RNA BIOLOGY IN CARDIOVASCULAR DISEASE

EDITED BY: Maarten M. G. van den Hoogenhof, Hamid El Azzouzi and Abdelaziz Beqqali

PUBLISHED IN: Frontiers in Genetics, Frontiers in Cell and Developmental Biology and Frontiers in Cardiovascular Medicine

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88971-808-5
DOI 10.3389/978-2-88971-808-5

About Frontiers

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

Frontiers Journal Series

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

Dedication to Quality

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

What are Frontiers Research Topics?

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

RNA BIOLOGY IN CARDIOVASCULAR DISEASE

Topic Editors:
Maarten M. G. van den Hoogenhof, Heidelberg University Hospital, Germany
Hamid El Azzouzi, Erasmus Medical Center, Netherlands
Abdelaziz Beqqali, University of Edinburgh, United Kingdom

Citation: van den Hoogenhof, M. M. G., El Azzouzi, H., Beqqali, A., eds. (2021). RNA Biology in Cardiovascular Disease. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88971-808-5

Table of Contents

frontiers in Genetics

EDITORIAL
published: 12 October 2021
doi: 10.3389/fgene.2021.775091

Editorial: RNA Biology in Cardiovascular Disease

Maarten M. G. van den Hoogenhof[1,2], Hamid El Azzouzi[3]* and Abdelaziz Beqqali[4]**

[1]Institute of Experimental Cardiology, Heidelberg University Clinic, Heidelberg, Germany, [2]Partner Site Heidelberg, German Center for Cardiovascular Research (DZHK), Heidelberg, Germany, [3]Department of Molecular Genetics, Erasmus University Medical Center, Rotterdam, Netherlands, [4]Centre for Cardiovascular Science, The Queen's Medical Research Institute, The University of Edinburgh, Edinburgh, United Kingdom

Keywords: RNA biology, heart, RNA splicing, cardiovasclar disease, non-coding RNAs

Editorial on the Research Topic

RNA Biology in Cardiovascular Disease

In the last years, RNA biology has gained significant attention in cardiovascular science. The rise of next-generation sequencing has not only uncovered various previously unknown types of RNA, but also a much greater extent of the function(s) of RNA than previously realized. For example, multiple pivotal cardiac splicing factors, such as RBM20 and RBM24, have recently been identified (Guo et al., 2012; Yang et al., 2014); the functional role of RNA modifications such as m6A are started to be elucidated (Dorn et al., 2019), and the roles of noncoding RNAs (ncRNAs), such as lincRNAs (Hobuss et al., 2019) and circRNAs (Aufiero et al., 2019), are being unravelled. Heart disease remains one of the major health problems in the world, and changes in RNA biology can both be a cause and a consequence of disease (van den Hoogenhof et al., 2016). These new insights in RNA biology underline the need to study the diverse roles for RNA in the healthy and diseased heart more comprehensively. In this Research Topic, we focus on novel roles of ncRNAs, but also review new insights in techniques and processes related to RNA biology in the heart.

One of the major advances in RNA sequencing, is single-cell RNA sequencing (as opposed to bulk RNA sequencing). Marin-Sedeño et al. review what single-cell RNA sequencing has meant for the cardiovascular field, more specifically how it has been used to shape the cardiac cellular landscape both in healthy and diseased hearts. In addition, they discuss drawbacks and future perspectives of this relatively new and powerful tool. Another advantage of the advance of next generation sequencing is the increased availability of transcriptomic data. This makes it possible to (re)analyze published datasets with new questions in mind, and use this as a starting point for new studies. Nie et al. used published transcriptomic data to gain insight in abdominal artery aneurisma (AAA), a condition that kills approximately 13.000 people every year. Their study showed that AAA samples could be distinguished from healthy samples by the infiltration of immune cells, and they identified 5 key immunological genes, SSTR1, GPER1, CCR10, PI3, and MAP2K1, that may be involved in AAA. Yan et al. used publicly available data to provide a weighted gene correlation network analysis of immune cells in atrial tissue of AF patients as compared to non-AF. They identified and confirmed high atrial expression of three genes (CTSS, CSF2RB and NCF2) which are strongly associated with AF. The authors propose these three candidate genes may play an important role in the development of AF and may serve as potential therapeutic targets for AF treatment. However, further in-depth studies are needed to unravel what role these three genes have in the molecular mechanisms underlying AF. Wang et al. used publicly available datasets to identify miRNAs and their targets in aortic dissection. They identified miR-193a-3p and its target ACTG2 to regulate vascular smooth muscle cell proliferation and migration and phenotypic switching. Taken together, they provide a novel molecular pathway in the development of aortic dissection that has the potential to be therapeutically targeted.

Abu-Halima et al. have characterized the blood miRNA profile of univentricular heart patients with and without Fontan palliation using a large panel of miRNAs for initial screening in order to identify those that may indicate advanced liver fibrosis/cirrhosis and thus might have clinical as well

OPEN ACCESS

Edited by:
William C. Cho, QEH, Hong Kong, SAR China

Reviewed by:
Alexander E. Berezin, Zaporizhia State Medical University, Ukraine

***Correspondence:**
Maarten M. G. van den Hoogenhof
maarten.vandenhoogenhof@med.uni-heidelberg.de
Hamid El Azzouzi
h.elazzouzi@erasmusmc.nl
Abdelaziz Beqqali
a.beqqali@ed.ac.uk

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: 13 September 2021
Accepted: 30 September 2021
Published: 12 October 2021

Citation:
van den Hoogenhof MMG, El Azzouzi H and Beqqali A (2021) Editorial: RNA Biology in Cardiovascular Disease. Front. Genet. 12:775091. doi: 10.3389/fgene.2021.775091

as prognostic impact in this cohort of patients. This screen led to the identification of microRNA-29b/c-3p and the authors propose that this miR represents an independent predictor of advanced liver fibrosis/cirrhosis and may be used in the risk assessment of these patients. Pham et al. identified a novel LncRNA, which they termed Aerrie, in HUVECs exposed to shear stress. They found that Aerrie is expressed in the endothelium, where it regulates endothelial function through binding with YBX1 and controlling DNA damage repair. Since Aerrie is upregulated during aging, this suggests that Aerrie is crucial to maintain proper endothelial cell function during life. Another important subclass of non-coding RNAs are circular RNAs (circRNAs), which are characterized by a stable structure without free ends. In the work of Xu et al., after circRNA-sequencing of atherosclerotic mouse aortas, a circRNA called circDENND1B was shown to have an important role in modulating atherosclerosis through regulation of cholesterol efflux. By acting as a sponge for miR-17-5p, circDENND1B induced an increase in ATP binding cassette subfamily A member 1 (Abca1) expression and thus promoting cholesterol efflux. Moreover, circDENND1B could alleviate foam cell formation, and is involved in modulating inflammation enhancing the antiatherosclerosis effect of IL-1b mAb in mice. While this work elegantly provides new insights into the interaction between inflammation and cholesterol transport, the therapeutic potential of circDENND1B needs further exploration.

While much of the RNA research of the last decade centred around its non-coding forms and the numerous roles these ncRNAs have in cellular physiology, the classical coding role of RNA in the heart is still fascinatingly complex. Duran et al. showcased this in their effort to sum-up the tremendous insights into CaMKII dependent signalling. Encompassing 12 subunits, the oligomerization of this holoenzyme is multifaceted with the existence of different splice variants adding to the complexity. Importantly, the authors highlight that lesser expression doesn't mean lesser importance and that next to the four isoforms described in the review other isoforms that are less abundant in the heart might have surprisingly important spatio-temporal effects during disease. Obviously, more work is needed to understand the role of the different splice variants that code this vital signalling complex. What is less obvious but important to realize, is that studying ncRNAs that target a complex without understanding this complex' true composition is less fruitful. Three-dimensional genome organization has been extensively studied and has emerged as an important regulator of gene expression. In his article, Bertero discusses the possibility that interchromosomal genome architecture can be regulated by RNA biogenesis, and uses the example of RNA polymerase II clusters and foci of the cardiac-specific splicing regulator RBM20 to show the existence and functionality of *trans*-interacting chromatin domains. With this, the author extends on an exciting but not completely understood paradigm of gene regulation, and proposes an experimental framework with which to test this hypothesis. And lastly, an elegant study by Boileau et al. shows a comprehensive integrated analysis of RNA-sequencing and Ribo-sequencing datasets in the same heart samples under physiological vs pathological remodeling. The authors have identified changes in transcriptome and translatome which are specific to physiological vs pathological changes using their hypertrophy mouse models. They found that transcriptome networks are only partially reproducible at the translatome level, providing further evidence of the importance of post-transcriptional control at the level of translation. Their results provide novel insights into the complexity of the organisation of *in vivo* cardiac regulatory networks.

The diverse collection of articles in this Research Topic shows that RNA biology is a vast and vibrant field of study in cardiovascular biology, and the immense advances in next-generation sequencing and its analysis methods will lead to new and exciting insights into the heart.

AUTHOR CONTRIBUTIONS

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

REFERENCES

Aufiero, S., Reckman, Y. J., Pinto, Y. M., and Creemers, E. E. (2019). Circular RNAs Open a New Chapter in Cardiovascular Biology. *Nat. Rev. Cardiol.* 16 (8), 503–514. doi:10.1038/s41569-019-0185-2

Dorn, L. E., Lasman, L., Chen, J., Xu, X., Hund, T. J., Medvedovic, M., et al. (2019). The N 6 -Methyladenosine mRNA Methylase METTL3 Controls Cardiac Homeostasis and Hypertrophy. *Circulation* 139 (4), 533–545. doi:10.1161/circulationaha.118.036146

Guo, W., Schafer, S., Greaser, M. L., Radke, M. H., Liss, M., Govindarajan, T., et al. (2012). RBM20, a Gene for Hereditary Cardiomyopathy, Regulates Titin Splicing. *Nat. Med.* 18 (5), 766–773. doi:10.1038/nm.2693

Hobuss, L., Bär, C., and Thum, T. (2019). Long Non-coding RNAs: At the Heart of Cardiac Dysfunction? *Front. Physiol.* 10, 30. doi:10.3389/fphys.2019.00030

van den Hoogenhof, M. M. G., Pinto, Y. M., and Creemers, E. E. (2016). RNA Splicing. *Circ. Res.* 118 (3), 454–468. doi:10.1161/circresaha.115.307872

Yang, J., Hung, L.-H., Licht, T., Kostin, S., Looso, M., Khrameeva, E., et al. (2014). RBM24 Is a Major Regulator of Muscle-specific Alternative Splicing. *Dev. Cell* 31 (1), 87–99. doi:10.1016/j.devcel.2014.08.025

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

frontiers in Genetics

BRIEF RESEARCH REPORT
published: 16 November 2020
doi: 10.3389/fgene.2020.583124

A Multi-Network Comparative Analysis of Transcriptome and Translatome Identifies Novel Hub Genes in Cardiac Remodeling

Etienne Boileau[1,2,3], *Shirin Doroudgar*[2,3], *Eva Riechert*[2,3], *Lonny Jürgensen*[2,3], *Thanh Cao Ho*[2], *Hugo A. Katus*[2,3], *Mirko Völkers*[2,3†] *and Christoph Dieterich*[1,2,3*†]

[1] Section of Bioinformatics and Systems Cardiology, Klaus Tschira Institute for Integrative Computational Cardiology, Heidelberg, Germany, [2] Department of Internal Medicine III (Cardiology, Angiology, and Pneumology), University Hospital Heidelberg, Heidelberg, Germany, [3] DZHK (German Centre for Cardiovascular Research), Partner Site Heidelberg/Mannheim, Berlin, Germany

Our understanding of the transition from physiological to pathological cardiac hypertrophy remains elusive and largely based on reductionist hypotheses. Here, we profiled the translatomes of 15 mouse hearts to provide a molecular blueprint of altered gene networks in early cardiac remodeling. Using co-expression analysis, we showed how sub-networks are orchestrated into functional modules associated with pathological phenotypes. We discovered unappreciated hub genes, many undocumented for their role in cardiac hypertrophy, and genes in the transcriptional network that were rewired in the translational network, and associated with semantically different subsets of enriched functional terms, such as *Fam210a*, a novel musculoskeletal modulator, or *Psmd12*, implicated in protein quality control. Using their correlation structure, we found that transcriptome networks are only partially reproducible at the translatome level, providing further evidence of post-transcriptional control at the level of translation. Our results provide novel insights into the complexity of the organization of *in vivo* cardiac regulatory networks.

Keywords: cardiovascular, cardiac hypertrophy, transcription/RNA-seq, translation/Ribo-seq, co-expression networks

OPEN ACCESS

Edited by:
Abdelaziz Beqqali, University of Edinburgh, United Kingdom

Reviewed by:
Yibin Wang, University of California, Los Angeles, United States
Thomas M. Vondriska, University of California, Los Angeles, United States
Christoph D. Rau, University of North Carolina at Chapel Hill, United States
Luc Furic, Peter MacCallum Cancer Centre, Australia

***Correspondence:**
Christoph Dieterich
christoph.dieterich@uni-heidelberg.de

†These authors share senior authorship

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *14 July 2020*
Accepted: *09 October 2020*
Published: *16 November 2020*

Citation:
Boileau E, Doroudgar S, Riechert E, Jürgensen L, Ho TC, Katus HA, Völkers M and Dieterich C (2020) A Multi-Network Comparative Analysis of Transcriptome and Translatome Identifies Novel Hub Genes in Cardiac Remodeling. Front. Genet. 11:583124. doi: 10.3389/fgene.2020.583124

1. INTRODUCTION

Exercise- and disease-induced cardiac growth are associated with different molecular profiles and differ in the signaling pathways that drive remodeling; yet both are characterized by an increase in size of cardiomyocytes, sarcomerogenesis, and overall increase in heart-weight-to-body-weight (HW/BW) ratio (Shimizu and Minamino, 2016). While adaptive exercise-induced hypertrophy allows the heart to maintain an adequate cardiac output with improved contractility, pathological hypertrophy is a maladaptive response, concurring with irreversible changes (e.g., cardiomyocyte loss, fibrosis, reduced cardiac function), and typically progressing to heart failure (Shimizu and Minamino, 2016; Bernardo et al., 2018).

Little is known about the molecular mechanisms controlling physiological hypertrophy, particularly from a multi-omics systems biology perspective. Yet our understanding of the *in vivo* transition from adaptive hypertrophy to cardiac dysfunction has important clinical

implications (Boström et al., 2010). Only recently have post-transcriptional regulatory networks been uncovered that are of central importance for morphological remodeling in fibrosis (Chothani et al., 2019), or for modulating the early reponse to cardiac stress (Doroudgar et al., 2019).

In this study, we adopt a systems biology approach to integrate multi-omics data through the use of co-expression networks to highlight higher-order relationships among gene programs that are expressed in the heart *in vivo* under growth stimuli. In such networks, genes are connected if there is a significant co-expression relationship between them (Langfelder and Horvath, 2008). Modules or sub-networks represent clusters of genes with related function or involved in common processes or pathways. Our analysis of 15 mouse left ventricular tissues from experimental models of exercise- and disease-induced cardiac hypertrophy showed, for the first time, the organization of the transcriptome and translatome into networks of biologically meaningful clusters of co-expressed genes. By correlating module expression and disease phenotypes, we were able to show the synchronized expression dynamics of genes encoding extracellular matrix, and cytoskeletal proteins, and a diminished contribution of electron transport complex genes, genes associated with oxidative phosphorylation and mitochondrial function. While concerted dynamic changes were observed in both transcriptome and translatome networks, transcriptome networks were only partially reproducible at the translatome level, reflecting the existence of RNA- or Ribo-specific modules. In contrast to differential expression analysis, co-expression and rewiring analysis led us to the identification of yet uncharacterized candidate genes, key to organizing the behavior of transcriptome and translatome networks. In particular, our results uncovered *Fam210a*, a novel musculoskeletal modulator, which we hypothesize to regulate the expression of mitochondrial encoded genes; and *Psmd12*, a regulatory subunit of the 26S proteasome, whose deregulation may act as a pathogenic factor compromising protein quality control in cardiomyocytes.

2. MATERIALS AND METHODS

2.1. Experimental Models

We acquired data from experimental models of pathological cardiac hypertrophy (transverse aortic constriction or TAC) and swimming-induced physiological hypertrophy. TAC (27 gauge needle) surgery was performed as previously described (Völkers et al., 2013), and animals were sacrificed after 2 days ($n = 4$) and 2 weeks ($n = 5$). For exercise training in the physiological hypertrophy model, mice swam regularly in a water tank for either 2 days ($n = 3$) or 2 weeks ($n = 3$) (Evangelista et al., 2003). The experiments were performed in 9-week-old male C57Bl6/N mice using the RiboTag system (Doroudgar et al., 2019; Kmietczyk et al., 2019). All animal experimental procedures were reviewed and approved by the Institutional Animal Care and Use Committees at the Animal Experiment Review Board of the government of the state of Baden-Württemberg, Germany.

2.2. Preparation of Sequencing Libraries

Mice were sacrificed, and their hearts were excised, washed in PBS containing 100 μg/ml cycloheximide (CHX), and snap frozen in liquid nitrogen. Left ventricular tissue was homogenized using a tissue homogenizer in 5 volumes of ice-cold polysome buffer (20 mM Tris pH 7.4, 10 mM MgCl, 200 mM KCl, 2 mM DTT, 1% Triton X-100, 1U DNase/μl) containing 100 μg/ml CHX. Ribo-seq and RNA-seq libraries were prepared for each biological replicate from the identical lysate. Ribosome protected fragments (RPFs) were generated after immunoprecipitation of cardiac myocyte-specific polysomes with anti-HA magnetic beads after treating the lysate with RNase I (Ambion). Libraries were generated according to the mammalian Ribo-seq kit (Illumina), and sequenced on the HiSeq 2000 platform using a 50-bp sequencing chemistry.

2.3. The RiboTag System

In the RiboTag mouse, the exon 4 of the Rpl22 gene is flanked by Loxp recombination sites, followed by an HA-tagged exon 4. When the RiboTag mouse is crossed to a Cre driver mouse, the Cre recombinase enzyme is activated resulting in the removal of the LoxP-flanked wild type Rpl22 exon 4 and replacement with the HA-tagged Rpl22 exon 4, which is incorporated into the ribosome particle. In mouse hearts, a cell-specific promotor (Myosin heavy chain, α isoform, or αMHC, encoded by the Myh6 gene) drives the expression of Cre which induces cardiomyocyte-specific HA-tagged ribosomes. RiboTag mice were purchased from Jackson Laboratory (JAX ID 011029) and bred to the αMHC-Cre mice line to obtain homozygous mice expressing Rpl22-HA in cardiomyocytes.

2.4. Detecting Active Translation

Translation prediction using Ribo-seq data was performed with Rp-Bp v2.0 (Malone et al., 2017), based on Ensembl release 96. We used evidence from uniquely mapped reads and periodic fragment lengths only. For each sample, the fragment lengths and ribosome P-site offsets were determined from a metagene analysis using the automatic Bayesian selection of read lengths and ribosome P-site offsets (BPPS). The final list of translation events includes, in addition to annotated open reading frames (ORFs), ORFs with evidence of translation outside of annotated coding sequences (**Supplementary Table 1**). For the analyses, translation in non-coding regions as well as variants of canonical coding sequences were discarded (**Supplementary Figures 1, 2**). We required ORFs to have a minimum length of 3 aa and more than 10 in-frame P-sites. The final list of translation events was further filtered to only include ORFs that were predicted in at least three samples and whose host gene was also annotated in APPRIS (Rodriguez et al., 2018), resulting in 9,129 unique genes (**Supplementary Table 2**).

2.5. Sequencing Data Alignment

Adapters removal and quality filtering was done with flexbar v3.0.3 (Dodt et al., 2012) using standard filtering parameters implemented in Rp-Bp. Reads aligning to a custom bowtie2 v2.3.0 (Langmead and Salzberg, 2012) ribosomal index were discarded. Remaining reads were then aligned in genomic

coordinates to the mouse genome (GRCm38.p6) with STAR v2.5.3a (Dobin et al., 2013). For the RNA-seq data, reads were trimmed from the 3' end after adapter removal, to match the maximum periodic fragment length, as determined with the BPPS method, for each sample. Finally, abundance estimates and read count to coding sequences were obtained using HTSeq-count (Anders et al., 2014), taking into account the strand-specific protocols.

2.6. Constructing Gene Co-expression Networks

Read counts to coding sequences were used, only including genes that were considered to be translated (9,129 genes), as explained in the section 2.4. We removed low variance genes and genes with the lowest sequencing-depth normalized average expression (first centile). From these, 7,976 genes with the highest connectivity were clustered on the basis of topological overlap (TO) to identify patterns of co-expression, using the WGCNA R package (Langfelder and Horvath, 2008, 2012) (**Supplementary Table 3**). The network construction was done separately for Ribo-seq and RNA-seq data, on this common set of genes. We first applied a regularized log transformation, and corrected for batch effects, where applicable (Johnson et al., 2006). Weighted adjacencies were defined based on signed co-expression similarity using biweight midcorrelation and a soft thresholding power of $\beta = 18$ (for both RNA-seq and Ribo-seq). For each network, a reference TO matrix was first calculated. To produce robust and reproducible clusters, we then performed bootstrap-resampling ($n = 100$) and computed the TO matrix for each of the resampled networks. In each case, resampling was done within the physiological (swim) or the pathological (TAC) group. The final consensus TO matrix was defined as the median of all scaled TO matrices, and used as input for hierarchical clustering. The consensus TO matrix can be viewed as a "smoothed" version of the adjacency matrix. Network modules whose eigengenes were highly correlated were merged, and characterized by their eigengene expression and significance. To validate module membership, we applied *post-hoc* resampling ($n = 100$) by subsetting TO of random modules matched by size with respect to the consensus TO for every module. A one-proportion Z-test was used to assess whether the mean TO of the random modules was higher than that of the module assigned by the hierarchical clustering and merging algorithm.

The module membership is defined as the correlation (biweight midcorrelation) between eigengene and gene expression values, and measures the importance of a gene within a cluster. Gene significance is defined as the correlation (biweight midcorrelation) between genes and biological traits or disease association. To create indicators for level contrasts, e.g., pathological vs. physiological, TAC 2d vs. swim 2d, or TAC 2w vs. swim 2w, categorical variables were binarized as input for correlation. The pathological model includes all TAC samples, and the physiological model includes all swim samples. For binary/discrete variable correlation, biweight midcorrelation was replaced by the standard Pearson correlation. Hub genes were defined as genes with the highest intramodular connectivity. We ranked genes in each module and selected as hub genes those having a number of interactions greater than two standard deviations above the average connectivity found in a given module, i.e., with a Z-score > 2.

Preservation statistics were derived using the RNA-seq network as a reference, using the correlation structure of the networks. We used the *medianRank* statistic, defined as the mean of the observed density and connectivity statistics, and the $Z_{summary}$, defined as the mean of Z-scores computed for density and connectivity measures (Langfelder et al., 2011). The *medianRank* was used to compare relative preservation among clusters. The $Z_{summary}$ was used to assess the significance of observed statistics by distinguishing preserved from non-preserved clusters via permutation testing ($n = 100$).

Differentially connectivity (DC) was defined as the $Log_{10}(k_{in}^{RNA}/k_{in}^{Ribo})$. It is not a robust measure, and it is based uniquely on intramodular connectivity. A gene was DC if this ratio was greater than two standard deviations above the average across all genes. All results are found in **Supplementary Table 4**.

2.7. Dynamic Neighborhoods Score

The Dn score, or dynamic neighborhoods score, of a given gene is calculated based on the variance of the state-space adjacency matrix over the network states (RNA-seq and Ribo-seq), relative to the mean centroid, as $\mathrm{Dn} = \sum_{i=1}^{RNA,Ribo} d\left(\mathbf{V_i}, centroid\right)^2$, where d is the Euclidean distance, and $\mathbf{V_i}$ is a vector of genes in the RNA-seq and Ribo-seq networks (Goenawan et al., 2016). To calculate the Dn score, we used the consensus TO matrix, thresholded at 0.1 (all interactions below this threshold were considered inexistant).

2.8. Differential Translational-Efficiency Analysis

Differential translational-efficiency analysis was performed using DESeq2 (Love et al., 2014). The calculation of change in translational efficiency was done using an interaction term (~assay + condition + assay:condition) with a likelihood-ratio test, accounting for variance and level of expression. Regulation status of a gene at the transcriptional and/or translational level can be integrated using the fold changes from standard Wald test (RNA-seq, Ribo-seq) and the likelihood-ratio test (translational efficiency). All analyses were performed on the same set of genes used as input for network clustering ($n = 7{,}976$). We used genome-wide significance threshold of FDR < 0.05, and a fold change (FC) of $\log_2(1.2)$. All results are found in **Supplementary Table 7**.

2.9. Gene Ontology (GO) Enrichment

GO enrichment (The Gene Ontology Consortium, 2018) to assign functional annotation to modules were performed with topGO v2.34.0 (Alexa and Rahnenfuhrer, 2018). To define the relevant gene sets corresponding to each clusters, we considered hub genes and genes with strong module membership and significance for a given trait (heart-weight-to-body-weight ratio, pathological vs. physiological, TAC2d vs. swim2d, or TAC2w vs. swim2w). The latter were identified by taking the upper quartile of genes with the highest module membership and gene

significance for a trait having the highest correlation between absolute values of module membership and gene significance. The universe of genes consisted of all translated genes (n = 9,129). All results are found in **Supplementary Table 5**.

3. RESULTS

3.1. Exercise- vs. Disease-Induced Cardiac Remodeling

To characterize stress-induced cardiac remodeling and the response to acute and chronic pressure overload, we used the swim model (Wang et al., 2010) and the transverse aortic constriction (TAC) model (Rockman et al., 1991). The swim model is referred to as the physiological or the healthy model, and represents exercise-induced hypertrophy. The TAC model is referred to as the pathological model, and represents disease-induced hypertrophy.

To study co-expression networks during cardiac remodeling, we used *in vivo* Ribo-seq and RNA-seq libraries from 15 mouse hearts (**Figure 1A**). We used the RiboTag approach to capture the cardiac translatome. Cardiomyocyte-specific analysis of ribosome protected fragments was achieved after affinity purification using the RiboTag mouse (Doroudgar et al., 2019). To catalog translation events, we performed an unsupervised search for actively translated open reading frames (ORFs) using Rp-Bp (Malone et al., 2017) (**Supplementary Figures 1, 2**). The final list of translated genes was used as background for co-expression and differential expression analyses. Co-expression networks were constructed separately for RNA-seq (transcriptome) and Ribo-seq (translatome) count data and used to identify hub genes associated with cardiac remodeling (**Supplementary Figure 3**). All translated ORFs used in this study, RNA-seq and Ribo-seq read counts can be found as **Supporting Information** and **Supplementary Tables 2**, **3**.

We monitored the acute response at an intermediate time point (2 days after TAC), and a chronic time point (2 weeks after TAC), when cellular and molecular remodeling has occurred, but cardiac function is preserved (Doroudgar et al., 2019). Matching time points were monitored in the physiological model (swim at 2 days and 2 weeks). At the RNA-seq level only, we observed an upregulation of *Nppa*, the fetal isoform of myosin heavy chain (*Myh7*) (Taegtmeyer et al., 2010), clinically relevant genes such as *Ankrd1* or *Synpo2l* (Ling et al., 2017; van Eldik et al., 2017), as well as a number of genes implicated in tissue remodeling (**Figure 1B**). These observations were corroborated with RT-PCR results for two markers of hypertrophy and fibrosis (**Figure 1C**). Increased HW/BW ratios were detected after 2 weeks in the swim and TAC models, with a larger increase in the pathological model (**Figure 1D**). Taken together, these results are consistent with graded, pathological cardiac hypertrophy in the TAC model. In the physiological model, we did not observe fetal gene re-expression, typically associated with metabolic remodeling in a variety of pathophysiologic conditions (Taegtmeyer et al., 2010). There was a significant increase in *Ppargc1a*, a master regulator of mitochondrial biogenesis associated with physiological hypertrophy (Boström et al., 2010), and *Agt*, key component of the renin-angiotensin system (RAS), suggesting that the swim model did not induce a pathological hypertrophy phenotype, but instead improved the cellular energetics of the heart.

3.2. Co-expression Networks of Cardiac Remodeling

We calculated topological overlap and clustered genes, identifying independently 17 distinct co-expression modules for each of the RNA-seq and Ribo-seq networks (**Supplementary Table 4**). The modules were labeled in order from RNA1 to RNA17, and from Ribo1 to Ribo17, using unsupervised hierarchical clustering based on co-expression correlation with disease association (**Figures 2A, 3A**). Our analysis revealed how gene expression programs in the heart are organized differently in transcriptome and translatome space into modules, or sub-networks, of highly connected genes. Gene Ontology (GO) enrichment analysis suggest that significant genes in a number of modules localize to common cellular components, such as the extracellular matrix (ECM) and associated proteins (RNA1, RNA7, Ribo1, Ribo2, and Ribo4), the cytoskeleton, related membrane ruffling (RNA4 and RNA6) and the cell cortex (Ribo8, Ribo10), the Golgi apparatus (RNA2), the nucleosome (Ribo12), or the various sub-compartments of the mitochondrion (RNA13, RNA14, RNA15, Ribo13, and Ribo16) (**Supplementary Table 5**). These modules had one or more related molecular function or were associated with shared biological processes.

3.2.1. Correlation With the Cardiac Pathophysiology of Remodeling

We calculated the module correlations to disease association and HW/BW ratios, and clustered them to identify sub-networks associated with patho-physiological features and pre-clinical symptoms of cardiac hypertrophy. Five RNA modules and five Ribo modules had a positive correlation to the pathological model and to increased HW/BW ratio, and almost all were significant (RNA1 to RNA5, Ribo1 to Ribo5, **Figures 2A**, **3A**). We also observed a negative correlation to the pathological model (i.e., a positive correlation to the physiological model) in five RNA modules and, to a moderate extent, in four Ribo modules (RNA13 to RNA17, Ribo14 to Ribo17). RNA1 to RNA5 (Ribo1 to Ribo5) were activated after pressure overload in the pathological model, whereas RNA13 to RNA17 (Ribo14 to Ribo17) were repressed, when looking at the module eigengenes (**Figures 2B,C**, **3B,C**, and **Supplementary Figures 4A,B 5A,B**). An eigengene is a representative of the standardized module expression values across all samples. Eigengenes have been largely regarded as robust biomarkers (Oldham et al., 2008; Johnson et al., 2018; Zhang et al., 2018; Di et al., 2019). The strong association between RNA1 to RNA5 (Ribo1 to Ribo5), on the one hand, and that of RNA13 to RNA17 (Ribo14 to Ribo17), on the other hand, suggests a synchronized expression dynamics characterized by an increased role of genes encoding ECM and cytoskeletal proteins, and a diminished or altered contribution from mitochondrial translation, metabolic pathways of carbohydrate, fat, and protein metabolism, as well as

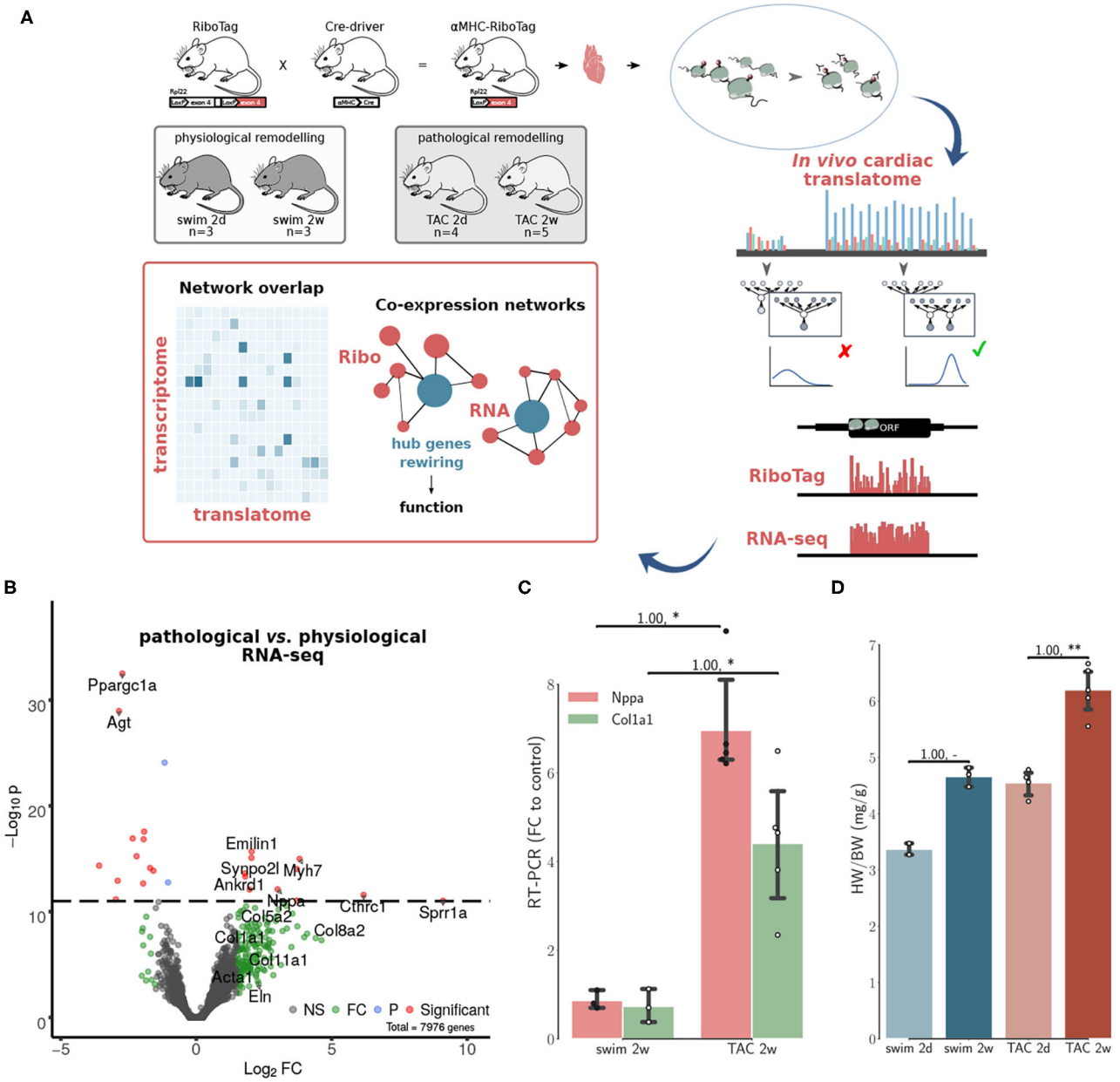

FIGURE 1 | Co-expression networks of transcriptional and translational regulation in cardiac remodeling. **(A)** The RiboTag system is used to find the cardiac translatome of 15 animals from two models at two time points. A final set of translated genes is inferred using graphical Bayesian models. Co-expression networks are then constructed separately for *in vivo* Ribo-seq (translatome from RiboTag data) and RNA-seq (transcriptome) data. The networks were used to identify hub genes associated with cardiac remodeling, and how changes in connectivity (network rewiring) are associated with differential functionality in transcriptome vs. translatome, see also **Supplementary Figure 3**. **(B)** Volcano plot displaying the distribution of all genes with relative abundance ($\log_2$ RNA-seq pathological vs. physiological) plotted against significance level, showing significantly increased and decreased genes during *in vivo* pathological stress. Top genes (red) are highlighted. A number of genes (green) significant at a threshold of 0.05 are also highlighted. NS, Non-significant; FC, Fold change only significant; P, *p*-value only significant; Significant, FC+P. **(C)** RT-PCR fold changes relative to control animals (2w) for two important markers of hypertrophy and fibrosis, corroborating evidence from RNA-seq shown in **(B)**. **(D)** Heart-weight-to-body-weight ratio (HW/BW) after 2d and 2w of exercise in the physiological model, 2d and 2w after transverse aortic constriction (TAC) surgery. For **(C,D)**, N≥3 at each time point. Significance was measured using Mann–Whitney *U*-test, at a threshold of * < 0.05 (** < 0.01, *p*-values may be affected by the small sample size, for this reason we also report effect size using non-directional rank-biserial correlation). For **(D)** only, Kruskall–Wallis test $H = 10.9$, $p = 0.01$.

oxidative phosphorylation. RNA6 to RNA12 (Ribo6 to Ribo13) showed a dynamic association pattern to either the 2d or the 2w time points, uncovering the transcriptional and translational heterogeneity in response to pressure overload (**Figures 2A**, **3A**, and **Supplementary Figures 4C**, **5C**). The significance of these associations was highly consistent, and when we looked

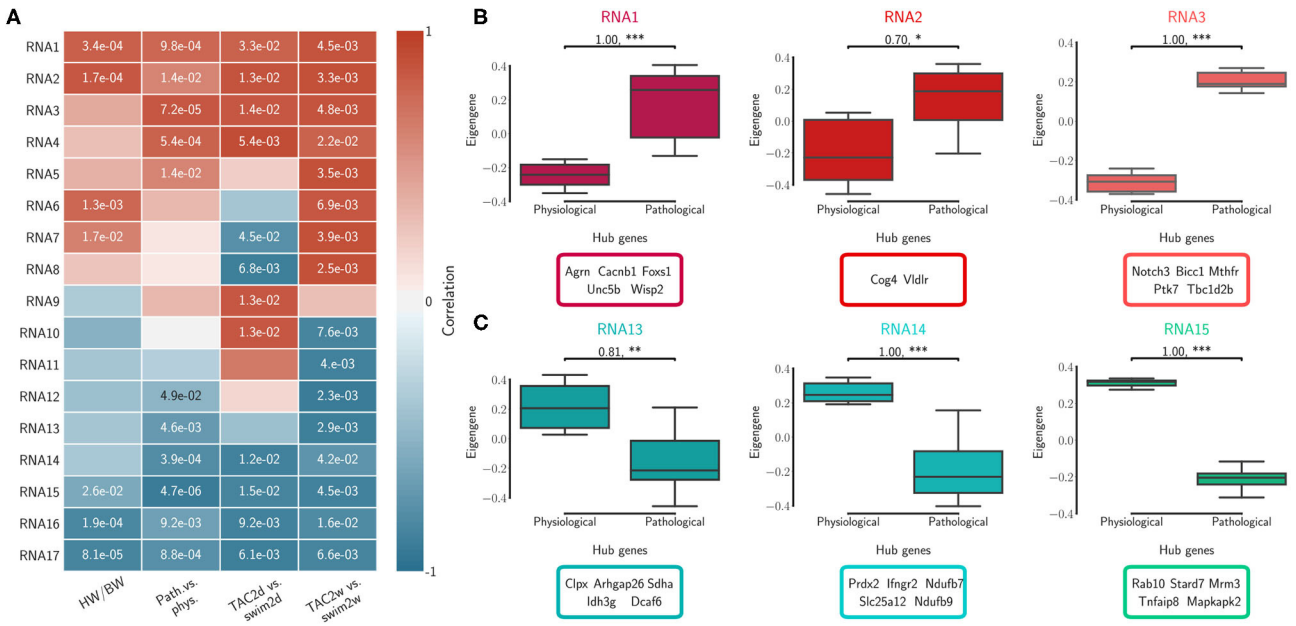

FIGURE 2 | Transcriptome (RNA) network modules were clustered to assess relatedness based on correlation of co-expressed eigengenes. **(A)** Biweight midcorrelation and Student correlation *p*-values between eigengene module expression and disease association as well as heart-weight-to-body-weight ratio. **(B)** The first three module expression profiles and top ranked hub genes that are positively correlated with pathological cardiac remodeling. **(C)** The first three module expression profiles and top ranked hub genes that are negatively correlated with pathological cardiac remodeling (i.e., positively correlated with the physiological model). Significance was measured using a one-sided Mann-Whitney *U*-test, at a threshold of * < 0.05 (** < 0.01, *** < 0.001, *p*-values may be affected by the small sample size, for this reason we also report effect size using non-directional rank-biserial correlation). Up to five hub genes are shown.

at the correlation between module membership and gene significance, we found that the strongest and most significant associations were for the top five modules, particularly for RNA (RNA1 to RNA5) (**Supplementary Table 6**). In addition, for modules RNA6 to RNA12, gene significance at 2w was more often and more strongly correlated with module membership, suggesting that driver genes are associated with the later time points. On the contrary, for most modules Ribo6 to Ribo13, the association was observed with the earlier time points, supporting a higher relative contribution of translational control at 2d, consistent with a rapid translational response to stress (Doroudgar et al., 2019). In each module, we also identified intramodular hub genes, which are highly co-expressed genes with respect to all other genes in the same module, and may thus function as key components of the hypertrophic response (**Figures 2B,C**, **3B,C**, and **Supplementary Figures 4**, **5**). Finally, markers of the fetal gene program were found in sub-networks correlated with the pathological models (Taegtmeyer et al., 2010) (**Supplementary Table 4**): *Nppa* (RNA1, Ribo1), *Nppb* (RNA3, Ribo7), or *Myh7* (RNA3, Ribo1). While *Myh6* was found in RNA15 and Ribo16, consistent with the observed known "gene switches," we observed the presence of several other genes clustered in RNA1, Ribo4 or Ribo2 (*Myh7b*, *Myh10*, *Myh11*, and *Myh14*), and whose clinical significance has not yet been described in the context of cardiac hypertrophy. These switches were also observed for *Glut1* and *Glut4*: *Slc2a8* (RNA1, Ribo1), *Slc2a1* (RNA10, Ribo10), *Slc2a4* (RNA14, Ribo3), or *Slc2a12* (RNA15, Ribo17); and for *Myc*: *Mycn* (RNA1, Ribo4), and *Myc* (RNA17, Ribo10).

3.2.2. Co-expression Networks Uncover Hub Genes Not Found by Differential Expression

Unsupervised hierarchical clustering based on hub gene expression showed that the top interacting genes serve as a molecular signature to differentiate physiological and pathological models of cardiac hypertrophy (**Supplementary Figures 6A**, **7A**). We then compared these observations with results from differential translational-efficiency (DTE) and differential connectivity (DC) analyses (**Supplementary Figure 8** and **Supplementary Table 7** for extended DTE results). Although a large number of non-significant genes in DTE showed a higher DC between RNA-seq and Ribo-seq networks, hub genes remain relatively unchanged in DC (**Supplementary Table 4**). Many hub genes from modules with positive (RNA1 to RNA5, and Ribo1 to Ribo5) or negative correlation (RNA13 to RNA17, and Ribo14 to Ribo17) were up-/down-regulated in the pathological model, respectively (**Supplementary Figure 6B**). Similar observations were made considering the different time points and the varying correlations, using hub genes from modules RNA6 to RNA12, and Ribo6 to Ribo13 (**Supplementary Figure 7B**). While the number of differentially regulated genes is much larger, we found hub genes from co-expression correlation only that were not identified in DTE.

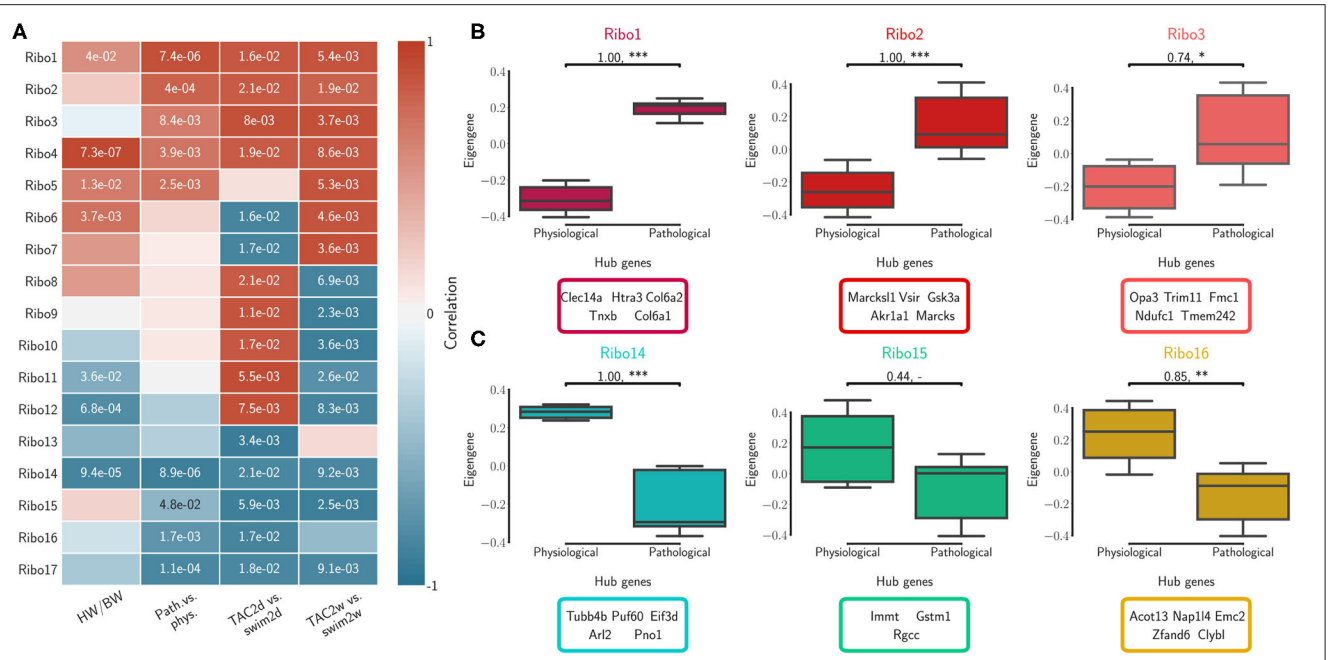

FIGURE 3 | Translatome (Ribo) network modules were clustered to assess relatedness based on correlation of co-expressed eigengenes. **(A)** Biweight midcorrelation and Student correlation *p*-values between eigengene module expression and disease association as well as heart-weight-to-body-weight ratio. **(B)** The first three module expression profiles and top ranked hub genes that are positively correlated with pathological cardiac remodeling. **(C)** The first three module expression profiles and top ranked hub genes that are negatively correlated with pathological cardiac remodeling (i.e., positively correlated with the physiological model). Significance was measured using a one-sided Mann-Whitney *U*-test, at a threshold of * < 0.05 (** < 0.01, *** < 0.001, *p*-values may be affected by the small sample size, for this reason we also report effect size using non-directional rank-biserial correlation). Up to five hub genes are shown.

3.2.3. Co-expression Networks Describe the Organization of the Heart Transcriptome and Translatome

We investigated the degree of preservation between RNA network structure and Ribo co-expression network, and the amount of overlap between sub-networks. Preservation is based on density and connectivity measures (Langfelder et al., 2011), and uses the correlation structure of the networks to identify differences between RNA-seq and Ribo-seq. We identified modules that were highly correlated/anti-correlated with the pathological model that were partially shared across transcriptome and translatome (RNA1, RNA3, and RNA4 overlap with Ribo1, Ribo2 and Ribo4; RNA14, RNA15, and RNA17 overlap with Ribo14, Ribo16, and Ribo17) (**Figure 4A**). These modules may represent ubiquitous processes and mechanisms of response to stress. Five RNA clusters were found to be highly preserved, and six moderately preserved, at the translatome level (**Figure 4B**). Six more RNA modules, one of which was activated (RNA5) after pressure overload in the pathological model, two of which were repressed (RNA13 and RNA16), as well as RNA7, RNA8, and RNA11, which showed differential activation/repression at 2d and 2w, had no or little gene overlap with translatome modules, and did not have a preserved network structure, suggesting that the transcriptome does not capture all key changes occurring in the heart during early hypertrophy.

To uncover how the heart transcriptome and translatome networks are rewired in response to stress, we highlighted genes which had the most dynamic neighborhoods. The Dn score, or dynamic neighborhoods score, captures changes in connectivity of a gene, even when its intramodular connectivity remains similar in RNA-seq and Ribo-seq networks, and is thus better suited than DC to highlight hub genes associated with potential regulatory mechanisms. The most rewired genes all belonged to the top preserved modules (**Figure 4C**). Among these, we highlighted two candidates, *Fam210a* and *Psmd12*, identified earlier (**Supplementary Figure 6**), which were strongly rewired, and whose immediate interactors were enriched in different GO terms between RNA-seq and Ribo-seq networks (**Figure 4D** and **Supplementary Figure 9**). *Fam210a* is a conserved transmembrane protein localized in the mitochondria, containing a mitochondrial targeting signal peptide (MitoCarta2.0 mouse), a DUF1279 (Domain of Unknown Function) domain with a transmembrane peptide, and a coiled coil at the C-terminus (InterPro). It is mostly expressed in the heart [1.72534] and skeletal muscle [1.57137] (Standardized values, BioGPS Mouse Cell Type and Tissue Gene Expression Profiles), and is thought to play a role in modulating muscle and bone biology (Tanaka et al., 2018), but its function in the heart and its molecular mechanisms are unknown. *Psmd12*, encoding the non-ATPase regulatory subunit 12 of the 26S proteasome, is better characterized. The 26S proteasome is a multiprotein complex involved in the

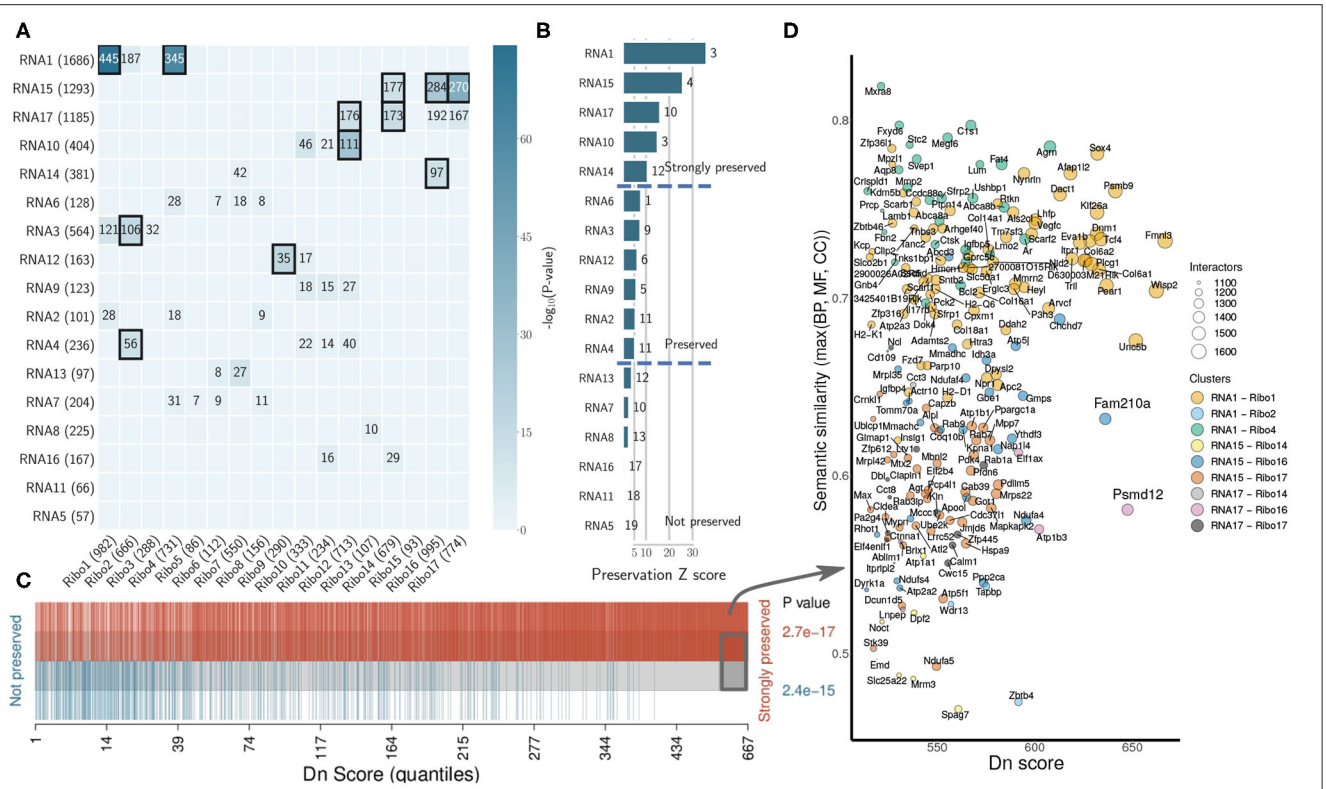

FIGURE 4 | The organization of the heart transcriptome and translatome under cardiac remodeling. **(A)** Cross-tabulation of RNA-seq (rows) and Ribo-seq (columns) modules. Each row and column is labeled by the corresponding module and its size (number of genes). Counts of genes in the intersection is shown for modules sharing a significant overlap ($p < 0.01$, two-tailed Fisher's exact test). Overlap with $p <$ 1e-10 are highlighted. **(B)** Preservation Z scores (mean of *Z*-scores computed for density and connectivity measures) for the RNA-seq modules in the Ribo-seq network. The vertical lines indicate the threshold for moderate preservation (5) and strong preservation (10). The *medianRank* statistic is indicated to the right of each bar. **(C)** Barcode plot showing the enrichment of genes based on their dynamic neighborhoods score (Dn score), compared to their association with modules that are either strongly or not preserved. Genes are ranked according to their Dn score, and colored based on their association. The gray box toward the right indicates genes for which Dn score is greater than two standard deviations above the average. To test whether genes belonging to preserved or not preserved clusters are highly ranked in terms of Dn score, the camera test from the limma R package was performed and *p*-values are shown. **(D)** Scatter plot of the most rewired genes (gray box in **C**). A semantic similarity analysis was performed for each ontology (BP, MF, and CC) and every gene using its immediate interactors in the RNA-seq and Ribo-seq networks, respectively, and the maximum semantic similarity is reported on the y-axis. The set of all translated genes was used as background. Dots are colored according to their cluster association, and the size represents the total number of interactors, or connected genes, in a hypothetical joint network combining the respective RNA-seq and Ribo-seq modules. Semantic similarity was calculated using the GOSemSim R package. Only genes in modules with significant overlap are reported. BP, Biological Process; MF, Molecular Function; CC, Cellular Component.

ATP-dependent degradation of ubiquitinated proteins, and thus plays a key role in protein homeostasis. *Psmd12* is associated with several pathways, including Regulation of Apoptosis, Stabilization of p53, and p53-Dependent/Independent G1 DNA Damage Response (WikiPathways, Reactome). The tumor suppressor Trp53 (p53) regulates cell growth and fate, and its role in the heart is well-known. Psmd12 is also associated with inflammation [2.13579] and hypertrophy [2.10896] (Standardized values, CTD Comparative Toxicogenomics Database). Overall, module-eigengene association to disease phenotypes has led to the identification of highly rewired hub genes that may function as drivers of cardiac remodeling. These hub genes are potentially involved in related, but different molecular pathways or functions, suggesting some form of translational control that may not be immediately apparent from DTE analyses.

4. DISCUSSION

In this study, a model of left ventricular pressure overload was used to mimic hypertrophy induced by systemic hypertension and aortic stenosis, and compared with a physiological model of exercise-induced cardiac growth. Transcriptional and translational co-expression networks uncovered *in vivo* changes in the heart occurring within 2 weeks of a transverse aortic constriction (TAC) surgery, revealing the complexity of the organization as well as unappreciated genes that may act as key drivers of the hypertrophic response.

Physiologic and pathophysiologic stimuli act upon the cell membrane and work their way through various cascades to mediate gene expression, translational control and protein levels (Haque and Wang, 2017). As expected, pressure overload was associated with profound changes in the composition

of the extracellular matrix (ECM), which were reflected by a sub-clustering and a synchronized expression dynamics of ECM-, and cytoskeletal-related genes, in both transcriptome and translatome networks (**Figures 2**, **3**, **Supplementary Figures 4**, **5**, and **Supplementary Table 4**). A marked upregulation of genes encoding ECM proteins has previously been observed during the transition from stable cardiac hypertrophy to heart failure (Boluyt et al., 1994). Our results indicate that concerted dynamic changes occur early *in vivo* after stimuli, and are likely to be implicated in transducing molecular signals driving the maladaptive response. Concurrently to these observations, modules anti-correlated to the pathological model showed a diminished expression or altered contribution of mitochondrial, electron transport complex and oxidative phosphorylation genes. These modules (1–4, and 14, 15, 17) were also among the most preserved (**Figure 4**), suggesting the existence of stable sub-network structures, which could be associated with ubiquitous mechanisms of response to stress, although genes associated with these may or may not show significant changes in translational/transcriptional efficiency (**Supplementary Figures 6**, **7**). Co-expression network and differential translational-efficiency (DTE) analyses are based on different assumptions (Langfelder and Horvath, 2008, 2012). In co-expression networks, the top genes are the most connected genes, based on the correlation structure.

In this study, we identified a number of hub genes that may function as molecular drivers of cardiac remodeling, many of which were recently described for their putative role in myofiber hypertrophy, cardiac inflammation or injury, such as *Gsk3a* (Ribo2) (Sugden et al., 2008; Zhou et al., 2016), *Cand2* (Ribo2) (Sandmann et al., 2018), *Rptor* (Ribo9) (Shende et al., 2011), *Lonp1* (Ribo9) (Venkatesh et al., 2019), *Ubr4* (Ribo9) (Hunt et al., 2019), *C5ar1* (RNA10) (Natarajan et al., 2018), *S100a4* (RNA10) (Doroudgar et al., 2016), or *Phb2* (Ribo13) (Wu D. et al., 2020). We also identified hub genes in the transcriptional network that were rewired in the translational network, and associated with semantically different subsets of enriched terms (**Figure 4**). Notably, we highlighted the presence of two hub genes that were rewired under hypertrophic stimuli, *Fam210a* (RNA15, Ribo16), and *Psmd12* (RNA17, Ribo16). *Fam210a*, a gene of previously unknown function, has been described as a musculoskeletal modulator (Tanaka et al., 2018). In humans, a prior study reported that *Fam210a* (*C18orf19*) was the strongest candidate partner protein of *Atad3a* (ATPase Family AAA Domain Containing 3A), which was also found in the same modules (RNA15 and/or Ribo16), along with 60 (out of 153) interacting proteins identified by Orbitrap MS analysis and quantified by SILAC labeling (He et al., 2012). *Atad3a* is essential for mitochondrial metabolism and translation, and has been implicated in several processes in mitochondria. More recent work, which we discovered while this manuscript was under review, has shown how the miR-574-*Fam210a* axis regulates mitonchondrial-encoded protein expression in cardiac pathological remodeling (Wu J. et al., 2020). Taken together, these results suggest that *Fam210a* could modulate translation of mitochondrial-encoded electron-transport chain proteins, and play a yet undescribed role in cardiac muscle adaptation and growth. The biological importance of *Psmd12* as a scaffolding subunit in proteasome function has been described earlier in the context of neuronal development, but remains un-documented in the heart. The ubiquitin proteasome system (UPS) is critical in preventing accumulation of damaged and misfolded proteins, and has been implicated in a number of cardiac proteinopathies and heart failure (Pagan et al., 2013; Cacciapuoti, 2014; Maejima, 2020). Our results support the existence of transcriptional/translational regulatory processes affecting the or affected by proteasome function in the pathogenesis of cardiac hypertrophy.

In summary, these results highlight the organization of distinct molecular processes into sub-networks of co-expressed genes, and describe how transcriptome and translatome signatures are orchestrated into functional modules associated with the early stages of cardiac remodeling. Our results constitute a valuable resource to study *in vivo* cardiac regulatory networks, and a first step toward the identification and characterization of novel proteins involved in cardiac remodeling, hypertrophy and heart failure.

DATA AVAILABILITY STATEMENT

The data generated for this study have been deposited in NCBI's Sequence Read Archive through the BioProject accession numbers PRJNA484227 and PRJNA543399. All raw counts and translation events used in this study are available as **Supporting Information**. RP-BP is publicly available at https://github.com/dieterich-lab/rp-bp under the MIT License. The code for generating co-expression networks is available as **Supplementary Information**.

ETHICS STATEMENT

The animal study was reviewed and approved by Ethikkommission der Med. Fakultät Heidelberg.

AUTHOR CONTRIBUTIONS

EB and CD conceptualized the project. EB conducted formal analyses and interpretation, maintained the software and was in charge of the original draft preparation. CD was in charge of project administration, supervision, contributed to resources, data analysis and interpretation. SD and MV contributed to resources, data acquisition, and investigation. ER, LJ, and TH performed animal experiments. HK provided funding acquisitions. EB, CD, SD, and MV contributed to review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

EB and CD acknowledge support by the Klaus Tschira Stiftung gGmbH [00.219.2013]. CD, SD, and MV acknowledge the DZHK (German Centre for Cardiovascular Research) Partner Site Heidelberg/Mannhein. MV was supported by

the Deutsche Forschungsgemeinschaft, Plus 3 Programme of the Boehringer Ingelheim Foundation (BISI) and Baden Württemberg Stiftung. SD is supported by the DZHK (German Centre for Cardiovascular Research) Excellence Program.

ACKNOWLEDGMENTS

CD and EB thank the Dieterich Lab group members for providing feedback on the manuscript. This manuscript has been released as a pre-print at bioRxiv (Boileau et al., 2020).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.583124/full#supplementary-material

Supplementary Figure 1 | A survey of active translation in 15 mouse hearts representing models of cardiac remodeling.

Supplementary Figure 2 | A survey of active translation in 15 mouse hearts: open reading frames (ORFs) translated in cardiac cells.

Supplementary Figure 3 | Co-expression networks of transcriptional and translational regulation in cardiac remodeling: schematic workflow.

Supplementary Figure 4 | Translatome (RNA) network modules were clustered to assess relatedness based on correlation of co-expressed eigengenes.

Supplementary Figure 5 | Translatome (Ribo) network modules were clustered to assess relatedness based on correlation of co-expressed eigengenes.

Supplementary Figure 6 | Hub genes provide a different molecular signature for differentiating physiological and pathological cardiac remodeling.

Supplementary Figure 7 | Hub genes provide a different molecular signature for differentiating physiological and pathological cardiac remodeling.

Supplementary Figure 8 | Differential connectivity and differential translational-efficiency.

Supplementary Figure 9 | Rewired genes are associated with different biological processes, molecular functions and cellular components in transcriptome vs. translatome.

Supplementary Table 1 | Open reading frame predictions per sample, all categories, Rp-Bp output.

Supplementary Table 2 | Open reading frame predictions, unique (translated genes used in this study).

Supplementary Table 3 | Transcriptome and translatome read counts used as input for network analyses.

Supplementary Table 4 | Transcriptome and translatome co-expression modules.

Supplementary Table 5 | Gene ontology enrichment of transcriptome and translatome co-expression modules.

Supplementary Table 6 | Transcriptome and translatome co-expression modules correlation association and significance.

Supplementary Table 7 | Extended differential translational-efficiency (DTE) results.

Supplementary Information | Code for generating co-expression networks.

REFERENCES

Alexa, A., and Rahnenfuhrer, J. (2018). *topGO: Enrichment Analysis for Gene Ontology*. R package version 2.28.0.

Anders, S., Pyl, P. T., and Huber, W. (2014). HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638

Bernardo, B. C., Ooi, J. Y. Y., Weeks, K. L., Patterson, N. L., and McMullen, J. R. (2018). Understanding key mechanisms of exercise-induced cardiac protection to mitigate disease: Current knowledge and emerging concepts. *Physiol. Rev.* 98, 419–475. doi: 10.1152/physrev.00043.2016

Boileau, E., Doroudgar, S., Riechert, E., Jürgensen, L., Ho, T. C., Katus, H. A., et al. (2020). A multi-network comparative analysis of transcriptome and translatome in cardiac remodeling. *bioRxiv [Preprint]*. doi: 10.1101/2020.07.01.181743

Boluyt, M., O'Neill, L., Meredith, A., Bing, O., Brooks, W., Conrad, C., et al. (1994). Alterations in cardiac gene expression during the transition from stable hypertrophy to heart failure. Marked upregulation of genes encoding extracellular matrix components. *Circ. Res.* 75, 23–32. doi: 10.1161/01.RES.75.1.23

Boström, P., Mann, N., Wu, J., Quintero, P. A., Plovie, E. R., Panáková, D., et al. (2010). C/ebpb; controls exercise-induced cardiac growth and protects against pathological cardiac remodeling. *Cell* 143, 1072–1083. doi: 10.1016/j.cell.2010.11.036

Cacciapuoti, F. (2014). Role of ubiquitin-proteasome system (UPS) in left ventricular hypertrophy (LVH). *Am. J. Cardiovasc. Dis.* 4, 1–5.

Chothani, S., Schofer, S., Adami, E., Viswanathan, S., Widjaja, A. A., Langley, S. R., et al. (2019). Widespread translational control of fibrosis in the human heart by RNA-binding proteins. *Circulation* 140, 937–951. doi: 10.1161/CIRCULATIONAHA.119.039596

Di, Y., Chen, D., Yu, W., and Yan, L. (2019). Bladder cancer stage-associated hub genes revealed by WGCNA co-expression network analysis. *Hereditas* 156:7. doi: 10.1186/s41065-019-0083-y

Dobin, A., Davis, C., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). Star: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635

Dodt, M., Roehr, J., Ahmed, R., and Dieterich, C. (2012). Flexbar-flexible barcode and adapter processing for next-generation sequencing platforms. *Biology* 1, 895–905. doi: 10.3390/biology1030895

Doroudgar, S., Hofmann, C., Boileau, E., Malone, B., Riechert, E., Gorska, A. A., et al. (2019). Monitoring cell-type specific gene expression using ribosome profiling *in vivo* during cardiac hemodynamic stress. *Circ. Res.* 125, 431–448. doi: 10.1161/CIRCRESAHA.119.314817

Doroudgar, S., Quijada, P., Konstandin, M., Ilves, K., Broughton, K., Khalafalla, F. G., et al. (2016). S100a4 protects the myocardium against ischemic stress. *J. Mol. Cell. Cardiol.* 100, 54–63. doi: 10.1016/j.yjmcc.2016.10.001

Evangelista, F., Brum, P., and Krieger, J. (2003). Duration-controlled swimming exercise training induces cardiac hypertrophy in mice. *Braz. J. Med. Biol. Res.* 36, 1751–1759. doi: 10.1590/S0100-879X2003001200018

Goenawan, I. H., Bryan, K., and Lynn, D. J. (2016). DyNet: visualization and analysis of dynamic molecular interaction networks. *Bioinformatics* 32, 2713–2715. doi: 10.1093/bioinformatics/btw187

Haque, Z. K., and Wang, D.-Z. (2017). How cardiomyocytes sense pathophysiological stresses for cardiac remodeling. *Cell. Mol. Life Sci.* 74, 983–1000. doi: 10.1007/s00018-016-2373-0

He, J., Cooper, H. M., Reyes, A., Di, Re, M., Sembongi, H., Litwin, T. R., et al. (2012). Mitochondrial nucleoid interacting proteins support mitochondrial protein synthesis. *Nucleic Acids Res.* 40, 6109–6121. doi: 10.1093/nar/gks266

Hunt, L. C., Stover, J., Haugen, B., Shaw, T. I., Li, Y., Pagala, V. R., et al. (2019). A key role for the ubiquitin ligase ubr4 in myofiber hypertrophy in drosophila and mice. *Cell Rep.* 28, 1268–1281. doi: 10.1016/j.celrep.2019.06.094

Johnson, E. C. B., Dammer, E. B., Duong, D. M., Yin, L., Thambisetty, M., Troncoso, J. C., et al. (2018). Deep proteomic network analysis of Alzheimer's disease brain reveals alterations in RNA binding proteins and RNA splicing associated with disease. *Mol. Neurodegen.* 13:52. doi: 10.1186/s13024-018-0282-4

Johnson, W. E., Li, C., and Rabinovic, A. (2006). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8, 118–127. doi: 10.1093/biostatistics/kxj037

Kmietczyk, V., Riechert, E., Kalinski, L., Boileau, E., Malovrh, E., Malone, B., et al. (2019). M6A-mrna methylation regulates cardiac gene expression and cellular growth. *Life Sci. Alliance* 2:e201800233. doi: 10.26508/lsa.201800233

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform*. 9:559. doi: 10.1186/1471-2105-9-559

Langfelder, P., and Horvath, S. (2012). Fast r functions for robust correlations and hierarchical clustering. *J. Stat. Softw*. 46, 1–17. doi: 10.18637/jss.v046.i11

Langfelder, P., Luo, R., Oldham, M. C., and Horvath, S. (2011). Is my network module preserved and reproducible? *PLoS Comput. Biol*. 7:e1001057. doi: 10.1371/journal.pcbi.1001057

Langmead, B., and Salzberg, S. (2012). Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Ling, S. S. M., Chen, Y.-T., Wang, J., Richards, A. M., and Liew, O. W. (2017). Ankyrin repeat domain 1 protein: a functionally pleiotropic protein with cardiac biomarker potential. International *J. Mol. Sci*. 18:1362. doi: 10.3390/ijms18071362

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DEseq2. *Genome Biol*. 15:550. doi: 10.1186/s13059-014-0550-8

Maejima, Y. (2020). The critical roles of protein quality control systems in the pathogenesis of heart failure. *J. Cardiol*. 75, 219–227. doi: 10.1016/j.jjcc.2019.09.019

Malone, B., Atanassov, I., Aeschimann, F., Li, X., Grosshans, H., and Dieterich, C. (2017). Bayesian prediction of RNA translation from ribosome profiling. *Nucleic Acids Res*. 45, 2960–2972. doi: 10.1093/nar/gkw1350

Natarajan, N., Abbas, Y., Bryant, D. M., Gonzalez-Rosa, J. M., Sharpe, M., Uygur, A., et al. (2018). Complement receptor C5AR1 plays an evolutionarily conserved role in successful cardiac regeneration. *Circulation* 137, 2152–2165. doi: 10.1161/CIRCULATIONAHA.117.030801

Oldham, M. C., Konopka, G., Iwamoto, K., Langfelder, P., Kato, T., Horvath, S., et al. (2008). Functional organization of the transcriptome in human brain. *Nat. Neurosci*. 11, 1271–1282. doi: 10.1038/nn.2207

Pagan, J., Seto, T., Pagano, M., and Cittadini, A. (2013). Role of the ubiquitin proteasome system in the heart. *Circ. Res*. 112, 1046–1058. doi: 10.1161/CIRCRESAHA.112.300521

Rockman, H., Ross, R., Harris, A., Knowlton, K., Steinhelper, M., Field, L., et al. (1991). Segregation of atrial-specific and inducible expression of an atrial natriuretic factor transgene in an *in vivo* murine model of cardiac hypertrophy. *Proc. Natl. Acad. Sci.U.S.A*. 88, 8277–8281. doi: 10.1073/pnas.88.18.8277

Rodriguez, J., Rodriguez-Rivas, J., Domenico, T., Vázquez, J., Valencia, A., and Tress, M. (2018). APPRIS 2017: principal isoforms for multiple gene sets. *Nucleic Acids Res*. 46, D213–D217. doi: 10.1093/nar/gkx997

Sandmann, C., Gorska, A., Hofmann, C., Katus, H., and Volkers, M. (2018). Abstract 15125: Cand2, a translationally18 regulated muscle specific protein, mediates mTORC1 dependent cardiac hypertrophy. *Circulation* 138(Suppl_1):A15125.

Shende, P., Plaisance, I., Morandi, C., Pellieux, C., Berthonneche, C., Zorzato, F., et al. (2011). Cardiac raptor ablation impairs adaptive hypertrophy, alters metabolic gene expression, and causes heart failure in mice. *Circulation* 123, 1073–1082. doi: 10.1161/CIRCULATIONAHA.110.977066

Shimizu, I., and Minamino, T. (2016). Physiological and pathological cardiac hypertrophy. *J. Mol. Cell. Cardiol*. 97, 245–262. doi: 10.1016/j.yjmcc.2016.06.001

Sugden, P. H., Fuller, S. J., Weiss, S. C., and Clerk, A. (2008). Glycogen synthase kinase 3 (gsk3) in the heart: a point of integration in hypertrophic signalling and a therapeutic target? A critical analysis. *Brit. J. Pharmacol*. 153(Suppl 1), S137–S153. doi: 10.1038/sj.bjp.0707659

Taegtmeyer, H., Sen, S., and Vela, D. (2010). Return to the fetal gene program. *Ann. N. Y. Acad. Sci*. 1188, 191–198. doi: 10.1111/j.1749-6632.2009.05100.x

Tanaka, K.-I., Xue, Y., Nguyen-Yamamoto, L., Morris, J. A., Kanazawa, I., Sugimoto, T., et al. (2018). Fam210a is a novel determinant of bone and muscle structure and strength. *Proc. Natl. Acad. Sci. U.S.A*. 115, E3759–E3768. doi: 10.1073/pnas.1719089115

The Gene Ontology Consortium (2018). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res*. 47, D330–D338. doi: 10.1093/nar/gky1055

van Eldik, W., den Adel, B., Monshouwer-Kloots, J., Salvatori, D., Maas, S., van der Made, I., et al. (2017). Z-disc protein CHAPb induces cardiomyopathy and contractile dysfunction in the postnatal heart. *PLoS ONE* 12:e0189139. doi: 10.1371/journal.pone.0189139

Venkatesh, S., Li, M., Saito, T., Tong, M., Rashed, E., Mareedu, S., et al. (2019). Mitochondrial lonp1 protects cardiomyocytes from ischemia/reperfusion injury *in vivo*. *J. Mol. Cell. Cardiol*. 128, 38–50. doi: 10.1016/j.yjmcc.2018.12.017

Völkers, M., Toko, H., Doroudgar, S., Din, S., Quijada, P., Joyo, A., et al. (2013). Pathological hypertrophy amelioration by pras40-mediated inhibition of mTORC1. *Proc. Natl. Acad. Sci. U.S.A*. 110, 12661–12666. doi: 10.1073/pnas.1301455110

Wang, Y., Wisloff, U., and Kemi, O. (2010). Animal models in the study of exercise-induced cardiac hypertrophy. *Physiol. Res*. 59, 633–44.

Wu, D., Jian, C., Peng, Q., Hou, T., Wu, K., Shang, B., et al. (2020). Prohibitin 2 deficiency impairs cardiac fatty acid oxidation and causes heart failure. *Cell Death Dis*. 11:181. doi: 10.1038/s41419-020-2374-7

Wu, J., Subbaiah, K. C. V., Jiang, F., Hadaya, O., Mohan, A., Yang, T., et al. (2020). Microrna-574 regulates fam210a expression and influences pathological cardiac remodeling. *bioRxiv*. doi: 10.1101/2020.01.09.900712

Zhang, Q., Ma, C., Gearing, M., Wang, P. G., Chin, L.-S., and Li, L. (2018). Integrated proteomics and network analysis identifies protein hubs and network alterations in Alzheimer's disease. *Acta Neuropathol. Commun*. 6:19. doi: 10.1186/s40478-018-0524-2

Zhou, J., Ahmad, F., Parikh, S., Hoffman, N. E., Rajan, S., Verma, V. K., et al. (2016). Loss of adult cardiac myocyte GSK-3 leads to mitotic catastrophe resulting in fatal dilated cardiomyopathy. *Circ. Res*. 118, 1208–1222. doi: 10.1161/CIRCRESAHA.116.308544

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 10 December 2020
doi: 10.3389/fgene.2020.579215

Combining Bioinformatics Techniques to Study the Key Immune-Related Genes in Abdominal Aortic Aneurysm

*Han Nie[1], Jiacong Qiu[2], Si Wen[3] and Weimin Zhou[1]**

[1] Department of Vascular Surgery, The Second Affiliated Hospital of Nanchang University, Nanchang, China, [2] Divison of Vascular Surgery, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China, [3] Xinjian District People's Hospital of Jiangxi Province, Jiangxi, China

OPEN ACCESS

Edited by:
Maarten M. G. van den Hoogenhof, Heidelberg University Hospital, Germany

Reviewed by:
Tobias Jakobi, Heidelberg University Hospital, Germany
Mirka Rabajdova, University of Pavol Jozef Šafárik, Slovakia

***Correspondence:**
Weimin Zhou
weiminzhouncu@163.com

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *02 July 2020*
Accepted: *10 November 2020*
Published: *10 December 2020*

Citation:
Nie H, Qiu J, Wen S and Zhou W (2020) Combining Bioinformatics Techniques to Study the Key Immune-Related Genes in Abdominal Aortic Aneurysm. Front. Genet. 11:579215. doi: 10.3389/fgene.2020.579215

Approximately 13,000 people die of an abdominal aortic aneurysm (AAA) every year. This study aimed to identify the immune response-related genes that play important roles in AAA using bioinformatics approaches. We downloaded the GSE57691 and GSE98278 datasets related to AAA from the Gene Expression Omnibus database, which included 80 AAA and 10 normal vascular samples. CIBERSORT was used to analyze the samples and detect the infiltration of 22 types of immune cells and their differences and correlations. The principal component analysis showed significant differences in the infiltration of immune cells between normal vascular and AAA samples. High proportions of CD4+ T cells, activated mast cells, resting natural killer cells, and 12 other types of immune cells were found in normal vascular tissues, whereas high proportions of macrophages, CD8+ T cells, resting mast cells, and six other types of immune cells were found in AAA tissues. In the selected samples, we identified 39 upregulated (involved in growth factor activity, hormone receptor binding, and cytokine receptor activity) and 133 downregulated genes (involved in T cell activation, cell chemotaxis, and regulation of immune response mediators). The key differentially expressed immune response-related genes were screened using the STRING database and Cytoscape software. Two downregulated genes, *PI3* and *MAP2K1*, and three upregulated genes, *SSTR1, GPER1,* and *CCR10*, were identified by constructing a protein–protein interaction network. Functional enrichment of the differentially expressed genes was analyzed, and the expression of the five key genes in AAA samples was verified using quantitative polymerase chain reaction, which revealed that *MAP2K1* was downregulated in AAA, whereas *SSTR1, GEPR1,* and *CCR10* were upregulated; there was no significant difference in PI3 expression. Our study shows that normal vascular and AAA samples can be distinguished via the infiltration of immune cells. Five genes, *PI3, MAP2K1, SSTR1, GPER1,* and *CCR10*, may play important roles in the development, diagnosis, and treatment of AAA.

Keywords: immune related genes, abdominal aortic aneurysm, bioinformatics, vascular, surgery

INTRODUCTION

With changes in lifestyle, the incidence rates of cardiovascular and cerebrovascular diseases have increased every year, making them a serious public health problem. According to World Health Organization statistics, cardiovascular disease is the leading cause of death worldwide. In 2011, coronary heart disease and cerebrovascular disease caused 13.2 million deaths, accounting for 24% of the total global deaths. It is estimated that by 2030, the number of deaths due to cardiovascular diseases will increase to 23.3 million, and cardiovascular diseases will continue to be the leading cause of death (World Health Organization, 2013). Reportedly, approximately 13,000 people die of abdominal aortic aneurysm (AAA) every year (Rooke et al., 2012).

Current studies suggest that age (more than 65 years), family history, sex (male), and smoking are important risk factors for AAA (Sakalihasan et al., 2018). AAA has become an important cause of death in the elderly who are more than 65 years old and is a serious aortic disease involving irreversible radial dilatation of the abdominal aorta more than 3 cm or 1.5 times the normal diameter due to various reasons (Umebayashi et al., 2018). Most AAA patients are asymptomatic and cannot be treated before the tumor ruptures, or the patient dies (Golledge and Norman, 2011). Aneurysm rupture is an important cause of mortality in patients with AAA (Golledge et al., 2006), which is reportedly as high as 80%. Open surgery and interventional surgery are the main methods to treat AAA. However, they have limited use because the prevalence of AAA is positively correlated with age, and patients with AAA often suffer from other cardio-cerebrovascular diseases, such as heart failure, atherosclerosis, and ischemic cardiomyopathy (Arya et al., 2015).

Studies have shown that if the increase in the aortic diameter can be slowed down by 50%, the annual rates of aortic reconstruction surgery for AAA can be halved. In clinical practice, only patients whose AAAs are greater than 5.5 cm and who are at risk of rupture are treated with open surgery (Golledge et al., 2006), and there is no drug to slow down the development of AAA. The occurrence and development of AAA is a complex process involving multiple factors. It is generally believed that AAA is directly related to atherosclerosis, hypertension, chronic obstructive pulmonary disease, and a variety of proteases, but there is no clear evidence to explain the roles of these factors in the pathogenesis of AAA. The pathophysiological processes of AAA include infiltration of inflammatory cells (Moxon et al., 2010; Gordon and Toursarkissian, 2014), degradation of elastic and collagen fibers, death of smooth muscle cells, defects of the arterial wall, and increased oxidative stress (Newman et al., 1994).

Vascular inflammation is the first event in the development of AAA. In the early stages of the disease, immune cells such as lymphocytes, macrophages, mast cells, neutrophils, and natural killer (NK) cells infiltrate and accumulate in the blood vessels and surrounding tissues, causing a series of inflammatory reactions in the vascular wall (Eliason et al., 2005; Forester et al., 2006; Rateri et al., 2011; Wang et al., 2014a; Yan et al., 2016). The invasion of inflammatory cells often stimulates smooth muscle cells to secrete matrix metalloproteinases, which degrade elastin and collagen, thus reducing the stability of the arterial wall and inducing the apoptosis of vascular smooth muscle cells (Rizzo et al., 1989; Sakalihasan et al., 1996; Culav et al., 1999; Longo et al., 2002, 2005; Keeling et al., 2005). Inflammatory response plays an important role in the immune system and is involved in the occurrence and development of AAA (Yamaguchi et al., 2000; Liu et al., 2015). Indeed, lowering interleukin (IL)-17 levels in animal models slows down the increase in aortic diameter in aneurysms (Chang et al., 2015). Therefore, this study aimed to analyze the infiltration of immune cells in AAA and identify genes related to the immune response that play a role in the development of AAA.

MATERIALS AND METHODS

Data Acquisition and Processing

We searched the Gene Expression Omnibus database[1] and obtained two gene expression datasets, GSE57691 (including 49 AAA samples and 10 normal aortic vessels samples) (Biros et al., 2015) and GSE98278 (Gäbel et al., 2017) (including 31 AAA samples), which included 80 samples from AAAs and 10 samples from normal aortic vessels. We used Limma (Ritchie et al., 2015) and SVA (Leek et al., 2019) packages in R (3.61) to correct the sample data.

Infiltration of Immune Cells in the Samples

To investigate the infiltration of immune cells in AAAs and normal aortic vessels and evaluate and predict the enrichment of immune cells in the samples, we used CIBERSORT (Newman et al., 2015). CIBERSORT is a tool used to deconvolute the expression matrix of immune cell subtypes based on the principle of linear support vector regression. RNA-Seq data were used to estimate the infiltration of immune cells. CIBERSORT analyzed the relative abundance of 22 types of immune infiltrating cells in each sample, including NK cells, T cells, B cells, and macrophages. The 69 samples were screened according to the *P*-value predicted by CIBERSORT ($P < 0.05$).

Principal Component Analysis

We reduced the dimensions of the samples and performed principal component analysis (PCA) of the infiltration of 22 types of immune cells in the samples.

Acquisition of Immune Response-Related Gene Expression Profiles

From ImmPort[2], we downloaded 2,498 immune response-related genes, including those related to antigen-presenting cells, chemokines and their receptors, cytokines and their receptors, interferons, and ILs. We used the Limma package (Ritchie et al., 2015) in R (3.61) to compare gene expression data of immune response-related genes in AAA and normal aortic blood samples downloaded from the Gene Expression Omnibus

[1]http://www.ncbi.nlm.nih.gov/geo/

[2]https://www.immport.org/home

database and extracted information on the expression levels of immune response-related genes in the samples.

Screening of Differentially Expressed Immune Response-Related Genes

After the data were standardized using the Limma software package (Ritchie et al., 2015) in R, we identified immune response-related, differentially expressed genes by comparing normal aortic blood vessel samples with AAA samples. If the change factor was greater than onefold (|fold change| $\geq$ 1) and the corrected *P*-value (false discovery rate) $\leq$ 0.05, the gene was considered to be differentially expressed.

Construction and Analysis of Protein–Protein Interaction Network

We used the STRING[3] database to construct a protein–protein interaction (PPI) network of the differentially expressed gene products. The PPI file was imported to Cytoscape 3.6.0[4], and the MCODE plug-in was used to map the PPIs. The degree, closeness, intermediate degree of each node in the network, and the average value of each protein's nodal degree were defined as the threshold of the PPI network nodes, and the proteins whose degrees were greater than the threshold value were selected. The key nodes of the PPI network were identified, and the correlation scores of the nodes and their interacting proteins were calculated.

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Enrichment Analyses

Gene ontology and Kyoto Encyclopedia of Genes and Genomes enrichment analyses were performed using clusterProfiler (Yu et al., 2012) package in R (3.6.1). The hypergeometric distribution was used to analyze and calculate the significance levels of these differentially expressed genes in each signaling pathway to identify the signaling pathways that were significantly affected ($P < 0.05$).

Identification of Differentially Expressed Immune Response-Related Genes

From June 2019 to February 2020, we recruited eight patients with AAA who underwent surgical resection in the Second Affiliated Hospital of Nanchang University and Sun Yat-sen University. All tissue samples were frozen in liquid nitrogen during the surgery. Two experienced pathologists confirmed AAA. This study was approved by the Ethics Committee of the Second Affiliated Hospital of Nanchang University and Sun Yat-sen University. Written informed consent was obtained from each patient before surgery. The samples were pretreated, and RNA was extracted with TRIzol reagent (German DBI) for real-time quantitative polymerase chain reaction (qPCR). Reverse transcription of RNA into complementary DNA was performed using the Bestar qPCR RT Kit (DBI, Germany) following the manufacturer's instructions. qPCR was performed to determine the expression levels of *PI3, MAP2K1, SSTR1, GPER1, CCR10, IL-6, IL-17,* tumor necrosis factor *(TNF)*-α, and *ACTB* (internal reference gene) in the samples. Primer sequences were obtained from PrimerBank[5] (**Table 1**).

Statistical Analysis

Data were analyzed using GraphPad Prism 7 (GraphPad Software Inc.) and IBM SPSS 17.00 (IBM Analytics, United States). Chi-square and Fisher's exact tests were used for the qualitative analysis of the data. Data were expressed as the mean ± standard deviation (x ± s) and compared using Student's *t*-test. PCR data were analyzed using GraphPad Prism 8. $P < 0.05$ was considered as statistically significant and $P < 0.01$ as very significant.

RESULTS

Infiltration of Immune Cells

We used CIBERSORT to analyze immune cell infiltration in the samples and selected 69 AAA samples and 10 normal aortic samples that met the standard according to the analysis ($P < 0.05$). We found a significant difference in the abundance of 22 types of immune cells in the 10 normal aortic samples and 69 AAA samples (**Figure 1A**). There were also significant differences in the infiltration ratio of plasma cells, T cells, naïve CD4+ T cells, resting memory CD4+ T cells, and macrophages M0 in normal aortic and AAA samples (**Figure 1B**). Correlation analysis of the 22 types of immune cells revealed that the negative correlation between resting mast cells and activated mast cells was the strongest and that between macrophages M0 and plasma cells, macrophages M0 and resting memory CD4+ T cells, and monocytes and plasma cells was strong. A negative correlation was observed between resting CD4+ memory T cells and naïve CD4+ T cells and between plasma cells and CD4+ T cells; a strong positive correlation was observed between naïve CD4+ T cells (**Figure 2A**). **Figure 2B** clearly shows that there are 18 types of immune cells with different proportions in normal and AAA samples, such as resting dendritic cells (**Figure 3A**), eosinophils (**Figure 3B**), macrophages M0 (**Figure 3C**), macrophages M1

[3]https://string-db.org

[4]http://www.cytoscape.org/

[5]https://pga.mgh.harvard.edu/primerbank/

TABLE 1 | Gene sequence.

	Forward primer	Reverse primer
PI3 (5′ -> 3′)	CACGGGAGTTCCTGTT AAAGG	TCTTTCAAGCAGCGGT TAGGG
MAP2K1 (5′ -> 3′)	TGTCGCCAGAAAGAC TCCAG	TCCATTCCGTATGAGC TAAGGG
SSTR1 (5′ -> 3′)	CCAGCATCTACTGTCTG ACTGT	ATGACGAGCAGCGA TAGCAC
GPER1 (5′ -> 3′)	CCTGCTTCTGTTTCG CGGAT	CAATGAGGGAGTAGC ACAGGC
CCR10 (5′ -> 3′)	GCAAACGCAAGGAT GTCGC	CGTAGAGAACGGGATT GAGGC
β-actin (5′ -> 3′)	ATCGTGCGTGACATTAA GGAGAAG	AGGAAGGAAGGCTGGAA GAGTG

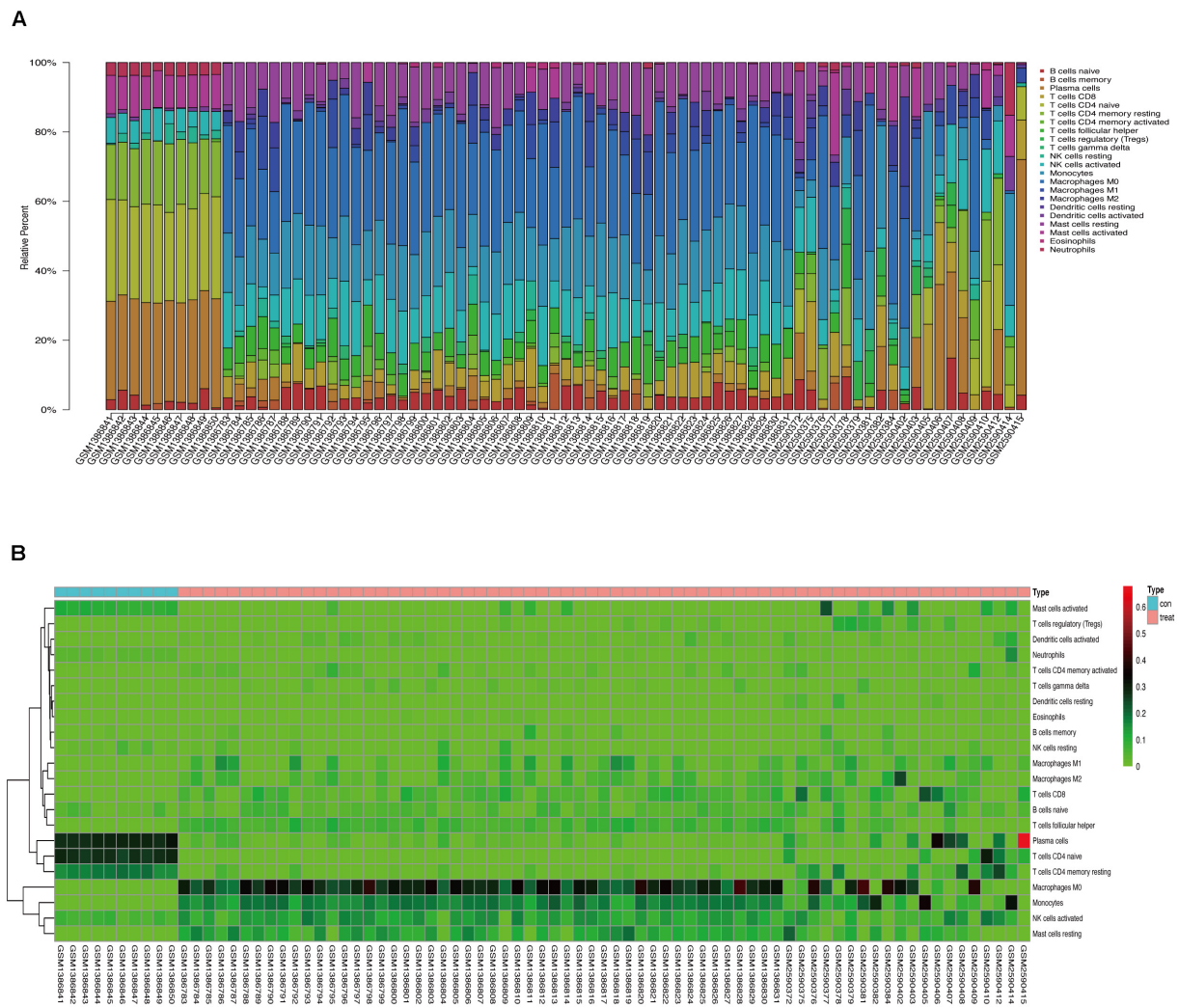

FIGURE 1 | (A) Proportion of 22 kinds of immune cells in the chip. **(B)** Expression heat map of differential immune cells in the sample.

(**Figure 3D**), macrophages M2 (**Figure 3E**), resting mast cells (**Figure 3F**), monocytes (**Figure 3G**), activated NK cells (**Figure 3H**), activated CD4+ memory T cells (**Figure 3I**), CD8+ T cells (**Figure 3J**), and volatile T helper cells (**Figure 3K**). The proportion of regulatory T cells (**Figure 3L**) was increased in AAA samples, whereas the proportions of activated mast cells (**Figure 4A**), neutrophils (**Figure 4B**), resting NK cells (**Figure 4C**), plasma cells (**Figure 4D**), resting memory CD4+ T cells (**Figure 4E**), and naïve CD4+ T cells (**Figure 4F**) were increased in normal aortic samples.

Principal Component Analysis of Samples

We reduced the dimensions of the samples and performed PCA based on the infiltration of 22 types of immune cells in the samples. The results showed that normal aortic and AAA samples could be clearly distinguished based on the infiltration of immune cells (**Figure 5A**).

Differentially Expressed Immune-Related Gene Screening

We extracted and analyzed differences in the expression of immune response-related genes in 69 AAA and 10 normal aortic samples. The results showed 39 upregulated and 133 downregulated immune response-related genes (**Figure 5B**).

Construction and Analysis of the Protein–Protein Interaction Network

We uploaded 39 upregulated and 133 downregulated genes into the STRING database (**Supplementary Figures 1A,B**) to construct the PPI network. Then, we imported the PPI network into Cytoscape and used MCODE to identify the related nodes. As shown in **Figures 5C,D**, the larger the degree of the node in the graph, the darker the color and the larger the diameter of the node. Among the upregulated genes, *SSTR1, GPER1,* and *CCR10* had the largest nodes (highest degrees), whereas, among

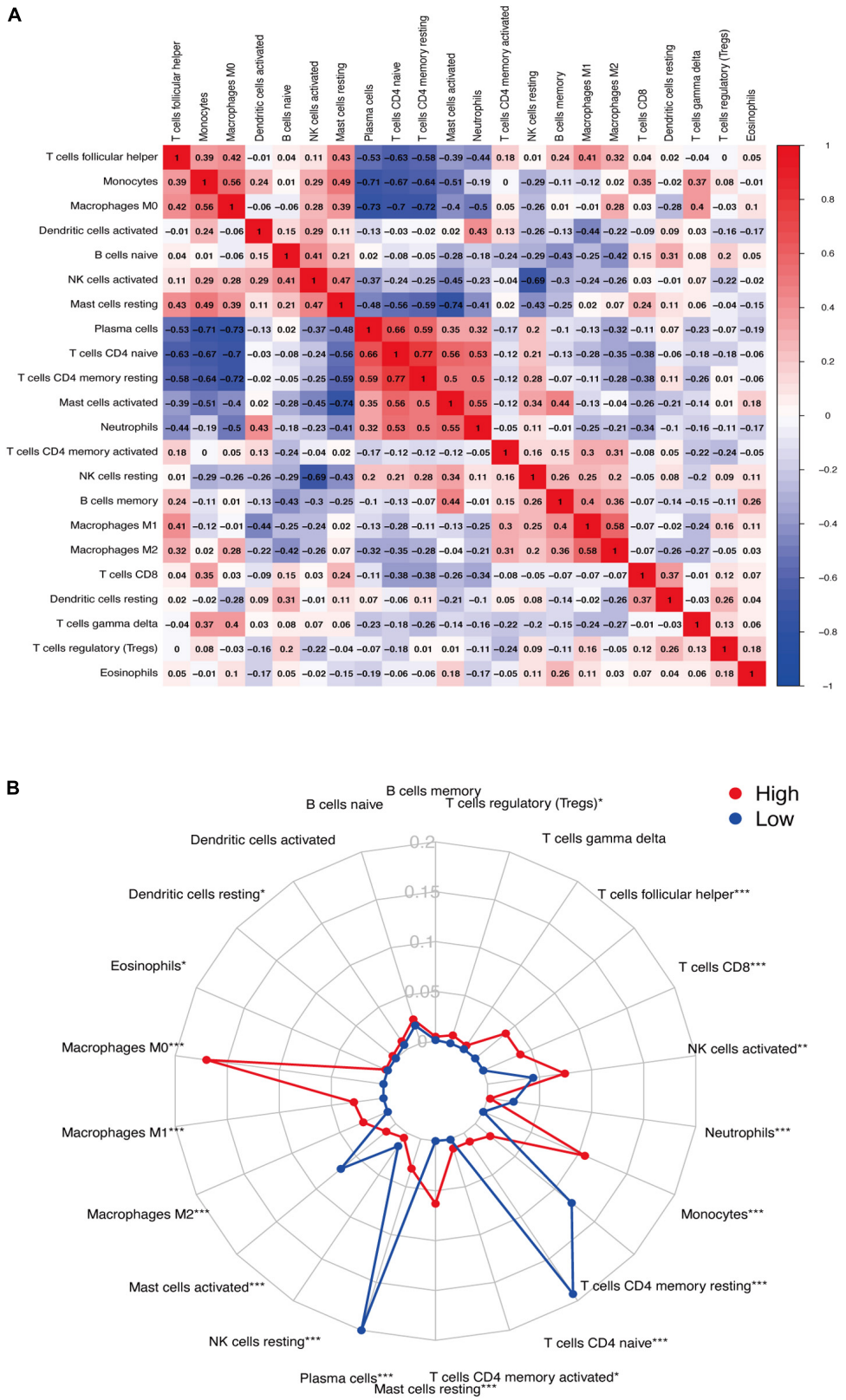

FIGURE 2 | (A) Correlation between immune cells. **(B)** Infiltration of immune cells (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$).

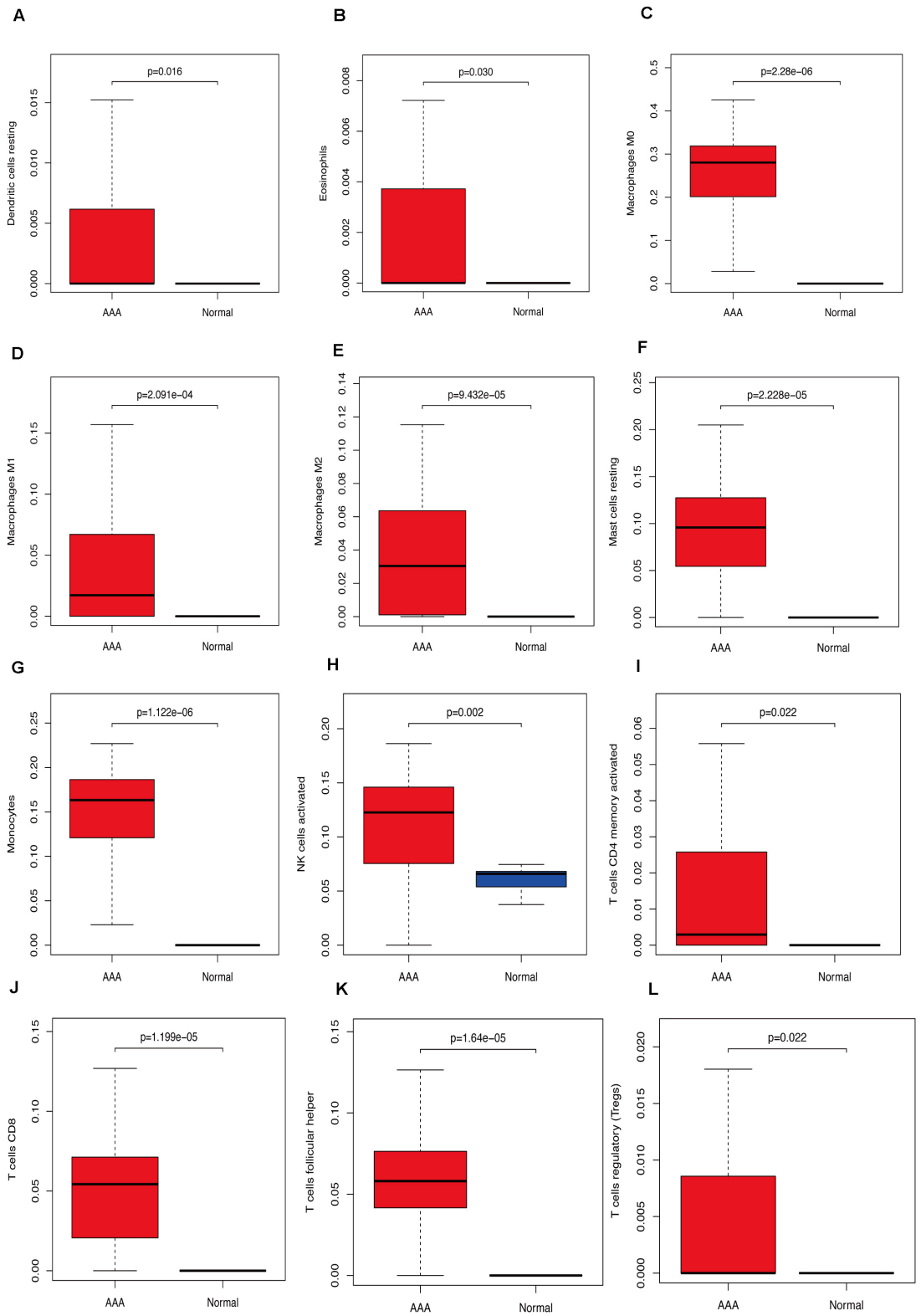

FIGURE 3 | Immunocyte expression ratio, red represents the expression of immune cells in abdominal aortic aneurysm, and blue represents the expression of immune cells in normal vascular tissues. **(A)** The expression of resting dendritic cells. **(B)** The expression of eosinophils. **(C)** The expression of macrophages M0. **(D)** The expression of macrophages M1. **(E)** The expression of macrophages M2. **(F)** The expression of resting mast cells. **(G)** The expression of monocytes, **(H)** Activated NK cell expression, **(I)** activated CD4 memory cell expression, **(J)** CD8 cell expression, **(K)** follicular helper cell expression, **(L)** T regulatory cell (Tregs) expression.

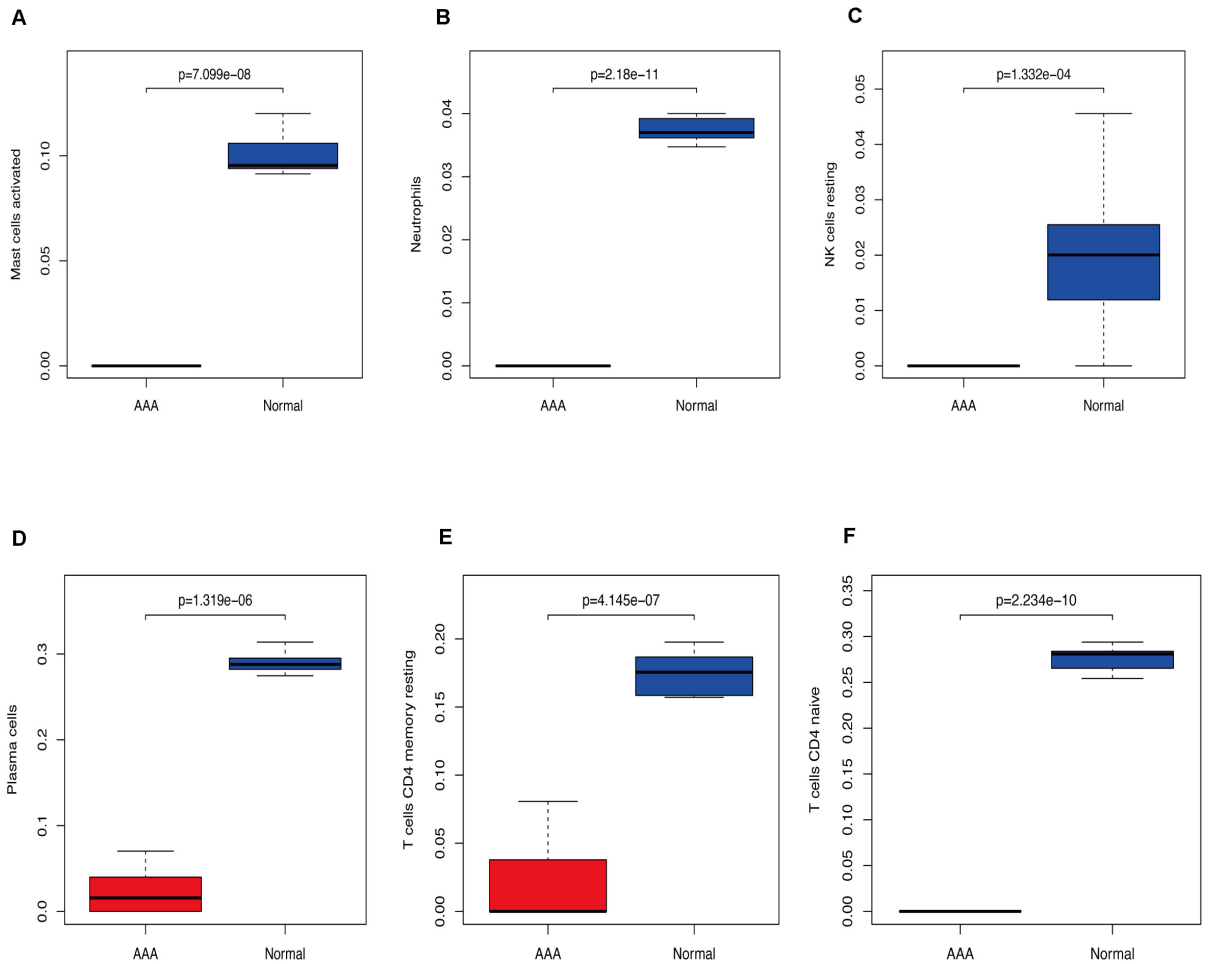

FIGURE 4 | Immunocyte expression ratio, red represents the expression of immune cells in abdominal aortic aneurysm, and blue represents the expression of immune cells in normal vascular tissues **(A)** Expression of activated mast cells. **(B)** Expression of neutrophils. **(C)** The expression of resting NK cells. **(D)** Expression of plasma cells. **(E)** Expression of resting CD4 memory cells. **(F)** Expression of CD4 naïve cells.

the downregulated genes, *PI3* and *MAP2K1* had the largest nodes (highest degrees) (**Table 2**).

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Enrichment Analyses

Figures 6A,B show that the biological processes regulated by the upregulated immune-response-related genes include response to growth hormone, peptide hormone response, growth factor activity, and cytokine receptor activity. These processes were closely related to the growth hormone response, positive regulation of the JAK-STAT cascade, growth hormone receptor signaling, regulation of tyrosine phosphorylation of STAT proteins (**Figures 6C,D**), and hormone-mediated signaling. **Figures 7A,B** show that the downregulated genes were mainly involved in T cell activation, peptide tyrosine phosphorylation, peptidyl tyrosine modification, negative regulation of external stimulus-response, response to TNF, cell chemotaxis, as well as regulation of immune response molecular mediators, platelet α granules, and protein tyrosine kinase activity. This was related to T cell activation, TNF-mediated positive regulation of STAT signaling cascade, negative regulation of external stimulation response, positive regulation of protein kinase B signal transduction, leukocyte migration, leukocyte proliferation, regulation of cell morphogenesis during differentiation, regulation of STAT cascade, leukocyte chemotaxis, and lymphocyte proliferation (**Figures 7C,D**).

Identification of Differentially Expressed Immune Genes

We collected eight AAA samples and six samples from the adjacent aortic aneurysm vessels at the Second Affiliated Hospital of Nanchang University. The expression levels of SSTR1, GPER1, CCR10, PI3, and MAP2K1 were measured using qPCR. The relative expression of each target gene in **Figure 5** was calculated as follows: relative expression = 2 – Δ CT, where Δ CT = CT value of target gene – CT value of the internal reference gene (actin). Compared with normal samples, the expression levels

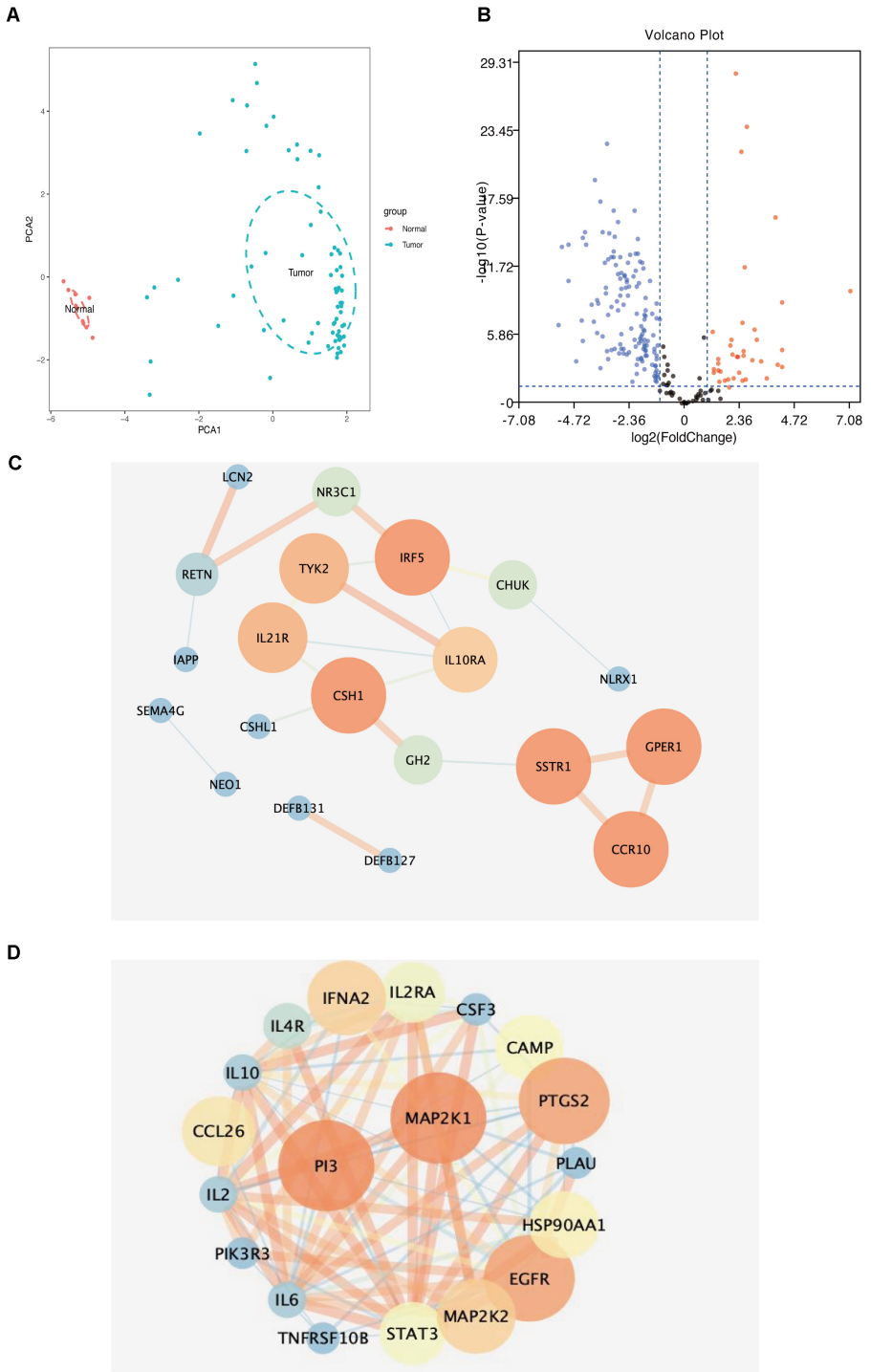

FIGURE 5 | (A) Principal component analysis distribution of normal aortic samples and abdominal aortic aneurysm samples. **(B)** Volcanic map of differential immune-related genes in samples (blue dots represent downregulated genes, red represents core genes in immune-related genes downregulated by upregulated genes). **(C)** Downregulated immune-related genes (the darker the color, the larger the circle, the higher the score of immune-related genes). **(D)** Upregulated immune-related genes (the darker the color, the larger the circle, the higher the score of immune-related genes).

of SSTR1, GPER1, and CCR10 were increased, whereas those of MAP2K1 were decreased in AAA samples; no significant difference was observed in the expression of PI3 (**Figure 8**). We found that the expression of vascular inflammatory factors IL-6, IL-17, and TNF-α increased significantly in AAA samples (**Supplementary Figure 2**).

TABLE 2 | Key of the immune related genes.

Symbol	Category		
SSTR1	Cytokine_Receptors		
GPER1	Cytokine_Receptors		
CCR10	Antimicrobials		
PI3	Antimicrobials		
MAP2K1	NaturalKiller_Cell_Cytotoxicity		
Name	**Score**	**Name**	**Score**
CSH1	6	MAP2K1	8.59090909
GPER1	6	PI3	8.59090909
CCR10	6	PTGS2	8.41666667
IRF5	2	EGFR	8.41666667
SSTR1	2	IFNA2	8
TYK2	1.66666667	MAP2K2	8
IL21R	1.66666667	CCL26	7.82222222
IL10RA	1.4	HSP90AA1	7.71428571
NR3C1	0.66666667	CAMP	7.64444444
GH2	0.66666667	STAT3	7.60233918
CHUK	0.66666667	IL2RA	7.56363636
RETN	0.5	IL4R	7.27272727
LCN2	0	IL10	7.0952381
DEFB127	0	IL2	7.0952381
DEFB131	0	IL6	7.0952381
CSHL1	0	CSF3	7.01470588
IAPP	0	PLAU	7
NEO1	0	TNFRSF10B	7
SEMA4G	0	PIK3R3	7
NLRX1	0		

DISCUSSION

Abdominal aortic aneurysm has always been the focus of vascular surgery research. Due to second-generation sequencing development, more and more researchers began to use bioinformatics technology to study AAA. Wang et al. thought that UBB, NFIA, sparcl1, and other genes play an important role in AAA by comparing gene expression levels and simply constructing a regulatory network (Wang et al., 2018). Gan et al. found that hsa-mir-30a-gng2 and hsa-mir-15b-acss2 may play a role in the development of AAA by screening differentially expressed miRNAs and mRNA and predicting the regulatory relationship through a database (Gan et al., 2019). Gäbel et al. (2017) showed that ccl4l1 and ANGPTL4 were closely related to AAA rupture. In addition, according to Biros et al. (2015), small AAA and large AAA have some unique immune characteristics. They found that cytotoxic T lymphocyte-associated protein 4 (CTLA4) was upregulated in small AAA, and CD8a was upregulated in large AAA. Subsequent studies found that the downregulation of CTLA4 could promote the immune response induced by T cells, leading to the occurrence of AAA. However, few studies have used bioinformatics technology to comprehensively analyze the infiltration of immune cells in AAA. Only one article analyzes the infiltration of immune cells in a ruptured abdominal aortic aneurysm and stable abdominal aortic aneurysm. The sample size is only 48 cases, and there is no laboratory data validation (Lei et al., 2020).

In this study, we analyzed the GSE57691 and GSE98278 datasets related to AAAs. A total of 90 samples were included to comprehensively analyze the status of immune cell infiltration in AAA and the correlation between immune cells. Furthermore, we screened the genes that play a key role in AAA by constructing a network and verified the results by collecting clinical samples for PCR analysis, which made our research results more reliable. We found a significant difference in the infiltration ratio of 22 types of immune cells between normal aortic samples and AAA samples (**Figure 1A**). Furthermore, we found that the proportions of cells related to M1 macrophages, M2 macrophages, mast cell quiescence, monocytes, NK cell activation, and activated memory CD4+ T cells increased in AAAs, whereas those related to mast cell activation, NK cell quiescence (**Figure 4C**), quiescent memory CD4+ T cells (**Figure 4E**), and naïve CD4+ T cells (**Figure 4F**) increased in normal aortic samples. Although the role of mast cells in AAAs is not clear (Sillesen et al., 2015), the recruitment of macrophages in aortic tissue marks the beginning of infiltration into the adventitia, which promotes the secretion of matrix degradants to contribute to the formation of AAAs (Blomkalns et al., 2013). Activated memory CD4+ T cells can play a pro-inflammatory role by differentiating into Th2 cells, and studies have shown that the proportion of Th2 cells in AAAs is increased (Wang et al., 2014b). Furthermore, NK cells can produce pro-inflammatory factors, such as IL-2 and interferon-γ, which can lead to increased cytotoxic activity and promote AAA formation (Chan et al., 2005a,b; Forester et al., 2006). In addition, there is evidence that monocyte depletion can inhibit the formation of AAAs (Wang et al., 2010). Moreover, the analysis of these same datasets also showed that the invasion of monocytes and CD4T cells into the vascular wall and expression of cytotoxic mediators might be the cause of AAA (Ritchie et al., 2015; Gäbel et al., 2017). Thus, our results are highly consistent with these findings. The results of the PCA analysis showed that normal aortic and AAA samples could be clearly distinguished by the infiltration of immune cells. Also, our study confirmed that the expression of inflammatory factors IL-6, IL-17, and TNF was significantly increased in abdominal aortic aneurysm samples compared with adjacent aorta samples. Therefore, searching for the key immune-related genes in AAA is a new direction to study the occurrence and development of AAA.

We screened 172 immune-response-related genes that were differentially expressed in normal aortic vascular samples and AAA samples, including 39 upregulated genes and 133 downregulated genes. These genes were used to construct a PPI network. *SSTR1*, *GPER1*, and *CCR10* were found to be important genes in the upregulated gene network, whereas *PI3* and *MAP2K1* were important genes in the downregulated gene network. SSTR1, one of the five somatostatin receptors (SSTRs), plays an important role in neuroendocrine tumors, such as lung carcinoid (Vesterinen et al., 2019), and the overexpression of SSTR1 can inhibit cell proliferation (Zou et al., 2015). Compared with normal uterine tissues, the expression of GPER1 in uterine leiomyoma is higher and increases cell migration (Kim et al., 2020), which may be one reason for its

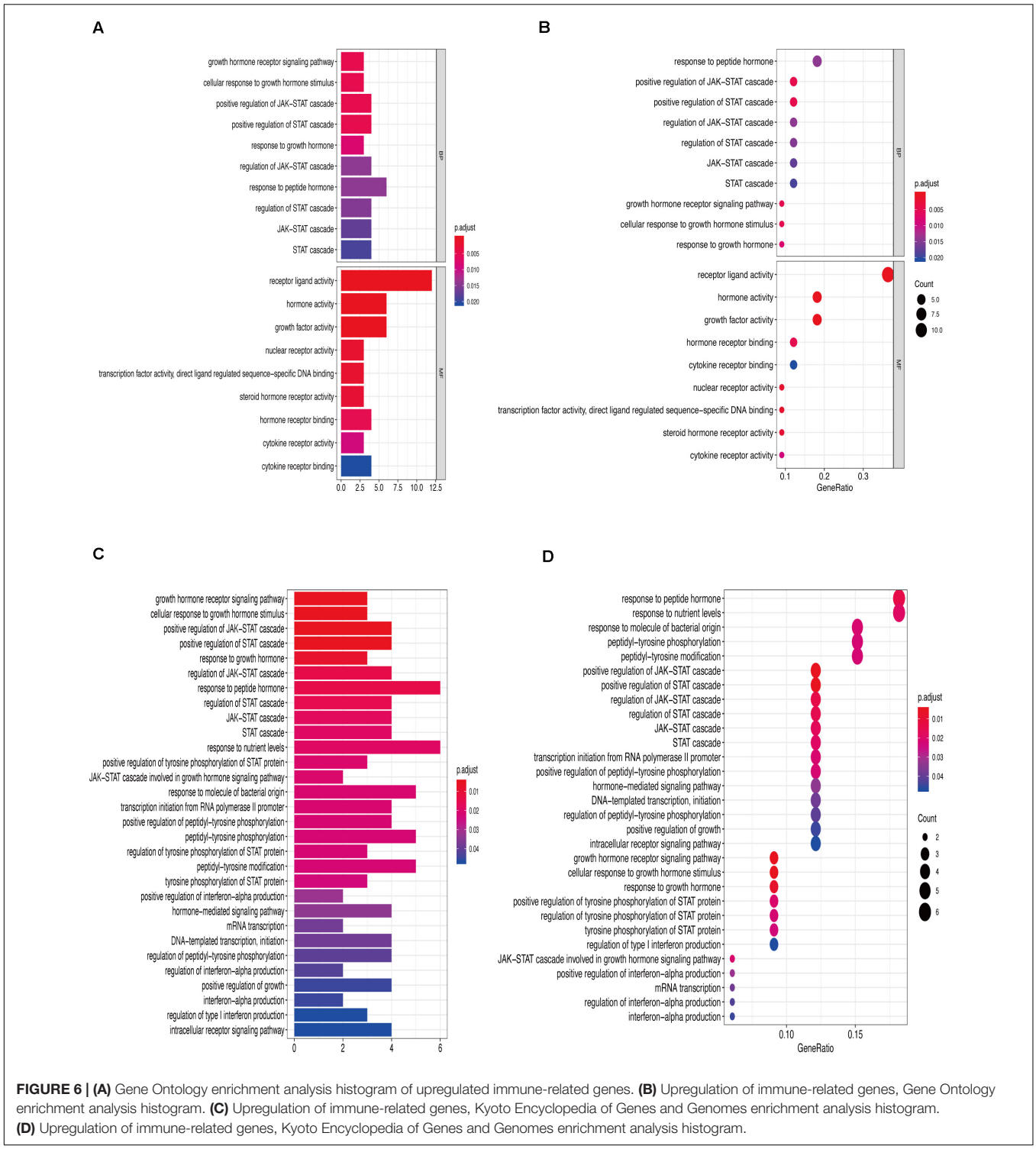

FIGURE 6 | (A) Gene Ontology enrichment analysis histogram of upregulated immune-related genes. **(B)** Upregulation of immune-related genes, Gene Ontology enrichment analysis histogram. **(C)** Upregulation of immune-related genes, Kyoto Encyclopedia of Genes and Genomes enrichment analysis histogram. **(D)** Upregulation of immune-related genes, Kyoto Encyclopedia of Genes and Genomes enrichment analysis histogram.

high expression in AAAs. CCR10 is an important receptor-mediating chemokine, which participates in angiogenesis by endothelial cells. It can promote angiogenesis and improve wound healing by inhibiting the reaction between CCL28 and CCR10 (Chen et al., 2020). CCR10 is expressed in the skin of most patients with psoriasis and atopic or allergic contact dermatitis and plays a key role in T-cell-mediated skin inflammation (Homey et al., 2002). Studies on the same datasets also showed that chemokines and their ligands were upregulated in AAA and could interact with other factors to play a role in AAA. For example, CCL4L1, CCL3L3, CXCL1, CXCL2, CXCL13, and CCR7 were all highly expressed in AAA

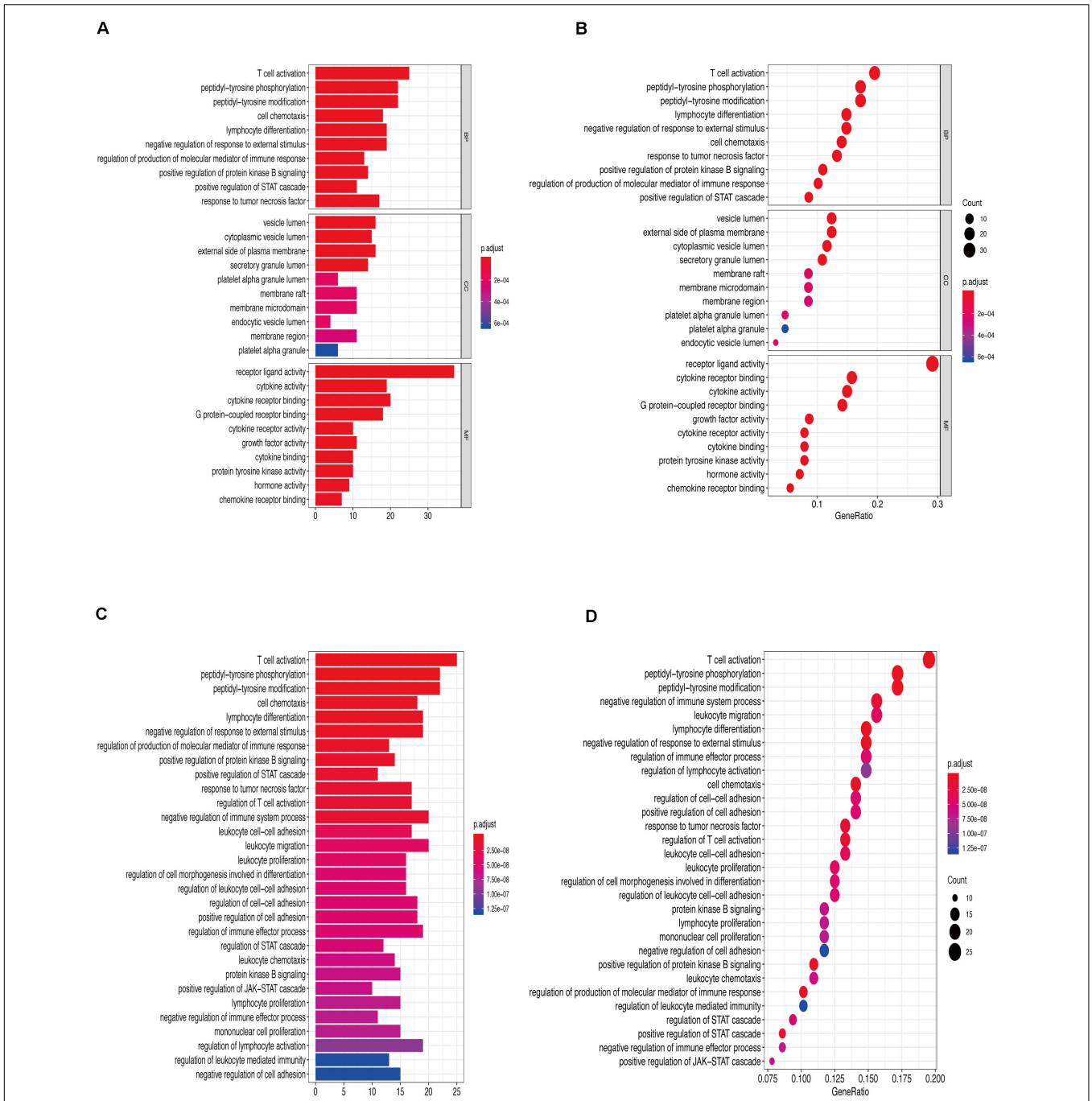

FIGURE 7 | (A) Downregulation of immune-related genes, Gene Ontology enrichment analysis histogram. **(B)** Downregulation of immune-related genes, Gene Ontology enrichment analysis histogram. **(C)** Downregulated immune-related genes Kyoto Encyclopedia of Genes and Genomes enrichment analysis histogram. **(D)** Downregulated immune-related genes Kyoto Encyclopedia of Genes and Genomes enrichment analysis dot map.

(Gäbel et al., 2017; Gan et al., 2019). In one study, GNG2, CXCL1, and CCR7 were considered as central genes in the AAA network, and GNG2 interacted with CXCL1 and CCR7 to participate in the chemokine signaling road (Gan et al., 2019). PI3 is an important mediator in the occurrence and development of inflammation and synthesized and secreted by infiltrating neutrophils. Lipopolysaccharide, elastase, and TNF-α can also promote its production, and although the expression of PI3 is increased in the airway and mucosa where inflammatory stimulation persists (Sallenave et al., 1994; Pfundt et al., 1996, 2000; Reid et al., 1999), this does not mean that PI3 is a pro-inflammatory factor. On the contrary, PI3 seems to play a

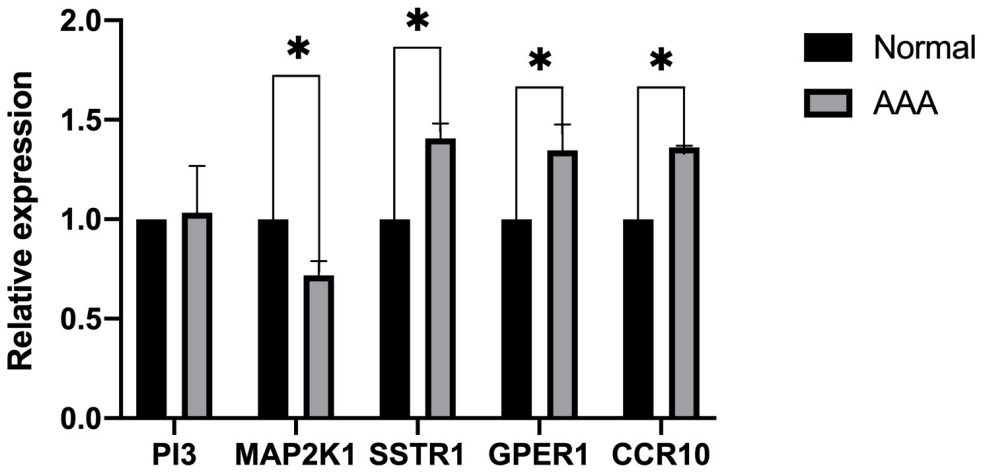

FIGURE 8 | Expression of five immune-related key genes in the samples was analyzed using PCR (*$P < 0.05$).

protective role in inflammation; for example, it is downregulated in the acute phase of acute respiratory distress syndrome, and in an experimental study, the plasma PI3 levels of control subjects were much higher than those of patients with respiratory distress syndrome. Therefore, a decrease in PI3 expression may lead to a decrease in the body's tolerance for inflammation (Tejera et al., 2009). The mutation of *MAP2K1* is considered the most common cause of extracranial arteriovenous malformations. *MAP2K1* encodes MAP-extracellular signal-regulated kinase 1 (MEK1) and affects cell development through the RAS/MAPK signaling pathway. Extracranial arteriovenous malformations are characterized by vascular endothelial cell dysfunction and the promotion of arteriovenous malformations (Couto et al., 2017; Konczyk et al., 2020). These findings are mostly consistent with our results.

In our study, Gene Ontology and Kyoto Encyclopedia of Genes and Genomes analyses of differentially expressed genes showed that the upregulated genes were mainly involved in growth hormone response, peptide hormone response, growth factor activity, cytokine receptor activity, positive regulation of JAK-STAT cascade, and hormone-mediated signaling. Downregulated genes were mainly involved in T cell activation, negative regulation of external stimulus–response, response to TNF, cell chemotaxis, regulation of immune response molecular mediators, T cell activation, and leukocyte migration and proliferation. These biological processes and pathways are closely related to inflammation and immunity. Other studies on AAA have also confirmed that the interaction between cytokines and cytokine receptors, chemokine signaling pathway and T cell receptor signaling pathway play an important role in the occurrence and development of AAA (Biros et al., 2015). Therefore, we believe that the five immune response-related genes, *SSTR1, GPER1, CCR10, PI3,* and *MAP2K1*, may play important roles in the formation of AAA.

This study has the limitation of a small sample size for the measurement of selected gene expression; indeed, this may be the reason why we found no significant difference in PI3 expression. Hence, we plan to perform more in-depth studies on the selected genes to elucidate the mechanisms of action of these genes in the development of AAA.

CONCLUSION

In this study, by analyzing an AAA microarray, we found that normal aorta and AAA samples could be clearly distinguished via the infiltration of immune cells and identified the key differentially expressed immune response-related genes, SSTR1, GPER1, CCR10, PI3, and MAP2K1.

Based on our results and the literature, we hypothesize that these five genes participate in the immune response and play important roles in the development of AAA. Thus, these genes may be key targets for the diagnosis and treatment of AAA.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: GSE57691 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE57691) and GSE98278 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE98278).

ETHICS STATEMENT

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

AUTHOR CONTRIBUTIONS

HN: designing research direction and writing the manuscript. JQ: searching for references and writing the manuscript.

SW: helping to revise manuscript. WZ: reviewing and revising the manuscript and guidance for manuscript preparation. All authors contributed to the article and approved the submitted version.

FUNDING

The authors' research was supported by the research fund of The National Natural Science Foundation of China (13008426) and the Key R&D Projects of Jiangxi Province (20171ACG70008).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.579215/full#supplementary-material

Supplementary Figure 1 | (A) Upregulation gene network map downloaded from STRING. **(B)** Downregulation gene network map downloaded from STRING.

Supplementary Figure 2 | The expression of IL-6, IL-17 and TNF-α in the samples was analyzed using PCR (* represents $P < 0.05$).

REFERENCES

Arya, S., Kim, S. I., Duwayri, Y., Brewster, L. P., Veeraswamy, R., Salam, A., et al. (2015). Frailty increases the risk of 30-day mortality, morbidity, and failure to rescue after elective abdominal aortic aneurysm repair independent of age and comorbidities. *J. Vasc. Surg.* 61, 324–331. doi: 10.1016/j.jvs.2014.08.115

Biros, E., Gäbel, G., Moran, C. S., Schreurs, C., Lindeman, J. H., Walker, P. J., et al. (2015). Differential gene expression in human abdominal aortic aneurysm and aortic occlusive disease. *Oncotarget* 6, 12984–12996. doi: 10.18632/oncotarget.3848

Blomkalns, A. L., Gavrila, D., Thomas, M., Neltner, B. S., Blanco, V. M., Benjamin, S. B., et al. (2013). CD14 directs adventitial macrophage precursor recruitment: role in early abdominal aortic aneurysm formation. *J. Am. Heart Assoc.* 2:e000065. doi: 10.1161/JAHA.112.000065

Chan, W. L., Pejnovic, N., Hamilton, H., Liew, T. V., Popadic, D., Poggi, A., et al. (2005a). Atherosclerotic abdominal aortic aneurysm and the interaction between autologous human plaque-derived vascular smooth muscle cells, type 1 NKT, and helper T cells. *Circ. Res.* 96, 675–683. doi: 10.1161/01.RES.0000160543.84254.f1

Chan, W. L., Pejnovic, N., Liew, T. V., and Hamilton, H. (2005b). Predominance of Th2 response in human abdominal aortic aneurysm: mistaken identity for IL-4-producing NK and NKT cells? *Cell. Immunol.* 233, 109–114. doi: 10.1016/j.cellimm.2005.04.020

Chang, T. W., Gracon, A. S., Murphy, M. P., and Wilkes, D. S. (2015). Exploring autoimmunity in the pathogenesis of abdominal aortic aneurysms. *Am. J. Physiol. Heart Circ. Physiol.* 309, H719–H727. doi: 10.1152/ajpheart.00273.2015

Chen, Z., Haus, J. M., Chen, L., Wu, S. C., Urao, N., Koh, T. J., et al. (2020). CCL28-induced CCR10/eNOS interaction in angiogenesis and skin wound healing. *FASEB J.* 34, 5838–5850. doi: 10.1096/fj.201902060R

Couto, J. A., Huang, A. Y., Konczyk, D. J., et al. (2017). Somatic MAP2K1 mutations are associated with extracranial arteriovenous malformation. *Am. J. Hum. Genet.* 100, 546–554. doi: 10.1016/j.ajhg.2017.01.018

Culav, E. M., Clark, C. H., and Merrilees, M. J. (1999). Connective tissues: matrix composition and its relevance to physical therapy. *Phy. Ther.* 79, 308–319. doi: 10.1093/ptj/79.3.308

Eliason, J. L., Hannawa, K. K., Ailawadi, G., Sinha, I., Ford, J. W., and Deogracias, M. P. (2005). Neutrophil depletion inhibits experimental abdominal aortic aneurysm formation. *Circulation* 112, 232–240. doi: 10.1161/CIRCULATIONAHA.104.517391

Forester, N. D., Cruickshank, S. M., Scott, D. J. A., and Carding, S. R. (2006). Increased natural killer cell activity in patients with an abdominal aortic aneurysm. *Br. J. Surg.* 93, 46–54. doi: 10.1002/bjs.5215

Gäbel, G., Northoff, B. H., Weinzierl, I., Ludwig, S., Hinterseher, I., Wilfert, W., et al. (2017). Molecular fingerprint for terminal abdominal aortic aneurysm disease. *J. Am. Heart Assoc.* 6:e006798. doi: 10.1161/JAHA.117.006798

Gan, S., Pan, Y., and Mao, J. (2019). miR-30a-GNG2 and miR-15b-ACSS2 interaction pairs may be potentially crucial for development of abdominal aortic aneurysm by influencing inflammation. *DNA Cell Biol.* 38, 1540–1556. doi: 10.1089/dna.2019.4994

Golledge, J., and Norman, P. E. (2011). Current status of medical management for abdominal aortic aneurysm. *Atherosclerosis* 217, 57–63. doi: 10.1016/j.atherosclerosis.2011.03.006

Golledge, J., Muller, J., Daugherty, A., and Norman, P. (2006). Abdominal aortic aneurysm: pathogenesis and implications for management. *Arterioscler Thrombosis Vascul. Biol.* 26, 2605–2613. doi: 10.1161/01.ATV.0000245819.32762.cb

Gordon, P. A., and Toursarkissian, B. (2014). Treatment of abdominal aortic aneurysms: the role of endovascular repair. *AORN* 100, 241–259. doi: 10.1016/j.aorn.2014.01.025

Homey, B., Alenius, H., Müller, A., Soto, H., Bowman, E. P., Yuan, W., et al. (2002). CCL27-CCR10 interactions regulate T cell-mediated skin inflammation. *Nat. Med.* 8, 157–165. doi: 10.1038/nm0202-157

Keeling, W. B., Armstrong, P. A., Stone, P. A., Bandyk, D. F., and Shames, M. L. (2005). An overview of matrix metalloproteinases in the pathogenesis and treatment of abdominal aortic aneurysms. *Vasc. Endovasc. Surg.* 39, 457–464. doi: 10.1177/153857440503900601

Kim, M., Kim, Y. S., Choi, J. I., Kim, J. M., Lee, H. H., and Kim, T. H. (2020). G protein-coupled estrogen receptor 1 expression in normal myometrium, leiomyoma, and adenomyosis tissues of premenopausal women. *Gynecol. Endocrinol.* 36, 599–604. doi: 10.1080/09513590.2020.1751108

Konczyk, D. J., Goss, J. A., Smits, P. J., Sudduth, C. L., Al-Ibraheemi, A., and Greene, A. K. (2020). Arteriovenous malformation MAP2K1 mutation causes local cartilage overgrowth by a cell-non autonomous mechanism. *Sci. Rep.* 10:4428. doi: 10.1038/s41598-020-61444-x

Leek, J. T., Johnson, W. E., Parker, H. S., Fertig, E. J., Jaffe, A. E., Storey, J. D., et al. (2019). *sva: Surrogate Variable.Analysis. R package version 3.34.0.*

Lei, C., Yang, D., Chen, S., Chen, W., Sun, X., Wu, X., et al. (2020). Patterns of immune infiltration in stable and raptured abdominal aortic aneurysms: A gene-expression-based retrospective study. *Gene* 762:145056. doi: 10.1016/j.gene.2020.145056

Liu, Y., Liao, J., Zhao, M., Wu, H., Yung, S., Chan, T. M., et al. (2015). Increased expression of TLR2 in CD4(+) T cells from SLE patients enhances immune reactivity and promotes IL-17 expression through histone modifications. *Eur. J. Immunol.* 45, 2683–2693. doi: 10.1002/eji.201445219

Longo, G. M., Buda, S. J., Fiotta, N., Xiong, W., Griener, T., Shapiro, S., et al. (2005). MMP-12 has a role in abdominal aortic aneurysms in mice. *Surgery* 137, 457–462. doi: 10.1016/j.surg.2004.12.004

Longo, G. M., Xiong, W., Greiner, T. C., Zhao, Y., Fiotti, N., and Baxter, B. T. (2002). Matrix metalloproteinases 2 and 9 work in concert to produce aortic aneurysms. *J. Clin. Invest.* 110, 625–632. doi: 10.1172/JCI0215334

Moxon, J. V., Parr, A., Emeto, T. I., Walker, P., Norman, P. E., and Golledge, J. (2010). Diagnosis and monitoring of abdominal aortic aneurysm: current status and future prospects. *Curr. Probl. Cardiol.* 35, 512–548. doi: 10.1016/j.cpcardiol.2010.08.004

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., and Xu, Y. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457. doi: 10.1038/nmeth.3337

Newman, K. M., Jean-Claude, J., Li, H., Ramey, W. G., and Tilson, M. D. (1994). Cytokines that activate proteolysis are increased in abdominal aortic aneurysms. *Circulation* 90, II224–II227.

Pfundt, R., van Ruissen, F., van Vlijmen-Willems, I. M., Alkemade, H. A., Zeeuwen, P. L., Jap, P. H., et al. (1996). Constitutive and inducible expression of SKALP/elafin provides anti-elastase defense in human epithelia. *J. Clin. Invest.* 98, 1389–1399. doi: 10.1172/jci118926

Pfundt, R., Wingens, M., Bergers, M., Zweers, M., Frenken, M., and Schalkwijk, J. T. N. F. - (2000). Alpha and serum induce SKALP/elafin gene expression in human keratinocytes by a p38 MAP kinase-dependent pathway. *Arch. Dermatol. Res.* 292, 180–187. doi: 10.1007/s004030050475

Rateri, D. L., Howatt, D. A., Moorleghen, J. J., Charnigo, R., Cassis, L. A., and Daugherty, A. (2011). Prolonged infusion of angiotensin II in apoE$^{-/-}$ mice promotes macrophage recruitment with continued expansion of abdominal aortic aneurysm. *Am. J. Pathol.* 179, 1542–1548. doi: 10.1016/j.ajpath.2011.05.049

Reid, P. T., Marsden, M. E., Cunningham, G. A., Haslett, C., and Sallenave, J. M. (1999). Human neutrophil elastase regulates the expression and secretion of elafin (elastase-specific inhibitor) in type II alveolar epithelial cells. *FEBS Lett.* 457, 33–37. doi: 10.1016/s0014-5793(99)01004-2

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gkv007

Rizzo, R. J., McCarthy, W. J., Dixit, S. N., Lilly, M. P., Shively, V. P., Flinn, W. R., et al. (1989). Collagen types and matrix protein content in human abdominal aortic aneurysms. *J. Vasc. Surg.* 10, 365–373. doi: 10.1067/mva.1989.13151

Rooke, T. W., Hirsch, A. T., Misra, S., Sidawy, A. N., Beckman, J. A., Findeiss, L. K., et al. (2012). American college of cardiology foundation., american heart association., society for cardiovascular angiography and interventions., society of interventional radiology., society for vascular medicine., society for vascular surgery.(2012). 2011 ACCF/AHA focused update of the guideline for the management of patients with peripheral artery disease (updating the 2005 guideline): a report of the american college of cardiology foundation/american heart association task force on practice guidelines: developed in collaboration with the society for cardiovascular angiography and interventions, society of interventional radiology, society for vascular medicine, and society for vascular surgery. *Catheter Cardiovasc. Interv.* 79, 501–531. doi: 10.1002/ccd.23373

Sakalihasan, N., Delvenne, P., Nusgens, B. V., Limet, R., Lapière, C. M., et al. (1996). Activated forms of MMP2 and MMP9 in abdominal aortic aneurysms. *J. Vasc. Surg.* 24, 127–133. doi: 10.1016/s0741-5214(96)70153-2

Sakalihasan, N., Michel, J. B., Katsargyris, A., Kuivaniemi, H., Defraigne, J. O., Nchimi, A., et al. (2018). Abdominal aortic aneurysms. *Nat. Rev. Dis. Primers* 4:34. doi: 10.1038/s41572-018-0030-7

Sallenave, J. M., Shulmann, J., Crossley, J., Jordana, M., and Gauldie, J. (1994). Regulation of secretory leukocyte proteinase inhibitor (SLPI) and elastase-specific inhibitor (ESI/elafin) in human airway epithelial cells by cytokines and neutrophilic enzymes. *Am. J. Respir. Cell Mol. Biol.* 11, 733–741. doi: 10.1165/ajrcmb.11.6.7946401

Sillesen, H., Eldrup, N., Hultgren, R., Lindeman, J., Bredahl, K., and Thompson, M. (2015). AORTA trial investigators. Randomized clinical trial of mast cell inhibition in patients with a medium-sized abdominal aortic aneurysm. *Br. J. Surg.* 102, 894–901. doi: 10.1002/bjs.9824

Tejera, P., Wang, Z., Zhai, R., Su, L., Sheu, C. C., Taylor, D. M., et al. (2009). Genetic polymorphisms of peptidase inhibitor 3 (elafin) are associated with acute respiratory distress syndrome. *Am. J. Respir. Cell Mol. Biol.* 41, 696–704. doi: 10.1165/rcmb.2008-0410OC

Umebayashi, R., Uchida, H. A., and Wada, J. (2018). Abdominal aortic aneurysm in aged population. *Aging* 10, 3650–3651. doi: 10.18632/aging.101702

Vesterinen, T., Leijon, H., Mustonen, H., Remes, S., Knuuttila, A., Salmenkivi, K., et al. (2019). Somatostatin receptor expression is associated with metastasis and patient outcome in pulmonary carcinoid tumors. *J. Clin. Endocrinol. Metab.* 104, 2083–2093. doi: 10.1210/jc.2018-01931

Wang, G., Bi, L., Wang, G., Huang, F., Lu, M., and Zhu, K. (2018). Microarray analysis to identify the similarities and differences of pathogenesis between aortic occlusive disease and abdominal aortic aneurysm. *Vascular* 26, 301–314. doi: 10.1177/1708538117736695

Wang, J., Lindholt, J. S., Sukhova, G. K., Shi, M. A., Xia, M., Chen, H., et al. (2014a). IgE actions on CD4+ T cells, mast cells, and macrophages participate in the pathogenesis of experimental abdominal aortic aneurysms. *EMBO Mol. Med.* 6, 952–969. doi: 10.15252/emmm.201303811 .

Wang, L., Gao, S., Xu, W., Zhao, S., Zhou, J., Wang, N., et al. (2014b). Allergic asthma accelerates atherosclerosis dependent on Th2 and Th17 in apolipoprotein E deficient mice. *J. Mol. Cell Cardiol.* 72, 20–27. doi: 10.1016/j.yjmcc.2014.02.005

Wang, Y., Ait-Oufella, H., Herbin, O., Bonnin, P., Ramkhelawon, B., Taleb, S., et al. (2010). TGF-beta activity protects against inflammatory aortic aneurysm progression and complications in angiotensin II-infused mice. *J. Clin. Invest.* 120, 422–432. doi: 10.1172/JCI38136

World Health Organization (2013). *Media centre,Cardiovascular diseases (CVDs)*. Available Online at: http://www.who.int/mediacentre/factsheets/fs317/zh/index.html. Accessed Jan 29, 2015

Yamaguchi, T., Yokokawa, M., Suzuki, M., Higashide, S., Katoh, Y., Sugiyama, S., et al. (2000). The effect of immunosuppression on aortic dilatation in a rat aneurysm model. *Surg. Today* 30, 1093–1099. doi: 10.1007/s005950070007

Yan, Y. W., Fan, J., Bai, S. L., Hou, W. J., Li, X., and Tong, H. (2016). Zinc prevents abdominal aortic aneurysm formation by induction of A20-mediated suppression of NF-κB pathway. *PLoS One* 11:e0148536. doi: 10.1371/journal.pone.0148536

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118

Zou, Y., Tan, H., Zhao, Y., Zhou, Y., Cao, L., et al. (2015). Expression and selective activation of somatostatin receptor subtypes induces cell cycle arrest in cancer cells. *Oncol Lett.* 17, 1723–1731. doi: 10.3892/ol.2018.9773

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

frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 08 January 2021
doi: 10.3389/fcvm.2020.619083

MicroRNA-29b/c-3p Indicate Advanced Liver Fibrosis/Cirrhosis in Univentricular Heart Patients With and Without Fontan Palliation

Masood Abu-Halima[1,2]*, *Eckart Meese*[1], *Mohamad Ali Saleh*[2], *Andreas Keller*[3], *Hashim Abdul-Khaliq*[2] *and Tanja Raedle-Hurst*[2]

[1] Institute of Human Genetics, Saarland University Medical Center, Homburg, Germany, [2] Department of Pediatric Cardiology, Saarland University Medical Center, Homburg, Germany, [3] Center for Clinical Bioinformatics, Saarland University, Saarbruecken, Germany

OPEN ACCESS

Edited by:
Abdelaziz Beqqali, University of Edinburgh, United Kingdom

Reviewed by:
Kumaravelu Jagavelu, Central Drug Research Institute (CSIR), India
Xiaohu Li, First Affiliated Hospital of Anhui Medical University, China

***Correspondence:**
Masood Abu-Halima
masood@daad-alumni.de

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

Received: *19 October 2020*
Accepted: *14 December 2020*
Published: *08 January 2021*

Citation:
Abu-Halima M, Meese E, Saleh MA, Keller A, Abdul-Khaliq H and Raedle-Hurst T (2021) MicroRNA-29b/c-3p Indicate Advanced Liver Fibrosis/Cirrhosis in Univentricular Heart Patients With and Without Fontan Palliation. Front. Cardiovasc. Med. 7:619083. doi: 10.3389/fcvm.2020.619083

Aim: The present study aims to identify those microRNAs (miRNAs) in patients with univentricular heart (UVH) disease with and without Fontan palliation that may be associated with advanced liver fibrosis/cirrhosis.

Materials and Methods: SurePrint^{TM} 8 × 60K Human v21 miRNA arrays were used to determine the miRNA abundance profiles in the blood of 48 UVH patients with and without Fontan palliation and 32 matched healthy controls. The abundance levels of selected miRNAs have been validated by quantitative reverse transcription-polymerase chain reaction (RT-qPCR).

Results: According to microarray analysis, 50 miRNAs were found to be significantly abundant in UVH patients of which miR-29b-3p and miR-29c-3p were significantly related to the model of end-stage liver disease (MELD)-Albumin and albumin-bilirubin (ALBI) score representing advanced liver fibrosis/cirrhosis. Relative expression levels of both miRNAs were significantly higher in patients with a higher collapsibility index representing venous hepatic congestion, a higher MELD-Albumin or ALBI score and incomplete or no Fontan palliation. In the logistic regression analysis, a MELD-Albumin score $\geq$ 11 or ALBI score $>$ −2.6 were best predicted by total bilirubin (OR 6.630, $P = 0.016$), albumin (OR 0.424, $P = 0.026$), and miR-29c-3p (OR 33.060, $P = 0.047$). After adjustment to the status of Fontan palliation, however, no statistical significance of these parameters was found thus underlining the importance of palliation status on progression of liver fibrosis/ cirrhosis in UVH patients.

Conclusions: In UVH patients with and without Fontan palliation, miR-29b-3p and miR-29c-3p seem to be markers of advanced liver fibrosis/cirrhosis and thus may be used in the risk assessment of these patients.

Keywords: microRNA, univentricular heart, MELD-Albumin score, ALBI score, collapsibility index

INTRODUCTION

Univentricular heart (UVH) disease is a rare and complex congenital heart disorder that is characterized by functionally single ventricle anatomy resulting in volume overload of the single ventricular chamber and systemic venous congestion. The only surgical option to improve hemodynamics in these patients is a palliative procedure called Fontan operation connecting both caval veins to the pulmonary artery and thus directing systemic venous blood return directly into the pulmonary circulation without the pulsatile function of a ventricular chamber. Depending on the pulmonary vascular resistance and diastolic function of the single ventricle, this passive venous blood flows into the pulmonary circulation results in chronic systemic venous congestion, particularly of the liver (1). In UVH patients with and without Fontan circulation, the development of liver fibrosis due to long-standing venous congestion of the liver is a well-known and challenging complication resulting finally in liver cirrhosis and in rare cases hepatocellular carcinoma (2–4). Thus, early detection of these complications in this cohort of patients is of diagnostic and prognostic importance. To date, liver biopsy still represents the gold standard in assessing the grade and stage of liver fibrosis. However, this technique is invasive and carries the potential risk of bleeding complications but also sampling and interpretation errors (5). Therefore, non-invasive diagnostic tools such as serological scoring or imaging-based methods should be used for screening and follow-up of congestive hepatopathy in these patients.

Using non-invasive methods, differentiation of reversible liver congestion from irreversible fibrosis and reliable prediction of liver fibrosis as assessed by biopsies remains challenging. Previous studies have shown that measurement of liver stiffness using transient elastography, shear wave elastography, or acoustic radiation force impulse imaging currently cannot differentiate reliably between congestion and fibrosis of the liver and thus may overestimate the extent of liver fibrosis in congestive hepatopathy (6, 7). Moreover, biochemical assessment of liver fibrosis using laboratory scores such as APRI, Forns, FIB4 or FibroSURE score is not feasible in these patients either because these scores do not correlate well with liver biopsy findings or are not related to the presence of sinusoidal or portal fibrosis on biopsy specimens that are typical for congestive hepatopathy (7, 8). However, the Model for End-Stage Liver Disease (MELD) excluding international normalized ratio (MELD-XI) score is a functional score indicating advanced liver disease that seems to be suitable in the risk assessment of patients with combined cardiac and hepatic dysfunction and even predicts clinical outcomes in congestive liver disease (6, 9, 10). Recently, a new MELD-Albumin score including albumin to replace INR in the conventional MELD score has been investigated in patients undergoing tricuspid annuloplasty for severe tricuspid regurgitation revealing to be a prognostic indicator as good as the MELD-XI score (11). Moreover, the ALBI score is applicable, robust, and superior to the conventional MELD score in the evaluation of the functional status and long-term prognosis of patients with liver cirrhosis (12, 13).

MicroRNAs (miRNAs) are known to be expressed in a cell type- or tissue-specific and stage-dependent manner in chronic liver disease (14–17). They also play a significant role in the regulation of liver fibrosis/cirrhosis representing the common end stage of most chronic liver diseases and being associated with tremendous morbidity and mortality (18, 19). Since miRNAs are known to modulate different steps in the pathophysiology of liver fibrosis, they might be used for early detection or progression of liver fibrosis (16). To our knowledge, no data are available on the specific miRNAs that are involved in the onset or progression of liver fibrosis in UVH patients. In this study, we aimed to identify miRNAs that indicate significant liver fibrosis in UVH patients using prognostic scores of advanced liver fibrosis/cirrhosis, specifically the new MELD-Albumin and ALBI score, and to assess the most important determinants of advanced liver fibrosis/cirrhosis in this cohort of patients.

MATERIALS AND METHODS

Patients

Between 02/05/2015 and 18/06/2018, 48 consecutive UVH patients who were regularly seen in our outpatient clinic were enrolled in the present study. 42/48 (87.5%) patients had a complete Fontan palliation and 6/48 (12.5%) patients an incomplete or no Fontan palliation. Regarding the morphology of the dominant systemic ventricle, 32/48 (66.7%) patients had a morphological dominant left, and 16/32 (50%) patients a morphological dominant right systemic ventricle. In the left systemic ventricle UVH group, 12 patients presented with tricuspid atresia, 12 patients with double inlet left ventricle, and eight patients with pulmonary atresia with or without ventricular septal defect. In the right systemic ventricle UVH group, eight patients had hypoplastic left heart syndrome or mitral atresia, and eight patients double outlet right ventricle with pulmonary stenosis. The mean age was 22.8 $\pm$ 10.1 years (range 11– 46 years) including 17 females and 31 males. At enrollment, a structured protocol including a 12-lead surface electrocardiogram, a physical examination, measurement of blood pressure and transcutaneous oxygen saturation at rest, two-dimensional echocardiography as well as a venous blood draw for routine laboratory parameters and blood sampling were performed. Additionally, ultrasonographic parameters of liver congestion such as the diameter of the inferior caval vein (IVC) at deep inspiration and expiration were measured to calculate the collapsibility index (20). Clinical characteristics of the patient cohort and UVH-specific data have already been presented in a previous study (21). Patients' specific liver parameters are illustrated in **Table 1**. Thirty-two healthy volunteers served as controls and were matched to UVH patients according to age and gender. In all volunteers, a physical examination, two-dimensional echocardiography to verify the absence of any heart abnormality, ultrasonography of the inferior caval vein to rule out liver congestion, and venous blood sampling were performed on the same day. The study complies with the Declaration of Helsinki, was approved by the local ethics committee and all participants or their guardians gave written and informed consent before enrollment.

TABLE 1 | Specific liver parameters of UVH patients according to the status of palliation.

Variables	All patients ($n = 48$)	Complete Fontan palliation ($n = 42$)	Incomplete or no Fontan palliation ($n = 6$)	*P*-value*
Age at enrolment (years)	22.8 ± 10.1	20.5 ± 8.5	38.2 ± 4.3	<0.001
Male gender (%)	31/48 (64.6%)	29/42 (69.0%)	2/6 (33.3%)	ns
NYHA class	1.5 ± 0.7	1.4 ± 0.6	2.7 ± 0.5	<0.001
Diameter of IVC at deep inspiration (mm)	14.9 ± 4.3	14.3 ± 3.9	18.8 ± 5.6	0.038
Collapsibility index of IVC	0.27 ± 0.09	0.28 ± 0.09	0.19 ± 0.06	0.018
Collapsibility index of IVC ≤ 0.15	6/48 (12.5%)	3/42 (7.1%)	3/6 (50%)	0.020
Total bilirubin (mg/dl)	0.75 (0.60–0.98)	0.70 (0.60–0.90)	1.05 (0.80–2.38)	0.021
Albumin (g/l)	48.0 (44.0–49.0)	48.0 (44.0–49.0)	43.5 (37.8–49.0)	ns
γGT (U/l)	67.0 (40.5–96.0)	67.5 (45.8–102.0)	36.5 (22.8–70.8)	0.025
Platelet count (per mm^3)	193.5 (152.0–230.0)	200.0 (166.3–238.3)	147.5 (123.5–187.0)	0.034
Creatinine (mg/dl)	0.79 (0.66–0.94)	0.78 (0.64–0.91)	0.99 (0.78–1.38)	ns
GFR (ml/min)	105.9 (85.3–122.5)	107.8 (96.8–125.3)	82.4 (60.8–92.1)	0.007
MELD-XI score	9.44 (9.44–10.51)	9.44 (9.44–9.44)	11.24 (9.44–15.95)	0.014
MELD-Albumin score	6.43 (6.43–7.40)	6.43 (6.43–7.17)	8.42 (7.11–12.89)	0.003
ALBI score	−3.25 (−3.41–−2.97)	−3.27 (−3.42–−3.03)	−2.79 (−3.33–−2.36)	0.057
MELD-Albumin score ≥ 11 or ALBI score > −2.6 (%)	5/48 (10.4%)	2/42 (4.8%)	3/6 (50%)	0.010

UVH, univentricular heart; GFR, glomerular filtration rate; MELD, model of end-stage liver disease; MELD-XI, model of end-stage liver disease without international normalized ratio; ALBI; albumin-bilirubin; ns, not significant.
Mean ± standard deviation or median (interquartile interval) are used.
**Complete compared to incomplete/no Fontan palliation subgroup.*

Sample Preparation and RNA Isolation

In all patients and controls, blood samples for miRNA detection were collected in PAXgene™ blood tubes (Becton–Dickinson, Heidelberg, Germany) shortly after echocardiographic evaluation. All PAXgene™ blood tubes were stored at room temperature for at least 24 h to ensure complete lysis of the blood cells, then stored at −20°C for several days and finally transferred to −80°C for long-term storage until RNA isolation. Total RNA including miRNAs was isolated from blood samples using PAXgene™ Blood miRNA Kit on the QIAcube™ robot (Qiagen, Hilden, Germany) following the manufacturer's recommendations and included DNase I treatment (Qiagen). To confirm the absence of genomic DNA contamination, a conventional PCR with exon spanning primers for Glyceraldehyde 3-Phosphate Dehydrogenase (*GAPDH*) as previously described (22). The concentration of isolated total RNA was measured using the NanoDrop ND-2000 spectrophotometer (Thermo Fisher Scientific, Massachusetts, United States). RNA purity was estimated by examining the OD 260/280 and the OD 260/230 ratios. The qualities of total RNA were assessed using the Agilent Bioanalyser 2100 Eukaryote Total RNA Nano Series II (Agilent Technologies, California, United States).

Analysis of miRNAs by Microarray

The miRNA abundance profile from the 48 patients with UVH and 32 age- and gender-matched healthy controls was obtained from our previously generated and uploaded raw data to the NCBI GEO database (Accession ID: GSE136547) using SurePrint™ 8X60K Human v21 miRNA platforms (Agilent Technologies) (21). These platforms contain probes for the detection of 2,549 human miRNAs. An input amount of 100 ng of isolated RNA including miRNAs was labeled and subsequently hybridized to the miRNA microarray chip and the procedure was completed as previously described (23).

Analysis of miRNAs by RT-qPCR

Real-time quantitative PCR (RT-qPCR) validation analysis was performed using the StepOnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA, United States) and the *mi*Script PCR System (Qiagen) as previously described (24). For the validation step, we used the same samples that were used for the microarray analysis (48 UVH patients and 32 healthy controls) (21). Fifteen miRNAs (miR-101-3p, miR-144-5p, miR-18a-5p, miR-15a-5p, miR-17-3p, miR-96-5p, miR-18b-5p, miR-140-5p, miR-148a-3p, miR-29b-3p, miR-29c-3p, miR-210-3p, miR-125a-5p, miR-150-5p, and miR-324-3p) were chosen for RT-qPCR validation based on the significant abundance changes on the microarray analysis. Of these miRNAs, miR-29b-3p and miR-29c-3p were significantly correlated to the prognostic scores of advanced liver fibrosis/cirrhosis (MELD-Albumin and ALBI score). For miRNA abundance level detection, an amount of 250 ng of the total RNA was converted into complementary DNA (cDNA). The resulting cDNA was then diluted to have 0.5 ng/μL input material for miRNA detection. All RT-qPCR experiments were carried out using the Liquid Handling Robot QIAgility™ (Qiagen) before performing RT-qPCR. All primer assays used in the current study were provided

by Qiagen. Moreover, miRNA reverse transcription control (miRTC) (Qiagen) was performed to assess the performance of the reverse transcription reaction. The melting curve analysis was used to control the specificity of RT-qPCR products. The specificity of amplicons was further confirmed by agarose gel electrophoresis.

Data Analysis

Clinical data of the patients were collected from medical records. Echocardiography and ultrasonography were performed using a Vivid™ E9 Ultrasound System (GE Healthcare, Horten, Norway). The echocardiographic loops and ultrasonographic images were stored digitally and analyzed on an Echopac server (Echopac Version 6, GE Healthcare). The collapsibility index of the IVC was calculated according to the formula: the difference of maximum and minimum diameter of the IVC divided by the maximum diameter of the IVC (20). The formulas used for calculating MELD and modified MELD scores have been described previously (11). ALBI score was calculated according to the equation: ALBI score = ($\log_{10}$ bilirubin × 0.66) + (albumin × −0.085) and has been previously published (25). Echocardiographic and ultrasonographic data sets were assessed by investigators blinded to the laboratory results. Investigators of miRNA signatures were blinded to the clinical, echocardiographic, ultrasonographic, and laboratory data of the patients.

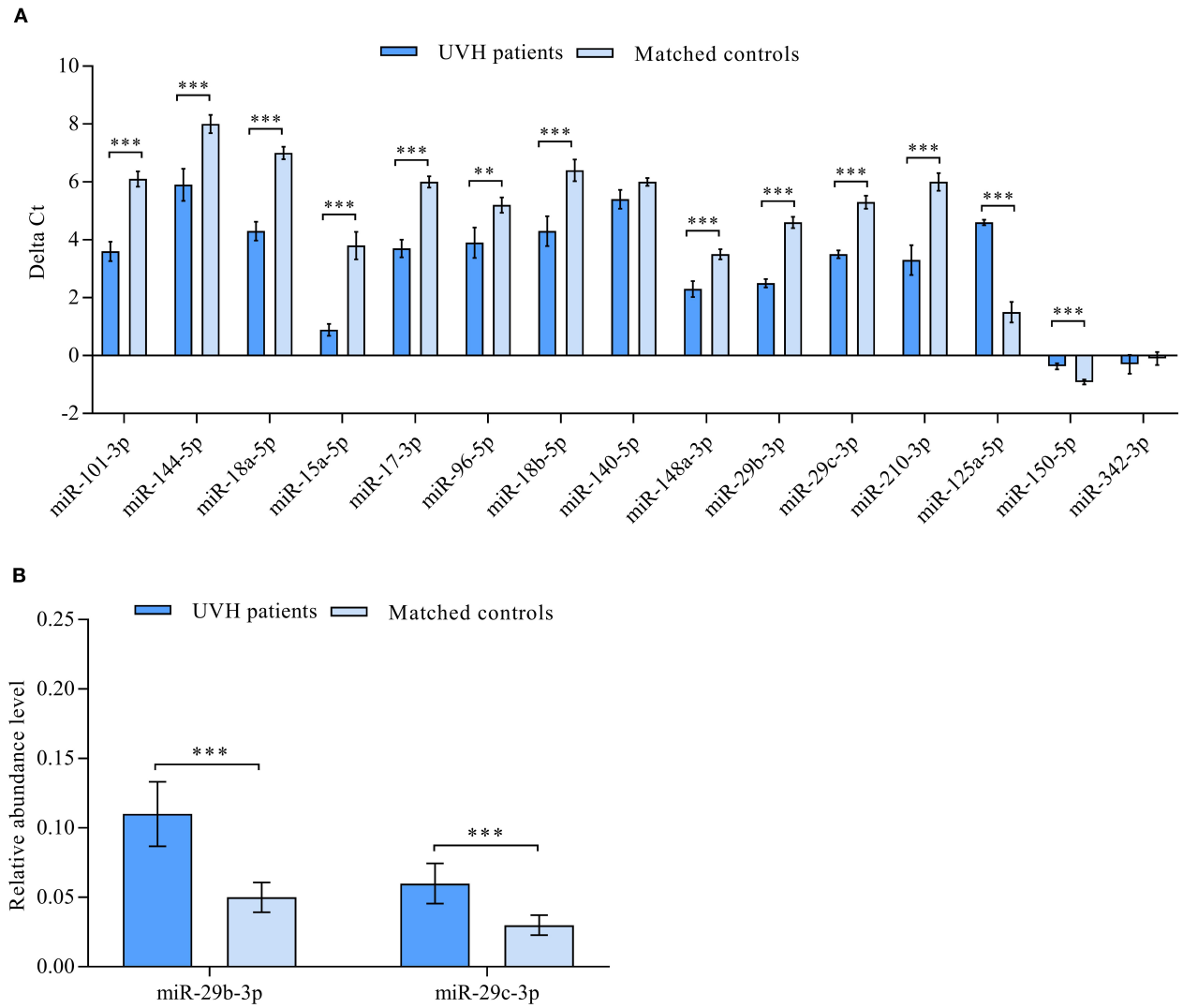

FIGURE 1 | (A) Validation of differentially abundant miRNAs in the blood of patients with univentricular heart (UVH) ($n = 48$) compared to age and gender-matched healthy controls ($n = 32$) as determined by RT-qPCR ($P < 0.05$). Mean ΔCt values of patients and healthy controls *(lower ΔCt values indicating higher abundance levels).* **(B)** Validation of clinically correlated differentially abundant miR-29b-3p and miR-29c-3p in the blood of UVH patients ($n = 48$) compared to age and gender-matched healthy controls ($n = 32$) as determined by RT-qPCR ($P < 0.05$). Mean relative abundance level ($2^{-\Delta Ct}$) of patients and healthy controls are shown with RNAU6B as an endogenous control for normalization. Un-paired-two-tailed *t*-tests and mean ± standard deviation (STDV) were used to evaluate differences in abundance levels. $^{**}P \leq 0.01$; $^{***}P \leq 0.001$.

Statistical Analysis

DataAssist™ Software v3.0 (Applied Biosystems) was used to calculate the relative abundance changes of miRNAs by the equation $2^{-\Delta Ct}$ with RNAU6B serving as an endogenous control as previously used and validated for this type of sample (21, 23, 24, 26–31). Adjustment for multiple testing was performed by controlling the false discovery rate (FDR) according to the approach of Benjamini and Hochberg (32). Clinical data were analyzed using standard statistical software (SPSS version 25; SPSS Inc., Chicago, Illinois). Continuous variables are expressed as mean ± standard deviation or median (interquartile interval) as appropriate. Differences between unpaired groups were analyzed using a Mann-Whitney-U test for continuous variables and a chi-square test (or Fisher exact test, if numbers were small) for nominal variables. Correlations were evaluated using Spearman's regression coefficient. Univariate and multivariate analysis was performed using binary logistic regression to identify the most significant factors of advanced liver fibrosis/cirrhosis. Variables entered into the multivariate model were those that gave statistically significant results in the univariate analysis and didn't show any multicollinearity. A two-tailed $P < 0.05$ was considered statistically significant.

RESULTS

Evaluation of Conventional Liver Parameters

Conventional ultrasonographic and laboratory liver parameters according to the status of Fontan palliation are presented in **Table 1.** As compared to patients with complete Fontan palliation, patients with incomplete or no Fontan palliation had a higher NYHA class, a lower glomerular filtration rate, and a lower collapsibility index of the IVC reflecting the presence of more severe global heart failure and liver congestion. Moreover, laboratory parameters of advanced liver fibrosis/cirrhosis such as total bilirubin or platelet count were also different between both subgroups. The calculated prognostic liver scores MELD-XI, MELD-Albumin, and ALBI scores were significantly higher in patients without Fontan palliation indicating more advanced stages of liver fibrosis/cirrhosis.

Identification of Abundant miRNAs

According to a miRNA microarray, a total of 50 miRNAs were found to be differentially abundant when comparing the samples from patients with UVH to healthy controls (21). RT-qPCR confirmed the direction of abundance changes and the significance of different abundances for 13 miRNAs. Specifically, out of 15 miRNAs used for the RT-qPCR validation, 11 miRNAs showed statistically significant higher abundance levels and two miRNAs showed statistically significant lower abundance levels in UVH patients compared with those in healthy controls ($P < 0.05$, FDR adjusted) (**Figure 1A**). There was no significant difference in the abundance level of miR-140-5p and miR-342-3p (**Figure 1A**). However, correlation analyses of significantly abundant miRNAs to prognostic scores of advanced liver fibrosis/cirrhosis (MELD-Albumin and ALBI score) identified miR-29b-3p and miR-29c-3p to be best related to these scores.

TABLE 2 | Correlation of miRNAs validated by RT-qPCR with liver-specific parameters and prognostic liver scores ($n = 48$).

Variables	miR-29b-3p		miR-29c-3p	
	r	*P*-value	*r*	*P*-value
Collapsibility index of IVC	−0.325	0.024	−0.359	0.012
Collapsibility index of IVC ≤ 0.15	0.336	0.020	0.340	0.018
Total bilirubin (mg/dl)	0.310	0.032	0.387	0.007
Albumin (g/l)	−0.366	0.010	−0.288	0.047
γGT (U/l)	–	ns	–	ns
Platelet count (per mm^3)	–	ns	−0.330	0.022
MELD-XI score	0.295	0.042	0.361	0.012
MELD-Albumin score	0.285	0.050	0.330	0.022
ALBI score	0.447	0.001	0.384	0.007
Death from any cause	0.283	0.051	0.276	0.058

IVC, inferior caval vein; MELD, model of end-stage liver disease; MELD-XI, model of end-stage liver disease without international normalized ratio; ALBI, albumin bilirubin; ns, not significant.

Thus, these two miRNAs were further validated by RT-qPCR indicating significantly higher abundance levels in UVH patients compared to healthy controls (**Figure 1B**).

Association of miRNAs With Measures of Liver Congestion and Prognostic Liver Fibrosis Scores

Correlation analysis revealed miR-29c-3p to be significantly related to the extent of liver congestion represented by a collapsibility index of the IVC ≤ 0.15, laboratory measures of severe liver dysfunction such as total bilirubin and platelet count as well as prognostic measures of advanced liver fibrosis/cirrhosis as indicated by the different MELD scores and ALBI score. In contrast, miR-29b-3p was only significantly related to albumin and ALBI score (**Table 2**). Relative abundance levels of miR-29b-3p and miR-29c-3p were significantly higher in patients with severe liver congestion as indicated by a collapsibility index of the IVC ≤ 0.15 (**Figure 2A**). Moreover, higher relative abundance levels of miR-29b-3p and miR-29c-3p were observed in patients with a higher MELD-Albumin score ≥ 11 and ALBI score > −2.6 both indicating advanced liver fibrosis/cirrhosis (**Figures 2B,C**) as well as in patients with incomplete or no Fontan palliation (**Figure 2D**).

Indicators of Advanced Liver Fibrosis/Cirrhosis

Logistic regression analysis was used to identify the most significant factors of advanced liver fibrosis/cirrhosis as indicated by a MELD-Albumin score ≥ 11 or an ALBI score > −2.6 in all patients (**Table 3**). Here, univariate analysis identified total bilirubin, albumin and miR-29c-3p to be the most important factors of advanced liver fibrosis/cirrhosis. After adjustment for palliation status, however, no statistical significance could be seen thus indicating the importance of palliation status for advanced liver fibrosis/cirrhosis in UVH patients.

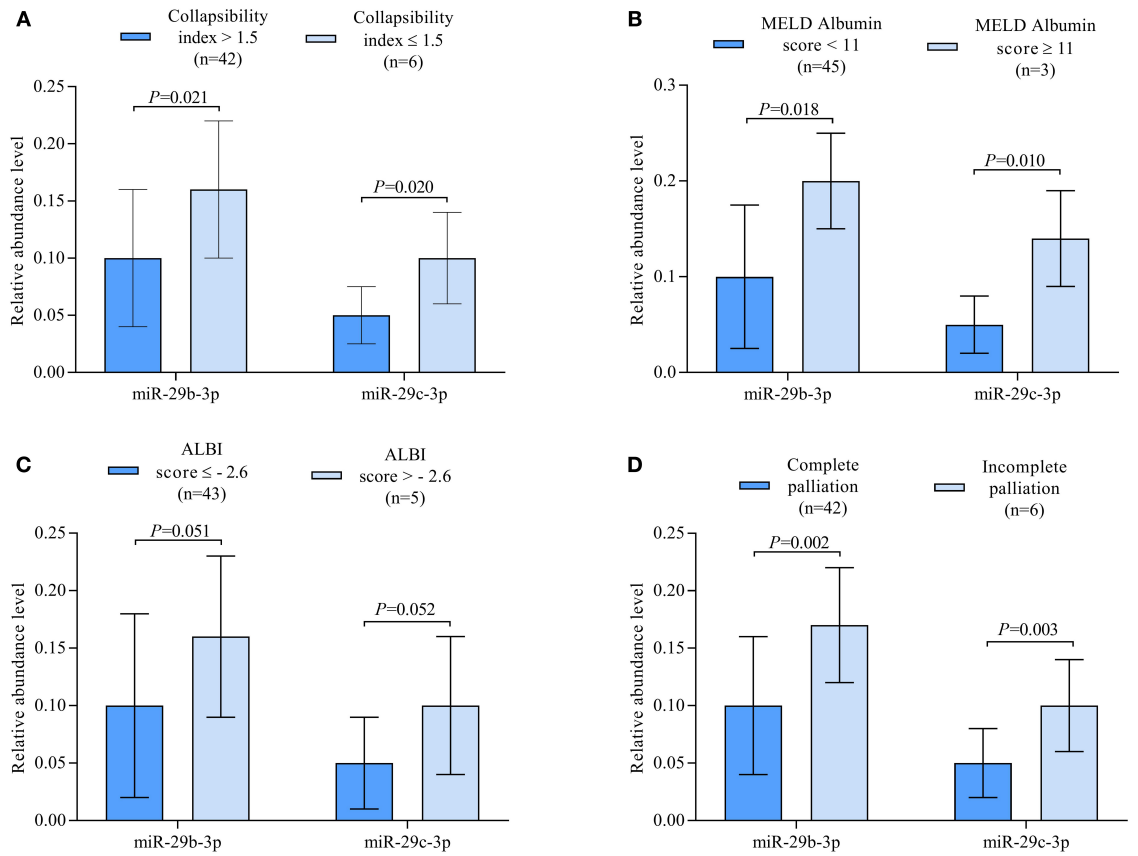

FIGURE 2 | (A) Boxplots illustrating relative abundance levels of miR-29b-3p and miR-29c-3p in UVH patients with ($n = 6$) and without ($n = 42$) severe liver congestion as indicated by a collapsibility index of the IVC ≤ 0.15. **(B)** Boxplots illustrating normalized expression levels of miR-29b-3p and miR-29c-3p in UVH patients with a MELD-Albumin score < 11 ($n = 45$) and ≥ 11 ($n = 3$) indicating advanced liver fibrosis/cirrhosis. **(C)** Boxplots illustrating normalized expression levels of miR-29b-3p and miR-29c-3p in UVH patients with an ALBI score ≤ −2.6 ($n = 43$) and > −2.6 ($n = 5$) indicating a higher degree of liver cirrhosis. **(D)** Boxplots illustrating normalized expression levels of miR-29b-3p and miR-29c-3p in UVH patients with complete ($n = 42$) and incomplete ($n = 6$) Fontan palliation. UVH, univentricular heart; IVC, inferior caval vein.

TABLE 3 | Determinants of advanced liver fibrosis/cirrhosis according to combined MELD-Albumin score ≥ 11 or ALBI score > −2.6.

	Univariate analysis			Multivariate analysis*		
Variables	**OR**	**95% CI**	***P*-value**	**OR**	**95% CI**	***P*-value**
Total bilirubin (mg/dl)	6.630	1.423–30.884	0.016	6.560	0.978–43.993	0.053
miR-29b-3p	20.944	0.671–653.432	0.083	6.279	0.072–549.233	0.421
Albumin (g/l)	0.424	0.199–0.901	0.026	0.280	0.051–1.532	0.142
Platelets (per mm^3)	0.991	0.975–1.008	0.291	1.000	0.981–1.020	0.990
miR-29c-3p	33.060	1.044–1046.655	0.047	26.803	0.343–2092.997	0.139

*OR, odds ratio; CI, confidence interval. *Adjusted for status of Fontan palliation.*

DISCUSSION

In UVH patients with and without Fontan palliation, the development of liver fibrosis/cirrhosis due to chronic liver congestion is a typical late complication (2, 3). These non-cardiac complications may limit potential heart transplantation candidacy and overall survival in this cohort of patients with advancing age. Thus, in addition to cardiorespiratory parameters, monitoring of liver status during the long-term follow-up of these patients is crucial. However, currently used methods to assess the grade or stage of liver fibrosis in these patients show significant limitations (6, 8). Recently, specific miRNAs are dysregulated in different liver diseases (18, 19) but are also involved in hepatic fibrosis progression (16) and thus might be used as valuable biomarkers. Therefore, our study aimed to identify those miRNAs in UVH patients with and without Fontan circulation

that are associated with advanced liver fibrosis/cirrhosis and to assess the most significant determinants of advanced liver fibrosis/cirrhosis in this cohort of patients.

Using conventional ultrasonographic and laboratory parameters such as collapsibility index and different scores of liver fibrosis/cirrhosis, we were able to demonstrate abnormal liver findings in our cohort of patients that were more severe in UVH patients with incomplete or no Fontan palliation as compared to patients with complete Fontan palliation (**Table 1**). Moreover, UVH patients without Fontan palliation were significantly older and presented with a higher NYHA class reflecting more severe systolic and diastolic dysfunction of the UVH in these patients. In contrast to published data indicating Fontan physiology as the leading cause of liver morbidities, we were able to show that UVH patients without Fontan palliation may also suffer from the same liver complications in older age.

In this study, miR-29b-3p and miR-29c-3p were found to be best associated with the prognostic liver scores MELD-Albumin and ALBI score in our patients using microarray analysis. After validation by RT-qPCR, miR-29b-3p turned out to be significantly related to ALBI score whereas miR-29c-3p was best related to MELD-Albumin, MELD-XI, and ALBI score (**Table 2**). It is of note that these liver scores represent advanced stages of liver fibrosis or even liver cirrhosis and also have prognostic value in patients with chronic liver and heart disease. In contrast to previous studies evaluating these miRNAs in other clinical settings, both miRNAs were upregulated in our patient cohort with or without Fontan palliation as compared to healthy controls (**Figure 1B**). However, it needs to be taken into consideration that the pathomechanism of liver fibrosis/cirrhosis after inflammatory or toxic liver damage may be different from congestive liver disease that was present in our studied patients. Moreover, higher abundance levels of both miRNAs were found in patients with a higher MELD-Albumin or ALBI score indicating advanced stages of liver fibrosis/cirrhosis (**Figures 2B,C**). Furthermore, abundance levels of both miRNAs were significantly elevated in UVH patients with incomplete or no Fontan palliation reflecting the more longstanding venous congestion of the liver, more advanced stages of liver fibrosis/cirrhosis, and a poorer prognosis in these patients as compared to patients with complete Fontan palliation (**Figure 2D**). Overall, both miRNAs were able to differentiate between lower and higher liver fibrosis/cirrhosis scores indicating the progression of liver pathology and worsening of prognosis.

Our results are in accordance with previous studies identifying the miR-29 member family as an indicator of advanced liver fibrosis/cirrhosis. In a previous liver biopsy study in patients with hepatitis B-associated liver fibrosis, a panel of 4 miRNAs was identified to be involved in the progression of liver fibrosis, namely miR-122, miR-146a-5p, miR-29c-3p, and miR-223 (33). In that study, miR-122 has been shown to differentiate between various fibrosis stages (F0-F4) and thus can be used as a precise fibrosis staging biomarker, especially for the early detection of liver fibrosis. In contrast, miR-29c-3p has been shown to be only significantly abundant in advanced liver fibrosis/cirrhosis in that study and is in agreement with previous studies identifying the miR-29 family as a marker of advanced liver fibrosis/cirrhosis (34, 35).

An increasing number of studies have demonstrated the down-regulation of miR-29 in human fibrotic disorders of multiple organs by transforming growth factor ß/*SMAD3* signaling (34–36). Moreover, miR-29 has been identified as a strong tumor suppressor rather than an oncogenic factor in most cancer studies (37) and many of these studies indicate that high miR-29 expression leads to increased survival (38, 39). However, there are also some studies reporting that higher levels of miR-29 in plasma are associated with lower survival rates (40, 41) and that up-regulation of miR-29c may be associated with higher relapse risk in patients with acute myeloid leukemia (42). In our study, the correlation of both miRNAs to all-cause mortality was statistically borderline (**Table 2**) what can be explained by the fact that mortality was mainly cardiac-related in this cohort of patients (21). Nevertheless, liver fibrosis and cirrhosis are important complications increasing with age and limiting survival in UVH patients with or without Fontan palliation. In our study, the elevated abundance levels of miR-29b-3p and miR-29c-3p in UVH patients with incomplete or no Fontan palliation may indicate the potential worse overall prognosis and limited survival in these patients (**Figure 2D**).

Study Limitations

This is the first study that aims to characterize signatures of miRNAs in UVH patients with and without Fontan palliation using a large panel of miRNAs for initial screening in order to identify those that may indicate advanced liver fibrosis/cirrhosis and thus might have clinical as well as prognostic impact in this cohort of patients. Since UVH disease is a rare congenital cardiac disorder, the sample size of our patient cohort is rather small. Moreover, non-invasive methods that reliably diagnose different stages of liver fibrosis in this cohort of patients are limited (6, 8). Thus, prognostic scores of advanced liver fibrosis/cirrhosis were used for miRNA screening in the present study. The cross-sectional rather than longitudinal design of the study may limit the predictive value of the measured miRNAs. Therefore, the measured miRNAs may have limited value in detecting earlier stages of liver fibrosis in such patients. Hence, a larger cohort of UVH patients should be evaluated and liver biopsies performed to assess the diagnostic value of miRNAs in the early and late stages of liver fibrosis in these patients.

CONCLUSIONS

In UVH patients with and without Fontan palliation, miR-29b-3p and miR-29c-3p are associated with measures of advanced liver fibrosis/cirrhosis and thus may be used as potential biomarkers in the risk assessment of these patients.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board approval/Ethikvotum Ärztekammer des Saarlandes: Ethical vote No. 73/09. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. The animal study was reviewed and approved by Ethical vote No. 73/09. Written informed consent was obtained from the owners for the participation of their animals in this study.

AUTHOR'S NOTE

This manuscript has been released as a pre-print at Research Square (43).

AUTHOR CONTRIBUTIONS

MA-H performed experimental work, particularly the miRNA isolation, array experiment, RT-qPCR validation, and helped in manuscript writing. EM designed the study, coordinated the molecular biology experiment, and edited the manuscript. MS helped in the experimental work and analyses. AK performed bioinformatics analysis. HA-K designed the study and diagnosed patients. TR-H designed the study, recruited and examined controls, diagnosed patients, collected blood samples, and helped in writing. All authors read and approved the final manuscript.

FUNDING

This study was supported by the German Heart Foundation for the research on adults with congenital heart disease and by the Hedwig Stalter Fund (2016).

REFERENCES

1. Ghaferi AA, Hutchins GM. Progression of liver pathology in patients undergoing the Fontan procedure: chronic passive congestion, cardiac cirrhosis, hepatic adenoma, and hepatocellular carcinoma. *J Thorac Cardiovasc Surg.* (2005) 129:1348–52. doi: 10.1016/j.jtcvs.2004.10.005
2. Wu FM, Kogon B, Earing MG, Aboulhosn JA, Broberg CS, John AS, et al. Liver health in adults with Fontan circulation: a multicenter cross-sectional study. *J Thorac Cardiovasc Surg.* (2017) 153:656–64. doi: 10.1016/j.jtcvs.2016.10.060
3. Goldberg DJ, Surrey LF, Glatz AC, Dodds K, O'Byrne ML, Lin HC, et al. Hepatic fibrosis is universal following fontan operation, and severity is associated with time from surgery: a liver biopsy and hemodynamic study. *J Am Heart Assoc.* (2017) 6:e004809. doi: 10.1161/JAHA.116.004809
4. Josephus Jitta D, Wagenaar LJ, Mulder BJ, Guichelaar M, Bouman D, van Melle JP. Three cases of hepatocellular carcinoma in Fontan patients: review of the literature and suggestions for hepatic screening. *Int J Cardiol.* (2016) 206:21–6. doi: 10.1016/j.ijcard.2015.12.033
5. Regev A, Berho M, Jeffers LJ, Milikowski C, Molina EG, Pyrsopoulos NT, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. *Am J Gastroenterol.* (2002) 97:2614–8. doi: 10.1111/j.1572-0241.2002.06038.x
6. Lemmer A, VanWagner L, Ganger D. Congestive hepatopathy: differentiating congestion from fibrosis. *Clin Liver Dis.* (2017) 10:139–43. doi: 10.1002/cld.676
7. Schachter JL, Patel M, Horton SR, Mike Devane A, Ewing A, Abrams GA. FibroSURE and elastography poorly predict the severity of liver fibrosis in Fontan-associated liver disease. *Congenit Heart Dis.* (2018) 13:764–70. doi: 10.1111/chd.12650
8. Wu FM, Earing MG, Aboulhosn JA, Johncilla ME, Singh MN, Odze RD, et al. Predictive value of biomarkers of hepatic fibrosis in adult Fontan patients. *J Heart Lung Transpl.* (2017) 36:211–9. doi: 10.1016/j.healun.2016.07.011
9. Abe S, Yoshihisa A, Takiguchi M, Shimizu T, Nakamura Y, Yamauchi H, et al. Liver dysfunction assessed by model for end-stage liver disease excluding INR (MELD-XI) scoring system predicts adverse prognosis in heart failure. *PLoS ONE.* (2014) 9:e100618. doi: 10.1371/journal.pone.0100618
10. Assenza GE, Graham DA, Landzberg MJ, Valente AM, Singh MN, Bashir A, et al. MELD-XI score and cardiac mortality or transplantation in patients after Fontan surgery. *Heart.* (2013) 99:491–6. doi: 10.1136/heartjnl-2012-303347
11. Chen Y, Liu YX, Seto WK, Wu MZ, Yu YJ, Lam YM, et al. Prognostic value of hepatorenal function by modified Model for End-stage Liver Disease (MELD) score in patients undergoing tricuspid annuloplasty. *J Am Heart Assoc.* (2018) 7:e009020. doi: 10.1161/JAHA.118.009020
12. Chen RC, Cai YJ, Wu JM, Wang XD, Song M, Wang YQ, et al. Usefulness of albumin-bilirubin grade for evaluation of long-term prognosis for hepatitis B-related cirrhosis. *J Viral Hepatitis.* (2017) 24:238–45. doi: 10.1111/jvh.12638
13. Lei Q, Zhang Y, Ke C, Yan C, Huang P, Shen H, et al. Value of the albumin-bilirubin score in the evaluation of hepatitis B virus-related acute-on-chronic liver failure, liver cirrhosis, and hepatocellular carcinoma. *Exp Therapeutic Med.* (2018) 15:3074–9. doi: 10.3892/etm.2018.5748
14. Krauskopf J, de Kok TM, Schomaker SJ, Gosink M, Burt DA, Chandler P, et al. Serum microRNA signatures as "liquid biopsies" for interrogating hepatotoxic mechanisms and liver pathogenesis in human. *PLoS ONE.* (2017) 12:e0177928. doi: 10.1371/journal.pone.0177928
15. Musaddaq G, Shahzad N, Ashraf MA. Circulating liver-specific microRNAs as noninvasive diagnostic biomarkers of hepatic diseases in human. *Biomarkers.* (2019) 24:103–9. doi: 10.1080/1354750X.2018.1528631
16. Ezhilarasan D. Role of microRNAs in hepatic fibrosis progression. *J Appl Pharm Sci.* (2018) 8:174–8. doi: 10.7324/JAPS.2018.8524
17. Alles J, Fehlmann T, Fischer U, Backes C, Galata V, Minet M, et al. An estimate of the total number of true human miRNAs. *Nucleic Acids Res.* (2019) 47:3353–64. doi: 10.1093/nar/gkz097
18. Hayes CN, Chayama K. MicroRNAs as biomarkers for liver disease and hepatocellular carcinoma. *Int J Mol Sci.* (2016) 17:280. doi: 10.3390/ijms17030280
19. Schueller F, Roy S, Vucur M, Trautwein C, Luedde T, Roderburg C. The role of miRNAs in the pathophysiology of liver diseases and toxicity. *Int J Mol Sci.* (2018) 19:261. doi: 10.3390/ijms19010261
20. Rudski LG, Lai WW, Afilalo J, Hua L, Handschumacher MD, Chandrasekaran K, et al. Guidelines for the echocardiographic assessment of the right heart in adults: a report from the American society of echocardiography endorsed by the European association of echocardiography, a registered branch of the European society of cardiology, and the Canadian society of echocardiography. *J Am Soc Echocardiogr.* (2010) 23:685–713. doi: 10.1016/j.echo.2010.05.010
21. Abu-Halima M, Meese E, Saleh MA, Keller A, Abdul-Khaliq H, Raedle-Hurst T. Micro-RNA 150-5p predicts overt heart failure in patients with univentricular hearts. *PLoS ONE.* (2019) 14:e0223606. doi: 10.1371/journal.pone.0223606
22. Abu-Halima M, Ludwig N, Hart M, Leidinger P, Backes C, Keller A, et al. Altered micro-ribonucleic acid expression profiles of extracellular microvesicles in the seminal plasma of patients with oligoasthenozoospermia. *Fertil Steril.* (2016) 106:1061–9.e3. doi: 10.1016/j.fertnstert.2016.06.030
23. Abu-Halima M, Weidinger J, Poryo M, Henn D, Keller A, Meese E, et al. Micro-RNA signatures in monozygotic twins discordant for congenital heart defects. *PLoS ONE.* (2019) 14:e0226164. doi: 10.1371/journal.pone.0226164

24. Abu-Halima M, Kahraman M, Henn D, Radle-Hurst T, Keller A, Abdul-Khaliq H, et al. Deregulated microRNA and mRNA expression profiles in the peripheral blood of patients with Marfan syndrome. *J Transl Med.* (2018) 16:60. doi: 10.1186/s12967-018-1429-3
25. Johnson PJ, Berhane S, Kagebayashi C, Satomura S, Teng M, Reeves HL, et al. Assessment of liver function in patients with hepatocellular carcinoma: a new evidence-based approach-the ALBI grade. *J Clin Oncol.* (2015) 33:550–8. doi: 10.1200/JCO.2014.57.9151
26. Abu-Halima M, Ludwig N, Radle-Hurst T, Keller A, Motsch L, Marsollek I, et al. Characterization of micro-RNA profile in the blood of patients with marfan's syndrome. *Thorac Cardiovasc Surg.* (2017) 66:116–24. doi: 10.1055/s-0037-1604083
27. Abu-Halima M, Meese E, Keller A, Abdul-Khaliq H, Radle-Hurst T. Analysis of circulating microRNAs in patients with repaired tetralogy of Fallot with and without heart failure. *J Transl Med.* (2017) 15:156. doi: 10.1186/s12967-017-1255-z
28. Abu-Halima M, Oberhoffer FS, El Rahman MA, Jung AM, Zemlin M, Rohrer TR, et al. Insights from circulating microRNAs in cardiovascular entities in turner syndrome patients. *PLoS ONE.* (2020) 15:e0231402. doi: 10.1371/journal.pone.0231402
29. Abu-Halima M, Poryo M, Ludwig N, Mark J, Marsollek I, Giebels C, et al. Differential expression of microRNAs following cardiopulmonary bypass in children with congenital heart diseases. *J Transl Med.* (2017) 15:117. doi: 10.1186/s12967-017-1213-9
30. Henn D, Abu-Halima M, Falkner F, Wermke D, Meme LG, Kuhner C, et al. Micro-RNA-regulated proangiogenic signaling in arteriovenous loops in patients with combined vascular and soft-tissue reconstructions: revisiting the nutrient flap concept. *Plast Reconstr Surg.* (2018) 142:489–502. doi: 10.1097/PRS.0000000000004750
31. Henn D, Abu-Halima M, Wermke D, Falkner F, Thomas B, Kopple C, et al. MicroRNA-regulated pathways of flow-stimulated angiogenesis and vascular remodeling *in vivo*. *J Transl Med.* (2019) 17:22. doi: 10.1186/s12967-019-1767-9
32. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol.* (1995) 57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
33. Wang TZ, Lin DD, Jin BX, Sun XY, Li N. Plasma microRNA: a novel non-invasive biomarker for HBV-associated liver fibrosis staging. *Exp Therap Med.* (2019) 17:1919–29. doi: 10.3892/etm.2018.7117
34. Roderburg C, Urban GW, Bettermann K, Vucur M, Zimmermann H, Schmidt S, et al. Micro-RNA profiling reveals a role for miR-29 in human and murine liver fibrosis. *Hepatology.* (2011) 53:209–18. doi: 10.1002/hep.23922
35. Noetel A, Kwiecinski M, Elfimova N, Huang J, Odenthal M. microRNA are central players in anti- and profibrotic gene regulation during liver fibrosis. *Front Physiol.* (2012) 3:49. doi: 10.3389/fphys.2012.00049
36. Cushing L, Kuang P, Lu J. The role of miR-29 in pulmonary fibrosis. *Biochem Cell Biol.* (2015) 93:109–18. doi: 10.1139/bcb-2014-0095
37. Kwon JJ, Factora TD, Dey S, Kota J. A systematic review of miR-29 in cancer. *Mol Ther Oncolytics.* (2019) 12:173–94. doi: 10.1016/j.omto.2018.12.011
38. Wu DW, Hsu NY, Wang YC, Lee MC, Cheng YW, Chen CY, et al. c-Myc suppresses microRNA-29b to promote tumor aggressiveness and poor outcomes in non-small cell lung cancer by targeting FHIT. *Oncogene.* (2015) 34:2072–82. doi: 10.1038/onc.2014.152
39. Zhu C, Wang Y, Kuai W, Sun X, Chen H, Hong Z. Prognostic value of miR-29a expression in pediatric acute myeloid leukemia. *Clin Biochem.* (2013) 46:49–53. doi: 10.1016/j.clinbiochem.2012.09.002
40. Joerger M, Baty F, Fruh M, Droege C, Stahel RA, Betticher DC, et al. Circulating microRNA profiling in patients with advanced non-squamous NSCLC receiving bevacizumab/erlotinib followed by platinum-based chemotherapy at progression (SAKK 19/05). *Lung Cancer.* (2014) 85:306–13. doi: 10.1016/j.lungcan.2014.04.014
41. Zhang L, Ye Y, Tu H, Hildebrandt MA, Zhao L, Heymach JV, et al. MicroRNA-related genetic variants in iron regulatory genes, dietary iron intake, microRNAs and lung cancer risk. *Ann Oncol.* (2017) 28:1124–9. doi: 10.1093/annonc/mdx046
42. Butrym A, Rybka J, Baczynska D, Poreba R, Kuliczkowski K, Mazur G. Clinical response to azacitidine therapy depends on microRNA-29c (miR-29c) expression in older acute myeloid leukemia (AML) patients. *Oncotarget.* (2016) 7:30250–7. doi: 10.18632/oncotarget.7172
43. Abu-Halima M, Meese E, Saleh MA, Keller A, Abdul-Khaliq H, Raedle-Hurst T. MicroRNA 29c-3p indicates advanced liver fibrosis/cirrhosis in univentricular heart patients with and without Fontan palliation. *Res Square.* (2020) *[preprint]*. doi: 10.21203/rs.2.22992/v1

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

frontiers in Cell and Developmental Biology

ORIGINAL RESEARCH
published: 11 January 2021
doi: 10.3389/fcell.2020.619079

Long Non-coding RNA Aerrie Controls DNA Damage Repair via YBX1 to Maintain Endothelial Cell Function

*Tan Phát Pham[1], Diewertje I. Bink[1], Laura Stanicek[1,2], Anke van Bergen[1], Esmee van Leeuwen[1], Yvonne Tran[1], Ljubica Matic[3], Ulf Hedin[3], Ilka Wittig[2], Stefanie Dimmeler[2,4] and Reinier A. Boon[1,2,4]**

[1] Department of Physiology, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, Netherlands, [2] Institute of Cardiovascular Regeneration, Goethe University, Frankfurt, Germany, [3] Vascular Surgery Division, Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden, [4] German Center for Cardiovascular Research DZHK, Partner Site Frankfurt Rhine-Main, Berlin, Germany

OPEN ACCESS

Edited by:
Maarten M. G. van den Hoogenhof, Heidelberg University Hospital, Germany

Reviewed by:
Kewal Asosingh, Cleveland Clinic, United States
Leon J. De Windt, Maastricht University, Netherlands

***Correspondence:**
Reinier A. Boon
boon@med.uni-frankfurt.de

Specialty section:
This article was submitted to Molecular Medicine, a section of the journal Frontiers in Cell and Developmental Biology

Received: 19 October 2020
Accepted: 07 December 2020
Published: 11 January 2021

Citation:
Pham TP, Bink DI, Stanicek L, van Bergen A, van Leeuwen E, Tran Y, Matic L, Hedin U, Wittig I, Dimmeler S and Boon RA (2021) Long Non-coding RNA Aerrie Controls DNA Damage Repair via YBX1 to Maintain Endothelial Cell Function. Front. Cell Dev. Biol. 8:619079. doi: 10.3389/fcell.2020.619079

Aging is accompanied by many physiological changes. These changes can progressively lead to many types of cardiovascular diseases. During this process blood vessels lose their ability to maintain vascular homeostasis, ultimately resulting in hypertension, stroke, or myocardial infarction. Increase in DNA damage is one of the hallmarks of aging and can be repaired by the DNA signaling and repair system. In our study we show that long non-coding RNA Aerrie (linc01013) contributes to the DNA signaling and repair mechanism. Silencing of Aerrie in endothelial cells impairs angiogenesis, migration, and barrier function. Aerrie associates with YBX1 and together they act as important factors in DNA damage signaling and repair. This study identifies Aerrie as a novel factor in genomic stability and as a binding partner of YBX1 in responding to DNA damage.

Keywords: LncRNA – long non-coding RNA, endothelial cell, aging, DNA damage, cardiovascular disease

INTRODUCTION

Cardiovascular disease (CVD) is the main cause of death worldwide at old age (Benjamin et al., 2019). Structural and functional changes of the vasculature accumulate throughout life resulting in increased arterial thickening, stiffness, and dysfunctional endothelium (Lakatta and Levy, 2003a). These changes result in increased systolic pressure, increased risk of development of atherosclerosis, myocardial infarction, hypertension, and stroke (Lakatta and Levy, 2003b). Aging is causally linked with many cellular changes leading to replicative senescence and inflammation. Endothelial cell senescence is associated with reduced activity of endothelial nitric oxide synthase, less responsiveness to hemodynamic forces, and inflammation (Wang et al., 2007; Kang et al., 2009; Collins and Tzima, 2011). Senescence begins with shortening of telomeres that ultimately exposes an uncapped free double-stranded chromosome end leading to activation of the DNA damage response pathway (Childs et al., 2015). Understanding the mechanisms may open new paths for intervention of CVDs.

The DNA damage response pathway is well described and is important to maintain the genomic integrity of the aging cell. In response to DNA damage, multiple intra-cellular signaling

cascades are activated in order to coordinate the appropriate repair of DNA damage (Shiloh and Ziv, 2013; Blackford and Jackson, 2017). This so-called DNA damage response (DDR) is controlled by three phosphoinositide 3-kinase (PI3K)-related kinases: Ataxia-telangiectasia mutated (ATM), ATM- and Rad3- related (ATR), and DNA-dependent protein kinase (DNA-PK) (Blackford and Jackson, 2017). Double-strand breaks (DSBs) can be recognized by the Ku70/80 heterodimer, leading to recruitment and activation of DNA-PK promoting non-homologous end joining (NHEJ) (Davis et al., 2014). DSBs are also recognized by the Mre11–Rad50–Nbs1 (MRN) complex leading to either homologous recombination which is not error-prone or NHEJ (Stracker and Petrini, 2011). Capture of DNA ends by the MRN complex leads to the rapid autophosphorylation of ATM (Bakkenist and Kastan, 2003). Once activated, ATM phosphorylates histone protein H2AX at serine 139 at the site of the DSB. H2AX is associated with the recruitment of necessary DNA repair factors to the site of DSB (Paull et al., 2000; Celeste et al., 2002). Additionally, ATM phosphorylates cell-cycle checkpoint protein CHK2 (Marechal and Zou, 2013). Both ATM and CHK2 are able to phosphorylate P53, thereby inhibiting the interaction of P53 with ubiquitin ligase MDM2, resulting in rapid P53 stabilization (Chehab et al., 2000). Due to this stabilization, P53 can control cell fate decisions, such as cell-cycle arrest, DNA repair, apoptosis, senescence, and metabolic reprogramming (Li et al., 2012; Moulder et al., 2018). Long non-coding RNAs (lncRNAs) are known to be involved in DNA repair. Unraveling their implication in DNA repair opens further possibilities in understanding the mechanisms underlying CVD (Thapar, 2018; Haemmig et al., 2020).

Long non-coding RNAs are longer than 200 nucleotides, transcribed by RNA pol-II, and are not translated into protein. These transcripts share several properties with protein coding RNAs, being 5'capped, spliced and 3' polyadenylated (Guttman et al., 2009). In contrast, some lncRNAs such as MALAT1 and MEN β are not polyadenylated, but a highly conserved triple helical structure is generated at the 3' end to stabilize their structure (Wilusz et al., 2012). LncRNAs contribute to a variety of physiological and pathophysiological cellular processes in the cardiovascular system (Uchida and Dimmeler, 2015). For example, they influence vessel outgrowth, remodeling, and angiogenesis by regulation of cellular processes like proliferation, apoptosis, and migration (Simion et al., 2019). The molecular functions of lncRNAs are diverse and dependent on the subcellular location. In the nucleus, lncRNAs are described as post-transcriptional regulators by binding to chromatin modifiers. Moreover, they regulate transcription by binding to transcription factors, mRNA splicing by binding to pre-mRNAs, and other various processes by binding to ribonucleic protein components. In the cytoplasm lncRNAs regulated microRNA function by acting as a miRNA sponge or they can bind to mRNA to alter its stability (Boon et al., 2016).

We identified the novel lncRNA Aerrie and provide experimental evidence for its involvement in the DNA damage repair pathway in the cardiovascular system. We show that Aerrie is expressed in the endothelium and is upregulated during aging. Loss of Aerrie reduces endothelial function due to increased DNA damage. Aerrie is associated with Y-Box protein 1 (YBX1) and both are crucial for DNA damage repair. Overexpression of Aerrie increased the efficiency of DNA repair upon genotoxic damage. Together, these results show a crucial role of lncRNA Aerrie in regulating endothelial cell aging.

MATERIALS AND METHODS

Cell Culture

Human umbilical vein endothelial cells (HUVECs) were purchased from Lonza (batch p1028 and p1032) and cultured in endothelial cell medium (ScienceCell, 1001), supplemented with ECGS (ScienceCell, 1052), penicillin/streptomycin (ScienceCell, 0503), and 5% fetal bovine serum (ScienceCell, 0025). HUVECs were used between passage 1 and 5 for experiments unless stated otherwise. HUVECs were stimulated with unidirectional flow with 20 dyn/cm^2 for 72 h to induce laminar shear stress and oscillatory flow with 20 dyn/cm^2 for 14 h to induce turbulent flow. For endothelial-to-mesenchymal (endMT) experiments, HUVECs were stimulated with IL-1β and TGF-β2 for 72 h. Hek293T cells were purchased from ATCC and cultured in Dulbecco's Modified Eagle Medium (Thermo Fisher, 31966021) supplemented with 10% FCS, 1% Pyruvate, 1% D-glucose, and 1% penicillin/streptomycin, 1% minimum essential media non-essential amino acid mix (Sigma-Aldrich, M7145). Cells were cultured at 37°C with 5% CO_2. Cell numbers were determined with the Countess II cell counter (Thermo Fisher). All cell types were cultured at 37°C in a 5% CO_2 atmosphere and tested negative for mycoplasma.

BiKE

Patients undergoing surgery for symptomatic or asymptomatic, high-grade (>50% NASCET) (Naylor et al., 2003) carotid stenosis at the Department of Vascular Surgery, Karolinska University Hospital, Stockholm, Sweden, were enrolled in the Biobank of Karolinska Endarterectomies (BiKE) study. Symptoms of plaque instability were defined as transitory ischemic attack, minor stroke, and amaurosis fugax. Patients without qualifying symptoms within 6 months prior to surgery were categorized as asymptomatic and indication for carotid endarterectomy based on results from the Asymptomatic Carotid Surgery Trial (ACST) (Halliday et al., 2010). Carotid endarterectomies (carotid plaques) were collected at surgery (totally 127) and normal artery controls were obtained from nine macroscopically disease-free iliac arteries and one aorta from organ donors without any history of cardiovascular disease (totally 10). All samples were collected with informed consent from patients or organ donors' guardians. All human studies were approved by the regional Ethics Committee. The BiKE study follows the principles outlined in the declaration of Helsinki. The BiKE study cohort demographics, details of sample collection, processing, and large-scale profiling analyses were previously extensively described (Perisic et al., 2013, 2016). The microarray dataset is available from Gene Expression Omnibus (GSE21545). Global gene expression profiles have been analyzed by Affymetrix HG-U133 plus 2.0 Genechip microarrays in 127 patients' plaque tissues

(n = 87 symptomatic and n = 40 asymptomatic) and n = 10 non-atherosclerotic (normal) arteries.

ISHD Samples

Left ventricular tissues from ISHD samples were acquired from the University of Sydney (Sydney, NSW, Australia), with the ethical approval of the Human Research Ethics Committee (number 2012/2814). Explanted left ventricular heart tissue of healthy donors was used as control samples; the donors died from a non-cardiac cause, typically motor vehicle accidents. These healthy donor samples were also acquired from the University of Sydney. Patient information are shown in **Supplementary Table 1**.

RT-qPCR

Total RNA from cultured HUVECs was isolated with direct-zol RNA miniprep (Zymo Research, R2052) according to the manufacturer's protocol. For Real Time Quantitative PCR (RT-qPCR) analysis, 100–1000 ng total RNA was reverse transcribed using iScript cDNA synthesis Kit (Bio-Rad, #1708891). RT-qPCR was performed with iQ SYBR Green Supermix (Bio-RAD, #170-8886) in the Bio-Rad CFX96 Touch Real-time PCR Detection system. Ribosomal protein, large, P0 (RPLP0) or glyceraldehyde-3-phosphate dehydrogenase (GAPDH) were used for normalization of the samples. Gene expression analysis was done using the $2^{-\Delta\ CT}$ method. Primer sequences are listed in the **Supplementary Table 2**.

Nuclear and Cytoplasmic Cell Fractionation

Nuclear and cytoplasmic RNA fractions were separated as described in the literature (Gagnon et al., 2014). All steps were performed on 4°C. Briefly, HUVECs were collected by cell scraping and centrifugation (5 min at 500 × g, 4°C). Cell pellets were treated with cytoplasmic lysis buffer [10 mM Tris–HCl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.5% Nonidet P-40 (NP-40)] and incubated on ice for 5 min. Cells were spun down and supernatants were collected as cytoplasmic fraction. Pelleted nuclei were washed with cytoplasmic lysis buffer and incubated with nucleic lysis buffer (10 mM Tris–HCl pH 7.5, 150 mM NaCl, 3 mM $MgCl_2$) for 5 min on ice. Cytoplasmic fraction and pelleted nuclei were treated with Trizol and RNA was extracted from both fractions. Equal volumes of RNA were used for cDNA synthesis.

LNA-GapmeRs/siRNAs

Cultured HUVECs were transfected at 50–70% confluence with 50 nM LNA-GapmeRs/siRNAs (Qiagen, Hilden, Germany) using Lipofectamine RNAiMax (Life Technologies) according to the manufacturer's protocol in serum reduced OptiMEM medium (Life Technologies). The medium was changed after 4 h of transfection to ECM. Sequences of the LNA-GapmeRs/siRNAs can be found in the **Supplementary Table 2**.

Lentiviral Constructs

Aerrie (NR_038981.1) full length cDNA was cloned into pLenti4/v5 (Life Technologies). Lentivirus stocks were produced in HEK293T cells using pCMVΔR8.91 as packaging plasmid and pMD2.G (Addgene#12259) as vesicular stomatitis virus G glycoprotein envelope expressing plasmid (Zufferey et al., 1997). Empty vectors were used as mock control. Transduction was done for 24 h.

In vitro Sprouting Assay

Human umbilical vein endothelial cells were seeded in ECM containing methylcellulose (20%) into a non-adherent round-bottom 96-well plate to allow one spheroid to be formed per well. Spheroids were collected after 24 h and embedded into a collagen type I gel (BD Biosciences) in a 12-well plate. After polymerization (30 min), ECM was added on top and the plate was incubated for 24 h. Spheroids were fixated with 10% formaldehyde for 15 min and analyzed by bright field microscopy (Olympus IX50, magnification: 10 ×) using the Optimas 6.5 imaging software (Media cybernetics). Cumulative and discontinuous sprout length of each spheroid was measured. Subtraction of the cumulative sprout length from the maximal distance of the migrated cell was defined as discontinuous sprout length. For the rescue experiments with lentiviral Aerrie, HUVECs were stimulated with 50 nM Doxorubicin (Sigma-Aldrich, D1515-10MG) for 2 and 4 h. Equal volumes of DMSO (Sigma-Aldrich, 472301) were used as control.

Scratch Wound Healing Assay

Cell migration chambers (Ibidi) were placed in a 24-well tissue culture dish. Cells were seeded into each half-chamber and grown overnight. After removal of the inserts, lateral cell migration was visualized by bright field microscopy (EVOS XL AMG, magnification: 4 ×). Pictures were taken at 0 and 6 h after removal of the inserts. Quantitative assay analysis was performed in ImageJ. The area covered by cells was determined for the indicated time points.

Endothelial Barrier and Wound Healing

Endothelial barrier function was measured by the ECIS system (Applied BioPhysics). 100,000 HUVECs were seeded per well into a gelatin-coated (1%) 10WE plate (Applied BioPhysics). Endothelial barrier integrity was analyzed after 24 h when cells formed a stable monolayer. Barrier resistance (R_b) was measured by applying an alternating current of 4000 Hz resulting in a potential which is detected by the ECIS instrument Zθ (Applied BioPhysics), impedance is determined according to Ohm's law. Cell migration was analyzed by inducing cell wounding through lethal electroporation. Wound repair was observed over a period of 4 h.

Western Blot Analysis

Human umbilical vein endothelial cells were lysed in Triton X-100 buffer containing benzonase (Santa Cruz Biotechnology, Cas 9025-65-4), protease inhibitors (Thermo Fisher, Halt) and phosphatase inhibitors (Thermo Fisher, Halt) for 1 h on a spinning wheel on 4°C. After centrifugation with 15000 × g, protein content was analyzed with Pierce BCA protein assay kit (Thermo Fisher). 10 μg of protein was loaded on Sodium

dodecyl sulfate (SDS) gels and blotted on 0.2 μm nitrocellulose membranes (GE healthcare). GAPDH was used as a loading control. Antibodies are listed in the **Supplementary Table 3**.

RNA Pulldown

Human umbilical vein endothelial cells were crosslinked with 50 mJ UV light and treated in lysis buffer (50 mM Tris–HCl, 150 mM NaCl, 0.5% NaCl, 0.5% NP-40, 80 U Ribolock protease inhibitor (Thermo Fisher) and volumes were adjusted to 1 ml. Streptavidin-coupled beads C1 (Thermo Fisher) were pre-blocked using ytRNA (Thermo Fisher, AM7119) and glycogen (Thermo Fisher, R0561). For selection of RNP complexes, lysates were pre-cleared with 50 μl pre-blocked beads for 2 h at 4°C and subsequently incubated with 100 pmol 2'O-MeRNA oligonucleotides targeting Aerrie or a scrambled control oligonucleotide for 1 h at 37°C. Sequence of the oligos are listed in the **Supplementary Table 2**. RNP-oligonucleotide complexes were captured using 25 μl pre-blocked (yeast tRNA, glycogen; both 0.2 mg/ml) streptavidin C1 beads (Thermo Fisher) for 1 h at 37°C. Beads were washed thoroughly with washing buffer (50 mM Tris–HCl pH8, 150 mM NaCl, 0.05% NP-40) and eluted with biotin (250 mM) for 1 h at RT. Eluates were analyzed by RT-qPCR and mass spectrometry.

Crosslinking RNA Immunoprecipitation

Human umbilical vein endothelial cells were crosslinked with 50 mJ UV light and lysed with total lysis buffer (50 mM, Tris–HCl pH8, 150 mM, NaCl, 0.5% NP-40, and protease inhibitors). The lysates were cleared by centrifugation at 15000 × *g* and incubated with 50 μl protein G magnetic beads (Life Technologies) coated with antibodies listed in the **Supplementary Table 3**, overnight at 4°C, and then washed with lysis buffer containing 0.05% NP-40. RNA was recovered after protein digestion with Proteinase K (Thermo Fisher, EO0491) by phenol/chloroform/isoamyl extraction and analyzed by RT-qPCR.

Mass Spectrometry

RNA pulldown eluates were analyzed by liquid chromatography/MS using a QExactive Plus (Thermo Fisher) mass spectrometer equipped with an ultra-high performance liquid chromatography unit (Dionex Ultimate 3000, Thermo Fisher) and a Nanospray Flex Ion-Source (Thermo Fisher). Data analysis was performed in MaxQuant 1.5.3.30, Perseus 1.5.6.0, and plotted in Excel (Microsoft Office).

Comet Assay

Human umbilical vein endothelial cells were stimulated with 50 nM Doxorubicin (Sigma-Aldrich, D1515-10MG) for 2 and 4 h ECM. As a positive control HUVECs were stimulated with 500 μM H_2O_2. Equal volumes of DMSO (Sigma-Aldrich, 472301) were used as control. HUVECs were incorporated in low melting agarose. Cells were lysed with lysis solution (Trevigen, #4250-050-01). DNA was unwinded with alkaline solution (200 mM NaOH, 1 mM EDTA). Electrophoresis was performed at 24V and at least 300 mA for 30 min. Comets were stained with SYBR gold (Thermo Fisher, S11494) for 30 min and visualized with a Zeiss DIMI fluorescent microscope. Comets were analyzed by Cometscore software (RexHoover).

SCRINSHOT RNA FISH

Subcellular RNA localization was analyzed by SCRINSHOT RNA FISH (Sountoulidis et al., 2020). Briefly, cells were seeded in 1%-gelatin-coated 12-well removable slides (81201, Ibidi) and fixated with 4% PFA after 24 h. Slides were treated with 0.1 M HCl for permeabilization and dehydrated with alcohol to mount hybridization chambers (Grace Bio-labs). After rehydration, the cells were incubated for 30 min with blocking solution containing 0.1 M oligo-dT, Ampligase buffer, 0.05 M KCl, 20% formamide, 0.2 μg/μl BSA, 1 U/μl Ribolock (Thermo Fisher) and 0.2 μg/μl tRNA's (Ambion). Padlock probes were incubated for 15 min 55°C and 120 min at 45°C, washed with 10% formamide in 2× SSC and ligated with SplintR (NEB) for 16–24 h at 25°C. Rolling circle amplification was performed with Φ29 polymerase (Lucigen) and the primer TAAATAGACGCAGTCAGT*A*A. The amplified products were fixated with 4% PFA and detection oligos were incubated at RT for 60 min. Coverslips were mounted with SlowFade^{TM} Gold Antifade mounting medium.

Statistical Analysis

Data are expressed as mean ± standard error of the mean (SEM), except the human data from BiKE which is expressed as mean standard deviation. Graphpad Prism 8 was used for statistical analysis. Data were tested with paired or unpaired Student's *t*-test or Mann-Whitney test when comparing two groups. Analysis of variance (ANOVA) followed by Tukey's post-test was performed for multiple comparisons. Statistical significance was depicted as follows: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, ns = not statistically significant.

RESULTS

Aerrie Is Expressed in Endothelial Cells and Is Regulated by Aging

We identified lncRNA linc01013 in previously published RNA sequencing analysis of HUVECs exposed to laminar shear stress (20 dyn/cm^2 for 72 h) compared to static conditions (Doddaballapur et al., 2015). Linc01013, which we called Aerrie (Age and EndMT Regulated RNA In Endothelium), is regulated by shear stress (**Figure 1A**). Aerrie is downregulated under laminar flow while it is upregulated by oscillatory flow. KLF2 is a well-known transcriptional regulator that relays effects of laminar shear stress (**Supplementary Figure 1A**), accordingly we observed that Aerrie is regulated by KLF2 (**Supplementary Figure 1B**). Cellular senescence or aging is known to regulate KLF2. In this sense, KLF2 is downregulated by aging in endothelial cells (Boon et al., 2011; Warboys et al., 2014). On the other hand, P21 levels are described to be increased in senescent, aged or disturbed flow exposed endothelial cells (Katsuumi et al., 2018). Therefore, we were interested whether Aerrie is regulated by aging and other age-related diseases, such as atherosclerosis and ischemic heart disease (ISHD). To this end, HUVECs

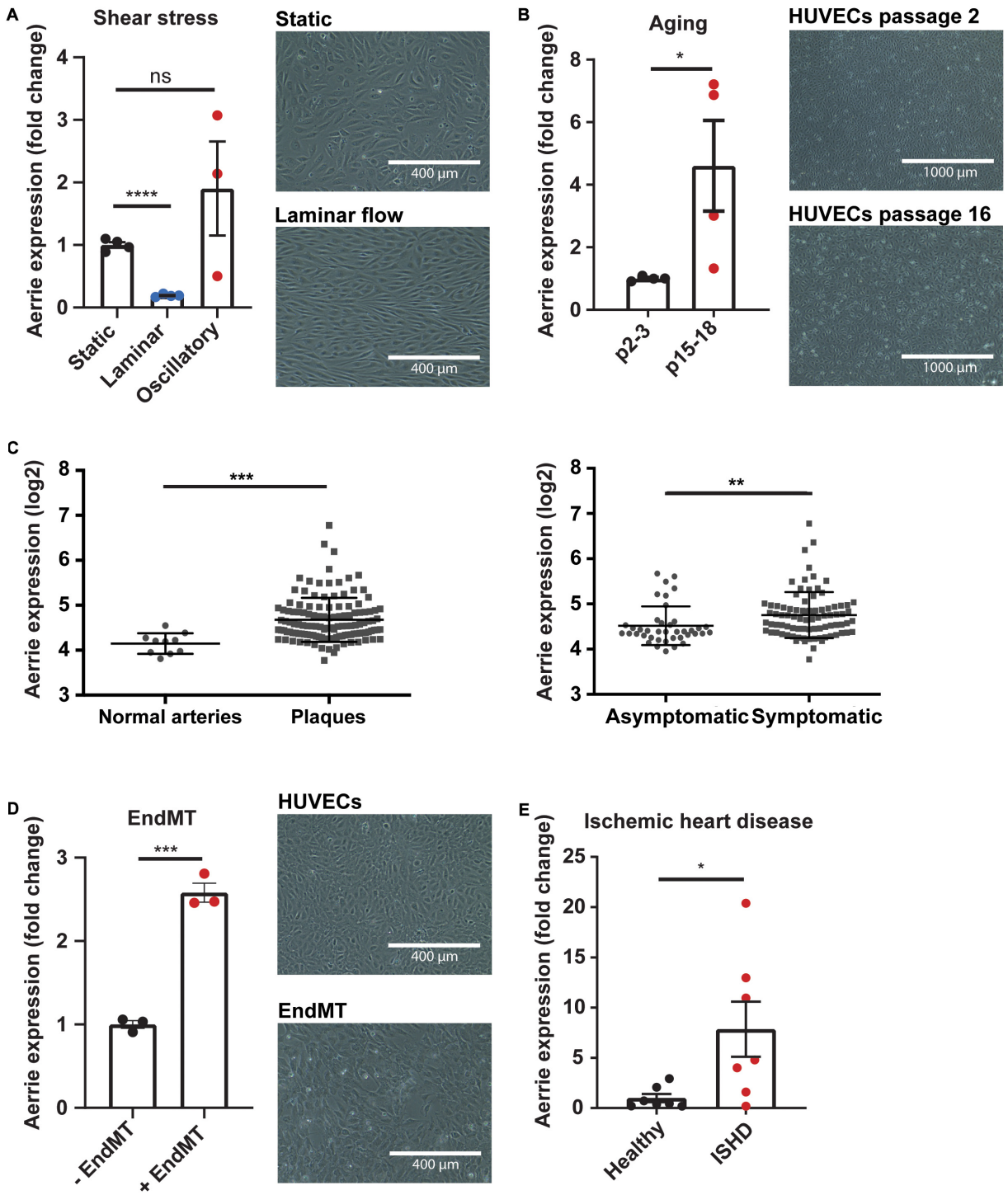

FIGURE 1 | The lncRNA Aerrie is expressed in endothelial cells and is regulated by aging, shear stress, and endothelial-to-mesenchymal transition (endMT). **(A)** Human umbilical vein endothelial cells (HUVECs) were exposed to laminar (72 h, 20 dyn/cm^2) and oscillatory flow (14 h, 20 dyn/cm^2). Expression levels of Aerrie were measured by real-time quantitative PCR (RT-qPCR). Expression values are relative to static condition and normalized to GAPDH mRNA ($n > 3$). **(B)** Human umbilical vein endothelial cells (HUVECs) were artificially aged by passing frequently until senescence is reached. Expression levels of Aerrie were measured by RT-qPCR in early and late passages (p) ($n = 4$). **(C)** Human atherosclerotic plaques were collected, and microarray profiled (GSE21545). Expression levels of Aerrie were plotted as normal healthy arteries ($n = 10$) versus atherosclerotic plaques ($n = 127$). From the atherosclerotic plaques, those from asymptomatic ($n = 40$) and symptomatic patients ($n = 87$) were distinguished and expression levels of Aerrie were plotted. **(D)** HUVECs were stimulated with IL-1β and TGF-β2 to induce endMT. Expression levels of Aerrie were measured by RT-qPCR ($n = 3$). **(E)** Human heart tissue of the left ventricle was isolated from healthy donor ($n = 7$) versus ischemic heart disease (ISHD) patients ($n = 7$). Expression levels of Aerrie were measured by RT-qPCR. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; *ns*, not statistically significant.

were artificially aged *in vitro* by serial passaging until cells reached senescence. These cells showed increased size, disturbed monolayer formation, and increased expression of Aerrie (**Figure 1B**). Additionally, we observed that these senescent or "aged" HUVECs have higher P21 levels, indicative of cellular aging (**Supplementary Figure 1C**). One of the most prominent age-associated vascular diseases, atherosclerosis, is known to occur at sites of flow disturbance in arteries (Wang and Bennett, 2012). We examined the expression levels of Aerrie in normal arteries, asymptomatic and symptomatic atherosclerotic plaques from the BiKE study (Naylor et al., 2003). This analysis showed that Aerrie is increased in atherosclerotic plaques. Additionally, the upregulated expression is also observed in plaques from symptomatic patients compared to asymptomatic atherosclerotic patients (**Figure 1C**). Another hallmark of vascular aging is endothelial-to-mesenchymal transition (endMT) (Fleenor et al., 2012). To induce endMT, HUVECs are stimulated with IL-1β and TGF-β2 cytokines for 72 h. This treatment resulted in morphological changes, such as increased cell size and elongation, loss of stable monolayer properties (**Figure 1D**), and increased expression of mesenchymal markers (**Supplementary Figure 1D**). Importantly, endMT induced the expression of Aerrie (**Figure 1D**). Ischemic heart disease (ISHD) is common in the older population and therefore of interest for this study (Benjamin et al., 2019). In line with the aging-induced expression of Aerrie, we observed increased expression of Aerrie in ISHD tissue from the left ventricle (**Figure 1E**). Taken together, these results identify Aerrie as an aging-induced lncRNA.

Loss of Aerrie Induces Endothelial Dysfunction

Aging impairs endothelial function such as proliferation and migration and is linked to cellular senescence (Heiss et al., 2005; Moriya and Minamino, 2017). To assess whether reduction of Aerrie contributes to endothelial cell dysfunction, several cellular assays were performed. Firstly, the localization of Aerrie was determined by nuclear and cytoplasmic fractionation. We used RPLP0 as a cytoplasmatic control and the lncRNA Malat1 as a nuclear control. We observed that lncRNA Aerrie is localized in both fractions (**Figure 2A**). We further confirmed the localization by SCRINSHOT RNA FISH (**Figure 2B**). To fully deplete both nuclear and cytoplasmic localized Aerrie, we used LNA-GapmeRs, after which Aerrie expression was reduced by 90% (**Figure 2B** and **Supplementary Figure 1E**). These oligonucleotides were designed according to our RNA-sequencing data from HUVECs (**Supplementary Figure 1F**). Depletion of Aerrie impaired HUVEC migration capabilities, as quantified in the scratch assay after 6 h (**Figure 2C**). This was further validated by performing a wound healing assay by electric cell-substrate impedance sensing (ECIS). Accordingly, loss of Aerrie resulted in slower barrier recovery of the endothelial monolayer after lethal electroporation (**Figure 2D**). Furthermore, ECIS was performed to assess monolayer integrity by measuring the resistance of the monolayer. Barrier function was significantly decreased upon Aerrie depletion (**Figure 2E**). Migration and intact cell-cell contacts are important for angiogenic sprouting. Therefore, we assessed the involvement of Aerrie in this process by performing a 3D endothelial cell sprouting assay. Silencing of Aerrie resulted in loss of cumulative sprout length and increased discontinuous sprout length (**Figures 2F,G**). This suggests that the loss of Aerrie may affect the ability of endothelial stalk cells to follow tip cells during angiogenic sprouting. Taken together, these results show that loss of Aerrie contributes to endothelial dysfunction.

Depletion of Aerrie Activates DNA Damage Response Mediated by YBX1(2/3)

To understand the molecular function of Aerrie in the endothelium, we aimed to identify protein interaction partners of Aerrie. Therefore, we performed an RNA pulldown of endogenous Aerrie followed by protein purification followed by mass spectrometry. Among the many binding proteins, Y-Box protein 1 (YBX1, also known as YB1) had the highest iBAQ value, a measure of absolute protein abundance (**Figure 3A**; Schwanhausser et al., 2011). The binding of YBX1 to Aerrie was validated by crosslinking RNA immunoprecipitation (CLIP) (**Supplementary Figure 2A**). YBX1 is a DNA/RNA binding protein that is involved in cell differentiation and embryonal development, stress response, and DNA repair in mammalian cells (Lu et al., 2005; Alidousty et al., 2014; Guarino et al., 2018). However, the exact mechanism of action of YBX1 remains elusive.

To further reveal the molecular function of Aerrie, we decided to profile activation of signaling pathways, using a phosphokinase assay (**Figure 3B** and **Supplementary Figure 2B**). A noticeable increase of phosphorylation was observed for P53 and CHK-2, indicating an increase in activation of the DNA damage signaling (**Figure 3B**). To confirm the activation of these pathways, the phosphorylation levels of the proteins involved in double- and single-strand break repair were determined. Loss of Aerrie showed induced phosphorylation of ATM at serine 1981, CHK-2 at tyrosine 68, H2AX at serine 139, and P53 at serine 15 indicating increased activation of the double-strand break repair (**Figure 3C**). At the same time increased phosphorylation of proteins involved in single-strand break repair was observed, such as ATR at serine 1989 and CHK1 at serine 317 (**Figure 3D**). To assess whether YBX1 plays a role in the activation of DNA damage signaling upon loss of Aerrie, we analyzed phosphorylation of DNA repair proteins after silencing YBX1 and its homologs YBX2 and YBX3. The rationale for this is that these homologs might compensate for loss of YBX1 and it is also described that YBX1 can regulate the expression of YBX3 (Lyabin et al., 2020). We indeed observed that a single knockdown of YBX1 is less robust than the knockdown of the entire family (**Supplementary Figure 2C**). The combination of silenced YBX2 or YBX3 with Aerrie does not result in a robust reduction of phosphorylated ATM compared to the total knockdown of the YBX family (**Supplementary Figure 2C, 3C**). Therefore, a knockdown of the entire YBX family was performed in our studies. Interestingly,

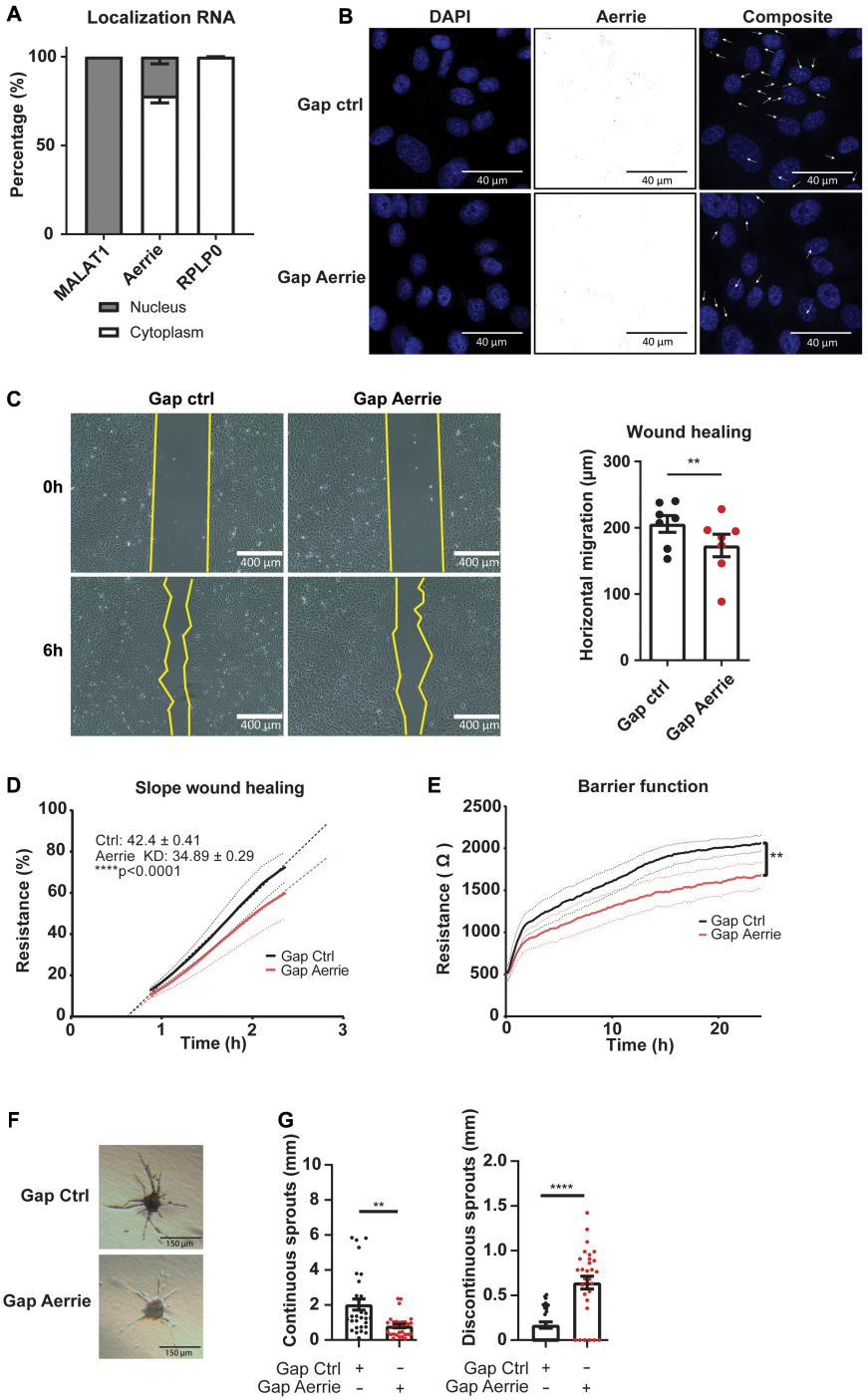

FIGURE 2 | The loss of Aerrie reduces migration, barrier function, and angiogenic sprouting *in vitro*. **(A)** HUVECs were fractionated into nucleus and cytoplasm. RT-qPCR was performed to measure lncRNA Aerrie localization. RPLP0 was used as cytoplasmatic control and lncRNA Malat1 as a nuclear control ($n = 3$). **(B–G)** HUVECs were treated with gapmeR (gap) targeting Aerrie or a respective control. **(B)** Subcellular localization of Aerrie in HUVECs was analyzed by SCRINSHOT RNA FISH. Nuclei were visualized with DAPI. Arrows indicate Aerrie localization. HUVECS were visualized with DAPI on the 405 nm channel and Aerrie on the 488 nm channel. **(C)** Scratch assay was performed with HUVECs for 6 h. The horizontal migration is calculated by the change in surface area ($n = 7$). **(D)** Electric cell-substrate impedance sensing (ECIS) was performed to measure migration of HUVECs. The endothelial monolayer was damaged through lethal electroporation. The recovery slope was determined between 1 to 2 h post wounding ($n = 3$). **(E)** ECIS was performed to measure the barrier function of the endothelial monolayer. The resistance of the endothelial monolayer was determined after 24 h ($n = 3$). **(F)** Brightfield images of silenced Aerrie sprouts and its control. **(G)** Cumulative sprout length was determined for continuous and discontinuous sprouts. ($n = 3$, at least 10 spheroids were measured per experiment). Quantification of the discontinuous sprouts measured by the distance from the tip cell to the stalk cell. $^{**}p < 0.01$; $^{****}p < 0.0001$; *ns*, not statistically significant.

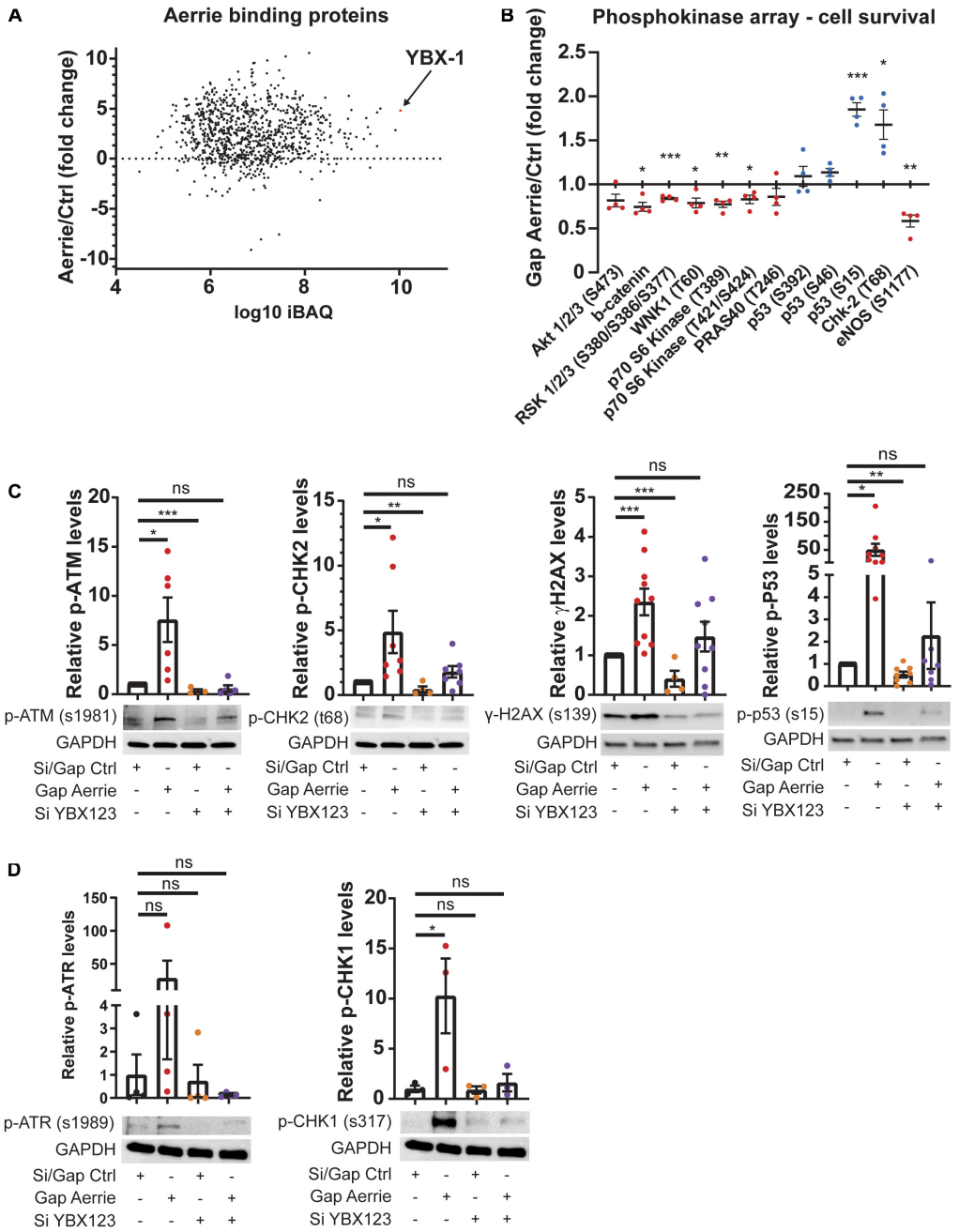

FIGURE 3 | Interaction of Aerrie with YBX1 is important for DNA damage signaling. **(A)** Aerrie-interacting proteins identified by mass spectrometry of RNA-antisense purifications in HUVECs. Identified proteins are plotted against respective IBAQ values. **(B–D)** HUVECs were treated with gapmeR (gap) targeting Aerrie, siRNA (si) targeting YBX1/2/3 or a respective control. **(B)** Phosphokinase levels of cell survival-related proteins ($n = 4$). Red dots are defined as the average level of phosphorylation after silencing knockdown is decreased compared to its respective control. Blue dots are defined as the average level of phosphorylation after silencing knockdown is increased compared to its respective control. **(C)** Phosphorylation levels of activated proteins in the double-strand break repair pathway were analyzed by western blotting. Phosphorylated ATM at serine 1981, CHK-2 at tyrosine 68, H2AX at serine 139, and P53 at serine 15 were analyzed upon depletion of Aerrie and YBX1 ($n > 4$). **(D)** Phosphorylation levels of activated proteins in the single-strand break repair pathway were analyzed by western blotting. Phosphorylated ATR at serine 1989, CHK-1 at serine 317 were analyzed upon depletion of Aerrie and YBX1 ($n = 3$). $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; *ns*, not statistically significant.

HUVECs without YBX1/2/3 showed no activation of the repair pathway proteins (**Figures 3C,D**). Moreover, when Aerrie and YBX1/2/3 were co-silenced, the activation of the repair pathway was abolished, indicating a crucial role of YBX1 in DNA repair signaling upon loss of Aerrie. YBX1 is described to be involved in DNA damage, however, its mechanism of action is not known (Kim et al., 2013; Alemasova et al., 2016). As mentioned above, we have shown that loss of Aerrie results

in impaired sprouting and that Aerrie is bound to YBX1. To assess whether YBX1/2/3 are involved in endothelial function, we analyzed angiogenic sprouting in endothelial cells with depletion of YBX1/2/3 and Aerrie (**Figure 4A**). The loss of YBX1/2/3 resulted in similar decrease of sprouting as the inhibition of Aerrie (**Figures 2F**, **4B**). However, discontinuous sprouts were significantly more present (**Figure 4B**, right). Co-silencing of Aerrie and YBX1/2/3 results in a reduced number of discontinuous sprouts compared to silencing YBX1/2/3 alone. This phenotype is comparable to knockdown of Aerrie alone, suggesting that YBX1 and Aerrie have a common function and that the phenotype can be explained due to inefficient DNA repair during angiogenic sprouting.

Aerrie Is Involved in DNA Damage and YBX1 Is Required for DNA Damage Signaling

The loss of Aerrie induces endothelial dysfunction and shows increased DNA damage signaling. Whether only DNA damage signaling is disturbed or whether DNA damage itself is affected is not known. To this end, we performed a comet assay. The comet assay is a well-established assay that can determine the relative amount of DNA damage of cells in all conditions (Olive and Banath, 2006). Doxorubicin is a chemotherapeutical drug that causes the DNA to remain open, damaged and unprotected, leading to apoptosis (Tacar et al., 2013). In line with increased DNA damage signaling upon Aerrie depletion (**Figures 3C,D**), we observed increased DNA damage upon loss of Aerrie in control and doxorubicin-treated cells (**Figure 5A**). However, whereas silencing YBX1 rescues DNA damage signaling induced by loss of Aerrie (**Figures 3C,D**), DNA damage itself is not affected by silencing YBX1. This indicates that loss of Aerrie induces DNA damage (and subsequent signaling), and that YBX1 is required only for subsequent signaling and is not involved in DNA damage induction after loss of Aerrie. Furthermore, DNA damage signaling in HUVECs treated with Doxorubicin is increased in all conditions (**Figures 5B,C**), as assessed by western blotting. Despite of the silencing of the YBX family, DNA damage signaling is still activated by doxorubicin. This effect can be explained by other DNA damage signaling activation that bypasses Aerrie and YBX1 (Paull, 2015). Taken together, the data suggests that knockdown of Aerrie induces DNA damage, and that DNA damage repair in turn requires Aerrie and YBX1 for efficient signaling to repair the DNA.

Lentiviral Overexpression of Aerrie Improves DNA Repair and Angiogenic Sprouting

Next, we determined whether Aerrie gain-of-function inhibits DNA damage and induces angiogenic sprouting. Therefore, we overexpressed Aerrie by lentiviral transduction. Without an exogenous DNA damage stimulus, Aerrie overexpression does not affect DNA damage, as measured by comet assay (**Figure 6A**). Inducing DNA damage using doxorubicin increases the tail moment by two-fold and, interestingly, is reduced upon overexpression of Aerrie (**Figure 6A**). Indeed, overexpression of Aerrie reduces DNA damage and suggests that DNA repair is enhanced. This effect is diminished when YBX1/2/3 is knocked down, implying that YBX1 and Aerrie together are necessary to have efficient DNA repair.

To identify whether impaired endothelial sprouting after silencing Aerrie is caused by defect DNA repair (**Figure 2F**), we stimulated endothelial spheroids overexpressing Aerrie with doxorubicin. Indeed, overexpression of Aerrie show improved angiogenic sprouting under doxorubicin stimulation compared to mock (**Figures 6B,C**). Unexpectedly, we also observed improved continuous sprouts when

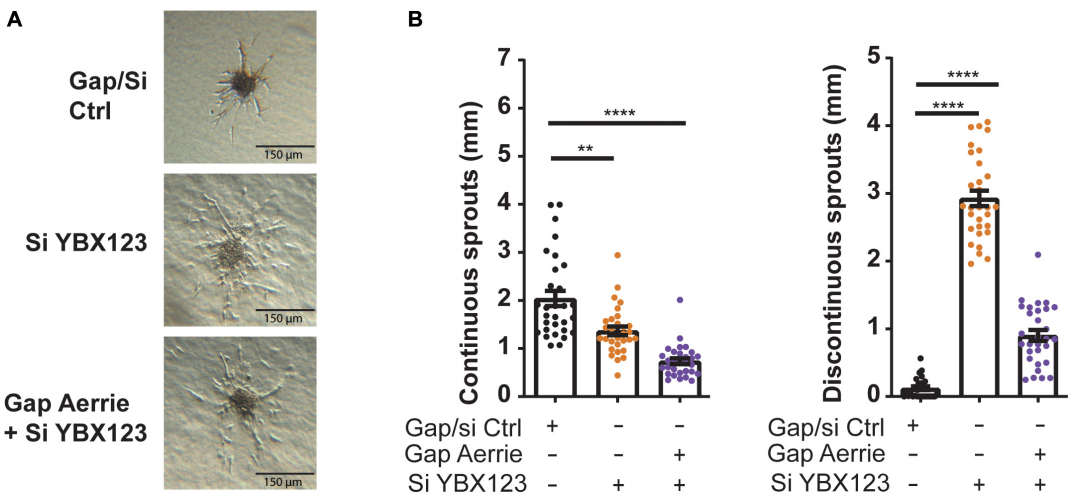

FIGURE 4 | Loss of Aerrie and YBX1/2/3 result in impaired sprouting. **(A,B)** HUVECs were treated with gapmeR (gap) targeting Aerrie, siRNA (si) targeting YBX1/2/3 or a respective control. **(A)** Sprouting assay **(B)** Quantification of the continuous and discontinuous sprouts per condition. ($n = 3$, at least 10 spheroids per experiment). Sprouts measured by the distance from the tip cell to the stalk cell. $^{**}p < 0.01$; $^{****}p < 0.0001$; *ns*, not statistically significant.

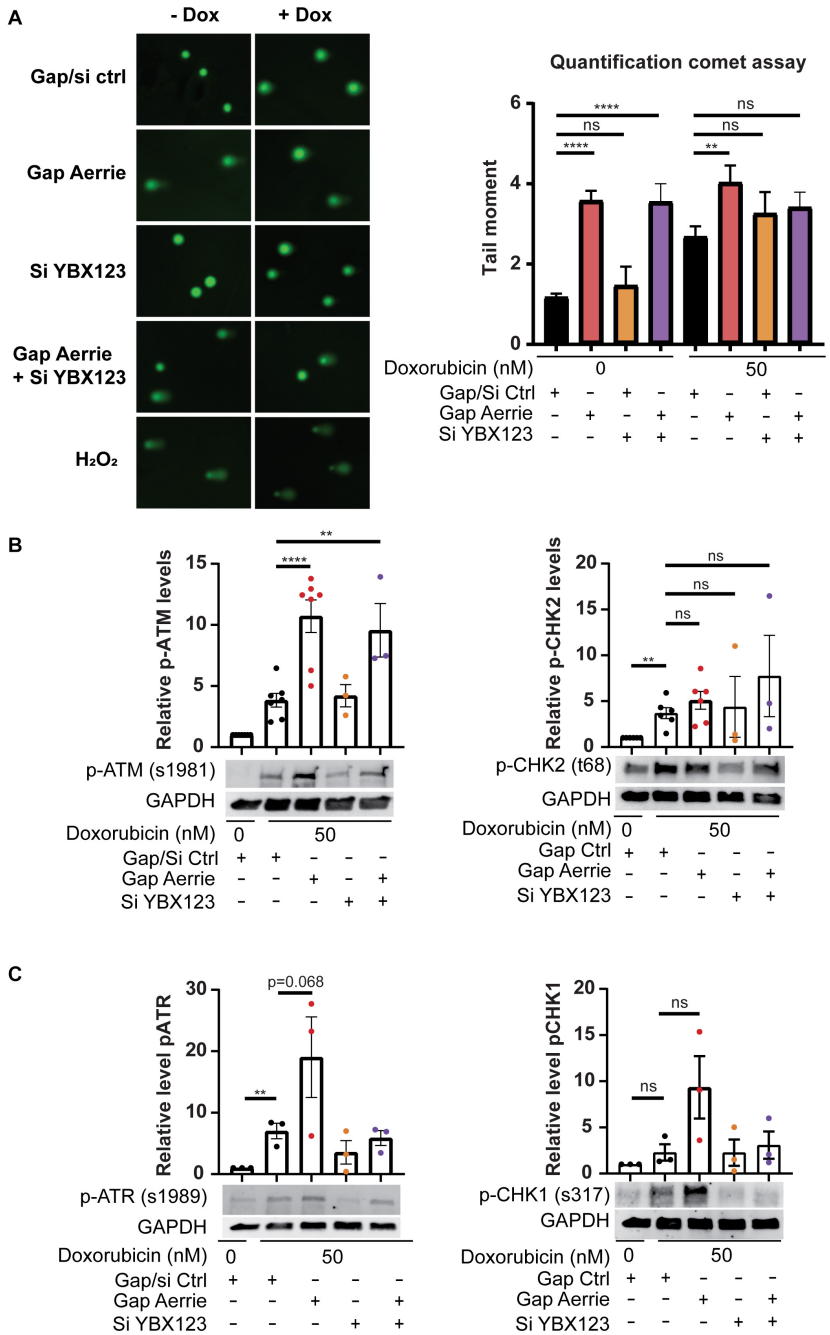

FIGURE 5 | Loss of Aerrie causes DNA damage and activation of signaling, while loss of YBX1/2/3 results in impaired DNA damage signaling. **(A–C)** HUVECs were treated with gapmeR (gap) targeting Aerrie, siRNA (si) targeting YBX1/2/3 or a respective control. **(A)** DNA damage was quantified by comet assay. HUVECs were stimulated with doxorubicin (Dox, 50 nM), embedded in low melting agarose, and stained with SYBR gold for imaging by fluorescence microscopy. Comets were quantified by CometScore ($n > 3$, at least 50 comets per experiment). **(B)** Phosphorylation levels of activated proteins in the double-strand break repair pathway were analyzed by western blotting. Phosphorylated ATM at serine 1981 and phosphorylated CHK-2 at tyrosine 68 were analyzed of Aerrie and YBX1 depleted HUVECs ($n > 3$). **(C)** Phosphorylation levels of activated proteins in the single-strand break repair pathway were analyzed by western blotting. Phosphorylated ATR at serine 1981 and CHK-1 at serine 317 were analyzed of Aerrie and YBX1 depleted HUVECs ($n = 3$). $^{**}p < 0.01$; $^{****}p < 0.0001$; *ns*, not statistically significant.

YBX1/2/3 is knocked down in combination with Aerrie overexpression. However, we did not observe a reduction of discontinuous sprouts in the same condition, suggesting that Aerrie overexpression is unable to rescue induction of discontinuous sprouts in the absence of YBX1/2/3. The increase of discontinuous sprouts by loss of YBX1/2/3 is the same effect we have observed in the previous experiment (**Figures 4A,B**). A possible explanation for this phenotype is a defect in proliferation on the long term (Kotake et al., 2017).

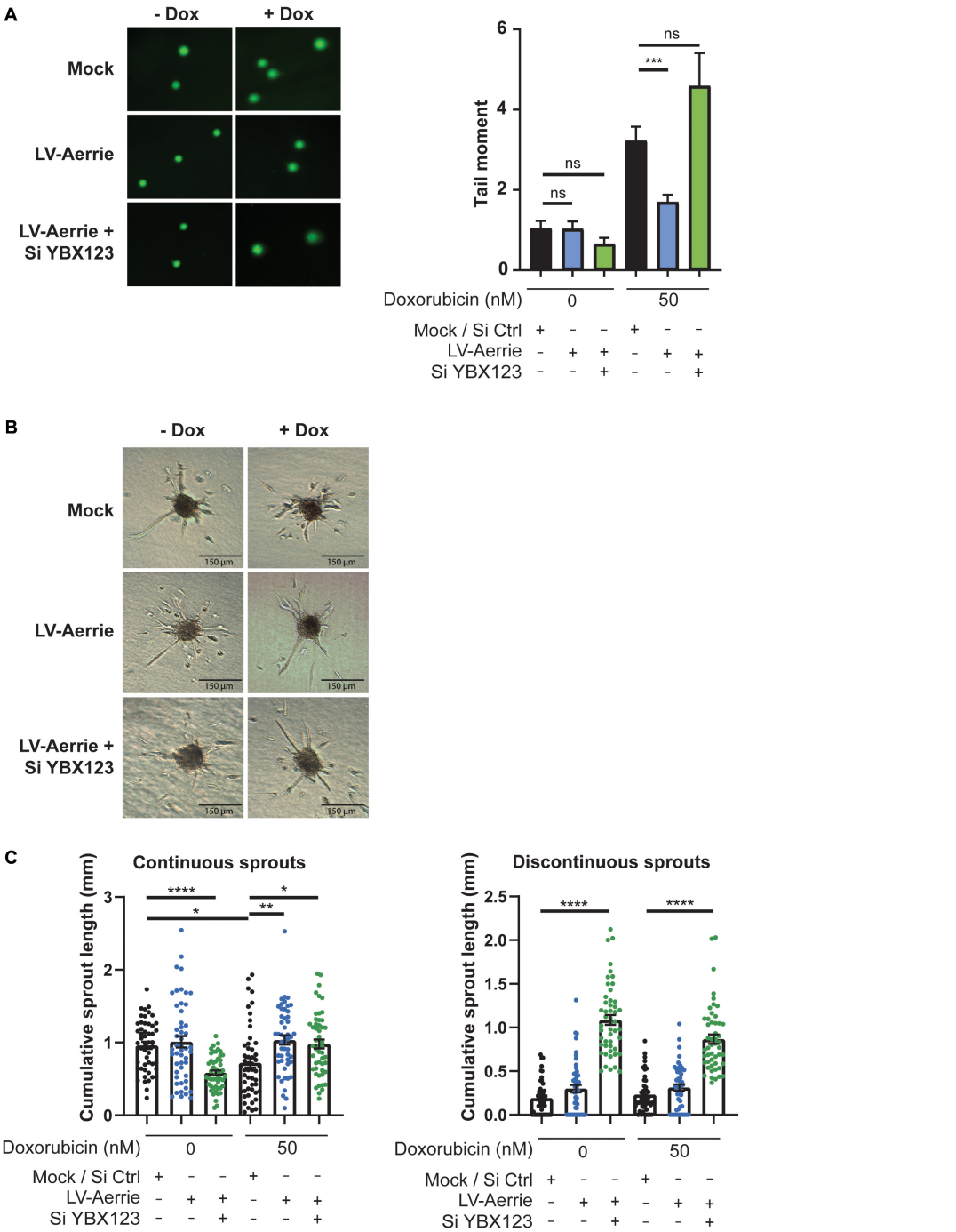

FIGURE 6 | Overexpression of Aerrie improves DNA repair and angiogenic sprouting and is impaired by loss of YBX1/2/3. **(A–C)** HUVECs were treated with lentivirus (LV) for Aerrie overexpression and siRNA (si) targeting YBX1/2/3 or a respective control. **(A)** HUVECs with depleted Aerrie or YBX1 are stimulated with Doxorubicin (50 nM) and embedded in low melting agarose stained with SYBR gold and imaged by fluorescence microscopy. Comets were quantified by CometScore. ($n > 3$, at least 50 comets per experiment) **(B)** Sprouting assay performed with or without stimulation of Doxorubicin (50 nM). **(C)** Quantification of the continuous and discontinuous sprouts per condition. Measured by the distance from the tip cell to the stalk cell. ($n = 3$, ≥10 spheroids per experiment). *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; *ns*, not statistically significant.

In summary, we show that the novel lncRNA Aerrie is required for normal endothelial function. Aerrie is upregulated by aging which may have implications for CVD. We show that depletion of Aerrie causes detrimental effects that results in impaired sprouting, migration, and barrier function. HUVECs overexpressing Aerrie are less prone to DNA damage suggesting

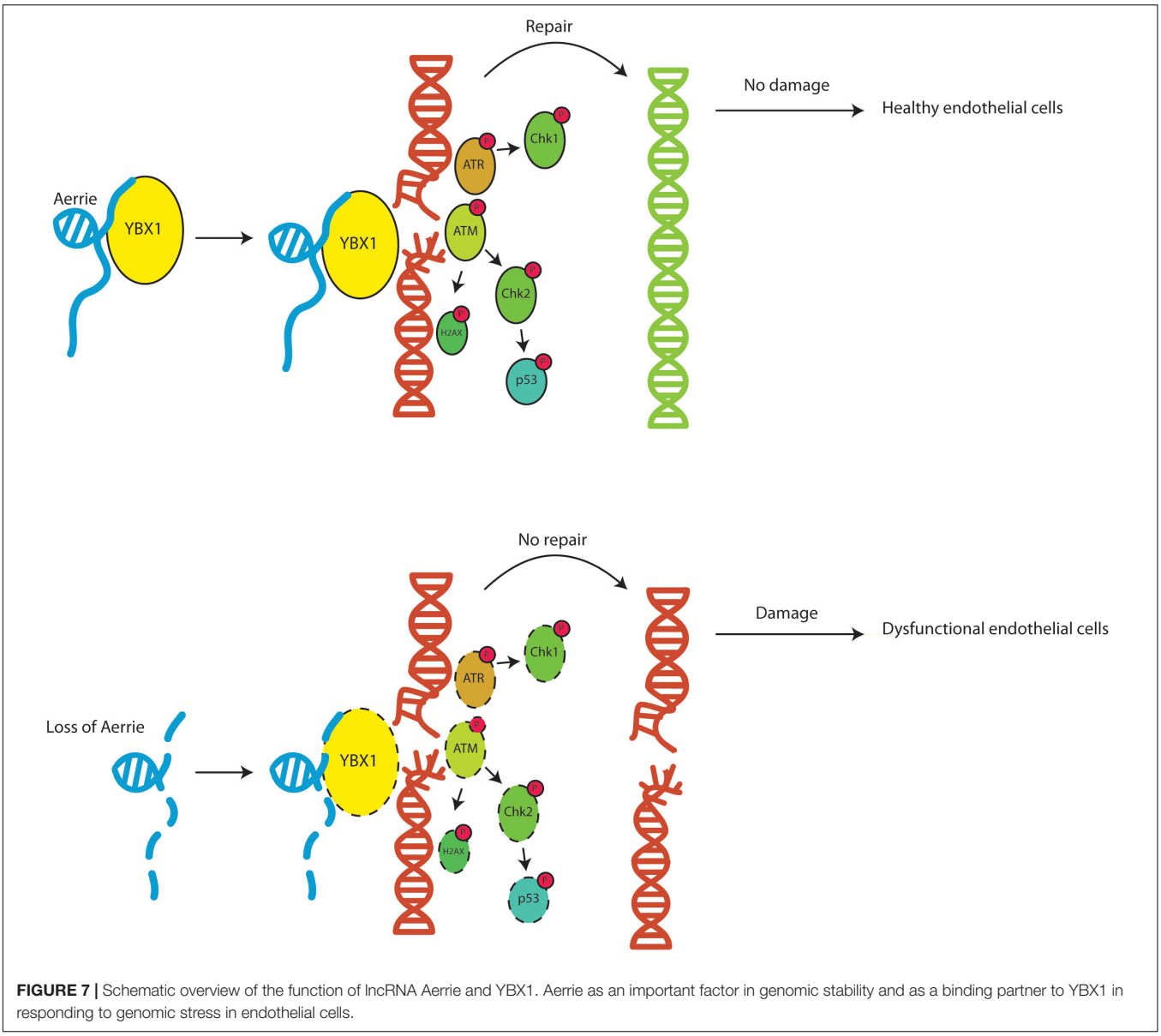

FIGURE 7 | Schematic overview of the function of lncRNA Aerrie and YBX1. Aerrie as an important factor in genomic stability and as a binding partner to YBX1 in responding to genomic stress in endothelial cells.

more efficient DNA repair. Additionally, Aerrie interacts with YBX1 and upon depletion of YBX1 and Aerrie, DNA damage signaling is impaired and increased DNA damage is observed. The interaction of Aerrie and YBX1 is needed for efficient DNA repair to maintain proper angiogenic sprouting. In conclusion, lncRNA Aerrie and YBX1 are important factors in DNA damage repair in endothelial cells that may be needed to maintain cardiovascular homeostasis (**Figure 7**).

DISCUSSION

Our results demonstrate that the novel lncRNA Aerrie is an essential regulator of endothelial DNA repair and that Aerrie expression is increased during aging. Loss of Aerrie is accompanied by impaired angiogenic sprouting, migration, and endothelial barrier function. Overexpression of Aerrie improved DNA repair efficiency when the genome is damaged by doxorubicin. YBX1 is an important binding partner of Aerrie and it is essential to activate signaling of key components of the DNA repair pathway (**Figure 7**).

Aerrie expression is increased with aging, disturbed flow, and endMT *in vitro* (**Figures 1A,B,D**). Aerrie is expressed in arteries and the heart and is increased in atherosclerotic plaques and in ischemic hearts (**Figures 1C,E**). However, the regulation of Aerrie by aging remains elusive. KLF2 is known to be regulated by shear stress and acts as an atheroprotective factor (Fledderus et al., 2007). Mechanistically, we observe that KLF2 regulates Aerrie expression (**Supplementary Figure 1B**). Overexpression of KLF2 decreases expression of Aerrie. Accordingly, we observe repressed expression of Aerrie *in vitro* under laminar flow (**Figure 1A**) when KLF2 expression is increased (**Supplementary**

Figure 1A), while Aerrie expression is increased under disturbed flow when KLF2 is decreased (**Figure 1A** and **Supplementary Figure 1A**). Future studies will address the potential role of KLF2 in controlling Aerrie expression during aging. Another mechanism that regulates Aerrie is endMT. We identify JMJD2B, a recently reported factor involved in endMT (Glaser et al., 2020), as a potential regulator of Aerrie expression (**Supplementary Figure 2D**), but whether JMJD2B is involved in endothelial aging remains unknown.

Aerrie or LINC01013 has already been studied in Anaplastic Large-Cell Lymphoma (ALCL). Chung et al. (2017) describes Aerrie as an invasion activator of epithelial-to-mesenchymal transition (EMT). EMT is a physiological process for embryonic development, but is also found to play a role in fibrosis and cancer (Stone et al., 2016). EndMT is described as an analogous process in endothelial cells sharing similar cellular and molecular events (Medici and Kalluri, 2012). In both processes, the cells lose their adherens junctions and increase mesenchymal markers to transition into a mesenchymal phenotype. Chung et al. has shown that Aerrie regulates Snail to promote ALCL invasion. However, we observed that Snail is not regulated by Aerrie in HUVECs (**Supplementary Figure 2E**). Suggesting that Aerrie have different functions in other cell types.

Aging is a complex process that has multiple effects on the endothelium. One of the hallmarks of aging is increased DNA damage (Donato et al., 2015). Our data show that loss of Aerrie results in DNA damage and markers such as ATM, CHK2, γH2AX, P53, ATR, CHK1 are activated. Interestingly, overexpression of Aerrie enhances the DNA repair capabilities and results in less DNA damage when stimulated with doxorubicin (**Figure 6A**). The model that emerges from these data is that an increase in Aerrie expression with aging likely contributes to protection from endothelial dysfunction, but this rescues mechanism is not sufficient to prevent aging-induced endothelial dysfunction.

We showed that Aerrie controls DNA repair *via* YBX1 interaction. Loss of YBX1 reduces efficient DNA repair even in the presence of Aerrie (**Figure 6A**). YBX1 has a broad function in various cell types and is recognized as an oncogenic factor in several tumors (Lim et al., 2019; Kim et al., 2020). YBX1 participates in post-transcriptional regulation, stress response signaling, inflammation, and DNA damage (Guarino et al., 2018; Wang et al., 2019; Hermert et al., 2020). YBX1 can bind to the DNA close to the damage site with its cold shock domain (Zhang et al., 2020). Together with Aerrie, this may be necessary to recruit DNA damage repair proteins to ensure proper DNA damage repair, such as ATM and ATR. It is unclear whether and how the Aerrie-YBX1 complex recruits the DNA repair proteins. However, reduction of YBX1 was sufficient to abolish the activation of ATM, CHK2, P53, γH2AX, ATR, and CHK1 in order to repair DNA damage (**Figures 3B,C**, **5B,C**). A possible mechanism would be that upon loss of Aerrie, YBX1 cannot be translocated to the nucleus (Koike et al., 1997; Das et al., 2007). Without the translocation, DNA repair proteins cannot be recruited or activated by Aerrie and YBX1. Slightly higher expression of YBX1 is observed upon knockdown of Aerrie, although the localization is not known (**Supplementary Figure 2F**). For example, it is described that the lncRNA TP53TG1 blocks the translocation of YBX1 to the nucleus to repair the DNA (Diaz-Lagares et al., 2016). The study provides evidence that epigenetic silencing of TP53TG1 results in increased DNA damage resistance by YBX1 and the activation of the PI3K/AKT pathway.

Mechanistically, YBX1 is known to be phosphorylated by CHK1 at threonine 80 (Blasius et al., 2011). We have detected an increase in phosphorylated CHK1 and other DNA damage markers with loss of Aerrie. A possible explanation is that CHK1 regulates YBX1 and thereby regulates DNA repair. Dissociation of YBX1 from the DNA repair mechanism can be achieved by ubiquitination. It has been described that YBX1 can be ubiquitinated by RBBP6 and its RING finger domain (Chibi et al., 2008). Thus, dissociation of YBX1 from the DNA repair machinery may be achieved by ubiquitination. LncRNA Aerrie and YBX1 may be involved in a more complex and versatile system than we have explored. A possible mechanism could be that YBX1 and Aerrie are being recruited by DNA damage sensing proteins such as ATM to enhance the DNA damage signaling. In these events, ATM is capable to modulate the function of CHK1 and CHK2 to slow down the cell cycle (Awasthi et al., 2015). CHK1 subsequently activates YBX1 to complement the DNA damage signaling to repair the DNA. Upon knockdown of YBX1, the DNA repair signaling would not be activated when DNA damage is present, which fits to our data.

Surprisingly, we discovered that overexpression of Aerrie and the loss of YBX1/2/3 results in rescue of continuous sprouts under stimulation of doxorubicin (**Figures 6B,C**). No rescue in the sprouting assay was expected due to the increase of DNA damage in the comet assay (**Figure 6A**). It is possible that the reduction of DNA damage signaling by loss of YBX1 and the inhibition of DNA damage by overexpressed Aerrie the cells can survive apoptosis for a short period of time. In this sense, the cells can undergo angiogenic sprouting and proliferation until the threshold of accumulated DNA damage is reached. Also, the amount of error prone repair can reach a limit until the cells undergo apoptosis (Khanna, 2015). With the same thought, we could argue that the rescue of continuous sprouts that we observe (**Figures 6B,C**) is due to accumulated amount of error-repaired DNA that results in cancerous cell growth.

In conclusion, this study provides evidence that lncRNA Aerrie is required for proper DNA damage signaling, via interaction YBX1, and for efficient DNA repair in endothelial cells.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE21545.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Stockholm Regional Ethics Committee.

The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TP designed and performed the experiments, analyzed the data, and drafted the manuscript. DB provided technical, conceptual advice, and performed the experiments. LS provided technical, conceptual advice, and provided samples. AB, EL, and YT performed the experiments. LM and UH provided data. IW performed the mass spectrometry measurements and analysis. SD provided data and gave conceptual advice. RB supervised the project, designed experiments, analyzed data, handled funding, and drafted the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the German Centre for Cardiovascular Research (DZHK), the Cardiopulmonary Institute (CPI), the Deutsche Forschungsgemeinschaft (SFB834 to RB and SD, SFB815 to IW, TRR267 to IW, SD, and RB), the European Research Council ("NOVA" to RB and "Angiolnc" to SD), the Netherlands Organisation for Scientific Research (NWO Vidi to RB), the European Union (Horizon 2020 grant no. 825670 to RB).

ACKNOWLEDGMENTS

We thank Chris Dos Remedios and the Sydney Heart Bank for the samples used in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2020.619079/full#supplementary-material

Supplementary Figure 1 | (A) KLF2 mRNA expression in HUVECs stimulated under static, laminar (20 dyn/cm^2 for 72 h), or oscillatory flow (20 dyn/cm^2 for 14 h). Expression levels of Aerrie were measured by real-time quantitative PCR (RT-qPCR). Expression values are relative to static condition and normalized to GAPDH mRNA ($n > 3$). **(B)** RNA expression levels of Aerrie and KLF2 measured in lentiviral KLF2 induced HUVECs by RT-qPCR. Expression values are normalized to RPLP0 RNA. **(C)** P21 RNA expression levels in early vs. late passaged HUVECs by RT-qPCR. Expression values are normalized to RPLP0 RNA. **(D)** EndMT markers SM22, snail, CNN1 expression levels of HUVECs and endMT induced HUVECs measured by RT-qPCR. Expression values are normalized to RPLP0 RNA. **(E)** Expression level of Aerrie after knockdown with LNA-GapmeR measured by RT-qPCR. Expression values are normalized to RPLP0 RNA. **(F)** RNA-sequencing of HUVECs at Aerrie region. The vertical lines indicate reads and the arches above and below indicate splice junctions. 4 exons of 2 known validated transcripts of Aerrie annotated in NCBI show reads, namely NR_038981.1, and NR146223.1.

Supplementary Figure 2 | (A) RT-qPCR of Aerrie from the elution of RNA crosslinking immunoprecipitation (CLIP) experiments of YBX1. **(B)** Quantified phosphorylation levels of proteins from the phosphokinase assay. **(C)** Phosphorylation levels of activated ATM after separate YBX1, YBX2, and YBX3 knockdowns analyzed by western blotting. **(D)** EndMT was induced by IL-1β and TGF-β2. Aerrie expression level were measured by RT-qPCR after JMJD2B knockdown in endMT stimulated HUVECs. Expression values are normalized to RPLP0 RNA. **(E)** Snail RNA expression levels in HUVECs by RT-qPCR. HUVECs were treated gapmeR targeting Aerrie (gap) or a respective control. Expression values are normalized to RPLP0 RNA. **(F)** YBX1 levels measured upon knockdown of Aerrie by western blotting ($n > 3$).

Supplementary Table 1 | ISHD patient IDs with their corresponding information. Left ventricular tissues from ISHD samples were acquired from the University of Sydney (Sydney, NSW, Australia), with the ethical approval of the Human Research Ethics Committee (number 2012/2814). Explanted left ventricular heart tissue of healthy donors was used as control samples; the donors died from a non-cardiac cause, typically motor vehicle accidents. The healthy donor samples were also acquired from the University of Sydney.

Supplementary Table 2 | Gene silencing, qPCR primer and RNA pulldown oligo sequences. (ST3) Antibody list used for Western blots and RNA IPs.

REFERENCES

Alemasova, E. E., Moor, N. A., Naumenko, K. N., Kutuzov, M. M., Sukhanova, M. V., Pestryakov, P. E., et al. (2016). Y-box-binding protein 1 as a non-canonical factor of base excision repair. *Biochim. Biophys. Acta* 1864, 1631–1640. doi: 10.1016/j.bbapap.2016.08.012

Alidousty, C., Rauen, T., Hanssen, L., Wang, Q., Alampour-Rajabi, S., Mertens, P. R., et al. (2014). Calcineurin-mediated YB-1 dephosphorylation regulates CCL5 expression during monocyte differentiation. *J. Biol. Chem.* 289, 21401–21412. doi: 10.1074/jbc.m114.562991

Awasthi, P., Foiani, M., and Kumar, A. (2015). ATM and ATR signaling at a glance. *J. Cell Sci.* 128, 4255–4262. doi: 10.1242/jcs.169730

Bakkenist, C. J., and Kastan, M. B. (2003). DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature* 421, 499–506. doi: 10.1038/nature01368

Benjamin, E. J., Muntner, P., Alonso, A., Bittencourt, M. S., Callaway, C. W., Carson, A. P., et al. (2019). Heart disease and stroke statistics-2019 update: a report from the American Heart Association. *Circulation* 139, e56–e66.

Blackford, A. N., and Jackson, S. P. (2017). ATM, ATR, and DNA-PK: the trinity at the heart of the DNA damage response. *Mol. Cell* 66, 801–817. doi: 10.1016/j.molcel.2017.05.015

Blasius, M., Forment, J. V., Thakkar, N., Wagner, S. A., Choudhary, C., and Jackson, S. P. (2011). A phospho-proteomic screen identifies substrates of the checkpoint kinase Chk1. *Genome Biol.* 12:R78.

Boon, R. A., Hofmann, P., Michalik, K. M., Lozano-Vidal, N., Berghauser, D., Fischer, A., et al. (2016). Long noncoding RNA Meg3 controls endothelial cell aging and function: implications for regenerative angiogenesis. *J. Am. Coll. Cardiol.* 68, 2589–2591. doi: 10.1016/j.jacc.2016.09.949

Boon, R. A., Urbich, C., Fischer, A., Fontijn, R. D., Seeger, F. H., Koyanagi, M., et al. (2011). Kruppel-like factor 2 improves neovascularization capacity of aged proangiogenic cells. *Eur. Heart J.* 32, 371–377. doi: 10.1093/eurheartj/ehq137

Celeste, A., Petersen, S., Romanienko, P. J., Fernandez-Capetillo, O., Chen, H. T., Sedelnikova, O. A., et al. (2002). Genomic instability in mice lacking histone H2AX. *Science* 296, 922–927. doi: 10.1126/science.1069398

Chehab, N. H., Malikzay, A., Appel, M., and Halazonetis, T. D. (2000). Chk2/hCds1 functions as a DNA damage checkpoint in G(1) by stabilizing p53. *Genes Dev.* 14, 278–288.

Chibi, M., Meyer, M., Skepu, A., Dj, G. R., Moolman-Smook, J. C., and Pugh, D. J. (2008). RBBP6 interacts with multifunctional protein YB-1 through its RING finger domain, leading to ubiquitination and proteosomal degradation of YB-1. *J. Mol. Biol.* 384, 908–916. doi: 10.1016/j.jmb.2008.09.060

Childs, B. G., Durik, M., Baker, D. J., and van Deursen, J. M. (2015). Cellular senescence in aging and age-related disease: from mechanisms to therapy. *Nat. Med.* 21, 1424–1435. doi: 10.1038/nm.4000

Chung, I. H., Lu, P. H., Lin, Y. H., Tsai, M. M., Lin, Y. W., Yeh, C. T., et al. (2017). The long non-coding RNA LINC01013 enhances invasion of human anaplastic large-cell lymphoma. *Sci. Rep.* 7:295.

Collins, C., and Tzima, E. (2011). Hemodynamic forces in endothelial dysfunction and vascular aging. *Exp. Gerontol.* 46, 185–188. doi: 10.1016/j.exger.2010.09.010

Das, S., Chattopadhyay, R., Bhakat, K. K., Boldogh, I., Kohno, K., Prasad, R., et al. (2007). Stimulation of NEIL2-mediated oxidized base excision repair via YB-1 interaction during oxidative stress. *J. Biol. Chem.* 282, 28474–28484. doi: 10.1074/jbc.m704672200

Davis, A. J., Chen, B. P., and Chen, D. J. (2014). DNA-PK: a dynamic enzyme in a versatile DSB repair pathway. *DNA Repair (Amst)* 17, 21–29. doi: 10.1016/j.dnarep.2014.02.020

Diaz-Lagares, A., Crujeiras, A. B., Lopez-Serra, P., Soler, M., Setien, F., Goyal, A., et al. (2016). Epigenetic inactivation of the p53-induced long noncoding RNA TP53 target 1 in human cancer. *Proc. Natl. Acad. Sci. U.S.A.* 113, E7535–E7544.

Doddaballapur, A., Michalik, K. M., Manavski, Y., Lucas, T., Houtkooper, R. H., You, X., et al. (2015). Laminar shear stress inhibits endothelial cell metabolism via KLF2-mediated repression of PFKFB3. *Arterioscler. Thromb. Vasc. Biol.* 35, 137–145. doi: 10.1161/atvbaha.114.304277

Donato, A. J., Morgan, R. G., Walker, A. E., and Lesniewski, L. A. (2015). Cellular and molecular biology of aging endothelial cells. *J. Mol. Cell. Cardiol.* 89, 122–135. doi: 10.1016/j.yjmcc.2015.01.021

Fledderus, J. O., van Thienen, J. V., Boon, R. A., Dekker, R. J., Rohlena, J., Volger, O. L., et al. (2007). Prolonged shear stress and KLF2 suppress constitutive proinflammatory transcription through inhibition of ATF2. *Blood* 109, 4249–4257. doi: 10.1182/blood-2006-07-036020

Fleenor, B. S., Marshall, K. D., Rippe, C., and Seals, D. R. (2012). Replicative aging induces endothelial to mesenchymal transition in human aortic endothelial cells: potential role of inflammation. *J. Vasc. Res.* 49, 59–64. doi: 10.1159/000329681

Gagnon, K. T., Li, L., Chu, Y., Janowski, B. A., and Corey, D. R. (2014). RNAi factors are present and active in human cell nuclei. *Cell Rep.* 6, 211–221. doi: 10.1016/j.celrep.2013.12.013

Glaser, S. F., Heumuller, A. W., Tombor, L., Hofmann, P., Muhly-Reinholz, M., Fischer, A., et al. (2020). The histone demethylase JMJD2B regulates endothelial-to-mesenchymal transition. *Proc. Natl. Acad. Sci. U.S.A.* 117, 4180–4187. doi: 10.1073/pnas.1913481117

Guarino, A. M., Troiano, A., Pizzo, E., Bosso, A., Vivo, M., Pinto, G., et al. (2018). Oxidative stress causes enhanced secretion of YB-1 protein that restrains proliferation of receiving cells. *Genes (Basel)* 9:513. doi: 10.3390/genes9100513

Guttman, M., Amit, I., Garber, M., French, C., Lin, M. F., Feldser, D., et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 458, 223–227. doi: 10.1038/nature07672

Haemmig, S., Yang, D., Sun, X., Das, D., Ghaffari, S., Molinaro, R., et al. (2020). Long noncoding RNA SNHG12 integrates a DNA-PK-mediated DNA damage response and vascular senescence. *Sci. Transl. Med.* 12:eaaw1868. doi: 10.1126/scitranslmed.aaw1868

Halliday, A., Harrison, M., Hayter, E., Kong, X., Mansfield, A., Marro, J., et al. (2010). 10-year stroke prevention after successful carotid endarterectomy for asymptomatic stenosis (ACST-1): a multicentre randomised trial. *Lancet* 376, 1074–1084. doi: 10.1016/s0140-6736(10)61197-x

Heiss, C., Keymel, S., Niesler, U., Ziemann, J., Kelm, M., and Kalka, C. (2005). Impaired progenitor cell activity in age-related endothelial dysfunction. *J. Am. Coll. Cardiol.* 45, 1441–1448. doi: 10.1016/j.jacc.2004.12.074

Hermert, D., Martin, I. V., Reiss, L. K., Liu, X., Breitkopf, D. M., Reimer, K. C., et al. (2020). The nucleic acid binding protein YB-1-controlled expression of CXCL-1 modulates kidney damage in liver fibrosis. *Kidney Int.* 97, 741–752. doi: 10.1016/j.kint.2019.10.024

Kang, L. S., Reyes, R. A., and Muller-Delp, J. M. (2009). Aging impairs flow-induced dilation in coronary arterioles: role of NO and H(2)O(2). *Am. J. Physiol. Heart Circ. Physiol.* 297, H1087–H1095.

Katsuumi, G., Shimizu, I., Yoshida, Y., and Minamino, T. (2018). Vascular senescence in cardiovascular and metabolic diseases. *Front. Cardiovasc. Med.* 5:18. doi: 10.3389/fcvm.2018.00018

Khanna, A. (2015). DNA damage in cancer therapeutics: a boon or a curse? *Cancer Res.* 75, 2133–2138. doi: 10.1158/0008-5472.can-14-3247

Kim, A., Shim, S., Kim, Y. H., Kim, M. J., Park, S., and Myung, J. K. (2020). Inhibition of Y box binding protein 1 suppresses cell growth and motility in colorectal cancer. *Mol. Cancer Ther.* 19, 479–489. doi: 10.1158/1535-7163.mct-19-0265

Kim, E. R., Selyutina, A. A., Buldakov, I. A., Evdokimova, V., Ovchinnikov, L. P., and Sorokin, A. V. (2013). The proteolytic YB-1 fragment interacts with DNA repair machinery and enhances survival during DNA damaging stress. *Cell Cycle* 12, 3791–3803. doi: 10.4161/cc.26670

Koike, K., Uchiumi, T., Ohga, T., Toh, S., Wada, M., Kohno, K., et al. (1997). Nuclear translocation of the Y-box binding protein by ultraviolet irradiation. *FEBS Lett.* 417, 390–394. doi: 10.1016/s0014-5793(97)01296-9

Kotake, Y., Arikawa, N., Tahara, K., Maru, H., and Naemura, M. (2017). Y-box binding protein 1 is involved in regulating the G2/M phase of the cell cycle. *Anticancer Res.* 37, 1603–1608. doi: 10.21873/anticanres.11490

Lakatta, E. G., and Levy, D. (2003a). Arterial and cardiac aging: major shareholders in cardiovascular disease enterprises: part I: aging arteries: a "set up" for vascular disease. *Circulation* 107, 139–146. doi: 10.1161/01.cir.0000048892.83521.58

Lakatta, E. G., and Levy, D. (2003b). Arterial and cardiac aging: major shareholders in cardiovascular disease enterprises: part II: the aging heart in health: links to heart disease. *Circulation* 107, 346–354. doi: 10.1161/01.cir.0000048893.62841.f7

Li, T., Kon, N., Jiang, L., Tan, M., Ludwig, T., Zhao, Y., et al. (2012). Tumor suppression in the absence of p53-mediated cell-cycle arrest, apoptosis, and senescence. *Cell* 149, 1269–1283. doi: 10.1016/j.cell.2012.04.026

Lim, J. P., Nair, S., Shyamasundar, S., Chua, P. J., Muniasamy, U., Matsumoto, K., et al. (2019). Silencing Y-box binding protein-1 inhibits triple-negative breast cancer cell invasiveness via regulation of MMP1 and beta-catenin expression. *Cancer Lett.* 452, 119–131. doi: 10.1016/j.canlet.2019.03.014

Lu, Z. H., Books, J. T., and Ley, T. J. (2005). YB-1 is important for late-stage embryonic development, optimal cellular stress responses, and the prevention of premature senescence. *Mol. Cell. Biol.* 25, 4625–4637. doi: 10.1128/mcb.25.11.4625-4637.2005

Lyabin, D. N., Eliseeva, I. A., Smolin, E. A., Doronin, A. N., Budkina, K. S., Kulakovskiy, I. V., et al. (2020). YB-3 substitutes YB-1 in global mRNA binding. *RNA Biol.* 17, 487–499. doi: 10.1080/15476286.2019.1710050

Marechal, A., and Zou, L. (2013). DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harb. Perspect. Biol.* 5:a012716. doi: 10.1101/cshperspect.a012716

Medici, D., and Kalluri, R. (2012). Endothelial-mesenchymal transition and its contribution to the emergence of stem cell phenotype. *Semin. Cancer Biol.* 22, 379–384. doi: 10.1016/j.semcancer.2012.04.004

Moriya, J., and Minamino, T. (2017). Angiogenesis, cancer, and vascular aging. *Front. Cardiovasc. Med.* 4:65. doi: 10.3389/fcvm.2017.00065

Moulder, D. E., Hatoum, D., Tay, E., Lin, Y., and McGowan, E. M. (2018). The roles of p53 in mitochondrial dynamics and cancer metabolism: the pendulum between survival and death in breast cancer? *Cancers (Basel)* 10:189. doi: 10.3390/cancers10060189

Naylor, A. R., Rothwell, P. M., and Bell, P. R. (2003). Overview of the principal results and secondary analyses from the European and North American randomised trials of endarterectomy for symptomatic carotid stenosis. *Eur. J. Vasc. Endovasc. Surg.* 26, 115–129. doi: 10.1053/ejvs.2002.1946

Olive, P. L., and Banath, J. P. (2006). The comet assay: a method to measure DNA damage in individual cells. *Nat. Protoc.* 1, 23–29. doi: 10.1038/nprot.2006.5

Paull, T. T. (2015). Mechanisms of ATM activation. *Annu. Rev. Biochem.* 84, 711–738. doi: 10.1146/annurev-biochem-060614-034335

Paull, T. T., Rogakou, E. P., Yamazaki, V., Kirchgessner, C. U., Gellert, M., and Bonner, W. M. (2000). A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Curr. Biol.* 10, 886–895. doi: 10.1016/s0960-9822(00)00610-2

Perisic, L., Aldi, S., Sun, Y., Folkersen, L., Razuvaev, A., Roy, J., et al. (2016). Gene expression signatures, pathways and networks in carotid atherosclerosis. *J. Intern. Med.* 279, 293–308. doi: 10.1111/joim.12448

Perisic, L., Hedin, E., Razuvaev, A., Lengquist, M., Osterholm, C., Folkersen, L., et al. (2013). Profiling of atherosclerotic lesions by gene and tissue microarrays reveals PCSK6 as a novel protease in unstable carotid atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* 33, 2432–2443. doi: 10.1161/atvbaha.113.301743

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., et al. (2011). Global quantification of mammalian gene expression control. *Nature* 473, 337–342. doi: 10.1038/nature10098

Shiloh, Y., and Ziv, Y. (2013). The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat. Rev. Mol. Cell Biol.* 14, 197–210.

Simion, V., Haemmig, S., and Feinberg, M. W. (2019). LncRNAs in vascular biology and disease. *Vascul. Pharmacol.* 114, 145–156. doi: 10.1016/j.vph.2018.01.003

Sountoulidis, A., Liontos, A., Nguyen, H. P., Firsova, A. B., Fysikopoulos, A., Qian, X., et al. (2020). SCRINSHOT enables spatial mapping of cell states in tissue sections with single-cell resolution. *PLoS Biol.* 18:e3000675.

Stone, R. C., Pastar, I., Ojeh, N., Chen, V., Liu, S., Garzon, K. I., et al. (2016). Epithelial-mesenchymal transition in tissue repair and fibrosis. *Cell Tissue Res.* 365, 495–506. doi: 10.1007/s00441-016-2464-0

Stracker, T. H., and Petrini, J. H. (2011). The MRE11 complex: starting from the ends. *Nat. Rev. Mol. Cell Biol.* 12, 90–103. doi: 10.1038/nrm3047

Tacar, O., Sriamornsak, P., and Dass, C. R. (2013). Doxorubicin: an update on anticancer molecular action, toxicity and novel drug delivery systems. *J. Pharm. Pharmacol.* 65, 157–170. doi: 10.1111/j.2042-7158.2012.01567.x

Thapar, R. (2018). Regulation of DNA double-strand break repair by non-coding RNAs. *Molecules* 23:2789. doi: 10.3390/molecules23112789

Uchida, S., and Dimmeler, S. (2015). Long noncoding RNAs in cardiovascular diseases. *Circ. Res.* 116, 737–750. doi: 10.1161/circresaha.116.302521

Wang, J. C., and Bennett, M. (2012). Aging and atherosclerosis: mechanisms, functional consequences, and potential therapeutics for cellular senescence. *Circ. Res.* 111, 245–259. doi: 10.1161/circresaha.111.261388

Wang, M., Zhang, J., Jiang, L. Q., Spinetti, G., Pintus, G., Monticone, R., et al. (2007). Proinflammatory profile within the grossly normal aged human aortic wall. *Hypertension* 50, 219–227. doi: 10.1161/hypertensionaha.107.089409

Wang, S., He, F., Li, Z., Hu, Y., Huangfu, N., and Chen, X. (2019). YB1 protects cardiac myocytes against H2O2 induced injury via suppression of PIAS3 mRNA and phosphorylation of STAT3. *Mol. Med. Rep.* 19, 4579–4588.

Warboys, C. M., de Luca, A., Amini, N., Luong, L., Duckles, H., Hsiao, S., et al. (2014). Disturbed flow promotes endothelial senescence via a p53-dependent pathway. *Arterioscler. Thromb. Vasc. Biol.* 34, 985–995. doi: 10.1161/atvbaha.114.303415

Wilusz, J. E., JnBaptiste, C. K., Lu, L. Y., Kuhn, C. D., Joshua-Tor, L., and Sharp, P. A. (2012). A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.* 26, 2392–2407. doi: 10.1101/gad.204438.112

Zhang, J., Fan, J. S., Li, S., Yang, Y., Sun, P., Zhu, Q., et al. (2020). Structural basis of DNA binding to human YB-1 cold shock domain regulated by phosphorylation. *Nucleic Acids Res.* 48, 9361–9371. doi: 10.1093/nar/gkaa619

Zufferey, R., Nagy, D., Mandel, R. J., Naldini, L., and Trono, D. (1997). Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat. Biotechnol.* 15, 871–875. doi: 10.1038/nbt0997-871

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 12 January 2021
doi: 10.3389/fgene.2020.572707

Bioinformatics Analysis Reveals MicroRNA-193a-3p Regulates ACTG2 to Control Phenotype Switch in Human Vascular Smooth Muscle Cells

Weitie Wang, Yong Wang, Hulin Piao, Bo Li, Zhicheng Zhu, Dan Li, Tiance Wang and Kexiang Liu**

Department of Cardiovascular Surgery of the Second Hospital of Jilin University, The Second Hospital of Jilin University, Changchun, China

OPEN ACCESS

Edited by:
Maarten M. G. van den Hoogenhof, Heidelberg University Hospital, Germany

Reviewed by:
Monika Gladka, Hubrecht Institute (KNAW), Netherlands
Li Zhang, Xinhua Hospital, School of Medicine, Shanghai Jiao Tong University, China

***Correspondence:**
Kexiang Liu
kxliu64@hotmail.com
Weitie Wang
iloveyuiyui@163.com

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *15 June 2020*
Accepted: *30 November 2020*
Published: *12 January 2021*

Citation:
Wang W, Wang Y, Piao H, Li B, Zhu Z, Li D, Wang T and Liu K (2021) Bioinformatics Analysis Reveals MicroRNA-193a-3p Regulates ACTG2 to Control Phenotype Switch in Human Vascular Smooth Muscle Cells. Front. Genet. 11:572707. doi: 10.3389/fgene.2020.572707

Aortic dissection (AD) is among the most fatal cardiovascular diseases. However, the pathogenesis of AD remains poorly understood. This study aims to integrate the microRNAs (miRNA) and mRNA profiles and use bioinformatics analyses with techniques in molecular biology to delineate the potential mechanisms involved in the development of AD. We used the human miRNA and mRNA microarray datasets GSE98770, GSE52093, and GEO2R, Venn diagram analysis, gene ontology, and protein–protein interaction networks to identify target miRNAs and mRNAs involved in AD. RNA interference, western blotting, and luciferase reporter assays were performed to validate the candidate miRNAs and mRNAs in AD tissues and human vascular smooth muscle cells (VSMCs). Furthermore, we studied vascular smooth muscle contraction in AD. *In silico* analyses revealed that miR-193a-3p and ACTG2 were key players in the pathogenesis of AD. miR-193a-3p was upregulated in the AD tissues. We also found that biomarkers for the contractile phenotype in VSMCs were downregulated in AD tissues. Overexpression and depletion of miR-193a-3p enhanced and suppressed VSMC proliferation and migration, respectively. Dual luciferase reporter assays confirmed that ACTG2 was a target of miR-193a-3p. ACTG2 was also downregulated in human AD tissues and VMSCs overexpressing miR-193a-3p. Taken together, miR-193a-3p may be a novel regulator of phenotypic switching in VSMCs and the miR-193a-3p/ACTG2 axis may serve as a promising diagnostic biomarker and therapeutic candidate for AD.

Keywords: mRNA, miRNA, vascular disease, aortic dissection, cardiovascular

Abbreviations: AD, aortic dissection; VSMCs, vascular smooth muscle cells; miRNAs, microRNAs; GEO, gene expression omnibus; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEGs, the differentially expressed genes; CTA, computed tomography angiography; PPI, protein-protein interaction; qRT-PCR, quantitative reverse transcription-PCR; CCK, Cell Counting Kit; FDR, false discovery rate; DEMs, differential expression profile of miRNAs; CAD, coronary heart disease; CC, cellular component; BP, biological processes; MF, molecular function.

INTRODUCTION

Type A aortic dissection (AD) is a fatal cardiovascular disease associated with high morbidity and mortality, and requires a complex treatment regimen (Elsayed et al., 2017; Silaschi et al., 2017). Vascular smooth muscle cells (VSMCs) are involved in vascular function and have been implicated in the pathogenesis of AD (Zhang et al., 2016; Wei et al., 2017). However, the precise mechanisms involved in AD remain to be fully understood. Therefore, it is important to delineate the roles of VSMCs in AD and phenotypic plasticity. This will help identify new modes of treatment, especially targeted drug therapy. Owing to advances in sequencing technology, differential genes expression between normal and damaged tissues have been widely used to identify candidate pathogenic genes (Wang et al., 2019a). Bioinformatics tools can analyze such high-throughput data, while omitting "junk" data to provide reliable analyses.

Vascular smooth muscle cells phenotypic remodeling is mainly relevant with intima-media thickening which acts important roles during the pathological progression for vascular disease, such as AD (Gomez and Owens, 2012). Although the pathogenesis of AD remains unclear, VSMCs phenotypic remodeling alternating from contractile to synthetic in response to stimulation are essential in AD, which regulates vascular remodeling. Contractile VSMCs generally demonstrate reduced ability in proliferation and migration, whereas synthetic VSMCs present enhanced viability in proliferation and migration (Alexander and Owens, 2012). Downregulation of differentiation markers such as SMA, SM22, and MYH11 always happen in synthetic VSMCs. Therefore, detection of differentiation markers suggests the VSMCs phenotypic remodeling. Numerous factors including growth factor and cell adhesion molecules promote the phenotypic switch to synthetic VSMCs and enhance cell proliferation and migration (Gomez and Owens, 2012). However, the molecular mechanisms between VSMC phenotypic switch and AD remain unclear.

MicroRNAs (miRNAs) are non-coding RNAs that function by binding to target mRNAs. miRNAs regulate various cellular functions, including proliferation and migration, and phenotypic switch in VSMCs (Tang et al., 2017; Li et al., 2019). Therefore, it is imperative to understand the importance of miRNAs and their targets in VSMC function in patients with AD.

In this study, we combined bioinformatics analysis with techniques in molecular biology to elucidate key miRNAs and mRNAs involved in AD. GEO2R analysis, target gene prediction, gene ontology (GO) and pathway analysis, Venn diagrams, and protein–protein interaction (PPI) networks were used to assess the roles of miR-193a-3p/ACTG2 as candidates associated with phenotypic switching in VSMCs. Our study presented that *in vitro* cell culture-based experiments revealed miR-193a-3p/ACTG2 participated in the development of AD.

MATERIALS AND METHODS

Datasets and Workflow

Human datasets GSE98770 (GPL14550) and GSE52093 (GPL10558) were obtained from the GEO database. We analyzed 11 samples from GSE98770 (six AD and five normal ascending tissues) and 12 samples from GSE52093 (seven AD and five normal ascending tissues). **Figure 1** shows the workflow used in the study.

Analyzing the Differential Expression of RNAs

GEO2R with R software was used to determine the gene expression profiles of AD and normal control (NC) samples and identify genes that were differentially expressed. Data from normal and AD samples were arranged in order before using the R software. $P < 0.05$ and $|\text{Log fold change}| \geq 1.0$ were used as the screening thresholds for datasets GSE98770 and GSE52093.

Bioinformatic Analysis

All differentially expressed genes (DEGs) were analyzed using the DAVID online tool that included GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment. The top pathways associated with the DEGs were further analyzed. We had previously determined the mRNA profiles for six AD and normal tissues each (Wang et al., 2019a) that were used to compare within dataset GSE52093 to identify the DEGs.

Target mRNA Identification and Generation of the PPI Network

We generated the PPI network using geneMANIA[1] for the DEGs associated with vascular smooth muscle contraction in dataset GSE52093 and our mRNA profiles. A confidence score >0.4 was considered significant.

Prediction of Differentially Expressed miRNAs (DEMs)

All miRNAs expressed in GSE52093 and the predicted mRNA profiles correlating with vascular smooth muscle contraction were selected to predict the DEMs using TargetScan.[2]

[1]https://genemania.org
[2]www.targetscan.org

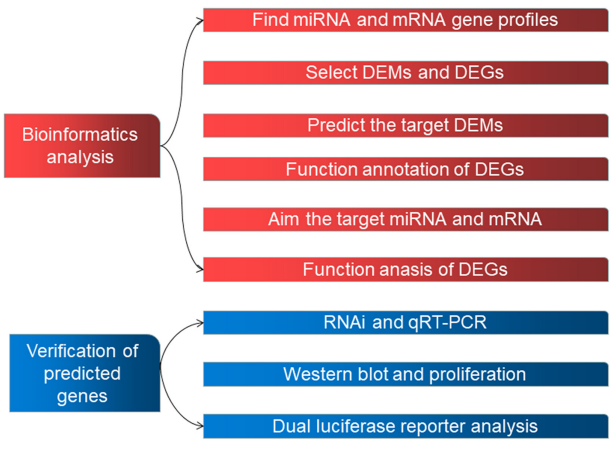

FIGURE 1 | The workflow used in this study.

Venn Diagram Analysis

Differentially expressed miRNAs from the GSE98770 dataset and miRNAs predicted in this study were analyzed using Venn diagrams generated by using FunRich (Zhang and Wang, 2019).

Human Aortic Samples

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethical Committee at Jilin University (IRB:2019018). All patients agreed to participate in this study and provided written informed consent. Human ascending aorta specimens were acquired from AD patients (n = 10) and ischemic heart disease patients (n = 10) during surgery (**Table 1**). All AD patients were diagnosed using computed tomography angiography and patients with hereditary disease were excluded.

Cell Culture

Human aortic VSMCs (Lonza, Walkersville, MD, United States) were cultured in smooth muscle culture medium (SMCM) (America Sciencell) supplemented with 2% fetal bovine serum and 1% smooth muscle cell growth supplement at 37°C in a humidified 5% CO_2 incubator. Cells used in this study were between passage number 5 and 7. VSMCs were serum starved (0.5% fetal bovine serum) for 24 h and stimulated using PDGF-BB.

Quantitative Reverse Transcription-Polymerase Chain Reaction

Total RNA was extracted using TRIzol (Invitrogen) according to the protocol provided. After measuring RNA concentration at 260 nm, the total RNA was reverse transcribed to cDNA (Takara Bio Inc., Japan) to analyze the mRNA and miRNA content in samples. **Table 2** lists all the primers used in this study. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and U6 were used as endogenous controls. We calculated relative gene expression using the $2^{-\Delta\Delta Ct}$ method.

TABLE 1 | Information of aortic dissection and ischemic heart disease patients.

	AD group	control group
	(n = 10)	(n = 10)
Age (years old)	48.3 ± 7.31	58.22 ± 5.91
Male	6(60.00%)	6(60.00%)
Maximal diameter, cm	5.81 ± 2.09	2.98 ± 0.34
Smoking	2(20.00%)	6(60.00%)
NYHA class III-IV	1(10.00%)	4(20.00%)
Hypertension	7(70.00%)	6(60.00%)
Diabetes mellitus	0(00.00%)	6(60.00%)
Chronic renal dysfunction	0	0

AD, aortic dissection; NYHA, New York Heart Association.

TABLE 2 | The primers in this work.

Gene name	Forward primer	Reverse primer
miR-193a-3p	ATGCTCAAACTGGCCTACAAG	TATGGTTGTTCTGCTCTCT GTCTC
ACTG2	GCGTGTAGCACCTGAAGAG	GAATGGCGACGTACATGGCA
SMA	GCGTGGCTATTCCTTCGTTA	ATGAAGGATGGCTGGAACAG
Calponin	AGCTAAGAGAAGGGCGGAAC	CATCTGCAGGCTGACATTGA
SM22a	AACAGCCTGTACCCTGATGG	CGGTAGTGCCCATCATTCTT
MMP-2	ACCCATTTACACCTACACCAAG	TGTTTGCAGATCTCAGGAGTG
MMP-9	CGAACTTTGACAGCGACAAG	CACTGAGGAATGATCTAAGCCC
MYH11	TGGAACTTCATCGACTTTGGG	ACAGCTTCTCCACGAAAGAC
U6	GCGCGTCGTGAAGCGTTC	GTGCAGGGTCCGAGGT
GAPDH	CGGACCAATACGACCAAATCCG	AGCCACATCGCTCAGACACC

miR-193, microRNA-193; ACTG2, actin gamma 2; SMA, smooth muscle – actin; SM22, smooth muscle 22; MYH11, smooth muscle myosin heavy chain.

VSMC Transfection

The miR-193a-3p mimic and inhibitor were designed by GenePharma (Shanghai, China). VSMCs were transfected with 250 pmol of the mimic or inhibitor and 2 μl Transfection Reagent (TransDetect, Beijing, China) for ∼5 h. Cellular miRNA expression were analyzed 48 h after transfection.

Western Blot Analysis

Total protein content of the cell was separated using sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transferred to nitrocellulose membranes. Blots were blocked using 5% non-fat milk supplemented with 0.1% Tween 20. Subsequently, the blots were incubated with primary antibodies against ACTG2 (1:500, bioss), SM22a (1:500, Proteintech), SMA (1:800, bioss), calponin (1: 2,000, bioss), MYH (1:1,000, Proteintech), MMP-2 (1:500, abcam), and MMP-9 (1:500, abcam) followed by incubation with secondary antibodies. The images were analyzed using ImageJ.

Dual Luciferase Reporter Assay

Wild or mutant 3′-untranslated region (UTR) of ACTG2 were cloned into pGL6-miR (Beyotime, China). We seeded 2×10^4 HEK293T cells onto six-well plates for 24 h before transfection. The luciferase kit (Beyotime) was used after transfection (pGL6-ACTG2-wt and pGL6-ACTG2-mut) using Lipofectamine 2000 (Invitrogen, Life Technologies). Luciferase activity was measured at 560 nm. Signal from Renilla luciferase was normalized to that from firefly luciferase.

Cell Proliferation Assay

We seeded ∼3×10^4 cells VSMC in each well of a 96-well plate for 15–20 h before transfection. Cells were then transfected with the miR-193a-3p mimic and inhibitor and incubated for 20 h. Subsequently, the cells were incubated with medium containing the Cell Counting Kit (CCK) Solution (TransDetect, Beijing, China) for 4 h following which we measured absorbance at 450 nm.

Wound Healing Assay

We seeded 2.5 × 10^5 cells into each well of a six-well plate and transfected them with miR-193a-3p. After starving for 48 h, we generated a linear wound using a 200-μl tip. Cells were then incubated in medium for 48 and 72 h. Finally, cells that migrated into the wounded area were counted.

Statistical Analysis

Continuous data are expressed as mean ± standard deviation. One-way analysis of variance (ANOVA) was performed between-group differences and multiple groups. Comparisons among multiple groups were performed using one-way ANOVAs, followed by *post hoc* Tukey tests. Relative expression of RT-qPCR was calculated using the $\Delta\Delta$Cq method.

RESULTS

DEGs in AD

All the raw data were first analyzed using the GEO2R tool. From the GSE98770 dataset, we identified 34 and 47 upregulated and downregulated miRNAs, respectively. From the GSE52093 dataset, we identified 1,409 and 1,235 upregulated and downregulated mRNAs, respectively. The volcano plots showed the DEMs (**Figure 2A**) and DEGs (**Figure 2B**) in the AD and NC samples.

GO and KEGG Pathway Analysis

DAVID was used to functionally analyze the DEGs. The downregulated mRNAs correlated with muscle

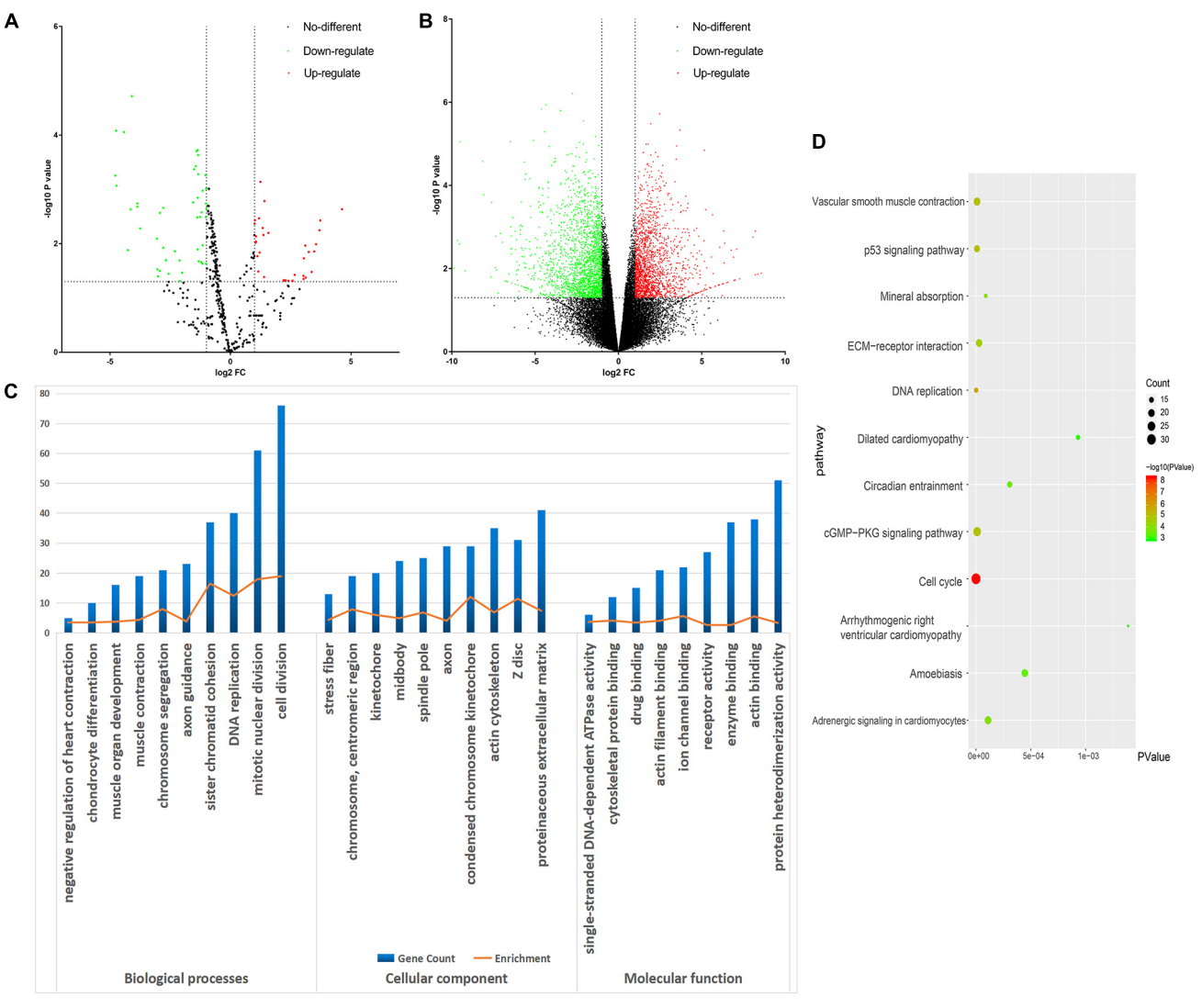

FIGURE 2 | Bioinformatic analysis of the differentially expressed genes (DEGs). **(A)** Volcano plot of the miRNA profiles. **(B)** Volcano plot for mRNA expression. We used green and red to represent decrease and increase in expression, respectively. **(C)** Gene Ontology analysis of the DEGs. **(D)** Kyoto Encyclopedia of Genes and Genomes analysis of the DEGs. Gradual color means P value and the size means gene number.

contraction and muscle organ development in AD, suggesting that downregulated genes primarily localized to the aortic media. Upregulated DEGs were involved in cell division and mitotic nuclear division, indicating the importance of cell division in AD (**Table 3** and **Figure 2C**).

Kyoto Encyclopedia of Genes and Genomes analysis revealed that the downregulated and upregulated genes were involved in vascular smooth muscle contraction and cell cycle, respectively (**Figure 2D**). Thus, GO and KEGG analyses highlighted the importance of vascular smooth muscle contraction in the aortic media during AD.

PPI Network

We generated a PPI network with genes related to the vascular smooth muscle contraction pathway using GeneMANIA bioinformatic analysis. Results showed that 8 genes (ACTG2 and PPP1R12B predicted by DAVID tool. MYH11, EDNRA, MYL2, ROCK2, ROCK1, and KCNQ1 predicted by GeneMANIA) were involved in muscular system processes (FDR:1.40e-7, nine gene count), muscle contraction (FDR:1.26e-6, eight gene count), and smooth muscle contraction (FDR:1.65e-3, four gene count) (**Figure 3A**). Then ACTG2 and PPP1R12B were used for PPI network analysis by GeneMANIA, which also indicated changing muscle function was relevant in AD (**Figure 3B**). Thus, combined with PPI analyses and previous GO and KEGG analyses, ACTG2 and PPP1R12B may be crucial genes involved in AD.

Target Gene Prediction and Validation

ACTG2 and PPP1R12B were used to predict potential target miRNAs by TargetScan. About 881 predicted miRNAs for ACTG2 and 776 predicted miRNAs for PPP1R12B (score $\geq$ 80) were collected and compared to 81 candidate miRNAs from GSE98770 to identify DEMs. Venn diagram analysis showed 49 DEMs (**Figure 3C**) that were common to data from GSE98770 and the miRNAs predicted in this study. A miRNA–mRNA network was constructed using Cytoscape (**Figure 3D**). miR-193a-3p was upregulated and among the top 3 predicted miRNAs with high score and binding sites for ACTG2. Furthermore, miR-193a-3p is crucial for cell proliferation. Therefore, miR-193a-3p may regulate VSMC proliferation.

TABLE 3 | Enriched analysis.

Expression	Category	Term	Description	Gene count	*P*-value
UP-DEGs	BP	cell division	GO:0051301	76	1.31E-19
	BP	mitotic nuclear division	GO:0007067	61	1.36E-18
	BP	sister chromatid cohesion	GO:0007062	37	3.82E-17
	BP	DNA replication	GO:0006260	40	4.61E-13
	BP	chromosome segregation	GO:0007059	21	1.29E-08
	CC	condensed chromosome kinetochore	GO:0000777	29	1.02E-12
	CC	chromosome, centromeric region	GO:0000775	19	1.61E-08
	CC	spindle pole	GO:0000922	25	1.73E-07
	CC	cytosol	GO:0005829	290	3.39E-07
	CC	membrane	GO:0016020	204	6.67E-07
	MF	protein binding	GO:0005515	643	2.32E-04
	MF	single-stranded DNA-dependent ATPase activity	GO:0043142	6	2.46E-04
	MF	drug binding	GO:0008144	15	4.36E-04
	MF	protein heterodimerization activity	GO:0046982	51	5.44E-04
	MF	ATP binding	GO:0005524	130	0.001165931
DOWN-DEGs	BP	muscle contraction	GO:0006936	19	5.16E-05
	BP	axon guidance	GO:0007411	23	1.69E-04
	BP	muscle organ development	GO:0007517	16	2.09E-04
	BP	chondrocyte differentiation	GO:0002062	10	3.58E-04
	BP	negative regulation of heart contraction	GO:0045822	5	3.67E-04
	CC	Z disc	GO:0030018	31	4.67E-12
	CC	proteinaceous extracellular matrix	GO:0005578	41	4.59E-08
	CC	actin cytoskeleton	GO:0015629	35	1.60E-07
	CC	plasma membrane	GO:0005886	296	4.28E-05
	CC	stress fiber	GO:0001725	13	5.30E-05
	MF	calcium ion binding	GO:0005509	78	2.79E-07
	MF	ion channel binding	GO:0044325	22	2.68E-06
	MF	actin binding	GO:0003779	38	3.43E-06
	MF	cytoskeletal protein binding	GO:0008092	12	9.05E-05
	MF	actin filament binding	GO:0051015	21	9.99E-05

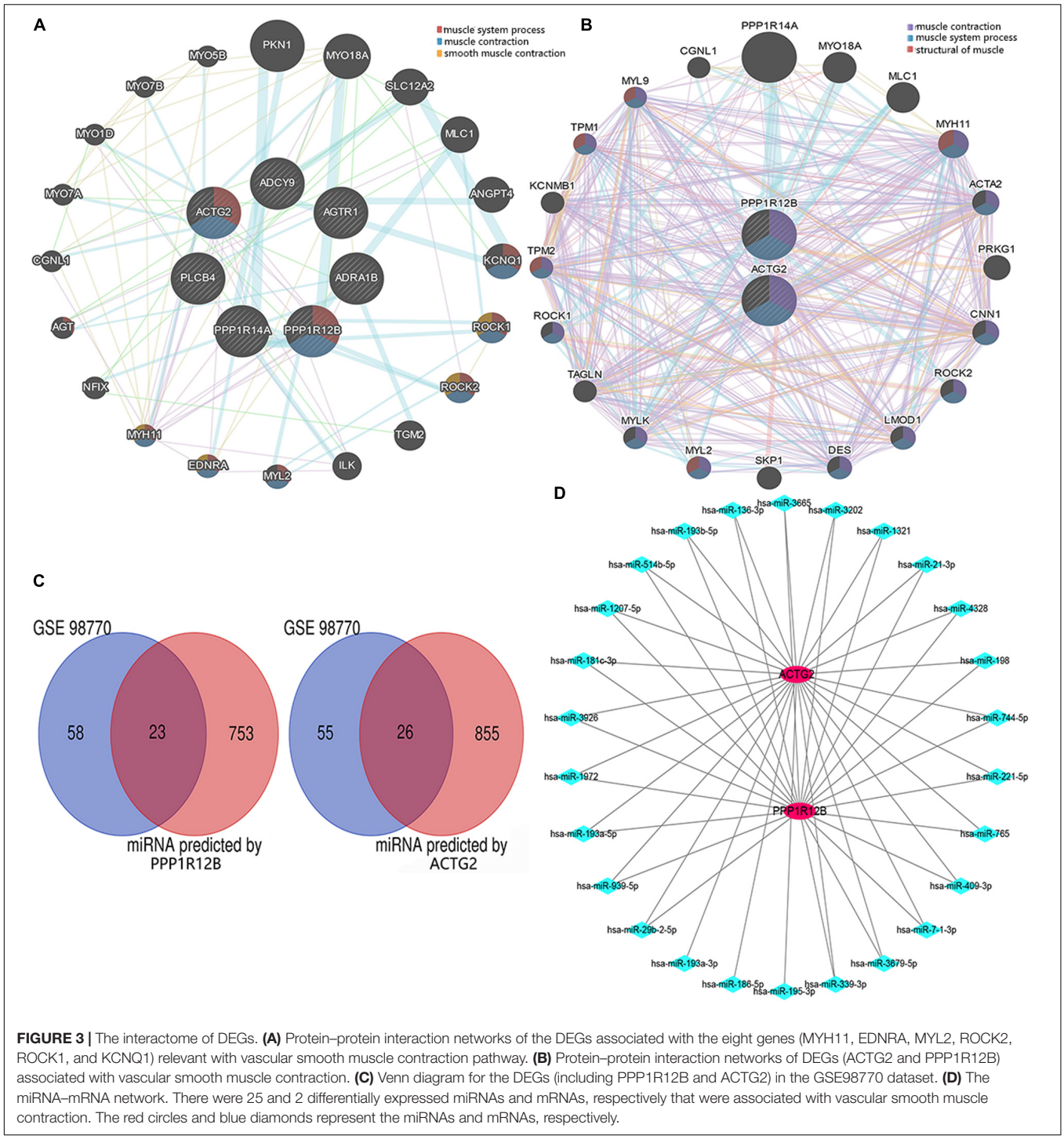

FIGURE 3 | The interactome of DEGs. **(A)** Protein–protein interaction networks of the DEGs associated with the eight genes (MYH11, EDNRA, MYL2, ROCK2, ROCK1, and KCNQ1) relevant with vascular smooth muscle contraction pathway. **(B)** Protein–protein interaction networks of DEGs (ACTG2 and PPP1R12B) associated with vascular smooth muscle contraction. **(C)** Venn diagram for the DEGs (including PPP1R12B and ACTG2) in the GSE98770 dataset. **(D)** The miRNA–mRNA network. There were 25 and 2 differentially expressed miRNAs and mRNAs, respectively that were associated with vascular smooth muscle contraction. The red circles and blue diamonds represent the miRNAs and mRNAs, respectively.

miR-193a-3p Regulates VSMC Phenotypes

Tissues near the intimal tear isolated from patients with AD exhibited decreased expression of differentiation biomarkers (SMA, SM22, calponin, and MYH11) and contained highly proliferative VSMCs compared to those in tissues from healthy individuals (**Figures 4A,B**). At the meantime, miR-193a-3p also increased in the AD tissues (**Figure 4C**). These results highlight the involvement of miR-193a-3p in human VSMC function and phenotypic plasticity.

Overexpression of miR-193a-3p in Proliferative VSMCs

To identify that miR-193a-3p is specifically involved in the regulation of VSMC, quantitative reverse

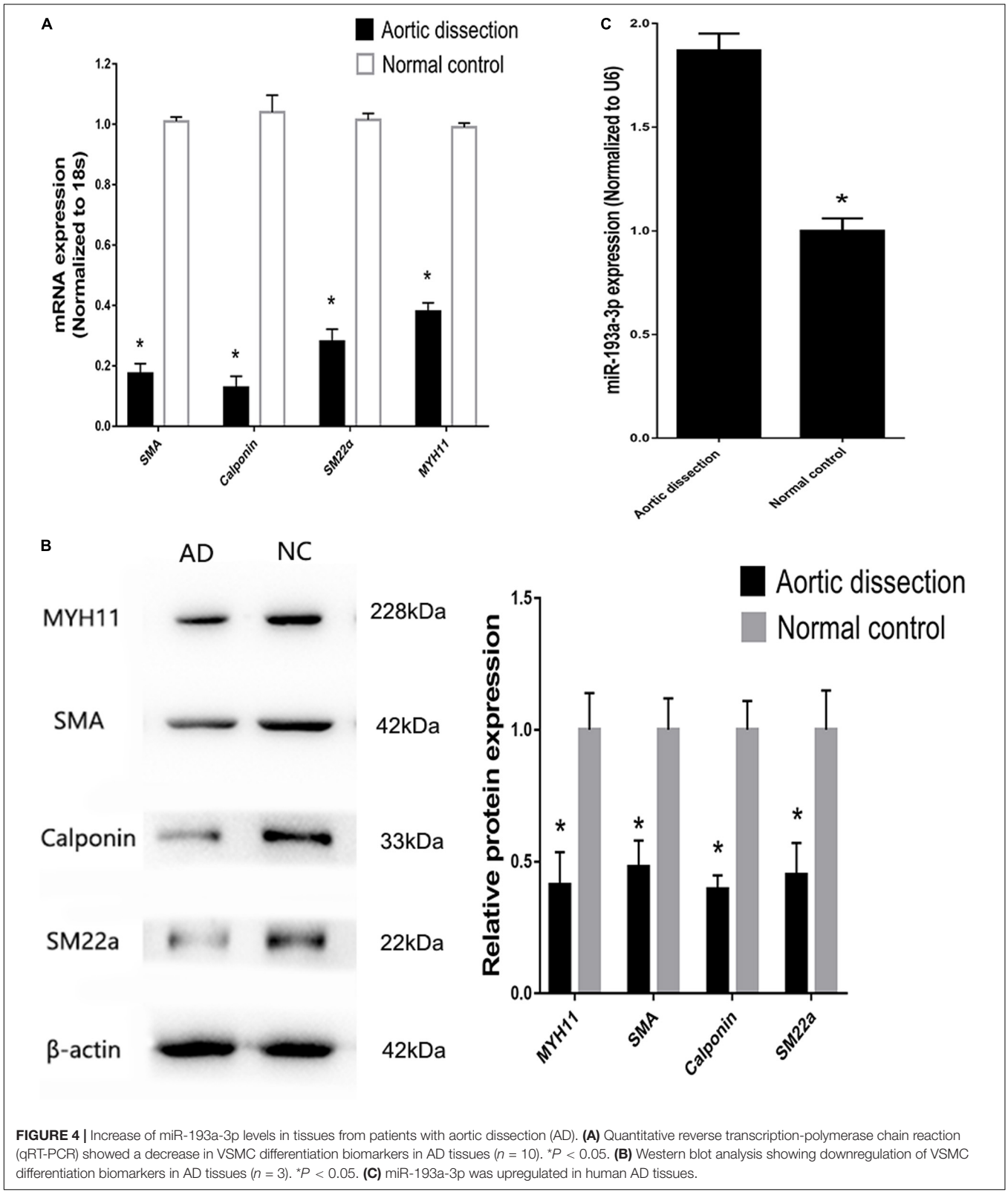

FIGURE 4 | Increase of miR-193a-3p levels in tissues from patients with aortic dissection (AD). **(A)** Quantitative reverse transcription-polymerase chain reaction (qRT-PCR) showed a decrease in VSMC differentiation biomarkers in AD tissues (n = 10). *P < 0.05. **(B)** Western blot analysis showing downregulation of VSMC differentiation biomarkers in AD tissues (n = 3). *P < 0.05. **(C)** miR-193a-3p was upregulated in human AD tissues.

transcription-polymerase chain reaction (qRT-PCR) was used to detect miR-193a-3p expression levels in VSMCs with increasing duration of proliferation (2, 4, 6, 8, 12, and 24 h). The qRT-PCR results showed that miR-193a-3p was upregulated in VSMCs with an increase in cell proliferation. As shown in **Figure 5**, miR-193a-3p was upregulated in

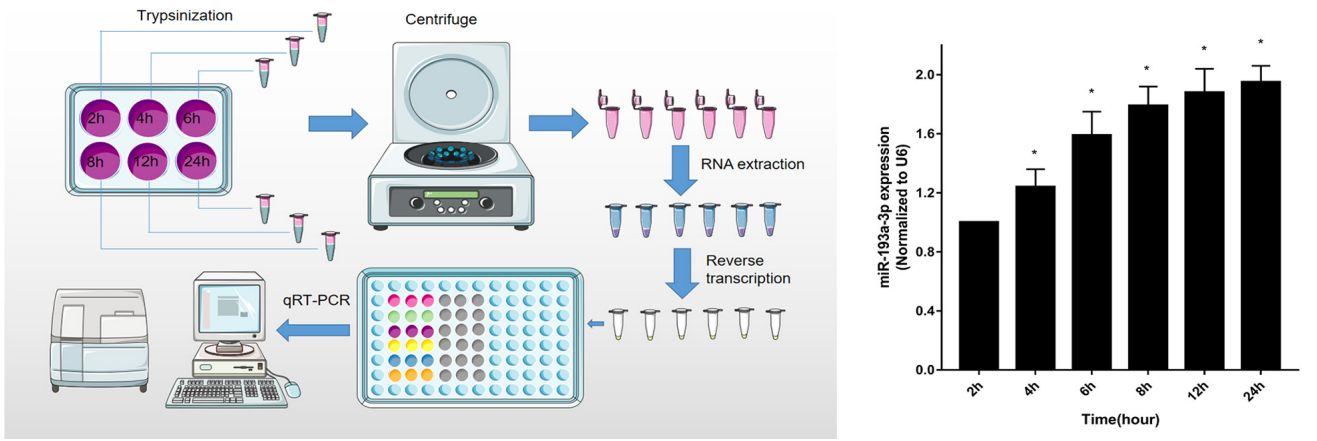

FIGURE 5 | Upregulation of miR-193a-3p in proliferating VSMCs. qRT-PCR showed the time-dependent increase of miR-193a-3p in VSMCs ($n = 5$). $^*P < 0.05$ compared to the 24 h cells.

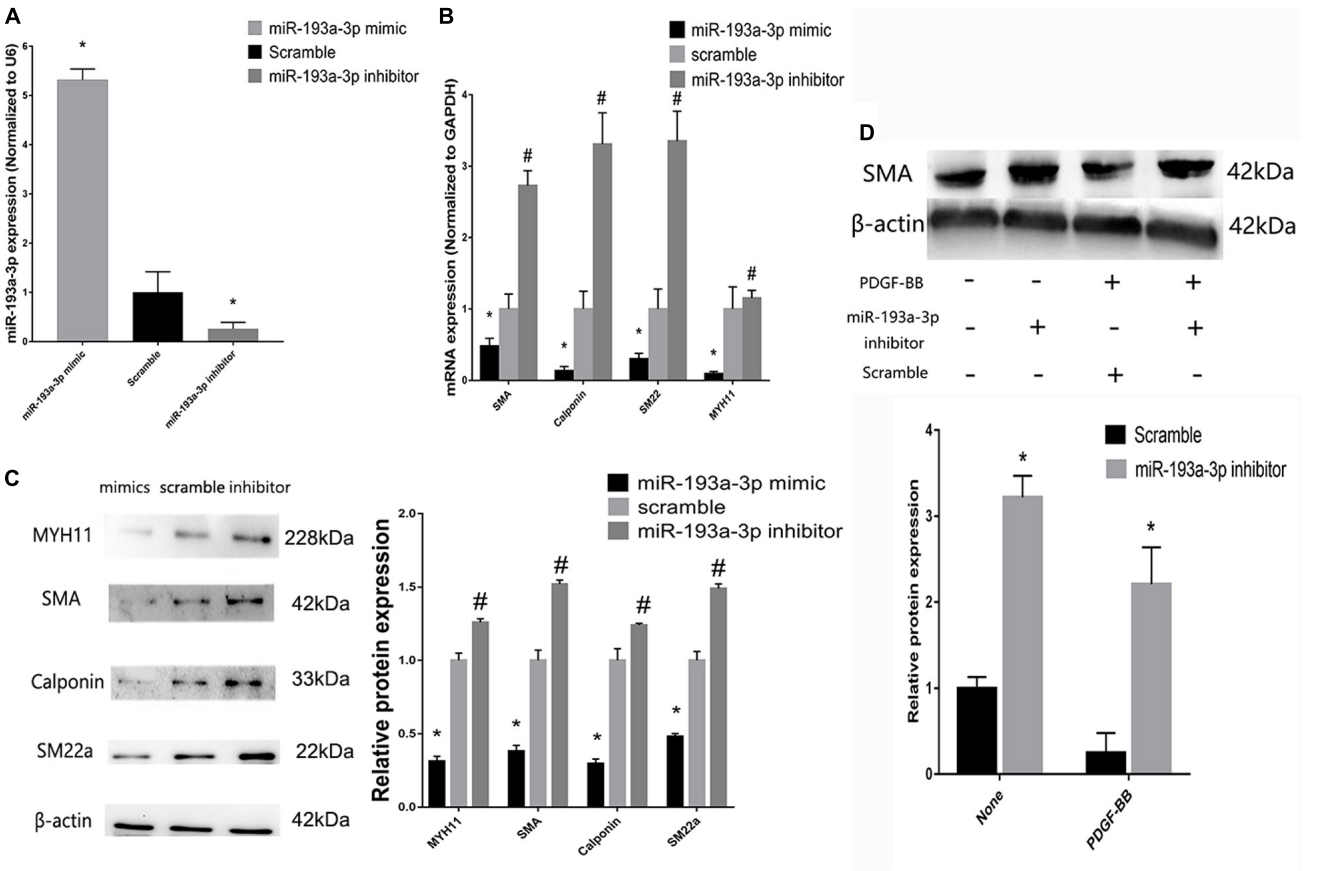

FIGURE 6 | miR-193a-3p regulates smooth muscle biomarkers *in vitro*. **(A)** qRT-PCR showing the overexpression and downregulation of miR-193a-3p after transfection with the miR-193a-3p mimic and inhibitor, respectively ($n = 10$, $^*P \leq 0.05$ vs. Scramble). **(B)** miR-193a-3p inhibitor increased the expression of the differentiation biomarkers, whereas miR-193a-3p mimic downregulated the biomarkers ($n = 10$, $^*P \leq 0.05$ vs. Scramble, $^\#P < 0.05$ vs. Scramble). **(C)** Western blots showing the decrease and increase in differentiation biomarker expression in miR-193a-3p inhibitor and mimic-transfected cells, respectively ($n = 3$, $^*P < 0.05$ vs. Scramble, $^\#P \leq 0.05$ vs. Scramble). **(D)** Effect of miR-193a-3p on SMA expression in PDGF-BB stimulated VSMCs by densitometric analysis of SMA protein levels as measured by Western blot analysis ($n = 5$). $^*P < 0.05$ vs. scramble with or without PDGF-BB treatment (20 ng/ml).

a time-dependent manner together with an increase in VSMC proliferation.

miR-193a-3p Regulates Smooth Muscle Biomarkers in VSMCs

To investigate the role of miR-193a-3p in human aortic VSMC phenotype switch, we transiently transfected miR-193a-3p mimic and miR-193a-3p inhibitor into human aortic VSMCs. qRT-PCR was used to confirm the transfection efficiency of the miR-193a-3p mimic, inhibitor, and controls. **Figure 6A** showed that the miR-193a-3p mimic elevated miR-193a-3p levels, whereas miR-193a-3p inhibitor markedly reduced the endogenous levels of miR-193a-3p in VSMCs. Western blotting and qRT-PCR showed the decrease in biomarkers involved in VSMC differentiation, such as SMA, SM22a, calponin, and MYH11 (**Figures 6B,C**),

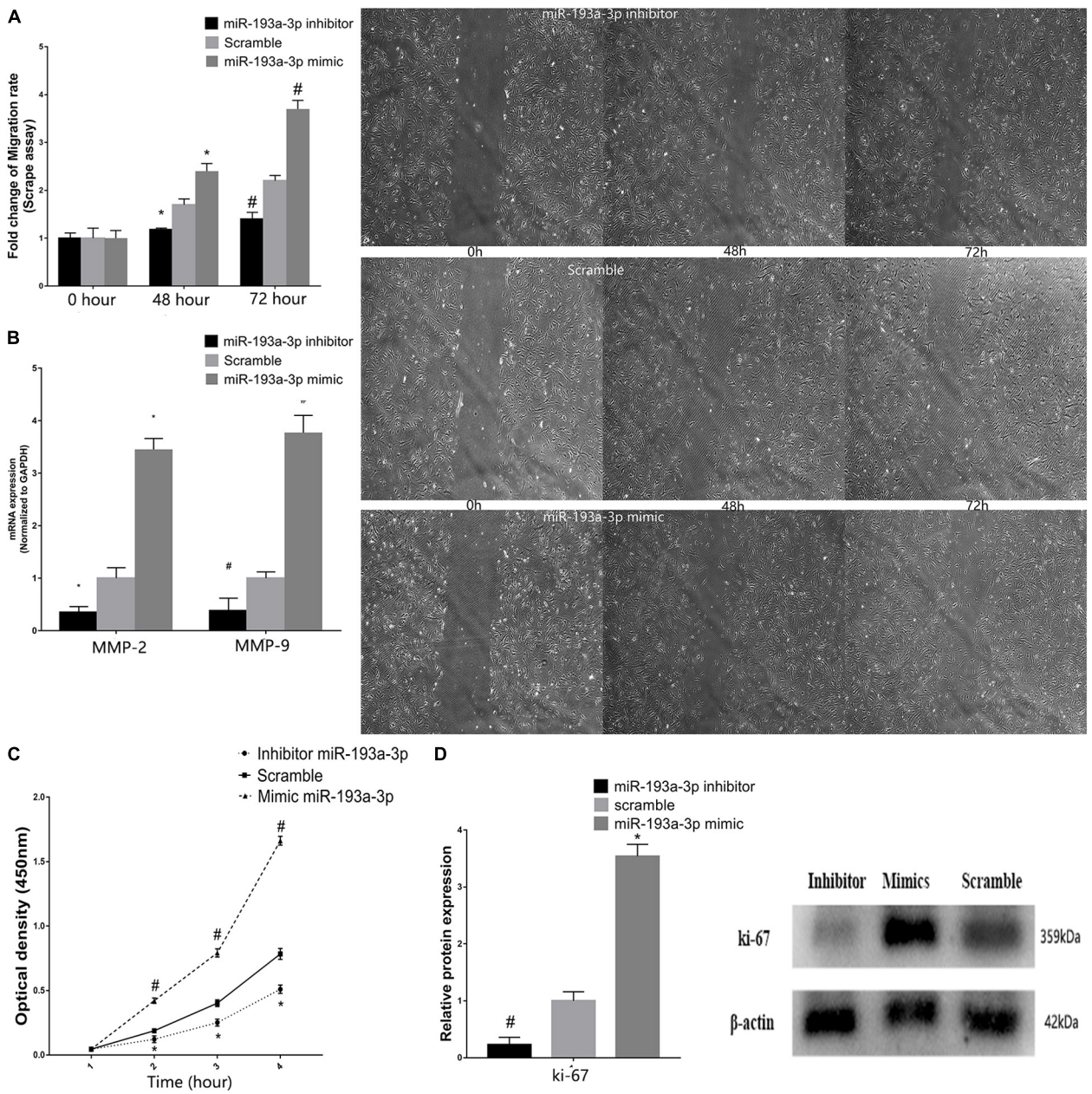

FIGURE 7 | miR-193a-3p in VSMC proliferation and migration. **(A)** Scrape assay showing the reduction in the migration of human aortic VSMCs transfected with miR-193a-3p inhibitor. Cells transfected with the miR-193a-3p mimic exhibited enhanced cell migration (*$P \leq 0.05$ vs. Scramble, #$P \leq 0.05$ vs. Scramble). **(B)** qRT-PCR showing the increase in mRNA levels of MMP-2 and MMP-9 in VSMCs transfected with the miR-193a-3p mimic. **(C)** The miR-193a-3p mimic enhanced VSMC proliferation as measured by the cell counting kit-8 ($n = 10$, *$P < 0.05$ vs. 1 h, #$P \leq 0.05$ vs. Scramble). **(D)** Western blots showing the upregulation of Ki-67 in cells transfected with the miR-193a-3p mimic (*$P \leq 0.05$ vs. Scramble, #$P \leq 0.05$ vs. Scramble).

whereas lower levels of miR-193a-3p were associated with upregulation of these biomarkers in human VSMCs. Western blot analysis also demonstrated that the SMA protein levels were dramatically increased after ectopic downregulation of miR-193a-3p under both basal and PDGF-BB-stimulated conditions. Thus, miR-193a-3p levels regulated VSMC differentiation.

miR-193a-3p Regulates VSMC Proliferation and Migration

Vascular smooth muscle cells were transfected with miR-193a-3p mimic, inhibitor, or control. Subsequently, the scratch assay was used to evaluate the effect of miR-193a-3p on cell migration. **Figure 7A** showed the marked increase in migration in cells transfected with the miR-193a-3p mimic. In addition, we also detected the migration biomakers such as MMP-2 and MMP-9, which have been implicated in VSMC migration. MMP-2 and MMP-9 RNA levels were also higher in the miR-193a-3p mimic-transfected cells (**Figure 7B**). These results showed that high expression of miR-193a-3p promoted VSMC migration. As shown in **Figure 7C**, we assayed the effect of miR-193a-3p on VSMC proliferation using CCK-8. Cell proliferation was recorded at 1, 2, 3, and 4 h. Cells overexpressing miR-193a-3p exhibited increased proliferation as compared to the control cells. However, VSMCs depleted of miR-193a-3p showed reduced proliferation. Moreover, we also detected the proliferation biomaker (Ki-67). Western blots showed the upregulation of Ki-67 in miR-193a-3p-overexpressing VSMCs, whereas inhibition of miR-193a-3p in VSMCs downregulated Ki-67 (**Figure 7D**). Taken together, these results indicated that miR-193a-3p was a promoter of VSMC proliferation and migration.

ACTG2 Is a Target of miR-193a-3p

Figure 8A shows the predicted binding sites for miR-193a-3p within the 3′ UTR of ACTG2. We used the dual luciferase reporter assay to detect the effect of miR-193a-3p on the wild-type and mutant 3′ UTR of ACTG2. HEK293T cells overexpressing miR-193a-3p decreased luciferase activity to a greater extent in cells expressing wildtype ACTG2, as compared to that in cells transfected with miR-193a-3p control or expressing mutant

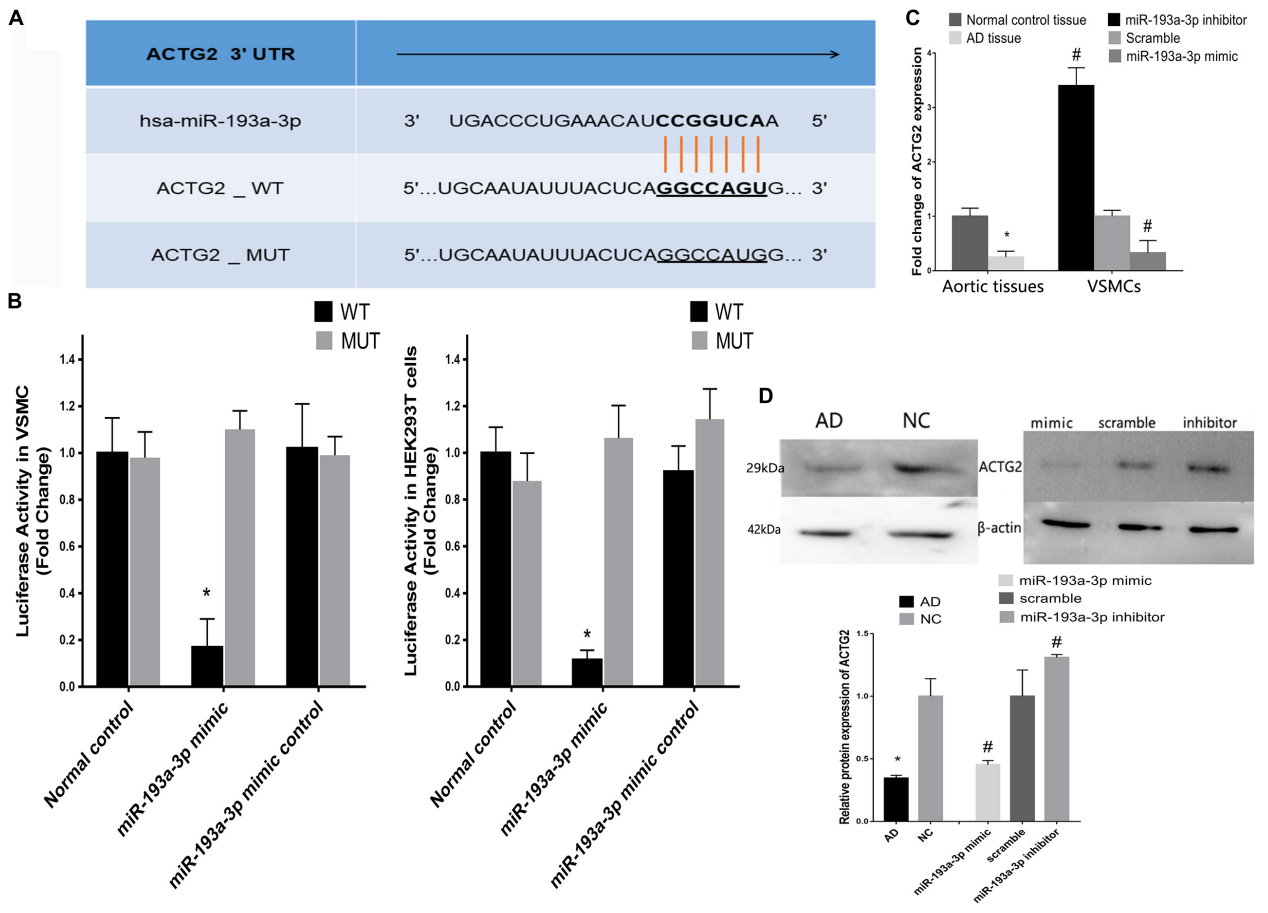

FIGURE 8 | Correlation between miR-193a-3p and ACTG2 in AD. **(A)** miR-193a-3p targets the 3′ untranslated region of ACTG2. **(B)** Dual luciferase assay showing the effect of miR-193a-3p on the wildtype and mutant 3′untranslated region of ACTG2 in HEK293T cells ($n = 6$, $^{*}P < 0.05$). **(C)** Reduction in the mRNA levels of ACTG2 in tissues from patients with AD and VSMCs transfected with the miR-193a-3p mimic ($^{*}P < 0.05$, $^{\#}P < 0.05$). **(D)** Western blots showing the decrease in the protein levels of ACTG2 in tissues from patients with AD and miR-193a-3p mimic-transfected VSMCs ($^{*}P < 0.05$, $^{\#}P < 0.05$).

ACTG2 (**Figure 8B**). Moreover, qRT-PCR revealed that ACTG2 was downregulated in tissues from AD patients. However, VSMCs transfected with the miR-193a-3p mimic exhibited upregulation of ACTG2 in tissues from healthy individuals and VSMCs depleted of miR-193a-3p (**Figures 8C,D**).

pcDNA3.1-ACTG2 vector and ACTG2 siRNA were transfected into VSMC for rescue experiments. qRT-PCR was used to verify the transfection efficiency of pcDNA3.1-ACTG2 vector and ACTG2 siRNA. pcDNA3.1-ACTG2 vector could elevate the expression of ACTG2 in VSMC, whereas ACTG2 siRNA reduced ACTG2 expression (**Figure 9A**). miR-193a-3p inhibitor + ACTG2 siRNA showed rescue proliferation and migration (**Figure 9B**) ability compared to that of VSMCs transfected with miR-193a-3p inhibitor only. Phenotypic transition biomarkers involved in VSMC differentiation were lower in miR-193a-3p inhibitor + ACTG2 siRNA than VSMCs transfected with miR-193a-3p inhibitor only (**Figures 9C,D**).

DISCUSSION

Vascular smooth muscle cells in aortic media maintain the biological properties of the aortic wall (Rzucidlo et al., 2007; Ren et al., 2017; Chen et al., 2019; Wang et al., 2019b). Proliferation and migration of VSMCs are regulated by phenotypic switching and correlate with the initial stages involved in the development of AD (An et al., 2017a,b; Li et al., 2018; Zhang and Wang, 2019). However, the mechanism(s) of phenotypic switching in VSMCs remain to be understood in detail. In this study, we observed an upregulation of miR-193a-3p in highly proliferative VSMCs and AD tissues. Up- or downregulating miR-193a-3p resulted in the decrease or increase of the differentiation biomarkers (SMA, MYH11, SM22a, and calponin) in VSMCs, respectively. Therefore, miR-193a-3p stimulates cell proliferation and may be a novel modulator of phenotype switching in VSMCs.

In this study, integrating the miRNA and mRNA profiles and using bioinformatics analyses revealed that the miR-193a-3p/ACTG2 axis played an essential role in the pathogenesis of AD. The reason we choose miR-193a-3p/ACTG2 as the target axis is base on the bioinformatic analyses. In addition, PPP1R12B was also predicted as a potential gene in vascular smooth muscle contraction pathway. However, few references reported the interaction between PPP1R12B and vascular smooth muscle contraction pathway as well as AD. miR-193a-3p inhibits the formation, proliferation, and migration of tumor cells in the lung by directly binding to KRAS (Fan et al., 2017). It also functions as a tumor suppressor in colon cancer by interacting with IL17RD (Pekow et al., 2017). However, studies have reported that miR-193a-3p promotes the formation, proliferation, and migration of tumor cells. miR-193a-3p enhances the proliferation and migration of renal carcinoma cells by directly targeting PTEN (Liu et al., 2017). Furthermore, it promotes the development of

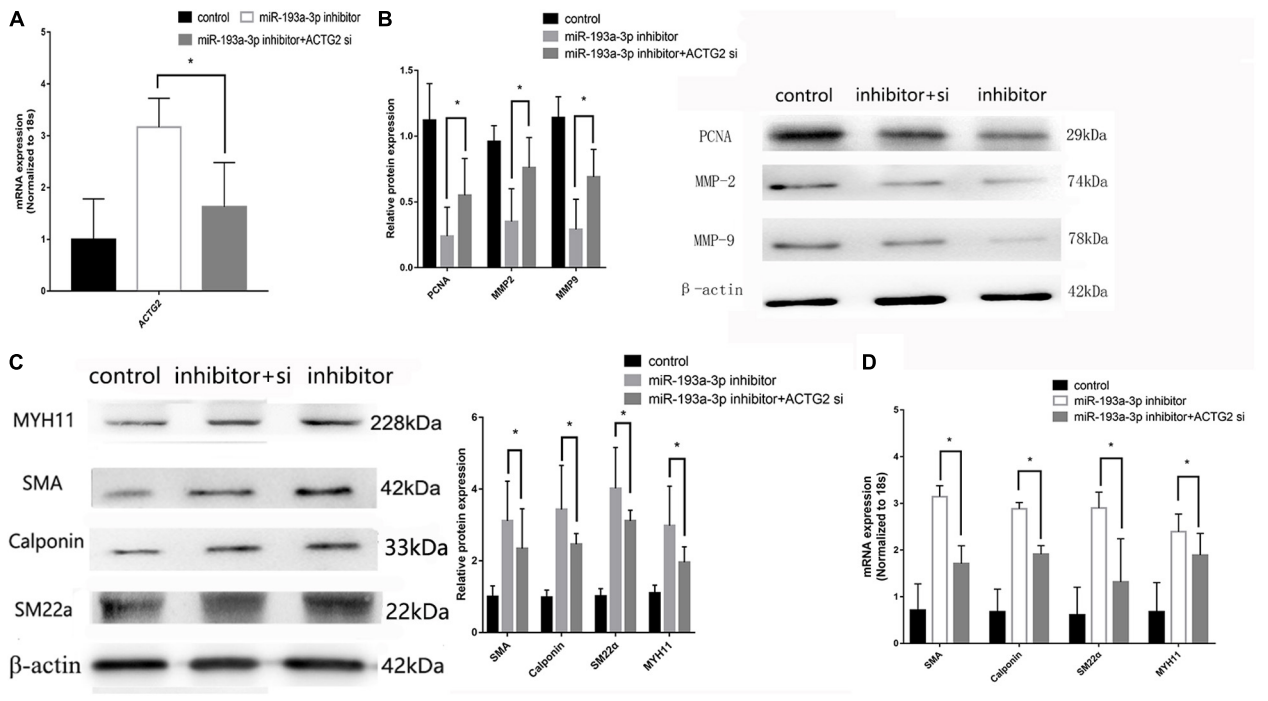

FIGURE 9 | Rescue experiments of ACTG2. **(A)** qRT-PCR presented over/downregulation of ACTG2 after transfection with ACTG2 vector and siRNA, respectively ($n = 3$, $^*P \leq 0.05$ vs. control). **(B)** Western blot showed remedy PCNA, MMP2, and MMP9 after transfection with miR-193a-3p inhibitor + ACTG2 siRNA than transfection with miR-193a-3p inhibitor only ($n = 10$, * vs. miR-193a-3p inhibitor). **(C)** Western blot analysis showed that low expression of VSMCs differentiation biomarkers in cells transfected with miR-193a-3p inhibitor + ACTG2 siRNA comparing to transfection with miR-193a-3p inhibitor only ($n = 3$, * vs. miR-193a-3p inhibitor). **(D)** qRT-PCR presented remedy expression of VSMC differentiation biomarkers.

bladder cancer by targeting HOXC9 (Faugeroux et al., 2015). However, the role of miR-193a-3p on cardiovascular disease, especially AD, has not been studied. Therefore, the identification of miR-193a-3p as a modifier of VSMC phenotype and proliferation will increase our understanding of the pathogenesis of heart disease.

We observed that increased expression of miR-193a-3p (in AD tissues) correlated with the reduced expression of differentiation biomarkers in VSMCs. Furthermore, miR-193a-3p was overexpressed in proliferating human aortic VSMCs that had been proliferating for longer durations. CCK-8 and scrape assays showed that the miR-193a-3p mimic and inhibitor increased and suppressed the proliferation and migration of human VSMCs, respectively. These results help strengthen the inference that miR-193a-3p may be a modulator of phenotypic switching in VSMCs.

MicroRNAs regulate translation by binding to the 3′ UTRs of target mRNAs (Liu et al., 2019). Using various molecular biology techniques and bioinformatics analyses, we have demonstrated that ACTG2 is a target of miR-193a-3p. Overexpression of miR-193a-3p resulted in the downregulation of ACTG2 in transfected VSMCs. These VSMCs exhibited enhanced proliferation and reduced expression of the differentiation biomarkers SMA, MYH11, SM22a, and calponin.In addition, we found the converse regulation between miR-193a-3p and the differentiation markers. So targetscan bioinformatic analyses toold was used to predict binding sites of miR-193a-3p. The results showed no binding sites between miR-193a-3p and SMA, MYH11, SM22a, and calponin. Therefore, we concluded that miR-193a-3p could not directly regulate these differentiation markers.

However, this study also contains severel limitations. All *in vitro* experiments were performed in VSMCs, the different expression of ACTG2 and miR-193a-3p in endothelial remained unclear. In addition, two gene profiles were used in our study, which might cause bias. More sequencing results need to be used for co-expression analysis. Finally, the potential roles of PPP1R12B in VSMCs phenotypic transition need further verification experiments.

In summary, our study revealed that miR-193a-3p was increased in ascending aortic tissues from AD patients. Moreover, miR-193a-3p targeted the 3′ UTR of ACTG2 and may be a novel regulator of phenotypic switching, proliferation, and migration in VSMCs. Thus, the miR-193a-3p/ACTG2 axis may provide mechanistic insight into the pathogenesis of AD and serve as a promising diagnostic biomarker and therapeutic target for AD.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: GSE98770 and GSE52093.

ETHICS STATEMENT

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethical Committee of Jilin University. All patients agreed to participate this program and provided written informed consent.

AUTHOR CONTRIBUTIONS

WW designed and supervised the study. TW, KL, HP, YW, and BL performed the analysis work. ZZ and DL contributed to the data analysis. WW and KL organized, designed, and wrote the manuscript. All authors reviewed the final manuscript.

FUNDING

This work was supported by the project of National Natural Science Foundation of China, China (Grant no. 81970399), the project of special health project of Jilin Provincial Department of Finance, China (Grant no. 20180101), the project of Jilin Key Laboratory Construction Project, China (Grant no. 20190901008JC).

REFERENCES

Alexander, M. R., and Owens, G. K. (2012). Epigenetic control of smooth muscle cell differentiation and phenotypic switching in vascular development and disease. *Annu. Rev. Physiol.* 74, 13–40. doi: 10.1146/annurev-physiol-012110-142315

An, Z., Liu, Y., Song, Z. G., Tang, H., Yuan, Y., Xu, Z.-Y., et al. (2017a). Mechanisms of aortic dissection smooth muscle cell phenotype switch. *J. Thorac. Cardiovasc. Surg.* 154, 1511–1521. doi: 10.1016/j.jtcvs.2017.05.066

An, Z., Qiao, F., Lu, Q., Ma, Y., Liu, Y., Lu, F., et al. (2017b). Interleukin-6 downregulated vascular smooth muscle cell contractile proteins via ATG4B-mediated autophagy in thoracic aortic dissection. *Heart Vessels* 32, 1523–1535. doi: 10.1007/s00380-017-1054-8

Chen, D., Zang, Y. H., Qiu, Y., Zhang, F., Chen, A.-D., Wang, J.-J., et al. (2019). BCL6 attenuates proliferation and Oxidative stress of vascular smooth Muscle Cells in Hypertension. *Oxid. Med. Cell Longev.* 2019:5018410.

Elsayed, R. S., Cohen, R. G., Fleischman, F., and Bowdish, M. E. (2017). Acute Type A Aortic dissection. *Cardiol. Clin.* 35, 331–345.

Fan, Q., Hu, X., and Zhang, H. (2017). MiR-193a-3p is an important tumour suppressor in lung cancer and directly targets KRAS. *Cell Physiol. Biochem.* 44, 1311–1324. doi: 10.1159/000485491

Faugeroux, J., Lv, L., Li, Y., Deng, H., Zhang, C., Pu, Y., et al. (2015). MiR-193a-3p promotes the multi-chemoresistance of bladder cancer by targeting the HOXC9 gene. *Cancer Lett.* 357, 105–113. doi: 10.1016/j.canlet.2014.11.002

Gomez, D., and Owens, G. K. (2012). Smooth muscle cell phenotypic switching in atherosclerosis. *Cardiovasc. Res,* 95, 156–164. doi: 10.1093/cvr/cvs115

Li, R., Yi, X., Wei, X., Huo, B., Guo, X., Cheng, C., et al. (2018). EZH2 inhibits autophagic cell death of aortic vascular smooth muscle cells to affect aortic dissection. *Cell Death Dis.* 9:180.

Li, X. L., de Molfetta, G. A., and Ramão, A. (2019). miR-450a acts as a tumor suppressor in ovarian cancer by regulating energy metabolism. *Cancer Res.* 79, 3294–3305. doi: 10.1158/0008-5472.can-19-0490

Liu, L., Li, Y., Liu, S., Duan, Q., Chen, L., Wu, T., et al. (2017). Downregulation of miR-193a-3p inhibits cell growth and migration in renal cell carcinoma by targeting PTEN. *Tumour. Biol.* 39:1010428317711951.

Liu, Y., Chen, G., Liu, H., Li, Z., Yang, Q., Gu, X., et al. (2019). Integrated bioinformatics analysis of miRNA expression in Ewing sarcoma and potential

regulatory effects of miR-21 via targeting ALCAM/CD166. *Artif. Cells Nanomed.Biotechnol.* 47, 2114–2122. doi: 10.1080/21691401.2019.1620760

Pekow, J., Meckel, K., Dougherty, U., Huang, Y., Chen, X., Almoghrabi, A., et al. (2017). miR-193a-3p is a Key tumor suppressor in ulcerative colitis-associated colon cancer and promotes carcinogenesis through upregulation of IL17RD. *Clin. Cancer Res.* 3, 5281–5291. doi: 10.1158/1078-0432.ccr-17-0171

Ren, X. S., Tong, Y., Ling, L., Chen, D., Sun, H.-J., Zhou, H., et al. (2017). NLRP3 gene deletion attenuates angiotensin II-induced phenotypic transformation of vascular smooth Muscle Cells and vascular remodeling. *Cell Physiol. Biochem.* 44, 2269–2280. doi: 10.1159/000486061

Rzucidlo, E. M., Martin, K. A., and Powell, R. J. (2007). Regulation of vascular smooth muscle cell differentiation. *J. Vasc. Surg.* 45, A25–A32.

Silaschi, M., Byrne, J., and Wendler, O. (2017). Aortic dissection: medical, interventional and surgical management. *Heart* 103, 78–87. doi: 10.1136/heartjnl-2015-308284

Tang, Y., Yu, S., Liu, Y., Zhang, J., Han, L., Xu, Z., et al. (2017). MicroRNA-124 controls human vascular smooth muscle cell phenotypic switch via Sp1. *Am. J. Physiol. Heart Circ. Physiol.* 313, 641–649.

Wang, W., Liu, Q., Wang, Y., Piao, H., Li, B., Zhu, Z., et al. (2019a). Verification of hub genes in the expression profile of aortic dissection. *PLoS One* 14:e0224922. doi: 10.1371/journal.pone.0224922

Wang, W., Wang, T., Wang, Y., Piao, H., Li, B., Zhu, Z., et al. (2019b). Integration of gene expression profile data to verify hub genes of patients with stanford A Aortic dissection. *Biomed. Res. Int,* 2019:3629751.

Wei, X., Sun, Y., Wu, Y., Zhu, J., Gao, B., Yan, H., et al. (2017). Downregulation of Talin-1 expression associates with increased proliferation and migration of vascular smooth muscle cells in aortic dissection. *BMC Cardiovasc. Disord.* 17:162. doi: 10.1186/s12872-017-0588-0

Zhang, M., and Wang, Z. (2019). Downregulation of miR143/145 gene cluster expression promotes the aortic media degeneration process via the TGF-β1 signaling pathway. *Am. J. Transl. Res.* 11, 370–378.

Zhang, P., Hou, S., Chen, J., Zhang, J., Lin, F., Ju, R., et al. (2016). Smad4 Deficiency in Smooth Muscle Cells Initiates the Formation of Aortic Aneurysm. *Circ. Res.* 118, 388–399. doi: 10.1161/circresaha.115.308040

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

frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 21 January 2021
doi: 10.3389/fcvm.2020.631775

Integrative Identification of Hub Genes Associated With Immune Cells in Atrial Fibrillation Using Weighted Gene Correlation Network Analysis

Tao Yan†, Shijie Zhu†, Miao Zhu†, Chunsheng Wang and Changfa Guo**

Department of Cardiovascular Surgery, Zhongshan Hospital, Fudan University, Shanghai, China

OPEN ACCESS

Edited by:
Abdelaziz Beqqali, University of Edinburgh, United Kingdom

Reviewed by:
Mingxian Chen, Central South University, China
Elisa Cairrão, Universidade da Beira Interior, Portugal

***Correspondence:**
Chunsheng Wang
wangchunsheng@fudan.edu.cn
Changfa Guo
guo.changfa@zs-hospital.sh.cn

†These authors have contributed equally to this work

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

Received: *20 November 2020*
Accepted: *30 December 2020*
Published: *21 January 2021*

Citation:
Yan T, Zhu S, Zhu M, Wang C and Guo C (2021) Integrative Identification of Hub Genes Associated With Immune Cells in Atrial Fibrillation Using Weighted Gene Correlation Network Analysis. Front. Cardiovasc. Med. 7:631775. doi: 10.3389/fcvm.2020.631775

Background: Atrial fibrillation (AF) is the most common tachyarrhythmia in the clinic, leading to high morbidity and mortality. Although many studies on AF have been conducted, the molecular mechanism of AF has not been fully elucidated. This study was designed to explore the molecular mechanism of AF using integrative bioinformatics analysis and provide new insights into the pathophysiology of AF.

Methods: The GSE115574 dataset was downloaded, and Cibersort was applied to estimate the relative expression of 22 kinds of immune cells. Differentially expressed genes (DEGs) were identified through the limma package in R language. Weighted gene correlation network analysis (WGCNA) was performed to cluster DEGs into different modules and explore relationships between modules and immune cell types. Functional enrichment analysis was performed on DEGs in the significant module, and hub genes were identified based on the protein-protein interaction (PPI) network. Hub genes were then verified using quantitative real-time polymerase chain reaction (qRT-PCR).

Results: A total of 2,350 DEGs were identified and clustered into eleven modules using WGCNA. The magenta module with 246 genes was identified as the key module associated with M1 macrophages with the highest correlation coefficient. Three hub genes (*CTSS*, *CSF2RB*, and *NCF2*) were identified. The results verified using three other datasets and qRT-PCR demonstrated that the expression levels of these three genes in patients with AF were significantly higher than those in patients with SR, which were consistent with the bioinformatic analysis.

Conclusion: Three novel genes identified using comprehensive bioinformatics analysis may play crucial roles in the pathophysiological mechanism in AF, which provide potential therapeutic targets and new insights into the treatment and early detection of AF.

Keywords: atrial fibrillation, WGCNA, immune cells, bioinformatics, hub genes

INTRODUCTION

Atrial fibrillation, the most common sustained arrhythmias in the clinic (1), affects ∼1 to 2% of the population (2), which increases the morbidity, mortality, and medical burden worldwide (3). The risk of stroke in patients with AF is approximately five-fold higher than that in healthy people (4). AF is a complex disease (5) associated with common risk factors, including advancing age, male sex, hypertension, obesity, and diabetes (6). However, the pathophysiological mechanism leading to AF is still unclear. Increasing evidence demonstrates that immune cells play a significant role in the pathogenesis of AF (7). Several immune-mediated serum inflammatory markers such as *CRP* and *IL-6* have been confirmed to be elevated in patients with AF (8, 9). Nevertheless, the association between immune cells and the biological molecular mechanism of AF still needs further research to clarify the pathogenesis of AF and find potential therapeutic targets.

Weighted gene correlation network analysis (WGCNA) is a system biology method used to cluster genes into modules according to expression patterns among different samples (10). Based on the interconnectivity of genes, WGCNA can explore the relationships between gene modules and the clinical phenotypes to identify candidate biomarkers or therapeutic targets. It has been applied to numerous kinds of diseases (11–14).

In this study, we aimed to explore the association between immune cells and AF using comprehensive bioinformatics analysis. Cibersort was used to evaluate immune cell composition, and WGCNA was used to identify the hub gene module. We then analyzed the functional enrichment of the genes in the hub module. Based on the protein-protein interaction (PPI) network, hub genes were identified for further analysis and validation. We hope that this research can provide potential targets and new research ideas for the treatment of AF.

METHODS AND MATERIALS

Data Acquisition and Processing

The GSE115574 dataset, including fourteen gene expression profiles from left atrial tissues of AF patients and fifteen gene expression profiles from left atrial tissues of sinus rhythm (SR) patients, were downloaded in the Gene Expression Omnibus (GEO) database. The R language was used to process the original expression profile of the GSE115574 dataset. Cibersort, a bioinformatics algorithm that could estimate 22 immune cell types, was applied to evaluate immune cell composition based on the gene expression matrix (15). The Linear Models for Microarray data (limma) package in the R language (16) was utilized to identify differentially expressed genes (DEGs) with a $p < 0.05$ between patients with AF and SR. The GSE31821, GSE41177, and GSE79768 datasets were used for validation. The batch effect was removed using the SVA package in the R language.

WGCNA

The WGCNA (10) was applied to construct the mRNA co-expression network based on the DEGs. Briefly, an appropriate soft-thresholding power β was determined to realize the scale-free topology. Then DEGs were clustered into modules, which were labeled with different colors using the average linkage hierarchical clustering method. The minimum number of genes in each module was twenty, and the threshold for module merging was 0.25. Pearson's correlation method was utilized to calculate the correlation between each module and the relative expression of immune cells identified by the Cibersort. The module with the highest correlation coefficient was selected for further analyses.

TABLE 1 | Lists of primer sequences used for quantitative real-time PCR.

Genes	Sequences
GAPDH	Forward: GGAGCGAGATCCCTCCAAAAT Reverse: GGCTGTTGTCATACTTCTCATGG
CTSS	Forward: TGACAACGGCTTTCCAGTACA Reverse: GGCAGCACGATATTTTGAGTCAT
CSF2RB	Forward: CTCCTTTGGCCTATTCTACAAGC Reverse: TGAACAGAGACGATGTATTGGC
NCF2	Forward: CCAGAAGCATTAACCGAGACAA Reverse: CCTCGAAGCTGAATCAAGGC

Functional Enrichment Analyses

The Database for Annotation, Visualization and Integrated Discovery (DAVID, v6.8) (17) was used to perform the Gene Ontology (GO) (18, 19) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses, which revealed the biological processes (BPs), cellular components (CCs), molecular functions (MFs), and pathways related to genes in the module identified above. GO terms and KEGG maps with a $p < 0.05$ were considered significant enrichment.

Construction of PPI Network and Hub Genes Identification

The DEGs in the selected module were imported into the Search Tool for the Retrieval of Interacting Genes (STRING, v11.0) (20) to generate the PPI network identifying the interactions between the genes with the threshold of interaction score >0.4. Nodes represent proteins, and edges represent protein-protein associations in the PPI network. The results downloaded from the STRING database was then visualized utilizing Cytoscape software (v3.8.1). CytoHubba (21), a plug-in of the Cytoscape software, was used to identified hub genes. The intersection of five algorithms in CytoHubba was generated for hub genes identification to ensure the accuracy and robustness of the results.

Sample Collection

Adult patients with persistent AF undergoing cardiac surgery in Zhongshan Hospital were included in this study. Persistent AF is defined as AF which is continuously sustained beyond 7 days, including episodes terminated by cardioversion (drugs or electrical cardioversion) after more than seven days (3). Excluded from the research were patients with coronary artery heart disease, hypertension, diabetes, or obstructive sleep apnea syndrome, whose ejection fractions were <30%, and those who

Input Sample	B cells naive	B cells memory	Plasma cells	T cells CD8	T cells CD4 naive	T cells CD4 memory resting	T cells CD4 memory activated	T cells follicular helper	T cells regulatory (Tregs)	T cells gamma delta	NK cells resting	NK cells activated	Monocytes	Macrophages M0	Macrophages M1	Macrophages M2	Dendritic cells resting	Dendritic cells activated	Mast cells resting	Mast cells activated	Eosinophils	Neutrophils
GSM3182680	0	0	0.132	0.093	0	0.192	0	0.03	0	0	0.04	0.046	0.014	0	0.006	0.303	0.052	0.079	0	0.001	0	0.014
GSM3182682	0	0.03	0.041	0	0.055	0.187	0	0.006	0.079	0	0.115	0	0.055	0.056	0.055	0.185	0.021	0	0.098	0	0.019	0
GSM3182684	0.006	0	0.162	0.16	0	0	0	0.077	0.025	0	0	0.053	0	0	0.02	0.341	0.036	0.061	0.032	0	0	0.028
GSM3182686	0	0.07	0.123	0.056	0	0	0	0.144	0	0	0	0.094	0	0	0.011	0.36	0	0.094	0.046	0	0	0
GSM3182688	0	0.044	0.063	0.103	0	0	0	0.12	0.017	0	0	0.065	0.036	0	0.07	0.38	0.007	0	0.096	0	0	0
GSM3182690	0	0.038	0.041	0.151	0	0	0	0.118	0.048	0.032	0	0.055	0.033	0	0.089	0.196	0.043	0	0.136	0	0.02	0
GSM3182692	0	0.095	0.05	0	0.168	0	0	0.033	0	0	0.039	0.037	0.03	0.009	0.051	0.346	0.014	0	0.091	0	0.035	0
GSM3182694	0	0.109	0.098	0.084	0	0.02	0	0.195	0.004	0	0.034	0.05	0	0	0.03	0.229	0.021	0	0.118	0	0	0.007
GSM3182696	0.013	0	0.525	0	0.13	0.111	0	0.022	0	0	0.054	0	0	0.019	0.034	0.071	0.015	0	0	0.008	0.001	0
GSM3182698	0	0	0.097	0.115	0.022	0	0	0.102	0.022	0	0.044	0.049	0	0	0.046	0.405	0.053	0	0.039	0	0	0.007
GSM3182700	0	0.071	0.066	0.061	0	0.105	0	0.091	0	0	0.024	0.071	0.003	0	0.036	0.358	0.03	0	0.084	0	0	0
GSM3182702	0	0.072	0.098	0.049	0.004	0	0	0.073	0.021	0	0.007	0.051	0	0	0.019	0.408	0.037	0.071	0.061	0	0	0.03
GSM3182704	0	0.098	0.093	0	0	0.078	0	0.156	0	0	0.007	0.084	0	0	0.06	0.331	0.066	0	0	0.018	0.009	0
GSM3182707	0	0	0.169	0.025	0	0.124	0	0.07	0	0	0	0.117	0	0	0	0.237	0.123	0.047	0.088	0	0	0
GSM3182708	0	0.023	0.112	0.157	0	0	0	0.137	0.011	0	0.039	0.029	0.012	0	0.072	0.306	0.052	0.02	0.031	0	0	0
GSM3182710	0	0.043	0.079	0.084	0	0.075	0	0.091	0	0	0.079	0	0	0	0.072	0.364	0.022	0.047	0	0.023	0	0
GSM3182712	0	0.012	0.085	0.019	0	0.168	0	0.104	0.026	0	0	0.092	0.049	0	0.018	0.342	0	0.026	0	0.058	0.002	0
GSM3182714	0	0.053	0.094	0.138	0	0.004	0	0.12	0.069	0	0.054	0.021	0	0	0.04	0.194	0	0.113	0	0.101	0	0
GSM3182716	0	0.155	0.096	0.106	0	0.106	0	0.067	0.078	0	0.065	0.025	0.013	0	0	0.152	0.034	0.066	0.017	0.01	0.01	0
GSM3182718	0	0.066	0.093	0.123	0	0.02	0	0.057	0.03	0	0	0.105	0.006	0	0.014	0.355	0.015	0.072	0.064	0	0	0.002
GSM3182720	0	0.077	0.111	0.082	0.034	0	0	0.049	0.014	0	0	0.081	0.039	0	0.029	0.254	0.077	0.017	0	0.136	0	0
GSM3182722	0	0	0.04	0.006	0	0	0	0.181	0	0.059	0	0.137	0	0	0.009	0.171	0.124	0.016	0	0.239	0	0.019
GSM3182724	0	0.034	0.068	0.115	0	0	0	0.206	0	0	0.034	0.031	0	0	0.024	0.278	0.023	0.11	0.016	0.017	0	0.044
GSM3182726	0	0	0.047	0.115	0.123	0	0	0.134	0.045	0	0	0.142	0.023	0	0	0.24	0.041	0	0.073	0	0.017	0
GSM3182728	0	0.021	0.08	0	0.047	0	0	0.115	0.006	0	0.036	0.048	0.173	0	0.041	0.249	0.015	0	0.155	0	0.014	0
GSM3182730	0	0.084	0.088	0.11	0	0.113	0	0.166	0.013	0	0.011	0.057	0.024	0	0	0.188	0.003	0.111	0.031	0	0	0
GSM3182732	0	0.093	0.093	0	0	0.17	0	0.146	0.028	0	0.048	0.035	0.008	0	0.044	0.228	0.022	0.018	0.066	0	0	0
GSM3182734	0	0	0.125	0.104	0	0.066	0	0.086	0.04	0	0.015	0.025	0.005	0	0	0.324	0.074	0.105	0.019	0	0	0
GSM3182736	0	0.108	0.11	0.168	0	0	0	0.119	0	0	0	0.097	0	0	0.026	0.249	0.021	0.059	0.006	0.037	0	0

FIGURE 1 | The relative expression of 22 immune cell subtypes in each sample estimated using Cibersort. The relative expression was higher from blue to red.

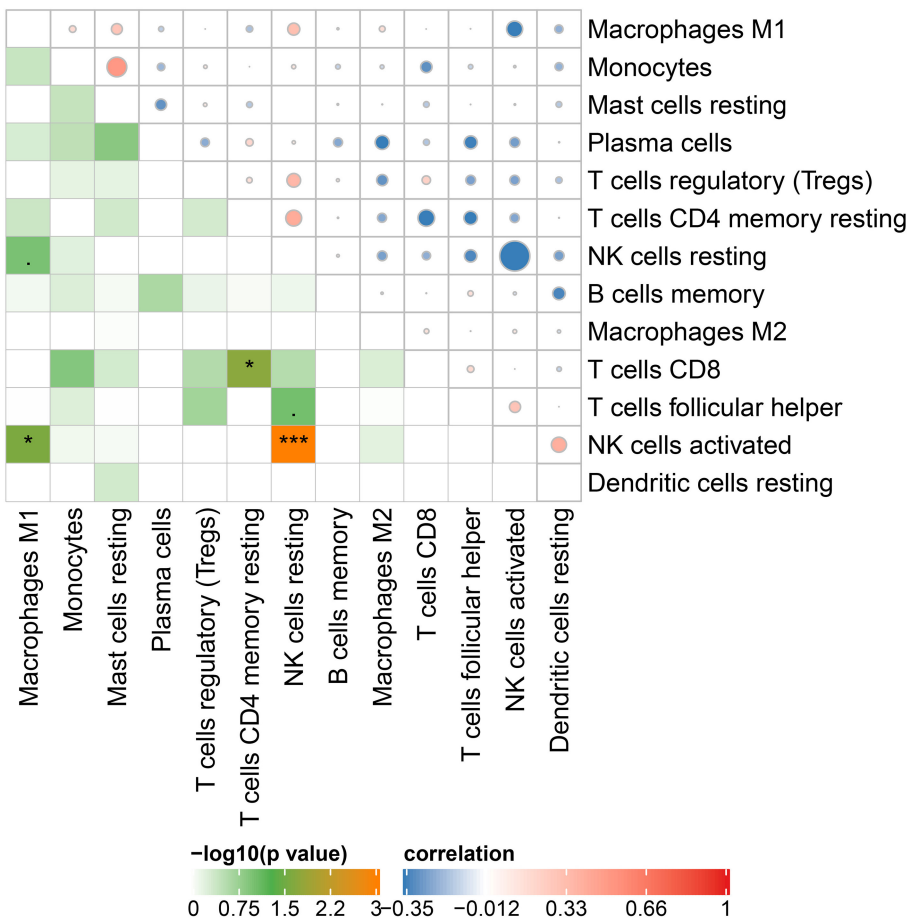

FIGURE 2 | Correlation matrix of 13 immune cell subtype compositions. Blue dots represent negative correlation, and red dots represent positive correlation. The size of the dot is positively correlated with the correlation coefficient. $^{*}p < 0.05$ and $^{***}p < 0.001$.

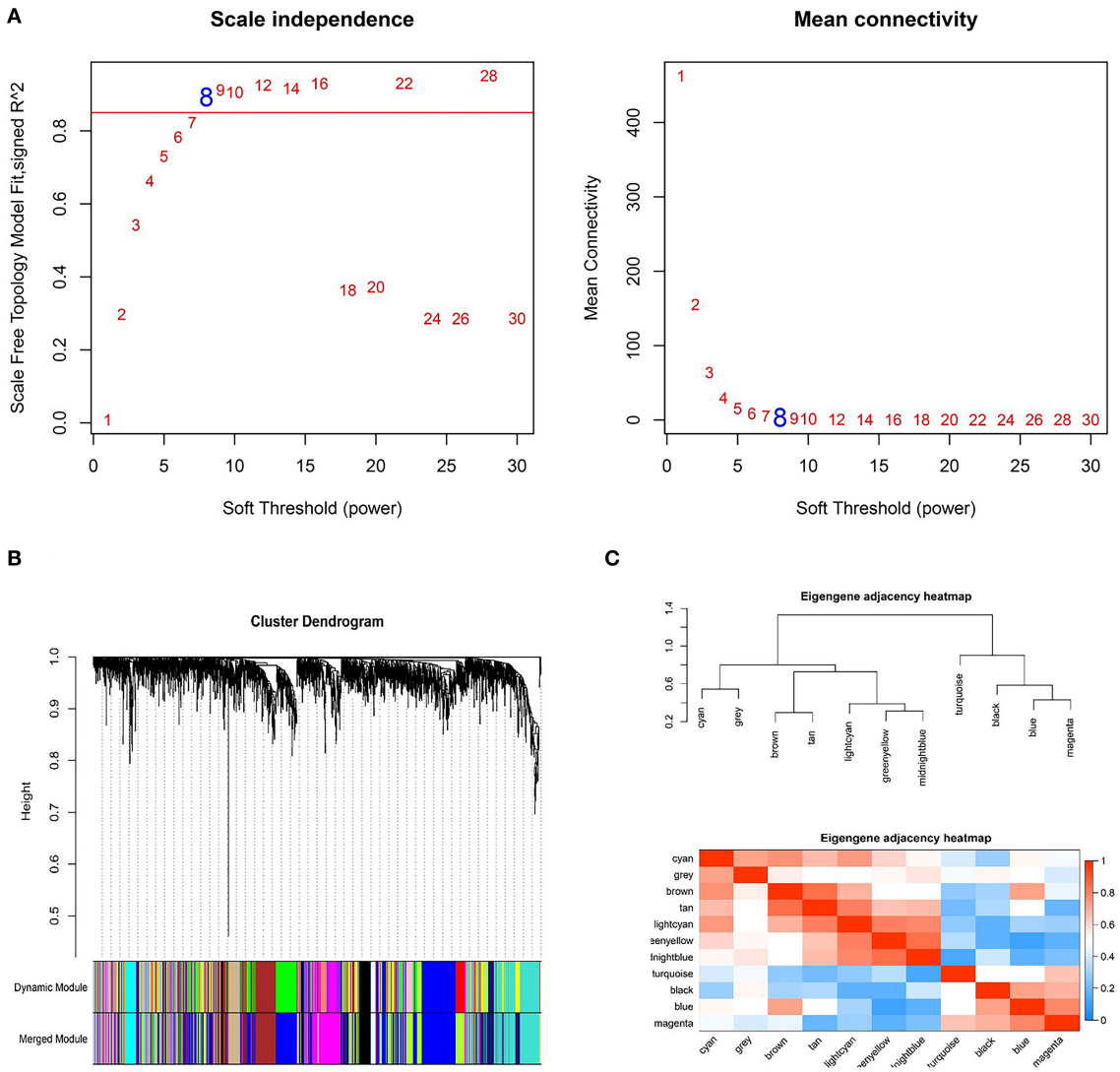

FIGURE 3 | Weighted genes correlation network analysis to cluster genes into different modules. **(A)** The selection of the soft-thresholding power β. **(B)** Dendrogram of all differentially expressed genes. **(C)** The clustering heat map between modules. Red means closer similarity, and blue means farther similarity.

had contraindications to surgery. The left atrium tissues and blood samples of twenty patients with persistent AF and ten healthy donors with SR were collected during cardiac surgery. The samples were then immediately preserved in liquid nitrogen for the later experiment. This study was in full compliance with the Declaration of Helsinki and approved by the Medical Ethics Committee of Zhongshan Hospital, Fudan University (Approval No. B2019-198R). All patients participating in this study have signed written informed consent before surgery.

Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

Total left atrium tissue RNA was extracted with the RNeasy^{TM} Mini Kit (QIAGEN, Frankfurt, Germany) following the manufacturer's instruction. The PrimeScript^{TM} RT reagent Kit (Takara, Otsu, Japan) was used to conduct reverse transcription. QRT-PCR was performed with the TB Green® Premix Ex Taq^{TM} II (Takara, Otsu, Japan) on QuantStudio^{TM} 5 System (Thermo Fisher Science, Waltham, MA, USA). The expression data was normalized by GAPDH, and the $2^{-\Delta\Delta CT}$ method was applied to analyze the results. All sequences for RNA primers (Sangon Biotech, Shanghai, China) are shown in **Table 1**.

RESULTS

Identification of DEGs

A total of 2,350 genes, including 1,115 upregulated and 1,235 downregulated, which were differentially expressed between AF samples and SR samples, were identified in the GSE115574 dataset with the limma package.

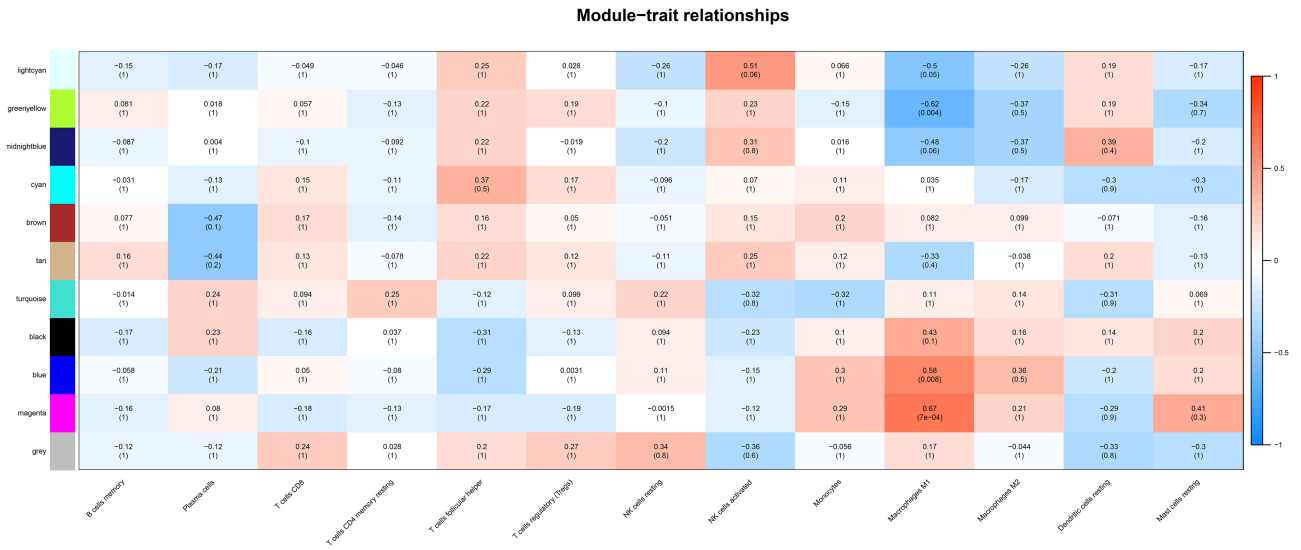

FIGURE 4 | The heatmap showing module-trait correlations. Blue represent negative correlation, and red represent positive correlation. The magenta module had the strongest correlation with M1 macrophages.

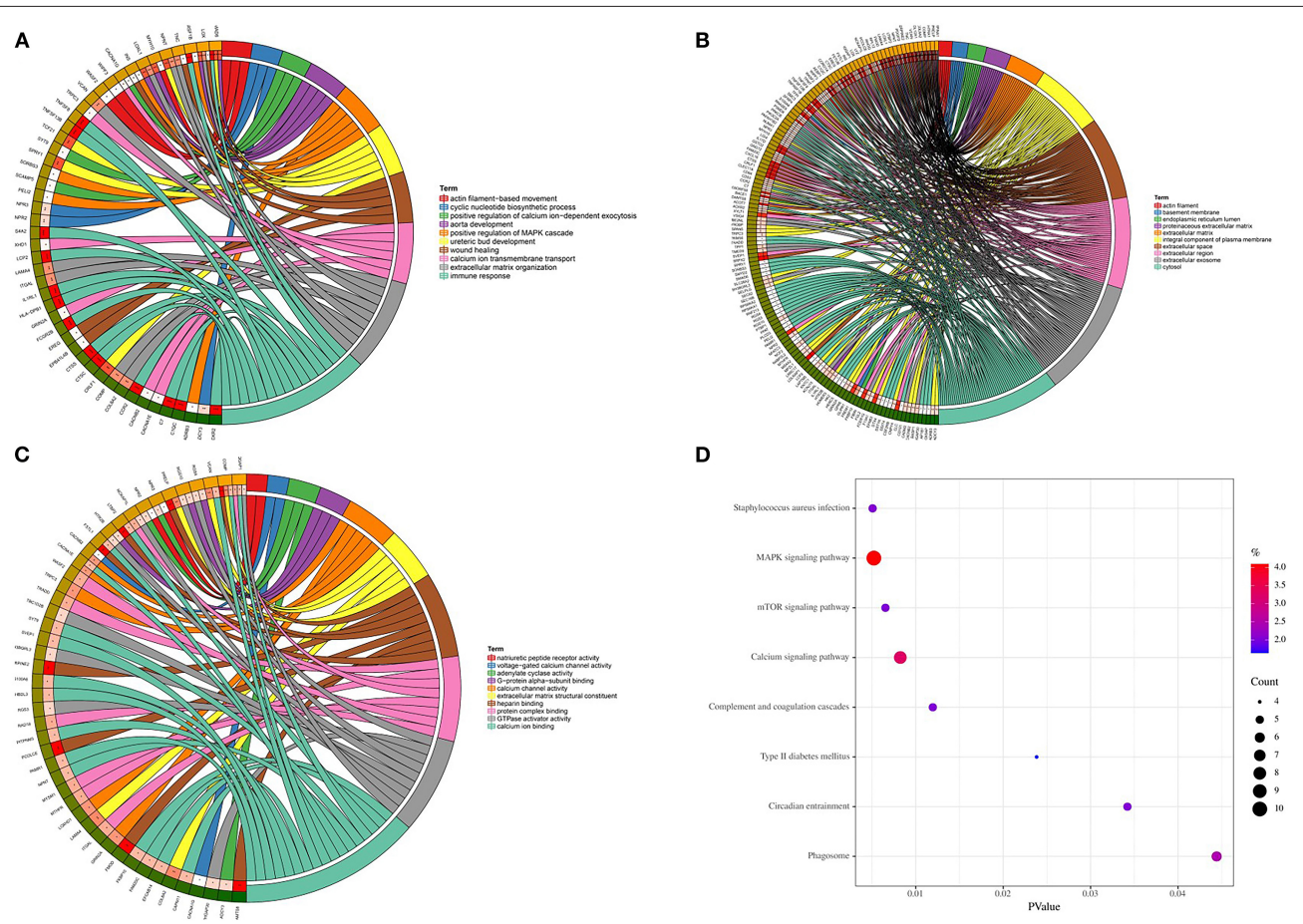

FIGURE 5 | GO and KEGG enrichment analyses. **(A)** Biological process. **(B)** Cellular component. **(C)** Molecular function. **(D)** KEGG pathways.

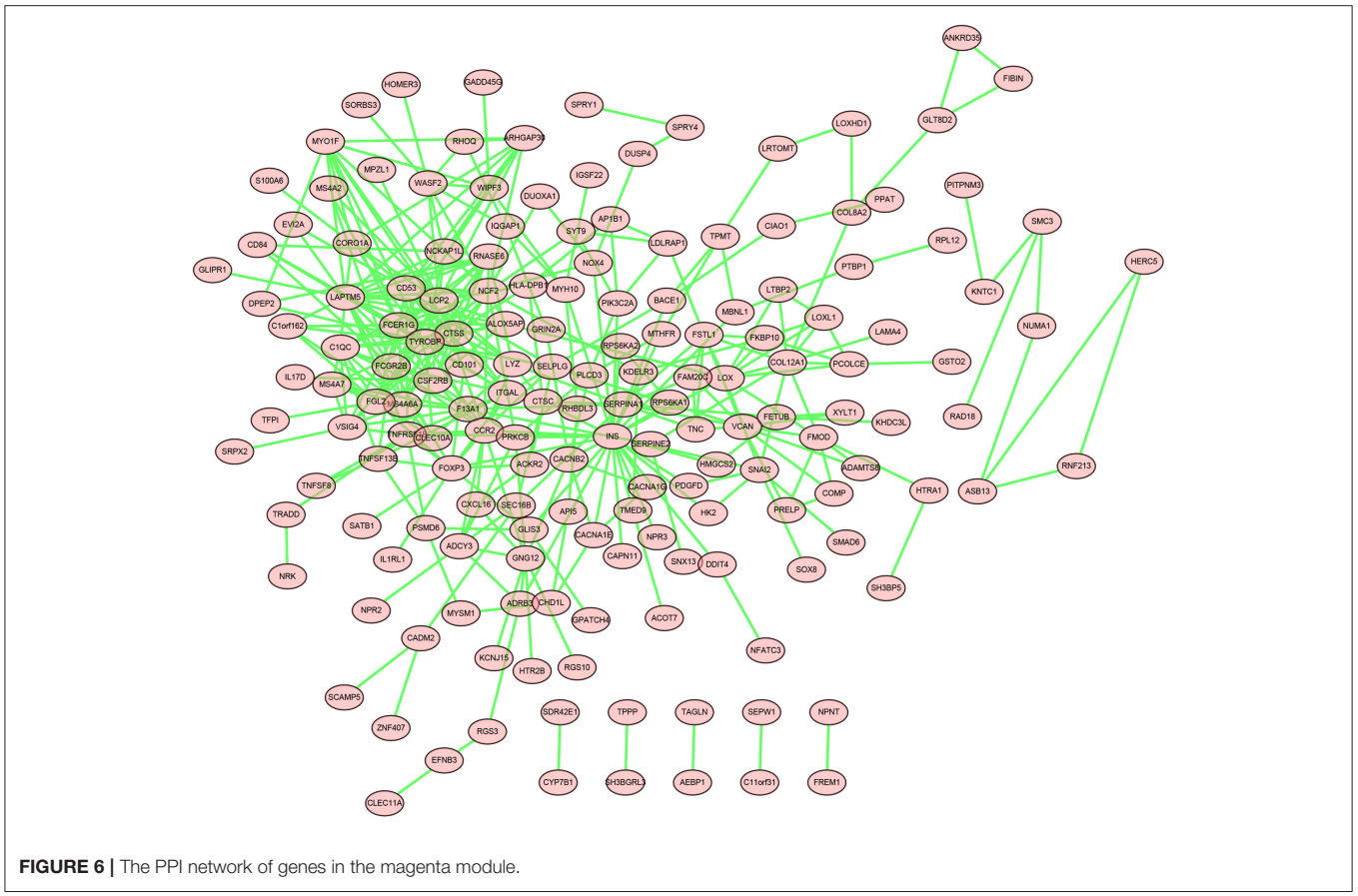

FIGURE 6 | The PPI network of genes in the magenta module.

Relative Immune Cells Expression

The Cibersort with 22 types of immune cell subtypes was applied to estimate the putative relative expression of immune cells. The results are shown in **Figure 1**. Naïve B cells, CD4 naïve T cells, CD4 memory activated T cells, gamma delta T cells, resting NK cells, M0 macrophages, activated mast cells, eosinophils, and neutrophils were eliminated because most of the samples were not inferred to express in these immune cells. The correlation of remaining immune cells was calculated using the Pearson correlation coefficient, as shown in **Figure 2**. M1 macrophages were significantly negatively correlated with activated NK cells ($r = -0.42$, $p = 0.0247$). CD4 memory resting T cells were significantly negatively correlated with CD8 T cells ($r = -0.43$, $p = 0.0191$). Resting NK cells were significantly negatively correlated with activated NK cells ($r = -0.78$, $p < 0.001$).

Construction of the Weighted Coexpression Networks

Based on the DEGs identified above, a total of 2,350 genes were subjected to WGCNA. We then established a scale-free (scale-free $R^2 > 0.85$) coexpression network with the soft-thresholding power $\beta = 8$. Using the average linkage hierarchical clustering method, DEGs were clustered into eleven modules with different colors, including black, blue, brown, cyan, greenyellow, gray, lightcyan, magenta, midnightblue, tan, and turquoise (**Figure 3**).

Correlation Between Modules and Immune Cell Types

Correlation analysis was performed between each module and immune cell types selected above, including memory B cells, plasma cells, CD8 T cells, CD4 memory T cells, follicular helper T cells, regulatory T cells (Tregs), resting NK cells, activated NK cells, monocytes, M1 macrophages, M2 macrophages, resting dendritic cells, resting mast cells. The results demonstrated that the magenta module ($r = 0.67$, $p < 0.001$) and the blue module ($r = 0.58$, $p = 0.008$) were significantly positively correlated with M1 macrophages, while the greenyellow module ($r = -0.62$, $p = 0.004$) was significantly negatively correlated with M1 macrophages, as shown in **Figure 4**. The magenta module with 246 genes was identified as the key module associated with M1 macrophages with the highest correlation coefficient.

Functional Enrichment Analyses

Genes in the magenta module were selected to perform GO and KEGG functional enrichment analyses utilizing the DAVID online tool to investigate the biological effects, as shown in **Figure 5**. The significant enriched BPs included immune response, ureteric bud development, aorta development, extracellular matrix organization, cyclic nucleotide biosynthetic process, and calcium ion transmembrane transport. In addition, extracellular matrix, extracellular space, extracellular region,

TABLE 2 | Top ten genes calculated by five algorithms of CytoHubba.

Ranks	MCC	DMNC	MNC	Degree	EPC
1	*TYROBP*	*ALOX5AP*	*TYROBP*	*INS*	*TYROBP*
2	*FCER1G*	*F13A1*	*LCP2*	*TYROBP*	*LCP2*
3	*CTSS*	*MS4A7*	*FCER1G*	*LCP2*	*LAPTM5*
4	*LCP2*	*VSIG4*	*LAPTM5*	*LAPTM5*	*FCER1G*
5	*CSF2RB*	*CTSS*	*FCGR2B*	*FCER1G*	*CSF2RB*
6	*LAPTM5*	*CORO1A*	*CD53*	*FCGR2B*	*CD53*
7	*CD53*	*ARHGAP30*	*CTSS*	*CD53*	*FCGR2B*
8	*NCF2*	*CSF2RB*	*CSF2RB*	*NCF2*	*CTSS*
9	*FCGR2B*	*NCF2*	*FGL2*	*CSF2RB*	*NCF2*
10	*ALOX5AP*	*C1QC*	*NCF2*	*CTSS*	*C1QC*

basement membrane, and actin filament were significant enriched in CC. For MF, the most significant entries were heparin binding, extracellular matrix structural constituent, calcium channel activity, protein complex binding, and calcium ion binding. Furthermore, the KEGG pathway analysis suggested that DEGs were mainly enriched in staphylococcus aureus infection, MAPK signaling pathway, mTOR signaling pathway, calcium signaling pathway, and complement and coagulation cascades.

Construction of the PPI Network and Hub Genes Identification

DEGs in the magenta module were imported into the STRING online tool to evaluate the interaction between these genes, and a total of 167 nodes and 393 edges were identified from the PPI network, as shown in **Figure 6**. Five algorithms of the CytoHubba, including MCC, DMNC, MNC, Degree, and EPC, were then applied to process the PPI network to identify the top ten genes, which are shown in **Table 2**. A Venn diagram (**Figure 7**) was generated to establish the intersection of genes identified by five algorithms, and *CTSS*, *CSF2RB*, and *NCF2* were determined as hub genes. These three genes may play considerable roles in the pathophysiology of AF.

Validation of the Hub Genes

The expression levels of three hub genes were detected in LAs and blood samples, respectively, by qRT-PCR. The results showed that the expression levels of *CTSS*, *CSF2RB*, and *NCF2* in AF were significantly higher than those in SR both in LAs and blood samples, which were consistent with the bioinformatic analysis (**Figure 8**). Moreover, we combined three datasets, GSE31821, GSE41177, and GSE79768, to verify the expression levels of these three genes between AF and SR. The results also demonstrated that the expression levels of *CTSS*, *CSF2RB*, and *NCF2* were significantly higher in AF than those in SR (**Figure 9**).

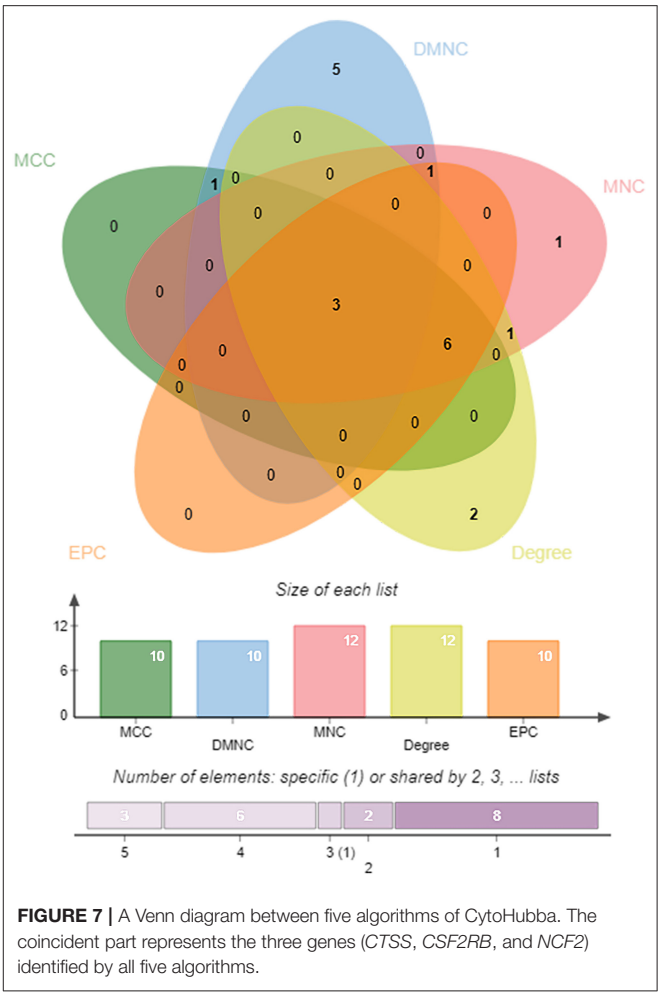

FIGURE 7 | A Venn diagram between five algorithms of CytoHubba. The coincident part represents the three genes (*CTSS*, *CSF2RB*, and *NCF2*) identified by all five algorithms.

DISCUSSION

AF is the most common tachyarrhythmia in the clinic. The typical clinical manifestations of AF are palpitations, fatigue, chest tightness, and decreased exercise tolerance, which seriously affect the life quality of patients (22). AF increases the risk of ischemic stroke by five times higher than healthy people and leads to high morbidity and mortality (23). Despite the fact that lots of efforts have been made, however, the molecular mechanism of AF development is still not completely understood. Therefore, it is significant urgent to clarify the pathogenesis of AF and find potential therapeutic targets.

In this study, we downloaded the GSE115574 dataset from the GEO database and estimated the composition of the immune cells using Cibersort based on the expression matrix. WGCNA was performed to determine the module with the most robust relationship between genes in the module and immune cell types. A total of eleven modules were identified, and the magenta module was significantly correlated with M1 macrophages. To the best of our knowledge, it is the first time that WGCNA has been used to analyze the relationships between immune cell types and AF. We then performed enrichment analysis on genes in the magenta module. According to GO analysis, genes were mainly enriched in immune response, ureteric bud development, aorta development, and extracellular matrix organization. KEGG pathway analysis demonstrated that genes were mainly enriched in staphylococcus aureus infection, MAPK signaling pathway,

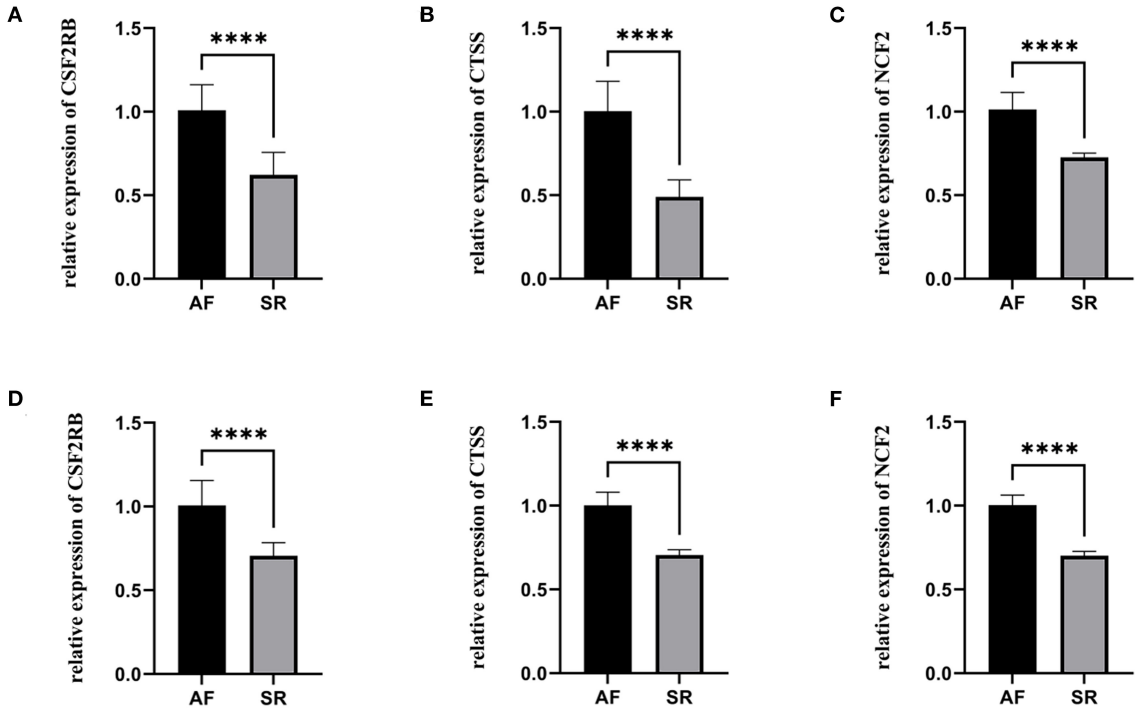

FIGURE 8 | The relative expression of three hub genes between atrial fibrillation (AF) and sinus rhythm (SR). **(A–C)** The relative expression of *CSF2RB*, *CTSS*, and *NCF2* in left atriums. **(D–F)** The relative expression of *CSF2RB*, *CTSS*, and *NCF2* in blood samples. ****$p < 0.0001$.

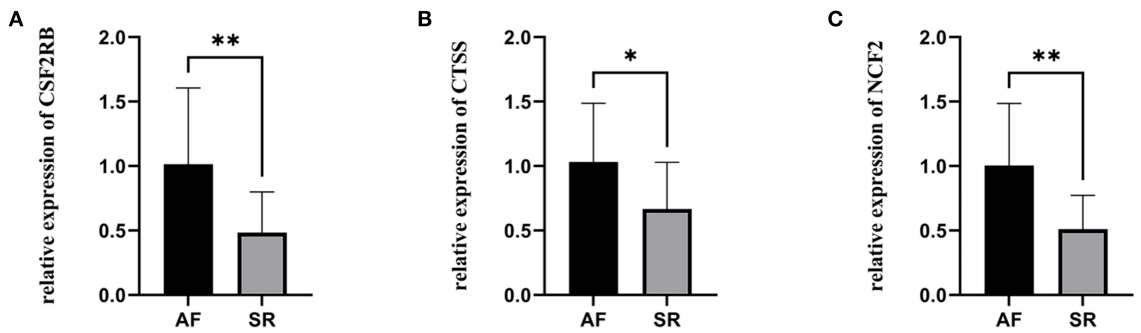

FIGURE 9 | The relative expression of three hub genes verified by three datasets. **(A)** The relative expression of *CSF2RB*. **(B)** The relative expression of *CTSS*. **(C)** The relative expression of *NCF2*. *$p < 0.05$, **$p < 0.01$.

and mTOR signaling pathway. Based on the PPI network and CytoHubba, we identified three hub genes, including *CTSS*, *CSF2RB*, and *NCF2*. As we know, the relationships between these three genes and the molecular mechanism of AF has not been studied, which is worth further research. Finally, qRT-PCR was performed to verify the expression of *CTSS*, *CSF2RB*, and *NCF2*. The expression levels of these three genes in patients with AF were significantly higher than those in patients with SR, which were consistent with the bioinformatic analysis. The expression levels of these three genes were also verified in three other datasets, GSE31821, GSE41177, and GSE79768.

Previous studies have shown that macrophages are associated with atrial fibrosis, leading to structural remodeling in the process of AF (2, 7, 24). Cytokines released by macrophages such as tumor necrosis factor-α (*TNF-α*) and interleukin-1β (*IL-1β*) can activate fibroblast proliferation, leading to fibrous tissue formation (25). However, few studies have investigated the precise role of these cytokines in the molecular mechanism of AF.

CTSS is a lysosomal cysteine proteinase, playing a crucial role in the degradation of antigenic proteins on major histocompatibility complex (MHC) class II molecules (26). *CTSS* is associated with many inflammatory and autoimmune diseases. *CTSS* expressed by intimal macrophages was involved in atherogenesis, and deficiency of *CTSS* could reduce atherosclerosis in LDL receptor-deficient mice

(27). *miR4498*/*CTSS* might polarize macrophages into pro-inflammatory phenotype and accelerate chronic atherosclerotic inflammation (28). *CTSS* participated in the abdominal aortic aneurysm (AAA) formation, and inhibition of *CTSS* suppressed AAA formation in mice (29). A previous study showed that *CTSS*-mediated induction of *CX3CL1* might contribute to the ocular surface and lacrimal glands inflammation in Sjögren's syndrome with a 4.5-fold increase in *CX3CR1*-expressing macrophages (30). In chronic obstructive pulmonary disease, reduction in *CTSS* expression prevents loss of lung function, reduces inflammation, and slows the lung tissue remodeling (31). Moreover, *CTSS* was identified as novel biomarkers for diseases and physiological processes, including triple-negative breast cancer, sarcoidosis, and particulate-induced lysosomal disruption in macrophages (32–34).

CSF2RB is the common beta chain of the high affinity receptor for interleukin-3, interleukin-5, and colony-stimulating factor. The role of *CSF2RB* has been studied in many diseases. Runt-related transcription factor 1 directly bound to the promoters of *CSF2RB*, which regulated apoptosis of neuroblastoma (35). The activating hotspot mutation in *CSFR2B* was identified in myeloid leukemia in Down syndrome (36). The mutation in *CSF2RB* can also cause hereditary pulmonary alveolar proteinosis (PAP) (37). In *CSF2RB*-/- mice, statin therapy reduces cholesterol accumulation in alveolar macrophages and ameliorates PAP (38). Moreover, *CSF2RB* was found overexpressed on monocytes from Alzheimer's disease patients, which contributed to granulocyte-macrophage colony-stimulating factor-induced monocyte migration (39). However, the role of *CSF2RB* has never been studied in AF. Further research is required to determine whether *CSF2RB* can become a novel therapeutic target for AF.

NCF2, encoding neutrophil cytosolic factor 2, mainly results in autoimmune diseases. In a previous case-control study, four single-nucleotide polymorphisms within the *NCF2* gene were genotyped, and the rs10911362 variants were associated with a decreased TB risk in the Western Chinese Han population (40). *NCF2* deficiency resulted in granulomas, and the *NCF2* mutation caused diverse and unusual clinical phenotype of chronic granulomatous disease (41, 42). In multiple sclerosis, *NCF2* was identified to be associated with eleven single nucleotide polymorphisms (43). Nevertheless, the relationship between *NCF2* and AF has not been elucidated.

There are some limitations to our study. First, the data we used was from public databases, which were limited in the sample size. Further prospective studies on more patients should be carried out to support our results. Second, although we have performed qRT-PCR to verify the expression levels of genes, mechanistic studies need to be conducted.

CONCLUSION

In this study, we performed WGCNA to analyze the relationships between immune cell types and AF for the first time. Three novel genes (*CTSS*, *CSF2RB*, and *NCF2*) which have never been studied in AF were identified. These three genes are worthy of further study and may become potential therapeutic targets in AF.

DATA AVAILABILITY STATEMENT

The dataset presented in this study can be found at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE115574, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31821, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE41177, and https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE79768.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Medical Ethics Committee of Zhongshan Hospital, Fudan University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CG and CW conceived and designed this study. TY, SZ, and MZ contributed to data analysis and prepared the main manuscript. All authors reviewed the final manuscript.

FUNDING

This study was supported by the General Program of the National Natural Science Foundation of China (No. 81770408).

ACKNOWLEDGMENTS

We thanked Dr. Zhengyang Hu and Qi Wu for their proofreading and valuable advice.

REFERENCES

1. January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, Cleveland JC, et al. AHA/ACC/HRS Focused Update of the 2014. AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American college of cardiology/American heart association task force on clinical practice guidelines and the heart rhythm society in collaboration with the society of thoracic surgeons. *Circulation.* (2019) 140:104–32. doi: 10.1161/CIR.0000000000000719
2. Andrade J, Khairy P, Dobrev D, Nattel S. The clinical profile and pathophysiology of atrial fibrillation: relationships among clinical features, epidemiology, and mechanisms. *Circ Res.* (2014) 114:1453–68. doi: 10.1161/CIRCRESAHA.114.303211
3. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomström-Lundqvist C, et al. (2020). ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association of Cardio-Thoracic Surgery (EACTS). *Eur Heart J.* (2020). doi: 10.1093/eurheartj/ehaa612. [Epub ahead of print].
4. Camm AJ, Kirchhof P, Lip GY, Schotten U, Savelieva I, Ernst S, et al. Guidelines for the management of atrial fibrillation: the task force for the management of atrial fibrillation of the European Society of Cardiology (ESC). *Eur Heart J.* (2010) 31:2369–429. doi: 10.1093/eurheartj/ehq278
5. Wyse DG, Van Gelder IC, Ellinor PT, Go AS, Kalman JM, Narayan SM, et al. Lone atrial fibrillation: does it exist? *J Am Coll Cardiol.* (2014) 63:1715–23. doi: 10.1016/j.jacc.2014.01.023

6. Andrade JG, Aguilar M, Atzema C, Bell A, Cairns JA, Cheung CC, et al. The 2020 Canadian cardiovascular society/Canadian heart rhythm society comprehensive guidelines for the management of atrial fibrillation. *Can J Cardiol.* (2020) 36:1847–948. doi: 10.1016/j.cjca.2020.09.001
7. Liu Y, Shi Q, Ma Y, Liu Q. The role of immune cells in atrial fibrillation. *J Mol Cell Cardiol.* (2018) 123:198–208. doi: 10.1016/j.yjmcc.2018.09.007
8. Hernandez MA, Moro C. Atrial fibrillation and C-reactive protein: searching for local inflammation. *J Am Coll Cardiol.* (2007) 49:1649–50. doi: 10.1016/j.jacc.2007.02.009
9. Liu L, Zheng Q, Lee J, Ma Z, Zhu Q, Wang Z. PD-1/PD-L1 expression on CD(4+) T cells and myeloid DCs correlates with the immune pathogenesis of atrial fibrillation. *J Cell Mol Med.* (2015) 19:1223–33. doi: 10.1111/jcmm.12467
10. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol.* (2005) 4:e17. doi: 10.2202/1544-6115.1128
11. Huang Q, Deng G, Wei R, Wang Q, Zou D, Wei J. Comprehensive identification of key genes involved in development of diabetes mellitus-related atherogenesis using weighted gene correlation network analysis. *Front Cardiovasc Med.* (2020) 7:580573. doi: 10.3389/fcvm.2020.580573
12. Cheng Y, Liu C, Liu Y, Su Y, Wang S, Jin L, et al. Immune microenvironment related competitive endogenous RNA network as powerful predictors for melanoma prognosis based on WGCNA analysis. *Front Oncol.* (2020) 10:577072. doi: 10.3389/fonc.2020.577072
13. Qu Y, Zhang S, Qu Y, Guo H, Wang S, Wang X, et al. Novel gene signature reveals prognostic model in acute myeloid leukemia. *Front Genet.* (2020) 11:566024. doi: 10.3389/fgene.2020.566024
14. Chen L, Shi L, Ma Y, Zheng C. Hub Genes Identification in a Murine Model of Allergic Rhinitis Based on Bioinformatics Analysis. *Front Genet.* (2020) 11:970. doi: 10.3389/fgene.2020.00970
15. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* (2015) 12:453–7. doi: 10.1038/nmeth.3337
16. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* (2015) 43:e47. doi: 10.1093/nar/gkv007
17. Dennis GJ, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* (2003) 4:P3. doi: 10.1186/gb-2003-4-5-p3
18. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still going strong. *Nucleic Acids Res.* (2019) 47:D330–8. doi: 10.1093/nar/gky1055
19. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology the gene ontology consortium. *Nat Genet.* (2000) 25:25–9. doi: 10.1038/75556
20. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* (2019) 47:D607–13. doi: 10.1093/nar/gky1131
21. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol.* (2014) 8(Suppl. 4):S11. doi: 10.1186/1752-0509-8-S4-S11
22. Piccini JP, Fauchier L. Rhythm control in atrial fibrillation. *Lancet.* (2016) 388:829–40. doi: 10.1016/S0140-6736(16)31277-6
23. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, et al. Heart disease and stroke statistics-2020 update: a report from the American heart association. *Circulation.* (2020) 141:e139–596. doi: 10.1161/CIR.0000000000000746
24. Burstein B, Nattel S. Atrial fibrosis: mechanisms and clinical relevance in atrial fibrillation. *J Am Coll Cardiol.* (2008) 51:802–9. doi: 10.1016/j.jacc.2007.09.064
25. Wynn TA, Ramalingam TR. Mechanisms of fibrosis: therapeutic translation for fibrotic disease. *Nat Med.* (2012) 18:1028–40. doi: 10.1038/nm.2807
26. Campden RI, Zhang Y. The role of lysosomal cysteine cathepsins in NLRP3 inflammasome activation. *Arch Biochem Biophys.* (2019) 670:32–42. doi: 10.1016/j.abb.2019.02.015
27. Sukhova GK, Zhang Y, Pan JH, Wada Y, Yamamoto T, Naito M, et al. Deficiency of cathepsin S reduces atherosclerosis in LDL receptor-deficient mice. *J Clin Invest.* (2003) 111:897–906. doi: 10.1172/JCI200314915
28. Li X, He X, Wang J, Wang D, Cong P, Zhu A, et al. The Regulation of Exosome-Derived miRNA on Heterogeneity of Macrophages in Atherosclerotic Plaques. *Front Immunol.* (2020) 11:2175. doi: 10.3389/fimmu.2020.02175
29. Lai CH, Chang JY, Wang KC, Lee FT, Wu HL, Cheng TL. Pharmacological inhibition of cathepsin S suppresses abdominal aortic aneurysm in Mice. *Eur J Vasc Endovasc Surg.* (2020) 59:990–9. doi: 10.1016/j.ejvs.2020.01.008
30. Fu R, Guo H, Janga S, Choi M, Klinngam W, Edman MC, et al. Cathepsin S activation contributes to elevated CX3CL1 (fractalkine) levels in tears of a Sjogren's syndrome murine model. *Sci Rep.* (2020) 10:1455. doi: 10.1038/s41598-020-58337-4
31. Doherty DF, Nath S, Poon J, Foronjy RF, Ohlmeyer M, Dabo AJ, et al. Protein phosphatase 2A reduces cigarette smoke-induced cathepsin S and loss of lung function. *Am J Respir Crit Care Med.* (2019) 200:51–62. doi: 10.1164/rccm.201808-1518OC
32. Hughes CS, Colhoun LM, Bains BK, Kilgour JD, Burden RE, Burrows JF, et al. Extracellular cathepsin S and intracellular caspase 1 activation are surrogate biomarkers of particulate-induced lysosomal disruption in macrophages. *Part Fibre Toxicol.* (2016) 13:19. doi: 10.1186/s12989-016-0129-5
33. Tanaka H, Yamaguchi E, Asai N, Yokoi T, Nishimura M, Nakao H, et al. Cathepsin S, a new serum biomarker of sarcoidosis discovered by transcriptome analysis of alveolar macrophages. *Sarcoidosis Vasc Diffuse Lung Dis.* (2019) 36:141–7. doi: 10.36141/svdld.v36i2.7620
34. Wilkinson R, Burden RE, McDowell SH, McArt DG, McQuaid S, Bingham V, et al. A novel role for cathepsin S as a potential biomarker in triple negative breast cancer. *J Oncol.* (2019) 2019:3980273. doi: 10.1155/2019/3980273
35. Hong M, He J, Li D, Chu Y, Pu J, Tong Q, et al. Runt-related transcription factor 1 promotes apoptosis and inhibits neuroblastoma progression *in vitro* and *in vivo*. *J Exp Clin Cancer Res.* (2020) 39:52. doi: 10.1186/s13046-020-01558-2
36. Labuhn M, Perkins K, Matzk S, Varghese L, Garnett C, Papaemmanuil E, et al. Mechanisms of progression of myeloid preleukemia to transformed myeloid leukemia in children with down syndrome. *Cancer Cell.* (2019) 36:123–38. doi: 10.1016/j.ccell.2019.06.007
37. Arumugam P, Suzuki T, Shima K, McCarthy C, Sallese A, Wessendarp M, et al. Long-term safety and efficacy of gene-pulmonary macrophage transplantation therapy of PAP in Csf2ra(-/-) Mice. *Mol Ther.* (2019) 27:1597–611. doi: 10.1016/j.ymthe.2019.06.010
38. McCarthy C, Lee E, Bridges JP, Sallese A, Suzuki T, Woods JC, et al. Statin as a novel pharmacotherapy of pulmonary alveolar proteinosis. *Nat Commun.* (2018) 9:3127. doi: 10.1038/s41467-018-05491-z
39. Shang S, Yang YM, Zhang H, Tian L, Jiang JS, Dong YB, et al. Intracerebral GM-CSF contributes to transendothelial monocyte migration in APP/PS1 Alzheimer's disease mice. *J Cereb Blood Flow Metab.* (2016) 36:1978–91. doi: 10.1177/0271678X16660983
40. Jiao L, Song J, Ding L, Liu T, Wu T, Zhang J, et al. A novel genetic variation in NCF2, the core component of NADPH oxidase, contributes to the susceptibility of tuberculosis in western Chinese Han population. *DNA Cell Biol.* (2020) 39:57–62. doi: 10.1089/dna.2019.5082
41. Roth IL, Salamon P, Freund T, Gadot YB, Baron S, Hershkovitz T, et al. Novel NCF2 mutation causing chronic granulomatous disease. *J Clin Immunol.* (2020) 40:977–86. doi: 10.1007/s10875-020-00820-8
42. van der Weyden L, Speak AO, Swiatkowska A, Clare S, Schejtman A, Santilli G, et al. Pulmonary metastatic colonisation and granulomas in NOX2-deficient mice. *J Pathol.* (2018) 246:300–10. doi: 10.1002/path.5140
43. Cardamone G, Paraboschi EM, Solda G, Duga S, Saarela J, Asselta R. Genetic association and altered gene expression of CYBB in multiple sclerosis patients. *Biomedicines.* (2018) 6:117. doi: 10.3390/biomedicines6040117

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

frontiers in Genetics

HYPOTHESIS AND THEORY
published: 01 March 2021
doi: 10.3389/fgene.2021.645863

RNA Biogenesis Instructs Functional Inter-Chromosomal Genome Architecture

*Alessandro Bertero**

Department of Laboratory Medicine and Pathology, Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, WA, United States

Three-dimensional (3D) genome organization has emerged as an important layer of gene regulation in development and disease. The functional properties of chromatin folding within individual chromosomes (i.e., intra-chromosomal or in *cis*) have been studied extensively. On the other hand, interactions across different chromosomes (i.e., inter-chromosomal or in *trans*) have received less attention, being often regarded as background noise or technical artifacts. This viewpoint has been challenged by emerging evidence of functional relationships between specific *trans* chromatin interactions and epigenetic control, transcription, and splicing. Therefore, it is an intriguing possibility that the key processes involved in the biogenesis of RNAs may both shape and be in turn influenced by inter-chromosomal genome architecture. Here I present the rationale behind this hypothesis, and discuss a potential experimental framework aimed at its formal testing. I present a specific example in the cardiac myocyte, a well-studied post-mitotic cell whose development and response to stress are associated with marked rearrangements of chromatin topology both in *cis* and in *trans*. I argue that RNA polymerase II clusters (i.e., transcription factories) and foci of the cardiac-specific splicing regulator RBM20 (i.e., splicing factories) exemplify the existence of *trans*-interacting chromatin domains (TIDs) with important roles in cellular homeostasis. Overall, I propose that inter-molecular 3D proximity between co-regulated nucleic acids may be a pervasive functional mechanism in biology.

Keywords: transcription, splicing, chromatin, genome organization, cardiomyocyte

OPEN ACCESS

Edited by:
Maarten M.G. van den Hoogenhof, Heidelberg University Hospital, Germany

Reviewed by:
Hiroshi Kimura, Tokyo Institute of Technology, Japan
Ralf Gilsbach, Goethe University Frankfurt, Germany

***Correspondence:**
Alessandro Bertero
abertero@uw.edu

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *24 December 2020*
Accepted: *08 February 2021*
Published: *01 March 2021*

Citation:
Bertero A (2021) RNA Biogenesis Instructs Functional Inter-Chromosomal Genome Architecture. Front. Genet. 12:645863. doi: 10.3389/fgene.2021.645863

INTRODUCTION

RNA biogenesis is a complex, multi-step process that largely takes place in the nucleus. Following transcription by RNA polymerases, most RNA species undergo extensive chemical modifications that are vital to their function. In eukaryotes, primary transcripts of messenger RNAs (pre-mRNAs) are 5'-capped, spliced, and finally 3'-polyadenylated and cleaved, altogether generating mature mRNAs. It is well established that these central steps of mRNA biogenesis occur in close proximity and at the same time and place as transcription (i.e., "co-transcriptionally"; reviewed in Perales and Bentley, 2009). Over a hundred additional RNA modifications are known (Boccaletto et al., 2018). At least some of these can be found on eukaryotic mRNAs, most notably for its abundance N^6-methyladenosine (m^6A), but also N^1-methyladenosine (m^1A),

5-methylcytosine (m^5C), and pseudouridine (reviewed in Zhao et al., 2017). Emerging evidence suggests that many of these so-called "post-transcriptional" mRNA modifications actually take place largely co-transcriptionally. For instance, the dynamic writing and erasing of m^6A is essentially complete before mRNAs are released from chromatin (Bartosovic et al., 2017; Ke et al., 2017a). This may be explained by recruitment of m^6A "writers" by DNA polymerase II (Slobodin et al., 2017), epigenetic histone marks of transcriptional elongation (Huang et al., 2019), and/or specific transcription factors (Barbieri et al., 2017; Bertero et al., 2018). In turn, m^6A deposition can affect alternative splicing and polyadenylation (Dominissini et al., 2012; Ke et al., 2015; Yue et al., 2018). Overall, mRNA biogenesis pathways are closely intertwined both spatially and functionally, and largely take place at the site of transcription.

The complexity of nuclear regulations of gene expression is staggering. Taking *Homo sapiens* as an example, there are at least 1,600 proteins likely functioning as transcription factors (Lambert et al., 2018), and over 250 splicing regulators (Jurica and Moore, 2003; Barbosa-Morais et al., 2006). Each of these molecules can affect the expression of tens to hundreds of genes with remarkable specificity despite having to locate their targets in a huge genomic and transcriptomic sampling space (~6.4 billion DNA nucleotides, up to 10% of which may be transcribed at a given time; Pertea, 2012), all the while being expressed at relatively modest copy numbers (estimated to be as low as a few thousand per cell in the case of transcription factors; Biggin, 2011). How this is all achieved in a time- and energy-effective manner remains a key open question.

An emerging topic in cell biology is that certain biochemical reactions are spatially organized in membraneless subcompartments formed by biomolecular condensates (reviewed in Banani et al., 2017). One of the first described and perhaps the best example of such structures is the nucleolus, the main site of ribosomal RNA (rRNA) biogenesis where most rDNA genes colocalize with highly concentrated factors implicated in rRNA transcription and processing (reviewed in Pederson, 2011). Many other conceptually similar nuclear structures have been described, including Cajal bodies, promyelocytic leukemia (PML) bodies, nuclear speckles, and histone locus bodies, some of which are still very poorly understood (reviewed in Mao et al., 2011).

With regards to mRNA biogenesis, over 20 years ago Peter Cook and Ana Pombo proposed that transcription may be largely spatially restricted to "transcription factories" (Cook, 1995; Pombo et al., 2000). These correspond to concentrated foci of RNA polymerase II (Pol II) and various other transcriptional regulators that collectively support highly productive transcriptional initiation and elongation. It was readily recognized that transcription factories may have a strong effect on genome organization, driving the assembly of intra- and inter-chromosomal clusters involving active loci (Sexton et al., 2007; Cook, 2010). Here, I broaden this hypothesis arguing that not just transcription but also splicing and possibly other layers of RNA biogenesis may collectively shape chromatin architecture. Specifically, I propose the existence of RNA factories-nucleated "*trans*-interacting chromatin domains" (TIDs) involving loci on different chromosomes whose nucleic acids are co-regulated by the same set of protein factors. I further propose that the resulting architecture is not just practical (i.e., it facilitates and stabilizes chromatin folding) but also functional, by which I mean that it promotes the efficacy and fidelity of RNA biogenesis.

I begin by extensively reviewing the main features of inter-chromosomal genome architecture. I follow this with a shorter discussion of the possible role of genetic variants in *trans*, and with some specific examples borrowed from cardiomyocyte biology. Having outlined the supporting evidence, I then flesh out my hypothesis and propose an experimental framework for its formal testing. I conclude by discussing open questions in the area. Throughout I focus on mammalian genome regulation and architecture since it is well established that it involves unique mechanisms not seen in other model organisms such as invertebrates, plants, and yeasts (which in turn have their own peculiarities). Nevertheless, I present selected examples from the broader spectrum of eukaryotic life, as I speculate that the general principle behind the notion of TIDs may be broadly applicable.

INTER-CHROMOSOMAL GENOME ARCHITECTURE

Over the last decade, technological advancements in the field of chromatin architecture have firmly established that mammalian interphase chromosomes are not randomly dispersed in the nucleus, but rather highly organized (reviewed in Rowley and Corces, 2018; Kempfer and Pombo, 2020). Most of our current understanding of genome topology is focused on intra-chromosomal (*cis*) interactions resulting from a variety of hierarchical features at different genomic scales. These include various kinds of DNA loops (such as promoter-enhancer pairing), topologically-associating domains (TADs; sub-megabase regions of preferential self-interaction), and A/B compartmentalization (chromosome-wide separation of active/inactive chromatin due to inter-TAD contacts). A large body of work has recently emerged to show that while a large fraction of intra-chromosomal genome structure is largely invariant across cell types and states, dynamic changes do occur at all levels and can contribute to gene regulation both during development and in disease (reviewed in Krumm and Duan, 2018; Zheng and Xie, 2019).

In contrast to this, inter-chromosomal (*trans*) interactions are much less understood and studied. The motivations are at least three-fold. First, there is a widespread belief in the community that most inter-chromosomal interactions are just noise resulting from Brownian movements of chromatin, and/or technical artifacts arising from proximity ligation-based methods (which dominate the field of chromatin conformation capture, 3C). Indeed, one commonly used quality control measure for 3C-type studies is the ratio of *cis*/*trans* interactions, which is expected to be high. While *trans* interactions are certainly more noisy than *cis* ones, as discussed below, abundant evidence indicates that there is nevertheless signal worth measuring. Which brings us to the second problem, namely the relative immaturity of statistical and computational approaches

to confidently identify *trans* interactions. Indeed, due to their different biophysical nature *trans* interactions cannot be simply treated analytically with the now well-established methods developed for *cis* ones (which for instance rely on normalizing contact counts based on the linear genomic distance to account for the varying probability of random interactions). Finally, compounding on this limitation is the fact that *trans* interactions are less frequent than *cis* ones by a factor of ~2–5 (depending on the cell type and assay), while being also spread across a much wider contact matrix 2D space (and even a broader one if we consider three-way interactions). This results in data that is very sparse unless a prohibitively expensive sequencing depth is utilized, particularly in the case of genome-wide studies. Statistical approaches to detect *trans* interactions have therefore either low power (few data points) or low resolution (genomic bins of large size are used to aggregate the sparse interaction counts). As a result of these three main factors, *trans* interactions are often dismissed early during the analysis or only used for very basic broad assessments of chromosome territories.

That all being said, some key advances have already been made in our understanding of inter-chromosomal genome architecture. These are summarized in the subsections that follow.

Chromosome Territories

Mammalian interphase chromosomes are not fully intermixed yet occupy mostly distinct nuclear domains referred to as "chromosome territories" (reviewed in Cremer and Cremer, 2010). The territorialization of chromatin was hypothesized in the late 19th and early 20th century by Carl Rabl and Theodor Boveri, and experimentally demonstrated through microscopical observations by Thomas and Christoph Cremer in the 1980s (Cremer et al., 1982). Some 30 years later, genome-wide 3C analysis (Hi-C) by Lieberman-Aiden et al. (2009) showed that the intra-chromosomal contact probability is much greater than the average contact probability between different chromosomes, consistently with the existence of chromosome territories. Chromosome territories have been observed in multiple species including yeasts (Duan et al., 2010) and plants (Dong et al., 2018), indicating that this is perhaps a universal feature of nuclear organization in multicellular organisms. Chromosome territories are not randomly positioned in the nucleus yet present preferential relative placement (Croft et al., 1999; Nagele et al., 1999; Cremer et al., 2001), a property that may partially explain the reproducible outcome of certain common chromosomal translocations (Roix et al., 2003). The distribution of chromosome territories is also somewhat tissue-specific (Parada et al., 2004; Bolzer et al., 2005), and is transmitted through mitosis (Gerlich et al., 2003). Chromosome territory dynamics may thus influence (or be influenced by) cell identity, and contribute to the epigenetic memory stabilizing a given state.

The well-established concept of chromosome territories is perhaps the motivation leading many to dismiss inter-chromosomal architecture, relegating its possible relevance to "exotic" examples such as sperm cells (which are characterized by a high frequency of both extra-long-range *cis* interactions and *trans* contacts; Ke et al., 2017b) or to non-mammalian cells (such as the KNOT structure in *Arabidopsis thaliana*, an established heat-shock responsive inter-chromosomal structure; Dong et al., 2018; Sun et al., 2020). The general skepticism about early reports of *trans* interactions frequency in Hi-C experiments as high as 60–70% was well funded: it was later shown that in-solution proximity ligation leads to many spurious ligation events, a problem that can be bypassed for instance by performing in-nucleus proximity ligation (Nagano et al., 2015). Nevertheless, even using this refined protocol the frequency of *trans* interactions remained consistently between 10 and 15% for both human and mice samples (Nagano et al., 2015). An even more rigorous assessment of multi-way chromosomal conformation through chromosomal walks (C-walks) confirmed a frequency of 7–10% inter-chromosomal interactions (Olivares-Chauvet et al., 2016). Even assuming that this is still an overestimation of *bona fide* interactions that may in truth be as limited as to ~5% of the total (and as discussed below this fraction seems to be much larger in certain cell types), this would still account for 1/20th of the data. Is it not worth to analyze it beyond just confirming the well-established concept of chromosome territories?

Nuclear Subcompartments

When the question posed above has been answered in the affirmative, important insights have been made. Lieberman-Aiden et al. (2009) leveraged their aforementioned Hi-C data to show that small, gene-rich chromosomes and large, gene-poor chromosomes form distinct clusters that preferentially interact with each other. These findings confirmed earlier analyses by DNA fluorescent *in situ* hybridization (FISH; Boyle et al., 2001; Tanabe et al., 2002). This was an important hint that *trans* interactions happening at the borders of chromosome territories are not simply random, but may for instance relate to transcriptional activity.

Building upon this concept, Rao et al. (2014) found that incorporating *trans* interactions when determining chromatin compartmentalization features allowed to go beyond the simple A/B paradigm by identifying at least six subcompartments, each bearing a distinctive pattern of epigenetic features. While the dataset in question was exceptional for its sequencing depth (4.9 billion contacts), this same feat can now be achieved from "conventionally-sized" datasets with moderate coverage using the SNIPER algorithm (Xiong and Ma, 2019). This computational approach leverages once again on inter-chromosomal interactions, imputing them based on a statistical model when the data is too sparse, and identifies subcompartments with defined epigenetic characteristics and, in some cases, strong cell specificity. This is another strong piece of evidence indicating that *trans* interactions correlate with gene expression regulation dynamics.

Kaufmann et al. (2015) performed a comprehensive analysis of inter-chromosomal interaction networks comparing Hi-C data for mouse and human embryonic stem cells (mESCs and hESCs). Remarkably, this pointed at ~70 and ~40% of the genome as being involved in at least one inter-chromosomal interaction in mESCs and hESCs, respectively, corresponding

to ~30% of genes in both species. Such contacts are not homogeneously distributed: similarly to what previously observed for *Saccharomyces cerevisiae* (Kruse et al., 2013), the authors found a high degree of clustering within the interaction networks. In hESCs, small, gene dense chromosomes dominate the network, interacting predominantly with themselves in agreement with the aforementioned findings of Lieberman-Aiden et al. (2009). Additionally, both mouse and human chromosomes interact more strongly at centromeres. Domains involved in human intra-chromosomal interactions (but not mouse ones) are enriched in active histone marks, suggesting their involvement with human-specific transcriptional dynamics. Further supporting this notion, inter-chromosomal interactions between human gene pairs significantly correlate with similarity in both function and co-expression. In this context, approximately half of all human *trans*-interacting genes are bound by the CTCF/cohesin complex, an established master regulator of both intra- and inter-chromosomal interactions (reviewed in Ong and Corces, 2014). Remarkably, human and mouse inter-chromosomal network seem to be very poorly conserved (i.e., not more significantly that would be expected at random), in sharp contrast with the strong conservation of intra-chromosomal architecture between these two species (Dixon et al., 2012). It is therefore tempting to speculate that rearrangements in inter-chromosomal architecture may have played a role in human-specific evolutionary changes, perhaps by altering gene co-regulation through novel nuclear "hubs."

Aiming to increase the signal to noise ratio of Hi-C measurements, particularly in the context of inter-chromosomal interactions, Kalhor et al. (2012) developed tethered conformation capture (TCC), which prevents Brownian interactions between crosslinked DNA complexes during proximity ligation. This approach resulted in a proportion of inter-chromosomal interactions of 25–30% of the total contacts, an approximately two-fold reduction compared to non-tethered, in solution Hi-C. To bypass the inherent inability of bulk Hi-C to detect multivalent chromatin interactions happening in individual cells, the authors also developed a modeling approach that constructs a population of three-dimensional (3D) genome structures that collectively best explain the observed interactions. Analyses of *trans* contacts revealed three major aspects: (1) each human chromosome possesses a few regions collectively explaining the majority of *trans* interactions; (2) such regions are characterized by high transcriptional activity; and (3) they interact with each other in a largely indiscriminate fashion. This led the authors to speculate that such domains intermingle chiefly because they happen to all be accessible within the nuclear interior, "escaping" the constraints of their respective chromosome compartments. However, refinements in the quantitative analyses of 3D structure population models showed that some inter-chromosomal interactions are actually relatively specific and liked to gene co-regulation (Dai et al., 2016). Indeed, the authors identified nearly four thousand "regulatory communities": clusters of chromatin domains enriched for specific regulatory factors. These include Pol II and transcription factors, RNA polymerase III (Pol III), and the polycomb binding protein YY1, suggesting that the relevant regulatory communities may represent transcription factories, transfer RNA factories, and polycomb domains, respectively. Notably, ~80% of regulatory communities contain domains from multiple chromosomes, indicating an extensive degree of inter-chromosomal architecture.

Proximity ligation-based approaches to map genome organization are designed to capture close proximity between DNA fragments. Therefore, they are relatively blind to intra- and inter-chromosomal architecture revolving around relatively large nuclear bodies, which can range from ~0.5 to ~2 μm. To bypass this limitation, Quinodoz et al. (2018) developed a ligation-independent approach called split-pool recognition of interactions by tag extension (SPRITE), which allows the identification of very large chromatin clusters with even more than a thousand DNA fragments. By applying this approach to human cells, they revealed how a large fraction of inter-chromosomal interactions revolve around two main hubs: a "nucleolar hub" enriched for peri-centrosomal, gene-poor domains stably interacting with rDNA genes and rRNA transcripts, and a "speckle hub" partial to highly transcribed, gene-rich regions dynamically coming into contact with spliceosomal RNAs and other mRNA processing factors (all characteristics of so-called "nuclear speckles"; Galganski et al., 2017). This elegant study extended earlier observation that actively transcribed, gene-dense regions can loop out from the center of chromosome territories (Mahy et al., 2002; Branco and Pombo, 2006), indicating that at least some of these converge onto nuclear speckles. Notably, similar results have been obtained with other ligation-free methods such as genome architecture mapping (GAM; Beagrie et al., 2017) and tyramide signal amplification sequencing (TSA-seq; Chen et al., 2018).

Interestingly, a subsequent study showed that inter-chromosomal interactions mapped by both TCC and SPRITE have a strong GC sequence bias (Jabbari et al., 2019), in agreement with earlier observations on Hi-C data (Yaffe and Tanay, 2011). Therefore, it is possible that aggregation into nuclear bodies may favor flexible, nucleosome poor, and loosely packed genomic regions, all characteristics of GC-rich domains, while rigid and compacted AT-rich regions are instead enriched at heterochromatic sites at the nuclear lamina (Jabbari and Bernardi, 2017).

The overall picture emerging from these seminal studies is that while chromosome territories greatly limit the possibility for inter-chromosomal interactions, they do not present hard boundaries. Among the regions able to overcome chromosome territories-mediated topological restrictions, several genomic domains (particularly flexible, GC-rich regions characterized by high transcriptional activity), can extrude to the surface and engage with each other in the proximity of transcription factories, tRNA factories, polycomb domains, the nucleolus, nuclear speckles, and/or other yet-to-be clarified nuclear subcompartments (**Figure 1**).

In the following subsections, I provide a more in-depth review of the nuclear subcompartments that are most relevant to the hypothesis of this article. For this, I focus on structures involved in the biogenesis of mRNAs. The relationship between chromatin organization and other nuclear subcompartments involved in non-coding RNA biogenesis, such as the nucleolus and tRNA

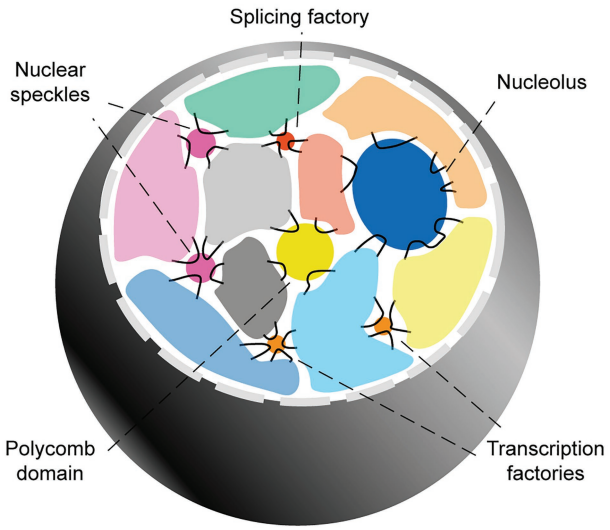

FIGURE 1 | Inter-chromosomal genome organization. Interphase mammalian nuclei are characterized by chromosome territories. Inter-chromosomal interactions mainly involve genomic domains that extrude into the inter-chromosomal space and engage with a variety of membranelles structures involved in gene regulation and RNA biogenesis.

factories, is reviewed extensively elsewhere (Pederson, 2011; Kirkland et al., 2013; Bersaglieri and Santoro, 2019).

Transcription Factories

As just discussed, highly transcribed genomic domains are those most often involved in inter-chromosomal interactions. Accordingly, inhibition of transcription reduces chromosome intermingling (Branco and Pombo, 2006). Moreover, mapping of Pol II-mediated chromatin interactions showed that these are not just intra- but also inter-chromosomal (Li et al., 2012). *Trans* interactions between active genes may be indiscriminate and dictated by the random motion of open chromatin in the euchromatic nuclear interior, or could be highly specific. A few examples of the latter category have accumulated over the years, indicating that both mechanisms likely contribute to inter-chromosomal architecture.

Following early indirect evidence and the formulation of the "transcription factory" hypothesis by Cook (1995) and Pombo et al. (2000), the laboratory of Peter Fraser was the first to formally show that transcribed genes in *cis*- and in *trans*- often colocalize at the site of transcription factories (Osborne et al., 2004). Through RNA FISH analyses the frequency of this colocalization was estimated to be 40–60% for several intra-chromosomal interactions involving the *Hbb* (beta-globin) gene, and 7% for the inter-chromosomal interaction between *Hbb* and *Hba* (alpha-globin), which is substantially higher than the expected frequency for a random interaction (~1%). In this context *cis* interactions involving *Hbb* require initiation of transcription, arguing for an active recruitment model (Mitchell and Fraser, 2008). Subsequent genome-wide studies using 3C-based approaches validated the various *cis* interactions of the *Hbb* locus (Simonis et al., 2006), and found a strong proportion of *trans* interactions for *Hbb* in the fetal liver (~40% of all contacts, a rate ~10% higher than what is seen in non-*Hbb*-expressing brain tissue; Pink et al., 2010). Enhanced chromatin immunoprecipitation followed by 3C-on-chip (e4C) indicated that a striking ~90% of all interactions involving Pol II-enriched *Hbb* or *Hba* involve genes on other chromosomes (Schoenfelder et al., 2010). While these *trans*-interacting gene networks are for the most part distinct, some genes were shown to significantly interact with both *Hbb* and *Hba* at the same transcription factory (*Slca1*, *Kel*, and *Tfrc*). Remarkably, both the recruitment of *Hbb* to transcription factories and its colocalization with its *trans* partners *Hba* and *Epb4.9* requires the erythroid transcription factor Klf1 (Schoenfelder et al., 2010). Klf1 is enriched in ~40 foci per nucleus, corresponding to only ~10–20% of transcription factories. These findings strongly argued for the existence of specialized transcription factories "bookmarked" by high levels of Klf1.

Compelling evidence for another specialized transcription factory came once again from the Fraser lab, which showed that induction of immediate early genes in B cells leads to frequent repositioning of the proto-oncogene *Myc* to the same transcription factory occupied by *Igh*, a well-known inter-chromosomal translocation partner (Osborne et al., 2007). Such association was reported at a frequency of ~25%, which is similar to that of the intra-chromosomal colocalization of *Igh* with *Fos*, and up to 10-fold higher than that for *Igh* and other genes on the same chromosome as *Myc*, overall suggesting a strong specificity. Other notable inter-chromosomal gene-gene interactions in either putative or *bona fide* transcription factories include the Oct4 and Nanog-dependent interactions between the *Nanog* gene itself and several other pluripotency factors in mESCs (de Wit et al., 2013), and the Brg1- and Stat3-dependent clustering of *Gfap* and other co-regulated genes such as *Osmr* in differentiating astrocytes (Takizawa et al., 2008; Ito et al., 2016, 2018).

One final but particularly remarkable example of specialized transcription factory is the one involving TNFα-responsive genes in endothelial cells (**Figure 2A**; Papantonis et al., 2010, 2012). It was shown that 10 min after stimulation with TNFα the promoters of two genes located ~50 mb apart on the same chromosome (*SAMD4A* and *TNFAIP2*) and a third gene on a different chromosome (*SLC6A5*) can associate as part of a NF-κB-dependent multigene cluster (Papantonis et al., 2010, 2012). This happens in ~5% of the cells, indicating that it is a relatively reproducible event (Fanucchi et al., 2013). Using this model system, Fanucchi et al. (2013) set out to test whether the formation of these intra- and inter-chromosomal contacts are required for cotranscription of the interacting genes. By perturbing, in turn, each of the three chromatin interaction sites through cleavage of the relevant genomic DNA, they uncovered important *trans*-acting effects on NF-κB-dependent gene expression. Indeed, cleavage of a single allele of *SAMD4A* resulted in virtually no transcription of both *TNFAIP2* and *SLC6A5* after TNFα stimulation. Importantly, this was observed in cells where the remaining copy of *SAMD4A* was sufficient to maintain

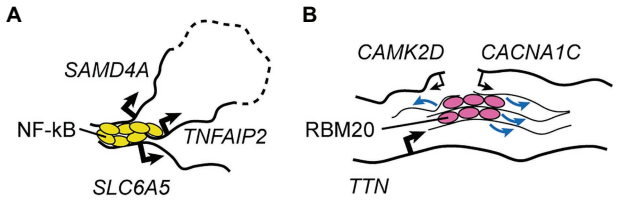

FIGURE 2 | RNA Factories. **(A)** A representative, well-understood transcription factory involved in the control of TNFα-responsive genes in endothelial cells (Papantonis et al., 2010, 2012; Fanucchi et al., 2013). Rapidly after TNFα treatment, binding of NF-κB leads to clustering of two genes located on the same chromosome but separated by ~50 Mb (dashed line), and of a third gene found on a different chromosome. This mechanism is key to transcriptional activation of all three factors in the same cell. **(B)** The muscle-specific RBM20 splicing factory (Bertero et al., 2019a). Binding of RBM20 to the many sites on its core target, the *TTN* pre-mRNA, nucleates foci that lead to clustering of other targets, promoting the regulation of their splicing. Black arrows indicate transcriptional activity (thin lines: weak; thick lines: strong), and blue arrows indicate regulation of splicing.

wild type-like levels of SAMD4A protein throughout the duration of the experiment, excluding the possibility that this factor is required for the transcription of *TNFAIP2* and *SLC6A5*. The authors further observed that not all genes in the cluster were equally important for each other's transcription: cleavage of *TNFAIP2* only reduced *SLC6A5* expression, while cleavage of *SLC6A5* did not affect the other two genes. Overall, the authors proposed a hierarchical model whereby the dominant member of the NF-κB multigene complex, *SAMD4A*, organizes transcription of subordinate genes through the establishment of both long-range intra-chromosomal interactions and inter-chromosomal contacts. To my knowledge, this is the only mechanistic study to date that demonstrated how a specialized transcriptional factory can be more than just a structural feature of the nucleus, representing a key functional entity that promotes gene co-regulation.

RNA factories may not just characterize interphase chromosomes. Peter Cook first theorized that transcription factories from homologue chromosomes play a key role in their pairing during mitosis (Cook, 1997). In this context, the formation of inter-chromosomal interactions between loci sharing the same set of protein regulators would stabilize the association of partially-condensed homologue chromosomes. Fulfilling this prophecy, recent work demonstrated that homologue chromosome pairing in yeast relies on ribonucleoprotein (RNP) complexes involving locally-transcribed, meiosis-specific long non-coding RNAs (lncRNAs; Ding et al., 2019).

Besides these examples in mammalian systems, inter-chromosomal interactions involving co-regulated genes have been observed in organism as different as yeasts (Homouz and Kudlicki, 2013), wheat (Concia et al., 2020), and sea urchins (Matsushita et al., 2017). This suggests that the formation of transcription factories that reproducibly involve the same loci (henceforward referred to as "zip coded transcription factories") is either an evolutionarily conserved mechanism or an emergent property of gene regulation across genomes of different complexity and size. However, the pervasiveness of this mechanism remains to be established.

Nuclear Speckles

Inter-chromosomal interactions between co-regulated active genes are not limited to those occurring at transcription factories. As discussed, nuclear speckles (which are functionally very distinct from Pol II clusters; Galganski et al., 2017), represent another key subcompartment enriched for *trans* interactions (Quinodoz et al., 2018). Nevertheless, only a handful of specific cases have been studied in detail to date.

One such example is the estrogen receptor α-induced interaction of *TFF1* and *GREB1* in mammary epithelial cells (Hu et al., 2008). This involves a two-step process relying first on actin/myosin1/DLC1-mediated cytoskeletal dynamics to reorganize chromosome territories, and then on the histone lysine demethylase LSD1 to induce the *trans* association of *TFF1* and *GREB1* in the context of nuclear speckles (Hu et al., 2008). Another example is the aforementioned interaction of alpha- and beta-globin genes with each other as well as with other erythrocyte genes. Challenging the notion that these often share a transcription factory, Brown et al. (2006, 2008) found that a large fraction of these inter-chromosomal associations happen around nuclear speckles.

Whether colocalization at the same nuclear speckle for co-regulated genes promotes transcriptional and/or post-transcriptional regulation has not, to my knowledge, been yet formally tested.

Splicing Factories

While much less established, the role of mRNA biogenesis steps besides transcription in inter-chromosomal regulation is beginning to emerge. Indeed, my colleagues and I recently showed that in cardiomyocytes derived from hESCs (hESC-CMs) foci of the muscle-specific splicing factor RBM20 mediate inter-chromosomal interactions between some of its key target loci (**Figure 2B**; Bertero et al., 2019a). We also mechanistically dissected the process by showing that RBM20 foci form only when scaffolded by its primary target, the pre-mRNA of the giant sarcomeric factor *TTN*, which is known to contain more than 100 RBM20 binding sites (Maatz et al., 2014). In cells lacking the *TTN* pre-mRNA due to a promoter deletion, and therefore with no RBM20 foci, the *TTN* gene no longer engages in *trans* interactions with the RBM20 target genes *CAMK2D* and *CACNA1C*. While in these conditions the expression and nuclear localization of RBM20 are unchanged, the RBM20-dependent alternative splicing events on both *CAMK2D* and *CACNA1C* is markedly impaired. Importantly, deletion of the titin protein product but not of the *TTN* pre-mRNA (through the insertion of a premature stop codon that does not lead to nonsense-mediated decay) had no effect on the formation of RBM20 foci and on the splicing of *CAMK2D* and *CACNA1C*. This excludes the possibility that titin is somehow involved in the control of RBM20 localization and/or activity, and indicates that the *TTN* pre-mRNA has specific *trans*-acting effects.

Overall, we concluded that RBM20 foci represent the first example, to the best of our knowledge, of a nuclear subcompartment specialized in the regulation of splicing for a defined subset of genes located across multiple chromosomes, in other words, a *trans*-acting "zip coded splicing factory." Whether this is an exceptional case or the first of many examples of this type of regulation is yet to be determined.

Trans-Acting Regulatory Regions

On top of engagement of co-regulated genes in transcription and/or RNA processing factories, there are several instances of specific regulatory *trans* interactions involving a single gene and one or more regulatory sequences that either repress or promote its expression. Repressive regulations include the inter-chromosomal interaction between the T helper cell 2 locus control region (TH2 LCR) and the *IFN-γ* gene (Spilianakis et al., 2005). This interaction is very strong in naïve CD4 positive T cells, and is relieved after specification of either TH2 or TH1 cells, the latter leading to a novel intra-chromosomal interaction of *IFN-γ* that promotes its expression. Thus, the TH2 LCR is believed to create a repressive-yet-poised chromatin hub key to the rapid activation of one of the two T helper specification programs. Initiation of mesendoderm differentiation is similarly regulated by an Oct4-dependent switch of the *Sox17* enhancer from a locked inter-chromosomal conformation (involving the *Sox2* gene) to an active intra-chromosomal engagement with *Sox17* (Abboud et al., 2015). Myogenesis is also temporally regulated by the inter-chromosomal sequestration of key regulatory regions for late muscle genes, which relies on the master transcription factor MyoD1 and is required to prevent their premature activation (Harada et al., 2015). Another interesting example is the interaction between the well-studied Igf2/H19 imprinting control region (ICR) and over 100 sequences from all autosomes, leading to epigenetic regulation in *trans* (Zhao et al., 2006). One such interaction involving the maternal ICR and the mouse Wsb1/Nsf1 Locus relies on CTCF for its stabilization, and represses Wsb1/Nsf1 expression (Ling et al., 2006). In porcine cells, the IGF2/H19 ICR is found in the proximity of other imprinted genes such as *DLK1* and *MEG3* (Lahbib-Mansais et al., 2016), in agreement with the strong overrepresentation of imprinted domains in the genome-wide interactions involving this ICR in the mouse (Zhao et al., 2006).

Besides these repressive *trans* chromatin structures, there are multiple examples of inter-chromosomal regulations that promote gene expression. Activation of the *IFN-β* gene relies on *trans* interactions with up to three genomic regions that bring limiting amounts of the key viral-induced transcription factor NF-κB onto the *IFN-β* enhancer, allowing the formation of an enhanceosome (Apostolou and Thanos, 2008). Similar inter-chromosomal enhancers promote expression of *Pax5* specifically in B cells (Fujita et al., 2017), and of *Tead4* in trophoblast stem cells (Tomikawa et al., 2020). One final example is perhaps the most extraordinary *trans* chromatin regulation described to date, namely the formation of the mouse olfactory receptor multi-chromosomal super-enhancer (Lomvardas et al., 2006; Markenscoff-Papadimitriou et al., 2014; Monahan et al., 2017). This structure, which includes 63 enhancers from 18 chromosomes (named Greek islands), is key to remove heterochromatin marks from a single olfactory receptor chosen stochastically among more than 1,000 such genes dispersed in various heterochromatin clusters. Formation and maintenance of this remarkable inter-chromosomal enhancer cluster requires the active involvement of the Greek islands-bound transcription factor Lhx2 and of its adaptor protein Ldb1 (Monahan et al., 2019).

Overall, it is now clear that inter-chromosomal chromatin architecture is not just restricted to active genes in more or less defined nuclear subcompartments or neighborhoods, but can extend to include complex multifactorial regulatory interactions generally mediated and stabilized by specific protein co-factors.

Polycomb Domains

A specific subtype of *trans*-interacting regulatory structure is presented by regions brought in proximity by the Polycomb protein complex (PcG), an important regulator of gene silencing through the deposition of the repressive histone mark histone 3 lysine 27 trimethylation (H3K27me3). PcG factors such as the catalytic subunit EZH2 are involved in both intra- and inter-chromosomal interactions (Tiwari et al., 2008). For instance, in mESCs primed for differentiation the homeobox gene clusters (Hox; key developmental regulators of antero-posterior patterning) become bivalently marked by H3K4me3 and H3K27me3, corresponding to the onset of specific long-range intra- and inter-chromosomal interactions (Joshi et al., 2015). Notably, these Hox "polycomb domains" are localized away from the nuclear lamina and within the A compartment (Vieux-Rochas et al., 2015), suggesting that the formation of a specific bivalently-marked chromatin sub-microenvironment within active chromatin may be key to maintain selective gene silencing while allowing for rapid activation of Hox factors upon developmental cues.

TRANS-ACTING EXPRESSION QUANTITATIVE TRAIT LOCI

Genetic variants associated with disease phenotypes are overwhelmingly non-coding, with <5% of trait-related single-nucleotide polymorphisms (SNPs) identified by genome-wide association studies (GWAS) representing non-synonymous substitutions in protein-coding genes (Eicher et al., 2015). It is generally thought that a large fraction of disease-associated non-coding variants may be explained by alterations in inter-chromosomal regulations such as promoter-enhancer interactions (Dekker and Mirny, 2016; Delaneau et al., 2019), possibly due to altered binding of transcription factors (Wong et al., 2017). Indeed, GWAS SNPs are very often associated with the expression levels of nearby genes, referred to as *cis* expression quantitative trait loci (*cis* eQTLs; Michaelson et al., 2009). However, eQTL analyses have also revealed a growing number of genetic variants that affect gene targets (eGenes) either far away on the same chromosome (separated by at

1–5 Mb, depending on the study definition) or located on a different chromosome altogether (Heinig et al., 2010; Fehrmann et al., 2011; Innocenti et al., 2011; Fairfax et al., 2012; Grundberg et al., 2012; Liang et al., 2013; Battle and Montgomery, 2014; Kirsten et al., 2015). Such variants are therefore referred to as *trans* eQTL. These are difficult to identify with statistical confidence both because of the large multiple testing burden involved and due to a generally weaker effect on gene expression compared to *cis* eQTLs (McKenzie et al., 2014). Nevertheless, an eQTL meta-analysis including 5,311 patients was able to identify 1,513 significant *trans* eQTLs involving 346 SNPs and 430 genes (Westra et al., 2013). An even larger analysis of a single cohort of 6,111 individuals found 5,749 lead *trans* eQTLs affecting 4,958 genes (Joehanes et al., 2017).

The mechanism of action of *trans* eQTLs remains unclear, but three major possibilities have been proposed so far (**Figure 3A**). First, the gene variant may directly affect the expression of the distal gene (*trans* eGene), perhaps as a result of an inter-chromosomal regulatory interaction such as those described in the above section "*Trans*-Acting Regulatory Regions". Secondly, the gene variant may act indirectly through a third actor: a proximal gene (*cis* eGene) whose expression is directly regulated by the SNP and that in turn can directly modulate expression of the *trans* eGene. Proposed mechanisms behind the mediating action of such a *cis* eGene include its function as a transcription factor, micro RNA, or another regulatory non-coding RNA (Joehanes et al., 2017). Finally, the gene variant may be regulating the *trans* eGene even more indirectly as a result of "reverse causality": a feedback mechanism arising from a diseased state caused directly by the SNP (for instance in the case of non-synonymous substitutions) or through modulation of a *cis* eGene. Indirect regulation or reverse causality can be tested through mediation analysis, which assesses whether a third factor (either a *cis* eGene or a phenotype) significantly accounts for the observed *trans* eQTL-*trans* eGene relationship. Using this approach it was found that 20–35% of *trans* eQTLs can be explained by the alteration of a mediating *cis* eGene (Pierce et al., 2014; Yao et al., 2017). Notably, over 80% of *trans* eQTLs also have an associated *cis* eGene (Yao et al., 2017), indicating that the proportion of "indirect *trans* eQTLs" may be even larger than what could be assessed with confidence by mediation analysis. The comprehensive genotype-tissue expression (GTEx) consortium reported similar observations, as 31.6% of lead *trans* eQTL identified across 27 tissues proved to be also *cis* eQTL, and 77% of these were predicted to contribute to the *trans* effect by mediation analysis (Aguet et al., 2020). Interestingly, Yao et al. (2017) found no evidence for the possibility of reverse causality as an underlying mechanism explaining *trans* eQTLs, suggesting that if this mechanism does contribute it is not a common and/or a strong factor.

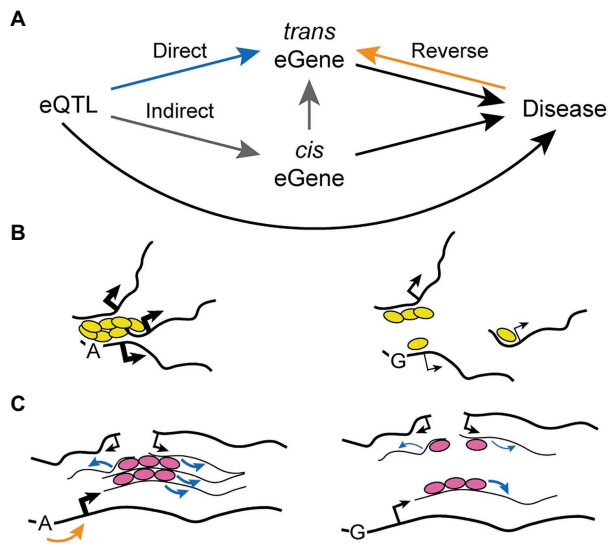

FIGURE 3 | Mechanism of Action of *Trans*-Acting Expression Quantitative Loci. **(A)** The connection between a *trans* eQTL and its *trans* eGene can be mediated by three possibilities: direct regulation (blue path), indirect nuclear regulation through an intermediate *cis* eGene (gray path), or reverse regulation from a diseased state (black plus orange pathways). **(B,C)** Indirect regulation of *trans* eGenes by *trans* eQTL may in some cases be mediated by disruption of transcription factories **(B)** or splicing factories **(C)**. An allelic variant may decrease binding of *trans*-acting regulators (i.e., transcription factors, yellow ovals, or splicing factors, pink ovals) to core nucleic acids, disrupting their clustering and the resulting formation of *trans*-interacting chromatin domains and RNA factories (see **Figure 4A**). Black and blue arrows indicate transcriptional and splicing activity, respectively (thin lines: weak; thick lines: strong). The orange arrow indicates regulation of transcriptional initiation.

In the study of Yao et al. (2017) some *trans* eQTLs were found to significantly regulate numerous genes and were thus called "*trans* eQTL hotspots." These were also notable because the resulting gene regulation had a clear directional bias, with 65% of *trans* eGenes for a given *trans* eQTL hotspot being all up- or-down-regulated. This may be explained by the indirect alteration of a *cis* eGene with function of transcription factor activating or repressing a gene network containing the *trans* eGenes. For instance, this kind of mechanism was proposed for a *trans* eQTL which controls the expression of IFN-α-responsive genes through indirect alteration of the transcription factor *IKZF1* (Westra et al., 2013). However, Yao et al. (2017) found no significant enrichment for transcription factors in the *cis* eGenes linked to *trans* eQTL hotspots (only 2 out of 37). On the other hand, the authors found that approximately one third (13/37) of these *cis* eGenes shared the same regulatory motif with the *trans* eGenes. A similar enrichment of transcription factor motifs in the promoter of *trans* eGenes linked to *trans* eQTL has also been reported elsewhere (Joehanes et al., 2017). Thus, it is an intriguing possibility that a substantial fraction of *trans* eQTLs may act through an alternative mechanism that does not involve the protein product of a *cis* eGene, but rather shared regulatory mechanisms at the level of mRNA biogenesis. For instance, disruption of RNA factories due to impaired binding of *trans*-regulatory factors and/or to reduced abundance of core nucleic acids may impair *trans*-acting transcriptional and/or splicing defects (**Figures 3B,C**; refer to the model discussed in the below section "Mechanisms

Leading to the Formation of RNA Factories"). Of note, a similar mechanism could in principle be involved in the action of *trans* eQTLs involving *cis* eGenes located far away on the same chromosome.

A recent study found evidence that human inter-individual correlation of chromatin regulatory activity driven by genetic variation occurs both intra- and inter-chromosomally, leading to the formation of *trans*-regulatory hubs, TRHs (Delaneau et al., 2019). The authors found several lines of evidence suggesting a link between this kind of regulation and inter-chromosomal genome architecture: (1) correlation in chromatin activity is enriched for genomic domains involved in inter-chromosomal interactions; (2) TRHs belong to the A or B compartment; (3) *trans* allelic correlation was observed in TRHs; and (4) some TRHs mediate *trans* eQTL effects. Along these lines, *trans* eQTLs identified by the GTEx consortium are specifically enriched for CTCF binding sites (Aguet et al., 2020). Overall, these findings add further evidence to the possibility that genetic variation may influence inter-chromosomal genome architecture through modulation of RNA biogenesis in *trans*.

THE CARDIOMYOCYTE NUCLEUS: A CASE IN POINT

While in the previous sections I introduced evidence for the interplay between RNA biogenesis and inter-chromosomal genome organization across many different biological systems, here I provide a cohesive case in point for a specific model: the mammalian cardiac myocyte. This is chosen for several reasons. Cardiomyocytes are highly important from a medical perspective, as cardiovascular disease remains the number one killer worldwide (Virani et al., 2020). As a result, developmental and disease-associated mechanisms controlling cardiomyocyte gene expression have been extensively studied for several decades (reviewed in Chang and Bruneau, 2012; Akerberg and Pu, 2019). Of particular relevance to this manuscript, cardiomyocyte 3D nuclear organization has been the focus of numerous recent investigations, an aspect my colleague and I recently reviewed (Bertero and Rosa-Garrido, 2020). Moreover, being a highly specialized, postmitotic cell type, the cardiomyocyte has a relatively stable 3D chromatin architecture that does not undergo the substantial rearrangements associated to the cell cycle (Nagano et al., 2013; Ramani et al., 2017), thus representing an ideal model to test the hypothesis outlined in this manuscript. On a practical level, the development of robust protocols to differentiate and mature cardiomyocytes from hESCs and induced pluripotent stem cells (iPSCs) has greatly simplified the study of cardiac biology, making this a widely accessible model (reviewed in Mummery et al., 2012; Protze et al., 2019; Karbassi et al., 2020).

Cardiac Inter-Chromosomal Genome Organization

My colleagues and I found that differentiation of hESCs into cardiomyocytes leads to an almost two-fold increase of inter-chromosomal interactions, which represent ~50% of all interactions in hESC- and hiPSC-derived cardiomyocytes (Bertero et al., 2019a). These results were obtained using *in situ* DNase Hi-C, a genome-wide 3C approach that has less sequence bias and higher effective resolution than conventional, restriction enzyme-based Hi-C (Ramani et al., 2016). By detecting chromatin interactions within fixed and minimally altered nuclei, this method also minimizes the chances of detecting spurious inter-chromosomal interactions resulting from intermixing of chromatin (a problem characteristic of in-solution proximity ligation-based approaches, as discussed in the above section "Nuclear Subcompartments"; Nagano et al., 2015). Notably, the proportion of *trans* interactions in hESC-CMs remained consistently high over a broad range of fixation conditions (1–4% PFA), suggesting that *trans* contacts were not inflated by suboptimal capture of the native chromatin environment. A high proportion of *trans* interactions in cardiomyocytes was observed also in other *in situ* Hi-C experiments in cardiomyocytes from other hESC lines (~46%; Zhang et al., 2019), hiPSC lines (~60%; Bertero et al., 2019b), primary mouse cardiomyocytes (~39%; Rosa-Garrido et al., 2017), and fetal human hearts (~45%; Bertero et al., 2019a). Overall, cardiomyocytes appear to have marked inter-chromosomal genome architecture. This could be connected to their postmitotic state, which may facilitate the acquisition of a stable inter-chromosomal organization that is no longer periodically perturbed by the cell cycle. Whether the formation inter-chromosomal interactions follows changes in intra-chromosomal genome architecture or *vice versa* has not been yet clarified: establishing this will require sampling of differentiating cardiomyocytes at a much finer interval than what has been done so far.

Similarly to other cell types, cardiomyocytes present defined chromosome territories whereby small, gene-rich chromosomes are in closer proximity to each other than with large, gene-poor chromosomes (Bertero et al., 2019a,b). Accordingly, inter-chromosomal interactions are enriched for domains within the A compartment, and depleted for B compartment regions (Bertero et al., 2019a,b). Significant inter-chromosomal interactions are also markedly and specifically enriched for cardiac-specific genes (Chapski et al., 2019), suggesting that some *trans* interactions may involve the formation of functional chromatin subcompartments such as those described in the above section "Nuclear Subcompartments".

One such example has already been presented in the above section "Splicing Factories", namely the muscle-specific RBM20 "splicing factory" (**Figure 2B**), which is nucleated by its key target, the *TTN* pre-mRNA, and promotes alternative splicing of other targets on different chromosomes as they come in proximity with the *TTN* locus (Bertero et al., 2019a). Here it is worth adding that this structure is not just notable from a mechanistic standpoint, but may also play an important role in the regulation of cardiac function. Indeed, mutations in RBM20 are an established cause of dilated cardiomyopathy (DCM), a condition that leads to progressive impairment of cardiac output and ultimately to heart failure (Brauch et al., 2009). RBM20 mutations affect ~3% of DCM patients and lead to a particularly malignant form of the disease characterized

also by conduction system disorders and/or life-threatening arrhythmias (in ~30 and ~45% of patients, respectively; Refaat et al., 2012; Haas et al., 2015; Kayvanpour et al., 2017; van den Hoogenhof et al., 2018). This is associated to a faster progression to heart failure and to the need of cardiac transplantation at a younger age than other DCM patients (Kayvanpour et al., 2017). Contractile dysfunction has been attributed primarily to splicing alterations in *TTN* (Guo et al., 2012; Maatz et al., 2014). Indeed, mutations in RBM20 lead to a longer, more compliant isoform of titin, a protein that plays a key role in maintaining cardiac stiffness by working like a molecular spring (Herman et al., 2012). The mechanism leading to electrical disturbances in the heart of RBM20 DCM patients is less understood, but may rely on other splicing alterations in ion-handling genes (Wyles et al., 2016a,b; van den Hoogenhof et al., 2018). In this context, a recent study by Schneider et al. (2020) found that a mutation of RBM20 in an arginine- and serine-rich hotspot leads to impaired nuclear localization and accumulation of mutant RBM20 in cytoplasmic RNP granules. Overall, it is tempting to speculate that maintenance of efficient RBM20 splicing factories is key to maintain cardiac homeostasis.

The cardiomyocyte nucleus is also characterized by numerous transcription factories. Analyses in mouse cardiomyocytes revealed that these respond dynamically to stress induced by humoral or mechanical stimuli (Karbassi et al., 2019). Moreover, this response is distinct in neonatal and adult cardiomyocytes: the former increase the number of actively elongating Pol II clusters, while the latter increase the size of existing ones. These transcription factory dynamics parallel the repositioning of cardiac genes that are up- or down-regulated in response to stress, which move closer to or away from the nuclear interior, respectively. Notably, promoter enrichment of Pol II binding is higher for loci closer to the interior. Moreover, integration of Hi-C data showed that genes upregulated after stress have a greater fraction of interaction with each other, and *vice versa* for downregulated genes (Rosa-Garrido et al., 2017; Karbassi et al., 2019). This collective evidence points towards the conclusion that the cardiomyocyte nucleus possesses transcription factories that are dynamically modulated in response to stress. Whether "zip coded transcription factories" (as defined in the above section "Transcription Factories") are a common mechanism in this context remains, however, unclear.

It is important to mention that cardiomyocytes are well known as one of the few cell types that become polyploid in physiological conditions. Such process is a hallmark of postnatal maturation following withdrawal from the cell cycle (reviewed in Marchianò et al., 2019). This process is species-specific: while ~90% of rodent adult cardiomyocytes are tetraploid (4N) through bi-nucleation, with each nucleus being diploid (2N), ~75% of human adult cardiomyocytes remain mononucleated but grow in ploidy to 4N or, more rarely, 8N and even 16N. Stress can further augment human cardiomyocyte ploidy, particularly the proportion of 8N and 16N nuclei (Gilsbach et al., 2018). Therefore, it is conceivable that both developmental- and disease-associated ploidy changes may impose substantial changes on inter-chromosomal nuclear architecture. This aspect is, however, not straightforward to study using hESC- or hiPSC-CMs since these largely immature cells (resembling fetal or early postnatal developmental stages) are for the most part diploid (reviewed in Karbassi et al., 2020). Moreover, the aforementioned differences in rodent and human ploidy limits the predictive power of experiments in mouse models. Studies of ploidy in primary human cardiomyocytes have begun to emerge (Hesse et al., 2021), and may be the best approach to elucidate its impact on inter-chromosomal genome architecture.

Cardiac *Trans*-Acting Expression Quantitative Trait Loci

As discussed in the above section "*Trans*-Acting Expression Quantitative Trait Loci", the study of *trans* eQTLs requires a large sample size. Therefore, the identification of *trans* eQTLs in human cardiomyocytes has proved challenging due to the complexity associated to acquiring the necessary material from living individuals. Accordingly, the GTEx consortium only found a single *trans* eQTL from the analysis of 386 left ventricle samples (Aguet et al., 2020). On the other hand, the study of recombinant inbred rodent strains has provided valuable insights. For instance, analysis of 24 recombinant inbred mouse strains identified 1,357 *trans* eQTLs, as well as three clusters of *trans* eQTLs hotspots each regulating over 50 genes (Imholte et al., 2013). A similar analysis of a panel of 29 recombinant inbred rat strains led to the identification of 2,140 *trans* eQTLs for left ventricle tissue samples (Grieve et al., 2008). Notably, left ventricle *trans* eQTLs outnumbered by more than two-fold those found in the same study for fat, kidney, and adrenal tissue samples. The number of *trans* eQTL clusters was similarly highest for the left ventricle. Notably, left ventricle *trans* eQTL showed a larger degree of correlation in the expression of matched *trans* eGenes compared to that observed for *cis* eQTL-*cis* eGene pairs (3.4 vs. 0.6%, respectively). Such correlation was even more remarkable for left ventricle *trans* eQTL gene clusters, reaching 77.2%. Therefore, cardiomyocyte gene expression seems to be particularly affected by genetic variation in *trans*, providing further evidence for the possible role of inter-chromosomal genome organization in this cell type.

A BIDIRECTIONAL LINK BETWEEN RNA BIOGENESIS AND NUCLEAR REGULATIONS IN *TRANS*?

Having introduced the key relevant observations in the previous sections, here I present and elaborate on the central thesis of this manuscript, namely that RNA biogenesis and regulation of inter-chromosomal genome organization are closely linked in a bidirectional fashion. As introduced earlier, this is by no means a novel idea *per se*. In fact, the contribution of transcription factories to 3D genome topology has been hypothesized at least a decade ago (Sexton et al., 2007; Cook, 2010). This concept has been extended by recent

studies that coined terms such as "nuclear speckle hubs" or "regulatory communities" to describe emerging inter-chromosomal structures in various models (Dai et al., 2016; Quinodoz et al., 2018). In this context, my goal here is three-fold: (1) to generalize the "transcription factory hypothesis" of Cook, Pombo, Fraser, and others by also incorporating other RNA biogenesis steps beyond transcription; (2) to provide a partially distinct mechanistic framework; (3) and to outline the resulting predictions, to be used to test the overall hypothesis.

A Unifying Hypothesis for RNA Factories and *Trans*-Interacting Domains

This hypothesis can be articulated in three central statements:

I. All key steps involved in RNA biogenesis have the *potential* to contribute to inter-chromosomal genome architecture by leading to the formation of specialized RNA factories.
II. RNA factories are sites where co-regulated nucleic acids, DNA and/or RNA, are found in close proximity to each other and to highly concentrated clusters of the factors that are key to their regulation.
III. Engagements of nucleic acids with an RNA factory increase the fidelity and efficiency of their regulation.

Despite the evidence presented in the previous sections focused on mRNA biogenesis, the hypothesis described here may be generalizable also to other types of non-coding RNAs (such as rRNA in the nucleolus) since some of the underlying mechanisms (described in the next subsection) likely translate to the specific regulations involved in the biogenesis of various classes of RNAs. Moreover, while this hypothesis is articulated specifically for inter-chromosomal regulations, it likely also extends for intra-chromosomal ones. Indeed, there is abundant evidence that *cis* genome architecture involves clustering of co-regulated genes (Schoenfelder et al., 2010; de Wit et al., 2013; Delaneau et al., 2019).

With regards to the first statement, relevant RNA biogenesis steps include the regulation of transcription, as exemplified by the existence of specific regulatory inter-chromosomal interactions (sections "*Trans*-Acting Regulatory Regions" and "Polycomb Domains") and of "zip coded" transcription factories (section "Transcription Factories"), but are not limited to this. Further regulation can be achieved by pre-mRNA processing steps in both specialized nuclear speckles (section "Nuclear Speckles") and tissue-specific splicing factories (section "Splicing Factories"). Moreover, I anticipate that future studies will reveal how other aspects of RNA biogenesis, such as RNA post-transcriptional modifications, are similarly involved in regulating inter-chromosomal architecture.

With reference to the second statement, relevant regulatory factors include transcription factors (many examples; section "Transcription Factories"), chromatin modifiers (i.e., PRC2), genome organizers (i.e., CTCF), the transcriptional machinery (i.e., Pol II and its cofactors), splicing regulators (i.e., RBM20), and other protein factors involved in various steps of RNA biogenesis. RNA species with a structural and/or enzymatic role are also likely involved, though this will need to be determined by further studies.

While the term "RNA factory" describes the overall subnuclear structure, I propose that the genomic loci involved could be referred to as a "*trans*-interacting (chromatin) domain" (TID). This acronym purposely reminisces TAD in order to stress the conceptual similarity between these structural features of the genome. Similarly to TADs, TIDs are statistical constructs that can be observed only at a population level (Rowley and Corces, 2018; Beagan and Phillips-Cremins, 2020). In individual cells, interactions between loci within a given TAD or TID are necessarily limited by steric hindrance and are transient. Thus, both of these terms do not aim to define a stable structure, yet they describe a set of loci that are more likely to interact with each other than with different loci outside of the domain but within the same genomic context. Such definition is applied in *cis* for TADs, and in *trans* for TIDs. Notably, however, TIDs and TADs may partially intersect in that multiple *cis*-interacting loci within a TAD may contribute to a TID by engaging in inter-chromosomal interactions with the same set of *trans* loci. On average, the interaction probability for loci within a TAD is expected to be higher than that for loci within a TID, as *cis* interactions are generally favored. Importantly, beyond a certain genomic distance (10–50 Mb, depending on the cell cycle stage), the probability of intra- and inter-chromosomal interactions becomes similar. Therefore, very long-range intra-chromosomal interactions may be effectively considered analogous to inter-chromosomal ones when computing and studying TIDs.

Mechanisms Leading to the Formation of RNA Factories

I propose a three-step process (**Figure 4A**):

I. *Trans*-acting regulatory factors bind onto core co-regulated nucleic acids.
II. Regulatory factors from multiple nucleic acids aggregate to form new clusters and/or enrich pre-existing ones.
III. Accessory nucleic acids are brought into proximity the cluster until a dynamic equilibrium is reached.

Step one involves different factors for distinct classes of RNA factories. For transcription factories, *trans*-acting regulators include transcription factors, epigenetic remodelers, genome organizing proteins, and Pol II and its cofactors, which engage with DNA either directly or through its associated histones. For splicing factories, the key *trans*-acting factors, namely splicing regulators, require prior initiation of transcription to generate the target RNA molecules. In both cases, multiple binding sites are found on either the same or multiple molecules (in the case of RNAs). The model stipulates that one or few nucleic acids with the highest density of binding sites function as "core" element(s), dictating the local accumulation of the regulators.

Step two is the result of either active mechanisms (i.e., cytoskeletal dynamics) and/or Brownian movements of nucleic acids extruding from the core of chromosomal territories. Changes in the local

chromatin environment may facilitate this process and increase intermixing in the inter-chromosomal space. Stabilization of core inter-chromosomal interactions involve either high-affinity interactions between factors that can form dimers and/or higher-order multimers (i.e., CTCF and many transcription factors), as well as weaker interactions such as those involving low-complexity regions characteristic of proteins that can undergo liquid-liquid phase separation (Patel et al., 2015). In some cases, regulators may join larger clusters that are pre-existing independently of nucleic acids.

In step three, large and growing clusters of regulators stabilize additional inter-chromosomal interactions with co-regulated nucleic acids characterized by lower binding density for the regulators. These "accessory" DNA or RNA molecules are those that most strongly benefit from interfacing with the RNA factory, which augments the efficiency of their regulation. As the RNA factory grows in size, it eventually hits steric and thermodynamic limitations that stabilize its further growth to a state of dynamic equilibrium.

Throughout the process, regulators can be either general and housekeeping, binding to nucleic acids with low degree of specificity or no specificity at all (i.e., CTCF and Pol II), or highly specific and responsive to cell states (i.e., tissue-specific transcription and splicing regulators). The resulting RNA factory is more or less cell state-specific depending on the relative balance of these two classes. "Zip coded" transcription and/or splicing factories result from highly specific regulators enriching for their target nucleic acids.

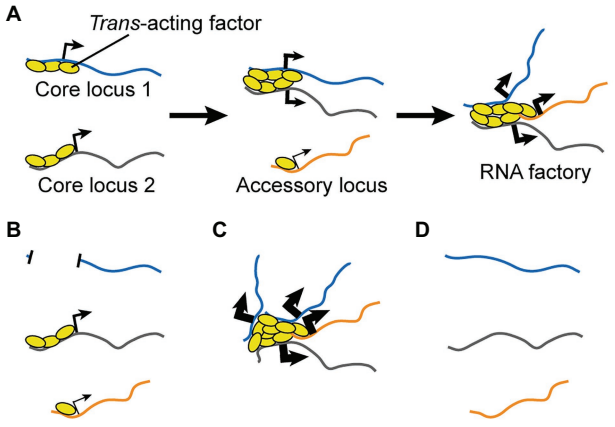

FIGURE 4 | A Testable Model for the Formation of RNA Factories. **(A)** Multiple copies of *trans*-acting regulatory factors bind onto core nucleic acids, aggregate to form new clusters and/or enrich pre-existing ones, and recruit accessory co-regulated nucleic acids. RNA factories promote the efficacy and accuracy of RNA biogenesis processes (thicker black arrows). **(B-D)** The model can be tested by depleting **(B)** or adding **(C)** core binding sites for *trans*-acting regulatory factors, or by depleting the *trans*-acting regulatory factors. Loss-of-function perturbations are expected to impair inter-chromosomal interactions **(B,D)** and RNA biogenesis mechanisms **(B)**, while a gain-of-function experiment should have opposite effects (very thick arrows, **C**).

Experimental Testing Strategies and Predicted Outcomes

Besides descriptive investigations, at least three types of perturbations can be used to test the model (**Figures 4B-D**):

I. Depletion of core nucleic acids.
II. Addition of clustered binding sites for *trans*-acting regulatory factors.
III. Depletion of *trans*-acting regulatory factors.

The first experiment involves reducing or completely preventing the binding of *trans*-acting regulatory factor(s) to one or more of the core co-regulated nucleic acids. For DNA loci, this can be achieved by mutations or larger deletions. For RNAs, this can be achieved by mutations of the originating genes to prevent their transcription, or through epigenetic silencing. For both DNA and RNA perturbations, an important control experiment is to reduce the protein levels of the gene(s) of interest by a similar magnitude as the test perturbation, but without affecting the abundance of binding sites for the *trans*-acting regulatory factors. This is key to exclude that any observed effect is not simply due to impairment of intermediate protein-mediated regulations. For instance, one can generate knockout and/or hypomorph mutations that do not alter the DNA sequence in the regions of interest and/or do not alter RNA levels. The predicted outcomes of this perturbation are: (1) impaired clustering of the *trans*-acting regulators; (2) impaired inter-chromosomal interactions between loci in the TID; and (3) decreased efficiency in the regulation for accessory nucleic acids (**Figure 4B**).

The second experiment has the opposite goal, namely to increase the clustered binding of the *trans*-acting regulatory factor(s). For instance, one can express transgenic copies of a core nucleic acid or an artificial sequence containing arrayed binding sites. In the former case, the transgene can be mutated to prevent protein translation and bypass the need to control for this aspect. An RNA can also be upregulated through epigenetic mechanisms. The predicted outcomes of this perturbation are opposite to those outlined above (**Figure 4C**). Of note, this experiment can be also performed in combination to the previous perturbation in order to rescue the function of the RNA factory, so as to formally prove the sequence(s) that are necessary and sufficient for its formation and activity.

The third experiment is a simple loss-of-function perturbation for the *trans*-acting regulatory factor. In this context, the only testable aspect of the model is the expectation that this perturbation should reduce inter-chromosomal interactions for loci involved in the TID (**Figure 4D**). Indeed, loss-of-function of the regulatory factor is expected to reduce or abolish both its clustering and the regulation of its target nucleic acids independently of its function in an RNA factory.

TESTING THE HYPOTHESIS: CHALLENGES AND IDEAS

To the best of my knowledge, only a handful of published studies on RNA factories have performed functional experiments

along the lines of those proposed in the previous subsection. Examples include the aforementioned analyses of the TNFα-responsive transcription factory (section "Transcription Factories"; Fanucchi et al., 2013), and of the RBM20 splicing factory (section "Splicing Factories"; Bertero et al., 2019a). While the results of such investigations have been consistent with the expectations outlined above, much more work will be needed to thoroughly test the model proposed in the previous section and to identify its limitation. Moreover, the pervasiveness of this mechanism in the regulation of gene expression and inter-chromosomal genome organization will have to be established. Here I outline some of the key challenges towards meeting these two goals, and I propose some potential solutions.

Identification of RNA Factories

Perhaps the most important roadblock in the field is the difficulty in predicting which inter-chromosomal interactions are functionally related to specific regulations of RNA biogenesis and not just background noise. Overcoming this issue will likely require refinements in both the experimental and analytical pipelines.

Methodological Consideration

From a methodological perspective, it is now clear that in-solution Hi-C is too noisy and therefore inadequate to probe for inter-chromosomal interactions (Nagano et al., 2015). Alternatives that rely on in-nucleus or substrate-tethered proximity ligation are therefore preferable (Kalhor et al., 2012; Nagano et al., 2015; Dai et al., 2016). An open limitation remains the lack of approaches that specifically detect *trans* chromatin interactions: since these can account for as low as 10% of the total ligation fragments and are spread over a large 2D genomic space, the costs to achieve a reasonable sequencing depth can be prohibitive. This is an even larger concern for single-cell Hi-C approaches, which result in data that is sometimes so sparse that even the analysis of *cis* interactions becomes a challenge (Nagano et al., 2013; Ramani et al., 2017). The development of strategies to enrich for inter-chromosomal proximity ligation fragments, for instance using capture probe sets for different chromosomes, may be a worthy pursuit. Along these same lines, it may be valuable to enrich for chromatin regions interacting with specific factors thought to be implicated in RNA factories, for instance using chromatin interaction analysis by paired-end tag sequencing (ChIA-PET; Fullwood et al., 2009) or proximity ligation-assisted ChIP-seq (PLAC-seq; Fang et al., 2016). Importantly, direct probing of nuclear RNA-RNA interactions in the context of RNA factories may be an important complementary approach, which may be enabled by recent methodological advances (reviewed in Kudla et al., 2020).

Another important technical consideration is the inability of Hi-C to reliably detect inter-chromosomal architecture when loci are not in very close and stable proximity, such as it may be the case for many RNA factories. Alternative ligation-free methods such as SPRITE, GAM, and TSA-seq have begun to elucidate inter-chromosomal architecture at varying spatial resolution (Beagrie et al., 2017; Chen et al., 2018; Quinodoz et al., 2018). Extension of these and similar approaches to also detect nuclear RNA-RNA interactions will be an important achievement. The application of emerging imaging approaches that allow simultaneous and/or sequential labeling of tens of loci will be key to validating the existence of RNA factories (Bintu et al., 2018; Mateo et al., 2019). Moreover, the recent development of high-throughput DNA FISH approaches gives hope that such methods may soon allow the necessary scalability required for discovery-type studies (Shachar et al., 2015; Finn et al., 2019).

Analytical Considerations

The identification of significant inter-chromosomal interactions must rely on statistical approaches that are specifically designed for this task. Indeed, approaches originally devised for the analysis of *cis* interactions control for biases that are not necessarily applicable to *trans* ones (for instance by correcting for the linear genomic distance between the interacting loci). Unfortunately, most of the analytical tools for 3C-type data developed so far have focused only on intra-chromosomal interactions (reviewed in Lin et al., 2019). The most recent implementation of FitHiC is one of the few publicly available tools that allows the user to interrogate bulk Hi-C data for both *cis* and *trans* interactions (Ay et al., 2014; Kaul et al., 2020). This approach makes minimal assumptions about genome size or structure, and generates an empirical null model to estimate the significance of interactions. While robust and flexible, this strategy does not account for known biases of *trans* interactions in mammalian systems, such as their enrichment for loci in the A compartment and their different distribution across small and large chromosomes. A similar limitation applies to existing network-based strategies to study inter-chromosomal interactions both from bulk Hi-C (Kaufmann et al., 2015) and single-cell Hi-C data (Bulathsinghalage and Liu, 2020). Statistical frameworks that explicitly model chromosome territorialization and chromatin (sub) compartmentalization may lead to more streamlined identification of inter-chromosomal interactions arising from metastable RNA factories, which are expected to "stand out" from other *trans* interactions resulting from the random intermixing of neighboring chromatin domains.

Another promising venue is the modeling and study of 3D genome structure populations, which as discussed in the section "Nuclear Subcompartments" has already provided important insights in inter-chromosomal genome architecture (Kalhor et al., 2012; Dai et al., 2016). Nevertheless, these models are currently limited in their spatial resolution due to the computational complexity involved in finely modeling large mammalian genomes. Besides presenting its own important statistical challenges, analysis of imaging-based chromatin structure data is also currently not easily integrated with sequencing-based data modalities. This remains is one of the most important areas for further analytical development in the wider 3D chromatin organization field.

Identification of *Trans*-Acting Regulators in RNA Factories

A second key challenge towards clarifying the role of RNA factories is the identification of the *trans*-acting factors that contribute to their formation and/or function. For transcription factories, the enrichment of DNA binding motifs within the regulatory regions proximal to the involved loci may provide testable hypotheses as to the transcriptional regulators involved. RNA binding motifs enrichment on pre-mRNAs may similarly help in the context of RNA splicing factories. Further evidence could be gathered by integration of data modalities directly probing for DNA and/or RNA binding of candidate factors. However, these analyses are limited by the availability of well-annotated binding motifs and/or of experimentally determined binding profiles for the cell state of interest.

Robust methods to unbiasedly identify factors that interact with specific nucleic acids are direly needed. Biochemical purification approaches relying on direct capture of nucleic acids have been developed, but so far showed limited sensitivity and specificity (reviewed in Machyna and Simon, 2018). A potential alternative is the application of *in situ* proximity labeling, for instance by relying on promiscuous biotinylating enzymes to covalently tag nearby proteins for subsequent streptavidin-based purification and identification through mass spectrometry. These methods have been successfully used to determine protein-protein interactions, and their implementation to study locus-specific regulations are also emerging (reviewed in Ummethum and Hamperl, 2020).

Testing the Function of RNA Factories

Last but not least, the functional characterization of RNA factories is key to determining their physiological relevance. As elaborated in the section "Experimental Testing Strategies and Predicted Outcomes", mechanistic studies may involve various means of perturbing RNA factories through loss- and/or gain-of-function. While these studies do not pose unsurmountable challenges due to recent advancements in genome editing technologies, they remain time-consuming. Characterization of putative RNA factories may be performed with more throughput by leveraging on functional screening strategies based on the combination of CRISPR/Cas9 perturbations and transcriptional phenotyping through single-cell RNA-sequencing (reviewed in Yao and Dai, 2018).

CONCLUSION

In summary, emerging evidence suggests a functional bidirectional link between the regulation of RNA biogenesis and inter-chromosomal 3D genome architecture. In other words, the formation of RNA factories organized around TIDs may be not just an emergent property of gene regulation, but also a feature needed to maximize the fidelity and efficiency of such regulations. This is a testable hypothesis based on a clear mechanistic model: the clustered binding of *trans*-acting regulators onto core nucleic acid targets, leading to their aggregation and to the subsequent recruitment of accessory co-regulated nucleic acids. Future studies will prove, disprove, or modify this model. It will also be key to test the pervasiveness of such regulation to determine whether it involves only a few specific nucleic acids or it is a general property controlling many loci.

Many key open questions remain in this area. First off, what is the biochemical and/or biophysical mechanism that explains the co-localization of certain nucleic acids from different chromosomes? Besides transcription factors (Dai et al., 2016, 2018), several DNA binding proteins may be involved, including CTCF (Ling et al., 2006; Botta et al., 2010), Condensin II (Rowley et al., 2019), ARID1A (Wu et al., 2019), and the Pol II machinery (Nagashima et al., 2019). Regulatory RNAs may also play a key role. Such factors may form higher-order clusters *via* a combination of specific and non-specific interactions, possibly leading to liquid-liquid phase separation (reviewed in Banani et al., 2017; Langdon and Gladfelter, 2018). Clustering may happen after progressive concentration onto specific nucleic acids with high avidity for the factors. Alternatively, pre-existing clusters may "trap" target nucleic acids as they come into contact through Brownian motions. Active repositioning of loci through the nuclear cytoskeleton may be also involved (Hu et al., 2008). Likely, multiple mechanisms contribute to different extents in various cases.

Is dysfunction of RNA factories and/or TIDs involved in the pathogenesis of human disease? Cancer has been linked to alteration of chromosome territories (Barutcu et al., 2015) and of specific inter-chromosomal interactions (Patel et al., 2014; Du et al., 2015). Disruption of inter-chromosomal chromatin interactions has been observed in a human model of 15q duplication syndrome, an aneuploidy linked to 1–3% of autism cases (Meguro-Horike et al., 2011). Viral genomes can also engage in specific inter-chromosomal associations, as shown for Epstein-Barr Virus (EBV; Moquin et al., 2017) and Kaposi's Sarcoma-Associated Herpesvirus (KSHV; Kang et al., 2011). Whether any of these or other diseases involve the dysregulation of RNA biogenesis factories remains to be tested. Notably, disruption of RBM20 splicing factories may play a key role in the pathogenesis of DCM (Schneider et al., 2020), though whether RBM20 mutations disrupt 3D chromatin organization is unclear.

Moving beyond RNA biogenesis, is proximity between co-regulated nucleic acids a general phenomenon in biology? It is tempting to speculate that other key biochemical reactions such as DNA repair and replication may also benefit from the increased efficiency and fidelity led by local accumulation of relevant regulators. Overall, increasing our spatial awareness of nuclear regulations may prove a worthy and impactful pursuit.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This work was supported by an Innovation Pilot Award from the Institute for Stem Cell and Regenerative Medicine of the University of Washington, and by an administrative supplement from the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health (U54DK107979-05S1).

ACKNOWLEDGMENTS

The author thanks Dr. Charles E. Murry and the past and current members of his laboratory for the stimulating discussions in this area, particularly Dr. Aidan Fenix for providing critical feedback on the manuscript.

REFERENCES

Abboud, N., Morris, T. M., Hiriart, E., Yang, H., Bezerra, H., Gualazzi, M. -G., et al. (2015). A cohesin–OCT4 complex mediates sox enhancers to prime an early embryonic lineage. *Nat. Commun.* 6:6749. doi: 10.1038/ncomms7749

Aguet, F., Barbeira, A. N., Bonazzola, R., Jo, B., Kasela, S., Liang, Y., et al. (2020). The GTEx consortium atlas of genetic regulatory effects across human tissues the genotype tissue expression consortium. *Science* 369, 1318–1330. doi: 10.1101/787903

Akerberg, B. N., and Pu, W. T. (2019). Genetic and epigenetic control of heart development. *Cold Spring Harb. Perspect. Biol.* 12:a036756. doi: 10.1101/cshperspect.a036756

Apostolou, E., and Thanos, D. (2008). Virus infection induces NF-κB-dependent interchromosomal associations mediating monoallelic IFN-β gene expression. *Cell* 134, 85–96. doi: 10.1016/j.cell.2008.05.052

Ay, F., Bailey, T. L., and Noble, W. S. (2014). Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. *Genome Res.* 24, 999–1011. doi: 10.1101/gr.160374.113

Banani, S. F., Lee, H. O., Hyman, A. A., and Rosen, M. K. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* 18, 285–298. doi: 10.1038/nrm.2017.7

Barbieri, I., Tzelepis, K., Pandolfini, L., Shi, J., Millán-Zambrano, G., Robson, S. C., et al. (2017). Promoter-bound METTL3 maintains myeloid leukaemia by m6A-dependent translation control. *Nature* 552, 126–131. doi: 10.1038/nature24678

Barbosa-Morais, N. L., Carmo-Fonseca, M., and Aparício, S. (2006). Systematic genome-wide annotation of spliceosomal proteins reveals differential gene family expansion. *Genome Res.* 16, 66–77. doi: 10.1101/gr.3936206

Bartosovic, M., Molares, H. C., Gregorova, P., Hrossova, D., Kudla, G., and Vanacova, S. (2017). N6-methyladenosine demethylase FTO targets pre-mRNAs and regulates alternative splicing and 3'-end processing. *Nucleic Acids Res.* 45, 11356–11370. doi: 10.1093/nar/gkx778

Barutcu, A. R., Lajoie, B. R., McCord, R. P., Tye, C. E., Hong, D., Messier, T. L., et al. (2015). Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. *Genome Biol.* 16:214. doi: 10.1186/s13059-015-0768-0

Battle, A., and Montgomery, S. B. (2014). Determining causality and consequence of expression quantitative trait loci. *Hum. Genet.* 133, 727–735. doi: 10.1007/s00439-014-1446-0

Beagan, J. A., and Phillips-Cremins, J. E. (2020). On the existence and functionality of topologically associating domains. *Nat. Genet.* 52, 8–16. doi: 10.1038/s41588-019-0561-1

Beagrie, R. A., Scialdone, A., Schueler, M., Kraemer, D. C. A., Chotalia, M., Xie, S. Q., et al. (2017). Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* 543, 519–524. doi: 10.1038/nature21411

Bersaglieri, C., and Santoro, R. (2019). Genome organization in and around the nucleolus. *Cell* 8:579. doi: 10.3390/cells8060579

Bertero, A., Brown, S., Madrigal, P., Osnato, A., Ortmann, D., Yiangou, L., et al. (2018). The SMAD2/3 interactome reveals that TGFβ controls m6A mRNA methylation in pluripotency. *Nature* 555, 256–259. doi: 10.1038/nature25784

Bertero, A., Fields, P. A., Ramani, V., Bonora, G., Yardımcı, G. G., Reinecke, H., et al. (2019a). Dynamics of genome reorganization during human cardiogenesis reveal an RBM20-dependent splicing factory. *Nat. Commun.* 10:1538. doi: 10.1038/s41467-019-09483-5

Bertero, A., Fields, P. A., Smith, A. S., Leonard, A., Beussman, K., Sniadecki, N. J., et al. (2019b). Chromatin compartment dynamics in a haploinsufficient model of cardiac laminopathy. *J. Cell Biol.* 218, 2919–2944. doi: 10.1101/555250

Bertero, A., and Rosa-Garrido, M. (2020). Three-dimensional chromatin organization in cardiac development and disease. *J. Mol. Cell. Cardiol.* 151, 89–105. doi: 10.1016/j.yjmcc.2020.11.008

Biggin, M. D. (2011). Animal transcription networks as highly connected, quantitative continua. *Dev. Cell* 21, 611–626. doi: 10.1016/j.devcel.2011.09.008

Bintu, B., Mateo, L. J., Su, J. H., Sinnott-Armstrong, N. A., Parker, M., Kinrot, S., et al. (2018). Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* 362:eaau1783. doi: 10.1126/science.aau1783

Boccaletto, P., Machnicka, M. A., Purta, E., Piatkowski, P., Baginski, B., Wirecki, T. K., et al. (2018). MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* 46, D303–D307. doi: 10.1093/nar/gkx1030

Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., et al. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol.* 3:e157. doi: 10.1371/journal.pbio.0030157

Botta, M., Haider, S., Leung, I. X. Y., Lio, P., and Mozziconacci, J. (2010). Intra-and inter-chromosomal interactions correlate with CTCF binding genome wide. *Mol. Syst. Biol.* 6:426. doi: 10.1038/msb.2010.79

Boyle, S., Gilchrist, S., Bridger, J. M., Mahy, N. L., Ellis, J. A., and Bickmore, W. A. (2001). The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum. Mol. Genet.* 10, 211–219. doi: 10.1093/hmg/10.3.211

Branco, M. R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol.* 4, 780–788. doi: 10.1371/journal.pbio.0040138

Brauch, K. M., Karst, M. L., Herron, K. J., de Andrade, M., Pellikka, P. A., Rodeheffer, R. J., et al. (2009). Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. *J. Am. Coll. Cardiol.* 54, 930–941. doi: 10.1016/j.jacc.2009.05.038

Brown, J. M., Green, J., Neves, R., Das, P., Wallace, H. A. C., Smith, A. J. H., et al. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *J. Cell Biol.* 182, 1083–1097. doi: 10.1083/jcb.200803174

Brown, J. M., Leach, J., Reittie, J. E., Atzberger, A., Lee-Prudhoe, J., Wood, W. G., et al. (2006). Coregulated human globin genes are frequently in spatial proximity when active. *J. Cell Biol.* 172, 177–187. doi: 10.1083/jcb.200507073

Bulathsinghalage, C., and Liu, L. (2020). Network-based method for regions with statistically frequent interchromosomal interactions at single-cell resolution. *BMC Bioinformatics* 21, 1–15. doi: 10.1186/s12859-020-03689-x

Chang, C. -P., and Bruneau, B. G. (2012). Epigenetics and cardiovascular development. *Annu. Rev. Physiol.* 74, 41–68. doi: 10.1146/annurev-physiol-020911-153242

Chapski, D. J., Rosa-Garrido, M., Hua, N., Alber, F., and Vondriska, T. M. (2019). Spatial principles of chromatin architecture associated with organ-

specific gene regulation. *Front. Cardiovasc. Med.* 5:186. doi: 10.3389/fcvm.2018.00186

Chen, Y., Zhang, Y., Wang, Y., Zhang, L., Brinkman, E. K., Adam, S. A., et al. (2018). Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. *J. Cell Biol.* 217, 4025–4048. doi: 10.1083/jcb.201807108

Concia, L., Veluchamy, A., Ramirez-Prado, J. S., Martin-Ramirez, A., Huang, Y., Perez, M., et al. (2020). Wheat chromatin architecture is organized in genome territories and transcription factories. *Genome Biol.* 21:104. doi: 10.1186/s13059-020-01998-1

Cook, P. R. (1995). A chromomeric model for nuclear and chromosome structure. *J. Cell Sci.* 108, 2927–2935.

Cook, P. R. (1997). The transcriptional basis of chromosome pairing. *J. Cell Sci.* 110, 1033–1040.

Cook, P. R. (2010). A model for all genomes: the role of transcription factories. *J. Mol. Biol.* 395, 1–10. doi: 10.1016/j.jmb.2009.10.031

Cremer, T., and Cremer, M. (2010). Chromosome territories. *Cold Spring Harb. Perspect. Biol.* 2:a003889. doi: 10.1101/cshperspect.a003889

Cremer, T., Cremer, C., Baumann, H., Luedtke, E. K., Sperling, K., Teuber, V., et al. (1982). Rabl's model of the interphase chromosome arrangement tested in Chinise hamster cells by premature chromosome condensation and laser-UV-microbeam experiments. *Hum. Genet.* 60, 46–56. doi: 10.1007/BF00281263

Cremer, M., Hase, J.von, Volm, T., Brero, A., Kreth, G., Walter, J., et al. (2001). Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosom. Res.* 9, 541–567. doi: 10.1023/A:1012495201697

Croft, J. A., Bridger, J. M., Boyle, S., Perry, P., Teague, P., and Bickmore, W. A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. *J. Cell Biol.* 145, 1119–1131. doi: 10.1083/jcb.145.6.1119

Dai, Y., Li, C., Pei, G., Dong, X., Ding, G., Zhao, Z., et al. (2018). Multiple transcription factors contribute to inter-chromosomal interaction in yeast. *BMC Syst. Biol.* 12:140. doi: 10.1186/s12918-018-0643-1

Dai, C., Li, W., Tjong, H., Hao, S., Zhou, Y., Li, Q., et al. (2016). Mining 3D genome structure populations identifies major factors governing the stability of regulatory communities. *Nat. Commun.* 7:11549. doi: 10.1038/ncomms11549

de Wit, E., Bouwman, B. A. M., Zhu, Y., Klous, P., Splinter, E., Verstegen, M. J. A. M., et al. (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. *Nature* 501, 227–231. doi: 10.1038/nature12420

Dekker, J., and Mirny, L. (2016). The 3D genome as moderator of chromosomal communication. *Cell* 164, 1110–1121. doi: 10.1016/j.cell.2016.02.007

Delaneau, O., Zazhytska, M., Borel, C., Giannuzzi, G., Rey, G., Howald, C., et al. (2019). Chromatin three-dimensional interactions mediate genetic effects on gene expression. *Science* 364:eaat8266. doi: 10.1126/science.aat8266

Ding, D. -Q., Okamasa, K., Katou, Y., Oya, E., Nakayama, J. -I., Chikashige, Y., et al. (2019). Chromosome-associated RNA-protein complexes promote pairing of homologous chromosomes during meiosis in Schizosaccharomyces pombe. *Nat. Commun.* 10:5598. doi: 10.1038/s41467-019-13609-0

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., et al. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380. doi: 10.1038/nature11082

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112

Dong, Q., Li, N., Li, X., Yuan, Z., Xie, D., Wang, X., et al. (2018). Genome-wide Hi-C analysis reveals extensive hierarchical chromatin interactions in rice. *Plant J.* 94, 1141–1156. doi: 10.1111/tpj.13925

Du, M., Yuan, T., Schilter, K. F., Dittmar, R. L., Mackinnon, A., Huang, X., et al. (2015). Prostate cancer risk locus at 8q24 as a regulatory hub by physical interactions with multiple genomic loci across the genome. *Hum. Mol. Genet.* 24, 154–166. doi: 10.1093/hmg/ddu426

Duan, Z., Andronescu, M., Schutz, K., McIlwain, S. Y., Kim, Y. J., Lee, C., et al. (2010). A three-dimensional model of the yeast genome. *Nature* 465, 363–367. doi: 10.1038/nature08973

Eicher, J. D., Landowski, C., Stackhouse, B., Sloan, A., Chen, W., Jensen, N., et al. (2015). GRASP v2.0: an update on the genome-wide repository of associations between SNPs and phenotypes. *Nucleic Acids Res.* 43, D799–D804. doi: 10.1093/nar/gku1202

Fairfax, B. P., Makino, S., Radhakrishnan, J., Plant, K., Leslie, S., Dilthey, A., et al. (2012). Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat. Genet.* 44, 502–510. doi: 10.1038/ng.2205

Fang, R., Yu, M., Li, G., Chee, S., Liu, T., Schmitt, A. D., et al. (2016). Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. *Cell Res.* 26, 1345–1348. doi: 10.1038/cr.2016.137

Fanucchi, S., Shibayama, Y., Burd, S., Weinberg, M. S., and Mhlanga, M. M. (2013). Chromosomal contact permits transcription between coregulated genes. *Cell* 155, 606–620. doi: 10.1016/j.cell.2013.09.051

Fehrmann, R. S. N., Jansen, R. C., Veldink, J. H., Westra, H. J., Arends, D., Bonder, M. J., et al. (2011). Trans-eqtls reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the hla. *PLoS Genet.* 7:e1002197. doi: 10.1371/journal.pgen.1002197

Finn, E. H., Pegoraro, G., Brandão, H. B., Valton, A. L., Oomen, M. E., Dekker, J., et al. (2019). Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* 176, 1502–1515.e10. doi: 10.1016/j.cell.2019.01.020

Fujita, T., Kitaura, F., Yuno, M., Suzuki, Y., Sugano, S., and Fujii, H. (2017). Locus-specific ChIP combined with NGS analysis reveals genomic regulatory regions that physically interact with the Pax5 promoter in a chicken B cell line. *DNA Res.* 24, 537–548. doi: 10.1093/dnares/dsx023

Fullwood, M. J., Liu, M. H., Pan, Y. F., Liu, J., Xu, H., Mohamed, Y. B., et al. (2009). An oestrogen-receptor-α-bound human chromatin interactome. *Nature* 462, 58–64. doi: 10.1038/nature08497

Galganski, L., Urbanek, M. O., and Krzyzosiak, W. J. (2017). Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic Acids Res.* 45, 10350–10368. doi: 10.1093/nar/gkx759

Gerlich, D., Beaudouin, J., Kalbfuss, B., Daigle, N., Eils, R., and Ellenberg, J. (2003). Global chromosome positions are transmitted through mitosis in mammalian cells. *Cell* 112, 751–764. doi: 10.1016/s0092-8674(03)00189-2

Gilsbach, R., Schwaderer, M., Preissl, S., Grüning, B. A., Kranzhöfer, D., Schneider, P., et al. (2018). Distinct epigenetic programs regulate cardiac myocyte development and disease in the human heart *in vivo*. *Nat. Commun.* 9:391. doi: 10.1038/s41467-017-02762-z

Grieve, I. C., Dickens, N. J., Pravenec, M., Kren, V., Hubner, N., Cook, S. A., et al. (2008). Genome-wide co-expression analysis in multiple tissues. *PLoS One* 3:e4033. doi: 10.1371/journal.pone.0004033

Grundberg, E., Small, K. S., Hedman, Å. K., Nica, A. C., Buil, A., Keildson, S., et al. (2012). Mapping cis-and trans-regulatory effects across multiple tissues in twins. *Nat. Genet.* 44, 1084–1089. doi: 10.1038/ng.2394

Guo, W., Schafer, S., Greaser, M. L., Radke, M. H., Liss, M., Govindarajan, T., et al. (2012). RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat. Med.* 18, 766–773. doi: 10.1038/nm.2693

Haas, J., Frese, K. S., Peil, B., Kloos, W., Keller, A., Nietsch, R., et al. (2015). Atlas of the clinical genetics of human dilated cardiomyopathy. *Eur. Heart J.* 36, 1123–1135. doi: 10.1093/eurheartj/ehu301

Harada, A., Mallappa, C., Okada, S., Butler, J. T., Baker, S. P., Lawrence, J. B., et al. (2015). Spatial re-organization of myogenic regulatory sequences temporally controls gene expression. *Nucleic Acids Res.* 43, 2008–2021. doi: 10.1093/nar/gkv046

Heinig, M., Petretto, E., Wallace, C., Bottolo, L., Rotival, M., Lu, H., et al. (2010). A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk. *Nature* 467, 460–464. doi: 10.1038/nature09386

Herman, D. S., Lam, L., Taylor, M. R. G., Wang, L., Teekakirikul, P., Christodoulou, D., et al. (2012). Truncations of titin causing dilated cardiomyopathy. *N. Engl. J. Med.* 366, 619–628. doi: 10.1056/NEJMoa1110186

Hesse, M., Bednarz, R., Carls, E., Becker, C., Bondareva, O., Lother, A., et al. (2021). Proximity to injury, but neither number of nuclei nor ploidy define pathological adaptation and plasticity in cardiomyocytes. *J. Mol. Cell. Cardiol.* 152, 95–104. doi: 10.1016/j.yjmcc.2020.11.012

Homouz, D., and Kudlicki, A. S. (2013). The 3D organization of the yeast genome correlates with co-expression and reflects functional relations between genes. *PLoS One* 8:e54699. doi: 10.1371/journal.pone.0054699

Hu, Q., Kwon, Y. S., Nunez, E., Cardamone, M. D., Hutt, K. R., Ohgi, K. A., et al. (2008). Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin

granules. *Proc. Natl. Acad. Sci. U. S. A.* 105, 19199–19204. doi: 10.1073/pnas.0810634105

Huang, H., Weng, H., Zhou, K., Wu, T., Zhao, B. S., Sun, M., et al. (2019). Histone H3 trimethylation at lysine 36 guides m6A RNA modification co-transcriptionally. *Nature* 567, 414–419. doi: 10.1038/s41586-019-1016-7

Imholte, G. C., Scott-Boyer, M. P., Labbe, A., Deschepper, C. F., and Gottardo, R. (2013). IBMQ: A R/Bioconductor package for integrated Bayesian modeling of eQTL data. *Bioinformatics* 29, 2797–2798. doi: 10.1093/bioinformatics/btt485

Innocenti, F., Cooper, G. M., Stanaway, I. B., Gamazon, E. R., Smith, J. D., Mirkov, S., et al. (2011). Identification, replication, and functional fine-mapping of expression quantitative trait loci in primary human liver tissue. *PLoS Genet.* 7:e1002078. doi: 10.1371/journal.pgen.1002078

Ito, K., Noguchi, A., Uosaki, Y., Taga, T., Arakawa, H., and Takizawaa, T. (2018). Gfap and Osmr regulation by BRG1 and STAT3 *via* interchromosomal gene clustering in astrocytes. *Mol. Biol. Cell* 29, 209–219. doi: 10.1091/mbc.E17-05-0271

Ito, K., Sanosaka, T., Igarashi, K., Ideta-Otsuka, M., Aizawa, A., Uosaki, Y., et al. (2016). Identification of genes associated with the astrocyte-specific gene Gfap during astrocyte differentiation. *Sci. Rep.* 6:23903. doi: 10.1038/srep23903

Jabbari, K., and Bernardi, G. (2017). An isochore framework underlies chromatin architecture. *PLoS One* 12:e0168023. doi: 10.1371/journal.pone.0168023

Jabbari, K., Chakraborty, M., and Wiehe, T. (2019). DNA sequence-dependent chromatin architecture and nuclear hubs formation. *Sci. Rep.* 9:14646. doi: 10.1038/s41598-019-51036-9

Joehanes, R., Zhang, X., Huan, T., Yao, C., Ying, S. -X., Nguyen, Q. T., et al. (2017). Integrated genome-wide analysis of expression quantitative trait loci aids interpretation of genomic association studies. *Genome Biol.* 18:16. doi: 10.1186/s13059-016-1142-6

Joshi, O., Wang, S. Y., Kuznetsova, T., Atlasi, Y., Peng, T., Fabre, P. J., et al. (2015). Dynamic reorganization of extremely long-range promoter-promoter interactions between two states of pluripotency. *Cell Stem Cell* 17, 748–757. doi: 10.1016/j.stem.2015.11.010

Jurica, M. S., and Moore, M. J. (2003). Pre-mRNA splicing: awash in a sea of proteins. *Mol. Cell* 12, 5–14. doi: 10.1016/S1097-2765(03)00270-3

Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F., and Chen, L. (2012). Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nat. Biotechnol.* 30, 90–98. doi: 10.1038/nbt.2057

Kang, H., Wiedmer, A., Yuan, Y., Robertson, E., and Lieberman, P. M. (2011). Coordination of KSHV latent and lytic gene control by CTCF-cohesin mediated chromosome conformation. *PLoS Pathog.* 7:e1002140. doi: 10.1371/journal.ppat.1002140

Karbassi, E., Fenix, A., Marchiano, S., Muraoka, N., Nakamura, K., Yang, X., et al. (2020). Cardiomyocyte maturation: advances in knowledge and implications for regenerative medicine. *Nat. Rev. Cardiol.* 17, 341–359. doi: 10.1038/s41569-019-0331-x

Karbassi, E., Rosa-Garrido, M., Chapski, D. J., Wu, Y., Ren, S., Wang, Y., et al. (2019). Direct visualization of cardiac transcription factories reveals regulatory principles of nuclear architecture during pathological remodeling. *J. Mol. Cell. Cardiol.* 128, 198–211. doi: 10.1016/j.yjmcc.2019.02.003

Kaufmann, S., Fuchs, C., Gonik, M., Khrameeva, E. E., Mironov, A. A., and Frishman, D. F. (2015). Inter-chromosomal contact networks provide insights into mammalian chromatin organization. *PLoS One* 10:e0126125. doi: 10.1371/journal.pone.0126125

Kaul, A., Bhattacharyya, S., and Ay, F. (2020). Identifying statistically significant chromatin contacts from hi-C data with FitHiC2. *Nat. Protoc.* 15, 991–1012. doi: 10.1038/s41596-019-0273-0

Kayvanpour, E., Sedaghat-Hamedani, F., Amr, A., Lai, A., Haas, J., Holzer, D. B., et al. (2017). Genotype-phenotype associations in dilated cardiomyopathy: meta-analysis on more than 8000 individuals. *Clin. Res. Cardiol.* 106, 127–139. doi: 10.1007/s00392-016-1033-6

Ke, S., Alemu, E. A., Mertens, C., Gantman, E. C., Fak, J. J., Mele, A., et al. (2015). A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev.* 29, 2037–2053. doi: 10.1101/gad.269415.115

Ke, S., Pandya-Jones, A., Saito, Y., Fak, J. J., Vågbø, C. B., Geula, S., et al. (2017a). m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. *Genes Dev.* 31, 990–1006. doi: 10.1101/gad.301036.117

Ke, Y., Xu, Y., Chen, X., Feng, S., Liu, Z., Sun, Y., et al. (2017b). 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell* 170, 367–381.e20. doi: 10.1016/j.cell.2017.06.029

Kempfer, R., and Pombo, A. (2020). Methods for mapping 3D chromosome architecture. *Nat. Rev. Genet.* 21, 207–226. doi: 10.1038/s41576-019-0195-2

Kirkland, J. G., Raab, J. R., and Kamakaka, R. T. (2013). TFIIIC bound DNA elements in nuclear organization and insulation. *Biochim. Biophys. Acta* 1829, 418–424. doi: 10.1016/j.bbagrm.2012.09.006

Kirsten, H., Al-Hasani, H., Holdt, L., Gross, A., Beutner, F., Krohn, K., et al. (2015). Dissecting the genetics of the human transcriptome identifies novel trait-related trans-eQTLs and corroborates the regulatory relevance of non-protein coding loci. *Hum. Mol. Genet.* 24, 4746–4763. doi: 10.1093/hmg/ddv194

Krumm, A., and Duan, Z. (2018). Understanding the 3D genome: emerging impacts on human disease. *Semin. Cell Dev. Biol.* 90, 62–77. doi: 10.1016/j.semcdb.2018.07.004

Kruse, K., Sewitz, S., and Madan Babu, M. (2013). A complex network framework for unbiased statistical analyses of DNA-DNA contact maps. *Nucleic Acids Res.* 41, 701–710. doi: 10.1093/nar/gks1096

Kudla, G., Wan, Y., and Helwak, A. (2020). RNA conformation capture by proximity ligation. *Annu. Rev. Genomics Hum. Genet.* 21, 81–100. doi: 10.1146/annurev-genom-120219-073756

Lahbib-Mansais, Y., Barasc, H., Marti-Marimon, M., Mompart, F., Iannuccelli, E., Robelin, D., et al. (2016). Expressed alleles of imprinted IGF2, DLK1 and MEG3 colocalize in 3D-preserved nuclei of porcine fetal cells. *BMC Cell Biol.* 17:35. doi: 10.1186/s12860-016-0113-9

Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., et al. (2018). The human transcription factors. *Cell* 172, 650–665. doi: 10.1016/j.cell.2018.01.029

Langdon, E. M., and Gladfelter, A. S. (2018). A new lens for RNA localization: liquid-liquid phase separation. *Annu. Rev. Microbiol.* 72, 255–271. doi: 10.1146/annurev-micro-090817-062814

Li, G., Ruan, X., Auerbach, R. K., Sandhu, K. S., Zheng, M., Wang, P., et al. (2012). Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* 148, 84–98. doi: 10.1016/j.cell.2011.12.014

Liang, L., Morar, N., Dixon, A. L., Lathrop, G. M., Abecasis, G. R., Moffatt, M. F., et al. (2013). A cross-platform analysis of 14,177 expression quantitative trait loci derived from lymphoblastoid cell lines. *Genome Res.* 23, 716–726. doi: 10.1101/gr.142521.112

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293. doi: 10.1126/science.1181369

Lin, D., Bonora, G., Yardımcı, G. G., and Noble, W. S. (2019). Computational methods for analyzing and modeling genome structure and organization. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 11:e1435. doi: 10.1002/wsbm.1435

Ling, J. Q., Li, T., Hu, J. F., Vu, T. H., Chen, H. L., Qiu, X. W., et al. (2006). CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science* 312, 269–272. doi: 10.1126/science.1123191

Lomvardas, S., Barnea, G., Pisapia, D. J., Mendelsohn, M., Kirkland, J., and Axel, R. (2006). Interchromosomal interactions and olfactory receptor choice. *Cell* 126, 403–413. doi: 10.1016/j.cell.2006.06.035

Maatz, H., Jens, M., Liss, M., Schafer, S., Heinig, M., Kirchner, M., et al. (2014). RNA-binding protein RBM20 represses splicing to orchestrate cardiac pre-mRNA processing. *J. Clin. Invest.* 124, 3419–3430. doi: 10.1172/JCI74523

Machyna, M., and Simon, M. D. (2018). Catching RNAs on chromatin using hybridization capture methods. *Brief. Funct. Genomics* 17, 96–103. doi: 10.1093/BFGP/ELX038

Mahy, N. L., Perry, P. E., and Bickmore, W. A. (2002). Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. *J. Cell Biol.* 159, 753–763. doi: 10.1083/jcb.200207115

Mao, Y. S., Zhang, B., and Spector, D. L. (2011). Biogenesis and function of nuclear bodies. *Trends Genet.* 27, 295–306. doi: 10.1016/j.tig.2011.05.006

Marchianò, S., Bertero, A., and Murry, C. E. (2019). Learn from your elders: developmental biology lessons to guide maturation of stem cell-derived

cardiomyocytes. *Pediatr. Cardiol.* 40, 1367–1387. doi: 10.1007/s00246-019-02165-5

Markenscoff-Papadimitriou, E., Allen, W. E., Colquitt, B. M., Goh, T., Murphy, K. K., Monahan, K., et al. (2014). Enhancer interaction networks as a means for singular olfactory receptor expression. *Cell* 159, 543–557. doi: 10.1016/j.cell.2014.09.033

Mateo, L. J., Murphy, S. E., Hafner, A., Cinquini, I. S., Walker, C. A., and Boettiger, A. N. (2019). Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* 568, 49–54. doi: 10.1038/s41586-019-1035-4

Matsushita, M., Ochiai, H., Suzuki, K. -I. T., Hayashi, S., Yamamoto, T., Awazu, A., et al. (2017). Dynamic changes in the interchromosomal interaction of early histone gene loci during development of sea urchin. *J. Cell Sci.* 130, 4097–4107. doi: 10.1242/jcs.206862

McKenzie, M., Henders, A. K., Caracella, A., Wray, N. R., and Powell, J. E. (2014). Overlap of expression quantitative trait loci (eQTL) in human brain and blood. *BMC Med. Genet.* 7:31. doi: 10.1186/1755-8794-7-31

Meguro-Horike, M., Yasui, D. H., Powell, W., Schroeder, D. I., Oshimura, M., LaSalle, J. M., et al. (2011). Neuron-specific impairment of inter-chromosomal pairing and transcription in a novel model of human 15q-duplication syndrome. *Hum. Mol. Genet.* 20, 3798–3810. doi: 10.1093/hmg/ddr298

Michaelson, J. J., Loguercio, S., and Beyer, A. (2009). Detection and interpretation of expression quantitative trait loci (eQTL). *Methods* 48, 265–276. doi: 10.1016/j.ymeth.2009.03.004

Mitchell, J. A., and Fraser, P. (2008). Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev.* 22, 20–25. doi: 10.1101/gad.454008

Monahan, K., Horta, A., and Lomvardas, S. (2019). LHX2- and LDB1-mediated trans interactions regulate olfactory receptor choice. *Nature* 565, 448–453. doi: 10.1038/s41586-018-0845-0

Monahan, K., Schieren, I., Cheung, J., Mumbey-Wafula, A., Monuki, E. S., and Lomvardas, S. (2017). Cooperative interactions enable singular olfactory receptor expression in mouse olfactory neurons. *elife* 6:e28620. doi: 10.7554/eLife.28620

Moquin, S. A., Thomas, S., Whalen, S., Warburton, A., Fernandez, S. G., McBride, A. A., et al. (2017). The Epstein-Barr virus episome maneuvers between nuclear chromatin compartments during reactivation. *J. Virol.* 92;e01413–17. doi: 10.1128/jvi.01413-17

Mummery, C. L., Zhang, J., Ng, E. S., Elliott, D. A., Elefanty, A. G., and Kamp, T. J. (2012). Differentiation of human embryonic stem cells and induced pluripotent stem cells to cardiomyocytes: a methods overview. *Circ. Res.* 111, 344–358. doi: 10.1161/CIRCRESAHA.110.227512

Nagano, T., Lubling, Y., Stevens, T. J., Schoenfelder, S., Yaffe, E., Dean, W., et al. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502, 59–64. doi: 10.1038/nature12593

Nagano, T., Várnai, C., Schoenfelder, S., Javierre, B. M., Wingett, S. W., and Fraser, P. (2015). Comparison of Hi-C results using in-solution versus in-nucleus ligation. *Genome Biol.* 16;175. doi: 10.1186/s13059-015-0753-7

Nagashima, R., Hibino, K., Ashwin, S. S., Babokhov, M., Fujishiro, S., Imai, R., et al. (2019). Single nucleosome imaging reveals loose genome chromatin networks *via* active RNA polymerase II. *J. Cell Biol.* 218, 1511–1530. doi: 10.1083/jcb.201811090

Nagele, R. G., Freeman, T., McMorrow, L., Thomson, Z., Kitson-Wind, K., and Lee, H. Y. (1999). Chromosomes exhibit preferential positioning in nuclei of quiescent human cells. *J. Cell Sci.* 112, 525–535.

Olivares-Chauvet, P., Mukamel, Z., Lifshitz, A., Schwartzman, O., Elkayam, N. O., Lubling, Y., et al. (2016). Capturing pairwise and multi-way chromosomal conformations using chromosomal walks. *Nature* 540, 296–300. doi: 10.1038/nature20158

Ong, C. T., and Corces, V. G. (2014). CTCF: an architectural protein bridging genome topology and function. *Nat. Rev. Genet.* 15, 234–246. doi: 10.1038/nrg3663

Osborne, C. S., Chakalova, L., Brown, K. E., Carter, D., Horton, A., Debrand, E., et al. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet.* 36, 1065–1071. doi: 10.1038/ng1423

Osborne, C. S., Chakalova, L., Mitchell, J. A., Horton, A., Wood, A. L., Bolland, D. J., et al. (2007). Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biol.* 5:e192. doi: 10.1371/journal.pbio.0050192

Papantonis, A., Kohro, T., Baboo, S., Larkin, J. D., Deng, B., Short, P., et al. (2012). TNFα signals through specialized factories where responsive coding and miRNA genes are transcribed. *EMBO J.* 31, 4404–4414. doi: 10.1038/emboj.2012.288

Papantonis, A., Larkin, J. D., Wada, Y., Ohta, Y., Ihara, S., Kodama, T., et al. (2010). Active RNA polymerases: mobile or immobile molecular machines? *PLoS Biol.* 8:e1000419. doi: 10.1371/journal.pbio.1000419

Parada, L. A., McQueen, P. G., and Misteli, T. (2004). Tissue-specific spatial organization of genomes. *Genome Biol.* 5:R44. doi: 10.1186/GB-2004-5-7-R44

Patel, B., Kang, Y., Cui, K., Litt, M., Riberio, M. S. J., Deng, C., et al. (2014). Aberrant TAL1 activation is mediated by an interchromosomal interaction in human T-cell acute lymphoblastic leukemia. *Leukemia* 28, 349–361. doi: 10.1038/leu.2013.158

Patel, A., Lee, H. O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M. Y., et al. (2015). A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* 162, 1066–1077. doi: 10.1016/j.cell.2015.07.047

Pederson, T. (2011). The nucleolus. *Cold Spring Harb. Perspect. Biol.* 3:a000638. doi: 10.1101/cshperspect.a000638

Perales, R., and Bentley, D. (2009). "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Mol. Cell* 36, 178–191. doi: 10.1016/j.molcel.2009.09.018

Pertea, M. (2012). The human transcriptome: an unfinished story. *Genes* 3, 344–360. doi: 10.3390/genes3030344

Pierce, B. L., Tong, L., Chen, L. S., Rahaman, R., Argos, M., Jasmine, F., et al. (2014). Mediation analysis demonstrates that trans-eQTLs are often explained by cis-mediation: a genome-wide analysis among 1,800 south Asians. *PLoS Genet.* 10:e1004818. doi: 10.1371/journal.pgen.1004818

Pink, R. C., Eskiw, C. H., Caley, D. P., and Carter, D. R. F. (2010). Analysis of β-globin chromatin micro-environment using a novel 3C variant, 4Cv. *PLoS One* 5:e13045. doi: 10.1371/journal.pone.0013045

Pombo, A., Jones, E., Iborra, F. J., Kimura, H., Sugaya, K., Cook, P. R., et al. (2000). Specialized transcription factories within mammalian nuclei. *Crit. Rev. Eukaryot. Gene Expr.* 10, 21–29.

Protze, S. I., Lee, J. H., and Keller, G. M. (2019). Human pluripotent stem cell-derived cardiovascular cells: from developmental biology to therapeutic applications. *Cell Stem Cell* 25, 311–327. doi: 10.1016/j.stem.2019.07.010

Quinodoz, S. A., Ollikainen, N., Tabak, B., Palla, A., Schmidt, J. M., Detmar, E., et al. (2018). Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* 174, 744–757.e24. doi: 10.1016/j.cell.2018.05.024

Ramani, V., Cusanovich, D. A., Hause, R. J., Ma, W., Qiu, R., Deng, X., et al. (2016). Mapping 3D genome architecture through *in situ* DNase hi-C. *Nat. Protoc.* 11, 2104–2121. doi: 10.1038/nprot.2016.126

Ramani, V., Deng, X., Qiu, R., Gunderson, K. L., Steemers, F. J., Disteche, C. M., et al. (2017). Massively multiplex single-cell Hi-C. *Nat. Methods* 14, 263–266. doi: 10.1038/nmeth.4155

Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680. doi: 10.1016/j.cell.2014.11.021

Refaat, M. M., Lubitz, S. A., Makino, S., Islam, Z., Frangiskakis, J. M., Mehdi, H., et al. (2012). Genetic variation in the alternative splicing regulator RBM20 is associated with dilated cardiomyopathy. *Hear. Rhythm* 9, 390–396. doi: 10.1016/j.hrthm.2011.10.016

Roix, J. J., McQueen, P. G., Munson, P. J., Parada, L. A., and Misteli, T. (2003). Spatial proximity of translocation-prone gene loci in human lymphomas. *Nat. Genet.* 34, 287–291. doi: 10.1038/ng1177

Rosa-Garrido, M., Chapski, D. J., Schmitt, A. D., Kimball, T. H., Karbassi, E., Monte, E., et al. (2017). High-resolution mapping of chromatin conformation in cardiac myocytes reveals structural remodeling of the epigenome in heart failure. *Circulation* 136, 1613–1625. doi: 10.1161/CIRCULATIONAHA.117.029430

Rowley, M. J., and Corces, V. G. (2018). Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* 19, 789–800. doi: 10.1038/s41576-018-0060-8

Rowley, M. J., Lyu, X., Rana, V., Ando-Kuri, M., Karns, R., Bosco, G., et al. (2019). Condensin II counteracts cohesin and RNA polymerase II in the establishment of 3D chromatin organization. *Cell Rep.* 26, 2890–2903.e3. doi: 10.1016/j.celrep.2019.01.116

Schneider, J. W., Oommen, S., Qureshi, M. Y., Goetsch, S. C., Pease, D. R., Sundsbak, R. S., et al. (2020). Dysregulated ribonucleoprotein granules

promote cardiomyopathy in RBM20 gene-edited pigs. *Nat. Med.* 26, 1788–1800. doi: 10.1038/s41591-020-1087-x

Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N. F., Horton, A., Andrews, S., et al. (2010). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat. Genet.* 42, 53–61. doi: 10.1038/ng.496

Sexton, T., Umlauf, D., Kurukuti, S., and Fraser, P. (2007). The role of transcription factories in large-scale structure and dynamics of interphase chromatin. *Semin. Cell Dev. Biol.* 18, 691–697. doi: 10.1016/j.semcdb.2007.08.008

Shachar, S., Voss, T. C., Pegoraro, G., Sciascia, N., and Misteli, T. (2015). Identification of gene positioning factors using high-throughput imaging mapping. *Cell* 162, 911–923. doi: 10.1016/j.cell.2015.07.035

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., et al. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* 38, 1348–1354. doi: 10.1038/ng1896

Slobodin, B., Han, R., Calderone, V., Vrielink, J. A. F. O., Loayza-Puch, F., Elkon, R., et al. (2017). Transcription impacts the efficiency of mRNA translation *via* co-transcriptional N6-adenosine methylation. *Cell* 169, 326–337.e12. doi: 10.1016/j.cell.2017.03.031

Spilianakis, C. G., Lalioti, M. D., Town, T., Lee, G. R., and Flavell, R. A. (2005). Interchromosomal associations between alternatively expressed loci. *Nature* 435, 637–645. doi: 10.1038/nature03574

Sun, L., Jing, Y., Liu, X., Li, Q., Xue, Z., Cheng, Z., et al. (2020). Heat stress-induced transposon activation correlates with 3D chromatin organization rearrangement in Arabidopsis. *Nat. Commun.* 11:1886. doi: 10.1038/s41467-020-15809-5

Takizawa, T., Gudla, P. R., Guo, L., Lockett, S., and Misteli, T. (2008). Allele-specific nuclear positioning of the monoallelically expressed astrocyte marker GFAP. *Genes Dev.* 22, 489–498. doi: 10.1101/gad.1634608

Tanabe, H., Habermann, F. A., Solovei, I., Cremer, M., and Cremer, T. (2002). Non-random radial arrangements of interphase chromosome territories: evolutionary considerations and functional implications. *Mutat. Res.* 504, 37–45. doi: 10.1016/s0027-5107(02)00077-5

Tiwari, V. K., Cope, L., McGarvey, K. M., Ohm, J. E., and Baylin, S. B. (2008). A novel 6C assay uncovers polycomb-mediated higher order chromatin conformations. *Genome Res.* 18, 1171–1179. doi: 10.1101/gr.073452.107

Tomikawa, J., Takada, S., Okamura, K., Terao, M., Ogata-Kawata, H., Akutsu, H., et al. (2020). Exploring trophoblast-specific Tead4 enhancers through chromatin conformation capture assays followed by functional screening. *Nucleic Acids Res.* 48, 278–289. doi: 10.1093/nar/gkz1034

Ummethum, H., and Hamperl, S. (2020). Proximity labeling techniques to study chromatin. *Front. Genet.* 11:450. doi: 10.3389/fgene.2020.00450

van den Hoogenhof, M. M. G., Beqqali, A., Amin, A. S., van der Made, I., Aufiero, S., Khan, M. A. F., et al. (2018). RBM20 mutations induce an arrhythmogenic dilated cardiomyopathy related to disturbed calcium handling. *Circulation* 138, 1130–1342. doi: 10.1161/CIRCULATIONAHA.117.031947

Vieux-Rochas, M., Fabre, P. J., Leleu, M., Duboule, D., and Noordermeer, D. (2015). Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain. *PNAS* 112, 4672–4677. doi: 10.1073/pnas.1504783112

Virani, S. S., Alonso, A., Benjamin, E. J., Bittencourt, M. S., Callaway, C. W., Carson, A. P., et al. (2020). Heart disease and stroke statistics—2020 update: a report from the American Heart Association. *Circulation* 141, E139–E596. doi: 10.1161/CIR.0000000000000757

Westra, H. J., Peters, M. J., Esko, T., Yaghootkar, H., Schurmann, C., Kettunen, J., et al. (2013). Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* 45, 1238–1243. doi: 10.1038/ng.2756

Wong, E. S., Schmitt, B. M., Kazachenka, A., Thybert, D., Redmond, A., Connor, F., et al. (2017). Interplay of cis and trans mechanisms driving transcription factor binding and gene expression evolution. *Nat. Commun.* 8;1092. doi: 10.1038/s41467-017-01037-x

Wu, S., Fatkhutdinov, N., Rosin, L., Luppino, J. M., Iwasaki, O., Tanizawa, H., et al. (2019). ARID1A spatially partitions interphase chromosomes. *Sci. Adv.* 5;eaaw5294. doi: 10.1126/sciadv.aaw5294

Wyles, S. P., Hrstka, S. C., Reyes, S., Terzic, A., Olson, T. M., and Nelson, T. J. (2016a). Pharmacological modulation of calcium homeostasis in familial dilated cardiomyopathy: an *in vitro* analysis from an RBM20 patient-derived iPSC model. *Clin. Transl. Sci.* 9, 158–167. doi: 10.1111/cts.12393

Wyles, S. P., Li, X., Hrstka, S. C., Reyes, S., Oommen, S., Beraldi, R., et al. (2016b). Modeling structural and functional deficiencies of RBM20 familial dilated cardiomyopathy using human induced pluripotent stem cells. *Hum. Mol. Genet.* 25, 254–265. doi: 10.1093/hmg/ddv468

Xiong, K., and Ma, J. (2019). Revealing Hi-C subcompartments by imputing inter-chromosomal chromatin interactions. *Nat. Commun.* 10;5069. doi: 10.1038/s41467-019-12954-4

Yaffe, E., and Tanay, A. (2011). Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. *Nat. Genet.* 43, 1059–1065. doi: 10.1038/ng.947

Yao, J., and Dai, H. L. (2018). Is pooled CRISPR-screening the dawn of a new era for functional genomics. *Adv. Exp. Med. Biol.* 1068, 171–176. doi: 10.1007/978-981-13-0502-3_14

Yao, C., Joehanes, R., Johnson, A. D., Huan, T., Liu, C., Freedman, J. E., et al. (2017). Dynamic role of trans regulation of gene expression in relation to complex traits. *Am. J. Hum. Genet.* 100, 571–580. doi: 10.1016/j.ajhg.2017.02.003

Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., et al. (2018). VIRMA mediates preferential m6A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov.* 4:10. doi: 10.1038/s41421-018-0019-0

Zhang, Y., Li, T., Preissl, S., Amaral, M. L., Grinstein, J. D., Farah, E. N., et al. (2019). Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. *Nat. Genet.* 51, 1380–1388. doi: 10.1038/s41588-019-0479-7

Zhao, B. S., Roundtree, I. A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* 18, 31–42. doi: 10.1038/nrm.2016.132

Zhao, Z., Tavoosidana, G., Sjölinder, M., Göndör, A., Mariano, P., Wang, S., et al. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.* 38, 1341–1347. doi: 10.1038/ng1891

Zheng, H., and Xie, W. (2019). The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.* 20, 535–550. doi: 10.1038/s41580-019-0132-4

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

frontiers in Cell and Developmental Biology

REVIEW
published: 11 March 2021
doi: 10.3389/fcell.2021.644630

CaMKIIδ Splice Variants in the Healthy and Diseased Heart

*Javier Duran[1,2], Lennart Nickel[1,2], Manuel Estrada[3], Johannes Backs[1,2] and Maarten M. G. van den Hoogenhof[1,2]**

[1] Institute of Experimental Cardiology, Heidelberg University, Heidelberg, Germany, [2] German Center for Cardiovascular Research (DZHK), Partner Site Heidelberg/Mannheim, Heidelberg, Germany, [3] Faculty of Medicine, Institute of Biomedical Sciences, University of Chile, Santiago, Chile

RNA splicing has been recognized in recent years as a pivotal player in heart development and disease. The Ca^{2+}/calmodulin dependent protein kinase II delta (CaMKIIδ) is a multifunctional Ser/Thr kinase family and generates at least 11 different splice variants through alternative splicing. This enzyme, which belongs to the CaMKII family, is the predominant family member in the heart and functions as a messenger toward adaptive or detrimental signaling in cardiomyocytes. Classically, the nuclear CaMKIIδB and cytoplasmic CaMKIIδC splice variants are described as mediators of arrhythmias, contractile function, Ca^{2+} handling, and gene transcription. Recent findings also put CaMKIIδA and CaMKIIδ9 as cardinal players in the global CaMKII response in the heart. In this review, we discuss and summarize the new insights into CaMKIIδ splice variants and their (proposed) functions, as well as CaMKII-engineered mouse phenotypes and cardiac dysfunction related to CaMKIIδ missplicing. We also discuss RNA splicing factors affecting CaMKII splicing. Finally, we discuss the translational perspective derived from these insights and future directions on CaMKIIδ splicing research in the healthy and diseased heart.

Keywords: RNA splicing, heart, CaMKII delta, splice variant, therapeutics

OPEN ACCESS

Edited by:
Paola Rizzo,
University of Ferrara, Italy

Reviewed by:
Joan Heller Brown,
University of California, San Diego, United States
Yan Zhang,
Peking University, China

***Correspondence:**
Maarten M. G. van den Hoogenhof
maarten.vandenhoogenhof@med.uni-heidelberg.de

Specialty section:
This article was submitted to Molecular Medicine, a section of the journal Frontiers in Cell and Developmental Biology

Received: *21 December 2020*
Accepted: *22 February 2021*
Published: *11 March 2021*

Citation:
Duran J, Nickel L, Estrada M, Backs J and van den Hoogenhof MMG (2021) CaMKIIδ Splice Variants in the Healthy and Diseased Heart. Front. Cell Dev. Biol. 9:644630. doi: 10.3389/fcell.2021.644630

INTRODUCTION

Heart disease is a major cause of death around the world, but the mechanisms underlying its development are not completely understood (Maggioni, 2015). In recent years, (dys)regulation of RNA splicing has been implicated in heart disease, but its contribution to the development and progression of heart disease is not fully known. RNA splicing, a conserved posttranscriptional mechanism, can generally be divided into two processes: constitutive and alternative splicing. In constitutive splicing, which happens in all intron-containing genes and is necessary for the maturation of pre-mRNAs into mRNA, introns are removed from the pre-mRNA and exons are joined together. In alternative splicing, exons of a gene may be included or excluded in the final mature mRNA, producing several different mRNA transcripts from only one gene. More than 95% of mRNAs are subjected to alternative splicing, promoting an increase in the diversity of the transcriptome, which subsequently leads to an increased proteome (Wang et al., 2008). The transcriptome, however, not only includes protein-coding mRNAs, but also non-coding mRNAs that are involved in a multitude of processes, thereby even further increasing the range of functions that are exerted by alternative splicing. Aberrations of alternative splicing can be the cause of multiple diseases including myotonic dystrophy, spinal bulbar atrophy, Prader-Willi syndrome, tauopathies, among others (Tazi et al., 2009).

In the heart, alternative splicing participates in pre- and postnatal development, as well as in the development and progression of heart disease (van den Hoogenhof et al., 2016). The function(s) of multiple splicing factors have been studied by using genetically modified mice, for example by knocking out RNA Binding Motif Protein 20 and 24 (Rbm20 and Rbm24), RNA Binding Fox-1 Homolog 1 and 2 (Rbfox1 and Rbfox2), and the Serine/arginine-rich splicing factor 1 (Srsf1 or Asf/sf2; van den Hoogenhof et al., 2016). Additionally, mutations that impact RNA splicing and induce heart disease have been documented. For example, mutations in splice sites for the genes encoding the Cardiac type Troponin T2 or for Myosin Binding Protein C, which lead to exon skipping or activation of cryptic splice sites, lead to truncated mRNA variants and subsequently impaired sarcomere contractions or hypertrophic cardiomyopathy (Thierfelder et al., 1994; Bonne et al., 1995).

Mutations in splicing factors as a cause for heart disease are only described for RBM20, and these often lead to an arrhythmogenic form of dilated cardiomyopathy (Brauch et al., 2009; Refaat et al., 2012; Parikh et al., 2019). Titin, encoded by the *TTN* gene, is a sarcomeric protein that acts as a scaffolding filament, a signaling platform, and is a principal regulator of contraction in striated muscle. Mutations located in the *RBM20* gene are linked to Titin missplicing toward a larger isoform termed N2BA-Giant, both in animal models and human dilated cardiomyopathy (Guo et al., 2012). In addition to Titin, RBM20 also regulates splicing of other transcripts such as the Ca^{2+} channel ryanodine receptor type 2 (RYR2), the L-type Ca^{2+} channel (LTCC) alpha 1C subunit (CACNA1C/Cav1.2), and the Ca^{2+}/calmodulin-dependent protein kinase II delta (CAMK2δ). Missplicing of these genes may contribute to the severity of the cardiomyopathy (Beraldi et al., 2014; Maatz et al., 2014; van den Hoogenhof et al., 2018). In this regard, Rbm20 deficiency induces an intracellular Ca^{2+} overload, similar to what is seen in mice lacking the alternative splice factor Asf/sf2 (Xu et al., 2005; van den Hoogenhof et al., 2018). Interestingly, mice lacking either Rbm20 or Asf/sf2 display a similar splicing switch in Ca^{2+}/calmodulin dependent protein kinase II delta (CaMKIIδ) in the heart, from the CaMKIIδB and CaMKIIδC isoform to the CaMKIIδA and CaMKIIδ9 isoform (Xu et al., 2005; Guo et al., 2012; van den Hoogenhof et al., 2018). Traditionally, studies on CaMKII focus on overexpression or activation of CaMKII, but these new findings situate CaMKIIδ splice isoform changes as a focus into the regulators of functional remodeling in the heart.

As an example, mice lacking Asf/sf2 or Rbm20 both missplice CaMKIIδ to the same extent, and it is hypothesized that the switch in CaMKIIδ splice isoforms is one of the causes of Ca^{2+} handling deregulation in these mice (Xu et al., 2005; van den Hoogenhof et al., 2018). Moreover, the functional redundancy of splice factors that alternatively splice CaMKIIδ in the heart, as well as the identification of potential other splice factors deserves additional attention. The (partial) overlap in targets of many cardiac splicing factors, including Asf/sf2 and Rbm20, suggests a coordinated process that needs to be unraveled. Lastly, the functional differences of (increased expression or activation of) the different CaMKIIδ splice variants remain to be investigated. In this review, we will discuss recent discoveries around CaMKIIδ splice variants, how these might control positive and negative changes in the heart, the therapeutic potential derived from these studies, and we will end with potential future directions of CaMKIIδ research.

Ca^{2+}/CALMODULIN DEPENDENT PROTEIN KINASES

Ca^{2+}/calmodulin-dependent protein kinase (CaMK) belongs to the multifunctional CaMK family. In mammals, CaMKs are divided into 3 classes: CaMKI, CaMKII, and CaMKIV; which in total comprise more than 27 proteins (Swulius and Waxham, 2008; Takemoto-Kimura et al., 2017). The three CaMK classes are expressed in a broad range of cell types, including cardiomyocytes. Activated CaMKI and IV can induce cardiomyocyte hypertrophy, and can activate distinct transcriptional targets in the heart that participate in cardiac remodeling, such as myocyte enhancer factor 2 (MEF2), or synergize with other players involved in cardiac hypertrophy, such as Nuclear factor of activated T cells (NFAT; Passier et al., 2000). In the heart, CaMKII is involved in the control of Ca^{2+} handling and gene transcription, and is suggested as a mediator of mostly (but not only) maladaptive and detrimental effects on cardiac integrity (Bers and Grandi, 2009; Anderson et al., 2011; Beckendorf et al., 2018). Although less is known compared to its function in disease, CaMKII regulates multiple physiological processes such as excitation–contraction coupling (ECC) and excitation–transcription coupling (ETC), the flight or fight response, and contractile force generation (Beckendorf et al., 2018). For example, CaMKII is necessary to increase the heart rate after β-adrenergic stimulation (a.k.a. the fight or flight response). This was first observed in mice with genetic CaMKII inhibition, which surprisingly had a lower heart rate after stress. This could be explained by the effect of CaMKII on the Ca^{2+} uptake and release from the SR in the sinoatrial node (SAN; Wu et al., 2009). Similarly, the LTCC is another important component in the fight or flight response, and this is independent from β-adrenergic stimulation. CaMKII inhibition prevents the increase in heart rate after LTCC stimulation with its agonist BayK, even though I_{Cal} similarly increased in both control and CaMKII-inhibited SAN cells. This suggests that CaMKII activation is necessary for the heart rate increase in the fight or flight response, also without β-adrenergic stimulation (Gao et al., 2011). In addition to its role in the fight or flight response, CaMKII is also involved in the muscle response to exercise. Multiple studies using CaMKII inhibitors showed the involvement of CaMKII in contractile adaptation and in physiological hypertrophy in response to exercise (Kemi et al., 2007; Burgos et al., 2017). However, these experiments were all performed with complete CaMKII inhibition, and did not focus on specific CaMKII proteins or splice isoforms. Therefore, no conclusions can be made on what CaMKII proteins or splice isoforms are necessary in these processes. However, one could, for example, speculate that the effects on Ca^{2+} handling in the heart during the fight or flight response are likely mediated by CaMKIIδC (see **Table 1**). In general, even though CaMKII is necessary for some physiological

TABLE 1 | Phosphorylation targets of CaMKII.

Phosphorylation target	Splice variant	References
PDE4	n.s.	Mika et al., 2015
NOX	n.s.	Pandey et al., 2011
eNOS	n.s.	Murthy et al., 2017
DrP1	n.s.	Bo et al., 2018
McU	n.s.	Joiner et al., 2012
Titin	n.s.	Hidalgo et al., 2013
MyBP-C	n.s.	Hartzell and Glass, 1984
NHE	n.s.	Vila-Petroff et al., 2010
LTCC	n.s.	Anderson et al., 1994; Koval et al., 2010
RyR	CaMKIIδC, but not CaMKIIδB	Zhang et al., 2007; Di Carlo et al., 2014
PLN	CaMKIIδC, but not CaMKIIδB	Zhang et al., 2003
Calcineurin	n.s.	Kreusser et al., 2014
HDAC	CaMKIIδA, CaMKIIδB, CaMKIIδC	Backs et al., 2006; Zhang et al., 2007
UBE2T	CaMKIIδ9, but not CaMKIIδA or CaMKIIδB	Zhang et al., 2019
HSF1	CaMKIIδB	Holmberg et al., 2001; Peng et al., 2010
CREB	n.s.	Sun et al., 1994
NF-κB	CaMKIIδC, but not CaMKIIδB	Gray et al., 2017
IKK	n.s.	Martin et al., 2018
Histone 3	CaMKIIδB	Awad et al., 2013
Different ion channels		See review Hegyi et al., 2019

List of known CaMKII phosphorylation targets, and which CaMKII splice isoforms are able to phosphorylate these targets.
n.s., not specified.

processes, the loss of CaMKII is thought to be protective against heart disease, and sustained activation of CaMKII is therefore proposed to induce maladaptive remodeling.

Ca^{2+}/CALMODULIN DEPENDENT PROTEIN KINASE II

Ca^{2+}/calmodulin dependent protein kinase II proteins are encoded by four separate genes, *CaMKII*α/β/γ/δ, each one with a different expression pattern. CaMKIIα and CaMKIIβ are mainly expressed in the brain and mediate synaptic functions underlying learning, memory, and cognition (Cook et al., 2018). Their expression in the heart is still under debate, since some authors suggest that the α and β genes are not detectable in the heart (Kreusser et al., 2014), while others report the presence of CaMKIIα and β in ventricular cardiomyocytes (Cipolletta et al., 2015). CaMKIIγ and CaMKIIδ are expressed and found in the healthy and diseased heart, with CaMKIIδ being the highest expressed, and only these CaMKII genes have been knocked out specifically in the heart (Edman and Schulman, 1994; Colomer et al., 2003; Kreusser et al., 2014). The basic structure of CaMKIIδ consists of a specific Ser/Thr kinase domain at the *N*-terminus (exons 1–10), a regulatory segment (exons 11–12), a variable linker (exons 13–19), and a hub domain (exons 20–22; **Figure 1**). Under basal conditions, the kinase is in an autoinhibited state, with the regulatory domain acting like a substrate for the catalytic domain (Kelly et al., 1988; Hoffman et al., 2011). Binding of calmodulin induces a conformational change, which releases the association of the regulatory and catalytic domain, rendering the enzyme active (Kolodziej et al., 2000; Gaertner et al., 2004; Myers et al., 2017). Under physiological conditions, CaMKIIδ activity is typically driven by the presence of Ca^{2+}/CaM. However, posttranslational modifications can result in the prevention of reassociation between the regulatory and catalytic domains, thereby converting the enzyme to a persistently active state. These posttranslational modifications include autophosphorylation at Thr-287 (Lai et al., 1987) and Thr-306/307 (Patton et al., 1990; Lu et al., 2003), oxidation at Met-281/282 (Erickson et al., 2008), and Met-308 (Konstantinidis et al., 2020), O-GlcNAc modification at Ser-280 (Erickson et al., 2013), and nitrosylation at Cys-290 (Erickson et al., 2015; **Figure 1**). While posttranslational modifications in between the catalytic domain and the CaM binding site lead to autonomous activity, modifications within the CaM binding site, in contrast, appear to have an inhibitory effect. For example, phosphorylation of Thr-306/307 has an inhibitory effect on CaMKII activity, as this modification prevents binding of CaM (Lu et al., 2003). CaMKII proteins can form large oligomeric structures of 12 subunits from one or a combination of different isoforms, with the carboxy-terminal hub domains centrally located and the kinase domains arranged in a circle around the hub center, connected by the linker region. Oligomerization is necessary for rapid inter-subunit autophosphorylation at Thr-287 (Thr-286 in CaMKIIα) by catalytic domains of other subunits localized in the same oligomer during longer periods of Ca^{2+}/CaM binding (Lai et al., 1987). In the phosphorylated state, the affinity of Ca^{2+}/CaM binding is increased while the release is slowed down, a process known as Calmodulin trapping (Meyer et al., 1992). This enables the activity of CaMKII even with decreasing Ca^{2+} concentrations, until the phosphate group is removed by a phosphatase (Strack et al., 1997). Oxidation of CaMKIIδ at Met-281/282 works as a sensor system for reactive oxygen species (ROS) and therefore oxidative stress in the heart. Altered oxidation has been shown in a variety of cardiac disease models and can lead to arrhythmias, suggesting that CaMKIIδ-specific antioxidants could serve as a potential future therapeutic agent (Luczak and Anderson, 2014). Nitrosylation also appears to be an important modification of key elements of cardiac function, including CaMKIIδ. *S*-nitrosylation of Cys-290 site leads to sustained autonomous CaMKIIδ activation, whereas *S*-nitrosylation at Cys-273 inhibits CaMKIIδ activation when NO donors are present before Ca^{2+}/CaM is available (Erickson et al., 2015). The existing dual mechanisms of activating and deactivating modifications of the same nature (phosphorylation and nitrosylation) suggest that the regulation of CaMKIIδ activity by posttranslational modifications is much more complex than previously assumed. Especially the dynamics of the presence or absence of different modifications in relation

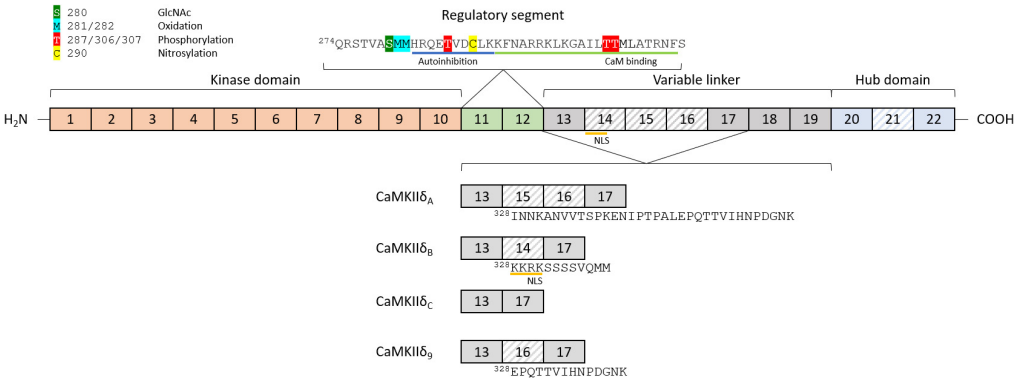

FIGURE 1 | Structure and features of the *CaMKIIδ* gene and exon composition of different splice isoforms. The basic structure of the *CaMKIIδ* gene consists of a kinase domain (exons 1–10), a regulatory segment (exons 11–12), a variable linker (exons 13–19), and a hub domain (exons 20–22). The variable regions in different splice isoforms are located between exons 13 and 17 (variable domain 1) and exons 20 and 22 (variable domain 2). Different posttranslational modifications can lead to activation or inactivation of CaMKIIδ, including phosphorylation, oxidation, nitrosylation, and *O*-linked *N*-acetylglucosamine. CaM-binding and autoinhibition regions are marked in the regulatory segment. The four major cardiac splice isoforms are shown with their different exon compositions in the variable region 1. Differentially spliced exons are striped, and amino acid sequences of the differently spliced region are added starting at position 328.

to varying Ca^{2+}/CaM concentrations should be the aim of future experiments. Ser-280 can be O-GlcNAcylated in the presence of elevated glucose concentrations, leading to activation of CaMKIIδ (Erickson et al., 2013). The enzyme *O*-GlcNAc transferase catalyzes this modification in the presence of UDP-*N*-acetylglucosamine, which is formed as a product of the of the hexosamine biosynthetic pathway (Hart and Akimoto, 2009). This modification is particularly important in the diabetic heart, where the ratio of *O*-GlcNAc-modified CaMKII to total CaMKII is increased due to changes in glucose signaling (Erickson et al., 2013; Lu et al., 2020). Whether alternative splicing increases the susceptibility of specific splice isoforms to these PTMs, or whether alternative splicing induces new potential PTM sites is currently unknown.

Ca^{2+}/calmodulin dependent protein kinase II gamma, the second most expressed CaMKII protein in the heart, has 4 described splice variants (Singer et al., 1997). However, the relationship between CaMKIIγ and cardiovascular disease is poorly explored. In this regard, CaMKIIγ is involved in atherosclerotic plaque development, which could drive myocardial infarction (Doran et al., 2017), and its mRNA levels are increased in pressure overload-induced cardiac hypertrophy in mice and in Rbm20-deficient rats (Guo et al., 2012; Kreusser et al., 2014). Cardiomyocyte specific deletion of CaMKIIγ shows a decrease in cardiomyocyte apoptosis and a reduction of cardiac hypertrophy induced by transverse aortic constriction and isoproterenol treatment in mice (Kreusser et al., 2014).

Ca^{2+}/calmodulin dependent protein kinase II delta is best known to participate in the regulation of Ca^{2+} handling (Zhang et al., 2007), through the phosphorylation of the LTCC, which facilitates intracellular Ca^{2+} entry (Koval et al., 2010), the sarcoplasmic reticulum (SR) membrane protein Phospholamban (PLN), with the consequent increase in Ca^{2+} uptake from cytoplasm into SR lumen by the Sarco/endoplasmic reticulum Ca^{2+}-ATPase (SERCA2a; Mattiazzi and Kranias, 2014), and RyR2 which increases SR Ca^{2+} leak into the cytoplasm (Witcher et al., 1991; van Oort et al., 2010). Additionally, CaMKIIδ regulates gene transcription, e.g., through phosphorylation of histone deacetylase 4 (HDAC4; Backs et al., 2006, 2009) and Histone 3 (Awad et al., 2013; Saadatmand et al., 2019), mitochondrial reprogramming (Westenbrink et al., 2015), and inflammasome activation (Suetomi et al., 2018; Willeford et al., 2018). These effects are closely related to each other and are involved in the development of physiological and pathophysiological effects of CaMKIIδ in the heart (**Figure 2** and **Table 1**). However, it cannot be ruled out that other CaMKIIδ functions, independent of calcium handling and transcriptional regulation, contribute significantly to heart disease since many other phosphorylation targets of CaMKII have been described.

Both global and cardiomyocyte specific deletion of CaMKIIδ protects against adverse cardiac remodeling (Backs et al., 2009; Ling et al., 2009, 2013), while transgenic overexpression of CaMKIIδ promotes cardiac hypertrophy or dilated cardiomyopathy (Zhang et al., 2002, 2003, 2019; Xu et al., 2005). Similar to CaMKIIγ deletion in the heart, CaMKIIδ cardiomyocyte knockout (KO) mice show an only partial decrease in known target phosphorylation, and reduction of cardiac hypertrophy in mice subjected to transverse aortic constriction surgery. While this is less evident in some CaMKIIδ KO mouse models (Ling et al., 2009) than in others (Kreusser et al., 2014), the data suggests redundancy of these two genes.

Cardiomyocyte specific deletion of CaMKIIδ attenuates the increase in the inflammatory mediators CCL2/CCL3, NFkB and its downstream targets IL-1β, IL-6, CXCL1, TNF-α, and NLRP3 in non-ischemic and ischemic cardiac disease (Weinreuter et al., 2014; Suetomi et al., 2018; Willeford et al., 2018). In line with this, hearts from cardiomyocyte specific CaMKIIδ KO mice demonstrate that cardiac infarct formation in ischemia/reperfusion injury, as well as the upregulation of the NFκB and gene expression of pro-inflammatory TNF-α and IL-6 are dependent on CaMKIIδ specifically in cardiomyocytes (Ling et al., 2013; Gray et al., 2017). Due to the redundant contribution

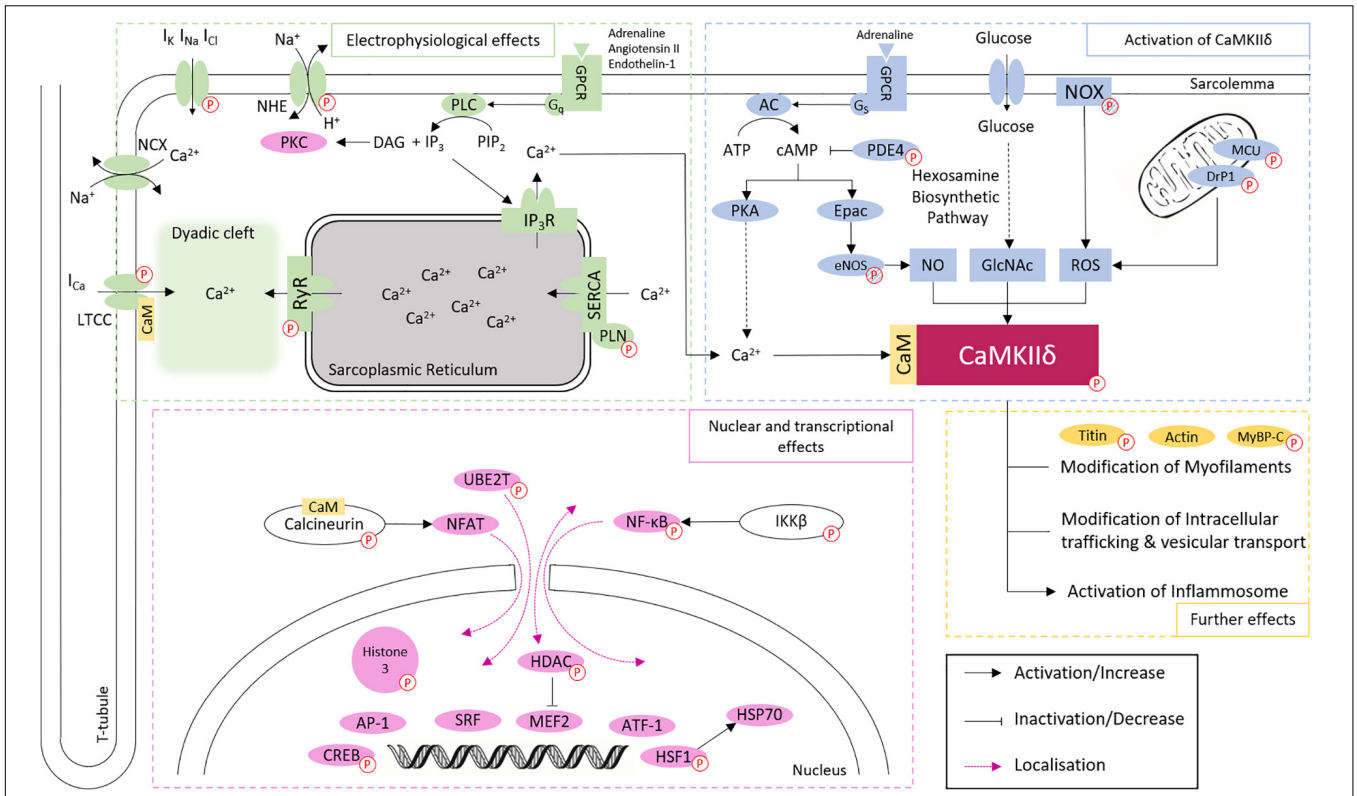

FIGURE 2 | Overview of CaMKIIδ activation and downstream signaling in cardiomyocytes. CaMKIIδ can be activated via Ca^{2+}/CaM binding, autophosphorylation, and different posttranslational modifications. The phosphorylation of different proteins (indicated by *P* symbol) by CaMKIIδ can have both activating and inactivating functions. This affects cardiac electrophysiology, especially Ca^{2+} handling, gene expression, intracellular trafficking, mitochondrial processes, energetics, cell-cell coupling, inflammatory reactions, and the cytoskeletal network. Different phosphorylation targets are listed with a corresponding reference in **Table 1**. Targets linked to phosphorylation by CaMKIIδ, but without proven direct phosphorylation, are indicated by striped *P* symbols.

of CaMKIIγ and CaMKIIδ in cardiomyocyte apoptosis and cardiac hypertrophy, the generation of a double KO (DKO) for CaMKIIγ and CaMKIIδ was necessary to clarify the role of CaMKII in the heart. DKO mice show that both genes contribute redundantly to phosphorylation of PLN (Thr-17), RyR2 (Ser-2814), HDAC4 (Ser-632), and calcineurin (Ser-411). Moreover, DKO mice show protection against cardiac dysfunction and interstitial fibrosis induced by pressure overload and chronic β-adrenergic stimulation (Kreusser et al., 2014). These mice, similar to CaMKIIδ KO generated by another group (Ling et al., 2013), are protected against post-infarct remodeling and inflammatory processes in the heart (Weinreuter et al., 2014). Unexpectedly, the deletion of both genes does not inhibit cardiac hypertrophy, due the hypo-phosphorylation of calcineurin, which in turn increases calcineurin activity (Kreusser et al., 2014). This uncovered CaMKIIδ to be not only a maladaptive effector in the heart, but also a regulator of physiological calcineurin-induced cardiac hypertrophy.

In short, CaMKIIδ is a central transducer of intra- and extracellular signaling in physiological processes, but is also involved in a multitude of maladaptive processes (**Figure 2** and **Table 1**). The question arises how a protein can have this degree of diversity in its function. We hypothesize that this is, at least in part, due to the different splice isoforms of CaMKIIδ .

CaMKIIδ SPLICE VARIANTS

The first description of different splice isoforms of CaMKIIδ was made by Schworer et al. (1993), after which additional splice isoforms were discovered and described in different tissues (Edman and Schulman, 1994; Srinivasan et al., 1994) [for an overview see Beckendorf et al. (2018)]. Until now, 11 different splice isoforms have been described, but not all of them are expressed in the heart (Zhang et al., 2019). The isoforms differ in inclusion of exons between exons 13–17, or exons 20–22 (**Figure 1**). Initially, research focus lay on the CaMKIIδC (a.k.a. CaMKIIδ2) and CaMKIIδB (a.k.a. CaMKIIδ3) splice isoforms in the heart. Only later CaMKIIδA, and very recently CaMKIIδ9 were described as mediators in the pathogenesis of cardiac disease (Xu et al., 2005; Zhang et al., 2019).

Ca^{2+}/calmodulin dependent protein kinase heteromultimerization can include different CaMKII proteins and splice isoforms from those proteins. The ratio of CaMKII proteins and splice isoforms can regulate the localization of the holoenzyme (Mishra et al., 2011). Thus, it is hypothesized that the switch in CaMKIIδ isoform ratio might determine the effect on phosphorylation sites in proteins present in different subcellular locations, e.g., high relative expression of nuclear CaMKIIδB might phosphorylate proteins involved in gene transcription,

such as HDAC4/5, and high relative expression of cytoplasmatic CaMKIIδC on Ca^{2+}-handling proteins, such as PLN or RyR2.

The RNA binding proteins Rbfox1 and Rbfox2, Asf/sf2, Sc35, and Rbm20 can regulate CaMKIIδ splicing, suggesting that CaMKIIδ splicing is a dynamic and regulated process, which should be explored extensively (Gray and Heller Brown, 2014; Beckendorf et al., 2018). The first study that showed a shift in CaMKIIδ splicing, from the CaMKIIδB/δC isoforms toward CaMKIIδA, characterized the cardiomyocyte specific KO of the splicing factor Asf/sf2. These mice presented with cardiomyopathy and cardiomyocyte Ca^{2+} handling disturbances, but this phenotype was not seen in the cardiomyocyte specific KO of the Asf/sf2-related splicing regulator Sc35 (Xu et al., 2005). *In vitro*, Sc35 has been described as a CaMKIIδ splicing regulator, but its *in vivo* relevance is not entirely clear (Han et al., 2011). Cardiac specific KO of Rbfox1 or Rbfox2 induces CaMKIIδ missplicing and these mice present with cardiomyopathy. However, this phenotype does not seem to depend on CaMKIIδ missplicing, at least in the Rbfox1 KO, as re-expression of proper Mef2 isoforms rescues the phenotype (Wei et al., 2015; Gao et al., 2016). The last known CaMKIIδ splicing regulator is Rbm20, of which KO leads to increased expression of CaMKIIδA and CaMKIIδ9, and disturbed intracellular Ca^{2+}-handling, similar to what is seen in cardiomyocyte specific Asf/sf2 KO mice (van den Hoogenhof et al., 2018).

Patients with mutations in RBM20 are at high risk of suffering from sustained ventricular arrhythmias, which might be related to changes in Ca^{2+} handling due to CaMKIIδ misplicing into CaMKIIδA and CaMKIIδ9 (van den Hoogenhof et al., 2018). Increased CaMKIIδA expression is associated with enhanced L-type Ca^{2+} current density (Xu et al., 2005; Koval et al., 2010). Later on, we will discuss the discoveries related to this isoform, but it illustrates the relevance to understand the functional consequences of different CaMKIIδ splice isoforms.

CaMKIIδB

Ca^{2+}/calmodulin dependent protein kinase II deltaB, characterized traditionally as the nuclear CaMKIIδ splice isoform, is together with CaMKIIδC the most studied CaMKIIδ splice isoform in the heart. It is the only CaMKIIδ splice isoform comprising exon 14, which is translated in a 11 amino acid insert. The first four amino acids of this exon correspond to the nuclear localization signal ^{332}KKRK, which locates CaMKIIδ to the nucleus (**Figure 2**; Ramirez et al., 1997). However, phosphorylation of CaMKIIδB on Thr-287 curiously localizes the protein in the cytoplasm, but this can be overcome by changing Ser-332 (immediately adjacent to the NLS) to a phosphoresistant Ala, which restores nuclear localization of CaMKIIδB (Backs et al., 2006). In addition, CaMKIIδB can be phosphorylated at this residue, possibly by CaMKI or CaMKIV, which induces nucleocytoplasmic shuttling and limits its nuclear localization (Heist et al., 1998; Backs et al., 2006).

Similar to other splice isoforms, CaMKIIδB expression varies in cardiac development and remodeling. For example, in cultured cardiac progenitor cells CaMKIIδB is upregulated during differentiation and compartmentalizes mostly in the nucleus, after which CaMKIIδB expression decreases in adulthood (Quijada et al., 2015). Interestingly, CaMKIIδB overexpression in cardiac progenitor cells reduces proliferation rate and increases cell size, which suggests that CaMKIIδB drives cardiac progenitor cell commitment (Quijada et al., 2015). However, *in vivo* lineage tracing experiments are required to further clarify the role of CaMKIIδB in this process. Furthermore, CaMKIIδB is upregulated in mouse hearts after myocardial infarction, and in failing human hearts (Hoch et al., 1999; Little et al., 2009; Quijada et al., 2015). In doxorubicin-induced cardiomyopathy and Rbm20 cardiomyopathy, on the other hand, CaMKIIδB is downregulated (Little et al., 2009; van den Hoogenhof et al., 2018). These studies show the versatility in CaMKIIδB expression in different stages in cardiomyocytes.

Because of its subcellular localization, CaMKIIδB has been linked to gene transcription regulation. In cardiomyocytes, CaMKIIδB overexpression in cardiomyocytes increases the activity of an atrial natriuretic factor (*ANP*) gene reporter as well as Anf protein expression (Ramirez et al., 1997). Interestingly, the use of chimera proteins of different CaMKII splice isoforms with the NLS encoded in exon 14 (which is normally only included in CaMKIIδB) similarly increased activity of the ANP gene reporter, suggesting that nuclear localization is sufficient to drive gene expression downstream of CaMKII activation (Ramirez et al., 1997). Furthermore, CaMKIIδB can phosphorylate nuclear epigenetic targets such as HDAC4 and HDAC5 and Histone 3 (at Ser-10 and Ser-28; Backs et al., 2006; Saadatmand et al., 2019). Phosphorylation of HDAC4 and HDAC5 increases their nuclear export, and is related with activation of a hypertrophic program downstream of MEF2 and GATA4 (Lu et al., 2010; Awad et al., 2013, 2015). Increased phosphorylation of Histone 3 is correlated with increased hemoglobin expression, but the functional relevance of this event is not yet known (Saadatmand et al., 2019). However, the premise that hemoglobin is not only expressed in cardiomyocytes, but also has a function in these cells is intriguing and deserves additional attention.

Multiple works also showed the protective effect of CaMKIIδB (Little et al., 2009; Peng et al., 2010). For example, overexpression of CaMKIIδB in primary neonatal cardiomyocytes suppresses apoptosis through increased expression of the anti-apoptotic modulator B-cell lymphoma 2 (Bcl-2) and phosphorylation and activation of HSF1 (Peng et al., 2010). Similarly, overexpression of CaMKIIδB in cardiac progenitor cells protects from oxidative stress damage, while knocking down CaMKIIδB increases cell death and expression of senescence markers (Quijada et al., 2015). In rats with Doxorubicin-induced cardiomyopathy, CaMKIIδB expression decreases, and together this suggests that CaMKIIδB is necessary to limit cardiac cell death (Little et al., 2009). It would be interesting to restore CaMKIIδB expression in this model, e.g., though AAV-mediated gene transfer, and see if this is protective. In line with this, CaMKIIδ-KO mice are protected from cardiac dysfunction in an ischemia/reperfusion injury model in mice, and transgenic CaMKIIδB overexpression in the CaMKIIδ-KO further protects by inhibiting the upregulation of pro-inflammatory cytokines and decreasing infarct size (Gray et al., 2017). Whether

CaMKIIδB is also involved in the regulation of the inflammatory response after non-ischemic stress remains to be seen, but this is clinically interesting as cardiac inflammation commonly occurs in heart failure patients (Elster et al., 1956; Suetomi et al., 2019; Adamo et al., 2020). CaMKIIδB overexpression is not only protective in ischemic stress, but also in cultured cardiomyocytes exposed to Angiotensin II. CaMKIIδB phosphorylates the transcription factor HSF1, and this increases the expression of its downstream target gene inducible (i)HSP70, which in turn suppresses stress-induced apoptotic signals of multiple Bcl-2 members and therefore acts as an anti-apoptotic chaperone (Peng et al., 2010).

Cardiomyocyte specific transgenic overexpression of CaMKIIδB *in vivo* promotes the development of cardiac hypertrophy at 4 months of age, even without an experimental intervention (Zhang et al., 2002). However, even though CaMKII activity is increased in these hearts, CaMKII-dependent phosphorylation of PLN (Thr-17) and RyR2 (Ser-2814) was not, and cardiac Ca^{2+} handling was equally not affected (Zhang et al., 2002, 2007). However, transgenic CaMKIIδB overexpression does increase HDAC4 phosphorylation, which shuttles HDAC4 out of the nucleus. Subsequently, MEF2 is derepressed, and transcription of pro-hypertrophic genes, such as Anf, Bnp, β-mhc and skeletal actin is increased (Backs et al., 2006; Little et al., 2007; Zhang et al., 2007). Summarizing, CaMKIIδB can be both detrimental and protective, depending on what downstream processes are activated. On the one hand, CaMKIIδB can activate pro-hypertrophic and detrimental HDAC4-MEF2 signaling, and CaMKIIδB transgenic mice have a, albeit relatively mild, cardiac phenotype (Zhang et al., 2002; See **Table 2**). On the other hand, CaMKIIδB expression protects against *in vivo* ischemic injury (Gray et al., 2017). Further research is needed to clarify when CaMKIIδB expression is protective, and when it is not. Additionally, the simultaneous activation of detrimental transcription factors such as MEF2 and protective transcription factors such as HSF1 suggests that transcription factors downstream CaMKIIδB do not function in only one direction, and that their interaction should be further explored.

CaMKIIδC

Classically known as the cytoplasmic CaMKIIδ isoform, CaMKIIδC lacks exons 14–16, and is therefore the smallest CaMKIIδ splice variant (**Figure 2**; Schworer et al., 1993). Although initial studies described this isoform exclusively in the cytoplasm (Edman and Schulman, 1994), in cardiomyocytes it is also associated with the SR membrane, the plasma membrane, and the nuclear membrane (Mishra et al., 2011; Ljubojevic-Holzer et al., 2020), where it possible translocates together with CaMKIIδB (Mishra et al., 2011). Therefore, even though this splice isoform lacks the NLS, it is nevertheless found in the nuclear compartment, which most likely relates to the composition of the holoenzyme, which includes (different ratios) of all CaMKII proteins and isoforms (Mishra et al., 2011). Together with CaMKIIδB, CaMKIIδC was initially thought to be the most expressed isoform in the heart, but recent results challenge this view and situate CaMKIIδC as only the third most highly expressed isoform in the human, rhesus monkey and rat heart (Zhang et al., 2019). CaMKIIδC expression increases in the postnatal heart after day 2 (Quijada et al., 2015), in mice exposed to TAC surgery (Zhang et al., 2003), and in patients with heart failure (Ljubojevic-Holzer et al., 2020).

The first study with CaMKIIδC overexpression showed that adenoviral CaMKIIδC overexpression in adult cardiomyocytes increases activation of β1-adrenergic-induced apoptosis (Zhu et al., 2003, 2007). Moreover, inhibition of CaMKII activity through the use of KN-93 or AIP, or the use of a dominant negative mutant of CaMKIIδC, blocks the increase in apoptosis induced by both the overexpression of CaMKIIδC or cell death stimuli (Zhu et al., 2007). CaMKIIδC also upregulates the proapoptotic transcription factor p53 *in vivo*, which in part could explain the effects of CaMKIIδC in cardiomyocyte apoptosis (Toko et al., 2010). Interestingly, these data suggest that CaMKIIδC and CaMKIIδB have opposite roles when it comes to apoptosis, as CaMKIIδC induces apoptosis while CaMKIIδB protects against apoptosis.

Cardiomyocyte specific overexpression of CaMKIIδC in mice leads to rapidly progressing cardiac hypertrophy that transitions to heart failure at 3 months of age (see **Table 2**), as well as a dysregulation of Ca^{2+} handling. This is accompanied by an increase in PLN (Thr-17) and RyR2 (Ser-2814) phosphorylation (Zhang et al., 2003), and a reduction in total protein of PLN, RyR2, SERCA2, and upregulation of the Na^{+}-Ca^{2+} exchanger (Maier et al., 2003). Even though increased PLN phosphorylation should increase SERCA function, this is offset by the decrease in SERCA and PLN protein levels. Therefore, decreased SERCA function and increased NCX function leads to a lower SR Ca^{2+} content and Ca^{2+} transient amplitude, demonstrating the involvement of CaMKIIδC in ECC regulation (Maier et al., 2003). Interestingly, cardiomyocyte specific overexpression of CaMKIIδC in a CaMKIIδ-KO background induces a more severe phenotype than overexpression of CaMKIIδC in a wildtype background (Ljubojevic-Holzer et al., 2020). This suggests that (the ratio of) different CaMKIIδ splice isoforms may modulate long-term dysfunction induced by CaMKIIδC overexpression in cardiomyocytes.

The detrimental effects of CaMKIIδC activation seems to depend on timing as well. Early activation of CaMKIIδC after TAC (i.e., 5 days after TAC) increases both SR Ca^{2+} content and release, possibly to compensate the afterload imposed by TAC (Ljubojevic-Holzer et al., 2020). Notably, in mice subjected to acute pressure overload, pharmacological inhibition of CaMKII shows an impairment of ECC and decreased survival, suggesting that inhibiting the early CaMKII response can be detrimental (Baier et al., 2020). These early adaptive effects are also seen in CaMKIIδC transgenic mice. For example, young mice overexpressing CaMKIIδC (6–8 weeks) show similar compensatory Ca^{2+} transient effects and do not present with cardiac dysfunction (Zhang et al., 2003; Ljubojevic-Holzer et al., 2020). However, over time these compensatory effects are reversed, both after TAC and in CaMKIIδC TG mice, and these mice progress to HF (at 45 days and 11–13 weeks, respectively; Ljubojevic-Holzer et al., 2020). This suggests that CaMKIIδC

TABLE 2 | Overview of *in vivo* studies of CaMKIIδ splice variant overexpression in cardiomyocytes.

Splicevariant	Mouse model	Intervention/Cardiac Phenotype	Ca^{2+} handling/Major signaling	References
CaMKIIδA	Cardiomyocyte specific transgenic overexpression (αMHC-driven)	• Cardiac hypertrophy at 4 weeks of age • Contraction defects • Death at 8 weeks of age	• Ca^{2+} handling defects	Xu et al., 2005
	Cardiomyocyte specific transgenic overexpression (αMHC-driven)	• Activation of hypertrophic gene program at 4 weeks of age • Cardiac hypertrophy and moderate cardiac dysfunction at 4 months of age	• Downregulation of PLN phosphorylation (Ser16/Thr17) • Decrease in SR Ca^{2+} uptake (*in vitro*) • Repression of HDAC4 and upregulation of MEF2 activity • Increase in PP2A activity	Zhang et al., 2002, 2007
CaMKIIδB	AAV9-mediated overexpression in cardiomyocytes in CaMKIIδ/γ (αMHC-driven) KO background	• Protection against myocardial damage and dysfunction following 1 hour *ex vivo* I/R	-	Weinreuter et al., 2014
	Cardiomyocyte specific transgenic overexpression (αMHC-driven) in cardiomyocytes (αMHC-driven) CaMKIIδ KO background	• Cardiac damage not affected after I/R injury for 1 day	• Upregulation of IL-6 mRNA	Gray et al., 2017
	Cardiomyocyte specific transgenic (αMHC-driven)	• Cardiac hypertrophy at 6 weeks of age • Heart failure at 12 weeks of age	• Ca^{2+} handling defects • Repression of HDAC4 and upregulation of MEF2 activity	Maier et al., 2003; Zhang et al., 2003, 2007; Ljubojevic-Holzer et al., 2020
	AAV9-mediated overexpression in cardiomyocytes in CaMKIIδ/γ (αMHC-driven) KO background	• Cardiac damage not affected after I/R injury for 1 day	-	Weinreuter et al., 2014
CaMKIIδC	Cardiomyocyte specific transgenic overexpression (αMHC-driven) in cardiomyocyte (αMHC-driven) CaMKIIδ KO background	• Exacerbated myocardial damage and dysfunction following 1 hour *ex vivo* I/R	• Upregulation of IKK-α/β phosphorylation and NF-κB protein content • Upregulation of TNF-α and IL-6 mRNA	Gray et al., 2017
	Cardiomyocyte specific transgenic (αMHC-driven) plus mitochondrial localization signal	• Modest cardiac hypertrophy (Without cardiomyocyte cross-sectional area changes) • Dilated cardiomyopathy	• MCU phosphorylation • Decrease in ATP content • Decreased mitochondrial complex I and increased complex II activity	Luczak et al., 2020
CaMKIIδ9	Cardiomyocyte specific transgenic (αMHC-driven)	• Cardiac hypertrophy at 6 weeks of age • Heart failure at 10 weeks of age	• Disruption of UBE2T-dependent DNA repair pathway • DNA damage	Zhang et al., 2019

is necessary for an early adaptive response, but detrimental at later stages. It raises the questions whether CaMKIIδ splicing itself is affected differently in different disease stages, and if it could be beneficial to redirect splicing toward isoforms that are specifically needed during different phases of disease (e.g., increased expression of CaMKIIδC in the early stage of disease).

Crossing CaMKIIδC transgenic mice with PLN KO mice (CaMKIIδC-TG/PLN-KO) attenuates the Ca^{2+} handling defects seen in transgenic mice overexpressing CaMKIIδC, which is in line with the hypothesis that the decreased SR Ca^{2+} uptake through SERCA underlies the Ca^{2+} handling defects. However, while the SR Ca^{2+}content and Ca^{2+} transient amplitude were normalized, left ventricular dilation, ventricular function, apoptosis, and mortality were exacerbated (Zhang et al., 2010). This can be explained in part by the increased phosphorylation of RyR2, which together with the restored SR Ca^{2+}content, increased the SR Ca^{2+} spark frequency and SR Ca^{2+} leak. Interestingly, inhibiting SR Ca^{2+} leak through the use of Ryanodine in CaMKIIδC-TG/PLN-KO cardiomyocytes improved the viability and prevented the increase in apoptosis (Zhang et al., 2010). On the other hand, CaMKIIδC transgenic mice crossed with mice expressing a SR-targeted autocamtide-2-related inhibitory peptide (SR-AIP) show a reduction in PLN and RyR2 phosphorylation, but no improvement in cardiac function, suggesting that the cardiac dysfunction in CaMKIIδC transgenic mice is not only dependent on phosphorylation of its SR-associated targets (Huke et al., 2011).

These effects on Ca^{2+} handling are distinct from the effects of CaMKIIδB overexpression, but overexpression of both splice variants similarly leads to HDAC4 phosphorylation (Ser-632), which inhibits nuclear import, and subsequently induces MEF2 transcriptional activity (Backs et al., 2006; Zhang et al., 2007). Additionally, Ljubojevic-Holzer et al. demonstrated recently that in mice subjected to TAC surgery and in failing human hearts the perinuclear and nuclear CaMKII population correspond majorly to CaMKIIδC, which suggests that CaMKIIδC also regulates gene expression in cardiac remodeling (Ljubojevic-Holzer et al., 2020). CaMKII phosphorylates class IIa HDACs, but overexpression of CaMKIIδC also affects class I HDAC activity (Zhang et al., 2020). CaMKIIδC increases HDAC1 activity, and mice overexpressing

CaMKIIδC show an increase in HDAC1 and HDAC3 expression. Interestingly, class I HDAC inhibition attenuates hypertrophy and maladaptive remodeling in CaMKIIδC transgenic mice, suggesting that, at least in part, the pro-hypertrophic effect of increased CaMKIIδC expression is dependent on class I HDAC activity (Zhang et al., 2020). Opposite to the protective effect of CaMKIIδB in I/R in mice, CaMKIIδC overexpression in a CaMKIIδ-KO background exacerbates cardiac dysfunction and increases infarct size. This effect is mediated by increased phosphorylation of IKK and nuclear localization of NF-κB, which is accompanied by an upregulation of TNFα and IL-6 in *ex vivo* hearts exposed to I/R (Gray et al., 2017). This suggests that CaMKIIδ mediates inflammasome activation and macrophage recruitment in non-ischemic injury through a transcriptional mechanism involving NF-κB, which is supported by multiple other studies (Weinreuter et al., 2014; Suetomi et al., 2018; Willeford et al., 2018). Thus, NF-κB activation represents another transcriptional pathway through which CaMKIIδ exerts its pathophysiological effects on the heart. However, while it is shown that CaMKIIδB and CaMKIIδC have opposite effects in this regard, the specific action of different CaMKIIδ splice isoforms is not entirely understood.

Ca^{2+}/calmodulin dependent protein kinase II is also implicated in the regulation of mitochondrial processes (Odagiri et al., 2009). After the observation that mitochondrial Ca^{2+} and apoptosis are elevated in mice overexpressing CaMKIIδC (Zhu et al., 2007, Zhang et al., 2010), the Anderson group demonstrated that mice with specific mitochondrial CaMKII inhibition are resistant to I/R injury and myocardial infarction (Joiner et al., 2012). Mitochondrial CaMKII inhibition decreased phosphorylation of the inner membrane mitochondrial Ca^{2+} uniporter (MCU), which is the transmembrane protein that allows the passage of Ca^{2+} from cytoplasm into the inner mitochondria (Joiner et al., 2012). The phosphorylation of MCU by CaMKII increases I_{MCU}, which in turn promotes mitochondrial permeability transition pore (mPTP) opening, triggering programmed cell death (Joiner et al., 2012; Halestrap and Richardson, 2015). However, this result is in discrepancy with a study from the Kirichok group, which using similar conditions, reported that I_{MCU} in cardiomyocytes is very small and is not directly regulated by CaMKII (Fieni et al., 2012, 2014). These studies were performed in mitoplasts with perfusion of constitutively active CaMKII, and the contribution of CaMKII in mitochondria *in vivo* will clarify whether the specific action of this enzyme on I_{MCU} underlies the protective effect of mitochondrial CaMKII inhibition in the prevention of cardiac dysfunction. Another recent study reports that CaMKIIγ/δ deletion does not affect mitochondrial I_{MCU} upon β-adrenergic or electrical stimulation in isolated cardiomyocytes, nor in cardiomyocyte-isolated mitochondria subjected to oxidative stress (Nickel et al., 2020). Mitochondria-targeted CaMKIIδC overexpression causes dilated cardiomyopathy, but with modest cardiac hypertrophy and no changes in cardiomyocyte cross-sectional area or cell death. It also reduces expression of assembled complex I, the mitochondrial isoform of creatine kinase and increases the production of NADH (Luczak et al., 2020). However, the finding that CaMKIIγ/δ deletion does not change redox state of mitochondria, measured by NADH accumulation, argues that while CaMKIIδC seems sufficient to increase NADH production, CaMKIIγ/δ as a whole is not necessary (Nickel et al., 2020). Another question that remains is related with the equal or differential action of cytoplasmic and mitochondrial CaMKII pools. Independent groups demonstrated the upregulation of CaMKIIδC in the mitochondrial fraction derived from hearts exposed to ischemia/reperfusion and the involvement of CaMKIIδC in mitochondrial Ca^{2+} content (Zhang et al., 2010; Weinreuter et al., 2014). In addition, in transgenic mice with cardiomyocyte sustained activation of Gαq signaling, which present with mitochondrial dysfunction and upregulation of mitochondrial ROS, the additional cardiomyocyte-specific deletion of CaMKIIδ attenuates these changes in mitochondrial function (Westenbrink et al., 2015). This suggests the involvement of CaMKIIδ in mitochondrial (dys)function upon sustained Gαq signaling. Overall, it is clear that both cytoplasmic and mitochondrial CaMKII pools can affect mitochondrial function, but the underlying mechanisms, and the specific actions of different CaMKIIδ isoforms, remain to be investigated. In conclusion, CaMKIIδC is involved in a multitude of processes in the development of heart failure and arrhythmias (Gray and Heller Brown, 2014). However, even though there are no studies on the physiological functions of CaMKIIδC specifically, it likely is the splice isoform that precisely modulates intracellular Ca^{2+} handling in ECC and ETC (Maier and Bers, 2007; Bers, 2011) and the heart rate in fight or flight response (Wu et al., 2009), suggesting that CaMKIIδC is also needed for normal heart function. In addition, the response of cardiomyocytes to the adaptive stimuli Insulin-like growth factor 1, exercise, or the vasoactive peptide alamandine show the dependence of positive effects on Ca^{2+} cycling and contractility on CaMKII activity (Beckendorf et al., 2018; Jesus et al., 2020). Therefore, also with this splice isoform, the question remains when it is detrimental and when it is not. It is hypothesized that spatio-temporal status plays a cardinal role here, and the answer is likely complex. Further research to understand the different functions of CaMKIIδC in different stages of remodeling are needed to fully understand this splice isoform.

CaMKIIδA

Ca^{2+}/calmodulin dependent protein kinase II delta A comprises exon 13 and 15–17 (**Figure 2**), was first identified as a neuronal CaMKIIδ splice isoform (Schworer et al., 1993), and was only later found in the heart (Xu et al., 2005). The expression of CaMKIIδA is developmentally regulated, with high expression in the neonatal heart (Xu et al., 2005). In the early postnatal period, CaMKIIδA is downregulated, and the adult CaMKIIδB and CaMKIIδC are upregulated. Additionally, CaMKIIδA also is upregulated in hearts from rats with chronic heart failure, in hypoxic cardiomyocytes, and in isoproterenol-treated mice (Li et al., 2011; Gui et al., 2018). CaMKIIδA is preferentially expressed at the T-tubules, in the perinuclear region, and at the intercalated disks (Xu et al., 2005). Unlike CaMKIIδB and CaMKIIδC, the localization of

CaMKIIδA only gave limited hints toward its possible functions. However, due to its expression at the *T*-tubules, it has been hypothesized that CaMKIIδA is involved in Ca^{2+} handling by regulating the LTCC. Developmentally, this makes sense, since the developing heart mostly relies on *L*-type calcium current for contraction, and the increased expression of CaMKIIδA in cardiac development supports this line of reasoning (Haddock et al., 1999). However, it must be noted that this hypothesis has not been proven. CaMKIIδA is also implicated in regulating Ca^{2+} handling through regulating RyR2 phosphorylation. In cultured cardiomyocytes, hypoxia increases the expression of CaMKIIδA, which increases RyR2 phosphorylation and SR Ca^{2+} leak, and decreases SERCA2a protein levels (Gui et al., 2018). Also in the case of gene transcription regulation, there is overlap between the different splice isoforms. CaMKIIδA overexpression can, similar to CaMKIIδB and CaMKIIδC, activate the HDAC4-MEF2 axis *in vitro* (Li et al., 2011), and thus induce MEF2-dependent hypertrophy. Whether this effect is reached in a similar way as by the other splice isoforms, or whether this works through a different mechanism, would be interesting to investigate. It is possible that CaMKIIδA, since it can be localized primarily at the *T*-tubules and the perinuclear space, increases transcription through regulating nuclear Ca^{2+}. Nuclear and perinuclear Ca^{2+} domains have been described to mediate transcriptional effects (Ibarra et al., 2013), and the involvement of the *T*-tubule associated CaMKIIδA might be an interesting target to study the relation between *T*-tubule signaling and transcription upon Ca^{2+} mobilization.

In the cardiomyocyte specific K.O. of the splicing factor Asf/sf2, CaMKIIδ is misspliced into CaMKIIδA, and this KO mouse presents with Ca^{2+} handling defects, a hypercontractile phenotype, and cardiomyopathy (Xu et al., 2005). To show that increased CaMKIIδA expression is sufficient to induce the Ca^{2+} handling defects seen in the Asf/sf2 KO mice, the authors engineered a mouse model with transgenic overexpression of CaMKIIδA in the heart, and this mouse phenocopied the Asf/sf2-deficient mice (Xu et al., 2005). RBM20 is another splicing regulator that regulates the splicing of CaMKIIδ. Rbm20 KO mice show a shift of CaMKIIδ toward the CaMKIIδA and CaMKIIδ9 splice variants, similar to the Asf/sf2 KO mouse model, and also present Ca^{2+} handling defects, including an increased *L*-type calcium current density (van den Hoogenhof et al., 2018). Overall, these data point toward a causal role for CaMKIIδA in the disturbed Ca^{2+}handling in these mouse models. However, unlike the changes seen in hypoxic cardiomyocytes with increased CaMKIIδA expression, in both the Asf/sf2 and Rbm20 KO mouse model, Ryr2 phosphorylation was not increased at the CaMKII-dependent phosphorylation site (Xu et al., 2005; van den Hoogenhof et al., 2018). Therefore, it remains to be investigated how increased CaMKIIδA leads to these Ca^{2+} handling changes. In conclusion, even though some of the downstream effects of CaMKIIδA seem similar to the effects of CaMKIIδC and CaMKIIδB, and there is a significant overlap in function, the phenotype of the transgenic mouse models is clearly different (see **Table 2**). Therefore, the how, when and where CaMKIIδA is upregulated in the heart, and what the functional consequences are, needs additional attention.

CaMKIIδ9

The CaMKIIδ9 splice isoform includes exon 16, but not exon 15, in the first variable domain (see **Figure 2**). It has long been overlooked, but some recent findings have put CaMKIIδ9 onto the radar of CaMKIIδ research. First, a recent paper demonstrated that CaMKIIδ9 is in fact the highest expressed isoform in the human heart, and is also highly expressed in hearts of other mammals such as mice, rabbits, and rhesus monkeys (Zhang et al., 2019). Second, in human RBM20 cardiomyopathy it is CaMKIIδ9 that is upregulated, and not CaMKIIδA, since it seems that CaMKIIδA is not expressed in humans (at least not in adulthood; van den Hoogenhof et al., 2018). CaMKIIδ missplicing is hypothesized to play a major role in the development of DCM caused by RBM20 mutations, and therefore it is important to delineate the functional relevance of CaMKIIδ9 (van den Hoogenhof et al., 2018). The expression of CaMKIIδ9 is upregulated in cultured cardiomyocytes exposed to doxorubicin, oxidative stress, in mice exposed to TAC surgery, and in human heart tissue from dilated cardiomyopathy patients (Zhang et al., 2019). CaMKIIδ9 is localized in the cytoplasm, like CaMKIIδC, and overexpression in cardiomyocytes induces cardiomyocyte cell death through DNA damage and genome instability. This process is mediated by phosphorylation of the ubiquitin E2 enzyme UBE2T, a ubiquitin ligase involved in DNA repair pathways (Machida et al., 2006), which induces its downregulation through proteasomal degradation (Zhang et al., 2019). In the physiological state of the cell, the cytosolic (but not nuclear) UBE2T is phosphorylated by CaMKIIδ9 and thus marked for degradation. Under stress conditions, CaMKIIδ9 expression increases in the heart, which in turn increases the degradation of UBE2T. This disrupts the UBE2T balance between cytosol and nucleus, which negatively affects the repair machinery of DNA and leads to increased DNA damage, genome instability and cell death (Zhang et al., 2019). Transgenic mice overexpressing CaMKIIδ9 present with cardiomyopathy and heart failure, which is attenuated by overexpression of UBE2T (see **Table 2**). Interestingly, other splice variants of CaMKIIδ do not show regulation of UBE2T signaling, suggesting that although CaMKIIδ9 has almost the same primary structure and cellular localization, it is more pathologically relevant than CaMKIIδA, -δB or -δC in DNA-repair pathways. It would be interesting to further investigate the regulation of other known CaMKIIδ-dependent processes by CaMKIIδ9, such as calcium handling and gene transcription. Overall, these new insights further highlight the importance of splice variant-specific signaling in heart disease.

CaMKIIδ AND ITS SPLICE VARIANTS AS THERAPEUTIC TARGETS

Approximately 518 kinases are encoded in the human genome (Cohen, 2000; Ficarro et al., 2002; Manning et al., 2002), phosphorylating one third of the proteome, so it is not surprising that they are involved in the pathogenesis of various autoimmune, inflammatory, nervous and cardiovascular

diseases, and cancer. The treatment of various types of cancer with small kinase inhibitors has been shown to be successful in clinical therapy, but this success has not been achieved in the cardiac field or experimentally tested for CaMKII. In total, the U.S. FDA has approved 52 small molecule protein kinase inhibitors by January 1, 2020, of which 46 are used in the treatment of neoplastic diseases, which shows the rather one-sided success (Roskoski, 2020). Several reasons for this restricted success in cardiac medicine can be outlined, starting with the historically preferred development of ion channel blockers against arrhythmias up to extremely high costs for clinical studies in the cardiovascular field (Fordyce et al., 2015). For CaMKII, the development of specific inhibitors is further complicated by the difference in function but similarity in structure of the different splice isoforms. Nevertheless, CaMKII inhibition has been viewed as a promising therapeutic approach, and multiple studies have shown promising effects. The use of CaMKII inhibitors, such as KN-93, AIP and GS-680, leads to improvement in cardiac function in maladaptive cardiac remodeling animal models and in human heart failure samples (Lebek et al., 2018). However, complete CaMKII inhibition might induce unwanted side effects if it does not specifically target the heart but also many other tissues. In addition, complete CaMKII inhibition affects all downstream pathways, and not only detrimental ones. Therefore, targeting CaMKII more selectively, either by inhibiting specific interactions (e.g., the interaction between CaMKII and a specific phosphorylation target), targeting specific tissues, or by targeting specific CaMKII isoforms, could be of great value. While small molecules and proteins dominate the previous approvals of therapeutics including in the cardiovascular area, the question arises whether these are suitable for CaMKII-directed therapy. Small molecules require a high research effort to achieve the required target specificity and efficiency. Since CaMKII-therapeutics might require an isoform-specific mode of action, the hurdles are even greater because the kinase domains are largely homologous. Protein-based therapeutics, on the other hand, can achieve high specificity, but size, stability and form of administration are challenging (Fosgerau and Hoffmann, 2015). However, the stability *in vivo* is weak due to proteolytic degradation, and their membrane impermeability makes it difficult to find a suitable administration route, as the bioavailable concentration is quickly reduced. While both approaches have been further advanced in recent years, new therapeutic concepts offer potential.

ANTISENSE OLIGONUCLEOTIDES TARGETING CaMKIIδ

One promising method is the use of antisense oligonucleotides (ASOs). ASOs in their original form are short, single-stranded, synthetically produced oligodeoxynucleotides that bind to mRNA targets in a complementary manner and thus modulate mRNA and protein expression (Dias and Stein, 2002). The first "naked" ASOs induced toxicity, and had an insufficient specificity as well as low biological activity (Rinaldi and Wood, 2018). In the last years, the design of RNA therapeutics has made significant advances leading to new generations of chemically modified ASOs. While early ASOs were only active via endonuclease (RNase H) mediated mRNA degradation (Eder et al., 1991), new ASOs can, for example, sterically block splicing factors or prevent ribosome recruitment to inhibit translation (Muntoni and Wood, 2011). This makes the use of ASOs a promising approach to not only decrease expression of targets, but also to redirect splicing. The first ASO therapy was approved by the FDA for the treatment of chorioretinitis, an inflammation of the retina caused by the cytomegalovirus (Roehr, 1998). The method has also successfully been tested in other diseases, such as Duchenne muscular dystrophy (Goemans et al., 2011), spinal muscular atrophy (Hua et al., 2011), and myotonic dystrophy (Wheeler et al., 2012). For the treatment of Duchenne muscular dystrophy, the therapeutic agent Eteplirsen was developed and approved by FDA in 2016, which promotes the production of a shortened, but active form of the dystrophin protein (Mendell et al., 2013). In theory, RNAi-based therapeutics are suitable for a CaMKII-directed treatment. Both systems, the knockdown of detrimental isoforms by mRNA degradation as well as targeting specific splice sites to redirect splicing could lead to success. Since loss of CaMKIIδ improves cardiac function, an RNAi-mediated knockdown of the entire gene would be the straightforward attempt for a therapy. In RBM20 cardiomyopathy, the overall level of CaMKIIδ is increased and at the same time CaMKIIδ splicing shifts toward CaMKIIδA and CaMKIIδ9. An RNAi-mediated knockdown of CaMKIIδA and CaMKIIδ9, e.g., by targeting exon 16, would be an appropriate approach. CaMKIIδ KO mice are protected from cardiac dysfunction in an ischemia/reperfusion injury model, with additional protective effects in case of parallel overexpression of CaMKIIδB. For patients suffering from acute myocardial infarction, RNAi therapeutics could be developed to redirect splicing in favor of CaMKIIδB, e.g., by blocking corresponding splice sites/factors. Alternatively, expressing exogenous CaMKIIδB through the use of gene therapy could have a protective effect and reduce the extent of dysfunction.

Probably the most important problem, however, is the delivery *in vivo*, since ASOs circulate in the bloodstream after intravenous or subcutaneous injection and can accumulate in the liver, kidney or spleen (Roberts et al., 2020) and may also be toxic. In addition, ASOs are large, hydrophilic polyanions that cannot easily cross the plasma membrane and must resist a variety of mechanisms, such as nuclease degradation in the extracellular space (Tsui et al., 2002) or removal by renal clearance (Iversen et al., 2013). Finally, it is currently not possible to direct ASOs to specific tissues such as the heart, but treatment is rather systemic. Due to these hurdles, it is not surprising that most approved RNAi therapeutics to date are limited to local administration (e.g., pegaptanib in the eye) or targeting the liver (due to the discontinuous sinusoidal endothelium, which allows for easy uptake of ASOs). Therefore, alternative and more advanced delivery systems have been developed in recent years and the pharmacokinetic properties of RNAi therapeutics have been improved by chemical modifications. Spherical nucleic acids

(Kapadia et al., 2018), exosome loading, and nanotechnological systems or "intelligent materials" are promising approaches (Roberts et al., 2020). Despite all innovations, it remains unclear whether the progress made in RNAi technology to date is sufficient to develop a successful CaMKIIδ-directed therapy. Apart from cardiomyocytes, CaMKIIδ is also the dominant isoform in smooth muscle cells (House et al., 2007) and endothelial cells (Wang et al., 2010; McCluskey et al., 2019; Dalal et al., 2021), so differences in cellular function need to be investigated to predict and avoid potential side effects. In general, it is important to further investigate the precise role of the different splice isoforms of CaMKIIδ, because the complex and diverse involvement of CaMKIIδ in various cellular and intracellular processes will require precise fine-tuning of therapeutics. In addition, CaMKII-directed therapeutics should consider the spatio-temporal distribution of the various splice isoforms. RNAi based therapeutics look more promising due to their high specificity, but the problem of cardiac-specific delivery needs further investigation and new solutions.

Overall, CaMKIIδ-directed therapeutics show great promise for three main reasons: (1) its levels are upregulated in both experimental models of heart failure and human samples which suggests a causal role for increased CaMKIIδ activation in disease, (2) the transgenic overexpression of CaMKIIδA, -δC and -δ9 leads to adverse cardiac remodeling accompanied by contractile dysfunction, and most importantly (3) CaMKIIδ inhibition is generally associated with protection against cardiac dysfunction and a reduction in cardiac damage in animal models subjected to detrimental stimuli. Nevertheless, the importance of CaMKIIδ in physiological processes, as well as the proposed protective effects of, e.g., CaMKIIδB suggest that CaMKIIδ-directed therapeutics require high specificity [For further details see (Nassal et al., 2020)].

CONCLUSION AND FUTURE DIRECTIONS

The first report of the role of Ca^{2+} in the heart came from work of Ringer in the nineteenth century who accidentally used tap water instead of distilled water in his NaCl solution, and then found out that Ca^{2+} was necessary for contraction (Ringer, 1883; Eisner, 2014). Since then, Ca^{2+} has been demonstrated to be crucial in numerous processes in the heart, one of which being Ca^{2+}-dependent signaling. CaMKII is one of the major transducers of specific spatio-temporal Ca^{2+} signaling, and has been the focus of many (cardiac) studies for decades (Clapham, 2007). These studies have led to tremendous insight into CaMKII-dependent signaling, but also to many new questions. One basic question that remains, is the function of the holoenzyme and how this is affected by the different splice variants. The dodecameric holoenzyme formation of CaMKII is distinct from structures of other protein kinases (Rosenberg et al., 2005). Based on the evolutionary conservation of this feature, it is thought that oligomerization plays a central role in the functioning of CaMKII. Phosphorylation spreading (autophosphorylation) within the holoenzyme and the exchange of activated and inactivated subunits between holoenzyme formations (Stratton et al., 2014) are only two special features offered by this structural setup. Investigating the structural features of CaMKII, and specifically CaMKIIδ, is essential for accurately addressing future therapeutics, and the comprehensive structure-function links have yet to be adequately discovered. However, modern technologies, such as molecular dynamics, offer novel tools and initial experiments have indicated, for example, that regulatory segments can spontaneously dock to the interfaces between hub subunits or that the length of the linker region controls the balance between activating or inhibitory autophosphorylation (Bhattacharyya et al., 2020; Karandur et al., 2020). It would be interesting to study how the dynamics of the enzyme structure changes during different periods in the cell, for example, during changing Ca^{2+}/CaM concentrations between Ca^{2+} pulse events. In addition, different splice isoform ratios within the oligomer and their effect on the structure and dynamics are not fully understood and should be investigated. In addition, the functional difference of the different splice isoforms of CaMKIIδ, and how to exploit these differences, is only partly understood. In line with this, the differential effects of the CaMKIIδ isoforms could be related with its interactome. Hence, unraveling of their possible interaction partners (such as PKA) could clarify why these isoforms exert different and/or redundant functions. Moreover, the 4 isoforms described in this review are not the only CaMKIIδ splice isoforms expressed in the heart. Even though less highly expressed, other splice isoforms such as CaMKIIδ4 are also found in the heart, and the function and relevance of these isoforms has not been explored. The fact that CaMKIIδ9, after being overlooked for a long time, turns out to be one of the highest expressed isoforms with a relevant function in the heart, illustrates that lesser known isoforms should not be neglected (Zhang et al., 2019). Another interesting question relates to the spatio-temporal control of activation of (splice variants of) CaMKIIδ. For example, CaMKIIδ's early response to detrimental stimuli is compensatory, but turns maladaptive in later stages. A more specific approach to CaMKIIδ inhibition, be it through targeting single splice isoforms, focusing on specific interactions, and on specific stages of CaMKIIδ activation will therefore be extremely valuable. In conclusion, tremendous progress has been made in our understanding of CaMKIIδ in the heart, but the picture is not yet complete. More specific studies will likely lead to new insights into this multifunctional and versatile protein, and could yield new therapeutic possibilities.

AUTHOR CONTRIBUTIONS

All authors contributed to writing and revising the manuscript.

FUNDING

MH was supported by grants from the German Center for Cardiovascular Research (DZHK; grant# 81X3500122), the German Research Foundation (DFG; grant# HO 6446/1-1), and the German Society of Cardiology (DGK; grant# DGK02/2019).

REFERENCES

Adamo, L., Rocha-Resende, C., Prabhu, S. D., and Mann, D. L. (2020). Reappraising the role of inflammation in heart failure. *Nat. Rev. Cardiol.* 17, 269–285. doi: 10.1038/s41569-019-0315-x

Anderson, M. E., Braun, A. P., Schulman, H., and Premack, B. A. (1994). Multifunctional Ca^{2+}/calmodulin-dependent protein kinase mediates Ca(2+)-induced enhancement of the L-type Ca^{2+} current in rabbit ventricular myocytes. *Circ. Res.* 75, 854–861. doi: 10.1161/01.res.75.5.854

Anderson, M. E., Brown, J. H., and Bers, D. M. (2011). CaMKII in myocardial hypertrophy and heart failure. *J. Mol. Cell Cardiol.* 51, 468–473. doi: 10.1016/j.yjmcc.2011.01.012

Awad, S., Al-Haffar, K. M., Marashly, Q., Quijada, P., Kunhi, M., Al-Yacoub, N., et al. (2015). Control of histone H3 phosphorylation by CaMKIIdelta in response to haemodynamic cardiac stress. *J. Pathol.* 235, 606–618. doi: 10.1002/path.4489

Awad, S., Kunhi, M., Little, G. H., Bai, Y., An, W., Bers, D., et al. (2013). Nuclear CaMKII enhances histone H3 phosphorylation and remodels chromatin during cardiac hypertrophy. *Nucleic Acids Res.* 41, 7656–7672. doi: 10.1093/nar/gkt500

Backs, J., Backs, T., Neef, S., Kreusser, M. M., Lehmann, L. H., Patrick, D. M., et al. (2009). The delta isoform of CaM kinase II is required for pathological cardiac hypertrophy and remodeling after pressure overload. *Proc. Natl. Acad. Sci. U.S.A.* 106, 2342–2347. doi: 10.1073/pnas.0813013106

Backs, J., Song, K., Bezprozvannaya, S., Chang, S., and Olson, E. N. (2006). CaM kinase II selectively signals to histone deacetylase 4 during cardiomyocyte hypertrophy. *J. Clin. Invest.* 116, 1853–1864. doi: 10.1172/JCI27438

Baier, M. J., Klatt, S., Hammer, K. P., Maier, L. S., and Rokita, A. G. (2020). Ca(2+)/calmodulin-dependent protein kinase II is essential in hyperacute pressure overload. *J. Mol. Cell Cardiol.* 138, 212–221. doi: 10.1016/j.yjmcc.2019.12.002

Beckendorf, J., van den Hoogenhof, M. M. G., and Backs, J. (2018). Physiological and unappreciated roles of CaMKII in the heart. *Basic Res. Cardiol.* 113:29. doi: 10.1007/s00395-018-0688-8

Beraldi, R., Li, X., Martinez Fernandez, A., Reyes, S., Secreto, F., Terzic, A., et al. (2014). Rbm20-deficient cardiogenesis reveals early disruption of RNA processing and sarcomere remodeling establishing a developmental etiology for dilated cardiomyopathy. *Hum. Mol. Genet.* 23, 3779–3791. doi: 10.1093/hmg/ddu091

Bers, D. M. (2011). Ca(2)(+)-calmodulin-dependent protein kinase II regulation of cardiac excitation-transcription coupling. *Heart Rhythm* 8, 1101–1104. doi: 10.1016/j.hrthm.2011.01.030

Bers, D. M., and Grandi, E. (2009). Calcium/calmodulin-dependent kinase II regulation of cardiac ion channels. *J. Cardiovasc. Pharmacol.* 54, 180–187. doi: 10.1097/FJC.0b013e3181a25078

Bhattacharyya, M., Lee, Y. K., Muratcioglu, S., Qiu, B., Nyayapati, P., Schulman, H., et al. (2020). Flexible linkers in CaMKII control the balance between activating and inhibitory autophosphorylation. *eLife* 9:e53670. doi: 10.7554/eLife.53670

Bo, T., Yamamori, T., Suzuki, M., Sakai, Y., Yamamoto, K., and Inanami, O. (2018). Calmodulin-dependent protein kinase II (CaMKII) mediates radiation-induced mitochondrial fission by regulating the phosphorylation of dynamin-related protein 1 (Drp1) at serine 616. *Biochem. Biophys. Res. Commun.* 495, 1601–1607. doi: 10.1016/j.bbrc.2017.12.012

Bonne, G., Carrier, L., Bercovici, J., Cruaud, C., Richard, P., Hainque, B., et al. (1995). Cardiac myosin binding protein-C gene splice acceptor site mutation is associated with familial hypertrophic cardiomyopathy. *Nat. Genet.* 11, 438–440. doi: 10.1038/ng1295-438

Brauch, K. M., Karst, M. L., Herron, K. J., de Andrade, M., Pellikka, P. A., Rodeheffer, R. J., et al. (2009). Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. *J. Am. Coll Cardiol.* 54, 930–941. doi: 10.1016/j.jacc.2009.05.038

Burgos, J. I., Yeves, A. M., Barrena, J. P., Portiansky, E. L., Vila-Petroff, M. G., and Ennis, I. L. (2017). Nitric oxide and CaMKII: critical steps in the cardiac contractile response To IGF-1 and swim training. *J. Mol. Cell Cardiol.* 112, 16–26. doi: 10.1016/j.yjmcc.2017.08.014

Cipolletta, E., Rusciano, M. R., Maione, A. S., Santulli, G., Sorriento, D., Del Giudice, C., et al. (2015). Targeting the CaMKII/ERK interaction in the heart prevents cardiac hypertrophy. *PLoS One* 10:e0130477. doi: 10.1371/journal.pone.0130477

Clapham, D. E. (2007). Calcium signaling. *Cell* 131, 1047–1058. doi: 10.1016/j.cell.2007.11.028

Cohen, P. (2000). The regulation of protein function by multisite phosphorylation–a 25 year update. *Trends Biochem. Sci.* 25, 596–601. doi: 10.1016/s0968-0004(00)01712-6

Colomer, J. M., Mao, L., Rockman, H. A., and Means, A. R. (2003). Pressure overload selectively up-regulates Ca^{2+}/calmodulin-dependent protein kinase II in vivo. *Mol. Endocrinol.* 17, 183–192. doi: 10.1210/me.2002-0350

Cook, S. G., Bourke, A. M., O'Leary, H., Zaegel, V., Lasda, E., Mize-Berge, J., et al. (2018). Analysis of the CaMKIIalpha and beta splice-variant distribution among brain regions reveals isoform-specific differences in holoenzyme formation. *Sci. Rep.* 8:5448. doi: 10.1038/s41598-018-23779-4

Dalal, P. J., Sullivan, D. P., Weber, E. W., Sacks, D. B., Gunzer, M., Grumbach, I. M., et al. (2021). Spatiotemporal restriction of endothelial cell calcium signaling is required during leukocyte transmigration. *J. Exp. Med.* 218:e20192378. doi: 10.1084/jem.20192378

Di Carlo, M. N., Said, M., Ling, H., Valverde, C. A., De Giusti, V. C., Sommese, L., et al. (2014). CaMKII-dependent phosphorylation of cardiac ryanodine receptors regulates cell death in cardiac ischemia/reperfusion injury. *J. Mol. Cell Cardiol.* 74, 274–283. doi: 10.1016/j.yjmcc.2014.06.004

Dias, N., and Stein, C. A. (2002). Antisense oligonucleotides: basic concepts and mechanisms. *Mol. Cancer Ther.* 1, 347–355.

Doran, A. C., Ozcan, L., Cai, B., Zheng, Z., Fredman, G., Rymond, C. C., et al. (2017). CAMKIIgamma suppresses an efferocytosis pathway in macrophages and promotes atherosclerotic plaque necrosis. *J. Clin. Invest.* 127, 4075–4089. doi: 10.1172/JCI94735

Eder, P. S., DeVine, R. J., Dagle, J. M., and Walder, J. A. (1991). Substrate specificity and kinetics of degradation of antisense oligonucleotides by a 3' exonuclease in plasma. *Antisense Res. Dev.* 1, 141–151. doi: 10.1089/ard.1991.1.141

Edman, C. F., and Schulman, H. (1994). Identification and characterization of delta B-CaM kinase and delta C-CaM kinase from rat heart, two new multifunctional Ca^{2+}/calmodulin-dependent protein kinase isoforms. *Biochim. Biophys. Acta* 1221, 89–101. doi: 10.1016/0167-4889(94)90221-6

Eisner, D. (2014). Calcium in the heart: from physiology to disease. *Exp. Physiol.* 99, 1273–1282. doi: 10.1113/expphysiol.2013.077305

Elster, S. K., Braunwald, E., and Wood, H. F. (1956). A study of C-reactive protein in the serum of patients with congestive heart failure. *Am. Heart J.* 51, 533–541. doi: 10.1016/0002-8703(56)90099-0

Erickson, J. R., Joiner, M. L., Guan, X., Kutschke, W., Yang, J., Oddis, C. V., et al. (2008). A dynamic pathway for calcium-independent activation of CaMKII by methionine oxidation. *Cell* 133, 462–474. doi: 10.1016/j.cell.2008.02.048

Erickson, J. R., Nichols, C. B., Uchinoumi, H., Stein, M. L., Bossuyt, J., and Bers, D. M. (2015). S-Nitrosylation induces both autonomous activation and inhibition of Calcium/Calmodulin-dependent protein kinase II delta. *J. Biol. Chem.* 290, 25646–25656. doi: 10.1074/jbc.M115.650234

Erickson, J. R., Pereira, L., Wang, L., Han, G., Ferguson, A., Dao, K., et al. (2013). Diabetic hyperglycaemia activates CaMKII and arrhythmias by O-linked glycosylation. *Nature* 502, 372–376. doi: 10.1038/nature12537

Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., et al. (2002). Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 20, 301–305. doi: 10.1038/nbt0302-301

Fieni, F., Johnson, D. E., Hudmon, A., and Kirichok, Y. (2014). Mitochondrial Ca^{2+} uniporter and CaMKII in heart. *Nature* 513, E1–E2. doi: 10.1038/nature13626

Fieni, F., Lee, S. B., Jan, Y. N., and Kirichok, Y. (2012). Activity of the mitochondrial calcium uniporter varies greatly between tissues. *Nat. Commun.* 3:1317. doi: 10.1038/ncomms2325

Fordyce, C. B., Roe, M. T., Ahmad, T., Libby, P., Borer, J. S., Hiatt, W. R., et al. (2015). Cardiovascular drug development: is it dead or just hibernating? *J. Am. Coll. Cardiol.* 65, 1567–1582. doi: 10.1016/j.jacc.2015.03.016

Fosgerau, K., and Hoffmann, T. (2015). Peptide therapeutics: current status and future directions. *Drug Discov. Today* 20, 122–128. doi: 10.1016/j.drudis.2014.10.003

Gaertner, T. R., Kolodziej, S. J., Wang, D., Kobayashi, R., Koomen, J. M., Stoops, J. K., et al. (2004). Comparative analyses of the three-dimensional structures and enzymatic properties of alpha, beta, gamma and delta isoforms of

Ca^{2+}-calmodulin-dependent protein kinase II. *J. Biol. Chem.* 279, 12484–12494. doi: 10.1074/jbc.M313597200

Gao, C., Ren, S., Lee, J. H., Qiu, J., Chapski, D. J., Rau, C. D., et al. (2016). RBFox1-mediated RNA splicing regulates cardiac hypertrophy and heart failure. *J. Clin. Invest.* 126, 195–206. doi: 10.1172/JCI84015

Gao, Z., Singh, M. V., Hall, D. D., Koval, O. M., Luczak, E. D., Joiner, M. L., et al. (2011). Catecholamine-independent heart rate increases require Ca^{2+}/calmodulin-dependent protein kinase II. *Circ. Arrhythm Electrophysiol.* 4, 379–387. doi: 10.1161/CIRCEP.110.961771

Goemans, N. M., Tulinius, M., van den Akker, J. T., Burm, B. E., Ekhart, P. F., Heuvelmans, N., et al. (2011). Systemic administration of PRO051 in Duchenne's muscular dystrophy. *N. Engl. J. Med.* 364, 1513–1522. doi: 10.1056/NEJMoa1011367

Gray, C. B., and Heller Brown, J. (2014). CaMKIIdelta subtypes: localization and function. *Front. Pharmacol.* 5:15. doi: 10.3389/fphar.2014.00015

Gray, C. B., Suetomi, T., Xiang, S., Mishra, S., Blackwood, E. A., Glembotski, C. C., et al. (2017). CaMKIIdelta subtypes differentially regulate infarct formation following ex vivo myocardial ischemia/reperfusion through NF-kappaB and TNF-alpha. *J. Mol. Cell Cardiol.* 103, 48–55. doi: 10.1016/j.yjmcc.2017.01.002

Gui, L., Guo, X., Zhang, Z., Xu, H., Ji, Y. W., Wang, R. J., et al. (2018). Activation of CaMKIIdeltaA promotes Ca(2+) leak from the sarcoplasmic reticulum in cardiomyocytes of chronic heart failure rats. *Acta Pharmacol. Sin.* 39, 1604–1612. doi: 10.1038/aps.2018.20

Guo, W., Schafer, S., Greaser, M. L., Radke, M. H., Liss, M., Govindarajan, T., et al. (2012). RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat. Med,* 18, 766–773. doi: 10.1038/nm.2693

Haddock, P. S., Coetzee, W. A., Cho, E., Porter, L., Katoh, H., Bers, D. M., et al. (1999). Subcellular $[Ca^{2+}]$i gradients during excitation-contraction coupling in newborn rabbit ventricular myocytes. *Circ. Res.* 85, 415–427. doi: 10.1161/01.res.85.5.415

Halestrap, A. P., and Richardson, A. P. (2015). The mitochondrial permeability transition: a current perspective on its identity and role in ischaemia/reperfusion injury. *J. Mol. Cell Cardiol,* 78, 129–141. doi: 10.1016/j.yjmcc.2014.08.018

Han, J., Ding, J. H., Byeon, C. W., Kim, J. H., Hertel, K. J., Jeong, S., et al. (2011). SR proteins induce alternative exon skipping through their activities on the flanking constitutive exons. *Mol. Cell Biol.* 31, 793–802. doi: 10.1128/MCB.01117-10

Hart, G. W., and Akimoto, Y. (2009). "The O-GlcNAc modification," in *Essentials of Glycobiology*, eds A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, et al. (Cold Spring Harbor NY: Cold Spring Harbor Laboratory Press).

Hartzell, H. C., and Glass, D. B. (1984). Phosphorylation of purified cardiac muscle C-protein by purified cAMP-dependent and endogenous Ca^{2+}-calmodulin-dependent protein kinases. *J. Biol. Chem.* 259, 15587–15596.

Hegyi, B., Bers, D. M., and Bossuyt, J. (2019). CaMKII signaling in heart diseases: emerging role in diabetic cardiomyopathy. *J. Mol. Cell Cardiol.* 127, 246–259. doi: 10.1016/j.yjmcc.2019.01.001

Heist, E. K., Srinivasan, M., and Schulman, H. (1998). Phosphorylation at the nuclear localization signal of Ca^{2+}/calmodulin-dependent protein kinase II blocks its nuclear targeting. *J. Biol. Chem.* 273, 19763–19771. doi: 10.1074/jbc.273.31.19763

Hidalgo, C. G., Chung, C. S., Saripalli, C., Methawasin, M., Hutchinson, K. R., Tsaprailis, G., et al. (2013). The multifunctional Ca(2+)/calmodulin-dependent protein kinase II delta (CaMKIIdelta) phosphorylates cardiac titin's spring elements. *J. Mol. Cell Cardiol.* 54, 90–97. doi: 10.1016/j.yjmcc.2012.11.012

Hoch, B., Meyer, R., Hetzer, R., Krause, E. G., and Karczewski, P. (1999). Identification and expression of delta-isoforms of the multifunctional Ca^{2+}/calmodulin-dependent protein kinase in failing and nonfailing human myocardium. *Circ. Res.* 84, 713–721. doi: 10.1161/01.res.84.6.713

Hoffman, L., Stein, R. A., Colbran, R. J., and McHaourab, H. S. (2011). Conformational changes underlying calcium/calmodulin-dependent protein kinase II activation. *EMBO J.* 30, 1251–1262. doi: 10.1038/emboj.2011.40

Holmberg, C. I., Hietakangas, V., Mikhailov, A., Rantanen, J. O., Kallio, M., Meinander, A., et al. (2001). Phosphorylation of serine 230 promotes inducible transcriptional activity of heat shock factor 1. *EMBO J.* 20, 3800–3810. doi: 10.1093/emboj/20.14.3800

House, S. J., Ginnan, R. G., Armstrong, S. E., and Singer, H. A. (2007). Calcium/calmodulin-dependent protein kinase II-delta isoform regulation of vascular smooth muscle cell proliferation. *Am. J. Physiol. Cell Physiol.* 292, C2276–C2287. doi: 10.1152/ajpcell.00606.2006

Hua, Y., Sahashi, K., Rigo, F., Hung, G., Horev, G., Bennett, C. F., et al. (2011). Peripheral SMN restoration is essential for long-term rescue of a severe spinal muscular atrophy mouse model. *Nature* 478, 123–126. doi: 10.1038/nature10485

Huke, S., Desantiago, J., Kaetzel, M. A., Mishra, S., Brown, J. H., Dedman, J. R., et al. (2011). SR-targeted CaMKII inhibition improves SR Ca(2)+ handling, but accelerates cardiac remodeling in mice overexpressing CaMKIIdeltaC. *J. Mol. Cell Cardiol.* 50, 230–238. doi: 10.1016/j.yjmcc.2010.10.014

Ibarra, C., Vicencio, J. M., Estrada, M., Lin, Y., Rocco, P., Rebellato, P., et al. (2013). Local control of nuclear calcium signaling in cardiac myocytes by perinuclear microdomains of sarcolemmal insulin-like growth factor 1 receptors. *Circ. Res.* 112, 236–245. doi: 10.1161/CIRCRESAHA.112.273839

Iversen, F., Yang, C., Dagnaes-Hansen, F., Schaffert, D. H., Kjems, J., and Gao, S. (2013). Optimized siRNA-PEG conjugates for extended blood circulation and reduced urine excretion in mice. *Theranostics* 3, 201–209. doi: 10.7150/thno.5743

Jesus, I. C. G., Mesquita, T. R. R., Monteiro, A. L. L., Parreira, A. B., Santos, A. K., Coelho, E. L. X., et al. (2020). Alamandine enhances cardiomyocyte contractility in hypertensive rats through a nitric oxide-dependent activation of CaMKII. *Am. J. Physiol. Cell Physiol.* 318, C740–C750. doi: 10.1152/ajpcell.00153.2019

Joiner, M. L., Koval, O. M., Li, J., He, B. J., Allamargot, C., Gao, Z., et al. (2012). CaMKII determines mitochondrial stress responses in heart. *Nature* 491, 269–273. doi: 10.1038/nature11444

Kapadia, C. H., Melamed, J. R., and Day, E. S. (2018). Spherical nucleic acid nanoparticles: therapeutic potential. *BioDrugs* 32, 297–309. doi: 10.1007/s40259-018-0290-5

Karandur, D., Bhattacharyya, M., Xia, Z., Lee, Y. K., Muratcioglu, S., McAffee, D., et al. (2020). Breakage of the oligomeric CaMKII hub by the regulatory segment of the kinase. *eLife* 9:e57784. doi: 10.7554/eLife.57784

Kelly, P. T., Weinberger, R. P., and Waxham, M. N. (1988). Active site-directed inhibition of Ca^{2+}/calmodulin-dependent protein kinase type II by a bifunctional calmodulin-binding peptide. *Proc. Natl. Acad. Sci. U.S.A.* 85, 4991–4995. doi: 10.1073/pnas.85.14.4991

Kemi, O. J., Ellingsen, O., Ceci, M., Grimaldi, S., Smith, G. L., Condorelli, G., et al. (2007). Aerobic interval training enhances cardiomyocyte contractility and Ca^{2+} cycling by phosphorylation of CaMKII and Thr-17 of phospholamban. *J. Mol. Cell Cardiol.* 43, 354–361. doi: 10.1016/j.yjmcc.2007.06.013

Kolodziej, S. J., Hudmon, A., Waxham, M. N., and Stoops, J. K. (2000). Three-dimensional reconstructions of calcium/calmodulin-dependent (CaM) kinase IIalpha and truncated CaM kinase IIalpha reveal a unique organization for its structural core and functional domains. *J. Biol. Chem.* 275, 14354–14359. doi: 10.1074/jbc.275.19.14354

Konstantinidis, K., Bezzerides, V. J., Lai, L., Isbell, H. M., Wei, A. C., Wu, Y., et al. (2020). MICAL1 constrains cardiac stress responses and protects against disease by oxidizing CaMKII. *J. Clin. Invest.* 130, 4663–4678. doi: 10.1172/JCI133181

Koval, O. M., Guan, X., Wu, Y., Joiner, M. L., Gao, Z., Chen, B., et al. (2010). CaV1.2 beta-subunit coordinates CaMKII-triggered cardiomyocyte death and afterdepolarizations. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4996–5000. doi: 10.1073/pnas.0913760107

Kreusser, M. M., Lehmann, L. H., Keranov, S., Hoting, M. O., Oehl, U., Kohlhaas, M., et al. (2014). Cardiac CaM Kinase II genes delta and gamma contribute to adverse remodeling but redundantly inhibit calcineurin-induced myocardial hypertrophy. *Circulation* 130, 1262–1273. doi: 10.1161/CIRCULATIONAHA.114.006185

Lai, Y., Nairn, A. C., Gorelick, F., and Greengard, P. (1987). Ca^{2+}/calmodulin-dependent protein kinase II: identification of autophosphorylation sites responsible for generation of Ca^{2+}/calmodulin-independence. *Proc. Natl. Acad. Sci. U.S.A.* 84, 5710–5714. doi: 10.1073/pnas.84.16.5710

Lebek, S., Plossl, A., Baier, M., Mustroph, J., Tarnowski, D., Lucht, C. M., et al. (2018). The novel CaMKII inhibitor GS-680 reduces diastolic SR Ca leak and prevents CaMKII-dependent pro-arrhythmic activity. *J. Mol. Cell Cardiol.* 118, 159–168. doi: 10.1016/j.yjmcc.2018.03.020

Li, C., Cai, X., Sun, H., Bai, T., Zheng, X., Zhou, X. W., et al. (2011). The deltaA isoform of calmodulin kinase II mediates pathological cardiac hypertrophy by

interfering with the HDAC4-MEF2 signaling pathway. *Biochem. Biophys. Res. Commun.* 409, 125–130. doi: 10.1016/j.bbrc.2011.04.128

Ling, H., Gray, C. B., Zambon, A. C., Grimm, M., Gu, Y., Dalton, N., et al. (2013). Ca^{2+}/Calmodulin-dependent protein kinase II delta mediates myocardial ischemia/reperfusion injury through nuclear factor-kappaB. *Circ. Res.* 112, 935–944. doi: 10.1161/CIRCRESAHA.112.276915

Ling, H., Zhang, T., Pereira, L., Means, C. K., Cheng, H., Gu, Y., et al. (2009). Requirement for Ca^{2+}/calmodulin-dependent kinase II in the transition from pressure overload-induced cardiac hypertrophy to heart failure in mice. *J. Clin. Invest.* 119, 1230–1240. doi: 10.1172/JCI38022

Little, G. H., Bai, Y., Williams, T., and Poizat, C. (2007). Nuclear calcium/calmodulin-dependent protein kinase IIdelta preferentially transmits signals to histone deacetylase 4 in cardiac cells. *J. Biol. Chem.* 282, 7219–7231. doi: 10.1074/jbc.M604281200

Little, G. H., Saw, A., Bai, Y., Dow, J., Marjoram, P., Simkhovich, B., et al. (2009). Critical role of nuclear calcium/calmodulin-dependent protein kinase IIdeltaB in cardiomyocyte survival in cardiomyopathy. *J. Biol. Chem.* 284, 24857–24868. doi: 10.1074/jbc.M109.003186

Ljubojevic-Holzer, S., Herren, A. W., Djalinac, N., Voglhuber, J., Morotti, S., Holzer, M., et al. (2020). CaMKIIdeltaC drives early adaptive Ca(2+) change and late eccentric cardiac hypertrophy. *Circ. Res.* 127, 1159–1178. doi: 10.1161/CIRCRESAHA.120.316947

Lu, C. S., Hodge, J. J., Mehren, J., Sun, X. X., and Griffith, L. C. (2003). Regulation of the Ca^{2+}/CaM-responsive pool of CaMKII by scaffold-dependent autophosphorylation. *Neuron* 40, 1185–1197. doi: 10.1016/s0896-6273(03)00786-4

Lu, S., Liao, Z., Lu, X., Katschinski, D. M., Mercola, M., Chen, J., et al. (2020). Hyperglycemia acutely increases cytosolic reactive oxygen species via O-linked GlcNAcylation and CaMKII activation in mouse ventricular myocytes. *Circ. Res.* 126, e80–e96. doi: 10.1161/CIRCRESAHA.119.316288

Lu, Y. M., Shioda, N., Yamamoto, Y., Han, F., and Fukunaga, K. (2010). Transcriptional upregulation of calcineurin Abeta by endothelin-1 is partially mediated by calcium/calmodulin-dependent protein kinase IIdelta3 in rat cardiomyocytes. *Biochim. Biophys. Acta* 1799, 429–441. doi: 10.1016/j.bbagrm.2010.02.004

Luczak, E. D., and Anderson, M. E. (2014). CaMKII oxidative activation and the pathogenesis of cardiac disease. *J. Mol. Cell Cardiol.* 73, 112–116. doi: 10.1016/j.yjmcc.2014.02.004

Luczak, E. D., Wu, Y., Granger, J. M., Joiner, M. A., Wilson, N. R., Gupta, A., et al. (2020). Mitochondrial CaMKII causes adverse metabolic reprogramming and dilated cardiomyopathy. *Nat. Commun.* 11:4416. doi: 10.1038/s41467-020-18165-6

Maatz, H., Jens, M., Liss, M., Schafer, S., Heinig, M., Kirchner, M., et al. (2014). RNA-binding protein RBM20 represses splicing to orchestrate cardiac pre-mRNA processing. *J. Clin. Invest.* 124, 3419–3430. doi: 10.1172/JCI74523

Machida, Y. J., Machida, Y., Chen, Y., Gurtan, A. M., Kupfer, G. M., D'Andrea, A. D., et al. (2006). UBE2T is the E2 in the Fanconi anemia pathway and undergoes negative autoregulation. *Mol. Cell* 23, 589–596. doi: 10.1016/j.molcel.2006.06.024

Maggioni, A. P. (2015). Epidemiology of heart failure in Europe. *Heart Fail Clin.* 11, 625–635. doi: 10.1016/j.hfc.2015.07.015

Maier, L. S., and Bers, D. M. (2007). Role of Ca^{2+}/calmodulin-dependent protein kinase (CaMK) in excitation-contraction coupling in the heart. *Cardiovasc. Res.* 73, 631–640. doi: 10.1016/j.cardiores.2006.11.005

Maier, L. S., Zhang, T., Chen, L., DeSantiago, J., Brown, J. H., and Bers, D. M. (2003). Transgenic CaMKIIdeltaC overexpression uniquely alters cardiac myocyte Ca^{2+} handling: reduced SR Ca^{2+} load and activated SR Ca^{2+} release. *Circ. Res.* 92, 904–911. doi: 10.1161/01.RES.0000069685.20258.F1

Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science* 298, 1912–1934. doi: 10.1126/science.1075762

Martin, T. P., McCluskey, C., Cunningham, M. R., Beattie, J., Paul, A., and Currie, S. (2018). CaMKIIdelta interacts directly with IKKbeta and modulates NF-kappaB signalling in adult cardiac fibroblasts. *Cell Signal* 51, 166–175. doi: 10.1016/j.cellsig.2018.07.008

Mattiazzi, A., and Kranias, E. G. (2014). The role of CaMKII regulation of phospholamban activity in heart disease. *Front. Pharmacol.* 5:5. doi: 10.3389/fphar.2014.00005

McCluskey, C., Mooney, L., Paul, A., and Currie, S. (2019). Compromised cardiovascular function in aged rats corresponds with increased expression and activity of calcium/calmodulin dependent protein kinase IIdelta in aortic endothelium. *Vascul Pharmacol.* 11:106560. doi: 10.1016/j.vph.2019.04.002

Mendell, J. R., Rodino-Klapac, L. R., Sahenk, Z., Roush, K., Bird, L., Lowes, L. P., et al. (2013). Eteplirsen for the treatment of duchenne muscular dystrophy. *Ann. Neurol.* 74, 637–647. doi: 10.1002/ana.23982

Meyer, T., Hanson, P. I., Stryer, L., and Schulman, H. (1992). Calmodulin trapping by calcium-calmodulin-dependent protein kinase. *Science* 256, 1199–1202. doi: 10.1126/science.256.5060.1199

Mika, D., Richter, W., and Conti, M. (2015). A CaMKII/PDE4D negative feedback regulates cAMP signaling. *Proc. Natl. Acad. Sci. U.S.A.* 112, 2023–2028. doi: 10.1073/pnas.1419992112

Mishra, S., Gray, C. B., Miyamoto, S., Bers, D. M., and Brown, J. H. (2011). Location matters: clarifying the concept of nuclear and cytosolic CaMKII subtypes. *Circ. Res.* 109, 1354–1362. doi: 10.1161/CIRCRESAHA.111.248401

Muntoni, F., and Wood, M. J. (2011). Targeting RNA to treat neuromuscular disease. *Nat. Rev. Drug Discov.* 10, 621–637. doi: 10.1038/nrd3459

Murthy, S., Koval, O. M., Ramiro Diaz, J. M., Kumar, S., Nuno, D., Scott, J. A., et al. (2017). Endothelial CaMKII as a regulator of eNOS activity and NO-mediated vasoreactivity. *PLoS One* 12:e0186311. doi: 10.1371/journal.pone.0186311

Myers, J. B., Zaegel, V., Coultrap, S. J., Miller, A. P., Bayer, K. U., and Reichow, S. L. (2017). The CaMKII holoenzyme structure in activation-competent conformations. *Nat. Commun.* 8:15742. doi: 10.1038/ncomms15742

Nassal, D., Gratz, D., and Hund, T. J. (2020). Challenges and opportunities for therapeutic targeting of calmodulin kinase II in heart. *Front. Pharmacol.* 11:35. doi: 10.3389/fphar.2020.00035

Nickel, A. G., Kohlhaas, M., Bertero, E., Wilhelm, D., Wagner, M., Sequeira, V., et al. (2020). CaMKII does not control mitochondrial Ca(2+) uptake in cardiac myocytes. *J. Physiol.* 598, 1361–1376. doi: 10.1113/JP276766

Odagiri, K., Katoh, H., Kawashima, H., Tanaka, T., Ohtani, H., Saotome, M., et al. (2009). Local control of mitochondrial membrane potential, permeability transition pore and reactive oxygen species by calcium and calmodulin in rat ventricular myocytes. *J. Mol. Cell Cardiol.* 46, 989–997. doi: 10.1016/j.yjmcc.2008.12.022

Pandey, D., Gratton, J. P., Rafikov, R., Black, S. M., and Fulton, D. J. (2011). Calcium/calmodulin-dependent kinase II mediates the phosphorylation and activation of NADPH oxidase 5. *Mol. Pharmacol.* 80, 407–415. doi: 10.1124/mol.110.070193

Parikh, V. N., Caleshu, C., Reuter, C., Lazzeroni, L. C., Ingles, J., Garcia, J., et al. (2019). Regional variation in RBM20 causes a highly penetrant arrhythmogenic cardiomyopathy. *Circ. Heart Fail.* 12:e005371. doi: 10.1161/CIRCHEARTFAILURE.118.005371

Passier, R., Zeng, H., Frey, N., Naya, F. J., Nicol, R. L., McKinsey, T. A., et al. (2000). CaM kinase signaling induces cardiac hypertrophy and activates the MEF2 transcription factor in vivo. *J. Clin. Invest.* 105, 1395–1406. doi: 10.1172/JCI8551

Patton, B. L., Miller, S. G., and Kennedy, M. B. (1990). Activation of type II calcium/calmodulin-dependent protein kinase by Ca^{2+}/calmodulin is inhibited by autophosphorylation of threonine within the calmodulin-binding domain. *J. Biol. Chem.* 265, 11204–11212.

Peng, W., Zhang, Y., Zheng, M., Cheng, H., Zhu, W., Cao, C. M., et al. (2010). Cardioprotection by CaMKII-deltaB is mediated by phosphorylation of heat shock factor 1 and subsequent expression of inducible heat shock protein 70. *Circ. Res.* 106, 102–110. doi: 10.1161/CIRCRESAHA.109.210914

Quijada, P., Hariharan, N., Cubillo, J. D., Bala, K. M., Emathinger, J. M., Wang, B. J., et al. (2015). Nuclear calcium/Calmodulin-dependent protein kinase II signaling enhances cardiac progenitor cell survival and cardiac lineage commitment. *J. Biol. Chem.* 290, 25411–25426. doi: 10.1074/jbc.M115.657775

Ramirez, M. T., Zhao, X. L., Schulman, H., and Brown, J. H. (1997). The nuclear deltaB isoform of Ca^{2+}/calmodulin-dependent protein kinase II regulates atrial natriuretic factor gene expression in ventricular myocytes. *J. Biol. Chem.* 272, 31203–31208. doi: 10.1074/jbc.272.49.31203

Refaat, M. M., Lubitz, S. A., Makino, S., Islam, Z., Frangiskakis, J. M., Mehdi, H., et al. (2012). Genetic variation in the alternative splicing regulator RBM20 is associated with dilated cardiomyopathy. *Heart Rhythm* 9, 390–396. doi: 10.1016/j.hrthm.2011.10.016

Rinaldi, C., and Wood, M. J. A. (2018). Antisense oligonucleotides: the next frontier for treatment of neurological disorders. *Nat. Rev. Neurol.* 14, 9–21. doi: 10.1038/nrneurol.2017.148

Ringer, S. (1883). A further contribution regarding the influence of the different constituents of the blood on the contraction of the heart. *J. Physiol.* 4, 29–42.3. doi: 10.1113/jphysiol.1883.sp000120

Roberts, T. C., Langer, R., and Wood, M. J. A. (2020). Advances in oligonucleotide drug delivery. *Nat. Rev. Drug Discov.* 19, 673–694. doi: 10.1038/s41573-020-0075-7

Roehr, B. (1998). Fomivirsen approved for CMV retinitis. *J. Int. Assoc. Physic. AIDS Care* 4, 14–16.

Rosenberg, O. S., Deindl, S., Sung, R. J., Nairn, A. C., and Kuriyan, J. (2005). Structure of the autoinhibited kinase domain of CaMKII and SAXS analysis of the holoenzyme. *Cell* 123, 849–860. doi: 10.1016/j.cell.2005.10.029

Roskoski, R. Jr. (2020). Properties of FDA-approved small molecule protein kinase inhibitors: a 2020 update. *Pharmacol. Res.* 152:104609. doi: 10.1016/j.phrs.2019.104609

Saadatmand, A. R., Sramek, V., Weber, S., Finke, D., Dewenter, M., Sticht, C., et al. (2019). CaM kinase II regulates cardiac hemoglobin expression through histone phosphorylation upon sympathetic activation. *Proc. Natl. Acad. Sci. U.S.A.* 116, 22282–22287. doi: 10.1073/pnas.1816521116

Schworer, C. M., Rothblum, L. I., Thekkumkara, T. J., and Singer, H. A. (1993). Identification of novel isoforms of the delta subunit of Ca^{2+}/calmodulin-dependent protein kinase II. Differential expression in rat brain and aorta. *J. Biol. Chem.* 268, 14443–14449.

Singer, H. A., Benscoter, H. A., and Schworer, C. M. (1997). Novel Ca^{2+}/calmodulin-dependent protein kinase II gamma-subunit variants expressed in vascular smooth muscle, brain, and cardiomyocytes. *J. Biol. Chem.* 272, 9393–9400. doi: 10.1074/jbc.272.14.9393

Srinivasan, M., Edman, C. F., and Schulman, H. (1994). Alternative splicing introduces a nuclear localization signal that targets multifunctional CaM kinase to the nucleus. *J. Cell Biol.* 126, 839–852. doi: 10.1083/jcb.126.4.839

Strack, S., Barban, M. A., Wadzinski, B. E., and Colbran, R. J. (1997). Differential inactivation of postsynaptic density-associated and soluble Ca^{2+}/calmodulin-dependent protein kinase II by protein phosphatases 1 and 2A. *J. Neurochem.* 68, 2119–2128. doi: 10.1046/j.1471-4159.1997.68052119.x

Stratton, M., Lee, I. H., Bhattacharyya, M., Christensen, S. M., Chao, L. H., Schulman, H., et al. (2014). Activation-triggered subunit exchange between CaMKII holoenzymes facilitates the spread of kinase activity. *eLife* 3:e01610. doi: 10.7554/eLife.01610

Suetomi, T., Miyamoto, S., and Brown, J. H. (2019). Inflammation in nonischemic heart disease: initiation by cardiomyocyte CaMKII and NLRP3 inflammasome signaling. *Am. J. Physiol. Heart Circ. Physiol.* 317, H877–H890. doi: 10.1152/ajpheart.00223.2019

Suetomi, T., Willeford, A., Brand, C. S., Cho, Y., Ross, R. S., Miyamoto, S., et al. (2018). Inflammation and NLRP3 inflammasome activation initiated in response to pressure overload by Ca(2+)/Calmodulin-dependent protein kinase II delta signaling in cardiomyocytes are essential for adverse cardiac remodeling. *Circulation* 138, 2530–2544. doi: 10.1161/CIRCULATIONAHA.118.034621

Sun, P., Enslen, H., Myung, P. S., and Maurer, R. A. (1994). Differential activation of CREB by Ca^{2+}/calmodulin-dependent protein kinases type II and type IV involves phosphorylation of a site that negatively regulates activity. *Genes Dev.* 8, 2527–2539. doi: 10.1101/gad.8.21.2527

Swulius, M. T., and Waxham, M. N. (2008). Ca(2+)/calmodulin-dependent protein kinases. *Cell Mol. Life Sci.* 65, 2637–2657. doi: 10.1007/s00018-008-8086-2

Takemoto-Kimura, S., Suzuki, K., Horigane, S. I., Kamijo, S., Inoue, M., Sakamoto, M., et al. (2017). Calmodulin kinases: essential regulators in health and disease. *J. Neurochem.* 141, 808–818. doi: 10.1111/jnc.14020

Tazi, J., Bakkour, N., and Stamm, S. (2009). Alternative splicing and disease. *Biochim. Biophys. Acta* 1792, 14–26. doi: 10.1016/j.bbadis.2008.09.017

Thierfelder, L., Watkins, H., MacRae, C., Lamas, R., McKenna, W., Vosberg, H. P., et al. (1994). Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere. *Cell* 77, 701–712. doi: 10.1016/0092-8674(94)90054-x

Toko, H., Takahashi, H., Kayama, Y., Oka, T., Minamino, T., Okada, S., et al. (2010). Ca^{2+}/calmodulin-dependent kinase IIdelta causes heart failure by accumulation of p53 in dilated cardiomyopathy. *Circulation* 122, 891–899. doi: 10.1161/CIRCULATIONAHA.109.935296

Tsui, N. B., Ng, E. K., and Lo, Y. M. (2002). Stability of endogenous and added RNA in blood specimens, serum, and plasma. *Clin. Chem.* 48, 1647–1653. doi: 10.1093/clinchem/48.10.1647

van den Hoogenhof, M. M., Pinto, Y. M., and Creemers, E. E. (2016). RNA splicing: regulation and dysregulation in the heart. *Circ. Res.* 118, 454–468. doi: 10.1161/CIRCRESAHA.115.307872

van den Hoogenhof, M. M. G., Beqqali, A., Amin, A. S., van der Made, I., Aufiero, S., Khan, M. A. F., et al. (2018). RBM20 mutations induce an arrhythmogenic dilated cardiomyopathy related to disturbed calcium handling. *Circulation* 138, 1330–1342. doi: 10.1161/CIRCULATIONAHA.117.031947

van Oort, R. J., McCauley, M. D., Dixit, S. S., Pereira, L., Yang, Y., Respress, J. L., et al. (2010). Ryanodine receptor phosphorylation by calcium/calmodulin-dependent protein kinase II promotes life-threatening ventricular arrhythmias in mice with heart failure. *Circulation* 122, 2669–2679. doi: 10.1161/CIRCULATIONAHA.110.982298

Vila-Petroff, M., Mundina-Weilenmann, C., Lezcano, N., Snabaitis, A. K., Huergo, M. A., Valverde, C. A., et al. (2010). Ca(2+)/calmodulin-dependent protein kinase II contributes to intracellular pH recovery from acidosis via Na(+)/H(+) exchanger activation. *J. Mol. Cell Cardiol.* 49, 106–112. doi: 10.1016/j.yjmcc.2009.12.007

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., et al. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476. doi: 10.1038/nature07509

Wang, Z., Ginnan, R., Abdullaev, I. F., Trebak, M., Vincent, P. A., and Singer, H. A. (2010). Calcium/Calmodulin-dependent protein kinase II delta 6 (CaMKIIdelta6) and RhoA involvement in thrombin-induced endothelial barrier dysfunction. *J. Biol. Chem.* 285, 21303–21312. doi: 10.1074/jbc.M110.120790

Wei, C., Qiu, J., Zhou, Y., Xue, Y., Hu, J., Ouyang, K., et al. (2015). Repression of the central splicing regulator RBFox2 Is functionally linked to pressure overload-induced heart failure. *Cell Rep.* 10, 1521–1533. doi: 10.1016/j.celrep.2015.02.013

Weinreuter, M., Kreusser, M. M., Beckendorf, J., Schreiter, F. C., Leuschner, F., Lehmann, L. H., et al. (2014). CaM Kinase II mediates maladaptive post-infarct remodeling and pro-inflammatory chemoattractant signaling but not acute myocardial ischemia/reperfusion injury. *EMBO Mol. Med.* 6, 1231–1245. doi: 10.15252/emmm.201403848

Westenbrink, B. D., Ling, H., Divakaruni, A. S., Gray, C. B., Zambon, A. C., Dalton, N. D., et al. (2015). Mitochondrial reprogramming induced by CaMKIIdelta mediates hypertrophy decompensation. *Circ. Res.* 116, e28–e39. doi: 10.1161/CIRCRESAHA.116.304682

Wheeler, T. M., Leger, A. J., Pandey, S. K., MacLeod, A. R., Nakamori, M., Cheng, S. H., et al. (2012). Targeting nuclear RNA for in vivo correction of myotonic dystrophy. *Nature* 488, 111–115. doi: 10.1038/nature11362

Willeford, A., Suetomi, T., Nickle, A., Hoffman, H. M., Miyamoto, S., and Heller Brown, J. (2018). CaMKIIdelta-mediated inflammatory gene expression and inflammasome activation in cardiomyocytes initiate inflammation and induce fibrosis. *JCI Insight* 3:e97054. doi: 10.1172/jci.insight.97054

Witcher, D. R., Kovacs, R. J., Schulman, H., Cefali, D. C., and Jones, L. R. (1991). Unique phosphorylation site on the cardiac ryanodine receptor regulates calcium channel activity. *J. Biol. Chem.* 266, 11144–11152.

Wu, Y., Gao, Z., Chen, B., Koval, O. M., Singh, M. V., Guan, X., et al. (2009). Calmodulin kinase II is required for fight or flight sinoatrial node physiology. *Proc. Natl. Acad. Sc.i U.S.A.* 106, 5972–5977. doi: 10.1073/pnas.0806422106

Xu, X., Yang, D., Ding, J. H., Wang, W., Chu, P. H., Dalton, N. D., et al. (2005). ASF/SF2-regulated CaMKIIdelta alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. *Cell* 120, 59–72. doi: 10.1016/j.cell.2004.11.036

Zhang, M., Gao, H., Liu, D., Zhong, X., Shi, X., Yu, P., et al. (2019). CaMKII-delta9 promotes cardiomyopathy through disrupting UBE2T-dependent DNA repair. *Nat. Cell Biol.* 21, 1152–1163. doi: 10.1038/s41556-019-0380-8

Zhang, M., Yang, X., Zimmerman, R. J., Wang, Q., Ross, M. A., Granger, J. M., et al. (2020). CaMKII exacerbates heart failure progression by activating class I HDACs. *J. Mol. Cell Cardiol* 149, 73–81. doi: 10.1016/j.yjmcc.2020.09.007

Zhang, T., Guo, T., Mishra, S., Dalton, N. D., Kranias, E. G., Peterson, K. L., et al. (2010). Phospholamban ablation rescues sarcoplasmic reticulum Ca(2+) handling but exacerbates cardiac dysfunction in CaMKIIdelta(C) transgenic mice. *Circ. Res.* 106, 354–362. doi: 10.1161/CIRCRESAHA.109.207423

Zhang, T., Johnson, E. N., Gu, Y., Morissette, M. R., Sah, V. P., Gigena, M. S., et al. (2002). The cardiac-specific nuclear delta(B) isoform of Ca^{2+}/calmodulin-dependent protein kinase II induces hypertrophy and dilated cardiomyopathy associated with increased protein phosphatase 2A activity. *J. Biol. Chem.* 277, 1261–1267. doi: 10.1074/jbc.M108525200

Zhang, T., Kohlhaas, M., Backs, J., Mishra, S., Phillips, W., Dybkova, N., et al. (2007). CaMKIIdelta isoforms differentially affect calcium handling but similarly regulate HDAC/MEF2 transcriptional responses. *J. Biol. Chem.* 282, 35078–35087. doi: 10.1074/jbc.M707083200

Zhang, T., Maier, L. S., Dalton, N. D., Miyamoto, S., Ross, J. Jr., Bers, D. M., et al. (2003). The deltaC isoform of CaMKII is activated in cardiac hypertrophy and induces dilated cardiomyopathy and heart failure. *Circ. Res.* 92, 912–919. doi: 10.1161/01.RES.0000069686.31472.C5

Zhu, W., Woo, A. Y., Yang, D., Cheng, H., Crow, M. T., and Xiao, R. P. (2007). Activation of CaMKIIdeltaC is a common intermediate of diverse death stimuli-induced heart muscle cell apoptosis. *J. Biol. Chem.* 282, 10833–10839. doi: 10.1074/jbc.M611507200

Zhu, W. Z., Wang, S. Q., Chakir, K., Yang, D., Zhang, T., Brown, J. H., et al. (2003). Linkage of beta1-adrenergic stimulation to apoptotic heart cell death through protein kinase A-independent activation of Ca^{2+}/calmodulin kinase II. *J. Clin. Invest.* 111, 617–625. doi: 10.1172/JCI16326

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

frontiers in Cell and Developmental Biology

ORIGINAL RESEARCH
published: 29 April 2021
doi: 10.3389/fcell.2021.652032

circDENND1B Participates in the Antiatherosclerotic Effect of IL-1β Monoclonal Antibody in Mouse by Promoting Cholesterol Efflux via miR-17-5p/Abca1 Axis

Fei Xu[1,2], Li Shen[1,2], Han Chen[1,2], Rui Wang[1,2], Tongtong Zang[1,2], Juying Qian[1,2]* and Junbo Ge[1,2]**

[1] Department of Cardiology, Zhongshan Hospital, Fudan University, Research Unit of Cardiovascular Techniques and Devices, Chinese Academy of Medical Sciences, Shanghai, China, [2] National Clinical Research Center for Interventional Medicine, Shanghai, China

OPEN ACCESS

Edited by:
Hamid El Azzouzi, Erasmus Medical Center, Netherlands

Reviewed by:
Zhiguang Su, Sichuan University, China
Ting Ni, Fudan University, China

***Correspondence:**
Junbo Ge
ge.junbo2@zs-hospital.sh.cn
Juying Qian
qian.juying@zs-hospital.sh.cn
Li Shen
nt_sunny@163.com

Specialty section:
This article was submitted to Molecular Medicine, a section of the journal Frontiers in Cell and Developmental Biology

Received: *11 January 2021*
Accepted: *06 April 2021*
Published: *29 April 2021*

Citation:
Xu F, Shen L, Chen H, Wang R, Zang T, Qian J and Ge J (2021) circDENND1B Participates in the Antiatherosclerotic Effect of IL-1β Monoclonal Antibody in Mouse by Promoting Cholesterol Efflux via miR-17-5p/Abca1 Axis. Front. Cell Dev. Biol. 9:652032. doi: 10.3389/fcell.2021.652032

Inflammation is a crucial mediator of atherosclerosis, and several therapeutic methods that focus on inflammatory cytokines, including interleukin-1β (IL-1β), have proven effective in preventing atherogenesis. Circular RNAs (circRNAs) are a subclass of non-coding RNAs (ncRNAs) that can exert critical functions in the regulation of atherosclerosis. Here, using circRNA sequencing, we revealed that circRNA circDENND1B (mmu_circ_0000081) is a promising novel mediator of atherosclerosis in mouse. The expression of circDENND1B is negatively related to the progression of atherosclerosis and foam cell formation, and the upregulation of circDENND1B significantly alleviates foam cell formation induced by ox-LDL by promoting cholesterol efflux. Moreover, circDENND1B participates in the anti-atherosclerotic effect of IL-1β monoclonal antibody (IL-1β mAb), both *in vivo* and *in vitro*. With bioinformatic prediction and RNA pull-down assays, we determined that circDENND1B sponges mmu-miR-17-5p to promote *Abca1* expression in cells treated with IL-1β mAb. Our study revealed that circDENND1B, a novel regulator of cholesterol efflux, is a potential therapeutic target in atherosclerosis and provides new insights into the interaction between inflammation and cholesterol transport.

Keywords: atherosclerosis, circular RNA, inflammation, IL-1β mAb, cholesterol efflux

INTRODUCTION

Atherosclerosis, a chronic inflammatory disease characterized by fibrofatty lesions formed in the artery walls, is considered a major cause of myocardial infarction (MI) and stroke worldwide (Libby et al., 2019). Foam cell formation is a hallmark of atherosclerosis progression. Most foam cells are cholesterol ester-enriched macrophages originating from monocytes, while others are macrophage-like cells originating from vascular smooth muscle cells (Ouimet et al., 2019). The formation of foam cells is determined by the integrated effect of lipid uptake, cholesterol esterification, and cholesterol efflux (Wang D. et al., 2019).

Therapeutically targeting proinflammatory cytokines is a potential method to tackle atherosclerosis. Cytokines such as IFN-γ, TNF-α, and IL-1β are promoters of macrophage foam cell formation (McLaren et al., 2011), and activated macrophages can exacerbate inflammatory responses by releasing proinflammatory cytokines. In apolipoprotein E-deficient ($ApoE^{-/-}$) mice, neutralizing IL-1β significantly reduces the area of aortic atheroma, increase the plaque-free lumen area, and shift blood monocytes to a less inflammatory state (Bhaskar et al., 2011; Vromman et al., 2019). Several clinical studies concerning blocking IL-1β to prevent cardiovascular diseases have been conducted, although none has an indication for large-scale clinical use currently (Buckley and Abbate, 2018). Canakinumab (Ilaris), a fully humanized mAb against IL-1β, significantly lowered the rate of recurrent cardiovascular events without interfering with lipid levels (Ridker et al., 2017), indicating that IL-1β mAb is a promising antiatherosclerotic drug, although the mechanism is not thoroughly understood.

Circular RNAs (circRNAs) are a subclass of non-coding RNAs (ncRNAs) characterized by a stable structure without free 5′ or 3′ ends. circRNAs can function by modulating the activity of microRNAs (miRNAs) as competing endogenous RNAs (ceRNAs), regulating RNA-binding proteins, and could be directly translated into proteins (Han et al., 2018; Xiao et al., 2020). Previous studies have indicated that circRNAs are highly associated with the progression of atherosclerosis. The plasma level of hsa_circ_0001445 is negatively related to coronary atherosclerotic burden in patients (Vilades et al., 2020). The combination of hsa_circ_0001879 and hsa_circ_0004104 is capable of differentiating CAD patients (Wang L. et al., 2019). As an antiatherogenic circRNA, circANRIL modulates ribosomal RNA maturation and influences atherogenic pathways (Holdt et al., 2016). circCHFR and hsa_circ_0029589 were reported to participate in atherogenesis by promoting the proliferation and migration of vascular smooth muscle cells (Yang et al., 2019; Huang et al., 2020). In oxidized low-density lipoprotein (ox-LDL)-treated human aorta endothelial cells (HAECs), circ_0003204 sponges miR-370 and increases TGFβR2 expression, thus inhibiting the proliferation, migration, and tube formation of HAECs (Zhang et al., 2020). To summarize, these results imply that circRNAs play a crucial role in regulating atherosclerosis.

To date, whether circRNAs can be used in anti-atherosclerotic therapies is widely unknown. In the current study, we revealed that a novel circRNA, circDENND1B, could alleviate foam cell formation, and is involved in the antiatherosclerotic effect of IL-1β mAb in mice. circDENND1B could sponge miR-17-5p and increase ATP binding cassette subfamily A member 1 (*Abca1*) expression, thus inhibiting atherosclerosis by promoting cholesterol efflux.

MATERIALS AND METHODS

Animal Experiments

All animal experiments were approved by the ethics committee of Zhongshan Hospital, Fudan University, and the experiments took place in the animal laboratory of Zhongshan Hospital, Fudan University (Shanghai, China). Six-week-old (20.7 ± 2 g) male apolipoprotein E-deficient ($ApoE^{-/-}$) mice on a C57BL/6 background were purchased from GemPharmatech (China). The mice were randomly allocated to the experimental group or control group and were fed an HFD with 40% kcal fat, 1.25% cholesterol and 0.5% cholic acid (Research Diets, United States) or chow diet for 12 weeks. For the study involving IL-1β mAb, mice were treated with InVivoMab anti-mouse/rat IL-1β (i.p., twice weekly, 1.0 mg/kg, Bio X Cell, United States) or PBS (i.p., twice weekly, 0.1 ml) from the seventh week of atherogenic diet for 6 weeks. The mice were maintained on a 12:12 h light/dark cycle, with controlled temperature (22 ± 2°C) and humidity (55 ± 5%) conditions, and had free access to water and food. At the end of the 12-week-period, all mice were anesthetized by inhaling isoflurane and euthanized by collecting blood through cardiac puncture.

Cell Culture

The RAW264.7 cell line was purchased from Shanghai Zhong Qiao Xin Zhou Biotechnology Co., Ltd. (China) and cultured in DMEM (Gibco, United States) supplemented with 10% fetal bovine serum (HyClone, United States) and penicillin/streptomycin solution (HyClone, United States). For foam cell induction, RAW264.7 cells were incubated with 100 mg/ml ox-LDL (Yiyuan Biotech, China) for 48 h.

Total RNA/circRNA/miRNA Isolation

Total RNA was isolated and purified from cell or tissue samples using TRIzol reagent (Invitrogen, United States) following the manufacturer's protocol. RNase R (Lucigen, United States) was used to isolate circRNA from total RNA. The miRcute miRNA Isolation Kit (Tiangen, China) was utilized to isolate miRNA. The concentration and purity of each isolated RNA sample were confirmed with a NanoDrop 2000 (Thermo Fisher Scientific, United States).

RNA Library Construction, Sequencing and Data Analyses

Approximately 5 μg of total RNA was used in the RNA sequencing assay. Ribosomal RNA was depleted using the Ribo-Zero^{TM} rRNA Removal Kit (Illumina, United States), and the remaining RNA was fragmented into small pieces with divalent cations. We reverse-transcribed the cleaved RNA fragments into cDNA and synthesized U-labeled second-strand DNA with reverse-transcribed cDNA, *E. coli* DNA polymerase I, RNase H and dUTP. An A-base was then added to the blunt ends of each strand, preparing for ligation to the indexed adapters. Single- or dual-index adapters were ligated to the fragments, and AMPureXP beads were used for size selection. After heat-labile UDG enzyme treatment of the U-labeled second-strand DNA, the ligated products were amplified by PCR. The average insert size for the final cDNA library was 300 bp (±50 bp). Finally, we performed paired-end sequencing on an Illumina HiSeq 4000 (LC Bio, China) following the manufacturer's recommended protocol. To analyze the

sequencing data, Cutadapt (Martin, 2011) was used to remove the reads that contained adaptor contamination, low-quality bases or undetermined bases. The sequence quality was then verified with Fast QC. CIRCExplorer (Zhang et al., 2016) and CIRI (Gao et al., 2015) were used to assemble the mapped reads to circRNAs, and the back-splicing reads were identified in the unmapped reads by TopHat-fusion. The differentially expressed circRNAs were selected with $\log_2$ (fold change) >1 or $\log_2$ (fold change) <-1 and with statistical significance (p-value < 0.05).

BODIPY-Cholesterol Efflux Assay

To measure cholesterol efflux in RAW264.7 cells, a stock solution of BODIPY-cholesterol (Sigma, United States) was prepared at 5 mM in DMSO. Cells were loaded with 2.5 μM BODIPY-cholesterol in culture medium for 1 h at 37°C and rinsed twice with MEM (HyClone, United States) supplemented with 10 mM HEPES. Then, the cells were incubated with 10 μg/ml apoA-I (Sangon Biotech, China) or 1% (w/v) cholesterol acid (Sangon Biotech, China) for 4 h at 37°C, and the cell supernatant was filtered with a 0.45 μm filter. The BODIPY fluorescence intensity in the supernatants was measured using a fluorescence spectrophotometer. The images of BODIPY-cholesterol were taken under the same exposure time with a fluorescence microscope (Leica, Germany).

Reverse Transcription and Quantitative Real-Time PCR (qRT-PCR)

The PrimeScript™ RT reagent Kit with gDNA Eraser (TAKARA, China) was used to synthesize cDNA from isolated total RNA or circRNA, and the cDNA was then processed for qRT-PCR using SYBR® Premix Ex Taq™ (TAKARA, China) following the manufacturer's protocols. Reverse transcription and quantitative detection of miRNA were performed with the miRcute Plus miRNA First-Strand cDNA Kit (Tiangen, China) and miRcute Plus miRNA qPCR Kit (SYBR Green) (Tiangen, China) following the manufacturer's instructions. Primer sequences are available in **Supplementary Table 1**.

Generation of Overexpressing Cell Lines

Lentiviruses overexpressing circDENND1B or containing the control vector were purchased from Hanbio Biotechnology (China). RAW264.7 cells were cultured in 6-well plates, and culture medium containing lentivirus was added when the cells reached 70% confluence. The culture medium was replaced 24 h later, and stable cell lines were selected using puromycin (2 mg/ml, Beyotime, China).

siRNA and miRNA Mimic Transfection

siRNA targeting circDENND1B, non-specific negative control siRNA, miRNA-17-5p mimics and negative control mimics were purchased from GenePharma (China). Transfection of siRNAs or miRNA mimics into RAW264.7 cells were conducted with Lipofectamine RNAiMAX Transfection Reagent (Thermo Fisher Scientific, United States) according to the manufacturer's protocols. Six hours after transfection, the medium was replaced. The sequences of the siRNAs and miRNA mimics are provided in **Supplementary Table 2**.

Oil Red O Staining

For the Oil Red O (ORO) staining of cells, 24 h after siRNA or miRNA-mimic treatment, RAW264.7 cells were incubated with ox-LDL for 48 h. The cells were washed with PBS and fixed with 4% paraformaldehyde (PFA) for 20 min at room temperature. Then, the cells were incubated with filtered ORO staining solution (Sangon Biotech, China) for 30 min at room temperature and processed for hematoxylin staining for 5 min. For the staining of aorta roots, mouse aorta root tissues were immediately snap-frozen in liquid nitrogen and placed in OCT cryostat embedding compound (Tissue-Tek, United States) after sacrifice. Frozen aorta root sections were then stained with ORO staining solution. The intracellular lipid droplets were observed and assessed under a bright-field microscope (Leica, Germany).

Western Blot (WB)

Forty-eight hours after ox-LDL and IL-1β mAb treatment, cells were lysed with ice-cold RIPA lysis buffer (Beyotime, China) following the manufacturer's instructions. The concentration of protein was measured using a BCA Protein Assay Kit (Beyotime, China). Extracted proteins were separated by SDS-PAGE, transferred to polyvinylidene fluoride (PVDF) membranes (Millipore), blocked in 5% BSA in 0.05% Tween 20/TBS for 2 h, and incubated overnight with the following primary antibodies at the indicated dilutions: anti-ABCA1 (1:800; Abcam, #66217, United States), anti-LPL (1:1000; Santa Cruz, #sc0373759, United States) and anti-GAPDH (1:2000; Cell Signaling Technology, #5147, United States). PVDF membranes were exposed to horseradish peroxidase (HRP)-conjugated secondary antibodies (Cell Signaling Technology, United States), and signals were detected with the Luminata™ Forte Western HRP Substrate (Millipore, United States).

ELISA

Mouse IL-1β, IL-6, and TNF-α in lung and liver homogenates were measured by ELISA. All ELISA tests were performed using DuoSet kits (R&D Systems, United States) following the manufacturer's instructions, and the results were read on an Epoch™ 2 Microplate Spectrophotometer (BioTek, United States). The quantitation of total protein was measured using a BCA Protein Assay Kit (Beyotime, China).

Fluorescence *in situ* Hybridization (FISH) Assay

The *in vitro* localization of circDENND1B was measured using a FISH kit (RiboBio, China) following the manufacturer's instructions, and the FISH probe was synthesized by RiboBio. The localization of circDENND1B was detected under a fluorescence microscope (Leica, Germany).

Immunohistochemistry (IHC)

IL-1β, IL-6 and TNF-α in aorta roots were measured by IHC. The aorta root sections were immersed in 4% PFA and embedded

into paraffin blocks. After dewaxing, the sections were rinsed with PBS, followed by antigen retrieval. After treatment with 3% H2O2 for 25 min, 3% BSA was used to prevent non-specific antibody binding for 30 min, followed by incubation with the following primary antibodies at 4°C overnight: rabbit anti-mouse IL-1β pAb (Abcam, #205924, United States); rabbit anti-mouse IL-6 pAb (Abcam, #208113, United States); and mouse anti-mouse TNF-α mAb (Abcam, #1793, United States). The sections were washed with PBS and then incubated with secondary antibodies for 50 min. Next, DAB buffer was added for development, followed by hematoxylin counterstaining, HCl differentiation, water rinsing and drying in gradient ethanol. The samples were immersed in xylene and mounted with neutral buffered resin. Images were analyzed with ImageJ software.

RNA Pull-Down Assay

For the RNA pull-down assay, the specific biotinylated probe for circDENND1B was designed and synthesized by Cloudseq (China). Cell lysate was incubated with 3 μg biotinylated probes at room temperature for 1 h. Then, the biotin-coupled RNA complex was pulled down by incubating with streptavidin magnetic beads (Thermo Fisher, United States) for 2 h. Afterward, the magnetic beads were washed with lysis buffer five times, and the samples were resuspended in 50 μl elution buffer, followed by RNA extraction and qRT-PCR assay of the target miRNA.

Statistical Analysis

Statistical analyses were performed with Prism7 (GraphPad, United States). Unpaired two-tailed *t*-test was used to compare data between two groups. For experiments involving more than two groups, data were analyzed with one-way ANOVA followed by Dunnett's *post hoc* test. Data are presented as the mean ± SEM unless otherwise mentioned. *P*-values < 0.05 were considered statistically significant.

RESULTS

A High-Fat Diet Induced an Inflammatory State in *ApoE*$^{-/-}$ Mice

Six-week-old male *ApoE*$^{-/-}$ mice were randomly allocated to the chow diet group or high-fat diet (HFD) group. After an HFD for 12 weeks, atherosclerotic lesions could be detected in the aorta, and the lesion area was significantly larger in mice fed an HFD than in mice fed a normal diet (**Supplementary Figures 1A–C**).

To confirm that HFD-fed mice were in a systemic inflammatory state (Bäck et al., 2019), ELISA was conducted to measure the levels of IL-1β, IL-6, and TNF-α in the lungs and livers, which showed that the HFD-fed mice were in a systemic inflammatory state (**Supplementary Figures 2A–F**). The immunohistochemistry (IHC) positive staining areas of IL-1β, IL-6, and TNF-α in the aortic roots were also larger in atherosclerotic mice (**Supplementary Figure 2G**). In addition, the expression of IL-1β and IL-6 mRNA clearly increased in atherosclerotic aortas, while the expression of TNF-α in aortas increased but did not reach statistical significance (**Supplementary Figure 2H**). IL-1β and TNF-α were significantly highly expressed in the lungs and livers of atherosclerotic mice (**Supplementary Figures 2I,J**). The above results indicate that HFD-fed mice were in a highly inflammatory state, which stresses the necessity of anti-inflammation therapy toward atherosclerosis.

Differentially Expressed mRNA/circRNA in Atherosclerotic *ApoE*$^{-/-}$ Mice Were Identified With mRNA/circRNA Sequencing

To investigate the roles of circRNAs in the initialization and development of atherosclerotic plaques, we extracted RNA from the aortas of chow diet-fed or HFD-fed mice and conducted second-generation sequencing to identify the differentially expressed mRNAs and circRNAs.

From the mRNA sequencing results, 1732 upregulated genes and 1047 downregulated genes were identified ($p < 0.05$) (**Figure 1A**). Gene ontology (GO) analyses of the differentially expressed genes (DEGs) revealed that innate immune response, inflammatory response, and immune system process were significantly affected (**Figure 1B**). Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses showed that the DEGs were significantly enriched in pathways related to inflammation, including the chemokine signaling pathway, cytokine-cytokine receptor interaction, and B cell receptor signaling pathway (**Figure 1C**). These results confirmed that modulating inflammation is a promising strategy for the treatment of atherosclerosis.

Furthermore, 56 differentially expressed circRNAs (DECs) were identified out of 10024 circRNAs. The numbers of increased and decreased circRNAs in HFD-fed mice were 15 and 41, respectively (**Figure 1D**). Among the DECs with a *P* value lower than 0.01, we confirmed that the expression of circDENND1B, circRNA690, circRNA994, mmu-circ-0015075 and circRNA1984 decreased in the aortas of HFD-fed mice (**Figure 1E**). The selected DECs target over 100 miRNAs, as predicted by RegRNA (Chang et al., 2013) (**Figure 1F**).

circDENND1B Participates in the Antiatherosclerotic Effect of IL-1β mAb *in vitro*, and Its Expression Was Affected by IL-1β mAb *in vivo*

As atherosclerotic mice were in an inflammatory state, we wondered whether blocking certain cytokines could affect atherosclerosis via modulation of circRNAs. Thus, we intraperitoneally injected IL-1β mAb, which was proven effective in preventing MI in the CANTOS study (Ridker et al., 2017), into HFD-fed *ApoE*$^{-/-}$ male mice biweekly (**Figure 2A**). IL-1β mAb treatment relieved atherosclerotic plaques and attenuated local inflammation in aortic roots (**Figures 2B,C** and **Supplementary Figures 3A,B**). Besides, the level of collagen around atherosclerotic plaques was detected using Masson' trichrome staining, indicating that the plaque in IL-1β mAb-treated mice is more stable (**Figure 2D**).

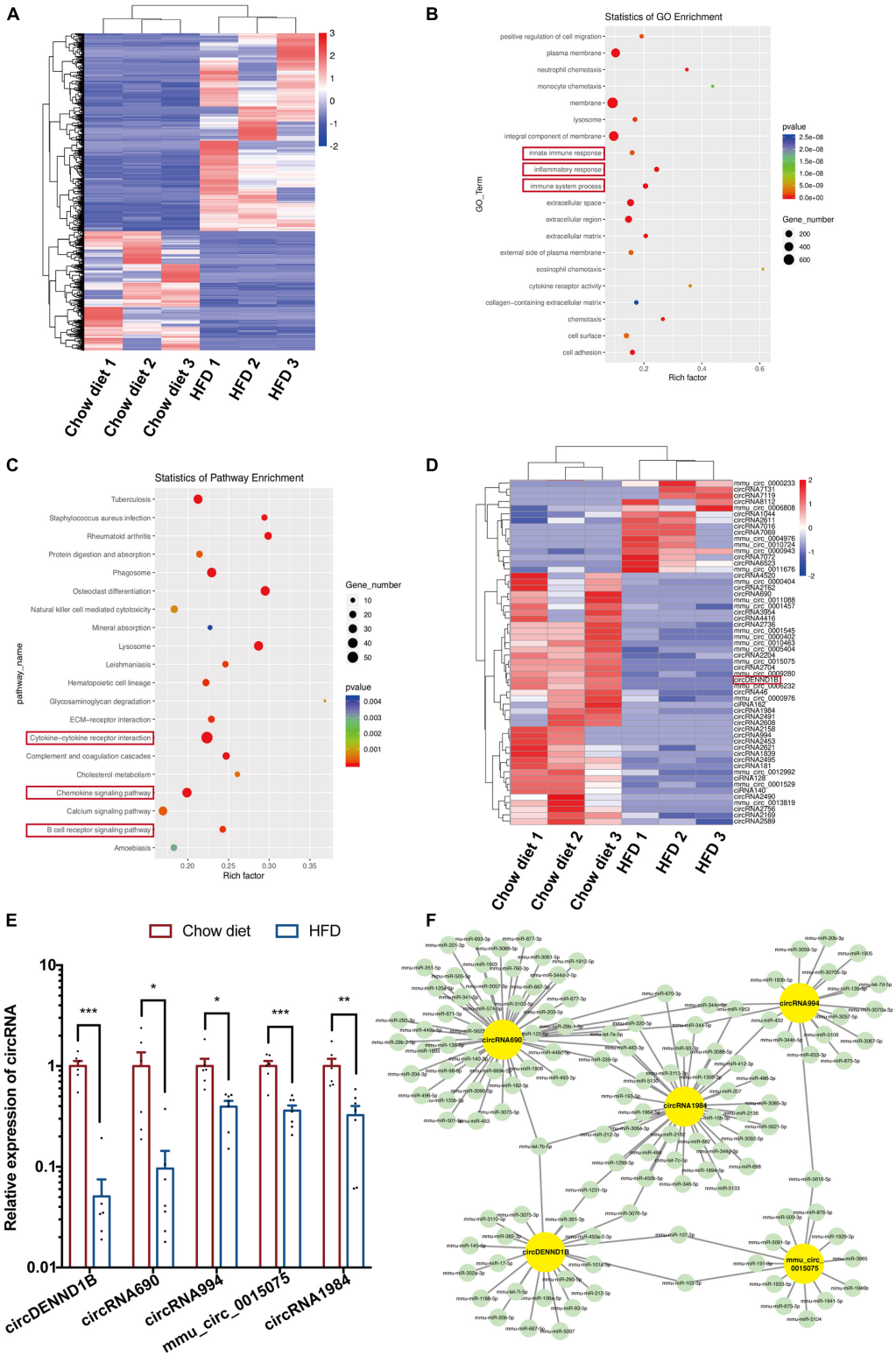

FIGURE 1 | Differentially expressed circular RNA/mRNA in mouse aortas. **(A)** Heatmap of DEGs in the aortas of HFD-fed compared to chow diet-fed mice. **(B)** The top 20 GO terms of detected DEGs. The enriched GO terms are ordered alphabetically. **(C)** Scatterplots of the top 20 differentially regulated pathways identified in KEGG. The KEGG terms are ordered alphabetically. **(D)** Heatmap of the 56 identified DECs in mouse aortas. **(E)** qRT-PCR results confirming that five DECs are downregulated in atherosclerotic aortas, Data were normalized to GAPDH (n = 6 per group). **(F)** The ceRNA module network of five identified DECs. The yellow nodes represent circRNAs, and the green nodes represent predicted miRNAs. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

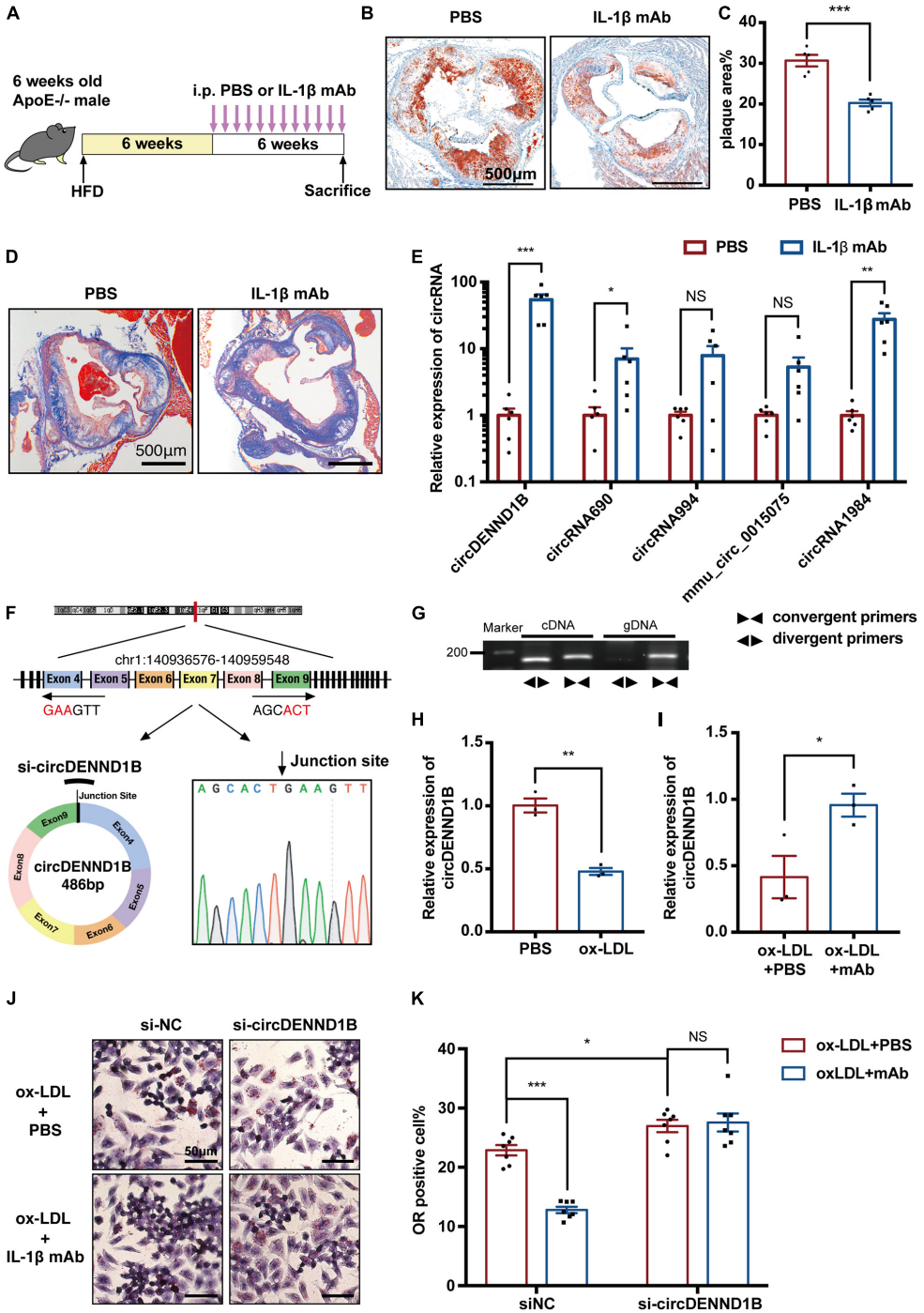

FIGURE 2 | circDENND1B participates in the antiatherosclerotic effect of IL-1β mAb *in vitro*, and its expression was affected by IL-1β mAb *in vivo*. **(A)** Schematic depiction of the study in which ApoE$^{-/-}$ mice were fed a HFD for 12 weeks and treated with PBS or IL-1β mAb. **(B)** ORO staining showing that IL-1β mAb injection resulted in a smaller plaque area in the aortic roots. **(C)** Quantification of the percentage of plaque area compared to the total area of aortic roots (n = 5 per group). **(D)** Representative picture of Masson's trichrome staining of aorta roots. **(E)** The expression of five circRNAs in aortas was detected with qRT-PCR. Data were normalized to GAPDH (n = 6 per group). **(F,G)** Schematic depiction of circDENND1B's structure, showing it is comprised of the fourth to ninth exon of *Dennd1b*. Agarose gel electrophoresis showing that circDENND1B is a circular RNA. The circRNA was amplified by divergent primers (◀▶) from cDNA but not from gDNA. *Dennd1b*, the host gene of circDENND1B, was amplified by convergent primers (▶◀) from both cDNA and gDNA. Sanger sequencing validated its splicing junction. **(H)** qRT-PCR shows that circDENND1B expression decreased in RAW264.7 cells treated with ox-LDL. **(I)** qRT-PCR shows that circDENND1B expression increased after IL-1β mAb treatment. **(J)** Representative image of ORO staining, depicting that fewer foam cells are formed in RAW264.7 cells treated with ox-LDL and IL-1β mAb. **(K)** The percentage of foam cells was lower after IL-1β mAb treatment. Approximately 1000 cells were counted per treatment over seven separate experiments. gDNA, genomic DNA; NS, no significant difference, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

To determine the participation of circRNAs in the antiatherosclerotic process, we conducted qRT-PCR assay on circRNAs. Among the DECs identified between the aortas of HFD-fed and chow diet-fed mice, the expression of circDENND1B increased markedly after IL-1β mAb treatment (**Figure 2E**), suggesting the possibility of association between circDENND1B and the antiatherosclerotic effect of IL-1β mAb *in vivo*.

Furthermore, we established an *in vitro* foam cell model by incubating RAW264.7 cells with 100 mg/ml ox-LDL for 48 h. Foam cell formation was demonstrated by ORO staining (**Supplementary Figure 4A**). We confirmed the structure of circDENND1B using PCR, agarose gel electrophoresis, and Sanger's sequencing (**Figures 2F,G**). The expression of circDENND1B was downregulated in ox-LDL-treated cells (**Figure 2H**). In cells incubated with IL-1β mAb, while the expression of circDENND1B was upregulated (**Figure 2I**), an alleviating trend in lipid accumulation was observed, indicating that IL-1β mAb could interfere with lipid accumulation or efflux (**Figures 2J,K**). Moreover, although IL-1β mAb treatment significantly alleviated foam cell formation, this effect was inhibited after knocking down circDENND1B, indicating that circDENND1B is necessary in the anti-foam-cell-formation process of IL-1β mAb. These results suggest that circDENND1B protects atherogenesis, and participates in the antiatherosclerotic effect of IL-1β mAb *in vitro*.

circDENND1B Inhibits Foam Cell Formation by Promoting *Abca1* Expression

To explore the possible roles of circDENND1B in the formation of macrophage-derived foam cells, we used siRNA to knock down circDENND1B, and established a circDENND1B-overexpressing RAW264.7 cell line using lentivirus (**Supplementary Figures 5A,B**). Cells overexpressing circDENND1B exhibit a significantly lower percentage of ORO-positive cells, while circDENND1B-knock down cells show an opposite trend (**Figures 3A–D**), indicating that circDENND1B regulates foam cell formation.

Foam cell formation is mainly determined by the combined effect of lipid uptake and cholesterol efflux (Wang D. et al., 2019). As illustrated in **Figures 3E,F**, the uptake of ox-LDL remains unchanged in cells overexpressing circDENND1B or transfected with si-circDENND1B, as compared with control group. However, the cholesterol efflux rate significantly increased in OE group (**Figures 3G,H**), suggesting that circDENND1B possibly regulates foam cell formation by promoting cholesterol efflux. Therefore, we conducted qRT-PCR on genes related to foam cell formation, including *Abca1*, *Abcg1*, *Acat1*, *Sr-a1*, *Sr-b1*, *Msr-1*, *and Lpl* (McLaren et al., 2011). In cells treated with ox-LDL, *Abca1* expression increased significantly after overexpressing circDENND1B, which is the most significant among these genes (**Figure 3I**). Moreover, the increase of ABCA1 protein expression after circDENND1B overexpression was further identified using WB (**Figure 3J**). On the contrary, the expression of *Abca1*, as well as *Msr-1* and *Lpl*, decreased in cells transfected with si-circ-DENND1B (**Figure 3K**). These results suggested that circDENND1B modulates foam cell formation via promoting cholesterol efflux, and *Abca1* is a potential target of circDENND1B.

circDENND1B Participates in the Antiatherosclerotic Effect of IL-1β mAb by Promoting Cholesterol Efflux

As circDENND1B is correlated with the antiatherosclerotic effect of IL-1β mAb, we then conducted experiments to identify whether this effect is also related to cholesterol efflux. From qRT-PCR results on genes related to foam cell formation, we found that *Abca1* was most significantly differentially regulated in cells treated with IL-1β mAb and ox-LDL (**Figure 4A**), suggesting that IL-1β mAb is capable of mediating *Abca1*.

Moreover, in si-circDENND1B-treated cells, the increase in cholesterol efflux induced by IL-1β mAb was less significant than in the si-NC group (**Figures 4B,C**) (Sankaranarayanan et al., 2011). qRT-PCR and Western blotting also showed that the increase of *Abca1* mRNA expression and ABCA1 protein expression induced by IL-1β mAb is compromised after knocking down circDENND1B (**Figures 4D,E**). These data suggested that circDENND1B modulates the antiatherosclerotic effect of IL-1β mAb by promoting cholesterol efflux.

circDENND1B/miR-17-5p/*Abca1* Regulates the Inhibition of Foam Cell Formation Induced by IL-1β mAb

MicroRNA sponging is one of the most predominant functions of circRNAs. As circDENND1B is predominantly localized in the cytoplasm of RAW264.7 cells (**Figure 5A**), which is similar to the location of most miRNAs (Li et al., 2013), we hypothesized that circDENND1B promotes *Abca1* expression through a miRNA sponging mechanism.

To confirm this, we used RegRNA to predict the target miRNAs of circDENND1B. Among the 23 miRNAs predicted to bind with circDENND1B, three miRNAs, let-7b-5p (Wei et al., 2018), miRNA-17-5p (miR-17-5p) (Huang et al., 2018; An et al., 2019; Tan et al., 2019) and miRNA-212-5p (miR-212-5p) (Miao et al., 2018), have been reported to affect foam cell formation. The expression of miR-17-5p and miR-212-5p increased in cells treated with ox-LDL, and miR-17-5p decreased after IL-1β mAb treatment (**Figures 5B,C**). As circDENND1B harbors eight binding sites for miR-17-5p, and miR-17-5p has been reported to modulate the progress of atherosclerosis via *Abca1* (He et al., 2015; Tan et al., 2019; Liu L. et al., 2020) (**Figure 5D**), we investigated the relationship between circDENND1B and miR-17-5p. Upregulating circDENND1B resulted in a decrease in miR-17-5p expression, while knocking-down had a converse effect (**Figures 5E,F**). A biotinylated RNA pull-down assay further confirmed that circDENND1B is a miRNA-17-5p sponge (**Figure 5G**).

In RAW264.7 cells overexpressing miR-17-5p, qRT-PCR results showed a decrease in *Abca1* expression (**Supplementary Figure 5C** and **Figures 5H,I**). The increase in *Abca1* induced by

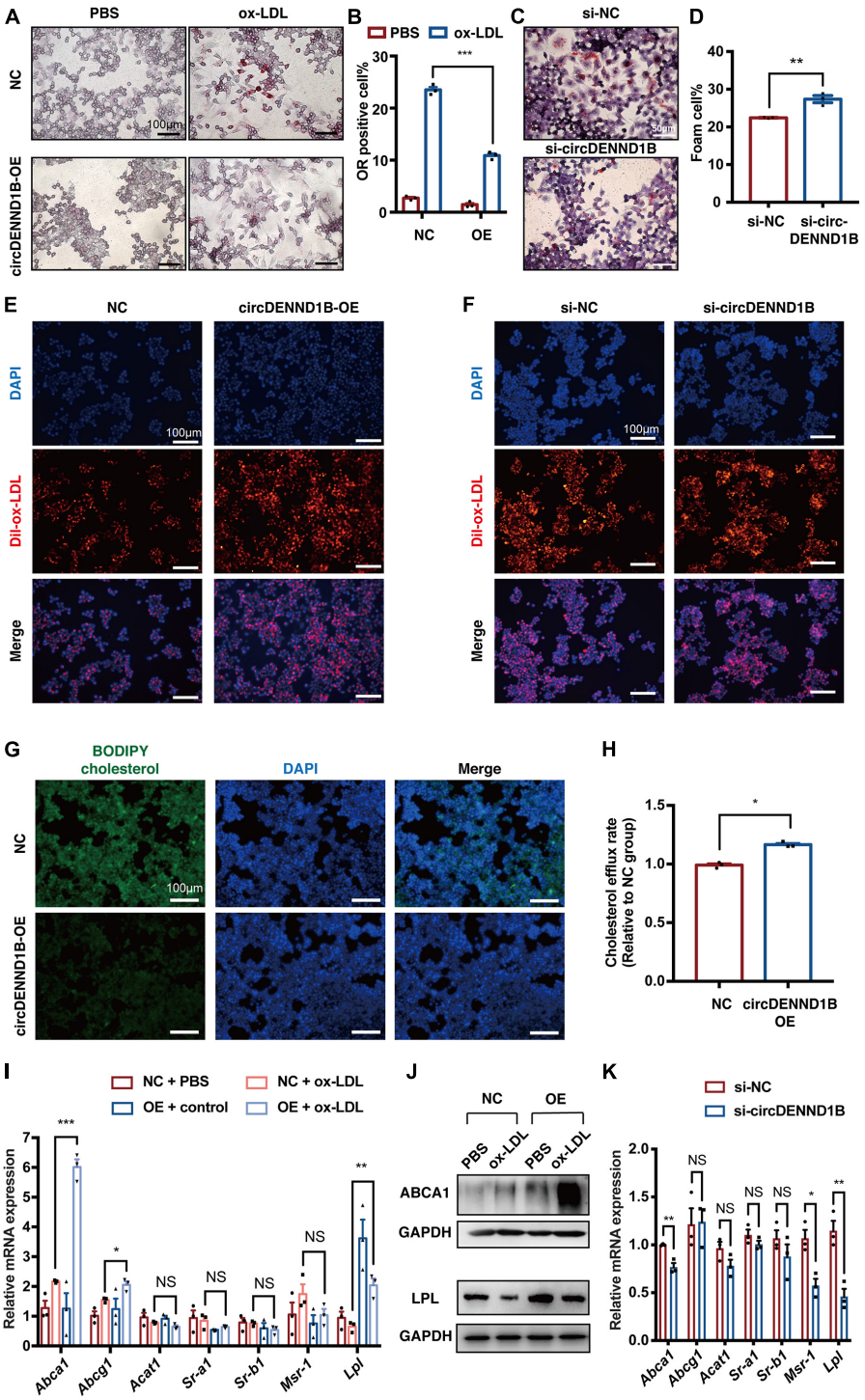

FIGURE 3 | circDENND1B regulates foam cell formation by promoting cholesterol efflux. **(A)** ORO staining showing that fewer foam cells are formed in RAW264.7 cells overexpressing circDENND1B. **(B)** The percentage of foam cells was lower after overexpressing circDENND1B. Approximately 1000 cells were counted per treatment over three separate experiments. **(C)** ORO staining showing that fewer foam cells are formed after knocking down circDENND1B. **(D)** The percentage of foam cells was higher after knocking down circDENND1B. Approximately 1000 cells were counted per treatment over three separate experiments. **(E,F)** Representative fluorescent images of cells treated with dil-ox-LDL for 48 h. Dil-ox-LDL positive cells were stained in red. **(G)** Representative images of the BODIPY assay showing overexpressing circDENND1B lowered cellular cholesterol. **(H)** Relative cholesterol efflux rate in cells overexpressing circDENND1B, measured by BODIPY assay. **(I)** qRT-PCR results of genes related to foam cell formation. Data were normalized to GAPDH (n = 3 per group). **(J)** WB results of proteins related to foam cell formation after ox-LDL treatment. **(K)** qRT-PCR results of genes related to foam cell formation in RAW264.7 cells. Data were normalized to GAPDH (n = 3 per group). NC: negative control. OE, overexpress; NS, no significant difference, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

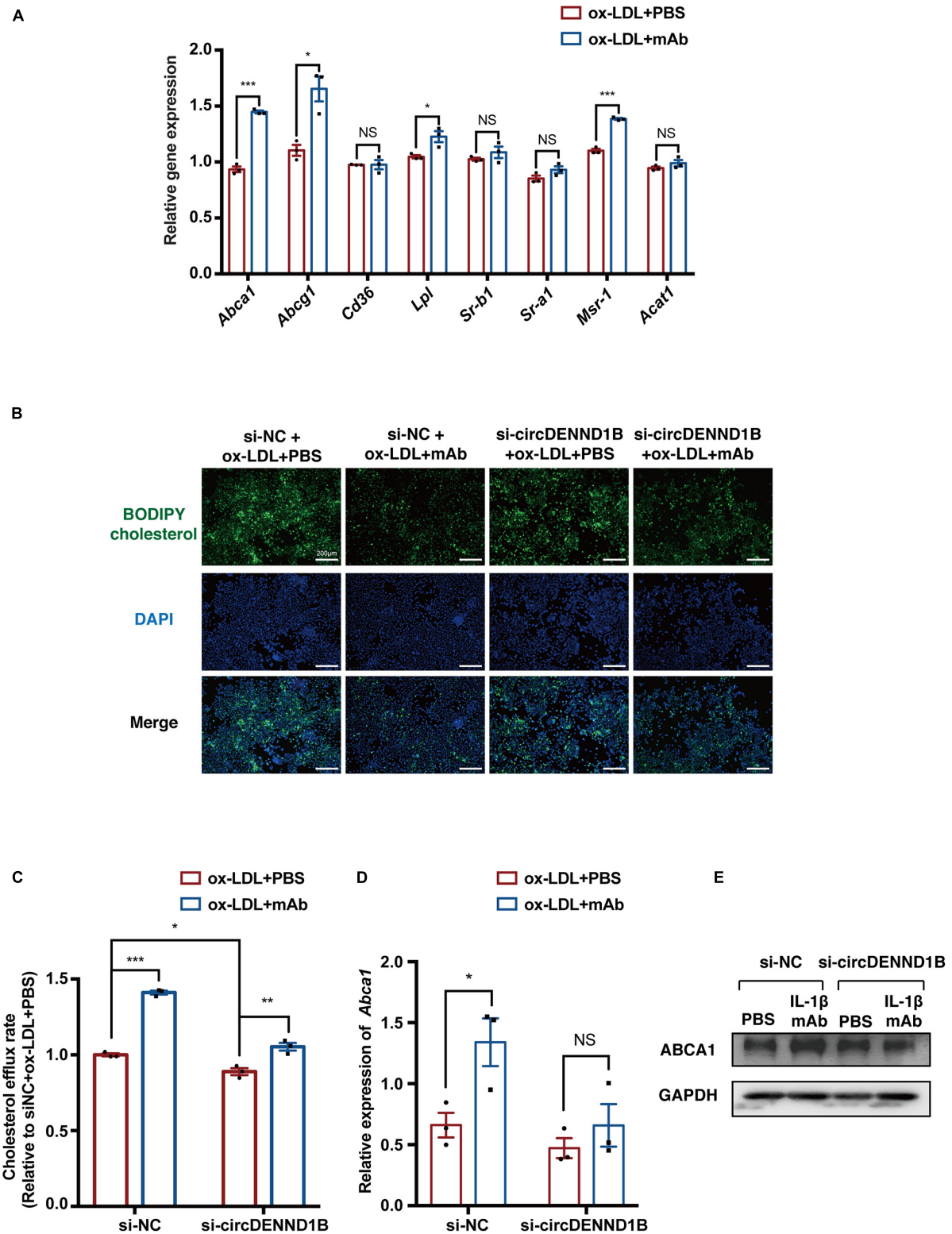

FIGURE 4 | circDENND1B is crucial in the pro-cholesterol efflux effect of IL-1β mAb *in vitro.* **(A)** qRT-PCR results of genes related to foam cell formation in IL-1β mAb-treated foam cells. **(B)** Representative images of BODIPY assay showing circDENND1B mediates the pro-cholesterol efflux effect of IL-1β mAb. **(C)** Relative cholesterol efflux rate in RAW264.7 cells, measured by BODIPY assay (n = 3 per group). **(D)** qRT-PCR results of *Abca1* expression in si-circDENND1B-treated cells, data were normalized to GAPDH data (n = 3 per group) **(E)** WB results of ABCA1 in cells transfected with si-circDENND1B and treated with ox-LDL plus PBS or IL-1β mAb.

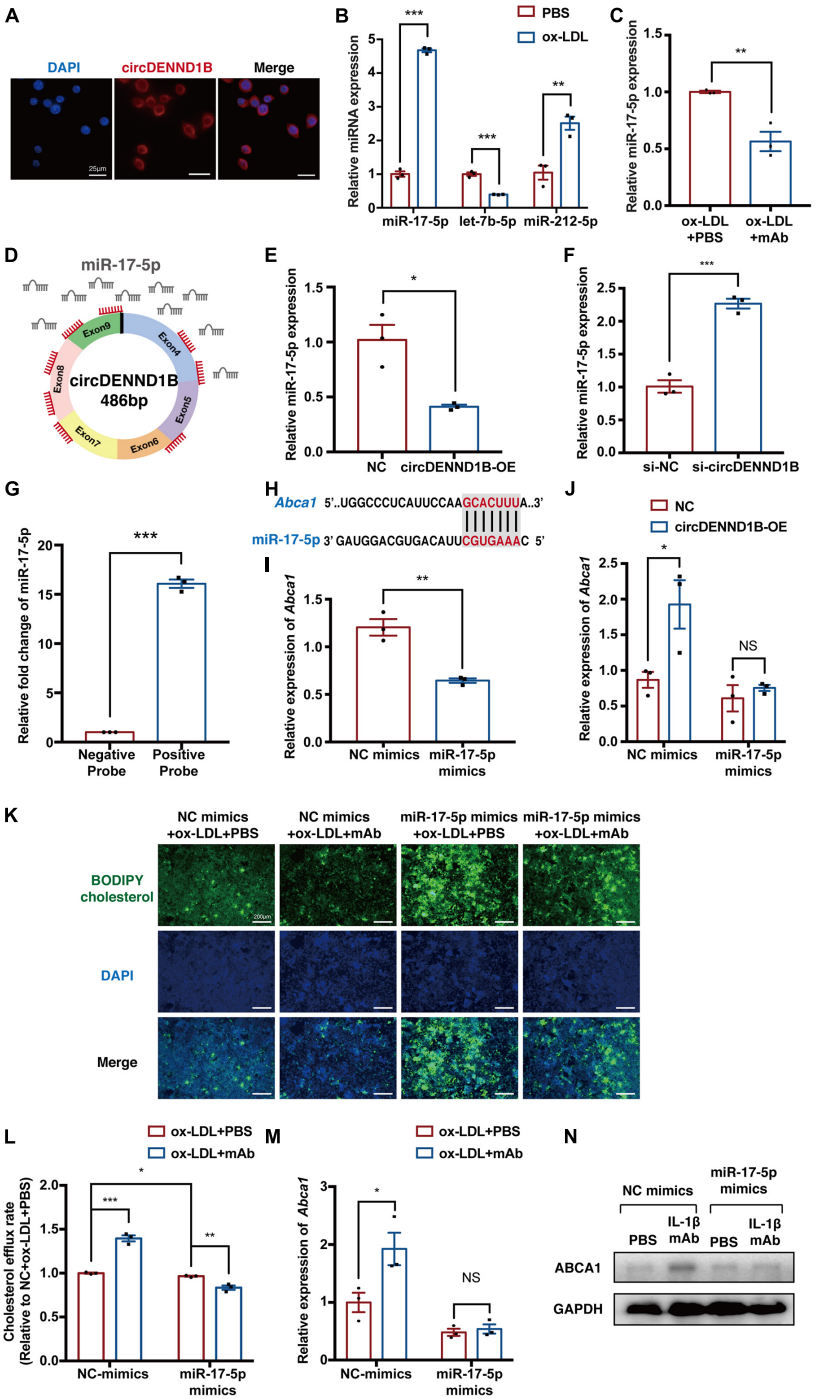

FIGURE 5 | circDENND1B/miR-17-5p/*Abca1* inhibits foam cell formation induced by IL-1β mAb. **(A)** Fluorescence *in situ* hybridization (FISH) assay showing that circDENND1B is localized in the cytoplasm. **(B)** qRT-PCR results of let-7b-5p, miR-17-5p and miR-212-5p in ox-LDL-treated cells. **(C)** qRT-PCR results of miR-17-5p in cells treated with ox-LDL and PBS or IL-1β mAb. **(D)** Schematic of the miR-17-5p sponge mechanism of circDENND1B, showing eight binding sites (red) of miR-17-5p (gray) in circDENND1B. **(E)** qRT-PCR shows miR-17-5p expression decreased after overexpressing circDENND1B. **(F)** qRT-PCR shows miR-17-5p expression increased after knocking down circDENND1B. **(G)** RNA pull-down assay confirmed the binding between circDENND1B and miR-17-5p. **(H)** miR-17-5p was predicted to bind the 3′UTR of *Abca1*. **(I)** qRT-PCR results of *Abca1* in foam cells treated with NC mimics or miR-17-5p mimics. **(J)** qRT-PCR results of *Abca1* in cells transfected with miRNA mimics and overexpressing circDENND1B. **(K)** Representative images of the BODIPY assay showing the impact of miR-17-5p mimics and IL-1β mAb on cholesterol efflux. **(L)** Cholesterol efflux rate in miR-17-5p-mimics-treated cells, measured by BODIPY assay. **(M)** qRT-PCR results of *Abca1* in cells transfected with miRNA mimics ($n = 3$ per group). **(N)** WB results of ABCA1 in cells transfected with miR-17-5p and treated with ox-LDL plus PBS or IL-1β mAb. The qRT-PCR data of mRNA were normalized to GAPDH data ($n = 3$ per group), and miRNA expression levels were normalized to U6 levels ($n = 3$ per group). NC, negative control; OE, overexpression; NC mimics, negative control miRNA mimics. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

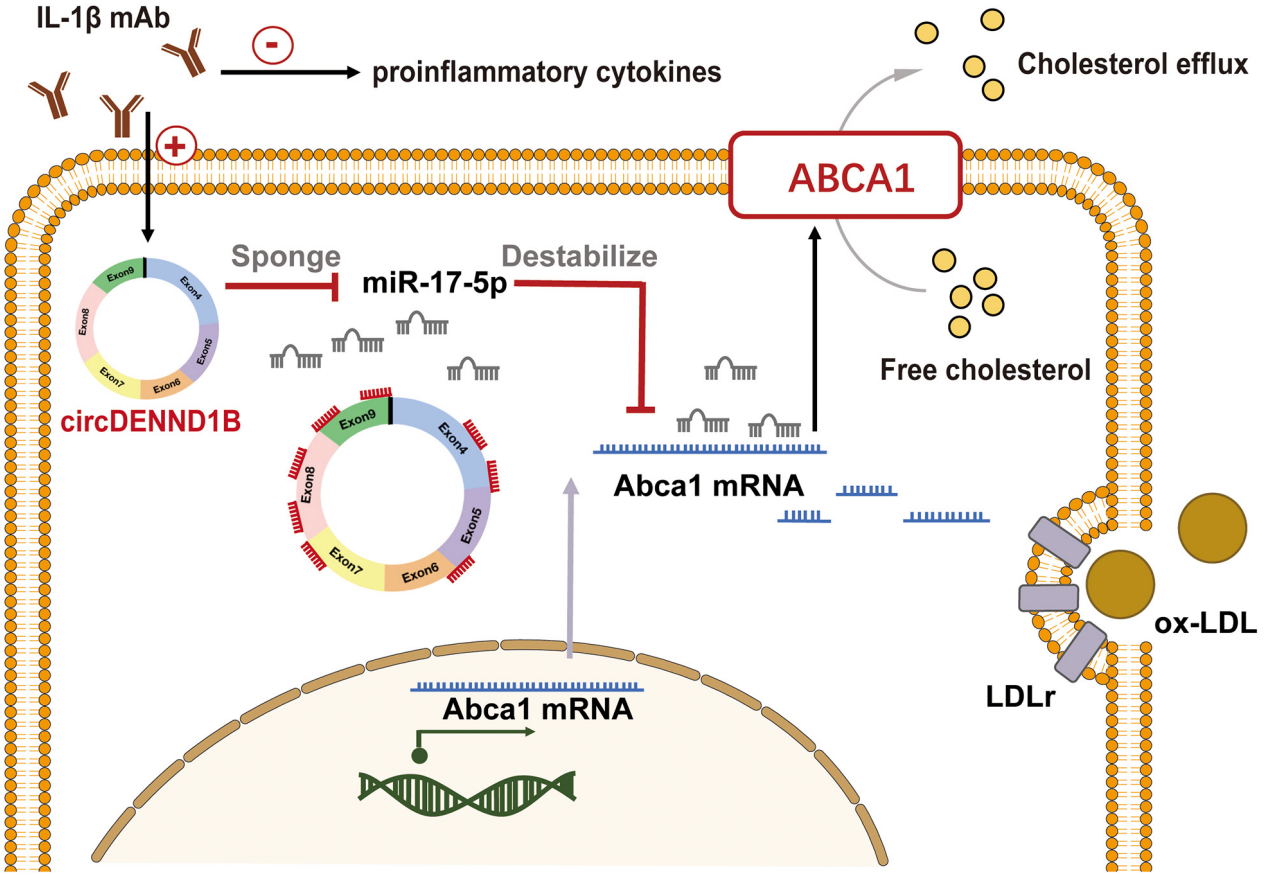

FIGURE 6 | Proposed model of circDENND1B/miR-17-5p/*Abca1* inhibits foam cell formation induced by IL-1β mAb. Treating macrophages with IL-1β mAb may increase the expression of circDENND1B, thus inhibit the miR-17-5p-*Abca1* pathway. The increased ABCA1 protein would promote cholesterol efflux, and regulate the formation of foam cells.

overexpressing circDENND1B was compromised in the presence of miR-17-5p mimics (**Figure 5J**). As Tan et al. (2019) have confirmed the binding property between miR-17-5p and *Abca1*, these findings suggested that circDENND1B sponges miR-17-5p to increase *Abca1* in foam cells.

Next, we identified the role of miR-17-5p in the antiatherosclerotic effect. Overexpressing miR-17-5p with miRNA mimics in foam cells attenuated the IL-1β mAb-induced changes in cholesterol efflux, foam cell formation and *Abca1* expression, suggesting that overexpressing miR-17-5p has a similar influence on foam cell formation as silencing circDENND1B (**Figures 5K–N** and **Supplementary Figures 6A,B**).

In conclusion, these results indicate that both circDENND1B and miR-17-5p modulate the impact of IL-1β mAb on *Abca1*, thus promoting cholesterol efflux (**Figure 6**).

DISCUSSION

Modulating inflammation is crucial in treating atherosclerosis, and blocking proinflammatory cytokines is a promising strategy to prevent atherogenesis. Inflammatory responses influence lipid metabolism in atherosclerosis and vice versa. IL-1β, a soluble cytokine that plays a pivotal role in innate immune responses, was reported to mediate lipid transport by inhibiting ABCA1 (Tumurkhuu et al., 2018; Liu H. et al., 2020), but the mechanism has not been thoroughly elucidated. In this work, we conducted circRNA-sequencing in atherosclerotic mouse aortas for the first time and identified a novel circRNA, circDENND1B, through which IL-1β mAb inhibits atherosclerosis and promotes cholesterol efflux.

circRNAs are a novel subclass of ncRNAs that play important roles in multiple physiological processes, but the roles of circRNAs in atherosclerosis are not well understood. In the present study, we identified 56 DECs in the aortas of HFD-fed and chow-diet-fed mice. Among them, an exonic circRNA, circDENND1B, was shown to modulate *Abca1* expression by sponging miR-17-5p, thus inhibiting foam cell formation. A significant negative correlation between circDENND1B and miR-17-5p was observed by qRT-PCR, suggesting that detectable cellular miR-17-5p decreased after being sponged. Moreover, we confirmed that miR-17-5p is physically bound to circDENND1B using RNA pull-down assay. These results are consistent with previous studies which showed that

miR-17-5p accelerates atherogenesis by modulating lipid accumulation (Tan et al., 2019) and endothelial cell ferroptosis (Xiao et al., 2019).

ABCA1 is an essential receptor for the initial step of cholesterol efflux from foam cells in atherosclerotic plaques. Approximately 50% of the net cholesterol efflux from cholesterol-laden mouse macrophages is attributed to ABCA1, while another member of the ATP binding cassette superfamily, ATP binding cassette subfamily G member 1 (ABCG1), accounts for 20% of cholesterol efflux (Adorni et al., 2007). Moreover, various proinflammatory cytokines, including IL-1β, TNF-α, and IFN-γ, can regulate *Abca1* expression and promote foam cell formation (McLaren et al., 2011). Consistent with prior reports showing that IL-1β could suppress *Abca1* expression in macrophages (Tumurkhuu et al., 2018), we confirmed that blocking IL-1β elevated the expression of *Abca1* and promoted cholesterol efflux in foam cells and showed that the pro-cholesterol-efflux function of IL-1β mAb is partially mediated through circDENND1B.

Nevertheless, other possible mechanisms of the antiatherogenic effect of IL-1β mAb should be considered. According to former reports, IL-1β competes for access to ABCA1 with cholesterol (Tumurkhuu et al., 2018) and downregulates ABCA1 through the ROS-NF-κB pathway (Chen et al., 2007). Blocking IL-1β with mAbs disrupts the IL-1β positive feedback loop, thus promoting cholesterol efflux and preventing the production of IL-1β itself. Moreover, IL-1β contributes greatly to smooth muscle cell (SMC) proliferation, endothelial cell activation, and the modulation of various immune cells (Tall and Yvan-Charvet, 2015; Williams et al., 2019). Whether circDENND1B plays a role in these processes still needs further investigation.

A limitation of this study is that the expression and function of circDENND1B were not identified in other species except for mouse, especially in humans. According to previous studies, a proportion of circRNAs are conserved among mammals in organs including the heart, brain, and liver (Rybak-Wolf et al., 2015; Werfel et al., 2016; Xia et al., 2017). In a recent study concerning atherosclerosis (Zeng et al., 2021), circMAP3K5 was found to prevent hyperplasia of vascular smooth muscle cells conservatively in both human and mouse. Although The mouse circRNA circDENND1B was not detected in human macrophages, it was found to be conserved with the human circRNA hsa_circ_0111650 in the brain (Rybak-Wolf et al., 2015), but the effect of the latter circRNA on human foam cell formation is still unknown. Therefore, to prove the therapeutic potential of circDENND1B, further research is necessary to identify whether it could exert its antiatherosclerotic effect in human.

In conclusion, our results suggest that circDENND1B is a ceRNA through which IL-1β mAb inhibits atherosclerosis. This newly identified circRNA could be a promising therapeutic target of atherosclerosis.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI GEO (accession: GSE159379).

ETHICS STATEMENT

The animal study was reviewed and approved by The Ethics Committee of Zhongshan Hospital.

AUTHOR CONTRIBUTIONS

FX and LS came up with the research idea and designed the experiments. FX, HC, RW, and TZ conducted the experiments and analyzed the data. LS, JQ, and JG supervised the project. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (81521001 and 81670319).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.652032/full#supplementary-material

REFERENCES

Adorni, M. P., Zimetti, F., Billheimer, J. T., Wang, N., Rader, D. J., Phillips, M. C., et al. (2007). The roles of different pathways in the release of cholesterol from macrophages. *J. Lipid Res.* 48, 2453–2462. doi: 10.1194/jlr.M700274-JLR200

An, J., Chen, Z., Ma, Q., Wang, H., Zhang, J., and Shi, F. (2019). LncRNA SNHG16 promoted proliferation and inflammatory response of macrophages through miR-17-5p/NF-κB signaling pathway in patients with atherosclerosis. *Eur. Rev. Med. Pharmacol. Sci.* 23, 8665–8677. doi: 10.26355/eurrev_201910_19184

Bäck, M., Yurdagul, A., Tabas, I., Öörni, K., and Kovanen, P. (2019). Inflammation and its resolution in atherosclerosis: mediators and therapeutic opportunities. *Nat. Rev. Cardiol.* 16, 389–406. doi: 10.1038/s41569-019-0169-2

Bhaskar, V., Yin, J., Mirza, A. M., Phan, D., Vanegas, S., Issafras, H., et al. (2011). Monoclonal antibodies targeting IL-1 beta reduce biomarkers of atherosclerosis in vitro and inhibit atherosclerotic plaque formation in Apolipoprotein E-deficient mice. *Atherosclerosis* 216, 313–320. doi: 10.1016/j.atherosclerosis.2011.02.026

Buckley, L., and Abbate, A. (2018). Interleukin-1 blockade in cardiovascular diseases: a clinical update. *Eur. Heart J.* 39, 2063–2069. doi: 10.1093/eurheartj/ehy128

Chang, T.-H., Huang, H.-Y., Hsu, J. B.-K., Weng, S.-L., Horng, J.-T., and Huang, H.-D. (2013). An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs. *BMC Bioinformatics* 14 (Suppl 2):S4.

Chen, M., Li, W., Wang, N., Zhu, Y., and Wang, X. (2007). ROS and NF-kappaB but not LXR mediate IL-1beta signaling for the downregulation of ATP-binding cassette transporter A1. *Am. J. Physiol. Cell Physiol.* 292, C1493–C1501. doi: 10.1152/ajpcell.00016.2006

Gao, Y., Wang, J., and Zhao, F. (2015). CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol.* 16:4.

Han, B., Chao, J., and Yao, H. (2018). Circular RNA and its mechanisms in disease: from the bench to the clinic. *Pharmacol. Ther.* 187, 31–44. doi: 10.1016/j.pharmthera.2018.01.010

He, Y., Lin, L., Cao, J., Mao, X., Qu, Y., and Xi, B. (2015). Up-regulated miR-93 contributes to coronary atherosclerosis pathogenesis through targeting ABCA1. *Int. J. Clin. Exp. Med.* 8, 674–681.

Holdt, L. M., Stahringer, A., Sass, K., Pichler, G., Kulak, N. A., Wilfert, W., et al. (2016). Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. *Nat. Commun.* 7:12429. doi: 10.1038/ncomms12429

Huang, C., Yu, X., Zheng, X., Ou, X., and Tang, C. (2018). Interferon-stimulated gene 15 promotes cholesterol efflux by activating autophagy via the miR-17-5p/Beclin-1 pathway in THP-1 macrophage-derived foam cells. *Eur. J. Pharmacol.* 827, 13–21. doi: 10.1016/j.ejphar.2018.02.042

Huang, Z., Li, P., Wu, L., Zhang, D., Du, B., Liang, C., et al. (2020). Hsa_circ_0029589 knockdown inhibits the proliferation, migration and invasion of vascular smooth muscle cells via regulating miR-214-3p and STIM1. *Life Sci.* 259:118251. doi: 10.1016/j.lfs.2020.118251

Li, S., Liu, L., Zhuang, X., Yu, Y., Liu, X., Cui, X., et al. (2013). MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis. *Cell* 153, 562–574. doi: 10.1016/j.cell.2013.04.005

Libby, P., Buring, J. E., Badimon, L., Hansson, G. K., Deanfield, J., Bittencourt, M. S., et al. (2019). Atherosclerosis. *Nat. Rev. Dis. Primers* 5:56. doi: 10.1038/s41572-019-0106-z

Liu, H., Deng, Y., Wu, L., Li, Y., Lin, N., Li, W., et al. (2020). Interleukin-1beta regulates lipid homeostasis in human glomerular mesangial cells. *J. Nutr. Health Aging* 24, 246–250. doi: 10.1007/s12603-019-1302-y

Liu, L., Tan, L., Yao, J., and Yang, L. (2020). Long non-coding RNA MALAT1 regulates cholesterol accumulation in ox-LDL-induced macrophages via the microRNA-17-5p/ABCA1 axis. *Mol. Med. Rep.* 21, 1761–1770. doi: 10.3892/mmr.2020.10987

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12.

McLaren, J. E., Michael, D. R., Ashlin, T. G., and Ramji, D. P. (2011). Cytokines, macrophage lipid metabolism and foam cells: implications for cardiovascular disease therapy. *Prog. Lipid Res.* 50, 331–347.

Miao, H., Zeng, H., and Gong, H. (2018). microRNA-212 promotes lipid accumulation and attenuates cholesterol efflux in THP-1 human macrophages by targeting SIRT1. *Gene* 643, 55–60. doi: 10.1016/j.gene.2017.11.058

Ouimet, M., Barrett, T., and Fisher, E. (2019). HDL and reverse cholesterol transport. *Circ. Res.* 124, 1505–1518. doi: 10.1161/circresaha.119.312617

Ridker, P. M., Everett, B. M., Thuren, T., MacFadyen, J. G., Chang, W. H., Ballantyne, C., et al. (2017). Antiinflammatory therapy with canakinumab for atherosclerotic disease. *N. Engl. J. Med.* 377, 1119–1131. doi: 10.1056/NEJMoa1707914

Rybak-Wolf, A., Stottmeister, C., Glazar, P., Jens, M., Pino, N., Giusti, S., et al. (2015). Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Mol. Cell* 58, 870–885. doi: 10.1016/j.molcel.2015.03.027

Sankaranarayanan, S., Kellner-Weibel, G., de la Llera-Moya, M., Phillips, M. C., Asztalos, B. F., Bittman, R., et al. (2011). A sensitive assay for ABCA1-mediated cholesterol efflux using BODIPY-cholesterol. *J. Lipid Res.* 52, 2332–2340. doi: 10.1194/jlr.D018051

Tall, A., and Yvan-Charvet, L. (2015). Cholesterol, inflammation and innate immunity. *Nat. Rev. Immunol.* 15, 104–116. doi: 10.1038/nri3793

Tan, L., Liu, L., Jiang, Z., and Hao, X. (2019). Inhibition of microRNA-17-5p reduces the inflammation and lipid accumulation, and up-regulates ATP-binding cassette transporterA1 in atherosclerosis. *J. Pharmacol. Sci.* 139, 280–288. doi: 10.1016/j.jphs.2018.11.012

Tumurkhuu, G., Dagvadorj, J., Porritt, R. A., Crother, T. R., Shimada, K., Tarling, E. J., et al. (2018). Chlamydia pneumoniae hijacks a host autoregulatory IL-1beta loop to drive foam cell formation and accelerate atherosclerosis. *Cell Metab.* 28, 432–448.e4. doi: 10.1016/j.cmet.2018.05.027

Vilades, D., Martínez-Camblor, P., Ferrero-Gregori, A., Bär, C., Lu, D., Xiao, K., et al. (2020). Plasma circular RNA hsa_circ_0001445 and coronary artery disease: performance as a biomarker. *FASEB J.* 4403–4414. doi: 10.1096/fj.201902507R

Vromman, A., Ruvkun, V., Shvartz, E., Wojtkiewicz, G., Santos, M. G., Tesmenitsky, Y., et al. (2019). Stage-dependent differential effects of interleukin-1 isoforms on experimental atherosclerosis. *Eur. Heart J.* 40, 2482–2491. doi: 10.1093/eurheartj/ehz008

Wang, D., Yang, Y., Lei, Y., Tzvetkov, N., Liu, X., Yeung, A., et al. (2019). Targeting foam cell formation in atherosclerosis: therapeutic potential of natural products. *Pharmacol. Rev.* 71, 596–670.

Wang, L., Shen, C., Wang, Y., Zou, T., Zhu, H., Lu, X., et al. (2019). Identification of circular RNA Hsa_circ_0001879 and Hsa_circ_0004104 as novel biomarkers for coronary artery disease. *Atherosclerosis* 286, 88–96. doi: 10.1016/j.atherosclerosis.2019.05.006

Wei, Y., Corbalán-Campos, J., Gurung, R., Natarelli, L., Zhu, M., Exner, N., et al. (2018). Dicer in macrohages prevents atherosclerosis by promoting mitochondrial oxidative metabolism. *Circulation* 138, 2007–2020. doi: 10.1161/circulationaha.117.031589

Werfel, S., Nothjunge, S., Schwarzmayr, T., Strom, T., Meitinger, T., and Engelhardt, S. (2016). Characterization of circular RNAs in human, mouse and rat hearts. *J. Mol. Cell. Cardiol.* 98, 103–107. doi: 10.1016/j.yjmcc.2016.07.007

Williams, J., Huang, L., and Randolph, G. (2019). Cytokine circuits in cardiovascular disease. *Immunity* 50, 941–954. doi: 10.1016/j.immuni.2019.03.007

Xia, S., Feng, J., Lei, L., Hu, J., Xia, L., Wang, J., et al. (2017). Comprehensive characterization of tissue-specific circular RNAs in the human and mouse genomes. *Brief. Bioinform.* 18, 984–992. doi: 10.1093/bib/bbw081

Xiao, F. J., Zhang, D., Wu, Y., Jia, Q. H., Zhang, L., Li, Y. X., et al. (2019). miRNA-17-92 protects endothelial cells from erastin-induced ferroptosis through targeting the A20-ACSL4 axis. *Biochem. Biophys. Res. Commun.* 515, 448–454. doi: 10.1016/j.bbrc.2019.05.147

Xiao, M. S., Ai, Y., and Wilusz, J. E. (2020). Biogenesis and functions of circular RNAs come into focus. *Trends Cell Biol.* 30, 226–240. doi: 10.1016/j.tcb.2019.12.004

Yang, L., Yang, F., Zhao, H., Wang, M., and Zhang, Y. (2019). Circular RNA circCHFR facilitates the proliferation and migration of vascular smooth muscle via miR-370/FOXO1/Cyclin D1 pathway. *Mol. Ther. Nucleic Acids* 16, 434–441. doi: 10.1016/j.omtn.2019.02.028

Zeng, Z., Xia, L., Fan, S., Zheng, J., Qin, J., Fan, X., et al. (2021). Circular RNA CircMAP3K5 acts as a MicroRNA-22-3p sponge to promote resolution of intimal hyperplasia via TET2-mediated smooth muscle cell differentiation. *Circulation* 143, 354–371. doi: 10.1161/circulationaha.120.049715

Zhang, S., Song, G., Yuan, J., Qiao, S., Xu, S., Si, Z., et al. (2020). Circular RNA circ_0003204 inhibits proliferation, migration and tube formation of endothelial cell in atherosclerosis via miR-370-3p/TGFβR2/phosph-SMAD3 axis. *J. Biomed. Sci.* 27:11. doi: 10.1186/s12929-019-0595-9

Zhang, X.-O., Dong, R., Zhang, Y., Zhang, J.-L., Luo, Z., Zhang, J., et al. (2016). Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res.* 26, 1277–1287.

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

frontiers in Cell and Developmental Biology

REVIEW
published: 12 May 2021
doi: 10.3389/fcell.2021.645276

Understanding the Adult Mammalian Heart at Single-Cell RNA-Seq Resolution

Ernesto Marín-Sedeño[1,2†], Xabier Martínez de Morentin[3†], Jose M. Pérez-Pomares[1,2], David Gómez-Cabrero[3,4,5‡] and Adrián Ruiz-Villalba[1,2*‡]*

[1] Department of Animal Biology, Faculty of Sciences, Instituto Malagueño de Biomedicina, University of Málaga, Málaga, Spain, [2] BIONAND, Centro Andaluz de Nanomedicina y Biotecnología, Junta de Andalucía, Universidad de Málaga, Málaga, Spain, [3] Traslational Bioinformatics Unit, Navarrabiomed, Complejo Hospitalario de Navarra, Instituto de Investigación Sanitaria de Navarra (IdiSNA), Universidad Pública de Navarra, Pamplona, Spain, [4] Centre of Host-Microbiome Interactions, Faculty of Dentistry, Oral & Craniofacial Sciences, King's College London, London, United Kingdom, [5] Biological and Environmental Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

During the last decade, extensive efforts have been made to comprehend cardiac cell genetic and functional diversity. Such knowledge allows for the definition of the cardiac cellular interactome as a reasonable strategy to increase our understanding of the normal and pathologic heart. Previous experimental approaches including cell lineage tracing, flow cytometry, and bulk RNA-Seq have often tackled the analysis of cardiac cell diversity as based on the assumption that cell types can be identified by the expression of a single gene. More recently, however, the emergence of single-cell RNA-Seq technology has led us to explore the diversity of individual cells, enabling the cardiovascular research community to redefine cardiac cell subpopulations and identify relevant ones, and even novel cell types, through their cell-specific transcriptomic signatures in an unbiased manner. These findings are changing our understanding of cell composition and in consequence the identification of potential therapeutic targets for different cardiac diseases. In this review, we provide an overview of the continuously changing cardiac cellular landscape, traveling from the pre-single-cell RNA-Seq times to the single cell-RNA-Seq revolution, and discuss the utilities and limitations of this technology.

Keywords: single-cell RNAseq, heart, infarction, cardiac cell heterogeneity, transcriptomics

OPEN ACCESS

Edited by:
Maarten M. G. van den Hoogenhof, Heidelberg University Hospital, Germany

Reviewed by:
Carlos Talavera-Lopez, Wellcome Sanger Institute (WT), United Kingdom
Michela Noseda, Imperial College London, United Kingdom
Marco Osterwalder, University of Bern, Switzerland

***Correspondence:**
David Gómez-Cabrero
david.gomez.cabrero@navarra.es
Adrián Ruiz-Villalba
adruiz@uma.es

[†] These authors have contributed equally to this work and share first authorship
[‡] These authors have contributed equally to this work and share senior authorship

Specialty section:
This article was submitted to Molecular Medicine, a section of the journal Frontiers in Cell and Developmental Biology

Received: *22 December 2020*
Accepted: *09 April 2021*
Published: *12 May 2021*

Citation:
Marín-Sedeño E, de Morentin XM, Pérez-Pomares JM, Gómez-Cabrero D and Ruiz-Villalba A (2021) Understanding the Adult Mammalian Heart at Single-Cell RNA-Seq Resolution. Front. Cell Dev. Biol. 9:645276. doi: 10.3389/fcell.2021.645276

INTRODUCTION

The heart is the first fully functional organ forming during embryonic development [Hamburger and Hamilton stages (HH) 9–10 in the chicken, at embryonic day (E) 8 in the mouse, and Carnegie stages (CS) 9–10 in the human] (Hamburger and Hamilton, 1951; Theiler, 1972; Müller and O'Rahilly, 1987). In mammals, the cardiac primordium starts beating very soon (around E8.5 in the mouse, and between 26 and 32 days after conception in the human) (Howe et al., 1991; Porter and Rivkees, 2001), and progressively transforms into a four-chambered heart that does not stop its contractile activity until the death of the organism (Buckingham et al., 2005). Coordination is, no doubt, one of the most remarkable features of cardiac cell function. Indeed, cardiac pumping activity requires the continuous interaction of a great variety of cell types, including contractile

cardiomyocytes, endocardial cells, vascular endothelial and smooth muscle cells, cardiac fibroblasts and pacemaker cells, among many others. Cardiac cell interactions, mediated by juxtacrine, paracrine and endocrine signals, sustain cardiac homeostasis and are essential to articulate heart responses to pathologic stimuli (Buckingham et al., 2005; Rog-Zielinska et al., 2016; Pogontke et al., 2019). These interactions are, at least in part, modulated by the specific functional profile of different cardiac chambers (e.g., intra-chamber pressure and resistance to blood flow are significantly different in the left versus the right ventricle). Moreover, in a pathologic context, cardiac function can be severely modified. Such functional changes may result from altered inter-cellular signaling, loss of cell activity or cell death, but it is also evident that changes in cardiac performance can then result in further changes in cell-to-cell communication.

Surprisingly, our knowledge of cardiac cell diversity in the mammalian heart has remained limited, which hampers our understanding of the heart as a complex cellular system requiring the finely orchestrated and accurate activity of its cellular components. Quantifying cardiac cellular components was always a goal of classical organographists (Roberts and Wearn, 1941), while researchers studying cardiac embryonic development have provided a continuous stream of relevant information about the cell types that build up the heart. Cardiogenesis, which involves the patterned differentiation of different cell types from multiple embryonic heart fields (Buckingham et al., 2005; Srivastava, 2006; Dunwoodie, 2007; Meilhac and Buckingham, 2018), is a unique context for the study of cardiac cell diversity. The genuine interest of cardiac embryologists in understanding the diversification of cardiac cell types from their mesodermal progenitors has generated a considerable volume of research in this field over the last two decades.

Cardiac cell differentiation and diversification have been studied using a great variety of cell tracing techniques in different experimental models. Genetic cell tracing methods, based on the known specific activity of one or several genes in certain cell types, have been extensively used. Unfortunately, this approach often involves the use of one gene at the time. Therefore, the complexity underlying a specific cell lineage is unlikely to be revealed using such methods, as we will further discuss in this review (see Meilhac and Buckingham, 2018).

Gene expression profiling at single-cell RNA sequencing (scRNA-Seq) resolution is a revolutionary and robust high throughput technology that allows for the understanding of complex biological systems from the analysis of their components via the unbiased identification of new cell types, cell lineage progression, and cellular plasticity in dynamic processes (Massaia et al., 2018; Meilhac and Buckingham, 2018; Mereu et al., 2020). One of the most relevant properties of this technology is its capacity to identify cellular heterogeneity at single cell resolution as based on the prevalence and/or co-expression of genes. On the other hand, scRNA-Seq allows for the unbiassed approach to the massive analysis of gene expression, and has a remarkable ability to compare data from several species in an integrative manner. Moreover, due to its excellent resolution power, scRNA-Seq is breaking new ground in several biological disciplines (Stubbington et al., 2017; Zheng et al., 2017; Xia and Yanai, 2019; Imdahl and Saliba, 2020; Jakab and Augustin, 2020; Rich-Griffin et al., 2020; Zhou and Wang, 2020). In fact, the use of scRNA-Seq technology constitutes a milestone in the history of research on the human body, leading to the generation of a human organ atlas at single-cell resolution (Haber et al., 2017; Han et al., 2018; Vanlandewijck et al., 2018; Aizarani et al., 2019; Kalucka et al., 2020; Litviňuková et al., 2020; Travaglini et al., 2020).

In the context of cardiac research, the high resolution of scRNA-Seq technology has been key to the dissection of cardiac cell diversity. Such diversity is evident from the early stages of cardiac development. Indeed, the precardiac mesoderm has been shown to be heterogeneous, comprising cells characterized by the activation of significantly different transcriptional programs that affect their developmental fate. Early in mammalian development, cardiac precursors organize into two cardiac progenitor domains (the First Heart Field, FHF and the Second Heart Field, SHF) that sequentially contribute to the formation of the embryonic cardiac primordium. FHF cells characteristically express the transcription factors Nkx2.5, Mef2C, Srf and Gata4, whereas SHF cells can be primarily identified by the expression of the transcription factors Isl1 and Tbx1 (Kelly et al., 2014). Multiple studies have characterized the properties of these two morphogenetic fields, it has been only since the advent of scRNA-Seq approaches that we have really started to understand the real heterogeneity of these cell populations. Lescroart et al. (2018) have recently identified that the first cardiovascular progenitor cells (CPCs), Mesp1^{POS} cells derived from the primitive streak around E6.5, are heterogeneous at scRNA-Seq resolution. This study provides robust data indicating that the capacity of these cells to differentiate into specific cardiovascular lineages is temporal and spatial determined at early gastrulation stages (Lescroart et al., 2018). In the same line, Jia et al. (2018) recently described three Nkx2.5^{POS} and five Isl1^{POS} different subpopulations of CPCs with significant differences in their trajectorial patterns and chromatin accessibility, two features that are thought to determine their differentiation capacity. These two studies highlight the enormous complexity that underlies the regulation of decisions during heart development, and are a good example of the remarkable resolution of scRNA-Seq analysis even when dealing with small number of cells.

In this review, we summarize current knowledge regarding the heterogeneity of the main cellular components of the adult mammal heart and their dynamics in both healthy and pathological conditions. In order to do so, we compare the vision of cardiac cell composition before and after the appearance of scRNA-Seq technology. In addition, we also describe the main drawbacks and limitations of this powerful tool and discuss the future perspectives it opens in the cardiovascular research field.

CELLULAR LANDSCAPE OF THE ADULT HEART BEFORE SINGLE CELL RNA-SEQ TECHNOLOGY

The systematic analysis of cardiac cell composition was initiated during the first half of the 20th century

(Roberts and Wearn, 1941). Using stereological and morphometric approaches in adult rats, cardiomyocytes (CMs) were found to cover 75% of the total volume of the heart, although it soon became clear they accounted for 25-43% of the cardiac cells only (Zak, 1973; Anversa et al., 1980; Nag, 1980). Some years later, an important study based on flow cytometry and immunohistochemistry suggested that significant difference in the cellular composition of the mouse versus the rat heart existed (26.4 ± 5.8 vs. 55.9 ± 8.3% of CMs, respectively) (Banerjee et al., 2007). Notwithstanding this, a consensus was reached on this topic, and cardiomyocytes were estimated to represent around 30% of total cardiac cells (Walsh et al., 2010; Bergmann et al., 2015). In contrast with studies on CMs, research on non-myocyte cells, including cell types as important as cardiac fibroblasts and endothelial cells, has remained scarce and inconclusive.

The careful evaluation of these studies suggests that their different conclusions can be accounted for by their extensive use of histological technologies to quantify cell numbers (Zak, 1973; Vliegen et al., 1991; Banerjee et al., 2007; Tang et al., 2009). To overcome the bias inherent in histological approaches and the occasional lack of specificity of markers for specific cell types, cell lineage tracing methods in combination with other technologies were used. Primarily, cell-tracing techniques, ranging from the use of carbon particles to auto-radiographic tattooing and advanced fluorescent dyes, were applied to describe cell migration or conformational changes in the developing heart (Castro-Quezada et al., 1972; Thompson et al., 1987; Patwardhan et al., 2000). Unfortunately, these methods frequently lack single-cell resolution, and provide cell tagging that can be lost over time (e.g., by dilution of the dye). Replication-defective retroviruses have also been effectively used to trace the origin and differentiation of cardiac cells (Mikawa and Fischman, 1992; Mikawa et al., 1992), but the retrovirus production process, and the need for highly accurate local viral delivery, limit the scope of this method. Paradoxically, lessons from the experimental use of retroviruses to label cells using viral genome integration in the host cell DNA strongly boosted the use of transgenic mouse lines to unambiguously track and identify cells.

The first transgenic mouse lines generated to tag cardiovascular cells were based on constructs containing a reporter cassette (most typically *LacZ* or *GFP*) under the control of a specific gene promoter sequence (Fadel et al., 1999; Didié et al., 2013; Velecela et al., 2013). In these animal models, the reporter expression is dependent on tissue/cell type-specific promoter activity, thus preventing the permanent genetic tracing of cell progenies. The advent of the Cre/LoxP (Soriano, 1999) or Tet OFF (Harding et al., 1998) genetic technologies addressed this issue, and the number of studies using these methods to study cell progenies increased enormously. Genetic cell tracing methods are extremely powerful and still have a deep impact on our understanding of heart cell composition, structure, and responses to pathologic stimuli (Moses et al., 2001; Huynh et al., 2007; Engleka et al., 2012; Ruiz-Villalba et al., 2015; Cano et al., 2016). However, these technologies have drawbacks, such as the association of a specific cell type with the expression of a single gene or the erroneous interpretation of the concept of cell lineage as based on the expression of such a gene (for an authoritative review on this topic see Meilhac and Buckingham, 2018).

Most cardiomyocyte subtypes are anatomically patterned. Contractile working CMs present in the left ventricle are mainly derived from cardiac progenitors of the first heart field, whereas the principal origin of right ventricular working CMs is the secondary heart field (Kelly et al., 2014; Meilhac and Buckingham, 2018). Endothelial cells (ECs) in the coronary vascular system mainly originate from the sinus venosus and ventricular endocardium (Red-Horse et al., 2010; Wu et al., 2012), with the contribution of septum transversum/epicardium (Cano et al., 2016) and other extracardiac endothelial cells (Palmquist-Gomes et al., 2018). The vast majority of cardiac ECs occupy the cardiac interstitium, a unique harboring compartment hosting a large variety of cells (Pogontke et al., 2019). Cardiac fibroblasts (CFs) are prototypical interstitial cells, and most CFs derive from the embryonic epicardium (Wessels et al., 2012; Moore-Morris et al., 2014; Ruiz-Villalba et al., 2015; Kanisicak et al., 2016), but endocardial-derived fibroblasts have also been described (Zeisberg et al., 2007). The cardiac interstitium is also the home of blood-borne cells that are recruited to the heart during embryonic development and postnatal stages. These cells, which will become cardiac resident, include different populations of monocytes/macrophages (Epelman et al., 2014; Ruiz-Villalba et al., 2015; Dick et al., 2019; Sampaio-Pinto et al., 2020).

In order to characterize and quantify cardiac cell composition, a sophisticated study was carried out using a FACS-based approach in murine and human hearts (Pinto et al., 2016). In this paper, the authors suggest that ECs are the most abundant cells in the adult heart, and that CFs and immune cells contribute less than 20% and 10%, respectively, to the non-myocyte cardiac cell fraction. These data differ from those discussed above, highlighting how differences in the experimental design of these type of analyses can result in significantly different results (Lescroart et al., 2012; Meilhac and Buckingham, 2018). It is anyway evident that all these studies have significantly helped to increase our knowledge of cardiac cell heterogeneity, especially in some experimental animal models like the mouse, even if this knowledge cannot be always extrapolated to humans.

In this regard, single-cell "-omics" appear as a set of revolutionary, unbiased methods allowing for the characterization of cardiac cell heterogeneity. In the last 5 years, the cardiovascular field has moved from the "cell-marker" concept, mostly used to define a "specific" cell type or lineage in the mouse, to the development of a first map draft of cardiac cell diversity based on single-nuclei RNA-Seq data and specifically annotated for the human heart (Litviňuková et al., 2020). In the following sections, we describe the contribution of scRNA-Seq analysis to (1) the state of knowledge about cardiac cell heterogeneity prior to the single-cell "-omics" era; and (2) how these techniques are significantly changing our understanding and experimental strategies to study the heart. To do that, we describe the most recent published scRNA-Seq data related to cardiac cell heterogeneity, both in mice and humans.

CARDIOMYOCYTES

Cardiomyocytes are required for cardiac contraction and the consequent distribution of blood throughout the organism. This crucial function is based on the coordinated depolarization of electrically coupled CMs in the atria and ventricles regulated by the cardiac conduction system (CCS) (Ng et al., 2010; van Eif et al., 2018). For this reason, CMs have been historically classified in terms of their contractile properties and anatomical location. In the pre scRNA-Seq era, CMs were identified as based on their expression of several common markers associated with the contractile machinery and coordination (*Tnnt2*, *Tpma*, *Tnnc1*, *Tnni3, Actc1, Ryr2*). However, atrial and ventricular compartments display significant differences in their gene expression profiles (Sehnert et al., 2002; Ng et al., 2010). Adult atrial CMs widely express the light and heavy myosin chains *Myh4*/ALC1 and *Myh6*/αMHC, whereas the light chain *Myl7*, the calcium regulator *Atp2a2*/SERCA2a, and the channel *Gja5*/CX40 present a more restricted expression. Ventricular CMs predominantly transcribe *Myh3*/ELC and *Myh7*/βMHC myosin genes, and specifically express *Myl2* and the potassium channel *Kcne1*. The transcriptional identity of each chamber has been associated with specific genes: *Hey1* in the atria and *Hey2* and *Irx4* in the ventricles (Ng et al., 2010; **Figure 1**).

The cardiac pacemaker or CCS is formed by specialized cells able to transduce the electrical stimulus that triggers cardiomyocyte contraction. These cells derive from primitive cardiomyocytes patterned in space to form a complex network of nodes (e.g., sinoatrial node; atrioventricular) and conduction tracts of different size (e.g., Bachmann's bundle; bundle of His; right and left bundle branches; Purkinje fibers) (van Weerd and Christoffels, 2016). Primary functions of these different elements include the cyclic initiation of the electrical impulse (sinoatrial node), the fast propagation of the electric signal (CCS tracts or the delay of this same stimulus to guarantee atrial and ventricular systolic phases are asynchronous). Important parts of the CCS are protected by cardiac fibrous tissue, and recent studies have shown that macrophages play an important role in the normal functioning of the cardiac pacemaker (Anderson and Ho, 1998; Hulsmans et al., 2017). Adult CMs from the CCS have been traditionally identified as based on the expression of *Hcn4* (Liang et al., 2013) and *Cntn2*/Contactin-2 (Pallante et al., 2010). However, there are other markers such as *Gja1*/CX43, *Gja5*/CX40, *Smoc2*, *Isl1* (van Eif et al., 2019) or *Tbx3* (Hoogaars et al., 2004) that show a heterogeneous expression pattern within the different parts of the CCS (nodes, atrioventricular bundle, and the Purkinje fiber network) (**Figure 1**).

In this context, scRNA-Seq has significantly contributed to the study of cardiomyocyte heterogeneity. For example, different laboratories have recently identified a specific subpopulation of CMs characterized by enrichment in the expression of *Myoz2*, both in the adult murine and the developing human heart (Gladka et al., 2018; Asp et al., 2019). In adult human samples, transcriptomic profiling of CMs isolated from different anatomic compartments suggests functional specialization (Wang et al., 2020). Accordingly, small sets of CMs isolated from atria and ventricles showed significant differences in their metabolic activity, retinoic acid-responsive capacity and smooth muscle cell transcriptomic profiles (Litviňuková et al., 2020). Interestingly, the expression of *Smarca4*, a gene that defined the atrioventricular CM cluster, has been associated with cardiac hypertrophy (Wang et al., 2020). In contrast to their working myocardium counterparts, the single-cell identity of the adult CCS remains to be fully elucidated. Until now, scRNA-Seq has uncovered a complex landscape within cell-type heterogeneity and cellular transitional phenotypes in the human CCS during development (Goodyer et al., 2019). In the same study, the authors identified different cell populations in all the CCS components, and found that while the expression of *Igfbp5*, *Cpne5,* and *Ntm* is enriched in the entire CCS, *Smoc2* is exclusively expressed in the sinoatrial node.

Studies based on the scRNA-Seq technology have also contributed relevant data to the contentious topic of adult cardiomyocyte proliferation abilities. CMs have been convincingly shown to display some turnover during adult life (Bergmann et al., 2015). However, mammalian CMs are unable to regenerate the heart upon injury even if their proliferation under stress conditions seems to be significantly increased (Kimura et al., 2015). Proliferative CMs have been found in both embryonic and neonatal mouse hearts, but scRNA-Seq studies have been unable to identify proliferating cardiomyocytes, either in the adult homeostatic or the infarcted heart (Kretzschmar et al., 2018; Li G. et al., 2020). Only an integrative analysis using data from the "Tabula Muris" compendium (Schaum et al., 2018) and single-nuclei RNA-Seq (snRNA-seq) approaches has revealed the existence of a rare cardiomyocyte subpopulation (0,4% of total cardiac cells) displaying proliferative markers (Galow et al., 2020). Using the same single-cell technology, a previous report revealed that long intergenic non-coding RNAs were key regulators of the cell cycle in a subpopulation of CMs, suggesting the presence of CMs with an inherent proliferative ability in the adult heart (See et al., 2017). Taken together, these results emphasize that scRNA-Seq has not been able to definitively identify proliferative CMs. However, this technology can be used in combination with different high throughput technologies, such as ATAC-Seq, proteomics, or metabolomics, to properly address this search for future regenerative interventions (Hu et al., 2018; Kretzschmar et al., 2018; Li G. et al., 2020).

Deciphering the dynamics of cardiac cell populations upon injury has been another important aim of single-cell studies in the field. Hypertrophic CMs display high expression of the disease-related genes *Nppa*, *Nppb,* and *Vegfa* (Yekelchyk et al., 2019; McLellan et al., 2020). In the case of *Nppa* and *Nppb*, although their relevance in cardiac development and disease had been previously reported (Sergeeva et al., 2014), scRNA-Seq technology has revealed the transcriptomic signature that describes specific subpopulations of CMs involved in cardiac disease. Mono- and multi-nucleated myocytes have been shown to have similar transcriptional profiles in both homeostatic and pathological conditions (Yekelchyk et al., 2019). In human hearts with end-stage cardiac failure (dilated cardiomyopathy and coronary disease), a recent scRNA-Seq study has pointed out that the left ventricle is always the most severely affected compartment (Wang et al., 2020). Besides the commonly downregulated genes

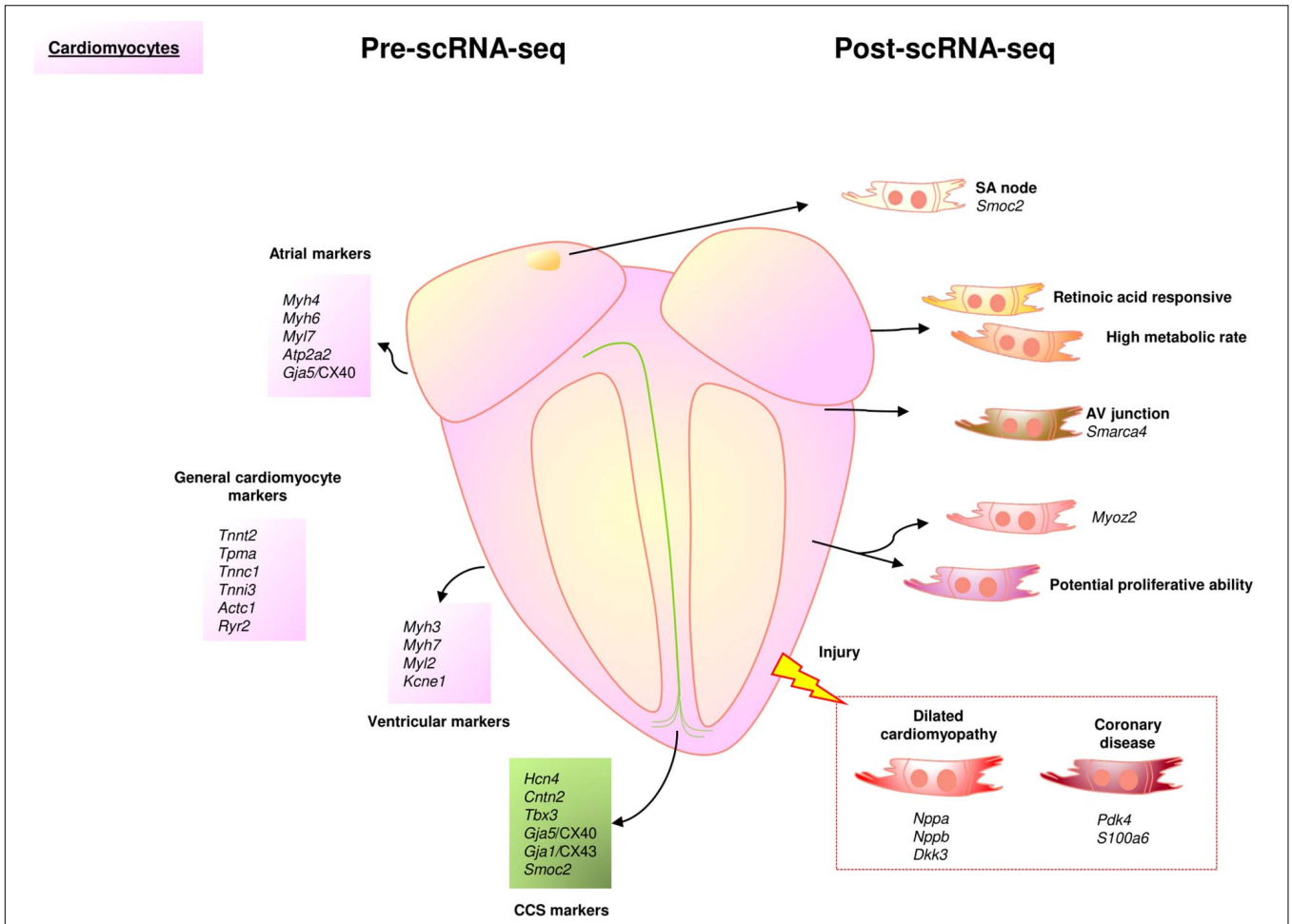

FIGURE 1 | Cardiomyocytes landscape before and after scRNA-Seq. In the scRNA-Seq part, each cell represents an identified cluster characterized by one or more genes, a specific cell function or other cell aspects. *Nppa* and *Nppb* gene expression identifies characteristic cardiomyocyte subpopulations in the pathological context of dilated cardiomyopathy. CCS, cardiac conduction system; SA, sinoatrial; AV, atrioventricular.

like *Spp1* or the transcription factors *Tcf7l2* and *Cebpd*, other differentially expressed genes (*Pdk4* and *S100a6* in coronary disease and *Nppa*, *Nppb*, and *Dkk3* in dilated cardiomyopathy) may help to identify specific mechanisms associated with these two cardiac diseases (Wang et al., 2020; **Figure 1**).

The analysis of CMs at single-cell resolution also has its drawbacks (Li G. et al., 2020). These include the use of protocols for tissue dissociation (which can damage and destroy cells) and the variable size of adult CMs, as this can alter cell capture in single-cell platforms (Gladka et al., 2018; Zhou and Wang, 2020). In any case, scRNA-Seq has been proven to be a very robust technology. Indeed, although the ploidy of CMs and the nuclear transcripts may affect the results obtained in these experiments (Zhou and Wang, 2020), similar transcriptional profiles were obtained when comparing data from single-cell and single-nuclei RNA-Seq (Selewa et al., 2020). Regarding heart composition, snRNA-Seq revealed that CMs represent between 23 and 49% of cardiac cells (Galow et al., 2020; Litviňuková et al., 2020; Tucker et al., 2020; Wolfien et al., 2020a,b) in contrast to the 9% indicated from the scRNA-Seq of the "Tabula Muris" project (Schaum et al., 2018). Different proportions of CMs have been found between the atria (30%) and ventricles (49%) in the human heart (Litviňuková et al., 2020), suggesting additional chamber-specific differences in cardiac cell distribution with functional implications. From a technical standpoint, we believe it is necessary to explore the limitations of scRNA-Seq in order to improve our understanding of cardiomyocyte diversity. This will require the simultaneous analysis of spatial gene expression patterns and the evaluation of the biological roles of genes, not only in homeostatic, but also in pathologic contexts.

ENDOTHELIAL CELLS

Cardiac ECs line the inner surface of cardiac chambers, blood and lymphatic vessels. They play a role in essential cardiovascular cell functions like permeability, leukocyte trafficking, hemostasis, thermoregulation or angiogenesis (Aird, 2007a; Talman and Kivelä, 2018). These cells display different morphological features across the vascular tree, but also respond to different gene

expression programs (vascular versus lymphatic; arterial versus venous; and large versus microvascular endothelial phenotypes) (Chi et al., 2003; Nolan et al., 2013). Classically, the majority of endothelial markers found in the literature are related to the unique properties and functions of ECs, as is the case with *Pecam1*/CD31, *Cdh5*/VE-cadherin, *Sele*/E-selectin, *Icam2*/CD102, *Flt1*/VEGFR1, *Tie2*/Tek, *Eng*/CD105 (Garlanda and Dejana, 1997; Brutsaert, 2003; Aird, 2007a; Banerjee et al., 2012). Some other genes, however, have been found to be preferentially expressed in arterial (*Efnb2*, *Dll4*, *Hey1/2*, *Nrp1*), venous (*Ephb4*, *Nrp2*, *Nr2f2*/COUP-TFII) (Aird, 2007b), or lymphatic ECs (*Prox1*, *Lyve1, Flt4*/VEGFR3) (Wigle and Oliver, 1999; Mäkinen et al., 2001; **Figure 2**).

Despite the established knowledge on the heterogeneity of ECs, scRNA-Seq has allowed us to gain a deeper understanding of this complex population of cardiac cells. A very recent study has summarized the discoveries of endothelial cell diversity (or "angiodiversity" as the authors call it) resulting from the use of scRNA-Seq (Jakab and Augustin, 2020). Even though this technology has confirmed the role of some of the genes quoted above as markers for ECs, specific metabolic and gene expression programs were spotted and analyzed across the vascular tree. In a recent paper, Kalucka et al. (2020) have described several subsets of ECs with specialized phenotypes in the homeostatic heart, including capillary ECs with an interferon-induced gene program and some subpopulations with a characteristic angiogenic signature. Interestingly, the authors of this work also described how microvascular ECs were found to be the most heterogeneous endothelial cell pool within the same organ and even among different ones. Other authors proposed *Rgcc* as a specific marker for capillaries (Schaum et al., 2018; Kalucka et al., 2020). The endocardium, the special endothelium that covers internal cardiac chamber walls, is another interesting case, as it displays a high expression of *Npr3* and *Cytl1* (Feng et al., 2019), which are also expressed in some capillaries and arteries (Hu et al., 2018; Kalucka et al., 2020), and cannot be therefore regarded as endocardial markers. Interestingly, scRNA-Seq has confirmed previous bulk RNA-Seq data describing a specific transcriptomic signature for cardiac ECs (Coppiello et al., 2015). This signature includes genes involved in fatty acid uptake and metabolism like *Fabp4*, *Cd36*, *Pparg*, *Tcf15, Aqp7* and *Meox2* (Feng et al., 2019; Jambusaria et al., 2020; Kalucka et al., 2020), all of which could be relevant to our understanding of cardiomyocyte bioenergetics (Lother et al., 2018). In humans, some clusters of ECs were even identified on the basis of their propensity to secrete cytokines or their implication in immune response and cell-to-cell assembly (Wang et al., 2020; **Figure 2**).

Unlike CMs, ECs of adult hearts can proliferate (Kretzschmar et al., 2018). The use of scRNA-Seq technology has refined this concept by showing that a cluster of pre-existing ECs drives vasculogenesis in a clonal way in the healthy heart (Li et al., 2019). Rare subpopulations of ECs that co-express cardiomyocyte markers such as *Tnnt2* or *Actc1* have also been identified in data from sc/snRNA-Seq experiments (Hu et al., 2018; Nomura et al., 2018; Li et al., 2019; Lukowski et al., 2019; Galow et al., 2020; Jambusaria et al., 2020; Wolfien et al., 2020a). Whether these cells are involved in the modulation of the activity of CMs, as suggested by some authors (Jambusaria et al., 2020) is still unclear.

Single-cell RNA-Seq technology has also been used to study neovascularization phenomena after myocardial infarction (MI). Recent work has described the rise of several endothelial cell subsets enriched in proliferation markers (*Mki67*, *Cdc20*), showing either interferon or retinoic acid pathway signatures (Li et al., 2019). It is noteworthy that these "response to damage" endothelial subsets, enriched in *Plvap* expression, have been found to accumulate preferentially in the border zone of mouse and human infarcted hearts (Li et al., 2019). This finding led the authors to propose *Plvap* as a marker for neovasculogenesis and a target for future therapeutic approaches. Following a similar rationale, other groups have suggested that pre-existing pools of ECs are the primary source of new blood vessels in ischemic tissues (He et al., 2017; Manavski et al., 2018). As proof of concept, the administration *Ackr1*$^{+}$ ECs, a subpopulation of ECs identified in healthy hearts exclusively, improved cardiac function when applied to a mouse infarct model. Because of that, these authors also considered these cells a potential therapeutic target to treat cardiac diseases (Wang et al., 2020).

FIBROBLASTS AND MURAL CELLS

The study of CFs has been one major objective in the study of cardiac cells at single cell resolution. CFs are classically associated with the synthesis, deposition, and remodeling of the extracellular matrix (ECM). Moreover, they also contribute to cardiac homeostasis, communicate with immune cells, sustain cardiomyocyte electrical coupling and are essential to stress sensing (Camelliti et al., 2005; Souders et al., 2009; Frangogiannis, 2020). There is a set of "classical" markers defining CFs both under normal (CD90.1/THY1, PDGFRA, S100A4/FSP1, DDR2, SCA1, VIMENTIN, COL1A1) or pathological conditions (PERIOSTIN, αSMA*)*. However, most of these markers are not exclusive to CFs (Tallquist and Molkentin, 2017). For example, collagens are produced, to a certain extent, by many non-fibroblastic cells (Hynes, 2012; Fidler et al., 2017). Other molecules closely associated with CFs like VIMENTIN, CD90 or FSP1 are also found in endothelial and immune cells (Tallquist and Molkentin, 2017; **Figure 3**). As expected, the scRNA-Seq technology has enormously contributed to the identification and understanding of CFs heterogeneity in normal and pathological conditions.

First, scRNA-Seq has unequivocally confirmed that CFs are a heterogeneous cell type. Cardiac fibroblast populations can be characterized by their variable expression of classical markers such as *Pdgfra* and *Tcf21*, and different functions have been suggested for some of these fibroblast types (Kretzschmar et al., 2018; Skelly et al., 2018; Farbehi et al., 2019; McLellan et al., 2020; Ruiz-Villalba et al., 2020). For example, a subpopulation of *Pdgfrα*^{POS} CFs with a high expression of *Ly6a* and *Thy1* has been related to the response to interferon (Farbehi et al., 2019), whereas Muhl and colleagues describe a subpopulation of *Wif1*^{POS}/*Comp*^{POS} CFs associated with valve interstitial cells

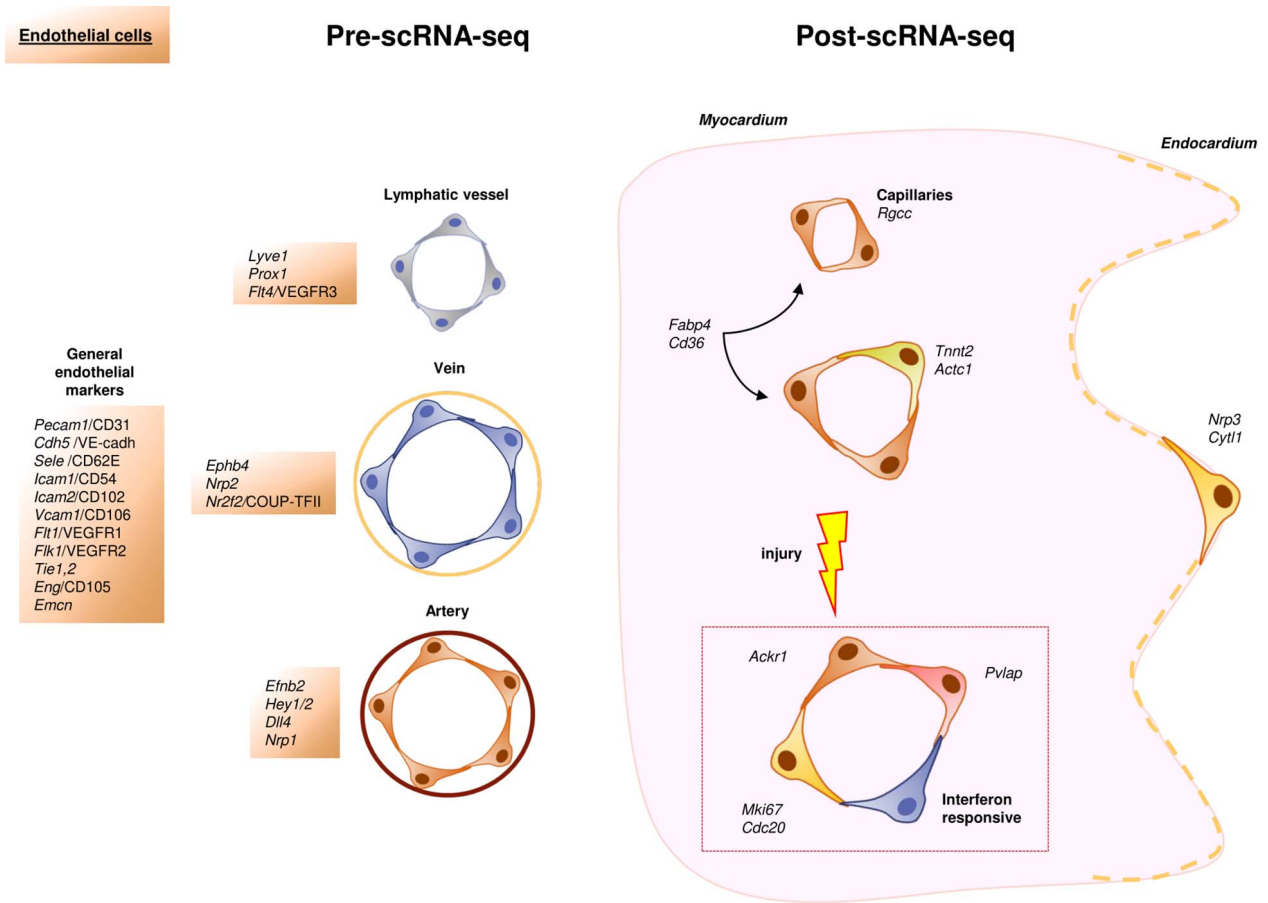

FIGURE 2 | Schematic representation of the cardiac endothelial cellular landscape before and after scRNA-Seq. In the scRNA-Seq part, each color represents an identified cell subpopulation characterized by specific genes, cellular functions or other relevant aspects.

and suggest that these CFs modulate ECM in an organ and location-specific manner (Muhl et al., 2020).

This association between anatomical regions and specific fibroblast functions has been recently reported for human CFs isolated from different regions of homeostatic hearts (Litviňuková et al., 2020). In this paper, seven subpopulations of CFs are identified as based on the expression of molecules such as DCN, GSN, and PDGFRA, FB1 and FB2, showed a basal, chamber-specific fibroblast expression profile; FB3 CFs are stress-responsive and could contribute to sustain cardiac homeostasis; FB4 CFs seem to be more responsive to TGFβ signaling; and FB5 CFs are characterized by the expression of genes involved in the production, remodeling, and degradation of ECM. However, further detailed studies are needed to analyze differences in cardiac fibroblast heterogeneity between species, taking into account the technical limitations related to each one of them (e.g., in contrast to murine hearts, the isolation of the whole human CFs population is not possible, so that we have to assume that small pieces of tissue are representative enough of the complete organ).

Second, the study of CFs in the context of heart disease is of extreme relevance considering the impact of fibrosis in the progression of many cardiac ailments. Different studies have found interesting patterns of fibroblast activation in injured tissues that are associated with different functional properties and stages of disease progression in several pathologies (Frangogiannis, 2020). A perfect example of the crucial role played by CFs in disease is MI. The first 48 h after MI are characterized by a rapid inflammation and leukocyte recruitment partially promoted by CFs (Mouton et al., 2019). Following this first step, a subpopulation of fibroblast-like or stromal cells with a pro-inflammatory and pro-survival gene signature appears in the heart roughly 24 h after the damage (*IR* in Forte et al., 2020). Within the first 3 days post-infarction (dpi), several subpopulations of CFs, described by different groups, appear in the heart. In a pioneer scRNA-Seq study, a cardiac fibroblast subpopulation enriched in ECM-related genes (*Col1a1*, *Cthrc1*, *Postn*, *Fn1*, *Tnc*) was described 3 dpi (Gladka et al., 2018). These authors identified *Ckap4* gene as a marker for these activated CFs, in both mice and human cardiac ischemic samples. Between 3 and 7 dpi, several fibroblast clusters with proliferation abilities and significant expression of *Acta2*, appear near the infarct zone (Fu et al., 2018; Farbehi et al., 2019; Forte et al., 2020). In this "proliferative" period of ventricular remodeling,

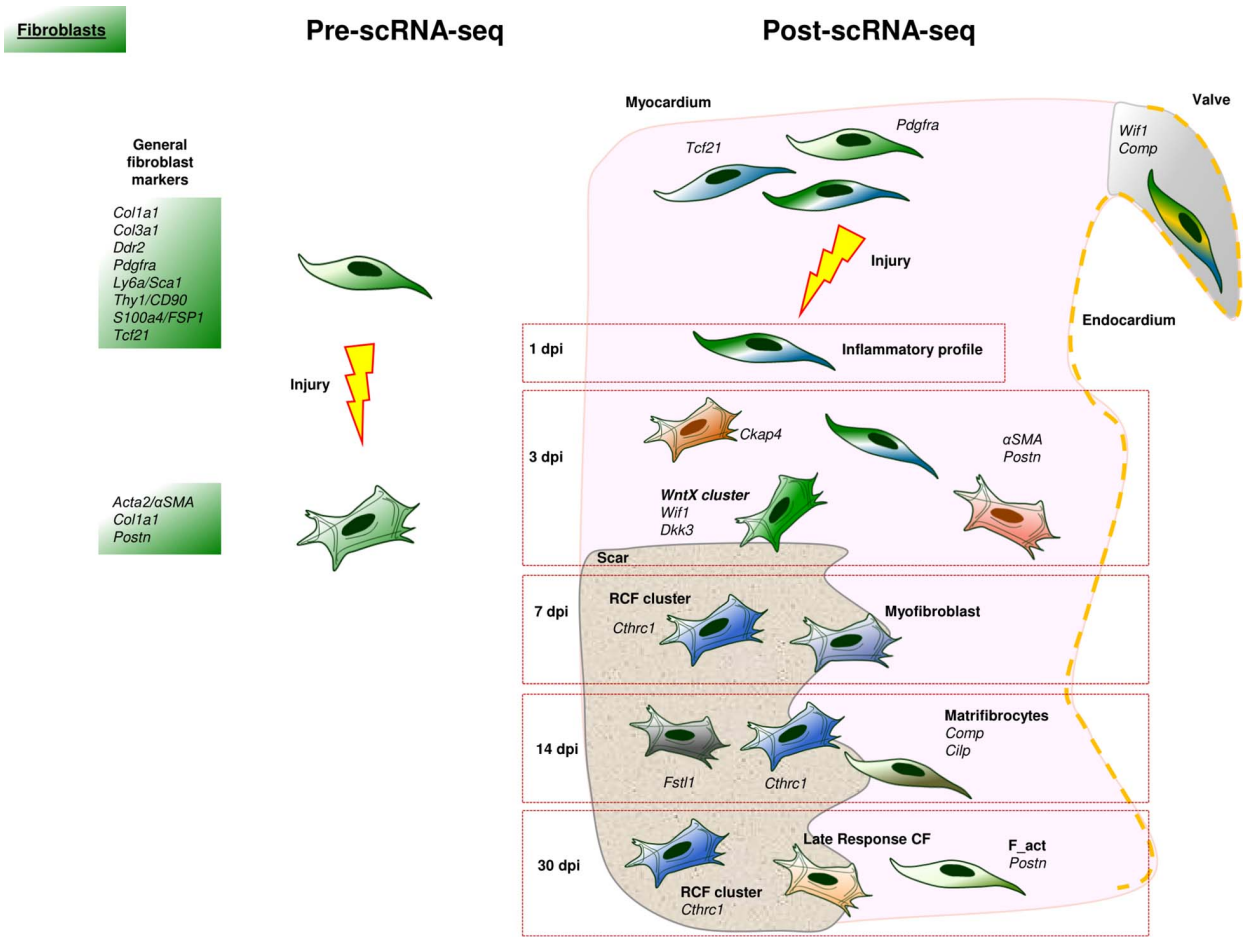

FIGURE 3 | Cardiac fibroblast landscape before and after scRNA-Seq. In the scRNA-Seq part, each cell represents an identified subpopulation of cardiac fibroblasts characterized by a specific transcriptomic signature. Some of these populations have been named after the original study that identified them.

minor CF-like interstitial subpopulations were identified by their myeloid-phagocytic profile (Forte et al., 2020). Furthermore, a Wnt-related cluster (*Wif1*^{POS}, *Dkk3*^{POS}) was described and found to correspond to cells located in the scar and border zone (*WntX* in Farbehi et al., 2019). A similar cardiac fibroblast subpopulation was reported to be endocardial-derived (*EndD*, see Forte et al., 2020).

This scRNA-Seq transcriptomic study indicates that the majority of CFs responding to cardiac damage generate in the embryo from the epicardium, confirming the results from previous studies using genetic cell-tracing tools (Ruiz-Villalba et al., 2015; Moore-Morris et al., 2018). In accordance with this discovery, it is the current consensus that resident interstitial mesenchymal populations are the main source of activated CFs, and that the embryonic epicardium is the main origin of these stromal cells (Tallquist and Molkentin, 2017). The fundamental question, however, is whether all the CFs are equal. Recent studies suggest this is not the case, as shown by our own research. We have recently described a subpopulation of CFs with a crucial role in cardiac repair (*Reparative Cardiac Fibroblasts*, RCFs) that closely cluster after scRNA-Seq analysis. These RCFs appear in the infarct and border zones from 7 to 30 dpi in mice and are also present in ventricular remodeling tissues in both in pigs and human infarcted hearts. *Cthrc1* is the top marker gene of RCFs, which also express genes of the non-canonical TGFβ1/PI3K-Akt pathway (Ruiz-Villalba et al., 2020). RCFs have a gene expression pattern similar to that of other CF clusters described by Forte and colleagues, who regarded these cells as *myofibroblasts* (Forte et al., 2020). On a final note, RCFs were also detected in murine fibrotic hearts after continuous treatment with angiotensin-II and patients suffering cardiac hypertrophy (Ruiz-Villalba et al., 2020). The dynamics of these cells in both murine models of cardiac fibrosis suggests a similar role in the early stages of the pathology.

Differences in cardiac fibroblast heterogeneity are also found in the maturation phase of ventricular remodeling. A scRNA-Seq experiment has revealed a high expression of *Fstl1* in activated CFs at 14 dpi (Kretzschmar et al., 2018). At these stages of the remodeling process, the scar contained "matrifibrocytes," a subset of activated CFs described by two different research teams (Fu et al., 2018; Forte et al., 2020). These CFs are characterized by their reduced secretory activity, restricted contractility, and low

proliferative capacity, as well as by the expression of extracellular matrix and tendon related genes (such as *Comp* and *Cilp*). These data suggest that these cells are acquiring a more specialized, structural, and supporting phenotype in the mature scar (Fu et al., 2018). Other two subpopulations of activated CFs, with different transcriptomic profiles, persist in the mature scar around 30 dpi (Forte et al., 2020). These have been termed as *"late response" fibroblasts (LR)* and "matrifibrocytes" (MFC) (**Figure 3**).

Vascular mural (mainly smooth muscle cells, SMCs) and perivascular cells (pericytes, PCs) are abundant and important cell types in the adult heart. They are in close relation with the blood vessel endothelium and play a key role in the regulation of vascular function (Armulik et al., 2011). SMCs have a contractile phenotype traditionally characterized by the expression of *Acta2*/αSMA, *Tagln* and *Cnn1* (Yoshida and Owens, 2005; Rensen et al., 2007), while PCs are identified by their *Cspg4*/NG2 and *Pdgfrb* expression, although they have also been shown to express *Acta2* in specific locations (Nehls and Drenckhahn, 1991; O'Farrel et al., 2017). The single-cell RNA profiling of healthy heart tissues revealed that mural cells share the expression of *Cspg4*, *Pdgfrb*, *Itga7*, *Mcam*/CD146 and *Rgs5* genes (Skelly et al., 2018; Muhl et al., 2020), and that PCs, but not SMCs, express *Vtn* (Skelly et al., 2018; Farbehi et al., 2019). SMCs show less heterogeneity as compared to CFs, but they can be discretely clustered too. These important findings might be relevant to fully understand the clonal nature of atherosclerotic lesions (Bennett et al., 2016; Chappell et al., 2016), but this point has not been extensively addressed as yet (Muhl et al., 2020).

IMMUNE CELLS

The heart comprises a small population of resident immune cells from the myeloid and lymphoid lineages (Hart and Fabre, 1981; Bönner et al., 2012; Epelman et al., 2014; Pinto et al., 2016). *Ptprc*/CD45 is a pan-leukocyte cell marker that is widely expressed in circulating/bone marrow-derived cells and has been traditionally used to also identify these cells in the heart (Nakano et al., 1990; Haudek et al., 2006; Ruiz-Villalba et al., 2015; Pinto et al., 2016). However, the characterization of bone-marrow-derived cell heterogeneity has proven to be too complex, as expected from cells deriving from multiple lineages, displaying wide functional versatility, and dynamically expressing molecular markers through time. In this section, we discuss the most relevant blood-borne cell types.

Cardiac Monocytes/Macrophages

Macrophages (MPs) are the epitome of cardiac immune cells. These cells can be isolated using classical markers, such as F4/80 or CD64 (Zaman et al., 2021). However, MPs can be "polarized" in response to pathologies including myocardial infarction, so that cardiac macrophage heterogeneity increases. After a cardiac injury, two classes of MPs are located in the myocardium: the classically activated M1 (Ly6C^{HIGH}/MRC1^{NEG}/CX3CR1^{LOW}), with a pro-inflammatory profile, and the alternatively activated M2 type, which consists of anti-inflammatory and reparative MPs (Ly6C^{LOW}/MRC1^{POS}/CX3CR1^{HIGH}) (Arnold et al., 2007; Nahrendorf et al., 2007). There is a clear trend to consider this classification as the oversimplification of a complex cellular maturation process, since unique and clear transcriptomic signatures for M1 and M2 are not evident. However, the interest in MP polarization has contributed positively to the study of the role of these cells in the adult heart (Frangogiannis, 2012; Pinto et al., 2014). During macrophage maturation, the monocyte chemoattractant protein-1 (MCP-1), and its receptor (CCR2), play an important role in the control of the process (Frangogiannis et al., 2007; Chen B. et al., 2019). Cardiac CCR2^{NEG} and CCR2^{POS} macrophages have distinct functional properties in both, murine and human hearts: CCR2^{NEG} are considered as reparative (and are thus close to the M1 classic MP phenotype), in contrast to CCR2^{POS}, which are considered as inflammatory MPs (Bajpai et al., 2018, 2019). Remarkably, a relation between cardiac macrophage heterogeneity and their embryonic origin has been pointed out: CCR2^{NEG}/MHC-II$^{POS(LOW/HIGH)}$ MPs derive from embryonic progenitors during the development, and CCR2^{POS}/MHC-II$^{POS(HIGH/LOW)}$ correspond to minor macrophage subpopulations derived from circulating monocytes and monocytes (Epelman et al., 2014; Lavine et al., 2014). It has been described that many resident CCR2^{NEG} MPs die soon after MI (Leuschner et al., 2012; Heidt et al., 2014), but thanks to scRNA-Seq studies, it is now clear that the population of CCR2^{NEG}MHC-II^{LOW}TIMD4^{POS}LYVE1^{POS} cardiac resident MPs is maintained after injury (Dick et al., 2019; Farbehi et al., 2019; **Figure 4**). Interestingly, a population of LYVE1^{POS} tissue-resident macrophages associated with cardiovascular remodeling has been recently described in human homeostatic hearts, although they were found to be TIMD4^{NEG} cells (Litviňuková et al., 2020).

Remarkably, scRNA-Seq studies have yielded results that are quite similar to a macrophage cellular map resulting from the use of a set of well-known MP surface markers (*Cx3cr1*, *H2-Aa/Ab*/MHC-II, *Ccr2*, *Mrc1*/CD206, *Fcgr1*/CD64, *Adgre1*/F4/80, *Cd68*, *Itgam*/CD11b) (King et al., 2017; Skelly et al., 2018; Dick et al., 2019; Farbehi et al., 2019). However, scRNA-Seq has identified three novel subpopulations of MPs by their unique transcriptomic profile: TIMD4^{POS}/LYVE-1^{POS}, MHC-II^{HIGH}/CCR2^{NEG}, and MHC-II^{HIGH}/CCR2^{POS} MPs (Zaman et al., 2021). Other authors have identified only two subsets of resident MPs using scRNA-Seq in combination with cell fate mapping approaches that studied samples from several healthy organs including the heart. One of the relevant findings of this study is that one of these macrophage populations (LYVE-1^{HIGH}/MHC-II^{LOW}) resides in the vicinity of blood vessels, while the other one (LYVE-1^{LOW}MHC-II^{HIGH}) is preferentially found close to nerves (Chakarov et al., 2019).

Finally, an additional subpopulation of mononuclear cells with an intermediate M1-M2 profile has been described. This subpopulation, called interferon-inducible cells (IFNICs), is characterized by the expression of F4-80^{POS}/MHC-II^{POS}/CCR2^{POS}/CXCL10^{POS}/Ly6C^{NEG}/CD11c^{NEG}. IFNICs seem to have a relevant role in the amplification of inflammation 4 days after MI via the Interferon regulatory factor 3 (IRF3)/type I interferons (IFNs) axis. In an attempt to summarize multiple

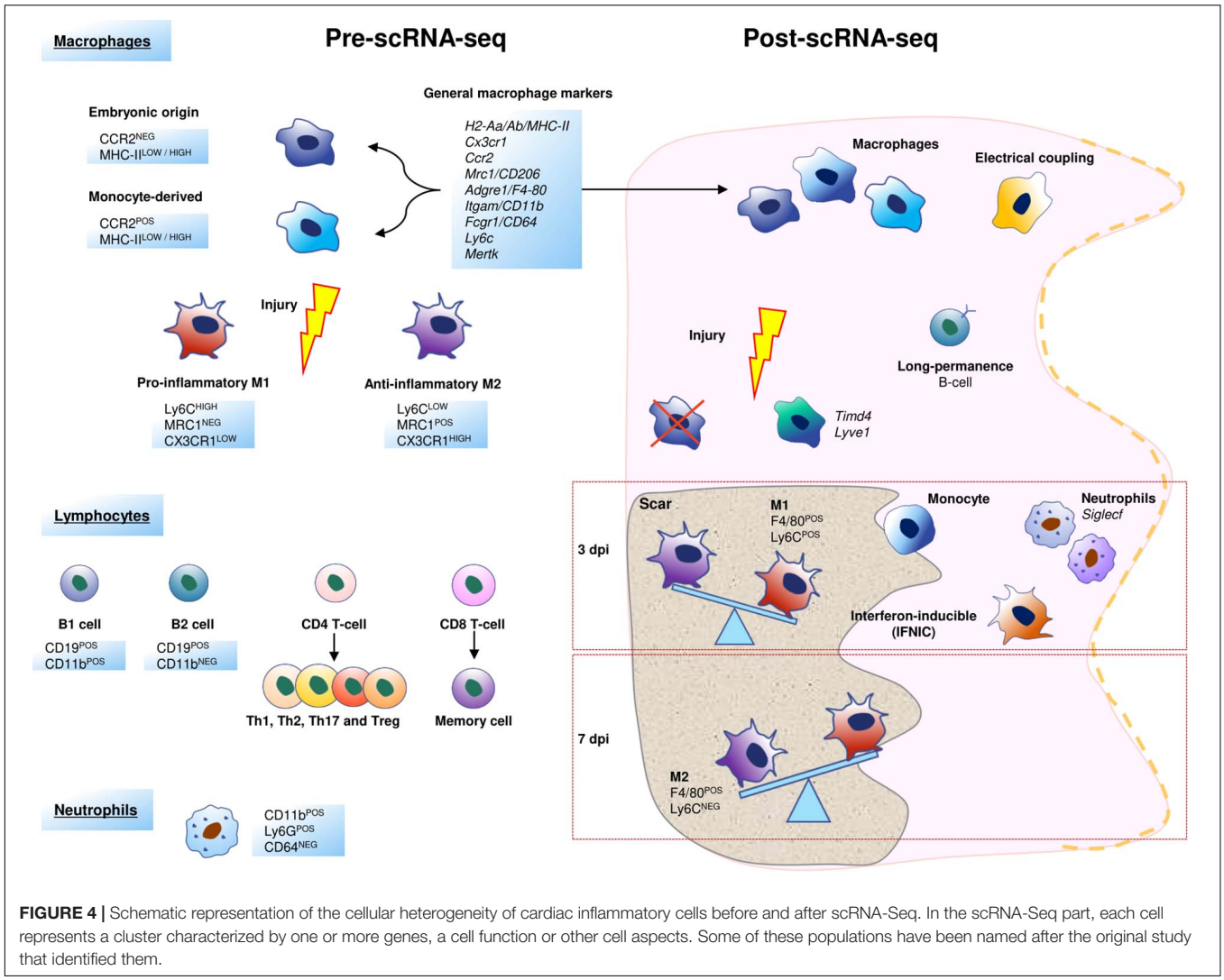

FIGURE 4 | Schematic representation of the cellular heterogeneity of cardiac inflammatory cells before and after scRNA-Seq. In the scRNA-Seq part, each cell represents a cluster characterized by one or more genes, a cell function or other cell aspects. Some of these populations have been named after the original study that identified them.

studies, we can conclude that only 7 macrophage subpopulations show different transcriptional profiles after the 1st week post-infarction when compared to MPs isolated from healthy hearts. In contrast, circulating MPs almost disappear from the heart around 28 days after MI, and the proportion of resident MPs is not recovered in comparison with the homeostatic heart (Dick et al., 2019). These data suggest a permanent modification in the cellular landscape because of the infarction damage. Moreover, all these findings illustrate the degree of complexity and subtlety of the maturation of MPs in the infarcted heart (King et al., 2017; **Figure 4**).

Neutrophils, Dendritic, and Mast Cells: Other Myeloid Cells Found in the Heart

Single-cell transcriptomics has also revealed an increase in the heterogeneity in other inflammatory cell populations in comparison with the pre scRNA-Seq landscape (Stubbington et al., 2017; Villani et al., 2017), including cardiac resident ones. This heterogeneity is dynamic and changes in different pathological settings (Farbehi et al., 2019). After a MI, neutrophils (NPs) are the first inflammatory cells recruited in the infarcted heart, and their numbers peak between 1 and 3 dpi (Ma et al., 2013). In a very recent scRNA-Seq study, six different clusters of NPs are shown to sequentially appear after MI (Vafadarnejad et al., 2020). The two predominant NP subpopulations at 3 dpi display enrichment in *Siglecf* expression, while between 3 and 5 dpi, *Siglecf*^{HIGH} neutrophils represent a distinctive state, exhibiting more phagocytosis and ROS-production than other neutrophils (**Figure 4**).

Dendritic cells (DCs) are another relevant cell type. It is currently assumed that there are three types of DCs: plasmacytoid or pre-DCs, derived from the lymphoid lineage, and two additional types of conventional DCs (cDCs), derived from the myeloid lineage. This classification is nonetheless controversial, since DCs are known to express different sets of surface markers depending on the organ in which they reside. In the heart, the general dendritic cell population expresses *Dpp4*/CD26 and *Zbtb46* (Guilliams et al., 2016; Clemente-Casares et al., 2017), cDCs type 1 have been recently

identified as CD103^{POS}/CD172α^{NEG}/CD11b^{NEG}, and cDCs type 2 as CD103^{NEG}/CD172α^{POS}/CD11b^{POS} (Van der Borght et al., 2017; Lee et al., 2018). Unfortunately, a plethora of bone-marrow-derived cells share some of these surface markers, limiting the unambiguous identification of these cells. In recent times, scRNA-Seq has helped to characterize CDCs in-depth (Dick et al., 2019; Farbehi et al., 2019; Martini et al., 2019). Whilst some authors propose beneficial functions of these cells regulating reparative actions from MPs and T-cells (Nagai et al., 2014; Choo et al., 2017), others have demonstrated a negative contribution of DCs to the infarct outcome (Lee et al., 2018).

Cardiac mast cells represent a minor cardiac cellular population characterized by the expression of *Fcer1g*/IgE, c-Kit/CD117, and CD9 (Sperr et al., 1994). Mast cells have been shown to increase their number and their secretion of the vasodilator and pro-fibrotic factors histamine and tryptase in failing human hearts (Patella et al., 1998). However, the role of mast cells in the mouse, the preferred animal model for the study and assessment of cells involved in the repair of the damaged heart, is not fully understood. After MI, mast cells have been reported to have a positive contribution to cardiac muscle functionality (Ngkelo et al., 2016), but have also been implicated in cardiac fibrosis (Legere et al., 2019). Thanks to a recent study using scRNA-Seq, it is now known that mast cells distributed through the epicardium and myocardium express high levels of *Mcp8* and pro-inflammatory cytokine genes like *Il6*. The level of the activation marker *Cd69* in a context of pressure overload led these authors to suggest an involvement of these cells in the early inflammatory response associated with cardiac hypertrophy (Martini et al., 2019).

Lymphoid Cells in the Heart

Lymphocytes, comprising classic B-cells, T-cells and the innate lymphoid cells (ILCs), such as NK cells, account for a small proportion of the immune cells found in the healthy cardiac interstitium (Pinto et al., 2016). B-lymphocytes are the most frequent leukocytes in the naïve murine heart, and they are known to play a role in the modulation of inflammation and cardiac remodeling after MI (Adamo et al., 2020a). B-cells can be divided into two major populations, B-1 and B-2. Although it is still unclear which one is their origin and lineage relationship, it is known that B-2 cells (CD19^{POS}CD11b^{NEG}) represent the majority of cardiac lymphocytes in the adult heart (Montecino-Rodriguez and Dorshkind, 2012). Other functions have been claimed for some of these cells. For example, it was recently reported that some circulating B-cells are prone to adhere to cardiac endothelial cells and take up long-term residence in the myocardium; scRNA-Seq analysis has revealed that these lymphocytes express a distinct gene expression profile as compared with other circulating counterparts (Adamo et al., 2020a; **Figure 4**).

T-lymphocytes include the naïve subtypes CD4 (helper; Th) and CD8 (cytotoxic; CTL), their differentiated states such as Th1, Th2, Th17 or Treg and memory CD8^{POS} cells, and the NK cells (Zhu et al., 2010; Casey et al., 2012). These T-lymphocyte subtypes play an important role in modulating inflammation and cardiac repair (Wang et al., 2019), but disagreement exists concerning their beneficial or detrimental contribution to heart recovery after damage (Hofmann and Frantz, 2015; Li J. et al., 2020). The use of single-cell transcriptomics has revealed the existence of previously unknown states (Stubbington et al., 2017; Villani et al., 2017), but the functions of these cells are not known as yet.

UNDERSTANDING scRNA-SEQ DATA FROM A BIOLOGICAL PERSPECTIVE

As discussed above, cardiomyocytes, endothelial, interstitial and immune cells display significant differences in gene expression. Still, it is important to note that several variables such as the genetic background, spatial location or cell-to-cell interaction may affect cell behavior and thus drive changes in the composition of cell populations (**Table 1**). Forte et al. (2020) have recently demonstrated that a mouse strain prone to hypertension (129S1/SvlmJ) has more susceptibility to cardiac rupture after MI than the inbred strain C57BL/6J, even if it is clear that 129S1/SvlmJ mice have more significant numbers of CFs in their ventricular walls. On the other hand, the specific cellular expression of ligands and receptors suggests the existence of molecular crosstalk patterns specific for atria or ventricles (Tucker et al., 2020; Wang et al., 2020). Moreover, the combination of single-cell and spatial transcriptomics is a powerful additional tool for the elucidation of such cellular interactions. For instance, two specific clusters of human arterial ECs and SMCs previously identified by scRNA-Seq were shown to co-localize in space and directly interact through JAG1 and NOTCH2 (Litviňuková et al., 2020). This analytical strategy based on receptor-ligand interactions can be highly relevant for the identification of paracrine signaling mechanisms existing among specific cell types, and may even suggest novel roles for CFs (Skelly et al., 2018; Litviňuková et al., 2020). The sex of individuals may also have an impact in cardiac cellular composition and susceptibility to disease (**Table 1**). Female mice present more CFs, T-cells and fewer granulocytes as compared with males, and these differences have gonadal hormone support (Squiers et al., 2020). Additionally, sex seems to have an impact not only on the abundance of the clusters mentioned above, but also on the gene expression profile of different cell types (McLellan et al., 2020). These differences, again, seem to be related to some degree to hormone response. Since CFs are the cell type with the highest expression of sex hormone receptors in the heart, it is very likely that these cells are the key to the articulation of such differential responses (McLellan et al., 2020). MPs have been found to differentially express genes associated with inflammation in males and genes with anti-inflammatory effects in females (Skelly et al., 2018). Of note, similar differences have been observed in human hearts. CMs and CFs were the cell types with more differentially expressed genes between the sexes (Tucker et al., 2020). In addition, ventricles from women have been found to have higher numbers of CMs than those of men (Litviňuková et al., 2020). During cardiac repair after infarct, female mice have been shown to have lower ventricular

rupture rates, a higher influx of reparative leukocytes, and different regulation of immune mediators (Pullen et al., 2020). All these findings highlight the need for the evaluation of cell dynamics in both sexes, as this might have a positive impact on pharmacological interventions.

UNDERSTANDING scRNA-SEQ DATA FROM A DATA-PERSPECTIVE

As previously shown, single-cell technologies are expanding our understanding of biological systems, and cardiac research certainly benefits from this revolution (Abplanalp et al., 2020; Ruiz-Villalba et al., 2020; Zhou and Zhang, 2020). In addition to the opportunities that this technology has made available to the scientific community, scRNA-Seq has also presented some challenges related to both experimental approaches and data-analysis routines. We discuss these aspects below.

All these technologies began with the dissociation of the tissues of interest in individual cells, a step that has been detailed in several experimental protocols in order to guarantee the reproducibility of the results. This is especially relevant in the cardiovascular field, where differences derived from the sample preparation procedure have been previously reported (Pinto et al., 2016; Zhou and Wang, 2020). After the sample preparation, cells need to be isolated, and to do this there are multiple options (Lafzi et al., 2018). Two relevant examples are the C1 microfluidic platform from Fluidigm, which isolates single cells into individual reaction chambers (Wu et al., 2014), and the Chromium platform from 10XGenomics, which uses droplets to capture single cells (Zheng et al., 2017). In both cases, cells are lysated, the RNA is reversely transcribed to complementary DNA, amplified, and processed to build up sequencing-ready libraries. There are no definitive conclusions on which protocol is the best, but criteria for selection are: "*the number of cells profiled per sample*," "*the sequencing depth required*," and "*tag-based – which only provides 3′ or 5′ reads – vs. full transcript sequencing*," among others. Importantly, methods are continuously under development to improve their scale, accuracy and sensitivity, with marked differences between them (Shiroguchi et al., 2012; Brennecke et al., 2013; Grün et al., 2014; Islam et al., 2014; Ziegenhain et al., 2017). Additional considerations are the possible effects of storage (e.g., frozen samples), tissue dissociation and "single-cell vs. single-nucleus" biases (Denisenko et al., 2020). A general characteristic of the single-cell experimental protocols aimed to limit the impact of PCR-based mRNA amplification is the use of unique molecular identifies (UMI). UMIs are used to tag the fragments of mRNA (Kivioja et al., 2012; Islam et al., 2014); as a result, during the bioinformatic analysis, it is possible to characterize both mRNA and UMI signals and, therefore filter duplicate reads (identified as pairs with same mRNA sequence and same UMI).

Regarding the data-analysis part, many of the lessons learned in the analysis of RNA-Seq data (Conesa et al., 2016) need to be reviewed in the context of single-cell data (Stegle et al., 2015; Vallejos et al., 2017). Interestingly, specific aspects of bulk RNA-Seq analysis that can be imported into the single-cell RNA-Seq analysis (Soneson and Robinson, 2018). In this review we go over the different data analysis steps, with the aim of providing a broad overview on the topic and key references for the interested reader. Generally, the single-cell analysis field is advancing so rapidly that many of the state-of-the-art references and tools are published as not-peer reviewed preprints.

The very first step in the analysis of scRNA-Seq data is to generate count matrices (e.g., gene as rows and cell as columns), where every matrix cell contains the total number of mRNA reads or UMIs for a given gene and a given biological cell. The second step, *to identify what part of the data should be used for the analysis*, is (even) more crucial. It is thus necessary to *adequately* filter those genes and cells that do not provide sufficient signal or data quality. An initial filter is aimed at discarding genes with low UMIs and cells expressing a small number of genes (Ilicic et al., 2016; Soneson and Robinson, 2018). Additional filters are required, but they require us to take the nature and physiological state of the cells into consideration. For instance, cells with high levels of reads in mitochondrial genes may be dead or dying cells (Ilicic et al., 2016), but also cells at a defined metabolic state (Denisenko et al., 2020). Interestingly, it is frequent to apply "data-set or even sample-specific thresholds" to the previous filtering criteria (Luecken et al., 2020), as no standard criterion for this filtering exists (Soneson and Robinson, 2018). A third data filter is intended to identify doublets, i.e., two or more cells sharing the same cell-identifying barcode (McGinnis et al., 2019; Wolock et al., 2019; Bernstein et al., 2020; DePasquale et al., 2020). Importantly, filtering scRNA-Seq data often requires several rounds of analysis, as doublet identification requires data preprocessing and, as a result of the filtering, the data-analysis process may require to be started again.

The second step is the pre-processing of the count matrix. This step cannot be separated from the following objectives of the analysis, namely cell-subtype identification (Trapnell, 2015), differential gene expression (Soneson and Robinson, 2018), marker identification (Soneson and Robinson, 2018), or visualization (Cakir et al., 2020) among others (Stegle et al., 2015). It is not within the scope of this review to provide a comprehensive overview of all these methods, but we use the Seurat package to picture the associated steps and challenges in the analysis of scRNA-Seq data. Seurat was developed as an unified framework for scRNA-Seq analysis (Satija et al., 2015) and it has evolved to include multiple data-sets (Butler et al., 2018) and multiple data-types (Hao et al., 2020) (e.g., scATAC-Seq). In the first version, Seurat v1, the data matrix (*UMI counts per gene per cell*) was normalized in a cell-based manner as the "*number of unique UMIs per 200,000 unique UMIs*" and the data was log-transformed for the downstream analysis, which included identification of highly variable genes, Principal Component Analysis and data imputation (Satija et al., 2015). In Seurat v2 (Butler et al., 2018), the normalization was modified to "*number of unique UMIs per 10,000 unique UMIs*" and to apply natural log-transformation. Furthermore, in Seurat v2, a strategy was developed that uses Canonical Correlation Analysis (CCA) to identify the most highly correlated features of the data sets in order to align several batches

TABLE 1 | Summary of single-cell RNA-Seq analyses performed in the adult mammalian heart.

Heart sample	Cell target	Strain	Sex	Condition	Injury stage	Platform	Cell/Nuclei sequenced number	Single-cell or nuclei	References
Complete	Cardiac cells	C57BL/6NRj, Fzt:DU	Male	Normal	–	10X Genomics	11,672	Nuclei	Galow et al., 2020
Ventricles	Cardiac cells	C57BL/6J	Not specified	Normal and ischemia reperfusion	3 dpi	SORT-seq	932	Cell	Gladka et al., 2018
Ventricles	Cardiac cells	Mixed C57BL/6J	Mixed	Neonatal, Normal and MI (LAD)	3, 7, and 14 dpi	SORT-seq	1,939	Cell	Kretzschmar et al., 2018
Atria and ventricle	Cardiac cells	Human	Female and male	Normal	–	10X Genomics	45,870/363,213	Cell and nuclei	Litviňuková et al., 2020
Complete	Cardiac cells	Human	Mixed	Normal, HF and recovery	–	ICELL8	12,266 (normal)/5,933 (HF)	Cell	Wang et al., 2020
Ventricles	Cardiac cells	C57BL/6J	Mixed	Normal and hypertension AngII-induced	14 dpi	10X Genomics	29,558	Cell and nuclei	McLellan et al., 2020
Not specified	Cardiac cells	C57BL/6J	Mixed	Normal	–	10X Genomics	>4,000	Cell	Schaum et al., 2018
Atria and ventricle	Cardiac cells	Human	Mixed	MI	–	10X Genomics	287,269	Nuclei	Tucker et al., 2020
Complete	Cardiac cells	Fzt:DU	Male	Normal	–	10X Genomics	8,635	Nuclei	Wolfien et al., 2020a
LV	CMs	C57BL/6J/Human	Male	TAC/Normal and dilated cardiomyopathy	8 weeks post-TAC/End stage	Fluidigm C1	359/116	Nuclei	See et al., 2017
Ventricles	CMs	C57BL/6J	Male	Normal and TAC	8 weeks post-infarction	ICELL8	586	Cell	Yekelchyk et al., 2019
Complete	ECs	C57BL/6J	Male	Normal	–	10X Genomics	4,612	Cell	Kalucka et al., 2020
Ventricles	ECs	Mixed C57BL/6J	Mixed	Normal and MI (LAD)	7 dpi	10X Genomics	3,200-4,000	Cell	Li et al., 2019
LV	Interstitial cells	129S4/SvJaeSor	Male	Normal and MI (LAD)	3 and 7 dpi	10X Genomics	>30,000	Cell	Farbehi et al., 2019
Ventricles	Interstitial cells	C57BL/6J and 129S1/SvlmJ	Male	Normal and MI (LAD)	1, 3, 5, 7, 14, and 28 dpi	10X Genomics	36,847	Cell	Forte et al., 2020
Ventricles	Interstitial cells	C57BL/6J	Female and male	Normal	–	10X Genomics	12,000	Cell	Skelly et al., 2018
Not specified	CFs and mural cells	C57BL/6J	Male	Normal	–	Smart-Seq2	< 6,158	Cell	Muhl et al., 2020
Ventricles	CFs	C57BL/6J	Not specified	Normal and MI (LAD)	7, 14, and 30 dpi	10X Genomics	32,669	Cell	Ruiz-Villalba et al., 2020
Complete	Immune cells	C57BL/6J	Female	Normal	–	10X Genomics	17,500	Cell	Adamo et al., 2020b
Complete	Immune cells	C57BL/6J	Not specified	Normal and MI	11 dpi	10X Genomics	1,780 (normal)/6,503 (MI)	Cell	Dick et al., 2019
Complete	Immune cells	C57BL/6J	Male	Normal and heart failure (TAC)	1- and 4-weeks post-infarction	10X Genomics	>17,853	Cell	Martini et al., 2019

LV, left ventricle; CMs, cardiomyocyte; ECs, endothelial cell; CFs, cardiac fibroblast; MI, myocardial infarction; TAC, transverse aortic constriction; LAD, left anterior descending artery; HF, heart failure. " – ": means no data; "dpi" means days post-infarction; "complete" indicates whole heart analysis; "cardiac cells" means the main heart cell types; "interstitial cells" do not include cardiomyocytes and "atria and ventricle" indicate separated analysis.

(Butler et al., 2018). Seurat v3 (Stuart et al., 2019) replaced the CCA-based integrative strategy in order to include the concept of "*anchors*," pairs of cells that can be paired across data-sets. Importantly, this anchoring strategy allows for the integration among data modalities, specifically scRNA-Seq and scATAC-Seq. Variance-stabilizing transformation to take into account the mean-variance relationship was also included in this software (Hafemeister and Satija, 2019). Seurat v4 includes a methodology for leveraging over the paired nature of multi-omic single-cell data (Hao et al., 2020). As can be concluded from this brief snapshot-based review of the Seurat package, the bioinformatic analysis pipelines are in continuous evolution, and require frequent review. Furthermore, the data pre-processing also has an impact on the differential gene expression analysis as the two steps cannot be disentangled (Soneson and Robinson, 2018).

A third challenge in the analysis of scRNA-Seq data is the identification of the different cell populations. To this end, clustering methodologies are available, as well as several systematic reviews on these procedures which might be helpful for the interpretation of the data (Duò et al., 2020; Peyvandipour et al., 2020). Once cell groups have been identified, the next step is to *label* such clusters. To do so, differential gene expression between cell groups is conducted in order to identify markers, and those markers are used to label cell groups (Zeisel et al., 2015). Manual annotation is possible, but time-consuming. Automatic annotation is the obvious alternative, but it has limitations, and its use remains an open challenge (Abdelaal et al., 2019). Fortunately, for specific tissues such as blood, the annotation strategies are maturing, and tools such as *Azimuth*[1] (from the Satija lab) allow for the automatic annotation of peripheral blood mononuclear cells (PBMCs). Importantly, although a first draft of the cardiac cellular landscape has been described (Litviňuková et al., 2020), the use of automatic annotation in this context is still an open challenge. Another challenge is the identification of known cell subtypes within a given cell population (for instance, to identify Th1 $CD4^{POS}$ T-cells within a $CD4^{POS}$ T-cell population). In some cases, the data obtained may not have enough resolution (or enough number of cells) to discriminate between sub-types. A second additional challenge is the robust validation of rare cells types that can be identified because no quality control can provide sufficient evidence for these cells to be discarded (Wegmann et al., 2019).

Single-cell RNA-Seq has been frequently used for the characterization of cell differentiation processes. In this type of experiment, the samples are obtained at predetermined time-points. However, by making use of the variability of "*speeds*" of differentiation in cells, it is possible to order these cells so that such ordering can define a *pseudotime* (Trapnell et al., 2014). There are, again, many available methodologies to achieve this goal (Saelens et al., 2019) and they can be used for the characterization of complex phenomena such as hematopoiesis (Athanasiadis et al., 2017) or human heart development (Cui et al., 2019). Importantly, *pseudotime* methods are also limited in their ability to accurately order cells. Several methods have been developed as a complementary approach to estimate the dynamics of every cell by using the comparison between unspliced and spliced mRNA signal to estimate a vector of differentiation (RNA velocity) (La Manno et al., 2018). This method, known as *velocyto*, has completely changed the analysis of single-cell dynamics (e.g., SIB 2019 Bioinformatics Award), so that updated approaches for such a strategy such as *scVelo* (Bergen et al., 2020) and very recently *CellRank* (Lange et al., 2020) have appeared. At this point, we believe it is important to highlight that all these methods are powerful exploratory tools supporting model-generation approaches. Nevertheless, we cannot forget that all these tools are based on specific (mathematical) assumptions. Therefore, any new insights into biological entities or events derived from this kind of analysis will require a wet-lab validation and a careful biological interpretation (Everaert et al., 2017; Zhou and Wang, 2020).

A significant aspect of single-cell RNA-Seq data is the large number of data-points (cells) available for every sample. When every sample may contain up to 10 000 cells, and every cell contains the profile of between 2000 and 5000 genes, then "studies with many samples" enter a data-rich environment rapidly. Such an environment is ideally suited for using Machine Learning methodologies, and specific attention has been given to Deep Learning methods (Eraslan et al., 2019a). These methodologies cover a wide range of applications such as *in silico* data generation (e.g., cscGAN) (Marouf et al., 2020), data imputation (e.g., scIGANSs) (Xu et al., 2020) based on Generative Adversarial Networks (Goodfellow et al., 2014), data integration on unpaired datasets (e.g., totalVI) (Gayoso et al., 2021) or paired datasets (e.g., LIBRA) (Martinez-De-Morentin et al., 2021), among others. In summary, many of the methodologies associated with single-cell RNA-Seq analysis are using Machine Learning tools and new applications are appearing to refine many steps of the analysis framework described (Oller-Moreno et al., 2021).

Finally, while we have reviewed the most frequently used tools to analyze single-cell RNA-Seq data and discussed their limitations as well as the technical challenges that still need to be addressed in their use in a biological context, there are additional methods that may require attention (Poirion et al., 2016). Among those we would highlight: (i) data-analysis tools designed for non-bioinformaticians (Franzén and Björkegren, 2020); (ii) visualization tools for single-cell RNA-Seq (Cakir et al., 2020); (iii) identification of Gene-Regulatory Networks (Chen and Mar, 2018; van Dijk et al., 2018); (iv) new *gene expression imputation* methods, aiming at evaluating the increased sparsity observed in scRNA-Seq data (Wagner et al., 2017; Huang et al., 2018; van Dijk et al., 2018; Eraslan et al., 2019b); (v) integration of multiple (and possible massive) data-sets (Luecken et al., 2020); and (vi) the implementation of multi-omic (e.g., scRNA-Seq, Baek and Lee, 2020) and scATAC-Seq analysis (Chen H. et al., 2019; Baek and Lee, 2020). It is worth mentioning that single-cell multi-omic analysis also benefits from machine learning techniques; for instance "Latent Semantic Indexing" and "Latent Dirichlet Allocation" – both used in natural language processing – are implemented in Signac (Stuart et al., 2020) and cisTopic (Bravo González-Blas et al., 2019) scATAC-Seq packages respectively.

[1] https://satijalab.org/azimuth/

FUTURE PERSPECTIVES: TOWARD DATA INTEGRATION

While scRNA-Seq has allowed for an unprecedented level of detail and characterization of heart cellular components, it is becoming evident that scRNA-Seq alone cannot fully disclose cardiac cell complexity (Lähnemann et al., 2020) as scRNA-Seq provides a *just-a-transcriptomical view* and additional regulatory layers are necessary to understand the system (Gomez-Cabrero et al., 2014, 2019). Fortunately, single-cell profiling is now available for chromatin accessibility (Buenrostro et al., 2015), proteomics (Cheung et al., 2020), DNA methylation (Galvão and Kelsey, 2021) or chromatin conformation assays (Ramani et al., 2020). However, the analysis of data-sets from these different analyses requires specific developments for each one of them, as well as an increased effort to provide a multi-omic integrative analysis (Jansen et al., 2019; Argelaguet et al., 2020). As the multi-omics approaches are still under development, protocols are being improved continuously, allowing for the profiling of more-than-one-omic analysis of the same cell (Chen S. et al., 2019). It is anyway clear we will have to compromise, and accept that a certain decrease in the quality of the data obtained from these approaches can be balanced with the advantages provided by the analysis of paired profiles (Lee et al., 2020).

The profiling of a cell-population is a relevant issue because it allows us to understand how cells interact and organize in space and time. Therefore, scRNA-Seq approaches can be complemented by the use of other techniques. For instance, *Spatial Transcriptomics* (ST) allows for the 2D characterization of tissue transcriptomics (Asp et al., 2019). ST and scRNA-Seq are indeed complementary, because scRNA-Seq identifies the cell-types and their markers, while ST contributes to illustrate their spatial organization within the tissue (Andersson et al., 2020). Importantly, cell-to-cell interactions can be computationally predicted by combining ligand and receptor information, their expression in each cell-type, and the available information on protein-protein interactions (Efremova et al., 2020; Hou et al., 2020). The ultimate challenge in the use of these techniques is to bring all this knowledge on cell characterization into a clinical setting (Haque et al., 2017; Keener, 2019). Therefore, it is not surprising that the cross-referencing between scRNA-Seq and other high throughput technologies, such as proteomics or metabolomics, is regarded as crucial for determining and prioritizing the molecular candidates associated with prevalent cardiac complex conditions such as heart failure (Chan et al., 2020).

CONCLUSION

Despite all the relevant discoveries made using the scRNA-Seq technology, much more is needed to understand the intrinsic complexity of cell communities. We believe it is especially important to characterize in detail the temporal dimension of cell differentiation or incorporation into tissues. The paradigmatic case to illustrate this point would be that of the circulating (blood-borne) cells recruited to tissues after damage. In any case, and regardless of the biological system we choose to study via scRNA-Seq, we should always consider the conceptual limitation of the "cell marker" concept, the plasticity of molecular cell phenotypes, and all the caveats of a technology that heavily depends on bioinformatics and mathematical routines for the analysis of the data it yields. In the cardiovascular research field, scRNA-Seq has been instrumental to progress in the understanding of cardiac progenitor cell dynamics, the characterization of specific subpopulations of poorly studied cardiac cell types, and the cardiac events in which they participate. Future refinements of this knowledge are likely to derive from the improvement of protocols for cell extraction, isolation, purification, and the transcriptomic analysis itself, together with the continuous development of bioinformatic analytical tools.

AUTHOR CONTRIBUTIONS

EM-S and AR-V contributed to the conception, designed the study, and wrote the first draft of the manuscript. XM, JP-P, and DG-C wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported by funds from the Spanish Ministry of Science, Innovation and Universities (RTI2018-095410-B-I00); University of Málaga (UMA18-FEDERJA-146), and Carlos III Institute of Health (RD16/0011/0030). EM-S was supported by funds from the Spanish Ministry of Science, Innovation and Universities FPU fellowship (FPU18/05219). AR-V was supported by funds from University of Málaga (*Incorporación de doctores* from the *I Plan Propio de Incorporación de Doctores, 2020*).

REFERENCES

Abdelaal, T., Michielsen, L., Cats, D., Hoogduin, D., Mei, H., Reinders, M. J. T., et al. (2019). A comparison of automatic cell identification methods for single-cell RNA sequencing data. *Genome Biol.* 20:194. doi: 10.1186/s13059-019-1795-z

Abplanalp, W. T., John, D., Cremer, S., Assmus, B., Dorsheimer, L., Hoffmann, J., et al. (2020). Single-cell RNA-sequencing reveals profound changes in circulating immune cells in patients with heart failure. *Cardiovasc. Res.* 117, 484–494. doi: 10.1093/cvr/cvaa101

Adamo, L., Rocha-Resende, C., and Mann, D. L. (2020a). The Emerging Role of B Lymphocytes in Cardiovascular Disease. *Annu. Rev. Immunol.* 38, 99–121. doi: 10.1146/annurev-immunol-042617-053104

Adamo, L., Rocha-Resende, C., Lin, C. Y., Evans, S., Williams, J., Dun, H., et al. (2020b). Myocardial B cells are a subset of circulating lymphocytes with delayed transit through the heart. *JCI Insight* 5:e134700. doi: 10.1172/jci.insight.134700

Aird, W. C. (2007a). Phenotypic heterogeneity of the endothelium: I. Structure, function, and mechanisms. *Circul. Res.* 100, 158–173. doi: 10.1161/01.RES.0000255691.76142.4a

Aird, W. C. (2007b). Phenotypic heterogeneity of the endothelium: II. Representative vascular beds. *Circul. Res.* 100, 174–190. doi: 10.1161/01.RES.0000255690.03436.ae

Aizarani, N., Saviano, A., Sagar, Mailly, L., Durand, S., Herman, J. S., et al. (2019). A human liver cell atlas reveals heterogeneity and epithelial progenitors. *Nature* 572, 199–204. doi: 10.1038/s41586-019-1373-2

Anderson, R. H., and Ho, S. Y. (1998). The Architecture of the Sinus Node, the Atrioventricular Conduction Axis, and the Internodal Atrial Myocardium. *J. Cardiovas. Electrophysiol.* 9, 1233–1248. doi: 10.1111/j.1540-8167.1998.tb00097.x

Andersson, A., Bergenstråhle, J., Asp, M., Bergenstråhle, L., Jurek, A., Fernández Navarro, J., et al. (2020). Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography. *Commun. Biol.* 3:565. doi: 10.1038/s42003-020-01247-y

Anversa, P., Olivetti, G., Melissari, M., and Loud, A. V. (1980). Stereological measurement of cellular and subcellular hypertrophy and hyperplasia in the papillary muscle of adult rat. *J. Mole. Cell. Cardiol.* 12, 781–795. doi: 10.1016/0022-2828(80)90080-2

Argelaguet, R., Arnol, D., Bredikhin, D., Deloro, Y., Velten, B., Marioni, J. C., et al. (2020). MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol.* 21:111. doi: 10.1186/s13059-020-02015-1

Armulik, A., Genové, G., and Betsholtz, C. (2011). Pericytes: Developmental, Physiological, and Pathological Perspectives, Problems, and Promises. *Devel. Cell* 21, 193–215. doi: 10.1016/j.devcel.2011.07.001

Arnold, L., Henry, A., Poron, F., Baba-Amer, Y., Van Rooijen, N., Plonquet, A., et al. (2007). Inflammatory monocytes recruited after skeletal muscle injury switch into antiinflammatory macrophages to support myogenesis. *J. Exp. Med.* 204, 1057–1069. doi: 10.1084/jem.20070075

Asp, M., Giacomello, S., Larsson, L., Wu, C., Fürth, D., Qian, X., et al. (2019). A Spatiotemporal Organ-Wide Gene Expression and Cell Atlas of the Developing Human Heart. *Cell* 179, 1647–1660.e19. doi: 10.1016/j.cell.2019.11.025

Athanasiadis, E. I., Botthof, J. G., Andres, H., Ferreira, L., Lio, P., and Cvejic, A. (2017). Single-cell RNA-sequencing uncovers transcriptional states and fate decisions in haematopoiesis. *Nat. Commun.* 8:2045. doi: 10.1038/s41467-017-02305-6

Baek, S., and Lee, I. (2020). Single-cell ATAC sequencing analysis: From data preprocessing to hypothesis generation. *Comput. Struct. Biotechnol. J.* 18, 1429–1439. doi: 10.1016/j.csbj.2020.06.012

Bajpai, G., Bredemeyer, A., Li, W., Zaitsev, K., Koenig, A. L., Lokshina, I., et al. (2019). Tissue Resident CCR2- and CCR2+ Cardiac Macrophages Differentially Orchestrate Monocyte Recruitment and Fate Specification Following Myocardial Injury. *Circulat. Res.* 124, 263–278. doi: 10.1161/CIRCRESAHA.118.314028

Bajpai, G., Schneider, C., Wong, N., Bredemeyer, A., Hulsmans, M., Nahrendorf, M., et al. (2018). The human heart contains distinct macrophage subsets with divergent origins and functions. *Nat. Med.* 24, 1234–1245. doi: 10.1038/s41591-018-0059-x

Banerjee, I., Fuseler, J. W., Price, R. L., Borg, T. K., and Baudino, T. A. (2007). Determination of cell types and numbers during cardiac development in the neonatal and adult rat and mouse. *Am. J. Physiol. Heart Circul. Physiol.* 293, 1883–1891. doi: 10.1152/ajpheart.00514.2007

Banerjee, S., Dhara, S. K., and Bacanamwo, M. (2012). Endoglin is a novel endothelial cell specification gene. *Stem Cell Res.* 8, 85–96. doi: 10.1016/j.scr.2011.08.006

Bennett, M. R., Sinha, S., and Owens, G. K. (2016). Vascular Smooth Muscle Cells in Atherosclerosis. *Circul. Res.* 118, 692–702. doi: 10.1161/CIRCRESAHA.115.306361

Bergen, V., Lange, M., Peidli, S., Wolf, F. A., and Theis, F. J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* 38, 1408–1414. doi: 10.1038/s41587-020-0591-3

Bergmann, O., Zdunek, S., Felker, A., Salehpour, M., Alkass, K., Bernard, S., et al. (2015). Dynamics of Cell Generation and Turnover in the Human Heart. *Cell* 161, 1566–1575. doi: 10.1016/j.cell.2015.05.026

Bernstein, N. J., Fong, N. L., Lam, I., Roy, M. A., Hendrickson, D. G., and Kelley, D. R. (2020). Solo: Doublet Identification in Single-Cell RNA-Seq via Semi-Supervised Deep Learning. *Cell Systems* 11, 95–101.e5. doi: 10.1016/j.cels.2020.05.010

Bönner, F., Borg, N., Burghoff, S., and Schrader, J. (2012). Resident cardiac immune cells and expression of the ectonucleotidase enzymes CD39 and CD73 after ischemic injury. *PLoS One* 7:e34730. doi: 10.1371/journal.pone.0034730

Bravo González-Blas, C., Minnoye, L., Papasokrati, D., Aibar, S., Hulselmans, G., Christiaens, V., et al. (2019). cisTopic: cis-regulatory topic modeling on single-cell ATAC-seq data. *Nat. Methods* 16, 397–400. doi: 10.1038/s41592-019-0367-1

Brennecke, P., Anders, S., Kim, J. K., Kołodziejczyk, A. A., Zhang, X., Proserpio, V., et al. (2013). Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* 10, 1093–1098. doi: 10.1038/nmeth.2645

Brutsaert, D. L. (2003). Cardiac endothelial-myocardial signaling: Its role in cardiac growth, contractile performance, and rhythmicity. *Physiol. Rev.* 83, 59–115. doi: 10.1152/physrev.00017.2002

Buckingham, M., Meilhac, S., and Zaffran, S. (2005). Building the mammalian heart from two sources of myocardial cells. *Nat. Rev. Genet.* 6, 826–835. doi: 10.1038/nrg1710

Buenrostro, J. D., Wu, B., Litzenburger, U. M., Ruff, D., Gonzales, M. L., Snyder, M. P., et al. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 523, 486–490. doi: 10.1038/nature14590

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 36, 411–420. doi: 10.1038/nbt.4096

Cakir, B., Prete, M., Huang, N., van Dongen, S., Pir, P., and Kiselev, V. Y. (2020). Comparison of visualization tools for single-cell RNAseq data. *NAR Genom. Bioinform.* 2:lqaa052. doi: 10.1093/nargab/lqaa052

Camelliti, P., Borg, T. K., and Kohl, P. (2005). Structural and functional characterisation of cardiac fibroblasts. *Cardiov. Res.* 65, 40–51. doi: 10.1016/j.cardiores.2004.08.020

Cano, E., Carmona, R., Ruiz-Villalba, A., Rojas, A., Chau, Y. Y., Wagner, K. D., et al. (2016). Extracardiac septum transversum/proepicardial endothelial cells pattern embryonic coronary arterio-venous connections. *Proc. Natl. Acad. Sci. U. S. A.* 113, 656–661. doi: 10.1073/pnas.1509834113

Casey, K. A., Fraser, K. A., Schenkel, J. M., Moran, A., Abt, M. C., Beura, L. K., et al. (2012). Antigen-Independent Differentiation and Maintenance of Effector-like Resident Memory T Cells in Tissues. *J. Immunol.* 188, 4866–4875. doi: 10.4049/jimmunol.1200402

Castro-Quezada, A., Nadal-Ginard, B., and De La Cruz, M. V. (1972). Experimental study of the formation of the bulboventricular loop in the chick. *J. Embryol. Exp. Morphol.* 27, 623–637.

Chakarov, S., Lim, H. Y., Tan, L., Lim, S. Y., See, P., Lum, J., et al. (2019). Two distinct interstitial macrophage populations coexist across tissues in specific subtissular niches. *Science* 363:eaau0964. doi: 10.1126/science.aau0964

Chan, M. Y., Efthymios, M., Tan, S. H., Pickering, J. W., Troughton, R., Pemberton, C., et al. (2020). Prioritizing Candidates of Post–Myocardial Infarction Heart Failure Using Plasma Proteomics and Single-Cell Transcriptomics. *Circulation* 142, 1408–1421. doi: 10.1161/CIRCULATIONAHA.119.045158

Chappell, J., Harman, J. L., Narasimhan, V. M., Yu, H., Foote, K., Simons, B. D., et al. (2016). Extensive Proliferation of a Subset of Differentiated, yet Plastic, Medial Vascular Smooth Muscle Cells Contributes to Neointimal Formation in Mouse Injury and Atherosclerosis Models. *Circul. Res.* 119, 1313–1323. doi: 10.1161/CIRCRESAHA.116.309799

Chen, B., Brickshawana, A., and Frangogiannis, N. G. (2019). The Functional Heterogeneity of Resident Cardiac Macrophages in Myocardial Injury: CCR2+ Cells Promote Inflammation, whereas CCR2- Cells Protect. *Circul. Res.* 124, 183–185. doi: 10.1161/CIRCRESAHA.118.314357

Chen, H., Lareau, C., Andreani, T., Vinyard, M. E., Garcia, S. P., Clement, K., et al. (2019). Assessment of computational methods for the analysis of single-cell ATAC-seq data. *Genome Biol.* 20:241. doi: 10.1186/s13059-019-1854-5

Chen, S., and Mar, J. C. (2018). Evaluating methods of inferring gene regulatory networks highlights their lack of performance for single cell gene expression data. *BMC Bioinform.* 19:232. doi: 10.1186/s12859-018-2217-z

Chen, S., Lake, B. B., and Zhang, K. (2019). High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nat. Biotechnol.* 37, 1452–1457. doi: 10.1038/s41587-019-0290-0

Cheung, T. K., Lee, C.-Y., Bayer, F. P., McCoy, A., Kuster, B., and Rose, C. M. (2020). Defining the carrier proteome limit for single-cell proteomics. *Nat. Methods* 18, 76–83. doi: 10.1038/s41592-020-01002-5

Chi, J. T., Chang, H. Y., Haraldsen, G., Jahnsen, F. L., Troyanskaya, O. G., Chang, D. S., et al. (2003). Endothelial cell diversity revealed by global expression profiling. *Proc. Natl. Acad. Sci. U. S. A.* 100, 10623–10628. doi: 10.1073/pnas.1434429100

Choo, E. H., Lee, J. H., Park, E. H., Park, H. E., Jung, N. C., Kim, T. H., et al. (2017). Infarcted Myocardium-Primed Dendritic Cells Improve Remodeling and Cardiac Function after Myocardial Infarction by Modulating the Regulatory T Cell and Macrophage Polarization. *Circulation* 135, 1444–1457. doi: 10.1161/CIRCULATIONAHA.116.023106

Clemente-Casares, X., Hosseinzadeh, S., Barbu, I., Dick, S. A., Macklin, J. A., Wang, Y., et al. (2017). A CD103+ Conventional Dendritic Cell Surveillance System Prevents Development of Overt Heart Failure during Subclinical Viral Myocarditis. *Immunity* 47, 974–989.e8. doi: 10.1016/j.immuni.2017.10.011

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., et al. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biol.* 17, 1–19. doi: 10.1186/s13059-016-0881-8

Coppiello, G., Collantes, M., Sirerol-Piquer, M. S., Vandenwijngaert, S., Schoors, S., Swinnen, M., et al. (2015). Meox2/Tcf15 heterodimers program the heart capillary endothelium for cardiac fatty acid uptake. *Circulation* 131, 815–826. doi: 10.1161/CIRCULATIONAHA.114.013721

Cui, Y., Zheng, Y., Liu, X., Yan, L., Fan, X., Yong, J., et al. (2019). Single-Cell Transcriptome Analysis Maps the Developmental Track of the Human Heart. *Cell Rep.* 26, 1934–1950.e5. doi: 10.1016/j.celrep.2019.01.079

Denisenko, E., Guo, B. B., Jones, M., Hou, R., de Kock, L., Lassmann, T., et al. (2020). Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. *Genome Biol.* 21:130. doi: 10.1186/s13059-020-02048-6

DePasquale, E. A. K., Schnell, D., Chetal, K., and Salomonis, N. (2020). Protocol for Identification and Removal of Doublets with DoubletDecon. *STAR Protocols* 1:100085. doi: 10.1016/j.xpro.2020.100085

Dick, S. A., Macklin, J. A., Nejat, S., Momen, A., Clemente-Casares, X., Althagafi, M. G., et al. (2019). Self-renewing resident cardiac macrophages limit adverse remodeling following myocardial infarction. *Nat. Immunol.* 20, 29–39. doi: 10.1038/s41590-018-0272-2

Didié, M., Christalla, P., Rubart, M., Muppala, V., Döker, S., Unsöld, B., et al. (2013). Parthenogenetic stem cells for tissue-engineered heart repair. *J. Clin. Investig.* 123, 1285–1298. doi: 10.1172/JCI66854

Dunwoodie, S. L. (2007). Combinatorial signaling in the heart orchestrates cardiac induction, lineage specification and chamber formation. *Sem. Cell Devel. Biol.* 18, 54–66. doi: 10.1016/j.semcdb.2006.12.003

Duò, A., Robinson, M. D., and Soneson, C. (2020). A systematic performance evaluation of clustering methods for single-cell RNA-seq data. *F1000Research* 7:1141. doi: 10.12688/f1000research.15666.3

Efremova, M., Vento-Tormo, M., Teichmann, S. A., and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Prot.* 15, 1484–1506. doi: 10.1038/s41596-020-0292-x

Engleka, K. A., Manderfield, L. J., Brust, R. D., Li, L., Cohen, A., Dymecki, S. M., et al. (2012). Islet1 Derivatives in the Heart Are of Both Neural Crest and Second Heart Field Origin. *Circulat. Res.* 110, 922–926. doi: 10.1161/CIRCRESAHA.112.266510

Epelman, S., Lavine, K. J., Beaudin, A. E., Sojka, D. K., Carrero, J. A., Calderon, B., et al. (2014). Embryonic and adult-derived resident cardiac macrophages are maintained through distinct mechanisms at steady state and during inflammation. *Immunity* 40, 91–104. doi: 10.1016/j.immuni.2013.11.019

Eraslan, G., Avsec, Ž, Gagneur, J., and Theis, F. J. (2019a). Deep learning: new computational modelling techniques for genomics. *Nat. Rev. Genet.* 20, 389–403. doi: 10.1038/s41576-019-0122-6

Eraslan, G., Simon, L. M., Mircea, M., Mueller, N. S., and Theis, F. J. (2019b). Single-cell RNA-seq denoising using a deep count autoencoder. *Nat. Commun.* 10:390. doi: 10.1038/s41467-018-07931-2

Everaert, C., Luypaert, M., Maag, J. L. V., Cheng, Q. X., Dinger, M. E., Hellemans, J., et al. (2017). Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data. *Scientific Rep.* 7:1559. doi: 10.1038/s41598-017-01617-3

Fadel, B. M., Boutet, S. C., and Quertermous, T. (1999). Octamer-dependent in Vivo Expression of the Endothelial Cell-specific TIE2 Gene. *J. Biol. Chem.* 274, 20376–20383. doi: 10.1074/jbc.274.29.20376

Farbehi, N., Patrick, R., Dorison, A., Xaymardan, M., Janbandhu, V., Wystub-Lis, K., et al. (2019). Single-cell expression profiling reveals dynamic flux of cardiac stromal, vascular and immune cells in health and injury. *ELife* 8, 1–39. doi: 10.7554/eLife.43882

Feng, W., Chen, L., Nguyen, P. K., Wu, S. M., and Li, G. (2019). Single Cell Analysis of Endothelial Cells Identified Organ-Specific Molecular Signatures and Heart-Specific Cell Populations and Molecular Features. *Front. Cardiovasc. Med.* 6:165. doi: 10.3389/fcvm.2019.00165

Fidler, A. L., Darris, C. E., Chetyrkin, S. V., Pedchenko, V. K., Boudko, S. P., Brown, K. L., et al. (2017). Collagen IV and basement membrane at the evolutionary dawn of metazoan tissues. *ELife* 6, 1–24. doi: 10.7554/elife.24176

Forte, E., Skelly, D. A., Chen, M., Daigle, S., Morelli, K. A., Hon, O., et al. (2020). Dynamic Interstitial Cell Response during Myocardial Infarction Predicts Resilience to Rupture in Genetically Diverse Mice. *Cell Rep.* 30, 3149–3163.e6. doi: 10.1016/j.celrep.2020.02.008

Frangogiannis, N. G. (2012). Regulation of the inflammatory response in cardiac repair. *Circulat. Res.* 110, 159–173. doi: 10.1161/CIRCRESAHA.111.243162

Frangogiannis, N. G. (2020). Fact and fiction about fibroblast to endothelium conversion: Semantics and substance of cellular identity. *Circulation* 142, 1663–1666. doi: 10.1161/CIRCULATIONAHA.120.050875

Frangogiannis, N. G., Dewald, O., Xia, Y., Ren, G., Haudek, S., Leucker, T., et al. (2007). Critical role of monocyte chemoattractant protein-1/CC chemokine ligand 2 in the pathogenesis of ischemic cardiomyopathy. *Circulation* 115, 584–592. doi: 10.1161/CIRCULATIONAHA.106.646091

Franzén, O., and Björkegren, J. L. M. (2020). alona: a web server for single-cell RNA-seq analysis. *Bioinformatics* 36, 3910–3912. doi: 10.1093/bioinformatics/btaa269

Fu, X., Khalil, H., Kanisicak, O., Boyer, J. G., Vagnozzi, R. J., Maliken, B. D., et al. (2018). Specialized fibroblast differentiated states underlie scar formation in the infarcted mouse heart. *J. Clin. Investig.* 128, 2127–2143. doi: 10.1172/JCI98215

Galow, A. M., Wolfien, M., Müller, P., Bartsch, M., Brunner, R. M., Hoeflich, A., et al. (2020). Integrative Cluster Analysis of Whole Hearts Reveals Proliferative Cardiomyocytes in Adult Mice. *Cells* 9, 1–16. doi: 10.3390/cells9051144

Galvão, A., and Kelsey, G. (2021). Profiling Genome-Wide in Single Cells. *Method. Mole. Biol.* 2214, 221–240. doi: 10.1007/978-1-0716-0958-3_15

Garlanda, C., and Dejana, E. (1997). Heterogeneity of Endothelial Cells. *Arterioscl. Thromb. Vasc. Biol.* 17, 1193–1202. doi: 10.1161/01.ATV.17.7.1193

Gayoso, A., Steier, Z., Lopez, R., Regier, J., Nazor, K. L., Streets, A., et al. (2021). Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat. Method.* 18, 272–282. doi: 10.1038/s41592-020-01050-x

Gladka, M. M., Molenaar, B., De Ruiter, H., Van Der Elst, S., Tsui, H., Versteeg, D., et al. (2018). Single-Cell Sequencing of the Healthy and Diseased Heart Reveals Cytoskeleton-Associated Protein 4 as a New Modulator of Fibroblasts Activation. *Circulation* 138, 166–180. doi: 10.1161/CIRCULATIONAHA.117.030742

Gomez-Cabrero, D., Abugessaisa, I., Maier, D., Teschendorff, A., Merkenschlager, M., Gisel, A., et al. (2014). Data integration in the era of omics: current and future challenges. *BMC Syst. Biol.* 8:I1. doi: 10.1186/1752-0509-8-S2-I1

Gomez-Cabrero, D., Tarazona, S., Ferreirós-Vidal, I., Ramirez, R. N., Company, C., and Schmidt, A. (2019). STATegra, a comprehensive multi-omics dataset of B-cell differentiation in mouse. *Scientific Data* 6:256. doi: 10.1038/s41597-019-0202-7

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). *Generative Adversarial Networks*. URL: http://arxiv.org/abs/1406.2661 (accessed December 19, 2020).

Goodyer, W. R., Beyersdorf, B. M., Paik, D. T., Tian, L., Li, G., Buikema, J. W., et al. (2019). Transcriptomic profiling of the developing cardiac conduction system at single-cell resolution. *Circul. Res.* 125, 379–397. doi: 10.1161/CIRCRESAHA.118.314578

Grün, D., Kester, L., and Van Oudenaarden, A. (2014). Validation of noise models for single-cell transcriptomics. *Nat. Methods* 11, 637–640. doi: 10.1038/nmeth.2930

Guilliams, M., Dutertre, C. A., Scott, C. L., McGovern, N., Sichien, D., and Chakarov, S. (2016). Unsupervised High-Dimensional Analysis Aligns Dendritic Cells across Tissues and Species. *Immunity* 45, 669–684. doi: 10.1016/j.immuni.2016.08.015

Haber, A. L., Biton, M., Rogel, N., Herbst, R. H., Shekhar, K., Smillie, C., et al. (2017). A single-cell survey of the small intestinal epithelium. *Nature* 551, 333–339. doi: 10.1038/nature24489

Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* 20:296. doi: 10.1186/s13059-019-1874-1

Hamburger, V., and Hamilton, H. L. (1951). A series of normal stages in the development of the chick embryo. *J. Morphol.* 88, 49–92. doi: 10.1002/jmor.1050880104

Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., and Lai, S. (2018). Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell* 172, 1091–1107.e17. doi: 10.1016/j.cell.2018.02.001

Hao, Y., Hao, S., Andersen-nissen, E., Zheng, S., and Lee, M. J. (2020). Integrated analysis of multimodal single-cell data. *bioRxiv* [preprint]. doi: 10.1101/2020.10.12.335331

Haque, A., Engel, J., Teichmann, S. A., and Lönnberg, T. (2017). A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med.* 9:75. doi: 10.1186/s13073-017-0467-4

Harding, T. C., Geddes, B. J., Murphy, D., Knight, D., and Uney, J. B. (1998). Switching transgene expression in the brain using an adenoviral tetracycline-regulatable system. *Nat. Biotechnol.* 16, 553–555. doi: 10.1038/nbt0698-553

Hart, D. N. J., and Fabre, J. W. (1981). Demonstration and characterization of ia-positive dendritic cells in the interstitial connective tissues of rat heart and other tissues, but not brain. *J. Exp. Med.* 154, 347–361. doi: 10.1084/jem.154.2.347

Haudek, S. B., Xia, Y., Huebener, P., Lee, J. M., Carlson, S., Crawford, J. R., et al. (2006). Bone marrow-derived fibroblast precursors mediate ischemic cardiomyopathy in mice. *Proc. Natl. Acad. Sci. U. S. A.* 103, 18284–18289. doi: 10.1073/pnas.0608799103

He, L., Huang, X., Kanisicak, O., Li, Y., Wang, Y., Li, Y., et al. (2017). Preexisting endothelial cells mediate cardiac neovascularization after injury. *J. Clin. Investig.* 127, 2968–2981. doi: 10.1172/JCI93868

Heidt, T., Courties, G., Dutta, P., Sager, H. B., Sebas, M., Iwamoto, Y., et al. (2014). Differential contribution of monocytes to heart macrophages in steady-state and after myocardial infarction. *Circulat. Res.* 115, 284–295. doi: 10.1161/CIRCRESAHA.115.303567

Hofmann, U., and Frantz, S. (2015). Role of lymphocytes in myocardial injury, healing, and remodeling after myocardial infarction. *Circulat. Res.* 116, 354–367. doi: 10.1161/CIRCRESAHA.116.304072

Hoogaars, W. M. H., Tessari, A., Moorman, A. F. M., De Boer, P. A. J., Hagoort, J., Soufan, A. T., et al. (2004). The transcriptional repressor Tbx3 delineates the developing central conduction system of the heart. *Cardiovasc. Res.* 62, 489–499. doi: 10.1016/j.cardiores.2004.01.030

Hou, R., Denisenko, E., Ong, H. T., Ramilowski, J. A., and Forrest, A. R. R. (2020). Predicting cell-to-cell communication networks using NATMI. *Nat. Commun.* 11:5011. doi: 10.1038/s41467-020-18873-z

Howe, R. S., Isaacson, K. J., Albert, J. L., and Coutifaris, C. B. (1991). Embryonic heart rate in human pregnancy. *J. Ultras. Med.* 10, 367–371. doi: 10.7863/jum.1991.10.7.367

Hu, P., Liu, J., Zhao, J., Wilkins, B. J., Lupino, K., Wu, H., et al. (2018). Single-nucleus transcriptomic survey of cell diversity and functional maturation in postnatal mammalian hearts. *Genes Devel.* 32, 1344–1357. doi: 10.1101/gad.316802.118

Huang, M., Wang, J., Torre, E., Dueck, H., Shaffer, S., Bonasio, R., et al. (2018). SAVER: gene expression recovery for single-cell RNA sequencing. *Nat. Methods* 15, 539–542. doi: 10.1038/s41592-018-0033-z

Hulsmans, M., Clauss, S., Xiao, L., Aguirre, A. D., King, K. R., Hanley, A., et al. (2017). Macrophages Facilitate Electrical Conduction in the Heart. *Cell* 169, 510–522.e20. doi: 10.1016/j.cell.2017.03.050

Huynh, T., Chen, L., Terrell, P., and Baldini, A. (2007). A fate map of Tbx1 expressing cells reveals heterogeneity in the second cardiac field. *Genesis* 45, 470–475. doi: 10.1002/dvg.20317

Hynes, R. O. (2012). The evolution of metazoan extracellular matrix. *J. Cell Biol.* 196, 671–679. doi: 10.1083/jcb.201109041

Ilicic, T., Kim, J. K., Kolodziejczyk, A. A., Bagger, F. O., McCarthy, D. J., Marioni, J. C., et al. (2016). Classification of low quality cells from single-cell RNA-seq data. *Genome Biol.* 17, 1–15. doi: 10.1186/s13059-016-0888-1

Imdahl, F., and Saliba, A.-E. (2020). Advances and challenges in single-cell RNA-seq of microbial communities. *Curr. Opin. Microbiol.* 57, 102–110. doi: 10.1016/j.mib.2020.10.001

Islam, S., Zeisel, A., Joost, S., La Manno, G., Zajac, P., Kasper, M., et al. (2014). Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat. Method* 11, 163–166. doi: 10.1038/nmeth.2772

Jakab, M., and Augustin, H. G. (2020). Understanding angiodiversity: insights from single cell biology. *Development* 147:dev146621. doi: 10.1242/dev.146621

Jambusaria, A., Hong, Z., Zhang, L., Srivastava, S., Jana, A., Toth, P. T., et al. (2020). Endothelial heterogeneity across distinct vascular beds during homeostasis and inflammation. *ELife* 9, 1–32. doi: 10.7554/eLife.51413

Jansen, C., Ramirez, R. N., El-Ali, N. C., Gomez-Cabrero, D., Tegner, J., Merkenschlager, M., et al. (2019). Building gene regulatory networks from scATAC-seq and scRNA-Seq using Linked Self Organizing Maps. *PLoS Comput. Biol.* 15:e1006555. doi: 10.1371/journal.pcbi.1006555

Jia, G., Preussner, J., Chen, X., Guenther, S., Yuan, X., Yekelchyk, M., et al. (2018). Single cell RNA-seq and ATAC-seq analysis of cardiac progenitor cell transition states and lineage settlement. *Nat. Commun.* 9:4877. doi: 10.1038/s41467-018-07307-6

Kalucka, J., de Rooij, L. P. M. H., Goveia, J., Rohlenova, K., Dumas, S. J., and Meta, E. (2020). Single-Cell Transcriptome Atlas of Murine Endothelial Cells. *Cell* 180, 764–779.e20. doi: 10.1016/j.cell.2020.01.015

Kanisicak, O., Khalil, H., Ivey, M. J., Karch, J., Maliken, B. D., Correll, R. N., et al. (2016). Genetic lineage tracing defines myofibroblast origin and function in the injured heart. *Nat. Commun.* 7, 1–14. doi: 10.1038/ncomms12260

Keener, A. B. (2019). Single-cell sequencing edges into clinical trials. *Nat. Med.* 25, 1322–1326. doi: 10.1038/d41591-019-00017-6

Kelly, R. G., Buckingham, M. E., and Moorman, A. F. (2014). Heart Fields and Cardiac Morphogenesis. *Cold Spring Harbor. Perspect. Med.* 4, a015750. doi: 10.1101/cshperspect.a015750

Kimura, W., Xiao, F., Canseco, D. C., Muralidhar, S., Thet, S., Zhang, H. M., et al. (2015). Hypoxia fate mapping identifies cycling cardiomyocytes in the adult heart. *Nature* 523, 226–230. doi: 10.1038/nature14582

King, K. R., Aguirre, A. D., Ye, Y. X., Sun, Y., Roh, J. D., Ng, R. P., et al. (2017). IRF3 and type i interferons fuel a fatal response to myocardial infarction. *Nat. Med.* 23, 1481–1487. doi: 10.1038/nm.4428

Kivioja, T., Vähärautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S., et al. (2012). Counting absolute numbers of molecules using unique molecular identifiers. *Nat. Methods* 9, 72–74. doi: 10.1038/nmeth.1778

Kretzschmar, K., Post, Y., Bannier-Hélaouët, M., Mattiotti, A., Drost, J., Basak, O., et al. (2018). Profiling proliferative cells and their progeny in damaged murine hearts. *Proc. Natl. Acad. Sci. U. S. Am.* 115:E12245. doi: 10.1073/pnas.1805829115

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., et al. (2018). RNA velocity of single cells. *Nature* 560, 494–498. doi: 10.1038/s41586-018-0414-6

Lafzi, A., Moutinho, C., Picelli, S., and Heyn, H. (2018). Tutorial: guidelines for the experimental design of single-cell RNA sequencing studies. *Nat. Protocols* 13, 2742–2757. doi: 10.1038/s41596-018-0073-y

Lähnemann, D., Köster, J., Szczurek, E., McCarthy, D. J., Hicks, S. C., and Robinson, M. D. (2020). Eleven grand challenges in single-cell data science. *Genome Biol.* 21:31. doi: 10.1186/s13059-020-1926-6

Lange, M., Bergen, V., Klein, M., Setty, M., Reuter, B., Bakhti, M., et al. (2020). CellRank for directed single-cell fate mapping. *BioRxiv* [preprint]. doi: 10.1101/2020.10.19.345983 2020.10.19.345983,

Lavine, K. J., Epelman, S., Uchida, K., Weber, K. J., Nichols, C. G., Schilling, J. D., et al. (2014). Distinct macrophage lineages contribute to disparate patterns of cardiac recovery and remodeling in the neonatal and adult heart. *Proc. Natl. Acad. Sci. U. S. A.* 111, 16029–16034. doi: 10.1073/pnas.1406508111

Lee, J. S., Jeong, S.-J., Kim, S., Chalifour, L., Yun, T. J., Miah, M. A., et al. (2018). Conventional Dendritic Cells Impair Recovery after Myocardial Infarction. *J. Immunol.* 201, 1784–1798. doi: 10.4049/jimmunol.1800322

Lee, J., Hyeon, D. Y., and Hwang, D. (2020). Single-cell multiomics: technologies and data analysis methods. *Exp. Mole. Med.* 52, 1428–1442. doi: 10.1038/s12276-020-0420-2

Legere, S. A., Haidl, I. D., Légaré, J. F., and Marshall, J. S. (2019). Mast cells in cardiac fibrosis: New insights suggest opportunities for intervention. *Front. Immunol.* 10:580. doi: 10.3389/fimmu.2019.00580

Lescroart, F., Mohun, T., Meilhac, S. M., Bennett, M., and Buckingham, M. (2012). Lineage Tree for the Venous Pole of the Heart. *Circulat. Res.* 111, 1313–1322. doi: 10.1161/CIRCRESAHA.112.271064

Lescroart, F., Wang, X., Lin, X., Swedlund, B., Gargouri, S., Sànchez-Dànes, A., et al. (2018). Defining the earliest step of cardiovascular lineage segregation by single-cell RNA-seq. *Science* 359, 1177–1181. doi: 10.1126/science.aao4174

Leuschner, F., Rauch, P. J., Ueno, T., Gorbatov, R., Marinelli, B., Lee, W. W., et al. (2012). Rapid monocyte kinetics in acute myocardial infarction are sustained by extramedullary monocytopoiesis. *J. Exp. Med.* 209, 123–137. doi: 10.1084/jem.20111009

Li, G., Tian, L., Goodyer, W., Kort, E. J., Buikema, J. W., Xu, A., et al. (2020). Correction: Single cell expression analysis reveals anatomical and cell cycle-dependent transcriptional shifts during heart development. *Development* 147:dev190819. doi: 10.1242/dev.190819

Li, J., Liang, C., Yang, K. Y., Huang, X., Han, M. Y., Li, X., et al. (2020). Specific ablation of CD4+ T-cells promotes heart regeneration in juvenile mice. *Theranostics* 10, 8018–8035. doi: 10.7150/thno.42943

Li, Z., Solomonidis, E. G., Meloni, M., Taylor, R. S., Duffin, R., Dobie, R., et al. (2019). Single-cell transcriptome analyses reveal novel targets modulating cardiac neovascularization by resident endothelial cells following myocardial infarction. *Eur. Heart J.* 40, 2507–2520. doi: 10.1093/eurheartj/ehz305

Liang, X., Wang, G., Lin, L., Lowe, J., Zhang, Q., Bu, L., et al. (2013). HCN4 dynamically marks the first heart field and conduction system precursors. *Circulat. Res.* 113, 399–407. doi: 10.1161/CIRCRESAHA.113.301588

Litviňuková, M., Talavera-López, C., Maatz, H., Reichart, D., Worth, C. L., and Lindberg, E. L. (2020). Cells of the adult human heart. *Nature* 588, 466–472. doi: 10.1038/s41586-020-2797-4

Lother, A., Bergemann, S., Deng, L., Moser, M., Bode, C., and Hein, L. (2018). Cardiac endothelial cell transcriptome. *Arterioscl. Thromb. Vasc. Biol.* 38, 566–574. doi: 10.1161/ATVBAHA.117.310549

Luecken, M., Büttner, M., Chaichoompu, K., Danese, A., Interlandi, M., Mueller, M., et al. (2020). Benchmarking atlas-level data integration in single-cell genomics. *bioRxiv* [preprint]1–61. doi: 10.1101/2020.05.22.111161

Lukowski, S. W., Patel, J., Andersen, S. B., Sim, S. L., Wong, H. Y., Tay, J., et al. (2019). Single-Cell Transcriptional Profiling of Aortic Endothelium Identifies a Hierarchy from Endovascular Progenitors to Differentiated Cells. *Cell Rep.* 27, 2748–2758.e3. doi: 10.1016/j.celrep.2019.04.102

Ma, Y., Yabluchanskiy, A., and Lindsey, M. L. (2013). Neutrophil roles in left ventricular remodeling following myocardial infarction. *Fibrogen. Tissue Rep.* 6:1. doi: 10.1186/1755-1536-6-11

Mäkinen, T., Veikkola, T., Mustjoki, S., Karpanen, T., Catimel, B., Nice, E. C., et al. (2001). Isolated lymphatic endothelial cells transduce growth, survival and migratory signals via the VEGF-C/D receptor VEGFR-3. *EMBO J.* 20, 4762–4773. doi: 10.1093/emboj/20.17.4762

Manavski, Y., Lucas, T., Glaser, S. F., Dorsheimer, L., Günther, S., Braun, T., et al. (2018). Clonal expansion of endothelial cells contributes to ischemia-induced neovascularization. *Circulat. Res.* 122, 670–677. doi: 10.1161/CIRCRESAHA.117.312310

Marouf, M., Machart, P., Bansal, V., Kilian, C., Magruder, D. S., Krebs, C. F., et al. (2020). Realistic in silico generation and augmentation of single-cell RNA-seq data using generative adversarial networks. *Nat. Commun.* 11:166. doi: 10.1038/s41467-019-14018-z

Martinez-De-Morentin, X., Khan, S. A., Lehmann, R., Tegner, J., and Gomez-Cabrero, D. (2021). Machine Translation between paired Single Cell Multi Omics Data. *BioRxiv* [preprint]. doi: 10.1101/2021.01.27.428400

Martini, E., Kunderfranco, P., Peano, C., Carullo, P., Cremonesi, M., Schorn, T., et al. (2019). Single-Cell Sequencing of Mouse Heart Immune Infiltrate in pressure Overload-Driven Heart Failure Reveals Extent of Immune Activation. *Circulation* 140, 2089–2107. doi: 10.1161/CIRCULATIONAHA.119.041694

Massaia, A., Chaves, P., Samari, S., Miragaia, R. J., Meyer, K., Teichmann, S. A., et al. (2018). Single Cell Gene Expression to Understand the Dynamic Architecture of the Heart. *Front. Cardiovasc. Med.* 5:167. doi: 10.3389/fcvm.2018.00167

McGinnis, C. S., Murrow, L. M., and Gartner, Z. J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* 8, 329–337.e4. doi: 10.1016/j.cels.2019.03.003

McLellan, M. A., Skelly, D. A., Dona, M. S. I., Squiers, G. T., Farrugia, G. E., Gaynor, T. L., et al. (2020). High-Resolution Transcriptomic Profiling of the Heart During Chronic Stress Reveals Cellular Drivers of Cardiac Fibrosis and Hypertrophy. *Circulation* 142, 1448–1463. doi: 10.1161/CIRCULATIONAHA.119.045115

Meilhac, S. M., and Buckingham, M. E. (2018). The deployment of cell lineages that form the mammalian heart. *Nat. Rev. Cardiol.* 15, 705–724. doi: 10.1038/s41569-018-0086-9

Mereu, E., Lafzi, A., Moutinho, C., Ziegenhain, C., McCarthy, D. J., and Álvarez-Varela, A. (2020). Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. *Nat. Biotechnol.* 38, 747–755. doi: 10.1038/s41587-020-0469-4

Mikawa, T., and Fischman, D. A. (1992). Retroviral analysis of cardiac morphogenesis: discontinuous formation of coronary vessels. *Proc. Natl. Acad. Sci. U. S. A.* 89, 9504–9508. doi: 10.1073/pnas.89.20.9504

Mikawa, T., Borisov, A., Brown, A. M. C., and Fischman, D. A. (1992). Clonal analysis of cardiac morphogenesis in the chicken embryo using a replication-defective retrovirus: I. Formation of the ventricular myocardium. *Devel. Dynam.* 193, 11–23. doi: 10.1002/aja.1001930104

Montecino-Rodriguez, E., and Dorshkind, K. (2012). B-1 B Cell Development in the Fetus and Adult. *Immunity* 36, 13–21. doi: 10.1016/j.immuni.2011.11.017

Moore-Morris, T., Cattaneo, P., Guimarães-Camboa, N., Bogomolovas, J., Cedenilla, M., Banerjee, I., et al. (2018). Infarct fibroblasts do not derive from bone marrow lineages. *Circul. Res.* 122, 583–590. doi: 10.1161/CIRCRESAHA.117.311490

Moore-Morris, T., Guimarães-Camboa, N., Banerjee, I., Zambon, A. C., Kisseleva, T., Velayoudon, A., et al. (2014). Resident fibroblast lineages mediate pressure overload-induced cardiac fibrosis. *J. Clin. Investig.* 124, 2921–2934. doi: 10.1172/JCI74783

Moses, K. A., DeMayo, F., Braun, R. M., Reecy, J. L., and Schwartz, R. J. (2001). Embryonic expression of anNkx2-5/Cre gene usingROSA26 reporter mice. *Genesis* 31, 176–180. doi: 10.1002/gene.10022

Mouton, A. J., Ma, Y., Rivera Gonzalez, O. J., Daseke, M. J., Flynn, E. R., Freeman, T. C., et al. (2019). Fibroblast polarization over the myocardial infarction time continuum shifts roles from inflammation to angiogenesis. *Basic Res. Cardiol.* 114, 1–16. doi: 10.1007/s00395-019-0715-4

Muhl, L., Genové, G., Leptidis, S., Liu, J., He, L., and Mocci, G. (2020). Single-cell analysis uncovers fibroblast heterogeneity and criteria for fibroblast and mural cell identification and discrimination. *Nat. Commun.* 11:740. doi: 10.1038/s41467-020-17740-1

Müller, F., and O'Rahilly, R. (1987). The development of the human brain, the closure of the caudal neuropore, and the beginning of secondary neurulation at stage 12. *Anat. Embryol.* 176, 413–430. doi: 10.1007/BF00310083

Nag, A. (1980). Study of non-muscle cells of the adult mammalian heart: a fine structural analysis and distribution. *Cytobios* 28, 41–61.

Nagai, T., Honda, S., Sugano, Y., Matsuyama, T. A., Ohta-Ogo, K., Asaumi, Y., et al. (2014). Decreased myocardial dendritic cells is associated with impaired reparative fibrosis and development of cardiac rupture after myocardial infarction in humans. *J. Am. Heart Assoc.* 3, 1–8. doi: 10.1161/JAHA.114.000839

Nahrendorf, M., Swirski, F. K., Aikawa, E., Stangenberg, L., Wurdinger, T., Figueiredo, J. L., et al. (2007). The healing myocardium sequentially mobilizes two monocyte subsets with divergent and complementary functions. *J. Exp. Med.* 204, 3037–3047. doi: 10.1084/jem.20070885

Nakano, A., Harada, T., Morikawa, S., and Kato, Y. (1990). Expression of Leukocyte Common Antigen (CD45) on Various Human Leukemia/Lymphoma Cell Lines. *Pathol. Int.* 40, 107–115. doi: 10.1111/j.1440-1827.1990.tb01549.x

Nehls, V., and Drenckhahn, D. (1991). Heterogeneity of microvascular pericytes for smooth muscle type alpha-actin. *J. Cell Biol.* 113, 147–154. doi: 10.1083/jcb.113.1.147

Ng, S. Y., Wong, C. K., and Tsang, S. Y. (2010). Differential gene expressions in atrial and ventricular myocytes: Insights into the road of applying embryonic

stem cell-derived cardiomyocytes for future therapies. *Am. J. Physiol. Cell Physiol.* 299, 1234–1249. doi: 10.1152/ajpcell.00402.2009

Ngkelo, A., Richart, A., Kirk, J. A., Bonnin, P., Vilar, J., Lemitre, M., et al. (2016). Mast cells regulate myofilament calcium sensitization and heart function after myocardial infarction. *J. Exp. Med.* 213, 1353–1374. doi: 10.1084/jem.20160081

Nolan, D. J., Ginsberg, M., Israely, E., Palikuqi, B., Poulos, M. G., and James, D. (2013). Molecular Signatures of Tissue-Specific Microvascular Endothelial Cell Heterogeneity in Organ Maintenance and Regeneration. *Devel. Cell* 26, 204–219. doi: 10.1016/j.devcel.2013.06.017

Nomura, S., Satoh, M., Fujita, T., Higo, T., Sumida, T., Ko, T., et al. (2018). Cardiomyocyte gene programs encoding morphological and functional signatures in cardiac hypertrophy and failure. *Nat. Commun.* 9, 1–17. doi: 10.1038/s41467-018-06639-7

O'Farrel, F. M., Mastitskaya, S., Hammond-Haley, M., Freitas, F., Wah, W. R., and Attwell, D. (2017). Capillary pericytes mediate coronary no-reflow after myocardial ischaemia. *ELife* 6, 1–16. doi: 10.7554/eLife.29280

Oller-Moreno, S., Kloiber, K., Machart, P., and Bonn, S. (2021). Algorithmic advances in Machine Learning for single cell expression analysis. *Curr. Opin. Syst. Biol.* 25, 27–33. doi: 10.1016/j.coisb.2021.02.002

Pallante, B. A., Giovannone, S., Fang-Yu, L., Zhang, J., Liu, N., Kang, G., et al. (2010). Contactin-2 expression in the cardiac Purkinje fiber network. *Circul. Arrhyth. Electrophysiol.* 3, 186–194. doi: 10.1161/CIRCEP.109.928820

Palmquist-Gomes, P., Guadix, J. A., and Pérez-Pomares, J. M. (2018). Avian embryonic coronary arterio-venous patterning involves the contribution of different endothelial and endocardial cell populations. *Devel. Dynam.* 247, 686–698. doi: 10.1002/dvdy.24610

Patella, V., Marinò, I., Arbustini, E., Lamparter-Schummert, B., Versa, L., Adt, M., et al. (1998). Stem cell factor in mast cells and increased mast cell density in idiopathic and ischemic cardiomyopathy. *Circulation* 97, 971–978. doi: 10.1161/01.CIR.97.10.971

Patwardhan, V., Fernandez, S., Montgomery, M., and Litvin, J. (2000). The rostro-caudal position of cardiac myocytes affect their fate. *Devel. Dynam.* 218, 123–135. doi: 10.1002/(SICI)1097-0177(200005)218:1<123::AID-DVDY11<3.0.CO;2-6

Peyvandipour, A., Shafi, A., Saberian, N., and Draghici, S. (2020). Identification of cell types from single cell data using stable clustering. *Scient. Rep.* 10:12349. doi: 10.1038/s41598-020-66848-3

Pinto, A. R., Godwin, J. W., and Rosenthal, N. A. (2014). Macrophages in cardiac homeostasis, injury responses and progenitor cell mobilisation. *Stem. Cell Res.* 13, 705–714. doi: 10.1016/j.scr.2014.06.004

Pinto, A. R., Ilinykh, A., Ivey, M. J., Kuwabara, J. T., D'antoni, M. L., Debuque, R., et al. (2016). Revisiting cardiac cellular composition. *Circulat. Res.* 118, 400–409. doi: 10.1161/CIRCRESAHA.115.307778

Pogontke, C., Guadix, J. A., Ruiz-Villalba, A., and Pérez-Pomares, J. M. (2019). Development of the Myocardial Interstitium. *Anatom. Record* 302, 58–68. doi: 10.1002/ar.23915

Poirion, O. B., Zhu, X., Ching, T., and Garmire, L. (2016). Single-Cell Transcriptomics Bioinformatics and Computational Challenges. *Front. Genet.* 7:163. doi: 10.3389/fgene.2016.00163

Porter, G. A., and Rivkees, S. A. (2001). Ontogeny of humoral heart rate regulation in the embryonic mouse. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 281, R401–R407. doi: 10.1152/ajpregu.2001.281.2.R401

Pullen, A. B., Kain, V., Serhan, C. N., and Halade, G. V. (2020). Molecular and Cellular Differences in Cardiac Repair of Male and Female Mice. *J. Am. Heart Assoc.* 9:e015672. doi: 10.1161/JAHA.119.015672

Ramani, V., Deng, X., Qiu, R., Lee, C., Disteche, C. M., Noble, W. S., et al. (2020). Sci-Hi-C: A single-cell Hi-C method for mapping 3D genome organization in large number of single cells. *Methods* 170, 61–68. doi: 10.1016/j.ymeth.2019.09.012

Red-Horse, K., Ueno, H., Weissman, I. L., and Krasnow, M. A. (2010). Coronary arteries form by developmental reprogramming of venous cells. *Nature* 464, 549–553. doi: 10.1038/nature08873

Rensen, S. S. M., Doevendans, P. A. F. M., and Van Eys, G. J. J. M. (2007). Regulation and characteristics of vascular smooth muscle cell phenotypic diversity. *Netherlands Heart J.* 15, 100–108. doi: 10.1007/BF03085963

Rich-Griffin, C., Stechemesser, A., Finch, J., Lucas, E., Ott, S., and Schäfer, P. (2020). Single-Cell Transcriptomics: A High-Resolution Avenue for Plant Functional Genomics. *Trends Plant Sci.* 25, 186–197. doi: 10.1016/j.tplants.2019.10.008

Roberts, J. T., and Wearn, J. T. (1941). Quantitative changes in the capillary-muscle relationship in human hearts during normal growth and hypertrophy. *Am. Heart J.* 21, 617–633. doi: 10.1016/S0002-8703(41)90726-3

Rog-Zielinska, E. A., Norris, R. A., Kohl, P., and Markwald, R. (2016). The Living Scar – Cardiac Fibroblasts and the Injured Heart. *Trends Mole. Med.* 22, 99–114. doi: 10.1016/j.molmed.2015.12.006

Ruiz-Villalba, A., Romero, J. P., Hernández, S. C., Vilas-Zornoza, A., Fortelny, N., Castro-Labrador, L., et al. (2020). Single-Cell RNA Sequencing Analysis Reveals a Crucial Role for CTHRC1 (Collagen Triple Helix Repeat Containing 1) Cardiac Fibroblasts After Myocardial Infarction. *Circulation* 142, 1831–1847. doi: 10.1161/CIRCULATIONAHA.119.044557

Ruiz-Villalba, A., Simón, A. M., Pogontke, C., Castillo, M. I., Abizanda, G., Pelacho, B., et al. (2015). Interacting resident epicardium-derived fibroblasts and recruited bone marrow cells form myocardial infarction scar. *J. Am. College Cardiol.* 65, 2057–2066. doi: 10.1016/j.jacc.2015.03.520

Saelens, W., Cannoodt, R., Todorov, H., and Saeys, Y. (2019). A comparison of single-cell trajectory inference methods. *Nat. Biotechnol.* 37, 547–554. doi: 10.1038/s41587-019-0071-9

Sampaio-Pinto, V., Ruiz-Villalba, A., Nascimento, D. S., and Pérez-Pomares, J. M. (2020). Bone marrow contribution to the heart from development to adulthood. *Sem. Cell Devel. Biol.* 112, 16–26. doi: 10.1016/j.semcdb.2020.06.008

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* 33, 495–502. doi: 10.1038/nbt.3192

Schaum, N., Karkanias, J., Neff, N. F., May, A. P., Quake, S. R., Wyss-Coray, T., et al. (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* 562, 367–372. doi: 10.1038/s41586-018-0590-4

See, K., Tan, W. L. W., Lim, E. H., Tiang, Z., Lee, L. T., Li, P. Y. Q., et al. (2017). Single cardiomyocyte nuclear transcriptomes reveal a lincRNA-regulated de-differentiation and cell cycle stress-response in vivo. *Nat. Commun.* 8, 1–13. doi: 10.1038/s41467-017-00319-8

Sehnert, A. J., Huq, A., Weinstein, B. M., Walker, C., Fishman, M., and Stainier, D. Y. R. (2002). Cardiac troponin T is essential in sarcomere assembly and cardiac contractility. *Nat. Genet.* 31, 106–110. doi: 10.1038/ng875

Selewa, A., Dohn, R., Eckart, H., Lozano, S., Xie, B., Gauchat, E., et al. (2020). Systematic Comparison of High-throughput Single-Cell and Single-Nucleus Transcriptomes during Cardiomyocyte Differentiation. *Scientific Rep.* 10, 1–13. doi: 10.1038/s41598-020-58327-6

Sergeeva, I. A., Hooijkaas, I. B., Van Der Made, I., Jong, W. M. C., Creemers, E. E., and Christoffels, V. M. (2014). A transgenic mouse model for the simultaneous monitoring of ANF and BNP gene activity during heart development and disease. *Cardiovasc. Res.* 101, 78–86. doi: 10.1093/cvr/cvt228

Shiroguchi, K., Jia, T. Z., Sims, P. A., and Xie, X. S. (2012). Digital RNA sequencing minimizes sequence-dependent bias and amplification noise with optimized single-molecule barcodes. *Proc. Natl. Acad. Sci. U. S. A.* 109, 1347–1352. doi: 10.1073/pnas.1118018109

Skelly, D. A., Squiers, G. T., McLellan, M. A., Bolisetty, M. T., Robson, P., Rosenthal, N. A., et al. (2018). Single-Cell Transcriptional Profiling Reveals Cellular Diversity and Intercommunication in the Mouse Heart. *Cell Rep.* 22, 600–610. doi: 10.1016/j.celrep.2017.12.072

Soneson, C., and Robinson, M. D. (2018). Bias, robustness and scalability in single-cell differential expression analysis. *Nat. Methods* 15, 255–261. doi: 10.1038/nmeth.4612

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat. Genet.* 21, 70–71. doi: 10.1038/5007

Souders, C. A., Bowers, S. L. K., and Baudino, T. A. (2009). Cardiac fibroblast: The renaissance cell. *Circulat. Res.* 105, 1164–1176. doi: 10.1161/CIRCRESAHA.109.209809

Sperr, W. R., Bankl, H. C., Mundigler, G., Klappacher, G., Großschmidt, K., Agis, H., et al. (1994). The human cardiac mast cell: Localization, isolation, phenotype, and functional characterization. *Blood* 84, 3876–3884. doi: 10.1182/blood.v84.11.3876.bloodjournal84113876

Squiers, G. T., McLellan, M. A., Ilinykh, A., Branca, J., Rosenthal, N. A., and Pinto, A. R. (2020). Cardiac Cellularity is Dependent upon Biological Sex and is Regulated by Gonadal Hormones. *Cardiovasc. Res.* 17:cvaa265. doi: 10.1093/cvr/cvaa265

Srivastava, D. (2006). Making or Breaking the Heart: From Lineage Determination to Morphogenesis. *Cell* 126, 1037–1048. doi: 10.1016/j.cell.2006.09.003

Stegle, O., Teichmann, S. A., and Marioni, J. C. (2015). Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* 16, 133–145. doi: 10.1038/nrg3833

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., et al. (2019). Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888–1902.e21. doi: 10.1016/j.cell.2019.05.031

Stuart, T., Srivastava, A., Lareau, C., and Satija, R. (2020). Multimodal single-cell chromatin analysis with Signac. *BioRxiv* [preprint] 1–17. doi: 10.1101/2020.11.09.373613

Stubbington, M. J. T., Rozenblatt-Rosen, O., Regev, A., and Teichmann, S. A. (2017). Single-cell transcriptomics to explore the immune system in health and disease. *Science* 358, 58–63. doi: 10.1126/science.aan6828

Tallquist, M. D., and Molkentin, J. D. (2017). Redefining the identity of cardiac fibroblasts. *Nat. Rev. Cardiol.* 14, 484–491. doi: 10.1038/nrcardio.2017.57

Talman, V., and Kivelä, R. (2018). Cardiomyocyte—Endothelial Cell Interactions in Cardiac Remodeling and Regeneration. *Front. Cardiovasc. Med.* 5:101. doi: 10.3389/fcvm.2018.00101

Tang, Y., Nyengaard, J. R., Andersen, J. B., Baandrup, U., and Gundersen, H. J. G. (2009). The application of stereological methods for estimating structural parameters in the human heart. *Anatom. Rec.* 292, 1630–1647. doi: 10.1002/ar.20952

Theiler, K. (1972). *The house mouse. Development and normal stages from fertilization to 4 weeks of age.tle*. Berlin: Springer-Verlag.

Thompson, R. P., Abercrombie, V., and Wong, M. (1987). Morphogenesis of the truncus arteriosus of the chick embryo heart: Movements of autoradiographic tattoos during septation. *Anatom. Rec.* 218, 434–440. doi: 10.1002/ar.1092180411

Trapnell, C. (2015). Defining cell types and states with single-cell genomics. *Genome Res.* 25, 1491–1498. doi: 10.1101/gr.190595.115

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., et al. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* 32, 381–386. doi: 10.1038/nbt.2859

Travaglini, K. J., Nabhan, A. N., Penland, L., Sinha, R., Gillich, A., Sit, R. V., et al. (2020). A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* 587, 619–625. doi: 10.1038/s41586-020-2922-4

Tucker, N. R., Chaffin, M., Fleming, S. J., Hall, A. W., Parsons, V. A., Bedi, K. C., et al. (2020). Transcriptional and Cellular Diversity of the Human Heart. *Circulation* 142, 466–482. doi: 10.1161/circulationaha.119.045401

Vafadarnejad, E., Rizzo, G., Krampert, L., Arampatzi, P., Arias-Loza, P. A., Nazzal, Y., et al. (2020). Dynamics of Cardiac Neutrophil Diversity in Murine Myocardial Infarction. *Circulat. Res.* 127, e232–e249. doi: 10.1161/circresaha.120.317200

Vallejos, C. A., Risso, D., Scialdone, A., Dudoit, S., and Marioni, J. C. (2017). Normalizing single-cell RNA sequencing data: Challenges and opportunities. *Nat. Methods* 14, 565–571. doi: 10.1038/nmeth.4292

Van der Borght, K., Scott, C. L., Nindl, V., Bouché, A., Martens, L., Sichien, D., et al. (2017). Myocardial Infarction Primes Autoreactive T Cells through Activation of Dendritic Cells. *Cell Rep.* 18, 3005–3017. doi: 10.1016/j.celrep.2017.02.079

van Dijk, D., Sharma, R., Nainys, J., Yim, K., Kathail, P., Carr, A. J., et al. (2018). Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* 174, 716–729.e27. doi: 10.1016/j.cell.2018.05.061

van Eif, V. W. W., Devalla, H. D., Boink, G. J. J., and Christoffels, V. M. (2018). Transcriptional regulation of the cardiac conduction system. *Nat. Rev. Cardiol.* 15, 617–630. doi: 10.1038/s41569-018-0031-y

van Eif, V. W. W., Stefanovic, S., van Duijvenboden, K., Bakker, M., Wakker, V., de Gier-de Vries, C., et al. (2019). Transcriptome analysis of mouse and human sinoatrial node cells reveals a conserved genetic program. *Development* 146, 1–15. doi: 10.1242/dev.173161

van Weerd, J. H., and Christoffels, V. M. (2016). The formation and function of the cardiac conduction system. *Development* 143, 197–210. doi: 10.1242/dev.124883

Vanlandewijck, M., He, L., Mäe, M. A., Andrae, J., Ando, K., Del Gaudio, F., et al. (2018). A molecular atlas of cell types and zonation in the brain vasculature. *Nature* 554, 475–480. doi: 10.1038/nature25739

Velecela, V., Lettice, L. A., Chau, Y.-Y., Slight, J., Berry, R. L., Thornburn, A., et al. (2013). WT1 regulates the expression of inhibitory chemokines during heart development. *Hum. Mole. Genet.* 22, 5083–5095. doi: 10.1093/hmg/ddt358

Villani, A. C., Satija, R., Reynolds, G., Sarkizova, S., Shekhar, K., Fletcher, J., et al. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* 356:eaah4573. doi: 10.1126/science.aah4573

Vliegen, H. W., Van der Laarse, A., Cornelisse, C. J., and Eulderink, F. (1991). Myocardial changes in pressure overload-induced left ventricular hypertrophy. *Eur. Heart J.* 12, 488–494. doi: 10.1093/oxfordjournals.eurheartj.a059928

Wagner, F., Yan, Y., and Yanai, I. (2017). K-nearest neighbor smoothing for high-throughput single-cell RNA-Seq data. *BioRxiv* [preprint]. doi: 10.1101/217737

Walsh, S., Pontén, A., Fleischmann, B. K., and Jovinge, S. (2010). Cardiomyocyte cell cycle control and growth estimation in vivo-An analysis based on cardiomyocyte nuclei. *Cardiovasc. Res.* 86, 365–373. doi: 10.1093/cvr/cvq005

Wang, J., Duan, Y., Sluijter, J. P. G., and Xiao, J. (2019). Lymphocytic subsets play distinct roles in heart diseases. *Theranostics* 9, 4030–4046. doi: 10.7150/thno.33112

Wang, L., Yu, P., Zhou, B., Song, J., Li, Z., Zhang, M., et al. (2020). Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. *Nat. Cell Biol.* 22, 108–119. doi: 10.1038/s41556-019-0446-7

Wegmann, R., Neri, M., Schuierer, S., Bilican, B., Hartkopf, H., Nigsch, F., et al. (2019). CellSIUS provides sensitive and specific detection of rare cell populations from complex single-cell RNA-seq data. *Genome Biol.* 20:142. doi: 10.1186/s13059-019-1739-7

Wessels, A., van den Hoff, M. J. B., Adamo, R. F., Phelps, A. L., Lockhart, M. M., Sauls, K., et al. (2012). Epicardially derived fibroblasts preferentially contribute to the parietal leaflets of the atrioventricular valves in the murine heart. *Devel. Biol.* 366, 111–124. doi: 10.1016/j.ydbio.2012.04.020

Wigle, J. T., and Oliver, G. (1999). Prox1 function is required for the development of the murine lymphatic system. *Cell* 98, 769–778. doi: 10.1016/S0092-8674(00)81511-1

Wolfien, M., Galow, A. M., Müller, P., Bartsch, M., Brunner, R. M., Goldammer, T., et al. (2020b). Single nuclei sequencing of entire mammalian hearts: strain-dependent cell-type composition and velocity. *Cardiovasc. Res.* 116, 1249–1251. doi: 10.1093/cvr/cvaa054

Wolfien, M., Galow, A.-M., Müller, P., Bartsch, M., Brunner, R. M., Goldammer, T., et al. (2020a). Single-Nucleus Sequencing of an Entire Mammalian Heart: Cell Type Composition and Velocity. *Cells* 9:318. doi: 10.3390/cells9020318

Wolock, S. L., Lopez, R., and Klein, A. M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* 8, 281–291.e9. doi: 10.1016/j.cels.2018.11.005

Wu, A. R., Neff, N. F., Kalisky, T., Dalerba, P., Treutlein, B., Rothenberg, M. E., et al. (2014). Quantitative assessment of single-cell RNA-sequencing methods. *Nat. Methods* 11, 41–46. doi: 10.1038/nmeth.2694

Wu, B., Zhang, Z., Lui, W., Chen, X., Wang, Y., Chamberlain, A. A., et al. (2012). Endocardial cells form the coronary arteries by angiogenesis through myocardial-endocardial VEGF signaling. *Cell* 151, 1083–1096. doi: 10.1016/j.cell.2012.10.023

Xia, B., and Yanai, I. (2019). A periodic table of cell types. *Development* 146:dev169854. doi: 10.1242/dev.169854

Xu, Y., Zhang, Z., You, L., Liu, J., Fan, Z., and Zhou, X. (2020). scIGANs: single-cell RNA-seq imputation using generative adversarial networks. *Nucl. Acids Res.* 48:e85. doi: 10.1093/nar/gkaa506

Yekelchyk, M., Guenther, S., Preussner, J., and Braun, T. (2019). Mono- and multi-nucleated ventricular cardiomyocytes constitute a transcriptionally homogenous cell population. *Basic Res. Cardiol.* 114, 1–13. doi: 10.1007/s00395-019-0744-z

Yoshida, T., and Owens, G. K. (2005). Molecular determinants of vascular smooth muscle cell diversity. *Circulat. Res.* 96, 280–291. doi: 10.1161/01.RES.0000155951.62152.2e

Zak, R. (1973). Cell proliferation during cardiac growth. *Am. J. Cardiol.* 31, 211–219. doi: 10.1016/0002-9149(73)91034-5

Zaman, R., Hamidzada, H., and Epelman, S. (2021). Exploring cardiac macrophage heterogeneity in the healthy and diseased myocardium. *Curr. Opin. Immunol.* 68, 54–63. doi: 10.1016/j.coi.2020.09.005

Zeisberg, E. M., Tarnavski, O., Zeisberg, M., Dorfman, A. L., McMullen, J. R., Gustafsson, E., et al. (2007). Endothelial-to-mesenchymal transition contributes to cardiac fibrosis. *Nat. Med.* 13, 952–961. doi: 10.1038/nm1613

Zeisel, A., Munoz-Manchado, A. B., Codeluppi, S., Lonnerberg, P., La Manno, G., Jureus, A., et al. (2015). Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* 347, 1138–1142. doi: 10.1126/science.aaa1934

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., et al. (2017). Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 8:14049. doi: 10.1038/ncomms14049

Zhou, B., and Wang, L. (2020). Journal of Molecular and Cellular Cardiology Reading the heart at single-cell resolution. *J. Mole. Cell. Cardiol.* 148, 34–45. doi: 10.1016/j.yjmcc.2020.08.010

Zhou, Y., and Zhang, J. (2020). Single-Cell Transcriptomics: New Insights in Heart Research. *Circulat* 141, 1720–1723. doi: 10.1161/CIRCULATIONAHA.120.046043

Zhu, J., Yamane, H., and Paul, W. E. (2010). Differentiation of effector CD4+ T cell populations. *Annu. Rev. Immunol.* 28, 445–489. doi: 10.1146/annurev-immunol-030409-101212

Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., et al. (2017). Comparative Analysis of Single-Cell RNA Sequencing Methods. *Mole. Cell* 65, 631–643.e4. doi: 10.1016/j.molcel.2017.01.023

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

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership