Genomic analysis of circular RNAs in heart

Kunzhe Dong
Augusta University
Xiangqin He
Augusta University
Huabo Su
Augusta University
David J.R. Fulton
Augusta University
Jiliang Zhou (JIZHOU@augusta.edu)
Augusta University https://orcid.org/0000-0002-2375-4636

Research article

Keywords: circular RNA, dilated cardiomyopathy, read-through circRNA, miRNA sponge

Posted Date: August 13th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-50151/v1

Version of Record: A version of this preprint was published on November 7th, 2020. See the published version at https://doi.org/10.1186/s12920-020-00817-7.

Abstract

Background: Heart failure (HF) is a leading cause of human morbidity and mortality. Circular RNAs (circRNAs) are a newly discovered class of RNA that have been found to have important physiological and pathological roles. In the current study, we *de novo* analyzed existing whole transcriptome data from 5 normal and 5 dilated cardiomyopathy (DCM) human heart samples and compared the results with circRNAs that have been previously reported in human, mouse and rat hearts.

Results: Our analysis identifies a list of cardiac circRNAs that are reliably detected in multiple studies. We have also defined the top 30 most abundant circRNAs in healthy human hearts which include some with previously un-appreciated cardiac roles such as circHIPK3_11 and circTULP4_1. We further found that many circRNAs are dysregulated in DCM, particularly transcripts originating from DCM-related gene loci, such as *TTN* and *RYR2*. In addition, we predict the potential of cardiac circRNAs to sponge miRNAs that have reported roles in heart disease. We found that circALMS1_6 has the highest potential to bind miR-133, a regulator of cardiac remodeling. Interestingly, we detected a novel class of circRNAs, referred to as read-though (rt)-circRNAs which are produced from exons of two different neighboring genes. Specifically, rt-circRNAs from *SCAF8* and *TIAM2* were observed to be dysregulated in DCM and both have the potential to sponge multiple heart disease-related miRNAs.

Conclusions: In summary, this study provides a valuable resource for exploring the function of circRNAs in human heart disease and establishes a functional paradigm for identifying novel circRNAs in other tissues.

Background

Heart failure (HF) is a leading cause of human morbidity and mortality worldwide. Despite advances in the management of HF, it remains a tremendous and growing public health burden in an aging population with an estimated prevalence of > 37.7 million individuals globally [1]. Dilated cardiomyopathy (DCM), characterized by the dilation and impaired contraction of the left or both ventricles, is a primary cause of HF and sudden cardiac death [2]. Understanding the etiology of DCM is vitally important in the search for better therapeutic approaches for HF.

In addition to non-genetic causes, such as hypertension, inflammation, and toxins, genetic factors are increasingly recognized for their role in HF susceptibility [3, 4]. Currently, mutations in more than 100 genes have been linked to DCM [5]. However, these DCM relevant gene variants are only detected in ~37% of DCM patients [6]. Therefore, it is likely that other unknown genetic factors await discovery. The aberrant expression of different types of RNA transcripts, including those transcribed from protein coding genes [7-9], long non-coding genes [10, 11] and miRNAs [12-14], have been observed in DCM.

Circular RNAs (circRNAs) are a novel class of RNA that are generated by the covalent ligation of the 5’ and 3’ ends of a single-stranded RNA transcript to form a circular transcript [15]. CircRNAs are stable and resistant to RNase R, widespread throughout the plant and animal kingdoms and conserved across

multiple species. Furthermore, circRNAs can be expressed in a tissue-specific manner [16, 17]. These properties make circRNAs excellent candidates as potential biomarkers for the diagnosis of disease and therapeutic intervention. Emerging evidence has demonstrated that circRNAs play important roles in various physiological and pathological processes, such as myogenesis [18], phenotypic modulation of smooth muscle cells [19, 20], atherosclerosis [21, 22], and cancer [23-25].

Although several efforts have been made to characterize the circRNAs expressed in the heart [26-29], our understanding of the importance of cardiac circRNAs is still very limited and several fundamental questions remain to be addressed. First, there is a large discrepancy between the presence and abundance of cardiac circRNAs that have been identified in previous studies [26, 27], which limits any interpretation of their importance. To pinpoint truly important cardiac circRNAs, a cross-reference analysis between multiple independent studies is required. Second, it has been documented that circRNAs function primarily as molecular sponges to competitively bind miRNAs or encode small peptides. While the coding ability of cardiac circRNAs have been investigated recently by analyzing translational Ribo-seq data from human hearts [30], the potential of cardiac circRNAs to function as miRNA sponges remains incompletely understood. Furthermore, read through-circRNAs (rt-circRNAs), a recently discovered type of circRNAs that are produced from the exons of two adjacent genes on the same strand, have been implicated in both normal and tumor tissues [28], yet whether this type of rt-circRNAs are expressed in the human heart remains to be identified.

To address these questions, we *de novo* analyzed publicly available RNA-seq datasets from human hearts. These data sets were of sufficient sample size (5 normal controls and 5 DCM patients) [10], to enable us to investigate the general landscape of human cardiac circRNAs under homeostatic (normal) and disease settings. By integrating the analysis of circRNAs identified in human hearts with that of mouse and rat hearts in previous studies [26, 27, 29], we: (i) characterized the general features of circRNAs in human hearts; (ii) determined the key circRNAs expressed in healthy hearts as defined by the top 30 most abundant circRNAs; (iii) identified circRNAs that are dysregulated in DCM hearts; (iv) predicted the potential of high-confidence cardiac circRNAs to act as a sponge for previously identified heart disease-related miRNAs; Lastly, we identified a list of rt-circRNAs, especially those from the *SCAF8* and *TIAM2* gene loci that are dysregulated in DCM and show strong potential of being miRNA sponges. Our work identifies a set of highly reliable circRNAs with potentially important roles in maintaining cardiac health which is a valuable resource for exploring the function of circRNAs in human heart disease and we establish a paradigm for identifying novel circRNAs in other tissues.

Results

General features of circRNAs in both healthy and DCM human hearts

A *de novo* analysis of public RNA-seq data sets generated using left ventricular tissue obtained from 5 healthy individuals and 5 DCM patients [10] was performed. These data sets were selected because they

were generated using a ribosome-depletion library, which is the gold standard method to identify back-spliced junctions for circRNAs detection [31]. We first removed low-quality reads and then utilized Bowtie2 [32] to align clean reads with the human reference genome. Reads continually aligned to the genome were used for linear gene expression analysis and the remaining un-mapped reads were further used as input for find_circ [33] to detect both linear and back-spliced junctions (Additional file 1: Figure S1). We first performed a quality control validation of this data by performing principal component analysis (PCA) based on the expression of linear genes with raw count > 10 in at least 5 samples. PCA analysis confirmed that samples from each group were clustered respectively as expected (Additional file 2: Figure S2A). We further examined the expression of a marker of HF, atrial natriuretic peptide α (*NPPA*). Consistent with the reported HF phenotype [34], our analysis revealed that the expression of *NPPA* was significantly increased in DCM specimens (Additional file 2: Figure S2B). Taken together, these initial analyses confirm the quality of this RNA-seq dataset and its suitability for the further analysis of circRNAs.

With a threshold cutoff of ³ 2 unique back-spliced reads present in at least one sample, we identified a total of 20,214 circRNAs in the 10 samples (Figure 1A and Additional file 3: Table S1). The identified circRNAs originated primarily from the coding-sequence (CDS) regions of exons of single linear genes (67.2%), followed by exons spanning coding and untranslated regions (UTR) (UTR-CDS) (11.6%) (Figure 1B). Notably, a total of 252 circRNAs (1.2%) were assigned to the exons of two adjacent genes on the same strand (Additional file 4: Table S2), likely representing a novel class of circRNAs termed as rt-circRNA [23]. In particular, the genes *MYH6* and *MYH7* produced the largest number (n = 22) of circRNAs, which were also observed in human, mouse and rat hearts as reported by two earlier studies (Additional file 5: Figure S3A-C) [26, 27]. In addition, these *MYH6/MYH7*-derived circRNAs were highly abundant and conserved in hearts of human, mouse and rat (Additional file 5: Figure S3D). However, close inspection of the sequence between the exons where they are back-spliced suggests that the back-spliced junctions of these circRNAs are probably artifactual due to read misalignment caused by the high sequence homology between *MYH6* and *MYH7* (Additional file 5: Figure S3E). This result prompted us to employ more stringent criteria to filter circRNAs originating from two neighboring genes with highly homological sequences. Using this revised filtering strategy, we identified 110 putative rt-circRNAs. Among them, *SCAF8* and *TIAM2* contain the largest number of rt-circRNAs (n = 4), all of which were also previously identified in the human heart (Figure 1C).

We next examined whether a correlative relationship between the expression of circRNAs and their parent linear genes. We found that there was a significantly higher correlation between circRNAs and their linear parent genes compared to that between circRNAs and randomly chosen linear genes (Figure 1D), suggesting that the generally high-abundance circRNAs tend to have high-abundance linear counterparts. We further found that the length of the majority of identified circRNAs was around 300-400 bp (Figure 1E). In accordance with previous studies [35-37], the most common splice accepting circularized exon was exon 2 (Figure 1F). Furthermore, we found that the majority of circRNA-producing genes generated 1 circRNA (41.7%, 2,231/5,350) and only 6.0% (n = 319) produced more than 10 circRNAs (Figure 1G). The top two genes producing the largest number of circRNAs were *TTN* (Titin) and *RYR2* (Ryanodine receptor

2), which generated 197 and 173 circRNAs, respectively (Additional file 6: Figure S4A-B). Because the *TTN* and *RYR2* genes have the largest number of exons in human genome, we hypothesize that the number of circRNAs produced is positively correlated to the number of exons within their parent genes. Indeed, a proportional increase in the number of circRNAs was observed to correlate with the number of exons per gene (Additional file 6: Figure S4C). Furthermore, we found 17,733 identified circRNAs, in total, contain exons. Among these, single exonic circRNAs were found in 8.4% (1,494/17,733) while the rest of circRNAs were encoded by more than two exons (Additional file 6: Figure S4D).

Next we compared the circRNAs that we identified with others that were previously identified in human hearts [26, 27]. We found that 10,990 circRNAs (54.4%) were also detected by at least one of the two studies, and 6,225 (30.8%) were detected by all three studies (Figure 1H). In addition, 2,135 and 2,297 circRNAs have orthologs in the mouse and rat, respectively, and 1,203 were conserved in all three species (Figure 1I). We found that 7,585 and 6,211 circRNAs were specifically identified in normal and DCM hearts, respectively, and that 6,418 were identified in both (Figure 1J). Further analysis revealed that the majority of circRNAs were only detected in one human sample, while only 729 circRNAs were present in all the 10 samples (Figure 1K), suggesting that many circRNAs exhibit a high degree of variation in expression among human individuals, likely due to low fidelity and/or the low expression levels of most circRNAs. Collectively, these results indicate that the expression of circRNAs, including rt-circRNAs, is widespread in the human heart with a certain degree of conservation between species.

Identification of the most abundant circRNAs in healthy human heart

CircRNAs that are abundantly expressed in normal heart likely contribute to the maintenance of cardiac homeostasis. We therefore focused on the top 30 circRNAs with the highest average expression level across the 5 healthy heart samples (Figure 2A). The top 2 most abundant circRNAs (*circSLC8A1_11* and *circSLC8A1_12*) originate from the *SLC8A1* (sodium/calcium exchanger 1) gene. While most of the top circRNAs were also identified in the studies of Tan et al. [27] and Werfel et al. [26], which found high expression levels of circRNAs such as circSLC8A1_12, circHIPK3_1 and circALPK2_2 in healthy hearts, some exceptions were observed. These include circSLC8A1_11 and circYY1AP1_3 that were not detected in study of Tan et al. [27], as well as circZNF91_2 and circPCMTD1_6 whose expression was identified to be abundant in the heart in only one of the previous studies (Figure 2A). Comparison of the results obtained from different species revealed that half of these highly abundant circRNAs in normal human hearts were conserved in mouse or rat hearts, but only 2 of them including circSLC8A1_11 and circHIPK3_1 exhibited high levels of expression in all species, and 2 of them including circQKI_4 and circTULP4_1 were abundantly expressed in human and mouse hearts, but not rat. The expression levels of the remaining conserved circRNAs were moderate or low in rodent hearts. We then calculated the correlation between the abundance of top circRNAs and their host genes and found a highly positive correlation for most of these circRNAs (Additional file 7: Figure S5A), including the top 3 most abundant ones (i.e., circSLC8A1_11, circSLC8A1_12 and circHIPK3_1) (Figure 2B). In contrast, circSETD3_3,

circMB_1 and circCORO1C_1, together with some others (Additional file 7: Figure S5B), appeared to be independently transcribed with their linear parent genes (Figure 2C). By integrating results from multiple studies, our analyses define the most abundant circRNAs in the healthy human heart.

Identification of dysregulated circRNAs in DCM

We next performed differential analysis to identify circRNAs that are dysregulated in DCM. Given the observation that circRNA abundance was not universally distributed across samples (Figure 1K), we restricted our analysis to 2,461 circRNAs which were considered to be high-confidence circRNAs as they had ³ 2 reads spanning the back-spliced junction in ³ 4 samples in at least one group (Additional file 8: Table S3). Using a threshold of $p < 0.05$ and fold change (FC) > 2, a total of 392 circRNAs, including 101 up-regulated and 291 down-regulated circRNAs, were identified to be differentially expressed in DCM hearts as compared to normal hearts (Figure 3A and Additional file 8: Table S3). These differentially expressed circRNAs were derived from 290 unique host genes. Pathway analysis revealed that these host genes are significantly over-represented in pathways associated with diseases of the heart, including arrhythmogenic right ventricular cardiomyopathy (ARVC), hypertrophic cardiomyopathy, and DCM (Figure 3B). Interestingly, 2 rt-circRNAs originating from *SCAF8* and *TIAM2*, including SCAF8_e4:TIAM2_e1 and SCAF8_e4:TIAM2_e2, were remarkably downregulated in DCM (Figure 3C). Many of the significantly down- and up-regulated circRNAs were positively correlated with the expression of their host genes such as circFBLN1_5 with *FBLN1* and circNLGN1_1 with *NLGN1* (Figure 3D and E). However, the expression of some circRNAs in DCM was independent of changes in their host gene expression (Figure 3F). Examples of this discordant relationship included circABCC1_9, which was significantly up-regulated in DCM, while the expression of its host gene mRNA had no obvious change in DCM. Together, these results indicate that dysregulation of circRNAs in DCM tend to originate from heart disease-related gene loci, which include the rt-circRNAs from *SCAF8* and *TIAM2*.

Screening circRNAs for potential as miRNA sponges

To screen for cardiac circRNAs with the potential to act as miRNA sponges, we analyzed the sequences of each high-confidence circRNAs for the presence of putative binding sites for 50 miRNAs that have established roles in heart disease (Additional file 9: Table S4). In total, 142 circRNAs were predicted to harbor at least 5 predicted binding sites for at least one heart disease-related miRNA (Additional file 10: Table S5). The top 32 circRNAs that contained more than 10 putative binding sites for at least one miRNA are listed in Figure 4A. Notably, 23 of these are derived from one gene, *TTN* (Figure 4A). In particular, circALMS1_6 had the strongest potential to function as miRNA sponge, containing 28 and 20 putative binding sites for miR-133a-5p and miR-133a-3p, respectively (Figure 4B). Another notable finding is that the rt-circRNA, circSCAF8_e4:TIAM2_e2, was predicted to be capable of sponging multiple miRNAs including miR-150-5p and miR-17 that has been reported to be induced in the rat heart of myocardial ischemia reperfusion injury [38] and in the mouse heart of chronic hypoxia-induced pulmonary

hypertension mouse model [39]. (Figure 4C). Interestingly, some of the miR-17 downstream targets, such as *SCN2B* [40, 41], *F2R* [42], *KCNB1* [43] and *EGLN3* [44] that have been implicated in heart dysfunction, were down-regulated in the DCM hearts revealed by expression analysis of linear RNAs, likely through the unleashed miR-17 that resulted from the down-regulation of circSCAF8_e4:TIAM2_e2 in DCM (Figure 4C). Collectively, these analyses provide a systematic evaluation of the potential of cardiac circRNAs to sponge miRNAs and provide a resource for studying cardiac miRNA inhibitors.

Experimental validation of identified circRNAs

Due to lack of availability of human samples, we next sought to validate the conserved circRNAs with mouse orthologs that we identified above using mouse heart samples. We designed divergent primers for the murine orthologs of 6 identified human cardiac circRNAs (Additional file 11: Table S6) and performed PCR amplification using mouse heart cDNA as template. Mouse heart genomic DNA was used as a negative control and mouse *Yap1* gene primers were used as an internal control for the PCR amplification of genomic DNA. All 6 of the selected circRNAs were successfully amplified in heart cDNA but not in genomic DNA (Figure 5A). Sanger sequencing of the PCR products verified back-spliced junctions (Figure 5B-G). Taken together, all the circRNAs examined were confirmed to be bona fide, suggesting the reliability of our *in silico* circRNA identification.

Discussion

Herein, we have analyzed rRNA-depleted RNA sequencing datasets from 10 human hearts. By integrating circRNAs that were previously identified in human, mouse and rat hearts [26, 27, 29], we have systematically interrogated the global landscape of circRNAs in human heart and have revealed the most abundant circRNAs in healthy hearts, and those that are dysregulated in DCM hearts. Furthermore, we have investigated the potential of human cardiac circRNAs to act as sponges for heart disease-related miRNAs and have generated a list of rt-circRNAs expressed in human heart. Our study adds new information to the previous annotated circRNAs and provides a valuable resource for further functional study of circRNAs in the human heart.

Two earlier studies found that circRNAs are expressed in the human heart and identified more than 15,000 circRNAs [26, 27]. The number of circRNAs identified in our study (n = 20,214) is larger than that of previous studies. This could be partially explained by the differences in the genetic background of the human subjects, the algorithms employed and the thresholds used for detecting and filtering circRNAs, and sample size. When comparing the degree of overlap of circRNAs identified in previous studies, congruency was lower than expected despite the use of the same tissue (heart) and species (human). For example, only about half of the circRNAs reported by the two previous studies are in common [26, 27] (Figure 1H). This observation suggests that either the expression pattern of circRNAs varies significantly among different individuals, or that the identified circRNAs in the limited samples were artifactual, which underscores the need to identify a list of reliable cardiac circRNAs by integrating multiple independent

studies. Our comparative analysis revealed a total 10,990 cardiac circRNAs that are reproducible in at least two independent studies (Figure 1H), which represent a bona fide category of cardiac circRNAs that can be functionally explored in the future.

Our analysis revealed that a significant number of cardiac circRNAs are derived from two different genes, which were also reported by previous studies of circRNAs in heart [26, 27] and brain [16, 33, 35]. Due to the fact that many such circRNAs are derived from neighboring genes with highly homologous sequences, a common practice in circRNA detection is to discard circRNAs from multiple genes [30, 35]. However, a recent breakthrough study showed that a significant portion of circRNAs indeed originate from the exons of two adjacent genes on the same strand. These circRNA are referred to as rt-circRNAs [28] and suggests that circRNAs originated from different genes should not simply be dismissed as false positives, especially in studies with a major focus on circRNA discovery. One striking result of this study is that a number of circRNAs are identified to be produced from the exons of *MYH6* and *MYH7*, which were also detected, and display high abundance, in hearts of human, mouse and rat by two previous studies (Additional file 5: Figure S3A-D) [26, 27], albeit using different algorithms and cohorts for circRNA detection. This is particularly interesting because these two genes play important roles in healthy and diseased hearts [45]. In-depth examination of sequences reveals that these circRNAs are derived from highly homologous region of *MYH6* and *MYH7*, suggesting that these circRNAs are likely artifacts due to the read misalignment as concerned by previous studies [30, 35]. Nevertheless, after filtering out additional circRNAs from the homologous regions of two nearby genes, we identified a total of 110 potentially bona fide rt-circRNAs (Additional file 4: Table S2). Especially, *SCAF8* and *TIAM2* gene produce the largest number of rt-circRNAs (Figure 1C). Notably, all of the identified putative rt-circRNAs are not conserved in mouse or rat, suggesting that they are human-specific and recently evolved novel non-coding molecules. Future efforts are needed to validate these rt-circRNAs in greater detail and explore their functional roles in the heart.

Our analysis yielded a list of the top 30 most abundant circRNAs expressed in normal hearts (Figure 2A). All of the identified top circRNAs, except for circSPHKAP_1 and circRMST_3, are within the top 100 most abundant circRNAs in human or mouse hearts as revealed by at least one previous study (Figure 2A) [26, 27], suggesting a high degree of fidelity of this top circRNA list. In previous studies, discordant findings have been observed for some of these top circRNAs with regard to their presence and abundance [26, 27], such as circTULP4_1 that is absent in the study of Tan et al. [27] but is abundantly expressed in human and mouse heart. Our cross-reference analysis of multiple independent studies helps to reduce confusion over the importance of the identified circRNAs. The function of one of the most abundant circRNAs, circSLC8A1_11, in heart function has been assessed by a recent study [46]. CircHIPK3_1, formed from the circularization of the second exon of *HIPK3*, is another conserved circRNA that is highly expressed in both human and murine heart. An earlier study has shown that circHIPK3_1 plays a critical role in promoting retinal vascular dysfunction in diabetes mellitus [47]. In future studies, it will be interesting to explore the function of circHIPK3_1, as well as some of the other top candidates we have identified such as circTULP4_1 (Figure 2A), in maintaining cardiac hemostasis. Furthermore, half of the most abundant human circRNAs are conserved in mouse and/or rat. However, only several of these, including

circSLC8A1_11, circHIPK3_1, circQKI_4 and circTULP4_1 are observed to be highly expressed in murine heart as well, indicating differences in the circRNA landscape between humans and rodents. In addition, the expression levels of the three human-specific circRNAs including circMB_1, circCORO1C_1 and circSETD3_3 tends to be independent of their host genes, suggesting potentially important roles of these circRNAs in human heart.

Differential analysis revealed that most of the aberrant circRNAs were down-regulated in DCM as compared to normal hearts (Figure 3A), a feature that is similar to observations in cancer [28, 48]. In cancer, decreased expression of circRNAs was speculated to result from the dilution of circRNA upon cancer cell division [28, 48]. However, this mechanism cannot account for the reduced expression in DCM hearts due to the fact that adult cardiac myocytes have a very limited replicative capacity and extensive cell death leads to a reduction in myocyte number in DCM [49]. Another interesting finding is that many of the abundant circRNAs originate from already-established cardiovascular disease-related gene loci, as indicated by the results of pathway analysis of their host genes (Figure 3B). For example, four circRNAs were generated from the *TTN* locus, including circTTN_34, circTTN_52, circTTN_70 and circTTN_132 and were significantly down-regulated in DCM as compared to normal hearts (Additional file 5: Table S3). These data are consistent with an earlier study that reported on the dysregulation of circRNAs synthesized from the *TTN* gene in DCM [50]. The *TTN* gene encodes the protein titin, which is the largest known protein expressed in the heart. The *TTN* gene was also shown to generate the largest number of circRNAs in the heart (Additional file 6: Figure S4A). Another interesting gene which generated two significantly down-regulated circRNAs (circRYR2_71 and circRYR2_95) in DCM, was the *RYR2* gene. RYR2 is the major Ca^{2+} release channel on the sarcoplasmic reticulum and plays a critical role in regulating excitation-contraction coupling and Ca^{2+} homeostasis in the heart [51]. Sequence variations of these two genes have been linked to DCM and resulting in altered amino acid sequence and disruption of protein structure and function [52, 53]. There was no evidence to support the presence of mutations in *TTN* and *RYR2* in the 5 DCM patients analyzed. Furthermore, RNA-seq analysis revealed that their expression did not change at the mRNA level as compared to controls. Thus the downregulation of circRNAs observed from the *TTN* and *RYR2* genes in DCM likely represent a novel mechanism that awaits further identification. We observed that multiple circRNAs were differentially expressed in DCM, with no obvious change in the parent gene expression (Figure 3D and F), confirming previous findings that the expression of circRNAs can be independent of their parent genes [54].

One of the best documented functions of circRNAs is to sponge miRNAs [55] and the regulatory role of circRNAs via the circRNA-miRNA-mRNA axis has been increasingly implicated in human diseases [56]. Our results in predicting the ability of circRNAs to bind heart disease-related miRNAs provides a guideline for further exploration of the regulatory role of cardiac circRNAs (Additional file 10: Table S5). Notably, circRNAs produced from the *TTN* gene show a strong potential to bind miRNAs, suggesting that the function of the cardiac *TTN* gene may have added complexity. In addition, circALMS1_6 has strong potential to sponge miR-133, a well-known miRNA that is abundantly expressed in skeletal muscle and cardiomyocytes [57]. Given the important roles of miR-133 in regulating cardiac hypertrophy, it will be

interesting to determine whether circALMS1_6 influences cardiac hemostasis via inhibition of miRNAs. Furthermore, we identified a rt-circRNA, circSCAF8_e4:TIAM2_e2, with the potential to bind multiple miRNAs including miR-17 that have been shown to be induced in diseased hearts [38, 39] (Figure 4C). In DCM hearts, the expression of circSCAF8_e4:TIAM2_e2 is decreased (Figure 3C). It is therefore possible that the down-regulation of this rt-circRNA contributes to the up-regulation of miR-17 via releasing the sequestration of miR-17 interacted with, thereby inhibiting the down-stream targets of miR-17. Supporting this hypothesis, several miR-17 targets with known cardiac function are found to be down-regulated in DCM heart (Figure 4D). One interesting target is *SCN2B* that encodes voltage-gated sodium channel b2-subunits. The sequence variations of *SCN2B* have been associated with human cardiac arrhythmias [41] and genetic deletion of *Scn2b* in mice leads to ventricular and atrial arrhythmias [40]. Our study suggests a possible regulatory role of circSCAF8_e4:TIAM2_e2-miR-17-*SCN2B* axis in heart and expands our understanding of the function of the newly discovered rt-circRNAs.

Conclusions

In conclusion, our study provides a list of putative key circRNAs that are expressed in normal and DCM human hearts that will be helpful to guide future functional studies. Given their abundance and multiple functional roles, circRNAs, especially rt-circRNAs, are a group of emerging players in the regulation of human heart diseases.

Methods

Identification of circRNAs

RNA-seq raw reads from the left ventricle of 5 DCM patients and 5 normal controls were obtained from SRA database (SRP108571) [10]. Low-quality reads were first removed by Trimmomatic 3.0 [58]. The clean reads were then mapped to the GRCh38 human genome reference by Bowtie2 [32]. For the linear RNA analysis, the raw count of each gene was calculated by using HTSeq [59] and Fragments Per Kilobase of exon per Million fragments mapped (FPKM) values were calculated from raw counts to quantify expression of genes [60]. Next, circRNAs were identified following the instruction of find_circ [33]. Briefly, by using SAMtools the unmapped reads were retrieved [61]. Then these unmapped reads were cut off 20 mers from both ends and mapped to the human genome independently in order to find unique anchor positions. Anchors that aligned in the reversed orientation suggest a back-spliced junction. CircRNAs with ≥ 2 unique back spliced reads in at least one sample were saved for analysis of circRNA features. The expression of circRNAs was normalized to Spliced Reads per Billion Mapped Reads (SRPBM). For identification of the differentially expressed circRNAs between normal and DCM samples, only those high confidence circRNAs, with ≥ 2 unique reads in at least four samples from at least one group were used. Differentially expressed circRNAs were identified with edgeR package based on a negative binomial distribution model [62]. CircRNAs with fold change > 2 and p value < 0.05 were considered significant. Pathway analysis was performed using Metascape [63].

Comparison with other studies

Results from three previous studies [26, 27, 29], which identified circRNAs in human, mouse or rat hearts, were used in comparison with our results. The R package liftOver (https://www.bioconductor.org/packages/release/workflows/html/liftOver.html) was used to convert start and end coordinates (hg19, mm9, mm10 and rn5) to hg38 first and then the converted coordinates were used for cross-reference and cross-species comparison.

Analysis of circRNAs from two different genes

The 50 bp sequences from each side of the back-spliced junction of the cricRNAs arising from the exons of two different genes (gene 1 and gene 2), were retrieved and used as query sequences to perform BLAST analysis against human cDNA sequences downloaded from ENSEMBL database with default parameters. If any of the two sequences could be aligned to *both* gene 1 and gene 2, the circRNA was considered as possible false positive due to the sequence homologue of gene 1 and gene 2 (as illustrated in Additional file 5: Figure S3E). If the sequence from each side of the back-spliced junction returned BLAST hits only matching its expected host gene, the circRNA was considered as true rt-circRNA.

miRNA sponge analysis

A list of 50 miRNAs which have been implicated in heart disease were obtained (References were provided in Additional file 10: Table S5). Seed sequence of these miRNAs was obtained from TargetScan database (http://www.targetscan.org/). CircRNA sequence was retrieved using custom R script and TargetScan was used for prediction of binding sites for the selected miRNAs. CircRNAs with at least 5 binding sites for at least one miRNA were considered to have potential to sponge miRNAs and were deposited in Supplemental Table S5.

Experimental validation of circRNAs

Due to the lack of availability of human samples, the mouse orthologs of 6 identified human circRNAs were used for experimental validation. Mouse heart RNA was extracted with TRIzol reagent (Invitrogen) and reverse transcribed using random hexamers and High Capacity RNA-to-cDNA Kit (Invitrogen). Convergent primers with opposite orientation were designed (Additional file 11: Table S6) and used for PCR amplification on mouse heart cDNA and genomic DNA. Mouse *Yap1* primers were used as a positive control for PCR amplification of genomic DNA. PCR products were then subjected to Sanger sequencing.

Abbreviations

circRNA: circular RNA

rt-circRNA: read-through circular RNA

HF: heart failure

DCM: dilated cardiomyopathy

PCA: principle component analysis

UTR: untranslated region

CDS: coding sequence

FC: fold change

ARVC: arrhythmogenic right ventricular cardiomyopathy

FPKM: fragments per kilobase of exon per million fragments mapped

SRPBM: spliced reads per billion mapped reads

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable

Availability of data and materials

The RNA-seq data used in this study was downloaded from NCBI database under accession number SRP108571.

Competing interests

The authors declare that they have no competing interests.

Funding

The work at the J.Z. laboratory is supported by a grant from the National Heart, Lung, and Blood Institute, NIH (R01HL132164). J.Z. is a recipient of Established Investigator Award (17EIA33460468) and Transformational Project Award (19TPA34910181) from the American Heart Association. K.D. is supported by a postdoctoral fellowship (19POST34450071) from the American Heart Association. The funders provided the financial support to this research, but had no role in the design of the study, analysis, interpretations of the data and in writing the manuscript.

Authors' contributions

J.Z. and K.D. designed the study. K.D. performed the bioinformatics analysis. K.D. and X.H. performed experimental validation of circRNAs. K.D. wrote the manuscript. H.S., D.F. and J.Z. edited the manuscript. All authors have read and approved the manuscript.

Acknowledgements

We gratefully thank all members at the J.Z. lab for their stimulating discussions of this work.

Supplementary Information

Additional file 1: Figure S1. Flowchart for *de novo* identification of circRNAs in human heart.

Additional file 2: Figure S2. Validation of the public RNA-seq dataset. (**A**) Principal Component Analysis (PCA) of the 10 samples (5 normal and 5 DCM hearts) using expression of genes with raw count > 10 in at least 5 samples. (**B**) Expression of *NPPA* in normal and DCM hearts from the RNA-seq dataset. PFKM: Fragments Per Kilobase of exon per Million fragments mapped. *p < 0.05 (edgeR).

Additional file 3: Table S1. Information of the 20,214 circRNAs identified in this study. Chromosome coordinates are in hg38 annotation.

Additional file 4: Table S2. List of circRNAs originated from two different genes identified in 10 human heart samples. Chromosome coordinates are in hg38 annotation. In the last column, “yes” indicates that the circRNA is from two neighboring genes with highly similar sequence, and “no” indicates no sequence similarity was observed and the circRNA is likely a bona fide rt-circRNA.

Additional file 5: Figure S3. Analysis of circRNAs from *MYH6* and *MYH7* gene. (A) Schematic diagram showing a number of rt-circRNAs originate from *MYH6* and *MYH7* genes in human heart identified in this study, study of Tan et al. [12] and Werfel et al. [11], and in **(B)** mouse heart identified in this study and study of Werfel et al. [11], as well as in **(C)** rat heart revealed in the study of Werfel et al. [11]. Arrows point to the transcriptional directions of *MYH6* and *MYH7* genes. **(D)** Heatmap illustrates the rank of the average expression across all studied samples for circRNAs produced from *MYH6* and *MYH7* genes in this study, and in human, mouse and rat hearts as identified by Tan et al. [12] and Werfel et al. [11]. The

numbers in the colored boxes indicate the rank of the circRNAs within the indicated species. **(E)** Representative example of circMYH6_e35:MYH7_e37 that is likely artifact. The upper schematic diagram illustrating the formation of circMYH6_e35:MYH7_e37 due to read misalignment. The exon 35 of *MYH6* (red, the donor exon) and exon 36 of *MYH7* (yellow, upstream of the acceptor exon) are homologous. The exon 36 of *MYH6* (green, downstream of the donor exon) and exon 37 of *MYH7* (blue, the acceptor exon) are homologues. Considering a chimeric sequencing read that is supposedly to be aligned back to exon 35 (red line without arrow) and 36 (green line with arrow) of *MYH6*, due to the sequence similarity, the down-stream part of the read (green line with arrow) may be misaligned to exon 37 of *MYH7* (blue line with arrow), leading to the identification of back-spliced junction, and versa vice. Dashed lines indicate splicing junctions. The bottom panel shows the sequence alignment of linear MYH6_e35:MYH6_e36, MYH7_e36:MYH7_e37, and the predicted sequence of the circular MYH6_e35:MYH7_e37.

Additional file 6: Figure S4. Relationship between numbers of circRNAs and exons of their host genes. (**A**) Schematic diagram of all the circRNAs originated from *TTN* and (**B**) *RYR2* gene. Arrows point to the transcriptional directions of *TTN* and *RYR2* genes. (**C**) The numbers of circRNAs increased proportionally with the numbers of exons per gene (binned to 10). (**D**) Distribution of exon numbers to produce circRNAs.

Additional file 7: Figure S5. Correlation of expression between 24 top most abundant circRNAs and their host genes. (A) CircRNAs with expression significantly correlated with the expression of host genes (Pearson correlation, $p < 0.05$**). (B)** CircRNAs with expression independent of their host genes (Pearson correlation, $p > 0.05$).

Additional file 8: Table S3. List of high-abundance and significantly dysregulated circRNAs in DCM compared to normal heart samples. Chromosome coordinates are in hg38 annotation. The threshold for significance was $p < 0.05$ and fold change > 2.

Additional file 9: Table S4. Information of DCM-related miRNAs used for analysis of sponge potential. The seed sequence of each miRNA was obtained from TargetScan database. PMID is the PubMed number for studies which reported the miRNAs.

Additional file 10: Table S5. CircRNAs with potential to sponge heart disease-related miRNAs. CircRNAs with at least 5 predicted binding sites for at least one miRNA were listed.

Additional file 11: Table S6. Primers used for validation of the mouse ortholog of selected human circRNAs. Mouse *Yap1* primers were used as internal control for PCR reaction using genomic DNA.

References

1. Ziaeian B, Fonarow GC: Epidemiology and aetiology of heart failure. *Nat Rev Cardiol* 2016, 13(6):368-378.

2. Lakdawala NK, Winterfield JR, Funke BH: Dilated cardiomyopathy. *Circ Arrhythm Electrophysiol* 2013, 6(1):228-237.
3. Bruneau BG: The developmental genetics of congenital heart disease. *Nature* 2008, 451(7181):943-948.
4. Bozkurt B, Colvin M, Cook J, Cooper LT, Deswal A, Fonarow GC, Francis GS, Lenihan D, Lewis EF, McNamara DM *et al*: Current Diagnostic and Treatment Strategies for Specific Dilated Cardiomyopathies: A Scientific Statement From the American Heart Association. *Circulation* 2016, 134(23):e579-e646.
5. McNally EM, Mestroni L: Dilated Cardiomyopathy: Genetic Determinants and Mechanisms. *Circ Res* 2017, 121(7):731-748.
6. Pugh TJ, Kelly MA, Gowrisankar S, Hynes E, Seidman MA, Baxter SM, Bowser M, Harrison B, Aaron D, Mahanta LM *et al*: The landscape of genetic variation in dilated cardiomyopathy as surveyed by clinical DNA sequencing. *Genet Med* 2014, 16(8):601-608.
7. Sweet ME, Cocciolo A, Slavov D, Jones KL, Sweet JR, Graw SL, Reece TB, Ambardekar AV, Bristow MR, Mestroni L *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.
8. Molina-Navarro MM, Trivino JC, Martinez-Dolz L, Lago F, Gonzalez-Juanatey JR, Portoles M, Rivera M: Functional networks of nucleocytoplasmic transport-related genes differentiate ischemic and dilated cardiomyopathies. A new therapeutic opportunity. *PLoS One* 2014, 9(8):e104709.
9. Barth AS, Kuner R, Buness A, Ruschhaupt M, Merk S, Zwermann L, Kaab S, Kreuzer E, Steinbeck G, Mansmann U *et al*: Identification of a common gene expression signature in dilated cardiomyopathy across independent microarray studies. *J Am Coll Cardiol* 2006, 48(8):1610-1617.
10. Qiu Z, Ye B, Yin L, Chen W, Xu Y, Chen X: Downregulation of AC061961.2, LING01-AS1, and RP11-13E1.5 is associated with dilated cardiomyopathy progression. *J Cell Physiol* 2019, 234(4):4460-4471.
11. Zhang X, Nie X, Yuan S, Li H, Fan J, Li C, Sun Y, Zhao Y, Hou H, Wang DW *et al*: Circulating Long Non-coding RNA ENST00000507296 Is a Prognostic Indicator in Patients with Dilated Cardiomyopathy. *Mol Ther Nucleic Acids* 2019, 16:82-90.
12. Satoh M, Minami Y, Takahashi Y, Tabuchi T, Nakamura M: Expression of microRNA-208 is associated with adverse clinical outcomes in human dilated cardiomyopathy. *J Card Fail* 2010, 16(5):404-410.
13. Satoh M, Minami Y, Takahashi Y, Tabuchi T, Nakamura M: A cellular microRNA, let-7i, is a novel biomarker for clinical outcome in patients with dilated cardiomyopathy. *J Card Fail* 2011, 17(11):923-929.
14. Onrat ST, Onrat E, Ercan Onay E, Yalim Z, Avsar A: The Genetic Determination of the Differentiation Between Ischemic Dilated Cardiomyopathy and Idiopathic Dilated Cardiomyopathy. *Genet Test Mol Biomarkers* 2018, 22(11):644-651.

15. Memczak S, Papavasileiou P, Peters O, Rajewsky N: Identification and Characterization of Circular RNAs As a New Class of Putative Biomarkers in Human Blood. *PLoS One* 2015, 10(10):e0141214.
16. Rybak-Wolf A, Stottmeister C, Glazar P, Jens M, Pino N, Giusti S, Hanan M, Behm M, Bartok O, Ashwal-Fluss R *et al*: Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed. *Mol Cell* 2015, 58(5):870-885.
17. Qu S, Yang X, Li X, Wang J, Gao Y, Shang R, Sun W, Dou K, Li H: Circular RNA: A new star of noncoding RNAs. *Cancer Lett* 2015, 365(2):141-148.
18. Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, Fatica A, Santini T, Andronache A, Wade M *et al*: Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. *Mol Cell* 2017, 66(1):22-37 e29.
19. Hall IF, Climent M, Quintavalle M, Farina FM, Schorn T, Zani S, Carullo P, Kunderfranco P, Civilini E, Condorelli G *et al*: Circ_Lrp6, a Circular RNA Enriched in Vascular Smooth Muscle Cells, Acts as a Sponge Regulating miRNA-145 Function. *Circ Res* 2019, 124(4):498-510.
20. Chen J, Cui L, Yuan J, Zhang Y, Sang H: Circular RNA WDR77 target FGF-2 to regulate vascular smooth muscle cells proliferation and migration by sponging miR-124. *Biochem Biophys Res Commun* 2017, 494(1-2):126-132.
21. Holdt LM, Stahringer A, Sass K, Pichler G, Kulak NA, Wilfert W, Kohlmaier A, Herbst A, Northoff BH, Nicolaou A *et al*: Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. *Nature communications* 2016, 7:12429.
22. Burd CE, Jeck WR, Liu Y, Sanoff HK, Wang Z, Sharpless NE: Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet* 2010, 6(12):e1001233.
23. Guarnerio J, Bezzi M, Jeong JC, Paffenholz SV, Berry K, Naldini MM, Lo-Coco F, Tay Y, Beck AH, Pandolfi PP: Oncogenic Role of Fusion-circRNAs Derived from Cancer-Associated Chromosomal Translocations. *Cell* 2016, 165(2):289-302.
24. Chen S, Huang V, Xu X, Livingstone J, Soares F, Jeon J, Zeng Y, Hua JT, Petricca J, Guo H *et al*: Widespread and Functional RNA Circularization in Localized Prostate Cancer. *Cell* 2019, 176(4):831-843 e822.
25. Hu W, Bi ZY, Chen ZL, Liu C, Li LL, Zhang F, Zhou Q, Zhu W, Song YY, Zhan BT *et al*: Emerging landscape of circular RNAs in lung cancer. *Cancer Lett* 2018, 427:18-27.
26. Werfel S, Nothjunge S, Schwarzmayr T, Strom TM, Meitinger T, Engelhardt S: Characterization of circular RNAs in human, mouse and rat hearts. *J Mol Cell Cardiol* 2016, 98:103-107.
27. Tan WL, Lim BT, Anene-Nzelu CG, Ackers-Johnson M, Dashi A, See K, Tiang Z, Lee DP, Chua WW, Luu TD *et al*: A landscape of circular RNA expression in the human heart. *Cardiovasc Res* 2017, 113(3):298-309.
28. Vo JN, Cieslik M, Zhang Y, Shukla S, Xiao L, Zhang Y, Wu YM, Dhanasekaran SM, Engelke CG, Cao X *et al*: The Landscape of Circular RNA in Cancer. *Cell* 2019, 176(4):869-881 e813.

29. Jakobi T, Czaja-Hasse LF, Reinhardt R, Dieterich C: Profiling and Validation of the Circular RNA Repertoire in Adult Murine Hearts. *Genomics Proteomics Bioinformatics* 2016, 14(4):216-223.
30. van Heesch S, Witte F, Schneider-Lunitz V, Schulz JF, Adami E, Faber AB, Kirchner M, Maatz H, Blachut S, Sandmann CL *et al*: The Translational Landscape of the Human Heart. *Cell* 2019, 178(1):242-260 e229.
31. Hansen TB, Veno MT, Damgaard CK, Kjems J: Comparison of circular RNA prediction tools. *Nucleic Acids Res* 2016, 44(6):e58.
32. Langmead B, Salzberg SL: Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012, 9(4):357-359.
33. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M *et al*: Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 2013, 495(7441):333-338.
34. Song W, Wang H, Wu Q: Atrial natriuretic peptide in cardiovascular biology and disease (NPPA). *Gene* 2015, 569(1):1-6.
35. Gruner H, Cortes-Lopez M, Cooper DA, Bauer M, Miura P: CircRNA accumulation in the aging mouse brain. *Sci Rep* 2016, 6:38907.
36. Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO: Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* 2012, 7(2):e30733.
37. Westholm JO, Miura P, Olson S, Shenker S, Joseph B, Sanfilippo P, Celniker SE, Graveley BR, Lai EC: Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep* 2014, 9(5):1966-1980.
38. Du W, Pan Z, Chen X, Wang L, Zhang Y, Li S, Liang H, Xu C, Zhang Y, Wu Y *et al*: By targeting Stat3 microRNA-17-5p promotes cardiomyocyte apoptosis in response to ischemia followed by reperfusion. *Cell Physiol Biochem* 2014, 34(3):955-965.
39. Pullamsetti SS, Doebele C, Fischer A, Savai R, Kojonazarov B, Dahal BK, Ghofrani HA, Weissmann N, Grimminger F, Bonauer A *et al*: Inhibition of microRNA-17 improves lung and heart function in experimental pulmonary hypertension. *Am J Respir Crit Care Med* 2012, 185(4):409-419.
40. Bao Y, Willis BC, Frasier CR, Lopez-Santiago LF, Lin X, Ramos-Mondragon R, Auerbach DS, Chen C, Wang Z, Anumonwo J *et al*: Scn2b Deletion in Mice Results in Ventricular and Atrial Arrhythmias. *Circ Arrhythm Electrophysiol* 2016, 9(12).
41. Watanabe H, Darbar D, Kaiser DW, Jiramongkolchai K, Chopra S, Donahue BS, Kannankeril PJ, Roden DM: Mutations in sodium channel beta1- and beta2-subunits associated with atrial fibrillation. *Circ Arrhythm Electrophysiol* 2009, 2(3):268-275.
42. Gigante B, Vikstrom M, Meuzelaar LS, Chernogubova E, Silveira A, Hooft FV, Hamsten A, de Faire U: Variants in the coagulation factor 2 receptor (F2R) gene influence the risk of myocardial infarction in men through an interaction with interleukin 6 serum levels. *Thromb Haemost* 2009, 101(5):943-953.
43. Arnett DK, Li N, Tang W, Rao DC, Devereux RB, Claas SA, Kraemer R, Broeckel U: Genome-wide association study identifies single-nucleotide polymorphism in KCNB1 associated with left

ventricular mass in humans: the HyperGEN Study. *BMC Med Genet* 2009, 10:43.

44. Frank D, Gantenberg J, Boomgaarden I, Kuhn C, Will R, Jarr KU, Eden M, Kramer K, Luedde M, Mairbaurl H *et al*: MicroRNA-20a inhibits stress-induced cardiomyocyte apoptosis involving its novel target Egln3/PHD3. *J Mol Cell Cardiol* 2012, 52(3):711-717.
45. Hill JA, Olson EN: Cardiac plasticity. *N Engl J Med* 2008, 358(13):1370-1380.
46. Lim TB, Aliwarga E, Luu TDA, Li YP, Ng SL, Annadoray L, Sian S, Ackers-Johnson MA, Foo RS: Targeting the highly abundant circular RNA circSlc8a1 in cardiomyocytes attenuates pressure overload induced hypertrophy. *Cardiovasc Res* 2019.
47. Shan K, Liu C, Liu BH, Chen X, Dong R, Liu X, Zhang YY, Liu B, Zhang SJ, Wang JJ *et al*: Circular Noncoding RNA HIPK3 Mediates Retinal Vascular Dysfunction in Diabetes Mellitus. *Circulation* 2017, 136(17):1629-1642.
48. Bachmayr-Heyda A, Reiner AT, Auer K, Sukhbaatar N, Aust S, Bachleitner-Hofmann T, Mesteri I, Grunt TW, Zeillinger R, Pils D: Correlation of circular RNA abundance with proliferation–exemplified with colorectal and ovarian cancer, idiopathic lung fibrosis, and normal human tissues. *Sci Rep* 2015, 5:8057.
49. Beltrami CA, Finato N, Rocco M, Feruglio GA, Puricelli C, Cigola E, Sonnenblick EH, Olivetti G, Anversa P: The cellular basis of dilated cardiomyopathy in humans. *J Mol Cell Cardiol* 1995, 27(1):291-305.
50. Khan MA, Reckman YJ, Aufiero S, van den Hoogenhof MM, van der Made I, Beqqali A, Koolbergen DR, Rasmussen TB, van der Velden J, Creemers EE *et al*: RBM20 Regulates Circular RNA Production From the Titin Gene. *Circ Res* 2016, 119(9):996-1003.
51. Bers DM: Cardiac excitation-contraction coupling. *Nature* 2002, 415(6868):198-205.
52. Herman DS, Lam L, Taylor MR, Wang L, Teekakirikul P, Christodoulou D, Conner L, DePalma SR, McDonough B, Sparks E *et al*: Truncations of titin causing dilated cardiomyopathy. *N Engl J Med* 2012, 366(7):619-628.
53. Tang Y, Tian X, Wang R, Fill M, Chen SR: Abnormal termination of Ca2+ release is a common defect of RyR2 mutations associated with cardiomyopathies. *Circ Res* 2012, 110(7):968-977.
54. Salzman J, Chen RE, Olsen MN, Wang PL, Brown PO: Cell-type specific features of circular RNA expression. *PLoS Genet* 2013, 9(9):e1003777.
55. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J: Natural RNA circles function as efficient microRNA sponges. *Nature* 2013, 495(7441):384-388.
56. Rong D, Sun H, Li Z, Liu S, Dong C, Fu K, Tang W, Cao H: An emerging function of circRNA-miRNAs-mRNA axis in human diseases. *Oncotarget* 2017, 8(42):73271-73281.
57. Chen JF, Mandel EM, Thomson JM, Wu Q, Callis TE, Hammond SM, Conlon FL, Wang DZ: The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat Genet* 2006, 38(2):228-233.
58. Bolger AM, Lohse M, Usadel B: Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014, 30(15):2114-2120.

59. Anders S, Pyl PT, Huber W: HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015, 31(2):166-169.
60. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 2008, 5(7):621-628.
61. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S: The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009, 25(16):2078-2079.
62. Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010, 26(1):139-140.
63. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK: Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019, 10(1):1523.

Figures

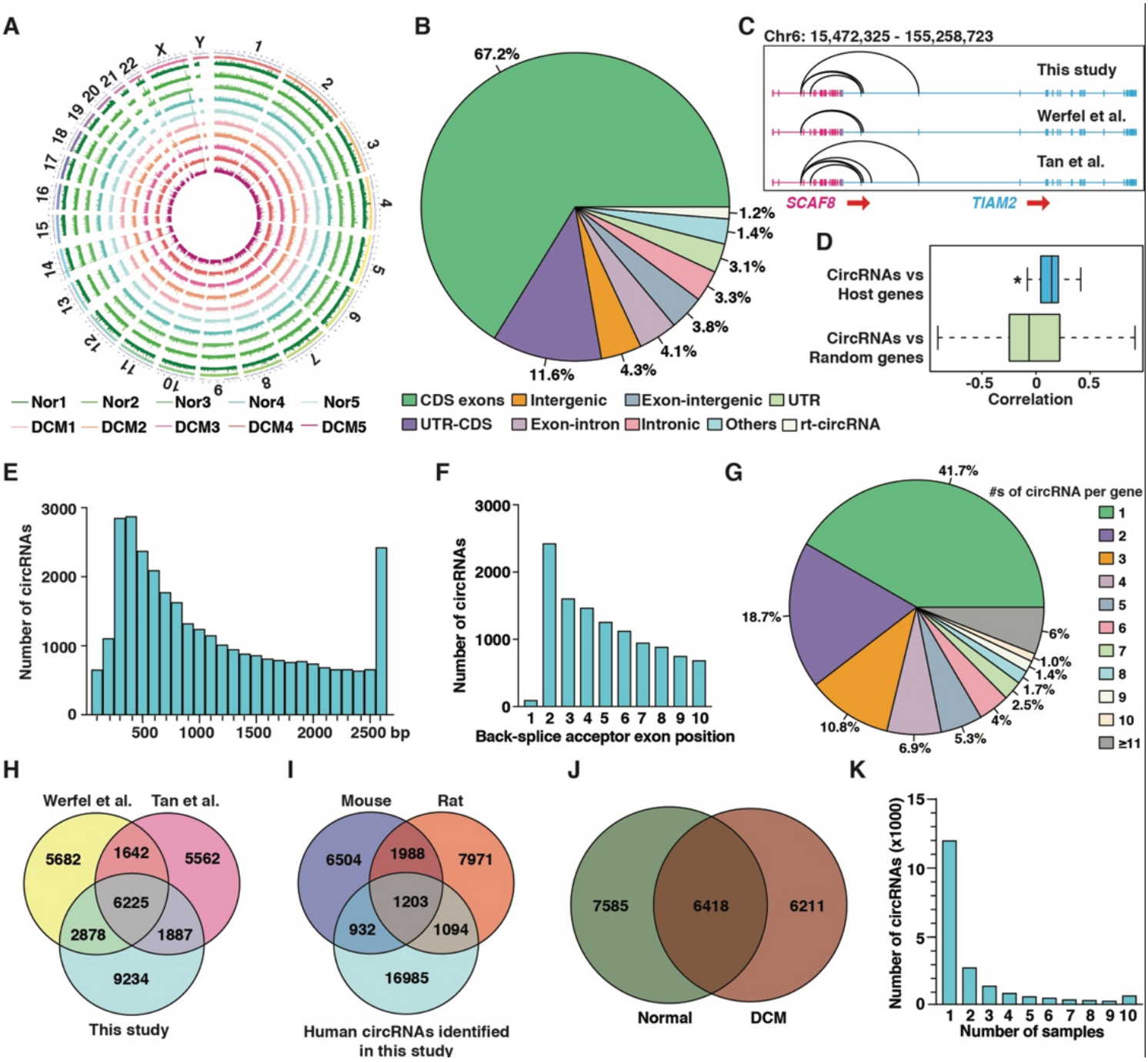

Figure 1

General feature of the circRNAs identified in human heart. (A) The distribution of 20,214 circRNAs identified in human hearts from 5 normal (Nor1-5) and 5 DCM (DCM1-5) patients, was mapped in different chromosomes. The circus plot tracks represent the normalized expression (spliced reads per billion mapping, denoted as SRPBM) of circRNAs in the different samples. (B) Percentage of genomic origins of the identified circRNAs. (C) Schematic diagram showing that rt-circRNAs originate from SCAF8 (in red) and TIAM2 (in cyan) genes identified in human heart in this study (upper) and in studies by Werfel et al. (middle) [26] and Tan et al. [27] (bottom). Arrows point to the transcriptional directions of SCAF8 and TIAM2 genes. (D) Correlation between the expression of circRNAs and their host genes or randomly chosen linear genes. *p < 0.05. (E) The size distribution of identified circRNAs. (F) Distribution of

circRNAs with respect to most upstream circularized exon. (G) Distribution of numbers of circRNAs produced per gene. (H) Venn diagram showing the over-lapping circRNAs identified in human heart in this study and in studies by Tan et al. and Werfel et al.. (I) Venn diagram showing the over-lapping circRNAs identified in human heart in this study and circRNAs identified in mouse and rat heart by Tan et al. and Werfel et al.. (J) Venn diagram showing the number of circRNAs unique or common between normal and DCM human samples. (K) Number of human heart samples in which circRNAs with different abundance were identified.

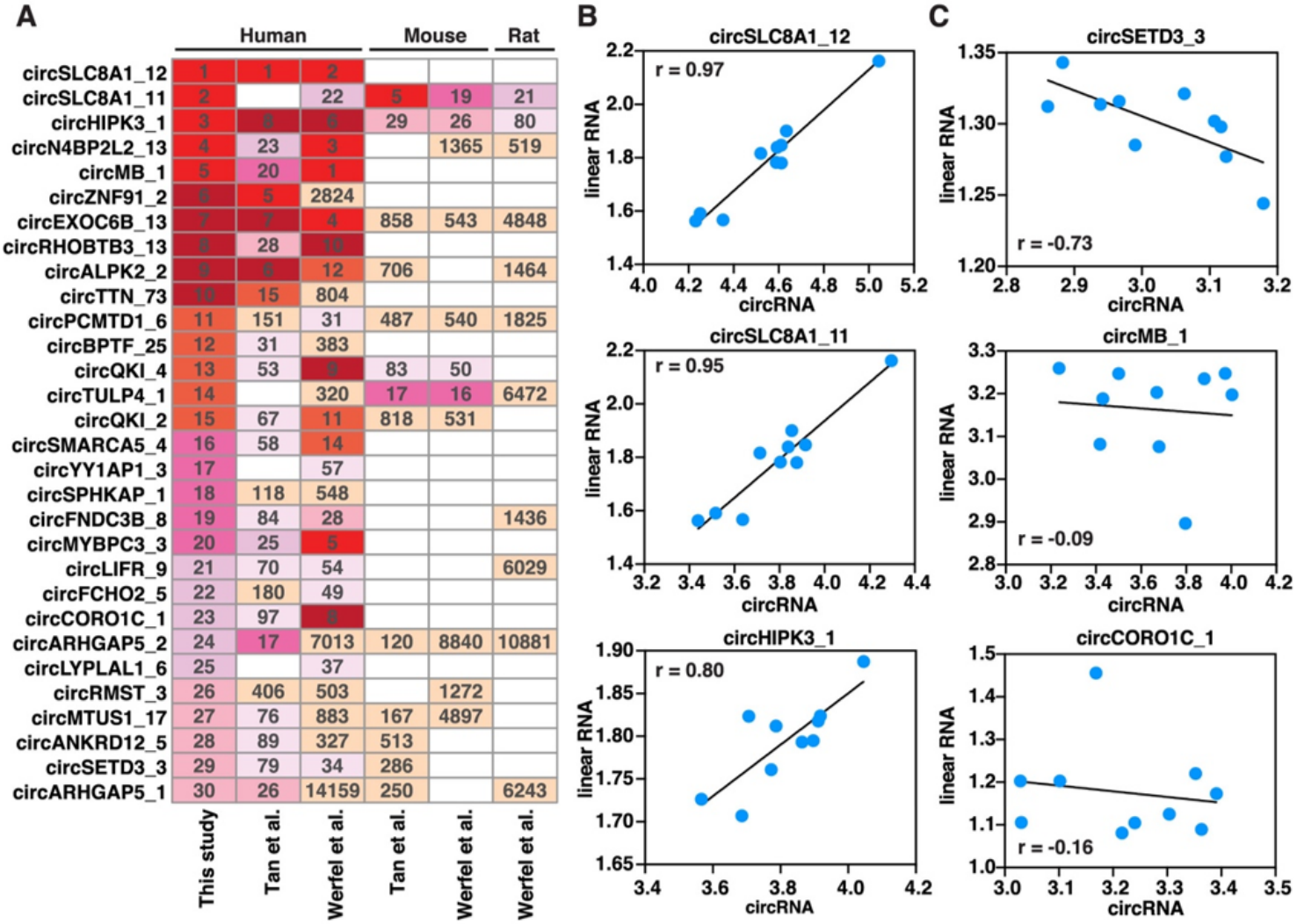

Figure 2

Identification of the most abundant circRNAs in normal human hearts. (A) Heatmap illustrates the ranking of the top 30 most abundant circRNAs identified in normal hearts in this study, and in normal human, mouse and rat hearts that were identified by Tan et al. and Werfel et al. [26, 27]. The numbers in the colored boxes indicate the ranking of the circRNAs within the indicated species and blank boxes represent the circRNA was not detected in the indicated species. (B) Correlation between expression of the top 3 most abundant circRNAs including circSLC8A1_12, circSLC8A1_11 and circHIPK3_1 and their host genes which show positively correlation, as well as (C) circSETD3_3, circMB_1, circCORO1C_1 and their host genes which are independent of each other.

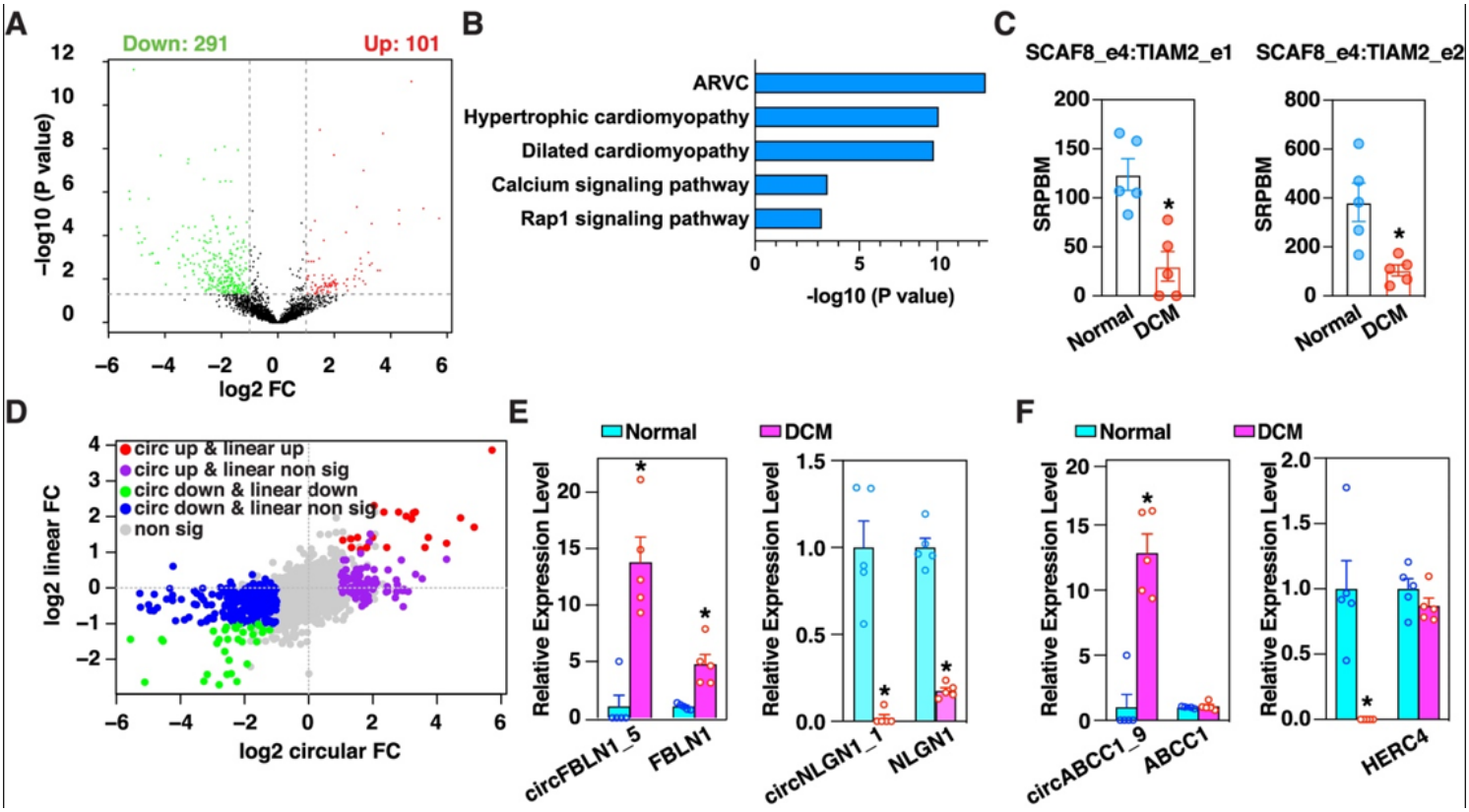

Figure 3

Identification of dysregulated circRNAs in DCM. (A) Volcano plot showing differentially expressed circRNAs in DCM as compared to normal samples ($p < 0.05$ and Fold Change (FC) > 2). (B) Enriched pathway analysis for the host genes of the significantly differentially expressed circRNAs. ARVC: arrhythmogenic right ventricular cardiomyopathy. (C) Down-regulation of 2 rt-circRNAs originated from SCAF8 and TIAM2 gene including SCAF8_e4:TIAM2_e1 and SCAF8_e4:TIAM2_e2 in DCM. (D) Correlation of log FC of circRNA versus log FC of linear RNA expression. (E) Examples of significantly up- and down-regulated circRNAs with expression that is dependent or (F) independent of their host genes. *$p < 0.05$ (edgeR).

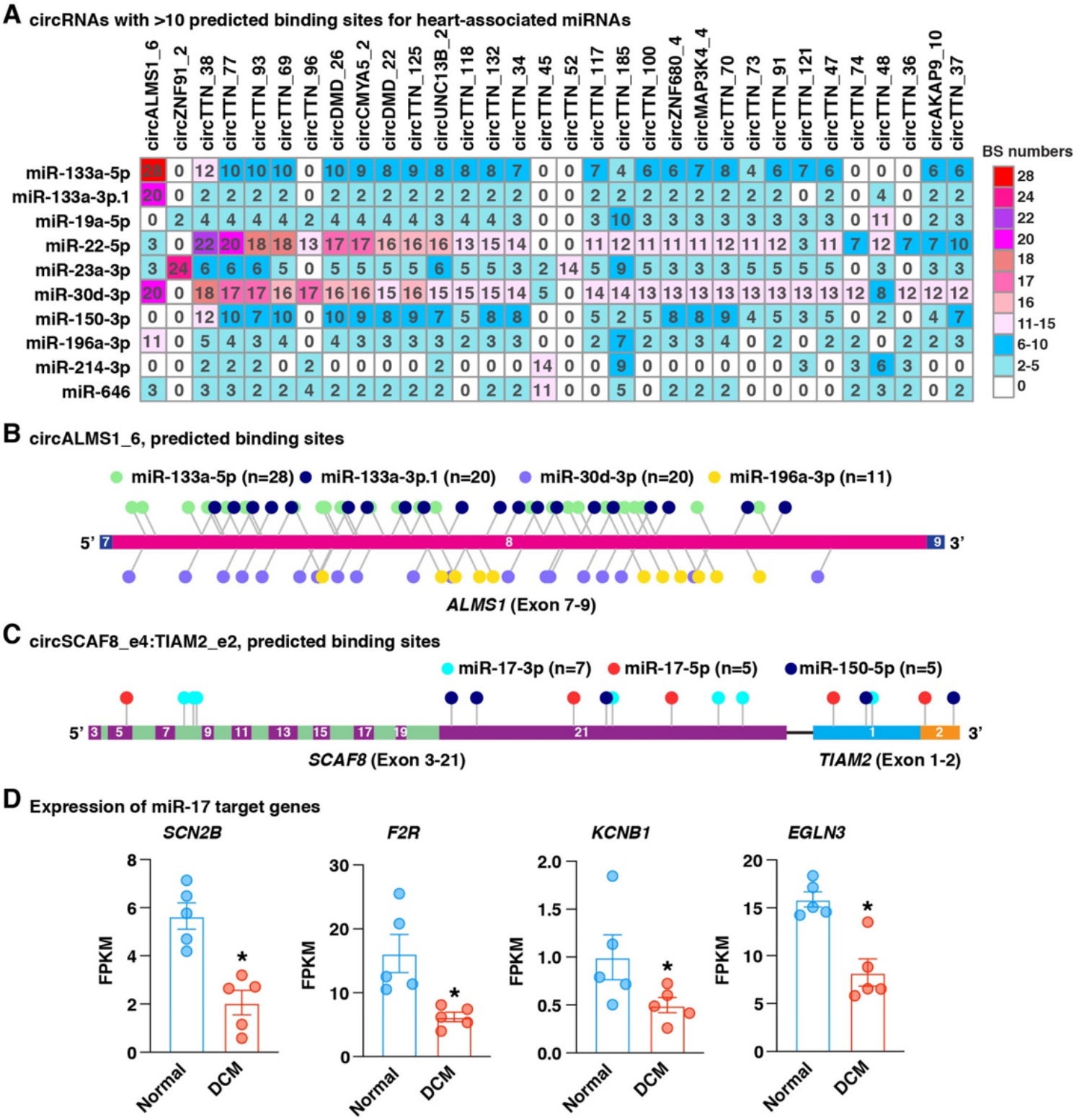

Figure 4

Prediction of circRNA potential for sponging miRNAs. (A) Heatmap showing the number of binding sites (BS) of circRNAs with at least 10 BS for at least one heart disease-related miRNA. (B) Schematic representation of circALMS1_6 which was predicted to have the strongest potential to sponge miR-133. (C) Schematic representation of a rt-circRNAs, circSCAF8_e4:TIAM2_e2, which was predicted to have potential to sponge multiple heart disease-related miRNAs. Colored rectangles indicate exons and the

numbers within them represent the ID of exons. (D). Expression of miR-17 target genes in normal and DCM human hearts. *p < 0.05 (edgeR).

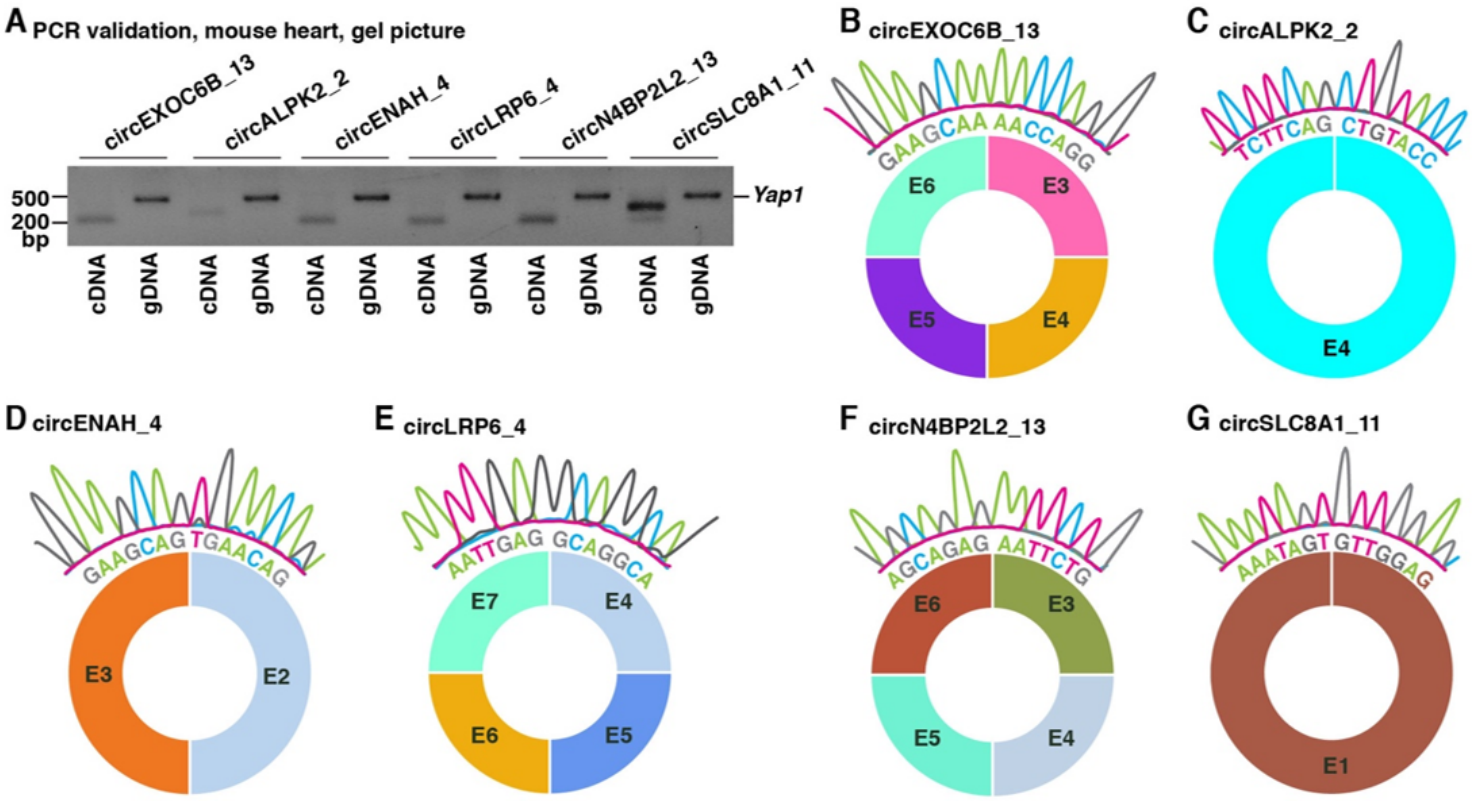

Figure 5

Experimental validation of identified circRNAs. (A) Representative gel pictures showing PCR amplification of the murine orthologs of 6 selected human circRNAs using divergent primers and mouse heart cDNA. Mouse heart genomic DNA served as negative control. Mouse Yap1 primers served as a positive control for the PCR amplification of genomic DNA. (B) Sanger sequencing results of the mouse homolog of circEXOC6B_13, (C) circALPK2_2, (D) circENAH_4, (E) circLRP6_4, (F) circN4BP2L2_13 and (G) circSLC8A1_11.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- additionalfile12originalgelpicture.pdf
- Additionalfile11TableS6.xlsx
- Additionalfile10TableS5.xlsx
- Additionalfile9TableS4.xlsx
- Additionalfile8TableS3.xlsx
- Additionalfile7FigureS5top30circRNAcorrelationV2.pdf

- Additionalfile6FigureS4exonusageofcircRNAsV4.pdf
- Additionalfile5FigureS3MYH6MYH7rtCircRNAsV2.pdf
- Additionalfile4TableS2.xlsx
- Additionalfile3TableS1.xlsx
- Additionalfile2FigureS2validationofRNAseqdataV4.pdf
- Additionalfile1FigureS1WokingflowchartV5.pdf