scientific reports

Check for updates

OPEN

Deciphering the role of cuproptosis in the development of intimal hyperplasia in rat carotid arteries using single cell analysis and machine learning techniques

Miao He[1,4], Hui Chen[3,4], Zhengli Liu[2,4], Boxiang Zhao[2], Xu He[2], Qiujin Mao[1,5]✉, Jianping Gu[2,5]✉ & Jie Kong[2,5]✉

This study aims to explore the regulatory role of cuproptosis in carotid intimal hyperplasia (IH), providing new insights into its pathophysiological mechanisms and potential diagnostic and therapeutic strategies. Methods: We downloaded single-cell sequencing and bulk transcriptome data from the GEO database to screen for copper-growth-associated genes (CAGs) using machine-learning algorithms, including Random Forest and Support Vector Machine. After identifying relevant genes, we verified CAGs expression in IH and control groups using a rat model of carotid balloon strain. We analyzed the immune infiltration characteristics of carotid intimal hyperplasia and used electron microscopy to observe mitochondrial structural changes in cuproptosis. Additionally, we performed subgroup analyses of carotid balloon strains. The cuproptosis activity of VSMCs was explored in a single-cell dataset. Immunohistochemistry was applied to validate the expression of CAGs. Results: By means of machine learning algorithms, we identified several genes, including Pdhx and Fdx1, as novel therapeutic targets for carotid intimal hyperplasia. Meanwhile, immunohistochemistry results observed decreased expression of Pdhx and Fdx1 in the Neointimal hyperplasia(Neo) group. Immunohistochemical results showed a difference in cellular infiltration between Dendritic cells resting and Mast cells resting. By calculating cuproptosis activity in vascular smooth muscle cells (VSMCs), we found increased cuproptosis activity in normal vascular smooth muscle cells which was also observed in the electron microscopy. Microscopy revealed less mitochondrial swelling characteristic of cuproptosis in Neo group. Conclusion: The CAGs identified may regulate intimal hyperplasia in rat carotid arteries by modulating cuproptosis and represent potential targets for treatment.

Keywords Cuproptosis, Neointimal hyperplasia, Single-cell analysis, Carotid balloon strains, Machine-learning

Background

Carotid stenosis is characterized by the proliferation of smooth muscle cells in the carotid arteries. According to the American College of Chest Physicians in 2014, more than 795,000 people have a first or recurrent stroke each year. Of these, 87% are ischemic strokes, 10% are cerebral hemorrhages, and 3% are subarachnoid hemorrhages[1]. Carotid artery stenosis, primarily caused by the over-proliferation of carotid VSMCs, plays a significant role

[1]Department of General Surgery, Nanjing First Hospital, Nanjing Medical University, Nanjing 210006, Jiangsu, People's Republic of China. [2]Department of Interventional Radiology, Nanjing First Hospital, Nanjing Medical University, Nanjing 210006, Jiangsu, People's Republic of China. [3]Department of Pharmacy, The Affiliated Jiangning Hospital of Nanjing Medical University, Nanjing 211100, Jiangsu, People's Republic of China. [4]Miao He, Hui Chen and Zhengli Liu contributed equally to this work and share the co-first author. [5]Qiujin Mao, Jianping Gu, and Jie Kong contributed equally to this work and share the co-corresponding author. ✉email: njsdyyyjrk@163.com; cjr.gujianping@163.com; kilogram@163.com

in this condition[2]. It can be categorized into symptomatic and asymptomatic types, with risk factors including hypertension, hyperlipidemia, impaired glucose metabolism, obesity, physical inactivity, and sleep apnea[3]. The pathophysiological mechanisms underlying carotid artery stenosis are complex, involving multiple cell types and molecular pathways.

Current treatment options for carotid artery stenosis include surgical interventions such as angioplasty, carotid endarterectomy (CEA), and bypass grafting, as well as pharmacological treatments targeting underlying conditions and improving metabolic status[4,6]. However, understanding the pathogenesis of carotid artery stenosis and identifying preventive targets remainreserch focus.

Established methods for constructing animal models of carotid intimal hyperplasia include carotid artery balloon strain model. This model stimulates smooth muscle cell proliferation, leading to carotid artery stenosis[7]. Phenotypic transformation and overproliferation of vascular smooth muscle cells characterize carotid artery stenosis, and promotion of apoptosis seems to be an effective solution to prevent intimal hyperplasia. Promoting apoptosis while inhibiting smooth muscle cell over-proliferation may delay the onset and progression of carotid artery stenosis[8]. Cuproptosis, a newly identified form of programmed cell death triggered by copper, is directly related to the acylation of tricarboxylic acid (TCA) cycle components, resulting in toxic protein stress and cell death[9,10]. However, studies on cuproptosis in vascular smooth muscle cells are limited. This research aims to elucidate the molecular mechanisms of carotid artery stenosis and explore new preventive or therapeutic strategies.

In this study, we investigate the relationship between intimal hyperplasia and cuproptosis in rat carotid arteries using single-cell and bulk transcriptome sequencing for the first time. This approach enhances our understanding of the pathophysiological mechanisms underlying carotid intimal hyperplasia and offers new therapeutic strategies to inhibit the formation of carotid artery stenosis.

Materials and methods

Ethics declaration

This experiment was approved by the Animal Ethics Committee of Nanjing Hospital of Nanjing Medical University in accordance with institutional guidelines (approval number DWSY-23080536). The study complied with the ARRIVE guidelines, and all the methodology described in this section was performed according to the relevant guidelines and regulations.

Downloading and processing of dataset

The Gene Expression Omnibus (GEO) is a database created by NCBI[11] that stores a large amount of gene expression data. We obtained the dataset GSE164050 from GEO (https://www.ncbi.nlm.nih.gov/geo/), which includes four endothelial hyperplasia (Neointimal hyperplasia, Neo) and four normal arterial tissues (Control, Con), and organized this dataset for functional enrichment and modeling; and 6 Neo tissues and 6 Con tissues from GSE126627, collating this dataset for external validation dataset for diagnostic modeling. The single-cell dataset GSE174098[12]of carotid Neo specimens and normal control specimens was downloaded from the GEO database[11], which contains a total of 1 carotid NIH sample and 1 normal carotid tissue sample. Cuproptosis-associated genes (CAGs)[13] were obtained from the GeneCards database (https://www.genecards.org/), which collects and organizes CAGs, and a total of 26 copper mortality-related genes were obtained.

Differential expression gene function and pathway enrichment analysis

In the data processing and differential gene expression analysis, we employed R for various tasks such as dataset annotation, normalization, and log2 transformations[14,15]. From the GSE164050 dataset, we extracted the Differential Expressed Genes (DEGs)[16]. The "limma" software package was utilized to analyze the processed dataset, allowing us to perform differential gene analysis and filter the differentially expressed genes based on predefined criteria[17,18]. Specifically, we set the thresholds at $|\log2 \text{ fold change (FC)}| > 0.3$ and p-value < 0.05.

Following identifying DEGs, we conducted functional enrichment analyses using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG). The GO analysis provides a widely-used approach for large-scale functional enrichment of genes, categorizing them into biological processes (BP), molecular functions (MF), and cellular components (CC)[19]. On the other hand, KEGG is a comprehensive database storing valuable information about genomes, biological pathways, diseases, and drugs, making it an important resource for pathway enrichment analysis[20].

Machine learning and external dataset validation

We performed SVM-RFE and Random Forest (RF) feature selection on the expression values (TPM) of the 8 selected DEGs. We perform Random Forest (RF) feature selection on the dataset using the "Random Forest" R package. The RF algorithm is an integrated learning method that uses multiple decision trees to form regressions[20]. It is used to screen the importance of genes in diseases. Its main advantages are good performance in handling high dimensional data and insensitivity to outliers. We then perform Support Vector Machine-Recursive Feature Elimination (SVM-RFE) feature selection on the set of genes using the "e 1071" R package. SVM-RFE uses a binary classification method to score and rank the gene features and selects the top genes with the lowest error rate. The k parameter was set to 1, which corresponds to the standard SVM-RFE method where one feature is eliminated at a time. To enhance the efficiency of the feature elimination process, especially when dealing with a relatively small number of features, we set the halve above parameter to 2. We intersected the results of these two machine learning methods with the differentially expressed genes we previously screened and used the intersected genes as pivotal genes for further analysis. We selected six Neo samples and six Con samples from the dataset GSE126627 in GEO for external differential expression validation of hub genes[21–23].

We employed box plots to visualize the differentially expressed genes between the two groups.We selected the validated genes and analyzed them further.

Construction and evaluation of diagnostic models

Multivariate logistic regression modeling was performed for the two selected genes. A generalized linear model with binary classification was constructed using the GSE164050 dataset as a training set. The pROC package made ROC curves for model validation[24]. The ROC curves were plotted and the area under the curve (AUC) was calculated. External Neo dataset GSE126627 was downloaded via GEO to validate the model. The sensitivity and specificity of the model were calculated. The stability and reliability of the model were evaluated.

Molecular subtyping of carotid balloon strains

Consensus clustering is an algorithm that utilizes resampling techniques to identify cluster membership and subgroup assignments while validating the reliability of clustering results. In our study, we employed consensus clustering methods using the ConsensusClusterPlus package in R[25]. These methods consistently clustered gene expression profiles based on the previously identified key genes associated with cuproptosis. By selecting the optimal cluster from the consensus results, we were able to identify distinct patterns related to cuproptosis.

Immune infiltration analysis

CIBERSORT is an algorithm for deconvolution of the expression matrix of immune cell subtypes based on the principle of linear support vector regression, which uses RNA-Seq data to estimate immune cell infiltration in carotid balloon strain specimens[26]. We performed an immune infiltration analysis of the Neo group compared to the Con group in the GSE164050 dataset using the CIBERSORT algorithm to identify immune cells that were differentially enriched in the GSE164050 data between the two groups. We calculated Pearson correlation coefficients between the expression levels of cuproptosis-related genes and various immune cell populations.

Single-cell RNA statistical processing

In our study, we used the "Seurat" R package to create a Seurat object, followed by data quality control. We retained cells with less than 20% mitochondrial genes, less than 3% erythrocyte genes, more than 200 genes in total, expression ranging from 200 to 3000 genes, and genes expressed in at least three cells. The number of High Variable Genes (HVG) was set to 5000. The two samples were normalized by the "LogNormalize" method. The TSNE method was used to Dimensionality Reduction of the data, and the KNN method was used to cluster the cells with a resolution of 0.2. We used the "FindAllMarkers" function and the Wilcoxon rank sum test algorithm to identify marker genes for each cell cluster, with conditions including $|\log2\ FC|>0.25$, $p<0.05$, and the minimum percentage > 0.1. Cells were labeled by cell surface markers. To better identify vascular smooth muscle cell clusters, we selected clusters of vascular smooth muscle cells for further TSNE analysis, graph-based clustering, and marker gene analysis.

The cuproptosis gene was imported by the "AddModuleScore" function to obtain the cuproptosis score (CupScore) in each cell. The smooth muscle cells were divided into cup-high and cup-low groups according to the CupScore, and the differentially expressed genes (DEGs) were obtained and analyzed by GSVA[27], which is a kind of gene set analysis that calculates the variation scores of a specific set of genes in each sample based on the expression matrix. A database of expression matrices and gene sets with genes as rows can be input to GSVA, and a matrix of variant scores with gene set names as rows can be output[27]. We subsequently performed Single-Cell Regulatory Network Inference And Clustering (SCENIC)[28]on single-cell data. The SCENIC algorithm, published in Nature Methods in 2017, utilizes single-cell RNA-seq data and allows for simultaneous gene regulatory network reconstruction and cell state identification[28]. We performed a comprehensive analysis of intercellular communication molecules using a recently developed tool, the "CellChat" R software package. This tool allows in-depth analysis of intercellular communication by generating and mapping the probabilities and interaction strengths of intercellular communication from single-cell transcriptomic data. Seurat's normalized counts and cell types were used for this analysis.

Validation of a rat carotid balloon injury model

A carotid balloon straining intimal hyperplasia model was conducted on six rats to validate the expression levels of the two cuproptosis-associated genes (CAGs). Six Neo arteries and six Con arteries were collected for analysis.

Male SD rats weighing between 300 and 350 g (Vital River, Beijing, China) were used for the experiment. The rats were anesthetized by intraperitoneal injection of 2% sodium pentobarbital (55 mg/kg) and intravenously injected with sodium heparin (100 U/kg). The neck area was shaved and sterilized with 70% alcohol. The underlying glandular tissue was dissected from the skin, and the left central and adjacent musculature were bluntly dissected longitudinally. The left common carotid artery (CCA) was exposed, and the posterior blunt entrapment was directed towards the carotid artery bifurcation, left internal carotid artery (ICA), and external carotid artery (ECA). The distal ECA was ligated, and temporary occlusion of the CCA and ICA was achieved using arterial clips. A transverse arteriotomy was performed on the ECA, and a 1.5 F balloon catheter was inserted through the ECA into the CCA.

The balloon was slowly inflated to 4.0 atmospheres and then withdrawn by rotation. This process was repeated three times. After removing the catheter, the ECA was ligated near the bifurcation, and the arterial clamps in the ICA and CCA were released to restore normal blood flow. Hemostasis was ensured, and once normal blood flow was confirmed, the skin was closed. The intimal hyperplastic artery on the left side and the normal artery on the right side were sampled two weeks later for subsequent projection electron microscopy experiments and immunohistochemical experiments.

Methods of electron microscopy

The methods of electron microscopy in this study included the process of embedding-filming-photographing. The main methods were:

Sampling and fixation: Fresh rat carotid artery tissue is used to determine the sampling site, minimize mechanical damage such as pulling, contusion, and extrusion, and sampling within 1–3 minutes, sampling tissue 1 mm3 size. A petri dish with electron microscope fixative (2% glutaraldehyde) was prepared in advance before sampling, and the small tissue blocks were put into the petri dish immediately after being removed from the body, and then cut into $1mm^3$ small tissue blocks with a scalpel in the fixative of the petri dish. The cut small tissue blocks were then transferred to EP tubes containing new electron microscopy fixative to continue fixation, and fixed at 4 °C for preservation and transportation. We used 0.1 M phosphate buffer PB (PH7.4) rinsed 3 times, 15 min each time. then 1% osmium acid prepared with 0.1 M phosphate buffer PB (PH7.4) was used to fix at room temperature away from light for 2 h. 0.1 M phosphate buffer PB (PH7.4) was rinsed three times, 15 min each time. Room temperature dehydration: tissues were sequentially dehydrated into 30–50%–70-80%–95-100%–100% alcohol upstream for 20 min each time and 100% acetone twice for 15 min each time. Infiltration embedding: acetone: 812 embedding agent = 1: 1 37 ℃ 2–4 h, acetone: 812 embedding agent = 1: 2 37 ℃ infiltration overnight, pure 812 embedding agent 37 ℃ 5–8 h. Pure 812 embedding agents poured into the embedding plate, the samples will be inserted into the embedding plate and then 37 ℃ oven overnight. Polymerization: Embedding plate in 60 ℃ oven polymerization 48 h, remove the resin block spare. Ultrathin sectioning: the resin block should be sliced at 60–80 nm in an ultrathin slicing machine, and the slices should be fished by 150 mesh copper mesh of Fanghua film. Staining: the copper mesh was stained with 2% uranyl acetate saturated alcohol solution for 8 min, washed with 70% alcohol for 3 times, washed with ultrapure water for 3 times, stained with 2.6% lead citrate solution for 8 min, washed with ultrapure water for 3 times, and then dried with filter paper. The copper mesh sections were put into the copper mesh box and dried at room temperature overnight. Observe under the transmission electron microscope and collect images for analysis.

Methods of immunohistochemistry

In this study, immunohistochemical labeling was performed using Pdhx and Fdx1 monoclonal antibodies to assess specific markers. The Pdhx antibody (provided by GeneTex, Inc., North America) was preserved in PBS with 50% Glycerol and applied at a concentration of 1:100 for immunohistochemistry. The Fdx1 recombinant monoclonal antibody (JE63-56), provided by Thermofisher scientific company, was preserved in TBS (pH 7.4) with 0.05% BSA and 40% glycerol. It was applied at a concentration of 1:200 for immunohistochemistry.

Paraffin-embedded rat carotid artery tissues were processed for staining and microscopic examination after relevant treatments. The main procedures included deparaffinization, antigen retrieval, treatment with 3% hydrogen peroxide, circle drawing, serum blocking, overnight incubation with primary antibodies at 4℃, incubation with secondary antibodies at room temperature, color development, hematoxylin staining of nuclei, dehydration, transparency, sealing, and subsequent microscopic examination. Relevant indicators were also screened for quantitative analysis.

Statistical analysis

All data calculations and statistical analyses were performed using R programming (https://www.r-project.org/, version 4.0.2). For the comparison of continuous variables between the two groups, the statistical significance of normally distributed variables was estimated by independent Student t-test, and differences between non-normally distributed variables were analyzed by Mann-Whitney U-test (i.e. Wilcoxon rank sum test). All statistical p-values were two-sided, with $P < 0.05$ considered statistically significant.

Results

Analysis of differential gene expression

The research process in this study is illustrated in Fig. 1. To identify differentially expressed genes (DEGs) in the dataset, we utilized the "Limma" R software package. The screening criteria were set as $|log2fc| > 0.3$ and p-value < 0.05. A total of 4,194 differentially expressed genes (DEGs) were identified and were shown by volcano plots (Fig. 2B). We identified 12 genes by intersecting the 4,194 DEGs with our a priori list of 26 cuproptosis-related genes (CRGs) for further analysis. The 12 genes were show by Box-and-line plots (Fig. 2A). Out of the 26 CRGs, six genes exhibited significant differential expression, including four down-regulated genes and two up-regulated genes. Additionally, Fig. 2C displays the correlation between these 12 genes based on the correlation coefficients.

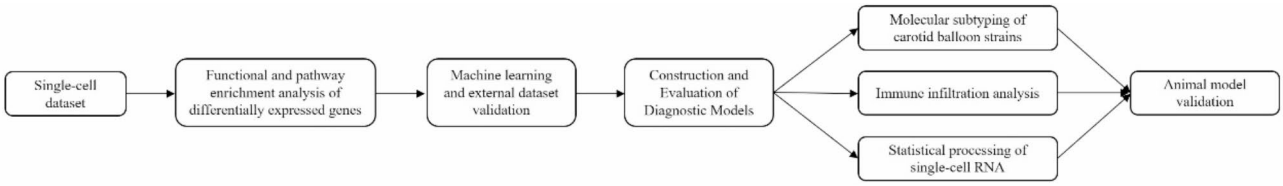

Fig. 1. Flowchart.

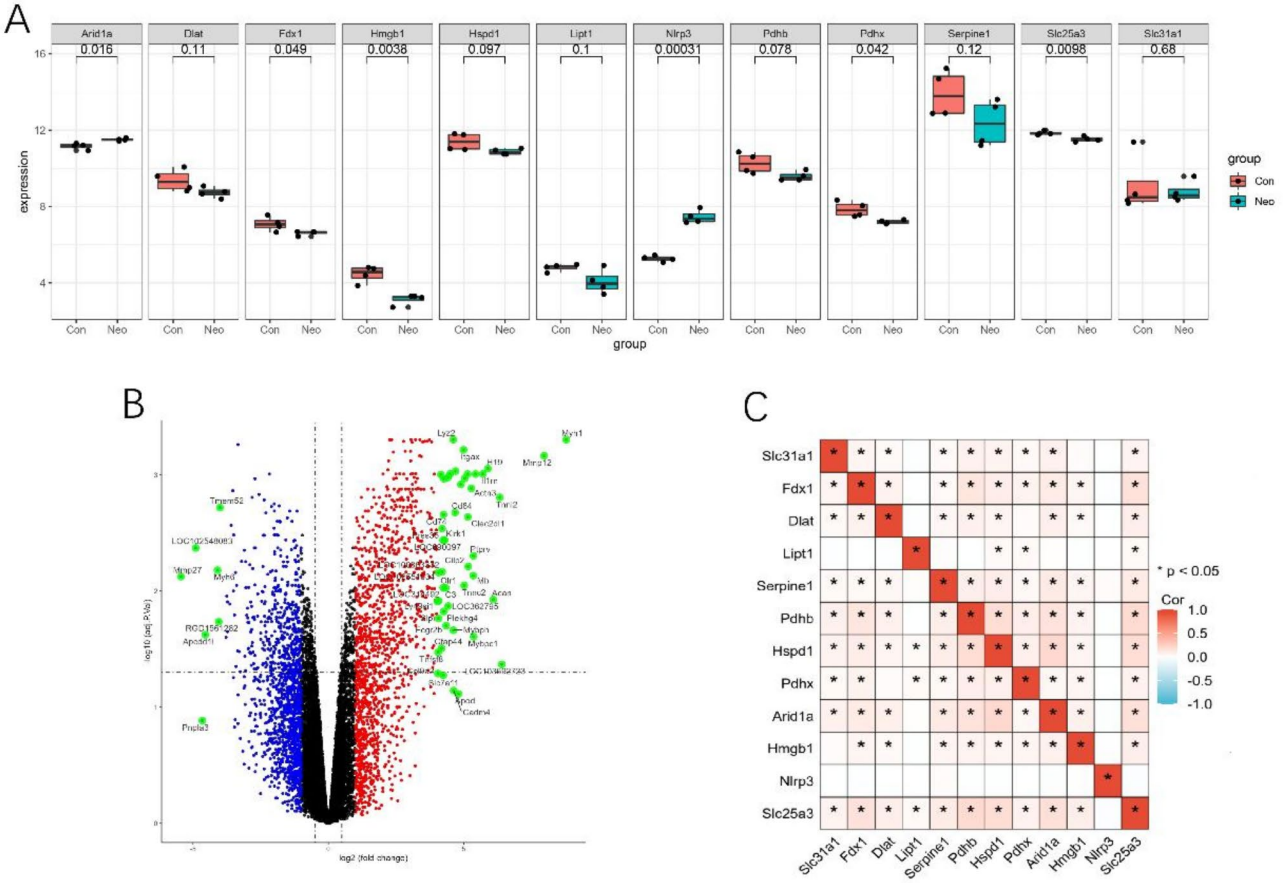

Fig. 2. Differential gene expression analysis. Box-and-line and volcano plots of differentially expressed genes (**A**, **B**), and correlations between differentially expressed genes (**C**).

Functional enrichment analysis of differentially expressed genes

To analyze the relationship between the differentially expressed genes associated with carotid balloon strain and biological processes, molecular functions, cellular components, biological pathways, and diseases, we first carried out a functional enrichment analysis of the differentially expressed genes associated with carotid balloon strain. The differentially expressed genes associated with carotid balloon strain were mainly enriched in NADH dehydrogenase complex assembly, mitochondrial respiratory chain complex I assembly, negative regulation of immune system process, immune response-regulating signaling pathway, regulation of inflammatory response, and other biological processes (Fig. 3A). It is also enriched in oxidoreductase complex, mitochondrial protein-containing complex, respiratory chain complex, mitochondrial respirasome, mitochondrial matrix (Fig. 3B), and other cellular components, as well as in GTPase regulator activity, nucleoside-triphosphatase regulator activity, sulfur compound binding, extracellular matrix structural constituent, collagen binding, electron transfer activity, and other molecular functions were also expressed (Fig. 3C). We next performed pathway enrichment analysis of DEGs, which showed that DEGs were enriched in Oxidative phosphorylation, ECM-receptor interaction, PI3K-Akt signaling pathway, Rap1 signaling pathway, AGE-RAGE signaling pathway in diabetic complications, and other biological pathways (Fig. 3D).

Machine learning-based central gene selection

Machine learning methods are widely used in vascular diseases. SVM-RFE and RF machine learning methods were used to further screen candidate hub genes. First, the RF algorithm was used to rank the importance of each gene (Fig. 4A), and the top 8 genes with the highest ranked importance were selected: Hmgb1, Pdhx, Fdx1, Nlrp3, Slc25a3, Arid1a, Lipt1, and Pdhb. Then, the top 6 genes with the highest accuracy in the SVM-RFE results were selected and intersected with the cuproptosis related genes to take the intersection to obtain 4 CAGs, ("Pdhx", "Hmgb1", "Slc25a3", "Fdx1") after machine learning intersection (Fig. 4B), in verified with the external dataset GSE126627 (Fig. 4C). CAGs ("Pdhx", "Fdx1") were included in further studies.

Animal experiments to validate CAGs

The carotid vessels of the rats were dissected in two groups on the 14th postoperative days, and the rats were sampled and executed at the same time. The right common carotid artery specimen was collected for further analysis. Pathological sections of the common carotid artery were prepared for H&E staining,

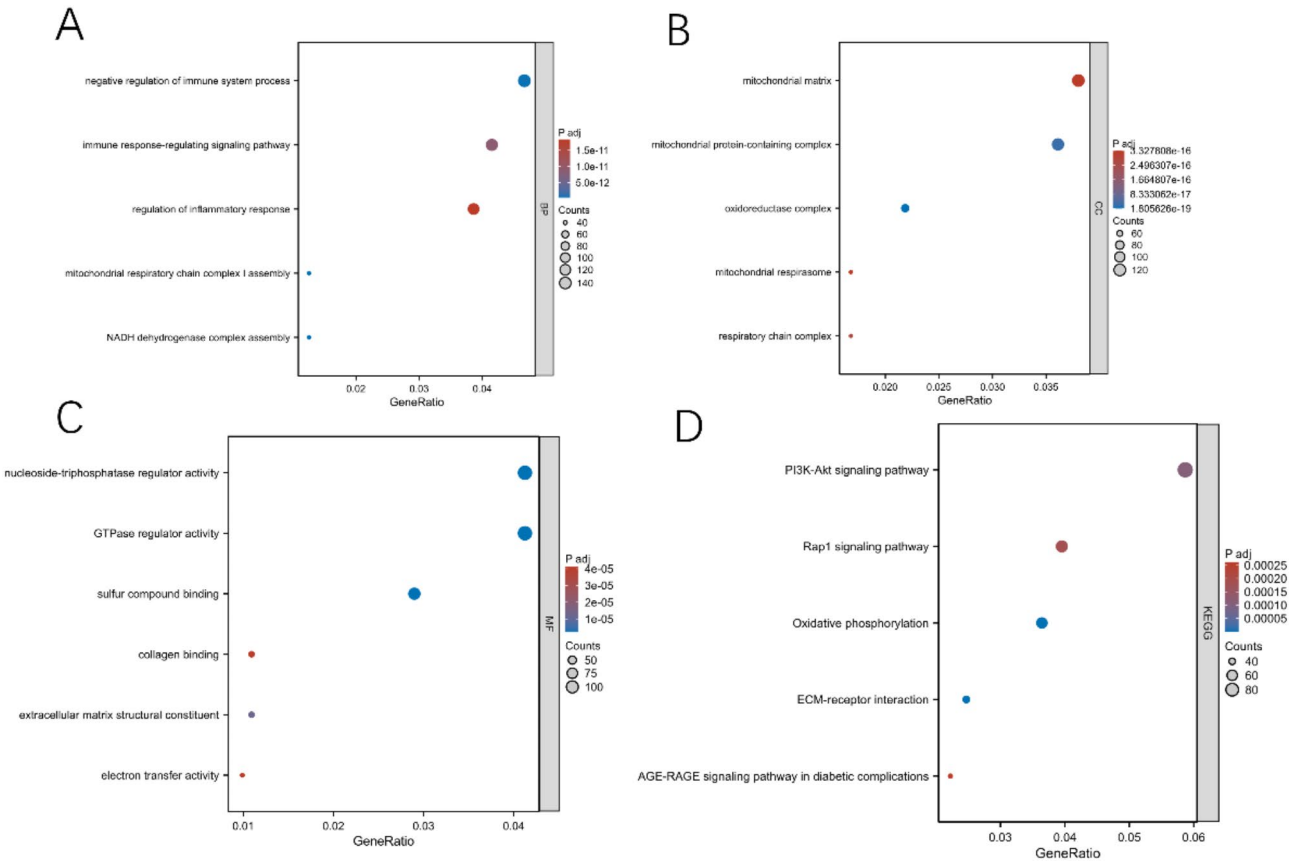

Fig. 3. Functional enrichment analysis of differentially expressed genes. (**A**, **B**) Biological processes underlying the major enrichment of differentially expressed genes associated with carotid balloon strains. (**C**) Expression of relevant differentially expressed genes in molecular functions. (**D**) KEGG analysis.

immunohistochemical analysis, and transmission electron microscopy. The sections covered a range of 15 mm proximal to the heart, starting from the bifurcation of the common carotid artery. At the same time, we retained the left carotid artery at the same site, and H&E staining was performed to examine the cross-section of the carotid artery. We measured the diameter of the common carotid artery using CaseViewer 2.4 software (3DHISTECH Ltd., Budapest, Hungary). We observed and analyzed the expression and distribution of Pdhx and Fdx1 in the common carotid artery to assess the expression in the Neo and Con groups. Among them, Fig. 5A1 and A2 shows the expression in Pdhx and Fdx1 in the Con group, and Fig. 5B1 and B2 shows the expression in Pdhx and Fdx1 in the Neo group. Immunohistochemical images of carotid arteries were analyzed using computerized quantitative image analysis software Image-Pro Plus 6 (Media Cybernetics Inc., Rockville, MD, United States). We measured and analyzed mean integrated optical density (IOD) to quantify the level of cellular value added. In the Neo group, we can see more intimal hyperplasia of vascular smooth muscle. Also in the Neo group, we could observe more narrowing of the carotid lumen. Pdhx and Fdx1 had significantly lower mean IOD values in the Neo group than in the Con group (Fig. 5C). Immunohistochemistry suggested reduced expression of the Neo group ("Pdhx", "Fdx1"). Low expression of Fdx1 and Pdhx reduces the occurrence of cuproptosis, and we used electron microscopy to observe mitochondria to verify this conjecture. Figure 5D and E show the characteristic expression of mitochondria in transmission electron microscopy in the Con and Neo groups. As critical genes for cuproptosis, Fdx1 and Pdhx also have effects on mitochondrial morphology. Electron microscopy observations revealed distinct mitochondrial changes. In the Con group, characteristic manifestations of cuproptosis, including mitochondrial swelling and rupture of mitochondrial organelle membranes, were frequently observed. In contrast, such cellular mitochondria with cuproptosis characteristics were less commonly observed in the Neo group.

Establishment and evaluation of diagnostic models

Establishment and evaluation of diagnostic models

First, we plotted the ROC curves of the 2 CAGs separately using the dataset GSE164050 to evaluate each gene's diagnostic specificity and sensitivity to calculate the area under the curve (AUC). ("Pdhx",) (AUC = 1.00), "Fdx1" (AUC = 0.812) were obtained (Fig. 6A, B). We performed multivariate logistic regression modeling for the 2 hub genes and evaluated the model using ROC curves (Fig. 6C). The calculated AUC was 1. We then validated the diagnostic model with the dataset GSE126627 by plotting the ROC curve with AUC = 1, (Fig. 6D). These results indicate that the diagnostic model has a high diagnostic value.

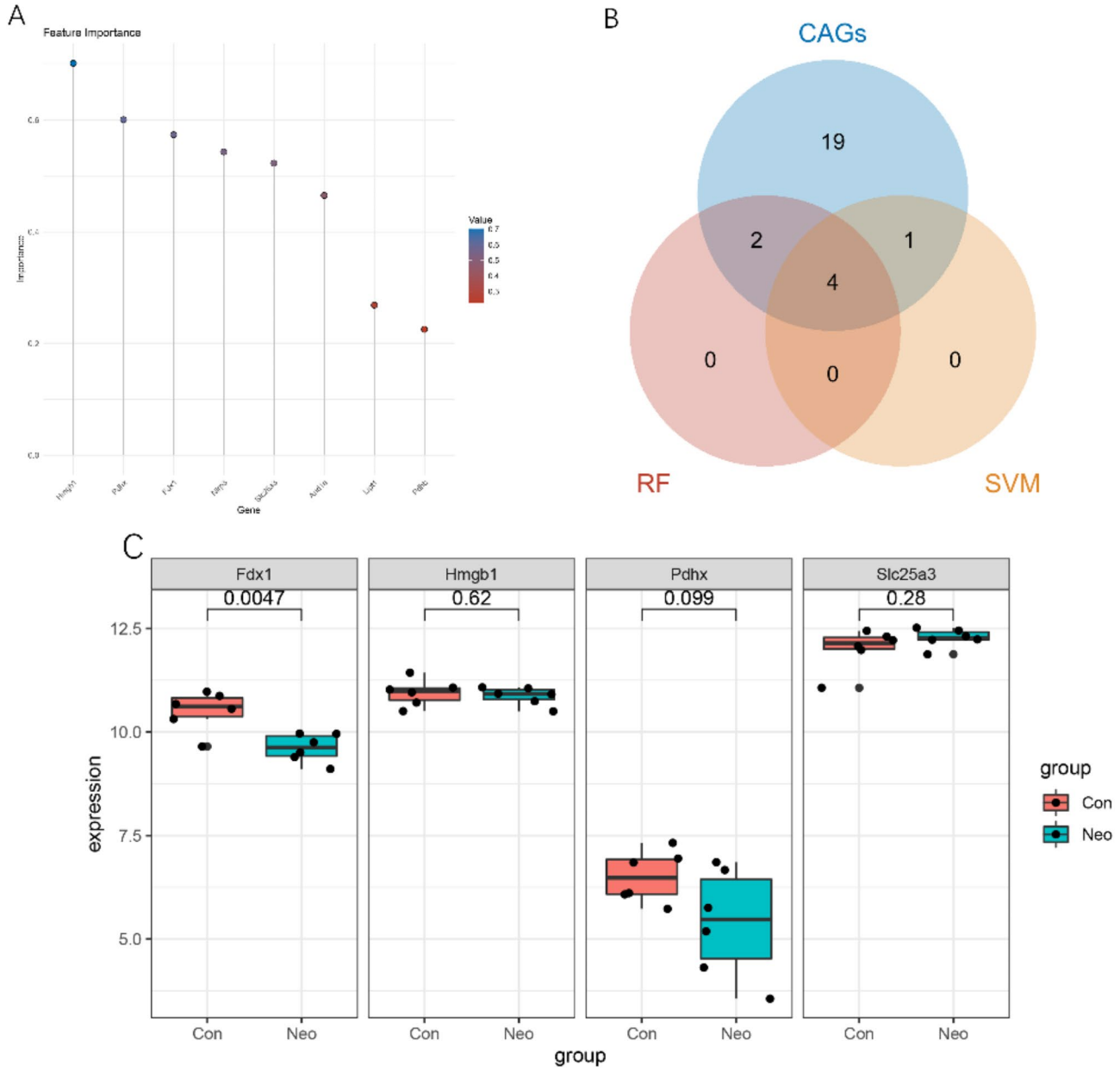

Fig. 4. Machine learning-based selection of hub genes. The importance of each gene was ranked using the RF algorithm (**A**). The gene positions were selected ("Pdhx", "Hmgb1", "Slc25a3", "Fdx1") after machine learning intersection (**B**), validated with the external dataset GSE1266274C (**C**).

Two different cuproptosis patterns identified by characterized genes

Using the "ConsensusClusterPlus" package in R software, we performed cluster analysis based on eight cuproptosis apoptosis-related genes using the consistency clustering method and plotted the consistency clusters at k = 2 and k = 3 (Fig. 7A, B). Relative change of the area under the CDF curves from k 2 to 5 (Fig. 7C) as well as the cumulative distribution function (Fig. 7D) and tracking plot (Fig. 7E) of the consistency clustering were plotted.

Immune infiltration analysis

To analyze the degree of immune infiltration in carotid balloon strain specimens, we calculated the degree of infiltration of 22 immune cells in carotid balloon strain specimens and normal specimens using the cibersort algorithm in GSE164050 data (Fig. 8A). There were 2 types of immune cells in the GSE164050 dataset that were significantly different between the con and neo groups when using the Wilcox test algorithm (Fig. 8B): dendritic cells resting, and mast cells resting. we analyzed the correlation between the expression levels of the 2 cuproptosis-related genes and the 22 types of immune cells. correlations, the correlation coefficient between

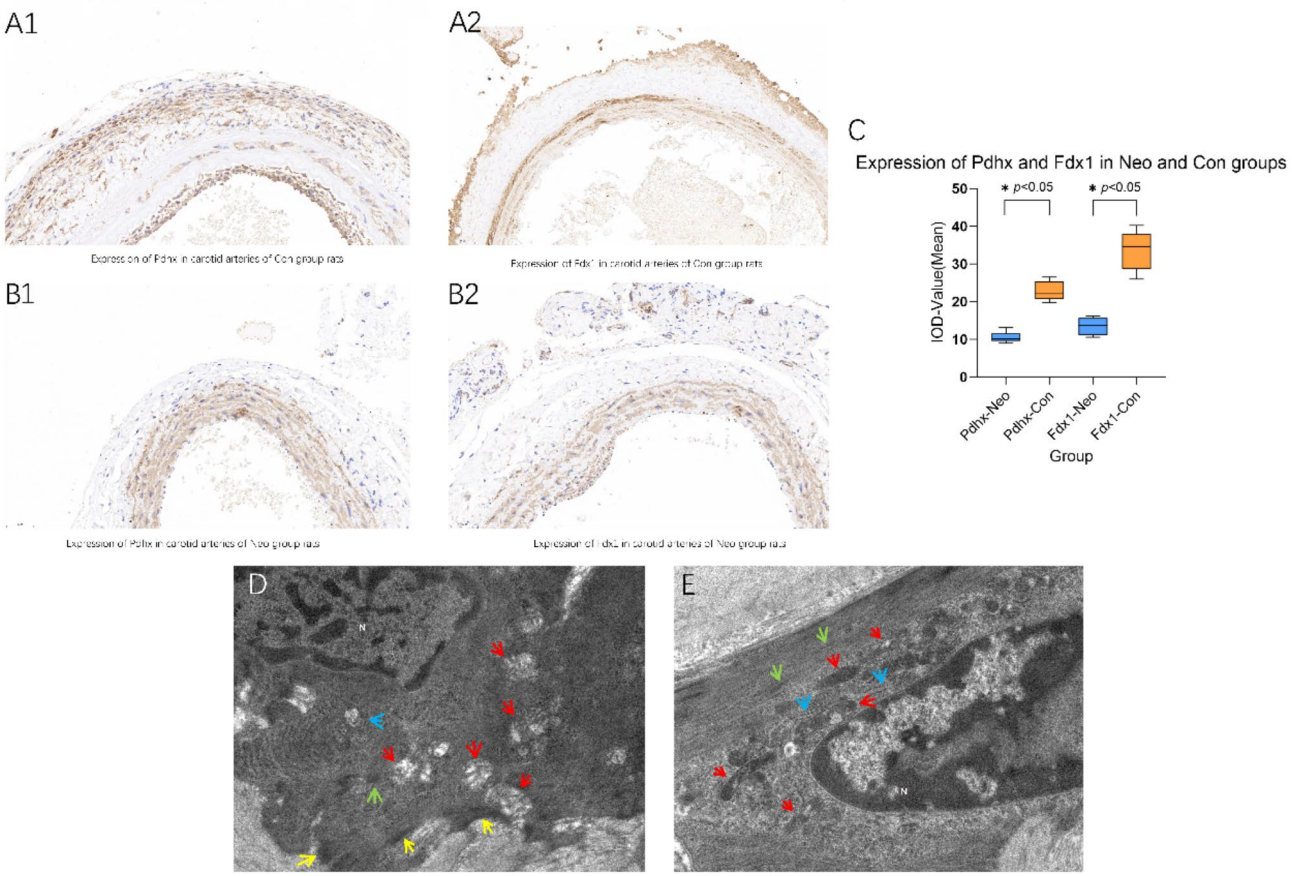

Fig. 5. Immunohistochemical results and electron microscopy analysis. Expression in Pdhx and Fdx1 in the Con group (**A1** and **A2**), Pdhx and Fdx1 in the Neo group (**B1** and **B2**). Pdhx and Fdx1 had significantly lower mean IOD values in the Neo group than in the Con group (**C**). Characteristic mitochondrial expression in transmission electron microscopy in the Con and Neo groups (**D** and **E**).

PDHX and T cells gamma delta was greater than 0.5, and the *p*-value of each correlation coefficient was greater than 0.05, which was not statistically significant (Fig. 8C).

Single-cell data analysis

We analyzed NEO's single-cell sequencing dataset GSE174098, integrating different samples. After filtering, 7934 cells were retained, including 2491 NEO tissue cells and 5443 normal tissue cells. The expression characteristics of each set of samples are shown in Fig. 9A-D. The nCount_RNA representing the number of unique molecular identifiers (UMI) was positively correlated with the nFeature_RNA representing the number of genes, and cells with less than 20% of mitochondrial genes and 3% of erythrocyte genes were retained (Fig. 9E-G). The top 10 most variable HVGs were identified (Fig. 9H). PCA grouping showed no significant batch effect in both groups (Fig. 9I). A heat map was used to show the top 5 marker genes for each cell cluster (Fig. 9J).

We clustered all cells into 16 clusters using the KNN (k-Nearest Neighbor) clustering algorithm, and then annotated the cell type of each cell cluster according to the surface marker genes of different cell types and labeled the different subgroups of the cells, which were shown by TSNE plots (Fig. 10A, B). The VSMCs cells were extracted for further analysis, and the heatmap showed the top10 markers of each cluster (Fig. 10C), and the VSMCs were displayed with TSNE downscaling plots (Fig. 10D), and the TSNE plots were displayed (Fig. 10D).

In this study, we incorporated the cuproptosis gene using the "AddModuleScore" function. This allowed us to calculate the cuproptosis score (CupScore) for each vascular smooth muscle cell (VSMC). Based on the median value of the CupScore, we divided the VSMCs into two groups: cup-high and cup-low groups, GSVA analysis was performed between the smooth muscle cells in the cup high and cup low groups, and the heatmap showed that the signaling pathways with statistically significant differences in GSVA analysis included PI3K AKT MTOR SIGNALING, IL6 JAK STAT3 SIGNALING and _GLYCOLYSIS (Fig. 11A). We demonstrated with bubble plots that the cellular interaction signaling pathways of smooth muscle cells, macrophages and endothelial cells analyzed by SCENIC in the cup-high and cup-low groups included GAS6 - AXL, PDGFA - PDGFRA, PDGFA - PDGFRB, SPP1 - (ITGA5 + ITGB1), SPP1 - CD44, TNFSF12 - TNFRSF12A, VEGFA - VEGFR1 (Fig. 11B-D).

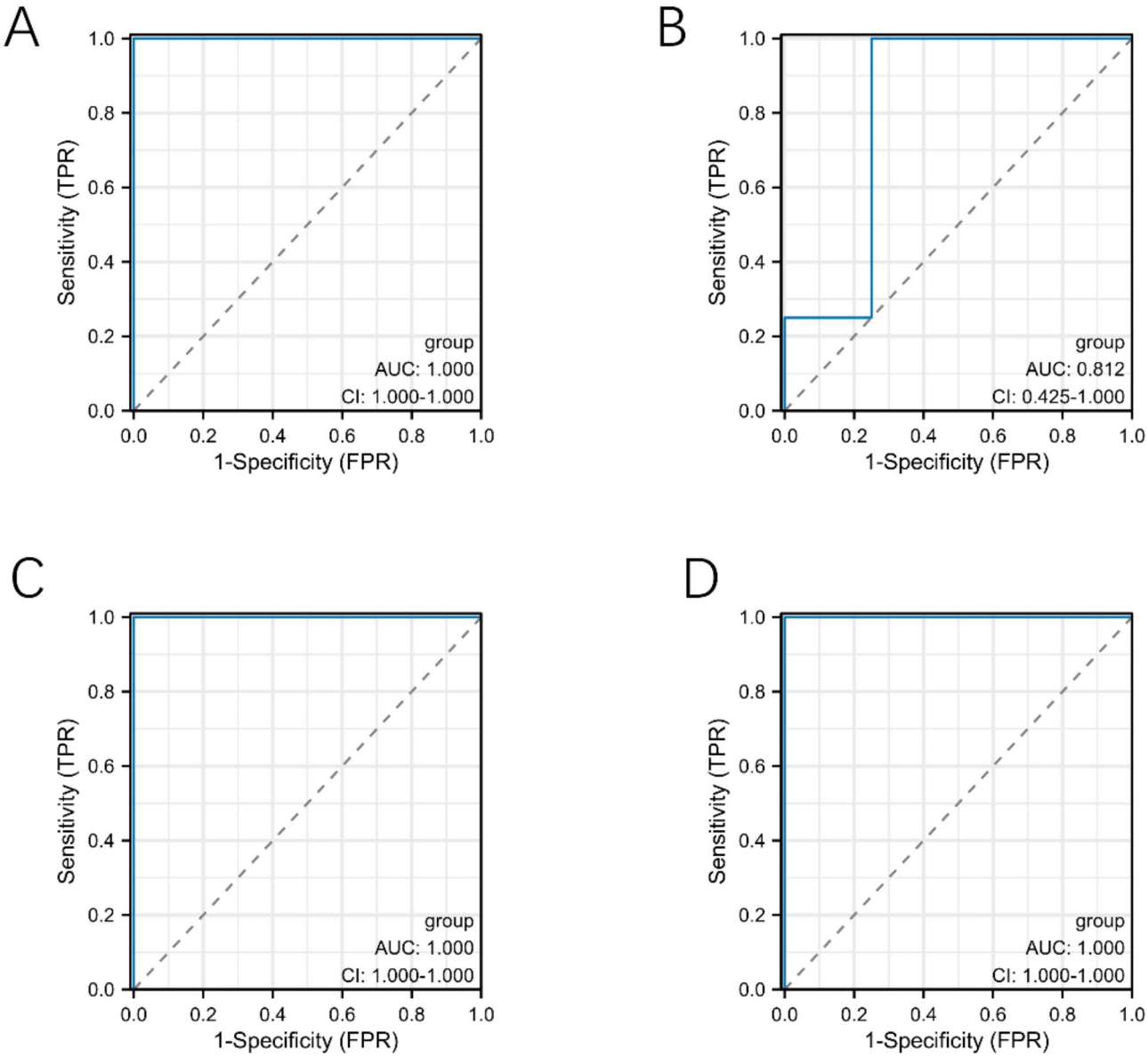

Fig. 6. Diagnostic modeling and evaluation. The diagnostic specificity and sensitivity of each gene were assessed by calculating the area under the curve (AUC). ("Pdhx") (AUC = 1.00), "Fdx1" (AUC = 0.812) were obtained (**A**, **B**). The model was evaluated using ROC curves after multivariate logistic regression modeling (**C**). The calculated AUC was 1. The dataset GSE126627 validated this diagnostic model by plotting the ROC curve with AUC = 1, (**D**).

Discussion

Carotid artery stenosis, characterized by excessive proliferation of vascular smooth muscle cells, involves complex pathophysiological processes. Treatment of carotid artery stenosis usually includes pharmacological therapy, endoluminal angioplasty, and vascular bypass reconstruction[5,6]. However, both pharmacologic and surgical interventions are physical methods to open stenotic vessels and achieve hemodynamic remodeling. There is still a lack of understanding of the pathophysiological mechanisms of VSMCs proliferation. Therefore, early diagnosis and targeted intervention of vascular smooth muscle cell overproliferation are important, and identifying novel biomarkers to inhibit smooth muscle cell proliferation is crucial.

Copper can cause cell death by directly binding to the lipid-acylated components of the tricarboxylic acid (TCA) cycle and increasing hydrotoxic stress on proteins, leading to aggregation of lipid-acylated proteins and loss of iron-sulfur (Fe-S) cluster proteins, and ultimately to cell death. Cuproptosis, as a new form of death, is different from known forms of death such as necrosis, apoptosis, pyroptosis, and iron death. The processes involved in copper play an important role in the development of lungs, oral mucosa, and myocardium[29,30]. While the effects of cuproptosis on vascular smooth muscle cell proliferation remain underexplored, our study provides insights into its role through the analysis of the single-cell dataset GSE174098.

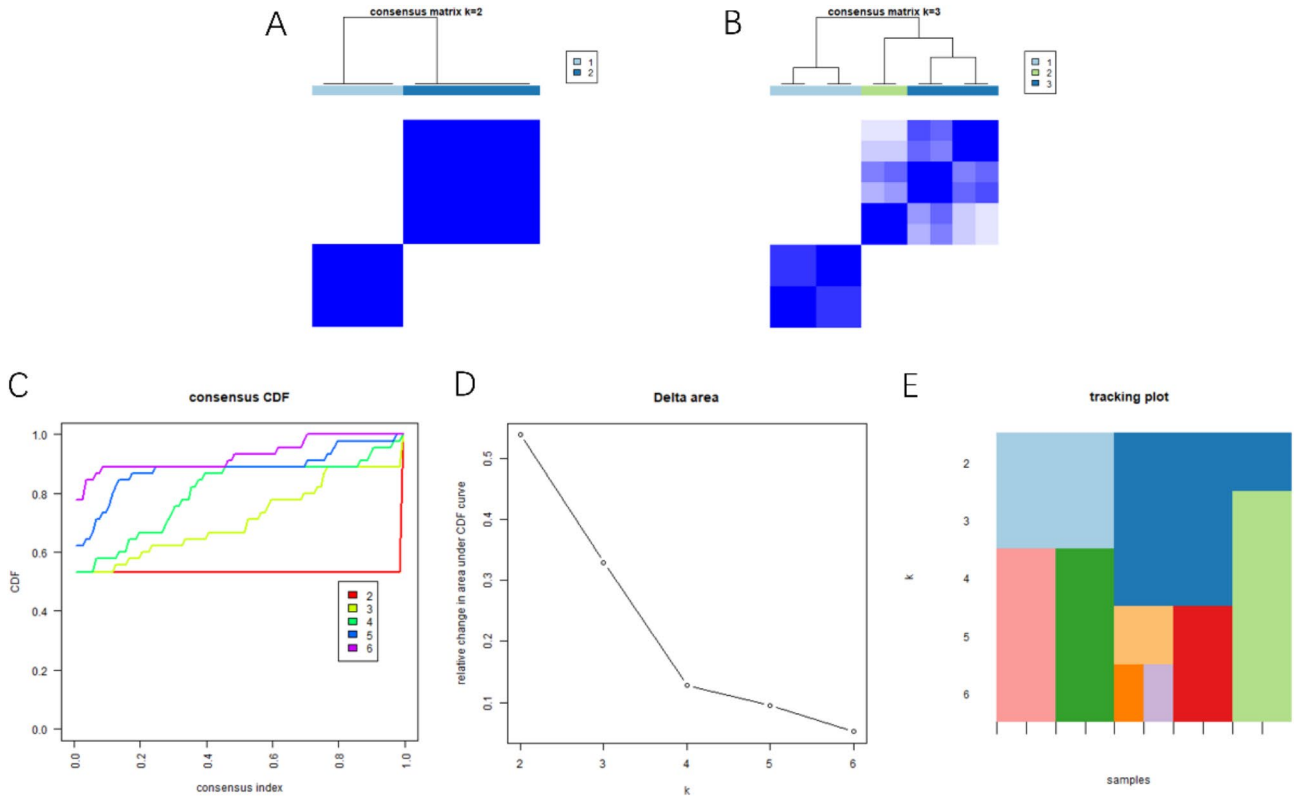

Fig. 7. Two different patterns of cuproptosis identified by characteristic genes Consistent clustering method for cluster analysis, plotting consistent clusters at k = 2 and k = 3 (**A**, **B**), relative change in area under the CDF curve from k of 2 to 5 (**C**), cumulative distribution function of consistent clusters (**D**) and tracking plot (**E**).

We imported the cuproptosis gene through the "AddModuleScore" function and obtained the cuproptosis score (CupScore) in each cell. We found that VSMCs from the Con group exhibited higher cuproptosis scores than those from the Neo group, suggesting that cuproptosis may mitigate intimal hyperplasia. Our findings indicate that low expression levels of Pdhx and Fdx1 in the Neo group may inhibit cuproptosis, promoting vascular smooth muscle cell proliferation.

The Pdhx gene encodes E3-binding proteins critical for the pyruvate dehydrogenase complex, which plays a role in mitochondrial function and cellular metabolism. Previous studies have linked Pdhx to cuproptosis and metabolic pathways. Similarly, Fdx1 is involved in electron transfer and lipidation processes, with limited studies examining its role in cuproptosis.

The pyruvate dehydrogenase (PDH) complex is a vital component located within the mitochondrial matrix. The gene Pdhx plays a crucial role in encoding proteins known as E3-binding proteins. These proteins are essential constituents of the larger pyruvate dehydrogenase complex. Comprising various enzymes and other associated proteins, this complex includes the prominent enzyme E3. E3-binding proteins attach E3 to the complex and provide the correct structure for the complex to function.

In recent studies, the association between Pdhx and cuproptosis has been explored. For instance, in a clinical investigation on esophageal cancer tumorigenesis and development, Pdhx has been found to be linked to heat stress and cuproptosis. It has been implicated in various cellular processes, including cell survival, proliferation, migration, metabolism, and immunosuppression. Furthermore, differential gene expression analysis revealed that patients with high and low Pdhx expression exhibited distinct metabolic and immune pathway enrichments. These pathways included fatty acid metabolism, glutathione metabolism, NK cytotoxicity, and T cell receptor signaling. Notably, elevated Pdhx expression has been associated with abnormal cellular metabolism and compromised immune responses, leading to inhibition of immune cell proliferation and promoting tumor progression[31].

The Fdx1 gene encodes a small iron-sulfur protein, which transfers electrons from NADPH to the mitochondrial cytochrome P450 through the enzyme ferric oxidoreductase. Fdx1 serves as a key factor in enhancing protein lipidation processes. Additionally, it is associated with the induction of cell death triggered by copper[9]. The expression profile of genes related to Fdx1 is linked to the enrichment of genes involved in the tricarboxylic acid (TCA) cycle, NOTCH pathway, and other related pathways. Despite being a core regulatory gene in cuproptosis, limited studies have been conducted on Fdx1, and its regulation in tumor biology remains unclear. However, some studies have shown that knockdown of the Fdx1 gene leads to the downregulation of cuproptosis in renal tumors, subsequently promoting tumor proliferation and development[32,33].

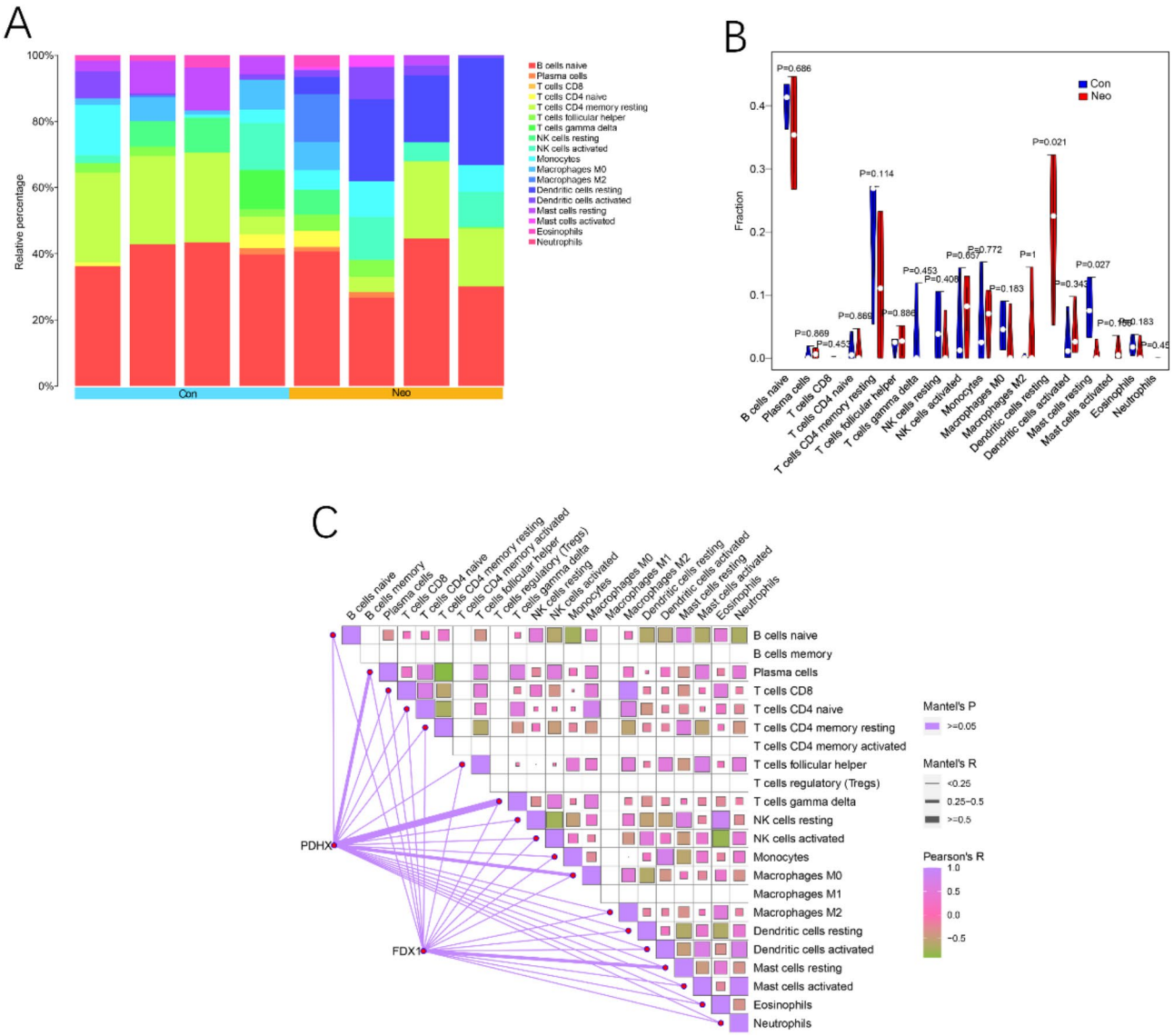

Fig. 8. Immune infiltration analysis Degree of infiltration of 22 immune cells in carotid balloon strain specimens and normal specimens (**A**). 2 types of immune cells were significantly different between the Con group and the Neo group (**B**). Correlation between the expression levels of the 2 cuproptosis-related genes and the 22 types of immune cells (**C**).

In our study, we conducted differential expression gene function and pathway enrichment analyses, as well as constructed diagnostic models, to investigate the role of Pdhx and Fdx1 genes. Validation was performed through animal experiments. Our findings revealed that the Neo group exhibited lower expression levels of both Pdhx and Fdx1 genes. Electron microscopy analysis further indicated that mitochondria in the cuproptosis, characterized by swollen mitochondria, intact membranes, vacuolated intra-membrane matrix, and reduced or disappeared cristae, were less prevalent in the Neo group compared to the Con group. Immunohistochemistry results suggested that the Neo group displayed enhanced vascular smooth muscle proliferation, potentially attributed to the decreased expression of Pdhx and Fdx1.

Although our study yielded valuable insights into the relationship between Pdhx and Fdx1 and carotid smooth muscle cell proliferation, certain limitations still need to be recognized. Firstly, the sample size available in public databases was relatively small, which may introduce statistical errors and limit the generalizability of our results. Moreover, the validation of our animal model was limited, emphasizing the need for additional animal experiments to validate the roles of these candidate genes at a later stage. Furthermore, further investigation is warranted to elucidate the molecular mechanisms through which these relevant candidate genes modulate the effects of cuproptosis on the formation and development of carotid intimal hyperplasia.

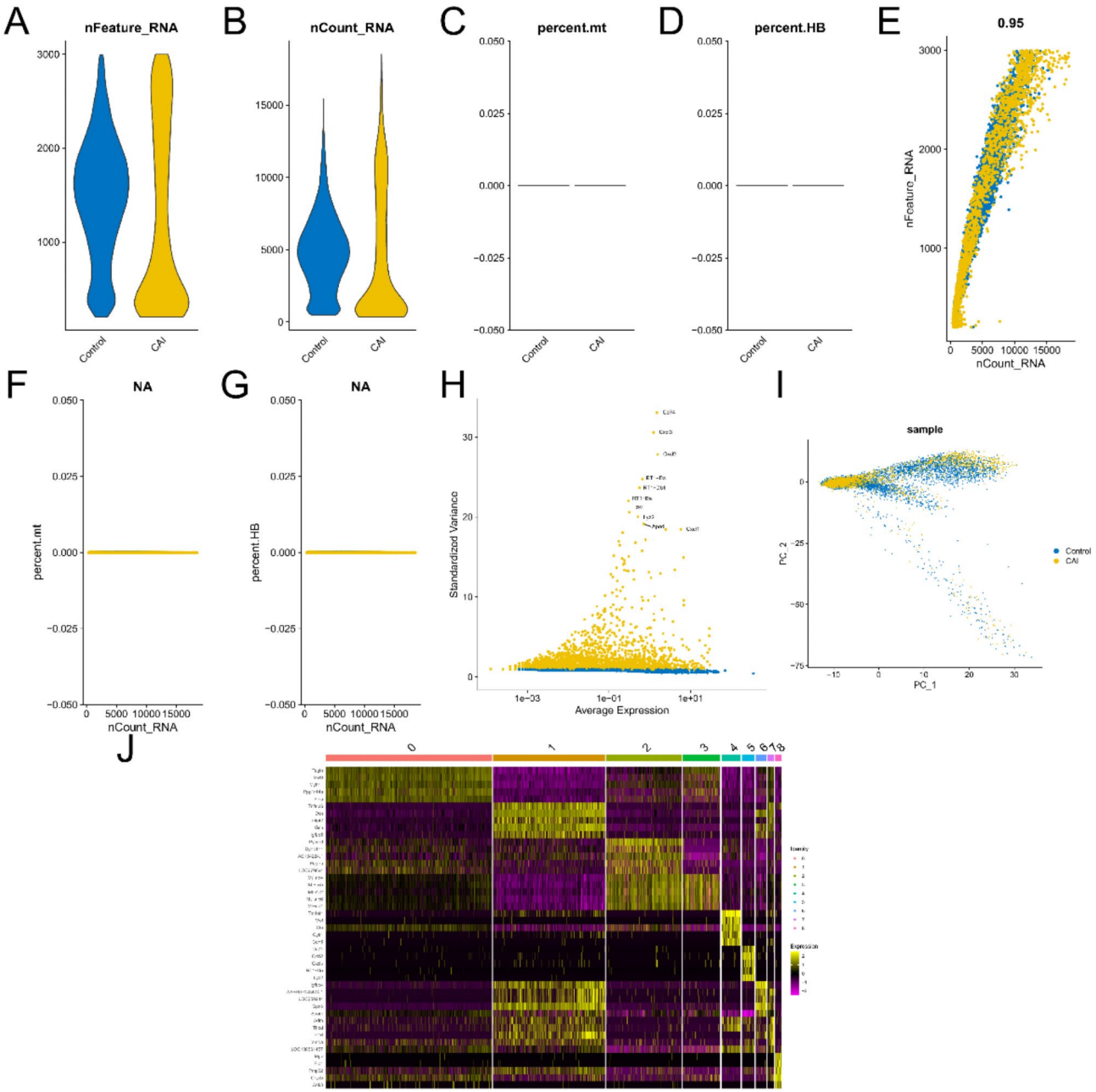

Fig. 9. Single-cell data analysis. The expression characteristics of each group of samples are shown in (**A**–**D**). The nCount_RNA representing the number of unique molecular identifiers (UMI) was positively correlated with the nFeature_RNA representing the number of genes, retaining cells with less than 20% of mitochondrial genes and 3% of erythrocyte genes (**E**–**G**). The top 10 most variable HVGs (**H**). The PCA grouping showed no significant batch effect for the two groups (**I**). The heatmap demonstrated the top 5 marker genes for each cell cluster (**J**).

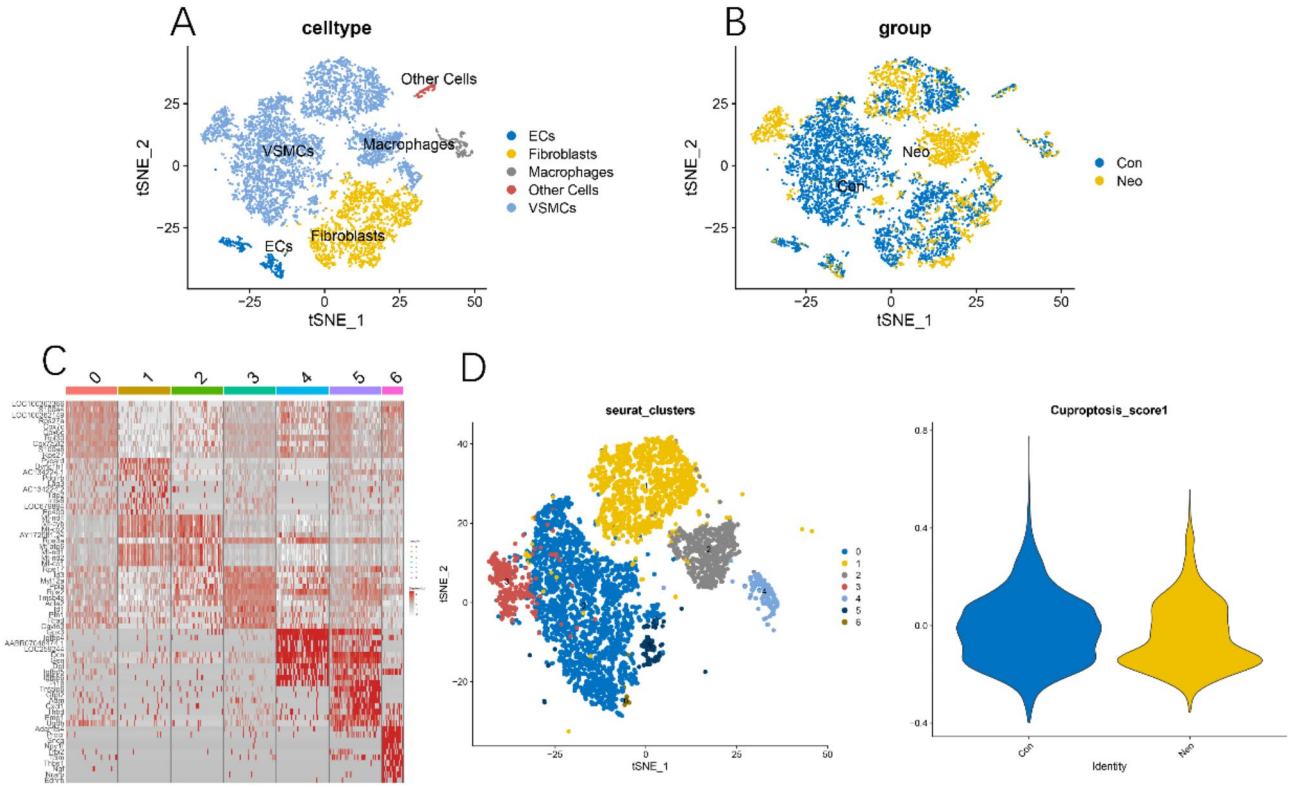

Fig. 10. Analysis of different subgroups of cells. cell type of each cell cluster was annotated according to the surface marker genes of the different cell types, and different subgroups of cells were labeled and shown in TSNE plots (**A**, **B**). The heatmap demonstrated the 10 markers of each cluster (**C**), and the VSMCs were displayed with TSNE plots (**D**).

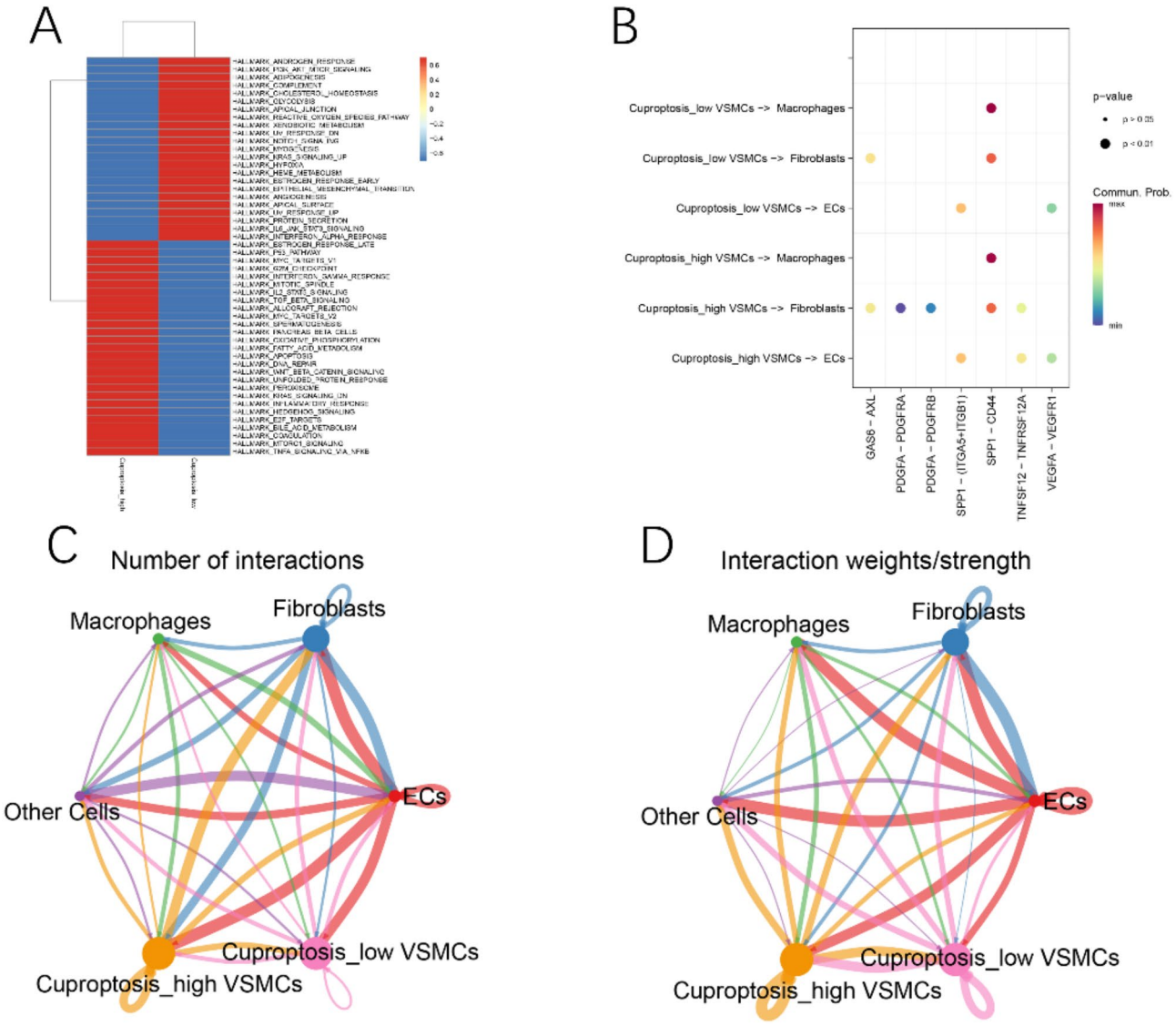

Fig. 11. Cuproptosis score and GSVA analysis. Heatmaps demonstrating the signaling pathways where GSVA analysis showed statistically significant differences included (**A**). Bubble plots demonstrating cellular interaction signaling pathways of smooth muscle cells with macrophages, endothelial cells, and fibroblasts in the cup-high and cup-low groups of SCENIC analysis (**B**–**D**).

Data availability

The dataset supporting the conclusions of this article is available in the GSE164050 repository, https://www.ncbi.nlm.nih.gov/geo/.

Received: 2 October 2024; Accepted: 5 February 2025
Published online: 12 February 2025

References

1. Go, A. S. et al. Heart disease and stroke statistics—2014 update: a report from the American Heart Association. *Circulation* **129**, e28–e292 (2014).
2. Lian, N. & Li, T. Growth factor pathways in hypertrophic scars: molecular pathogenesis and therapeutic implications. *BioMed. Pharmacother.* **84**, 42–50 (2016).
3. Furie, K. L. et al. Guidelines for the prevention of stroke in patients with stroke or transient ischemic attack: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* **42**, 227–276 (2011).
4. Brott, T. G. et al. Long-term results of stenting versus endarterectomy for carotid-artery stenosis [J]. *N Engl. J. Med.* **374**, 1021–1031 (2016).
5. Cronenwett, J. L. & Johnston, W. *Rutherford's Vascular Surgery* 8th edn (Saunders, 2014).
6. Furie, K. L. et al. Guidelines for the prevention of stroke in patients with stroke or transient ischemic attack: a guideline for healthcare professionals from the American heart association / American stroke association [J]. *Stroke* **42**, 227–276 (2011).

7. Tsuruta, W. et al. Simple new method for making a rat carotid artery post-angioplasty stenosis model. *Neurol. Med. Chir. (Tokyo)* **47**, 525–529 (2007).
8. Hinz, B. & Lagares, D. Evasion of apoptosis by myofibroblasts: a hallmark of fibrotic diseases. *Nat. Rev. Rheumatol.* **16**, 11–31. https://doi.org/10.1038/s41584-019-0324-5 (2020).
9. Tsvetkov, P. et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science* **375**, 1254–1261. https://doi.org/10.1126/science.abf0529 (2022).
10. Tang, D., Chen, X. & Kroemer, G. Cuproptosis: a copper-triggered modality of mitochondrial cell death. *Cell. Res.* **32** (5), 417–418 (2022).
11. Barrett, T. et al. NCBI GEO: mining tens of millions of expression profiles–database and tools update. *Nucleic Acids Res.* **35** (Database issue), D760–D765 (2007).
12. Gao, X. F. et al. Single-cell RNA sequencing of the rat carotid arteries uncovers potential cellular targets of neointimal hyperplasia. *Front. Cardiovasc. Med.* **8**, 751525 (2021).
13. Zhou, N. & Bao, J. FerrDb: a manually curated resource for regulators and markers of ferroptosis and ferroptosis-disease associations. *Database (Oxford)* **2020** (2020).
14. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 559 (2008).
15. Botling, J. et al. Biomarker discovery in non-small cell lung cancer: integrating gene expression profiling, meta-analysis, and tissue microarray validation. *Clin. Cancer Res.* **19** (1), 194–204 (2013).
16. Yu, G. et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* **16** (5), 284–287 (2012).
17. 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** (D1), D607–D613 (2019).
18. Smoot, M. E. et al. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27** (3), 431–432 (2011).
19. Ashburner, M. et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* **25** (1), 25–29 (2000).
20. Chin, C. H. et al. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* **8** (Suppl 4), S11 (2014).
21. Yu, G. Gene ontology semantic similarity analysis using GOSemSim. *Methods Mol. Biol.* **2117**, 207–215 (2020).
22. Mazumder, R. & Hastie, T. The graphical Lasso: new insights and alternatives. *Electron. J. Stat.* **6**, 2125–2149 (2012).
23. Engebretsen, S. & Bohlin, J. Statistical predictions with glmnet. *Clin. Epigenet.* **11** (1), 123 (2019).
24. Robin, X. et al. pROC: an open-source package for R and S + to analyze and compare ROC curves. *BMC Bioinform.* **12**, 77 (2011).
25. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26** (12), 1572–1573 (2010).
26. Chen, B. et al. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* **1711**, 243–259 (2018).
27. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* **14**, 7 (2013).
28. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14** (11), 1083–1086 (2017).
29. Pandiar, D., Krishnan, R. P., Ramani, P., Anand, R. & Sarode, S. Oral submucous fibrosis and the malignancy arising from it, could best exemplify the concepts of cuproplasia and cuproptosis. *J. Stomatol. Oral Maxillofac. Surg.* **124**, 101368 (2023).
30. Cui, X. et al. The molecular mechanisms of defective copper metabolism in diabetic cardiomyopathy. *Oxid. Med. Cell. Longev.* **2022**, 5418376. (2022).
31. Jiang, R. et al. Cuproptosis-related gene PDHX and heat stress-related HSPD1 as potential key drivers associated with cell stemness, aberrant metabolism and immunosuppression in esophageal carcinoma. *IntImmunopharmacol* **117**, 109942. https://doi.org/10.1016/j.intimp.2023.109942 (2023).
32. Wang, T. et al. Cuproptosis-related gene FDX1 expression correlates with the prognosis and tumor immune microenvironment in clear cell renal cell carcinoma. *Front. Immunol.* **13**, 999823. https://doi.org/10.3389/fimmu.2022.999823 (2022).
33. Xu, J. et al. Multiomics pan-cancer study of cuproptosis core gene FDX1 and its role in kidney renal clear cell carcinoma. *Front. Immunol.* **13**, 981764 (2022).

Acknowledgements

We are grateful that this study was supported by Animal Experiment Center of Nanjing First Hospital, Nanjing Medical University.

Author contributions

KJ, GJP and MQJ designed, guided, and funded the study. HM, CH and LZL performed most of the operations, and ZBX provided a lot of constructive advice in the operation. LZL and HX helped with the software calculation, CH assisted with drawing pictures, and KJ helped with the revised picture. Data analysis was performed by KJ and LZL. Critical revision of the manuscript and final approval for publication were done by KJ, GJP and MQJ.

Funding

This work was supported by grants from the Jiangsu Medical Association Special Fund Project [SYH-3201140-0094(2023041)]

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

This experiment was approved by the Animal Ethics Committee of Nanjing Hospital of Nanjing Medical University in accordance with institutional guidelines (approval number DWSY-23080536). The study complied with the ARRIVE guidelines, and all the methodology described in this section was performed according to the relevant guidelines and regulations.

Consent for publication
Not applicable.

Additional information
Correspondence and requests for materials should be addressed to Q.M., J.G. or J.K.

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.