

REFINED METHODS FOR TRANSCRIPTOME ANALYSIS IN CARDIOMYOPATHY

A. Cullen, (MD, SOM), M. Taylor, L. Mestroni, S. Graw, D. Grine, G. Storm, S. Gao, and S. Chen, Human Medical Genetics and Genomics, Cardiovascular Institute and Adult Medical Genetics Program, Department of Cardiology, and MD Program, University of Colorado, Denver, CO

Heart Failure (HF):

- Affects ~6.2 million Americans (2013-2016), with projected 46% prevalence increase 2012-2030^{1,2}
- Significant cause of death
- High financial cost to HC system^{3,4}

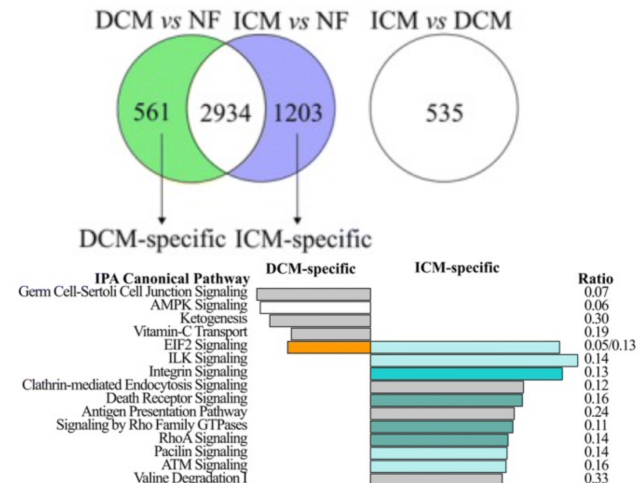
Impetus for Study:

- Currently HF treatment uses a “common pathway” assumption
- Want personalized gene approach^{3,4}

Context:

- Sweet et al. (2018) analyzed RNA expression profiles for non-failing (NF) v. Ischemic Cardiomyopathy (ICM) v. Dilated Cardiomyopathy (DCM) tissue samples from human hearts
 - Searching for Differentially Expressed Genes (DEG's) → pathway analysis

Results from Sweet et al. (2018):



My Contribution:

- Matlab script that can be easily adjusted to alter DEG selection criteria
- RNA-seq from 13 ICM, 37 DCM, and 14 NF human LV samples; 57974 genes

My Adjustable Algorithm:

- Rid data of missing/nonsensical entries
 - Can require $\geq \frac{1}{2}$ of entries to exceed 0 for each group (NF/DCM/ICM)
- Flexibly set row mean cutoffs, requiring at least one (or both) of two row means to exceed the cutoff
- Sift for genes with change in mean expression in one of two ways:
 - Δ mean > threshold (i.e. > 5)
 - |Fold change (FC)| b/w row means > threshold (i.e. 1.5)
- T-test on $\log_2(\text{RPKM} + \text{cutoff})$
- Benjamini-Hochberg procedure to sift out false positives
- Work in progress to account for covariates of age & sex
- Pathway analysis in IPA

Effort to first replicate 2018 DCM results:

- Some dissimilarity b/w sets; working to incorporate covariate adjustment
- 543 DEG's compared to 561; many genes are same, some different

Top Canonical Pathways		
Name	p-value	Overlap
Hypoxia Signaling in the Cardiovascular System	8.97E-04	9.5 % 7/74
Angiotensin Signaling	9.72E-04	9.3 % 7/75
Integrin Signaling	1.83E-03	5.6 % 12/213
NRF2-mediated Oxidative Stress Response	2.17E-03	5.8 % 11/189
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	2.17E-03	5.8 % 11/189

DCM Gene Results:

Expr Log Ratio	
Molecules	Expr. Value
PROS2P	↑ 1.668
BEX1	↑ 1.385
GSTM1	↑ 1.336
DIRAS3	↑ 1.285
OMD	↑ 1.253
MYOC	↑ 1.223
SEMA3B-AS1	↑ 1.168
miR-126	↑ 1.038
RN7SL2	↑ 0.967
VLDLR-AS1	↑ 0.965

Expr Log Ratio	
Molecules	Expr. Value
SPP1	↓ -2.520
RP1_317E236	↓ -2.493
AC055036	↓ -2.042
RPL23AP7	↓ -1.842
AC05943.2	↓ -1.719
COXMD3-BM1	↓ -1.632
RP11-49K24.6	↓ -1.447
RPL17-C18orf32	↓ -1.313
RP11_603J249	↓ -1.200

Effort to replicate 2018 ICM results:

- Closer similarity, 1204 v. 1203 DEG's

Top Canonical Pathways		
Name	p-value	Overlap
Clathrin-mediated Endocytosis Signaling	9.65E-06	12.4 % 24/193
EIF2 Signaling	4.09E-05	11.2 % 25/224
Integrin Signaling	4.97E-05	11.3 % 24/213
Actin Cytoskeleton Signaling	5.10E-05	11.0 % 25/227
Antigen Presentation Pathway	5.44E-05	23.1 % 9/39

ICM Gene Results:

Expr Log Ratio	
Molecules	Expr. Value
ICAM1	↑ 3.131
SOC3S	↑ 3.055
FOS	↑ 2.619
C11orf96	↑ 2.442
ACKR1	↑ 2.120
NR4A1	↑ 2.011
NNMT	↑ 1.895
TNC	↑ 1.875
SNORA48	↑ 1.738
MTLXP1	↑ 1.738

Expr Log Ratio	
Molecules	Expr. Value
HMOCS2	↓ -2.911
C1orf105	↓ -1.972
RP11_1081M51	↓ -1.710
LINC02137	↓ -1.344
SDHAP3	↓ -1.243
RP1_206D155	↓ -1.199
PIGA	↓ -1.156
BCLAF1P1	↓ -1.149
AKR1C1/AKR1C2	↓ -1.115
MYO9A	↓ -1.085

Discussion and Next Steps:

- Flexible algorithm to adjust parameters
- Repeat 2018 work, expand w/ new DEG sets, and verify results w/ larger HF set