Identification of key pathways and core genes involved in atherosclerotic plaque progression

Rong Yang¹, Linpeng Yao¹, Chengli Du², Yihe Wu²

¹Department of Radiology, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China; ²Department of Thoracic Surgery, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

Contributions: (I) Conception and design: Y Wu; (II) Administrative support: R Yang, Y Wu; (III) Provision of study materials or patients: R Yang, L Yao, C Du; (IV) Collection and assembly of data: R Yang, L Yao, C Du; (V) Data analysis and interpretation: R Yang, Y Wu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Yihe Wu, PhD, MD. Department of Thoracic Surgery, the First Affiliated Hospital, Zhejiang University School of Medicine, 79 Qingchun Road, Hangzhou 310003, China. Email: drwuyihe@zju.edu.cn.

Background: Atherosclerosis leads to the occurrence of cardiovascular diseases. However, the molecular mechanisms that contribute to atherosclerotic plaque rupture are incompletely characterized. We aimed to identify the genes related to atherosclerotic plaque progression that could serve as novel biomarkers and interventional targets for plaque progression.

Methods: The datasets of GSE28829 in early vs. advanced atherosclerotic plaques and those of GSE41571 in stable vs. ruptured plaques from Gene Expression Omnibus (GEO) were analyzed by using bioinformatics methods. In addition, we used quantitative reverse transcription polymerase chain reaction (qRT-PCR) to verify the expression level of core genes in a mouse atherosclerosis model.

Results: There were 29 common differentially expressed genes (DEGs) between the GSE28829 and GSE41571 datasets, and the DEGs were mainly enriched in the chemokine signaling pathway and the Staphylococcus aureus infection pathway (P<0.05). We identified 6 core genes (FPR3, CCL18, MS4A4A, CXCR4, CXCL2, and C1QB) in the protein-protein interaction (PPI) network, 3 of which (CXCR4, CXCL2, and CCL18) were markedly enriched in the chemokine signaling pathway. qRT-PCR analysis showed that the messenger RNA levels of two core genes (CXCR4 and CXCL2) increased significantly during plaque progression in the mouse atherosclerosis model.

Conclusions: In summary, bioinformatics techniques proved useful for the screening and identification of novel biomarkers of disease. A total of 29 DEGs and 6 core genes were linked to atherosclerotic plaque progression, in particular the CXCR4 and CXCL2 genes.

Keywords: Atherosclerosis; molecular mechanism; bioinformatics; core genes; differentially expressed genes (DEGs)

doi: 10.21037/atm-21-193

View this article at: http://dx.doi.org/10.21037/atm-21-193

Introduction

Atherosclerosis, defined as the formation of fibrofatty lesions in the artery wall, causes considerable morbidity and mortality worldwide (1). Typically, atherosclerotic plaque have superficial fibrous cap composed of smooth muscle cells and relatively dense collagen. Beneath and side of cap shows more cellular area containing macrophages, T cells and smooth muscle cells. Deep to the fibrous cap is a necrotic core containing lipid debris from dead cells, foam cells fibrin, variably organized thrombus and other plasma proteins (1). Atherosclerotic plaque formation, progression, and rupture lead to later symptomatic disease such as myocardial infarction, stroke, disabling peripheral artery disease, and eventually death (2). Atherosclerosis seriously
reduces people’s quality of life and life expectancy, and greatly increases health costs (3). Therefore, it is urgent to study the pathogenesis of atherosclerosis in order to find suitable targets for the prevention and treatment of this condition.

Currently, there is a lack of understanding regarding the molecular pathways of atherosclerotic plaque rupture (4). However, high-throughput gene profiling has become an effective tool for revealing the pathogenesis of cardiovascular diseases. Using this method, researchers have already discovered several genes and pathways that are closely related to coronary artery disease and atherosclerosis (5,6). Most of these studies have focused on comparing the differential gene expression between atherosclerotic plaques and normal tissues (7-11). Fewer studies exist, however, which compare the differential gene expression between early and advanced plaques, or between stable and ruptured plaques (4,12,13). In this study, we identified common differentially expressed genes (DEGs) in early and advanced plaques (GSE28829), and in stable and ruptured plaques (GSE41571), and used bioinformatics methods to determine the key pathways and core genes involved in atherosclerotic plaque progression. It is hoped our findings will provide new potential targets for the prevention, diagnosis, and treatment of atherosclerosis. We present the following article in accordance with the MDAR reporting checklist (available at http://dx.doi.org/10.21037/atm-21-193).

Methods

Microarray data information

The microarray data of the GSE28829 dataset for early and advanced plaques and the GSE41571 dataset for stable and ruptured atherosclerotic plaques, were downloaded from the National Center for Biotechnology Information Gene Expression Omnibus (NCBI-GEO) public database. The GSE28829 and GSE41571 datasets were both on the GPL570 platforms (HG-U133_Plus_2) which included 13 early plaque cases and 16 advanced plaque cases, and 6 stable plaque cases, and 5 ruptured plaque cases, respectively. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Data processing of DEGs

DEGs were identified via the GEO2R interactive web tool with |log FC| >1.5 and adjusted P value <0.05 (14). We then used Venn diagram software to check the raw data in TXT format in order to identify the common DEGs between the two sets of microarray data (14). The DEGs with log FC <0 and log FC >0 were considered downregulated and upregulated genes, respectively.

Gene ontology (GO) and pathway analysis

GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were used to study the biological characteristics of these DEGs (15). GO and KEGG pathway analyses were performed by DAVID online software (https://david.ncifcrf.gov/), which can visualize the DEG enrichment of molecular function (MF), biological process (BP), cellular component (CC), and KEGG pathways (15). A P value <0.05 in the results was considered statistically significant (15).

Protein-protein interaction (PPI) network and core gene identification

A PPI helped to clarify the regulatory functions among proteins, which were evaluated using the Search Tool for the Retrieval of Interacting Genes (STRING) (14,16). We uploaded the common DEGs obtained from the GSE28829 and GSE41571 datasets onto STRING’s website to construct a PPI network, and then used Cytoscape software (version 3.2.1) to obtain a visualized interactive network. Finally, we used Molecular Complex Detection (MCODE) application in Cytoscape to screen the core genes (16).

Animal model

Male wild-type (WT) and atherosclerosis-prone apolipoprotein E-deficient ApoE−/− mice (6 weeks old, 20±2 g) with a C57BL/6J genetic background were obtained from Xipuer-Bikai Experimental Animal Co., Ltd. (Shanghai, China). The WT mice were given a normal diet and used as controls, and the ApoE−/− mice were given a Western diet (high-fat diet of 21% butter fat and 0.15% cholesterol) for 20 weeks and used as atherosclerosis models. All mice were fed in the Zhejiang University Animal Center. After the mice were euthanized, the aortas were excised and processed for histology. All animal experiments were approved by the Ethics Committee of the First Affiliated Hospital, Zhejiang University School of Medicine (Hangzhou, China). All animal experiments
were performed in accordance with Zhejiang university guidelines for the care and use of animals.

**Verification by quantitative reverse transcription quantitative polymerase chain reaction (qRT-PCR)**

The extent of atherosclerosis was assessed in mouse aortas by staining for lipid depositions with Oil Red O (12). The mouse aorta samples were divided into three groups: normal aortas from controls, atherosclerotic artery segments with early lesions (pathological intimal thickening and intimal xanthoma) from ApoE<sup>−/−</sup> mice, and atherosclerotic artery segments with advanced lesions (thin or thick fibrous cap atheroma) from ApoE<sup>−/−</sup> mice (12). Tri<sup>®</sup>-Reagent (Sigma-Aldrich, St. Louis, MO, USA) was used to extract total RNA from the aortic samples and PrimeScript RT reagent Kit (T<sup>a</sup>KaRa Biotechnology, Dalian, China) was used to reverse transcribe total RNA into complementary DNA (cDNA). The polymerase chain reaction (PCR) reaction was performed using SYBR Green (Thermo Fisher Scientific, USA) and specific primer pairs (sequences in Table 1). Amplification was performed in duplicate using a 7900 HT Fast Real-Time PCR system (Applied Biosystems, Thermo Fisher Scientific, USA). The experiment was repeated three times for each gene, and GAPDH was used as the endogenous reference gene.

### Statistical analysis

The 2<sup>−ΔΔCT</sup> method was used to determine the expression level of all messenger RNAs (mRNAs) (17). One-way analysis of variance was used for comparisons among three groups, and a P value <0.05 was considered statistically significant. All statistical analyses were conducted using SPSS 25.0 software (IBM Corp., Armonk, NY, USA).

### Results

#### Identification of DEGs

There were 96 DEGs in advanced atherosclerotic plaque compared with early plaque, including 92 upregulated genes and 4 downregulated genes. In addition, there were 838 DEGs in ruptured atherosclerotic plaque compared with stable plaque, including 173 upregulated genes and 665 downregulated genes. Finally, there were 29 DEGs overlapped between GSE28829 and GSE41571 in the Venn diagram, which were all upregulated genes (log FC >0) (Table 2 and Figure 1).

#### DEGs GO and KEGG pathway analyses

GO and KEGG pathway analyses of 29 DEGs were performed by DAVID software. GO analysis showed the following: (I) for BP, DEGs were particularly enriched in regulation of immune response, complement activation, receptor-mediated endocytosis, innate immune response, phagocytosis, and proteolysis; (II) for CC, DEGs were significantly enriched in the plasma membrane, extracellular exosome, extracellular space, extracellular region, and blood microparticle; and (III) for MF, DEGs were enriched in

---

### Table 1 Primer sequences for qRT-PCR

<table>
<thead>
<tr>
<th>Targets</th>
<th>Forward primer (5'-3')</th>
<th>Reverse primer (3'-5')</th>
</tr>
</thead>
<tbody>
<tr>
<td>CXCR4</td>
<td>GAGGCCAAGGAAACTGCTG</td>
<td>GCCGTCACAGATGTACCTGT</td>
</tr>
<tr>
<td>CXCL2</td>
<td>CTCAGACAGCGAGGCACATC</td>
<td>CCTCAACGGAAAGACAAAGAG</td>
</tr>
<tr>
<td>GAPDH</td>
<td>CACCATCTCCAGGAGCAG</td>
<td>CTTTCTCCATGTGGAAGAGAC</td>
</tr>
</tbody>
</table>

qRT-PCR, quantitative reverse transcription quantitative polymerase chain reaction.

### Table 2 A total of 29 commonly differentially expressed genes (DEGs) identified from the GSE28829 and GSE41571 datasets

<table>
<thead>
<tr>
<th>DEGs</th>
<th>Gene symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upregulated</td>
<td>IGKV10R2-108, CXCR4, CD37, CD36, IGHM, IGLC1, CD52, CYAT1, AQP9, CCL18, IGLJ3, FPR3, IGHV3-23, NPL, IGLL3P, RAC2, 1Q0B, IGKC, MS4A4A, HLA-DQB1, JCHAIN, MIR8071-2, CXCL2, IGK, SLC16A10, IGLV1-44, MLIP, IGH2, CTSC</td>
</tr>
<tr>
<td>Downregulated</td>
<td>-</td>
</tr>
</tbody>
</table>
antigen binding, immunoglobulin receptor binding, serine-type endopeptidase activity, peptidoglycan binding, and phosphatidylcholine binding (Table 3).

The results of KEGG analysis showed that 29 DEGs were particularly enriched in the chemokine signaling pathway and the staphylococcus aureus infection pathway (P<0.05) (Table 3). Four genes (RAC2, CXCR4, CXCL2, and CCL18) were markedly enriched in the chemokine signaling pathway (Table 4 and Figure 2).

**PPI network and modular analysis**

After the pathway analyses, 18 of the 29 DEGs were imported into the PPI network complex, and the other 11 DEGs were not contained in the PPI network complex (Figure 3A). We then used Cytoscape MCODE to screen the core genes, and identified 6 hub genes (FPR3, CCL18, MS4A4A, CXCR4, CXCL2, and C1QB) among 18 nodes (Figure 3B). Three of the six core genes (CXCR4, CXCL2,
Table 4 KEGG pathway analysis results of differentially expressed genes during the progression of atherosclerotic plaques

<table>
<thead>
<tr>
<th>Pathway ID</th>
<th>Name</th>
<th>Count</th>
<th>%</th>
<th>P value</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa04062</td>
<td>Chemokine signaling pathway</td>
<td>4</td>
<td>19.2</td>
<td>0.003574</td>
<td>RAC2, CXCR4, CXCL2, CCL18</td>
</tr>
<tr>
<td>hsa05150</td>
<td>Staphylococcus aureus infection</td>
<td>3</td>
<td>14.4</td>
<td>0.003796</td>
<td>HLA-DQB1, C1QB, FPR3</td>
</tr>
</tbody>
</table>

Figure 2 Four genes (RAC2, CXCR4, CXCL2, and CCL18) were significantly enriched in the chemokine signaling pathway. RAC means RAC2; Chemokine means CXCL2 and CCL18; Chemokine R means CXCR4.

and CCL18) were markedly enriched in the chemokine signaling pathway (Table 4).

**Corroboration of two core genes using qRT-PCR**

Mouse aortic samples were divided into three groups: normal aortas, early atherosclerotic plaque, and advanced atherosclerotic plaque. According to the PPI network analysis, the three core genes markedly enriched in the chemokine signaling pathway were CXCR4, CXCL2, and CCL18. qRT-PCR was performed in mouse aortic samples to confirm the expression levels of the three core genes. The results showed that the expression of two core genes (CXCR4 and CXCL2) increased significantly during the progression of atherosclerotic plaques (Figure 4). Because no homologs of CCL18 have been found in rodents (18),...
Figure 3 The PPI network of common DEGs was constructed by the STRING online database and module analysis. (A) A total of 18 DEGs were included in the DEG PPI network; (B) core genes in the PPI network. PPI, protein-protein interaction; DEGs, differentially expressed genes.

Figure 4 Corroboration of two core genes using qRT-PCR. (A) Oil Red O-stained mice aortas; (B) Relative expression of the CXCR4 measured by qRT-PCR; (C) Relative expression of the CXCL2 measured by qRT-PCR. qRT-PCR, quantitative reverse transcription quantitative polymerase chain reaction.

CCL18 was not verified by qRT-PCR.

Discussion

Although great progress has been made in the prevention, diagnosis, and treatment of cardiovascular disease, it is still the leading cause of death worldwide (19). The main cause of cardiovascular disease is atherosclerosis, which is itself a chronic inflammatory disease (19). The molecular mechanisms of atherosclerotic plaque formation and development are very complex. Although some studies have found that cytokines participate in all stages of atherosclerosis, the detailed pathophysiology of the atherosclerosis is still unclear (19,20). In recent decades, the rapid development of high-throughput detection technology and the progress of bioinformatics technology have brought new methods into the study of pathophysiological mechanism of atherosclerosis.

In this study, the GSE28829 dataset that was related to early vs. advanced atherosclerotic plaques and the GSE41571 that was dataset related to stable vs. ruptured plaques from GEO were analyzed by using bioinformatics methods. We found that there were 29 common DEGs between GSE28829 and GSE41571 and that the DEGs were mainly enriched in the chemokine signaling pathway and the Staphylococcus aureus infection pathway (P<0.05). Our research suggests that the two pathways play important roles in the progression of atherosclerotic plaque and
were markedly enriched in the CXCR4. Recently reported that CXCR4-
Ann Transl Med CXCL18 were the top six core genes in the progression of atherosclerotic plaques. Three of the six core genes (C1QB, CCL2, CCL4, CCL5, C-X-C motif chemokine ligand 10 (CXCL10), CXCL1, and CXCL16, have been identified to play an important role during the progression of atherosclerotic plaque (26-30). Inflammation within the atherosclerotic plaque is mediated by their recruitment of target inflammatory cells (25). CXCL1 and CXCL2, the common ligands of CXCR2, are the major leukocyte chemoattractants and activators (31,32). CXCL1 is a chemoattractant of neutrophils, T lymphocytes, and monocytes and is expressed in human atherosclerotic plaques (33,34). It can induce the production of free radicals and lead to the injury of vascular endothelial cells (34). CXCL2, secreted by activated monocytes, has the same activity as CXCL1 (25). It attracts neutrophils during the inflammatory process, but its role in atherosclerosis is not clear (35-36). High CXCL2 expression has been reported in patients with rheumatoid arthritis, in whom neutrophils predominate in the joints (37). Rouault et al. reported that CXCL2 contributes to the adhesion of neutrophils to endothelial cells in vitro which can then the promote inflammation and senescence of endothelial cells in turn (38). Our results showed that CXCL2 is upregulated at the mRNA level during the progression of atherosclerotic plaques, which has not been reported in previous studies. Further study of CXCL2 will help us to understand the pathophysiological mechanism of atherosclerosis and facilitate its utilization as a biomarker and intervention target for atherosclerosis.

C-X-C motif chemokine ligand 18 (CCL18) is also a chemokine, secreted by type 2 macrophages (39). It is a chemoattractant factor for T cells and serves as a marker of anti-inflammatory activation (18). The accumulation of macrophages in the arterial intima is an important factor in promoting the progression of atherosclerosis. CCL18 may participate in human atherosclerotic plaque formation (40). CCL18 was found to be positively expressed in human atherosclerotic plaques, and its mRNA was limited to the macrophage-rich region (41). Chen et al. reported that the expression levels of CCL18 in ruptured plaques were significantly higher than that in asymptomatic stable plaques (18).

Formyl peptide receptors (FPRs) are G protein–coupled receptors and are mainly expressed in white blood cells (42). Details regarding FPR signaling are scarce (43). Three FPRs have been identified in humans, FPR1–FPR3, and putative counterparts in mice (43). FPR1 and FPR2 are positively expressed on the surfaces of leukocytes, and are closely related to the control of inflammation induction (44). However, the overall function of FPR3 remains unclear, and whether or not it is related to inflammation has not been reported (42).
The membrane spanning 4-domains A (MS4A) family of proteins includes well-known members, MS4A1-MS4A13 (45). These proteins are expressed in certain leukocyte subsets and play a key role in regulating cell activation, growth, and development (45). MS4A4A is also a member of the MS4A family and is selectively expressed by macrophage-lineage cells (46). However, our understanding of its biological roles is rather limited. MS4A4A has not been previously studied in atherosclerosis. Mattiola et al. reported that MS4A4A was induced during the differentiation of monocytes into macrophages (46), while genome-wide association studies have correlated MS4A4A with Alzheimer’s disease and cutaneous systemic sclerosis (47,48).

Finally, complement C1q B chain (C1QB) is part of the complement system, and plays an important role in the regulation of immune response (49). The activation of the complement system can lead to endothelial dysfunction, which is considered to be an important mechanism for the formation and development of atherosclerosis (50). Tian et al. reported that C1QB in particular is associated with inflammatory response (51), and Bos et al. have reported that it is significantly associated with the stage of coronary artery disease (49).

The bioinformatics analysis in our study showed that the high expression of FPR3, CCL18, MS4A4A, CXCR4, CXCL2, and C1QB had appreciable value in distinguishing plaques and even identifying advanced and ruptured stages. However, this study had several limitations. First, due to the heterogeneity of atherosclerosis and the relatively small sample size of this study, the interpretation of the findings needs should be conducted with caution. Second, this study was only a preliminary screening of core genes during the progression of atherosclerosis, and the specific pathophysiological mechanism of these core genes in the progression of atherosclerosis requires further study. In the future, we must further verify these results in a larger number of samples, and clarify the role of these core genes in atherosclerosis through in vitro and in vivo studies.

Conclusions

This study preliminarily screened and identified the key pathways and genes closely related to the progression of atherosclerotic plaques. We speculated that these core genes (FPR3, CCL18, MS4A4A, CXCR4, CXCL2, and C1QB) may play an important role during the progression of atherosclerotic plaques, influencing atherosclerosis progression through chemokine signaling pathway and Staphylococcus aureus infection pathway. These findings may provide needed insights into the pathophysiological mechanism of atherosclerosis, and offer new potential targets for the prevention, diagnosis, and treatment of atherosclerosis in the future.

Acknowledgments

Funding: This research was funded by the Natural Science Foundation of Zhejiang Province, China (No. LQ18H180002) and the National Natural Science Foundation of China (No. 31700690).

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at http://dx.doi.org/10.21037/atm-21-193

Data Sharing Statement: Available at http://dx.doi.org/10.21037/atm-21-193

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-21-193). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was approved by the Ethics Committee of the First Affiliated Hospital, Zhejiang University School of Medicine (Hangzhou, China). All animal experiments were performed in accordance with Zhejiang university guidelines for the care and use of animals. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license).
References


© Annals of Translational Medicine. All rights reserved.
