Spring 4-17-2023

# Identification of Hub Clacification Proteins in Plaque formation in Atherosclerosis using Multi Omnics and Network Based Approach

Charmy Shah

# Investigating the Role of Hub Calcification Proteins in Atherosclerosis via Integrated Multiomics and Network-Based Approach

by

Charmy Dharmesh Shah

Under the Direction of Chunying Li, M.D., Ph.D.

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Interdisciplinary Studies

in the Institute for Biomedical Sciences at

Georgia State University

2023

ABSTRACT

**Background:** Atherosclerosis is a chronic inflammatory condition that affects the arteries and is distinguished by the formation of plaques. This can lead to constricted blood flow and conceivably lethal cardiovascular events. The disease progressed over the years and is impacted by various risk factors, including high blood pressure, diabetes, smoking, obesity, and high cholesterol. Atherosclerosis contributes to a quarter of the world's deaths globally. Therefore, in this article, we put together an impartial, systematic, and organized bioinformatics system to understand the biomarkers and possible regulatory targets involved in treating Atherosclerosis and plaque development to understand better the disease's pathogenesis, which can help reduce the global risk of cardiovascular disease.

**Material/Methods:** Microarray datasets GSE43292 and GSE28829 were obtained from the Gene Expression Omnibus database (GEO). Limma analysis was done in R version 4.2.2 using the Bioconductor package version 3.16 by dividing GSE28829 into two groups, advanced and early plaque, to find the DEGS, whereas GSE43292 was divided into two groups, namely advanced plaque, and microscopically intact tissue to find the DEGS. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of these modules were performed by Enrichr. Differentially expressed genes (DEGs) were mapped into calcification-related genes using the STRING database; two PPI networks were generated. These PPI genes were used to extract. Transcription factors (Tf) from CHea3 and MiRNA from MIRwalk were sorted

about Atherosclerosis to do the FFL analysis. Further experimental studies were done to verify the key miRNAs, i.e., that showed a strong association with the gene and transcription factor (TF).

**Results:** The genes were found in the DEGS list in GSE28829 and GSE43292. The PPI genes were identified as. GO and KEGG enrichment analysis revealed that genes in GSE28829 were enriched. At the same time, the genes in GSE43292 were enriched. The miRNA and Tf for FFL analysis were identified Let-7e-5p (a member let-7 family) as the top-ranked miRNAs regulating the feed-forward loop (FFL) that includes SOX7 and PECAM1 for GSE28829, whereas we found that TF: GATA2, miRNA: miR-550a-3p and Gene: MEF2C as the regulating FFL loop for GSE43292.

**Conclusions:** Overall, the DEG analysis of GSE28829 and GSE43292 provides insights into the molecular mechanisms underlying the progression of Atherosclerosis from early to advanced stages. The upregulation of genes involved in an extracellular matrix organization and adhesion and the downregulation of genes involved in immune and inflammatory responses suggest that changes in the plaque microenvironment may contribute to disease progression. These findings have important implications for the development of new therapeutic strategies for the treatment of Atherosclerosis. We have identified Let-7e-5p (a member let-7 family) as the top-ranked miRNAs regulating the feed-forward loop (FFL) that includes SOX7 and PECAM1 for GSE28829, whereas we found that TF: GATA2, miRNA: miR-550a-3p and Gene: MEF2C as the regulating FFL loop for GSE43292 related to advanced plaques/early plaque about calcification for

GSE43292 and GSE28829, which may be a potential therapeutic target to treat atherosclerotic plaques in the future.

INDEX WORDS: Atherosclerosis, Bioinformatics, DEGS

Investigating the Role of Hub Calcification Proteins in Atherosclerosis via Integrated Multiomics and Network-Based Approach

by

Charmy Dharmesh Shah

Committee Chair: Dr. Chunying Li

Committee: Dr. Ping Song

Electronic Version Approved:

Office of Academic Assistance – Graduate Programs

Institute for Biomedical Sciences

Georgia State University

May 3, 2023

## ACKNOWLEDGEMENTS

I extend my heartfelt gratitude to my committee chair and principal investigator Dr. Chunying Li, for his immense wisdom and support throughout my project. I want to thank Dr. Ping Song for agreeing to join my Thesis Defense Committee. This project would not have been possible without the continued guidance of my mentor Shaligram Sharma, and for his infinite patience in teaching me and being there for me. A special thank you to all the other members of the lab for their constant help and support. Finally, I thank my family and friends for pushing me to perform my best.

TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1.INTRODUCTION

## 1.1 Atherosclerosis

Atherosclerosis (AS) is a chronic and progressive disease of the arterial wall that leads to plaque accumulation and narrowing of the arterial lumen. Atherosclerosis is a significant contributor to cardiovascular disease worldwide; the prevalence of it is increasing due to bad lifestyle choices and an increase in the age of the population. According to the World Health Organization (WHO), approximately 31% of world deaths globally can be accounted to cardiovascular disease, with atherosclerosis being the underlying reason in most cases (4). Several risk factors are responsible for the development of atherosclerosis, like smoking, obesity, diabetes, hypertension, and dyslipidemia. Clinically, the disease progresses through the continuous development of arterial wall lesions due to lipid retention in the intima matrix, which causes arterial lumen stenosis and obstruction of blood flow or plaque rupture that can lead to atherosclerotic thromboembolism, stroke, and peripheral artery disease (1,2,3).

## 1.2 Plaque formation

The process of plaque formation is progressive and complex, involving multiple cellular and molecular functions. The formation of plaque begins with the accumulation of low-density lipoproteins (LDL) in the intima of the artery due to some injury of the endothelial layer of the artery; this leads to the expression of Reactive oxygen species (ROS) by the endothelial cells in the intima and the LDL gets oxidized to oxidized LDL(OX-LDL). This leads to the expression of receptor proteins to which the monocytes attach and go inside the intima, called diapedesis. This monocyte transforms into Macrophages and engulfs the OX-LDL to make foam cells (6). The foam cells, along with other inflammatory cells, make more proinflammatory cytokines and chemokines that lead to the recruitment of more immune cells; hence, the immune response

amplifies, leading to even more inflammation at the site of infection/disruption. This activation of endothelial cells and inflammatory response is the first step in developing atherosclerotic plaque, after this, a fatty streak develops, the earliest stage of plaque formation. In intermediate lesions, there is an extracellular matrix (ECM) accumulation, collagen secreted by Smooth muscle cells (SMCs), and its proliferation. This leads to the thickening of the arterial wall. The more collagen accumulation along with calcification, the more stable the plaque is. After intermediate plaque, it progresses to the fibrous plaque by more accumulation of SMCs, ECM, and inflammatory cells; the plaque contains a fibrous cap with a necrotic core which contains foam cells, cholesterol crystals, calcium deposition, and inflammatory cells (5).

**1.3 Calcification**

Calcification of atherosclerotic plaque plays a significant role in the progression of atherosclerosis, which can affect the stability of plaque and the chances of its rupture. Vascular calcium detection by computed tomography can be used as a subclinical marker of atherosclerosis. The fact that coronary artery and aortic valve calcification often occur concurrently reinforces the suggestion that atherosclerosis is a precursor to these variants of calcification. The deposition of calcium in plaque leads to an increase in the stiffness of the arteries and hardening of the plaque; this makes the plaque even less suspectable to rupture. There are two types of calcifications of atherosclerotic plaque: intimal and medial calcification. Intimal calcification is a common type characterized by the deposition of calcium phosphate crystals in the necrotic core and fibrous cap. Medial calcification can be characterized by the accumulation of calcium phosphate crystals in the media of the arterial wall and cause arterial stiffening. The calcification in atherosclerosis is intimal; the calcification process is complex and is controlled by various factors like inflammatory cytokines, growth factors, and ECM. The

expression of these is controlled by TFs like Runx2 and Sox9, which play a significant role in osteoblast differentiation and bone formation (9). Several studies have found that the expression of specific proteins is associated with calcification progression in the plaque. Some of these proteins are osteopontin, osteocalcin, and bone sialoprotein. The formation of plaque calcification is also associated with the expression of various miRNAs (10). One such example is miRNA mir-204, which inhibits the expression of RUNX2 and stops plaque calcification (8). This advanced plaque formation with calcification is irreversible, and no drugs are available to treat this plaque formation (7). So, there needs to be the development of new technology to treat and detect the calcification of plaque.

**1.4 Multiomics and Network-Based Approach**

The use of high-throughput sequencing technology in recent years has been an effective tool in understanding the underlying genes and biological processes during atherosclerotic plaque formation. Transcription factors (TFs) and miRNAs are major in regulating gene expression. TFs are proteins that control the rate at which genes are transcribed into mRNA by binding to specialized DNA sequences. The miRNAs are small RNA molecules that bind to mRNA and suppress or rapid protein synthesis. This complex interaction between miRNAs and TFs can either be inhibitory or synergistic. This synergy creates a feedback loop that balances proper gene expression levels and regulation of TF and miRNAs. This relationship between miRNAs and TFs creates Feed-forward loops (FFLs) or feedback loops (FBLs), in which TFs regulated miRNA, or miRNA inhibits the expression of TFs; both together coregulate the face of the standard target, which is a gene. Three types of FFLs categorized based on the master regulators TF-FFL, miRNA-FFL, and composite FFL (11,12).

The expression of genes is disrupted if this balance between TFs and miRNAs is dysregulated, as seen in many diseases, like cardiovascular disease, cancer, and other diseases. Recently many studies using bioinformatics and gene expression profiling found various genes, Transcription factors (TFs). Few studies have recognized the role of miRNA and TFs in regulating gene expression in atherosclerosis and calcification. MiR-29a and miR-29b inhibited the calcification of VSMCs by suppressing the expression of a disintegrin and metalloproteinase with thrombospondin motifs-7 (ADAMTS-7) (13).In another study, they found that miR-125b downregulation can facilitate the calcification of VSMCs by targeting Ets1, a transcription factor protein.RUNX2 a TF is found to be the critical osteogenic regulator for osteoblast differentiation and chondrocyte maturation by repressing MYOCD-induced differentiation and promoting the calcification in plaque (14).Transcription Factor MAFF (MAF Basic Leucine Zipper Transcription Factor F) was found to Regulate Atherosclerosis Inflammation and Cholesterol Metabolism (15), in which, gene expression profiling is more helpful in revealing the disease mechanism and its progression. Previous studies have shown that miRNAs and TFs may play a role in treating Atherosclerosis plaque formation and calcification; as the miRNAs and TFs have stability, specificity, and detectability, they can become excellent biomarkers. Therefore, identifying miRNA and TFs networks will take us one step closer to identifying the gene regulatory pathways involved in atherosclerosis and potential therapeutic targets for plaque formation and calcification.

## 2. Materials and Methods

### 2.1 Study Design

The series matrix data and phenodata for GSE28829 and GSE43292 downloaded from the Gene Expression Omnibus (GEO) database (16,17). The overall research design flow chart is

shown in Figure 1. First, gene expression profiles of plaques in GSE28829 were used to identify DEGs in control and Advanced plaque. GSE43292 were used to identify DEGs in early and later atherosclerotic plaque. Both the data sets were normalized, duplicates were removed, after that the variances of every gene in all samples were calculated and sorted in descending order, and the top 50% of genes were selected as candidates. After getting the DEGs list for GSE ids we filtered them using the Calcification genes. Then these genes for the GSE ids was used to construct a PPI network. The PPI network and performed molecular complex detection (MCODE) analysis using the Cystoscope software to obtain the important subnetworks of these genes (18). The hub genes from the most important subnetwork were further analyzed using Enrichr to do GO and KEGG enrichment analysis (19). Furthermore, analysis of genes was done using three-node miRNA FFL, revealing the key TF, miRNA, and gene associated with the development of atherosclerosis plaque. The study identified targets for plaque formation, which can be used as an effective biomarker tool against plaque formation.

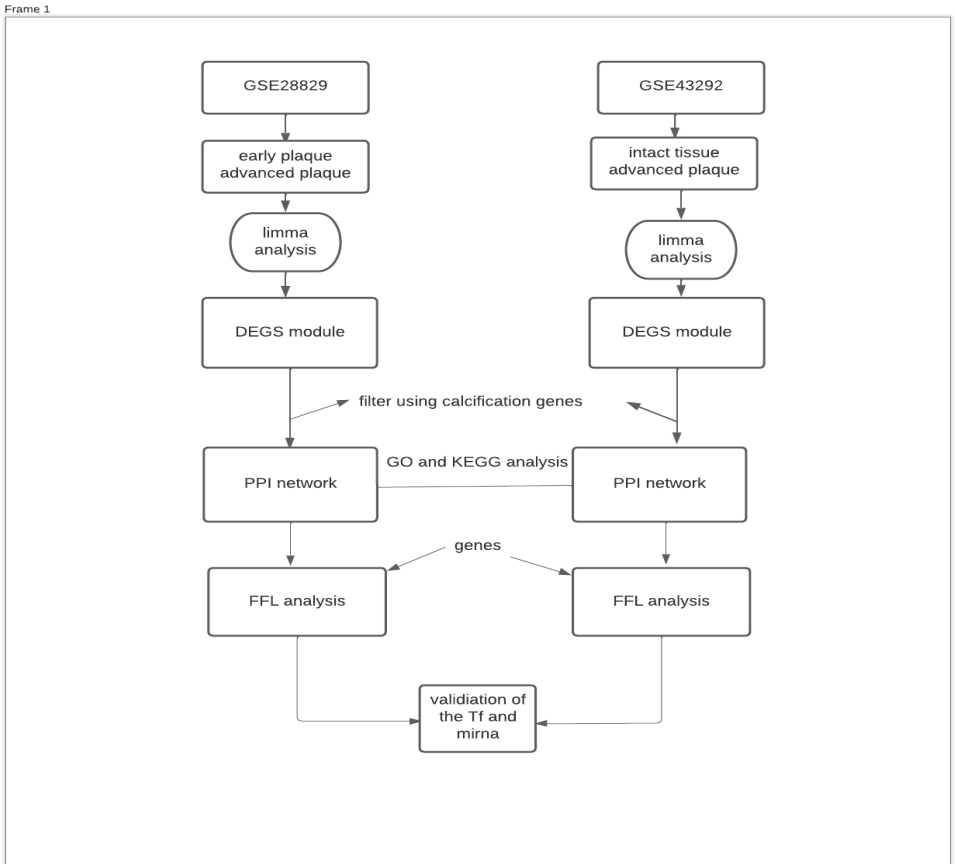



***Figure 1: The Flowchart for Investigating the role of hub calcification proteins in Atherosclerosis via integrated multiomics and network-based approach.***

## 2.2 Atherosclerosis mRNA Expression Profile Extraction (DEA)

We accessed the NCBI-GEO (https://www.ncbi.nlm.nih.gov/geo/, accessed on 2 December 2022) utilizing "Atherosclerosis" as appropriate keyword for Atherosclerosis associated mRNA expression profile retrieval. The search results were furthermore streamlined in compliance with the following selection criteria:(1) The datasets of "expression profiling by array" type and with Atherosclerosis in "Homo Sapiens" as the host;(2) The dataset(s) must have both raw and processed microarray data;(3) The dataset require both the control and infected sample type; (4) The data must have been submitted within the range of the last ten year (i.e., 2012-2022) (5) The data set should have a sample size of 25 or more.

Studies lacking case reports, review articles, abstracts, non-human samples, and cell-line-based experimental design were excluded.

Based on the above-mentioned criteria we selected two data sets for analysis GSE28829 and GSE43292.GSE28829 is a gene expression dataset titled "Gene expression in early and advanced atherosclerotic plaque from human carotid." This study aimed to identify genes and pathways involved in progressing atherosclerotic plaques in humans from early to advanced stages. The dataset includes gene expression data from carotid atherosclerotic plaques of 29 samples. The sample was collected. Out of 29 collected samples, 13 were early plaque, and 16 were late plaque. The dataset was made using the Affymetrix Human Genome U133A 2.0 microarray platform. GSE43292 is a gene expression dataset that contains the transcriptome profiles of human carotid atheroma samples. The study aimed to identify differentially expressed genes and molecular pathways associated with atherosclerosis and plaque vulnerability. The dataset includes 64 carotid atheroma samples from patients undergoing carotid endarterectomy surgery. The samples were classified into stable and unstable plaques based on histological criteria, and gene expression profiles were generated using microarray analysis. The dataset includes both raw and processed gene expression data.

| | **Dataset 1** | **Dataset 2** |
|---|---|---|
| **GEO Accession No.** | GSE28829 | GSE43292 |
| **Type of study** | Expression profiling by array | Expression profiling by array |
| **Platform type** | Affymetrix Human Genome U133 Plus 2.0 Array | Affymetrix Human Genome U133A 2.0 |

| **Species** | Homo sapiens | Homo sapiens |
|---|---|---|
| **No. of samples** | 29 samples (13=early plaque,16=Advanced plaque) | 64 samples(control = 32, Advanced plaque = 32) |

***Table 1:***

***Characteristics of Atherosclerosis-associated datasets retrieved from GEO.***

### 2.3 Identification of Atherosclerosis-specific DEGs

The series matrix expression file for the selected datasets GSE28829 and GSE43292 were obtained from GEO Datasets and was subjected to quality checks. The data sets were divided into two groups early and late plaque for GSE43292 and control and advanced plaque for GSE28829. The ARSyNseq function (The batch settings for the dataset were unknown) was available within the NOISeq package and was used to acquire batch-corrected expression values. The HUGO Gene Nomenclature Committee (HGNC) was mapped with their respective probe IDs using GEO2R (21). The expression values were averaged to avoid redundancy in the expression values of genes mapping to several probe IDS. The p-values and the log2(fold change) values of all the genes in both the sample of both GSE IDs were subsequently calculated using the Limma package using the two sample T-test in R version 4.2.2(20). This was followed by p-value correction using Benjamini-Hochberg (BH) method. Genes with a p-value of <0.05 and an absolute log2(fold change) >0.1 were deemed to be differentially expressed genes among (DEGs) in the two samples. The HGNC multi-symbol checker (https://www.genenames.org/tools/multi-symbolchecker/,Accessed on 10 December 2022) was then used to check and obtain the DEGs with the official and approved HGNC symbols. DEGs with log2(fold change)>0.5 and log2(fold change)<−0.5 were classified as up and downregulated, respectively.

### 2.4 PPI Network Construction and Modular Analysis

The DEGs of both Control and plaque groups for GSE43292 and GSE28829 advanced and early plaque groups were segregated using calcification genes. These filtered genes were given as input one after the other to Search Tool for the Retrieval of Interacting Genes (STRING, https://string-db.org/,accessed on 25 January 2023) v11.5 web-based Tool (22) to develop a PPI network corresponding to highest confidence (i.e., interaction score > 0.9) and later visualized utilizing Cytoscape v3.91 (18). The MCODE Cytoscape app43 was used to identify densely correlated modules within our PPI network. The default setting of MCODE (i.e., "Degree cutoff = 2", "node score cutoff = 0.2", "k-score = 2", "max. depth = 100", and "cut style = haircut") was used for network scoring and module detection. The module that scored the highest was selected as the hub module.

### 2.5 Pathway and Gene Ontology (GO) Term Enrichment Analyses for Hub Gene Selection

All the genes for both the data sets from the PPI network were compiled using the ReactomePA package39 and GO-BP, MF, and CC libraries in the Enrichr database40. GO terms and pathways corresponding to a p-value of <0.05 were taken as statistically significant. (https://maayanlab.cloud/Enrichr/, accessed on 26 February 2023), and top 10 significantly (p-value < 0.05) enriched pathways along with GO terms were collected from Bio Planet, GO Biological Process (BP), GO-Molecular Function (MF), and GO-Cellular Compartment (CC) libraries.

### 2.6 Atherosclerosis-Specific Three-Node miRNA FFL Construction and Analysis

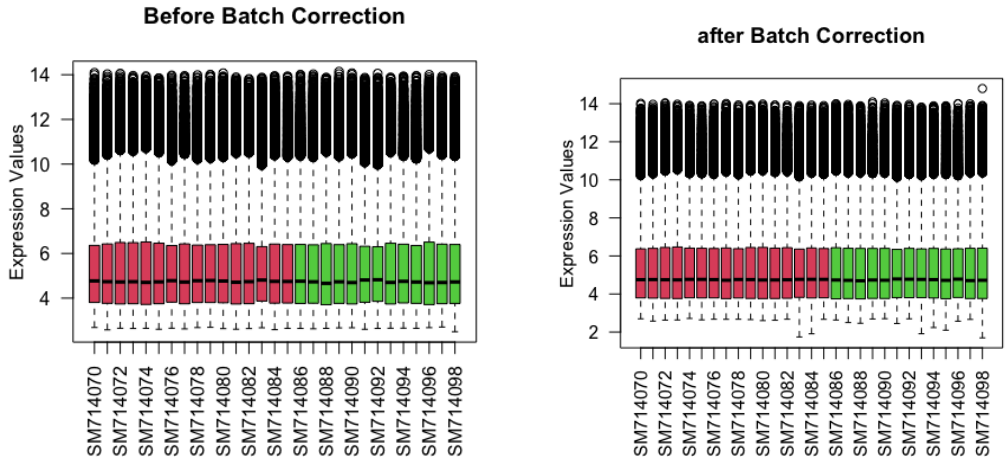
Significant TFs corresponding to humans with a p-value score of <0.05 that we are regulating the hub mRNA obtained from the PPI network were retrieved from the ChEA v3.0

database (23). Then miRNAs (with a score >0.95 and binding only on 3′UTR region) which were regulating the mRNA and extracted significant TFs (collected from ChEA), were retrieved from miRWalk v3.0 database (24). These miRNAs were then validated using literature studies, and those related to atherosclerosis and plaque formation were retained for future analysis. To create a closed three-node miRNA FFL, all three interaction pairs involving TF-mRNA, miRNA-mRNA, and miRNA-TF were modified so that the FFL consists solely of overlapping miRNAs, TFs, and mRNAs. The resulting miRNA FFL was then represented graphically using Cytoscape (18)

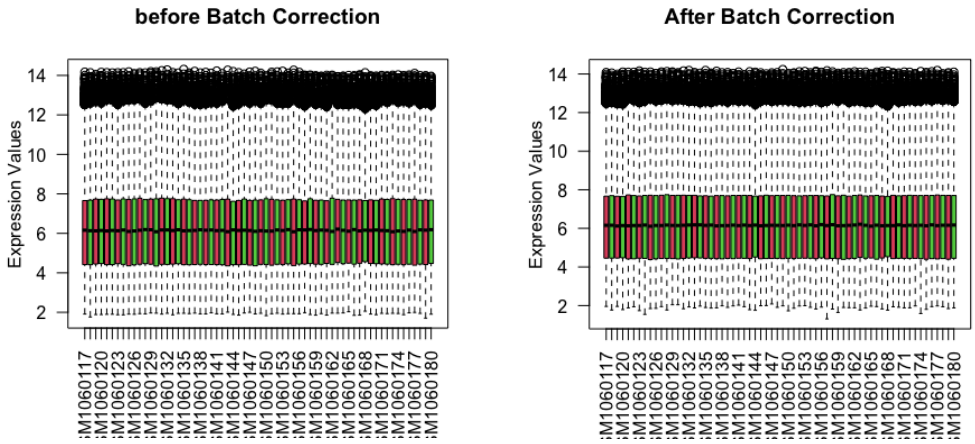
## 3. Results

### 3.1 Data Preprocessing and Identification of DEGs

The gene expression distribution of samples in GSE28829 and GSE43292 before data processing is shown in (Figure 2(a) and 2(c) respectively. We can see here that their median distribution was not in a straight line. Whereas after normalization the median value of gene expression is seen in straight line at the same level (Figure 2(b) and 2(d)).



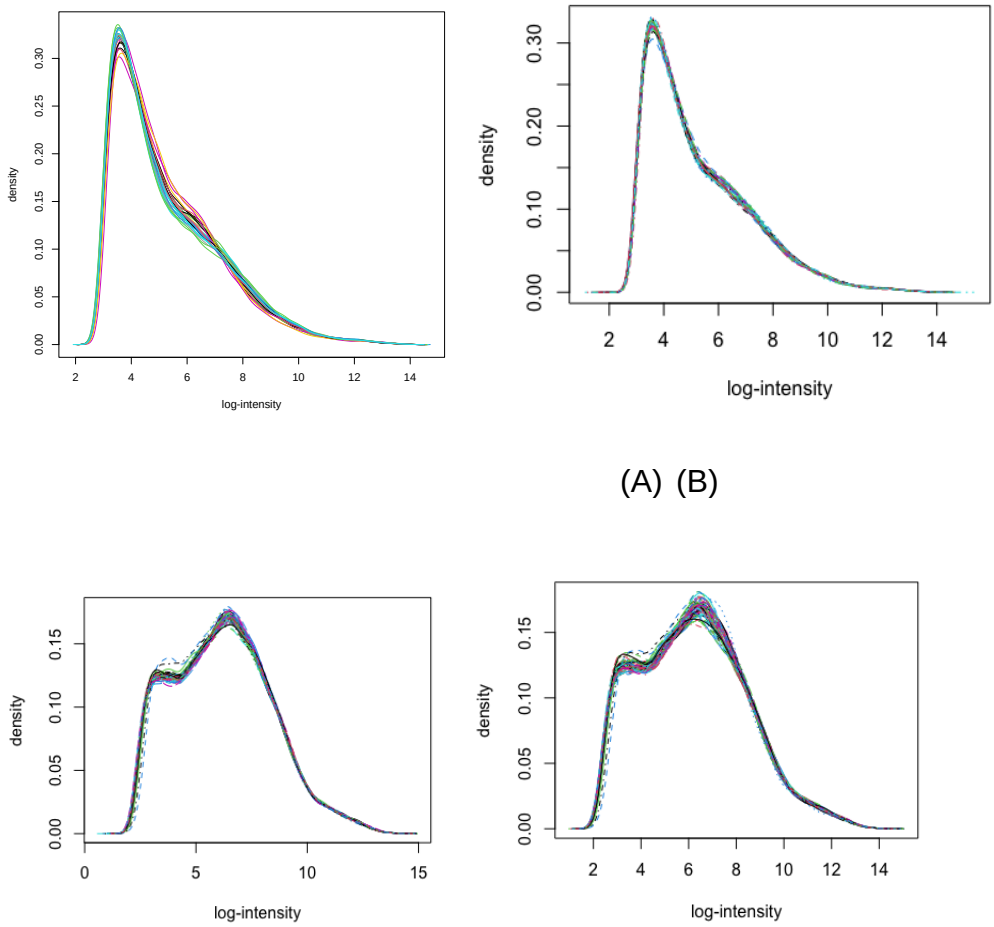
(A) (B)




(C) (D)

***Figure 2: Box plots for the gene expression data. Red bars represent early atherosclerotic plaque samples, and green bars represent advanced atherosclerotic plaque samples for GSE28829 whereas Red bars for GSE43292 is early plaque and Green bars represents Advanced plaque. The black lines in each box represent the median gene expression level. (a,c) The black lines of raw data are not at the same level. (b,d) After data processing, the black lines are almost at the same level.***

The density plots of GSE28829 and GSE43292 was done to check the channel intensities of the microarray data, it represents the level of gene expression for particular genes, and uneven channel intensities can be problematic during data analysis as it can give biased results. So, here we checked the channel intensities before and after normalization for both the datasets.




(A) (B)

(C) (D)

***Figure 3:Density plots for the gene expression data. The plot is a log-intensity vs density plot, the x-axis represents the log-transformed intensity values, while the y-axis represents the density of the observations at each intensity value.The distribution of intensities is expected to be approximately symmetric and bell-shaped, with most observations clustered around the center of the distribution.As you can see in Figure 3(a) and 3(c) the intensities are not clustered around the center so it show that there is some discrupancies in the datasets.Whereas 3(b) and 3(d) shows that the intensitis are centered and shows that the data is normalized properly for both the datasets***

After checking that normalization is done properly the duplicates are removed from both the datasets using R. duplicacy removal is an important step in data cleaning, as duplicate records can lead to a biased result. To prepare the data further AGG function in R was used since sometimes one gene can have more than one value of expression. After that the data

was cleared of variance and it was set at 50% for both datasets. Since selecting the genes that have variance is necessary to get the genes that are most differentially expressed.

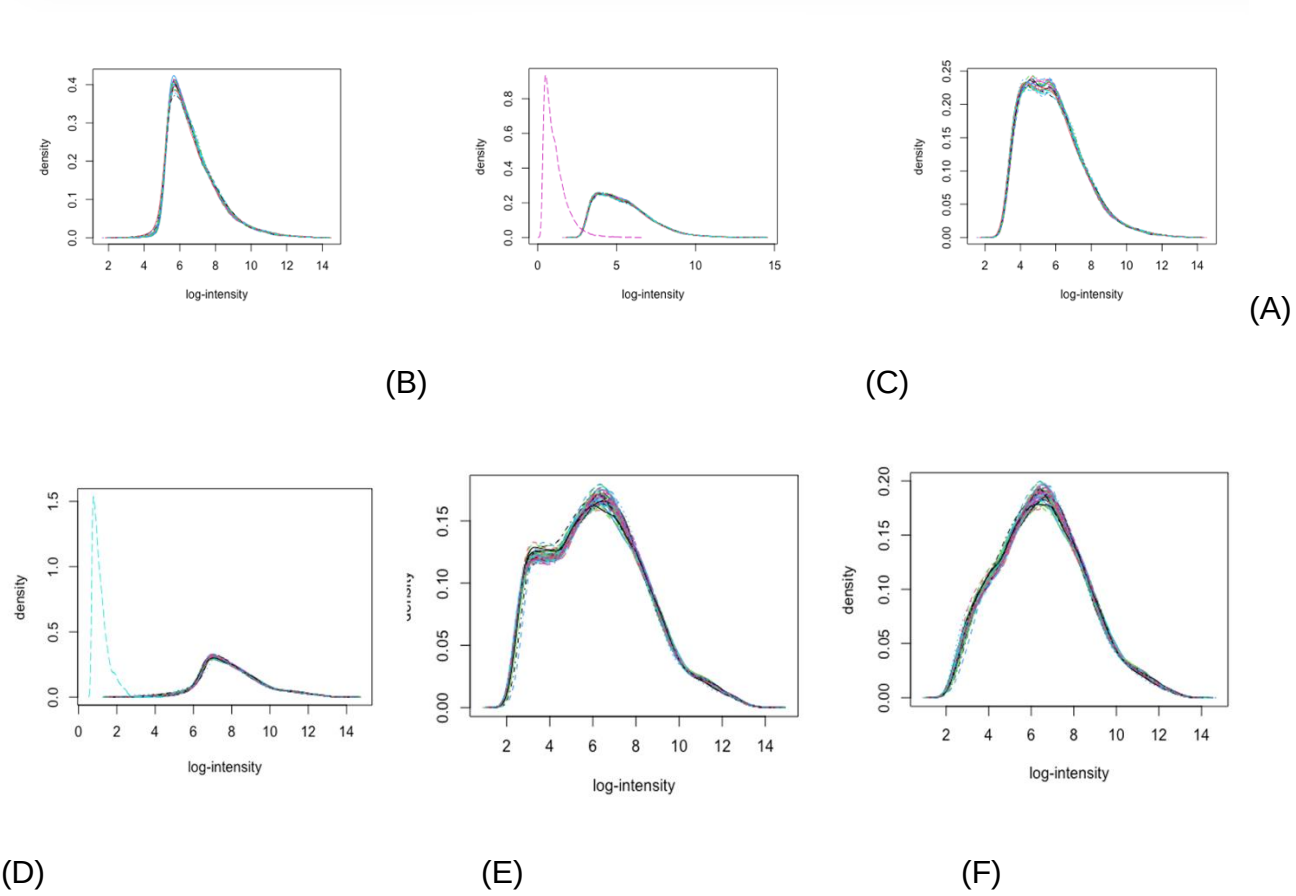



***Figure 4: On basis of the density plots and their channel intensities the variance threshold was selected for 4(A,D)it is just normalized ,duplicacy removed and aggerated data shows that the channel intensities are still distorted.4(B,E) the channel intensities are distorted at 25% variance and so ultimately the variance was chosen to be 50% as channel intensities were accurate for both the datasets as shown in figure 4 (C,F).***

Following the threshold [i.e., p-value <0.05 and ||log2,(fold change)||>0.1], 3786 DEGs were identified for GSE28829 out of which 2111 upregulated and 1675 genes downregulated and 6177 DEGs were identified for GSE43292 out of which 4175 upregulated and 2002 downregulated.

**Table 2:**

GSE28829

| Gene | logFC | adj.P.Va |
|---|---|---|
| ATP1A2 | -1.6519529 | 9.04E-07 |
| ITLN1 | -1.4363462 | 0.04754792 |
| ACADL | -1.3557004 | 2.18E-05 |
| TMEM35A | -1.305323 | 3.81E-08 |
| MBNL1-AS1 | -1.2707273 | 8.33E-08 |
| LGR6 | -1.1729011 | 0.00012763 |
| SH3BGR | -1.1703176 | 1.06E-07 |
| KCNK17 | -1.1680352 | 0.00012055 |
| CASQ2 | -1.1644604 | 0.00028388 |
| ANGPTL1 | -1.1329336 | 4.47E-07 |

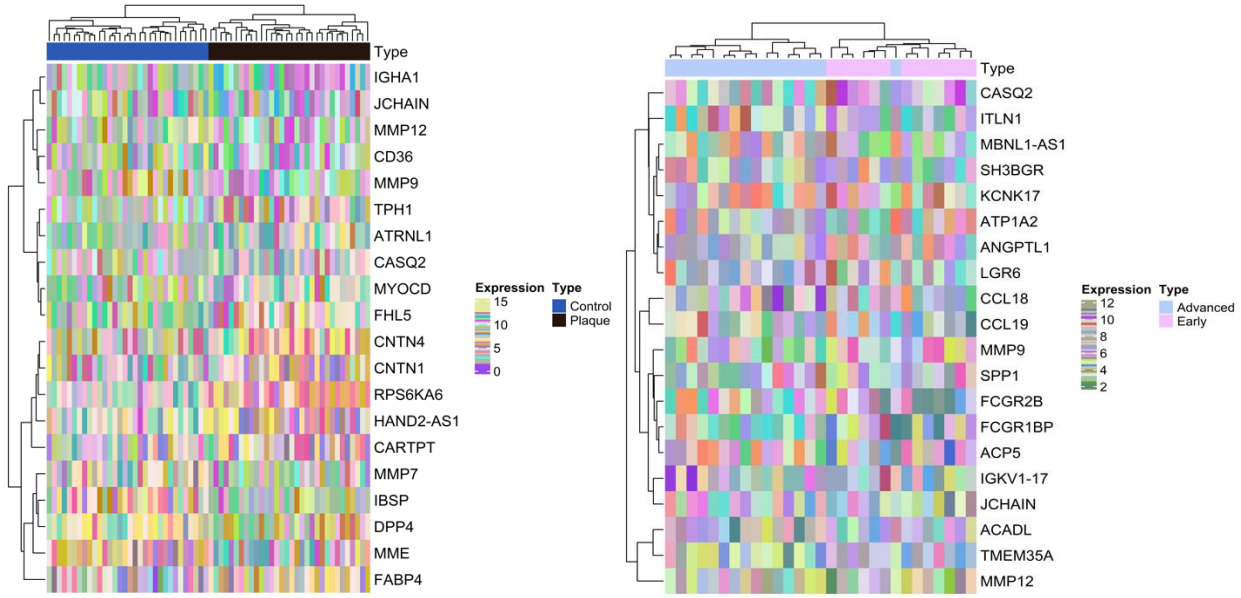
| Gene | logFC | adj.P.Va |
|---|---|---|
| MMP9 | 2.01113047 | 0.00027454 |
| FCGR2B | 2.04459216 | 3.62E-07 |
| ACP5 | 2.05226723 | 1.66E-06 |
| CCL19 | 2.13176836 | 1.39E-06 |
| FCGR1BP | 2.15675635 | 2.57E-07 |
| SPP1 | 2.38458632 | 4.95E-06 |
| IGKV1-17 | 2.41133028 | 9.94E-06 |
| MMP12 | 2.57619745 | 0.00212257 |
| CCL18 | 2.62564254 | 1.62E-06 |
| JCHAIN | 2.6866487 | 1.61E-06 |

GSE43292

| Gene | logFC |
|---|---|
| CNTN1 | -1.9110322 |
| TPH1 | -1.8866263 |
| CNTN4 | -1.7923319 |
| CASQ2 | -1.6676641 |
| MYOCD | -1.6284488 |
| FHL5 | -1.5646522 |
| HAND2-AS1 | -1.5043553 |
| CARTPT | -1.4704869 |
| ATRNL1 | -1.4374975 |
| RPS6KA6 | -1.3692 |

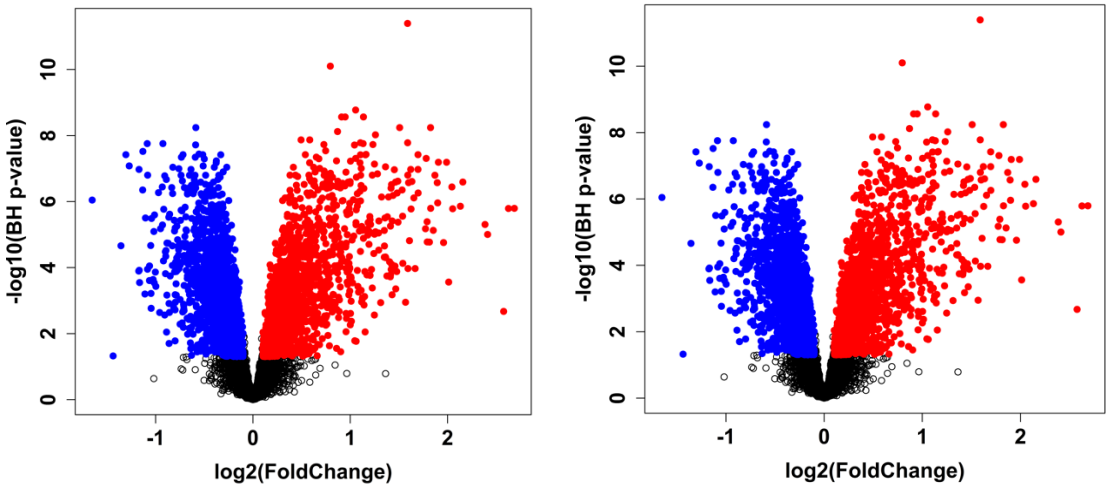
| Gene | logFC |
|---|---|
| FABP4 | 2.45446094 |
| JCHAIN | 1.89314938 |
| MMP7 | 1.84023063 |
| MMP9 | 1.81780406 |
| CD36 | 1.80220531 |
| IBSP | 1.79498219 |
| DPP4 | 1.61087281 |
| MMP12 | 1.57410563 |
| MME | 1.48500031 |
| IGHA1 | 1.45345672 |

***Table 2: This table shows the top up and down regulated genes for GSE28829 and GSE43292 respectively. For GSE28829 and GSE43292 the blue ones are downregulated as they have a negative fold change value whereas the green one is upregulated with a positive fold change value***




(A) (B)

***Figure 5: The Clustered heat map of gene expression heat map demonstrates an overview of significant (up versus downregulated) and nonsignificant genes in early and late plaque for GSE22829 5(A) and Advanced plaque and control for GSE43292 5(B) respectively.***




(A) (B)

***Figure 6: This table shows the volcano plot for the DEGs for both the data sets. The volcano plot displays the log2 fold change (x-axis) versus the statistical significance or p-value (y-axis) of each gene in a dataset.6(A) is for GSE28829 whereas 6(B) is volcano plot for GSE43292.***

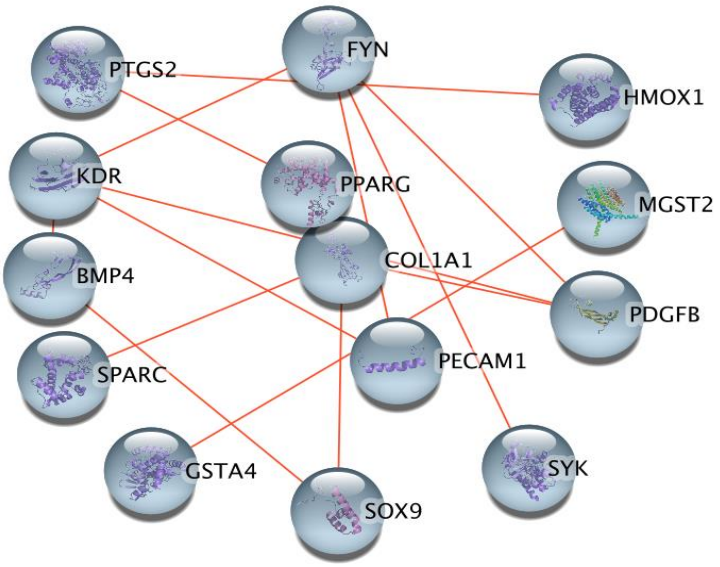
### 3.2 PPI Construction and Multidata base Analysis of Modules

| | | |
|---|---|---|
| ACP5 | FOSL1 | MSX2 |
| AKT | Gas6 | NA |
| ALK | GSTA3 | OPG |
| ALPL | GSTA4 | PECAM1 |
| FYN | HMOX1 | PLCG2 |
| BMP2 | KDR | PPARG |
| BMP2K | MAPK | PTGS2 |
| BMP4 | MEF2C | RUNX2 |
| BSP | MGP | SOST |
| COL1A1 | PDGFB | SOX9 |
| COL1A2 | PDGFA | SPARC |
| SRY | SRY | TNF |
| TNFAIP3 | TNFRSF11A | TNFSF13B |
| TRPV4 | | |

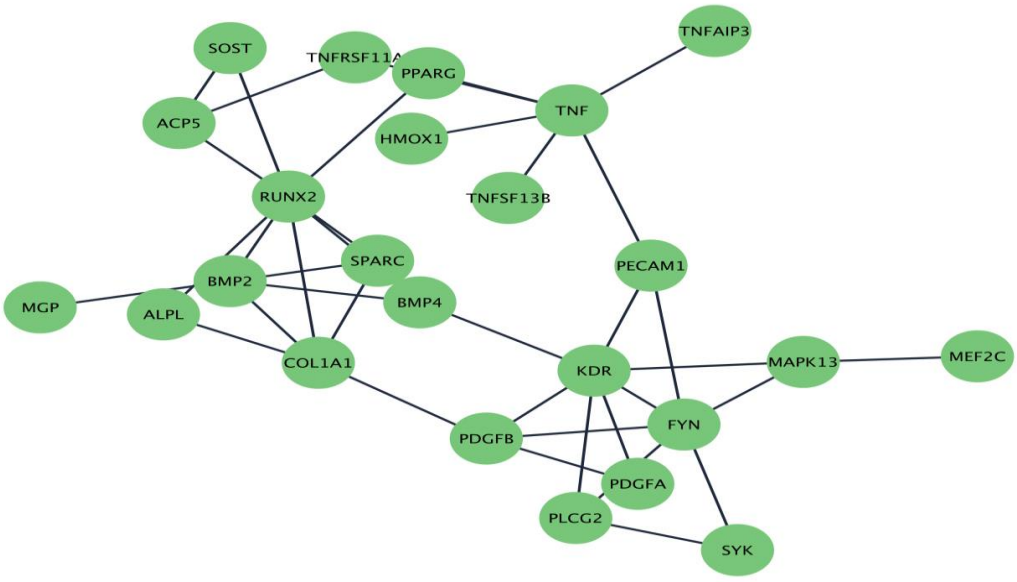
***Table 3: List of the genes collected using literature search that are related to atherosclerosis and vascular calcification.***

Using the DEGs list for both GSE IDs and the calcification list a new list was generated that was used to generate PPI network. The PPI network of the genes that were classified was constructed using the STRING app in Cystoscope. The confidence score was set to be 0.7(17).

Figure:7




(A)




(B)

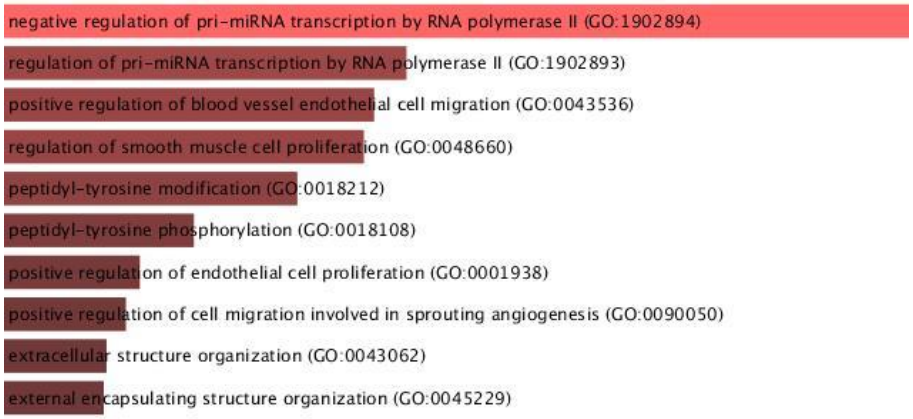
***Figure 7: The figure number 7 (a) and 7(b) shows the PPI network of GSE28829 and GSE43292 respectively.***

### 3.3 Functional Enrichment Analysis of the PPI list

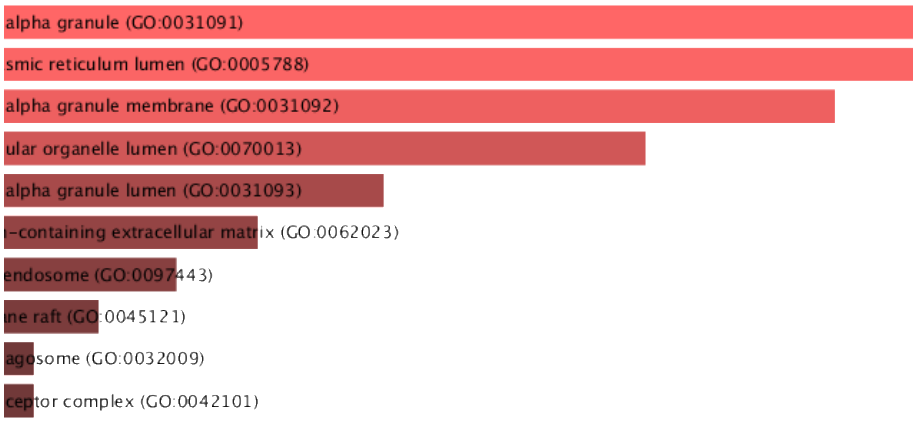
To understand how the genes obtained from PPI analysis play a role in Vascular calcification and plaque formation. We did a multi-database analysis of those genes for both datasets using Enrichr and got the GO: BP, GO: CC, GO: MF, Reactome, and KEGG pathways. The GO terms of the biological process (BP) analysis showed that the BPs of the GSE28829 were mainly enhanced in white cell differentiation, ventricular morphogenesis, vasculogenic, vascular wound healing, vascular endothelial growth factor signaling which showed that they play a role not only in immune cells stimulation but vascular growth and calcification in patients with atherosclerosis plaque development.(figure 8(a))The GO terms of the cellular component (CC) were mainly enhanced in Platelet alpha granule, Endoplasmic reticulum lumen, Platelet alpha granule membrane, Intracellular organelle lumen, Platelet alpha granule lumen, Collagen-containing extracellular matrix, Sorting endosome, Membrane raft, Early phagosome, T cell receptor complex(figure 8(b)). The GO terms of molecular function (MF) were enhanced Platelet-derived growth factor binding, phospholipase binding, protein tyrosine kinase activity, glutathione transferase activity, chemoattractant activity, non-membrane spanning protein tyrosine kinase activity, protein homodimerization activity, heme binding, growth factor receptor binding, leukotriene-C4 synthase activity (figure 8(c)). The KEGG pathway and the gene regulating the atherosclerosis vascular calcification are also shown (Figure 8 (d),e). Further, the Disease enrichment analysis in the DisGeNET database revealed that these genes were mainly present in Arteriosclerosis, Pulmonary Hypertension, Myeloid Leukemia Chronic, Fibrosis, Retinal Diseases, Mesothelioma, Pancreatic Neoplasm, Liver Cirrhosis, Glioma, Atherosclerosis.

Whereas the GO terms of the biological process (BP) analysis showed that the BPs of the GSE43292 were mainly enhanced in the Regulation of smooth muscle cell proliferation, Positive regulation of cell population proliferation, Positive Regulation of nucleic acid-templated transcription, Regulation of blood vessel endothelial cell migration, Positive Regulation of protein phosphorylation, Positive Regulation of cell differentiation, Regulation of glomerular mesangial cell proliferation, Regulation of protein phosphorylation, Positive Regulation of cell migration, Cellular response to growth factor stimulus. This shows how these genes play a role in plaque formation. (figure 9(A)). The GO terms of the cellular component (CC) were prominent in the Platelet alpha granule, Platelet alpha granule lumen, Collagen-containing extracellular matrix, Platelet alpha granule membrane, Endoplasmic reticulum lumen, Membrane raft, Intracellular organelle lumen, Secretory granule lumen, Sorting endosome.(figure 9(B)).The GO terms of the Molecular function (MF) were prominent in the expression of Platelet-derived growth factor binding, receptor-ligand activity, platelet-derived growth factor receptor binding, BMP receptor binding, phospholipase binding, transmembrane receptor protein serine/threonine kinase binding, growth factor receptor binding, protein tyrosine kinase activity, protease binding, chemoattractant activity. (figure 9(C)Disease enrichment analysis in the DisGeNET database revealed that these genes were mainly related to Bone Diseases, Osteopenia, Osteoporosis, Arteriosclerosis, Atherosclerosis, Degenerative polyarthritis, Uremia, Pulmonary Hypertension, Neoplasm Metastasis, Inflammation. (9 f)

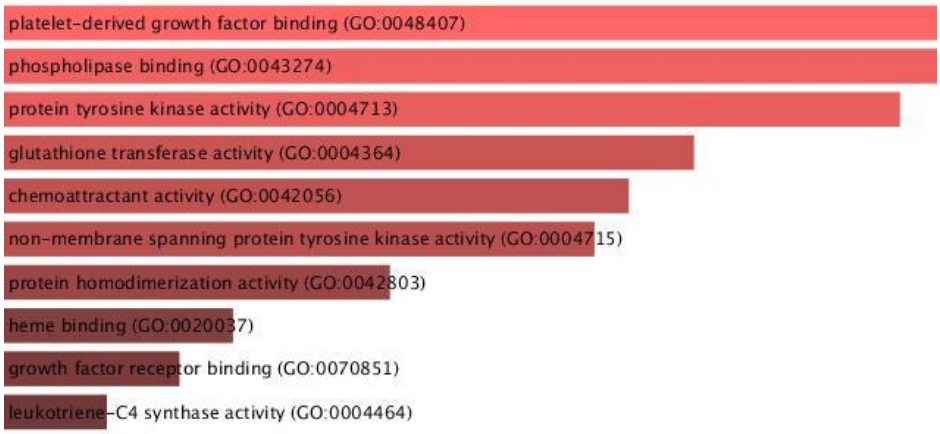
Figure 8:


negative regulation of pri-miRNA transcription by RNA polymerase II (GO:1902894)
regulation of pri-miRNA transcription by RNA polymerase II (GO:1902893)
positive regulation of blood vessel endothelial cell migration (GO:0043536)
regulation of smooth muscle cell proliferation (GO:0048660)
peptidyl-tyrosine modification (GO:0018212)
peptidyl-tyrosine phosphorylation (GO:0018108)
positive regulation of endothelial cell proliferation (GO:0001938)
positive regulation of cell migration involved in sprouting angiogenesis (GO:0090050)
extracellular structure organization (GO:0043062)
external encapsulating structure organization (GO:0045229)

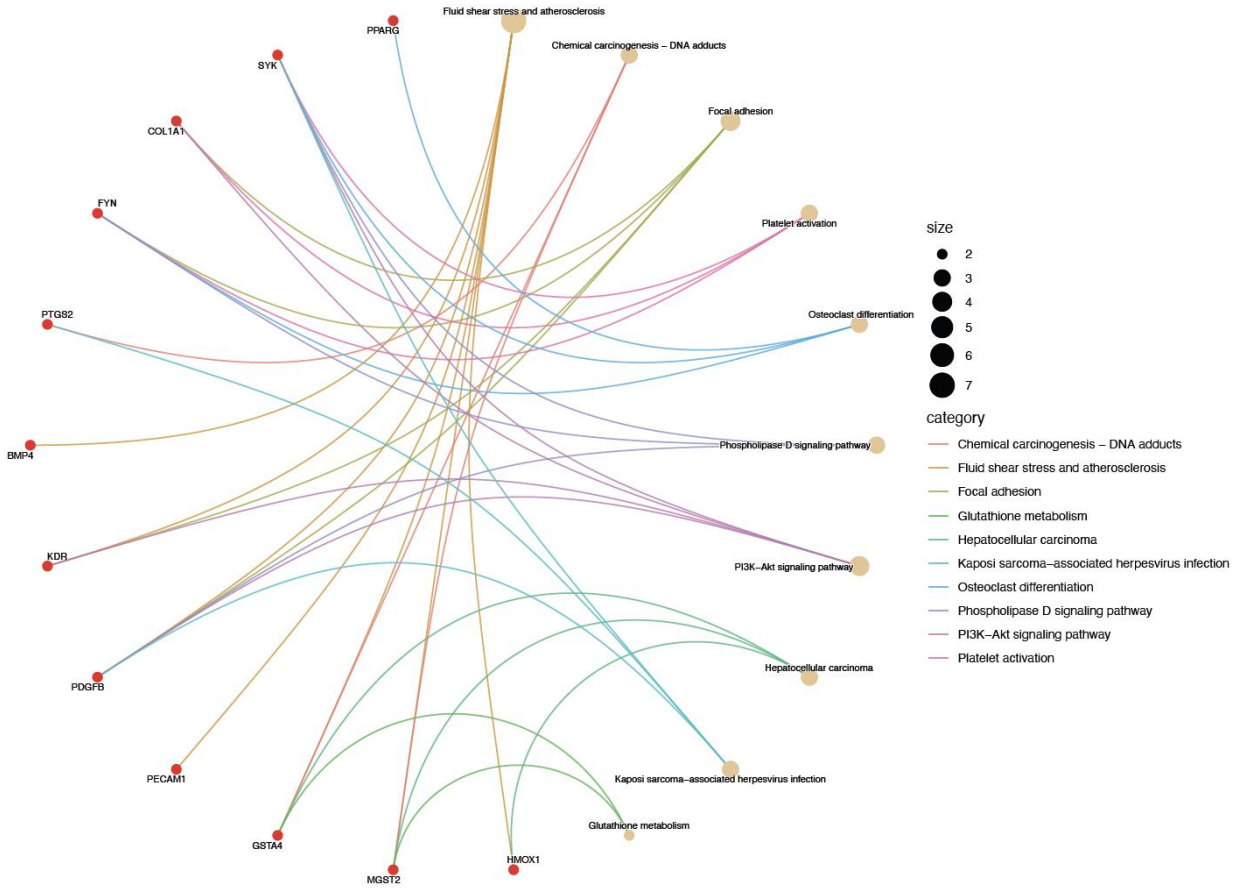
(A)


alpha granule (GO:0031091)
smic reticulum lumen (GO:0005788)
alpha granule membrane (GO:0031092)
ular organelle lumen (GO:0070013)
alpha granule lumen (GO:0031093)
-containing extracellular matrix (GO:0062023)
endosome (GO:0097443)
ne raft (GO:0045121)
agosome (GO:0032009)
ceptor complex (GO:0042101)

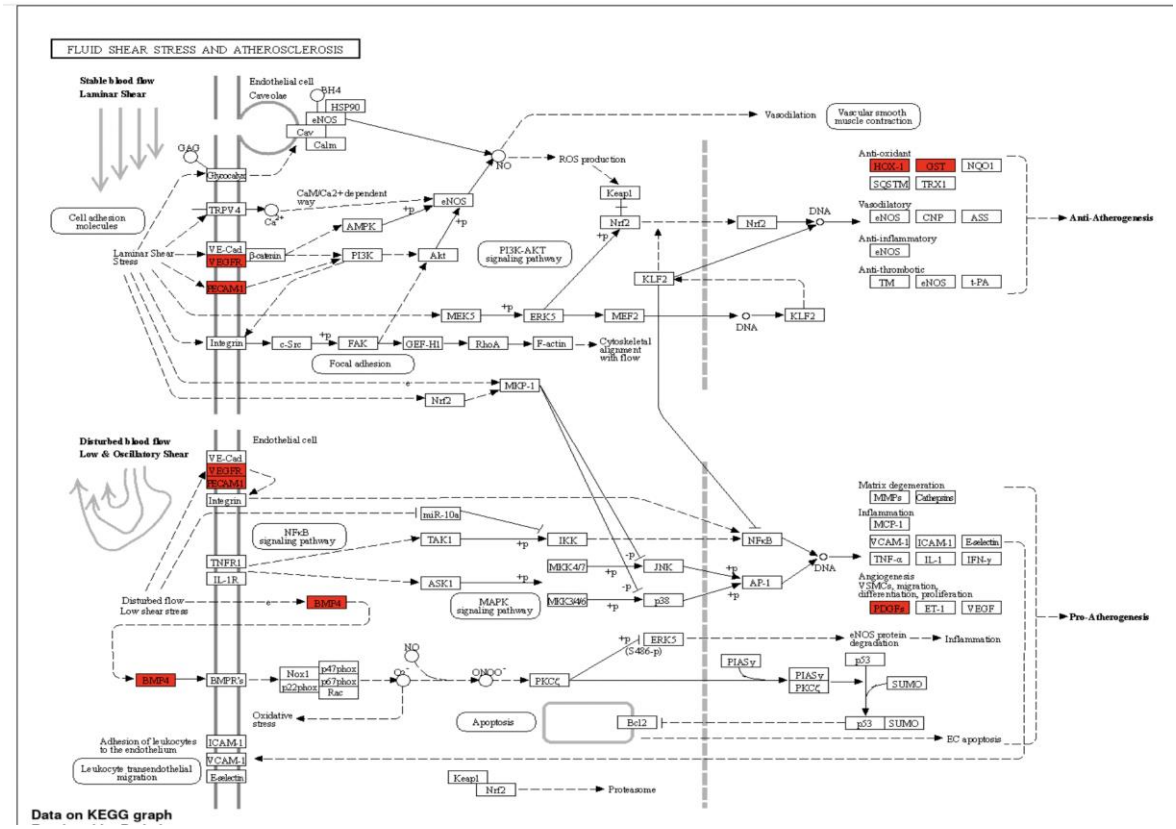
(B)


platelet-derived growth factor binding (GO:0048407)
phospholipase binding (GO:0043274)
protein tyrosine kinase activity (GO:0004713)
glutathione transferase activity (GO:0004364)
chemoattractant activity (GO:0042056)
non-membrane spanning protein tyrosine kinase activity (GO:0004715)
protein homodimerization activity (GO:0042803)
heme binding (GO:0020037)
growth factor receptor binding (GO:0070851)
leukotriene-C4 synthase activity (GO:0004464)

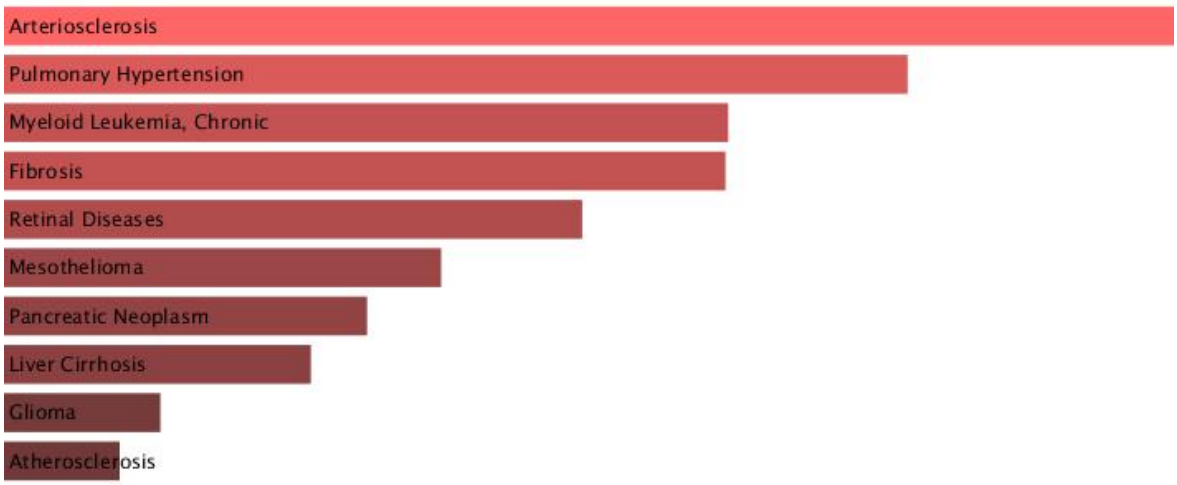
(C)


Fluid shear stress and atherosclerosis
PPARG
Chemical carcinogenesis – DNA adducts
SYK
Focal adhesion
COL1A1
FYN
Platelet activation
PTGS2
Osteoclast differentiation
BMP4
Phospholipase D signaling pathway
KDR
PI3K–Akt signaling pathway
Hepatocellular carcinoma
PDGFB
PECAM1
Kaposi sarcoma–associated herpesvirus infection
Glutathione metabolism
GSTA4
HMOX1
MGST2
size
2
3
4
5
6
7
category
Chemical carcinogenesis – DNA adducts
Fluid shear stress and atherosclerosis
Focal adhesion
Glutathione metabolism
Hepatocellular carcinoma
Kaposi sarcoma–associated herpesvirus infection
Osteoclast differentiation
Phospholipase D signaling pathway
PI3K–Akt signaling pathway
Platelet activation

(D)



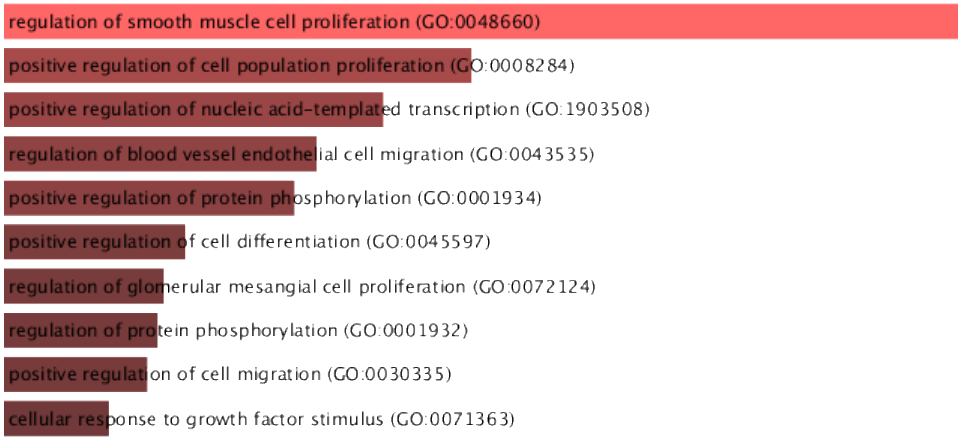

(E)



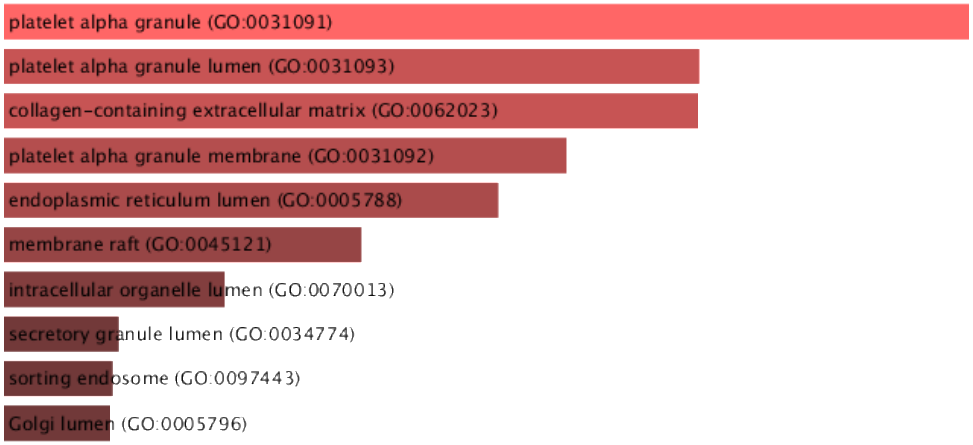

(F)

***Figure 8: Multidata base analysis of hub genes :(A) Biological functions of hub genes,(B ) Cellular components of hub genes,(C) Molecular functions of hub genes (D ) (E) pathways regulated by hub genes,(F) Disease enrichment related to hub genes involved in AA.***

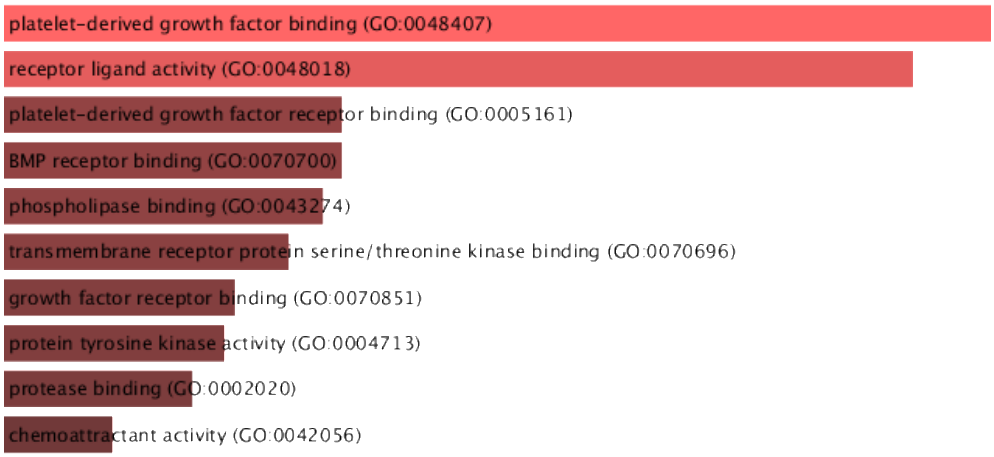
Figure 9:


regulation of smooth muscle cell proliferation (GO:0048660)
positive regulation of cell population proliferation (GO:0008284)
positive regulation of nucleic acid-templated transcription (GO:1903508)
regulation of blood vessel endothelial cell migration (GO:0043535)
positive regulation of protein phosphorylation (GO:0001934)
positive regulation of cell differentiation (GO:0045597)
regulation of glomerular mesangial cell proliferation (GO:0072124)
regulation of protein phosphorylation (GO:0001932)
positive regulation of cell migration (GO:0030335)
cellular response to growth factor stimulus (GO:0071363)

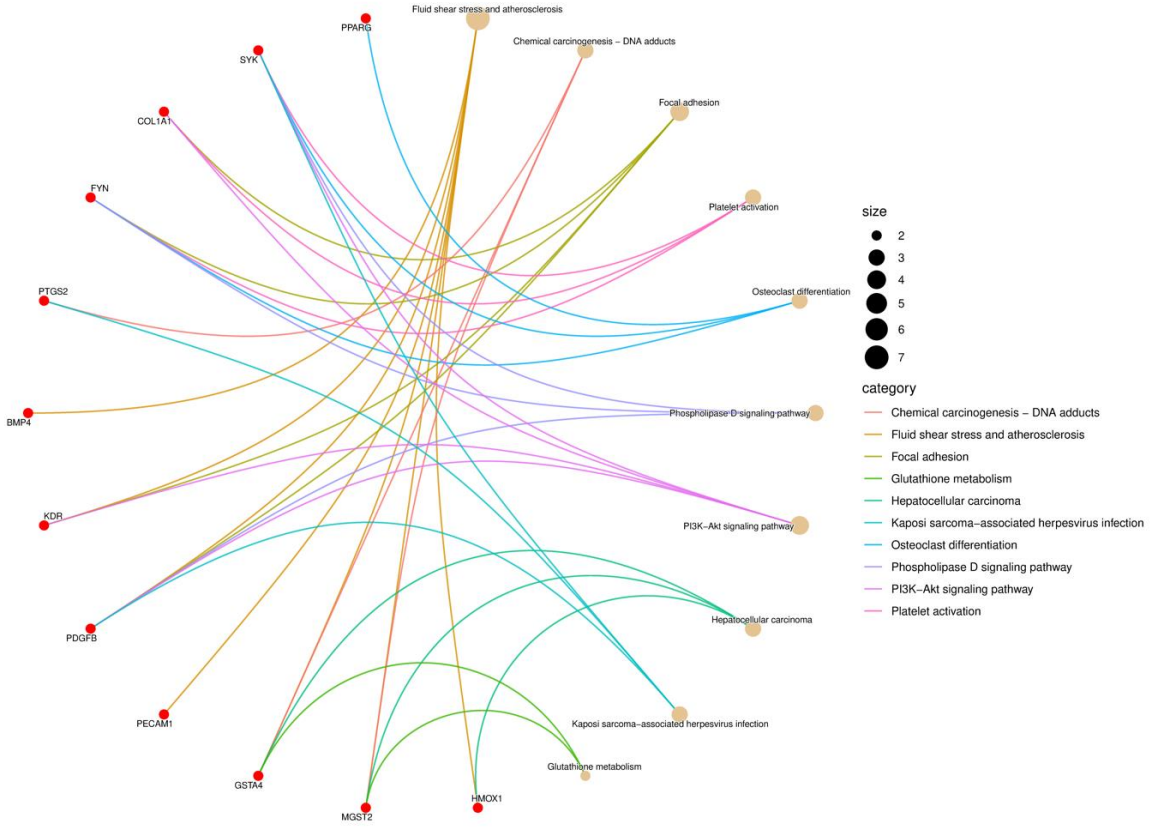
(A)


platelet alpha granule (GO:0031091)
platelet alpha granule lumen (GO:0031093)
collagen-containing extracellular matrix (GO:0062023)
platelet alpha granule membrane (GO:0031092)
endoplasmic reticulum lumen (GO:0005788)
membrane raft (GO:0045121)
intracellular organelle lumen (GO:0070013)
secretory granule lumen (GO:0034774)
sorting endosome (GO:0097443)
Golgi lumen (GO:0005796)

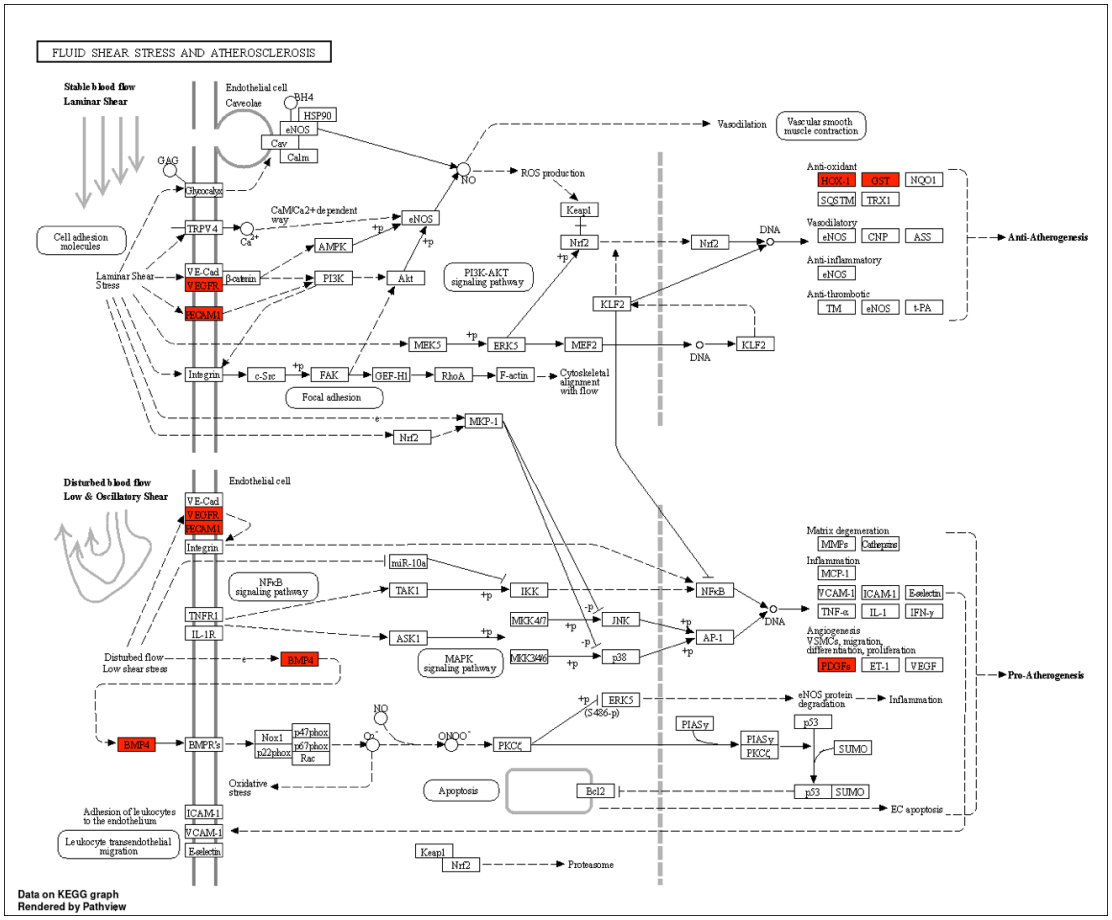
(B)


platelet-derived growth factor binding (GO:0048407)
receptor ligand activity (GO:0048018)
platelet-derived growth factor receptor binding (GO:0005161)
BMP receptor binding (GO:0070700)
phospholipase binding (GO:0043274)
transmembrane receptor protein serine/threonine kinase binding (GO:0070696)
growth factor receptor binding (GO:0070851)
protein tyrosine kinase activity (GO:0004713)
protease binding (GO:0002020)
chemoattractant activity (GO:0042056)

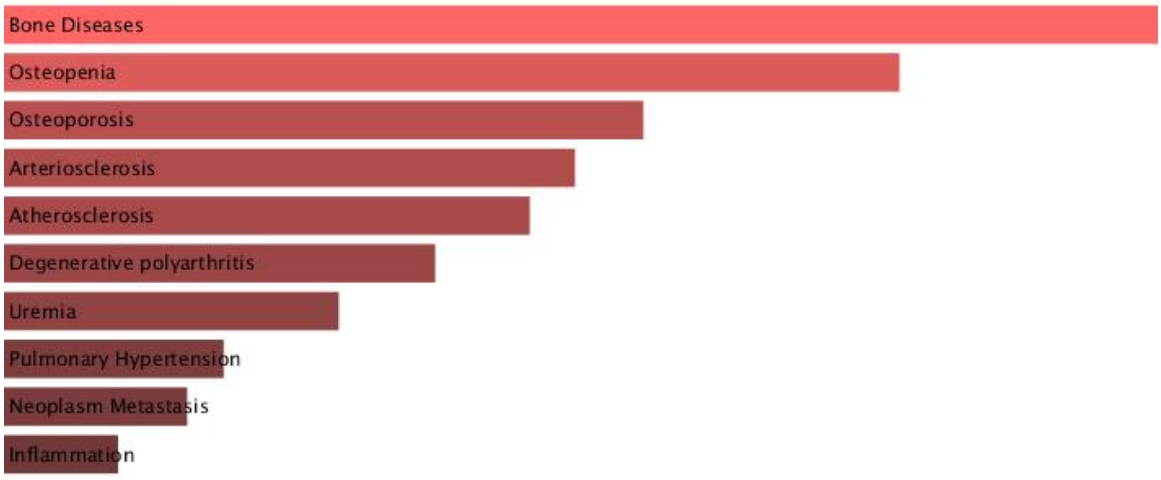
(C)


Fluid shear stress and atherosclerosis
PPARG
Chemical carcinogenesis – DNA adducts
SYK
Focal adhesion
COL1A1
FYN
Platelet activation
PTGS2
Osteoclast differentiation
BMP4
Phospholipase D signaling pathway
KDR
PI3K–Akt signaling pathway
Hepatocellular carcinoma
PDGFB
PECAM1
Kaposi sarcoma–associated herpesvirus infection
Glutathione metabolism
GSTA4
HMOX1
MGST2
size
2
3
4
5
6
7
category
Chemical carcinogenesis – DNA adducts
Fluid shear stress and atherosclerosis
Focal adhesion
Glutathione metabolism
Hepatocellular carcinoma
Kaposi sarcoma–associated herpesvirus infection
Osteoclast differentiation
Phospholipase D signaling pathway
PI3K–Akt signaling pathway
Platelet activation

(D)




(E)




(F)

***Figure 9: Multidata base analysis of hub genes :(A) Biological functions of hub genes ,(B ) Cellular components of hub genes,(C) Molecular functions of hub genes (D ) pathways regulated by hub genes ,(F) Disease enrichment related to hub genes involved in AA.***

### 3.4 Atherosclerosis-Specific Three-Node miRNA FFL Analysis

Figure 10A presents a miRNA-focused feed-forward loop (FFL) specific to the Atherosclerosis for GSE28829. Figure 11A presents a miRNA-focused feed-forward loop (FFL) specific to the Atherosclerosis for GSE43292.This FFL consists pairs of transcription factors (TFs) and messenger RNAs (mRNAs), miRNAs and mRNAs, and miRNAs and TFs.

Figure 10(A) 10(B)

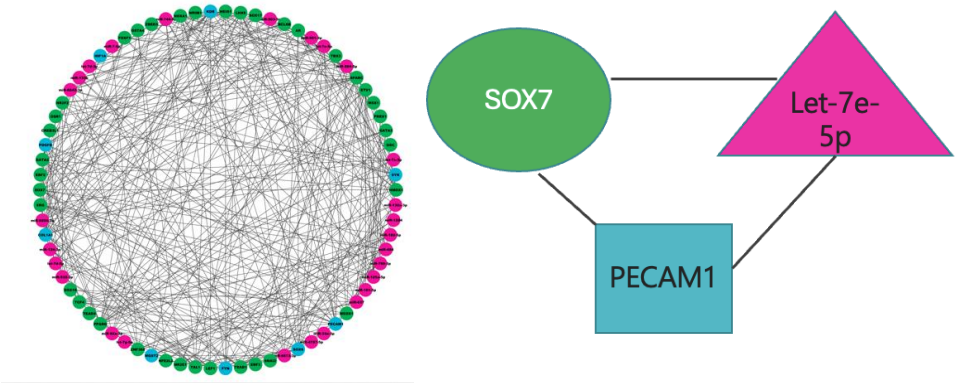



***Figure 10: (A)Atherosclerosis-specific three nodes miRNA FFL has ----- nodes and ---- edges. (B)Highest order for subnetwork motif is compromised of one TF(SOX7), one miRNA(let-7p-5p), and one mRNA (PECAM1). The Pink-colored circle represents miRNA, the blue color represent mRNA and whereas blue represent TFs.***

Figure 11(A) 11(B)

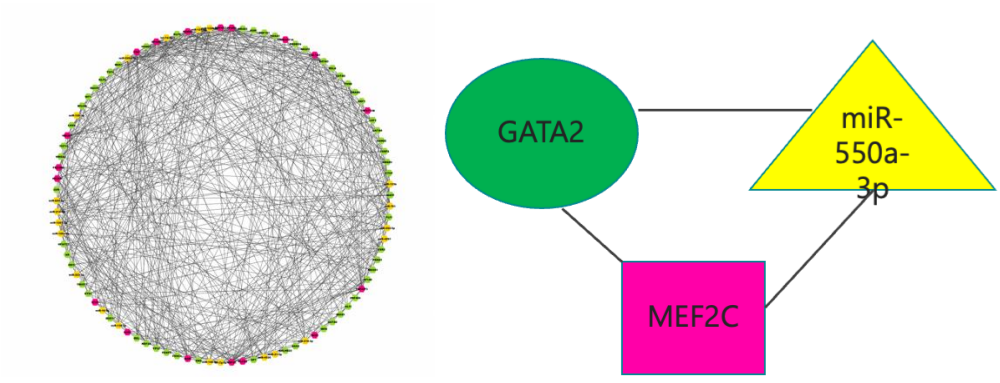



***Figure 11: (A)Atherosclerosis-specific three nodes miRNA FFL has ----- nodes and ---- edges. (B)Highest order for subnetwork motif is comprised of one TF(GATA2), one miRNA (miR-550a-3p, and one mRNA (MEF2C). The Pink-colored circle represents mRNA, the green color represents TF and whereas yellow represent miRNA.***

Table S1 summarizes the three regulatory relationships between miRNAs, mRNAs, and TFs.The three regulatory relationships between miRNAs, mRNAs, and TFs were summarized. Tables S2–S4 show the top three miRNAs, mRNAs, and TFs within FFL ranked based on degree, betweenness, closeness, eigenvector, and radiality. Whereas Table S5 summarizes the three regulatory relationships between miRNAs, mRNAs, and TFs.The three regulatory relationships between miRNAs, mRNAs, and TFs were summarized. Tables S6–S8 show the top three miRNAs, mRNAs, and TFs within FFL ranked based on degree, betweenness, closeness, eigenvector, and radiality. After analyzing the tables, it was found that the highest-order subnetwork motif consisted of one TF SOX7, one miRNA (let-7p-5p), and one mRNA PECAM-1 for GSE28829. For GSE43292 it was found that the highest-order subnetwork motif consisted of one TF GATA2, one miRNA (miR-550a-3p), and one mRNA ME2FC.

| Relationship | No. of edges | No. of miRNA | No. of TFs | No. of mRNA |
|---|---|---|---|---|
| **miRNA-mRNA** | 200 | 50 | 0 | 150 |
| **TF-mRNA** | 100 | 0 | 10 | 70 |
| **miRNA-TF** | 20 | 10 | 10 | 0 |

***Table S1: The relationship between miRNA-mrna, Tf-mRNA and miRNA-Tf***

| miRNA | Degree | Betweenness | Closeness |
|---|---|---|---|
| **let-7e-5p** | 17 | 0.37 | 0.52 |
| **miR-484** | 15 | 0.02 | 0.54 |
| **miR-6873-3p** | 13 | 0.01 | 0.48 |
| **miR-584-5p** | 12 | 0.01 | 0.52 |
| **miR-6842-3p** | 12 | 0.01 | 0.52 |

***Table S2: Top 5 miRNAs ranked based on centrality measures such as degree, betweenness and closeness.***

| Gene | Degree | Betweenness | Closeness |
|---|---|---|---|
| **SPARC** | 27 | 0.09 | 0.57 |
| **COL1A1** | 24 | 0.07 | 0.55 |
| **PPARG** | 22 | 0.06 | 0.53 |
| **KDR** | 22 | 0.05 | 0.53 |
| **PECAM1** | 19 | 0.03 | 0.53 |

***Table S3: Top 5 mRNAs ranked based on centrality measures such as degree, betweenness, and closeness.***

| TF | Degree | Betweenness | Closeness |
|---|---|---|---|
| **LHX6** | 16 | 0.03 | 0.51 |
| **SOX7** | 15 | 0.03 | 0.54 |
| **ERG** | 15 | 0.02 | 0.50 |
| **TCF4** | 14 | 0.03 | 0.50 |
| **TAL1** | 14 | 0.01 | 0.53 |

***Table S4: Top 5 TFs ranked based on centrality measures such as degree, betweenness and closeness.***

| Relationship | No. of edges | No. of miRNA | No. of TFs | No. of mRNA |
|---|---|---|---|---|
| **miRNA-mRNA** | 300 | 80 | 0 | 220 |
| **TF-mRNA** | 150 | 0 | 50 | 100 |
| **miRNA-TF** | 87 | 25 | 25 | 0 |

***Table S5: The relationship between miRNA-mrna, Tf-mRNA and miRNA-Tf***

| miRNA | Degree | Betweenness | Closeness |
|---|---|---|---|
| **miR-550a-3p** | 29 | 0.08 | 0.58 |
| **miR-1228-3p** | 21 | 0.03 | 0.52 |
| **miR-124-3p** | 21 | 0.03 | 0.50 |
| **miR-214-3p** | 19 | 0.02 | 0.51 |
| **miR-761** | 19 | 0.02 | 0.49 |

***Table S6: Top 5 miRNAs ranked based on centrality measures such as degree, betweenness and closeness.***

| Gene | Degree | Betweenness | Closeness |
|---|---|---|---|
| **MEF2C** | 30 | 0.07 | 0.57 |
| **COL1A1** | 30 | 0.05 | 0.55 |
| **TNFAIP3** | 26 | 0.05 | 0.52 |
| **SPARC** | 28 | 0.04 | 0.54 |
| **PPARG** | 29 | 0.04 | 0.52 |

***Table S7: Top 5 genes ranked based on centrality measures such as degree, betweenness and closeness.***

| TF | Degree | Betweenness | Closeness |
|---|---|---|---|
| **ESR2** | 17 | 0.02 | 0.50 |
| **MEF2A** | 13 | 0.01 | 0.50 |
| **SP7** | 12 | 0.01 | 0.47 |
| **TCF4** | 11 | 0.01 | 0.48 |
| **GATA2** | 12 | 0.01 | 0.50 |

***Table S8: Top 5 TFs ranked based on centrality measures such as degree, betweenness and closeness***

## 4.Discussion

Our primary approach involved searching GEO for a dataset that contained mRNA samples from patients of GSE28829 and GSE43292. We included representatives from both control and advanced plaque for GSE28829 in our analysis and early/late plaque for GSE43292. Our findings revealed DEGs for the 3786 GSE28829 and 6177 GSE43292, respectively. Additionally, a modular analysis of the PPI network revealed modules 14 genes

for GSE28829 and 24 genes for GSE4329. Further pathway and GO term enrichment analyses of this module led to identifying the role of these genes for both datasets between the top 10 significant tracks, GO-BP, MF, CC gene sets, pathways, and diseases. We also identified Atherosclerosis-associated miRNAs and TFs related to these hub genes. Our study focused on transcriptional interactions, such as miRNA → gene, miRNA → TF, and TF → gene regulation, to create a miRNA-centered feed-forward loop (FFL). This FFL plays a crucial role in Atherosclerosis pathogenesis. There is a scarcity of studies investigating miRNA regulation during Atherosclerosis. Using centrality measures, we identified Let-7e-5p (a member let-7 family) as the top-ranked miRNAs regulating the feed-forward loop (FFL) that includes SOX7 and PECAM1 for GSE28829, whereas we found that TF: GATA2, miRNA: miR-550a-3p and Gene: MEF2C as the regulating FFL loop for GSE43292.

The let-78-5p has already been implicated in being downregulated in the cellular calcification model (25). The INF-alpha pathway contributes to the development of Atherosclerosis by triggering inflammation, activating the immune system, and causing immune cells to stick to the walls of arteries. A study has found that let 7e-5p is found to be downregulated in the IFN-α signaling pathway in B cells (26). Another study found that let-7e-5p promoted osteoblast differentiation but regulated the expression levels of the RUNX2(27). This RUNX2 has been found to play a role in Vascular calcification, so it can be hypothesized that let-7e-5p plays a role in calcification (28). Additionally, let-7e-5p has also been found to target TLR4, which is involved in Atherosclerosis directly (29). This shows how the let-7e-5p is downregulated in the process of Atherosclerosis and calcification and may be a potential biomarker in the case of atherosclerotic calcification. The Let-7 family has also been found to affect inflammation-related diseases like stroke, myocardial infarction, cardiac fibrosis, and AS

(30). Other studies reported that Let-7e-5p might regulate the MAPK signaling pathway in IS (31). The functional identification of specific miRNAs may help provide new insights into the underlying mechanisms of AS. Additionally Li 140 et al. showed that let-7e-5p, miR-222-3p, and miR-433 might be the underlying cause for abnormalities since they target NOTCH1, HAND1, GATA3, and ZFPM2 resulting in altered morphogenesis and VSD. let-7e-5p was also significantly downregulated in calcified aortic valves (32).

The TF Sox7 has been found to play a role in vascular calcification, and it was found that it had a higher expression (upregulated) with a higher stenosis rate in atherosclerotic tissue (33).Sox7 is found to regulate the WNT pathway during angiogenesis (34).The Wnt signaling pathway is linked to several processes that contribute to the narrowing of arteries, such as the attachment of monocytes, inflammation, cell growth, the formation of foam cells, and the calcification of atherosclerotic plaques. This can eventually lead to the development of arterial stenosis. Vascular smooth muscle cells (VSMCs) are involved in these processes. This shows how Sox7 may be over-expressed in vascular calcification and might play a role in atherosclerosis plaque formation (35).

PECAM-1 has been found in the inner lining of blood vessels and contributes to the development of Atherosclerosis in areas of turbulent blood flow by controlling the expression of genes through a pathway called NF-κB (36). The protein PECAM-1 promotes the development of atherosclerotic lesions and raises the levels of substances that contribute to Atherosclerosis. The results of this study suggest that the absence of PECAM-1 delays the onset of Atherosclerosis but does not entirely prevent it. The PECAM-1 is found to be upregulated in atherosclerotic plaque formation. The mechanisms that depend on PECAM-1 are essential for

the entire progression of atherosclerotic lesions, indicating that PECAM-1 has significant potential as a target to limit the course of the disease (37,38).

The first hypothesis revolves around the association of miR-let-7e-5p. According to previous studies, the finalized miRNAs play a significant role in developing atherosclerotic plaque. It has been shown that let-7e-5p has been downregulated in the atherosclerotic plaque in VSMCs.in addition, we can hypothesize that PECAM-1 from mentioned studies that PECAM-1 is upregulated as the earlier studies provided evidence to support this hypothesis by showing that miRNAs can impact the development of Atherosclerosis by either directly altering gene expression or promoting cellular responses. Let-7e-5p can either downregulate by upregulating the function of PECAM-1.

Another hypothesis revolves around the connection between PECAM-1 and SOX7.PECAM1 is also expressed more in Atherosclerosis during plaque formation, and SOX7 is more. According to studies, we can hypothesize that SOX7 gets upregulated and, in turn, upregulated PECAM-1 and down-regulates let-7e-5p.

Our third hypothesis connects the miR-let-7e-5p with SOX7. Since the miRNAs are hypothesized based on the studies to upregulate the expression of PECAM-1, there are chances that miR-let-7e-5p is inhibiting the expression of SOX7.To prove these hypotheses, further validation is required through in vitro studies.

GATA2 which is a transcription factor has been found to play a major role in the proper functioning and making of endothelial cells, these cells make the lining of the interior surface of blood vessels which when disrupted leads to atherosclerosis (39). Studies have shown that GATA2 is upregulated in early atherosclerotic plaques, so it leads to the development of plaque and the disease MiR-550a-3p has been shown to have a protective role when it comes

to the development of atherosclerosis. MiR-550a-3p has been found to be downregulated which contributes to protection against the disease by inhibiting the proliferation, invasion, migration, and apoptosis of vascular smooth muscle cells. In doing so, miR-550a-3p normalizes the function of the injured vascular endothelial cells, hence providing protection against plaque formation and progression. MEF2C has been found to be downregulated in advanced plaque which shows that it also has a protective role against the disease. MEF2C is found to have to regulate the expression of genes which is involved in smooth muscle cell differentiation and proliferation. MEF2C protects against the development of atherosclerosis by inhibiting TLR/NF-κB activation, smooth muscle cell (SMC) migration and proliferation (40).

Based on this information, Hypothesis can be made that Upregulation of GATA2 expression in atherosclerotic plaques may lead to upregulation of miR-550a-3p expression, via direct or indirect mechanisms. For example, GATA2 may bind to the promoter region of miR-550a-3p and increase the expression of factors that would lead to increase in expression of miR-550a-3p. This would lead to the progression of atherosclerosis by allowing proliferation, invasion, migration, and apoptosis of vascular smooth muscle cells, as miR-550a-3p has been shown to target genes involved in these processes. Another hypothesis is that the downregulation of miR-550a-3p expression in atherosclerotic plaques can lead to downregulation of MEF2C, which can lead to protection against plaque formation by stopping the TNF pathway or smooth muscle cells proliferation. Lastly, the upregulation of MEF2C in atherosclerotic plaque may lead to the upregulation of GATA2.This can lead to progression of plaque by increasing endothelial dysfunction since GATA2 is shown to contribute to the endothelial dysfunction. This all could potentially lead to development of plaque or protection against the plaque development, but it further needs to be validated by wet lab analysis.

## 5.Conclusion

Using an integrated multiomics approach to invest the hub calcification proteins in atherosclerosis has provided essential insight into atherosclerosis's molecular mechanism and plaque formation. The analysis found the miRNA, Tf, and genes that may act as a potential biomarker for the calcification in plaque formation. The role of TF GATA2 and SOX7, genes MEF2C and PECAM-1, and MiRNA let-7e-5p and MiRNA mir-550a-3p has not yet been verified by molecular biology, which can further help in understanding the roles of them during calcification in plaque. In the future, they may act as bio makers that can help in understanding and healing the plaque formation process for atherosclerosis patients.

**Data availability**

The datasets used in our study were already available from Gene Expression Omnibus (GEO) under accession numbersGSE28829(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28829) and GSE43292(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE43292).

## 6. References

1. Mohd Nor, N. S., Al-Khateeb, A. M., Chua, Y. A., Mohd Kasim, N. A., & Mohd Nawawi, H. (2019). Heterozygous familial hypercholesterolaemia in a pair of identical twins: a case report and updated review. BMC pediatrics, 19(1), 106. https://doi.org/10.1186/s12887-019-1474-y

2. Paul, S., Lancaster, G. I., & Meikle, P. J. (2019). Plasmalogens: A potential therapeutic target for neurodegenerative and cardiometabolic disease. Progress in lipid research, 74, 186–195. https://doi.org/10.1016/j.plipres.2019.04.003

3. Dichgans, M., Pulit, S. L., & Rosand, J. (2019). Stroke genetics: discovery, biology, and clinical applications. The Lancet. Neurology, 18(6), 587–599. https://doi.org/10.1016/S1474-4422(19)30043-2

4. World Health Organization. (n.d.). Cardiovascular diseases. Retrieved April 12, 2023, from https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1

5. Jebari-Benslaiman, S., Galicia-García, U., Larrea-Sebal, A., Olaetxea, J. R., Alloza, I., Vandenbroeck, K., Benito-Vicente, A., & Martín, C. (2022). Pathophysiology of Atherosclerosis. International journal of molecular sciences, 23(6), 3346. https://doi.org/10.3390/ijms23063346

6. Javadifar, A., Rastgoo, S., Banach, M., Jamialahmadi, T., Johnston, T. P., & Sahebkar, A. (2021). Foam Cells as Therapeutic Targets in Atherosclerosis with a Focus on the Regulatory Roles of Non-Coding RNAs. International journal of molecular sciences, 22(5), 2529. https://doi.org/10.3390/ijms22052529

7. Wilkins, J. T., Gidding, S. S., & Robinson, J. G. (2019). Can atherosclerosis be cured?. Current opinion in lipidology, 30(6), 477–484. https://doi.org/10.1097/MOL.0000000000000644

8. Cheng, H. S., & Wu, J. (2018). MicroRNAs and cardiovascular disease: From pathogenesis to therapeutic implications. Biomedicine & Pharmacotherapy, 109, 1183-1192. doi: 10.1016/j.biopha.2018.10.025

9. Shi, X., Gao, J., Lv, Q., Cai, H., Wang, F., Ye, R., & Liu, X. (2020). Calcification in Atherosclerotic Plaque Vulnerability: Friend or Foe?. Frontiers in physiology, 11, 56. https://doi.org/10.3389/fphys.2020.00056

10. Sommer, B., Bickel, M., Hofstetter, W., & Wetterwald, A. (1996). Expression of matrix proteins during the development of mineralized tissues. Bone, 19(4), 371–380. https://doi.org/10.1016/s8756-3282(96)00218-9

11. Arora, S., Singh, P., Dohare, R., Jha, R., & Ali Syed, M. (2020). Unravelling host-pathogen interactions: ceRNA network in SARS-CoV-2 infection (COVID-19). Gene, 762, 145057. https://doi.org/10.1016/j.gene.2020.145057

12. Xie, G. Y., Xia, M., Miao, Y. R., Luo, M., Zhang, Q., & Guo, A. Y. (2020). FFLtool: a web server for transcription factor and miRNA feed forward loop analysis in human. Bioinformatics (Oxford, England), 36(8), 2605–2607. https://doi.org/10.1093/bioinformatics/btz929

13. Du, Y., Gao, C., Liu, Z., Wang, L., Liu, B., He, F., Zhang, T., Wang, Y., Wang, X., Xu, M., Luo, G. Z., Zhu, Y., Xu, Q., Wang, X., & Kong, W. (2012). Upregulation of a disintegrin and metalloproteinase with thrombospondin motifs-7 by miR-29 repression mediates

vascular smooth muscle calcification. Arteriosclerosis, thrombosis, and vascular biology, 32(11), 2580–2588. https://doi.org/10.1161/ATVBAHA.112.300206

14. Goettsch, C., Rauner, M., Pacyna, N., Hempel, U., Bornstein, S. R., & Hofbauer, L. C. (2011). miR-125b regulates calcification of vascular smooth muscle cells. The American journal of pathology, 179(4), 1594–1600. https://doi.org/10.1016/j.ajpath.2011.06.016
15. von Scheidt, M., Zhao, Y., de Aguiar Vallim, T. Q., Che, N., Wierer, M., Seldin, M. M., Franzén, O., Kurt, Z., Pang, S., Bongiovanni, D., Yamamoto, M., Edwards, P. A., Ruusalepp, A., Kovacic, J. C., Mann, M., Björkegren, J. L. M., Lusis, A. J., Yang, X., & Schunkert, H. (2021). Transcription Factor MAFF (MAF Basic Leucine Zipper Transcription Factor F) Regulates an Atherosclerosis Relevant Network Connecting Inflammation and Cholesterol Metabolism. Circulation, 143(18), 1809–1823. https://doi.org/10.1161/CIRCULATIONAHA.120.050186
16. Barrett, T., Troup, D. B., Wilhite, S. E., Ledoux, P., Rudnev, D., Evangelista, C., & Kim, I. F. (2009). NCBI GEO: archive for high-throughput functional genomic data. Gene Expression Omnibus. GSE28829.
17. McDermott, M. M., Ferrucci, L., Liu, K., Criqui, M. H., Greenland, P., Guralnik, J. M., & Tian, L. (2013). Genome-wide expression study of human carotid atheroma [Data set]. Gene Expression Omnibus. GSE43292.
18. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Research, 13(11), 2498-2504. doi: 10.1101/gr.1239303

19. Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S. L., Jagodnik, K. M., Lachmann, A., McDermott, M. G., Monteiro, C. D., Gundersen, G. W., & Ma'ayan, A. (2016). Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Research, 44(W1), W90-W97. doi: 10.1093/nar/gkw377

20. R Core Team. (2020). R: A language and environment for statistical computing (Version 4.2.2) [Software]. R Foundation for Statistical Computing. https://www.R-project.org/

21. Gray, K. A., Yates, B., Seal, R. L., Wright, M. W., & Bruford, E. A. (2015). Genenames.org: The HGNC resources in 2015. Nucleic Acids Research, 43(Database issue), D1079-D1085. doi: 10.1093/nar/gku1071

22. Szklarczyk, D., Morris, J. H., Cook, H. V., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N. T., Roth, A., Bork, P., Jensen, L. J., & von Mering, C. (2017). The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Research, 45(D1), D362-D368. doi: 10.1093/nar/gkw937

23. Lachmann, A., Xu, H., Krishnan, J., Berger, S. I., Mazloom, A. R., & Ma'ayan, A. (2020). ChEA v3.0: Transcription factor enrichment analysis by orthogonal omics integration. Nucleic Acids Research, 48(W1), W212-W219. doi: 10.1093/nar/gkaa299

24. Sticht, C., De La Torre, C., Parveen, A., Gretz, N., (2018). miRWalk: An online resource for prediction of microRNA binding sites. PLoS ONE, 13(10), e0206239. doi: 10.1371/journal.pone.0206239

25. Pan, W., Liang, J., Tang, H., Fang, X., Wang, F., Ding, Y., Huang, H., & Zhang, H. (2020). Differentially expressed microRNA profiles in exosomes from vascular smooth muscle

cells associated with coronary artery calcification. The international journal of biochemistry & cell biology, 118, 105645. https://doi.org/10.1016/j.biocel.2019.105645

26.Dong, G., Fan, H., Yang, Y., Zhao, G., You, M., Wang, T., & Hou, Y. (2019). 17β-Estradiol enhances the activation of IFN-α signaling in B cells by down-regulating the expression of let-7e-5p, miR-98-5p and miR-145a-5p that target IKKε. Cellular Signalling, 54, 155-165. doi: 10.1016/j.cellsig.2018.11.012

27. Wang, C., Liu, S., Li, J., Cheng, Y., Wang, Z., Feng, T., Lu, G., Wang, S., Song, J., Xia, P., & Hao, L. (2021). Biological functions of Let-7e-5p in promoting the differentiation of MC3T3-E1 cells. Frontiers in Cell and Developmental Biology, 9, 671170. doi: 10.3389/fcell.2021.671170

28. Lin, M. E., Chen, T., Leaf, E. M., Speer, M. Y., & Giachelli, C. M. (2015). Runx2 Expression in Smooth Muscle Cells Is Required for Arterial Medial Calcification in Mice. The American journal of pathology, 185(7), 1958–1969. https://doi.org/10.1016/j.ajpath.2015.03.020

29. Bao, M. H., Feng, X., Zhang, Y. W., Lou, X. Y., Cheng, Y., & Zhou, H. H. (2013). Let-7 in cardiovascular diseases, heart development and cardiovascular differentiation from stem cells. International journal of molecular sciences, 14(11), 23086–23102. https://doi.org/10.3390/ijms141123086

30. He, L., & Hannon, G. J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. Nature reviews. Genetics, 5(7), 522–531. https://doi.org/10.1038/nrg1379

31. Huang, S., Lv, Z., Guo, Y., Li, L., Zhang, Y., Zhou, L., Yang, B., Wu, S., Zhang, Y., Xie, C., Li, S., & Cheng, J. (2016). Identification of Blood Let-7e-5p as a Biomarker for Ischemic Stroke. PloS one, 11(10), e0163951. https://doi.org/10.1371/journal.pone.0163951

32. Chen, L., Wei, K., Li, J., Li, Y., Cao, H., & Zheng, Z. (2022). Integrated analysis of LncRNA-mediated ceRNA network in calcific aortic valve disease. Cells, 11(14), 2204. doi: 10.3390/cells11142204

33. Jasper lundstrom. (2008). The effects of exercise on cardiovascular health (Doctoral dissertation). Karolinska Institutet, Stockholm, Sweden. Retrieved from https://openarchive.ki.se/xmlui/bitstream/handle/10616/39609/thesis.pdf;sequence=1

34. Doyle, M. J., Magli, A., Estharabadi, N., Amundsen, D., Mills, L. J., & Martin, C. M. (2019). Sox7 Regulates Lineage Decisions in Cardiovascular Progenitor Cells. Stem cells and development, 28(16), 1089–1103. https://doi.org/10.1089/scd.2019.0040

35. Albanese, I., Khan, K., Barratt, B., Al-Kindi, H., & Schwertani, A. (2018). Atherosclerotic Calcification: Wnt Is the Hint. Journal of the American Heart Association, 7(4), e007356. https://doi.org/10.1161/JAHA.117.007356

36. Nguyen, M-A., Karunakaran, D., Geoffrion, M., Cheng, H. S., Tandoc, K., Perisic Matic, L., Hedin, U., Maegdefessel, L., Fish, J. E., & Rayner, K. J. (2018). Extracellular vesicles secreted by atherogenic macrophages transfer microRNA to inhibit cell migration. Arteriosclerosis, Thrombosis, and Vascular Biology, 38, 49-63. doi: 10.1161/ATVBAHA.117.309795

37. Stevens, H. Y., Melchior, B., Bell, K. S., Yun, S., Yeh, J. C., & Frangos, J. A. (2008). PECAM-1 is a critical mediator of atherosclerosis. Disease models & mechanisms, 1(2-3), 175–179. https://doi.org/10.1242/dmm.000547

38. Goel, R., Schrank, B. R., Arora, S., Boylan, B., Fleming, B., Miura, H., Newman, P. J., Molthen, R. C., & Newman, D. K. (2008). Site-specific effects of PECAM-1 on

atherosclerosis in LDL receptor-deficient mice. Arteriosclerosis, Thrombosis, and Vascular Biology, 28, 1996-2002. doi: 10.1161/ATVBAHA.108.172270

39. Kanki, Y., Kohro, T., Jiang, S., Tsutsumi, S., Mimura, I., Suehiro, J., Wada, Y., Ohta, Y., Ihara, S., Iwanari, H., Naito, M., Hamakubo, T., Aburatani, H., Kodama, T., & Minami, T. (2011). Epigenetically coordinated GATA2 binding is necessary for endothelium-specific endomucin expression. The EMBO journal, 30(13), 2582–2595. https://doi.org/10.1038/emboj.2011.173

40. He, L., Wang, T., Lu, F-M., Xu, J., & Cong, H-L. (2016). MEF2C protects against the development of atherosclerosis via inhibiting TLR/NF-κB activation. International Journal of Clinical and Experimental Medicine, 9(4), 7067-7074.

## 7. Vitae

Charmy Shah
Oakwood, Georgia -30566

### Education

**GEORGIA STATE UNIVERISTY**

Master of Interdisciplinary Science in Biomedical Science and Enterprise Atlanta, GA
Relevant Coursework: Graduate research assistantship awarded. Aug 2021-current

**UNIVERSITY OF ILLINOIS AT CHICAGO** Chicago, IL
Bachelor of Science in Biology Jan2019- Dec 2020
Minor in Art and Nutrition

### Experience

**GEORGIA STATE UNIVERSITY** Atlanta, GA
**Graduate Research Assistant** August 2021-current

- Performed wet-lab and molecular biology scientific research, using methods including enzyme-linked immunosorbent assay (ELISA), Western blot analysis, cell fractionation, gene analysis and polymerase chain reaction (PCR).
- Carried out NGS and network analysis using bioinformatics work flow to find hub calcification markers.
- Served on research team that is carrying out the research to find the gene affecting formation of atherosclerosis plaque.
- Carried out cell culture preparation to study the effect of SYMD2 on translocation of p65 in HEK293 cells.
- Maintained lab and chemical inventories, safely disposing of waste and accurately calibrating laboratory equipment.
- Prepared materials for reports, presentations, and submission to peer-reviewed journal publications.
- Observed and assisted in performing basic animal treatments for mice such as handling, restraining, blood sample collection and injections.

**UNIVERSITY OF ILLINOIS AT CHICAGO** Chicago, Illinois
**EMT student** Aug 2020-Dec 2020

- Worked at emergency room to get experience in patient care as an EMT student.
- Leveraged comprehensive knowledge of emergency medicine to assess, treat, stabilize and transport seriously ill or injured patients to area hospitals and trauma centers.

- Maintained levelheadedness and efficiency in high-pressure situations and effectively prioritizing tasks to save lives and provide medical care.

**OAKTON COMMUNITY COLLEGE** Des Plaines, Illinois
**Chemistry Lab Assistant** March 2017-Dec 2018

- Served as a chemistry lab assistant for approximately 4 sections of Chemistry per week, with a minimum of 20 students in each section.
- Cleaned and organized lab equipment including 100s of chemicals and glassware, and prepared lab devices such as computer analyzers for experiments.
- Achieved a level of experience in teaching while assisting and being involved in the department of chemistry

**HIRALAL AND SONS** Ahmedabad, Gujarat
**Assistant Manager** January 2013-Dec 2015

- Increased quarterly sales by 15% after implementing proper customer feedback working mechanism.
- Identified and developed new business opportunities- expanded from selling locally to internationally, bringing in 8 new clients on an average every month.
- Prospecting potential new clients- by collecting customer satisfaction data and work towards acquiring and retaining many new clients
- Trained the workers for better communication and written skills.

**GUJARAT BLOOD BANK** Ahmedabad, Gujarat
**Volunteer** March 2014-Dec 2015

- Assisted the receptionist with answering multi-line phone systems, greeted patriots, providing directions and information.
- Shadowed during a variety of routine blood drawing procedures such as venipuncture; a technique using standard equipment includes vacutainer tubes and sleeves, tourniquets, syringes, and butterfly needles.
- Used Composite Health Care System (CHCS) to log inpatient lab orders and results of 50 patients each day.
- Protected patients' rights by maintaining the confidentiality of personal and financial information.

**SHYAM EAR NOSE AND THROAT HOSPITAL** Ahmedabad, Gujarat
**Volunteer** June 2015-Dec 2015

- Shadowed doctors during surgery
- Directed individuals and families around the hospital and provided them with necessary information
- Provided medical interpretation to people who do not understand medical terminology

## Leadership & Activities

**UNIVERSITY OF ILLINOIS AT CHICAGO** Chicago, Illinois
**Premedical Student Representative** Jan 2020 – Dec 2020

- Aided the society in finding the differences in the available health care for underprivileged people.
- Organized blood drives during covid.
- Represented the premedical community for minorities.

## Skills & Interests

**Technical:** basic proficiency in SQLand Python, Ms. Excel, MS. Word, Proficient in MS Excel including VBA, Adobe Photoshop,
**Language:** fluent in English, Hindi, and Guajarati
**Interests:** stock market, finance, Art, and photography
**Other**: certified in Laboratory Animal Research Support

## Certifications

- SAS Certified Associate: Programming Fundamentals Using SAS 9.4
- Harvard Core: Business Analysis, issued by Harvard Business School Online
- Creating Powerful Dashboard with PowerBI, issued by GrowthSchool
- Foundations: Data, Data, Everywhere, issued by Coursera
- SQL Essential Training, issued by LinkedIn
- SQL for Non-Programmers, issued by LinkedIn
- SQL Programming, issued by LinkedIn
- Conflict of Interest, issued by CITI Program
- Group 1 - Biomedical, issued by CITI Program
- Social and Behavioral Responsible Conduct of Research Course 1, issued by CITI Program

## Projects

- Identification of key genes and pathways involved in atherosclerosis using WCGNA analysis.
- Investigating the role of hub calcification proteins in Atherosclerosis (Early and late plaque formation) via integrated multiomics and network-based approach
- Identification of atherosclerosis prevalence in population using data from NIH and SAS.

## Publications

- Investigating the role of hub calcification proteins in Atherosclerosis (Early and late plaque formation) via integrated multiomics and network-based approach (In process)