Advances in Bioresearch
Adv. Biores., Vol 11 (5) September 2020: 177-180

Print ISSN 0976-4585; Online ISSN 2277-1573
Journal's URL:http://www.soeagra.com/abr.html
CODEN: ABRDC3
DOI: 10.15515/abr.0976-4585.11.5.177180

Advances in Bioresearch

CASE REPORT

A bioinformatics approach for deciphering the pathogenic processes of Myocardial Infarction

[1]**Aditya Saxena**
[1]Department of Biotechnology, Institute of Applied Sciences & Humanities, GLA University, Mathura, UP, India

ABSTRACT
Myocardial infarction (MI), also known as a heart attack is a prevalent cardiovascular problem which is generally found secondary to obesity, and Type 2 Diabetes. Its pathogenesis is attributed to arthrosclerosis in coronary artery that escalate into atherothrombosis and eventual onset of infarction. Prognosis of MI is difficult and the first attack set the stage for functional limitations of heart. It is therefore a pressing need to elucidate its pathogenic mechanisms at the molecular level to identify drug targets, and biomarkers for MI. To meet this end, various DNA-microarray studies have been conducted and gene expression datasets are available at NCBI-GEO database. However, microarray technology suffers with inherent problem of low-reproducibility due to study design, and biological variability. Integrative analysis of multiple datasets is an attractive and cost effective approach to obtain a robust gene-signature of the disease that can be utilized to unravel new biological insights using pathway- or network-based enrichment approach. In this study, I have carried out a statistical meta-analysis across three microarray studies, comprising a total of 157 individuals. These gene expression profiles were extracted from circulating endothelial cells and peripheral blood. Series matrix files for these datasets were downloaded from GEO, normalized, filtered for low quality gene expression measures, and their probe ids were collapsed to official gene symbols. After this preprocessing, datasets were checked for integrity and adjusted for batch effects. Obtained meta-gene signature was then search against MSigDB knowledge base through Gene Set Enrichment Analysis (GSEA). I found enrichment of various immune- and inflammation- as well as cardiac fibrosis-related cellular pathways and literature search expose some of the known pathophysioloigcal aspects of MI.
Keywords; *Mayocardial Infarction, GSEA, Meta analysis, microarray, GEO*

Received 21.04.2020 Revised 01.06.2020 Accepted 21.07.2020

How to cite this article:
Saxena A. A bioinformatics approach for deciphering the pathogenic processes of Myocardial Infarction. Adv. Biores. Vol 11 [5] September 2020: 177-180.

INTRODUCTION

Myocardial Infarction is the most common coronary heart disease that account to over 32.4 million cases worldwide every year [1]. Its pathogenesis involves coronary artery occlusion due to rupture of atherosclerotic plaque which results in necrosis of myocardium due to low oxygen supply. Despite significant improvements in the clinical management of its primary factors such as diabetes, hypertension, obesity etc., we are still unable to accurately predict the occurrence of acute myocardial infarction. It is therefore necessary to elucidate its pathophysiology at the molecular level to identify novel disease mechanisms and biomarkers for its early detection.

Various global gene expression studies have been conducted and datasets are available at NCBI-GEO database [2]. However, microarray technology suffers with inherent problem of low-reproducibility due to study design, and biological variability [3]. Integrative analysis of multiple datasets is an attractive and cost effective approach to obtain a robust gene-signature of the disease that can be utilized to unravel new biological insights using pathway- or network-based enrichment approach.

MATERIAL & METHODS

Relevant MI studies were searched using 'Datasets2Tools' repository which indexes human gene expression studies from GEO database [4]. A total of three studies, comprising 157 gene expression

datasets were found relevant. Table 1 present the included studies and number of datasets in each of them.

Table 1. Microarray samples used in meta- analysis.

S. No.	Series	Tissue	Platform	Place	Number of Samples	
					Normal	MI
1	GSE48060	Blood cells	Affymetrix Human Genome U133 Plus 2.0 Array	Mayo Clinic, Health Sciences Research, Rochester, USA	21	31
2	GSE97320	Peripheral Blood		The Department of Cardiology, China–Japan Union Hospital, Jilin University,Changchun,China	3	3
3	GSE66360	Circulating Endothelial Cells		The Scripps Research Institute,La Jolla,USA	50	49

These studies profiled gene expression from peripheral blood of these individuals who have encountered atleast one heart attack or were had normal cardiac function.

Suresh *et al.* [5] conducted a microarray study to predict long term clinical outcomes following first-time myocardial infarction (AMI) by profiling gene expression from peripheral blood. Sample from MI patients were drawn after 48 hours post MI. these patents were monitored for 18-months and five out of twenty two individual had encountered second MI. They identified decreased epithelial-to-mesenchymal transition, and modulation of cholesterol transport genes.

Another gene expression study was conducted to identify differentially regulated genes in the peripheral blood of Northeast Chinese Han people who have encountered their first myocardial infarction. The study remains unpublished.

Muse *et al.* [6] isolated circulating endothelial cells from patients experiencing acute myocardial infarction and healthy cohorts, and measured gene expression using the HG-133U Plus 2 microarray platform [6]. Their objective was to identify a robust gene signature for the diagnosis of MI.

Series matrix files of these studies were downloaded using Bioconductor package – GEO query in R environment. All metadata associated with each file was removed and sample levels were annotated with each sample.

For conducting meta analysis, a web-based tool 'NetworkAnalyst' [7] was used that allow to integrate multiple gene expression datasets using their summary-level data (i.e., P-values, fold changes or effect sizes, vote counting, and direct merging. To adjust batch effect among these datasets, it provides function ComBat in the SVA R package. Fisher's method was used on gene-level log-transformed P-values to obtain a meta-signature of MI studies.

To determine the biological functions related to the results of meta-analysis in human COPD datasets, Gene Ontology (GO) analysis using ClueGO was carried out [8]. For pathway enrichment analysis, gene-wise combined Z-scores of meta-signature were applied to Gene Set Enrichment Analysis (GSEA) program [9] using the GSEA Pre-ranked method that uses MSigDB gene sets as its knowledge base. I used canonical pathways subset of curated gene set C2 comprising 2199 gene sets which includes genes sets from Reactome, KEGG, Biocarta, and NCI-Nature Pathway Interaction Database.

RESULTS

Network Analyst identified a Meta signature of 496 genes ($p < 0.05$). Among the meta-signature, 224 genes were up-regulated (Z-score > 0) and 272 genes were down-regulated (Z-score < 0).

CluGO-based enriched various terms related to biological processes that converge into various immune- and inflammatory processes ($p < 0.05$): cell chemotaxis, cellular defense response, inflammatory response, interleukin-1-mediated signaling pathway, intracellular signal transduction, leukocyte activation, macrophage activation, negative regulation of mature B cell apoptotic process, pattern recognition receptor signaling pathway, platelet degranulation, positive regulation of immune system process regulation of cell activation, regulation of cellular component movement, regulation of response to external stimulus, response to other organism, and response to wounding. (Figure 1)

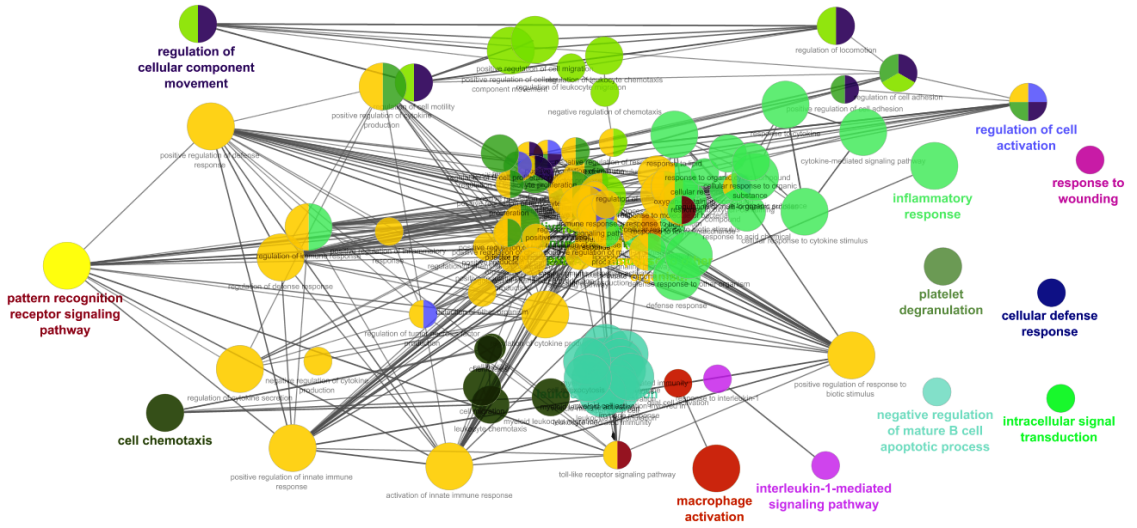

Figure 1. ClueGO functionally grouped networks of GO – BP terms enriched by MI meta signature

This analysis clearly indicates that MI-implicated narcosis evidently set a systemic inflammatory state that induces cardiac tissue modeling, derangement of cardiac function and subsequent escalation toward cardiac arrest.

Results of GSEA-based enrichment analysis indicate that meta-analysis provides valuable information for generalizing the results of multiple studies (Figure 2).

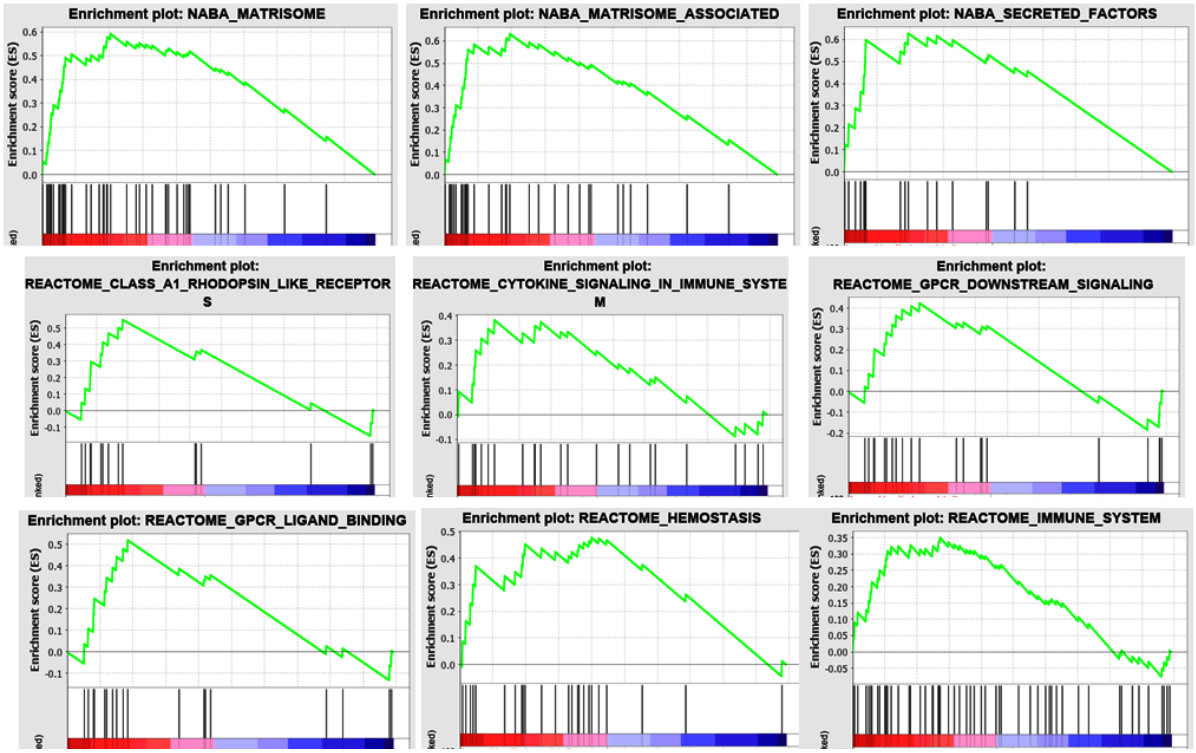

Figure 2. GSEA plots for enriched gene set by MI meta signature

MSigDB contain an inventory of Extracellular matrix proteins – *NABA Matrisome*. Enrichment of matrisome, and its associated proteins by MI meta signature further supports that MI leads to extensive cardiac modeling [10]. Two other enriched gene sets: *Cytokine signaling* and *Immune System* gene sets are related to inflammation due to necrosis.

Another gene sets *Reactome Class A1 Rhodopsin like Receptors, Reactome GPCR Ligand Binding*, and *Reactome GPCR Downstream Signaling* are also implicated in cardiac biology and many of the G-protein

coupled receptors are targeted for the treatment of cardiovascular disease (CVD), including hypertension, arrhythmias and heart failure.

CONCLUSION

Present bioinformatics study was conducted to ascertain pathophysiology of MI through integration of multiple gene expression studies. A meta-signature was derived and functional analysis of genes constituting the signature enriched relevant processes. The method is generic and can be applied to decipher the pathophysiology of other biomedical settings.

REFERENCES

1. Krishnan M. (2012). Coronary heart disease and risk factors in India – On the brink of an epidemic? Indian Heart Journal. ;64(4):364–7.
2. Clough E, Barrett T. (2016). The Gene Expression Omnibus Database. Methods in Molecular Biology Statistical Genomics. ;:93–110.
3. Ramasamy A, Mondry A, Holmes CC, Altman DG.(2008). Key Issues in Conducting a Meta-Analysis of Gene Expression Microarray Datasets. PLoS Medicine. ;5(9). 9-11
4. Torre D, Krawczuk P, Jagodnik KM, Lachmann A, Wang Z, Wang L, et al. (2018). Datasets2Tools, repository and search engine for bioinformatics datasets, tools and canned analyses. Scientific Data. 2018;5(1). 1-8
5. Suresh R, Li X, Chiriac A, Goel K et al. (2014). Transcriptome from circulating cells suggests dysregulated pathways associated with long-term recurrent events following first-time myocardial infarction. J Mol Cell Cardiol Sep;74:13-21.
6. Muse ED, Kramer ER, Wang H, Barrett P et al. (2017). A Whole Blood Molecular Signature for Acute Myocardial Infarction. Sci Rep 25;7(1):12268.
7. Xia, J.; Gill, E.E.; Hancock, R.E.W. (2015). NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. Nat. Protoc.,10(6), 823-844.
8. Bindea G., Mlecnik B., Hackl H., Charoentong P., Tosolini M., Kirilovsky A., Fridman W.H., Pagès F., Trajanoski Z., Galon J., Clue G.O. (2009). A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics. 25(8):1091–1093.
9. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102: 15545-15550
10. Foster SR, Roura E, Molenaar P, (2015). Thomas WG. G protein-coupled receptors in cardiac biology: old and new receptors. Biophysical Reviews. ;7(1):77–89.