

Accepted Manuscript

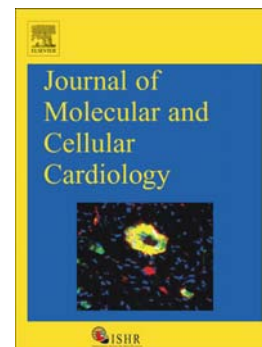
Characterization of circular RNAs in human, mouse and rat hearts

Stanislas Werfel, Stephan Nothjunge, Thomas Schwarzmayr, Tim-Matthias Strom, Thomas Meitinger, Stefan Engelhardt

PII: S0022-2828(16)30229-2
DOI: doi: [10.1016/j.yjmcc.2016.07.007](https://doi.org/10.1016/j.yjmcc.2016.07.007)
Reference: YJMCC 8425

To appear in: *Journal of Molecular and Cellular Cardiology*

Received date: 1 June 2016
Revised date: 19 July 2016
Accepted date: 26 July 2016



Please cite this article as: Werfel Stanislas, Nothjunge Stephan, Schwarzmayr Thomas, Strom Tim-Matthias, Meitinger Thomas, Engelhardt Stefan, Characterization of circular RNAs in human, mouse and rat hearts, *Journal of Molecular and Cellular Cardiology* (2016), doi: [10.1016/j.yjmcc.2016.07.007](https://doi.org/10.1016/j.yjmcc.2016.07.007)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Characterization of circular RNAs in human, mouse and rat hearts

Stanislas Werfel^{1,2}, Stephan Nothjunge¹, Thomas Schwarzmayr³, Tim-Matthias Strom^{3,4},
Thomas Meitinger^{3,4}, Stefan Engelhardt^{1,2}

¹Institut für Pharmakologie und Toxikologie, Technische Universität München (TUM),
Munich, Germany;

²DZHK (German Center for Cardiovascular Research), partner site Munich Heart Alliance,
Munich, Germany;

³Institute of Human Genetics, German Research Center for Environmental Health (GmbH),
Helmholtz Zentrum München, Neuherberg, Germany;

⁴Institute of Human Genetics, Technische Universität München, Munich, Germany.

Abstract

Deep sequencing techniques and advanced data analysis methods recently enabled the characterization of thousands of circular RNA isoforms (circRNAs) from a number of tissues and organisms. There is emerging evidence that some circRNAs may have important biological functions or serve as diagnostic biomarkers in disease conditions.

In order to analyze circRNA expression in the heart and its changes in different conditions we performed RNA-Seq analysis of ribosome-depleted libraries from rats (neonatal and adult), mice (sham or after transverse aortic constriction, TAC) and humans (failing, non-failing). All samples were sequenced after a treatment with exonuclease RNase R or a mock treatment and >9,000 candidate circRNAs were detected for each species.

Additionally, we performed separate isolation of nuclear and cytoplasmic RNA and co-immunoprecipitated RNA interacting with endogenous argonaute 2 (Ago2) in primary cardiac myocytes. We found circRNAs to be significantly enriched in the cytoplasm compared to linear transcripts and to have a similar level of association with Ago2.

Notably in all three species we observed dozens of circRNAs arising from the titin (Ttn) gene, which is known to undergo highly complex alternative splicing during heart maturation. Correspondingly we observed extensive differential regulation of Ttn circRNAs between neonatal and adult rat hearts, suggesting that circRNA formation could be involved in the regulation of titin splicing.

We expect that our inventory of cardiac circRNAs, as well as the information on their conservation and differential expression will provide an important basis for further studies addressing their function and suitability as biomarkers.

Keywords

Circular RNA; circRNA; heart; titin

1. Introduction

Although circular RNA molecules (circRNAs) have been first described a long time ago, it was not until recently, that it was discovered how widespread their expression is [1–3]. CircRNAs typically arise through splicing of a splice donor to an upstream localized splice acceptor, thus forming a closed circle consisting of regular 3'-5' carbon linkages throughout. Molecules formed this way most often arise at known exons and annotated splice sites and are also termed “exonic” circRNAs to distinguish them from “intronic” forms that may arise from introns and contain a 2'-5' carbon linkage at the branch point (similar to lariat RNA)[3]. It is currently assumed that due to their circular structure and the absence of a 5' cap, circRNAs are not translated into proteins (though there might be exceptions [4]). It has been reported for several circRNAs that they can exert cellular functions by binding to and inhibiting microRNAs [2,5], but this mechanism could be reserved for circRNAs with exceptionally high number of target sites for a specific microRNA. It appears plausible that they could assume further functions, similar to long non-coding RNAs (lncRNAs).

CircRNAs have significantly longer half-lives than linear RNA molecules, possibly due to reduced exonuclease susceptibility [2] and rising evidence suggests that they can be secreted and identified in plasma, which makes them promising candidate disease biomarkers [6].

A typical feature of circRNAs is the “backsplice” or “head-to-tail” splice junction, where exons in the RNA appear in reversed order compared to their chromosomal localization (Fig. 1A, bottom). Deep sequencing reads spanning this junction allow recognition of circRNA candidates. These can be further confirmed through enrichment after treatment with the exonuclease RNase R, which digests linear but not circular RNA.

A characterization of cardiac circRNAs has previously been performed in the hearts of healthy adult mice, resulting in the discovery of several hundred candidates [7], however, their expression in humans and rats and changes in disease conditions have so far not been reported.

Here we characterize circRNAs in the hearts of humans, mice and rats in native and disease conditions and assess further characteristics of this class of molecules, such as their cellular localization and association with microRNAs.

2. Materials and Methods

Materials and methods are described in the online data supplement.

3. Results and Discussion

3.1 Description of analyzed samples

We performed deep sequencing analysis of ribosomal-RNA(rRNA)-depleted heart samples from rat (neonatal, NRH, and adult, ARH), mouse (sham and 3 weeks after pressure overload using transverse aortic constriction, TAC) and human (non-failing hearts, HN, heart-failure patients, HF). Each sample was treated with RNase R or mock treated to assess enrichment of exonuclease-resistant circRNAs (Fig. 1A).

Additionally, we performed separate isolation of cytoplasmic and nuclear RNA (2 samples each) and ribonucleoprotein immunoprecipitation (RIP) with antibodies against endogenous argonaute 2 (Ago2) and heat shock protein 90 (Hsp90) as control (2 samples each) in neonatal rat cardiac myocytes (NRCMs). Overall 38 RNA-Seq libraries were generated and sequenced on an Illumina HiSeq 2500 machine. Suppl. Table 1 shows an overview of all samples and obtained reads that passed quality control. To confirm the presence of cardiac failure, levels of atrial natriuretic peptide (NPPA) and β -myosin heavy chain (MYH7) mRNA were measured using qPCR in the human samples and showed strong elevation in failing hearts (Suppl. Table 2A). Mice were assessed using the ventricle weight normalized to body weight ratio, Nppa (both strongly elevated) and ejection fraction (strongly reduced, Suppl. Table 2B).

We applied the algorithm established by Memczak et al. [2] to quantify backsplices and linear junctions. Only candidates supported by at least three unique reads were reported for each sample. We determined >9,000 candidate circRNAs for each species, of which $\approx 30\%$ were conserved between mouse and rat and $\approx 10\%$ were conserved in all three species (Fig. 1B). To our knowledge this is the first extensive analysis of circRNA conservation in a specific tissue between several species.

The candidates were indexed in ascending order based on their genomic location and full results are reported in Suppl. Table 3 (rat), Suppl. Table 4 (mouse) and Suppl. Table 5 (human).

3.2 CircRNAs are highly enriched in the cytoplasm of NRCMs

To find a cutoff for expected enrichment of circRNAs after RNase R digestion with our protocol, we compared the enrichment of robustly expressed rat circRNA candidates (≥ 3 backsplice-spanning reads detected in either all mock NRH or all mock ARH samples, $n=988$) to linear junctions (Fig. 1C). Most circRNA candidates localized to a population distinct from linear junctions with few (likely false positives) showing low or no enrichment. Based on the distributions of the two populations we defined a threshold for relative enrichment ($2^{1.1} \approx 2.14$) with an expected false-positive rate $< 5\%$, to select high-probability and high expression rat circRNAs that were retained for further analyses.

To assess cytoplasmic vs. nuclear localization we first confirmed successful separate isolation of the two fractions by quantifying in our libraries the reads falling on representative RNAs with known localization. We found (as previously reported [8]) that the long non-coding RNAs Malat1 and RGD1566401(Meg3) were highly enriched in the nuclear fraction and the mRNA of Gapdh in the cytoplasmic fraction (Fig. 1D).

Although it has been previously described that circRNAs can localize to the cytoplasm [1,2], we were surprised to find that they were highly enriched there compared to linear transcripts (Fig. 1E). To our knowledge this is the first deep sequencing analysis comparing the subcellular localization of circRNAs to linear RNAs and it remains currently unknown how circular RNAs exit the nucleus. It appears unlikely that a selective export mechanism exists for circRNAs that is more efficient than the regular mRNA export. A possible explanation could be selective depletion of circRNAs in the nucleus (e.g. through degradation), which is absent or much weaker in the cytoplasm.

3.3 CircRNAs show a similar extent of interaction with the RNA-induced silencing complex (RISC) as do linear transcripts in NRCMs

We indirectly assessed the overall extent of circRNA-microRNA interactions on transcriptome level using (endogenous) Ago2-RIP-sequencing in NRCMs, with an Anti-Hsp90 antibody serving as control. To confirm successful immunoprecipitation and microRNA target enrichment, we quantified sequencing reads falling on known targets, the lncRNA H19 and Pten, and found strong enrichment in the Ago2-RIP but not in the Hsp90-RIP, while Gapdh was depleted (Fig. 1F).

A comparison of circRNAs to linear transcripts in the Ago2-RIP (normalized to baseline expression in NRH) showed a similar relative enrichment of the two RNA forms (Fig. 1G), suggesting that circRNAs, on average, interact with microRNAs to a similar extent as do linear transcript. Interestingly the most abundant circRNA detected in the Ago2-RIP stemmed from the Slc8a1 gene, which codes for the sodium-calcium exchanger (NCX1). It was also among the most abundant circRNAs in all three species and is one of the very few mammalian circRNAs that has been discovered before the advent of deep sequencing technology [4].

3.4 Confirmation of select circRNA candidates

To further confirm the circularity of select candidates we quantified their enrichment in RNase R treated vs. mock treated samples using qPCR (Fig. 2A). We observed that circRNAs with higher predicted lengths showed lower enrichment, possibly due to minimal endonuclease contamination or small degree of RNA degradation during or after RNA isolation (m004295: 1.8 kb, m005505: 6.2 kb, m005504: 6,5 kb vs. m001237: 0.4 kb, m003757: 0.6 kb). In addition, we quantified relative efficiency of reverse transcription (RT) with random hexamer vs. oligo(dT) primers. Circular RNAs do not have poly-A tails and therefore an RT with random hexamers is expected to show an apparent enrichment of the circular form compared to oligo(dT), when normalized to the linear mRNA (further described in supplementary methods). As expected the selected circRNAs showed marked enrichment in this assay (Suppl. Fig. 1, Ttn circRNAs not included due to very long distance of circRNA exons from the poly-A tail).

3.5 Differential expression analysis of cardiac circRNAs

To calculate overall circRNA expression levels, normalized backsplice-spanning read counts of all candidates meeting the RNase R enrichment threshold ($n > 9,000$ in each species) were summed up. We observed significantly lower overall circRNA expression in adult compared to neonatal rat hearts (Fig. 2B). Interestingly human hearts showed a baseline level comparably high to NRH, about double of that in adult rats and mice.

We then selected robustly expressed circRNA candidates for each species (detected with ≥ 3 backsplice-spanning reads in each mock-treated sample of at least one treatment group, e.g. in all NRH_mock or all ARH_mock samples, and passing the RNase R threshold) and performed a statistical analysis for differential expression in the analyzed conditions (Suppl. Tables 6A, B and C for rat, mouse and human respectively). Replicates from individual groups clustered well together in a hierarchical clustering analysis (Suppl. Fig. 2).

As expected from the summary analysis (Fig. 2B) we found a large number of strongly and significantly (with a false discovery rate, $FDR < 0.1$) regulated circRNAs between neonatal and adult rat hearts (Suppl. Table 6A and Fig. 2C).

We detected over 40 backsplices stemming from the sense strand of the titin (Ttn) gene in all three species, most of which were enriched in RNase R treated samples. Titin transcripts undergo highly complex, developmentally controlled differential splicing, resulting in many thousands of possible isoforms [9]. Correspondingly we found several titin-derived circRNAs which were specifically and strongly enriched in adult and others in neonatal rat hearts, respectively (Fig. 2C, Suppl. Table 6A). We further confirmed this finding by qPCR in independent samples (Fig. 2F). Also one titin-derived circRNA was significantly downregulated in mice subjected to pressure overload (Fig. 2F, Suppl. Table 6B).

It has been reported that the formation of circRNAs can influence mRNA processing by competing with linear splice events [10]. It appears therefore possible that circRNA formation could be one of the driving forces behind the diversity of titin splicing. Furthermore missplicing of titin has been causally linked to hereditary cardiomyopathy in rats and humans [11]. It is therefore tempting to speculate that perturbation of circRNA levels could be contributing to the disease phenotype in rodents and patients.

We also found 2 other genes (Eya3 and Scmh1) that, similarly to titin, produced differentially regulated circRNAs, of which some were enriched by at least 2-fold in neonatal and others in adult rat hearts with an $FDR < 0.1$ (Fig. 2C and F, Suppl. Table 6A). These examples provide strong evidence for specific regulation of circRNA formation, independent of transcriptional regulation.

Under disease conditions a reexpression of fetal genes such as NPPA and MYH7 typically takes place in cardiomyocytes. Both in mouse and human failing hearts we observed a tendency towards increased circRNA expression (Fig. 2B and visible as a shift in the point clouds in Fig. 2D and E), however the overall changes were lower and less circRNAs were differentially expressed compared to neonatal vs. adult rat hearts (Suppl. Tables 6B and C).

Apart from titin, several other genes showed a high number of different circRNA isoforms, notably ryanodine receptor 2 (RYR2), which expressed over 100 circRNA isoforms in human hearts that were enriched after RNase R treatment (Suppl. Table 7).

4. Conclusions

Taken together our data provide an extensive catalogue of circRNAs in human, mouse and rat hearts and their conservation between the three species.

CircRNAs appear to be robustly expressed and show differential regulation in postnatal development and cardiac disease.

Glossary

Backsplice (or head-to-tail splice junction): a splice junction that produces circular RNA and can be identified in RNA-Seq data due to exchanged order of exons.

RNA-Seq: Deep sequencing of RNA samples

Acknowledgements

We thank Roger Hajjar (Mount Sinai School of Medicine, New York) for kindly providing patient heart samples.

We thank Nikolaus Rajewsky, Sebastian Memczak and Antigoni Elefsinioti (Max-Delbrück-Center for Molecular Medicine, Berlin) for providing help with RNA-Seq library preparation and analysis.

Funding

This work was supported by the Bavarian Ministry of Sciences, Research and the Arts in the framework of the Bavarian Molecular Biosystems Research Network.

Disclosures

None.

References

- [1] J. Salzman, C. Gawad, P.L. Wang, N. Lacayo, P.O. Brown, Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types., *PLoS One*. 7 (2012) e30733. doi:10.1371/journal.pone.0030733.
- [2] S. Memczak, M. Jens, A. Elefsinioti, F. Torti, J. Krueger, A. Rybak, L. Maier, S.D. Mackowiak, L.H. Gregersen, M. Munschauer, A. Loewer, U. Ziebold, M. Landthaler, C. Kocks, F. le Noble, N. Rajewsky, Circular RNAs are a large class of animal RNAs with regulatory potency., *Nature*. 495 (2013) 333–8. doi:10.1038/nature11928.
- [3] W.R. Jeck, N.E. Sharpless, Detecting and characterizing circular RNAs., *Nat. Biotechnol.* 32 (2014) 453–61. doi:10.1038/nbt.2890.
- [4] X.-F. Li, J. Lytton, A Circularized Sodium-Calcium Exchanger Exon 2 Transcript, *J. Biol. Chem.* 274 (1999) 8153–8160. doi:10.1074/jbc.274.12.8153.
- [5] T.B. Hansen, T.I. Jensen, B.H. Clausen, J.B. Bramsen, B. Finsen, C.K. Damgaard, J. Kjems, Natural RNA circles function as efficient microRNA sponges., *Nature*. 495 (2013) 384–8. doi:10.1038/nature11993.
- [6] W. Koh, W. Pan, C. Gawad, H.C. Fan, G.A. Kerchner, T. Wyss-Coray, Y.J. Blumenfeld, Y.Y. El-Sayed, S.R. Quake, Noninvasive in vivo monitoring of tissue-specific global gene expression in humans., *Proc. Natl. Acad. Sci. U. S. A.* 111 (2014) 7361–6. doi:10.1073/pnas.1405528111.
- [7] T. Jakobi, L.F. Czaja-Hasse, R. Reinhardt, C. Dieterich, Profiling and Validation of the

- Circular RNA Repertoire in Adult Murine Hearts., *Genomics. Proteomics Bioinformatics.* (2016). doi:10.1016/j.gpb.2016.02.003.
- [8] K.M. Michalik, X. You, Y. Manavski, A. Doddaballapur, M. Zörnig, T. Braun, D. John, Y. Ponomareva, W. Chen, S. Uchida, R. a Boon, S. Dimmeler, Long Noncoding RNA MALAT1 Regulates Endothelial Cell Function and Vessel Growth., *Circ. Res.* 114 (2014) 1389–97. doi:10.1161/CIRCRESAHA.114.303265.
- [9] M. Krüger, W.A. Linke, The giant protein titin: a regulatory node that integrates myocyte signaling pathways., *J. Biol. Chem.* 286 (2011) 9905–12. doi:10.1074/jbc.R110.173260.
- [10] R. Ashwal-Fluss, M. Meyer, N.R. Pamudurti, A. Ivanov, O. Bartok, M. Hanan, N. Evtantal, S. Memczak, N. Rajewsky, S. Kadener, circRNA Biogenesis Competes with Pre-mRNA Splicing, *Mol. Cell.* (2014) 1–12. doi:10.1016/j.molcel.2014.08.019.
- [11] W. Guo, S. Schafer, M.L. Greaser, M.H. Radke, M. Liss, T. Govindarajan, H. Maatz, H. Schulz, S. Li, A.M. Parrish, V. Dauksaite, P. Vakeel, S. Klaassen, B. Gerull, L. Thierfelder, V. Regitz-Zagrosek, T. a Hacker, K.W. Saupe, G.W. Dec, P.T. Ellinor, C. a MacRae, B. Spallek, R. Fischer, A. Perrot, C. Özcelik, K. Saar, N. Hubner, M. Gotthardt, RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing., *Nat. Med.* 18 (2012) 766–73. doi:10.1038/nm.2693.

Fig. 1. Identification and characterization of cardiac circRNAs. (A) Processing of heart samples from rats (neonatal, NRH, and adult, ARH, n=3 each), mouse (pressure overload, TAC, n=2 and sham, n=3) and humans (failing, HF, non-failing, HN, n=2 each) for circRNA identification. Deep sequencing reads (shown in gray) spanning the backsplice allow identification of circRNA candidates in RNA-Seq data. (B) Venn diagram of identified backsplices and their conservation. (C) Density histogram of relative enrichment of backsplice-spanning reads from highly expressed rat circRNA candidates (detected with ≥ 3 reads in either all NRH or all ARH mock treated samples, n=988) and regular splice junction spanning reads after RNase R digestion. Dashed line indicates the selected threshold of $2^{1.1}$ (n=767 at $\geq 2^{1.1}$). (D) Relative enrichment of reads from the known nuclear lncRNAs Malat1 and Meg3 and cytoplasm-enriched Gapdh in the respective compartments confirms successful separation of the two fractions. (E) CircRNAs are highly enriched in the cytoplasm of NRCMs compared to linear transcripts. (F) Enrichment of the known microRNA targets H19 and Pten (compared to Gapdh) in the Ago2-RIP, which is much lower in the Hsp90-RIP, confirms successful enrichment of microRNA targets. (G) CircRNAs show a similar extent of enrichment in the Ago2-RIP in NRCMs as do linear transcripts. (All p-values were calculated using a Mann–Whitney U test)

Fig. 2. Confirmation of select circRNA candidates and differential expression analysis. (A) Relative enrichment of some of the top expressed circRNAs in adult mouse heart after digestion with RNase R (normalized to linear transcripts), quantified with qPCR (n ≥ 3). Values >1 indicate increased stability compared to linear transcripts of the respective genes. (B) Sum of all reads falling on backsplices of high-confidence circRNAs in mock treated samples. High-confidence circRNAs were defined as those showing an average enrichment $\geq 2^{1.1}$ in RNase R treated samples. *p<0.05 in a two-way ANOVA analysis with Bonferroni's correction for multiple testing. (C) Overview of the differential expression of circRNAs from Suppl. Table 6A in adult and neonatal rat hearts (n=767). CircRNAs from select loci (indicated below) are highlighted in respective colors. Dashed lines indicate the interval of $\pm 2x$ fold change. (D) As in C for circRNAs in human samples (Suppl. Table 6C, n=1363). (E) As in C for circRNAs in mouse samples (Suppl. Table 6B, n=675). (F) Confirmation of differential expression of selected circRNAs in neonatal and adult rat hearts (n=6 each) and sham and TAC treated mouse hearts (n=4-5 each) using qPCR. *p<0.05 in a t-test for independent samples.

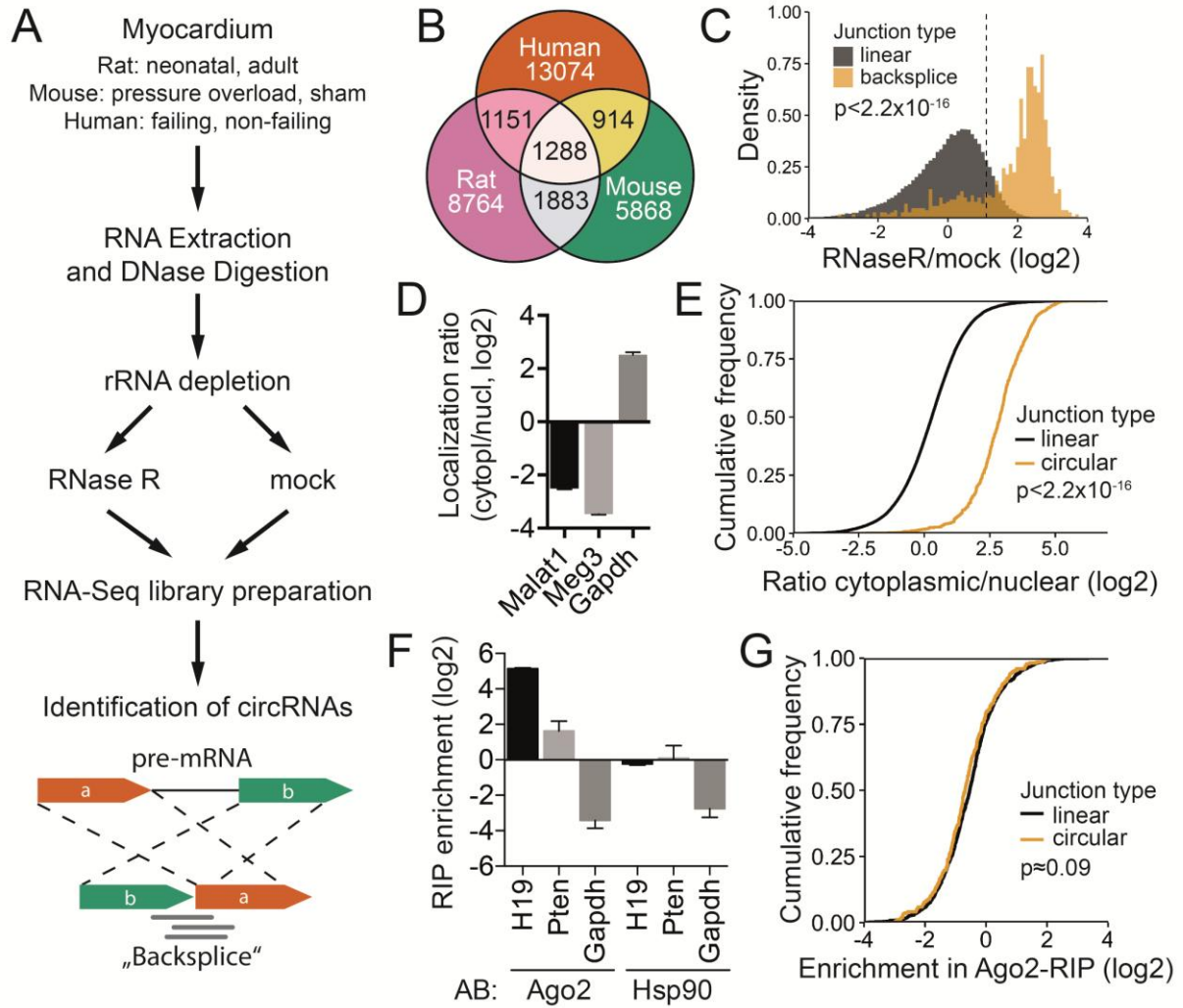


Figure 1

ACCE

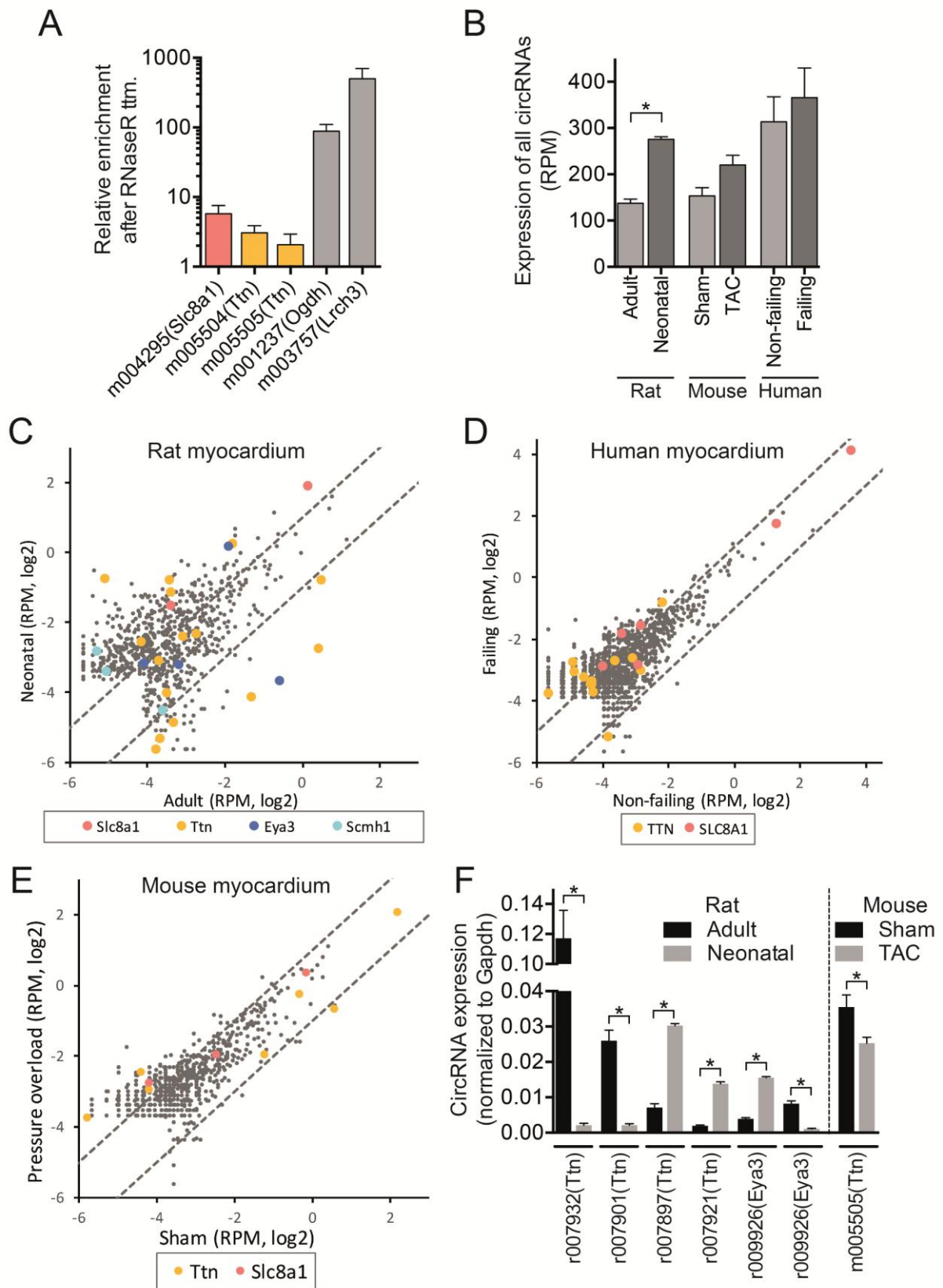


Figure 2