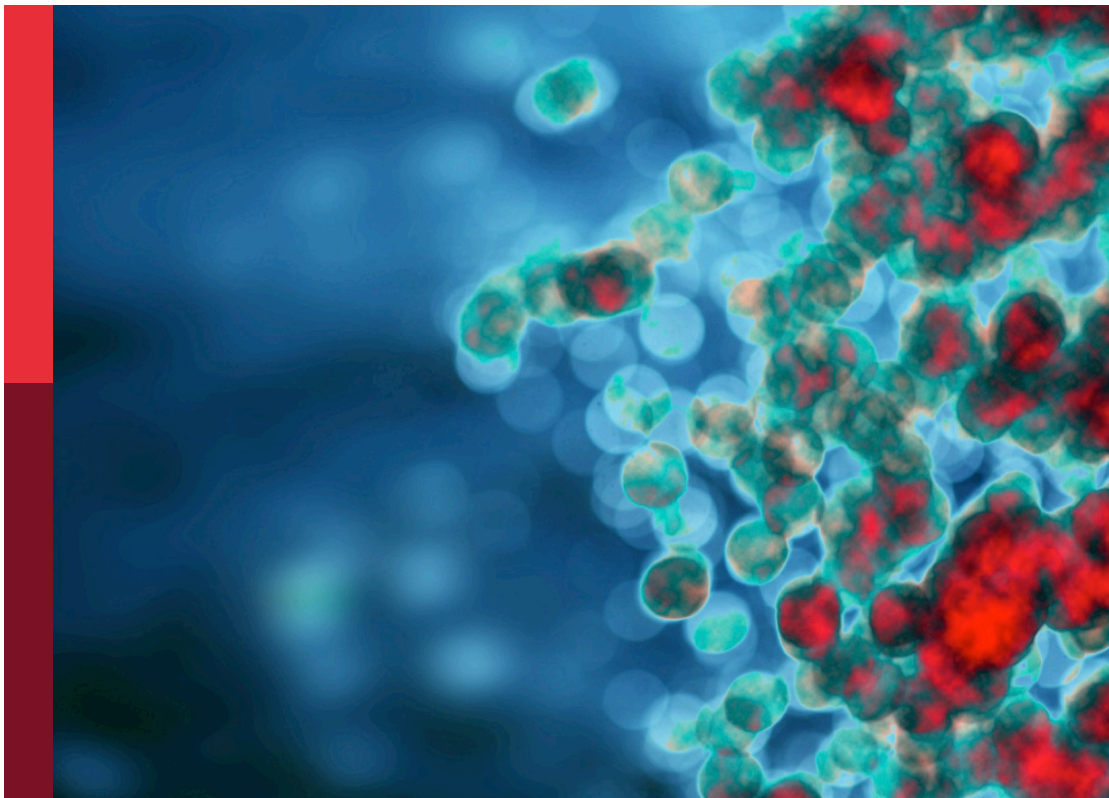
frontiers | Research Topics

Unraveling immune metabolism: single-cell & spatial transcriptomics illuminate disease dynamics

Edited by
Yejun Tan, Zhengtao Liu and Yafeng Zhu

Published in
Frontiers in Immunology
Frontiers in Endocrinology



ISSN 1664-8714
ISBN 978-2-8325-7480-5
DOI 10.3389/978-2-8325-7480-5

Generative AI statement
Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Unraveling immune metabolism: single-cell & spatial transcriptomics illuminate disease dynamics

Topic editors

Yejun Tan — Hong Kong Polytechnic University, Hong Kong, SAR China

Zhengtao Liu — Zhejiang University, China

Yafeng Zhu — Sun Yat-sen Memorial Hospital, China

Citation

Tan, Y., Liu, Z., Zhu, Y., eds. (2026). *Unraveling immune metabolism: single-cell & spatial transcriptomics illuminate disease dynamics*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-7480-5

Table of contents

frontiers | Frontiers in Endocrinology

TYPE Editorial
PUBLISHED 23 January 2026
DOI 10.3389/fendo.2026.1779505

OPEN ACCESS

EDITED AND REVIEWED BY
Ralf Jockers,
Université Paris Cité, France

*CORRESPONDENCE
Zhengtao Liu
liuzhengtao@zjsru.edu.cn

†These authors have contributed equally to this work

RECEIVED 02 January 2026
ACCEPTED 14 January 2026
PUBLISHED 23 January 2026

CITATION
Tan Y, Zhu Y and Liu Z (2026) Editorial: Unraveling immune metabolism: single-cell & spatial transcriptomics illuminate disease dynamics. *Front. Endocrinol.* 17:1779505. doi: 10.3389/fendo.2026.1779505

Editorial: Unraveling immune metabolism: single-cell & spatial transcriptomics illuminate disease dynamics

Yejun Tan[1†], Yafeng Zhu[2†] and Zhengtao Liu[3,4*]

[1]Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong, Hong Kong SAR, China, [2]Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Guangdong-Hong Kong Joint Laboratory for RNA Medicine, Medical Research Center, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, China, [3]Key Laboratory of Artificial Organs and Computational Medicine in Zhejiang Province, Shulan International Medical College, Zhejiang Shuren University, Hangzhou, China, [4]NHC Key Laboratory of Combined Multi-Organ Transplantation, Key Laboratory of the Diagnosis and Treatment of Organ Transplantation, First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China

KEYWORDS

immune microenvironment, immunometabolism, metabolic reprogramming, scRNA sequencing, spatial transcriptomics

Editorial on the Research Topic
Unraveling immune metabolism: single-cell & spatial transcriptomics illuminate disease dynamics

Introduction

The interplay between cellular metabolism and immune function—immunometabolism—has emerged as a cornerstone of modern pathology (1). Immune cells are not static entities; they continuously adapt their metabolic programs to survive and function within hostile microenvironments, whether in the hypoxic core of a tumor, the inflamed synovium of an arthritic joint, or the fibrotic tissue of a failing kidney (1–5). Historically, our understanding of these processes was limited by bulk analyses that averaged metabolic signals across heterogeneous cell populations (6). However, the advent of single-cell RNA sequencing (scRNA-seq) and spatial transcriptomics has precipitated a paradigm shift (6, 7). We can now dissect the metabolic heterogeneity of immune cells at high resolution, mapping how specific metabolic pathways drive disease progression, resistance to therapy, and tissue remodeling (6).

This Research Topic, Unraveling Immune Metabolism: Single-Cell & Spatial Transcriptomics Illuminate Disease Dynamics, was curated to bridge the gap between static metabolic profiling and dynamic disease pathology. The Research Topic published here spans a diverse spectrum of conditions—from solid tumors and renal disease to autoimmune disorders and cardiovascular failure. Collectively, they demonstrate how

metabolic rewiring is not merely a consequence of disease, but a fundamental driver of the immune landscape.

Reshaping the tumor microenvironment

Nowhere is metabolic competition more fierce than in the tumor microenvironment (TME), where cancer cells and immune cells vie for limited nutrients. Several contributions to this topic highlight how spatial and single-cell technologies are decoding this competition.

In the context of colorectal cancer, Wang et al. utilized single-cell and spatial transcriptomics to construct a high-resolution map of tumor heterogeneity. Their work reveals distinct molecular programs that govern the spatial distribution of immune cells, offering new targets for disrupting the tumor-supportive niche. Similarly, Fu et al. investigated lung adenocarcinoma, identifying the Midkine (MDK)-Nucleolin (NCL) pathway as a critical regulator of the immunosuppressive environment. By integrating spatial data, they demonstrated how this pathway orchestrates immune exclusion, suggesting that metabolic or signaling interventions targeting MDK-NCL could reinvigorate anti-tumor immunity.

Two comprehensive reviews further elucidate the metabolic hurdles within the TME. Chen et al. focused on gastric cancer, detailing how aberrant lipid metabolism reshapes the immune microenvironment to favor tumor growth. Chen et al. extended this discussion to Triple-Negative Breast Cancer (TNBC), synthesizing evidence on how metabolic plasticity limits the efficacy of immunotherapy and proposing metabolic vulnerabilities that could be exploited for combined treatment strategy.

Metabolic reprogramming in renal and systemic disease

Beyond oncology, this topic emphasizes the critical role of immunometabolism in chronic inflammatory and metabolic diseases. The progression from Acute Kidney Injury (AKI) to Chronic Kidney Disease (CKD) represents a complex metabolic shift. Zeng et al. applied integrated transcriptomics to identify key genes—CLCNKB, KLK1, and PLEKHA4—that mark this transition, providing potential biomarkers for early intervention. Complementing this, Li et al. employed a multi-omics and network pharmacology approach to validate the Jianpi-Yishen formula, a traditional intervention, revealing its capacity to modulate metabolic networks in CKD.

In the realm of systemic metabolic disorders, Li et al. utilized scRNA-seq to explore Type 2 Diabetes Mellitus (T2DM). Their study uncovers distinct immunometabolic alterations in peripheral blood mononuclear cells, linking specific immune subtypes to the systemic metabolic dysregulation characteristic of diabetes.

Autoimmunity, inflammation, and stress responses

The plasticity of macrophages and T cells is central to autoimmune pathology. Jiang et al. provided a compelling analysis of Rheumatoid Arthritis (RA), specifically the ACPA-negative subtype. Their scRNA-seq analysis highlighted a unique macrophage expansion driven by metabolic reprogramming, distinguishing the pathogenesis of this subtype from classical RA and suggesting that metabolic inhibition could be a viable therapeutic avenue for these patients.

Finally, the Research Topic addresses how immune metabolism responds to systemic stress and hypoxia. Wang et al. probed heart failure through the lens of immunogenic cell death (ICD), identifying transcriptomic biomarkers that link cell death pathways to immune activation in cardiac tissue. In a study connecting hypoxia to systemic inflammation, Ye et al. used interpretable machine learning to decode the "hypoxia-exosome-immune triad" in Obstructive Sleep Apnea (OSA). They revealed how the PRCP/UCHL1/BTG2 axis drives metabolic dysregulation, offering a novel mechanistic view of how sleep-disordered breathing impacts immune health.

Conclusion

The studies presented in Unraveling Immune Metabolism collectively reinforce the concept that metabolism is not merely the energy source for immune cells, but the instruction manual for their function. By leveraging single-cell and spatial technologies, these authors have moved beyond static snapshots to reveal the dynamic, location-specific metabolic engines driving disease. As we look to the future, the integration of these transcriptomic maps with direct metabolite sensing and flux analysis will be the next frontier, promising precision therapies that target the metabolic heartbeat of pathology.

Author contributions

YT: Writing – original draft, Writing – review & editing. YZ: Writing – original draft, Writing – review & editing. ZL: Writing – original draft, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. Funding for this study was provided by the Hangzhou Natural Science Foundation (2025SZRJJ1736), Tianqing Liver Diseases Research Fund (TQGB20200114), Chinese Society of Clinical Oncology Bayer Tumor Research Fund (Y-bayer202001/zb-0003), Chen XiaoPing

Foundation for the Development of Science and Technology of Hubei Province (CXPJJH122002-078), Beijing iGandan Foundation (1082022-RGG022), Zhejiang Shuren University Basic Scientific Research Special Funds (2023XZ010) and Key Laboratory of Artificial Organs and Computational Medicine of Zhejiang Province (SZD2025B014).

Acknowledgments

We extend our gratitude to all the authors, reviewers, and editors who contributed to this Research Topic.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

1. O'Neill LA, Kishton RJ, Rathmell J. A guide to immunometabolism for immunologists. *Nat Rev Immunol*. (2016) 16:553–65. doi: 10.1038/nri.2016.70

2. Kaymak I, Williams KS, Cantor JR, Jones RG. Immunometabolic interplay in the tumor microenvironment. *Cancer Cell*. (2021) 39:28–37. doi: 10.1016/j.ccell.2020.09.004

3. Weyand CM, Goronzy JJ. Immunometabolism in the development of rheumatoid arthritis. *Immunol Rev*. (2020) 294:177–87. doi: 10.1111/imr.12838

4. Miguel V, Shaw IW, Kramann R. Metabolism at the crossroads of inflammation and fibrosis in chronic kidney disease. *Nat Rev Nephrol*. (2025) 21:39–56. doi: 10.1038/s41581-024-00889-z

5. Basso PJ, Andrade-Oliveira V, Câmara N. Targeting immune cell metabolism in kidney diseases. *Nat Rev Nephrol*. (2021) 17:465–80. doi: 10.1038/s41581-021-00413-7

6. Artyomov MN, Van den Bossche J. Immunometabolism in the single-cell era. *Cell Metab*. (2020) 32:710–25. doi: 10.1016/j.cmet.2020.09.013

7. Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*. (2016) 353:78–82. doi: 10.1126/science.aaf2403

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 03 January 2025
DOI 10.3389/fimmu.2024.1512483

OPEN ACCESS

EDITED BY
Zhengtao Liu,
Zhejiang University, China

REVIEWED BY
Jie Ding,
Shanghai Jiao Tong University, China
Pei Bo,
Wuhan University, China

*CORRESPONDENCE
Jin Kang
kangjin2015@csu.edu.cn

RECEIVED 16 October 2024
ACCEPTED 26 November 2024
PUBLISHED 03 January 2025

CITATION
Jiang Y, Hu Z, Huang R, Ho K, Wang P and Kang J (2025) Metabolic reprogramming and macrophage expansion define ACPA-negative rheumatoid arthritis: insights from single-cell RNA sequencing. *Front. Immunol.* 15:1512483. doi: 10.3389/fimmu.2024.1512483

Metabolic reprogramming and macrophage expansion define ACPA-negative rheumatoid arthritis: insights from single-cell RNA sequencing

Yafeng Jiang[1], Zhaolan Hu[2], Roujie Huang[3], Kaying Ho[4], Pengfei Wang[2] and Jin Kang[5,6*]

[1]Department of Hematology, the Second Xiangya Hospital of Central South University, Changsha, China, [2]Department of Anesthesiology, The Second Xiangya Hospital, Central South University, Changsha, China, [3]Department of Obstetrics and Gynecology, National Clinical Research Center for Obstetric & Gynecologic Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [4]School of Nursing, The Hong Kong Polytechnic University, Hong Kong, Hong Kong SAR, China, [5]Department of Rheumatology and Immunology, the Second Xiangya Hospital of Central South University, Changsha, China, [6]Department of Rheumatology and Immunology, Clinical Medical Research Center for Systemic Autoimmune Diseases in Hunan Province, Changsha, China

Background: Anti-citrullinated peptide antibodies (ACPA)-negative (ACPA–) rheumatoid arthritis (RA) presents significant diagnostic and therapeutic challenges due to the absence of specific biomarkers, underscoring the need to elucidate its distinctive cellular and metabolic profiles for more targeted interventions.

Methods: Single-cell RNA sequencing data from peripheral blood mononuclear cells (PBMCs) and synovial tissues of patients with ACPA– and ACPA+ RA, as well as healthy controls, were analyzed. Immune cell populations were classified based on clustering and marker gene expression, with pseudotime trajectory analysis, weighted gene co-expression network analysis (WGCNA), and transcription factor network inference providing further insights. Cell-cell communication was explored using CellChat and MEBOCOST, while scFEA enabled metabolic flux estimation. A neural network model incorporating key genes was constructed to differentiate patients with ACPA– RA from healthy controls.

Results: Patients with ACPA– RA demonstrated a pronounced increase in classical monocytes in PBMCs and C1QChigh macrophages ($p < 0.001$ and $p < 0.05$). Synovial macrophages exhibited increased heterogeneity and were enriched in distinct metabolic pathways, including complement cascades and glutathione metabolism. The neural network model achieved reliable differentiation between patients with ACPA– RA and healthy controls (AUC = 0.81). CellChat analysis identified CD45 and CCL5 as key pathways facilitating macrophage-monocyte interactions in ACPA– RA, prominently involving iron-mediated metabolite communication. Metabolic flux analysis indicated elevated beta-alanine and glutathione metabolism in ACPA– RA macrophages.

Conclusion: These findings underscore that ACPA-negative rheumatoid arthritis is marked by elevated classical monocytes in circulation and metabolic

reprogramming of synovial macrophages, particularly in complement cascade and glutathione metabolism pathways. By integrating single-cell RNA sequencing with machine learning, this study established a neural network model that robustly differentiates patients with ACPA– RA from healthy controls, highlighting promising diagnostic biomarkers and therapeutic targets centered on immune cell metabolism.

KEYWORDS

rheumatoid arthritis, single-cell RNA sequencing, ACPA, synovial macrophage, beta-alanine and glutathione metabolism

Introduction

Rheumatoid arthritis (RA) is a chronic autoimmune disorder marked by persistent synovial inflammation, leading to joint destruction and impaired functionality. Its pathogenesis is driven by a multifaceted interaction of genetic, environmental, and immunological factors that promote immune dysregulation and chronic synovial inflammation (1, 2). A central feature of RA is the presence of autoantibodies, notably anti-citrullinated peptide antibodies (ACPA), which exhibit high specificity for the disease and serve as important diagnostic and prognostic markers (3, 4). Patients with ACPA-positive (ACPA+) RA typically experience a more aggressive disease course, characterized by accelerated joint damage and systemic involvement (5).

Nevertheless, approximately 20–30% of patients with RA are ACPA-negative (ACPA–), lacking these specific autoantibodies (6). ACPA– RA presents distinct clinical challenges, as it may follow unique disease trajectories and exhibit variable therapeutic responses compared to ACPA+ RA (7). The absence of ACPA complicates early diagnosis, potentially delaying treatment onset and impacting long-term patient outcomes (8). Furthermore, the immunopathological mechanisms underlying ACPA– RA remain incompletely characterized, posing a barrier to the development of targeted treatments for this subgroup (9).

Abbreviations: ACPA, RA: Anti-Citrullinated Protein Antibody-negative Rheumatoid Arthritis; ACPA+ RA, Anti-Citrullinated Protein Antibody-positive Rheumatoid Arthritis; PBMCs, Peripheral Blood Mononuclear Cells; scRNA-seq, Single-Cell RNA Sequencing; MIF, Macrophage Migration Inhibitory Factor; DMARDs, Disease-Modifying Anti-Rheumatic Drugs; TNF-α, Tumor Necrosis Factor Alpha; IL-6, Interleukin-6; JAK, Janus Kinase; STAT, Signal Transducer and Activator of Transcription; WGCNA, Weighted Gene Co-expression Network Analysis; HLA-DRA, Major Histocompatibility Complex, Class II, DR Alpha; CD74, Cluster of Differentiation 74; FCER1G, High Affinity Immunoglobulin E Receptor Subunit Gamma; MRC1, Mannose Receptor C-Type 1; AUC, Area Under the Curve; C9, Complement Component 9; GLUT1, Glucose Transporter 1; NAC, N-Acetylcysteine; PCR, Polymerase Chain Reaction; TGF-β, Transforming Growth Factor Beta.

Recent findings suggest that ACPA– RA represents a distinct clinical entity with unique immunological characteristics (10). Variations in genetic predisposition, cytokine profiles, and immune cell composition differentiate ACPA– RA from its ACPA+ counterpart (11, 12). Notably, alterations in monocyte and macrophage populations have been implicated in RA pathogenesis (13). Monocytes and macrophages are pivotal in inflammation and immune modulation, driving synovial hyperplasia and joint destruction through the release of pro-inflammatory cytokines and matrix-degrading enzymes (14). However, the precise roles of these immune cells in ACPA– RA remain inadequately elucidated.

Metabolic reprogramming in immune cells is increasingly recognized as a pivotal factor in autoimmune diseases, including RA (15). During immune activation, differentiation, and effector functions, immune cells reconfigure their metabolic pathways to meet heightened energetic and biosynthetic demands (16). Dysregulated metabolic processes can profoundly impact immune cell function, fostering chronic inflammation (17). In RA, research has demonstrated that altered glucose and lipid metabolism in both synovial fibroblasts and immune cells accelerates disease progression (15, 18). However, the metabolic characteristics of immune cells in ACPA– RA remain largely unexamined.

Advancements in single-cell RNA sequencing (scRNA-seq) now enable precise profiling of cellular heterogeneity, facilitating the identification of novel cell subtypes and disease-associated pathways (19). Utilizing scRNA-seq on peripheral blood mononuclear cells (PBMCs) and synovial tissue mononuclear cells (STMCs) from patients with RA allows researchers to delineate the complex cellular interactions and metabolic pathways underlying inflammation (20). Coupling scRNA-seq data with computational models further supports the estimation of metabolic fluxes and the construction of cell-cell communication networks (21).

This study investigates the cellular composition, metabolic reprogramming, and intercellular communication specific to ACPA– RA. scRNA-seq analysis was performed on PBMCs and STMCs from both patients with ACPA– RA and those with ACPA+ RA, with a focus on monocyte and macrophage subsets. Our

hypothesis posits that patients with ACPA– RA exhibit distinctive immune cell profiles and metabolic pathways that underlie their unique clinical features. By identifying differentially expressed genes, metabolic modules, and signaling pathways, this research aims to pinpoint potential biomarkers and therapeutic targets for ACPA– RA. Our findings offer new insights into ACPA– RA pathogenesis and underscore the critical role of metabolism in modulating immune responses within this patient subgroup.

Methods

Data acquisition

The sequence data used for this study have been deposited in the Genome Sequence Archive at the BIG Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under accession code HRA000155 (22). Researchers seeking access must submit an application for approval to utilize this dataset for further analysis.

Single-cell RNA sequencing alignment and quality control

Raw 10x Genomics sequencing data were processed with CellRanger v2.2.0 using the human transcriptome GRCh38-1.2.0 as a reference (23). Additional quality control measures were applied to remove low-quality cells, specifically excluding cells with mitochondrial gene expression exceeding 5%. Single-cell read counts from all samples were analyzed with the Seurat package (v5.0.1) in R (v4.3.1), where data were transformed into Seurat objects (24). Filtering criteria included retaining cells with unique molecular identifier (UMI) counts between 1000 and 25000 and genes detected in at least five cells while restricting cells to those expressing between 500 and 3500 genes. Post-filtering, data normalization was executed with Seurat's NormalizeData function, followed by the identification of highly variable genes using FindVariableFeatures.

Integration of scRNA-seq data from the same tissue

For tissue-specific scRNA-seq data integration (PBMC or synovial tissue), the Harmony package was employed. Downstream analyses, including dimensionality reduction and clustering, leveraged highly variable gene correlations.

Dimensionality reduction and major cell type annotation

Separate analyses were conducted for PBMC and synovial tissue datasets, with adjustments for confounders such as UMI counts, mitochondrial gene percentage, and cell cycle genes. Gene expression was scaled to unit variance, and dimensionality was reduced using principal component analysis (PCA), selecting the top 20 principal components (PCs) based on the elbow plot and variance explained. Cell clusters were visualized in two-dimensional space *via* Uniform Manifold Approximation and Projection (UMAP), and unsupervised clustering was executed with Seurat's FindClusters function, applying the Louvain algorithm for community detection. Resolution parameters were set to 0.5 for PBMC and 0.8 for synovial tissue.

Resolution settings were determined through an iterative approach, evaluating cluster stability and biological significance by varying resolution from 0.2 to 1.5 in 0.2 increments. Silhouette scores and modularity metrics were utilized to assess cluster cohesion and separation. The final resolutions provided an optimal balance, capturing distinct subpopulations without excessive clustering of biologically similar cells. Cell identities were assigned based on known marker genes for each cell type, as illustrated in Figure 1A and Supplementary Figure 1B, with validation through cross-referencing published datasets and established cell type annotations. For ambiguous marker expression, differential expression analysis was applied to confirm cell identity.

Differential expressed genes and pathway analysis

Differentially expressed gene (DEG) analysis was conducted using the FindMarkers function in Seurat with the Wilcoxon test. Bonferroni correction was applied to adjust p-values, and DEGs were filtered at a significance threshold of $p < 0.05$. For this study, the mini pct was set to 0.1, meaning at least 10% of cells in either group must express the gene for it to be included in the analysis. Enrichment analysis of DEGs was carried out using the clusterProfiler package (v3.12.0), examining Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (25). Specific parameters used in the analysis include a p-value cutoff of 0.05, a q-value cutoff of 0.2, and a gene set size range of 10 to 500. Multiple testing correction was performed using the Benjamini-Hochberg method. These parameter choices were guided by established practices to ensure biologically meaningful and statistically reliable results. To elucidate the functional roles of each macrophage subset, gene set variation analysis (GSVA) was performed with standard settings in the GSVA R package (v1.32.0).For this single-cell analysis, log-normalized expression data from Seurat were used as input. Pathways were selected from the MSigDB KEGG gene set collection, ensuring a comprehensive evaluation of biological processes. Specific parameters for the gsva() function included method = "gsva" (default kernel-based density estimation), mx.diff = TRUE (to calculate enrichment scores based on maximum difference between conditions), and a min.sz = 10 and max.sz = 500 to ensure only biologically relevant pathways were considered while accounting for sparsity in single-cell datasets. These parameter choices were optimized for single-cell data to maintain robustness and biological interpretability.

A

Sample Type	Source	RA Subtype	Sample Size	Number of Cells Analyzed
Peripheral Blood Mononuclear Cells (PBMCs)	Healthy Controls (HCs)	N/A	4	Part of 135,429
Peripheral Blood Mononuclear Cells (PBMCs)	RA Patients	ACPA+	10	Part of 135,429
Peripheral Blood Mononuclear Cells (PBMCs)	RA Patients	ACPA-	10	Part of 135,429
Synovial Tissue Mononuclear Cells (STMCs)	RA Patients	ACPA+	10	Part of 71,073
Synovial Tissue Mononuclear Cells (STMCs)	RA Patients	ACPA-	10	Part of 71,073

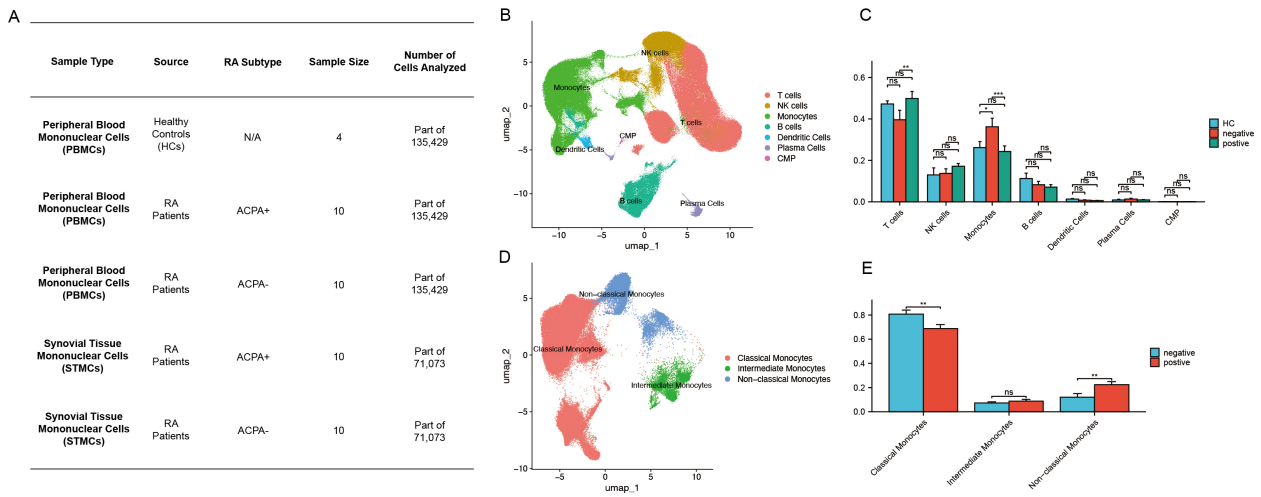


FIGURE 1
(A) Sample size distribution of scRNA-seq data from HC, ACPA+ RA, and ACPA– RA individuals. **(B)** UMAP clustering of immune cells from PBMCs. **(C)** Wilcoxon test comparing immune cell proportions between ACPA+ and ACPA– groups. **(D)** UMAP clustering of monocyte subpopulations. **(E)** Wilcoxon test comparing monocyte subpopulations between ACPA+ and ACPA– groups. (P-values are expressed as follows: * $p \leq 0.05$, ** $p \leq 0.01$, and *** $p \leq 0.001$.).

Additionally, AUCell analysis was performed to evaluate the activity of gene sets in individual cells, complementing GSVA results by providing cell-level resolution. This analysis used the AUCell package (v1.12.0) with the AUCell_buildRankings() function to rank genes based on expression levels across individual cells. The AUCell_calcAUC() function was then applied to calculate the Area Under the Curve (AUC) scores for predefined gene sets, with thresholds determined based on empirical distributions. Parameters included a ranking threshold of 5% and the use of log-normalized data to ensure compatibility with single-cell datasets. These details enhance the transparency and reproducibility of the methods used in this study.

Trajectory inference

The Monocle2 algorithm was applied to explore differentiation trajectories within selected clusters (26). Cells of interest were subset using the Seurat subset function, and a CellDataSet object was generated with Monocle2's newCellDataSet function, setting the lowerDetectionLimit parameter to 0.5. Low-quality cells and genes were removed with min_expr = 0.1, and dimensionality reduction was conducted using the DDRTree method. Visualization of trajectories was achieved through plot_cell_trajectory and plot_genes_in_pseudotime functions.

SCENIC analysis

To identify regulons (transcription factors [TFs], their modules, and potential targets) and assess their activity, this study employed the single-cell regulatory network inference and clustering (SCENIC) approach (27). This workflow began with the inference of co-expression modules using GRNBoost2, followed by motif pruning with cisTarget. Regulon activity was quantified with AUCell scores, and TF activity was evaluated using the Python-based tool pySCENIC (28). Leveraging the cis-target and motif databases, all TFs with motifs were analyzed to identify cell-type-specific regulons with high regulon specificity scores (RSS) (29).

HdWGCNA analysis

High-dimensional weighted gene co-expression network analysis (hdWGCNA) was employed to identify key macrophage-related genes (30). Monocyte and macrophage populations were extracted from scRNA-seq data, gene expression correlation matrices were computed, and gene co-expression modules were identified. Critical parameters were carefully optimized during the hdWGCNA process to ensure robust network construction and module detection. In the initial step, gene expression correlation matrices were calculated, and the soft-thresholding power was determined to optimize scale-free network topology. A soft-thresholding power of 7 was selected to ensure that the network exhibited scale-free properties, a hallmark of biological networks. This selection was guided by plotting the scale-free topology model fit against various power values and choosing the point where the network's R-squared value reached a plateau. Following network construction, co-expression modules—clusters of genes with similar expression patterns across the macrophage population—were identified. The relevance of these modules was assessed *via* module-trait relationship analysis, correlating each module with specific traits related to macrophage activation and inflammation. For each trait-related module, hub genes—genes with high intramodular connectivity central to the network structure—were identified. Hub genes were defined based on their connectivity scores (kME values) within their respective modules, following the approach outlined in previous studies.

Deep neural network construction

A deep neural network (DNN) was constructed using PyTorch to define and optimize the network architecture. Based on PBMC expression data from 20 DEGs and corresponding cell-type data, the DNN was developed to distinguish patients with ACPA– RA from healthy controls (HC). Data were divided into a 70% training set and a 30% test set, with training performed over 1000 epochs using mini-batch gradient descent. The DNN architecture consisted of an input layer with 21 features, followed by four hidden layers containing 128, 64, 32, and 16 neurons, each employing Sigmoid activation functions, and concluded with a single Sigmoid neuron in the output layer for binary classification (ACPA-negative or healthy).In clinical settings, the characterization of macrophage populations plays a crucial role in diagnosing and understanding rheumatoid arthritis (RA) subtypes. If over 50% of a patient's macrophages are found to be ACPA-negative, this could strongly suggest an ACPA-negative RA diagnosis. Otherwise, the patient is likely classified as healthy.

To prevent overfitting, early stopping was applied based on validation loss, and each hidden layer included a dropout rate of 0.2. Key model parameters, including learning rate, number of layers, and dropout rates, were optimized *via* grid search, exploring learning rates from 0.001 to 0.01. A learning rate of 0.005 was ultimately selected based on improved validation accuracy. Model performance was assessed through accuracy metrics and ROC curve analysis, with the ROC curve generated using Scikit-learn's roc_curve function. Additionally, cross-validation was implemented to reinforce model robustness, averaging performance metrics across five folds to ensure generalizability.

Cell communication and signaling pathways

Cell communication analysis was performed using the CellChat package in R with default parameters, focusing on PBMC monocyte and synovial macrophage subsets independently (31). The analysis utilized the human CellChatDB and enabled a comparative assessment of interactions between ACPA+ and ACPA– macrophage subpopulations and PBMC monocytes.

MEBOCOST analysis

MEBOCOST, a Python-based tool, inferred metabolite-mediated cell communication from scRNA-seq data. This tool, which leverages a curated database of metabolite sensors and partners, identified sender and receiver cells based on metabolite outflow/inflow rates and enzyme/sensor expression levels. scRNA-seq expression data were first loaded into a Python pandas DataFrame, integrated with cell annotations, and then used to infer metabolic communications. Results were visualized to illustrate communication events, sender-to-receiver flows, and sensor expression levels.

Construction of single-cell metabolic flux curves

Single-cell metabolic flux profiles were derived using the single-cell flux estimation analysis (scFEA) algorithm, a graph neural network-based approach (21). The algorithm utilized 168 metabolic modules, obtained from scFEA's official GitHub repository (https://github.com/changwn/scFEA). KEGG enrichment analysis was conducted on input and output modules using MetaboAnalyst (https://www.metaboanalyst.ca/home.xhtml).

Statistical analysis

All statistical analyses were performed using R software (v4.3.1), with visualizations generated in R Studio. Data were pre-processed to meet the assumptions for each statistical test, and appropriate transformations were applied when necessary. Statistical tests were selected based on data distribution and study design. For comparisons between two groups with normally distributed data and equal variances, Student's t-test was used. The Wilcoxon Rank-Sum Test was applied for non-parametric data, providing a robust method for comparing medians between two independent groups, especially suitable for small sample sizes or skewed distributions. The Kruskal-Wallis Test was employed for comparisons across more than two independent groups with non-parametric data. To control the family-wise error rate, p-values were adjusted using the Holm-Bonferroni method. Statistical significance was set as follows: "ns" for $p > 0.05$, * for $p \leq 0.05$, ** for $p \leq 0.01$, *** for $p \leq 0.001$, and **** for $p \leq 0.0001$.

Results

Identification of distinct immune cell types in patients with RA

Single-cell sequencing data of immune cells from patients with ACPA– RA and patients with ACPA+ RA were obtained from the Genome Sequence Archive at the Big Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences. The dataset comprised 44 samples, including CD45+ PBMCs isolated from HC (n = 4) and from ACPA+ (n = 10) and ACPA– (n = 10) RA individuals (Figure 1A). Additionally, synovial tissue mononuclear cells (STMCs) were obtained from ACPA+ (n = 10) and ACPA– (n = 10) RA individuals (Figure 1A). None of the patients were receiving disease-modifying antirheumatic drugs (DMARDs), corticosteroids, or targeted therapies at the time of sampling, though some opted for physical therapies, such as thermotherapy or acupuncture, to manage pain. A graph-based unsupervised clustering method was applied to identify cell types by examining typical marker genes. Cell populations identified included T cells, B cells, monocytes, dendritic cells, plasma cells, NK cells, and common myeloid progenitors (CMP). Each cell type was annotated according to well-characterized marker genes (Figure 1B, Supplementary Figure 1A). Specifically, T cells were defined by high expression of CD3D and CD3E, while NK cells were distinguished by

NKG7 and GNLY. Monocytes were annotated by CD14 and FCGR3A, and dendritic cells by ITGAX and HLA-DQA1. B cells and plasma cells were characterized by distinct marker profiles, with B cells expressing CD19, MS4A1, and CD79A, and plasma cells marked by SDC1 and MZB1. CMPs were identified using CD34, KIT, and FLT3, established indicators of progenitor populations. Marker selection was based on specificity for each cell type, validated by previous research in the field. This rigorous marker selection and clustering approach enabled robust and precise classification of cell types within the dataset.

Increased monocyte proportions in patients with ACPA– RA

The Wilcoxon test was applied to assess differences in immune cell type proportions across ACPA-positive, ACPA-negative, and HC groups. Results indicated a statistically significant increase in monocyte proportions within the ACPA-negative group compared to both ACPA-positive and HC groups ($p < 0.001$). Additionally, a significant difference was detected in T cell proportions between ACPA-negative and ACPA-positive groups ($p < 0.01$). No significant differences were observed for NK cells, B cells, dendritic cells, plasma cells, or CMPs across the groups (Figure 1C).

Identification of monocyte subpopulations

Further dimensionality reduction and clustering analysis of monocytes identified three distinct subpopulations: classical, non-classical, and intermediate monocytes (Figure 1D). Classical monocytes were characterized by CD14 expression, non-classical monocytes by CD16, and intermediate monocytes by the co-expression of CD14 and CD16 (Supplementary Figure 1B).

Patients with ACPA+ RA show increased classical monocytes and reduced non-classical monocytes

The Wilcoxon test was subsequently conducted to compare the proportions of monocyte subpopulations between ACPA-positive and ACPA-negative groups. This analysis revealed a statistically significant reduction in the proportion of non-classical monocytes in the ACPA-positive group relative to the ACPA-negative group ($p < 0.01$) (Figure 1E). Conversely, the ACPA-positive group exhibited a significant increase in classical monocyte proportions ($p < 0.01$) (Figure 1E). No significant difference was identified in intermediate monocyte proportions between the two groups (Figure 1E).

Macrophages and fibroblasts are increased in ACPA-positive synovial tissue

Recognizing synovial inflammation as a hallmark of RA, dimensionality reduction and clustering analysis were performed on scRNA-seq data from synovial cells. This approach identified eight distinct cell populations within synovial tissue: T cells, plasma cells, NK cells, B cells, macrophages and myeloid cells, endothelial cells, mast cells, and fibroblasts (Figure 2A, Supplementary Figure 1C). Each cell type was annotated based on classical markers, selected for their established involvement in RA-related inflammation and immune response. Specifically, T cells were characterized by CD3D and CD3E expression, B cells by CD19 and CD79A, and plasma cells by markers such as MZB1 and IGLC2. Macrophages and myeloid cells showed high CD68 and LYZ levels, while endothelial cells were identified by PECAM1 and VWF. Mast cells were marked by TPSAB1, and fibroblasts by ACTA2 and DCN expression. These cell types are well-documented contributors to the inflammatory cascade and tissue damage observed in RA, providing insights into the cellular landscape of synovial inflammation. Using the Wilcoxon test, immune cell proportions were compared between patients with ACPA+ RA and those with ACPA– RA. Results demonstrated a statistically significant increase in macrophages and myeloid cells ($p < 0.01$) and fibroblasts ($p < 0.05$) in the ACPA+ group (Figure 2E), aligning with the roles of macrophages and fibroblasts in sustaining inflammation and facilitating joint destruction in RA. No significant differences were found for T cells, plasma cells, NK cells, B cells, endothelial cells, or mast cells (Figure 2E).

Identification of 11 distinct macrophage subtypes with differential roles in RA

Focusing on macrophages, which are central to chronic inflammation, tissue destruction, and immune dysregulation in RA, further dimensionality reduction and clustering analysis identified 11 distinct macrophage subtypes based on gene expression profiles: C1: C1QChigh(T), C2:NAMPT/NFKBIAhigh(H), C3:FN1high(H), C4: LYZhigh(H), C5:CD163high(T), C6:HLAhigh(H), C7:APOEhigh(T), C8:MARCOhigh(H), C9:MIFhigh(H), C10:BIRC3high(T), and C11: C1QBhigh(H) (Figure 2B, Supplementary Figure 1D). These subtypes reflect macrophage populations with diverse roles in RA. Here, T denotes tissue-resident macrophages, which sustain local inflammation in synovial tissue, while H represents hematogenous macrophages, recruited from the bloodstream in response to inflammatory signals. Tissue-resident macrophages were identified by CXCR6, ITGAE, and CD69 markers, while hematogenous macrophages were marked by S1PR1, KLF2, and CCR7, following marker definitions from prior studies (Figure 2C) (32–34).

Specific macrophage subtypes are enriched in ACPA-negative and ACPA-positive RA

KEGG enrichment analysis on differentially expressed genes across 11 macrophage clusters revealed that genes downregulated in ACPA-positive samples (i.e., upregulated in ACPA-negative samples) were enriched in RA-related subgroups, particularly clusters C1 and C7. These genes were associated with immune pathways such as Th17, Th1, and Th2 cell differentiation. Conversely, ACPA-positive samples showed lower counts and

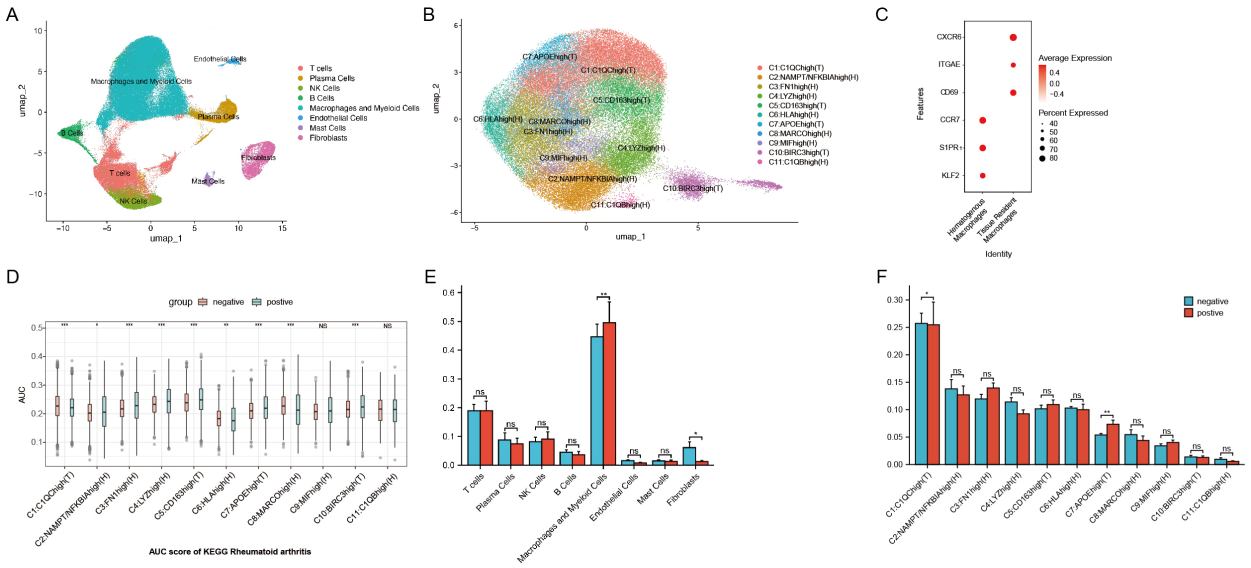


FIGURE 2
(A) UMAP plot of scRNA-seq data from synovial cells. **(B)** UMAP plot of macrophage subtypes from synovial cells. **(C)** Markers for hematogenous macrophages and tissue-resident macrophages. **(D)** AUCell activity scoring of the 11 macrophage subtypes in the KEGG Rheumatoid Arthritis pathway. **(E)** Wilcoxon test comparing synovial cell populations between ACPA+ and ACPA− groups. **(F)** Wilcoxon test comparing macrophage subpopulations from synovial cells between ACPA+ and ACPA− groups. (P-values are expressed as follows: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, and NS indicates no significance.).

higher p-values in the upregulated differentially expressed genes for C1 and C7, indicating less enrichment compared to ACPA-negative samples (Supplementary Figures 2A, B).

GSVA identified several key pathways in C1 and C7, including complement and coagulation cascades, allograft rejection, alcoholic liver disease, phagosome, antigen processing and presentation, cholesterol metabolism, pertussis, lysosome, and Staphylococcus aureus infection (Figure 3F). Further Wilcoxon test analysis indicated a significant increase in the proportion of C1 macrophages in the ACPA-negative group ($p < 0.05$), suggesting these cells contribute to local inflammation and synovial hyperplasia in ACPA-negative RA. In contrast, a significant decrease in C7 proportions was observed in the ACPA-negative group ($p < 0.01$), suggesting that C7 macrophages may have a regulatory or protective function that is diminished in ACPA-positive RA (Figure 2F).

AUCell activity scoring for rheumatoid arthritis pathways in the KEGG database across the 11 macrophage subtypes revealed distinct activity patterns between ACPA-positive and ACPA-negative groups. Specifically, C1, C2, C5, C6, C7, and C10 exhibited significantly higher activity scores in the ACPA-positive group, whereas C1 and C8 had notably higher activity in the ACPA-negative group. These results suggest that C1 and C8 may play pivotal roles in ACPA-negative RA, while other subtypes are more active in ACPA-positive RA (Figure 2D).

ACPA-negative RA macrophages display more complex developmental trajectories

To elucidate the dynamic roles of macrophage subtypes in RA progression and immune responses, pseudotime trajectory analysis was conducted on 11 macrophage subtypes to investigate their developmental paths (Figures 3A–D). Separate analyses were performed for macrophages from ACPA− RA and ACPA+ RA individuals. Results indicated that macrophages in ACPA− RA exhibited a more intricate developmental trajectory, forming four distinct branches (Figure 3A), whereas ACPA+ RA macrophages formed only two branches (Figure 3C). This suggests greater diversity in developmental and activation processes among macrophages in ACPA− RA, possibly reflecting increased heterogeneity in macrophage function compared to ACPA+ RA.

Within ACPA− RA, certain macrophage subtypes displayed distinct patterns along the developmental path. Subtypes C10(T) and C11(H), for instance, appeared primarily in early developmental stages, indicating a role in initial macrophage activation or differentiation. In contrast, C1(T) and C7(T) spanned both early and late stages but were absent from intermediate stages, suggesting that these subtypes may have specialized roles at the onset and resolution phases of the macrophage lifecycle, potentially involved in initiating and resolving inflammation. In ACPA+ RA, macrophage subtypes were more uniformly distributed along the trajectory, indicating less developmental complexity, which may reflect a more sustained and homogeneous inflammatory response in ACPA+ RA. The increased developmental complexity and unique pathway involvement in ACPA− RA highlight a higher degree of macrophage heterogeneity, which could contribute to the variable clinical presentation and disease progression observed in ACPA − RA.

To elucidate the biological relevance of pseudotime-related changes, KEGG enrichment analysis was performed on genes associated with pseudotime trajectories for ACPA+ RA and

10.3389/fimmu.2024.1512483

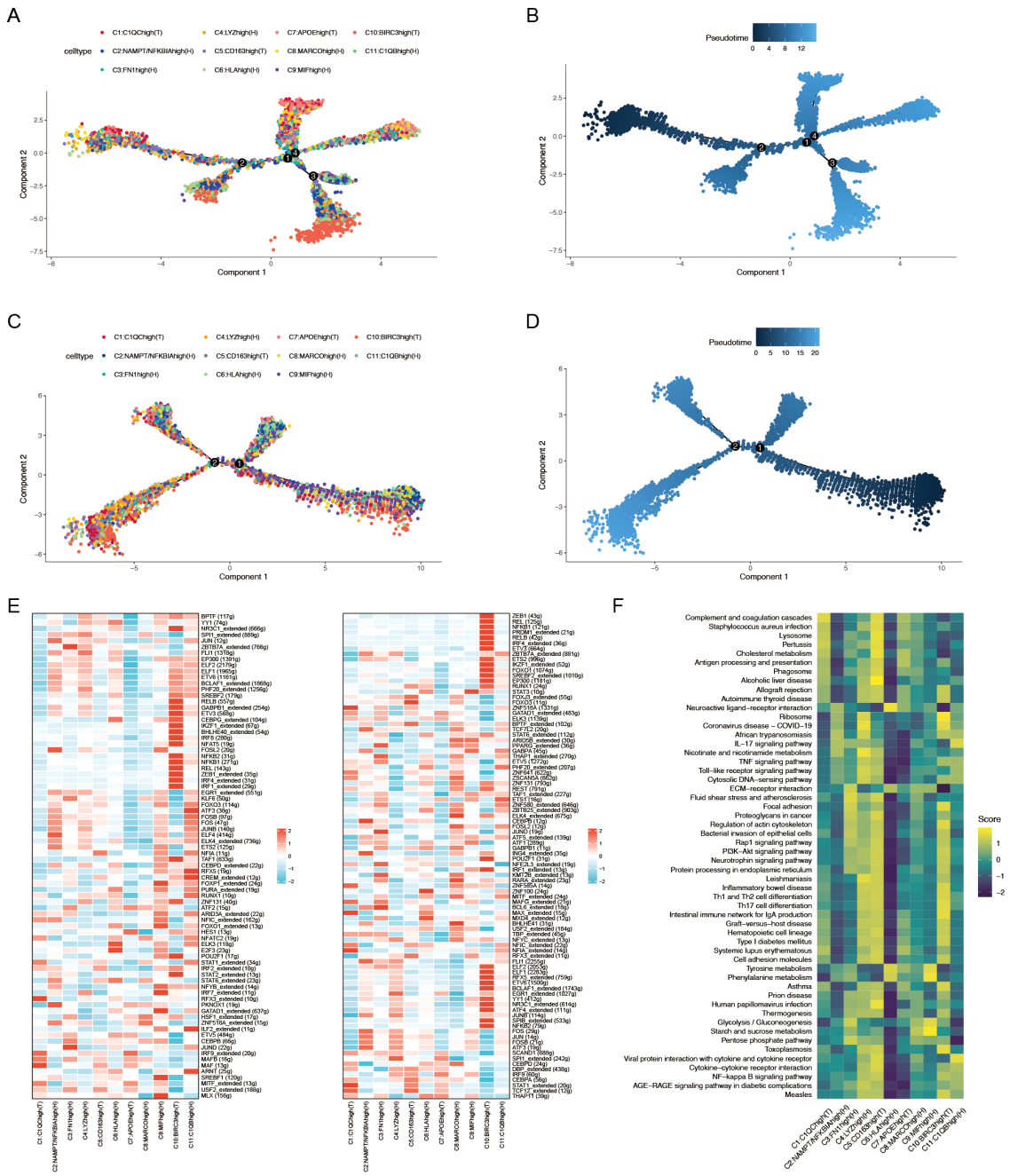


FIGURE 3
(A) Pseudotime trajectory analysis of macrophage subtypes in patients with ACPA– RA. **(B)** Pseudotime progression of macrophage subtypes in patients with ACPA– RA. **(C)** Pseudotime trajectory analysis of macrophage subtypes in patients with ACPA+ RA. **(D)** Pseudotime progression of macrophage subtypes in patients with ACPA+ RA. **(E)** Heatmap of transcription factor activity in ACPA– (left panel) and ACPA+ (right panel) RA macrophages. **(F)** GSVA analysis of macrophage subpopulations.

ACPA– RA macrophages. Both groups shared 141 pathways, including key inflammatory and RA-related pathways such as Rheumatoid arthritis and Osteoclast differentiation, which are fundamental to RA pathology (Supplementary Table 1).

Distinctly, ACPA– RA macrophages were enriched in seven pathways, including Complement and coagulation cascades, Antifolate resistance, and Glycosphingolipid biosynthesis – ganglion series. These pathways suggest specific roles in the development and progression of ACPA– RA (Supplementary Table 1). Enrichment in the Complement and coagulation cascades pathway implies a role in heightened inflammation and immune activation, potentially exacerbating joint damage. Antifolate resistance indicates an altered response to treatments such as methotrexate, suggesting the potential need for alternative therapeutic strategies in patients with ACPA– RA. Furthermore, enrichment in Glycosphingolipid biosynthesis suggests unique lipid metabolism influencing macrophage activity and immune regulation, further distinguishing ACPA– RA from ACPA+ RA. These pathways

underscore critical biological differences that may impact both treatment response and disease progression in ACPA– RA.

Conversely, ACPA+ RA macrophages were enriched in 22 unique pathways, including key signaling pathways such as Sphingolipid signaling pathway, JAK-STAT signaling pathway, mTOR signaling pathway, and Adipocytokine signaling pathway (Supplementary Table 1). These pathways are pivotal in immune regulation and inflammation, with their enrichment in ACPA+ RA macrophages pointing to distinct molecular mechanisms underlying the more aggressive disease phenotype commonly observed in patients with ACPA+ RA.

More extensive transcription factor networks in ACPA-positive RA macrophages

To further elucidate the gene regulatory mechanisms underlying these differences, SCENIC analysis was conducted to infer TF regulatory networks. This analysis identified 80 active TFs regulating macrophage subtypes in ACPA– RA and 90 active TFs in ACPA+ RA (Figure 3E). Notably, 43 TFs were shared between the two groups, indicating common regulatory mechanisms in macrophage activation across both ACPA+ and ACPA– RA (Figure 3E). However, the number of genes regulated by these shared TFs was greater in ACPA+ RA, suggesting a more extensive and complex gene regulatory network in this group. This expanded network in ACPA+ RA likely reflects a more robust and uniform activation of regulatory pathways, consistent with the severe and sustained inflammatory phenotype frequently observed in patients with ACPA+ RA.

Gene modules associated with ACPA-negative RA identified by hdWGCNA

To investigate the molecular mechanisms of macrophage subtypes associated with ACPA-negative (ACPA–) RA, high-dimensional weighted gene co-expression network analysis (hdWGCNA) was employed. While traditional WGCNA and other dynamic network analysis tools are effective for bulk RNA-seq data, hdWGCNA provides distinct advantages for high-dimensional single-cell RNA-seq, being optimized to address unique challenges such as data sparsity, high noise levels, and the need for granularity in capturing cell-type-specific networks. Unlike standard WGCNA, hdWGCNA preserves cellular-level data structure, making it well-suited to the complex heterogeneity present in RA macrophage populations.

The hdWGCNA approach enabled the identification of modules of highly co-expressed genes, offering biological insights through enrichment analysis and integration with known pathways. An optimal soft threshold of 7 was chosen to ensure a scale-free network topology, facilitating robust co-expression analysis. Using this threshold, a co-expression network was constructed, identifying seven distinct gene co-expression modules, each representing a unique set of interconnected genes with potential regulatory roles in macrophage function.

Correlation analysis between these modules and ACPA+/- RA showed that the brown, red, and black modules were associated with ACPA–, while the yellow, turquoise, and blue modules were linked to ACPA+ (Figure 4A, Supplementary Table 2). The brown module, in particular, exhibited high expression in macrophage subtypes C1:C1QChigh(T), C5:CD163high(T), and C7:APOEhigh (T) (Figure 4B). Enrichment analysis on the brown, red, and black modules revealed that the brown module was enriched in critical immune-related pathways, such as MHC class II-related pathways, Rheumatoid arthritis, Complement and coagulation cascades, Antigen processing and presentation, and Th1 and Th2 cell differentiation (Figure 4C).

By intersecting the 426 genes in the brown module with differentially expressed genes in PBMCs and macrophage subtypes, 20 intersecting genes were identified (Figure 4D), indicating their differential expression in both PBMCs and macrophage subtypes. Given that anti-cyclic citrullinated peptide (anti-CCP) antibodies serve as an important diagnostic marker for RA but are absent in patients with ACPA– RA, complicating diagnosis relative to ACPA+ RA, these 20 differentially expressed genes were leveraged to construct a neural network model.

Neural network model distinguishes ACPA-negative RA from healthy controls

Using PBMC expression data and cell-type annotations, a deep neural network was constructed to distinguish patients with ACPA– RA from healthy controls. The data was split into a 70% training set and a 30% test set, with the model trained over 1000 epochs using mini-batch gradient descent (Figure 4E). To ensure robustness and prevent overfitting, the ROC curve was evaluated for both training and test sets, achieving an AUC of 0.92 on the training set and 0.81 on the test set. To further validate the robustness and generalizability, five-fold cross-validation was applied, with the average AUC across folds reaching 0.87 and individual AUCs ranging from 0.84 to 0.89. These results indicate stable model performance, supporting the potential clinical application of single-cell transcriptomics for RA diagnostics (Figure 4F).

ACPA-positive RA exhibits stronger macrophage-monocyte communication

Examining PBMC monocyte and macrophage subtype interactions in RA is essential to understanding systemic immune responses that contribute to local joint inflammation and tissue damage. These interactions highlight mechanisms driving chronic inflammation, reveal biomarkers for disease progression, and identify therapeutic targets by isolating specific pathways involved in monocyte-to-macrophage differentiation.

CellChat was utilized to analyze cell communication between monocytes and macrophages in ACPA+ and ACPA– RA. ACPA+

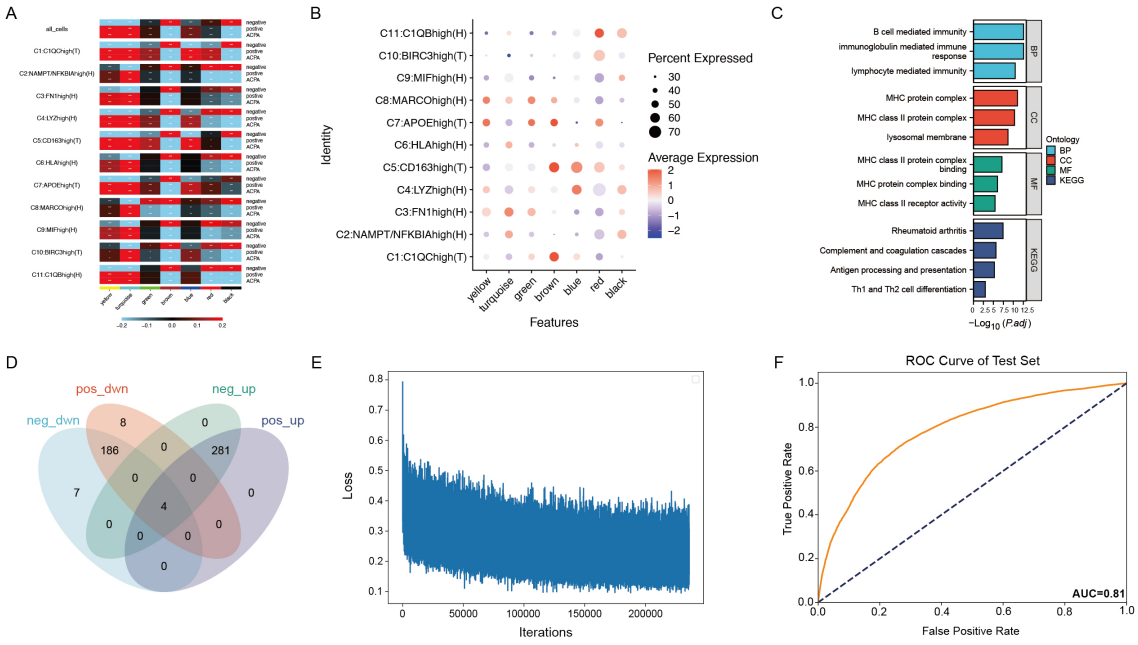


FIGURE 4
(A) Correlation heatmap of seven gene co-expression modules identified by WGCNA in macrophage subtypes. **(B)** Dot plot showing gene expression within the brown module across macrophage subtypes. **(C)** Enrichment analysis of genes in the brown module. **(D)** Venn diagram showing the overlap of 426 genes from the brown module with differentially expressed genes in PBMCs and macrophage subtypes. **(E)** Training loss curve for the deep neural network model distinguishing patients with ACPA– RA from healthy controls. **(F)** ROC curve displaying the neural network model's performance on the test set.

RA showed 2,199 inferred interactions, higher than the 1,789 interactions observed in ACPA– RA. Interaction strength was also significantly higher in ACPA+ RA (0.467) compared to ACPA– RA (0.196) (Figure 5A), suggesting that macrophage-monocyte communication in ACPA+ RA is more intense, potentially contributing to the aggressive inflammatory response and severe clinical presentation typically seen in ACPA+ RA.

Regarding relative information flow, CD45 and CCL5 emerged as primary pathways mediating macrophage-monocyte communication in ACPA– RA (Figure 5B, Supplementary Tables 3, Supplementary Table 4). The CD45 pathway was particularly critical for cross-organ communication between classical monocytes and C1:C1QChigh(T) macrophages (Figure 5C), indicating its role in macrophage activation and recruitment to inflamed tissues in ACPA– RA. In contrast, fewer interactions were observed between classical monocytes and C1:C1QChigh(T) macrophages in ACPA+ RA, suggesting alternative pathways may drive immune responses in ACPA+ RA.

To further explore, the Wilcoxon test was employed to compare gene expression levels of CD45 pathway mediators, specifically PTPRC and MRC1, between the two groups. PTPRC expression was significantly higher in C1(T) and C9:MIFhigh(H) macrophages in ACPA+ RA ($P < 0.001$ and $P < 0.05$, respectively) (Figure 5D). This elevated expression of PTPRC, a key component of the CD45 pathway, suggests sustained macrophage activation in ACPA+ RA. Additionally, MRC1 showed significantly higher expression in C1(T) macrophages ($P < 0.001$) (Figure 5E), implying a role in modulating immune responses through alternative pathways in this macrophage subtype.

Iron-mediated cell communication is prominent in ACPA-negative RA

Previous CellChat analysis suggested that macrophage-monocyte communication in PBMCs might be mediated by the CD45 pathway. Given the importance of cell metabolism in RA pathogenesis—particularly in shaping immune cell functions and inflammatory responses—the role of metabolite-mediated interactions between macrophages and monocytes was considered. Metabolic factors such as lipids and iron play significant roles in RA by influencing cellular energy balance, signaling, and differentiation, thereby driving inflammation and disease progression.

To explore this further, MEBOCOST, a Python-based computational tool for inferring metabolite-mediated cell-cell communication from single-cell RNA sequencing data, was employed. Analysis showed a higher number of metabolite-mediated communication events in ACPA+ RA (Figure 6B) compared to ACPA– RA (Figure 6A). However, focusing on the communication flow from sender metabolite to sensor in the receiver, key interactions were identified in ACPA– RA between monocytes and C1(T) macrophages, as predicted by CellChat. Specifically, the metabolic communication pathways included classical monocytes (sender) – Iron (metabolite) – TFRC (sensor) – C1:C1QChigh(T) macrophages (receiver) and classical monocytes (sender) – Iron (metabolite) – SLC40A1 (sensor) – C1:C1QChigh(T) macrophages (receiver). Additionally, non-classical monocytes displayed similar

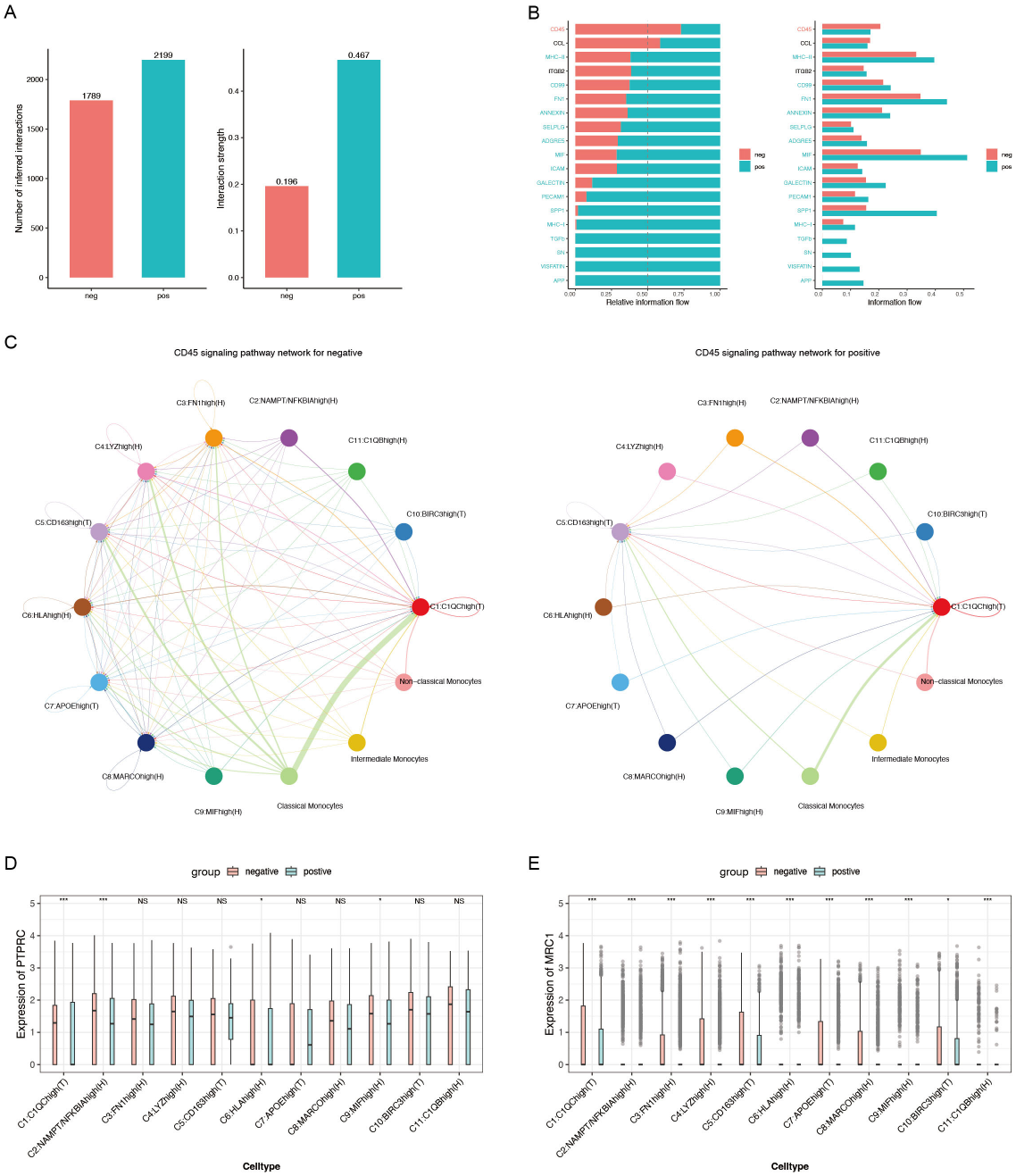


FIGURE 5
(A) Bar chart depicting the number of inferred interactions and interaction strengths. **(B)** Relative information flow of key signaling pathways mediating macrophage-monocyte communication. **(C)** CD45 signaling pathway network for ACPA– (left) and ACPA+ (right) RA. **(D)** Box plot of PTPRC expression, a critical component of the CD45 signaling pathway. **(E)** Box plot of MRC1 expression, another essential component of the CD45 signaling pathway. (P-values are expressed as follows: * $p \leq 0.05$, *** $p \leq 0.001$, and NS indicates no significance.).

iron-mediated communication pathways in ACPA– RA (Figure 6C, Supplementary Tables 5, Supplementary Table 6).

In contrast, metabolite-mediated communication between monocytes and macrophages was less prominent in ACPA+ RA (Figure 6D), consistent with CellChat findings, indicating that macrophage-monocyte communication may not be as central in ACPA+ RA. This suggests that iron-mediated interactions may be more critical in ACPA– RA, while alternative communication mechanisms could be more relevant in ACPA+ RA.

Visualizing the mean abundance of communication-mediating metabolites revealed higher levels of iron and L-glutamine in ACPA – RA, indicating their roles in macrophage-monocyte interactions (Figure 6E). Conversely, cholesterol abundance was higher in ACPA+ RA, pointing to a shift towards lipid-related metabolic pathways in this group (Figure 6F). These results underscore distinct metabolic profiles in ACPA– and ACPA+ RA, with iron and glutamine as key mediators in ACPA– RA, while cholesterol may be more influential in the immune response of ACPA+ RA.

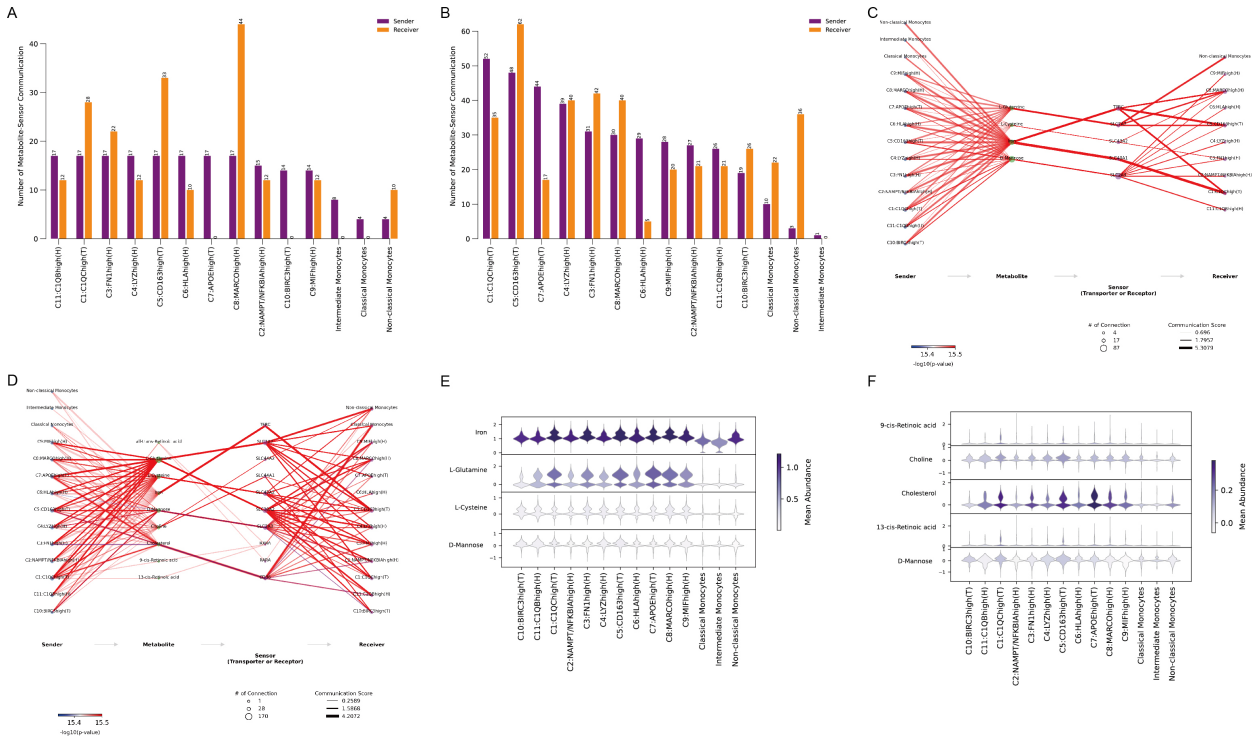


FIGURE 6
(A) Number of metabolite-sensor communication events in ACPA– RA. **(B)** Number of metabolite-sensor communication events in ACPA+ RA **(C)** Metabolite-mediated communication pathways in ACPA– RA. **(D)** Metabolite-mediated communication pathways in ACPA+ RA. **(E)** Violin plots showing the mean abundance of metabolites in ACPA– RA. **(F)** Violin plots showing the mean abundance of metabolites in ACPA+ RA.

Metabolic flux analysis reveals upregulated pathways in ACPA-negative RA macrophages

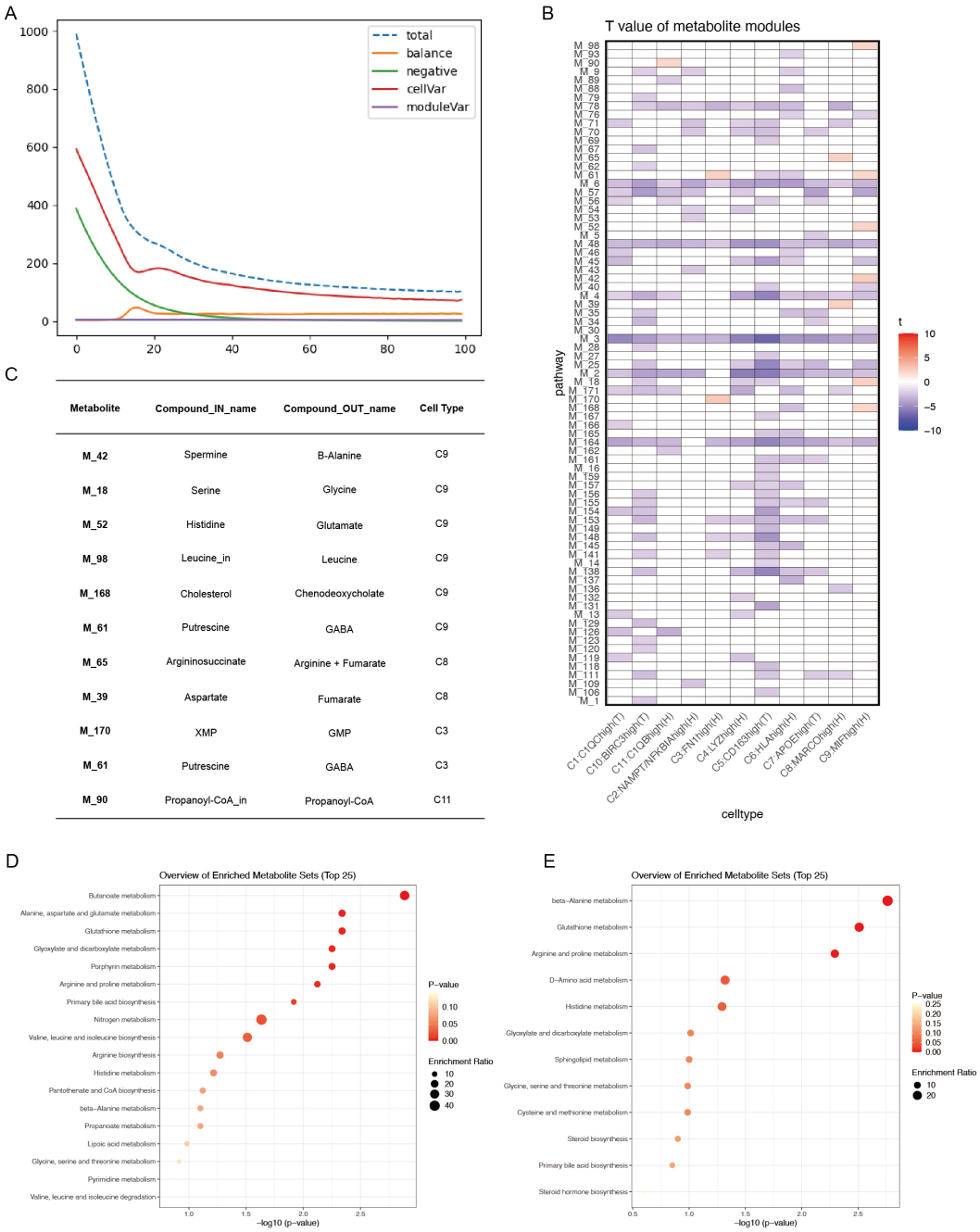
To further explore metabolic flux in macrophage subtypes within ACPA– RA, scFEA—a graph neural network model tailored for estimating cell metabolism using scRNA-seq data—was employed. scFEA leverages a reconstructed human metabolic map, utilizing a probabilistic model with flux balance constraints and an optimization solver within a graph neural network to capture the intricate relationships from transcriptomics to metabolomics. This model reflects the non-linear dependencies between enzyme gene expression and reaction rates, using gene expression profiles of macrophage subtypes as input data.

Figure 7A shows model convergence through the loss function, confirming its accuracy. Given the continuous and normally distributed output data, the limma package was used to compare ACPA– and ACPA+ RA samples, considering $p < 0.05$ as statistically significant, with $t > 0$ indicating upregulation in ACPA– RA and $t < 0$ indicating downregulation. This analysis identified 11 metabolic modules upregulated in ACPA– RA (Figures 7B, C), with each module corresponding to in and out metabolites. Notably, the C9: MIFhigh(H) macrophage subtype exhibited a substantial number of upregulated metabolites. KEGG enrichment analysis of these metabolites was performed using MetaboAnalyst. The input metabolites were predominantly enriched in pathways such as beta-alanine metabolism, Glutathione metabolism, Arginine and proline metabolism, D-amino acid metabolism, and Histidine metabolism (Figure 7E). The output metabolites were enriched in pathways including Butanoate metabolism, Alanine, aspartate and glutamate metabolism, Glutathione metabolism, Glyoxylate and dicarboxylate metabolism, Porphyrin metabolism, Arginine and proline metabolism, Primary bile acid biosynthesis, Nitrogen metabolism, and Valine, leucine and isoleucine biosynthesis (Figure 7D).

These results suggest that macrophage subtypes, particularly C9:MIFhigh(H), may significantly contribute to RA pathogenesis by promoting key metabolic processes, highlighting distinct metabolic pathways active in ACPA– RA.

Discussion

This study offers a detailed analysis of the cellular and molecular distinctions of ACPA– RA, particularly focusing on metabolic alterations. scRNA-seq identified unique immune cell compositions, metabolic pathways, and intercellular communication patterns that set ACPA– RA apart from ACPA+ RA. Notably, a marked increase in monocytes, especially classical monocytes, was observed in the PBMCs of patients with ACPA– RA patients compared to patients with ACPA+ RA and healthy controls. This elevation suggests a pivotal role for monocytes in the systemic inflammation that characterizes ACPA– RA. Classical monocytes, known for their potent pro-inflammatory cytokine production and their capacity to differentiate into macrophages



Metabolite	Compound_IN_name	Compound_OUT_name	Cell Type
M_42	Spermine	B-Alanine	C9
M_18	Serine	Glycine	C9
M_52	Histidine	Glutamate	C9
M_98	Leucine_in	Leucine	C9
M_168	Cholesterol	Chenodeoxycholate	C9
M_61	Putrescine	GABA	C9
M_65	Argininosuccinate	Arginine + Fumarate	C8
M_39	Aspartate	Fumarate	C8
M_170	XMP	GMP	C3
M_61	Putrescine	GABA	C3
M_90	Propanoyl-CoA_in	Propanoyl-CoA	C11

FIGURE 7
(A) Convergence of the loss function during scFEA analysis. **(B)** T-values of metabolite modules across macrophage subtypes, visualized with color codes: red for t-values > 0 (indicating upregulation in the ACPA– group) and blue for t-values < 0 (indicating downregulation in the ACPA– group or upregulation in the ACPA+ group). **(C)** Summary table of the top in-and-out metabolites for significant metabolic modules in ACPA– RA. **(D)** KEGG enrichment analysis of output metabolites from macrophage subtypes. **(E)** KEGG enrichment analysis of input metabolites from macrophage subtypes.

and dendritic cells, likely contribute significantly to disease pathology. While previous studies have reported elevated monocyte levels in patients with RA (35), our findings emphasize their increased presence specifically in ACPA– RA, indicating a subtype-specific inflammatory mechanism. This suggests that therapies aimed at monocyte recruitment or activation might be particularly beneficial. Monocyte-targeted interventions, such as inhibitors of monocyte chemoattractant proteins or their receptors, may hold promise for reducing systemic inflammation in ACPA– RA. Additionally, therapies that modulate monocyte differentiation into pro-inflammatory macrophages could help in slowing disease progression in these patients.

In synovial tissue, macrophages from patients with ACPA– RA demonstrated greater heterogeneity and more complex developmental trajectories, forming four distinct branches in pseudotime analysis compared to only two branches in ACPA+

RA. This increased heterogeneity suggests varied activation and differentiation processes, potentially leading to diverse disease courses and therapeutic responses. Notably, macrophages in ACPA− RA were enriched in distinct metabolic pathways, including complement and coagulation cascades, antifolate resistance, and glycosphingolipid biosynthesis. The complement and coagulation cascades are central to immune responses and inflammation. Within the RA context, the complement system contributes to synovial inflammation by promoting opsonization, chemotaxis, and membrane attack complex formation, which drives tissue damage (36). The activation of the coagulation cascade leads to thrombin generation and fibrin deposition in the synovium, intensifying inflammation and encouraging pannus formation (37). These processes establish a pro-inflammatory environment that supports the infiltration and activation of immune cells, such as macrophages and T cells, thereby sustaining joint destruction. Targeting components of the complement and coagulation pathways could thus be a promising therapeutic approach in ACPA− RA, potentially reducing synovial inflammation and preventing joint damage.

The enrichment of metabolic pathways in ACPA− RA macrophages emphasizes the critical role of altered metabolism in disease pathogenesis. The upregulation of complement and coagulation cascades, for example, may intensify inflammatory responses, as components of these pathways act as chemoattractants and immune cell activators (38). Antifolate resistance presents potential treatment challenges, given that methotrexate, a folate antagonist, remains central to RA therapy (39). The identification of glycosphingolipid biosynthesis pathways aligns with evidence that lipid metabolism influences immune cell function and inflammation (40), potentially impacting macrophage activation and cytokine production in ACPA− RA. These insights suggest that therapies targeting metabolic pathways, such as inhibitors of specific enzymes in glycosphingolipid biosynthesis like glucosylceramide synthase, may modulate macrophage function and reduce inflammation in ACPA− RA. Addressing antifolate resistance with alternative disease-modifying antirheumatic drugs (DMARDs) or combination therapies could further enhance treatment efficacy. Some inhibitors, like eliglustat, are already approved for Gaucher disease (41), though their viability in RA requires further investigation. Elucidating macrophage metabolic dependencies in ACPA− RA could guide the development of selective therapies that target pathogenic immune cell subsets while sparing normal immune function.

scFEA further identified 11 upregulated metabolic modules in ACPA− RA macrophages, enriched in pathways like beta-alanine and glutathione metabolism. Beta-alanine metabolism is linked to carnosine synthesis, an antioxidant dipeptide that can modulate inflammatory responses (42). Glutathione metabolism is essential for redox balance and cellular protection against oxidative stress, which is elevated in RA (43). The pronounced role of the C9:MIFhigh(H) macrophage subtype in driving these metabolic pathways suggests that specific macrophage populations contribute to the metabolic reprogramming seen in ACPA− RA. Targeting these metabolic pathways could provide novel therapeutic approaches. Enhancing glutathione levels or modulating its metabolism might alleviate oxidative stress and inflammation in the synovial environment. N-acetylcysteine, a glutathione precursor, is already used clinically for other indications and could be repurposed for RA treatment (44). Similarly, interventions targeting beta-alanine metabolism and carnosine synthesis could influence macrophage activation or cytokine production (45). Identifying the C9:MIFhigh(H) macrophage subtype as a driver of these metabolic alterations highlights it as a potential therapeutic target. Agents that inhibit MIF (macrophage migration inhibitory factor) or its downstream signaling could reduce inflammation and tissue damage in patients with ACPA− RA (46, 47).

Our cell-cell communication analysis revealed that macrophage-monocyte interactions in ACPA− RA are primarily mediated by CD45 and CCL5 signaling pathways. CD45, a receptor tyrosine phosphatase encoded by PTPRC, is critical for T-cell and B-cell receptor signaling and can modulate macrophage activation (48). The involvement of CD45 and its ligands, such as MRC1, suggests a shift in immune regulation in ACPA− RA. Additionally, metabolite-mediated communication analysis highlighted significant engagement of iron-mediated pathways. Elevated iron and L-glutamine levels in patients with ACPA− RA point to a pivotal role for iron metabolism in immune cell interactions. Iron can drive macrophage polarization towards a pro-inflammatory phenotype (49), while the increased abundance of L-glutamine, a key amino acid for immune cell proliferation and function, underscores the metabolic demands of activated immune cells in ACPA− RA (50). Therapeutically, targeting the CD45 pathway may offer a means to modulate macrophage activation and reduce inflammation (48). CD45 inhibitors, already explored in other inflammatory conditions, hold potential for repurposing in RA (51). Modulating iron metabolism presents another promising strategy; iron chelators or agents that regulate iron homeostasis could influence macrophage polarization and attenuate pro-inflammatory responses (52). Additionally, interventions that restrict glutamine availability or inhibit glutamine metabolism could limit immune cell proliferation and activation, providing another therapeutic approach (50). Such strategies may be especially beneficial for patients with ACPA− RA, who often respond suboptimally to standard treatments.

These metabolic alterations may underlie the distinct clinical features of ACPA− RA. Unlike ACPA+ RA, typically associated with more severe joint damage and systemic manifestations, ACPA− RA may follow a different trajectory due to these metabolic distinctions. Our findings support previous research suggesting that metabolic reprogramming of immune cells is a hallmark of autoimmune diseases (53). Future studies should aim to validate these metabolic pathways as biomarkers for disease progression and treatment response in ACPA− RA. Longitudinal studies examining metabolic profile changes pre- and post-therapy could further clarify their clinical utility. Additionally, clinical trials evaluating agents that target these metabolic pathways could assess their efficacy and safety in ACPA− RA, paving the way for more personalized treatment strategies.

Weighted gene co-expression network analysis (WGCNA) identified key gene modules associated with ACPA− RA,

particularly the brown module, which is enriched in immune-related pathways. Intersecting genes from this module with differentially expressed genes in PBMCs and macrophage subtypes pinpointed 20 genes differentially expressed in both compartments. Notably, genes such as HLA-DRA, CD74, and FCER1G, which are involved in antigen presentation and immune activation, emerged as potential biomarkers or therapeutic targets. Modulating HLA-DRA and CD74 could influence antigen presentation to T cells, potentially mitigating autoimmune responses (54). Small molecules or antibodies targeting these proteins could be developed, though this would require extensive research and development. A neural network model utilizing these genes was constructed, effectively distinguishing patients with ACPA– RA from healthy controls with an area under the curve (AUC) of 0.81. This outcome underscores the potential of integrating scRNA-seq data with machine learning to enhance ACPA– RA diagnosis, especially given the absence of specific serological markers in these patients. Early and precise diagnosis is essential for initiating timely treatment and improving patient outcomes. To advance these results into clinical practice, further validation of the neural network model is required. Prospective studies with larger, independent cohorts are necessary to confirm its diagnostic accuracy and reliability. Additionally, integrating this model into clinical workflows would necessitate developing accessible assays or platforms to measure the identified genes, potentially through targeted PCR panels or immunoassays. Considerations around regulatory approval and cost-effectiveness would also be essential. Ultimately, this approach holds promise for enabling earlier diagnosis and more personalized treatment strategies for patients with ACPA– RA.

Limitation

While this study provides valuable insights, certain limitations exist. The cross-sectional design precludes evaluation of temporal changes in immune cell metabolism and function. Future studies with larger, longitudinal cohorts are needed to validate these findings and further investigate the therapeutic potential of targeting metabolic pathways in ACPA– RA.

Conclusion

In conclusion, this study underscores the significant role of altered metabolism in ACPA– RA pathogenesis. The identification of distinct immune cell compositions, metabolic pathways, and intercellular communication patterns enhances understanding of the disease and suggests new avenues for therapeutics targeting metabolic processes. By pinpointing specific metabolic pathways and immune cell interactions unique to ACPA– RA, these findings highlight potential biomarkers and therapeutic targets that could support the development of more effective, personalized treatments. Future research should focus on clinically validating these targets and examining their impact on patient outcomes. Targeting the metabolic reprogramming of immune cells, particularly macrophages, may enable the creation of precise interventions aimed at modulating inflammation and improving clinical outcomes for patients with ACPA– RA.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

YJ: Writing – original draft, Methodology. ZH: Funding acquisition, Resources, Writing – review & editing. RH: Software, Supervision, Writing – review & editing. KH: Project administration, Visualization, Writing – original draft, Writing – review & editing. PW: Formal analysis, Software, Writing – review & editing. JK: Writing – review & editing, Writing – original draft, Funding acquisition.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by the Science Foundation of Hunan Province for Young Scholars (2024JJ6597 to JK); the Science Foundation of Hunan Province for Distinguished Young Scholars (2023JJ10088 to ZH); the National Natural Science Foundation of China (82271379 to ZH); and the Open Project of Guangxi Key Laboratory of Reproductive Health and Birth Defect Prevention (Maternal and Child Health Hospital of Guangxi Zhuang Autonomous Region, GXWCH-ZDKF-2023-01 to ZH).

Acknowledgments

Acknowledgment is extended to the Genome Sequence Archive at the BIG Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, for data repository services. Special thanks go to Prof. Zhang Xuan from the Department of Rheumatology, Beijing Hospital, for his invaluable contributions to this study.

Conflict of interest

The authors declare that this research was conducted without any commercial or financial relationships that could present a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1512483/full#supplementary-material

SUPPLEMENTARY FIGURE 1
(A) Top 3 markers of cell types from PBMC. **(B)** Expression of CD14 and CD16 in monocytes from PBMC. **(C)** Top 3 markers of cell types from STMCs. **(D)** Top 3 markers of cell types in macrophage subpopulations from STMCs.

SUPPLEMENTARY FIGURE 2
(A) Enrichment analysis of DEGs upregulated in ACPA+ macrophage subpopulations. **(B)** Enrichment analysis of DEGs upregulated in ACPA– macrophage subpopulations.

SUPPLEMENTARY TABLE 1
Pathways enriched by pseudotime-related gene intersections.

SUPPLEMENTARY TABLE 2
Genes from WGCNA modules in macrophage populations.

SUPPLEMENTARY TABLE 3
Cell-cell communications between STMC macrophages and PBMC monocytes in patients with ACPA– RA.

SUPPLEMENTARY TABLE 4
Cell-cell communications between STMC macrophages and PBMC monocytes in patients with ACPA+ RA.

SUPPLEMENTARY TABLE 5
Metabolite-mediated cell-cell communications between STMC macrophages and PBMC monocytes in patients with ACPA– RA.

SUPPLEMENTARY TABLE 6
Metabolite-mediated cell-cell communications between STMC macrophages and PBMC monocytes in patients with ACPA+ RA.

References

1. Alivernini S, Firestein GS, McInnes IB. The pathogenesis of rheumatoid arthritis. *Immunity*. (2022) 55:2255–70. doi: 10.1016/j.immuni.2022.11.009

2. Firestein GS. Evolving concepts of rheumatoid arthritis. *Nature*. (2003) 423:356–61. doi: 10.1038/nature01661

3. van Venrooij WJ, van Beers JJ, Pruijn GJ. Anti-CCP antibodies: the past, the present and the future. *Nat Rev Rheumatol*. (2011) 7:391–8. doi: 10.1038/nrrheum.2011.76

4. Nishimura K, Sugiyama D, Kogata Y, Tsuji G, Nakazawa T, Kawano S, et al. Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. *Ann Intern Med*. (2007) 146:797–808. doi: 10.7326/0003-4819-146-11-200706050-00008

5. Rönnelid J, Hansson M, Mathsson-Alm L, Cornillet M, Reed E, Jakobsson PJ, et al. Anticitrullinated protein/peptide antibody multiplexing defines an extended group of ACPA-positive rheumatoid arthritis patients with distinct genetic and environmental determinants. *Ann Rheum Dis*. (2018) 77:203–11. doi: 10.1136/annrheumdis-2017-211782

6. van der Helm-van Mil AH, Huizinga TW. Advances in the genetics of rheumatoid arthritis point to subclassification into distinct disease subsets. *Arthritis Res Ther*. (2008) 10:205. doi: 10.1186/ar2384

7. van Steenbergen HW, Aletaha D, Beaart-van de Voorde LJ, Brouwer E, Codreanu C, Combe B, et al. EULAR definition of arthralgia suspicious for progression to rheumatoid arthritis. *Ann Rheum Dis*. (2017) 76:491–6. doi: 10.1136/annrheumdis-2016-209846

8. van Nies JA, de Jong Z, van der Helm-van Mil AH, Knevel R, Le Cessie S, Huizinga TW. Improved treatment strategies reduce the increased mortality risk in early RA patients. *Rheumatol (Oxford)*. (2010) 49:2210–6. doi: 10.1093/rheumatology/keq250

9. Deane KD, Demoruelle MK, Kelmenson LB, Kuhn KA, Norris JM, Holers VM. Genetic and environmental risk factors for rheumatoid arthritis. *Best Pract Res Clin Rheumatol*. (2017) 31:3–18. doi: 10.1016/j.berh.2017.08.003

10. Pratt AG, Isaacs JD, Mattey DL. Current concepts in the pathogenesis of early rheumatoid arthritis. *Best Pract Res Clin Rheumatol*. (2009) 23:37–48. doi: 10.1016/j.berh.2008.08.002

11. Raza K, Falciani F, Curnow SJ, Ross EJ, Lee CY, Akbar AN, et al. Early rheumatoid arthritis is characterized by a distinct and transient synovial fluid cytokine profile of T cell and stromal cell origin. *Arthritis Res Ther*. (2005) 7:R784–95. doi: 10.1186/ar1733

12. Guo Q, Wang Y, Xu D, Nossent J, Pavlos NJ, Xu J. Rheumatoid arthritis: pathological mechanisms and modern pharmacologic therapies. *Bone Res*. (2018) 6:15. doi: 10.1038/s41413-018-0016-9

13. Edilova MI, Akram A, Abdul-Sater AA. Innate immunity drives pathogenesis of rheumatoid arthritis. *BioMed J*. (2021) 44:172–82. doi: 10.1016/j.bj.2020.06.010

14. Weyand CM, Goronzy JJ. The immunology of rheumatoid arthritis. *Nat Immunol*. (2021) 22:10–8. doi: 10.1038/s41590-020-00816-x

15. Hultqvist M, Olofsson P, Holmberg J, Bäckström BT, Tordsson J, Holmdahl R. Enhanced autoimmunity, arthritis, and encephalomyelitis in mice with a reduced oxidative burst due to a mutation in the Ncf1 gene. *Proc Natl Acad Sci U S A*. (2004) 101:12646–51. doi: 10.1073/pnas.0403831101

16. Bantug GR, Galluzzi L, Kroemer G, Hess C. The spectrum of T cell metabolism in health and disease. *Nat Rev Immunol*. (2018) 18:19–34. doi: 10.1038/nri.2017.99

17. Weyand CM, Goronzy JJ. Immunometabolism in early and late stages of rheumatoid arthritis. *Nat Rev Rheumatol*. (2017) 13:291–301. doi: 10.1038/nrrheum.2017.49

18. Yang Z, Fujii H, Mohan SV, Goronzy JJ, Weyand CM. Phosphofructokinase deficiency impairs ATP generation, autophagy, and redox balance in rheumatoid arthritis T cells. *J Exp Med*. (2013) 210:2119–34. doi: 10.1084/jem.20130252

19. Papalexi E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol*. (2018) 18:35–45. doi: 10.1038/nri.2017.76

20. Zhang F, Wei K, Slowikowski K, Fonseka CY, Rao DA, Kelly S, et al. Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nat Immunol*. (2019) 20:928–42. doi: 10.1038/s41590-019-0378-1

21. Alghamdi N, Chang W, Dang P, Lu X, Wan C, Gampala S, et al. A graph neural network model to estimate cell-wise metabolic flux using single-cell RNA-seq data. *Genome Res*. (2021) 31:1867–84. doi: 10.1101/gr.271205.120

22. Wu X, Liu Y, Jin S, Wang M, Jiao Y, Yang B, et al. Single-cell sequencing of immune cells from anticitrullinated peptide antibody positive and negative rheumatoid arthritis. *Nat Commun*. (2021) 12:4977. doi: 10.1038/s41467-021-25246-7

23. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun*. (2017) 8:14049. doi: 10.1038/ncomms14049

24. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, et al. Comprehensive integration of single-cell data. *Cell*. (2019) 177:1888–1902.e21. doi: 10.1016/j.cell.2019.05.031

25. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*. (2021) 2:100141. doi: 10.1016/j.xinn.2021.100141

26. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. (2017) 14:979–82. doi: 10.1038/nmeth.4402

27. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. (2017) 14:1083–6. doi: 10.1038/nmeth.4463

28. Van de Sande B, Flerin C, Davie K, De Waegeneer M, Hulselmans G, Aibar S, et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat Protoc*. (2020) 15:2247–76. doi: 10.1038/s41596-020-0336-2

29. Suo S, Zhu Q, Saadatpour A, Fei L, Guo G, Yuan GC. Revealing the critical regulators of cell identity in the mouse cell atlas. *Cell Rep*. (2018) 25:1436–1445.e3. doi: 10.1016/j.celrep.2018.10.045

30. Morabito S, Reese F, Rahimzadeh N, Miyoshi E, Swarup V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep Methods*. (2023) 3:100498. doi: 10.1016/j.crmeth.2023.100498

31. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. (2021) 12:1088. doi: 10.1038/s41467-021-21246-9

32. Park MD, Silvin A, Ginhoux F, Merad M. Macrophages in health and disease. *Cell*. (2022) 185:4259–79. doi: 10.1016/j.cell.2022.10.007

33. Gordon S, Plüddemann A. Tissue macrophages: heterogeneity and functions. *BMC Biol*. (2017) 15:53. doi: 10.1186/s12915-017-0392-4

34. Koda Y, Teratani T, Chu PS, Hagihara Y, Mikami Y, Harada Y, et al. CD8(+) tissue-resident memory T cells promote liver fibrosis resolution by inducing apoptosis of hepatic stellate cells. *Nat Commun*. (2021) 12:4474. doi: 10.1038/s41467-021-24734-0

35. Carmona-Rivera C, Kaplan MJ. Low-density granulocytes in systemic autoimmunity and autoinflammation. *Immunol Rev*. (2023) 314:313–25. doi: 10.1111/imr.13161

36. Hietala MA, Nandakumar KS, Persson L, Fahlén S, Holmdahl R, Pekna M. Complement activation by both classical and alternative pathways is critical for the effector phase of arthritis. *Eur J Immunol*. (2004) 34:1208–16. doi: 10.1002/eji.200424895

37. Macfarlane SR, Seatter MJ, Kanke T, Hunter GD, Plevin R. Proteinase-activated receptors. *Pharmacol Rev*. (2001) 53:245–82.

38. Ricklin D, Lambris JD. Complement-targeted therapeutics. *Nat Biotechnol*. (2007) 25:1265–75. doi: 10.1038/nbt1342

39. Siddiqui A, Totonchian A, Jabar Ali JB, Ahmad I, Kumar J, Shiwlani S, et al. Risk factors associated with non-respondence to methotrexate in rheumatoid arthritis patients. *Cureus*. (2021) 13:e18112. doi: 10.7759/cureus.18112

40. Peters L, Kuebler WM, Simmons S. Sphingolipids in atherosclerosis: chimeras in structure and function. *Int J Mol Sci*. (2022) 23:11948. doi: 10.3390/ijms231911948

41. Cox TM, Drelichman G, Cravo R, Balwani M, Burrow TA, Martins AM, et al. Eliglustat maintains long-term clinical stability in patients with Gaucher disease type 1 stabilized on enzyme therapy. *Blood*. (2017) 129:2375–83. doi: 10.1182/blood-2016-12-758409

42. Perim P, Marticorena FM, Ribeiro F, Barreto G, Gobbi N, Kerksick C, et al. Can the skeletal muscle carnosine response to beta-alanine supplementation be optimized. *Front Nutr*. (2019) 6:135. doi: 10.3389/fnut.2019.00135

43. Mateen S, Moin S, Khan AQ, Zafar A, Fatima N. Increased reactive oxygen species formation and oxidative stress in rheumatoid arthritis. *PloS One*. (2016) 11: e0152925. doi: 10.1371/journal.pone.0152925

44. Luo A, Liu X, Hu Q, Yang M, Jiang H, Liu W. Efficacy of N-acetylcysteine on idiopathic or postinfective non-cystic fibrosis bronchiectasis: a systematic review and meta-analysis protocol. *BMJ Open*. (2022) 12:e053625. doi: 10.1136/bmjopen-2021-053625

45. Artioli GG, Gualano B, Smith A, Stout J, Lancha AH Jr. Role of beta-alanine supplementation on muscle carnosine and exercise performance. *Med Sci Sports Exerc*. (2010) 42:1162–73. doi: 10.1249/MSS.0b013e3181c74e38

46. Al-Abed Y, VanPatten S. MIF as a disease target: ISO-1 as a proof-of-concept therapeutic. *Future Med Chem*. (2011) 3:45–63. doi: 10.4155/fmc.10.281

47. Cho Y, Crichlow GV, Vermeire JJ, Leng L, Du X, Hodsdon ME, et al. Allosteric inhibition of macrophage migration inhibitory factor revealed by ibudilast. *Proc Natl Acad Sci U S A*. (2010) 107:11313–8. doi: 10.1073/pnas.1002716107

48. Hermiston ML, Xu Z, Weiss A. CD45: a critical regulator of signaling thresholds in immune cells. *Annu Rev Immunol*. (2003) 21:107–37. doi: 10.1146/annurev.immunol.21.120601.140946

49. Recalcati S, Locati M, Marini A, Santambrogio P, Zaninotto F, De Pizzol M, et al. Differential regulation of iron homeostasis during human macrophage polarized activation. *Eur J Immunol*. (2010) 40:824–35. doi: 10.1002/eji.200939889

50. Newsholme P, Curi R, Pithon Curi TC, Murphy CJ, Garcia C, Pires de Melo M. Glutamine metabolism by lymphocytes, macrophages, and neutrophils: its importance in health and disease. *J Nutr Biochem*. (1999) 10:316–24. doi: 10.1016/S0955-2863(99)00022-4

51. Rider DA, Bayley R, Clay E, Young SP. Does oxidative inactivation of CD45 phosphatase in rheumatoid arthritis underlie immune hyporesponsiveness. *Antioxid Redox Signal*. (2013) 19:2280–5. doi: 10.1089/ars.2013.5458

52. Recalcati S, Locati M, Gammella E, Invernizzi P, Cairo G. Iron levels in polarized macrophages: regulation of immunity and autoimmunity. *Autoimmun Rev*. (2012) 11:883–9. doi: 10.1016/j.autrev.2012.03.003

53. O'Neill LA, Kishton RJ. Rathmell J. *A guide to immunometabolism immunologists. Nat Rev Immunol*. (2016) 16:553–65. doi: 10.1038/nri.2016.70

54. Su H, Na N, Zhang X, Zhao Y. The biological function and significance of CD74 in immune diseases. *Inflammation Res*. (2017) 66:209–16. doi: 10.1007/s00011-016-0995-1

frontiers | Frontiers in Endocrinology

TYPE Mini Review
PUBLISHED 09 January 2025
DOI 10.3389/fendo.2024.1528248

OPEN ACCESS

EDITED BY
Yejun Tan,
University of Minnesota Health Twin Cities, United States

REVIEWED BY
Wei Zhou,
National Clinical Research Center For Child Health, China
Wei Jiang,
Nanjing Agricultural University, China

*CORRESPONDENCE
Xiao-xiao Yang
yangxiaoxiao2057@163.com
Yi-bo He
20173626@zcmu.deu.cn

†These authors have contributed equally to this work

RECEIVED 14 November 2024
ACCEPTED 17 December 2024
PUBLISHED 09 January 2025

CITATION
Chen S-l, Fei Y-R, Cai X-x, Wang C, Tong S-y, Zhang Z-z, Huang Y-x, Bian D-d, He Y-b and Yang X-x (2025) Exploring the role of metabolic pathways in TNBC immunotherapy: insights from single-cell and spatial transcriptomics. *Front. Endocrinol.* 15:1528248. doi: 10.3389/fendo.2024.1528248

Exploring the role of metabolic pathways in TNBC immunotherapy: insights from single-cell and spatial transcriptomics

Shi-liang Chen[1†], Yi-Ran Fei[1,2†], Xin-xian Cai[1,3], Cong Wang[1], Shi-yuan Tong[4], Zhe-zhong Zhang[1], Yan-xia Huang[1], Dan-dan Bian[1], Yi-bo He[1*] and Xiao-xiao Yang[1*]

[1]The First Affiliated Hospital of Zhejiang Chinese Medical University (Zhejiang Provincial Hospital of Chinese Medicine), Hangzhou, China, [2]The First Clinical Medical College, Zhejiang Chinese Medical University, Hangzhou, China, [3]School of Medical Technology and Informmation Engineering, Zhejiang Chinese Medical University, Hangzhou, China, [4]State Key Laboratory of Medical Neurobiology and MOE Frontiers Center for Brain Science, Institutes of Brain Science, Fudan University, Shanghai, China

The article provides an overview of the current understanding of the interplay between metabolic pathways and immune function in the context of triple-negative breast cancer (TNBC). It highlights recent advancements in single-cell and spatial transcriptomics technologies, which have revolutionized the analysis of tumor heterogeneity and the immune microenvironment in TNBC. The review emphasizes the crucial role of metabolic reprogramming in modulating immune cell function, discussing how specific metabolic pathways, such as glycolysis, lipid metabolism, and amino acid metabolism, can directly impact the activity and phenotypes of various immune cell populations within the TNBC tumor microenvironment. Furthermore, the article explores the implications of these metabolic-immune interactions for the efficacy of immune checkpoint inhibitor (ICI) therapies in TNBC, suggesting that strategies targeting metabolic pathways may enhance the responsiveness to ICI treatments. Finally, the review outlines future directions and the potential for combination therapies that integrate metabolic modulation with immunotherapeutic approaches, offering promising avenues for improving clinical outcomes for TNBC patients.

KEYWORDS

triple-negative breast cancer, metabolic reprogramming, immune cell function, immune checkpoint inhibitors, combination therapies, tumor microenvironment

Introduction

Triple-negative breast cancer (TNBC) treatment has historically relied on chemotherapy due to the absence of targeted therapies, limiting effective options. Recent advancements in immunotherapy, particularly immune checkpoint inhibitors (ICIs) like PD-L1 inhibitors, have shown potential, as evidenced by trials such as IMpassion130 (1, 2). However, the modest efficacy of ICIs, benefiting only a subset of patients, highlights the challenges TNBC's heterogeneity poses. Identifying predictive biomarkers and exploring combination strategies, including metabolic interventions, are critical to improving therapeutic outcomes and addressing TNBC's metabolic pathways (3).

Although TNBC is generally considered a "cold" tumor with limited immune cell infiltration, emerging evidence suggests it has antigenic properties conducive to immunotherapy (4). TNBC generally shows low levels of tumor-infiltrating lymphocytes (TILs); the presence of specific immune cell types can correlate with better patient outcomes. Specific immune markers, such as granzyme B^+ $CD8^+$ T cells (5), sometimes correlate with improved prognosis. Additionally, plasma cells and other immune subsets have been linked to survival benefits (6), challenging the traditional view of TNBC as uniformly immunologically inactive. Understanding this heterogeneity is crucial for tailoring immunotherapies to re-engage the immune system effectively (7).

ICIs, which block proteins like PD-1 that suppress immune responses, have emerged as promising therapies for TNBC. Cytotoxic T lymphocytes (CTLs) play a pivotal role in anti-tumor immunity, while regulatory T cells (Tregs) can hinder these responses. $PD\text{-}1^+$ CTLs and other tumor-infiltrating lymphocytes (TILs) significantly impact ICI efficacy. A nuanced understanding of the interactions between immune cell populations within TNBC is vital for optimizing immunotherapeutic strategies. Advancements in single-cell RNA sequencing (scRNA-seq) and spatial transcriptomics have revolutionized TNBC research (8). These technologies provide unprecedented insights into tumor heterogeneity and immune microenvironments by analyzing gene expression at single-cell resolution and mapping spatial interactions (9–11). Studies reveal diverse immune cell subsets and spatial relationships, offering new biomarkers and therapeutic targets (12). Integrating these technologies enables researchers to uncover immune evasion mechanisms and develop tailored immunotherapeutic strategies.

Applying these advanced techniques has deepened the understanding of TNBC's tumor microenvironment and immune interactions. Researchers can better predict therapeutic responses and personalize treatment (13) by identifying cellular diversity and spatial organization. These insights hold significant potential for identifying novel targets, improving clinical outcomes, and advancing precision medicine in TNBC.

Metabolic influences on immune cells in the tumor microenvironment

Link between metabolic pathways and immune function

The metabolic landscape within the tumor microenvironment (TME) significantly impacts the behavior and functionality of immune cells. Tumors often exhibit altered metabolic pathways, producing specific metabolites that can modulate immune responses (14, 15). For instance, it has been observed that the accumulation of lactate, a byproduct of glycolysis, can create an immunosuppressive environment by inhibiting the function of cytotoxic T cells and promoting regulatory T cells (16). Furthermore, tumor-derived metabolites such as adenosine can disrupt T cell activation and promote immune evasion mechanisms (17). This dynamic interplay creates a feedback loop in which tumor cells' metabolic state affects their proliferation and survival and influences the immune landscape, leading to an environment conducive to tumor progression (18).

Recent studies have elucidated how specific metabolic pathways in tumor cells can directly alter the immune response. For example, it was found that activating IDO pathway in tumors results in tryptophan catabolism, leading to T cell dysfunction and promoting an immune-suppressive environment (19). Moreover, the Warburg effect, characterized by increased aerobic glycolysis in tumor cells, has created an environment that favors the recruitment of immunosuppressive cell types while inhibiting effector T cell functions (20). Understanding these metabolic interactions is crucial for developing strategies to reprogram the TME to reinvigorate anti-tumor immunity (21).

Importance of metabolic pathways in modulating immune response

Metabolic reprogramming is emerging as a critical factor influencing immune cell functionality and their therapeutic responses. Immune cells adapt their metabolism to fulfill their bioenergetic and biosynthetic needs during activation. For instance, T cells require metabolic reprogramming towards glycolysis to sustain their proliferation and effector functions. However, a skewed metabolic environment can lead to dysfunction (16). Furthermore, studies have shown that targeting metabolic pathways enhances the immune response against TNBC. By inhibiting metabolic checkpoints like mTOR and AMPK, it is possible to improve T cell activation and restore anti-tumor immunity (5).

In the context of TNBC, therapeutic strategies focusing on metabolic reprogramming show promise in enhancing the efficacy of existing treatments. Combining metabolic inhibitors with immunotherapy has been proposed as a novel approach to improve the anti-tumor immune response. For instance, recent research highlights the potential of using metabolic modulators to enhance the effectiveness of immune checkpoint inhibitors, which could lead to better clinical outcomes for TNBC patients (19). Overall, understanding the intricate relationship between metabolic pathways and immune function presents an opportunity to develop innovative strategies to augment the effectiveness of therapies to TNBC.

Metabolic pathways and their effects on immune cells in TNBC

Overview of key metabolic pathways

Metabolic pathways are critical determinants of immune cell function and can significantly influence the efficacy of anti-tumor

responses in TNBC. Key metabolic processes, including glycolysis, lipid metabolism, and amino acid metabolism, orchestrate the activities of various immune cells (Figure 1). Glycolysis, for instance, is vital for T cell activation and proliferation. Increased glycolytic activity in T cells correlates with enhanced effector functions, allowing them to respond effectively to tumor cells (22). In contrast, fatty acid oxidation is crucial for the developing and maintaining memory T cells, ensuring long-lasting immune protection against recurrent tumors (23).

Amino acid metabolism also plays a pivotal role in immune responses. The availability of specific amino acids, such as glutamine, influences T cell metabolism and function. Tumor cells often deplete local amino acids, leading to T cell dysfunction and impaired anti-tumor activity (24). Understanding these metabolic pathways provides insights into how metabolic reprogramming in immune cells can enhance their functionality and effectiveness against TNBC.

Effects on different immune cells

Regulatory T cells

Lactate accumulation enhances the immunosuppressive function of Treg cells by activating FOXP3 gene expression (25). Lactate also induces Treg cell proliferation, allowing them to dominate within the tumor microenvironment, further diminishing the activity of $CD8^+$ T cells, NK cells, and thereby supporting immune evasion by the tumor. Adenosine, catalyzed by ADA2 (Adenosine Deaminase 2), activates the A2A receptor on Treg cells, enhancing their immunosuppressive functions (26). Elevated adenosine levels strengthen Treg cell function and reduce effector T cell activation, fostering an immunosuppressive environment. Treg cell metabolism relies on specific lipid metabolic pathways, with these metabolites promoting Treg immunosuppressive abilities through particular lipid transport proteins, such as FABP5 (27). Lipid accumulation in Treg cells facilitates their proliferation and survival, further diminishing the activity of effector T cells within the tumor microenvironment (28).

Macrophages

The metabolic profiles of macrophages are critical in dictating their pro-tumor or anti-tumor functions. In TNBC, metabolic reprogramming within macrophages can lead to polarization towards a tumor-promoting M2 phenotype characterized by immunosuppressive properties (29). Conversely, promoting metabolic shifts towards an M1-like state can enhance their anti-tumor capabilities. Understanding these metabolic dynamics could lead to novel strategies for reprogramming macrophages to adopt anti-tumor phenotypes, potentially improving therapeutic outcomes in TNBC. Lactate induces macrophage polarization towards an immunosuppressive M2 phenotype, giving rise to tumor-associated macrophages (TAMs) (30, 31). These M2-polarized macrophages secrete elevated levels of immunosuppressive factors, such as IL-10 and TGF-β, which inhibit the antitumor responses of T cells and NK cells (32). PRMT5, by modulating iron metabolism, restricts the pro-inflammatory activity of M1 macrophages, thereby allowing the immunosuppressive properties of the M2 phenotype to predominate (33). Reducing in iron ions further promotes M2 polarization by inhibiting the NRF2/HMOX1 pathway. CAFs upregulate lipid metabolism, driving macrophages toward a lipid-associated macrophage (LAM) phenotype (34), forming immunosuppressive macrophages. These macrophages enhance immunosuppressive effects through lipid signaling molecules, reducing the functional infiltration of effector immune cells within the tumor.

$CD8^+$ T cells

Within the tumor microenvironment, the accumulation of high concentrations of lactate results in functional impairment of $CD8^+$ T cells by lowering the local pH. TNFR2 enhances immunosuppressive capacity in endothelial cells by inhibiting the glycolytic pathway, resulting in decreased $CD8^+$ T cell activity. Blocking TNFR2, however, can restore antitumor immunity (30, 35). Lactate interferes with the mTOR signaling pathway, inhibiting T cell proliferation and diminishing the secretion of key effector cytokines, such as IFN-γ, further compromising antitumor immunity (31). Through the Warburg effect, TNBC cells preferentially consume glucose, leading to glucose deprivation in the surrounding environment, which hampers $CD8^+$ T cells' ability to maintain the glucose levels required for efficient glycolysis. This glucose deficiency directly reduces the activity of $CD8^+$ T cells and, by limiting energy supply through the PI3K/AKT/mTOR pathway (36), decreases their proliferation and cytotoxicity. In the TNBC microenvironment, high glutamine uptake exhausts the glutamine needed by immune cells, adversely affecting particularly the antioxidant-dependent $CD8^+$ T cells (37). Glutamine scarcity restricts glutathione synthesis in T cells, reducing their tolerance to oxidative stress and weakening their antioxidative and antitumor functions within the tumor.

Natural killer cells

Lactate significantly diminishes the cytotoxicity of NK cells, reducing their tumor-killing capacity. The accumulation of lactate also compromises NK cell survival and proliferation by acidifying the environment, further weakening their immune clearance functions (38). In lung cancer, lactate has been found to upregulate PD-L1 expression on tumor cell surfaces via the Warburg effect (39), suggesting a close link between metabolic reprogramming of tumor cells and immune evasion mechanisms. This pathway promotes NK cell exhaustion, enhancing the tumor's ability to evade immune detection (40, 41).

Dendritic cells

Dendritic cells are pivotal in antigen presentation and the initiation of T cell responses. Tumor-derived metabolites can profoundly influence DC function, affecting their ability to activate T cells effectively. Exosomes secreted by TNBC cells, rich in pro-inflammatory molecules, activate the cGAS/STING pathway in dendritic cells (42, 43), thereby enhancing the

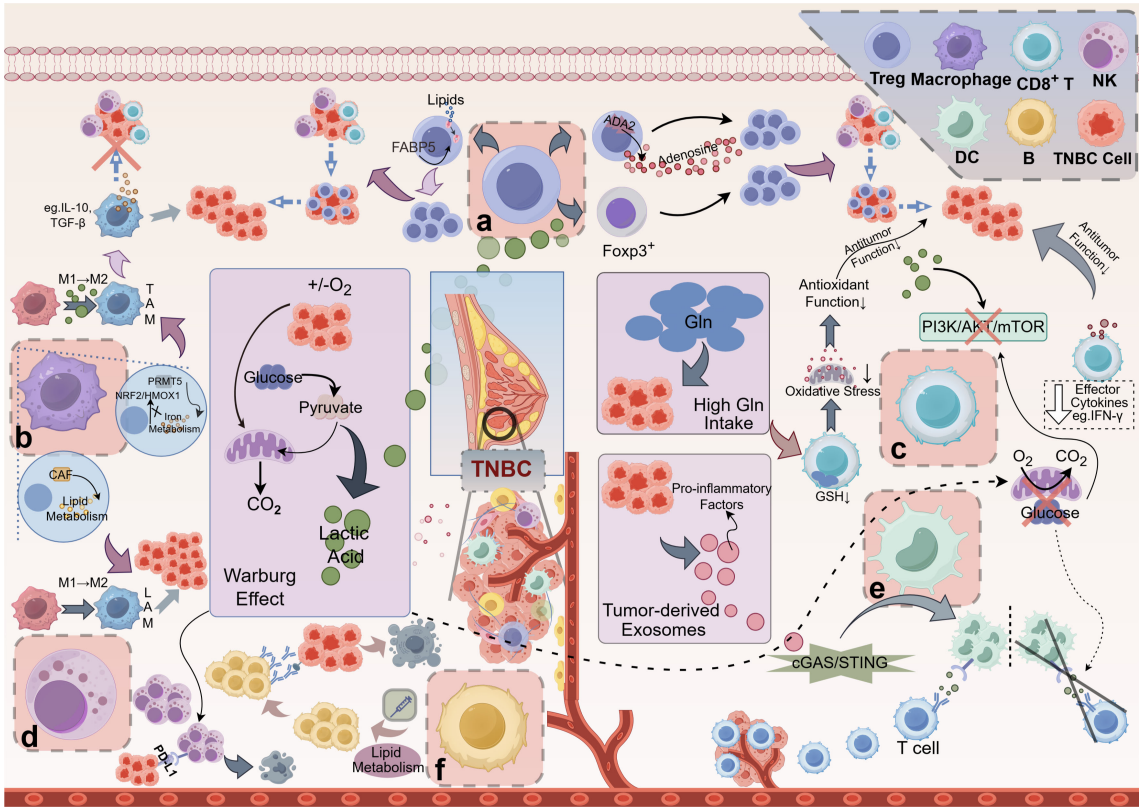


FIGURE 1
Mechanisms of Metabolite-Mediated Immunosuppression in the TNBC. **(A)** Tregs: Lactic acid boosts FOXP3 expression in Tregs, enhancing their proliferation and dominance in the TME, which inhibits $CD8^+$ T cells and NK cells, promoting tumor immune evasion. Adenosine, produced via ADA2 on Tregs, further reinforces Treg predominance, contributing to the immunosuppressive TME. Lipid metabolites, transported by proteins like FABP5, also support Treg proliferation and function. **(B)** Macrophages: Lactic acid drives macrophage polarization to the immunosuppressive M2 phenotype, forming TAMs that secrete IL-10 and TGF-β, suppressing antitumor responses. PRMT5 regulates iron metabolism, inhibiting pro-inflammatory M1 macrophages and favoring M2 polarization. CAF-induced lipid metabolism upregulation further promotes M2 macrophage transformation. (C) $CD8^+$ T cells: High lactic acid in the TNBC microenvironment lowers local pH and disrupts the PI3K/AKT/mTOR pathway, impairing $CD8^+$ T cell proliferation and cytokine secretion (e.g., IFN-γ), weakening antitumor immunity. Glucose consumption via the Warburg effect depletes glucose needed for glycolysis in $CD8^+$ T cells, exacerbating this inhibition. Additionally, high Gln uptake by TNBC cells reduces GSH synthesis in $CD8^+$ T cells, impairing oxidative stress tolerance and antitumor function. **(D)** NK cells: Lactic acid increases PD-L1 expression on tumor cells, binding PD-1 on NK cells and inhibiting their cytotoxicity, leading to NK cell exhaustion and tumor immune escape. **(E)** DCs: TNBC cells release exosomes that activate dendritic cells via the cGAS/STING pathway, enhancing T cell activation and immune responses. **(F)** B cells: Lipid metabolism significantly impacts B cell function, particularly in the context of antibody production and memory formation. However, metabolic reprogramming, such as the Warburg effect, reduces glycolysis in DCs, impairing their maturation and antigen presentation, thus weakening T cell activation and effector immune cell infiltration in the tumor. TNBC, Triple-Negative Breast Cancer; Treg, Regulatory T Cell; FOXP3, Forkhead Box P3; TME, Tumor Microenvironment; NK cell, Natural Killer Cell; ADA2, Adenosine Deaminase 2; FABP5, Fatty Acid-Binding Protein 5; TAM, Tumor-Associated Macrophage; IL-10, Interleukin 10; TGF-β, Transforming Growth Factor Beta; PRMT5, Protein Arginine Methyltransferase 5; CAF, Cancer-Associated Fibroblast; LAM, Lipid-Associated Macrophage; PI3K, Phosphoinositide 3-Kinase; AKT, Protein Kinase B (often referred to as AKT); mTOR, Mechanistic Target of Rapamycin; IFN-γ, Interferon Gamma; Gln, glutamine; GSH, Glutathione; PD-L1, Programmed Death-Ligand 1; PD-1, Programmed Death-1; DC, Dendritic Cell; cGAS, Cyclic GMP-AMP Synthase; STING, Stimulator of Interferon Genes. This figure was created using the Figdraw online drawing tool.

initiation of antitumor immune responses. The release of these exosomes bolsters DC activity, facilitating downstream T cell activation and strengthening immune responses within the tumor microenvironment. However, metabolic reprogramming in TNBC suppresses the glycolytic pathway in dendritic cells (44), impairing their maturation and activation capabilities and consequently weakening their efficacy in T cell activation. This metabolic inhibition directly impacts the antigen-presenting capacity of DCs, resulting in reduced infiltration and activity of T cells within the tumor. Researchers can explore therapeutic strategies to enhance DC-mediated T cell activation by understanding how tumor metabolism affects DC function. Targeting metabolic pathways in DCs may help restore their function and improve the overall anti-tumor immune response (45, 46).

B cells

Lipid metabolism significantly impacts B cell function, particularly in antibody production and memory formation. Research indicates that B cells rely on fatty acid metabolism for optimal antibody responses (47). In TNBC, modulating lipid metabolic pathways could enhance the effectiveness of therapeutic vaccines by promoting robust B cell activation and differentiation. For example, interventions that enhance lipid uptake and utilization by B cells might increase their ability to produce high-affinity

antibodies against tumor antigens (48). This approach could represent a novel strategy to improve vaccine efficacy in TNBC patients.

Future directions and clinical implications

The interaction between metabolic pathways and immune responses offers promising avenues for combination therapies in TNBC. Studies suggest that metabolic reprogramming can significantly enhance immune cell efficacy against tumors. Combining ICIs with metabolic pathway-targeting agents could boost anti-tumor immunity by reinvigorating T cells or enhancing TIL populations. These strategies may overcome the limitations of ICIs alone and lead to more personalized treatments based on individual metabolic profiles. Advanced techniques like single-cell functional enzymatic assays (scFEA) and metabolic profiling tools (e.g., Mebocost, scMetabolism) enable more insights into immune cell metabolism in TNBC. These tools allow single-cell analysis of metabolic activity, helping researchers understand how immune cells metabolize nutrients within the tumor environment. By mapping the metabolic landscape, researchers can identify critical metabolic checkpoints as therapeutic targets, which could lead to optimized immune function therapies. The goal of immune metabolism research in TNBC is to translate findings into clinical practice. Targeting specific metabolic pathways, such as glycolysis, could lead to personalized therapies that improve survival and quality of life for TNBC patients. Collaboration between researchers and clinicians is essential, with clinical trials for combination therapies already underway, signaling a shift towards personalized medicine in TNBC and improved treatment outcomes.

Conclusion

Metabolic products, such as lactate and adenosine, are pivotal in establishing an immunosuppressive tumor microenvironment by modulating immune cell functions. Lactate has been shown to promote the proliferation of Tregs while impairing the functionality of cytotoxic $CD8^+$ T cells, and adenosine disrupts T cell activation via the A2A receptor. Recent research further highlights the impact of glutamine depletion on T cell oxidative stress tolerance, demonstrating the intricate connection between metabolic reprogramming and immune responses in TNBC. These findings underscore the potential of combining ICIs with metabolic modulators targeting pathways such as glycolysis and fatty acid oxidation. Preclinical models indicate that such combination therapies can effectively reinvigorate exhausted T cells and enhance antitumor immunity, paving the way for improved therapeutic strategies.

Author contributions

SC: Conceptualization, Data curation, Formal analysis, Validation, Writing – original draft. YF: Conceptualization, Data curation, Formal analysis, Validation, Visualization, Writing – original draft. XC: Conceptualization, Data curation, Formal analysis, Validation, Writing – original draft. CW: Data curation, Project administration, Visualization, Writing – review & editing. ST: Data curation, Formal analysis, Project administration, Supervision, Writing – original draft. ZZ: Formal analysis, Methodology, Supervision, Writing – review & editing. YH: Data curation, Project administration, Visualization, Writing – review & editing. DB: Data curation, Project administration, Visualization, Writing – review & editing. YH: Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing. XY: Conceptualization, Formal analysis, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The study design, data collection, data analysis, manuscript preparation, and publication decisions of this work were supported by Zhejiang Province Traditional Chinese Medicine Science and Technology Project (No. 2025ZR123 by SC, No. 2023ZL409 by XY, No. 2024ZR015 by YH, No. 2023ZL056 by ZZ), Zhejiang Provincial Natural Science Foundation(No. QN25H270030 by YH), Zhejiang Province Medical and Health Science and Technology Project (No. 2024KY1201 by YH, 2024KY1225 by CW, 2024KY1213 by ZZ), the Foundation Project of Zhejiang Chinese Medical University (No. 2022JKZKTS26 by YH, No. 2022JKJNTZ16 by SC, No. 2022JKJNTZ23 by CW).

Acknowledgments

Figure in our manuscript was created using Figdraw online tools (https://www.figdraw.com/). We would like to express our gratitude for these invaluable resources.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. In the course of preparing this work, the authors utilized ChatGPT-4.0 to refine the language. Following the use of this tool, the authors thoroughly reviewed and revised the content as necessary and assume full responsibility for the final publication.

Publisher's note

References

1. Gagliato D de M, Buzaid A, Pérez-García J, Cortés J. Immunotherapy in breast cancer: current practice and clinical challenges. *BioDrugs*. (2020) 34:611–23. doi: 10.1007/s40259-020-00436-9

2. Kim I, Sanchez K, McArthur H, Page D. Immunotherapy in triple-negative breast cancer: present and future. *Curr Breast Cancer Rep*. (2019) 11:259–71. doi: 10.1007/s12609-019-00345-z

3. Li L, Zhang F, Liu Z, Fan Z. Immunotherapy for triple-negative breast cancer: combination strategies to improve outcome. *Cancers*. (2023) 15:321. doi: 10.3390/cancers15010321

4. Deng L, Lu D, Bai Y, Wang Y, Bu H, Zheng H. Immune profiles of tumor microenvironment and clinical prognosis among women with triple-negative breast cancer. *Cancer Epidemiol Biomarkers Prev*. (2019) 28:1977–85. doi: 10.1158/1055-9965.EPI-19-0469

5. Gruosso T, Gigoux M, Manem VSK, Bertos N, Zuo D, Perlitch I, et al. Spatially distinct tumor immune microenvironments stratify triple-negative breast cancers. *J Clin Invest*. (2019) 129:1785–800. doi: 10.1172/JCI96313

6. He TF, Yost SE, Frankel PH, Dagis A, Cao Y, Wang R, et al. Multi-panel immunofluorescence analysis of tumor infiltrating lymphocytes in triple negative breast cancer: Evolution of tumor immune profiles and patient prognosis. *PloS One*. (2020) 15: e0229955. doi: 10.1371/journal.pone.0229955

7. Kim JY, Jung HH, Sohn I, Woo SY, Cho H, Cho E, et al. Prognostication of a 13-immune-related-gene signature in patients with early triple-negative breast cancer. *Breast Cancer Res Treat*. (2020) 184:325–34. doi: 10.1007/s10549-020-05874-1

8. Andersson A, Larsson L, Stenbeck L, Salmén F, Ehinger A, Wu SZ, et al. Spatial deconvolution of HER2-positive breast cancer delineates tumor-associated cell type interactions. *Nat Commun*. (2021) 12:6012. doi: 10.1038/s41467-021-26271-2

9. Bassiouni R, Idowu MO, Gibbs LD, Robila V, Grizzard PJ, Webb MG, et al. Spatial transcriptomic analysis of a diverse patient cohort reveals a conserved architecture in triple-negative breast cancer. *Cancer Res*. (2022) 83:34–48. doi: 10.1158/0008-5472.CAN-22-2682

10. Li J, Long S, Yang Z, Wei W, Yu S, Liu Q, et al. Single-cell transcriptomics reveals IRF7 regulation of the tumor microenvironment in isocitrate dehydrogenase wild-type glioma. *MedComm*. (2024) 5:e754. doi: 10.1002/mco2.754

11. Li J, Zhang Y, Liang C, Yan X, Hui X, Liu Q. Advancing precision medicine in gliomas through single-cell sequencing: unveiling the complex tumor microenvironment. *Front Cell Dev Biol*. (2024) 12:1396836. doi: 10.3389/fcell.2024.1396836

12. Monjo T, Koido M, Nagasawa S, Suzuki Y, Kamatani Y. Efficient prediction of a spatial transcriptomics profile better characterizes breast cancer tissue sections without costly experimentation. *Sci Rep*. (2021) 12:4133. doi: 10.1038/s41598-022-07685-4

13. Yoosuf N, Navarro JF, Salmén F, Ståhl PL, Daub C. Identification and transfer of spatial transcriptomics signatures for cancer diagnosis. *Breast Cancer Res: BCR*. (2020) 22:6. doi: 10.1186/s13058-019-1242-9

14. Feng D. Energy metabolism-related gene prognostic index predicts biochemical recurrence for patients with prostate cancer undergoing radical prostatectomy. *Front Immunol*. (2022) 13:839362. doi: 10.3389/fimmu.2022.839362

15. Feng D, Xiong Q, Wei Q, Yang L. Cellular landscape of tumour microenvironment in prostate cancer. *Immunology*. (2023) 168:199–202. doi: 10.1111/imm.13456

16. Zheng H, Siddharth S, Parida S, Wu X, Sharma D. Tumor microenvironment: key players in triple negative breast cancer immunomodulation. *Cancers*. (2021) 13:3357. doi: 10.3390/cancers13133357

17. Xiao Y, Ma D, Zhao S, Suo C, Shi J, Xue MZ, et al. Multi-omics profiling reveals distinct microenvironment characterization and suggests immune escape mechanisms of triple-negative breast cancer. *Clin Cancer Res*. (2019) 25:5002–14. doi: 10.1158/1078-0432.CCR-18-3524

18. Zhang Y, Qin N, Wang X, Liang R, Liu Q, Geng R, et al. Glycogen metabolism-mediated intercellular communication in the tumor microenvironment influences liver cancer prognosis. *OR*. (2024) 32:563–76. doi: 10.32604/or.2023.029697

19. Sun X, Wang M, Wang M, Yu X, Guo J, Sun T, et al. Metabolic reprogramming in triple-negative breast cancer. *Front Oncol*. (2020) 10:428. doi: 10.3389/fonc.2020.00428

20. Ramzy A, ElSafy S, Elshoky HA, Soliman A, Youness R, Mansour S, et al. Drugless nanoparticles tune-up an array of intertwined pathways contributing to immune checkpoint signaling and metabolic reprogramming in triple-negative breast cancer. *Biomed Mater*. (2022) 18:015023. doi: 10.1088/1748-605X/aca85d

21. Xiong Q. Fatty acid synthase is the key regulator of fatty acid metabolism and is related to immunotherapy in bladder cancer. *Front Immunol*. (2022) 13. doi: 10.3389/fimmu.2022.836939

22. Li G, Ma X, Sui S, Chen Y, Li H, Liu L, et al. NAT10/ac4C/JunB facilitates TNBC malignant progression and immunosuppression by driving glycolysis addiction. *J Exp Clin Cancer Res*. (2024) 43:278.

23. Raud B, Roy DG, Divakaruni AS, Tarasenko TN, Franke R, Ma EH, et al. Etomoxir actions on regulatory and memory T cells are independent of Cpt1a-mediated fatty acid oxidation. *Cell Metabolism*. (2018) 28:504–15.e7.

24. Zou J, Mai C, Lin Z, Zhou J, Lai G. Targeting metabolism of breast cancer and its implications in T cell immunotherapy. *Front Immunol*. (2024) 15:1381970.

25. Angelin A, Gil-de-Gómez L, Dahiya S, Jiao J, Guo L, Levine MH, et al. Foxp3 reprograms T cell metabolism to function in low-glucose, high-lactate environments. *Cell Metab*. (2017) 25:1282–1293.e7. doi: 10.1016/j.cmet.2016.12.018

26. Leone RD, Sun IM, Oh MH, Sun IH, Wen J, Englert J, et al. Inhibition of the adenosine A2a receptor modulates expression of T cell coinhibitory receptors and improves effector function for enhanced checkpoint blockade and ACT in murine cancer models. *Cancer Immunol Immunother*. (2018) 67:1271–84. doi: 10.1007/s00262-018-2186-0

27. Field CS, Baixauli F, Kyle RL, Puleston DJ, Cameron AM, Sanin DE, et al. Mitochondrial integrity regulated by lipid metabolism is a cell-intrinsic checkpoint for treg suppressive function. *Cell Metab*. (2020) 31:422–437.e5. doi: 10.1016/j.cmet.2019.11.021

28. Pacella I, Procaccini C, Focaccetti C, Miacci S, Timperi E, Faicchia D, et al. Fatty acid metabolism complements glycolysis in the selective regulatory T cell expansion during tumor growth. *Proc Natl Acad Sci*. (2018) 115:E6546–55. doi: 10.1073/pnas.1720113115

29. Murray PJ, Wynn TA. Protective and pathological functions of macrophage subsets. *Nat Rev Immunol*. (2021) 21:421–36. doi: 10.1038/s41577-021-00551-4

30. Deng X, Zhu Y, Dai Z, Liu Q, Song Z, Liu T, et al. A bimetallic nanomodulator to reverse immunosuppression via sonodynamic-ferroptosis and lactate metabolism modulation. *Small*. (2024) 20:2404580. doi: 10.1002/smll.202404580

31. Naik A, Decock J. Lactate metabolism and immune modulation in breast cancer: A focused review on triple negative breast tumors. *Front Oncol*. (2020) 10:598626. doi: 10.3389/fonc.2020.598626

32. Santoni M, Romagnoli E, Saladino T, Foghini L, Guarino S, Capponi M, et al. Triple negative breast cancer: Key role of Tumor-Associated Macrophages in regulating the activity of anti-PD-1/PD-L1 agents. *Biochim Biophys Acta (BBA) Rev Cancer*. (2018) 1869:78–84. doi: 10.1016/j.bbcan.2017.10.007

33. Wang Z, Li R, Hou N, Zhang J, Wang T, Fan P, et al. PRMT5 reduces immunotherapy efficacy in triple-negative breast cancer by methylating KEAP1 and inhibiting ferroptosis. *J Immunother Cancer*. (2023) 11:e006890. doi: 10.1136/jitc-2023-006890

34. Timperi E, Gueguen P, Molgora M, Magagna I, Kieffer Y, Lopez-Lastra S, et al. Lipid-associated macrophages are induced by cancer-associated fibroblasts and mediate immune suppression in breast cancer. *Cancer Res*. (2022) 82:3291–306. doi: 10.1158/0008-5472.CAN-22-1427

35. Wiggs A, Molina S, Sumner SJ, Rushing BR. A review of metabolic targets of anticancer nutrients and nutraceuticals in pre-clinical models of triple-negative breast cancer. *Nutrients*. (2022) 14:1990. doi: 10.3390/nu14101990

36. Huang M, Yu X, Wang Q, Jiang Z, Li X, Chen W, et al. The immune checkpoint TIGIT/CD155 promotes the exhaustion of CD8 + T cells in TNBC through glucose metabolic reprogramming mediated by PI3K/AKT/mTOR signaling. *Cell Commun Signal*. (2024) 22:35. doi: 10.1186/s12964-023-01455-z

37. Huang R, Wang H, Hong J, Wu J, Huang O, He J, et al. Targeting glutamine metabolic reprogramming of SLC7A5 enhances the efficacy of anti-PD-1 in triple-negative breast cancer. *Front Immunol*. (2023) 14:1251643. doi: 10.3389/fimmu.2023.1251643

38. Long Y, Gao Z, Hu X, Xiang F, Wu Z, Zhang J, et al. Downregulation of MCT4 for lactate exchange promotes the cytotoxicity of NK cells in breast carcinoma. *Cancer Med*. (2018) 7:4690–700. doi: 10.1002/cam4.1713

39. Feng J, Yang H, Zhang Y, Wei H, Zhu Z, Zhu B, et al. Tumor cell-derived lactate induces TAZ-dependent upregulation of PD-L1 through GPR81 in human lung cancer cells. *Oncogene*. (2017) 36:5829–39. doi: 10.1038/onc.2017.188

40. Hsu J, Hodgins JJ, Marathe M, Nicolai CJ, Bourgeois-Daigneault MC, Trevino TN, et al. Contribution of NK cells to immunotherapy mediated by PD-1/PD-L1 blockade. *J Clin Invest*. (2018) 128:4654–68. doi: 10.1172/JCI99317

41. Liu Y, Cheng Y, Xu Y, Wang Z, Du X, Li C, et al. Increased expression of programmed cell death protein 1 on NK cells inhibits NK-cell-mediated anti-tumor function and indicates poor prognosis in digestive cancers. *Oncogene*. (2017) 36:6143–53. doi: 10.1038/onc.2017.209

42. Kitai Y, Kawasaki T, Sueyoshi T, Kobiyama K, Ishii KJ, Zou J, et al. DNA-containing exosomes derived from cancer cells treated with topotecan activate a STING-dependent pathway and reinforce antitumor immunity. *J Immunol*. (2017) 198:1649–59. doi: 10.4049/jimmunol.1601694

43. Zhang H, Tang K, Zhang Y, Ma R, Ma J, Li Y, et al. Cell-free Tumor Microparticle Vaccines Stimulate Dendritic Cells via cGAS/STING Signaling. *Cancer Immunol Res*. (2015) 3:196–205. doi: 10.1158/2326-6066.CIR-14-0177

44. Kelly B, O'Neill LA. Metabolic reprogramming in macrophages and dendritic cells in innate immunity. *Cell Res*. (2015) 25:771–84. doi: 10.1038/cr.2015.68

45. Jneid B, Bochnakian A, Hoffmann C, Delisle F, Djacoto E, Sirven P, et al. Selective STING stimulation in dendritic cells primes antitumor T cell responses. *Sci Immunol*. (2023) 8:eabn6612. doi: 10.1126/sciimmunol.abn6612

46. Lv M, Chen M, Zhang R, Zhang W, Wang C, Zhang Y, et al. Manganese is critical for antitumor immune responses via cGAS-STING and improves the efficacy of clinical immunotherapy. *Cell Res*. (2020) 30:966–79. doi: 10.1038/s41422-020-00395-4

47. Li M, Quintana A, Alberts E, Hung MS, Boulat V, Ripoll MM, et al. B cells in breast cancer pathology. *Cancers*. (2023) 15:15173.

48. Gu J, Wang M, Zhang H, Xu X. Enhancing antibody responses through modulation of lipid metabolism in B cells. *Cancer Immunol Res*. (2023) 11:1011–23. doi: 10.1158/2326-6066.CIR-22-0545

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 14 January 2025
DOI 10.3389/fimmu.2024.1537909

OPEN ACCESS

EDITED BY
Yejun Tan,
University of Minnesota Health Twin Cities, United States

REVIEWED BY
Qi Fu,
Nanjing Medical University, China
Huang Chuiguo,
The Chinese University of Hong Kong, China

*CORRESPONDENCE
Junkun Zhan
zhanjunkun@csu.edu.cn

RECEIVED 02 December 2024
ACCEPTED 23 December 2024
PUBLISHED 14 January 2025

CITATION
Li H, Zou L, Long Z and Zhan J (2025) Immunometabolic alterations in type 2 diabetes mellitus revealed by single-cell RNA sequencing: insights into subtypes and therapeutic targets.
Front. Immunol. 15:1537909.
doi: 10.3389/fimmu.2024.1537909

Immunometabolic alterations in type 2 diabetes mellitus revealed by single-cell RNA sequencing: insights into subtypes and therapeutic targets

Huahua Li[1], Lingling Zou[1], Zhaowei Long[2] and Junkun Zhan[3,4]*

[1]Department of Geriatric, Hunan Provincial People's Hospital, The First Affiliated Hospital of Hunan Normal University, Changsha, China, [2]Department of Geriatric, The Third Affiliated Hospital of Kunming Medical University, Kunming, China, [3]Department of Geriatrics, The Second Xiangya Hospital, Central South University, Changsha, China, [4]Institute of Aging and Age-related Disease Research, Central South University, Changsha, China

Background: Type 2 Diabetes Mellitus (T2DM) represents a major global health challenge, marked by chronic hyperglycemia, insulin resistance, and immune system dysfunction. Immune cells, including T cells and monocytes, play a pivotal role in driving systemic inflammation in T2DM; however, the underlying single-cell mechanisms remain inadequately defined.

Methods: Single-cell RNA sequencing of peripheral blood mononuclear cells (PBMCs) from 37 patients with T2DM and 11 healthy controls (HC) was conducted. Immune cell types were identified through clustering analysis, followed by differential expression and pathway analysis. Metabolic heterogeneity within T cell subpopulations was evaluated using Gene Set Variation Analysis (GSVA). Machine learning models were constructed to classify T2DM subtypes based on metabolic signatures, and T-cell-monocyte interactions were explored to assess immune crosstalk. Transcription factor (TF) activity was analyzed, and drug enrichment analysis was performed to identify potential therapeutic targets.

Results: In patients with T2DM, a marked increase in monocytes and a decrease in CD4+ T cells were observed, indicating immune dysregulation. Significant metabolic diversity within T cell subpopulations led to the classification of patients with T2DM into three distinct subtypes (A-C), with HC grouped as D. Enhanced intercellular communication, particularly through the MHC-I pathway, was evident in T2DM subtypes. Machine learning models effectively classified T2DM subtypes based on metabolic signatures, achieving an AUC > 0.84. Analysis of TF activity identified pivotal regulators, including NF-kB, STAT3, and FOXO1, associated with immune and metabolic disturbances in T2DM. Drug enrichment analysis highlighted potential therapeutic agents targeting these TFs and related pathways, including Suloctidil, Chlorpropamide, and other compounds modulating inflammatory and metabolic pathways.

Conclusion: This study underscores significant immunometabolic dysfunction in T2DM, characterized by alterations in immune cell composition, metabolic pathways, and intercellular communication. The identification of critical TFs and the development of drug enrichment profiles highlight the potential for personalized therapeutic strategies, emphasizing the need for integrated immunological and metabolic approaches in T2DM management.

KEYWORDS

type 2 diabetes mellitus (T2DM), single-cell RNA sequencing, immunometabolism, T cells, machine learning models

Introduction

Type 2 Diabetes Mellitus (T2DM) represents a growing global health crisis, with prevalence rates increasing rapidly. The International Diabetes Federation estimates that 537 million adults were living with diabetes in 2021, a number projected to rise to 643 million by 2030 and 783 million by 2045 (1). T2DM accounts for 90-95% of all diabetes cases and is a leading contributor to morbidity and mortality, with associated complications such as cardiovascular disease, neuropathy, nephropathy, and retinopathy (2) The rising incidence of T2DM is driven by a combination of genetic predisposition and lifestyle factors, including obesity, sedentary behavior, and poor dietary habits (3).

Beyond its metabolic consequences, T2DM is increasingly recognized for its significant immunological components, characterized by chronic low-grade inflammation and immune dysregulation (4). Peripheral blood mononuclear cells (PBMCs), including T cells and monocytes, play pivotal roles in the inflammatory processes of T2DM (5). Alterations in immune cell populations have been documented in patients with T2DM, with changes observed in the proportions and functions of various immune cell subsets (6).

T cells and monocytes are particularly implicated in T2DM pathogenesis through their contribution to systemic inflammation and insulin resistance (7). Chronic activation of these immune cells results in the secretion of pro-inflammatory cytokines, which disrupt insulin signaling pathways (8). However, the precise mechanisms by which these immune cells contribute to T2DM, particularly at the single-cell level, remain poorly understood.

Recent advancements in single-cell RNA sequencing (scRNA-seq) have enabled high-resolution analysis of cellular heterogeneity, facilitating the characterization of individual cell types and states within complex tissues (9). This technology offers a unique opportunity to explore the immunological landscapes of PBMCs in T2DM at an unprecedented level of detail. By analyzing gene expression profiles at the single-cell level, it is possible to identify specific cellular subpopulations and uncover new insights into the disease mechanisms.

Metabolic reprogramming of immune cells is a critical aspect of their activation and function (10). In the context of T2DM, metabolic disturbances can influence immune cell behavior, contributing to disease progression (10). Metabolic reprogramming in T cells and monocytes plays a pivotal role in the pathogenesis of T2DM (11). Immune cells, like T cells and monocytes, undergo metabolic shifts in T2DM, which affect their activation and function, thereby exacerbating chronic inflammation and insulin resistance (11). These metabolic alterations can promote the secretion of pro-inflammatory cytokines, further driving disease progression (12). Understanding how metabolic reprogramming influences immune cell behavior could identify novel therapeutic targets for T2DM.

Furthermore, cell-cell communication, mediated by signaling pathways and cytokines, is essential for orchestrating immune responses (13). Dysregulation of these communication networks can intensify inflammation and insulin resistance in T2DM (14). Investigating intercellular signaling dynamics may reveal potential therapeutic targets for modulating immune responses.

In this study, publicly available scRNA-seq data were used to analyze PBMCs from patients with T2DM and healthy controls (HC). This study aimed to characterize the immune cell composition, metabolic heterogeneity, and cell-cell communication networks at the single-cell level. Additionally, advanced machine learning models were employed to classify T2DM subtypes based on metabolic signatures. The findings offer comprehensive insights into the immunometabolic alterations in T2DM, providing a foundation for the development of personalized therapeutic strategies.

Methods

Data collection

The sequencing data used in this study are publicly available from the Gene Expression Omnibus (GEO) database. scRNA-seq data for PBMCs from 11 HC individuals (GSE244515) (15) and 37 patients diagnosed with T2DM (GSE268210) (16) were utilized.

Single-cell RNA sequencing alignment and quality control

All single-cell read counts were analyzed using the Seurat package (v5.0.1) in R (v4.3.1), converting each dataset into individual Seurat objects. Data filtering was performed based on unique molecular identifiers (UMIs) and the number of detected genes (17). Specifically, cells with between 500 and 3,500 detected genes, and those expressed in at least five cells, were retained. Cells exhibiting mitochondrial gene expression greater than 5% were excluded to ensure data quality. Following filtering, data normalization was carried out using Seurat's NormalizeData function, and highly variable genes were identified using the FindVariableFeatures function.

Integration of scRNA-seq data from multiple datasets

To integrate scRNA-seq data from multiple datasets, the Harmony package was employed, focusing on highly variable genes. This integration enabled subsequent dimensionality reduction and clustering analyses, correcting for batch effects and other technical variations across datasets.

Dimensionality reduction and major cell type annotation

For the PBMC dataset, clustering resolution was set to 0.5. Principal component analysis (PCA) was used for dimensionality reduction, followed by Uniform Manifold Approximation and Projection (UMAP) for visualization. Clusters were identified and annotated based on known cell type markers, as shown in Figures 1B, 1E, and 1H

Differential gene expression and pathway analysis

Differential gene expression analysis was conducted using the FindMarkers function of the Seurat package, employing the Wilcoxon rank-sum test. Genes were considered differentially expressed if detected in at least 25% of cells (min.pct = 0.25) and had an adjusted p-value below 0.05 after Bonferroni correction. Significant differentially expressed genes (DEGs) were subjected to Gene Ontology (GO) term and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses using the clusterProfiler package (v3.12.0) (18). Drug enrichment analysis was performed using the Drug-Gene Interaction Database (DGIdb) as the reference, selecting enriched drugs with an adjusted p-value threshold of $P < 0.05$ after multiple testing correction.

Gene set variation analysis

Gene Set Variation Analysis (GSVA) was employed to assess pathway activity across single cells using 42 KEGG pathways as predefined gene sets. The GSVA method was implemented with the GSVA package in R, specifying appropriate gene set indices and kernel-based distribution functions (kcdf). To optimize computational efficiency, parallel processing was utilized, with parameter adjustments based on available processor cores. This approach allowed for scalable analysis, reduced processing time, and preserved result integrity. GSVA provided pathway activity scores for each cell, enabling the exploration of pathway heterogeneity and functional states within the single-cell populations.

Calculation of transcription factor activity

To assess transcription factor (TF) activity, the DoRothEA package was used to retrieve human regulon data, selecting regulons with confidence levels A, B, and C (19). TF activity scores were calculated using the VIPER method, with normalization performed *via* the "scale" method and a minimum regulon size of 4. These scores were stored in the "dorothea" assay of the Seurat object. Dimensionality reduction was conducted using PCA, followed by clustering with the top 10 principal components, and UMAP was applied for cluster visualization. Differential TF activity between clusters was evaluated using Seurat's FindAllMarkers function, with significant TFs identified based on log fold change and expression percentage. The VIPER activity scores were summarized by cell type, and the three most variable TFs across cell types were identified. These TFs were visualized in a heatmap, with color intensities reflecting TF activity.

Unsupervised clustering (consensus clustering)

To classify patients with T2DM based on T cell metabolic patterns, consensus clustering was applied, a robust and reproducible method that aggregates multiple clustering results to enhance stability and reliability using the ConsensusClusterPlus package. Initially, the mean GSVA scores for the 42 pathways were calculated for each sample. Consensus clustering mitigates inherent variability in individual clustering runs by repeatedly subsampling the data and aggregating clustering results, ensuring the identification of consistent and biologically meaningful clusters.

The optimal number of clusters (k) was determined by calculating the incremental area, which measures changes in the cumulative distribution function (CDF) curve area between consecutive k values. The incremental area quantifies improvements in cluster stability as the number of clusters increases. A significant drop in the incremental area suggests that additional clusters contribute minimally to cluster stability, aiding in the selection of the optimal k. Consensus clustering was performed across a range of k values (from $k = 2$ to $k = 9$), and incremental area plots were generated to visualize changes in the CDF curve areas. Using the "elbow method," where the k value at which the incremental area plateaus is selected (indicating diminishing returns from adding more clusters), we identified $k = 4$ as the optimal number. From $k = 4$ onward, the reduction in

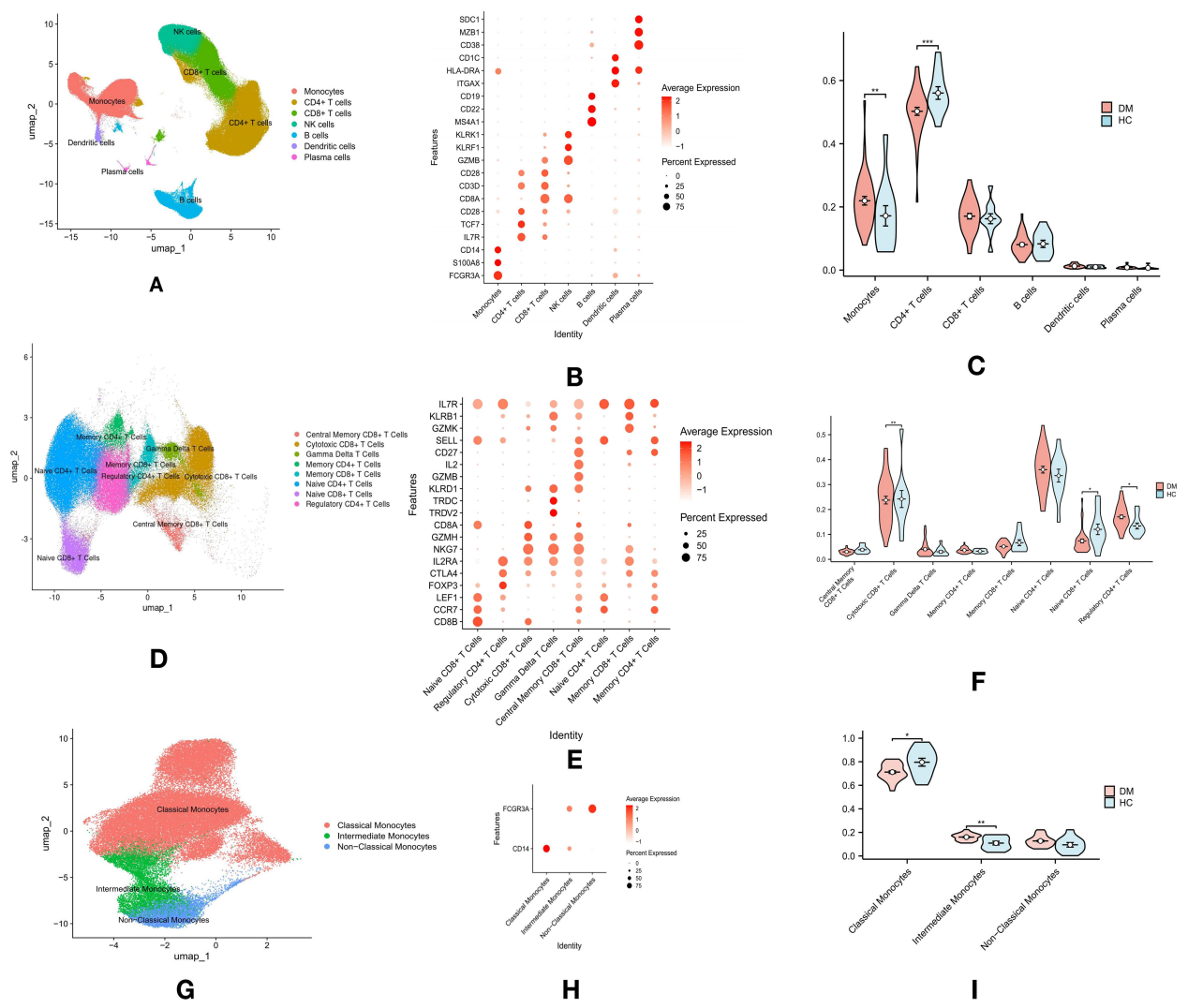


FIGURE 1
Overview of Immune Cell Profiling in T2DM and Healthy Control (HC) Groups: **(A)** Single-cell RNA sequencing and clustering analysis identified seven major immune cell types in PBMCs from both T2DM and HC groups. **(B)** The top three marker genes for each of the seven major immune cell types in PBMCs. **(C)** Violin plots comparing the proportions of these seven immune cell types in PBMCs across T2DM and HC groups. **(D)** Dimensional reduction analysis of T cell clusters, revealing eight distinct T cell subtypes. **(E)** The top three marker genes for each of the eight T cell subtypes. **(F)** Violin plots comparing the proportions of the eight T cell subtypes in PBMCs. **(G)** Monocyte subpopulation analysis identified three distinct subtypes: classical, intermediate, and non-classical monocytes. **(H)** Expression of marker genes (CD14 and FCG3RA) used for classifying monocyte subpopulations. **(I)** Violin plots comparing the proportions of the three monocyte subtypes. *p-values are indicated as follows: *p ≤ 0.05, **p ≤ 0.01, and ***p ≤ 0.001.*

incremental area was significantly less, indicating that k = 4 struck a balance between minimizing the metric and maintaining manageable cluster numbers. Clustering at k = 4 was subsequently visualized using heatmaps and PCA plots.

Cell communication and signaling pathways

Cell communication analysis was performed using the CellChat package in R with default parameters (20). The pathways mediating cell communication between three T cell subtypes and monocytes were analyzed independently, utilizing the human CellChatDB as a reference. The rankNet function was modified to output scaled contribution values for each pathway within each subtype. Differences in the strength of cell communication pathways between the T cell subtypes and monocytes were compared and visualized with bar charts generated by ggplot2. Additionally, specific signaling patterns for each pathway within each subtype were illustrated using the netVisual_bubble function.

Machine learning algorithms

An integrated machine learning model incorporating multiple algorithms was developed to enhance predictive accuracy.

A comprehensive dataset of 196,623 T cells was divided into a training set (70%) and a test set (30%). A total of 75 different combinations of machine learning models were evaluated. Independent predictive models included Support Vector Machines (SVM) and Ridge regression. Boosting methods such as glmBoost, Elastic Net (Enet) with varying alpha values, and Gradient Boosting Machines (GBM) were sequentially applied to correct errors from previous models. Stepwise regression (Stepglm), utilizing forward, backward, or both selection criteria, was combined with models like Ridge, Enet, and Lasso to optimize predictive performance. Additional models, including XGBoost, Linear Discriminant Analysis (LDA), Random Forest (RF), and Naive Bayes, were integrated to leverage the unique strengths of each algorithm in different scenarios.

For multiclass classification adjustments, both one-vs-rest (OvR) and multinomial classification approaches were employed. The OvR strategy decomposes the multiclass problem into multiple binary classifiers, each distinguishing one class from all others. This method was applied to SVM and Logistic Regression algorithms to establish binary decision boundaries within a multiclass framework. Multinomial classification methods, such as GBM and RF, handle all classes simultaneously within a single model, allowing for direct modeling of class probabilities. These algorithms natively support multinomial classification, enabling the simultaneous prediction of multiple classes without decomposing them into separate binary tasks. The choice of methods was guided by the algorithm's native support for multiclass classification and empirical performance during model tuning.

At the patient level, individuals were classified based on the distribution of cell subtypes within their samples. If the majority of a patient's cells were assigned to a specific subtype, the patient was classified into that subtype. This strategy enabled the extension of single-cell classification to predict subtypes at the patient level.

Models were configured to identify the one with the highest average concordance index (C-index) across all validation datasets. The accuracy of the resulting risk scores was validated by calculating the area under the curve (AUC) using the "timeROC" package.

Statistical analysis

All statistical analyses were performed using R software (v4.3.1), and visualizations were generated through R Studio. The selection of statistical tests was determined by the data distribution and characteristics. For normally distributed data, Student's t-test was used to compare means between two groups. For non-normally distributed data, the Wilcoxon rank-sum test was applied for two-group comparisons, and the Kruskal-Wallis test was utilized for comparisons across multiple groups. P-values > 0.05 were considered not statistically significant and were marked as "ns." P-values ≤ 0.05 were considered statistically significant, with the following indications: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, and **** $p \leq 0.0001$.

Results

Significant increase in monocytes and decrease in CD4+ T cells in patients with T2DM

Single-cell sequencing data of PBMCs from 11 HC (GSE244515) and 37 patients with T2DM (GSE268210) were obtained from the GEO database. Cluster analysis revealed seven major immune cell types, annotated by specific marker genes: CD4+ T cells (CD3D, IL7R), CD8+ T cells (CD3D, CD8B), NK cells (KLRF1), B cells (MS4A1), monocytes (CD14, FCG3RA), dendritic cells (ITGAX, CD1C), and plasma cells (SDC1, MZB1) (Figure 1A). Cell types were annotated using established marker genes, a method validated in prior studies (Figure 1B). Rigorous marker selection and clustering methods were applied to ensure accurate and consistent categorization of cell types within the datasets. Proportions of immune cell types between T2DM and HC groups were compared using the Wilcoxon test. The analysis revealed a significant increase in monocyte proportions ($p < 0.01$) and a decrease in CD4+ T cells ($p < 0.01$) in the T2DM group compared to the HC group, while no significant differences were observed in CD8+ T cells, B cells, dendritic cells, or plasma cells (Figure 1C).

Altered proportions of T cell subtypes in patients with T2DM

Dimensional reduction and cluster analysis of T cells based on gene expression profiles identified eight distinct subtypes: Central Memory CD8+ T cells (IL7R, CD27, SELL), Cytotoxic CD8+ T cells (CD8A, GZMH, NKG7), Gamma Delta T cells (TRDC, TRDV2), Memory CD4+ T cells (IL7R, CD27), Memory CD8+ T cells (IL7R, CD27), Naive CD4+ T cells (LEF1, SELL, CCR7), Naive CD8+ T cells (CD8A, LEF1, CCR7), and Regulatory CD4+ T cells (FOXP3) (Figure 1D, E). SELL expression was utilized to distinguish between Central Memory and Memory CD8+ T cells. Differences in T cell subtype proportions between T2DM and HC groups were assessed using the Wilcoxon test. Significant increases in the proportions of Cytotoxic CD8+ T cells ($p < 0.01$) and Naive CD8+ T cells ($p < 0.05$) were observed in the T2DM group, alongside a significant reduction in Regulatory CD4+ T cells ($p < 0.05$). No significant differences were found in Central Memory CD8+ T cells, Gamma Delta T cells, Memory CD4+ T cells, Memory CD8+ T cells, or Naive CD4+ T cells (Figure 1F).

Changes in monocyte subpopulations in patients with T2DM

The interaction between monocytes and T cells plays a critical role in the inflammatory mechanisms driving T2DM progression (21). Monocytes modulate T cell responses and are central to the immune dysregulation observed in T2DM (22). Further analysis of monocytes revealed three subgroups: classical monocytes, non-classical

10.3389/fimmu.2024.1537909

monocytes, and intermediate monocytes (Figure 1G). Classical monocytes were defined by CD14 expression, non-classical by CD16, and intermediate by both CD14 and CD16 (Figure 1H). The Wilcoxon test revealed a significant increase in intermediate monocytes ($p < 0.01$) and a decrease in classical monocytes ($p < 0.05$) in the T2DM group compared to the HC group (Figure 1I).

Metabolic heterogeneity in T cell subpopulations in T2DM

To investigate the metabolic heterogeneity within T cell subpopulations in T2DM, each cell within these subpopulations was scored for 42 metabolic-related pathways from the KEGG database using GSVA. Unsupervised consensus clustering, based on the mean pathway values for each sample, was performed. The optimal number of clusters (k = 4) was determined using the delta area value and the "elbow method," partitioning the samples into four groups (Figure 2A). The clustering heatmap clearly distinguished the samples into four groups, with T2DM samples assigned to groups A-C and HC samples grouped in D (Figure 2B). This segregation was further validated by the PCA plot, which highlighted a distinct separation between group D (HC) and groups A-C (Figure 2D). Specifically, group A included 12 patients, group B included 14 patients, group C included 12 patients, and group D contained 11 HC.

Group A exhibited elevated expression across various metabolic pathways, including sulfur, ether lipid, and sphingolipid metabolism; nicotinate and nicotinamide metabolism; xenobiotic and drug metabolism by cytochrome P450; tryptophan, porphyrin, and chlorophyll metabolism; glycine, serine, and threonine metabolism; linoleic and alpha-linolenic acid metabolism; taurine and hypotaurine metabolism; histidine metabolism; ascorbate and aldarate metabolism; retinol metabolism; arachidonic acid metabolism; and starch and sucrose metabolism, among others (Figure 2C). This broad metabolic profile, encompassing lipid, amino acid, and complex carbohydrate pathways, suggests an adaptive metabolic response in T cells within Group A.

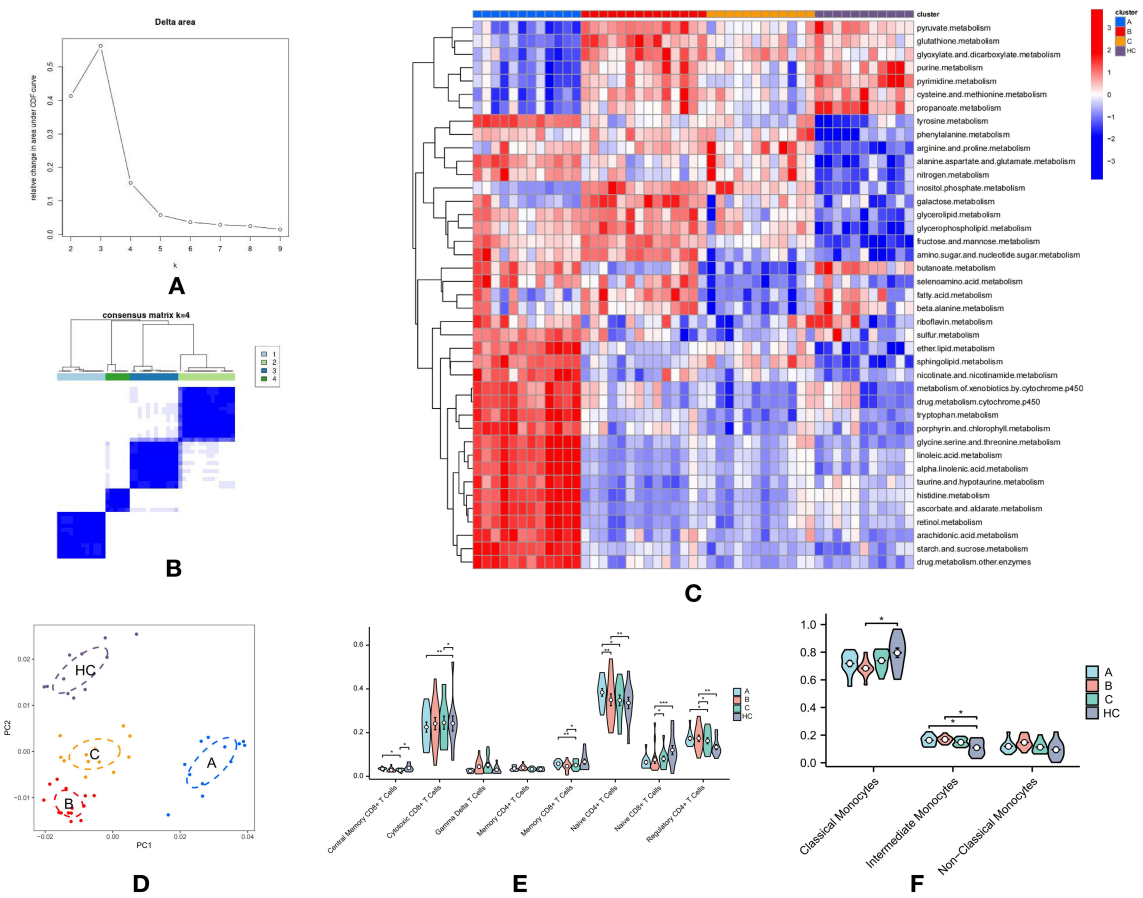


FIGURE 2
Metabolic Heterogeneity in T Cell Subpopulations and Immunological Differences in T2DM: **(A)** Delta area plot showing k values from 2 to 9 used for selecting the optimal k in consensus clustering. **(B)** Consensus clustering heatmap of metabolic pathway scores, dividing samples into four distinct groups. Groups A-C consist of T2DM samples, while Group D represents healthy control (HC) samples. **(C)** Clustering heatmap displaying the expression levels of 42 metabolic-related pathways in T cell subpopulations across the four groups, emphasizing differences in metabolic activity. **(D)** PCA plot demonstrating clear separation between the HC group **(D)** and T2DM groups **(A-C)**. **(E)** Violin plots showing the proportions of eight T cell subtypes across the four groups, with significant differences observed between groups. **(F)** Violin plots illustrating the proportions of three monocyte subtypes across the four groups, highlighting further immune profile differences. *p-values are indicated as follows: *p ≤ 0.05, **p ≤ 0.01, and ***p ≤ 0.001.*

Group B was distinguished by high expression in pathways such as pyruvate, glutathione, glyoxylate and dicarboxylate, purine, pyrimidine, cysteine and methionine, nitrogen, inositol phosphate, galactose, glycerolipid, glycerophospholipid, fructose and mannose, and amino sugar and nucleotide sugar metabolism (Figure 2C). This unique metabolic signature suggests a specific adaptation in Group B, likely reflecting a distinct functional or activation state compared to other groups.

Group C shared a metabolic profile with Group B, marked by high expression in nitrogen, inositol phosphate, galactose, glycerolipid, glycerophospholipid, fructose and mannose, and amino sugar and nucleotide sugar metabolism (Figure 2C). However, the metabolic reprogramming in Group C appeared more targeted or restricted, suggesting a more specific metabolic shift in the T cells.

Group D, representing the HC group, displayed strong expression in pyruvate, glutathione, glyoxylate and dicarboxylate, purine, pyrimidine, cysteine and methionine, propanoate, butanoate, fatty acid, and beta-alanine metabolism (Figure 2C). This metabolic profile aligns with basic cellular metabolism and energy homeostasis, contrasting with the altered metabolic states observed in the T2DM groups.

Immunological differences between T2DM subtypes and HC group

The Kruskal-Wallis test was performed to examine immunological differences in T cell and monocyte subtypes across the groups, revealing significant alterations indicative of substantial immune modulation in T2DM. Notably, Groups A and HC displayed increased proportions of Central Memory CD8+ T Cells, essential for long-term immune memory, suggesting potential immune adaptation or ongoing immune responses. A significant reduction in Cytotoxic CD8+ T Cells was observed in Groups A and C compared to the HC group, indicating an impaired cytotoxic response critical for targeting infected or dysfunctional cells (Figure 2E).

Additionally, a decrease in Memory and Naive CD8+ T Cells in Group C suggests a compromised adaptive immune response, essential for effective long-term immunity. The reduction in Regulatory CD4+ T Cells, especially in Group C, suggests diminished regulatory function, potentially contributing to unchecked immune responses and inflammation characteristic of chronic conditions like T2DM (Figure 2E).

Moreover, a significant reduction in classical monocytes in Group B ($P < 0.05$) was observed, while proportions of intermediate monocytes were significantly increased in Groups A and B ($P < 0.05$) compared to the HC group (Figure 2F).

These findings underscore the intricate interplay between metabolic and immune shifts in T2DM, illustrating how metabolic disturbances may impact immune function and potentially exacerbate the disease. The distinct metabolic profiles observed in T2DM subgroups suggest that targeted metabolic or immunomodulatory therapies could be tailored to address specific dysregulations in these patients.

The communication between T-cells and monocytes in type 2 diabetes

The communication between T-cells and monocytes in T2DM plays a critical role, profoundly influencing immune regulation, inflammation, and autoimmunity, all pivotal in the disease's progression and management. Understanding these interactions offers insights into how immune dysregulation contributes to chronic inflammation and insulin resistance in T2DM.

CellChat was employed to analyze communication differences between T-cells and monocytes across three T2DM subtypes and an HC group (Supplementary Tables 1–4). Regarding the number of inferred interactions, Subtypes A, B, C, and HC had 1418, 1841, 1537, and 1531 interactions, respectively (Figure 3B). Interaction strength values were 0.995 for Subtype A, 1.133 for Subtype B, 0.93 for Subtype C, and 0.793 for HC (Figure 3C). These results highlight variability in communication intensity and complexity across diabetic subtypes compared to HC (Figure 3A), indicating stronger cellular interactions in patients with T2DM, suggesting an enhanced immune response in diabetic conditions.

Intensive pathway mediation in subtype B

In Subtype B of T2DM, multiple pathways actively mediate communication between immune cells. The CD30 pathway facilitates interactions from Naive CD4+ T Cells and Regulatory CD4+ T Cells to Non-Classical Monocytes, which serve as receptors (Figure 3D). This pathway is pivotal as it involves T cells that are essential for maintaining immune tolerance and preventing autoimmune responses while interacting with monocytes that play a central role in inflammation. Activation of this pathway in Subtype B suggests a specific immune regulatory mechanism that could significantly impact the inflammatory environment characteristic of T2DM.

Similarly, the CD48 pathway orchestrates communication between three monocyte subtypes and various T-cell subtypes to Central Memory CD8+ T Cells, also functioning as receptors, and extends this interaction to include Non-Classical Monocytes (Figure 3F). This pathway underscores a robust exchange of signals, enhancing immune memory and responsiveness, which is essential for managing recurrent or chronic antigen exposure in T2DM.

Additionally, the Transforming Growth Factor Beta (TGF-β) pathway mediates interactions from multiple T-cell and monocyte subtypes to Central Memory CD8+ T Cells (Figure 3E). TGF-β, a key cytokine in regulating immune responses, cell growth, and inflammation, suggests a dual role in promoting immune homeostasis and potentially contributing to immune tolerance in T2DM.

The Interferon Type II (IFN-II) pathway is prominently active in Subtypes B and C, facilitating signals from Cytotoxic CD8+ T Cells to Classical and Non-Classical Monocytes, and from Central Memory CD8+ T Cells to Intermediate Monocytes (Figure 3G). The engagement of this pathway highlights an active antiviral and antitumor response, which may be dysregulated in T2DM,

10.3389/fimmu.2024.1537909

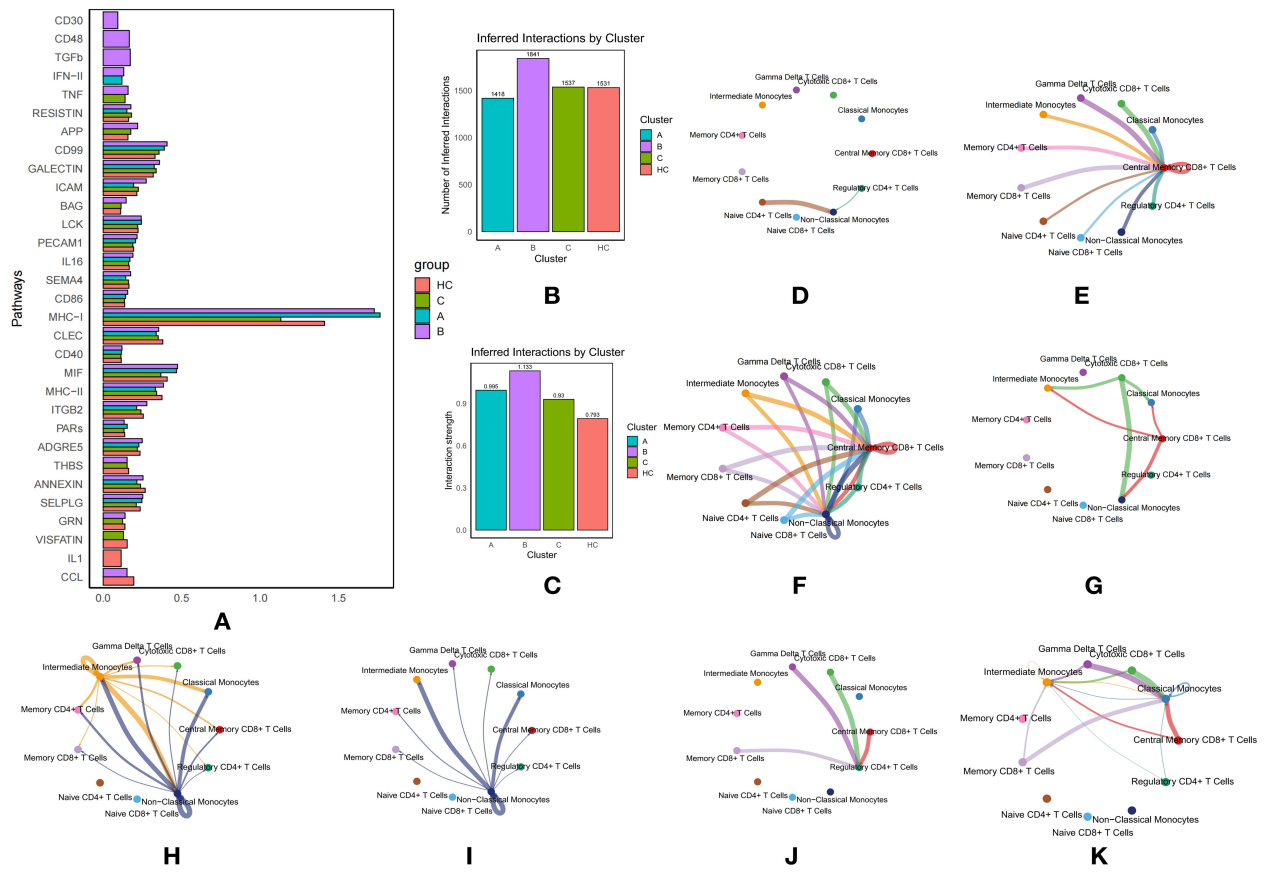


FIGURE 3
Communication Between T Cells and Monocytes in Type 2 Diabetes Mellitus (T2DM): **(A)** Comparison of pathway activity across T2DM subtypes and healthy controls (HC). **(B)** Number of inferred interactions by cluster. **(C)** Interaction strength by cluster. **(D)** Network diagram illustrating key pathway-mediated interactions between T cell and monocyte subtypes for the CD30 pathway. **(E)** Network diagram illustrating key pathway-mediated interactions between T cell and monocyte subtypes for the TGF-β pathway. **(F)** Network diagram illustrating key pathway-mediated interactions between T cell and monocyte subtypes for the CD48 pathway. **(G)** Network diagram illustrating key pathway-mediated interactions between T cell and monocyte subtypes for the IFN-γ pathway. **(H)** Network diagram illustrating TNF pathway interactions in T2DM subtype B. **(I)** Network diagram illustrating TNF pathway interactions in T2DM subtype C. **(J)** Network diagram illustrating CCL pathway interactions in T2DM subtype B. **(K)** Network diagram illustrating CCL pathway interactions in HC.

contributing to altered immune cell activation and cytokine production and influencing disease progression.

The extensive involvement of these pathways in Subtype B reveals a complex and distinct immune modulation pattern that may significantly influence the clinical manifestations and progression of T2DM. The differential activation of these pathways underscores the intricate interplay between immune cells in diabetes, providing a foundation for the development of targeted therapeutic strategies.

TNF and CCL pathway involvement

The Tumor Necrosis Factor (TNF) pathway was particularly active in Subtype B, mediating communication from Intermediate Monocytes to other monocyte and T-cell subtypes (acting as receptors), and from Non-Classical Monocytes to various monocyte and T-cell subtypes (acting as receptors) (Figure 3H). In contrast, in Subtype C, the TNF pathway exclusively mediated communication from Non-Classical Monocytes to other monocyte and T-cell subtypes (as receptors), suggesting its involvement in promoting inflammatory processes that may exacerbate diabetes complications (Figure 3I).

The CCL pathway in Subtype B specifically mediated interactions with Regulatory CD4+ T Cells as receptors and Gamma Delta T Cells, Cytotoxic CD8+ T Cells, Central Memory CD8+ T Cells, and Memory CD8+ T Cells as ligands (Figure 3J). In the HC group, the CCL pathway significantly mediated communication with Classical Monocytes and Intermediate Monocytes as receptors (Figure 3K). The differential involvement of this pathway highlights its potential role in modulating immune responses differently in diabetic patients versus healthy individuals.

MHC-I pathway dominance

The Major Histocompatibility Complex Class I (MHC-I) pathway contributed extensively across all three subtypes, mediating nearly all communication between T-cell and monocyte subtypes (Figure 3A). Subtype A exhibited the highest

activity, followed by HC, with Subtype C showing the least. This underscores the pivotal role of antigen presentation in T2DM, which could influence autoimmune responses and overall immune function in these patients.

VISFATIN pathway specificity

The VISFATIN pathway, uniquely present in Subtypes C and HC, was involved exclusively in mediating communication among T-cell subtypes, without interactions between T-cells and monocytes. This selective engagement suggests a distinct metabolic or inflammatory state inherent to these subtypes and indicates that VISFATIN may play a role in unique disease progression pathways or therapeutic resistance mechanisms in T2DM. The focused activity of VISFATIN offers insights into subtype-specific immune functions, potentially guiding more personalized treatment approaches.

Additional pathways mediated T-cell and monocyte communication

Further analysis revealed additional pathways—GRN, SELPLG, ANNEXIN, THBS, ADGRE5, PARs, ITGB2, MHC-II, MIF, CD40, CLEC, CD86, SEMA4, IL16, PECAM1, LCK, BAG, ICAM, GALECTIN, CD99, APP, and RESISTIN—that mediate communication across various T-cell and monocyte subtypes (Figure 3A). These pathways are involved in a range of regulatory and signaling processes, such as adhesion, immune response modulation, and inflammation. Their involvement across multiple subtypes highlights the complexity and dynamic nature of cellular communication in T2DM, emphasizing the potential for targeted therapeutic interventions based on these specific molecular interactions.

Analysis of transcription factor activity across diabetes subtypes

We also analyzed TF activity across the three T2DM subtypes, identifying 126 active TFs. Key examples include IRF1, GATA6, SPI1, EPAS1, NFKB2, and STAT5B, which are involved in immune response, cell differentiation, and metabolic regulation, all of which are critical in diabetes pathogenesis. A heatmap was generated to visualize the top three TFs for each cell type across the subtypes (Figures 4A–C).

Subtype A: activation of transcription factors in immune cells

In Subtype A, TFs were notably active in Central Memory CD8+ T Cells, Memory CD8+ T Cells, Cytotoxic CD8+ T Cells, and Gamma Delta T Cells, indicating an enhanced immune response (Figure 4A). Of particular interest, HNF4A was uniquely active in Naive CD4+ T Cells and Naive CD8+ T Cells, suggesting a role in early immune cell activation. EPAS1, a key factor involved in oxygen sensing and cellular stress responses, was active in Memory CD8+ T Cells, Cytotoxic CD8+ T Cells, and Gamma Delta T Cells, highlighting its involvement in regulating immune cell function during inflammatory or stress-induced conditions (Figure 4A).

Subtype B: immune modulation and inflammation

In Subtype B, similar TF activity was observed in Central Memory CD8+ T Cells, Memory CD8+ T Cells, Cytotoxic CD8+ T Cells, and Gamma Delta T Cells, with additional unique findings (Figure 4B). NFKB1, known for its role in immune modulation and inflammation, was specifically active in Regulatory CD4+ T Cells, suggesting its contribution to immune tolerance and the prevention of autoimmunity in this subtype. Additionally, SMAD4, a central player in the TGF-β signaling pathway, was active across several T cell types, indicating its role in immune response regulation and tissue remodeling in Subtype B (Figure 4B).

Subtype C: strong immune activation and differentiation

In Subtype C, TFs such as SPI1, STAT4, SMAD1, BCL11A, IKZF1, LYL1, REST, and TBX21 were highly active in Central Memory CD8+ T Cells, suggesting robust immune activation and differentiation (Figure 4C). BCL11A, active in Central Memory CD8+ T Cells, Cytotoxic CD8+ T Cells, and Gamma Delta T Cells, plays a critical role in these cell types. Moreover, BHLHE22, active in Naive CD4+ T Cells and Naive CD8+ T Cells, may regulate early-stage immune responses (Figure 4C). KLF6, active in Central Memory CD8+ T Cells and Memory CD4+ T Cells, likely governs immune cell differentiation and survival. Lastly, TBX21, essential for T cell differentiation and function, was active in Central Memory CD8+ T Cells, underscoring its role in shaping long-term immune responses in this subtype (Figure 4C).

Differential gene expression in subtype A

For Subtype A, further analysis revealed 436 DEGs, highlighting significant involvement in pathways related to microRNA (miRNA) transcription and immune system regulation (Figure 4D). The enrichment of miRNA-related pathways, such as positive regulation of miRNA transcription, regulation of miRNA transcription, and miRNA transcription itself, suggests that miRNAs play a critical role in controlling gene expression that modulates T-cell function and overall immune responses (Figure 4D). This subtype also exhibited significant enrichment in immune-related pathways, including the MAPK signaling pathway, TNF signaling pathway, and Th1/Th2 cell differentiation (Figure 4D). These pathways are pivotal in mediating immune responses and likely contribute to the inflammatory state observed

10.3389/fimmu.2024.1537909

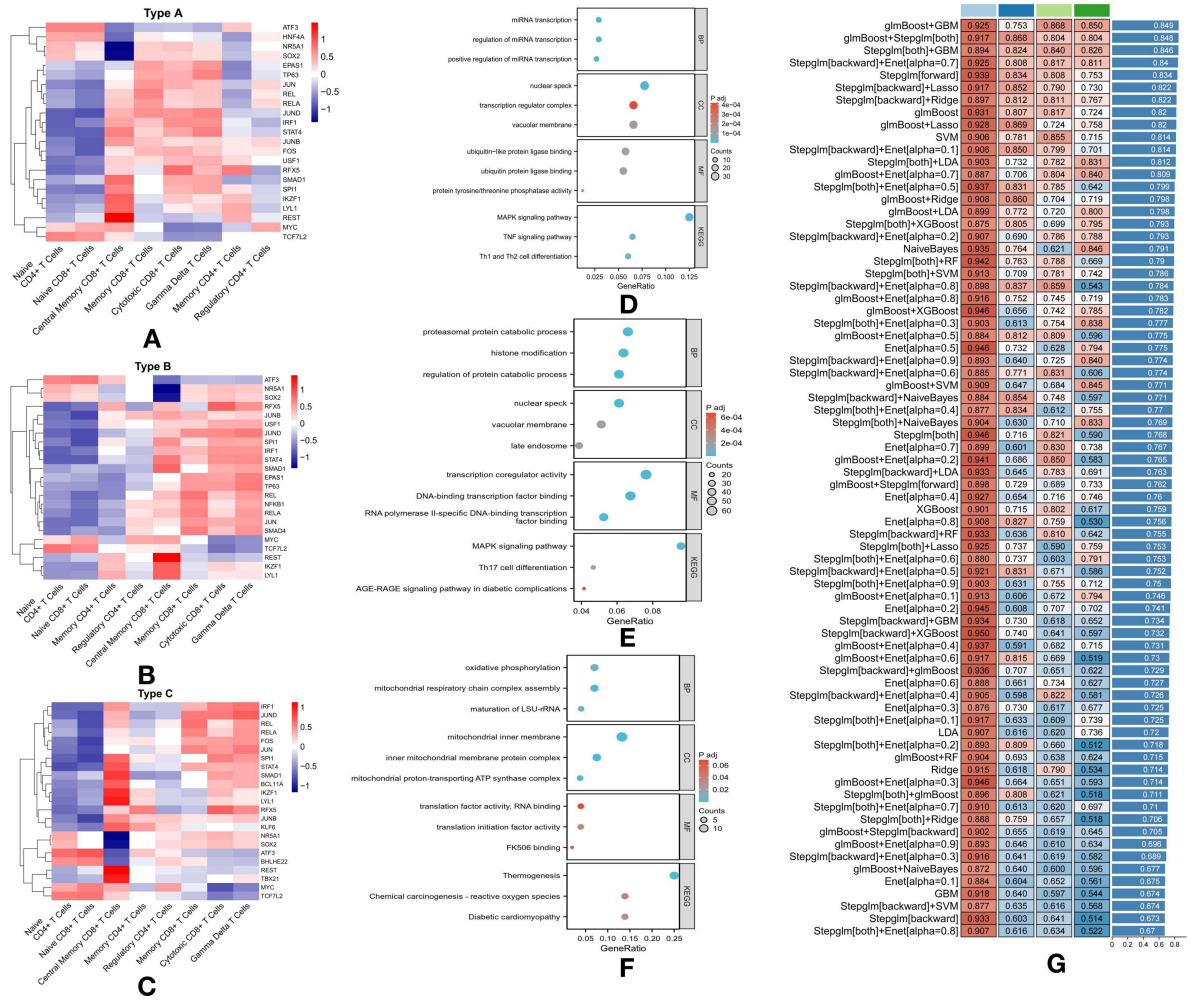


FIGURE 4
Transcription Factor Activation and Pathway Enrichment Analysis in T2DM Subtypes with Machine Learning Classification: **(A)** Heatmap showing logFC of the top three transcription factors in each cell type of Subtype A. **(B)** Heatmap showing logFC of the top three transcription factors in each cell type of Subtype B. **(C)** Heatmap showing logFC of the top three transcription factors in each cell type of Subtype C. **(D)** Pathway enrichment analysis of 436 differentially expressed genes in Subtype A. **(E)** Pathway enrichment analysis of 845 differentially expressed genes in Subtype B. **(F)** Pathway enrichment analysis of 122 differentially expressed genes in Subtype C. **(G)** Performance of various machine learning models for classifying T2DM subtypes based on T-cell metabolic characteristics.

in diabetes and its associated complications in Subtype A. The presence of these pathways highlights the intricate interplay between genetic regulation and immune responses, offering potential insights for developing targeted therapeutic strategies for this subtype.

Differential gene expression in subtype B

Subtype B, distinguished by 845 DEGs, is characterized by a broad range of enriched pathways primarily related to protein metabolism and modification (Figure 4E). Pathways such as the regulation of protein catabolic processes, proteasomal protein catabolism, and histone modification highlight an increased focus on protein turnover and post-translational modifications, both critical for cellular function and signaling. Immune-related pathways, including the MAPK signaling pathway, AGE-RAGE signaling in diabetic complications, and Th17 cell differentiation, are also prominently represented (Figure 4E). The AGE-RAGE pathway is particularly notable for linking metabolic dysregulation to inflammatory responses, a hallmark of diabetes-related complications (Figure 4E). Furthermore, the Th17 differentiation pathway suggests the involvement of a specific T-cell subset known for its role in inflammation and autoimmunity, potentially contributing to the pathophysiological complexity observed in Subtype B (Figure 4E).

Differential gene expression in subtype C

In Subtype C, the 122 DEGs are significantly enriched in pathways related to metabolic processes, with a particular focus on oxidative phosphorylation, a key energy production mechanism in cells (Figure 4F). The inclusion of pathways such as chemical carcinogenesis—reactive oxygen species and diabetic cardiomyopathy

points to an increased susceptibility to oxidative stress and its associated cardiac complications, common challenges in diabetes management (Figure 4F). The prominence of oxidative phosphorylation suggests altered metabolic function that may exacerbate energy deficits in diabetic cells, potentially driving cellular dysfunction and cardiomyopathy progression (Figure 4F). The emphasis on metabolic and oxidative stress pathways in this subtype underscores the importance of metabolic control and highlights potential therapeutic targets for addressing these specific challenges.

Advanced machine learning models for subtype classification

This study further developed machine learning models to differentiate T2DM subtypes based on the metabolic characteristics of T-cells, derived from the GSVA results of KEGG metabolic pathways for each individual cell. A total of 75 model combinations were evaluated, with particular emphasis on high-performing models such as glmBoost+GBM, glmBoost +Stepglm (both combinations), Stepglm+GBM, and Stepglm (backward)+Enet [alpha = 0.7] (Figure 4G).

GlmBoost, or Generalized Linear Model Boosting, enhances prediction accuracy by combining multiple weak models, typically linear, into a stronger predictive ensemble. Stepglm, or Stepwise Generalized Linear Model, refines model accuracy by iteratively adding or removing predictors based on their statistical significance, optimizing the model for maximum performance (Figure 4G). These models demonstrated robust predictive power, achieving AUC values between 0.894 and 0.925 in the training set (Figure 4G). Notably, this high performance extended to the validation set, where all selected models achieved AUC values exceeding 0.8, with an average AUC of over 0.84 across both sets (Figure 4G). The strong accuracy of these models underscores the utility of advanced computational techniques in improving our understanding and management of T2DM, enabling precise subtype classification based on the metabolic profiles of T-cells.

Drug enrichment analyses for personalized treatment

To facilitate the application of the three subtypes of T2DM for personalized treatment, a drug enrichment analysis was conducted on the upregulated DEGs (logFC > 0.5) for each subtype. This approach identifies potential drugs tailored to the specific needs of each T2DM subtype, offering a foundation for more targeted therapeutic strategies.

Subtype A: suloctidil and inflammation pathways

In subtype A, suloctidil emerged as the most promising drug for diabetes treatment (Figure 5A) (23). This drug was linked to genes involved in inflammation and immune regulation, including NR4A2, IFITM1, PPP1R15A, FOSB, TNFAIP3, FOS, ZFP36, MCL1, DUSP1, NFKBIA, JUN, KLF6, KLF2, and FTH1 (Figure 5D). These genes are critical in regulating inflammatory responses, which are central to insulin resistance and diabetes-related complications. The enrichment of suloctidil with these genes suggests its potential in managing the inflammation-associated aspects of T2DM in subtype A.

Subtype B: suloctidil as a key drug

For subtype B, suloctidil was again identified as the key drug associated with diabetes treatment (Figure 5B) (23). This drug was linked to genes such as NR4A2, FOSB, GADD45B, IFITM1, PPP1R15A, TNFAIP3, DDIT4, IER2, ZFP36, FOS, MCL1, NFKBIA, HSPA5, JUN, CD69, DUSP1, KLF2, and FTH1 (Figure 5E), which are involved in stress responses, immune regulation, and cell survival. While suloctidil remained the most relevant drug for this subtype, other medications, such as Fendiline, Prenylamine, and Perhexiline—though primarily used for cardiovascular issues—may have indirect effects on diabetes, but are not specifically designed for its treatment.

Subtype C: chlorpropamide for insulin regulation

In subtype C, chlorpropamide was identified as the key drug associated with diabetes treatment (Figure 5C). As a sulfonylurea, chlorpropamide stimulates insulin secretion, which plays a pivotal role in improving glucose control in patients with T2DM. This drug was associated with genes such as GADD45B, PPP1R15A, TNFAIP3, DDIT4, IER2, ZFP36, FOS, HSPA5, JUN, DUSP1, and KLF2 (Figure 5F), which are involved in stress response and metabolic regulation. These associations suggest that chlorpropamide may be particularly effective in managing insulin secretion and glucose metabolism in subtype C.

Discussion

T2DM is a complex metabolic disorder marked by chronic hyperglycemia resulting from insulin resistance and impaired insulin secretion (24). This study sought to investigate the immunological and metabolic alterations in T2DM by analyzing single-cell RNA sequencing data from PBMCs of patients with T2DM and HC. Our findings highlighted significant immune cell alterations, including an increase in monocytes and a decrease in CD4+ T cells in patients with T2DM. Furthermore, we observed metabolic heterogeneity within T cell subpopulations and enhanced cell-cell communication pathways in T2DM.

The observed increase in monocytes in patients with T2DM reflects heightened chronic inflammation and immune activation (4). These monocytes contribute to insulin resistance by secreting pro-inflammatory cytokines such as TNF-α and IL-6 (8). Previous studies have shown that monocyte-derived macrophages infiltrate adipose tissue in T2DM, where they play a pivotal role in promoting inflammation and exacerbating insulin resistance (25, 26). In contrast, the decrease in CD4+ T cells, which are critical for

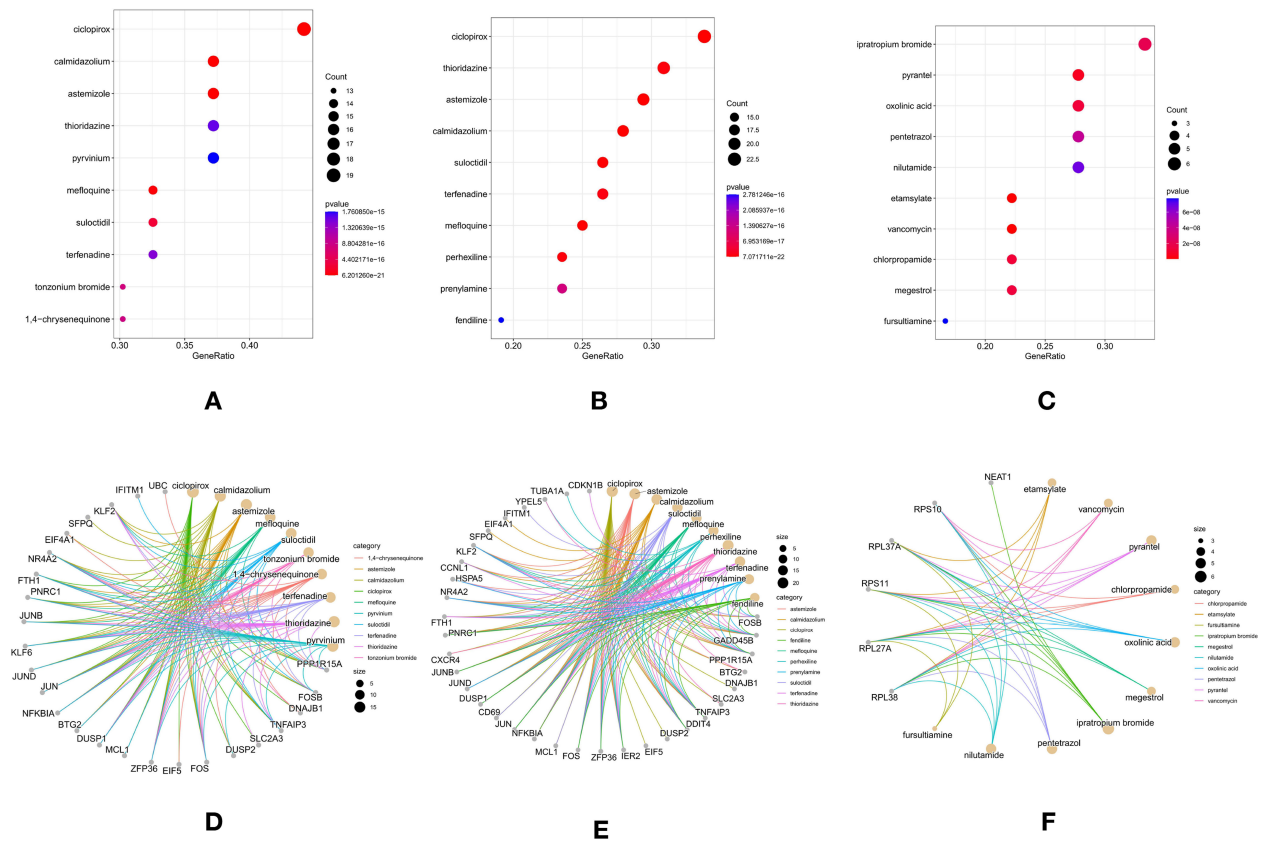


FIGURE 5
Drug Enrichment Analyses in T2DM Subtypes: **(A)** Dot plot showing the top 10 lowest P-value enriched drugs in Subtype A. **(B)** Dot plot showing the top 10 lowest P-value enriched drugs in Subtype B. **(C)** Dot plot showing the top 10 lowest P-value enriched drugs in Subtype C. **(D)** Cnetplot showing the correlated genes with enriched drugs in Subtype A. **(E)** Cnetplot showing the correlated genes with enriched drugs in Subtype B. **(F)** Cnetplot showing the correlated genes with enriched drugs in Subtype C.

coordinating adaptive immune responses, may impair immune regulation (27). This reduction in CD4+ T cells potentially undermines the body's ability to control inflammation, exacerbating insulin resistance and beta-cell dysfunction (28).

Further alterations in T cell subtypes underscore the immune dysregulation associated with T2DM. An increase in cytotoxic CD8 + T cells and naive CD8+ T cells likely reflects an overactive immune surveillance mechanism (29). Elevated cytotoxic CD8+ T cells can induce beta-cell apoptosis, impairing insulin secretion (30). The rise in naive CD8+ T cells indicates ongoing recruitment and activation in response to chronic metabolic stress (31). These changes suggest an altered immune response, compromising the body's ability to regulate inflammation and immune tolerance, thereby contributing to the pathogenesis of T2DM (32, 33). The reduction in CD4+ T cells, which are critical for orchestrating adaptive immune responses, may compromise immune regulation. Previous studies have linked decreased CD4+ T cell counts in patients with T2DM to impaired immune tolerance and increased autoimmunity (34, 35). This decline may result in unregulated inflammatory responses, thereby exacerbating the chronic low-grade inflammation characteristic of T2DM (36).

Analysis of metabolic heterogeneity within T cell subpopulations revealed distinct metabolic profiles in patients with T2DM. Subtype A T cells exhibited high expression of lipid and amino acid metabolism pathways, suggesting an adaptive metabolic response to the diabetic environment (37). This subtype demonstrated broad metabolic activity, particularly in lipid, amino acid, and carbohydrate metabolism, indicating an adaptive response to chronic stress. However, it may also reflect a hyperactivated or exhausted T cell state. Such metabolic reprogramming likely enables T cells to survive in conditions of altered nutrient availability, but it could also promote a pro-inflammatory phenotype (38). Subtype B, with its emphasis on oxidative stress-related pathways, indicates a heightened immune response, while Subtype C displays more targeted metabolic reprogramming, suggesting a potentially less generalized immune activation (39). These metabolic alterations may affect T cell activation and function, potentially exacerbating immune dysfunction in T2DM (40). In contrast, the HC group exhibited baseline metabolic activity, emphasizing the metabolic disturbances present in patients with T2DM. Metabolic reprogramming significantly impacts T cell function (41), and understanding these shifts is critical for developing targeted therapies aimed at restoring normal T cell function and improving metabolic control.

Enhanced cell-cell communication pathways were also observed in patients with T2DM, indicating intensified immune responses. CellChat analysis revealed heightened activity of pathways such as CD30, CD48, TGF-β, and IFN-γ in subtype B (42). These pathways are pivotal in immune regulation, T cell activation, and cytokine

signaling. CD30 activation, for instance, can drive pro-inflammatory responses and immune dysregulation, while TGF-β plays a key role in balancing immune tolerance and inflammation. In the context of T2DM, the activation of these pathways may impair immune function, exacerbate insulin resistance, and contribute to beta-cell dysfunction (43). Furthermore, alterations in TNF and CCL pathway engagement, critical for inflammation and immune cell recruitment, suggest significant changes in chemokine signaling, which could further influence immune cell interactions and disease progression (44).

The dominance of the MHC-I pathway emphasizes the importance of antigen presentation in T2DM (45). Increased antigen presentation may enhance autoimmune responses, potentially contributing to beta-cell destruction (46). The selective activation of the visfatin pathway in specific T2DM subtypes may reflect unique metabolic and inflammatory states, providing potential targets for subtype-specific interventions (47). Focusing on the MHC-I and visfatin pathways could offer targeted therapeutic opportunities for more effective management of T2DM. Collectively, these findings underscore the complexity of immune cell interactions in T2DM and highlight potential pathways for therapeutic targeting.

This study also highlighted the critical role of TFs in regulating immune cell function and metabolic processes in T2DM. We identified several TFs that are differentially expressed across T2DM subtypes, including those involved in immune response regulation and insulin resistance. Notably, TFs such as NF-kB and STAT3, key players in inflammatory pathways, were upregulated in patients with T2DM, highlighting the persistent immune activation and inflammatory environment characteristic of the disease (48). In contrast, TFs associated with insulin signaling, such as PAX6 and FOXO1, were downregulated, potentially contributing to impaired insulin secretion and resistance (49, 50). The dysregulation of TF activity in T2DM thus opens novel therapeutic avenues, as targeting specific TFs could help restore immune homeostasis and improve metabolic control, offering a more tailored approach to treatment.

The drug enrichment analysis further reinforces the potential for personalized T2DM therapy based on TF activity and metabolic alterations. For example, suloctidil, identified as associated with specific immune-related TFs and inflammatory pathways, could serve as a promising candidate for managing inflammation and immune dysfunction in T2DM (8). Similarly, targeting pathways regulated by TFs like NF-kB and STAT3 may help reverse the chronic inflammation that drives T2DM pathogenesis (51, 52). Drugs such as chlorpropamide, which influence insulin secretion, may be especially effective for subtypes with dysregulated insulin signaling pathways (53). These findings emphasize the importance of integrating TF activity and drug enrichment data into personalized treatment strategies, potentially improving therapeutic outcomes by addressing the underlying molecular mechanisms specific to each patient's disease profile.

Clinically, the development of advanced machine learning models enabled accurate classification of T2DM subtypes based on T cell metabolic profiles. These models achieved high AUC values, demonstrating their potential application in clinical settings for patient stratification and personalized treatment planning. By identifying distinct metabolic and immunological signatures linked to different T2DM subtypes, clinicians can better tailor interventions to address the underlying dysfunctions of each patient.

Limitation

Despite the strengths of this study, several limitations remain. These include potential biases arising from the use of public datasets, a limited sample size that may impact generalizability, and the cross-sectional design, which limits causal inference regarding immune changes and T2DM progression. Additionally, findings may not be universally applicable due to demographic variations in T2DM influences. Inherent limitations of single-cell sequencing, such as dropout events and batch effects, may also impact data interpretation. Future research should validate these results using larger, more diverse cohorts, incorporate longitudinal studies to explore disease progression, and evaluate targeted therapies through clinical trials, with predictive models supporting personalized treatment strategies.

Conclusion

In conclusion, this study highlights significant immune and metabolic dysregulation in T2DM, marked by elevated monocytes, reduced CD4+ T cells, and distinct metabolic profiles within T cell subpopulations. Enhanced cell-cell communication pathways, particularly those involving the MHC-I pathway, further highlight the complexity of the immune landscape in T2DM. The analysis of TF activity, in conjunction with drug enrichment findings, identifies promising therapeutic targets for personalized treatment. Integrating these immunological and metabolic insights—along with key TFs and drug candidates—into clinical practice could optimize T2DM management and improve patient outcomes, reinforcing the critical role of personalized medicine in addressing the multifaceted nature of metabolic disorders.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

HL: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. LZ: Writing – original draft, Writing – review & editing. ZL: Writing – original draft, Writing – review & editing. JZ: Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by the Science Research Foundation of Education Bureau of Hunan Province, China(22B0041).

Acknowledgments

Acknowledgments are given to the GEO database for providing the publicly available data utilized in this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1537909/full#supplementary-material

SUPPLEMENTARY TABLE 1
Cell-cell interactions between T cells and monocytes in Subtype A.

SUPPLEMENTARY TABLE 2
Cell-cell interactions between T cells and monocytes in Subtype B.

SUPPLEMENTARY TABLE 3
Cell-cell interactions between T cells and monocytes in Subtype C.

SUPPLEMENTARY TABLE 4
Cell-cell interactions between T cells and monocytes in Healthy Controls.

References

1. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res Clin Pract*. (2022) 183:109119. doi: 10.1016/j.diabres.2021.109119

2. Fan W. Epidemiology in diabetes mellitus and cardiovascular disease. *Cardiovasc Endocrinol*. (2017) 6:8–16. doi: 10.1097/XCE.0000000000000116

3. Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nat Rev Endocrinol*. (2018) 14:88–98. doi: 10.1038/nrendo.2017.151

4. Donath MY, Shoelson SE. Type 2 diabetes as an inflammatory disease. *Nat Rev Immunol*. (2011) 11:98–107. doi: 10.1038/nri2925

5. Eguchi K, Nagai R. Islet inflammation in type 2 diabetes and physiology. *J Clin Invest*. (2017) 127:14–23. doi: 10.1172/JCI88877

6. Jagannathan-Bogdan M, McDonnell ME, Shin H, Rehman Q, Hasturk H, Apovian CM, et al. Elevated proinflammatory cytokine production by a skewed T cell compartment requires monocytes and promotes inflammation in type 2 diabetes. *J Immunol*. (2011) 186:1162–72. doi: 10.4049/jimmunol.1002615

7. Kolb H, Mandrup-Poulsen T. The global diabetes epidemic as a consequence of lifestyle-induced low-grade inflammation. *Diabetologia*. (2010) 53:10–20. doi: 10.1007/s00125-009-1573-7

8. Hotamisligil GS. Inflammation and metabolic disorders. *Nature*. (2006) 444:860–7. doi: 10.1038/nature05485

9. Stuart T, Satija R. Integrative single-cell analysis. *Nat Rev Genet*. (2019) 20:257–72. doi: 10.1038/s41576-019-0093-7

10. O'Neill LA, Kishton RJ, Rathmell J. A guide to immunometabolism for immunologists. *Nat Rev Immunol*. (2016) 16:553–65. doi: 10.1038/nri.2016.70

11. Sun L, Yang X, Yuan Z, Wang H. Metabolic reprogramming in immune response and tissue inflammation. *Arterioscler Thromb Vasc Biol*. (2020) 40:1990–2001. doi: 10.1161/ATVBAHA.120.314037

12. Ferreira AV, Domínguez-Andrés J, Merlo Pich LM, Joosten L, Netea MG. Metabolic regulation in the induction of trained immunity. *Semin Immunopathol*. (2024) 46:7. doi: 10.1007/s00281-024-01015-8

13. Armingol E, Officer A, Harismendy O, Lewis NE. Deciphering cell-cell interactions and communication from gene expression. *Nat Rev Genet*. (2021) 22:71–88. doi: 10.1038/s41576-020-00292-x

14. Wensveen FM, Valentić S, Šestan M, Turk Wensveen T, Polić B. The "Big Bang" in obese fat: Events initiating obesity-induced adipose tissue inflammation. *Eur J Immunol*. (2015) 45:2446–56. doi: 10.1002/eji.201545502

15. Kang J, Lee H, Joo JY, Song JM, Kim HJ, Kim YH, et al. Comparison of genetic and epigenetic profiles of periodontitis according to the presence of type 2 diabetes. *MedComm*. (2020) 5:e620. doi: 10.1002/mco2.620

16. Gu D, Lim J, Han KY, Seo IH, Jee JH, Cho SJ, et al. Single-cell analysis of human PBMCs in healthy and type 2 diabetes populations: dysregulated immune networks in type 2 diabetes unveiled through single-cell profiling. *Front Endocrinol (Lausanne)*. (2024) 15:1397661. doi: 10.3389/fendo.2024.1397661

17. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive integration of single-cell data. *Cell*. (2019) 177:1888–1902.e21. doi: 10.1016/j.cell.2019.05.031

18. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*. (2021) 2:100141. doi: 10.1016/j.xinn.2021.100141

19. Garcia-Alonso L, Holland CH, Ibrahim MM, Turei D, Saez-Rodriguez J. Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome Res*. (2019) 29:1363–75. doi: 10.1101/gr.240663.118

20. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. (2021) 12:1088. doi: 10.1038/s41467-021-21246-9

21. Yan WJ, Sun P, Wei DD, Wang SX, Yang JJ, Li YH, et al. T cell immunoglobulin and mucin domain-containing molecule 3 on CD14(+) monocytes serves as a novel biological marker for diabetes duration in type 2 diabetes mellitus. *J Diabetes Investig*. (2016) 7:867–73. doi: 10.1111/jdi.2016.7.issue-6

22. Girard D, Vandiedonck C. How dysregulation of the immune system promotes diabetes mellitus and cardiovascular risk complications. *Front Cardiovasc Med*. (2022) 9:991716. doi: 10.3389/fcvm.2022.991716

23. Marcheva B, Weidemann BJ, Taguchi A, Perelis M, Ramsey KM, Newman MV, et al. P2Y1 purinergic receptor identified as a diabetes target in a small-molecule screen to reverse circadian β-cell failure. *Elife*. (2022) 11:e75132. doi: 10.7554/eLife.75132.sa2

24. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. (2014) 37 Suppl 1:S81–90. doi: 10.2337/dc14-S081

25. Lumeng CN, Saltiel AR. Inflammatory links between obesity and metabolic disease. *J Clin Invest*. (2011) 121:2111–7. doi: 10.1172/JCI57132

26. Weisberg SP, McCann D, Desai M, Rosenbaum M, Leibel RL, Ferrante AW Jr. Obesity is associated with macrophage accumulation in adipose tissue. *J Clin Invest*. (2003) 112:1796–808. doi: 10.1172/JCI200319246

27. Palmer MT, Weaver CT. Autoimmunity: increasing suspects in the CD4+ T cell lineup. *Nat Immunol*. (2010) 11:36–40. doi: 10.1038/ni.1802

28. Winer DA, Winer S, Shen L, Wadia PP, Yantha J, Paltser G, et al. B cells promote insulin resistance through modulation of T cells and production of pathogenic IgG antibodies. *Nat Med*. (2011) 17:610–7. doi: 10.1038/nm.2353

29. Shimabukuro-Vornhagen A, Gödel P, Subklewe M, Stemmler HJ, Schlößer HA, Schlaak M, et al. Cytokine release syndrome. *J Immunother Cancer*. (2018) 6:56. doi: 10.1186/s40425-018-0343-9

30. Mathis D, Shoelson SE. Immunometabolism: an emerging frontier. *Nat Rev Immunol*. (2011) 11:81. doi: 10.1038/nri2922

31. Bluestone JA, Trotta E, Xu D. The therapeutic potential of regulatory T cells for the treatment of autoimmune disease. *Expert Opin Ther Targets*. (2015) 19:1091–103. doi: 10.1517/14728222.2015.1037282

32. Rohm TV, Meier DT, Olefsky JM, Donath MY. Inflammation in obesity, diabetes, and related disorders. *Immunity*. (2022) 55:31–55. doi: 10.1016/j.immuni.2021.12.013

33. Tsalamandris S, Antonopoulos AS, Oikonomou E, Papamikroulis GA, Vogiatzi G, Papaioannou S, et al. The role of inflammation in diabetes: current concepts and future perspectives. *Eur Cardiol*. (2019) 14:50–9. doi: 10.15420/ecr

34. Zhou T, Hu Z, Yang S, Sun L, Yu Z, Wang G. Role of adaptive and innate immunity in type 2 diabetes mellitus. *J Diabetes Res*. (2018) 2018:7457269. doi: 10.1155/2018/7457269

35. Xia C, Rao X, Zhong J. Role of T lymphocytes in type 2 diabetes and diabetes-associated inflammation. *J Diabetes Res*. (2017) 2017:6494795. doi: 10.1155/2017/6494795

36. Zeng C, Shi X, Zhang B, Liu H, Zhang L, Ding W, et al. The imbalance of Th17/Th1/Tregs in patients with type 2 diabetes: relationship with metabolic factors and complications. *J Mol Med (Berl)*. (2012) 90:175–86. doi: 10.1007/s00109-011-0816-5

37. Pearce EL, Pearce EJ. Metabolic pathways in immune cell activation and quiescence. *Immunity*. (2013) 38:633–43. doi: 10.1016/j.immuni.2013.04.005

38. MacIver NJ, Michalek RD, Rathmell JC. Metabolic regulation of T lymphocytes. *Annu Rev Immunol*. (2013) 31:259–83. doi: 10.1146/annurev-immunol-032712-095956

39. Wang R, Green DR. Metabolic checkpoints in activated T cells. *Nat Immunol*. (2012) 13:907–15. doi: 10.1038/ni.2386

40. Ganeshan K, Chawla A. Metabolic regulation of immune responses. *Annu Rev Immunol*. (2014) 32:609–34. doi: 10.1146/annurev-immunol-032713-120236

41. Buck MD, Sowell RT, Kaech SM, Pearce EL. Metabolic instruction of immunity. *Cell*. (2017) 169:570–86. doi: 10.1016/j.cell.2017.04.004

42. Charo IF, Ransohoff RM. The many roles of chemokines and chemokine receptors in inflammation. *N Engl J Med*. (2006) 354:610–21. doi: 10.1056/NEJMra052723

43. Wellen KE, Hotamisligil GS. Inflammation, stress, and diabetes. *J Clin Invest*. (2005) 115:1111–9. doi: 10.1172/JCI25102

44. Mollica Poeta V, Massara M, Capucetti A, Bonecchi R. Chemokines and chemokine receptors: new targets for cancer immunotherapy. *Front Immunol*. (2019) 10:379. doi: 10.3389/fimmu.2019.00379

45. Bru-Tari E, Oropeza D, Herrera PL. Cell heterogeneity and paracrine interactions in human islet function: A perspective focused in β-cell regeneration strategies. *Front Endocrinol (Lausanne)*. (2020) 11:619150. doi: 10.3389/fendo.2020.619150

46. Arif S, Moore F, Marks K, Bouckenooghe T, Dayan CM, Planas R, et al. Peripheral and islet interleukin-17 pathway activation characterizes human autoimmune diabetes and promotes cytokine-mediated β-cell death. *Diabetes*. (2011) 60:2112–9. doi: 10.2337/db10-1643

47. Dakroub A, Nasser S A, Younis N, Bouckenooghe T, Dayan CM, Planas R, et al. Visfatin: A possible role in cardiovasculo-metabolic disorders. *Cells*. (2020) 9:2444. doi: 10.3390/cells9112444

48. Lu S, Li Y, Qian Z, Zhao T, Feng Z, Weng X, et al. Role of the inflammasome in insulin resistance and type 2 diabetes mellitus. *Front Immunol*. (2023) 14:1052756. doi: 10.3389/fimmu.2023.1052756

49. Teaney NA, Cyr NE. FoxO1 as a tissue-specific therapeutic target for type 2 diabetes. *Front Endocrinol (Lausanne)*. (2023) 14:1286838. doi: 10.3389/fendo.2023.1286838

50. Gosmain Y, Katz LS, Masson MH, Cheyssac C, Poisson C, Philippe J. Pax6 is crucial for β-cell function, insulin biosynthesis, and glucose-induced insulin secretion. *Mol Endocrinol*. (2012) 26:696–709. doi: 10.1210/me.2011-1256

51. Xu G, Zhao Y, Bai Y, Lin Y. Study of hub nodes of transcription factor-target gene regulatory network and immune mechanism for type 2 diabetes based on chip analysis of GEO database. *Front Mol Biosci*. (2024) 11:1410004. doi: 10.3389/fmolb.2024.1410004

52. Saengboonmee C, Phoomak C, Supabphol S, et al. NF-κB and STAT3 co-operation enhances high glucose induced aggressiveness of cholangiocarcinoma cells. *Life Sci*. (2020) 262:118548. doi: 10.1016/j.lfs.2020.118548

53. Hecht A, Gershberg H, Hulse M. Effect of chlorpropamide treatment on insulin secretion in diabetics: its relationship to the hypoglycemic effect. *Metabolism*. (1973) 22:723–33. doi: 10.1016/0026-0495(73)90244-8

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 14 January 2025
DOI 10.3389/fimmu.2024.1512519

OPEN ACCESS

EDITED BY
Yafeng Zhu,
Sun Yat-sen Memorial Hospital, China

REVIEWED BY
Ying-Yong Zhao,
Northwest University, China
Xin Chen,
University of Macau, China

*CORRESPONDENCE
Yueming Luo
yuemingluo@163.com
Shunmin Li
zyylishunmin@126.com
Hong Cheng
bchlhy@163.com

†These authors share first authorship

RECEIVED 16 October 2024
ACCEPTED 19 December 2024
PUBLISHED 14 January 2025

CITATION
Li Y, Luo Y, Hu Y, Li S, Li G, Zhang W, Gu X, Wang J, Li S and Cheng H (2025) Network pharmacology and multi-omics validation of the Jianpi-Yishen formula in the treatment of chronic kidney disease.
Front. Immunol. 15:1512519.
doi: 10.3389/fimmu.2024.1512519

Network pharmacology and multi-omics validation of the Jianpi-Yishen formula in the treatment of chronic kidney disease

Yuyan Li[1†], Yueming Luo[2*†], Yilan Hu[3], Siting Li[4], Guandong Li[3], Wanyangchuan Zhang[5], Xiufen Gu[1], Jianting Wang[1], Shunmin Li[1*] and Hong Cheng[2*]

[1]Department of Nephrology, Shenzhen Traditional Chinese Medicine Hospital, The Fourth Clinical Medical College of Guangzhou University of Chinese Medicine, Shenzhen, China, [2]Department of Geriatrics, Shenzhen Traditional Chinese Medicine Hospital, The Fourth Clinical Medical College of Guangzhou University of Chinese Medicine, Shenzhen, China, [3]The Fourth Clinical Medical College, Guangzhou University of Chinese Medicine, Shenzhen, China, [4]Beijing Tongrentang Hospital of Traditional Chinese Medicine, Beijing, China, [5]Department of Minimally Invasive Intervention and Vascular Surgery, Chongqing Red Cross Hospital (People's Hospital of Jiangbei District), Chongqing, China

Objective: Chronic kidney disease (CKD) is a major global health problem. In clinical practice, the Chinese patent herbal medicine Jianpi-Yishen (JPYS) formula is commonly used to treat CKD. However, the molecular mechanisms by which JPYS targets and modulates the host immune response remain unclear.

Methods: This study utilized network pharmacology, RNA sequencing (RNA-seq), and metabolic analyses using *in vivo* and *in vitro* models to investigate the impact of the JPYS formula on inflammation and the immune system. Specifically, the study focused on macrophage polarization and metabolic changes that may slow down the progression of CKD.

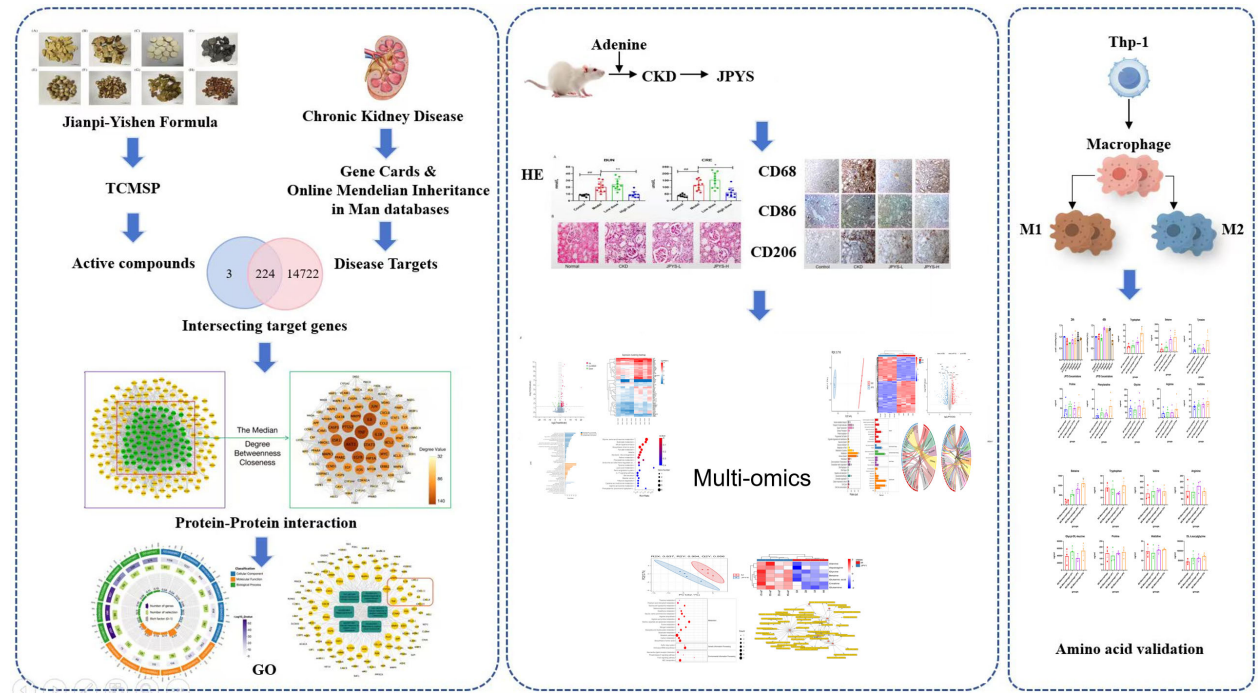
Results: A total of 14,946 CKD-related targets were identified from the GeneCards and Online Mendelian Inheritance in Man (OMIM) databases through network pharmacology analyses. 227 potential targets of the JPYS formula were predicted using the TCMSP database. Additionally, network diagram demonstrated that 11 targets were associated with macrophage activity. *In vivo* studies indicated that the JPYS formula could reduce blood urea nitrogen and serum creatinine in adenine-induced CKD rats. Furthermore, the formula inhibited inflammatory damage and abnormal macrophage infiltration in this CKD model. RNA-seq, proteomic and metabolic analyses identified the regulation of amino acid metabolism by betaine, specifically referring to glycine, serine, and threonine metabolism, as a key target of the JPYS formula in slowing the progression of CKD. In addition, *in vitro* studies suggested that JPYS may enhance tryptophan metabolism in M1 macrophage polarization and betaine metabolism in M2 macrophage polarization.

10.3389/fimmu.2024.1512519

Conclusions: The JPYS formula has been shown to have beneficial impact on CKD; a key mechanism is the mitigation of inflammatory damage through the interaction between amino acid metabolism and macrophage polarization. Of specific importance in this context are the roles of tryptophan in M1 polarization and betaine in M2 polarization.

KEYWORDS

Jianpi-Yishen formula, chronic kidney disease, network pharmacology, macrophage polarization, multi-omics



GRAPHICAL ABSTRACT

1 Introduction

Chronic kidney disease (CKD) is characterized by long-lasting abnormalities in renal structure or function. These abnormalities last longer than three months and have serious health repercussions (1). The global prevalence of CKD is estimated to be approximately 10–14% (1). Furthermore, CKD accelerates the aging process and promotes the development of end-stage renal disease. This leads to increased disability, decreased life expectancy, and a high annual mortality rate, all of which are important contributors to the worldwide burden of disease (2, 3). However, there are no specific treatment modalities available that can entirely arrest the progression of CKD, and coping with CKD poses challenges for both patients and their caregivers (4).

Chinese herbal medicines (CHMs) are characterized by their intricate chemical compositions, which complicates the process of identifying the specific constituents that collectively contribute to the therapeutic effects of these herbal remedies, as these are typically applied in a multi-ingredient manner (5). Previous studies have demonstrated that substances originating from Chinese herbal remedies can ameliorate CKD via multiple molecular pathways (6–9). In addition, certain formulations, such as the Sanqi oral solution and the Bupi Yishen formula, have demonstrated positive impact on kidney function (10, 11). Herbal medicine also has numerous advantages over chemical agents in the management of CKD, not least because of its diverse ingredients (12). The JPYS formula, which translates to "strengthen the spleen and kidney", is a patented traditional Chinese medicine (TCM) formulation developed by

Professor Li Shunmin, a distinguished physician of traditional Chinese medicine in Guangdong Province, drawing upon decades of comprehensive clinical experience. In previous research, the effectiveness and safety of the JPYS formula in CKD patients have been investigated (13). More recently, there are randomized controlled trials being conducted to further explore its clinical application. Previous studies have identified multi targeted effects of the JPYS formula in slowing the progression of CKD, including anti-inflammatory properties, protection against iron deficiency anemia (14), inhibition of mitochondrial fission, promotion of mitochondrial fusion, and suppression of oxidative stress, among others (15). Notably, the anti-inflammatory effects of the JPYS formula have been recognized as therapeutically significant (16) . However, the specific molecular mechanisms through which the JPYS formula targets the endogenous immune response remain unclear.

Macrophages play essential roles in immune surveillance and the maintenance of kidney homeostasis (17). Throughout the progression of CKD, macrophage polarization has been implicated in the mechanisms of inflammatory injury, renal interstitial fibrosis, and kidney repair (18). Various stimuli can influence the functional phenotype of macrophages, leading to the differentiation towards classically activated macrophages (M1) or alternatively activated macrophages (M2). M1 macrophages are recognized as pro-inflammatory cells that contribute to the progression of kidney injury, whereas M2 macrophages are traditionally known as anti-inflammatory cells. Some TCM therapies, encompassing both formulated remedies and herbal active ingredients, have demonstrated efficacy in modulating macrophage polarization towards either M1 or M2 phenotypes in kidney disease (19, 20). As such, the immunoregulatory properties of herbal medicine, particularly its anti-inflammatory effects, present a novel approach for the management of kidney diseases (21).

In this study, network pharmacology, RNA sequencing (RNA-seq), proteomics and metabolic analyses were employed to examine the impact of the JPYS formula on inflammation, the immune response, macrophage polarization, and metabolic changes that may inhibit the progression of CKD. The findings reveal that the interaction between amino acid metabolism and macrophage polarization serves as a key mechanism through which the JPYS formula mitigates inflammatory injury in CKD.

2 Materials and methods

2.1 Network pharmacology

2.1.1 Screening and prediction of effective chemical constituents in the JPYS formula

The active chemical constituents of the JPYS formula (*Astragali Radix, Atractylodis Macrocephalae Rhizoma, Dioscoreae Rhizoma, Cistanches Herba, Amomi Fructus Rotundus, Salviae Miltiorrhizae Radix et Rhizoma, Rhei Radix et Rhizoma, Glycyrrhizae Radix et Rhizoma*) were obtained from the Traditional Chinese Medicine System Pharmacology Database (old.tcmsp-e.com/tcmsp.php, updated until September 2023) (22). The criteria of oral bioavailability (OB) ≥ 30% and drug likeness (DL) ≥ 0.18 were applied to assess the active ingredients of the JPYS formula and identify the pertinent effective active ingredients. The active ingredients of the JPYS formula were then converted into the corresponding human gene names using the Uniprot database. By utilizing "chronic kidney disease" as the keyword, the GeneCards and Online Mendelian Inheritance in Man databases were utilized to retrieve genes associated with CKD. The identified CKD-related genes and drug target genes were mapped to identify the common target genes of the "JPYS formula-CKD".

2.1.2 "Drug-Ingredient-Target-Disease" visualization network construction

The active components of the JYPS formula and the common target genes of the "JYPS formula–CKD" were imported into Cytoscape 3.7.1 software for visualization. Subsequently, a network diagram of "drug–ingredient–target–disease" was then established. Each node in the diagram symbolized a disease, drug, bioactive ingredient of a drug, or target, with the connections between nodes indicating the interrelations among the disease, drug, bioactive ingredient, and target.

2.1.3 Protein-protein interaction network and, gene ontology functional analyses

The shared targets of the JPYS formula and CKD were entered into the STRING database (https://string-db.org/) using specific parameters to extract the PPI network. The analysis focused on the human species (Homo sapiens) with a protein relationship score threshold of 0.4. The presence of free proteins was concealed to obtain the protein interaction network. The protein-protein interaction network data was downloaded and imported into Cytoscape 3.7.1 software. Utilizing the Network Analyzer tool, a topological analysis was conducted on the relevant parameters of drug-disease common targets, which included connectivity (Degree), betweenness centrality, and closeness centrality. Targets exceeding the median values of the aforementioned parameters were designated as core targets.

The Gene Ontology Biological Process (BP), Molecular Function (MF), and Cellular Component enrichment analysis data, were obtained from the STRING database. The GO analysis conditions were set to include observed gene count and strength both greater than the median. Subsequently, the top 6 significantly enriched items in BP, MF and CC were selected and import into ChiPlot (https://chiplot.online/) to generate a circular enrichment plot. Additionally, the top 6 significantly enriched items in GO-BP and their associated targets were imported into Cytoscape 3.7.1 software for visualization processing, resulting in the creation of a network diagram titled "BP Entry - Target."

2.2 JPYS formula preparation

Medicinal herbs for JPYS formula were gained from the Pharmaceutical Department of Shenzhen Traditional Chinese Medicine Hospital. The JPYS formula comprises the following eight herbs: *Astragali Radix, Atractylodis Macrocephalae Rhizoma, Dioscoreae Rhizoma, Cistanches Herba, Amomi Fructus Rotundus, Salviae Miltiorrhizae Radix et Rhizoma, Rhei Radix et Rhizoma, Glycyrrhizae Radix et Rhizoma.* These raw herbs were weighed and boiled twice for 1 h each time in 8 times of water. Our earlier research detailed the preparation and quality control of JPYS formula extract (15).

2.3 Animals and experimental treatments

Male Sprague Dawley rats(ethics approval reference number: TOP-IACUC-2021-0112) aged 6–8 weeks were randomly assigned to one of four groups: control (n = 8), CKD (n = 10), CKD + JPYS-L (n = 10), and CKD + JPYS-H (n = 10). Rats in the CKD and CKD + JPYS were fed a diet containing 0.75% adenine for 3 weeks, followed by a normal diet for 1 week. Rats in the control group were fed a normal diet for 4 weeks. The CKD + JPYS groups were administered with 5.44 g/kg/day of JPYS extract (CKD + JPYS-L, low-dose group) and 10.89 g/kg/day of JPYS extract (JPYS-H, high-dose group) via gastric irrigation for 4 weeks during the study period.

2.4 Biochemical analysis

Serum creatinine and urea nitrogen levels were measured using a Roche automatic biochemistry analyzer (Tokyo, Japan) in accordance with the manufacturer's instructions.

2.5 Histological analysis and immunohistochemistry

Paraffin-embedded kidney tissues extracted from four groups of rats were cut into 3-μm sections, dewaxed, and rehydrated. Sections were stained with hematoxylin and eosin (H&E) stain and visualized. Immunohistochemistry was performed according to the established protocol as described previously (23). Antibodies used are in Supplementary Table 1.

2.6 RNA-seq

The kidney samples from the CKD and CKD+JPYS-H groups underwent analysis at the Beijing Genomics Institute (BGI, Shenzhen, China). The samples were purified and amplified through polymerase chain reaction (PCR). The PCR yield was quantified using Qubit, and the samples were combined to produce a single-stranded DNA circle (ssDNA circle) which generated DNA nanoballs. These nanoballs were then loaded into patterned nanoarrays. Subsequent data analysis was conducted using the BGISEQ500 platform.

2.7 Kidney proteomics analysis

Proteomics analysis was conducted at the Climb Technology Co., Ltd. Briefly, kidney tissue were thoroughly lysed using a protein lysis buffer, followed by the measurement of protein concentration. Based on the results of these measurements, an appropriate volume of protein was extracted from each sample for enzymatic hydrolysis. The subsequent day, the ultrafiltration tubes were centrifuged at 13000g at room temperature for 10 minutes. The liquid collected in the collectiong tube was then transferred to a new centrifuge tube and subjected to vacuum drying. The samples were subsequently desalinated utilizing a C18 desalination column.

For mass spectrometry detection, the chromatographic mobile phase A consisted of 0.1% formic acid, while phase B comprised 80% acetonitrile and 0.1% formic acid. The freeze-dried peptide segments were completely dissolved in solution A (0.1% formic acid) and centrifuged at 17,000 g for 15 minutes. The supernatant was then added to the built-in tube and placed in the automatic sampling device. The sample was introduced into the C18 analytical column (inner diameter 150 μm, 25 cm) from the automatic sampler at a flow rate of 1.2 μL/min using the EASY nLC 1200 liquid chromatography system (Thermo, USA) for elution. The elution conditions for the liquid chromatography were set at a flow rate of 600 nL/min, with the B solution (acetonitrile containing 0.1% formic acid) increasing linearly from 6% to 30% over 0 to 42 minutes, followed by a further increase from 30% to 42% between 42 and 51 minutes, and finally rising to 95% within 5 minutes, which was maintained for 60 minutes.

The Thermo Scientific Q Exactive HF mass spectrometer, equipped with a Nanospray Flex ion source, was utilized, with the ion spray voltage set to 2.3 kV and the temperature of the ion transfer tube maintained at 320°C. The mass spectrometer operated in Data-Independent Acquisition (DIA) mode. Following the collection of DIA data, the Spectronaut 18.0 software (Biognosys) was employed to search the human database downloaded from Uniprot.

2.8 Kidney metabolome analysis

Kidney samples from the CKD and JPYS-H groups (n=4 in each group) were processed by combining them with a standard chromatography and mass spectrometry protocol. Briefly, The procedure involved several methodical steps: (1) Kidney samples were processed through homogenization in 80% methanol and subsequently incubated at -80°C for a duration of two hours. After the incubation, the mixture was subjected to centrifugation, and the supernatant was collected and evaporated using nitrogen gas. To facilitate reconstitution, 100μl of an acetonitrile-water solution (in a 4:1 ratio) was added. The resulting mixture was vortexed, centrifuged, and the supernatant was transferred to a

liquid phase vial for further analysis; (2) The mobile phase A consisted of 0.1% formic acid in ultrapure water, while mobile phase B was composed of methanol. The flow rate was maintained at 0.30 mL/min, with the oven temperature set at 40°C. The autosampler temperature was regulated at 10°C, and the injection volume was 2 μL utilizing a full loop injection method; (3) Targeted profiling was conducted using a QTRAP® 5500 System (SCIEX) operating in positive mode, employing the Multiple Reaction Monitoring (MRM) technique. The electrospray ionization parameters were optimized for a flow rate of 0.30 mL/min, with the following specifications: electrospray voltage of 5500 V, temperature of 500°C, curtain gas at 40, CAD gas at 12, and gases 1 and 2 set at 50 psi each.

2.9 Cell culture

THP-1 cells were cultured in a 6-well plate and incubated with 10 ng/ml of Phorbol 12-myristate 13-acetate (PMA) for 48 hours. Followed by incubated in different doses of JPYS formula for 24h and 48h. Cell Counting Kit-8 (CCK-8) was used to assess cell viability.

Afterwards, the macrophages were stimulated into M1 and M2 polarization, respectively. In M1 polarization, LPS (100 ng/ml) and IFN-γ (20 ng/ml) were used to stimulate M1 macrophages. The groups were divided into different categories, including Macrophage, M1-Macrophage, and JPYS formula doses (M1 Macrophage incubated with different JPYS formula doses) for 48h. In M2 polarization, IL-4 (25ng/ml) and IL-13 (25ng/ml) were used to stimulate M2 macrophages. The groups were divided into different categories, including Macrophage, M2-Macrophage, and JPYS formula doses (M2 Macrophage incubated with different JPYS formula doses) for 48h.

2.10 Macrophage ultra-performance liquid chromatography tandem mass spectrometry

The methodology employed for the extraction of metabolites from cells, chemicals, reagents, and the UPLC-MS/MS conditions adhered to the procedures outlined in previous studies (24) .The protocol involved several steps: (1) Cell samples were treated with Methanol for shaking and lysis, followed by incubation at -80°C for 30 minutes. Subsequently, the samples were subjected to shaking and centrifugation, and the resulting supernatant was dried using a nitrogen blower. Prior to sample running, re-dissolution was performed. Additionally, the preparation and optimization of an amino acid standard solution were carried out, including the determination of parent and daughter ions of the standard. (2) Standard curves were created at various concentrations (1000%, 500%, 200%, 100%, 80%, 40%, 20%, 10%, 5%) along with the configuration of the mobile phase, liquid phase method, and mass spectrometry method. The samples were then analyzed using LC-MS. (3) Experimental sample concentration involved running the samples on a C18 chromatography column, with adjustments made for samples with low amino acid content. Samples that did not produce peaks were rerun using Glycan columns with corresponding adjustments to the mobile phase and mass spectrometry method. (4) Machine testing was conducted by mixing 50μl of supernatant from each sample in a centrifuge tube, followed by randomization and interspersion of Quality Control (QC) samples with cell samples. The samples were numbered and sorted based on random numbers, with QC values used to calculate Coefficient of Variation (CV) values. A CV value within 15% indicated acceptable system deviation. QC samples 1-10 were configured to control sample quality, and new standard curve ranges were established based on sample concentration test results. Post-sample run, Multi Quant software was utilized for result analysis.

2.11 Statistical analysis

The measurement data was presented as the mean ± SEM. The one-way analysis of variance (ANOVA) or the Kruskal-Wallis test was employed to assess significant differences among groups. Statistical analysis was conducted using GraphPad Prism software, with a significance level set at $P < 0.05$.

3 Results

3.1 Network pharmacology showed that JPYS formula might reduce CKD progression via different targets and pathways

933 active ingredients of the JPYS formula were retrieved through the TCMSP database and further screening performed using the parameters of OB ≥ 30% and DL ≥ 0.18 revealed 224 potential active ingredients (Supplementary Table 2). Afterwards, 227 potential targets of the JPYS formula were predicted by TCMSP database, and 14,946 CKD-related targets were collected via the GeneCards and OMIM databases. The comparison of the targets identified via these two methods revealed 224 overlapping targets (Figure 1A). Imported 224 common targets of JPYS formula and chronic kidney disease into the STRING database to obtained PPI protein interaction network data. Afterwards, the data were imported into Cytoscape 3.7.1 software to analyze and obtain the connectivity (Degree), BC and CC of drug-disease common targets. Core targets were considered with targets greater than the median of the above parameters, totaling 88 (Figure 1A).

3.2 GO analysis of "JPYS formula–CKD"

2117 GO enrichment analysis entries were obtained in the STRING database, including 1784 for BP analysis, 203 for MF

10.3389/fimmu.2024.1512519

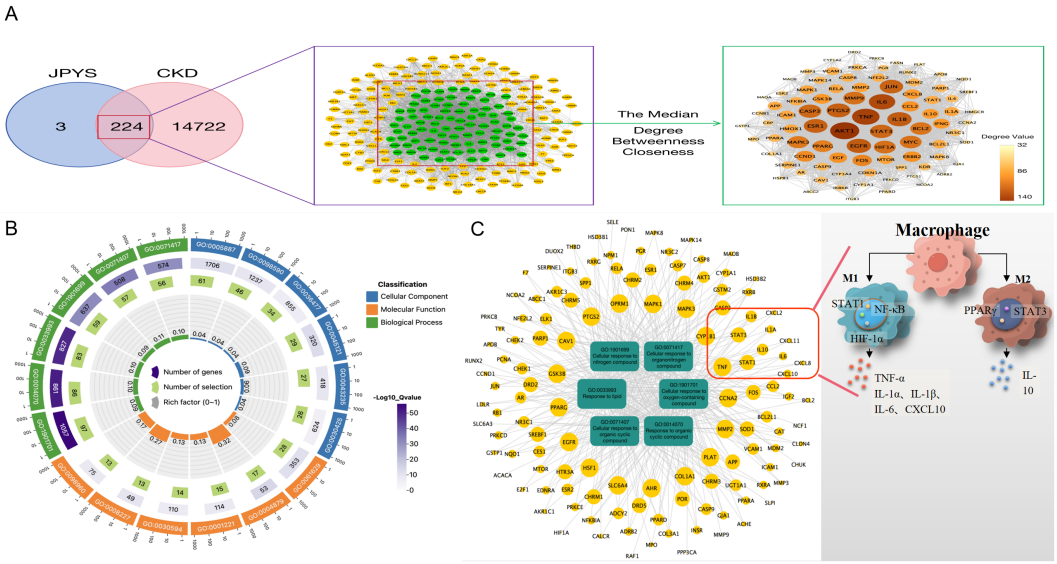


FIGURE 1
Network pharmacology screening and prediction of effective chemical constituents in the JPYS formula. **(A)** 227 potential targets of the JPYS formula were predicted by TCMSP database, and 14,946 CKD-related targets were collected via the GeneCards and OMIM databases. The comparison of the targets identified via these two methods revealed 224 overlapping targets. Core targets were considered with targets greater than the median of the above parameters, totaling88. **(B)** Top 6 significantly enriched items in BP/MF/CC; **(C)** "BP entry target" network diagram demonstrated that 11 targets were associated with macrophage activity.

analysis, and 130 for CC analysis. Set the analysis conditions to observed gene count and strength both greater than the median, and obtain the top 6 significantly enriched items in BP/MF/CC (Figure 1B). Establishing a "BP entry target" network diagram, it was found that 11 targets were associated with macrophage activity, as shown in Figure 1C.

3.3 JPYS formula exhibits renoprotective effects in CKD rats

Rats in the CKD group exhibited higher levels of serum creatinine (Cre) and blood urea nitrogen (BUN), which were restored after JPYS formula treatment (Figure 2A). H&E staining revealed the CKD group exhibited inflammatory injuries and fibrotic changes, while the CKD + JPYS group showed a significant reduction in pathological injuries, consistent with the improvement in renal function (Figure 2B).

To confirm whether the renoprotective effect of JPYS formula was associated with the modulation of macrophages, immunohistochemistry was performed to measure the expression of CD68 and CD86 in the kidney tissue. The CKD group exhibited higher levels of CD68 and CD86 expression than the control group, while the JPYS group showed lower expression levels than the CKD group (Figure 3A). Additionally, there were notable statistical differences in Integrated Optical Density (IOD) values among various groups (Figure 3B). These results indicate that JPYS formula therapy down-regulates macrophages, including M1 and M2 macrophages in the kidneys of CKD rats.

3.4 Amino acid metabolism may be the core targets for JPYS formula to delay the progression of CKD

RNA enrichment analysis was conducted to compare the differences between CKD and CKD+JPYS groups in RNA-seq (n = 4 per group). There's 132 different genes between the CKD group and JPYS groups (Figure 4A). And the heatmap showed different cluster between CKD and JPYS groups (Figure 4B). The cluster analysis of the GO classification in RNA-seq revealed a significant enrichment of the metabolic process (Figure 4C) between CKD and JPYS formula. In the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, the Glycine, serine, and threonine metabolism (Amino acid metabolism), Butanoate metabolism (Carbohydrate metabolism), Biosynthesis of amino acids (Global and overview maps), Pyruvate metabolism (Carbohydrate metabolism), and Glycolysis/Gluconeogenesis (Carbohydrate metabolism) were enriched. This suggests that JPYS formula adjusted the metabolic function in CKD rats (Figure 4D).

Moreover, the proteomics analysis, which included PCA and heatmap visualization, suggested that JPYS rats could be separated from CKD rats (Figures 5A, B). Volcanoplot revealed that JPYS formula exhibited up-regulated 260 proteins while down-regulated 339 proteins compared to CKD (Figure 5C). An enrichment analysis of the pathway functional entries within the Reactome database, where differential proteins are identified, indicates that the immune system and metabolic pathways are critical for interventions involving JPYS (Figure 5D). Additionally, we performed an extensive analysis of the differences in metabolic pathways through proteomics, which

10.3389/fimmu.2024.1512519

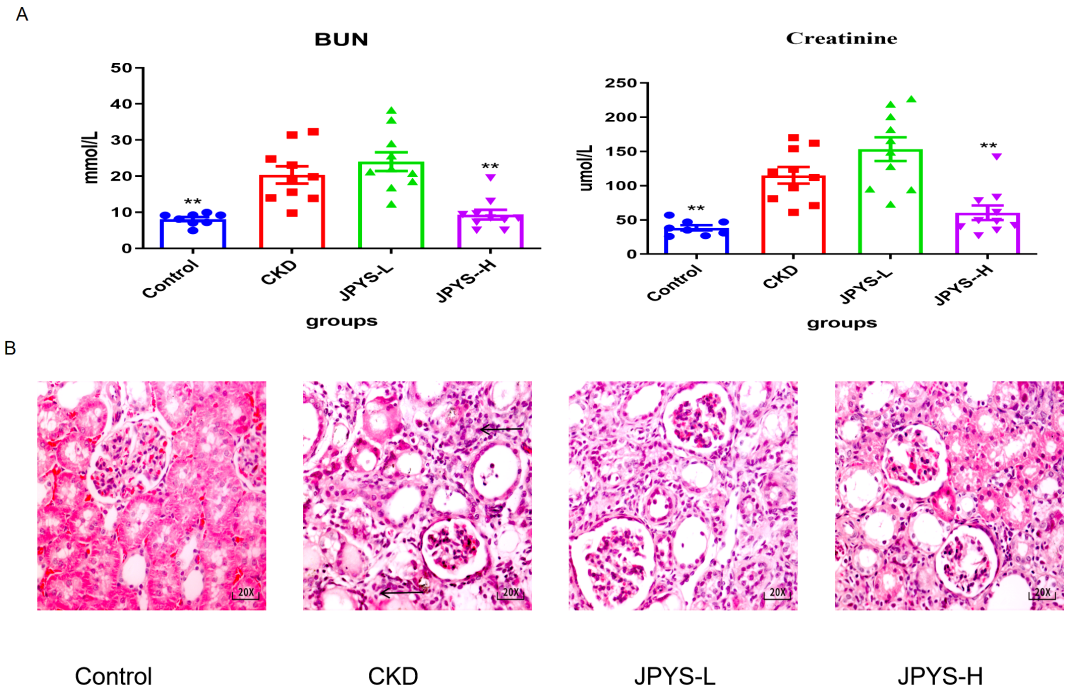


FIGURE 2
Effects of JPYS formula on renal function and pathological injury in CKD rats. **(A)** Blood urea nitrogen levels and serum creatinine levels. **(B)** Representative HE staining images in each groups. The arrows indicated the usual pathological alterations associated with CKD. **Represents a significant variation compare with the CKD group among the multiple comparisons. $^{**}P < 0.05$.

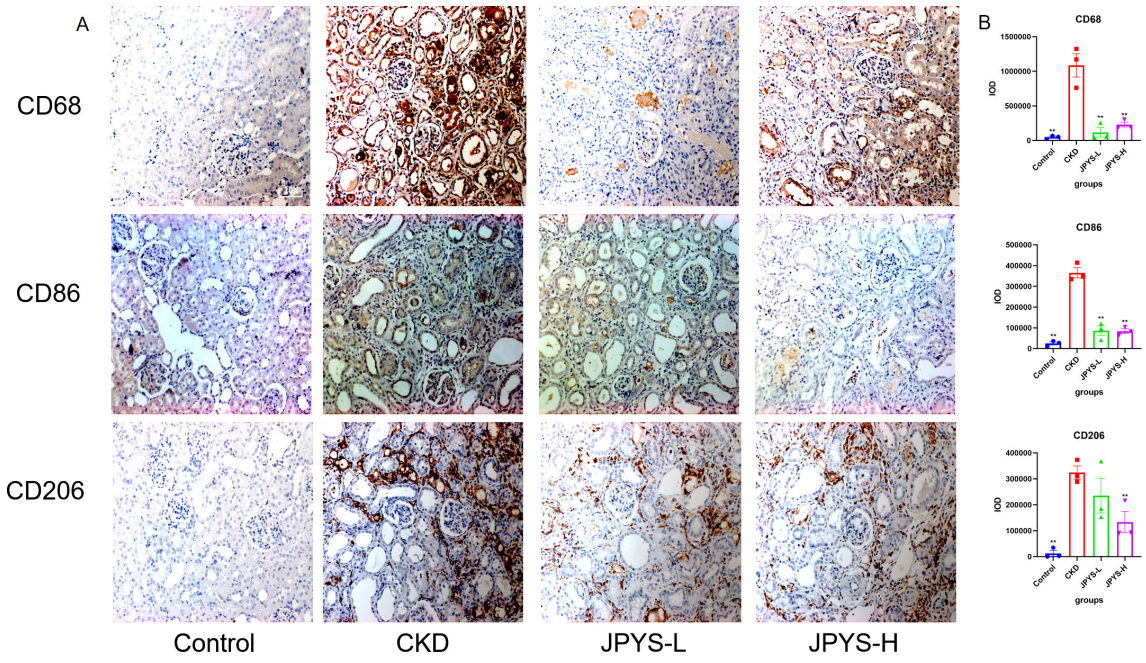


FIGURE 3
Representative immunohistochemistry images and IOD values of CD68, CD86 and CD206 expression in the kidneys of rats. **(A)** All images are shown at identical magnification, ×200. **(B)** The IOD values among various groups. **Represents a significant variation compare with the CKD group among the multiple comparisons. $^{**}P < 0.05$.

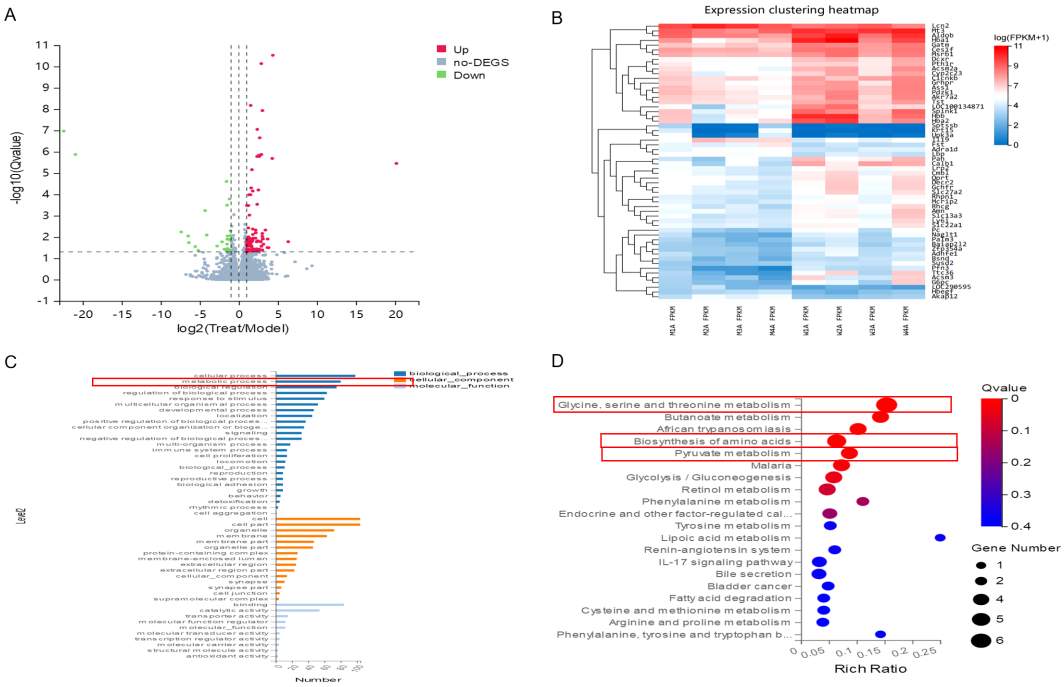


FIGURE 4
RNA-seq between CKD and JPYS groups. **(A)** Volcano plot analysis showed 132different genes between the CKD group and JPYS groups. **(B)** Heatmap showed different cluster between CKD and JPYS groups. **(C)** The cluster analysis of the GO classification revealed a significant enrichment of the metabolic process. **(D)** In the KEGG pathway, the Glycine, serine, and threonine metabolism (Amino acid metabolism), Butanoate metabolism (Carbohydrate metabolism), Biosynthesis of aminoacids (Global and overview maps), Pyruvate metabolism (Carbohydrate metabolism), and Glycolysis/Gluconeogenesis (Carbohydrate metabolism) were enriched.

demonstrated that JPYS could upregulate various metabolic pathways, particularly highlighting the significance of amino acid metabolism and the metabolism of other amino acids in this upregulation (Figure 5E). These results further validated our transcriptomic findings. A comprehensive KEGG analysis of cellular processes revealed that the regulation of the actin cytoskeleton, phagosome, and lysosome are significant biological processes influencing the differential proteins associated with JPYS (Figure 5F). And the tryptophan played a pivotal role in the metabolism pathway (Figure 5G). This finding corroborated our previous network pharmacology hypothesis that macrophages may served as vital cellular targets for JPYS in the context of delaying CKD.

3.5 The regulation of amino acid metabolism by Betaine in macrophage polarization may serve as a potential target for the JPYS formula in delaying the progression of CKD

To identify metabolic pathways, we conducted metabolomic in kidney sample to further investigate the metabolic between JPYS rats and CKD rats. The Partial Least Squares Discriminant Analysis (PLSDA) conducted in the field of metabolomics indicates a distinct separation between JPY rats and CKD rats, as illustrated in Figure 6A. An examination of renal metabolism post-JPYS intervention identified seven metabolites that exhibited significantly elevated levels, namely Betaine, Glycine, Alanine, Asparagine, Glutamic acid, Creatine, and Glutamine, as depicted in Figure 6B. The identification of KEGG pathways associated with the differential metabolic functions observed between the two groups (Figure 6C). The relevant differential pathways encompassed Glycine, Serine, and Threonine metabolism, as well as amino acid biosynthesis, aligning with the results obtained from RNA sequencing. Additionally, an interactive network graph analysis indicated that Betaine plays a regulatory role in Glycine, Serine, and Threonine metabolism, metabolic pathways, and ABC transporters (Figure 6D).

To confirm the amino acid metabolism in macrophages, we utilized THP-1 cells and incubated into macrophages, followed by stimulation to generate M1 and M2 polarization macrophages for further metabolic analysis. In CCK-8, we observed that 1mg/ml, 2mg/ml and 4mg/ml shows positive influence in THP-1 cells while 8mg/ml JPYS formula downregulated THP-1 proliferation (Figures 7A, B). With the macrophage UPLC-MS/MS, we observed changes in amino acid metabolic pathways (Figures 7, 8). The tryptophan was up-regulated after treatment with JPYS formula (Figure 7C). And betaine was up-regulated after treatment with JPYS formula (Figure 8A).

4 Discussion

In previous research, the JPYS formula has been shown to have convincing effects in anti-inflammation, anti-fibrosis, and the

10.3389/fimmu.2024.1512519

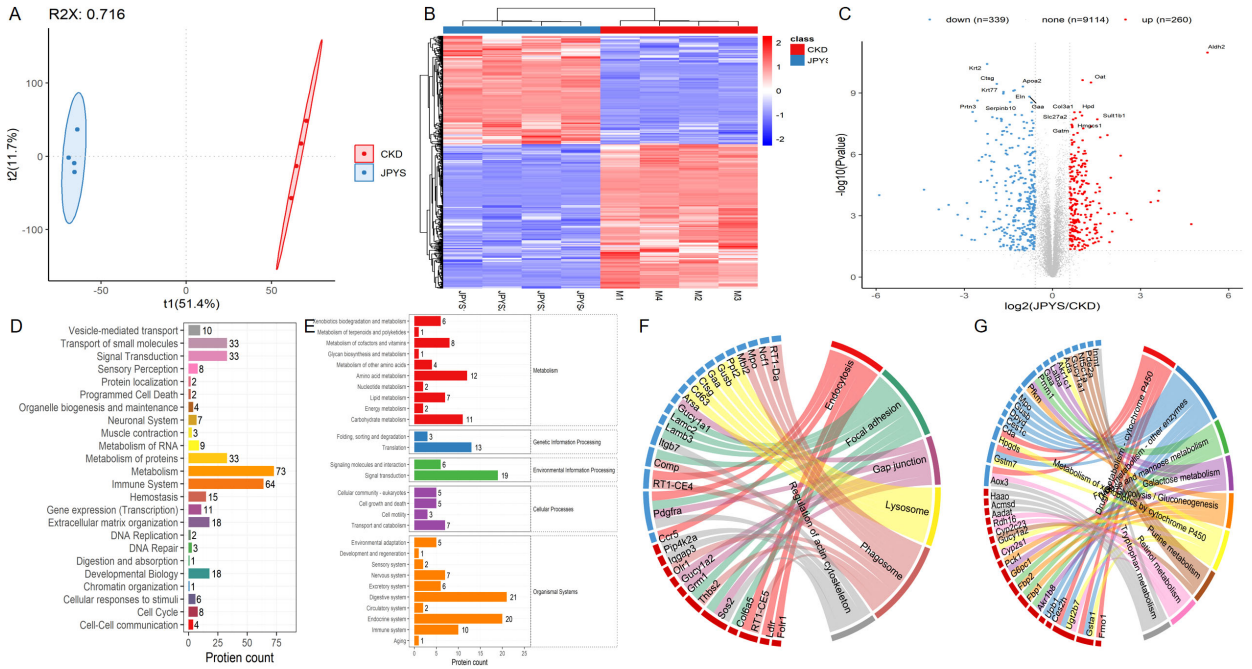


FIGURE 5
The proteomics analysis in kidney tissue. **(A, B)** The PCA and heatmap suggested that JPYS rats could be separated from CKD rats. **(C)** Volcanoplot revealed that JPYS formula exhibited up-regulated 260 proteins while down-regulated 339 proteins than CKD. **(D)** An enrichment analysis of the pathway functional entries within the Reactome database, where differential proteins were identified, indicated that the immune system and metabolic pathways were critical for interventions involving JPYS. **(E)** In metabolic pathways, the importance of amino acid metabolism, as well as the metabolism of other amino acids, played a crucial role in the process of upregulation. **(F)** A comprehensive KEGG analysis of cellular processes revealed that the regulation of the actin cytoskeleton, phagosome, and lysosome were significant biological processes influencing the differential proteins associated with JPYS. **(G)** And the tryptophan played a pivotal role in the metabolism pathway.

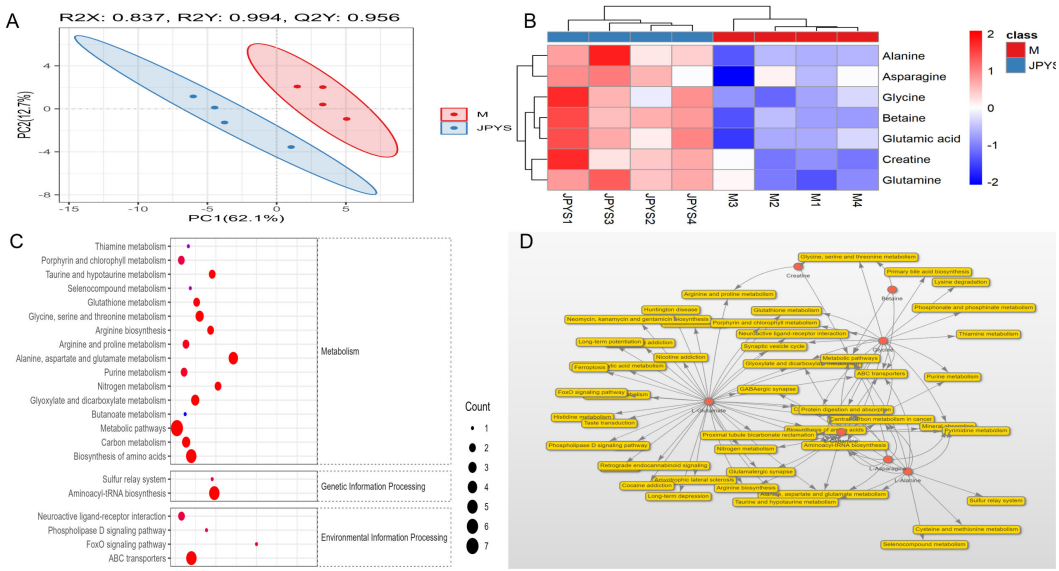


FIGURE 6
The metabolomic analysis in kidney tissue to further investigate the metabolic between JPYS rats and CKD rats. **(A)** The PLSDA indicated a distinct separation between JPY rats and CKD rats. **(B)** An examination of renal metabolism post-JPYS intervention identified seven metabolites that exhibited significantly elevated levels. **(C)** The identification of KEGG pathways associated with the differential metabolic functions observed between the two groups. **(D)** The interactive network graph analysis indicated that Betaine plays a regulatory role in Glycine, Serine, and Threonine metabolism, metabolic pathways, and ABC transporters.

10.3389/fimmu.2024.1512519

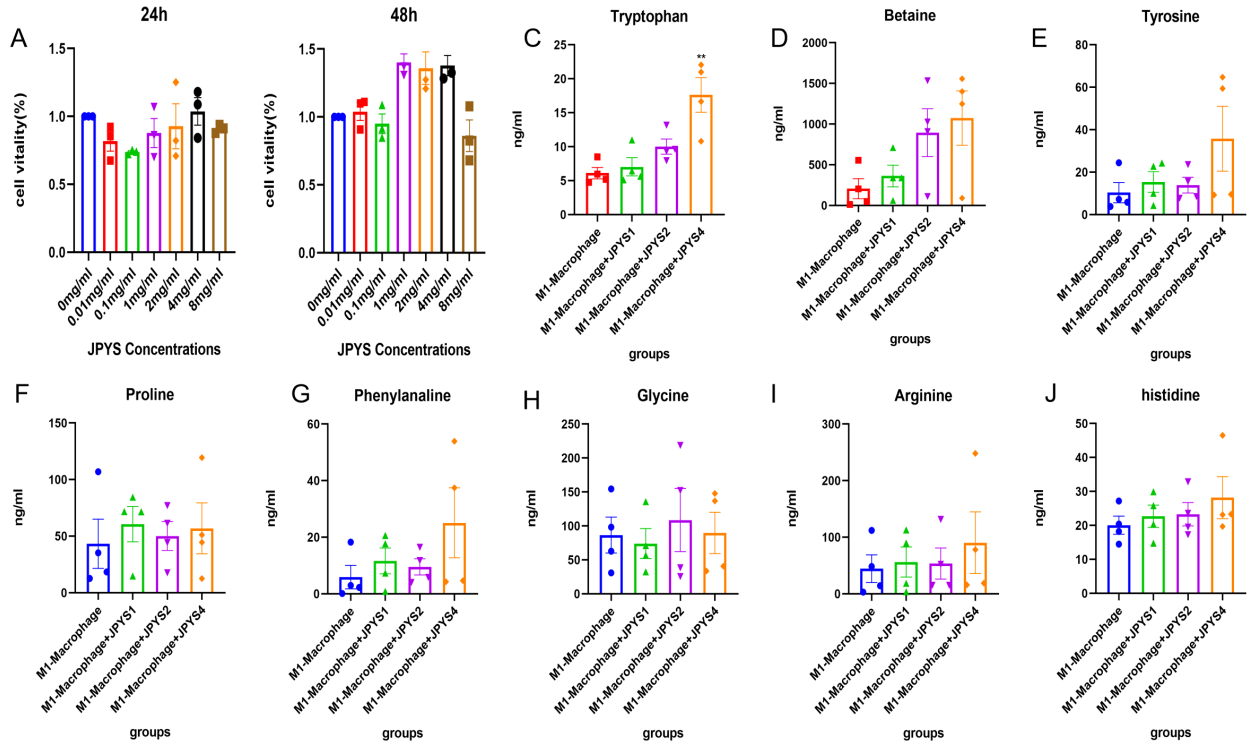


FIGURE 7
Metabolic pathway changed in M1 macrophage polarization incubated with JPYS. **(A, B)** THP-1 cells and incubated into macrophages, followed by incubated in different doses of JPYS for 24h and 48h. **(C-J)** Changes in amino acid metabolic pathways in M1 macrophage polarization incubated with JPYS, particularly the tryptophan was up-regulated after treatment with JPYS. **Represents a significant variation compare with the CKD group among the multiple comparisons. $^{**}P < 0.05$.

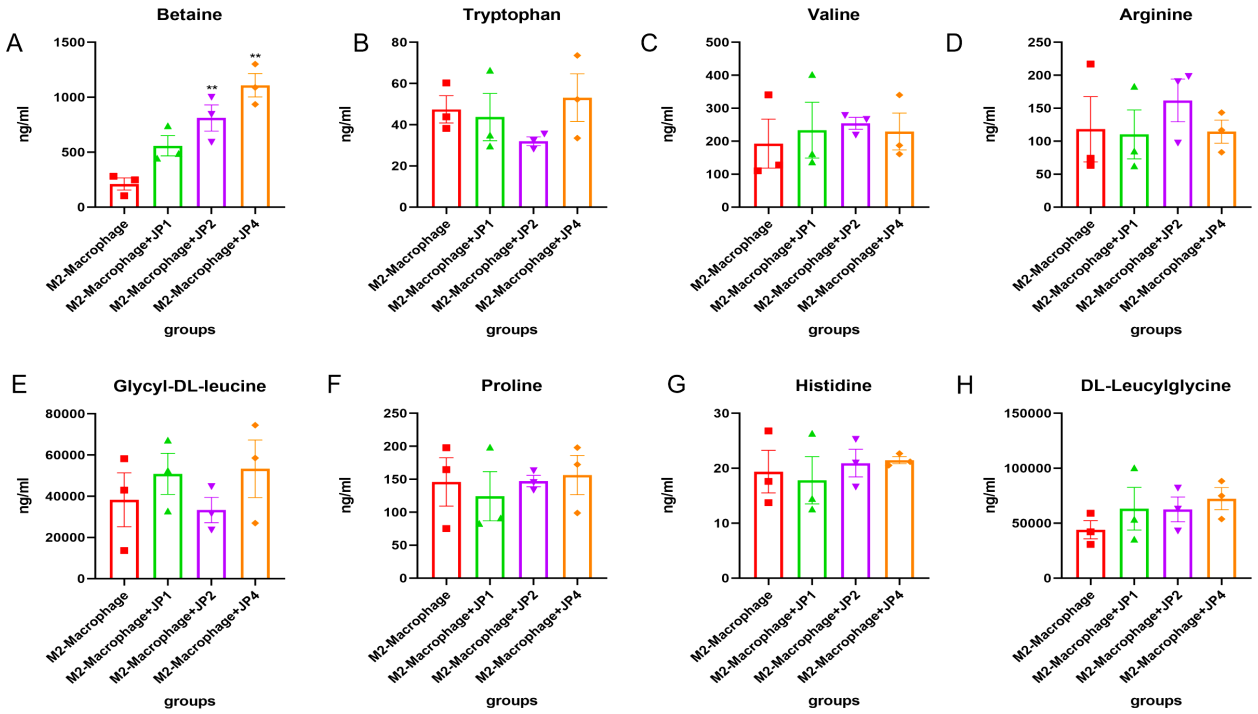


FIGURE 8
Metabolic pathway changed in M2 macrophage polarization incubated with JPYS. **(A-H)** metabolic pathways particularly Betaine was up-regulated after treatment with JPYS. **Represents a significant variation compare with the CKD group among the multiple comparisons. $^{**}P < 0.05$.

restoration of iron metabolism in CKD rats (14, 25). However, the underlying mechanisms by which the JPYS formula inhibits inflammation, especially via macrophage activity, remain unclear.

In the study, we initially employed network pharmacology to predict the active ingredients of the JPYS formula that are pharmacologically effective in CKD. This was followed by the construction and analysis of protein interaction networks, as well as conducting GO and KEGG enrichment analyses. Within the enriched pathways, we identified macrophage polarization as a potential target pathway of the JPYS formula, in relation to inflammatory injuries associated with CKD. This is consistent with the inflammatory injury that occurs during the natural progression of CKD. In line with this, in the *in vivo* model of CKD, the JPYS formula was shown to improve kidney function and alleviate kidney histopathological inflammatory damage. To validate the network pharmacology results, we looked further into macrophage immunophenotype expression and found that CKD rats had higher levels of CD68, CD86, and CD206, while the JPYS formula may have caused downregulation of macrophage surface marker expression in CKD. Furthermore, transcriptomic profiling of kidney tissue has indicated that the metabolic pathways linked to "Glycine, serine, and threonine metabolism" and "Biosynthesis of amino acids" are enriched as potential pathways of interest from, through which the JPYS formula provides renoprotection.

We then carried out additional validation using kidney proteomics, which identified the "metabolism" and "immune system" pathways as important mechanisms via which the JPYS formula has its therapeutic effects in CKD. The analyses also identified cellular functions, such as the regulation of the actin cytoskeleton, phagosomes, and lysosomes, in addition to significant associations with macrophages, which play a pivotal role in immune function. Additionally, tryptophan was found to be essential for the metabolism pathway. This finding is consistent with the network pharmacology analysis, indicating that macrophages may serve as important therapeutic targets in the treatment of CKD with JPYS.

Building further on this hypothesis, the metabolomics analysis from kidney tissue showed that significant differential pathways, such as the metabolism of glycine, serine, and threonine as well as the biosynthesis of amino acids, were enriched after JPYS intervention. Additionally, we observed a significant overexpression of betaine in the kidney tissue, which is probably related to the inclusion of Astragalus membranaceus and Cistanche deserticola (26, 27). Enrichment of the pathways related to glycine, serine, and threonine metabolism appear to correlate with the increased levels of betaine. Macrophage polarization was also impacted by the JPYS intervention, according to data from the *in vitro* studies. In particular, treatment with the JPYS formula increased tryptophan levels in the context of M1 macrophage polarization. JPYS formula treatment also increased the expression of betaine during the process of M2 macrophage polarization. Together, the JPYS formula may have protective effects against CKD injury by reducing inflammatory damage through the interaction of macrophage polarization and amino acid metabolism.

Macrophages are essential parts of kidney tissue which play critical roles in renal inflammation, the immune response, and the maintenance of kidney homeostasis (28, 29). In CKD, the persistent activation of pro-inflammatory monocytes and the presence of reparative macrophages contribute to conditions like glomerulosclerosis and tubulointerstitial fibrosis (30). Macrophage polarization, characterized by M1 pro-inflammatory and M2 reparative phenotypes, is a response to inflammatory stimuli, and the transition from M1 to M2 macrophages has been observed during the progression of CKD (31). Although certain herbal ingredients have demonstrated potential in regulating macrophage polarization and lowering inflammation, further research is needed to fully understand the impact of herbal treatments on aberrant macrophage-driven inflammation (32).

Current research indicates that the JPYS formula has the ability to influence both M1 and M2 macrophage polarization. Considering the role of amino acids, tryptophan metabolism can be improved by the JPYS formula in the context of M1 macrophage polarization. As an essential aromatic amino acid, tryptophan plays a key role in cellular synthesis, homeostasis maintenance, and it has been implicated in CKD progression (33). Disturbances in tryptophan metabolism are frequently reported in CKD patients, leading to worsening renal fibrosis and the progression of CKD, by causing metabolites to activate the aryl hydrocarbon receptor. As CKD advances, uremic toxins accumulate due to inadequate renal excretion, further resulting in deterioration of the condition (34). Conversely, disturbances in tryptophan metabolism can affect the kynurenine pathway, influencing the production of serotonin, indole-pyruvate derivatives, and tryptamine (34).

This study's findings indicate that the JPYS formula offers protective effects against inflammation-induced damage driven by M1 macrophages. This therapeutic effect may be due to the modulation of tryptophan levels by JPYS. Moreover, it has been shown that the JPYS formula raises levels of betaine, which is a neutral amino acid derivative that is associated with maintaining organ homeostasis and halting the progression of disease. Previous studies have demonstrated the betaine in lowering steatosis, inflammation, and fibrosis in metabolism-associated fatty liver disease as well as oxidative stress and inflammation linked to alcoholic liver disease. Among other benefits, betaine has also been demonstrated to maintain the integrity of the intestinal epithelial barrier, control adipose function, and prevent the development of cancer (35–37). In the context of kidney health, betaine plays a crucial role in protecting cells against osmotic stress, exhibiting anti-inflammatory and antioxidant properties. Furthermore, low betaine levels have been linked to increased intestinal dysbiosis, oxidative stress, inflammation, and kidney damage, underscoring its significance as a metabolite for assessing the stages of CKD (38, 39). Therefore, patients with CKD may benefit from incorporating betaine-rich diets into their diets. Together, the available evidence generally supports using the JPYS formula as an effective modulator of amino acid metabolism during macrophage polarization.

One previous study demonstrated that amino acids play a significant role in modulating the inflammatory resolution processes, particularly through their interaction with macrophages, specifically in terms of polarization and secretion (40). In the current study's enrichment analyses, the "Glycine, Serine, and Threonine Metabolism" pathway (KEGG map 00260) was shown to play a crucial role in the mechanism of action in JPYS treatment by

regulating immunity and mitigating inflammatory damage. Furthermore, betaine in JPYS promotes the upregulation of glycine, which further enhances the expression of serine and threonine. Elevated levels of serine also facilitate tryptophan metabolism. Finally, the activation of amino acid metabolism, particularly the glycine, serine, and threonine metabolism pathways, regulates macrophage polarization and ultimately alleviates renal immune and inflammatory damage.

Treatment with JPYS, which stands for "strengthening the spleen and kidney," is a traditional approach in Chinese medicine that aims to enhance blood circulation, eliminate dampness, and detoxify the body (41, 42). The JPYS formula has been widely applied in clinical settings and previous research has highlighted its various effects in delaying CKD progression. However, there is a research gap regarding the impact of the JPYS formula on immune function, which is crucial in understanding its potential therapeutic effects.

This study employed network pharmacology, RNA-seq, proteomic and metabolic analysis both *in vivo* and *in vitro*. We hypothesize that the JPYS formula elevates betaine levels in the kidney, thereby impacting amino acid synthesis and metabolism, particularly in pathways related to glycine, serine, and threonine metabolism. Ultimately, this modulation appears to influence macrophage polarization, which may represent a potential target for the JPYS formula in order to mitigate inflammatory injury and provide protection against CKD. Additionally, the study explored the formula's role in immune regulation, inflammation modulation, in macrophage polarization, and its impact on metabolic changes to inhibit the progression of CKD.

5 Conclusion

Taken together, our findings suggest that the JPYS formula exerts its therapeutic effects through multiple mechanisms. These mechanisms include modulating inflammation, immune response, and macrophage polarization, as well as influencing metabolic changes. The interaction between amino acid metabolism and polarization, specifically the involvement of tryptophan in M1 polarization and betaine in M2 polarization, is a crucial mechanism of the JPYS formula in reducing inflammatory damage in CKD and decelerating its progression.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, bioproject/1090638 and www.iprox.cn. ID: IPX0010760000.

Ethics statement

The animal study was reviewed and approved by the institutional ethics review board of Shenzhen TopBiotech Co., Ltd. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

YYL: Funding acquisition, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. YML: Data curation, Formal analysis, Funding acquisition, Investigation, Writing – original draft. YH: Data curation, Investigation, Writing – original draft. SL: Data curation, Writing – original draft, Formal analysis. GL: Data curation, Writing – original draft, Investigation. WZ: Investigation, Writing – original draft. XG: Investigation, Writing – original draft. JW: Investigation, Writing – original draft. SL: Conceptualization, Methodology, Project administration, Supervision, Writing – review & editing. HC: Conceptualization, Project administration, Supervision, Writing – review & editing, Funding acquisition, Validation.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by grants from the Natural Science Foundation of Guangdong Province (Grant No.2022A1515011957) and Shenzhen Science and Technology Program (Grant No. RCBS20221008093346102). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1512519/full#supplementary-material

References

1. Huang R, Fu P, Ma L. Kidney fibrosis: from mechanisms to therapeutic medicines. *Signal Transduct Target Ther*. (2023) 8:129. doi: 10.1038/s41392-023-01379-7

2. Kishi S, Kadoya H, Kashihara N. Treatment of chronic kidney disease in older populations. *Nat Rev Nephrol*. (2024) 20:586–602. doi: 10.1038/s41581-024-00854-w

3. Yamamoto T, Isaka Y. Pathological mechanisms of kidney disease in ageing. *Nat Rev Nephrol*. (2024) 20:603–15. doi: 10.1038/s41581-024-00868-4

4. Kalantar-Zadeh K, Li PK-T, Tantisattamo E, Kumaraswami L, Liakopoulos V, Lui S-F, et al. Living well with kidney disease by patient and care partner empowerment: kidney health for everyone everywhere. *Transpl Int*. (2021) 34:391–7. doi: 10.1111/tri.13811

5. Li C, Jia WW, Yang JL, Cheng C, Olaleye OE. Multi-compound and drug-combination pharmacokinetic research on Chinese herbal medicines. *Acta Pharmacol Sin*. (2022) 43:3080–95. doi: 10.1038/s41401-022-00983-7

6. Miao H, Liu F, Wang Y-N, Yu X-Y, Zhuang S, Guo Y, et al. Targeting Lactobacillus johnsonii to reverse chronic kidney disease. *Signal Transduct Target Ther*. (2024) 9:195. doi: 10.1038/s41392-024-01913-1

7. Guo Z-Y, Wu X, Zhang S-J, Yang J-H, Miao H, Zhao Y-Y. Poria cocos: traditional uses, triterpenoid components and their renoprotective pharmacology. *Acta Pharmacol Sin*. (2024). doi: 10.1038/s41401-024-01404-7

8. He JY, Hong Q, Chen BX, Cui SY, Liu R, Cai GY, et al. Ginsenoside Rb1 alleviates diabetic kidney podocyte injury by inhibiting aldose reductase activity. *Acta Pharmacol Sin*. (2022) 43:342–53. doi: 10.1038/s41401-021-00788-0

9. Huang D, Shen P, Wang C, Gao J, Ye C, Wu F. Calycosin plays a protective role in diabetic kidney disease through the regulation of ferroptosis. *Pharm Biol*. (2022) 60:990–6. doi: 10.1080/13880209.2022.2067572

10. Mao W, Yang N, Zhang L, Li C, Wu Y, Ouyang W, et al. Bupi yishen formula versus losartan for non-diabetic stage 4 chronic kidney disease: A randomized controlled trial. *Front Pharmacol*. (2021) 11:627185. doi: 10.3389/fphar.2020.627185

11. Hu D, Wang L, Zhang Y, Liu X, Lu Z, Li H. Sanqi oral solution ameliorates renal fibrosis by suppressing fibroblast activation via HIF-1α/PKM2/glycolysis pathway in chronic kidney disease. *J Ethnopharmacol*. (2024) 335:118679. doi: 10.1016/j.jep.2024.118679

12. Zhao M, Yu Y, Wang R, Chang M, Ma S, Qu H, et al. Mechanisms and efficacy of chinese herbal medicines in chronic kidney disease. *Front Pharmacol*. (2021) 11:619201. doi: 10.3389/fphar.2020.619201

13. Bo F, Airong Q, Guoliang X, Yihou Z, Shunmin L. Clinical observation on the treatment of 146 cases of chronic renal failure with kidney failure no.1 formula. *World Chinese Medicine*. (2017) 12:2999–3001.

14. Li C, Huang H, Wang R, Zhang C, Huang S, Wu J, et al. Jian-Pi-Yi-Shen formula restores iron metabolism from dysregulation in anemic rats with adenine-induced nephropathy. *J Ethnopharmacol*. (2023) 312:116526. doi: 10.1016/j.jep.2023.116526

15. Liu X, Chen J, Liu X, Wang D, Zheng P, Qi A, et al. Jian-Pi-Yi-Shen Formula ameliorates chronic kidney disease: involvement of mitochondrial quality control network. *BMC Complement Altern Med*. (2018) 18:340. doi: 10.1186/s12906-018-2395-2

16. Zhao Y, Li X, Wang F, Huang S, Du H, Li S, et al. Network pharmacology and experimental verification strategies to illustrate the mechanism of jian-pi-yi-shen formula in suppressing epithelial-mesenchymal transition. *Front Pharmacol*. (2022) 13:873023. doi: 10.3389/fphar.2022.873023

17. Islamuddin M, Qin X. Renal macrophages and NLRP3 inflammasomes in kidney diseases and therapeutics. *Cell Death Discovery*. (2024) 10:229. doi: 10.1038/s41420-024-01996-3

18. Rayego-Mateos S, Marquez-Expósito L, Rodrigues-Diez R, Sanz AB, Guiteras R, Doladé N, et al. Molecular mechanisms of kidney injury and repair. *Int J Mol Sci*. (2022) 23:1542. doi: 10.3390/ijms23031542

19. Tan RZ, Liu J, Zhang YY, Wang HL, Li JC, Liu YH, et al. Curcumin relieved cisplatin-induced kidney inflammation through inhibiting Mincle-maintained M1 macrophage phenotype. *Phytomedicine*. (2019) 52:284–94. doi: 10.1016/j.phymed.2018.09.210

20. Liu Y, Wu J, Liang S, Xu J, Wei M, Du Z, et al. Guben Xiezhuo Decoction inhibits M1 polarization through the Raf1/p-Elk1 signaling axis to attenuate renal interstitial fibrosis. *J Ethnopharmacol*. (2024) 319:117189. doi: 10.1016/j.jep.2023.117189

21. Geng X, Fu Z, Geng G, Chi K, Liu C, Hong H, et al. Astilbin improves the therapeutic effects of mesenchymal stem cells in AKI-CKD mice by regulating macrophage polarization through PTGS2-mediated pathway. *Stem Cell Res Ther*. (2024) 15:427. doi: 10.1186/s13287-024-04025-3

22. Ru J, Li P, Wang J, Zhou W, Li B, Huang C, et al. TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. *J Cheminform*. (2014) 6:13. doi: 10.1186/1758-2946-6-13

23. Liu X, Zhang B, Huang S, Wang F, Zheng L, Lu J, et al. Metabolomics analysis reveals the protection mechanism of huangqi-danshen decoction on adenine-induced chronic kidney disease in rats. *Front Pharmacol*. (2019) 10:992. doi: 10.3389/fphar.2019.00992

24. Rong G, Weng W, Huang J, Chen Y, Yu X, Yuan R, et al. Artemether alleviates diabetic kidney disease by modulating amino acid metabolism. *BioMed Res Int*. (2022) 2022:7339611. doi: 10.1155/2022/7339611

25. Gao L, Huang X, Deng R, Wu S, Peng Y, Xiong G, et al. Jian-Pi-Yi-Shen formula alleviates renal fibrosis by restoring NAD+ biosynthesis *in vivo* and *in vitro*. *Aging (Albany NY)*. (2023) 16:106–28. doi: 10.18632/aging.205352

26. Li Z-Y, He P, Sun H-F, Qin X-M, Du G-H. (1)H NMR based metabolomic study of the antifatigue effect of Astragali Radix. *Mol Biosyst*. (2014) 10:3022–30. doi: 10.1039/c4mb00370e

27. Li Z, Li J, Li Y, Guo L, Xu P, Du H, et al. The role of Cistanches Herba and its ingredients in improving reproductive outcomes: A comprehensive review. *Phytomedicine*. (2024) 129:155681. doi: 10.1016/j.phymed.2024.155681

28. Tang PM-K, Nikolic-Paterson DJ, Lan H-Y. Macrophages: versatile players in renal inflammation and fibrosis. *Nat Rev Nephrol*. (2019) 15:144–58. doi: 10.1038/s41581-019-0110-2

29. Meng X-M, Nikolic-Paterson DJ, Lan HY. Inflammatory processes in renal fibrosis. *Nat Rev Nephrol*. (2014) 10:493–503. doi: 10.1038/nrneph.2014.114

30. Bell R, Conway BR. Macrophages in the kidney in health, injury and repair. *Int Rev Cell Mol Biol*. (2022) 367:101–47. doi: 10.1016/bs.ircmb.2022.01.005

31. Zeng J, Zhang Y, Huang C. Macrophages polarization in renal inflammation and fibrosis animal models (Review). *Mol Med Rep*. (2024) 29:29. doi: 10.3892/mmr.2023.13152

32. Kim M, Kim JY, Yang HS, Choe J-S, Hwang IG. Nepetoidin B from salvia plebeia R. Br. Inhibits inflammation by modulating the NF-κB and nrf2/HO-1 signaling pathways in macrophage cells. *Antioxidants (Basel)*. (2021) 10:1208. doi: 10.3390/antiox10081208

33. Grifka-Walk HM, Jenkins BR, Kominsky DJ. Amino acid trp: the far out impacts of host and commensal tryptophan metabolism. *Front Immunol*. (2021) 12:653208. doi: 10.3389/fimmu.2021.653208

34. Hui Y, Zhao J, Yu Z, Wang Y, Qin Y, Zhang Y, et al. The role of tryptophan metabolism in the occurrence and progression of acute and chronic kidney diseases. *Mol Nutr Food Res*. (2023) 67:e2300218. doi: 10.1002/mnfr.202300218

35. Olli K, Lahtinen S, Rautonen N, Tiihonen K. Betaine reduces the expression of inflammatory adipokines caused by hypoxia in human adipocytes. *Br J Nutr*. (2013) 109:43–9. doi: 10.1017/S0007114512000888

36. Abdelmalek MF, Sanderson SO, Angulo P, Soldevila-Pico C, Liu C, Peter J, et al. Betaine for nonalcoholic fatty liver disease: results of a randomized placebo-controlled trial. *Hepatology*. (2009) 50:1818–26. doi: 10.1002/hep.23239

37. Arumugam MK, Paal MC, Donohue TM Jr, Ganesan M, Osna NA, Kharbanda KK. Beneficial effects of betaine: A comprehensive review. *Biol (Basel)*. (2021) 10:456. doi: 10.3390/biology10060456

38. Alvarenga L, Ferreira MS, Kemp JA, Mafra D. The role of betaine in patients with chronic kidney disease: a narrative review. *Curr Nutr Rep*. (2022) 11:395–406. doi: 10.1007/s13668-022-00426-z

39. Guo F, Dai Q, Zeng X, Liu Y, Tan Z, Zhang H, et al. Renal function is associated with plasma trimethylamine-N-oxide, choline, L-carnitine and betaine: a pilot study. *Int Urol Nephrol*. (2021) 53:539–51. doi: 10.1007/s11255-020-02632-6

40. Zheng X, Zhu Y, Zhao Z, Chu Y, Yang W. The role of amino acid metabolism in inflammatory bowel disease and other inflammatory diseases. *Front Immunol*. (2023) 14:1284133. doi: 10.3389/fimmu.2023.1284133

41. Liu X, Deng R, Chen Y, Huang S, Lu J, Zheng L, et al. Jian-pi-yi-shen formula improves adenine-induced chronic kidney disease via regulating tryptophan metabolism and aryl hydrocarbon receptor signaling. *Front Pharmacol*. (2022) 13:922707. doi: 10.3389/fphar.2022.922707

42. Liu X, Deng R, Wei X, Wang Y, Weng J, Lao Y, et al. Jian-Pi-Yi-Shen formula enhances perindopril inhibition of chronic kidney disease progression by activation of SIRT3, modulation of mitochondrial dynamics, and antioxidant effects. *Biosci Rep*. (2021) 41:BSR20211598. doi: 10.1042/BSR20211598

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 12 March 2025
DOI 10.3389/fimmu.2025.1556386

OPEN ACCESS

EDITED BY
Yejun Tan,
Hong Kong Polytechnic University,
Hong Kong SAR, China

REVIEWED BY
Chunji Li,
Guangzhou National Laboratory, China
Hong Zhou,
Shanghai General Hospital, China

*CORRESPONDENCE
Shenglong Li
lishenglong@cqmu.edu.cn

RECEIVED 06 January 2025
ACCEPTED 24 February 2025
PUBLISHED 12 March 2025

CITATION
Wang T, Chen Z, Wang W, Wang H and Li S (2025) Single-cell and spatial transcriptomic analysis reveals tumor cell heterogeneity and underlying molecular program in colorectal cancer. *Front. Immunol.* 16:1556386. doi: 10.3389/fimmu.2025.1556386

Single-cell and spatial transcriptomic analysis reveals tumor cell heterogeneity and underlying molecular program in colorectal cancer

Teng Wang[1], Zhaoming Chen[1], Wang Wang[2,3], Heng Wang[1] and Shenglong Li[1]*

[1]Department of Bioinformatics, School of Basic Medical Sciences, Chongqing Medical University, Chongqing, China, [2]Department of Immunology, School of Basic Medical Sciences, Chongqing Medical University, Chongqing, China, [3]Chongqing Key Laboratory of Tumor Immune Regulation and Immune Intervention, Chongqing Medical University, Chongqing, China

Background: Colorectal cancer (CRC) is a highly heterogeneous tumor, with significant variation in malignant cells, posing challenges for treatment and prognosis. However, this heterogeneity offers opportunities for personalized therapy.

Methods: The consensus non-negative matrix factorization algorithm was employed to analyze single-cell transcriptomic data from CRC, which helped identify malignant cell expression programs (MCEPs). Subsequently, a crosstalk network linking MCEPs with immune/stromal cell trajectory development was constructed using Monocle3 and NicheNet. Additionally, bulk RNA-seq data were utilized to systematically explore the relationships between MCEPs, clinical features, and genetic mutations. A prognostic model was then established through Lasso and Cox regression analyses, integrating clinical data into a nomogram for personalized risk prediction. Furthermore, key genes associated with MCEPs and their potential therapeutic targets were identified using protein-protein interaction networks, followed by molecular docking to predict drug-binding affinity.

Results: We classified CRC malignant cell transcriptional states into eight distinct MCEPs and successfully constructed crosstalk networks between these MCEPs and immune or stromal cells. A prognostic model containing 15 genes was developed, demonstrating an AUC greater than 0.8 for prognostic evaluation over 1 to 10 years when combined with clinical features. A key drug target gene TIMP1 was identified, and several potential targeted drugs were discovered.

Conclusion: This study demonstrated that characterization of the malignant cell transcriptional programs could effectively reveal the biological features of highly heterogeneous tumors like CRC and exhibit significant potential in tumor prognosis assessment. Our research provides new theoretical and practical directions for CRC prognosis and targeted therapy.

KEYWORDS

colorectal cancer, tumor heterogeneity, prognosis, therapy, single-cell transcriptomics, spatial transcriptomics

1 Introduction

Colorectal cancer (CRC) is one of the three most common cancers worldwide and the second leading cause of cancer-related deaths, driven by its profound molecular and cellular heterogeneity (1–3). CRC is primarily classified into two genetic subtypes—chromosomal instability (CIN) and microsatellite instability (MSI)—with distinct biological behaviors and therapeutic responses (4–7). Immune checkpoint blockade (ICB) therapy has shown efficacy in advanced MSI-H tumors, yet most patients remain unresponsive, underscoring the need for novel biomarkers (8–10). Molecular subtyping approaches, such as the Consensus Molecular Subtypes (CMS) classification, integrate bulk transcriptomic and genomic data to stratify CRC into four prognostic subtypes (CMS1-4) (11). However, these bulk-level analyses fail to resolve the continuum of malignant cell states or their dynamic crosstalk with the tumor microenvironment (TME) (12, 13).

Recent advances in single-cell and spatial transcriptomics have revolutionized cancer research by enabling high-resolution dissection of tumor heterogeneity. Single-cell RNA sequencing (scRNA-seq) and ATAC-seq reveal transcriptional and epigenetic diversity within malignant cells, while spatial technologies map cellular interactions in TME niches (14–17). Despite these advances, existing studies often categorize malignant cells into discrete subtypes or focus on isolated TME components, neglecting the continuum of transcriptional plasticity and bidirectional stromal-immune interactions (18–20). Traditional methods like PCA or clustering impose rigid structures on transcriptional data: PCA reduces variance to orthogonal components but obscures transitional states, while clustering forces discrete boundaries on inherently continuous programs. In contrast, consensus non-negative matrix factorization (cNMF) decodes continuous transcriptional dynamics, as demonstrated by its ability to resolve plastic cell states in lung cancer (21).

To advance beyond these limitations, this study integrates single-cell and spatial multi-omics data, applying cNMF to decode CRC heterogeneity. We identified eight continuous transcriptional programs (MCEPs) in malignant cells, encompassing dynamic phenotypes such as hypoxia adaptation, partial EMT plasticity, and glandular differentiation. By combining spatial co-localization with pseudotime trajectory analysis of stromal and immune cells, we uncovered how MCEPs remodel the TME through specific regulatory nodes (e.g., TGFB1-mediated fibroblast activation, HMGB2-dependent angiogenesis). Furthermore, we developed a prognostic model integrating MCEP-TME interactions, validated through protein-protein network analysis and experimental databases to prioritize therapeutic targets.

The eight MCEPs delineate critical biological dimensions in colorectal cancer progression (1): Inflammatory-Hypoxia Stress Program (IHS-P) coordinates hypoxic adaptation and immune modulation within immune-enriched niches (2); Wnt Signaling Stress Program (Wnt-S-P) drives canonical Wnt activation in tumor cores (3); Proliferation Stress Program (PS-P) governs cell cycle progression through MYC/mTORC1 signaling (4); Inflammatory Epithelial pEMT Program (IE-pEMT-P) bridges interferon responses with partial EMT plasticity (5); Intermediate pEMT Program (I-pEMT-P) mediates TGFB1-dependent stromal activation (6); Mesenchymal pEMT Program (M-pEMT-P) executes ECM remodeling in stromal compartments (7); Cell Cycle Program (CC-P) regulates pan-tumoral mitotic processes (8); Glandular Secretion Program (GS-P) maintains epithelial differentiation near normal tissues. This framework deciphers CRC heterogeneity through malignant cell state dynamics and their spatial-ecological networks, enabling prognostic prediction and therapeutic target discovery for precision oncology.

2 Materials and methods

2.1 Download and preprocessing of single-cell and spatial transcriptomics sequencing data

Single-cell RNA sequencing data were processed using Seurat (v5.1.0) with rigorous quality control. Three publicly available human colorectal cancer datasets were analyzed: GSE166555 (13 tumors, 12 normals) (22), GSE200997 (16 tumors, 7 normals) (23) from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/), and syn26844071 (141 tumors, 39 normals) (24) from the Synapse database (https://www.synapse.org/). Doublets were removed using Scrublet (v0.2.3), followed by gene/cell filtering criteria: genes detected in ≥3 cells, cells expressing ≥250 genes, UMI counts <15,000, mitochondrial gene percentage <20%, and erythrocyte gene ratio <1%.

Spatial transcriptomics data were obtained from the 10x Genomics Visium HD platform (8 μm resolution) and downloaded from the official 10x Genomics website (https://www.10xgenomics.com/), comprising a total of three samples (25). Quality control was performed on the spatial transcriptomics data, with spots retained for downstream analysis meeting the following thresholds: detection of ≥10 genes, UMI counts >20, and mitochondrial gene ratio <25%.

2.2 Cell annotation for single-cell and spatial transcriptomics data

scRNA-seq data underwent log-normalization and identification of highly variable genes (vst method). Batch correction was performed using Harmony (v0.1.0). Cell types were annotated through a two-step approach: 1) Initial classification using SingleR (v2.6.0) and CellTypist (v1.6.3) with canonical markers; 2) Refinement via secondary dimensionality

reduction and iterative CellTypist-based annotation, followed by removal of misclassified cells.

For spatial data, we implemented memory-efficient processing by subsampling 50,000 points using SketchData. Cell type deconvolution was performed using RCTD (v2.2.1) with scRNA-seq data as reference. Each spatial sample underwent independent dimensionality reduction and annotation.

2.3 Identification of malignant epithelial cells and gene expression program profiling

Epithelial cells were isolated from the full cell atlas and subjected to chromosomal copy number variation (CNV) analysis using inferCNV (v1.18.1), with normal colorectal epithelial cells as the reference. A CNV score matrix was generated, and unsupervised K-means clustering partitioned cells into malignant or normal clusters based on CNV-driven cluster purity.

For malignant cell subtyping, consensus high-variance genes were identified through 200 iterations of 75% subsampling. Genes recurrently ranked among the top 2,500 highly variable genes in ≥150 iterations were retained. These genes underwent non-negative matrix factorization (cNMF) to decompose the expression matrix into gene expression programs (GEPs) and their corresponding activity scores. The optimal number of GEPs was determined by minimizing reconstruction error and maximizing stability via elbow plot analysis.

To define high-weight genes within each MCEP, genes were ranked by their absolute weights in the cNMF gene coefficient matrix. The top 100 genes per program, exhibiting the strongest association with each transcriptional module, were selected for downstream spatial mapping. Spatial enrichment scores for these gene sets were computed using the AUCell R package (v1.24.0), enabling visualization of MCEP distribution patterns across tissue sections.

2.4 Pseudotime analysis

Developmental trajectories were reconstructed using Monocle3 (v1.3.5) with UMAP for dimensionality reduction. Cell subtypes were pre-annotated through immune and stromal cell clustering, which revealed preliminary developmental hierarchies. To resolve ambiguous differentiation origins arising from complex branching trajectories, we implemented a hybrid strategy for root node selection (1): For lineages with biologically established progenitor-differentiated cell relationships (e.g., T cell and B cell hierarchies), root nodes were manually assigned to progenitor states based on canonical marker expression and prior biological knowledge (2); For cell types lacking definitive developmental origins, root nodes were computationally determined by selecting the subpopulation with the highest transcriptional immaturity index, as quantified by CytoTRACE2 (v1.0.0). Trajectory-associated genes were identified using Monocle3's graph_test function with "neighbor_graph= principal_graph" to evaluate gene expression dynamics along reconstructed paths.

2.5 Expression program crosstalk networks

Intercellular crosstalk networks were constructed by defining trajectory-associated genes (Moran's $|I| > 0.25$, $q < 0.05$) from each malignant cell population as target gene sets. For each MCEP, the top 100 weighted genes in expression programs were selected as candidate regulators. Ligand-target interactions were predicted using NicheNet (v2.1.5), generating regulatory potential matrices where malignant cell regulators were prioritized based on their capacity to modulate target gene sets. Potential interactions in the lowest tertile of regulatory scores were nullified to eliminate spurious associations. Final immune and stromal interaction networks were reconstructed in Cytoscape (v3.10.2) using thresholded matrices for edge weighting.

2.6 Bulk sequencing data sources

Bulk RNA-seq data and simple nucleotide variation (SNV) data for colorectal cancer were obtained from The Cancer Genome Atlas (TCGA) database (https://www.cancer.gov/ccg/research/genome-sequencing/tcga). Using the R package TCGAbiolinks (v2.30.4), we retrieved RNA-seq data from 581 colorectal cancer patients and 51 normal colorectal control samples, along with SNV data for 538 patients. Clinical data for TCGA patients and pan-cancer gene expression profiles were additionally acquired from the UCSC Xena database (https://xena.ucsc.edu/).

To complement TCGA data, gene expression microarray datasets and corresponding clinical information were downloaded from the GEO database. Datasets included GSE39582 (26), GSE17536 (27), GSE17537 (27), GSE29621 (28), GSE38832 (29), GSE143985 (30), and GSE161158 (31), all generated on the GPL570 platform. From GSE39582, GSE17536, GSE17537, GSE29621, and GSE38832, overall survival (OS) data were extracted. After filtering samples with missing survival time, status, or non-positive survival time, 573, 177, 55, 65, and 122 samples were retained, respectively. Disease-free survival (DFS) and recurrence/survival status data were obtained from GSE143985 and GSE161158. Following similar quality control, 91 and 174 samples were retained, respectively.

2.7 Differential and enrichment analyses

To further investigate the changes in expression program-related genes at the bulk level, we integrated two distinct gene cohorts: 1) the top 100 weighted genes from each MCEP module, and 2) computationally predicted target genes in the MCEP-immune/stromal cell crosstalk network. Differential gene expression analysis was performed on this merged gene set using bulk RNA-seq data from the TCGA cohort through the R package

DESeq2 (version 1.42.1). Statistical significance was defined as absolute Fold Change > 1.5 and padj < 0.05. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were subsequently conducted on the identified differentially expressed genes (DEGs) using the clusterProfiler package (version 4.2.2) to characterize their functional roles.

2.8 Consensus clustering and intra-cluster comparison

Differentially expressed genes from TCGA were subjected to univariate Cox regression analysis (survival package v3.5-8, $p<0.05$) to identify survival-associated genes. Consensus clustering via ConsensusClusterPlus (v1.66.0) with 500 bootstraps (80% sample resampling) and K-means (Euclidean distance) identified optimal clusters (k=2-10) by evaluating consensus matrices and cumulative distribution functions (CDF). Subtype-specific survival differences were assessed by Kaplan-Meier analysis, while chi-square tests evaluated clinical characteristics (gender, age, stage). Mutation landscapes were visualized using maftools (v2.18.0), highlighting the top 15 recurrently mutated genes per subtype.

2.9 Construction of the prognostic model

Gene expression data were obtained from TCGA and seven GEO datasets (GSE39582, GSE17536, GSE17537, GSE29621, GSE38832, GSE143985, GSE161158). Batch effects were mitigated through z-score normalization followed by batch correction using the `removeBatchEffect` function (limma package v3.58.1). The TCGA and GSE39582 cohorts were partitioned into a training set (70% of samples) and an internal validation set (30%), while remaining datasets served as external validation cohorts.

To address feature redundancy, genes identified by univariate Cox regression ($p < 0.05$) were subjected to Lasso regression (glmnet v4.1-4) for dimensionality reduction. A stepwise backward Cox regression was then applied to optimize model complexity by minimizing the Akaike Information Criterion (AIC).

Risk scores were computed for all samples across training and validation cohorts. Survival differences between high- and low-risk groups (stratified by median risk scores) were evaluated using Kaplan-Meier analysis with log-rank tests. Predictive performance was quantified via time-dependent ROC curves and AUC values. Model robustness and clinical applicability were systematically validated across internal and external datasets using survival outcomes and AUC consistency.

2.10 Bulk immune landscape and calculation of single-cell and spatial risk scores

To explore the biological relevance of our prognostic model, we performed tumor immune microenvironment analysis on the TCGA cohort using the IOBR package (v0.99.9). Immune cell composition was quantified by integrating eight computational algorithms (MCPcounter, EPIC, xCell, CIBERSORT, IPS, quanTIseq, ESTIMATE, and TIMER). Spearman correlation analysis was then applied to evaluate associations among immune infiltration scores, prognostic feature gene expression, and sample risk scores.

For single-cell and spatial transcriptomic data, we adapted our risk scoring approach to address inherent data sparsity. Based on the regression coefficients from the linear prognostic model, feature genes were partitioned into two subsets: a positive-coefficient subset (PosRisk genes) and a negative-coefficient subset (NegRisk genes). The AddModuleScore function was employed to calculate PosRiskScore and NegRiskScore for each subset independently. Final RiskScore was derived as PosRiskScore minus NegRiskScore. This strategy enabled robust quantification of model-associated biological processes at cellular and spatial resolutions while mitigating technical limitations of sparse transcriptomic data.

2.11 Construction of a nomogram

Univariate Cox regression analysis was performed on TCGA cohort data to preliminarily identify variables (risk score, age, gender, tumor stage, and other clinical features) associated with overall survival. Subsequently, multivariate Cox regression analysis incorporating all candidate variables without prior feature selection was conducted to evaluate their independent prognostic contributions while adjusting for potential confounders.

A nomogram integrating the risk score and significant clinical predictors was developed using the regplot package (v1.1) to visualize survival probability estimates. Time-dependent receiver operating characteristic (ROC) analyses spanning 1-10 years were implemented to quantify predictive accuracy through area under the curve (AUC) calculations. Model calibration was validated using the rms package (v6.8-1) by comparing predicted versus observed survival probabilities via bootstrapped calibration curves (1,000 resamples). Clinical utility was further assessed through decision curve analysis (DCA) using the rmda package (v1.6), which quantified net benefits across threshold probabilities ranging from 0% to 100%. This comprehensive validation framework ensures methodological rigor and supports clinical translation of the prognostic model.

2.12 Key genes identification with malignant cell expression programs and drug screening

Differential expression analysis was performed on prioritized genes derived from malignant cell expression programs and their microenvironment-associated targets. Resultant genes were analyzed through the STRING database (https://cn.string-db.org/) to construct protein-protein interaction (PPI) networks, which were

further visualized and analyzed in Cytoscape (v3.9.1). Core hub genes were systematically identified using the cytoHubba plugin (v0.1) with four topology-based algorithms: MNC, MCC, DMNC, and Degree.

Expression differences of candidate genes between tumor and adjacent normal tissues were statistically validated using the Wilcoxon rank-sum test. Immunohistochemical images from The Human Protein Atlas (HPA, https://www.proteinatlas.org/) were utilized as supporting evidence.

For therapeutic exploration, three-dimensional structures of key targets were retrieved from UniProt (https://www.uniprot.org/), and 2,391 FDA-approved small-molecule drugs were sourced from DrugBank (https://go.drugbank.com/). Structural data standardization was implemented using rdkit (v2023.9.6) and meeko (v0.5.1), followed by protein active site prediction via the Prankweb database (https://prankweb.cz/). Molecular docking simulations were executed with AutoDock Vina (v1.2.5), prioritizing compounds based on binding affinity (ΔG, kcal/mol). The top two ligands exhibiting optimal docking scores were selected for binding conformation visualization using PyMOL (v3.1.0a0).

2.13 Software and data analysis tools

Single-cell and spatial transcriptomic analyses were performed using R (v4.3.2), with the cNMF algorithm (https://github.com/dylkot/cNMF) implemented in Python (v3.8.19). Drug virtual screening was conducted using Python (v3.10.14). Data visualization was facilitated by R packages, including SCP (v0.5.6), ggplot2 (v3.5.1), and ComplexHeatmap (v2.18.0). Univariate and multivariate Cox regression analyses were executed using the survival package (v3.5-8), while time-dependent AUC values were computed with the timeROC package (v0.4). Kaplan-Meier survival curves were generated using the survminer package (v0.4.9).

3 Results

3.1 Identification of malignant cells and characterization of heterogeneous expression programs

In this study, we integrated single-cell transcriptomic data from three datasets (GSE166555, GSE200997, and syn26844071), comprising 58 normal colorectal samples and 170 CRC samples. Following rigorous quality control and dimensionality reduction, a total of 320,475 cells were classified into 10 major cell types: B cells, T/NK cells, epithelial cells, plasma cells, fibroblasts, myeloid cells, endothelial cells, mast cells, mural cells, and enteric glial cells. Among these, T/NK cells were the most abundant (135,789 cells), followed by myeloid cells and fibroblasts (Figure 1A, Supplementary Figure S1-Supplementary Figure S2, and Supplementary Figure S3A-G). These refined annotations were applied to three high-resolution spatial transcriptomic datasets (ST1, ST2, ST3), enabling the visualization of the spatial distribution of different cell types within colorectal cancer tumors (Figure 1B).

To further investigate CRC heterogeneity, epithelial cell data were extracted from the comprehensive cell atlas. To ensure the purity of the epithelial cells, we re-annotated them using the SingleR and CellTypist algorithms, removing incorrectly classified cells (Supplementary Figure S4A-B). CNV scoring was performed on epithelial cells from tumor samples using the inferCNV algorithm, with normal epithelial cells serving as the reference. K-means clustering of the CNV score matrix revealed that epithelial cells from normal samples predominantly clustered in clusters 10, 15, and 25, exhibiting no significant CNV alterations. In contrast, epithelial cells from tumor samples showed clear gene copy number alterations, distinguishing them as malignant cells (Figure 1C, Supplementary Figure S4C). Malignant epithelial cells were identified by excluding clusters 10, 15, and 25 from the tumor samples.

Given the high heterogeneity of CRC cells, traditional clustering methods were insufficient to fully capture their complexity. Therefore, we applied the cNMF algorithm, which demonstrated high stability and low error when set to eight expression programs (Figure 1D). Consensus analysis confirmed the robustness of these eight expression programs, with substantial consistency across repeated experiments and outliers identified using a threshold of 0.05 (Figure 1E, Supplementary Figure S4D). These eight stable expression programs effectively captured the transcriptional characteristics of malignant CRC cells, providing a reliable framework for further analysis of CRC heterogeneity.

To visualize the spatial distribution of these MCEPs, we applied the AUCell algorithm to spatial transcriptomic data, scoring each sample based on the top 100 weight genes of each program. Enrichment analysis of the top 100 weight genes from each program was conducted, primarily referencing a gene set from the study by Barkley, D. et al. on pan-cancer tumor cell heterogeneity, supplemented with enrichment results from Hallmark Gene Sets and KEGG Pathways (32). This analysis revealed that MCEP 1, 2, and 7 were associated with stress responses. MCEP 1 was enriched in pathways related to hypoxia, antigen processing and presentation, chemokine signaling, and IL-17 signaling, while MCEP 2 was enriched in Wnt signaling. MCEP 7 was enriched in cell proliferation-related pathways, including the G2M checkpoint, mTORC1 signaling, and Myc targets V1. These programs were categorized as Inflammatory-Hypoxia Stress Expression Program (IHS-P), Wnt Signaling Stress Expression Program (Wnt-S-P), and Proliferation Stress Expression Program (PS-P), respectively. The spatial distribution of these MCEPs showed that IHS-P was prevalent in malignant and immune cell-rich regions, while Wnt-S-P and PS-P were more confined to malignant cells (Figures 1F, G).

Additionally, MCEP 3, 4, and 6 were associated with pEMT states. MCEP 3 was enriched in pEMT states and interferon responses, with higher spatial scores observed in both malignant and normal epithelial cells. MCEP 6, enriched in mesenchymal,

10.3389/fimmu.2025.1556386

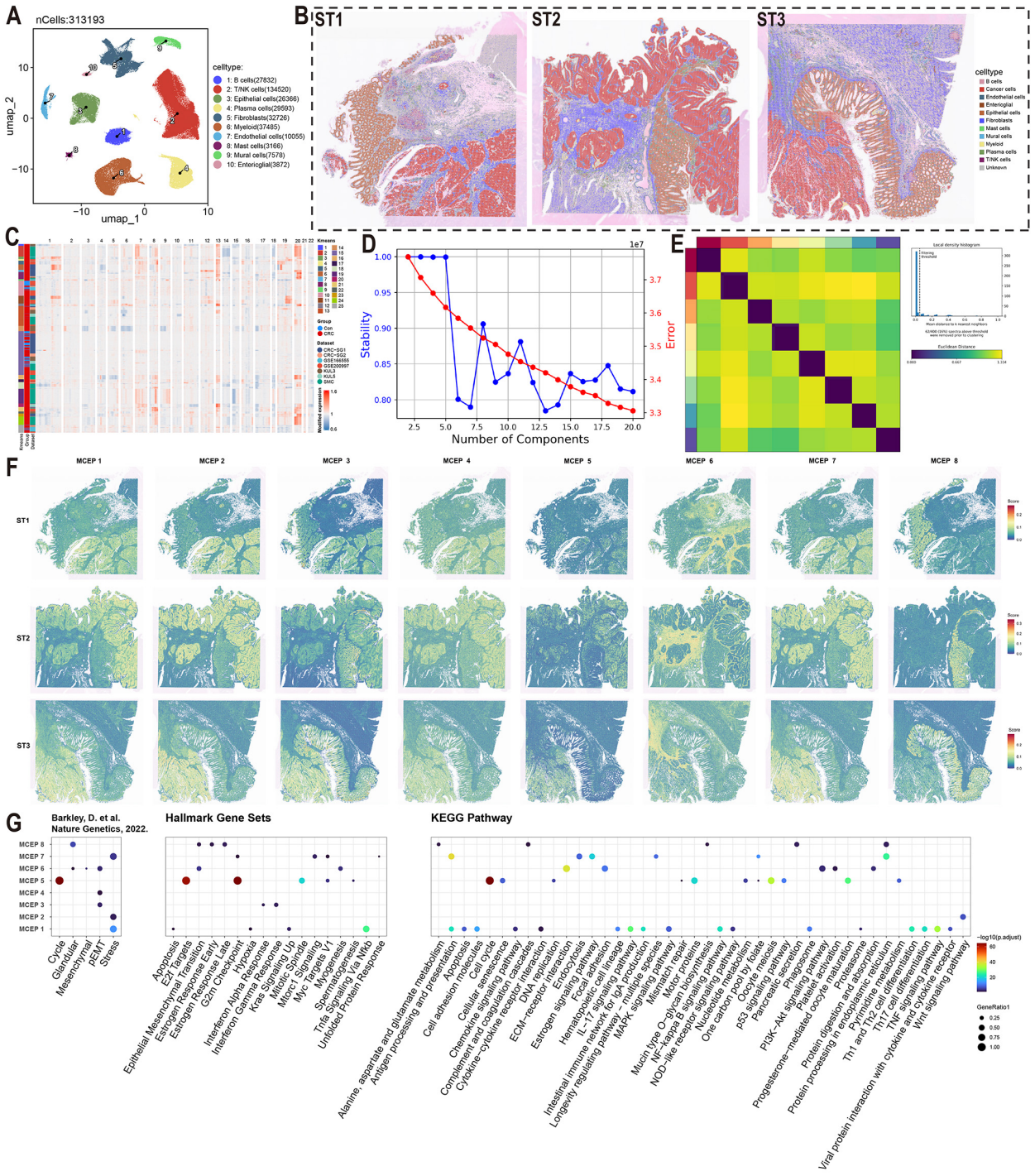


FIGURE 1
Functional characterization of malignant cell expression programs in colorectal cancer. **(A)** UMAP visualization of major cell types color-coded by cell lineage following quality control. **(B)** Spatial mapping of cell type distributions across three independent colorectal cancer specimens (ST1-3) using spatial transcriptomics. **(C)** Copy number variation (CNV) heatmap of epithelial cells stratified by k-means clustering (left panel). Tumor-derived cells (red) and normal counterparts (blue) are segregated based on chromosomal amplification (red) and deletion (blue) patterns. **(D)** Model selection curve demonstrating the optimal number of expression programs determined by consensus non-negative matrix factorization (CNMF), balancing stability and reconstruction error. **(E)** Consensus matrix establishing robust program identification. **(F)** Spatial activation patterns of MCEPs across tumor sections (ST1-3). **(G)** Functional enrichment analysis integrating pan-cancer malignant cell states (Barkley et al.), Hallmark gene sets, and KEGG pathways.

myogenesis, and ECM-receptor interaction pathways, displayed preferential spatial scores in the stromal compartment. Based on these findings, MCEPs 3, 4, and 6 were categorized as Inflammatory Epithelial-type pEMT Program (IE-pEMT-P), Intermediate Type pEMT Expression Program (I-pEMT-P), and Mesenchymal Type pEMT Expression Program (M-pEMT-P), respectively. The spatial distributions and enrichment results for these programs are shown in Figures 1F, G.

MCEP 5, enriched in cell cycle-related pathways such as Cell Cycle, E2F Targets, and G2M checkpoint, exhibited a dispersed spatial distribution across malignant and epithelial cells, and was categorized as the Cell Cycle Expression Program (CC-P). MCEP 8, primarily enriched in glandular and protein processing pathways in the endoplasmic reticulum, showed a preference for normal epithelial cells and was categorized as the Glandular Secretion Expression Program (GS-P). The spatial distributions and enrichment analyses for MCEP 5 and MCEP 8 are also shown in Figures 1F, G.

3.2 Crosstalk networks between malignant cells and immune cells mediated by differential MCEPs

To investigate the cell-cell interactions between malignant cells and immune cells, we first extracted each immune cell type (T/NK cells, B/plasma cells, and myeloid cells) from the comprehensive cell atlas for further detailed cell type annotation. T/NK cells were subdivided into 16 subpopulations, including CD4 Naive, CD4 Effector/Memory, and ILC; B/plasma cells were further categorized into 6 subpopulations, such as Naive B, Memory B, and IgA Plasma; Myeloid cells were divided into 10 subpopulations, including Macro_C1QC, Mast cells, and Mono_CD16 (Figure 2A, Supplementary Figure S5-7). Subsequently, pseudotime analysis was performed based on the secondary annotation results of each immune cell type and the stemness scores of each cell type, leading to the identification of genes associated with developmental trajectories in each immune cell population (Figure 2B, Supplementary Figure S8A).

These genes, associated with the pseudotime developmental trajectory of immune cell subsets, were used as target gene sets. For each MCEP, we selected the top 100 weighted genes in the expression programs as candidate regulators (Figure 2C). Among the three stress-related MCEPs, IHS-P had the highest number of regulatory factors, with HLA-DMA and PLAU affecting more target genes than other factors. In the three pEMT-related MCEPs, I-pEMT-P had the most regulatory factors, with TGFB1 having the greatest potential impact. Regulatory factors EDN1 and AREG were also abundant and shared between I-pEMT-P and IHS-P. In CC-P, HMGB1 had the most target genes, while TFF1 and WNT4 were more prominent in GS-P.

Regarding immune cell responses to MCEP crosstalk, TGFB1 and CALR were the main regulatory factors influencing T/NK cells, with TGFB1 originating from I-pEMT-P and CALR from IHS-P (Figure 2C, Supplementary Figure S9A, Supplementary Figure S10A). Notable downstream target genes of TGFB1 in T/NK cells included CCL3, FOXP3, and GZMB. For B/plasma cells, EDN1 and TGFB1 were the main regulatory factors, with EDN1 shared between IHS-P and I-pEMT-P (Figure 2C, Supplementary Figure S9B, Supplementary Figure S10B). Potential target genes of EDN1 in B cells included NCF1, PTPRC, and SLC2A3, while TGFB1 target genes included TIMP1, VIM, and CD38. In myeloid cells, the primary regulatory factors were TGFB1 and ANXA1, with ANXA1 originating from I-pEMT-P (Figure 2C, Supplementary Figure S9C, Supplementary Figure S10C). Potential target genes of TGFB1 in myeloid cells included ASB2, IGF1, and MMP9.

KEGG pathway enrichment analysis of the potential target genes in these immune cell subsets revealed significant biological insights (Figure 2D). The target genes of T/NK cells regulated by malignant cells were enriched in pathways such as Cytokine −cytokine receptor interaction, Th17 cell differentiation, and Chemokine signaling pathway, indicating a key role of cytokine networks in anti-tumor immune responses. The potential target genes of B/plasma cells were enriched in pathways such as Fc gamma R−mediated phagocytosis and Leukocyte transendothelial migration, suggesting their role in tumor-associated immunosuppression. In myeloid cells, the target genes regulated by malignant cells were enriched in IL−17 signaling pathway and TNF signaling pathway, highlighting their involvement in immune regulation and inflammation within the tumor microenvironment. These findings provide valuable biological insights for the development of future cancer therapies.

3.3 Crosstalk networks between malignant cells and stromal cells mediated by differential MCEPs

To investigate the effects of malignant cells on stromal cells, we performed detailed cell type annotation and stemness analysis on four stromal cell types: endothelial cells, mural cells, fibroblasts, and enteric glial cells, using methods similar to those employed for immune cell analysis (Figure 3A, Supplementary Figure S8B, Supplementary Figure S11-14). By integrating detailed annotations and stemness analysis, we reconstructed the developmental trajectories of these stromal cells and identified genes associated with their development (Figure 3B). We used high-weight genes from each MCEP as ligands to identify potential target genes in stromal cells associated with pseudotime trajectories, constructing a crosstalk network between malignant and stromal cells (Figure 3C).

Regarding regulatory factors in MCEPs affecting stromal cells, IHS-P had the highest number of potential regulatory factors, with PLAU affecting the most target genes. In Wnt-S-P, MIF was the only potential regulatory factor, while HSP90B1 and CDH1 were found in PS-P. Among the pEMT-related MCEPs, M-pEMT-P had more potential regulatory factors than the others, with BMP4 having the most target genes. I-pEMT-P's top regulatory factor was TGFB1, with AREG and EDN1 also shared with IHS-P. CC-P had two regulatory factors, HMGB1 and HMGB2, with HMGB1

10.3389/fimmu.2025.1556386

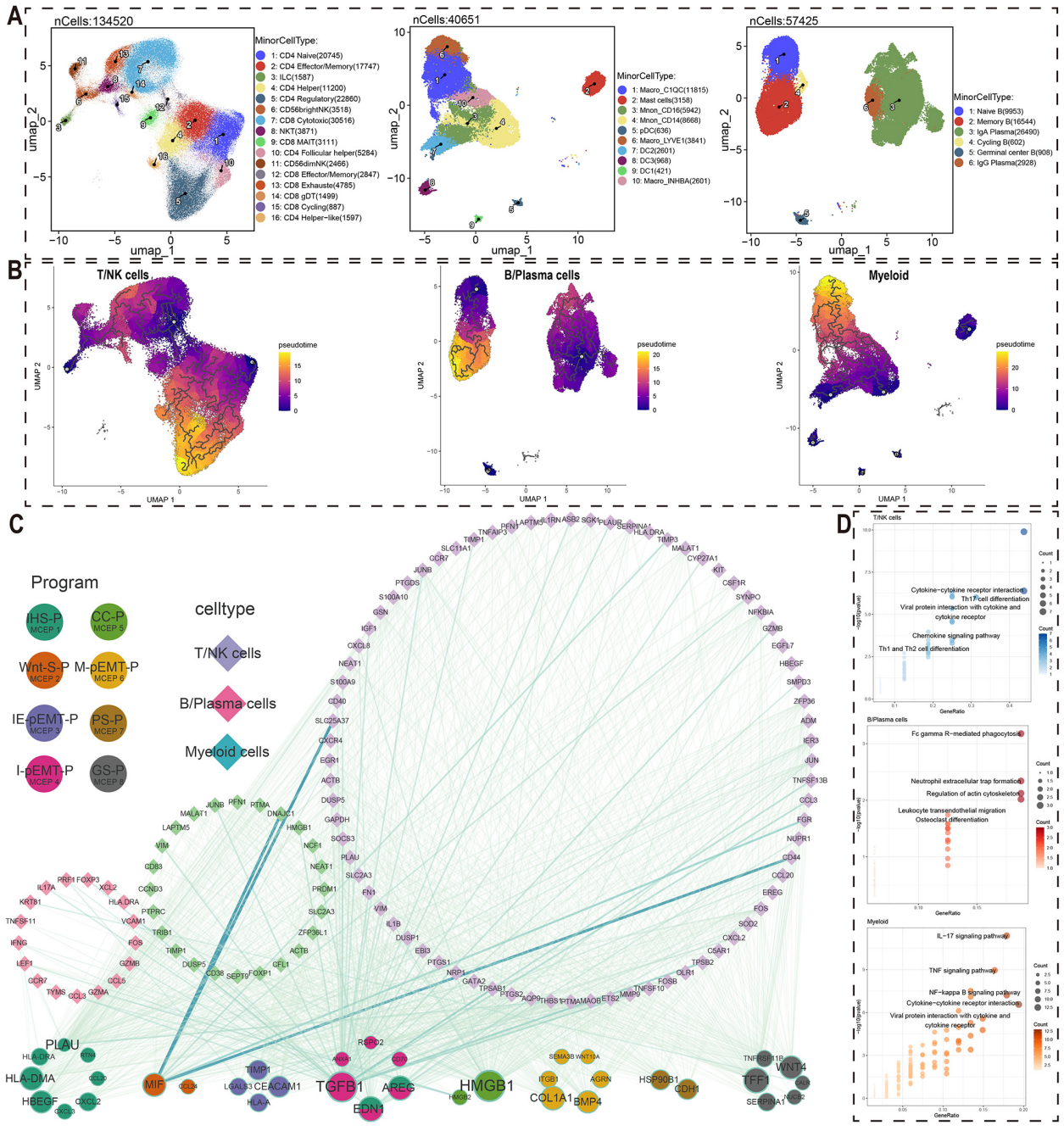


FIGURE 2
Crosstalk networks between MCEPs and immune cells. **(A)** Secondary dimensionality reduction annotation of three immune cell types (T/NK cells, B/Plasma cells, Myeloid cells). **(B)** Pseudotemporal trajectories reconstructed by Monocle3 for T/NK cells, plasma cells, and myeloid cells. **(C)** Ligand-receptor interaction network between MCEP-derived factors (circles, size scaled by target connectivity) and immune cell targets (diamonds, line width reflecting interaction strength). **(D)** Pathway enrichment analysis of target genes using hypergeometric testing, showing top five KEGG pathways per immune subset (point size: gene count; color intensity: -log10[P-value]).

affecting more target genes, although HMGB2 exhibited stronger interactions with certain stromal targets. TFF1 was the top regulatory factor in GS-P.

From a stromal cell perspective, the key regulatory factors for endothelial cells were HMGB2, TGFB1, and EDN1. HMGB2 target genes, associated with proliferative endothelial cells, included ASPM, AURKB, and BIRC5 (Figure 3C, Supplementary Figure S15A, Supplementary Figure S16A). TGFB1 and EDN1 target genes, including CTGF, EDN1, IGF1, and CALCRL, are mainly involved in angiogenesis. For mural cells, HMGB2, TGFB1, and EDN1 were the main regulatory factors, with HMGB2 targets such as FOXM1, KIF20A, and KIF2C, expressed in proliferative mural cells. TGFB1 and EDN1 targets included CDKN1A, CNN1, COL1A1, and EDNRB, contributing to cell proliferation and

10.3389/fimmu.2025.1556386

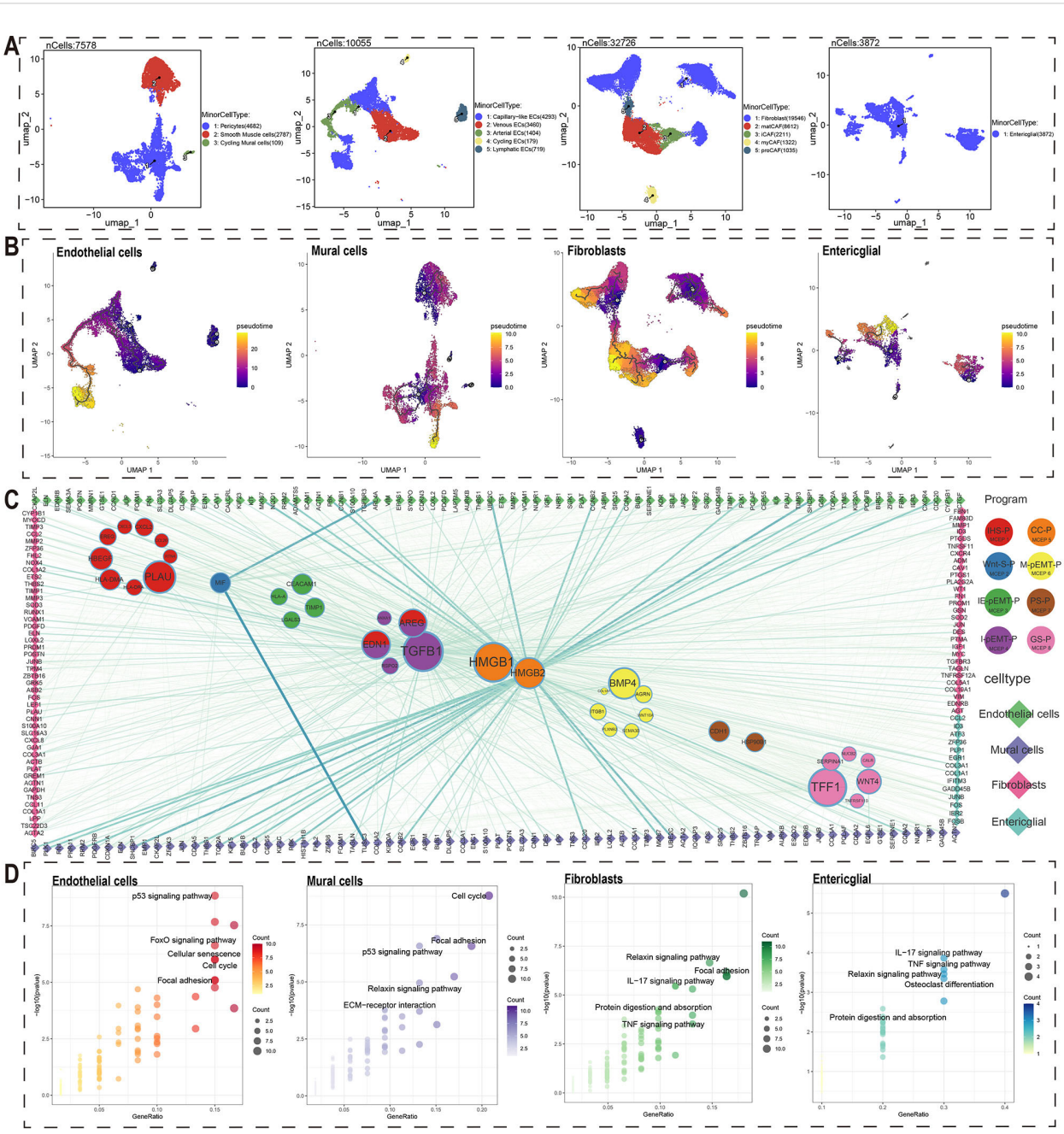


FIGURE 3
Crosstalk networks between MCEPs and stromal cells. **(A)** Secondary dimensionality reduction annotation of four stromal cell types (endothelial cells, mural cells, fibroblasts, and enteric glial cells). **(B)** Pseudo-temporal trajectory analysis of four stromal cell subtypes (endothelial cells, mural cells, fibroblasts, and enteric glial cells) shown through UMAP visualization. Color gradient (purple to yellow) indicates developmental progression from early to late stages. **(C)** Ligand-receptor interaction network between stromal cell-derived ligands (circles) and immune cell targets (diamonds). Node size corresponds to ligand-associated target quantity, line thickness represents interaction strength. **(D)** KEGG pathway enrichment of stromal cell target genes. Top five non-disease related pathways are displayed with point size indicating gene count and color intensity showing significance level (-log10[P-value]).

stromal stability (Figure 3C, Supplementary Figure S15B, Supplementary Figure S16B). In fibroblasts, HMGB2, TGFB1, and EDN1 were also key regulatory factors, with TGFB1 target genes including NOX4, THBS2, and DES (Figure 3C, Supplementary Figure S15C, Supplementary Figure S16C). Enteric glial cells had fewer potential crosstalk genes, with top regulatory factors ANXA1, TIMP1, and HLA-A, and target genes such as COL1A1 and COL3A1, which may support tumor structure and growth (Figure 3C, Supplementary Figure S15D, Supplementary Figure S16D). Overall, the primary regulatory factors influencing stromal cell crosstalk were HMGB2, TGFB1, and EDN1, with HMGB2 regulating cell cycle-related targets.

Additionally, KEGG pathway enrichment analysis of potential crosstalk target genes for each stromal cell type revealed significant

biological insights (Figure 3D). Endothelial cell targets were enriched in pathways such as the p53 signaling pathway, FoxO signaling pathway, Cellular Senescence, and Cell Cycle, suggesting their adaptability in the tumor microenvironment. Mural cell targets were enriched in the Cell Cycle, p53 signaling pathway, Focal Adhesion, Relaxin signaling pathway, and ECM-receptor interaction, emphasizing their roles in cell proliferation and matrix remodeling. Fibroblast targets were enriched in the Relaxin signaling pathway, Focal Adhesion, IL-17 signaling pathway, Protein Digestion and Absorption, and TNF signaling pathway, reflecting their dual role in immune regulation and matrix homeostasis. Enteric glial cell targets were enriched in IL-17 signaling, TNF signaling, Relaxin signaling, Osteoclast differentiation, and Protein Digestion and Absorption pathways, indicating their role in immune function and matrix support in the gut microenvironment.

3.4 MCEPs validation in CRC progression and development of MCEPs-related prognostic model

We conducted a validation study using the TCGA CRC cohort to explore the relationship between the 8 MCEPs and CRC progression. First, we merged two gene sets: 1) the top 100 weighted genes from each MCEP module, and 2) predicted target genes from the MCEP-immune/stromal cell interaction network. Differential expression analysis was then performed comparing tumor versus normal tissues. This analysis identified 323 upregulated genes and 215 downregulated genes (Figure 4A).

To validate the relationship between these MCEPs and CRC onset and progression, we conducted univariate Cox regression analysis and identified 75 differentially expressed genes (DEGs) associated with survival, including 26 risk genes and 49 protective genes (Figure 4B). Clustering analysis based on these genes divided the TCGA cohort into two subtypes (Figure 4C, Supplementary Figure S17). Survival analysis revealed significant differences between the subtypes, with patients in subtype C1 showing significantly higher survival rates compared to those in subtype C2 (Figure 4D). Chi-square tests indicated significant differences in tumor stage, lymph node metastasis, and distant metastasis, suggesting that tumors in the C2 subtype progressed more rapidly and were more prone to metastasis compared to those in the C1 subtype (Figure 4E).Genomic analysis revealed that the most frequently mutated genes in subtype C1 were APC (70%), KRAS (48%), and TP53 (47%) (Supplementary Figure S18A), while in subtype C2, the most frequently mutated genes were APC (81%), TP53 (74%), and TTN (44%) (Supplementary Figure S18B).

A prognostic model for assessing CRC patient survival was developed using the identified genes. LASSO regression analysis was performed to reduce the feature set from the 75 survival-related DEGs identified in the previous study to 29 genes at the minimum λ value ($\lambda = 0.0153$), including genes such as CLCA1, NPDC1, and MUC16 (Figures 4F, G). A backward stepwise Cox regression method was then applied to further reduce the feature set to 15 genes, with the regression coefficients visualized in a lollipop plot. The combination of LASSO and backward stepwise Cox regression methods enabled the identification of the most robust prognostic markers, minimizing overfitting while ensuring the model's predictive accuracy. Thus, these 15 genes were selected to establish the final prognostic model. Seven features had positive coefficients, with CLCA1 having the largest coefficient, while eight features had negative coefficients, with ITLN1 showing the largest absolute coefficient (Figure 4H).

In the training set, internal testing set, and external independent validation set, samples were divided into high-risk and low-risk groups based on the median risk score for each dataset. Significant survival differences were observed between the two groups (Figures 4J, K). In the training set, the AUC values for 1-year, 3-year, and 5-year survival were all greater than 0.7; in the internal testing set, the AUC value for 1-year survival was greater than 0.7, while those for 3-year and 5-year survival were above 0.65 (Figure 4I). The model also demonstrated excellent predictive performance in the independent validation set, with only GSE17536 showing a 5-year survival AUC value lower than 0.65. For all other datasets, the AUC values for 1-year, 3-year, and 5-year survival were all greater than 0.65. Notably, the GSE29621 dataset showed AUC values for 1-year, 3-year, and 5-year survival above 0.7, and the GSE38832 dataset exhibited even higher AUC values for all three survival endpoints, with values exceeding 0.75 (Figure 4I).

To further validate the prognostic prediction capability of this model, we assessed its ability to predict disease-free survival (DFS) in the GSE143985 and GSE161158 datasets. Samples were divided into risk groups based on the median predicted risk score, and significant differences in DFS were observed between the groups (Figure 4L). In GSE143985, the AUC values for 1-year and 3-year DFS were above 0.65, with the 5-year DFS AUC value approaching 0.65. In GSE161158, the corresponding AUC values for DFS were above 0.65 (Figure 4I). The model was further validated in the TCGA cohort for disease-specific survival (DSS), progression-free interval (PFI), and disease-free interval (DFI), showing excellent predictive performance for DSS and PFI, with significant differences in median survival times (Figure 4M). For DSS, the AUC values for 1-year, 3-year, and 5-year survival were all above 0.7, and for PFI, the AUC values were above 0.65 (Figure 4I). Notably, the model consistently achieved stable predictive accuracy across six independent validation cohorts (GSE17536, GSE17537, GSE29621, GSE38832, GSE143985, and GSE161158) and multiple clinical endpoints (OS, DFS, DSS, PFI), highlighting its strong generalizability to diverse patient populations and survival outcomes.

3.5 Multidimensional biological interpretation of the prognostic model

To gain further insights into the biological underpinnings of the prognostic model, the cellular abundance of various cell types in the TCGA cohort was first calculated using deconvolution methods. Next, the correlation between each gene in the prognostic model

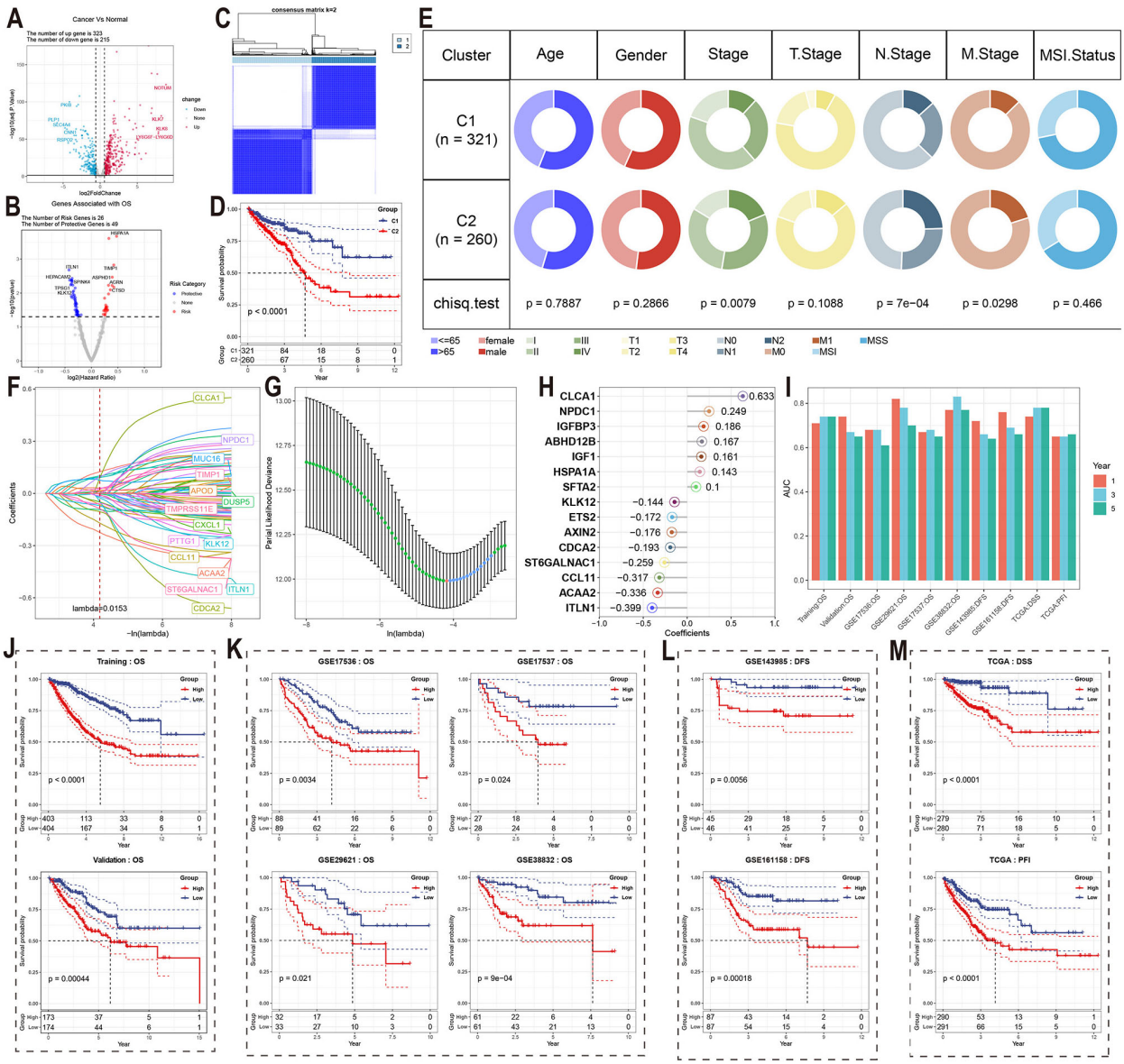


FIGURE 4
Validation of MCEPs in CRC progression and development of MCEPs-related prognostic models. **(A)** Volcano plot of differentially expressed genes (|FC| > 1.5, adjusted p < 0.05). **(B)** Univariate Cox survival analysis of prognostic genes (HR>1, red: risk factors; HR<1, blue: protective factors; p<0.05). **(C)** Consensus clustering matrix for two molecular subtypes. **(D)** Kaplan-Meier survival comparison between subtypes. **(E)** Clinical feature distribution across subtypes (χ^2 test). **(F)** LASSO coefficient profiles of candidate genes. **(G)** Optimal λ selection through 10-fold cross-validation (minimum deviance criterion). **(H)** Final model features with corresponding regression coefficients. **(I)** AUC values of the prognostic model across different datasets and prognostic indicators at 1-, 3-, and 5-year time points. **(J-M)** Kaplan-Meier survival curves stratified by median risk score in: **(J)** Training and internal validation sets; **(K)** External independent validation set; **(L)** DFS-specific dataset; **(M)** TCGA cohort with distinct survival endpoints.

and the cell scores was computed, revealing that CCL11, IGF1, and IGFBP3 were significantly correlated with multiple cell types. Specifically, these genes were positively correlated with cancer-associated fibroblasts, stromal score, and Tregs, while negatively correlated with tumor purity (Figure 6A).

The model was then further dissected at the single-cell level. Using genes with positive coefficients, a PosRiskScore for each cell was calculated, and similarly, a NegRiskScore was calculated using genes with negative coefficients. The total RiskScore for each cell was derived by computing the difference between PosRiskScore and NegRiskScore. The distribution of these scores was first visualized, and distinct distribution patterns for PosRiskScore and NegRiskScore were observed (Figure 6B). Specifically, PosRiskScore was found to be higher in endothelial cells and pre-cancer-associated fibroblasts (preCAFs), potentially linked to angiogenesis and epithelial-mesenchymal transition. In contrast, NegRiskScore was elevated in iCAFs, epithelial cells, normal fibroblasts, and myeloid immune cells, with NegRiskScore correlating with iCAFs and myeloid immune cells, which might reflect the inflammatory characteristics of the tumor microenvironment. Higher scores in epithelial cells were also

10.3389/fimmu.2025.1556386

observed, which could be indicative of a more epithelial-like phenotype associated with partial EMT processes (Figure 6C).

Furthermore, the analysis was extended to the spatial transcriptomics level. It was shown that PosRiskScore was predominantly localized in the stromal regions of malignant cell areas, while NegRiskScore was mainly concentrated in the epithelial regions. Consequently, the final RiskScore had the lowest score in the epithelial areas and the highest score in the stromal regions, with similar distribution patterns observed across three samples (Figure 6D). Overall, the positive coefficient features in the prognostic model were likely to represent higher levels of mesenchymal traits associated with pEMT, while the negative coefficient features were likely linked to a more inflammatory microenvironment and epithelial characteristics of pEMT. Thus, the final RiskScore reflected the relative balance between epithelial-mesenchymal features and the degree of inflammation in the

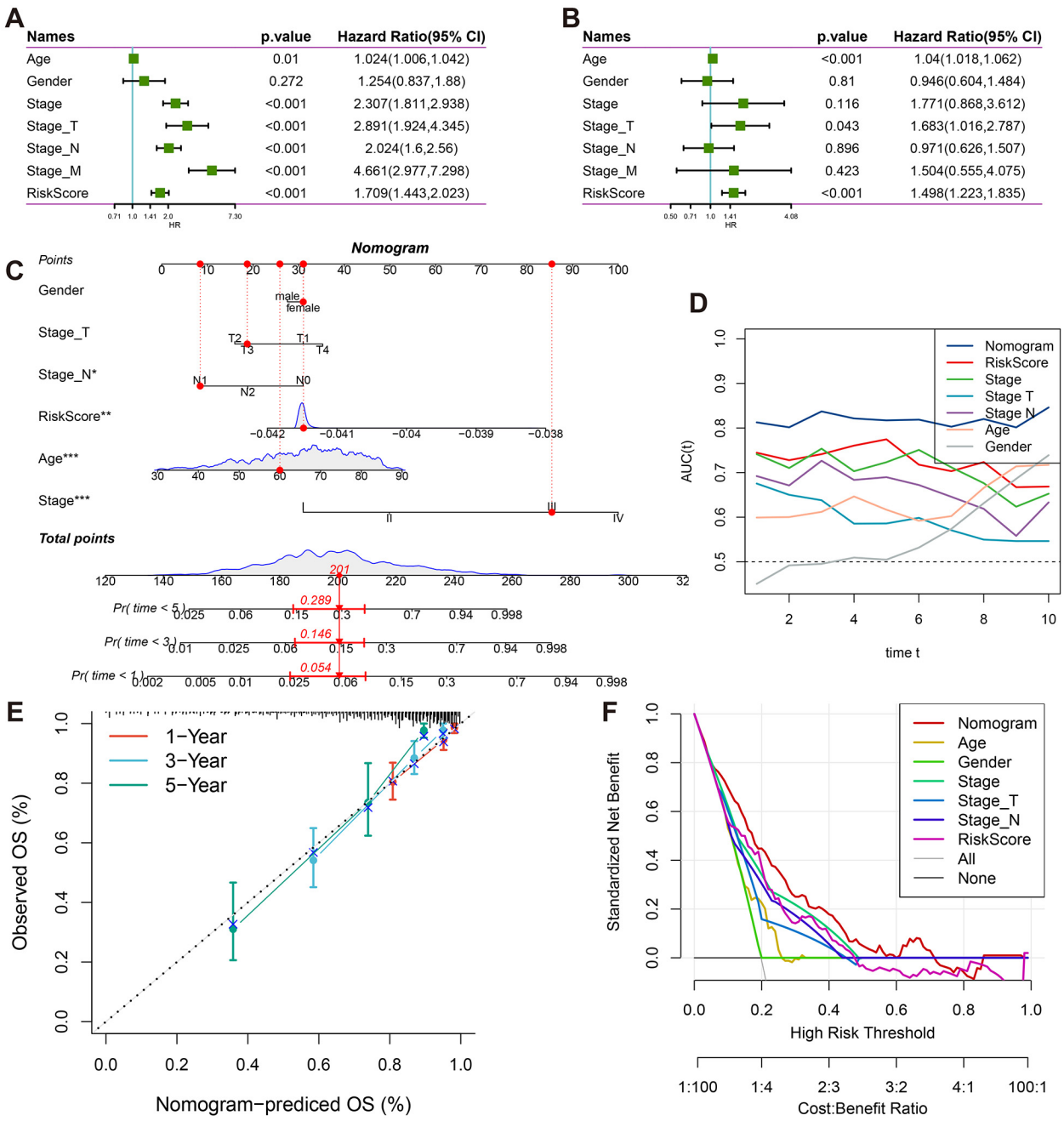


FIGURE 5
Clinical prognostic value and nomogram construction. **(A)** Forest plots of univariate Cox regression analyses for RiskScore and clinicopathological parameters (gender, age, tumor stage). **(B)** Forest plots of multivariate Cox regression analyses for RiskScore and clinicopathological parameters (gender, age, tumor stage). **(C)** Clinical nomogram integrating T/N staging, tumor stage, age, gender, and RiskScore. **(D)** Time-dependent ROC analysis (1-10 years) for nomogram performance. **(E)** Calibration curves comparing predicted vs observed survival probabilities at 1/3/5 years. **(F)** Decision curve analysis evaluating clinical utility across threshold probabilities.

10.3389/fimmu.2025.1556386

tumor microenvironment, offering valuable insights into patient prognosis.

3.6 Integration of risk score and clinical features to construct a nomogram for prognosis prediction

To enhance the prognostic accuracy and clinical applicability of the model, univariate Cox regression analysis was performed on age, gender, clinical stage, Stage_T, Stage_N, Stage_M, and RiskScore (Figure 5A). Significant survival risk factors were identified for all features except gender. In multivariate Cox regression analysis, age, Stage_T, and RiskScore were found to be independently associated with survival, confirming RiskScore as an independent prognostic factor (Figure 5B).

A nomogram was subsequently constructed, incorporating age, gender, clinical stage, Stage_T, Stage_N, and RiskScore (Figure 5C). It was demonstrated that the nomogram improved clinical decision-making compared to traditional staging systems through three key mechanisms: First, continuous risk quantification allowed

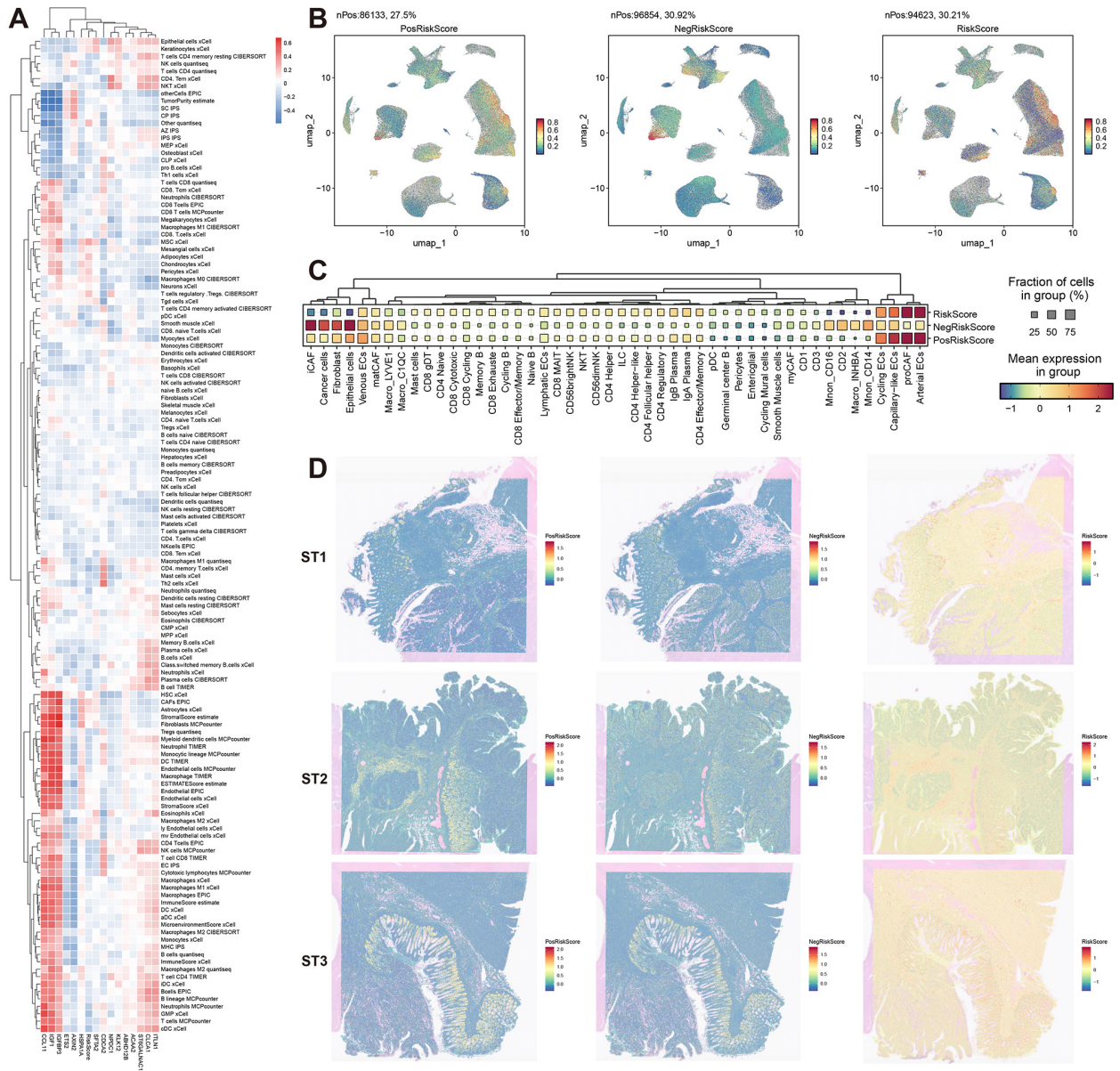


FIGURE 6
Multidimensional biological interpretation of prognostic signatures. **(A)** Spearman correlation heatmap between model genes and deconvoluted immune cell populations (red: positive, blue: negative). **(B)** Single-cell UMAP projections visualizing risk-associated signatures: PosRiskScore (positive coefficient genes), NegRiskScore (negative coefficient genes), and composite RiskScore. **(C)** Dot plot displaying cell type-specific enrichment of risk signatures (dot size: scoring cell proportion; color intensity: score magnitude). **(D)** Spatial distribution patterns of risk signatures across three representative specimens (ST1-ST3).

for more precise stratification of patient outcomes than categorical staging classifications. Second, the multidimensional integration of molecular risk scores with clinicopathological parameters provided complementary prognostic information that surpassed the limitations of anatomical staging alone. Third, the dynamic estimation of survival probability for specific timepoints (1-10 years) facilitated personalized follow-up planning and therapeutic decision-making. Stage_M was excluded from the analysis due to collinearity with overall stage.

Excellent predictive performance was demonstrated by the nomogram, with AUC values exceeding 0.8 for survival predictions at 1, 3, 5, and 10 years (Figure 5D). Strong agreement between predicted and actual survival probabilities was observed in calibration curves for 1, 3, and 5 years (Figure 5E). Clinical decision curve analysis revealed that the nomogram consistently provided higher net benefits across various threshold probabilities when compared to both individual clinical parameters and traditional staging systems (Figure 5F). The enhanced clinical utility of the nomogram was attributed to its ability to synthesize molecular biomarkers with conventional staging data, addressing the heterogeneity within traditional stage categories and enabling more individualized risk assessment. These findings collectively validated the effectiveness and clinical applicability of the proposed model.

3.7 Potential drug therapeutic targets based on MCEPs

To identify actionable therapeutic targets in CRC, we systematically analyzed 538 DEGs through PPI network construction. Four distinct topological algorithms (MNC, MCC, DMNC, Degree) were employed to prioritize the top 100 hub genes from the PPI network. Subsequent survival impact analysis revealed that TIMP1 and IGF1 emerged as prognostic risk genes among these hub genes. Notably, TIMP1 exhibited consistent identification across all four algorithms, whereas IGF1 was only captured by MNC and Degree algorithms (Figure 7A). Based on its algorithm-independent prioritization and significant association with poor prognosis, TIMP1 was selected as the principal therapeutic target for further investigation.

Pan-cancer expression profiling demonstrated significant TIMP1 upregulation in 15 malignancies (including colorectal adenocarcinoma [COAD], breast invasive carcinoma [BRCA], and cholangiocarcinoma [CHOL] as representative examples), while downregulation was observed in 10 cancer types (exemplified by kidney chromophobe [KICH] and lung squamous cell carcinoma [LUSC]) with no significant alterations detected in other malignancies (Figure 7B). Immunohistochemical validation via the Human Protein Atlas confirmed elevated TIMP1 protein levels in CRC, breast cancer, glioma, hepatocellular carcinoma, and gastric adenocarcinoma (Figure 7C), underscoring its pan-cancer relevance.

Virtual screening of 2,000 bioactive compounds against the TIMP1 structure identified Venetoclax (ΔG = -12.236 kcal/mol) and Lumacaftor (ΔG = -12.129 kcal/mol) as top candidates with superior binding affinities (Figure 7D). Molecular docking simulations predicted stable interactions between these compounds and key TIMP1 functional domains.

These findings computationally nominate TIMP1 as a multi-cancer therapeutic target, with the identified small-molecule inhibitors warranting preclinical evaluation for targeted therapy development in CRC and other TIMP1-driven malignancies.

4 Discussion

In this study, we re-examined the biological characteristics of CRC by leveraging prior research on malignant cell transcriptional signatures and identified eight major MCEPs (32). These programs encompass three stress-related categories (hypoxia-inflammation, Wnt-related, and proliferation), three EMT subtypes (inflammatory epithelial, intermediate, and mesenchymal), one cell cycle category, and one glandular secretion category. Each program is critically linked to functional roles in regulating malignant cell proliferation, migration, drug resistance, metastasis, and patient prognosis (33–36). Traditional molecular subtyping approaches, such as those based on hypoxic metabolism, cellular senescence, or microenvironmental cell markers (37–39), often oversimplify tumor heterogeneity. Solid tumors are multifactorial systems, and reliance on binary phenotypic classifications risks underestimating inter-individual variability and obscuring underlying biological processes, thereby limiting the molecular interpretability of subtypes.

To address this, we employed a programmatic state-based framework to characterize CRC gene expression, accounting for potential confounders and mutual exclusivity between states. Importantly, we emphasized continuity within each state rather than discrete isolation. For instance, malignant cell partial EMT was defined as a tripartite continuum (mesenchymal, intermediate, and epithelial), aligning with the evolving concept of "epithelial-mesenchymal plasticity" endorsed by the International EMT Association (40). The tumor microenvironment, a complex ecosystem sculpted predominantly by malignant cells, has historically been analyzed by grouping tumor cells homogeneously or partitioning them into static clusters. In contrast, our crosstalk analysis originated from malignant cell expression programs, enabling simultaneous exploration of heterogeneity in both malignant and stromal/immune compartments.

In our analysis of the eight MCEPs, we identified critical regulators with potential crosstalk interactions in immune/stromal compartments, including TGFβ1 and HMGB1. Functional annotation of downstream target genes in immune/stromal cells revealed biological roles consistent with established mechanisms. Specifically, TGFβ1 signaling dysregulation plays a pivotal role in colorectal carcinogenesis by governing cell growth, differentiation, migration, and apoptosis (41–43). Pathological overexpression of

10.3389/fimmu.2025.1556386

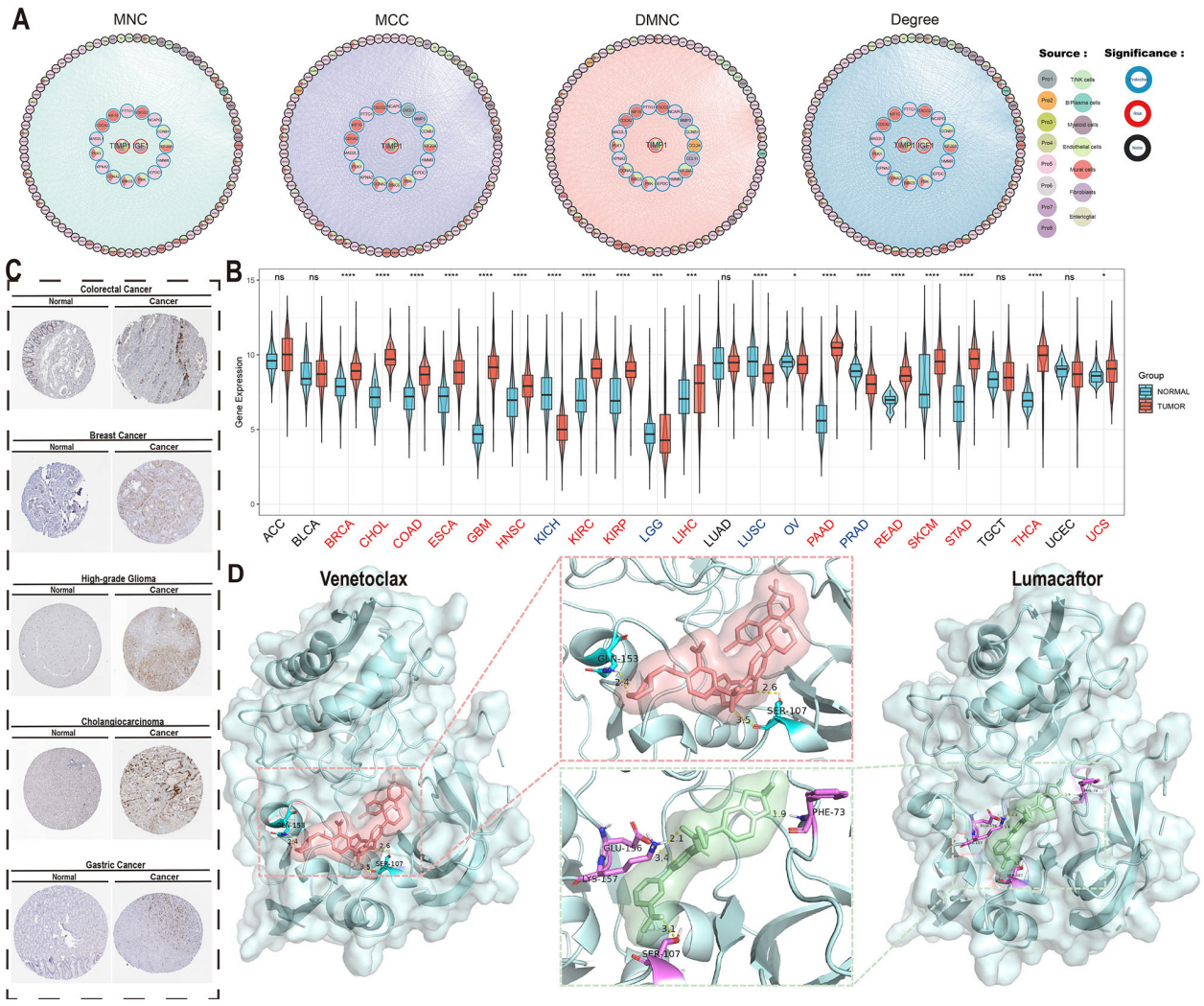


FIGURE 7
Core determinant identification and therapeutic exploration. **(A)** Protein-protein interaction (PPI) network of top 100 survival-associated genes identified through MNC/MCC/DMNC/Degree algorithms (border color: survival association; fill color: algorithm source). **(B)** TIMP1 differential expression across TCGA tumor types versus normal tissues (ns, P > 0.05; *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 0.0001). **(C)** TIMP1 immunohistochemical validation in multiple carcinomas and paired normal tissues. **(D)** Molecular docking of TIMP1 (PDB:3V96) with Lumacaftor (ΔG = -12.13 kcal/mol) and Umbralisib (ΔG = -11.80 kcal/mol), showing binding pocket configurations.

TGFβ1 drives epithelial-mesenchymal transition, extracellular matrix remodeling, and cancer-associated fibroblast activation (44–46). Notably, TGFβ1 emerged as a key regulator in the I-pEMT-P program, targeting immune cell genes including FOXP3, CD38, and MMP9—established mediators of immune evasion and immunosuppressive TME remodeling (47–49). In stromal compartments, TGFβ1 may further facilitate CAF transformation and immunosuppressive functions through NOX4-mediated pathways (50).Meanwhile, nuclear HMGB1 functions as a chromatin-binding factor regulating nucleosome organization, transcriptional control, and genomic stability, whereas extracellular HMGB1 modulates cell differentiation, metastatic dissemination, and apoptosis (51). Concurrently, HMGB2 within the CC-P program demonstrated regulatory effects on mesenchymal-like cells, modulating pro-angiogenic genes such as AURKB, BIRC5, and FOXM1 that coordinate endothelial and vascular smooth muscle cell proliferation (52–54). This integrated regulatory network analysis reveals how malignant cell-derived signals orchestrate multicellular ecosystem dynamics through conserved molecular pathways, providing mechanistic insights into TME reprogramming during CRC progression.

CRC prognosis remains challenging due to pronounced tumor heterogeneity. Existing prognostic models, often anchored to singular features (e.g., immune, EMT, or metabolic signatures), provide incomplete assessments. Our integrative model, combining immune and stromal features, offers enhanced biological interpretability. Risk stratification revealed that high-risk scores correlate with mesenchymal-like, immunosuppressive TMEs

enriched in CAFs, Tregs, and inflammatory markers. Conversely, low-risk scores associate with epithelial-like phenotypes marked by partial EMT, reduced stromal activation, and preserved epithelial integrity. The model incorporates 15 genes, with CLCA1 and ITLN1 exhibiting the strongest prognostic weights. CLCA1, a tumor suppressor, inhibits CRC progression by suppressing Wnt/β-catenin signaling and EMT, consistent with its reduced expression in advanced tumors and inverse correlation with metastasis (55). ITLN1, conversely, antagonizes tumor neovascularization and MDSC accumulation via IL-17D/CXCL2 axis modulation, thereby reshaping the immunosuppressive TME—a mechanism aligning with its prognostic significance in both CRC and ovarian cancer (56, 57). Additional contributors, such as IGFBP3 and ACAA2, further underscore the multifactorial nature of CRC heterogeneity. Elevated IGFBP3, driven by genetic predisposition, may enhance CRC risk through IGF1-mediated mitogenic signaling, as supported by Mendelian randomization analyses (58). ACAA2, a fatty acid metabolism enzyme, inversely correlates with cetuximab resistance, particularly in KRAS-mutant CRC, suggesting its role in metabolic adaptation and therapy response regulation (59). This framework bridges molecular mechanisms to clinical outcomes, providing biological interpretability to the prognostic model.

As an independent prognostic factor, our model achieved an AUC >0.8 for 10-year outcome prediction when combined with clinical variables. Integration with TNM staging via a nomogram improves CRC management by enabling dynamic survival probability estimation (1–10 years), optimizing adjuvant therapy selection, surveillance intervals, and resource allocation.

PPI network analysis identified TIMP1 as a hub gene within the I-pEMT-P program. TIMP1, a matrix metalloproteinase inhibitor, exhibits context-dependent roles in cancer. In brain metastases, astrocyte-derived TIMP1 suppresses $CD8^+$ T cell activity (60), while in pancreatic cancer, TIMP1-CD63-ERK signaling drives neutrophil extracellular trap formation and tumor progression (61). In CRC, TIMP1 correlates with tumor cell proliferation, invasion, and poor prognosis (62). Our data suggest that the I-pEMT-P program may remodel the stromal niche via TIMP1, influencing tumor progression and clinical outcomes.

4.1 Limitations and future directions

Despite the significant findings, this study has some limitations. Although single-cell data from over 100 samples were analyzed, the lack of clinical annotations, such as tumor stage, survival time, and survival status, restricted our ability to directly correlate expression programs with tumor progression and patient outcomes. Therefore, we relied on bulk RNA-seq datasets, which included complete clinical information. Additionally, while computational predictions identified key regulators, such as TGFβ1 and HMGB2, in stromal/immune modulation, their mechanistic roles remain unvalidated experimentally. Future studies should employ co-culture models or *in vivo* systems to confirm these interactions.

5 Conclusion

This study identified eight distinct MCEPs that characterize the transcriptional states of CRC malignant cells. We constructed interaction networks between these MCEPs and immune or stromal cells, which led to the development of a prognostic model consisting of 15 genes. Furthermore, TIMP1 was identified as a key gene, and two potential drugs, Venetoclax and Lumacaftor, were highlighted for targeted therapeutic strategies. In summary, this study provides new insights and references for CRC heterogeneity and prognostic therapy.

Data availability statement

Publicly available datasets were analyzed in this study. Gene Expression Omnibus (GEO): Datasets such as GSE166555, GSE200997, GSE39582, GSE17536, GSE17537, GSE29621, GSE38832, GSE143985, and GSE161158 are available in the GEO database. Synapse Database: The dataset with accession number syn26844071 is hosted on Synapse.10x Genomics: The Visium HD colorectal cancer sequencing datasets can be accessed on the 10x Genomics website. The Cancer Genome Atlas (TCGA) and UCSC Xena: TCGA provides bulk RNA-seq data, SNV data, pan-cancer clinical data, and pan-cancer gene expression data. The Human Protein Atlas: Immunohistochemistry images are available from The Human Protein Atlas.

Ethics statement

This study exclusively used publicly available data; therefore, ethical approval was not required. The studies were conducted in accordance with the local legislation and institutional requirements. The data used in this study are all from publicly available databases, and the specific data sources are listed in the Materials and Methods. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements. Ethical approval was not required for the study involving animals in accordance with the local legislation and institutional requirements because This study exclusively used publicly available data; therefore, ethical approval was not required.

Author contributions

TW: Data curation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. ZC: Data curation, Writing – review & editing. WW: Writing – review & editing. HW: Writing – review & editing. SL: Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Science Foundation of Chongqing City [grant numbers CSTB2023NSCQ-MSX0241], Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJQN202300433).

Acknowledgments

The computing work in this paper was partly supported by the Supercomputing Center of Chongqing Medical University.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2025.1556386/full#supplementary-material

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. (2021) 71(3):209–49. doi: 10.3322/caac.21660

2. Dienstmann R, Vermeulen L, Guinney J, Kopetz S, Tejpar S, Tabernero J. Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nat Rev Cancer*. (2017) 17(2):79–92. doi: 10.1038/nrc.2016.126

3. Network CGA. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. (2012) 487(7407):330–7. doi: 10.1038/nature11252

4. Bakhoum SF, Cantley LC. The multifaceted role of chromosomal instability in cancer and its microenvironment. *Cell*. (2018) 174(6):1347–60. doi: 10.1016/j.cell.2018.08.027

5. Lengauer C, Kinzler KW, Vogelstein B. Genetic instability in colorectal cancers. *Nature*. (1997) 386(6625):623–7. doi: 10.1038/386623a0

6. Taieb J, Svrcek M, Cohen R, Basile D, Tougeron D, Phelip JM. Deficient mismatch Repair/Microsatellite unstable colorectal cancer: Diagnosis, prognosis and treatment. *Eur J Cancer*. (2022) 175:136–57. doi: 10.1016/j.ejca.2022.07.020

7. Mei WJ, Mi M, Qian J, Xiao N, Yuan Y, Ding PR. Clinicopathological characteristics of high microsatellite Instability/Mismatch repair-deficient colorectal cancer: A narrative review. *Front Immunol*. (2022) 13:1019582. doi: 10.3389/fimmu.2022.1019582

8. Li J, Wu C, Hu H, Qin G, Wu X, Bai F, et al. Remodeling of the immune and stromal cell compartment by PD-1 blockade in mismatch repair-deficient colorectal cancer. *Cancer Cell*. (2023) 41(6):1152–69.e7. doi: 10.1016/j.ccell.2023.04.011

9. Weng J, Li S, Zhu Z, Liu Q, Zhang R, Yang Y, et al. Exploring immunotherapy in colorectal cancer. *J Hematol Oncol*. (2022) 15(1):95. doi: 10.1186/s13045-022-01294-4

10. Westcott PMK, Muyas F, Hauck H, Smith OC, Sacks NJ, Ely ZA, et al. Mismatch repair deficiency is not sufficient to elicit tumor immunogenicity. *Nat Genet*. (2023) 55(10):1686–95. doi: 10.1038/s41588-023-01499-4

11. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med*. (2015) 21(11):1350–6. doi: 10.1038/nm.3967

12. Lei Y, Tang R, Xu J, Wang W, Zhang B, Liu J, et al. Applications of single-cell sequencing in cancer research: Progress and perspectives. *J Hematol Oncol*. (2021) 14(1):91. doi: 10.1186/s13045-021-01105-2

13. Ren X, Zhang L, Zhang Y, Li Z, Siemers N, Zhang Z. Insights gained from single-cell analysis of immune cells in the tumor microenvironment. *Annu Rev Immunol*. (2021) 39:583–609. doi: 10.1146/annurev-immunol-110519-071134

14. Ding S, Chen X, Shen K. Single-cell RNA sequencing in breast cancer: Understanding tumor heterogeneity and paving roads to individualized therapy. *Cancer Commun (Lond)*. (2020) 40(8):329–44. doi: 10.1002/cac2.12078

15. Liu Y, Zhang Q, Xing B, Luo N, Gao R, Yu K, et al. Immune phenotypic linkage between colorectal cancer and liver metastasis. *Cancer Cell*. (2022) 40(4):424–37.e5. doi: 10.1016/j.ccell.2022.02.013

16. Wang R, Li J, Zhou X, Mao Y, Wang W, Gao S, et al. Single-cell genomic and transcriptomic landscapes of primary and metastatic colorectal cancer tumors. *Genome Med*. (2022) 14(1):93. doi: 10.1186/s13073-022-01093-z

17. Liu Z, Hu Y, Xie H, Chen K, Wen L, Fu W, et al. Single-cell chromatin accessibility analysis reveals the epigenetic basis and signature transcription factors for the molecular subtypes of colorectal cancers. *Cancer Discov*. (2024) 14(6):1082–105. doi: 10.1158/2159-8290.Cd-23-1445

18. Long F, Wang W, Li S, Wang B, Hu X, Wang J, et al. The potential crosstalk between tumor and plasma cells and its association with clinical outcome and immunotherapy response in bladder cancer. *J Transl Med*. (2023) 21(1):298. doi: 10.1186/s12967-023-04151-1

19. Song H, Weinstein HNW, Allegakoen P, Wadsworth MH 2nd, Xie J, Yang H, et al. Single-cell analysis of human primary prostate cancer reveals the heterogeneity of tumor-associated epithelial cell states. *Nat Commun*. (2022) 13(1):141. doi: 10.1038/s41467-021-27322-4

20. Li C, Song W, Zhang J, Luo Y. Single-cell transcriptomics reveals heterogeneity in esophageal squamous epithelial cells and constructs models for predicting patient prognosis and immunotherapy. *Front Immunol*. (2023) 14:1322147. doi: 10.3389/fimmu.2023.1322147

21. Marjanovic ND, Hofree M, Chan JE, Canner D, Wu K, Trakala M, et al. Emergence of a high-plasticity cell state during lung cancer evolution. *Cancer Cell*. (2020) 38(2):229–46.e13. doi: 10.1016/j.ccell.2020.06.012

22. Uhlitz F, Bischoff P, Peidli S, Sieber A, Trinks A, Lüthen M, et al. Mitogen-activated protein kinase activity drives cell trajectories in colorectal cancer. *EMBO Mol Med*. (2021) 13(10):e14123. doi: 10.15252/emmm.202114123

23. Khaliq AM, Erdogan C, Kurt Z, Turgut SS, Grunvald MW, Rand T, et al. Refining colorectal cancer classification and clinical stratification through a single-cell atlas. *Genome Biol*. (2022) 23(1):113. doi: 10.1186/s13059-022-02677-z

24. Joanito I, Wirapati P, Zhao N, Nawaz Z, Yeo G, Lee F, et al. Single-cell and bulk transcriptome sequencing identifies two epithelial tumor cell states and refines the consensus molecular classification of colorectal cancer. *Nat Genet*. (2022) 54(7):963–75. doi: 10.1038/s41588-022-01100-4

25. Oliveira MF, Romero JP, Chung M, Williams S, Gottscho AD, Gupta A, et al. Characterization of immune cell populations in the tumor microenvironment of colorectal cancer using high definition spatial profiling. *bioRxiv*. (2024) 2024.06.04.597233. doi: 10.1101/2024.06.04.597233

26. Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, et al. Gene expression classification of colon cancer into molecular subtypes: Characterization,

validation, and prognostic value. *PloS Med*. (2013) 10(5):e1001453. doi: 10.1371/journal.pmed.1001453

27. Smith JJ, Deane NG, Wu F, Merchant NB, Zhang B, Jiang A, et al. Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology*. (2010) 138(3):958–68. doi: 10.1053/j.gastro.2009.11.005

28. Chen DT, Hernandez JM, Shibata D, McCarthy SM, Humphries LA, Clark W, et al. Complementary strand microRNAs mediate acquisition of metastatic potential in colonic adenocarcinoma. *J Gastrointest Surg*. (2012) 16(5):905–12. doi: 10.1007/s11605-011-1815-0

29. Tripathi MK, Deane NG, Zhu J, An H, Mima S, Wang X, et al. Nuclear factor of activated T-cell activity is associated with metastatic capacity in colon cancer. *Cancer Res*. (2014) 74(23):6947–57. doi: 10.1158/0008-5472.Can-14-1592

30. Shinto E, Yoshida Y, Kajiwara Y, Okamoto K, Mochizuki S, Yamadera M, et al. Clinical significance of a gene signature generated from tumor budding grade in colon cancer. *Ann Surg Oncol*. (2020) 27(10):4044–54. doi: 10.1245/s10434-020-08498-3

31. Szeglin BC, Wu C, Marco MR, Park HS, Zhang Z, Zhang B, et al. A SMAD4-modulated gene profile predicts disease-free survival in stage II and III colorectal cancer. *Cancer Rep (Hoboken)*. (2022) 5(1):e1423. doi: 10.1002/cnr2.1423

32. Barkley D, Moncada R, Pour M, Liberman DA, Dryg I, Werba G, et al. Cancer cell states recur across tumor types and form specific interactions with the tumor microenvironment. *Nat Genet*. (2022) 54(8):1192–201. doi: 10.1038/s41588-022-01141-9

33. Nunes L, Li F, Wu M, Luo T, Hammarström K, Torell E, et al. Prognostic genome and transcriptome signatures in colorectal cancers. *Nature*. (2024) 633 (8028):137–46. doi: 10.1038/s41586-024-07769-3

34. Sui Q, Zhang X, Chen C, Tang J, Yu J, Li W, et al. Inflammation promotes resistance to immune checkpoint inhibitors in high microsatellite instability colorectal cancer. *Nat Commun*. (2022) 13(1):7316. doi: 10.1038/s41467-022-35096-6

35. Zhao H, Ming T, Tang S, Ren S, Yang H, Liu M, et al. Wnt signaling in colorectal cancer: Pathogenic role and therapeutic target. *Mol Cancer*. (2022) 21(1):144. doi: 10.1186/s12943-022-01616-7

36. Sabouni E, Nejad MM, Mojtabavi S, Khoshduz S, Mojtabavi M, Nadafzadeh N, et al. Unraveling the function of epithelial-mesenchymal transition (EMT) in colorectal cancer: Metastasis, therapy response, and revisiting molecular pathways. *BioMed Pharmacother*. (2023) 160:114395. doi: 10.1016/j.biopha.2023.114395

37. Huang A, Sun Z, Hong H, Yang Y, Chen J, Gao Z, et al. Novel hypoxia- and lactate metabolism-related molecular subtyping and prognostic signature for colorectal cancer. *J Transl Med*. (2024) 22(1):587. doi: 10.1186/s12967-024-05391-5

38. Feng J, Fu F, Nie Y. Comprehensive genomics analysis of aging related gene signature to predict the prognosis and drug resistance of colon adenocarcinoma. *Front Pharmacol*. (2023) 14:1121634. doi: 10.3389/fphar.2023.1121634

39. Bu F, Zhao Y, Zhao Y, Yang X, Sun L, Chen Y, et al. Distinct tumor microenvironment landscapes of rectal cancer for prognosis and prediction of immunotherapy response. *Cell Oncol (Dordr)*. (2022) 45(6):1363–81. doi: 10.1007/s13402-022-00725-1

40. Yang J, Antin P, Berx G, Blanpain C, Brabletz T, Bronner M, et al. Guidelines and definitions for research on epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol*. (2020) 21(6):341–52. doi: 10.1038/s41580-020-0237-9

41. Itatani Y, Kawada K, Sakai Y. Transforming growth factor-β signaling pathway in colorectal cancer and its tumor microenvironment. *Int J Mol Sci*. (2019) 20(23):5822. doi: 10.3390/ijms20235822

42. Jung B, Staudacher JJ, Beauchamp D. Transforming growth factor β superfamily signaling in development of colorectal cancer. *Gastroenterology*. (2017) 152(1):36–52. doi: 10.1053/j.gastro.2016.10.015

43. Soleimani A, Pashirzad M, Avan A, Ferns GA, Khazaei M, Hassanian SM. Role of the transforming growth factor-β signaling pathway in the pathogenesis of colorectal cancer. *J Cell Biochem*. (2019) 120(6):8899–907. doi: 10.1002/jcb.28331

44. Su J, Morgani SM, David CJ, Wang Q, Er EE, Huang YH, et al. Tgf-β orchestrates fibrogenic and developmental EMTs *via* the RAS effector RREB1. *Nature*. (2020) 577 (7791):566–71. doi: 10.1038/s41586-019-1897-5

45. Chakravarthy A, Khan L, Bensler NP, Bose P, De Carvalho DD. Tgf-β-Associated extracellular matrix genes link cancer-associated fibroblasts to immune evasion and immunotherapy failure. *Nat Commun*. (2018) 9(1):4692. doi: 10.1038/s41467-018-06654-8

46. Meng XM, Nikolic-Paterson DJ, Lan HY. Tgf-β: The master regulator of fibrosis. *Nat Rev Nephrol*. (2016) 12(6):325–38. doi: 10.1038/nrneph.2016.48

47. Saito T, Nishikawa H, Wada H, Nagano Y, Sugiyama D, AtaRAShi K, et al. Two FOXP3(+)CD4(+) T cell subpopulations distinctly control the prognosis of colorectal cancers. *Nat Med*. (2016) 22(6):679–84. doi: 10.1038/nm.4086

48. Zhu H, Xu J, Wang W, Zhang B, Liu J, Liang C, et al. Intratumoral CD38(+) CD19(+)B cells associate with poor clinical outcomes and immunosuppression in patients with pancreatic ductal adenocarcinoma. *EBioMedicine*. (2024) 103:105098. doi: 10.1016/j.ebiom.2024.105098

49. Chen Y, Ouyang D, Wang Y, Pan Q, Zhao J, Chen H, et al. EBV promotes TCR-T-cell therapy resistance by inducing CD163+M2 macrophage polarization and MMP9 secretion. *J Immunother Cancer*. (2024) 12(6):e008375. doi: 10.1136/jitc-2023-008375

50. Tang P, Sheng J, Peng X, Zhang R, Xu T, Hu J, et al. Targeting NOX4 disrupts the resistance of papillary thyroid carcinoma to chemotherapeutic drugs and lenvatinib. *Cell Death Discov*. (2022) 8(1):177. doi: 10.1038/s41420-022-00994-7

51. Tang D, Kang R, Zeh HJ, Lotze MT. The multifunctional protein HMGB1: 50 years of discovery. *Nat Rev Immunol*. (2023) 23(12):824–41. doi: 10.1038/s41577-023-00894-6

52. Zhang Y, Che N, Wang S, Meng J, Zhao N, Han J, et al. Nrf2/ASPM axis regulated vasculogenic mimicry formation in hepatocellular carcinoma under hypoxia. *J Gastroenterol*. (2024) 59(10):941–57. doi: 10.1007/s00535-024-02140-9

53. Sanhueza C, Wehinger S, Castillo Bennett J, Valenzuela M, Owen GI, Quest AF. The twisted survivin connection to angiogenesis. *Mol Cancer*. (2015) 14:198. doi: 10.1186/s12943-015-0467-1

54. Zanin R, Pegoraro S, Ros G, Ciani Y, Piazza S, Bossi F, et al. HMGA1 promotes breast cancer angiogenesis supporting the stability, nuclear localization and transcriptional activity of FOXM1. *J Exp Clin Cancer Res*. (2019) 38(1):313. doi: 10.1186/s13046-019-1307-8

55. Li X, Hu W, Zhou J, Huang Y, Peng J, Yuan Y, et al. CLCA1 suppresses colorectal cancer aggressiveness *via* inhibition of the Wnt/beta-catenin signaling pathway. *Cell Commun Signal*. (2017) 15(1):38. doi: 10.1186/s12964-017-0192-z

56. Chen L, Jin XH, Luo J, Duan JL, Cai MY, Chen JW, et al. ITLN1 inhibits tumor neovascularization and myeloid derived suppressor cells accumulation in colorectal carcinoma. *Oncogene*. (2021) 40(40):5925–37. doi: 10.1038/s41388-021-01965-5

57. Au-Yeung CL, Yeung TL, Achreja A, Zhao H, Yip KP, Kwan SY, et al. ITLN1 modulates invasive potential and metabolic reprogramming of ovarian cancer cells in omental microenvironment. *Nat Commun*. (2020) 11(1):3546. doi: 10.1038/s41467-020-17383-2

58. Murphy N, CarreRAS-Torres R, Song M, Chan AT, Martin RM, Papadimitriou N, et al. Circulating levels of insulin-like growth factor 1 and insulin-like growth factor binding protein 3 associate with risk of colorectal cancer based on serologic and mendelian randomization analyses. *Gastroenterology*. (2020) 158(5):1300–12.e20. doi: 10.1053/j.gastro.2019.12.020

59. Yuan Y, Sun X, Liu M, Li S, Dong Y, Hu K, et al. Negative correlation between acetyl-CoA acyltransferase 2 and cetuximab resistance in colorectal cancer. *Acta Biochim Biophys Sin (Shanghai)*. (2023) 55(9):1467–78. doi: 10.3724/abbs.2023111

60. Priego N, de Pablos-Aragoneses A, Perea-García M, Pieri V, Hernández-Oliver C, Álvaro-Espinosa L, et al. TIMP1 mediates astrocyte-dependent local immunosuppression in brain metastasis acting on infiltrating CD8+ T cells. *Cancer Discov*. (2025) 15(1):179–201. doi: 10.1158/2159-8290.Cd-24-0134

61. Schoeps B, Eckfeld C, Prokopchuk O, Böttcher J, Häußler D, Steiger K, et al. TIMP1 triggers neutrophil extracellular trap formation in pancreatic cancer. *Cancer Res*. (2021) 81(13):3568–79. doi: 10.1158/0008-5472.Can-20-4125

62. Ma B, Ueda H, Okamoto K, Bando M, Fujimoto S, Okada Y, et al. TIMP1 promotes cell proliferation and invasion capability of right-sided colon cancers *via* the FAK/Akt signaling pathway. *Cancer Sci*. (2022) 113(12):4244–57. doi: 10.1111/cas.15567

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 06 May 2025
DOI 10.3389/fimmu.2025.1546382

OPEN ACCESS

EDITED BY
Yafeng Zhu,
Sun Yat-sen Memorial Hospital, China

REVIEWED BY
Shivam Priya,
The Ohio State University, United States
Guoshuai Cai,
University of Florida, United States

*CORRESPONDENCE
Ran Xu
xur2@sj-hospital.org

†The authors have contributed equally to this work

RECEIVED 16 December 2024
ACCEPTED 21 April 2025
PUBLISHED 06 May 2025

CITATION
Fu Y, Li S, Zhao Y, Zhang X, Mao X and Xu R (2025) Integrative single-cell and spatial transcriptomics analysis reveals MDK-NCL pathway's role in shaping the immunosuppressive environment of lung adenocarcinoma.
Front. Immunol. 16:1546382.
doi: 10.3389/fimmu.2025.1546382

Integrative single-cell and spatial transcriptomics analysis reveals MDK-NCL pathway's role in shaping the immunosuppressive environment of lung adenocarcinoma

Yu Fu†, Song Li†, Yikang Zhao, Xiran Zhang, Xiaolu Mao and Ran Xu*

Department of Thoracic Surgery, Shengjing Hospital of China Medical University, Shenyang, China

Objectives: The tumor microenvironment (TME) plays a critical role in the progression of lung adenocarcinoma (LUAD). This study aims to investigate the cellular composition of the TME in LUAD and assess the role of the MDK-NCL signaling pathway.

Methods: We employed a multi-omics strategy to investigate LUAD, combining single-cell RNA sequencing (scRNA-seq), spatial transcriptomics (ST), and bulk RNA-seq datasets. Publicly available scRNA-seq data and ST data were utilized. scRNA-seq data underwent quality control, dimensionality reduction, and clustering to characterize cell populations and identify malignant epithelial subtypes using the Seurat and inferCNV packages. Spatial transcriptomics data facilitated the identification of distinct tumor niches, while immune infiltration and ligand-receptor interactions were analyzed using MCPcounter and Niches. Experimental validation was performed via real-time PCR and western blotting on paired LUAD and adjacent normal tissue samples.

Results: scRNA-seq revealed the presence of multiple immune and stromal cell populations, with malignant epithelial cells being subdivided into six clusters. The MDK-NCL axis demonstrated high activity in malignant cells, showing strong interactions with immune and stromal components. Spatial transcriptomics revealed nine distinct tumor niches, with MDK-NCL signaling notably upregulated at the tumor-immune interface, highlighting its role in establishing an immunosuppressive microenvironment. In both the TCGA-LUAD cohort and in-house cohort, MDK and NCL were significantly upregulated at the mRNA and protein levels in tumor samples compared to normal tissues. High MDK-NCL expression in the TCGA-LUAD cohort correlated with increased TMB, MSI, and reduced immune cell infiltration. Elevated levels of immune checkpoint genes, including PD-1 and CTLA-4, in patients with high MDK-NCL expression suggested a potential resistance to immune checkpoint inhibitors. Moreover, patients with high MDK-NCL expression exhibited poorer survival outcomes, underscoring the pathway's role in tumor progression and immune evasion.

10.3389/fimmu.2025.1546382

Conclusion: Our findings reveal that LUAD cells use the MDK-NCL signaling pathway to shape the TME, suppressing immune activity and promoting malignancy in epithelial cells. This study highlights the MDK-NCL axis as a potential therapeutic target for LUAD, particularly for patients with high MDK-NCL expression.

KEYWORDS

lung adenocarcinoma, tumor microenvironment, MDK-NCL, single-cell transcriptomics, spatial transcriptomics

Introduction

Lung adenocarcinoma (LUAD) is the most common subtype of non-small cell lung cancer (NSCLC), accounting for approximately 40% of lung cancer cases (1). While targeted therapies and immunotherapies have significantly improved the survival rates of some LUAD patients, the overall prognosis remains poor. This is primarily attributed to the tumor's heterogeneity and the complexity of its tumor microenvironment (TME) (2). The TME, which consists of immune cells, stromal cells, extracellular matrix, and various signaling molecules, plays a critical role in tumor progression, immune evasion, and therapeutic resistance (3). Therefore, gaining a deeper understanding of the interactions between the tumor and its microenvironment is crucial for uncovering the mechanisms underlying cancer development and for the development of new therapeutic strategies.

Single-cell RNA sequencing (scRNA-seq) technology offers an unprecedented level of detail for deciphering the cellular heterogeneity and dynamic changes within tumors, enabling the identification of distinct cell types and their specific roles in the TME (4). In recent years, scRNA-seq has been widely employed in LUAD research, leading to the discovery of multiple heterogeneous cell subpopulations, including tumor cells, immune cells, and stromal cells, further elucidating mechanisms of immune evasion and the interactions between tumors and their microenvironment (5, 6). Moreover, the application of spatial transcriptomics has allowed for a more comprehensive understanding of the spatial distribution of these cell populations within tumors and their interactions, providing a more complete view of the TME (7, 8).

Among the many signaling pathways that influence the TME, the Midkine (MDK)-Nucleolin (NCL) axis has garnered significant attention in recent years. MDK, a pro-tumor growth factor, is highly expressed in various types of cancer and has been shown to promote cell proliferation, migration, and survival (9–11).In LUAD, MDK expression correlates with poor prognosis, yet its potential role in modulating immune suppression remains unclear (13). Unlike TGF-β-mediated immunosuppression, which primarily acts via Treg activation and myeloid suppression, MDK-NCL signaling may establish a distinct immunosuppressive niche by interacting with tumor-associated macrophages (TAMs) and fibroblasts. Research has demonstrated that the MDK-NCL axis facilitates the formation of an immunosuppressive microenvironment, thereby promoting immune evasion by tumor cells and contributing to tumor progression (12). Given these unique properties, investigating the MDK-NCL axis may reveal novel mechanisms of immune evasion in LUAD.

In this study, we utilized scRNA-seq and spatial transcriptomics to deeply analyze the TME in LUAD and further classify malignant cell populations. We identified that MDK-NCL signaling plays a critical role in the interactions between maliganant cells and immune cells, potentially driving immune evasion and reshaping the microenvironment. Through spatial transcriptomic data, we further revealed the differential spatial distribution of MDK-NCL signaling across various tumor niches. Moreover, by integrating bulk RNA-seq data from the TCGA-LUAD cohort, we investigated the relationship between MDK-NCL expression, immune cell infiltration, and clinical outcomes. This study provides new insights into the role of the MDK-NCL axis in LUAD, particularly regarding its involvement in microenvironmental remodeling and immune evasion. Our findings offer a theoretical foundation for considering MDK-NCL as a potential therapeutic target, with significant implications for enhancing the efficacy of immunotherapy in clinical settings.

Results

ScRNA-seq and cell type identification of LUAD

After correcting for batch effects, performing dimensionality reduction, and clustering, we analyzed several key aspects of the single-cell data (GSE131907). We visualized sample origins (Figure 1A), transcript counts (Figure 1B), cell clusters (Figure 1C), and cell type annotations (Figure 1D). Marker gene expression patterns, used to identify different cell types, are depicted in Figure 1E. Specifically, T cells were identified by TRAC, monocyte-macrophages by LYZ, NK cells by NKG7, epithelial cells by EPCAM, B cells by CD79A, fibroblasts by COL1A1, mast cells by MS4A2, endothelial cells by PECAM1, conventional dendritic cells (cDCs) by CD1C, and plasmacytoid dendritic cells

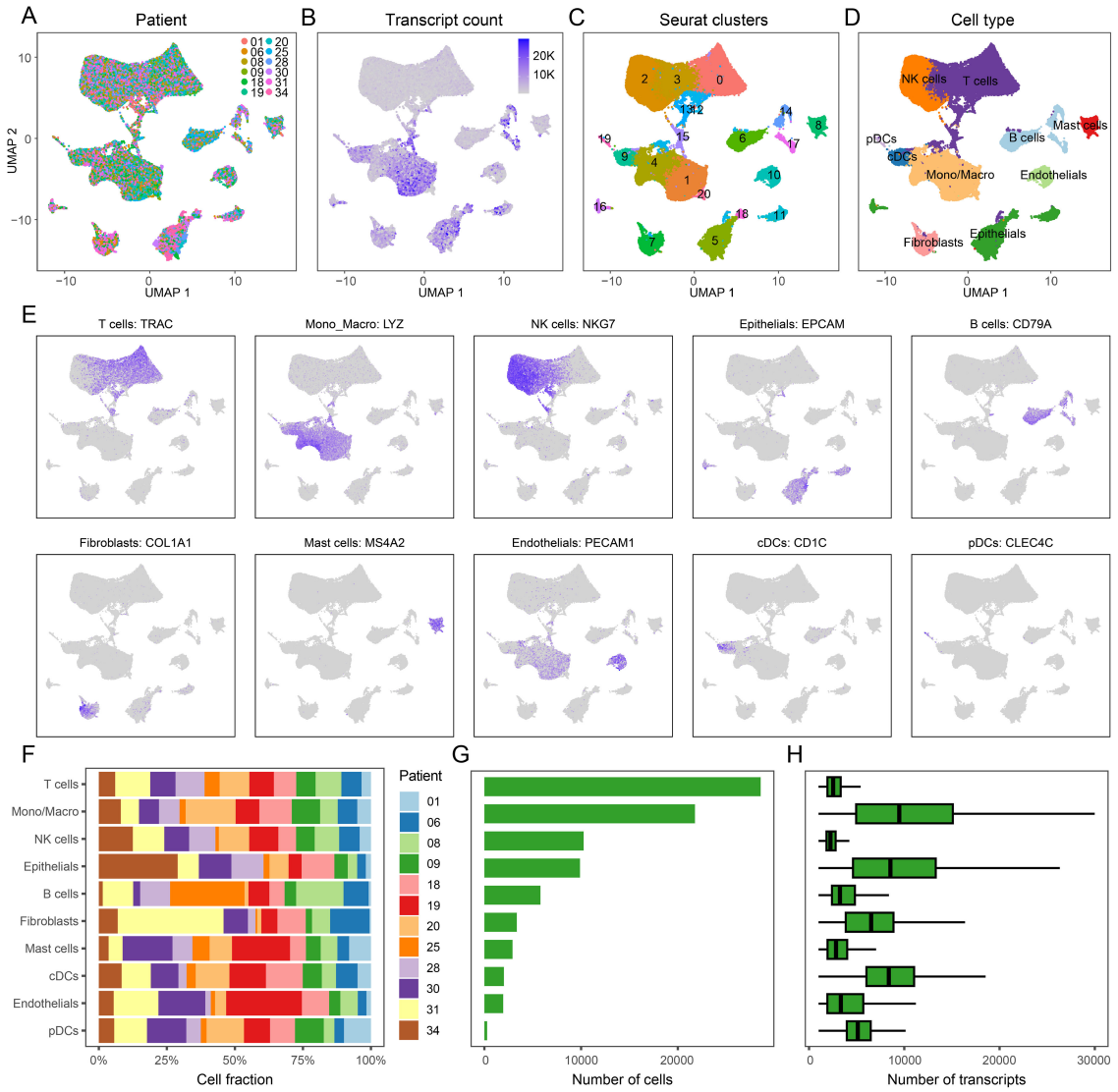


FIGURE 1
Annotation Results of scRNA-seq for LUAD. **(A)** Sample origin of the single-cell data, 12 samples were identified without batch effect. **(B)** Transcript counts in the single-cell dataset. **(C)** Clustering results of the single-cell data, totally 21 clusters were presented. **(D)** Cell type annotation based on marker gene expression, including T cells, monocyte-macrophages, NK cells, epithelial cells, B cells, fibroblasts, mast cells, endothelial cells, conventional dendritic cells (cDCs) and plasmacytoid dendritic cells (pDCs). **(E)** Expression profiles of representative markers for ten distinct cell types. **(F)** Proportion of each cell type across samples. **(G)** Total number of cells for each identified cell type. **(H)** Transcript counts per cell type, reflecting transcriptional activity at the single-cell level.

(pDCs) by CLEC4C. The proportions of each cell type across samples are shown in Figure 1F, with the absolute numbers in Figure 1G, and transcript counts for each cell type detailed in Figure 1H. The same cell annotation procedure was also performed on the single-cell validation data GSE153935 (Supplementary Figure S1A-E).

Malignant cell subpopulations in LUAD and their characteristics

To infer malignancy within the epithelial cell populations, we applied inferCNV analysis in GSE131907 and GSE153935 (Figure 2A & Supplementary Figure S1F). In the GSE131907 dataset, malignant epithelial cells were subsequently extracted for further dimensionality reduction and clustering, revealing six distinct malignant cell subpopulations (clusters 0-5, Figure 2B). Sample distribution across these malignant clusters is illustrated in Figure 2C, highlighting both intra- and inter-sample heterogeneity within LUAD tumors.

Using the FindAllMarkers function, we identified cluster-specific markers for each malignant subpopulation (Figure 2D). Functional enrichment analysis via ssGSEA using hallmark gene sets revealed distinct biological pathways across clusters (Figure 2E). For example, clusters 0 and 1 were enriched in pathways related to metabolism and mitosis. Univariate Cox

10.3389/fimmu.2025.1546382

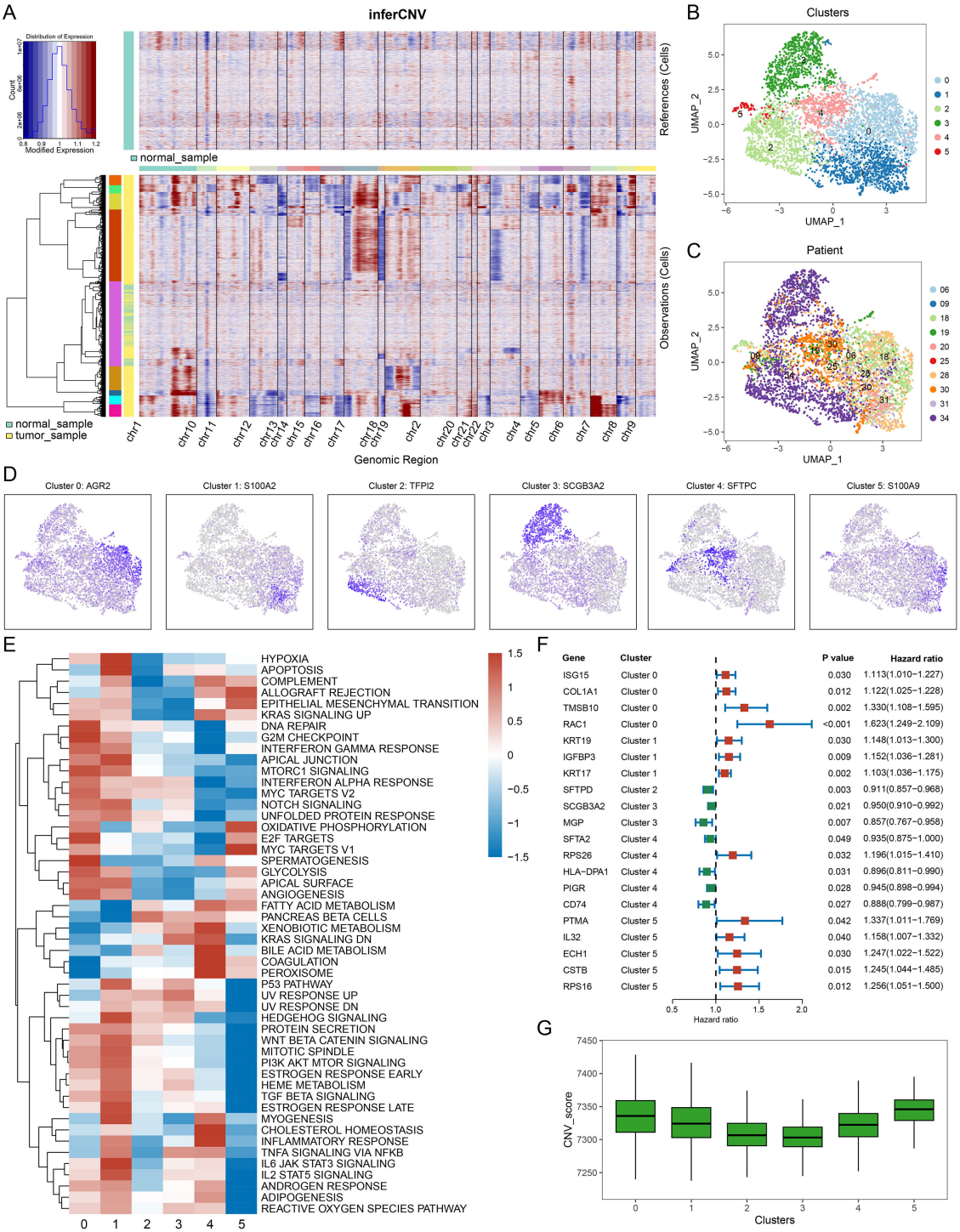


FIGURE 2
Identification of malignant cell subtypes. **(A)** inferCNV heatmap displaying copy number variations (CNVs) across cells, with normal samples in the upper panel and tumor samples in the lower panel. Red and blue indicate CNV gains and losses, respectively. **(B)** UMAP plot showing clustering of malignant cells, revealing distinct subpopulations. **(C)** UMAP plot indicating the patient origin of malignant cells, highlighting inter-sample heterogeneity. **(D)** Expression of representative marker genes for each malignant cluster: AGR2 (Cluster 0), S100A2 (Cluster 1), TPP2 (Cluster 2), SCGB3A2 (Cluster 3), SFTPC (Cluster 4), and S100A9 (Cluster 5). **(E)** Heatmap of hallmark pathway activities across clusters, with red indicating upregulation and blue indicating downregulation of pathways, such as hypoxia response and interferon signaling, cluster0 and cluster1 have more upregulated pathways. **(F)** Univariate Cox analysis of key marker genes, with hazard ratios, confidence intervals, and P-values showing their prognostic significance. Red indicates higher risk associations, while green indicates lower risk. **(G)** CNV scores of different malignant cell subtypes.

regression survival analysis, based on the top five markers per cluster, showed that markers from clusters 0, 1, and 5 were associated with higher risk, while clusters 2, 3, and 4 were more protective (Figure 2F). Meanwhile, clusters 0, 1, and 5 have higher CNV scores, suggesting greater genomic instability (Figure 2G). This suggests that clusters 0, 1, and 5 exhibit more aggressive, malignant phenotypes.

The result of single-cell communication analysis

Cell-to-cell communication was analyzed for both GSE131907 and GSE153935 using the CellChat package, which identified receptor-ligand signaling pathways received (Figure 3A & Supplementary Figure S2A) and emitted (Figure 3B & Supplementary Figure S2B) by different cell types. Notably, the MIF, MK, and CXCL signaling pathways were highly active. Figure 3C and Supplementary Figure S2C illustrates the overall communication strength between cell types, while Figure 3D and Supplementary Figure S2D shows the intensity of signals emitted and received by each cell type. Malignant cells exhibited the highest signal emission strength, underscoring their dominant role in influencing the TME.

Among receptor-ligand pairs, four of the top ten interactions belonged to the MK pathway, with the MDK-NCL interaction being the most significant (Supplementary Figure S3 & Supplementary Figure S4). The strength of MK pathway communication across different cell types is presented in Figure 3E and Supplementary Figure S2E, with malignant cells being the primary senders and receivers of these signals. We further analyzed MDK-NCL interactions between malignant and immune/stromal cells, finding significant interaction strengths (Figure 3F & Supplementary Figure S2F). Malignant cells exhibited extensive interactions with all immune and stromal cell types through the MDK-NCL axis. Expression levels of genes involved in the MK pathway are shown in Figure 3G, with higher expression of MDK in malignant cells and broad expression of NCL across all cell types. These findings highlight the critical role of the MDK-NCL interaction in shaping the TME.

Spatial transcriptomic niche communication analysis

Following dimensionality reduction and clustering, we identified nine distinct spatial niches (niche 0-8, Figure 4A). Based on the expression of key marker genes-MUC1 (tumor region), LYZ (immune region), COL14A1 (stromal region), and SFTPC (normal region)—we classified the niches into tumor, immune-stromal, and normal regions across all spatial transcriptomic samples (Figures 4B, C, Supplementary Figure S5). To validate our classification, we performed MCPcounter immune infiltration analysis (Figure 4D), identifying six distinct cell types-endothelial cells, fibroblasts, monocytes, T cells, B cells, and neutrophils-within the niches. The distribution of these cell types across the spatial niches is depicted in Figure 4E, showing a clear division into tumor, immune-stromal, and normal epithelial regions.

We then conducted spatial communication analysis to assess the spatial distribution of MDK-NCL receptor-ligand signaling across the niches (Figure 4F, Supplementary Figure S5). This included examining MDK-NCL ligand-receptor binding, NCL receptor levels, and MDK ligand expression. These spatial analyses revealed that the MDK-NCL axis plays a significant role in mediating communication between malignant cells and the surrounding immune and stromal cells, further contributing to the spatial organization of the TME.

Pseudotime analysis of single cells

To explore the developmental trajectory of malignant epithelial cells and the changes in the MK signaling pathway during tumor progression, we conducted a pseudotime analysis using spatial transcriptomics data from LUAD. Figures 5A-C illustrate the differentiation states, cell subtypes, and pseudotime scores obtained from the analysis. In Figure 5D, pseudotime scores are visualized using a UMAP dimensionality reduction plot, while a box plot (Figure 5E) compares the pseudotime scores of different malignant cell clusters, revealing that clusters 0, 1, and 5 have higher pseudotime scores. Additionally, the differentiation states of these clusters are shown in Figure 5F, and the proportion of cells in each state is presented in Figure 5G, with clusters 0, 1, and 5 primarily occupying differentiation state 6, which is associated with a more advanced pseudotime score. These findings indicate that clusters 0, 1, and 5, which are negatively correlated with prognosis, not only have higher pseudotime scores but also reside in more differentiated states, suggesting a higher level of tumor progression and malignancy. Finally, we analyzed the expression trends of MK pathway genes along the pseudotime trajectory (Figure 5H), which showed a gradual upregulation of MDK and NCL expression with increasing pseudotime scores.

The impact of MDK-NCL on the LUAD immune microenvironment

Using single-cell and spatial transcriptomic analyses, MDK-NCL communication between tumor cells and other cells was identified as a critical mechanism in shaping the TME. Analysis of bulk transcriptomic data from the TCGA-LUAD cohort revealed that MDK and NCL expression levels were significantly higher in tumor samples compared to control samples (Figure 6A). Similarly, GSVA enrichment scores for the MDK-NCL pathway were also markedly elevated in tumor samples (Figure 6B). Three validation public cohorts were corresponding to the same results (Supplementary Figure S6A-C). Consistent with these findings, in our cohort, the relative mRNA expression levels of MDK and NCL were significantly higher in tumor tissues than in adjacent normal

10.3389/fimmu.2025.1546382

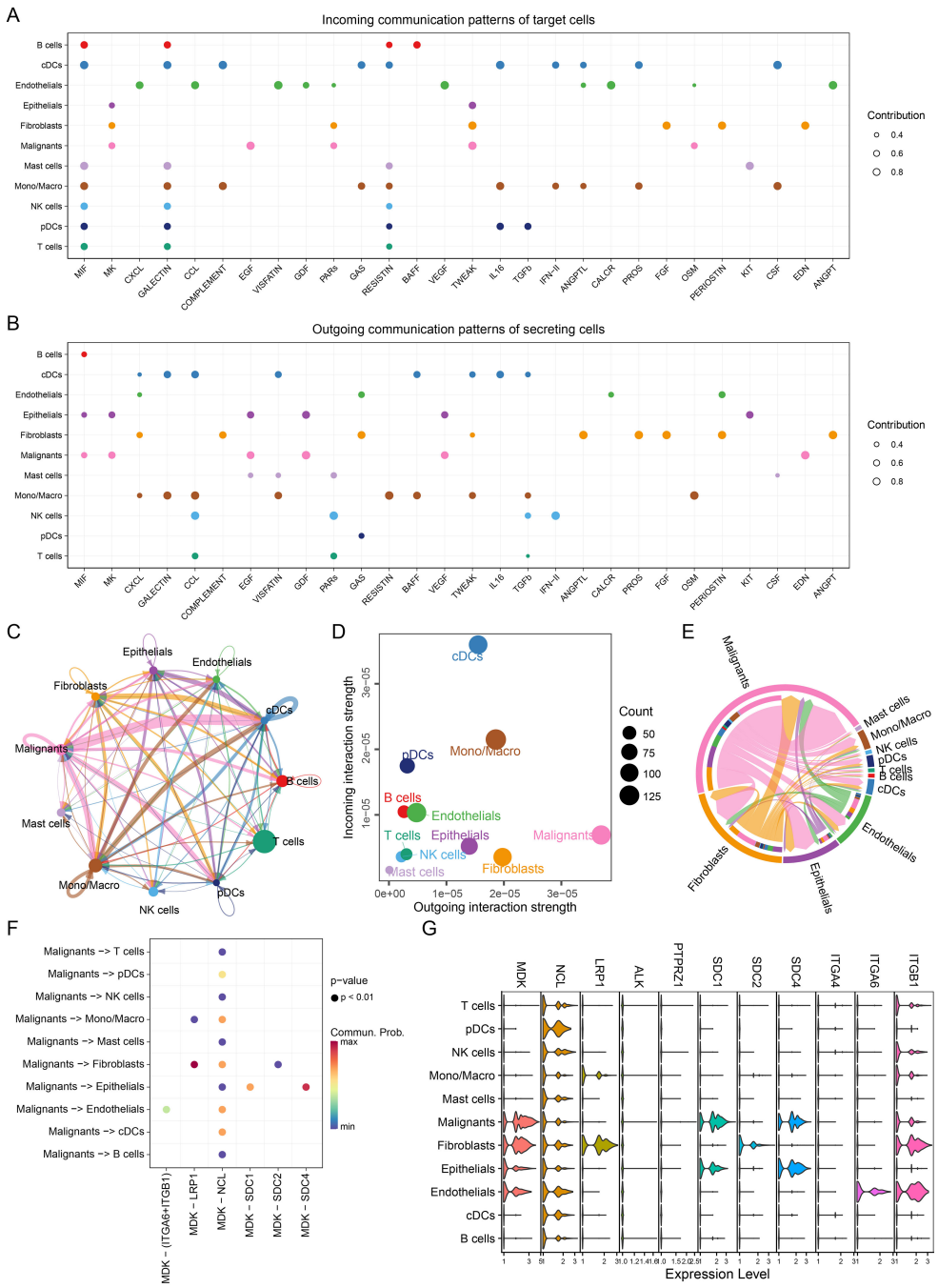


FIGURE 3
Single-cell communication networks. **(A)** Incoming communication patterns of target cells, showing pathways to which each cell type responds. **(B)** Outgoing communication patterns of secreting cells, illustrating the pathways through which cells send signals, MIF, MK and CXCL pathway exhibit high activity. **(C)** Network diagram showing the strength of intercellular communication, with connections between various cell types. **(D)** Scatter plot comparing outgoing and incoming communication strengths across cell populations, with bubble size indicating the number of interactions, malignant cells have higher strength of intercellular communication. **(E)** Chord diagram depicting communication via the MK pathway between different cell types. **(F)** Ligand-receptor interaction probabilities within the MK pathway between malignant and other cell types. Dot size represents significance (P-value), and color represents communication probability highlighting the MDK-NCL signaling pathway. **(G)** Violin plots of MK pathway gene expression levels across cell types, showing gene activity variations, MDK has advancer expression level in malignant cells.

tissues (Figure 6C). Western blot analysis further confirmed that protein expression levels of MDK and NCL were significantly upregulated in tumor samples compared to controls (Figures 6D-F). To explore the impact of MDK and NCL on the immune microenvironment, we performed ESTIMATE analysis using the TCGA-LUAD dataset. The results demonstrated a negative correlation between MDK and NCL expression levels and immune-related scores, including the ImmuneScore,

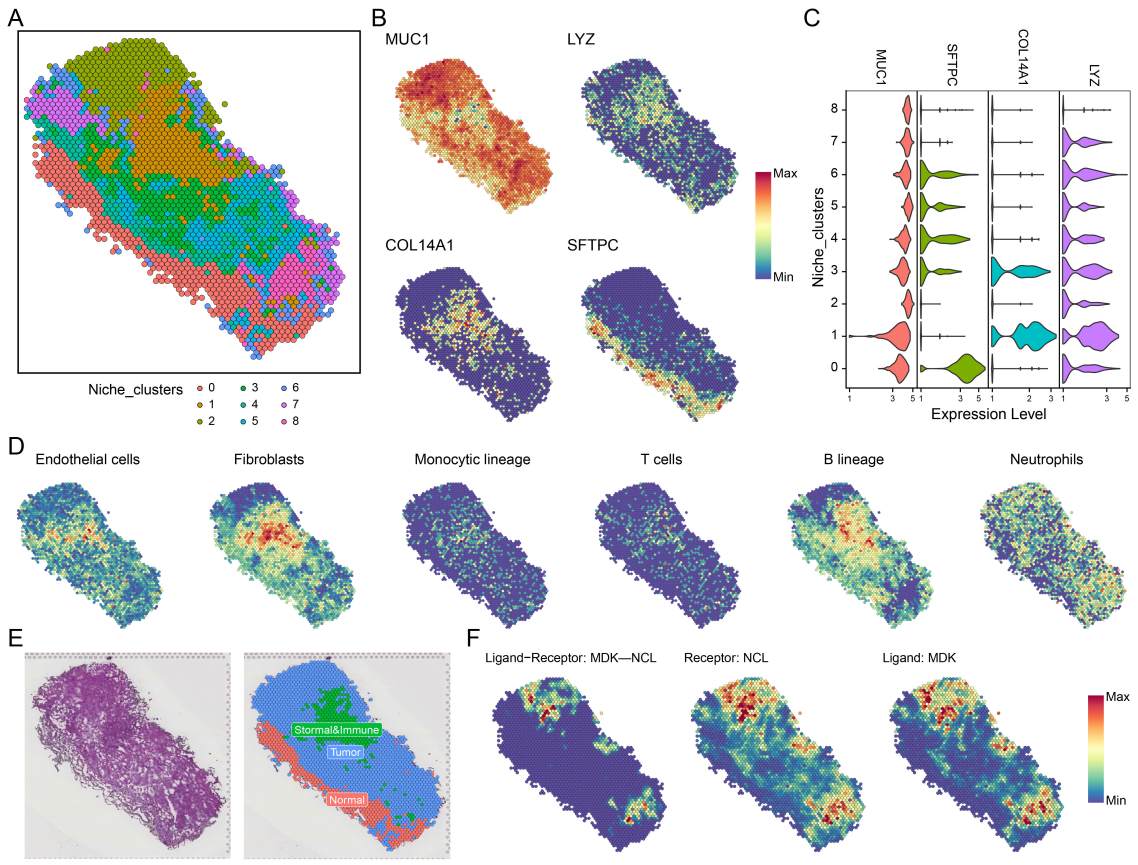


FIGURE 4
Spatial transcriptomics and MDK-NCL signal communication. **(A)** Niche clustering in spatial transcriptomics samples, identifying distinct ecological zones. **(B)** Spatial expression of representative markers in key regions: MUC1 (tumor region), LYZ (immune region), COL14A1 (stromal region), and SFTPC (normal region). **(C)** Violin plots displaying the expression of MUC1, LYZ, COL14A1, and SFTPC across different niches. **(D)** MCPcounter analysis showing the infiltration of six cell types (e.g., endothelial cells, fibroblasts, immune lineages) across spatial regions. **(E)** Spatial niche classification, distinguishing tumor, immune-stromal, and normal regions. **(F)** MDK-NCL ligand-receptor interaction analysis, spatially mapping MDK ligands, NCL receptors, and their binding regions.

StromalScore, and ESTIMATEScore. Conversely, a positive correlation was observed between MDK and NCL expression and TumorPurity (Figures 6G, H). These findings suggest that MDK and NCL are associated with the development of an immunosuppressive TME. Further analysis divided tumor samples into high and low MDK-NCL expression groups based on the median enrichment score. Immune infiltration analysis revealed that immune cell scores for various cell types were significantly lower in the high MDK-NCL expression group compared to the low-expression group. Patients with high MDK-NCL expression groups exhibit increased infiltration of regulatory T cells (Tregs), myeloid-derived suppressor cells (MDSCs), and M2-like macrophages, which are known to promote immune evasion and tumor progression. Additionally, the MDK-NCL pathway suppresses cytotoxic immunity by reducing activated and effector memory CD8+ T cells while promoting an immunosuppressive microenvironment through increased Tregs and altering helper T cell differentiation, facilitating tumor immune evasion. This supports the conclusion that MDK-NCL activity suppresses immune cell infiltration and activity, contributing to immune evasion in LUAD.

The association of MDK-NCL with immunotherapy

We observed that the high MDK-NCL expression group exhibited higher tumor mutation burden (TMB) (Figure 7A) and microsatellite instability (MSI) scores (Figure 7B), indicating increased genomic instability. TIDE analysis revealed that the high MDK-NCL group had lower Dysfunction scores (Figure 7C) and higher Exclusion scores (Figure 7D), suggesting that although T cell dysfunction was lower, there was a higher degree of T cell exhaustion. This supports that MDK-NCL may promote an immune-resistant TME through T cell exclusion rather than direct T cell exhaustion, a mechanism distinct from PD-1/PD-L1, which primarily induces T cell dysfunction at the tumor-immune interface. Additionally, we analyzed the expression patterns of immunogenic cell death (ICD)-related genes (Figure 7E), finding that the high MDK-NCL group had higher expression of several ICD genes, while toll-like receptors TLR3 and TLR4 showed lower expression. These findings suggest that MDK-NCL may contribute to immune evasion by promoting T cell exclusion and downregulating innate immune sensing, similar to TGF-β.

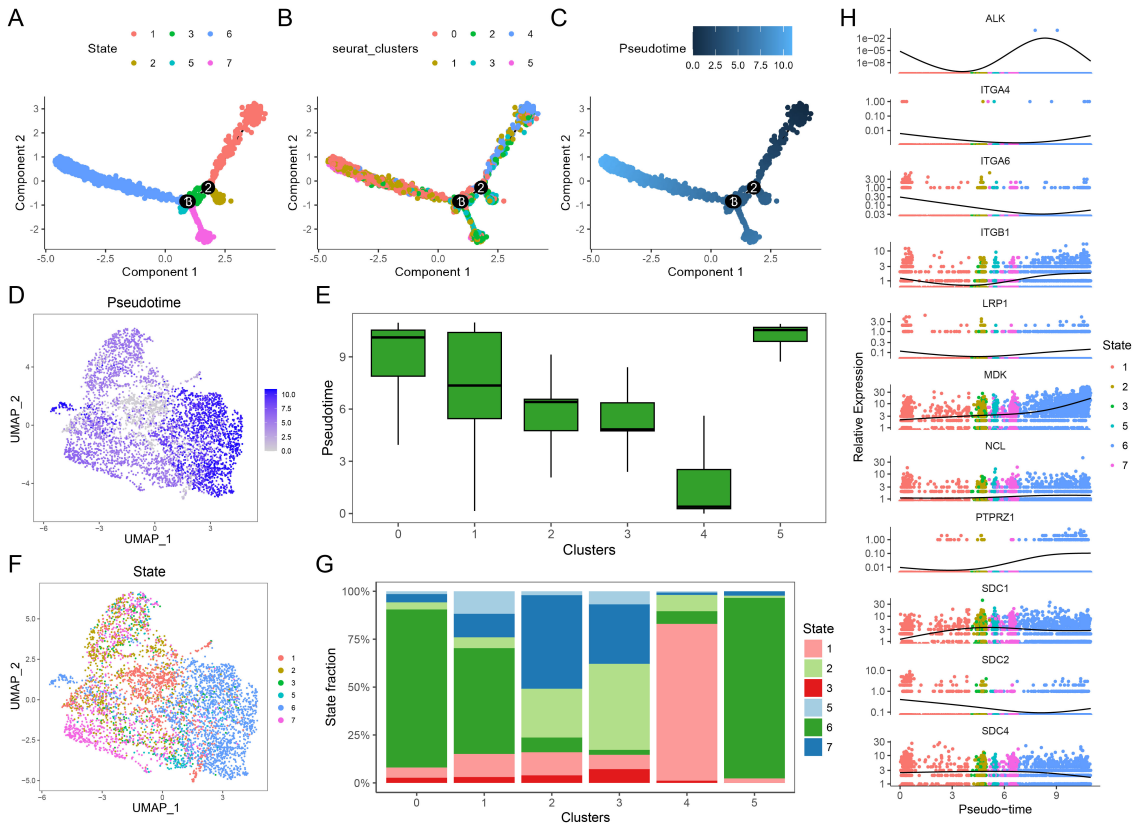


FIGURE 5
Single-cell pseudotime analysis. **(A)** Pseudotime trajectory analysis showing the 6 differentiation states of cells. **(B)** Subtype classification of malignant cells along the pseudotime trajectory. **(C)** Pseudotime scores mapped along the differentiation trajectory. **(D)** UMAP plot visualizing pseudotime scores across individual cells. **(E)** Box plots comparing pseudotime scores across different malignant cell clusters, cluster 0, 1, and 5 had higher pseudotime scores. **(F)** UMAP plot of differentiation states, with colors representing distinct states. **(G)** Stacked bar plots showing the proportion of differentiation states within each malignant cell cluster, cluster 0, 1, and 5 have larger proportion of state 6. **(H)** Expression dynamics of MK pathway genes (e.g., MDK, NCL, ITG genes) along the pseudotime trajectory, highlighting gene expression changes during differentiation, MDK and NCL express more in the later time.

Predictions from the TCIA database indicated that the high MDK-NCL group had fewer patients with dual-negative CTLA4 and PD1 status, as well as fewer patients with PD1 single positivity but more patients with CTLA4 single positivity (Figure 7F). This suggests that MDK-NCL may enhance CTLA-4 mediated immune suppression, potentially influencing the response to anti-CTLA-4 therapy. Similarly, most of these factors got the same trends in validation cohorts (Supplementary Figure S6D-F). Finally, we compared the expression profiles of immune checkpoint-related genes between the two groups (Figure 7G), revealing that the high MDK-NCL group had elevated expression of checkpoint genes such as LAG3 and PDCD1, suggesting that these patients may respond more favorably to immune checkpoint inhibitors. Overall, our findings suggest that high MDK-NCL expression may predict poor ICI response by fostering an immune-excluded tumor microenvironment. Despite high TMB/MSI, MDK-NCL-high tumors show low CD8+ T-cell infiltration and increased Tregs/MDSCs, potentially negating the benefits of increased neoantigens. This highlights MDK-NCL as a negative predictor of ICI response and a potential target to enhance ICI efficacy.

Discussion

In this study, we systematically investigated the role of the MDK-NCL signaling axis in the TME of LUAD through the integration of scRNA-seq and spatial transcriptomics. Our findings shed light on the mechanisms by which the MDK-NCL pathway contributes to immune suppression and tumor immune evasion. offering novel insights into the potential of targeting this axis as a therapeutic strategy. This work deepens our understanding of TME remodeling in LUAD.

MDK, a pro-tumorigenic growth factor, is highly expressed in various cancers (10, 14). It exerts its oncogenic effects primarily by binding to its receptor, Nucleolin (NCL), through which it modulates various signaling pathways critical for the regulation of tumor progression and the maintenance of the TME (15, 16). Through scRNA-seq and spatial transcriptomics, we elucidated the role of MDK-NCL signaling in LUAD at the cellular level. Our results indicate that MDK-NCL plays a pivotal role in the interaction between malignant, immune, and stromal cells, particularly by fostering an immunosuppressive environment that

10.3389/fimmu.2025.1546382

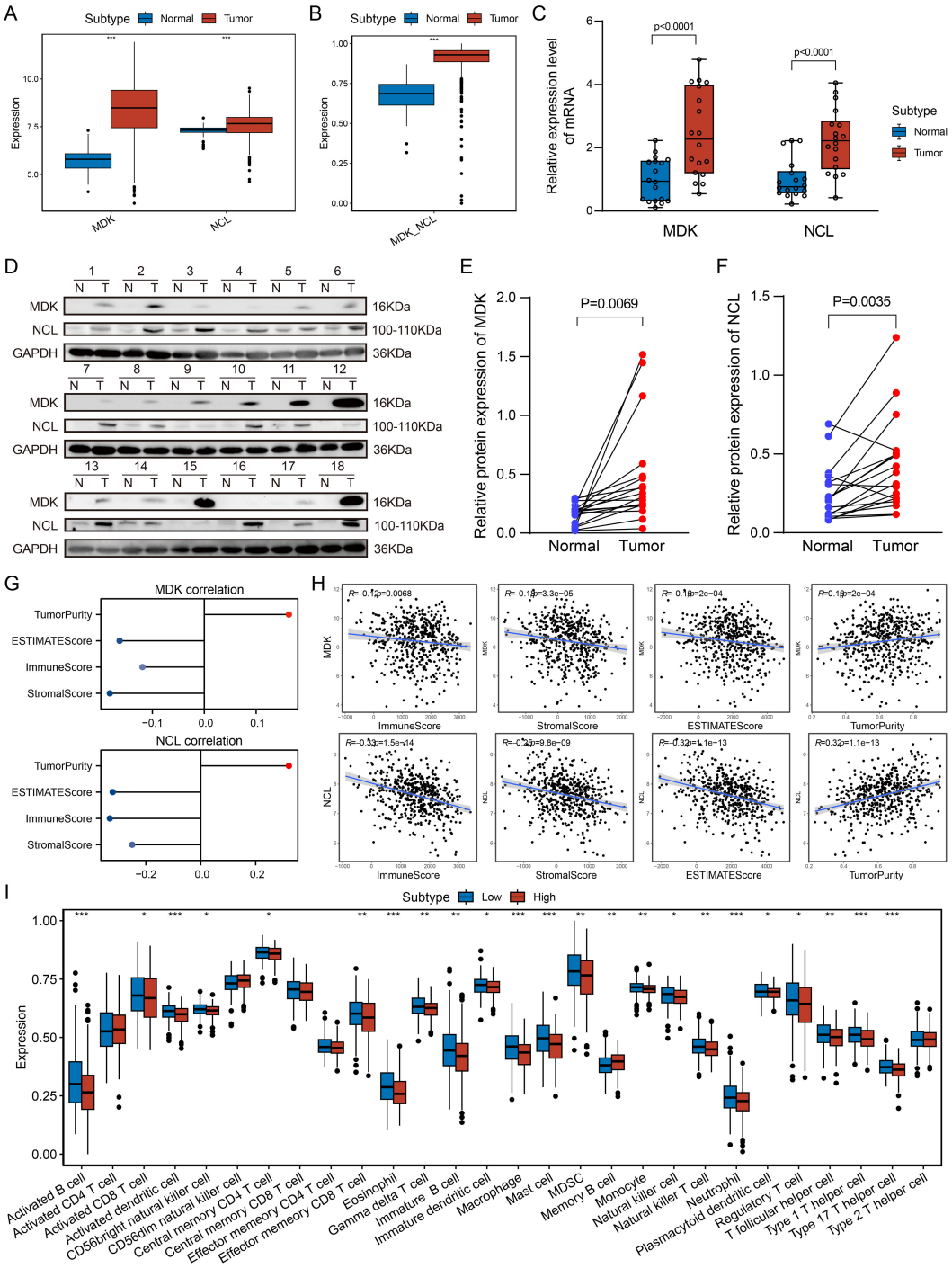


FIGURE 6
Association of MDK-NCL with the immune microenvironment. **(A)** Boxplot shows the expression levels of MDK and NCL genes in tumor and control groups, it exhibit higher activity in tumor group. **(B)** MDK-NCL enrichment scores in tumor and control groups. **(C)** Relative mRNA expression levels of MDK and NCL in tumor and control groups from in-house data. **(D)** Relative protein expression levels of MDK and NCL in tumor and control groups from in-house data. **(E)** Comparison of MDK protein expression levels between tumor and control groups. **(F)** Comparison of NCL protein expression levels between tumor and control groups. **(G)** Correlation of MDK and NCL expression with ImmuneScore, StromalScore, ESTIMATEScore, and TumorPurity. **(H)** Scatter plots depicting the relationship between MDK and NCL expression and immune-related scores (ImmuneScore, StromalScore, ESTIMATEScore) as well as TumorPurity. **(I)** Comparison of immune cell infiltration scores across high and low MDK-NCL expression groups for 28 immune cell types. *P < 0.05, **P < 0.01, ***P < 0.001.

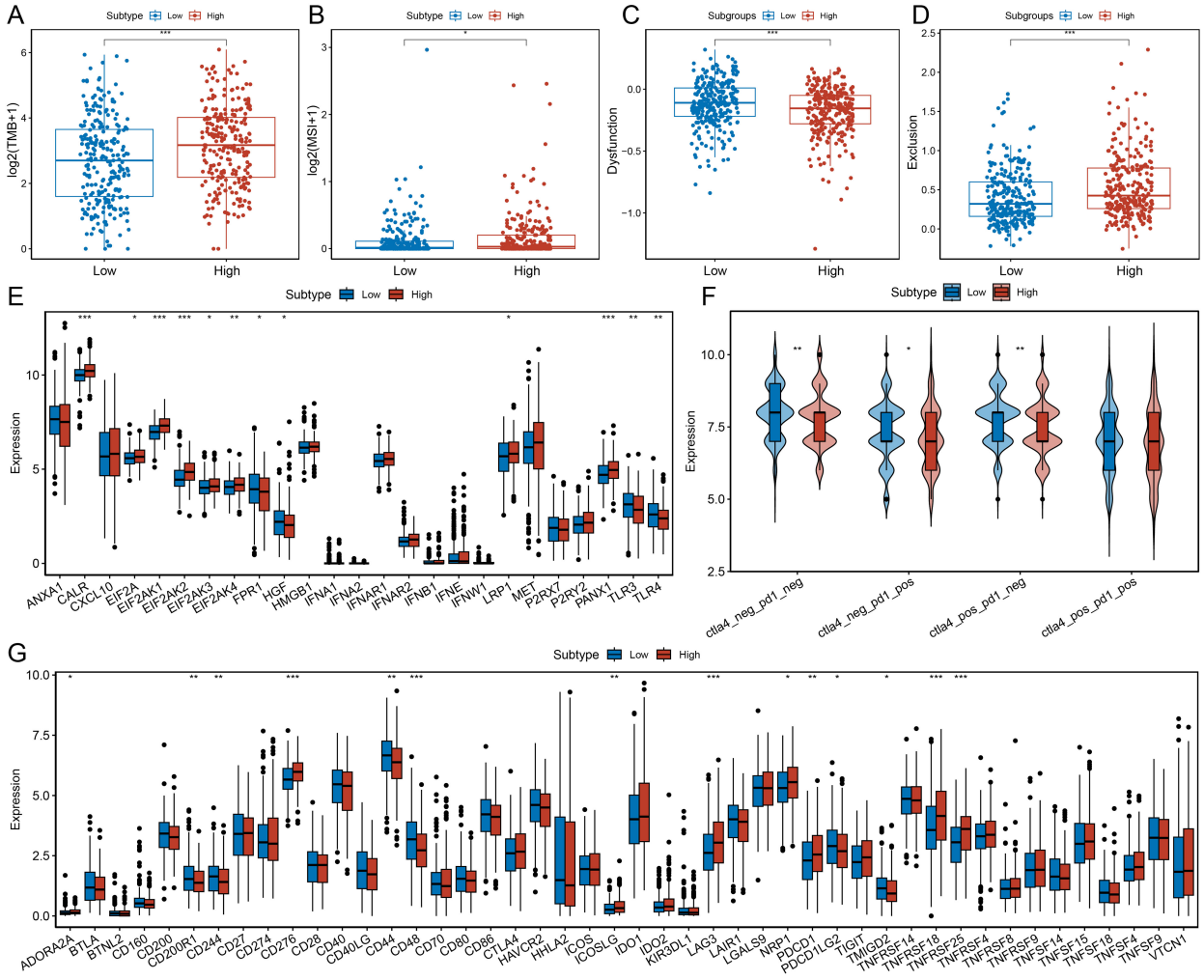


FIGURE 7
Association of MDK-NCL with immunotherapy response. **(A)** Comparison of tumor mutation burden (TMB) between high and low MDK-NCL expression groups. **(B)** Comparison of microsatellite instability (MSI) between high and low MDK-NCL groups. **(C)** Comparison of dysfunction scores between high and low MDK-NCL groups. **(D)** Comparison of exclusion scores between high and low MDK-NCL groups. **(E)** Expression of immunogenic cell death (ICD)-related genes in high and low MDK-NCL groups. **(F)** Expression levels of CTLA4 and PD1 in high and low MDK-NCL groups. **(G)** Comparison of immune checkpoint gene expression between high and low MDK-NCL expression groups. *P < 0.05, **P < 0.01, ***P < 0.001.

supports tumor immune evasion. This mechanism is further supported by its observed spatial heterogeneity across distinct tumor regions. Spatial transcriptomics revealed that MDK-NCL signaling activity was markedly elevated at the tumor-immune interface, a region characterized by high cellular density and active immune-tumor interactions. This enrichment suggests that MDK-NCL may serve as a defensive mechanism for tumor cells at immune hotspots, preventing effective immune cell infiltration and cytotoxic activity. The differential expression across tumor niches underscores the biological importance of spatial heterogeneity in shaping TME architecture and influencing immune evasion strategies. For example, in low-immune regions, MDK-NCL may facilitate stromal remodeling, whereas in high-immune regions, it likely plays a more direct role in immune cell suppression. These observations align with previous studies emphasizing the role of spatial heterogeneity in defining TME functions (17). Understanding this spatial regulation not only highlights the complexity of MDK-NCL signaling but also opens avenues for spatially targeted therapeutic strategies, such as local delivery of inhibitors to high-activity regions within the TME.

Another important finding of this study is the potential role of the MDK-NCL axis in immunotherapy. By analyzing TMB and MSI data from the TCGA-LUAD dataset, we found that patients with high MDK-NCL expression tend to have higher TMB and MSI levels, indicating a potential association between MDK-NCL signaling and genomic instability, which may impact the response to immune checkpoint inhibitors (ICIs). While high TMB/MSI tumors are generally considered more immunogenic and respond better to ICIs, tumors with high MDK-NCL expression exhibit immune exclusion, despite their increased TMB/MSI levels. Specifically, we hypothesize that MDK-NCL blockade could enhance the effectiveness of anti-PD-1/PD-L1 and anti-CTLA-4

therapies by reversing T-cell exclusion and promoting immune cell infiltration. These findings highlight the potential dual role of MDK-NCL in influencing ICI responses. On one hand, increased TMB and MSI levels suggest heightened neoantigen production, which is typically associated with improved ICI efficacy. On the other hand, the observed high T-cell exclusion scores in patients with elevated MDK-NCL expression reflect an immunosuppressive phenotype, indicating that this axis might hinder the infiltration or activation of T cells despite a high mutational load. Additionally, elevated expression of immune checkpoint genes (e.g., PDCD1 and CTLA-4) in the high MDK-NCL group suggests that this pathway might promote immune evasion by enhancing the tumor's dependence on checkpoint mechanisms. Therefore, targeting MDK-NCL signaling could potentially synergize with ICIs, alleviating immune suppression and restoring effective T-cell-mediated immunity. Moreover, the ability to stratify LUAD patients based on MDK-NCL expression into groups with distinct immune profiles and ICI responses could provide valuable insights for personalized therapy. Negative correlation between MDK-NCL expression and IFN-γ response genes, suggests that MDK-NCL may suppress IFN-γ-mediated antitumor immunity. Additionally, given the known role of TGF-β in promoting immune exclusion, MDK-NCL may interact with this pathway to reinforce immune suppression. For instance, patients with high MDK-NCL expression may benefit from combination therapies targeting both MDK-NCL signaling and immune checkpoints, improving response rates and reducing resistance to treatment. Future studies should focus on preclinical models to validate this hypothesis and assess the feasibility of such combination strategies in LUAD.

Preclinical studies have demonstrated the efficacy of MDK and NCL inhibitors in cancers such as glioblastoma (18) and pancreatic cancer (19), where MDK signaling is implicated in tumor progression and immune suppression. However, their clinical efficacy in LUAD remains unexplored. Our study highlights the critical role of the MDK-NCL axis in LUAD immune evasion and tumor progression, providing a theoretical basis for targeting this pathway as a novel therapeutic strategy. Our findings indicated that high MDK-NCL expression correlates with reduced infiltration of antigen-presenting cells (APCs), such as dendritic cells and MHC class I/II expression levels. This suggests that MDK-NCL signaling may downregulate antigen presentation, reducing tumor immunogenicity. However, further functional studies are required to confirm this hypothesis. The high expression of MDK-NCL signaling in LUAD patients is associated with an unfavorable immune microenvironment and increased immune exclusion, suggesting that targeting this axis may enhance the efficacy of existing immunotherapies. Unlike PD-1/PD-L1, which primarily induces T-cell exhaustion, our data suggest that MDK-NCL drives immune suppression through T-cell exclusion and stromal remodeling. Additionally, MDK-NCL-high tumors show increased infiltration of regulatory T cells (Tregs) and myeloid-derived suppressor cells (MDSCs), highlighting a distinct mechanism of immune evasion. Developing MDK-NCL pathway inhibitors holds promise as a novel treatment option for refractory LUAD, particularly for patients unresponsive to conventional immunotherapy.

Despite these significant findings, our study has several limitations. Although we integrated multi-omics data to explore the role of MDK-NCL signaling in LUAD and conducted preliminary experimental validation, further mechanistic studies are required, particularly in larger clinical cohorts. Additionally, our analysis relies heavily on publicly available datasets, which lack extensive clinical sample support. Prospective clinical studies are needed to validate our conclusions. Moreover, the development of MDK-NCL pathway inhibitors should be prioritized and evaluated in clinical trials. Finally, integrating additional multi-omics approaches, such as proteomics and metabolomics, could provide a more comprehensive understanding of the complex regulatory roles of MDK-NCL signaling in tumor progression and immune evasion.

Methods

Data acquisition

We downloaded spatial transcriptomics data from a LUAD patient sample using the 10x Visium technology from the BioStudies database (20) (https://www.ebi.ac.uk/biostudies/) (Accession number: E-MTAB-13530, This dataset includes a total of 40 lung tissue or NSCLC samples. For our analysis, we selected 8 tumor samples from this cohort.

From the Gene Expression Omnibus (GEO) database (21) (https://www.ncbi.nlm.nih.gov/geo/), we obtained the GSE131907 dataset (22) and GSE153935 dataset (23), which includes scRNA-seq data generated using the 10x Genomics platform and Drop-seq platform. GSE131907 dataset comprises 22 single-cell samples, including 11 primary tumor samples and 11 normal lung tissue samples, and was used for experimental analysis in this study. GSE153935 dataset comprises 18 single-cell samples, including 12 primary tumor samples and 6 normal lung tissue samples, and was used for validation analysis in this study.

We also retrieved bulk gene expression data (TPM) and clinical information such as patient gender, age, stage, grade, and survival outcomes from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/). Additionally, tumor mutation burden (TMB) and microsatellite instability (MSI) data for LUAD patients were obtained from cBioPortal (24) (https://www.cbioportal.org/). Meanwhile, GSE11969 (25) (including 94 LUAD and 5 normal samples), GSE43458 (26) (including 80 LUAD and 30 normal samples), GSE116959 (27) (including 57 LUAD and 11 normal samples) were obtained from GEO database as well for validation.

Single-cell RNA-seq data processing for LUAD

We utilized the Seurat package (version 4.3.0) (4) to process and analyze the scRNA-seq data. Quality control was performed by filtering out cells with fewer than 200 or more than 8,000 genes,

those with fewer than 50,000 transcripts, cells with greater than 20% ribosomal gene content, and cells with more than 3% hemoglobin gene expression.

Next, SCTransform (28) was applied for normalization and variance stabilization, followed by the Harmony algorithm (29) correct batch effects. Principal component analysis (PCA) was performed, and the first 30 principal components were used for clustering with the Louvain algorithm (resolution = 0.5) UMAP embedding was generated using default parameters (n.neighbors = 30) for visualization. Using characteristic gene markers, we classified the single-cell populations into T cells (CD3D, CD3E, TRAC), monocyte-macrophages (LYZ, CSF3R), NK cells (NKG7), epithelial cells (EPCAM), B cells (CD79A, MS4A1), fibroblasts (COL1A1, FN1), mast cells (MS4A2, TPSB2), endothelial cells (VWF, PECAM1), cDC cells (CD1C), and pDC cells (CLEC4C). Finally, we visualized the clinical information, clustering results, marker gene expression, and cell annotations using UMAP plot to display the reduced dimensions of the single-cell data.

Identification of benign and malignant epithelial cells and subtyping of malignant epithelial cells

To distinguish malignant from benign epithelial cells, we applied inferCNV analysis (https://github.com/broadinstitute/inferCNV). We randomly selected 1,000 normal epithelial cells from control samples and inserted them into the tumor epithelial cell dataset. The remaining normal epithelial cells served as the reference. CNVs were inferred based on expression intensity across genomic regions, using denoise=TRUE and default settings. Cells displaying significant CNV patterns distinct from normal epithelial cells were classified as malignant, while those resembling reference cells were categorized as benign. The CNV scores of epithelial cells were also utilized to assist in distinguishing between benign and malignant epithelial cells.

After isolating all malignant cells, we performed further clustering to categorize them into distinct malignant cell clusters. Using the Seurat package's "FindAllMarkers" function, we identified highly expressed marker genes for each cluster (log2FoldChange > 1, p value < 0.05). Subsequently, with hallmark gene sets from the MsigDB database (30), we applied single-sample gene set enrichment analysis (ssGSEA) via the GSVA package (31) to explore the biological functional characteristics of the malignant cell clusters. Additionally, univariate Cox regression analysis was performed to assess the prognostic significance of marker genes in each malignant cell cluster.

Cell-cell communication analysis

To explore intercellular communication within the tumor microenvironment, we used the CellChat package (32). Receptor-ligand interactions were inferred using the computeCommunProb() function, with a minimum interaction probability threshold of 0.05 to filter out weak interactions. Pathway activity scores were generated using computeCommunProbPathway(), and the rankNet() function was applied to identify the most active signaling pathways. Significant interactions were visualized using netVisual_circle() and netVisual_aggregate(), highlighting key intercellular communication networks. The MDK-NCL signaling pathway emerged as a central interaction hub, particularly enriched in malignant epithelial and stromal cells, and was selected for further spatial and functional analysis.

Processing of spatial transcriptomics data for LUAD

The spatial transcriptomics data were generated using the 10x Genomics Visium platform and processed using the Seurat package (4). Quality control was performed by removing spots with fewer than 500 detected genes or over 10% mitochondrial gene expression. Normalization and variance stabilization were conducted using SCTransform, followed by PCA for dimensionality reduction. The top 30 principal components were used for Louvain clustering (resolution = 0.5). After dimensionality reduction and clustering, we identified nine distinct spatial niches. Based on the expression of MUC1 (tumor region), LYZ (immune region), COL14A1 (stromal region), and SFTPC (normal region), we classified the niches into tumor, immune-stromal, and normal regions.

We then applied MCPcounter analysis (33) to assess the infiltration levels of various cell types (including T cells, B cells, neutrophils, monocytes, fibroblasts, and endothelial cells) in each spot of the spatial transcriptomics data. This allowed us to map the spatial distribution of immune infiltration and compare it with the defined niche regions.

Lastly, using the niches R package (34), we conducted spatial ligand-receptor interaction analysis, which integrates gene expression with spatial proximity. Interaction scores were computed for each ligand-receptor pair between neighboring spots, and only statistically significant pairs (adjusted $p < 0.05$) were retained for downstream analysis. Compared to single-cell analysis, spatial transcriptomics data incorporates spatial localization, providing more biologically accurate ligand-receptor interactions.

Pseudotime analysis

Monocle (35) was used to construct pseudotime trajectories. The "orderCells" function assigned pseudotime values to each cell, and branching events were analyzed to assess transitions between malignant cell states. The MK signaling pathway activity was overlaid on the trajectory to observe its temporal dynamics.

Immune-related analysis of TCGA-LUAD

Using the ESTIMATE package (36), we performed ESTIMATE analysis to assess the overall tumor immune microenvironment in

each LUAD sample. This was achieved by calculating tumor purity, immune score, and stromal score. Additionally, the infiltration levels of 28 different immune cell types in each sample were evaluated using ssGSEA, and detail of 28 immune signature genes were shown in Supplementary Table S1.

Therapy-related analysis of TCGA-LUAD

TMB and MSI are critical factors that influence the interaction between immune cells and tumor cells during immunotherapy. Numerous studies have demonstrated their role in predicting responses to immunotherapy. In this study, we explored the relationship between the MDK-NCL pathway and immunotherapy by comparing TMB and MSI between groups. We also conducted TIDE analysis (Tumor Immune Dysfunction and Exclusion) (37), a widely used method to assess the functional state of T cells in transcriptomic samples, and obtained two key metrics, Dysfunction and Exclusion, which reflect T cell functionality.

In addition, immunogenic cell death (ICD) is another key factor influencing immunotherapy efficacy. We compared the expression patterns of ICD-related genes between groups in TCGA and validation datasets. From the TCIA database (https://www.tcia.at/home), we retrieved predictions of CTLA4 and PDCD1 expression levels in TCGA-LUAD patients and conducted comparisons between the two groups. Lastly, we examined the differential expression profiles of immune checkpoint-related genes, which are closely associated with the response to immune checkpoint inhibitors, between the two groups.

Sample collection

A total of 18 paired LUAD (lung adenocarcinoma) tissues and corresponding adjacent normal tissues were collected from patients undergoing surgical resection at Department of Thoracic Surgery, Shengjing Hospital of China Medical University. All patients included in the study had not received neoadjuvant therapy prior to surgery. The study was approved by the Ethics Committee of Shengjing Hospital, China Medical University (Approval No. 2024PS1727K).

Real-time quantitative PCR

Total RNA was extracted from tissues using the Trizol reagent (R401-01, Vazyme, Nanjing, China) following the manufacturer's protocol. Complementary DNA (cDNA) was synthesized from the extracted RNA using the reverse transcription kit (RR047A, TAKARA, Japan) according to the kit instructions. The relative expression levels of the target genes were determined using β-actin as the internal reference gene. Primer sequences for all genes are listed in Supplementary Table S2. All target gene expression analyses were performed in triplicate to ensure reproducibility.

Western blot

LUAD and control tissue samples were homogenized using ultrasonic disruption and lysed for 30 minutes in RIPA lysis buffer (BL504A, Biosharp, China) containing PMSF (1:100, BL507A, Biosharp, China) and a protease inhibitor cocktail (1:50, P1082, Beyotime, China). Lysates were centrifuged at 12,000 rpm for 20 minutes, and protein concentrations were determined using the BCA protein assay kit (PC0020, Solarbio, China). Proteins were separated by SDS-PAGE (10% gel for MDK and 6% gel for NCL) and transferred onto PVDF membranes (IPVH00010, Millipore, USA). Membranes were blocked with 5% non-fat milk at room temperature for 2 hours and incubated overnight at 4°C with primary antibodies. Afterward, membranes were incubated with secondary antibodies for 2 hours at room temperature. Protein bands were visualized using enhanced chemiluminescence (ECL) reagent (BMU102, Abbkine, USA). Primary antibodies included MDK (1:1000, BM4392, BOSTER, Wuhan, China), NCL (1:1000, A00228-1, BOSTER, Wuhan, China), and GAPDH (1:1000, Sigma, USA), which was used as an internal control. The secondary antibody used was BA1039 (BOSTER, Wuhan, China). All protein bands were quantified using ImageJ software (Rawak Software Inc., Stuttgart, Germany).

Statistical analysis

All data processing and statistical analyses were performed using R software (version 4.1.1). The Mann-Whitney U test (also known as the Wilcoxon rank-sum test) was used to evaluate differences between non-normally distributed variables. Spearman correlation analysis was employed to calculate correlation coefficients between non-normally distributed data. A p-value of less than 0.05 was considered statistically significant.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Ethics Committee of Shengjing Hospital, China Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent

for participation from the participants or the participants' legal guardians/next of kin because We only used human waste tissue.

Author contributions

YF: Data curation, Writing – original draft. SL: Data curation, Writing – original draft. YZ: Visualization, Writing – original draft. XZ: Visualization, Writing – original draft. XM: Visualization, Writing – original draft. RX: Conceptualization, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the 345 Talent Project of Shengjing Hospital of China Medical University to RX.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2025.1546382/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Annotation results of scRNA-seq dataset GSE153935 for LUAD. **(A)** Sample origin of the GSE153935. **(B)** Transcript counts in the GSE153935. **(C)** Clustering results of the GSE153935. **(D)** Cell type annotation based on marker gene expression. **(E)** Expression profiles of representative markers for ten distinct cell types. **(F)** InferCNV heatmap displaying CNVs across cells, with normal samples in the upper panel and tumor samples in the lower panel. Red and blue indicate CNV gains and losses, respectively.

SUPPLEMENTARY FIGURE 2
Single-Cell communication networks for validation dataset. **(A)** Incoming communication patterns of target cells, showing pathways to which each cell type responds. **(B)** Outgoing communication patterns of secreting cells, illustrating the pathways through which cells send signals. **(C)** Network diagram showing the strength of intercellular communication, with connections between various cell types. **(D)** Scatter plot comparing outgoing and incoming communication strengths across cell populations, with bubble size indicating the number of interactions. **(E)** Chord diagram depicting communication via the MK pathway between different cell types. **(F)** Ligand-receptor interaction probabilities within the MK pathway between malignant and other cell types. Dot size represents significance, and color represents communication probability. **(G)** Violin plots of MK pathway gene expression levels across cell types, showing gene activity variations.

SUPPLEMENTARY FIGURE 3
Contribution of each ligand-receptor pair for GSE131907 dataset.

SUPPLEMENTARY FIGURE 4
Contribution of each ligand-receptor pair for GSE153935 dataset.

SUPPLEMENTARY FIGURE 5
The spatial distribution of MDK-NCL receptor-ligand signaling across the niches.

SUPPLEMENTARY FIGURE 6
MDK-NCL ligand-receptor interaction analysis, spatially mapping MDK ligands, NCL receptors, and their binding regions in validation data.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin.* (2021) 71:209–49. doi: 10.3322/caac.21660

2. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature.* (2018) 553:446–54. doi: 10.1038/nature25183

3. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med.* (2013) 19:1423–37. doi: 10.1038/nm.3394

4. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol.* (2018) 36:411–20. doi: 10.1038/nbt.4096

5. Lambrechts D, Wauters E, Boeckx B, Aibar S, Nittner D, Burton O, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med.* (2018) 24:1277–89. doi: 10.1038/s41591-018-0096-5

6. Maynard A, McCoach CE, Rotow JK, Harris L, Haderk F, Kerr DL, et al. Therapy-induced evolution of human lung cancer revealed by single-cell RNA sequencing. *Cell.* (2020) 182:1232–1251.e1222. doi: 10.1016/j.cell.2020.07.017

7. Moncada R, Barkley D, Wagner F, Chiodin M, Devlin JC, Baron M, et al. Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat Biotechnol.* (2020) 38:333–42. doi: 10.1038/s41587-019-0392-8

8. Asp M, Bergenstråhle J, Lundeberg J. Spatially resolved transcriptomes-next generation tools for tissue exploration. *BioEssays: News Rev Mol Cell Dev Biol.* (2020) 42:e1900221. doi: 10.1002/bies.201900221

9. Muramatsu T. Midkine, a heparin-binding cytokine with multiple roles in development, repair and diseases. *Proc Japan Acad Ser B Phys Biol Sci.* (2010) 86:410–25. doi: 10.2183/pjab.86.410

10. Zhang Y, Zuo C, Liu L, Hu Y, Yang B, Qiu S, et al. Single-cell RNA-sequencing atlas reveals an MDK-dependent immunosuppressive environment in ErbB pathway-mutated gallbladder cancer. *J Hepatol.* (2021) 75:1128–41. doi: 10.1016/j.jhep.2021.06.023

11. Filippou PS, Karagiannis GS, Constantinidou A. Midkine (MDK) growth factor: a key player in cancer progression and a promising therapeutic target. *Oncogene.* (2020) 39:2040–54. doi: 10.1038/s41388-019-1124-8

12. Kadomatsu K, Kishida S, Tsubota S. The heparin-binding growth factor midkine: the biological activities and candidate receptors. *J Biochem*. (2013) 153:511–21. doi: 10.1093/jb/mvt035

13. Győrffy B. Transcriptome-level discovery of survival-associated biomarkers and therapy targets in non-small-cell lung cancer. *Br J Pharmacol*. (2024) 181:362–74. doi: 10.1111/bph.16257

14. Carvalho RF, do Canto LM, Abildgaard C, Aagaard MM, Tronhjem MS, Waldstrøm M, et al. Single-cell and bulk RNA sequencing reveal ligands and receptors associated with worse overall survival in serous ovarian cancer. *Cell Communication Signaling: CCS*. (2022) 20:176. doi: 10.1186/s12964-022-00991-4

15. Yu X, Xie L, Ge J, Li H, Zhong S, Liu X. Integrating single-cell RNA-seq and spatial transcriptomics reveals MDK-NCL dependent immunosuppressive environment in endometrial carcinoma. *Front Immunol*. (2023) 14:1145300. doi: 10.3389/fimmu.2023.1145300

16. Chen S, Liao J, Li J, Wang S. GPC2 promotes prostate cancer progression via MDK-mediated activation of PI3K/AKT signaling pathway. *Funct Integr Genomics*. (2024) 24:127. doi: 10.1007/s10142-024-01406-y

17. Du Y, Shi J, Wang J, Xun Z, Yu Z, Sun H, et al. Integration of pan-cancer single-cell and spatial transcriptomics reveals stromal cell features and therapeutic targets in tumor microenvironment. *Cancer Res*. (2024) 84:192–210. doi: 10.1158/0008-5472.CAN-23-1418

18. Muramatsu T. Midkine: a promising molecule for drug development to treat diseases of the central nervous system. *Curr Pharm Design*. (2011) 17:410–23. doi: 10.2174/138161211795164167

19. Neumaier EE, Rothhammer V, Linnerbauer M. The role of midkine in health and disease. *Front Immunol*. (2023) 14:1310094. doi: 10.3389/fimmu.2023.1310094

20. Sarkans U, Füllgrabe A, Ali A, Athar A, Behrangi E, Diaz N, et al. From arrayExpress to bioStudies. *Nucleic Acids Res*. (2021) 49:D1502–6. doi: 10.1093/nar/gkaa1062

21. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res*. (2013) 41:D991–995. doi: 10.1093/nar/gks1193

22. Kim N, Kim HK, Lee K, Hong Y, Cho JH, Choi JW, et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat Commun*. (2020) 11:2285. doi: 10.1038/s41467-020-16164-1

23. Hanley CJ, Waise S, Ellis MJ, Lopez MA, Pun WY, Taylor J, et al. Single-cell analysis reveals prognostic fibroblast subpopulations linked to molecular and immunological subtypes of lung cancer. *Nat Commun*. (2023) 14:387. doi: 10.1038/s41467-023-35832-6

24. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signaling*. (2013) 6:pl1. doi: 10.1126/scisignal.2004088

25. Takeuchi T, Tomida S, Yatabe Y, Kosaka T, Osada H, Yanagisawa K, et al. Expression profile-defined classification of lung adenocarcinoma shows close relationship with underlying major genetic changes and clinicopathologic behaviors. *J Clin Oncol: Off J Am Soc Clin Oncol*. (2006) 24:1679–88. doi: 10.1200/JCO.2005.03.8224

26. Kabbout M, Garcia MM, Fujimoto J, Liu DD, Woods D, Chow CW, et al. ETS2 mediated tumor suppressive function and MET oncogene inhibition in human non-small cell lung cancer. *Clin Cancer Res: Off J Am Assoc Cancer Res*. (2013) 19:3383–95. doi: 10.1158/1078-0432.CCR-13-0341

27. Moreno Leon L, Gautier M, Allan R, Ilié M, Nottet N, Pons N, et al. The nuclear hypoxia-regulated NLUCAT1 long non-coding RNA contributes to an aggressive phenotype in lung adenocarcinoma through regulation of oxidative stress. *Oncogene*. (2019) 38:7146–65. doi: 10.1038/s41388-019-0935-y

28. Choudhary S, Satija R. Comparison and evaluation of statistical error models for scRNA-seq. *Genome Biol*. (2022) 23:27. doi: 10.1186/s13059-021-02584-9

29. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods*. (2019) 16:1289–96. doi: 10.1038/s41592-019-0619-0

30. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. (2015) 1:417–25. doi: 10.1016/j.cels.2015.12.004

31. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinf*. (2013) 14:7. doi: 10.1186/1471-2105-14-7

32. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. (2021) 12:1088. doi: 10.1038/s41467-021-21246-9

33. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol*. (2016) 17:218. doi: 10.1186/s13059-016-1070-5

34. Raredon MSB, Yang J, Kothapalli N, Lewis W, Kaminski N, Niklason LE, et al. Comprehensive visualization of cell–cell interactions in single-cell and spatial transcriptomics with NICHES. *Bioinformatics*. (2023) 39:btac775. doi: 10.1093/bioinformatics/btac775

35. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. (2014) 32:381–6. doi: 10.1038/nbt.2859

36. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. (2013) 4:2612. doi: 10.1038/ncomms3612

37. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med*. (2018) 24:1550–8. doi: 10.1038/s41591-018-0136-1

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 24 June 2025
DOI 10.3389/fimmu.2025.1560903

OPEN ACCESS

EDITED BY
Yafeng Zhu,
Sun Yat-sen Memorial Hospital, China

REVIEWED BY
Jie Zhang,
Peking University Third Hospital, China
Chenxi Xia,
Peking University, China
Yangpo Cao,
Southern University of Science and Technology, China

*CORRESPONDENCE
Qinghua Han
hqh@sxmu.edu.cn

RECEIVED 15 January 2025
ACCEPTED 05 June 2025
PUBLISHED 24 June 2025

CITATION
Wang H, Wu D, Han G, Yan J, Wang Z, He X, Chen Y, Wang Y and Han Q (2025) Immunogenic cell death-related biomarkers in heart failure probed by transcriptome and single-cell sequencing.
Front. Immunol. 16:1560903.
doi: 10.3389/fimmu.2025.1560903

Immunogenic cell death-related biomarkers in heart failure probed by transcriptome and single-cell sequencing

Haoyue Wang[1,2], Dongdong Wu[2,3], Gangfei Han[1,2], Jingjing Yan[4], Zehui Wang[1], Xing He[3,5], Yuxiang Chen[2,3], Yan Wang[3] and Qinghua Han[1,6,7]*

[1]Department of Cardiology, The First Hospital of Shanxi Medical University, Taiyuan, Shanxi, China, [2]The First Clinical Medical College, Shanxi Medical University, Taiyuan, Shanxi, China, [3]Department of Hepatobiliary Surgery and Liver Transplantation Center, The First Hospital of Shanxi Medical University, Taiyuan, Shanxi, China, [4]Department of Health Statistics, School of Public Health, Shanxi Provincial Key Laboratory of Major Diseases Risk Assessment, Shanxi Medical University, Taiyuan, Shanxi, China, [5]College of Basic Medicine, Shanxi Medical University, Taiyuan, Shanxi, China, [6]Key laboratory of Cellular Physiology at Shanxi Medical University, Ministry of Education, Taiyuan, Shanxi, China, [7]Shanxi Innovation Center for Integrated Management of Hypertension, Hyperlipidemia and Hyperglycemia Correlated with Cardiovascular and Cerebrovascular Diseases, Taiyuan, Shanxi, China

Background: Heart failure (HF) represents the terminal stage of various cardiovascular disorders, with immunogenic cell death (ICD) potentially influencing HF progression through modulation of immune cell activity. This study aimed to identify ICD-associated biomarkers in patients with HF and explore their underlying mechanisms.

Methods: Data from GSE57338, GSE3586 and GSE5406 were retrieved from the Gene Expression Omnibus (GEO) database. Differential expression analysis and weighted gene co-expression network analysis (WGCNA) were employed to identify candidate genes, followed by enrichment analysis and Protein-Protein Interaction (PPI) network construction. Candidate biomarkers were selected using two machine learning approaches and validated for expression levels, with receiver operating characteristic (ROC) curve analysis determining the final biomarkers. A nomogram model was built based on the biomarkers, followed by molecular regulatory network analysis, gene set enrichment analysis (GSEA), immune infiltration assessment, and drug prediction. Additionally, key cells were selected for pseudo-time and cell communication analysis using the GSE183852 dataset. Next, pseudotemporal analysis was also performed on key cell subpopulations. Real-time quantitative PCR (RT-qPCR) was employed to validate the biomarkers.

Results: Three biomarkers, CD163, FPR1, and VSIG4, were identified as having significant diagnostic value for HF. GSEA revealed their enrichment in ribosomal and immune cell-related pathways. These biomarkers were notably correlated with CD8 T cells and M2 macrophages. Carbachol and etynodiol were predicted to interact with all three biomarkers. Single-cell RNA sequencing identified nine cell types, with expression of the biomarkers confined to monocytes and macrophages. Strong cell communication was observed between these cell types and fibroblasts. Expression of CD163 and VSIG4 decreased over time in monocytes and macrophages, whereas FPR1 showed an upward trend. In

addition, the expression levels of CD163 and VSIG4 increased in subpopulations of monocytes and macrophages, whereas FPR1 showed a decreasing trend. RT-qPCR results confirmed significant down-regulation of CD163, FPR1, and VSIG4 in patients with HF and animal models.

Conclusions: This study identified and validated three ICD-related biomarkers in HF—CD163, FPR1, and VSIG4—offering a novel theoretical foundation for the clinical diagnosis and treatment of HF.

KEYWORDS

immunogenic cell death, heart failure, biomarker, single-cell RNA sequencing analysis, monocytes and macrophages

1 Introduction

Heart failure (HF), the terminal stage of various cardiovascular diseases, affects approximately 56.2 million people worldwide (1, 2). Despite lifestyle changes and advances in medical care that have stabilized age-adjusted incidence rates, the prevalence and mortality rates of HF remain high, highlighting the need for further research to identify improved management strategies (3). Although HF was once considered non-immune-mediated, recent studies have demonstrated the involvement of the immune system in its pathophysiology, and clinical trials on immune modulation therapy for HF have been conducted (4). Consequently, modulating immune responses to maintain stability may serve as a promising strategy to delay HF progression.

Immunogenic cell death (ICD), a unique form of regulated cell death that occurs as a downstream effect of tumor-specific immune responses, has been extensively studied in cancer immunotherapy (5, 6), with emerging research in cardiovascular diseases. Endothelial cell ICD in atherosclerosis has been linked to the initiation of adaptive immune responses, sustaining chronic inflammation within plaques (7). In coronary artery disease, stratification based on ICD-related genes (IRGs) enables the development of risk models and immune subtypes that facilitate treatment decisions (8). Moreover, ICD has been explored as a diagnostic tool for ischemic stroke in elderly women, identifying key biomarkers for diagnosis (9). However, the mechanisms underlying ICD in HF remain unexplored.

This study utilized machine learning techniques to identify ICD biomarkers in HF, followed by immune infiltration analysis, targeted drug prediction, gene set enrichment analysis (GSEA), single-cell data clustering and annotation, cell communication analysis, and pseudotime analysis. The findings revealed the functional and potential molecular mechanisms of these biomarkers at both the transcriptomic and cellular levels, providing a novel theoretical framework for the clinical diagnosis and treatment of HF.

Abbreviations: HF, Heart failure; ICD, Immunogenic cell death; GEO, Gene Expression Omnibus; PPI, Protein-Protein Interaction; ROC, Receiver operating characteristic; GSEA, Gene set enrichment analysis; DEGs, Differential expression genes; ssGSEA, Single sample gene set enrichment analysis; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; STRING, Search tool for the retrieval of interacting genes; MCC, Maximum Connectivity Component; MNC, Minimum Network Connectivity; DMNC, Degree of Minimum Network Connectivity; SVM-RFE, Support vector machine-recursive feature elimination; AUC, Area under the curve; TFs, Transcription factors; HVGs, Highly variable genes; PCA, Principal component analysis; UMAP, Uniform Manifold Approximation and Projection.

2 Materials and methods

2.1 Data collection

RNA data from GSE57338 (sequencing platform: GPL11532) was obtained from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/), comprising 136 normal left ventricular tissue samples and 177 left ventricular tissue samples from patients with HF (10). Additionally, RNA data from GSE3586 (sequencing platform: GPL3050) was downloaded, containing 15 normal left ventricular tissue samples and 13 left ventricular tissue samples from patients with HF (11). Moreover, the GSE5406 dataset contained 16 normal and 194 HF patients' heart tissue samples. The data were obtained from the GPL96 platform using chip sequencing technology, mainly for biomarkers expression validation. The single-cell dataset GSE183852 was retrieved from the GEO website (sequencing platform: GPL24676), including heart tissue samples from 5 patients with HF and 2 normal heart tissue samples (12). A total of 34 ICD-associated genes were obtained from the literature (13) (Additional file 1).

2.2 Differential expression analysis

Differential expression analysis was conducted using the R package "limma" (v 3.58.1) (14), applying the screening criteria of

$|log_2 fold\ change\ (FC)| > 0.5$ and $P < 0.05$ to compare HF and control samples in the GSE57338 dataset. Volcano plots of the differentially expressed gene (DEGs) were visualized using the R package "ggplot2" (v 3.4.1) (15), highlighting the top 10 up- and down-regulated DEGs. Heatmaps of the top 10 DEGs were generated using the R package "ComplexHeatmap" (v 2.4.0) (16).

2.3 Weighted gene co-expression networks analysis

To calculate the single-sample gene set enrichment analysis (ssGSEA) scores for ICD-related genes across 313 samples, the ssGSEA algorithm from the R package "GSVA" (v 1.46.0) (17) was applied, and box plots were created using "ggplot2" (v 3.4.1).

WGCNA was performed on the GSE57338 dataset using the R package "WGCNA" (v 1.72.5) (18), with ssGSEA scores as the feature. Initial clustering of samples identified and excluded abnormal samples. The soft threshold (power) was determined based on an $R^2 > 0.85$ and mean connectivity = 0. The dynamic tree cutting algorithm, with a minimum gene number of 50 per module and a module merging threshold of 0.3, was applied to define gene modules. Genes were color-coded, and the "grey" module (containing unclassified genes) was excluded. Pearson correlation coefficients were calculated between the modules and ssGSEA scores, with a heatmap generated to highlight modules with significant correlation ($|cor| > 0.5$, $P < 0.05$). Genes within these modules were identified as key module genes.

2.4 Enrichment analysis of candidate genes and protein-protein interactions network analysis

The R package "ggvenn" (v 1.7.3) (19) was employed to identify the intersection between DEGs and key module genes, resulting in the selection of candidate genes. These genes were then converted from SYMBOL to ENTREZID using the human genome database org.Hs.eg.db (v 3.18.0) (20). Candidate genes underwent Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analysis with the R package "ClusterProfiler" (v 3.16.0) (21), with a threshold of $P < 0.05$. To construct PPI networks, candidate genes were analyzed using the search tool for the retrieval of interacting genes (STRING) database (https://string-db.org) with a confidence score of 0.4. PPI networks were then visualized with Cytoscape (v 3.10.0) (22). The Cytohubba plugin in Cytoscape (v 3.10.0) was utilized to rank candidate genes using six algorithms: Maximum Connectivity Component (MCC), Minimum Network Connectivity (MNC), Degree of Minimum Network Connectivity (DMNC), Degree, Closeness, and Betweenness. Based on the ranking results, the top 20 genes from each algorithm were extracted, and their intersection was used to identify the final candidate key genes. UpSet plots were generated using the R package (v 1.4.0) (23).

2.5 Screening candidate biomarkers by machine learning

Candidate key genes were further screened based on sample grouping information from GSE57338 using the support vector machine-recursive feature elimination (SVM-RFE) algorithm (10-fold cross validation) (v 4.1.4) (24) to obtain feature genes. The R package "randomForest" (v 3.2.2) (25) was used for random forest algorithm analysis of the feature genes, incorporating sample grouping information from GSE57338. A total of 500 decision trees were computed using the randomForest function, and the MeanDecreaseGini values for each feature gene were visualized in a bar chart. The median of the MeanDecreaseGini values (MeanDecreaseGini measures the effect of each variable on the heterogeneity of observations at each node in the classification tree, thus assessing the importance of the variable. The larger the value, the higher the importance of the variable) was calculated, and genes with values above the median were selected as candidate biomarkers. Correlation analysis of the candidate biomarkers was performed using the R package "corrplot" (v 0.92) (26), with thresholds of $|cor| > 0.3$ and $P < 0.05$.

2.6 Expression validation of candidate biomarkers

Expression differences of candidate biomarkers between HF and normal samples were analyzed using the grouping information from GSE3586 and GSE57338, with a threshold of $P < 0.05$. Box plots were constructed using the R package "ggplot2" (v 3.4.1). Candidate biomarkers showing differential expression between groups and consistent trends across both datasets were selected for receiver operating characteristic (ROC) analysis. ROC curves for candidate biomarkers were generated using the R package "pROC" (v 1.18.0) (27), and the area under the curve (AUC) was calculated, with biomarkers defined as those having an AUC > 0.7. To validate biomarkers expression, differential expression analysis was performed in the GSE5406 dataset.

2.7 Construction of a nomogram

In the GSE57338 dataset, a nomogram was constructed using the R package "rms" (v 5.1.4) (28) to evaluate the risk of developing HF, based on the expression of identified biomarkers. The predictive performance of the nomogram was assessed by plotting the ROC curve with the R package "pROC" (v 1.18.0).

2.8 Gene set enrichment analysis

Spearman correlation analysis was performed between each biomarker and the remaining genes across all GSE57338 samples using the R package "psych" (v 2.2.9) (29), generating correlation

coefficients. Genes were then ranked according to these coefficients, yielding gene lists associated with each biomarker. GSEA was performed using the sorted results and the R package "ClusterProfiler" (v 3.16.0), with "c2.kegg.v7.4.symbols.gmt" and "c5.go.v7.4.symbols.gmt" from the Molecular Signatures Database (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb/index.jsp) as reference gene sets. The top 5 most significant signaling pathways were visualized using the enrichplot package ($P < 0.05$ and $|$ Normalized Enrichment Score (NES)$| > 1$) (v 1.18.3) (20).

2.9 Immune infiltration analysis

The CIBERSORT algorithm (v 1.03) (30) was employed to calculate the relative abundance of 22 immune cell types (31) in HF and normal samples from the GSE57338 dataset. Immune cells with a result of 0 were excluded. Differential immune cells ($P < 0.05$) were identified, and box plots were constructed for visualization. Spearman correlation analysis was used to assess the relationships among differential immune cells and between biomarkers and immune cells ($|cor| > 0.3$ and $P < 0.05$). A correlation matrix was created using the R package "corrplot" (v 0.92) (26), and a heatmap was plotted using the R package "pheatmap" (v 1.0.12) (32).

2.10 Regulatory network analysis

MiRNAs targeting the biomarkers were predicted using the microRNA database (miRDB, http://mirdb.org) and the starBase database (http://starbase.sysu.edu.cn/), and the intersection of miRNAs from both databases was extracted. Based on these predictions, a miRNA-biomarker network was constructed using Cytoscape (v 3.10.0). Transcription factors (TFs) related to the biomarkers were identified using the TRRUST database (http://www.grnpedia.org/trrust/), while the disease signatures database (DSigDB, https://www.dsigdb.org/) was used to identify drugs targeting the biomarkers. A biomarker-drug network was then created and visualized.

2.11 Single-cell RNA sequencing analysis

The single-cell RNA sequencing data from GSE183852 were processed into Seurat objects using the R package "Seurat" (v 4.4.0) (33). Quality control was performed by applying the following parameters: 200 < nFeature_RNA < 4,000, nCount_RNA < 10,000, and Mt < 10%. Genes covered by fewer than three cells were removed. Hypervariable genes were selected using variance stabilization transformation (vst), and the highly variable genes (HVGs) were retained for further analysis. The LabelPoints function was applied to identify the top 10 most variable genes, and the Scale Data function was used for normalization. Principal component analysis (PCA) was performed on the HVGs for dimensionality reduction. The p-value for PCs 1 to 15 was calculated using the Jackstraw function, and variance drop values for PCs were computed using the Elbowplot function. Based on the elbow plot, appropriate PCs were selected for subsequent analysis ($P < 0.05$). Uniform Manifold Approximation and Projection (UMAP) clustering analysis was applied to identify cell clusters (resolution = 0.5). Cellular annotation was performed according to the literature (12). The Dotplot function was used to visualize the expression of the three biomarkers in the cells, and cells expressing all three biomarkers were selected as key cells. Enrichment analysis for each cell subtype was conducted using the analyze_sc_clusters function from the R package "ReactomeGSA" (v 1.12.0) (34). The pathways function was used to extract enrichment results, and a heatmap displayed the top ten enriched pathways in each cell subtype. Cell subtype interactions were explored using the R package "CellChat" (v 1.6.1) (35) to conduct communication analysis. Trajectory differentiation of key cell clusters was simulated using the R package "Moncle" (v 2.30.1) (36). The dynamic trend of biomarker expression during cell differentiation was plotted using the plot_pseudotime_heatmap function. Next, the marker genes of key cell subpopulations were selected for annotation based on the CellMarker 2.0 database (https://ngdc.cncb.ac.cn/databasecommons/database/id/6110), and the final key cell subpopulations were identified based on the specific expression of these genes in different clusters. To further explore the expression dynamics and temporal trajectories of biomarkers in the key cells, the annotated key cell subpopulations were analyzed by the proposed timeline trajectory analysis. Using the R package Monocle2 (v 2.24.1) (37), the distribution of biomarkers in each key cell subtype was projected onto a root and multiple branches, a single-cell trajectory map was constructed, and the dynamic trend of biomarker expression during cell differentiation was plotted. Subsequently, in order to analyze the relationship between differentiation states and subtypes of key cells, stacked maps of cell subpopulations in different differentiation states were drawn. Based on the subtype annotation results, the proportions of cell types under different groupings were first visualized. Wilcoxon test. Finally, the differences in the expression of NOS2, TNF, ARG1, and MRC1 genes in Monocyte&Macrophage between HF and control samples were analyzed and statistically analyzed using the Wilcoxon test.

2.12 Human Subjects and Extraction of PBMC

Patients with HF admitted to the First Hospital of Shanxi Medical University were selected as the HF group, and a control group was matched with the HF group based on age, gender, and other underlying diseases besides HF. Based on the expression of biomarkers obtained through bioinformatics, the sample size was calculated using PASS.15, resulting in a total of 15 pairs of samples. In the morning of the second day after admission, venous blood was collected into EDTA tubes, and peripheral blood lymphocytes were isolated within 2 hours using human peripheral blood lymphocyte separation liquid (Solarbio, China). The trial protocol was approved by the Scientific Research Ethics Review Committee of the First Hospital of Shanxi Medical University (NO. KYLL-2024-236), and all patients provided written informed consent.

2.13 Animal model (echocardiography)

SSPF-grade male Sprague-Dawley rats (180–200 g, 6–8 weeks old) were used to establish a chronic HF model (38). HF was induced by permanently ligating the left coronary artery in rats, while sham-operated rats underwent the same surgical procedure without artery ligation. Six weeks post-ligation, high-resolution echocardiography was performed using the Vevo 770 system (Visualsonics) with a 40 MHz RMV 704 scanhead to assess cardiac function. Rats with an ejection fraction (EF) < 40% were considered to have successfully developed HF, and those that did not develop HF were excluded. After completing echocardiography, the animals were euthanized, and tissues were collected for analysis. The experimental protocol was approved by the Animal Experimental Center Ethics Committee of Beijing Yongxinkangtai Science and Technology Development Co., Ltd. (NO. YXKT2024L010).

2.14 Staining

Hearts were fixed in 4% paraformaldehyde at room temperature for 48 hours, followed by dehydration and embedding. The samples were sectioned at 5μm thickness, dewaxed, rehydrated, and stained with Hematoxylin and Eosin (HE) and Masson stains. For IHC staining, primary antibodies targeting CD163 (1:200, Selleck, F1548) was incubated overnight at 4 °C. Then, second antibody was incubated at 37°C for 1 hour. Chromogen development was accomplished with DAB. Images were captured under a microscope (Olympus, Japan).

2.15 Real-time quantitative PCR

Following tissue homogenization, total RNA was extracted using Trizol (Thermo Fisher Scientific, USA). cDNA synthesis was carried out using PrimeScript RT Master Mix (Takara, Japan) according to the manufacturer's protocol. Real-time quantitative PCR (qPCR) analysis was performed with SYBR Green Master Mix (DBI Bioscience, Germany) on a QuantStudio3 real-time PCR instrument (Thermo Fisher Scientific, USA), with GAPDH as an internal control. Relative mRNA expression levels were quantified using the $2^{-\Delta\Delta Ct}$ method. Primer sequences are provided in Table 1.

2.16 Statistical analysis

Statistical analyses were conducted using R software (v 4.2.2) and GraphPad Prism 9. Differences between two groups were assessed using the Wilcoxon rank sum test, with statistical significance defined as $P < 0.05$.

3 Results

3.1 Acquisition of key module genes

A total of 441 DEGs were identified, including 236 up-regulated and 205 down-regulated genes in HF (Additional files 2a-b). The ssGSEA scores for ICD-related genes significantly differed between

TABLE 1 Primer sequences for quantitative real-time PCR.

Species	Target gene		Primer sequence (5'to3')
Human	VSIG4	Forward	AAGCAACATCTACAGTGAAGCAGTC
		Reverse	ATGATGAGGATGATGGCAAAGACAG
	FPR1	Forward	AGTGGACATCAACTTGTTCGGAAG
		Reverse	ACGGTGCGGTGGTTCTGG
	CD163	Forward	ACAATGAAGATGCTGGCGTGAC
		Reverse	TCTCTGAATCTCCACCTCAACTGTC
	GAPDH	Forward	CGTATCGGACGCCTGGTT
		Reverse	AGGTCAATGAAGGGGTCGTT
Rat	VSIG4	Forward	AGCTGCCGATCTTTGCCATAATC
		Reverse	TCCTGCTCACCTCATAGACATACTC
	FPR1	Forward	CCGTGAACACTTGAGGAACATACC
		Reverse	GGATTGGGTTGAGGCAGCTATTG
	CD163	Forward	GAATCACAGCATGGCACAGGTC
		Reverse	CACAAGAGGAAGGCAATGAGAAGG
	GAPDH	Forward	GACATGCCGCCTGGAGAAAC
		Reverse	AGCCCAGGATGCCCTTTAGT

HF and normal samples (Additional file 2c). In the WGCNA analysis of the GSE57338 dataset, no outlier samples were detected (Additional file 2d). The soft threshold was determined to be 7 (Additional file 2e). Similar modules were merged from the co-expression matrix, resulting in 11 identified gene modules (excluding the gray module for unclassified genes), with each module represented by a different color (Figure 1a). The yellow module (cor = 0.72, $P = 5.8 \times 10^{-17}$) demonstrated the strongest correlation with ICD-related gene ssGSEA scores. Consequently, the 432 genes within the yellow module were designated as key module genes (Figure 1b).

3.2 Identification and enrichment analysis of candidate genes and PPI

In this study, 47 candidate genes were identified through the intersection of DEGs and key module genes (Figure 2a). The obtained candidate genes were subject to gene ID conversion, though FCGR1B could not be successfully converted. GO enrichment analysis revealed 272 GO terms, comprising 224 biological processes (BP), 24 cellular components (CC), and 24 molecular functions (MF) ($P < 0.05$) (Figure 2b). The candidate genes were significantly enriched in pathways such as the positive regulation of inflammatory response, secretory granule membrane, and RAGE receptor binding. Additionally, the candidate genes were enriched in 26 KEGG pathways ($P < 0.05$), including staphylococcus aureus infection, phagosome, and neutrophil extracellular trap formation (Figure 2c). These results implied that candidate genes may play important roles in antimicrobial immunity, inflammatory response and cellular damage repair.

The candidate genes were further subjected to PPI network construction, resulting in 42 genes, such as TLR2, FPR1, and MRC1, and 240 gene-to-gene pairs, including TLR2-CD163 and VSIG4-CD14 (Figure 2d).

To optimize the screening of candidate genes, the genes were ranked using different algorithms. The top 20 genes from each algorithm were extracted, and the intersection of these top 20 genes was taken. Finally, 16 genes were identified as the candidate key genes for further analysis (Figure 2e).

3.3 Machine learning for candidate biomarker screening

Based on the sample grouping information from GSE57338, the SVM-RFE algorithm was applied for screening, resulting in 13 feature genes: CD163, VSIG4, FCER1G, CCR1, CCL5, FPR1, TLR2, C1QB, CD14, MSR1, CD68, MRC1, and CYBB (Figure 3a). The MeanDecreaseGini values for each feature gene ranged from 0 to 30, with notable differences observed between the genes (Figure 3b). By calculating the median of the MeanDecreaseGini values, six genes greater than the median were selected as candidate biomarkers: CD163, VSIG4, FCER1G, CCR1, CCL5, and FPR1. Among these, CCL5 showed a negative correlation with VSIG4 and CD163, while the remaining five genes exhibited positive correlations with each other ($P < 0.01$) (Figure 3c). The correlation between these genes suggested that they may work in concert at different stages of the immune response or in different types of immune cells.

3.4 Diagnosis and evaluation of biomarkers

In GSE57338, the six candidate biomarkers demonstrated significant differences between HF and normal samples ($P < 0.05$), with CD163, FPR1, and VSIG4 showing decreased expression in HF samples (Figure 4a). In GSE3586, only CD163, VSIG4, CCR1, and FPR1 were expressed, with CD163, FPR1, and VSIG4 levels significantly reduced in HF samples, consistent with the expression patterns observed in GSE57338 (Figure 4b). Consequently, CD163, FPR1, and VSIG4 were selected for ROC analysis, which revealed that the AUC for all three biomarkers exceeded 0.7 in both datasets, confirming their potential as HF biomarkers (Figures 4c-h). Next, the expression analysis of CD163, FPR1, and VSIG4 in the GSE5406 dataset showed that all three were significantly under-expressed in the HF group compared to the

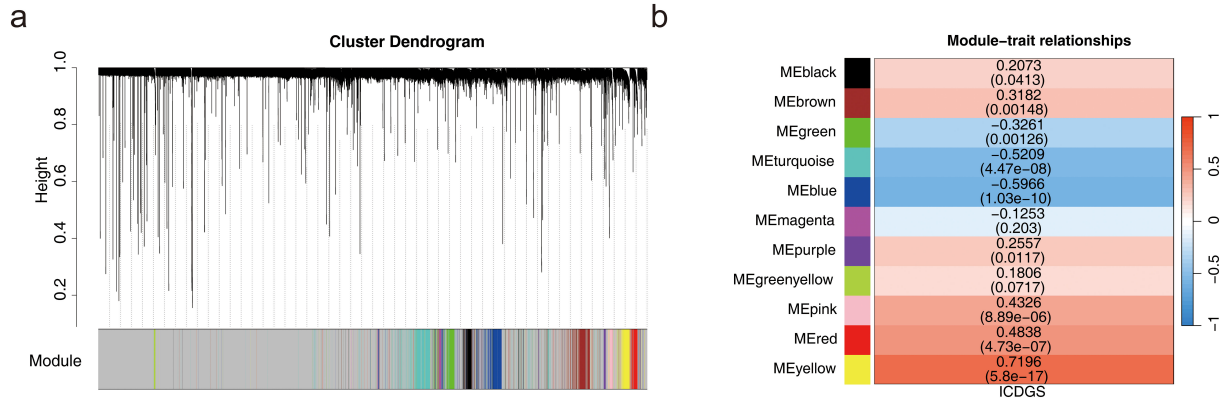


FIGURE 1
Acquisition of key module genes. **(a)** Co-expression module identification. **(b)** Heatmap showing the correlation between modules and phenotypes.

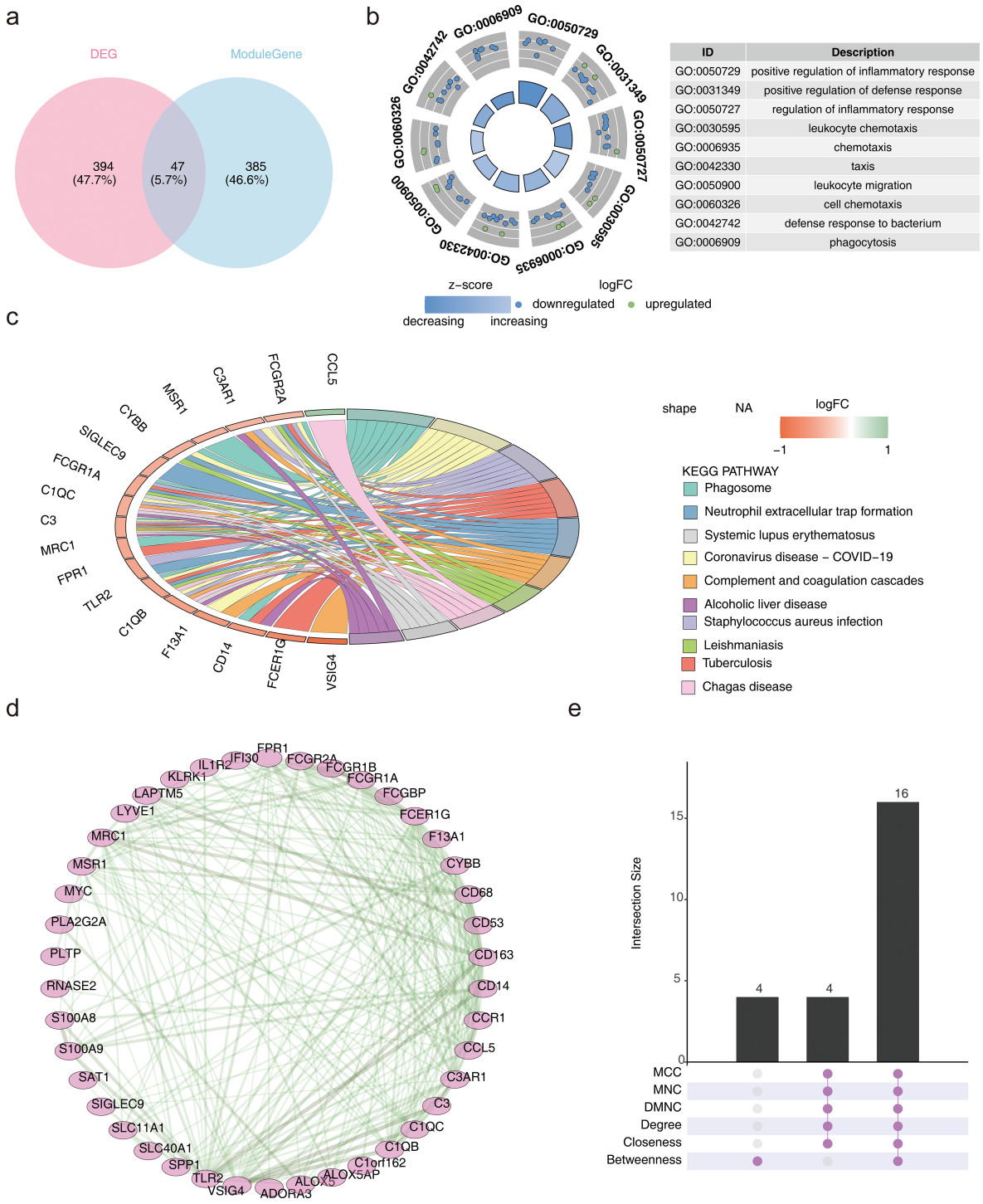


FIGURE 2
Identification and enrichment analysis of candidate genes and PPI. **(a)** Venn diagram depicting the overlap between differentially expressed genes (DEGs) and key module genes. **(b)** Gene Ontology (GO) enrichment analysis results. **(c)** Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis results. **(d)** Protein-Protein Interaction (PPI) network. **(e)** Upset plot representing the PPI network.

normal group (Additional file 3). The expression patterns were consistent with those in the GSE57338 and GSE3586 datasets.

The nomogram model demonstrated that these three biomarkers could accurately predict the risk of HF occurrence. ROC analysis of the nomogram yielded an AUC of 0.913, indicating that the predictive accuracy of the nomogram model was significantly superior to single-gene predictions (Additional file 4). It also suggested that the onset and progression of HF may involve complex interactions of multiple genes or biological pathways.

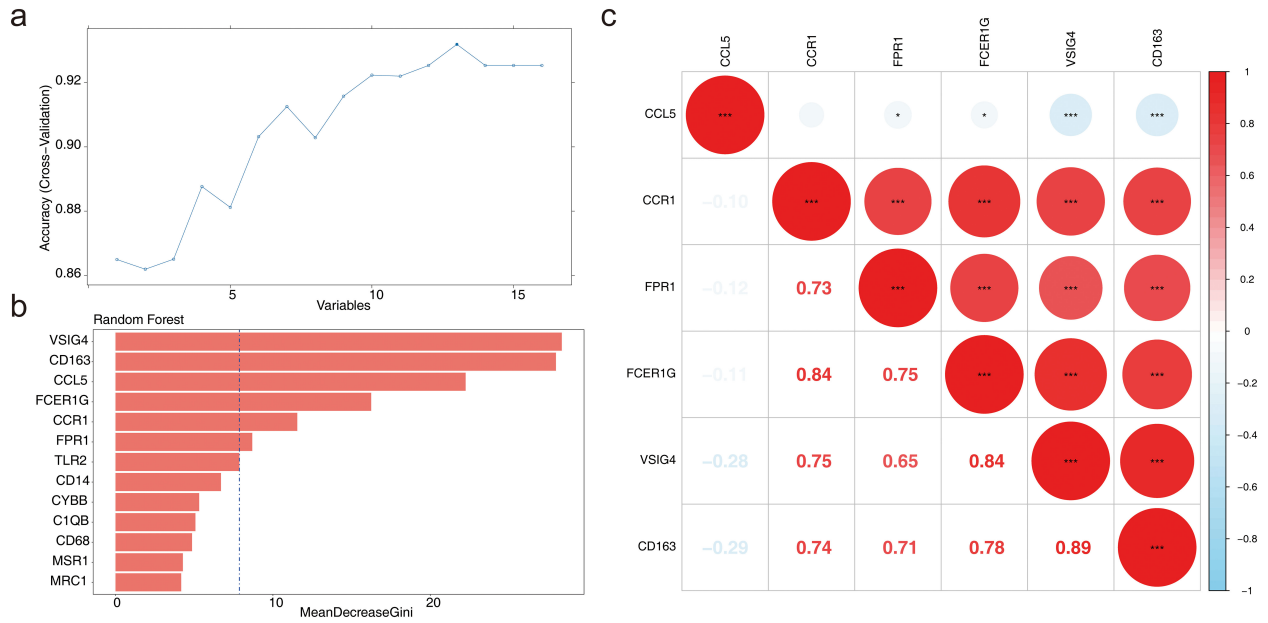


FIGURE 3
Machine learning for candidate biomarker screening. **(a)** Results of the Support Vector Machine-Recursive Feature Elimination (SVM-RFE) model. **(b)** Bar chart depicting the MeanDecreaseGini scores for candidate genes. **(c)** Correlation analysis of candidate biomarkers. *$P < 0.05$, ***$P < 0.001$.

3.5 Functional analysis of biomarkers

Further analysis of the signaling pathways involving CD163, FPR1, and VSIG4 revealed that CD163 was enriched in 76 pathways, including ribosome, Parkinson's disease, leishmania infection, Fc gamma R-mediated phagocytosis, and B cell receptor signaling (Figure 5a). FPR1 was enriched in 79 pathways, including ribosome, leishmania infection, Parkinson's disease, cytokine-cytokine receptor interaction, and chemokine signaling (Figure 5b). VSIG4 was enriched in 85 pathways, including ribosome, Fc gamma R-mediated phagocytosis, B cell receptor signaling, leishmania infection, and chemokine signaling (Figure 5c). Notably, all three biomarkers were enriched in pathways related to ribosome function, immune cells, and immune factors. These findings provided a basis for further investigation of the potential applications of biomarkers in immunomodulation, disease diagnosis and therapy.

3.6 Analysis of immune cell infiltration

To further explore immune status differences between HF and normal samples, immune infiltration analysis was performed on GSE57338 samples, revealing differences in the abundance of 22 immune cell types between samples from patients with HF and normal samples (Additional file 5). Immune cells with a result of 0 in 30% of the samples were excluded, leaving 12 immune cell types for subsequent analysis. Five immune cell types showed significant differences between the groups: M2 macrophages, resting mast cells, plasma cells, $CD8^+$ T cells, and T regulatory cells (Tregs) ($P < 0.05$) (Figure 6a). Correlation analysis among these five immune cell types revealed a strong positive correlation between $CD8^+$ T cells and Tregs, while plasma cells exhibited negative correlations with Tregs, $CD8^+$ T cells, M2 macrophages, and resting mast cells ($|cor| > 0.3$, $P < 0.05$) (Figure 6b). The correlation heatmap between biomarkers and the five immune cell types showed that VSIG4 had a strong positive correlation with M2 macrophages, and M2 macrophages positively correlated with CD163 and FPR1. In contrast, $CD8^+$ T cells and plasma cells negatively correlated with CD163, FPR1, and VSIG4, respectively. Resting mast cells demonstrated an inverse correlation with CD163 and FPR1 ($|cor| > 0.3$, $P < 0.05$) (Figure 6c). The above results suggested that biomarkers may be involved in disease onset and progression by modulating immune responses and cellular functions.

3.7 Molecular regulatory network and drug prediction

Prediction of miRNA interactions with the three biomarkers revealed that VSIG4 was regulated by four miRNAs, including hsa-miR-665; CD163 was regulated by 11 miRNAs, including hsa-miR-4262; while no miRNA regulatory relationships were found for FPR1 (Figure 7a). TFs regulating the biomarkers were also analyzed, revealing that no TFs regulated VSIG4 or FPR1, but eight TFs, including SOX9, were found to regulate CD163 (Figure 7b). These findings provided important clues for further understanding of immune markers and their regulatory networks in HF.

A total of 74 biomarker-drug/compound relationships were identified. The network analysis suggested that carbachol and

10.3389/fimmu.2025.1560903

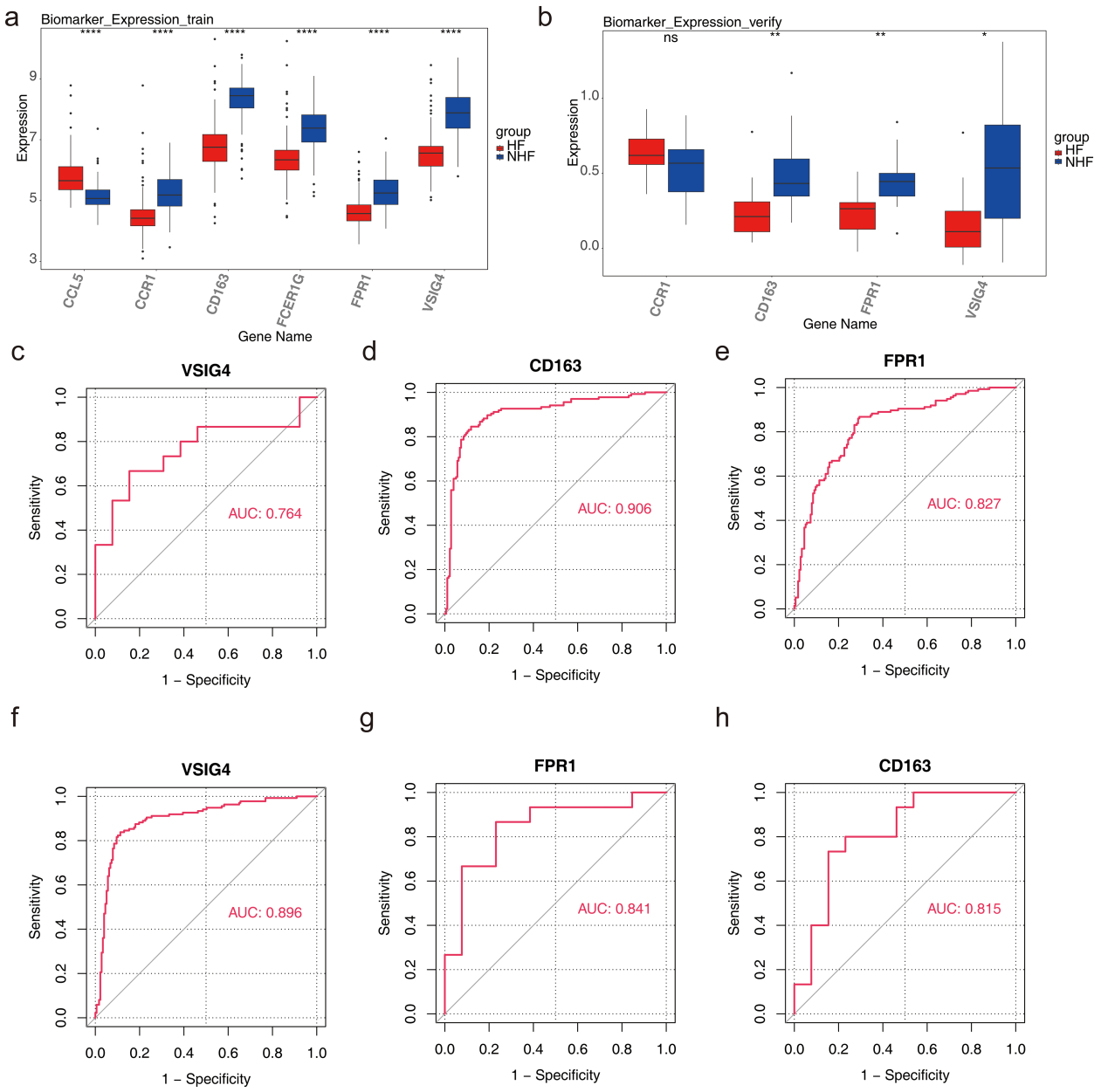


FIGURE 4
Diagnosis and evaluation of biomarkers. **(a)** Expression levels of candidate genes in the training set, with the horizontal axis representing genes and the vertical axis indicating gene expression levels (Wlicoxon rank sum test, ****$P < 0.0001$). **(b)** Expression levels of candidate genes in the validation set, with similar axis labels and significance markers (Wlicoxon rank sum test, *$P < 0.05$, **$P < 0.01$, ns: $P > 0.05$). **(c)** ROC curve analysis of the VSIG4 biomarker in the validation set. **(d)** ROC curve analysis of the CD163 biomarker in the training set. **(e)** ROC curve analysis of the FPR1 biomarker in the training set. **(f)** ROC curve analysis of the VSIG4 biomarker in the training set. **(g)** ROC curve analysis of the FPR1 biomarker in the validation set. **(h)** ROC curve analysis of the CD163 biomarker in the validation set.

etynodiol may have potential effects on all three biomarkers. Additionally, six compounds were shared between CD163 and FPR1—prednisolone, flunisolide, fludroxycortide, halcinonide, ribavirin, and isoflupredone—while five compounds were shared between FPR1 and VSIG4, including anisomycin, trichostatin A, cephaeline, emetine, and beclometasone (Figure 7c). By understanding the role of these drugs in regulating the expression of immune markers, more effective therapeutic strategies may be developed in the future.

3.8 Single-cell RNA sequencing analysis

Following quality control, 23,963 genes and 49,042 cells were identified (Additional file 6). The top 2000 HVGs were selected, and the 10 genes exhibiting the greatest variation were identified (Additional file 7a). PCA was performed on the selected HVGs, and the top 10 principal components (PCs) were chosen for further analysis ($P < 0.05$) (Additional files 7b-c). UMAP clustering analysis was conducted prior to cell annotation, resulting in the

10.3389/fimmu.2025.1560903

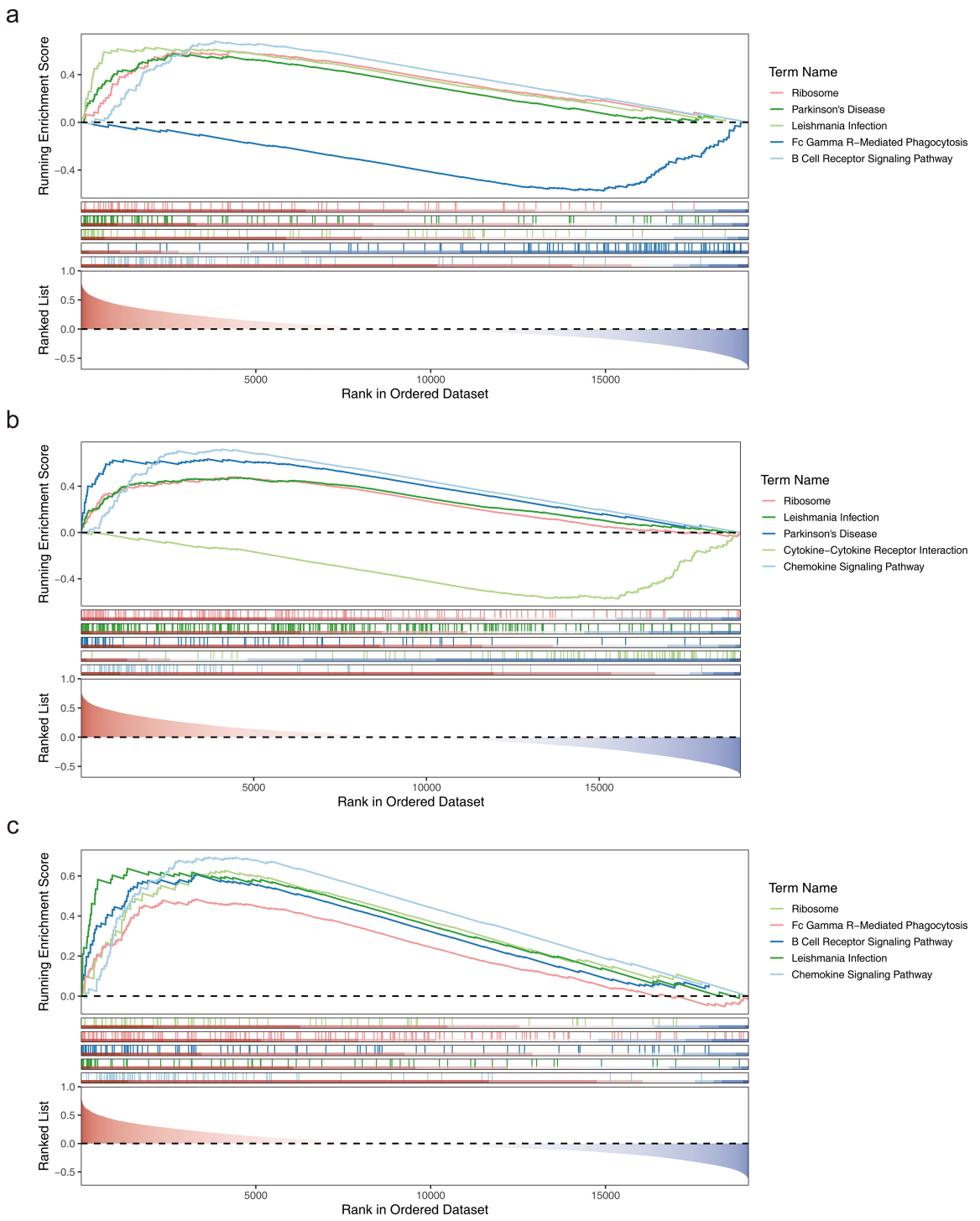


FIGURE 5
Functional analysis of biomarkers. **(a)** GSEA enrichment analysis of the CD163 gene. **(b)** GSEA enrichment analysis of the FPR1 gene. **(c)** GSEA enrichment analysis of the VSIG4 gene.

identification of 14 distinct cell clusters (Additional file 7d). Nine cell types and their corresponding markers were extracted for annotation based on the reference (12). Subsequently, cell annotation revealed eight distinct cell types: endothelium, fibroblasts, pericytes, monocytes and macrophages, natural killer and T lymphocytes (NK&T cells), neurons, B cells, and smooth muscle cells (Figure 8a; Additional file 8). Monocytes and macrophages expressing all three biomarkers were designated as key cells (Figure 8b). To explore the biological pathways and functions of these cell subtypes in HF development, enrichment analysis revealed that pericytes and smooth muscle cells were significantly associated with ATP-sensitive potassium channels

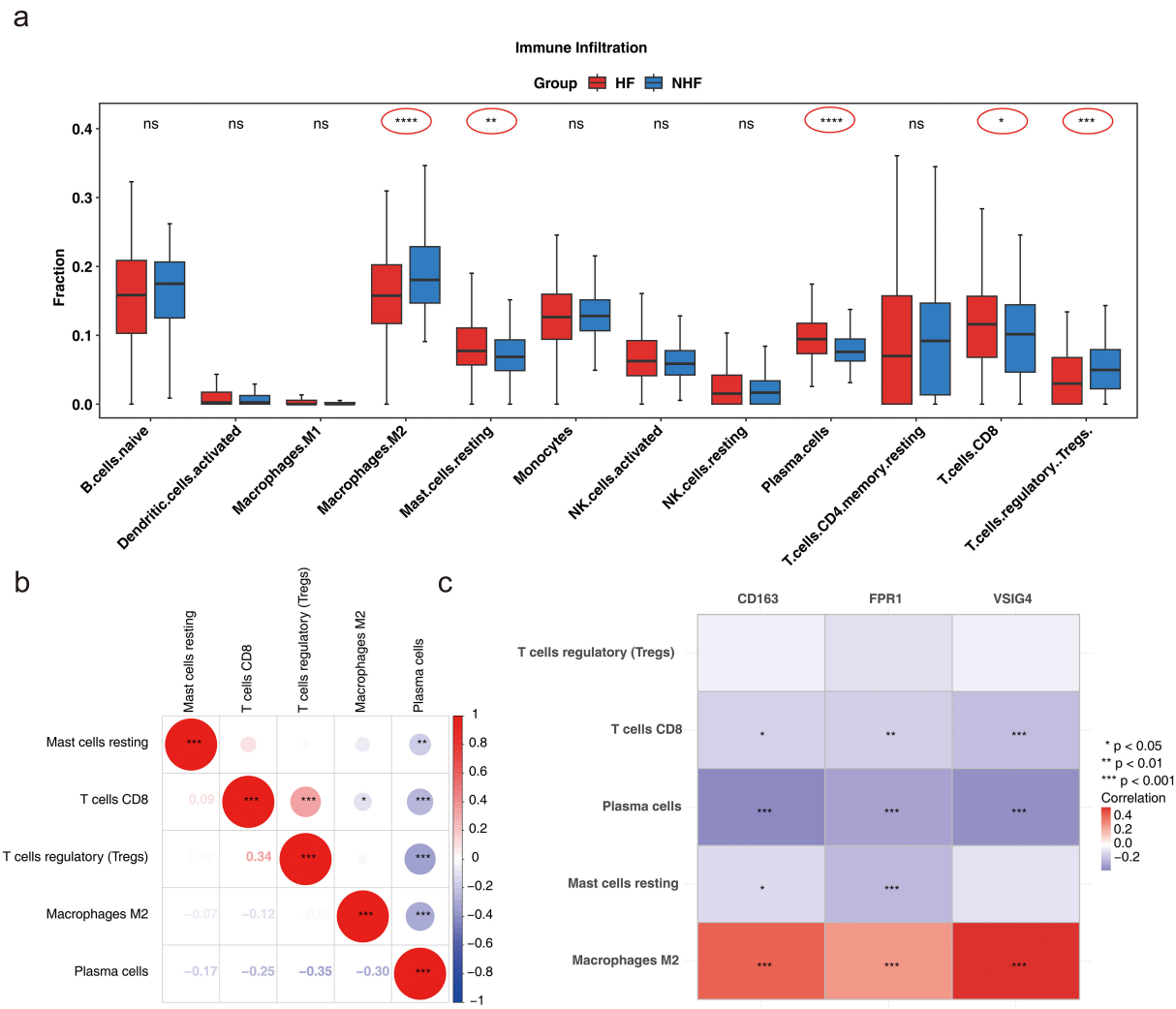


FIGURE 6
Analysis of immune cell infiltration. **(a)** Box plot illustrating immune cell infiltration differences (Wlicoxon rank sum test, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$, ns: $P > 0.05$). **(b)** Correlation of differential immune cell types (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$). **(c)** Correlation between biomarkers and differential immune cells (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

and BDNF activation of NTRK2 (TRKB) signaling, while NK&T cells and B cells were predominantly enriched for activation of Na-permeable kainate receptors and hydroxycarboxylic acid-binding receptors (Figure 8c). The above results implied that these cell types act synergistically through multiple mechanisms and may provide new targets and ideas for the treatment of HF.

Analysis of cell communication between the eight cell types showed that fibroblasts and neurons exhibited the highest number of ligand-receptor pairs, indicating the strongest interaction between these two cell types. Fibroblasts also demonstrated a higher probability of communication with monocytes and macrophages, NK&T cells, and B cells (Figures 9a, b; a: plot of probability of cellular communication, b: plot of number of cellular communications). The high-frequency interaction of fibroblasts with these immune cells suggested that they may play an important role in tissue repair and remodeling in immune responses, and inflammation.

Monocytes and macrophages were projected onto a root with 9 branches, traversing 9 nodes along their developmental trajectory. Clusters 0 and 3 marked the initial stages of monocyte and macrophage development, while clusters 4 and 6 were primarily located at the final stages of cellular differentiation (Figures 9c, d). This dynamic developmental trajectory may reflected how immune cells progressively differentiate and regulate their functions in the body according to different needs.

Given the specific expression of the biomarkers in monocytes and macrophages, the gene expression of the three biomarkers was analyzed across the pseudo-time series. The expression of CD163 showed a decreasing trend over time, with slight increases at certain nodes of the developmental cycle, but overall, the expression in the cells declined. In contrast, FPR1 exhibited an upward trend, indicating its potential significant role in cellular development and differentiation. The expression pattern of VSIG4 mirrored that of CD163 (Figure 9e). This expression pattern suggested that

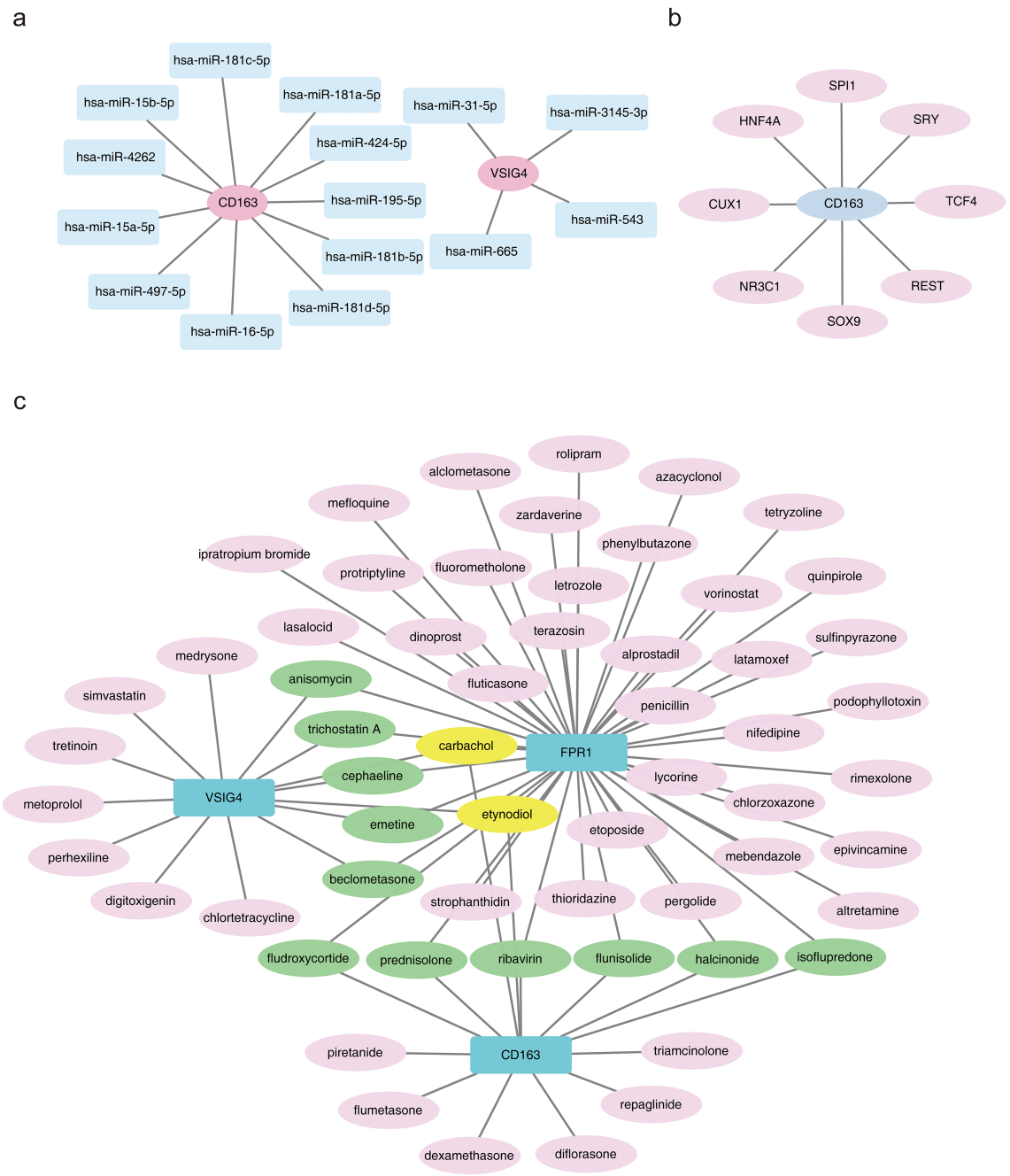


FIGURE 7
Molecular regulatory network and drug prediction. **(a)** Regulatory network between the CD163 gene and miRNAs, where pink nodes represent biomarkers and blue nodes represent interacting miRNAs. **(b)** Regulatory network between the CD163 gene and transcription factors (TFs). **(c)** Diagram of drug prediction for biomarkers.

their immunosuppressive or reparative functions may be gradually replaced by other functions.

To further explore the biomarker expression of monocyte and macrophage subpopulations at different stages of differentiation, 13 cells were first clustered and annotated into 5 subpopulations based on marker genes (Table 2; Additional file 9a-c). Subsequently, the five cell subpopulations were analyzed in a proposed time series. As shown in Additional file 9d, cells gradually differentiated over time, with darker blue representing earlier differentiation. Each cell subpopulation mapped to a different differentiation time and corresponded to a different differentiation state, with darker red indicating the earliest type of differentiation. As cells differentiated, the expression of CD163 and VSIG4 in key cell subpopulations gradually increased, while the expression of FPR1 slowly decreased (Additional file 9e). Next, stacked plots of cell subpopulations in different differentiation states (Additional file 9f) showed that M1 macrophages were distributed in all differentiation states, especially more in state 2 and state 5; Intermediate monocytes were

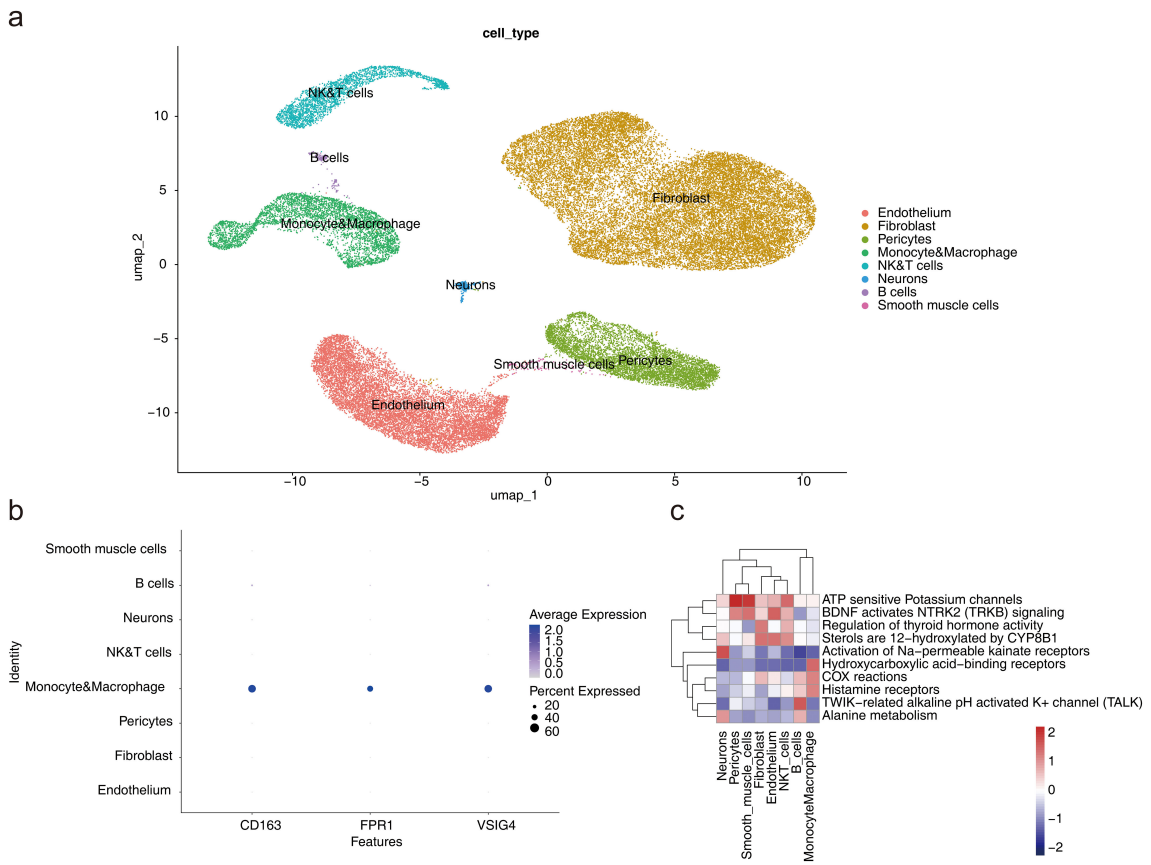


FIGURE 8
Single-cell RNA sequencing analysis. **(a)** UMAP plot for different cell types. **(b)** Expression profile plot of biomarkers. **(c)** Pathway enrichment analysis of cell subtypes.

distributed only in state 1; and Non-classical monocytes were mainly distributed in state 1 and state 3; M2 macrophages were concentrated in state 3 and state 4 in the later stages of differentiation; Classical monocytes were found mainly in state 1 and state 5. Subsequently, the proportions of cell subtypes under different groupings were visualized (Additional file 10a). By comparing NOS2, TNF, ARG1 and MRC1 gene expression in Monocyte&Macrophage between HF and control samples, TNF and MRC1 were found to be significantly different between the two groups (Additional files 10b-e). This provided important clues to a deeper understanding of the function of monocytes and macrophages and their role in disease.

3.9 Clinical and animal validation of Hub genes

To validate the expression levels of ICD-related hub genes in HF, PBMCs were extracted from 15 clinical patients with HF and controls for RT-qPCR analysis. Results revealed significant down-regulation of CD163, FPR1, and VSIG4 in patients with HF (Figure 10a). Further investigation was conducted in heart tissues using the HF rat model. Echocardiography showed reduced left ventricular ejection fraction (LVEF) and left ventricular fractional shortening index (LVFS), alongside increased left ventricular end-systolic diameter (LVIDs) and left ventricular end-diastolic diameter (LVIDd) in HF rats (Figures 10b-c). The ratios of heart weight to body weight and lung weight to tibia length were significantly elevated (Figure 10d). HE staining revealed prominent cardiomyocyte hypertrophy, with inflammatory cell infiltration in the HF group (Figure 10e). Masson staining indicated severe fibrosis in the HF group (Figure 10f), and the difference in fibrosis between the two groups was significant (Figure 10g). Cardiac tissue RT-qPCR results confirmed that CD163, FPR1, and VSIG4 were significantly down-regulated in HF rats (Figure 10h). The results of immunohistochemistry showed that the expression of $CD163^{+}$ cells was decreased in the myocardial tissue of HF mice (Additional file 11). These results suggested that down-regulation of CD163, FPR1, and VSIG4 expression in HF patients and HF rat models may be closely associated with dysregulation of the immune system, decreased cardiac function, and tissue damage.

4 Discussion

Cardiac immunology has recently emerged as a focal area of research. While some aspects of immune regulation in HF are

10.3389/fimmu.2025.1560903

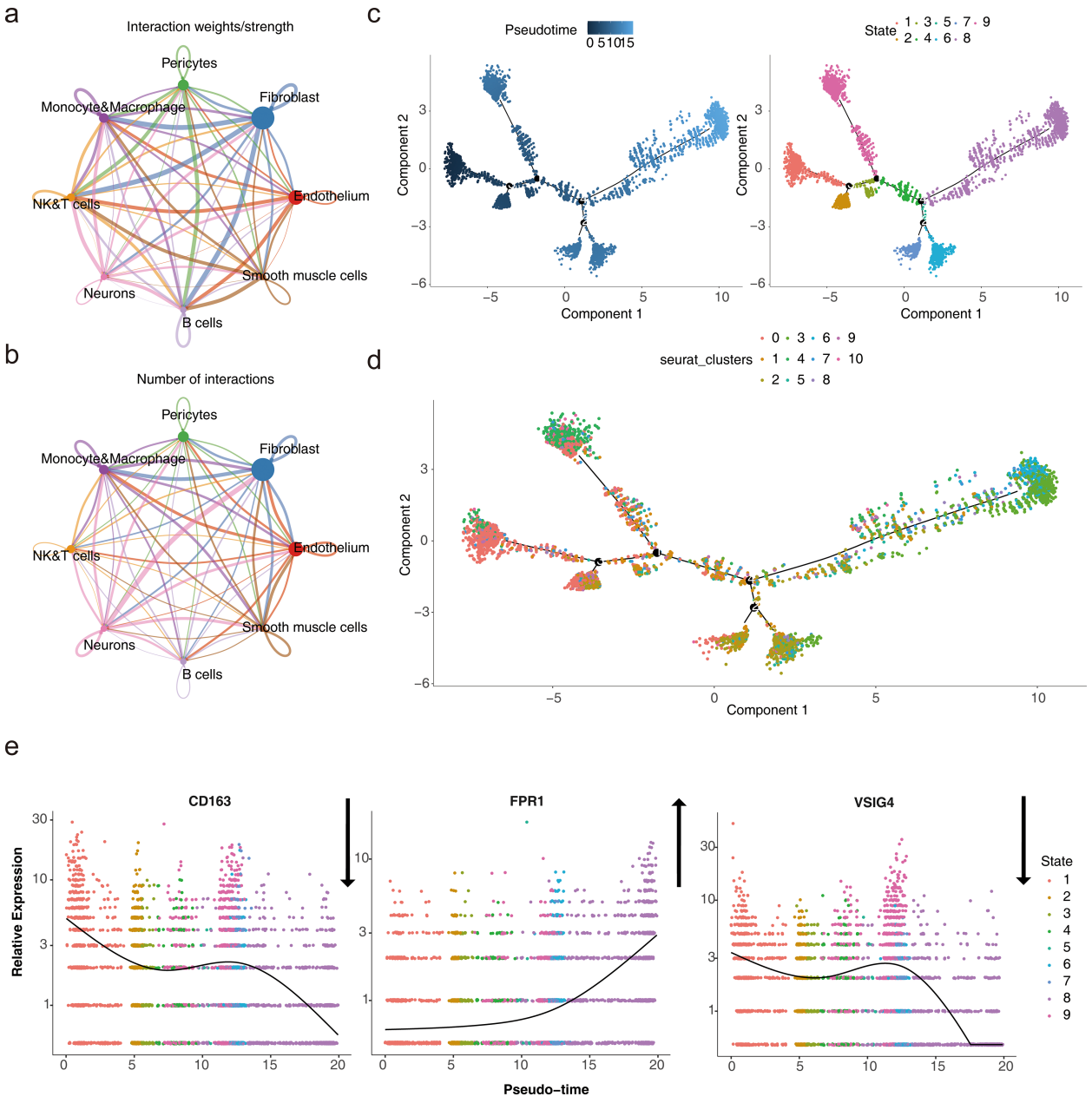


FIGURE 9
Cell subtype communication analysis. **(a-b)** Diagrams of cell communication results. **(c)** Results of pseudotime and state analysis in single-cell pseudotemporal analysis. Darker colors indicate the most advanced state of development, while lighter colors indicate more mature development. The cells were divided into 9 different periods according to their developmental state. **(d)** Cell pseudotime analysis of Seurat clusters. Different cell clusters presented different positions at various nodes of the developmental trajectory. **(e)** Dynamic atlas of biomarkers.

understood, many questions remain to be addressed. ICD is a form of programmed cell death induced by antigens and adjuvants, triggering downstream immune responses. However, the role and mechanisms of ICD in HF pathophysiology remain unclear. In this study, three ICD-related biomarkers—CD163, FPR1, and VSIG4—were identified in patients with HF using transcriptomic and single-cell dataset analyses (Additional file 12). Previous studies have shown that these three genes, as combined markers, may act synergistically to affect the occurrence and development of HF and non-alcoholic fatty liver disease by regulating mechanisms such as immune response and monocyte migration. In addition, their association with natural killer (NK) cells and macrophages was also found, further supporting their important role in the immune response (39).

Single-cell sequencing data in this study were obtained from the research by Koenig et al. (12). Unlike the study by Koenig, our work systematically integrated multiomics analyses (including transcriptomes and single-cell sequencing), machine-learning approaches (e.g., SVM-RFE and random forests), and immune infiltration assessments, which were not comprehensively

10.3389/fimmu.2025.1560903

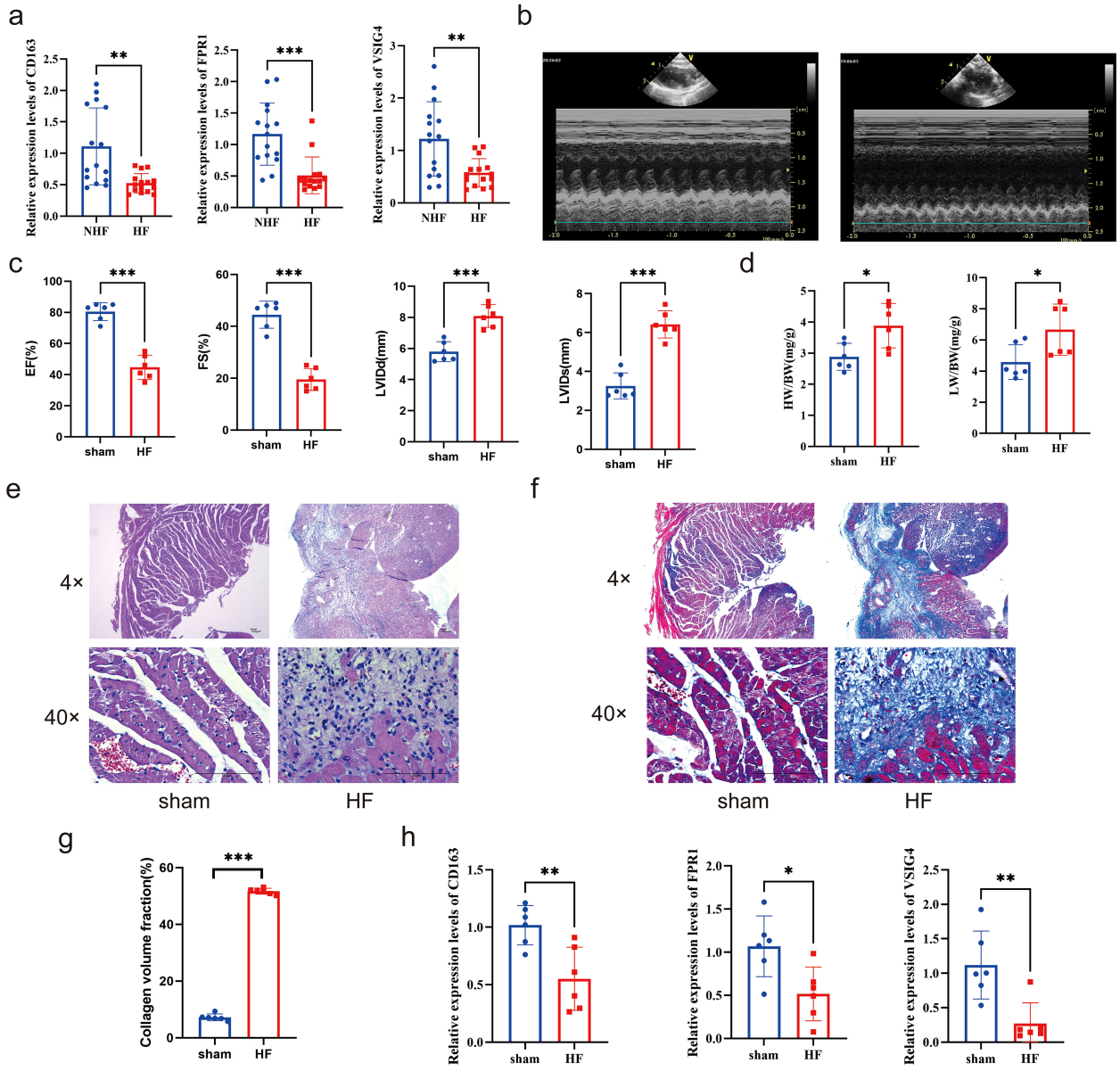


FIGURE 10
Clinical and animal validation of hub genes. **(a)** Expression of CD163, FPR1, and VSIG4 in peripheral blood mononuclear cells of patients with HF and NHF individuals (Unpaired t test, $^{**}P < 0.01$, $^{***}P < 0.001$). **(b-c)** Echocardiograms of the HF rat model and sham group (Unpaired t test, $^{***}P < 0.001$). **(d)** The ratios of heart weight to body weight and lung weight to tibia length in rat model (Unpaired t test, $^{*}P < 0.05$). **(e-f)** HE and Masson staining of rat hearts. **(g)** Collagen volume fraction(%) calculated by Masson staining (Unpaired t test, $^{***}P < 0.001$). **(h)** Expression of CD163, FPR1, and VSIG4 in the hearts of HF and sham groups (Unpaired t test, $^{*}P < 0.05$, $^{**}P < 0.01$).

TABLE 2 Marker gene annotation information for key cell subpopulations.

TNF	M1 macrophage
MERTK, CD163, STAB1, MRC1	M2 macrophage
BASP1, CXCL8, GPR183	Classical monocyte
FCN1	Non-classical monocyte
FCGR3A	Intermediate monocyte

combined in their study. Furthermore, this study identified the association of CD163, FPR1, and VSIG4 with ICD, a connection that Koenig et al. did not investigate. Specifically, during ICD, certain molecules, especially ANXA1, may enhance local inflammatory responses by binding to FPR1 receptors and activating macrophages and monocytes. At the same time, the activation of fibroblasts may promote vascular wall structural changes and fibrosis (40). Therefore, targeting FPR1 or ICD-related pathways may be a potential strategy for the treatment of ascending aortic aneurysm. When tumor cells develop ICD through

radiotherapy or other therapeutic modalities, macrophages recognize tumor cell death signals through CD163 receptors. $CD163^+$ macrophages are normally in an immunosuppressive state and help tumors evade immune surveillance by promoting Treg cell infiltration and inhibiting effector T cell function (41). In addition, carbon ion radiotherapy has been shown to effectively reduce fiber deposition in scar tissue by inducing ICD of fibroblasts, slowing their proliferation and promoting their death (42). Another study pointed out that ICD may affect cancer-associated fibroblasts by regulating immune responses, thereby altering tumor progression and patient survival prognosis. Although no association between ICD and macrophages or fibroblasts has been found in HF, these immune cells may affect the occurrence and development of HF through the ICD process. Additionally, potential therapeutic targets were proposed via drug prediction, such as carbachol and etynodiol, which target all three biomarkers. Collectively, this study not only extends the findings of Koenig et al. but also offers novel insights and references for future research in HF.

CD163 (Cluster of Differentiation 163), a 130 kDa cell surface glycoprotein, is predominantly expressed on monocytes and macrophages. It plays significant roles in metabolic diseases and immune regulation and is considered a promising target for drug development (43, 44). Soluble CD163 (sCD163) is a soluble inflammatory mediator produced through the enzymatic hydrolysis of CD163 (45). CD163 expression tends to be low in conditions such as non-alcoholic fatty liver (39, 46) and ischemic cardiomyopathy (47), whereas sCD163 tends to be elevated in hypertension (48) and diabetes (49, 50). Additionally, sCD163 has been linked to increased cardiovascular mortality in diabetic patients. In HF, CD163 expression is down-regulated in cardiac tissues (39, 51), consistent with both bioinformatics and experimental findings in this study. CD163 expression in cardiac tissue is also associated with hyperlipidemia (52) and cellular stemness (51). Moreover, sCD163 is highly expressed in the blood of patients with HF (53), though the mechanisms driving this increase remain under investigation. Some studies suggest that sCD163 levels are influenced by left ventricular diastolic volume (53), while others have linked sCD163 to monocyte activation, particularly activation related to the M2 phenotype (54), which warrants further exploration. In addition, research has demonstrated that CD163 serves as a critical link between the immune system, inflammatory response, and cardiovascular disease by not only reflecting the activation of immune cells, particularly macrophages, but also modulating immune responses (54). Furthermore, in another study, CD163, acting as a macrophage marker, was found to play a significant role in regulating inflammation and the tumor microenvironment (44). This study also found a reduction in CD163 expression in macrophages in HF, suggesting that the progression of HF may be linked to decreased CD163 expression in macrophages.

FPR1 (Formyl Peptide Receptor 1), a key member of the G protein-coupled receptor family, plays a critical role in the inflammatory process and immune cell recruitment. It is highly expressed in macrophages (55) and mediates macrophage chemotactic motility and functional activation by binding to TAFA4 (56, 57). FPR1 is a well-established target for clinical therapeutic drugs, with various agonists and inhibitors developed for its modulation (58). Studies have demonstrated that FPR1 modulates the immune response and repair process of the heart by regulating macrophage activity, and dysregulation of the immune response following cardiac injury may contribute to the development of HF (59). Moreover, FPR1 may mitigate inflammatory responses and facilitate cardiac repair and recovery in HF through the regulation of macrophage function (60). Studies suggest that FPR1 may be a promising drug target for cardiovascular diseases, aiding both diagnosis and treatment (61). It plays a negative regulatory role in myocardial ischemia-reperfusion and coronary atherosclerosis but a positive regulatory role in myocardial infarction. FPR1 contributes to atherosclerotic lesions by modulating the number of blood neutrophils under hypercholesterolemia (62) and exacerbates myocardial cell apoptosis and inflammation during ischemia-reperfusion through the MAPK signaling pathway (63). However, FPR1 activation has been shown to improve left ventricular remodeling after myocardial infarction in mice and rats, potentially by promoting early neutrophil migration and infiltration, thus accelerating wound healing (64). In the present study, decreased expression of FPR1 was observed in PBMCs from patients with HF and in the hearts of HF rats through both bioinformatics and experimental validation. However, no significant difference in FPR1 expression was found in macrophages in HF. Notably, FPR1 expression gradually increased during macrophage differentiation, suggesting its potential as a therapeutic target for HF.

VSIG4 (V-set and immunoglobulin domain containing 4) is a type I transmembrane receptor that inhibits T cell activation and induces the differentiation of regulatory T cells, thus suppressing immune-mediated inflammatory diseases (65). Soluble VSIG4, shed from the surface of macrophages, serves as a biomarker for diseases associated with macrophage activation (66). VSIG4 has a protective role in cardiovascular diseases and can alleviate age-related insulin resistance and hypertension (67). Additionally, research has highlighted that VSIG4, as a critical immune marker, is strongly associated with macrophage function and plays a pivotal role in both the immune response and the diagnosis of HF (39, 68). In myocardial ischemia/reperfusion (I/R) injury, VSIG4 inhibits M1 macrophage polarization by blocking TLR4/NF-κB signaling, thus preventing cardiomyocyte apoptosis (69). However, VSIG4 expression in M2 macrophages promotes fibrosis after acute myocardial infarction, suggesting its potential as an immunomodulatory therapeutic target (70). In HF, VSIG4 expression is significantly down-regulated in patients with right ventricular HF (71), while serum levels of VSIG4 are elevated in patients with left ventricular HF, with high levels correlating with poor prognosis (72). In the present study, VSIG4 expression was decreased in macrophages in HF, and its expression showed a decreasing trend during macrophage differentiation, further suggesting that HF progression may be linked to the expression of VSIG4 in macrophages.

GSEA enrichment analysis reveals that the three biomarkers are significantly enriched in ribosomes. The enhanced translation function of ribosomes is a hallmark of cardiac hypertrophy, and inhibiting ribosomal translation can effectively mitigate hypertrophy (73). However, systemic inhibition of ribosomal translation may cause adverse effects in organs outside the heart. For example, while rapamycin effectively inhibits cardiac hypertrophy, it can lead to severe consequences such as immune suppression (74). Recent studies have identified the cardiac-specific nuclear ribonucleoprotein (RNP)-binding long non-coding RNA (lncRNA) CARDINAL, which alleviates cardiac hypertrophy *in vivo* and *in vitro* by inhibiting the translation of hypertrophy-related proteins (75). In the study by Koji Kasahara et al. (76), FPR1 indirectly influenced ribosomal function through the regulation of ribosomal protein gene expression. Additionally, VSIG4, an up-regulated gene, is linked to ribosome function, implying its potential significance in protein synthesis or cellular function regulation (77). Prior research has demonstrated that CD163 expression correlates with the mTOR signaling pathway (78), which governs translation initiation and ribosome biogenesis (79). The biomarkers identified in this study are all associated with ribosomes, offering a new avenue for basic research. Single-cell analysis highlights the pivotal role of monocytes and macrophages in HF progression, with cardiac macrophages regulating both survival and adaptive remodeling in patients with HF. However, these macrophages are highly infiltrated in the hearts of patients with HF, potentially due to the elevated expression of Ang II, which mobilizes macrophages (80). Macrophages are categorized into M1 and M2 types based on their secreted factors and functions. Promoting the conversion of M1 to M2 macrophages and maintaining a balance between these two subtypes may provide an effective strategy for treating HF (81). It has been demonstrated that sodium-glucose cotransporter 2 (SGLT2) inhibitors can reduce fibrosis markers by promoting M2 macrophage polarization and enhancing angiogenic factors (82), while nicorandil can suppress the production of pro-inflammatory cytokines by inhibiting M1 polarization (83). Furthermore, this study found a positive correlation between the expression levels of these three biomarkers and M2 macrophages, suggesting that targeting these biomarkers to modulate macrophage homeostasis in HF may offer a promising therapeutic strategy.

Cell subtype communication analysis revealed that fibroblasts likely engage in frequent interactions with monocytes, macrophages, NK cells, T cells, and B cells. Previous studies have demonstrated that macrophages influence cardiac function by modulating fibroblast activity and affecting the remodeling and excessive deposition of extracellular matrix (ECM) (84). During cardiac inflammation and remodeling, macrophages and fibroblasts exhibit a close interconnection. Notably, M1 macrophages release pro-inflammatory cytokines, activate fibroblasts, and drive the progression of fibrosis (85). Additionally, research has shown that macrophages interact with TWEAK via the receptor CD163, playing a critical role in cardiac fibrosis and HF (86). VSIG4 promotes cardiac fibrosis repair during acute myocardial infarction (AMI) by regulating M2-type macrophage function and interacting with immune factors such as TGF-β1 and IL-10 (70). Furthermore, the FPR1 receptor is crucial for the aggregation and activation of immune cells, including monocytes and macrophages, which subsequently impacts fibroblast activation and fibrosis, thereby promoting inflammatory and fibrotic responses in the heart and lung (87). Collectively, the intricate crosstalk between immune cells and fibroblasts plays a pivotal role in the pathogenesis of cardiac inflammation and fibrosis, offering potential therapeutic targets and novel strategies for treating cardiac fibrosis.

In this paper, drug prediction was performed based on three biomarkers, and it was found that carbachol and etynodiol may have potential roles for all three biomarkers. Carbachol, a structural analogue of acetylcholine that acts on muscarinic and nicotinic receptors, is used clinically to treat glaucoma (88). Only a few literatures have found that carbachol increases phagocytosis of macrophages *in vitro* (89). Progestin is the first progestin with moderate progestogen activity, and progestin has some effect on macrophages. In a clinical study of adolescent endometriosis, one-year progestin treatment increased the number of $CD206^+$ monocytes ($P < 0.001$) but decreased the number of $CD163^+$ monocytes ($P = 0.017$) (90). The specific effects of the above two drugs on macrophages are still superficial, and the relevant mechanisms are not deeply studied. In addition, the effects of the above two drugs on heart failure are lack of relevant research support and still need to be further explored.

This study has several limitations. First, the dataset is relatively small, necessitating the inclusion of larger, multi-center datasets (e.g. UK Biobank, HF registry study data) for more robust conclusions. Furthermore, validation in human and animal models is preliminary; additional functional experiments, such as gene knockout or overexpression studies, are needed to clarify the roles of these biomarkers in HF progression. Simultaneously, further experimental evidence is required to clarify the relationship between biomarkers and ribosomes. Moreover, existing studies have predominantly focused on monocytes/macrophages, while the interactions with other cell types, such as fibroblasts and cardiomyocytes, remain underexplored. Future investigations could leverage spatial transcriptome technologies, like Visium, to map co-localization regions and deepen our understanding of macrophage-fibroblast interactions. Lastly, the absence of experimental validation for drug predictions restricts their direct clinical application. In subsequent studies, carbachol or etynodiol could be administered in HF rat models to monitor changes in CD163/VSIG4 expression levels, cardiac function parameters, and inflammatory/fibrosis markers. Despite these limitations, the study identifies novel mechanisms underlying HF and highlights potential biomarkers, offering valuable insights for the prevention and treatment of HF and establishing a foundation for future research.

5 Conclusions

This study identified three biomarkers—CD163, FPR1, and VSIG4—associated with immunogenic cell death in patients with HF, integrating transcriptomic data with single-cell datasets. The

functions and biological pathways of these biomarkers were examined, and the potential links between immunogenic cell death-related genes and HF pathophysiology were explored. Additionally, the expression of these biomarkers was validated in both human and animal models, providing a novel theoretical framework for clinical diagnosis and treatment of HF.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Scientific Research Ethics Review Committee of the First Hospital of Shanxi Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. The animal study was approved by the Animal Experimental Center Ethics Committee of Beijing Yongxinkangtai Science and Technology Development. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

HW: Formal Analysis, Investigation, Writing – original draft. DW: Resources, Writing – original draft. GH: Resources, Writing – original draft. JY: Methodology, Writing – original draft. ZW: Data curation, Writing – original draft. XH: Data curation, Writing – original draft. YC: Data curation, Writing – original draft. YW: Funding acquisition, Supervision, Writing – review & editing. QH: Funding acquisition, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported by the National Nature Science Foundation of China (grant numbers 81970204, 32200717) and Key Research and Development Program of Shanxi Province (grant numbers 2022ZDYF089). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

We would like to express our sincere gratitude to all individuals and organizations who supported and assisted us throughout this research.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2025.1560903/full#supplementary-material

ADDITIONAL FILE 1
Thirty-four ICD-related genes.

ADDITIONAL FILE 2
Analysis of differential gene expression and sample clustering in HF. **(a)** Volcano plot illustrating differentially expressed genes. **(b)** Heatmap displaying the top 10 differentially expressed genes. **(c)** Comparison of ssGSEA scores between heart failure and normal samples. The horizontal axis represents the ssGSEA score, and the vertical axis represents sample groupings (Wlicoxon rank sum test, ** $P < 0.01$). **(d)** Hierarchical clustering of samples. Each branch in the clustering tree corresponds to a sample, with the vertical coordinate representing the Euclidean distance of sample expression levels. **(e)** Soft-thresholding analysis for network construction.

ADDITIONAL FILE 3
Expression analysis of biomarkers in GSE5406 dataset.

ADDITIONAL FILE 4
ROC curves of single-gene and nomogram models.

ADDITIONAL FILE 5
Stacked column chart depicting differences in immune cell infiltration.

ADDITIONAL FILE 6
Quality control results of single-cell sequencing.

ADDITIONAL FILE 7
Cellular heterogeneity and dimensionality reduction analysis. **(a)** Diagram for screening highly variable genes. **(b-c)** Principal component analysis (PCA) plots of cells. **(d)** UMAP plot for different cell clusters.

ADDITIONAL FILE 8
Bubble chart of cell marker genes.

ADDITIONAL FILE 9
Characterization of Monocyte and Macrophage subpopulations and their differentiation dynamics. **(a-b)** Expression levels of Marker genes in different

cell populations. **(c)** Annotated cellular map of monocyte and macrophage subpopulations. **(d)** Proposed temporal analysis plot of key cell subpopulations. **(e)** Changes in biomarker expression during cell differentiation. **(f)** The stacked plots of cell subpopulations in different differentiation states.

ADDITIONAL FILE 10

TNF and MRC1 were significantly different between HF and control samples. **(a)** Visualization of the proportions of cell subtypes under different groupings. **(b-e)** Expression of NOS2, TNF, ARG1, and MRC1 genes in Monocyte&Macrophage between HF and control samples.

ADDITIONAL FILE 11

Expression of $CD163^{+}$macrophages in rat myocardium.

ADDITIONAL FILE 12

The flowchart of this research.

References

1. Khan MS, Shahid I, Bennis A, Rakisheva A, Metra M, Butler. J. Global epidemiology of heart failure. *Nat Rev Cardiol*. (2024) 21:717–34. doi: 10.1038/s41569-024-01046-6

2. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats. AJS. Global burden of heart failure: a comprehensive and updated review of epidemiology. *Cardiovasc Res*. (2023) 118:3272–87. doi: 10.1093/cvr/cvac013

3. Ostrominski JW, DeFilippis EM, Bansal K, Riello RJ, Bozkurt 3B, Heidenreich PA, et al. Contemporary american and european guidelines for heart failure management: JACC: heart failure guideline comparison. *JACC Heart Fail*. (2024) 12:810–25. doi: 10.1016/j.jchf.2024.02.020

4. Markousis-Mavrogenis G, Baumhove L, Al-Mubarak AA, Aboumsallem JP, Bomer N, Voors AA, et al. Immunomodulation and immunopharmacology in heart failure. *Nat Rev Cardiol*. (2024) 21:119–49. doi: 10.1038/s41569-023-00919-6

5. Meier P, Legrand AJ, Adam D, Silke. J. Immunogenic cell death in cancer: targeting necroptosis to induce antitumour immunity. *Nat Rev Cancer*. (2024) 24:299–315. doi: 10.1038/s41568-024-00674-x

6. Arimoto KI, Miyauchi S, Liu M, Zhang. DE. Emerging role of immunogenic cell death in cancer immunotherapy. *Front Immunol*. (2024) 15:1390263. doi: 10.3389/fimmu.2024.1390263

7. Tian Z, Li X, Jiang. D. Analysis of immunogenic cell death in atherosclerosis based on scRNA-seq and bulk RNA-seq data. *Int Immunopharmacol*. (2023) 119:110130. doi: 10.1016/j.intimp.2023.110130

8. Zhang YJ, Huang C, Zu XG, Liu JM, Li. YJ. Use of machine learning for the identification and validation of immunogenic cell death biomarkers and immunophenotypes in coronary artery disease. *J Inflammation Res*. (2024) 17:223–49. doi: 10.2147/jir.S439315

9. Qin L, Li S, Cao X, Huang T, Liu Y, Chen. O. Potential diagnostic biomarkers for immunogenic cell death in elderly female patients with ischemic stroke: identification and analysis. *Sci Rep*. (2024) 14:14553. doi: 10.1038/s41598-024-65390-w

10. Liu Y, Morley M, Brandimarto J, Hannenhalli S, Hu Y, Ashley EA, et al. RNA-Seq identifies novel myocardial gene expression signatures of heart failure. *Genomics*. (2015) 105:83–9. doi: 10.1016/j.ygeno.2014.12.002

11. Barth AS, Kuner R, Buness A, Ruschhaupt M, Merk S, Zwermann L, et al. Identification of a common gene expression signature in dilated cardiomyopathy across independent microarray studies. *J Am Coll Cardiol*. (2006) 48:1610–7. doi: 10.1016/j.jacc.2006.07.026

12. Koenig AL, Shchukina I, Amrute J, Andhey PS, Zaitsev K, Lai L, et al. Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure. *Nat Cardiovasc Res*. (2022) 1:263–80. doi: 10.1038/s44161-022-00028-6

13. Garg AD, De Ruysscher D, Agostinis. P. Immunological metagene signatures derived from immunogenic cancer cell death associate with improved survival of patients with lung, breast or ovarian Malignancies: A large-scale meta-analysis. *Oncoimmunology*. (2016) 5:e1069938. doi: 10.1080/2162402x.2015.1069938

14. Zhang L, Zhang X, Liu H, Yang C, Yu J, Zhao W, et al. MTFR2-dependent mitochondrial fission promotes HCC progression. *J Transl Med*. (2024) 22:73. doi: 10.1186/s12967-023-04845-6

15. Gustavsson EK, Zhang D, Reynolds RH, Garcia-Ruiz S, Ryten. M. ggtranscript: an R package for the visualization and interpretation of transcript isoforms using ggplot2. *Bioinformatics*. (2022) 38:3844–6. doi: 10.1093/bioinformatics/btac409

16. Gu Z, Eils R, Schlesner. M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. (2016) 32:2847–9. doi: 10.1093/bioinformatics/btw313

17. Hänzelmann S, Castelo R, Guinney. J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf*. (2013) 14:7. doi: 10.1186/1471-2105-14-7

18. Langfelder P, Horvath. S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf*. (2008) 9:559. doi: 10.1186/1471-2105-9-559

19. Mao W, Ding J, Li Y, Huang R, Wang. B. Inhibition of cell survival and invasion by Tanshinone IIA via FTH1: A key therapeutic target and biomarker in head and neck squamous cell carcinoma. *Exp Ther Med*. (2022) 24:521. doi: 10.3892/etm.2022.11449

20. Wang L, Wang D, Yang L, Zeng X, Zhang Q, Liu G, et al. Cuproptosis related genes associated with Jab1 shapes tumor microenvironment and pharmacological profile in nasopharyngeal carcinoma. *Front Immunol*. (2022) 13:989286. doi: 10.3389/fimmu.2022.989286

21. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*. (2021) 2:100141. doi: 10.1016/j.xinn.2021.100141

22. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. (2003) 13:2498–504. doi: 10.1101/gr.1239303

23. Conway JR, Lex A, Gehlenborg. N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*. (2017) 33:2938–40. doi: 10.1093/bioinformatics/btx364

24. Yang L, Pan X, Zhang Y, Zhao D, Wang L, Yuan G, et al. Bioinformatics analysis to screen for genes related to myocardial infarction. *Front Genet*. (2022) 13:990888. doi: 10.3389/fgene.2022.990888

25. Zhao P, Zhen H, Zhao H, Huang Y, Cao. B. Identification of hub genes and potential molecular mechanisms related to radiotherapy sensitivity in rectal cancer based on multiple datasets. *J Transl Med*. (2023) 21:176. doi: 10.1186/s12967-023-04029-2

26. Liu Z, Wang L, Xing Q, Liu X, Hu Y, Li W, et al. Identification of GLS as a cuproptosis-related diagnosis gene in acute myocardial infarction. *Front Cardiovasc Med*. (2022) 9:1016081. doi: 10.3389/fcvm.2022.1016081

27. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf*. (2011) 12:77. doi: 10.1186/1471-2105-12-77

28. Xu J, Yang T, Wu F, Chen T, Wang A, Hou. S. A nomogram for predicting prognosis of patients with cervical cerclage. *Heliyon*. (2023) 9:e21147. doi: 10.1016/j.heliyon.2023.e21147

29. Unger-Plasek B, Temesi Á, Lakner. Z. Towards understanding the motivators of sustainable consumer behavior-validation of the food eco-guilt scale. *Nutrients*. (2024) 16:(21). doi: 10.3390/nu16213695

30. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh. AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol*. (2018) 1711:243–59. doi: 10.1007/978-1-4939-7493-1_12

31. Chen F, Yang Y, Zhao Y, Pei L, Yan. H. Immune infiltration profiling in nonsmall cell lung cancer and their clinical significance: study based on gene expression measurements. *DNA Cell Biol*. (2019) 38:1387–401. doi: 10.1089/dna.2019.4899

32. Gu Z, Hübschmann. D. Make interactive complex heatmaps in R. *Bioinformatics*. (2022) 38:1460–2. doi: 10.1093/bioinformatics/btab806

33. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng 3S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell*. (2021) 184:3573–3587.e29. doi: 10.1016/j.cell.2021.04.048

34. Griss J, Viteri G, Sidiropoulos K, Nguyen V, Fabregat A, Hermjakob. H. ReactomeGSA - efficient multi-omics comparative pathway analysis. *Mol Cell Proteom*. (2020) 19:2115–25. doi: 10.1074/mcp.TIR120.002155

35. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. (2021) 12:1088. doi: 10.1038/s41467-021-21246-9

36. Qiu X, Hill A, Packer J, Lin D, Ma YA, Trapnell. C. Single-cell mRNA quantification and differential analysis with Census. *Nat Methods*. (2017) 14:309–15. doi: 10.1038/nmeth.4150

37. Jiang Y, Yu W, Hu T, Peng H, Hu F, Yuan Y, et al. Unveiling macrophage diversity in myocardial ischemia-reperfusion injury: identification of a distinct lipid-associated macrophage subset. *Front Immunol*. (2024) 15:1335333. doi: 10.3389/fimmu.2024.1335333

38. Hu W, Tu H, Wadman MC, Li YL, Zhang. D. Renal denervation achieves its antiarrhythmic effect through attenuating macrophage activation and neuroinflammation in stellate ganglia in chronic heart failure. *Cardiovasc Res*. (2024) 120(18):2420–33. doi: 10.1093/cvr/cvae196

39. Zhang Y, Feng L, Guan X, Zhu Z, He Y, Li. X. Non-alcoholic fatty liver disease and heart failure: A comprehensive bioinformatics and Mendelian randomization analysis. *ESC Heart Fail*. (2024) 11:4185–200. doi: 10.1002/ehf2.15019

40. Tian Z, Zhang P, Li X, Jiang. D. Analysis of immunogenic cell death in ascending thoracic aortic aneurysms based on single-cell sequencing data. *Front Immunol*. (2023) 14:1087978. doi: 10.3389/fimmu.2023.1087978

41. Lip H, Zetrini A, Park E, Cai P, Abbasi AZ, Huyan T, et al. Mitigating radioresistance mechanisms by polymer-lipid manganese dioxide nanoparticles enhances immunogenic cell death and antitumor immune response to facilitate abscopal effect in breast tumor models. *Drug Delivery Transl Res*. (2025). doi: 10.1007/s13346-025-01873-1

42. Zhou H, Yang P, Zhang T, Kepp O, Ren Y, Jiang N, et al. The role of apoptosis, immunogenic cell death, and macrophage polarization in carbon ion radiotherapy for keloids: Targeting the TGF-β1/SMADs signaling pathway. *Biochim Biophys Acta Mol Basis Dis*. (2025) 1871:167499. doi: 10.1016/j.bbadis.2024.167499

43. Ratajczak W, Atkinson SD, Kelly. C. The TWEAK/Fn14/CD163 axis-implications for metabolic disease. *Rev Endocr Metab Disord*. (2022) 23:449–62. doi: 10.1007/s11154-021-09688-4

44. Skytthe MK, Graversen JH, Moestrup. SK. Targeting of CD163(+) macrophages in inflammatory and Malignant diseases. *Int J Mol Sci*. (2020) 21:(15). doi: 10.3390/ijms21155497

45. Plevriti A, Lamprou M, Mourkogianni E, Skoulas N, Giannakopoulou M, Sajib MS, et al. The role of soluble CD163 (sCD163) in human physiology and pathophysiology. *Cells*. (2024) 13:(20). doi: 10.3390/cells13201679

46. Li G, Lu Z, Chen Z. Identification of common signature genes and pathways underlying the pathogenesis association between nonalcoholic fatty liver disease and heart failure. *Front Immunol*. (2024) 15:1424308. doi: 10.3389/fimmu.2024.1424308

47. Wang J, Xie S, Cheng Y, Li X, Chen J, Zhu M. Identification of potential biomarkers of inflammation-related genes for ischemic cardiomyopathy. *Front Cardiovasc Med*. (2022) 9:972274. doi: 10.3389/fcvm.2022.972274

48. Al-Daghri NM, Al-Attas OS, Bindahman LS, Alokail MS, Alkharfy KM, Draz HM, et al. Soluble CD163 is associated with body mass index and blood pressure in hypertensive obese Saudi patients. *Eur J Clin Invest*. (2012) 42(11):1221–6. doi: 10.1111/j.1365-2362.2012.02714.x

49. Siwan E, Twigg SM, Min D. Alterations of CD163 expression in the complications of diabetes: A systematic review. *J Diabetes Complications*. (2022) 36 (4):108150. doi: 10.1016/j.jdiacomp.2022.108150

50. Semnani-Azad Z, Blanco Mejia S, Connelly PW, Bazinet RP, Retnakaran R, Jenkins DJA, et al. The association of soluble CD163, a novel biomarker of macrophage activation, with type 2 diabetes mellitus and its underlying physiological disorders: A systematic review. *Obes Rev*. (2021) 22(8):e13257. doi: 10.1111/obr.13257

51. Yan W, Li Y, Wang G, Huang Y, Xie. P. Clinical application and immune infiltration landscape of stemness-related genes in heart failure. *ESC Heart Fail*. (2024) 12(1):250–70. doi: 10.1002/ehf2.15055

52. Wang C, Yang H, Gao. C. Potential biomarkers for heart failure. *J Cell Physiol*. (2019) 234:9467–74. doi: 10.1002/jcp.27632

53. Ptaszynska-Kopczynska K, Marcinkiewicz-Siemion M, Lisowska A, Waszkiewicz E, Witkowski M, Jasiewicz M, et al. Alterations of soluble TWEAK and CD163 concentrations in patients with chronic heart failure. *Cytokine*. (2016) 80:7–12. doi: 10.1016/j.cyto.2016.02.005

54. Durda P, Raffield LM, Lange EM, Olson NC, Jenny NS, Cushman M, et al. Circulating soluble CD163, associations with cardiovascular outcomes and mortality, and identification of genetic variants in older individuals: the cardiovascular health study. *J Am Heart Assoc*. (2022) 11:e024374. doi: 10.1161/jaha.121.024374

55. Pei X, Liu L, Wang J, Guo C, Li Q, Li J, et al. Exosomal secreted SCIMP regulates communication between macrophages and neutrophils in pneumonia. *Nat Commun*. (2024) 15:691. doi: 10.1038/s41467-024-44714-4

56. Wang Z, Wang Y, Yan Q, Cai C, Feng Y, Huang Q, et al. FPR1 signaling aberrantly regulates S100A8/A9 production by CD14(+)FCN1(hi) macrophages and aggravates pulmonary pathology in severe COVID-19. *Commun Biol*. (2024) 7:1321. doi: 10.1038/s42003-024-07025-4

57. Zhu S, Hu X, Bennett S, Mai Y, Xu. J. Molecular structure, expression and role of TAFA4 and its receptor FPR1 in the spinal cord. *Front Cell Dev Biol*. (2022) 10:911414. doi: 10.3389/fcell.2022.911414

58. Yi X, Tran E, Odiba JO, Qin CX, Ritchie RH, Baell. JB. The formyl peptide receptors FPR1 and FPR2 as targets for inflammatory disorders: recent advances in the development of small-molecule agonists. *Eur J Med Chem*. (2024) 265:115989. doi: 10.1016/j.ejmech.2023.115989

59. Vafadarnejad E, Rizzo G, Krampert L, Arampatzi P, Arias-Loza AP, Nazzal Y, et al. Dynamics of cardiac neutrophil diversity in murine myocardial infarction. *Circ Res*. (2020) 127:e232–49. doi: 10.1161/circresaha.120.317200

60. Asahina Y, Wurtz NR, Arakawa K, Carson N, Fujii K, Fukuchi K, et al. Discovery of BMS-986235/LAR-1219: A potent formyl peptide receptor 2 (FPR2) selective agonist for the prevention of heart failure. *J Med Chem*. (2020) 63:9003–19. doi: 10.1021/acs.jmedchem.9b02101

61. Zhangsun Z, Dong Y, Tang J, Jin Z, Lei W, Wang C, et al. FPR1: A critical gatekeeper of the heart and brain. *Pharmacol Res*. (2024) 202:107125. doi: 10.1016/j.phrs.2024.107125

62. Döring Y, Bender A, Soehnlein. O. Lack of formyl-peptide receptor 1 mitigates atherosclerosis in hyperlipidemic mice. *Thromb Haemost*. (2024) 124:986–9. doi: 10.1055/s-0044-1787264

63. Zhou QL, Teng F, Zhang YS, Sun Q, Cao YX, Meng. GW. FPR1 gene silencing suppresses cardiomyocyte apoptosis and ventricular remodeling in rats with ischemia/reperfusion injury through the inhibition of MAPK signaling pathway. *Exp Cell Res*. (2018) 370:506–18. doi: 10.1016/j.yexcr.2018.07.016

64. García RA, Ito BR, Lupisella JA, Carson NA, Hsu MY, Fernando G, et al. Preservation of post-infarction cardiac structure and function via long-term oral formyl peptide receptor agonist treatment. *JACC Basic Transl Sci*. (2019) 4:905–20. doi: 10.1016/j.jacbts.2019.07.005

65. Li Y, Wang Q, Li J, Li A, Wang Q, Zhang Q, et al. Therapeutic modulation of V Set and Ig domain-containing 4 (VSIG4) signaling in immune and inflammatory diseases. *Cytotherapy*. (2023) 25(6):561–72. doi: 10.1016/j.jcyt.2022.12.004

66. Liu B, Cheng L, Gao H, Zhang J, Dong Y, Gao W, et al. The biology of VSIG4: Implications for the treatment of immune-mediated inflammatory diseases and cancer. *Cancer Lett*. (2023) 553:215996. doi: 10.1016/j.canlet.2022.215996

67. Liu MA, Shahabi S, Jati S, Tang K, Gao H, Jin Z, et al. Gut microbial DNA and immune checkpoint gene Vsig4/CRIg are key antagonistic players in healthy aging and age-associated development of hypertension and diabetes. *Front Endocrinol (Lausanne)*. (2022) 13:1037465. doi: 10.3389/fendo.2022.1037465

68. Li S, Ge T, Xu X, Xie L, Song S, Li R, et al. Integrating scRNA-seq to explore novel macrophage infiltration-associated biomarkers for diagnosis of heart failure. *BMC Cardiovasc Disord*. (2023) 23:560. doi: 10.1186/s12872-023-03593-1

69. Wang Y, Ding J, Song H, Teng Y, Fang. X. VSIG4 regulates macrophages polarization and alleviates inflammation through activating PI3K/AKT and inhibiting TLR4/NF-κB pathway in myocardial ischemia-reperfusion injury rats. *Physiol Int*. (2022). doi: 10.1556/2060.2022.00055

70. Wang Y, Zhang Y, Li J, Li C, Zhao R, Shen C, et al. Hypoxia induces M2 macrophages to express VSIG4 and mediate cardiac fibrosis after myocardial infarction. *Theranostics*. (2023) 13:2192–209. doi: 10.7150/thno.78736

71. di Salvo TG, Yang KC, Brittain E, Absi T, Maltais S, Hemnes. A. Right ventricular myocardial biomarkers in human heart failure. *J Card Fail*. (2015) 21:398–411. doi: 10.1016/j.cardfail.2015.02.005

72. Xie Z, Shen Y, Huang S, Shen W, Liu. J. Abnormal ADAMTS2 and VSIG4 in serum of HF patients and their relationship with CRP, UA, and HCY. *Clin Lab*. (2022) 68:(5). doi: 10.7754/Clin.Lab.2021.210811

73. Milenkovic I, Santos Vieira HG, Lucas MC, Ruiz-Orera J, Patone G, Kesteven S, et al. Dynamic interplay between RPL3- and RPL3L-containing ribosomes modulates mitochondrial activity in the mammalian heart. *Nucleic Acids Res*. (2023) 51:5301–24. doi: 10.1093/nar/gkad121

74. Gu J, Hu W, Song ZP, Chen YG, Zhang DD, Wang. CQ. Rapamycin inhibits cardiac hypertrophy by promoting autophagy via the MEK/ERK/beclin-1 pathway. *Front Physiol*. (2016) 7:104. doi: 10.3389/fphys.2016.00104

75. He X, Yang T, Lu YW, Wu G, Dai G, Ma Q, et al. The long noncoding RNA CARDINAL attenuates cardiac hypertrophy by modulating protein translation. *J Clin Invest*. (2024) 134:(13). doi: 10.1172/jci169112

76. Kasahara K, Nakayama R, Shiwa Y, Kanesaki Y, Ishige T, Yoshikawa H, et al. Fpr1, a primary target of rapamycin, functions as a transcription factor for ribosomal protein genes cooperatively with Hmo1 in Saccharomyces cerevisiae. *PloS Genet*. (2020) 16:e1008865. doi: 10.1371/journal.pgen.1008865

77. Textoris J, Ivorra D, Ben Amara A, Sabatier F, Ménard JP, Heckenroth H, et al. Evaluation of current and new biomarkers in severe preeclampsia: a microarray approach reveals the VSIG4 gene as a potential blood biomarker. *PloS One*. (2013) 8: e82638. doi: 10.1371/journal.pone.0082638

78. Lund NC, Kayode Y, McReynolds MR, Clemmer DC, Hudson H, Clerc I, et al. mTOR regulation of metabolism limits LPS-induced monocyte inflammatory and procoagulant responses. *Commun Biol*. (2022) 5:878. doi: 10.1038/s42003-022-03804-z

79. Simcox J, Lamming. DW. The central moTOR of metabolism. *Dev Cell*. (2022) 57:691–706. doi: 10.1016/j.devcel.2022.02.024

80. Rudi WS, Molitor M, Garlapati V, Finger S, Wild J, Münzel T, et al. ACE inhibition modulates myeloid hematopoiesis after acute myocardial infarction and reduces cardiac and vascular inflammation in ischemic heart failure. *Antioxid (Basel)*. (2021) 10:(3). doi: 10.3390/antiox10030396

81. Shi M, Yuan H, Li Y, Guo Z, Wei. J. Targeting macrophage phenotype for treating heart failure: A new approach. *Drug Des Devel Ther*. (2024) 18:4927–42. doi: 10.2147/dddt.S486816

82. Lee TM, Chang NC, Lin. SZ. Dapagliflozin, a selective SGLT2 Inhibitor, attenuated cardiac fibrosis by regulating the macrophage polarization via STAT3 signaling in infarcted rat hearts. *Free Radic Biol Med*. (2017) 104:298–310. doi: 10.1016/j.freeradbiomed.2017.01.035

83. Zhang F, Xuan Y, Cui J, Liu X, Shao Z, Yu. B. Nicorandil modulated macrophages activation and polarization via NF-κb signaling pathway. *Mol Immunol*. (2017) 88:69–78. doi: 10.1016/j.molimm.2017.06.019

84. Yang B, Qiao Y, Yan D, Meng. Q. Targeting interactions between fibroblasts and macrophages to treat cardiac fibrosis. *Cells*. (2024) 13:(9). doi: 10.3390/cells13090764

85. Barcena ML, Niehues MH, Christiansen C, Estepa M, Haritonow N, Sadighi AH, et al. Male macrophages and fibroblasts from C57/BL6J mice are more susceptible to inflammatory stimuli. *Front Immunol.* (2021) 12:758767. doi: 10.3389/fimmu.2021.758767

86. Watany MM, Elhosary. MM. Clinical utility of circulating TWEAK and CD163 as biomarkers of iron-induced cardiac decompensation in transfusion dependent thalassemia major. *Cytokine*. (2024) 173:156443. doi: 10.1016/j.cyto.2023.156443

87. Margraf A, Chen J, Christoforou M, Claria-Ribas P, Henriques Schneider A, Cecconello C, et al. Formyl-peptide receptor type 2 activation mitigates heart and lung damage in inflammatory arthritis. *EMBO Mol Med*. (2025) 17(5):1153–83. doi: 10.1038/s44321-025-00227-1

88. Kato N, Kambe T, Chiba T, Taguchi K, Abe. K. Analgesic effect of a cholinergic agonist (carbachol) in a sural nerve ligation-induced hypersensitivity mouse model. *Neurol Res*. (2024) 46:505–15. doi: 10.1080/01616412.2024.2337512

89. Moussa AT, Rabung A, Reichrath S, Wagenpfeil S, Dinh T, Krasteva-Christ G, et al. Modulation of macrophage phagocytosis *in vitro*-A role for cholinergic stimulation? *Ann Anat*. (2017) 214:31–5. doi: 10.1016/j.aanat.2017.07.007

90. Khashchenko EP, Krechetova LV, Vishnyakova PA, Fatkhudinov TK, Inviyaeva EV, Vtorushina VV, et al. Altered monocyte and lymphocyte phenotypes associated with pathogenesis and clinical efficacy of progestogen therapy for peritoneal endometriosis in adolescents. *Cells*. (2024) 13:(14). doi: 10.3390/cells13141187

frontiers | Frontiers in Immunology

TYPE Mini Review
PUBLISHED 18 August 2025
DOI 10.3389/fimmu.2025.1639823

OPEN ACCESS

EDITED BY
Yejun Tan,
Hong Kong Polytechnic University, China

REVIEWED BY
Jie Ren,
Dalian Medical University, China

*CORRESPONDENCE
Shengliang Qiu
shengliang.qiu@zcmu.edu.cn
Yibo He
heyb20173626@126.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 02 June 2025
ACCEPTED 23 July 2025
PUBLISHED 18 August 2025
CORRECTED 12 December 2025

CITATION
Chen S, Chen W, Xu T, Li J, Yu J, He Y and Qiu S (2025) The impact of aberrant lipid metabolism on the immune microenvironment of gastric cancer: a mini review.
Front. Immunol. 16:1639823.
doi: 10.3389/fimmu.2025.1639823

The impact of aberrant lipid metabolism on the immune microenvironment of gastric cancer: a mini review

Shuangyu Chen[1†], Wenqian Chen[1†], Tinghui Xu[1†], Jiayang Li[1†], Jianghao Yu[2], Yibo He[1*] and Shengliang Qiu[1*]

[1]The First Affiliated Hospital of Zhejiang Chinese Medical University (Zhejiang Provincial Hospital of Chinese Medicine), Hangzhou, Zhejiang, China, [2]School of Medical Technology and Information Engineering, Zhejiang Chinese Medical University, Hangzhou, Zhejiang, China

Gastric cancer (GC) remains one of the leading causes of cancer-related mortality worldwide, with limited responses to immune checkpoint blockade (ICB) therapies in most patients. Increasing evidence indicates that the tumor immune microenvironment (TIME) plays a crucial role in immunotherapy outcomes. Among various metabolic abnormalities in the TIME, dysregulated lipid metabolism has emerged as a critical determinant of immune cell fate, differentiation, and function. In this review, we comprehensively summarize the current understanding of the immune landscape in GC, focusing on how altered lipid metabolism reshapes immune cell populations—including tumor-associated macrophages (TAMs), dendritic cells (DCs), regulatory T cells (Tregs), myeloid-derived suppressor cells (MDSCs), and cytotoxic $CD8^+$ T cells. We highlight key metabolic pathways such as fatty acid oxidation(FAO), cholesterol homeostasis, and lipid uptake that impact immune cell activity, contributing to immune evasion and therapeutic resistance. Importantly, we explore emerging therapeutic strategies targeting lipid metabolism, including inhibitors of cluster of differentiation 36 (CD36), fatty acid synthase (FASN), and sterol regulatory element-binding protein 1 (SREBP1) and discuss their synergistic potential when combined with ICB therapies. In conclusion, lipid metabolic reprogramming represents a promising yet underexplored axis in modulating antitumor immunity in GC. Integrating metabolic intervention with immunotherapy holds potential to overcome current treatment limitations and improve clinical outcomes. Future studies incorporating spatial omics and single-cell profiling will be essential to elucidate cell-type specific metabolic dependencies and foster translational breakthroughs.

KEYWORDS

gastric cancer, lipid metabolism, tumor immune microenvironment, CD8+ T cells, tumor-associated macrophages, immunotherapy resistance, fatty acid oxidation, immune checkpoint blockade

1 Introduction

According to GLOBOCAN 2022 statistics, in 2022, more than 968,000 new cases of gastric cancer (GC) were added, with nearly 660,000 deaths, ranking fifth globally both in terms of incidence and mortality. The region with the highest incidence rate is East Asia, which imposes a significant burden on cancer (1). Consequently, an urgent exploration and development of new therapeutic approaches has become imperative.

The tumor microenvironment (TME) is a complex system that can inhibit immune responses while promoting tumor progression. The composition of the TME differs across different tumor types, but its defining features include immune cells, stromal cells, vasculature, and extracellular matrix (2, 3). The complexity and dynamic interactions within the TME contribute significantly to the aggressive nature of GC and the development of therapeutic resistance (4). Therefore, understanding the intricate characteristics of the TME, particularly metabolic reprogramming within this milieu, is of substantial clinical importance for developing effective treatments for GC patients.

Metabolic reprogramming is widely recognized as a hallmark of cancer, allowing tumor cells to sustain proliferation, evade immune surveillance, and survive under stressful conditions. Among various metabolic alterations, abnormal lipid metabolism has emerged as a pivotal player in cancer progression, influencing energy metabolism, membrane biosynthesis, and signaling pathways (5–7). Cancer cells undergo significant lipid metabolic reprogramming, including increased lipid uptake, enhanced fatty acid synthesis (FAS), and elevated fatty acid oxidation (FAO). These alterations not only provide essential metabolic substrates but also enable cancer cells to resist oxidative stress, promoting tumor survival and resistance to conventional therapies (8).

Key enzymes involved in lipid metabolism, such as fatty acid synthase (FASN), ATP citrate lyase (ACLY), and stearoyl-CoA desaturase (SCD), are upregulated in GC (9–11), indicating their potential as therapeutic targets. Aberrant lipid metabolic pathways influence the recruitment, differentiation, and function of key immune cell populations including tumor-associated macrophages (TAMs), regulatory T cells (Tregs),and myeloid-derived suppressor cells (MDSCs), dendritic cells (DCs), CD8+ T cells,contributing to an immunosuppressive microenvironment that facilitates tumor progression.

Abbreviations: ACC, acetyl CoA carboxylase; ACLY, ATP citrate lyase; ACSLs, acyl-CoA synthetase long-chain family members; CAFs, cancer-associated fibroblasts; CD36, cluster of differentiation 36; CPT1, carnitine palmitoyl-transferase 1; DCs, dendritic cells; FAO, fatty acid oxidation; FAS, fatty acid synthesis; FASN, fatty acid synthase; FABP5, fatty acid-binding protein 5; GC, gastric cancer; ICIs, immune checkpoint inhibitors; LSR, lipolysis-stimulated lipoprotein receptor; MDSCs, myeloid-derived suppressor cells; MHC, major histocompatibility complex; MIF, migration inhibitory factor; PPAR-γ, peroxisome proliferator-activated receptor γ; ROS, reactive oxygen species; SCD, stearoyl-CoA desaturase; SREBP1, sterol regulatory element-binding protein 1; TAMs, tumor-associated macrophages; TME, tumor microenvironment; Tpex, progenitor-exhausted T cells; Trm, tissue-resident memory T cells; Tregs, regulatory T cells.

2 Lipid metabolic pathways and molecular mechanisms

A key energy-generating pathway in lipid metabolism is mitochondrial fatty acid β-oxidation, which is mediated by carnitine palmitoyl-transferase 1 (CPT1), especially the isoform CPT1a (12, 13). This enzyme facilitates the transport of long-chain fatty acids to the mitochondria for oxidative breakdown and ATP production, particularly under nutrient-deprived conditions (14). Simultaneously, cancer cells exploit exogenous lipid sources through dietary uptake, with cluster of differentiation 36 (CD36) functioning as a major fatty acid translocase (15). CD36 is frequently overexpressed in malignant cells, contributing to enhanced fatty acid uptake, intracellular lipid accumulation, and increased metabolic plasticity (16–18). This metabolic architecture is tightly regulated by oncogenic signaling cascades, especially the PI3K/Akt/mTOR axis. This axis activates sterol regulatory element-binding protein 1 (SREBP1), a master transcriptional regulator of lipid biosynthesis (19, 20). When SREBP1 is activated, the expression of key enzymes involved in fat production, such as FASN and acetyl CoA carboxylase (ACC), is enhanced. This can promote *de novo* fat generation and support the promotion of membrane biogenesis and proliferation (21, 22). The uptake of extracellular lipids via CD36 and the lipolysis-stimulated lipoprotein receptor (LSR) is often upregulated in tumors and is also responsive to PI3K/mTOR signaling, reinforcing the lipid supply for cancer progression (23–25). Enzymes like acyl-CoA synthetase long-chain family members (ACSLs) activate imported fatty acids and channel them into biosynthetic and storage pathways, while lipogenesis induced by SREBP1 inhibits ferroptosis and improves tumor cell survival (20, 26). Uptake of lipids by CD36 enhances metastatic potential and contributes to adaptation to the TME (27). Additionally, reorganization of lipid metabolism can alter antigen presentation and inhibit T-cell activation, leading to impairment of immune surveillance (28). Phospholipid remodeling represents another critical branch of lipid metabolism. This metabolic adaptation highlights the key function of lipid metabolism in coordinating cellular bioenergetics with tumor invasiveness and immune escape, laying the mechanistic foundation for its involvement in the formation of an immunosuppressive TME (29).

3 Overview of the immune microenvironment in gastric cancer

TME of GC is composed of various immune cell subsets and non-immune components, and is characterized by prominent immunosuppressive features. Single-cell analyses have revealed a highly heterogeneous pattern of immune cell infiltration within the TME of GC. Immunosuppressive components such as Tregs,

MDSCs, and TAMs are widely distributed and are closely associated with ineffective antitumor immune responses (30–33). Tregs suppress $CD8^+$ T cell activity and the antigen presentation process through multiple mechanisms, serving as key regulatory factors in the progression of GC (34, 35). MDSCs exacerbate the immunosuppressive state by secreting inhibitory factors and modulating macrophage polarization (36). Moreover, M2 polarization of TAMs in GC has been shown to be closely associated with immune evasion and poor prognosis (37–39). Another key mechanism underlying the immunosuppressive TME is the upregulation of immune checkpoints, such as PD-L1 and the CD39/CD73 axis, which inhibit T cell effector functions and promote tumor immune evasion (40, 41). Studies have indicated that the TME in GC patients often exhibits a "cold tumor" phenotype—characterized by low immune cell infiltration and weak immune activation—which not only predicts poor prognosis but also correlates with low responsiveness to immunotherapy (42, 43).

Immune infiltration patterns exhibit dynamic changes across different GC subtypes and treatment contexts. Neoadjuvant chemotherapy can significantly remodel the TME by enhancing CD8+T cell infiltration and reducing immunosuppressive cells, highlighting the plasticity of the immune landscape (44, 45). High-throughput analyses and multiplex immunofluorescence have revealed complex interactions among different immune cells within the TME, such as exosome-mediated communication between TAMs and cancer cells (46, 47).

Furthermore, the degree of immune cell infiltration is closely associated with clinical outcomes. For instance, high PD-L1 expression often coexists with an "immune-excluded" infiltration pattern, suggesting that patients may benefit from immune checkpoint inhibitor therapy (48, 49). Key molecular features of the TME significantly shape immune infiltration and immunotherapy responses in GC, highlighting new avenues for enhancing antitumor immunity (50–52). Among these features, spatial metabolic heterogeneity — particularly lipid gradients within the TME — has recently gained attention as a critical factor influencing immune cell behavior.

4 Interactions between aberrant lipid metabolism and immune cells

4.1 TAMs

TAMs, one of the most abundant immune cells in the GC immune microenvironment, exhibit significant metabolic plasticity. Under the stimulation of various cytokines, macrophages can be polarized into two phenotypes with different functions: M1 macrophages, which have pro-inflammatory and tumor-inhibiting effects; And M2 macrophages, which have anti-inflammatory and tumor-promoting effects Their functional state is closely linked to their lipid metabolic program. In gastric cancer, scavenger receptors such as CD36 mediate the endocytosis of fatty acids and cholesterol from the tumor microenvironment, leading to intracellular lipid accumulation and promoting the establishment of a highly immunosuppressive TME (53, 54). This process further activates the peroxisome proliferator-activated receptor γ (PPAR-γ) signaling pathway, upregulating FAO, promoting TAM towards a m2 polarized phenotype, and enhancing its oncogenic function (55, 56). Moreover, lipid uptake promotes enhanced FAO, providing a stable energy supply for M2-polarized TAMs and augmenting their secretion of immunosuppressive factors such as IL-10 and TGF-β (57–59). These alterations collectively contribute to the formation of a microenvironment that favors tumor survival and immune evasion (60, 61). Mechanistically, lipid uptake via CD36 facilitates intracellular fatty acid accumulation, which activates PPAR-γ signaling and upregulates key enzymes of FAO, such as CPT1A.

Further studies have revealed that the metabolic state of TAMs is a key determinant of their spatial distribution and functional heterogeneity. For example, lipid-rich TAMs are predominantly located in hypoxic regions, where they respond to tumor-derived factors such as IL-34 and signals associated with p53 inactivation, exhibiting enhanced immunosuppressive capabilities (62, 63). At the metabolic level, lipid metabolic reprogramming is closely regulated by the TRAF3/STAT6 pathway, which governs key transcriptional programs involved in the polarization process (64). Meanwhile, signaling molecules such as CD40 have been shown to promote the reprogramming of TAMs toward an antitumor phenotype by remodeling fatty acid and glutamine metabolism, highlighting the potential of metabolic interventions in reshaping TAM function (65). Overall, lipid uptake and metabolism determine the fate of TAMs, representing a critical regulatory axis within the GC immune microenvironment and a promising therapeutic target for future treatment strategies (66). These findings highlight the central role of TAM lipid metabolism in promoting immune evasion and progression of gastric cancer.

4.2 Dendritic cells

DCs within the GC immune microenvironment is often markedly suppressed by dysregulated lipid metabolism. In gastric cancer, this metabolic dysfunction contributes to impaired tumor antigen presentation and weakened immune surveillance.The lipid-rich tumor environment leads to lipid accumulation in DCs, particularly the formation of lipid droplets enriched with cholesterol and triglycerides, which significantly impairs their antigen-presenting capacity (67, 68). Lipid overload not only diminishes the expression of major histocompatibility complex (MHC) class I and II molecules but also suppresses the expression of costimulatory molecules such as CD80 and CD86, thereby limiting T cell activation (69, 70). Studies have shown that Epstein-Barr virus–associated GC exacerbates antigen presentation impairment by secreting exosomes that interfere with DC maturation (70). Moreover, tumor-induced lipid metabolic reprogramming can suppress mitochondrial function and glucose metabolism in DCs, driving them toward an immunotolerant phenotype (67, 68). A decline in cross-presentation capacity is another critical defect of lipid-laden DCs, particularly impairing

their ability to elicit CD8$^+$ T cell responses (71, 72). Some studies have reported that lipid accumulation hinders the ability of DCs to uptake and process extracellular antigens, thereby weakening their effectiveness in activating tumor-specific T cells (73, 74). Furthermore, Tregs form immunosuppressive complexes with DCs through a CXCR3-mediated chemotactic mechanism, further limiting the ability of DCs to activate CD8+ T cells (75). In recent years, engineered dendritic cell (DC) systems have been developed to bypass the metabolic impairments of natural DCs, offering new avenues for tumor vaccines and targeted immunotherapy (76, 77). Therefore, targeting lipid metabolic regulatory pathways is considered a potential strategy to restore DC immune function and enhance immune responses in gastric cancer (78, 79).

4.3 Tregs and MDSCs

Tregs are abundantly infiltrated in the GC immune microenvironment and rely on lipid metabolism to maintain their stability and immunosuppressive function. Studies have shown that within the tumor environment, Tregs gain an energetic advantage by enhancing FAO, which sustains their Foxp3 expression and suppressive capacity (80, 81). PD-1 deficiency disrupts the metabolic stability of Tregs, suggesting that their metabolic adaptability is a critical factor in the establishment of immune tolerance (80). Moreover, fatty acid-binding protein 5 (FABP5) and the SIRT1–CX3CL1 axis play important roles in regulating lipid metabolism in Tregs, influencing their distribution within the TME and their immunosuppressive capacity (82, 83). In lipid-rich microenvironments, Tregs exhibit enhanced stability and activity, representing one of the major obstacles to the efficacy of immune checkpoint inhibition therapy (84, 85).

Similar to Tregs, MDSCs exhibit potent immunosuppressive properties regulated by lipid metabolism. In high-lipid microenvironments, they sustain their survival through FAS and cholesterol metabolism, while secreting a range of immunosuppressive factors (29, 86). Ginger polysaccharide–induced lipid metabolic disruption can promote apoptosis of MDSCs, indicating that targeting lipid metabolism holds potential for enhancing immune responses (86). Within the GC TME, MDSCs cooperate with Tregs to establish a metabolically coupled immunosuppressive network (87, 88). Recent studies have shown that cancer-associated fibroblasts (CAFs) influence the metabolic activity of MDSCs through CD36 and the secretion of macrophage migration inhibitory factor (MIF), further exacerbating immune evasion (87). In summary, targeting lipid metabolism has emerged as a key strategy for modulating the functions of Tregs and MDSCs and overcoming immune tolerance (29, 85).

4.4 CD8$^+$ T cells

CD8$^+$ T cells are the central effector cells in antitumor immune responses, and their functional state is significantly influenced by dysregulated lipid metabolism within the TME. In the GC microenvironment, fatty acid uptake and cholesterol metabolism reshape the metabolic programming of CD8$^+$ T cells, leading to metabolic imbalance, enhanced exhaustion phenotypes, and reduced cytotoxic function (89). Tumor cells secrete lipid metabolism–regulating factors such as SCD1 and FABP5, which elevate levels of free fatty acids and oxidized lipids in the TME. This induces the accumulation of reactive oxygen species (ROS) in CD8$^+$ T cells, leading to lipid peroxidation and mitochondrial damage (90). This process is accompanied by the upregulation of inhibitory receptors such as PD-1 and TIGIT, ultimately leading to T cell exhaustion and the loss of sustained cytotoxic activity (91). Moreover, excess cholesterol can accumulate in the membranes of CD8$^+$ T cells, disrupting immunological synapse formation and TCR signaling, thereby further suppressing their effector functions (92).

Studies have also indicated that certain lipid metabolic pathways exert bidirectional regulatory effects on CD8$^+$ T cells. Tissue-resident CD8$^+$ T cells rely on FAO to sustain energy supply and long-term survival; however, in the nutrient-deprived and competitive TME, this metabolic dependency may actually constrain the sustained activation of their effector functions (89). Under high-lipid conditions, tumor cells compete with CD8$^+$ T cells for nutritional substrates, leading to energy deprivation in CD8$^+$ T cells. This results in a state of "functional starvation," characterized by reduced expression of effector molecules such as Granzyme B and IFN-γ (27, 91). Therefore, targeting lipid metabolic pathways—such as CD36 inhibition, FAO blockade, or cholesterol metabolism modulation—is considered a promising strategy to restore CD8$^+$ T cell function and enhance the efficacy of immunotherapy (90, 93) (Figure 1).

5 Clinical and therapeutic implications

Lipid metabolic reprogramming is not only a key mechanism in shaping the TME of GC, but also offers multidimensional therapeutic targets for clinical intervention. High expression of key lipid metabolic molecules such as CD36, FASN, and SREBP1 is closely associated with the infiltration of immunosuppressive cells and T-cell exhaustion, and is considered one of the major contributors to immunotherapy resistance (94–96). For instance, Li et al. found that lipid metabolic imbalance can promote symbiotic signaling pathways between CAFs and TAMs, which significantly impairs the efficacy of immune checkpoint inhibitors (ICIs) (97). Emerging lipid-targeted strategies —such as FASN inhibitors, FAO pathway blockers, and cholesterol metabolism modulators—are being actively explored to enhance CD8$^+$T cell function, inhibit TAM polarization, and reduce Treg-mediated immunosuppression (94, 98, 99). Moreover, lipid metabolism–related genes have also been identified as potential predictive biomarkers of immune response. Genes such as RGS2, APOD, and MTTP have demonstrated promising prognostic and therapeutic response prediction value in multiple studies (94, 96, 98).

Combination therapy strategies are emerging as a key approach to overcoming the bottlenecks of immunotherapy in GC Several clinical trials—such as ATTRACTION-2, ATTRACTION-4, KEYNOTE-859, KEYNOTE-061 and CheckMate-649—have validated the efficacy of combining ICIs with chemotherapy (100–104). Combination

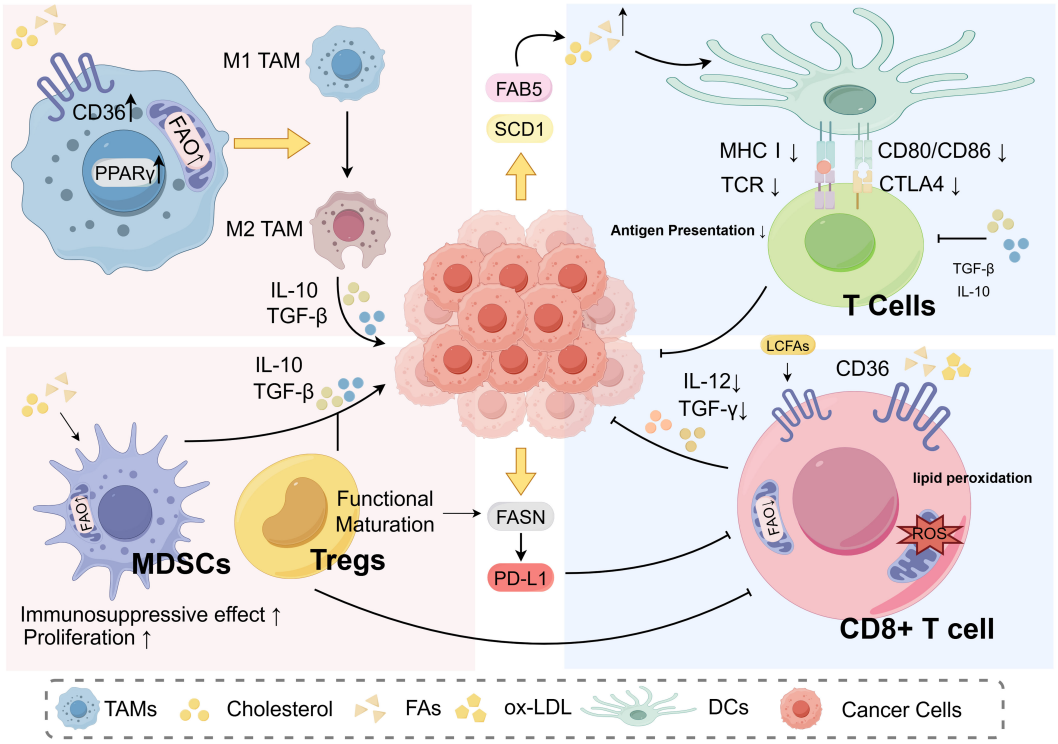


FIGURE 1
Interactions between aberrant lipid metabolism and immune cells.

strategies involving CD36 antagonists or cholesterol synthase inhibitors have significantly enhanced antitumor immune responses in preclinical models (94). Meanwhile, lipid metabolism–based immune subtyping approaches are increasingly being employed to guide the selection of GC patients for immunotherapy (101, 105). In summary, the role of lipid metabolism in precision immunotherapy for GC is becoming increasingly prominent. Existing clinical trials combining immune checkpoint inhibitors with chemotherapy have demonstrated heterogeneous outcomes, which may partially reflect underlying metabolic states of the tumor immune microenvironment (106–110). Aberrant expression of lipid metabolism–related molecules such as FASN, CD36, and SREBP1 has been associated with immune cell exhaustion, Treg enrichment, and impaired dendritic cell function, suggesting their potential value as both therapeutic targets and predictive biomarkers (111–114). Integrating lipidomic analysis into future clinical trial designs may enhance stratification strategies and optimize combination regimens to overcome resistance (Table 1).

6 Research gaps and future perspectives

Although the role of lipid metabolism in regulating the immune microenvironment of GC has been progressively elucidated, many gaps remain in understanding its mechanistic network. Current research primarily focuses on classical lipid metabolism regulators such as CD36 and FASN, while the roles of non-coding RNAs and RNA modifications (e.g. m^6A) in the cross-regulation of lipid metabolism remain largely underexplored (115–117). Moreover, how lipid metabolism specifically affects different immune cell subsets—such as tissue-resident memory T cells (Trm) and progenitor-exhausted T cells (Tpex)—remains insufficiently investigated at the single-cell resolution level (118, 119). Most current mechanistic studies are based on *in vitro* cell experiments and traditional animal models, with a lack of application of emerging technologies—such as spatial transcriptomics, spatial metabolomics, and single-cell lipidomics—for constructing a "functional lipid map" within the immune microenvironment (120, 121).

In future research, a primary focus should be the expanded systematic screening of lipid metabolism regulators, including transporters, enzymes, and intermediate metabolites, to evaluate their immunological effects (122, 123). Secondly, integrating clinical cohorts to perform lipid metabolic phenotyping and establishing a biomarker system capable of predicting immunotherapy response and resistance risk will be critical for advancing personalized treatment (124–126). Moreover, constructing *in vitro* microenvironment models—such as organoid–immune cell co-culture systems—or developing novel drug delivery platforms targeting lipid metabolism will help bridge the gap between basic

TABLE 1 Clinical trials of immunotherapy-based combination strategies in gastric cancer.

Trial	Phase	Drugs	Actual enrollment	Study period	Reference	Lipid metabolism/ immune remodeling findings
NCT02872116 (CHECKMATE-649)	III	Nivolumab + Ipilimumab or Nivolumab in Combination With Oxaliplatin + Fluoropyrimidine vs Oxaliplatin + Fluoropyrimidine	2031	May 27, 2020- May 31, 2024	(100)	↑ $CD8^+$ T cells, ↓ PD-L1 immune evasion; lipid modifications regulate PD-L1.
NCT02746796 (ATTRCTION-04)	II/III	SOX/Capecitabine + Oxaliplatin with vs without Nivolumab	724	March 7, 2017 – May 10, 2018	(101)	↑ $CD8^+$ T cells; enhanced tumor microenvironment immune activation.
NCT03675737 (KEYNOTE-859)	III	Pembrolizumab+ Chemotherapy vs Placebo + Chemotherapy	1579	November 8, 2018 – September 28, 2024	(102)	↑ PD-L1 expression, immune activation linked to lipid gene co-signatures.
NCT03878472	II	Camrelizumab + Apatinib + S-1 ± Oxaliplatin	25	April 1, 2019 – May 31, 2024	(105)	↑ $CD8^+$ T cells, ↓ PD-L1 immune evasion.
NCT04082364 (MAHOGANY)	II/III	Combination Margetuximab, Retifanlimab, Tebotelimab, and Chemotherapy	81	September 30, 2019 - December 2023	(106)	↑ T-cell activation via PD-1 and LAG-3 blockade; HER2–PD-L1 immune crosstalk implicated.
NCT03335540 (ADVISE)	I	Nivolumab + Ipilimumab vs Nivolumab	20	May 7, 2018 – August 25, 2021	(107)	↑ Immune markers in low/ intermediate PD-L1 tumors; ↑ T-cell and macrophage activation.
NCT03662659 (RELATIVITY-060)	II	Relatlimab + Nivolumab + XELOX/FOLFOX/SOX vs. Nivolumab + XELOX/ FOLFOX/SOX	274	October 16, 2018 – January 16, 2024	(108)	↑ T-cell activation via PD-1 and LAG-3 blockade
NCT04908566	II	PD-1 inhibitor + mFOLFIRINOX vs. mFOLFIRINOX	30	August 2023 – May 2025	(109)	↑ $CD8^+$ T and NK cells, ↓ macrophages and $FOXP3^+$ Tregs; dynamic immune remodeling predicts response
NCT04997837	III	Chemotherapy + PD-1 inhibitor + Radiotherapy VS Chemotherapy	433	July 21, 2021 – July 21, 2027	(110)	Radiation-induced PD-L1 upregulation
NCT03615326 (KEYNOTE-811)	III	Pembrolizumab/Trastuzumab/ Chemotherapy vs Trastuzumab/Chemotherapy	698	October 5, 2018 - March 20, 2024	(111)	↑ T-cell activation; HER2–PD-L1 crosstalk enhances immune response with pembrolizumab.
NCT02589496	II	Pembrolizumab	45	March 26, 2016 – December 2021	(113)	↑ Immune activation; metabolism pathways and epigenetic features linked to tumor microenvironment score (TMEscore) predicting ICB response.
NCT04182724 (KEYNOTE-061)	II	PD-1 inhibitor + albumin-bound paclitaxel + apatinib	43	July 11, 2019 – October 13, 2022	(114)	↑ PD-L1 expression; VEGFR inhibition and immune activation via PD-1 blockade.
NCT02267343 (ATTRACTION-2)	III	Nivolumab vs Placebo	493	October 2014 – January 2021	(115)	↑ PD-L1–dependent immune response; lipid metabolism not reported.
NCT05008783	III	Cadonilimab + Oxaliplatin + Capecitabine (XELOX) vs. Placebo + Oxaliplatin + Capecitabine (XELOX)	610	September 17, 2021 – October 18, 2025	(116)	↑ PD-L1 expression; enhanced immune activation via dual PD-1/ CTLA-4 blockade.

↑, upregulated; ↓, downregulated.

research and clinical application in metabolic immune regulation (127). Building on this foundation, conducting multicenter prospective clinical studies to evaluate the efficacy and safety of lipid metabolism–targeted interventions combined with immunotherapy will be a key pathway toward the clinical translation of metabolism-based immunotherapies (128, 129).

7 Conclusion

Lipid metabolism plays a central regulatory role in the TME of GC. Lipid competition between tumor cells and immune cells not only reshapes energy metabolism patterns but also alters immune cell functional states, inducing immunosuppressive phenotypes such as M2 polarization of TAMs, impaired antigen presentation by DCs, enhanced Treg functionality, and exhaustion of $CD8^+$ T cells (30, 32, 34). Lipid metabolic reprogramming mechanisms—including CD36-mediated lipid uptake, enhanced FAO, and cholesterol accumulation—have been shown to play critical roles in GC progression and immune evasion by regulating immune checkpoint expression, immune cell metabolic adaptation, and the secretion of immunosuppressive factors (40, 48, 101). Targeting lipid metabolic pathways—such as FASN, CPT1A, CD36, or cholesterol metabolism—can enhance immunotherapeutic responses and alleviate the immunosuppressive nature of the TME, demonstrating promising translational potential (123). However, the cell-specific functions of lipid metabolism across different immune cell subsets, its spatial heterogeneity, and the interplay between metabolic and epigenetic regulation axes remain to be further investigated (119, 130, 131). Future research should integrate emerging technologies such as spatial transcriptomics, single-cell lipidomics, and multi-omics analyses, while establishing clinical cohorts to explore predictive biomarkers and novel strategies for metabolism-targeted therapies (127, 132).

Author contributions

SC: Writing – original draft, Writing – review & editing. WC: Writing – original draft. TX: Writing – original draft. JL: Writing – original draft. JY: Writing – original draft. YH: Writing – original draft, Writing – review & editing. SQ: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Acknowledgments

We utilized the Figdraw online platform (https://www.figdraw.com/) to create the figures in this manuscript and extend our heartfelt thanks for the support and functionality it offers.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Correction note

A correction has been made to this article. Details can be found at: 10.3389/fimmu.2025.1761409.

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. The authors utilized ChatGPT-4.0 to assist with language refinement during the preparation of this work. All content was subsequently reviewed and revised by the authors, who take full responsibility for the final version of the publication.

Publisher's note

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin*. (2024) 74:229–63. doi: 10.3322/caac.21834

2. Anderson NM, Simon MC. The tumor microenvironment. *Curr biol: CB*. (2020) 30:R921–r5. doi: 10.1016/j.cub.2020.06.081

3. Yasuda T, Wang YA. Gastric cancer immunosuppressive microenvironment heterogeneity: implications for therapy development. *Trends cancer*. (2024) 10:627–42. doi: 10.1016/j.trecan.2024.03.008

4. Liu Y, Li C, Lu Y, Liu C, Yang W. Tumor microenvironment-mediated immune tolerance in development and treatment of gastric cancer. *Front Immunol*. (2022) 13:1016817. doi: 10.3389/fimmu.2022.1016817

5. Snaebjornsson MT, Janaki-Raman S, Schulze A. Greasing the wheels of the cancer machine: the role of lipid metabolism in cancer. *Cell Metab*. (2020) 31:62–76. doi: 10.1016/j.cmet.2019.11.010

6. Broadfield LA, Pane AA, Talebi A, Swinnen JV, Fendt SM. Lipid metabolism in cancer: New perspectives and emerging mechanisms. *Dev Cell*. (2021) 56:1363–93. doi: 10.1016/j.devcel.2021.04.013

7. Jin HR, Wang J, Wang ZJ, Xi MJ, Xia BH, Deng K, et al. Lipid metabolic reprogramming in tumor microenvironment: from mechanisms to therapeutics. *J Hematol Oncol*. (2023) 16:103. doi: 10.1186/s13045-023-01498-2

8. Bacci M, Lorito N, Smiriglia A, Morandi A. Fat and furious: lipid metabolism in antitumoral therapy response and resistance. *Trends cancer*. (2021) 7:198–213. doi: 10.1016/j.trecan.2020.10.004

9. El-Rifai W, Frierson HF Jr., Moskaluk CA, Harper JC, Petroni GR, Bissonette EA, et al. Genetic differences between adenocarcinomas arising in Barrett's esophagus and gastric mucosa. *Gastroenterology*. (2001) 121:592–8. doi: 10.1053/gast.2001.27215

10. Ezzeddini R, Taghikhani M, Somi MH, Samadi N, Rasaee MJ. Clinical importance of FASN in relation to HIF-1α and SREBP-1c in gastric adenocarcinoma. *Life Sci*. (2019) 224:169–76. doi: 10.1016/j.lfs.2019.03.056

11. Gao Y, Li J, Xi H, Cui J, Zhang K, Zhang J, et al. Stearoyl-CoA-desaturase-1 regulates gastric cancer stem-like properties and promotes tumour metastasis via Hippo/YAP pathway. *Br J cancer*. (2020) 122:1837–47. doi: 10.1038/s41416-020-0827-5

12. Wang J, Xiang H, Lu Y, Wu T, Ji G. The role and therapeutic implication of CPTs in fatty acid oxidation and cancers progression. *Am J Cancer Res*. (2021) 11:2477–94.

13. Wang C, Zhang C, Li X, Shen J, Xu Y, Shi H, et al. CPT1A-mediated succinylation of S100A10 increases human gastric cancer invasion. *J Cell Mol Med*. (2019) 23:293–305. doi: 10.1111/jcmm.13920

14. Terry AR, Hay N. Emerging targets in lipid metabolism for cancer therapy. *Trends Pharmacol Sci*. (2024) 45:537–51. doi: 10.1016/j.tips.2024.04.007

15. Glatz JFC, Heather LC, Luiken J. CD36 as a gatekeeper of myocardial lipid metabolism and therapeutic target for metabolic disease. *Physiol Rev*. (2024) 104:727–64. doi: 10.1152/physrev.00011.2023

16. Guerrero-Rodríguez SL, Mata-Cruz C, Pérez-Tapia SM, Velasco-Velázquez MA. Role of CD36 in cancer progression, stemness, and targeting. *Front Cell Dev Biol*. (2022) 10:1079076. doi: 10.3389/fcell.2022.1079076

17. Terry AR, Nogueira V, Rho H, Ramakrishnan G, Li J, Kang S, et al. CD36 maintains lipid homeostasis via selective uptake of monounsaturated fatty acids during matrix detachment and tumor progression. *Cell Metab*. (2023) 35:2060–76.e9. doi: 10.1016/j.cmet.2023.09.012

18. Pascual G, Avgustinova A, Mejetta S, Martín M, Castellanos A, Attolini CS, et al. Targeting metastasis-initiating cells through the fatty acid receptor CD36. *Nature*. (2017) 541:41–5. doi: 10.1038/nature20791

19. Yi J, Zhu J, Wu J, Thompson CB, Jiang X. Oncogenic activation of PI3K-AKT-mTOR signaling suppresses ferroptosis via SREBP-mediated lipogenesis. *Proc Natl Acad Sci United States America*. (2020) 117:31189–97. doi: 10.1073/pnas.2017152117

20. Chen M, Li H, Zheng S, Shen J, Chen Y, Li Y, et al. Nobiletin targets SREBP1/ACLY to induce autophagy-dependent cell death of gastric cancer cells through PI3K/Akt/mTOR signaling pathway. *Phytomed: Int J phytother phytopharmacol*. (2024) 128:155360. doi: 10.1016/j.phymed.2024.155360

21. Chen J, Ye M, Gu D, Yu P, Xu L, Xue B, et al. FTO-induced APOE promotes the Malignant progression of pancreatic neuroendocrine neoplasms through FASN-mediated lipid metabolism. *Int J Biol Sci*. (2025) 21:1478–96. doi: 10.7150/ijbs.103428

22. Zhou X, Li Y, Yang C, Chen D, Wang T, Liu T, et al. Cordycepin reprogramming lipid metabolism to block metastasis and EMT via ERO1A/mTOR/SREBP1 axis in cholangiocarcinoma. *Life Sci*. (2023) 327:121698. doi: 10.1016/j.lfs.2023.121698

23. Kawabata K, Takahashi T, Tanaka K, Kurokawa Y, Yamamoto K, Saito T, et al. Lipolysis-stimulated lipoprotein receptor promote lipid uptake and fatty acid oxidation in gastric cancer. *Gastric Cancer*. (2024) 27:1258–72. doi: 10.1007/s10120-024-01552-z

24. Yang H, Zhao H, Ren Z, Yi X, Zhang Q, Yang Z, et al. Overexpression CPT1A reduces lipid accumulation via PPARα/CD36 axis to suppress the cell proliferation in ccRCC. *Acta Biochim Biophys Sinica*. (2022) 54:220–31. doi: 10.3724/abbs.2021023

25. Zhao S, Cheng L, Shi Y, Li J, Yun Q, Yang H. MIEF2 reprograms lipid metabolism to drive progression of ovarian cancer through ROS/AKT/mTOR signaling pathway. *Cell Death disease*. (2021) 12:18. doi: 10.1038/s41419-020-03336-6

26. Zheng YN, Lou SY, Lu J, Zheng FL, Tang YM, Zhang EJ, et al. Selective PI3Kδ inhibitor TYM-3–98 suppresses AKT/mTOR/SREBP1-mediated lipogenesis and promotes ferroptosis in KRAS-mutant colorectal cancer. *Cell Death disease*. (2024) 15:474. doi: 10.1038/s41419-024-06848-7

27. Tang Y, Chen Z, Zuo Q, Kang Y. Regulation of CD8+ T cells by lipid metabolism in cancer progression. *Cell Mol Immunol*. (2024) 21:1215–30. doi: 10.1038/s41423-024-01224-z

28. Wang Y, Guo Z, Isah AD, Chen S, Ren Y, Cai H. Lipid metabolism and tumor immunotherapy. *Front Cell Dev Biol*. (2023) 11:1187989. doi: 10.3389/fcell.2023.1187989

29. Zhang H, Li S, Wang D, Liu S, Xiao T, Gu W, et al. Metabolic reprogramming and immune evasion: the interplay in the tumor microenvironment. *biomark Res*. (2024) 12:96. doi: 10.1186/s40364-024-00646-1

30. Chen J, Liu K, Luo Y, Kang M, Wang J, Chen G, et al. Single-cell profiling of tumor immune microenvironment reveals immune irresponsiveness in gastric signet-ring cell carcinoma. *Gastroenterology*. (2023) 165:88–103. doi: 10.1053/j.gastro.2023.03.008

31. Kang B, Camps J, Fan B, Jiang H, Ibrahim MM, Hu X, et al. Parallel single-cell and bulk transcriptome analyses reveal key features of the gastric tumor microenvironment. *Genome Biol*. (2022) 23:265. doi: 10.1186/s13059-022-02828-2

32. Li Y, Hu X, Lin R, Zhou G, Zhao L, Zhao D, et al. Single-cell landscape reveals active cell subtypes and their interaction in the tumor microenvironment of gastric cancer. *Theranostics*. (2022) 12:3818–33. doi: 10.7150/thno.71833

33. Zhou X, Yang J, Lu Y, Ma Y, Meng Y, Li Q, et al. Relationships of tumor differentiation and immune infiltration in gastric cancers revealed by single-cell RNA-seq analyses. *Cell Mol Life sciences: CMLS*. (2023) 80:57. doi: 10.1007/s00018-023-04702-1

34. Negura I, Pavel-Tanasa M, Danciu M. Regulatory T cells in gastric cancer: Key controllers from pathogenesis to therapy. *Cancer Treat Rev*. (2023) 120:102629. doi: 10.1016/j.ctrv.2023.102629

35. Su J, Mao X, Wang L, Chen Z, Wang W, Zhao C, et al. Lactate/GPR81 recruits regulatory T cells by modulating CX3CL1 to promote immune resistance in a highly glycolytic gastric cancer. *Oncoimmunology*. (2024) 13:2320951. doi: 10.1080/2162402X.2024.2320951

36. Tsutsumi C, Ohuchida K, Katayama N, Yamada Y, Nakamura S, Okuda S, et al. Tumor-infiltrating monocytic myeloid-derived suppressor cells contribute to the development of an immunosuppressive tumor microenvironment in gastric cancer. *Gastric Cancer*. (2024) 27:248–62. doi: 10.1007/s10120-023-01456-4

37. Zhang C, Wei S, Dai S, Li X, Wang H, Zhang H, et al. The NR_109/FUBP1/c-Myc axis regulates TAM polarization and remodels the tumor microenvironment to promote cancer development. *J immunother Cancer*. (2023) 11. doi: 10.1136/jitc-2022-006230

38. Li Y, Jiang L, Chen Y, Li Y, Yuan J, Lu J, et al. Specific lineage transition of tumor-associated macrophages elicits immune evasion of ascitic tumor cells in gastric cancer with peritoneal metastasis. *Gastric Cancer*. (2024) 27:519–38. doi: 10.1007/s10120-024-01486-6

39. He Y, Hong Q, Chen S, Zhou J, Qiu S. Reprogramming tumor-associated macrophages in gastric cancer: a pathway to enhanced immunotherapy. *Front Immunol*. (2025) 16:1558091. doi: 10.3389/fimmu.2025.1558091

40. Miao Z, Li J, Wang Y, Shi M, Gu X, Zhang X, et al. Hsa_circ_0136666 stimulates gastric cancer progression and tumor immune escape by regulating the miR-375/PRKDC Axis and PD-L1 phosphorylation. *Mol cancer*. (2023) 22:205. doi: 10.1186/s12943-023-01883-y

41. Huang T, Ren X, Tang X, Wang Y, Ji R, Guo Q, et al. Current perspectives and trends of CD39-CD73-eAdo/A2aR research in tumor microenvironment: a bibliometric analysis. *Front Immunol*. (2024) 15:1427380. doi: 10.3389/fimmu.2024.1427380

42. Khan M, Lin J, Wang B, Chen C, Huang Z, Tian Y, et al. A novel necroptosis-related gene index for predicting prognosis and a cold tumor immune microenvironment in stomach adenocarcinoma. *Front Immunol*. (2022) 13:968165. doi: 10.3389/fimmu.2022.968165

43. Lv K, Sun M, Fang H, Wang J, Lin C, Liu H, et al. Targeting myeloid checkpoint Siglec-10 reactivates antitumor immunity and improves anti-programmed cell death 1 efficacy in gastric cancer. *J immunother Cancer*. (2023) 11. doi: 10.1136/jitc-2023-007669

44. Xing X, Shi J, Jia Y, Dou Y, Li Z, Dong B, et al. Effect of neoadjuvant chemotherapy on the immune microenvironment in gastric cancer as determined by multiplex immunofluorescence and T cell receptor repertoire analysis. *J immunother Cancer*. (2022) 10. doi: 10.1136/jitc-2021-003984

45. Huan X, Zou K, Zhang P, Ding H, Luo C, Xiang C, et al. Neoadjuvant chemotherapy is linked to an amended anti-tumorigenic microenvironment in gastric cancer. *Int immunopharmacol*. (2024) 127:111352. doi: 10.1016/j.intimp.2023.111352

46. Qiu Y, Lu G, Li N, Hu Y, Tan H, Jiang C. Exosome-mediated communication between gastric cancer cells and macrophages: implications for tumor microenvironment. *Front Immunol*. (2024) 15:1327281. doi: 10.3389/fimmu.2024.1327281

47. Groen-van Schooten TS, Franco Fernandez R, van Grieken NCT, Bos EN, Seidel J, Saris J, et al. Mapping the complexity and diversity of tertiary lymphoid structures in primary and peritoneal metastatic gastric cancer. *J immunother Cancer*. (2024) 12. doi: 10.1136/jitc-2024-009243

48. Zhao Q, Yu H, Shi M, Wang X, Fan Z, Wang Z. Tumor microenvironment characteristics of lipid metabolism reprogramming related to ferroptosis and EndMT influencing prognosis in gastric cancer. *Int immunopharmacol*. (2024) 137:112433. doi: 10.1016/j.intimp.2024.112433

49. Zhong M, Yu Z, Wu Q, Lu B, Sun P, Zhang X, et al. PCDHGA10 as a potential prognostic biomarker and correlated with immune infiltration in gastric cancer. *Front Immunol*. (2024) 15:1500478. doi: 10.3389/fimmu.2024.1500478

50. Lin Y, Jing X, Chen Z, Pan X, Xu D, Yu X, et al. Histone deacetylase-mediated tumor microenvironment characteristics and synergistic immunotherapy in gastric cancer. *Theranostics*. (2023) 13:4574–600. doi: 10.7150/thno.86928

51. Bos J, Groen-van Schooten TS, Brugman CP, Jamaludin FS, van Laarhoven HWM, Derks S. The tumor immune composition of mismatch repair deficient and Epstein-Barr virus-positive gastric cancer: A systematic review. *Cancer Treat Rev*. (2024) 127:102737. doi: 10.1016/j.ctrv.2024.102737

52. Li X, Wu N, Wang C, Pei B, Ma X, Xie J, et al. NALCN expression is down-regulated and associated with immune infiltration in gastric cancer. *Front Immunol*. (2025) 16:1512107. doi: 10.3389/fimmu.2025.1512107

53. Li C, Zhang L, Qiu Z, Deng W, Wang W. Key molecules of fatty acid metabolism in gastric cancer. *Biomolecules*. (2022) 12:706. doi: 10.3390/biom12050706

54. Aoki T, Kinoshita J, Munesue S, Hamabe-Horiike T, Yamaguchi T, Nakamura Y, et al. Hypoxia-induced CD36 expression in gastric cancer cells promotes peritoneal metastasis via fatty acid uptake. *Ann Surg Oncol*. (2023) 30:3125–36. doi: 10.1245/s10434-022-12465-5

55. Liu S, Zhang H, Li Y, Zhang Y, Bian Y, Zeng Y, et al. S100A4 enhances protumor macrophage polarization by control of PPAR-γ-dependent induction of fatty acid oxidation. *J immunother Cancer*. (2021) 9. doi: 10.1136/jitc-2021-002548

56. Zeng W, Li F, Jin S, Ho PC, Liu PS, Xie X. Functional polarization of tumor-associated macrophages dictated by metabolic reprogramming. *J Exp Clin Cancer research: CR*. (2023) 42:245. doi: 10.1186/s13046-023-02832-9

57. Gao J, Liang Y, Wang L. Shaping polarization of tumor-associated macrophages in cancer immunotherapy. *Front Immunol*. (2022) 13:888713. doi: 10.3389/fimmu.2022.888713

58. Chen D, Zhang X, Li Z, Zhu B. Metabolic regulatory crosstalk between tumor microenvironment and tumor-associated macrophages. *Theranostics*. (2021) 11:1016–30. doi: 10.7150/thno.51777

59. Pan Y, Yu Y, Wang X, Zhang T. Tumor-associated macrophages in tumor immunity. *Front Immunol*. (2020) 11:583084. doi: 10.3389/fimmu.2020.583084

60. Li C, Xu X, Wei S, Jiang P, Xue L, Wang J. Tumor-associated macrophages: potential therapeutic strategies and future prospects in cancer. *J immunother Cancer*. (2021) 9. doi: 10.1136/jitc-2020-001341

61. Chu X, Tian Y, Lv C. Decoding the spatiotemporal heterogeneity of tumor-associated macrophages. *Mol cancer*. (2024) 23:150. doi: 10.1186/s12943-024-02064-1

62. Nian Z, Dou Y, Shen Y, Liu J, Du X, Jiang Y, et al. Interleukin-34-orchestrated tumor-associated macrophage reprogramming is required for tumor immune escape driven by p53 inactivation. *Immunity*. (2024) 57:2344–61.e7. doi: 10.1016/j.immuni.2024.08.015

63. Liang Y, Bu Q, You W, Zhang R, Xu Z, Gan X, et al. Single-cell analysis reveals hypoxia-induced immunosuppressive microenvironment in intrahepatic cholangiocarcinoma. *Biochim Biophys Acta Mol basis disease*. (2024) 1870:167276. doi: 10.1016/j.bbadis.2024.167276

64. Shi JH, Liu LN, Song DD, Liu WW, Ling C, Wu FX, et al. TRAF3/STAT6 axis regulates macrophage polarization and tumor progression. *Cell Death differentiation*. (2023) 30:2005–16. doi: 10.1038/s41418-023-01194-1

65. Liu PS, Chen YT, Li X, Hsueh PC, Tzeng SF, Chen H, et al. CD40 signal rewires fatty acid and glutamine metabolism for stimulating macrophage anti-tumorigenic functions. *Nat Immunol*. (2023) 24:452–62. doi: 10.1038/s41590-023-01430-3

66. Xiang X, Wang J, Lu D, Xu X. Targeting tumor-associated macrophages to synergize tumor immunotherapy. *Signal transduction targeted Ther*. (2021) 6:75. doi: 10.1038/s41392-021-00484-9

67. Sun Z, Zhang L, Liu L. Reprogramming the lipid metabolism of dendritic cells in tumor immunomodulation and immunotherapy. *Biomed pharmacother = Biomed pharmacotherapie*. (2023) 167:115574. doi: 10.1016/j.biopha.2023.115574

68. Peng X, He Y, Huang J, Tao Y, Liu S. Metabolism of dendritic cells in tumor microenvironment: for immunotherapy. *Front Immunol*. (2021) 12:613492. doi: 10.3389/fimmu.2021.613492

69. Chen J, Duan Y, Che J, Zhu J. Dysfunction of dendritic cells in tumor microenvironment and immunotherapy. *Cancer Commun (London England)*. (2024) 44:1047–70. doi: 10.1002/cac2.12596

70. Hinata M, Kunita A, Abe H, Morishita Y, Sakuma K, Yamashita H, et al. Exosomes of epstein-barr virus-associated gastric carcinoma suppress dendritic cell maturation. *Microorganisms*. (2020) 8:1776. doi: 10.3390/microorganisms8111776

71. MacNabb BW, Tumuluru S, Chen X, Godfrey J, Kasal DN, Yu J, et al. Dendritic cells can prime anti-tumor CD8(+) T cell responses through major histocompatibility complex cross-dressing. *Immunity*. (2022) 55:982–97.e8. doi: 10.1016/j.immuni.2022.04.016

72. Murphy TL, Murphy KM. Dendritic cells in cancer immunology. *Cell Mol Immunol*. (2022) 19:3–13. doi: 10.1038/s41423-021-00741-5

73. Hilligan KL, Ronchese F. Antigen presentation by dendritic cells and their instruction of CD4+ T helper cell responses. *Cell Mol Immunol*. (2020) 17:587–99. doi: 10.1038/s41423-020-0465-0

74. Chudnovskiy A, Castro TBR, Nakandakari-Higa S, Cui A, Lin CH, Sade-Feldman M, et al. Proximity-dependent labeling identifies dendritic cells that drive the tumor-specific CD4(+) T cell response. *Sci Immunol*. (2024) 9:eadq8843. doi: 10.1126/sciimmunol.adq8843

75. Moreno Ayala MA, Campbell TF, Zhang C, Dahan N, Bockman A, Prakash V, et al. CXCR3 expression in regulatory T cells drives interactions with type I dendritic cells in tumors to restrict CD8(+) T cell antitumor immunity. *Immunity*. (2023) 56:1613–30.e5. doi: 10.1016/j.immuni.2023.06.003

76. Mateus D, Sebastião AI, Frasco MF, Carrascal MA, Falcão A, Gomes CM, et al. Artificial dendritic cells: A new era of promising antitumor immunotherapy. *Small (Weinheim an der Bergstrasse Germany)*. (2023) 19:e2303940. doi: 10.1002/smll.202303940

77. Zheng J, Wang M, Pang L, Wang S, Kong Y, Zhu X, et al. Identification of a novel DEC-205 binding peptide to develop dendritic cell-targeting nanovaccine for cancer immunotherapy. *J Controlled release*. (2024) 373:568–82. doi: 10.1016/j.jconrel.2024.07.056

78. Marciscano AE, Anandasabapathy N. The role of dendritic cells in cancer and anti-tumor immunity. *Semin Immunol*. (2021) 52:101481. doi: 10.1016/j.smim.2021.101481

79. Zhou J, Nie RC, Yin YX, Wang Y, Yuan SQ, Zhao ZH, et al. Genomic analysis uncovers the prognostic and immunogenetic feature of pyroptosis in gastric carcinoma: indication for immunotherapy. *Front Cell Dev Biol*. (2022) 10:906759. doi: 10.3389/fcell.2022.906759

80. Kim MJ, Kim K, Park HJ, Kim GR, Hong KH, Oh JH, et al. Deletion of PD-1 destabilizes the lineage identity and metabolic fitness of tumor-infiltrating regulatory T cells. *Nat Immunol*. (2023) 24:148–61. doi: 10.1038/s41590-022-01373-1

81. Kumar A, Das JK, Peng HY, Wang L, Ballard DJ, Ren Y, et al. Metabolic fitness of NAC1-deficient Tregs in the tumor microenvironment fuels tumor growth. *JCI Insight*. (2025) 10. doi: 10.1172/jci.insight.186000

82. Kobayashi S, Wannakul T, Sekino K, Takahashi Y, Kagawa Y, Miyazaki H, et al. Fatty acid-binding protein 5 limits the generation of Foxp3(+) regulatory T cells through regulating plasmacytoid dendritic cell function in the tumor microenvironment. *Int J cancer*. (2022) 150:152–63. doi: 10.1002/ijc.33777

83. Zi R, Zhao X, Liu L, Wang Y, Zhang R, Bian Z, et al. Metabolic-immune suppression mediated by the SIRT1-CX3CL1 axis induces functional enhancement of regulatory T cells in colorectal carcinoma. *Advanced Sci (Weinheim Baden-Wurttemberg Germany)*. (2025) 12:e2404734. doi: 10.1002/advs.202404734

84. van Hooren L, Handgraaf SM, Kloosterman DJ, Karimi E, van Mil L, Gassama AA, et al. CD103(+) regulatory T cells underlie resistance to radio-immunotherapy and impair CD8(+) T cell activation in glioblastoma. *Nat cancer*. (2023) 4:665–81. doi: 10.1038/s43018-023-00547-6

85. Yang L, Wang X, Wang S, Shen J, Li Y, Wan S, et al. Targeting lipid metabolism in regulatory T cells for enhancing cancer immunotherapy. *Biochim Biophys Acta Rev cancer*. (2025) 1880:189259. doi: 10.1016/j.bbcan.2025.189259

86. Song C, Ji Y, Wang W, Tao N. Ginger polysaccharide promotes myeloid-derived suppressor cell apoptosis by regulating lipid metabolism. *Phytother research: PTR*. (2023) 37:2894–901. doi: 10.1002/ptr.7784

87. Zhu GQ, Tang Z, Huang R, Qu WF, Fang Y, Yang R, et al. CD36(+) cancer-associated fibroblasts provide immunosuppressive microenvironment for hepatocellular carcinoma via secretion of macrophage migration inhibitory factor. *Cell discovery*. (2023) 9:25. doi: 10.1038/s41421-023-00529-z

88. Yan Y, Huang L, Liu Y, Yi M, Chu Q, Jiao D, et al. Metabolic profiles of regulatory T cells and their adaptations to the tumor microenvironment: implications for antitumor immunity. *J Hematol Oncol*. (2022) 15:104. doi: 10.1186/s13045-022-01322-3

89. Wang R, Liu Z, Fan Z, Zhan H. Lipid metabolism reprogramming of CD8(+) T cell and therapeutic implications in cancer. *Cancer letters*. (2023) 567:216267. doi: 10.1016/j.canlet.2023.216267

90. Wu D, Chen Y. Lipids for CD8(+) TILs: beneficial or harmful? *Front Immunol*. (2022) 13:1020422. doi: 10.3389/fimmu.2022.1020422

91. Chen W, Teo JMN, Yau SW, Wong MY, Lok CN, Che CM, et al. Chronic type I interferon signaling promotes lipid-peroxidation-driven terminal CD8(+) T cell exhaustion and curtails anti-PD-1 efficacy. *Cell Rep*. (2022) 41:111647. doi: 10.1016/j.celrep.2022.111647

92. Zheng M, Zhang W, Chen X, Guo H, Wu H, Xu Y, et al. The impact of lipids on the cancer-immunity cycle and strategies for modulating lipid metabolism to improve cancer immunotherapy. *Acta Pharm Sin B*. (2023) 13:1488–97. doi: 10.1016/j.apsb.2022.10.027

93. Cui MY, Yi X, Zhu DX, Wu J. The role of lipid metabolism in gastric cancer. *Front Oncol*. (2022) 12:916661. doi: 10.3389/fonc.2022.916661

94. Yang S, Sun B, Li W, Yang H, Li N, Zhang X. Fatty acid metabolism is related to the immune microenvironment changes of gastric cancer and RGS2 is a new tumor biomarker. *Front Immunol*. (2022) 13:1065927. doi: 10.3389/fimmu.2022.1065927

95. Li Y, Zeng Z. Investigating the dysregulation of genes associated with glucose and lipid metabolism in gastric cancer and their influence on immunity and prognosis. *BioFactors (Oxford England)*. (2025) 51:e2138. doi: 10.1002/biof.2138

96. Wang Z, Chen H, Sun L, Wang X, Xu Y, Tian S, et al. Uncovering the potential of APOD as a biomarker in gastric cancer: A retrospective and multi-center study. *Comput Struct Biotechnol J*. (2024) 23:1051–64. doi: 10.1016/j.csbj.2024.02.015

97. Li Y, Zheng Y, Huang J, Nie RC, Wu QN, Zuo Z, et al. CAF-macrophage crosstalk in tumour microenvironments governs the response to immune checkpoint blockade in gastric cancer peritoneal metastases. *Gut*. (2025) 74:350–63. doi: 10.1136/gutjnl-2024-333617

98. Wang W, Gao Y, Liu Y, Xia S, Xu J, Qin L, et al. Pan-cancer analysis reveals MTTP as a prognostic and immunotherapeutic biomarker in human tumors. *Front Immunol*. (2025) 16:1549965. doi: 10.3389/fimmu.2025.1549965

99. Hou W, Zhao Y, Zhu H. Predictive biomarkers for immunotherapy in gastric cancer: current status and emerging prospects. *Int J Mol Sci*. (2023) 24:15321. doi: 10.3390/ijms242015321

100. Kang YK, Morita S, Satoh T, Ryu MH, Chao Y, Kato K, et al. Exploration of predictors of benefit from nivolumab monotherapy for patients with pretreated advanced gastric and gastroesophageal junction cancer: *post hoc* subanalysis from the

ATTRACTION-2 study. *Gastric Cancer*. (2022) 25:207–17. doi: 10.1007/s10120-021-01230-4

101. Kang YK, Chen LT, Ryu MH, Oh DY, Oh SC, Chung HC, et al. Nivolumab plus chemotherapy versus placebo plus chemotherapy in patients with HER2-negative, untreated, unresectable advanced or recurrent gastric or gastro-oesophageal junction cancer (ATTRACTION-4): a randomised, multicentre, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol*. (2022) 23:234–47. doi: 10.1016/S1470-2045(21)00692-6

102. Rha SY, Oh DY, Yañez P, Bai Y, Ryu MH, Lee J, et al. Pembrolizumab plus chemotherapy versus placebo plus chemotherapy for HER2-negative advanced gastric cancer (KEYNOTE-859): a multicentre, randomised, double-blind, phase 3 trial. *Lancet Oncol*. (2023) 24:1181–95. doi: 10.1016/S1470-2045(23)00515-6

103. Gou M, Zhang Y, Wang Z, Qian N, Dai G. PD-1 inhibitor combined with albumin paclitaxel and apatinib as second-line treatment for patients with metastatic gastric cancer: a single-center, single-arm, phase II study. *Investigational New Drugs*. (2024) 42:171–8. doi: 10.1007/s10637-024-01425-3

104. Janjigian YY, Shitara K, Moehler M, Garrido M, Salman P, Shen L, et al. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial. *Lancet (London England)*. (2021) 398:27–40. doi: 10.1016/S0140-6736(21)00797-2

105. Butterfield LH, Najjar YG. Immunotherapy combination approaches: mechanisms, biomarkers and clinical observations. *Nat Rev Immunol*. (2024) 24:399–416. doi: 10.1038/s41577-023-00973-8

106. Catenacci DV, Rosales M, Chung HC, HY H, Shen L, Moehler M, et al. MAHOGANY: margetuximab combination in HER2+ unresectable/metastatic gastric/gastroesophageal junction adenocarcinoma. *Future Oncol (London England)*. (2021) 17:1155–64. doi: 10.2217/fon-2020-1007

107. Luke JJ, Bever K, Hodi FS, Taube J, Massey A, Yao D, et al. Rationale and feasibility of a rapid integral biomarker program that informs immune-oncology clinical trials: the ADVISE trial. *J immunother Cancer*. (2025) 13. doi: 10.1136/jitc-2024-011170

108. Hegewisch-Becker S, Mendez G, Chao J, Nemecek R, Feeney K, Van Cutsem E, et al. First-line nivolumab and relatlimab plus chemotherapy for gastric or gastroesophageal junction adenocarcinoma: the phase II RELATIVITY-060 study. *J Clin Oncol*. (2024) 42:2080–93. doi: 10.1200/JCO.23.01636

109. Ji Z, Wang X, Xin J, Ma L, Zuo D, Li H, et al. Multiomics reveals tumor microenvironment remodeling in locally advanced gastric and gastroesophageal junction cancer following neoadjuvant immunotherapy and chemotherapy. *J immunother Cancer*. (2024) 12. doi: 10.1136/jitc-2024-010041

110. Yang W, Zhou M, Li G, Zhou C, Wang L, Xia F, et al. Adjuvant chemoradiotherapy plus PD-1 inhibitor for pN3 gastric cancer: a randomized, multicenter, Phase III trial. *Future Oncol (London England)*. (2024) 20:3389–96. doi: 10.1080/14796694.2024.2421156

111. Janjigian YY, Kawazoe A, Bai Y, Xu J, Lonardi S, Metges JP, et al. Pembrolizumab plus trastuzumab and chemotherapy for HER2-positive gastric or gastro-oesophageal junction adenocarcinoma: interim analyses from the phase 3 KEYNOTE-811 randomised placebo-controlled trial. *Lancet (London England)*. (2023) 402:2197–208. doi: 10.1016/S0140-6736(23)02033-0

112. Bagaev A, Kotlov N, Nomie K, Svekolkin V, Gafurov A, Isaeva O, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell*. (2021) 39:845–65.e7. doi: 10.1016/j.ccell.2021.04.014

113. Zeng D, Wu J, Luo H, Li Y, Xiao J, Peng J, et al. Tumor microenvironment evaluation promotes precise checkpoint immunotherapy of advanced gastric cancer. *J immunother Cancer*. (2021) 9. doi: 10.1136/jitc-2021-002467

114. Li S, Yu W, Xie F, Luo H, Liu Z, Lv W, et al. Neoadjuvant therapy with immune checkpoint blockade, antiangiogenesis, and chemotherapy for locally advanced gastric cancer. *Nat Commun*. (2023) 14:8. doi: 10.1038/s41467-022-35431-x

115. Shitara K, Ajani JA, Moehler M, Garrido M, Gallardo C, Shen L, et al. Nivolumab plus chemotherapy or ipilimumab in gastro-oesophageal cancer. *Nature*. (2022) 603:942–8. doi: 10.1038/s41586-022-04508-4

116. Shen L, Zhang Y, Li Z, Zhang X, Gao X, Liu B, et al. First-line cadonilimab plus chemotherapy in HER2-negative advanced gastric or gastroesophageal junction adenocarcinoma: a randomized, double-blind, phase 3 trial. *Nat Med*. (2025) 31:1163–70. doi: 10.1038/s41591-024-03450-4

117. Chen L, Deng J. Role of non-coding RNA in immune microenvironment and anticancer therapy of gastric cancer. *J Mol Med (Berlin Germany)*. (2022) 100:1703–19. doi: 10.1007/s00109-022-02264-6

118. Jing Q, Yao H, Li H, Yuan C, Hu J, Zhang P, et al. A novel RNA modification prognostic signature for predicting the characteristics of the tumor microenvironment in gastric cancer. *Front Oncol*. (2023) 13:905139. doi: 10.3389/fonc.2023.905139

119. Yang P, Yang W, Wei Z, Li Y, Yang Y, Wang J. Novel targets for gastric cancer: The tumor microenvironment (TME), N6-methyladenosine (m6A), pyroptosis, autophagy, ferroptosis and cuproptosis. *Biomed pharmacother = Biomed pharmacotherapie*. (2023) 163:114883. doi: 10.1016/j.biopha.2023.114883

120. He X, Guan XY, Li Y. Clinical significance of the tumor microenvironment on immune tolerance in gastric cancer. *Front Immunol*. (2025) 16:1532605. doi: 10.3389/fimmu.2025.1532605

121. Chang J, Wu H, Wu J, Liu M, Zhang W, Hu Y, et al. Constructing a novel mitochondrial-related gene signature for evaluating the tumor immune microenvironment and predicting survival in stomach adenocarcinoma. *J Trans Med*. (2023) 21:191. doi: 10.1186/s12967-023-04033-6

122. Su C, Yu R, Hong X, Zhang P, Guo Y, Cai JC, et al. CXCR4 expressed by tumor-infiltrating B cells in gastric cancer related to survival in the tumor microenvironment: an analysis combining single-cell RNA sequencing with bulk RNA sequencing. *Int J Mol Sci*. (2023) 24:12890. doi: 10.3390/ijms241612890

123. Pang MJ, Burclaff JR, Jin R, Adkins-Threats M, Osaki LH, Han Y, et al. Gastric organoids: progress and remaining challenges. *Cell Mol Gastroenterol hepatol*. (2022) 13:19–33. doi: 10.1016/j.jcmgh.2021.09.005

124. Bin YL, Hu HS, Tian F, Wen ZH, Yang MF, Wu BH, et al. Metabolic reprogramming in gastric cancer: trojan horse effect. *Front Oncol*. (2021) 11:745209. doi: 10.3389/fonc.2021.745209

125. Chen X, Zhou B, Wang S, Jiang X, Ping Y, Xia J, et al. Intestinal metaplasia key molecules and UPP1 activation via Helicobacter pylori/NF-kB: drivers of Malignant progression in gastric cancer. *Cancer Cell Int*. (2024) 24:399. doi: 10.1186/s12935-024-03598-6

126. Shang Z, Ma Z, Wu E, Chen X, Tuo B, Li T, et al. Effect of metabolic reprogramming on the immune microenvironment in gastric cancer. *Biomed pharmacother = Biomed pharmacotherapie*. (2024) 170:116030. doi: 10.1016/j.biopha.2023.116030

127. Jiang T, Zhang J, Zhao S, Zhang M, Wei Y, Liu X, et al. MCT4: a key player influencing gastric cancer metastasis and participating in the regulation of the metastatic immune microenvironment. *J Trans Med*. (2025) 23:276. doi: 10.1186/s12967-025-06279-8

128. Liu K, Yuan S, Wang C, Zhu H. Resistance to immune checkpoint inhibitors in gastric cancer. *Front Pharmacol*. (2023) 14:1285343. doi: 10.3389/fphar.2023.1285343

129. Liu J, Yuan Q, Guo H, Guan H, Hong Z, Shang D. Deciphering drug resistance in gastric cancer: Potential mechanisms and future perspectives. *Biomed pharmacother = Biomed pharmacotherapie*. (2024) 173:116310. doi: 10.1016/j.biopha.2024.116310

130. Zhang Y, Yang Y, Chen Y, Lin W, Chen X, Liu J, et al. PD-L1: Biological mechanism, function, and immunotherapy in gastric cancer. *Front Immunol*. (2022) 13:1060497. doi: 10.3389/fimmu.2022.1060497

131. Yang T, Guo L. Advancing gastric cancer treatment: nanotechnology innovations and future prospects. *Cell Biol toxicol*. (2024) 40:101. doi: 10.1007/s10565-024-09943-9

132. Keshavjee SH, Moy RH, Reiner SL, Ryeom SW, Yoon SS. Gastric cancer and the immune system: the key to improving outcomes? *Cancers*. (2022) 14:5940. doi: 10.3390/cancers14235940

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 30 September 2025
DOI 10.3389/fimmu.2025.1628962

OPEN ACCESS

EDITED BY
Yafeng Zhu,
Sun Yat-sen Memorial Hospital, China

REVIEWED BY
Lin Chen,
Northwest University, China
Vladimir T. Todorov,
Witten/Herdecke University, Germany

*CORRESPONDENCE
Zhenhua Yang
zhenhua1124@126.com
Zufeng Wang
wangzf4552@sr.gxmu.edu.cn

RECEIVED 15 May 2025
ACCEPTED 12 September 2025
PUBLISHED 30 September 2025

CITATION
Zeng F, Yang Z and Wang Z (2025) Integrated transcriptomic and single-cell RNA-seq analysis identifies CLCNKB, KLK1 and PLEKHA4 as key gene of AKI-to-CKD progression.
Front. Immunol. 16:1628962.
doi: 10.3389/fimmu.2025.1628962

Integrated transcriptomic and single-cell RNA-seq analysis identifies CLCNKB, KLK1 and PLEKHA4 as key gene of AKI-to-CKD progression

Fanhua Zeng, Zhenhua Yang* and Zufeng Wang*

Department of Nephrology, The First Affiliated Hospital of Guangxi Medical University, Nanning, GuangXi, China

Background: Studies have demonstrated a significant connection between acute kidney injury (AKI) and chronic kidney disease (CKD). The purpose of this study was to identify biomarkers linked to the advancement of AKI and CKD, aiming to offer new targets and insights for treating and intervening in these conditions.
Methods: Initially, candidate genes were identified by overlapping the results from differential expression analyses of AKI and CKD. Biomarkers were subsequently identified using machine learning algorithms, receiver operating characteristic curve analysis, expression analysis and experimental verification. Functional enrichment, drug prediction analyses and immune cells infiltration were conducted to investigate the functional mechanisms of the identified biomarkers. Furthermore, single-cell analyses were performed to examine the trends of biomarker expression across different cell types.
Results: CLCNKB, KLK1 and PLEKHA4 were identified as biomarkers by the screening. Subsequently, enrichment analysis showed that CLCNKB was notably enriched in oxidative phosphorylation and the degradation of valine, leucine, and isoleucine in both AKI and CKD datasets. CLCNKB, KLK1 and PLEKHA4 were found to be significantly associated with multiple immune cell types. The regulatory network indicated that PLEKHA4 might play a more important role in the progression of AKI and CKD. Furthermore, it was discovered that CLCNKB, KLK1, and PLEKHA4 are commonly targeted by tetrachlorodibenzodioxin. Finally, in the single-cell data analysis, Type A intercalated cell and Collecting duct-principal cell were identified as the key cells. It was observed that the expression trends of these biomarkers were different under different differentiation states of the key cell subpopulations.
Conclusion: CLCNKB, KLK1 and PLEKHA4 were identified as biomarkers related to the development of AKI and CKD in this study, and new ideas were provided for the research on the potential mechanisms of the progression of AKI and CKD.

KEYWORDS

acute kidney injury, chronic kidney disease, biomarkers, drug prediction, single-cell RNA sequencing

1 Background

Acute kidney injury (AKI) is identified by a quick loss of kidney function happening within a 48-hour period (1). In medical environments, AKI is typically marked by a rapid increase in serum creatinine and a significant decrease in urine output, often resulting from renal tubular cell necrosis and tissue damage following acute insults such as renal ischemia-reperfusion, exposure to nephrotoxic medications, and sepsis, among other causes (2). Currently, AKI affects 10-15% of all patients in hospitals and up to 50% of those in ICUs, with its prevalence growing annually (3). Furthermore, AKI contributes to long-term chronic kidney damage and accelerates the onset of various complications, including hypertension and cardiovascular disease (4). Alternatively, chronic kidney disease (CKD) is a progressive and lasting disorder identified by the degeneration of renal units, tubular atrophy, interstitial fibrosis, glomerulosclerosis, vascular thinning, and arteriosclerosis (5). A glomerular filtration rate under 60 mL/min/1.73m^2 for a period exceeding three months defines CKD (6). In addition, the incidence rate of CKD globally is roughly 9.1% (7). Due to its high incidence, significant economic impact, and strong association with morbidity and mortality, CKD represents a major public health concern (8). The clinical management of CKD is hindered by several limitations (9). Therefore, novel insights into the mechanism of AKI and CKD are urgently required to enhance CKD treatment strategies.

Clinically, AKI and CKD are closely interrelated. Atrophy of the tubules and fibrosis in the interstitial area are pathological changes that arise due to inadequate repair mechanisms following AKI, ultimately leading to the development of CKD (10). In China, there are at least 3 million cases of AKI annually, with approximately 50% of survivors subsequently developing CKD (11). Furthermore, individuals with CKD have a higher chance of developing AKI due to pre-existing renal lesions (12). The transition from AKI to CKD is thought to be significantly influenced by the immune-inflammatory response and kidney fibrosis, both of which contribute to persistent renal damage (13). Despite this, the transition from AKI to CKD remains largely unexplored, with the key genes and pathways involved in this intermediary process not yet clearly identified. Hence, it is essential to pinpoint biomarkers related to the progression from AKI to CKD to uncover possible therapeutic targets.

This study utilized transcriptomic and single-cell datasets from public repositories related to AKI and CKD to evaluate biomarkers associated with the progression of these conditions. The assessment was conducted through differential expression analysis, machine learning algorithms, Receiver Operating Characteristic (ROC) curve evaluation, and expression validation. Subsequently, the potential mechanisms of action of these biomarkers in AKI and CKD were explored using biomarker enrichment analysis, immune infiltration analysis, molecular regulatory network construction, and drug prediction. The single-cell data enabled the examination of intercellular communication, leading to the identification of key cellular entities and additional experiments were undertaken to verify the results. Furthermore, we investigated the expression trends of biomarkers in specific cell subsets to elucidate the molecular roles of these biomarkers and their mechanisms in the progression of AKI and CKD. This research aims to offer new perspectives for the early detection and personalized treatment of AKI and CKD patients, thereby reducing the societal burden of kidney diseases.

2 Materials and methods

2.1 Data source

Datasets related to both AKI and CKD were sourced from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/). GSE139061 (GPL20301) consisted of 39 renal tissue samples from AKI patients and 9 normal renal tissue samples, while GSE30718 (GPL570) included samples included 28 from AKI patients and 11 from healthy kidney tissues, functioning respectively as the training set and validation set of AKI. Similarly, GSE66494 (GPL6480) consisted of 53 renal tissue samples from CKD patients and 8 normal renal tissue samples, while GSE104948 (GPL22945) included 50 CKD patients' renal tissue samples and 18 normal renal tissue samples, serving respectively as the training set and validation set of CKD. Furthermore, GSE183277 (GPL24676) comprised single-cell RNA sequencing (scRNA-seq) data from kidney cortex tissue samples of 5 AKI patients, 2 CKD patients and 11 normal individuals.

2.2 Differential expression analysis

Differentially expressed genes1 (DEGs1) between AKI and normal samples in the GSE139061 dataset were pinpointed by employing the DEseq2 (v 1.38.0) package (14).The dataset was normalized using the estimateSizeFactors function, and genes with counts ≤ 1 were filtered out. DEGs1 were selected with the thresholds of $|\log_2$fold-change (FC)$| > 1.5$ and $P < 0.05$, and the false discovery rate (FDR) was applied to control for multiple comparisons. For the GSE66494 dataset, differential expression analysis between CKD and control samples was performed using the limma package (v3.44.3) (15). Genes with missing values were removed using the na.omit() function. DEGs2 were identified with the same thresholds of $\log_2$FC > 1.5 and $P < 0.05$, and FDR correction was applied. Subsequently, DEGs1 and DEGs2 were visualized as volcano plots and heatmaps, displaying only the top 10 in descending order of log2FC for both up- and down-regulated genes. The visualizations were generated through the ggplot2 (v 3.3.2) package (16) and the pheatmap (v 0.7.7) package (17), respectively.

2.3 Identification and functional analysis of candidate genes

To identify candidate genes in AKI and CKD progression, the up-regulated genes from DEGs1 and DEGs2, as well as the down-regulated genes from DEGs1 and DEGs2 were separately overlapped using (v 1.7.3) ggvenn package (18). Subsequently, the clusterProfiler package (version 3.16.0) was used to perform Gene Ontology (GO) and Kyoto Encyclopedia of the Genome (KEGG) enrichment analyses on the candidate genes (19). The Benjamini-Hochberg (BH) method was applied to control the FDR, with a significance threshold of pvalueCutoff = 0.05. The top 10 most significantly enriched terms (ranked in ascending order of p-value) from the GO and KEGG analyses were visualized using the enrichplot package (v 1.14.2) (20). To explore protein interactions among the candidate genes, the protein-protein interaction (PPI) network (interaction score > 0.15) was constructed using the Searching for Interacting Genes (STRING, https://www.string-db.org) database and the results were visualized using Cytoscape (v 3.10.2) software (21).

2.4 Biomarkers identification and expression analysis

The glmnet (v 4.1.4) package was used to apply the least absolute shrinkage and selection operator (LASSO) method to the candidate genes in the GSE139061 and GSE66494 datasets (22). The parameter family was set as binomial, and 10-fold cross-validation (nfolds = 10) was performed to determine the optimal lambda (λ) value. Potential feature genes were screened based on the lambda.min value for each dataset. Moreover, feature genes were obtained by overlapping the potential feature genes obtained from the GSE139061 and GSE66494 datasets, respectively. Immediately, to evaluate the potential of the feature genes to distinguish AKI samples from control samples, and CKD samples from control samples, these feature genes were subjected to ROC curve analysis using pROC (v 1.18.0) package (23) in the AKI training set and the AKI validation set, the CKD training set and the CKD validation set, respectively, and feature genes with area under the curve (AUC) >0.7 in all four datasets were named as candidate biomarkers. Simultaneously, the candidate biomarkers were subjected to gene expression analysis in the AKI training set and AKI validation set, the CKD training set and the CKD validation set, respectively, and the candidate biomarkers showing a notable difference ($P<0.05$) between the disease samples and the control samples in the four datasets and a consistent expression trend were selected as the biomarkers for the subsequent analyses.

2.5 Gene set enrichment analysis

To investigate the biological roles of biomarkers involved in AKI and CKD, GSEA was performed in the GSE139061 and GSE66494 datasets, respectively. For the analysis, the c2.cp.kegg.v2023.1.Hs.symbols.gmt gene set was acquired from the Molecular Signatures Database (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb/) to act as the background set. First, Spearman correlations between the biomarkers and other genes were calculated using the psych (v 2.2.5) package (24) in the GSE139061 dataset. Subsequently, GSEA for each biomarker was constructed using clusterProfiler (v 3.16.0) package, with significance determined at $P <0.05$ and |normalized enrichment score (NES)| > 1. The top five pathways in descending order of P-value were visualized using the enrichplot package (v 1.14.2). Similarly, GSEA of the biomarkers was carried out using the same methods and thresholds in the GSE66494 dataset.

2.6 Analysis of immune infiltration and cytokines expression

To assess the infiltration of 64 immune cells in disease samples and control samples in the GSE139061 and GSE66494 datasets, respectively. In the case of the GSE139061 dataset, relative abundance was calculated using the xCell (v 1.1.0) package (25), and the proportionate distribution of the 64 immune cells of the AKI samples versus the control samples was visualized using the ggplot2 (v 3.3.2) package. Differences in infiltration scores between AKI samples and control samples in the GSE139061 dataset were then assessed using Wilcoxon test to screen for immune cell types with a significant difference in infiltration ($P<0.05$), which were named differential immune cells. Subsequently, Spearman correlation analysis was performed using corrplot (v 0.92) package (26) to explore the relationship between differential immune cells and the association between diverse immune cells and biomarkers($|cor| > 0.30$, $P < 0.05$), and correlation heatmaps were plotted to show the results. In addition, immune infiltration and correlation analyses were carried out in the GSE66494 dataset with the same methods and thresholds.

2.7 Construction of regulatory networks and drug prediction

Biomarkers targeted by miRNAs were forecasted using the TargetScan (http://www.targetscan.org/) and miRDB (http://mirdb.org/) databases. The transcription factors (TFs) that regulate biomarkers were predicted through the ChEA3 (https://maayanlab.cloud/chea3/) database. Then, the lncRNAs targeting the aforementioned miRNAs were predicted by means of the LncBase (http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=lncba) database. The miRNA-mRNA, TF-mRNA and TF-mRNA-miRNA networks were visualized by using the Cytoscape (v 3.10.2) software, and using the ggplot2 package (v 3.3.2), the lncRNA-miRNA-mRNA network was visualized. Additionally, the Comparative Toxicogenomics Database (CTD, http://ctdbase.org/) was employed to predict drugs targeting biomarkers and Cytoscape version 3.10.2 was employed to plot the biomarker-drug network.

2.8 scRNA-seq analysis

Firstly, 5 AKI samples and 11 control samples were selected from the GSE183277 dataset as the AKI single-cell dataset. The AKI single-cell dataset was the "Seurat" package (v 4.1.0) was utilized for quality control (QC) to filter out cells with exceeded 20% of mitochondrial genes, cells with nCount_RNA under 200 and surpassed 30,000 genes, and cells with nFeature_RNA > 200 (27). Then, in light of the GSE183277 dataset, data were normalized by the "NormalizeData" function in the "Seurat" package (v 4.1.0), and highly variable genes (HVGs) were selected by the "FindVariableFeatures" function. Next, the "ScaleData" function in the "Seurat" package (v 4.1.0) was applied to scale data before principal components analysis (PCA). Subsequently, the "JackStraw" function within the "Seurat" package (v 5.0.1) was applied to execute PCA on HVGs. The "ElbowPlot" function within the "Seurat" package (v 4.1.0) was thereafter applied to draw a scree plot of the top 30 principal components (PCs), aiming to identify PCs that notably contributed to variation for subsequent analysis ($p < 0.05$). Afterward, cell cluster analysis was conducted on cells after dimensionality reduction utilizing "FindNeighbors" and "FindClusters" functions (resolution = 0.2, dimension = 30). Finally, the Seurat package's FindNeighbors and FindClusters functions were employed to categorize all high-quality cells into various cell clusters using the uniform manifold approximation and projection (UMAP) clustering technique.The FindAllMarkers function was used to identify key marker genes for various populations, and the classical marker genes of relevant cells in the CellMarker (http://xteam.xbio.top/CellMarker/) database were used as the reference gene set to annotate each cell cluster (Supplementary Table 1). Additionally, 2 CKD samples and 11 control samples were selected from the GSE183277 dataset as the CKD single-cell dataset and analyzed by scRNA-seq in the same way, with marker genes shown in Supplementary Table 2.

2.9 Cell communication analysis and identification of key cells

Cellular communication networks between cell types of AKI samples and control samples as well as those between cell types of CKD samples and control samples were analyzed respectively using the CellChat (v 1.6.1) package (28) based on the AKI single-cell dataset and the CKD single-cell dataset. And visualization was carried out by using the patchwork (v 1.3.0) package (29). In addition, key cells were screened and obtained based on the expression situation of biomarkers in cell types within the 2 single-cell datasets.

2.10 Pseudotime analysis

To explore the expression changes of biomarkers during the process of cell state transformation, key cells were first extracted respectively based on the AKI single-cell dataset and the CKD single-cell dataset for secondary dimensionality reduction and clustering, and the key cells were reclustered and divided into different cell subpopulations. Following this, the Monocle (v 2.30.0) package was used to conduct cell pseudo-time trajectory analysis on both the AKI and CKD single-cell datasets (30).

2.11 Mice models

In this study, male C57BL/6J mice, approximately 8 weeks of age, were utilized. The strain was sourced from the University Model Animal Research Center at Guangxi Medical University. Ethical approval for the use of animals in this research was obtained in compliance with the Guidelines for the Management of Laboratory Animals as stipulated by the Ministry of Science and Technology of the People's Republic of China, as well as the Guidelines for Ethical Review of Laboratory Animals according to the National Standard GB/T35892–2018 of the People's Republic of China, and the protocols of the Animal Care and Welfare Committee at Guangxi Medical University (No:202506002). The mice were provided with food and water ad libitum, and the housing environment was maintained at a temperature of 25 ± 2°C with a 12-hour light/dark cycle. The experimental design included three groups of mice, with the model being established through renal artery ischemia-reperfusion surgery. For the intervention study, the C57BL/6J mice were divided into three groups (n = 5 or 6 per group): (1) normal control group; (2) AKI group; and (3) CKD group. Ischemic AKI was experimentally induced using a bilateral ischemia-reperfusion injury (BIRI) model. In this model, mice were anesthetized, and bilateral dorsal incisions were performed to access the kidneys. Both kidneys were then clamped to occlude blood flow for a duration of 30 minutes. CKD was simulated through a unilateral ischemia-reperfusion procedure combined with a contralateral total nephrectomy. Following anesthesia, a left dorsal incision was made to clamp the left kidney, obstructing blood flow for 30 minutes. Subsequently, 14 days post-procedure, a right dorsal incision was executed to facilitate the complete removal of the right kidney (31).

2.12 Immunohistochemistry

Kidney tissues were paraffin-embedded and sectioned into 4 μm slices. After deparaffinization and rehydration, antigen retrieval was conducted with EDTA buffer at pH 9.0 for 25 minutes. A 15-minute treatment with 10% hydrogen peroxide was used to block endogenous peroxidase activity, and secondary antibodies were blocked with 5% serum for 30 minutes at room temperature.The kidney tissues underwent overnight incubation at 4°C with primary antibodies (PLEKHA4,BD-PB3919, 1:300, Biodragon, Jiangsu, China; KLKI,YP-AB-02871, 1:200, UpingBio, Zhejiang, China; CLCNKB, DF9376, 1:150, Biodragon, Jiangsu, China) targeting the candidate biomarkers. Horseradish peroxidase (HRP)-conjugated antibodies were applied to the sections on the subsequent day. 3,3'-diaminobenzidine (DAB) (G1212-200T, Servicebio, Wuhan, China), a substrate specific to HRP, was used

to highlight the stained areas in kidney tissue. Subsequently, counterstaining is performed using hematoxylin (G1004-100ML, Servicebio, Wuhan, China). Representative images were captured using an Olympus microscope, and ImageJ (NIH, USA) was employed to quantify the average optical density of the images to assess the expression levels of candidate biomarkers.

2.13 Immunofluorescent staining

Immunofluorescence staining was conducted on 5 μm-thick paraffin-embedded sections of mice kidney tissue. Following deparaffinization and antigen retrieval using EDTA (pH 9.0), the sections were blocked with goat serum and incubated overnight at 4°C with primary antibodies targeting SLC4A1 (A17391, 1:150, ABclonal, Wuhan, China) and CA II (EM1801-08, 1:150, HuaAn, Zhejiang, China). Subsequently, the sections were treated with iFluor™ 647-conjugated goat anti-rabbit IgG and iFluor™ 488-conjugated goat anti-mouse IgG (HA1125 and HA1123, 1:300, HuaAn) for one hour at room temperature. Nuclei were counterstained with DAPI, and imaging was performed using a Zeiss Axio-Imager A2 confocal microscope (Carl Zeiss, Jena, Germany).

2.14 Reverse-transcription polymerase chain reaction

In summary, total RNA was extracted from renal tissues using the Trizol method (15596026, Invitrogen, USA). Equivalent amounts of mRNA were reverse transcribed into cDNA utilizing the HiScript RT SuperMix kit (R122-01; Vazyme, China). Quantitative real-time PCR (qRT-PCR) was conducted with the ChamQ Universal SYBR qPCR Master Mix (Q711-02; Vazyme, China) on a Viia 7 quantitative real-time PCR instrument (Thermo-Fisher Scientific, USA). The PCR amplification protocol consisted of 35 cycles at 95°C for 30 seconds, 58°C for 30 seconds, and 72°C for 30 seconds. The following primers were employed: Aqp6 forward: GCCGTCATTGTTGGGAAGTTC and reverse: GGCT CCAGGTCTACCACTTTC; Kit forward: GAACAGGACCTC GGCTAACAA and reverse: CCTTTGCTCTGCTCCTGTACA; Slc4a1 forward: CCTCGTCCAATACATCTCCCG and reverse: CGTCATGGCAAGTAGGAAGGT. RT-PCR products were separated on a 1.5% agarose gel and visualized under UV light. The quantification of qRT-PCR was performed using the $2^{-\Delta\Delta Ct}$ method and expressed as relative fold changes.

2.15 Patient samples

To investigate the expression of candidate biomarkers in patients with AKI and CKD), we selected a cohort comprising five patients with AKI and five with CKD. Additionally, we included five patients diagnosed with renal malignancy, from whom normal renal tissue adjacent to the tumor was obtained during surgical procedures. Patients were identified as having AKI if they fulfilled any of these conditions: (a) a rise in serum creatinine (Scr) exceeding 26.5 μmol/L within 48 hours; (b) a 50% increase in Scr over the course of one week; or (c) urine output below 0.5 mL per kilogram per hour lasting over 6 hours. CKD patients were recognized by an estimated glomerular filtration rate (eGFR) under 60 mL/min/1.73 m^2.Approval for this study was granted by the ethics committee of The First Affiliated Hospital of Guangxi Medical University, with informed consent obtained from the patients (No. 2024-E0918).The data analysis design of this study was showed in Figure 1.

2.16 Statistical analysis

Bioinformatic analysis were performed in the R (v 4.2.2). GraphPad Prism statistical software was used for experiment statistical analyses in the study. We employed the unpaired t test to compare continuous variables between two groups. Values are shown as mean ± SEM, with statistical significance set at $P<0.05$.

3 Results

3.1 Candidate genes were ascertained

In the GSE139061 dataset, 1153 differentially expressed genes (DEGs1) were screened out, among which 913 were up-regulated and 240 were down-regulated. Similarly, in the GSE66494 dataset, 153 differentially expressed genes (DEGs2) were screened out, with 60 being up-regulated and 93 being down-regulated. The top 10 up- and down-regulated DEGs in both datasets and their expression profiles were labeled on the volcano plots and heatmaps respectively (Figures 2A–D). Subsequently, by overlapping the 913 up-regulated DEGs1 with the 60 up-regulated DEGs2, 6 common up-regulated genes were identified (Figure 2E). And by overlapping the 240 down-regulated DEGs1 with the 93 down-regulated DEGs2, 13 common down-regulated genes were obtained (Figure 2F). The 6 common up-regulated genes and the 13 common down-regulated genes were combined, and 19 candidate genes were determined. In conclusion, this analysis focused on the discovery of candidate genes that might play important roles in the progression of AKI and CKD.

3.2 Function and pathways of candidate genes were explored

Enrichment analyses of the 19 candidate genes showed that they were enriched in 22 GO entries, such as organic anion transport (Figure 2G; Supplementary Table 3), whereas KEGG analyses revealed that the candidate genes were significantly enriched in the Renin-angiotensin system (Figure 2H; Supplementary Table 4). In addition, in the constructed PPI network, genes such as ALB, SLC22A6 and SLC12A3 were highly associated with other genes (Figure 2I).

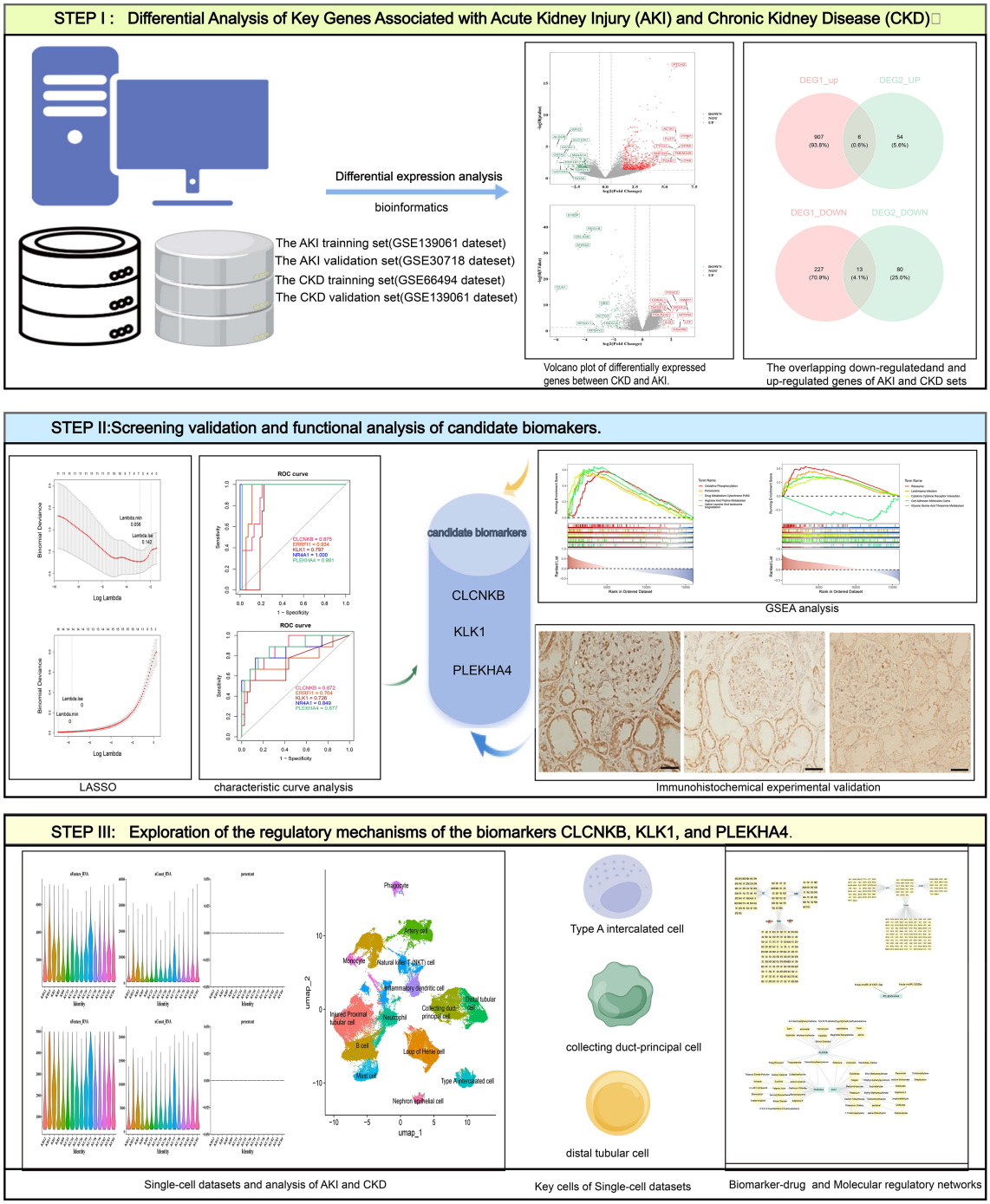


FIGURE 1
The overall workflow of this study. AKI, Acute kidney injury; CKD, Chronic kidney disease; GSEA, Gene set enrichment analysis.

3.3 CLCNKB, KLK1, and PLEKHA4 were deemed as biomarkers

Based on the candidate genes, 7 potential feature genes in the AKI training set and 16 potential feature genes in the CKD training set were obtained respectively through the LASSO regression analysis (Figures 3A, B). Then, 5 feature genes were finally obtained by overlapping (Figure 3C). Subsequently, it was found in the AKI training set, validation set as well as the CKD training set and validation set that the AUC values of CLCNKB, KLK1 and PLEKHA4 were all greater than 0.7, and thus they could be regarded as the candidate biomarkers for this study (Figures 3D, E). Moreover, the expression analysis of the candidate biomarkers showed that the expression trends of CLCNKB, KLK1 and PLEKHA4 were consistent in the four datasets. Among them, CLCNKB and KLK1 were significantly down-regulated in AKI

10.3389/fimmu.2025.1628962

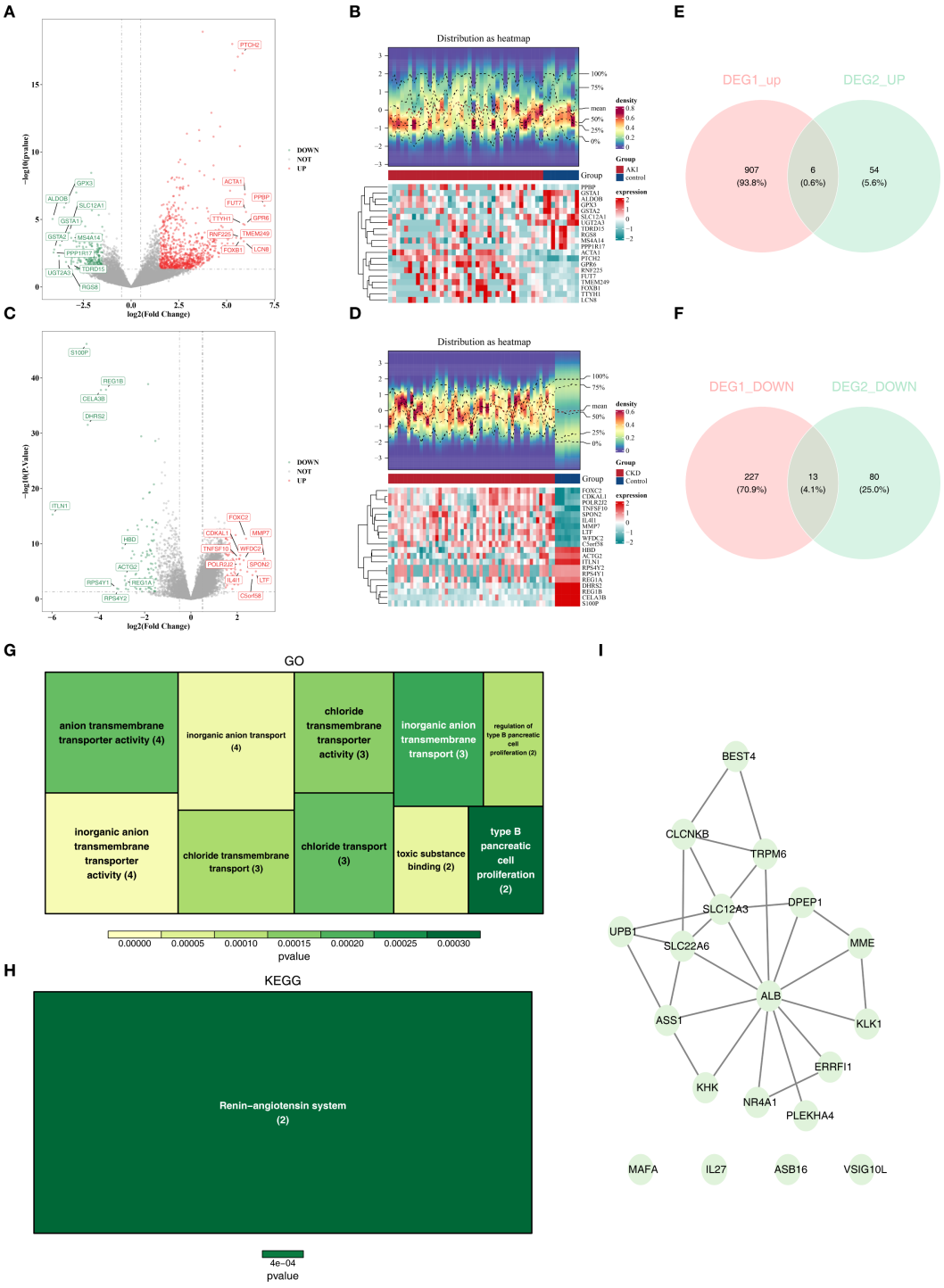


FIGURE 2
The different expressed genes(DEGs) of AKI and CKD sets and function and pathways of candidate genes. **(A, B)** The volcano plots of DEGs in AKI set. **(B)** The heatmaps of DEGs in CKD set. **(C)** The volcano plots of DEGs in AKI set. **(D)** The heatmaps of DEGs in CKD set. **(E)** The overlapping up-regulated genes of AKI and CKD sets. **(F)** The overlapping down-regulated genes of AKI and CKD sets. **(G)** GO analysis of candidate genes. **(H)** KEGG analysis of candidate genes. **(I)** PPI network of candidate genes. AKI, Acute kidney injury; CKD, Chronic kidney disease; DEGs, Differentially expressed genes1; GO, Gene Ontology (GO); KEGG, Kyoto Encyclopedia of the Genome; PPI, protein-protein interaction.

10.3389/fimmu.2025.1628962

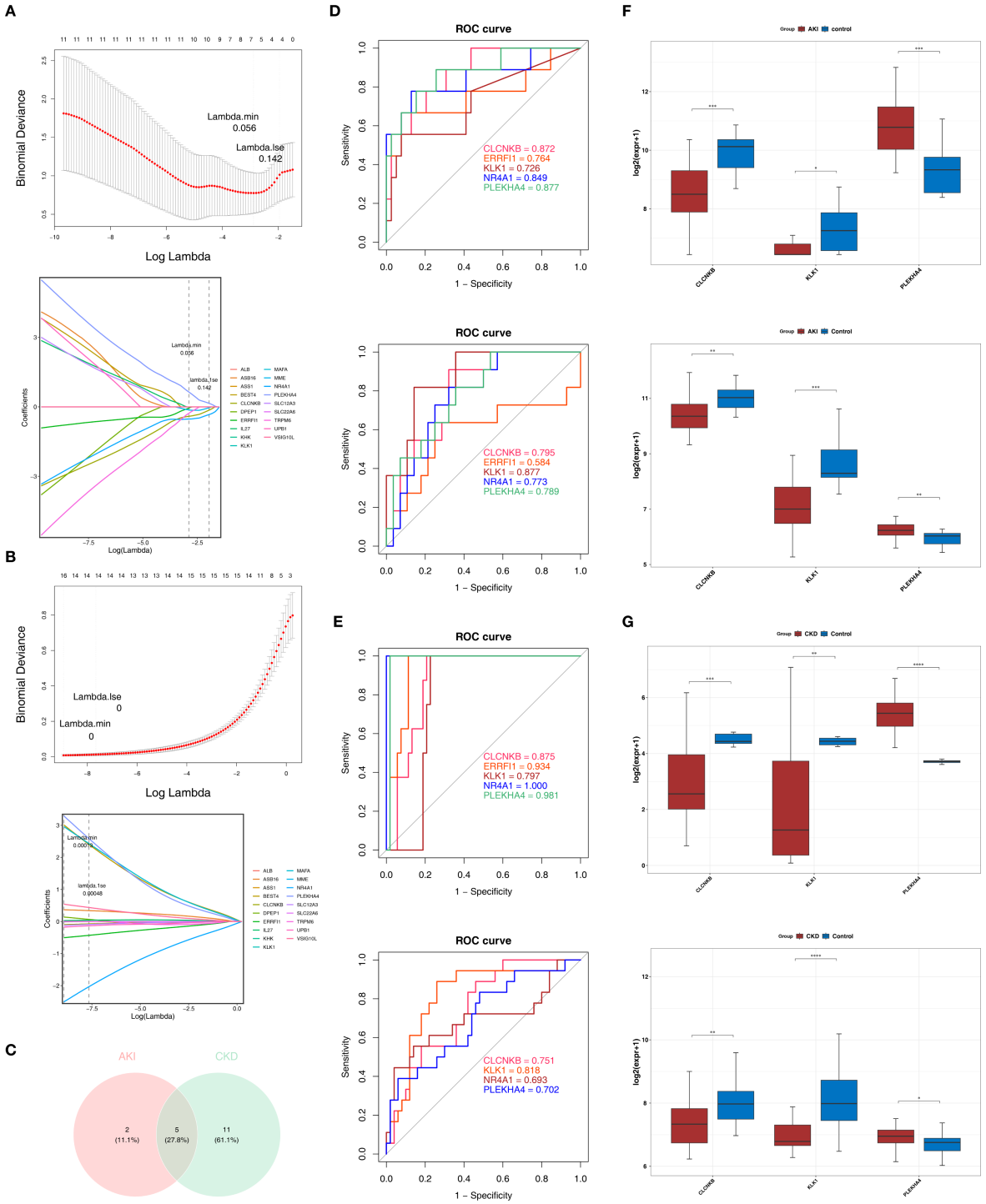


FIGURE 3
CLCNKB, KLK1 and PLEKHA4 were regarded as the candidate biomarkers and the GSEA analysis of candidate biomarkers. **(A, B)** LASSO regression analysis of AKI and CKD training sets. **(C)** Overlapping genes of AKI and CKD training sets. **(D, E)** Receiver operating characteristic curve analysis of AKI and CKD training sets and validation sets. **(F, G)** Expression trends of CLCNKB, KLK1 and PLEKHA4 in AKI and CKD training sets and validation sets. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. AKI, Acute kidney injury; CKD, Chronic kidney disease; LASSO, Least absolute shrinkage and selection operator.

and CKD samples, while PLEKHA4 was significantly up-regulated (Figures 3F, G). Moreover, the expression trends of CLCNKB, KLK1 and PLEKHA4 in renal tissues of different groups of patients and different groups of mice models were consistent with our results (Figure 4).

3.4 Functional analysis helps explore potential mechanisms of AKI and CKD progression

GSEA was performed on the GSE139061 and GSE6649 datasets to investigate the biological roles of the biomarkers. In the GSE139061 dataset of AKI, CLCNKB, KLK1 and PLEKHA4 were significantly enriched in 50, 15 and 51 pathways respectively (Supplementary Tables 5–7). It was worth noting that CLCNKB and PLEKHA4 were co-enriched in the valine leucine and isoleucine degradation pathways and oxidative phosphorylation, and KLK1 and PLEKHA4 were co-enriched in the neuroactive ligand receptor interaction pathway (Figures 5A–C). In the GSE6649 dataset of CKD, CLCNKB, KLK1 and PLEKHA4 were significantly enriched in 39, 60 and 44 pathways respectively (Supplementary Tables 8–10). Among them, CLCNKB and KLK1 were jointly enriched in the oxidative phosphorylation and valine, leucine, and isoleucine degradation pathways, as well as in the cytokine-cytokine receptor interaction pathway (Figures 5D–F).

3.5 GYG1 and PPP1R3D were associated with immune infiltrating cells

Figures 6A, B illustrated the infiltration levels of 64 immune cells in AKI versus control samples, and CKD versus control samples, respectively. Among them, the infiltration levels of 6 types of immune cells (differential immune cells 1) were significantly different in AKI and control samples (Figure 6C), and the infiltration levels of 26 types of immune cells (differential immune cells 2) were significantly different in CKD and control samples (Figure 6D), and the common differential immune cells included Astrocytes, Th2 cells. Furthermore, among the differential immune cells 1 in AKI, Fibroblasts had the most significant positive correlation with aDC (cor = 0.32), and Astrocytes had the most significant negative correlation with Fibroblasts (cor = -0.36) (Figure 6E; Supplementary Table 11). Whereas PLEKHA4 had the strongest positive relationship with pDC (cor = 0.58) and the strongest negative relationship with Astrocytes (cor = -0.38), CLCNKB had the strongest positive relationship with Astrocytes (cor = 0.30) and the strongest negative relationship with Th2 cells (cor = -0.49), but KLK1 was significantly correlated with Differential Immune Cells 1 (Figure 6F; Supplementary Table 12). Subsequently, among the differential immune cells 2 in CKD, cDC had the most significant positive correlation with DC (cor = 0.75) and the highest positive association with Macrophages M2 and Neurons (cor = -0.74) (Figure 6G; Supplementary Table 13). In contrast, PLEKHA4 had the greatest positive connection with Th2 cells (cor = 0.58) and the greatest negative linkage with MEP (cor = -0.40), CLCNKB had the most significant positive correlation with Th1 cells (cor = 0.79) and the strongest inverse relationship with NKT (cor = -0.67), and KLK1 had the greatest positive connection with MEP (cor = 0.62) and the most prominent negative connection with Th2 cells (cor = -0.43) (Figure 6H; Supplementary Table 14).

3.6 Molecular regulatory networks probe regulatory mechanisms of biomarkers

Initially, merely 2 miRNAs were predicted for PLEKHA4, whereas no miRNAs could be predicted for CLCNKB and KLK1 (Figure 7A). Subsequently, the TF-mRNA networks consisting of 133, 34 and 56 TFs corresponding to PLEKHA4, CLCNKB and KLK1 respectively were acquired from the ChEA3 database (Figure 7B). Then, a miRNA-mRNA-TF network was established by integrating the 2 miRNAs (Figure 7C). Eventually, 8 lncRNAs upstream of miRNAs were predicted and a lncRNA-miRNA-mRNA network was constructe, such as EBLN3P-hsa-miR-3187-3p-PLEKHA4 (Figure 7D). In a nutshell, this analysis centered around the finding that PLEKHA4 was likely to play a more crucial role in the progression of AKI and CKD.

3.7 CLCNKB, PLEKHA4 and KLK1 were simultaneously targeted by Tetrachlorodibenzodioxin

Drugs were screened for activation of CLCNKB and KLK1, which are down-regulated in expression, and inhibition of PLEKHA4, which is up-regulated in expression, including 27 drugs targeting KLK1, 19 drugs targeting CLCNKB and 19 drugs targeting PLEKHA4 (Supplementary Tables 15–17). A biomarker-drug network was constructed accordingly (Figure 7E). It was noteworthy that CLCNKB, PLEKHA4 and KLK1 were simultaneously targeted by Tetrachlorodibenzodioxin.

3.8 Annotation in AKI and CKD yielded 14 and 13 cell types, respectively

In the AKI single-cell dataset, a total of 78,791 cells were retained after quality control (Supplementary Figure S1). Subsequently, the top 2,000 highly variable genes and the top 30 PCs were applied to UMAP clustering (Figures 8A, B). All high-quality cells were divided into 17 different cell clusters (Figure 8C). In addition, marker genes had high specificity in different cell clusters (Figures 8D, E). The cell clusters were annotated and 14 cell types were determined, such as Injured Proximal tubular cell and Loop of Henle cell (Figure 8F). Subsequently, in the CKD single-cell dataset, a total of 58,561 cells were retained after quality control (Supplementary Figure S2). Next, the top 2,000 highly

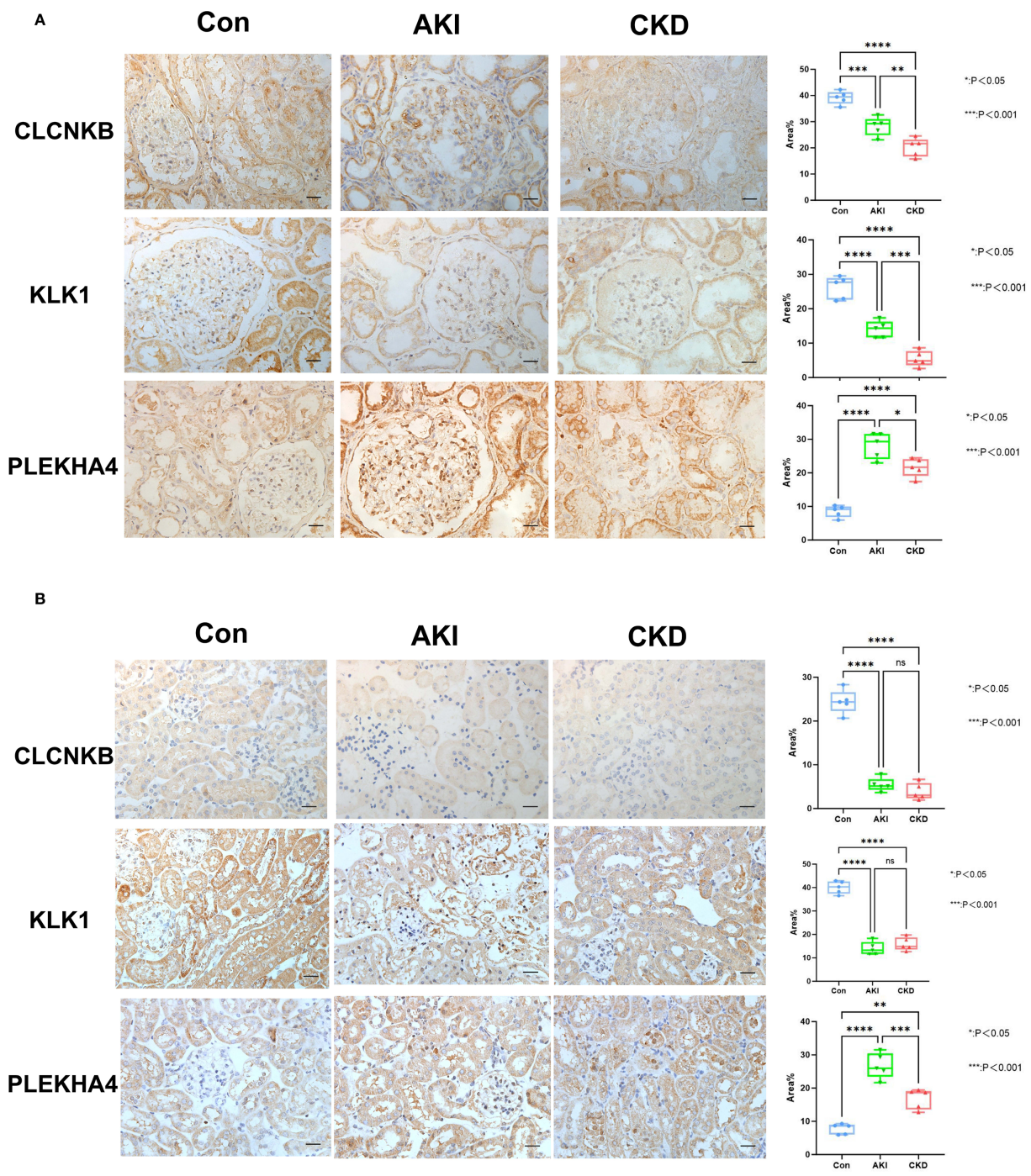


FIGURE 4
The verification of CLCNKB, KLK1 and PLEKHA4 in AKI and CKD samples. **(A)** The expression level of CLCNKB, KLK1 and PLEKHA4 in the kidneys of patients with AKI and CKD. **(B)** The expression level of CLCNKB, KLK1 and PLEKHA4 in the kidneys of AKI and CKD mice models. ns means not significant, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.Bar:50um.

variable genes and the top 30 PCs were applied to UMAP clustering (Figures 9A, B). All high-quality cells were divided into 16 different cell clusters (Figure 9C). Moreover, marker genes also had high specificity in different cell clusters (Figures 9D, E). The cell clusters were annotated and 13 cell types were determined, such as Nephron epithelial cell and Loop of Henle cell (Figure 9F).

3.9 Type A intercalated cell and collecting duct-principal cell identified as key cells

In the AKI single-cell dataset, the injured proximal tubular cells in AKI samples had a relatively large number of interactions and a relatively high intensity with other cells (Figure 10A), while

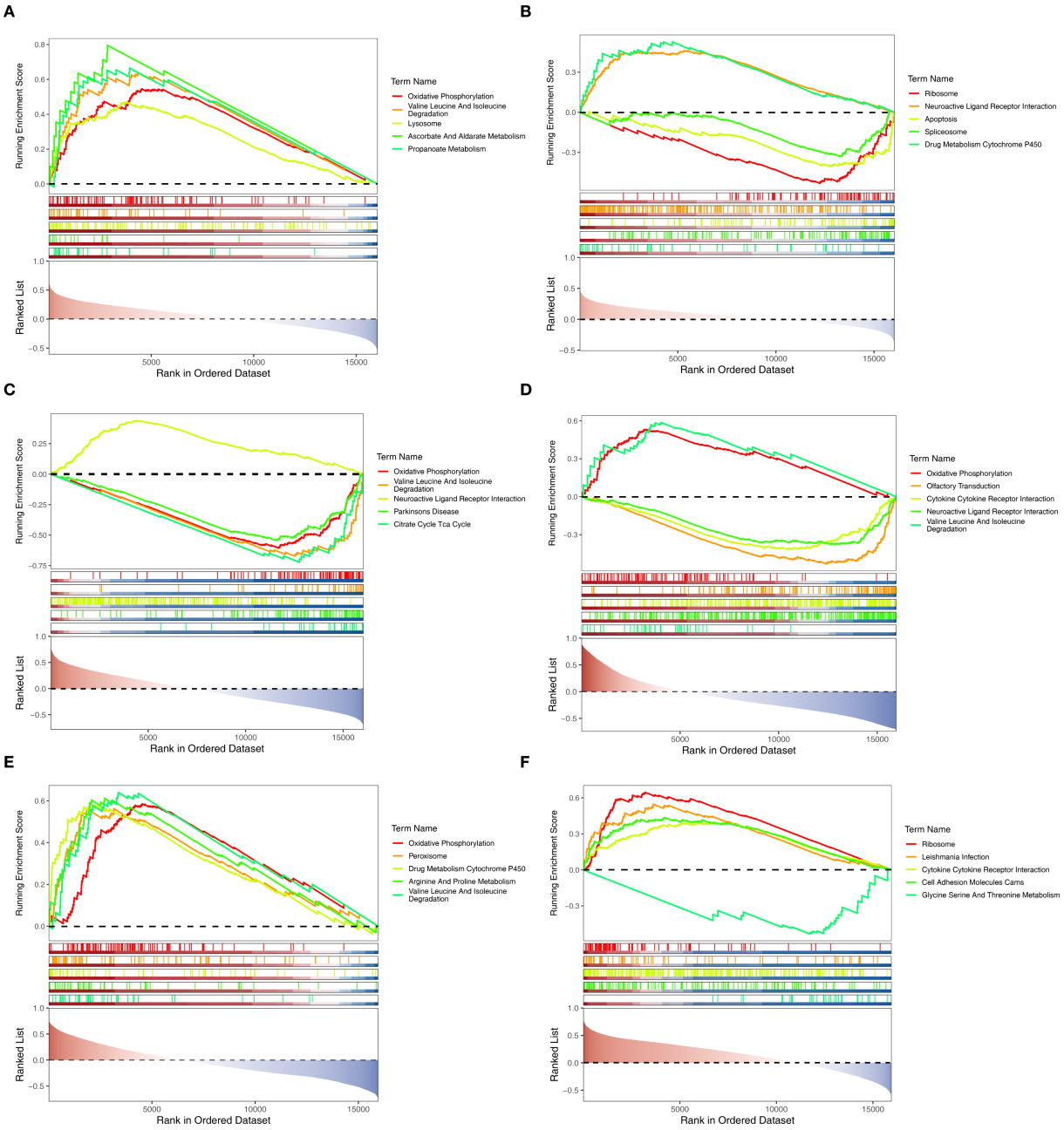


FIGURE 5
GSEA analysis of biomarkers. **(A–C)** GSEA analysis of CLCNKB, KLK1 and PLEKHA4 in AKI set. **(D–F)** GSEA analysis of CLCNKB, KLK1 and PLEKHA4 in CKD set. GSEA, Gene set enrichment analysis.

in the control samples, B cells had a relatively large number of interactions and a relatively high intensity with other cells (Figure 10B). Interestingly, in the CKD single-cell dataset, Nephron epithelial cells and B cells had a relatively large number of interactions and a relatively high intensity with other cells both in CKD and control samples (Figures 10C, D). In addition, KLK1 and CLCNKB had relatively high expression levels in Type A intercalated cells and Collecting duct-principal cells in both single-cell datasets (Figures 10E, F). To evaluate the abundance of type A intercalated cells in kidney disease, immunofluorescence staining was performed on kidney sections of AKI/CKD mice. Consistent with the reduced expression of SLC4A1 (a specific marker for this cell type) in diseased kidneys (Figure 11A), the expressions of additional markers (Aqp6, Kit and Slc4a1) were also significantly downregulated (Figures 11B–D), confirming the loss of type A intercalated cells in AKI/CKD. Consequently, Type A intercalated cells could have been part of the disease's development.

10.3389/fimmu.2025.1628962

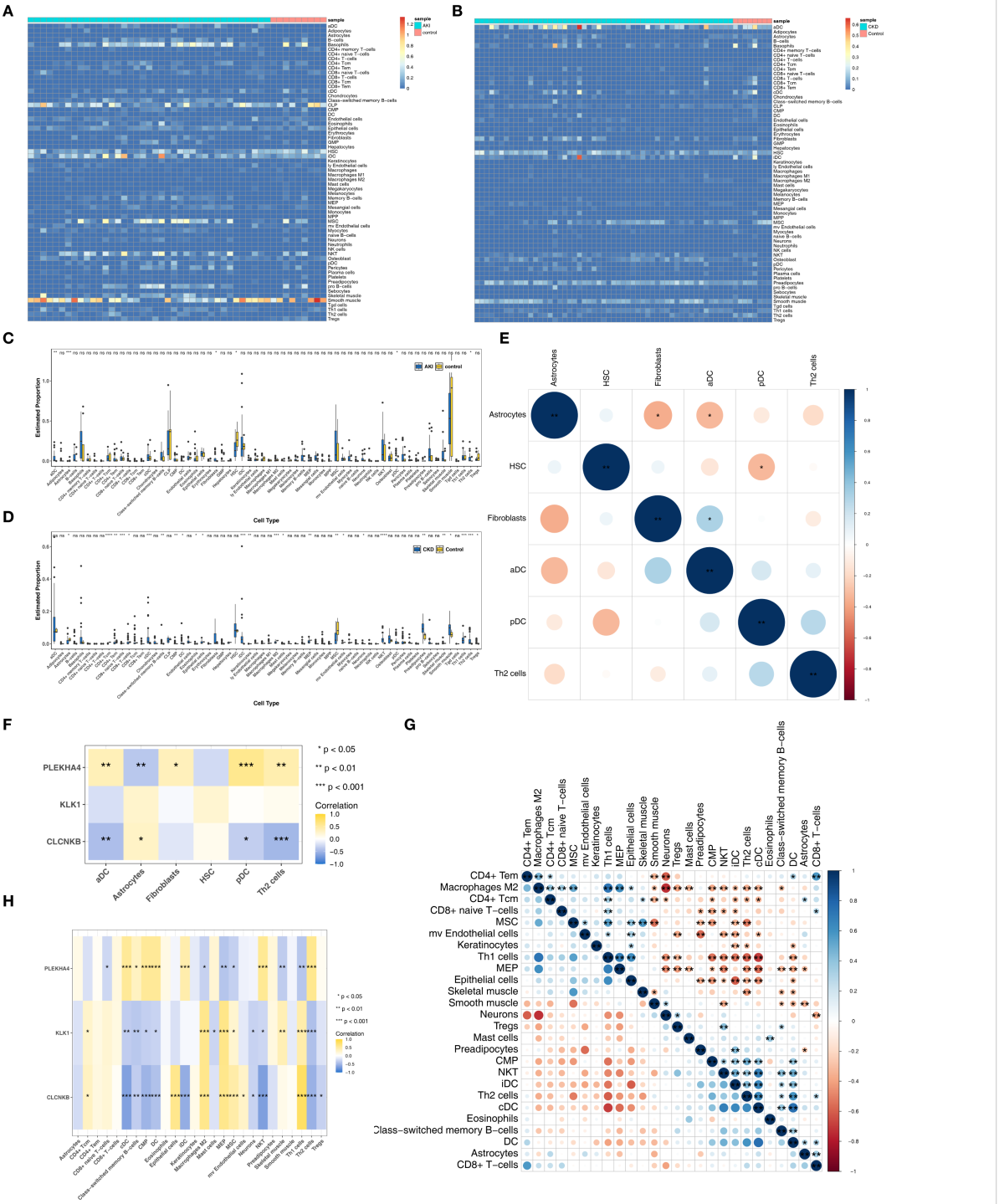


FIGURE 6
Immune infiltrating cells of AKI and CKD. **(A, B)** Infiltration levels of 64 immune cells in AKI and CKD set. **(C)** The different immune cells in AKI and control samples. **(D)** The different immune cells in CKD and control samples. **(E)** The correlation of different immune cells in AKI set. **(F)** Correlation of immune cells and candidate biomarkers in AKI set. **(G)** The correlation of different immune cells in CKD set. **(H)** Correlation of immune cells and candidate biomarkers in CKD set. AKI, Acute kidney injury; CKD, Chronic kidney disease. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

10.3389/fimmu.2025.1628962

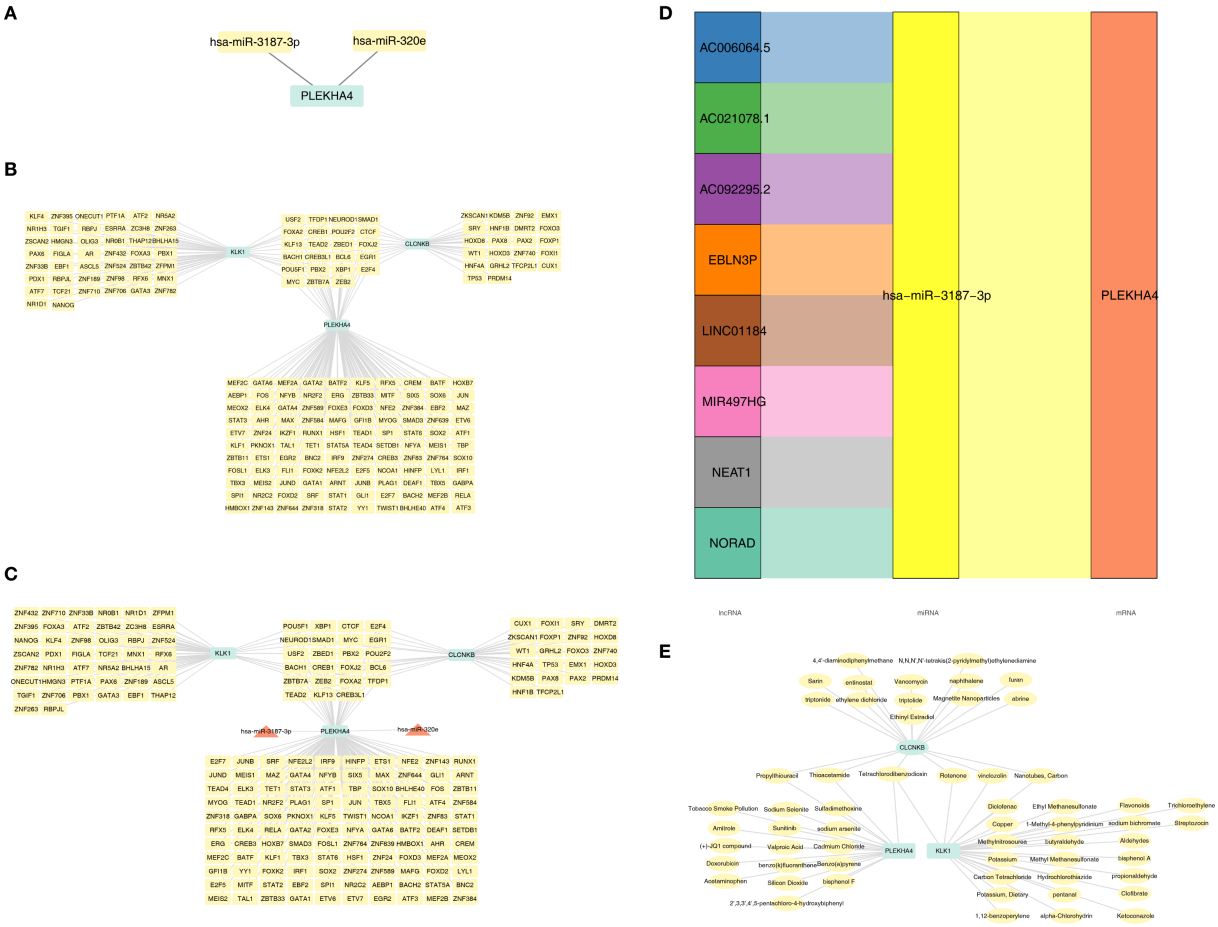


FIGURE 7
Molecular regulatory networks and biomarker-drug network of candidate biomarkers. **(A)** The miRNAs were predicted for candidate biomarkers. **(B)** TF-mRNA networks of candidate biomarkers. **(C)** miRNA-mRNA-TF network of candidate biomarkers. **(D)** lncRNA-miRNA-mRNA network of candidate biomarkers. **(E)** Biomarker-drug network of CLCNKB, KLK1 and PLEKHA4. TF, Transcription factors.

3.10 CLCNKB, KLK1, and PLEKHA4 expression changes during development of key cell subpopulations

Secondary dimensionality reduction clustering analysis was performed on Type A intercalated cell and Collecting duct-principal cell. It was found that Type A intercalated cell and Collecting duct-principal cell were divided into 10 and 8 subgroups respectively in the AKI single-cell dataset (Figures 12A, B). Whereas in the CKD single-cell dataset, Type A intercalated cell was divided into 10 subgroups and Collecting duct-principal cell was divided into 9 subgroups (Figures 12C, D). Subsequently, the different subgroups within Type A intercalated cell and Collecting duct-principal cell were arranged on the developmental trajectory according to the differentiation time. A darker blue indicates earlier cell differentiation. In addition, after different cell subgroups were mapped to the pseudo-time trajectory plot, it was found that they exhibited different differentiation states. In the AKI single-cell dataset, Type A intercalated cell had 10 differentiation states, with State 4 being the earliest and most specific in differentiation. Collecting duct-principal cell had 8 differentiation states, and State 4 was also the earliest and most specific (Figures 12E, F). In the CKD single-cell dataset, Type A intercalated cell also had 10 differentiation states, with State 9 being the earliest in differentiation. Collecting duct-principal cell had 8 differentiation states, and State 0 was the earliest and most specific (Figures 12G, H). In the AKI single-cell dataset, with the differentiation of Type A intercalated cells, the expressions of KLK1 and PLEKHA4 had no significant changes. The expression of CLCNKB showed a trend of first decreasing, then increasing and finally decreasing again (Figures 12I). With the development of Collecting duct-principal cells, PLEKHA4 had no significant change. The expression of CLCNKB showed a trend of first increasing and then decreasing, and the expression of KLK1 showed a trend of first remaining unchanged, then increasing, then decreasing and finally remaining unchanged (Figures 12J) In the CKD single-cell dataset, with the development of Type A intercalated cells, PLEKHA4 had no significant change. The expression of CLCNKB showed a trend of first decreasing and then increasing, and the expression of KLK1 showed a trend of first decreasing, then remaining unchanged and finally increasing (Figures 12K). With the development of Collecting duct-principal cells, PLEKHA4 had no significant

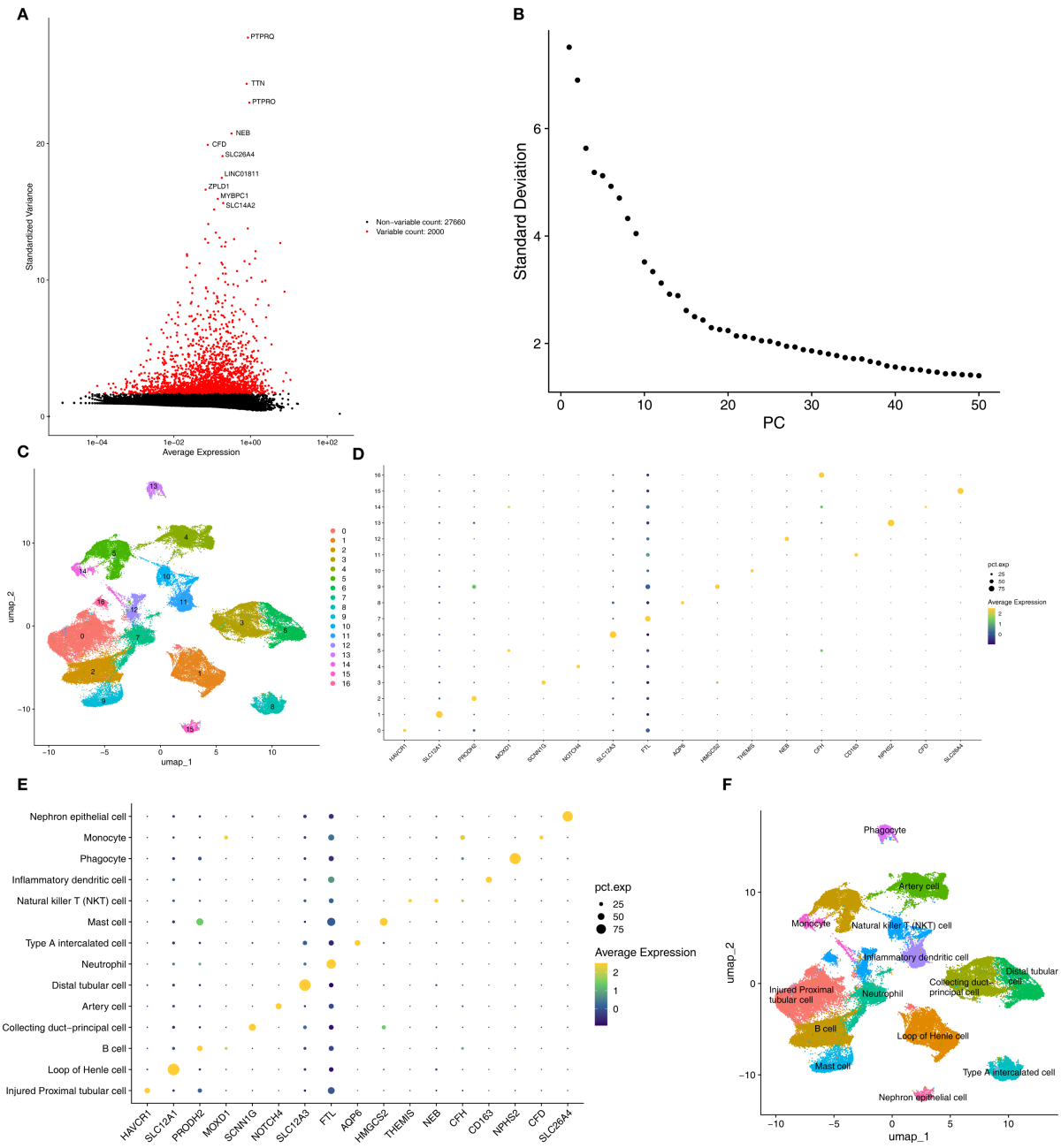


FIGURE 8
Annotated cell types in AKI groups. **(A, B)** Top variable genes and PCs were applied to UMAP clustering. **(C)** Different cell cluster of high-quality cells of AKI single-cell dataset. **(D, E)** Marker genes identified different cell clusters. **(F)** 14 cell types were determined in AKI single-cell dataset. AKI, Acute kidney injury; UMAP, Uniform manifold approximation and projection.

change. CLCNKB expression consistently declined, while KLK1 expression initially decreased and then stabilized (Figures 12L).

4 Discussion

AKI is marked by a swift reduction in kidney function over a brief period, and the transition from AKI to CKD is a widely recognized clinical occurrence. Our study identified three biomarkers (CLCNKB, KLK1, and PLEKHA4) through a combination of machine learning algorithms, ROC curve analysis, and expression validation. The potential mechanisms associated with these biomarkers in AKI and CKD were explored through enrichment analysis, regulatory network construction, immune infiltration analysis, and drug target prediction. By integrating single-cell data, we identified key cell types and investigated the expression of these biomarkers at the cellular level. Thus, our investigation discovered some new perspectives on the potential pathogenesis and progression of AKI to CKD, which might provide therapeutic targets to avert the transition.

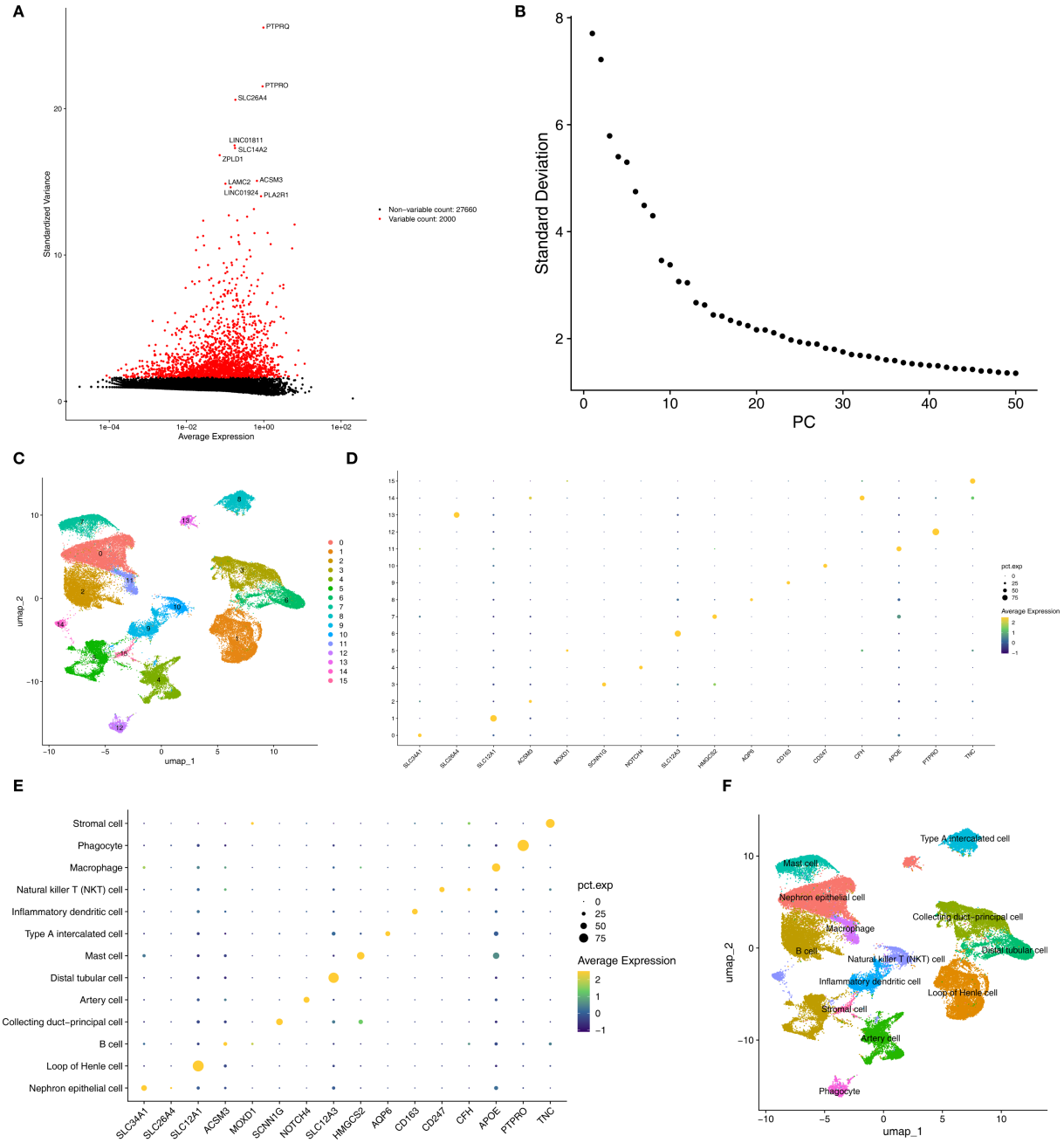


FIGURE 9
Annotated cell types in CKD groups. **(A, B)** Top variable genes and PCs were applied to UMAP clustering. **(C)** Different cell cluster of high-quality cells of CKD single-cell dataset. **(D, E)** Marker genes identified different cell clusters. **(F)** 13 cell types were determined in AKI single-cell dataset. CKD, Chronic kidney disease; UMAP, Uniform manifold approximation and projection.

CLCNKB is responsible for encoding the CLC-Kb protein, a component of the CLC chloride channel family, that helps with chloride ion reabsorption in the renal tubules (32, 33).Genetic mutations in CLCNKB can impair the glycosylation of the CLC-Kb protein, compromising its functionality and resulting in reduced uptake of sodium and chloride ions in the kidney tubules (34). Mutations in the CLCNKB gene are notably linked to Bartter syndrome type III, a rare hereditary renal tubular disorder characterized by salt loss and electrolyte imbalances, frequently culminating in CKD (35, 36).Our study demonstrated a significant down-regulation of CLCNKB in renal tissue samples from patients with both AKI and CKD, aligning with the loss-of-function effect indicated by the aforementioned genetic evidence. Moreover, this substantial loss of sodium and chloride ions triggers the activation of the renin – angiotensin – aldosterone system (RAAS), which may exacerbate kidney injury in AKI and facilitate the development to CKD (37).During the acute phase, persistent activation of the RAAS may exacerbate AKI-induced renal damage by promoting

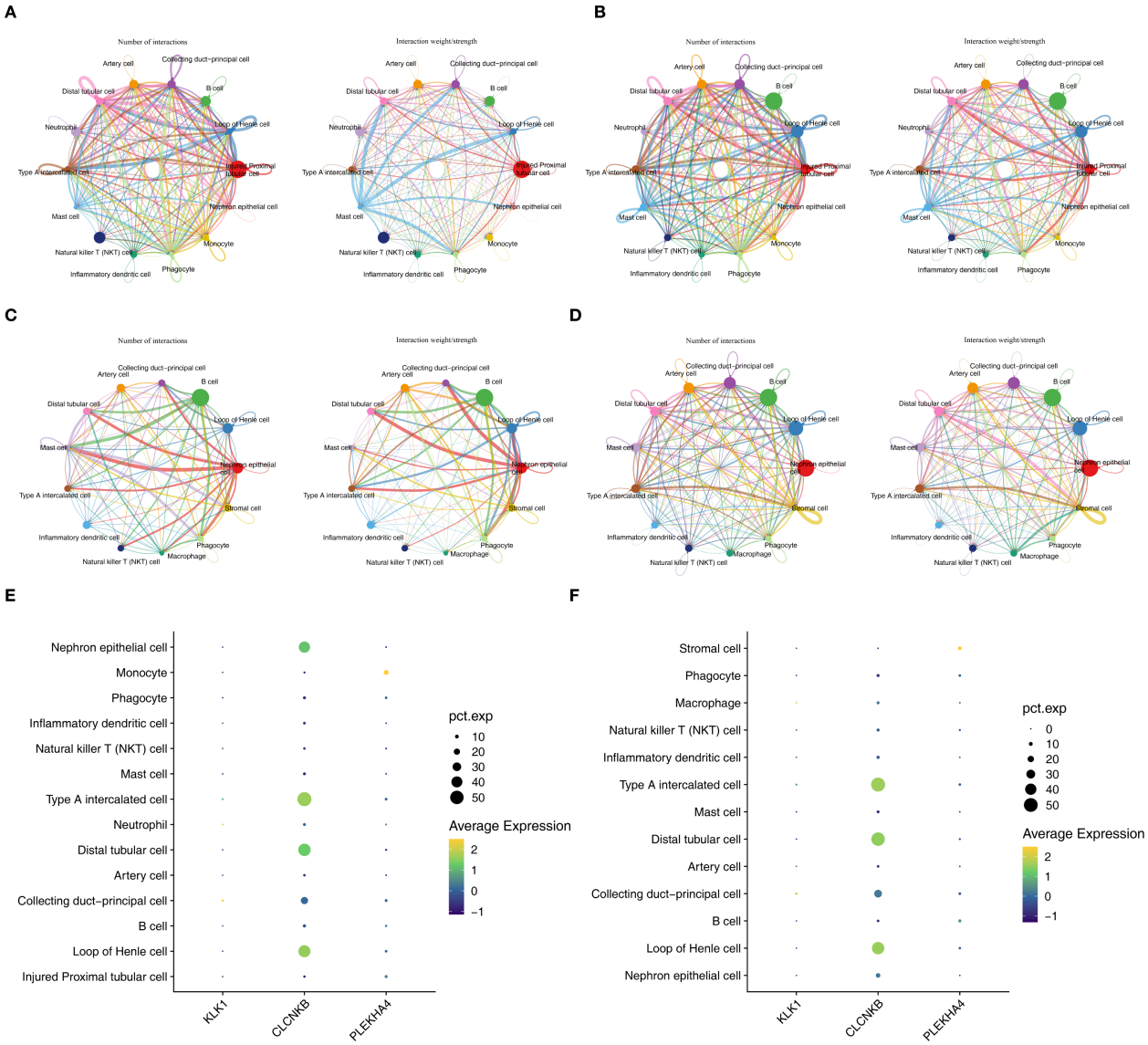


FIGURE 10
Key cells and expression trends of candidate biomarkers during development of key cell sub-populations of AKI and CKD single-cell datasets. **(A, B)** The cell interactions of AKI and control samples. **(C, D)** The cell interactions of CKD and control samples. **(E, F)** Cells with high expression of candidate biomarkers in AKI and CKD datasets.

vasoconstriction and inflammatory responses (37). Over the long term, this mechanism is pivotal in driving renal fibrosis and glomerulosclerosis, thereby expediting the progression from AKI to CKD (38). Functional enrichment analysis corroborated this mechanism. Furthermore, CLCNKB was found to be significantly associated with metabolic pathways, such as oxidative phosphorylation and branched-chain amino acid degradation, suggesting that its down-regulation may also be implicated in energy metabolism disorders within renal tubular cells, collectively facilitating the chronic progression of the disease.

KLK1 is a serine protease that plays a pivotal role in the kininase-kinin system (KKS) by breaking down low molecular weight kininogen to yield bradykinin (BK) (39). The KKS is intricately associated with several physiological processes, including inflammation, coagulation, pain, and increased vascular permeability, all of which are mediated by kinin production (40). KLK1 is notably involved in the signaling pathways of the B1 receptor for bradykinin (B1R) and the B2 receptor for bradykinin (B2R), thereby triggering a series of physiological responses that produce anti-apoptotic, anti-inflammatory, anti-fibrotic, and antioxidant effects. These actions collectively contribute to tissue protection, underscoring the multifaceted beneficial roles of KLK1 in maintaining tissue homeostasis (41). Furthermore, previous research has demonstrated that Klk1 ameliorates lupus nephritis in murine models (42, 43). The functional enrichment analysis conducted in this study revealed a significant association between the down-regulation of KLK1 expression and the neuroactive ligand-receptor interaction and cytokine-cytokine receptor

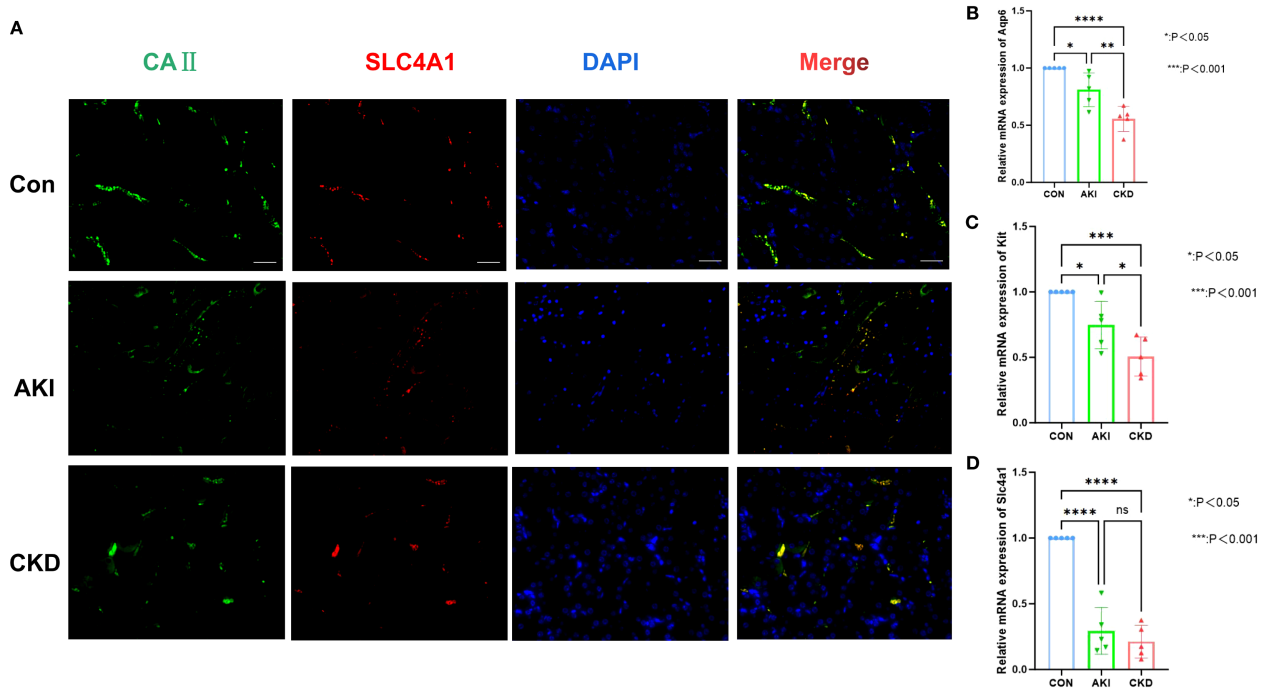


FIGURE 11
Abundance of type A intercalated cells in AKI/CKD mice. **(A)** Immunostaining of intercalated cell maker CAII (green) and the type A intercalated cell maker SLC4A1 (red) in renal collecting duct of AKI/CKD mice. **(B–D)** Relative expression level of type A intercalated cell maker (Aqp6, Kit and Slc4a1) in the kidneys of AKI/CKD mice.ns means not significant, *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.Bar:50um.

interaction signaling pathways. This suggests that reduced KLK1 expression may compromise the protective function of renal tubular cells by disrupting bradykinin signaling and exacerbating the inflammatory microenvironment.Consequently, the absence of KLK1 may be implicated in the development from AKI to CKD.

PLEKHA4 encodes a protein characterized by a Pleckstrin homology domain near its N-terminus and has an important function in cancer biology, particularly in gliomas. Furthermore, PLEKHA4 regulates the Wnt/β-catenin signaling pathway. *In vitro* downregulation of PLEKHA4 resulted in decreased dishevelled protein levels and a later diminishment of Wnt/β-catenin signaling (44).Conversely, overexpression of PLEKHA4 activated the Wnt/β-catenin pathway, facilitating the transfer of β-catenin to the nucleus and promoting signaling activity (45). The Wnt/β-catenin pathway, a developmental signaling cascade typically inactive in the adult kidney, becomes reactivated in various renal pathologies and plays a pivotal role in the pathogenesis of CKD (46, 47). Continuous activation of the Wnt/β-catenin signaling pathway has been linked to the advancement of kidney fibrosis, podocyte injury, and proteinuria in CKD (48–50), as well as contributing to AKI and sustained tissue damage in cystic kidney disease (51, 52). Furthermore, molecular regulatory networks suggest that PLEKHA4 may play a significant role in the progression from AKI to CKD. Consequently, the overexpression of PLEKHA4 could potentially exacerbate kidney damage in AKI and expedite the progression from AKI to CKD, warranting further investigation into the underlying mechanisms.

Our study indicates that during the progression from AKI to CKD, the oxidative phosphorylation pathway and the degradation pathways of valine, leucine, and isoleucine are pivotal. Following acute kidney injury, the renal repair process is often maladaptive, resulting in the dedifferentiation of tubular cells and the intensification of inflammatory responses. This maladaptive repair mechanism is intricately linked to the dysregulation of oxidative phosphorylation, which subsequently impacts long-term kidney function (53, 54). Research has demonstrated a significant association between valine degradation disorder and renal fibrosis, a critical pathological feature of CKD (55).Similarly, amino acid metabolism assumes a pivotal role in CKD (56). Amino acids can influence renal lesions and fibrosis through the aryl hydrocarbon receptor (AhR) signaling pathway (57). Certain amino acids, such as taurine, exhibit renoprotective properties by safeguarding the mitochondrial membrane and inhibiting cell apoptosis, thereby mitigating structural damage to the renal cortex (58, 59). The significance of amino acid metabolism in disease mechanisms positions it as a potential target for the early diagnosis and treatment of CKD (60). In CKD, amino acid metabolism is markedly disrupted, typically evidenced by elevated levels of arginine and citrulline and a decreased ornithine/citrulline ratio, indicating that citrulline may serve as a potent biomarker of renal metabolism (61). This metabolic disturbance interacts with systemic inflammation and metabolic acidosis, disrupting amino acid and protein homeostasis. As CKD progresses, glomerular filtration and renal tubular reabsorption functions are further

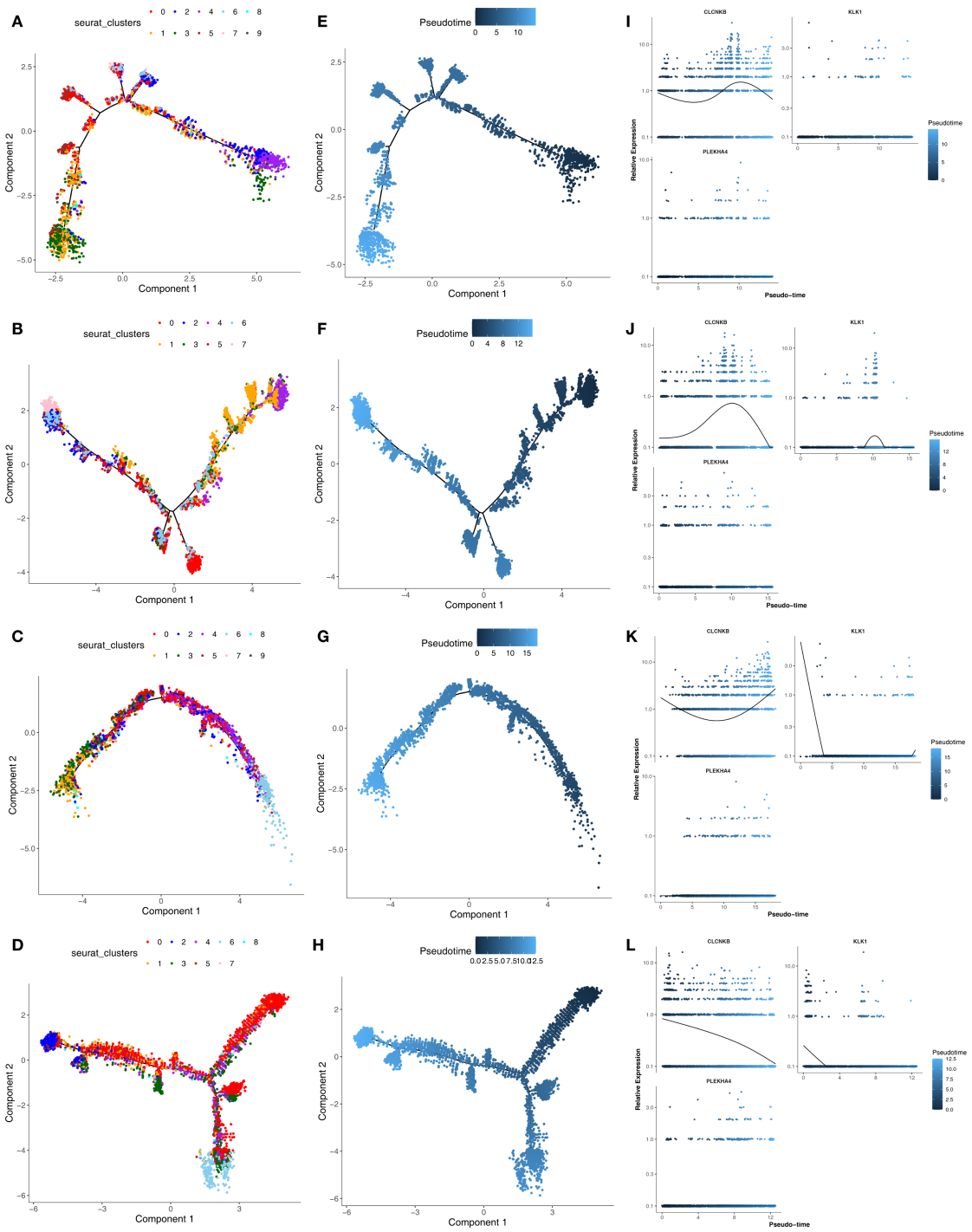


FIGURE 12
(A, B) The subgroups of Type A intercalated cell and Collecting duct-principal cell in AKI single-cell dataset. **(C, D)** The subgroups of Type A intercalated cell and Collecting duct-principal cell in CKD single-cell dataset. **(E, F)** Different states of Type A intercalated cell and Collecting duct-principal cell in AKI single-cell dataset. **(G, H)** Different states of Type A intercalated cell and Collecting duct-principal cell in CKD single-cell dataset. **(I, J)** Expression trends of CLCNKB、KLK1 and PLEKHA4 in Type A intercalated cell and Collecting duct-principal cell in AKI single-cell dataset. **(K, L)** Expression trends of CLCNKB、KLK1 and PLEKHA4 in Type A intercalated cell and Collecting duct-principal cell in CKD single-cell dataset. AKI, Acute kidney injury; CKD, Chronic kidney disease.

compromised, exacerbating amino acid depletion and proteinuria, thereby perpetuating a detrimental cycle (62). Therefore, interventions targeting oxidative phosphorylation pathways and amino acid metabolism may offer advanced therapeutic techniques to decelerate the progression from AKI to CKD.

During the transition from AKI to CKD, various cell types including fibroblasts, Th2 cells, astrocytes, DCs, and M2 macrophages, play pivotal roles, aligning with our findings. Studies indicate that fibroblasts differentiate into myofibroblasts following kidney injury, thereby promoting extracellular matrix

accumulation and contributing to renal fibrosis (63). Moreover, M2 macrophages exhibit a dual role in this process, engaging in tissue repair while potentially facilitating fibrosis progression in chronic inflammation (64). Post-AKI, Th2 cell activation may mitigate inflammatory responses and promote renal repair.However, an excessive Th2 cell response can also exacerbate fibrosis (65). In addition, suppressing PLEKHA4 might obstruct the M2 polarization process in macrophages (66). Thus, the positive association of Th2 cells with PLEKHA4 may expedite the progression of kidney fibrosis. Dendritic cells modulate T cell activation and differentiation through antigen presentation and cytokine secretion, thereby impacting the inflammatory and reparative mechanisms of kidney (67). Meanwhile, astrocytes are integral to the neuroimmune regulation of the kidney, potentially influencing the inflammatory response and fibrotic processes via the release of neurotransmitters and cytokines (68). It is noteworthy that interstitial cells, as a crucial component of the renal microenvironment, play a significant role in renal inflammation. They amplify local inflammatory signals by releasing proinflammatory factors, thereby inducing increased apoptosis of renal parenchymal cells and exacerbating renal fibrosis through the promotion of myofibroblast activation and extracellular matrix deposition (69). This process is intricately linked to abnormal oxidative stress, which not only results from the inflammatory response but also exacerbates inflammation and apoptosis by impairing mitochondrial function and activating the NF-κB and Nrf2 signaling pathways (70–72). These pathways are central to the regulation of apoptosis, inflammation (73), and oxidative stress in kidney diseases and are pivotal in determining the progression and outcomes of CKD (74, 75). Furthermore, Type A intercalated cells and collecting duct principal cells are identified as pivotal in the transition from AKI to CKD. This process encompasses a variety of complex pathophysiological mechanisms, including inflammation, fibrosis, and renal tubular injury. Type A interstitial cells, a distinct group of cells located in the kidney's collecting duct, are crucial for maintaining acid-base equilibrium and facilitating ion transport (76). AKI is frequently associated with an inflammatory response, which stimulates the release of pro-inflammatory cytokines and chemokines (77). Type A interstitial cells may exacerbate renal fibrosis by promoting fibroblast activation and collagen synthesis (78). Furthermore, the dysfunction of intercalated cells is intricately associated with alterations in the renal microenvironment, which may encompass hypoxia, modifications in the extracellular matrix, and dysregulation of intercellular signaling pathways (79). Collecting duct principal cells, another predominant cell type in the collecting duct, are responsible for the regulation of sodium and water reabsorption, thereby maintaining fluid balance (77). Dysfunction in the collecting duct principal cells results in compromised water and sodium reabsorption, further exacerbating kidney damage. Collecting duct principal cells demonstrate considerable proliferative capacity following acute kidney injury, a response likely aimed at compensating for tubular damage and facilitating renal repair (80). Consequently, these cellular types may represent potential therapeutic targets in the progression from AKI to CKD, warranting further in-depth investigation into the interactions among different cell types.

In this study, a biomarker-drug network was developed, revealing that CLCNKB, PLEKHA4, and KLK1 are concurrently targeted by tetrachlorodibenzodioxin. However, tetrachlorobiphenyldioxin is recognized as an environmental pollutant that induces toxicity across multiple tissues, including the kidneys (81). Research has demonstrated that exposure to tetrachlorobiphenyldioxin can result in oxidative stress, leading to cellular damage and dysfunction within the kidneys (82). However, the therapeutic effect of tetrachlorobiphenyldioxin are poorly studied. These findings are contrary to our results, indicating that the mechanisms of action of related drugs require further exploration in future studies.

In this study, miRNAs and immune cells synergistically influence the progression from AKI to CKD by targeting specific biomarkers and engaging in the inflammation-fibrosis axis. Regarding miRNAs, although only two miRNAs, such as hsa-miR-3187-3p, were predicted to target PLEKHA4, the constructed lncRNA-miRNA-mRNA network indicates its regulatory role. MiRNAs may negatively regulate PLEKHA4 expression by promoting mRNA degradation or inhibiting its translation. The downregulation or loss of function of miRNAs can lead to PLEKHA4 overexpression, which subsequently activates the Wnt/β-catenin pathway. This activation promotes fibroblast activation, epithelial-mesenchymal transition, and extracellular matrix deposition, thereby accelerating renal fibrosis (51). These findings illuminate the intricate mechanisms underlying immunometabolic regulation in kidney diseases and provide a rationale for therapeutic strategies targeting miRNAs or immune cells.

Among the biomarkers identified in this study, CLCNKB demonstrates significant novelty. Prior research has predominantly concentrated on the relationship between CLCNKB variants and inherited renal tubular disorders, such as Bartter syndrome (35). However, to date, no investigations have reported an association between CLCNKB and AKI or CKD. This study is the first to reveal that CLCNKB plays a crucial role in the transition from AKI to CKD, potentially offering a novel perspective on the mechanisms underlying AKI-CKD progression. In contrast, KLK1 and PLEKHA4 are established targets in AKI and CKD research. KLK1 has been demonstrated to play a significant role in kidney disease (83). Regarding PLEKHA4, the continuous activation of the Wnt/β-catenin signaling pathway is implicated in the progression of renal fibrosis in CKD, contributing to ongoing tissue damage in kidney disease (52). Through comprehensive bioinformatics analysis, this study systematically examined the expression patterns and potential regulatory networks of KLK1 and PLEKHA4 in the AKI-CKD transition, thereby enhancing the understanding of their mechanisms in kidney diseases. The identification of these biomarkers not only provides potential molecular indicators for early diagnosis but also enriches the current understanding of kidney disease pathophysiology.

Three biomarkers, CLCNKB, KLK1, and PLEKHA4, were identified through bioinformatics methods as being associated with the progression of AKI to CKD. Functional enrichment analysis was

conducted based on these biomarkers to elucidate the biological pathways involved in AKI and CKD. Additionally, correlation analysis between differential immune cells and the identified biomarkers was performed to explore potential regulatory relationships. Single-cell analysis provided insights into the cellular-level expression of these biomarkers, offering new perspectives for early diagnosis and the development of novel therapeutic strategies for AKI and CKD. This study is subject to several limitations. Firstly, the retrospective analysis based on public databases is unable to fully eliminate batch effects, and the sample sizes are constrained (for instance, the scRNA-seq dataset includes only five cases of AKI and two cases of CKD), which impedes the effective application of multivariate statistical analysis to control for confounding factors. Secondly, the study lacks gene function experiments, such as gene knockout or overexpression, which are necessary to directly validate the causal mechanisms of the candidate genes. Furthermore, the clinical translation of target-related compounds, such as tetrachlorodibenzo-dioxins, is severely limited due to their toxicity. Future research should aim to expand the sample size through multi-center prospective cohort studies to acquire comprehensive clinical information. Additionally, animal models and cellular experiments, including gene editing and inhibitor or agonist treatments, should be employed to further elucidate the specific mechanisms by which CLCNKB, KLK1, and PLEKHA4 regulate fibrosis and the immune microenvironment. Moreover, flow cytometry and RNA sequencing (RNA-seq) technologies will be employed to assess the dynamic expression and functional status of type A interstitial cells within a kidney injury model, thereby elucidating their potential role in the disease pathology. Ultimately, these insights are intended to be translated into early intervention and targeted therapies for kidney disease through drug repositioning or the development of novel inhibitors or agonists.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

The studies involving humans were approved by The ethics committee of The First Affiliated Hospital of Guangxi Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. The animal study was approved by Animals as stipulated by the Ministry of Science and Technology of the People's Republic of China, as well as the Guidelines for Ethical Review of Laboratory Animals according to the National Standard GB/T35892-2018 of the People's Republic of China, and the protocols of the Animal Care and Welfare Committee at Guangxi Medical University (No:202506002). The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

FZ: Validation, Visualization, Writing – original draft, Writing – review & editing, Conceptualization, Methodology. ZY: Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing, Formal analysis. ZW: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Resources, Software, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. Project supported by the Young Scientists Fund of the Guangxi Medical University (RC2300016245,provided by ZW); the fifth level Talent Fund of the First Affiliated Hospital of Guangxi Medical University (No. 2022020,provided by ZW).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2025.1628962/full#supplementary-material

SUPPLEMENTARY FIGURE 1
The quality control of AKI single-cell dataset. AKI: Acute kidney injury.

SUPPLEMENTARY FIGURE 2
The quality control of CKD single-cell dataset. CKD: Chronic kidney disease.

References

1. Chawla LS, Bellomo R, Bihorac A, Goldstein SL, Siew ED, Bagshaw SM, et al. Acute kidney disease and renal recovery: consensus report of the Acute Disease Quality Initiative (ADQI) 16 Workgroup. *Nat Rev Nephrol*. (2017) 13:241–57. doi: 10.1038/nrneph.2017.2

2. Almazmomi MA, Esmat A, Naeem A. Acute kidney injury: definition, management, and promising therapeutic target. *Cureus*. (2023) 15:e51228. doi: 10.7759/cureus.51228

3. Hoste EA, Bagshaw SM, Bellomo R, Cely CM, Colman R, Cruz DN, et al. Epidemiology of acute kidney injury in critically ill patients: the multinational AKI-EPI study. *Intensive Care Med*. (2015) 41:1411–23. doi: 10.1007/s00134-015-3934-7

4. See EJ, Jayasinghe K, Glassford N, Bailey M, Johnson DW, Polkinghorne KR, et al. Long-term risk of adverse outcomes after acute kidney injury: a systematic review and meta-analysis of cohort studies using consensus definitions of exposure. *Kidney Int*. (2019) 95:160–72. doi: 10.1016/j.kint.2018.08.036

5. Chen TK, Knicely DH, Grams ME. Chronic kidney disease diagnosis and management: A review. *JAMA*. (2019) 322:1294–304. doi: 10.1001/jama.2019.14745

6. Liu P, Quinn RR, Lam NN, Elliott MJ, Xu Y, James MT, et al. Accounting for age in the definition of chronic kidney disease. *JAMA Intern Med*. (2021) 181:1359–66. doi: 10.1001/jamainternmed.2021.4813

7. GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. (2020) 395:709–33. doi: 10.1016/S0140-6736(20)30045-3

8. Liyanage T, Toyama T, Hockham C, Ninomiya T, Perkovic V, Woodward M, et al. Prevalence of chronic kidney disease in Asia: a systematic review and analysis. *BMJ Glob Health*. (2022) 7:e007525. doi: 10.1136/bmjgh-2021-007525

9. Wang YN, Ma SX, Chen YY, Chen L, Liu BL, Liu QQ, et al. Chronic kidney disease: Biomarker diagnosis to therapeutic targets. *Clin Chim Acta*. (2019) 499:54–63. doi: 10.1016/j.cca.2019.08.030

10. Chawla LS, Eggers PW, Star RA, Kimmel PL. Acute kidney injury and chronic kidney disease as interconnected syndromes. *N Engl J Med*. (2014) 371:58–66. doi: 10.1056/NEJMra1214243

11. Kurzhagen JT, Dellepiane S, Cantaluppi V, Rabb H. AKI: an increasingly recognized risk factor for CKD development and progression. *J Nephrol*. (2020) 33:1171–87. doi: 10.1007/s40620-020-00793-2

12. Pickkers P, Darmon M, Hoste E, Joannidis M, Legrand M, Ostermann M, et al. Acute kidney injury in the critically ill: an updated review on pathophysiology and management. *Intensive Care Med*. (2021) 47:835–50. doi: 10.1007/s00134-021-06454-7

13. Zhang T, Widdop RE, Ricardo SD. Transition from acute kidney injury to chronic kidney disease: mechanisms, models, and biomarkers. *Am J Physiol Renal Physiol*. (2024) 327:F788–805. doi: 10.1152/ajprenal.00184.2024

14. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. (2014) 15:550. doi: 10.1186/s13059-014-0550-8

15. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. (2004) 3. doi: 10.2202/1544-6115.1027

16. Xie ZW, He Y, Feng YX, Wang XH. Identification of programmed cell death-related genes and diagnostic biomarkers in endometriosis using a machine learning and Mendelian randomization approach. *Front Endocrinol (Lausanne)*. (2024) 15:1372221. doi: 10.3389/fendo.2024.1372221

17. Zhang X, Chao P, Zhang L, Xu L, Cui X, Wang S, et al. Single-cell RNA and transcriptome sequencing profiles identify immune-associated key genes in the development of diabetic kidney disease. *Front Immunol*. (2023) 14:1030198. doi: 10.3389/fimmu.2023.1030198

18. Jia A, Xu L, Wang Y. Venn diagrams in bioinformatics. *Brief Bioinform*. (2021) 22:bbab108. doi: 10.1093/bib/bbab108

19. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. (2012) 16:284–7. doi: 10.1089/omi.2011.0118

20. Zhang C, Zheng Y, Li X, Hu X, Qi F, Luo J. Genome-wide mutation profiling and related risk signature for prognosis of papillary renal cell carcinoma. *Ann Transl Med*. (2019) 7:427. doi: 10.21037/atm.2019.08.113

21. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape stringApp: network analysis and visualization of proteomics data. *J Proteome Res*. (2019) 18:623–32. doi: 10.1021/acs.jproteome.8b00702

22. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Software*. (2010) 33:1–22. doi: 10.18637/jss.v033.i01

23. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf*. (2011) 12:77. doi: 10.1186/1471-2105-12-77

24. Robles-Jimenez LE, Aranda-Aguirre E, Castelan-Ortega OA, Shettino-Bermudez BS, Ortiz-Salinas R, Miranda M, et al. Worldwide traceability of antibiotic residues from livestock in wastewater and soil: A systematic review. *Anim (Basel)*. (2021) 12:60. doi: 10.3390/ani12010060

25. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol*. (2017) 18:220. doi: 10.1186/s13059-017-1349-1

26. Su J, Zhou W, Yuan H, Wang H, Zhang H. Identification and functional analysis of novel biomarkers in adenoid cystic carcinoma. *Cell Mol Biol (Noisy-le-grand)*. (2023) 69:203–7. doi: 10.14715/cmb/2023.69.6.31

27. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol*. (2024) 42:293–304. doi: 10.1038/s41587-023-01767-y

28. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. (2021) 12:1088. doi: 10.1038/s41467-021-21246-9

29. Mayrhofer M, DiLorenzo S, Isaksson A. Patchwork: allele-specific copy number analysis of whole-genome sequenced tumor tissue. *Genome Biol*. (2013) 14:R24. doi: 10.1186/gb-2013-14-3-r24

30. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. (2017) 14:979–82. doi: 10.1038/nmeth.4402

31. Zhang YL, Tang TT, Wang B, Wen Y, Feng Y, Yin Q, et al. Identification of a novel ECM remodeling macrophage subset in AKI to CKD transition by integrative spatial and single-cell analysis. *Adv Sci (Weinh)*. (2024) 11:e2309752. doi: 10.1002/advs.202309752

32. Seys E, Andrini O, Keck M, Mansour-Hendili L, Courand PY, Simian C, et al. Clinical and genetic spectrum of bartter syndrome type 3. *J Am Soc Nephrol*. (2017) 28:2540–52. doi: 10.1681/ASN.2016101057

33. Zaika O, Tomilin V, Mamenko M, Bhalla V, Pochynyuk O. New perspective of ClC-Kb/2 Cl- channel physiology in the distal renal tubule. *Am J Physiol Renal Physiol*. (2016) 310:F923–30. doi: 10.1152/ajprenal.00577.2015

34. Sharma Y, Lo R, Tomilin VN, Ha K, Deremo H, Pareek AV, et al. ClC-Kb pore mutation disrupts glycosylation and triggers distal tubular remodeling. *JCI Insight*. (2024) 9:e175998. doi: 10.1172/jci.insight.175998

35. Roodaki N, Salinas LM, Maceda EBG, Frias J. Novel CLCNKB mutation in two siblings with classic bartter syndrome. *Case Rep Genet*. (2025) 2025:8862780. doi: 10.1155/crig/8862780

36. Feng J, Chen Z, Wang Y, Liu Y, Zhao D, Gu X. Identification of chromatin remodeling-related gene signature to predict the prognosis in breast cancer. *Clin Exp Med*. (2025) 25:137. doi: 10.1007/s10238-025-01661-8

37. Gupta S, Mandal S, Banerjee K, Almarshood H, Pushpakumar SB, Sen U. Complex pathophysiology of acute kidney injury (AKI) in aging: epigenetic regulation, matrix remodeling, and the healing effects of H2S. *Biomolecules*. (2024) 14:1165. doi: 10.3390/biom14091165

38. Sacks D, Baxter B, Campbell BCV, Carpenter JS, Cognard C, Dippel D, et al. Multisociety consensus quality improvement revised consensus statement for endovascular therapy of acute ischemic stroke. *Int J Stroke*. (2018) 13:612–32. doi: 10.1177/1747493018778713

39. Devetzi M, Goulielmaki M, Khoury N, Spandidos DA, Sotiropoulou G, Christodoulou I, et al. Genetically–modified stem cells in treatment of human diseases: Tissue kallikrein (KLK1)–based targeted therapy (Review). *Int J Mol Med*. (2018) 41:1177–86. doi: 10.3892/ijmm.2018.3361

40. Kakoki M, Smithies O. The kallikrein-kinin system in health and in diseases of the kidney. *Kidney Int*. (2009) 75:1019–30. doi: 10.1038/ki.2008.647

41. Chao J, Shen B, Gao L, Xia CF, Bledsoe G, Chao L. Tissue kallikrein in cardiovascular, cerebrovascular and renal diseases and skin wound healing. *Biol Chem*. (2010) 391:345–55. doi: 10.1515/BC.2010.042

42. Lindgren S, Anzén B, Bohlin AB, Lidman K. HIV and child-bearing: clinical outcome and aspects of mother-to-infant transmission. *AIDS*. (1991) 5:1111–6. doi: 10.1097/00002030-199109000-00009

43. Li QZ, Zhou J, Yang R, Yan M, Ye Q, Liu K, et al. The lupus-susceptibility gene kallikrein downmodulates antibody-mediated glomerulonephritis. *Genes Immun*. (2009) 10:503–8. doi: 10.1038/gene.2009.7

44. Shami Shah A, Cao X, White AC, Baskin JM. PLEKHA4 promotes wnt/β-catenin signaling-mediated G1-S transition and proliferation in melanoma. *Cancer Res*. (2021) 81:2029–43. doi: 10.3892/mmr.2024.13395

45. Yue Y, An G, Cao S, Li X, Du L, Xu D, et al. PLEKHA4 upregulation regulates KIRC cell proliferation through β-catenin signaling. *Mol Med Rep*. (2025) 31:30. doi: 10.3892/mmr.2024.13395

46. Tian Y, Chen J, Huang W, Ren Q, Feng J, Liao J, et al. Myeloid-derived Wnts play an indispensible role in macrophage and fibroblast activation and kidney fibrosis. *Int J Biol Sci*. (2024) 20:2310–22. doi: 10.7150/ijbs.94166

47. Fu D, Senouthai S, Wang J, You Y. FKN facilitates HK-2 cell EMT and tubulointerstitial lesions via the wnt/β-catenin pathway in a murine model of lupus nephritis. *Front Immunol*. (2019) 10:784. doi: 10.3389/fimmu.2019.00784

48. Long Y, Song D, Xiao L, Xiang Y, Li D, Sun X, et al. m6A RNA methylation drives kidney fibrosis by upregulating β-catenin signaling. *Int J Biol Sci*. (2024) 20:3185–200. doi: 10.7150/ijbs.96233

49. Miao H, Wang YN, Su W, Zou L, Zhuang SG, Yu XY, et al. Sirtuin 6 protects against podocyte injury by blocking the renin-angiotensin system by inhibiting the Wnt1/β-catenin pathway. *Acta Pharmacol Sin*. (2024) 45:137–49. doi: 10.1038/s41401-023-01148-w

50. Jin B, Liu J, Zhu Y, Lu J, Zhang Q, Liang Y, et al. Kunxian capsule alleviates podocyte injury and proteinuria by inactivating β-catenin in db/db mice. *Front Med (Lausanne)*. (2023) 10:1213191. doi: 10.3389/fmed.2023.1213191

51. Schunk SJ, Floege J, Fliser D, Speer T. WNT-β-catenin signalling - a versatile player in kidney injury and repair. *Nat Rev Nephrol*. (2021) 17:172–84. doi: 10.1038/s41581-020-00343-w

52. Bienaimé F, Canaud G, El Karoui K, Gallazzini M, Terzi F. Molecular pathways of chronic kidney disease progression. *Nephrol Ther*. (2016) 12 Suppl 1:S35–8. doi: 10.1016/j.nephro.2016.02.009

53. Venkatachalam MA, Griffin KA, Lan R, Geng H, Saikumar P, Bidani AK. Acute kidney injury: a springboard for progression in chronic kidney disease. *Am J Physiol Renal Physiol*. (2010) 298:F1078–94. doi: 10.1152/ajprenal.00017.2010

54. Console L, Scalise M, Giangregorio N, Tonazzi A, Barile M, Indiveri C. The link between the mitochondrial fatty acid oxidation derangement and kidney injury. *Front Physiol*. (2020) 11:794. doi: 10.3389/fphys.2020.00794

55. Sato Y, Yanagita M. Resident fibroblasts in the kidney: a major driver of fibrosis and inflammation. *Inflammation Regen*. (2017) 37:17. doi: 10.1186/s41232-017-0048-3

56. Miao H, Liu F, Wang YN, Yu XY, Zhuang S, Guo Y, et al. Targeting Lactobacillus johnsonii to reverse chronic kidney disease. *Signal Transduct Target Ther*. (2024) 9:195. doi: 10.1038/s41392-024-01913-1

57. Miao H, Wang YN, Yu XY, Zou L, Guo Y, Su W, et al. Lactobacillus species ameliorate membranous nephropathy through inhibiting the aryl hydrocarbon receptor pathway via tryptophan-produced indole metabolites. *Br J Pharmacol*. (2024) 181:162–79. doi: 10.1111/bph.16219

58. Ma J, Yang Z, Jia S, Yang R. A systematic review of preclinical studies on the taurine role during diabetic nephropathy: focused on anti-oxidative, anti-inflammation, and anti-apoptotic effects. *Toxicol Mech Methods*. (2022) 32:420–30. doi: 10.1080/15376516.2021.2021579

59. Chen WQ, Jin H, Nguyen M, Carr J, Lee YJ, Hsu CC, et al. Role of taurine in regulation of intracellular calcium level and neuroprotective function in cultured neurons. *J Neurosci Res*. (2001) 66:612–9. doi: 10.1002/jnr.10027

60. Li C, Gao L, Lv C, Li Z, Fan S, Liu X, et al. Active role of amino acid metabolism in early diagnosis and treatment of diabetic kidney disease. *Front Nutr*. (2023) 10:1239838. doi: 10.3389/fnut.2023.1239838

61. Zhang W, Zheng J, Zhang J, Li N, Yang X, Fang ZZ, et al. Associations of serum amino acids related to urea cycle with risk of chronic kidney disease in Chinese with type 2 diabetes. *Front Endocrinol (Lausanne)*. (2023) 14:1117308. doi: 10.3389/fendo.2023.1117308

62. Liu X, Li Q, Zhang L, He Y, Tan S, Chen X, et al. Impacts of amino acid supplementation on renal function and nutritional parameters in patients with renal insufficiency: bibliometric analysis and meta-analysis. *Front Nutr*. (2025) 12:1594507. doi: 10.3389/fnut.2025.1594507

63. Kim MG, Kim SC, Ko YS, Lee HY, Jo SK, Cho W. The role of M2 macrophages in the progression of chronic kidney disease following acute kidney injury. *PLoS One*. (2015) 10:e0143961. doi: 10.1371/journal.pone.0143961

64. Liu L, Kou P, Zeng Q, Pei G, Li Y, Liang H, et al. CD4+ T Lymphocytes, especially Th2 cells, contribute to the progress of renal fibrosis. *Am J Nephrol*. (2012) 36:386–96. doi: 10.1159/000343283

65. Li HD, You YK, Shao BY, Wu WF, Wang YF, Guo JB, et al. Roles and crosstalks of macrophages in diabetic nephropathy. *Front Immunol*. (2022) 13:1015142. doi: 10.3389/fimmu.2022.1015142

66. He Y, Zheng W, Huo Y, Sa L, Zhang H, He G, et al. PLEKHA4 promotes glioblastoma progression through apoptosis inhibition, tumor cell migration, and macrophage infiltration. *Immunobiology*. (2023) 228:152746. doi: 10.1016/j.imbio.2023.152746

67. Lee B, Ines I, Je J, Park EJ, Seong H, Jo MG, et al. Effect of renal ischemia reperfusion on brain neuroinflammation. *Biomedicines*. (2022) 10:2993. doi: 10.3390/biomedicines10112993

68. Gewin LS. Transforming growth factor-β in the acute kidney injury to chronic kidney disease transition. *Nephron*. (2019) 143:154–7. doi: 10.1159/000500093

69. Huang J, Liu Y, Shi M, Zhang X, Zhong Y, Guo S, et al. Empagliflozin attenuating renal interstitial fibrosis in diabetic kidney disease by inhibiting lymphangiogenesis and lymphatic endothelial-to-mesenchymal transition via the VEGF-C/VEGFR3 pathway. *BioMed Pharmacother*. (2024) 180:117589. doi: 10.1016/j.biopha.2024.117589

70. Jiang XS, Liu T, Xia YF, Gan H, Ren W, Du XG. Activation of the Nrf2/ARE signaling pathway ameliorates hyperlipidemia-induced renal tubular epithelial cell injury by inhibiting mtROS-mediated NLRP3 inflammasome activation. *Front Immunol*. (2024) 15:1342350. doi: 10.3389/fimmu.2024.1342350

71. Wang YN, Miao H, Yu XY, Guo Y, Su W, Liu F, et al. Oxidative stress and inflammation are mediated via aryl hydrocarbon receptor signalling in idiopathic membranous nephropathy. *Free Radic Biol Med*. (2023) 207:89–106. doi: 10.1016/j.freeradbiomed.2023.07.014

72. Li J, Guo Q, Wei X, Zhu Y, Luo M, Luo P. Association of serum Nrf2 protein levels with disease activity and renal impairment in lupus nephritis. *Front Immunol*. (2024) 15:1304167. doi: 10.3389/fimmu.2024.1304167

73. Zhang J, Luan ZL, Huo XK, Zhang M, Morisseau C, Sun CP, et al. Direct targeting of sEH with alisol B alleviated the apoptosis, inflammation, and oxidative stress in cisplatin-induced acute kidney injury. *Int J Biol Sci*. (2023) 19:294–310. doi: 10.7150/ijbs.78097

74. Li XJ, Shan QY, Wu X, Miao H, Zhao YY. Gut microbiota regulates oxidative stress and inflammation: a double-edged sword in renal fibrosis. *Cell Mol Life Sci*. (2024) 81:480. doi: 10.1007/s00018-024-05532-5

75. Zhao BR, Hu XR, Wang WD, Zhou Y. Cardiorenal syndrome: clinical diagnosis, molecular mechanisms and therapeutic strategies. *Acta Pharmacol Sin*. (2025) 46:1539–55. doi: 10.1038/s41401-025-01476-z

76. Wall SM. Regulation of blood pressure and salt balance by pendrin-positive intercalated cells: donald seldin lecture 2020. *Hypertension*. (2022) 79:706–16. doi: 10.1161/HYPERTENSIONAHA.121.16492

77. Wu ST, Feng Y, Song R, Qi Y, Li L, Lu D, et al. Foxp1 is required for renal intercalated cell differentiation and acid-base regulation. *J Am Soc Nephrol*. (2024) 35:533–48. doi: 10.1681/ASN.0000000000000319

78. Baudoux T, Jadot I, Declèves AE, Antoine MH, Colet JM, Botton O, et al. Experimental aristolochic acid nephropathy: A relevant model to study AKI-to-CKD transition. *Front Med (Lausanne)*. (2022) 9:822870. doi: 10.3389/fmed.2022.822870

79. Rao R, Bhalla V, Pastor-Soler NM. Intercalated cells of the kidney collecting duct in kidney physiology. *Semin Nephrol*. (2019) 39:353–67. doi: 10.1016/j.semnephrol.2019.04.005

80. Li Y, Wang Z, Xu H, Hong Y, Shi M, Hu B, et al. Targeting the transmembrane cytokine co-receptor neuropilin-1 in distal tubules improves renal injury and fibrosis. *Nat Commun*. (2024) 15:5731. doi: 10.1038/s41467-024-50121-6

81. Vial R, Poitevin S, McKay N, Burtey S, Cerini C. Tryptophan metabolites regulate neuropentraxin 1 expression in endothelial cells. *Int J Mol Sci*. (2022) 23:2369. doi: 10.3390/ijms23042369

82. Vijaya Padma V, Kalai Selvi P, Sravani S. Protective effect of ellagic acid against TCDD-induced renal oxidative stress: modulation of CYP1A1 activity and antioxidant defense mechanisms. *Mol Biol Rep*. (2014) 41:4223–32. doi: 10.1007/s11033-014-3292-5

83. Liu K, Li QZ, Delgado-Vega AM, et al. Kallikrein genes are associated with lupus and glomerular basement membrane-specific antibody-induced nephritis in mice and humans. *J Clin Invest*. (2009) 119:911–23. doi: 10.1172/JCI36728

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 27 October 2025
DOI 10.3389/fimmu.2025.1587522

OPEN ACCESS

EDITED BY
Yejun Tan,
Hong Kong Polytechnic University,
Hong Kong SAR, China

REVIEWED BY
Jinwei Li,
Sichuan University, China
Zoya Serebrovska,
Institute of Gerontology (NAN Ukraine),
Ukraine

*CORRESPONDENCE
Riken Chen
chenriken@126.com
Weimin Yao
490296443@qq.com

†These authors have contributed equally to this work

RECEIVED 05 March 2025
ACCEPTED 06 October 2025
PUBLISHED 27 October 2025

CITATION
Ye W, Yang Y, Chen F, Lin X, Wang Y, Du L, Pan J, Liao W, Chen B, Chen R and Yao W (2025) Decoding the hypoxia-exosome-immune triad in OSA: PRCP/UCHL1/BTG2-driven metabolic dysregulation revealed by interpretable machine learning.
Front. Immunol. 16:1587522.
doi: 10.3389/fimmu.2025.1587522

Decoding the hypoxia-exosome-immune triad in OSA: PRCP/UCHL1/BTG2-driven metabolic dysregulation revealed by interpretable machine learning

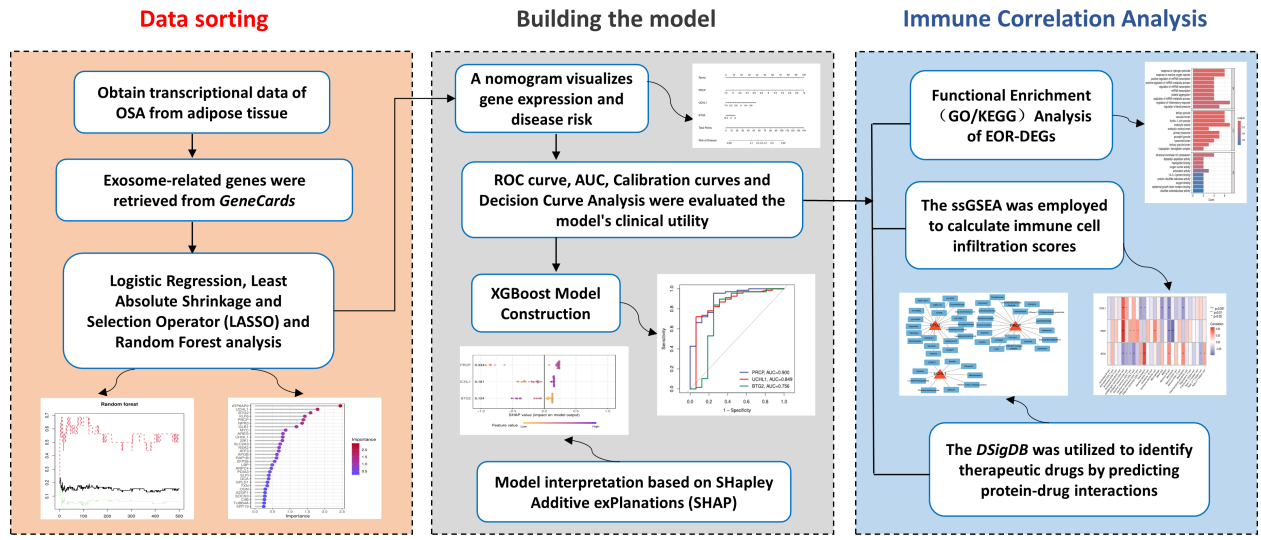
Weilong Ye†, Yitian Yang†, Feiju Chen†, Xiaoxi Lin, Yunan Wang, Lianfang Du, Jingjing Pan, Weifeng Liao, Bainian Chen, Riken Chen* and Weimin Yao*

The Second Affiliated Hospital of Guangdong Medical University, Zhanjiang, Guangdong, China

Background: Obstructive sleep apnea (OSA) is a prevalent disorder characterized by significant metabolic and immune dysregulation. This study aims to uncover exosome-related biomarkers implicated in immune-metabolic disturbances in OSA and explore their potential as diagnostic and therapeutic targets.
Methods: Transcriptomic data from two GEO datasets (GSE135917 and GSE38792) were integrated and analyzed using differential expression analysis via the *limma* package. Key biomarkers were identified using feature selection techniques including LASSO and Random Forest. Machine learning models, specifically XGBoost, were trained to evaluate biomarker performance, with model accuracy assessed by ROC curve analysis and AUC values. Immune cell infiltration was evaluated using single-sample Gene Set Enrichment Analysis (ssGSEA). Drug enrichment predictions were made through the Drug Signatures Database (DSigDB). Vivo and Vitro Experimental Validation on Multiple Independent cohorts.
Results: Three exosome-related biomarkers—PRCP, UCHL1, and BTG2—were identified as central to OSA's immune-metabolic dysregulation. XGBoost modeling demonstrated robust predictive power (AUC = 0.968). Immune analysis revealed significant correlations between gene expression and immune cell subsets, particularly CD56 bright natural killer cells and Memory B cells. Drug enrichment analysis identified potential therapeutic compounds, including Pentaphenate and Delphinidin, which target these biomarkers. OSA is associated with a reproducible transcriptional signature characterized by increased PRCP and UCHL1 expression and decreased BTG2 expression.
Conclusions: This study identifies PRCP, UCHL1, and BTG2 as key exosome-related biomarkers in OSA that regulate immune-metabolic disruption. By integrating transcriptomic data, machine learning, and immune analysis, we uncover an "exosome-immune" axis in OSA pathophysiology.

KEYWORDS

exosome signaling, obstructive sleep apnea (OSA), immune infiltration, machine learning, biomarkers



GRAPHICAL ABSTRACT

1 Introduction

An estimated 1 billion people worldwide are affected by OSA (1), and its prevalence continues to rise (2), primarily due to the increasing global rates of obesity (3–5). OSA is characterized by repeated partial or complete obstruction (collapse) of the upper airway during sleep, leading to hypercapnia, intermittent hypoxia (IH), and a reduction in blood oxygen saturation (2). The clinical diagnostic standard for OSA relies on polysomnography (PSG) (6). However, the high cost and time-intensive nature of this diagnostic method limit its application in the early screening and long-term treatment monitoring of OSA. As a result, identifying reliable biomarkers has become a research focus in the field of sleep medicine over the past decade (7). Pathophysiological studies suggest that IH, a core pathological feature of OSA, activates the sympathetic nervous system, induces metabolic disturbances, and promotes systemic inflammation and oxidative stress (8). Notably, IH exposure significantly upregulates the transcriptional activity of hypoxia-inducible factor-1α (HIF-1α), which then regulates a variety of downstream signaling pathways (9, 10). In terms of immune regulation, OSA patients exhibit characteristic proliferation of natural killer (NK) cells and natural killer T (NKT) cells (11). Further analysis shows that in mild to moderate cases, the proportion of CD4+ effector T cell subsets is abnormally elevated, while the numbers of effector memory T cells (TEM) and central memory T cells (TCM) are significantly reduced (12). Severe OSA cases display also pronounced immune dysregulation: the ratio of T helper (Th) cells to cytotoxic T lymphocytes (CTLs) decreases, while the number of B lymphocytes, which mediate humoral immunity, is significantly reduced (13). These findings suggest that the pathological progression of OSA involves complex immune cell dynamic imbalances, with characteristic immune phenotype changes observed at different stages of the disease. This provides potential targets for the development of novel diagnostic and therapeutic strategies.

Exosomes, key components of adipose-derived extracellular vesicles, play a crucial role in systemic metabolic regulation (14). These nanometer-sized vesicles, ranging from 30 to 150 nm in diameter, are rich in proteins and nucleic acids (including mRNA, miRNA, and lncRNA) derived from their parent cells (15). By mediating intercellular communication, metabolic waste clearance, and the maintenance of microenvironment homeostasis, exosomes significantly contribute to metabolic processes (16). Notably, exosome-carried metabolic regulatory factors can specifically bind to lipid transport proteins, modulating inflammatory cascades, immune response networks, and programmed cell death pathways (17, 18). This ultimately leads to pathological changes associated with metabolic disorders (19). Based on these functions, this study proposes an innovative hypothesis: intermittent hypoxia may alter the exosome secretion profile of adipose tissue, which in turn changes immune cell infiltration patterns, ultimately driving the pathological processes of OSA.

Current research has yet to fully elucidate the molecular mechanisms by which adipose-derived exosomes interact with metabolic regulation. Experimental evidence has shown that adipose tissue macrophages (ATMs) deliver miR-155 to adipocytes via exosomes, and this microRNA plays a significant role in improving obesity-related metabolic abnormalities by inhibiting the expression of peroxisome proliferator-activated receptor γ (PPARγ) (20). On the other hand, exosome-derived miR-34a from adipocytes has been shown to suppress M2 macrophage polarization, exacerbating the chronic inflammatory state induced by obesity (17). These findings suggest a bidirectional regulatory network between adipocytes and immune cells mediated by exosomes, offering a new perspective on the mechanistic study of metabolic diseases.

Building on this background, this study aims to adopt a comprehensive bioinformatics approach. First, it will screen OSA-specific exosome biomarkers and establish a machine learning-assisted diagnostic model. Second, the study will analyze the immune microenvironment of adipose tissue using the ssGSEA (single-sample Gene Set Enrichment Analysis) algorithm. Finally, we conducted *in vivo* and *in vitro* experimental validations across multiple independent cohorts and established a theoretical framework for the "hypoxia–exosome–immune" regulatory axis, thereby providing a solid foundation for the development of precise therapeutic targets.

2 Materials and methods

2.1 Collection and preprocessing of OSA transcriptomic data

The mRNA expression profiles for OSA were obtained from the GEO database, specifically datasets GSE135917 (21) and GSE38792 (22), both generated using the GPL6244 platform (Affymetrix Human Gene 1.0 ST Array). In the GSE135917 dataset, the control group included 8 samples, while the OSA group comprised 34 samples, with total RNA extracted from subcutaneous adipose tissue. Similarly, the GSE38792 dataset consisted of 8 control samples and 10 OSA patient samples, with RNA extracted from visceral adipose tissue biopsies collected during surgery. Log transformation was applied to both datasets to standardize expression values, followed by correction of distribution differences across samples. The datasets were then merged, and batch correction was performed to mitigate technical variations. Principal component analysis (PCA) was employed to visualize the differences between the two datasets before and after batch correction, ensuring improved data comparability.

2.2 Differential gene expression analysis and intersection with exosome-related genes

After data preprocessing, differential expression analysis was conducted using the *limma* package to compare gene expression profiles between control and disease groups, aiming to investigate the molecular mechanisms underlying sleep apnea. The *normalizeBetweenArrays()* function was applied to standardize the data. Subsequently, further analysis was performed using linear modeling: the *lmFit()* function was employed to fit a generalized linear model, *makeContrasts()* was used to construct a contrast matrix defining specific comparisons, followed by *contrasts.fit()* for contrast analysis, and finally, *eBayes()* was applied for empirical Bayesian adjustment to enhance the robustness and accuracy of statistical inference. The filtering criteria included an adjusted p-value < 0.05 and $|\log2FC| > 0.5$ (approximately corresponding to a 1.41-fold change). This threshold was chosen based on established practices in similar studies (23, 24), as microarray data typically reveal subtle expression changes, with a $|\log2FC| > 0.5$ regarded as a meaningful difference. The resulting differentially expressed genes were visualized using a heatmap. Exosome-related genes were retrieved from the *GeneCards* database, a publicly available resource for human gene information (https://www.genecards.org/). We selected genes with Relevance Score > 2 as strongly associated genes, which accounted for more than 50 percent of the total. A Venn diagram was then constructed to visualize the intersection between exosome-related genes and differentially expressed genes, highlighting those with potential relevance to the study.

2.3 Functional enrichment analysis of EOR-DEGs

To explore the functional roles of exosome-related differentially expressed genes (EOR-DEGs), Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were performed using the *clusterProfiler* package (25). Enrichment was considered significant when both p-values and adjusted p-values were less than 0.05. GO analysis encompassed 3 domains: biological processes (BP), cellular components (CC), and molecular functions (MF). The results of these enrichment analyses were visualized using bar-plots to highlight significant pathways and cnet-plots to illustrate the relationships between genes and their associated terms.

2.4 Logistic regression analysis and feature selection of EOR-DEGs

To assess the prognostic and diagnostic value of EOR-DEGs, univariate logistic regression was first applied, with the odds ratio (OR) and p-value used to identify genes significantly associated with prognosis and diagnosis ($p < 0.05$). Genes meeting this threshold were then subjected to feature selection using Least Absolute Shrinkage and Selection Operator (LASSO) regression (λ. min) and Random Forest (RF) analysis (Importance > 4) (26). The overlap of selected genes from both methods was visualized using a Venn diagram, identifying a set of key biomarkers for further clinical and mechanistic analysis. Subsequently, box plot was used to illustrate the expression levels of feature genes across different groups, and correlation plot was employed to visualize their interrelationships.

2.5 Construction and evaluation of a diagnostic model

A nomogram was developed to visualize the relationship between feature gene expression and disease risk, with coefficients derived from multivariate logistic regression. The model's performance was evaluated using the Receiver Operating Characteristic (ROC) curve, with the area under the curve (AUC) indicating predictive accuracy. Calibration curves were constructed to assess the agreement between predicted and observed outcomes, while Decision Curve Analysis (DCA) evaluated the model's clinical utility by assessing net benefit at various threshold probabilities.

2.6 XGBoost model construction

The XGBoost algorithm (27) was selected for its efficiency and robust performance in binary classification tasks. The feature genes were set as the predictors, with occurrence of OSA acting as response variable. The model's predictive performance was evaluated using ROC curves. To minimize overfitting, 5-fold cross-validation was performed during model validation, alongside a reduced learning rate and limited maximum depth.

2.7 Model interpretation based on SHAP

We calculated SHAP (SHapley Additive exPlanations) (28) values to interpret the XGBoost model. The SHAP summary plot visualized their relative importance. Dependency plots were generated to illustrate the relationship between gene expression levels and disease risk. Additionally, SHAP force plots were used to analyze individual patient predictions, offering detailed insights into the gene-specific contributions to the probability of OSA occurrence.

2.8 Immune correlation analysis

The ssGSEA was employed to calculate immune cell infiltration scores, which were subsequently correlated with the expression of feature genes. Spearman's correlation method was used to assess the relationship between immune cell activity and gene expression, with statistical significance determined for each correlation. The results were visualized in a heatmap, where the strength and significance of the correlations were clearly represented. The 28 immune cell–related gene sets were obtained from previously published studies (29, 30).

2.9 Drug enrichment analysis

The Drug Signatures Database (DSigDB) was utilized to identify potential therapeutic drugs by predicting protein-drug interactions. The DSigDB online platform (https://dsigdb.tanlab.org/), a publicly accessible database that integrates drug-associated gene expression data, was employed to explore drug-gene relationships, mechanisms of drug action, and opportunities for drug repurposing (31). Candidate drugs were identified by comparing the database's drug gene expression signatures with disease-related gene expression profiles. The results were visualized using Cytoscape software (https://cytoscape.org/).

2.10 Vivo and vitro experimental validation

Human SW872 liposarcoma cells (n=6) and murine 3T3-L1 preadipocytes (n=6) were cultured under standard conditions, with 3T3-L1 cells induced to differentiate into mature adipocytes using a commercial induction kit. Male C57BL/6J mice (8 weeks of age; n=10) were randomly assigned to normoxia or chronic CIH exposure. This experiment was reviewed and approved by the Animal Welfare and Ethics Committee under review number: IACUC-20250701-299. Cells and mice were exposed to intermittent hypoxia (IH/CIH) with cyclic oxygen fluctuations, while controls were maintained under normoxia. Total RNA was extracted from cells and mouse adipose tissues, reverse-transcribed into cDNA, and analyzed by SYBR Green-based qRT-PCR. Gene expression was quantified using the 2^(-ΔΔCT) method with GAPDH as the internal control, and all reactions were performed in triplicate to ensure reliability. IHC was performed on FFPE iWAT sections using antibodies against PRCP, BTG2, and UCHL1, with DAB visualization, and staining was quantified as percentage positive area using ImageJ. The detailed methodological section has been added in the supplementary file. A brief overview of the process is presented in the Graphical Abstract.

2.11 Statistical analysis

The entire analysis was conducted using R software (version 4.4.2). During data collection, the *GEOmirror* and *idmap2* packages facilitated data retrieval and annotation. The *limma* and *sva* packages were utilized for dataset organization, correction, merging, and differential expression analysis. To ensure uniform distribution of expression values across all samples, quantile normalization was applied using the *normalize-Between-Arrays* function. After merging the datasets, the *ComBat* method was employed to correct for batch effects. For visualization, box plots were generated, and the Wilcoxon rank-sum test was applied for group comparisons. Correlation analysis was performed using Pearson's correlation coefficient to assess the relationships between gene expression levels. Data are presented as mean ± SEM; qPCR was analyzed using per-sample ΔCt values and reported as 2^−ΔΔCt, while IHC results were quantified as the percentage of positive area (area%) per sample across predefined fields, with two groups compared using unpaired two-tailed t-tests, multiple groups analyzed by one-way ANOVA with appropriate *post-hoc* tests, and non-parametric alternatives applied when assumptions of normality or homoscedasticity were not met; $P < 0.05$ was considered statistically significant.

3 Results

3.1 Data integration and differential expression analysis

PCA demonstrated that batch correction effectively mitigated batch effects, thereby preserving the integrity of the biological signal (Figures 1A, B). Addressing batch effects is crucial to minimize non-biological variations that could otherwise compromise the reliability of downstream analyses. Utilizing thresholds of p-value < 0.05 and |log2FC| > 0.5, 245 differentially expressed genes were identified (Figure 1C). A heatmap showcasing the top 50 upregulated or downregulated genes, ranked by |log2FC|, provides a visual representation of the key expression changes (Figure 1D). Using the keyword "exosome," 5,293 protein-coding genes were identified, of

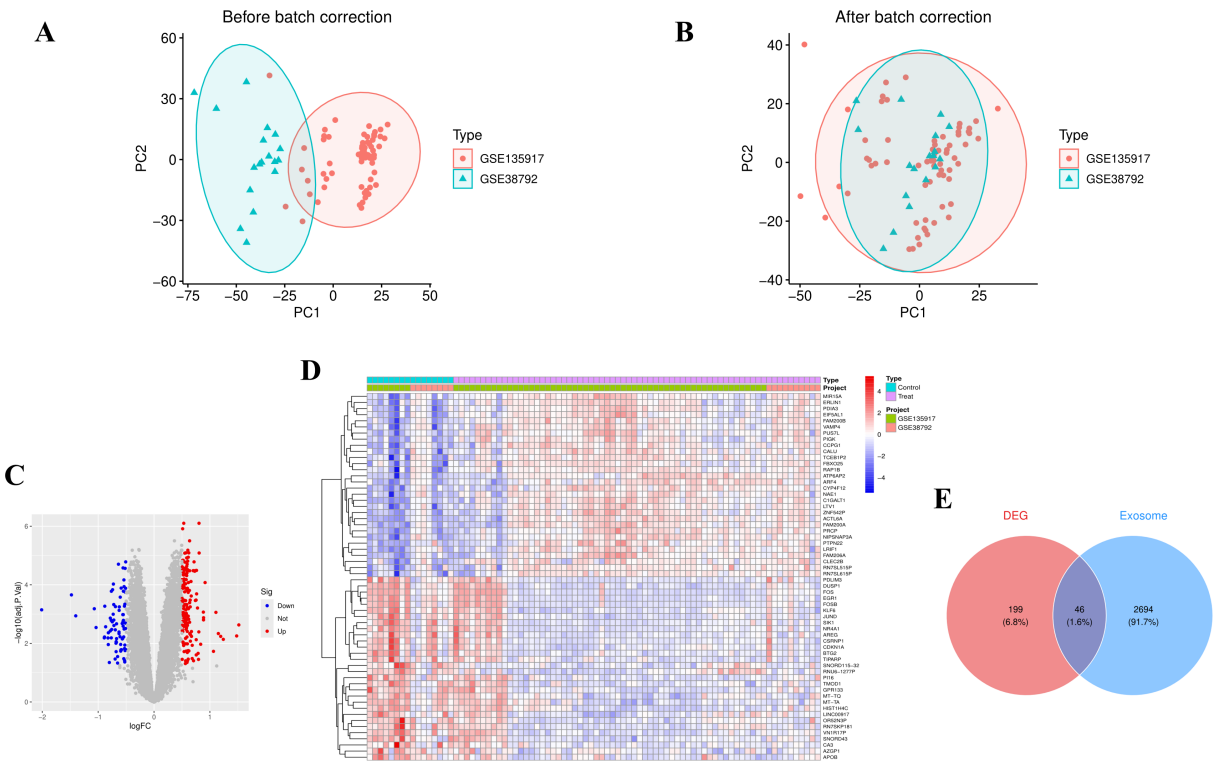


FIGURE 1
Data preprocessing and differential gene screening. **(A)** PCA plot before batch correction showed clustering by dataset origin. **(B)** PCA plot after batch correction, demonstrating clustering by disease status, indicating the removal of technical variations. **(C)** Volcano plot of differentially expressed genes (red dots: upregulated genes; blue dots: downregulated genes; thresholds: p-value < 0.05 and |log2FC| > 0.5), identifying 245 differentially expressed genes (DEGs). **(D)** Heatmap displaying the top 50 DEGs (ranked by |log2FC|), with row-normalized expression values (Z-score) reflecting expression patterns between OSA and control groups. **(E)** Venn diagram showing 46 exosome-related differentially expressed genes (EOR-DEGs, intersection of Gene Cards exosome gene set and DEGs).

which 2,740 genes with a Relevance Score > 2 were selected for further analysis. The Venn diagram displayed 46 EOR-DEGs (Figure 1E).

3.2 Functional enrichment analysis for EOR-DEGs

The GO terms with the highest number of enriched genes in BP, CC, and MF were: regulation of inflammatory response, endocytic vesicle, and structural constituent of the cytoskeleton (Figure 2A). In the KEGG analysis, relatively few pathways were enriched (with a p-value < 0.05 and an adjusted p-value < 0.05 with a primary focus on lipid metabolism and atherosclerosis pathways (Figure 2B). The GO analysis network plot highlights the top 10 most significant functional enrichment categories (Figure 2C). The KEGG path view suggests that LBP, MMP9, APOB, IL6, and RAP1B are involved in lipid metabolism and atherosclerosis (Figure 2D).

3.3 Logistic regression analysis and feature selection of EOR-DEGs

In univariate logistic regression analysis, all 46 EOR-DEGs had p-values less than 0.05. Among them, 20 genes had odds ratios (ORs) less than 1, while the remaining genes showed ORs greater than 1 (Table 1). These 46 EOR-DEGs were further included in LASSO analysis to address potential collinearity, resulting in the selection of 10 genes (Figures 3A, B). Random forest analysis was employed to determine gene importance, with genes having importance scores greater than 4 being highlighted (Figures 3C, D). The intersection of genes identified through LASSO and RF analyses revealed three feature genes for model construction: PRCP, UCHL1, and BTG2. The box plot indicated that PRCP and UCHL1 were highly expressed in the OSA group, while BTG2 showed lower expression (Figure 3E). Correlation analysis revealed that UCHL1 was negatively correlated with both BTG2 and PRCP (Figure 3F).

3.4 Construction and evaluation of a diagnostic model

The nomogram visually represents a diagnostic model constructed through multivariate logistic regression analysis, leveraging the expression levels of hub genes to predict the risk of OSA (Figure 4A). The ROC curve (Bootstrapping method) demonstrates the model's superior diagnostic performance, with an AUC value exceeding that of individual genes, confirming its

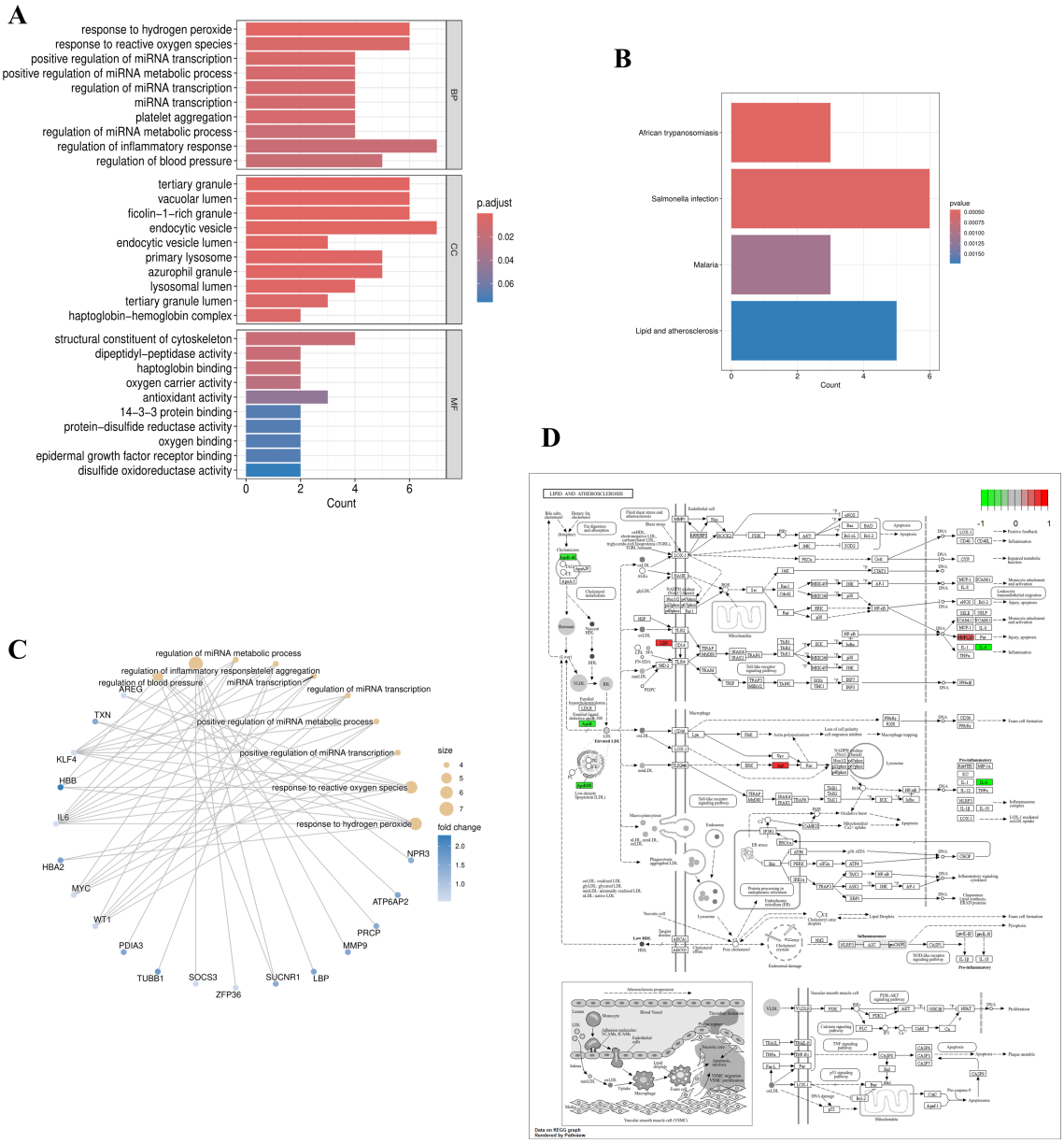


FIGURE 2
Functional and pathway enrichment of EOR-DEGs. **(A)** GO enrichment bar plot with significant terms ($p < 0.05$) including "regulation of inflammatory response" (BP), "endocytic vesicle" (CC), and "structural constituent of cytoskeleton" (MF). **(B)** KEGG pathway enrichment highlighted significant pathways such as "lipid metabolism" and "atherosclerosis". **(C)** GO network diagram displaying the top 10 enriched terms, where node size represents the number of genes and edge width indicates gene overlap. **(D)** KEGG pathway map (lipid metabolism) highlighting key genes (LBP, MMP9, APOB, IL6, RAP1B).

robustness (Figures 4B, C). Model evaluation through the calibration curve indicates that the bias-corrected curve closely parallels the ideal curve, with only minor deviations observed in the high-probability range (approaching 1.0) (Figure 4D). Additionally, the Decision Curve Analysis (DCA) reveals that employing the model for prediction and intervention provides a higher net benefit (Figure 4E). These findings underscore the model's reliability and practical utility in diagnostic applications. In the XGBoost model, the AUC value reached 0.968 (Figure 4F). To mitigate overfitting, 5-fold cross-validation was performed, further validating the model's robustness (AUC = 0.989) (Figure 4G). Additionally, a feature importance plot was generated to illustrate the contribution of each gene to the model's predictions (Figure 4H). For the XGBoost model, we set the following hyperparameters: learning rate (eta) = 0.01, maximum tree depth = 2, minimum child weight = 2, gamma = 0.1, subsample = 0.8, colsample_bytree = 0.8, lambda = 1, and alpha = 0. The objective was to use a smaller learning rate and limit model

10.3389/fimmu.2025.1587522

TABLE 1 Univariate logistic regression analysis.

Gene	OR	OR.95L	OR.95H	p-value
PRCP	986.0791	60.17411	43735.75	2.90E-05
ATP6AP2	114.3371	14.29222	1758.272	8.39E-05
PDIA3	75.6924	11.11494	951.5929	0.000113
ARF4	47.53213	8.308052	428.4111	9.42E-05
GLB1	38.61957	7.375862	314.5474	0.000108
EXOSC3	27.94962	5.829075	190.7486	0.000158
UCHL1	27.675	6.2062	175.6567	7.80E-05
RAP1B	22.04958	4.769966	176.006	0.000579
ARPC4	15.57668	4.089505	76.78485	0.000201
TUBB4A	11.25975	3.354366	44.70034	0.000205
GCA	10.92942	3.585175	44.0934	0.000151
NSA2	8.126889	2.39504	39.09922	0.003046
SUCNR1	6.522213	2.137218	24.80007	0.002313
ALCAM	6.172922	2.180943	21.61758	0.001649
TUBB1	5.6868	2.185966	17.86299	0.001091
GPLD1	5.651804	2.053642	18.85715	0.00191
TXN	5.265804	1.945368	18.08795	0.003045
LBP	5.037635	1.938829	15.97191	0.002282
GPC4	4.393301	1.675899	13.27509	0.004583
HLA-DRB5	3.279293	1.375657	10.54385	0.019915
LYZ	2.892675	1.350791	6.975785	0.01005
CHI3L1	2.632013	1.409485	5.399355	0.004322
NPR3	2.523628	1.203417	5.78103	0.018065
MMP9	2.396216	1.199165	5.348592	0.020294
HBA2	2.236278	1.21276	4.4752	0.014438
HBB	2.092857	1.330592	3.531524	0.002651
ITLN1	0.62736	0.378592	0.949145	0.039053
C4B	0.477884	0.214379	0.907881	0.036145
IL6	0.477017	0.278475	0.779596	0.004177
SLPI	0.432136	0.19901	0.824733	0.017656
SLC2A3	0.392045	0.206861	0.702392	0.002305
KLF4	0.365831	0.186386	0.678436	0.00198
SOCS3	0.319353	0.147249	0.636311	0.001906
ZFP36	0.31906	0.148147	0.64201	0.001953
OGN	0.303189	0.124932	0.647185	0.003842
AREG	0.292776	0.124983	0.604413	0.001861
ATF3	0.283908	0.127676	0.582859	0.000993
MYC	0.270612	0.108965	0.590503	0.002069
WT1	0.259777	0.092083	0.626412	0.005114
DPP4	0.245084	0.073703	0.637639	0.009457
KRT19	0.215046	0.059646	0.574421	0.007162
APOB	0.158699	0.050603	0.423645	0.000567
AZGP1	0.138617	0.036829	0.396253	0.000937
SIK1	0.099512	0.026518	0.308981	0.000188
BTG2	0.068947	0.014537	0.254142	0.000202
KLF6	0.050211	0.009535	0.202541	9.92E-05

complexity to prevent overfitting when analyzing relatively small sample sizes.

3.5 Interpreting the machine learning model with SHAP analysis

The SHAP Summary Plot illustrates the contributions of the 3 hub genes (PRCP, UCHL1, and BTG2) to the overall model prediction (Figure 5A). Among them, PRCP shows the highest average SHAP value (0.334), indicating its strongest influence on the model's predictions, while UCHL1 and BTG2 have relatively smaller contributions. The Dependence Plot visualizes the relationship between the expression levels of the feature genes and their corresponding SHAP values. For instance, the SHAP value for BTG2 peaks around an expression level of 8.5, decreasing as the expression level increases, suggesting a nonlinear relationship that may reflect BTG2's complex regulatory role in the model's output (Figure 5B). The figure also shows the relationships for PRCP and UCHL1 (Figures 5C, D). According to the SHAP Force Plot, in one control sample, the hub genes exhibit a negative contribution to the occurrence of OSA (Figure 5E). These results suggest that BTG2 might act as a protective gene, with low expression increasing risk, while UCHL1 and PRCP may serve as risk genes, with higher expression correlating with an increased risk of OSA.

3.6 Correlation analysis of immunity

The ssGSEA analysis was employed to calculate the immune cell infiltration scores, which were subsequently correlated with the expression levels of the hub genes. UCHL1 and PRCP demonstrated a strong positive correlation with CD56 bright natural killer cells and a significant negative correlation with Memory B cells. In contrast, BTG2 exhibited an inverse correlation pattern. A heatmap was generated to visualize the correlations between the other immune cells and the feature genes, providing a comprehensive overview of the immune landscape associated with these feature genes (Figure 6A).

10.3389/fimmu.2025.1587522

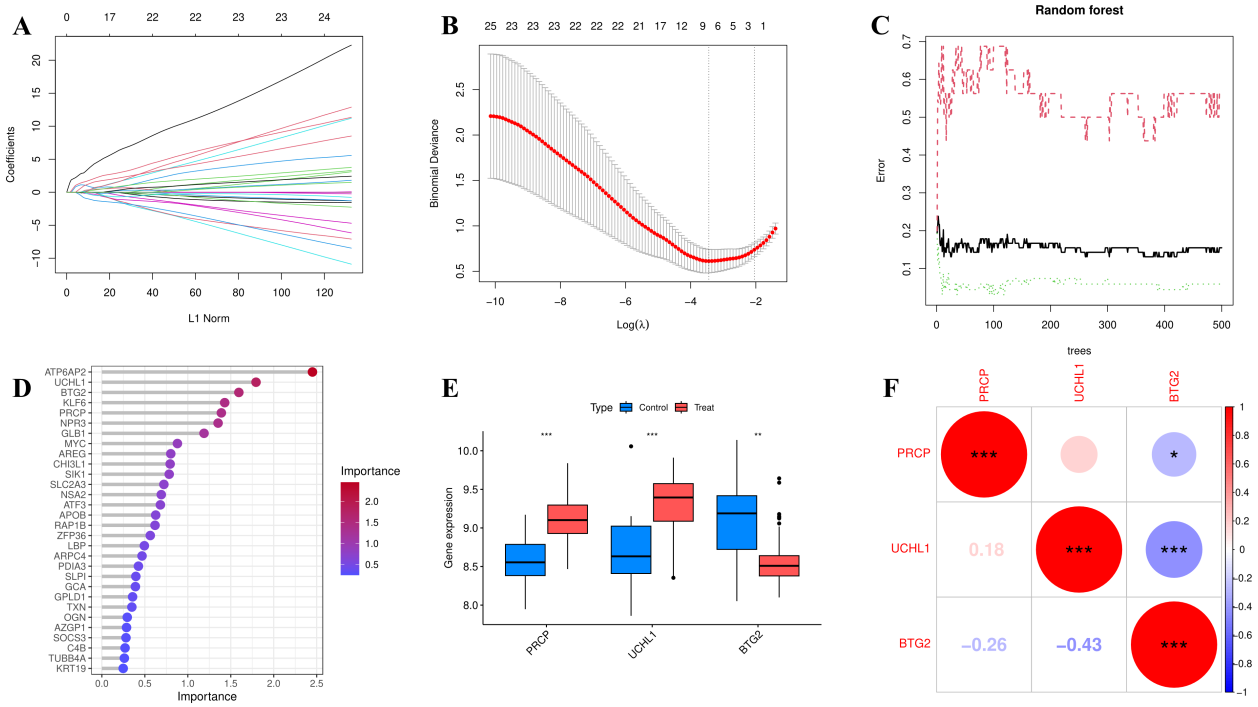


FIGURE 3
Feature selection was performed using LASSO (10-fold CV, lambda. min) and random forest (optimal trees, Mean Decrease Gini > 4). **(A)** coefficient path (lasso). **(B)** cross-validation error (lasso). **(C)** error rate curve (RF). **(D)** variable importance ranking (RF). **(E)** Boxplot showing the expression differences of feature genes (PRCP, UCHL1, BTG2) between the OSA group and the control group (*p<0.05, **p<0.01, ***p<0.001). **(F)** Heatmap of feature gene expression correlations (Pearson correlation coefficient).

3.7 Drug enrichment analysis

Using the gene IDs PRCP, UCHL1, and BTG2 as input in an online platform, 69 potential drugs were identified (p-value < 0.05). Among these, compounds such as Pentaphenate and Delphinidin exhibited significant associations with specific genes like PRCP and BTG2. Functional enrichment analysis highlighted their fold enrichment, z-scores, and adjusted p-values, suggesting that these compounds may exert critical biological effects on the related genes. Moreover, they could play regulatory roles in specific biological processes, providing insights into potential therapeutic applications. The network diagram illustrates the connections between identified drugs and their corresponding feature genes (Figure 6B). Each node represents a drug or a gene, with a maximum of 20 drugs displayed per gene.

3.8 Vivo and vitro experimental validation on multiple independent cohorts

To determine whether OSA induces transcriptional alterations in stress and metabolism-related genes, we performed qRT-PCR in multiple independent cohorts. We focused on PRCP, UCHL1, and BTG2, given their previously reported roles in proteolytic regulation, protein homeostasis, and cell cycle control. As shown in Figures 7A–C (Differentiated 3T3-L1 murine adipocytes), PRCP and UCHL1 mRNA levels were significantly elevated in the OSA group compared with controls, whereas BTG2 mRNA expression was markedly reduced. These findings were consistently reproduced in independent experimental sets (Figures 7D–F, SW872 human adipocytes; Figures 7G–I, eWAT), where PRCP upregulation was highly significant (Figures 7D, G), UCHL1 expression was robustly increased (Figures 7E, H), and BTG2 levels were consistently downregulated across all comparisons (Figures 7F, I). Notably, the concordant results across independent replicates underscore the stability and reproducibility of these transcriptional changes. Collectively, these data indicate that OSA is associated with a reproducible transcriptional signature characterized by increased PRCP and UCHL1 expression and decreased BTG2 expression. IHC: Representative micrographs (Figures 7Ja–f) and semi-quantitative analysis of DAB-positive area (%) (Figures 7Jg–i) revealed group-dependent differences (n = 3 per group). Compared with the control (CON) group, OSA samples exhibited significantly higher UCHL1 and PRCP expression and markedly lower BTG2 levels.

4 Discussion

This study elucidates the potential pathogenesis of OSA through adipose tissue transcriptomics, revealing PRCP, UCHL1, and BTG2 as exosome-associated hub genes that orchestrate metabolic-immune dysregulation. By synergizing cross-platform data integration (GSE135917/GSE38792), machine learning-

10.3389/fimmu.2025.1587522

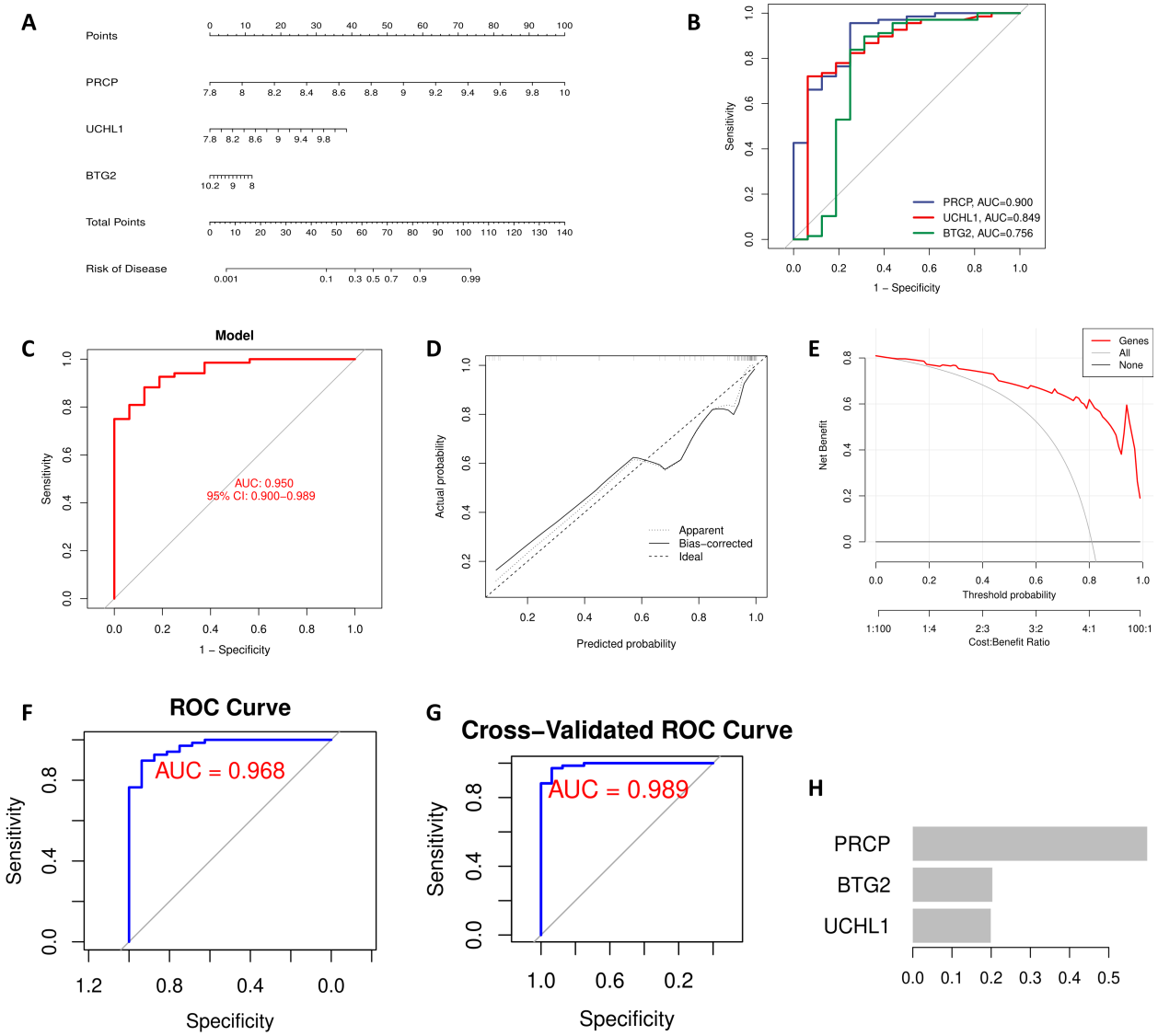


FIGURE 4
Feature gene-based OSA diagnostic model. **(A)** The nomogram integrates the expression levels of PRCP, UCHL1, and BTG2 to predict OSA risk, with the total score corresponding to the right-side risk axis. **(B)** ROC curve showing the performance of individual genes in predicting OSA. **(C)** ROC curve showing the performance of the combined diagnostic model based on feature genes. **(D)** Calibration curve with Bootstrap = 1000 iterations. The dashed line represents the ideal fit, and the solid line represents the model's bias-corrected prediction. **(E)** DCA showing the net clinical benefit of the model when the threshold probability exceeds 10%. XGBoost Model Validation: **(F)** ROC curve. **(G)** 5-fold cross-validation. **(H)** Feature importance plot.

driven biomarker discovery (XGBoost AUC = 0.968), and single-sample immune deconvolution, we reveal an unprecedented "exosome-immune" axis in OSA pathophysiology. Our robust feature selection pipeline—incorporating LASSO regularization (λ. min) and random forest permutation importance—convergently identified PRCP (prolyl carboxypeptidase), UCHL1 (Ubiquitin C-Terminal Hydrolase L1), and BTG2 (B-cell translocation gene 2) as key non-redundant classifiers, validated through SHAP interpretability to dissect nonlinear gene-disease interactions (SHAP value for PRCP: 0.334). These results not only demonstrate the diagnostic potential of these biomarkers but also highlight the utility of interpretable machine learning techniques in elucidating complex biological relationships (32).

Prolyl carboxypeptidase (PRCP), a serine protease, exerts regulatory effects across multiple endocrine axes including the renin-angiotensin system (RAS), kallikrein-kinin system (KKS), and pro-opiomelanocortin (POMC) (33). Study had demonstrated that PRCP plays a crucial role in the onset and progression of obesity, regulating the balance between energy intake and expenditure through an α-MSH1-mediated mechanism (34). The coexistence of obesity and OSA is commonly observed, with a bidirectional relationship between the two conditions (35). UCHL1, a key member of the deubiquitinating enzyme family, influences cell proliferation, differentiation, and damage by modulating both ubiquitination and non-ubiquitination pathways (36, 37). Notably, HIF-1α has been identified as a potential target

10.3389/fimmu.2025.1587522

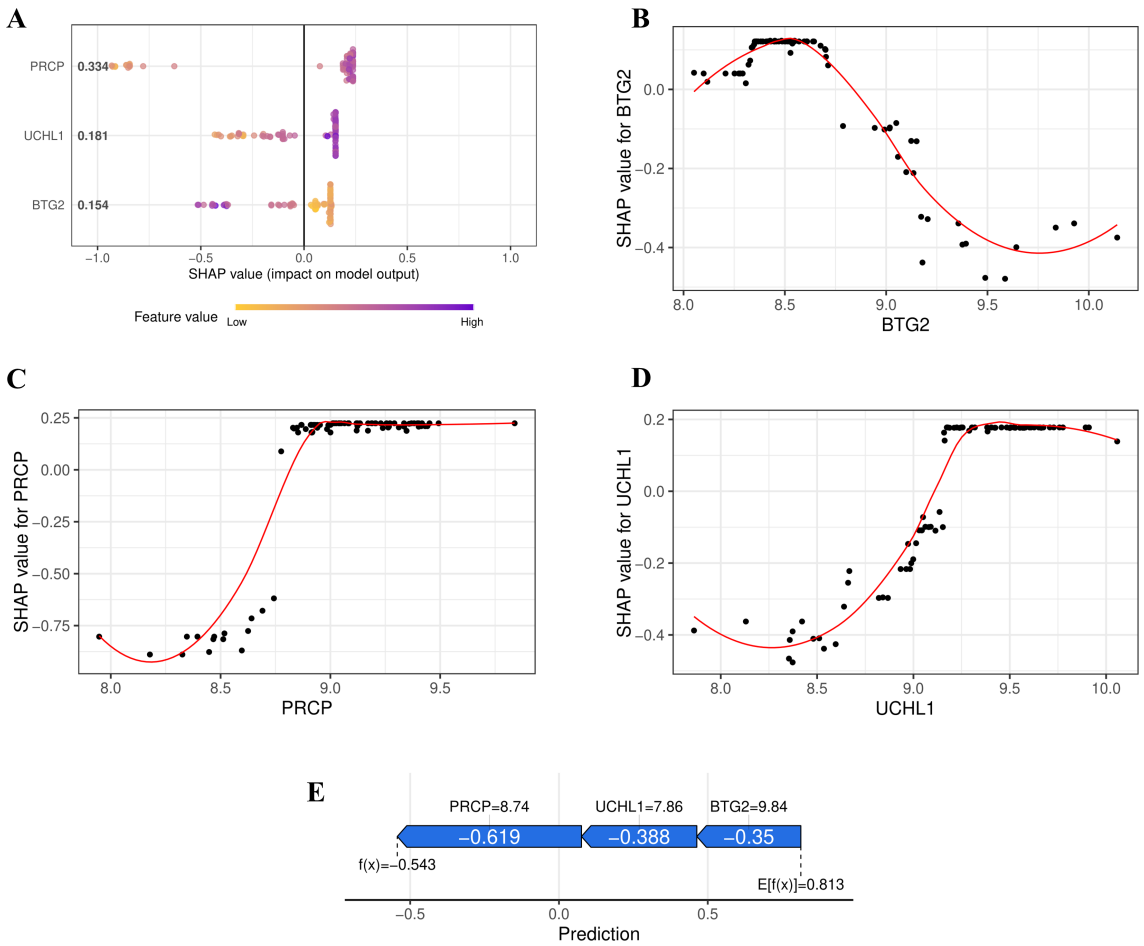


FIGURE 5
SHAP Interpretation of the XGBoost Model. **(A)** Summary plot illustrating the contributions of PRCP, UCHL1, and BTG2 to the overall prediction of OSA risk. **(B)** Relationship between BTG2 Expression Level and SHAP Value (LOESS fitting curve). **(C)** Relationship between PRCP Expression Level and SHAP Value. **(D)** Relationship between UCHL1 Expression Level and SHAP Value. **(E)** SHAP force plot illustrating the contribution of feature genes to the OSA risk prediction for a control sample.

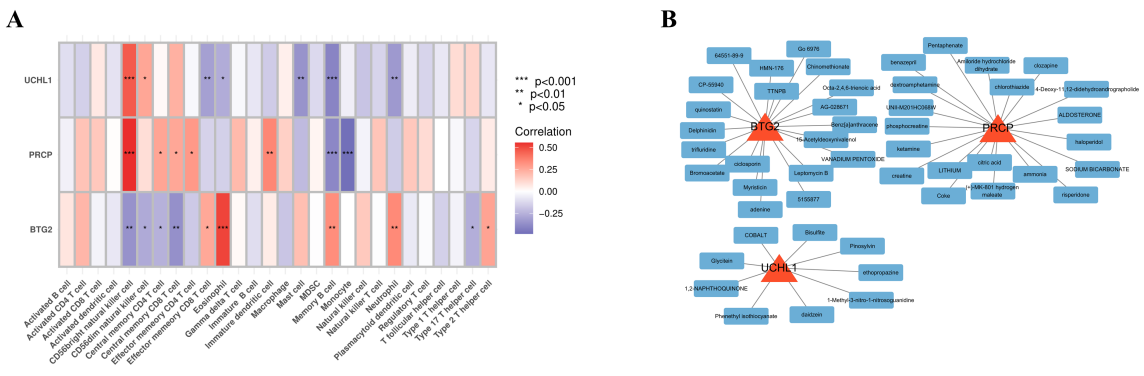


FIGURE 6
Immune association and potential targeted drugs of feature genes. **(A)** Heatmap depicting the correlations between immune cell infiltration and feature gene expression. **(B)** Network of drug-gene interactions, visualized using Cytoscape, showing potential therapeutic drugs targeting the feature genes.

10.3389/fimmu.2025.1587522

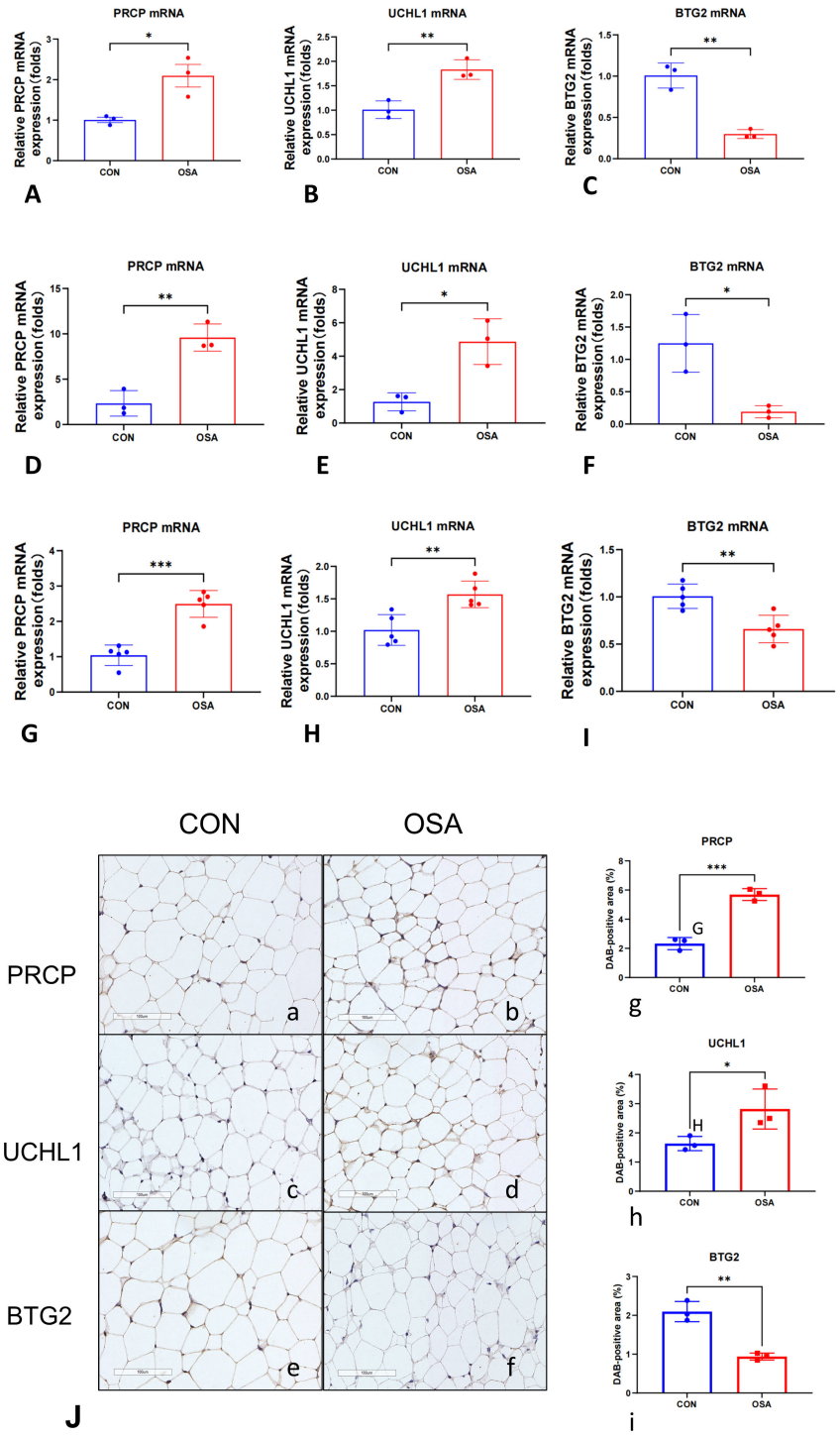


FIGURE 7
Intermittent hypoxia induces a conserved cellular stress signature across adipocyte models *in vitro* and *in vivo*. Relative mRNA expression of PRCP, UCHL1, and BTG2 was measured by qRT-PCR. **(A–C)** Differentiated 3T3-L1 murine adipocytes (n = 3 per group). **(D–F)** SW872 human adipocytes (n = 3 per group). Cells were exposed to 24 hours of normoxia or intermittent hypoxia (IH). **(G–I)** Epididymal white adipose tissue (eWAT) from mice exposed to 4 weeks of normoxia (Control, n = 5) or chronic intermittent hypoxia (CIH, n = 5). Gene expression was normalized to Actb. Data are presented as mean ± SEM. Statistical significance was assessed by unpaired, two-tailed Student's t-test on ΔCt values. (**J**, a, b) Representative IHC images of PRCP, (**J**, c, d) UCHL1, (**J**, e, f) BTG2. Brown DAB precipitate indicates positive staining; nuclei are counterstained blue (scale bar = 100 μm). (**J**, g–i) Quantification of DAB-positive area (%). Each dot represents one independent sample; bars denote mean ± SD. *p < 0.05, **p < 0.01, ***p < 0.001.

interacting with UCHL1, and under hypoxic conditions, UCHL1 may regulate the nuclear translocation of HIF-1α, influencing its role in cellular responses to low oxygen levels (38). In OSA patients, IH activates HIF-1α, which in turn triggers systemic inflammation and disrupts hepatic lipid metabolism (39–41). BTG2, a member of the ERBB2 (BTG/TOB) family, functions as a B-cell transducer and regulator (42). Research has shown that Btg2 expression is elevated in the subcutaneous adipose tissue of obese mice on a high-fat diet, highlighting its involvement in lipid metabolism during obesity and metabolic disorders (43). Specifically, Btg2 reduces interleukin-6 expression by inhibiting the Stat3 signaling pathway, which plays a pivotal role in adipocyte differentiation (44, 45).

Moreover, our immune correlation analysis using ssGSEA revealed significant associations between the expression of the hub genes and various immune cell populations. Specifically, UCHL1 and PRCP showed strong positive correlations with CD56 bright natural killer cells and significant negative correlations with Memory B cells, whereas BTG2 exhibited an opposing pattern. During the differentiation process of monocytes into M1 macrophages, a significant upregulation of PRCP activity is observed (46). Studies have shown that human blood-derived alveolar macrophages exhibit higher PRCP activity (47, 48). Given that M1 macrophages are defined as pro-inflammatory macrophages, this suggests that PRCP plays a key role in the inflammatory response mechanism (46). Additionally, PRCP is also highly expressed in human neutrophils (49). UCHL1 primarily promotes the polarization of M1 macrophages by regulating the PI3K/AKT signaling pathway (50). It can also modulate the inflammatory response in lipopolysaccharide (LPS)-activated macrophages through MAPK and NF-κB signaling pathways (51). BTG2 mainly by controlling cell proliferation and activation processes to maintain T cell quiescence (42). Moreover, the protein complex formed by BTG2 and PRMT1 can effectively counteract the proliferation activity of pre-B cells, thus promoting the development of B cells (52). These findings provide solid evidence supporting the theory that exosome-related genes are involved in immune regulation, fully revealing their key positions and mechanisms of action within the immune regulation network.

In addition, drug enrichment analysis using the DSigDB platform identified several candidate compounds, such as Pentaphenate and Delphinidin, that significantly interact with the hub genes. Previous study had shown that PRCP, through its involvement in the pro-opiomelanocortin (POMC) system, makes it a highly promising target in the treatment of obesity and related diseases (34, 53). *In vitro* and *in vivo* experiments indicate that myricetin may influence the lipid metabolic process in the adipose tissue of obese mice by regulating the expression levels of miR-222 and its target gene BTG2 (54). These potential therapeutic agents may modulate exosome-mediated signaling and immune responses, offering promising avenues for targeted intervention in OSA.

While our machine learning approaches provides novel insights, limitations warrant consideration. First, the analyses were based exclusively on adipose tissue transcriptomic data, which may not fully reflect the systemic pathophysiology of OSA involving airway, liver, and circulating immune cells. Second, the relatively small sample size (n=60) may limit the generalizability of the results, underscoring the need for validation in larger, multi-center cohorts. Finally, the DSigDB-based drug predictions require experimental confirmation of target engagement and efficacy.

Based on previous research, we have developed an innovative hypothesis: the "Exosome-Immune Axis in the Pathogenesis of OSA." During the progression of OSA, IH likely activates the HIF-1α signaling pathway in adipose tissue, leading to the release of pathological exosomes. These exosomes carry key regulatory molecules such as PRCP, UCHL1, and BTG2, initiating a vicious cycle of "hypoxia-exosome-immune metabolic disorder." In terms of specific mechanisms, PRCP in the exosomes may enhance the differentiation of M1 macrophages and disrupt the normal metabolism of α-MSH, thereby triggering a systemic inflammatory response. UCHL1 may regulate the nuclear translocation of HIF-1α and activate the PI3K/AKT signaling pathway, further exacerbating M1 macrophage polarization and suppressing NK cell activity. BTG2 primarily affects lipid metabolism via the STAT3 signaling pathway and, through the BTG2-PRMT1 protein complex, promotes the differentiation and maturation of B cells. This model comprehensively integrates interactions involving "hypoxia-exosome"-mediated signaling, immune cell functional remodeling, and metabolic disruption, offering a promising new research direction for a deeper understanding of the systemic pathological mechanisms of OSA. Finally, the proposed "Hypoxia-Exosome-Immune Axis" represents a hypothesis derived from bioinformatics associations rather than demonstrated causal relationships, and its mechanistic details await functional validation.

In summary, this study is the first to identify PRCP, UCHL1, and BTG2 as exosome-based biomarkers associated with the diagnosis of OSA. These biomarkers are closely linked to immune-metabolic imbalance in the body. The findings not only uncover key molecular nodes involved in immune-metabolic disruption in the pathogenesis of OSA but also provide potential theoretical support and direction for the development of targeted therapeutic strategies based on the OSA exosome-immune axis.

5 Conclusion

This study identifies PRCP, UCHL1, and BTG2 as key exosome-related biomarkers in OSA that contribute to immune–metabolic dysregulation. By integrating transcriptomic data, machine learning, immune profiling, and *in vivo* and *in vitro* validations across multiple independent cohorts, we reveal an "exosome–immune" axis underlying OSA pathophysiology.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Ethics statement

The animal study was approved by Mudanjiang Medical University welfare and ethical review of experimental animals. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

WYe: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Project administration, Writing – original draft. YY: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Project Administration, Writing – original draft. FC: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Project administration, Writing – original draft. XL: Conceptualization, Resources, Investigation, Writing – original draft. LD: Conceptualization, Resources, Investigation, Writing – original draft. YW: Conceptualization, Software, Writing – original draft. JP: Conceptualization, Software, Writing – original draft. WL: Conceptualization, Software, Writing – original draft. BC: Conceptualization, Software, Writing – original draft. WYa: Conceptualization, Funding acquisition, Project administration, Writing – review & editing. RC: Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This study was funded by the Natural Science Foundation of Basic and Applied Basic Research Fund of Guangdong Province (2022A1515012375), the Guangdong Medical Research Fund Project (A2024728, A2024723), the Science and Technology Development Special Project of Zhanjiang City (2022A01142, 2022A01110, 2021A05086, 2021A05088), the clinical and Basic Technology Innovation Program of Guangdong Medical University (GDMULCJC2024063, GDMULCJC2024064, GDMULCJC2025096).

Acknowledgments

We thank Dr. Jianming Zeng (University of Macau), and all the members of his bioinformatics team, biotrainee, for generously sharing their experience and codes. We also thank Huan Li, and Dongjie Huang (The Second Affiliated Hospital of Guangdong Medical University) assisted in reviewing the Methods and Results sections related to the in vivo and in vitro models and contributed to the selection of representative immunohistochemistry images. They also provided support in the statistical analysis of the immunohistochemistry data. In addition, Liu Yong (The Second Affiliated Hospital of Guangdong Medical University) prepared the documentation for animal welfare and ethical review.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2025.1587522/full#supplementary-material

References

1. Lyons MM, Bhatt NY, Pack AI, Magalang UJ. Global burden of sleep-disordered breathing and its implications. *Respirology (Carlton Vic)*. (2020) 25:690–702. doi: 10.1111/resp.13838

2. Gottlieb DJ, Punjabi NM. Diagnosis and management of obstructive sleep apnea: A review. *Jama*. (2020) 323:1389–400. doi: 10.1001/jama.2020.3514

3. Afshin A, Forouzanfar MH, Reitsma MB, Sur P, Estep K, Lee A, et al. Health effects of overweight and obesity in 195 countries over 25 years. *New Engl J Med*. (2017) 377:13–27. doi: 10.1056/NEJMoa1614362

4. Romero-Corral A, Caples SM, Lopez-Jimenez F, Somers VK. Interactions between obesity and obstructive sleep apnea: implications for treatment. *Chest*. (2010) 137:711–9. doi: 10.1378/chest.09-0360

5. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM. Increased prevalence of sleep-disordered breathing in adults. *Am J Epidemiol*. (2013) 177:1006–14. doi: 10.1093/aje/kws342

6. Veasey SC, Rosen IM. Obstructive sleep apnea in adults. *New Engl J Med*. (2019) 380:1442–9. doi: 10.1056/NEJMcp1816152

7. Zhang X, Wang S, Xu H, Yi H, Guan J, Yin S. Metabolomics and microbiome profiling as biomarkers in obstructive sleep apnoea: a comprehensive review. *Eur Respir review: an Off J Eur Respir Soc*. (2021) 30. doi: 10.1183/16000617.0220-2020

8. Lv R, Liu X, Zhang Y, Dong N, Wang X, He Y, et al. Pathophysiological mechanisms and therapeutic approaches in obstructive sleep apnea syndrome. *Signal transduction targeted Ther*. (2023) 8:218. doi: 10.1038/s41392-023-01496-3

9. Hunyor I, Cook KM. Models of intermittent hypoxia and obstructive sleep apnea: molecular pathways and their contribution to cancer. *Am J Physiol Regulatory Integr Comp Physiol*. (2018) 315:R669–r87. doi: 10.1152/ajpregu.00036.2018

10. Prabhakar NR, Peng YJ, Nanduri J. Hypoxia-inducible factors and obstructive sleep apnea. *J Clin Invest*. (2020) 130:5042–51. doi: 10.1172/jci137560

11. Ludwig K, Huppertz T, Radsak M, Gouveris H. Cellular immune dysfunction in obstructive sleep apnea. *Front surgery*. (2022) 9:890377. doi: 10.3389/fsurg.2022.890377

12. Polasky C, Steffen A, Loyal K, Lange C, Bruchhage KL, Pries R. Redistribution of monocyte subsets in obstructive sleep apnea syndrome patients leads to an imbalanced PD-1/PD-L1 cross-talk with CD4/CD8 T cells. *J Immunol (Baltimore Md: 1950)*. (2021) 206:51–8. doi: 10.4049/jimmunol.2001047

13. Domagała-Kulawik J, Osińska I, Piechuta A, Bielicki P, Skirecki TT. B, and NKT cells in systemic inflammation in obstructive sleep apnoea. *Mediators inflammation*. (2015) 2015:161579. doi: 10.1155/2015/161579

14. Kita S, Maeda N, Shimomura I. Interorgan communication by exosomes, adipose tissue, and adiponectin in metabolic syndrome. *J Clin Invest*. (2019) 129:4041–9. doi: 10.1172/jci129193

15. Pegtel DM, Gould SJ. Exosomes. *Annu Rev Biochem*. (2019) 88:487–514. doi: 10.1146/annurev-biochem-013118-111902

16. Krylova SV, Feng D. The machinery of exosomes: biogenesis, release, and uptake. *Int J Mol Sci*. (2023) 24. doi: 10.3390/ijms24021337

17. Pan Y, Hui X, Hoo RLC, Ye D, Chan CYC, Feng T, et al. Adipocyte-secreted exosomal microRNA-34a inhibits M2 macrophage polarization to promote obesity-induced adipose inflammation. *J Clin Invest*. (2019) 129:834–49. doi: 10.1172/jci123069

18. Flaherty SE 3rd, Grijalva A, Xu X, Ables E, Nomani A, Ferrante AW Jr. A lipase-independent pathway of lipid release and immune modulation by adipocytes. *Sci (New York NY)*. (2019) 363:989–93. doi: 10.1126/science.aaw2586

19. Wang W, Zhu N, Yan T, Shi YN, Chen J, Zhang CJ, et al. The crosstalk: exosomes and lipid metabolism. *Cell communication signaling: CCS*. (2020) 18:119. doi: 10.1186/s12964-020-00581-2

20. Ying W, Riopel M, Bandyopadhyay G, Dong Y, Birmingham A, Seo JB, et al. Adipose tissue macrophage-derived exosomal miRNAs can modulate *in vivo* and *in vitro* insulin sensitivity. *Cell*. (2017) 171:372–84.e12. doi: 10.1016/j.cell.2017.08.035

21. Gharib SA, Hurley AL, Rosen MJ, Spilsbury JC, Schell AE, Mehra R, et al. Obstructive sleep apnea and CPAP therapy alter distinct transcriptional programs in subcutaneous fat tissue. *Sleep*. (2020) 43. doi: 10.1093/sleep/zsz314

22. Gharib SA, Hayes AL, Rosen MJ, Patel SR. A pathway-based analysis on the effects of obstructive sleep apnea in modulating visceral fat transcriptome. *Sleep*. (2013) 36:23–30. doi: 10.5665/sleep.2294

23. Rashid H, Ullah A, Ahmad S, Aljahdali SM, Waheed Y, Shaker B, et al. Identification of novel genes and pathways of ovarian cancer using a comprehensive bioinformatic framework. *Appl Biochem Biotechnol*. (2024) 196:3056–75. doi: 10.1007/s12010-023-04702-8

24. Chen H, Su X, Li Y, Dang C, Luo Z. Identification of metabolic reprogramming-related genes as potential diagnostic biomarkers for diabetic nephropathy based on bioinformatics. *Diabetol Metab syndrome*. (2024) 16:287. doi: 10.1186/s13098-024-01531-5

25. Ye W, Li H, Zhao J, Lu D, Tao T, Zhu X. Graphene therapy-related lncRNAs as prognostic and immune microenvironmental biomarkers in hepatocellular carcinoma. *Trans Oncol*. (2024) 43:101915. doi: 10.1016/j.tranon.2024.101915

26. Wu X, Pan Z, Liu W, Zha S, Song Y, Zhang Q, et al. The discovery, validation, and function of hypoxia-related gene biomarkers for obstructive sleep apnea. *Front Med*. (2022) 9:813459. doi: 10.3389/fmed.2022.813459

27. Wang Y, Huang Z, Xiao Y, Wan W, Yang X. The shared biomarkers and pathways of systemic lupus erythematosus and metabolic syndrome analyzed by bioinformatics combining machine learning algorithm and single-cell sequencing analysis. *Front Immunol*. (2022) 13:1015882. doi: 10.3389/fimmu.2022.1015882

28. Hsu WH, Ko AT, Weng CS, Chang CL, Jan YT, Lin JB, et al. Explainable machine learning model for predicting skeletal muscle loss during surgery and adjuvant chemotherapy in ovarian cancer. *J cachexia sarcopenia muscle*. (2023) 14:2044–53. doi: 10.1002/jcsm.13282

29. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep*. (2017) 18:248–62. doi: 10.1016/j.celrep.2016.12.019

30. Sun Y, Ding S, Shen F, Yang X, Sun W, Wan J. Employ machine learning to identify NAD+ metabolism-related diagnostic markers for ischemic stroke and develop a diagnostic model. *Exp gerontology*. (2024) 196:112584. doi: 10.1016/j.exger.2024.112584

31. Yoo M, Shin J, Kim J, Ryall KA, Lee K, Lee S, et al. DSigDB: drug signatures database for gene set analysis. *Bioinf (Oxford England)*. (2015) 31:3069–71. doi: 10.1093/bioinformatics/btv313

32. Ye P, Qin H, Zhan X, Wang Z, Liu C, Song B, et al. Diagnosis of obstructive sleep apnea in children based on the XGBoost algorithm using nocturnal heart rate and blood oxygen feature. *Am J otolaryngology*. (2023) 44:103714. doi: 10.1016/j.amjoto.2022.103714

33. Graham TH. Prolylcarboxypeptidase (PrCP) inhibitors and the therapeutic uses thereof: a patent review. *Expert Opin Ther patents*. (2017) 27:1077–88. doi: 10.1080/13543776.2017.1349104

34. Shariat-Madar B, Kolte D, Verlangieri A, Shariat-Madar Z. Prolylcarboxypeptidase (PRCP) as a new target for obesity treatment. *Diabetes Metab syndrome obesity: Targets Ther*. (2010) 3:67–78. doi: 10.2147/dmsott.s7290

35. Meyer EJ, Wittert GA. Approach the patient with obstructive sleep apnea and obesity. *J Clin Endocrinol Metab*. (2024) 109:e1267–e79. doi: 10.1210/clinem/dgad572

36. Suong DN, Thao DT, Masamitsu Y, Thuoc TL. Ubiquitin carboxyl hydrolase L1 significance for human diseases. *Protein Pept letters*. (2014) 21:624–30. doi: 10.2174/09298665216614040312595 9

37. Bishop P, Rocca D, Henley JM. Ubiquitin C-terminal hydrolase L1 (UCH-L1): structure, distribution and roles in brain function and dysfunction. *Biochem J*. (2016) 473:2453–62. doi: 10.1042/bcj20160082

38. Geng B, Wang X, Park KH, Lee KE, Kim J, Chen P, et al. UCHL1 protects against ischemic heart injury via activating HIF-1α signal pathway. *Redox Biol*. (2022) 52:102295. doi: 10.1016/j.redox.2022.102295

39. Ryan S, Taylor CT, McNicholas WT. Predictors of elevated nuclear factor-kappaB-dependent genes in obstructive sleep apnea syndrome. *Am J Respir Crit Care Med*. (2006) 174:824–30. doi: 10.1164/rccm.200601-066OC

40. Ryan S, Taylor CT, McNicholas WT. Selective activation of inflammatory pathways by intermittent hypoxia in obstructive sleep apnea syndrome. *Circulation*. (2005) 112:2660–7. doi: 10.1161/circulationaha.105.556746

41. Rey E, Del Pozo-Maroto E, Marañón P, Beeler B, García-García Y, Landete P, et al. Intrahepatic expression of fatty acid translocase CD36 is increased in obstructive sleep apnea. *Front Med*. (2020) 7:450. doi: 10.3389/fmed.2020.00450

42. Hwang SS, Lim J, Yu Z, Kong P, Sefik E, Xu H, et al. mRNA destabilization by BTG1 and BTG2 maintains T cell quiescence. *Sci (New York NY)*. (2020) 367:1255–60. doi: 10.1126/science.aax0194

43. Kim S, Hong JW, Park KW. B cell translocation gene 2 (Btg2) is regulated by Stat3 signaling and inhibits adipocyte differentiation. *Mol Cell Biochem*. (2016) 413:145–53. doi: 10.1007/s11010-015-2648-z

44. Quy LN, Choi YW, Kim YH, Chwae YJ, Park TJ, Lim IK. TIS21(/BTG2/PC3) inhibits interleukin-6 expression via downregulation of STAT3 pathway. *Cell signalling*. (2013) 25:2391–9. doi: 10.1016/j.cellsig.2013.07.024

45. Wang D, Zhou Y, Lei W, Zhang K, Shi J, Hu Y, et al. Signal transducer and activator of transcription 3 (STAT3) regulates adipocyte differentiation via peroxisome-proliferator-activated receptor gamma (PPARgamma). *Biol Cell*. (2009) 102:1–12. doi: 10.1042/bc20090070

46. Kehoe K, Noels H, Theelen W, De Hert E, Xu S, Verrijken A, et al. Prolyl carboxypeptidase activity in the circulation and its correlation with body weight and adipose tissue in lean and obese subjects. *PloS One*. (2018) 13:e0197603. doi: 10.1371/journal.pone.0197603

47. Jackman HL, Tan F, Schraufnagel D, Dragović T, Dezsö B, Becker RP, et al. Plasma membrane-bound and lysosomal peptidases in human alveolar macrophages. *Am J Respir Cell Mol Biol*. (1995) 13:196–204. doi: 10.1165/ajrcmb.13.2.7626287

48. Kumamoto K, Stewart TA, Johnson AR, Erdös EG. Prolylcarboxypeptidase (angiotensinase C) in human lung and cultured cells. *J Clin Invest*. (1981) 67:210–5. doi: 10.1172/jci110015

49. Xu S, Lind L, Zhao L, Lindahl B, Venge P. Plasma prolylcarboxypeptidase (angiotensinase C) is increased in obesity and diabetes mellitus and related to cardiovascular dysfunction. *Clin Chem*. (2012) 58:1110–5. doi: 10.1373/clinchem.2011.179291

50. Huang Y, He S, Chen Y, Sheng J, Fu Y, Du X, et al. UCHL1 promoted polarization of M1 macrophages by regulating the PI3K/AKT signaling pathway. *J Inflammation Res*. (2022) 15:735–46. doi: 10.2147/jir.S343487

51. Zhang Z, Liu N, Chen X, Zhang F, Kong T, Tang X, et al. UCHL1 regulates inflammation via MAPK and NF-κB pathways in LPS-activated macrophages. *Cell Biol Int*. (2021) 45:2107–17. doi: 10.1002/cbin.11662

52. Wu GS, Bassing CH. Flip the switch: BTG2-PRMT1 protein complexes antagonize pre-B-cell proliferation to promote B-cell development. *Cell Mol Immunol*. (2018) 15:808–11. doi: 10.1038/cmi.2017.156

53. Palmiter RD. Reduced levels of neurotransmitter-degrading enzyme PRCP promote a lean phenotype. *J Clin Invest*. (2009) 119:2130–3. doi: 10.1172/jci40001

54. Gan M, Shen L, Wang S, Guo Z, Zheng T, Tan Y, et al. Genistein inhibits high fat diet-induced obesity through miR-222 by targeting BTG2 and adipor1. *Food Funct*. (2020) 11:2418–26. doi: 10.1039/c9fo00861f

Frontiers in Immunology

Explores novel approaches and diagnoses to treat immune disorders.

The official journal of the International Union of Immunological Societies (IUIS) and the most cited in its field, leading the way for research across basic, translational and clinical immunology.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact