

RESEARCH

Open Access



Molecular landscape of atherosclerotic plaque progression: insights from proteomics, single-cell transcriptomics and genomics

Chaonan Wang^{1,2†}, Yuyao Feng^{2†}, Xiaoyan Liu^{3†}, Haidan Sun³, Zhengguang Guo³, Jiang Shao², Kang Li², Junye Chen^{2,4}, Keqiang Shu², Deqiang Kong², Jiaxian Wang⁵, Yiran Li⁶, Xiangling Lei⁵, Chen Li⁷, Bao Liu^{2*}, Wei Sun^{4*} and Zhichao Lai^{2*}

Abstract

Backgrounds Atherosclerosis is a major contributor to cardiovascular diseases worldwide. Despite advancements in understanding its pathology, significant gaps remain in the molecular characterization of atherosclerotic plaques. This study addresses this gap by extensively profiling the proteomic landscape of carotid atherosclerotic plaques, classified under the American Heart Association (AHA) types IV to VI, to identify potential biomarkers and therapeutic targets.

Methods The study employed an integrated approach using data-independent acquisition (DIA) proteomics, single-cell RNA sequencing, and Mendelian randomization (MR). A total of 87 human carotid plaques were analyzed to identify and quantify protein expression. These proteins were then mapped to specific regions within the plaques, such as the fibrous cap and lipid core, and further validated in independent samples and single-cell datasets. Furthermore, Mendelian randomization techniques were employed to assess causal relationships between identified proteins levels and ischemic stroke.

Results The proteomic analysis of the 87 carotid plaques revealed 6143 proteins, highlighting diverse expression profiles across different plaque stages. Notably, proteins like CD44 and GAL-1 were predominantly expressed in the fibrous cap, suggesting a role in plaque stability, while TREM2, SMAD3, and IL-6R showed higher expression in the lipid core, indicating involvement in inflammatory processes. These findings were further corroborated by single-cell RNA sequencing, revealing cell-specific expression patterns that align with the observed proteomic data. Additionally, MR analysis indicated the causal role of IL6R, CD44, and SMAD3 in ischemic stroke.

Conclusions This study provides valuable insights into the progression of atherosclerotic plaques, identifying key proteins that could serve as potential biomarkers and therapeutic targets. It enhances our molecular understanding

[†]Chaonan Wang, Yuyao Feng and Xiaoyan Liu contributed equally to this work.

*Correspondence:

Bao Liu

liubao@pumch.cn

Wei Sun

sunwei@ibms.pumc.edu.cn

Zhichao Lai

zhichao_lai@126.com

Full list of author information is available at the end of the article



of atherosclerosis and opens up new avenues for treatment. Additionally, our study demonstrates the accuracy and robustness of proteomics in prioritizing genes associated with plaque-related traits.

Keywords Carotid artery stenosis, Atherosclerosis, Proteomics, Mendelian randomization, Therapeutic targets

Background

Atherosclerosis (AS) is a chronic disease that progresses slowly and is often symptomless in its early stages. However, it can ultimately lead to severe clinical events such as myocardial infarction or stroke, which are major causes of morbidity and mortality worldwide [1, 2]. In 2019, approximately 15 million people globally lost their lives due to ischemic heart disease and stroke, accounting for 27.2% of all deaths [3]. Plaque rupture, plaque erosion, and calcified nodules are the three primary types of vulnerable lesions that are prone to thrombosis and major adverse cardiovascular events [4, 5]. Understanding the molecular mechanisms of plaque development is essential for early identification of high-risk plaques and targeted therapies.

AS lesions naturally progress and can be classified into 8 types based on the American Heart Association (AHA) classification [4–6]. Among these, type IV–VI plaques are complex, linked to ischemic events and strokes, and are our primary focus for intervention. Type IV lesions are characterized by a lipid core, while type V lesions have a thick fibrous connective tissue layer surrounding the lipid core. Type VI lesions exhibit intimal erosion, hemorrhage, and thrombus formation. Accurately identifying molecular markers for type IV–VI plaques is significant for assessing patient conditions and choosing the most appropriate treatment. Moreover, understanding the mechanisms underlying the progression of type IV–VI plaques can aid in developing new therapeutic approaches to stabilize these plaques and prevent their progression to type VI, which would result in poorer clinical outcomes.

In recent years, advances in proteomics techniques have proven instrumental in identifying potential drug targets and disease biomarkers, particularly in the realm of cancer research [7, 8]. Likewise, proteomic inquiries into AS plaques offer the potential to deepen our understanding of this disease at both the molecular and functional levels. For example, one study identified four potential protein biomarkers from samples obtained during carotid endarterectomy surgery and secreted proteins from human smooth muscle cells (SMCs). These biomarkers were validated through long-term follow-up and detailed monitoring of atherosclerosis progression and manifestations [9]. Another recent large cohort study used proteomics to identify distinct plaque inflammation and calcification signatures, unveiling sex differences

in atherosclerosis [10]. Our previous work employed proteomics to differentiate between stable and unstable carotid plaques based on pathological assessment [11]. However, the binary classification of stable versus unstable plaques limits understanding of plaque development. For example, type IV and type V lesions, though considered stable, represent different stages of plaque progression. Understanding their differences and identifying potential interventions in the transition process is crucial for deeper insight into atherosclerosis pathology. Until now, a comprehensive proteomic characterization of carotid plaque progression across stages remains unexplored.

In this study, we employed DIA-based quantitative proteomic analysis on carotid artery plaques with varying pathological classifications. We identified critical proteins, conducted functional and pathway analyses, and validated our discoveries in an independent clinical cohort. Furthermore, we integrated single-cell transcriptomic data to unveil distinct molecular expression profiles across various cell types. We also utilized Mendelian randomization to establish causal associations between protein expression and stroke risk (Fig. 1). Our research presents the first comprehensive proteomics atlas of plaques, spanning from early to advanced stages, with the potential to aid in the discovery of biomarkers and therapeutic targets aimed at preventing plaque development and rupture.

Methods

Patient cohort

The Carotid Artery Stenosis Study (COAS-CAS, Clinical trial: NCT05629000) is a prospective investigation conducted at Peking Union Medical College Hospital (PUMCH). Its aim is to gather atherosclerotic carotid plaques and associated clinical characteristics, including blood parameters, liver and kidney functions, and imaging outcomes from patients undergoing carotid endarterectomy (CEA). The study received approval from the PUMCH Ethics Committee (Approval No: JS-2966), ensuring adherence to ethical guidelines. Prior to participation, written informed consent was acquired from all patients.

Between May 2015 and March 2022, we enrolled patients at PUMCH who underwent either standard CEA or reverse CEA (eversion, eCEA). Eligibility criteria

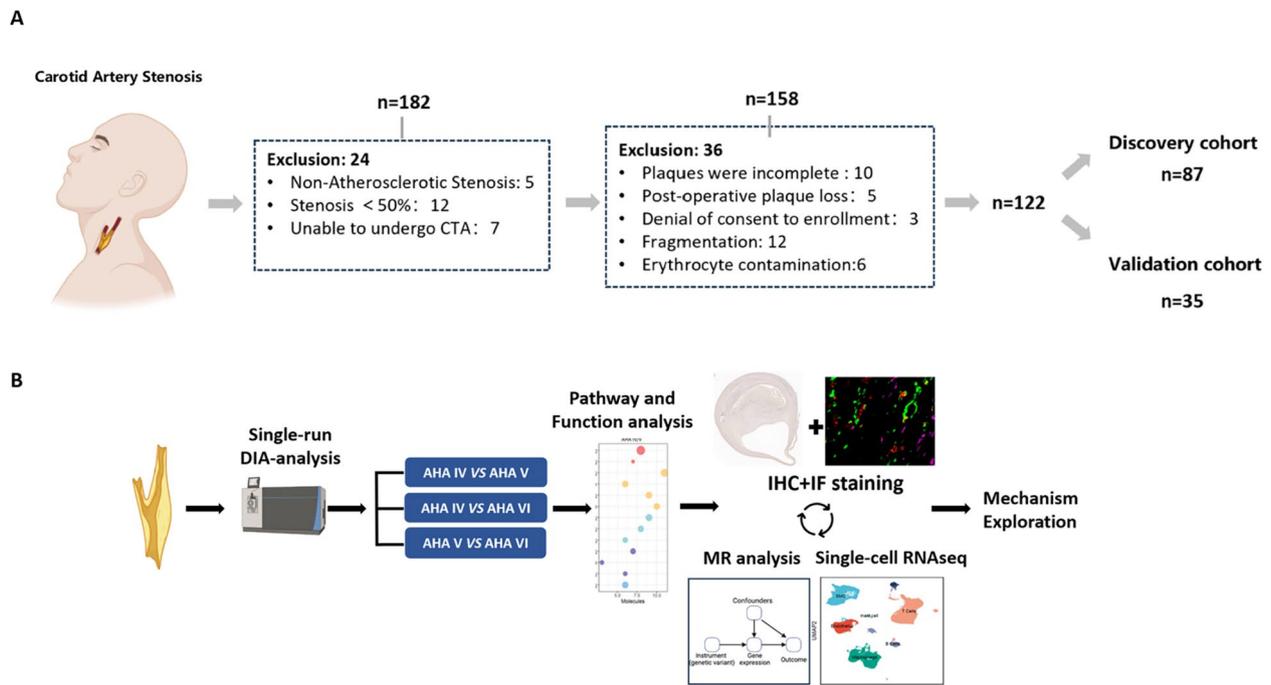


Fig. 1 DIA-based proteomic analysis of carotid plaque cohorts. **A** Study cohort inclusion and exclusion flowchart. **B** Overview of the carotid plaque proteomic detection and analysis workflow. CTA: computed tomography angiography; DIA: data-independent acquisition; AHA: American Heart Association; IHC: immunohistochemistry; IF: immunofluorescence; MR: Mendelian randomization

included asymptomatic carotid stenosis ranging from 70 to 99% and symptomatic carotid stenosis from 50 to 99%, as defined by the North American Symptomatic Carotid Endarterectomy Trial (NASCET) criteria. Upon admission, clinical data were systematically recorded for each participant. Additionally, all enrolled patients underwent preoperative carotid duplex ultrasound scans confirming a minimum of 50% carotid stenosis. Comprehensive clinical data and diagnostic indicators are detailed in Supplementary Table S1. This study strictly adhered to the ethical principles outlined in the Declaration of Helsinki.

Collection and processing of human carotid endarterectomy samples

During surgery, samples of carotid plaque were obtained. Throughout the procedure, efforts were made to ensure complete excision. Subsequently, the samples were gently rinsed with saline to eliminate blood and other contaminants. Following this, they were immediately immersed in 10% formalin for fixation and subjected to decalcification using 10% acetic acid. The plaque was divided into proximal and distal regions, with a focus on the carotid bifurcation and the artery's transverse axis. Wax block production utilized a

section from the proximal part, extending a center distance of 20 mm, and from the distal segment, extending a center distance of 15 mm. If necessary, the range was expanded until the entire plaque was encompassed.

Sample preparation prior to LC-MS analysis

Tissue sections were collected into an EP tube using a scalpel, followed by the addition of 20 μ l of protein extraction solvent. After incubating the samples in a 100 $^{\circ}$ C water bath for 20 min with intermittent shaking, protein digestion was facilitated using the Filter-Aided Sample Preparation (FASP) method. The proteins were initially reduced with 20 mM DTT at 95 $^{\circ}$ C for 5 min and subsequently carboxyamidomethylated with 50 mM IAA in the dark at room temperature for 45 min. The processed sample was then loaded onto a 30-KD ultracentrifugation filter, washed twice with UA buffer (containing 7 M urea and 50 mM Tris), and rinsed twice with 25 mM NH_4HCO_3 . Digestion was achieved using trypsin (2 μ g per 100 μ g protein) in 25 mM NH_4HCO_3 at 37 $^{\circ}$ C overnight. Following digestion, peptides were eluted from the 30-KD filter, desalted using C18 columns (3 cc, 60 mg, Oasis, Waters Corporation, Milford, MA), and subsequently lyophilized via vacuum centrifugation before storage at -80 $^{\circ}$ C.

LC–MS analysis workflow for tissue extracts

Tissue extracts underwent analysis using the Ulti-Mate3000 system (Thermo Fisher Scientific) equipped with an integrated monolithic C18 capillary column (ID, 75 μm ; Length, 50 cm; Uritech, Beijing). Throughout the experiment, the column temperature was consistently maintained at 60 °C. Separation was executed using stepped linear solvent gradients, with a fixed flow rate of 1.5 $\mu\text{l}/\text{min}$ over a 25-min duration. The gradient protocol began with an organic modifier content (acetonitrile acidified with 0.1% v/v formic acid) of 5%, which increased to 20% over 15.5 min. Subsequently, it progressed from 20 to 30% in 5 min, further to 50% in 1 min, reached 90% in 0.1 min, held for 1.3 min, and finally reverted to 5% over 0.1 min, maintaining this level for 2 min.

The mass spectrometry analysis utilized the Orbitrap Exploris 480 MS (Thermo Fisher Scientific) operated in positive mode. For full scans, parameters included a resolution of 120,000, a normalized AGC target of 300%, and a scan range spanning from 350 to 1200 m/z. In DIA mode, the resolution was set at 30,000 with a normalized AGC target of 200%, accompanied by a maximum inject time of 50 ms. A normalized HCD collision energy of 30% was employed for optimal results.

Spectral library generation

Following data acquisition, the obtained results were imported into the Spectronaut Pulsar software (version 14, Biognosys, Switzerland) for spectral library generation. The data were searched against the human SwissProt database. Specific parameters included allowance for a maximum of two missed cleavages for trypsin digestion. Cysteine carbamidomethylation was designated as a fixed modification, while variable modifications encompassed methionine oxidation, lysine deamination, and carbamylation (+ 43). Parent and fragment ion mass tolerances were established at 10 ppm and 0.02 Da, respectively. An applied false discovery rate (FDR) cutoff of 0.01 was set at the protein level to ensure the reliability of the generated library.

Quality control and differentially expressed protein (DEP) identification

All data underwent filtering using a Q value cutoff set at 0.01, corresponding to an FDR of 1%, to eliminate outliers. Protein intensities were determined by summing the intensities of their individual peptides. Proteins identified in more than 80% of samples within each group were selected for subsequent analysis. Missing values were imputed using the *k*-nearest neighbor method. Principal component analysis (PCA) was conducted through the web-based platform MetaboAnalyst 5.0 (<https://www.metaboanalyst.ca/>). For assessing inter-group differences,

non-parametric tests, specifically the Wilcoxon rank-sum test, were employed. Proteins were deemed DEPs if they exhibited a fold change greater than 2 and a *p*-value less than 0.05.

Proteomic data analysis and network construction

Using MetaboAnalyst 5.0, volcano plots were generated to visualize data patterns. For further visualization, a cluster heatmap was crafted utilizing the “ggplots” package in R. To elucidate the biological significance of differential proteins, Ingenuity Pathway Analysis (IPA) software (QIAGEN, Ingenuity Systems, Mountain View, CA) was employed. Within IPA, proteins were categorized based on disease associations, functional roles, and canonical pathways derived from the Ingenuity Knowledge Base and other relevant databases, subsequently ranked by their respective *P*-values.

To explore potential protein interactions, the differential proteins were inputted into the STRING database. A protein–protein interaction (PPI) network analysis was then conducted with a minimal interaction score threshold set at 0.4. Visualization and interpretation of the biomolecular interaction networks were achieved using Cytoscape software (version 3.7.1).

Furthermore, the relationship between clinical characteristics and DEPs was assessed using Pearson correlation coefficients. Associations were deemed significant if the absolute value of the Pearson correlation coefficient exceeded 0.4 and the associated *p*-value was less than 0.05. Correlation outcomes were graphically represented as heatmaps, facilitated by the “pheatmap” package in R.

Immunohistochemistry (IHC) and fluorescent multiplex immunohistochemistry (mIHC) staining

Formalin-fixed, paraffin-embedded CEA samples were used for immunostaining. Initially, tissue sections underwent deparaffinization and rehydration using xylene followed by a graded ethanol series. Subsequently, antigen retrieval was achieved by boiling sections in sodium citrate buffer (pH = 6.0) for 2 min. Sections were treated with 0.3% H₂O₂ for a duration of 15 min to inhibit endogenous peroxidase activity. A blocking step with 2% fetal calf serum ensued for 20 min. Primary antibodies were applied and allowed to incubate overnight at 4 °C, followed by secondary antibodies, specifically HRP-conjugated anti-mouse and anti-rabbit immunoglobulins, for 1 h at room temperature. Post-incubation, sections were counterstained using Mayer’s hematoxylin, dehydrated, and examined microscopically. For visualization and analysis, stained tissues were scanned using a Panoramic Desk Scanner (3DHistech, Hungary) and viewed through Case Viewer (version 2.3, 3DHistech, Hungary). Quantification of

the density of positively stained cells within cross-sections of carotid plaques was conducted using Image-Pro Plus software (Ver. 6.0, Media Cybernetics, USA).

For mIHC staining, Tyramide Signal Amplification (TSA) was applied following incubation with primary and secondary antibodies. After each staining step, stripping was performed by boiling the tissue sections in sodium citrate buffer (pH 6.0). Once each protein target was stained, the slides were mounted with DAPI.

Primary antibodies utilized included CD44 (rabbit polyclonal, 15,675–1-AP), LGALS1 (rabbit polyclonal, 11,858–1-AP), TREM2 (rabbit polyclonal, 13,483–1-AP), IL-6R (rabbit polyclonal, 23,457–1-AP), and SMAD3 (mouse monoclonal, 66,516–1-Ig), all sourced from Proteintech (China). Additionally, α SMA (rabbit polyclonal, GB111364) and CD68 (rabbit polyclonal, GB113150) primary antibodies were obtained from Servicebio.

Single-cell RNA sequencing and analysis

The investigation in this study aimed to assess how specific genes are expressed in diverse cell types within human atherosclerotic plaques. To achieve this objective, a reanalysis of previously published single-cell RNA sequencing dataset (GSE159677) [12, 13] was conducted. These datasets included 3 samples from the core of human carotid plaques and 3 paired peripheral plaque samples from the same individuals, obtained from the Gene Expression Omnibus (GEO).

The raw single-cell sequencing data for each sample were independently processed using the 10X Genomic Chromium platform and Cell Ranger with default parameters, as thoroughly described in the original publication [12]. Subsequently, quality filtering was applied by Seurat V4.1.1 [14] to eliminate cells with mitochondrial mRNA exceeding 10% of total mRNA, as well as cells expressing fewer than 500 or more than 4000 genes. Following this preprocessing, the data underwent normalization and transformation into scaled values. Linear dimensional reduction, cluster identification, and nonlinear dimensional reduction were carried out in accordance with the Seurat manual.

To identify distinctive clusters across multiple combined datasets, data integration was performed post-normalization and before scaling. Clustering was executed using the “FindClusters” function with a resolution parameter set to 0.4. Subsequently, the clusters were annotated with specific cell types based on established molecular markers commonly used in single-cell data analysis (Fig. 4B).

Methods for Mendelian randomization analysis

Study design and objective

We conducted a two-sample Mendelian randomization (MR) analysis [15, 16] to assess the causal relationship between blood protein expression (CD44, TREM2, IL6R, GAL-1, and SMAD3) and the risk of ischemic stroke. These proteins were selected based on their potential involvement in the transition of atherosclerotic plaque types from type IV to type VI, which are associated with plaque instability and ischemic stroke. We used summary data from genome-wide association studies (GWAS) of blood proteins and ischemic stroke to identify genetic instruments and estimate causal effects.

Genetic instruments (instrumental variables)

We selected genetic variants as instrumental variables (IVs) for five proteins of interest: IL6R, CD44, TREM2, GAL-1, and SMAD3. Genetic variants were selected based on their association with protein levels in previously published protein quantitative trait loci (pQTL) studies.

For IL6R, CD44, and GAL-1, we utilized summary data from the INTERVAL study [17], which provides pQTLs for 2965 blood proteins measured in 3301 individuals of European ancestry. Due to the absence of data on TREM2 and SMAD3 in the INTERVAL study, pQTLs for these proteins were obtained from the UK Biobank Pharma Proteomics Project (UKB-PPP), which includes protein data for 2922 proteins from 34,557 participants of European ancestry [18].

These candidate IVs rely on three key assumptions for MR analysis [19]: (1) The IVs are associated with gene expressions ($p < 1 \times 10^{-5}$). (2) The IVs are independent of any potential confounders. (3) The IVs are independent of the outcome variable (ischemic stroke).

To ensure the independence of IVs, we remove all SNPs that were present in the 1000 Genomes European population with an $r^2 > 0.1$ and located within 10,000 kb of the top SNPs [20]. This step was taken to strike a balance between considering the correlation between variants, which can enhance statistical power compared to selecting strictly independent variants, and preventing potential instability in Mendelian randomization estimators when variants are highly correlated [21]. The pruning r^2 threshold was chosen to address this balance. To evaluate the risk of weak instrument bias, we employed F statistics to assess the strength of the association between the allele and the exposure [22]. The F-statistic was > 10 for each variant confirming the validity of our selected IVs (Supplementary table S6).

Outcome data and study population

The outcome of interest was ischemic stroke, for which we used summary data from the MEGASTROKE consortium, a meta-analysis of 40,585 ischemic stroke cases and 406,111 controls of European ancestry [23]. The ischemic stroke data were harmonized with the exposure data to ensure that the genetic variants for proteins were aligned with those used in the ischemic stroke GWAS.

Mendelian randomization analysis

We performed two-sample MR analysis using the inverse variance weighted (IVW) method, which is the primary method for estimating causal effects in MR studies. We used the IVs identified for each protein (IL6R, CD44, TREM2, GAL- 1, and SMAD3) to estimate the causal relationship between blood protein levels and ischemic stroke risk. The IVW method provides a consistent estimate of the causal effect, assuming no pleiotropy and that the instrumental variables are valid. The effect size for the MR analysis is expressed as the effect of a one-standard deviation (1-SD) change in protein levels. Multiple testing correction was applied, and statistical significance was defined at an FDR threshold of < 0.05.

Sensitivity analysis

To minimize heterogeneity and horizontal pleiotropy, we conducted several tests.

Leave-one-out analysis sequentially removed each genetic variant to check if any single SNP disproportionately influenced the overall MR estimate. Cochran's Q test was used to evaluate the heterogeneity of the SNPs included as instruments for each protein. MR Egger regression was performed to test for evidence of horizontal pleiotropy, with a non-zero intercept suggesting potential bias in the causal estimates [24]. MR-PRESSO (Mendelian Randomization Pleiotropy RESidual Sum and Outlier) test was used to identify and correct for outliers in the IVs, particularly for GAL- 1, where heterogeneity and pleiotropy were observed.

Results from the IVW method were interpreted alongside those from the weighted median method, which is more robust to violations of instrument validity, particularly when pleiotropy is suspected [25].

Statistical analysis

Baseline demographic data were presented as mean \pm SD for continuous variables and as frequencies for categorical variables. Shapiro–Wilk test was used to test normality. The Brown-Forsythe test (or F-test) was used to assess variance homogeneity. For the data passed normality test, an unpaired Student's *t* test was applied to compare differences between the two independent groups. IHC staining quantification data are expressed as mean

\pm standard error of the mean (mean \pm SEM). Post-IHC staining comparisons between groups were executed using a two-tailed Student's *t* test. Pearson correlation analysis was employed to evaluate associations between clinical characteristics and DEPs. A threshold of $P < 0.05$ denoted statistical significance. Statistical computations were carried out utilizing GraphPad Prism software (ver. 7.0, GraphPad, USA) and R software (Version 4.1.0) (<https://www.r-project.org/>). TwoSampleMR (Version 0.6.6) and MR-PRESSO (Version 1.0) packages were used for MR analysis.

Results

Clinical characteristics of carotid plaques

A total of 182 subjects with carotid stenosis were consecutively recruited for this study at PUMCH between May 2015 and March 2022. After excluding 24 patients, 158 eligible patients remained for the subsequent study cohort. Hematoxylin and eosin (HE) staining was performed on tissue samples from these patients. These plaques were classified by two pathologists independently according to the American Heart Association (AHA) classification [4, 6]. After excluding plaques with contamination or severe damage that hindered accurate grading, a total of 122 plaques were included in the study. These were then randomly divided into a discovery cohort and a validation cohort in a 7:3 ratio. The discovery cohort consisted of 87 plaques, which were used for proteomic analysis in DIA mode, while the validation cohort included 35 plaques, which underwent IHC analysis. The flowchart of the study cohort inclusion and exclusion is shown in Fig. 1A.

The baseline clinical information of the enrolled patients is provided in Table 1 and Additional file 1. This included 100 males and 22 females, with a mean age of 66.79 ± 7.98 years. AS-related clinical indices were quantitatively analyzed for all enrolled patients. These analyses covered AS-related clinical indicators such as risk factors, AS-related diseases, degree of disease stenosis, and blood biochemical indicators of AS severity. There were no statistically significant differences between the groups in clinical indicators, including blood routine, liver function, and kidney function tests.

Common changes of proteome landscape throughout the plaque development

We utilized data-independent acquisition (DIA) technology to analyze protein expression in carotid artery plaques, specifically focusing on AHA types IV, V, and VI. Our analysis successfully identified a total of 6143 proteins, maintaining a strict 1% false discovery rate (FDR). On average, each sample revealed the presence of 3777 proteins, with varying numbers detected within

Table 1 Demographic and clinical characteristics of the subjects

	Discovery cohort (n = 87)		Validation cohort (n = 35)				P ¹ value	P ² value	P ³ value
	AHA IV (n = 29)	AHA V (n = 20)	AHA VI (n = 38)	AHA IV (n = 10)	AHA V (n = 10)	AHA VI (n = 15)			
Age (years)	66.79 ± 9.67	66.60 ± 8.70	67.74 ± 6.97	64.90 ± 5.45	65.60 ± 9.18	66.07 ± 3.69	0.838 ^a	0.860 ^a	0.528 ^a
Gender (male, %)	75.86	80.00	86.84	80.00	60	73.33	0.329 ^b	0.484 ^b	0.702 ^b
Surgical side (%)							0.371 ^b	0.734 ^b	0.188 ^b
Left	41.38	50.00	44.74	60.00	40.00	33.33			
Right	58.62	50.00	55.26	40.00	60.00	66.67	0.067 ^b	0.004 ^b	0.433 ^b
Degree of left carotid artery stenosis (%)									
Mild	55.17	25.000	28.95	10.00	40.00	20.00			
Moderate	24.14	20.000	15.79	10.00	20.00	33.33			
Severe	20.69	45.000	39.47	80.00	40.00	46.67			
Occlusion	0.00	10.000	15.79	0.00	0.00	0.00			
Degree of right carotid artery stenosis (%)							0.422 ^b	0.828 ^b	0.423 ^b
Mild	27.59	30.00	23.68	50.00	20.00	13.33	0.121 ^a	0.275 ^a	0.504 ^a
Moderate	10.35	15.00	10.53	10.00	0.00	20.00	0.329 ^a	0.702 ^a	0.483 ^a
Severe	48.28	40.00	60.53	10.00	80.00	66.67	> 0.999 ^a	0.027 ^a	0.027 ^a
Occlusion	13.79	15.00	5.26	30.00	0.00	0.00	0.159 ^a	0.405 ^a	0.467 ^a
Hypertension (%)	68.97	75.00	63.16	90.00	60.00	73.33	0.632 ^a	0.513 ^a	0.249 ^a
Diabetes (%)	41.38	30.00	44.74	40.00	20.00	26.67	0.639 ^a	0.861 ^a	0.733 ^a
Hyperlipidemia (%)	20.69	35.00	31.58	90.00	90.00	46.67	> 0.999 ^a	> 0.999 ^a	> 0.999 ^a
Coronary heart disease (%)	24.14	15.00	31.58	20.00	50.00	33.33	0.275 ^a	0.702 ^a	0.483 ^a
Smoking (%)	31.03	60.00	47.37	70.00	60.00	46.67	0.159 ^a	0.405 ^a	0.467 ^a
Cerebral infarction (%)	17.24	40.00	36.84	40.00	30.00	33.33	0.632 ^a	0.513 ^a	0.249 ^a
Subclavian artery stenosis (%)	17.24	15.00	7.90	20.00	20.00	20.00	> 0.999 ^a	> 0.999 ^a	> 0.999 ^a
Lower limb artery stenosis (%)	3.45	15.00	7.90	10.00	10.00	6.67	> 0.999 ^a	> 0.999 ^a	> 0.999 ^a
Total cholesterol (mmol/L)	3.70 ± 0.87	7.90 ± 16.53	6.96 ± 15.48	3.83 ± 1.97	3.57 ± 1.13	3.75 ± 0.67	0.725 ^a	0.548 ^a	0.945 ^a
Triglycerides (mmol/L)	1.28 ± 0.93	1.82 ± 1.12	1.40 ± 0.66	2.15 ± 1.52	1.98 ± 1.32	1.14 ± 0.45	0.795 ^a	0.030 ^a	0.022 ^a
HDL-C (mmol/L)	1.03 ± 0.20	1.00 ± 0.24	1.04 ± 0.25	1.05 ± 0.36	0.87 ± 0.34	1.08 ± 0.20	0.285 ^a	0.067 ^a	0.773 ^a
LDL-C (mmol/L)	2.05 ± 0.61	2.37 ± 0.79	1.73 ± 0.71	2.17 ± 0.83	2.08 ± 0.95	2.03 ± 0.59	0.824 ^a	0.870 ^a	0.624 ^a
ApoA1 (g/L)	1.25 ± 0.18	1.21 ± 0.18	1.20 ± 0.16	1.26 ± 0.28	1.22 ± 0.13	1.21 ± 0.16	0.757 ^a	0.847 ^a	0.631 ^a
ApoB (g/L)	0.750 ± 0.21	0.91 ± 0.25	0.70 ± 0.23	0.74 ± 0.15	0.83 ± 0.39	0.94 ± 0.38	0.458 ^a	0.488 ^a	0.109 ^a
Lipoprotein(a) (mg/L)	164.03 ± 190.49	255.77 ± 273.30	215.49 ± 295.08	200.10 ± 181.42	118.10 ± 98.35	312.53 ± 360.67	0.225 ^a	0.112 ^a	0.373 ^a
hsCRP (mg/L)	2.74 ± 3.63	3.87 ± 6.96	3.03 ± 6.17	1.83 ± 0.89	2.02 ± 1.71	2.89 ± 4.25	0.759 ^a	0.543 ^a	0.444 ^a

^a Two-sided P value for numerical values were determined using unpaired T-test

^b Two-sided P value for categories were calculated using chi-square test

P¹-value = AHA IV/AHA V; P²-value = AHA V/AHA VI; P³-value = AHA IV/AHA VI

Data are presented as the mean ± standard deviation or number (%). P value < 0.05 are indicated as significant

each AHA group (Additional file 2: Fig. S1a-e, Additional file 3). During mass spectrometry, we included a total of 7 quality control (QC) replicates, which were randomly interspersed among the experimental samples. To assess the technical variation in our experiment, we calculated the coefficient of variation of protein abundance between these samples. The results, as illustrated in Additional file 2: Fig. S1b, demonstrated excellent technical reproducibility. Plus, a panel of quality markers was used to evaluate erythrocyte contamination [26], leading to the exclusion of six samples from the study due to significant contamination (Additional file 2: Fig. S1c and Additional file 4).

Our analysis then focused on comparing protein levels across the three AHA groups to uncover pathological and molecular differences. Using principal component analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA), we observed clear separation and significant differences between the groups (Additional file 2: Fig. S1 d-e). Proteins with a fold change >2 and a Benjamini–Hochberg adjusted p -value <0.05 were classified as differentially expressed. This led to the identification of 453, 302, and 438 differentially expressed proteins in the AHA IV vs. V, V vs. VI, and IV vs. VI comparisons, respectively, highlighting distinct proteomic patterns among the different types of plaques (Additional file 2: Fig. S2a-b, Additional file 5: Table S4a-c).

In our analysis comparing protein expression across AHA types IV, V, and VI, we identified several biological functions that were commonly altered, highlighting their importance in plaque development (Fig. 2A). The most significant functions and pathways (with the lowest p -values) were selected for presentation in Fig. 2. These biological functions were primarily related to the regulation of cell death and survival (including apoptosis, necrosis, and autophagy), cellular structures (such as the cytoskeleton and cytoplasm), and metabolism (Fig. 2A). However, there were notable differences in the extent of activation across these processes.

One of the most significant differences between AHA types IV and V was the regulation of cell death mechanisms (Fig. 2A). Specifically, apoptosis was upregulated in grade V plaques compared to grade IV, while necrosis was downregulated in grade V. Interestingly, apoptosis was further suppressed in grade VI plaques compared to grade V, while necrosis was markedly

upregulated in grade VI. These findings suggest that cell death processes are more regulated in grade V plaques, with upregulated apoptosis (programmed cell death) and downregulated necrosis. In contrast, necrosis is most pronounced in grade VI plaques, indicating a more detrimental microenvironment in grade VI.

Another striking difference was observed in molecular transport processes, which were significantly downregulated in grade V plaques compared to both grades IV and VI (Fig. 2A). This may be related to the thick fibrous cap in grade V plaques, which could hinder molecular exchange between the plaque and the lumen.

The regulation of signaling pathways contributes to the differences in these critical functions. To explore the shared signaling pathways involved in plaque development, we performed signaling pathway enrichment analysis, which revealed a predominant activation of endocytosis signaling (Fig. 2B). This suggests active intercellular signaling and molecular transport. Consistent with this, actin cytoskeleton signaling was also found to be involved (Fig. 2B). Additionally, we identified the activation of several inflammation- and metabolism-related signaling pathways (Fig. 2B).

Distinct proteome characteristics of plaques in different AHA types

Type IV plaques represent a key bifurcation point leading to type V or VI. Studying the differences between type IV/V and IV/VI helps us understand how plaques turn to stable or unstable states.

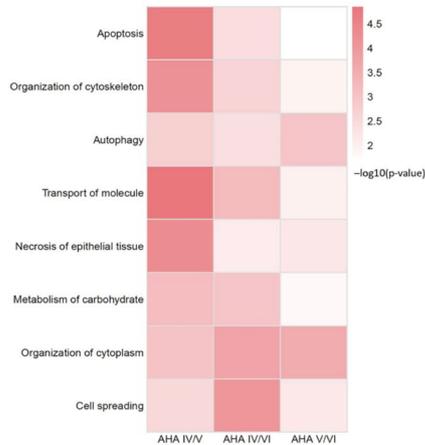
In our comparison of the proteomic characteristics of AHA type IV and type V plaques, we observed significant involvement of cellular transport processes, including both caveolar and clathrin-mediated endocytosis. Another key aspect is the regulation of the cytoskeleton and intercellular junctions, which encompasses actin cytoskeleton signaling (e.g., actin nucleation) and signaling by Rho family GTPases (Fig. 2C). Additionally, critical signaling pathways such as PI3 K/AKT, eIF4, and FAK signaling are implicated in these processes. Notably, metabolic pathways like PTEN signaling, insulin receptor signaling, and D-myo-inositol-trisphosphate biosynthesis exhibit varying degrees of activation between AHA type IV and type V plaques.

In type IV vs. type VI plaques, more pathways are altered, especially those related to metabolism and the

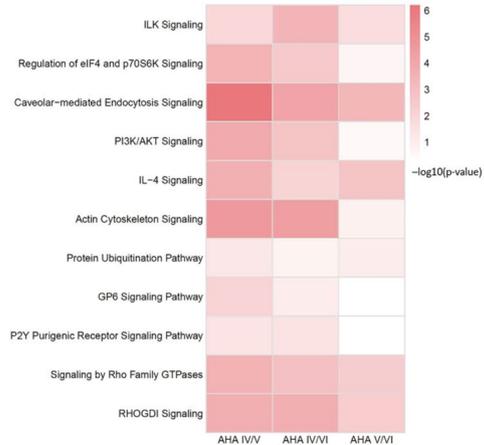
(See figure on next page.)

Fig. 2 Ingenuity pathway analysis (IPA) of functions and pathways associated with differentially expressed proteins (DEPs) between AHA types IV, V, and VI. **A** Analysis of commonly regulated biological functions among DEPs differentiating AHA types IV, V, and VI. **B** IPA canonical pathway analysis of DEPs involved in differentiating AHA types IV, V, and VI. **C** IPA functional analysis of DEPs involved in differentiating AHA types IV and V. **D** IPA functional analysis of DEPs involved in differentiating AHA types IV and VI. **E** IPA functional analysis of DEPs involved in differentiating AHA types V and VI. **F** Distinctly regulated biological functions between AHA types IV, V, and VI. DEPs, differentially expressed proteins

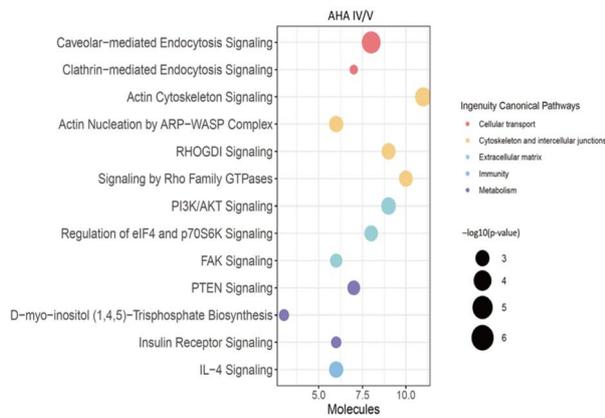
A



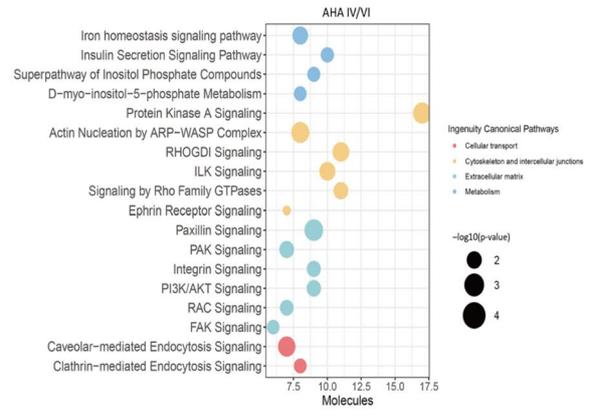
B



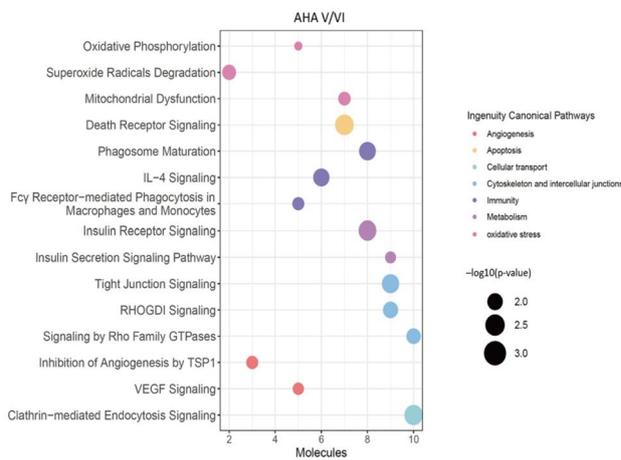
C



D



E



F

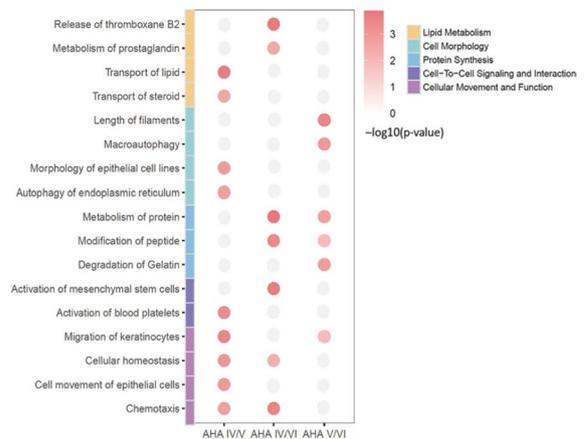


Fig. 2 (See legend on previous page.)

extracellular matrix (Fig. 2D). Changes in iron homeostasis, insulin secretion, and inositol metabolism are prominent, with integrin-linked kinase (ILK) playing a role in cell adhesion. The differences between type IV and type VI plaques will be explored in greater depth in the following sections.

Type V and type VI plaques show significant differences in plaque integrity, with type VI exhibiting unstable features such as intimal erosion, hemorrhage, and thrombus formation. These differences are underpinned by alterations in key biological processes, including disrupted metabolic pathways (e.g., oxidative phosphorylation, insulin signaling), heightened oxidative stress, immune responses, and angiogenesis (Fig. 2E). Notably, the activation of inflammatory pathways, such as IL-4 signaling and Fc gamma receptor-mediated phagocytosis in macrophages and monocytes, differs significantly between type V and type VI plaques (Fig. 2E).

Our investigation further examines the differentially regulated functions during the progression of AHA types IV, V, and VI plaques. Notably, lipid and steroid transport is significantly reduced in type V plaques compared to type IV plaques (Fig. 2F). Chemotaxis is more pronounced in type IV plaques, indicating more active cellular migration in the early stages of plaque formation, which may shift to different signaling mechanisms in later stages. Interestingly, autophagy is upregulated in type V plaques and downregulated in type VI plaques, suggesting that autophagy may play a critical role in maintaining cellular homeostasis in type V plaques (Fig. 2F).

Some biological processes are exclusively associated with the progression from type IV to type VI plaques, such as platelet aggregation (including the release of thromboxane B2 and prostaglandin metabolism) and the activation of mesenchymal stem cells (Fig. 2F). Additionally, processes related to lipid metabolism and cellular migration distinguish type V and type VI plaques (Fig. 2F). In summary, each stage of plaque development is characterized by distinct functional and signaling pathways.

Identification of key proteins in plaque deterioration from type IV to VI

The type IV lesion represents the initial phase of advanced plaque development, characterized by the presence of a lipid core and the absence of a significant increase in fibrous tissue or related complications. This stage is considered relatively primary and stable [6, 27]. Over time, this plaque can evolve in two distinct directions. It may progressively accumulate more fibrous tissue, leading to the development of a type V lesion. Alternatively, it can transform into a complicated plaque marked by hematoma, thrombosis, or surface defects.

Such complications are the primary contributors to the morbidity and mortality associated with atherosclerosis [6, 27]. Therefore, understanding the differences between type IV and type VI lesions is of paramount importance, as it can offer insights into preventive measures.

To achieve this goal, we deliberately selected proteins that demonstrated significant differences exclusively between the AHA type IV and type VI groups, which were used for further comprehensive investigation (Fig. 3A). Functional enrichment analysis uncovered that these proteins are associated with multiple functional pathways, notably those related to inflammatory responses, cell death and survival, and metabolism (Fig. 3B). Moreover, pathway enrichment analysis unveiled significant activation in pathways related to the glucocorticoid receptor, neuroinflammation, and iron homeostasis signaling (Additional file 2: Fig. S3a). Recognizing the intricate network of protein interactions, we proceeded to construct an undirected protein–protein interaction (PPI) network graph through the STRING database. Within this network, we identified a primary hub comprising 27 proteins, mainly associated with inflammatory responses (Fig. 3C). Additionally, a secondary hub emerged, comprising 21 proteins primarily linked to cellular metabolism and movement (Additional file 2: Fig. S3b). These proteins within the hub networks were identified as having more centralized characteristics.

To further understand if the key DEPs identified above are related to certain biochemical and clinical characteristics, a Pearson correlation analysis was conducted. A total of nine differential proteins were selected, including TF (tissue factor), MTOR (serine/threonine-protein kinase mTOR), TFR1 (Transferrin receptor protein 1), CD44, SMAD3 (mothers against decapentaplegic homolog 3), DPP4 (dipeptidyl peptidase 4), TREM2 (triggering receptor expressed on myeloid cells 2), IL6R (Interleukin-6 receptor), and GAL-1 (Galectin-1), along with 36 clinical features (Additional file 2: Fig. S3c). The intrinsic correlations among the nine differential proteins were also performed (Additional file 2: Fig. S3d). None of these relationships showed significant correlations ($r^2 > 0.4$), suggesting that these proteins primarily influence AS plaque development in a localized manner, without substantially impacting the overall metabolic status of the entire body. However, it is noteworthy that mTOR expression exhibited a correlation with alanine aminotransferase (ALT) at an r^2 value of 0.35, while TF expression displayed a similar correlation with total cholesterol (TC), also at an r^2 value of 0.35. Furthermore, DPP4 demonstrated a correlation with direct bilirubin (Dbil) at an r^2 value of 0.33. These intriguing correlations provide valuable insights into the potential mechanisms through which these proteins contribute to AS plaque development.

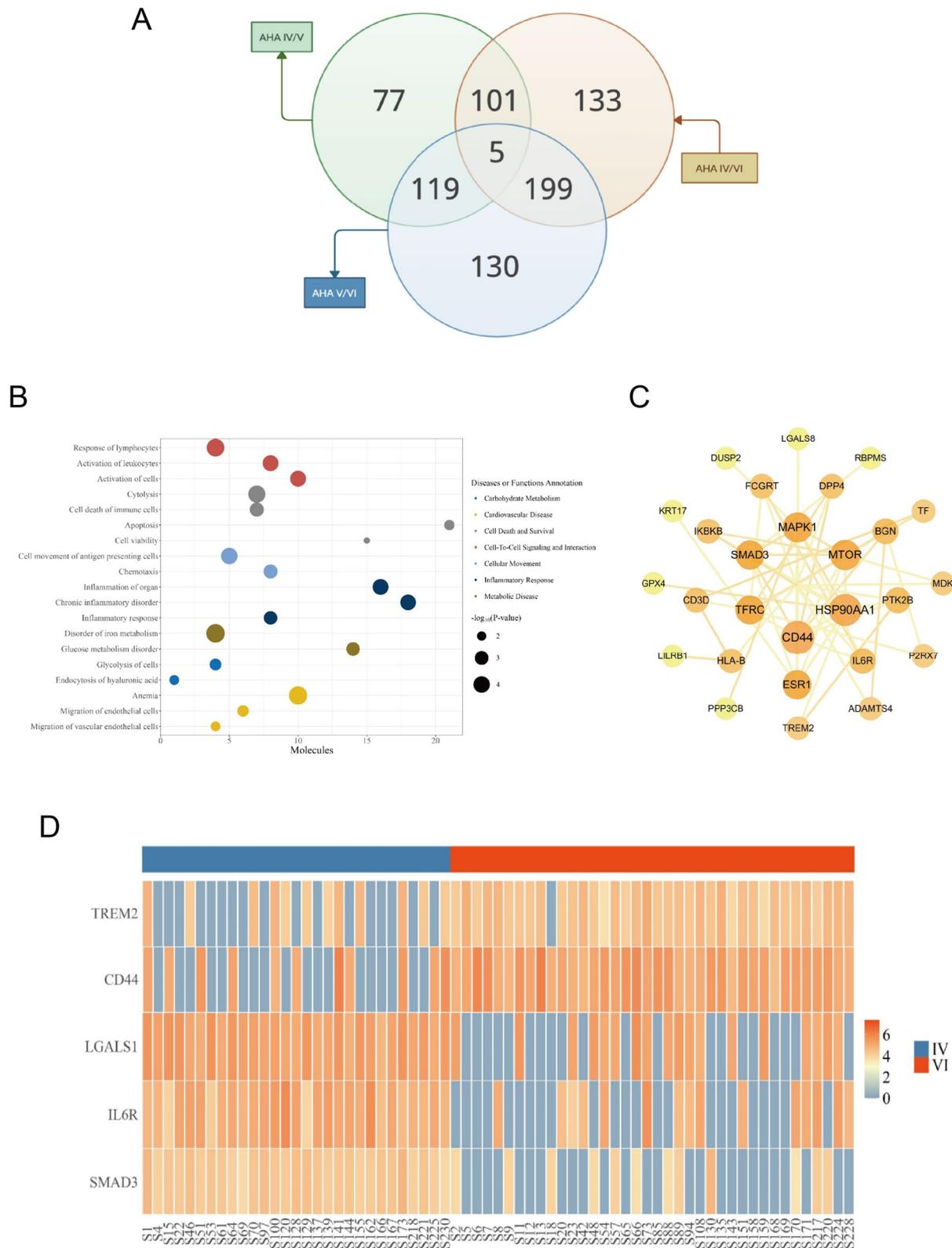


Fig. 3 Network analysis of DEPs and validation of inflammation-associated hub proteins in a validation cohort. **A** Venn diagram showing the DEPs exclusively different between the AHA type IV and type VI groups. **B** IPA disease or function annotations of DEPs exclusively involved in the differentiation between AHA type IV and type VI groups. **C** Protein-protein interaction (PPI) network of inflammation-associated hub proteins. **D** Heatmap of the key proteins expression between the two groups

Table 2 The differentially expressed proteins (DEPs) that were validated using immunohistochemistry (AHA IV/VI)

Protein name	Accession	Gene symbol	FC (discovery)	FC (validation)	P-value (discovery)	P-value (validation)
Cell surface adhesion receptor 44	B9A6J2	CD44	- 1.6649	- 1.2924	4.95E -03	0.0287
Galectin 1	A0A0B6XK00	GAL-1	- 1.2110	- 1.4311	4.66E -05	0.0355
Triggering receptor expressed on myeloid cells 2	Q5TCX1	TREM2	2.3160	1.7664	7.06E -05	0.0221
SMAD Family Member 3	A0A024R5Z3	SMAD3	- 2.4386	- 1.7348	1.73E -10	0.1851
Interleukin 6 receptor	A0A087WTB5	IL-6R	2.0382	1.6389	4.66E -06	0.0404

IHC validates the expression level of key proteins in plaque advancement from type IV to VI

Through our analysis involving IPA and PPI networks of the previously mentioned differential proteins, we prioritized pivotal proteins closely linked to inflammatory responses, serving as hub proteins (CD44, SMAD3, TREM2, IL6R, GAL- 1) (Additional file 6). Specifically, when comparing type VI plaques to type IV plaques, TREM2 and CD44 exhibited elevated expression levels, whereas SMAD3, IL6R, and GAL- 1 displayed reduced expression levels (Fig. 3D, Table 2). To validate these proteomic findings, we performed IHC staining on an independent cohort of 35 human carotid atherosclerotic plaques.

IHC results confirmed the proteomics data, showing increased expression of TREM2 and CD44 in type VI plaques (Fig. 4A, B). The density of CD44-positive cells was significantly higher in type VI plaques compared to type IV plaques (318.3 ± 33.7 vs. 246.3 ± 27.5 , $p < 0.05$). Similarly, TREM2-positive cell density was greater in type VI plaques (824.7 ± 169.4 vs. 632.0 ± 113.9 , $p < 0.05$).

For the proteins showing decreased expression in type VI plaques, IHC staining also revealed a consistent trend (Fig. 4A, B). The expression of SMAD3 showed a decreasing trend in type VI plaques, though the difference was not statistically significant (318.3 ± 33.7 vs. 246.3 ± 27.5 , $p = 0.1851$). GAL- 1 expression was lower in type VI plaques compared to type IV plaques (737.7 ± 85.8 vs. 515.5 ± 70.9 , $p < 0.05$). IL6R expression was higher in type IV plaques (908.5 ± 171.8 vs. 554.3 ± 109.0 , $p < 0.05$).

Overall, IHC results are consistent with the proteomics data, demonstrating the high reliability of the proteomic

approach in assessing protein expression in plaque progression.

Spatial expression pattern of hub proteins in carotid plaques

To investigate the spatial distribution of hub proteins and their association with SMCs and macrophages in carotid atherosclerotic plaques, we employed triple immunofluorescence staining. This approach allowed us to visualize the expression of hub proteins alongside α SMA (SMC marker) and CD68 (macrophage marker), enabling the examination of their co-localization within different regions of the plaque and providing insights into their roles in plaque pathology. Intriguingly, different proteins exhibit distinct expression patterns.

IL6R and CD44 are expressed in both macrophages and SMCs (Fig. 5A, B). In type IV lesions, IL6R primarily co-localizes with α SMA, reflecting the low macrophage content. In type V lesions, IL6R expression decreases in SMCs and increases in macrophages. In type VI lesions, IL6R expression is further reduced in SMCs and predominantly localized around macrophages, indicating its association with macrophage-driven inflammation. CD44, which co-localizes with CD68 in macrophages, is highly expressed in regions with high macrophage content, particularly in the plaque shoulder near the lumen, fibrous cap, and media in type VI plaques. It is also expressed in some media SMCs in type VI plaques. Despite abundant macrophages in the necrotic core of type V plaques, CD44 expression remains low in the macrophages in these areas, while its expression in media SMCs is comparable to that in type VI plaques. In type

(See figure on next page.)

Fig. 4 Immunohistochemical (IHC) staining of the hub proteins in human carotid plaques. **A** Representative images showing the immunohistochemical (IHC) staining of CD44, GAL- 1, TREM2, SMAD3, and IL6R proteins in plaques in the plaques of AHA type IV and type VI. **B** Quantification results (density of positive cells) of IHC staining ($n = 10$ in AHA type IV group, $n = 15$ in in AHA type VI group). Unpaired Student's t test was used to compare differences between two groups as the data passed Shapiro–Wilk test normality test. CD44, Cluster of Differentiation 44; SMAD3: Mothers Against Decapentaplegic Homolog 3; TREM2: Triggering Receptor Expressed on Myeloid Cells 2; IL6R: Interleukin- 6 Receptor; GAL- 1: Galectin- 1; DEPs, differentially expressed proteins. * $P < 0.05$

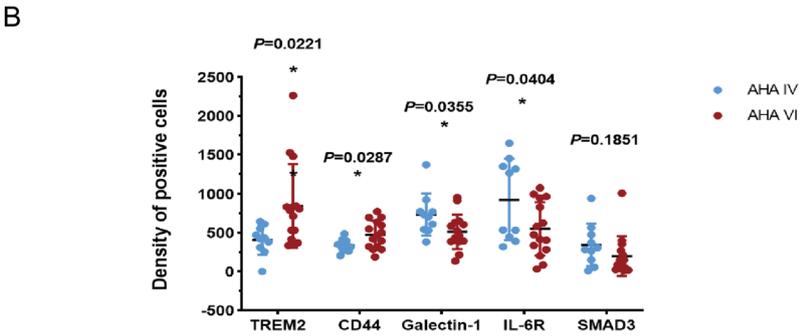
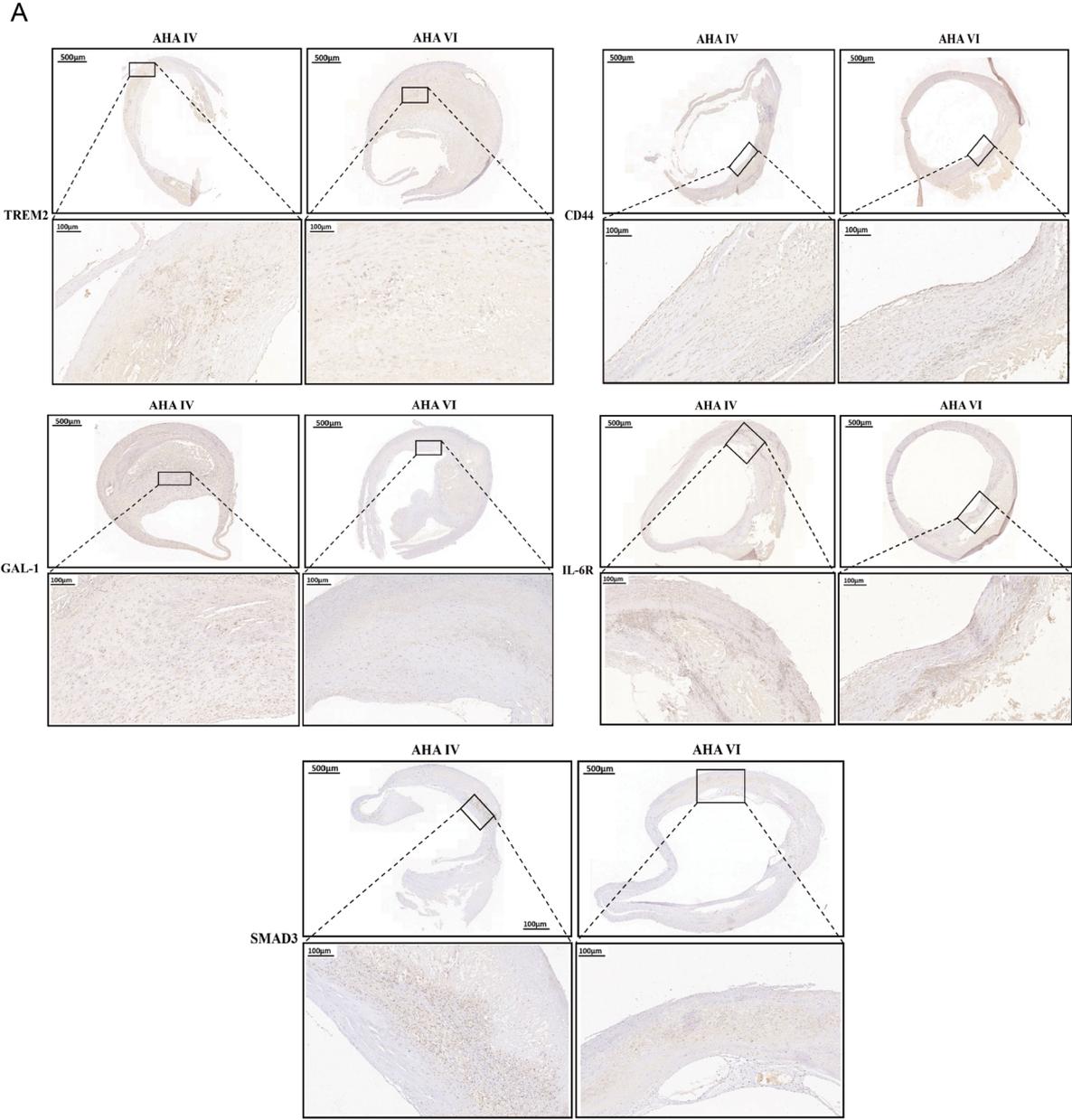


Fig. 4 (See legend on previous page.)

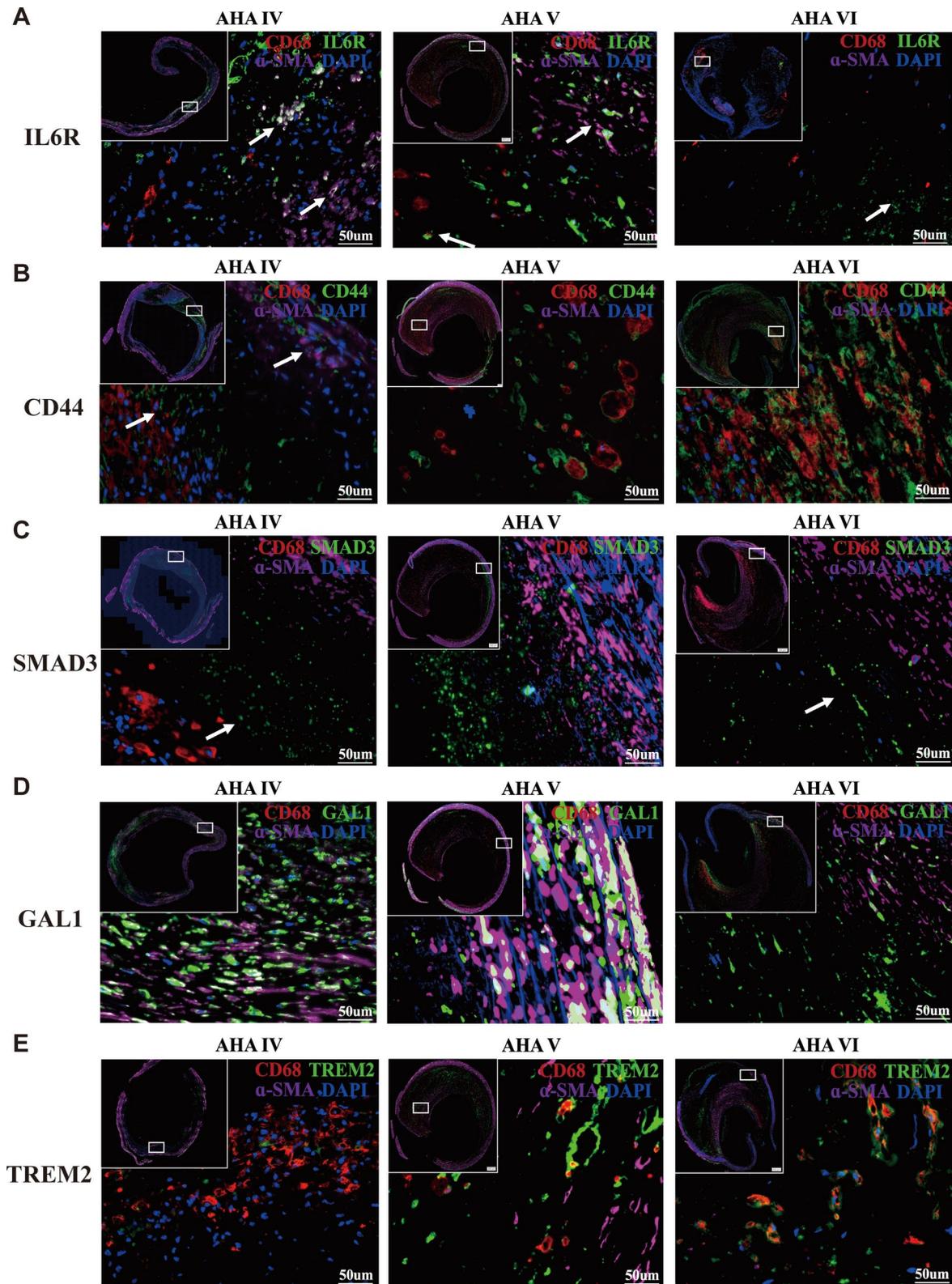


Fig. 5 Fluorescent multiplex immunohistochemistry (mIHC) staining of the hub proteins in human carotid plaques. **A–E** Representative images showing mIHC staining of IL6R (**A**), CD44 (**B**), SMAD3 (**C**), GAL1 (**D**), or TREM2 (**E**), along with αSMA and CD68 proteins, in AHA type IV, V, and VI carotid plaque lesions. CD44, Cluster of Differentiation 44; SMAD3: Mothers Against Decapentaplegic Homolog 3; TREM2: Triggering Receptor Expressed on Myeloid Cells 2; IL6R: Interleukin- 6 Receptor; GAL1: Galectin- 1

IV plaques, CD44 expression is more restricted, reflecting the limited macrophage presence.

SMAD3 and GAL1 are primarily expressed in SMCs (Fig. 5C, D). SMAD3 shows high expression in the region between the media and the atherotic core, where modulated SMCs are located. It is also sparsely expressed in SMCs within the fibrous cap and media. Notably, SMAD3 expression is most prominent in atherosclerotic plaques of AHA grades IV and V, with reduced expression in grade VI plaques. In type VI plaques, modulated SMCs are largely replaced by macrophages, likely accounting for the lower SMAD3 expression. GAL1 exhibits variable expression across lesion types. In type IV lesions, GAL1 is mainly expressed in SMCs near the lipid core. In type V lesions, GAL1 expression remains high in media SMCs but is also upregulated in macrophages. Sparse expression is seen in SMCs within the fibrous cap. In type VI lesions, GAL1 is highly expressed at the plaque shoulder, where macrophages are concentrated, and in the region between the necrotic core and media, where modulated SMCs reside. These modulated SMCs, however, have lost the α -SMA marker, indicating a shift toward a more macrophage-like phenotype.

TREM2, which co-localizes almost exclusively with CD68, is specifically expressed in macrophages (Fig. 5E). TREM2 expression is low in type IV plaques, likely due to the relatively low macrophage content. In contrast, TREM2 expression is significantly higher in types V and VI plaques, reflecting the increased macrophage presence. TREM2-positive macrophages are dispersed throughout the lipid core, with higher concentrations at the plaque shoulder, particularly near the lumen, where they may influence plaque stabilization. Overall, each protein exhibits a distinct spatial expression pattern, closely associated with the cell types they exist and the regulation of their levels at different plaque stages.

Integrated analysis of the carotid plaque single-cell transcriptome atlas

To explore the cellular mechanisms influencing carotid plaque advancement, we analyzed scRNA-seq data derived from human carotid plaque samples, including atherosclerotic core (AC) plaques and patient-matched proximal adjacent (PA) regions ($n = 3$), yielding a total

of 51,981 cells [12]. After rigorous quality control (see “Methods” and Additional file 2: Fig. S4a-c), we removed contaminants and doublets and performed unsupervised clustering, identifying 21 distinct cell groups (Additional file 2: Fig. S4 d). These were further classified into seven major cellular subsets: smooth muscle cells (SMCs), macrophages, T cells, B cells, NK cells, mast cells, and endothelial cells, which were visualized using UMAP (Fig. 6A). Figure 6B presents a dot plot showing marker genes for each cell type.

scRNA-seq provides information on the cellular mRNA expression patterns of genes, which differs from proteomics that focuses on protein expression levels. The five key proteins identified through proteomics analysis exhibited distinct patterns of cellular expression (Fig. 6C). CD44 exhibited widespread expression across various immune cell types, including T cells, B cells, NK cells, macrophages, and mast cells (Fig. 6D). Intriguingly, its expression was notably elevated in mast cells within the atherosclerotic core, compared to the peripheral plaque region. As expected, TREM2 was exclusively expressed in macrophages and showed increased expression in the atherosclerotic core, aligning with proteomics and IHC findings (Fig. 6E).

In contrast, IL6R, identified as downregulated in AHA type VI plaques, showed reduced expression in macrophages within the plaque core (Fig. 6F). SMAD3 is primarily expressed in SMCs, with minor expression in endothelial cells and mast cells. Its expression significantly decreased in SMCs but increased in endothelial cells and mast cells in atherosclerotic core, indicating differential effects across cell types (Fig. 6G). LGALS1 was widely expressed across cell types, with higher levels in SMCs and macrophages in the atherosclerotic core compared to the adjacent region. However, protein expression data indicated lower levels of LGALS1 in AHA type VI plaques compared to AHA type IV (Fig. 6H). This discrepancy may be due to the different differences in mRNA and protein expression, detection methods, or individual variability. Furthermore, it is important to note that the atherosclerotic core (AC) may contain plaques of varying stages (AHA types IV–VI), as the original samples were not categorized by pathological presentation. Therefore, the AC-PA comparison primarily reflects

(See figure on next page.)

Fig. 6 Single-cell transcriptome analysis of the hub genes. **A** UMAP plot showing cell types of human carotid plaques. **B** Dot plot of different cell type markers. **C** Violin plots showing the single-cell mRNA expression levels of key genes across different cell types. **D** Dot plot showing the percentage and mRNA expression levels of CD44. **E** Dot plot showing the percentage and mRNA expression levels of TREM2. **F** Dot plot showing the percentage and mRNA expression levels of IL6R. **G** Dot plot showing the percentage and mRNA expression levels of SMAD3. **H** Dot plot showing the percentage and mRNA expression levels of LGALS1. CD44, Cluster of Differentiation 44; SMAD3: Mothers Against Decapentaplegic Homolog 3; TREM2: Triggering Receptor Expressed on Myeloid Cells 2; IL6R: Interleukin-6 Receptor; GAL-1: Galectin-1

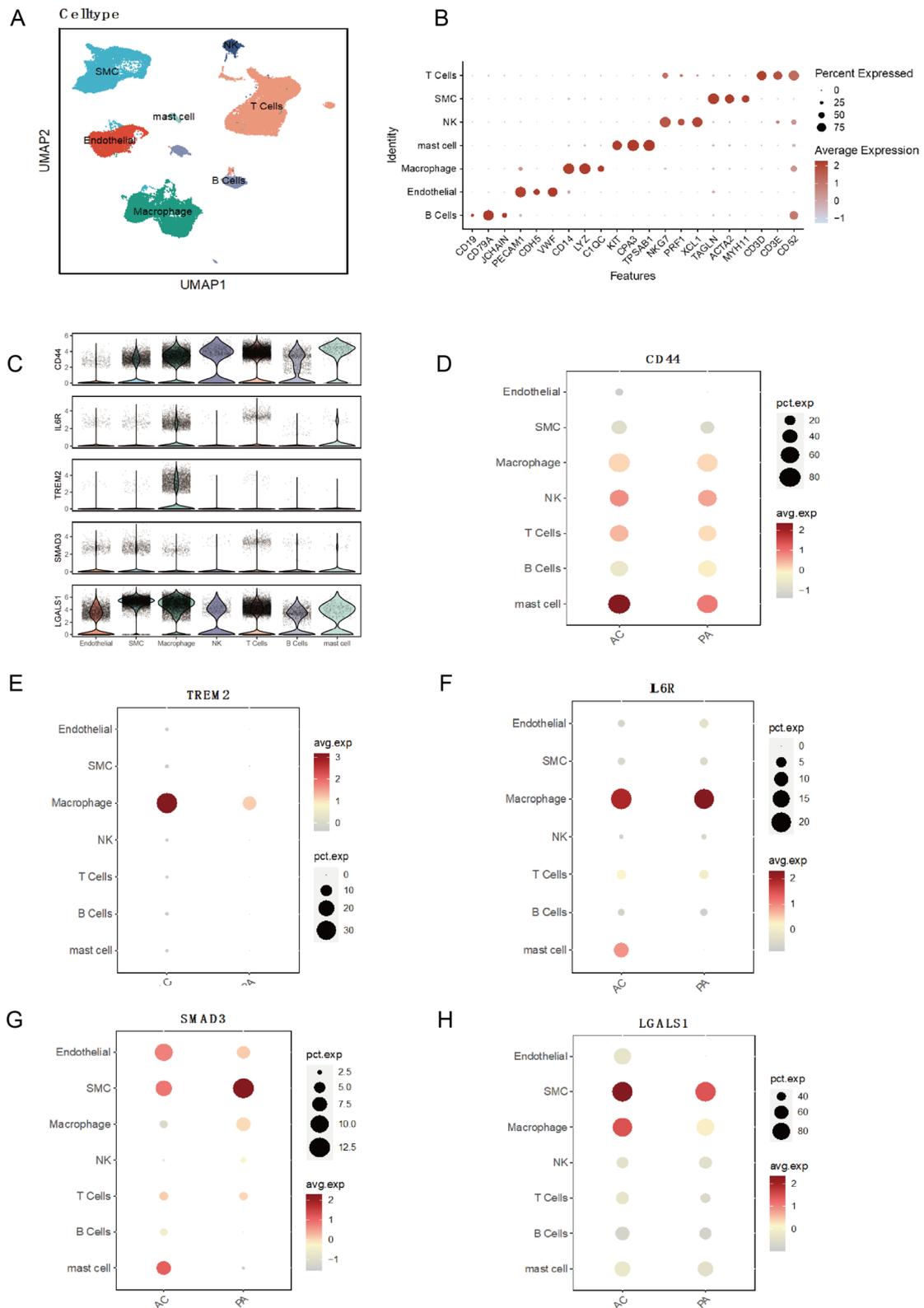


Fig. 6 (See legend on previous page.)

severe versus minor lesions, rather than directly correlating to the AHA grades IV and VI.

Genome-wide MR analysis suggested potential causal effects of key proteins on the risk of ischemic stroke

Mendelian randomization (MR) offers cost-effective analysis of gene expression-outcome causality using SNP summary data from separate datasets [16, 28]. Our analysis of human plaque samples identified CD44, TREM2, IL6R, GAL- 1, and SMAD3 as potential biomarkers for distinguishing between AHA types IV and VI plaques. We hypothesize that these proteins may play a role in driving the transition from AHA type IV to type VI plaques. Since ischemic stroke is often the clinical manifestation of unstable plaques, typically of AHA type VI, we estimated the causal effects of five blood proteins (IL6R, CD44, TREM2, GAL- 1, and SMAD3) on the risk of ischemic stroke using a two-sample MR approach. The analysis flowchart is shown in Fig. 7A.

To explore the causality of these proteins with ischemic stroke, we leveraged summary-level data from two separate datasets. First, we used data from the INTERVAL study [17], which included protein quantitative trait loci (pQTLs) for 2965 blood proteins measured in 3301 European-ancestry individuals, to identify instrumental variables (IVs) for IL6R, CD44, and GAL- 1. Due to the absence of pQTL data for SMAD3 and TREM2 in the INTERVAL dataset, we extracted IVs for these proteins from the UK Biobank Pharma Proteomics Project (UKB-PPP) [18], which included data from 34,557 European ancestry participants and 2922 proteins.

To test for the validity of the selected instrumental variables, we performed a series of sensitivity analyses, including leave-one-out analysis, Cochran's Q test, and the Egger intercept test to detect heterogeneity and horizontal pleiotropy. The Q test revealed significant heterogeneity in the IVs for GAL- 1 ($p < 0.01$), and the Egger intercept test showed significant evidence of horizontal pleiotropy for GAL- 1 ($p = 0.013$). MR-PRESSO and the leave-one-out analysis identified rs10174453 as an outlier, which was subsequently excluded from the analysis. After excluding this SNP, the remaining IVs for GAL- 1 no longer exhibited significant heterogeneity or pleiotropy. For the other four proteins (IL6R, SMAD3, TREM2,

and CD44), none of the tests indicated significant heterogeneity or pleiotropy. Supplementary Table S6 presents the associations between genetic variants and exposure, as well as between genetic variants and outcome. Leave-one-out plots are presented in Additional file 2: Fig. S5 f-j.

We assessed the causal relationship between these proteins and ischemic stroke using GWAS summary data from MEGASTROKE, a meta-analysis of 40,585 ischemic stroke cases and 406,111 controls of European ancestry [23]. Since all the summary data are derived from populations of European ancestry, the SNP-exposure associations are assumed to be similar across the exposure and outcome datasets. In the inverse variance weighted (IVW) MR analysis, three proteins (IL6R, SMAD3, and CD44) were significantly negatively associated with ischemic stroke (Fig. 7B–E; Additional file 2: Fig. S5a–e; Additional file 7). Specifically, IL6R showed an odds ratio (OR) per 1 SD increase of 0.98 (95% CI 0.96–1.00, $p < 0.001$), indicating a protective effect. After correcting for multiple testing using the false discovery rate (FDR), IL6R remained significantly associated with ischemic stroke (adjusted $p < 0.05$). CD44 showed an OR of 0.95 (95% CI 0.91–0.99, $p = 0.020$), also indicating a protective effect, and this association remained significant after FDR correction (adjusted $p < 0.05$). SMAD3 had an OR of 0.95 (95% CI 0.90–1.00) with a p -value of 0.047, but after FDR correction, the association was no longer significant (adjusted $p = 0.059$).

In the weighted median MR analysis, which is robust to invalid instruments, the effects of IL6R and CD44 were consistent with the IVW MR results. Both proteins remained significantly associated with ischemic stroke (Fig. 7B). However, after correcting for multiple comparisons using FDR, no protein remained significantly associated with ischemic stroke.

Taken together, lower levels of soluble IL- 6R, which reflect upregulated IL- 6 signaling, are associated with an increased risk of ischemic stroke. Additionally, reduced blood expression of CD44 or SMAD3 may elevate ischemic stroke risk. However, it is important to note that protein expression levels may differ between blood and plaque tissues, so the underlying clinical significance and mechanisms should be interpreted with caution.

(See figure on next page.)

Fig. 7 Mendelian randomization (MR) analysis assessing the causal effects of blood expression of key proteins on ischemic stroke. **A** Flowchart of MR and subsequent sensitivity analysis. **B** Forest plot showing the calculated causal effect of blood expression level of TREM2, SMAD3, IL6R, GAL- 1, and CD44 proteins on ischemic stroke (using two methods, IVW and WM). **C–E** Scatter plots showing the effect of instrumental variables from SMAD3 (**C**), IL6R (**D**), and CD44 (**E**) pQTLs on the ischemic stroke. IVW, inverse-variance weighting; WM, weighted median; OR, odds ratio. CD44, Cluster of Differentiation 44; SMAD3: Mothers Against Decapentaplegic Homolog 3; TREM2: Triggering Receptor Expressed on Myeloid Cells 2; IL6R: Interleukin- 6 Receptor; GAL- 1: Galectin- 1

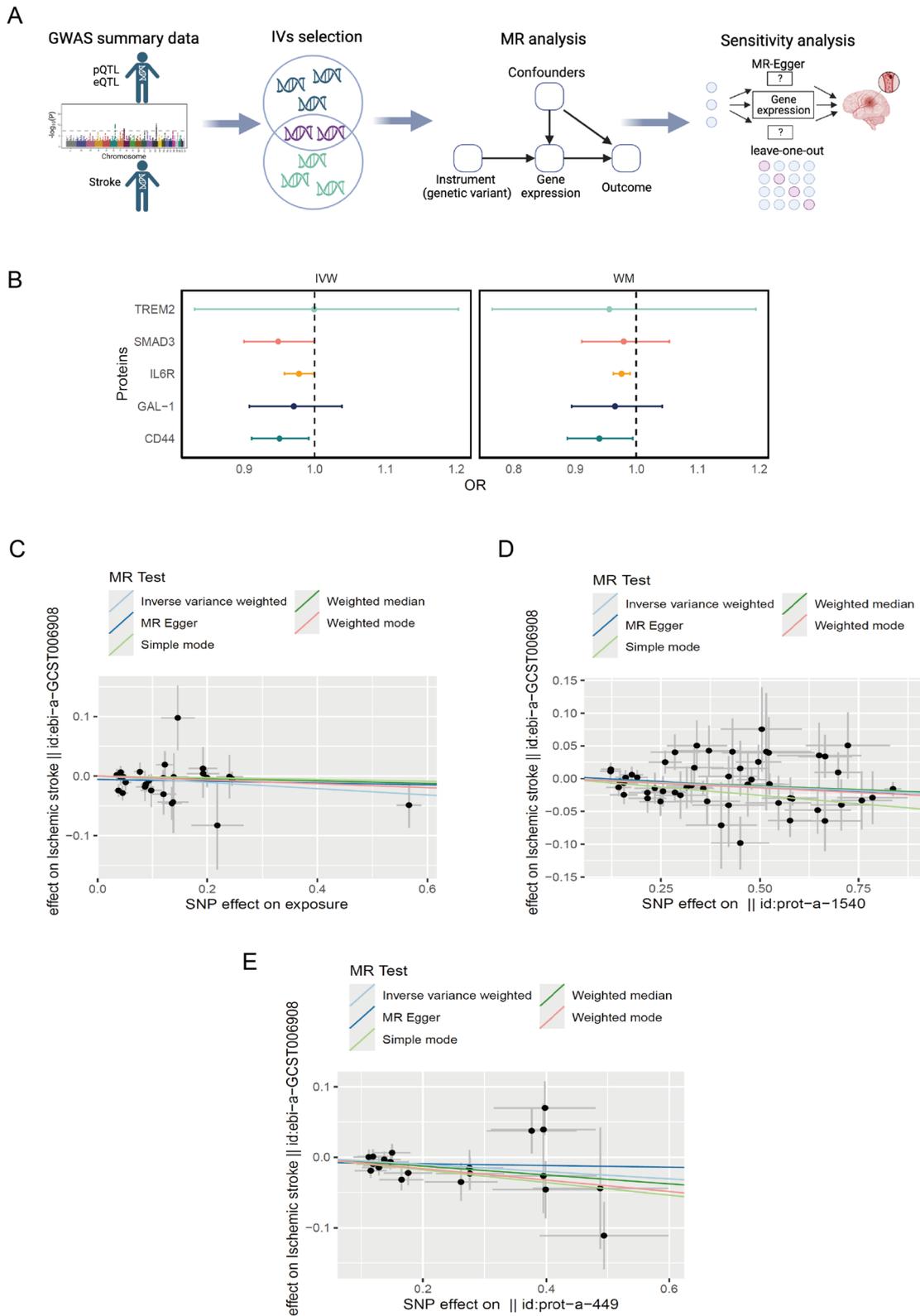


Fig. 7 (See legend on previous page.)

Discussion

Carotid atherosclerosis poses a significant health challenge, often leading to stroke and other associated complications. Understanding the proteomic differences among various plaque types is crucial for developing more targeted therapeutic strategies. In this study, we applied high-throughput proteomics to investigate protein expression patterns within carotid atherosclerotic plaques classified according to the AHA types IV, V, and VI. Our analysis revealed distinct protein changes between plaque types, particularly between types IV and V, and types IV and VI. Notably, the transition from type IV to type VI, which marks the shift from a stable to an unstable plaque, was associated with key proteomic alterations. We identified five hub proteins that play pivotal roles in this transition. Immunostaining confirmed the expression patterns of these proteins, while single-cell transcriptomics provided further insights into their cellular distribution. These findings contribute to our understanding of the molecular mechanisms driving plaque progression and evolution (Fig. 8).

Previous studies have explored the differences in carotid plaque protein composition between symptomatic and asymptomatic patients [10]. In our earlier work, we investigated the proteomic differences between

stable and unstable carotid plaques based on pathological assessments [11]. However, detailed proteomic profiles associated with more precise classifications of carotid plaques have not been fully explored. To address this gap, we utilized the American Heart Association (AHA) classification system to more accurately categorize carotid plaques. This allowed us to classify plaques into three stages: type IV, representing an early stage of advanced plaque where a lipid core is first forming; type V, characterized by a thick fibrous cap, indicating a more stable but still developing lesion; and type VI, the most complex and advanced form, representing late-stage plaques.

In this study, we identified distinct protein expression patterns in carotid atherosclerotic plaques according to AHA classifications, providing new insights into the molecular mechanisms underlying plaque development. Clinical samples offer valuable data that are difficult to replicate accurately in animal models due to inherent species differences and limitations in modeling human disease [29]. Notably, large-cohort proteomics studies have yielded promising results, such as a protein signature's superior predictive value for cardiovascular mortality compared to imaging and histology [10]. Our study and similar proteomics investigations emphasize the complementary role of proteomics alongside RNA,

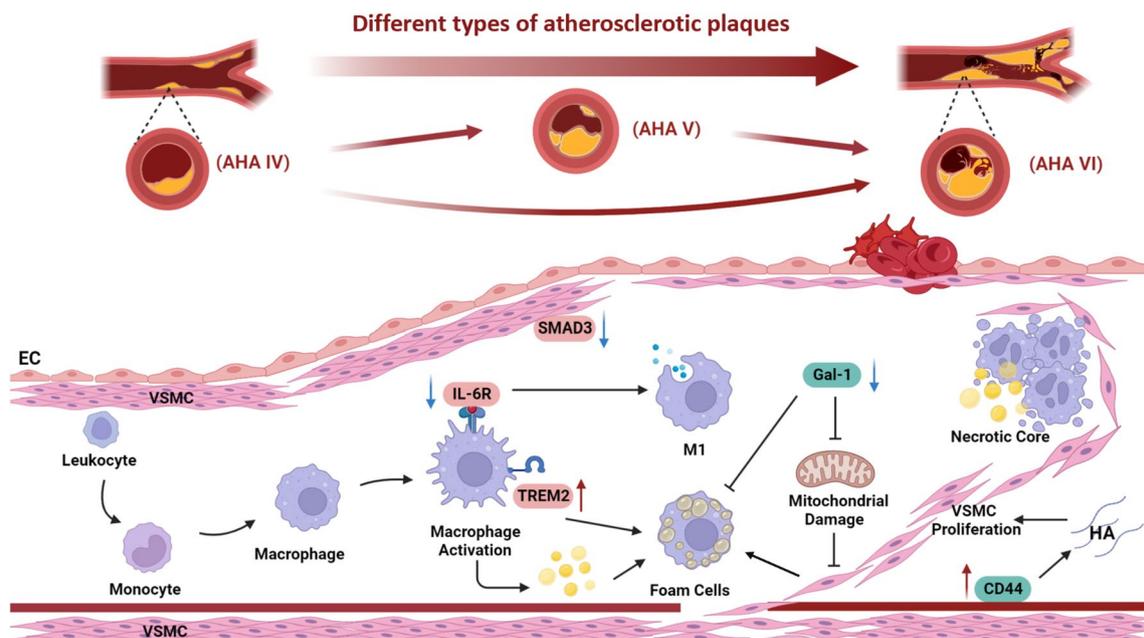


Fig. 8 Hypothetical molecular mechanisms underlying plaque progression from AHA types IV to VI. As the atherosclerotic plaque advances, there is a notable shift in the expression levels of key proteins (CD44, GAL-1, TREM2, SMAD3, and IL6R) associated with inflammatory processes. Specifically, IL-6R and TREM2 play crucial roles in modulating macrophage activation and foam cell formation. The reduction in GAL-1 levels may exacerbate cellular mitochondrial damage. SMAD3 exhibits decreased expression, particularly in vascular smooth muscle cells (VSMCs), where it governs the transition of VSMC phenotype. CD44's interaction with hyaluronic acid (HA) is implicated in regulating VSMC proliferation. AHA, American Heart Association; VSMC, vascular smooth muscle cell; HA, hyaluronic acid; CD44, Cluster of Differentiation 44; SMAD3: Mothers Against Decapentaplegic Homolog 3; TREM2: Triggering Receptor Expressed on Myeloid Cells 2; IL6R: Interleukin-6 Receptor; GAL-1: Galectin-1

histology, and imaging modalities, enhancing our ability to phenotype various plaque types and discover underlying mechanisms [30, 31].

Our enrichment analysis of differential proteins involved in plaque progression among various subtypes revealed their significant engagement in signaling pathways related to metabolism, immunity, cytoskeleton formation, extracellular matrix, cell transport, and cell proliferation during plaque development. These findings align with prior research in carotid artery plaque proteomics. For instance, a prospective study from the Karolinska Carotid Endarterectomy (BiKE) cohort in Sweden, utilizing transcriptome sequencing on 127 symptomatic and 96 asymptomatic plaque samples, identified enriched expression of genes associated with inflammation, cell proliferation, apoptosis, vascular genesis, and extracellular matrix degradation in symptomatic patient plaques [32]. Similarly, a proteomic analysis of mature carotid artery plaques employing DIA-MS technology emphasized the pivotal role of smooth muscle cells (SMCs), particularly in the regulation of cytoskeleton-related proteins [33]. Furthermore, a proteomic analysis of carotid artery plaques from symptomatic and asymptomatic patients underscored the significance and predictive potential of extracellular matrix-related protein markers in cardiovascular risk events, thus enhancing risk prediction and diagnosis in cardiovascular disease management [9]. These consistent findings underscore the robustness of our study results.

We further pinpointed several key proteins (CD44, SMAD3, TREM2, IL6R, GAL-1) when focusing on the transition from type IV to type VI, given that type IV marks the initial stage of advanced plaque development, detectable in clinical settings, and type VI represents an unstable form typically evolving from type IV. The integration of single-cell RNA data and immunostaining data provide spatial and cellular expression pattern of these proteins, indicating potential molecular mechanisms. However, due to the absence of standardized clinical symptom data, it remains unclear whether these protein expression patterns are correlated with patient symptoms.

In our study, we observed an upregulation of CD44 in type VI plaques, which aligns with previous research showing that CD44 levels are significantly higher in atheromatous tissues compared to normal arteries [34]. CD44, a receptor for hyaluronic acid (HA), is widely expressed in endothelial cells, smooth muscle cells (SMCs), and various immune cells. Our results further suggest that CD44 expression correlates with the progression of plaques from a stable to an unstable state, potentially linked to processes such as fibrous cap formation and endothelial erosion [35, 36]. In the apoE^{-/-} mouse model, CD44 expression in both vascular and bone marrow-derived compartments has been shown

to contribute to atherogenesis, indicating that CD44 on both resident and recruited cells may be necessary for its full pro-atherogenic effect in vivo [35, 37]. However, the role of CD44 expression in hematopoietic cells remains controversial, as one study suggested that atherogenesis does not require CD44 expression on hematopoietic cells in the LDL receptor-deficient mouse model [38]. Our study corroborates previous findings by showing CD44 expression in multiple cell types within atheromatous plaques. Interestingly, we also observed notably high CD44 expression in mast cells (Fig. 4D), a novel finding that provides new insights into the functional role of CD44 and warrants further investigation.

SMAD3, a pivotal transcription factor in the TGF- β signaling pathway, has sparked controversy regarding its role in atherosclerosis. While some studies suggest that SMAD3 promotes vascular smooth muscle cell de-differentiation [39, 40], others report contrasting findings. In our DIA-MS analysis, we observed significantly higher SMAD3 expression in AHA type IV plaques compared to AHA type VI plaques. Our MR analysis supports a protective role for SMAD3, in line with prior research. SMAD3's pro-differentiation function inhibits the dedifferentiation necessary for human coronary artery smooth muscle cells to respond to vascular pressure, thereby stabilizing AS plaques [41]. Moreover, recent single-cell transcriptome analysis unveiled that SMAD3 expression in smooth muscle cells (SMCs) suppresses the emergence of specific SMC phenotype transition cells linked to adverse plaque characteristics [42]. Mouse AS models corroborate these findings, indicating that SMC-specific deletion of SMAD3 leads to increased plaque burden, outward remodeling, and heightened vascular calcification [42]. Our research consistently reveals SMAD3's predominant expression in SMCs, which is inhibited in advanced plaques.

The use of two-sample Mendelian randomization (MR) allowed us to analyze the causal relationships and effect sizes between protein expression and disease outcomes. We examined the causal relationships between the five hub proteins identified in our proteomics analysis of carotid plaque and ischemic stroke, as ischemic stroke is the ultimate consequence of carotid plaque progression. Our MR analysis was conducted in accordance with the STROBE-MR guidelines, ensuring transparency and rigor (Additional file 8). The MR analysis revealed that IL6R expression is inversely associated with ischemic stroke risk, a finding that aligns with previous randomized controlled trials (RCTs) showing that the interleukin-6 receptor inhibitor tocilizumab improved myocardial salvage in patients with acute myocardial infarction [43]. This provides further support for the potential of targeting IL6R or its signaling pathway as a

strategy to prevent ischemic stroke in at-risk populations. Furthermore, decreased blood expression of CD44 or SMAD3 may increase the risk of ischemic stroke, indicating that these proteins may have a protective role in the development of carotid plaques.

However, several limitations should be noted. First, our analysis primarily uses summary statistics from studies with individuals of European ancestry, which may limit the generalizability of our findings to other populations. Second, while MR is a powerful tool for assessing causality, it relies on the assumption that genetic instruments are valid and free from violations, such as horizontal pleiotropy. Although we performed sensitivity analyses to detect pleiotropy and heterogeneity, unmeasured pleiotropic effects may still confound the results. Thirdly, while MR identifies causal relationships, it does not provide mechanistic insights into how these proteins influence ischemic stroke risk at the molecular or cellular level. While the single-cell RNA sequencing and immunostaining data offer some clues, further experimental studies are needed to better understand the underlying biological mechanisms. Finally, both the discovery and validation phases of the proteomics study were conducted using a cross-sectional design, which lacked prospective clinical outcome observation.

The integration of single-cell transcriptomics data has limitations. Comparisons between scRNA-seq, proteomics, and immunostaining data are challenging due to differences in grouping criteria. The scRNA-seq samples include plaques at various stages (AHA types IV–VI), and the AC-PA comparison primarily contrasts severe versus minor lesions, rather than directly comparing AHA stage IV lesions to stage VI. Additionally, while scRNA-seq provides gene expression insights, it does not show cell locations within the plaque. Immunofluorescence partially addresses this by visualizing protein expression and key markers like CD68 and α -SMA (Fig. 5), though some cell types remain unmarked.

Conclusions

In conclusion, this study significantly advances our understanding of atherosclerotic plaque progression by comprehensively profiling the proteomic landscape of carotid atherosclerotic plaques. The identification of key proteins, such as IL6R, CD44, and SMAD3, highlights their potential as biomarkers and therapeutic targets, shedding light on new avenues for treatment. Moreover, the integration of single-cell RNA sequencing and Mendelian randomization techniques further strengthens the validity of our findings. This research not only deepens our molecular insight into atherosclerosis but also underscores the precision and reliability of proteomics in prioritizing genes linked to plaque-related traits.

Abbreviations

AHA	American Heart Association
ALT	Alanine aminotransferase
CD44	Cluster of Differentiation 44
Dbil	Direct bilirubin
DEPs	Differentially expressed proteins
DIA	Data-independent acquisition
DPP4	Dipeptidyl Peptidase 4
GAL-1	Galectin-1
IHC	Immunohistochemistry
IL6R	Interleukin-6 Receptor
IPA	Ingenuity Pathway Analysis
MR	Mendelian randomization
MTOR	Serine/Threonine-Protein Kinase mTOR
PPI	Protein-protein interaction
SMCs	Smooth muscle cells
TC	Total cholesterol
TF	Tissue factor
TFR1	Transferrin Receptor Protein 1
TREM2	Triggering Receptor Expressed on Myeloid Cells 2

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12916-025-04058-2>.

Additional file 1. Table S1a–e. Table S1a. Summary of demographics of discovery cohort. Table S1b. Summary of demographics of validation cohort. Table S1c. Clinical Features of All Patients. Table S1 d. Clinical Features of DIA-MS. Table S1e. Clinical Features of Patients for Immunohistochemical Staining. Additional file 2: Figures S1–S6. Figure S1. Proteome quality control results. Figure S2. Differentially expressed proteins between AHA IV, V and VI. Figure S3. IPA pathway analysis, Protein–protein-interaction and pearson correlation analysis of the differentially expressed proteins. Figure S4. Single-cell transcriptomics quality control analysis. Figure S5. Forest plots and leave-one-out plots generated from MR analysis. Figure S6. Negative control for immunohistochemistry and immunofluorescence staining. Additional file 3. Table S2. Carotid plaque proteins quantified by DIA analysis. Additional file 4. Table S3. This table includes quality marker panels to determine the degree of contamination with erythrocytes. Additional file 5. Table S4. Proteome differential analysis of AHA IV/V and VI. Additional file 6. Table S5. PPI network analysis between differentially expressed proteins. Additional file 7. Table S6. MR analysis instrumental variables information. Additional file 8. STROBE-MR checklist.

Acknowledgements

We would like to acknowledge the Biologic Medicine Information Center of China, National Scientific Data Sharing Platform for Population and Health.

Authors' contributions

C.W., Y.F., X.L., Z.L., W.S., and B.L. designed all studies. J.S., K.L., J.C., J.W., K.S., Y.F., and D.K. collected and prepared carotid plaque samples. X.L., H.S., and Z.G. performed proteomics analysis, including LC-MS/MS analyses. C.W., X.L., D.K., Z.G., C.L., Y.L. performed database searching and proteomics downstream and statistical analyses. Y.F. and X.L. conducted immunohistochemistry and immunofluorescence staining. Y.F. carried out single-cell RNA sequencing and Mendelian randomization analyses, with Y.L. verifying and enhancing the Mendelian randomization analysis. J.S., K.L. selected patients and collected their clinical information. Y.F. and C.W. wrote and revised the manuscript. All authors read and approved the final manuscript.

Authors' twitter handles

Twitter handles: @Raina_yuyaoFeng (Yuyao Feng).

Funding

This study received funding from the National Natural Science Foundation of China (82470516, 82170524, and 31901039), Beijing Municipal Natural Science Foundation (L248071), National High Level Hospital Clinical CAMS Innovation Fund for Medical Sciences (CIFMS) (2023-I2M-C&T-B-017, National

Key Research and Development Program of China (2023YFC2415100) and Plastic Medicine Research Fund of Chinese Academy of Medical Sciences (2024-ZX-1-01).

Data availability

All GWAS summary statistics and the single-cell dataset were publicly available, and the sources of these data are described in the Methods section. The codes for MR analysis are deposited in the Github (<https://github.com/Rainna-F/CAS-proteomics-MR>). The proteome data is deposited in iProx (<https://www.iprox.cn/page/project.html?id=IPX0011499000>) with accession number PXD062283. The other raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. Requests to access these datasets should be directed to ZL, WS and BL.

Declarations

Ethics approval and consent to participate

The study was reviewed and approved by the ethics committee of PUMCH (No. JS- 2966), and all methods were performed by approved ethical guidelines. Written informed consent was obtained from all patients. All GWAS summary statistics were publicly available and did not require ethical approval.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Hemangiomas & Vascular Malformations, Plastic Surgery Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China. ²Department of Vascular Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Shuaifuyuan 1, Dongcheng District, Beijing 100730, China. ³Proteomics Research Center, Core Facility of Instrument, Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, School of Basic Medicine Peking Union Medical College, 5 Dong Dan San Tiao, Beijing 100005, China. ⁴State Key Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, Department of Pathophysiology, Peking Union Medical College, Beijing, China. ⁵Eight-Year Program of Clinical Medicine, Peking Union Medical College Hospital, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing, China. ⁶Bioinformatics Facility, National Infrastructure for Translational Medicine, Institute of Clinical Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China. ⁷Department of Emergency Medicine, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China.

Received: 2 June 2024 Accepted: 8 April 2025

Published: 1 May 2025

References

- Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med*. 2005;352(16):1685–95. <https://doi.org/10.1056/NEJMr043430>. PubMed PMID: 15843671.
- Libby P, Ridker PM, Hansson GK. Progress and challenges in translating the biology of atherosclerosis. *Nature*. 2011;473(7347):317–25. <https://doi.org/10.1038/nature10146>. PubMed PMID: 21593864.
- Ungvari Z, Tarantini S, Sorond F, Merkely B, Csiszar A. Mechanisms of vascular aging, a geroscience perspective: JACC Focus Seminar. *J Am Coll Cardiol*. 2020;75(8):931–41. <https://doi.org/10.1016/j.jacc.2019.11.061>. PubMed PMID:32130929;PMCID:PMC8559983.
- Finn AV, Nakano M, Narula J, Kolodgie FD, Virmani R. Concept of vulnerable/unstable plaque. *Arterioscler Thromb Vasc Biol*. 2010;30(7):1282–92. <https://doi.org/10.1161/atvbaha.108.179739>. PubMed PMID: 20554950.
- Gaba P, Gersh BJ, Muller J, Narula J, Stone GW. Evolving concepts of the vulnerable atherosclerotic plaque and the vulnerable patient: implications for patient care and future research. *Nat Rev Cardiol*. 2022. Epub 20220923. <https://doi.org/10.1038/s41569-022-00769-8>. PubMed PMID: 36151312.
- Stary HC. Natural history and histological classification of atherosclerotic lesions: an update. *Arterioscler Thromb Vasc Biol*. 2000;20(5):1177–8. <https://doi.org/10.1161/01.atv.20.5.1177>. (PubMed PMID: 10807728).
- Ge S, Xia X, Ding C, Zhen B, Zhou Q, Feng J, Yuan J, Chen R, Li Y, Ge Z, Ji J, Zhang L, Wang J, Li Z, Lai Y, Hu Y, Li Y, Li Y, Gao J, Chen L, Xu J, Zhang C, Jung SY, Choi JM, Jain A, Liu M, Song L, Liu W, Guo G, Gong T, Huang Y, Qiu Y, Huang W, Shi T, Zhu W, Wang Y, He F, Shen L, Qin J. A proteomic landscape of diffuse-type gastric cancer. *Nat Commun*. 2018;9(1):1012. <https://doi.org/10.1038/s41467-018-03121-2>. Epub 20180308 PubMed PMID: 29520031; PMCID: PMC5843664.
- Jiang Y, Sun A, Zhao Y, Ying W, Sun H, Yang X, Xing B, Sun W, Ren L, Hu B, Li C, Zhang L, Qin G, Zhang M, Chen N, Zhang M, Huang Y, Zhou J, Zhao Y, Liu M, Zhu X, Qiu Y, Sun Y, Huang C, Yan M, Wang M, Liu W, Tian F, Xu H, Zhou J, Wu Z, Shi T, Zhu W, Qin J, Xie L, Fan J, Qian X, He F. Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma. *Nature*. 2019;567(7747):257–61. <https://doi.org/10.1038/s41586-019-0987-8>. Epub 20190227. PubMed PMID: 30814741.
- Langley SR, Willeit K, Didangelos A, Matic LP, Skroblin P, Barallobre-Barreiro J, Lengquist M, Rungger G, Kapustin A, Kedenko L, Molenaar C, Lu R, Barwari T, Suna G, Yin X, Iglseider B, Paulweber B, Willeit P, Shalhoub J, Pasterkamp G, Davies AH, Monaco C, Hedin U, Shanahan CM, Willeit J, Kiechl S, Mayr M. Extracellular matrix proteomics identifies molecular signature of symptomatic carotid plaques. *J Clin Invest*. 2017;127(4):1546–60. <https://doi.org/10.1172/jci86924>. Epub 20170320 PubMed PMID: 28319050; PMCID: PMC5373893.
- Theofilatos K, Stojkovic S, Hasman M, van der Laan SW, Baig F, Barallobre-Barreiro J, Schmidt LE, Yin S, Yin X, Burnap S, Singh B, Popham J, Harkot O, Kampf S, Nackenhorst MC, Strassl A, Loewe C, Demyanets S, Neumayer C, Bilban M, Hengstenberg C, Huber K, Pasterkamp G, Wojta J, Mayr M. Proteomic atlas of atherosclerosis: the contribution of proteoglycans to sex differences, plaque phenotypes, and outcomes. *Circ Res*. 2023;133(7):542–58. <https://doi.org/10.1161/circresaha.123.322590>. Epub 20230830 PubMed PMID: 37646165; PMCID: PMC10498884.
- Lai Z, Wang C, Liu X, Sun H, Guo Z, Shao J, Li K, Chen J, Wang J, Lei X, Shu K, Feng Y, Kong D, Sun W, Liu B. Characterization of the proteome of stable and unstable carotid atherosclerotic plaques using data-independent acquisition mass spectrometry. *J Transl Med*. 2024;22(1):247. <https://doi.org/10.1186/s12967-023-04723-1>. Epub 20240307 PubMed PMID: 38454421; PMCID: PMC10921703.
- Alsaigh T, Evans D, Frankel D, Torkamani A. Decoding the transcriptome of calcified atherosclerotic plaque at single-cell resolution. *Commun Biol*. 2022;5(1):1084. <https://doi.org/10.1038/s42003-022-04056-7>.
- Alsaigh T, Evans D, Frankel D, Torkamani A. Decoding the transcriptome of calcified atherosclerotic plaque at single-cell resolution. 2020. GSE159677. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=gse159677>.
- Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoerckius M, Smibert P, Satija R. Comprehensive integration of single-cell data. *Cell*. 2019;177(7):1888–902.e21. <https://doi.org/10.1016/j.cell.2019.05.031>. Epub 20190606 PubMed PMID: 31178118; PMCID: PMC6687398.
- Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol*. 2013;178(7):1177–84. <https://doi.org/10.1093/aje/kwt084>. PubMed PMID: 23863760; PMCID: PMC3783091.
- Holmes MV, Ala-Korpela M, Smith GD. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nat Rev Cardiol*. 2017;14(10):577–90. Epub 20170601. <https://doi.org/10.1038/nrcardio.2017.78>. PubMed PMID: 28569269; PMCID: PMC5600813.
- Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P, Oliver-Williams C, Kamat MA, Prins BP, Wilcox SK, Zimmerman ES, Chi A, Bansal N, Spain SL, Wood AM, Morrell NW, Bradley JR, Janjic N, Roberts DJ, Ouweland WH, Todd JA, Soranzo N, Suhre K, Paul DS, Fox CS, Plenge RM, Danesh J, Runz H, Butterworth AS. Genomic atlas of the human plasma proteome. *Nature*. 2018;558(7708):73–9. Epub 20180606. <https://doi.org/10.1038/s41586-018-0175-2>. PubMed PMID: 29875488; PMCID: PMC6697541.
- Sun BB, Chiou J, Traylor M, Benner C, Hsu YH, Richardson TG, Surendran P, Mahajan A, Robins C, Vasquez-Grinnell SG, Hou L, Kvikstad EM, Burren

- OS, Davitte J, Ferber KL, Gillies CE, Hedman Å K, Hu S, Lin T, Mikkilineni R, Pendergrass RK, Pickering C, Prins B, Baird D, Chen CY, Ward LD, Deaton AM, Welsh S, Willis CM, Lehner N, Arnold M, Wörheide MA, Suhre K, Kastenmüller G, Sethi A, Cule M, Raj A, Burkitt-Gray L, Melamud E, Black MH, Fauman EB, Howson JMM, Kang HM, McCarthy MI, Nioi P, Petrovski S, Scott RA, Smith EN, Szalma S, Waterworth DM, Mitnau L, Szustakowski JD, Gibson BW, Miller MR, Whelan CD. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature*. 2023;622(7982):329–38. Epub 20231004. <https://doi.org/10.1038/s41586-023-06592-6>. PubMed PMID: 37794186; PMCID: PMC10567551.
19. Skrivankova VW, Richmond RC, Woolf BAR, Davies NM, Swanson SA, VanderWeele TJ, Timpson NJ, Higgins JPT, Dimou N, Langenberg C, Loder EW, Golub RM, Egger M, Davey Smith G, Richards JB. Strengthening the reporting of observational studies in epidemiology using mendelian randomization (STROBE-MR): explanation and elaboration. *Bmj*. 2021;375:n2233. Epub 2021 1026. n2233. <https://doi.org/10.1136/bmj.n2233>. PubMed PMID: 34702754; PMCID: PMC8546498.
 20. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R, Tan VY, Yarmolinsky J, Shihab HA, Timpson NJ, Evans DM, Relton C, Martin RM, Davey Smith G, Gaunt TR, Haycock PC. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7. Epub 20180530. <https://doi.org/10.7554/eLife.34408>. PubMed PMID: 29846171; PMCID: PMC5976434.
 21. Burgess S, Zuber V, Valdes-Marquez E, Sun BB, Hopewell JC. Mendelian randomization with fine-mapped genetic data: Choosing from large numbers of correlated instrumental variables. *Genet Epidemiol*. 2017;41(8):714–25. Epub 20170925. <https://doi.org/10.1002/gepi.22077>. PubMed PMID: 28944551; PMCID: PMC5725678.
 22. Davies NM, von Hinke Kessler Scholder S, Farbmacher H, Burgess S, Windmeijer F, Smith GD. The many weak instruments problem and Mendelian randomization. *Stat Med*. 2015;34(3):454–68. Epub 20141110. <https://doi.org/10.1002/sim.6358>. PubMed PMID: 25382280; PMCID: PMC4305205.
 23. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Ruttan-Jacobs L, Giese AK, van der Laan SW, Gretarsdottir S, Anderson CD, Chong M, Adams HHH, Ago T, Almgren P, Amouyel P, Ay H, Bartz TM, Benavente OR, Bevan S, Boncoraglio GB, Brown RD Jr, Butterworth AS, Carrera C, Carty CL, Chasman DI, Chen WM, Cole JW, Correa A, Cotlarciuc I, Cruchaga C, Danesh J, de Bakker PIW, DeStefano AL, den Hoed M, Duan Q, Engelter ST, Falcone GJ, Gottesman RF, Grewal RP, Gudnason V, Gustafsson S, Haessler J, Harris TB, Hassan A, Havulinna AS, Heckbert SR, Holliday EG, Howard G, Hsu FC, Hyacinth HI, Ikram MA, Ingelsson E, Irvin MR, Jian X, Jiménez-Conde J, Johnson JA, Jukema JW, Kanai M, Keene KL, Kissela BM, Kleindorfer DO, Kooperberg C, Kubo M, Lange LA, Langefeld CD, Langenberg C, Launer LJ, Lee JM, Lemmens R, Leys D, Lewis CM, Lin WY, Lindgren AG, Lorentzen E, Magnusson PK, Maguire J, Manichaikul A, McArdle PF, Meschia JF, Mitchell BD, Mosley TH, Nalls MA, Ninomiya T, O'Donnell MJ, Psaty BM, Pulit SL, Rannikmäe K, Reiner AP, Rexrode KM, Rice K, Rich SS, Ridker PM, Rosst NS, Rothwell PM, Rotter JJ, Rundek T, Sacco RL, Sakae S, Sale MM, Salomaa V, Sapkota BR, Schmidt R, Schmidt CO, Schminke U, Sharma P, Slowik A, Sudlow CLM, Tanislav C, Tatlisumak T, Taylor KD, Thijs VNS, Thorleifsson G, Thorsteinsdottir U, Tiedt S, Trompet S, Tzourio C, van Duijn CM, Walters M, Wareham NJ, Wassertheil-Smoller S, Wilson JG, Wiggins KL, Yang Q, Yusuf S, Bis JC, Pastinen T, Ruusalepp A, Schadt EE, Koplev S, Björkegren JLM, Codoni V, Civelek M, Smith NL, Tréguët DA, Christophersen IE, Roselli C, Lubitz SA, Ellinor PT, Tai ES, Koener JS, Kato N, He J, van der Harst P, Elliott P, Chambers JC, Takeuchi F, Johnson AD, Sanghera DK, Melander O, Jern C, Strbian D, Fernandez-Cadenas I, Longstreth WT Jr, Rolfs A, Hata J, Woo D, Rosand J, Pare G, Hopewell JC, Saleheen D, Stefansson K, Worrall BB, Kittner SJ, Seshadri S, Fornage M, Markus HS, Howson JMM, Kamatani Y, Dobbie S, Dichgans M. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet*. 2018;50(4):524–37. <https://doi.org/10.1038/s41588-018-0058-3>. Epub 20180312. PubMed PMID: 29531354; PMCID: PMC5968830.
 24. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44(2):512–25. Epub 20150606. <https://doi.org/10.1093/ije/dyv080>. PubMed PMID: 26050253; PMCID: PMC4469799.
 25. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol*. 2016;40(4):304–14. Epub 20160407. <https://doi.org/10.1002/gepi.21965>. PubMed PMID: 27061298; PMCID: PMC4849733.
 26. Geyer PE, Voytik E, Treit PV, Doll S, Kleinhempel A, Niu L, Müller JB, Buchholtz ML, Bader JM, Teupser D, Holdt LM, Mann M. Plasma proteome profiling to detect and avoid sample-related biases in biomarker studies. *EMBO Mol Med*. 2019;11(11):e10427. <https://doi.org/10.15252/emmm.201910427>. Epub 20190930. PubMed PMID: 31566909; PMCID: PMC6835559.
 27. Stary HC, Chandler AB, Dinsmore RE, Fuster V, Glagov S, Insull W, Jr, Rosenfeld ME, Schwartz CJ, Wagner WD, Wissler RW. A definition of advanced types of atherosclerotic lesions and a histological classification of atherosclerosis. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association. *Circulation*. 1995;92(5):1355–74. <https://doi.org/10.1161/01.cir.92.5.1355>. PubMed PMID: 7648691.
 28. Zheng J, Baird D, Borges MC, Bowden J, Hemani G, Haycock P, Evans DM, Smith GD. Recent Developments in Mendelian Randomization Studies. *Curr Epidemiol Rep*. 2017;4(4):330–45. <https://doi.org/10.1007/s40471-017-0128-6>. Epub 20171122. PubMed PMID: 29226067; PMCID: PMC5711966.
 29. Emini Veseli B, Perrotta P, De Meyer GRA, Roth L, Van der Donck C, Martinet W, De Meyer GRY. Animal models of atherosclerosis. *Eur J Pharmacol*. 2017;816:3–13. <https://doi.org/10.1016/j.ejphar.2017.05.010>. Epub 20170505. PubMed PMID: 28483459.
 30. Nurmohamed NS, Kraaijenhof JM, Mayr M, Nicholls SJ, Koenig W, Catapano AL, Stroes ESG. Proteomics and lipidomics in atherosclerotic cardiovascular disease risk prediction. *Eur Heart J*. 2023;44(18):1594–607. <https://doi.org/10.1093/eurheartj/ehad161>. PubMed PMID: 36988179; PMCID: PMC10163980.
 31. Nieddu G, Formato M, Lepedda AJ. Searching for Atherosclerosis Biomarkers by Proteomics: A Focus on Lesion Pathogenesis and Vulnerability. *Int J Mol Sci*. 2023;24(20) <https://doi.org/10.3390/ijms242015175>. Epub 20231014. PubMed PMID: 37894856; PMCID: PMC10607641.
 32. Perisic L, Aldi S, Sun Y, Folkersen L, Razuvaev A, Roy J, Lengquist M, Åkeson S, Wheelock CE, Maegdefessel L, Gabrielsen A, Odeberg J, Hansson GK, Paulsson-Berne G, Hedin U. Gene expression signatures, pathways and networks in carotid atherosclerosis. *J Intern Med*. 2016;279(3):293–308. <https://doi.org/10.1111/joim.12448>. Epub 20151130. PubMed PMID: 26620734.
 33. Hansmeier N, Buttigieg J, Kumar P, Pelle S, Choi KY, Kopriva D, Chao TC. Identification of mature atherosclerotic plaque proteome signatures using data-independent acquisition mass spectrometry. *J Proteome Res*. 2018;17(1):164–76. <https://doi.org/10.1021/acs.jproteome.7b00487>. Epub 20171128. PubMed PMID: 29129081.
 34. Krettek A, Sukhova GK, Schönbeck U, Libby P. Enhanced expression of CD44 variants in human atheroma and abdominal aortic aneurysm: possible role for a feedback loop in endothelial cells. *Am J Pathol*. 2004;165(5):1571–81. [https://doi.org/10.1016/s0002-9440\(10\)63414-1](https://doi.org/10.1016/s0002-9440(10)63414-1). PubMed PMID: 15509527; PMCID: PMC1618684.
 35. Zhao L, Lee E, Zukas AM, Middleton MK, Kinder M, Acharya PS, Hall JA, Rader DJ, Puré E. CD44 expressed on both bone marrow-derived and non-bone marrow-derived cells promotes atherogenesis in ApoE-deficient mice. *Arterioscler Thromb Vasc Biol*. 2008;28(7):1283–9. Epub 20080417. <https://doi.org/10.1161/atvbaha.108.165753>. PubMed PMID: 18420998.
 36. Kolodgie FD, Burke AP, Farb A, Weber DK, Kutys R, Wight TN, Virmani R. Differential accumulation of proteoglycans and hyaluronan in culprit lesions: insights into plaque erosion. *Arterioscler Thromb Vasc Biol*. 2002;22(10):1642–8. <https://doi.org/10.1161/01.atv.0000034021.92658.4c>. PubMed PMID: 12377743.
 37. Cuff CA, Kothapalli D, Azonobi I, Chun S, Zhang Y, Belkin R, Yeh C, Secreto A, Assoian RK, Rader DJ, Puré E. The adhesion receptor CD44 promotes atherosclerosis by mediating inflammatory cell recruitment and vascular cell activation. *J Clin Invest*. 2001;108(7):1031–40. <https://doi.org/10.1172/jci12455>. PubMed PMID: 11581304; PMCID: PMC200948.
 38. Sjöberg S, Eriksson EE, Tivesten A, Carlsson A, Klasson A, Levin M, Borén J, Krettek A. CD44-deficiency on hematopoietic cells limits T-cell number but does not protect against atherogenesis in LDL receptor-deficient mice. *Atherosclerosis*. 2009;206(2):369–74. <https://doi.org/10.1016/j.atherosclerosis.2009.03.002>. Epub 20090319. PubMed PMID: 19345948.

39. Shi X, DiRenzo D, Guo LW, Franco SR, Wang B, Seedial S, Kent KC. TGF- β /Smad3 stimulates stem cell/developmental gene expression and vascular smooth muscle cell de-differentiation. *PLoS One*. 2014;9(4):e93995. <https://doi.org/10.1371/journal.pone.0093995>. Epub 20140409 PubMed PMID: 24718260; PMCID: PMC3981734.
40. Zhu Y, Takayama T, Wang B, Kent A, Zhang M, Binder BY, Urabe G, Shi Y, DiRenzo D, Goel SA, Zhou Y, Little C, Roenneburg DA, Shi XD, Li L, Murphy WL, Kent KC, Ke J, Guo LW. Restenosis inhibition and re-differentiation of TGF β /Smad3-activated smooth muscle cells by resveratrol. *Sci Rep*. 2017;7:41916. <https://doi.org/10.1038/srep41916>. Pub 20170206. PubMed PMID: 28165488; PMCID: PMC5292946.
41. Iyer D, Zhao Q, Wirka R, Naravane A, Nguyen T, Liu B, Nagao M, Cheng P, Miller CL, Kim JB, Pjanic M, Quertermous T. Coronary artery disease genes SMAD3 and TCF21 promote opposing interactive genetic programs that regulate smooth muscle cell differentiation and disease risk. *PLoS Genet*. 2018;14(10):e1007681. <https://doi.org/10.1371/journal.pgen.1007681>. Epub 20181011 PubMed PMID: 30307970; PMCID: PMC6198989.
42. Cheng P, Wirka RC, Kim JB, Kim HJ, Nguyen T, Kundu R, Zhao Q, Sharma D, Pedroza A, Nagao M, Iyer D, Fischbein MP, Quertermous T. Smad3 regulates smooth muscle cell fate and mediates adverse remodeling and calcification of the atherosclerotic plaque. *Nat Cardiovasc Res*. 2022;1(4):322–33. <https://doi.org/10.1038/s44161-022-00042-8>. Epub 20220413. PubMed PMID: 36246779; PMCID: PMC9560061.
43. Broch K, Anstensrud AK, Woxholt S, Sharma K, Tøllefsen IM, Bendz B, Aakhus S, Ueland T, Amundsen BH, Damås JK, Berg ES, Bjørkelund E, Bendz C, Hopp E, Kleveland O, Stensæth KH, Opdahl A, Kløw NE, Seljeflot I, Andersen G, Wiseth R, Aukrust P, Gullestad L. Randomized trial of interleukin-6 receptor inhibition in patients with Acute ST-segment elevation myocardial infarction. *J Am Coll Cardiol*. 2021;77(15):1845–55. <https://doi.org/10.1016/j.jacc.2021.02.049>. PubMed PMID: 33858620.

Publisher's Note

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