

# Bioinformatics analysis of common differential genes of coronary artery disease and carotid atherosclerosis

**Luxia Song**

Graduate School, Beijing University of Chinese Medicine

**Jie Zhang**

Graduate School, Beijing University of Chinese Medicine

**Yixuan Fan**

Graduate School, Beijing University of Chinese Medicine

**Qiyu Liu**

Graduate School, Beijing University of Chinese Medicine

**Baoyi Guan**

Graduate School, China Academy of Chinese Medical Sciences

**Hao Xu** (✉ [xuhaotcm@hotmail.com](mailto:xuhaotcm@hotmail.com))

National Clinical Research Center for Chinese Medicine Cardiology, Xiyuan Hospital, China Academy of Chinese Medical Sciences

---

## Research Article

**Keywords:** Coronary Artery Disease, Carotid Plaque, Biomarkers, Hub Genes, Immune Response, Inflammatory Response, Cell Chemotaxis

**Posted Date:** March 5th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-273213/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Coronary artery disease(CAD) is one of the most fatal diseases in the world, which seriously threatens human health. Studies have demonstrated that the appearance of carotid plaque is related to the risk of CAD, but the common differential genes and mechanism between these two conditions are still unclear. Our study identified the common differential genes between carotid atherosclerosis tissues and blood samples of CAD patients, aiming to search promising biomarkers in CAD predicting and diagnosing. We obtain datasets of GSE100927 and GSE56885 from GENE EXPRESSION OMNIBUS (GEO) database. Through scanning their mutual differentially expressed genes(DEGs), we performed Gene Ontology (GO), Kyoto Encyclopedia of Genes, Genomes (KEGG) analysis, and PPI analysis to get hub genes between these two conditions. We found that both CAD blood samples and carotid atherosclerotic plaque tissues were related to immune response, inflammatory response and cell chemotaxis. Followed by PPI network construction, MCODE analysis found that 1 subnetwork, including CCR5, CCR2, CXCR4 and C5AR1, was extracted, which concerned as hub genes of the two datasets. Indicating that CCR5, CCR2, CXCR4 and C5AR1 maybe potential candidate biomarkers for CAD prediction in patients with carotid plaques.

## Introduction

Clinical studies have shown that carotid atherosclerotic plaque is closely related to the risk of CAD, and the application of carotid ultrasound can be used as an independent factor to predict CAD<sup>1,2</sup>. In order to improve the prediction of CAD, Vijay Namb et al<sup>3</sup> found that adding common carotid artery intima–media thickness (CCA-IMT) and plaque information to traditional risk factors could be a better choice for risk prediction. So far, with the increasing use of bioinformatics methods in medical field, a large number of researchers have used genetic information to excavate genetic differences in patients with carotid atherosclerosis or CAD to screen for disease related markers. It was found that CCA-IMT and plaques are also genetically related to CAD and stroke in a genome-wide association study (GWAS) <sup>4</sup>. However, most of the researches only focused on screening differential genes in carotid atherosclerosis or CAD alone, ignoring the relationship between these two conditions. Therefore, we aim to jointly explore the common differential genes between carotid atherosclerotic tissues and blood samples of CAD patients through downloading two microarray datasets from GEO database, intending to investigate the underlying molecular mechanisms of the mutual occurrence and the development of the two diseases and providing candidate biomarkers for CAD prediction in patients with carotid plaques.

## Results

### Datasets validation

The main purpose of the standardization of gene expression profile is to eliminate deviation in expression caused by experimental technology. Through observing the boxplots of the two datasets (Figure1A and Figure1B) showing below, we believed that the data in GSE100927 and GSE56885 was comparable, and could truly reflect the differences in biological functions. PCA analysis is a process of information

concentration, which outputs with new variables like PC1, PC2...and so on. In the PCA cluster graph of PC1 and PC2 of GSE100927 (Figure1C), case group (carotid atherosclerotic tissues) and control group (controlled carotid artery tissues) were distinguished. In PCA cluster graph of GSE56885(Figure1D), there was a close correlation in control group (healthy control) samples, and samples in case group (CAD patients) showed a strong correlation as well.

### **DEGs in GSE100927, GSE56885, and their common DEGs**

With NetworkAnalyst 3.0 software, we scanned the DEGs in GSE100927 and GSE56885, drawing volcano maps to visualize distribution of the DEGs (Figure2A and Figure2B). There were 733 DEGs between the carotid atherosclerotic tissues and controlled carotid artery tissues in GSE100927 dataset. And in GSE56885 dataset, which contained blood samples of healthy controls and CAD patients, a total of 183 DEGs were identified. Venn diagram, as shown in Figure2C, demonstrated that there were 23 genes simultaneously expressed in both GSE100927 and GSE56885 datasets.

### **GO enrichment analysis and KEGG pathway analysis**

In our study, we applied DAVID to perform GO function annotation and KEGG analysis on GSE100927 datasets, GSE56885 datasets and their common DEGs. Table 1 recorded the top five functional enrichment results of BP, CC, MF, KEGG analysis, P value and the DEGs corresponding to each term (the complete gene list is in Supplementary Table 1) of GSE100927. The biological process differences between the carotid atherosclerotic tissues and controlled carotid artery tissues were mainly related to immunity and inflammation, which contained "immune response", "inflammatory response", "innate immune response", and so on. The cellular component differences were "plasma membrane" and "integral component of plasma membrane". In terms of molecular functions, the distinction between the two groups were "MHC class II receptor activity" and "structural constituent of muscle". In KEGG analysis, DEGs were mainly enriched in "Staphylococcus aureus infection" and "Leishmaniasis" pathways. The relevant results in GSE100927 were displayed in circle diagram (Figure3A) and network diagram (Figure3D). Table 2 showed the GO enrichment and KEGG analysis results of GSE56885(the complete gene list is in Supplementary Table 2). In Biological Process, the enrichment results were also concentrated on immunity and inflammation in blood samples of healthy volunteers and CAD patients. However, "chemotaxis" and "chemokine-mediated signaling pathway" were prominent differences between the two groups. In Cellular Component, "cytoplasm" and "nucleus" were top 2 enriched terms in dataset. In terms of Molecular Function, the differences between the two groups were mainly concentrated on "protein binding" and "transcription factor binding". And in KEGG analysis, the pathways of DEGs enrichment were mainly focused on "NOD-like receptor signaling pathway" and "TNF signaling pathway". The relevant results in GSE56885 were shown in circle diagram (Figure3B) and network diagram (Figure3E). Table 3 displays the GO enrichment and KEGG analysis results of the common DEGs

of ges100927 and GSE56885. The Biological Process of the common DEGs enrichment involved immune response, inflammatory response and cell chemotaxis. In terms of cell composition, the top two results of enrichment were "integral component of plasma membrane" and "external side of plasma membrane". "C-C chemokine receptor activity", "chemokine receptor activity", and " coreceptor activity" were major molecular function enriched in common DEGs. And KEGG analysis revealed that "Cytokine-cytokine receptor interaction" and " Chemokine signaling pathway " were main enriched pathway in common DEGs.

**Table1:** GO and KEGG pathway analysis of GSE100927

classification	Term	P-Value	Genes
Biological Process	GO:0006955:immune response	2.96E-20	CSF3, ADAMDEC1, IL1RN, GPR65, LST1, NCF4, AQP9, CXCL14, TNF, CTSS...C1QC, total64 gene targets.
	GO:0006954:inflammatory response	3.36E-19	GPR68, ITGB2, ITGAL, LIPA, TNF, PIK3CG, PYCARD, CCRL2, BLNK, C3AR1...IL17B, total59 gene targets.
	GO:0045087:innate immune response	7.11E-12	CD84, NCF1, NCF2, TREM2, TREM1, PIK3CG, PYCARD, CLEC5A, TNFAIP8L2, JAK3...C1QC, total51 gene targets.
	GO:0002504:antigen processing and presentation of peptide or polysaccharide antigen via MHC class II	1.30E-10	HLA-DMA, HLA-DRB5, HLA-DMB, HLA-DPB1, HLA-DRA, HLA-DQA2, HLA-DQB2, HLA-DRB1, HLA-DQA1, HLA-DPA1, HLA-DQB1
	GO:0002250:adaptive immune response	1.36E-10	LILRA6, CD84, UNC93B1, LILRA2, PIK3CG, CTSS, LILRA4, LAMP3, SLAMF7, CSK...TNFRSF21, total27 gene targets.
Cellular Component	GO:0005886:plasma membrane	2.22E-17	IL1RN, VIPR2, HPSE2, CPNE7, AQP9, MLC1, PKD2L1, IFI30, TBC1D10C... MYO1G, total255 gene targets.
	GO:0005887:integral component of plasma membrane	2.76E-15	CD84, VIPR2, CD83, CSF3R, PLXND1, CD82, GPR68, GPR65, AQP9, PCDH11X... SLC28A3, total117 gene targets.
	GO:0070062:extracellular exosome	2.62E-11	FCGBP, ALDH1L1, CD84, IL1RN, CD82, CPNE7, PCDH11X, HP, LIPA, ACTG2...C1QC, total175 gene targets.
	GO:0042613:MHC class II protein complex	1.19E-10	CD74, HLA-DMA, HLA-DRB5, HLA-DMB, HLA-DPB1, HLA-DRA, HLA-DQA2, HLA-DQB2, HLA-DRB1, HLA-DQA1, HLA-DPA1, HLA-DQB1
	GO:0009986:cell surface	4.31E-10	SRPX, ITGAM, ITGB2, PKD2L1, ITGAL, TNF, PLAU, CLEC5A, ITGAX, ITGB7...FERMT2, total54 gene targets.
Molecular Function	GO:0032395:MHC class II receptor activity	2.21E-08	HLA-DMA, HLA-DPB1, HLA-DRA, HLA-DQA2, HLA-DQB2, HLA-DRB1, HLA-DQA1, HLA-DPA1, HLA-DQB1
	GO:0008307:structural constituent of muscle	3.38E-07	MYOM1, SYNM, SMTN, PDLIM3, ACTN2, TPM2, MYL2, MYBPH,

			NEXN, TCAP, MYH11, MYL9
	GO:0004872:receptor activity	9.03E-07	RTN4R, CSF3R, ITGB2, TREM2, LILRA2, TREM1, ANPEP, ITGAX, ITGB7, LDLR...TLR2, total26 gene targets.
	GO:0005102:receptor binding	2.14E-06	SELPLG, GRP, NXPH3, KCNA5, ARRB2, PKD2, DPP4, C3, NTF3, CSK...ANGPTL1, total34 gene targets.
	GO:0003779:actin binding	2.53E-05	TAGLN, FKBP15, SYNPO2, LMOD1, CXCR4, CAPG, LSP1, NOD2, CORO1A, CNN1... MYO1G, total27 gene targets.
KEGG_PATHWAY	hsa05150:Staphylococcus aureus infection	1.11E-16	C1QB, C1QA, HLA-DRB5, ITGAM, SELPLG, ITGB2, C5AR1, ITGAL, C2, C3...HLA-DQB1, total24 gene targets.
	hsa05140:Leishmaniasis	1.18E-15	ITGAM, NCF1, NCF2, NCF4, ITGB2, TNF, C3, HLA-DMA, FCGR3A, HLA-DMB...TLR2, total26 gene targets.
	hsa05152:Tuberculosis	7.70E-13	ITGAM, ITGB2, LSP1, NOD2, CORO1A, TNF, CTSS, C3, CYP27B1, HLA-DMA...TLR2, total36 gene targets.
	hsa04145:Phagosome	8.94E-13	ATP6V1A, ITGAM, NCF1, NCF2, NCF4, ITGB2, CORO1A, CTSS, C3, HLA-DMA...TLR2, total33 gene targets.
	hsa05323:Rheumatoid arthritis	2.65E-12	ATP6V1A, CCL3L3, ITGB2, ITGAL, TNF, HLA-DMA, HLA-DMB, CCL5, CCL3, ACP5...TLR2, total25 gene targets.

**Table2:** GO and KEGG pathway analysis of GSE56885.

classification	Term	P-Value	Genes
Biological Process	GO:0006954:inflammatory response	1.19E-07	CEBPB, CXCL8, RIPK2, CD180, C5AR1, CXCR4, TNFAIP3, CXCL1, PTGS2, CXCL2, BCL6, IL1B, NFKBIZ, CCL3, OLR1, NLRP3, CCR5, CCR2
	GO:0006935:chemotaxis	2.39E-06	CXCL8, C5AR1, PLAUR, CCR9, CCL3, CXCR4, CXCL1, CCR5, CXCL2, CCR2
	GO:0070098:chemokine-mediated signaling pathway	4.78E-06	CXCL8, CCR9, CCL3, CXCR4, CXCL1, CCR5, CXCL2, CCR2
		7.20E-06	CD83, RIPK2, PDE4D, PDE4B, CCR2
	GO:0032743:positive regulation of interleukin-2 production		
	GO:0006955:immune response	1.15E-05	HLA-DRB5, CEBPB, CXCL8, TNFSF14, C5AR1, CXCL1, CXCL2, NFIL3, CD8B...CCR2, total 16 target genes.
	GO:0071222:cellular response to lipopolysaccharide	1.19E-05	ZFP36, CEBPB, CXCL8, PDE4D, CD180, PDE4B, NLRP3, TNFAIP3, CCR5
	GO:0030593:neutrophil chemotaxis	3.73E-05	CXCL8, IL1B, PDE4D, C5AR1, PDE4B, CCL3, TREM1
Cellular Component	GO:0005737:cytoplasm	3.59E-05	ZNF174, TNFAIP3, KIF11, PRDM1, BZW1, CDC20, EPB41L4A, RGS2, ZFP36, FTH1...CDKN3, total 74 target genes.
		2.37E-04	C2ORF42, ZNF174, TNFAIP3, PRDM1, CDC20, RGS2, ZFP36, FTH1, ENC1, NFKBIZ...CDKN3, total 73 target genes.
	GO:0005634:nucleus		
	GO:0005829:cytosol	0.004103	BTG2, PFKFB3, TACSTD2, TNFAIP3, IRS2, KIF11, TYMS, AURKA, CDC20, DNAJB1...KBTBD7, total 46 target genes.
	GO:0005515:protein binding	8.86E-06	SIRPG, IRS2, KLHL35, PRDM1, GIMAP7, CDC20, RGS2, ZFP36,

Molecular Function			CNST... SIT1, total112 gene targets.
	GO:0008134:transcription factor binding	4.04E-04	PURB, NFKBIA, PER1, CEBPB, JUN, MAFB, ID1, NLRP3, ID3, TRIB1, JUNB
	GO:0042803:protein homodimerization activity	0.001671	JUN, CEBPB, ZNF174, BCL2A1, RIPK2, TYMS, PTGS2, GIMAP7, ADRA2A, NR4A2, NR4A3, NACC2, ID1, NAMPT, MTPAP, PSPH, CCR2
	GO:0000978:RNA polymerase II core promoter proximal region sequence-specific DNA binding	0.002188	PER1, EGR2, CEBPB, JUN, NR4A3, MAFB, NACC2, PRDM1, MXD1, JUNB, FOSL2
	GO:0035259:glucocorticoid receptor binding	0.004699	NR4A2, CEBPB, NR4A3
KEGG_PATHWAY	hsa04621:NOD-like receptor signaling pathway	3.23E-06	NFKBIA, CXCL8, RIPK2, IL1B, NLRP3, TNFAIP3, CXCL1, CXCL2
	hsa04668:TNF signaling pathway	3.02E-05	NFKBIA, CEBPB, JUN, IL1B, TNFAIP3, CXCL1, PTGS2, CXCL2, JUNB
	hsa04064:NF-kappa B signaling pathway	6.10E-05	NFKBIA, CXCL8, TNFSF14, BCL2A1, DDX58, IL1B, TNFAIP3, PTGS2
	hsa04380:Osteoclast differentiation	7.74E-04	NFKBIA, JUN, SOCS1, IL1B, SIRPG, GAB2, JUNB, FOSL2
	hsa04062:Chemokine signaling pathway	0.00135	NFKBIA, CXCL8, CCR9, CCL3, CXCR4, CXCL1, CCR5, CXCL2, CCR2

**Table3:** GO and KEGG pathway analysis of Common DEGs

classification	Term	P-Value	Genes
Biological Process	GO:0006955:immune response	7.30E-07	HLA-DRB5, NFIL3, RGS1, IL1B, C5AR1, CCL3, CCR5, CCR2
	GO:0006954:inflammatory response	6.98E-06	IL1B, CD180, C5AR1, CCL3, CXCR4, CCR5, CCR2
	GO:0006935:chemotaxis	1.75E-05	C5AR1, CCL3, CXCR4, CCR5, CCR2
	GO:0030593:neutrophil chemotaxis	8.47E-05	IL1B, C5AR1, CCL3, TREM1
	GO:0070098:chemokine-mediated signaling pathway	1.05E-04	CCL3, CXCR4, CCR5, CCR2
	GO:0002407:dendritic cell chemotaxis	2.20E-04	CXCR4, CCR5, CCR2
	GO:0007204:positive regulation of cytosolic calcium ion concentration	6.85E-04	C5AR1, CXCR4, CCR5, CCR2
	GO:0019722:calcium-mediated signaling	0.00201	CCL3, CXCR4, CCR5
GO:0006968:cellular defense response	0.002955	C5AR1, CCR5, CCR2	
Cellular Component	GO:0005887:integral component of plasma membrane	0.024235	CD83, CD180, C5AR1, CCR5, PAG1, CCR2
	GO:0009897:external side of plasma membrane	0.026937	HLA-DRB5, CD83, CCR5
	GO:0005886:plasma membrane	0.042248	HLA-DRB5, CD83, RGS1, C5AR1, CXCR4, CCR5, SGK1, TREM1, PAG1, CCR2
Molecular Function		0.014129	CCR5, CCR2
	GO:0016493:C-C chemokine receptor activity		
	GO:0004950:chemokine receptor activity	0.019961	CCR5, CCR2
	GO:0015026:coreceptor activity	0.034969	CXCR4, CCR5
KEGG_PATHWAY		4.12E-04	IL1B, CCL3, CXCR4, CCR5, CCR2

hsa04060:Cytokine-cytokine receptor interaction		
hsa04062:Chemokine signaling pathway	0.002734	CCL3, CXCR4, CCR5, CCR2
hsa05323:Rheumatoid arthritis	0.008256	HLA-DRB5, IL1B, CCL3

## PPI network analysis

Previously researches have proved that the interaction between proteins determine the biological functions in human organism. In-depth exploration of protein interactions exerts broad research prospects. We created PPI networks by string software (Figure4A, Figure4B and Figure4C) for DEGs in GSE100927, GSE56885 and their common DEGs. To discover the dense regions interacting in the PPI network, we used MCODE to make further exploration. MCODE is the most commonly used algorithm for mining protein complexes. According to MCODE analysis, GSE100927 had 13 subnetwork (Figure4D) in PPI network, and their annotation mainly included: activation of immune response, human phagosome, and human adaptive immune system (table4). Totally 2 subnetworks were found in GSE56885 (Figure4E), which mainly concerned with cell chemotaxis (table5). And in common DEGs, 1 subnetwork existed in PPI network(Figure4F), relating to dendritic cell chemotaxis, dendritic cell migration and peptide ligand-binding receptors (table6).

**Table4:** MCODE annotation of GSE100927

Network	Annotation
Symbols	GO:0002274 myeloid leukocyte activation -44.8; GO:0002366 leukocyte activation involved in immune response -41.2; GO:0002263 cell activation involved in immune response -41.0
Symbols_MCODE_ALL	GO:0002253 activation of immune response -27.8; hsa04145 Phagosome -27.7; R-HSA-1280218 Adaptive Immune System -27.4
Symbols_SUB1_MCODE_1	R-HSA-373076 Class A/1 (Rhodopsin-like receptors) -28.3; R-HSA-418594 G alpha (i) signaling events -26.6; R-HSA-500792 GPCR ligand binding -25.5
Symbols_SUB1_MCODE_2	R-HSA-390522 Striated Muscle Contraction -20.4; GO:0033275 actin-myosin filament sliding -20.1; GO:0030049 muscle filament sliding -20.1
Symbols_SUB1_MCODE_3	M17200 SA B CELL RECEPTOR COMPLEXES -9.9; ko04662 B cell receptor signaling pathway -7.9; hsa04662 B cell receptor signaling pathway -7.7
Symbols_SUB1_MCODE_4	R-HSA-416476 G alpha (q) signaling events -8.8; GO:0007200 phospholipase C-activating G protein-coupled receptor signaling pathway -7.9;R-HSA-373076 Class A/1 (Rhodopsin-like receptors) -7.9
Symbols_SUB1_MCODE_5	R-HSA-9660821 ADORA2B mediated anti-inflammatory cytokines production -14.0; R-HSA-9664433 Leishmania parasite growth and survival -13.4; R-HSA-9662851 Anti-inflammatory response favoring Leishmania parasite infection -13.4
Symbols_SUB1_MCODE_6	

	<p>R-HSA-1236973 Cross-presentation of particulate exogenous antigens (phagosomes) -18.4;</p> <p>R-HSA-5668599 RHO GTPases Activate NADPH Oxidases -15.5;</p> <p>R-HSA-1222556 ROS and RNS production in phagocytes -14.6</p>
Symbols_SUB1_MCODE_7	<p>R-HSA-8856828 Clathrin-mediated endocytosis -9.1;</p> <p>R-HSA-8856825 Cargo recognition for clathrin-mediated endocytosis -6.7;</p> <p>R-HSA-199991 Membrane Trafficking -6.6</p>
Symbols_SUB1_MCODE_8	<p>R-HSA-166663 Initial triggering of complement -12.5;</p> <p>CORUM:6418 C1q complex -12.0;</p> <p>R-HSA-977606 Regulation of Complement cascade -11.2</p>
Symbols_SUB1_MCODE_9	<p>M88 PID CD8 TCR PATHWAY -10.9;</p> <p>M34 PID TCR PATHWAY -10.6;</p> <p>GO:0050852 T cell receptor signaling pathway -8.5</p>
Symbols_SUB1_MCODE_10	<p>R-HSA-445355 Smooth Muscle Contraction -11.4;</p> <p>R-HSA-397014 Muscle contraction -8.5;</p> <p>GO:0006936 muscle contraction -7.6</p>
Symbols_SUB1_MCODE_11	<p>R-HSA-5602498 MyD88 deficiency (TLR2/4) -14.1;</p> <p>R-HSA-5603041 IRAK4 deficiency (TLR2/4) -13.9;</p> <p>R-HSA-5686938 Regulation of TLR by endogenous ligand -12.8</p>
Symbols_SUB1_MCODE_12	<p>M141 PID PI3KCI PATHWAY -8.3</p>
Symbols_SUB1_MCODE_13	<p>GO:0048011 neurotrophin TRK receptor signaling pathway -8.8;</p> <p>GO:0038179 neurotrophin signaling pathway -8.6;</p> <p>M187 PID TRKR PATHWAY -8.0</p>

**Table5:** MCODE annotation of GSE56885

Network	Annotation
gene_symbol	GO:0006935 chemotaxis -14.6; GO:0042330 taxis -14.6; GO:0060326 cell chemotaxis -14.2
gene_symbol_MCODE_ALL	R-HSA-373076 Class A/1 (Rhodopsin-like receptors) -18.7; R-HSA-418594 G alpha (i) signaling events -17.7; R-HSA-500792 GPCR ligand binding -17.1
gene_symbol_SUB1_MCODE_1	R-HSA-373076 Class A/1 (Rhodopsin-like receptors) -21.2; R-HSA-418594 G alpha (i) signaling events -20.3; R-HSA-500792 GPCR ligand binding -19.6
gene_symbol_SUB1_MCODE_2	ko04110 Cell cycle -7.1; ko04114 Oocyte meiosis -7.1; R-HSA-2500257 Resolution of Sister Chromatid Cohesion -7.1

**Table6:** MCODE annotation of common genes

Network	Annotation
hits	R-HSA-375276 Peptide ligand-binding receptors -9.4; R-HSA-380108 Chemokine receptors bind chemokines -9.2; GO:0030595 leukocyte chemotaxis -9.1
hits_MCODE_ALL	GO:0002407 dendritic cell chemotaxis -8.7; R-HSA-375276 Peptide ligand-binding receptors -8.6; GO:0036336 dendritic cell migration -8.4
hits_SUB1_MCODE_1	GO:0002407 dendritic cell chemotaxis -8.7; R-HSA-375276 Peptide ligand-binding receptors -8.6; GO:0036336 dendritic cell migration -8.4

## Discussion

By analyzing the common DEGs in carotid atherosclerotic plaque tissues and blood samples of patients with CAD, we obtained the following 23 common genes: CD83, RGS1, CD180, C5AR1, CCNB2, PAG1, HLA-DRB5, EGR2, MAFB, ID1, TREM1, TNFAIP8L2, CCR5, CCR2, ASPM, IL1B, SGK1, BCL2A1, CXCR4, NFIL3, TCEAL2, CCL3, and TMEM176B. The biological functions of these genes mainly focused on immune response, inflammatory response, chemotaxis, neutrophil chemotaxis, chemokine-mediated signaling pathway, and dendritic cell chemotaxis. While the molecular functions were enriched in chemokine receptor activity and the results of pathway enrichment were cytokine-cytokine receptor interaction and chemokine signaling pathway.

Similarly, we obtained the hub genes CCR5, CCR2, CXCR4 and C5AR1 through mining PPI network by MCODE. Their functional enrichment also focused on cell chemotaxis, cell migration and peptide ligand-binding receptors, which were related to inflammation. Immune response was also inseparable.

According to the bioinformatics analysis results mentioned above, inflammation and immune response are mutual primary process in CAD patients and carotid atherosclerosis patients. Atherosclerosis is a chronic inflammatory disease. There are multiple types of cells existing in the plaques, including monocytes, macrophages, neutrophils, dendritic cells, T cell and B cell. Chemokines are cytokines that could induce targeted cells chemotaxis. In atherosclerosis, chemokines regulate the chemotaxis of leukocytes by binding to G protein-coupled receptors (GPCRs), and thus participate in the process of atherosclerosis and cardiovascular diseases<sup>5,6</sup>.

CXC-motif chemokine ligand 12(CXCL12, stromal cell-derived factor SDF-1), a chemokine for T cells and monocytes, has the function of activating platelet activity and is involved in the progression of atherosclerosis after binding to CXCR4<sup>7</sup>. Compared with healthy artery, CXCR4 and CXCL12 showed an up-regulated trend in both mRNA and protein levels, indicating an increase expression in both stable and unstable carotid atherosclerotic plaques, especially in macrophages<sup>8</sup>. Investigation of CXCR4 mimic revealed that CXCR4 mimic could block macrophage migration-inhibitory factor (MIF), an atypical chemokine that promotes atherosclerosis through CXCR4, and inhibit inflammation process in atherosclerosis<sup>9</sup>. Through applying iPS-CRISPR/Cas9 technology to explore the role of platelet chemokine CXCL14 in atherosclerosis, researchers found that CXCL14 and its receptor CXCR4 had a direct interaction in thrombosis formation and platelet migration. This result may provide a new treatment strategy for controlling the thrombosis formation in atherosclerotic diseases<sup>10</sup>.

Target of CXCR4 is also used in atherosclerotic plaque analysis and disease diagnosis. Positron emission tomography (PET) is an advanced medical imaging technology that performs human functional metabolism imaging at the cellular and molecular level, playing an important role in the diagnosis of various diseases, such as tumors and CAD. CXCR4 could be a biomarker of damage in plaque endothelial cells and is related to the accumulation of monocytes in plaques. A PET tracer labeled CXCR4 can monitor the status of plaque damage and monocyte recruitment<sup>11</sup>. Combination use of CXCR4 and PET Ga-Pentixafor (a new CXCR4 targeting PET tracer) may be a promising non-invasive imaging method for detecting inflammation in atherosclerosis plaques. In arteries, a highly expression of CXCR4 in inflammatory plaques is often accompanied by a high [68Ga]Pentixafor uptake rate, showing a significantly higher incidence of history of cardiovascular disease<sup>12</sup>. PET/magnetic resonance imaging (MRI) is a large-scale functional metabolic and molecular imaging diagnostic device that combines PET and MRI, achieving complementary advantages in great sense. It is feasible to use [Ga]Pentixa PET/MRI to evaluate the expression of CXCR4 in carotid atherosclerosis in vivo. In atherosclerotic plaque tissue, the expression of CXCR4 can be used as a surrogate marker of inflammatory atherosclerosis<sup>13</sup>. In a retrospective study of patients with non-cardiovascular indications undergoing PET/CT examinations showed that in all patients with 1411 sites, Ga-Pentixa intake was significantly associated with calcified plaque burden and cardiovascular risk factors<sup>14</sup>. Ga-Pentixafor accumulation in arterial walls could assess the prevalence, pattern, and clinical correlation with cardiovascular diseases. It can be seen that it is necessary to define the biological mechanism of CXCR4 and explore the clinical application value of chemokine receptor-directed atherosclerosis imaging<sup>15</sup>.

However, discussion on the exact role of CXCR4 in atherosclerosis and cardiovascular remains controversial. Studies have demonstrated that CXCR4 can inhibit plaque progression in atherosclerosis<sup>16</sup>. In vascular endothelial cells, CXCR4 reduces vascular permeability through the wnt/ $\beta$ -catenin pathway, maintaining vascular endothelial function and arterial integrity; in smooth muscle cells (SMCs), CXCR4 regulates plaque cholesterol efflux and vascular NO dependence, showing anti-atherosclerosis effect in plaque development<sup>16</sup>. The protective effect of CXCR4 is also related to the increase of IgM secreted by B-1 cells, which also plays a positive role in anti-inflammatory and anti-atherosclerosis<sup>17,18</sup>. A risk variant

C-allele at rs2322864 in CXCR4 reduces the normal expression of CXCR4 in human carotid artery endarterectomy specimens<sup>16</sup>, which closely related to the risk of CAD.

C-C Motif Chemokine Receptor 2 (CCR2) is a chemokine which specifically mediates monocyte chemotaxis in monocytes, immature dendritic cells and T cell subpopulations. CCR2 and its ligands, such as CCL2, participate in monocyte infiltration and inflammation process, playing a role in development of atherosclerosis<sup>19</sup>. Increasing the expression of BH4 in vascular endothelial cells of ApoE-KO mice can reduce cell migration mediated by CCR2, impair inflammation response, vascular remodeling, and vein graft atherosclerosis<sup>20</sup>. Researchers also make efforts to explore treatment methods for the CCR2-chemokine axis. For example, CCR2 orthostatic antagonists and allosteric antagonists were developed to block the effect of CCR2 structural site, aiming to intervene the transition of CCR2 to an active state<sup>21,22</sup>. Injection of nanoparticle RNAi to interfere CCR2 expression in ApoE-KO mice could inhibit the recruitment of Ly-6C (high) monocytes and reduce the inflammatory response after myocardial infarction<sup>23</sup>.

CCR5 is expressed in a variety of cells and participating in the proliferation of vascular smooth muscle cells (VSMC), arterial vascular remodeling<sup>24</sup> and inflammation process. It is believed that the development of CCR5 blockers has certain prospective<sup>25</sup>. In clinical research, the relationship between CCR5 and the risk of CAD is controversial. Some studies believe that CCR5 and CCR2 are related to the occurrence of CAD in American women<sup>26</sup>. However, there's contrast conclusions in North Indian CAD patients<sup>27</sup>.

Complement C5a Receptor 1 (C5AR1) combines with complement C5a to recruit monocytes and neutrophils, which are involved in a variety of inflammatory disease processes. In patients with carotid artery plaque, C5a, C5AR1, and the release of IL-1 $\beta$ , IL-18 and IL-1 $\alpha$ <sup>34</sup><sup>28</sup> are significantly increased compared with healthy carotid artery tissue, indicating C5AR1 may be a promoter in progression of atherosclerosis.

DEGs scanning is a promising and meaningful method for disease diagnosis and treatment.

However, some researchers discover that it is the epicardial adipose tissue (EAT) causes inflammatory DEGs expression in blood samples of CAD patients, not the CAD itself<sup>29</sup>. So, there are still controversies and challenges in judging and identifying CAD from the perspective of figuring differential genes. Even so, the role of DEGs cannot be easily denied.

Biomarkers play a significant role in the diseases diagnosis and can be used to judge or monitor treatment response by detecting the deviation of biomarkers. In clinical practice, coronary angiography is the gold standard for the diagnosis of CAD. However, when the patients detected carotid plaques through carotid ultrasound, due to the contrast media allergy or the patients' nervousness about the examination, coronary angiography may not be the first choice for CAD detection. Thus, combination use of detecting blood biomarkers with clinical medical detection methods maybe a quicker and easier method to judge and predict the disease. Likewise, patients with carotid plaques combined with blood biomarkers

scanning maybe promising to predict the risk of CAD, and then the clinicians decide whether to use coronary angiography or coronary CT for the diagnosis of CAD or not.

Our analysis found that the common DEGs between carotid plaques and CAD are mainly related to inflammation and immune process, which concerned with chemokine functions. CCR5, CCR2, CXCR4 and C5AR1 may provide targets for CAD prediction and treatment in patients with carotid plaques.

## Materials And Methods

### Data Acquisition

The GEO (database <http://www.ncbi.nlm.nih.gov/geo>) is an open gene expression database created and maintained by the National Center for Biotechnology Information (NCBI). It is an open high-throughput gene expression data resource that submitted by research institutions all around the world. GSE100927, which contains carotid atherosclerotic tissues and controlled carotid artery tissues gene expression data and GSE56885, which concludes the stable coronary artery disease and healthy controls, were downloaded from the GEO database. The GSE100927 dataset (GPL17077, Agilent-039494 SurePrint G3 Human GE v2 8x60K Microarray 039381 (Probe Name version)) contained 29 carotid atherosclerotic tissues and 12 controlled carotid artery tissues. The GSE56885 dataset (GPL15207 [PrimeView] Affymetrix Human Gene Expression Array) contained 2 healthy controls and 4 CAD patients. The probes were transformed into the corresponding gene symbol according to the annotation information in the platform.

### Data quality control and normalization

NetworkAnalyst3.0 (<https://www.networkanalyst.ca/NetworkAnalyst/home.xhtml>) is a comprehensive network visual analytics platform for gene expression analysis<sup>30</sup>. We used NetworkAnalyst3.0 to draw box plots of the two datasets to observe data distribution and determine whether there were abnormal values in the datasets. Principal components analysis (PCA) is a method that reveal the internal structure of the data and better explain the variables of the data. Meanwhile, PCA is a common method of sample clustering which can detect the data repeatability of the datasets. In this study, PCA analysis was also applied by NetworkAnalyst3.0. Upload gene expression data in "Gene Expression Table", selecting specify organism "human", data type "Microarray data", ID type "official gene symbol", and gene level summarization "mean", then click "submit" to start PCA analysis. According to the data type in the original dataset, select "None" in the normalization option in data filtering & normalization to standardize the data.

### Identification of DEGs and common DEGs

We applied NetworkAnalyst 3.0 online software to screen genes of GSE100927 and GSE56885 datasets separately to obtain the DEGs. Here we chose  $FDR < 0.05$  and  $|\log_2 \text{fold change (FC)}| > 1$  as the criteria to determine the significant DEGs. Shengxinren software (<https://shengxin.ren>) were used to visualize the

DEGs through drawing Volcano maps. Then we used TBtools<sup>31</sup> to visualize common DEGs of the two datasets.

### Functional and pathway enrichment analysis

DAVID Bioinformatics Resources 6.8 (<https://david.ncifcrf.gov/home.jsp>)<sup>32,33</sup> (version provides 6.8), an online analysis tool suite with the function of integrated discovery and annotation is used for GO enrichment and KEGG analysis. In DAVID database operation interface, first submit the differential gene lists. And then in "Select Identifier", we chose "OFFICIAL\_GENE\_SYMBOL" and select "homo sapiens" as background. At last, we submitted the list.  $P < 0.05$  was considered to indicate a statistically significant difference. OmicShare tools (<http://www.omicshare.com/tools>) is an online data visualization platform. We used OmicShare tools to select biological processes (BP), cellular components (CC), molecular functions (MF) and KEGG pathways for data visualization. The metascape.org (<http://metascape.org/gp/index.html#/main/step1>)<sup>34</sup> were used to identify all statistically enriched terms to calculate clusters. The size of the point in each cluster was proportional to the number of genes in the corresponding entry. The nodes of the same color represented the same type of function. The network was visualized through cytoscape v3.1.2

### PPI network construction and module analysis

Protein-protein interaction (PPI) refers to the study of the proteins interaction network, which may provide assistance to mine the core regulatory genes. String (<https://string-db.org/>)<sup>35</sup> is a powerful online software that integrates research on protein interaction research, genome research and proteome research. Select "multiple proteins", upload the gene list, select "homo sapiens" for the organization, and click "search" to construct the PPI network. MCODE algorithm is the most commonly used algorithm for mining protein complexes, which could find closely-connected regions in the PPI network. The metascape.org are used to apply MCODE analysis. At last, GO enrichment and KEGG analysis to each MCODE network (top 3 best p-value clusters)

### Data availability

The datasets during and/or analyzed during the current study available from the corresponding author on reasonable request

## References

- 1 Plichart, M. *et al.* Carotid intima-media thickness in plaque-free site, carotid plaques and coronary heart disease risk prediction in older adults. The Three-City Study. *Atherosclerosis* **219**, 917-924, doi:10.1016/j.atherosclerosis.2011.09.024 (2011).
- 2 Wang, X. *et al.* Carotid Atherosclerosis Detected by Ultrasonography: A National Cross-Sectional Study. *J Am Heart Assoc* **7**, doi:10.1161/JAHA.118.008701 (2018).

- 3 Nambi, V. *et al.* Common carotid artery intima-media thickness is as good as carotid intima-media thickness of all carotid artery segments in improving prediction of coronary heart disease risk in the Atherosclerosis Risk in Communities (ARIC) study. *Eur Heart J* **33**, 183-190, doi:10.1093/eurheartj/ehr192 (2012).
- 4 Franceschini, N. *et al.* GWAS and colocalization analyses implicate carotid intima-media thickness and carotid plaque loci in cardiovascular outcomes. *Nat Commun* **9**, 5141, doi:10.1038/s41467-018-07340-5 (2018).
- 5 Doring, Y., Pawig, L., Weber, C. & Noels, H. The CXCL12/CXCR4 chemokine ligand/receptor axis in cardiovascular disease. *Front Physiol* **5**, 212, doi:10.3389/fphys.2014.00212 (2014).
- 6 Sbrana, S. *et al.* Blood Monocyte Phenotype Fingerprint of Stable Coronary Artery Disease: A Cross-Sectional Substudy of SMARTool Clinical Trial. *Biomed Res Int* **2020**, 8748934, doi:10.1155/2020/8748934 (2020).
- 7 S. Abi-Younes, A. S., F. Mach, G.K. Sukhova, P. Libby, A.D. Luster. The Stromal Cell–Derived Factor-1 Chemokine Is a Potent Platelet Agonist Highly Expressed in Atherosclerotic Plaques. *Circulation Research* **82**, 131-138 (2000).
- 8 Merckelbach, S. *et al.* Expression and Cellular Localization of CXCR4 and CXCL12 in Human Carotid Atherosclerotic Plaques. *Thromb Haemost* **118**, 195-206, doi:10.1160/TH17-04-0271 (2018).
- 9 Kontos, C. *et al.* Designed CXCR4 mimic acts as a soluble chemokine receptor that blocks atherogenic inflammation by agonist-specific targeting. *Nat Commun* **11**, 5981, doi:10.1038/s41467-020-19764-z (2020).
- 10 Witte, A. *et al.* The chemokine CXCL14 mediates platelet function and migration via direct interaction with CXCR4. *Cardiovasc Res*, doi:10.1093/cvr/cvaa080 (2020).
- 11 Baba, O. *et al.* CXCR4-Binding Positron Emission Tomography Tracers Link Monocyte Recruitment and Endothelial Injury in Murine Atherosclerosis. *Arterioscler Thromb Vasc Biol*, ATVBAHA120315053, doi:10.1161/ATVBAHA.120.315053 (2020).
- 12 Li, X. *et al.* [<sup>68</sup>Ga]Pentixafor-PET/MRI for the detection of Chemokine receptor 4 expression in atherosclerotic plaques. *Eur J Nucl Med Mol Imaging* **45**, 558-566, doi:10.1007/s00259-017-3831-0 (2018).
- 13 Li, X. *et al.* [(<sup>68</sup>Ga)]Pentixafor PET/MR imaging of chemokine receptor 4 expression in the human carotid artery. *Eur J Nucl Med Mol Imaging* **46**, 1616-1625, doi:10.1007/s00259-019-04322-7 (2019).
- 14 Weiberg, D. *et al.* Clinical Molecular Imaging of Chemokine Receptor CXCR4 Expression in Atherosclerotic Plaque Using (<sup>68</sup>Ga)-Pentixafor PET: Correlation with Cardiovascular Risk Factors and Calcified Plaque Burden. *J Nucl Med* **59**, 266-272, doi:10.2967/jnumed.117.196485 (2018).

- 15 Kircher, M. *et al.* Imaging Inflammation in Atherosclerosis with CXCR4-Directed (68)Ga-Pentixafor PET/CT: Correlation with (18)F-FDG PET/CT. *J Nucl Med* **61**, 751-756, doi:10.2967/jnumed.119.234484 (2020).
- 16 Doring, Y. *et al.* Vascular CXCR4 Limits Atherosclerosis by Maintaining Arterial Integrity: Evidence From Mouse and Human Studies. *Circulation* **136**, 388-403, doi:10.1161/CIRCULATIONAHA.117.027646 (2017).
- 17 Doring, Y. *et al.* B-Cell-Specific CXCR4 Protects Against Atherosclerosis Development and Increases Plasma IgM Levels. *Circ Res* **126**, 787-788, doi:10.1161/CIRCRESAHA.119.316142 (2020).
- 18 Upadhye, A. *et al.* Diversification and CXCR4-Dependent Establishment of the Bone Marrow B-1a Cell Pool Governs Atheroprotective IgM Production Linked to Human Coronary Atherosclerosis. *Circ Res* **125**, e55-e70, doi:10.1161/CIRCRESAHA.119.315786 (2019).
- 19 Li, X. *et al.* The Regulation of Exosome-Derived miRNA on Heterogeneity of Macrophages in Atherosclerotic Plaques. *Front Immunol* **11**, 2175, doi:10.3389/fimmu.2020.02175 (2020).
- 20 (!!! INVALID CITATION !!! 26).
- 21 Zheng, Y. *et al.* Structure of CC chemokine receptor 2 with orthosteric and allosteric antagonists. *Nature* **540**, 458-461, doi:10.1038/nature20605 (2016).
- 22 Taylor, B. C., Lee, C. T. & Amaro, R. E. Structural basis for ligand modulation of the CCR2 conformational landscape. *Proc Natl Acad Sci U S A* **116**, 8131-8136, doi:10.1073/pnas.1814131116 (2019).
- 23 Maulik D Majmudar, E. J. K., Timo Heidt, Florian Leuschner, Jessica Truelove, Brena F Sena, Rostic Gorbatov, Yoshiko Iwamoto, Partha Dutta, Gregory Wojtkiewicz, Gabriel Courties, Matt Sebas, Anna Borodovsky, Kevin Fitzgerald, Marc W Nolte, Gerhard Dickneite, John W Chen, Daniel G Anderson, Filip K Swirski, Ralph Weissleder, Matthias Nahrendorf. Monocyte-directed RNAi targeting CCR2 improves infarct healing in atherosclerosis-prone mice. *Circulation* **127**, 2038-2046, doi:10.1161/CIRCULATIONAHA (2013).
- 24 Oh, S. *et al.* Pyrogallol-Phloroglucinol-6,6-Bieckolon Attenuates Vascular Smooth Muscle Cell Proliferation and Phenotype Switching in Hyperlipidemia through Modulation of Chemokine Receptor 5. *Mar Drugs* **18**, doi:10.3390/md18080393 (2020).
- 25 Vangelista, L. & Vento, S. The Expanding Therapeutic Perspective of CCR5 Blockade. *Front Immunol* **8**, 1981, doi:10.3389/fimmu.2017.01981 (2017).
- 26 Pai, J. K. *et al.* Polymorphisms in the CC-chemokine receptor-2 (CCR2) and -5 (CCR5) genes and risk of coronary heart disease among US women. *Atherosclerosis* **186**, 132-139, doi:10.1016/j.atherosclerosis.2005.06.041 (2006).

- 27 Sharda, S. *et al.* Chemokine receptor 5 (CCR5) deletion polymorphism in North Indian patients with coronary artery disease. *Int J Cardiol* **124**, 254-258, doi:10.1016/j.ijcard.2006.12.021 (2008).
- 28 Niyonzima, N. *et al.* Cholesterol crystals use complement to increase NLRP3 signaling pathways in coronary and carotid atherosclerosis. *EBioMedicine* **60**, 102985, doi:10.1016/j.ebiom.2020.102985 (2020).
- 29 Fitzgibbons, T. P. *et al.* Coronary disease is not associated with robust alterations in inflammatory gene expression in human epicardial fat. *JCI Insight* **4**, doi:10.1172/jci.insight.124859 (2019).
- 30 Xia, J., Gill, E. E. & Hancock, R. E. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. *Nat Protoc* **10**, 823-844, doi:10.1038/nprot.2015.052 (2015).
- 31 Chen, C. *et al.* TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol Plant* **13**, 1194-1202, doi:10.1016/j.molp.2020.06.009 (2020).
- 32 Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1-13, doi:10.1093/nar/gkn923 (2009).
- 33 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).
- 34 Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523, doi:10.1038/s41467-019-09234-6 (2019).
- 35 Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**, D607-D613, doi:10.1093/nar/gky1131 (2019).

## Declarations

### Acknowledgments

The authors of this article are grateful for the datasets provided by GEO database and data providers.

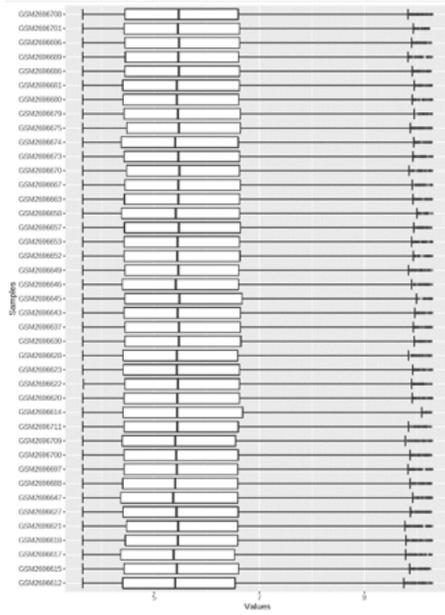
### Author Contributions

Luxia Song designed the research. Luxia Song, Jie Zhang and Yixuan Fan performed data analysis. Qiyu Liu and Baoyi Guan search references about this article. Luxia Song and Hao Xu wrote the manuscript. All authors read and approved the final manuscript.

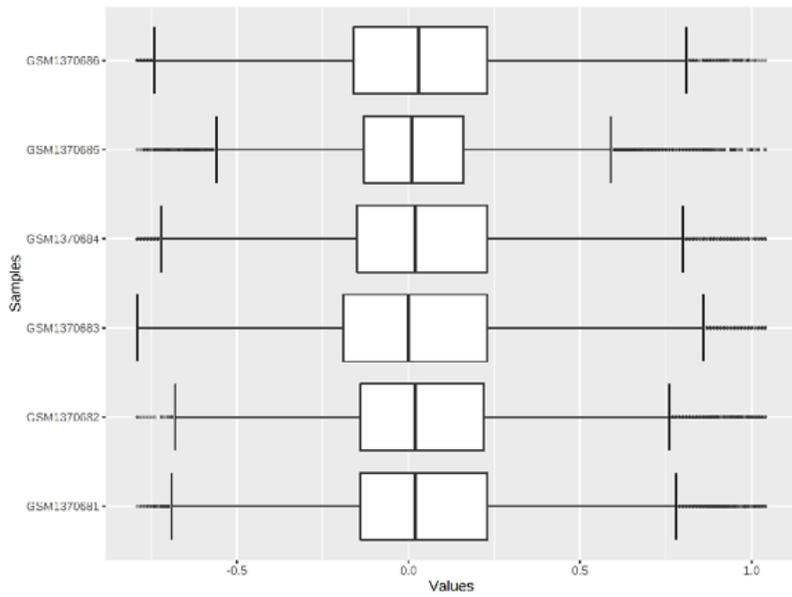
### Competing interests

The authors have no conflict of interests to this work

# Figures

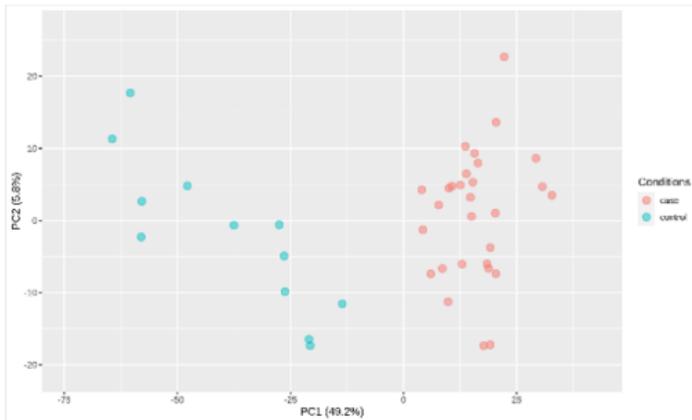


**a**

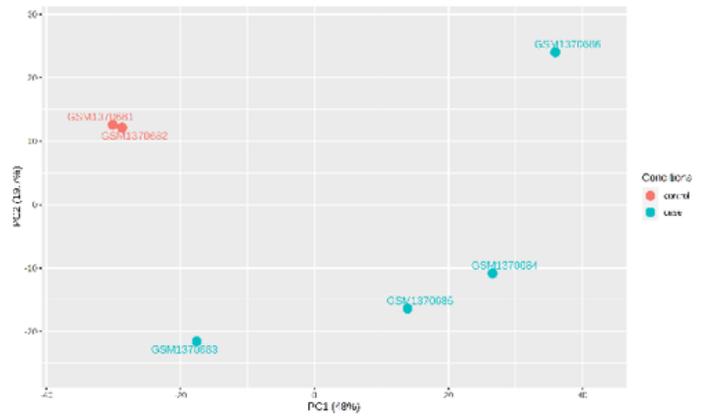


**b**

## PCA



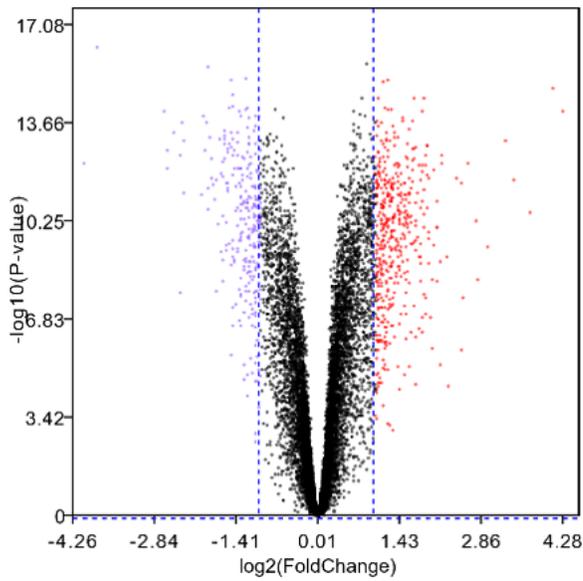
**c**



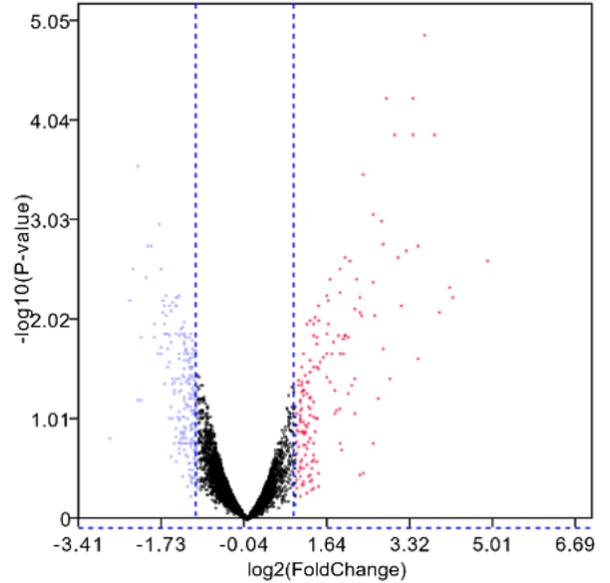
**d**

**Figure 1**

a Boxplot of ges100927 dataset after normalization, each boxplot represents a sample. b Boxplot of ges56885 dataset after normalization. c PCA in ges100927 dataset after normalization, each point represents a sample and pink points refer to case group (carotid atherosclerotic tissues), and green points refer to control group (controlled carotid artery tissues). PCA diagram could show data distribution in different groups, displaying whether there are differences between the two groups of data. d PCA in ges56885 dataset after normalization. Pink points refer to control group (healthy control), and green points refer to case group (CAD patients).



**a**



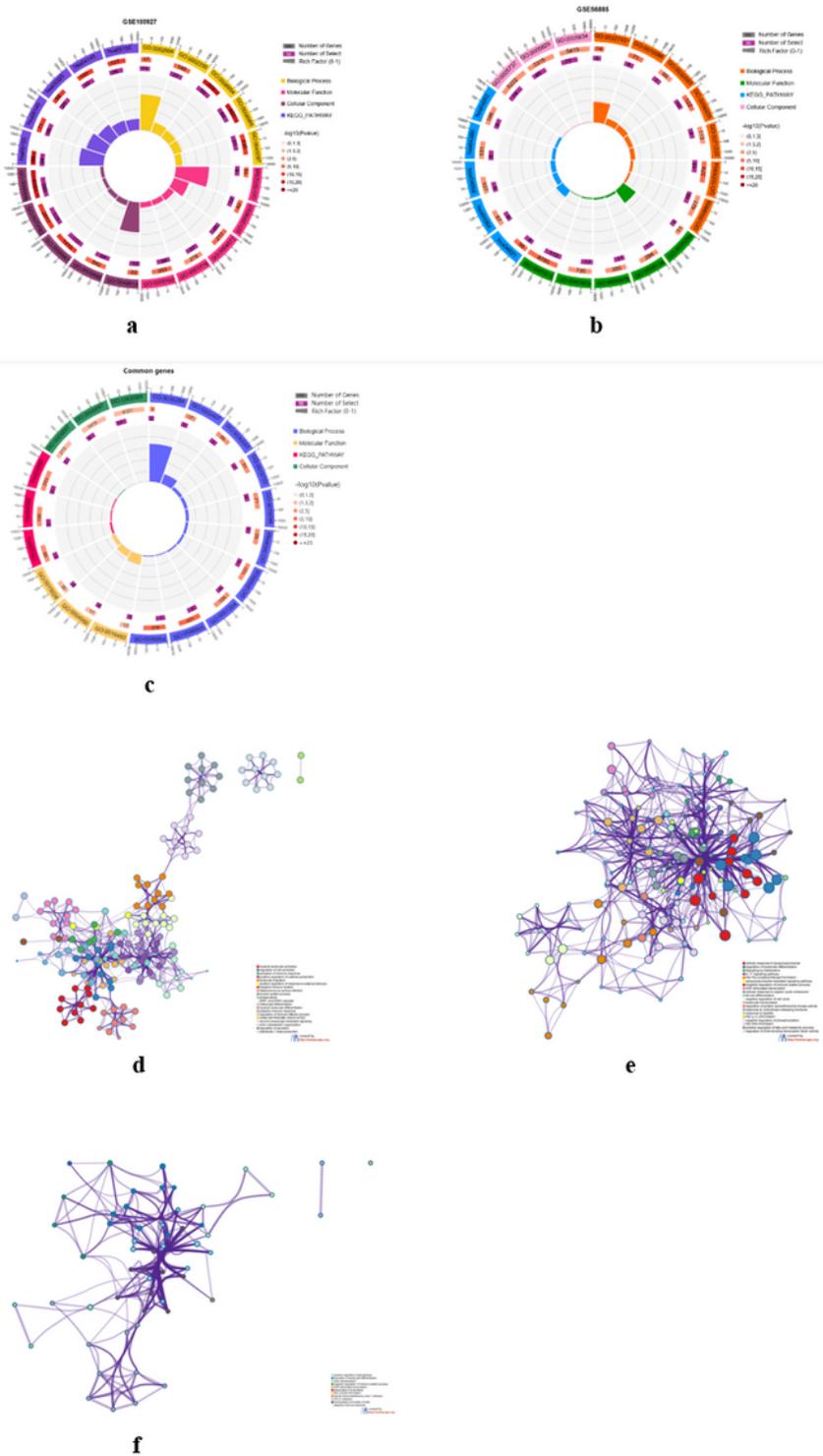
**b**



**c**

## Figure 2

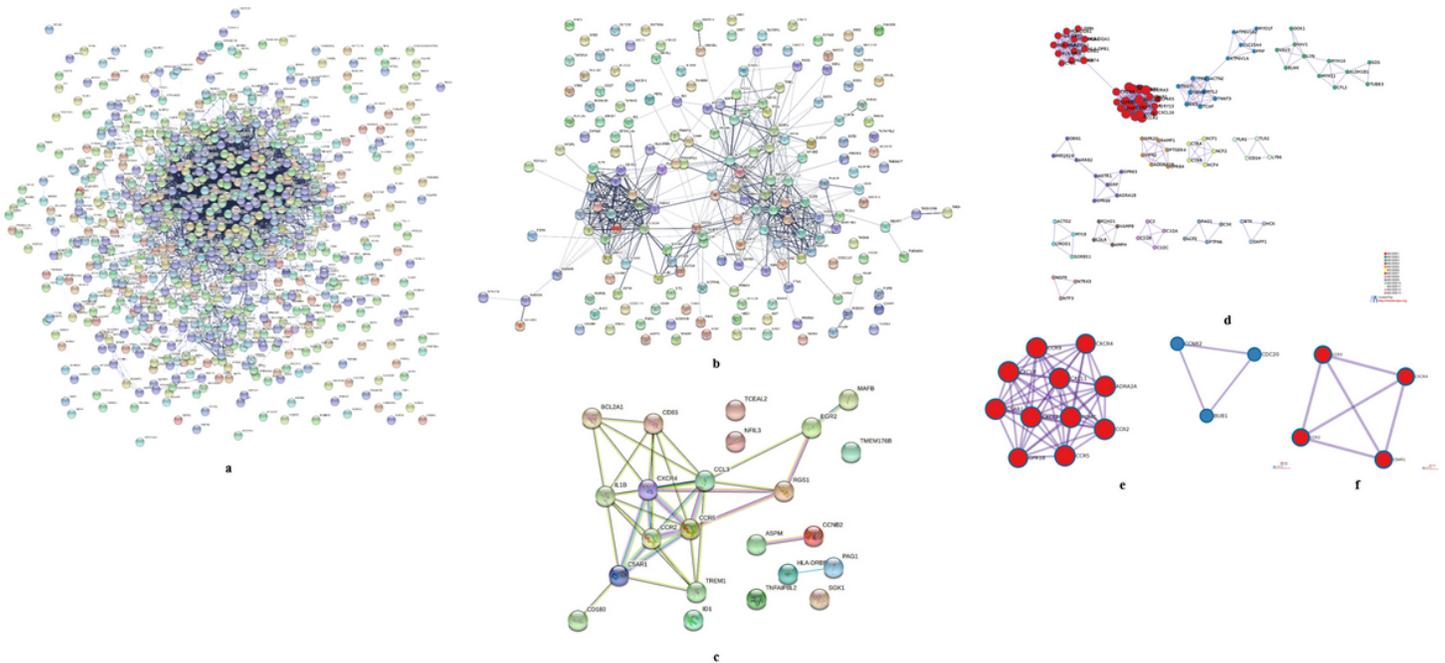
a Volcano plot of ges100927. Each dot in the figure represents the differential gene screened by the current chip. Red and purple dots refer to significantly different up-regulated and down-regulated genes. b Volcano plot of ges56885. c Venn gram of the two datasets. Venn gram shows DEGs in the two datasets separately and their common DEGs.



**Figure 3**

a GO enrichment and KEGG pathway analysis of GSE100927. BP, MF, CC, and KEGG pathway analysis were shown by the four parts of the outer circle in the circle diagram. b GO enrichment and KEGG pathway analysis of GSE56885. c GO enrichment and KEGG pathway analysis of the common DEGs. d Network visualization of GO enrichment and KEGG pathway analysis of GSE100927. In this network, each circle node represented a term, and nodes of the same color refer to the same cluster. When

similarity score > 0.3, the nodes would be linked by an edge. The thickness of the line is proportional to the similarity between the two terms. e Network visualization of GO enrichment and KEGG pathway analysis of GSE56885. f Network visualization of GO enrichment and KEGG pathway analysis of common DEGs.



**Figure 4**

a PPI network of GSE100927. PPI network shows interaction between proteins. The stronger the interaction between protein nodes, the thicker the connection lines are. b PPI network of GSE56885. c PPI network of common DEGs. d MCODE of GSE100927. Further analysis of potential subnets in the PPI network. e MCODE of GSE56885. f MCODE of common DEGs.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementtables.docx](#)