j.gene.2005.12.022.

45. Zhang, J.; Zheng, B.; Zhou, P.; Zhang, R.-N.; He, M.; Yang, Z.; Wen, J.-K. Vascular Calcification Is Coupled with Phenotypic Conversion of Vascular Smooth Muscle Cells through Klf5-Mediated Transactivation of the Runx2 Promoter. Biosci Rep **2014**, 34, doi:10.1042/BSR20140103.

46.   Petsophonsakul, P.; Furmanik, M.; Forsythe, R.; Dweck, M.; Schurink, G.W.; Natour, E.; Reutelingsperger, C.; Jacobs, M.; Mees, B.; Schurgers, L. Role of Vascular Smooth Muscle Cell Phenotypic Switching and Calcification in Aortic Aneurysm Formation. Arterioscler Thromb Vasc Biol **2019**, 39, 1351–1368, doi:10.1161/ATVBAHA.119.312787.

47.   Paavola, J.; Alakoski, T.; Ulvila, J.; Kilpiö, T.; Sirén, J.; Perttunen, S.; Narumanchi, S.; Wang, H.; Lin, R.; Porvari, K.; et al. Vezf1 Regulates Cardiac Structure and Contractile Function. EBioMedicine **2020**, 51, 102608, doi:10.1016/j.ebiom.2019.102608.

48.   de Albuquerque Dias, R.; Balbinot, K.M.; da Silva Kataoka, M.S.; de Melo Alves Júnior, S.; de Jesus Viana Pinheiro, J. Expression of Stem Cell Markers SALL4, LIN28A, and KLF4 in Ameloblastoma. Diagn Pathol **2023**, 18, 92, doi:10.1186/s13000-023-01379-9.

49.   Chiplunkar, A.R.; Lung, T.K.; Alhashem, Y.; Koppenhaver, B.A.; Salloum, F.N.; Kukreja, R.C.; Haar, J.L.; Lloyd, J.A. Krüppel-Like Factor 2 Is Required for Normal Mouse Cardiac Development. PLoS One **2013**, 8, e54891, doi:10.1371/journal.pone.0054891.