Bioinformation Analysis of Key Genes and Pathways of Acute Myocardial Infarction to Predict Potential Medicine

Zien Xu
Xi'an Jiaotong University, Xi'an 710061, China.

Abstract: This study was aimed to retrieval differentially expressed genes and biological signaling pathways of Acute Myocardial Infarction (AMI) based on integrated bioinformatics analysis[1]. Using Gene Expression Omnibus (GEO) dataset was used to download mRNA expression profile GSE66360 containing 50 healthy cohorts and 49 patients experiencing Acute Myocardial Infarction to study the the potential mechanism and predict medicine target in Acute Myocardial Infarction (AMI). The consistently differentially expressed genes(DEGs) were identified,and functional annotation and pathway analysis of these genes were excavated with GO, KEGG and Reactome. The protein-protein interaction network(PPI) of DEGs was created with Cytoscape and STRING to screen the hub genes including TNF, IL1B, FN1, CD4, PLEK, JUN, TLR4, FOS, LRRK2, TLR2. On the ground of discovered information, the potential drugs were calculated. The functional genes enrich mainly in Cytokine-cytokine receptor interaction, osteoclast differentiation, neutrophil degranulation and signaling by interleukins and so on. This study digs the pathological process of AMI and may aid in learning the deeper molecular mechanism of AMI and predict potential drugs of this disease.

Keywords: Acute Myocardial Infarction; Bioinformatics Analysis; Gene Expression Omnibus

1. Introduction

Although the burden cardialvascular disease (CVD)brings to us has been reduced sharply, CVD still accounts for a third of all deaths in China and worldwide each year.This desease can be the leading cause of death all over the world[2]. The most basic cause of myocardial infarction is coronary atherosclerosis. Plaques in the coronary artery cause stenosis of the coronary artery lumen, which leads to insufficient blood supply to the myocardium. One or more coronary atherosclerotic plaques can lead to coronary stenosis. On this basis, once the blood supply is sharply reduced or interrupted, it will cause serious myocardial infarction. Acute ischemia lasting for a long time reaches about half an hour, and AMI will occur.

Genome-wide Association Studies (GWAS) refers to the research on the potential relationship between genes and diseases that is carried out in multiple centers, large samples and repeated verification on the whole genome level. The Increased sample size of genome-wide association studies (GWAS) will revitalize the statistical power to identify the missing causal variants and may highlight additional disease mechanisms[3].There has already been experiments to prove that certain genes are related to the pathology of AMI. However, the exact mechanism is still not clear.

To further learn and dig the mechanism of the disease with gene relations and find potential small molecular medicine to accurately target certain signaling pathways related to Acute Myocardial Infarction, inhibiting the action of related signaling molecules, reducing inflammatory response and myocardialcyte damage after AMI may help improve the prognosis of myocardial infarction treatment.Bioinformatics was employed in this study to screen disease up and down regulated genes in order to obtain a better understand of the micromechanism and predict target drug to improve the clinical treatment.

2. Research methods and materials

2.1 Statistic download with Gene Expression Omnibus

The Gene Expression Omnibus (GEO) database is a gene expression database created and maintained by the National Center for Biotechnology Information (NCBI) in the United States. It collects high-throughput gene expression data submitted by research institutions around the world[4]. GEO (https://www.ncbi.nlm.nih.gov/geo/) was used to select samples composed of blood samples from 49 acute myocardial infection subjects and 50 healthy subjects without AMI. After downloading the GSE66360 series matrix files and GPL570, the perl script was employed to integrate and obtain a gene expression profile.

2.2 Evaluation of differentially expressed genes

To gain an insight of the differently expressed genes, R limma package of bioconductor was used to get the gene expression profile standaraized and adopted values of log2. Differential expressed genes (DEGs) were obtained from acute AMI and healthy subjects using P-value <0.05 and logFoldChange> 1[5]. and the related visualized heat map and gene loci map were drawn.

2.3 Protein and protein interaction and hub genes screening

Protein–protein interactions (PPIs) are of vital importance for obtaining mechanistic insights into the functional organization of the proteome. The resolution of PPI functions can do us an aid in the identification of novel diagnostic and therapeutic targets with medical utility, which is significant in learning the potential relationship of certain molecules[6]. DEGs were used in the STRING database to reveal their relationship with a minimum required interaction score>0.4. In order to make the data based on the number of connections between DEGs, the hub genes were determined by the betweenness centrality.

2.4 Enrichment analysis with GO/Rectome/KEGG

In order to make the differential genes functioning mechanism clear, the gotten differently expressed genes were employed to do GO, KEGG and Rectome enrichment analysis under the asistent of DAVID Bioinformatics Resources (LHRI) (https://david.ncifcrf.gov/). KEGG is a database for systematic analysis of gene function and genomic information, which helps researchers to connect key genes and expression information as a whole network for research. The database provides integrated metabolic pathway queries, including almost all possible metabolic pathways such as nucleotides, amino acids, carbohydrates, etc. Therefore, it is a powerful database for whole genome and metabolic pathway exploration.

2.5 Small molecular drugs prediction with Connectivity Map (https://clue.io/query)

Connective map was employed to analyse dig small molecular drugs. Connectivity Map is an online analysis tool used in the predction process of small molecular drugs[7], this comprehensive, large-scale perturbation database comprehends 1.5 million gene expression profiles from cultured human cells, can be used to identify potential therapeutic targets or drugs for the submitted gene signature[8].

3. Results

3.1 DEGs identification

Acocording to the threshold value P<0.05 and | logFoldChange|>1, differently expressed genes were identified in GSE66360 gene expression profile data.351 DEGs were found, containing 289 upregulated genes and 62 downregulated genes. NR4A2, S100A12, ITLN1, CSTA, CCL20, FCER1G, IL1B, TREM1, PTX3, NFIL3A was top 10 upregulated genes and TSIX, XIST, CTD-2528L19.6, GIMAP7, GIMAP4, CCR2, CRTAM, B3GALT2, GIMAP6, CCR5 was top 10 down regulated genes. Heat map and volcano map were respectively drawn.

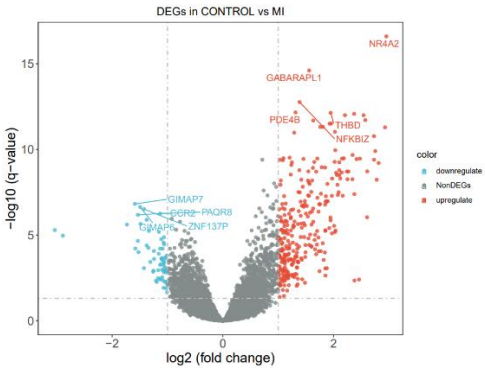

Figure 1. The distribution volcanic map of 351consistene DEGs of AMI (red:upregulated;blue:downregulated)

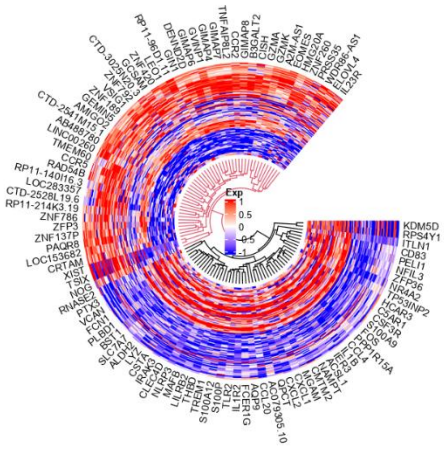

Figure 2. The distribution circle heat map of 100 consistent DEGs containing top 50 upregulated and 50 downregulated of AMI

(red:upregulated;blue:downregulated)

3.2 GO annotation analysis of DEGs

The 351 DEGs were analyzed by GO which divides genes functions into three main parts including molecular function(MF), cellular component(CC) and biological process(BP).The most enriched term was response to positive regulation of cytokine production in BP,tertiarygranule in CC and pattern recognition receptor MF,which proved a series of lipid metabolism disorders and inflammatory response.

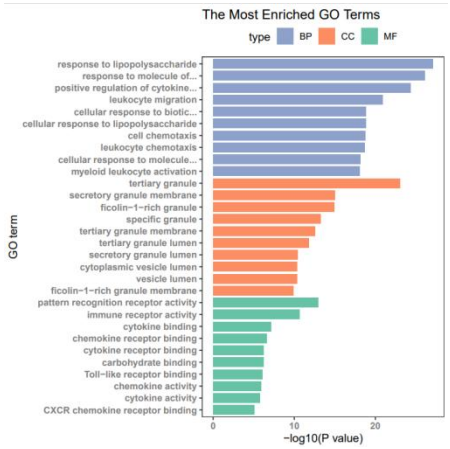

Figure 3. The most enriched GO terms classified by BP,CC and MF

3.3 KEGG and Rectome pathway analysis of DEGs

On the basis of KEGG database, 10 significant enriched cell signaling pathways were found out. The 10 most significantly enriched signaling pathways were Cytoline-cytokine receptor interaction, Osteoclast differentiation, Chemokine signaling pathway, Leishmaniasis, Amoebiasis, Hematopoietic cell lineage, Rheumatoid arthritis, Malaria, NOD-like receptor sigmaling pathway.

The Rectome enrichment showed that the top 10 pathways were Neutropjil degranulation, Signaling by interleukins, Toll-like Receptor Cascades, Interleukin-10 signaling, MyD88:MAL(TIRAP) cascade initiated on plasma membrane, Toll like Receptor (TLR) TLR6:TLR2 cascade,TLR2,TLR4 and Interleukin-4 and interleukin-13 signaling.

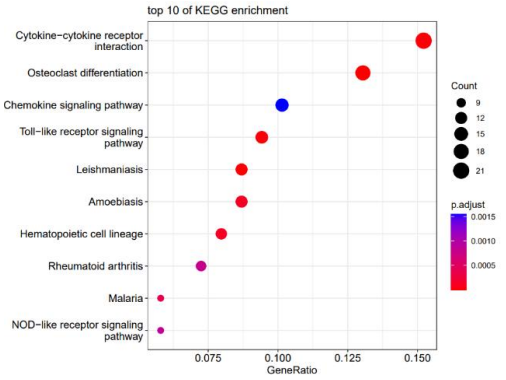

Figure 4. Top 10 enriched pathways in Rectome pathway

3.4 PPI network construction and identification of hub genes

351 DEGs in AMI were employed to builds a protein-protein interaction with STRING, as it presents, there are 307 nodes and 2374 edges in the PPI network. It has an average 15.5 node degree and 0.474 avg.local clustering coefficient. In order to gain an insight into the complex network and find the hub genes. A useful method calles betweenness centrality (BC) was adopted in the process of visual graph construction to identify core nodes and edges[9]. In gragh theory,intermediate centrality is one of the measurement criteria for network graph centrality based on the shortest path. For a fully connected network graph, there is at least one shortest path for any two nodes. In an unweighted network graph, the shortest path is the sum of the number of edges included in the path, while in a weighted network graph, the shortest path is the sum of the weights of the edges included in the path. The betweenness centrality of each node is the number of times these shortest paths pass through that node.The ten DEGs with the highest number of closely related genes are regarded as heb genes, they are tumor necrosis factor (TNF), interleukin 1, beta (IL1B), cluster of differentiation 4 (CD4), Pleckstrin (PLEK), Jun proto-oncogene (JUN), Toll-likereceptor4 (TLR4), Fructooligosaccharide, (FOS), Leucine-rich repeat kinase 2 (LRRK2), toll-like receptor 2 (TLR2). These genes may play an important administrative function in the pathological process of AMI.

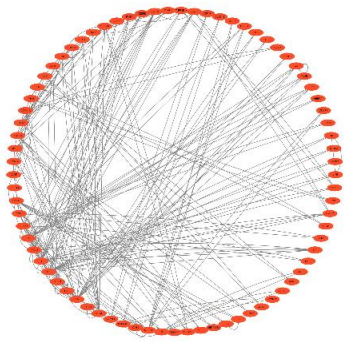

Figure 5. The PPI image with CB norm

3.5 Small molecular drugs prediction with Connectivity Map (https://clue.io/query)

The main feature of iron death is the lipid peroxidation of intact cell membrane, which leads to the deposition of lethal dose reactive oxygen species, which is an iron dependent process and can be inhibited by iron chelators[10].The Cmap analyse could see a decrease in GPX4 protein levels in myocardial cells and early induction of lipid peroxide accumulation, which can be Fer-1 inhibition[11], proves a decrease in GPX4 expression during MI, causing iron death and inducing myocardial damage.

4. Discussion

The two modes of myocardial cell death in Acute Myocardial Infarction are inflammation centered necrosis and gene controlled apoptosis under ischemia and hypoxia, respectively[12]. Apoptosis is a programmed cell death that can be regulated at the genetic level. Understanding and regulating certain signaling pathways involved in genes is an effective means for us to understand and reduce the occurrence of apoptosis during myocardial infarction[12].

Under the circumstance of the statistically high enrichment pathways of TLRs and NF- κ B. A particular focus on these two pathways was paied. It was estimated that asiatic-acid (AA) may potentially function in inhibiting the release of inflammatory factors by Cmap. In a new study, the expression of SIRT3 which can activate PPARγ, thus protecting the myocardium was significantly improved after administration of AA to AMI rats. Under this circumstance, AA development can be considered for SIRT3/β- catenin / PPARγ specific inhibition of signaling pathways. Simlilar experiment observed that after oxalic acid intervention lastlt after 4 weeks, TUNEL staining results showed a significant reduce of apoptotic cardiomyocytes, with orderly arrangement of cells and muscle bundles.At the same time,the complete cell morphology he infiltration of inflammatory cells was relatively low. The myocardial apoptotic cells were significantly reduced in each dose group of asiatic acid, and the levels of LDH, CK-MB, MDA in serum of rats were significantly reduced. The levels of SOD in serum were significantly increased, and the expression of Nrf2, HO-1 mRNA and protein was significantly increased. NF-κ B mRNA and protein expression were significantly reduced. These changes suggest that asiatic acid may regulate Nrf2/HO-1 and NF- κ B pathway to reduce impairment Acute Myocardial Infarction.

The CXCL16 was highlighted statistically in the upregulated genes analysis, this gene belongs to the CXC chemokine subfamily, which includes transmembrane domains and mucin like structures. It is characterized by transmembrane and there are two forms of dissolution[14]. CXCL16 can exacerbate myocardial infarction by three aspectsor progresses: Firstly, it functions as chemokines and can bind to phosphatidylserine -oxidized low-density lipoprotein scavenger receptors, participating in inflammatory reactions.Secondly, CXCL16 produces chemotaxis on T lymphocytes and promote T lymphocytes to metastasize like ischemic sites, thereby severer the development of atherosclerosis and myocardial infarction. The third CXCL16/CXCR6 axis can increase tumor necrosis factor-α(TNF-α) and promote myocardial fibrosis process, and this also echoes the HUB GENE with the highest centrality in the number of PPI intermediaries.

TREM1 gene is a member of the Ig superfamily and is expressed in myeloid cells. TREM1 can up regulate the expression of cell surface activation markers while stimulating the release of proinflammatory chemokines and cytokine,which can also be observed in the KEGG signaling pathway enrichment in this study.

The two genes have a relatively strong correlationship with each other, and it was also clear that there were other inflammatory related genes upregulated by TREM1.So it is significant to find a ideal dtug like AA to reduce the inflammatory symptoms and inhibit the apoptosis of myocardialcytes. In subsequent experiments, physical experiments will also be conducted to detect the expression of related genes and measure the protective effect of drugs on myocardial cell apoptosis through pathological changes in apoptosis markers such as mitochondrial morphology changes.

5. Conclusion

Acute Myocardial Infarction is the most common in Europe and the United States. Every year there is about 1.5 million incidence rate in the United States, while the incidence rate in China is also on the rise. At least 500000 new cases occur every year, and at least 2 million patients now, which accounts for a significant proportion of the population. It is a major factor affecting the health of Chinese people and cannot be ignored. Therefore, it is urgent to do a good job in disease prevention, treatment and prognosis rehabilitation.

There still exists some shortcomings in this study such as lack of survival curve analysis and animal experiments makes it impossible to accurately apply the analysis results and drug predictions to patients with Acute Myocardial Infarction. In addition, the side effects of this drug and conditions for drug application also need to be explored, which will be the next step of work. So in the new era of healthy China and the healthy world, preventing myocardial infarction plays a crucial role. When symptoms such as arrhythmia, shortness of breath, body pain, or chest tightness occur, one should immediately seek treatment from the hospital, regardless of the age group.

References

[1] Zhao XY, et al. (2018).Identification of key genes and pathways associated with osteogenic differentiation of adipose stem cells.. Journal of cellular physiology(12).

[2] GonzálezRamírez Javier, et al.(2022). Acute Myocardial Infarction and Periodontitis: Importance of Awareness and Prevention in Latin America. Applied Sciences(6).

[3] Wightman DP., et al. (2021). A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. Nature Genetics(9).

[4] Clough, E., & Barrett, T. (2016). The Gene Expression Omnibus Database. Methods in molecular biology (Clifton, N.J.), 1418, 93–110.

[5] Liu, Z., et al. (2023). Identification of key genes and small molecule drugs in osteoarthritis by integrated bioinformatics analysis. Biochemistry and biophysics reports, 34, 101450.

[6] Yang, L., et al. (2022). Identification of protein-protein interaction associated functions based on gene ontology and KEGG pathway. Frontiers in genetics, 13, 1011659.

[7] Lim DW, Choi MS & Kim SM.(2023). Bioinformatics and Connectivity Map Analysis Suggest Viral Infection as a Critical Causative Factor of Hashimoto's Thyroiditis. International Journal of Molecular Sciences(2).

[8] Justin Lamb, et al. (2006). The Connectivity Map: Using Gene-Expression Signatures to Connect Small Molecules, Genes, and Disease. Science(5795).

[9] Daniel C, et al. (2021). Fast cluster-based computation of exact betweenness centrality in large graphs. Journal of Big Data(1).

[10] Scott JD, et al. (2012). Ferroptosis: An Iron-Dependent Form of Nonapoptotic Cell Death. Cell(5).

[11] Park Tae-Jun, et al. (2019). Quantitative proteomic analyses reveal that GPX4 down regulation during myocardial infarction contributes to ferroptosis in cardiomyocytes. Cell death & disease(11).

[12] Liu, H., et al. (2018). NR4A2 protects cardiomyocytes against myocardial infarction injury by promoting autophagy. Cell death discovery, 4, 27.

[13] Shimamoto Akira, et al. (2006). Inhibition of Toll-like receptor 4 with eritoran attenuates myocardial ischemia-reperfusion injury. Circulation(1 Suppl).

[14] Hammond MD, et al(2014). CCR2+ Ly6C(hi) inflammatory monocyte recruitment exacerbates acute disability following intracerebral hemorrhage. The Journal of neuroscience : the official journal of the Society for Neuroscience(11).