A Gene Expression Atlas of Embryonic Cardiac Genes Generated by Developmental Stage-Specific Analysis

Dwayne Tally, Josh Soto, Laura Cochran, Michaela Ward, Hayden Fell, Garett Oxford, Joseph Dalloul, Andrew Williamson, Rusty Gonser, Jeff Kinne, and Kristopher Schwab

THE CENTER FOR GENOMIC ADVOCACY AT INDIANA STATE UNIVERSITY

Introduction

Heart development is a complex process that requires the step-wise activation of a complex genetic regulatory network. Perturbation of the cardiac regulatory network can result in severe developmental malformations of the cardiovascular system. Previous published high-throughput gene expression data sets of normal and abnormal mammalian heart development, provide the research community with the opportunity to re-investigate important biological processes and develop new hypotheses and experiments.

Method

In order to investigate previously published high-throughput gene expression data studying heart development, and develop a methodology for future gene expression experiments:

- We developed a mammalian embryonic heart gene expression atlas of genes that are enriched in the developing mouse heart.
- For these experiments, we identified a robust expression microarray (RMA) data set published by Li and colleagues (2014). This data set uniquely incorporates samples from the earliest stages of heart development to the mature adult heart, while also profiling mouse embryonic stem cells (mESCs) and the Embryonic Day 7 mouse embryo (Figure 1).

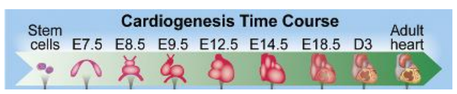

Figure 1. Cardiac developmental expression data set from Li et al. 2014 Physiol Genomics 46: 482-495.

Raw data was obtained from the Gene Expression Omnibus repository (GSE51483) and imported into GeneSpring software for analysis.

- Genespring performs a series of microarray analyses and normalization including: RMA analysis, log2 expression data transformation, and sample normalization.
- All samples were normalized to the E7.5 embryo sample group.
- Each of the embryonic heart stages (E8.5, E9.5, and E12.5) were compared to both the adult heart and mESC samples individually.
- Statistically significant genes were identified between the individual sample comparison using a statistical criteria (P <0.05 with a FDR) and a fold change cut-off (Fold change > or = 2X).

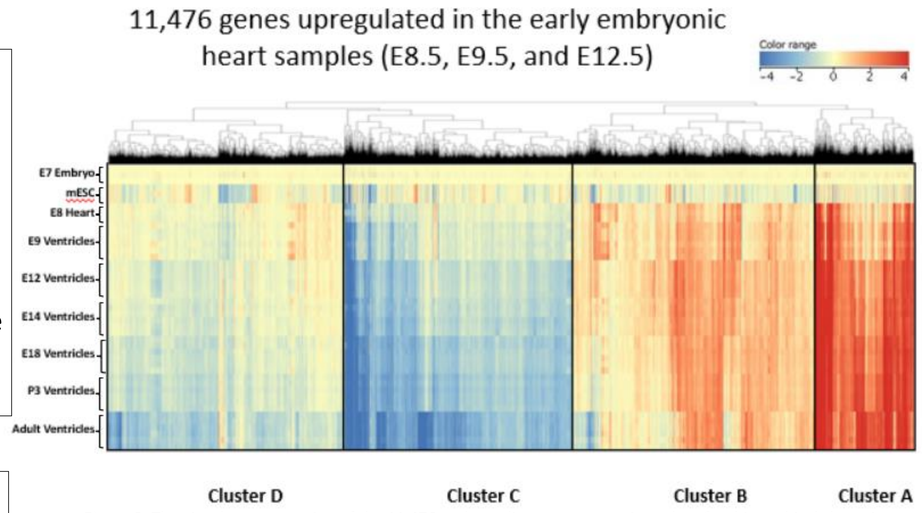

Figure 2: The visual representation of the 11,476 early cardiac gene expression atlas after hierarchical clustering analysis based on gene expression patterns.

	Cluster A		Cluster B		Cluster C		Cluster D	
	Symbol	Gene Name	Symbol	Gene Name	Symbol	Gene Name	Symbol	Gene Name
Signal Transduction	Tek	TEK receptor tyrosine kinase	Acvr1	activin A receptor, type 1	Acvr2b	activin receptor IIB	Adipor2	adiponectin receptor 2
	Ednra	endothelin receptor type A	Bmpr1a	bone morphogenetic protein receptor, type 1A	Aplnr	apelin receptor	Bmpr1a	bone morphogenetic protein receptor, type 1A
	Ramp2	receptor (calcitonin) activity modifying protein 2	Ednra	endothelin receptor type A	Lrp2	low density lipoprotein receptor-related protein 2	Cxadr	coxsackie virus and adenovirus receptor
	Tgfbr2	transforming growth factor, beta receptor II	Erbb2	erb-b2 receptor tyrosine kinase 2	Ptk7	PTK7 protein tyrosine kinase 7	Ncoa6	nuclear receptor coactivator 6
	Zfpm2	zinc finger protein, multitype 2	Akap13	A kinase (PRKA) anchor protein 13	Bmp2	bone morphogenetic protein 2	Ptpn11	protein tyrosine phosphatase, non-receptor type 11
	Mbd2	methyl-CpG binding domain protein 2	Gata4	GATA binding protein 4	Ctnnb1	catenin (cadherin associated protein), beta 1	Tgfbr1	transforming growth factor, beta receptor I
	Rbm20	RNA binding motif protein 20	Gata5	GATA binding protein 5	Cenpf	centromere protein F	Bicc1	BicC family RNA binding protein 1
	Zfpm2	zinc finger protein, multitype 2	Tab2	TGF-beta activated kinase 1/MAP3K7 binding protein 2	Dvl2	dishevelled segment polarity protein 2	Gata1	GATA binding protein 1
	Sparc	secreted acidic cysteine rich glycoprotein	Gjc1	gap junction protein, gamma 1	Firt3	fibronectin leucine rich transmembrane protein 3	Gata2	GATA binding protein 2
			Gnaq	guanine nucleotide binding protein, alpha q polypeptide	Lrp2	low density lipoprotein receptor-related protein 2	Ap2b1	adaptor-related protein complex 2, beta 1 subunit
			Heg1	heart development protein with EGF-like domains 1	Prkdc	protein kinase, DNA activated, catalytic polypeptide	Bmp4	bone morphogenetic protein 4
			Hspb11	heat shock protein family B (small), member 11	Rbp4	retinol binding protein 4, plasma	Ctnnb1	catenin (cadherin associated protein), beta 1
			Mbd1	methyl-CpG binding domain protein 1	Trp53bp2	transformation related protein 53 binding protein 2	Gnaq	guanine nucleotide binding protein, alpha q polypeptide
			Mbd2	methyl-CpG binding domain protein 2	Trp53	transformation related protein 53	Mbd3	methyl-CpG binding domain protein 3
			Pcsk5	proprotein convertase subtilisin/kexin type 5			Mib1	mindbomb E3 ubiquitin protein ligase 1
			Prkar1a	protein kinase, cAMP dependent regulatory, type I, alpha			Mapk1	mitogen-activated protein kinase 1
			Rps6ka2	ribosomal protein S6 kinase, polypeptide 2			Map2k1	mitogen-activated protein kinase kinase 1
			Sparc	secreted acidic cysteine rich glycoprotein			Shc1	src homology 2 domain-containing transforming protein C1
			Zfp36l1	zinc finger protein 36, C3H type-like 1				
Transcriptional Regulation	Symbol	Gene Name	Symbol	Gene Name	Symbol	Gene Name	Symbol	Gene Name
	Nkx2-5	NK2 homeobox 5	Gata4	GATA binding protein 4	Cited2	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2	Bcor	BCL6 interacting corepressor
	Hopx	HOP homeobox	Gata5	GATA binding protein 5	Gli3	GLI-Kruppel family member GLI3	Gata1	GATA binding protein 1
	Irx4	Iroquois homeobox 4	Nkx2-6	NK2 homeobox 6	Isl1	ISL1 transcription factor, LIM/homeodomain	Gata2	GATA binding protein 2
	Smyd1	SET and MYND domain containing 1	Smyd1	SET and MYND domain containing 1	Nkx3-1	NK3 homeobox 1	Smad2	SMAD family member 2
	Sox18	SRY (sex determining region Y)-box 18	Smyd2	SET and MYND domain containing 2	Senp2	SUMO/sentrin specific peptidase 2	Sox9	SRY (sex determining region Y)-box 9
	Tbx5	T-box 5	Sox4	SRY (sex determining region Y)-box 4	Smarca4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily	Axil2	additional sex combs like 2, transcriptional regulator
	Eng	endoglin	Tbx5	T-box 5	Tbx1	T-box 1	Bmp4	bone morphogenetic protein 4
	Hey2	hairy/enhancer-of-split related with YRPW motif 2	Tead1	TEA domain family member 1	Acvr2b	activin receptor IIB	Bmpr1a	bone morphogenetic protein receptor, type 1A
	Hdac9	histone deacetylase 9	Wt1	Wilms tumor 1 homolog	Bmp2	bone morphogenetic protein 2	Ctnnb1	catenin (cadherin associated protein), beta 1
	Mbd2	methyl-CpG binding domain protein 2	Acvr1	activin A receptor, type 1	Ctnnb1	catenin (cadherin associated protein), beta 1	Foxc1	forkhead box C1
	Myocd	myocardin	Bmpr1a	bone morphogenetic protein receptor, type 1A	Cenpf	centromere protein F	Id2	inhibitor of DNA binding 2
	Myl2	myosin, light polypeptide 2, regulatory, cardiac, slow	Eng	endoglin	Dvl2	dishevelled segment polarity protein 2	Id3	inhibitor of DNA binding 3
	Pam	peptidylglycine alpha-amidating monooxygenase	Erbb2	erb-b2 receptor tyrosine kinase 2	Foxc2	forkhead box C2	Jmjd6	jumonji domain containing 6
	Ripply3	ripply transcriptional repressor 3	Foxp1	forkhead box P1	Foxf1	forkhead box F1	Mbd3	methyl-CpG binding domain protein 3
	Zfpm2	zinc finger protein, multitype 2	Gjc1	gap junction protein, gamma 1	Id1	inhibitor of DNA binding 1	Mapk1	mitogen-activated protein kinase 1
			Hand2	heart and neural crest derivatives expressed 2	Ift74	intraflagellar transport 74	Map2k1	mitogen-activated protein kinase kinase 1
			Mbd1	methyl-CpG binding domain protein 1	Msx1	msh homeobox 1	Notch1	notch 1
			Mbd2	methyl-CpG binding domain protein 2	Pitx2	paired-like homeodomain transcription factor 2	Nfatc3	nuclear factor of activated T cells, cytoplasmic, calcineurin dependent 3
			Mical2	microtubule associated monooxygenase, calponin and LIM domain containing 2	Prkdc	protein kinase, DNA activated, catalytic polypeptide	Ncoa6	nuclear receptor coactivator 6
			Myocd	myocardin	Sall1	spalt like transcription factor 1	Osr1	odd-skipped related transcription factor 1
			Nfatc1	nuclear factor of activated T cells, cytoplasmic, calcineurin dependent 1	Sall4	spalt like transcription factor 4	Pitx2	paired-like homeodomain transcription factor 2
			Srf	serum response factor	Shh	sonic hedgehog	Pbrm1	polybromo 1
			Trps1	transcriptional repressor GATA binding 1	Trp53bp2	transformation related protein 53 binding protein 2	Shc1	src homology 2 domain-containing transforming protein C1
							Tgfbr1	transforming growth factor, beta receptor I
Muscle Contraction	Symbol	Gene Names	Symbol	Gene Name	Symbol	Gene Name	Symbol	Gene Name
	Myh7	myosin, heavy polypeptide 7, cardiac muscle, beta						
	Myh6	myosin, heavy polypeptide 6, cardiac muscle, alpha						
	Ttn	titin						
	Hopx	HOP homeobox						
	Hdac9	histone deacetylase 9						
	Tnni3	troponin I, cardiac 3						
	Tnnt2	troponin T2, cardiac						

Table 1: Using Gene Ontologies available through Mouse Genome Informatics we identified several genes involved in cardiac development and cardiac function. .

Cluster A			Cluster B			Cluster C			Cluster D		
Term	Count	P. Value	Term	Count	P. Value	Term	Count	P. Value	Term	Count	P. Value
cardiac muscle contraction	27	2.6E-22	protein binding	618	1.5E-24	poly(A) RNA binding	391	2.8E-118	poly(A) RNA binding	280	3.3E-43
sarcomere organization	19	1.4E-16	metal ion binding	470	1.1E-11	RNA binding	270	7.6E-78	mRNA processing	115	1.2E-31
regulation of blood pressure	22	3.6E-13	oxidation-reduction process	128	2.8E-10	mRNA processing	139	2E-52	RNA binding	194	5.8E-29
muscle contraction	20	4.4E-13	NADH dehydrogenase activity	11	0.000000039	cell cycle	199	2E-51	transcription, DNA-templated	353	7.1E-26
actin binding	51	5.5E-13	angiogenesis	56	0.000000044	nucleotide binding	407	4.5E-49	cell cycle	159	1.5E-25
angiogenesis	42	6.8E-13	transport	267	0.000000097	RNA splicing	114	7.5E-48	nucleotide binding	357	2.5E-25
heart development	44	7.8E-13	nervous system development	75	0.00000027	DNA binding	369	8.9E-39	RNA splicing	87	2E-24
ion channel binding	29	1.3E-12	cell migration	46	0.00000035	transcription, DNA-templated	374	1.8E-38	protein binding	628	3.4E-23
calcium ion binding	78	2.7E-12	heart development	57	0.0000004	cell division	132	2.3E-38	regulation of transcription, DNA-templated	392	1.3E-21
regulation of muscle contraction	13	1E-10	NADH dehydrogenase (ubiquinone) activity	16	0.0000001	mitotic nuclear division	110	2.2E-37	cell division	107	7.8E-21
regulation of heart rate	15	3.8E-10	response to hypoxia	45	0.0000011	regulation of transcription, DNA-templated	412	6.2E-33	mitotic nuclear division	86	2.5E-19
regulation of the force of heart contraction	12	2.1E-09	actin binding	66	0.0000017	cellular response to DNA damage stimulus	132	1.9E-32	DNA repair	88	3.6E-16
protein binding	265	4.1E-09	positive regulation of cell migration	46	0.0000021	rRNA processing	67	4.3E-32	nucleic acid binding	229	4.7E-16
cell adhesion	55	0.000000005	magnesium ion binding	45	0.0000029	helicase activity	67	2.7E-31	covalent chromatin modification	72	7.2E-13
oxidoreductase activity	61	0.000000038	cell adhesion	86	0.0000049	mRNA transport	54	4.1E-30	cellular response to DNA damage stimulus	98	7.7E-13
oxidation-reduction process	66	0.000000047	cellular response to hypoxia	28	0.0000095	nucleic acid binding	258	1.7E-29	chromatin binding	104	2.1E-12
collagen binding	16	0.000000063	positive regulation of transcription, DNA-templated	97	0.000011	DNA replication	63	5.3E-29	protein folding	44	7.8E-12
response to hypoxia	29	0.00000012	cartilage development	24	0.000013	chromatin binding	134	1.1E-28	transferase activity	241	7.1E-11
collagen fibril organization	13	0.00000014	mitochondrial electron transport, ubiquinol to cytochrome c	9	0.000015	DNA repair	104	3.4E-27	mRNA transport	34	7.5E-11
cardiac muscle tissue morphogenesis	9	0.00000023	kinase activity	108	0.000015	ATP binding	283	9.8E-25	mRNA binding	45	8.3E-11
regulation of cardiac muscle cell contraction	9	0.00000023	multicellular organism development	155	0.000016	covalent chromatin modification	84	5.3E-21	cadherin binding involved in cell-cell adhesion	68	4.1E-10
brown fat cell differentiation	12	0.00000027	transferase activity	207	0.000017	ribosome biogenesis	45	1.5E-20	DNA binding	286	4.3E-10
ventricular cardiac muscle tissue morphogenesis	11	0.00000033	intracellular signal transduction	72	0.000019	protein binding	582	3.2E-19	ATP binding	239	1.8E-09
regulation of heart rate by cardiac conduction	11	0.00000033	electron carrier activity	18	0.000021	chromosome segregation	43	1E-18	mRNA splicing, via spliceosome	36	5.9E-09
						negative regulation of transcription from RNA			positive regulation of transcription from RNA		
skeletal muscle tissue development	15	0.00000038	sprouting angiogenesis	13	0.000022	polymerase II promoter	152	2.2E-17	polymerase II promoter	166	0.000000037

Table 2: Functional annotation enrichment analysis of the four clusters (A - D) utilizing based upon the from DAVID

Discussion

To visualize the gene list, the 11,476 early cardiac gene expression atlas underwent hierarchical cluster analysis based on gene expression patterns throughout all heart development (Figure 3).

- Hierarchical clustering revealed a total of 4 unique expression signatures at the highest organizational level (clusters A-D) with interesting developmental expression patterns.
- These cluster were then subjected to further investigation identifying genes involved in cardiac development according to Gene Ontologies (Table 1) and functional annotation based upon the bioinformatics suite, DAVID (Table 2).

Future Investigations

The development of this cardiac gene expression atlas will assist our future research investigations involving:

- Investigation of additional cardiac gene expression data sets
- Gene expression profiling of novel cardiac genes
- Cardiac gene discovery
- Transcriptome profiling of genetic loss-of-function studies of mammalian cardiac development

References and Acknowledgments

Busser, B. W., Lin, Y., Yang, Y., Zhu, J., Chen, G., & Michelson, A. M. (2015). An Orthologous Epigenetic Gene Expression Signature Derived from Differentiating Embryonic Stem Cells Identifies Regulators of Cardiogenesis. *PLoS ONE*, *10*(10), e0141066. http://doi.org/10.1371/journal.pone.0141066

Li, Xing & Martinez-Fernandez, Almudena & A Hartjes, Katherine & Kocher, Jean-Pierre & M Olson, Timothy & Terzic, Andre & J Nelson, Timothy. (2014). Transcriptional Atlas of Cardiogenesis Maps Congenital Heart Disease Interactome.. Physiological genomics. 46. 10.1152/physiolgenomics.00015.2014.

This work was supported by The Center for Genomic Advocacy (Schwab), American Heart Association (Ahmad), University Research Council (Schwab), and the NIH grant for BD4ISU (Gonser) and graduate assistants Nykara Brown and Naureen Aslam