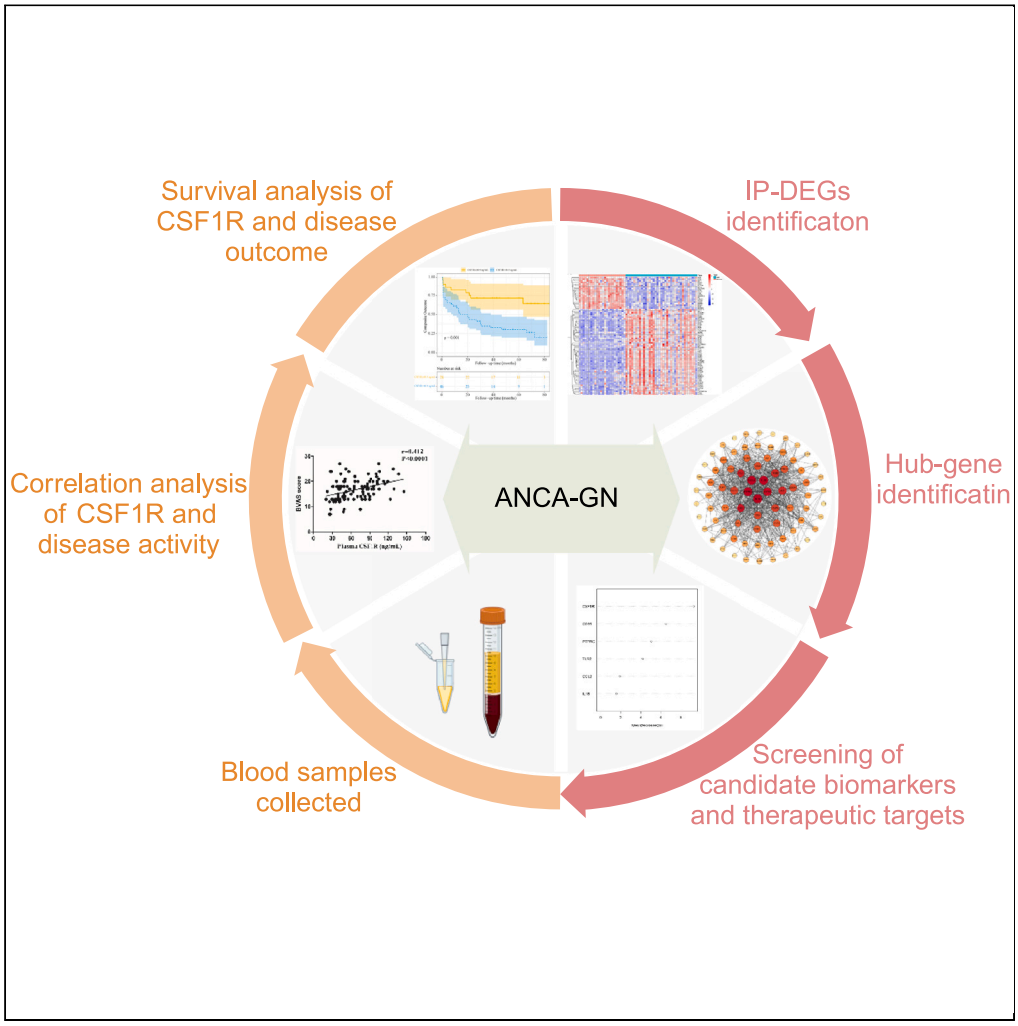
iScience

CellPress
OPEN ACCESS

Article

Identification of potential biomarkers and therapeutic targets for antineutrophil cytoplasmic antibody-associated glomerulonephritis



Yiru Wang, Chenlin Cao, Siyang Liu, Liu Hu, Yueliang Du, Yongman Lv, Qingquan Liu

qqliutj@163.com

Highlights

PTPRC, *CD86*, *TLR2*, *IL1B*, *CSF-1R*, and *CCL2* were hub genes for ANCA-GN

The *CSF-1*/*CSF-1R* axis may be more valuable in the pathological progression of ANCA-GN

CSF-1R might be a potential biomarker for ANCA-GN disease activity and prognosis

Wang et al., iScience *26*, 108157
November 17, 2023 © 2023 The Author(s).
https://doi.org/10.1016/j.isci.2023.108157

iScience

CellPress
OPEN ACCESS

Article

Identification of potential biomarkers and therapeutic targets for antineutrophil cytoplasmic antibody-associated glomerulonephritis

Yiru Wang,[1,5] Chenlin Cao,[1,2,5] Siyang Liu,[1] Liu Hu,[3] Yueliang Du,[4] Yongman Lv,[1] and Qingquan Liu[1,6,*]

SUMMARY

Exploring key genes for antineutrophil cytoplasmic antibody (ANCA)-associated glomerulonephritis (ANCA-GN) is of great significance. Through bioinformatics analysis, 79 immune protein-differentially expressed genes (IP-DEGs) were obtained. Six hub genes (*PTPRC*, *CD86*, *TLR2*, *IL1B*, *CSF-1R*, and *CCL2*) were identified and verified to be increased in ANCA-GN patients. Random forest algorithm and ROC analysis showed that *CSF-1R* was a potential biomarker. Plasma *CSF-1R* levels increased significantly in ANCA-GN-active patients compared with remission stage and control. Correlation analysis revealed that *CSF-1R* levels had positive relationship with serum creatinine and Birmingham scoring, while inversely correlated with eGFR. Multivariate analysis revealed that plasma *CSF-1R* were an independent poor prognostic variable for end-stage renal disease or death, after adjusting for age and gender (HR = 3.05, 95% CI = 1.45–6.43, p = 0.003). Overall, we revealed that the *CSF-1R* is related to disease activity and might be a vital gene associated with the pathogenesis of ANCA-GN.

INTRODUCTION

Antineutrophil cytoplasmic antibody (ANCA)-associated vasculitis (AAV) is a group of autoimmune disorder characterized by inflammation, systemic small blood vessel damage, and the presence of circulating ANCA.[1,2] More than 75% of AAV patients have kidney involvement, which is mainly manifested as rapidly progressive glomerulonephritis with a poor prognosis, also known as ANCA-associated glomerulonephritis (ANCA-GN).[3] Patients usually miss the best treatment opportunity because they are not diagnosed in time. Due to the fact that ANCA-GN has few pathognomonic features, the diagnosis is frequently delayed.[4] Without treatment, only 10% of patients survive for more than two years, indicating that its prognosis is worse than most cancers.[4,5] Immunosuppressive therapy could be effective if provided early, and the prognosis is similar to that of other chronic inflammatory diseases.[4] Thus, in-depth understanding of the immune and molecular mechanisms related to the progression of ANCA-GN renal damage is essential to the early diagnosis and treatment.

Growing clinical and experimental evidence indicates that the innate immune system or immune-related pathways are related to the pathogenesis of ANCA-GN.[6] However, due to the fact that the etiology is multifactorial, the pathophysiology mechanisms of ANCA-GN remain unknown. Most previous studies focused on ANCA-neutrophil interactions that resulted in degranulation of neutrophils, causing necroinflammation and destruction of the glomerular endothelium.[7] Recent studies indicated that monocytes and macrophages had a pivotal role in the pathogenesis of ANCA-GN.[6] Interactions between monocytes and endothelial cells are essential for adhesion and transmigration of monocytes into extravascular compartments and enhanced endothelial cell injury. Similar to monocytes, a host of markers have been identified that are related to macrophage polarization in AAV. $CD163^+/CD68^+$, one of the alternatively activated (M2) macrophage markers, is more prominent in early renal lesions in patients with AAV. The counts of $CD163^+$ and $CD206^+$ cells in renal tubulointerstitium were inversely related to glomerular filtration rate (GFR), and univariate analysis showed that the $CD163^+$ macrophages are a risk factor for developing end-stage renal disease (ESRD).[8] Colony-stimulating factor-1 (CSF-1) is an important molecule in the differentiation and survival of macrophages. CSF-1 binds to *CSF-1R* on the membrane of macrophage to regulate macrophage function. The inhibitor *CSF-1R* can reduce the infiltration of $Ly6C^+$ M2-like macrophages, thereby reducing ischemia-induced renal injury and fibrosis.[9] Therefore, it is promising to target specific biomarkers of macrophages to develop effective and novel therapeutic agents for patients with AAV.

In recent years, the combination of microarray technology and bioinformatics analysis has made outstanding contributions to the exploration of pathological mechanisms and the search for potential biomarkers.[10] By integrating and analyzing various omics data, we can find

[1]Department of Nephrology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[2]Department of the Second Clinical College, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[3]Health Management Center, Tongji Hospital, Tongji Medical college, Huazhong University of Science and Technology, Wuhan, Hubei, China
[4]Department of Nephrology, Luohe Central Hospital, Luohe, China
[5]These authors contributed equally
[6]Lead contact
*Correspondence: qqliutj@163.com
https://doi.org/10.1016/j.isci.2023.108157

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

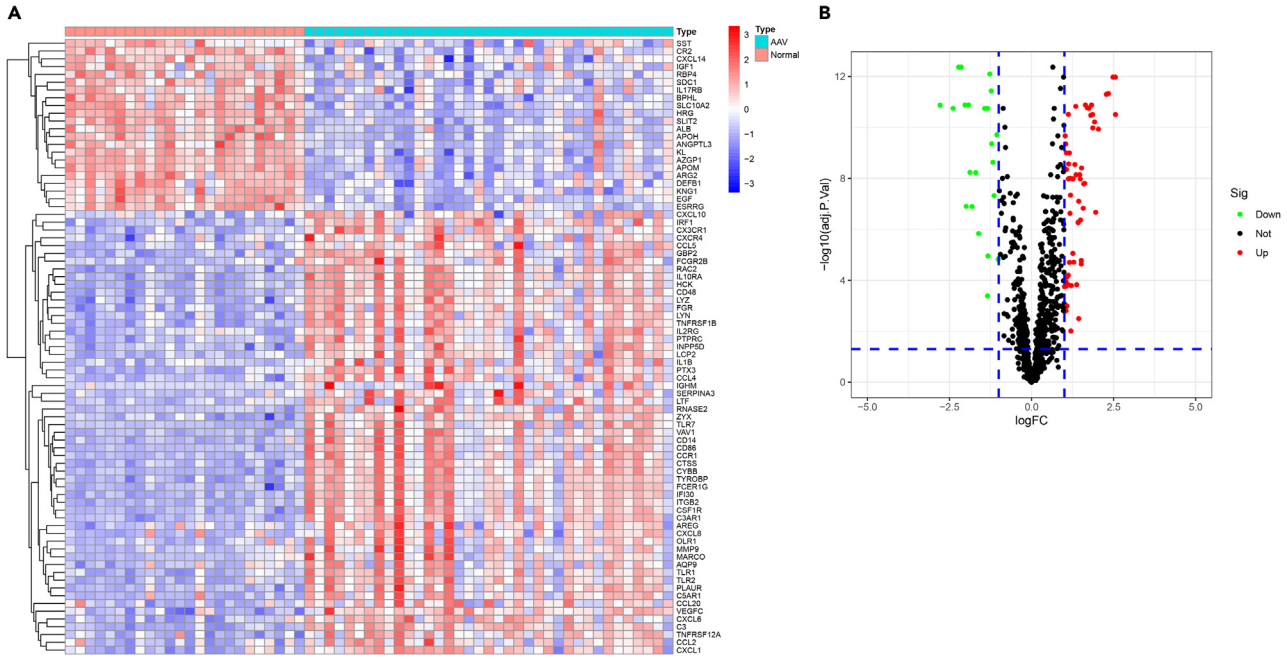


Figure 1. Differential expression analysis of immune-related genes in ANCA-GN patients (GSE108113 and GSE104948 datasets, micro-dissected glomerular tissue)

(A) Heatmap of immune-related genes. Legend on the top right indicates the log fold change of the genes. The vertical axis represents each differentially expressed gene. Blue and red colors represent low and high expression values, respectively.

(B) Volcano map of immune-related genes. Red dots indicate upregulated genes and green dots indicate downregulated genes. AAV, antineutrophil cytoplasmic antibody-associated vasculitis. ANCA-GN, antineutrophil cytoplasmic antibody-associated glomerulonephritis.

disease biomarkers at a new level, which may be conducive to accurate disease diagnosis, patient classification, and the implementation of precision medicine.[11] In this context, we downloaded the human kidney tissue microarray gene expression profile to perform bioinformatics analysis.[12] Because autoimmunity imbalance is the most important pathological process in ANCA-GN, we focused on immune-related genes for further investigation. Then the microarray data were used to screen out differentially expressed genes (DEGs). Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis, and gene set enrichment analysis (GSEA) were performed to in-depth comprehend the function of DEGs. Protein-protein interaction (PPI) analysis was used to screen the core genes of DEGs. Immune infiltration analysis, a model-based recursive partitioning algorithm, random forest (RF), and clinical online database were employed to further identify the characteristics of six hub genes. Subsequently, plasma *CSF-1R* was identified in ANCA-GN patients and controls to validate the expression of major disease biomarkers, and the associations between plasma *CSF-1R* levels, clinical characteristics, and prognosis were further analyzed.

RESULTS

Identification of IP-DEGs in ANCA-GN

We attempted to identify immune protein-DEGs (IP-DEGs) for analysis. First of all, we analyze the DEGs of GEO: GSE108113 and GEO: GSE104948 datasets. After consolidation and normalization of the microarray data, we found a total of 79 IP-DEGs involved in ANCA-GN. Heatmap showed the 56 upregulated genes and 23 downregulated genes (Figure 1A). The volcano map of IP-DEGs was shown in Figure 1B.

Functional enrichment analysis of IP-DEGs

GO and KEGG enrichment analyses were carried out by the "ClusterProfiler" package in R software. Both IP-DEGs in ANCA-GN and the whole immune-related genes were included in enrichment analyses to ensure the sensitivity. The top 10 GO items in ANCA-GN were shown in Figure 2A and top 5 GO items in the whole immune genes were shown in Figure S1A. In the top five biological processes terms, neutrophil activation, neutrophil-mediated immunity, and myeloid leukocyte migration were specific in ANCA-GN. Regarding the cellular component, the specific terms in ANCA-GN were secretory granule lumen, cytoplasmic vesicle lumen, and secretory granule membrane. Regarding their molecular function, the specific terms in ANCA-GN were G protein-coupled receptor binding, signaling receptor activator activity, and immune receptor activity. The top 10 KEGG terms in ANCA-GN were shown in Figure 2B and top 10 KEGG terms in the whole immune genes were shown in Figure S1B. The specific KEGG terms in ANCA-GN were lipid and

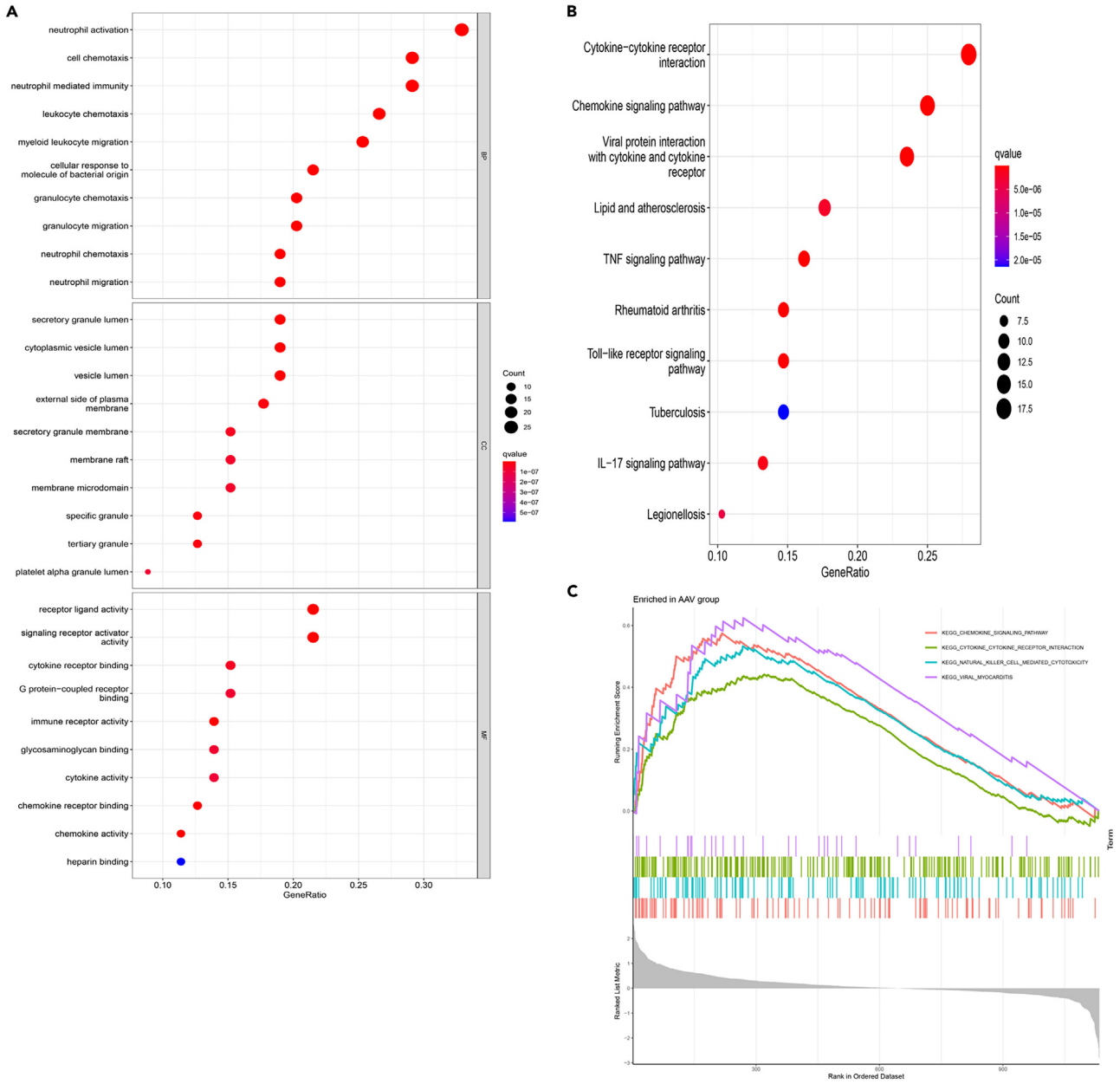


Figure 2. Functional enrichment analysis of immune-related DEGs (GSE108113 and GSE104948 datasets, micro-dissected glomerular tissue)

(A) GO enrichment analysis of immune-related DEGs. The x axis label represents the gene ratio, and the y axis label represents GO terms. The size of circle represents gene count. Different color of circles represents different adjusted q value.

(B) KEGG enrichment analysis of immune-related DEGs. The x axis label represents the gene ratio, and the y axis label represents KEGG terms. The size of circle represents gene count. Different color of circles represents different adjusted q value.

(C) Enriched KEGG terms in ANCA-GN patients via gene set enrichment analysis (GSEA). GO, gene ontology. KEGG, Kyoto Encyclopedia of Genes and Genomes. ANCA-GN, antineutrophil cytoplasmic antibody-associated glomerulonephritis. DEGs, differentially expressed genes.

atherosclerosis, TNF signaling pathway, rheumatoid arthritis, Toll-like receptor (TLR) signaling pathway, interleukin 17 (IL-17) signaling pathway, tuberculosis, and legionellosis. We further performed a GSEA analysis to evaluate the IP-DEGs based on the two categories of normal group and AAV group. Enrichment score>0.4 and $p < 0.05$ were filtering criteria for pathways. The chemokine signaling pathway (enrichment score = 0.58, $p < 0.001$), cytokine-cytokine receptor interaction (enrichment score = 0.44, $p = 0.001$), natural killer (NK) cell-mediated cytotoxicity (enrichment score = 0.53, $p < 0.001$), and viral myocarditis (enrichment score = 0.62, $p = 0.002$) were significantly enriched in ANCA-GN (Figure 2C).

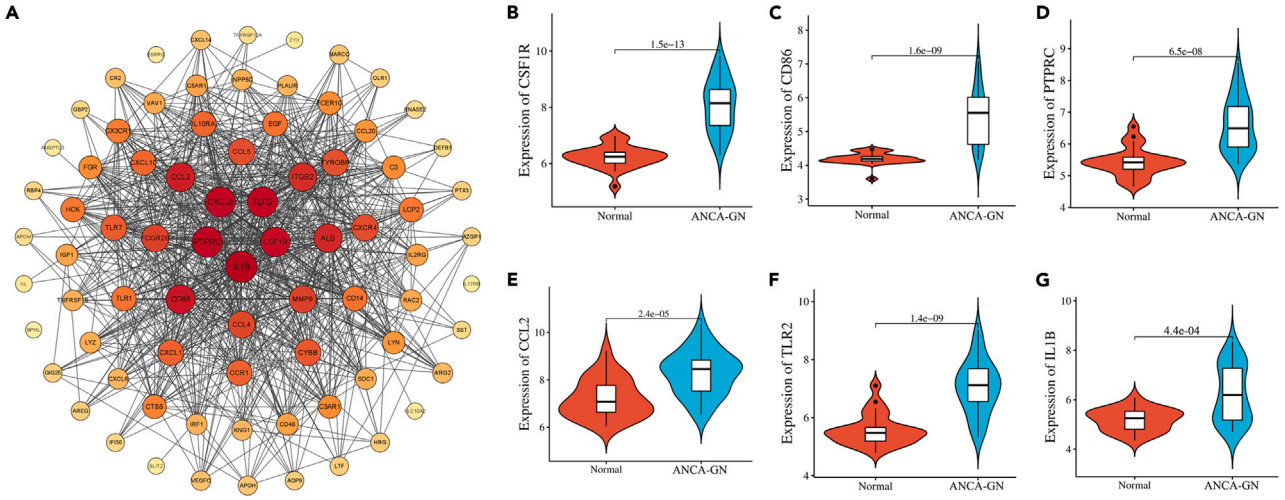


Figure 3. Construction of protein-protein interaction (PPI) network and expression levels of hub genes

(A) PPI network for immune-related DEGs visualized by Cytoscape. Nodes represent target protein, edges represent interactions among genes. The darker the color and the larger the node are, the greater the degree is. Showing the comparison of six hub genes in ANCA-GN tissues and healthy control, CSF-1R (B), CD86 (C), PTPRC (D), CCL2 (E), TLR2 (F), IL1B (G). DEGs, differentially expressed genes. ANCA-GN, antineutrophil cytoplasmic antibody-associated glomerulonephritis.

Establishment of PPI network and identification of hub genes

To screen out the most important genes, we constructed a PPI network to explore the mutual effect between the IP-DEGs. The PPI network consisted of 78 nodes and 765 edges. The nodes corresponded to genes, and the edges indicate co-expression between the two genes connected. *CSF-1R*, located in the center of the network, was one of the top 5 central genes in the PPI network (Figure 3A). To recognize potential hub genes in ANCA-GN, we applied the logical methods of MCC, MNC, Degree, EPC, Closeness, Radiality, Betweenness, and Stress via CytoHubba in Cytoscape to screen the top 10 hub genes (Table 1). We also screen overlapping genes using these logical methods and screened out six hub common genes, containing *PTPRC*, *CD86*, *TLR2*, *IL1B*, *CSF-1R*, and *CCL2*. Relative mRNA expressions of 6 hub genes in GEO: GSE108113 and GEO: GSE104948 were shown in Figures 3B–3G, and they were significantly upregulated in ANCA-GN glomeruli samples. Meanwhile, we compared the expressions of 6 hub genes between ANCA-associated glomerulonephritis and other glomerulonephritis. The data of membranous nephropathy (n = 44) and focal and segmental glomerulosclerosis (n = 30) were obtained from GEO: GSE108113 dataset. The data of IgA nephropathy (n = 27) and systemic lupus erythematosus (n = 32) were obtained from GEO: GSE104948 dataset. *PTPRC*, *CD86*, *TLR2*, *IL1B*, and *CSF-1R* were significantly upregulated in ANCA-associated glomerulonephritis compared with membranous nephropathy ($p < 0.001$, Figure S2A). Similarly, *PTPRC*, *CD86*, *TLR2*, *IL1B*, and *CSF-1R* were significantly upregulated in ANCA-associated glomerulonephritis compared with focal and segmental glomerulosclerosis ($p < 0.001$, Figure S2B). *TLR2* and *IL1B* were significantly upregulated in IgA nephropathy ($p < 0.001$), while *CCL2* was significantly upregulated in ANCA-associated glomerulonephritis ($p = 0.017$, Figure S2C). *PTPRC* ($p = 0.0019$), *CD86* ($p = 0.0031$), *TLR2* ($p < 0.001$), *IL1B* ($p < 0.001$), and *CSF-1R* ($p < 0.001$) were significantly upregulated in systemic lupus erythematosus compared with ANCA-associated glomerulonephritis (Figure S2D). Most hub genes had a high level of expression in glomerulonephritis caused by autoimmune diseases. So, we speculated that these hub genes may be closely related to the abnormal immune process.

External dataset and clinical validation

For proofing hub genes, we used ArrayExpress: E-MTAB-1944 dataset to evaluate the expression of the aforementioned six genes. The expression of these hub genes was displayed in a heatmap (Figure 4A). We next used the Nephroseq database to further verify the expression levels of hub genes and explore the relationship between their expression and the clinical characteristics of AAV. As shown in Figure 4B, their expression trends were consistent with the datasets of GEO: GSE108113, GEO: GSE104948, and ArrayExpress: E-MTAB-1944. The expression of *CSF-1R* was inversely correlated with the GFR in ANCA-GN glomeruli ($r = -0.573$, $p = 0.0074$) (Figure 4C), suggesting that the upregulation of *CSF-1R* was related to the damage of renal function and may exacerbate the development of ANCA-GN in renal tissues. Correlations between the expression of *CCL2*, *TLR2*, *CD86*, *IL1B*, and *PTPRC* in ANCA-GN glomeruli and the GFR were shown in Figures 4D–4H ($r = -0.558$, $p = 0.0057$; $r = -0.586$, $p = 0.0033$, $r = -0.494$, $p = 0.0165$; $r = -0.151$, $p = 0.492$; $r = -0.374$, $p = 0.0786$, respectively).

Immune cell infiltration in the ANCA-GN and its relationship with hub genes

The "CIBERSORT" algorithm was used to further investigate the relationship of immune infiltration between ANCA-GN group and healthy control group in GEO: GSE108113 and GEO: GSE104948. The results showed that the immune cells infiltration of B cell naive,

Table 1. Top 10 hub genes by topological analysis methods of CytoHubba

MCC	MNC	Degree	EPC	Closeness	Radiality	Betweenness	Stress
PTPRC	IL1B	IL1B	PTPRC	IL1B	IL1B	ALB	ALB
CD86	PTPRC	PTPRC	IL1B	PTPRC	PTPRC	IL1B	IL1B
TLR2	TLR2	TLR2	TLR2	TLR2	TLR2	TLR2	TLR2
CCL4	CXCL8	CXCL8	CXCL8	CXCL8	CXCL8	CXCL8	CXCL8
IL1B	CSF1R	CSF1R	CSF1R	CSF1R	CSF1R	CXCR4	PTPRC
CSF1R	CD86	CD86	CD86	CD86	CD86	PTPRC	CCL2
CCL2	CCL2	CCL2	CCL2	CCL2	CCL2	CCL2	CXCR4
TLR1	ALB	ALB	FCGR2B	ALB	ALB	CD86	CSF1R
FCGR2B	ITGB2	ITGB2	CCL4	ITGB2	FCGR2B	CSF1R	CD86
TLR7	FCGR2B	FCGR2B	CYBB	FCGR2B	ITGB2	HRG	ITGB2

memory resting CD4 T cells, T cells regulatory, NK cells resting, and neutrophils were lower in ANCA-GN than in the control group, while the immune cells infiltration of T cells gamma delta, NK cells activated, monocytes, and mast cells activated were higher in ANCA-GN than in healthy control group (Figures 5A and 5B), indicating that these immune cells may be involved in molecular mechanisms of ANCA-GN progression. We further used the "corrplot" package to calculate the correlation between immune cells and the hub genes and found negative correlations of B cell naive with *TLR2*, *PTPRC*, *CSF-1R*, and *CD86*. *CSF-1R* was positively correlated with resting dendritic cells, M0 macrophages, M2 macrophages, and gamma delta T cells, while negatively correlated with resting NK cells and memory resting CD4 T cells and CD8 T cells. Monocytes were positively correlated with *IL1B*. Mast cells activated had a positive relationship with *PTPRC*, *IL1B*, *CD86*, and *CCL2* (Figure 5C). These results further demonstrated that the hub genes may modulate immune processes during the progression of ANCA-GN.

Identification of candidate diagnostic biomarkers and gene set enrichment analysis

We used RF algorithms to identify candidate diagnostic biomarkers. The RF algorithm calculated the error rate of each tree of 1–2000 trees, and the trees with the lowest error rate (ntree = 60) were used for subsequent analysis (Figure 6A). The gene importance was calculated according to the precision decreasing method, and we found the score of *CSF-1R* has the highest rating (Figure 6B), further verifying the diagnostic value of the aforementioned six key genes in the GEO: GSE108113 and GEO: GSE104948 datasets to verify their clinical efficacy. The area under the curve (AUC) values of *IL1B*, *TLR2*, *PTPRC*, *CCL2*, *CSF-1R*, and *CD86* were 0.769, 0.918, 0.913, 0.801, 0.975, and 0.961, respectively. The *CSF-1R* gene had the highest AUC value of 0.975 (Figure 6C). Next, we used the GSEA to classify ANCA-GN glomeruli tissue into two categories based on the median *CSF-1R* levels. Chemokine signaling pathway (enrichment score = 0.54, $p < 0.001$), cytokine-cytokine receptor interaction (enrichment score = 0.44, $p = 0.007$), Fc-γ R-mediated phagocytosis (enrichment score = 0.65, $p = 0.001$), Leishmania infection (enrichment score = 0.58, $p = 0.005$), leukocyte transendothelial migration (enrichment score = 0.61, $p = 0.005$), NOD-like receptor signaling pathway (enrichment score = 0.70, $p < 0.001$), TLR signaling pathway (enrichment score = 0.55, $p = 0.005$), and viral myocarditis (enrichment score = 0.61, $p = 0.013$) were significantly enriched in the high *CSF-1R* subgroup (Figure 6D). Moreover, we used GSVA to classify ANCA-GN patients' glomeruli into two categories based on the median expression of CSF-1R. Glycosphingolipid biosynthesis lacto and neolacto series and the material metabolic pathway were significantly enriched in the high CSF-1R subgroup, whereas complement and coagulation cascades, TLR signaling pathways, NOD-like receptor signaling pathways, and primary immunodeficiency were significantly enriched in the low CSF-1R subgroup (Figure 6E). We also constructed a gene-gene interaction network in the GeneMANIA database with *CSF-1R* as the center gene. Genes related to *CSF-1R* are enriched in functions of mononuclear cell differentiation, myeloid leukocyte differentiation, regulation of monocyte differentiation, antigen receptor-mediated signaling pathway, positive regulation of macrophage migration, and regulation of macrophage chemotaxis (Figure 6F). These genes play important roles in the pathology of the disease and *CSF-1R* is the central gene of these genes. So, *CSF-1R* was selected for further clinical research.

CSF-1R characteristic of ANCA-GN patients

To explore the role of *CSF-1R* in ANCA-GN, we included 45 controls and 114 ANCA-GN patients, and then detected plasma *CSF-1R* levels. The basic features of patients are shown in Table 2. And the clinical parameters of healthy control group are shown in Table S2. Among patients, 48 were male and 66 were female, with an age of 52.80 ± 19.37 (range 15–82) years at diagnosis. Of the 114 patients, 87 had active ANCA-GN and 27 had partial or complete remission. Anti-myeloperoxidase (MPO) antibody was positive in 106 cases (93.0%) and anti-PR3 antibody was positive in 10 cases (8.8%). The initial serum creatinine concentration was 2.57 (1.32, 5.62 mg/dL) and the level of Birmingham Vasculitis Activity Score (BVAS) was 17 (13, 20). Plasma *CSF-1R* levels were significantly higher in patients with ANCA-GN in the active stage compared with patients in remission or partial remission stage and normal controls (71.13 ± 30.13 vs. 50.49 ± 25.90 ng/mL, $p = 0.0017$; 71.13 ± 30.13 vs. 40.85 ± 15.87 ng/mL, $p < 0.0001$, respectively). There was no significant difference in the plasma *CSF-1R* levels

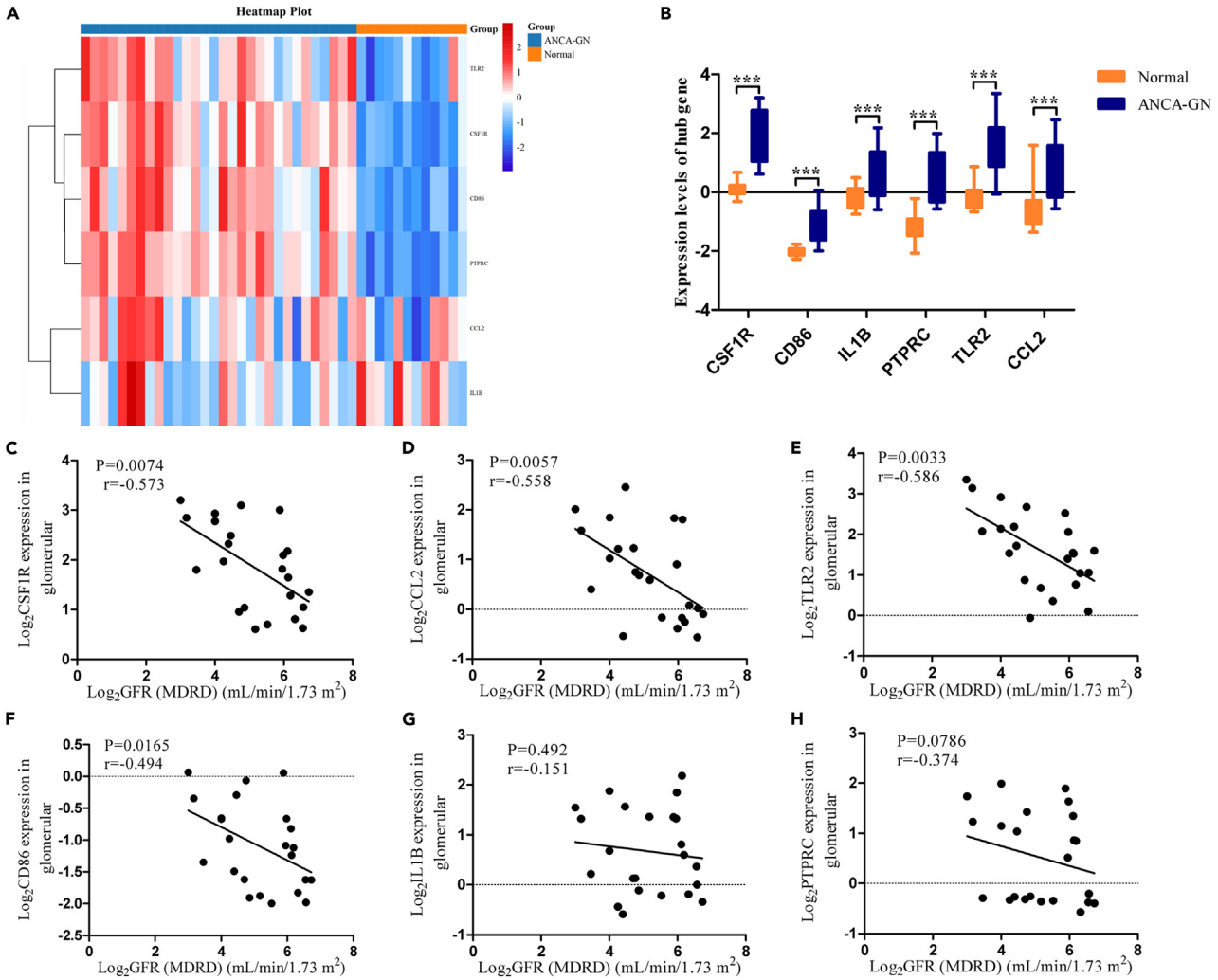


Figure 4. External dataset validation (E-MTAB-1944 dataset, micro-dissected kidney tissue) and correlation between the expression of hub genes in ANCA-GN glomeruli and GFR

(A) The clustering of six hub genes in heatmap of dataset E-MTAB-1944.

(B) Expression levels of CSF-1R, CD86, IL1B, PTRPC, TLR2, and CCL2 were significantly upregulated in glomeruli of ANCA-GN patient in Nephroseq database (normal n = 21, ANCA-GN patients n = 23). Correlation analysis of GFR and the expression of CSF-1R (C), CCL2 (D), TLR2 (E), CD86 (F), IL1B (G), and PTRPC (H) in ANCA-GN glomeruli from Nephroseq database (n = 23). ANCA-GN, antineutrophil cytoplasmic antibody-associated glomerulonephritis. GFR, glomerular filtration rate; MDRD, modification of diet in renal disease. ***p<0.001.

between ANCA-GN patients in normal controls and in remission or partial remission stage (40.85 ± 15.87 vs. 50.49 ± 25.90 ng/mL, p = 0.085) (Figure 7A). We further investigated the plasma *CSF-1R* levels in 15 of 114 patients ANCA-GN patients during different periods of time, including active stage and the remission or partial remission stage. Because the differences between paired data followed a normal distribution by Shapiro-Wilk test (p = 0.19), paired t-test was used for analysis. The plasma *CSF-1R* levels were significantly higher in the active stage than those in remission or partial remission stage (59.23 ± 31.64 vs. 48.66 ± 32.58 ng/mL, p = 0.0025) (Figure 7B). Moreover, immunohistochemical staining of biopsy specimens from 6 ANCA-GN patients and 6 normal kidney tissues revealed that *CSF-1R* levels were considerably increased in ANCA-GN patients compared with normal renal glomerular tissues (Figure 7C). Then we used ELISA kit to detect the plasma levels of CSF-1 and IL-34 which are the CSF-1R ligands to determine the main binding ligands of CSF-1R. The results showed that the levels of plasma CSF-1 and IL-34 in patients with ANCA-GN were higher than those in healthy controls (Figures S3A and S3B). Additionally, there was a positive correlation between CSF-1R and CSF-1 levels (r = 0.37, p = 0.019), but no correlation was observed between CSF-1R and IL-34 levels (Figures S3C and S3D). For tissue microarrays data, we conducted correlation analysis of CSF1, IL-34, and CSF-1R in GSE108113 and E-MTAB-1944 datasets (GSE104948 lacks the expression data of IL-34 so was excluded) (Figures S3E and S3F). This may indicate that plasma CSF-1R mainly plays a role by combining with CSF-1 in ANCA patients, which needs to be further verified.

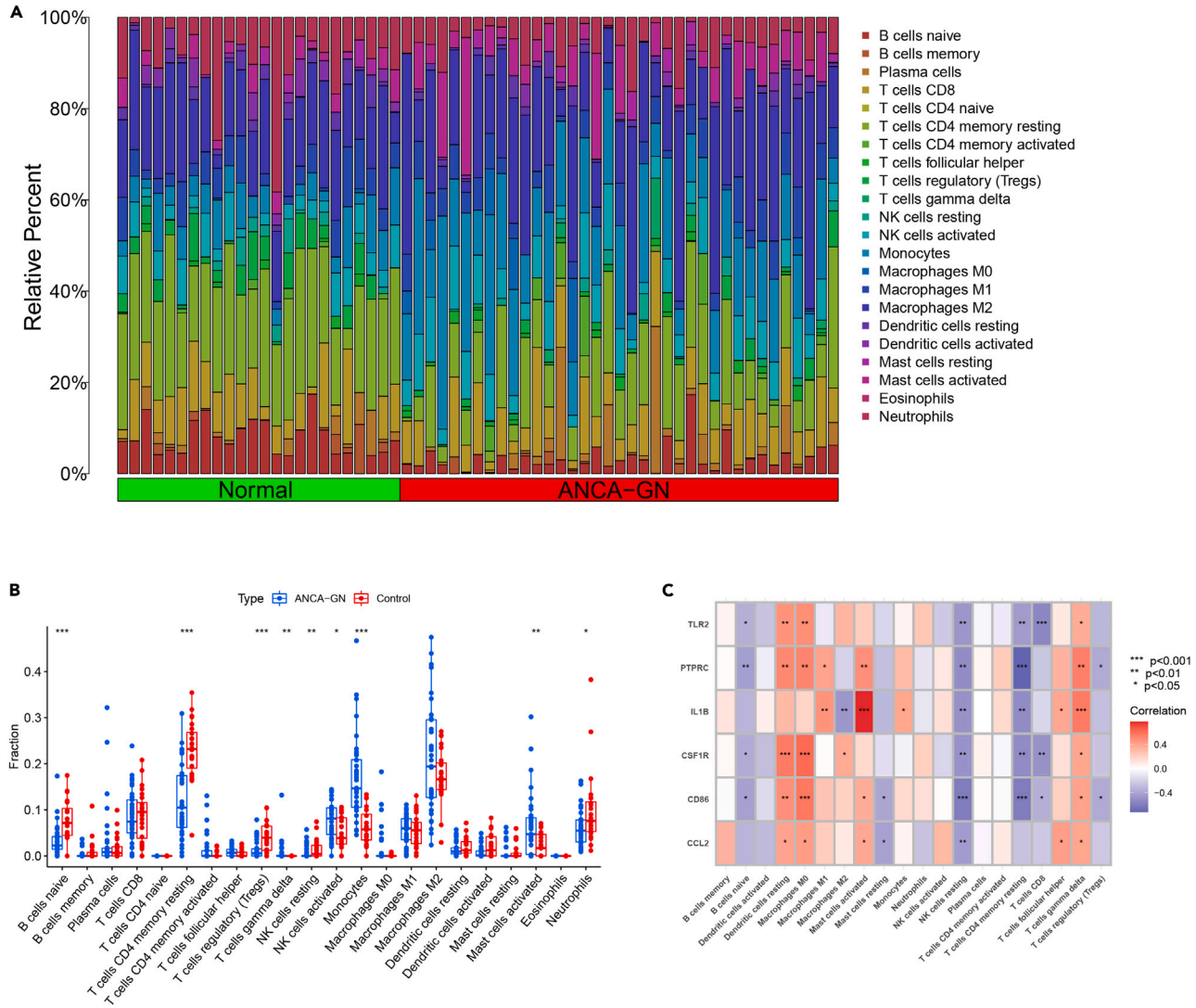


Figure 5. Analysis of immune landscape associated with ANCA-GN in GSE108113 and GSE104948

Heatmap (A) and boxplot (B) showing the distribution of 22 types of immune cells in ANCA-GN group and healthy control tissues.

(C) The relationship between six hub genes and immune cell infiltration. ANCA-GN, antineutrophil cytoplasmic antibody-associated glomerulonephritis. *p < 0.05, **p < 0.01, ***p < 0.001.

Relationship between plasma *CSF-1R* levels and clinical features in patients with active AAV

Correlation analysis of the *CSF-1R* level and clinical indicators of patients with active ANCA-GN revealed that the plasma level of *CSF-1R* was positively correlated with initial serum creatinine and BVAS ($r = 0.28$, $p = 0.002$; $r = 0.412$, $p < 0.0001$, respectively) and inversely correlated with GFR level ($r = -0.27$, $p = 0.0033$) (Figures 8A–8C). We found *CSF-1R* interaction with *C1QA*, *C1QB*, and *C1QC* by using the CytoHubba (Figure S4A). It has been confirmed in the research that there is a close relationship between *CSF-1R* and C1Q,[13] so the C1Q levels in the plasma of ANCA-GN patients were also assessed by ELISA. Plasma C1Q levels were increased in ANCA-GN patients (Figure S4B). The *C1QA*, *C1QB*, and *C1QC* expression levels were significantly increased in patients with ANCA-GN compared with healthy control by using the Nephroseq database, GEO: GSE108113 and GEO: GSE104948 dataset (Figures S4C–S4E). We further found the *CSF-1R* levels was positively related to C1Q in the plasma of active ANCA-GN patients ($r = 0.49$, $p < 0.0001$) (Figure 8D).

Association between plasma *CSF-1R* levels and outcomes of ANCA-GN patients

Patients with AAV included in the study were followed up with a median of 30 months and 16 patients died; 30 patients developed ESRD, while 44 patients reached the composite outcome during the follow-up period (Table S3). To verify the prediction efficiency of plasma *CSF-1R* for composite outcomes and determine the best cutoff value, receiver operating characteristic curve (ROC) analysis was used (AUC = 0.681, 95%

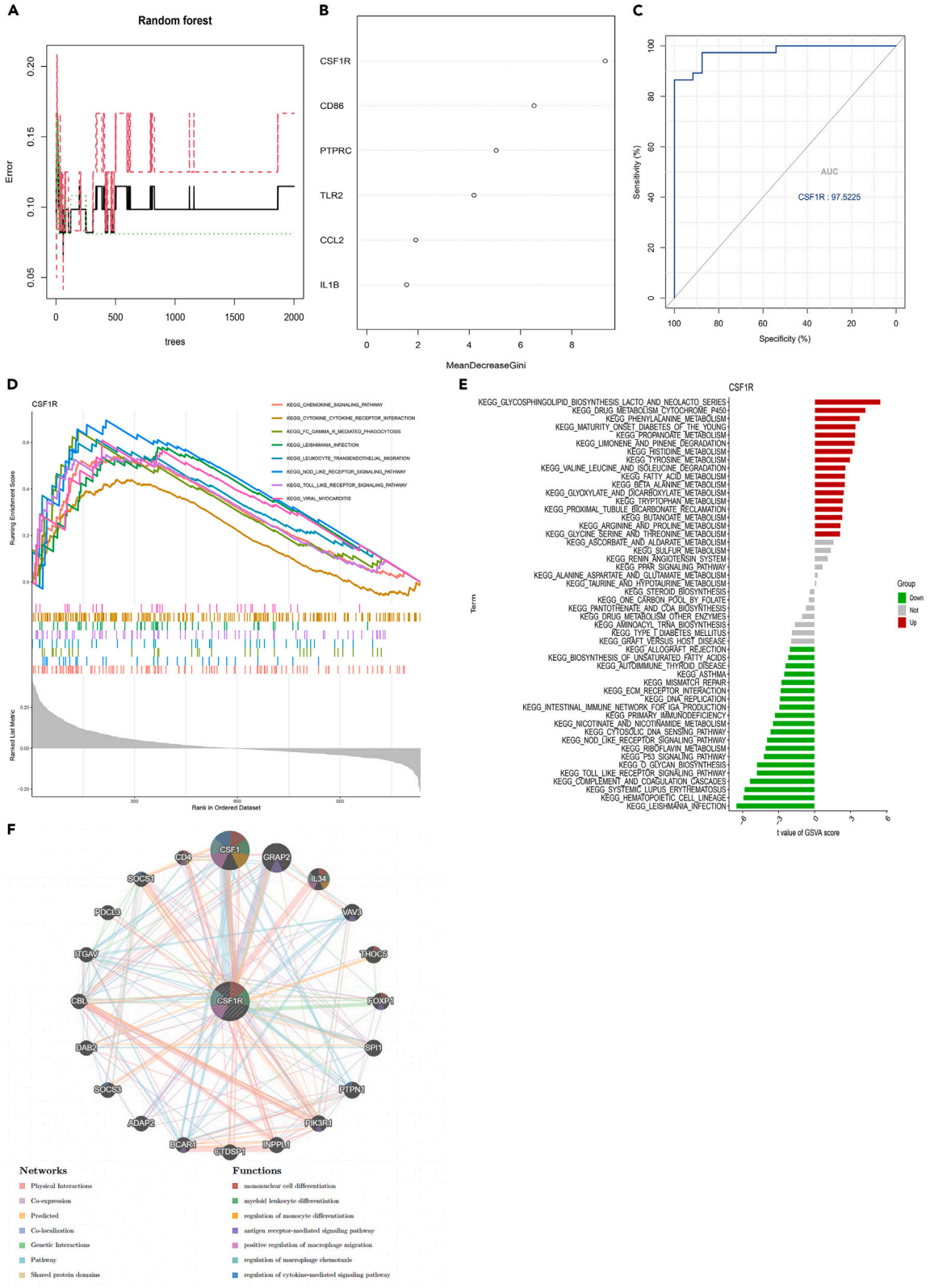

A
Random forest
Error
trees
B
CSF1R
CD86
PTPRC
TLR2
CCL2
IL1B
MeanDecreaseGini
C
Sensitivity (%)
Specificity (%)
AUC
CSF1R : 97.5225
D
CSF1R
Running Enrichment Score
Ranked List Metric
Rank in Ordered Dataset
KEGG_CHEMOKINE_SIGNALING_PATHWAY
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION
KEGG_FC_GAMMA_R_MEDIATED_PHAGOCYTOSIS
KEGG_LEISHMANIA_INFECTION
KEGG_LEUKOCYTE_TRANSENDOTHELIAL_MIGRATION
KEGG_NOD_LIKE_RECEPTOR_SIGNALING_PATHWAY
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY
KEGG_VIRAL_MYOCARDITIS
E
CSF1R
Term
t value of GSVA score
Group
Down
Not
Up
KEGG_GLYCOSPHINGOLIPID_BIOSYNTHESIS_LACTO_AND_NEOLACTO_SERIES
KEGG_DRUG_METABOLISM_CYTOCHROME_P450
KEGG_PHENYLALANINE_METABOLISM
KEGG_MATURITY_ONSET_DIABETES_OF_THE_YOUNG
KEGG_PROPANOATE_METABOLISM
KEGG_LIMONENE_AND_PINENE_DEGRADATION
KEGG_HISTIDINE_METABOLISM
KEGG_TYROSINE_METABOLISM
KEGG_VALINE_LEUCINE_AND_ISOLEUCINE_DEGRADATION
KEGG_FATTY_ACID_METABOLISM
KEGG_BETA_ALANINE_METABOLISM
KEGG_GLYOXYLATE_AND_DICARBOXYLATE_METABOLISM
KEGG_TRYPTOPHAN_METABOLISM
KEGG_PROXIMAL_TUBULE_BICARBONATE_RECLAMATION
KEGG_BUTANOATE_METABOLISM
KEGG_ARGININE_AND_PROLINE_METABOLISM
KEGG_GLYCINE_SERINE_AND_THREONINE_METABOLISM
KEGG_ASCORBATE_AND_ALDARATE_METABOLISM
KEGG_SULFUR_METABOLISM
KEGG_RENIN_ANGIOTENSIN_SYSTEM
KEGG_PPAR_SIGNALING_PATHWAY
KEGG_ALANINE_ASPARTATE_AND_GLUTAMATE_METABOLISM
KEGG_TAURINE_AND_HYPOTAURINE_METABOLISM
KEGG_STEROID_BIOSYNTHESIS
KEGG_ONE_CARBON_POOL_BY_FOLATE
KEGG_PANTOTHENATE_AND_COA_BIOSYNTHESIS
KEGG_DRUG_METABOLISM_OTHER_ENZYMES
KEGG_AMINOACYL_TRNA_BIOSYNTHESIS
KEGG_TYPE_I_DIABETES_MELLITUS
KEGG_GRAFT_VERSUS_HOST_DISEASE
KEGG_ALLOGRAFT_REJECTION
KEGG_BIOSYNTHESIS_OF_UNSATURATED_FATTY_ACIDS
KEGG_AUTOIMMUNE_THYROID_DISEASE
KEGG_ASTHMA
KEGG_MISMATCH_REPAIR
KEGG_ECM_RECEPTOR_INTERACTION
KEGG_DNA_REPLICATION
KEGG_INTESTINAL_IMMUNE_NETWORK_FOR_IGA_PRODUCTION
KEGG_PRIMARY_IMMUNODEFICIENCY
KEGG_NICOTINATE_AND_NICOTINAMIDE_METABOLISM
KEGG_CYTOSOLIC_DNA_SENSING_PATHWAY
KEGG_NOD_LIKE_RECEPTOR_SIGNALING_PATHWAY
KEGG_RIBOFLAVIN_METABOLISM
KEGG_P53_SIGNALING_PATHWAY
KEGG_O_GLYCAN_BIOSYNTHESIS
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY
KEGG_COMPLEMENT_AND_COAGULATION_CASCADES
KEGG_SYSTEMIC_LUPUS_ERYTHEMATOSUS
KEGG_HEMATOPOIETIC_CELL_LINEAGE
KEGG_LEISHMANIA_INFECTION
F
CSF1
CD4
GRAP2
SOCS1
IL34
PDCL3
VAV3
ITGAV
THOC5
CBL
CSF1R
FOXP1
DAB2
SPI1
SOCS3
PTPN1
ADAP2
PIK3R1
BCAR3
CTDSP1
INPPL1
Networks
Physical Interactions
Co-expression
Predicted
Co-localization
Genetic Interactions
Pathway
Shared protein domains
Functions
mononuclear cell differentiation
myeloid leukocyte differentiation
regulation of monocyte differentiation
antigen receptor-mediated signaling pathway
positive regulation of macrophage migration
regulation of macrophage chemotaxis
regulation of cytokine-mediated signaling pathway

Figure 6. Analysis of diagnostic biomarkers

(A) The influence of the number of decision trees on the error rate in random forest classifier. The x axis represents the number of decision trees and the y axis is the error rate.

(B) The Gini coefficient method scores for miRNAs in random forest classifier. The x axis represents the importance index, and the y axis represents the candidate biomarkers.

(C) ROC curve of six candidate biomarkers.

(D) Gene set enrichment analysis (GSEA) between group of CSF-1R high-expression and group of CSF-1R low-expression. Groups of high and low expression were divided according to the median expression value of CSF-1R.

(E) Gene Set Variation Analysis (GSVA) of in ANCA-GN patients. Red columns indicate upregulated terms and green columns indicate downregulated terms.

(F) Diagnostic biomarkers co-expression network constructed by the GeneMANIA database. ROC curve, receiver operating characteristic curve. ANCA-GN, antineutrophil cytoplasmic antibody-associated glomerulonephritis.

CI = 0.556–0.807, p = 0.008). According to the Jordan index Youden's index, the best cutoff value for *CSF-1R* was 48.9 ng/mL with a sensitivity of 60.0% and a specificity of 75.6% (Figure 9A). *CSF-1R* was divided into high and low groups according to the best cutoff value, and Kaplan-Meier survival analysis was performed to evaluate the correlation between plasma *CSF-1R* levels and the prognosis of ANCA-GN patients (Figures 9B and 9C). Results showed that the time to the composite outcome or ESRD in patients with *CSF-1R* levels>48.9 ng/mL was apparently shorter than that in patients with lower *CSF-1R* levels, indicating that the composite outcome and ESRD were worse in the high *CSF-1R* group (p = 0.001, p = 0.002, respectively). Furthermore, the association between *CSF-1R* and the composite outcome or ESRD was still significant when excluding patients with less than 3 months of follow-up in the Kaplan-Meier survival analysis (Figures S5A and S5B, p = 0.006, p < 0.001, respectively). We used competitive risk analysis to confirm that the high CSF-1R increases the risk of ESRD, even after controlling for the competing risk of death (p < 0.01) (Figure S5C).

Univariate Cox regression analysis suggested a role for *CSF-1R* level as a time-varying risk factor for a composite endpoint, and the risk of the outcome in the high *CSF-1R* group was 3.14 times higher than that in the low *CSF-1R* group (HR = 3.14, 95% CI = 1.50–6.56, p = 0.002). It was also found that *CSF-1R* was associated with the composite outcome of univariate Cox regression after excluding patients with less than 3 months of follow-up (Table S4, HR = 3.30, 95% CI = 1.33–8.20, p = 0.010). In addition to plasma *CSF-1R* levels, other candidate variables entered into the multivariate analysis are shown in Table 3. Adjusting a multivariate Cox regression analysis for age and gender alone had no effect on the significance of the *CSF-1R* level in predicting the composite outcome (HR = 3.05, 95% CI = 1.45–6.43, p = 0.003). After adjusting for age, gender, serum creatinine, urinary protein concentrations and C-reactive protein,[14] the CSF-1R level failed to be related to an endpoint outcome (HR = 1.28, 95% CI = 0.15–10.80, p = 0.819).

DISCUSSION

In this study, we analyzed IP-DEGs in glomeruli tissue of ANCA-GN patients from two independent microarray datasets, GEO: GSE108113 and GEO: GSE104948. The number of immune-related DEGs between ANCA-GN patients and control glomeruli was 79 in ANCA-GN. GO enrichment analysis showed IP-DEGs were primarily related to neutrophil activation, neutrophil-mediated immunity, and signaling receptor activator activity, while KEGG and GSEA enrichment analysis indicated that IP-DEGs mainly enriched in cytokine-cytokine receptor interaction, TLR signaling pathway, B cell receptor signaling pathway, complement, and coagulation cascades. Using the PPI network and logical methods for calculation, we discovered six hub genes: *PTPRC*, *CD86*, *TLR2*, *IL1B*, *CCL2*, and *CSF-1R* in the glomeruli. Furthermore, upon re-analyzing samples with an estimated GFR (eGFR)>60 mL/min/1.73 m^2, we observed a significant increase in the expression *IL1B*, *CSF-1R*, and *CCL2*, which were validated in the ArrayExpress: E-MTAB-1944 datasets and the Nephroseq database. Following immune infiltration analysis, RF algorithms, and ROC curves analysis, *CSF-1R* was identified as the most important hub gene. Therefore, *CSF-1R* was further applied to clinical study of ANCA-GN patients, while the results revealed that plasma *CSF-1R* levels were obviously higher in active ANCA-GN patients compared with ANCA-GN patients in remission and normal controls. Plasma *CSF-1R* levels are related to disease activity and composite outcome of MPO-ANCA-positive patients to a great extent.

AAV is the general name of a group of multi-system autoimmune vasculitis, which indicates systemic symptoms of chronic inflammatory diseases. The pathophysiology of AAV is complex, and the cellular immune system is involved in the pathological process. Previous studies mainly focused on identified ANCA-neutrophil interactions resulting in neutrophil respiratory burst with release of granules and neutrophil extracellular traps, and then activate the adaptive immune system to induce massive vascular injury.[15] Increasing evidence also verified an important role of monocytes and macrophages in the pathophysiology of AAV.[16] In this study, we used the GO enrichment analysis and KEGG signal pathway proved that IP-DEGs were primarily enriched in neutrophil activation, neutrophil-mediated immunity, and pathways related to immune- and inflammatory-related diseases (TLR signaling pathway and IL-17 signaling pathway). The number of B cells, monocytes, and activated mast cells was associated with ANCA-GN. ANCAs binds to MPO on neutrophils to activate neutrophils, which finally leads to upregulation of proinflammatory mediator genes expressing, and favors the neutrophils adhesion and migration through the endothelium. Activation of neutrophils also results in complement activation, which results in promoting inflammation and enhancing tissue damage.[17] TLRs are intrinsic pattern-recognition receptors that play a prominent role in inducing inflammatory responses. Previous studies have found elevated levels of TLR2, 3, 4, 7, and 9 in patients with AAV and intrarenal *TLR4* and *TLR2* expression correlated with renal injury.[18] Heeringa et al. found monocyte activation and whole blood IL-6 and IL-10 secretion increased after stimulation with TLRs agonists in AAV patients and healthy controls.[19] Since TLRs are a link between inflammation and autoimmunity, a study found Th17 autoimmunity induced by TLR2

Table 2. Clinical parameters of patients with AAV

Characteristic	Values
Number of patients	114
Age, (mean ± SD)	53.1 ± 19.5
Gender, (male/female)	48/66
MPO-ANCA/PR3-ANCA	106/8
Initial serum creatinine (mg/dL) (median, IQR)	3.6 (1.8, 6.5)
eGFR (mL/min/1.73 m^2)[a] (median, IQR)	24.9 (12.1, 63.6)
Urinary protein (g/24 h) (median, IQR)	0.8 (0.4, 1.7)
BVAS (median, IQR)	17.0 (13.0, 20.0)
Dialysis-dependent at presentation	28 (24.6%)
PCT (ng/mL) (median, IQR)	0.5 (0.2, 1.7)
C3 (g/L) (mean ± SD)	0.8 ± 0.3
MPO (RU/mL) (median, IQR)	138.0 (80.8, 210.8)
Total cholesterol (mmol/L) (median, IQR)	4.1 (3.3, 4.8)
TG (mmol/L) (median, IQR)	1.2 (0.9, 1.8)
HDL (mmol/L) (median, IQR)	0.9 (0.8, 1.3)
White blood cell count (*10^9/L) (median, IQR)	8.0 (6.2, 11.4)
Neutrophil count (*10^9/L) (median, IQR)	6.3 (4.4, 9.7)
Monocyte (*10^9/L) (median, IQR)	0.4 (0.2, 0.6)
Platelet count (*10^9/L) (median, IQR)	211.0 (164.0, 280.0)
Lymphocyte count (*10^9/L) (median, IQR)	1.0 (0.7, 1.6)

Normal values are mean ± SD; non-normal values are median (IQR); qualitative values are number (%).
AAV: antineutrophil cytoplasmic antibody-associated vasculitis; BVAS: Birmingham Vasculitis Activity Scores; eGFR: estimated glomerular filtration rate; IQR: interquartile range; SD: standard deviation; C3: complement 3; MPO: myeloperoxidase; PR3: proteinase 3; ANCA: antineutrophil cytoplasmic autoantibody; PCT: procalcitonin. HDL: High-density lipoprotein; TG, triglyceride.
[a]eGFR (mL/min per 1:73 m^2) = 175 × (plasma creatinine) $^{-1.234}$ × age$^{-0.179}$ × 0.79 (if female).

connection TLR2 was produced through retinoic acid receptor-related orphan nuclear receptor γt-dependent IL-17A, thus promoting glomerular injury in MPO-induced AAV.[20] In addition, IL-17 induces CXC chemokine release and promotes the synthesis and secretion of TNF-α and IL-1B, upregulates the expression of endothelial cell adhesion molecules, and activating and recruiting neutrophils.[21] Anti-IL-17 treatment or IL-17A deficiency can reduce neutrophil recruitment and accumulation of renal macrophages of anti-MPO glomerulonephritis.[22] Collectively, our results of DEGs enrichment revealed that the DEGs play an essential role in the pathological mechanism of ANCA-GN.

Over the last decade, high-throughput microarray technology has been applied to investigate new core genes associated with ANCA-GN.[23] According to these transcriptome data, we obtained six hub genes (*PTPRC*, *CD86*, *TLR2*, *IL1B*, *CSF-1R*, and *CCL2*) using logical methods. These hub genes expression levels are all significantly upregulated in ANCA-GN tissues. However, there are few studies on the role and molecular mechanisms of *PTPRC*, *CD86*, *IL1B*, *CSF-1R*, and *CCL2* in ANCA-GN. C-C motif chemokine 2 (CCL2) is a mononuclear cell chemoattractant and specifically regulates the migration of blood monocytes and tissue macrophages via binding its cell surface receptor, CCR2.[24] Little et al. found that urinary *CCL2* levels in patients with active ANCA-GN were higher than that in patients in remission and the AUC for distinguishing patients with an active stage or not was 0.794.[25] Another study suggested that initial serum *CCL2* levels were obviously increased in microscopic polyangiitis (MPA) interstitial lung disease group compared with MPA group, and initial serum *CCL2* levels were positively correlated with the degree of lung fibrosis.[26] We found *CCL2* expression levels in ANCA-GN glomeruli negatively correlated with the GFR, and the AUC for distinguishing ANCA-GN was 0.801. CD86 (B7-2) plays a decisive role in antigenic and allogeneic responses and is used as the M1 macrophage marker. Besides the neutrophils, the T cells, macrophages, and B cell were reported to play a key role in the pathogenesis of AAV.[27] Although few studies have focused on the potential role of B cells in the pathogenesis of AAV, B cell therapy has been argued as a promising treatment option for AAV in recent years.[28] Rituximab is an anti-CD20 antibody that plays a role in the treatment of B cell lymphoproliferative diseases by targeting B cell clearance. Rituximab treatment in patients with AAV can quickly clear B cells, resulting in making AAV complete remission in a short time.[29] Recently, a study reported abatacept, which works by binding to CD80 and CD86, blocking this key costimulatory signal and completely activates T lymphocytes; the study treated 20 granulomatosis with polyangiitis patients and found disease alleviation in 90% and remission (a BVAS score of 0) achievement in 80% at a median of 1.9 months.[30] Our study found that *CD86* mRNA levels were significantly higher in the ANCA-GN group

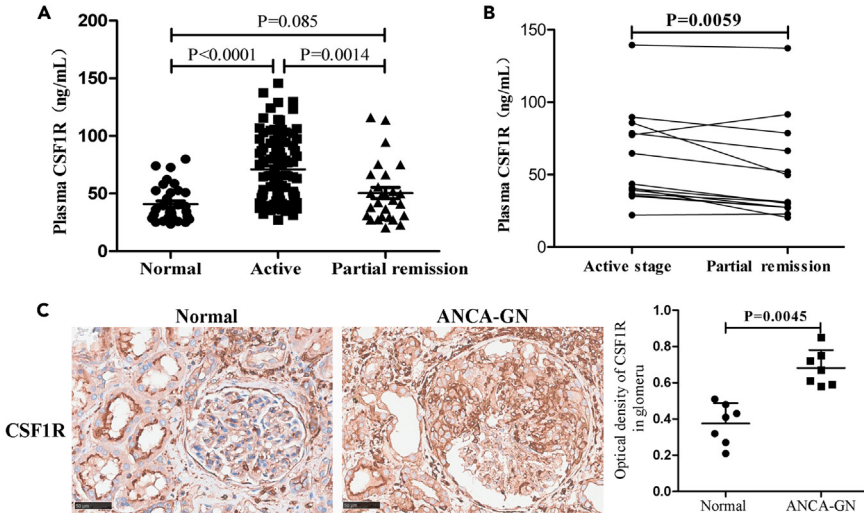


Figure 7. Plasma and renal tissues levels of CSF-1R in different groups of ANCA-GN patients

(A) Plasma levels of CSF-1R in ANCA-GN patients in the active stage (n = 87), partial remission (n = 27) and normal controls (n = 45).

(B) Changes of plasma CSF-1R levels in ANCA-GN patients with sequential plasma samples (n = 15).

(C) Representative of immunohistochemical analysis of CSF-1R expression in the glomerular from ANCA-GN patients (n = 6) and normal kidney tissues (n = 6), and the average optical density of CSF-1R was analyzed by ImageJ software. Data were presented as mean ± SD. Magnification 40×. Scale bars: 50 μm. ANCA-GN, antineutrophil cytoplasmic antibody-associated glomerulonephritis. SD, standard deviation.

compared with the control group, negatively correlated with the GFR, and correlated to immune cells according to the immune infiltration analysis.

CSF-1R is a transmembrane tyrosine kinase receptor mainly expressed in bone-marrow-derived macrophages, microglial cells, dendritic cells, and monocytes. The *CSF-1/CSF-1R* axis is involved in modulating the migration and activation of target immune cells. *CSF-1R* can also promote monocyte/macrophage chemotaxis of into inflammatory foci and initiate inflammatory functions.[31] *CSF-1R* exists in an auto-inhibited form and initiates a signal cascade through dimerization and autophosphorylation of several tyrosine residues. The activation of *CSF-1R* also induces the morphology and movement of naive macrophages through inducing actin polymerization, which are significantly related to PI3K/Akt pathway.[32] Additionally, *CSF-1R* signaling is important for the migration of macrophages, directing the cell fate toward monocytes or granulocytes, and then regulating host responses to pathogens through modulation of the expression of *TLR9*, *TLR1*, *TLR2*, and *TLR6*, but not *TLR4* or *TLR5*.[33,34] Over the past few decades, numerous studies have confirmed the potential of inhibiting *CSF-1R* signaling pathway in reducing immune response. Abnormal expression of *CSF-1R* might lead to multiple inflammatory diseases and lupus nephritis.[35] In addition, *CSF-1R* has been found to be excessively activated in inflammatory diseases, which may result in an inflammation response in the kidney and central nervous system.[32] Previous studies have shown that inhibiting *CSF-1R* by specific inhibitors, such as PLX3397 and GW2580, reduced the monocytes/macrophages infiltration and suppressed inflammatory cell expansion; thus, inhibition of the *CSF-1/CSF-1R* signaling pathway is a potential target for asthma, systemic lupus erythematosus, and rheumatoid arthritis treatment.[32] These results were consistent with the results of our GSEA analysis of *CSF-1R*, which indicated that it was involved in regulating TLR signaling pathway and leukocyte transendothelial migration. However, limited studies have focused on the role and pathological mechanism of *CSF-1R* in vasculitis and whether *CSF-1R* signaling can be identified as a biomarker for ANCA-GN. In this study, six hub genes were selected by the bioinformatics method; *CSF-1R* had a higher fold change of expression pattern, and the largest scoring for evaluating gene importance used the Gini coefficient method and AUC value for diagnostic ANCA-GN. A limitation is that both sensitivity and specificity of CSF-1R are not so high as a biomarker (sensitivity: 60.0% and specificity: 75.6%). Its clinical application may need to be combined with other biomarkers. In light of the previously mentioned background, the significance of *CSF-1R* in ANCA-GN has received less attention. Due to its exceptional clinical properties and molecular function, *CSF-1R* might be a potential therapeutic target for future clinical investigation.

In this study, we found that plasma CSF-1R levels were meaningfully higher in patients with ANCA-GN in the active stage than remission stage patients and normal controls. This result was consistent with the transcriptomic results from glomerular tissue of ANCA-GN patients. After further analysis, we found that plasma CSF-1R levels positively correlated with BVAS, initial serum creatinine, and C1Q levels but negatively correlated with GFR. These results showed that plasma CSF-1R levels could reflect disease activity in ANCA-GN patients. A previous study reported that CSF-1R expression was raised in the synovium of rheumatoid arthritis patients.[36] In addition, several studies demonstrated that the ligands of CSF-1R, CSF-1, and IL-34 were raised in the kidney tissues, serum, and urine of patients with systemic lupus erythematosus, and positively correlated with the disease activity parameters.[37,38] Considerable data identified complement activation through alternative pathways was one of the primary mechanisms behind AAV-induced glomerular damage. However, only a few investigations found C1q

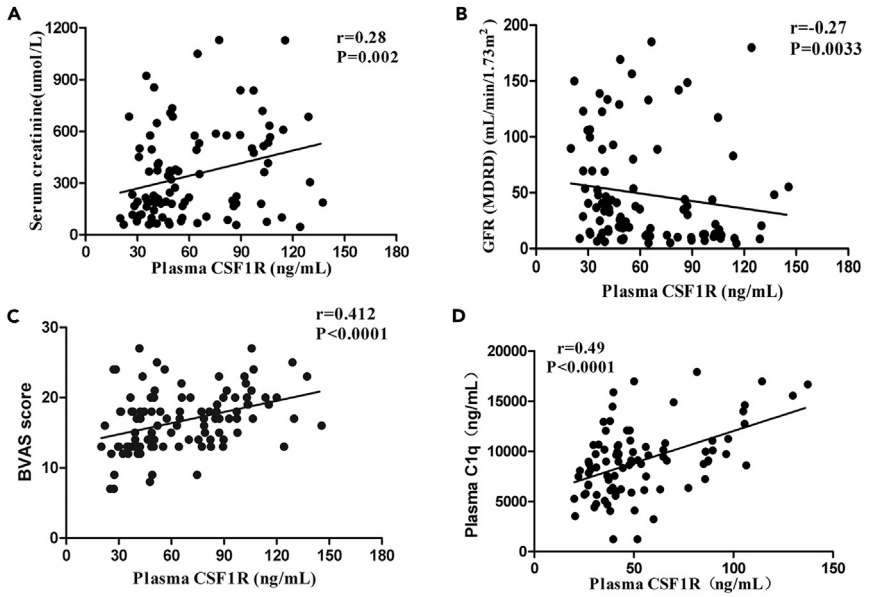


Figure 8. Plasma levels of CSF-1R correlated with clinical traits

(A) Initial serum creatinine.

(B) The glomerular filtration rate (GFR).

(C) Birmingham Vasculitis Activity Score (BVAS).

(D) Plasma levels of C1q. MDRD: modification of diet in renal disease.

and C3c in the glomerular capillary wall and/or mesangium of ANCA-GN patient specimens, which was correlated with more severe renal damage.[39] We first found that the levels of *C1QA*, *C1QB*, and *C1QC* expression were upregulated in glomerular and *CSF-1R* interactions with *C1QA*, *C1QB*, and *C1QC* by using bioinformatics analysis of ANCA-GN-related microarray data. What is more, we also found the plasma C1Q levels were higher in ANCA-GN patients compared with normal group, and there was a positive correlation in the levels of CSF-1R with C1Q levels in plasma. The aforementioned evidence shows that the abnormally expressed *CSF-1R* and *C1Q* may aggravate the development of ANCA-GN. Moreover, plasma CSF-1R levels were found to be related to the disease activity of patients with AAV, and to some extent, plasma CSF-1R levels were correlated with the renal survival and composite outcome in ANCA-GN patients.

In summary, we analyzed the gene expression profiles of ANCA-GN human glomerular tissue and screened out 79 IP-DEGs. We showed the biological process and pathway in which they were involved in the pathogenesis of ANCA-GN. Plasma levels of *CSF-1R* were significantly increased in ANCA-GN patients, correlated with disease activity, and associated with composite outcomes of ANCA-GN patients. This study

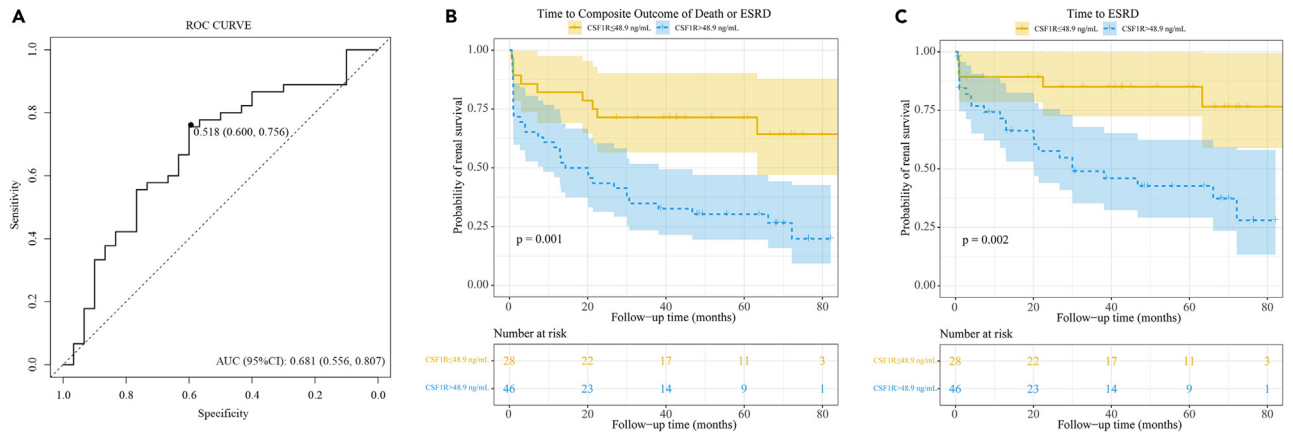


Figure 9. Association between plasma levels of CSF-1R and prognosis of patients with AAV

(A) Receiver operating characteristic (ROC) analyzed the prediction efficiency of plasma CSF-1R for composite outcomes (death or ESRD) in patients with AAV.

(B) Kaplan-Meier survival curves showed associations with composite outcomes (death or ESRD) according to CSF-1R status.

(C) Kaplan-Meier survival curves showed associations with ESRD according to CSF-1R status. ESRD, end-stage renal disease.

Table 3. Multivariate analysis of composite outcome in patients with AAV

	Univariate		Multivariate[a]		Multivariate[b]	
	HR (95%CI)	p value	HR (95%CI)	p value	HR (95%CI)	p value
Plasma CSF1R levels (CSF1R ≤ 48.9 ng/mL vs. >48.9 ng/mL)	3.14 (1.50, 6.56)	0.002	3.05 (1.45, 6.43)	0.003	1.28 (0.15, 10.80)	0.819
Age (year)[c]	1.03 (1.01, 1.06)	0.001	1.03 (1.01, 1.05)	0.001	1.07 (1.01, 1.13)	0.031
Gender (male vs. female)	1.11 (0.61, 2.02)	0.732	1.09 (0.59, 2.01)	0.780	0.15 (0.02, 1.08)	0.059
Urinary protein (per g/24 h)	0.93 (0.66, 1.31	0.678			1.59 (0.92, 2.75)	0.098
Initial serum creatinine (per mg/dL)	1.31 (1.18, 1.44)	<0.001			1.31 (1.01, 1.69)	0.042
CRP (mg/L)	1.01 (1.00, 1.01)	0.021			1.01 (0.99, 1.03)	0.362

AAV: antineutrophil cytoplasmic antibody-associated vasculitis; CRP: C-reactive protein; HR: hazard ratio; CI: confidence interval.
[a]Adjusted for age and gender.
[b]Adjusted for age, gender, initial serum creatinine, and urinary protein.
[c]per 1 unit increase.

provides new insights into potential roles of *CSF-1R* in the development of ANCA-GN, although the specific mechanism of *CSF-1R* in ANCA-GN remains unclear. Therefore, the specific mechanism of *CSF-1R* needs to be further studied.

Limitations of the study

In spite of its promising findings, this study still has some limitations. First, when 24 h urinary protein and initial serum creatinine were considered, CSF-1R was not an independent risk factor for death or ESRD of ANCA-GN patients, which may be closely related to plasma CSF-1R and renal function. Second, because this is retrospective research, it is restricted by the sample size and patient age distribution, and a small number of follow-up data are missing owing to other objective causes. Third, among Chinese AAV patients, MPO-ANCA-positive AAV patients are more common than PR3-ANCA-positive AAV patients,[40] which is consistent with our data, so our findings may be more helpful in reflecting MPO-ANCA-positive patients. Since the level of serum CSF-1R is positively correlated with eGFR, the main limitation of this paper is that it cannot be proved that the increase of serum CSF-1R level is not caused by the decrease of GFR. However, our tissue microarray analysis revealed an increase in the expression of CSF-1R in the glomeruli. Furthermore, upon reanalyzing samples with an eGFR>60 mL/min/1.73 m^2, we observed a significant increase in the expression of CSF-1R in the ANCA-GN patients (Figure S6). These findings suggest that the elevated levels of CSF-1R are not solely attributed to a decrease in GFR. This may require further collection of eGFR>60 mL/min/1.73 m^2 samples for verification. As a result, the significance of plasma CSF-1R levels as a predictor of ANCA-GN patient outcomes requires further investigation.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Patients and blood samples
  - Quantification of plasma CSF-1R, C1Q, CSF-1, IL-34 and CRP using ELISA
  - Immunohistochemistry staining and analysis
- METHOD DETAILS
  - Data acquisition and processing
  - Functional enrichment analysis
  - Protein-protein interaction (PPI) network construction and hub genes identification
  - Screening of candidate diagnostic biomarkers
  - Immune infiltration analysis
  - Differentially expressed analysis by nephroseq
  - Clinical evaluation and definitions
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108157.

ACKNOWLEDGMENTS

This study was financially supported by the foundation from the Health Commission of Hubei Province (No. ZY2021Q036).

The authors gratefully acknowledge the data provided by patients and researchers participating in GEO. We are very grateful to Mr. Du Yueliang, Director of Luohe Central Hospital, for providing the kidney biopsy specimen. The graphical abstract was created with BioRender.com.

AUTHOR CONTRIBUTIONS

Y.M.L., Y.L.D., and Q.Q.L. designed the study and experiments. Y.R.W. and C.L.C. conducted the experiments. L.H. and S.Y.L. collected and analyzed the data. Y.R.W. wrote the manuscript. Y.R.W. and C.L.C. interpreted the data and chart processing. Q.Q.L. revised the manuscript. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We worked to ensure gender balance in the recruitment of human subjects. We worked to ensure ethnic or other types of diversity in the recruitment of human subjects. We worked to ensure that the study questionnaires were prepared in an inclusive way. We worked to ensure diversity in experimental samples through the selection of the genomic datasets. While citing references scientifically relevant for this work, we also actively worked to promote gender balance in our reference list. We avoided "helicopter science" practices by including the participating local contributors from the region where we conducted the research as authors on the paper.

Received: April 15, 2023
Revised: August 21, 2023
Accepted: October 4, 2023
Published: October 7, 2023

REFERENCES

1. Kitching, A.R., Anders, H.J., Basu, N., Brouwer, E., Gordon, J., Jayne, D.R., Kullman, J., Lyons, P.A., Merkel, P.A., Savage, C.O.S., et al. (2020). ANCA-associated vasculitis. Nat. Rev. Dis. Primers *6*, 71. https://doi.org/10.1038/s41572-020-0204-y.
2. Nakazawa, D., Masuda, S., Tomaru, U., and Ishizu, A. (2019). Pathogenesis and therapeutic interventions for ANCA-associated vasculitis. Nat. Rev. Rheumatol. *15*, 91–101. https://doi.org/10.1038/s41584-018-0145-y.
3. Geetha, D., and Jefferson, J.A. (2020). ANCA-Associated Vasculitis: Core Curriculum 2020. Am. J. Kidney Dis. *75*, 124–137. https://doi.org/10.1053/j.ajkd.2019.04.031.
4. Hunter, R.W., Welsh, N., Farrah, T.E., Gallacher, P.J., and Dhaun, N. (2020). ANCA associated vasculitis. BMJ *369*, m1070. https://doi.org/10.1136/bmj.m1070.
5. Fauci, A.S., Haynes, B.F., Katz, P., and Wolff, S.M. (1983). Wegener's granulomatosis: prospective clinical and therapeutic experience with 85 patients for 21 years. Ann. Intern. Med. *98*, 76–85. https://doi.org/10.7326/0003-4819-98-1-76.
6. Wilde, B., van Paassen, P., Witzke, O., and Tervaert, J.W.C. (2011). New pathophysiological insights and treatment of ANCA-associated vasculitis. Kidney Int. *79*, 599–612. https://doi.org/10.1038/ki.2010.472.
7. Söderberg, D., and Segelmark, M. (2016). Neutrophil Extracellular Traps in ANCA-Associated Vasculitis. Front. Immunol. *7*, 256. https://doi.org/10.3389/fimmu.2016.00256.
8. Bitton, L., Vandenbussche, C., Wayolle, N., Gibier, J.B., Cordonnier, C., Verine, J., Humez, S., Bataille, P., Lenain, R., Ramdane, N., et al. (2020). Tubulointerstitial damage and interstitial immune cell phenotypes are useful predictors for renal survival and relapse in antineutrophil cytoplasmic antibody-associated vasculitis. J. Nephrol. *33*, 771–781. https://doi.org/10.1007/s40620-019-00695-y.
9. Deng, X., Yang, Q., Wang, Y., Zhou, C., Guo, Y., Hu, Z., Liao, W., Xu, G., and Zeng, R. (2020). CSF-1R inhibition attenuates ischemia-induced renal injury and fibrosis by reducing Ly6C(+) M2-like macrophage infiltration. Int. Immunopharmacol. *88*, 106854. https://doi.org/10.1016/j.intimp.2020.106854.
10. Hoheisel, J.D. (2006). Microarray technology: beyond transcript profiling and genotype analysis. Nat. Rev. Genet. *7*, 200–210. https://doi.org/10.1038/nrg1809.
11. Reel, P.S., Reel, S., Pearson, E., Trucco, E., and Jefferson, E. (2021). Using machine learning approaches for multi-omics data analysis: A review. Biotechnol. Adv. *49*, 107739. https://doi.org/10.1016/j.biotechadv.2021.107739.
12. Barrett, T., Troup, D.B., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., et al. (2011). NCBI GEO: archive for functional genomics data sets–10 years on. Nucleic Acids Res. *39*, D1005–D1010. https://doi.org/10.1093/nar/gkq1184.
13. Biundo, F., Chitu, V., Tindi, J., Burghardt, N.S., Shlager, G.G.L., Ketchum, H.C., DeTure, M.A., Dickson, D.W., Wszolek, Z.K., Khodakhah, K., and Stanley, E.R. (2023). Elevated granulocyte colony stimulating factor (CSF) causes cerebellar deficits and anxiety in a model of CSF-1 receptor related leukodystrophy. Glia *71*, 775–794. https://doi.org/10.1002/glia.24310.
14. Sada, K.E., Harigai, M., Amano, K., Atsumi, T., Fujimoto, S., Yuzawa, Y., Takasaki, Y., Banno, S., Sugihara, T., Kobayashi, M., et al. (2016). Comparison of severity classification in Japanese patients with antineutrophil cytoplasmic antibody-associated vasculitis in a nationwide, prospective, inception cohort study. Mod. Rheumatol. *26*, 730–737. https://doi.org/10.3109/14397595.2016.1140274.
15. Ramponi, G., Folci, M., De Santis, M., Damoiseaux, J.G.M.C., Selmi, C., and Brunetta, E. (2021). The biology, pathogenetic role, clinical implications, and open issues of serum anti-neutrophil cytoplasmic antibodies. Autoimmun. Rev. *20*, 102759. https://doi.org/10.1016/j.autrev.2021.102759.
16. Vegting, Y., Vogt, L., Anders, H.J., de Winther, M.P.J., Bemelman, F.J., and Hilhorst, M.L. (2021). Monocytes and macrophages in ANCA-associated vasculitis.

Autoimmun. Rev. *20*, 102911. https://doi.org/10.1016/j.autrev.2021.102911.

17. Huugen, D., van Esch, A., Xiao, H., Peutz-Kootstra, C.J., Buurman, W.A., Tervaert, J.W.C., Jennette, J.C., and Heeringa, P. (2007). Inhibition of complement factor C5 protects against anti-myeloperoxidase antibody-mediated glomerulonephritis in mice. Kidney Int. *71*, 646–654. https://doi.org/10.1038/sj.ki.5002103.
18. O'Sullivan, K.M., Ford, S.L., Longano, A., Kitching, A.R., and Holdsworth, S.R. (2018). Intrarenal Toll-like receptor 4 and Toll-like receptor 2 expression correlates with injury in antineutrophil cytoplasmic antibody-associated vasculitis. Am. J. Physiol. Renal Physiol. *315*, F1283–F1294. https://doi.org/10.1152/ajprenal.00040.2018.
19. Tadema, H., Abdulahad, W.H., Stegeman, C.A., Kallenberg, C.G.M., and Heeringa, P. (2011). Increased expression of Toll-like receptors by monocytes and natural killer cells in ANCA-associated vasculitis. PLoS One *6*, e24315. https://doi.org/10.1371/journal.pone.0024315.
20. Summers, S.A., Steinmetz, O.M., Gan, P.Y., Ooi, J.D., Odobasic, D., Kitching, A.R., and Holdsworth, S.R. (2011). Toll-like receptor 2 induces Th17 myeloperoxidase autoimmunity while Toll-like receptor 9 drives Th1 autoimmunity in murine vasculitis. Arthritis Rheum. *63*, 1124–1135. https://doi.org/10.1002/art.30208.
21. Nogueira, E., Hamour, S., Sawant, D., Henderson, S., Mansfield, N., Chavele, K.M., Pusey, C.D., and Salama, A.D. (2010). Serum IL-17 and IL-23 levels and autoantigen-specific Th17 cells are elevated in patients with ANCA-associated vasculitis. Nephrol. Dial. Transplant. *25*, 2209–2217. https://doi.org/10.1093/ndt/gfp783.
22. Gan, P.Y., Steinmetz, O.M., Tan, D.S.Y., O'Sullivan, K.M., Ooi, J.D., Iwakura, Y., Kitching, A.R., and Holdsworth, S.R. (2010). Th17 cells promote autoimmune anti-myeloperoxidase glomerulonephritis. J. Am. Soc. Nephrol. *21*, 925–931. https://doi.org/10.1681/asn.2009070763.
23. Yanaoka, H., Nagafuchi, Y., Hanata, N., Takeshima, Y., Ota, M., Suwa, Y., Shirai, H., Sugimori, Y., Okubo, M., Kobayashi, S., et al. (2021). Identifying the most influential gene expression profile in distinguishing ANCA-associated vasculitis from healthy controls. J. Autoimmun. *119*, 102617. https://doi.org/10.1016/j.jaut.2021.102617.
24. Raghu, H., Lepus, C.M., Wang, Q., Wong, H.H., Lingampalli, N., Oliviero, F., Punzi, L., Giori, N.J., Goodman, S.B., Chu, C.R., et al. (2017). CCL2/CCR2, but not CCL5/CCR5, mediates monocyte recruitment, inflammation and cartilage destruction in osteoarthritis. Ann. Rheum. Dis. *76*, 914–922. https://doi.org/10.1136/annrheumdis-2016-210426.
25. Moran, S.M., Monach, P.A., Zgaga, L., Cuthbertson, D., Carette, S., Khalidi, N.A., Koening, C.L., Langford, C.A., McAlear, C.A., Moreland, L., et al. (2020). Urinary soluble CD163 and monocyte chemoattractant protein-1 in the identification of subtle renal flare in anti-neutrophil cytoplasmic antibody-associated vasculitis. Nephrol. Dial. Transplant. *35*, 283–291. https://doi.org/10.1093/ndt/gfy300.
26. Matsuda, S., Kotani, T., Kuwabara, H., Suzuka, T., Kiboshi, T., Fukui, K., Ishida, T., Fujiki, Y., Shiba, H., Hata, K., et al. (2021). CCL2 produced by CD68+/CD163+ macrophages as a promising clinical biomarker of microscopic polyangiitis-interstitial lung disease. Rheumatology *60*, 4643–4653. https://doi.org/10.1093/rheumatology/keab064.
27. Dumoitier, N., Terrier, B., London, J., Lofek, S., and Mouthon, L. (2015). Implication of B lymphocytes in the pathogenesis of ANCA-associated vasculitides. Autoimmun. Rev. *14*, 996–1004. https://doi.org/10.1016/j.autrev.2015.06.008.
28. McClure, M., Gopaluni, S., Jayne, D., and Jones, R. (2018). B cell therapy in ANCA-associated vasculitis: current and emerging treatment options. Nat. Rev. Rheumatol. *14*, 580–591. https://doi.org/10.1038/s41584-018-0065-x.
29. Eriksson, P. (2005). Nine patients with anti-neutrophil cytoplasmic antibody-positive vasculitis successfully treated with rituximab. J. Intern. Med. *257*, 540–548. https://doi.org/10.1111/j.1365-2796.2005.01494.x.
30. Langford, C.A., Monach, P.A., Specks, U., Seo, P., Cuthbertson, D., McAlear, C.A., Ytterberg, S.R., Hoffman, G.S., Krischer, J.P., and Merkel, P.A.; Vasculitis Clinical Research Consortium (2014). An open-label trial of abatacept (CTLA4-IG) in non-severe relapsing granulomatosis with polyangiitis (Wegener's). Ann. Rheum. Dis. *73*, 1376–1379. https://doi.org/10.1136/annrheumdis-2013-204164.
31. Barca, C., Foray, C., Hermann, S., Herrlinger, U., Remory, I., Laoui, D., Schäfers, M., Grauer, O.M., Zinnhardt, B., and Jacobs, A.H. (2021). The Colony Stimulating Factor-1 Receptor (CSF-1R)-Mediated Regulation of Microglia/Macrophages as a Target for Neurological Disorders (Glioma, Stroke). Front. Immunol. *12*, 787307. https://doi.org/10.3389/fimmu.2021.787307.
32. Xiang, C., Li, H., and Tang, W. (2023). Targeting CSF-1R represents an effective strategy in modulating inflammatory diseases. Pharmacol. Res. *187*, 106566. https://doi.org/10.1016/j.phrs.2022.106566.
33. Chitu, V., and Stanley, E.R. (2006). Colony-stimulating factor-1 in immunity and inflammation. Curr. Opin. Immunol. *18*, 39–48. https://doi.org/10.1016/j.coi.2005.11.006.
34. Sweet, M.J., Campbell, C.C., Sester, D.P., Xu, D., McDonald, R.C., Stacey, K.J., Hume, D.A., and Liew, F.Y. (2002). Colony-stimulating factor-1 suppresses responses to CpG DNA and expression of toll-like receptor 9 but enhances responses to lipopolysaccharide in murine macrophages. J. Immunol. *168*, 392–399. https://doi.org/10.4049/jimmunol.168.1.392.
35. Chalmers, S.A., Wen, J., Shum, J., Doerner, J., Herlitz, L., and Putterman, C. (2017). CSF-1R inhibition attenuates renal and neuropsychiatric disease in murine lupus. Clin. Immunol. *185*, 100–108. https://doi.org/10.1016/j.clim.2016.08.019.
36. Toh, M.L., Bonnefoy, J.Y., Accart, N., Cochin, S., Pohle, S., Haegel, H., De Meyer, M., Zemmour, C., Preville, X., Guillen, C., et al. (2014). Bone- and cartilage-protective effects of a monoclonal antibody against colony-stimulating factor 1 receptor in experimental arthritis. Arthritis Rheumatol. *66*, 2989–3000. https://doi.org/10.1002/art.38624.
37. Wang, H., Cao, J., and Lai, X. (2016). Serum Interleukin-34 Levels Are Elevated in Patients with Systemic Lupus Erythematosus. Molecules *22*, 35. https://doi.org/10.3390/molecules22010035.
38. Menke, J., Amann, K., Cavagna, L., Blettner, M., Weinmann, A., Schwarting, A., and Kelley, V.R. (2015). Colony-stimulating factor-1: a potential biomarker for lupus nephritis. J. Am. Soc. Nephrol. *26*, 379–389. https://doi.org/10.1681/asn.2013121356.
39. Chen, M., Xing, G.Q., Yu, F., Liu, G., and Zhao, M.H. (2009). Complement deposition in renal histopathology of patients with ANCA-associated pauci-immune glomerulonephritis. Nephrol. Dial. Transplant. *24*, 1247–1252. https://doi.org/10.1093/ndt/gfn586.
40. Hong, Y., Shi, P., Liu, X., Yang, L., Li, K., Xu, F., Liang, S., Liu, Z., Zhang, H., Chen, Y., and Hu, W. (2019). Distinction between MPO-ANCA and PR3-ANCA-associated glomerulonephritis in Chinese patients: a retrospective single-center study. Clin. Rheumatol. *38*, 1665–1673. https://doi.org/10.1007/s10067-019-04458-9.
41. Grayson, P.C., Eddy, S., Taroni, J.N., Lightfoot, Y.L., Mariani, L., Parikh, H., Lindenmeyer, M.T., Ju, W., Greene, C.S., Godfrey, B., et al. (2018). Metabolic pathways and immunometabolism in rare kidney diseases. Ann. Rheum. Dis. *77*, 1226–1233. https://doi.org/10.1136/annrheumdis-2017-212935.
42. Brix, S.R., Stege, G., Disteldorf, E., Hoxha, E., Krebs, C., Krohn, S., Otto, B., Klätschke, K., Herden, E., Heymann, F., et al. (2015). CC Chemokine Ligand 18 in ANCA-Associated Crescentic GN. J. Am. Soc. Nephrol. *26*, 2105–2117. https://doi.org/10.1681/ASN.2014040407.
43. Otasek, D., Morris, J.H., Bouças, J., Pico, A.R., and Demchak, B. (2019). Cytoscape Automation: empowering workflow-based network analysis. Genome Biol. *20*, 185. https://doi.org/10.1186/s13059-019-1758-4.
44. Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A.L., Fang, T., Doncheva, N.T., Pyysalo, S., et al. (2023). The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. *51*, D638–D646. https://doi.org/10.1093/nar/gkac1000.
45. Villanueva, R.A.M., and Chen, Z.J. (2019). ggplot2: Elegant Graphics for Data Analysis, 2nd edition. Meas. Interdiscipl. Res. *17*, 160–167. https://doi.org/10.1080/15366367.2019.1565254.
46. Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M., and Alizadeh, A.A. (2015). Robust enumeration of cell subsets from tissue expression profiles. Nat. Methods *12*, 453–457. https://doi.org/10.1038/nmeth.3337.
47. Jennette, J.C., Falk, R.J., Bacon, P.A., Basu, N., Cid, M.C., Ferrario, F., Flores-Suarez, L.F., Gross, W.L., Guillevin, L., Hagen, E.C., et al. (2013). 2012 revised International

Chapel Hill Consensus Conference Nomenclature of Vasculitides. Arthritis Rheum. *65*, 1–11. https://doi.org/10.1002/art.37715.

48. Bhattacharya, S., Andorf, S., Gomes, L., Dunn, P., Schaefer, H., Pontius, J., Berger, P., Desborough, V., Smith, T., Campbell, J., et al. (2014). ImmPort: disseminating data to the public for the future of immunology. Immunol. Res. *58*, 234–239. https://doi.org/10.1007/s12026-014-8516-1.
49. Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287. https://doi.org/10.1089/omi.2011.0118.
50. Franz, M., Rodriguez, H., Lopes, C., Zuberi, K., Montojo, J., Bader, G.D., and Morris, Q. (2018). GeneMANIA update 2018. Nucleic Acids Res. *46*, W60–W64. https://doi.org/10.1093/nar/gky311.
51. Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. *47*, D607–D613. https://doi.org/10.1093/nar/gky1131.
52. Luqmani, R.A., Bacon, P.A., Moots, R.J., Janssen, B.A., Pall, A., Emery, P., Savage, C., and Adu, D. (1994). Birmingham Vasculitis Activity Score (BVAS) in systemic necrotizing vasculitis. QJM *87*, 671–678.
53. Huang, F., Li, Y., Xu, R., Cheng, A., Lv, Y., and Liu, Q. (2020). The Plasma Soluble Urokinase Plasminogen Activator Receptor Is Related to Disease Activity of Patients with ANCA-Associated Vasculitis. Mediators Inflamm. *2020*, 7850179. https://doi.org/10.1155/2020/7850179.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
CSF-1R	Beyotime	RRID: AB_3068527
Biological samples		
The blood samples of antineutrophil cytoplasmic antibody-associated glomerulonephritis patients and healthy controls.	Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology	N/A
Renal puncture tissues of antineutrophil cytoplasmic antibody-associated glomerulonephritis patients and healthy controls.	Luohe Central Hospital	N/A
Critical commercial assays		
Human CSF1R/M-CSFR ELISA Kit	Boster	EK0807
C1q Human ELISA Kit	Boster	BMS2099
Human M-CSF ELISA Kit	Boster	EK0444
Human IL-34 ELISA Kit	Boster	EK1365
Human CRP ELISA Kit	Boster	EK1316
Deposited data		
transcriptome data	Grayson et al.[41]	GEO: GSE108113
transcriptome data	Grayson et al.[41]	GEO: GSE104948
transcriptome data	Brix et al.[42]	ArrayExpress: E-MTAB-1944
transcriptome data	Nephroseq (https://nephroseq.org/)	Nephroseq :Ju CKD Glom
Software and algorithms		
R (version 4.1)	https://www.r-project.org/	N/A
Cytoscape (version 3.8.2)	http://www.cytoscape.org/ [43]	N/A
STRING database	https://string-db.org/ [44]	N/A
SPSS software (version 23)	https://www.ibm.com/spss	N/A
GraphPad Prism software (version 8)	https://www.graphpad-prism.cn/	N/A
ggplot2	R package[45]	N/A
CIBERSORT	R package[46]	N/A
randomForest	R package (https://journal.r-project.org/articles/RN-2002-022/RN-2002-022.pdf)	N/A
Image-Pro Plus software (Version 6.0)	https://mediacy.com/image-pro/	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Qingquan Liu, qqliutj@163.com.

Materials availability

This study did not generate new unique reagents.

Data and code availability

Data: This paper analyzes existing, publicly available data. These accession numbers for the datasets are listed in the key resources table.
Code: Requests for original code should be directed to the corresponding authors.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Patients and blood samples

This study included 114 patients with ANCA-GN diagnosed in the Department of Nephrology of Tongji Hospital from 2016 to 2018 (median age: 59 years old, range: 22-82; Male: Female = 48:66). The serotypes were PR3-ANCA or MPO-ANCA positive. All patients followed the definition criteria of AAV at the 2012 Chapel Hill Consensus Conference.[47] Survival follow-up was done every 3 months until patient death, loss of follow-up or the end of follow-up. The normal control group consisted of 45 health examiners selected from the Health Management Center of Tongji Hospital in 2018 (median age: 39 years old, range: 22-60; Male: Female = 25:20). The blood samples of each participant were collected and centrifuged by 3000 rpm for 10 minutes within 30 minutes. The blood samples both AAV patients and healthy controls were stored at -80°C and in single-use aliquots to prevent freeze-thaw cycles. This study was approved by the Medical Ethics Committee of Tongji Hospital of Huazhong University of Science and Technology (TJ-IRB20220159). The Medical Ethics Committee granted an exemption from the requirement for informed consent because the serum samples we collected were the samples left over from the participants' routine blood tests, and the study would not affect the rights or health of participants. All research activities were in line with the Declaration of Helsinki 2000.

Quantification of plasma CSF-1R, C1Q, CSF-1, IL-34 and CRP using ELISA

The plasma proteins levels of CSF-1R (Boster, EK0807), C1Q (Invitrogen, BMS209), CSF-1 (Boster, EK0444), IL-34 (Boster, EK1365) and CRP (Boster, EK1316) were measured using the Quantikine enzyme-linked immunosorbent assay (ELISA) according to the manufacturer's protocol. The levels of these protein from each sample were calculated using Curve Expert 1.3 (Hyams DG, Starkville, Mississippi, USA). Then the linear part of the standard curve was applied to the measurement of plasma CSF-1R, C1Q, CSF-1, IL-34 and CRP. Duplicate measurements were performed, and the coefficients of variation in all measurements were less than 10%.

Immunohistochemistry staining and analysis

Renal puncture tissues were fixed with 4% formalin and embedded in paraffin. After that, 3-mm-thick tissue slices were prepared, rehydrated and labeled with primary antibodies against CSF-1R (Beyotime, AG1701, 1:200) overnight at 4°C. The slides were washed in phosphate-buffered saline, and then the slides were incubated with peroxidase-labeled goat anti-rabbit secondary antibody for 30 minutes at room temperature. The slides were then counter-stained with hematoxylin stains. The mean optical density of CSF-1R expression on glomeruli was analyzed by Image-Pro Plus software (Version 6.0). The optical density of each glomeruli was estimated by the tissue IOD values per unit area of each measured field. We randomly selected 10 serial unduplicated high-power fields (400×) glomerulus from each slide, and the images were imported into the Image-Pro Plus software for analysis. The results were taken as an average of selected fields.

METHOD DETAILS

Data acquisition and processing

1793 immune-related genes were collected from the ImmPort database (https://www.immport.org/home).[48] There is a gene list containing immune-related genes in Resource project. We entered the website in order of "Resource - Additional Resources - Gene Lists" and downloaded the gene lists. The GEO database (https://www.ncbi.nlm.nih.gov/geo/) and the ArrayExpress database (https://www.ebi.ac.uk/arrayexpress/) were used to collected AAV gene expression data. The bioinformatics analysis only included glomerular transcriptome data. The GEO: GSE108113 and GEO: GSE104948 datasets were obtained from GEO database and contained 6 normal samples and 15 AAV samples, and 18 normal samples and 22 AAV samples, respectively.[41] The ArrayExpress: E-MTAB-1944 dataset[42] was from the ArrayExpress database and contained 12 normal samples and 30 AAV samples. The ArrayExpress: E-MTAB-1944 dataset was selected as the validation set. The detail information of these datasets was shown in Table S1. We used R package "sva" to merge data from GEO: GSE108113 and GEO: GSE104948. During the conversion process of probe ID and gene symbol, we convert the probe IDs into gene symbols according to the annotation files of their respective platforms, and deleted the probes that do not correspond to the gene symbols. If a gene has multiple probe sites, we use average value of the probe sites to represent the level of gene expression. We use "sva" R package to eliminate the batch effect of different datasets and microarray data is often plagued with a high false discovery rate. Due to some genes lack corresponding transcription data in the microarray, 1134 immune-related genes were used for follow-up analysis finally. R package "limma" was used to find DEGs between AAV and normal samples. Genes were considered as DEGs with adjust $p < 0.05$ and $|\log2(\text{fold change})| > 1$. R package "Pheatmap" was performed to make a heatmap of DEGs.

Functional enrichment analysis

R package "ClusterProfiler" was employed for GO, KEGG enrichment analysis and GSEA.[49] R package "GSVA" was applied for Gene set variation analysis (GSVA). The GeneMANIA database (http://genemania.org/) was used for clustering annotation of biological functions, which provided comprehensive understanding of the connection between hub genes and related genes.[50]

Protein-protein interaction (PPI) network construction and hub genes identification

The STRING database (https://string-db.org/) (version 11.5) was utilized to retrieve interacting genes for the PPI network.[51] We set the level of confidence to the medium confidence (0.4) and hid unconnected nodes in the network. Then Cytoscape software (version 3.8.2) was utilized for PPI network visualization. Cytohubba in Cytoscape was used to verify the significant hub genes within the PPI network. Here, Cytohubba included 8 advanced algorithms: Stress, Betweenness, Radiality, Closeness, EPC, Degree, MNC and MCC. Genes that appeared in all 8 algorithms were selected for follow-up analysis.

Screening of candidate diagnostic biomarkers

The RF algorithm was used to screen for candidate diagnostic biomarkers. The R package "random Forest" was used to perform the random forest algorithm. It computes the error rate of each tree in a set of 1-2000 trees, constructs random forest based on the number of trees with the lowest error rate and the highest stability, and selects important genes as candidate genes based on the precision decreasing method. The importance of genes greater than 1 is used as the screening standard in random forest algorithm.

Immune infiltration analysis

The "CIBERSORT" algorithm was employed to explore the levels of immune infiltration between normal and ANCA-GN-related kidney tissue. CIBERSORT collectively used the input matrix of reference gene expression signatures to estimate the relative proportions of each cell type of interest.[46]

Differentially expressed analysis by nephroseq

We used the Nephroseq database (https://nephroseq.org/) to verify the expression of hub genes in ANCA-GN glomeruli and the Spearman's correlation analysis was used to assess the correlation between the GFR and the hub genes. P-value <0.05 was statistically significant.

Clinical evaluation and definitions

The disease activity of ANCA-GN was assessed by the Birmingham Vasculitis Activity Score (BVAS).[52] The definitions of active stage, complete remission stage, and partial remission were same as we described previously.[53] In brief, patients with active disease were interpreted as new onset of ANCA-GN or relapse of ANCA-GN with BVAS $\geq$1. "Complete remission" was defined as "absence of disease activity after treatment". "Partial remission" was defined as a decrease in a patient's disease activity BVAS score of more than 50% and no new symptoms. Patients with ANCA-GN were followed up through the outpatient department. The composite endpoint was defined as the patient's renal function deteriorating to ESRD or all-cause death. ESRD was defined as the need for renal replacement therapy (such as hemodialysis or peritoneal dialysis).

QUANTIFICATION AND STATISTICAL ANALYSIS

Data of continuous normally distributed were expressed as mean $\pm$ standard deviation, nonnormally distributed continuous variables are expressed as medians and interquartile range, and categorical variables are shown as frequencies. T test or nonparametric test were used to analyze the differences between quantitative parameters of two groups. The one-way analysis of variance (ANOVA) followed by Bonferroni multiple-comparison test was performed for multiple comparisons. The paired t-test was used to assess of plasma CSF-1R levels of patients between acute and remission stage. Assumption of normal distribution was tested in the paired sample using the Shapiro–Wilk Test. The data were normally distributed except for psychoticism, somatization, and trauma symptomatic in the post-treatment measurement. Pearson's test or Spearman's test were used for correlation analysis. Using receiver operating characteristic (ROC) curve, the accuracy of CSF-1R for the prediction of ANCA-GN prognosis was evaluated. Based on the ROC curve, the optimal cutoff value was obtained using Youden's Index. The Kaplan-Meier curve was performed to analyze outcomes, and the log-rank test was used to determine the statistical significance. Univariate and multivariate Cox regression models were performed to analyze the relationship between CSF-1R levels and the composite outcome. SPSS software (SPSS, Chicago, IL) was used for statistical analysis. In addition, the GraphPad Prism software (Graph Software, San Diego, CA) was applied to drawing graphs. $p < 0.05$ was considered significant for differences.