Journal of Molecular and Cellular Cardiology Plus 11 (2025) 100282

Contents lists available at ScienceDirect

Journal of Molecular and Cellular Cardiology Plus

journal homepage: www.journals.elsevier.com/journal-of-molecular-andcellular-cardiology-plus

Integrated transcriptomic and regulatory RNA profiling reflects complex pathophysiology and uncovers a conserved gene signature in end stage heart failure

Amit Anand [a,*], Julius Punnen [b], U.M. Nagamalesh [b], Sabariya Selvam [c], Madhusudhan Bysani [c], Ramya Venkatesh [c], Kriti Nawin [c], Shilpa Garg [c], Bagirath Raghuraman [b], Varun Shetty [b], Senthil Kumaran [c], Manoj Dokania [c], Pradeep Narayan [b], Ankita Udwadia [d], Kushan Gunawardhana [d], David Gordon [d], Manjunath Ramarao [d], Lei Zhao [d], Jyoti Gulia [a,*]

[a] Bristol Myers Squibb, Bengaluru, India
[b] Narayana Institute of Cardiac Sciences, Narayana Health, Bengaluru, India
[c] Biocon Bristol Myers Squibb Research Center, Syngene International Limited, Bengaluru, India
[d] Bristol Myers Squibb, Princeton, NJ, United States

ARTICLE INFO

Keywords:
Heart failure with reduced ejection fraction (HFrEF)
Ischemic cardiomyopathy (ICM)
Non-ischemic cardiomyopathy (NICM)
Heart failure transcriptomics
miRNA heart failure
Machine learning transcriptomics
miRNA-seq
Heart failure gene signature

ABSTRACT

Background: Heart failure (HF) is a complex syndrome. Despite availability of multiple treatment options, the mortality remains high and the quality of life poor. Better understanding of the underlying pathophysiological processes can lead to development of novel therapies. Multiple comparative transcriptomics studies, which revealed gene level changes in the key pathophysiological pathways in failing hearts, point towards heterogeneity from interplay of disease stage, etiologies and ethnicity. Transcriptomic characterization of HF in patients from different ethnicities can potentially help in understanding the heterogeneity imparted by various factors and the core elements in heart failure.

Methods & results: An integrated analysis of bulk transcriptome and microRNA sequencing from the cardiac tissues of 30 South Asian (SA) patients having HF with reduced ejection fraction (HFrEF) and 19 control subjects was conducted. Plasma miRNAs from a subset of HFrEF and control patients were also sequenced to understand their biomarker potential. The altered transcriptome from the myocardium of SA HFrEF patients reflected cardiac muscle contraction, cellular energetics, altered immune signaling and extracellular matrix remodelling as predominant pathophysiological mechanisms. The SA HFrEF patients also showed dysregulation of multiple microRNAs in cardiac tissue like miR-216, miR-217, miR-184 and miR-9983. Many of these miRNAs, such as miR184 and few others, showed altered levels in both the plasma and cardiac tissue of HFrEF patients suggesting their biomarker potential. The diversity in the HFrEF transcriptomes from published studies led us to examine the core HF genes in our cohort. A gene signature generated using machine learning (ML) from the top dysregulated genes in SA HFrEF cohort stratified HF from controls in other cohorts. The sensitivity of the HF gene signature was further improved when union of two cohorts was used as a training set. Our ML analyses developed a core HF gene signature consisting of 21 genes that can stratify HF patients from controls with 98 % sensitivity in all the tested cohorts.

Conclusions: This study reveals molecular changes underlying the pathophysiology as reflected by coding and regulatory non-coding components of transcriptome from South Asian patients and uncovers a conserved gene signature for HF.

Abbreviations: AUCROC, area under curve of receiver operating characteristic; DEG, differentially expressed genes; ECM, extracellular matrix; FDR, false discovery rate; GSEA, gene set enrichment analysis; GDMT, guidance direct medical therapy; RF, random forest; SA, South Asian.

* Corresponding authors at: Bristol Myers Squibb, Bengaluru, India.

E-mail addresses: amit.anand@bms.com (A. Anand), Julius.punnen.dr@narayanahealth.org (J. Punnen), Sabariya.Selvam@Syngeneintl.com (S. Selvam), Madhu.Bysani@Syngeneintl.com (M. Bysani), Ramya.V@Syngeneintl.com (R. Venkatesh), Kriti.Nawin@Syngeneintl.com (K. Nawin), bagirath.raghuraman.dr@narayanahealth.org (B. Raghuraman), varun.shetty.dr@narayanahealth.org (V. Shetty), Senthil.Kumaran@Syngeneintl.com (S. Kumaran), Manoj.Dokania@Syngeneintl.com (M. Dokania), pradeep.narayan.dr@narayanahealth.org (P. Narayan), Ankita.Udwadia@bms.com (A. Udwadia), Jyoti.Gulia@bms.com (J. Gulia).

https://doi.org/10.1016/j.jmccpl.2025.100282
Received 24 October 2024; Received in revised form 18 December 2024; Accepted 2 January 2025
Available online 3 January 2025

1. Introduction

Heart failure (HF) is a global health issue with increasing prevalence due to ageing population across the world [1,2]. Among HF patients, about 30–60 % present with heart failure with reduced ejection fraction (HFrEF) depending on geography [2]. While the guideline directed medical therapy (GDMT) has reduced mortality in HFrEF [1–3], novel therapies are needed to bring down the hospitalization and residual mortality, which in turn requires deeper understanding of disease biology [1]. Multiple "omics" studies have been carried out at bulk and single cell level comparing diseased and normal human cardiac tissues to investigate HF pathophysiology. To date, most studies have focused on changes in protein coding transcripts [4–15] to delineate molecular changes underlying HF. However, the studies profiling the transcriptome in HFrEF also reflect significant variation, likely introduced by differences in ethnicities, disease heterogeneity and technical factors. Such variations have prompted investigation of core HF-related genes and signatures [16]. Significant variations imparted by these factors highlights the importance of examining gene expression in diverse disease cohorts to better understand the core elements of pathophysiology.

Last two decades have uncovered regulatory role of non-coding RNAs in physiological as well as disease conditions. The regulatory non-coding RNAs such as long non-coding RNAs (lncRNAs), small interfering RNAs (siRNAs), microRNAs (miRNAs), fine tune the gene expression transcriptionally and post-transcriptionally. Abnormal changes in non-coding transcriptome are now recognized as drivers of pathophysiology in HF [17–20]. These non-coding RNAs have been shown to have both therapeutic and biomarker potential in HF [17–19]. However, investigations on paired transcriptional and regulatory non-coding RNA dysregulation in cardiac tissue as well as matched plasma samples are lacking. Such comprehensive data sets can help reveal multiple layers of gene regulation in disease setting and also, provide strong biomarker leads.

The goals of our study were to i) investigate protein coding and regulatory non-coding RNA dysregulation from the bulk transcriptomic data generated using cardiac tissue obtained from South Asian (SA) HFrEF subjects, ii) identify a core HF stratifying gene signature, and iii) probe circulatory biomarker potential of cardiac tissue dysregulated miRNAs. We performed mRNA & miRNA sequencing in failing and non-failing left ventricular (LV) tissue and followed it up with matched plasma miRNA profiling from a subset of HFrEF patients. The integrated analysis of coding and non-coding components of transcriptome highlights the predominant pathways underlying the pathophysiological changes. We show multiple dysregulated non-coding RNAs out of which some could have biomarker potential; these analyses provide a snapshot of the RNA-based regulation of gene expression in the HFrEF patients. Finally, using machine learning, we uncover a core heart failure gene signature in our cohort.

2. Methods

A detailed list of methods is provided in the supplementary methods section.

2.1. Sample acquisition

Cardiac tissue and blood samples were collected from HFrEF patients and control subjects at Narayana Institute of Cardiac Sciences, Narayana Health, Bangalore, India. The study conformed to the Helsinki declaration and the clinical sample collection protocol was approved by the Ethics Committee (NHH/MEC-CL-2017-462) at Narayana Health, Bangalore, India. Informed consent was obtained from all subjects for participation in the study. Appropriate measures were taken to protect donor identity while collecting clinical information of study subjects. Failing heart samples were obtained from patients (NYHA class $\geq$ 2) undergoing heart transplant or left ventricular assist device (LVAD) implantation procedures. In addition, blood samples were collected from a subset of patients prior to cardiac transplant. Non-failing control heart samples were collected from brain dead patients whose hearts could not be used for organ transplant due to non-availability of a suitable recipient based on blood group mismatch, height, weight and size criteria. Control blood samples were collected from age-matched volunteers with no cardiovascular disease history (based on self-reporting and NT-proBNP test results). The cardiac tissue was collected from LV apex region and stored in RNAlater until further processing. The blood samples were collected in PAXgene® RNA tubes.

2.2. Cardiac tissue processing, RNA extraction from tissue and plasma

The cardiac tissue stored in RNAlater was directly added into 700 μL of QIAzol lysis reagent and homogenized using Qiagen TissueLyser II. Total RNA extraction including miRNAs was performed using miRNeasy Mini Kit as per the manufacturer's instructions. Plasma miRNA extraction was performed with 200 μL of plasma using miRNeasy Serum/Plasma kit.

2.3. RNA library preparation and sequencing

Total transcriptome libraries were prepared from 100 ng of total RNA from tissue. The RNA quality was examined using Tapestation (Agilent). The libraries were prepared using TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero Gold. Manufacturer's instructions were followed to prepare RNA-seq libraries. The libraries were sequenced on Illumina NovaSeq 6000 system. Global miRNA library was prepared using Qiagen's QiaSeq miRNA Library Kit. The miRNA libraries from plasma samples were prepared with the same kit using 5 μL of miRNA from plasma, following manufacturer's instructions.

2.4. Data analyses and statistics

RNA-seq data was aligned using splice-aware mapper. The genes represented by a very low number of reads in all the samples were filtered out. The downstream analyses of RNA-seq data was conducted in R Studio (version 4.02). The reads were normalised and differential gene expression analysis was performed using linear models. Description regarding read alignment, normalization and differential gene expression are provided in supplementary methods section. Pathway enrichment among the differentially expressed genes (DEGs) was investigated using gene set enrichment analysis (GSEA) and hypergeometric tests. Libraries were subjected to transcript prediction using a software package that implements network flow approach. The unannotated transcripts were examined for their coding potential to identify the novel long non-coding RNAs. The miRNA libraries were aligned to genome and differential expression of miRNAs was evaluated using linear models. The heart failure signature was developed using a random forest-based method. The stratification potential of genes in the signature was calculated using Area under Curve (AUC) of the Receiver Operating Characteristic (ROC) curve (AUCROC) analysis.

3. Results

3.1. Clinical characteristics of subjects

We investigated 30 and 19 cardiac tissue samples from HFrEF and control subjects, respectively. In the HFrEF group, 13 subjects had ischemic cardiomyopathy (ICM) and 17 subjects had non-ischemic cardiomyopathy (NICM) as underlying etiologies. The median age and gender distribution were similar across both the control and HFrEF groups (Fig. S1A). Detailed clinical information of HFrEF subjects in our study is summarized in Table S1. The median age in ICM group was significantly higher than in the NICM group, as expected (Table S1). The normal subjects in this study were defined by absence of cardiovascular

disease.

miRNA sequencing from matched plasma samples was conducted for 7 HFrEF patients. For the normal subjects recruited for plasma miRNA analyses, NT-proBNP levels were measured and all five subjects had levels lower than 60 pg/mL, which is well within the normal range. The median age of both normal ($n = 5$) and HFrEF ($n = 7$) subjects that were considered for plasma miRNA analyses was same (51 years).

3.2. Transcriptomic landscape of South Asian HFrEF patients

Principal component analysis of failing and non-failing cardiac transcriptomes, showed clear segregation between HF and control samples, suggesting that HF leads to significant changes in the transcriptome (Fig. 1A). Principal variable component analysis (PVCA) also showed HF as a predominant factor underlying variation in the gene expression and that the variations driven by other biological and technical factors like gender and batch were relatively minor (Fig. S1B).

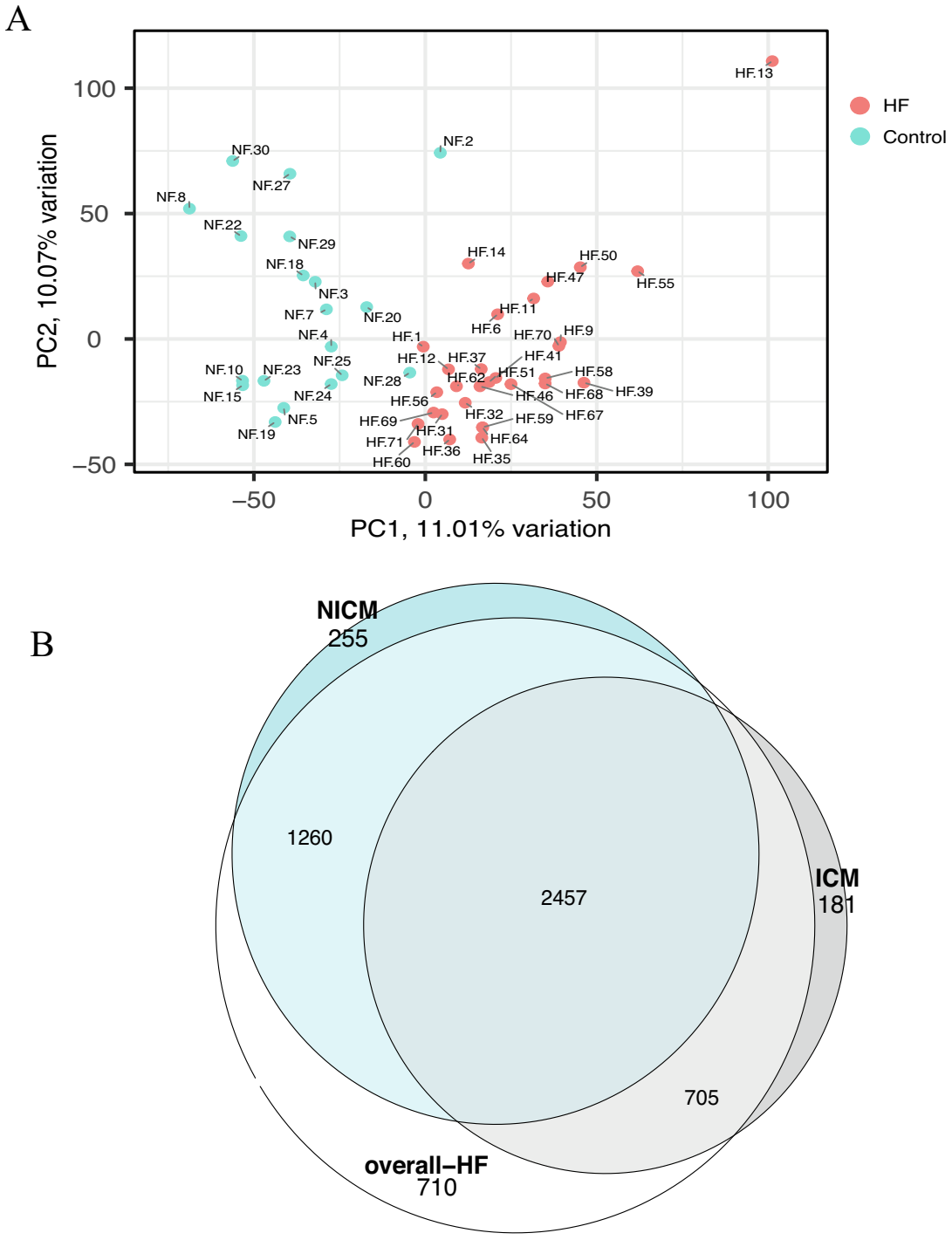

Fig. 1. Heart failure leads to large scale changes in transcriptome. (A) Principal component analysis of transcriptomes from 30 HFrEF (red) and 19 control (blue) heart samples. First two principal components (PC1 and PC2) are plotted. (B) Overlap of differentially expressed genes (DEGs), with 5 % FDR threshold, obtained after comparison between transcriptomes of all HFrEF and control, ICM (13 samples) and control and NICM (17 samples) and control samples. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

We obtained 5132 DEGs between HFrEF and control transcriptomes using 5 % false discovery rate (FDR) threshold. Of these, 751 genes showed >2-fold change in expression in HFrEF samples compared to controls (Table S2). Independent comparison of transcriptomes from NICM and ICM patients with control transcriptomes showed a large degree of overlap among the DEGs from the two etiologies with some unique DEGs for each etiology (Fig. 1B).

The top pathways enriched among the upregulated genes in the failing hearts were ECM remodelling, cytokine signaling, angiotensinogen metabolism, Wnt and Insulin-like growth factors signaling pathways (Fig. 2A). The GSEA with Reactome and Gene Ontology knowledge bases revealed significant upregulation of ECM pathways reflected by enrichment of terms like ECM reorganization, collagen biosynthesis and aggregation (Table S3). Similarly, enrichment of T cell and cytokine signaling pathways was shown by terms like T cell activation, T cell differentiation, cytokine signaling and chemotaxis (Table S3, Fig. 2A). Among the downregulated genes in HFrEF, the top enriched pathways were cardiac muscle contraction and actin and tubulin protein folding (Fig. 2B). We also observed downregulation of few innate immune pathway genes (Fig. 2B). Concordant with published

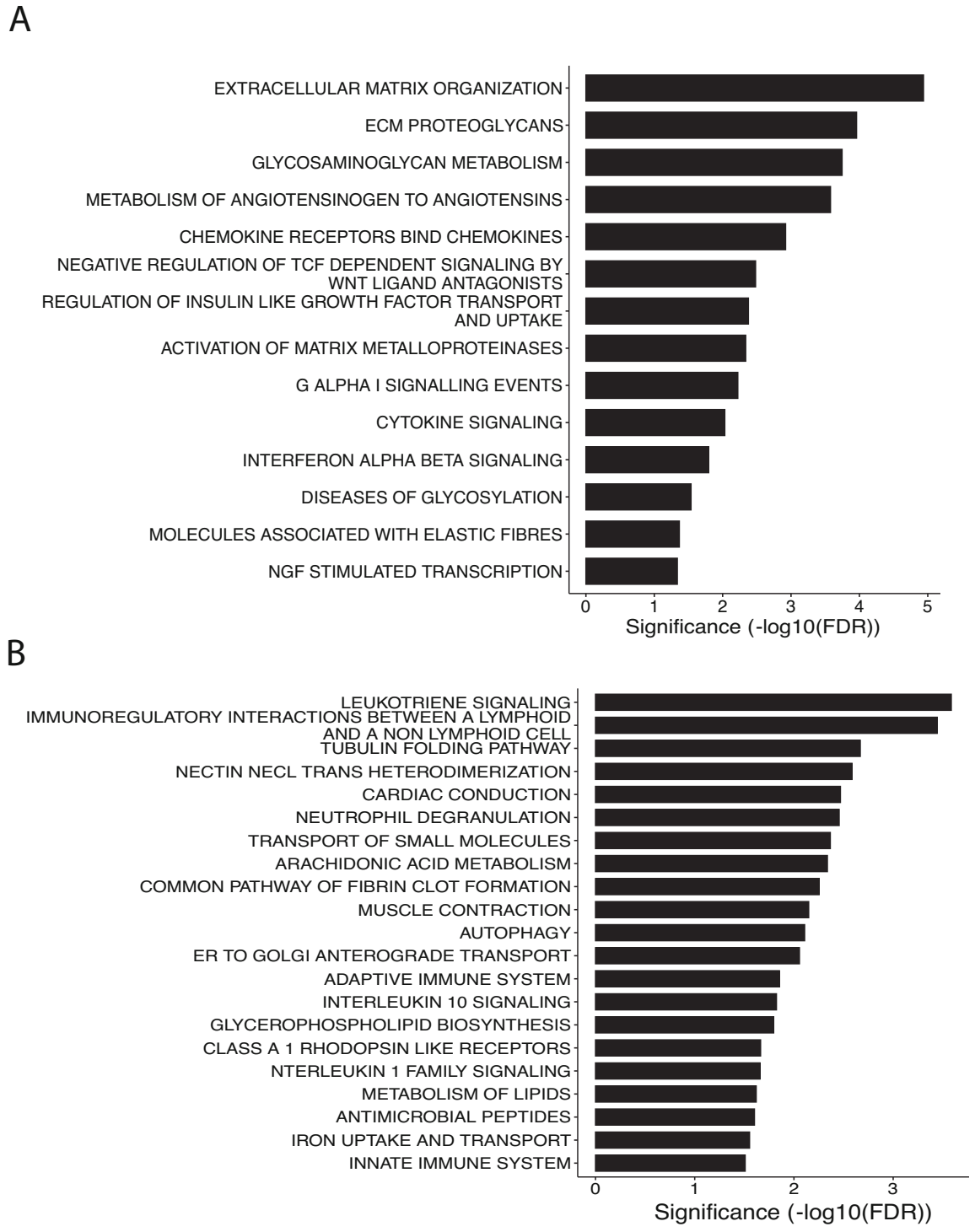

Fig. 2. Pathway enrichment analyses for DEGs in HFrEF samples. (A) Top enriched pathways obtained from upregulated genes. (B) Top enriched pathways from downregulated genes. Hypergeometric test was used for pathway analyses using DEGs with ≥1.5 fold expression change with FDR > 0.05.

findings [16], we observed downregulation of several cardiac contraction and conduction related genes such as *MYH6* and *ATP2A2* (Fig. S2). The GSEA also showed negative enrichment of pathways like electron transport chain and mitochondrial function (Table S3). Overall, these changes were suggestive of pathological ventricular remodelling and cardiac dysfunction.

Next, we examined the transcriptomic changes associated with underlying etiologies i.e. ICM and NICM. Multiple pathways were dysregulated in both ICM and NICM subjects. However, the extent of dysregulation, as reflected by differences in number of genes in a pathway or fold-change, varied between the etiologies. The NICM group showed greater downregulation of genes from multiple innate immune pathways compared to the ICM group (Fig. S3A, S3B and S3C). Multiple sarcomeric, ion channel and unfolded protein response (UPR) pathway related genes showed higher reduction in NICM group than ICM group (Fig. S3A and S3D). This suggests that altered muscle conduction and contraction are predominant pathophysiological mechanisms in NICM. The ICM group showed greater reduction in the expression of genes from cellular energetics pathways compared to NICM group (Fig. S3B and S3E). Specifically, many components of complex I and complex III of electron transport chain and related genes were downregulated in ICM group suggesting dysregulated mitochondrial function potentially due to chronic ischemic etiology. Some of the positively enriched overlapping pathways in both NICM and ICM groups included upregulation of T cell activation, cytokine signaling, ECM remodelling and angiotensinogen metabolism (Figs. S3A and S3B). ECM remodelling related genes including a few collagen genes exhibited increased upregulation in ICM group in comparison to NICM group (Fig. S3F). The dysregulation of angiotensinogen metabolism pathway in both ICM and NICM subjects (Fig. S3A and S3B) with overexpression of genes like *ACE*, *CMA1*, *CTSG* and *MME* highlighted potential myocardial production of angiotensin-II from angiotensin-I leading to localized Ang-II mediated fibrosis in heart (Table S2) [21,22]. Although upregulated in both NICM and ICM, the extent of upregulation of T cell activation and cytokine signaling pathways was higher in ICM subjects (Fig. S3A and S3B).

To further understand the immune component related changes in ICM and NICM, we examined the enrichment of gene sets representing signatures of different immune cells in our bulk transcriptomics data using single sample GSEA (ssGSEA). The NICM subgroup showed stronger downregulation of genes from neutrophil degranulation pathway like *TNFRSF10*, *MGAM*, and *CSF3R*, and increased expression of genes from mast cell related genes like CMA1, CPA3 and MS4A2 (Figs. 3 and S4A). In contrast, the ICM subgroup had significantly higher enrichment of activated CD8 T cells and cytotoxic cell signatures (Figs. 3 and S4B). We observed upregulation of genes like *GNLY* and *NKG7*, which are generally enriched in the cytotoxic cells, in ICM subjects (Fig. S4B). Overall, these findings suggest a greater involvement of activated CD8 T cells in ICM.

3.3. Dysregulation of regulatory RNAs in heart failure

To elucidate the role of short and long regulatory RNAs in HF pathophysiology, we performed miRNA profiling from the same cardiac tissue samples that were used for transcriptome sequencing. Additionally, we examined changes in the lncRNA transcripts, both known and novel, in the global transcriptomics data. Among the HF DEGs, 228 transcripts were classified as lncRNAs, including antisense and divergent transcripts from protein coding loci and other lncRNAs (Table S4). For multiple antisense lncRNAs, we observed dysregulation of the protein coding genes originating from the same locus. In many instances, dysregulation of antisense and protein coding transcripts was in the same direction; while for some loci, the antisense and protein coding transcripts expression were in opposite direction (Table S5). Notably, *C1QTNF1*, which upon downregulation leads to accelerated fibrosis and hypertrophy in heart [23], was downregulated and the *C1QTNF1-AS1* was upregulated (Table S5). Overall, we observed dysregulation of a number of annotated lncRNAs that are implicated in heart failure and atherosclerosis, for example, *H19*, *BANCR*, *MEG3*, *GASAL1*, *CERNA1*, *SENCR* and *HCG11* [24–30].

We then examined novel lncRNAs in the SA HFrEF cohort. We evaluated the coding potential of the unannotated transcripts from the overall assembled transcripts set. The analysis revealed 349 transcripts with very poor coding potential that could be putative novel lncRNAs. These predicted lncRNAs did not have any match in the public lncRNA databases and some of these had substantially higher expression in the HFrEF samples as indicated by the TPM values. We designed primers against 10 putative lncRNAs that showed overall high expression in control and HF samples and performed qPCR on a subset of HFrEF samples from the present cohort. We confirmed expression of these lncRNAs and significant upregulation of seven and downregulation of

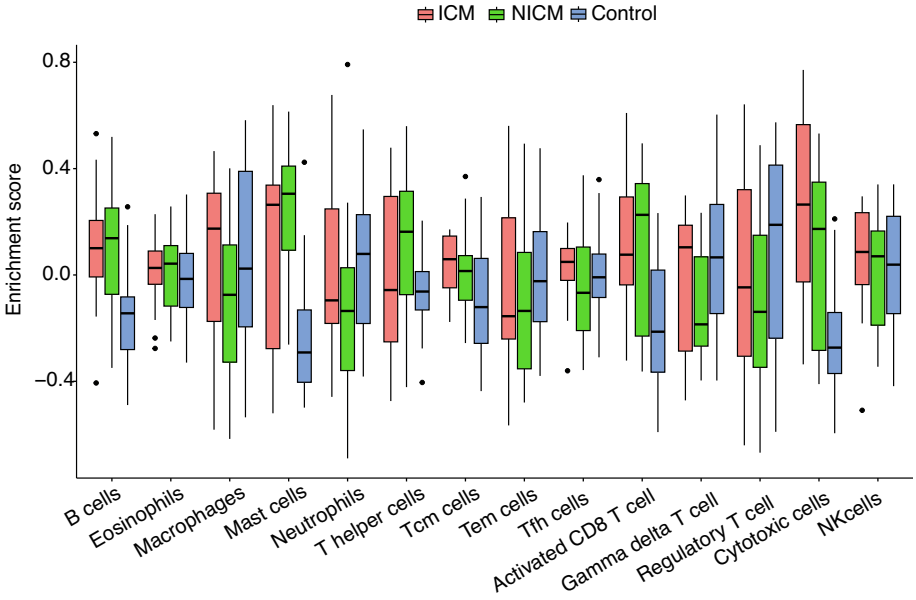

Fig. 3. Enrichment of immune cell signatures in the NICM, ICM and control heart transcriptomes. Single-sample GSEA was used to estimate the enrichment of the gene signatures from different immune cells in the respective transcriptomes. Enrichment scores for each cell type across control (blue), ICM (red) and NICM (green) subjects are represented by boxplot. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

one putative lncRNA in the failing hearts (Fig. S5). Dysregulation of multiple known and novel lncRNAs in HFrEF patients suggests the involvement of lncRNAs in HF pathophysiology.

As described, we also sequenced the miRNAs from the same RNA pool of control and HFrEF samples that was used to perform the transcriptomic profiling, and detected altered expression of 130 miRNAs in the HF group below 5 % FDR (Table S6). Some of the top upregulated miRNAs were miR-216, miR-217, miR-184, miR-204, miR-599 and miR-211. Dysregulation of many of these miRNAs has already been described in heart failure [19,31–33]. Among the downregulated miRNAs, miR-9983, miR-4284, miR-518, miR-122, miR-1261 showed significant change. In order to understand the regulatory potential of these miRNAs, we predicted the target transcripts of the miRNA using TargetScan tool [34]. For the downregulated miRNAs, we selected the putative target mRNAs that were upregulated; conversely, for the upregulated miRNAs, we selected putative targets that were downregulated. The downregulated putative targets with corresponding increase in the corresponding miRNAs showed enrichment of pathways like ionic transport, apoptosis, muscle conduction, mitochondrial functions, immune pathways and cellular integrity (Fig. 4A). The upregulated genes, which were predicted to be targets of downregulated miRNAs, showed enrichment of terms like PI3K-AKT signaling, IGF signaling, ECM remodelling, adaptive immune pathways, specifically some of the Th2 genes (Fig. 4B). These results suggest that dysregulation of a subset of genes in failing hearts could potentially be driven by altered miRNA expression. To examine the validity of miRNA regulation at cohort level, we looked at correlation between the expression levels of dysregulated miRNA and corresponding target mRNAs in failing and control heart samples. Indeed, we observed significant inverse correlation of 40 mRNA-miRNA pairs in HF patients. In comparison, these pairs showed insignificant correlation in the control subjects (data not shown). For example, expression of genes like *KCNN3*, *CXCR4*, *IGF1* was inversely correlated specifically with expression of miR-1-3p in HF subjects (Fig. S6A). Similar correlation was also observed for *FGF18* and miR-21-5p target-miRNA pair in HF group (Fig. S6A). To summarize, paired miRNA profiling revealed another layer of gene regulation in our cohort.

Next, we evaluated whether the differentially regulated cardiac miRNAs could also be detected in circulation with differences in their

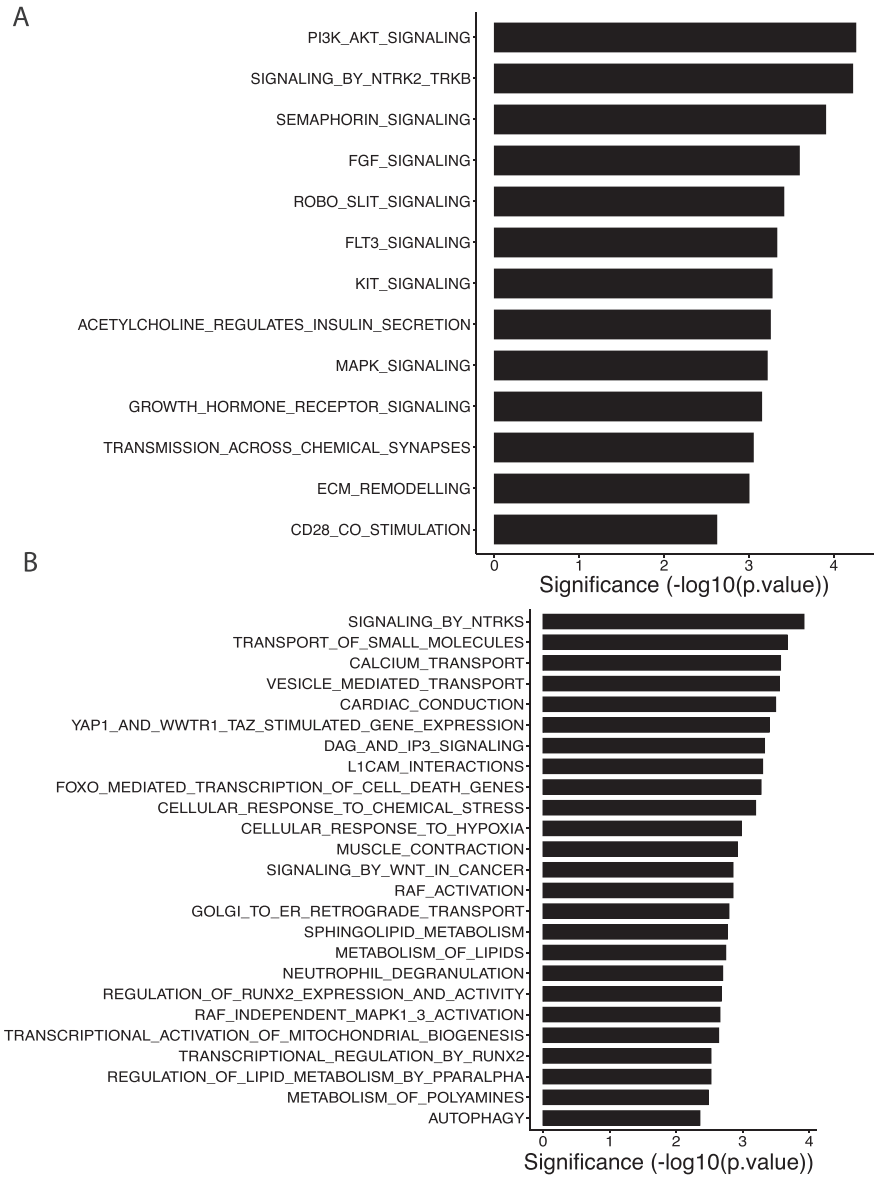

Fig. 4. Pathway enrichment analysis for putative miRNAs target genes dysregulated in HFrEF. (A) Top enriched pathways among upregulated genes that were putative targets of downregulated miRNAs. (B) Top enriched pathways from downregulated genes that were putative targets of upregulated miRNAs.

levels between the control and HF subjects. Circulatory miRNAs are altered in multiple disease indications including cardiovascular diseases and have shown biomarker potential for both diagnostic and prognosis purposes [20,35]. The miRNA sequencing from a subset of control and HF subject plasma samples revealed substantial miRNA dysregulation in plasma of HF subjects (429 miRNAs at 5 % FDR; Table S6). Among these, 12 miRNAs showed significantly robust dysregulation in the same direction in both myocardium and plasma samples. For example, miR-183 and miR-184 were robustly upregulated in the both myocardium and plasma samples, and miR-187, miR-187-3p, miR-1-3p were significantly downregulated in both tissues and plasma samples (Fig. S6B and S6C). These findings suggest a promising role of these miRNAs as biomarkers.

3.4. Machine learning derived HF gene signature from SA cohort

Meta analysis of published HFrEF transcriptomes has highlighted poor overlap among DEGs obtained in different studies [16]. We compared DEGs from the current study with those from three published studies - Sweet et al. 2018 [6], Hahn et al. 2021 (only HFrEF samples) [8] and Yang et al. 2014 [7] and found a poor overlap among their DEGs. This was reflected by Jaccard-Index, which ranged between 0.2 to 0.4 for top 500 DEGs from the aforementioned studies (Fig. S7A), reaffirming the findings of an earlier published study by Flores et al. [16]. In addition, <20 % of top 50 DEGs from any given cohort were differentially expressed in other cohorts (Fig. S7B). To delineate the core HF genes from SA cohort and the three published cohorts, we used a Random Forest (RF) based method. We trained the RF model on each cohort separately and examined the stratification ability of top RF-derived genes in other cohorts. The stratification ability of the top RF-derived genes was assessed solely on their expression. In contrast to top DEGs, the majority of top 50 RF-derived genes from any one cohort were DEGs in other three cohorts (Fig. S7C). The top 50 RF-derived genes after training each cohort separately could largely stratify the HF and control subjects in the remaining three cohorts (Fig. 5A). The average AUCROC of top 50 RF-derived genes from each cohort ranged between 0.7 and 0.9 (Fig. 5A). Among these, the signature derived using SA cohort as training data showed best average AUCROC followed by the signature from Hahn et al. 2021 (Figs. 5A and S7D).

While the top 50 RF-derived genes from individual cohorts exhibited reasonable stratification ability, the overlap of top 50 RF-derived genes from one cohort with the other three cohorts was between 60 and 90 %. To achieve better coverage from the training datasets, we combined two cohorts and used the union of their DEGs for RF model training. With this approach, the overlap of top 50 RF-derived genes with DEGs from test cohorts improved significantly compared to the overlap observed with single cohort training sets (Fig. S7C). In line with this observation, the signature generated using combined or dual training set showed improvement in AUCROC. This suggests that the signature generated using two cohorts increased chances of discovering core HF genes for disease stratification (Figs. 5B, S7D). The signatures generated from our study in combination with Yang et al. 2014 [7] and with Hahn et al. 2021 [8] showed highest average performance followed by the other combinations (Figs. 5B, S7D). These analyses show that machine learning (ML) based methods offer significant improvement in identifying the HF related genes overcoming the variation imparted by multiple factors including technical and disease heterogeneity. Finally, using the ML framework described above, we identified a core set of 21 gene signature that had high sensitivity, AUCROC >0.98, for all the cohorts (Fig. 5C). This set included genes like *MYH6*, *OGN*, *LUM*, *CCDC69*, *HMOX2*, *KCND3*, *FREM1*.

4. Discussion

In this study, we characterised the coding and regulatory non-coding compartments of the transcriptome from SA HFrEF patients. The coding component provided insights into HFrEF pathophysiology. The fibrotic changes and cardiomyocyte impairment in the failing hearts were well reflected by upregulation of multiple ECM reorganization, and collagen biosynthesis pathways and by downregulation of cardiac conduction, muscle contraction and cellular energetics pathways. We also observed significant changes in the adaptive and innate immune pathways. The expression of genes representing adaptive immune component, such as T cell activation, was markedly increased. Enrichment of a subset of T cell population expressing *GNLY* and *PRF1* has also been described previously in a single nuclear RNA-sequencing study from cardiac tissues obtained from end stage heart failure patients with hypertrophic and dilated cardiomyopathy as underlying etiologies [4]. In contrast to adaptive immune component, multiple genes from innate immune pathways were suppressed including neutrophil degranulation related genes like *CD14*, *LILRB3*, *S100A8*, *S100A9*, *CD177*. Reduction in neutrophil degranulation pathway has been reported in macrophages and other cell types in dilated cardiomyopathy [4,14]. Increased dysregulation of genes implicated in enhanced inflammation in this study, such as upregulation of *NLRP3*, *EGR1* and downregulation of *CD163*, *LYVE1* suggests a shift towards inflammatory macrophages, which has also been shown in dilated cardiomyopathy using single nuclear RNA-seq [14]. Multiple genes from cytokine signaling pathways like *IFIT1*, *IFIT2*, *MX1* and *MX2* and T cell migration and stimulation like *CCL21*, *CD3E*, *CD8A*, *HLA-DRA*, *PIK3R1* were upregulated in our study and the same has been reported for the other HFrEF cohorts as well [16]. With bulk transcriptomics, as in our study, it is difficult to ascertain these effects to a specific cell type. However, based on the commonality of underlying genes, it is likely that the dysregulated immune signature at bulk level in our study reflects the changes in subsets of lymphocytic cells and macrophages described at single cell level in failing hearts [4]. Overall, extensive dysregulation of inflammation related genes suggests that increased inflammation is pivotal in driving the pathophysiology in heart failure.

There was a substantial overlap in the DEGs between ICM and NICM groups. Similar convergence of disease-related changes at bulk or single cell level in end-stage heart failure patients with different underlying etiologies has also been reported previously [4]. We also observed that there were some distinctions that underscored the unique aspects of the underlying etiologies. Differential enrichment of pathophysiological pathways in ICM and NICM groups suggests that the net contribution of a disrupted biological mechanism towards driving disease progression may be different for different etiologies and also, highlights the unique pathophysiological features of each etiology. For example, we observed slightly higher dysregulation of genes from ECM remodelling, T cell activation, cytotoxic signature and cellular energetics pathways in the ICM subjects compared to NICM subjects. Enhanced expression of cytotoxic T cell signaling genes in the ICM patients like *GNLY*, *NKG7* and *CTSW* is reflective of potentially more aggressive fibrotic changes because of infarct related cardiomyocyte death. Santos-Zas et al. recently showed that in preclinical ischemia, activated CD8+ T cells release granzyme proteins leading to cardiomyocyte death and also, elevated granzyme B levels in acute MI patients are predictive of increased risk of death at 1 year [36]. Our findings suggest that T cell induced cardiotoxicity persists in individuals with ischemic cardiomyopathy. The NICM group showed negative enrichment of muscle conduction and contraction pathway and stronger downregulation of genes from innate immune pathways suggesting that the myocardial intrinsic factors potentially play a key role in NICM pathophysiology. Increased expression of human leukocyte antigen (HLA) genes was observed in both etiologies; HLA has been reported to be a risk locus for non-ischemic dilated cardiomyopathy [37,38].

We also observed dysregulation of several genes encoding for alpha- and beta-tubulins like *TUBA3D*, *TUBA3E*, *TUBB4B* in the myocardium of HF patients. Microtubules, formed by polymerization of alpha- and beta-tubulins, have been shown to play an important role in cardiac mechanics. Increased microtubular density and stability by post-translational modifications like detyrosination and acetylation are well

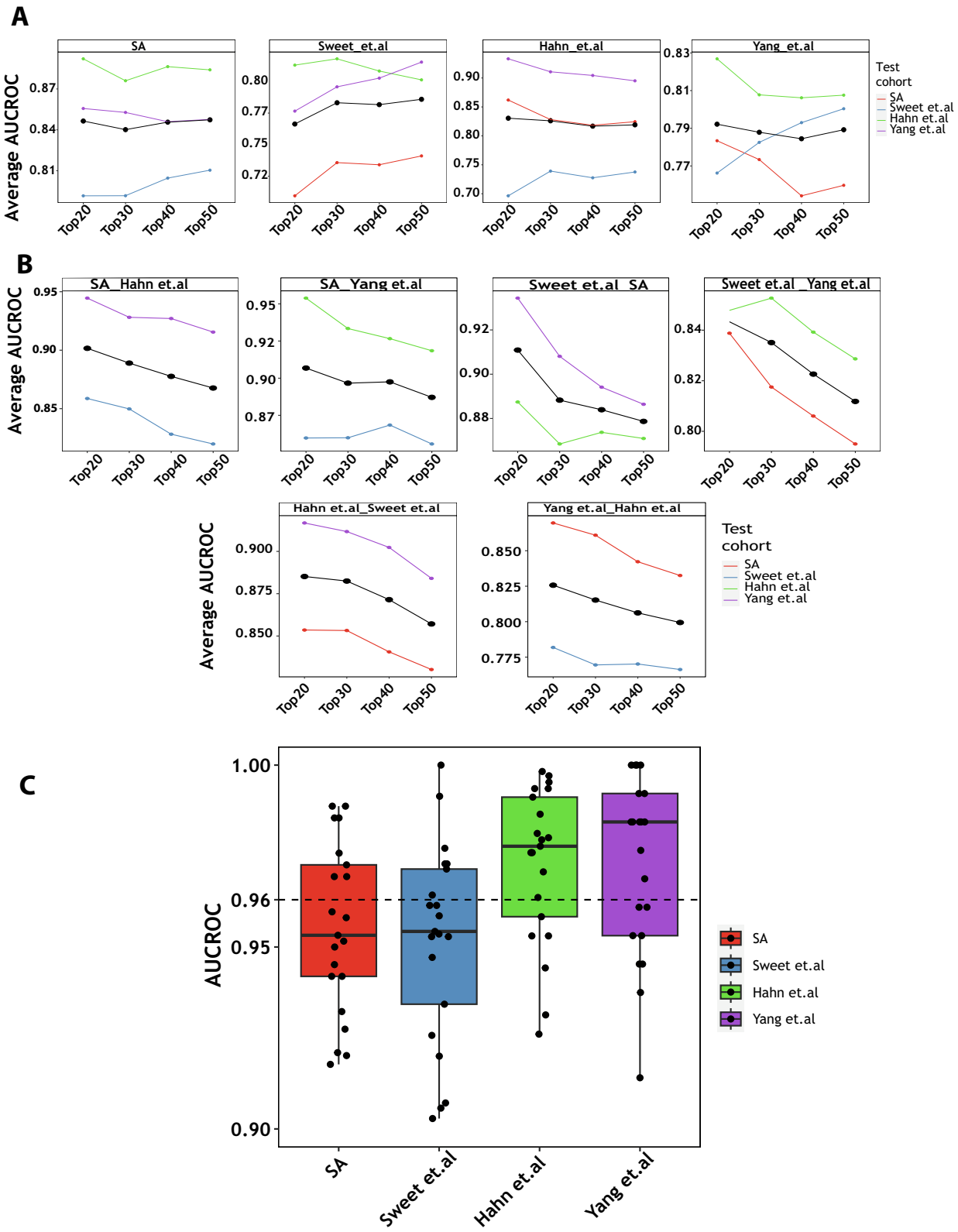

Fig. 5. Machine learning reveals gene signatures from SA and other HFrEF cohorts with strong stratification potential. (A) Each plot represents the mean AUCROC value for top 20 to top 50 gene signatures obtained by training from a single cohort. The AUCROC values were calculated based on the expression of the genes in independent cohorts used as test sets represented in different colors. The black line represents the mean of the mean AUCROC values for all the test cohorts for a given signature (B) Each plot represents the mean AUCROC for top 20 to top 50 gene signatures generated using a combination of two HFrEF cohorts. The mean AUCROC value for all the genes in the signature was calculated based on their expression in the test cohorts. Mean AUCROC values for each cohort is represented by different colors. The black line represents the mean of the mean AUROC values for all the test cohorts (C) AUCROC values for 21 gene core HF signature in different HFrEF datasets.

documented in heart failure [39,40]. Interestingly, like our study, the mRNA levels of various tubulins are often reported to be decreased in myocardium of HFrEF sample from multiple studies [16]. This discrepancy can be explained by autoinhibition of tubulin expression as suggested by reduced exonic RNA in comparison to intronic RNA levels in advanced heart failure [41].

Profiling of regulatory RNAs in cardiac tissue unravelled the complex regulation of genes in the pathogenesis of HF. There was altered expression of several known and novel lncRNAs in failing heart tissue. Among the annotated lncRNAs, many have already been implicated in cardiovascular diseases. *BANCR*, which is expressed in fetal cardiomyocytes and is responsible for cardiomyocyte migration and cardiac enlargement/dilation, was overexpressed [26]. The expression of *H19*, which is believed to be protective and has been found to be reduced in aortic stenosis, LVAD and failing heart samples [24], was also reduced in HF group in our study. LncRNAs like *SCIRT*, *GLIDR*, *SOCAR* that are known to be involved in proliferation or cancer [42–44] were also altered in the present study. The novel lncRNAs from our study that are enriched in HF group need to be taken up for further investigations to understand their role in heart failure pathophysiology and to assess their suitability as novel targets for therapy.

Changes in miRNA profile in HF further add to the complexity of disease pathophysiology as a single miRNA can regulate expression of multiple genes and a single gene can be regulated by multiple RNAs [20]. This enables miRNAs to regulate multiple cellular processes in a context dependent manner [20]. MiR-216 that was upregulated in our study has been implicated in cell proliferation and, under increased expression conditions, myofibroblast transdifferentiation [45]. Whether it has a protective or a deleterious role in heart, is not completely understood. The miR-216 null mice exhibit aggravated cardiac dysfunction due to impaired endothelial function in pressure-overload and myocardial infarction models of heart failure while in human cardiac fibroblasts, overexpression of miR-216 leads to profibrotic phenotype [45–47]. Similarly, miR-217, which is part of the same cluster, has also been shown to promote cardiac hypertrophy, fibrosis and dysfunction [48]. Our analyses also revealed multiple mRNA-miRNA pairs suggesting an overall miRNA-driven gene dysregulation in HF, which needs to be confirmed experimentally.

Plasma miRNA profiling of a small subset of HFrEF patients revealed many dysregulated miRNAs. However, only 12 of those showed dysregulation similar to that seen in heart, suggesting potential direct cardiac contribution to circulation. It is likely that the remaining dysregulated miRNA may be originating from other organs. This is not unexpected considering that severely impaired cardiac function leads to systemic pathological changes. In our study, miR-184 showed strong dysregulation in both cardiac tissue and plasma samples of HFrEF subjects. It has been linked with cellular apoptosis under oxidative stress and its oxidized form may worsen myocardial damage under ischemic conditions [49]. miR-184 levels in plasma have also been shown to be predictive of cardiac damage in a cohort of Anderson-Fabry disease [32]. While our data for miR-184 seem to be promising, its utility as a biomarker in heart failure needs to be confirmed in a larger patient cohort.

The ML framework successfully identified the core HF signature. The top genes from single training cohorts distinguished the HFrEF samples from controls based on only their expression in other cohorts. This ability to stratify based on only expression highlights its translatability as a better biomarker. Expanding the training set to two cohorts further improved the stratification ability of resulting signature in the remaining cohorts corroborating the effectiveness of ML based method in identifying the core HF gene signature. A 21 gene consensus HF stratifying signature from our study included conserved HF dysregulated genes like *MDK*, *SMOC2*, *KCND3*, *MYH6*, *OGN*, *LUM*, *CCDC69*, *HMOX2*, *FREM1* and others, which are known to be dysregulated across cohorts [16].

5. Limitations

Our study is the first integrated analysis of coding and non-coding transcriptome in the SA HFrEF patients. One of the limitations was the small size of the cohort but the findings have largely recapitulated core elements of HF pathophysiology. Earlier published meta analysis [16] indicates that the cohorts with smaller size are capable of reflecting the pathways or genes underlying HFrEF pathophysiology. We acknowledge that data from a larger patient pool would have been better in predicting the biomarker potential of multiple dysregulated genes and circulatory biomarkers. Another limitation was acquisition of normal samples. Our normal samples were obtained from autopsy patients. The timing of autopsy upon brain death may have had an impact on ischemia related genes. Indeed, *NPPA* was not differentially expressed in our data even though we observed a trend of higher expression in HF group. The ML based gene signature generated here shows that HFrEF samples in SA cohort show dysregulation of core HF genes and that the signature generated from SA cohort stratifies other cohorts well. Moreover, most of our findings are in line with the HF literature on other populations and also, with the consensus HF transcriptional signature generated by Flores et al. [16]. Lastly, the biological interpretation and hypotheses presented in this study are based on transcriptomic data and may not necessarily reflect the changes at protein level. Also, the cellular context is lost in bulk transcriptomic studies including this study.

6. Conclusions

To our knowledge, this is the first transcriptomic characterization of heart failure in South Asian subjects. Here, we present a detailed account of pathways underlying HF pathophysiology in combination with regulatory non-coding RNA component and a conserved HF gene signature using machine learning.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmccpl.2025.100282.

CRediT authorship contribution statement

Amit Anand: Writing – review & editing, Writing – original draft, Visualization, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Julius Punnen:** Writing – review & editing, Supervision, Conceptualization. **U.M. Nagamalesh:** Writing – review & editing, Supervision, Conceptualization. **Sabariya Selvam:** Writing – review & editing, Validation, Investigation. **Madhusudhan Bysani:** Writing – review & editing, Validation, Resources, Investigation. **Ramya Venkatesh:** Writing – review & editing, Validation, Investigation, Formal analysis. **Kriti Nawin:** Writing – review & editing, Validation, Investigation, Formal analysis. **Shilpa Garg:** Writing – review & editing, Validation, Investigation, Formal analysis. **Bagirath Raghuraman:** Writing – review & editing, Supervision, Conceptualization. **Varun Shetty:** Writing – review & editing, Supervision, Conceptualization. **Senthil Kumaran:** Writing – review & editing, Validation, Investigation. **Manoj Dokania:** Writing – review & editing, Investigation. **Pradeep Narayan:** Writing – review & editing, Supervision, Conceptualization. **Ankita Udwadia:** Writing – review & editing, Project administration. **Kushan Gunawardhana:** Writing – review & editing, Project administration. **David Gordon:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Manjunath Ramarao:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Lei Zhao:** Writing – review & editing, Supervision, Conceptualization. **Jyoti Gulia:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Project administration, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Funding

This work was funded by Bristol Myers Squibb.

Declaration of competing interest

AA, AU and GJ are employed by Bristol Myers Squibb. MR, DG, KG and LZ were employed by BMS during their work on this manuscript. Authors declare that they have no other competing interests.

Acknowledgments

We are thankful to clinical research group at Narayana Health for their support and Dr. Malini Venkata and Dr. Alben Sigamani for their help in establishing the clinical alliance.

Data availability

The transcriptomic data generated is available at EGA.

References

[1] Groenewegen A, Rutten FH, Mosterd A, Hoes AW. Epidemiology of heart failure. Eur J Heart Fail 2020;22(8):1342–56.
[2] Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats AJS. Global burden of heart failure: a comprehensive and updated review of epidemiology. Cardiovasc Res 2022;118(17):3272–87.
[3] Azizi Z, Golbus JR, Spaulding EM, Hwang PH, Ciminelli ALA, Lacar K, et al. Challenge of optimizing medical therapy in heart failure: unlocking the potential of digital health and patient engagement. J Am Heart Assoc 2024;13(2):e030952.
[4] Chaffin M, Papangeli I, Simonson B, Akkad AD, Hill MC, Arduini A, et al. Single-nucleus profiling of human dilated and hypertrophic cardiomyopathy. Nature 2022;608(7921):174–80.
[5] Liu Y, Morley M, Brandimarto J, Hannenhalli S, Hu Y, Ashley EA, et al. RNA-Seq identifies novel myocardial gene expression signatures of heart failure. Genomics 2015;105(2):83–9.
[6] Sweet ME, Cocciolo A, Slavov D, Jones KL, Sweet JR, Graw SL, et al. Transcriptome analysis of human heart failure reveals dysregulated cell adhesion in dilated cardiomyopathy and activated immune pathways in ischemic heart failure. BMC Genomics 2018;19(1):812.
[7] Yang K-C, Yamada KA, Patel AY, Topkara VK, George I, Cheema FH, et al. Deep RNA sequencing reveals dynamic regulation of myocardial noncoding RNAs in failing human heart and remodeling with mechanical circulatory support. Circulation 2014;129(9):1009–21.
[8] Hahn VS, Knutsdottir H, Luo X, Bedi K, Margulies KB, Haldar SM, et al. Myocardial gene expression signatures in human heart failure with preserved ejection fraction. Circulation 2021;143(2):120–34.
[9] Yi X, Jiang D-S, Feng G, Jiang X-j, Zeng H-L. An altered left ventricle protein profile in human ischemic cardiomyopathy revealed in comparative quantitative proteomics. Polish Heart J (Kardiologia Polska) 2019;77(10):951–9.
[10] Reitz Cristine J, Tavassoli M, Kim Da H, Shah S, Lakin R, Teng Allen CT, et al. Proteomics and phosphoproteomics of failing human left ventricle identifies dilated cardiomyopathy-associated phosphorylation of CTNNA3. Proc Natl Acad Sci 2023;120(19):e2212118120.
[11] van Heesch S, Witte F, Schneider-Lunitz V, Schulz JF, Adami E, Faber AB, et al. The translational landscape of the human heart. Cell 2019;178(1):242–260.e29.
[12] Li M, Parker BL, Pearson E, Hunter B, Cao J, Koay YC, et al. Core functional nodes and sex-specific pathways in human ischaemic and dilated cardiomyopathy. Nat Commun 2020;11(1):2843.
[13] Wang L, Yu P, Zhou B, Song J, Li Z, Zhang M, et al. Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. Nat Cell Biol 2020;22(1):108–19.
[14] Koenig AL, Shchukina I, Amrute J, Andhey PS, Zaitsev K, Lai L, et al. Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure. Nat Cardiovasc Res 2022;1(3):263–80.
[15] Kuppe C, Ramirez Flores RO, Li Z, Hayat S, Levinson RT, Liao X, et al. Spatial multi-omic map of human myocardial infarction. Nature 2022;608(7924):766–77.
[16] Flores ROR, Lanzer JD, Holland CH, Leuschner F, Most P, Schultz JH, et al. Consensus transcriptional landscape of human end-stage heart failure. J Am Heart Assoc 2021;10(7):e019667.
[17] Caporali A, Anwar M, Devaux Y, Katare R, Martelli F, Srivastava PK, et al. Non-coding RNAs as therapeutic targets and biomarkers in ischaemic heart disease. Nat Rev Cardiol 2024;21:556–73.
[18] Mably JD, Wang D-Z. Long non-coding RNAs in cardiac hypertrophy and heart failure: functions, mechanisms and clinical prospects. Nat Rev Cardiol 2024;21(5):326–45.
[19] Vegter EL, van der Meer P, de Windt LJ, Pinto YM, Voors AA. MicroRNAs in heart failure: from biomarker to target for therapy. Eur J Heart Fail 2016;18(5):457–68.
[20] Gargiulo P, Marzano F, Salvatore M, Basile C, Buonocore D, Parlati ALM, et al. MicroRNAs: diagnostic, prognostic and therapeutic role in heart failure—a review. ESC Heart Fail 2023;10(2):753–61.
[21] Urata H, Kinoshita A, Misono KS, Bumpus FM, Husain A. Identification of a highly specific chymase as the major angiotensin II-forming enzyme in the human heart. J Biol Chem 1990;265(36):22348–57.
[22] Nehme A, Zouein FA, Deris Zayeri Z, Zibara K. An update on the tissue renin angiotensin system and its role in physiology and pathology. J Cardiovasc Dev Dis 2019;6(2):14.
[23] Wu L, Gao L, Zhang D, Yao R, Huang Z, Du B, et al. C1QTNF1 attenuates angiotensin II-induced cardiac hypertrophy via activation of the AMPKa pathway. Free Radic Biol Med 2018;121:215–30.
[24] Viereck J, Bührke A, Foinquinos A, Chatterjee S, Kleeberger JA, Xiao K, et al. Targeting muscle-enriched long non-coding RNA H19 reverses pathological cardiac hypertrophy. Eur Heart J 2020;41(36):3462–74.
[25] Li J, Liu W, Peng F, Cao X, Xie X, Peng C. The multifaceted biology of lncR-Meg3 in cardio-cerebrovascular diseases. Front Genet 2023;14.
[26] Wilson KD, Ameen M, Guo H, Abilez OJ, Tian L, Mumbach MR, et al. Endogenous retrovirus-derived lncRNA BANCR promotes cardiomyocyte migration in humans and non-human Primates. Dev Cell 2020;54(6):694–709.e9.
[27] Deng H, Ouyang W, Zhang L, Xiao X, Huang Z, Zhu W. LncRNA GASL1 is downregulated in chronic heart failure and regulates cardiomyocyte apoptosis. Cell Mol Biol Lett 2019;24(1):41.
[28] Lu W, He X, Su L, Miao J. Long noncoding RNA-CERNA1 stabilized atherosclerotic plaques in apolipoprotein E−/− mice. J Cardiovasc Transl Res 2019;12(5):425–34.
[29] Tang Y, Li H, Chen C. Non-coding RNA-associated therapeutic strategies in atherosclerosis. Front Cardiovasc Med 2022;9.
[30] Zhou H, Song W-H. LncRNA HCG11 accelerates atherosclerosis via regulating the miR-224-3p/JAK1 Axis. Biochem Genet 2023;61(1):372–89.
[31] Li H, Fan J, Yin Z, Wang F, Chen C, Wen Wang D. Identification of cardiac-related circulating microRNA profile in human chronic heart failure. Oncotarget 2015;7(1).
[32] Salamon I, Biagini E, Kunderfranco P, Roncarati R, Ferracin M, Taglieri N, et al. Circulating miR-184 is a potential predictive biomarker of cardiac damage in Anderson–Fabry disease. Cell Death Dis 2021;12(12):1150.
[33] Fan X, Zhan E, Yao Y, Zhang R, Sun Y, Tian X. MiR-599 protects cardiomyocytes against oxidative stress-induced pyroptosis. Biomed Res Int 2021;2021:3287053.
[34] Agarwal V, Bell GW, Nam J-W, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. eLife 2015;4:e05005.
[35] Condrat CE, Thompson DC, Barbu MG, Bugnar OL, Boboc A, Cretoiu D, et al. miRNAs as biomarkers in disease: latest findings regarding their role in diagnosis and prognosis. Cells 2020;9(2):276.
[36] Santos-Zas I, Lemarié J, Zlatanova I, Cachanado M, Seghezzi J-C, Benamer H, et al. Cytotoxic CD8+ T cells promote granzyme B-dependent adverse post-ischemic cardiac remodeling. Nat Commun 2021;12(1):1483.
[37] Pankuweit S, Ruppert V, Jónsdóttir Þ, Müller H-H, Meyer T. The HLA class II allele DQB1*0309 is associated with dilated cardiomyopathy. Gene 2013;531(2):180–3.
[38] Meder B, Rühle F, Weis T, Homuth G, Keller A, Franke J, et al. A genome-wide association study identifies 6p21 as novel risk locus for dilated cardiomyopathy. Eur Heart J 2013;35(16):1069–77.
[39] Caporizzo MA, Chen CY, Prosser BL. Cardiac microtubules in health and heart disease. Exp Biol Med 2019;244(15):1255–72.
[40] Uchida K, Scarborough EA, Prosser BL. Cardiomyocyte microtubules: control of mechanics, transport, and remodeling. Annu Rev Physiol 2022;84(Volume 84, 2022):257–83.
[41] Phyo SA, Uchida K, Chen CY, Caporizzo MA, Bedi K, Griffin J, et al. Transcriptional, post-transcriptional, and post-translational mechanisms rewrite the tubulin code during cardiac hypertrophy and failure. Front Cell Dev Biol 2022;10.
[42] Zagorac S, de Giorgio A, Dabrowska A, Kalisz M, Casas-Vila N, Cathcart P, et al. SCIRT lncRNA restrains tumorigenesis by opposing transcriptional programs of tumor-initiating cells. Cancer Res 2021;81(3):580–93.
[43] Li S, Wang Y, Cao Q, Li H, Zhao Z, Wei B, et al. GLIDR promotes the aggressiveness progression of prostate cancer cells by sponging miR-128–3p. Pathol Res Pract 2023;242:154343.
[44] Guo Q, Wang L, Zhu L, Lu X, Song Y, Sun J, et al. The clinical significance and biological function of lncRNA SOCAR in serous ovarian carcinoma. Gene 2019;713:143969.
[45] Qu C, Liu X, Ye T, Wang L, Liu S, Zhou X, et al. miR-216a exacerbates TGF-β-induced myofibroblast transdifferentiation via PTEN/AKT signaling. Mol Med Rep 2019;19(6):5345–52.
[46] Juni RP, Kocken JMM, Abreu RC, Ottaviani L, Davalan T, Duygu B, et al. MicroRNA-216a is essential for cardiac angiogenesis. Mol Ther 2023;31(6):1807–28.
[47] Tao J, Wang J, Li C, Wang W, Yu H, Liu J, et al. MiR-216a accelerates proliferation and fibrogenesis via targeting PTEN and SMAD7 in human cardiac fibroblasts. Cardiovasc Diagn Ther 2019;9(6):535–44.
[48] Nie X, Fan J, Li H, Yin Z, Zhao Y, Dai B, et al. miR-217 promotes cardiac hypertrophy and dysfunction by targeting PTEN. Mol Ther - Nucleic Acids 2018;12:254–66.
[49] Wang J-X, Gao J, Ding S-L, Wang K, Jiao J-Q, Wang Y, et al. Oxidative modification of miR-184 enables it to target Bcl-xL and Bcl-w. Mol Cell 2015;59(1):50–61.