Heliyon 10 (2024) e32692

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Association of angiogenesis-associated genes with atherosclerotic plaque progression, intraplaque hemorrhage, and immune infiltration

Quanyou Chai [a,b,c], Chunling Guo [a], Long Li [a], Jimin Cao [b,**], Huimin Liu [d,***], Zhaoyang Lu [a,b,c,*]

[a] Department of Cardiology, The Second Hospital of Shanxi Medical University, Taiyuan, 030001, China
[b] Key Laboratory of Cellular Physiology at Shanxi Medical University, Ministry of Education, And the Department of Physiology, Shanxi Medical University, Taiyuan, 030001, China
[c] Key Laboratory of Cardiovascular Intervention and Regenerative Medicine of Zhejiang Province, Department of Cardiology, Sir Run Run Shaw Hospital, Zhejiang University, Hangzhou, 310016, China
[d] Liangzhu Laboratory, Zhejiang University Medical Center, Hangzhou, 311100, China

ARTICLE INFO

Keywords:
Atherosclerosis
Angiogenesis
Immune infiltration
Diagnosis
Intraplaque hemorrhage

ABSTRACT

Mounting evidence suggests that intraplaque angiogenesis is associated with the progression of atherosclerotic plaques and the development of intraplaque hemorrhage. The specificity of intraplaque immune cell infiltration may be associated with abnormalities in the structure and function of the nascent capillaries. Here, we analyzed expression levels of angiogenesis-associated genes in early and advanced carotid atheromatous plaque tissues as well as in stable and intraplaque hemorrhage plaques. Expression profiles of advanced arterial plaques based on angiogenesis-associated genes were classified into subtypes by performing a consensus clustering analysis. The correlation between the immune microenvironment of plaques and expression of angiogenesis-associated genes was also explored using the single sample gene set enrichment analysis method and the CIBERSORT algorithm. We identified hub angiogenesis-associated genes showing similar expression patterns throughout plaque adverse progression, and constructed a prediction model using the random forest algorithm. Receiver operating curves were constructed to evaluate efficacy in identification of intraplaque hemorrhage in a plaque. Our results suggest that heterogeneity of angiogenesis-related genes may promote the malignant development of plaques and cause plaque rupture. In conclusion, we propose a model based on expression of angiogenesis-related genes to predict the risk of plaque rupture.

1. Introduction

As a chronic inflammatory condition, atherosclerosis (AS) is a major cause of cardiovascular disease [1]. Acute thrombotic events due to atherosclerotic plaque rupture cause acute arterial vascular occlusion, and contribute to high levels of mortality and long-term

* Corresponding author.
** Corresponding author.
*** Corresponding author.
E-mail addresses: caojimin@sxmu.edu.cn (J. Cao), flysharon@zju.edu.cn (H. Liu), luzhaoyang@zju.edu.cn (Z. Lu).

https://doi.org/10.1016/j.heliyon.2024.e32692
Received 1 July 2023; Received in revised form 5 June 2024; Accepted 6 June 2024
Available online 7 June 2024

Abbreviations

Atherosclerosis (AS)
Intraplaque hemorrhage (IPH)
Therapeutic ultrasound (TUS)
Angiogenesis-associated genes (AAGs)
Molecular Signatures Database (MSigDB)
Gene Expression Omnibus (GEO)
Carotid endarterectomy (CEA)
Protein-Protein Interaction (PPI)
Gene Ontology (GO)
Kyoto Encyclopedia of Genes and Genomes (KEGG)
Principal component analysis (PCA)
Single sample gene set enrichment analysis (ssGSEA)
Gene Set Variant Analysis (GSVA)
Weighted Correlation Network Analysis (WGCNA)
Gene significance (GS)
Module membership (MM)
Random forest (RF)
Receiver operating curve (ROC)
Area under the ROC curve (AUC)
Cumulative Distribution Function (CDF)
Variable Importance Plot (varImpPlot)

disability around the world [2]. Therefore, a good understanding of the various mechanisms associated with plaque stability is essential to develop effective therapies against atherosclerosis. To this end, intraplaque hemorrhage (IPH) was found to be not only a signature of plaque instability but also a predisposing factor of plaque vulnerability [3]. Recent findings also highlighted the vital role of intraplaque angiogenesis in the development of atherosclerotic plaques and IPH formation. Angiogenesis has also been demonstrated to play a key role in rapid plaque progression and plaque vulnerability [4]. Intraplaque neovascularization is closely related to the development of IPH [5]. Angiogenesis occurs within the plaque under conditions of hypoxia, inflammation, and oxidative stress. Neovascularization is prone to rupture, causing intraplaque hemorrhage as well as atherosclerotic plaque rupture, which aggravates luminal narrowing and increases the incidence of cardiovascular events. This neovascularization in turn induces leakage of blood lipids, inflammatory cells, and red blood cells from the lumen, promoting macrophage infiltration, vessel wall thickening, lipid deposition, inflammation, and progression of atherosclerotic lesions [6–9].

HIF-1α, VEGF, and macrophages were previously found to co-exist in hypoxic areas within atherosclerotic plaques in patients. The expression of HIF-1α was found to be positively correlated with VEGF level, the density of neovascularization, and plaque progression [10]. Selective VEGF inhibitors were shown to attenuate the progression of atherosclerotic plaques in rabbit iliac arteries, which is associated with the inhibition of intraplaque revascularization [11]. Therefore, controlling the maturation of neovascularization was proposed to improve plaque stability via inhibition of intraplaque bleeding [12]. Blocking intraplaque revascularization at an early stage may thus avoid IPH and reverse plaque progression. To this end, therapeutic ultrasound (TUS) treatment was found to selectively deplete plaque neovascularization in a mouse model, and improve the stability of vulnerable plaques by reducing erythrocyte extravasation and inflammatory mediator inward flow. However, the effects of TUS treatment have not been investigated in large animals or humans [13]. Treatment of cancer patients with currently available anti-angiogenic agents (mainly anti-VEGF/VEGFR) in clinical trials has yielded adverse effects on the cardiovascular system such as hypertension, myocardial ischemia, and cardiomyopathy [14]. Therefore, identification of the specific mechanisms of intraplaque neovascularization is crucial to facilitate the discovery of new therapeutic targets as well as potential drugs.

The mechanisms involved in the angiogenesis of atherosclerotic plaques are still not fully elucidated. In this regard, an investigation of the expression profiles of angiogenesis-associated genes (AAGs), specifically in plaques, may provide new research directions and potential therapeutic targets for the prevention of cardiovascular events caused by angiogenesis-related plaque progression. Thus, a comprehensive analysis of AAGs in atherosclerotic plaques was performed in this work.

Our analysis was based on the GEO and ArrayExpress database, as well as the set of AAGs from the Molecular Signatures Database (MSigDB). We explored the specificity of angiogenesis-related genes expression throughout the development of atherosclerotic plaques and IPH plaques to identify biomarker genes. In addition, we also investigated the correlation between intraplaque immunocyte infiltration and AAG expression levels. Moreover, a random forest diagnostic model was constructed by screening hub AAGs through a series of bioinformatics analysis steps. Finally, the efficacy of the model was validated.

2. Materials and methods

2.1. Source of transcription data and feature gene set

Based on the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/gds/), we have compiled a set of 32 paired samples of early and advanced carotid atherosclerotic plaques. Each of these samples included the patient's atherosclerotic plaque samples (Stage IV and higher on the Stary scale), as well as one sample of a distant tissue (Stage I and II). The gene expression profiles from GSE43292 [15] were identified using an Affymetrix Human Gene 1.0 ST Array. Dataset GSE163154 [16,17] includes expression data from 43 carotid atherosclerotic plaques by transcriptome sequencing. These plaques were obtained from symptomatic patients who underwent carotid endarterectomy (CEA), and included 16 non-IPH plaques and 27 IPH plaques. All probes have been converted to gene symbols that match the probe annotation file. We further collected 36 AAGs from the hallmark gene set "HALLMARK_ANGIOGENESIS.v7.5.1" [18]. The gene transcription data of 24 stable plaques with advanced atherosclerosis with 24 ruptured plaques (including IPH) samples were acquired from the ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) dataset E-MTAB-2055 as validation of the results. Samples were obtained from carotid plaque specimens from patients undergoing CEA surgery. Two different portions of the same sample were paired. Sections designated as stable plaques were characterized by fibrous cap or pathological intimal thickening, whereas ruptured plaques were those including thrombus and/or intraplaque hemorrhage.

2.2. Expression analysis of AAGs

Expression analysis of AAGs was utilized on early-stage and advanced-stage samples with the limma (version 3.52.1) R (version 4.2.0) package [19]. "lmFit" function was chosen to fit a linear model to a given expression matrix, and "eBayes" function was used to incorporate a Bayesian adjustment to the fitted model obtained in the previous step to improve the accuracy of differential analysis. Genes showing the p-value lower than 0.05 were identified as AS-specific AAGs. Pearson's correlation test was applied to assess the relevance of AS-specific AAGs at the mRNA level.

2.3. Enrichment analysis and building of the Protein-Protein Interaction network

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were performed for AS-specific AAGs using the clusterProfiler package [20] (version 4.4.1) with p-values and q-values both limited to less than 0.05. We imported key genes into the STRING online analysis website (https://www.string-db.org/) for Protein-Protein Interaction (PPI) network analysis and screening.

2.4. Consensus clustering analysis

A total of 32 advanced-stage carotid atherosclerotic plaque specimens in GSE43292 were classified into subtypes using the ConsensusClusterPlus package (version 1.60.0) based on 19 AAGs with unsupervised k-means clustering analysis [21]. The function "ConsensusClusterPlus" was used with the following parameters: maxK = 7, pItem = 0.8, distance = "euclidean", clusterAlg = "km", reps = 50. Principal component analysis (PCA) was also performed to observe sample separation.

2.5. Analysis of immune cell infiltration

Immune infiltration was assessed in plaques based on different subtypes using the CIBERSORT [22] algorithm and the single sample gene set enrichment analysis (ssGSEA) method in the Gene Set Variant Analysis (GSVA) [23] package. The CIBERSORT analysis uses the LM22 gene file to estimate the relative percentage of 22 immune cell types in the sample. ssGSEA analysis predicts the enrichment of 28 immune cell subpopulations in the sample based on marker genes [24]. We acquired differential AAGs between subtypes and examined the correlation of genes with immune cell infiltration by Spearman's correlation test.

2.6. Weighted correlation network analysis

The Weighted Correlation Network Analysis (WGCNA) [25] package (version 1.71) was utilized to reveal genetic modules associated with subtype classification. The cutreeStatic function is used to plot sample clustering of subtypes with the parameter cutHeight = 70. The PickSoftThreshold function was used to calculate soft threshold power β values consistent with scale independence and average connectivity. Pearson's correlation coefficients between eigengene and subtype classification were identified for each module. After that, gene significance (GS) and module membership (MM) indicators were obtained by cor function. Genes most strongly associated with the subtype classification of plaques were identified by using a GS $>$ 0.5 and MM $>$ 0.8 as criteria. These key genes were enrichment analyzed using GO and KEGG terms, and PPI results were generated using the STRING (https://www.string-db.org/) database. Cytoscape is a visualization tool for complex biological networks. We used the CytoHubba plugin of Cytoscape (version 3.9.1) to calculate and analyze the network structure of PPI to identify biomarker genes. We applied the MCC algorithm via CytoHubba to identify relevant proteins from the PPI network. The top 10 genes were selected as the most significant biomarker genes.

2.7. Identification of hub AAGs and their correlation with immune infiltration of IPH plaques

The microarray data from GSE163154 was analyzed, and differentially expressed AAGs were identified between 16 stable plaques and 27 IPH plaques. We further combined differentially expressed AAGs from early and advanced plaques in the GSE43292 dataset and from the two plaque subtypes to identify hub AAGs. These hub AAGs were eventually identified with similar expression profiles throughout plaque progression and IPH occurrence. The correlation of between AAGs and immunocyte infiltration in plaques was assessed using Spearman's correlation analysis.

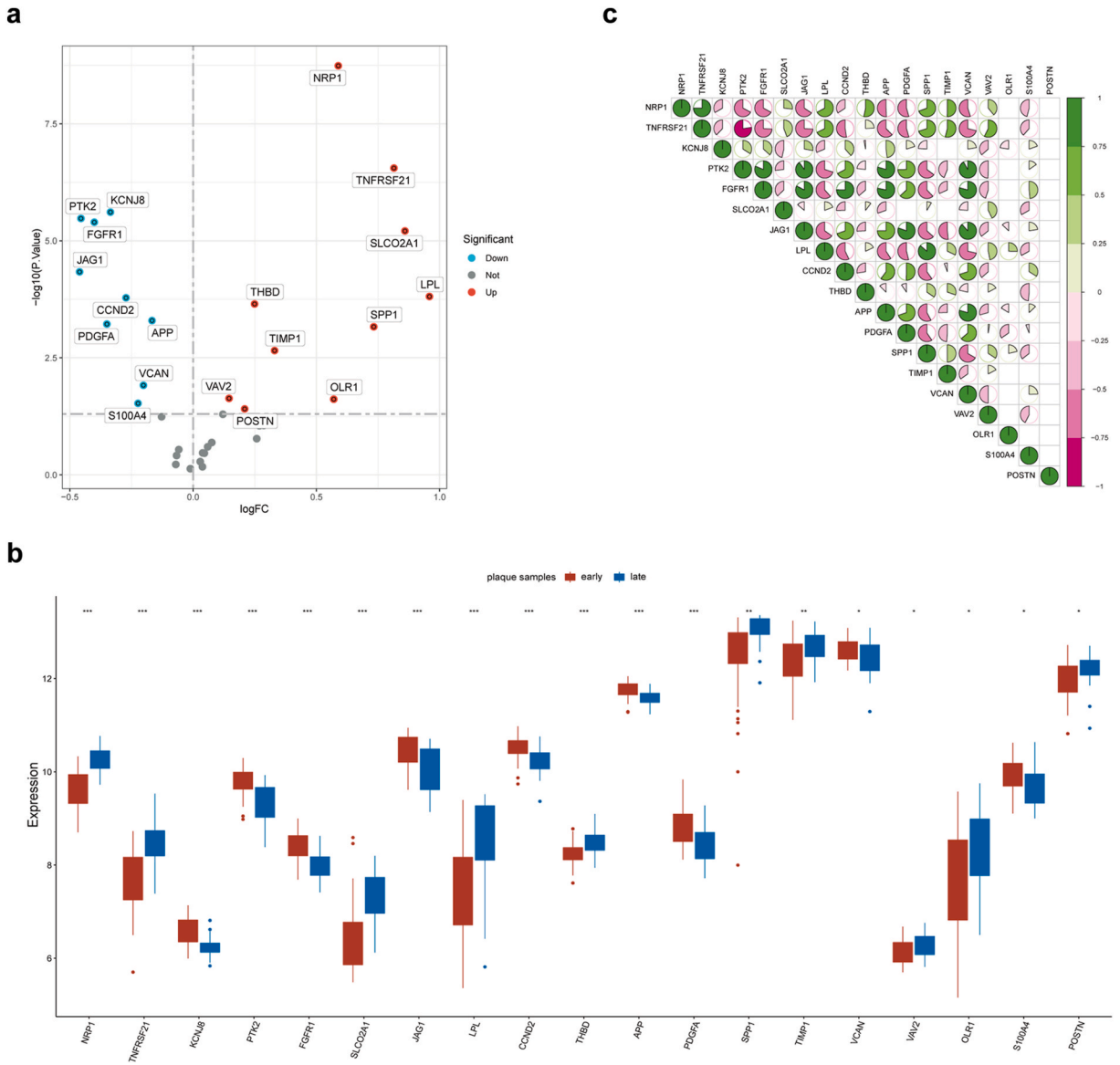


Fig. 1. Identification of AAGs in GSE43292 during atherosclerotic plaque progression. **a** Volcano plots showing expression patterns of AAGs in early and late stage carotid atherosclerotic plaque samples. Red dots indicate up-regulated genes, blue dots indicate down-regulated genes, and gray dots indicate non-significant genes. Differently expressed genes were analyzed for significance using p values < 0.05. **b** Box plots portraying the differentially expressed AAGs in early and late stage plaque samples. Expressions are presented as mean ± standard deviation. Student's t-test was applied to compare the data between the two groups. **c** Pearson's correlation test for the expression levels of AAGs in plaque samples, with positive correlations in green and negative correlations in rose. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

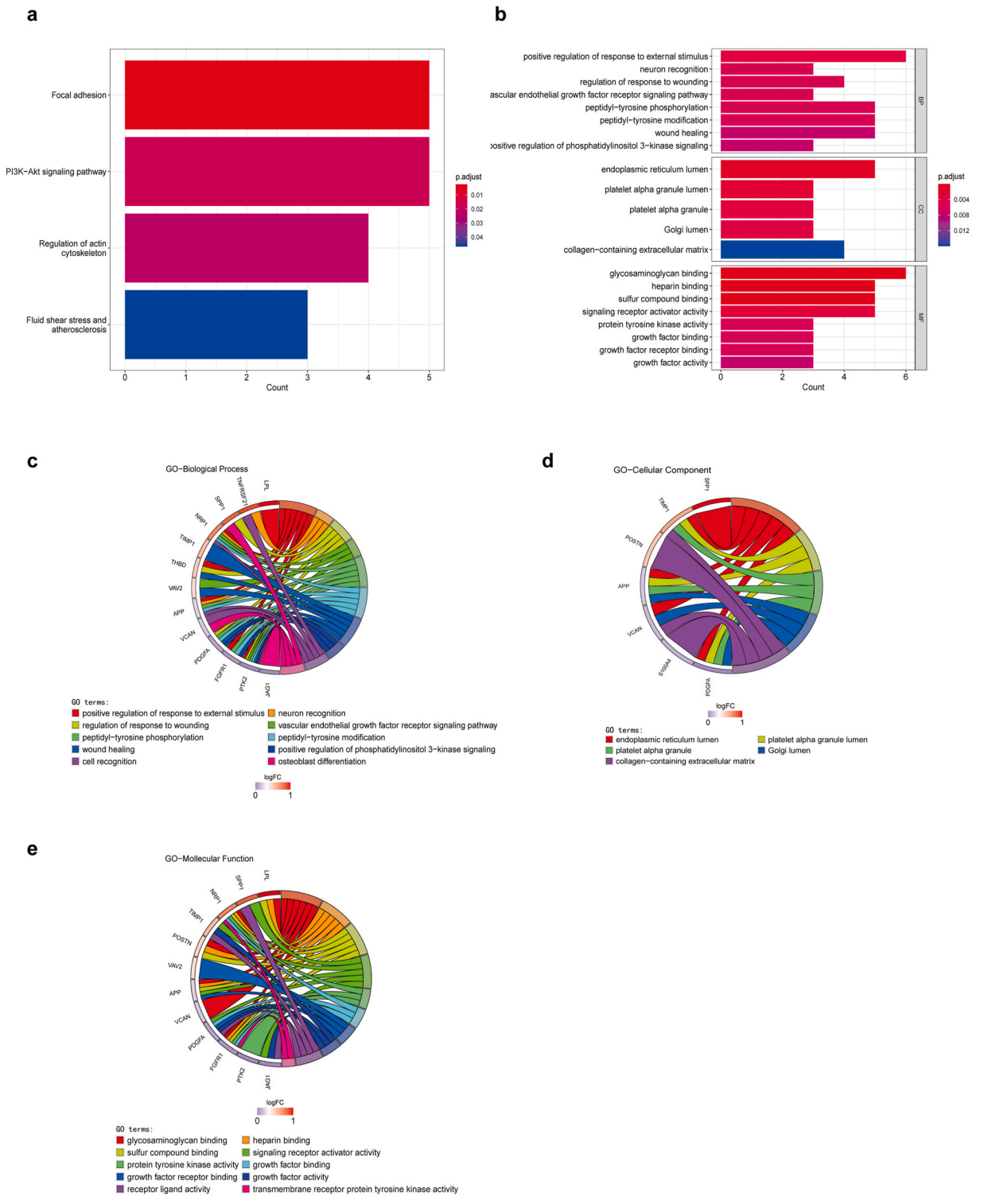


Fig. 2. Enrichment analysis of differentially expressed AAGs in plaque samples. **a, b** KEGG Pathway and GO enrichment analyses of differentially expressed AAGs among two stages in plaque specimens. **c-e** Major GO items enriched of AAGs in atherosclerotic plaque samples.

2.8. Development of a plaque classification model based on hub AAGs

The random forest (RF) algorithm was used to construct a plaque subtype classification model based on expression profiles of hub AAGs. The RF prediction model was developed using the "RandomForest" [26] R package (version 4.7–1.1). A receiver operating curve (ROC) was generated and the area under the ROC curve (AUC) was calculated to assess the predictive power of the model for prediction of the risk of intraplaque hemorrhagic events in the dataset GSE163154. The ArrayExpress database E-MTAB-2055 provided cross-validation of hub AAG-specific expression trends in IPH plaques.

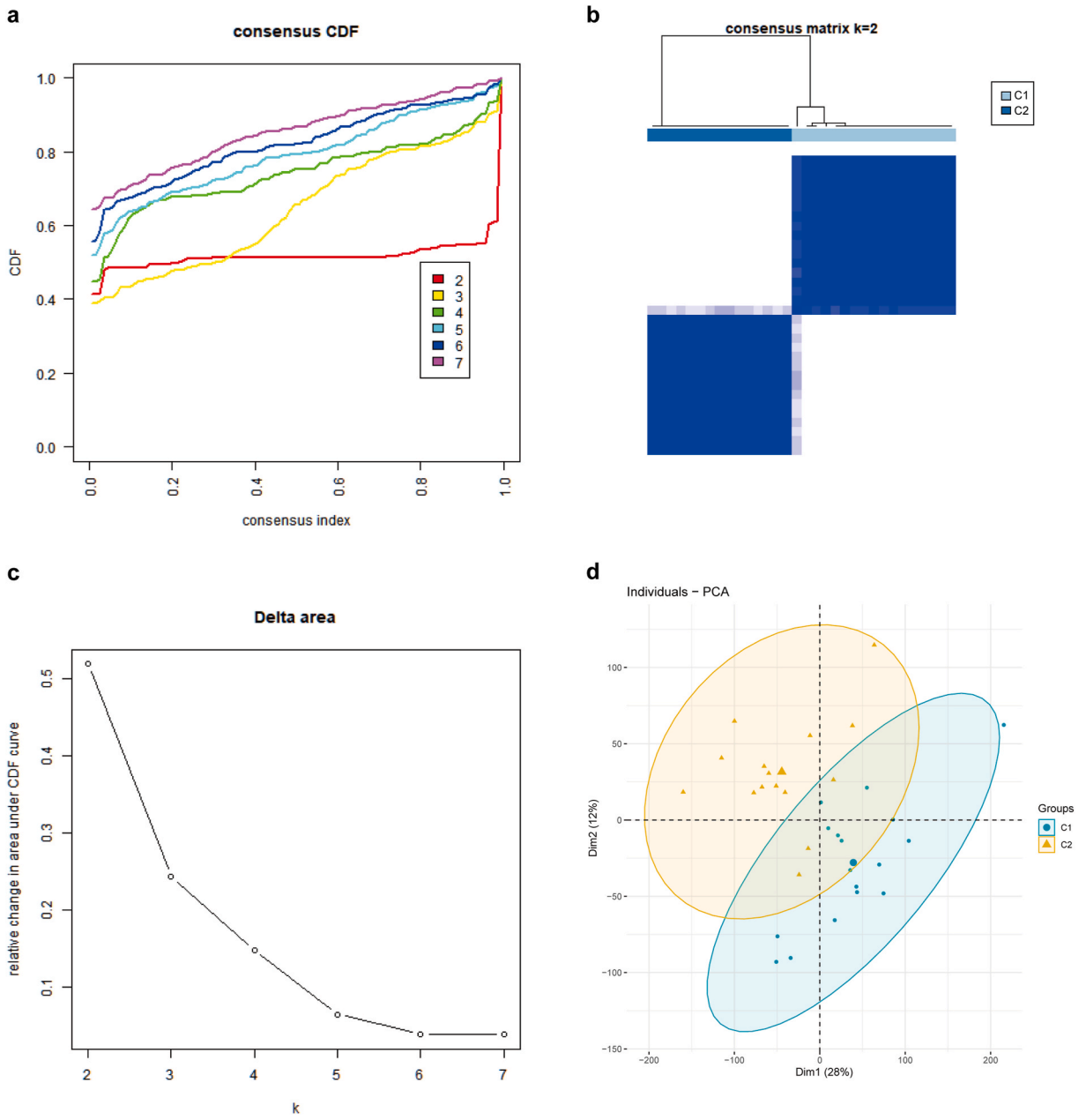


Fig. 3. Advanced-stage plaque samples were divided into two clusters based on expression profiles of AAGs. **a** The cumulative distribution function (CDF) plots showing the different results when k = 2–7. **b** Consensus matrix diagram portraying the consensus values with k = 2. **c** Relative variations of the area under the CDF curve. **d** PCA plots present late-stage atherosclerotic plaque specimens classified as two subtypes (C1, C2) according to AAGs expression pattern.

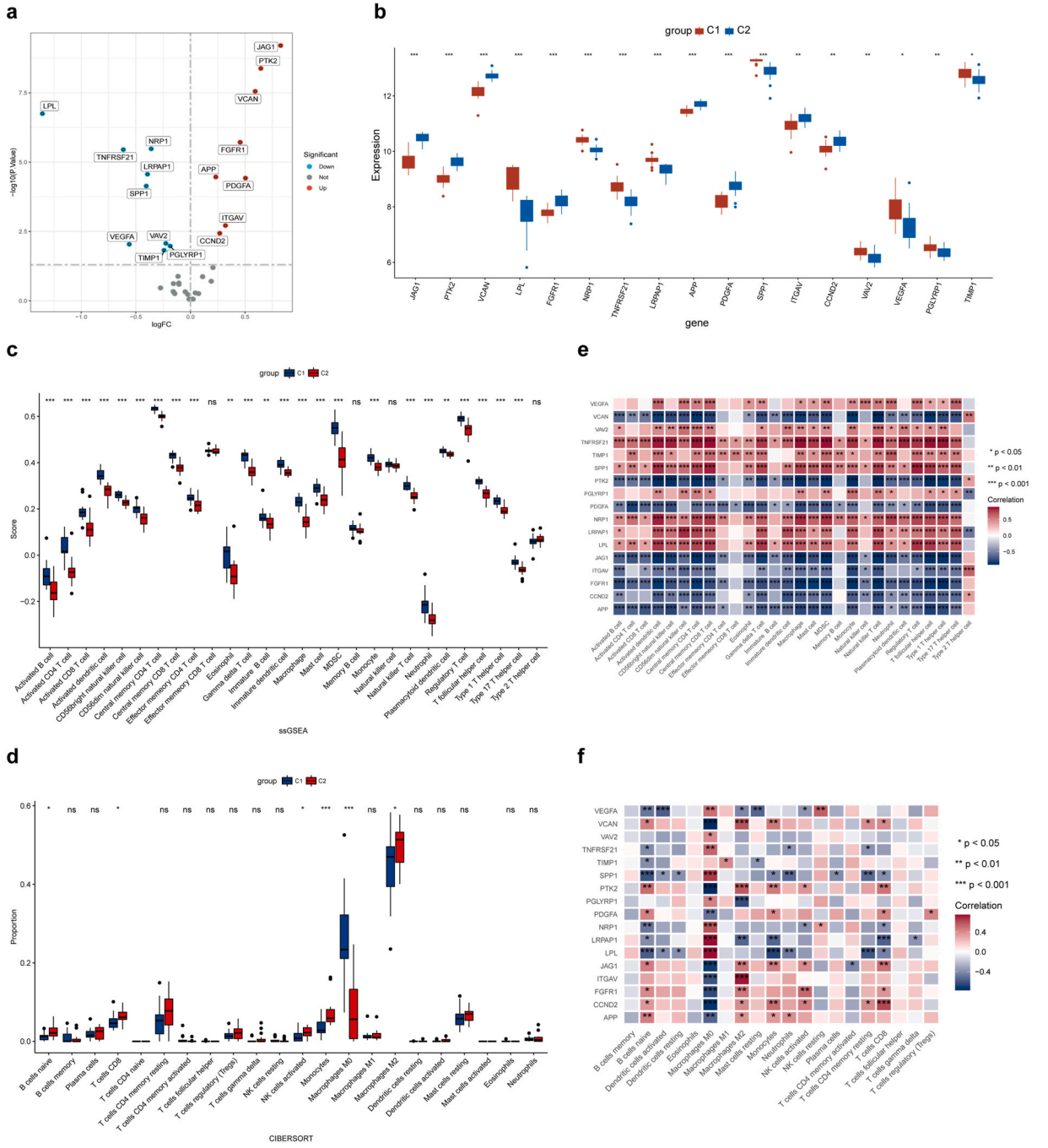
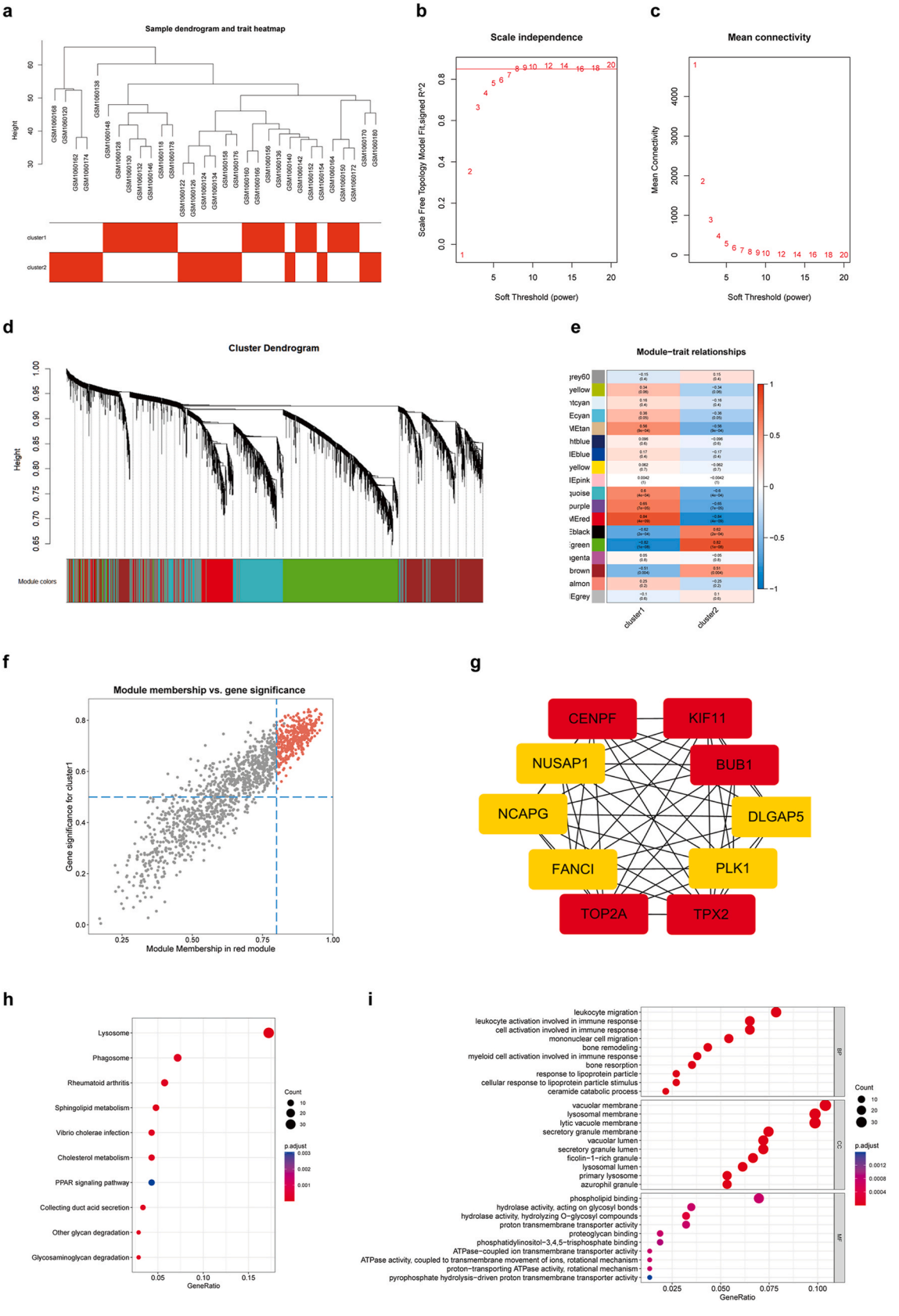


Fig. 4. Differential expression of AAGs and differential immunocyte infiltration among the two subtypes of plaques. **a** Volcano plot portraying the differential expression trends of AAGs between subtypes C1 and C2. **b** Box plots depicting the differences of AAGs in plaque samples. **c** Box plots depict the differences in immunocyte infiltration among C1 and C2 subtypes derived by ssGSEA method. **d** Box plots represent the differences of the immunological microenvironment between C1 and C2 subtypes obtained by using the CIBERSORT method. **e** Heatmaps showing the associations of immunocyte infiltration of ssGSEA with differentially expressed AAGs. **f** Heatmaps visualizing the correlations between immune cell infiltration of CIBERSORT and differentially expressed AAGs. *p < 0.05; **p < 0.01; ***p < 0.001.



(caption on next page)

Fig. 5. Analysis of the characteristics in plaque C1 subtype based on WGCNA. **a** Sample dendrogram and corresponding subtype classification. **b, c** Plots showing scale-free topological index and mean connectivity. **d** Gene hierarchical clustering tree and co-expression module based on optimal soft-thresholding powers. **e** Heatmaps of correlation between gene clustering modules and subtype classification features. **f** Scatter plots are constructed to distinguish key genes in the red module based on gene significance and module membership indicators. **g** Significant biomarker genes associated with subtype characteristics were acquired by the MCC method. **h, i** KEGG and GO enrichment analyses of key genes in the red module. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3. Results

3.1. Expression profiles and biological significance of AAGs during plaque progression

A total of 36 AAGs were obtained from the hallmark gene set (HALLMARK_ANGIOGENESIS.v7.5.1). Analysis of AAGs transcript levels in the dataset GSE43292 revealed upregulation of *VAV2*, *POSTN*, *THBD*, *TIMP1*, *OLR1*, *NRP1*, *SPP1*, *TNFRSF21*, *SLCO2A1*, and *LPL* expression in late-stage plaques, whereas *JAG1*, *PTK2*, *FGFR1*, *PDGFA*, *KCNJ8*, *CCND2*, *S100A4*, *VCAN*, and *APP* expression were found to be downregulated (Fig. 1a and b). These genes were considered to be AAGs associated with plaque progression. We detected a strong correlation between expression levels of these genes via Pearson's correlation analysis (Fig. 1c). The results of KEGG pathway enrichment analysis of these characteristic AAGs showed that the genes were involved in the Focal adhesion, PI3K-Akt signaling pathway, Regulation of actin cytoskeleton, and Fluid shear stress and atherosclerosis (Fig. 2a). Hence, these genes were significantly associated with positive regulation of response to external stimulus, endoplasmic reticulum lumen, and glycosaminoglycan binding (Fig. 2b–e).

3.2. Identification of plaque subtypes via consensus clustering analysis based on expression levels of plaque-specific AAGs

Based on 19 plaque-specific AAGs (AAG19), we classified the advanced plaques into two molecular subtypes, called cluster 1 (C1) and cluster 2 (C2), based on a consensus cluster analysis. This method is often used to discover new subtypes of a disease or for comparative analysis of different subtypes. The k-value with the smallest downward slope of the curve in the Cumulative Distribution Function (CDF) plot is usually considered optimal (Fig. 3a). The values in the consensus matrix diagram represent the likelihood of belonging to the same cluster. The colors range from white to dark blue, with the shade of the color indicating the level of cluster consistency (Fig. 3b). The Delta Area Plot shows the relative change in the area under the CDF curve for k and k-1. Larger values indicate a more pronounced improvement in the superiority of the clustering effect at that value of k compared to the clustering effect at k-1 (Fig. 3c). The results of PCA showed significant heterogeneity between subtypes (Fig. 3d). We also derived new inter-subtype differential expression profiles of AAGs (Fig. 4a and b).

3.3. Differential expression of AAGs between subtypes is associated with immune cell infiltration of plaques

We estimated the differences in the immune microenvironment among the two plaque subtypes using the ssGSEA method and the CIBERSORT algorithm. Our results showed higher levels of macrophages, monocytes, natural killer cells, and neutrophils infiltration in cluster 1 (C1) than in cluster 2 (C2) (Fig. 4c). High levels of M0 macrophages and low levels of M2 macrophages were found in cluster 1 (Fig. 4d). Our results confirm the immune heterogeneity between the two subtypes. We further found that these 17 differentially expressed AAGs (AAG17) (*JAG1, PTK2, VCAN, LPL, FGFR1, NRP1, TNFRSF21, LRPAP1, APP, PDGFA, SPP1, ITGAV, CCND2, VAV2, VEGFA, PGLYRP1,* and *TIMP1*) are remarkably correlated with the level of immunocyte infiltration (Fig. 4e and f).

3.4. Recognition of subtype classification-related modules by WGCNA

We used the WGCNA method to find the most relevant biomarker genes for distinguishing between subtypes (Fig. 5a). By calculating the scale-free fit index and the mean connectivity, we identified a rational soft threshold which yielded a scale-free R^2 value higher than 0.85. We finally identified $\beta = 8$ (scale-free R^2 value = 0.851) as the optimal value. (Fig. 5b and c). There were 18 gene modules (Supplementary Table 1) recognized (Fig. 5d). The red module was found to be highly correlated with subtype classification (Fig. 5e). Therefore, we selected this module for further analysis. Genes in this module with GS higher than 0.5 and MM higher than 0.8 were considered as key co-expressed genes (Fig. 5f). By constructing a PPI network (Supplementary Table 2) and using the cytoHubba plugin of Cytoscape, we identified 10 biomarker genes significantly associated with inter-subtype heterogeneity (Fig. 5g). Further KEGG and GO annotation analysis of key co-expressed genes revealed that they were mainly linked to lysosome (Fig. 5h). This suggests that these genes were associated with bone remodeling, lysosomal lumen, and hydrolase activity, hydrolyzing O-glycosyl compounds (Fig. 5i).

3.5. Hub AAGs is highly correlated with the progression of atherosclerotic plaques and the occurrence of IPH

We evaluated the differential immune cell infiltration profiles obtained by ssGSEA between stable and IPH plaque samples in the GSE163154 dataset, and found similar results to the above study between subtypes (Fig. 6a). Therefore, we identified 11 hub AAGs (*NRP1*, *TNFRSF21*, *LPL*, *SPP1*, *TIMP1*, *PTK2*, *FGFR1*, *JAG1*, *CCND2*, *APP*, and *VCAN*) by taking the intersection of 26 differentially

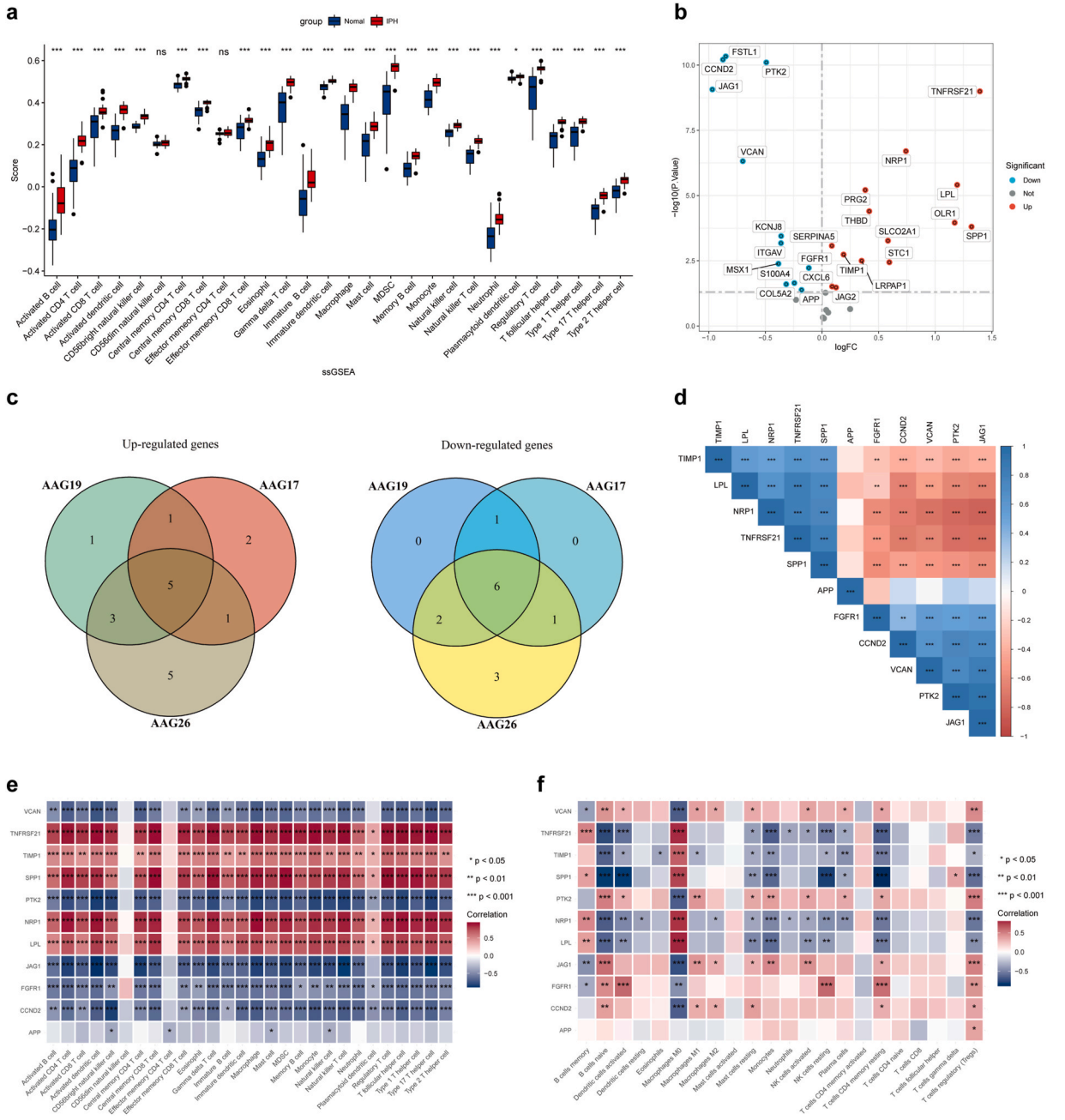


Fig. 6. Association between AAGs and immune infiltrative features in plaque samples from the GSE163154 dataset. **a** Box plots demonstrating the variation in immunocyte infiltration obtained by ssGSEA among non-IPH and IPH plaques in the GSE163154 dataset. **b** Volcano plots illustrating the differential expression trends of AAGs between normal and IPH plaques. **c** Venn plots presenting the hub AAGs among the three sets of genes obtained above. The left side displays the co-upregulated hub AAGs, and the right side shows the co-downregulated hub AAGs. **d** Spearman's correlation test between hub AAGs in the plaque samples. Blue represents positive correlations and red represents negative correlations. **e** Heatmaps presenting the correlation of hub AAGs with immune cell infiltration in plaques using the ssGSEA method. **f** Heatmaps showing the relevance of hub AAGs to immune cell infiltration with CIBERSORT in plaques. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

expressed AAGs (AAG26) (*FSTL1*, *CCND2, PTK2, JAG1, TNFRSF21, NRP1, VCAN, LPL, PRG2, THBD, OLR1, SPP1, KCNJ8, SLCO2A1, ITGAV, SERPINA5, TIMP1, LRPAP1, STC1, MSX1, FGFR1, S100A4, COL5A2, CXCL6, JAG2,* and *APP*) between stable and IPH plaques with the two sets of differentially expressed genes derived from the above-mentioned study (Fig. 6b and c). These genes showed similar expression trends, and were significantly overexpressed during the plaque progression, the C1 subtype in advanced-stage plaques, and the plaques with IPH ($p < 0.05$). These genes were also strongly associated with each other (Fig. 6d). We also noted a strong correlation

between hub AAGs and immunocyte infiltration in plaques in the GSE163154 dataset as well (Fig. 6e and f).

3.6. Subtype classification model constructed based on hub AAGs predicted the occurrence of IPH

To investigate the association of the two molecular subtypes identified above with the generation of IPH plaques, we constructed a random forest model for subtype classification based on hub AAGs. Random Forest is a machine learning algorithm based on decision trees that can be utilized for sample classification or regression tasks. Using this algorithm, the key components that distinguish the difference between two sets of tests can be identified. In the ERROR&TREES graph (Fig. 7a), the "trees" parameter refers to the number of decision trees in the model. If this parameter is too high, the model becomes complex and less efficient. Conversely, if its value is too low, the error rate will also be high. When the "trees" parameter was set to 200, the error rate within the model was found to be

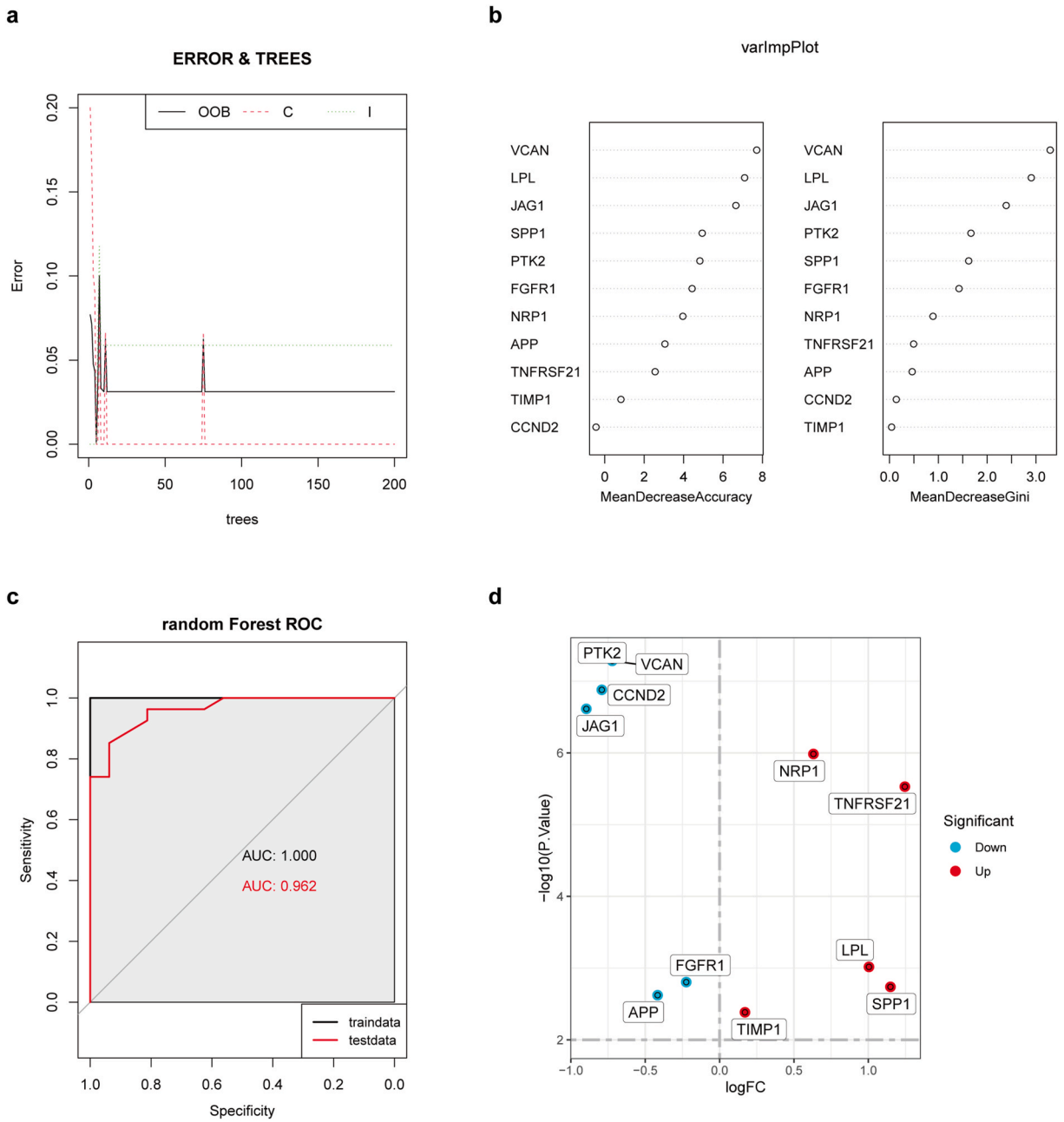


Fig. 7. Construction of a random forest diagnostic model. **a** Random Forest (RF) algorithm to construct a diagnostic model. **b** The RF algorithm evaluates the importance of hub AAGs. **c** ROC curve to assess the diagnostic efficiency of this model. **d** Volcano plots portraying the expression patterns of validated hub AAGs in the third-party dataset E-MTAB-2055. Significance was analyzed using the criterion of p-value < 0.05.

stabilized. Variable Importance Plot (varImpPlot) (Fig. 7b) shows the importance of the variable, including mean decrease accuracy and mean decrease Gini. Larger values indicate higher importance. We found that the AUC values for the model's specificity and sensitivity plot were 1.000 in distinguishing subtypes and 0.964 in identifying whether IPH occurred (Fig. 7c). Finally, we verified the hub AAGs' expression trends with an external dataset E-MTAB-2055 (containing 24 late stable plaques with 24 IPH plaque samples). For the recurring gene symbols in this dataset, we selected the ones with the smallest p-values in their differential expression analysis. The results were found to be consistent with the findings described above (Fig. 7d).

4. Discussion

Atherosclerotic plaque is a common pathological manifestation of atherosclerosis, and plaque rupture is the most serious type of complication in AS [2]. To identify markers that contribute to early diagnosis, we used various bioinformatics analysis approaches on the GSE43292, GSE163154, and E-MTAB-2055 datasets combined with the hallmark gene set to identify markers that may contribute to early diagnosis of AS. Based on the differential expression levels of AAGs, we separated advanced plaques into two molecular subtypes. Immune infiltration analysis resulted in similar immune features to IPH plaques in subtype C1. WGCNA analysis was utilized to obtain additional differentially expressed genes among the two molecular subtypes. Our analysis here yielded 11 hub AAGs that were consistently expressed during both plaque progression and IPH occurrence. Results showed that the hub AAGs-based random forest model had high diagnostic sensitivity and specificity in distinguishing subtypes and whether IPH occurred. Using GO and KEGG enrichment analysis, we found that focal adhesion, PI3K-Akt signaling pathway, actin cytoskeleton regulation, fluid shear stress, and atherosclerotic pathway are involved in the mechanism of angiogenesis-related plaque progression, and differential gene expression among plaque subtypes was majorly enriched in the lysosomal pathway.

Atherosclerosis is a disease characterized by aberrant removal of lipids from vessels, and one of its major pathological bases is the activation of specific cellular immune responses [27]. Inflammation controls all phases of the atherosclerotic plaque life cycle and inflammatory mediators regulate the entire process of plaque destruction and repair, and therefore can be targeted for disease alleviation [28]. In addition, the tissue-specific distribution of immune cells and types contributes to AS progression [29]. An in-depth investigation of immune infiltration in plaque specimens under different pathological conditions may lead to a better understanding of the AS progression. The density of trophoblastic vessels was previously noted to be strongly associated with the number of inflammatory cells within the plaque and is not related to the size of the plaque in ApoE−/− mice atherosclerotic plaques [30]. The immune microenvironment differed significantly among the plaque subtypes obtained in this study and was highly correlated with the expression of AAGs. In C1 subtype plaques, CIBERSORT analysis identified the top three enriched immune cell types (M0, M2 macrophages and CD4 memory resting T cells), in line with the infiltration trend of IPH plaques [31]. In the ssGSEA analysis, C1 showed a higher degree of infiltration of activated CD4 T cells and activated CD8 T cells which was significantly different compared to C2, whereas single-cell sequencing studies revealed a highly activated state of T cells specifically in the plaques. Hence, C1 may suggest a more aggressive state of progression in plaques [32].

GO enrichment results in premature and advanced plaques yielded that AAGs are mainly associated with focal adhesion, PI3K-Akt signaling pathway, regulation of actin cytoskeleton, fluid shear stress and atherosclerosis pathway. Various adhesion molecules are closely associated with intraplaque angiogenesis in atherosclerotic plaques, and thus may contribute to atherosclerotic plaque progression and instability [33]. Numerous studies have demonstrated that the PI3K-Akt signaling pathway is involved in the body's metabolism, cell survival, growth and value-added, as well as in angiogenesis within tissues. Calcium-activated proteases (Calpains) present in endothelial cells were also shown to activate the PI3K/AMPK/Akt/eNOS pathway through the induction of VEGF, leading to neovascularization within atherosclerotic plaques [34]. We identified key genes involved in molecular subtype heterogeneity in advanced plaques by WGCNA and found them to be significantly enriched in the lysosomal pathway. Lysosomes are involved in almost every step of atherosclerosis. Thus, lysosomal dysfunction caused by ox-LDL or lipase downregulation may contribute to atherosclerotic lesions by regulating autophagy, inflammasome, apoptosis, and lysosomal biogenesis. All of these factors may promote the development of atherosclerotic plaques [35].

Here, we identified 11 hub AAGs that are involved in a multitude of physiological and pathological processes in the human body. Among these, *Nrp1* is involved in postnatal angiogenesis and arteriogenesis in the heart and retina, pathological neovascularization in the retina, and angiogenesis-dependent tumor growth [36]. Studies have shown that TNFRSF21 induces endothelial mesenchymal transition (EndoMT) in coronary artery endothelial cells (CAECs) in type 2 diabetics, leading to enhanced permeability of CAECs [37]. Lipoprotein lipase (LpL)-associated lipid delivery may also alter endothelial barrier function and the operation of paracellular pathways around endothelial cells [38]. SPP1 and TIMP1 were shown to play a role in the inflammatory response during the acute phase of cerebral ischemia-reperfusion injury [39]. Inhibition of PTK2 (FAK) activity leads to decreased endothelial cell permeability [40] and reduced macrophage motility [41]. FGFR1 can promote angiogenesis and vascular endothelial growth factor secretion through the MAPK pathway, and FGFR1 can further enhance VEGF secretion by activating AKT signaling [42]. In contrast, in clear cell renal carcinoma, JAG1 promotes angiogenesis by increasing VEGFR3 and antagonizing notch signaling [43]. CCND2 is an important regulator in reducing oxidized-LDL-induced endothelial cell damage and increasing cell viability [44]. In addition, several animal experiments have also confirmed that loss of APP leads to endothelial diastolic dysfunction and reduced responsiveness of vascular smooth muscle cells to NO [45]. The direct local action of VCAN on blood endothelial progenitor cells has an integral role in angiogenesis and primitive hematopoiesis [46].

Hub AAGs were identified based on the combination of plaque progression, differential expression of genes among subtypes, and the occurrence of IPH in plaques. These genes are regarded as key genes to distinguish molecular subtypes during plaque progression and to identify the occurrence of IPH. To investigate their predictive performance, we used a random forest algorithm to construct a

model for subtype classification to determine their predictive probability for the occurrence of IPH in plaques. The resulting ROC curves demonstrated significant discriminatory power. The constructed model thus provides a new method for determining the probability of IPH during plaque development. The differential expression of hub AAGs was also consistent with these results after an independent external dataset validation was performed. This verification excluded as much error as possible due to the small sample size. In summary, our analysis features the entire process of plaque progression and IPH in the specific expression of AAGs, indicating that the occurrence of plaque-related cardiovascular events may be inseparable from the persistence of specific AAGs expression.

Atherosclerotic plaque formation is closely related to macrophage migration and lipid accumulation. We found that the vast majority of hub AAGs were significantly associated with M0 macrophages. Among them, the genes *VCAN*, *PTK2*, *JAG1*, *FGFR1*, and *CCND2*, which were downregulated during ongoing plaque progression, showed a significant negative correlation with M0-type macrophages, whereas the hub AAGs with increasing expression (*TNFRSF21*, *TIMP1*, *SPP1*, *NRP1*, and *LPL*) were found to be positively correlated. This suggests that with the adverse progression of plaques and the accumulation of macrophages, the respective expression trends of hub AAGs were simultaneously enhanced. M0-type macrophages can be readily activated in response to stimulation by inflammatory factors. For one thing, they will continue to promote the formation of atheromatous plaques, and for another, they may cause small inflammatory storms within the plaque under some occasional stimulation, which may be related to the occurrence of IPH. LPL allows macrophages to phagocytose lipids to transform them into foam cells, which directly promotes atherosclerosis. LPL expression increases during plaque progression, in line with our analysis results. The expression level of *VCAN* is a major parameter in our IPH prediction model. As an extracellular matrix component involved in angiogenesis, the steady downregulation of *VCAN* during adverse plaque progression may promote abnormal neovascularization and thus IPH.

There are also several limitations to this study. First, adjusted p-values are often used in differential gene expression analyses to achieve multiple hypothesis testing. However, due to the small differences in gene expression levels among tissues, we used the raw p-values to identify differentially expressed genes in order not to exclude any potential key genes. To avoid false positives, we used three independent IPH datasets to mutually verify whether the differences were indeed significant and whether the predictive models constructed using these genes improved their efficacy. The results of this analysis indicate that this approach does not affect the reliability of the conclusions in this study. Second, because of the inability to collect information on other clinical characteristics of the sample in the public dataset, we were unable to exclude the possible influence of factors such as patient's age and race on this study. Finally, the small sample size utilized for this study may limit the generalizability of the findings. Further research is needed with larger sample sizes of IPH plaques. In addition, the exact mechanisms of AAGs in IPH development need to be further investigated with *ex vivo* and *in vivo* experiments.

5. Conclusions

In our study, we revealed that changes in the expression levels of AAGs were strongly associated with malignant progression of atherosclerotic plaques and alterations in the immune microenvironment within plaques. Our findings revealed that heterogeneity in angiogenesis at the mRNA level is associated with plaque rupture. A random forest model based on expression profiles of AAGs was shown to provide a high-quality prediction of whether IPH events occur in advanced plaques. Further studies of the interaction between AAGs and IPH may provide new therapeutic targets for plaque rupture prevention.

Statements & declarations

Data availability statement

The datasets of GSE43292 and GSE163154 are available in the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/gds/) and E-MTAB-2055 is found in the ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) database. The list of angiogenesis-associated genes is obtained from the Molecular Signatures Database (MSigDB; https://www.gsea-msigdb.org/gsea/msigdb/).

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Ethical approval

The patient data analyzed in this work were acquired from publicly available datasets.

CRediT authorship contribution statement

Quanyou Chai: Software, Resources. **Chunling Guo:** Formal analysis. **Long Li:** Formal analysis. **Jimin Cao:** Formal analysis. **Huimin Liu:** Writing – original draft. **Zhaoyang Lu:** Writing – review & editing, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Chinese Natural Science Foundation grants [grant numbers 81800171, 81900275]; the Natural Science Foundation of Shanxi Province [grant number 201801D221273]; the Scientific and Technological Innovation Program of Shanxi Higher Education Institution [grant numbers 201804026, 201804027]; and Shanxi Provincial Commission of Health and Family Planning [grant number 2017053].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e32692.

References

[1] J. Frostegård, Immunity, atherosclerosis and cardiovascular disease, BMC Med. 11 (2013) 117, https://doi.org/10.1186/1741-7015-11-117.
[2] A. Gisterå, G.K. Hansson, The immunology of atherosclerosis, Nat. Rev. Nephrol. 13 (2017) 368–380, https://doi.org/10.1038/nrneph.2017.51.
[3] J.B. Michel, R. Virmani, E. Arbustini, G. Pasterkamp, Intraplaque haemorrhages as the trigger of plaque vulnerability, Eur. Heart J. 32 (1985) 1977–1985, https://doi.org/10.1093/eurheartj/ehr054(2011), 1985b, 1985c.
[4] B. Doyle, N. Caplice, Plaque neovascularization and antiangiogenic therapy for atherosclerosis, J. Am. Coll. Cardiol. 49 (2007) 2073–2080, https://doi.org/10.1016/j.jacc.2007.01.089.
[5] C. Van der Donckt, et al., Elastin fragmentation in atherosclerotic mice leads to intraplaque neovascularization, plaque rupture, myocardial infarction, stroke, and sudden death, Eur. Heart J. 36 (2015) 1049–1058, https://doi.org/10.1093/eurheartj/ehu041.
[6] D.A. Chistiakov, A.N. Orekhov, Y.V. Bobryshev, Contribution of neovascularization and intraplaque haemorrhage to atherosclerotic plaque progression and instability, Acta Physiol. 213 (2015) 539–553, https://doi.org/10.1111/apha.12438.
[7] J.B. Michel, J.L. Martin-Ventura, A. Nicoletti, B. Ho-Tin-Noé, Pathology of human plaque vulnerability: mechanisms and consequences of intraplaque haemorrhages, Atherosclerosis 234 (2014) 311–319, https://doi.org/10.1016/j.atherosclerosis.2014.03.020.
[8] L. Parma, F. Baganha, P.H.A. Quax, M.R. de Vries, Plaque angiogenesis and intraplaque hemorrhage in atherosclerosis, Eur. J. Pharmacol. 816 (2017) 107–115, https://doi.org/10.1016/j.ejphar.2017.04.028.
[9] P.R. Moreno, M. Purushothaman, K.R. Purushothaman, Plaque neovascularization: defense mechanisms, betrayal, or a war in progress, Ann. N. Y. Acad. Sci. 1254 (2012) 7–17, https://doi.org/10.1111/j.1749-6632.2012.06497.x.
[10] J.C. Sluimer, et al., Hypoxia, hypoxia-inducible transcription factor, and macrophages in human atherosclerotic plaques are correlated with intraplaque angiogenesis, J. Am. Coll. Cardiol. 51 (2008) 1258–1265, https://doi.org/10.1016/j.jacc.2007.12.025.
[11] Y. Wang, et al., Gene delivery of soluble vascular endothelial growth factor receptor-1 (sFlt-1) inhibits intra-plaque angiogenesis and suppresses development of atherosclerotic plaque, Clin. Exp. Med. 11 (2011) 113–121, https://doi.org/10.1007/s10238-010-0112-7.
[12] M.R. de Vries, et al., Blockade of vascular endothelial growth factor receptor 2 inhibits intraplaque haemorrhage by normalization of plaque neovessels, J. Intern. Med. 285 (2019) 59–74, https://doi.org/10.1111/joim.12821.
[13] X. Li, et al., Therapeutic ultrasound combined with microbubbles improves atherosclerotic plaque stability by selectively destroying the intraplaque neovasculature, Theranostics 10 (2020) 2522–2537, https://doi.org/10.7150/thno.39553.
[14] N. Maurea, et al., Pathophysiology of cardiotoxicity from target therapy and angiogenesis inhibitors, J. Cardiovasc. Med. 17 (Suppl 1) (2016) S19–S26, https://doi.org/10.2459/jcm.0000000000000377.
[15] H. Ayari, G. Bricca, Identification of two genes potentially associated in iron-heme homeostasis in human carotid plaque using microarray analysis, J. Biosci. 38 (2013) 311–315, https://doi.org/10.1007/s12038-013-9310-2.
[16] H. Jin, et al., Integrative multiomics analysis of human atherosclerosis reveals a serum response factor-driven network associated with intraplaque hemorrhage, Clin. Transl. Med. 11 (2021) e458, https://doi.org/10.1002/ctm2.458.
[17] H. Jin, B.M.E. Mees, E.A.L. Biessen, J.C. Sluimer, Transcriptional sex dimorphism in human atherosclerosis relates to plaque type, Circ. Res. 129 (2021) 1175–1177, https://doi.org/10.1161/circresaha.121.320099.
[18] A. Liberzon, et al., The Molecular Signatures Database (MSigDB) hallmark gene set collection, Cell systems 1 (2015) 417–425, https://doi.org/10.1016/j.cels.2015.12.004.
[19] M.E. Ritchie, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic acids research 43 (2015) e47, https://doi.org/10.1093/nar/gkv007.
[20] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS A J. Integr. Biol. 16 (2012) 284–287, https://doi.org/10.1089/omi.2011.0118.
[21] M.D. Wilkerson, D.N. Hayes, ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking, Bioinformatics 26 (2010) 1572–1573, https://doi.org/10.1093/bioinformatics/btq170.
[22] A.M. Newman, et al., Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (2015) 453–457, https://doi.org/10.1038/nmeth.3337.
[23] S. Hänzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (7) (2013), https://doi.org/10.1186/1471-2105-14-7.
[24] P. Charoentong, et al., Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade, Cell Rep. 18 (2017) 248–262, https://doi.org/10.1016/j.celrep.2016.12.019.
[25] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinf. 9 (2008) 559, https://doi.org/10.1186/1471-2105-9-559.
[26] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[27] J. Viola, O. Soehnlein, Atherosclerosis - a matter of unresolved inflammation, Semin. Immunol. 27 (2015) 184–193, https://doi.org/10.1016/j.smim.2015.03.013.
[28] P. Libby, Inflammation during the life cycle of the atherosclerotic plaque, Cardiovasc. Res. 117 (2021) 2525–2536, https://doi.org/10.1093/cvr/cvab303.

[29] K.L. Spiller, et al., The role of macrophage phenotype in vascularization of tissue engineering scaffolds, Biomaterials 35 (2014) 4477–4488, https://doi.org/10.1016/j.biomaterials.2014.02.012.
[30] K.S. Moulton, et al., Angiogenesis inhibitors endostatin or TNP-470 reduce intimal neovascularization and plaque growth in apolipoprotein E-deficient mice, Circulation 99 (1999) 1726–1732, https://doi.org/10.1161/01.cir.99.13.1726.
[31] X. Lv, et al., Differential gene expression and immune cell infiltration in carotid intraplaque hemorrhage identified using integrated bioinformatics analysis, Frontiers in cardiovascular medicine 9 (2022) 818585, https://doi.org/10.3389/fcvm.2022.818585.
[32] D.M. Fernandez, et al., Single-cell immune landscape of human atherosclerotic plaques, Nature medicine 25 (2019) 1576–1588, https://doi.org/10.1038/s41591-019-0590-4.
[33] S.W. Tas, C.X. Maracle, E. Balogh, Z. Szekanecz, Targeting of proangiogenic signalling pathways in chronic inflammation, Nat. Rev. Rheumatol. 12 (2016) 111–122, https://doi.org/10.1038/nrrheum.2015.164.
[34] Y. Zhang, N.M. Liu, Y. Wang, J.Y. Youn, H. Cai, Endothelial cell calpain as a critical modulator of angiogenesis, Biochim. Biophys. Acta (BBA) - Mol. Basis Dis. 1863 (2017) 1326–1335, https://doi.org/10.1016/j.bbadis.2017.03.021.
[35] Z. Zhang, et al., Role of lysosomes in physiological activities, diseases, and therapy, J. Hematol. Oncol. 14 (2021) 79, https://doi.org/10.1186/s13045-021-01087-1.
[36] A. Fantin, et al., Neuropilin 1 (NRP1) hypomorphism combined with defective VEGF-A binding reveals novel roles for NRP1 in developmental and pathological angiogenesis, Development (Cambridge, England) 141 (2014) 556–562, https://doi.org/10.1242/dev.103028.
[37] P.C. Hsu, et al., Tumor necrosis factor receptor superfamily member 21 induces endothelial-mesenchymal transition in coronary artery endothelium of type 2 diabetes mellitus, Biomedicines 10 (2022), https://doi.org/10.3390/biomedicines10061282.
[38] I.J. Goldberg, et al., Lipolytic enzymes and free fatty acids at the endothelial interface, Atherosclerosis 329 (2021) 1–8, https://doi.org/10.1016/j.atherosclerosis.2021.05.018.
[39] Q.Q. Nie, et al., SPP1/AnxA1/TIMP1 as essential genes regulate the inflammatory response in the acute phase of cerebral ischemia-reperfusion in rats, J. Inflamm. Res. 15 (2022) 4873–4890, https://doi.org/10.2147/jir.S369690.
[40] X.L. Chen, et al., VEGF-induced vascular permeability is mediated by FAK, Dev. Cell 22 (2012) 146–157, https://doi.org/10.1016/j.devcel.2011.11.002.
[41] M.Y. Abshire, K.S. Thomas, K.A. Owen, A.H. Bouton, Macrophage motility requires distinct α5β1/FAK and α4β1/paxillin signaling events, J. Leukoc. Biol. 89 (2011) 251–257, https://doi.org/10.1189/jlb.0710395.
[42] K. Golfmann, et al., Synergistic anti-angiogenic treatment effects by dual FGFR1 and VEGFR1 inhibition in FGFR1-amplified breast cancer, Oncogene 37 (2018) 5682–5693, https://doi.org/10.1038/s41388-018-0380-3.
[43] H. Yan, L. Zhu, J. Zhang, Z. Lin, Histone demethylase KDM4D inhibition suppresses renal cancer progression and angiogenesis through JAG1 signaling, Cell death discovery 7 (2021) 284, https://doi.org/10.1038/s41420-021-00682-y.
[44] Z. Wei, H. Ran, C. Yang, CircRSF1 contributes to endothelial cell growth, migration and tube formation under ox-LDL stress through regulating miR-758/CCND2 axis, Life Sci. 259 (2020) 118241, https://doi.org/10.1016/j.lfs.2020.118241.
[45] L.V. d'Uscio, Z.S. Katusic, Vascular phenotype of amyloid precursor protein-deficient mice, Am. J. Physiol. Heart Circ. Physiol. 316 (2019) H1297–h1308, https://doi.org/10.1152/ajpheart.00539.2018.
[46] S. Nandadasa, et al., The versican-hyaluronan complex provides an essential extracellular matrix niche for Flk1(+) hematoendothelial progenitors, Matrix Biol. : journal of the International Society for Matrix Biology 97 (2021) 40–57, https://doi.org/10.1016/j.matbio.2021.01.002.