Array and Computational Methods to investigate a role for circRNAs in complex traits in the Hybrid Rat Diversity Panel.

Jennifer H. Mahaffey, Lauren A. Vanderlinden, Spencer Mahaffey, Paula L. Hoffman, Laura Saba, and Boris Tabakoff.
Skaggs School of Pharmacy and Pharmaceutical Sciences and School of Medicine, University of Colorado Anschutz Medical Campus, Aurora CO 80045

University of Colorado Anschutz Medical Campus

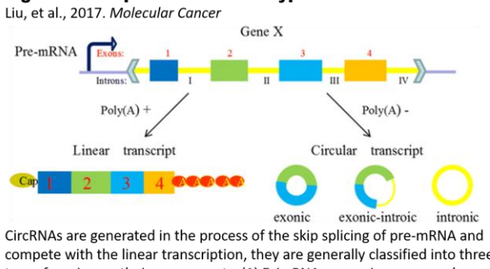
PhenoGen Informatics
The site for quantitative genetics of the transcriptome.

Abstract

The PhenoGen.org website was created to provide to the scientific community complete information on the genomes and the transcriptomes of several organs. This information is collected across a large panel of inbred and recombinant inbred (RI) rat strains (the Hybrid Rat Diversity Panel, HRDP) for use in quantitative genetic analysis of complex traits. The gathered transcriptome information contains quantitative data on both protein coding and non-coding RNA (lncRNA, miRNA, snoRNA and others). Recently, a new class of ncRNA, circular RNA (circRNA), has come to prominence as a regulator of physiology and behavior. circRNAs are thought to play important roles as miRNA sponges and through interaction with transcription complexes that affect gene expression. After deep sequencing RNA from various organs, we utilized the CIRI and CircExplorer2 circRNA prediction programs on ribosome-depleted total RNA-Seq data from BN-Lx and SHR strains, the parental strains of the HXB RI panel, one of the panels in the HRDP. Data were generated on brain, liver and left ventricle (LV) (n≥3). We identified 13,160 prospective circRNAs in brain, 2,934 in liver and 5,021 in LV. We therefore supplemented our RNA-Seq studies by analyzing the LV RNA using hybridization arrays designed to capture known and probable circRNA sequences (Arraystar, Inc). We quantified the expression of 12,132 circRNAs in at least 50% of samples tested. We found significant differential expression of 76 circRNAs between BN-Lx and SHR rats. The differential expression between these strains, coupled with the known functions of circRNAs, suggests that circRNAs may contribute to cardiovascular disease and other complex traits.

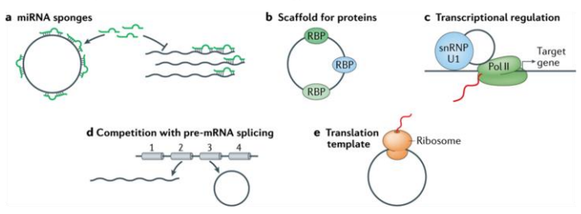
Background

Figure 1. Composition of three types of circRNAs.
Liu, et al., 2017. *Molecular Cancer*



CircRNAs are generated in the process of the skip splicing of pre-mRNA and compete with the linear transcription, they are generally classified into three types focusing on their components. (1) EcircRNAs comprise exons and are the most common circRNAs. (2) CiRNAs are composed by two or more connected introns and are detected rare in eukaryocyte. (3) ElciRNAs are circularized with introns 'retained' between the exons and play a part in gene regulation

Figure 2: Mechanisms of circular RNA functions.
Aufiero, et al., 2019. *Nature Reviews Cardiology*

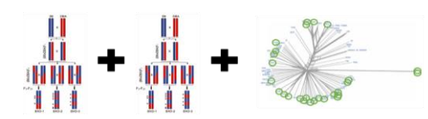


a | Circular RNAs (circRNAs) can function as microRNA (miRNA) sponges and reduce the capacity of miRNAs to target mRNAs. **b** | circRNAs can act as scaffolds for RNA-binding proteins (RBPs) to modulate their activity or localization. **c** | circRNAs can regulate transcription by interacting with the RNA polymerase II (Pol II) machinery and small nuclear ribonucleoprotein U1 (snRNP U1) in the nucleus. **d** | circRNA biogenesis affects the expression of the host gene by competing with linear mRNA splicing. **e** | circRNAs might associate with translating ribosomes and be subjected to translation in an internal ribosome entry site (IRES)-dependent and cap-independent manner. pre-mRNA, precursor mRNA.

Phenogen.org is a multifaceted and dynamic resource for the scientific community that aims to incorporate a tool set for circRNAs, an emerging class of non-coding RNAs (ncRNAs) generated primarily through "backsplicing" mechanisms. Three types of circRNAs: exonic, intronic and exonic-intronic (EIcircRNAs) (Figure 1), are thought to play important roles in genetic regulation in two main ways. Exonic circRNAs, located primarily in the cytoplasm, are mostly described to function as miRNA sponges, while nuclear interactions with transcription complexes that affect gene expression are more often described for EIcircRNAs and the rarer intronic RNAs (Figure 2). circRNAs have recently been linked to a number of cardiovascular diseases (Ref. 3).The normotensive BN-Lx and spontaneously hypertensive SHR rats used in our studies are widely used in hypertension research (Figure 3 & 4, Ref. 17). We pursued both computational detection methods on our existing RNAseq datasets in multiple tissues as well as circRNA arrays on left ventricle tissue from the parental strains of the HXB RI panel, one of two RI panels in the Hybrid Rat Diversity Panel (HRDP) (Figure 3). We hope to combine these two approaches with our broad, well established datasets to gain insight into the roles of circRNAs and eventually generate a circRNA pipeline and toolset for integration into Phenogen.org.

Figure 3. The Hybrid Rat Diversity Panel (HRDP)

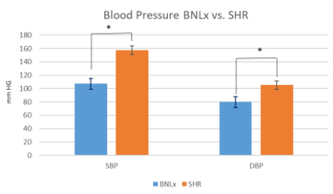
A Renewable Genetically Defined Population Consisting of Recombinant Inbred and Classical Inbred Rats for Cumulative Biology ●96 Strains ●Genetic and Omic-Analysis ●Ascertainment of Heritability ●Susceptibility/predisposition studies ●Mechanistic Genome Biology Studies ●Systems/Network Biology Studies ●Toxicologic Analysis ●Pharmaco/toxicokinetics ●Proof of Concept Studies for Medication Development



30 HXB/BXH Recombinant Inbred Strains
34 LEXF/FXLE Recombinant Inbred Strains
32 Classic Inbred Strains

Power to detect loci that contribute 10-20% to genetic variance
Power to detect significant genetic correlations as small as 0.28
Power to perform high resolution mapping (median haplotype block size = 225 Kb)

Figure 4. A Model for Hypertension



Systolic (Left) and Diastolic (Right) blood pressures are significantly different between brown normotensive and spontaneously hypertensive rats. * p= 3.79E-10 (Systolic) and p=3.57E-07 (Diastolic).

Methods

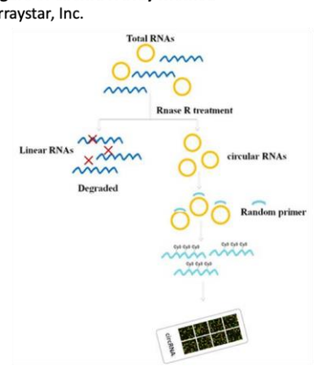
Table 1
circRNA Computational Detection Tools for RNASeq Datasets

Method	Approach	Genomic Origin	Dependencies	Performance
CIRI	Segmented read-based	Exonic, Intronic, Intergenic	bwa, perl	High Precision, Good Sensitivity
CirCExplorer2	Segmented read-based	Exonic, intronic	STAR, bedtools, python (pysam, numpy)	High Precision, Good Sensitivity, Intuitive Interface and Efficient

Adapted from Zeng, et al 2017 PLoS Comput Biol.

Telemetric daytime blood pressure (both systolic and diastolic blood pressure, SBP and DBP) was measured every day for 1 week on 7 rats/strain. Total RNA was isolated from BNLx and SHR Rat Brains, livers and left ventricles (n≥3) via the Qiagen RNeasy system. cDNA libraries were prepared from ribosome-depleted total RNA with the illumina Truseq Total Stranded Library preparation kit and sequenced by Genewiz on an illumina HiSeq sequencer. Raw reads were processed through CIRI, while trimmed and filtered reads were processed through CircExplorer2 (Table 1). Predicted circRNAs were merged based on circRNA annotation start and stop sites from each algorithm and number of exons (if applicable). The circRNA candidates were filtered based on detection in at least 2/3 samples in a strain. Due to the relevance of hypertension in our panel, we focused on the heart and supplemented our RNA-Seq studies with hybridization arrays designed to capture known and probable circRNA sequences (Arraystar, Inc). For the Arraystar arrays, Total RNA from left ventricle (n=3), prepared as described above, underwent RNAse L treatment to enrich for circRNAs, followed by labeling, and hybridization using conventional array methods by Arraystar, Inc (Fig 5). circRNAs that presented with significant differential expression by Arraystar, Inc (Fold Change ≥ 2, p ≤ 0.05) were predicted for miRNA binding sites in them by both Miranda and TargetScan algorithms (12, 16). To look at a subset not as dependent on fold change, we subsequently filtered circRNA candidates for an FDR <0.05. A genetic marker set (specifically SNPs) from the STAR consortium (http://www.snp-star.eu/; STAR Consortium et al 2008) was used for blood pressure quantitative trait loci (QTL) analyses. The QTLs were calculated using a weighted marker regression were the weight was defined as 1/within strain standard error. Genome-wide p-values were determined using permutation and 95% Bayesian credible intervals were estimated for both significant and suggestive (genome-wide p-value<0.05 and <0.63 respectively).

Figure 5. CircRNA Array Method
Arraystar, Inc.



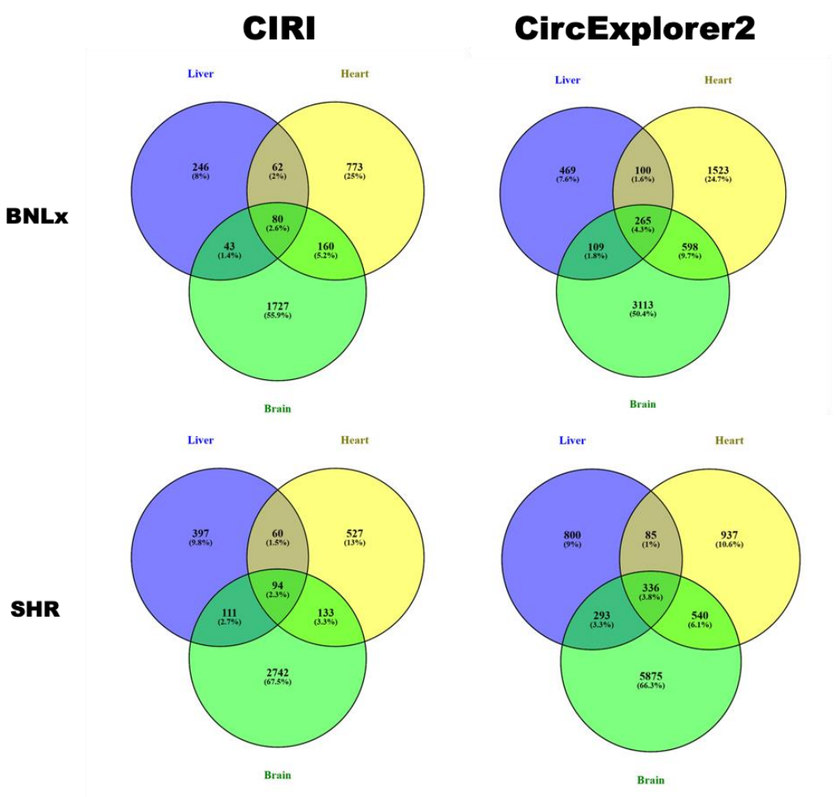
A random primer-based labeling system is coupled with RNase R-based sample pretreatment to efficiently remove linear RNAs, and specifically label circular RNAs.

Results

Table 2
circRNAs detected by Computational Methods

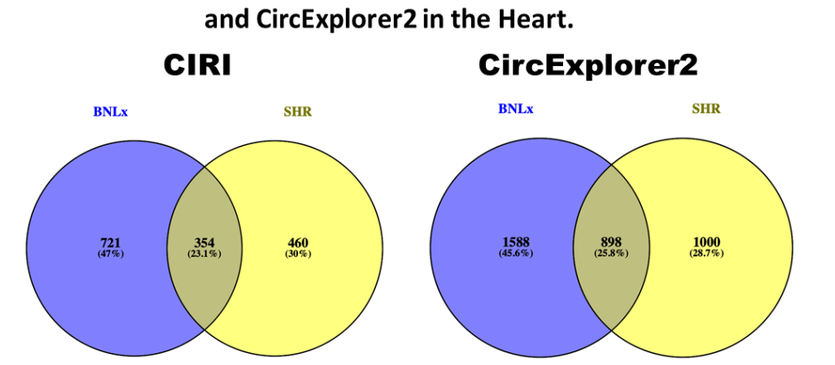
Tissue	CIRI	CircExplorer2
Liver	928	2006
Heart	1535	3486
Brain	4337	8823

Figure 6. circRNAs in Liver, Heart, and Brain.



The overlap of circRNAs detected by CIRI (left) and CircExplorer2 (right) in Liver (blue), Heart (yellow), and Brain (green) in BNLx (top) and SHR (lower) strains.
VENNS created with http://bioinfogp.cnb.csic.es/tools/venny/index.html

Figure 7. circRNAs identified in BNLx vs. SHR detected by CIRI and CircExplorer2 in the Heart.



Overlap of circRNAs detected between strains, by CIRI (left) and CircExplorer2 (right) in the left ventricle.
VENNS created with http://bioinfogp.cnb.csic.es/tools/venny/index.html

Table 3
circRNAs Detected by Array

circRNA_type variable	Number in Dataset Tested
Antisense	39
Exonic	9,729
Intergenic	249
Intronic	108
Sense Overlapping	2,009

Between the CIRI and CIRCExplorer2 predictions, we identified 13,160 prospective circRNAs in brain, 2,934 in liver and 5,021 in LV in the BNLx and SHR samples (Table 2). We also observed the overlap of circRNAs between tissues in these strains (Figure 6). Curiously, we were not able to detect any overlap of the predictions between CIRI and CircExplorer2. Subsequently, we focused on the heart and observed differences in the overlap of circRNAs between the BNLx and SHR strains (Figure 7). In our circRNA array, we quantified the expression of 12,132 circRNAs in at least 50% of samples tested in the left ventricle (Table 3) and found significant differential expression of 76 circRNAs between BN-Lx and SHR rats (Figure 8) and 12 of the circRNAs contained binding sites for miRNAs previously described in hypertension (Ref. 9-15). (Table 4). Additionally, after filtering independent of fold change we found 122 circRNAs with differential expression between strains (Figure 8). 15 of these circRNAs overlap with blood pressure QTL regions (Figure 8 & Table 5). Two of these circRNAs demonstrated differential expression in our array criteria, (thus miRNA binding sites were already mapped by Arraystar) with one, cirRNA_000109, containing a binding site for a miRNA associated with hypertension (Ref. 9). We also were able to detect this circRNA computationally with CircExplorer2 as well as several other circRNAs linked to hypertension (Table 6).

Figure 8. Differential Expression of circRNAs in SHR and BNLx Heart

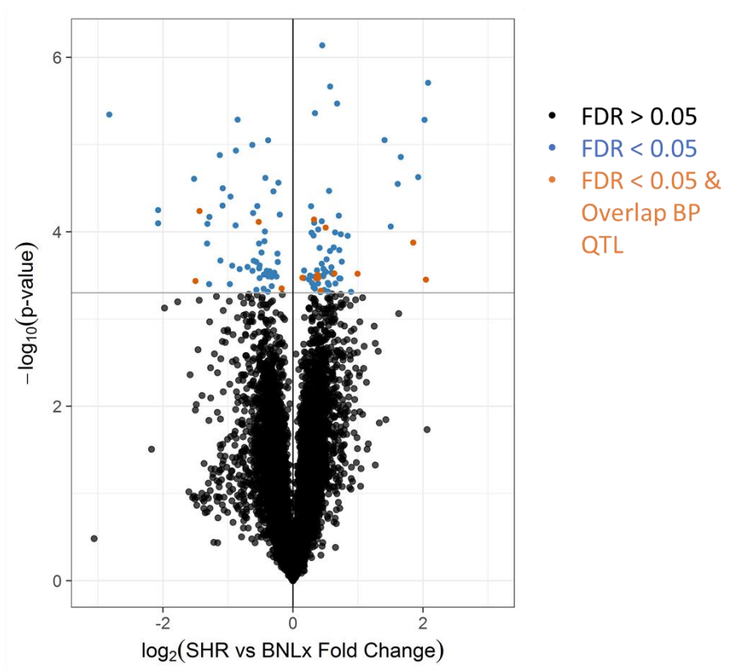


Table 4
circRNAs with differential expression associated with Hypertension by miRNA**

Regulation SHR vs. BNLx	circRNA	Chromosome	Strand	circRNA_type	Gene Symbol	miRNA Binding Sites
up	rno_circRNA_008880	chr2	+	exonic	Pde5a	rno-miR-362-5p
up	rno_circRNA_018022	chrX	-	exonic	Ophn1	rno-miR-495
up	mmu_circRNA_30337	chr10	-	exonic	Cacna1h	rno-miR-133c, rno-let-7g-5p
up	rno_circRNA_012827	chr5	-	sense overlapping	Zbtb8b	rno-miR-199a-5p
up	rno_circRNA_018020	chrX	-	exonic	Ophn1	rno-miR-21-5p
up	rno_circRNA_000109*	chr1	+	exonic	Ube3a	~~rno-miR-362-5p~~
down	rno_circRNA_014594	chr6	+	sense overlapping	LOC100359558	rno-miR-128-3p
down	rno_circRNA_013313	chr5	-	exonic	Gabbr2	rno-miR-199a-5p
down	rno_circRNA_008890	chr2	-	exonic	Ndst3	rno-miR-29b-5p
down	rno_circRNA_005195	chr14	-	exonic	Mtmr3	rno-miR-181b-5p
down	mmu_circRNA_21243	chr13	+	exonic	Tmco1	rno-miR-191b
down	rno_circRNA_002620	chr10	+	exonic	Spag9	rno-miR-29c-3p, rno-miR-29b-3p

*Overlaps with BP QTL ** Refs. 9-15

Table 5
15 circular RNAs physically located within Blood Pressure QTLs

circRNA	chromosome	strand	Gene Symbol
rno_circRNA_016461	chr8	-	Dis3l
rno_circRNA_005713	chr15	+	Pcdh17
rno_circRNA_005704	chr15	+	LOC686234
mmu_circRNA_41750	chr1	+	Sema4b
rno_circRNA_001530	chr1	+	Arid1b
rno_circRNA_016382	chr8	-	Neo1
mmu_circRNA_44470	chr8	+	Dpp8
rno_circRNA_005389	chr15	+	Cadps
rno_circRNA_000107	chr1	+	Ube3a
rno_circRNA_000727	chr1	+	Prkcb
rno_circRNA_000251	chr1	+	Abhd2
rno_circRNA_005708	chr15	+	LOC686234
rno_circRNA_000109*	chr1	+	Ube3a
mmu_circRNA_44300	chr8	-	Neo1
rno_circRNA_000459	chr1	-	Uvrag

*Overlaps with BP QTL

Table 6
Overlap of circRNA detection from Computational and Array Methods in Heart

circRNA	Regulation SHR vs. BNLx	Gene Symbol	miRNA Binding Sites	BNLx		SHR	
				CircExplorer2	CIRI	CircExplorer2	CIRI
rno_circRNA_008880	up	Pde5a	rno-miR-362-5p	✓		✓	
rno_circRNA_018022	up	Ophn1	rno-miR-495	✓			
rno_circRNA_000109*	up	Ube3a	rno-miR-362-5p	✓		✓	
rno_circRNA_005195	down	Mtmr3	rno-miR-181b-5p	✓		✓	
mmu_circRNA_21243	down	Tmco1	rno-miR-191b	✓		✓	
rno_circRNA_002620	down	Spag9	rno-miR-29c-3p, rno-miR-29b-3p	✓		✓	

*Differentially expressed and overlaps with BP QTL

Discussion and Conclusions

circRNAs have been demonstrated as important in cardiovascular disease and hypertension (3). We show here that computational methods may complement our existing RNAseq datasets in detecting circRNAs with biological relevance. We will need to further refine our approach to merge outputs from CIRI and CircExplorer2 as well as consider additional circRNA prediction methods. However, our circRNA array technologies coupled with analysis of our existing datasets may present new targets for understanding polygenic affects. In the case of hypertension, we identified differential expression in 76 circRNA candidates within Arraystar criteria and 122 circRNA candidates with an approach independent of fold change. 12 of these circRNAs may act as miRNA sponges implicated in hypertension (Refs. 9-15). In particular we identified circRNA_012827 which interacts with miR-199a-5p, a miR known to aggravate primary hypertension by promoting apoptosis and hindering autophagy (7). Also, we identified a circRNA (circRNA_018022) with a binding site for miR-495, a miR with a demonstrated role in vascular remodeling via cell cycle influence in mice with pulmonary hypertension (9). Many more circRNAs candidates and mechanisms may become clear as we predict miRNA binding sites for circRNAs differentially expressed independent of fold change through Miranda and TargetScan. Additionally, we would like to examine the overlap of these circRNAs in the brain, a tissue known to play an important role in hypertension (Refs.18-20). Overall, we were able to identify the potential of Phenogen.org coupled with computational resources on existing datasets to provide information on predicted circRNAs. Two probable candidates were identified that may function as either miRNA sponges or through mechanisms as yet undescribed, but possibly through other circRNA functions such as transcriptional regulation. Our overall goal is to use our existing datasets on Phenogen.org, complemented with circRNA detection algorithms and arrays, to provide a functional predictive and mapping circRNA tool for the HRDP available to researchers worldwide.

Future Directions

- Refine the merge and filter parameters for comparing CIRI and CircExplorer2 predictions and investigate additional computational tools.
- Comprehensively compare computational circRNA detection with circRNA Array detected circRNAs in the left ventricle.
- Identify miRNA binding sites for circRNAs with differential expression independent of fold change.
- Normalize CIRI and CircExplorer2 data and compare differential expression of circRNAs of interest from both the circRNA array and that overlap with the BP QTL.
- Identify overlap of differentially expressed circRNAs in heart with circRNAs in Brain and Liver tissues
- Collect data to complete the HRDP for heart to provide for expression QTL analysis.
- Develop and integrate circRNA tools and mapping for Phenogen.org

References

1. Liu, et al. 2017. Molecular Cancer
2. Aufiero, et al. 2019. Nature Reviews Cardiology
3. Fan, et al. 2017. BioMed Res Intl
4. Zeng, et al. 2017. PLoS Comput Biol
5. Gao, et al. 2015. Genome Biol.
6. Lusk, et al. 2018. Alcohol Clin Exp Res.
7. Saba, et al. 2017. Methods Mol Biol.
8. Harrall, K.K., et al. 2016. Mamm Genome
9. Han, et al. 2017. Medicine
10. Wei, et al. 2013. PLoS One.
11. Tian, et al. 2018. Exp Ther Med.
12. Lewis B.P. et al. 2005. Cell
13. Hromadnikova, et al. 2017. Int J Cardiol
14. Fu, et al. 2019. J Vasc Res.
15. Haddad, et al. 2019 Frontiers in Pharmacology
16. Miranda, et al. 2006. Cell
17. Doris, PA. 2017. Physiol Genomics
18. Hinton AO Jr et al. 2016 PLoS One.
19. do Carmo JM et al. 2017 Biochim Biophys Acta Mol Basis Dis.
20. Fujita M et al. 2016. Curr Hypertens Rep.
21. Oliveros, J.C. et al. 2007-2015. Venny. An interactive tool for comparing lists with Venn's diagrams. http://bioinfogp.cnb.csic.es/tools/venny/index.html

Acknowledgements

This project was supported by NIAAA (R24AA013162) and NIDA (P30DA044223).
Computing Resources: National Supercomputing Center & Dedicated Research Network, UNLV

Contact Information

Boris Tabakoff, PhD.
Skaggs School of Pharmacy and Pharmaceutical Sciences
University of Colorado Denver Anschutz Medical Campus
Email: Boris.Tabakoff@cuanschutz.edu
Phone: (303) 724-3668