scientific reports

Check for updates

OPEN

Single-cell transcriptome integrated with genome-wide association study reveals heterogeneity of carotid and femoral plaques and its association with plaque stability

Xinhuang Hou[1,2,6]✉, Zhipeng Li[3,6], Jun Lin[4,6], Wei Lin[1], Luyao Li[1], Xiaoqi Zheng[1], Xiaoling Lai[1], Lin Zhu[5], Pingfan Guo[1,2], Fanggang Cai[1,2], Jinchi Zhang[1,2], Wanglong Li[1,2], Changwei Yang[5]✉ & Yiquan Dai[1,2]✉

Carotid and femoral plaques exhibit varying degrees of stability; however, the relationships of different genes/cell types with plaque embolism are poorly understood. We evaluated differential gene/cell expression and investigated the cells/genes associated with carotid and femoral artery plaque embolism. sc-RNA-seq and bulk RNA data were obtained to identify differentially expressed genes (DEGs). Seven machine learning models were trained, and the top 10 DEGs across all models were selected. The most disturbed cells in carotid and femoral artery plaques were identified using Augur, while the genes and cells in the carotid plaque associated with embolism were analyzed through scPagwas. The differences in most disturbed cells and embolism-related cells were further analyzed. Compared with femoral plaques, carotid plaques had 80 downregulated and 90 upregulated genes. Machine learning identified the key DEGs between carotid and femoral plaques were predominantly from the HOX gene family. Natural Killer (NK) cells were the most significantly disturbed cells between carotid and femoral plaques, and they may be most strongly associated with plaque embolism. Among the differential genes in NK cells, *CD2* was most associated with embolism. Our research may offer new insights into atherosclerosis at different locations.

Keywords Single-cell sequencing, Carotid and femoral, Atherosclerosis, Plaque embolism

Atherosclerotic vascular disease represents a major global health challenge and is one of the leading causes of mortality worldwide[1]. Atherosclerosis is a chronic inflammatory disease characterized primarily by apolipoprotein B deposition in the arterial intima and inflammatory cell infiltration[2–4]. In recent years, the number of patients with peripheral atherosclerotic artery disease has been gradually increasing globally; therefore, it is becoming an increasingly serious issue[5–8]. However, previous research on the inflammatory mechanisms of atherosclerosis has primarily focused on cardiovascular and cerebrovascular diseases, with relatively less research focused on peripheral artery atherosclerosis[9–11].

Carotid artery and femoral artery plaques exhibit varying degrees of stability. Carotid plaques are more prone to rupture and form distal emboli, while femoral plaques tend to progress (tend to be more stable),

[1]Department of Vascular Surgery, The First Affiliated Hospital, Fujian Medical University, 20 Chazhong Road, Fuzhou 350005, China. [2]Department of Vascular Surgery, National Regional Medical Center, Binhai Campus of The First Affiliated Hospital, Fujian Medical University, Fuzhou 350212, China. [3]Department of Gastrointestinal Surgery, The First Affiliated Hospital, Fujian Medical University, Fuzhou 350005, China. [4]Trauma Center and Emergency Surgery Department, The First Affiliated Hospital, Fujian Medical University, Fuzhou 350005, China. [5]Department of Nutrition and Food Safety, School of Public Health, Fujian Medical University, Fuzhou 350122, China. [6]Xinhuang Hou, Zhipeng Li and Jun Lin contributed equally to this work. ✉email: fjfyhxh@163.com; vivianyang87@126.com; daiyiquan@sina.cn

ultimately leading to peripheral artery disease[12–14]. Furthermore, carotid and femoral plaques also differ in their inflammatory characteristics, with carotid plaques appearing to exhibit a more active inflammatory response[15–18]. Recent studies have used single-cell sequencing to compare the heterogeneity of plaque cells in the femoral artery and carotid artery[19–22] However, current research has only examined the differences in inflammation levels between carotid and femoral artery plaques, without conducting a joint analysis with plaque embolism events. Little is known about the relationship between different cell types, genes, and plaque embolism in carotid and femoral artery. The scPagwas method can be used to integrate genome-wide association study (GWAS) and sc-RNA seq to identify cell subpopulations and genes associated with specific traits, allowing the impact of cells in certain diseases to be evaluated.

To date, the relationship between the differential gene expression, inflammatory cell profiles, and risk of plaque embolism in carotid and femoral artery plaques has not been sufficiently explored. The purpose of this study is to use various machine learning approaches to screen for the key differentially expressed genes (DEGs) in carotid and femoral artery plaques, while also investigating the most disturbed cell type in these plaques. Furthermore, scPagwas was utilized to investigate the cell types and genes associated with plaque embolism, which were compared between carotid and femoral artery plaques to elucidate the relationships between gene/cell differences and the distinct vulnerability of these types of plaque.

Methods

Data availability

The single-cell sequencing data and bulk RNA data used in this study were obtained from the Gene Expression Omnibus (GEO) database. The single-cell data were from the GSE234077 dataset, which includes nine femoral artery plaque samples and four carotid artery plaque samples after $CD45^+$ selection[19]. The bulk RNA data were obtained from four datasets: GSE100927, GSE53274, GSE43292 and GSE23304. The GSE100927 dataset includes carotid and femoral artery plaque samples, as well as their respective arterial controls[12]. We used 29 carotid artery plaque samples and 26 femoral artery plaque samples from this dataset. The GSE53274 dataset includes primary femoral plaque samples and restenosis plaque samples[23]. We selected four primary femoral plaque samples for the analysis. The GSE43292 dataset includes carotid artery plaque samples and samples from neighboring arteries[24]. We selected 32 carotid artery plaque samples for the analysis. The GSE23304 dataset includes a comparison of 101 arterial plaques, from which this study selected 7 carotid plaques, 7 common femoral artery plaques, and 61 superficial femoral artery plaques (Table S1). The overall workflow of this study is shown in Fig. S1.

This study utilized publicly available datasets with pre-existing ethics approval. The ethics approval for the immunohistochemical (IHC) validation was granted by the Ethics Committee of the First Affiliated Hospital of Fujian Medical University (approval number: FMU [2023] 508). All participants had provided informed consent, and the research was conducted in accordance with the principles of the Declaration of Helsinki.

Analysis of DEGs and pathways between carotid and femoral artery plaques

We used the "sva" package of R to reduce batch effects from the GSE100927, GSE53274, and GSE43292 datasets. The initial differential gene expression analysis between carotid and femoral artery plaques was conducted at a cutoff FC = 1.

For inflammatory gene set enrichment, the "HALLMARK_INFLAMMATORY_RESPONSE" gene set from the Molecular Signatures Database (MSigDB) was used in a Gene Set Enrichment Analysis (GSEA) to assess differences in inflammatory pathways. The DEGs were further analyzed by Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment[25–27].

To identify key DEGs in carotid and femoral artery plaques, the bulk RNA dataset was randomly split into the training (70%) and validation (30%) sets. Seven machine learning models, including Least Absolute Shrinkage and Selection Operator, Gradient Boosting Machine, XGBoost, Support Vector Machine, Linear Discriminant Analysis, Elastic Net Regularization, and Logistic Regression, as well as their combinations, were trained and tuned on the training data to achieve the highest area under the curve for the validation set. We then evaluated the area under the curve of the models to assess their goodness of fit, retaining only those with values greater than 0.8 for subsequent analysis. We calculated the importance of each gene in each model. Based on the frequency with which a gene appeared across all models and its average importance rank within each model, we selected the top 10 genes as the DEGs.

Single-cell RNA-seq analysis

scRNA-seq analysis was performed using Seurat v4.2.1. All of the femoral artery plaque and carotid artery plaque data were combined into a single Seurat object, before extracting the carotid artery plaque data as separate Seurat objects for use in the scPagwas analysis. For initial quality control, cells with 1) > 3 expressed cells, 2) < 200 total genes, 3) mitochondrial genes accounting for > 10% of the total gene expression, 4) < 200 or > 10,000 detected RNA features, and 5) total molecule counts of < 1,000 were filtered. We used "SCTransform" to normalize and select highly variable genes. Then, for better interpretability, we employed Non-negative Matrix Factorization for dimensionality reduction instead of Principal Component Analysis. After performing Harmony batch correction, we used the "FindNeighbors" and "FindClusters" functions to identify clusters, setting the resolution to 0.6 for initial clustering. Uniform Manifold Approximation and Projection (UMAP) was used to visualize the cell clusters. In this study, we classified the initial cell clusters annotation using the following markers: T cells (CD3E, CD3D), monocytes/macrophages (CD14, CD68), B cells (CD79A, MS4A1), NK cells (NKG7, TYROBP, FCGR3A), and neutrophils (S100P, NAMPT, FCGR3B). Subsequently, we further classified the monocytes/macrophages into resident macrophages (LYVE1, FOLR2), M1 macrophages (IL1B, TNF), foamy macrophages (APOE, FABP5) and monocytes (CD14, LYZ, S100A9). The T cells were also divided

into the following subgroups: CD8 + T cell (CD8A), CD8 + effector (CD8A, FAS, FASLG), CD4 + memory T cell (CD44, S100A4, GPR183), CD4 + Th1 (IL2, IFNG), Regulatory T cells (IL2RA, FOXP3). For clusters without specific markers, we annotated them as "Others" and removed these cells. After annotation, we re-examined the marker genes for each major cluster in the UMAP and calculated the marker genes for each group.

BayesPrism deconvolution and differential analysis of cells in bulk RNA sequencing

To mitigate potential biases and improve the deconvolution process, highly expressed genes were removed, including ribosomal protein genes and mitochondrial genes, from the Seurat object. These genes tend to dominate the expression distribution and can introduce biases that may hinder deconvolution. Additionally, genes with low transcription were removed as their measurements are more susceptible to noise. Furthermore, because the gender proportions in the scRNA samples and the bulk RNA samples were not consistent, following the recommendation of the authors of BayesPrism, genes located on the X and Y chromosomes were excluded[21]. In the deconvolution results of bulk RNA, the cell proportions between carotid and femoral artery plaques were compared. At the same time, we analyzed the relationship between key DEGs and cell proportions in both carotid and femoral artery plaques.

Identify the most disturbed cell types in carotid and femoral artery plaques

We employed Augur, a machine learning-based classification algorithm, to investigate the disturbances of inflammatory cells in carotid and femoral artery plaques[28]. We utilized the random forest classifier in Augur to analyze cell disturbances. Then, the differential gene expression, KEGG pathways, and GO terms of the cell types with the highest AUC values in carotid and femoral artery plaques were analyzed.

Identifying genes and cells associated with carotid plaque embolism using ScPagwas

To investigate the genes and cells associated with plaque embolism, the scPagwas analysis was conducted. The scPagwas method uses polygenic regression and integrates pathway activity from scRNA sequencing data with GWAS summary data to prioritize trait-risk genes and identify associated cell subpopulations[29]. We selected "embolic stroke" (finngen_R9_I9_STR_EMBOLIC) from the Finnish database as the traits of plaque embolism for carotid plaques. (Table S2) The scPagwas analysis was then performed on samples from carotid artery plaques, and the bootstrap bias estimate was used to validate the statistical significance of the associations between each identified cell cluster and the embolic traits. In the scPagwas analysis, we used Pearson correlation coefficient (PCC) values to calculate the genes associated with the trait. By calculating trait-risk scores (TRS) for each cell, we were able to determine the overall trait risk within the cell population and identify the most relevant cells associated with the embolism. TRS were bias-corrected using 1,000 bootstrap iterations, subtracting bootstrap-estimated bias from original TRS. Significant TRS scores were defined as those exceeding the 95th percentile of a null distribution generated by 1,000 trait label permutations. (See supplement materials for details)

Integration analysis of disturbed cells, genes, and embolic stroke-related genes in carotid and femoral plaques

To investigate the differences between carotid and femoral artery plaques and their stability, we conducted a joint analysis of embolism-related genes/cells and differential genes/cells. We used PySCENIC to analyze the differences in transcription factors associated with embolism-related genes[30]. Additionally, we employed CellChat and NicheNet to investigate the communication differences between the most disturbed cells in carotid and femoral artery plaques, as well as those related to embolism-associated genes[31,32].

Validation

To validate the key DEGs identified by machine learning in the carotid and femoral plaque, we used an additional dataset, GSE23304. From this dataset, we selected specimens of common femoral artery and superficial femoral artery plaques for comparison with carotid artery plaques. Since the dataset compares peripheral artery specimens with 7 carotid artery plaques, we multiplied the results by 7 and presented the findings as a 1:1 comparison between femoral and carotid artery plaques. We then selected key DEGs from this additional dataset, calculated the log2 fold change values and p-values, and performed RNA-level validation.

Next, we selected the most differentially expressed genes from the key DEGs, focusing on those previously reported in clinical studies, for IHC analysis to validate protein-level differences. Thus, *HOXC9* and *HOXC6*, which have been reported in previous literature, were selected[33,34].

We matched patients who recently underwent carotid or femoral artery endarterectomy at our center, ensuring consistency in age (± 3 years), sex, hypertension, and diabetes status. The plaques were placed in neutral-buffered formalin and routinely processed at the Department of Pathology in our center. After fixation for approximately 24 h, the tissue was placed in cassettes, paraffin-embedded, and manually sectioned to a thickness of 6 μm. For IHC, *HOXC9* and *HOXC6* were detected using rabbit polyclonal anti-*HOXC9* (YN0785, Immunoway, dilution 1:200) and rabbit polyclonal anti-*HOXC6* (YT2218, Immunoway, dilution 1:200), respectively. The stained slides were evaluated, and immunostains were scored using integrated optical densitometry (IOD) analysis with the Image-Pro Plus 6.0® image processing program (Media Cybernetics, Inc., USA), utilizing the IOD measurement profile installed in the program. We removed low-quality samples from the IHC results, and the basic information of the patients is presented in Tables S3 and S4. The results shown represent the mean optical density of all measurements.

To verify that the increase in embolic stroke in carotid plaques represents a collective change rather than the effect of a few individual cells, we classified the most disturbed cells into high and low embolic stroke gene expression groups based on the average expression levels of embolic stroke genes. We then analyzed the differences in the expression of these genes in the most disturbed cells from carotid and femoral plaques using

scRNA-seq. Finally, we conducted additional scRNA analysis (GSE224273) to validate the expression differences of *CD2* and *ITGB1* between symptomatic and asymptomatic plaques.

Statistical analysis

The statistical analysis was conducted using R, version 4.3.0 (https://www.r-project.org/). To compare the two groups, Wilcoxon's signed-rank test was used. Pearson's or Spearman's correlation tests were used to determine the correlations between variables. The statistical p-values were two-sided, and $p < 0.05$ was considered statistically significant.

Results

Identification of the key DEGs and pathways between carotid and femoral artery plaques

The integration of three GEO datasets yielded a total of 61 carotid artery samples, 31 of which were included in the analysis. After batch effect correction (Fig. S2), compared to the femoral artery, the carotid artery had 80 downregulated and 90 upregulated genes (Fig. 1A and B). The GSEA suggested that enrichment of the inflammatory response was higher in the carotid artery than in the femoral artery ($p < 0.001$) (Fig. 1C). The GO analysis suggested that the upregulated DEGs in the carotid artery were primarily enriched in processes like regulation of plasma lipoprotein particle levels, while the downregulated genes were mainly enriched in muscle system process (Fig. 1D). The KEGG enrichment analysis showed that the DEGs were primarily enriched in the cholesterol metabolism pathway (Fig. 1E).

Multiple machine learning models indicated that most models performed well in both the training and validation sets, and all models had an area under the curve value of > 0.8 (Fig. 1F). Through the application of various machine learning approaches, we ultimately identified 10 genes that ranked the highest and appeared most frequently across the models, which we designated as the key DEGs, including *HOXC9*, *HOXC6*, *HOXC10*, *CACNA1H*, *HOXA9*, *HOXA7*, *HOXA6*, *HOXA5*, *APOC2*, and *HOXC8* (Fig. 1G). Among them, *APOC2* was upregulated in the carotid artery, while the rest were downregulated.

Single-cell analysis of carotid and femoral artery plaques

After initial quality control, four carotid artery plaque samples (containing 30,655 cells) and nine femoral artery plaque samples (containing 35,265 cells) were included. The cells were annotated into several clusters: residence macrophages, M1 macrophages, foamy macrophages, monocytes, CD8 + T cells, CD4 + memory T cells, CD4 + Th1 cells, Regulatory T cells, CD8 + cytotoxic T cells, B cells, NK cells, and neutrophils. (Fig. 2A) The UMAP plots of carotid and femoral artery cells are shown in Fig. 2B and D. The cell ratios for each sample are displayed in Fig. 2C. The markers of the cells were represented on the UMAP plots, indicating good cluster annotation. (Fig. 2E) The marker genes for each cell were shown in Fig. 2F.

Cell type proportions identified by bulk RNA bayesprism Deconvolution

Quality control of bulk RNA BayesPrism deconvolution resulted in the removal of most ribosomal and mitochondrial genes (Fig. 3A). The BayesPrism deconvolution showed strong correlations between single-cell and bulk RNA protein-coding genes, while long non-coding RNAs and pseudogenes were largely unselected, indicating that the deconvolution was predominantly based on protein-coding genes (Fig. 3B). Carotid artery plaques had a significantly higher proportion of CD8 + effector, CD4 + memory T cell, monocyte, CD8 + T cell, foamy macrophage and resident macrophage, while the proportions of NK cell, neutrophil and B cell were significantly lower in carotid artery plaques (Fig. 3C). In carotid artery plaques, key DEGs were more strongly associated with monocytes. (Fig. 3D) In contrast, in femoral artery plaques, NK cells and neutrophils showed a greater association with key DEGs. (Fig. 3E)

Identify the cell disturbances in the carotid and femoral plaques

Augur identified the AUC values for cell disturbances in carotid and femoral artery plaques, with the highest rankings for NK cells, B cells, and foamy macrophages, respectively. (Fig. 4A) We conducted a differential gene analysis on the most disturbed cells, specifically NK cells, which revealed significant upregulation of *MTRNR2L8*, *ITGB1*, and *SLFN5* in carotid plaques. (Fig. 4B)

GO enrichment analysis showed that NK cells in carotid plaques are enriched in leukocyte activation and adhesion to endothelial cells, while those in femoral plaques focus on regulating leukocyte-mediated immunity and natural killer cell activity. (Fig. 4C) KEGG analysis revealed that NK cells in carotid plaques are enriched in the tight junction and leukocyte transendothelial migration pathways, whereas those in femoral plaques are associated with apoptosis and ferroptosis pathways. (Fig. 4D)

The ScPagwas analysis identified genes and cells associated with embolism

scPaGwas integrated scRNA-seq and GWAS data, and the bootstrap results indicated that NK cells, CD4 + Th1 cells, and resident macrophages are associated with embolic stroke in carotid plaques. (Fig. 5A) The genes most closely related to embolic stroke in the plaques include *GNAS*, *FOS*, *IL32*, *CD2*, and *ITGB1*. (Fig. 5B) scPagwas integrated gene-related pathways and assigned TRS scores to the cells. The results showed that cell clusters with high TRS scores were mainly concentrated in T cells and NK cells, with NK cells having higher TRS scores than other immune cells. (Fig. 5C, D and E).

Integration analysis of disturbed cells and genes related to embolic stroke in carotid and femoral plaques

The above results suggested that the most disturbed cells in carotid and femoral plaques were NK cells, and scPaGwas indicated that NK cells were also the most relevant cells associated with embolic stroke in carotid

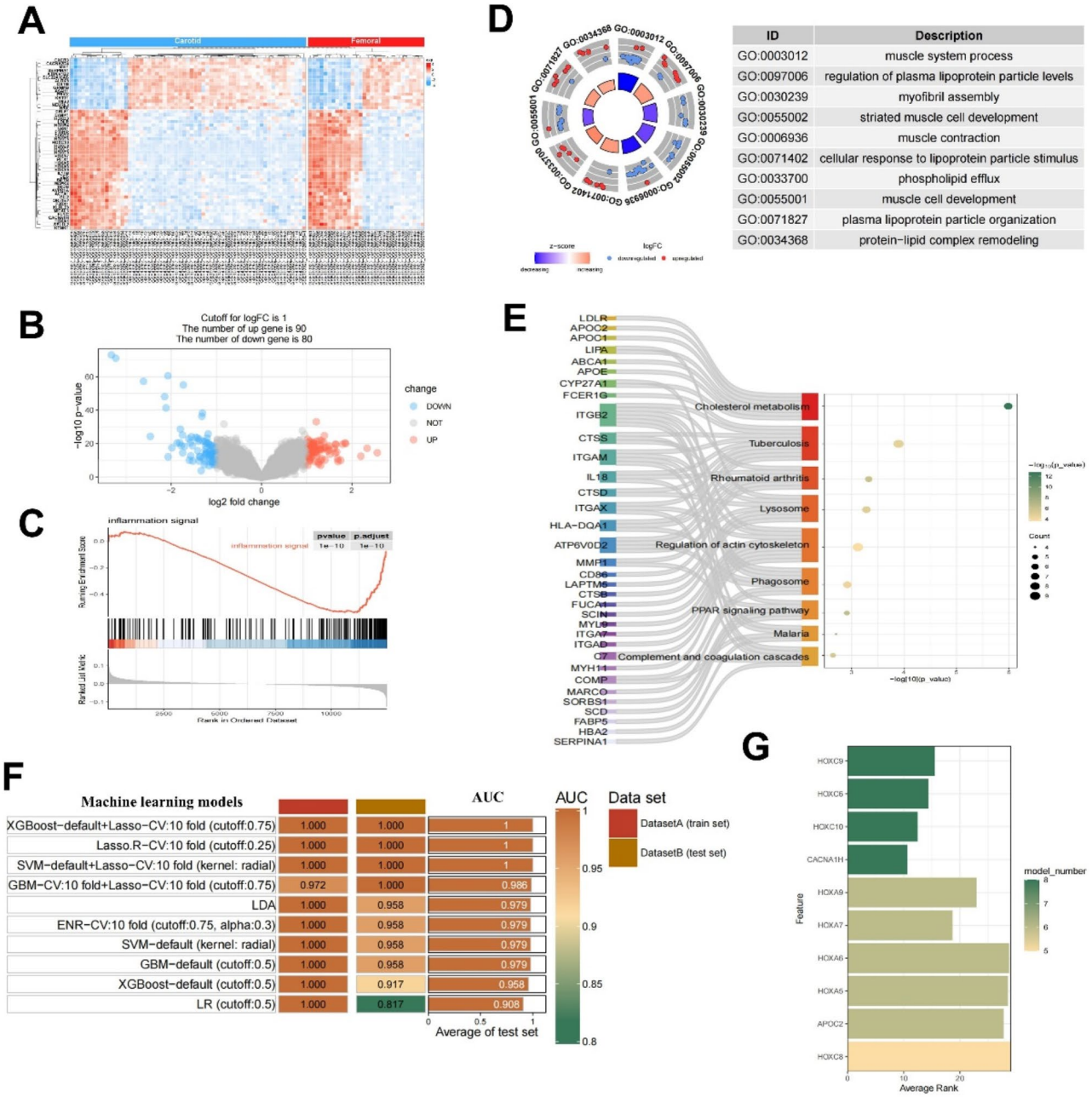

Fig. 1. Differential gene analysis in Bulk RNA of carotid and femoral artery plaques. (**A**) Heatmap of differential gene expression. Red represents genes that are upregulated, and blue represents genes that are downregulated. (**B**) Volcano plot of differential gene expression. Red represents genes that are upregulated in carotid plaques, and blue represents genes that are downregulated in carotid plaques. Gray represents genes with non-significant differential expression. (**C**) Enrichment analysis of differential inflammatory signaling using the gene set enrichment analysis. (**D**) Gene Ontology analysis of differential gene expression between carotid and femoral artery plaques. (**E**) Kyoto Encyclopedia of Genes and Genomes Sankey Enrichment Plot of differential gene expression. The leftmost column shows differentially expressed genes, the middle column shows enriched pathways, bubble size indicates the number of enriched genes, and color shows the -log10(*p*-value). (**F**) AUC results of machine learning models for the selection of key differentially genes between carotid and femoral artery plaques. (**G**) The top 10 most important genes (key differentially expressed genes) in machine learning models. The genes are sorted on the y-axis according to the number of times they appeared in the models. The color represents how many machine learning models retained the gene as an important feature (and did not filter it out), and the x-axis represents the average importance ranks of these genes across the retained models. AUC, area under the curve; LASSO, Least Absolute Shrinkage and Selection Operator. GBM, Gradient Boosting Machine. SVM, Support Vector Machine. LDA, Linear Discriminant Analysis. ENR, Elastic Net Regularization. LR, Logistic Regression.

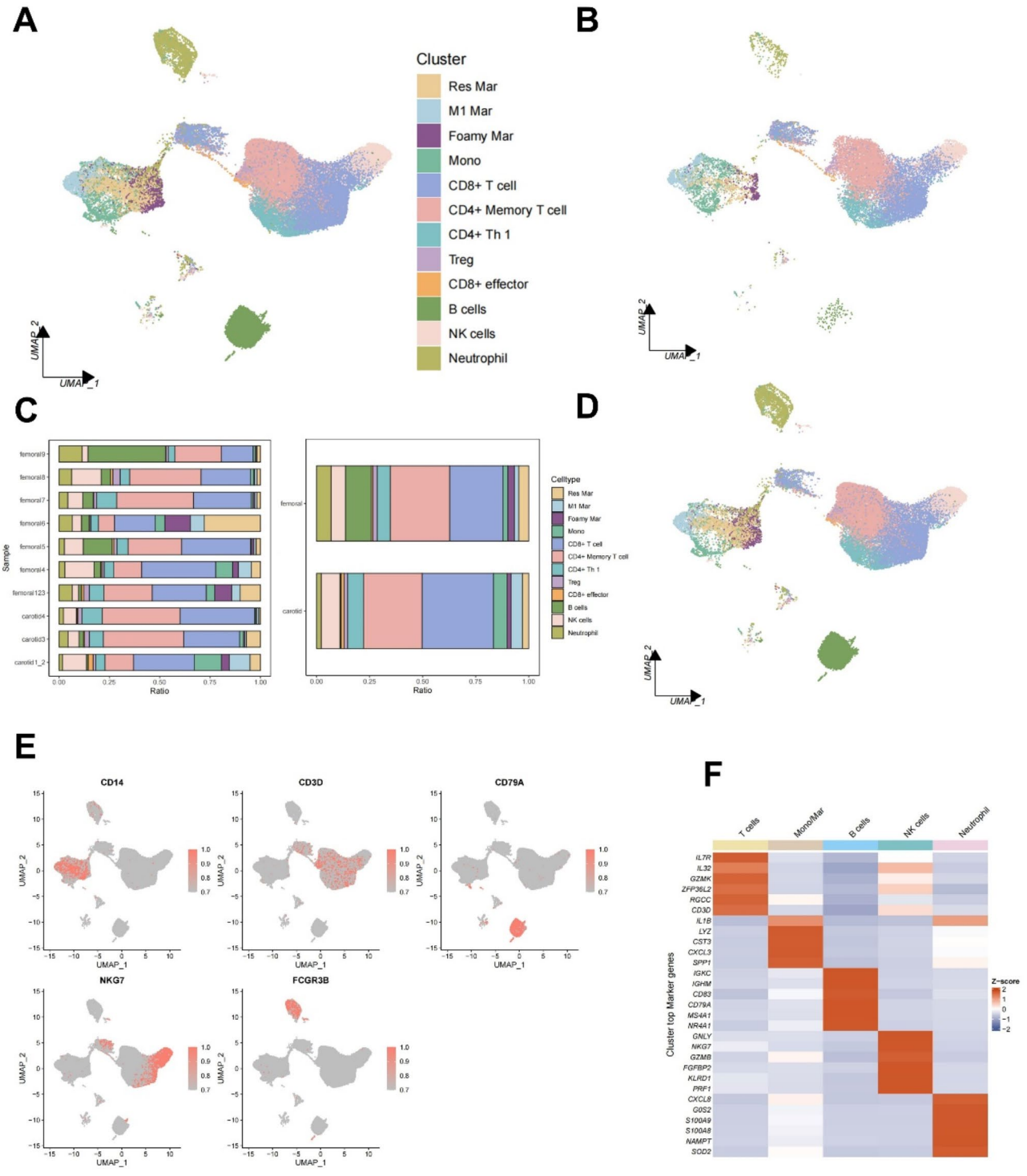

Fig. 2. Single-cell RNA analysis of carotid and femoral artery plaques. (**A**) UMAP plot of cell clusters. (**B**) UMAP plot of carotid cell clusters. (**C**) Proportional representation of each cell type. (**D**) UMAP plot of femoral cell clusters. (**E**) The distribution of each cell marker in the UMAP plot. (**F**) Heatmap of marker genes for each cell type. Res, residence. Mar, macrophage. Mono, monocyte. Treg, regulatory T cell. NK, natural killer. UMAP, uniform manifold approximation and projection.

plaques. The Venn diagram of the top 10 differentially expressed genes in NK cells from carotid and femoral plaques and the top 10 PCC genes related to embolic stroke is shown in Fig. 5F, with *CD2* and *ITGB1* being the common genes. We further investigated the transcriptional regulation of *CD2* and *ITGB1* and found that their transcription factors, *STAT2* and *BPTF* (Fig. 5G and H), were expressed at higher levels in the carotid plaque compared to the femoral plaque. (Fig. 5I and J)

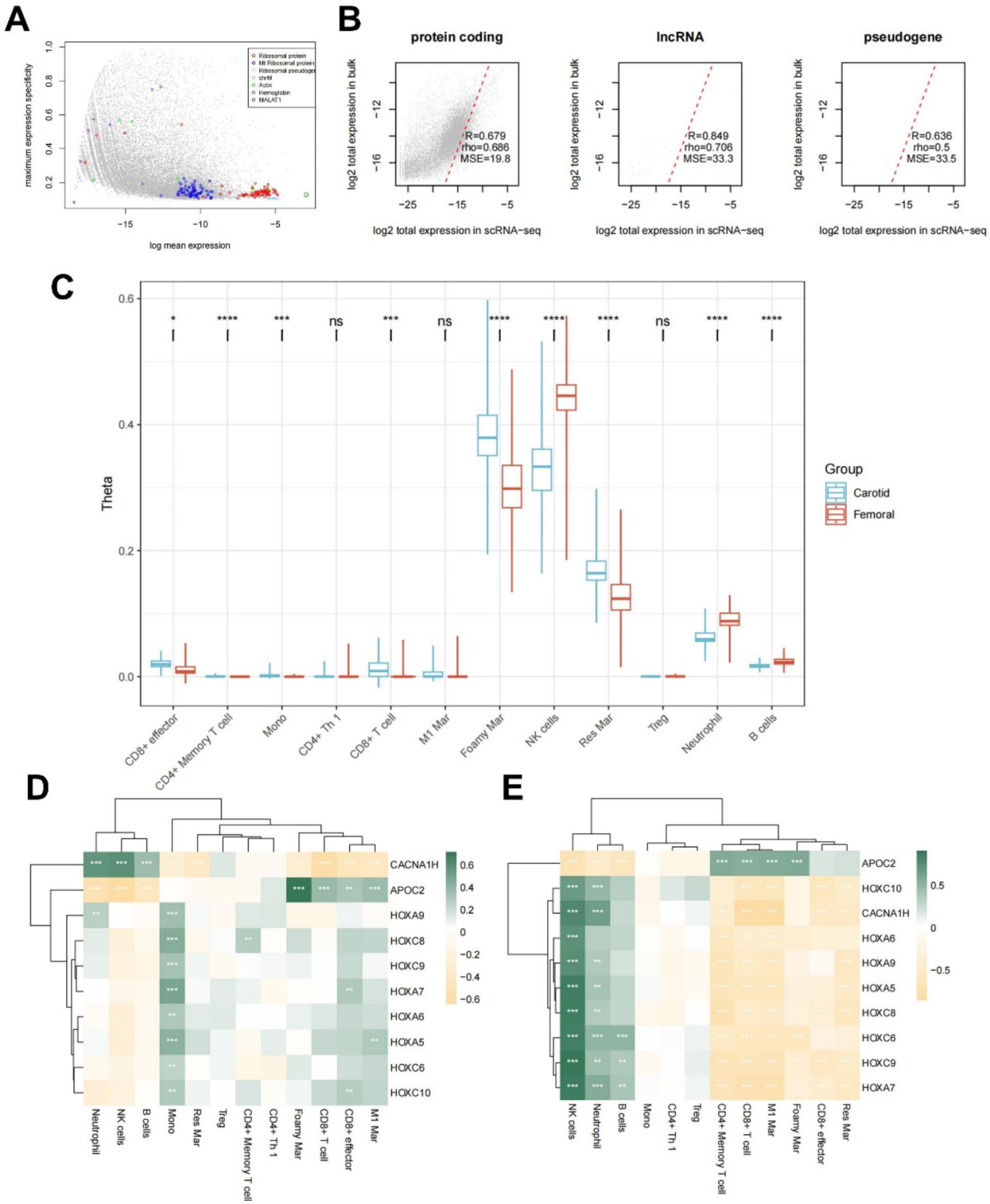

Fig. 3. Bulk RNA deconvolution using BayesPrism and cell proportions in carotid and femoral artery plaques. (**A**) Genes potentially generating noise were removed during Bayesian deconvolution. (**B**) Correlation between gene expression in bulk RNA and scRNA sequencing. (**C**) Differences in inflammatory cells between carotid and femoral artery plaques in bulk RNA data. (**D**) and (**E**) The correlation between key differentially expressed genes and inflammatory cells in carotid and femoral plaques. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$; ns: no significant.

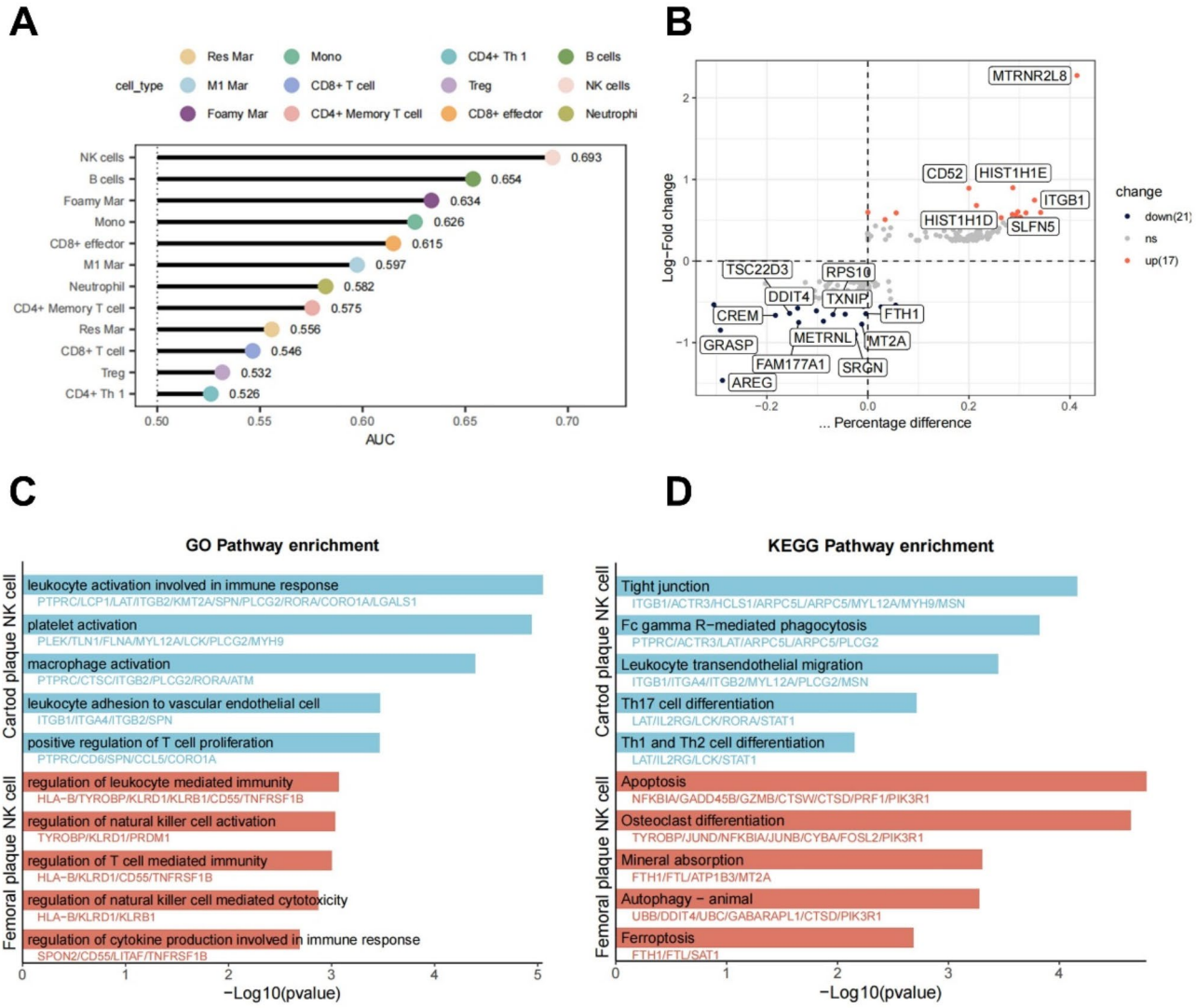

Fig. 4. The disturbance of cells in carotid and femoral artery plaques. (**A**) The disturbance of cells in carotid and femoral artery plaques. A larger AUC value indicates a more significant disturbance. (**B**) Volcano plot of differentially expressed genes in NK cells from carotid and femoral artery plaques. C and D, Gene Ontology and Kyoto Encyclopedia of Genes and Genomes enrichment analysis of differentially expressed genes in NK cells from carotid and femoral artery plaques.

We analyzed cell communication in carotid and femoral plaques and found that intercellular communication among immune cells is stronger in carotid plaques. (Fig. 6A) NK cells and M1 macrophages in carotid plaques communicated more actively, while foamy macrophages and neutrophils in femoral plaques showed higher communication levels. (Fig. 6B) NK cells in carotid plaques have more active signaling compared to those in femoral plaques, with the difference being particularly pronounced in their interaction with CD8 + T cells. (Fig. 6C and D) In the differences between NK cell signaling in carotid and femoral plaques, the alterations in APOE, VCAN, and SPP1 signals mediated by M1 macrophages and foamy macrophages are particularly notable. (Fig. 6E) In carotid plaques, CD8 + T cells exhibit increased GNAS signaling. Additionally, macrophages in these plaques release elevated levels of APOE signals, which target IL1B in NK cells, thereby promoting inflammation.

Validation

To validate the key DEGs in carotid and femoral artery plaques, we performed validation using an additional dataset. We found that in both the common femoral artery plaques and superficial femoral artery plaques, genes from the HOX family showed the largest differential expression, with *HOXC10*, *HOXC9*, *HOXA7*, and *HOXC6* exhibiting the most significant differences. (Fig. 7A and B)

To further validate the protein-level expression differences of the HOX family in carotid and femoral artery plaques, we selected *HOXC9* and *HOXC6* for IHC analysis. The results showed that the protein expression levels of *HOXC9* were higher in femoral artery plaques compared to carotid artery plaques (Fig. 7C and D), whereas no significant difference was observed for *HOXC6*. (Fig. 7E and F)

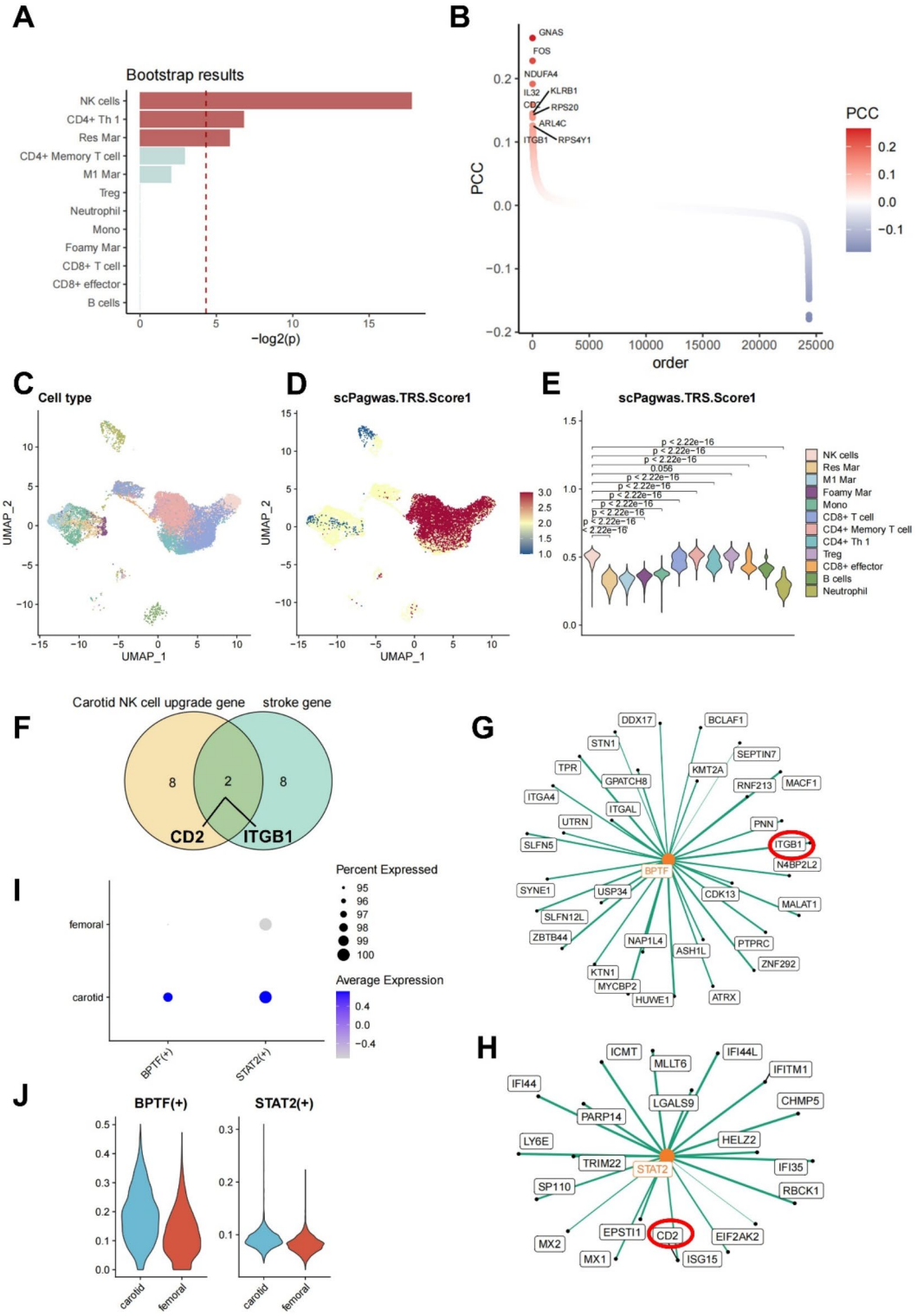

Fig. 5. Genes and cells associated with embolism in carotid and femoral artery plaques analyzed by scPagwas. (**A**) Bootstrap results of the scPagwas bias estimate for cell types in carotid plaque associated with embolic stroke. (**B**) The genes most correlated with embolic stroke in carotid artery plaques. (**C**) and (**D**), UMAP plot of cell types and TRS scores with scPagwas associated with embolic stroke in carotid artery plaques. E, Violin plot of TRS scores in carotid artery plaques associated with embolic stroke. (**F**) Venn diagram of upregulated genes in NK cells from carotid plaques and genes associated with embolism stroke. (**G**) and (**H**). Network diagram of genes regulated by transcription factors BPTF and STAT2. TRS, trait-risk scores; UMAP, uniform manifold approximation and projection. (**I**) and (**J**), Differences in transcription factors BPTF and STAT2 in NK cells from carotid and femoral artery plaques.

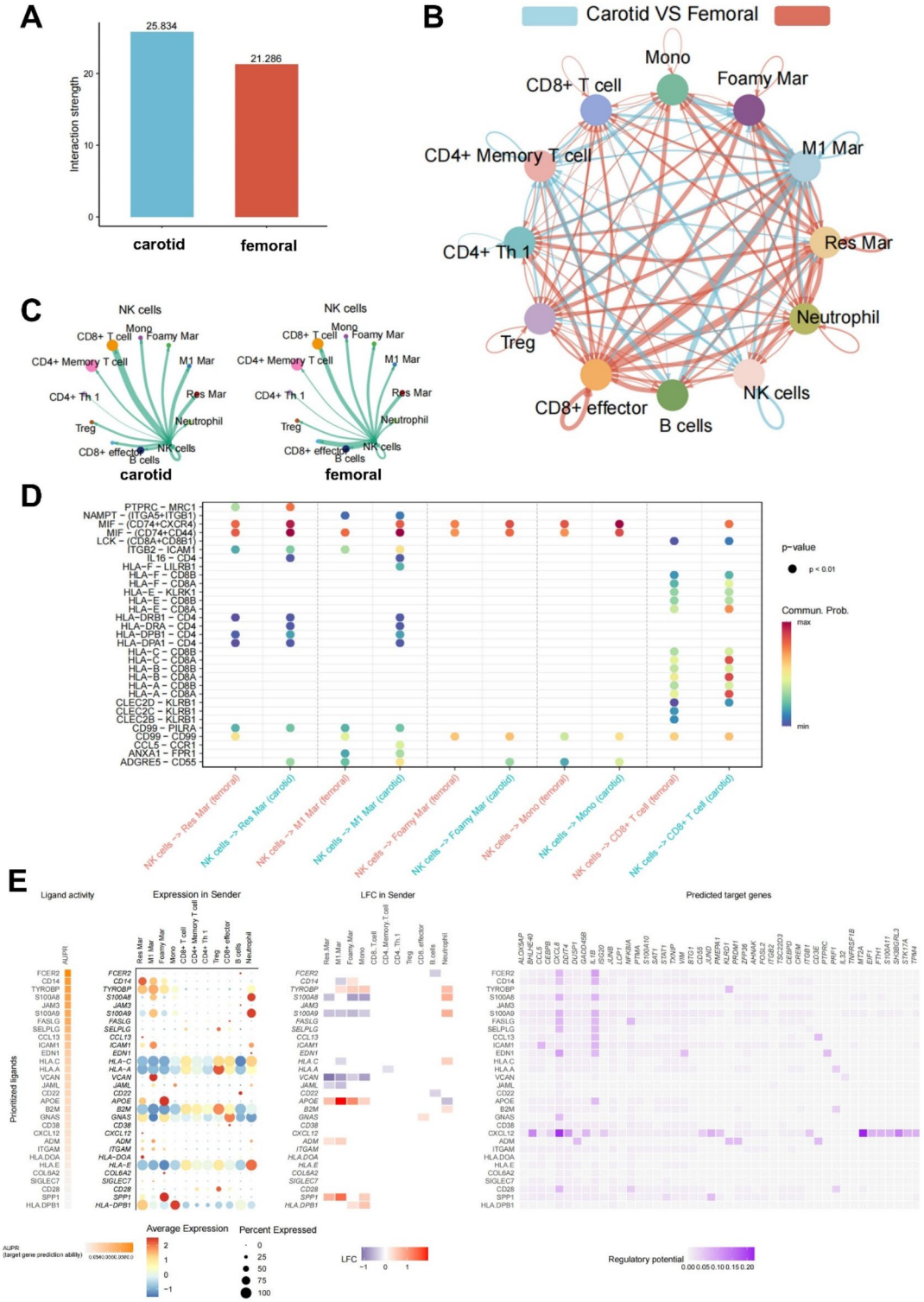

Fig. 6. Differences in NK cell communication in carotid and femoral artery plaques. (**A**) Differences in the number of communications in carotid and femoral artery plaques. (**B**) Cellular communication network in carotid and femoral artery plaques. (**C**) Differences in signal transmission of NK cells in carotid and femoral artery plaques. (**D**) Differences in communication of NK cells to other cells in carotid and femoral artery plaques. (**E**) Differences in the communication of NK cells in receiving signals in carotid and femoral artery plaques.

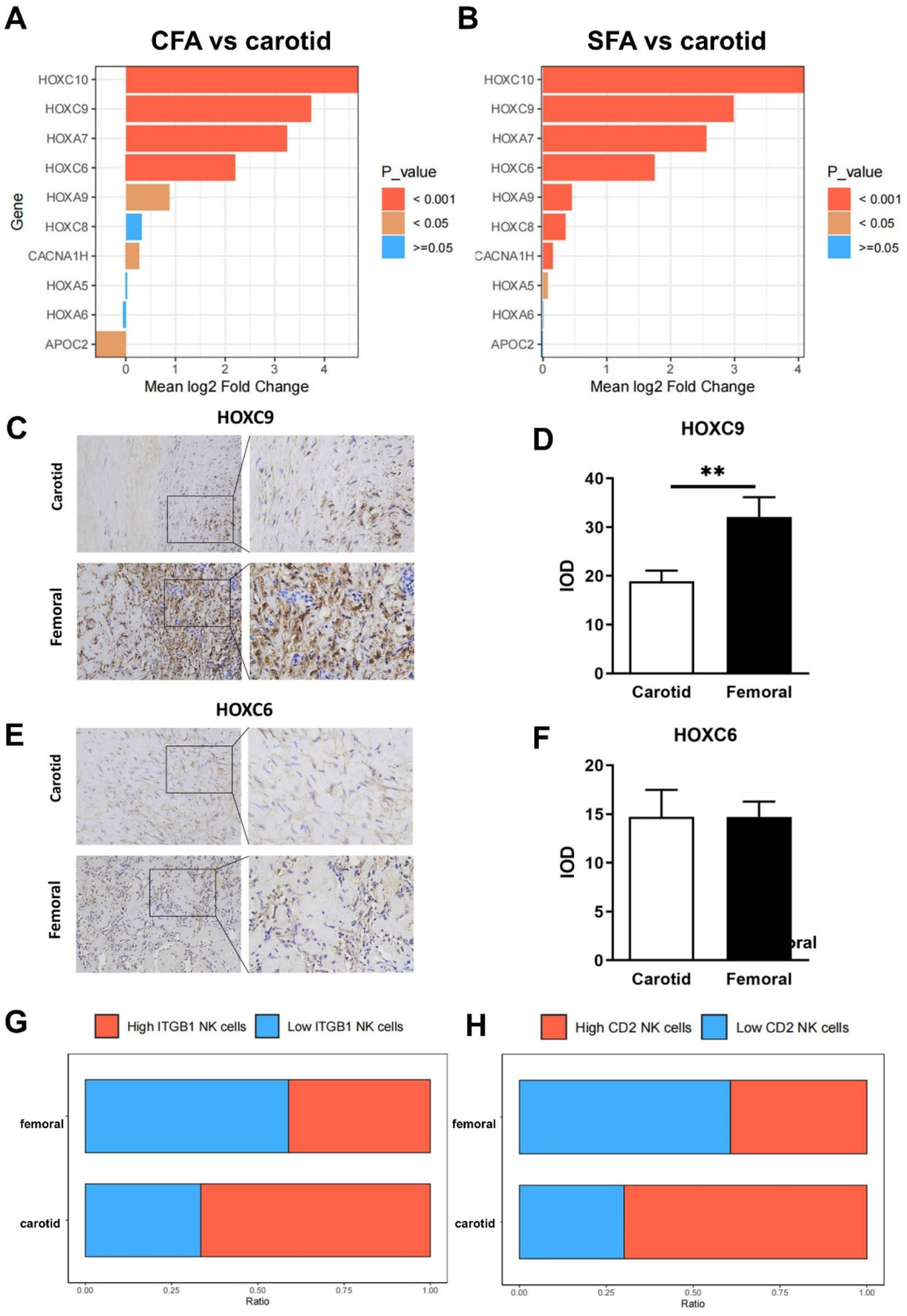

Fig. 7. Validation of key DEGs between carotid and femoral artery plaques. (**A**) and (**B**), Validation of the RNA expression differences of key DEGs in carotid and CFA/ SFA plaques using an additional dataset (GSE23304). (**C**) Representative IHC staining of *HOXC9*. (**D**) Quantitative expression level of *HOXC9* in carotid and femoral plaque. (**E**) Representative IHC staining of *HOXC6*. (**F**) Quantitative expression level of *HOXC6* in carotid and femoral plaque. (**G**) and (**H**), The proportion of NK cells with high/low expression of *ITGB1* and CD2 in carotid and femoral artery plaques. CFA, common femoral artery; IHC, immunohistochemical; SFA, superficial femoral artery; IOD, integrated optical density. **: $P < 0.01$.

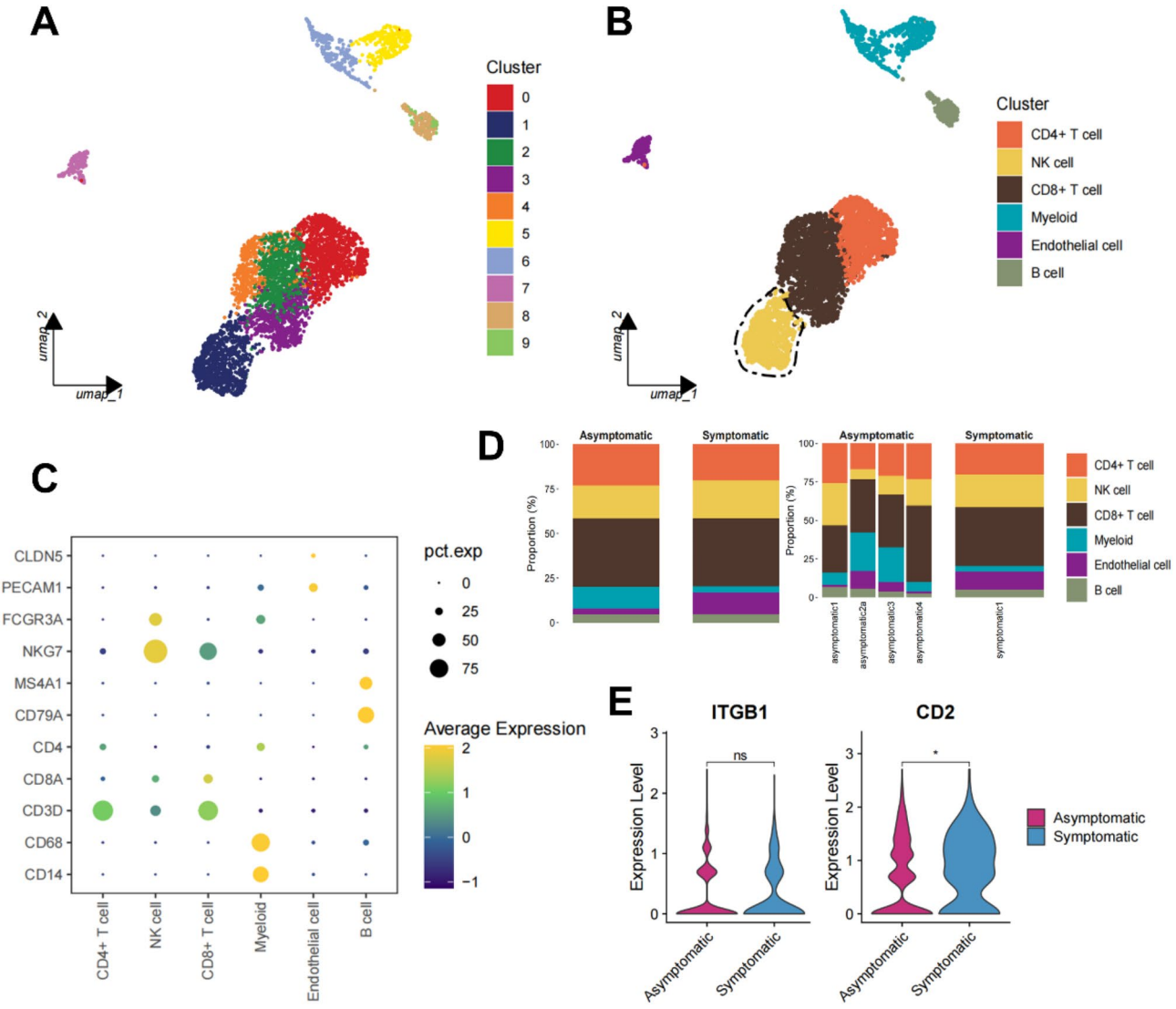

Fig. 8. Single-cell RNA analysis of symptomatic and asymptomatic plaques. (GSE224273). (**A**) UMAP plot of cell clusters. (**B**) UMAP plot of annotated cell clusters. (**C**) Marker gene expression of each cell cluster. D, Proportional representation of each cell type. E, The expression of ITGB1 and CD2 in NK cells in symptomatic and asymptomatic plaques.

In our scRNA-seq results, we observed that the number of NK cells exhibiting high *CD2* and high *ITGB1* expression in carotid plaques was significantly higher than in femoral plaques. (Fig. 7G and H) This indicates that the overall expression levels of *CD2* and *ITGB1* in NK cells within carotid plaques were elevated.

We conducted another scRNA analysis to validate the relationship between the expression of *CD2* and *ITGB1* in NK cells and plaque stability. The clustering, annotation, marker genes, and cell distribution of the scRNA were shown in Fig. 8A, B, C, and D. By comparing symptomatic (with embolic events) and asymptomatic (without embolic events) carotid plaques, we found that NK cells in symptomatic plaques exhibited significantly higher expression of *CD2*, while the difference in *ITGB1* expression was not statistically significant. (Fig. 8E)

Discussion

In this study, we used machine learning to identify key DEGs between carotid and femoral artery plaques. Combining scRNA-seq and bulk RNA analysis with the BayesPrism algorithm revealed differences in inflammatory cell proportions in carotid and femoral plaques. We found that NK cells were most relevant to embolic stroke in carotid plaques and showed the greatest disturbance compared to femoral plaques. Additionally, expression differences of *CD2* and *ITGB1* in NK cells may significantly associated with the stability differences between these plaques. Transcription factors regulating *CD2* and *ITGB1* were upregulated in carotid plaques compared to femoral plaques. NK cells may influence plaque instability by receiving signals from macrophages, mediating the cytotoxic effects of CD8 + T cells. The expression of *CD2* in NK cells was higher in plaques from symptomatic patients compared to NK cells in plaques from asymptomatic patients.

Although the carotid arteries generally exhibit stronger inflammatory characteristics, our research, through the combined use of machine learning and single-cell sequencing, found that the key DEGs between carotid and femoral artery plaques were not primarily related to inflammation; rather, these gene were from the *HOX* family[19,35]. In recent years, the relationship between the *HOX* family of genes and atherosclerosis has increasingly been recognized by researchers[36–39]. Experimental studies reveal distinct anti-inflammatory modalities within the HOX family: *HOXA5* exerts its endothelial-protective effects by blocking TNF-α-mediated monocyte-endothelial adhesion in vitro, while *HOXA9* ablation in atherosclerotic models improves coronary microvascular function through a triad mechanism involving PF4, E-selectin, and VCAM-1 suppression, illustrating isoform-selective transcriptional regulation of vascular inflammation[37,40]. In vitro experiments have shown that *HOXC9* overexpression blocks endothelial cell proliferation, migration, and tube formation[34,41]. Our study found that *HOXC9* is expressed at higher levels in femoral artery plaques compared to carotid artery plaques, both at the mRNA and protein levels. Previous studies have shown that high expression of *HOXC9* can stabilize plaques by inhibiting *IL-8*, which reduces the expression of *MMP9*[42,43]. Therefore, we hypothesize that the high expression of *HOXC9* may be one of the factors contributing to the greater stability of femoral artery plaques compared to carotid artery plaques.

The level of inflammatory cell infiltration in carotid and femoral artery plaques has been known to differ, and preliminary investigations have been conducted on this topic[44]. Our study validated this concept using the BayesPrism deconvolution approach on a large sample. The observed discrepancies between the cell proportions in single-cell samples and bulk RNA analysis after deconvolution using BayesPrism emphasize the importance of large-scale analyses. Despite BayesPrism's deconvolution results indicating a potentially higher proportion of NK cells in femoral artery plaques, KEGG enrichment analysis reveals that NK cells in carotid artery plaques exhibit greater inflammatory activity. In contrast, the pathways enriched in NK cells from femoral artery plaques are primarily related to the regulation of inflammatory stability.

Previous studies have mentioned the relationship between T cells, macrophages, and carotid plaque embolism[19,45–47]. Our study, combining scRNA-seq with GWAS data analysis, identified NK cells as a key factor associated with the occurrence of embolic stroke in carotid plaques. Previous studies have noted that patients with carotid plaques who experienced embolic strokes are associated with increased activity among foamy macrophages[19]. In our study, NK cells in carotid plaques received more APOE-IL1B signals from foamy macrophages and sent more signals to T cells. The scPagwas results indicated that T cells and NK cells exhibit relatively high TRS scores, highlighting a significant association with plaque-related embolic stroke. The well-known *APOE*, which is highly expressed in foam cells, may influence NK cells, leading to increased T cell activity and ultimately resulting in embolic stroke[48]. The communication between NK cells and M1 macrophages mainly focuses on the MIF-(CD74/CXCR4) signaling pathway. Some studies have confirmed that the activation of MIF signaling may lead to M1 cell polarization, promoting inflammation, which could be one of the reasons why NK cells are associated with plaque embolism[49]. Additionally, our study found that the high expression of *CD2* in NK cells may be one of the important factors associated with an increased likelihood of embolic events in carotid plaques. Currently, there are no studies on *CD2* in atherosclerotic plaques. *CD2* is an important adhesion molecule in NK cells that enhances their ability to kill target cells. In additional scRNA analysis, we found that NK cells with high *CD2* expression were significantly increased in symptomatic patients, which may support the hypothesis that these cells play an important role in plaque rupture. Recent studies suggest that high expression of *CD2* may be a key factor in various inflammatory diseases and can also strengthen NK cell cytotoxicity against tumor cells[50,51]. In carotid plaques, the high expression of *CD2* in NK cells may enhance their ability to target necrotic cells within the plaque, potentially leading to plaque instability. Emerging evidence suggests that targeting *CD2*—a therapeutic strategy implicated in chronic inflammatory diseases—may also hold promise for vascular pathologies. This is supported by observations in abdominal aortic aneurysms, where CD8 + cytotoxic T cell infiltration positively correlates with *CD2* expression. Such findings underscore the role of *CD2*-mediated immune mechanisms in disease pathogenesis and highlight novel research avenues for immunomodulatory interventions aimed at stabilizing arterial plaque integrity[50,52].

Our study identified several genes associated with embolism in carotid plaques, such as *GNAS*, *FOS*, and *IL32*, which are more highly expressed in NK cells and T cells and are closely linked to plaque rupture. Some population cohort studies have also found correlations between these genes and stroke; however, the specific molecular mechanisms remain unclear and warrant further investigation by future researchers[53,54].

This study has certain limitations. First, the single-cell sequencing data were subjected to $CD45^+$ cell selection, which makes it challenging to evaluate the situation of cells other than inflammatory cells in carotid and femoral artery plaques (such as endothelial cells and smooth muscle cells). Future studies should include non-immune cells (such as endothelial cells, fibroblasts, etc.) in analysis. Second, although this study identified certain genes that have not been mentioned in previous research, the mechanisms underlying their roles remain to be elucidated as experimental evidence was lacking in this study. Finally, both single-cell and bulk RNA samples were obtained from patients who underwent carotid/femoral artery endarterectomy, representing cases with severe conditions that warranted surgical intervention. Despite these limitations, this study may have identified differences in the genetic profiles of carotid and femoral plaques, as well as the reasons for their differing stability.

Conclusions

In summary, our research results indicate that carotid and femoral artery plaques exhibit differences in gene expression, cell–cell interactions, and inflammatory cell infiltration. Additionally, the cells associated with the differing stability of carotid and femoral plaques may be NK cells that express high levels of *CD2*. These findings provide important clues and serve as a basis for further research into the precise prevention and treatment of atherosclerosis in different vascular regions.

Data availability

All data in this study are available online. The following datasets can be accessed: GSE234077 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE234077); GSE100927 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE100927); GSE53274 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53274), GSE43292 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE43292); GSE23304 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE23304); GSE100927 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE100927); GSE224273 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi? acc=GSE224273);finngen_R9_I9_STR_EMBOLIC (https://storage.googleapis.com/finngen-public-data-r9/summary_stats/finngen_R9_I9_STR_EMBOLIC.gz).

Received: 28 January 2025; Accepted: 28 March 2025
Published online: 07 April 2025

References

1. Nedkoff, L., Briffa, T., Zemedikun, D., Herrington, S. & Wright, F. L. Global trends in atherosclerotic cardiovascular disease. *Clin. Ther.* **45**, 1087–1091. https://doi.org/10.1016/j.clinthera.2023.09.020 (2023).
2. Soehnlein, O. & Libby, P. Targeting inflammation in atherosclerosis - from experimental insights to the clinic. *Nat. Rev. Drug Discov.* **20**, 589–610. https://doi.org/10.1038/s41573-021-00198-1 (2021).
3. Libby, P. The changing landscape of atherosclerosis. *Nature* **592**, 524–533. https://doi.org/10.1038/s41586-021-03392-8 (2021).
4. Roy, P., Orecchioni, M. & Ley, K. How the immune system shapes atherosclerosis: Roles of innate and adaptive immunity. *Nat. Rev. Immunol.* **22**, 251–265. https://doi.org/10.1038/s41577-021-00584-1 (2022).
5. Global burden of peripheral artery disease and its risk factors. 1990–2019: a systematic analysis for the global burden of disease study 2019. *Lancet Global Health*. **11**, e1553–e1565. https://doi.org/10.1016/s2214-109x(23)00355-8 (2023).
6. McDermott, M. M. et al. Disparities in diagnosis, treatment, and outcomes of peripheral artery disease: JACC scientific statement. *J. Am. Coll. Cardiol.* **82**, 2312–2328. https://doi.org/10.1016/j.jacc.2023.09.830 (2023).
7. Smolderen, K. G. et al. Association between mental health burden, clinical presentation, and outcomes in individuals with symptomatic peripheral artery disease: A scientific statement from the American heart association. *Circulation* **148**, 1511–1528. https://doi.org/10.1161/cir.0000000000001178 (2023).
8. Allison, M. A. et al. Health disparities in peripheral artery disease: A scientific statement from the American heart association. *Circulation* **148**, 286–296. https://doi.org/10.1161/cir.0000000000001153 (2023).
9. Seijkens, T. T. P. et al. Deficiency of the T cell regulator Casitas B-cell lymphoma-B aggravates atherosclerosis by inducing CD8 + T cell-mediated macrophage death. *Eur. Heart J.* **40**, 372–382. https://doi.org/10.1093/eurheartj/ehy714 (2019).
10. Saigusa, R., Winkels, H. & Ley, K. T cell subsets and functions in atherosclerosis. *Nat. Rev. Cardiol.* **17**, 387–401. https://doi.org/10.1038/s41569-020-0352-5 (2020).
11. Fredman, G. & Serhan, C. N. Specialized pro-resolving mediators in vascular inflammation and atherosclerotic cardiovascular disease. *Nat. Rev. Cardiol.* https://doi.org/10.1038/s41569-023-00984-x (2024).
12. Steenman, M. et al. Identification of genomic differences among peripheral arterial beds in atherosclerotic and healthy arteries. *Sci. Rep.* **8**, 3940. https://doi.org/10.1038/s41598-018-22292-y (2018).
13. Pasterkamp, G., den Ruijter, H. M. & Giannarelli, C. False utopia of one unifying description of the vulnerable atherosclerotic plaque: A call for recalibration that appreciates the diversity of mechanisms leading to atherosclerotic disease. *Arterioscler. Thromb. Vasc. Biol.* **42**, e86–e95. https://doi.org/10.1161/atvbaha.121.316693 (2022).
14. Holder, T. A. & Aday, A. W. Symptom progression in peripheral artery disease: sounding the alarm for cardiovascular risk. *Circ. Cardiovasc. Interv.* **13**, e010021. https://doi.org/10.1161/circinterventions.120.010021 (2020).
15. Rymer, J. A. et al. Association of disease progression with cardiovascular and limb outcomes in patients with peripheral artery disease: insights from the EUCLID trial. *Circ. Cardiovasc. Interv.* **13**, e009326. https://doi.org/10.1161/circinterventions.120.009326 (2020).
16. Herisson, F. et al. Carotid and femoral atherosclerotic plaques show different morphology. *Atherosclerosis* **216**, 348–354. https://doi.org/10.1016/j.atherosclerosis.2011.02.004 (2011).
17. Depuydt, M. A. C. et al. Microanatomy of the human atherosclerotic plaque by single-cell transcriptomics. *Circul. Res.* **127**, 1437–1455. https://doi.org/10.1161/circresaha.120.316770 (2020).
18. Fernandez, D. M. et al. Single-cell immune landscape of human atherosclerotic plaques. *Nat. Med.* **25**, 1576–1588. https://doi.org/10.1038/s41591-019-0590-4 (2019).
19. Slysz, J. et al. Single-cell profiling reveals inflammatory polarization of human carotid versus femoral plaque leukocytes. *JCI Insight* **8** https://doi.org/10.1172/jci.insight.171359 (2023).
20. Wang, P. et al. A Single-Cell atlas of the atherosclerotic plaque in the femoral artery and the heterogeneity in macrophage subtypes between carotid and femoral atherosclerosis. *J. Cardiovasc. Dev. Disease*. **9** https://doi.org/10.3390/jcdd9120465 (2022).
21. Chu, T., Wang, Z., Pe'er, D. & Danko, C. G. Cell type and gene expression deconvolution with bayesprism enables bayesian integrative analysis across bulk and single-cell RNA sequencing in oncology. *Nat. cancer*. **3**, 505–517. https://doi.org/10.1038/s43018-022-00356-3 (2022).
22. Tran, K. A. et al. Performance of tumour microenvironment deconvolution methods in breast cancer using single-cell simulated bulk mixtures. *Nat. Commun.* **14**, 5758. https://doi.org/10.1038/s41467-023-41385-5 (2023).
23. Aavik, E. et al. Global DNA methylation analysis of human atherosclerotic plaques reveals extensive genomic hypomethylation and reactivation at imprinted locus 14q32 involving induction of a MiRNA cluster. *Eur. Heart J.* **36**, 993–1000. https://doi.org/10.1093/eurheartj/ehu437 (2015).
24. Ayari, H. & Bricca, G. Identification of two genes potentially associated in iron-heme homeostasis in human carotid plaque using microarray analysis. *J. Biosci.* **38**, 311–315. https://doi.org/10.1007/s12038-013-9310-2 (2013).
25. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. & Ishiguro-Watanabe M. KEGG: Biological systems database as a model of the real world. *Nucleic Acids Res.* https://doi.org/10.1093/nar/gkae909 (2024).
26. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.: Publ. Protein Soc.* **28**, 1947–1951. https://doi.org/10.1002/pro.3715 (2019).
27. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **51**, D587–d592. https://doi.org/10.1093/nar/gkac963 (2023).
28. Skinnider, M. A. et al. Cell type prioritization in single-cell data. *Nat. Biotechnol.* **39**, 30–34. https://doi.org/10.1038/s41587-020-0605-1 (2021).
29. Ma, Y. et al. Polygenic regression uncovers trait-relevant cellular contexts through pathway activation transformation of single-cell RNA sequencing data. *Cell. Genomics*. **3**, 100383. https://doi.org/10.1016/j.xgen.2023.100383 (2023).
30. Aibar, S. et al. SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086. https://doi.org/10.1038/nmeth.4463 (2017).

31. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: Modeling intercellular communication by linking ligands to target genes. *Nat. Methods*. **17**, 159–162. https://doi.org/10.1038/s41592-019-0667-5 (2020).
32. Jin, S., Plikus, M. V. & Nie, Q. CellChat for systematic analysis of cell-cell communication from single-cell transcriptomics. *Nat. Protoc.* https://doi.org/10.1038/s41596-024-01045-4 (2024).
33. Huang, C. et al. Long noncoding RNA HOXC-AS1 suppresses Ox-LDL-Induced cholesterol accumulation through promoting HOXC6 expression in THP-1 macrophages. *DNA Cell Biol.* **35**, 722–729. https://doi.org/10.1089/dna.2016.3422 (2016).
34. Stoll, S. J., Bartsch, S., Augustin, H. G. & Kroll, J. The transcription factor HOXC9 regulates endothelial cell quiescence and vascular morphogenesis in zebrafish via Inhibition of Interleukin 8. *Circul. Res.* **108**, 1367–1377. https://doi.org/10.1161/circresaha.111.244095 (2011).
35. Wu, X. et al. Identification of key genes for atherosclerosis in different arterial beds. *Sci. Rep.* **14**, 6543. https://doi.org/10.1038/s41598-024-55575-8 (2024).
36. Jing, Y., Gao, B., Han, Z., Xia, L. & Xin, S. The protective effect of HOXA5 on carotid atherosclerosis occurs by modulating the vascular smooth muscle cell phenotype. *Mol. Cell. Endocrinol.* **534**, 111366. https://doi.org/10.1016/j.mce.2021.111366 (2021).
37. Liu, S., Gao, J. & Wang, S. HOXA9 inhibitors promote microcirculation of coronary arteries in rats via downregulating E-selectin/VCAM-1. *Experimental Therapeutic Med.* **22**, 871. https://doi.org/10.3892/etm.2021.10303 (2021).
38. Jing, Y., Gao, B., Han, Z. & Xin, S. HOXA5 induces M2 macrophage polarization to attenuate carotid atherosclerosis by activating MED1. *IUBMB Life* **73**, 1142–1152. https://doi.org/10.1002/iub.2515 (2021).
39. Cao, W. et al. Hoxa5 alleviates obesity-induced chronic inflammation by reducing ER stress and promoting M2 macrophage polarization in mouse adipose tissue. *J. Cell. Mol. Med.* **23**, 7029–7042. https://doi.org/10.1111/jcmm.14600 (2019).
40. Lee, J. Y. et al. Human HOXA5 homeodomain enhances protein transduction and its application to vascular inflammation. *Biochem. Biophys. Res. Commun.* **410**, 312–316. https://doi.org/10.1016/j.bbrc.2011.05.139 (2011).
41. Stoll, S. J. & Kroll, J. HOXC9: a key regulator of endothelial cell quiescence and vascular morphogenesis. *Trends Cardiovasc. Med.* **22**, 7–11. https://doi.org/10.1016/j.tcm.2012.06.002 (2012).
42. Hanssen, N. M. et al. Higher levels of advanced glycation endproducts in human carotid atherosclerotic plaques are associated with a rupture-prone phenotype. *Eur. Heart J.* **35**, 1137–1146. https://doi.org/10.1093/eurheartj/eht402 (2014).
43. Gargiulo, S. et al. Relation between TLR4/NF-κB signaling pathway activation by 27-hydroxycholesterol and 4-hydroxynonenal, and atherosclerotic plaque instability. *Aging Cell* **14**, 569–581. https://doi.org/10.1111/acel.12322 (2015).
44. Mosser, D. M., Hamidzadeh, K. & Goncalves, R. Macrophages and the maintenance of homeostasis. *Cell Mol. Immunol.* **18**, 579–587. https://doi.org/10.1038/s41423-020-00541-3 (2021).
45. Gao, J. et al. Difference of immune cell infiltration between stable and unstable carotid artery atherosclerosis. *J. Cell. Mol. Med.* **25**, 10973–10979. https://doi.org/10.1111/jcmm.17018 (2021).
46. Tan, J. et al. Single-cell transcriptomics reveals crucial cell subsets and functional heterogeneity associated with carotid atherosclerosis and cerebrovascular events. *Arterioscler. Thromb. Vasc. Biol.* **43**, 2312–2332. https://doi.org/10.1161/atvbaha.123.318974 (2023).
47. Wang, J. et al. Identification of immune cell infiltration and diagnostic biomarkers in unstable atherosclerotic plaques by integrated bioinformatics analysis and machine learning. *Front. Immunol.* **13**, 956078. https://doi.org/10.3389/fimmu.2022.956078 (2022).
48. Guerrini, V. & Gennaro, M. L. Foam cells: One size doesn't fit all. *Trends Immunol.* **40**, 1163–1179. https://doi.org/10.1016/j.it.2019.10.002 (2019).
49. He, H. et al. HIF1α/MIF/CD74 signaling mediated OSA-induced atrial fibrillation by promoting M1 macrophages polarization. *Int. Immunopharmacol.* **149**, 114248. https://doi.org/10.1016/j.intimp.2025.114248 (2025).
50. Kashyap, M. P. et al. CD2 expressing innate lymphoid and T cells are critical effectors of immunopathogenesis in hidradenitis suppurativa. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2409274121. https://doi.org/10.1073/pnas.2409274121 (2024).
51. Kennedy, P. R. et al. Metabolic programs drive function of therapeutic NK cells in hypoxic tumor environments. *Sci. Adv.* **10**, eadn1849. https://doi.org/10.1126/sciadv.adn1849 (2024).
52. Li, T., Wang, T. & Zhao, X. Profiles of immune infiltration in abdominal aortic aneurysm and their associated marker genes: A gene expression-based study. *Brazilian J. Med. Biol. Res. Revista Brasileira De Pesquisas Medicas E Biologicas*. **54**, e11372. https://doi.org/10.1590/1414-431X2021e11372 (2021).
53. Li, G. et al. Identification of key genes associated with oxidative stress in ischemic stroke via bioinformatics integrated analysis. *BMC Neurosci.* **26** https://doi.org/10.1186/s12868-024-00921-9 (2025).
54. Zhu, H. et al. Interleukins and ischemic stroke. *Front. Immunol.* **13**, 828447. https://doi.org/10.3389/fimmu.2022.828447 (2022).

Acknowledgements

We thank Emily Woodhouse, PhD, from Liwen Bianji (Edanz) (http://www.liwenbianji.cn/) for editing the English text of a draft of this manuscript.

Author contributions

X.H.H. and Y.Q.D. designed the study. X.H.H. and W.L.L. drafted the manuscript. Z.P.L. and J.L. conducted the data analysis. W.L., L.Y.L., X.Q.Z., X.L.L., F.G.C., and J.C.Z. contributed to data collection and figure preparation. P.F.G. and Y.Q.D. performed the literature search. X.H.H., Z.P.L., W.L.L., L.Z. and Y.Q.D. contributed to data interpretation. C.W.Y. conducted external validation. All authors read and approved the final version of the manuscript.

Funding

This study was financially supported by Joint Funds for the Innovation of Science and Technology, Fujian Province (No. 2023Y9306), National Natural Science Foundation of China (No. 82204023) and Startup Fund for Scientific Research of Fujian Medical University (No. 2023QH1065).

Declarations

Competing interests

The authors declare no competing interests.

Ethical approval

This study was approved by the Ethics Committee of the First Affiliated Hospital of Fujian Medical University (approval number: FMU [2023] 508). All participants provided informed consent.

Consent for publication

The authors hereby declare that all participants involved in this study have provided informed consent for their data to be published.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-96434-4.

Correspondence and requests for materials should be addressed to X.H., C.Y. or Y.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.