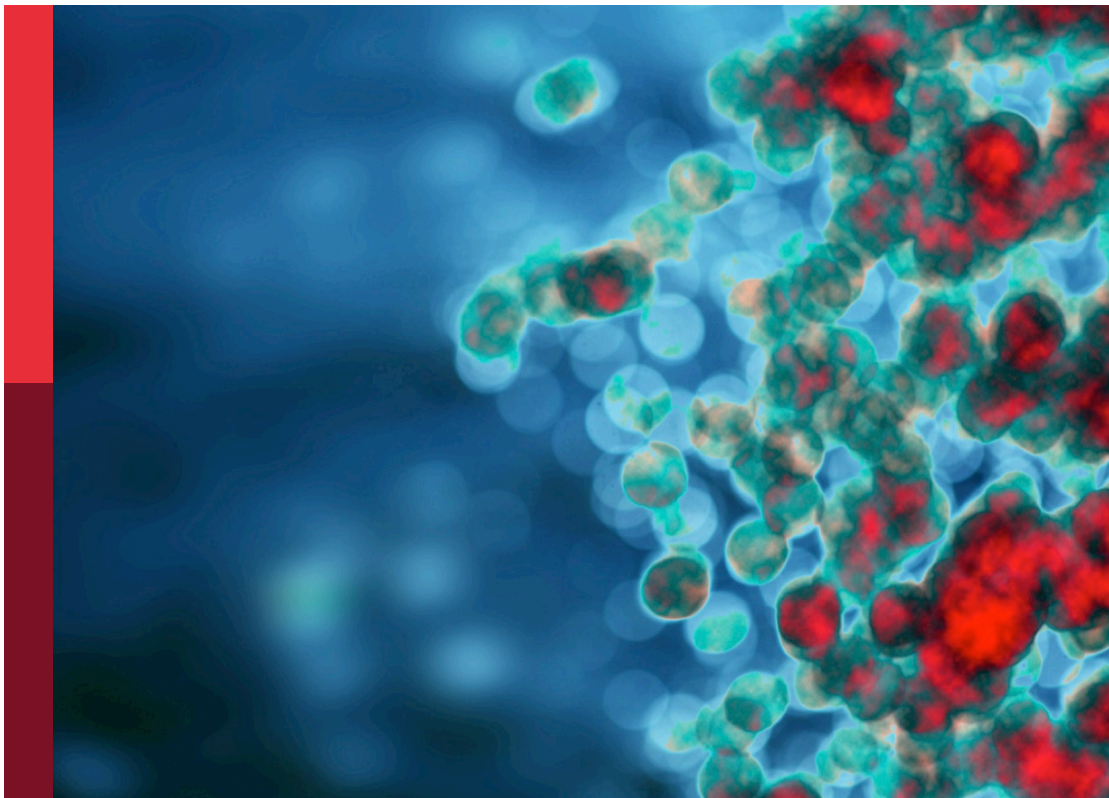
frontiers | Research Topics

Immunological precision therapeutics: integrating multi-omics technologies and comprehensive approaches for personalized immune intervention

Edited by
Wenyi Jin, Raquel Alarcon Rodriguez and Mar Requena Mullor

Published in
Frontiers in Immunology
Frontiers in Oncology



ISSN 1664-8714
ISBN 978-2-8325-6143-0
DOI 10.3389/978-2-8325-6143-0

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

Immunological precision therapeutics: integrating multi-omics technologies and comprehensive approaches for personalized immune intervention

Topic editors

Wenyi Jin — City University of Hong Kong, Hong Kong, SAR China

Raquel Alarcon Rodriguez — University of Almeria, Spain

Mar Requena Mullor — University of Almeria, Spain

Citation

Jin, W., Rodriguez, R. A., Mullor, M. R., eds. (2025). *Immunological precision therapeutics: integrating multi-omics technologies and comprehensive approaches for personalized immune intervention*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6143-0

Table of contents

frontiers | Frontiers in Immunology

TYPE Editorial
PUBLISHED 06 March 2025
DOI 10.3389/fimmu.2025.1581238

OPEN ACCESS

EDITED AND REVIEWED BY
Peter Brossart,
University of Bonn, Germany

*CORRESPONDENCE
Wenyi Jin
gin1994@whu.edu.cn

RECEIVED 21 February 2025
ACCEPTED 25 February 2025
PUBLISHED 06 March 2025

CITATION
Xue M and Jin W (2025) Editorial: Immunological precision therapeutics: integrating multi-omics technologies and comprehensive approaches for personalized immune intervention. *Front. Immunol.* 16:1581238. doi: 10.3389/fimmu.2025.1581238

Editorial: Immunological precision therapeutics: integrating multi-omics technologies and comprehensive approaches for personalized immune intervention

Mingyang Xue[1] and Wenyi Jin[2,3]*

[1]School of Medicine, Kunming University of Science and Technology, Kunming, China, [2]Department of Biomedical Sciences, City University of Hong Kong, Hong Kong, Hong Kong SAR, China, [3]Department of Orthopaedics, Renmin Hospital of Wuhan University, Wuhan University, Wuhan, China

KEYWORDS

immunotherapy, precision medicine, multi-omics analysis, immune signatures, personalized therapy

Editorial on the Research Topic
Immunological precision therapeutics: integrating multi-omics technologies and comprehensive approaches for personalized immune intervention

The integration of multi-omics technologies in immunology heralds a transformative era in the development of personalized therapeutic strategies. As immunotherapy continues to emerge as a cornerstone of precision medicine, this Research Topic brings together a rich collection of studies that explore how comprehensive multi-omics approaches—ranging from genomics and transcriptomics to proteomics, metabolomics, and single-cell profiling—can be harnessed to optimize immune interventions tailored to individual patients.

One of the key themes across the contributions is the application of single-cell sequencing to unravel the complexities of the immune landscape (Yu et al.). For example, the study on gliomas highlights how single-cell RNA sequencing (scRNA-seq) data can be leveraged to identify potential biomarkers like DAPK1, which may serve as a prognostic marker for glioma progression and therapeutic efficacy. Such insights are crucial, as gliomas remain one of the most difficult malignancies to treat due to their aggressive nature and the barriers posed by the blood-brain barrier (BBB). By utilizing multi-omics data, including scRNA-seq, the authors have demonstrated how specific immune subpopulations, such as DAPK1-expressing cells, could be used to predict patient outcomes, offering a new approach for precision therapy in glioma.

Similarly, the study on cancer stemness in ovarian cancer illustrates how multi-omics data can reveal the role of cancer stem cells (CSCs) in mediating resistance to immune checkpoint inhibitors (ICIs) (Liu et al.). Through the integration of RNA sequencing and CRISPR-based screens, the authors have identified critical genes involved in regulating

cancer stemness, such as SNRPE, which negatively affects ICI response. This approach exemplifies the power of combining machine learning and genomics to predict immune responses and guide treatment decisions in cancer immunotherapy (Sun et al.).

A consistent thread running through many of the contributions is the focus on immune microenvironment characterization. (Li et al.). The work on cuproptosis-related genes in ovarian cancer investigates how the deregulation of copper-dependent cell death pathways impacts the immune landscape, with the development of a robust risk score model for predicting prognosis and immunotherapy response (Xiaorong et al.). This study adds to the growing body of literature emphasizing the importance of not only tumor-intrinsic factors but also the immune environment in determining the success of immunotherapy. Similarly, the research on thyroid carcinoma highlights the role of the epithelial-mesenchymal transition (EMT) and immune cell infiltration in cancer progression, demonstrating how these pathways can be used to refine prognostic models (Wu et al.).

The studies on lung cancer immunotherapy (Yan et al.) and the gastric cancer microbiome (Qian et al.) offer further proof of the power of multi-omics in uncovering complex disease mechanisms. In the case of lung cancer, combining genomic, transcriptomic, and proteomic data provides insights into immune-related pathways, paving the way for more personalized and effective treatment options (Li et al.). The study on the gastric microbiome goes a step further by integrating microbiota data with immune-activated transcripts, suggesting that specific bacterial species may influence immune response and tumor progression, thus offering potential targets for therapeutic modulation (Qian et al.).

Across all studies, the integration of multi-omics data with machine learning algorithms is repeatedly showcased as a tool for predicting therapeutic efficacy and patient outcomes (Sun et al.). Whether through the development of prognostic risk scores or by uncovering previously unrecognized molecular interactions, machine learning serves as a bridge between vast amounts of complex biological data and actionable insights for personalized medicine.

One of the most compelling aspects of this Research Topic is its emphasis on synergistic approaches (Amiri et al). The review on CAR-T cell therapy and CRISPR/Cas9 exemplifies how combining cutting-edge gene-editing technologies with immunotherapy can enhance treatment specificity and efficacy (Amiri et al). As CAR-T cells are increasingly used in the treatment of hematologic malignancies, the integration of CRISPR/Cas9 holds the promise of overcoming some of the limitations, such as the tumor's ability to evade immune detection. By boosting CAR-T cell persistence and engineering them to overcome immune suppression, CRISPR-edited CAR-T therapies could expand the applicability of immunotherapy to solid tumors, offering new hope for patients with refractory cancers.

Together, these articles underscore the potential of personalized immune interventions (Wen et al.) and demonstrate the vast promise of multi-omics technologies in immunology (Sennikov et al.). They not only contribute to a deeper understanding of immune responses but also provide critical insights that will shape the future of immune-related disease treatments (Zhang et al.). As this Research Topic shows, the ability to integrate diverse types of data, from single-cell sequencing to machine learning models, allows us to build a more comprehensive and nuanced view of the immune system, its dysregulation in disease, and how best to tailor therapies to individual patients.

In conclusion, this Research Topic contributes to the growing momentum towards precision immunotherapy (Chen et al.). It reinforces the need for interdisciplinary approaches that combine the power of genomics, transcriptomics, proteomics, metabolomics, and computational techniques to drive the development of highly personalized therapies. As we move forward, these strategies will not only enhance the efficacy of existing therapies but also open the door to entirely new modalities of immune modulation that could transform the treatment of cancers, autoimmune diseases, and beyond.

Author contributions

MX: Writing – original draft, Writing – review & editing. WJ: Writing – original draft, Writing – review & editing.

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

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 14 June 2024
DOI 10.3389/fimmu.2024.1426474

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Biao Zhang,
Dalian Medical University, China
Jianfeng Chen,
Wuxi Traditional Chinese Medicine Hospital, China

*CORRESPONDENCE
MingHua Ren
002051@hrbmu.edu.cn
BeiChen Ding
dingbeichen@hrbmu.edu.cn
HaiXia Zhu
00zlingling@163.com

†These authors have contributed equally to this work

RECEIVED 01 May 2024
ACCEPTED 31 May 2024
PUBLISHED 14 June 2024

CITATION
Wang Y, Li C, He J, Zhao Q, Zhou Y, Sun H, Zhu H, Ding B and Ren M (2024) Multi-omics analysis and experimental validation of the value of monocyte-associated features in prostate cancer prognosis and immunotherapy. *Front. Immunol.* 15:1426474. doi: 10.3389/fimmu.2024.1426474

Multi-omics analysis and experimental validation of the value of monocyte-associated features in prostate cancer prognosis and immunotherapy

YaXuan Wang[1†], Chao Li[2†], JiaXing He[1†], QingYun Zhao[1], Yu Zhou[1], HaoDong Sun[1], HaiXia Zhu[3*], BeiChen Ding[1*] and MingHua Ren[1*]

[1]Department of Urology, The First Affiliated Hospital of Harbin Medical University, Harbin, China, [2]Department of General, Visceral, and Transplant Surgery, Ludwig-Maximilians-University Munich, Munich, Germany, [3]Clinical Laboratory, Tumor Hospital Affiliated to Nantong University, Nantong, China

Background: Monocytes play a critical role in tumor initiation and progression, with their impact on prostate adenocarcinoma (PRAD) not yet fully understood. This study aimed to identify key monocyte-related genes and elucidate their mechanisms in PRAD.

Method: Utilizing the TCGA-PRAD dataset, immune cell infiltration levels were assessed using CIBERSORT, and their correlation with patient prognosis was analyzed. The WGCNA method pinpointed 14 crucial monocyte-related genes. A diagnostic model focused on monocytes was developed using a combination of machine learning algorithms, while a prognostic model was created using the LASSO algorithm, both of which were validated. Random forest and gradient boosting machine singled out CCNA2 as the most significant gene related to prognosis in monocytes, with its function further investigated through gene enrichment analysis. Mendelian randomization analysis of the association of HLA-DR high-expressing monocytes with PRAD. Molecular docking was employed to assess the binding affinity of CCNA2 with targeted drugs for PRAD, and experimental validation confirmed the expression and prognostic value of CCNA2 in PRAD.

Result: Based on the identification of 14 monocyte-related genes by WGCNA, we developed a diagnostic model for PRAD using a combination of multiple machine learning algorithms. Additionally, we constructed a prognostic model using the LASSO algorithm, both of which demonstrated excellent predictive capabilities. Analysis with random forest and gradient boosting machine algorithms further supported the potential prognostic value of CCNA2 in PRAD. Gene enrichment analysis revealed the association of CCNA2 with the regulation of cell cycle and cellular senescence in PRAD. Mendelian randomization analysis confirmed that monocytes expressing high levels of HLA-DR may promote PRAD. Molecular docking results suggested a strong affinity of CCNA2 for drugs targeting PRAD. Furthermore, immunohistochemistry experiments validated the upregulation of CCNA2 expression in PRAD and its correlation with patient prognosis.

Conclusion: Our findings offer new insights into monocyte heterogeneity and its role in PRAD. Furthermore, CCNA2 holds potential as a novel targeted drug for PRAD.

KEYWORDS

prognosis, monocyte, machine learning, PRAD, multi-omics analysis

1 Introduction

Based on 2024 U.S. cancer statistics, prostate adenocarcinoma (PRAD) now surpasses lung cancer as the most common cancer among men (1). In China, recent data from the China National Cancer Center in 2022 revealed that PRAD incidence rates have exceeded those of kidney and bladder tumors based on 2016 data from 487 tumor registries nationwide (2). The incidence of PRAD has been on the rise in recent years due to economic and social development and increased life expectancy. Options for treating PRAD currently consist of radical radiotherapy, radical prostatectomy, chemotherapy, and androgen deprivation therapy, customized based on the progression of the individual patient's illness (3). Despite advancements in PRAD treatment, the 5-year survival rate for patients remains relatively low (4). Therefore, it is crucial to identify potential prognostic markers and assess therapeutic targets to improve the prognosis of PRAD patients.

Monocytes are vital components of the innate immune system and are indispensable for defending against foreign invaders (5). There are three primary subpopulations of monocytes: classical, nonclassical, and intermediate monocytes (6). Monocytes first mature into classical monocytes in the bone marrow, followed by differentiation into nonclassical monocytes in the bloodstream, with an intermediate monocyte phase in between. Numerous studies have shown that monocytes play a direct role in immune responses by initiating cell death and phagocytosis (7). Additionally, monocytes can engage with T cells and natural killer cells, impacting tumor progression by producing chemokines (8). Moreover, monocytes have the capability to transform into various immune cells such as tumor-associated macrophages and dendritic cells, critical components of the immune system that actively promote tumor growth and spread (9). Tumor-infiltrating immune cells play a crucial role in the pathogenesis of PRAD. Recent research indicates that prognostic markers linked to M2 macrophages can forecast biochemical recurrence in patients with PRAD (10). Furthermore, elevated levels of macrophages in prostate biopsies have been correlated with disease progression following hormone therapy (11). Moreover, there is evidence to suggest that circulating monocyte levels could serve as a biomarker for metastatic PRAD, indicating a notably unfavorable prognosis (12). Integrated multi-omics, machine learning, and artificial intelligence are being more frequently utilized in the field of medicine (13–17). It is essential to conduct further analysis on levels of tumor-infiltrating immune cells and identify genes related to immune cell infiltration using multi-omics and machine learning techniques to enhance the accuracy of diagnosis and treatment for PRAD. The objective of our study is to enhance researchers' comprehension of the mechanisms underlying tumor immune infiltration, progress in immunotherapy for PRAD patients, and offer novel insights for clinical immunotherapy.

The significance of immune cell infiltration in tumors and the exploration of its potential regulatory genes have been acknowledged based on existing research. The CIBERSORT algorithm provides a convenient method for evaluating immune cell infiltration levels in PRAD. By utilizing this algorithm, we calculated the infiltration levels of immune cells in TCGA-PRAD samples and grouped the samples accordingly. Our analysis revealed that only the infiltration level of monocytes significantly correlated with the prognosis of PRAD patients. Using the weighted correlation network analysis (WGCNA) method, we have discovered prognostic differential genes associated with monocytes in the PRAD dataset samples from the cancer genome atlas (TCGA) database. These genes exhibit correlations with patient stage, Gleason score, and PSA score. Subsequently, we developed diagnostic and prognostic models using various machine learning techniques, yielding positive results. By analyzing the TCGA-PRAD and GSE16560 datasets, CCNA2 emerged as the most promising prognostic gene related to monocytes in PRAD. Furthermore, we delved into the function of CCNA2 and its potential interactions with therapeutic drugs for PRAD. In conclusion, our study lays the groundwork for understanding the impact of monocytes on the prognosis of PRAD patients and identifies a novel drug target for PRAD treatment.

2 Materials and methods

2.1 Data acquisition

The TCGA database provided data on 52 normal prostate samples and 498 PRAD samples. The monocyte-related gene

diagnostic model was validated using the GSE62872 and GSE32571 datasets, while the GSE16560 dataset was used for the prognostic model of monocyte-related genes. In addition, 60 cases of prostate cancer tissue and paired para-cancerous tissue were obtained from Shanghai Outdo Biotech Company. The patients included in the tissue chip study underwent surgery between January 2011 and December 2014, with a follow-up period extending from November 2021, covering a span of 6 to 10 years.

2.2 Constructing diagnostic and prognostic models

Use multiple machine learning algorithms to combine into more than one hundred algorithm combinations to develop the best PRAD diagnostic model (18). The training set comprised the TCGA-PRAD data set, with validation sets GSE62872 and GSE32571. Area under curve (AUC) values were calculated for each algorithm combination, and the combination with the highest average AUC was selected as the best. The prognostic model was based on the least absolute shrinkage and selection operator (LASSO) regression algorithm and evaluated using 10-fold cross-validation in R software with the glmnet package (19, 20).

2.3 Functional analysis of candidate genes

The gene set cancer analysis (GSCA) and CancerSEA databases were used to analyze the functions of monocyte-related genes (21, 22). To better understand the oncogenic role of target genes, the ClusterProfiler package in R was used to analyze the potential functions of CCNA2 and enrich the Kyoto encyclopedia of genes and genomes (KEGG) pathway. The R packages "clusterProfiler" was utilized for the GSEA enrichment analysis of genes (23).

2.4 Analysis of the correlation between CCNA2 and immune cell infiltration

The GSCA database was utilized to examine the relationship between CCNA2 and monocytes. Furthermore, we investigated the correlation between CCNA2 and markers of monocytes using the TCGA-PRAD dataset. Additionally, the TISCH2 database was employed to analyze the association between CCNA2 and immune cell infiltration (24).

2.5 Immunohistochemical staining analysis of CCNA2 expression in PRAD tissues

The prostate cancer tissue chip was initially placed in an 85° C oven for 20 minutes, followed by soaking in xylene solution for 20 minutes for dewaxing. Subsequently, the tissue chips underwent a series of hydration steps involving immersion in 100%, 95%, 80%, and 70% ethanol for 2 minutes each. The tissue chip was then treated with citric acid solution and subjected to boiling in a pressure cooker for antigen retrieval, followed by cooling in ice water to reach room temperature. The chip was then rinsed with PBS, circled with a histochemistry pen, sealed with hydrogen peroxide solution, and cleaned with PBS. Next, CCNA2 antibody (BOSTER, PB9424) was applied dropwise to cover the tissue chip, which was left at room temperature for 2 hours. Post-reaction, the chip was rinsed with PBS and the immunohistochemistry secondary antibody was added dropwise, left for 20 minutes, and then cleaned with PBS. Finally, the tissue chip underwent DAB color development, dehydration in a series of ethanol solutions, sealing, and microscopic examination to conclude the experiment. The immunostaining intensity score ranges from 0 to 3, where 0, 1, 2, and 3 represent no reaction, weak reaction, moderate reaction, and strong reaction, respectively. Following this, a scale based on the proportion of positive staining is applied, with scores of 1, 2, 3, and 4 corresponding to 0%-25%, 26%-50%, 51%-75%, and 76%-100%, respectively. The final expression score is determined by multiplying the staining intensity score and the staining proportion score. This calculation results in a score ranging from 0 to 5, indicating low expression, and a score from 6 to 12, indicating high expression.

2.6 Mendelian randomization analysis

The Mendelian randomization analysis in this study investigated the impact of monocytes on prostate cancer patients using the MRBASE website (25). The exposure factor selected was HLA DR++ monocyte %monocyte (ebi-a-GCST90001475) from the MR Base GWAS catalog, with prostate cancer (EBI-A-GCST006085) as the outcome. The analysis criteria included a minimum LD Rsq value of 0.8, a MAF threshold of 0.01, and the exclusion of palindromic SNPs. Various methods such as MR Egger, Weighted median, Weighted mode, Simple mode, and Inverse variance weighted were employed for the analysis.

2.7 Statistical analyses

The level of immune cell infiltration and prognosis of TCGA-PRAD patients, along with the prognostic analysis of CCNA2 in prostate cancer tissue chips used in our experiments, were statistically analyzed using the Log-rank test. The prognostic analysis of ACSM3 and CCNA2 in TCGA-PRAD and GSE16560 datasets was conducted through COX regression. All correlation analyses in this study were performed using the Spearman method. Furthermore, the expression of monocyte-related genes at different stages, Gleason scores, and PSA scores in the TCGA-PRAD dataset, as well as the expression differences of CCNA2 in prostate cancer tissue chips, were analyzed using the Wilcoxon rank sum test.

3 Results

3.1 Analysis of the correlation between the level of immune cell infiltration and the prognosis of PRAD patients

The tumor immune microenvironment consists of tumor cells, immune cells, signaling molecules, extracellular matrix, and unique physical and chemical characteristics (26). This microenvironment significantly impacts tumor diagnosis, survival rates, and treatment responses. Immune cell infiltration in tumors is crucial as it can either help eliminate tumor cells or be manipulated by tumors to promote growth and metastasis (27, 28). The role of immune cells in cancer treatment and prevention, as well as their regulatory mechanisms, has garnered significant attention. An accurate understanding of the distribution and function of immune cells in tumor tissues is essential for effective treatment and prognosis assessment (29, 30). The CIBERSORT algorithm was used to calculate the proportion of 22 immune cells for each sample in the TCGA-PRAD dataset (31, 32). While the infiltration level of 8 types of immune cells in most samples was 0, our study focused on analyzing the relationship between the infiltration levels of the remaining 14 types of immune cells and the prognosis of PRAD patients. Our findings suggest that the infiltration level of monocytes is a significant factor in determining the prognosis of patients with PRAD. Specifically, a higher infiltration level of monocytes is associated with a poorer prognosis for PRAD patients (Figures 1A–N). Furthermore, based on monocyte infiltration levels, PRAD patient samples were classified into high and low monocyte groups. We then examined the percentage abundance of tumor-infiltrating immune cells in each sample (Figure 1O).

3.2 Screening of monocyte-associated differential genes based on the WGCNA method

WGCNA is an algorithm utilized for extracting module information from high-throughput expression data. Our objective was to identify genes highly correlated with monocytes in the TCGA-PRAD dataset using this algorithm. To achieve a scale-free network distribution, we carefully selected the value of the adjacency matrix weight parameter power. In our analysis, we determined the power value to be 20 (Figures 2A–D).

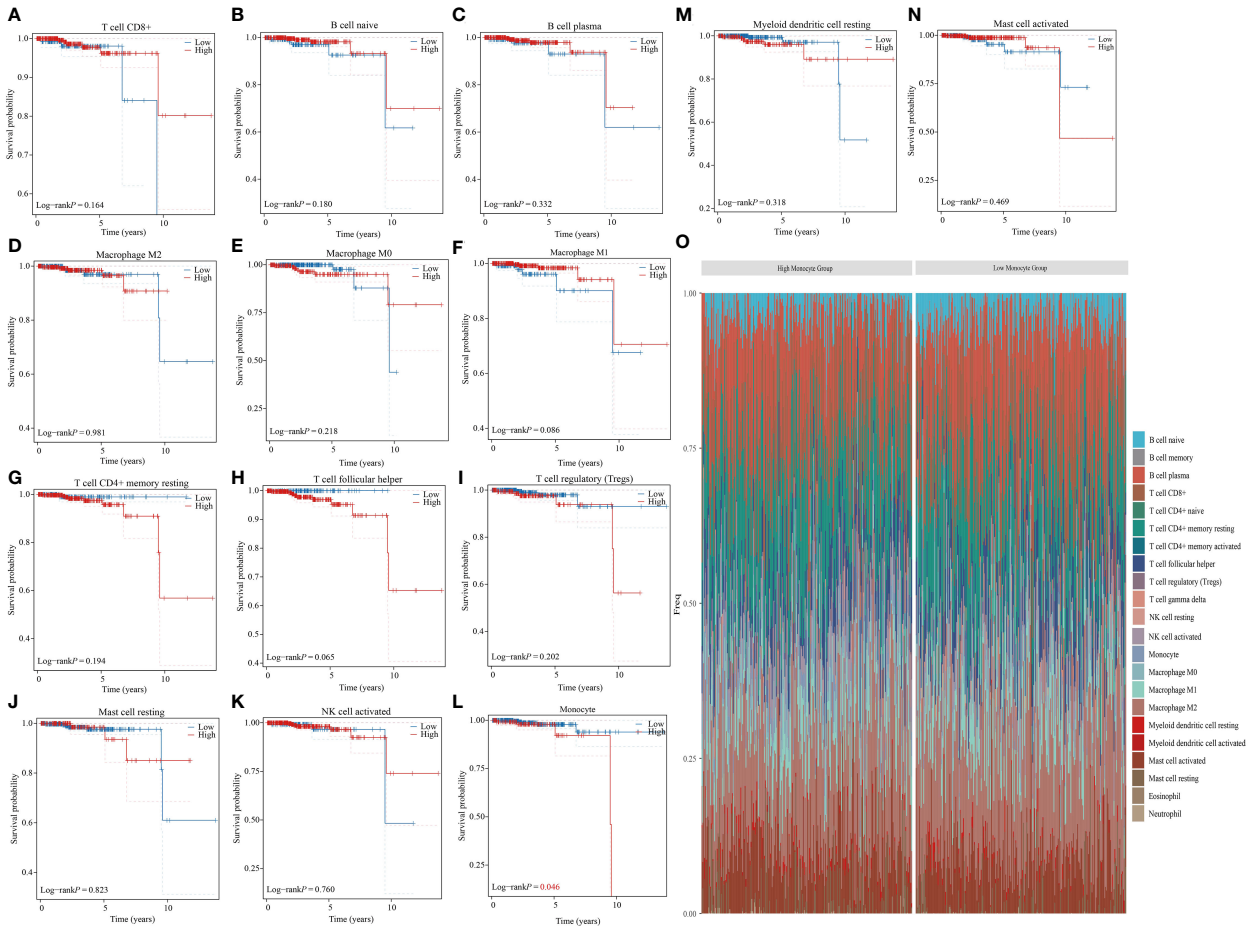


FIGURE 1
Monocyte infiltration correlates with prognosis in PRAD patients. **(A–N)** Analyzing the correlation between different immune cell infiltrations and prognosis in PRAD patients. **(O)** Percentage frequency of different tumor infiltrating immune cells in PRAD samples.

10.3389/fimmu.2024.1426474

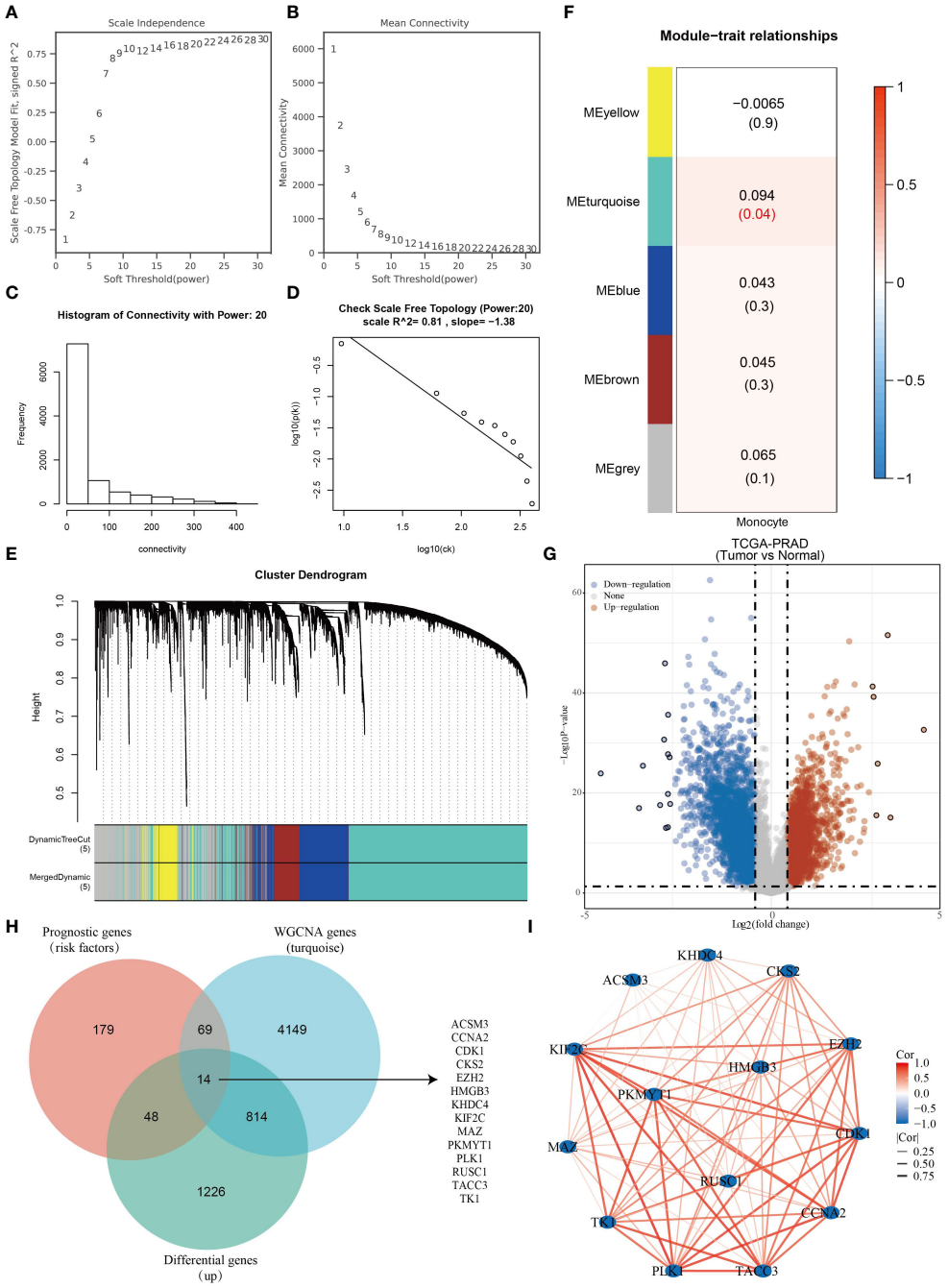


FIGURE 2
Fourteen monocyte-associated differential prognostic genes were identified. **(A–D)** WGCNA Network Construction Parameters. **(E)** Weighted co-expression network modeling based on selected power values. **(F)** Heatmap of trait module associations. **(G)** Analysis of differences in TCGA-PRAD dataset. **(H)** Venn diagram based on the intersection of TCGA-PRAD differential genes, prognostic genes and monocyte-associated genes. **(I)** Correlation network diagram of monocyte-associated prognostic differential genes.

Subsequently, a weighted co-expression network model was constructed based on this power value, leading to the division of the gene set into 5 modules. Notably, the gray module represents genes that do not align with any specific module and lack reference significance (Figure 2E). Using the Pearson correlation algorithm, we found that the turquoise module has the strongest correlation with monocytes (Figure 2F). We conducted differential analysis on TCGA-PRAD samples with a significance level of $P < 0.05$ and Log2 (Fold Change) >1.3 or Log2 (Fold Change) < -1.3 as the selection criteria. Subsequently, we generated a volcano plot to visualize the analysis outcomes (Figure 2G). Our findings indicated a link between high levels of mononuclear cell infiltration and poor prognosis in PRAD patients. Through the intersection of genes in the turquoise module with prognostic risk factors in the TCGA-PRAD dataset and genes highly expressed in PRAD, we identified a total of 14 monocyte-related prognostic differential genes (Figure 2H). Importantly, these 14 genes were found to be positively correlated with each other (Figure 2I).

10.3389/fimmu.2024.1426474

3.3 Functional analysis of monocyte-related prognostic differential genes

The study initially examined the relationship between 14 genes and clinicopathological characteristics of PRAD patients, illustrating this correlation through a heatmap (Figure 3A). Additionally, expression heatmaps were generated for the 14 genes in TCGA-PRAD samples and normal prostate tissue (Figure 3B). Friends analysis aimed to develop a gene interaction network, leveraging network topology to assess gene importance and identify key genes. Notably, TACC3 emerged as the central gene within this network (Figure 3C). A co-expression network diagram was constructed with TACC3 at its core, revealing that ACSM3 exhibited no correlation with TACC3, while the remaining 13 genes showed significant correlations with TACC3 (Figure 3D). The expression levels of CCNA2, CDK1, CKS2, EZH2, HMGB3, KHDC4, KIF2C, PKMYT1, and PLK1 were found to vary significantly across different T stages, N stages, Gleason scores, and PSA scores in TCGA-PRAD samples (Figures 3E–H). Utilizing the CancerSEA database, which is tailored to decode the diverse functional states of cancer cells at a single-cell level, we investigated the functions of these 14 genes in PRAD. Our analysis revealed that these genes play roles in DNA repair, cell cycle regulation, proliferation, inflammation, and stemness (Figure 3I). Furthermore, through enrichment analysis, we discovered that these genes are primarily associated with cell cycle processes (Figure 3J).

3.4 Multiple machine learning combinations to build PRAD diagnostic models

In order to develop a PRAD-related diagnosis model, we utilized three PRAD datasets: the TCGA-PRAD dataset for training, and the

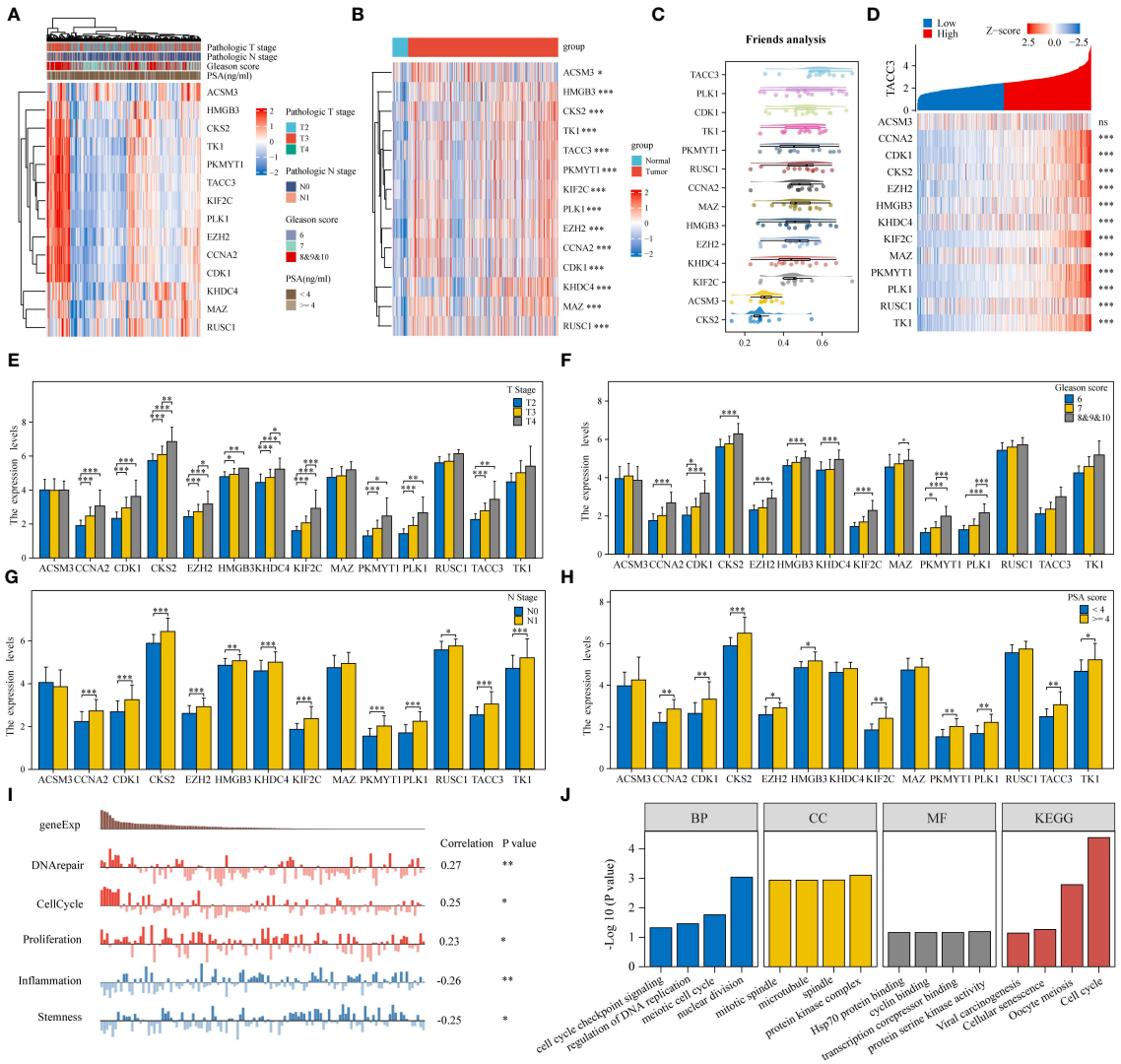


FIGURE 3
Monocyte-associated prognostic differential genes play an important role in PRAD. **(A)** Heatmap of monocyte-associated prognostic differential gene expression in different pathologic parameters. **(B)** Heatmap of monocyte-associated prognostic differential gene expression in PRAD and normal tissues. **(C)** Friends analysis explores key genes in monocyte-associated genes. **(D)** Heatmap of co-expression in monocyte-associated genes. **(E–H)** Histogram of monocyte-related gene expression in different clinicopathologic parameters. **(I, J)** Functional analysis of monocyte-related genes. *p< 0.05, **p< 0.01, ***p< 0.001.

GSE62872 and GSE32571 datasets for validation. Out of 94 algorithm combinations, the Enet[alpha=0.4] algorithm was identified as the most effective for constructing the diagnostic model (Figure 4A). The AUC value for the TCGA-PRAD training set was 0.9, while the AUC values for the validation sets GSE62872 and GSE32571 were 0.674 and 0.945. The diagnostic model built by the Enet[alpha=0.4] algorithm featured six genes: ACSM3, EZH2, HMGB3, KHDC4, MAZ, and TK1 (Figure 4B). Additionally, ROC curves for these six

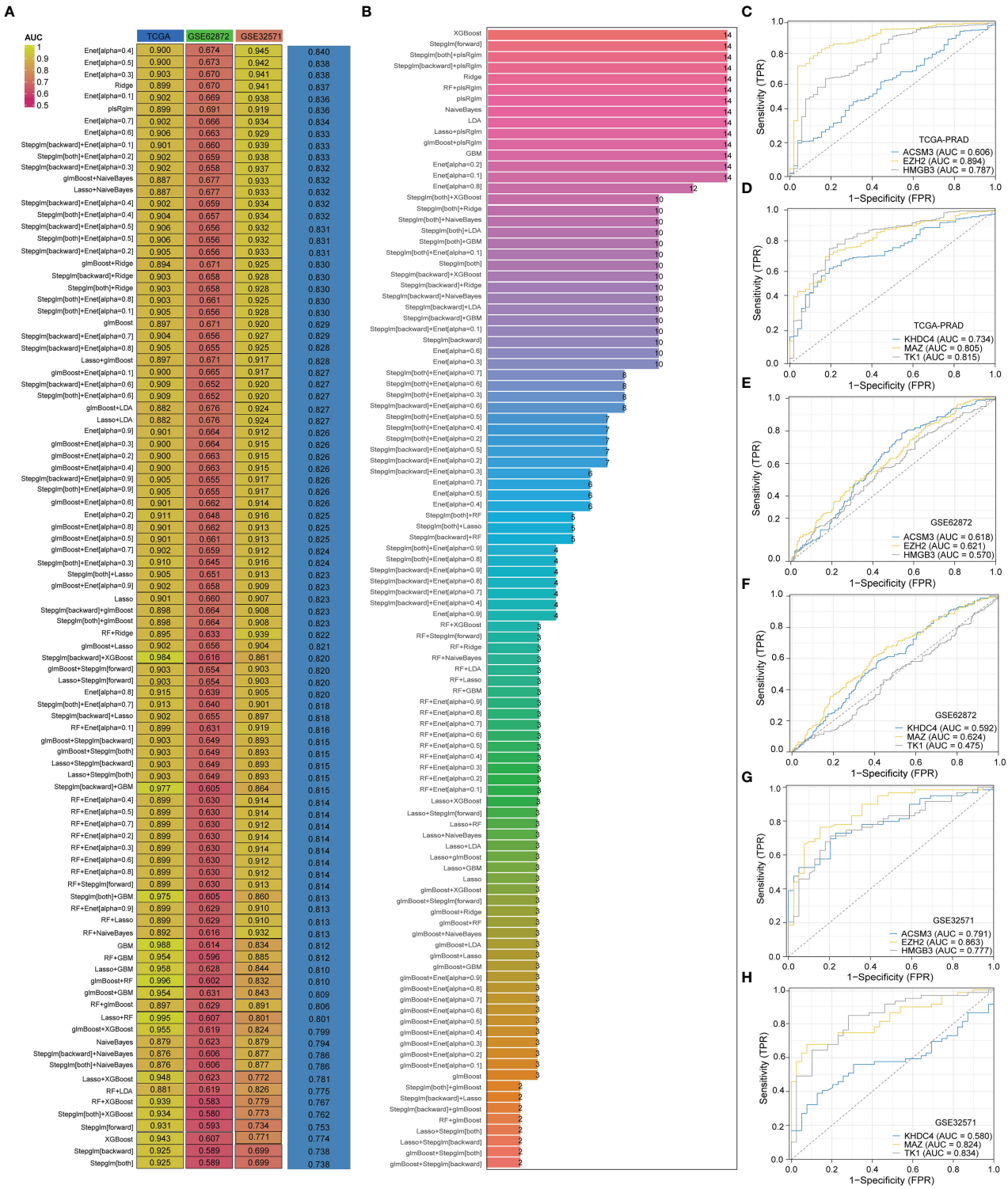


FIGURE 4
The model constructed by Enet [alpha=0.4] algorithm is the best PRAD diagnostic model. **(A)** AUC values of diagnostic models constructed with different combinations of algorithms. **(B)** Number of genes incorporated in diagnostic models constructed with different combinations of algorithms. **(C–H)** Diagnostic value of genes in diagnostic models in different datasets.

genes in the TCGA-PRAD, GSE62872, and GSE32571 datasets were presented (Figures 4C–H). The AUC values of ACSM3, EZH2, HMGB3, KHDC4, MAZ, and TK1 in the TCGA-PRAD dataset are 0.606, 0.894, 0.787, 0.734, 0.805, and 0.815, respectively. Similarly, in the GSE62872 dataset, these values are 0.618, 0.621, 0.570, 0.624, 0.624, and 0.475. While the diagnostic potential of these genes for PRAD in the initial dataset is significant, it lacks precision. To address this, we conducted further analysis using the GSE32571 dataset as a validation set, where the AUC values for the six genes were 0.791, 0.863, 0.777, 0.580, 0.824, and 0.834.

3.5 Constructing prognostic model

To enhance the prediction accuracy of PRAD patient prognosis, we developed a prognostic model utilizing monocyte-related genes through the LASSO algorithm. This model incorporated 9 genes, with corresponding risk scores calculated as follows: ACSM3*(-0.17156) +CCNA2*(0.11148) +CDK1*(0.05883) +CKS2*(0.09723) +EZH2*(0.20259) +KHDC4*(0.10915) +PLK1*(0.02555) +TACC3*(0.14386) +TK1*(0.08982) (Figures 5A, B). Initial validation in the GSE16560 dataset indicated a notably poorer prognosis for patients classified in the high-risk group compared to those in the low-risk group. Additionally, our prognostic model exhibited predictive abilities for 1-year, 5-year, and 7-year prognoses of PRAD patients, with corresponding AUC values of 0.667, 0.650, and 0.668, respectively (Figures 5C–E). Subsequent validation in the TCGA-PRAD dataset confirmed the accuracy of our prognostic model in predicting patient outcomes, particularly for 1-year and 7-year prognoses. However, the predictive ability for the 5-year prognosis of PRAD patients was found to be moderate (Figures 5F–H).

3.6 Multiple machine learning approaches to identify monocyte-associated prognostic genes

The key genes incorporated into the prognostic model were further analyzed. These genes were primarily associated with the cell cycle and activation of the hormone AR (Figure 6A). Validation from the CancerSEA database confirmed that these prognostic genes were linked to DNA repair, cell cycle, proliferation, angiogenesis, and inflammation (Figure 6B). Utilizing the GBM and Random Forest algorithms, we identified the top 5 genes most relevant to the prognosis of PRAD for display. CCNA2 and ACSM3 were found to have significant prognostic value in both the TCGA-PRAD and GSE16560 datasets (Figures 6C–F). Subsequently, prognostic KM curves for CCNA2 and ACSM3 were presented, revealing an opposite prognostic difference for ACSM3 in the two datasets, possibly due to insufficient sample size. However, the prognostic difference for CCNA2 in the two datasets remained consistent (Figures 6G–J). Thus, among monocyte-related genes, CCNA2 was identified as the gene with the highest prognostic correlation with PRAD.

3.7 CCNA2 is associated with monocyte infiltration in PRAD

In order to further investigate the relationship between the genes screened in PRAD and immune cell infiltration, we conducted an analysis on the correlation between CCNA2 expression in PRAD and monocytes using the GSCA database. Our results revealed a positive correlation between CCNA2 expression and the level of monocyte infiltration, with a correlation coefficient of 0.23 (Figure 7A). Furthermore, we conducted a correlation analysis on the TCGA-PRAD dataset to explore the relationship between CCNA2 expression and monocyte markers. Our findings indicated a significant association between CCNA2 and the monocyte markers CD14 and HLA-DRA (Figures 7B, C). In light of our research, we observed a strong correlation between CCNA2 and the monocyte marker HLA-DRA. Subsequently, we conducted further analysis to investigate the connection between monocytes expressing high levels of HLA-DR and prostate cancer using Mendelian randomization. Our results indicate that monocytes with elevated HLA-DR expression contribute to the progression of prostate cancer (Figure 7D). The correlation between CCNA2 and immune cell infiltration in PRAD was investigated using single cell analysis from the TISCH2 database. Our findings revealed that CCNA2 was linked to the levels of monocytes and macrophages infiltration in the GSE137829, GSE141445, GSE172301, and GSE176031 datasets (Figures 7E–I).

3.8 Gene enrichment analysis of CCNA2

KEGG analysis revealed that CCNA2 is associated with various pathways in PRAD, including the cell cycle, Human T-cell leukemia virus 1 infection, Proteoglycans in cancer, and Regulation of actin cytoskeleton. Additionally, it is linked to pathways like p53 signaling, TGF-beta signaling, and AGE-RAGE signaling in diabetic complications (Figure 8A). GSEA analysis further highlighted the role of CCNA2 in the immune microenvironment of PRAD, potentially influencing immunotherapy through the PD1 signaling pathway. The association of CCNA2 with transcription factors such as P53, HSF1, and MYC was noted, although experimental validation is needed. Furthermore, CCNA2 was found to regulate PRAD cell senescence, apoptosis, and ferroptosis (Figures 8B–J).

3.9 Analysis of CCNA2 and drug affinity in metastatic PRAD

In order to assess the binding affinity of the key gene CCNA2 with PRAD-targeted drugs, we utilized molecular docking methods for analysis. The CB-Dock2 website, known for its molecular docking analysis capabilities, facilitated our research. The Vina score was employed to measure the binding affinity between genes and drugs. A Vina score below -5 indicates strong binding activity, with lower

10.3389/fimmu.2024.1426474

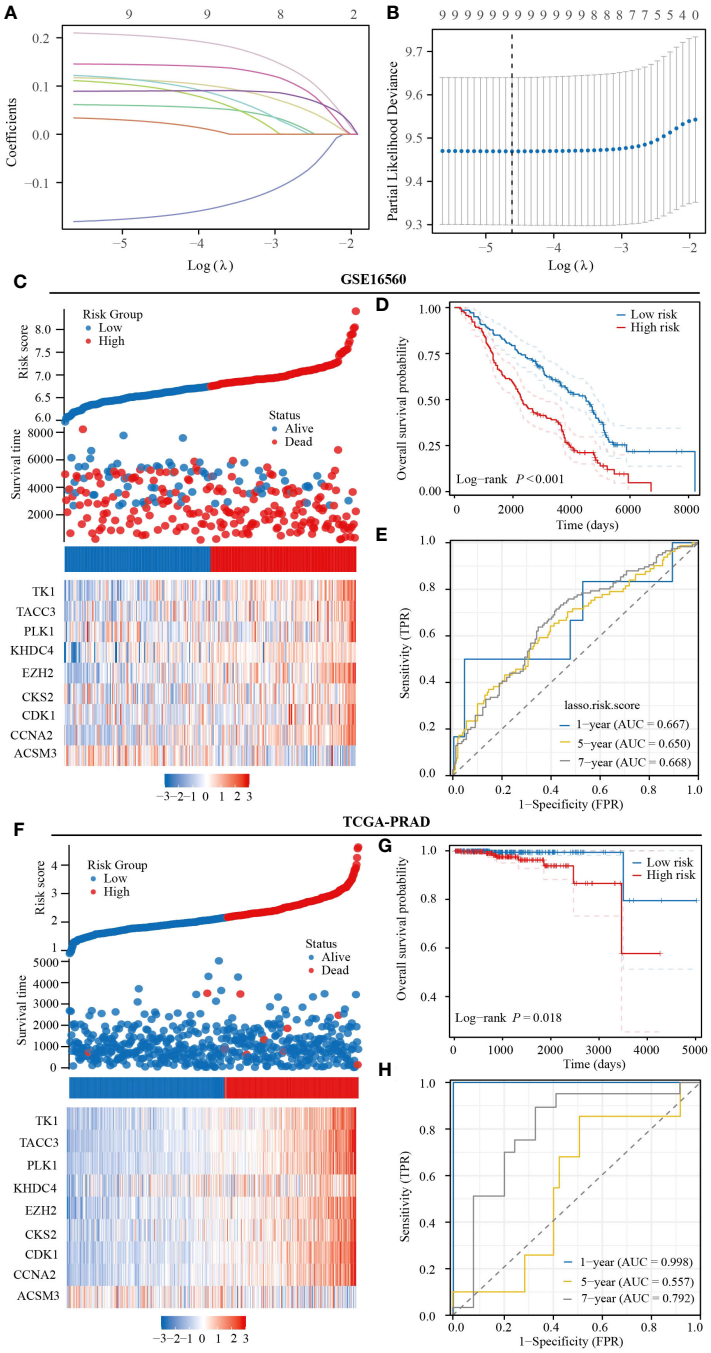


FIGURE 5
Prognostic model constructed based on monocyte-related genes has strong predictive value for the prognosis of PRAD patients. **(A, B)** Prognostic modeling based on the LASSO algorithm. **(C)** Heatmap of expression of prognostic model genes included in the GSE16560 dataset. **(D)** Prognostic differences between patients in the high- and low-risk groups in the GSE16560 dataset. **(E)** Predictive value of the GSE16560 dataset risk score for prognosis in patients with PRAD. **(F)** Heatmap of expression of prognostic model genes included in TCGA-PRAD dataset. **(G)** Prognostic differences between patients in the high- and low-risk groups in TCGA-PRAD dataset. **(H)** Predictive value of the TCGA-PRAD dataset risk score for prognosis in patients with PRAD.

scores indicating higher binding activity. The results of our GSEA analysis revealed a close relationship between CCNA2 and the PD1 signaling pathway. We further investigated the molecular binding affinity of CCNA2 with PD1 inhibitors and found a strong affinity in their molecular structures. With a vina score of -8.7, indicating high binding ability (Figure 9A). Additionally, we examined the binding ability of targeted drugs for metastatic PRAD - Bicalutamide, enzalutamide, and abiraterone - to CCNA2. Our findings demonstrated a strong binding ability of CCNA2 to these drugs at a molecular level (Figures 9B–D). These results not only suggest that CCNA2 may enhance the anti-cancer effects of these drugs but also support the potential of CCNA2 as a drug target.

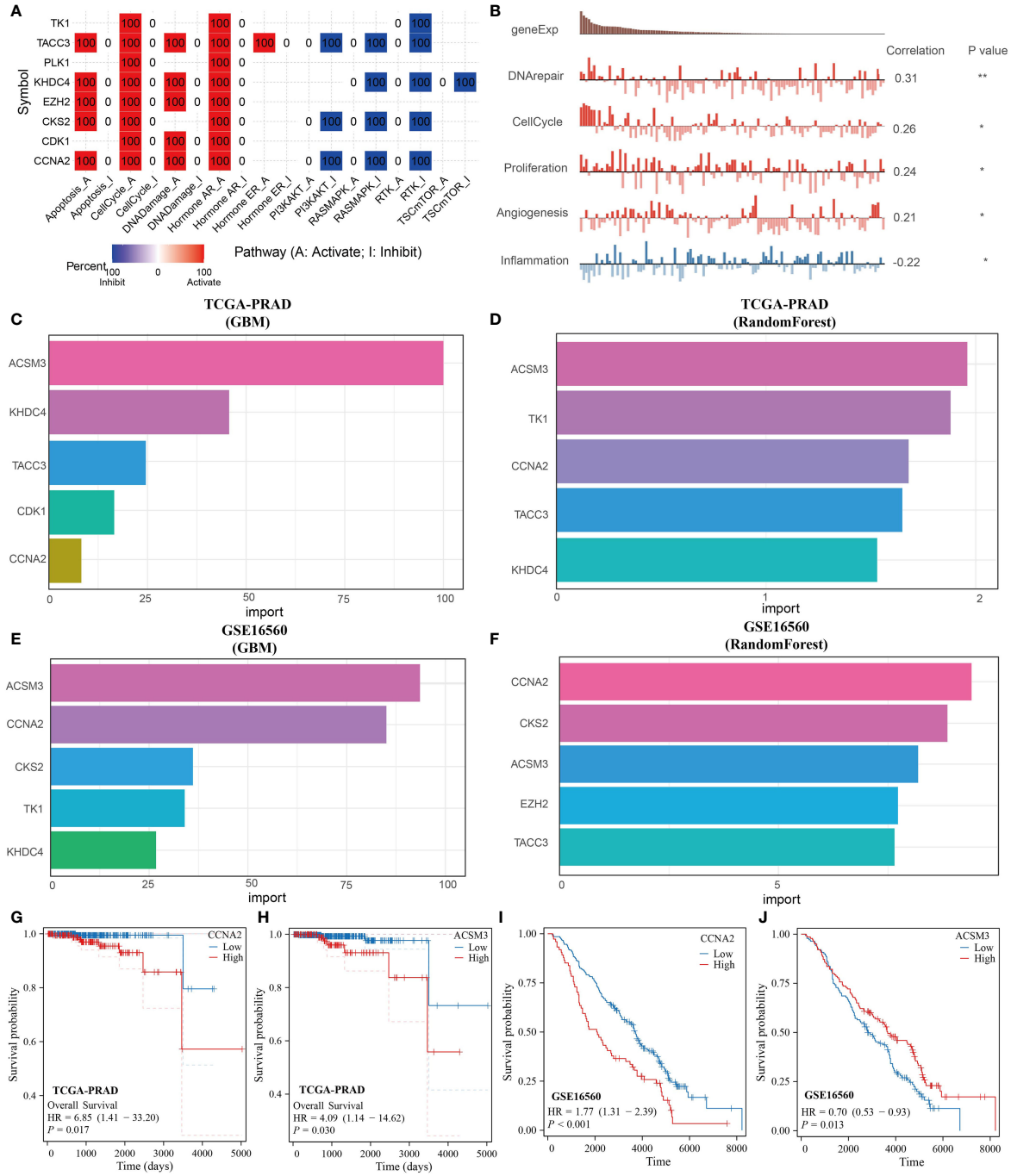


FIGURE 6
CCNA2 identified as the best prognostic gene among monocyte-associated genes. **(A, B)** Functional analysis of monocyte-related prognostic genes. **(C, D)** GBM and RandomForest algorithms to screen key prognostic genes in the TCGA-PRAD dataset. **(E, F)** GBM and RandomForest algorithms to screen key prognostic genes in the GSE16560 dataset. **(G, H)** KM curves of CCNA2 and ACSM3 in the TCGA-PRAD dataset. **(I, J)** KM curves of CCNA2 and ACSM3 in the GSE16560 dataset.

3.10 Expression and prognostic value of CCNA2 in PRAD

In this study, we investigated the role of CCNA2 as a monocyte-related gene in PRAD. A total of 60 paired PRAD samples and corresponding paracancerous samples were collected for analysis of CCNA2 expression differences using immunohistochemical staining. Our results showed a significantly higher expression of CCNA2 in PRAD compared to normal tissues (Figure 10A). Violin plots were also utilized to visually represent the expression variances of CCNA2 in PRAD and normal tissues (Figure 10B). Furthermore, we assessed the diagnostic potential of CCNA2 for PRAD and found promising results, although further validation with larger sample sizes and clinical experiments is necessary

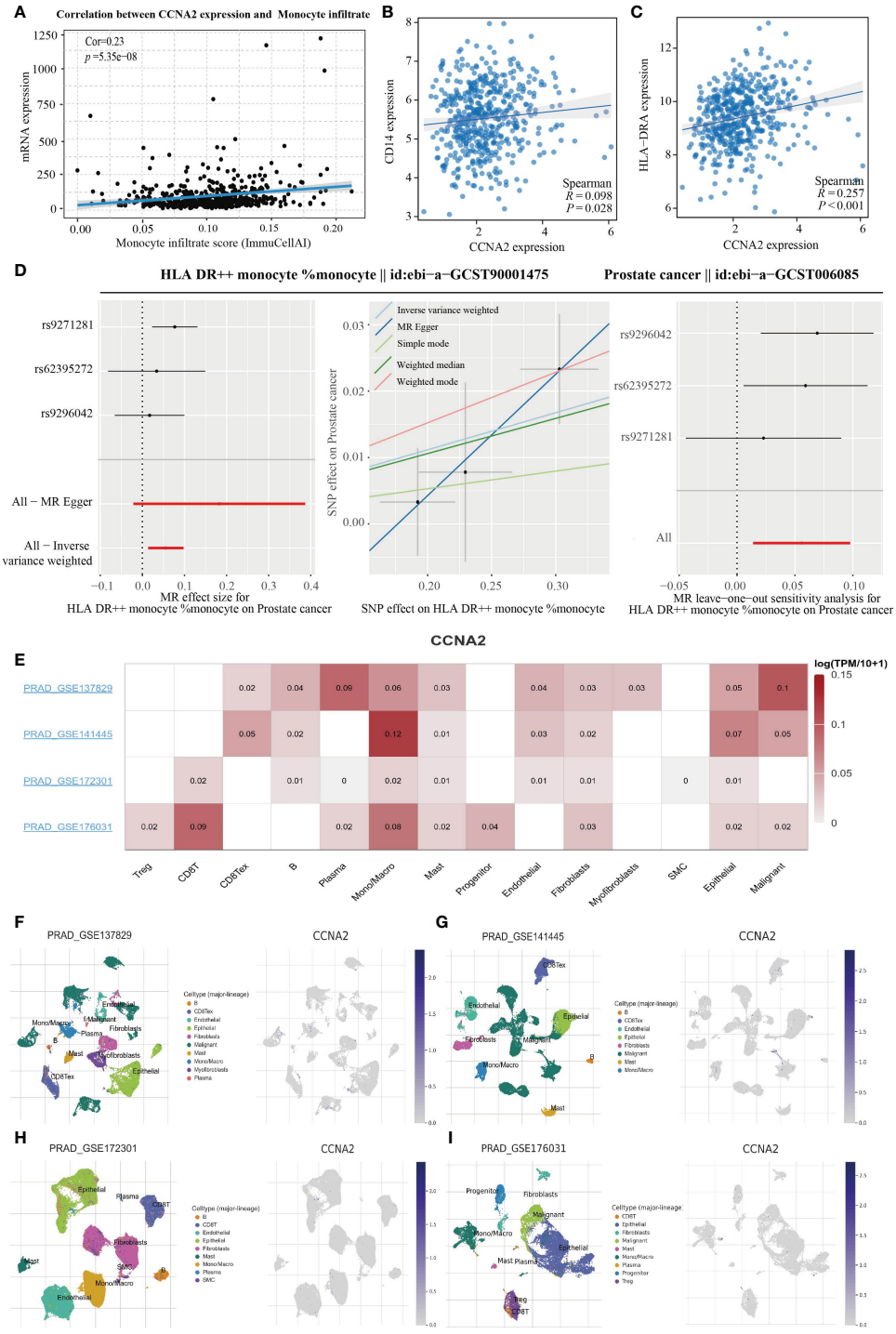


FIGURE 7
CCNA2 positively correlates with monocyte infiltration levels. **(A)** Correlation analysis of CCNA2 and monocyte infiltration levels. **(B, C)** Analysis of CCNA2 correlation with monocyte markers. **(D)** Mendelian randomization analysis of high HLA-DR expressing monocytes in relation to prostate cancer. **(E–I)** Single-cell analysis of the correlation between CCNA2 and immune cell infiltration.

(Figure 10C). Additionally, our analysis revealed a correlation between CCNA2 expression and the prognosis of PRAD patients, indicating a poorer prognosis for those with high CCNA2 expression levels (Figure 10D). Notably, CCNA2 showed strong predictive value for the prognosis of PRAD patients (Figure 10E). In conclusion, our experimental findings confirm the differential expression and prognostic implications of CCNA2 in PRAD.

4 Discussion

PRAD is a highly aggressive tumor with a poor prognosis, often being detected in advanced stages with metastasis (33). Biomarkers are essential in evaluating the therapeutic efficacy and prognosis of tumors and can be an essential component of precision medicine (34). Identifying PRAD and exploring new immune-related

10.3389/fimmu.2024.1426474

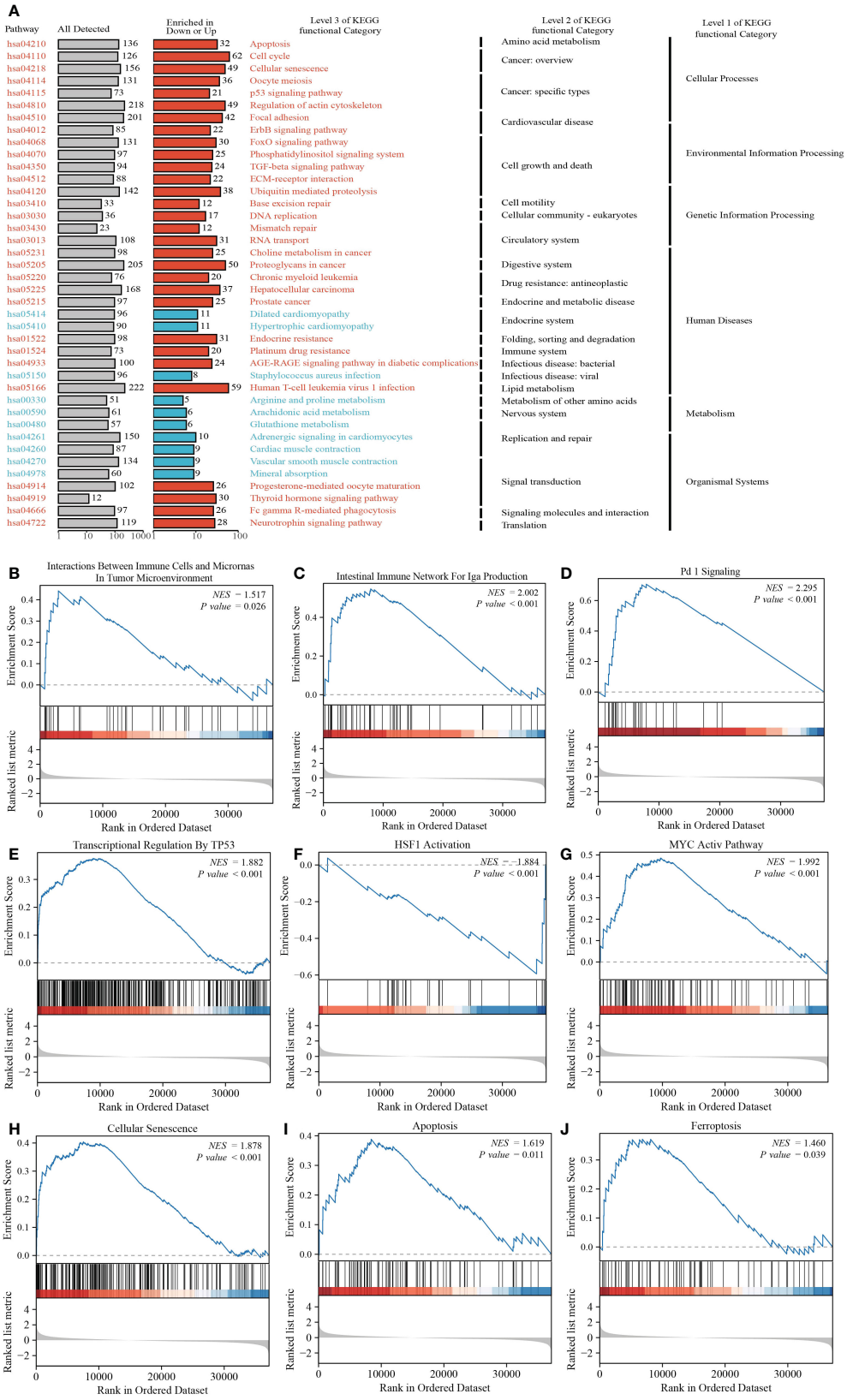


FIGURE 8
Functional analysis of CCNA2 in PRAD. **(A)** KEGG analysis of CCNA2 in PRAD. **(B–J)** GSEA analysis of CCNA2 in PRAD.

10.3389/fimmu.2024.1426474

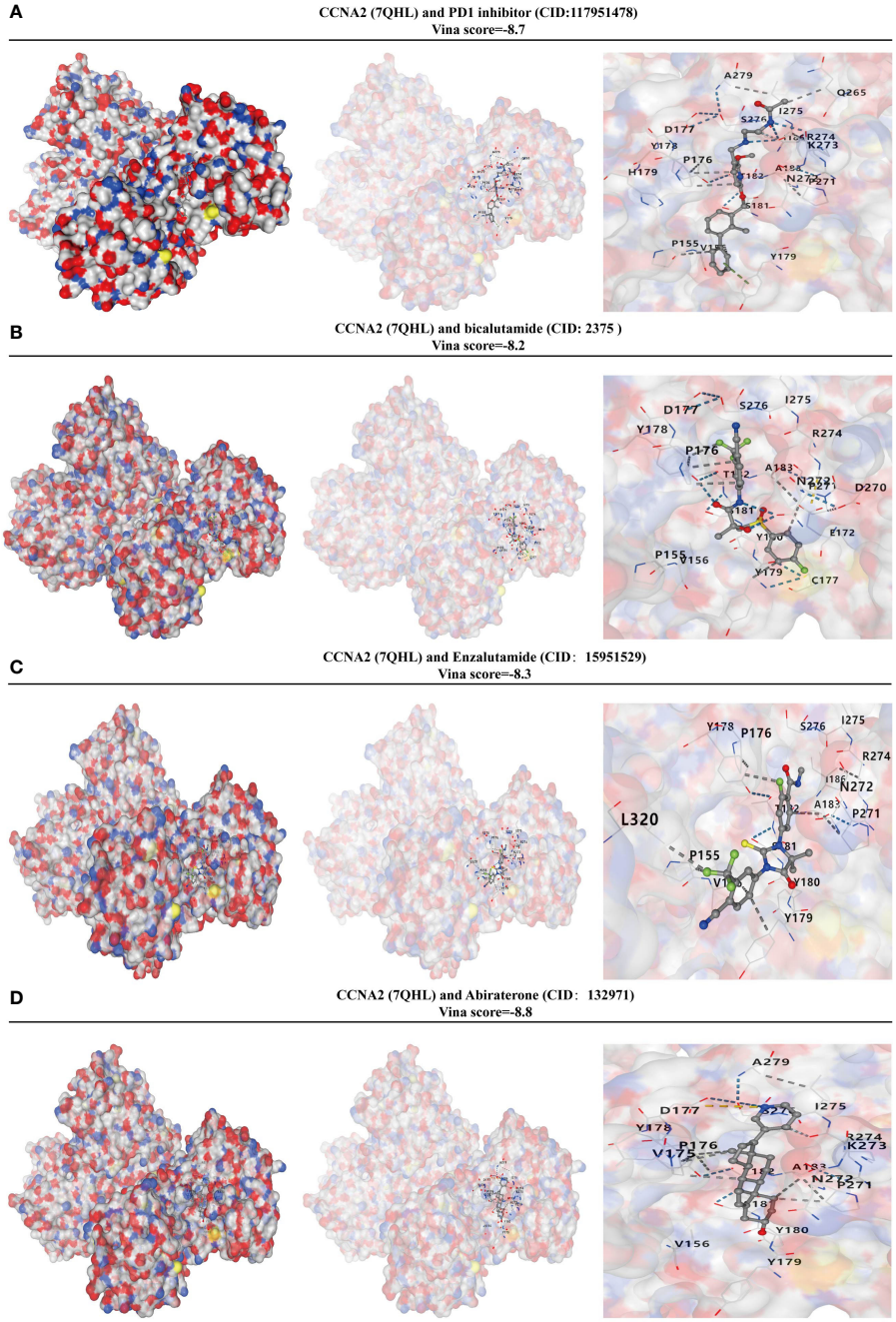


FIGURE 9
CCNA2 has a high binding capacity to PRAD-targeted drugs. **(A)** Analysis of the binding capacity of CCNA2 to PD1 inhibitors. **(B)** Analysis of the binding capacity of CCNA2 to bicalutamide. **(C)** Analysis of the binding capacity of CCNA2 to enzalutamide. **(D)** Analysis of the binding capacity of CCNA2 to abiraterone.

prognostic markers can significantly improve the efficacy of immunotherapy for individuals with this condition. The progression of tumors is closely connected to changes in the tumor microenvironment, where tumor cells impact their surroundings by releasing various chemokines and cytokines (35). Delving into the PRAD tumor microenvironment and discovering novel immune-related markers are essential for developing targeted therapeutic drugs and enhancing patient prognosis.

Research on monocytes in PRAD is increasing, with studies demonstrating their ability to stimulate PRAD cell invasion through pro-inflammatory cytokines (36). Circulating monocytes in metastatic PRAD patients have been found to secrete CHI3L1, promoting tumor growth (37). Our study revealed a correlation between higher levels of monocyte immune infiltration and poorer patient prognosis, aligning with previous findings on the carcinogenic role of monocytes. Through WGCNA analysis, we identified 14 monocyte-related genes. Among

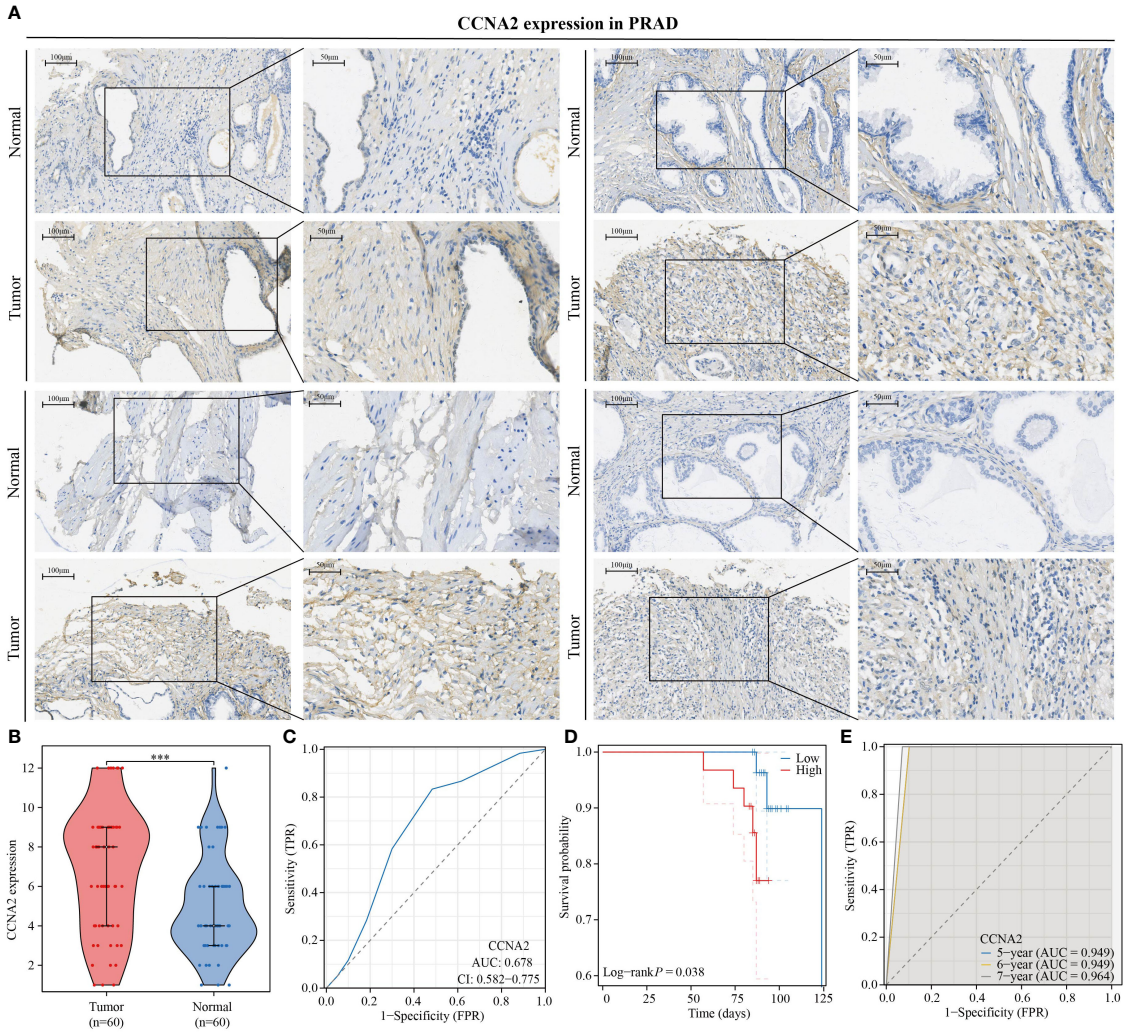


FIGURE 10
CCNA2 is highly expressed in PRAD and is associated with poor patient prognosis. **(A, B)** Differential expression of CCNA2 in PRAD. **(C)** Diagnostic predictive value of CCNA2 in PRAD. **(D)** KM curve of overall survival of CCNA2 in PRAD. **(E)** Prognostic predictive value of CCNA2 in PRAD.

these genes, CCNA2, CDK1, CKS2, EZH2, HMGB3, KHDC4, KIF2C, PKMYT1, and PLK1 were found to be associated with various PRAD stages, Gleason scores, and PSA scores, further highlighting their significance in PRAD. Previous studies have also highlighted the importance of CCNA2 in PRAD using WGCNA analysis, but its prognostic value and correlation with monocytes have not been confirmed with clinical samples (38). CDK1 has been identified as a key player in promoting tumor metastasis in prostate cancer cells through its influence on the epithelial-mesenchymal transition process. This is achieved by regulating the phosphorylation of the ERK/GSK3β/SNAIL pathway (39). Additionally, CDK1 has been found to modulate the phosphorylation of the androgen receptor, with its inhibitors demonstrating the ability to enhance the effectiveness of enzalutamide in targeting prostate cancer cells (40, 41). Aberrant expression of CKS2 promotes prostate tumorigenesis by promoting proliferation and inhibiting programmed cell death (42). EZH2, HMGB3, KIF2C, PKMYT1 and PLK1 have also been confirmed to be related to PRAD progression (43–47). Subsequently, we developed a diagnostic model for monocyte-related genes using machine learning techniques. The Enet[alpha=0.4] method was the most recent approach employed to construct a diagnostic model for PRAD. Our analysis revealed that the diagnostic models built on the training set TCGA-PRAD and validation set GSE32571 demonstrated strong predictive value. However, in the validation set GSE62872, the AUC value was only 0.674, potentially influenced by the expression of TK1. Due to the imbalanced distribution of samples in the TCGA-PRAD data set, with only 10 patients deceased out of 498 samples, we opted to develop a prognostic model using the GSE16560 data set and validate it with the TCGA-PRAD data set. Our findings indicate that the prognostic model we created demonstrated robust predictive capabilities for the prognosis of PRAD patients, particularly at the 1-year and 7-year. GBM and RF algorithms were utilized to identify the genes most pertinent to PRAD prognosis within the monocyte-related genes. CCNA2 and ACSM3 were initially identified as the most relevant genes, but due to inconsistencies with prognostic correlation results in the TCGA-PRAD and GSE16560 datasets, ACSM3 was subsequently excluded. Ultimately, CCNA2 was identified as the gene most relevant to PRAD prognosis among the monocyte-related genes.

Single-cell genomics offers a novel approach to investigate the tumor immune microenvironment by conducting genomic analysis at the single-cell level. An increasing number of studies are utilizing this method to gain valuable insights (48–50). The correlation between CCNA2 and monocytes was examined using data from the TISCH2 websites. Our analysis showed a positive relationship between CCNA2 and monocyte infiltration levels. Additionally, CCNA2 was found to be positively associated with the expression of monocyte markers in PRAD. Through KEGG and GSEA analysis, we uncovered the significant role of CCNA2 in PRAD, potentially regulating cell senescence, apoptosis, and ferroptosis. Investigation into the correlation between CCNA2 and therapeutic drugs for PRAD revealed a strong binding affinity between CCNA2 and three specific drugs targeting PRAD. Moreover, CCNA2 exhibited strong binding capabilities with PD1 inhibitors, suggesting its potential as a drug targeting PRAD. However, our study is limited by a small sample size, which may have impacted our findings. It is essential to expand the sample size and validate these conclusions through further experimentation.

5 Conclusion

Our study highlighted the significant roles of monocyte-related genes in PRAD. Furthermore, we created and tested models utilizing different machine learning techniques to forecast the diagnosis and prognosis of PRAD patients. These results enhance our comprehension of monocyte infiltration patterns and underscore the significance of the monocyte-related gene CCNA2 as a valuable prognostic and diagnostic indicator for PRAD. These insights pave the way for personalized treatment strategies for patients with PRAD.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Committee of Shanghai Outdo Biotech Company. 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. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

YW: Conceptualization, Data curation, Formal analysis, Validation, Writing – original draft. CL: Conceptualization, Data curation, Writing – original draft. JH: Data curation, Methodology, Writing – original draft. QZ: Software, Writing – original draft. YZ: Software, Writing – original draft. HS: Software, Writing – original draft. HZ: Validation, Writing – review & editing. BD: Funding acquisition, Writing – review & editing. MR: Funding acquisition, Project administration, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research received funding from the Heilongjiang Province Natural Science Foundation (LH2019H030) and the National Natural Science Foundation of China (82002680).

Acknowledgments

We would also like to thank the Shengxin Bean Sprout Platform (http://www.sxdyc.com/index) for its help in data statistical analysis.

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.

Publisher's note

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin.* (2024) 74:12–49. doi: 10.3322/caac.21820

2. Liu J, Dong L, Zhu Y, Dong B, Sha J, Zhu HH, et al. Prostate cancer treatment - China's perspective. *Cancer Lett.* (2022) 550:215927. doi: 10.1016/j.canlet.2022.215927

3. Sun Z, Wang J, Zhang Q, Meng X, Ma Z, Niu J, et al. Coordinating single-cell and bulk RNA-seq in deciphering the intratumoral immune landscape and prognostic stratification of prostate cancer patients. *Environ Toxicol.* (2024) 39:657–68. doi: 10.1002/tox.23928

4. Ananias HJ, van den Heuvel MC, Helfrich W, de Jong IJ. Expression of the gastrin-releasing peptide receptor, the prostate stem cell antigen and the prostate-specific membrane antigen in lymph node and bone metastases of prostate cancer. *Prostate.* (2009) 69:1101–8. doi: 10.1002/pros.20957

5. Ji Y, Sun K, Yang Y, Wu Z. Dihydroartemisinin ameliorates innate inflammatory response induced by Streptococcussuis-derived muramidase-released protein via inactivation of TLR4-dependent NF-κB signaling. *J Pharm Anal.* (2023) 13:1183–94. doi: 10.1016/j.jpha.2023.05.013

6. Narasimhan PB, Marcovecchio P, Hamers AAJ, Hedrick CC. Nonclassical monocytes in health and disease. *Annu Rev Immunol*. (2019) 37:439–56. doi: 10.1146/annurev-immunol-042617–053119

7. Yin J, Albers AJ, Smith TS, Riddell GT, Richards JO. Differential regulation of human monocytes and NK cells by antibody-opsonized tumors. *Cancer Immunol Immunother*. (2018) 67:1239–50. doi: 10.1007/s00262-018-2179-z

8. Hanna RN, Cekic C, Sag D, Tacke R, Thomas GD, Nowyhed H, et al. Patrolling monocytes control tumor metastasis to the lung. *Science*. (2015) 350:985–90. doi: 10.1126/science.aac9407

9. Olingy CE, Dinh HQ, Hedrick CC. Monocyte heterogeneity and functions in cancer. *J Leukoc Biol*. (2019) 106:309–22. doi: 10.1002/JLB.4RI0818-311R

10. Feng D, Shi X, Li D, Wu R, Wang J, Wei W, et al. M2 macrophage-related molecular subtypes and prognostic index for prostate cancer patients through integrating single-cell and bulk RNA sequencing analysis. *Genes Dis*. (2023) 11:101086. doi: 10.1016/j.gendis.2023.101086

11. Nonomura N, Takayama H, Nakayama M, Nakai Y, Kawashima A, Mukai M, et al. Infiltration of tumour-associated macrophages in prostate biopsy specimens is predictive of disease progression after hormonal therapy for prostate cancer. *BJU Int*. (2011) 107:1918–22. doi: 10.1111/j.1464-410X.2010.09804.x

12. Åberg AM, Bergström SH, Thysell E, Tjon-Kon-Fat LA, Nilsson JA, Widmark A, et al. High monocyte count and expression of S100A9 and S100A12 in peripheral blood mononuclear cells are associated with poor outcome in patients with metastatic prostate cancer. *Cancers (Basel)*. (2021) 13:2424. doi: 10.3390/cancers13102424

13. Zhang B, Liu J, Li H, Huang B, Zhang B, Song B, et al. Integrated multi-omics identified the novel intratumor microbiome-derived subtypes and signature to predict the outcome, tumor microenvironment heterogeneity, and immunotherapy response for pancreatic cancer patients. *Front Pharmacol*. (2023) 14:1244752. doi: 10.3389/fphar.2023.1244752

14. Wang J, Zheng T, Liao Y, Geng S, Li J, Zhang Z, et al. Machine learning prediction model for post- hepatectomy liver failure in hepatocellular carcinoma: A multicenter study. *Front Oncol*. (2022) 12:986867. doi: 10.3389/fonc.2022.986867

15. Pan C, Deng D, Wei T, Wu Z, Zhang B, Yuan Q, et al. Metabolomics study identified bile acids as potential biomarkers for gastric cancer: A case control study. *Front Endocrinol (Lausanne)*. (2022) 13:1039786. doi: 10.3389/fendo.2022.1039786

16. Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism*. (2017) 69S: S36–40. doi: 10.1016/j.metabol.2017.01.011

17. Chowdhury MA, Zhang JJ, Rizk R, Chen WCW. Stem cell therapy for heart failure in the clinics: new perspectives in the era of precision medicine and artificial intelligence. *Front Physiol*. (2024) 14:1344885. doi: 10.3389/fphys.2023.1344885

18. Liu J, Shi Y, Zhang Y. Multi-omics identification of an immunogenic cell death-related signature for clear cell renal cell carcinoma in the context of 3P medicine and based on a 101-combination machine learning computational framework. *EPMA J*. (2023) 14:275–305. doi: 10.1007/s13167–023-00327–3

19. Xu F, Huang X, Li Y, Chen Y, Lin L. m6A-related lncRNAs are potential biomarkers for predicting prognoses and immune responses in patients with LUAD. *Mol Ther Nucleic Acids*. (2021) 24:780–91. doi: 10.1016/j.omtn.2021.04.003

20. Guan H, Chen X, Liu J, Sun J, Guo H, Jiang Y, et al. Molecular characteristics and therapeutic implications of Toll-like receptor signaling pathway in melanoma. *Sci Rep*. (2023) 13:13788. doi: 10.1038/s41598-023-38850-y

21. Liu CJ, Hu FF, Xie GY, Miao YR, Li XW, Zeng Y, et al. GSCA: an integrated platform for gene set cancer analysis at genomic, pharmacogenomic and immunogenomic levels. *Brief Bioinform*. (2023) 24:bbac558. doi: 10.1093/bib/bbac558

22. Yuan H, Yan M, Zhang G, Liu W, Deng C, Liao G, et al. CancerSEA: a cancer single-cell state atlas. *Nucleic Acids Res*. (2019) 47:D900–8. doi: 10.1093/nar/gky939

23. Zhang S, Jiang C, Jiang L, Chen H, Huang J, Gao X, et al. Construction of a diagnostic model for hepatitis B-related hepatocellular carcinoma using machine learning and artificial neural networks and revealing the correlation by immunoassay. *Tumour Virus Res*. (2023) 16:200271. doi: 10.1016/j.tvr.2023.200271

24. Han Y, Wang Y, Dong X, Sun D, Liu Z, Yue J, et al. TISCH2: expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment. *Nucleic Acids Res*. (2023) 51:D1425–31. doi: 10.1093/nar/gkac959

25. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. (2018) 7:e34408. doi: 10.7554/eLife.34408

26. Xiong J, Chi H, Yang G, Zhao S, Zhang J, Tran LJ, et al. Revolutionizing anti-tumor therapy: unleashing the potential of B cell-derived exosomes. *Front Immunol*. (2023) 14:1188760. doi: 10.3389/fimmu.2023.1188760

27. Zhang B, Sun J, Guan H, Guo H, Huang B, Chen X, et al. Integrated single-cell and bulk RNA sequencing revealed the molecular characteristics and prognostic roles of neutrophils in pancreatic cancer. *Aging (Albany NY)*. (2023) 15:9718–42. doi: 10.18632/aging.205044

28. Yang H, Li Z, Zhu S, Wang W, Zhang J, Zhao D, et al. Molecular mechanisms of pancreatic cancer liver metastasis: the role of PAK2. *Front Immunol*. (2024) 15:1347683. doi: 10.3389/fimmu.2024.1347683

29. Sun Y, Yang X, Ren S, Lu Z, Liu Z, Kong F, et al. Stratification of risk based on immune signatures and prediction of the efficacy of immune checkpoint inhibitors in prostate cancer. *J Men's Health*. (2023) 19:16–33. doi: 10.22514/jomh.2023.113

30. Zhang B, Zhou YL, Chen X, Wang Z, Wang Q, Ju F, et al. Efficacy and safety of CTLA-4 inhibitors combined with PD-1 inhibitors or chemotherapy in patients with advanced melanoma. *Int Immunopharmacol*. (2019) 68:131–6. doi: 10.1016/j.intimp.2018.12.034

31. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. (2015) 12:453–7. doi: 10.1038/nmeth.3337

32. Zhang B, Chen X, Wang Z, Guo F, Zhang X, Huang B, et al. Identifying endoplasmic reticulum stress-related molecular subtypes and prognostic model for predicting the immune landscape and therapy response in pancreatic cancer. *Aging (Albany NY)*. (2023) 15:10549–79. doi: 10.18632/aging.205094

33. Yu Q, Gao Y, Dai W, Li D, Zhang L, Hameed MMA, et al. Cell membrane-camouflaged chitosan-polypyrrole nanogels co-deliver drug and gene for targeted chemotherapy and bone metastasis inhibition of prostate cancer. *Adv Healthc Mater*. (2024) 6:e2400114. doi: 10.1002/adhm.202400114

34. Zhu C, Sun Z, Wang J, Meng X, Ma Z, Guo R, et al. Exploring oncogenes for renal clear cell carcinoma based on G protein-coupled receptor-associated genes. *Discovery Oncol*. (2023) 14:182. doi: 10.1007/s12672-023-00795-z

35. Liu Q, Liao L. Identification of macrophage-related molecular subgroups and risk signature in colorectal cancer based on a bioinformatics analysis. *Autoimmunity*. (2024) 57:2321908. doi: 10.1080/08916934.2024.2321908

36. Begley LA, Kasina S, Mehra R, Adsule S, Admon AJ, Lonigro RJ, et al. CXCL5 promotes prostate cancer progression. *Neoplasia*. (2008) 10(3):244–54. doi: 10.1593/neo.07976

37. Cavassani KA, Meza RJ, Habiel DM, Chen JF, Montes A, Tripathi M, et al. Circulating monocytes from prostate cancer patients promote invasion and motility of epithelial cells. *Cancer Med*. (2018) 7:4639–49. doi: 10.1002/cam4.1695

38. Yang R, Du Y, Wang L, Chen Z, Liu X. Weighted gene co-expression network analysis identifies CCNA2 as a treatment target of prostate cancer through inhibiting cell cycle. *J Cancer*. (2020) 11:1203–11. doi: 10.7150/jca.38173

39. Zhang B, Zhang M, Li Q, Yang Y, Shang Z, Luo J. TPX2 mediates prostate cancer epithelial-mesenchymal transition through CDK1 regulated phosphorylation of ERK/GSK3β/SNAIL pathway. *Biochem Biophys Res Commun*. (2021) 546:1–6. doi: 10.1016/j.bbrc.2021.01.106

40. Ji G, He S, Huang C, Gong Y, Li X, Zhou L. Upregulation of ATP Binding Cassette Subfamily C Member 5 facilitates Prostate Cancer progression and Enzalutamide resistance via the CDK1-mediated AR Ser81 Phosphorylation Pathway. *Int J Biol Sci*. (2021) 17:1613–28. doi: 10.7150/ijbs.59559

41. Willder JM, Heng SJ, McCall P, Adams CE, Tannahill C, Fyffe G, et al. Androgen receptor phosphorylation at serine 515 by Cdk1 predicts biochemical relapse in prostate cancer patients. *Br J Cancer*. (2013) 108:139–48. doi: 10.1038/bjc.2012.480

42. Lan Y, Zhang Y, Wang J, Lin C, Ittmann MM, Wang F. Aberrant expression of Cks1 and Cks2 contributes to prostate tumorigenesis by promoting proliferation and inhibiting programmed cell death. *Int J Cancer*. (2008) 123:543–51. doi: 10.1002/ijc.23548

43. Zhang Z, Wang X, Kim M, He D, Wang C, Fong KW, et al. Downregulation of EZH2 inhibits epithelial-mesenchymal transition in enzalutamide-resistant prostate cancer. *Prostate*. (2023) 83:1458–69. doi: 10.1002/pros.24602

44. Xu Y, Xu M, Li X, Weng X, Su Z, Zhang M, et al. SOX9 and HMGB3 co-operatively transactivate NANOG and promote prostate cancer progression. *Prostate*. (2023) 83:440–53. doi: 10.1002/pros.24476

45. Zhang P, Gao H, Ye C, Yan R, Yu L, Xia C, et al. Large-scale transcriptome data analysis identifies KIF2C as a potential therapeutic target associated with immune infiltration in prostate cancer. *Front Immunol*. (2022) 13:905259. doi: 10.3389/fimmu.2022.905259

46. Wang J, Wang L, Chen S, Peng H, Xiao L, Du E, et al. PKMYT1 is associated with prostate cancer Malignancy and may serve as a therapeutic target. *Gene*. (2020) 744:144608. doi: 10.1016/j.gene.2020.144608

47. Liang H, Yang C, Zeng R, Song Y, Wang J, Xiong W, et al. Targeting CBX3 with a dual BET/PLK1 inhibitor enhances the antitumor efficacy of CDK4/6 inhibitors in prostate cancer. *Adv Sci (Weinh)*. (2023) 10:e2302368. doi: 10.1002/advs.202302368

48. Ma B, Qin L, Sun Z, Wang J, Tran LJ, Zhang J, et al. The single-cell evolution trajectory presented different hypoxia heterogeneity to reveal the carcinogenesis of genes in clear cell renal cell carcinoma: Based on multiple omics and real experimental verification. *Environ Toxicol*. (2024) 39:869–81. doi: 10.1002/tox.24009

49. Wang J, Zuo Z, Yu Z, Chen Z, Meng X, Ma Z, et al. Single-cell transcriptome analysis revealing the intratumoral heterogeneity of ccRCC and validation of MT2A in pathogenesis. *Funct Integr Genomics*. (2023) 23:300. doi: 10.1007/s10142–023-01225–7

50. Jiang S, Yang X, Lin Y, Liu Y, Tran LJ, Zhang J, et al. Unveiling Anoikis-related genes: A breakthrough in the prognosis of bladder cancer. *J Gene Med*. (2024) 26:e3651. doi: 10.1002/jgm.3651

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 17 June 2024
DOI 10.3389/fimmu.2024.1427475

OPEN ACCESS

EDITED BY
Raquel Alarcon Rodriguez,
University of Almeria, Spain

REVIEWED BY
Qihang Yuan,
Dalian Medical University, China
Chao Li,
LMU Munich University Hospital, Germany

*CORRESPONDENCE
Wanqing Wei
wanqingwei@tmu.edu.cn

†These authors have contributed equally to this work

RECEIVED 03 May 2024
ACCEPTED 05 June 2024
PUBLISHED 17 June 2024

CITATION
Wen X, Hou J, Qi T, Cheng X, Liao G, Fang S, Xiao S, Qiu L and Wei W (2024) Anoikis resistance regulates immune infiltration and drug sensitivity in clear-cell renal cell carcinoma: insights from multi omics, single cell analysis and *in vitro* experiment. *Front. Immunol.* 15:1427475. doi: 10.3389/fimmu.2024.1427475

Anoikis resistance regulates immune infiltration and drug sensitivity in clear-cell renal cell carcinoma: insights from multi omics, single cell analysis and *in vitro* experiment

Xiangyang Wen[1†], Jian Hou[2†], Tiantian Qi[3†], Xiaobao Cheng[2†], Guoqiang Liao[1], Shaohong Fang[1], Song Xiao[1], Longlong Qiu[1] and Wanqing Wei[4*]

[1]The Department of Surgery, Shenzhen Longgang Second People's Hospital, Shenzhen, China, [2]Department of Urology, The University of Hongkong-Shenzhen Hospital, Shenzhen, China, [3]Department of Bone & Joint Surgery, Peking University Shenzhen Hospital, Shenzhen, China, [4]Department of Urology, Lianshui People's Hospital of Kangda College Affiliated to Nanjing Medical University, Huaian, China

Background: Anoikis is a form of programmed cell death essential for preventing cancer metastasis. In some solid cancer, anoikis resistance can facilitate tumor progression. However, this phenomenon is underexplored in clear-cell renal cell carcinoma (ccRCC).

Methods: Using SVM machine learning, we identified core anoikis-related genes (ARGs) from ccRCC patient transcriptomic data. A LASSO Cox regression model stratified patients into risk groups, informing a prognostic model. GSVA and ssGSEA assessed immune infiltration, and single-cell analysis examined ARG expression across immune cells. Quantitative PCR and immunohistochemistry validated ARG expression differences between immune therapy responders and non-responders in ccRCC.

Results: ARGs such as CCND1, CDKN3, PLK1, and BID were key in predicting ccRCC outcomes, linking higher risk with increased Treg infiltration and reduced M1 macrophage presence, indicating an immunosuppressive environment facilitated by anoikis resistance. Single-cell insights showed ARG enrichment in Tregs and dendritic cells, affecting immune checkpoints. Immunohistochemical analysis reveals that ARGs protein expression is markedly elevated in ccRCC tissues responsive to immunotherapy.

Conclusion: This study establishes a novel anoikis resistance gene signature that predicts survival and immunotherapy response in ccRCC, suggesting that manipulating the immune environment through these ARGs could improve therapeutic strategies and prognostication in ccRCC.

KEYWORDS

renal cell carcinoma, anoikis, immune microenvironment, prognosis, signature

1 Introduction

Clear-cell renal cell carcinoma (ccRCC) represents the most prevalent subtype of renal carcinoma, accounting for approximately 75% of all kidney cancer cases (1). Surgical intervention and chemotherapy currently dominate the therapeutic landscape for this malignancy. Despite a relatively high overall survival rate associated with ccRCC, the occurrence of metastasis in advanced stages drastically reduces the five-year survival rate to below 8% (2). Due to the high recurrence rate and poor prognosis of kidney cancer, it is crucial to inhibit the distant metastasis of renal tumor cells. Whereas tumorigenesis and metastasis are closely related to changes in the tumor microenvironment and the migration ability of tumor cells (3).

Anoikis, a programmed cell death, is triggered by the loss of interactions between cells and the extracellular matrix (ECM) (4). In normal cells, these interactions are disrupted by molecules that initiate anoikis on the cell surface and by glycosylated ECM proteins, leading to apoptosis and cell death. The ECM confines tumor cells to a fixed site within the tissue. Tumor cells that acquire migratory capabilities and move to vascular sites develop resistance to anoikis, allowing them to metastasize to distant locations via the bloodstream, thus forming metastatic foci (5–7). Recent studies have uncovered molecular pathways and mechanisms that regulate resistance to anoikis, including cell adhesion molecules, growth factors, and signaling pathways that induce epithelial-to-mesenchymal transition (8). Downstream molecules in these pathways, such as PI3K/AKT (9) and ERK1/2 (10), play significant roles in apoptosis resistance and survival promotion. The latest research indicates that the Hippo pathway and collagen XIII are linked to anoikis resistance in breast cancer (11, 12).

T cells in the body perform surveillance functions, identifying and eliminating abnormal cells, thereby restricting the survival of tumor cells. The role of immune cell infiltration in shaping the tumor microenvironment and influencing tumor progression has been well recognized (13, 14). Numerous studies have highlighted the impact of immune cell apoptosis on the development and progression of various malignancies, including lung, breast, and endometrial cancers. For instance, research by K. Planells et al. suggested that silencing FAIM2 can inhibit the survival and drug resistance by regulating T cells (15). Additionally, the influence of L1CAM on the prognosis of endometrial cancer has been associated with its role in promoting Treg infiltration, thus impairing resistance to apoptosis (16). While existing research has elucidated the link between immune cell apoptosis and the prognosis of various cancers (17, 18), tumor cells can evade immune detection by acquiring resistance to anoikis (19). Although clinical treatments for kidney cancer include radical surgical interventions, chemotherapy and immunotherapy, there is still a lack of recognized and reliable standard predictors for the diagnosis and prognosis of early-stage kidney cancer. The relationship between immune cells and anoikis, as well as the impact of anoikis on the survival of ccRCC patients, has been minimally explored. Exploring the abnormal performance of immune cells and anoikis within renal cancer tissues holds the potential to uncover new molecular biomarkers that could enhance the accuracy of renal cancer diagnosis and prognosis assessment.

In this study, we developed a prognostic model related to anoikis that stratifies ccRCC patients into different risk categories. Through multi-omics and single-cell analyses, we elucidated the relationship between anoikis and immune cell infiltration across various risk groups. To gain insight into the role of anoikis in cancer immunotherapy, we further explored its relevance to the tumor microenvironment. We investigated its relationship with various immune processes and factors, including immune cell infiltration, immunosuppressive factors, and immunostimulatory factors. Moreover, using quantitative real-time PCR (qRT-PCR) and immunohistochemistry (IHC), we validated the expression patterns of four anoikis-related genes in ccRCC patients responding to immunotherapy. This reveals the potential role of anoikis in influencing the efficacy of immunotherapy and provides novel targets for immunotherapeutic strategies.

2 Materials and methods

2.1 Data collection process

The Figure 1 shows the flowchart of this study. The "TCGAbiolinks" R package was utilized to retrieve transcriptional data for clear cell renal carcinoma from the TCGA database (TCGA-KIRC; http://cancergenome.nih.gov). We downloaded the data of 542 ccRCC tumor tissues and 72 adjacent non-tumoral tissues. While the chi-square test was applied to compare the clinical characteristics of the two data sets to ensure that random matching did not bias the distribution of clinical characteristics. Anoikis-related genes (ARGs) were sourced from the GeneCards database (https://www.genecards.org/), resulting in the acquisition of 358 ARGs. Single-cell data and a validation cohort for ccRCC transcriptomes were procured from the GEO database.

2.2 Differential ARGs identification

We performed differential gene expression analysis on the TCGA-KIRC dataset (including tumor and adjacent non-tumor RNA transcriptomes) using the limma software package to identify differentially expressed genes (DEGs) ($P < 0.05$) (20). An intersection with the 358 ARGs yielded a subset of differential ARGs for subsequent analyses.

2.3 Anoikis functional enrichment

The R packages "GSVA" (21) and "GSEABase" (22) were employed to perform enrichment analysis on the DEGs. The focus of the analysis was to identify the enrichment levels of the anoikis-DEGs on KEGG (Kyoto Encyclopedia of Genes and Genomes) and GO (Gene Ontology) (23). In addition, we have

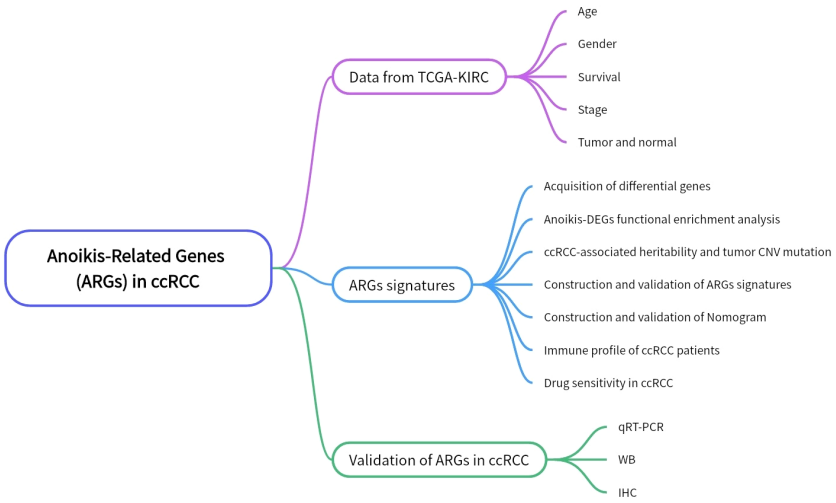


FIGURE 1
The workflow of this study.

applied Cytoscape software (24) and the STRING database (https://cn.string-db.org/) to explore the protein-protein interactions (25).

2.4 Construction of anoikis-related prognostic risk model

First, we selected core prognostic anoikis-related genes (ARGs) using the SVM-RFE algorithm (26). Then, we evaluated the relationship between ARG expression levels and survival in ccRCC patients with univariate Cox regression analysis. Further, we constructed a prognostic model using LASSO Cox regression, utilizing the R packages "survival" and "forestplot". Subsequently, multivariate Cox proportional hazards regression analysis was conducted to identify critical clinical phenotypes.

The risk score was calculated as the sum of the products of the coefficients and the expression levels of core ARGs (27). We categorized patients into high-risk and low-risk groups based on ARGs-Riskscore median value. In addition, we performed principal component analysis (PCA) mix with Kaplan-Meier analysis to investigate the relationship between the anoikis-based risk scoring and overall survival of ccRCC patients (28). The accuracy of the predictive model was further evaluated using ROC curves. At last, we applied univariate-multivariate Cox regression to validate the predictive power of this risk score model.

2.5 Construction and validation of nomogram

Using multivariate Cox and stepwise regression, we incorporated age, TMN staging, and risk scores to construct a prognostic Nomogram to predict overall survival in patients with ccRCC (27, 29). Calibration plots and decision curve analysis (DCA) were constructed to confirm the model's efficacy and clinical relevance (30). We evaluated the prognostic utility of this model sequentially in the training and test cohorts and in the entire TCGA dataset. This evaluation employed receiver operating characteristic (ROC) curves to assess 1, 3, and 5-year survival predictions (31). Independent prognostic determinants were delineated through sequential univariate and multivariate Cox regression analyses (32), which considered risk scores derived from patient age, gender, and comprehensive TNM staging. The development and assessment of the nomogram were facilitated by the "rms" package in R, ensuring robust discrimination and calibration capabilities within the training dataset.

2.6 Immune profile of ccRCC patients based on anoikis resistance

On the basis of the expression of the anoikis-related genes (ARGs), we divided ccRCC patients into two groups: high-risk and low-risk. Then we used the ssGSEA approach to profile the cellular composition of the tumor microenvironment (TME) in the two risk groups (33), while matrix scores and immune scores were assessed to identify differences between these categories. Spearman's analysis was employed to correlate immune cell characteristics with risk scores. And we examined the immune profiles of all patients through various computational techniques (including cibsort, timer, abs, quantitative, XCELL, and EPIC) (34, 35). Finally, we employed the ssGSEA methodology to assess the immune landscape and scrutinized checkpoint molecules to highlight differences between two groups.

2.7 Single-cell analysis of ccRCC based on ARGs

We processed single-cell RNA sequence data using the protocol of the "Seurat" software package (version 4.0.5), while gene expression levels were normalized using the LogNormalize method.

Subsequent analysis involved clustering cells and applying t-distributed Stochastic Neighbor Embedding (t-SNE) to identify cellular subpopulations. Our next studies focused on the expression of four anoikis-related genes (ARGs) in ccRCC immune cell subpopulations, while we used a "CellChat" to study intercellular communication between macrophages and dendritic cells.

2.8 Clinical ccRCC samples collection

Tissue specimens from ccRCC patients who underwent immunotherapy were acquired from the Department of Urology at the University of Hong Kong Shenzhen Hospital during the period from January 2022 to January 2024. Identifiable details

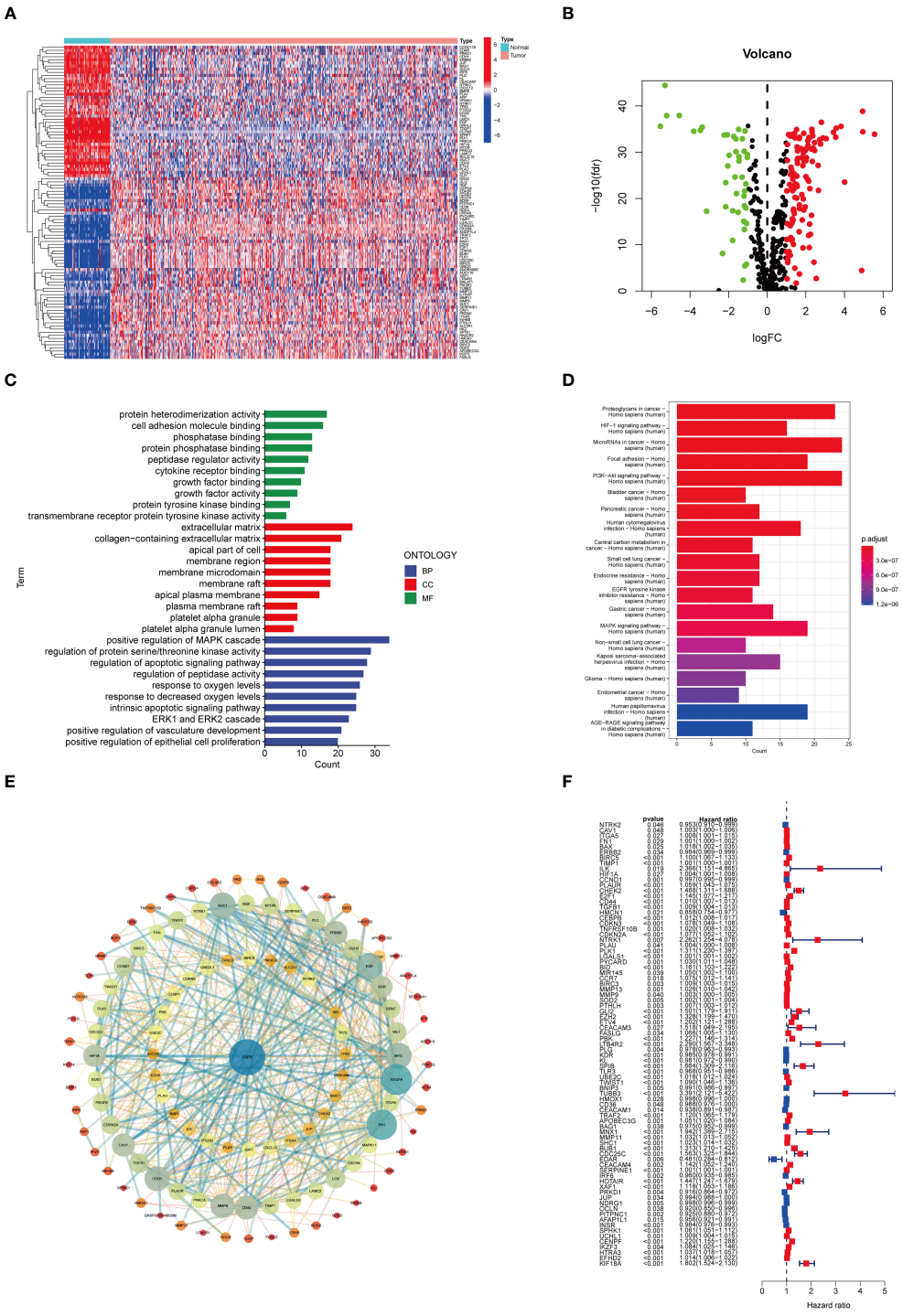


FIGURE 2
ARGs expression patterns in ccRCC. **(A)** Heat map of ARGs expression. **(B)** Volcano plot of ARGs differential genes. **(C)** Enrichment analysis. **(D)** KEGG analysis based on ARGs. **(E)** DEGs of ARGs. **(F)** Univariate Cox analysis on ARGs.

concerning the origins of these tissues were excluded, and the Ethics Committee of the hospital granted approval for this research.

2.9 cDNA production and PCR analysis

All cells were acquired from Procell Life Science & Technology Co., Ltd., and maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% Penicillin-Streptomycin. The culture conditions were set at 37°C with 5% CO2 (36, 37). Cells at a density of 5×10^5 were plated in six-well plates and incubated for 48 hours. Subsequently, cellular lysis was performed using TRIzol (Invitrogen). RNA was then extracted using a Total RNA Kit. A spectrophotometer was employed to assess the concentration and purity of the RNA. Following this, cDNA synthesis was conducted using an mRNA Reverse Transcription Kit (Roche). Finally, the quantification of target gene expression was achieved by employing a SYBR Green RNA Kit as per the manufacturer's instructions (38).

2.10 Immunohistochemistry

Immunohistochemistry (IHC) was used to validate differential expression levels of anoikis-related genes (ARGs) in ccRCC clinical samples. First, the ccRCC tissue sections were deparaffinized with

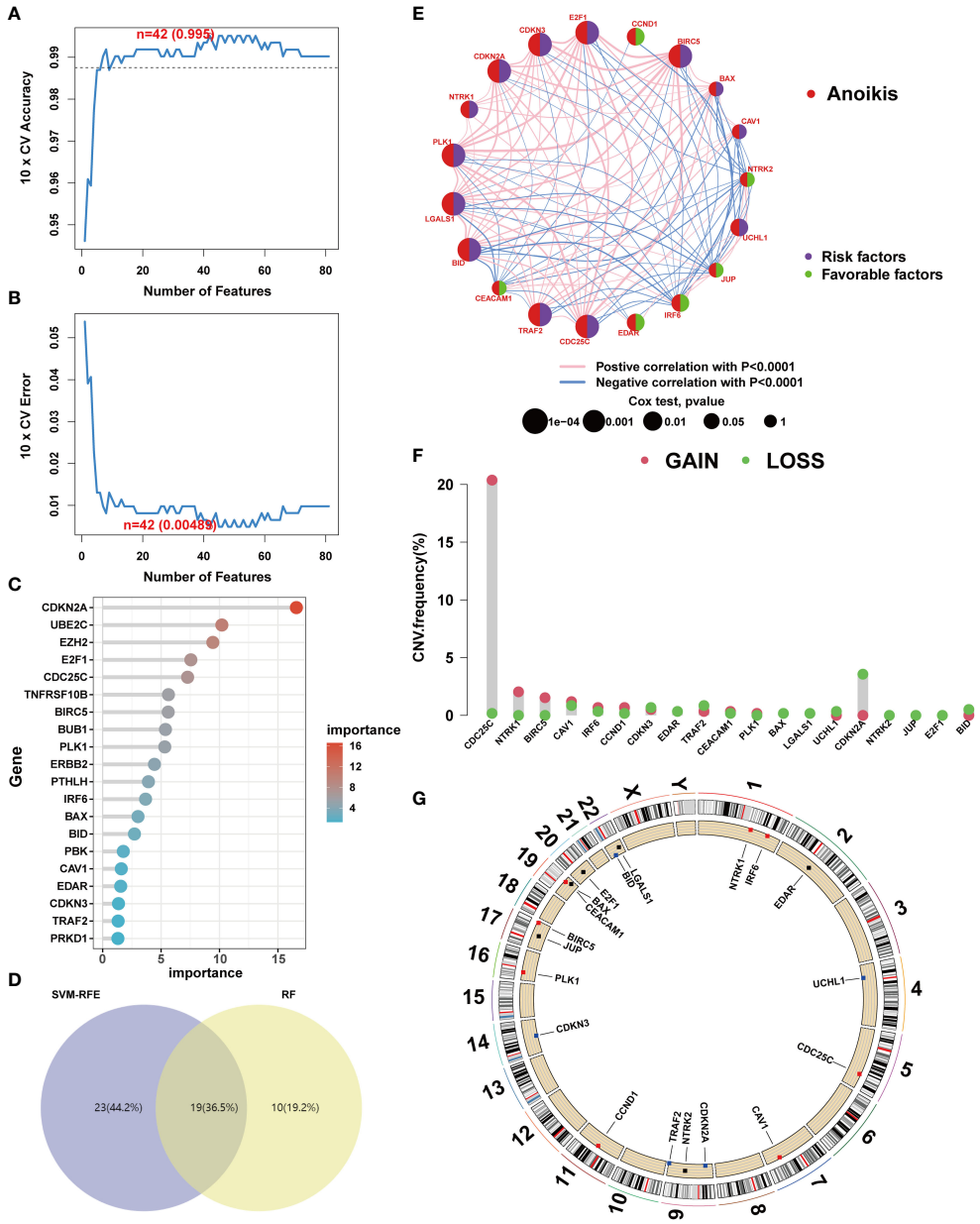


FIGURE 3
Core anoikis genes screening. **(A, B)** SVM–RFE screens the important anoikis ARGs. **(C)** Top 20 important ARGs. **(D)** Venn diagram. **(E)** ARGs interactions and correlations. **(F)** CNV frequency. **(G)** ARGs' location in chromosome.

xylene and washed stepwise with ethanol to rehydrate them. These sections were then treated with 3% hydrogen peroxide for 15 minutes to inhibit endogenous peroxidase activity, followed by antigen retrieval using 1 mM EDTA. Subsequently, the sections underwent overnight incubation at 4°C with antibodies against the model genes, diluted at a ratio of 1:200 (SANTA CRUZ). Following this, we applied PolyHRP Anti-Mouse/Rabbit IgG Detection System (Solarbio, China) and visualized the proteins using diaminobenzidine. Hematoxylin was used for counterstaining before the sections were dehydrated. The prepared slides were examined under a Zeiss microscope. For quantitative analysis, the staining intensity was measured and analyzed using ImageJ and GraphPad Prism version 7 software. Statistical significance was established at P-value < 0.05.

2.11 Statistical analysis

Statistical analyses in this study got executed utilizing the R software (release 4.0.2, https://www.r-project.org). Student's t-test was applied to calculate the DEGs. Comparisons of overall survival

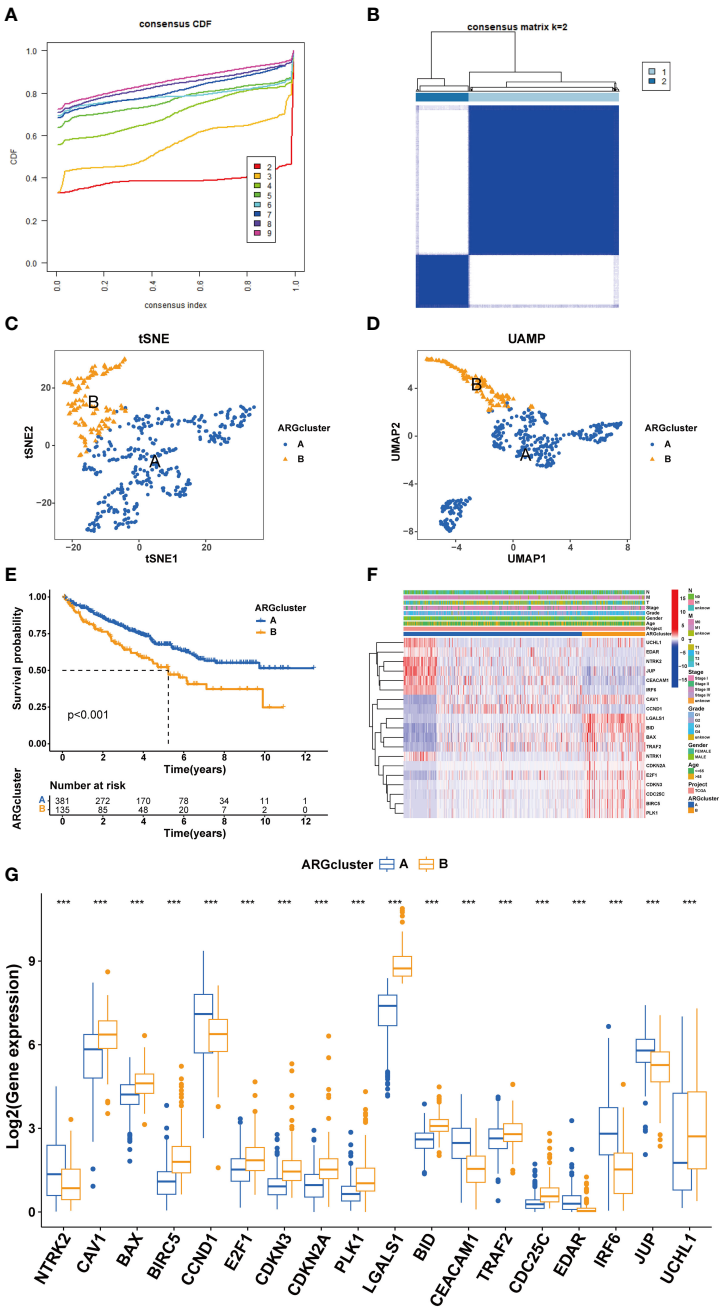


FIGURE 4
Cluster analysis for ccRCC patients. **(A, B)** cRCC patients were classified into two clusters based on ARGs profiles. **(C, D)** UAMP and tSNE analyses. **(E)** Survival analysis. **(F)** Heatmap based on ARGs and clinical characteristics of ccRCC patients. **(G)** ARGs expression between different clusters. (***P* < 0.001).

(OS) was performed using Kaplan–Meier analysis coupled with log-rank testing. $P < 0.05$ was established for statistical significance.

3 Results

3.1 Molecular insights into ARGs regulation and prognostic significance in ccRCC

The workflow of this study is depicted in Figure 1, with a comprehensive methodology described in the Methods section. Initially, transcriptome data and clinical information of 161 ARGs were extracted from the TCGA-KIRC cohort. Analysis revealed that 118 ARGs were upregulated, whereas 43 genes showed downregulation. Expression patterns of these differentially expressed genes are visualized in heatmaps and volcano plots (Figures 2A, B). Functional enrichment analysis indicated that the majority of ARGs are involved in processes such as the extracellular matrix, positive regulation of the MAPK cascade, regulation of the apoptotic signaling pathway, and the ERK1 and ERK2 cascade (Figure 2C), aligning with current insights into the mechanisms of anoikis resistance. Additionally, the HIF-1 signaling pathway and the PI3K-Akt signaling pathway were implicated in this context (Figure 2D). A protein interaction network further identified EGFR as a key upstream signaling molecule (Figure 2E). Finally, significant correlations between 81 ARGs and the prognosis of ccRCC were identified, with 23 ARGs serving as potential biomarkers for favorable prognosis (Figure 2F).

3.2 Anoikis genetic epigenetics and prognostic biomarkers in ccRCC

To elucidate the genetic underpinnings of ccRCC, a variety of machine learning techniques were employed for gene screening. Initially, the SVM-RFE algorithm was utilized to validate the screening of candidate genes following a 5-fold cross-validation process (Figures 3A, B). Subsequently, an ensemble RF algorithm pinpointed feature genes with a significance threshold exceeding 2, with CDKN2A exhibiting the highest importance (Figure 3C). An intersection of candidate genes identified by both SVM-RFE and RF algorithms highlighted 19 ARGs significantly impacting the prognosis of ccRCC patients (Figure 3D). In pursuit of understanding the relationships among these pivotal genes,

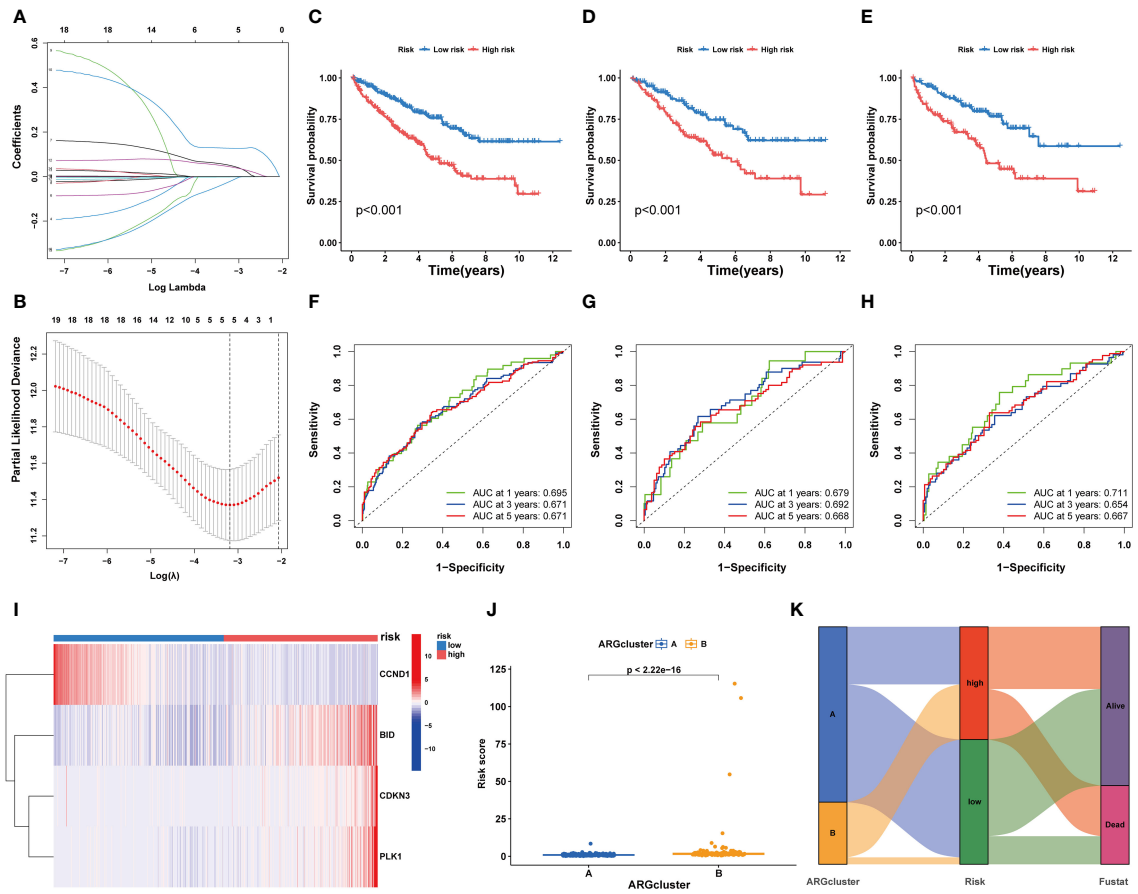


FIGURE 5
Development and validation of a risk prognostic model for ccRCC patients. **(A, B)** Lasso Cox regression analysis. **(C-E)** K-M curves of ccRCC patients under high and low risk. **(F-H)** ROC curves of ccRCC patients under high and low risk. **(I)** Heatmaps exhibit 4 core ARGs expression patterns. **(J)** The riskscore levels in two ARGclusters. **(K)** Alluvial plots.

correlation analyses were conducted. The results indicated that BID and PLK1 are markers of poor prognosis in ccRCC patients, and most ARGs demonstrated synergistic interactions (Figure 3E). The role of mutations in tumorigenesis was also investigated, particularly focusing on CNV mutation frequencies. Interestingly, a significant gain was only observed in CDC25C, suggesting that mutations might not be the primary mechanism influencing anoikis resistance (Figures 3F, G).

3.3 ARGs unsupervised cluster analysis

Clustering of ccRCC patients was performed based on the expression levels of 19 ARGs. When k equals 2, the patients were effectively stratified into two distinct groups (Figures 4A, B). Validation of clustering efficacy was provided by UMAP and tSNE scores (Figures 4C, D). Survival analysis indicated significant prognostic differences between these two subgroups of ccRCC (Figure 4E), with Group A exhibiting superior overall survival (OS) compared to Group B. An examination of clinical data revealed distinct ARG expression patterns and staging characteristics between the subgroups, where higher ARG expression was associated with advanced pathological stages (Figures 4F, G).

3.4 Development and validation of a risk prognostic model for ccRCC patients

Utilizing LASSO Cox regression and multivariate analysis, four core ARGs were identified from the 19 ARGs, and a prognostic model termed ARGs-Riskscore was established (Figures 5A, B). This model assigns a specific risk coefficient to each anoikis-related gene to calculate the riskscore, categorizing patients into high-risk and low-risk groups (Figure 5C). Kaplan-Meier curves demonstrated poorer survival outcomes for the high-risk group compared to the low-risk group (Figures 5D, E). ROC curve analysis revealed that the model's AUC value exceeded 0.6, indicating substantial accuracy (Figures 5F–H). Furthermore, in the high-risk group, the expression levels of CDKN3, PLK1, and BID were elevated, whereas CCND1 showed higher expression in the low-risk group (Figure 5I). A higher ARG expression level corresponded to an increased riskscore (Figures 5J, K).

Further, we integrated the risk associated with ARGs and clinical data from patients with ccRCC to develop a nomogram that estimates survival probabilities based on age, stage, and risk scores (Supplementary Figure S1A). Calibration plots confirmed the nomogram's accuracy in predicting 1-year, 3-year, and 5-year overall survival (OS) rates (Supplementary Figure S1B). Analysis

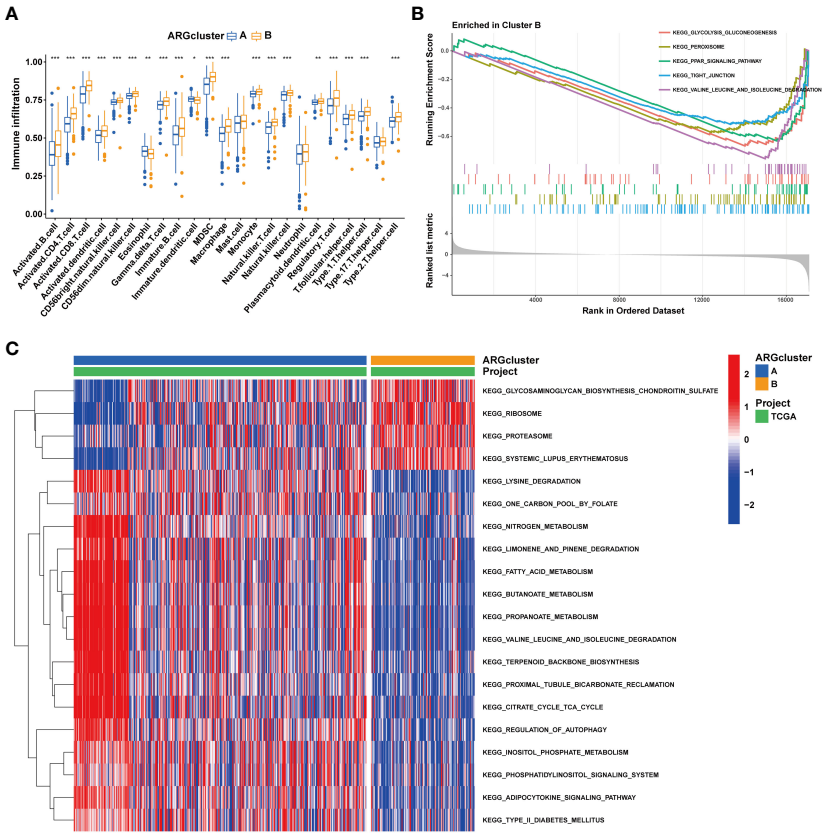


FIGURE 6
Immune landscape of two ARGcluster. **(A)** Immune cell infiltration levels at two ARGcluster patients. **(B)** GSEA analysis of the enrichment of ARGs. **(C)** GSVA enrichment analysis. (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$).

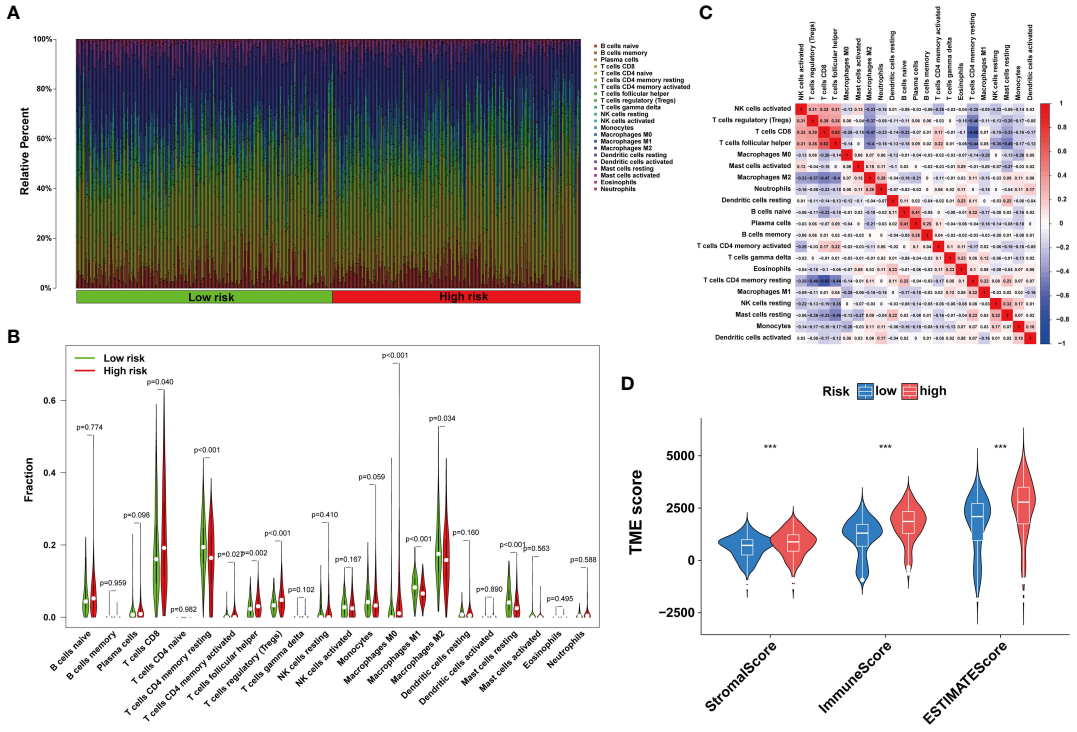


FIGURE 7
Immune infiltration in under different risk ccRCC patients **(A, B)** Immune cells infiltration analysis. **(C)** Correlation of 23 immune cells. **(D)** TME score. (Wilcox test, ***$P < 0.001$).

of hazard ratios indicated a strong correlation between age, cancer grade, risk scores, and tumor stage with OS (Supplementary Figure S1C). Furthermore, analysis of cumulative hazards revealed that patients with ccRCC who had higher nomorisk scores exhibited increased mortality risks (Supplementary Figure S1D).

3.5 Anoikis affects tumor immune microenvironment

The tumor immune microenvironment plays a pivotal role in the immune evasion processes of cancer. The onset of anoikis resistance is predicated on achieving immune escape. Consequently, we divided ccRCC patients into two subgroups based on the expression patterns of anoikis genes. The findings revealed that patients with high expression of anoikis genes exhibited significantly higher levels of immune infiltration, particularly with MDSC cells (Figure 6A), underscoring a close association between the anoikis process and the immunosuppressive microenvironment. Enrichment analyses indicated that, compared to cluster A, the tight junction and PPAR signaling pathways were significantly enriched in cluster B (Figures 6B, C), suggesting their major roles in shaping the immunosuppressive microenvironment. Patients with ccRCC of different risk levels demonstrated markedly distinct survival outcomes. Therefore, we further examined the levels of immune cell infiltration in patients classified into high-risk and low-risk groups. It was observed that the proportions of immune cell infiltration varied between different risk groups of ccRCC patients (Figure 7A). Notably, Tregs and macrophages M0 were significantly more prevalent in patients at higher risk (Figure 7B). Interestingly, there was a significant negative correlation between the infiltration of macrophages M0 and CD8 T cells, and between Tregs and memory CD4 T cells (Figure 7C), highlighting the crucial impact of the immunosuppressive microenvironment on anoikis resistance (Figure 7D).

Furthermore, as risk scores increased, changes were noted in the patterns of immune cell infiltration (Figure 8J). For example, the infiltration levels of macrophages M0 and Tregs gradually increased with rising risk scores (Figures 8A, F, H, I). In contrast, other cell types, such as macrophages M1 and NK cells, showed a decrease in infiltration as risk scores increased (Figures 8B–E, G).

3.6 Single-cell analysis reveals anoikis expression pattern in ccRCC

The advent of single-cell technologies has provided a crucial avenue for exploring cellular subtypes. Utilizing single-cell analysis and annotation, we categorized cell suspensions from patients treated with anti-PD-L1 into 24 immune cell subtypes and nine principal cell types (Figures 9A, C). Importantly, the proportions of

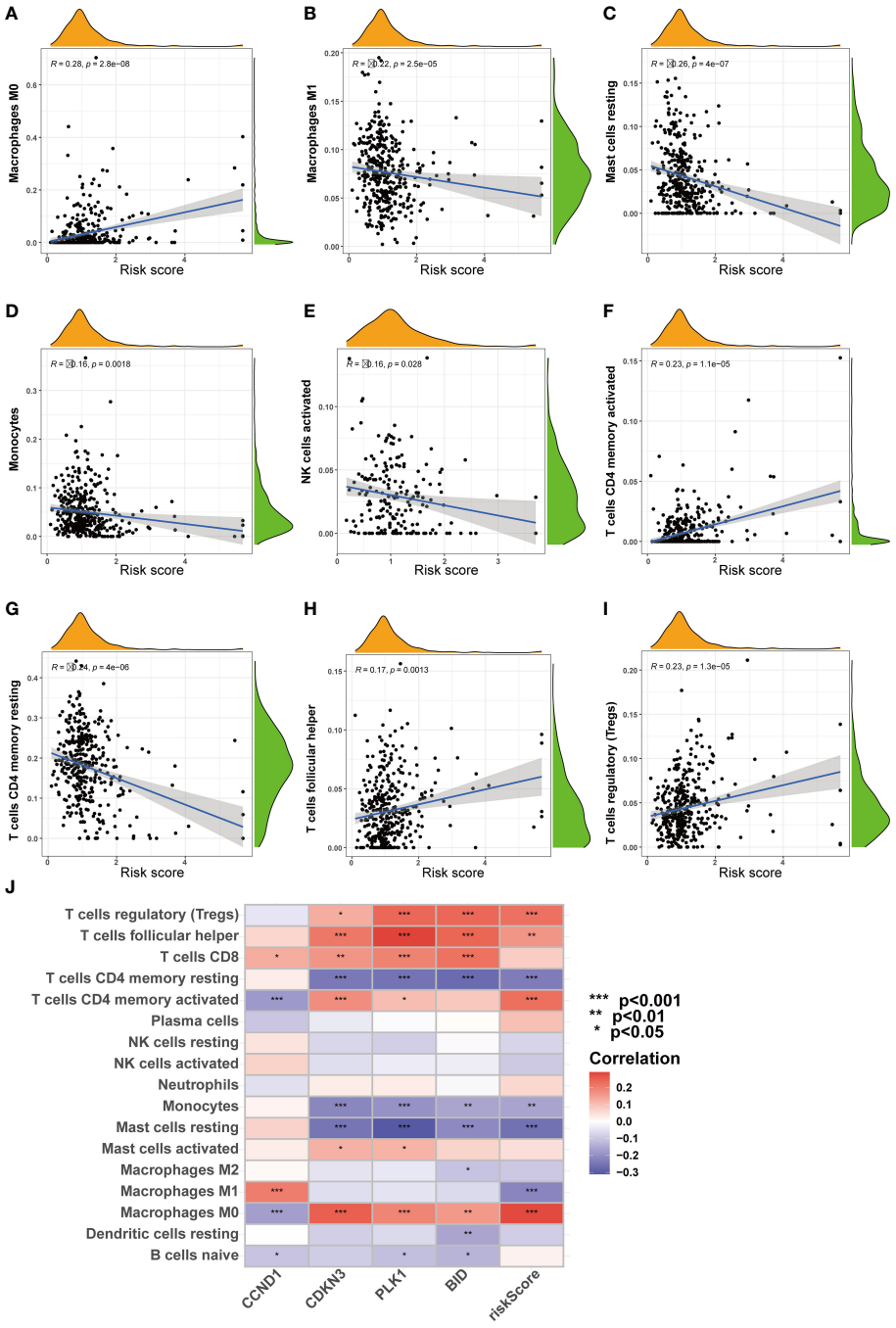


FIGURE 8
Correlation analysis of riskscore and immune cells. **(A-I)** Correlation of riskscore and immune cells. **(J)** Correlation of four ARGs and immune cells. (*P < 0.05; **P < 0.01; ***P < 0.001).

cellular components from samples of different patients demonstrated notable disparities. For instance, the Tumor 1 sample predominantly consisted of mono/macrophages, whereas CD8 T cells predominated in Blood4 (Figure 9D). We further elucidated the expression and distribution of four core ARGs constituting a prognostic model across various cell subtypes (Figure 9E). Our findings reveal that BID exhibits the highest expression in DC and T proliferation cells, with subsequent analyses revealing enhanced communication between DC, proliferative T cells, and other cell subtypes (Figures 9B, F).

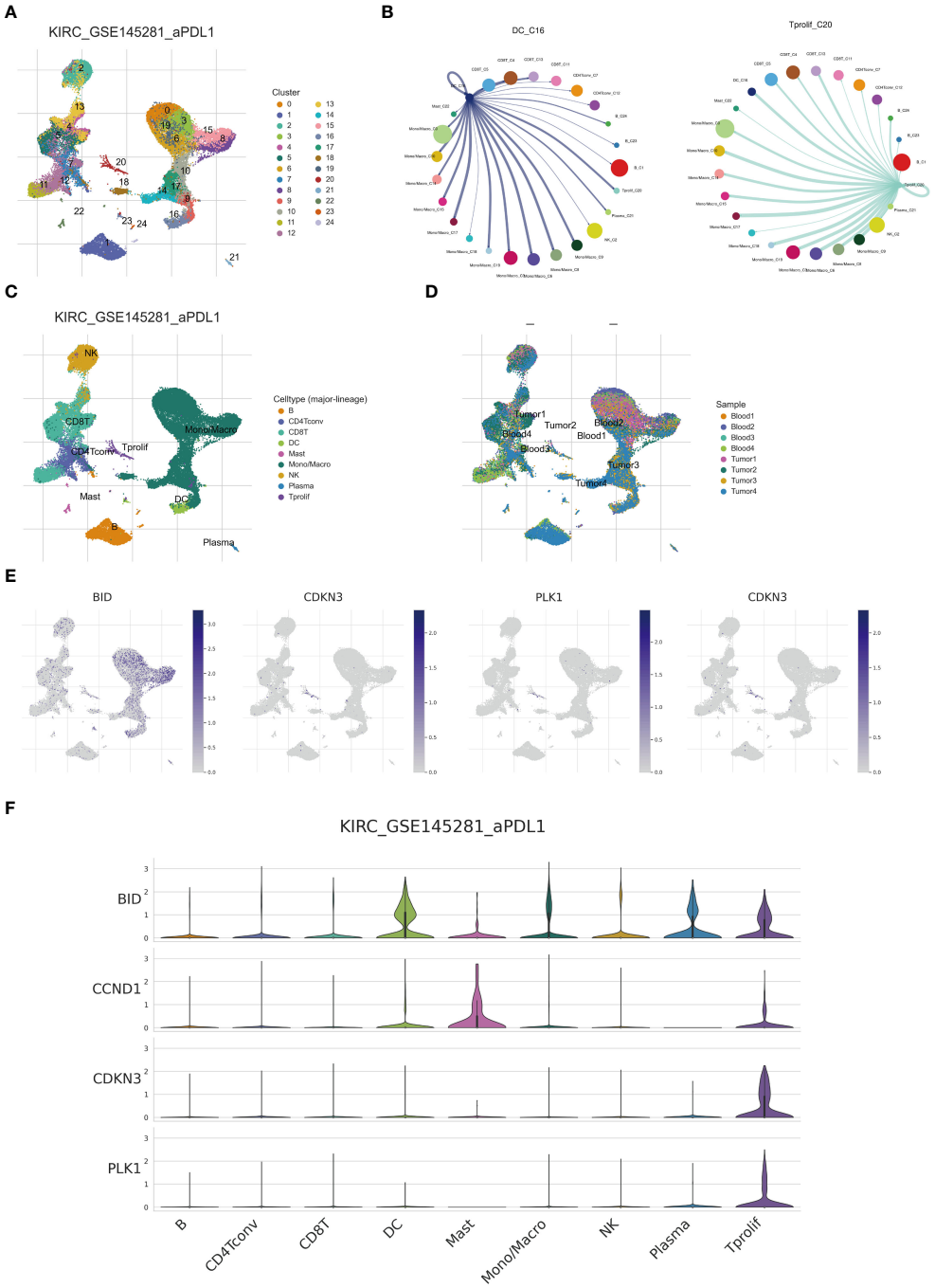


FIGURE 9
Single cell analysis reveals anoikis genes expression. **(A, C)** Umap of single cell clusters. **(B)** Cell communications of DCs and T proliferation cells. **(D)** Cell types distribution in each ccRCC sample. **(E, F)** 4 core ARGs expression patterns in immune cells.

3.7 Evaluation of anoikis gene expression by realtime PCR and IHC

To investigate the expression of anoikis genes, primary renal carcinoma cell lines and normal renal cell lines were cultured, and the expression levels of four central anoikis resistance genes (ARGs) were compared between these two cell types. The primary cell lines were assessed to ensure the reliability of the results (Figure 10A). Furthermore, upon culturing the cells up to the tenth passage, mRNA levels were re-evaluated (Figure 10B). Intriguingly, despite the expressions of BID, CDKN3, and PLK1 being consistently higher in the renal carcinoma cell lines than in the normal renal

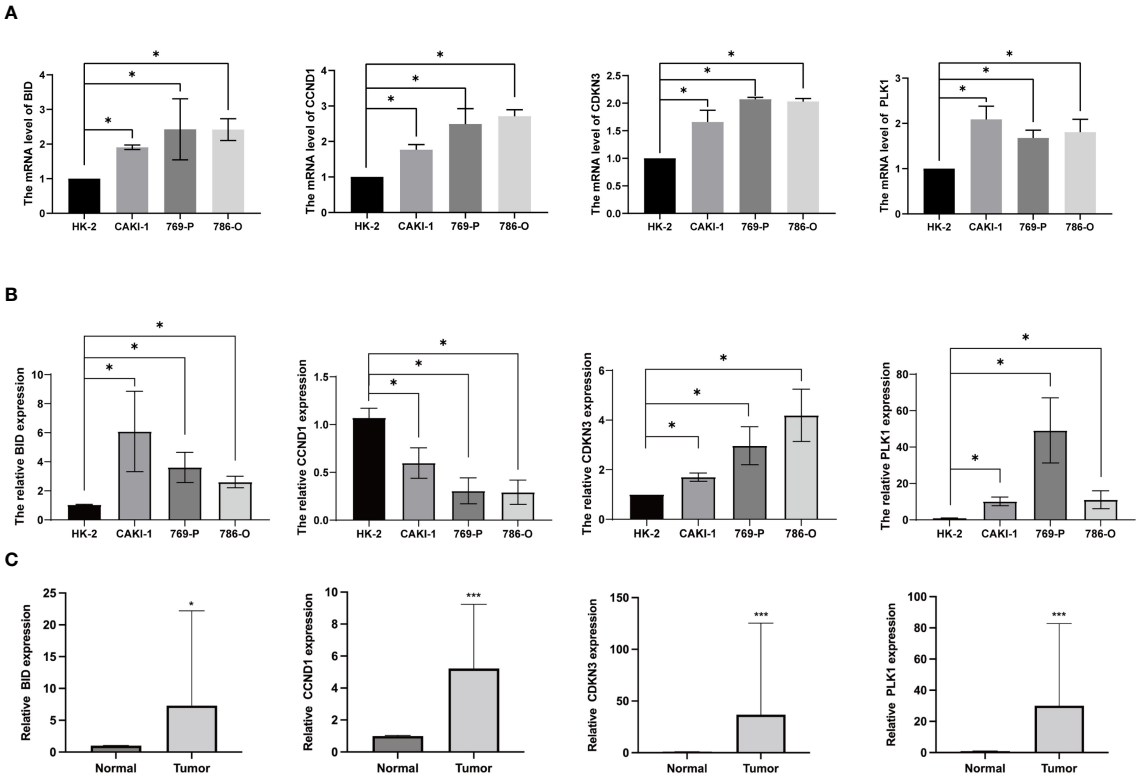


FIGURE 10
Anoikis genes' expression levels in ccRCC cells and tissues. **(A)** ARGs mRNA expression levels in primary ccRCC cells. **(B)** ARGs mRNA expression levels in ccRCC cells after culturing for 5 weeks **(C)** The mRNA levels of ARGs in ccRCC and normal kidney tissues. (*$P < 0.05$; ***$P < 0.001$).

cells across both the primary and tenth passages, shifts in the gene expression levels among the carcinoma lines were noted. Notably, the expression pattern of CCND1 demonstrated an inverse trend. Subsequent analyses involved examining the expression levels of these four core ARGs in cDNA extracted from normal renal tissues and renal carcinoma tissues. The results indicated a higher expression of all four ARGs in the carcinoma tissues (Figure 10C). Immunohistochemistry confirmed that the protein levels of ARGs corresponded with the trends observed at the gene expression level (Figure 11).

3.8 Potency of the anoikis signature in modulating drug resistance

In order to elucidate the association between the anoikis-related signature and drug responsiveness, the IC50 indices for various medications in ccRCC were evaluated (Supplementary Figure S2). This analysis implies that individuals with renal carcinoma who are categorized within the high-risk group could exhibit resistance to both chemotherapy and immunotherapeutic approaches. Conversely, this suggests opportunities for modulating drug efficacy through targeted interventions.

4 Discussion

As the complexity and diversity of clear cell renal cell carcinoma (ccRCC) become increasingly apparent (39), numerous therapeutic strategies have been introduced into clinical settings to address this condition. Cellular molecular-targeted therapy is the most effective method of treating metastatic ccRCC as patients suffering from kidney cancer do not respond to radiotherapy and chemotherapy. The European Urology Association (EUA) and the United States National Comprehensive Cancer Network (NCCN) recommended the molecular-targeted drugs as the first and second-line medicine for metastatic ccRCC. At present, there are no universally accepted and reliable predictors for the diagnosis and prognosis of ccRCC. The challenge of accurately predicting outcomes persists, highlighting the critical need for the discovery of new biomarkers. These biomarkers are crucial for enhancing the prognosis of ccRCC (40). The exploration of abnormally expressed genes in ccRCC tissues can potentially help identify new molecular biomarkers for the diagnosis and prognosis of ccRCC. Central to this endeavor is anoikis, a cellular process essential for controlling tumor proliferation, spread, and future outcomes (41, 42). Research has linked the development, advancement, and prognosis of ccRCC to specific genes involved in anoikis (43). The goal of this discussion is

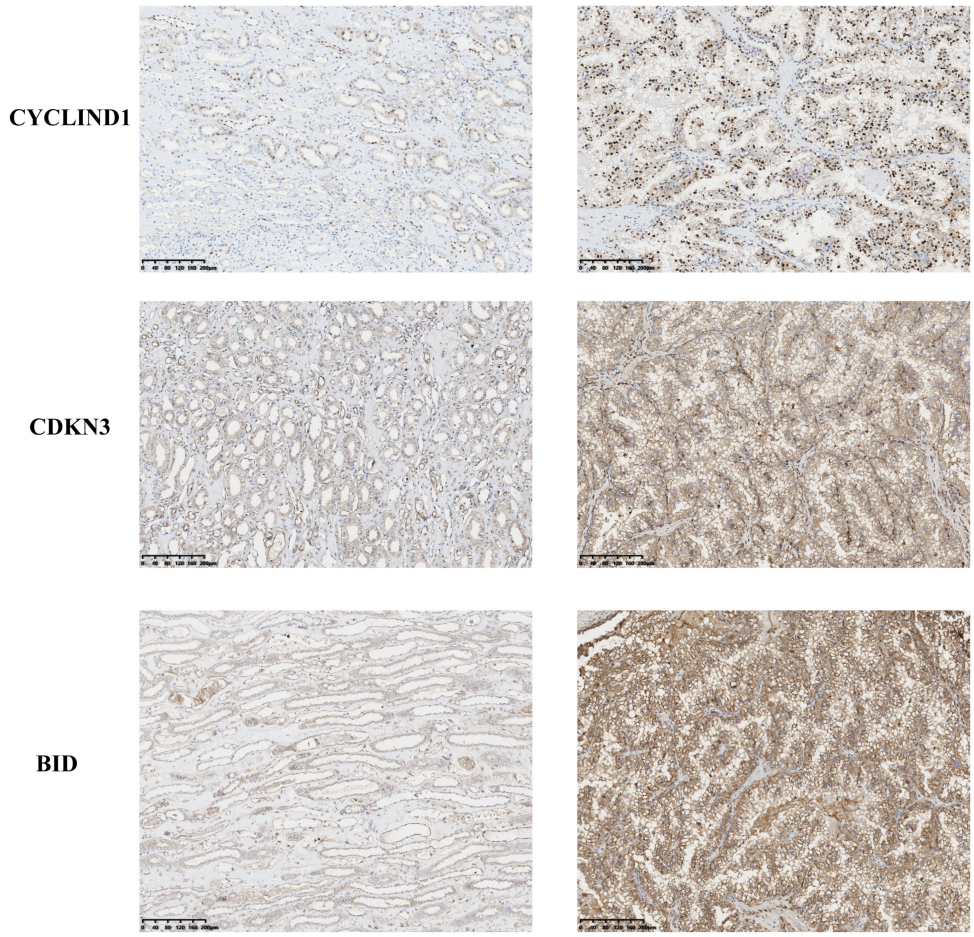


FIGURE 11
Immunohistochemical results of ARGs proteins.

to integrate findings on the connection between genes related to anoikis and the prognosis of ccRCC, examining their potential relevance in clinical practice and providing a novel theoretical and practical framework for tailored therapeutic approaches.

In this study, we identified critical roles of anoikis-associated genes in ccRCC and developed a predictive model. Herein, we described the differential expression of anoikis-associated genes in tumor tissues relative to normal samples and investigate the potential regulatory role of anoikis-associated genes in controlling the ccRCC immune microenvironment. In addition, we investigated the relationship between anoikis-associated genes expression levels and immunotherapy. An analysis of 161 anoikis-associated genes revealed four that were conclusively linked to the prognosis of ccRCC. Our study confirmed that BID, CCND1, CDKN3, and PLK1 showed upregulation in ccrCC tissues, with significantly higher expression compared to normal cells. It is reasonable to speculated that these four genes play critical roles in ccRCC tumorigenesis and progression. This study aimed to gain insights into the underlying mechanisms associated with the anoikis-associated gene that was associated with immune-related factors. BID, a pro-apoptotic protein in the Bcl-2 family, functions collaboratively with BAX to facilitate cellular apoptosis (44). Research by Ji Miao and colleagues showed that the expression levels of Bid correlate with the susceptibility of liver cancer cells to chemotherapeutic agents (45). Another key protein, Cyclin D1 (CCND1), a crucial component of the D-type cyclin group, regulates the progression of the cell cycle (46). Recent studies suggested the USP10/CCND1 pathway as a potential therapeutic target for glioblastoma (GBM) patients (47). Investigations by Hongying Zhang and team found that CCND1 suppression, achieved by gene silencing, impedes the differentiation of hepatic cancer stem cells by inhibiting autophagy (48). Furthermore, the expression patterns of CCND1 are strongly correlated with the initiation and progression of multiple cancer types (49–51).

CDKN3, a cyclin-dependent kinase inhibitor, is identified as a crucial therapeutic target for cervical cancer (52) and plays a role in the malignant advancement of pancreatic cancer by interacting with PSMD12 (53). Its expression in various cancers modulates

resistance to treatment. Aolin Li and colleagues demonstrated that ZNF677 represses the malignant evolution of renal cell carcinoma through the reduction of CDKN3 expression (54). Furthermore, the circular RNA circSDHC binds to miR-127–3p competitively, thereby diminishing CDKN3 expression in renal cell carcinoma and curbing its malignant advancement (55). These observations corroborate our analysis, thereby confirming the precision of our findings.

Polo-like kinase 1 (PLK1), an eminent serine/threonine kinase within the protein kinase superfamily, promotes the advancement of mitosis (56). Elevated levels of PLK1 are commonly observed in cancerous tissues, highlighting its potential as a target for therapeutic intervention (57). Suppression of PLK1 enhances the response of pancreatic cancer cells to immunotherapeutic strategies (58). Likewise, a reduction in PLK1 activity increases the sensitivity of breast cancer to radiation therapy (59), whereas enhanced expression of PLK1 contributes to the development and advancement of liver tumors (60).

Infiltration of Treg cells is frequently associated with poorer prognoses across various cancers, and a reduction in Treg cells has been observed to initiate and enhance antitumor immune responses. In this study, the risk score exhibited a significant positive correlation with the level of Treg cell infiltration, whereas an inverse trend was noted for NK cells, suggesting a critical role for the immunosuppressive microenvironment in anoikis resistance, subsequently impacting the overall survival of patients with ccRCC. The signaling cascade mediated by PPAR, documented to enhance angiogenesis within tumor matrices (61), is associated with the pathogenesis of both inflammatory and neoplastic conditions (62). Furthermore, this pathway has been shown to trigger anoikis in certain cell types under *in vitro* conditions (63). Our findings suggest that the PPAR pathway may serve as a primary mechanism by which Tregs orchestrate an immunosuppressive microenvironment, thereby facilitating anoikis resistance, which in turn supports distant metastasis and immune evasion in ccRCC. Conventional surgical treatment and radiotherapy and chemotherapy cannot be effective to treat patients suffering from late-stage ccRCC. Maybe more research should be conducted on the gene targets and immune checkpoint inhibitors associated with ccRCC as the results can potentially help predict the prognosis of antitumor immunotherapy. It is worth noting that the results of our research reflected the association of anoikis-associated genes with a substantial prognosis of ccRCC and confirmed the reliability of the analytical results obtained. We may infer that the modulation of the Anoikis-associated genes activity associated with ccRCC could potentially help obtain results that can help improve the therapeutic techniques.

As we know, there is a minor number of relevant researches currently available to explain the functions of anoikis in ccRCC. Our work provided valuable information on how the anoikis-associated gene participated in cancer immunotherapy, which may potentially help improve the processes of ccRCC targeting therapy. In our next step, we need to extend the existing database and mutually authenticate to larger database. Experiments should be performed at the molecular, cytological, and animal levels to investigate the relationship between the prognosis of the patients and the properties of the clinical tumor tissue samples. We believe that our results can potentially help for improving the efficiency of diagnosis, treatment methods, and survival prognosis of ccRCC patients.

5 Conclusion

In this study, we established the association between anoikis, immune cell infiltration, and the prognosis of clear cell renal cell carcinoma (ccRCC) patients through multi-omics and single-cell analyses. Furthermore, we elucidated their impact on the efficacy of immune therapy. These findings not only provide novel insights into the role of apoptosis in cancer progression but also highlight new research directions for immunotherapeutic strategies in ccRCC.

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 University of Hong Kong Shenzhen Hospital. 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.

Author contributions

XW: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. JH: Formal analysis, Methodology, Software, Visualization, Writing – original draft. TQ: Methodology, Software, Visualization, Writing – original draft. XC: Data curation, Resources, Visualization, Writing – original draft. GL: Formal analysis, Investigation, Writing – original draft. SF: Formal analysis, Investigation, Writing – original draft. SX: Formal analysis, Resources, Writing – original draft. LQ: Formal analysis, Resources, Writing – original draft. WW: Conceptualization, Funding acquisition, Investigation, 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. This manuscript was supported by a grant from the Guangdong Medical Science and Technology Research (B2023384).

Acknowledgments

We gratefully acknowledge the contributions from the TCGA and GEO project, which provided valuable data and resources for 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.

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.2024.1427475/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Nomogram establishment based on risk and clinical information. **(A)** A monogram graph to predict ccRCC patients' prognosis. **(B)** Calibration plots. **(C)** Multivariate analysis. **(D)** Cumulative risk values analysis. (*$P < 0.05$; ***$P < 0.001$).

SUPPLEMENTARY FIGURE 2
Drug sensitivity analysis.

References

1. Humphrey PA, Moch H, Cubilla AL, Ulbright TM, Reuter VE. The 2016 WHO classification of tumours of the urinary system and male genital organs-part B: prostate and bladder tumours. *Eur Urol.* (2016) 70:106–19. doi: 10.1016/j.eururo.2016.02.028

2. Choueiri TK, Motzer RJ. Systemic therapy for metastatic renal-cell carcinoma. *N Engl J Med.* (2017) 376:354–66. doi: 10.1056/NEJMra1601333

3. Wang J, Zuo Z, Yu Z, Chen Z, Meng X, Ma Z, et al. et al: Single-cell transcriptome analysis revealing the intratumoral heterogeneity of ccRCC and validation of MT2A in pathogenesis. *Funct Integr Genomics.* (2023) 23:300. doi: 10.1007/s10142-023-01225-7

4. Cancer Genome Atlas Research N, Linehan WM, Spellman PT, Ricketts CJ, Creighton CJ, Fei SS, et al. et al: comprehensive molecular characterization of papillary renal-cell carcinoma. *N Engl J Med.* (2016) 374:135–45. doi: 10.1056/NEJMoa1505917

5. Jin L, Chun J, Pan C, Kumar A, Zhang G, Ha Y, et al. et al: The PLAG1-GDH1 Axis Promotes Anoikis Resistance and Tumor Metastasis through CamKK2-AMPK Signaling in LKB1-Deficient Lung Cancer. *Mol Cell.* (2018) 69:87–99 e87. doi: 10.1016/j.molcel.2017.11.025

6. Buchheit CL, Angarola BL, Steiner A, Weigel KJ, Schafer ZT. Anoikis evasion in inflammatory breast cancer cells is mediated by Bim-EL sequestration. *Cell Death Differ.* (2015) 22:1275–86. doi: 10.1038/cdd.2014.209

7. Chen J, Gao F, Liu N. L1CAM promotes epithelial to mesenchymal transition and formation of cancer initiating cells in human endometrial cancer. *Exp Ther Med.* (2018) 15:2792–7. doi: 10.3892/etm

8. She K, Yang W, Li M, Xiong W, Zhou M. FAIM2 promotes non-small cell lung cancer cell growth and bone metastasis by activating the wnt/beta-catenin pathway. *Front Oncol.* (2021) 11:690142. doi: 10.3389/fonc.2021.690142

9. Manogaran P, Beeraka NM, Paulraj RS, Sathiyachandran P, Thammaiappa M. Impediment of cancer by dietary plant-derived alkaloids through oxidative stress: implications of PI3K/AKT pathway in apoptosis, autophagy, and ferroptosis. *Curr Top Med Chem.* (2023) 23:860–77. doi: 10.2174/1568026623666230111154537

10. Sugiura R, Satoh R, Takasaki T. ERK: A double-edged sword in cancer. ERK-dependent apoptosis as a potential therapeutic strategy for cancer. *Cells.* (2021) 10 (10):2509. doi: 10.3390/cells10102509

11. Liu X, Fu Y, Zhang G, Zhang D, Liang N, Li F, et al. et al: miR-424-5p Promotes Anoikis Resistance and Lung Metastasis by Inactivating Hippo Signaling in Thyroid Cancer. *Mol Ther Oncolytics.* (2019) 15:248–60. doi: 10.1016/j.omto.2019.10.008

12. Zhang H, Fredericks T, Xiong G, Qi Y, Rychahou PG, Li JD, et al. Membrane associated collagen XIII promotes cancer metastasis and enhances anoikis resistance. *Breast Cancer Res.* (2018) 20:116. doi: 10.1186/s13058-018-1030-y

13. Xia Z, Chen S, He M, Li B, Deng Y, Yi L, et al. Editorial: Targeting metabolism to activate T cells and enhance the efficacy of checkpoint blockade immunotherapy in solid tumors. *Front Immunol.* (2023) 14:1247178. doi: 10.3389/fimmu.2023.1247178

14. Gong X, Chi H, Xia Z, Yang G, Tian G. Advances in HPV-associated tumor management: Therapeutic strategies and emerging insights. *J Med Virol.* (2023) 95: e28950. doi: 10.1002/jmv.28950

15. Planells-Ferrer L, Urresti J, Soriano A, Reix S, Murphy DM, Ferreres JC, et al. et al: MYCN repression of Lifeguard/FAIM2 enhances neuroblastoma aggressiveness. *Cell Death Dis.* (2014) 5:e1401. doi: 10.1038/cddis.2014.356

16. Grage-Griebenow E, Jerg E, Gorys A, Wicklein D, Wesch D, Freitag-Wolf S, et al. et al: L1CAM promotes enrichment of immunosuppressive T cells in human pancreatic cancer correlating with Malignant progression. *Mol Oncol.* (2014) 8:982–97. doi: 10.1016/j.molonc.2014.03.001

17. Zhang X, Zhang P, Cong A, Feng Y, Chi H, Xia Z, et al. Unraveling molecular networks in thymic epithelial tumors: deciphering the unique signatures. *Front Immunol.* (2023) 14:1264325. doi: 10.3389/fimmu.2023.1264325

18. Lee YG, Yang N, Chun I, Porazzi P, Carturan A, Paruzzo L, et al. Apoptosis: a Janus bifrons in T-cell immunotherapy. *J Immunother Cancer.* (2023) 11(4):e005967. doi: 10.1136/jitc-2022-005967

19. Chaojun L, Pengping L, Yanjun L, Fangyuan Z, Yaning H, Yingbo S, et al. TJP3 promotes T cell immunity escape and chemoresistance in breast cancer: a comprehensive analysis of anoikis-based prognosis prediction and drug sensitivity stratification. *Aging (Albany NY).* (2023) 15:12890–906. doi: 10.18632/aging.v15i22

20. Li C, Liu T, Liu Y, Zhang J, Zuo D. Prognostic value of tumour microenvironment-related genes by TCGA database in rectal cancer. *J Cell Mol Med.* (2021) 25:5811–22. doi: 10.1111/jcmm.16547

21. Yuan Q, Lu X, Guo H, Sun J, Yang M, Liu Q, et al. Low-density lipoprotein receptor promotes crosstalk between cell stemness and tumor immune microenvironment in breast cancer: a large data-based multi-omics study. *J Transl Med.* (2023) 21:871. doi: 10.1186/s12967-023-04699-y

22. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. et al: STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* (2015) 43:D447–452. doi: 10.1093/nar/gku1003

23. Zhang J, Xiao J, Wang Y, Zheng X, Cui J, Wang C. A universal co-expression gene network and prognostic model for hepatic-biliary-pancreatic cancers identified by integrative analyses. *FEBS Open Bio.* (2022) 12:2006–24. doi: 10.1002/2211-5463.13478

24. 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

25. 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

26. Song G, Peng G, Zhang J, Song B, Yang J, Xie X, et al. Uncovering the potential role of oxidative stress in the development of periodontitis and establishing a stable diagnostic model via combining single-cell and machine learning analysis. *Front Immunol.* (2023) 14:1181467. doi: 10.3389/fimmu.2023.1181467

27. Wu Z, Zeng J, Wu M, Liang Q, Li B, Hou G, et al. Identification and validation of the pyroptosis-related long noncoding rna signature to predict the prognosis of patients with bladder cancer. *Med (Baltimore).* (2023) 102:e33075. doi: 10.1097/MD.0000000000033075

28. Zhang B, Huang B, Zhang X, Li S, Zhu J, Chen X, et al. PANoptosis-related molecular subtype and prognostic model associated with the immune microenvironment and individualized therapy in pancreatic cancer. *Front Oncol.* (2023) 13:1217654. doi: 10.3389/fonc.2023.1217654

29. Zhang S, Jiang C, Jiang L, Chen H, Huang J, Gao X, et al. et al: Construction of a diagnostic model for hepatitis B-related hepatocellular carcinoma using machine learning and artificial neural networks and revealing the correlation by immunoassay. *Tumour Virus Res*. (2023) 16:200271. doi: 10.1016/j.tvr.2023.200271

30. Liu Y, Wu Z, Feng Y, Gao J, Wang B, Lian C, et al. Integration analysis of single-cell and spatial transcriptomics reveal the cellular heterogeneity landscape in glioblastoma and establish a polygenic risk model. *Front Oncol*. (2023) 13:1109037. doi: 10.3389/fonc.2023.1109037

31. Li C, Wirth U, Schardey J, Ehrlich-Treuenstatt VV, Bazhin AV, Werner J, et al. An immune-related gene prognostic index for predicting prognosis in patients with colorectal cancer. *Front Immunol*. (2023) 14:1156488. doi: 10.3389/fimmu.2023.1156488

32. Zhang B, Sun J, Guan H, Guo H, Huang B, Chen X, et al. Integrated single-cell and bulk RNA sequencing revealed the molecular characteristics and prognostic roles of neutrophils in pancreatic cancer. *Aging (Albany NY)*. (2023) 15:9718–42. doi: 10.18632/aging.v15i18

33. Sun Z, Wang J, Zhang Q, Meng X, Ma Z, Niu J, et al. et al: Coordinating single-cell and bulk RNA-seq in deciphering the intratumoral immune landscape and prognostic stratification of prostate cancer patients. *Environ Toxicol*. (2024) 39:657–68. doi: 10.1002/tox.23928

34. Guan H, Chen X, Liu J, Sun J, Guo H, Jiang Y, et al. Molecular characteristics and therapeutic implications of Toll-like receptor signaling pathway in melanoma. *Sci Rep*. (2023) 13:13788. doi: 10.1038/s41598-023-38850-y

35. Liu T, Li C, Zhang J, Hu H, Li C. Unveiling efferocytosis-related signatures through the integration of single-cell analysis and machine learning: a predictive framework for prognosis and immunotherapy response in hepatocellular carcinoma. *Front Immunol*. (2023) 14:1237350. doi: 10.3389/fimmu.2023.1237350

36. Zhang B, Liu J, Li H, Huang B, Zhang B, Song B, et al. Integrated multi-omics identified the novel intratumor microbiome-derived subtypes and signature to predict the outcome, tumor microenvironment heterogeneity, and immunotherapy response for pancreatic cancer patients. *Front Pharmacol*. (2023) 14:1244752. doi: 10.3389/fphar.2023.1244752

37. Zhang H, Ni M, Wang H, Zhang J, Jin D, Busuttil RW, et al. Gsk3beta regulates the resolution of liver ischemia/reperfusion injury via MerTK. *JCI Insight*. (2023) 8(1): e151819. doi: 10.1172/jci.insight.151819

38. Ling B, Wei P, Xiao J, Cen B, Wei H, Feng X, et al. et al: Nucleolar and spindle –associated protein 1 promotes non–small cell lung cancer progression and serves as an effector of myocyte enhancer factor 2D. *Oncol Rep*. (2021) 45:1044–58. doi: 10.3892/or.2020.7918

39. Li X, Guan H, Ma C, Dai Y, Su J, Chen X, et al. Combination of bulk RNA sequencing and scRNA sequencing uncover the molecular characteristics of MAPK signaling in kidney renal clear cell carcinoma. *Aging (Albany NY)*. (2024) 16:1414–39. doi: 10.18632/aging.v16i2

40. Motzer RJ, Jonasch E, Boyle S, Carlo MI, Manley B, Agarwal N, et al. et al: NCCN guidelines insights: kidney cancer, version 1.2021. *J Natl Compr Canc Netw*. (2020) 18:1160–70. doi: 10.6004/jnccn.2020.0043

41. Paoli P, Giannoni E, Chiarugi P. Anoikis molecular pathways and its role in cancer progression. *Biochim Biophys Acta*. (2013) 1833:3481–98. doi: 10.1016/j.bbamcr.2013.06.026

42. Raeisi M, Zehtabi M, Velaei K, Fayyazpour P, Aghaei N, Mehdizadeh A. Anoikis in cancer: The role of lipid signaling. *Cell Biol Int*. (2022) 46:1717–28. doi: 10.1002/cbin.11896

43. Wang J, Qi X, Wang Q, Wu G. The role and therapeutic significance of the anoikis pathway in renal clear cell carcinoma. *Front Oncol*. (2022) 12:1009984. doi: 10.3389/fonc.2022.1009984

44. Billen LP, Shamas-Din A, Andrews DW. Bid: a bax-like BH3 protein. *Oncogene*. (2008) 27 Suppl 1:S93–104. doi: 10.1038/onc.2009.47

45. Miao J, Chen GG, Chun SY, Chak EC, Lai PB. Bid sensitizes apoptosis induced by chemotherapeutic drugs in hepatocellular carcinoma. *Int J Oncol*. (2004) 25:651–9. doi: 10.3892/ijo

46. Diakos CI, Charles KA, McMillan DC, Clarke SJ. Cancer-related inflammation and treatment effectiveness. *Lancet Oncol*. (2014) 15:e493–503. doi: 10.1016/S1470-2045(14)70263-3

47. Sun T, Xu YJ, Jiang SY, Xu Z, Cao BY, Sethi G, et al. Suppression of the USP10/CCND1 axis induces glioblastoma cell apoptosis. *Acta Pharmacol Sin*. (2021) 42:1338–46. doi: 10.1038/s41401-020-00551-x

48. Zhang H. CCND1 silencing suppresses liver cancer stem cell differentiation through inhibiting autophagy. *Hum Cell*. (2020) 33:140–7. doi: 10.1007/s13577-019-00295-9

49. Hussen BM, Hidayat HJ, Ghafouri-Fard S. Identification of expression of CCND1-related lncRNAs in breast cancer. *Pathol Res Pract*. (2022) 236:154009. doi: 10.1016/j.prp.2022.154009

50. Xie M, Zhao F, Zou X, Jin S, Xiong S. The association between CCND1 G870A polymorphism and colorectal cancer risk: A meta-analysis. *Med (Baltimore)*. (2017) 96: e8269. doi: 10.1097/MD.0000000000008269

51. Su Y, Zhou H, Ma Z, Liu J, Li C. CCND1-induced autophagy contributes to lymph node metastasis in endometrial cancer. *Horm Metab Res*. (2023) 55:413–9. doi: 10.1055/a-2044-9227

52. Berumen J, Espinosa AM, Medina I. Targeting CDKN3 in cervical cancer. *Expert Opin Ther Targets*. (2014) 18:1149–62. doi: 10.1517/14728222.2014.941808

53. Ma J, Zhou W, Yuan Y, Wang B, Meng X. PSMD12 interacts with CDKN3 and facilitates pancreatic cancer progression. *Cancer Gene Ther*. (2023) 30:1072–83. doi: 10.1038/s41417-023-00609-y

54. Li A, Cao C, Gan Y, Wang X, Wu T, Zhang Q, et al. ZNF677 suppresses renal cell carcinoma progression through N6-methyladenosine and transcriptional repression of CDKN3. *Clin Transl Med*. (2022) 12:e906. doi: 10.1002/ctm2.906

55. Cen J, Liang Y, Huang Y, Pan Y, Shu G, Zheng Z, et al. et al: Circular RNA circSDHC serves as a sponge for miR-127-3p to promote the proliferation and metastasis of renal cell carcinoma via the CDKN3/E2F1 axis. *Mol Cancer*. (2021) 20:19. doi: 10.1186/s12943-021-01314-w

56. Kalous J, Aleshkina D. Multiple roles of PLK1 in mitosis and meiosis. *Cells*. (2023) 12(1):187. doi: 10.3390/cells12010187

57. Liu Z, Sun Q, Wang X. PLK1, A potential target for cancer therapy. *Transl Oncol*. (2017) 10:22–32. doi: 10.1016/j.tranon.2016.10.003

58. Zhang Z, Cheng L, Li J, Qiao Q, Karki A, Allison DB, et al. et al: targeting plk1 sensitizes pancreatic cancer to immune checkpoint therapy. *Cancer Res*. (2022) 82:3532–48. doi: 10.1158/0008-5472.CAN-22-0018

59. Wang B, Huang X, Liang H, Yang H, Guo Z, Ai M, et al. et al: PLK1 Inhibition Sensitizes Breast Cancer Cells to Radiation via Suppressing Autophagy. *Int J Radiat Oncol Biol Phys*. (2021) 110:1234–47. doi: 10.1016/j.ijrobp.2021.02.025

60. Lin XT, Yu HQ, Fang L, Tan Y, Liu ZY, Wu D, et al. Elevated FBXO45 promotes liver tumorigenesis through enhancing IGF2BP1 ubiquitination and subsequent PLK1 upregulation. *Elife*. (2021) 10:e70715. doi: 10.7554/eLife.70715

61. Wagner KD, Du S, Martin L, Leccia N, Michiels JF, Wagner N. Vascular PPARbeta/delta promotes tumor angiogenesis and progression. *Cells*. (2019) 8 (12):1623. doi: 10.3390/cells8121623

62. Wagner N, Wagner KD. The role of PPARs in disease. *Cells*. (2020) 9(11):2367. doi: 10.3390/cells9112367

63. Choy JC, Hung VH, Hunter AL, Cheung PK, Motyka B, Goping IS, et al. Granzyme B induces smooth muscle cell apoptosis in the absence of perforin: involvement of extracellular matrix degradation. *Arterioscler Thromb Vasc Biol*. (2004) 24(11):2245–50. doi: 10.1161/01.ATV.0000147162.51930.b7

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 27 June 2024
DOI 10.3389/fimmu.2024.1400431

OPEN ACCESS

EDITED BY
Raquel Alarcon Rodriguez,
University of Almeria, Spain

REVIEWED BY
Liang-min Fu,
The First Affiliated Hospital of Sun Yat-sen University, China
Yuquan Chen,
Monash University, Australia
Aimin Jiang,
Fudan University, China
Zhifei Che,
First Affiliated Hospital of Hainan Medical University, China

*CORRESPONDENCE
Guanhu Yang
guanhuyang@gmail.com
Hao Chi
chihao7511@163.com
Jianhua Qin
2582536740@qq.com

†These authors have contributed equally to this work

RECEIVED 13 March 2024
ACCEPTED 14 June 2024
PUBLISHED 27 June 2024

CITATION
Jiang L, Ren X, Yang J, Chen H, Zhang S, Zhou X, Huang J, Jiang C, Gu Y, Tang J, Yang G, Chi H and Qin J (2024) Mitophagy and clear cell renal cell carcinoma: insights from single-cell and spatial transcriptomics analysis.
Front. Immunol. 15:1400431.
doi: 10.3389/fimmu.2024.1400431

Mitophagy and clear cell renal cell carcinoma: insights from single-cell and spatial transcriptomics analysis

Lai Jiang[1†], Xing Ren[2†], Jinyan Yang[3†], Haiqing Chen[1], Shengke Zhang[1], Xuancheng Zhou[1], Jinbang Huang[1], Chenglu Jiang[1], Yuheng Gu[1], Jingyi Tang[1], Guanhu Yang[4*], Hao Chi[1*] and Jianhua Qin[5,6*†]

[1]Clinical Medical College, Southwest Medical University, Luzhou, China, [2]Department of Oncology, Chongqing General Hospital, Chongqing, China, [3]School of Stomatology, Southwest Medical University, Luzhou, China, [4]Department of Specialty Medicine, Ohio University, Athens, OH, United States, [5]Department of Nephrology, Affiliated Hospital of Southwest Medical University, Luzhou, China, [6]Department of Nephrology, Sichuan Clinical Research Center for Nephropathy, The Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China

Background: Clear Cell Renal Cell Carcinoma (ccRCC) is the most common type of kidney cancer, characterized by high heterogeneity and complexity. Recent studies have identified mitochondrial defects and autophagy as key players in the development of ccRCC. This study aims to delve into the changes in mitophagic activity within ccRCC and its impact on the tumor microenvironment, revealing its role in tumor cell metabolism, development, and survival strategies.

Methods: Comprehensive analysis of ccRCC tumor tissues using single cell sequencing and spatial transcriptomics to reveal the role of mitophagy in ccRCC. Mitophagy was determined to be altered among renal clear cells by gene set scoring. Key mitophagy cell populations and key prognostic genes were identified using NMF analysis and survival analysis approaches. The role of UBB in ccRCC was also demonstrated by *in vitro* experiments.

Results: Compared to normal kidney tissue, various cell types within ccRCC tumor tissues exhibited significantly increased levels of mitophagy, especially renal clear cells. Key genes associated with increased mitophagy levels, such as UBC, UBA52, TOMM7, UBB, MAP1LC3B, and CSNK2B, were identified, with their high expression closely linked to poor patient prognosis. Particularly, the ubiquitination process involving the UBB gene was found to be crucial for mitophagy and its quality control.

Conclusion: This study highlights the central role of mitophagy and its regulatory factors in the development of ccRCC, revealing the significance of the UBB gene and its associated ubiquitination process in disease progression.

KEYWORDS

clear cell renal cell carcinoma, mitophagy, mitochondrial gene defects, multi-omics analysis, metabolic reprogramming, prognostic analysis, non-negative matrix factorization

1 Introduction

Renal cancer is a common malignant tumor, with its incidence continuously increasing worldwide (1). Despite some progress in treatment, many mysteries still remain regarding the pathogenesis of renal cancer (2). Clear Cell Renal Cell Carcinoma (ccRCC) is one of the most common types of renal cancer, occupying a major proportion of malignant kidney tumors (3, 4). This cancer typically originates from the epithelial cells of renal tubules and is characterized by high heterogeneity and complexity (5, 6). Compared to other tumor types, the treatment options for ccRCC are relatively limited, making it crucial to deepen our understanding of its pathogenesis for developing more effective treatment plans (7, 8).

Mitochondrial defects refer to structural or functional abnormalities in mitochondria, which can be caused by various factors, including genetic mutations, damage induced by environmental factors, increased oxidative stress, or damage to mitochondrial DNA (mtDNA) (9, 10). These defects often lead to an increased frequency of mitophagy. This is because mitochondrial defects, such as DNA damage, improper protein folding, increased oxidative stress, or insufficient energy production, can impair the normal function of mitochondria (11, 12).

In recent years, increasing evidence has suggested that mitophagy plays a key role in tumors (13, 14). Mitophagy is an intracellular self-degradation process through which cells can remove damaged mitochondria, thereby maintaining mitochondrial health (15). However, when mitophagy is dysregulated, it can lead to mitochondrial dysfunction, abnormal cell metabolism, and cell death (16). The anomalies in mitophagy associated with ccRCC suggest a close link between the two. In ccRCC, abnormalities in mitophagy may be caused by various factors, including changes in the intracellular and extracellular environment, genetic mutations, and dysregulation of regulatory pathways (17). These abnormalities not only affect the survival and proliferation of tumor cells but may also impact tumor development, invasion, and drug resistance (18).

This study aims to explore the connection between ccRCC and mitophagy genes through multi-omics analyses such as single-cell sequencing and spatial transcriptomics, revealing the importance of potential molecular aspects in the progression of renal cancer disease.

2 Materials and methods

2.1 Source of raw data

The single cell sequencing data of ccRCC used in this study were sourced from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) dataset GSE210038, which includes tumor samples from three patients with ccRCC (GSM6415686, GSM6415687, and GSM6415689) and one sample of normal adjacent tissue (GSM6415694). Through the analysis of these single-cell data, the study delves into the heterogeneity differences at the cellular level between renal cell carcinoma and adjacent normal tissues. The spatial transcriptomics data were also obtained from the GEO database (GSE210041), covering sequencing data for two formalin-fixed paraffin-embedded (FFPE) ccRCC tumor samples. This dataset provides a unique perspective for studying the spatial distribution heterogeneity of ccRCC and its surrounding environment. Additionally, RNA sequencing data for ccRCC were downloaded from the UCSC Xena platform (https://xena.ucsc.edu/), originating from the TCGA (The Cancer Genome Atlas) cohort, including sequencing information for 607 samples along with corresponding survival data for survival analysis, thereby enhancing our understanding of prognostic factors for ccRCC. Furthermore, genes related to mitophagy were sourced from the GSEA website (https://www.gsea-msigdb.org/gsea/index.jsp).

2.2 Processing of single-cell sequencing data

In this study, we analyzed the single-cell RNA-seq data of ccRCC using the Seurat package (version 4.3.0) in R (19). Through strict quality control, cells with a gene expression range of 200–4000 and mitochondrial gene expression ratio below 20% were selected. After standardization and normalization of the data, important principal components were determined using RunPCA and JackStraw analysis, followed by clustering and visualization with t-SNE to display the similarities and differences among cells. Differential expression analysis was conducted using the FindAllMarkers function, and cell types were annotated in conjunction with the CellMarker database (http://xteam.xbio.top/CellMarker/index.jsp), providing a data foundation for revealing the molecular mechanisms and potential therapeutic targets of ccRCC.

Five gene set scoring methods (AddModuleScore, ssGSEA, AUCell, UCell, singscore) were employed to score mitophagy-related genes in single-cell data. The mitophagy-related genes were obtained from the GSEA website and include 29 genes. The proteins encoded by these genes are involved in various processes including autophagosome formation, the composition of protein kinase CK2, mitochondrial fusion, mitochondrial fission, and ubiquitination processes. The use of multiple algorithms enhances the comprehensiveness, robustness, and biological interpretability of the assessments, allowing for more accurate determination of mitophagy in ccRCC. Additionally, clusterProfiler (4.6.2) and fgsea (1.24.0) were applied for enrichment analysis of single-cell transcriptomic data of ccRCC, precisely assessing gene set enrichment for cell types such as clear cells, supporting queries to various biological databases including GO, KEGG, and Reactome (20, 21). CellChat R package (version 1.6.1) was utilized to analyze cell communication patterns (22). CellChat simulates cell communication based on interactions between signaling ligands, receptors, and auxiliary factors, revealing how cells collaborate. To compare metabolic state differences between normal and tumor tissues, this study used the scMetabolism package (version 0.2.1) for quantitative analysis of single-cell metabolic pathway activity. We

also used the scFEA package to carry out flux studies to infer intracellular metabolites.

In this research, unsupervised non-negative matrix factorization (NMF) analysis of single-cell RNA sequencing data was applied using the NMF package (version 0.27) in R, aiming to explore the mitophagy characteristics of clear cell clusters (23). The component number was set to 10 to balance the granularity of different cell state distinctions and clustering interpretability. NMF results were integrated into the Seurat framework for dimensionality reduction clustering to identify different cell clusters. Key genetic markers were screened using the FindAllMarkers function, and each NMF cell cluster was categorized based on scores related to mitophagy-related genes and set thresholds. This method enhanced understanding of cell heterogeneity and tumor complexity, especially regarding mitophagy. Importantly, the ggplot2 package (version 3.4.2) served as our core tool for result visualization, offering a powerful and flexible way to create complex graphics based on the grammar of graphics.

2.3 Processing of spatial transcriptome sequencing data

In our study, the Seurat package (version 4.3.0) was used for the processing and analysis of spatial transcriptomics data, including normalization and feature selection of UMI counts with "SCTransform", and dimensionality reduction with "RunPCA". Additionally, the scMetabolism package was employed to assess metabolic features, while the "Monocle" package revealed cellular development and differentiation processes. In the Python environment, the Scanpy package processed spatial transcriptomics data through data preprocessing and dimensionality reduction with "SCTransform" and "RunPCA" (24). We also introduced the stLearn package, integrating gene expression, tissue morphology, and spatial location information to parse cell types, infer evolutionary paths, and identify cell interaction areas, providing a comprehensive spatial and functional perspective to understand tumor complexity (25).

2.4 Integrative analysis of spatial transcriptomics and single-cell sequencing data through deconvolution

Through deconvolution analysis, we inferred the proportions of cell types from mixed samples by combining single-cell and spatial transcriptomics data, revealing cellular and spatial heterogeneity within tissues. The "spacerxr" R package was used to perform RCTD analysis, constructing a reference model based on single-cell data and loading spatial data to form SpatialRNA objects. RCTD objects estimated the proportions of cell types in mixed samples through specific gene expression patterns, providing the distribution of cell types for each spot in the spatial data. Moreover, the "mistyR" package was employed to analyze cell interactions, revealing cellular interactions within tissues, inferring cell communication networks, and deepening the understanding of cell communication patterns in the tumor microenvironment (26).

2.5 Prognostic analysis of mitophagy-related clear cell subpopulations combined with bulk data

We explored the potential clinical prognostic value of newly identified mitophagy-related subpopulations of clear cells. For this purpose, we conducted an in-depth analysis using bulk sequencing data. Single-cell sequencing data were processed with the Seurat package, initially categorizing the identified mitophagy-related clear cells from patient tumor tissues into high and low expression subgroups based on their key gene expression levels. Next, the FindAllMarkers function was utilized to identify marker genes for these two subgroups. After obtaining the marker genes of key cell populations, we quantified these genes in bulk sequencing data, thus constructing high and low-risk groups. Lasso analysis was employed to filter out key prognostic genes for ccRCC, establishing a prognostic model based on mitophagy-related genes.

2.6 Cell culture and transient transfection

In our experimental studies, we utilized several cell lines, including the 786-O and 769-P renal clear cell carcinoma cells. These cell lines were obtained from the cell bank of the Central Laboratory at the Southwest Medical University Affiliated Hospital. To ensure the normal growth and maintenance of these cells, we cultured them in DMEM (HyClone) medium supplemented with 10% fetal bovine serum (HyClone), 100 U/L penicillin, and 100mg/L streptomycin (Thermo Fisher Scientific). We maintained standard culture conditions, including a 5% CO2 atmosphere, to provide an optimal environment for cell viability and experimental consistency. For the transient transfection experiments, we used Lipofectamine 3000 (Invitrogen, Carlsbad, CA, United States) as the transfection reagent. Negative control (NC) and UBB siRNA (RiboBio, Guangzhou, China) were transfected into the renal clear cell carcinoma cells according to the manufacturer's instructions. This involved preparing a transfection mixture containing the siRNA and transfection reagent and then adding it to the cells. The transfection process was generally conducted within the recommended time frame according to the manufacturer's protocol. By using Lipofectamine 3000 as the transfection reagent, our aim was to efficiently introduce the negative control or UBB siRNA into the renal clear cell carcinoma cells for subsequent analysis and research on the effects of gene knockdown or control on cellular processes and molecular pathways.

2.7 CCK-8 assay

We evaluated cell viability using the Cell Counting Kit-8 (CCK-8) assay. Twenty-four hours post-transfection, renal clear cell carcinoma cells were seeded into 96-well plates at a density of 1500 cells per well, and 200 μL of complete culture medium was added. The cells were then incubated at 37°C. For the CCK-8 assay, 10 μL of CCK-8

solution (Beyotime, Shanghai, China) was added to each well containing cells. After incubating for another 4 hours at 37°C, allowing the reagent to react with the cells, a colorimetric reaction related to cell viability occurs. At the end of the incubation period, the optical density (OD450) was measured using a microplate reader. The OD450 value reflects the absorbance of the formazan product generated by CCK-8, which is directly proportional to the metabolic activity and viability of the cells. By quantifying the OD450 values, we can assess the relative survival rate of the cells and compare them across different experimental conditions or treatment groups.

2.8 EdU-DAPI double staining assay

After 48 hours of transfection, 10 μM EdU was added and incubated for 4 hours, followed by fixation of cells with 4% paraformaldehyde for 10 minutes and permeabilization with 0.5% Triton X-100 for 5 minutes. EdU staining was performed using the Click-iT EdU Alexa Fluor 594 Imaging Kit according to the manufacturer's instructions, followed by staining of cell nuclei with 1 μg/mL DAPI for 10 min. Finally, the cells were observed and images were acquired using fluorescence microscopy. Merge images were used to analyze cell proliferation.

2.9 Wound healing experiment

To evaluate the migration ability of renal clear cell carcinoma cells, we employed a wound healing assay. The transfected cells were cultured in six-well plates and maintained at 37°C until they reached approximately 80% confluence. A uniform wound was introduced into the cell monolayer using a 200 μl sterile pipette tip. After wound formation, the cells were washed twice with PBS to remove any debris, and then the medium was supplemented with serum-free culture medium. The process of cell migration into the damaged area was recorded at 0 hours and 24 hours using an Olympus inverted microscope.

2.10 Transwell assay

The invasive ability of renal clear cell carcinoma cells was assessed using a well-established technique in cell biology research—the Transwell assay. In this assay, a specific number of renal clear cell carcinoma cells (approximately 1 × 10^5) were seeded into specialized chambers. To evaluate invasion potential, chambers coated with Matrigel were used. The upper chamber contained serum-free culture medium to create a chemotactic gradient, while the lower chamber was filled with complete DMEM culture medium, providing a favorable environment for cell movement. After 24 hours of culture, cells that had successfully invaded through the membrane were fixed with a 4% paraformaldehyde solution. To observe and quantify the invaded cells, they were stained with 0.1% crystal violet. The stained cells were then observed and counted under an optical microscope, allowing for the assessment of cell numbers and invasion capability.

2.11 Statistical analysis

The statistical analyses were conducted using R version 4.2.2, 64-bit, along with its support packages. The pycharm integrated development environment for Python was also utilized. The non-parametric Wilcoxon rank sum test was employed to assess the relationship between two groups for continuous variables. Spearman correlation analysis was conducted to examine correlation coefficients. A significance level of $P<0.05$ was considered statistically significant for all statistical investigations.

3 Results

3.1 Single-cell transcriptome atlas of clear cell renal cell carcinoma

In this study, we delved into the cellular heterogeneity and composition of ccRCC and its adjacent normal kidney tissue through scRNA-seq. To ensure the quality of data and rigor of analysis, we first performed meticulous quality control, quantifying multiple quality metrics including the assessment of the number of feature genes per cell, UMI counts, and the percentage of mitochondrial and hemoglobin gene expression, thereby eliminating the interference of senescent cells and erythrocytes (Figure 1A). Subsequently, we utilized the Harmony package for batch effect correction based on PCA analysis, which ensured the reliability of the analysis results while maximally preserving the original gene expression information of the cells (Figure 1B). By using the t-SNE algorithm, we performed a visualization of the cell clustering results, showing 22 cell clusters (Figure 1C). Based on the cell marker genes, we plotted bubble plots and feature plots to help us identify cell types by the expression and expression distribution of these genes (Figures 1D, E). After completing the cell type identification, we compared the cell distribution and number in ccRCC samples and normal kidney tissue samples, and observed that there was a significant increase in the proportion of T cells in ccRCC tissues (Figures 1F, G). The expression of cell marker genes in various cell types of cells is demonstrated by gene expression heatmap to check the accuracy of cell type identification (Figure 1H). To gain a preliminary understanding of the metabolic functionality of various cell types in tumor and normal tissues, we performed flux estimation analysis to infer intracellular fluxes of metabolites (Figure 1I).

3.2 Exploring mitophagy levels in ccRCC by gene scoring

To delve into the regulatory mechanisms of mitophagy in ccRCC and its role in the pathological process, this study quantitatively

10.3389/fimmu.2024.1400431

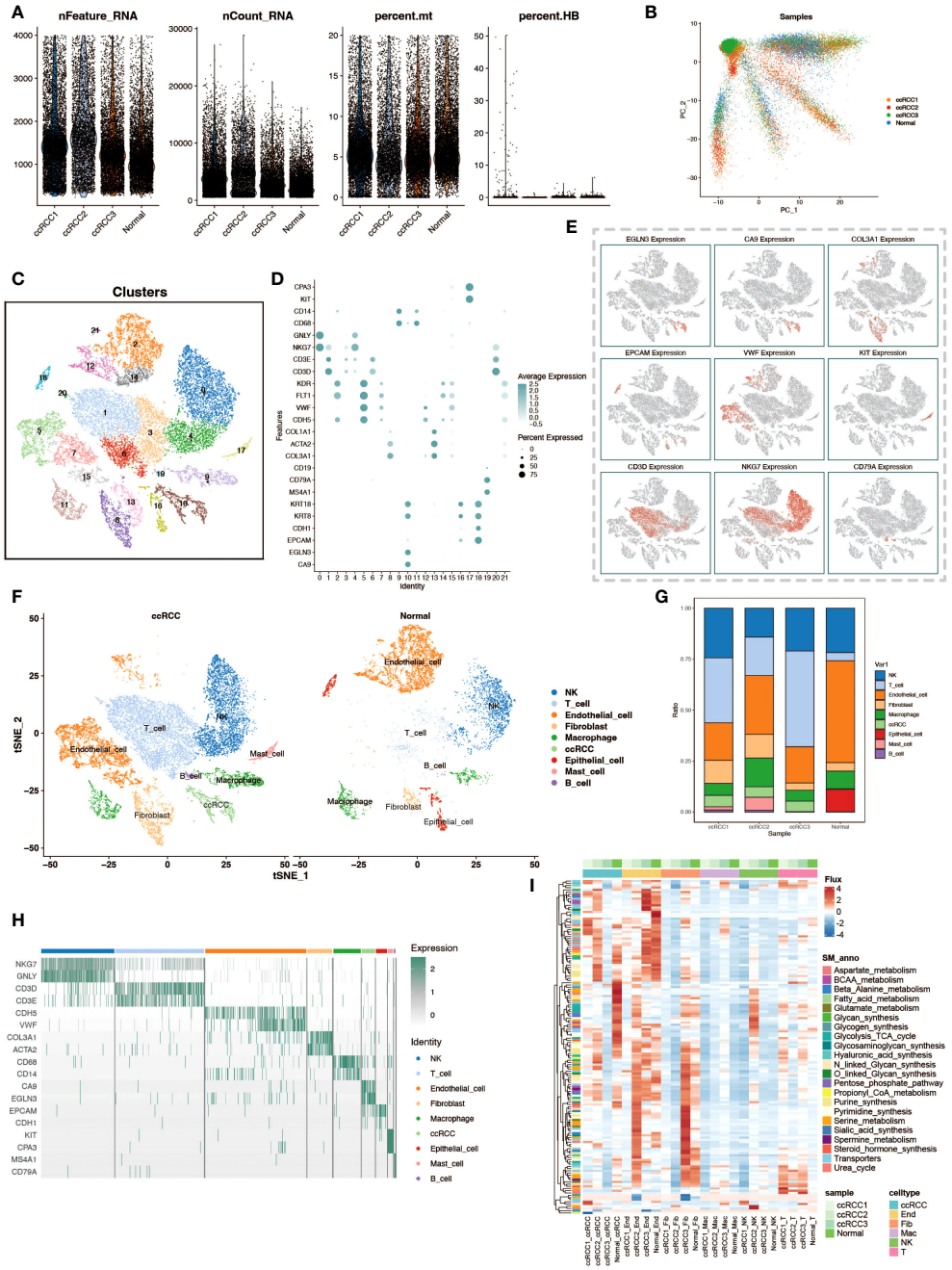


FIGURE 1
Single-cell transcriptomic atlas analysis of renal clear cell carcinoma. **(A)** Data quality control. Violin plots depict the number of genes per cell (nFeature_RNA), total transcript counts (nCount_RNA), percentage of mitochondrial genes (percent.mt), and percentage of hemoglobin genes (percent.HB) to evaluate sample quality. **(B)** PCA dimensionality reduction of patient samples. Principal component analysis (PCA) results based on expression profiles show the distribution of cell populations from different patients (ccRCC for tumor tissues of renal clear cell carcinoma patients, Normal for normal adjacent tissue samples). **(C)** t-SNE clustering visualization. The t-SNE dimensionality reduction technique reveals 22 distinct cell populations, each identified by a different color. **(D)** Marker genes of cell populations. Bubble charts display selected marker genes expressed in different cell populations. **(E)** Spatial expression patterns of cell marker genes. The t-SNE plot shows the expression patterns of selected marker genes. **(F)** Comparison of cell types between tumor and normal groups. The t-SNE plot shows the distribution of cell populations from the tumor and normal groups. **(G)** Stacked bar charts display the proportion of cell type distribution across different patient samples. **(H)** Heatmap of marker gene expression. Displays the expression levels of specific marker genes in different cell types. **(I)** Heatmap of metabolic levels. Shows how active various cell types are in different metabolic pathways in different samples.

evaluated the activity of mitophagy genes in ccRCC from multiple perspectives using various scoring algorithms, including AUCell, UCell, singcore, ssgsea, and AddModuleScore, revealing their potential role in tumor development. The analysis vividly presented the scores of mitophagy gene sets across various cell types through violin plots and bubble charts (Figures 2A, B). Heatmaps displayed the final scores for different cell types (Figure 2C). Comparing the scores of cells from different groups and performing Wilcox statistical

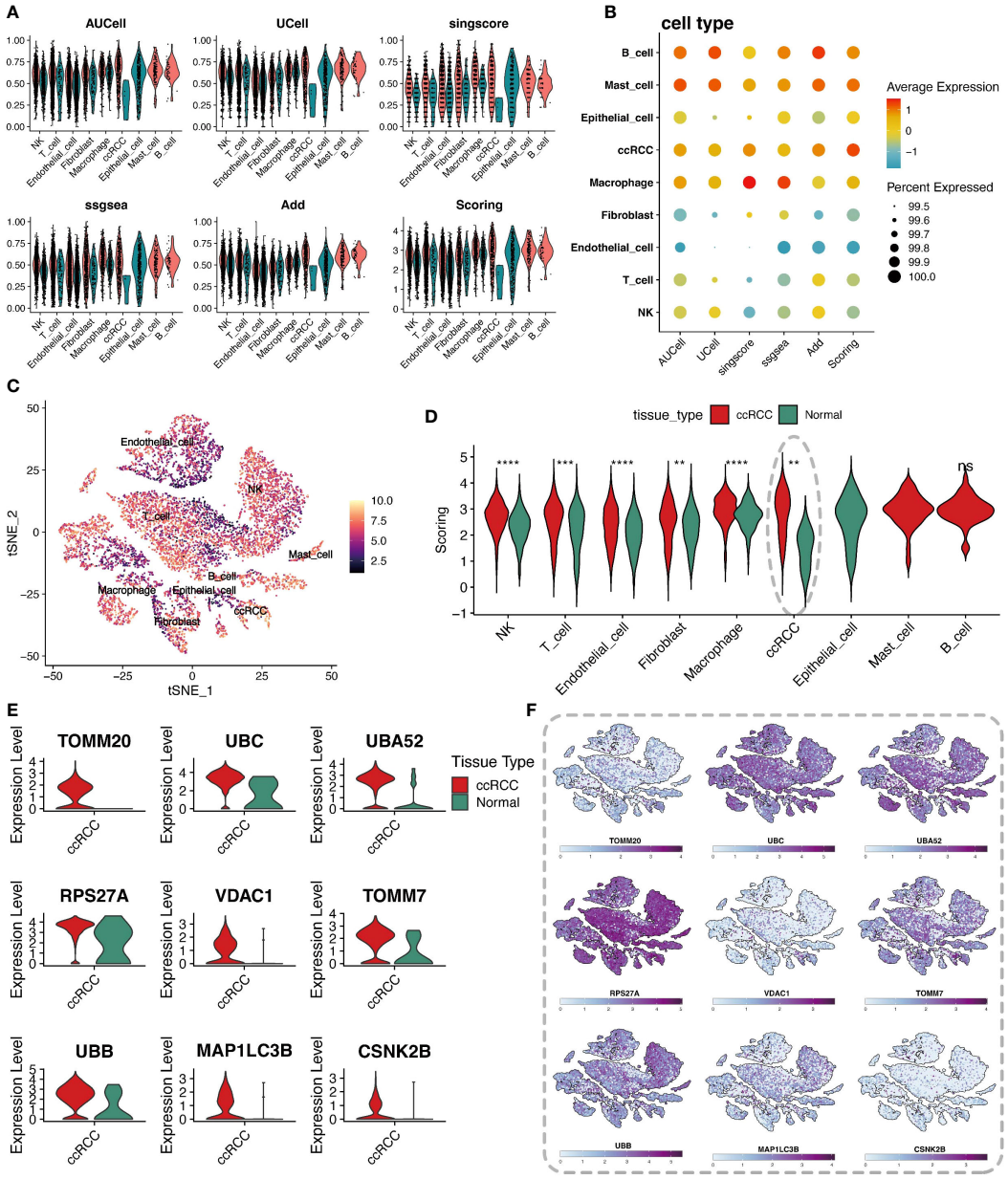


FIGURE 2
Analysis of cellular metabolic levels. **(A)** Violin plots of mitochondrial autophagy gene set scores through five gene set scoring methods and an integrated score. **(B)** Bubble chart of mitochondrial autophagy gene set expression scores in different cell types, based on the expression level of specific gene sets. **(C)** t-SNE plot showing the distribution of metabolic scores among cells, where the depth of color represents the level of scoring, revealing the metabolic heterogeneity of different cell types. **(D)** Violin plots comparing mitochondrial autophagy scores differences between tumor tissues and adjacent normal tissues for each cell type, showing changes in mitochondrial autophagy states in the tumor microenvironment. **(E)** Violin plots of differentially expressed mitochondrial autophagy genes between tumor and normal groups. **(F)** t-SNE plot showing the heatmap of differentially expressed mitochondrial autophagy genes expression levels in different cell types. "*" represents p-value less than 0.05, "**" represents p-value less than 0.01, "***" represents p-value less than 0.001. "****" represents p-value less than 0.0001. "ns" represents not statistically significant ($p \geq 0.05$).

analysis revealed that the scores of clear cells in the tumor group were significantly higher than those in the normal group, with statistically significant differences (p-value < 0.05) (Figure 2D). To further reveal which mitophagy genes play a key role in the pathogenesis of ccRCC, differential analysis was conducted between the tumor group and its normal counterpart, intersecting the resultant differential genes with mitophagy-related genes, and obtaining 9 key mitophagy-related genes with a logFC threshold of 0.5. The results showed that TOMM20, UBC, UBA52, RPS27A, and other genes were significantly upregulated in ccRCC cells (Figure 2E). Notably, these genes were not only universally upregulated in ccRCC cells but also widely distributed across various cell subpopulations (Figure 2F).

3.3 Characteristics of renal clear cells in the high and low mitophagy level group

We first analyzed the metabolic pathway activity in three ccRCC samples to determine the metabolic characteristics of ccRCC (Figure 3A). In these three tumor sample data, we divided renal clear cells into high and low groups according to the median value of his mitochondrial autophagy score to explore the effect of mitophagy levels on renal clear cell function and activity. Enrichment analyses showed a very significant difference in the functional activity of renal hyalinocytes between the high and low groups (Figure 3B). There were also differences in cellular communication between the high and low groups of renal hyalocytes, with the high mitophagy level group having a higher level of cellular communication than the low level group, both in terms of signaling efference and signaling reception, as well as differences in the structure of the communication patterns between the two groups (Figures 3C–E). The signaling pathways ligand receptors they involve also differ markedly in type and strength (Figure 3F). Differences between the two groups of cells were more clearly demonstrated by GSVA enrichment analyses, with renal hyalinocytes generally functioning more actively in the high-level group than in the low-level group (Figure 3G).

3.4 Application of non-negative matrix factorization (NMF) in revealing heterogeneity of mitophagy in renal clear cells

Non-negative matrix factorization (NMF) is a matrix decomposition method performed under the constraint that all elements of the output matrices are non-negative. Compared to principal component analysis (PCA), NMF has a natural advantage in analyzing tumor cell heterogeneity. By applying NMF technology and clustering ccRCC cells based on mitochondrial autophagy-related genes, we successfully identified five distinct subgroups (C0–4). To elucidate the potential link between the subgroups obtained by NMF analysis and mitochondrial autophagy, we performed differential expression analyses of cells in these subgroups. We obtained the differentially expressed genes for each subgroup and developed a series of rules to identify cell types: 1.differentially expressed genes were ranked according to logFC values. 2.If the first gene was a mitochondrial autophagy-related gene with a logFC value greater than 1 and a P value of less than 0.05, then the cell population was defined as a cell population marked by this gene. 3. If the first gene is a mitochondrial autophagy-related gene but its logFC value is less than 1 or its P value is greater than 0.05, then the cell population cannot be defined. 4. if the first gene is not a mitochondrial autophagy-related gene, then the cell population is defined as a non-mitochondrial autophagy cell population (Non-Mitophagy). With this approach, we were able to identify and categorize cell populations more clearly. The results yielded four subgroups: Unclear-ccRCC-C3, Non-Mitophagy-ccRCC-C4, CSNK2B+ccRCC-C1, MAP1LC3B+ccRCC-C2 (Figure 4A). Among them, Unclear-ccRCC-C3 was named due to the most significant gene logFC not meeting the selection criteria, and the Non-Mitophagy-ccRCC-C4 subgroup's most significant gene was not a mitochondrial autophagy gene. We performed a series of analyses on the four cell subpopulations of renal clear cells obtained to investigate differences in functional activity and biological heterogeneity. The results of the Hallmark enrichment analysis showed that the Non-Mitophagy cell population (C4) was significantly different from the remaining three populations, which happen to be the ones that are or may be related to mitophagy (Figure 4B). GSVA enrichment analysis also demonstrated concordant results, which further demonstrated the accuracy of the NMF analysis in identifying the mitophagy renal hyalinocyte subpopulation (Figure 4E). For the two identified populations of mitophagy -associated renal clear cells (CSNK2B +ccRCC-C1 and MAP1LC3B+ccRCC-C2), we performed separate GO enrichment analyses for further exploration of these two key cell types (Figures 4C, D). In the transcription factor analysis, opposite results were presented, with a stronger relationship between the Non-Mitophagy cell population and the transcription factors (Figure 4F). The cellular metabolic profiles in the four cell subpopulations demonstrated very clear differences between Non-Mitophagy and mitophagy -associated renal hyalinocytes, with mitophagy -associated renal hyalinocytes being much more advanced than the Non-Mitophagy cell population in a variety of metabolic pathways (Figure 4G).

3.5 Analysis of metabolic features in spatial transcriptomics data

Spatial transcriptomics data provided HE stained slice images of two ccRCC tumor tissue samples (Figures 5A, F). After dimensionality reduction clustering of spatial transcriptomics data, we mapped the clustering information onto the HE stained slices, obtaining dimensionality reduction clustering maps on the slices (Figures 5B, G). The differential expression of mitochondrial autophagy genes between tumor and normal groups in single-cell data was displayed on spatial transcriptomics data through bubble charts (Figures 5C, H). By using the scMetabolism package for metabolic analysis of spatial transcriptomics data, we showed the specific metabolic levels of each cell cluster in the two tumor samples (Figures 5D, I) and also mapped certain key metabolisms onto the slices. The heatmap colors allowed us to clearly see the high and low states of metabolism at different locations on the slices (Figures 5E, J).

3.6 Pseudotime analysis of spatial transcriptomics

In Figures 5E, J, we observed high metabolic areas on the slices of two tumor samples, with clusters 2 and 1 being the main high metabolic areas on the first slice, and clusters 10 and 4 on the second slice. Therefore, we selected the high metabolic areas and their surrounding cells for pseudotime analysis using the Monocle package. For the first slice, we conducted pseudotime analysis on cell clusters 2, 1, and 12 (Figure 6A). The heatmap showed the expression changes of mitochondrial autophagy-related genes over

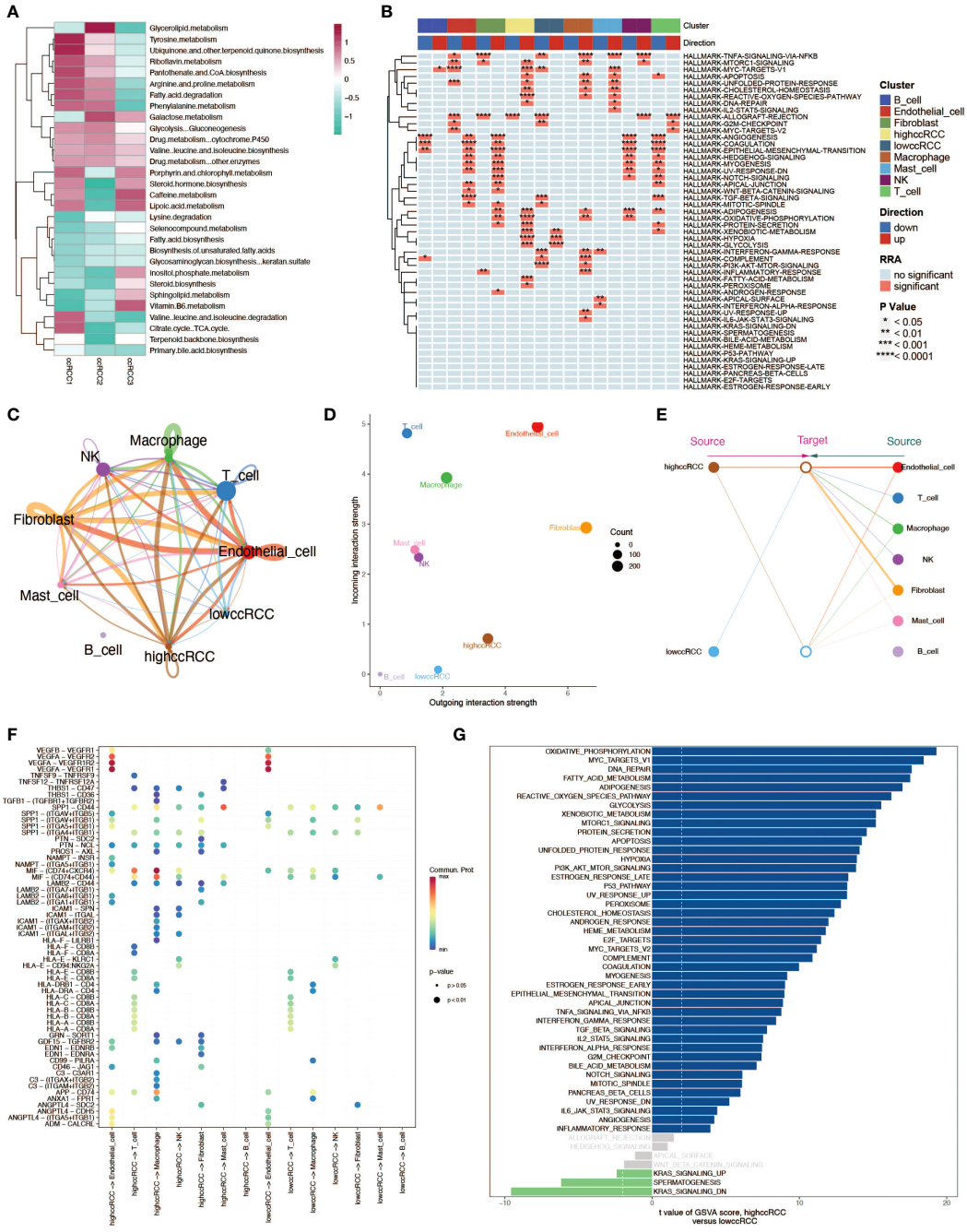


FIGURE 3
Multidimensional analysis of renal clear cells. **(A)** Heatmap of metabolic activity in renal clear cells from different sample sources, analyzed using the scMetabolism package, showing their expression patterns across different metabolic pathways. **(B)** HALLMARK gene set enrichment analysis. Individual cell types (including high and low group ccRCC) are shown to be up- and down-regulated in various pathways. **(C)** String diagram of cellular communication networks. Demonstrates the strength of cellular communication in tumor tissues. **(D)** Scatter plot demonstrating the average strength of signals received and sent by cells in each cell type. **(E)** Structural diagram demonstrating the communication patterns of various cell types in the cellular communication network, comparing the high and low groups of ccRCC. **(F)** Ligand receptor activation involved in cellular communication between high and low groups of ccRCC and other cell types. **(G)** GSVA enrichment analysis of evanescent bar graphs. Pathways with significant differences between the high-level group ccRCC and the low-level group are demonstrated.

pseudotime (Figure 6B). Cluster 1 occupied the earliest branch in the pseudotime analysis, while cluster 2 was on a later branch, which might indicate the developmental sequence of tumor cells (Figure 6C). The cell density map also hinted at the timing of cell appearances (Figure 6D). For the second slice, we analyzed cell clusters 10, 4, and 2, with the heatmap showing the expression changes of mitochondrial autophagy-related genes over pseudotime (Figures 6E, F). Cell clusters 10 and 4 were primarily in the early stages of the pseudotime sequence, while cluster 2 was mainly in the later stages (Figures 6G, H).

10.3389/fimmu.2024.1400431

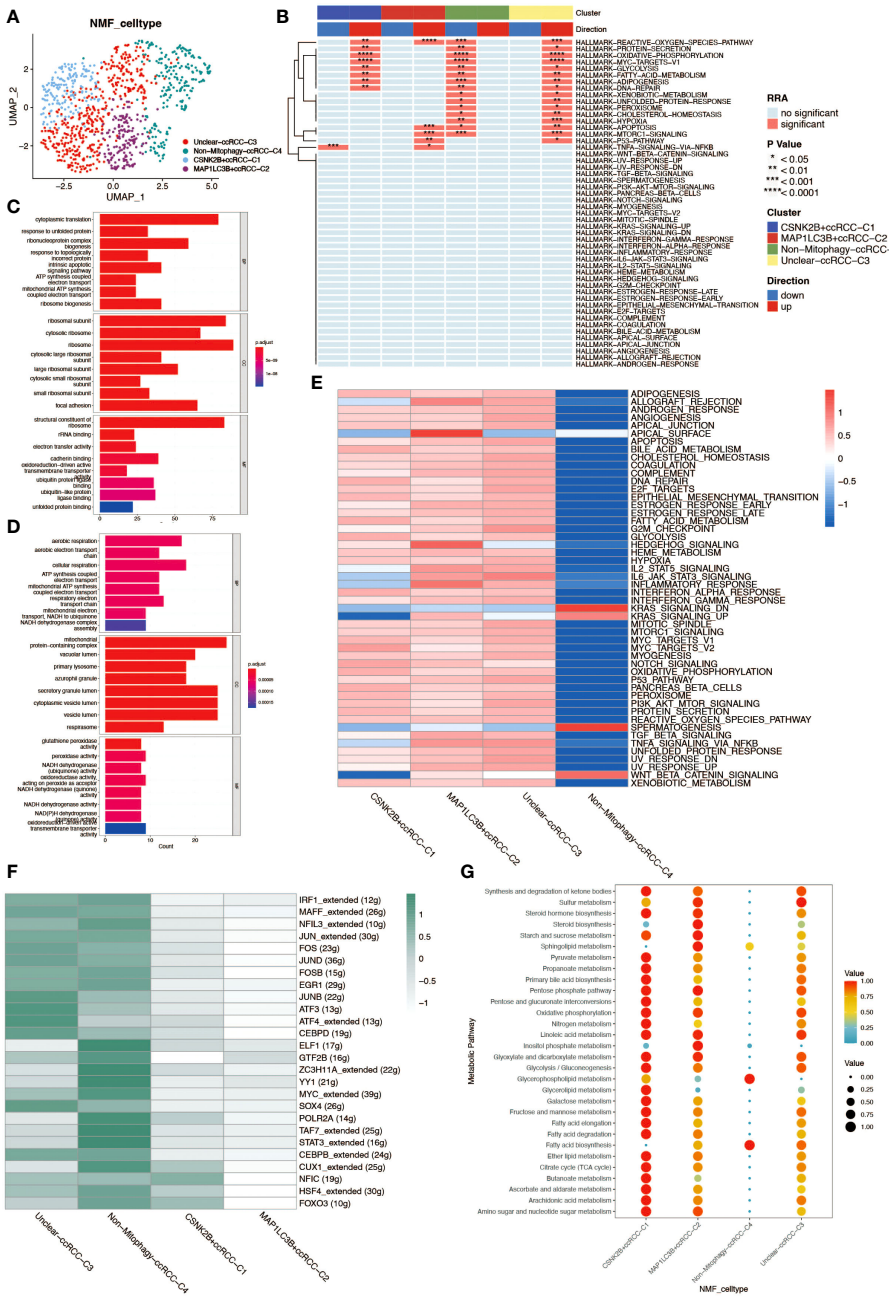


FIGURE 4
Non-negative matrix factorization (NMF) analysis of renal clear cells. **(A)** UMAP plot annotating renal clear cell types post-NMF, identifying two subtypes closely related to mitochondrial autophagy. **(B)** HALLMARK gene set enrichment analysis. The up- and down-regulation of the four ccRCC subgroups obtained from NMF analysis in various pathways is shown. **(C)** GO enrichment analysis of cell population C1. results of GO enrichment analysis of CSNK2B+ccRCC-C1 cell population demonstrating properties in BP, CC, and MF. **(D)** GO enrichment analysis of cell population C2. results of GO enrichment analysis of MAP1LC3B+ccRCC-C2 cell population, demonstrating the properties in three aspects: BP, CC, and MF. **(E)** GSVA enrichment analysis. The heatmap demonstrates the differences between the four ccRCC subpopulations in various pathways. **(F)** Heatmap of the transcription factor regulatory network in mitochondrial autophagy-related subtypes of renal clear cells. **(G)** Bubble chart analyzing the activity levels of renal clear cell subgroups in different metabolic pathways, where bubble size and color depth reflect the relative levels of metabolic activity.

3.7 Developmental trajectories revealed by spatial transcriptomics data

Spatial transcriptomics data provide transcriptional information on the precise location of cells within tissues. Using the stLearn toolkit, we conducted an in-depth analysis of spatial transcriptomics data to explore the developmental processes of tumors, including invasion and metastasis issues. By combining data quality control and dimensionality reduction with NumPy, and clustering with stLearn's Louvain method, we identified different cell clusters in ccRCC samples (Figures 7A, D). For cell clusters identified in the early stages of pseudotime sequence in the

10.3389/fimmu.2024.1400431

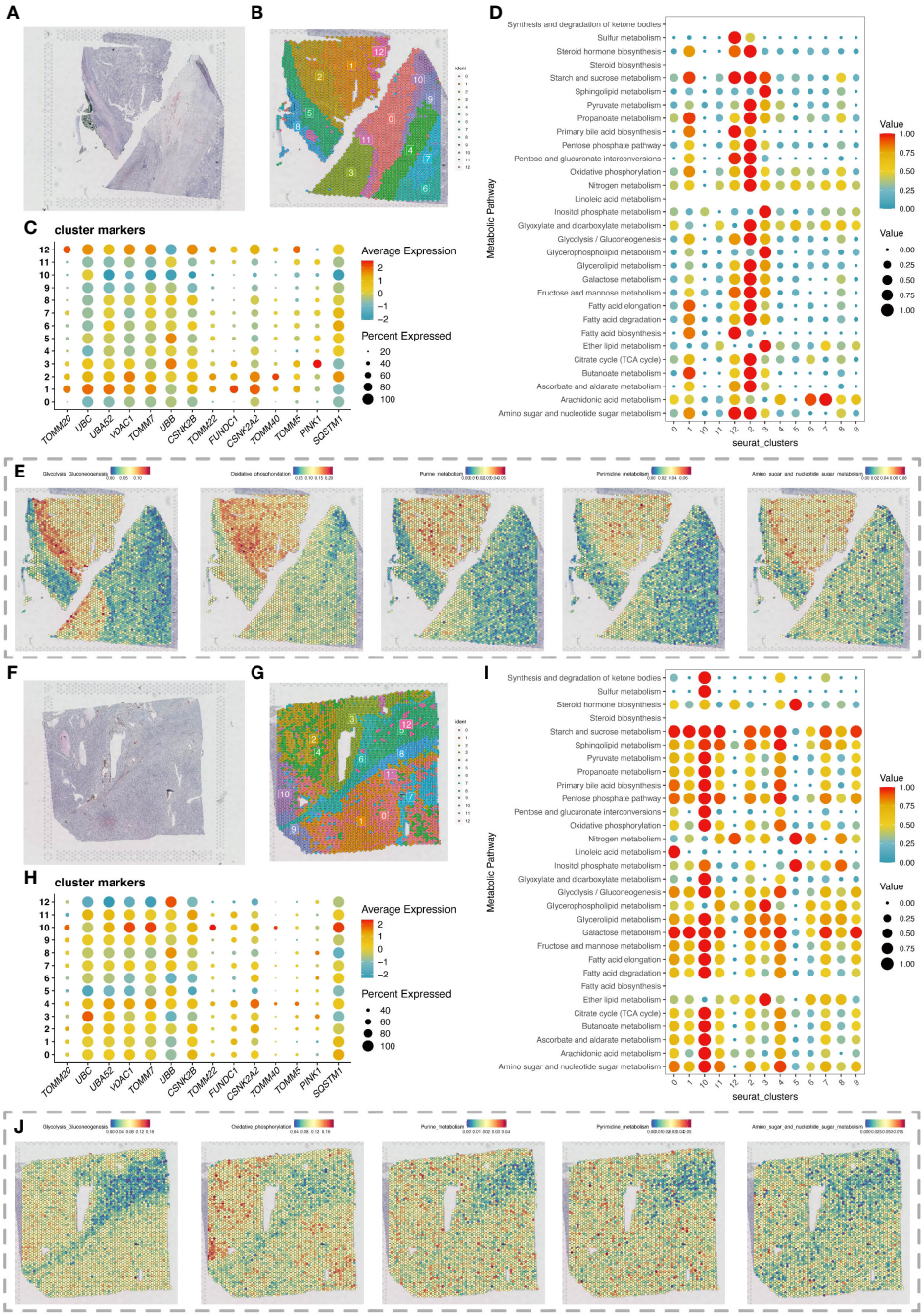


FIGURE 5
Spatial transcriptomics analysis revealing changes in metabolic activity. **(A)** H&E-stained section of renal clear cell carcinoma tumor tissue. **(B)** Spatial transcriptomics data of a renal clear cell carcinoma tumor tissue section, with cell clustering results obtained by dimensionality reduction clustering analysis using the Seurat package. **(C)** Expression of differentially expressed mitochondrial autophagy-related genes in the section data. **(D)** Bubble chart showing metabolic activity levels of different cell clusters in the renal clear cell carcinoma tumor tissue section analyzed with the scMetabolism package, highlighting each cluster's performance in various metabolic pathways. **(E)** Display of various metabolic levels on the section, including glycolysis, oxidative phosphorylation, purine metabolism, pyrimidine metabolism, and the metabolism of amino sugar and nucleotide sugar. **(F)** Second H&E-stained section of renal clear cell carcinoma tumor tissue. **(G)** Spatial transcriptomics data of a second renal clear cell carcinoma tumor tissue section, with cell clustering results obtained by dimensionality reduction clustering analysis using the Seurat package. **(H)** Expression of differentially expressed mitochondrial autophagy-related genes in the second section data. **(I)** Bubble chart showing metabolic activity levels of different cell clusters in the second renal clear cell carcinoma tumor tissue section analyzed with the scMetabolism package, highlighting each cluster's performance in various metabolic pathways. **(J)** Display of various metabolic levels on the second section, including glycolysis, oxidative phosphorylation, purine metabolism, pyrimidine metabolism, and the metabolism of amino sugar and nucleotide sugar.

10.3389/fimmu.2024.1400431

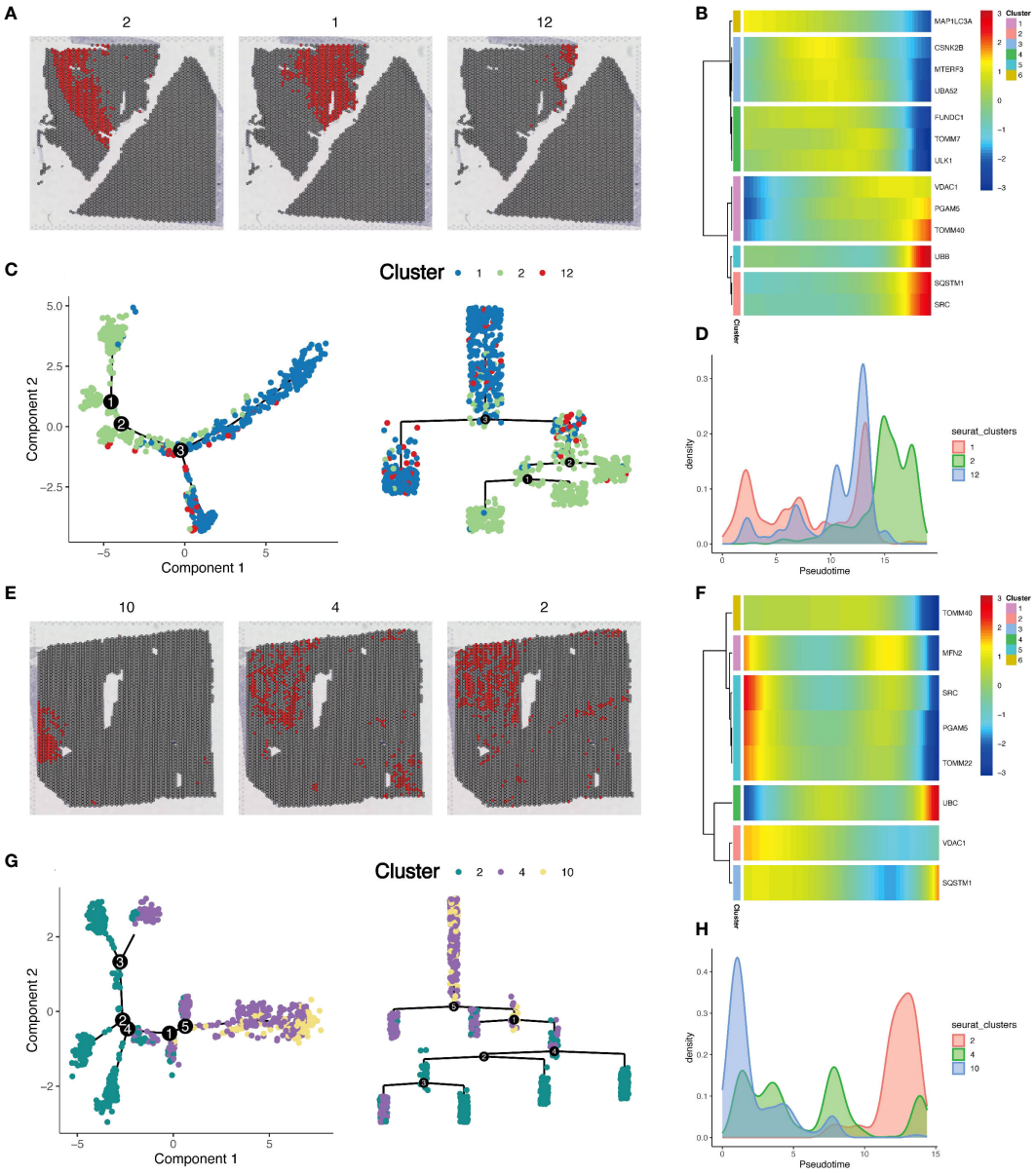


FIGURE 6
Pseudotime analysis of cells in local areas of tumor tissue sections. **(A)** Display of local cell populations on the section. **(B)** Expression of mitochondrial autophagy-related genes in pseudotime order. **(C)** Developmental trajectory map, showing the dynamic changes and differentiation paths of 3 cell populations in pseudotime development. **(D)** Density map explaining the distribution characteristics of each group of cells on the pseudotime axis. **(E)** Display of local cell populations on the section in a second renal clear cell carcinoma tumor slice. **(F)** Expression of mitochondrial autophagy-related genes in pseudotime order on the second slice. **(G)** Developmental trajectory map for the second slice data, showing the dynamic changes and differentiation paths of 3 cell populations in pseudotime development. **(H)** Density map explaining the distribution characteristics of each group of cells on the pseudotime axis for the second slice.

pseudotime analysis, we reconstructed the developmental trajectories using the Diffusion Pseudotime (DPT) algorithm, combined with spatial coordinates information, revealing the gradual invasion and metastasis process of tumor cells in the pseudotime sequence, consistent with Monocle pseudotime analysis (Figures 7B, E). The diverging bar charts of developmental trajectory analysis revealed gene expression changes based on trajectory differences, showing genes that were upregulated and downregulated throughout the tumor development process from start to end (Figures 7C, F).

3.8 Deconvolution and cell interaction analysis combining spatial transcriptomics with single-cell data

Due to the limitations of spatial transcriptomics sequencing technology, current spatial transcriptomics data do not achieve the same single-cell resolution as single-cell sequencing data. To address the limitations of spatial transcriptomics sequencing data, we employed deconvolution analysis methods to compensate for its lack of resolution. This analysis inferred the possible cell types and their

10.3389/fimmu.2024.1400431

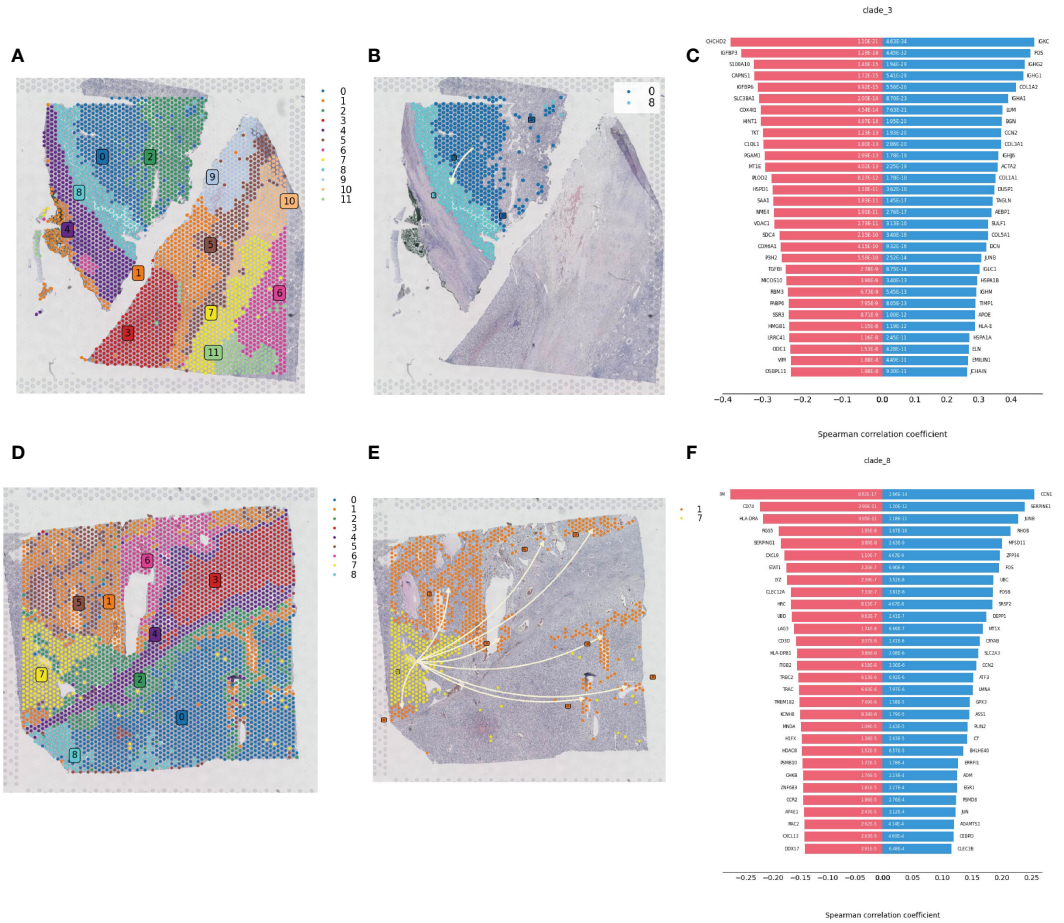


FIGURE 7
Spatial developmental trajectory analysis of renal clear cell carcinoma tumor tissue. **(A)** Clustering of renal clear cell carcinoma tumor tissue section sequencing data using the louvain method in the stLearn package, with the clustering map showing the spatial distribution of different cell populations. **(B)** Spatial developmental trajectory map of high metabolic area cells in the tumor tissue section, drawn using the stLearn package. **(C)** Diverging bar chart of developmental trajectory-related differentially expressed genes in the tumor tissue, performed statistical analysis using numpy, revealing key regulatory genes associated with the developmental trajectory. **(D)** Clustering map of the second slice data. **(E)** Spatial developmental trajectory map of high metabolic area cells. **(F)** Diverging bar chart of developmental trajectory-related differentially expressed genes in the tumor tissue.

proportions at each location in the spatial transcriptomics data based on the gene expression patterns of various cell types in ccRCC single cell sequencing data. This step allowed us to gain deeper insights into the spatial structure and function of tissues or cells, revealing interactions and communications between different cell types, and discovering spatial heterogeneity and state changes of cells. Through this method, we were able to provide more detailed information about cell types and proportions in ccRCC tumor samples, offering new perspectives and depth to the study (Figures 8A, G). Based on the deconvolution analysis of two tumor samples, we further applied the MISTy (Multiview Intercellular SpaTial modeling framework) framework for spatial transcriptomics cell interaction analysis. This framework is an interpretable machine learning framework for analyzing single-cell, highly multiplexed, spatially resolved data, enabling an in-depth understanding of the internal and intercellular relationships between markers. With MISTy, we could handle a custom number of views, each describing different spatial contexts such as intracellular regulation or paracrine regulation, and relationships between specific cell types. Our analysis results showed the contributions of three different views to cell interactions through bar charts, finding that intraview and paraview15 made the largest contributions in the two tumor samples (Figures 8B, H). This revealed the importance of intracellular regulation and paracrine regulation in tumor samples. Further heatmap and network graph analyses revealed the specific patterns of these two views in tumor samples, highlighting the significant interactions between two groups of clear cells with high and low mitochondrial autophagy states and other cell types (such as mast cells and fibroblasts) (Figures 8C–F, I–L).

3.9 Prognostic study of mitochondrial autophagy-related genes

In our study, nine key mitochondrial autophagy-related genes were significantly higher expressed in tumor tissues compared to normal tissues. We analyzed data of ccRCC from the TCGA database, first selecting positive cells with high expression of these nine genes, and compared them with negative cells with low

10.3389/fimmu.2024.1400431

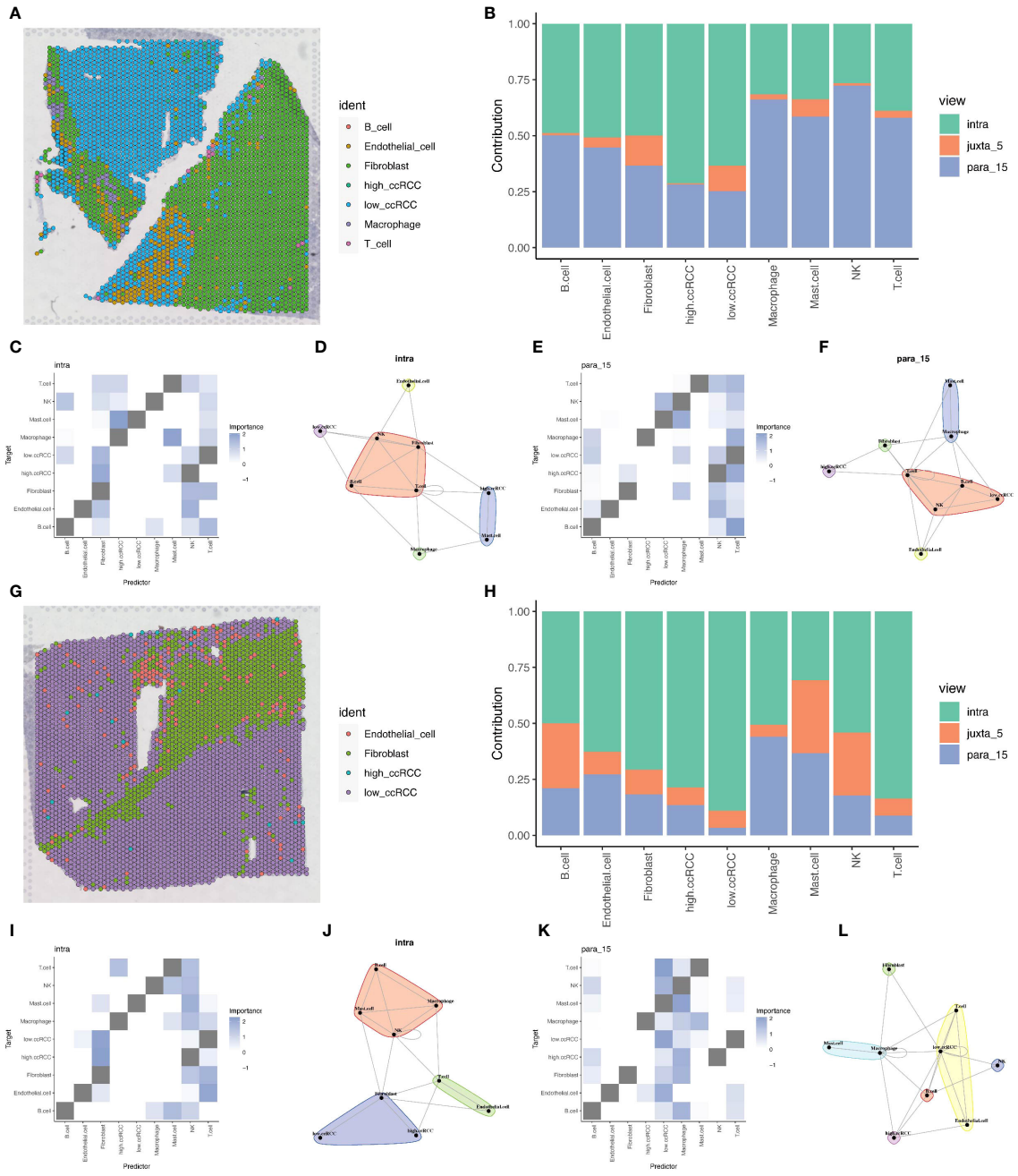


FIGURE 8
Deconvolution and cell interaction analysis based on spatial transcriptomics data. **(A)** Analysis of renal clear cell carcinoma tumor tissue section data using the RCTD deconvolution method, showing the spatial distribution probabilities of various cell types, including cells with high and low levels of mitochondrial autophagy. **(B)** Bar chart showing the contribution of different views to cell interactions assessed by the Mistyr package, demonstrating the relative importance of different views in cell interactions. **(C, D)** Heatmap and network diagram of cell interactions within the same view (intraview), revealing the interaction strength and patterns within the same cell type. **(E, F)** Heatmap and network diagram of cell interactions in the paraview15 view, showing the interaction strength and communication networks across cell types. **(G)** RCTD deconvolution analysis results of the second slice data, showing the probabilities and spatial distribution of different cell types, including cells with high and low levels of mitochondrial autophagy. **(H)** Bar chart showing the contribution of different views to cell interactions in prostate adenocarcinoma with infiltrating carcinoma tissue, assessing the relative contributions of each view. **(I, J)** Heatmap and network diagram of cell interactions (intraview) for the second slice data, showing the interaction relationships among the same cell type in the tumor environment. **(K, L)** Heatmap and network diagram of cell interactions in the paraview15 view of the same tissue, revealing the interaction strength and network structures across different cell types.

expression to identify unique marker genes of the positive cells. Subsequently, based on the expression levels of these marker genes, we divided patients into high and low expression groups and performed survival analysis. The results showed that patients with high expression of UBC, UBA52, TOMM7, UBB, MAP1LC3B, and CSNK2B had a poorer prognosis, with statistical significance (Figure 9A). Using LASSO regression model analysis, UBB and TOMM7 were identified as important prognostic factors for ccRCC

10.3389/fimmu.2024.1400431

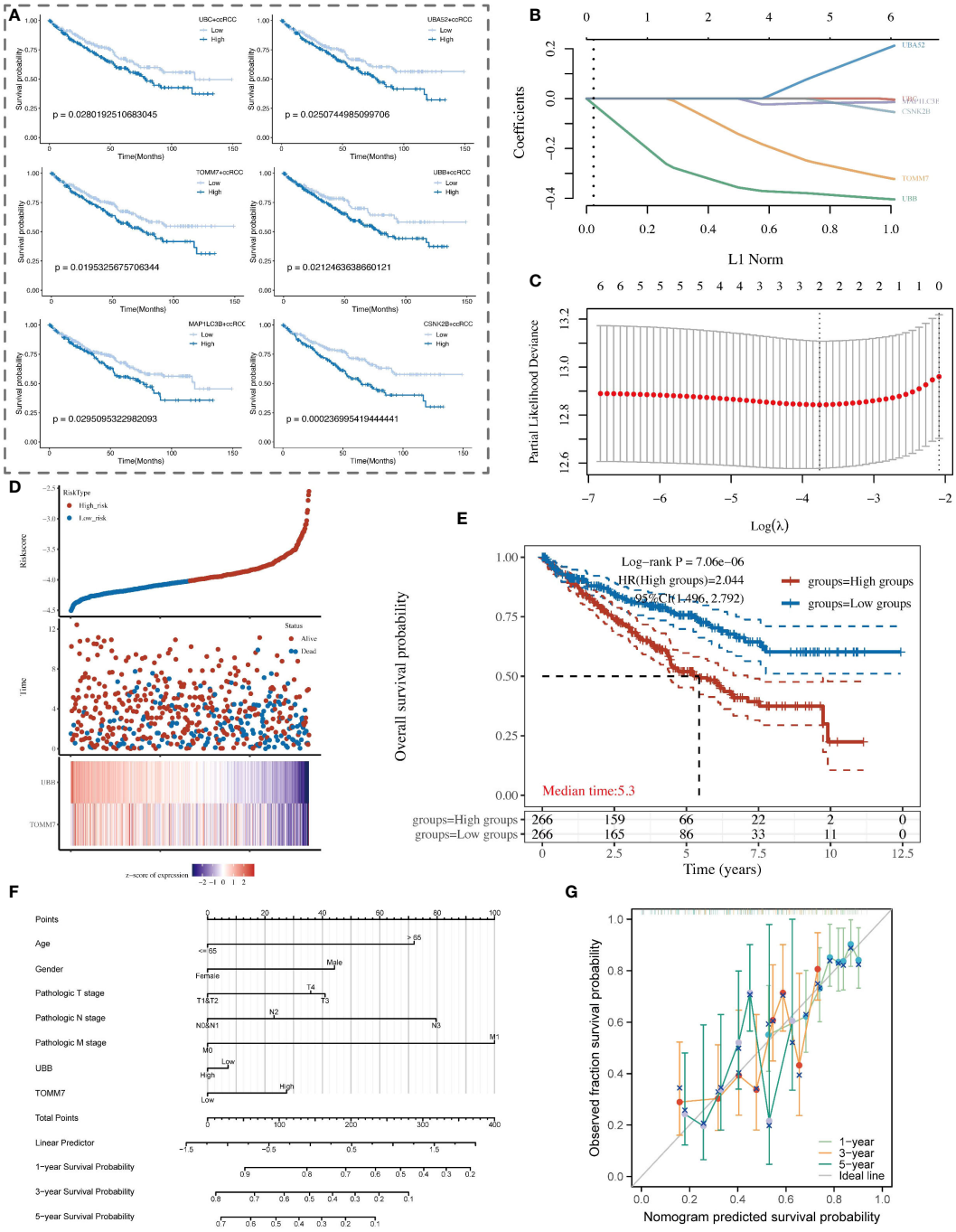


FIGURE 9
Analysis of the association between mitochondrial autophagy-related subtypes of renal clear cells and the clinical prognosis of patients with renal clear cell carcinoma. **(A)** Kaplan-Meier survival curve analysis showing the survival probability differences among renal clear cell subtypes with high and low expression of key mitochondrial autophagy genes in patients with renal clear cell carcinoma. **(B)** LASSO coefficient path graph. It illustrates how the LASSO coefficients of renal clear cell carcinoma prognosis-related genes change as the regularization strength (L1 norm) of the model increases. Selected genes maintain non-zero coefficients at high regularization levels, indicating their importance to the model. **(C)** Deviance plot of ten-fold cross-validation. Displays the performance of the LASSO model at different lambda values to determine the optimal lambda selection. The red dot identifies the lambda value providing the optimal prognosis model, determined by minimizing the validation error. **(D)** Risk score and survival status chart of patients with renal clear cell carcinoma based on the LASSO model. **(E)** Kaplan-Meier survival analysis curve of high and low-risk patient groups with renal clear cell carcinoma. **(F)** Nomogram model for the prognosis of patients with renal clear cell carcinoma, combining clinical variables such as age, gender, pathological staging, and the expression levels of UBB and TOMM7 genes. **(G)** Calibration curve of the nomogram prognosis model, showing the consistency between the nomogram-predicted 1-year, 3-year, and 5-year survival probabilities (X-axis) and the actual observed survival probabilities (Y-axis).

(Figures 9B–D). Kaplan-Meier curves showed that the survival rate of patients in the high-expression group was significantly lower than that of the low-expression group (Figure 9E). We also established a nomogram that includes these genes and clinical parameters to predict the survival probabilities of patients at 1, 3, and 5 years, and calibration curves validated the accuracy and reliability of this prognostic model (Figures 9F, G).

3.10 UBB promotes the proliferation and migration of renal clear cell carcinoma cells

To investigate the potential role of UBB in renal clear cell carcinoma, we conducted *in vitro* experiments. Initially, the CCK-8 assay indicated that silencing UBB significantly inhibited cell proliferation (Figure 10A). Silencing of the UBB gene resulted in a significant reduction in DAPI staining (blue) and EdU staining (red) signals in the 786 and 769 cell lines, indicating a decrease in both the number of cells and the number of DNA-synthesizing cells. The results showed that UBB gene knockdown significantly inhibited the proliferation of tumor cells (Figure 10B). Furthermore, wound healing assays, and transwell assays showed that knocking out UBB significantly reduced the cells' invasion and migration capabilities (Figures 10C, D). Taken together, these results suggest that the upregulation of UBB promotes the proliferation, invasion, and migration of renal clear cell carcinoma cells.

4 Discussion

Mitochondrial defects, including structural or functional abnormalities caused by genetic mutations, damage from environmental factors, increased oxidative stress, or mitochondrial DNA (mtDNA) damage, impact cellular proliferation, death, and metabolism and are closely linked to the development and progression of cancer (27, 28). These defects can trigger mitophagy—a cellular adaptive mechanism that maintains cell survival by removing dysfunctional mitochondria to prevent further cellular damage. As a quality control mechanism, mitophagy aids in the clearance of unhealthy or dysfunctional mitochondria, averting potential cellular damage caused by mitochondrial defects (12, 29). With an increase in mitochondrial defects, autophagy activity correspondingly intensifies to address these deficiencies. The process of mitophagy includes multiple steps: recognition of mitochondrial damage, formation of autophagosomes, fusion with lysosomes, degradation, and recycling (12). In our research, we observed changes in the levels of mitochondrial autophagy in various cell types within ccRCC tumor tissues compared to normal kidney tissue, especially a significant enhancement of mitochondrial autophagy levels in clear cells within tumor groups.

The enhancement of mitochondrial autophagy in ccRCC tissues can be understood from multiple perspectives: Firstly, tumor cells undergo metabolic reprogramming to adapt to the tumor microenvironment and promote survival and proliferation, activating more frequent mitochondrial autophagy to maintain intracellular metabolic balance. Secondly, ccRCC cells may experience increased oxidative stress, leading to mitochondrial damage, and enhance mitochondrial autophagy to clear damaged mitochondria, preventing the accumulation of oxidative damage that could lead to cell death. Additionally, mitochondrial autophagy may serve as a self-regulatory mechanism, helping tumor cells optimize survival strategies to adapt to stress conditions in the microenvironment (30, 31).

Through the analysis of multiple transcriptomic data, we identified several key genes closely related to mitochondrial autophagy, suggesting that these genes may be the main factors driving the changes in mitochondrial autophagy function in clear cells of renal cell carcinoma, especially in the prognostic analyses the high expression of six mitochondrial autophagy-related genes, namely, UBC, UBA52, TOMM7, UBB, MAP1LC3B, and CSNK2B was closely associated with poor patient prognosis. Among these genes, UBC (Ubiquitin C), UBA52 (Ubiquitin A-52 Residue Ribosomal Protein Fusion Product 1), and UBB (Ubiquitin B) are involved in the ubiquitination process—a critical protein modification mechanism that tags proteins for degradation or other fates (32). The ubiquitin-proteasome system plays a central role in regulating protein levels, maintaining protein homeostasis, and participating in cellular stress responses. TOMM7 (Translocase Of Outer Mitochondrial Membrane 7) is part of the mitochondrial protein import complex, responsible for transporting proteins from the cytosol into the mitochondria (33). MAP1LC3B (Microtubule Associated Protein 1 Light Chain 3 Beta) is a key protein in the autophagy process, involved in the formation of autophagosomes (34). CSNK2B (Casein Kinase 2 Beta), as part of the protein kinase CK2, is involved in various cellular processes including cell cycle regulation, cell survival, and DNA repair (35). Further analyses identified UBB and TOMM7 as important prognostic factors for ccRCC.

UBB is a protein-coding gene involved in the process of ubiquitination and is also associated with mitochondrial autophagy. The ubiquitination process plays a critical and widespread regulatory role within the cell, maintaining the stability of the intracellular environment and responding to environmental changes by controlling the fate of proteins. In mitochondrial autophagy, ubiquitination plays a central role, primarily by covalently attaching ubiquitin proteins to specific proteins on the surface of damaged or dysfunctional mitochondria, thereby marking these mitochondria for recognition and clearance by autophagosomes. The involvement of specific receptor proteins such as p62/SQSTM1, OPTN, and NBR1 allows these ubiquitinated mitochondria to interact with LC3 proteins on the autophagosome membrane, promoting the formation and expansion of autophagosomes to encapsulate and ultimately digest the damaged mitochondria (12, 29). Specifically, the ubiquitin B protein encoded by the UBB gene plays a core role in marking damaged or obsolete proteins for recognition and degradation by the 26S proteasome. By regulating the selective degradation of mitochondria, the UBB gene and its encoded ubiquitin B protein are crucial for maintaining mitochondrial quality control and intracellular environmental stability. High

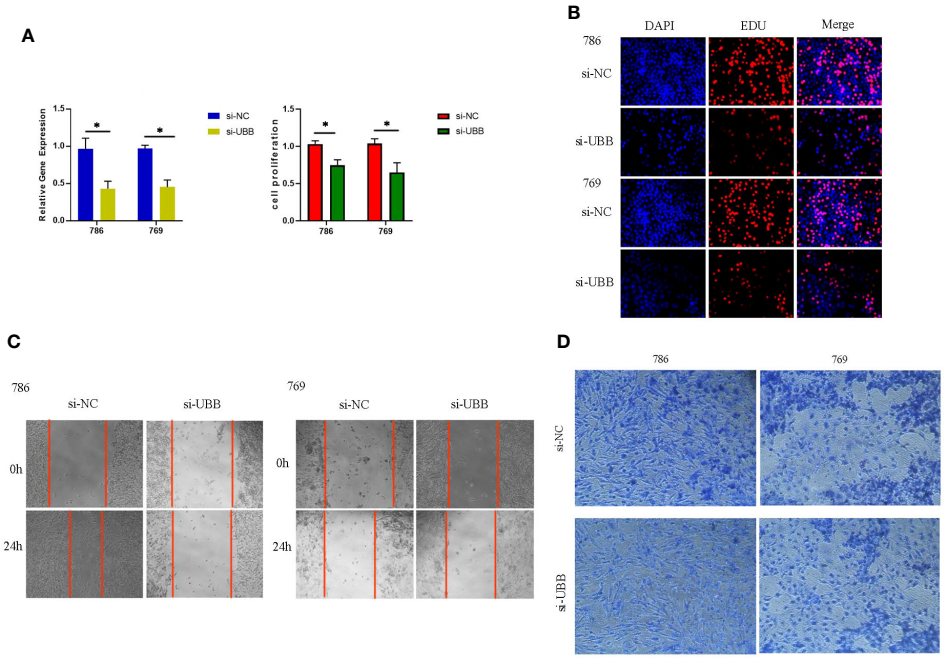


FIGURE 10
UBB has been demonstrated to promote the proliferation, invasion, and migration of renal clear cell carcinoma cells, as determined by a series of analytical results: **(A)** CCK-8 assay. **(B)** EdU-DAPI Double Staining Assay. **(C)** wound healing assay. **(D)** Transwell assay. * Indicates p-value < 0.05.

expression of the UBB gene may enhance the ubiquitination marking and rapid clearance of damaged mitochondria, helping tumor cells effectively remove damaged mitochondria to prevent cellular stress and death, thereby increasing the tumor cells' adaptability to adverse conditions. Furthermore, high expression of the UBB gene may also strengthen the adaptive response of the autophagy pathway under stress conditions such as nutrient deprivation or hypoxia, providing a survival advantage for tumor cells, especially in the challenging tumor microenvironment.

We have demonstrated through *in vitro* experiments that the proliferation, invasion and migration of tumor cells can be inhibited by decreasing the expression of the UBB gene in tumor cells. The results based on transcriptome data analysis and *in vitro* experiments demonstrated that UBB, a mitochondrial autophagy-related gene, has a very important role in renal clear cell carcinoma, which provides a new direction for potential clinical treatment. We can envisage the development of siRNA drugs or small molecule inhibitors based on the UBB gene, thereby reducing its expression level in tumor cells to inhibit tumor adaptability and growth. In conclusion, through in-depth research and clinical application of the UBB gene, we can provide more precise and effective therapeutic options for ccRCC patients and significantly improve their prognosis.

5 Conclusion

This study highlights the importance of increased mitochondrial autophagy in ccRCC and its impact on tumor behavior. By advanced analysis, key genes such as TOMM7 and UBB were associated with autophagy and prognosis, with the role of UBB in ubiquitination emphasizing its therapeutic potential. These findings highlight the central role of mitochondrial autophagy in ccRCC, suggesting new therapeutic targets and improving personalized treatment for ccRCC patients.

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.

Author contributions

LJ: Conceptualization, Data curation, Formal analysis, Visualization, Writing – original draft. XR: Conceptualization, Data curation, Writing – original draft. JY: Conceptualization, Data curation, Formal analysis, Writing – original draft. HQC: Formal analysis, Visualization, Writing – original draft. SZ: Formal analysis, Visualization, Writing – original draft. XZ: Formal analysis, Visualization, Writing – original draft. JH: Formal analysis, Visualization, Writing – original draft. CJ: Formal analysis, Visualization, Writing – original draft. YG: Formal analysis, Visualization, Writing – original draft. JT: Writing – review & editing. GY: Writing – review & editing, Conceptualization. HC: Writing – review & editing. JQ: Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We gratefully acknowledge the Gene Expression Omnibus (GEO), The Cancer Genome Atlas (TCGA), and the UCSC Xena website for their public datasets. Special thanks to the authors of "Mesenchymal-like Tumor Cells and Myofibroblastic Cancer-Associated Fibroblasts Are Associated with Progression and Immunotherapy Response of Clear Cell Renal Cell Carcinoma" for their ccRCC data contribution on GEO. We also appreciate the Gene Set Enrichment Analysis (GSEA) website for the mitochondrial autophagy-related gene sets and thank the R programming language community for the essential tools that facilitated our research. This work was made possible by the collective effort and resources of the scientific community.

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.

Publisher's note

References

1. Wang J, Tang J, Chen T, Yue S, Fu W, Xie Z, et al. A web-based prediction model for overall survival of elderly patients with early renal cell carcinoma: a population-based study. *J Transl Med*. (2022) 20:90. doi: 10.1186/s12967-022-03287-w

2. Gulati S, Vaishampayan U. Current state of systemic therapies for advanced renal cell carcinoma. *Curr Oncol Rep*. (2020) 22:26. doi: 10.1007/s11912-020-0892-1

3. Xu W, Lu J, Liu WR, Anwaier A, Wu Y, Tian X, et al. Heterogeneity in tertiary lymphoid structures predicts distinct prognosis and immune microenvironment characterizations of clear cell renal cell carcinoma. *J Immunother Cancer*. (2023) 11: e006667. doi: 10.1136/jitc-2023-006667

4. Jiang A, Ye J, Zhou Y, Zhu B, Lu J, Ge S, et al. Copper death inducer, FDX1, as a prognostic biomarker reshaping tumor immunity in clear cell renal cell carcinoma. *Cells*. (2023) 12:349. doi: 10.3390/cells12030349

5. Lih TM, Cho KC, Schnaubelt M, Hu Y, Zhang H. Integrated glycoproteomic characterization of clear cell renal cell carcinoma. *Cell Rep*. (2023) 42:112409. doi: 10.1016/j.celrep.2023.112409

6. Huang J, Liu M, Chen H, Zhang J, Xie X, Jiang L, et al. Elucidating the Influence of MPT-driven necrosis-linked LncRNAs on immunotherapy outcomes, sensitivity to chemotherapy, and mechanisms of cell death in clear cell renal carcinoma. *Front Oncol*. (2023) 13:1276715. doi: 10.3389/fonc.2023.1276715

7. von Roemeling CA, Marlow LA, Wei JJ, Cooper SJ, Caulfield TR, Wu K, et al. Stearoyl-CoA desaturase 1 is a novel molecular therapeutic target for clear cell renal cell carcinoma. *Clin Cancer Res*. (2013) 19:2368–80. doi: 10.1158/1078-0432.CCR-12-3249

8. Jiang A, Meng J, Gong W, Zhang Z, Gan X, Wang J, et al. Elevated SNRPA1, as a Promising Predictor Reflecting Severe Clinical Outcome via Effecting Tumor Immunity for ccRCC, Is Related to Cell Invasion, Metastasis, and Sunitinib Sensitivity. *Front Immunol*. (2022) 13:842069. doi: 10.3389/fimmu.2022.842069

9. Taylor RW, Turnbull DM. Mitochondrial DNA mutations in human disease. *Nat Rev Genet*. (2005) 6:389–402. doi: 10.1038/nrg1606

10. Schon EA, DiMauro S, Hirano M. Human mitochondrial DNA: roles of inherited and somatic mutations. *Nat Rev Genet*. (2012) 13:878–90. doi: 10.1038/nrg3275

11. Nadalutti CA, Stefanick DF, Zhao M-L, Horton JK, Prasad R, Brooks AM, et al. Mitochondrial dysfunction and DNA damage accompany enhanced levels of formaldehyde in cultured primary human fibroblasts. *Sci Rep*. (2020) 10:5575. doi: 10.1038/s41598-020-61477-2

12. Palikaras K, Lionaki E, Tavernarakis N. Mechanisms of mitophagy in cellular homeostasis, physiology and pathology. *Nat Cell Biol*. (2018) 20:1013–22. doi: 10.1038/s41556-018-0176-2

13. Gaude E, Frezza C. Defects in mitochondrial metabolism and cancer. *Cancer Metab*. (2014) 2:10. doi: 10.1186/2049-3002-2-10

14. Zou P, Liu L, Zheng LD, Payne KK, Manjili MH, Idowu MO, et al. Coordinated upregulation of mitochondrial biogenesis and autophagy in breast cancer cells: the role of dynamin related protein-1 and implication for breast cancer treatment. *Oxid Med Cell Longev*. (2016) 2016:4085727. doi: 10.1155/2016/4085727

15. Yan Y, Li M, Lin J, Ji Y, Wang K, Yan D, et al. Adenosine monophosphate activated protein kinase contributes to skeletal muscle health through the control of mitochondrial function. *Front Pharmacol*. (2022) 13:947387. doi: 10.3389/fphar.2022.947387

16. Yu R, Lendahl U, Nistér M, Zhao J. Regulation of mammalian mitochondrial dynamics: opportunities and challenges. *Front Endocrinol (Lausanne)*. (2020) 11:374. doi: 10.3389/fendo.2020.00374

17. Sun L, Zhang H, Gao P. Metabolic reprogramming and epigenetic modifications on the path to cancer. *Protein Cell*. (2022) 13:877–919. doi: 10.1007/s13238-021-00846-7

18. Yang Y, Jin M, Meng Y, Dai Y, Chen S, Zhou Y, et al. Involvement and targeted intervention of benzo(a)pyrene-regulated apoptosis related proteome modification and muti-drug resistance in hepatocellular carcinoma. *Cell Death Dis*. (2023) 14:265. doi: 10.1038/s41419-023-05771-7

19. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, 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

20. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. *bioRxiv*. (2021), 060012. doi: 10.1101/060012

21. 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

22. 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

23. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. *BMC Bioinf*. (2010) 11:367. doi: 10.1186/1471-2105-11-367

24. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol*. (2018) 19:15. doi: 10.1186/s13059-017-1382-0

25. Pham D, Tan X, Balderson B, Xu J, Grice LF, Yoon S, et al. Robust mapping of spatiotemporal trajectories and cell–cell interactions in healthy and diseased tissues. *Nat Commun*. (2023) 14:7739. doi: 10.1038/s41467-023-43120-6

26. Tanevski J, Flores ROR, Gabor A, Schapiro D, Saez-Rodriguez J. Explainable multiview framework for dissecting spatial relationships from highly multiplexed data. *Genome Biol*. (2022) 23:97. doi: 10.1186/s13059-022-02663-5

27. Carew JS, Huang P. Mitochondrial defects in cancer. *Mol Cancer*. (2002) 1:9. doi: 10.1186/1476-4598-1-9

28. Porporato PE, Filigheddu N, Pedro JMB-S, Kroemer G, Galluzzi L. Mitochondrial metabolism and cancer. *Cell Res*. (2018) 28:265–80. doi: 10.1038/cr.2017.155

29. Li A, Gao M, Liu B, Qin Y, L. chen H, Wu H, et al. Mitochondrial autophagy: molecular mechanisms and implications for cardiovascular disease. *Cell Death Dis*. (2022) 13:444. doi: 10.1038/s41419-022-04906-6

30. Liu Z, Fan M, Hou J, Pan S, Xu Y, Zhang H, et al. Serine hydroxymethyltransferase 2 knockdown induces apoptosis in ccRCC by causing lysosomal membrane permeabilization via metabolic reprogramming. *Cell Death Dis*. (2023) 14:144. doi: 10.1038/s41419-023-05677-4

31. Coffey NJ, Simon MC. Metabolic alterations in hereditary and sporadic renal cell carcinoma. *Nat Rev Nephrol*. (2024) 20(4):233–50. doi: 10.1038/s41581-023-00800-2

32. Deng L, Meng T, Chen L, Wei W, Wang P. The role of ubiquitination in tumorigenesis and targeted drug discovery. *Signal Transduct Target Ther*. (2020) 5:11. doi: 10.1038/s41392-020-0107-0

33. Wang W, Chen X, Zhang L, Yi J, Ma Q, Yin J, et al. Atomic structure of human TOM core complex. *Cell Discov*. (2020) 6:67. doi: 10.1038/s41421-020-00198-2

34. Xie Z, Klionsky DJ. Autophagosome formation: core machinery and adaptations. *Nat Cell Biol*. (2007) 9:1102–9. doi: 10.1038/ncb1007-1102

35. Rabalski AJ, Gyenis L, Litchfield DW. Molecular pathways: emergence of protein kinase CK2 (CSNK2) as a potential target to inhibit survival and DNA damage response and repair pathways in cancer cells. *Clin Cancer Res*. (2016) 22:2840–7. doi: 10.1158/1078-0432.CCR-15-1314

frontiers | Frontiers in Oncology

TYPE Systematic Review
PUBLISHED 08 July 2024
DOI 10.3389/fonc.2024.1432212

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Yinhui Yao,
Affiliated Hospital of Chengde Medical University, China
Kui Wang,
Shanghai Jiao Tong University, China

*CORRESPONDENCE
Hongcan Shi
shihongcan@163.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 13 May 2024
ACCEPTED 19 June 2024
PUBLISHED 08 July 2024

CITATION
Yuan L, Shen Z, Shan Y, Zhu J, Wang Q, Lu Y and Shi H (2024) Unveiling the landscape of pathomics in personalized immunotherapy for lung cancer: a bibliometric analysis. *Front. Oncol.* 14:1432212. doi: 10.3389/fonc.2024.1432212

Unveiling the landscape of pathomics in personalized immunotherapy for lung cancer: a bibliometric analysis

Lei Yuan[1,2,3†], Zhiming Shen[1,2,3†], Yibo Shan[1,2,3†], Jianwei Zhu[1,2,3], Qi Wang[1,2,3], Yi Lu[1,2,3] and Hongcan Shi[1,2,3*]

[1]Department of Thoracic Surgery, Northern Jiangsu People's Hospital Affiliated to Yangzhou University, Yangzhou, China, [2]Institute of Translational Medicine, Medical College, Yangzhou University, Yangzhou, China, [3]Jiangsu Key Laboratory of Integrated Traditional Chinese and Western Medicine for Prevention and Treatment of Senile Diseases, Yangzhou University, Yangzhou, China

Background: Pathomics has emerged as a promising biomarker that could facilitate personalized immunotherapy in lung cancer. It is essential to elucidate the global research trends and emerging prospects in this domain.

Methods: The annual distribution, journals, authors, countries, institutions, and keywords of articles published between 2018 and 2023 were visualized and analyzed using CiteSpace and other bibliometric tools.

Results: A total of 109 relevant articles or reviews were included, demonstrating an overall upward trend; The terms "deep learning", "tumor microenvironment", "biomarkers", "image analysis", "immunotherapy", and "survival prediction", etc. are hot keywords in this field.

Conclusion: In future research endeavors, advanced methodologies involving artificial intelligence and pathomics will be deployed for the digital analysis of tumor tissues and the tumor microenvironment in lung cancer patients, leveraging histopathological tissue sections. Through the integration of comprehensive multi-omics data, this strategy aims to enhance the depth of assessment, characterization, and understanding of the tumor microenvironment, thereby elucidating a broader spectrum of tumor features. Consequently, the development of a multimodal fusion model will ensue, enabling precise evaluation of personalized immunotherapy efficacy and prognosis for lung cancer patients, potentially establishing a pivotal frontier in this domain of investigation.

KEYWORDS

lung cancer, pathomics, artificial intelligence, deep learning, tumor microenvironment, immunotherapy

1 Introduction

Lung cancer remains one of the most prevalent malignancies and represents the foremost cause of cancer-related mortality worldwide (1, 2), the majority of lung cancers (80–90%) manifest as non-small cell lung cancer (NSCLC), often diagnosed at an advanced stage (65%), potentially with concurrent local or distant metastasis (3). Recent advances in immunotherapy, particularly the use of immune checkpoint inhibitors (ICIs), have shown promising outcomes in enhancing the prognosis of lung cancer patients (4). Nevertheless, not all patients experience the benefits of immunotherapy, highlighting the need for additional research into predictive biomarkers of immune response. These biomarkers, which may include substances, structures, or products of processes within the body, have the potential to facilitate personalized immunotherapy by enabling the monitoring of immune reactions.

Each lung cancer patient undergoes histopathological diagnosis, involving the preparation of biopsy tissues into pathological slides for examination. The traditional preservation method of using wax embedding techniques for pathological slides can now be digitized through computerization, archiving them as digital pathology images. This technological advancement serves as a foundation for applying big data analytics to digital pathology images. Consequently, the field of pathomics has emerged (5). Pathomics entails applying machine learning techniques to extract large-scale, objectively quantifiable, and readily analyzable datasets from digitally scanned pathological tissue images. Consistent with the pathological diagnostic requirements of diseases, morphological features, including size and shape of pathological images, along with multi-dimensional subtle features reflecting potential biological characteristics such as texture features and edge gradient features, are extracted. These features can be utilized for quantitative disease screening, diagnosis, prognosis prediction, and other applications (6).

In this study, CiteSpace (7) was utilized for the inaugural analysis of hotspots and trends in the application of pathomics in lung cancer. The objective is to provide valuable insights for scholars involved in research within this domain.

2 Materials and methods

2.1 Data collection

Web of Science Core Collection (WoSCC) database was chosen as the literature retrieval platform. The retrieval period spanned from 2018 to 2023, with the final search conducted on October 20, 2023. Subject terms were exclusively employed as the search method, and the search formula was: TS= ("Pathomics" OR "Pathomics" OR "Digital Pathology" OR "Whole-slide Imaging" OR "Whole Slide Imaging" OR "Computational Pathology") AND TS=("Lung Cancer" OR "Pulmonary Cancer" OR "Carcinoma of Lung" OR "Pulmonary Carcinoma" OR "Cancer of Lung" OR "Bronchogenic Carcinoma" OR "Bronchogenic" OR "Cancer of the Lung" OR "NSCLC" OR "SLC"), document type: Articles or Review Articles; a total of 109 documents were retrieved.

2.2 Statistical methods

Export the complete records and referenced bibliographies of the 109 documents retrieved from WoSCC in Text format, comprising 85 articles and 24 reviews. Conduct a comprehensive analysis of the literature using CiteSpace 6.2.R4 (64-bit) Basic, focusing on the country, institution, authorship, keywords, and cited references. The bibliometric online analysis platform, developed by the National Science Library of the Chinese Academy of Sciences, was employed to conduct a visual analysis of historical keywords and national collaborations.

3 Results

3.1 Annual publication volume in WoSCC

A total of 109 matching documents were retrieved, and the overall publication output exhibited a general upward trend, especially reaching a contribution rate of 26.61% in 2021 (Figure 1). The annual average publication output is approximately 21.8 articles. The results indicate a gradual increase in the attention to pathomics research in the context of lung cancer.

3.2 Distribution of source journals

The literature selected from the 109 studies on pathomics in the management of lung cancer has been indexed by 146 journals. For the top 10 journals in terms of publication output, detailed information on Journal Citation Reports (JCR) category, publication quantity, impact factor (IF), and their respective contribution percentages is provided in Table 1.

3.3 Visualization of collaborations between countries and institutions

Running the CiteSpace software for country analysis resulted in a knowledge graph with 35 nodes and 80 edges (Figure 2). Each

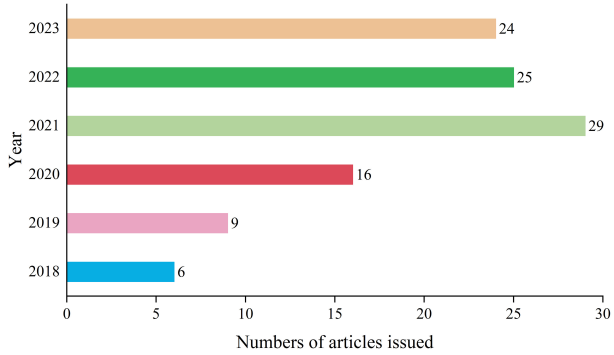


FIGURE 1
Annual analysis of the number of articles issued.

TABLE 1 Top 10 journals in terms of publication volume.

Journal Titles	JCR	Number	IF	Rate%
CANCERS	Q1	11	5.2	10.092
MODERN PATHOLOGY	Q1	6	7.5	5.505
EBIOMEDICINE	Q1	3	11.1	2.752
FRONTIERS IN ONCOLOGY	Q2	3	4.7	2.752
HISTOPATHOLOGY	Q1	3	6.4	2.752
IEEE ACCESS	Q2	3	3.9	2.752
MEDICAL IMAGE ANALYSIS	Q1	3	10.9	2.752
BIOINFORMATICS	Q1	2	5.8	1.835
COMPUTERS IN BIOLOGY AND MEDICINE	Q1	2	7.7	1.835
IEEE TRANSACTIONS ON MEDICAL IMAGING	Q1	2	10.6	1.835

circular node represents a country, with the size indicating the quantity of publications from that country. The connections between nodes represent collaborative relationships between countries, with the thickness of the connections reflecting the degree of collaboration. Different colors of nodes represent different time periods (8), the size of the purple circles reflects the centrality values indicating the influence of each country.

Leveraging the bibliometric online analysis platform, Figure 3 depicts the contributions of different countries in the field. Distinctly colored blocks represent the proportional contribution of each country. Table 2 presents the top 5 institutions in terms of publication output.

3.4 Visualization of author collaborations

Running the CiteSpace software, author analysis resulted in a knowledge graph with 200 nodes and 383 edges (Figure 4). Each circular node represents an author, and the connections between nodes represent collaborative relationships between authors. The thickness of the connections reflects the degree of collaboration. Different colors of nodes represent different time periods. Conducting a co-occurrence analysis on the author team collaboration network based on the literature retrieved from WoSCC, Table 3 is presented, listing the top 5 authors in terms of publication output along with their affiliated institutions in this research field.

3.5 Co-occurrence analysis of keywords

Keyword-related analysis, as manifested in the visualization of co-occurrence patterns, is crucial for delineating the research hotspots and frontiers within a given domain. Running the CiteSpace software with author keywords as node types, a co-occurrence network of keywords with 159 nodes and 334 edges

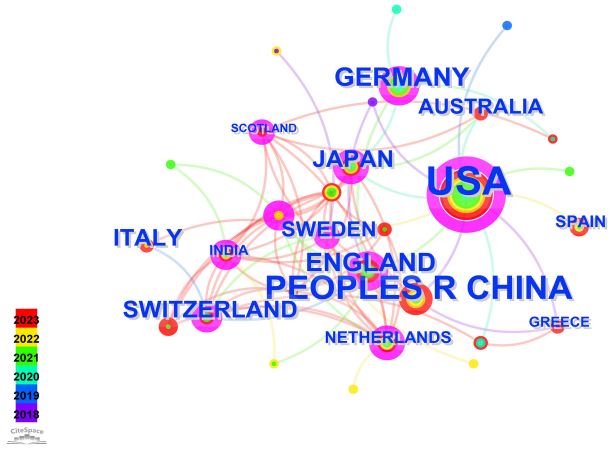


FIGURE 2
Visual map of countries.

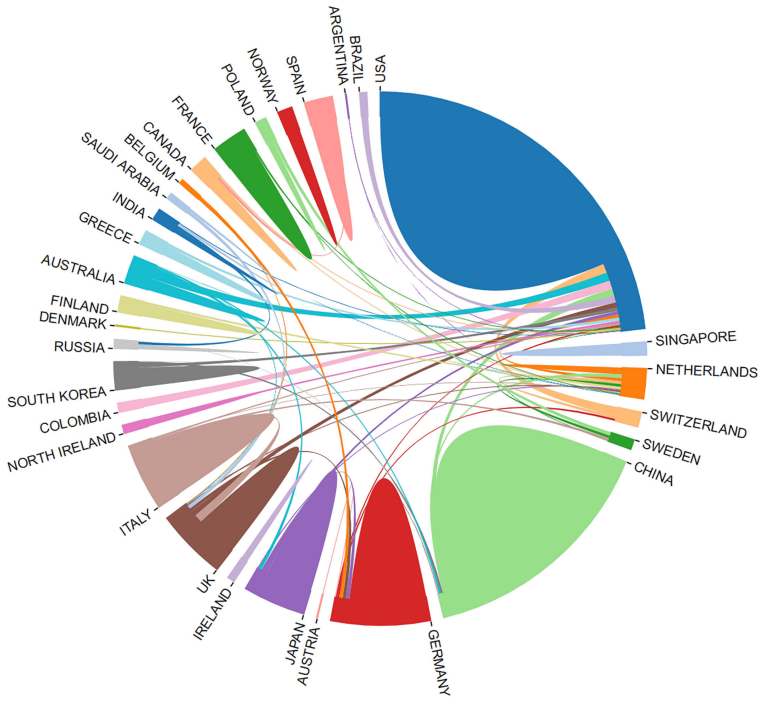


FIGURE 3
Proportion of national contribution.

was generated (Figure 5). After removing redundant terms that overlap with the search strategy, an analysis of the co-occurrence frequency and centrality values of keywords in this field (Table 4) reveals that the prominent keywords include: deep learning, artificial intelligence (AI), computer-aided diagnosis, tumor microenvironment, feature extraction, image analysis, tumor mutation burden, survival prediction, markov random field, mixture model. Furthermore, Figure 6 illustrates the temporal frequency changes of different keywords over time. It highlights the research focal points in the past few years related to the application of AI-based pathomics in the diagnosis and treatment of lung cancer. These themes reflect the proactive role of pathomics in aiding diagnosis, classification, predicting treatment efficacy, risk assessment, exploring emerging biomarkers, and analyzing gene expression levels in the context of lung cancer diagnosis and treatment.

TABLE 2 Top 5 institutions in terms of publication volume.

Rank	Number	Institution	Country
1	8	CASE WESTERN RESERVE UNIVERSITY	USA
2	8	UNIVERSITY OF TEXAS SYSTEM	USA
3	6	LOUIS STOKES CLEVELAND VETERANS AFFAIRS MEDICAL CENTER	USA
4	6	UNIVERSITY OF TEXAS SOUTHWESTERN MEDICAL CENTER DALLAS	USA
5	6	US DEPARTMENT OF VETERANS AFFAIRS	USA

3.6 Keyword cluster analysis

Keyword cluster analysis involves utilizing the log-likelihood rate (LLR) method to analyze the connection relationships among significant keyword nodes. This method reflects the hot topics within the research domain, with closely connected keywords in a cluster indicating higher research intensity. Larger node values within a cluster signify greater research interest. By examining these clusters, it is possible to predict the developmental patterns and emerging trends in the research field (9).

According to the keyword cluster analysis (Figure 7), researchers' studies are concentrated in the following 10 key areas: #0 parameter auto-tuning; #1 concordance study; #2 prognostic and predictive; #3 mixture model; #4 lung cancer slide cells; #5 non-small-cell lung cancer; #6 immunotherapy; #7 deep-learning microscopy; #8 telepathology; #9 radiology. By employing the clustering algorithm within CiteSpace software to organize title terms and visualize them (Figure 8), a clear sequential pattern emerges, encompassing: #0 spatial quantitative systems pharmacology platform spqsp-io; #1 adaptive radiotherapy; #2 patient survival; #3 pd-11 expression; #4 digital analysis; #5 Bayesian hidden Potts mixture model; #6 bayesian collaborative learning; #7 multi-stained feature matching; #8 oncology; #9 pathomics; Utilizing the clustering algorithm in CiteSpace to group subject categories and create a visual representation (Figure 9), a sequential progression of clusters is discernible,

FIGURE 4
Visual map of author network.

including: #0 Pathology; #1 Mathematics; #2 Medicine, Research & Experimental; #3 Computer Science, Theory & Methods; #4 Engineering, Multidisciplinary; #5 Statistics & Probability; #6 Imaging Science & Photographic Technology; #7 Biology; #8 Health Care Sciences Services; #9 Cell Biology. Employing the clustering algorithm for keywords and generating a graphical display (Figure 10), a sequential evolution of clusters is evident, incorporating: #0 digital pathology; #1 machine learning; #2 deep learning; #3 artificial intelligence; #4 lung cancer; # 5mixture model; #6 computational pathology; #7 scale invariant feature transform; #8 equity; #9 cancer immunopathology.

Each section is divided into 10 clustering modules, partial clustering blocks overlap with each other, suggesting close connections between these research areas. In addition to the research retrieval terms, other clusters demonstrate that pathomics in lung cancer research spans various fields, including medical experimental research, computer science, cell biology, statistics, and mathematics. Through advanced methods such as AI and machine learning, pathomics involves in-depth digital analysis of tumor tissues and the tumor microenvironment based on patients' pathological tissue sections. It aims to construct hybrid models, identify a multitude of pathological features, conduct precise assessments, and predict tumor-related indicators, including programmed death-ligand 1 Tumor cell Proportion Score (PD-L1 TPS). The goal is to assist in personalized diagnosis and treatment for patients and contribute to clinical decision-making by leveraging the synergies between AI and clinical medicine.

3.7 Cited references

A total of 426 relevant articles were retrieved from WoSCC, accumulating a total of 10,174 citations. The average number of citations per article is 24. The top 10 most cited articles are listed in Table 5.

TABLE 3 Top 5 authors in terms of publication volume.

Rank	Author	Institution	Country	Number
1	Xiao, Guanghua	University of Texas Southwestern Medical Center	USA	5
2	Xie, Yang	University of Texas Southwestern Medical Center Clinic Science	USA	4
3	Bera, Kaustav	CASE WESTERN RESERVE University	USA	4
4	Baxi, Vipul	BRISTOL MYERS SQUIBB	USA	3
5	Lu, Cheng	University of ALBERTA	Canada	3

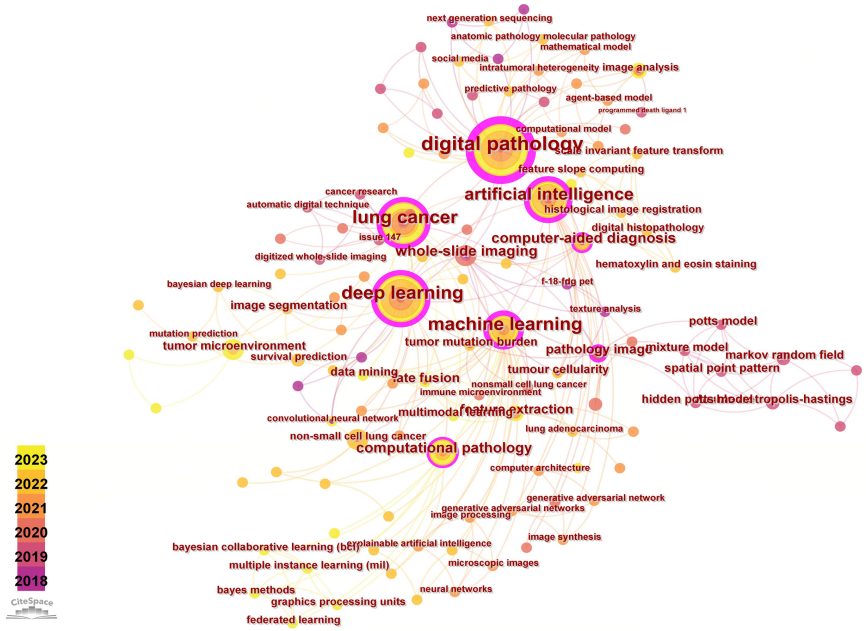


FIGURE 5
Visual map of author keywords.

4 Discussion

Pathomics is an innovative interdisciplinary field that combines digital pathology and AI. The rise of digital pathology has enabled the scanning of whole tissue slides, based on the fundamental principle of digitizing whole-slide images (WSI) using state-of-the-art whole-slide scanners. This technology can convert standard Hematoxylin-Eosin (H&E) staining glass slides into a digital format (WSI) (20). This allows for detailed spatial exploration of the entire tumor heterogeneity and its most invasive elements. It automatically extracts and classifies histological features, transforming this information into binary data. Finally, the extracted features are processed through sophisticated computer algorithms to perform tasks such as cancer classification and outcome prediction (21). Computational analysis of digitized histological slides through pathomics can extract valuable information. Some research primarily focuses on predicting the prognosis of lung cancer (22), including improving clinical decisions for cancer immunotherapy and exploring biomarkers related to potential benefits from ICIs, such as microsatellite instability (MSI), PD-L1 TPS, and inflammatory genes, among others (23). Another significant research area involves the integration of pathomics with multiple omics disciplines to explore the classification of lung cancer and other related aspects. Alvarez-Jimenez C et al. demonstrated the potential existence of cross-scale correlations between pathomics and CT imaging, which could be used to identify relevant imaging and histopathological features (24).

TABLE 4 High frequency and centrality keywords.

Rank	Keywords	Frequency	Centrality
1	deep learning	27	0.33
2	artificial intelligence	14	0.29
3	machine learning	10	0.22
4	computer-aided diagnosis	3	0.11
5	tumor microenvironment	4	0.07
6	feature extraction	3	0.04
7	image analysis	3	0.04
8	late fusion	2	0.02
9	tumor mutation burden	2	0.02
10	survival prediction	2	0.02

The escalating demand for personalized cancer treatment necessitates more precise biomarker assessments and quantitative tissue pathology for accurate cancer diagnosis. Pathologists must be equipped with new methodologies and tools to enhance diagnostic sensitivity and specificity, ultimately contributing to more informed and improved treatment decisions (13). Recently, significant success has been achieved in the analysis of medical images using AI due to the rapid advancement of “deep learning” algorithms (16).

Recent breakthroughs in AI hold the promise of significantly changing the way we diagnose and stratify cancer in pathology. Deep learning technology represents a milestone in this transformation, with numerous deep learning architectures applied to pathology-focused research. Various modeling objectives have been pursued, and recent studies demonstrate the application of deep learning in pathology aiming to predict

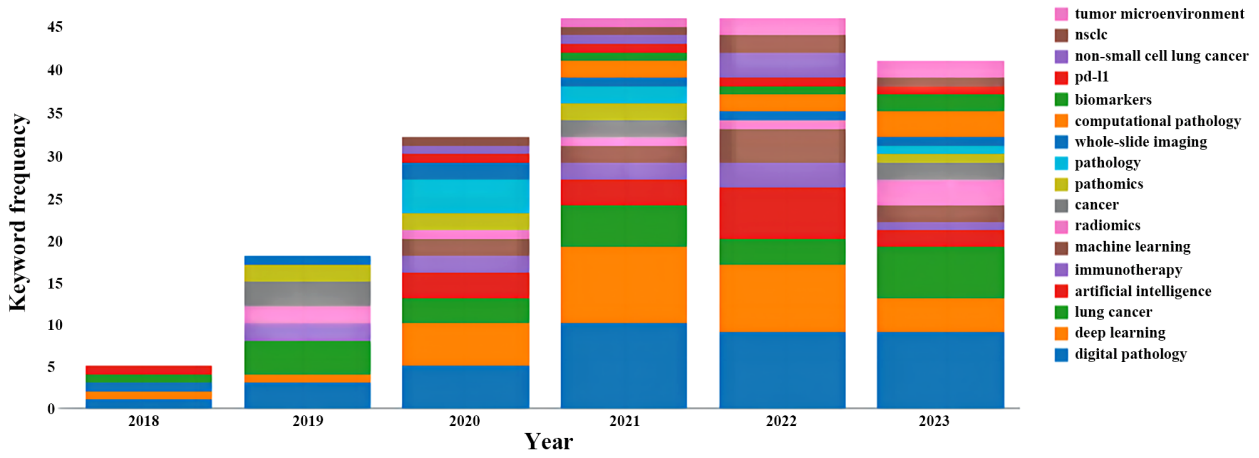


FIGURE 6
Variation in the number of keywords.

conventional diagnostic features used in pathology practice (such as distinguishing between diseases and normal tissues, defining tumor grades, and differentiating cancer types), leading to new insights into diseases (25, 26).

Deep learning encompasses various types of deep neural networks, and its application has achieved several breakthroughs in addressing current key challenges in pathology (27). Convolutional Neural Networks (CNN) are the most commonly used in the analysis of pathological images (28, 29). A standard CNN consists of an input layer, task-specific output layer, and multiple hidden layers. Each hidden layer is composed of numerous convolutional filters (parameters), which apply the same local transformation at different positions in their input images (30). Due to the shared parameters when applied locally in the image, effective parameterization of the CNN model is achieved. The typical implementation of CNN models offers a degree of translation invariance, allowing detected objects or patterns to appear at any position within the image. Pooling layers are often included between convolutional layers to down-sample the intermediate outputs (feature maps) of the convolution function. Following the convolutional layers are fully connected layers, which flatten the output of the convolutional layers and generate the final representation for the input-output layers (30, 31). Each neuron in a CNN calculates its output by applying a weight vector and bias (parameters) to the input values from the previous layer. The optimization (training) of the CNN model involves iteratively adjusting these biases and weights to minimize the loss function. One advantage of CNNs over other image classification algorithms is their suitability for end-to-end learning (32). Another major advantage of CNNs is their flexibility and efficiency in learning patterns from image data. Currently, they represent state-of-the-art technology in the field of image analysis and classification, consistently outperforming earlier generations of image analysis methods (29, 32). Kao Y-S et al. conducted a study on the

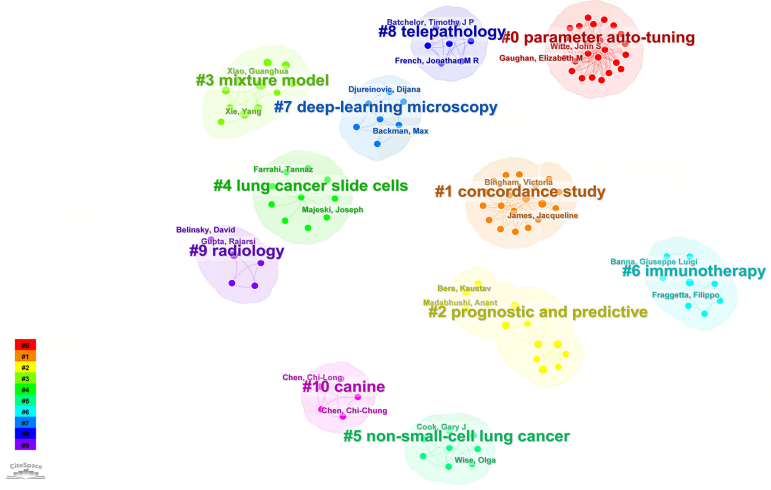


FIGURE 7
Visual map of author-generated keywords network.

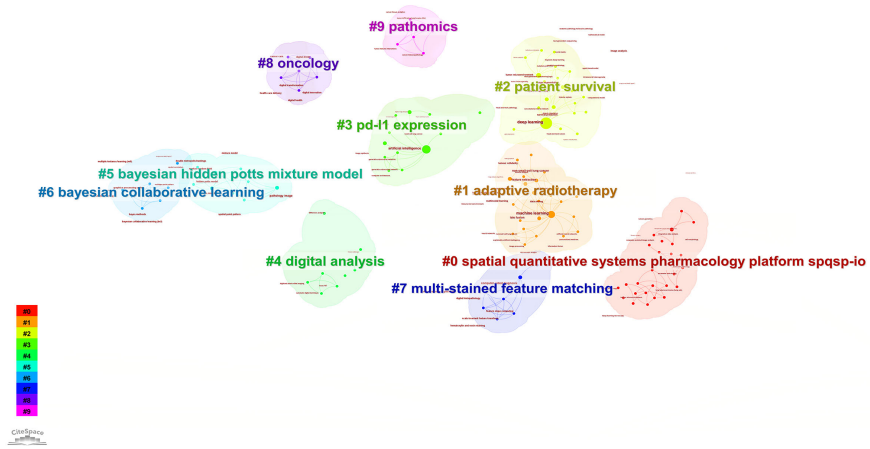


FIGURE 8
Visual map of title keywords network.

application of deep learning technology in histopathological tissue slices (deep pathomics) with the aim of predicting the response of stage III NSCLC to treatment (33). They assessed 35 digitalized tissue slices (biopsy or surgical specimens) from patients with stage IIIA or IIIB NSCLC. Based on the reduction in target volume observed in weekly CT scans during chemoradiotherapy, patients were categorized as responders (12/35, 34.7%) and non-responders (23/35, 65.7%). Employing a leave-two-out cross-validation method, they tested the digital tissue slices using 5 pre-trained CNNs-AlexNet, VGG, MobileNet, GoogLeNet, and ResNet, and evaluated the network performance. GoogLeNet was identified as the most effective CNN, accurately classifying 8/12 responders and 10/11 non-responders. Furthermore, deep pathomics exhibited a high level of specificity (True Negative Rate: 90.1) and considerable sensitivity (True Positive Rate: 0.75). Their data suggests that AI can surpass the capabilities of current diagnostic systems, providing additional insights beyond what is currently attainable in clinical practice.

Furthermore, there are studies attempting to apply AI to histological images with the aim of discovering novel image-based prognostic and predictive biomarkers. Cao R et al. proposed a deep learning model based on histopathological images to predict microsatellite status, achieving area under curve (AUC) of 0.88 and 0.85, respectively. It is noteworthy that this model can identify five distinct pathological imaging features, which are associated with the mutation burden in the genome, DNA damage repair-related genotypes, and the anti-tumor immune activation pathway in the transcriptome. The predictive model provides the potential for multi-omics correlations through interpretability associated with pathology, genomics, and transcriptomics phenotypes (34). Wang X et al. developed a system capable of identifying high-risk recurrence in early-stage NSCLC patients with an accuracy ranging from 75% to 82% (22). In another study, Wang S et al. characterized a group of high-risk NSCLC patients and identified image-based tumor shape features as an independent prognostic factor (35). Rakaee M et al. developed a machine learning-based scoring system

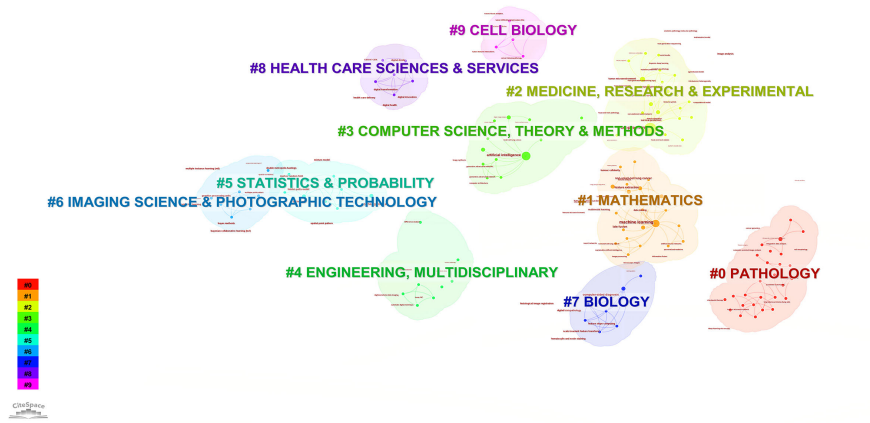


FIGURE 9
Visual map of subject categories keywords network.

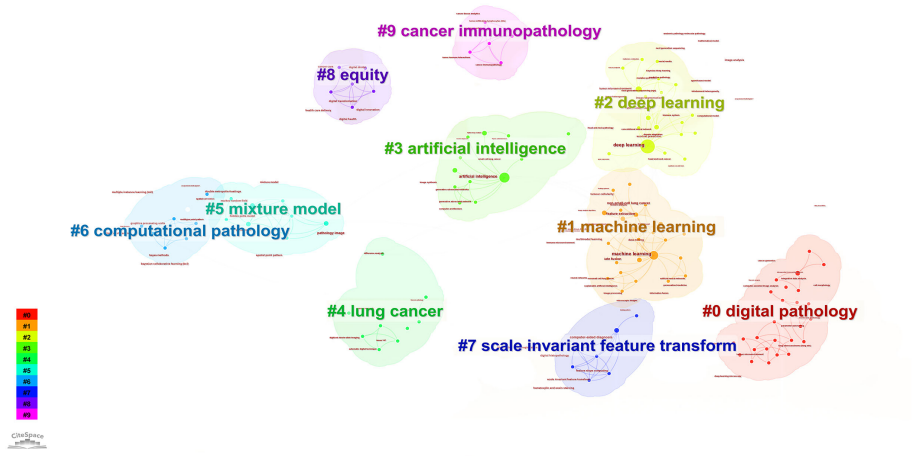


FIGURE 10
Visual map of keywords network.

for tumor-infiltrating lymphocytes (TILs) to predict the response of NSCLC to immune checkpoint inhibitor therapy (36). Additionally, Coudray N, Ocampo PS et al. applied AI to digital pathology slides to predict the presence of mutations in lung adenocarcinoma (37). In summary, the development of these advanced deep learning algorithms enhances the capability of analyzing lung cancer pathology images, assisting pathologists in challenging diagnostic tasks such as tumor identification, metastasis detection, and analysis of the tumor microenvironment.

TME is primarily composed of tumor cells, lymphocytes, stromal cells, macrophages, blood vessels, and other components. The composition of the TME varies based on the relative proportions of its different constituents, and its presence plays a crucial role in the growth and invasion of tumors.

Immune cells within the TME exhibit dual functions – on one hand, they identify and destroy tumor cells, while on the other hand, they also promote tumor growth and metastasis (38, 39). For instance, immune cells, including T cells, B cells, macrophages, and myeloid-derived suppressor cells, possess the ability to modulate the TME, thereby influencing tumor metastasis and pathological features (40, 41). Tumor Infiltrating Lymphocytes (TILs) in the TME involves a complex network of multiple cell types and cytokines and is a hallmark of immune recognition. Numerous studies have shown that activated $CD8^+$ T cells are the major players

TABLE 5 The top 10 cited articles.

Rank	Author	Year	Total Citations	Title
1	Pucci, Carlotta;	2019	343	Innovative approaches for cancer treatment: current perspectives and new challenges (10)
2	Lu, Ming Y;	2021	278	Data-efficient and weakly supervised computational pathology on whole-slide images (11)
3	Khosravi, Pegah;	2018	180	Deep Convolutional Neural Networks Enable Discrimination of Heterogeneous Digital Pathology Images (12)
4	Acs, B.;	2020	136	Artificial intelligence as the next step toward precision pathology (13)
5	Maibach, Fabienne;	2020	118	Tumor-Infiltrating Lymphocytes and Their Prognostic Value in Cutaneous Melanoma (14)
6	Mezheyeuski, Artur;	2018	114	Multispectral imaging for quantitative and compartment-specific immune infiltrates reveals distinct immune profiles that classify lung cancer patients (15)
7	Wang, Shidan;	2019	103	Artificial Intelligence in Lung Cancer Pathology Image Analysis (16)
8	Johnson, Douglas B.;	2018	91	Quantitative Spatial Profiling of PD-1/PD-L1 Interaction and HLA-DR/IDO-1 Predicts Improved Outcomes of Anti-PD-1 Therapies in Metastatic Melanoma (17)
9	Baxi, Vipul;	2022	80	Digital pathology and artificial intelligence in translational medicine and clinical practice (18)
10	Saw, Stephanie P. L.;	2021	56	Revisiting neoadjuvant therapy in non-small-cell lung cancer (19)

involved in anti-tumor immunity, and in a subset of tumors, cancer cells inhibit the activation of $CD8^+$ cytotoxic T cells through the expression of ligands such as PD-L1 that bind to inhibitory checkpoints, which has been suggested to be an important mechanism of immune escape for cancer cells (42). The expression of PD-L1 on TME immune cells, including myeloid cells (macrophages, dendritic cells) and T cells, appears to correlate more with the ICI response than expression on tumor cells. However, in NSCLC clinical practice, a limitation in histologically characterizing T lymphocyte infiltration is the scarcity of tumor tissue, which has hampered insight into the role of T lymphocytes in influencing the ICI response (43). Tumor-associated macrophages can promote angiogenesis and invasion by secreting cytokines, growth factors, and proteases (44). Cancer-associated fibroblasts (CAF) are pivotal in the formation of organs and the maintenance of tissue structure and function. They also play a significant role in tumor initiation, progression, metastasis, and the development of drug resistance through their potent immunosuppressive capabilities. Activated CAF possess the capability to secrete various substances, including extracellular matrix and vascular endothelial growth factor (VEGF), contributing to the complexity of the TME (45, 46). The markers associated with CAF are predominantly linked to T cell immunosuppression, inhibiting the functions of $CD8^+$ T cells and natural killer cells, particularly by secreting various chemokines and cytokines, notably interleukin-6 (IL-6), which leads to suboptimal clinical treatment outcomes. As research into the effects of CAF and the TME on immune cells and the efficacy of cancer immunotherapy advances, scientists can potentially develop novel compounds targeting these mechanisms, thereby offering innovative strategies for immunotherapy (47). It is noteworthy that research indicates a significant impact of the TME on the survival benefits of immunotherapy (48). The presence of immune cells in the TME, including the percentage of $CD8^+$ T cells, can serve as a predictive factor for the effectiveness of immunotherapy (49). The extracellular matrix can influence the mechanisms of tumorigenesis by affecting cell growth, metastasis, and immune evasion through the activation of signaling pathways. Additionally, tumor cells have the capability to release various growth factors, such as tumor growth factor, endothelial growth factor, and VEGF, contributing to the promotion of new blood vessel development (50). Angiogenesis is crucial for providing nutrients and oxygen to tumor cells, ultimately playing a critical role in tumor growth.

Therefore, TME plays a crucial role in tumor growth and metastasis. A comprehensive understanding of TME formation, investigating the interplay between immune cells and tumors, and exploring various genetic variations represent the future directions of TME research (51, 52). Additionally, selecting targeted therapeutic strategies based on TME subtypes can enhance the effectiveness of cancer treatment. To further emphasize this point, computer-assisted automatic detection of tumor cells in lymph nodes can significantly reduce the false-negative rate, thereby facilitating earlier detection and treatment of lung cancer, improving the accuracy of TNM staging, accelerating the examination process, and reducing the workload of pathologists. Moreover, tumor spread through air spaces (STAS) has been identified as an important clinical factor associated with tumor recurrence and poor prognosis in patient survival. The identification and quantification of STAS require experienced pathologists to perform detailed examinations of entire tissue sections. Therefore, pathological image analysis tools that rapidly and accurately identifies STAS would be useful for pathologists (16). Quantitative characterization of TME and accurate prediction and classification of important TME components are essential for targeted tumor therapy and prognosis assessment (53), necessitating advanced data processing and analysis approaches.

Quantitative characterization of TME involves a crucial step of segmenting different types of tissue substructures and cells from pathological images. This segmentation forms the foundation for various image analysis tasks, including cellular composition, spatial organization, and morphology specific to substructures. Previous studies in oncology primarily focused on tumor cells, overlooking the pivotal role of TME in the initiation and progression of cancer. The TME of lung cancer is primarily composed of tumor cells, lymphocytes, stromal cells, macrophages, blood vessels, and other components. Studies in lung cancer have indicated that TILs are positive prognostic factors, while angiogenesis is negatively associated with survival outcomes. The role of stromal cells in prognosis is complex. Traditional image processing methods encompass feature definition, feature extraction or segmentation. These techniques have been employed to segment lymphocytes and analyze the spatial organization of TILs and stromal cells within the TME (54). Research associated with the quantitative characterization of TME has the potential to predict treatment outcomes and provides insights for the development of targeted therapeutic strategies. Innovative studies in immunotherapy, in particular, heavily rely on understanding the interactions among various components within the TME and the mechanisms of immune evasion.

Accurate characterization of specific structures and features of TME is crucial for evaluating tumor prognosis (55), enhancing clinical decisions, and advancing precision medicine. Radiomics can unveil the heterogeneity of tumor cells and TME, while genomics and pathomics explore the biological significance of imaging histological features. The integration of these three approaches contributes to a comprehensive understanding and decoding of TME characteristics in tumors, facilitating prognostic predictions (56). The interconnection between radiomics, pathomics, and genomics contributes to establishing and deepening our understanding of cancer biology and imaging features. Concurrently, powerful machine learning techniques can decipher the complex interactions between tumors and cancer treatments. The integration of machine learning technologies with digital imaging and novel methods for assessing TME at the molecular level significantly enhances our comprehension of TME and cancer prognosis assessment. Vanguri RS et al. employed machine learning to integrate multimodal features into a risk prediction model (57). By combining radiological, histopathological, and genomic features, they assessed the predictive capability of immunotherapy response in NSCLC. Their study revealed that the AUC value of the multimodal model was 0.80, surpassing any single variable. These findings establish a quantitative foundation for enhancing

the accuracy of predicting immunotherapy response in NSCLC patients through the integration of multimodal features and machine learning.

Simultaneously, the quantitative characterization of TME in lung cancer poses certain challenges, including the following aspects: (1) Complexity and heterogeneity of lung cancer TME composition: In addition to the mentioned cell types, other structures such as bronchi, cartilage, and pleura often appear in pathological sections of the lung. This complexity and heterogeneity make segmentation and traditional feature definition challenging. (2) Cellular spatial organization (e.g., spatial distribution and interactions of different cell types): While playing a crucial role in TME, it is more challenging to capture than simply providing the quantity or ratio of different cell types. Current research mainly focuses on the proportion of different cell types, overlooking the intricate cellular spatial organization, which may result in limited and contradictory outcomes regarding the roles of different cell types in the TME. (3) For H&E-stained glass slides, there can be significant color variations based on staining conditions and the time gap between slide preparation and scanning. Traditional image processing methods based on manual feature extraction struggle to overcome these obstacles. (4) Multi-omics studies face the high dimensionality and heterogeneity of data, and integrating quantitative measurements of multi-modal data for prognosis prediction is a highly challenging task. In summary, pathomics, as a nascent research methodology, is presently undergoing preliminary investigation. Future studies utilizing extensive multi-omics datasets have the potential to advance the formulation of sophisticated integration strategies. These strategies would facilitate a more exhaustive evaluation, characterization, and elucidation of TME (58). Consequently, this advancement will yield profound insights into the imaging characteristics and the pathophysiological and biological underpinnings of tumor pathology.

In recent years, amidst the high incidence and mortality rates of lung cancer, the selection and implementation of treatment plans for advanced-stage lung cancer patients, as well as the creation of more precise platforms for predicting treatment responses, continue to face challenges. Pathomics not only synergizes with traditional pathological semantic information and clinical data to discover disease patterns but also interacts and integrates with various omics information, leveraging the unique advantages of each omics discipline. The development of these interdisciplinary approaches not only aids in identifying subtle lesions that may escape the naked eye and uncovering disease patterns beyond subjective judgment but also facilitates relatively objective and accurate assistance in disease screening, diagnosis, differential diagnosis, and prognosis assessment. Furthermore, it contributes to saving human and material resources, optimizing the utilization of limited medical resources to the maximum extent, and, on a broader scale, promoting the development of the personalized immune intervention.

5 Conclusion

In conclusion, this study systematically analyzed the literature on pathomics in the management of lung cancer indexed within the WoSCC. It offers an initial overview of recent research trends and forecasts potential hotspots and frontiers for future inquiry, aiming to provide valuable insights and references for scholars and researchers involved in personalized immunotherapy efficacy and prognosis for lung cancer.

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

LY: Validation, Visualization, Writing – original draft, Writing – review & editing. ZS: Validation, Writing – original draft, Writing – review & editing. YS: Data curation, Validation, Writing – original draft, Writing – review & editing. JZ: Data curation, Software, Validation, Visualization, Writing – review & editing. QW: Data curation, Software, Validation, Visualization, Writing – review & editing. YL: Data curation, Visualization, Writing – review & editing. HS: Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX23_3623) and National Natural Science Foundation of China (No. 82070020).

Acknowledgments

All authors acknowledge the contributions from the CiteSpace, the bibliometric online analysis platform and the WoSCC.

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.

Publisher's note

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. *CA: A Cancer J Clin*. (2021) 71:7–33. doi: 10.3322/caac.21654

2. 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

3. Brozos-Vazquez EM, Diaz-Pena R, Garcia-Gonzalez J, Leon-Mateos L, Mondelo-Macia P, Pena-Chilet M, et al. Immunotherapy in nonsmall-cell lung cancer: current status and future prospects for liquid biopsy. *Cancer Immunol Immunother*. (2021) 70:1177–88. doi: 10.1007/s00262-020-02752-z

4. Zhang Y, Yang Z, Chen R, Zhu Y, Liu L, Dong J, et al. Histopathology images-based deep learning prediction of prognosis and therapeutic response in small cell lung cancer. *NPJ Digital Med*. (2024) 7(1):15. doi: 10.1038/s41746-024-01003-0

5. Liu YH, Cheng X, Han XA, Cheng X, Jiang S, Lin YR, et al. Global research landscape and trends of lung cancer immunotherapy: A bibliometric analysis. *Front Immunol*. (2022) 13:1032747. doi: 10.3389/fimmu.2022.1032747

6. Dia AK, Ebrahimpour L, Yolchuyeva S, Tonneau M, Lamaze FC, Orain M, et al. The cross-Scale association between pathomics and radiomics features in immunotherapy-Treated NSCLC patients: a preliminary study. *Cancers*. (2024) 16 (2):348. doi: 10.3390/cancers16020348

7. Chen C. A glimpse of the first eight months of the COVID-19 literature on microsoft academic graph: themes, citation contexts, and uncertainties. *Front Res Metrics Analytics*. (2020) 5:607286. doi: 10.3389/frma.2020.607286

8. Miao L, Zhang J, Zhang Z, Wang S, Tang F, Teng M, et al. A bibliometric and knowledge-map analysis of CAR-T cells from 2009 to 2021. *Front Immunol*. (2022) 13:840956. doi: 10.3389/fimmu.2022.840956

9. Wang Q, Yang K-L, Zhang Z, Wang Z, Li C, Li L, et al. Characterization of global research trends and prospects on single-cell sequencing technology: bibliometric analysis. *J Med Internet Res*. (2021) 23(8):e25789. doi: 10.2196/25789

10. Pucci C, Martinelli C, Ciofani G. Innovative approaches for cancer treatment: current perspectives and new challenges. *Ecancermedicalscience*. (2019) 13:961. doi: 10.3332/ecancer.2019.961

11. Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng*. (2021) 5(6):555–70. doi: 10.1038/s41551-020-00682-w

12. Khosravi P, Kazemi E, Imielinski M, Elemento O, Hajirasouliha I. Deep convolutional neural networks enable discrimination of heterogeneous digital pathology images. *EBioMedicine*. (2018) 27:317–28. doi: 10.1016/j.ebiom.2017.12.026

13. Acs B, Rantalainen M, Hartman J. Artificial intelligence as the next step towards precision pathology. *J Internal Med*. (2020) 288:62–81. doi: 10.1111/joim.13030

14. Maibach F, Sadozai H, Seyed Jafari SM, Hunger RE, Schenk M. Tumor-infiltrating lymphocytes and their prognostic value in cutaneous melanoma. *Front Immunol*. (2020) 11:2105. doi: 10.3389/fimmu.2020.02105

15. Mezheyeuski A, Bergsland CH, Backman M, Djureinovic D, Sjöblom T, Bruun J, et al. Multispectral imaging for quantitative and compartmentelkiingnc immune infiltrates reveals distinct immune profiles that classify lung cancer patients. *J Pathol*. (2018) 244:421–31. doi: 10.1002/path.5026

16. Wang S, Yang DM, Rong R, Zhan X, Fujimoto J, Liu H, et al. Artificial intelligence in lung cancer pathology image analysis. *Cancers*. (2019) 11(11):1673. doi: 10.3390/cancers11111673

17. Johnson DB, Bordeaux J, Kim JY, Vaupel C, Rimm DL, Ho TH, et al. Quantitative spatial profiling of PD-1/PD-L1 interaction and HLA-DR/IDO-1 predicts improved outcomes of anti–PD-1 therapies in metastatic melanoma. *Clin Cancer Res*. (2018) 24:5250–60. doi: 10.1158/1078-0432.CCR-18-0309

18. Baxi V, Edwards R, Montalto M, Saha S. Digital pathology and artificial intelligence in translational medicine and clinical practice. *Modern Pathol*. (2022) 35:23–32. doi: 10.1038/s41379-021-00919-2

19. Saw SPL, Ong B-H, Chua KLM, Takano A, Tan DSW. Revisiting neoadjuvant therapy in non-small-cell lung cancer. *Lancet Oncol*. (2021) 22:e501–16. doi: 10.1016/S1470-2045(21)00383-1

20. Hanna MG, Parwani A, Sirintrapun SJ. Whole slide imaging: technology and applications. *Adv Anatomic Pathol*. (2020) 27:251–9. doi: 10.1097/PAP.0000000000000273

21. Classe M, Lerousseau M, Scoazec J-Y, Deutsch E. Perspectives in pathomics in head and neck cancer. *Curr Opin Oncol*. (2021) 33:175–83. doi: 10.1097/CCO.0000000000000731

22. Wang X, Janowczyk A, Zhou Y, Thawani R, Fu P, Schalper K, et al. Prediction of recurrence in early stage non-small cell lung cancer using computer extracted nuclear features from digital H&E images. *Sci Rep*. (2017) 7(1):13543. doi: 10.1038/s41598-017-13773-7

23. Prelaj A, Miskovic V, Zanitti M, Trovo F, Genova C, Viscardi G, et al. Artificial intelligence for predictive biomarker discovery in immuno-oncology: a systematic review. *Ann Oncol*. (2024) 35:29–65. doi: 10.1016/j.annonc.2023.10.125

24. Alvarez-Jimenez C, Sandino AA, Prasanna P, Gupta A, Viswanath SE, Romero E. Identifying cross-scale associations between radiomic and pathomic signatures of non-small cell lung cancer subtypes: preliminary results. *Cancers*. (2020) 12(12):3663. doi: 10.3390/cancers12123663

25. Hosseini MS, Bejnordi BE, Trinh VQ, Chan L, Hasan D, Li X, et al. et al: Computational pathology: A survey review and the way forward. *J Pathol Inform*. (2024) 15:100357. doi: 10.1016/j.jpi.2023.100357

26. Jin W, Yang Q, Chi H, Wei K, Zhang P, Zhao G, et al. Ensemble deep learning enhanced with self-attention for predicting immunotherapeutic responses to cancers. *Front Immunol*. (2022) 13:1025330. doi: 10.3389/fimmu.2022.1025330

27. Burrows L, Sculthorpe D, Zhang H, Rehman O, Mukherjee A, Chen K. Mathematical modelling and deep learning algorithms to automate assessment of single and digitally multiplexed immunohistochemical stains in tumoural stroma. *J Pathol Inform*. (2024) 15:100351. doi: 10.1016/j.jpi.2023.100351

28. LeCun Y, Kavukcuoglu K, Farabet C. Convolutional networks and applications in vision. *Proc 2010 IEEE Int Symposium Circuits Syst*. (2010), 253–6. doi: 10.1109/ISCAS.2010.5537907

29. Chang HY, Jung CK, Woo JI, Lee S, Cho J, Kim SW, et al. Artificial intelligence in pathology. *J Pathol Trans Med*. (2019) 53:1–12. doi: 10.4132/jptm.2018.12.16

30. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *J Pathol Inf*. (2016) 7:29. doi: 10.4103/2153-3539.186902

31. Robertson S, Azizpour H, Smith K, Hartman J. Digital image analysis in breast pathology—from image processing techniques to artificial intelligence. *Trans Res*. (2018) 194:19–35. doi: 10.1016/j.trsl.2017.10.010

32. Chaunzwa TL, Hosny A, Xu Y, Shafer A, Diao N, Lanuti M, et al. Deep learning classification of lung cancer histology using CT images. *Sci Rep*. (2021) 11(1):5471. doi: 10.1038/s41598-021-84630-x

33. Nibid L, Greco C, Cordelli E, Sabarese G, Fiore M, Liu CZ, et al. Deep pathomics: a new image-based tool for predicting response to treatment in stage III non-small cell lung cancer. *PLoS One*. (2023) 18(11):e0294259. doi: 10.1371/journal.pone.0294259

34. Cao R, Yang F, Ma S-C, Liu L, Zhao Y, Li Y, et al. et al: Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in Colorectal Cancer. *Theranostics*. (2020) 10:11080–91. doi: 10.7150/thno.49864

35. Wang S, Chen A, Yang L, Cai L, Xie Y, Fujimoto J, et al. Comprehensive analysis of lung cancer pathology images to discover tumor shape and boundary features that predict survival outcome. *Sci Rep*. (2018) 8(1):10393. doi: 10.1038/s41598-018-27707-4

36. Rakaee M, Adib E, Ricciuti B, Sholl LM, Shi W, Alessi JV, et al. Association of machine learning–based assessment of tumor-infiltrating lymphocytes on standard histologic images with outcomes of immunotherapy in patients with NSCLC. *JAMA Oncol*. (2023) 9(1):51–60. doi: 10.1001/jamaoncol.2022.4933

37. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, et al. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nat Med*. (2018) 24:1559–67. doi: 10.1038/s41591-018-0177-5

38. Deepak KGK, Vempati R, Nagaraju GP, Dasari VR, S N, Rao DN, et al. : Tumor microenvironment: Challenges and opportunities in targeting metastasis of triple negative breast cancer. *Pharmacol Res*. (2020) 153:104683. doi: 10.1016/j.phrs.2020.104683

39. Li L, Yu R, Cai T, Chen Z, Lan M, Zou T, et al. Effects of immune cells and cytokines on inflammation and immunosuppression in the tumor microenvironment. *Int Immunopharmacol*. (2020) 88:106939. doi: 10.1016/j.intimp.2020.106939

40. Xia Z, Chen S, He M, Li B, Deng Y, Yi L, et al. Editorial: Targeting metabolism to activate T cells and enhance the efficacy of checkpoint blockade immunotherapy in solid tumors. *Front Immunol*. (2023) 14:1247178. doi: 10.3389/fimmu.2023.1247178

41. Xiong J, Chi H, Yang G, Zhao S, Zhang J, Tran LJ, et al. Revolutionizing anti-tumor therapy: unleashing the potential of B cell-derived exosomes. *Front Immunol*. (2023) 14:1188760. doi: 10.3389/fimmu.2023.1188760

42. Fan J, Zhang T. Multi-index comprehensive evaluation of the efficacy and response mechanism of immunotherapy in non-small cell lung cancer. *Clin Transl Oncol*. (2024). doi: 10.1007/s12094-024-03519-z

43. Lei X, Lei Y, Li JK, Du WX, Li RG, Yang J, et al. Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy. *Cancer Lett*. (2020) 470:126–33. doi: 10.1016/j.canlet.2019.11.009

44. Vitale I, Manic G, Coussens LM, Kroemer G, Galluzzi L. Macrophages and metabolism in the tumor microenvironment. *Cell Metab*. (2019) 30:36–50. doi: 10.1016/j.cmet.2019.06.001

45. Mao X, Xu J, Wang W, Liang C, Hua J, Liu J, et al. Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: new findings and future perspectives. *Mol Cancer*. (2021) 20(1):131. doi: 10.1186/s12943-021-01428-1

46. Ren Q, Zhang P, Lin H, Feng Y, Chi H, Zhang X, et al. A novel signature predicts prognosis and immunotherapy in lung adenocarcinoma based on cancer-associated fibroblasts. *Front Immunol*. (2023) 14:1201573. doi: 10.3389/fimmu.2023.1201573

47. Akai M, Noma K, Kato T, Nishimura S, Matsumoto H, Kawasaki K, et al. Fibroblast activation protein-targeted near-infrared photoimmunotherapy depletes immunosuppressive cancer-associated fibroblasts and remodels local tumor immunity. *Br J Cancer*. (2024) 130:1647–58. doi: 10.1038/s41416-024-02639-1

48. Pitt JM, Marabelle A, Eggermont A, Soria JC, Kroemer G, Zitvogel L. Targeting the tumor microenvironment: removing obstruction to anticancer immune responses and immunotherapy. *Ann Oncol*. (2016) 27:1482–92. doi: 10.1093/annonc/mdw168

49. Sun D, Liu J, Zhou H, Shi M, Sun J, Zhao S, et al. Classification of tumor immune microenvironment according to programmed death-Ligand 1 expression and immune infiltration predicts response to immunotherapy plus chemotherapy in advanced patients with NSCLC. *J Thorac Oncol*. (2023) 18:869–81. doi: 10.1016/j.jtho.2023.03.012

50. de Visser KE, Joyce JA. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. *Cancer Cell*. (2023) 41:374–403. doi: 10.1016/j.ccell.2023.02.016

51. Fang X, Li D, Wan S, Hu J, Zhang P, Jie D, et al. Insights into the heterogeneity of the tumor microenvironment in lung adenocarcinoma and squamous carcinoma through single-cell transcriptomic analysis: implications for distinct immunotherapy outcomes. *J Gene Med*. (2024) 26:e3694. doi: 10.1002/jgm.3694

52. Chen X, Li R, Yin YH, Liu X, Zhou XJ, Qu YQ. Pan-cancer prognosis, immune infiltration, and drug resistance characterization of lung squamous cell carcinoma tumor microenvironment-related genes. *Biochem Biophys Rep*. (2024) 38:101722. doi: 10.1016/j.bbrep.2024.101722

53. Fukuda H, Arai K, Mizuno H, Nishito Y, Motoi N, Arai Y, et al. Molecular subtypes of lung adenocarcinoma present distinct immune tumor microenvironments. *Cancer Sci*. (2024) 115:1763–77. doi: 10.1111/cas.16154

54. Ichikawa T, Aokage K, Sugano M, Miyoshi T, Kojima M, Fujii S, et al. The ratio of cancer cells to stroma within the invasive area is a histologic prognostic parameter of lung adenocarcinoma. *Lung Cancer*. (2018) 118:30–5. doi: 10.1016/j.lungcan.2018.01.023

55. Huang M, Cha Z, Liu R, Lin M, Gafoor NA, Kong T, et al. Enhancing immunotherapy outcomes by targeted remodeling of the tumor microenvironment *via* combined cGAS-STING pathway strategies. *Front Immunol*. (2024) 15:1399926. doi: 10.3389/fimmu.2024.1399926

56. Kang W, Qiu X, Luo Y, Luo J, Liu Y, Xi J, et al. Application of radiomics-based multiomics combinations in the tumor microenvironment and cancer prognosis. *J Transl Med*. (2023) 21:598. doi: 10.1186/s12967-023-04437-4

57. Vanguri RS, Luo J, Aukerman AT, Egger JV, Fong CJ, Horvat N, et al. Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer. *Nat Cancer*. (2022) 3:1151–64. doi: 10.1038/s43018-022-00416-8

58. Lucena-Sánchez E, Hicke FJ, Clara-Trujillo S, García-Fernández A, Martínez-Máñez R. Nanoparticle-mediated tumor microenvironment remodeling favors the communication with the immune cells for tumor killing. *Advanced Ther*. (2024) 7(5). doi: 10.1002/adtp.202400004

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 31 July 2024
DOI 10.3389/fimmu.2024.1437068

OPEN ACCESS

EDITED BY
Wenyi Jin,
University of Hong Kong, Hong Kong SAR, China

REVIEWED BY
Hang Song,
Anhui University of Chinese Medicine, China
Jianzhou Cui,
National University of Singapore, Singapore
Zhijie Xu,
Central South University, China

*CORRESPONDENCE
Shiyi Chen
shiyichen_huashan@163.com
Shuo Chen
drchenshuo@136.com
Qing Wang
Doctorwq1983@163.com

†These authors have contributed equally to this work and share the first authorship

RECEIVED 23 May 2024
ACCEPTED 12 July 2024
PUBLISHED 31 July 2024

CITATION
Luo Z, Zhu J, Xu R, Wan R, He Y, Chen Y, Wang Q, Chen S and Chen S (2024) Exercise-downregulated CD300E acted as a negative prognostic implication and tumor-promoted role in pan-cancer.
Front. Immunol. 15:1437068.
doi: 10.3389/fimmu.2024.1437068

Exercise-downregulated CD300E acted as a negative prognostic implication and tumor-promoted role in pan-cancer

Zhiwen Luo[1†], Jinguo Zhu[2†], Rui Xu[3†], Renwen Wan[1], Yanwei He[1], Yisheng Chen[1], Qing Wang[4*], Shuo Chen[5*] and Shiyi Chen[1*]

[1]Department of Sports Medicine, Huashan Hospital, Fudan University, Shanghai, China, [2]Department of Orthopaedics, Nantong Tongzhou Hospital of Traditional Chinese Medicine, Tongzhou, Jiangsu, China, [3]The First Clinical Medicine College, Nanjing Medical University, Nanjing, China, [4]Department of Orthopaedics, Kunshan Hospital of Chinese Medicine, Kunshan, Jiangsu, China, [5]Department of Sports Medicine, Nanjing Hospital of Chinese Medicine Affiliated to Nanjing University of Chinese Medicine, Nanjing, China

Background: Breast cancer ranks as one of the most prevalent malignancies among women globally, with increasing incidence rates. Physical activity, particularly exercise, has emerged as a potentially significant modifier of cancer prognosis, influencing tumor biology and patient outcomes.

Methods: Using a murine breast cancer model, we established a control and an exercise group, where the latter was subjected to 21 days of voluntary running. RNA Sequencing, bioinformatics analysis, pan-cancer analysis, and cell experiments were performed to validate the underlying mechanisms.

Results: We observed that exercise significantly reduced tumor size and weight, without notable changes in body weight, suggesting that physical activity can modulate tumor dynamics. mRNA sequencing post-exercise revealed substantial downregulation of CD300E in the exercise group, accompanied by alterations in critical pathways such as MicroRNAs in cancers and the Calcium signaling pathway. Expanding our analysis to a broader cancer spectrum, CD300E demonstrated significant expression variability across multiple cancer types, with pronounced upregulation in myeloma, ovarian, lung, and colorectal cancers. This upregulation was correlated with poorer prognostic outcomes, emphasizing CD300E's potential role as a prognostic marker and therapeutic target. Moreover, CD300E expression was associated with cancer cell proliferation and apoptosis.

Conclusion: The study highlights the dual role of exercise in modulating gene expression relevant to tumor growth and the potential of CD300E as a target in cancer therapeutics. Further research is encouraged to explore the mechanisms by which exercise and CD300E influence cancer progression and to develop targeted strategies that could enhance patient outcomes in clinical settings.

KEYWORDS

physical exercise, breast cancer, CD300E, pan-cancer, bioinformatics, proliferation

1 Introduction

Breast cancer, a predominant malignancy among women, has witnessed an increasing global incidence (1, 2). The World Health Organization reports that it stands as one of the leading causes of cancer-related deaths among women worldwide (3, 4). The impacts of breast cancer extend beyond severe health threats; its cells invade surrounding tissues and metastasize via lymphatic and circulatory systems to distant organs such as bones, liver, lungs, and brain, complicating and escalating the complexity of treatment protocols (5, 6). Additionally, the socioeconomic repercussions are profound, imposing substantial financial burdens during treatment and straining familial and social relationships due to the psychological toll of the disease (7, 8). Therefore, deepening our understanding of the mechanisms underlying breast cancer pathogenesis and developing innovative targeted therapies are imperative (9–11).

The beneficial impacts of physical activity on health and cancer prevention are multifaceted (12). Exercise enhances cardiovascular efficiency and muscle strength, augments bone density, and aids in osteoporosis prevention (13). It also boosts metabolism, which helps maintain a healthy weight and physique. Immunologically, physical activity increases lymphocyte counts, thereby strengthening the immune system's defense against diseases, including cancer (14). Exercise also alleviates psychological stress and mitigates symptoms of anxiety and depression, enhancing overall mood and well-being, thereby indirectly reducing cancer risk (15–17). Persistently engaging in physical activities has been shown to correlate with lower cancer incidence rates, likely due to enhanced antioxidative capacity and expedited elimination of carcinogens (18–20). Recent research further underscores the therapeutic potentials of exercise in oncology. A study by Luo et al. revealed that physical activity could transform the immunological microenvironment of non-small cell lung cancer from a "cold" to a "hot" state, indicating that exercise not only increases the population of CD8+ T cells and M1 macrophages but also reduces immunosuppressive cells, thereby sensitizing tumors to immunotherapy (21). This transformative potential of exercise offers a promising adjunct to conventional cancer treatments, suggesting that integrating physical activity could significantly enhance therapeutic outcomes.

The CD300E gene encodes a protein that interacts with the TYRO protein tyrosine kinase binding protein, and is considered an activating receptor (22). Within the immune system, CD300E is posited to play a pivotal role in modulating the activity and functionality of immune cells (23–25). Research indicates that mCD300E can recognize sphingomyelin, thereby regulating the functions of atypical and intermediate monocytes through FcRγ and DAP12 (26). In the realm of oncology, the study of CD300E is garnering increasing attention due to its potential role in modulating tumor immune responses and facilitating immune escape (25, 27). Specifically, CD300E may promote tumor growth and dissemination by influencing the interactions between tumor cells and the immune system. In addition, one patent have reported that CD300E siRNA delays or halts cancer progression by blocking or knocking down cd300e to inhibit its activity or expression, and that the rate of tumor growth is significantly inhibited in mouse tumors compared to controls. Understanding the precise mechanisms of CD300E's involvement in tumor immunity is critical for the development of novel immunotherapeutic strategies, which could include modulating its expression or function to enhance the immune system's capacity to target tumors (28, 29).

This study has identified CD300E as a critical target through gene sequencing of voluntary running wheel exercises in mice as an anti-breast cancer initiative. By further analyzing CD300E through bioinformatics and cellular biology experiments, we aim to explore and demonstrate its role in tumor development and progression. This research not only sheds light on the mechanistic underpinnings of CD300E in cancer biology but also underscores the potential of exercise-induced molecular responses as a strategic approach in cancer prevention and treatment.

2 Materials and methods

2.1 Cell culture

The 4T1 mouse cancer cell line (catalog KGG2224-1) and MDAMB231 (catalog KGG3220-1) were procured from KeyGEN (Nanjing, China). MDA-MB-468 was procured from FengHui ShengWu, China. 4T1 cells were cultured in RPMI-1640 medium enriched with 10% fetal bovine serum (FBS) and sustained at 37°C in either an ambient atmosphere or one containing 5% CO_2. MDAMB231 and MDAMB468 cells were cultured in the MEM media with 1% non-essential amino acid and 1 mM sodium pyruvate. All media were added with 10% FBS at 37°C with or without 5% CO_2.

2.2 Animal interventions

Female BALB/c mice, aged 5-6 weeks, were obtained from the Shanghai Laboratory Animal Center (SLAC). To establish a triple-negative breast cancer (TNBC) model, 4T1 cells (5 × 10^6) were subcutaneously injected into the abdomen of BALB/c mice. The choice of this specific strain and demographic was based on its relevance to breast cancer research and its consistent response to exercise interventions. All mice were in good health, verified by a veterinarian prior to the commencement of the study. The mice were housed in a controlled environment with a 12-hour light/dark cycle, and were given free access to food and water. Tumor growth was monitored and measured regularly every 2-3 days using calipers. Mice were randomly divided into two groups: an exercise group (E) and a non-exercise group (NE), each comprising five animals. The exercise group underwent a 21-day regimen of voluntary running (no speed or distance limitation), whereas the non-exercise group was maintained under normal husbandry conditions without dietary restrictions. After 21 days, the mice were euthanized, and tumor tissues were collected for mRNA sequencing analysis. Animal experiments were granted by Ethics Committees at Nanjing Medical University.

2.3 mRNA sequencing and bioinformatics analysis

21 days subsequent to administering the treatments, tumor samples from mice were carefully collected for mRNA sequencing analysis (30). Following various treatments, cell samples were diligently harvested. The extraction of total RNA from these samples was performed using the highly regarded RNeasy Mini Kit (Qiagen, Hilden, Germany). After RNA extraction, the construction of paired-end libraries was carried out using the TruSeq RNA Sample Preparation Kit (Illumina, USA), adhering meticulously to the protocol provided by TruSeq RNA Sample Preparation. The Shanghai Biotechnology Corporation was tasked with the responsibility of constructing and sequencing the libraries. For the precise mapping of clean reads to the Rnor 6.0 reference genome, allowing up to two mismatches, the widely acclaimed Hisat2 software (version 2.0) was utilized. Subsequent to genome mapping, the esteemed Stringtie software (version 1.3.0) was employed to generate and annotate Fragments per kilobase of exon per million (FPKM) values. Gene expression data were normalized using the trimmed mean of M-values (TMM) method to correct for library size differences and compositional biases. Top-10 genes were shown.

Statistical significance was determined with a P-value threshold set according to the false discovery rate (FDR). mRNAs exhibiting a fold change of ≥ 2 and an FDR ≤ 0.05 were identified as differentially expressed. To further investigate the biological pathways involved, meticulous KEGG pathway analysis was performed using the revered KEGG database (http://www.genome.ad.jp/kegg) within the R environment. Additionally, Gene Set Enrichment Analysis (GSEA) was conducted using R BiocManager to delve deeper into the molecular mechanisms influenced by the treatments.

2.4 Pan-cancer analysis

2.4.1 Gene expression and datasets obtained

We utilized the Human Protein Atlas (HPA) to collate comprehensive RNA and protein expression profiles of CD300E in human samples. Furthermore, detailed data on CD300E expression across various tissues and cell lines were sourced from the Harmonizome database. We expanded our dataset by incorporating CD300E mRNA expression data from cancerous, paracancerous, and normal tissue samples provided by TCGA and GTEx databases. Our study spanned a diverse set of 33 cancer types including, Adrenocortical carcinoma (ACC), Bladder Urothelial Carcinoma (BLCA), Breast invasive carcinoma (BRCA), Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), Cholangiocarcinoma (CHOL), Colon adenocarcinoma (COAD), Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (DLBC), Esophageal carcinoma (ESCA), Glioblastoma multiforme (GBM), Head and Neck squamous cell carcinoma (HNSC), Kidney Chromophobe (KICH), Kidney renal clear cell carcinoma (KIRC), Kidney renal papillary cell carcinoma (KIRP), Acute Myeloid Leukemia (LAML), Brain Lower Grade Glioma (LGG), Liver hepatocellular carcinoma (LIHC), Lung adenocarcinoma (LUAD), Lung squamous cell carcinoma (LUSC), Mesothelioma (MESO), Ovarian serous cystadenocarcinoma (OV), Pancreatic adenocarcinoma (PAAD), Pheochromocytoma and Paraganglioma (PCPG), Prostate adenocarcinoma (PRAD), Rectum adenocarcinoma (READ), Sarcoma (SARC), Skin Cutaneous Melanoma (SKCM), Stomach adenocarcinoma (STAD), Testicular Germ Cell Tumors (TGCT), Thyroid carcinoma (THCA), Thymoma (THYM), Uterine Corpus Endometrial Carcinoma (UCEC), Uterine Carcinosarcoma (UCS), Uveal Melanoma (UVM).

For statistical analysis, we utilized R software (version 4.2.2) and employed the ggplot2 package to depict CD300E expression across the cancer spectrum. We adopted the median expression level as the threshold for differential expression analyses. Differences between expression groups were assessed using the Wilcoxon rank-sum test.

2.4.2 Survival analysis of CD300E in the 33 cancers

We also conducted survival analyses to explore the prognostic potential of CD300E expression in cancer (18). Using the survival package in R, we performed Kaplan-Meier analyses and employed Cox regression to compare survival outcomes between groups with high and low expression of CD300E. The impact of CD300E expression on survival was visually represented through forest plots using the survminer and ggplot2 packages.

2.4.3 Genetic alteration analysis of CD300E

An investigation into the genetic alterations associated with CD300E was conducted through the cBioPortal. This analysis included an examination of somatic mutation frequencies and detailed genomic information, helping to elucidate the mutation landscape of CD300E in various cancers.

2.4.4 Immunogenomic analyses of CD300E in the 33 cancers

In our immunogenomic analysis across 33 different cancers, we utilized the "GSVA" package and the "ssGSEA" algorithm to assess the relationship between CD300E expression and various immune components, including tumor-infiltrating lymphocytes, immunostimulators, immunoinhibitors, MHC molecules, chemokines, and chemokine receptors. The correlations were determined using Spearman's correlation coefficient, and p-values less than 0.05 were deemed significant. To effectively display these correlations, we generated heatmaps using the "ggplot2" package.

2.4.5 Functional enrichment analysis of CD300E

We carried out Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses to examine the functions and pathways associated with genes interacting closely with CD300E. These genes were identified through STRING and analyzed using the "clusterProfiler" and "org.Hs.eg.db" packages in R. We set a stringent cutoff threshold of a p-value < 0.01 for both GO and KEGG enrichment analyses.

The outcomes of these analyses were visually represented using bubble charts created with the "ggplot2" package.

2.5 Cellular experiments

2.5.1 Knockdown of CD300E gene

To knock down CD300E gene expression in tumor cells, we designed and synthesized small interfering RNAs (siRNAs) targeting CD300E using *In vivo*gen-based method (Detailed sequencing can be found in the Supplementary Table 1). These siRNA sequences were algorithmically predicted and selected as the most likely to effectively target CD300E mRNA. The specific steps are as follows: siRNA transfection: siRNA transfection was performed using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Briefly, cells were incubated with a mixture of Lipofectamine 2000 and siRNA to form a complex 24 hours after inoculation and then added to the cells. Gene knockdown efficiency assessment: 48 hours after transfection, CD300E mRNA and protein levels were analyzed by real-time quantitative PCR (qPCR) to verify the efficiency of siRNA knockdown.

2.5.2 Overexpression of CD300E gene

cDNAs of mouse CD300e (GenBankTM accession number NM_172050.3) were isolated by PCR from a cDNA library of mouse BM cells. To overexpress CD300E, we constructed a plasmid containing the complete CD300E coding region. This plasmid drives the expression of CD300E under the control of CMV promoter. The steps of the overexpression experiment are as follows:

Plasmid construction: the cDNA of CD300E was cloned into the expression vector pCMV, and the correctness of the insert sequence was verified by gene sequencing. Plasmid transfection: transfection of plasmid DNA was performed using Lipofectamine 2000. Cells were transfected 24 hours after inoculation, following similar steps as described above for siRNA transfection. Expression verification: 48 hours after transfection, mRNA and protein expression of CD300E were detected by qPCR to confirm the effect of gene overexpression.

2.5.3 Proliferation/apoptosis/migration/invision

To evaluate the proliferation of cancer cells, we cultured the cells in suspension and then seeded them at a density of 5 × 10^3 cells/mL (100 μL per well) in a 96-well plate. The plate was maintained at 37°C. Subsequently, we added 10 μL of CCK-8 reagent (catalog KGA9305, KeyGEN, Nanjing, China) to each well and allowed the plate to incubate for two hours before measuring the optical density at 450 nm using a microplate reader.

For the assessment of cell migration and invasion, we utilized Transwell chambers, applying a Matrigel coating for invasion assays and no coating for migration assays. We introduced cancer cells (5×10^4) in 200 μL of serum-free medium into the upper chamber, while the lower chamber was filled with 600 μL of medium supplemented with 10% FBS.

To determine levels of cell apoptosis, we analyzed the apoptosis rate using an Annexin V-FITC/PI Kit (Cat. KGA1102, KeyGEN, Nanjing, China), following the protocol provided by the manufacturer. This method facilitated a precise evaluation of the apoptotic stages within the cancer cell populations.

2.5.4 Real-time quantitative polymerase chain reaction

To assess mRNA abundance at the cellular level, total RNA was meticulously extracted from cells and muscle tissues using the Trizol reagent (Invitrogen) and was precisely quantified with a Nanodrop instrument (Thermo Scientific, USA). Following this, cDNA was synthesized and served as a template for quantifying mRNA expression levels in quantitative PCR (qPCR) assays. These assays were performed using the TB Green™ Premix Ex Taq™ II kit (Takara; RR820A), with GAPDH used as an internal control for normalization. Specific qPCR primers, essential for the amplification of mRNA, were synthesized by Bioengineering (Shanghai, China). The relative expression levels of the mRNA in each sample were calculated using the comparative Ct method (2^-ΔΔCt), ensuring the accuracy of the results through at least three independent experimental replicates. To provide a consistent baseline for comparison, all values were normalized against the control condition. Details of the primer sequences used are available in Supplementary Table 1.

2.6 Statistical methods

Statistical analysis and figure generation were performed with R language version 4.0.2 and Graphpad Prism 9.0. For the comparison of continuous variables between two groups, the choice between the Student t-test and the Mann-Whitney test depended on specific conditions. When comparing multiple groups, either one-way ANOVA or the Kruskal-Wallis test with subsequent multiple comparisons was used, depending on the circumstances. The prognostic significance of categorical variables was determined using the log-rank test. Statistical significance was set at a P value <0.05 across all analyses.

3 Results

3.1 Impact of voluntary running on tumor growth and gene expression

Following the intervention of exercise, a significant reduction in tumor size and weight was observed at day 21, with minimal changes in the body weight of the mice (Supplementary Figures 1A, B). We then conducted mRNA sequencing analysis on five matched pairs (Figure 1A). The quality control results confirmed normal parameters, with high intra-group consistency and notable expression differences between groups (Supplementary

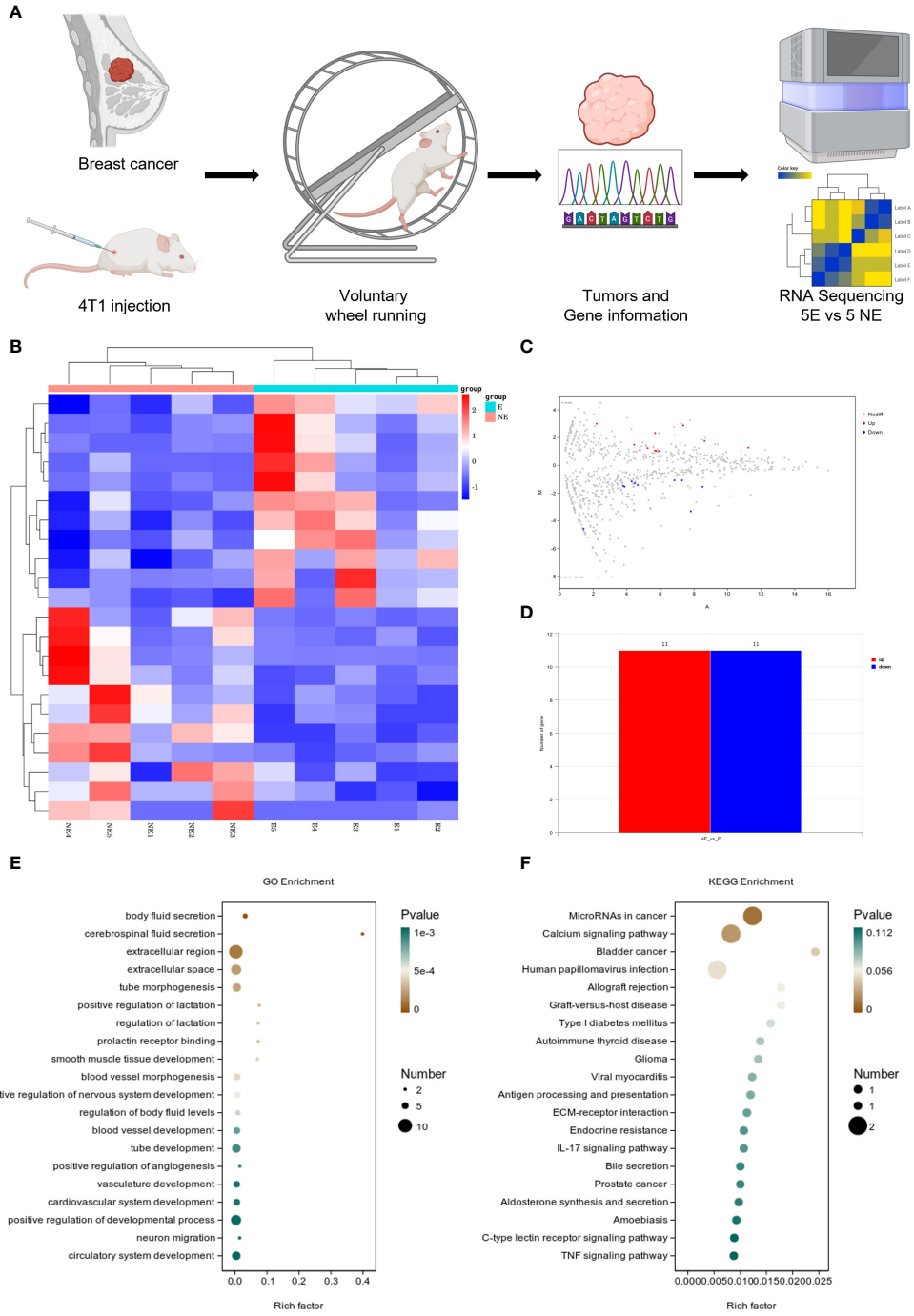


FIGURE 1
Voluntary wheel running exercise inhibits breast cancer growth. **(A)** Schematic diagram of the experiment. **(B)** Heatmap of Hierarchical clustering analysis of changed mRNAs. **(C, D)** Volcano plot and column of mRNAs differentially expressed between NE and E group. n = 5. **(E, F)** Bubble plot showing GO and KEGG enrichment by all the differentially expressed mRNAs expressed in tumors, including biological process, cellular component, and molecular function.

Figures 2A–E). Volcano plots and heatmaps revealed differential expression of 22 genes, among which CD300E expression was significantly reduced in the exercise group (E), representing only 46% of that in the non-exercise group (NE), with a p-value of 0.008 (Figures 1B–D). Gene enrichment analysis highlighted significant alterations in extracellular components, with the most pronounced changes observed in the MicroRNAs in cancers and Calcium signaling pathway (Figures 1E, F).

3.2 Pan-cancer analysis

3.2.1 Expression variability of CD300E in pan-cancer

To evaluate the expression of CD300E mRNA in normal human tissues, we analyzed data from the GTEx, HAP, and Consensus datasets. CD300E showed higher expression in tissues such as blood, lung, bone marrow, appendix, and bladder (Supplementary

Figure 1). Further in-depth evaluation using RNA-seq data from TCGA and GTEx databases revealed significant expression differences in CD300E across 33 types of cancer. In unmatched samples (Figure 2A), CD300E was notably upregulated in cancers like BRCA, COAD, ESCA, GBM, HNSC, KIRC, and STAD, and downregulated in KICH, LIHC, LUAD, LUSC, and PAAD. In matched samples (Figure 2B), upregulation was significant in BRCA, COAD, ESCA, HNSC, KIRC, and STAD, while downregulation was noted in COAD, KICH, LIHC, LUAD, and LUSC. The Human Atlas database further assessed the protein expression of CD300E across various cancers, showing upregulation in Myeloma, Diffuse large B-cell lymphoma, Ovarian cancer, Lung cancer, and Colorectal cancer without significant downregulation in any cancer type (Figure 2C).

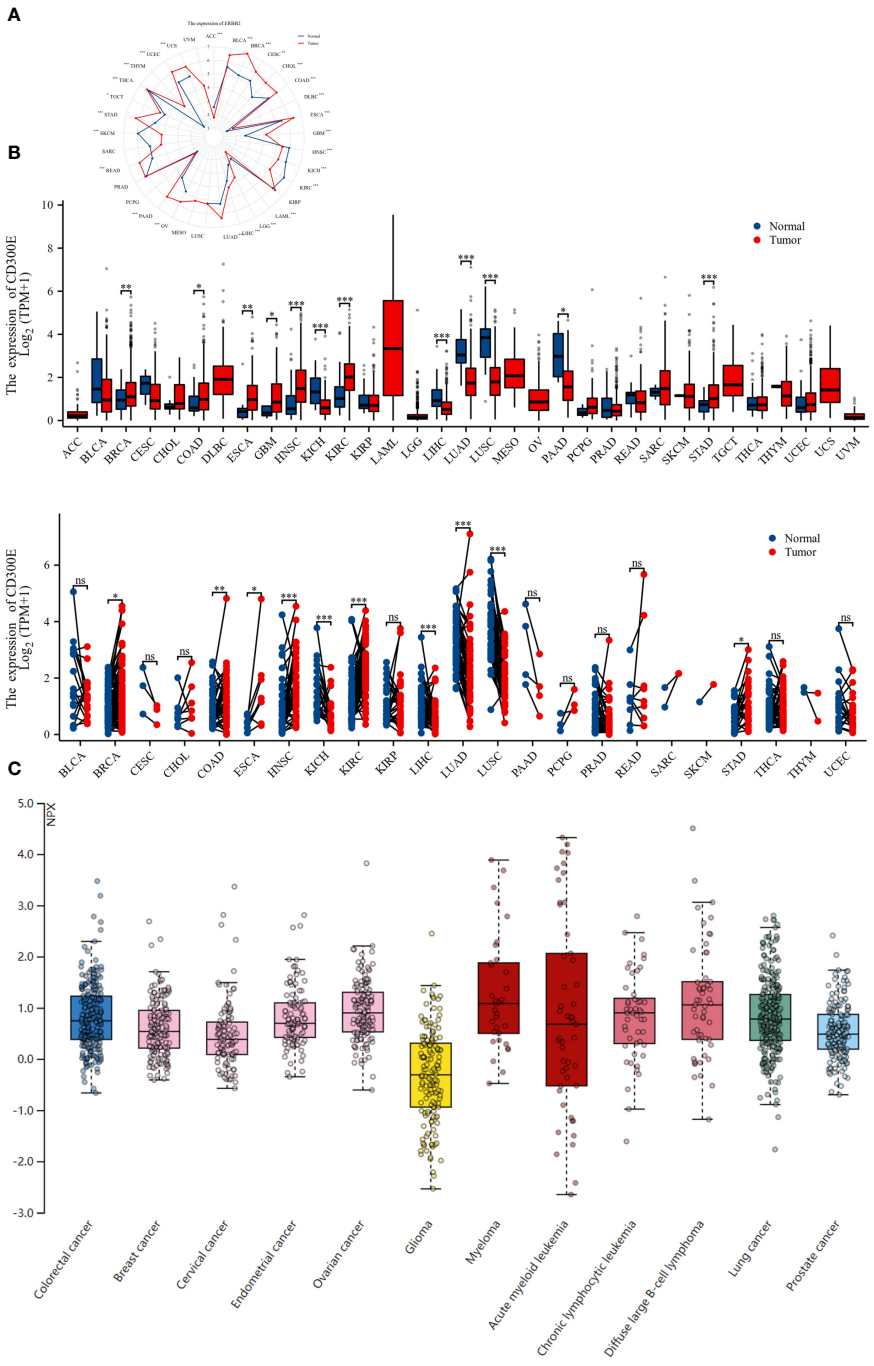


FIGURE 2
Differential expression pattern of CD300E. **(A)** Differential CD300E mRNA expression between paired samples in TCGA cancers. The red dot represents cancer samples, and the blue dot represents paired normal samples. Radargrams visualize and compare CD300E expression in different tumors. *p < 0.05, **p < 0.01, and ***p < 0.001. **(B)** Differential CD300E mRNA expression between TCGA cancers and GTEX normal tissues. The red column represents cancer samples, and the blue column represents normal samples. The normal group was normal tissue in TCGA and GTEX databases. *p < 0.05, **p < 0.01, and ***p < 0.001. **(C)** CD300E protein expression in different cancer types in Human Atlas.

10.3389/fimmu.2024.1437068

3.2.2 Prognostic impact of CD300E in pan-cancer

For overall survival (OS) and disease-specific survival (DSS), CD300E posed a risk factor in THCA, LUSC, LGG, LAML, KIRC, and GBM, while it acted as a protective factor only in SKCM (Figures 3A, B). For disease-free interval (DFI), progression-free interval (PFI), and disease-free survival (DFS), CD300E was a risk factor in KIRP, PAAD, and GBM, and a protective factor in LGG and CHOL (Figure 3A).

3.2.3 Correlation analysis of CD300E in pan-cancer

Copy number variations (CNVs), a common form of genomic instability in cancer, can lead to altered gene expression affecting

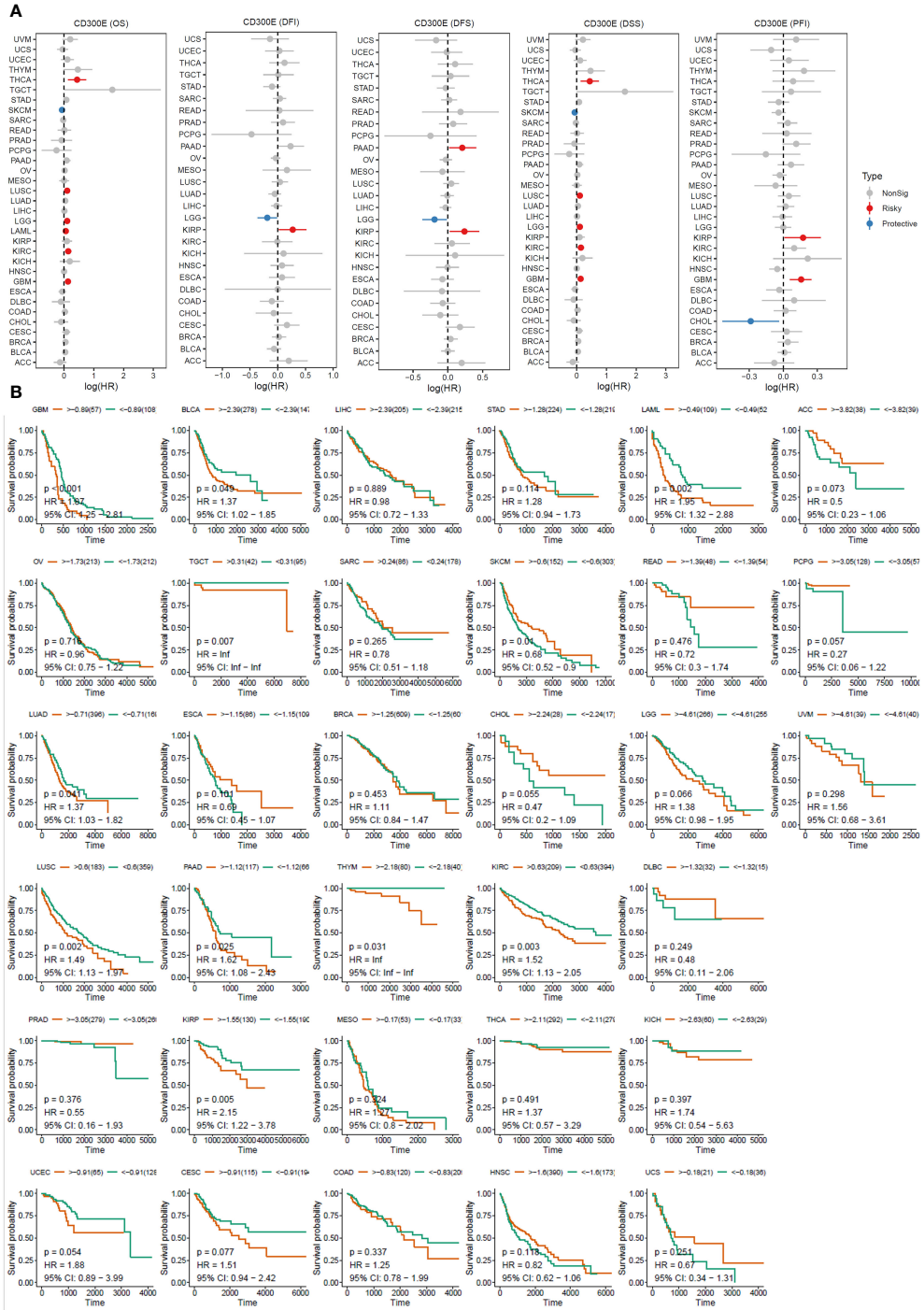


FIGURE 3
High expression of CD300E reduced patient survival period. **(A)** Forest plot of hazard ratios (HR) for overall survival (OS), PFI, DSS, DFS, and DFI for different cancer types associated with CD300E expression. Dots indicate log-transformed hazard ratios, red indicates significant risk, blue indicates protective associations, and gray indicates non-significant associations. **(B)** Individual OS figures for each cancer type.

cell proliferation, differentiation, and death. Bar graphs (Figure 4A) showed changes in CD300E copy numbers across various cancers, with significant variations in KICH and READ. Further correlation analysis indicated a negative relationship between CD300E copy numbers and cancer progression in KIRP and THCA, and a positive correlation in KICH and STAD (Figure 4B). Promoter methylation, a critical epigenetic regulatory mechanism affecting gene expression without altering the DNA sequence, was analyzed to explore its relationship with CD300E expression across multiple cancer types. Both unmatched and matched tumor samples showed a negative correlation between CD300E expression and methylation, particularly in KIRC and THCA (Figures 4C, D). Additionally, the relationship between tumor mutational burden (TMB) and CD300E expression was investigated, revealing a positive correlation in SARC, OV, COAD, BRCA, BLCA, and THYM, and a negative correlation in LAML, LIHC, and PAAD (Figures 4E, F).

3.2.4 Analysis of CD300E on the immune microenvironment across cancers

Heatmap analysis from Figure 5A intricately details the correlations between CD300E expression and various immune cell subtypes across different types of cancers. Notably, in cancers such as BRCA (Breast Cancer) and COAD (Colorectal Adenocarcinoma), a significant positive correlation exists between CD300E expression and M2 macrophages, typically associated with a tumor-promoting immunosuppressive environment. This

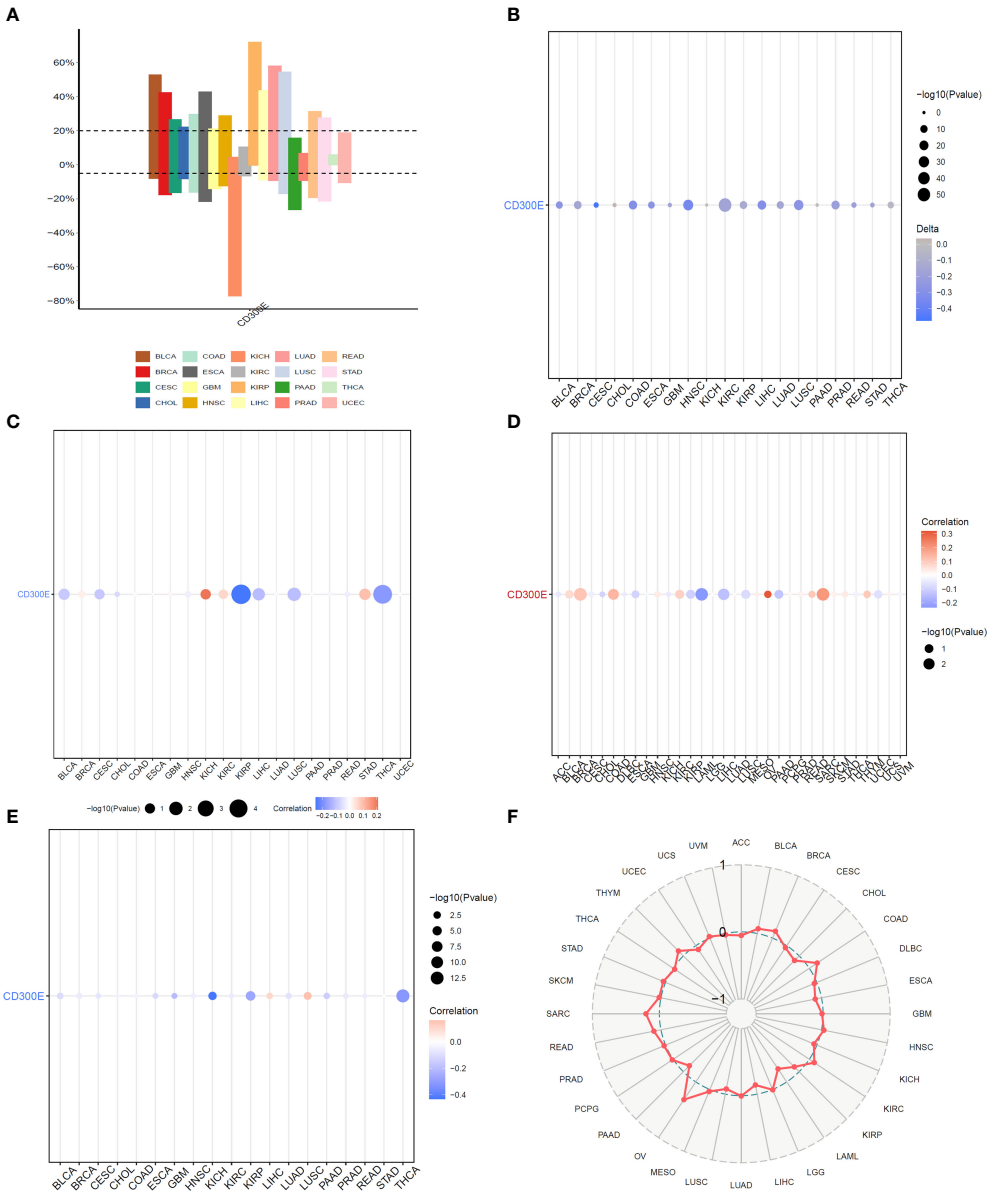


FIGURE 4
Correlation analysis of CD300E in pan-cancer. **(A)** Bar graphs illustrate CD300E copy number variation in different cancers. **(B)** CD300E copy number and pan-cancer direct correlation analysis. **(C, D)** The correlation between the methylation status of gene promoter regions and CD300E in multiple cancer types **(E, F)** The correlation between tumor mutational burden (TMB) and CD300E expression.

10.3389/fimmu.2024.1437068

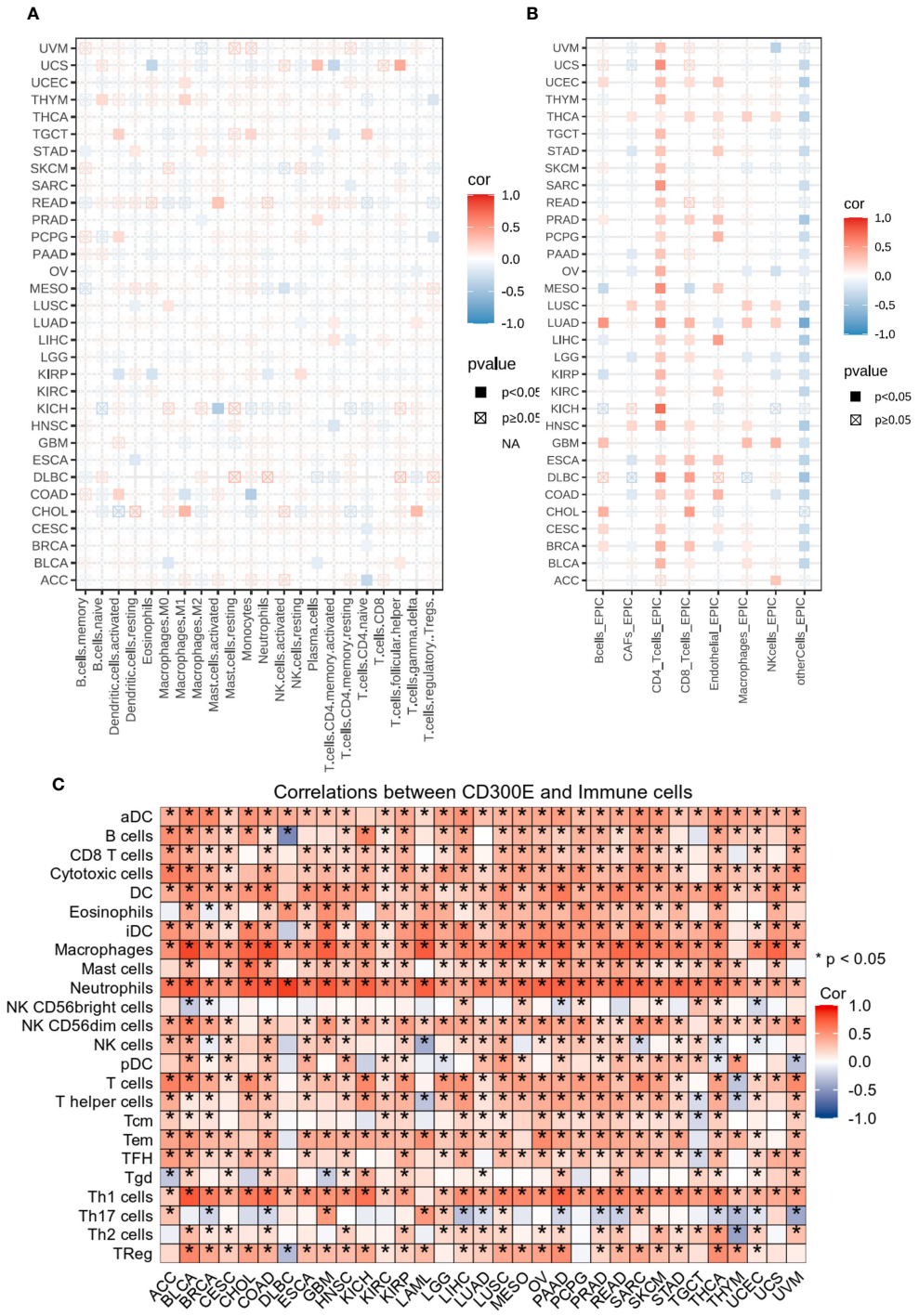


FIGURE 5
Analysis of immune microenvironmental cellular regulation of pan-cancer by CD300E. **(A)** Heatmap of immune cell infiltration in pan-cancer analyzed using the Cibersort method. Each cell represents the correlation between CD300E expression and the level of a specific immune cell type, and the intensity and sign of the color correspond to the strength and direction of the correlation, respectively. Statistical significance is indicated by the box around the cell. **(B)** CD300E pan-cancer immuno-infiltration analysis using Cibersort. **(C)** Gene Commons data analysis of correlations between single genes and immune infiltration results, using heatmap format to present results.

suggests that elevated expression of CD300E may foster an immunosuppressive state conducive to tumor growth and metastasis. Conversely, in Lung Adenocarcinoma (LUAD), CD300E exhibits a negative correlation with natural killer (NK) cells, although this association generally lacks statistical significance. This trend implies that in certain cancer contexts, CD300E expression may inversely affect the immunosurveillance capabilities of NK cells, potentially contributing to mechanisms of immune escape. Additionally, in certain cancer types like BRCA, CD300E shows a positive correlation with regulatory T cells (Tregs),

which play a critical role in modulating the immune system, particularly in maintaining immune tolerance and suppressing excessive immune responses. Increased CD300E expression might enhance the functionality of Tregs, thereby fostering an immune-suppressive tumor microenvironment favorable for tumor survival and progression.

EPIC analysis, a vital tool in studying the tumor microenvironment, enables researchers to understand the dynamic variations of different cell types within tumors, which is crucial for advancing tumor immunology and developing new therapeutic strategies (Figure 5B). From the heatmap, it is evident that CD300E's correlations with various immune cells vary, illustrating the heterogeneity of tumor microenvironments. For instance, in breast and colorectal cancers, Cancer-associated fibroblasts (CAFs) show a strong positive correlation with CD300E expression, suggesting their significant role in supporting or enhancing tumor growth and invasion, closely linked with the expression of this gene. Moreover, in cancers like LUAD, the activity of CD8+ T cells significantly correlates with CD300E expression, reflecting their importance in the tumor immune response and the potential regulatory role of this gene. Further analysis using the TCGA database's pan-cancer dataset revealed a broadly positive correlation between CD300E and various immune cells across different cancer types (Figure 5C).

3.2.5 Pathway enrichment and key gene mutation analysis of CD300E across cancers

Our further evaluation of CD300E's function in pan-cancer contexts revealed significant findings via the GSEA methodology. CD300E notably suppresses oxidative stress pathways, potentially facilitating conditions favorable for tumor growth. Additionally, CD300E significantly enhances pathways such as TNF-α signaling, inflammatory response pathways, IL6-JAK signaling, and epithelial-mesenchymal transition (EMT), all of which are documented to potentially promote tumor growth and metastasis (Figure 6).

A heatmap depicting the frequency of key gene mutations across various cancers highlights the high mutation rates of genes such as TP53 in LUAD, APC in COAD, and PTEN in UCEC,

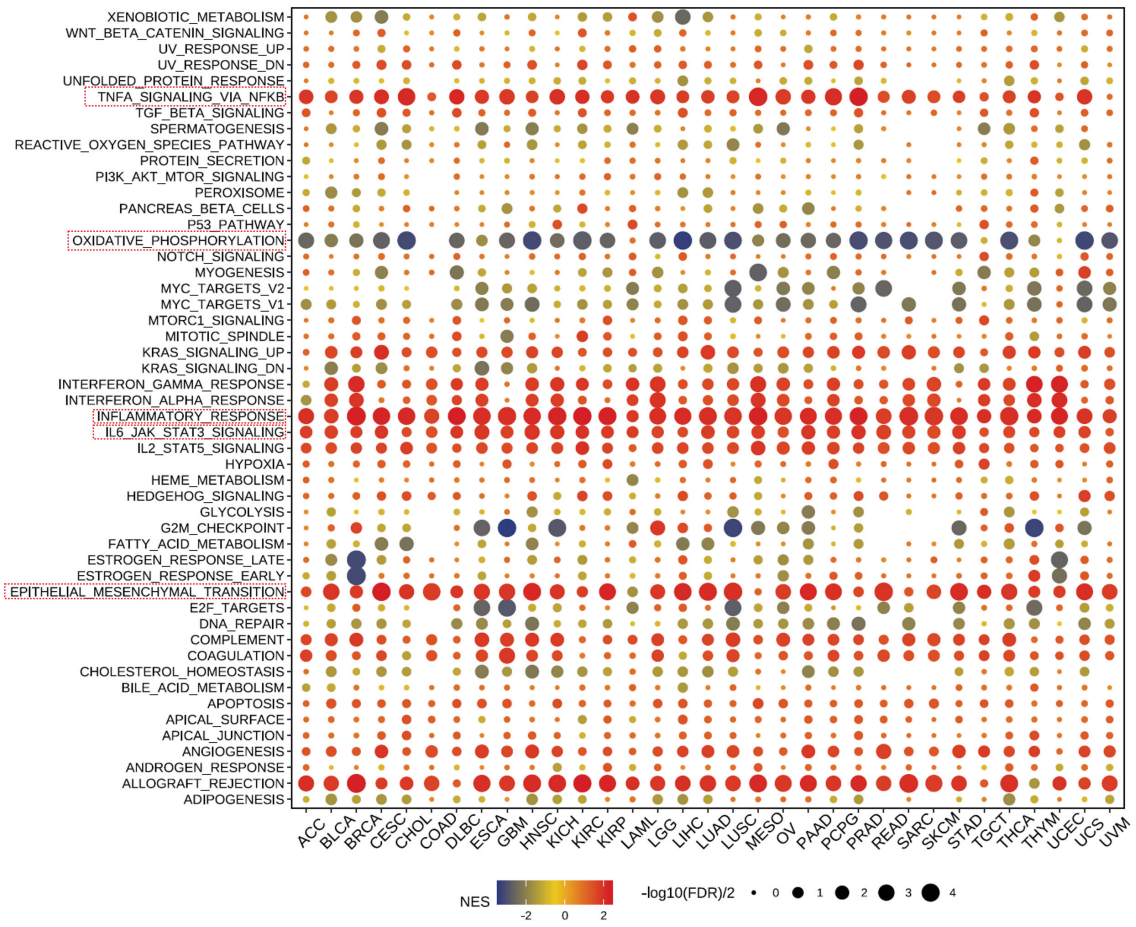


FIGURE 6
Pathway enrichment of CD300E in pan-cancer. Dot plots represent pan-cancer GSEA results using the official immunization gene set (GMT file) as a reference. Functional pathways are from GM7 files and are shown on the y-axis, with different cancer types shown on the x-axis. Dot color indicates correlation with CD300E expression; red indicates positive correlation and blue negative correlation. The size of the dots represents the -log10(FDR) value, indicating the significance of the enrichment.

indicating their common involvement in these cancers. Specific cancer types like BRCA, LGG, and HNSC show frequent mutations in genes like TP53, PIK3CA, and CDKN2A, providing insights that may guide therapeutic strategies (Supplementary Figure 5A).

3.3 Impact of CD300E on breast cancer cells

Finally, our study delves into the cellular functions of CD300E. We validated the expression of the CD300E gene after siRNA or plasmids intervention (Supplementary Figures 6A, B). Compared to control cells, overexpression of CD300E in MDAMB468 and MDAMB231 breast cancer cells leads to increased proliferation and cell viability, while suppression of CD300E expression reduces proliferation and cell viability (Figures 7A, B). Furthermore, overexpression of CD300E significantly promotes the migratory and invasive capabilities of these tumor cells, whereas its inhibition reduces these properties (Figures 7C, D). Overall, targeting CD300E could directly inhibit tumor cells, significantly impeding cancer progression and presenting a novel therapeutic target (Figure 8).

4 Discussion

This research explored the impact of exercise on tumor growth and gene expression within a murine model, focusing particularly on the expression patterns, functions, and potential clinical significance of the CD300E gene across various cancers. Our findings indicate that CD300E may adversely affect prognosis and promote tumor progression across a range of cancers. Additionally, exercise appears to inhibit breast cancer progression potentially by downregulating CD300E.

Exercise as well as widespread is believed to promote human health and improve a wide range of diseases (31–33). The phenomenon of exercise against cancer has been widely explored in recent years, but there is still a large number of exercise-responsive molecules whose roles need to be explored (34–38). Our study confirmed the positive impact of physical activity on inhibiting tumor growth. Exercise intervention significantly reduced tumor size and weight in the murine model without markedly affecting body weight. These outcomes suggest that moderate physical activity might suppress tumor growth by modifying the tumor microenvironment or regulating specific signaling pathways. Analysis of differentially expressed genes revealed significant downregulation of CD300E in the exercise group, indicating its role in tumor growth regulation, particularly within an active context. Furthermore, gene enrichment analysis showed significant changes in extracellular components and associated signaling pathways, such as MicroRNAs in cancers and the Calcium signaling pathway, providing clues on how exercise might influence tumor biology through molecular mechanisms.

In our pan-cancer analysis, CD300E exhibits significant expression variability across multiple cancer types, underscoring its potential role in various malignancies. Notably, CD300E is upregulated in cancers such as Myeloma, Diffuse Large B-cell Lymphoma, Ovarian Cancer, Lung Cancer, and Colorectal Cancer, suggesting its involvement in the progression of these diseases. Prognostic analyses reveal that CD300E acts as a risk factor in several cancers, providing valuable insights that may guide clinical prognostic assessments and therapeutic decision-making. Studies on the variability of CD300E copy numbers and their correlation with tumor mutational burden offer critical perspectives on its role in cancer progression. These findings support the notion that CD300E may promote cancer development by impacting genetic stability and the interactions within the immune microenvironment.

Analysis of the relationships between CD300E and various immune cell subpopulations indicates that CD300E may influence tumor growth and immune escape by modulating immune cells within the tumor microenvironment, particularly immunosuppressive M2 macrophages and regulatory T cells. Past studies have also shown that CD300E and T cells are associated with the regulation of immune function in macrophages, and more mechanistic studies are needed to explore this (22–24). But, the results also illustrated that CD300E showed a significant positive correlation with most other immune-promoting immune cells, including CD8 T cells, neutrophils, NK cells, etc. The literature reports that these cells more or less affect the heating and cooling of the immune microenvironment. It has been reported in the literature that these cells more or less affect the heat and cold of the immune microenvironment (26). The increased expression of CD300E resulted in both Immunosuppressive and immunopromoting cells, affecting the tumor microenvironment, which could potentially affect the prognosis and the degree of response to immunotherapy. These insights lay a theoretical foundation for targeting CD300E in immunotherapeutic strategies.

Furthermore, our analysis elucidates the role of CD300E in regulating key signaling pathways related to cancer progression, especially in suppressing oxidative stress pathways and activating several pathways that promote tumor progression. The inhibition of oxidative stress pathways may provide cancer cells with mechanisms to evade programmed cell death, thereby covertly supporting tumor growth and survival (39–41). Concurrently, CD300E significantly activates pathways such as the TNF-α pathway, inflammatory response pathways, the IL6-JAK pathway, and the epithelial-mesenchymal transition pathway, all closely associated with the invasiveness and metastatic potential of tumors (42–45). These pathways' activation might facilitate the dissemination of tumor cells within the host. Analysis of the frequency of key gene mutations reveals frequent mutations in genes such as TP53, APC, and PTEN across various cancers, highlighting these genes as critical factors in tumor development and progression (46). These mutations may affect cell cycle regulation, DNA repair mechanisms, and pathways of cell death, further substantiating the potential role of CD300E in pan-cancer contexts.

Additionally, our cellular experiments clearly demonstrate that the overexpression of CD300E in breast cancer cells is closely associated with enhanced cellular proliferation, reduced apoptosis rates, and increased migration and invasion capabilities. These findings not only confirm the role of CD300E as a tumor-promoting factor but

10.3389/fimmu.2024.1437068

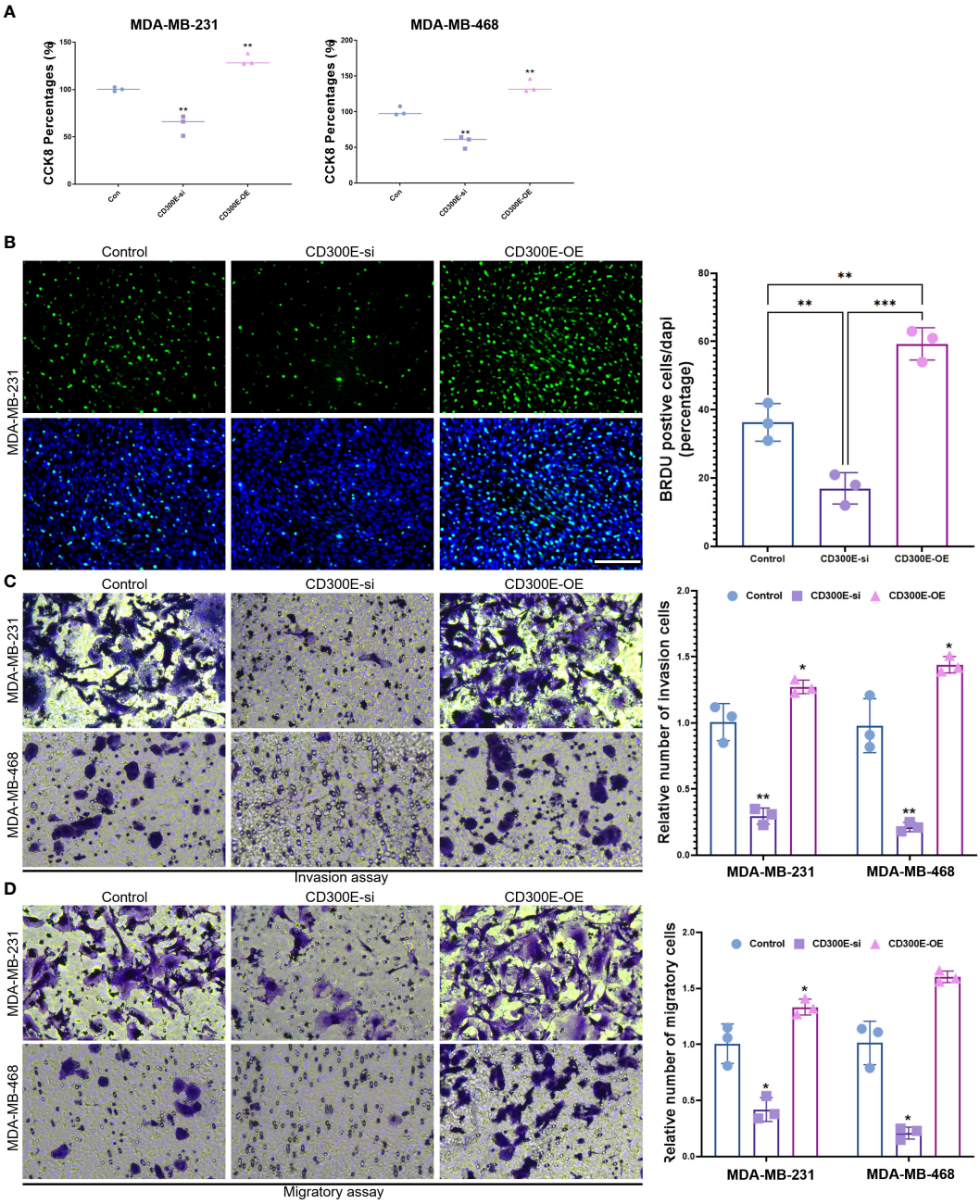


FIGURE 7
CD300E promotes breast cancer tumor cell growth. **(A)** The viability of control, CD300E-inhibited, and CD300E-overexpressed tumor cells was examined at 48h after transfection by CCK-8 assay. The statistical significance of the differences between various treatments is determined by one-way ANOVA with Bonferroni post-test (n = 3). Data are presented as mean ± SD. *P < 0.05 **P < 0.01. **(B)** The proliferative capacity of control, CD300E-inhibited, and CD300E-overexpressed tumor cells was examined at 24h after transfection by BRDU. The statistical significance of the differences between various treatments is determined by one-way ANOVA with Bonferroni post-test (n = 3). Data are presented as mean ± SD. *P < 0.05 **P < 0.01 ***P < 0.001. **(C, D)**. The migratory and invasive capacity of control, CD300E-inhibited, and CD300E-overexpressed tumor cells were examined at 24h after transfection by Boyden chamber assay. Total original magnification, 200×. The statistical significance of the differences between various treatments is determined by one-way ANOVA with Bonferroni post-test (n = 3). Data are presented as mean ± SD. *P < 0.05 **P < 0.01.

also highlight its potential as a therapeutic target. Experiments aimed at inhibiting CD300E expression further validate its significant role in tumor cell proliferation and survival, offering a potential therapeutic strategy to curb the progression of breast cancer. Studies have reported that CD300E can modulate apoptosis in monocytes by affecting calcium channels, which is consistent with our biological predictions. In addition, altered calcium signaling affects the behavior of immune cells (including T cells and macrophages), influencing their activation and cytokine production, thereby altering the immune microenvironment (47–49). Therefore, we hypothesize that the ability of CD300E to promote tumor cell value-addition and migration is reached by regulating calcium channels.

Mechanically, how exercise regulates CD300E lowering this process was not explored in this study. However, a large body of

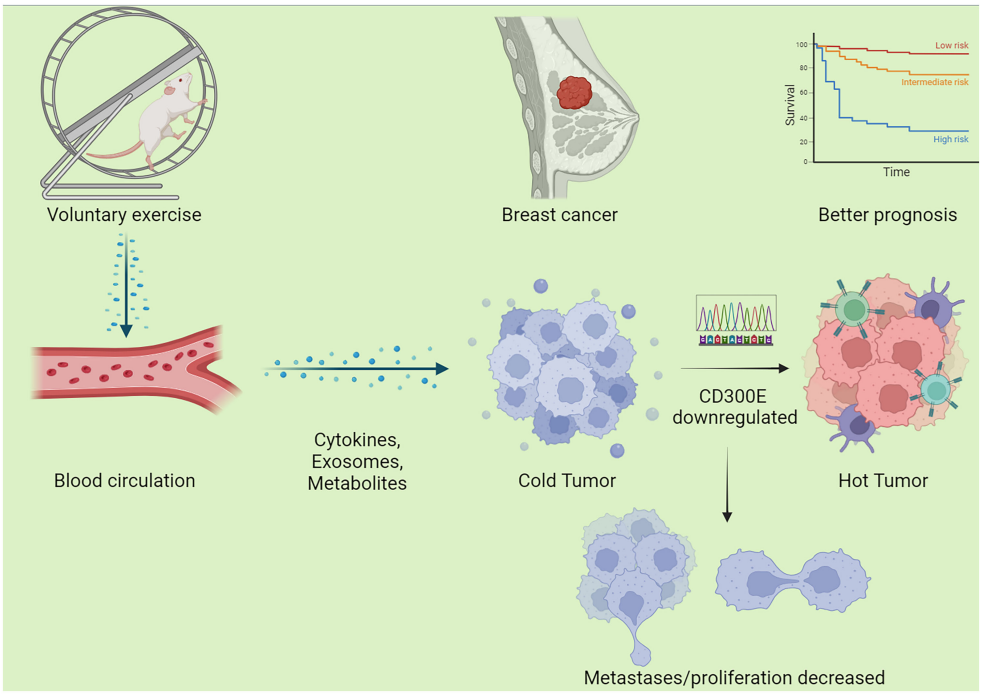


FIGURE 8
Schematic graph of this study. Exercise decreased CD300E expression of cells in breast cancer through a circulatory effect, which promotes immune cell infiltration, decreased tumor cell metastases/proliferation, warms the tumor microenvironment, and improves the prognosis of tumor patients.

literature has reported that exercise can bring about a series of physiological changes, including changes in the metabolome, proteins, and related molecules in the genome (12, 50, 51). Specifically, we hypothesize that exercise-induced changes in systemic factors, such as serum circulating exosome, muscle derived cytokines, and hormones, could impact transcription factors like NF-κB and STAT3, known regulators of gene transcription (41, 52–54). Additionally, the role of epigenetic modifications, including DNA methylation and histone acetylation, in the regulation of gene expression in response to physical activity could also influence the expression of CD300E (55). In addition, the direct upstream transcription factor(s) by which exercise regulates CD300E expression in tumor cells remains unknown. We proposed that exercise activates AMP-activated protein kinase (AMPK) and peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PGC-1alpha), which are central to tumor cell expression (56, 57). These molecules may influence the transcription factors and co-regulators that control CD300E expression. Furthermore, exercise can also modulate the expression of cellular miRNA, which may post-transcriptionally regulate CD300E (58, 59). For example, miR-4270 has been reported to directly target CD300E (60), but these are speculations based on the literature, and future studies will need to further explore the mechanisms by which exercise regulates CD300E.

As for the clinical translational perspective, we believe that patients with high CD300E expression may benefit from more intensive or specific types of exercise therapies that are particularly effective in downregulating CD300E. Conversely, patients with low CD300E expression may require different exercise regimens or adjunctive therapies to achieve optimal results. To test these hypotheses, we recommend that future studies design clinical trials that stratify patients according to CD300E expression levels. These trials should include a variety of exercise regimens from moderate to high intensity and monitor changes in CD300E expression, tumor progression, and clinical prognosis. In addition, patient-reported outcomes and quality-of-life measures should be included to assess the broader impact of tailored exercise interventions. Moreover, we recommend longitudinal studies to track CD300E expression and tumor progression in response to sustained exercise therapy. These studies will help determine the sustainability of exercise-induced changes in gene expression and their long-term impact on cancer prognosis (44, 61, 62).

Limitations and perspectives: Sample Size and Type Limitations: This study is primarily based on animal models and specific cancer cell lines, which may restrict the generalizability of the findings and their direct applicability to human cancer patients (63). While murine models provide valuable insights into tumor biology, they cannot fully replicate the complexity and heterogeneity of human tumors (64–66). Singular Focus of Study Design: Although we observed the impact of exercise on tumor growth and CD300E expression, there is a lack of exploration into variables such as exercise intensity, frequency, and duration. Moreover, the study focuses predominantly on the role of a single gene, CD300E, while tumor development involves multiple genes and signaling pathways interacting (67). Complexity in Data Interpretation: While gene expression and pathway enrichment

analyses have unveiled potential biological mechanisms, the exact causal relationships remain unclear. For instance, the direct link between changes in CD300E expression and specific tumor behaviors has not been fully established. Future experimental designs should consider the effects of various types and intensities of exercise on tumor growth and how these variables interact with gene expression and immune responses within the tumor microenvironment (68). Additionally, investigating the role of CD300E across different cancers and immune backgrounds may reveal its multifunctional potential as a therapeutic target. Further mechanistic studies should delve into how CD300E activates or inhibits cancer-related pathways, particularly how it influences key tumor behaviors such as cell cycle progression, apoptosis, migration, and invasion. While current research focuses on exploring tumor therapy at the level of a single gene, future studies could use single-cell sequencing and spatial transcriptome analysis to identify a broader range of genes affected by exercise (69–72). These studies could use integrated bioinformatics approaches to elucidate gene-gene interactions and pathways co-regulated by exercise.

5 Conclusions

In summary, CD300E not only plays a potentially crucial role in the process of exercise-mediated tumor growth inhibition but also exhibits viability as a therapeutic target based on its expression and function across various cancers. Future research should further explore the specific molecular mechanisms of CD300E and its role in different cancers to advance the development of novel anti-cancer strategies.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here GEO repository, accession number, GSE272213: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE272213.

Ethics statement

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used. The animal study was approved by Ethics Committees at Nanjing Medical University. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

ZL: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Writing – original draft, Writing – review & editing. JZ: Investigation, Methodology, Project administration, Resources, Writing – original draft. RX: Investigation, Methodology, Project administration, Writing – original draft. RW: Investigation, Writing – original draft. YH: Investigation, Writing – original draft. YC: Investigation, Software, Writing – original draft, Writing – review & editing. QW: Conceptualization, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft. SC (8th Author): Conceptualization, Funding acquisition, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. SC (9th Author): Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was financially supported by the National Natural Science Foundation of China (Grant No. 82172511,81972125 and 82172510). Shenzhen “San-Ming” Project of Medicine (No. SZSM202211019). Nanjing Municipal Bureau of Physical Education and Sports Research Bureau Management Subjects (NJTY2023-104) and Nanjing “Sports and Health Integration” New Model (JSYGY-3-2023-505).

Acknowledgments

We thank Dr. Alsaleh for the data analysis.

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.

Publisher’s note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1437068/full#supplementary-material

References

1. Hajmomeni P, Sisakhtnezhad S, Bidmeshkipour A. Thymoquinone-treated mouse mesenchymal stem cells-derived conditioned medium inhibits human breast cancer cells in vitro. *Chem Biol Interact*. (2023) 369:100283. doi: 10.1016/j.cbi.2022.110283

2. Xing P, Wang S, Cao Y, Liu B, Zheng F, Guo W, et al. Treatment strategies and drug resistance mechanisms in adenocarcinoma of different organs. *Drug Resist Updat*. (2023) 71:101002. doi: 10.1016/j.drup.2023.101002

3. Singh D, Assaraf YG, Gacche RN. Long non-coding RNA mediated drug resistance in breast cancer. *Drug Resistance Updates*. (2022) 63:100851. doi: 10.1016/j.drup.2022.100851

4. Zhang J, Sun J, Li C, Qiao H, Hussain Z. Functionalization of curcumin nanomedicines: a recent promising adaptation to maximize pharmacokinetic profile, specific cell internalization and anticancer efficacy against breast cancer. *J Nanobiotechnol*. (2023) 21:106. doi: 10.1186/s12951-023-01854-x

5. Zheng C, Zhang D, Kong Y, Niu M, Zhao H, Song Q, et al. Dynamic regulation of drug biodistribution by turning tumors into decoys for biomimetic nanoplatform to enhance the chemotherapeutic efficacy of breast cancer with bone metastasis. *Explor (Beijing China)*. (2023) 3:20220124. doi: 10.1002/EXP.20220124

6. Wang M, Sun R, Chen H, Liu X, Yoshitomi T, Takeguchi M, et al. Influence of hydrogel and porous scaffold on the magnetic thermal property and anticancer effect of Fe3O4 nanoparticles. *Microstructures*. (2023) 3:2023042. doi: 10.20517/microstructures.2023.46

7. Tang Y, Tian W, Zheng S, Zou Y, Xie J, Zhang J, et al. Dissection of FOXO1-Induced LYPLAL1-DT Impeding Triple-Negative Breast Cancer Progression *via* Mediating hnRNPK/β-Catenin Complex. *Res (Wash D C)*. (2023) 6:289. doi: 10.34133/research.0289

8. Ji X, Tian X, Feng S, Zhang L, Wang J, Guo R, et al. Intermittent F-actin perturbations by magnetic fields inhibit breast cancer metastasis. *Res (Wash D C)*. (2023) 6:80. doi: 10.34133/research.0080

9. Zeng R, Li Y, Li Y, Wan Q, Huang Z, Qiu Z, et al. Smartphone-based photoelectrochemical immunoassay with co(9)S(8)@ZnIn(2)S(4) for point-of-care diagnosis of breast cancer biomarker. *Res (Wash D C)*. (2022) 2022:9831521. doi: 10.34133/2022/9831521

10. Huo R, Zhan M, Zhu B, Zhi Q. Annual advances of traditional Chinese medicine on tumor immunity regulation in 2021. *Traditional Med Res*. (2022) 7:56. doi: 10.53388/TMR20220615001

11. Liao K, Gong L, Yang Y, He Y, Wang F, Huang Y, et al. A comprehensive review of research progress in Chinese medicines for primary liver cancer treatment. *Traditional Med Res*. (2022) 7:10. doi: 10.53388/TMR20220207263

12. Chen Y, Chen X, Luo Z, Kang X, Ge Y, Wan R, et al. Exercise-induced reduction of IGF1R sumoylation attenuates neuroinflammation in APP/PS1 transgenic mice. *J Adv Res*. (2024). doi: 10.1016/j.jare.2024.03.025

13. Luo Z, Zhang T, Chen S. Exercise prescription: pioneering the "Third pole" for clinical health management. *Research*. (2023) 6:284. doi: 10.34133/research.0284

14. Jones LW, Alfano CM. Exercise-oncology research: Past, present, and future. *Acta Oncol (Madr)*. (2013) 52:195–215. doi: 10.3109/0284186X.2012.742564

15. Chen Y, Sun Y, Luo Z, Chen X, Wang Y, Qi B, et al. Exercise modifies the transcriptional regulatory features of monocytes in alzheimer's patients: A multi-omics integration analysis based on single cell technology. *Front Aging Neurosci*. (2022) 14. doi: 10.3389/fnagi.2022.881488

16. Chen Y, Luo Z, Sun Y, Li F, Han Z, Qi B, et al. Exercise improves choroid plexus epithelial cells metabolism to prevent glial cell-associated neurodegeneration. *Front Pharmacol*. (2022) 13:1010785. doi: 10.3389/fphar.2022.1010785

17. Chen Y, Sun Y, Luo Z, Lin J, Qi B, Kang X, et al. Potential mechanism underlying exercise upregulated circulating blood exosome miR-215-5p to prevent necroptosis of neuronal cells and a model for early diagnosis of alzheimer's disease. *Front Aging Neurosci*. (2022) 14:860364. doi: 10.3389/fnagi.2022.860364

18. Luo Z, He Z, Qin H, Chen Y, Qi B, Lin J, et al. Exercise-induced IL-15 acted as a positive prognostic implication and tumor-suppressed role in pan-cancer. *Front Pharmacol*. (2022) 13:1053137. doi: 10.3389/fphar.2022.1053137

19. Luo Z, Wan R, Liu S, Feng X, Peng Z, Wang Q, et al. Mechanisms of exercise in the treatment of lung cancer – a mini-review. *Front Immunol*. (2023) 14:1244764. doi: 10.3389/fimmu.2023.1244764

20. Wan R, Chen Y, Feng X, Luo Z, Peng Z, Qi B, et al. Exercise potentially prevents colorectal cancer liver metastases by suppressing tumor epithelial cell stemness *via* RPS4X downregulation. *Heliyon*. (2024) 10:e26604. doi: 10.1016/j.heliyon.2024.e26604

21. Luo Z, Sun Y-Y, Xia W, Xu J-Y, Xie D-J, Jiao C-M, et al. Physical exercise reverses immuno-cold tumor microenvironment *via* inhibiting SQLE in non-small cell lung cancer. *Mil Med Res*. (2023) 10:39. doi: 10.1186/s40779-023-00474-8

22. Clark GJ, Jamriska L, Rao M, Hart DNJ. Monocytes immunoselected *via* the novel monocyte specific molecule, CD300e, differentiate into active migratory dendritic cells. *J Immunotherapy*. (2007) 30:303–11. doi: 10.1097/01.cji.0000211342.65964.9e

23. Chen K, Silveira P, Lo TH, Hart D, Clark G. Characterizing the function of the CD300e immunoregulatory molecule. *J Immunol*. (2018) 200:109.1. doi: 10.4049/jimmunol.200.Supp.109.1

24. Wang J, Manick B, Li J, Bi M, Kalabokis V, Person A, et al. CD300e acts as A ligand to inhibit T cell immunity. *J Immunol*. (2022) 208:54.16. doi: 10.4049/jimmunol.208.Supp.54.16

25. Isobe M, Izawa K, Sugiuchi M, Sakanishi T, Kaitani A, Takamori A, et al. The CD300e molecule in mice is an immune-activating receptor. *J Biol Chem*. (2018) 293:3793–805. doi: 10.1074/jbc.RA117.000696

26. Brckalo T, Calzetti F, Pérez-Cabezas B, Borràs FE, Cassatella MA, López-Botet M. Functional analysis of the CD300e receptor in human monocytes and myeloid dendritic cells. *Eur J Immunol*. (2010) 40:722–32. doi: 10.1002/eji.200939468

27. Coletta S, Trevellin E, Benagiano M, Romagnoli J, Della Bella C, D'Elios MM, et al. The antigen CD300e drives T cell inflammation in adipose tissue and elicits an antibody response predictive of the insulin sensitivity recovery in obese patients. *J Inflammation (United Kingdom)*. (2022) 19:21. doi: 10.1186/s12950-022-00318-7

28. Guo Y, Wang Z, Shi X, Shen M. Engineered cancer cell membranes: An emerging agent for efficient cancer theranostics. *Explor (Beijing China)*. (2022) 2:20210171. doi: 10.1002/EXP.20210171

29. Theivendran S, Lazarev S, Yu C. Mesoporous silica/organosilica nanoparticles for cancer immunotherapy. *Explor (Beijing China)*. (2023) 3:20220086. doi: 10.1002/EXP.20220086

30. Gu X, Cai L, Luo Z, Shi L, Peng Z, Sun Y, et al. Identification and validation of a muscle failure index to predict prognosis and immunotherapy in lung adenocarcinoma through integrated analysis of bulk and single-cell RNA sequencing data. *Front Immunol*. (2023) 13:1057088. doi: 10.3389/fimmu.2022.1057088

31. Ristow M, Zarse K, Oberbach A, Klöting N, Birringer M, Kiehntopf M, et al. Antioxidants prevent health-promoting effects of physical exercise in humans. *Proc Natl Acad Sci*. (2009) 106:8665–70. doi: 10.1073/pnas.0903485106

32. Mcleod JC, Currier BS, Lowisz CV, Phillips SM. The influence of resistance exercise training prescription variables on skeletal muscle mass, strength, and physical function in healthy adults: An umbrella review. *J Sport Health Sci*. (2023) 00:1–14. doi: 10.1016/j.jshs.2023.06.005

33. Lou J, Wu J, Feng M, Dang X, Wu G, Yang H, et al. Exercise promotes angiogenesis by enhancing endothelial cell fatty acid utilization *via* liver-derived extracellular vesicle miR-122-5p. *J Sport Health Sci*. (2022) 11:495–508. doi: 10.1016/j.jshs.2021.09.009

34. Cataldi S, Greco G, Mauro M, Fischetti F. Effect of exercise on cancer-related fatigue: A systematic review. *J Hum Sport Exercise*. (2020) 16:476–92. doi: 10.14198/jhse

35. Alves CRR, das Neves W, de Almeida NR, Eichelberger EJ, Jannig PR, Voltarelli VA, et al. Exercise training reverses cancer-induced oxidative stress and decrease in muscle COPS2/TRIP15/ALIEN. *Mol Metab*. (2020) 39:101012. doi: 10.1016/j.molmet.2020.101012

36. Thomas RJ, Kenfield SA, Jimenez A. Exercise-induced biochemical changes and their potential influence on cancer: a scientific review. *Br J Sports Med*. (2016) 51:640–4. doi: 10.1136/bjsports-2016-096343

37. Kim JS, Galvão DA, Newton RU, Gray E, Taaffe DR. Exercise-induced myokines and their effect on prostate cancer. *Nat Rev Urol*. (2021) 18:519–42. doi: 10.1038/s41585-021-00476-y

38. Gould DW, Lahart I, Carmichael AR, Koutedakis Y, Metsios GS. Cancer cachexia prevention *via* physical exercise: molecular mechanisms. *J Cachexia Sarcopenia Muscle*. (2012) 4:111–24. doi: 10.1007/s13539-012-0096-0

39. Valko M, Rhodes CJ, Moncol J, Izakovic M, Mazur M. Free radicals, metals and antioxidants in oxidative stress-induced cancer. *Chem Biol Interact*. (2006) 160:1–40. doi: 10.1016/j.cbi.2005.12.009

40. Reuter S, Gupta SC, Chaturvedi MM, Aggarwal BB. Oxidative stress, inflammation, and cancer: How are they linked? *Free Radic Biol Med*. (2010) 49:1603–16. doi: 10.1016/j.freeradbiomed.2010.09.006

41. Yang L, Hao J, Chen D. Role of oxidative stress in anti-cancer activity of Tripterygium wilfordii. *Traditional Med Res*. (2021) 6:50. doi: 10.53388/TMR20211011246

42. Mishra A, Sullivan L, Caligiuri MA. Molecular pathways: Interleukin-15 signaling in health and in cancer. *Clin Cancer Res*. (2014) 20:2044–50. doi: 10.1158/1078-0432.CCR-12-3603

43. Orlandella FM, De Stefano AE, Iervolino PLC, Buono P, Soricelli A, Salvatore G. Dissecting the molecular pathways involved in the effects of physical activity on breast cancers cells: A narrative review. *Life Sci*. (2021) 265:118790. doi: 10.1016/j.lfs.2020.118790

44. Guo S-B, Huang W-J, Tian X-P. Brusatol modulates diverse cancer hallmarks and signaling pathways as a potential cancer therapeutic. *Acta Materia Med*. (2022) 1:278–301. doi: 10.15212/AMM-2022-0014

45. Nilsson R, Jain M, Madhusudhan N, Sheppard NG, Strittmatter L, Kampf C, et al. Metabolic enzyme expression highlights a key role for {MTHFD2} and the mitochondrial folate pathway in cancer. *Nat Commun*. (2014) 5:3128. doi: 10.1038/ncomms4128

46. Cooks T, Pateras IS, Jenkins LM, Patel KM, Robles AI, Morris J, et al. Mutant p53 cancers reprogram macrophages to tumor supporting macrophages *via* exosomal miR-1246. *Nat Commun*. (2018) 9:771. doi: 10.1038/s41467-018-03224-w

47. Cui Y, Li J, Zhang P, Yin D, Wang Z, Dai J, et al. B4GALT1 promotes immune escape by regulating the expression of PD-L1 at multiple levels in lung adenocarcinoma. *J Exp Clin Cancer Res*. (2023) 42:146. doi: 10.1186/s13046-023-02711-3

48. Zhang P, Dong S, Sun W, Zhong W, Xiong J, Gong X, et al. Deciphering Treg cell roles in esophageal squamous cell carcinoma: a comprehensive prognostic and immunotherapeutic analysis. *Front Mol Biosci*. (2023) 10. doi: 10.3389/fmolb.2023.1277530

49. Zhang P, Pei S, Zhou G, Zhang M, Zhang L, Zhang Z. Purine metabolism in lung adenocarcinoma: A single-cell analysis revealing prognostic and immunotherapeutic insights. *J Cell Mol Med*. (2024) 28:e18284. doi: 10.1111/jcmm.18284

50. Sasso JP, Eves ND, Christensen JF, Koelwyn GJ, Scott J, Jones LW. A framework for prescription in exercise-oncology research. *J Cachexia Sarcopenia Muscle*. (2015) 6:115–24. doi: 10.1002/jcsm.12042

51. Hou Z, Qin X, Hu Y, Zhang X, Li G, Wu J, et al. Longterm exercise-derived exosomal miR-342-5p. *Circ Res*. (2019) 124:1386–400. doi: 10.1161/CIRCRESAHA.118.314635

52. Liu X, Jin X, Ou H, Qian C, Wu H, Zuo C, et al. The direct STAT3 inhibitor 6-ethoxydihydrosanguinarine exhibits anticancer activity in gastric cancer. *Acta Materia Med*. (2022) 1:365–80. doi: 10.15212/AMM-2022-0027

53. Dai S, Wen Y, Luo P, Ma L, Liu Y, Ai J, et al. Therapeutic implications of exosomes in the treatment of radiation injury. *Burns Trauma*. (2022) 10:tkab043. doi: 10.1093/burnst/tkab043

54. Hu T, Chang S, Qi F, Zhang Z, Chen J, Jiang L, et al. Neural grafts containing exosomes derived from Schwann cell-like cells promote peripheral nerve regeneration in rats. *Burns Trauma*. (2023) 11:tkad013. doi: 10.1093/burnst/tkad013

55. Gillman AS, Helmuth T, Koljack CE, Hutchison KE, Kohrt WM, Bryan AD. The effects of exercise duration and intensity on breast cancer-related dna methylation: A randomized controlled trial. *Cancers (Basel)*. (2021) 13:4128. doi: 10.3390/cancers13164128

56. Crane JD, Macneil LG, Lally JS, Ford RJ, Bujak AL, Brar IK, et al. Exercise-stimulated interleukin-15 is controlled by AMPK and regulates skin metabolism and aging. *Aging Cell*. (2015) 14:625–34. doi: 10.1111/acel.12341

57. Liu C, Qin Q, Cong H. Research progress on the relationship between mitochondrial deoxyguanosine kinase and apoptosis and autophagy in lung adenocarcinoma cells. *Cancer Insight*. (2022) 1:53–62. doi: 10.58567/ci01010004

58. Cui SF, Wang C, Yin X, Tian D, Lu QJ, Zhang CY, et al. Similar responses of circulating microRNAs to acute high-intensity interval exercise and vigorous-intensity continuous exercise. *Front Physiol*. (2016) 7:102. doi: 10.3389/fphys.2016.00102

59. Margolis LM, Hatch-McChesney A, Allen JT, DiBella MN, Carrigan CT, Murphy NE, et al. Circulating and skeletal muscle microRNA profiles are more sensitive to sustained aerobic exercise than energy balance in males. *J Physiol*. (2022) 600:3951–63. doi: 10.1113/JP283209

60. Pagliari M, Munari F, Toffoletto M, Lonardi S, Chemello F, Codolo G, et al. Helicobacter pylori affects the antigen presentation activity of macrophages modulating the expression of the immune receptor CD300E through miR-4270. *Front Immunol*. (2017) 8. doi: 10.3389/fimmu.2017.01288

61. Feng D, Li D, Wang J, Wu R, Zhang C. Senescence-associated lncRNAs indicate distinct molecular subtypes associated with prognosis and androgen response in patients with prostate cancer. *Acta Materia Med*. (2023) 2:299–309. doi: 10.15212/AMM-2023-0025

62. Wang J, Wei J, Inuzuka H. Aging and cancer hallmarks as therapeutic targets. *Acta Materia Med*. (2023) 2:281–4. doi: 10.15212/AMM-2023-0018

63. Conlon KC, Lugli E, Welles HC, Rosenberg SA, Fojo AT, Morris JC, et al. Redistribution, hyperproliferation, activation of natural killer cells and CD8 T cells, and cytokine production during first-in-human clinical trial of recombinant human interleukin-15 in patients with cancer. *J Clin Oncol*. (2015) 33:74–82. doi: 10.1200/JCO.2014.57.3329

64. Mu Y, Gao W, Zhou Y, Xiao L, Xiao Y. Physiological and pathological/ectopic mineralization: from composition to microstructure. *Microstructures*. (2023) 3:2023030. doi: 10.20517/microstructures

65. Wang X, Guo X. A rare case of a primary cardiac tumor presenting as fatal ventricular tachycardia. *Cardiovasc Innov Appl*. (2023) 7:996. doi: 10.15212/CVIA.2022.0014

66. Ghanta SN, Vallurupalli S. Carcinoid heart disease – A review of pathophysiology, clinical manifestations, diagnosis and management. *Cardiovasc Innov Appl*. (2023) 8:980. doi: 10.15212/CVIA.2023.0041

67. Zhang Y, Zhang D, An S, Liu Q, Liang C, Li J, et al. Development and characterization of nanobody-derived CD47 theranostic pairs in solid tumors. *Res (Wash D C)*. (2023) 6:77. doi: 10.34133/research.0077

68. Chen B, Guo K, Zhao X, Liu Z, Xu C, Zhao N, et al. Tumor microenvironment-responsive delivery nanosystems reverse immunosuppression for enhanced CO gas/immunotherapy. *Explor (Beijing China)*. (2023) 3:20220140. doi: 10.1002/EXP.20220140

69. Mukherjee D, Lahiri D, Nag M. Therapeutic effects of natural products isolated from different microorganisms in treating cervical cancer: A review. *Cancer Insight*. (2022) 1:31–46. doi: 10.58567/ci01020003

70. Chen H, Li H, Shi W, Qin H, Zheng L. The roles of m6A RNA methylation modification in cancer stem cells: new opportunities for cancer suppression. *Cancer Insight*. (2022) 1:1–18. doi: 10.58567/ci01020001

71. Mitchell H, Jeffrey R. Special healthcare services for lgbt cancer patients. *Cancer Insight*. (2022) 1:19–30. doi: 10.58567/ci01020002

72. Tong Z, Mang G, Wang D, Cui J, Yang Q, Zhang M. Single-cell RNA sequencing maps immune cell heterogeneity in mice with allogeneic cardiac transplantation. *Cardiovasc Innov Appl*. (2023) 8:988. doi: 10.15212/CVIA.2023.0023

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 20 August 2024
DOI 10.3389/fonc.2024.1454372

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Gang Tian,
Sichuan Province Engineering Technology Research Center of Molecular Diagnosis of Clinical Diseases, China
Ying Yang,
Jinhua Municipal Central Hospital, China

*CORRESPONDENCE
Qiming Wang
m13957885162_1@163.com

RECEIVED 25 June 2024
ACCEPTED 26 July 2024
PUBLISHED 20 August 2024

CITATION
Yang J, Yu H, Zhang Y, Zhu M, Zhang M and Wang Q (2024) Efficacy of PD-1 or PD-L1 inhibitors for the therapy of cervical cancer with varying PD-L1 expression levels: a single-arm meta-analysis.
Front. Oncol. 14:1454372.
doi: 10.3389/fonc.2024.1454372

Efficacy of PD-1 or PD-L1 inhibitors for the therapy of cervical cancer with varying PD-L1 expression levels: a single-arm meta-analysis

Jie Yang, Haizan Yu, Yilei Zhang, Mingli Zhu, Mengyu Zhang and Qiming Wang*

Department of Gynaecology, III, Women's and Children's Hospital of Ningbo University, Ningbo, Zhejiang, China

Objective: To assess the effectiveness and tolerability of both PD-1 and PD-L1 inhibitors in advanced cervical cancer (CC), focusing on varying PD-L1 levels.

Methods: A comprehensive exploration was carried out on EMBASE, PubMed, Cochrane Library databases as well as Web of Science up to May 25, 2024, for studies involving advanced CC patients receiving PD-1/PD-L1 inhibitors. Inclusion criteria were studies reporting objective response rate (ORR), disease control rate (DCR), median progression-free survival (PFS), as well as median overall survival (OS). Data extraction and quality assessment were performed by two reviewers using the JBI Case Series Critical Appraisal Checklist, followed by a meta-analysis via STATA/MP 16.0.

Results: Five eligible studies comprising 223 patients were chosen. ORR and DCR were 42% (95% CI: 17%-66%, P = 0.00) and 70% (95% CI: 22%-117%, P = 0.00), respectively, in the PD-L1 positive patients and were 36% (95% CI: 17%-54%, P = 0.00) and 47% (95% CI: 30%-63%, P = 0.00), respectively, in patients with PD-L1 negativity. For patients exhibiting PD-L1 positivity, median PFS and median OS were 3.98 months (95% CI: 0.80–7.16, P = 0.01) and 11.26 months (95% CI: 3.01–12.58, P = 0.00), respectively.

Conclusion: With PD-1/PD-L1 inhibitors, PD-L1 positive CC patients demonstrate superior ORR, DCR, median PFS, and median OS, underscoring PD-L1 as one biomarker for immunotherapy response.

KEYWORDS

PD-1inhibitors, PD-L1 expression, cervical cancer, PD-L1 inhibitors, meta-analysis

Introduction

Cervical cancer (CC) is still a significant contributor to cancer-related mortality in women worldwide, particularly in middle- and low-income countries (1). According to 2020 data, there were approximately 604,127 new cases of cervical cancer worldwide, and 341,831 deaths, with age-standardised incidence and mortality rates of 13.3 and 7.2 per 100,000 women, respectively (2). Despite great progress in both screening and vaccination, a majority of patients still experience serious disease or recurrence and have limited therapy options and unfavourable prognoses (3, 4). Traditional therapies, including chemotherapy, radiation as well as surgery, have presented limited efficacy in these stages of the disease, entailing the exploration of innovative therapy (5).

With the advent of immunotherapy, cancer treatment has been revolutionized bringing hope for patients suffering from advanced tumours. Programmed cell death protein 1 (PD-1) and programmed death-ligand 1 (PD-L1) inhibitors have presented encouraging results in cancers as one class of immune checkpoint inhibitors, including melanoma, bladder cancer as well as non-small cell lung cancer (6, 7). These inhibitors lift the immune system's capability of recognizing and eliminating cancer cells by disrupting the binding between PD-1 on T cells and PD-L1 on tumour cells (7). The PD-L1 quantification on tumour cells is commonly assessed using the Combined Positive Score (CPS). It has emerged as one potential biomarker for forecasting the reaction to PD-1/PD-L1 inhibitors (8). CPS is determined by assessing the proportion of PD-L1-positive tumour cells and immune ones relative to the total viable tumour ones (9). Preliminary clinical studies indicate a possibility of exhibiting better reactions to PD-1/PD-L1 inhibitors in patients having higher CPS, which implies a potential stratified therapy (10, 11).

The meta-analysis is to assess the effectiveness and tolerability of both PD-1 and PD-L1 inhibitors in treating advanced CC systematically, with a particular focus on different PD-L1 expressions. Data were integrated from various high-quality studies to comprehensively understand the potential of these immune therapies in improving the outcomes of advanced CC patients.

Methods

Based on implementation under the recommendations of the Cochrane Handbook for Systematic Reviews of Interventions, this study was reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (12). The current study was formally registered on the International Platform of Registered Systematic Review and Meta-analysis Protocols (INPLASY) (ID: INPLASY202460062).

Search strategy

We performed an extensive search across various databases like Web of Science, PubMed, EMBASE, as well as the Cochrane Library, encompassing articles published before May 25, 2024. The search was restricted to studies published exclusively in the English language with the following terms for search: "Uterine Cervical Neoplasms" OR "CC" AND "Immune Checkpoint Inhibitors" OR "PD-1 Inhibitor" OR "PD-L1 Inhibitor". We performed a manual review to the reference lists of the encompassed articles for identifying additional related research. The particular search process is detailed in Supplementary File 1.

Inclusion and exclusion criteria

Studies were encompassed if they met the criteria below:

1. Patients were confirmed with advanced or recurrent CC, regardless of subtype.
2. Patients received treatment by PD-1 or PD-L1 inhibitors alone or in conjunction with other therapies.
3. Retrospective analyses or stage II clinical trials.
4. Included studies assessed relevant clinical outcomes, such as PFS, ORR, OS, DCR, as well as AEs, using RECIST 1.1 criteria (13).
5. Tumour PD-L1 was assessed and quantified as one CPS, which was calculated as the percentage of PD-L1-stained cells divided by the sum of viable tumour cells multiplied by 100. The definition of positivity was established as having a CPS of 1 or higher.

The exclusion criteria were:

1. Animal research, meta-analyses, reviews, duplicate reports, letters or case reports.
2. Studies with fewer than 10 patients.

Two reviewers conducted a thorough screening of articles independently, assessing their eligibility according to pre-established criteria Disagreements/discrepan were resolved through consensus between the two reviewers or with the assessment of one-third reviewers if necessary.

Data extraction and quality evaluation

Through one predefined extraction form, two reviewers extracted data. The extracted data encompassed baseline patient characteristics, study characteristics, and predefined outcomes (ORR, DCR, PFS, OS). The quality of clinical studies was evaluated via the JBI Case Series Critical Appraisal Checklist (14).

Statistical analyses

Analyses were conducted via STATA/MP 16.0. Inter-study heterogeneity was judged via the chi-square test as well as the I^2 statistic. Random-effects models (REM) were adopted when $I^2 \geq 50\%$

(indicating high heterogeneity), and fixed-effects models (FEM) were adopted when I^2<50% (implying low heterogeneity) (15). The robustness of the pooled results was judged via sensitivity analyses. Egger's test was conduc to evaluate the possible publication bias.

Results

Literature search

The initial search strategy yielded 2,998 relevant articles. After removing 1,053 duplicate studies, we screened titles and abstracts, causing the exclusion of 1,894 studies not fulfilling the inclusion criteria. Subsequently, we performed a detailed review of the whole texts of the left 51 potentially eligible papers, and ultimately selected 5 trials for the final analysis (16–19). The process of selecting studies is depicted in Figure 1. All eligible research data were obtained from published manuscripts.

Study characteristics

Totally, 5 studies were included in the final analysis Table 1 presents their detailed characteristics.

Quality assessment

On the basis of the JBI Critical Appraisal Checklist for Case Series, five clinical studies were evaluated, comprising ten items that examine the quality of case reports including case selection, evaluation of the disease or health problem, and case data presentation. The assessment results are provided in Table 2.

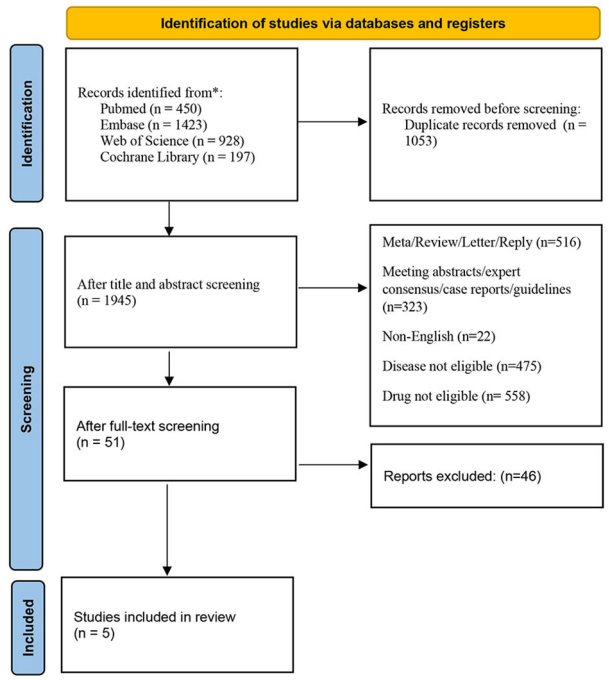


FIGURE 1
The flow diagram of studies is included in this meta-analysis.

Meta-analysis results

Comparison of ORR by PD-L1 CPS

Five studies (223 patients) analyzed ORR by PD-L1 CPS (16–20). In patients exhibiting PD-L1 positivity, a REM was used because of notable heterogeneity (I^2 = 89.53%, P = 0.00). The ORR was 42% (95% CI: 17%-66%, P = 0.00, Figure 2). In patients exhibiting PD-L1 negativity, a FEM was used because of low heterogeneity (I^2 = 0.00%, P = 0.45). The ORR was 36% (95% CI: 17%-54%, P = 0.00, Figure 3).

Comparison of DCR by PD-L1 CPS

Three studies (176 patients) analyzed DCR by PD-L1 CPS (17, 19, 21). In PD-L1 positive patients, a REM was used because of notable heterogeneity (I^2 = 98.15%, P = 0.00). The DCR was 70% (95% CI: 22%-117%, P = 0.00), as shown in Figure 4. In PD-L1 negative patients, a FEM was used because of low heterogeneity (I^2 = 10.25%, P = 0.33). The DCR was 47% (95% CI: 30%-63%, P = 0.00), as shown in Figure 5.

Median PFS in patients exhibiting PD-L1 positivity

Three studies (170 patients) analyzed PFS in Patients exhibiting PD-L1 positivity (16, 17, 20). A REM was used because of notable heterogeneity (I^2 = 78.54%, P = 0.01). The PFS was 3.98 months (95% CI: 0.80–7.16, P = 0.01), as shown in Figure 6.

Median OS in patients exhibiting PD-L1 positivity

Two studies (125 patients) analyzed OS in patients exhibiting PD-L1 positivity (16, 20). A FEM was used due to low heterogeneity (I^2 = 0.00%, P = 0.42). The OS was 11.26 months (95% CI: 3.01–12.58, P = 0.00, Figure 7).

Sensitivity analysis

By sequentially excluding each study, a sensitivity analysis was performed for assessing its impact on the summary results. According to the analysis results, no individual study significantly impacts the overall 95% CI of the summary results, indicating a relatively robust of the meta-analysis results. The results are presented in Supplementary File 2.

Publication bias

To ensure the validity of the meta-analysis, publication bias was judged via Egger's test. The p-value of 0.79 (> 0.05), indicates no notable publication bias.

TABLE 1 Characteristics of studies included in this meta-analysis.

Study	Year	Study type	Stage	Age	Intervention types	Number of patients	PD-L1 CPS≥1%	PD-L1 CPS<1%	PD-L1 CPS unknown	Follow-up (m), median (range)
Chunyan Lan	2024	NRCT single-arm, phase II	metastatic, recurrent, or persistent cervical cancer	51 (33–67)	Camrelizumab	45	10	30	5	6 (0.97–37.4)
Yin Wang	2023	NRCT single-arm, phase II	recurrent or metastatic cervical cancer	50 (34–68)	Sintilimab	27	18	5	4	10.2 (3.0–24.5)
Lingfang Xia	2022	NRCT single-arm, phase II	recurrent or metastatic cervical cancer	50 (43–55)	Camrelizumab	33	10	9	14	13.6 (10.0–23.6)
Hyun Cheol Chung	2019	international, open-label, multicohort	advanced Cervical Cancer	46 (24–75)	Pembrolizumab	98	82	15	1	10.2 (0.6–22.7)
Kenji Tamura	2019	prospective, multicenter, open-label	advanced or recurrent uterine cervical cancer	50 (32–68)	Nivolumab	20	5	15	/	5.4 (1.0–13.9)

Discussion

This study comparatively analyzed ORR and DCR among patients who had different PD-L1 CPS, focusing on assessing the efficacy disparity between groups exhibiting PD-L1 positivity and PD-L1 negativity. The results revealed an ORR of 42% (95% CI: 17%-66%) and 36% (95% CI: 17%-54%) in the group exhibiting PD-L1 positivity and group exhibiting PD-L1 negativity, respectively. This difference suggests a possibly larger response rate of PD-L1-positive patients to immunotherapy. The underlying mechanism for it can be explained by the interaction between PD-L1 with the immune system. PD-L1, a cell surface protein frequently found on tumour cells, binds to the PD-1 receptor on T cells, suppressing the activity of T cells and helping tumour cells evade immune system attacks (22, 23). In tumours expressing PD-L1, tumour cells can more effectively utilize this mechanism to evade immune surveillance. Thus, these patients possibly have a better response to immune checkpoint inhibitors like PD-1/PD-L1 inhibitors, as these drugs are able to disrupt the binding of PD-1/PD-L1 with restore T cell-mediated tumour attack (24, 25). DCR was also compared among patients who had different PD-L1 CPS. The group exhibiting PD-L1 positivity and group exhibiting PD-L1

TABLE 2 The JBI Critical Appraisal Checklist for Case Series.

Query	Chunyan Lan	Yin Wang	Lingfang Xia	Hyun Cheol Chung	Kenji Tamura
Were there clear criteria for inclusion in the case series?	YES	YES	YES	YES	YES
Was the condition measured in a standard, reliable way for all participants included in the case series?	YES	YES	YES	YES	YES
Were valid methods used for the identification of the condition for all participants included in the case series?	YES	YES	YES	YES	YES
Did the case series have consecutive inclusion of participants?	UNCLEAR	YES	YES	YES	UNCLEAR
Did the case series have a complete inclusion of participants?	YES	YES	YES	YES	YES
Was there clear reporting of the demographics of the participants in the study?	YES	YES	YES	YES	YES
Was there clear reporting of clinical information of the participants?	YES	YES	YES	YES	YES
Were the outcomes or follow-up results of cases clearly reported?	YES	YES	YES	YES	YES
Was there clear reporting of the presenting site(s)/clinic(s) demographic information?	YES	YES	YES	YES	YES
Was statistical analysis appropriate?	YES	YES	YES	YES	YES
Overall appraisal	Include	Include	Include	Include	Include

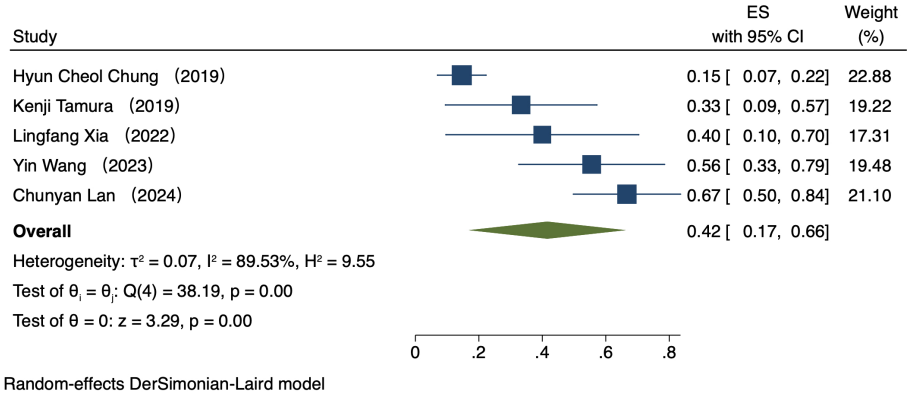


FIGURE 2
Forest plot of ORR in PD-L1 positive.

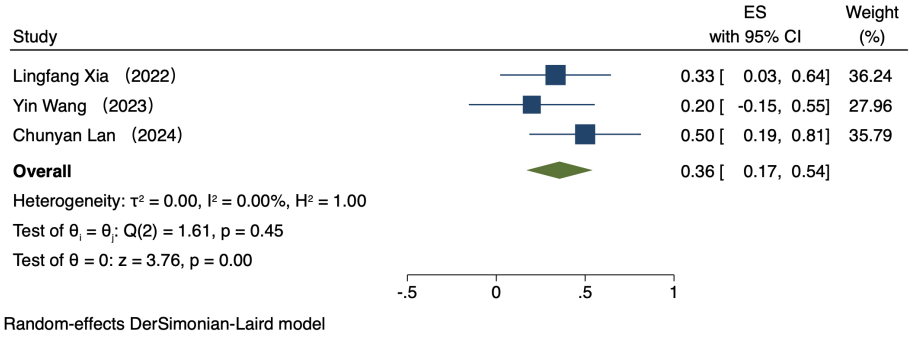


FIGURE 3
Forest plot of ORR in PD-L1 negative.

negativity had a DCR of 70% (95% CI: 22%-117%) and 47% (95% CI: 30%-63%), respectively. These findings imply the high value of PD-L1 expression in immune therapy response further (26). for more deeply probing into the survival outcomes of patients exhibiting PD-L1 CPS positivity, we analyzed the PFS and OS and found a PFS and OS of 3.98 months (95% CI: 0.80–7.16) and 7.80 months (95% CI: 3.01–12.58), respectively, in patients exhibiting PD-L1 CPS positivity. The findings imply the possibility of experiencing improved long-term survival rates among PD-L1 CPS-positive patients receiving immune therapy (27–29).

These results underscore the high value of PD-L1 in immune therapy. Patients exhibiting PD-L1 positivity demonstrated better efficacy in multiple key outcome measures in contrast to patients

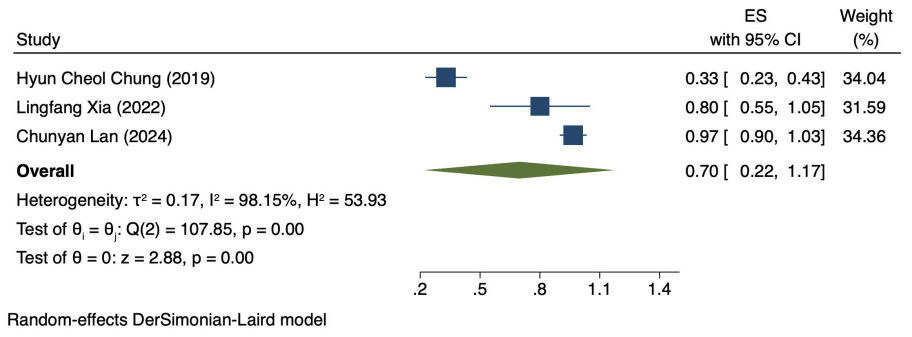


FIGURE 4
Forest plot of DCR in PD-L1 positive.

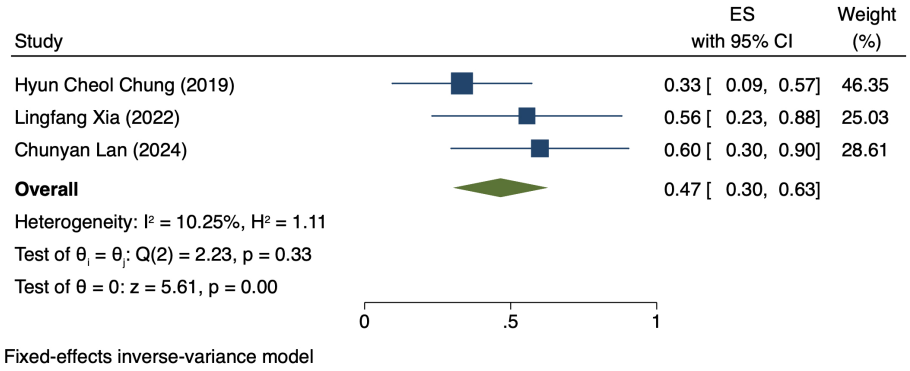


FIGURE 5
Forest plot of DCR in PD-L1 negative.

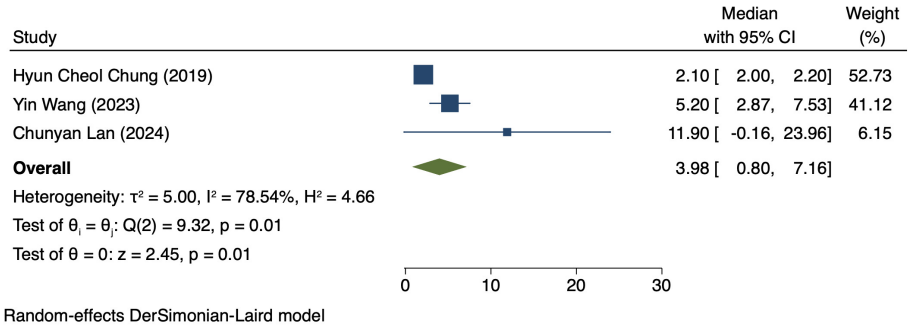


FIGURE 6
Forest plot of PFS in PD-L1 positive.

exhibiting PD-L1 negativity, indicating PD-L1 as an effective biomarker for identifying patients with a larger likelihood of favorable response to immune therapy in patients (30, 31).

Whereas, the current research also has certain limitations. First, a noticeable heterogeneity in the analysis could affect the stability of the results. Second, the included studies with relatively small sample sizes mostly consisted of non-controlled trials, limiting the generalizability and persuasiveness of the findings. Additionally, because of lack of enough pathological data, we could not further investigate the treatment response based on different types of CC tissue. studies included in this analysis predominantly involved Asian patients, raising uncertainty about the generalizability of these findings to other populations. Therefore, further validation of these findings is warranted through the implementation of large-scale randomized controlled trials (RCTs) in the future (32, 33).

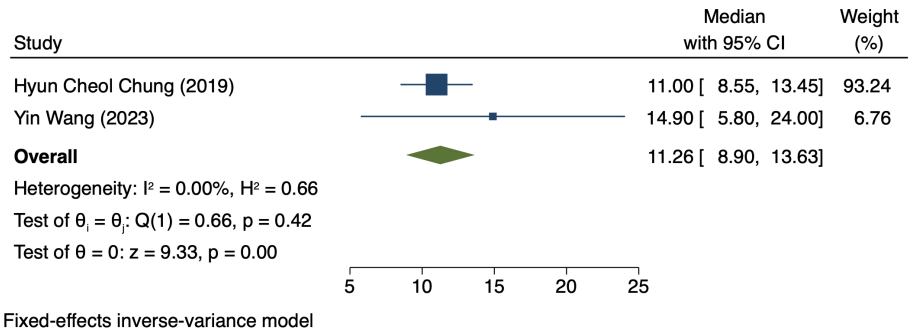


FIGURE 7
Forest plot of OS in PD-L1 positive.

In conclusion, PD-L1 expression is crucial in immune therapy, with PD-L1 CPS-positive patients demonstrating better efficacy in terms of ORR, DCR, median PFS, and median OS in contrast to patients exhibiting PD-L1 negativity. While the initial findings are encouraging, additional research is required to ascertain the wide applicability as well as long-term implications of these findings (34).

Conclusion

The meta-analysis verifies that CC patients exhibiting PD-L1 positivity have superior efficacy regarding ORR, DCR, median PFS, as well as median OS when receiving PD-1/PD-L1 inhibitor therapy in contrast to patients exhibiting PD-L1 negativity. These findings support the utilization of PD-L1 as one biomarker for forecasting the advanced CC patients' reaction to immunotherapy.

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

JY: Conceptualization, Data curation, Investigation, Writing – original draft. HY: Investigation, Methodology, Writing – original draft. YZ: Data curation, Formal analysis, Writing – original draft. MlZ: Formal analysis, Investigation, Writing – original draft. MyZ: Formal analysis, Investigation, Project administration, Writing – original draft. QW: Conceptualization, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2024.1454372/full#supplementary-material

SUPPLEMENTARY FILE 1
Search strategy.

SUPPLEMENTARY FILE 2
Sensitivity Analysis.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. (2018) 68:394–424. doi: 10.3322/caac.21492

2. Singh D, Vignat J, Lorenzoni V, Eslahi M, Ginsburg O, Lauby-Secretan B, et al. Global estimates of incidence and mortality of cervical cancer in 2020: a baseline analysis of the WHO Global Cervical Cancer Elimination Initiative. *Lancet Glob Health*. (2023) 11:e197–206. doi: 10.1016/S2214-109X(22)00501-0

3. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin*. (2024) 74:12–49. doi: 10.3322/caac.21820

4. Arbyn M, Weiderpass E, Bruni L, de Sanjosé S, Saraiya M, Ferlay J, et al. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *Lancet Glob Health*. (2020) 8:e191–203. doi: 10.1016/S2214-109X(19)30482-6

5. Marth C, Landoni F, Mahner S, McCormack M, Gonzalez-Martin A, Colombo N. Cervical cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. (2017) 28:iv72–83. doi: 10.1093/annonc/mdx220

6. Ribas A, Wolchok JD. Cancer immunotherapy using checkpoint blockade. *Science*. (2018) 359:1350–5. doi: 10.1126/science.aar4060

7. Sharma P, Allison JP. The future of immune checkpoint therapy. *Science*. (2015) 348:56–61. doi: 10.1126/science.aaa8172

8. Patel SP, Kurzrock R. PD-L1 expression as a predictive biomarker in cancer immunotherapy. *Mol Cancer Ther*. (2015) 14:847–56. doi: 10.1158/1535-7163.MCT-14-0983

9. Ilie M, Long-Mira E, Bence C, Butori C, Lassalle S, Bouhlel L, et al. Comparative study of the PD-L1 status between surgically resected specimens and matched biopsies of NSCLC patients reveal major discordances: a potential issue for anti-PD-L1 therapeutic strategies. *Ann Oncol*. (2016) 27:147–53. doi: 10.1093/annonc/mdv489

10. Chung H, Delord J-P, Perets R, Italiano A, Shapira-Frommer R, Manzuk L, et al. Pembrolizumab treatment of advanced cervical cancer: updated results from the phase II KEYNOTE-158 study. *Gynecologic Oncol*. (2021) 162:S27. doi: 10.1016/S0090-8258(21)00696-X

11. Rischin D, Gil-Martin M, González-Martin A, Braña I, Hou JY, Cho D, et al. PD-1 blockade in recurrent or metastatic cervical cancer: Data from cemiplimab phase I expansion cohorts and characterization of PD-L1 expression in cervical cancer. *Gynecol Oncol*. (2020) 159:322–8. doi: 10.1016/j.ygyno.2020.08.026

12. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PloS Med*. (2009) 6:e1000100.

13. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. (2009) 45:228–47. doi: 10.1016/j.ejca.2008.10.026

14. Munn Z, Barker TH, Moola S, Tufanaru C, Stern C, McArthur A, et al. Methodological quality of case series studies: an introduction to the JBI critical appraisal tool. *JBI Evid Synth*. (2020) 18:2127–33.

15. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. (2002) 21:1539–58. doi: 10.1002/sim.1186

16. Chung HC, Ros W, Delord JP, Perets R, Italiano A, Shapira-Frommer R, et al. Efficacy and safety of pembrolizumab in previously treated advanced cervical cancer:

results from the phase II KEYNOTE-158 study. *J Clin Oncol*. (2019) 37:1470–8. doi: 10.1200/JCO.18.01265

17. Lan C, Lu H, Zhou L, Liao K, Liu J, Xie Z, et al. Long-term survival outcomes and immune checkpoint inhibitor retreatment in patients with advanced cervical cancer treated with camrelizumab plus apatinib in the phase II CLAP study. *Cancer Commun (Lond)*. (2024) 44:654–69. doi: 10.1002/cac2.12547

18. Tamura K, Hasegawa K, Katsumata N, Matsumoto K, Mukai H, Takahashi S, et al. Efficacy and safety of nivolumab in Japanese patients with uterine cervical cancer, uterine corpus cancer, or soft tissue sarcoma: Multicenter, open-label phase 2 trial. *Cancer Sci*. (2019) 110:2894–904. doi: 10.1111/cas.14148

19. Xia L, Zhou Q, Gao Y, Hu W, Lou G, Sun H, et al. A multicenter phase 2 trial of camrelizumab plus famitinib for women with recurrent or metastatic cervical squamous cell carcinoma. *Nat Commun*. (2022) 13:7581. doi: 10.1038/s41467-022-35133-4

20. Wang Y, Zhao J, Liang H, Liu J, Huang S, Zou G, et al. Efficacy and safety of sintilimab plus albumin-bound-paclitaxel in recurrent or metastatic cervical cancer: a multicenter, open-label, single-arm, phase II trial. *EClinicalMedicine*. (2023) 65:102274. doi: 10.1016/j.eclinm.2023.102274

21. Shi Y. Regulatory mechanisms of PD-L1 expression in cancer cells. *Cancer Immunol Immunother*. (2018) 67:1481–9. doi: 10.1007/s00262-018-2226-9

22. Friedman JM. "Neurofibromatosis 1" In: Adam MP, Feldman J, Mirzaa GM, Pagon RA, Wallace SE, Bean LJH, et al, editors. *GeneReviews(®)*. Seattle, WA, United States: University of Washington, Seattle (1993).

23. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med*. (2012) 366:2443–54. doi: 10.1056/NEJMoa1200690

24. Powles T, Eder JP, Fine GD, Braiteh FS, Loriot Y, Cruz C, et al. MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature*. (2014) 515:558–62. doi: 10.1038/nature13904

25. Garon EB, Rizvi NA, Hui R, Leighl N, Balmanoukian AS, Eder JP, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med*. (2015) 372:2018–28. doi: 10.1056/NEJMoa1501824

26. Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N Engl J Med*. (2015) 373:1627–39. doi: 10.1056/NEJMoa1507643

27. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med*. (2010) 363:711–23. doi: 10.1056/NEJMoa1003466

28. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med*. (2015) 373:1803–13. doi: 10.1056/NEJMoa1510665

29. Robert C, Long GV, Brady B, Dutriaux C, Maio M, Mortier L, et al. Nivolumab in previously untreated melanoma without BRAF mutation. *N Engl J Med*. (2015) 372:320–30. doi: 10.1056/NEJMoa1412082

30. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature*. (2014) 515:568–71. doi: 10.1038/nature13954

31. Herbst RS, Soria JC, Kowanetz M, Fine GD, Hamid O, Gordon MS, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature*. (2014) 515:563–7. doi: 10.1038/nature14011

32. Bellmunt J, de Wit R, Vaughn DJ, Fradet Y, Lee JL, Fong L, et al. Pembrolizumab as second-line therapy for advanced urothelial carcinoma. *N Engl J Med*. (2017) 376:1015–26. doi: 10.1056/NEJMoa1613683

33. Ferris RL, Blumenschein G Jr., Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab for recurrent squamous-cell carcinoma of the head and neck. *N Engl J Med*. (2016) 375:1856–67. doi: 10.1056/NEJMoa1602252

34. Antonia SJ, Villegas A, Daniel D, Vicente D, Murakami S, Hui R, et al. Durvalumab after chemoradiotherapy in stage III non-small-cell lung cancer. *N Engl J Med*. (2017) 377:1919–29. doi: 10.1056/NEJMoa1709937

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 22 August 2024
DOI 10.3389/fimmu.2024.1427124

OPEN ACCESS

EDITED BY
Raquel Alarcon Rodriguez,
University of Almeria, Spain

REVIEWED BY
Jiahao Ma,
Xinxiang University, China
Yushun Gao,
Chinese Academy of Medical Sciences and Peking Union Medical College, China

*CORRESPONDENCE
Shenghua Liu
liushenghuafy@163.com
Haowen Jiang
urology_hs@163.com

†These authors have contributed equally to this work

RECEIVED 03 May 2024
ACCEPTED 29 July 2024
PUBLISHED 22 August 2024

CITATION
Chen Z, Hu J, Ou Y, Ye F, Li W, Liu S and Jiang H (2024) Integrating single-cell transcriptomics to reveal the ferroptosis regulators in the tumor microenvironment that contribute to bladder urothelial carcinoma progression and immunotherapy. *Front. Immunol.* 15:1427124. doi: 10.3389/fimmu.2024.1427124

Integrating single-cell transcriptomics to reveal the ferroptosis regulators in the tumor microenvironment that contribute to bladder urothelial carcinoma progression and immunotherapy

Ziang Chen[1,2†], Jia Hu[3†], Yuxi Ou[1,2], Fangdie Ye[1,2], Weijian Li[1,2], Shenghua Liu[1*] and Haowen Jiang[1,2,4*]

[1]Department of Urology, Huashan Hospital, Fudan University, Shanghai, China, [2]Fudan Institute of Urology, Huashan Hospital, Fudan University, Shanghai, China, [3]Department of Anesthesiology, The Second Affiliated Hospital of Anhui Medical University, Anhui Medical University, Hefei, Anhui, China, [4]National Clinical Research Center for Aging and Medicine, Huashan Hospital, Fudan University, Shanghai, China

Background: Ferroptosis, as a novel form of programmed cell death, plays a crucial role in the occurrence and development of bladder cancer (BCa). However, the regulatory mechanisms of ferroptosis in the tumor microenvironment (TME) of BCa remain to be elucidated.

Methods: Based on single-cell RNA (scRNA) transcriptomic data of BCa, we employed non-negative matrix factorization (NMF) dimensionality reduction clustering to identify novel ferroptosis-related cell subtypes within the BCa TME, aiming to explore the biological characteristics of these TME cell subtypes. Subsequently, we conducted survival analysis and univariate Cox regression analysis to explore the prognostic significance of these cell subtypes. We investigated the relationship between specific subtypes and immune infiltration, as well as their implications for immunotherapy. Finally, we discovered a valuable and novel biomarker for BCa, supported by a series of *in vitro* experiments.

Results: We subdivided cancer-associated fibroblasts (CAFs), macrophages, and T cells into 3-5 small subpopulations through NMF and further explored the biological features. We found that ferroptosis played an important role in the BCa TME. Through bulk RNA-seq analysis, we further verified that ferroptosis affected the progression, prognosis, and immunotherapy response of BCa by regulating the TME. Especially ACSL4+CAFs, we found that high-level infiltration of this CAF subtype predicted worse prognosis, more complex immune infiltration, and less response for immunotherapy. Additionally, we found that this type of CAF was associated with cancer cells through the PTN-SDC1 axis, suggesting that SDC1 may be crucial in regulating CAFs in cancer cells. A series of *in vitro* experiments confirmed these inferences: SDC1 promoted the progression of BCa. Interestingly, we also discovered FTH1+ macrophages, which were closely

related to SPP1+ macrophages and may also be involved in the regulation of BCa TME.

Conclusion: This study revealed the significant impact of ferroptosis on bladder cancer TME and identified novel ferroptosis-related TME cell subpopulations, ACSL4+CAFs, and important BCa biomarker SDC1.

KEYWORDS

single-cell, tumor microenvironment, bladder cancer, ferroptosis, immunotherapy, prognosi

Introduction

Bladder cancer ranks among the most prevalent malignant tumors affecting the urinary system. According to statistics, since 2023, there have been nearly 500,000 new cases of bladder cancer and 200,000 deaths globally each year (1, 2). Approximately 75% of the new cases occur in males. Smoking is the most common risk factor for BCa (3). With the rapid increase in tobacco consumption, the rising incidence of bladder cancer poses a significant burden on global healthcare (4). The emergence of immunotherapy has brought new hope for managing BCa patients, with immune checkpoint inhibitors (ICI) being the main treatment modality. They primarily work by inhibiting immune checkpoints to reduce immune suppression and promote anti-tumor immunity. Compared to traditional chemotherapy, ICI therapy offers higher precision and specificity (5).

Ferroptosis is a novel form of programmed cell death, characterized by iron-dependent lipid peroxidation and excessive reactive oxygen species (ROS) accumulation, leading to cell death (6). When intracellular glutathione (GSH) is depleted, glutathione peroxidase 4 (GPX4) becomes inactivated, resulting in the accumulation of lipid peroxides and subsequent cell death (7). Ferroptosis has been associated with the development and progression of multiple types of cancers (8). Some studies have reported that ferroptosis influences cancer development and progression by mediating cancer-associated fibroblasts (9). However, the regulatory mechanisms and targets of ferroptosis in the BCa tumor microenvironment remain unclear.

Single-cell RNA sequencing enables researchers to study tumors with precise details. We can identify novel tumor microenvironment cell subtypes based on single-cell RNA sequencing data and analyze their biological characteristics and prognostic significance. Cell-cell communication analysis reveals important signaling pathways between these cell subtypes and cancer cells, and identifies new targets.

By applying non-negative matrix factorization to single-cell RNA sequencing data, we identify novel ferroptosis-related cell subpopulations within the bladder cancer tumor microenvironment. Combining with classic biological function signatures, we explore the biological characteristics of these cell subpopulations. Cell-cell communication analysis reveals important signaling pathways between these cell subpopulations and cancer cells. Based on bulk RNA-seq data, we evaluate the prognostic significance of these cell subpopulations' infiltration. Finally, we identify a subtype of cancer-associated fibroblasts (CAFs), ACSL4+CAFs, which impact patients' overall survival (OS) and sensitivity to immunotherapy. Cell-cell communication analysis reveals SDC1 as an important target on cancer cells interacting with ACSL4+CAFs. Subsequent *in vitro* experiments confirm that SDC1 promotes the proliferation, migration, and invasion of BCa cells.

Materials and methods

Acquisition of data

To explore the microenvironment heterogeneity of bladder cancer, single-cell RNA transcriptome data was obtained from the Sequence Read Archive (SRA) (PRJNA662018) (https://www.ncbi.nlm.nih.gov/sra), which contained 8 bladder cancer and 3 normal bladder mucosa tissues (10). To investigate the relationship between ferroptosis and the TME of bladder cancer, we obtained the most frequently studied ferroptosis marker genes from the FerrDb database (http://www.zhounan.org/ferrdb/current/). Besides, six bulk-RNA sequencing datasets were employed to lucubrate the impact of specific ferroptosis subpopulations on patients' survival, which were obtained from The Cancer Genome Altas (TCGA-BLCA) (https://portal.gdc.cancer.gov/) and the GEO database (GSE48075, GSE32894, GSE31684, GSE160693, GSE13507) (https://www.ncbi.nlm.nih.gov/geo/) (11–15). All the data upon which this study is based are publicly available.

The processing and visualization of single-cell sequencing data

The CellRanger (v.3.0.1) software was employed to filter and read align the raw single-cell FASTQ data, and feature barcode

unique molecular identifier (UMI) matrices were generated based on the human reference genome GRCh38. The Seurat package (v.4.30.1) was used to process the single-cell RNA sequencing expression matrix, cells with gene expression counts less than 200 and cells where mitochondrial gene expression accounted for more than 15% were filtered out. The NormalizeData function was employed to normalize the expression matrix and the RunPCA function was applied to compute the principal components. The UMAP (Uniform Manifold Approximation and Projection) algorithm was utilized to visualize the single-cell RNA sequencing data. Finally, six cell types were identified.

Identification of the marker genes of ferroptosis-related cell types in TME

The non-negative matrix factorization algorithm was conducted to observe the effect of ferroptosis marker gene expression on TME cell types based on the NMF package (v.0.26). Next, The following criteria were used to determine representative markers for each NMF cell subtype in the FindAllMarkers function: logFC > 0.8, minimum proportion greater than 30%. The cell subpopulations with logFC of ferroptosis marker genes less than 0.5 will be defined as "Non-Ferr," while those with logFC greater than 0.5 but less than 0.8 will be defined as "Unclear."

Function enrichment analysis of ferroptosis-related cell subpopulations

To investigate the biological characteristics of ferroptosis-related cell subpopulations, we performed GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis based on the clusterProfiler package (v.4.8.3) (16). To explore the metabolic activity of macrophages, we calculated the metabolism enrichment scores based on the scMetabolism package (v.0.2.1). Besides, the AUCell package (v.1.22.0) was utilized to quantify the biology activities.

SCENIC analysis for ferroptosis-related cell subpopulations

To clarify the gene regulatory network of transcription factors (TFs) in TME cell subpopulations, the SCENIC package (v.1.3.1) was employed. Two gene-motif rankings (hg19-500bp-upstream-7species.mc9nr.feather and hg19-tss-centered-10kb-7species.mc9nr.feather) were downloaded from the RcisTarget database (https://github.com/aertslab/RcisTarget) to identify the transcription start site (TSS). Then, potential TF-target relationships were recognized and a co-expression gene network was constructed. Only TFs with False Discovery Rate (FDR) <0.05 were considered in this study.

Cell-cell communication analysis for ferroptosis-related cell subpopulations

The CellChat package (v.1.6.1) was utilized to construct the intratumor communications networks. CellChatDB.human was employed to evaluate the signaling pathway inputs and outputs between TME cell subpopulations and cancer cells. Next, the computeCommunProPathway and aggregateNet functions were used to calculate the cell-cell communication network and communication strength. Finally, the netVisual_bubble function was performed to visualize ligand-receptor interactions based on the human ligand-receptor pairs database.

Pseudotime trajectory analysis for ferroptosis-related cell subpopulations

To explore the role of the ferroptosis marker genes in the trajectory of cellular development and differentiation, we employed pseudotime trajectory analysis for TME cell subpopulations based on the Monocle package (v.2.22.0). Highly variable genes were filtered according to the following criteria: mean_expression ≥ 0.1 and dispersion_empirical ≥ 1*dispersion_fit. The method for dimensionality reduction was DDRTree. Next, the plot_pseudotime_heatmap function was employed to show the pseudotime heatmap, and the plot_cell_trajectory function was used to illustrate the dynamic expression of ferroptosis marker genes in TME.

Assessment of immune infiltration and ICI therapy

Four algorithms were utilized to compare immune cell infiltration across different groups, including CIBERSORT, XCell, EPIC, and Quantiseq. Subsequently, the ESTIMATE package (v.1.0.13) was utilized to calculate the abundance of TME components. The Tracking Tumor Immunophenotype (TIP) algorithm (http://biocc.hrbmu.edu.cn/TIP/) was employed to assess the cancer immunity cycle. The online website Tumor Immune Dysfunction and Exclusion (TIDE) (http://tide.dfci.harvard.edu/login/) was utilized to assess the ICI response of bladder cancer patients, as well as the Subclass Mapping (Submap) algorithm.

Survival analysis for specific TME cell subpopulations

The ssgsea function was employed to calculate the infiltration levels of specific cell subtypes in bladder cancer patients based on the GSVA package (v.1.42.0). To investigate the prognostic significance of specific TME cell subpopulations, the survival (v.3.2.13) and survminer (v.0.4.9) package were employed to conduct the survival analysis. Patients were divided into two

groups according to the optimal cutoff. All survival analyses in this study were subjected to log-rank tests.

Cell culture, transfection, and interference

The study utilized human bladder cancer cells (UM-UC-3 and T24) from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). UM-UC-3 was cultured in high glucose DMEM and T24 was cultured in RPMI-1640 with 10% fetal bovine serum (Gibco, USA) and 1% streptomycin/penicillin (Thermo Fisher Scientific, USA) in an incubator at 37°C and 5% CO2. The cells were transferred upon reaching a cell density of 70-80%, and the medium was changed daily.

The siRNA targeting SDC1 (siSDC1) lentivirus was purchased from GeneChem (shanghai, China) to suppress the SDC1 gene of T24 and UM-UC-3 cells. The siSDC1 was transfected into T24 and UM-UC-3 cells with polyethylene, then the cells were screened with puromycin. The siRNA sequences can be obtained in Supplementary Table 1.

Colony formation assay

UM-UC-3 and T24 cells were seeded onto 6-well plates at a density of 1000 cells per well and allowed to culture for 2 weeks until the formation of cell colonies. Subsequently, the cells were washed three times with phosphate-buffered solution (PBS) (Yeasen, China), fixed with 4% methanol for 15 minutes, stained with 0.5% crystal violet solution for 30 minutes, and analyzed using ImageJ software.

Wound-healing assay

To analyze cell direct migration, a wound-healing assay was conducted. UM-UC-3 and T24 cells were inoculated in a 6-well plate and cultured until reaching 70%-80% density. The cell monolayer was gently scratched using the tip of a sterile 200μL pipette after removing the medium. Subsequently, the wells were rinsed twice with PBS, and serum-free medium was added for continued culture. Images were captured at 0-, 12-, and 24-hours post-scratching and analyzed using ImageJ software.

Cell viability detection

A CCK-8 assay kit (Biosharp, China) was used to assess cell viability. After the intervention, cells were seeded on 96-well plates and incubated at 37°C with 5% CO2. UM-UC-3 and T24 cells were treated with a diluted CCK-8 solution for 2h. The absorbance values at 450nm were quantified using a microplate reader. (Thermo Fisher Scientific, USA).

EdU assay

Following the manufacturer's instructions, the EdU detection assay (Beyotime, China) was used to detect different cell states after a series of operations, DNA synthesis, and cell proliferation was observed through fluorescence microscopy (Olympus, USA).

Flow cytometric analysis

The Annexin V-FITC/PI Apoptosis (Beyotime, China) kit was used to detect the apoptosis rate of BCa cells, including early and terminal apoptosis. BD flow cytometry (BD FACSLyric, USA) was employed to analyze cell samples.

Transwell migration and invasion assay

The transwell assays were performed to observe the migration and invasion ability. T24 and UM-UC-3 cells were seeded at a density of 2×10^4 cells per well in the upper chambers, with 200μL serum-free medium, while the lower chambers contained 800μL of medium supplemented with 10% serum. To perform the invasion assay, 50mg/L Matrigel glue was covered in the upper chambers. After placing the transwell chambers (Corning, USA) in a 37°C, 5% CO2 incubator for 48 hours, 4% methanol was used to fix cells for 30 minutes, and 0.5% crystal violet was employed to stain for 30 minutes. Finally, the results can be obtained by taking photographs and counting.

Statistical analysis

All data processing and statistical analysis performed in this study were based on R software (v.4.3.1). To verify the differences among various groups, diverse tests (Wilcoxon rank-sum test, Fisher exact test, Student's t-test, Kruskal-Wallis test) were performed. In correlation analysis, the Pearson test was used to verify the statistical significance. All experiment data were presented as the mean ± SD, and GraphPad Prism 8 software was employed to analyze these experiment data. In this study, only a two-sided p-value below 0.05 was considered statistically significant.

Results

The landscape of ferroptosis-related genes in TME of BCa

A total of 83,146 cells were mapped onto the cell atlas and annotated into six major cell types (Figure 1A). Subsequently, the classical cell markers were displayed in the cell atlas according to their expression levels (Figure 1B). Additionally, the cell atlas comprised tissue samples from a total of 11 patients, including three normal tissues and eight cancer tissues, enabling observation

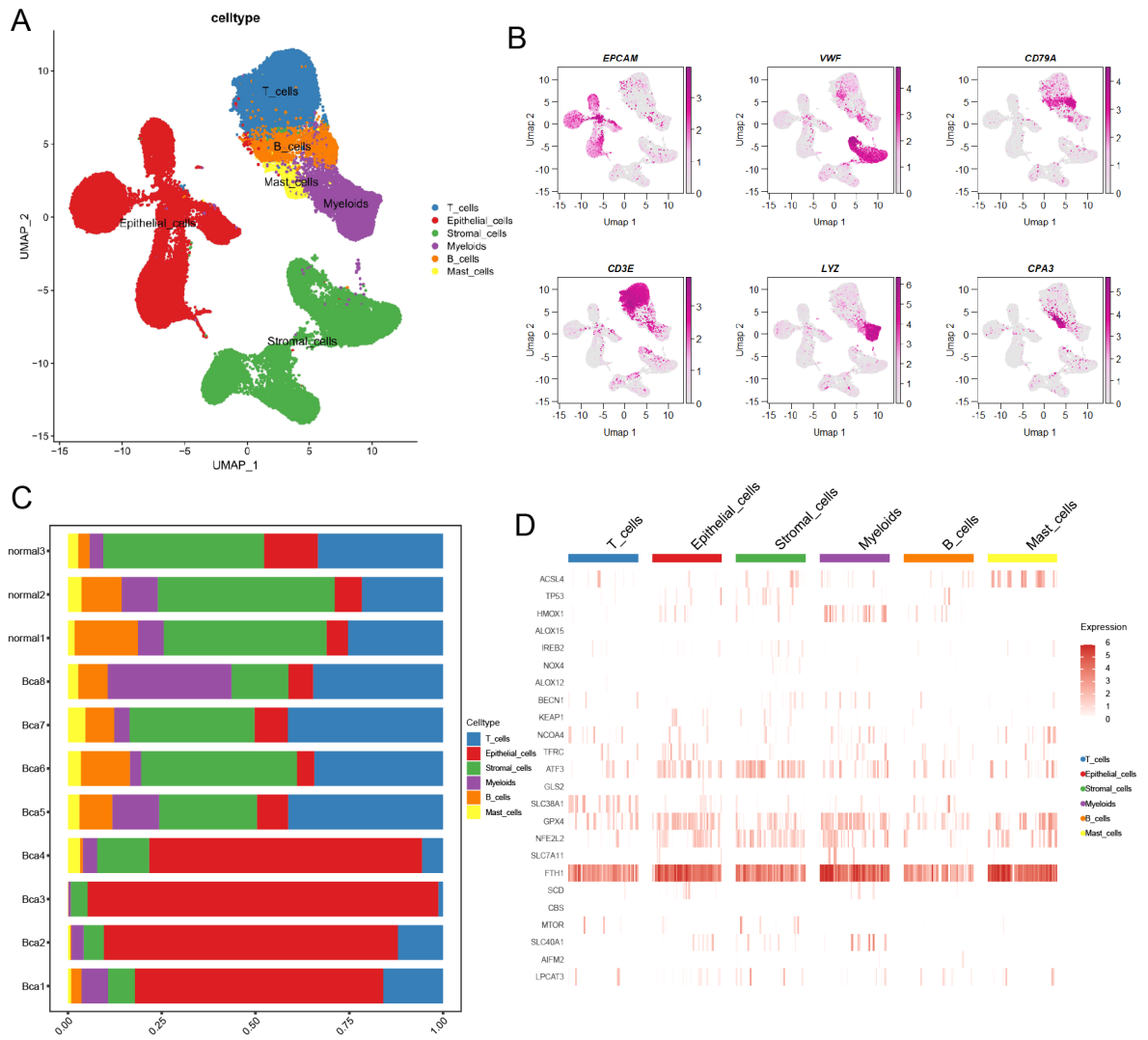


FIGURE 1
Overview of ferroptosis-related marker genes in scRNA transcriptome data of BCa. **(A)** The landscape of main cell types was illustrated by UMAP. **(B)** The heatmap showed the classical marker genes in the BCa landscape. **(C)** A bar chart illustrated the proportion of different cell types across various samples. **(D)** The heatmap displayed the distribution of ferroptosis-related marker genes in main cell types.

and comparison of the cellular composition proportions across different samples (Figure 1C). Furthermore, the differential expression of ferroptosis-related genes across different cell types can be observed through the heatmap (Figure 1D), as a ferroptosis suppressor gene, GPX4 had higher expression in epithelial cells and myeloid cells across major cell types, which meant these two cell types had lower ferroptosis level.

Novel ferroptosis-related CAFs mediated TME of BCa

According to classical marker genes, stromal cells were annotated into 3 major cell types: endothelial cells, smooth muscle cells, and CAFs (Figure 2A; Supplementary Figure 1A). In this section, we primarily focused on CAFs. After undergoing NMF dimensionality reduction and clustering, CAFs were subdivided into 5 novel cell subpopulations (Figure 2B). We found that ACSL4+CAFs interacted with glucocorticoid activities and myeloid leukocyte activation (Figure 2C). KEGG analysis yielded that ACSL4+CAFs were associated with estrogen signaling pathways, and ATF+CAFs exhibited active oxidative phosphorylation (Figure 2D). To explore the relationship between CAF subpopulations and cancer cells, cell-cell communication analysis yielded that ACSL4+CAFs had the strongest communication with cancer cells among these novel ferroptosis-related CAFs (Figure 2E). Furthermore, ligands-receptors analysis uncovered that ACSL4+CAFs were strongly associated with the EGF (Epidermal Growth Factor) and MIF (Macrophage migration Inhibitory Factor) signaling pathway (Supplementary Figures 1B, C).

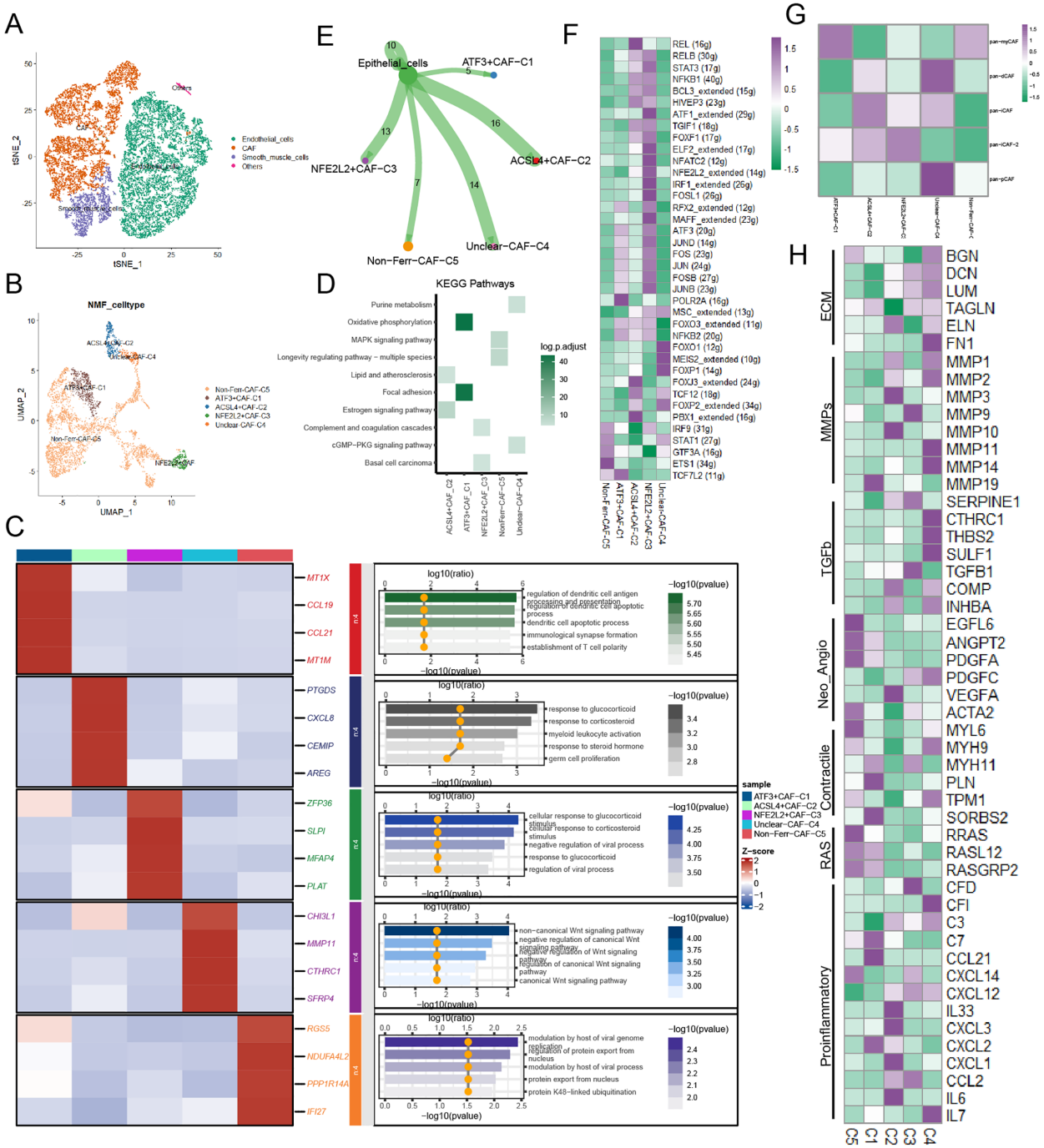


FIGURE 2
The landscape of ferroptosis-related marker genes in CAFs. **(A)** Presence of CAFs in stromal cells. **(B)** The ferroptosis-related CAF subpopulations were illustrated by UMAP. **(C)** The heatmap displayed the differential genes and biological functions of each subpopulation. **(D)** Activation of the KEGG pathway of each subpopulation. **(E)** The cell-cell communication strength between the ferroptosis-related CAF subpopulations and cancer cells. **(F)** Transcriptional regulatory factors for each cell subpopulation. **(G)** Correlations between the ferroptosis-related CAF subpopulations and classical CAF signatures ($p < 0.05$). **(H)** Heatmap illustrated the distinct average expression levels of prevalent signaling pathway genes among the five subpopulations of CAFs associated with ferroptosis, encompassing Proinflammatory, RAS, Contractile, Neo-Angio, TGFb, MMPs, and ECM.

TFs of these novel ferroptosis-related CAFs were illustrated in the gene regulatory network based on the SCENIC analysis, and FOXJ3, TCF12, FOXP2, and PBX1 were upregulated in ACSL4 +CAFs (Figure 2F). Next, to further investigate the biology characteristics, we assessed the correlation among Pan-CAF signatures and found that ACSL4+CAFs were similar to inflammatory CAF (pan-iCAF) (Figure 2G). Then, the heatmap highlighted the same conclusion: ACSL4+CAFs were closely related

to inflammation within BCa (Figure 2H). Pseudotime trajectory analysis exhibited the development of CAFs and dynamic expression of marker genes, the heatmap reflected ferroptosis-related genes played important roles in CAFs within BCa (Supplementary Figure 1D), we also found that ACSL4+CAFs were at the beginning of trajectory and ATF+CAFs were at the end of trajectory (Supplementary Figures 1E, F).

Ferroptosis-related macrophages exhibit particular biological features

Macrophages played a crucial role in TME in the progress of BCa, unique macrophage subpopulations exhibited distinct biological features. Firstly, myeloid cells were annotated into three major cell types based on classical marker genes (Supplementary Figures 2A, B). We identified three novel macrophage subpopulations through NMF dimensionality reduction (Figure 3A). By integrating specific macrophages-related signatures reported in previous literature, we gained a deeper understanding of the biological characteristics of these macrophage subpopulations and their potential roles in BCa, HMOX1+macrophages shared biological similarities with C1q+ macrophages and M2 macrophages, and FTH1+macrophages were closely related to SPP1+macrophages (Figures 3B, C). Then, pseudotime trajectory analysis yielded that these ferroptosis-related marker genes, especially FTH1, played a crucial role in the development of macrophages (Supplementary Figures 2C, D). Subsequently, cell-cell communication analysis indicated that HMOX1+macrophages had the most interactions with cancer cells, while FTH1+macrophages had the fewest (Figure 3D, Supplementary Figure 2E). Signaling pathway analysis uncovered that the SPP1 signaling pathway was upregulated in FTH1+macrophages, and IL6 and IL10 were overexpressed in HMOX1+macrophages (Supplementary Figures 2F, G). Next, the regulatory network

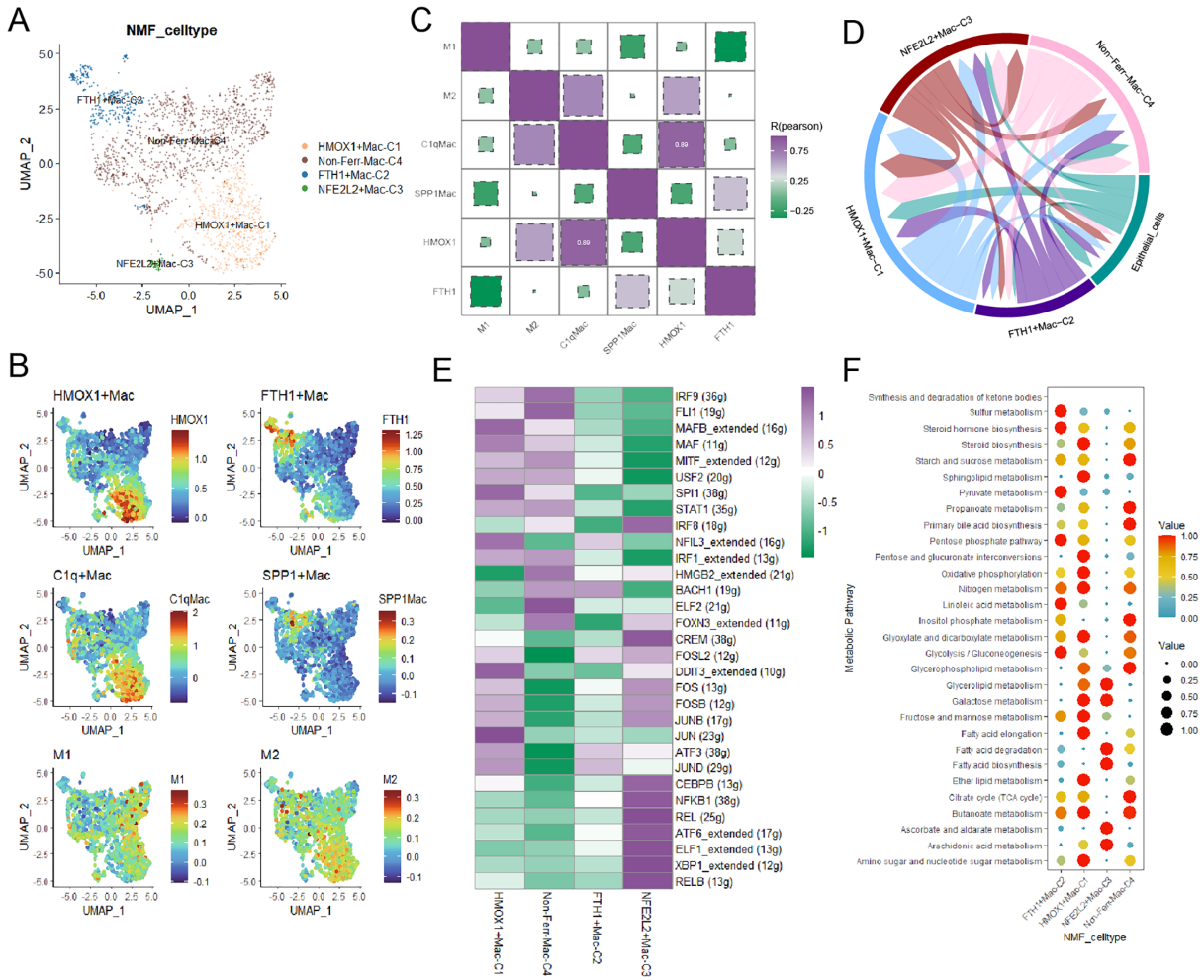


FIGURE 3
The landscape of ferroptosis-related marker genes in macrophages. **(A)** The ferroptosis-related macrophage subpopulations were illustrated by UMAP. **(B)** The heatmap demonstrated the distinct biological features in macrophages. **(C)** Correlations between the ferroptosis-related macrophage subpopulations and classical macrophage signatures. **(D)** The circular plot visualized the strength of cell-cell communications between macrophages and cancer cells. **(E)** Transcriptional regulatory factors for each cell subpopulation. **(F)** The metabolism landscape of ferroptosis-related macrophage subpopulations.

illustrated some TFs, such as JUN, JUNB, and ATF3, were upregulated in HMOX1+macrophages (Figure 3E).

To deeply understand the biology characteristics of these macrophage subpopulations, we performed GO enrichment analysis, and the results yielded that HMOX1+macrophages were associated with protein refolding, those genes upregulated in FTH1 +macrophages were enriched in chemokine-mediated signaling pathway and cellular response to chemokine (Supplementary Figure 2H). Finally, the scMetabolism package showed that FTH1 +macrophages had active sulfur and pyruvate metabolism, and HMOX1+macrophages exhibited vigorous steroid biosynthesis and oxidative phosphorylation (Figure 3F).

The landscape of ferroptosis-related T cells in TME

T cells were divided into four cell types according to their distinct gene expression (Figure 4A), and the bar plot illustrated their proportions in BCa patient samples (Figure 4B). These T cell

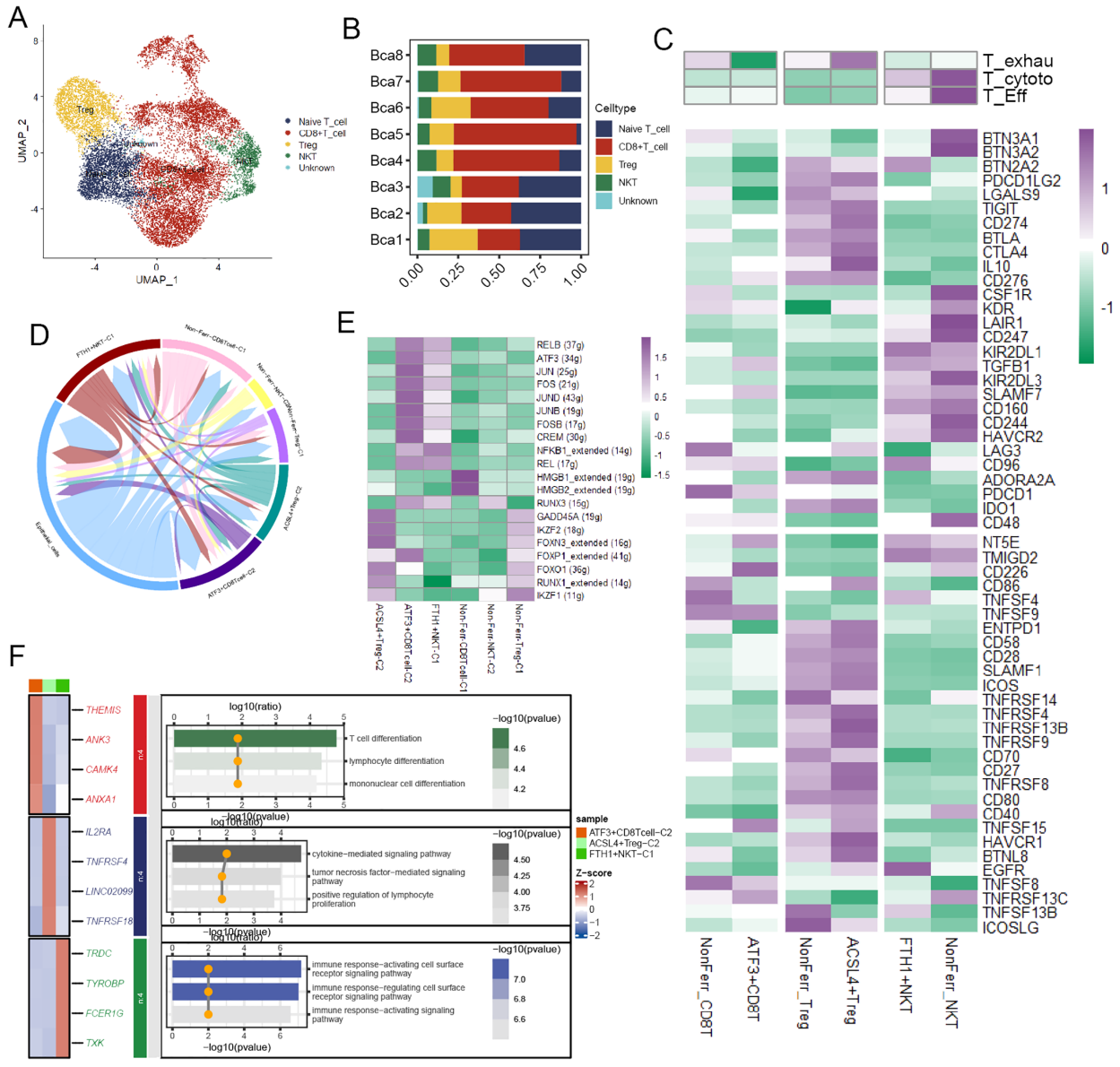


FIGURE 4
The biological features of the ferroptosis-related T cell subpopulations. **(A)** The landscape of T cells in BCa. **(B)** A bar chart illustrated the proportion of different T cell types across various samples. **(C)** The heatmap depicted the markedly distinct characteristics across ferroptosis-related subpopulations of T cells, encompassing CD8+T cells, NKTs, and Tregs. These features include T exhaustion score, T cytotoxic score, T effector score, various immune co-stimulators, and co-inhibitors (Kruskal–Wallis test, $p < 0.001$). **(D)** The circular plot visualized the strength of cell-cell communications between T cells and cancer cells. **(E)** Transcriptional regulatory factors for each cell subpopulation. **(F)** The heatmap displayed the differential genes and immune pathways of each subpopulation.

types were annotated into six cell subpopulations after undergoing NMF dimensionality reduction. To deeply unravel the molecular characteristics, we calculated some T cell-related signature scores in these cell subpopulations. The result demonstrated that compared to ATF3+CD8T cells, NonFerr_CD8T cells exhibited higher T cell exhaustion scores, while ACSL4+Tregs had higher scores than NonFerr_Tregs. Interestingly, compared to FTH1+NKT (Natural Killer T cells), NonFerr_NKT were biologically analogous to effector T cells and cytotoxic T cells (Figure 4C). Subsequently, cell-cell communication analysis unraveled that ATF3+CD8T cells and FTH1+NKT occupied most interactions with the cancer cells (Figure 4D). Next, we observed that some TFs, such as IKZF1, RUNX1, FOXO1, FOXP1, and FOXN3, were upregulated in ACSL4 +Tregs, and in ATF3+CD8T cells, JUN, FOS, JUND, JUNB, FOSB were overexpressed (Figure 4E). Finally, enrichment analysis yielded that ATF3+CD8T cells were associated with lymphocyte differentiation, ACSL4+Tregs were associated with tumor necrosis factor and cytokine regulation pathways, and FTH1+NKT were associated with cell surface receptor signaling pathways (Figure 4F).

Prognostic significance of specific ferroptosis-related TME cell subpopulations

The FindAllMarkers function was employed to calculate differential expression genes (DEGs) of those cell subpopulations, and we extracted the top 50 for each subpopulation to obtain the signatures of specific ferroptosis-related TME cell subpopulations. The survival analysis was performed to validate the prognostic significance of the signatures of those subpopulations, and we found that a few TME cell subpopulations were closely associated with overall survival (OS) in the TCGA cohort, including CAFs and macrophages (Figures 5A–C). Then, the same results were validated again in the GSE13507 cohort (Figures 5D–F). subsequently, the survival analysis performed in the GSE32894 yielded that only ACSL4+CAFs related to OS (Figure 5G).

To ensure the rigor of the results, Cox regression analysis was implemented using six BCa cohorts. The results yielded a high infiltration of ACSL4+CAFs was closely associated with adverse effects on patients' OS (Figure 5H).

ACSL4+CAFs impacted immune infiltration and ICI therapy

To explore the impact of ACSL4+CAFs on immune infiltration and response to ICI, we calculated immune infiltration in the TCGA cohort and evaluated ICI response. All patients were divided into two groups based on the optimal cutoff; a total of four algorithms were employed to assess the immune cell infiltration in BCa, and the heatmap illustrated a remarkable difference between the two groups, especially macrophages, Tregs and NKT (Figure 6A). the ESTIMATE algorithm was implemented to calculate the abundance of TME components, we found that high ACSL4+CAFs group had more complex TME than low ACSL4+CAFs group (Figures 6B–D). To further explore the difference between the two groups in the immunity cycle, the TIP algorithm yielded that the high ACSL4 +CAFs group exhibited a stronger capacity for immune cell recruitment (Figure 6E). To further investigate the difference between the two groups in ICI therapy response, the TIDE algorithm revealed that patients who did not respond to ICI therapy had higher ACSL4+CAFs infiltration (Figure 6F), similarly, there was a higher proportion of non-responsive patients to ICI treatment within high ACSL4+CAFs group (Figure 6G). Next, we discovered that the low ACSL4 +CAFs group appeared to be more sensitive to anti-PD-1 treatment (Figure 6H).

The close association between ACSL4 +CAFs and BCa cells

We performed cell-cell communication analysis between cancer cells and a few specific ferroptosis-related TME cell subpopulations, and the results yielded that SDC1 was a crucial target associated with ACSL4+CAFs (Figure 7A). The GEPIA website showed that BCa tissues had a higher SDC1 expression than normal tissues (Figure 7B), and we performed survival analysis within two cohorts to investigate the prognostic significance of SDC1, which revealed that higher SDC1 expression had an adverse impact on patients' OS (Figures 7C, D). Subsequently, the correlation analysis showed that SDC1 negatively correlated to CD8+ T cells in BCa (Figure 7E).

SDC1 promoted the proliferation and invasion ability of BCa cells

A series of *in vitro* experiments were performed to verify the function of SDC1 in BCa cells. CCK-8 assays yielded that SDC1 enhanced the viability of BCa cells (Figures 8A, B), and the colony formation assays showed that SDC1 remarkably increased the number of colonies (Figure 8C). To assess the different invasion abilities of BCa cells, the transwell assays were performed to display that the invasion ability was decreased after SDC1 was knocked down (Figure 8D). The wound-healing assays illustrated that SDC1 promoted the direct migration of BCa cells (Figure 8E).

The EdU assay was employed to quantify the proliferation level of cells, and the results indicated that knocking down SDC1 significantly inhibited the proliferation capability of BCa cells (Figure 9A). Furthermore, flow cytometric analysis revealed that knocking down SDC1 led to a significant increase in the number of apoptotic BCa cells, suggesting that SDC1 plays a role in inhibiting apoptosis in BCa cells (Figure 9B).

Discussion

Ferroptosis, as a novel form of cell death, has played a significant role in various cancers (17). Lipid peroxidation (LPO) and increased iron load have served as important signals for

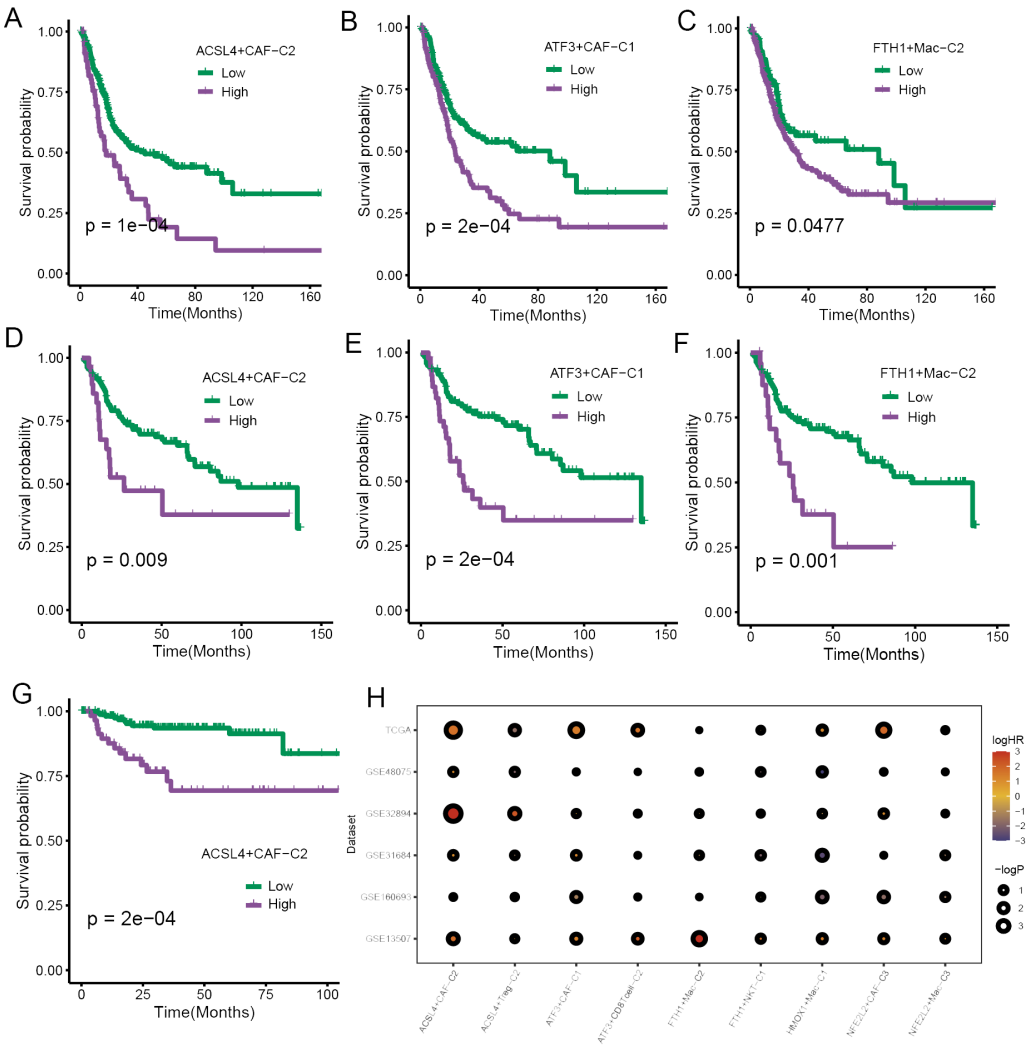


FIGURE 5
The prognostic significance of specific ferroptosis-related TME cell subtypes. **(A–C)** The K-M curve plots demonstrated the impact of specific TME cell infiltration on OS in the TCGA cohort. **(D–F)** The K-M curve plots demonstrated the impact of specific TME cell infiltration on OS in the GSE13507 cohort. **(G)** The K-M curve plots demonstrated the impact of specific TME cell infiltration on OS in the GSE32894 cohort. **(H)** The bubble heatmap illustrated the prognostic significance in univariate Cox regression analysis.

ferroptosis (18). Recent research suggests that ferroptosis could be a crucial target for cancer therapy (19). Apatinib, a tyrosine kinase inhibitor, has been reported to induce ferroptosis in gastric cancer cells by inhibiting GPX4 activity (20). P53 mutations, common events in tumor development, have been shown to inhibit ferroptosis and promote tumor progression (21). Additionally, the direct use of ferroptosis inducers as chemotherapy drugs holds promise. Erastin, a typical inducer of ferroptosis, can reduce the cellular synthesis of GSH by inhibiting SLC7A11, leading to increased LPO and subsequent ferroptosis (22). Some studies indicate that the anti-tumor efficacy of combined Erastin and platinum drugs is stronger than that of platinum drugs alone (23). Therefore, ferroptosis may offer a new therapeutic strategy for treating chemotherapy-resistant patients (24).

Some studies have indicated that when levels of free iron increase in bladder cancer cells, cell proliferation is inhibited, suggesting a close association between ferroptosis and bladder cancer cell proliferation (25). Recent studies have developed a novel targeted therapy approach for bladder cancer. CPNPs, a type of conjugated polymer nanoparticle carrying iron ions, can induce ferroptosis in cancer cells by releasing iron ions upon entry into tumor cells (26). It has been reported that CPNPs can kill 80% of cancer cells under high-dose conditions. These findings offer promising management strategies for patients with BCa.

TME, as a crucial component of bladder cancer, plays a significant role in tumor progression (27). CD8+ T cells, as the primary mediators of anti-tumor immunity, are also affected by ferroptosis (28). Literature suggests that tumor-derived CD8+ T

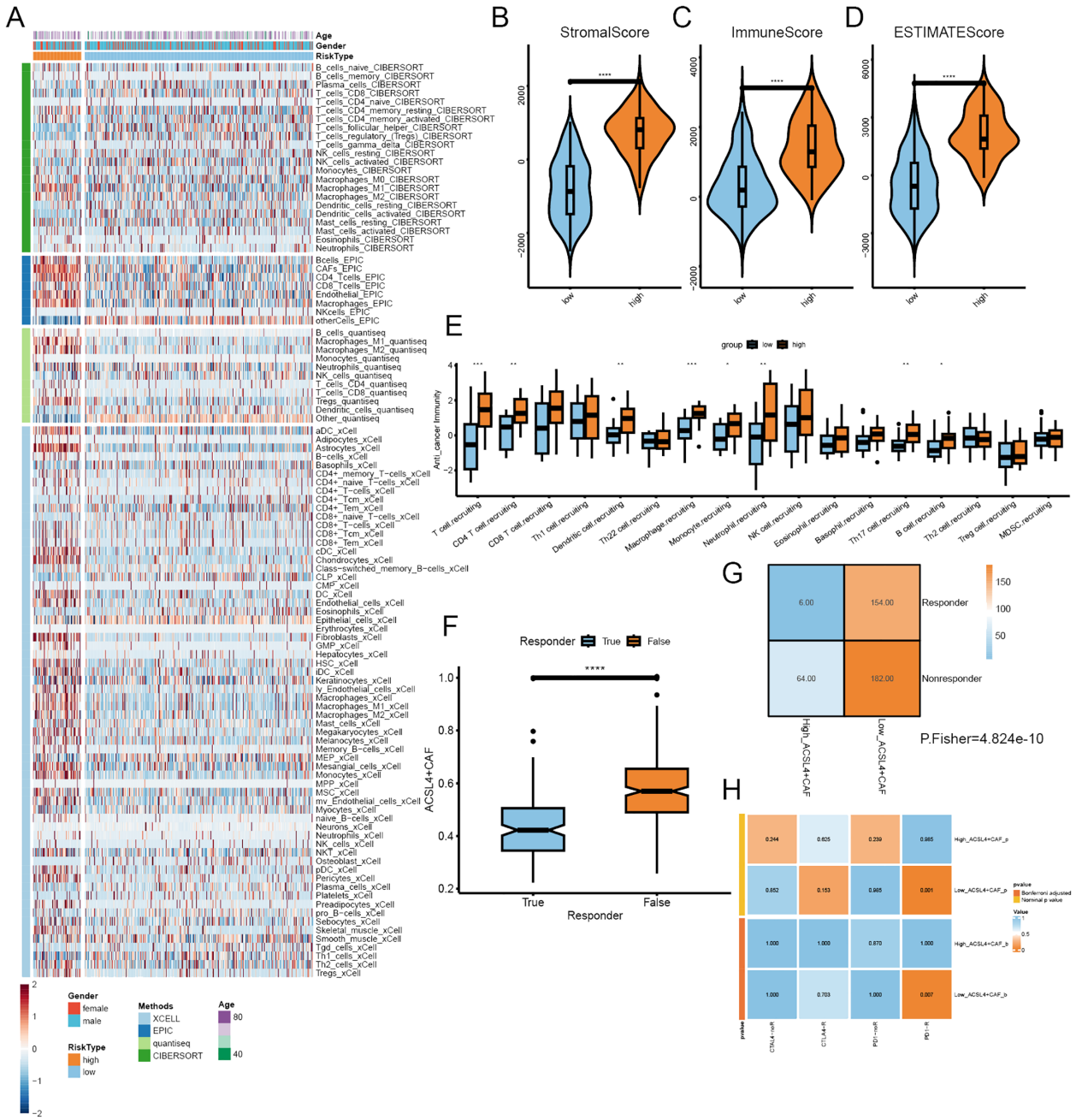


FIGURE 6
The immune infiltration of different ACSL4+CAFs groups. **(A)** The landscape of immune cells infiltration of TCGA cohort. **(B-D)** The distribution of immune components between the two groups. **(E)** Differences in the anti-tumor immunity cycle between the two groups. **(F)** Box plot showed the different ACSL4+CAFs infiltration levels between distinct immunotherapy responses in the TCGA cohort. **(G)** Contingency table between immunotherapy responses and ACSL4+CAFs infiltration based on TIDE in the TCGA cohort. **(H)** The contingency table presented the relationship between ACSL4+CAFs infiltration and optimal immunotherapy (anti-PD1, anti-CTLA4). *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

cells accumulate more LPO compared to lymph node-derived CD8 + T cells (29). Furthermore, studies have reported that overexpression of GPX4 in CD8+ T cells can protect them from the effects of ferroptosis, restore their secretion of cytotoxic factors, and increase the infiltration of CD8+ T cells within tumors (30). Tregs, classical immunosuppressive components in the TME, exhibit significantly lower levels of LPO than CD8+ T cells (31). Moreover, activation of TCR/CD28 in Tregs induces GPX4 expression, thereby inhibiting ferroptosis (32). These findings suggest that Tregs rarely undergo ferroptosis. Macrophages in bladder cancer are generally classified as anti-tumor M1 subtype or pro-tumor M2 subtype (33). Some iron-targeting nanoparticles have been developed to repolarize M2 macrophages into M1 subtype, assisting in tumor treatment (34). Recent studies indicate

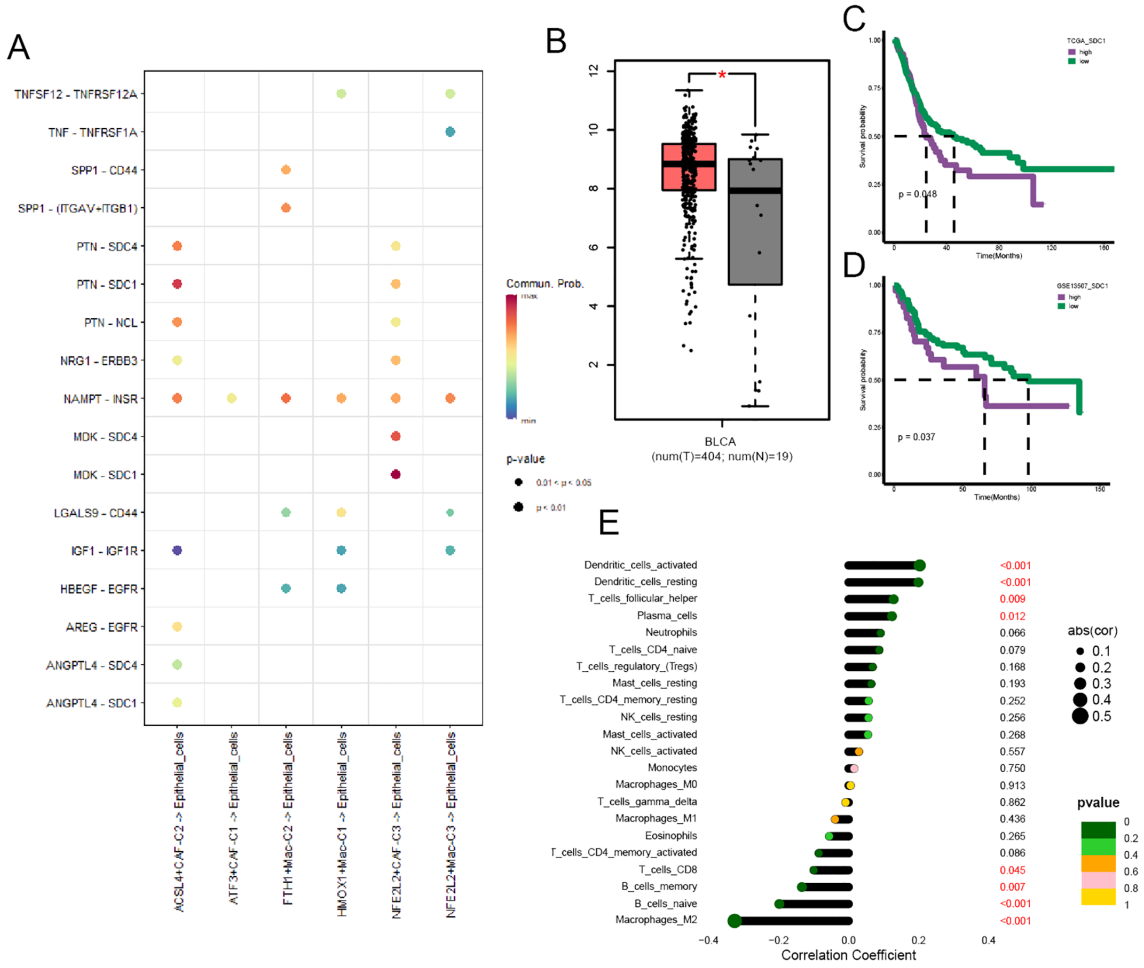


FIGURE 7
The prognostic significance of SDC1 in BCa. **(A)** The dot plot illustrated the communications between those specific TME cell subpopulations and cancer cells. **(B)** Box plot displaying the different expressions of SDC1 between tumor and normal tissue. **(C, D)** The K-M curve plot detected the prognostic significance of SDC1 among the TCGA and GSE13507 cohort. **(E)** The correlation between SDC1 and immune cell infiltration in BCa. *P < 0.05.

that tumor-derived NK cells express increased levels of proteins associated with ferroptosis and lipid peroxidation (35), with their mitochondria resembling those of ferroptotic cells. LPO leads to a metabolic imbalance in NK cells, causing functional impairment (36). In conclusion, ferroptosis is closely related to the bladder cancer TME.

Ferroptosis is closely linked to immunotherapy as well (37). Studies have found that cancer cells undergoing ferroptosis exhibit dual characteristics (38). On one hand, ferroptotic cancer cells release immunostimulatory signals, attracting macrophages, dendritic cells, and other immune cells to the tumor site (39). The enhanced immunogenicity also induces tumor-specific immune responses (40). On the other hand, it has been reported that ferroptotic cancer cells can release 8-hydroxy-2′-deoxyguanosine (8-OHdG), which promotes M2 polarization (41). These products originate from oxidative DNA damage (42). Cytotoxic CD8+ T cells secrete interferonγ (IFNγ), which inhibits SLC7A11 by activating the JAK/STAT1 pathway in cancer cells, thereby inducing ferroptosis in cancer cells (43). This reveals a new mechanism of anti-tumor immunity. However, cancer cells undergoing ferroptosis also release immunosuppressive signals, promoting the infiltration of immunosuppressive cells, and leading to feedback protection (44). In summary, there is a complex crosstalk between ferroptotic cancer cells and immune cells during anti-tumor immune processes (45). Some studies have reported that ferroptosis inducers can significantly enhance the efficacy of ICI therapy (46, 47). However, due to the diversity of ferroptosis pathways, the application of a single inducer in multiple cancers may not be practical. Therefore, selecting specific inducers for combination with ICI is worth considering (48).

In this study, utilizing NMF dimensionality reduction clustering, we identified a novel subpopulation of ACSL4+CAFs in the BCa TME. Subsequent survival analysis and immune infiltration assessment revealed that high infiltration of this subpopulation indicates poor prognosis and lack of response to ICI. Enrichment analysis of this subpopulation revealed its

10.3389/fimmu.2024.1427124

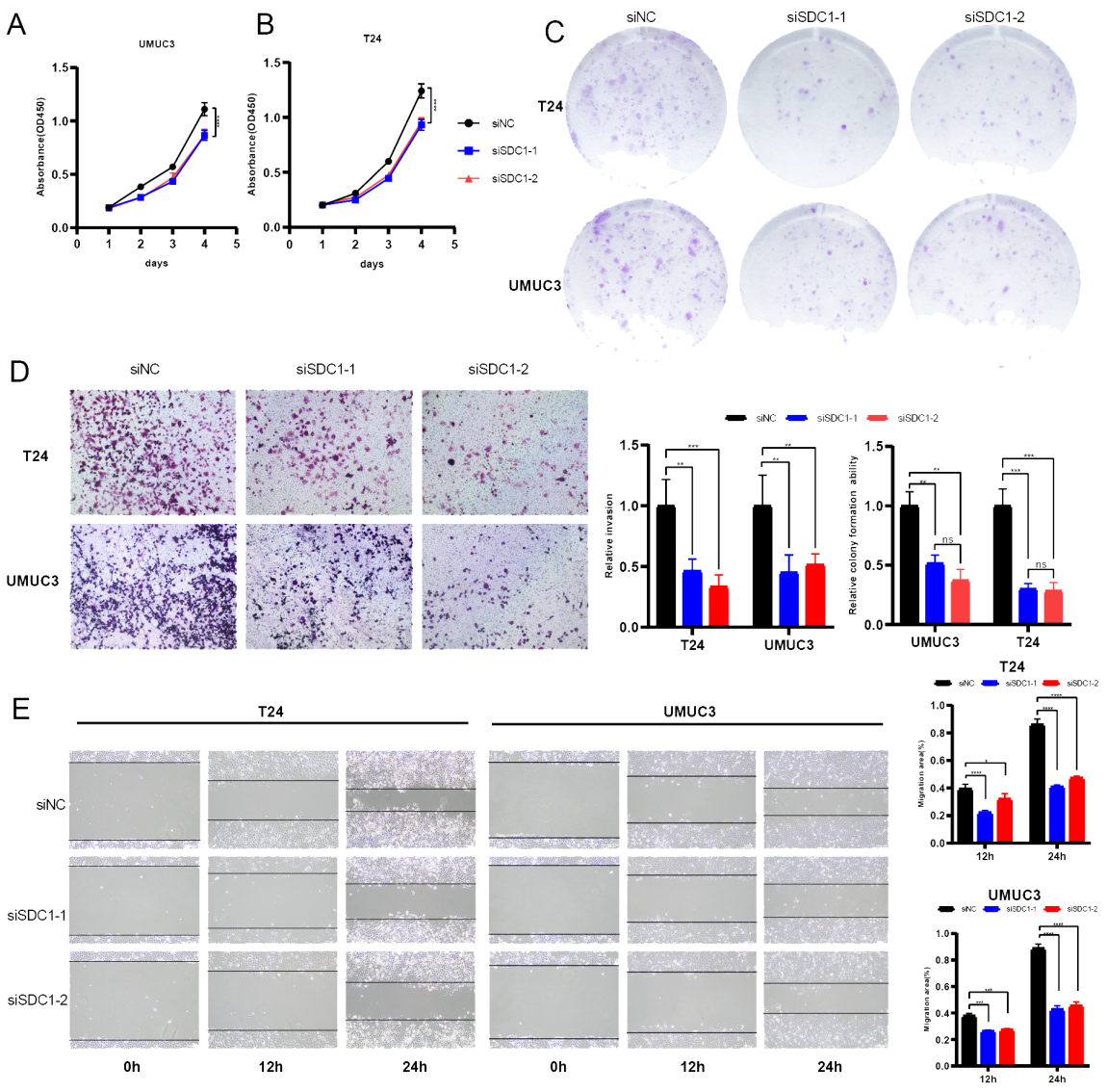


FIGURE 8
In vitro experiments. **(A, B)** Line plots showed that SDC1 enhanced the viability of BCa cells. **(C)** Colony formation assays displayed that SDC1 increased the colony numbers. **(D)** Transwell assays showed knocking down SDC1 inhibited the invasion ability of BCa cells. **(E)** Wound-healing assays showed that SDC1 promoted the direct migration of BCa cells. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. ns: no significance.

association with estrogen response, which was noteworthy given the clinical characteristic of poorer prognosis in female BCa patients compared to males (49). Furthermore, we observed frequent crosstalk between ACSL4+CAFs and cancer cells, consistent with previous research indicating that CAFs promote cancer cell proliferation and invasion (50). According to previous literature, we also found that the biological phenotype of ACSL4+CAFs is closer to iCAFs (51), which were associated with intra-tumoral inflammation. Inflammatory reactions within tumors can have a dual effect, recruiting more immune cell infiltration to enhance anti-tumor immunity while also leading to immune suppression due to chronic inflammation (52). It has been reported that iCAFs promote cancer cell proliferation, epithelial-mesenchymal transition (EMT), and the establishment of an immune-suppressive microenvironment (53). Research has shown that iCAFs promote immune suppression by releasing cytokines such as IL-6 and IL-10 to induce M2 polarization in macrophages (54). In our study, we found high expression of the ferroptosis marker gene ACSL4 on this subtype of CAFs, suggesting a role for ferroptosis in the growth and development of ACSL4+CAFs, which warrants further investigation. In the immune infiltration landscape, we also found a possible association between Tregs and ACSL4+CAFs, with increased infiltration of Tregs in the high ACSL4+CAFs group, providing further support for the notion of ACSL4+CAFs promoting immune suppression. Interestingly, we also discovered a noteworthy subpopulation of macrophages: FTH1

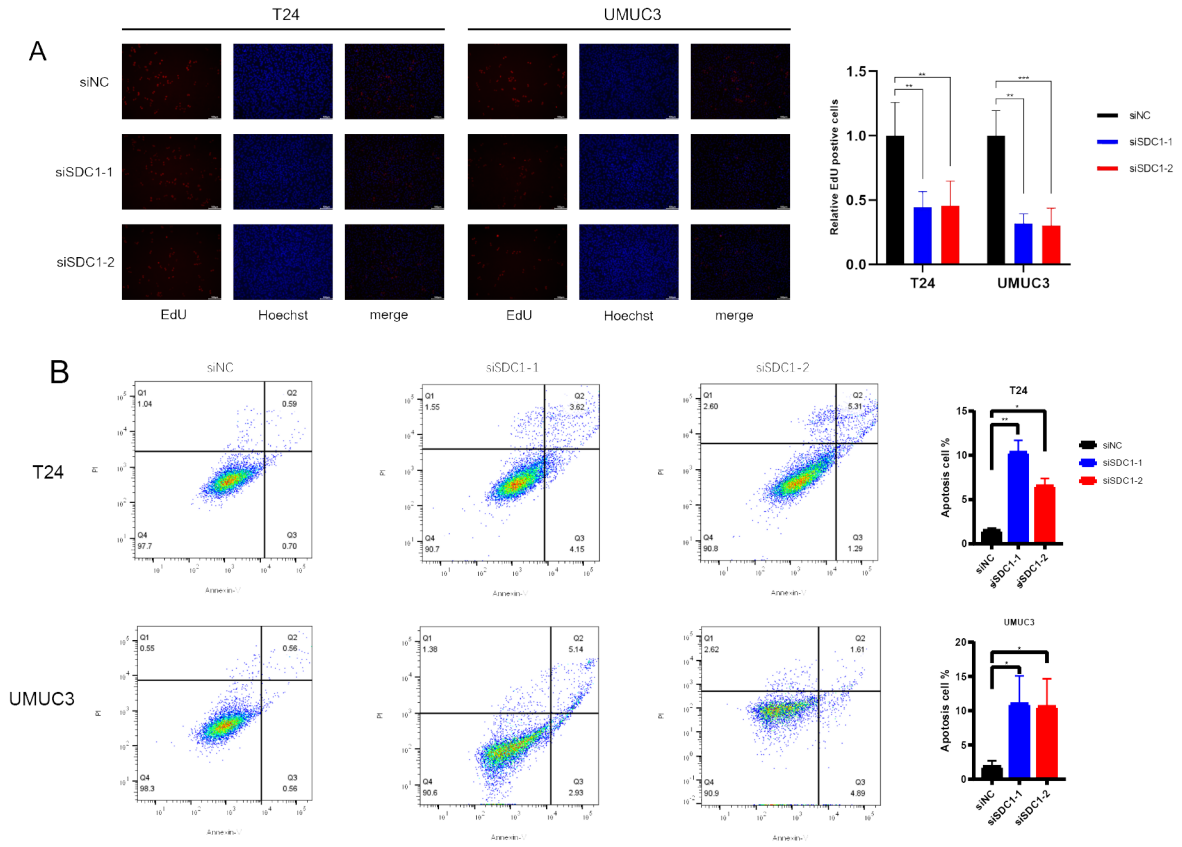


FIGURE 9
In vitro experiments. **(A)** EdU assays showed that SDC1 enhanced cell proliferation. **(B)** Flow cytometric analysis yielded that SDC1 suppressed apoptosis in BCa cells. *P < 0.05, **P < 0.01, ***P < 0.001.

+macrophages, closely associated with SPP1+macrophages. Studies have shown that SPP1+macrophages promote colorectal cancer cell proliferation and limit T cell infiltration, and their increased proportion in the TME is associated with worse patient prognosis (55, 56). FTH1+ macrophages also exhibited strong glycolysis metabolism. Some studies suggested that the metabolic reprogramming of glycolysis was crucial for macrophage polarization. We hypothesized that glycolytic activity was linked to the immunosuppressive environment in BCa and promoted the progression of BCa. In our study, we found that FTH1 +macrophages were associated with chemokine-regulated signaling pathways, which might be an important mechanism through which FTH1+macrophages regulated the immune microenvironment.

Through cell-cell communication analysis, we identified SDC1 as a target mediating crosstalk between BCa cells and ACSL4+CAFs. SDC1, also known as CD138, belongs to the syndecan proteoglycan family (57). It served as an important surface adhesion molecule involved in maintaining cell morphology and interacting with the surrounding microenvironment (58). Previous literature has reported varied expression of SDC1 in different cancers (59, 60), with decreased expression in gastric and colorectal cancers but increased expression in plasmacytoid urothelial carcinoma and pancreatic cancer (61). Particularly in pancreatic cancer, its silencing can inhibit cancer progression (62). al (63). Additionally, pancreatic cancer cells expressing SDC1 can interact with T cells expressing CCL5 in the TME, promoting tumor migration, and thereby providing a potential target for immunotherapy in pancreatic cancer (64). Overall, high expression of SDC1 presented in tumors generally predicted poor prognosis due to its association with cellular component or collagen matrix. This study revealed that SDC1 could be used as a potential marker and therapeutic target for bladder urothelial carcinoma.

Conclusion

We utilized NMF dimensionality reduction clustering to identify a novel ferroptosis-related TME cell subpopulation, ACSL4+CAFs, in BCa single-cell transcriptome data, uncovering its involvement in various phenotypes of bladder cancer. Subsequently, through integration with bulk RNA-seq data, we

validated its prognostic value. Finally, cell-cell communication analysis revealed a potential target, SDC1, providing new strategies for managing BCa patients.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Materials, further inquiries can be directed to the corresponding authors.

Ethics statement

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

Author contributions

ZC: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Writing – original draft. JH: Conceptualization, Data curation, Formal analysis, Writing – original draft. YO: Project administration, Validation, Writing – review & editing. FY: Data curation, Methodology, Project administration, Writing – review & editing. WL: Data curation, Methodology, Project administration, Writing – review & editing. SL: Funding acquisition, Supervision, Writing – review & editing. HJ: Funding acquisition, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (No.81802569, 82373222, 82202255), Leading Talent Program by Shanghai Municipal Health Commission (2022LJ008), Medical Innovation Research Special Project by Science and Technology Commission of Shanghai Municipality (22Y21900200), Clinical Scientific and Technological Innovation Project by Shanghai Hospital Development Center (SHDC12021104).

Acknowledgments

We thank the public database for providing data for our 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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1427124/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Trajectory analysis for ferroptosis-related CAF subpopulations. **(A)** The dot plot showed the marker genes of stromal cells. **(B)** Heatmap displaying the signaling pathways of CAF subpopulations. **(C)** Ligand-receptor analysis for CAF subpopulations. **(D)** Heatmap showed the role of ferroptosis-related marker genes in CAFs. **(E, F)** The pseudotime trajectory of CAF subpopulations.

SUPPLEMENTARY FIGURE 2
Trajectory analysis for ferroptosis-related macrophage subpopulations. **(A)** The t-SNE visualization for myeloid cells in BCa. **(B)** The dot plot showed the marker genes of myeloid cells. **(C, D)** The pseudotime trajectory analysis for macrophage subpopulations. **(E)** The cell-cell communication strength between the ferroptosis-related macrophage subpopulations and cancer cells. **(F)** Heatmap displaying the signaling pathways of macrophage subpopulations. **(G)** Ligand-receptor analysis for macrophage subpopulations. **(H)** The heatmap displayed the differential genes and enriched signaling pathways of each subpopulation.

References

1. Li F, Zheng Z, Chen W, Li D, Zhang H, Zhu Y, et al. Regulation of cisplatin resistance in bladder cancer by epigenetic mechanisms. *Drug resistance updates: Rev commentaries antimicrobial Anticancer chemotherapy*. (2023) 68:100938. doi: 10.1016/j.drup.2023.100938

2. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA: Cancer J Clin*. (2024) 74:12–49. doi: 10.3322/caac.21820

3. Antoni S, Ferlay J, Soerjomataram I, Znaor A, Jemal A, Bray F. Bladder cancer incidence and mortality: A global overview and recent trends. *Eur Urol*. (2017) 71:96–108. doi: 10.1016/j.eururo.2016.06.010

4. Lenis AT, Lec PM, Chamie K, Mshs MD. Bladder cancer: A review. *JAMA*. (2020) 324:1980–91. doi: 10.1001/jama.2020.17598

5. Patel VG, Oh WK, Galsky MD. Treatment of muscle-invasive and advanced bladder cancer in 2020. *CA: Cancer J Clin*. (2020) 70:404–23. doi: 10.3322/caac.21631

6. Li W, Shen Y, Yang C, Ye F, Liang Y, Cheng Z, et al. Identification of a novel ferroptosis-inducing micropeptide in bladder cancer. *Cancer Lett*. (2024) 582:216515. doi: 10.1016/j.canlet.2023.216515

7. Zhao L, Zhou X, Xie F, Zhang L, Yan H, Huang J, et al. Ferroptosis in cancer and cancer immunotherapy. *Cancer Commun (London England)*. (2022) 42:88–116. doi: 10.1002/cac2.12250

8. Battaglia AM, Chirillo R, Aversa I, Sacco A, Costanzo F, Biamonte F. Ferroptosis and cancer: mitochondria meet the "Iron maiden" Cell death. *Cells*. (2020) 9(6):1505. doi: 10.3390/cells9061505

9. Zhang H, Deng T, Liu R, Ning T, Yang H, Liu D, et al. Caf secreted Mir-522 suppresses ferroptosis and promotes acquired chemo-resistance in gastric cancer. *Mol Cancer*. (2020) 19:43. doi: 10.1186/s12943-020-01168-8

10. Chen Z, Zhou L, Liu L, Hou Y, Xiong M, Yang Y, et al. Single-cell rna sequencing highlights the role of inflammatory cancer-associated fibroblasts in bladder urothelial carcinoma. *Nat Commun*. (2020) 11:5077. doi: 10.1038/s41467-020-18916-5

11. Choi W, Porten S, Kim S, Willis D, Plimack ER, Hoffman-Censits J, et al. Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. *Cancer Cell*. (2014) 25:152–65. doi: 10.1016/j.ccr.2014.01.009

12. Sjödahl G, Lauss M, Lövgren K, Chebil G, Gudjonsson S, Veerla S, et al. A molecular taxonomy for urothelial carcinoma. *Clin Cancer Res*. (2012) 18:3377–86. doi: 10.1158/1078-0432.CCR-12-0077-T

13. Riester M, Taylor JM, Feifer A, Koppie T, Rosenberg JE, Downey RJ, et al. Combination of a novel gene expression signature with a clinical nomogram improves the prediction of survival in high-risk bladder cancer. *Clin Cancer Res*. (2012) 18:1323–33. doi: 10.1158/1078-0432.CCR-11-2271

14. Kardos J, Rose TL, Manocha U, Wobker SE, Damrauer JS, Bivalaqua TJ, et al. Development and validation of a nanostring Base47 bladder cancer gene classifier. *PLoS One*. (2020) 15:e0243935. doi: 10.1371/journal.pone.0243935

15. Kim WJ, Kim EJ, Kim SK, Kim YJ, Ha YS, Jeong P, et al. Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. *Mol Cancer*. (2010) 9:3. doi: 10.1186/1476-4598-9-3

16. Yu G, Wang LG, Han Y, He QY. Clusterprofiler: an R package for comparing biological themes among gene clusters. *Omics: J Integr Biol*. (2012) 16:284–7. doi: 10.1089/omi.2011.0118

17. Zhang C, Liu X, Jin S, Chen Y, Guo R. Ferroptosis in cancer therapy: A novel approach to reversing drug resistance. *Mol Cancer*. (2022) 21:47. doi: 10.1186/s12943-022-01530-y

18. Li J, Cao F, Yin HL, Huang ZJ, Lin ZT, Mao N, et al. Ferroptosis: past, present and future. *Cell Death Dis*. (2020) 11:88. doi: 10.1038/s41419-020-2298-2

19. Liang D, Minikes AM, Jiang X. Ferroptosis at the intersection of lipid metabolism and cellular signaling. *Mol Cell*. (2022) 82:2215–27. doi: 10.1016/j.molcel.2022.03.022

20. Zhao L, Peng Y, He S, Li R, Wang Z, Huang J, et al. Apatinib induced ferroptosis by lipid peroxidation in gastric cancer. *Gastric Cancer*. (2021) 24:642–54. doi: 10.1007/s10120-021-01159-8

21. Chu B, Kon N, Chen D, Li T, Liu T, Jiang L, et al. Alox12 is required for P53-mediated tumour suppression through a distinct ferroptosis pathway. *Nat Cell Biol*. (2019) 21:579–91. doi: 10.1038/s41556-019-0305-6

22. Koppula P, Zhuang L, Gan B. Cystine transporter Slc7a11/Xct in cancer: ferroptosis, nutrient dependency, and cancer therapy. *Protein Cell*. (2021) 12:599–620. doi: 10.1007/s13238-020-00789-5

23. Yamaguchi H, Hsu JL, Chen CT, Wang YN, Hsu MC, Chang SS, et al. Caspase-independent cell death is involved in the negative effect of Egf receptor inhibitors on cisplatin in non-small cell lung cancer cells. *Clin Cancer Res*. (2013) 19:845–54. doi: 10.1158/1078-0432.CCR-12-2621

24. Mou Y, Wang J, Wu J, He D, Zhang C, Duan C, et al. Ferroptosis, a new form of cell death: opportunities and challenges in cancer. *J Hematol Oncol*. (2019) 12:34. doi: 10.1186/s13045-019-0720-y

25. Shen L, Zhang J, Zheng Z, Yang F, Liu S, Wu Y, et al. Phgdh inhibits ferroptosis and promotes Malignant progression by upregulating Slc7a11 in bladder cancer. *Int J Biol Sci*. (2022) 18:5459–74. doi: 10.7150/ijbs.74546

26. Chang Y, Meng X, Zhao Y, Li K, Zhao B, Zhu M, et al. Novel water-soluble and ph-responsive anticancer drug nanocarriers: doxorubicin-pamam dendrimer conjugates attached to superparamagnetic iron oxide nanoparticles (Ionps). *J colloid Interface Sci*. (2011) 363:403–9. doi: 10.1016/j.jcis.2011.06.086

27. de Visser KE, Joyce JA. The evolving tumor microenvironment: from cancer initiation to metastatic outgrowth. *Cancer Cell*. (2023) 41:374–403. doi: 10.1016/j.ccell.2023.02.016

28. Liu J, Hong M, Li Y, Chen D, Wu Y, Hu Y. Programmed cell death tunes tumor immunity. *Front Immunol*. (2022) 13:847345. doi: 10.3389/fimmu.2022.847345

29. Ma X, Xiao L, Liu L, Ye L, Su P, Bi E, et al. Cd36-mediated ferroptosis dampens intratumoral Cd8(+) T cell effector function and impairs their antitumor ability. *Cell Metab*. (2021) 33:1001–12.e5. doi: 10.1016/j.cmet.2021.02.015

30. Wang W, Green M, Choi JE, Gijón M, Kennedy PD, Johnson JK, et al. Cd8(+) T cells regulate tumour ferroptosis during cancer immunotherapy. *Nature*. (2019) 569:270–4. doi: 10.1038/s41586-019-1170-y

31. Wang H, Jia Y, Gu J, Chen O, Yue S. Ferroptosis-related genes are involved in asthma and regulate the immune microenvironment. *Front Pharmacol*. (2023) 14:1087557. doi: 10.3389/fphar.2023.1087557

32. Lv M, Zheng Y, Wu J, Shen Z, Guo B, Hu G, et al. Evoking ferroptosis by synergistic enhancement of a cyclopentadienyl iridium-betulin immune agonist. *Angewandte Chemie (International ed English)*. (2023) 62:e202312897. doi: 10.1002/anie.202312897

33. Xu Y, Zeng H, Jin K, Liu Z, Zhu Y, Xu L, et al. Immunosuppressive tumor-associated macrophages expressing interlukin-10 conferred poor prognosis and therapeutic vulnerability in patients with muscle-invasive bladder cancer. *J immunotherapy Cancer*. (2022) 10(3):e003416. doi: 10.1136/jitc-2021-003416

34. Chen W, Schilperoort M, Cao Y, Shi J, Tabas I, Tao W. Macrophage-targeted nanomedicine for the diagnosis and treatment of atherosclerosis. *Nat Rev Cardiol*. (2022) 19:228–49. doi: 10.1038/s41569-021-00629-x

35. Yao L, Hou J, Wu X, Lu Y, Jin Z, Yu Z, et al. Cancer-associated fibroblasts impair the cytotoxic function of nk cells in gastric cancer by inducing ferroptosis via iron regulation. *Redox Biol*. (2023) 67:102923. doi: 10.1016/j.redox.2023.102923

36. Cui JX, Xu XH, He T, Liu JJ, Xie TY, Tian W, et al. L-kynurenine induces nk cell loss in gastric cancer microenvironment via promoting ferroptosis. *J Exp Clin Cancer research: CR*. (2023) 42:52. doi: 10.1186/s13046-023-02629-w

37. Gao W, Wang X, Zhou Y, Wang X, Yu Y. Autophagy, ferroptosis, pyroptosis, and necroptosis in tumor immunotherapy. *Signal transduction targeted Ther*. (2022) 7:196. doi: 10.1038/s41392-022-01046-3

38. Yin J, Meng X, Peng L, Xie W, Liu X, He W, et al. Ferroptosis and cancer immunotherapy. *Curr Mol Med*. (2023) 23:401–9. doi: 10.2174/1566524022666220509124608

39. Gong D, Chen M, Wang Y, Shi J, Hou Y. Role of ferroptosis on tumor progression and immunotherapy. *Cell Death Discovery*. (2022) 8:427. doi: 10.1038/s41420-022-01218-8

40. Tong X, Tang R, Xiao M, Xu J, Wang W, Zhang B, et al. Targeting cell death pathways for cancer therapy: recent developments in necroptosis, pyroptosis, ferroptosis, and cuproptosis research. *J Hematol Oncol*. (2022) 15:174. doi: 10.1186/s13045-022-01392-3

41. Hattori Y, Nishigori C, Tanaka T, Uchida K, Nikaido O, Osawa T, et al. 8-hydroxy-2'-deoxyguanosine is increased in epidermal cells of hairless mice after chronic ultraviolet B exposure. *J Invest Dermatol*. (1996) 107:733–7. doi: 10.1111/1523-1747.ep12365625

42. Weng SW, Wu JC, Shen FC, Chang YH, Su YJ, Lian WS, et al. Chaperonin counteracts diet-induced non-alcoholic fatty liver disease by aiding sirtuin 3 in the control of fatty acid oxidation. *Diabetologia*. (2023) 66:913–30. doi: 10.1007/s00125-023-05869-9

43. Ou Y, Dai X, Chen X, Chen Y, Wu S, Zhou Q, et al. Circ-afap1 promote clear cell renal cell carcinoma growth and angiogenesis by the Circ-Afap1/Mir-374b-3p/Vegfa signaling axis. *Cell Death Discovery*. (2022) 8:68. doi: 10.1038/s41420-022-00865-1

44. Ouyang S, Li H, Lou L, Huang Q, Zhang Z, Mo J, et al. Inhibition of stat3-ferroptosis negative regulatory axis suppresses tumor growth and alleviates chemoresistance in gastric cancer. *Redox Biol*. (2022) 52:102317. doi: 10.1016/j.redox.2022.102317

45. Liu T, Zhu C, Chen X, Guan G, Zou C, Shen S, et al. Ferroptosis, as the most enriched programmed cell death process in glioma, induces immunosuppression and immunotherapy resistance. *Neuro-oncology*. (2022) 24:1113–25. doi: 10.1093/neuonc/noac033

46. Ramadori P, Gallage S, Heikenwälder MF. Unique tumour microenvironment: when ferroptosis activation boosts ici of liver cancer. *Gut*. (2023) 72:1639–41. doi: 10.1136/gutjnl-2023-329472

47. Chen Z, Ou Y, Ye F, Li W, Jiang H, Liu S. Machine learning identifies the role of smad6 in the prognosis and drug susceptibility in bladder cancer. *J Cancer Res Clin Oncol*. (2024) 150:264. doi: 10.1007/s00432-024-05798-z

48. Lei H, Li Q, Pei Z, Liu L, Yang N, Cheng L. Nonferrous ferroptosis inducer manganese molybdate nanoparticles to enhance tumor immunotherapy. *Small (Weinheim an der Bergstrasse Germany)*. (2023) 19:e2303438. doi: 10.1002/smll.202303438

49. Dobruch J, Daneshmand S, Fisch M, Lotan Y, Noon AP, Resnick MJ, et al. Gender and bladder cancer: A collaborative review of etiology, biology, and outcomes. *Eur Urol*. (2016) 69:300–10. doi: 10.1016/j.eururo.2015.08.037

50. Zheng H, An M, Luo Y, Diao X, Zhong W, Pang M, et al. Pdgfrα(+)Itga11(+) fibroblasts foster early-stage cancer lymphovascular invasion and lymphatic metastasis via Itga11-sele interplay. *Cancer Cell*. (2024) 42:682–700.e12. doi: 10.1016/j.ccell.2024.02.002

51. Yoshida GJ. Regulation of heterogeneous cancer-associated fibroblasts: the molecular pathology of activated signaling pathways. *J Exp Clin Cancer research: CR*. (2020) 39:112. doi: 10.1186/s13046-020-01611-0

52. Michaud DS. Chronic inflammation and bladder cancer. *Urologic Oncol*. (2007) 25:260–8. doi: 10.1016/j.urolonc.2006.10.002

53. Goulet CR, Champagne A, Bernard G, Vandal D, Chabaud S, Pouliot F, et al. Cancer-associated fibroblasts induce epithelial-mesenchymal transition of bladder cancer cells through paracrine Il-6 signalling. *BMC Cancer*. (2019) 19:137. doi: 10.1186/s12885-019-5353-6

54. Zheng Y, Wang X, Yang X, Xing N. Single-cell rna sequencing reveals the cellular and molecular characteristics of high-grade and metastatic bladder cancer. *Cell Oncol (Dordrecht)*. (2023) 46:1415–27. doi: 10.1007/s13402-023-00820-x

55. Qi J, Sun H, Zhang Y, Wang Z, Xun Z, Li Z, et al. Single-cell and spatial analysis reveal interaction of Fap(+) fibroblasts and Spp1(+) macrophages in colorectal cancer. *Nat Commun*. (2022) 13:1742. doi: 10.1038/s41467-022-29366-6

56. Sathe A, Mason K, Grimes SM, Zhou Z, Lau BT, Bai X, et al. Colorectal cancer metastases in the liver establish immunosuppressive spatial networking between

tumor-associated Spp1+ Macrophages and fibroblasts. *Clin Cancer Res*. (2023) 29:244–60. doi: 10.1158/1078-0432.CCR-22-2041

57. Sheta M, Götte M. Syndecan-1 (Cd138) as a pathogenesis factor and therapeutic target in breast cancer. *Curr medicinal Chem*. (2021) 28:5066–83. doi: 10.2174/0929867328666210629122238

58. Ibrahim SA, Gadalla R, El-Ghonaimy EA, Samir O, Mohamed HT, Hassan H, et al. Syndecan-1 is a novel molecular marker for triple negative inflammatory breast cancer and modulates the cancer stem cell phenotype via the Il-6/Stat3, Notch and Egfr signaling pathways. *Mol Cancer*. (2017) 16:57. doi: 10.1186/s12943-017-0621-z

59. Palaiologou M, Delladetsima I, Tiniakos D. Cd138 (Syndecan-1) expression in health and disease. *Histol histopathology*. (2014) 29:177–89. doi: 10.14670/hh-29.177

60. Kind S, Kluth M, Hube-Magg C, Möller K, Makrypidi-Fraune G, Lutz F, et al. Increased cytoplasmic Cd138 expression is associated with aggressive characteristics in prostate cancer and is an independent predictor for biochemical recurrence. *BioMed Res Int*. (2020) 2020:5845374. doi: 10.1155/2020/5845374

61. Couchman JR. Syndecan-1 (Cd138), carcinomas and Emt. *Int J Mol Sci*. (2021) 22(8):4227. doi: 10.3390/ijms22084227

62. Chen K, Wang Y, Hou Y, Wang Q, Long D, Liu X, et al. Single cell rna-seq reveals the Ccl5/Sdc1 receptor-ligand interaction between T cells and tumor cells in pancreatic cancer. *Cancer Lett*. (2022) 545:215834. doi: 10.1016/j.canlet.2022.215834

63. Zhang CL, Shen Q, Liu FD, Yang F, Gao MQ, Jiang XC, et al. Sdc1 and Itga2 as novel prognostic biomarkers for pdac related to Ipmn. *Sci Rep*. (2023) 13:18727. doi: 10.1038/s41598-023-44646-x

64. Liao S, Liu C, Zhu G, Wang K, Yang Y, Wang C. Relationship between Sdc1 and cadherin signalling activation in cancer. *Pathology Res Pract*. (2020) 216:152756. doi: 10.1016/j.prp.2019.152756

frontiers | Frontiers in Immunology

TYPE Mini Review
PUBLISHED 13 September 2024
DOI 10.3389/fimmu.2024.1474770

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Yujie Ma,
The University of Manchester,
United Kingdom
Yongzhi Chen,
University of Massachusetts Medical School,
United States
Zhimin Hu,
University of California, San Diego,
United States
Haineng Xu,
University of Pennsylvania, United States

*CORRESPONDENCE
Qing Wang
Doctorwg1983@163.com;
doctorwq1983@njmu.edu.cn
Zhiwen Luo
zhiwen.luo_fudan@hotmail.com
Lijun Qu
lijunqu95276@163.com

RECEIVED 02 August 2024
ACCEPTED 28 August 2024
PUBLISHED 13 September 2024

CITATION
Feng Y, Feng X, Wan R, Luo Z, Qu L and Wang Q (2024) Impact of exercise on cancer: mechanistic perspectives and new insights. *Front. Immunol.* 15:1474770. doi: 10.3389/fimmu.2024.1474770

Impact of exercise on cancer: mechanistic perspectives and new insights

Ye Feng[1], Xingting Feng[2], Renwen Wan[2], Zhiwen Luo[2*], Lijun Qu[3*] and Qing Wang[3*]

[1]School of Stomatology, Xuzhou Medical University, Xuzhou, Jiangsu, China, [2]Department of Sports Medicine, Huashan Hospital, Fudan University, Shanghai, China, [3]Department of Orthopaedics, Kunshan Hospital of Chinese Medicine, Kunshan, Jiangsu, China

This review critically evaluates the substantial role of exercise in enhancing cancer prevention, treatment, and patient quality of life. It conclusively demonstrates that regular physical activity not only reduces cancer risk but also significantly mitigates side effects of cancer therapies. The key findings include notable improvements in fatigue management, reduction of cachexia symptoms, and enhancement of cognitive functions. Importantly, the review elucidates the profound impact of exercise on tumor behavior, modulation of immune responses, and optimization of metabolic pathways, advocating for the integration of exercise into standard oncological care protocols. This refined abstract encourages further exploration and application of exercise as a pivotal element of cancer management.

KEYWORDS

exercise, tumor microenvironment, cytokines, anti-cancer immunity, prevention, treatment

1 Introduction

The global burden of cancer continues to escalate, with millions of new cases diagnosed annually, which highlights the urgent need for effective prevention and treatment strategies. Recent statistics from major health organizations underscore a concerning rise in cancer incidence and mortality rates worldwide, compelling the medical community to explore innovative therapeutic modalities beyond traditional medical interventions (1, 2). Exercise oncology has emerged as a pivotal field of research, offering promising avenues for enhancing cancer prevention, treatment efficacy, and patient quality of life. The integration of exercise into oncological care is driven by a growing body of evidence that demonstrates the multiple benefits of physical activity for cancer patients. These benefits range from reducing the risk of cancer development and recurrence to alleviating the side effects of conventional cancer treatments such as chemotherapy and radiotherapy (3, 4). This review delves into the multifaceted advantages of exercise in the realm of cancer prevention and treatment. Consistent physical activity is demonstrated to not only mitigate

the incidence and recurrence of cancer but also augment the efficacy of various cancer therapies, including surgery, radiotherapy, chemotherapy, and immunotherapy. Moreover, exercise significantly alleviates the adverse effects associated with cancer treatments such as fatigue, cancer cachexia, and cognitive impairments. Building upon these therapeutic supports, the subsequent sections delve into the broader implications of exercise on the tumor microenvironment, showcasing its profound impact on tumor angiogenesis, cytokine modulation, and overall tumor behavior. Here, 'cancer cachexia' refers to a complex syndrome involving muscle and weight loss, while 'cognitive impairments' relate to difficulties with memory and concentration that some patients experience (5, 6). In the following sections, we explore the current state of exercise oncology, emphasizing how exercise is being integrated into cancer care protocols and highlighting the potential mechanisms through which physical activity exerts its beneficial effects. By providing healthcare professionals and researchers with a comprehensive overview of the latest insights and developments in this field, this review aims to foster a better understanding of the role of exercise in cancer care and encourage further research and clinical application of exercise as a standard component of oncological treatment strategies.

2 The positive impact of exercise on cancer prevention and treatment

The prevailing view was once that cancer survivors should refrain from exercise, but contemporary research underscores that with meticulous supervision, they can engage in exercise regimens safely. Physical activity proves beneficial throughout the phases of cancer prevention, treatment, and survivorship (Figure 1).

2.1 Exercise reduces cancer incidence and prevents recurrence

To enhance the manuscript's flow, discussions on exercise's role in reducing the incidence and recurrence of cancer are consolidated. Exercise through its multiple forms such as aerobic and resistance activities significantly lowers the risk of developing cancer. This effect is supported by numerous studies including systematic reviews and meta-analyses demonstrating reduced risks for cancers such as breast, colon, and prostate, influencing body weight, inflammation levels, and hormonal balances. Transitioning from prevention, the subsequent sections will explore how exercise also amplifies the therapeutic efficacy of cancer treatments, thus providing a dual benefit in oncology care. These exercises also modulate various metabolic pathways that are often implicated in cancer progression (7, 8). For example, a landmark study demonstrated that regular exercise reduces the risk of colon cancer by up to 24% compared to individuals who are inactive (9). These findings have helped shape current exercise guidelines for cancer prevention, which recommend at least 150 minutes of moderate-intensity or 75 minutes of vigorous-intensity aerobic physical activity per week (10).

2.2 Exercise enhances the efficacy of cancer treatments

Cancer patients typically undergo various treatments, including surgery, chemotherapy, radiotherapy, targeted therapy, and immunotherapy. There is ample evidence that exercise enhances the effectiveness of these treatments. It also improves patient outcomes.

2.2.1 Surgery

Surgery is a principal method for treating cancer, often utilized in clinical settings. However, surgical stress can induce significant acute systemic disturbances and local damage, which may lead to complications and promote cancer recurrence and metastasis through neuroendocrine, immune, and metabolic imbalances (10, 11). Clinical practice recognizes high cardiovascular reserve capacity and robust diastolic function as positive predictors for surgical outcomes (12). Preoperative exercise benefits include enhanced physical fitness, improved myocardial diastolic capacity, augmented contractile reserve, increased muscle mass, and reduced pulmonary congestion (13, 14). These benefits contribute to improved preoperative health, accelerated postoperative recovery, and reduced hospitalization duration (11). Research shows that postoperative rehabilitation training decreases complication rates post-gastric cancer surgery (15, 16), and home-based exercise programs are feasible for elderly cancer patients. There is a pressing need for large-scale, rigorous clinical trials to assess the role of exercise in surgical cancer treatment.

2.2.2 Radiotherapy

Radiotherapy, utilized by approximately 60% of cancer patients, targets malignant tumors effectively (17). The success of radiotherapy hinges on the oxygenation of tumor tissues since oxygen is essential for generating reactive oxygen species (ROS) that damage cancer cells (18). Smaller, well-vascularized tumors with minimal hypoxic cells respond better to radiotherapy, whereas larger, poorly vascularized tumors with central necrosis are less responsive (19). By normalizing tumor vasculature and improving blood flow and oxygen delivery, exercise reduces tumor hypoxia and enhances radiotherapy's efficacy. Experimental studies, such as those using 4T1 breast cancer and MC38 colorectal cancer (CRC) models, have demonstrated that when exercise is combined with radiotherapy, there is a significant reduction in tumor growth and metastasis (20). Furthermore, exercise is thought to bolster the immunological effects of radiotherapy. Animal studies have shown that exercise increases the secretion of endothelin, boosts natural killer (NK) cell infiltration, and enhances the expression of NK cell receptors such as Klrk1 and Il2rβ, with resistance training showing particular efficacy in these enhancements (21). Moreover, a combination of resistance and aerobic exercises has been observed to synergistically amplify these anticancer effects. In clinical settings, implementing exercise routines during radiotherapy has

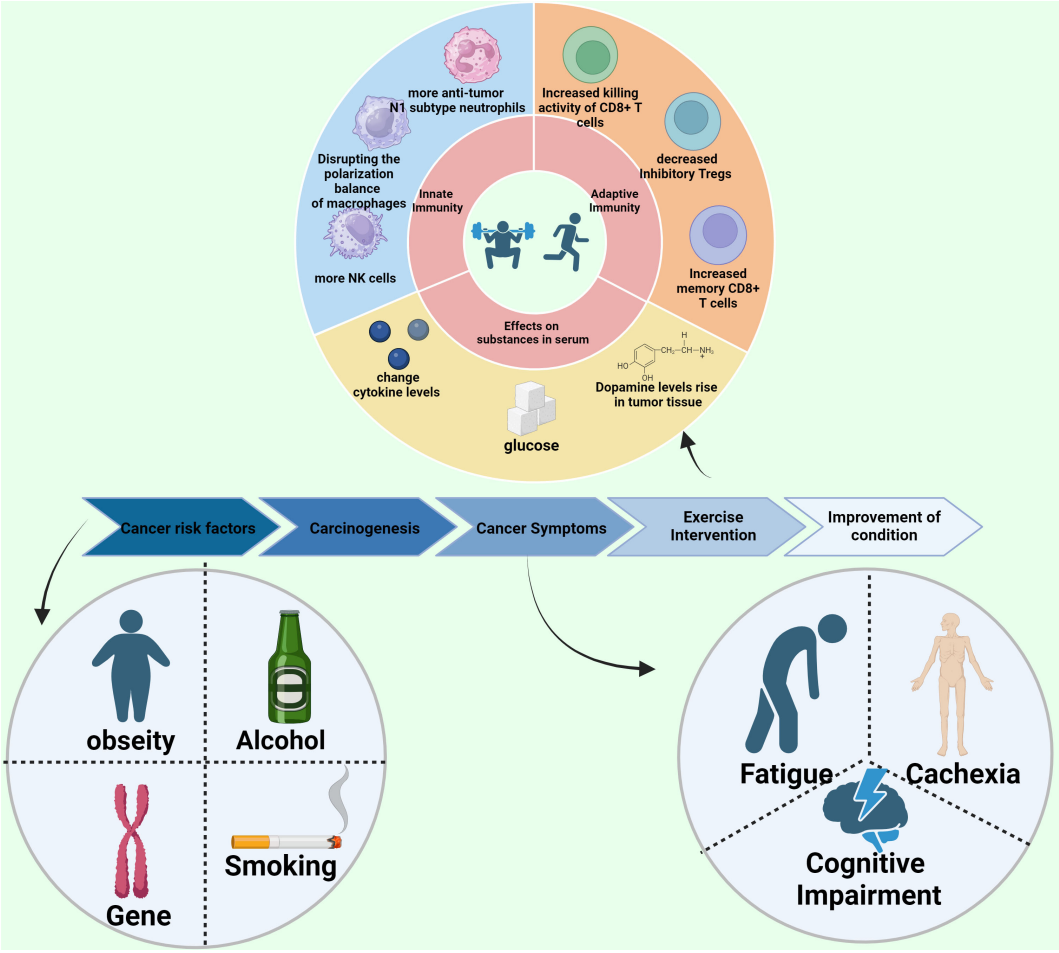


FIGURE 1
Exercise inhibits cancer. Schematic diagram of the review.

shown promising results; for example, resistance training significantly lowered serum kynurenine (KYN) levels in breast cancer patients, suggesting a non-pharmacological approach to improve radiotherapy outcomes (23).

2.2.3 Chemotherapy

Recent studies underscore the importance of incorporating exercise into the regimen of cancer patients undergoing chemotherapy. Exercise demonstrates significant benefits for these patients by countering the negative effects of cytotoxic drugs, which often increase systemic inflammation and local tissue damage. This is achieved by modulating inflammation markers such as interleukin-6 (IL-6) and tumor necrosis factor-alpha (TNF-α), thereby reducing treatment-related fatigue and enhancing overall physical well-being (24–26). Preclinical studies have shown that exercise promotes angiogenesis, normalizes tumor vasculature, and improves drug delivery efficacy, illustrating the potential for exercise to magnify the anticancer effects of chemotherapeutic agents such as gemcitabine and doxorubicin (27, 28). For example, enhanced blood perfusion facilitated by exercise in tumor areas has been shown to improve the efficacy of these drugs (29).

Clinical evidence supports these findings; for instance, studies involving pancreatic cancer patients who engaged in structured exercise programs during neoadjuvant chemotherapy observed improvements in tumor vascularization, which plays a crucial role in optimizing drug delivery and enhancing treatment efficacy (9). Additionally, exercise interventions before and after chemotherapy have been linked with significantly reduced risks of recurrence and mortality in breast cancer patients, showcasing the potential of exercise as a complementary therapy (9).

To integrate insights from animal studies with clinical trial results, we have now included comparative analyses and summary tables in our manuscript. These additions emphasize the translational potential of preclinical findings and spotlight areas where human trials could further explore the mechanistic bases of these exercise benefits. This integrated approach not only clarifies how exercise contributes to enhanced chemotherapeutic outcomes but also provides a blueprint for future research aimed at harnessing exercise as a standard adjunct in cancer treatment protocols.

2.2.4 Immunotherapy

Radiotherapy, utilized by approximately 60% of cancer patients, effectively targets malignant tumors. The success of radiotherapy

hinges crucially on the oxygenation of tumor tissues, as oxygen is essential for generating reactive oxygen species (ROS) that damage cancer cells (17, 18). Smaller, well-vascularized tumors with minimal hypoxic cells respond better to radiotherapy, whereas larger, poorly vascularized tumors with central necrosis are less responsive (19).

Exercise improves blood flow and oxygen delivery by normalizing tumor vasculature, which reduces tumor hypoxia and enhances the efficacy of radiotherapy. Experimental studies, such as those using 4T1 breast cancer and MC38 colorectal cancer (CRC) models, have demonstrated that when exercise is combined with radiotherapy, there is a significant reduction in tumor growth and metastasis (20). Furthermore, exercise is thought to bolster the immunological effects of radiotherapy. Animal studies have shown that exercise increases the secretion of endothelin, boosts natural killer (NK) cell infiltration, and enhances the expression of NK cell receptors such as Klrk1 and Il2rβ, with resistance training showing particular efficacy in these enhancements (21).

In response to the reviewer's comments, we have critically discussed the methodologies of the cited studies and expanded our examination of how variations in exercise protocols—such as intensity, duration, and type of exercise—impact the outcomes. This discussion now considers whether these differences have been adequately addressed in the literature and how they might contribute to discrepancies in study results. For example, we contrast the effects of aerobic versus resistance training in various cancer models and patient populations, and we highlight the need for standardized exercise protocols to facilitate more consistent and comparable results across studies.

This refined focus not only aligns with clinical observations but also sets a stage for future research to standardize exercise protocols as adjunct therapy in radiotherapy, ensuring more robust and reproducible benefits across diverse patient demographics.

2.3 Physical activity helps to lessen cancer symptoms and treatment-related adverse effects

Cancer and its array of treatments often result in substantial psychological and physiological distress, undermining both productivity and overall quality of life. Exercise is recognized for its role in mitigating these adverse effects, helping to sustain the physical vitality and mental resilience of patients, thereby enhancing their overall well-being. Numerous clinical guidelines advocate for physical activity in cancer care, supporting patients in maintaining a life enriched with activities beyond their identity as patients (37).

2.3.1 Fatigue

Cancer-related fatigue (CRF) is a common symptom experienced by patients following a cancer diagnosis or the start of treatment, differing from typical fatigue in that it is not alleviated simply by rest (38). Research underscores the effectiveness of exercise in reducing CRF. The American Society of Clinical Oncology (ASCO) advises cancer survivors to engage in 150 minutes of moderate aerobic exercise weekly, like brisk walking or swimming, supplemented by two to three strength-training sessions (37, 39, 40). This regimen, while seemingly modest, significantly diminishes the intensity of CRF. A comprehensive meta-analysis involving 11,525 participants across 113 trials highlighted that exercise outperforms pharmacological interventions in reducing fatigue during and post cancer therapy (38). Moreover, a clinical trial segmented pancreatic cancer patients into two cohorts—one receiving standard care and the other standard care plus a structured exercise program. Results indicated that the exercise group exhibited notable enhancements in physical function, quality of life, and clinical symptoms, thus affirming the role of physical activity in augmenting exercise capacity and overall health status (41).

2.3.2 Cancer cachexia

Cancer cachexia (CC) is a multifaceted syndrome prevalent among cancer patients, characterized by significant metabolic changes leading to progressive weight loss, primarily due to skeletal muscle atrophy, sometimes accompanied by fat loss. This syndrome is propelled by an inflammatory response that drives insulin resistance, hyperlipidemia, and mitochondrial dysfunction, thus perpetuating a deteriorating cycle (42). Although nutritional support is critical, it alone is insufficient to reverse the progression of CC (44). CC is particularly common in individuals with lung, colorectal, and gastrointestinal cancers, affecting over 85% of such patients (45, 46). While pharmaceutical solutions are being explored, standardized treatments remain limited.

Physical exercise serves as an effective non-pharmacological intervention for CC, significantly enhancing the survival rates and quality of life for affected patients. Most patients undergoing cancer treatment report a reduction in muscle mass and strength (47). This loss is often attributed to circulating tumor-derived factors that promote muscle degradation. Furthermore, the secretion of inflammatory adipokines in cancer patients may exacerbate insulin resistance, while the accumulation of intramuscular fat can impede blood flow within muscles, further aggravating metabolic imbalances and promoting tumor growth. Research has demonstrated that tumor-derived substances like parathyroid hormone-related protein and myostatin are key contributors to muscle atrophy and weight loss (47–50). Exercise has been shown to effectively reverse these effects. For example, engaging in voluntary wheel running during chemotherapy not only prevents weight loss but also maintains lean body mass and muscle strength, counteracting treatment-induced anorexia (51).

Despite promising results, the need for large-scale clinical trials to validate the effectiveness of exercise in managing CC remains. A particular study demonstrated the feasibility of exercise training among advanced cancer patients, although it was limited by its statistical power. Such multimodal approaches, which combine exercise, nutritional support, and pharmacotherapy, appear promising in addressing the multifaceted challenges of CC (52). Moving forward, research should focus on elucidating the molecular mechanisms through which exercise mitigates muscle atrophy and enhances metabolic functions in CC, potentially offering new avenues for treatment and patient care (43).

2.3.3 Cognitive impairment

The causes of cognitive impairment in cancer patients are diverse and complex. Beyond the direct cognitive decline due to brain metastases from certain cancers, a multitude of factors contribute to this condition. These include the stress and psychological impact of a cancer diagnosis, the systemic effects of the cancer itself, various treatments, and genetic predispositions. Intriguingly, some breast cancer patients demonstrate cognitive dysfunction even prior to receiving any treatment, suggesting that specific characteristics inherent to breast cancer may predispose individuals to cognitive impairments. The Apolipoprotein E4 (APOE) gene, a well-documented risk factor for Alzheimer's disease, has also been implicated in this context (53, 54). A prospective cohort study involving 943 breast cancer patients revealed that those adhering to national physical activity guidelines exhibited superior cognitive function both six months before and after undergoing chemotherapy, compared to their less active counterparts (55). Nonetheless, other studies report no significant correlation between higher self-reported physical activity levels during chemotherapy or follow-up and improved cognitive outcomes (56), highlighting the complexity of factors influencing cognitive health in cancer patients.

Larger clinical trials are underway to assess whether exercise can ameliorate cognitive impairment in cancer patients, focusing also on the underlying molecular mechanisms involved. These studies are designed to refine exercise protocols with the goal of enhancing cognitive functions, thereby improving the quality of life for those affected by cancer. By identifying the specific impacts of various forms of exercise on cognitive health, researchers hope to develop targeted strategies that effectively combat cognitive decline and boost overall mental capabilities in cancer patients. This approach could lead to more personalized exercise recommendations, tailored to the needs and health statuses of individual patients, maximizing the therapeutic benefits of physical activity in oncology settings.

3 Effects of exercise on tumor microenvironment regulation and molecular mechanisms

3.1 Effects of exercise on tumor angiogenesis

Angiogenesis is crucial for the progression, spread, and treatment of cancer. It involves the formation of new blood vessels, which is essential for supplying nutrients to tumors and facilitating the spread of cancer cells (57). Additionally, the oxygen carried through these new vessels enhances tumor perfusion, helping to reduce the effects of tumor hypoxia (58). Many cancer therapies target angiogenesis, using inhibitors to prevent the formation of these blood vessels within tumors. Improving vascular conductivity can alleviate tumor hypoxia, enhance the infiltration of immune cells into tumor tissues, inhibit tumor growth, and increase the effectiveness of cancer therapies (59).

However, the blood vessels in tumors are often abnormal and underdeveloped, leading to insufficient oxygen transport and significant tumor hypoxia (60). This hypoxic environment activates the production of hypoxia-inducible factor 1-α (HIF1α), which not only suppresses anti-cancer immune responses but also promotes cancer invasion and metastasis (57). Additionally, HIF1α drives abnormal angiogenesis, further deteriorating blood perfusion and oxygenation, thus reinforcing a cycle of hypoxia and increasingly aggressive cancer behaviors. Under these low-oxygen conditions, glycolysis is enhanced, resulting in an accumulation of lactate that negatively impacts the tumor's immune microenvironment.

Exercise influences tumor angiogenesis, or the process by which new blood vessels form to supply the tumor, through several key mechanisms. Think of it as building new roads to improve the delivery of goods to a city—except in this case, the 'goods' are nutrients that unfortunately help the tumor grow. Firstly, it enhances the density and maturity of blood vessels within tumor tissues, promoting vascular normalization. This helps improve the delivery of oxygen and nutrients, which can affect tumor growth and treatment response. Secondly, exercise increases interaction between endothelial cells and other cells, boosting pericyte coverage and enhancing the expression of angiogenic factors like VEGFA in endothelial cells (61–63). Thirdly, it enhances nitric oxide utilization, a promoter of angiogenesis, by upregulating endothelial nitric oxide synthase, improving blood flow and oxygen delivery to tumor areas (62, 64). Additionally, exercise-induced secretion of myokines from skeletal muscles and adipokines from fat tissue plays a crucial role in angiogenesis (65, 66). Within the tumor immune microenvironment, exercise can decrease the number of M2 tumor-associated macrophages (TAMs) and prevent tumor-associated neutrophils (TANs) from shifting to a pro-angiogenic phenotype, thereby reducing the support for tumor growth and spread.

Alleviating tumor hypoxia and increasing oxygen levels can independently exert anti-cancer effects. An animal study showed that inhaling high concentrations of oxygen reduced tumor metastasis and improved survival rates (67). Enhanced oxygen supply can boost the efficacy of immune cells in attacking tumor cells by elevating pro-inflammatory cytokines and reducing immunosuppressive molecules and regulatory T cells in mouse tumors (67). However, further research is needed to fully understand how exercise impacts tumor vasculature in cancer patients. A notable study involving potential cancer patients demonstrated that exercise could significantly remodel human tumor vasculature. Moderate aerobic or anaerobic training increased both the number and density of blood vessels within tumor tissues, providing new avenues for targeting anti-cancer drugs more effectively through the enhanced vasculature (28).

3.2 Effects of exercise on cytokines and growth factors

The tumor microenvironment (TME) is a critical factor in cancer progression and treatment response. Exercise exerts a multifaceted impact on the TME through various mechanisms that influence tumor growth, immune responses, and overall disease trajectory. This section explores how physical activity

compares with other non-pharmacological interventions, highlighting its unique contributions to cancer care. Exercise enhances the regulation of the TME primarily by improving systemic metabolism and reducing inflammation. Regular physical activity leads to an increased infiltration of immune cells into the tumor, enhancing anti-tumor immune responses (20). It also affects the production of cytokines and growth factors that can either inhibit or promote tumor growth, depending on their balance and the context of their release (21, 22). Comparatively, dietary modifications can also influence the TME but typically focus more on altering the nutrient supply to tumors and modifying systemic metabolic pathways that cancer cells exploit for growth and survival (23, 24). For example, ketogenic diets have been shown to reduce glucose availability to tumors, potentially slowing their growth. Psychological support, another crucial non-pharmacological intervention, primarily affects cancer outcomes by improving patients' mental health, which can indirectly influence the TME by reducing stress-induced alterations in immune function and hormone levels (25). Stress reduction has been shown to decrease the production of pro-inflammatory cytokines and stress hormones that can promote tumor growth and metastasis (26). The integration of exercise with dietary changes and psychological support can provide a comprehensive approach that maximizes the therapeutic potential of each modality. While exercise directly modifies the physical and immune landscape of the TME, dietary interventions can starve tumors of necessary nutrients, and psychological support can maintain a healthier systemic environment less conducive to cancer progression (68). Future research should focus on creating integrated treatment protocols that combine these non-pharmacological interventions to optimize cancer treatment outcomes. By doing so, it is possible to leverage the unique advantages of exercise alongside dietary and psychological interventions, creating a multi-faceted strategy that addresses the complex nature of cancer and enhances patient quality of life.

3.2.1 Myokines

Myokines (proteins released by muscle cells during exercise that have various biological effects) are a group of proteins and peptides secreted by skeletal muscles during exercise, including IL-6, irisin, decorin, IL-15, BDNF, IL-10, and IL-8. These can also be released from other organs and tissues. Myokines play dual roles in cancer biology, exhibiting either anti-tumor or pro-tumor activities depending on their nature and the surrounding environment. For instance, irisin can directly curtail tumor growth by inducing G2/M cell cycle arrest, escalating p21 levels, and simultaneously inhibiting cell proliferation and migration, while promoting apoptosis in glioblastoma cells (66, 69). Other myokines like IL-6 and IL-15 contribute to tumor suppression by hindering adipogenesis, while IL-6, IL-10, and IL-8 can bolster immune cell activity, enhancing their numbers and cytotoxic capabilities, thereby fostering a "hot" immune microenvironment conducive to fighting cancer.

IL-15, a prevalent myokine in skeletal muscle, is particularly important for its role in immunoregulation, supporting the proliferation and maturation of T cells and NK cells, crucial for the body's defense against malignancies. Exercise stimulates the release of adrenaline, which can trigger a cascade leading to an acute anti-inflammatory/immunoregulatory state, resulting in the production of myokines such as IL-6 and IL-8. These myokines are essential for modulating NK cells and CD8+ T cells to combat tumor growth (36). Research has shown that exercise-induced IL-6 possesses both direct and indirect anti-cancer properties. For example, serum collected immediately after moderate-intensity aerobic interval exercise from men with lifestyle risk factors significantly reduced the proliferation of human colorectal cancer (CRC) cells, hinting at the potent inhibitory influence of IL-6 on these cells. This suppression of CRC cell proliferation by exercise could be partially attributed to IL-6-driven DNA damage and repair dynamics. Animal studies have illustrated that adrenaline and IL-6 released during exercise facilitate NK cell mobilization, redistribution, activation, and enhanced infiltration into tumor sites. Additionally, IL-6 might also alleviate cancer-related fatigue, potentially via the actions of pro-inflammatory cytokines IL-1β and TNF-α.

3.2.2 Adipokines

Exercise influences adipogenesis and the metabolism of adipose tissue, with adipokines—proteins secreted by fat cells—having their secretion levels modified by physical activity. Dysregulated adipogenesis is a key contributor to cancer progression. Adipokines such as leptin, resistin, estrogen, macrophage migration inhibitory factor (MIF), and monocyte chemoattractant protein-1 (MCP-1) are instrumental in this context (70, 71). Leptin, for instance, promotes the growth of breast cancer cells, tumor angiogenesis, and inhibits apoptosis, whereas adiponectin exhibits opposing effects by reducing tumor cell proliferation and angiogenesis, thus restricting nutrient supply to tumors (72). Lower levels of adiponectin have been noted in patients with various cancers, including endometrial, esophageal, and liver cancers (65). Most other adipokines tend to facilitate cancer progression and metastasis by enhancing cell proliferation and migration, inhibiting apoptosis, and fostering inflammation.

A recently identified adipokine, kisspeptin, enhances the sensitivity of organs to glucose, lipids, and oxygen, thereby augmenting fat utilization during exercise and maintaining a balance between fat production and consumption (73). Exercise modulates adipose tissue by affecting adipokine levels, reducing adipogenesis, enhancing lipolysis, increasing glucose uptake and insulin sensitivity, and facilitating the conversion of white adipose tissue to brown adipose tissue (66, 74, 75). Studies, such as one involving a high-risk breast cancer population, have demonstrated that aerobic exercise training can reduce breast cancer risk by lowering body fat and modulating levels of leptin and adiponectin (76).

3.3 Effects of exercise on nutritional components and metabolism in cancer patients

3.3.1 Nutritional components and metabolic pathways in cancer

Emerging research highlights the intricate relationship between nutrition, exercise, and cancer recovery, demonstrating how these elements interact to significantly influence patient outcomes.

Studies involving post-surgical rats have underscored this interaction, revealing a notable decline in muscle strength and mass linked to the upregulation of genes associated with the ubiquitin-proteasome system, autophagy-lysosome system, and fast-twitch muscle fibers (77). Despite moderate exercise and amino acid supplementation, these rats exhibited reduced muscle strength. However, their gastrocnemius muscle mass increased, muscle atrophy was slowed, and genes related to fast-twitch fibers were downregulated, suggesting that postoperative intravenous amino acid and calcium ion supplementation combined with moderate exercise may help mitigate muscle loss.

Prolonged inactivity can disrupt the body's nutrient balance, affecting crucial elements like glucose and calcium ions, as well as growth factors. Elevated levels of certain growth factors can activate significant metabolic pathways such as the PI3K/AKT/mTOR pathway, enhancing nutrient absorption and utilization, potentially facilitating tumor growth and progression (78). Regular, long-term exercise has been shown to reduce the levels of these growth factors in the bloodstream, improve overall metabolic rates, and decrease the stimulation of cancerous tissues by these growth factors. Chronic physical activity substantially increases glucose uptake by skeletal muscles, reduces circulating glucose levels, and decreases both insulin and insulin-like growth factor (IGF) concentrations (79).

The general effects of exercise on growth factors, cytokines, nutrients, and metabolites are well-documented. However, more targeted clinical and basic research is required to verify these impacts specifically in the tumor tissues of cancer patients. For instance, a study involving prostate cancer patients who underwent a 12-week exercise program revealed increased serum levels of oncostatin M and myokines, decreased IGF levels, and a slowdown in tumor cell growth, supporting the potential tumor-inhibitory effects of exercise (80). Post-treatment exercise interventions have also shown beneficial impacts on IGF1 and inflammatory biomarkers in breast cancer patients (81). Additionally, research on breast cancer survivors demonstrated that a combined regimen of aerobic and anaerobic exercises effectively ameliorated metabolic disorders, reduced circulating biomarkers related to insulin resistance and inflammatory responses—such as insulin, IGF-1, IL-6, IL-8, and TNF-α—and significantly decreased endothelin levels, which are associated with muscle loss and degeneration. Concurrently, these exercises increased adiponectin levels, further illustrating the multifaceted benefits of physical activity in managing cancer-related metabolic disruptions (80, 82).

3.4 Effects of exercise on anti-cancer immunity

Exercise plays a pivotal role in modulating the immune landscape within cancer patients, impacting both innate and adaptive immune responses. This section delves into how exercise influences these responses and highlights the potential for personalizing exercise regimens to enhance their efficacy based on individual patient profiles.

Physical activity has been shown to significantly improve the functionality and number of various immune cells, which are crucial for combatting cancer. For example, exercise boosts the number and activity of natural killer (NK) cells, T cells, and dendritic cells, all of which play roles in recognizing and destroying cancer cells (30, 31). Regular exercise also reduces systemic inflammation, a common contributor to immunosuppression in cancer patients, thereby enhancing the overall immune surveillance and response to tumors (32).

Personalizing exercise regimens can maximize these immunological benefits by tailoring the intensity, duration, and type of exercise to individual patient needs. Factors such as the patient's cancer type, treatment stage, overall health, and genetic makeup should guide the customization of exercise programs. For instance, patients with solid tumors might benefit more from moderate-intensity aerobic exercises, which have been shown to improve blood flow and oxygenation to the tumor site, enhancing the efficacy of other treatments like chemotherapy and radiotherapy (33).

Additionally, understanding the genetic and metabolic profiles of cancer patients can further refine exercise prescriptions. Genetic markers related to inflammation and immune cell functionality, such as variations in cytokine genes, can indicate how a patient might respond to different forms of exercise (34). Similarly, metabolic profiling can reveal how exercise could influence cancer metabolism directly or support the body's natural anti-cancer mechanisms (35).

Ongoing research is increasingly supporting the idea of integrating biomarker analysis into routine clinical practice to guide exercise recommendations. By assessing markers of inflammation, immune cell activity, and metabolic function, clinicians can develop more effective, personalized exercise plans that not only support the patient's general health but also directly contribute to cancer treatment and recovery.

Future studies should focus on longitudinal analyses to better understand the long-term effects of personalized exercise on cancer prognosis. Such research will provide deeper insights into the optimal exercise modalities for different cancer types and stages, potentially leading to standardized yet customizable exercise guidelines within oncology.

3.4.1 Innate immunity

Natural killer (NK) cells are essential players in the body's innate immune response, and their activity and numbers can be significantly influenced by exercise. Research using mouse tumor models has demonstrated that interventions such as wheel-running increase NK cell infiltration into tumor tissues, which considerably slows cancer growth. This effect is primarily mediated by adrenaline and muscle-derived interleukin-6 (IL-6) (48, 83). Although exercise does not directly enhance the cytotoxicity of NK cells, it upregulates ligands for activating NK cell receptors in both mouse cancer models and human studies, thereby enhancing their cytotoxic potential (48, 84). Furthermore, combining exercise with radiation therapy has shown to increase NK cell infiltration and upregulate gene expression of NK cell receptors, boosting the effectiveness of the radiation treatment (21, 85).

Macrophages also play a crucial role in innate immunity and anti-cancer responses, with the pro-inflammatory M1 phenotype exhibiting anti-tumor effects, while the anti-inflammatory M2 phenotype supports tumor growth by releasing factors like IL-10

and TGF-β. Exercise can influence macrophage polarization towards the M1 phenotype, enhancing anti-cancer effects. Studies suggest that long-term exercise disrupts the balance of macrophage polarization, increasing differentiation towards the M1 phenotype, and thus contributing to the delay in cancer progression (86). However, the detailed mechanisms through which exercise influences macrophage activity remain largely unexplored.

Neutrophils, particularly tumor-associated neutrophils (TANs), play dual roles in cancer progression. The pro-tumor N2 subtype and the anti-tumor N1 subtype of TANs directly and indirectly regulate cancer cell survival, migration, immune function, and angiogenesis (87). Preclinical studies indicate that both swimming and running can significantly delay tumor growth, associated with a reduction in neutrophil counts (88–90). Furthermore, exercise-induced release of high-mobility group box 1 (HMGB1) has been observed to enhance citric acid metabolism in the tricarboxylic acid cycle, thereby improving immunosurveillance of senescent cells in a mechanism dependent on nuclear factor erythroid 2–related factor 2 (NRF2) (11). These findings underline the significant role of exercise in modulating the innate immune response against cancer, suggesting potential therapeutic benefits for cancer patients.

3.4.2 Adaptive immunity

Exercise has demonstrated a positive influence on adaptive anti-cancer immunity as well. In various studies, particularly with mouse models of breast cancer, physical activity has been shown to not only increase the number of CD8+ T cells infiltrating tumors but also to enhance their cytotoxic capabilities. This boost in CD8+ T cell activity due to exercise may be mediated through the CXCL9/11-CXCR3 signaling pathway, which is crucial for T cell recruitment and function (91). Another research finding suggests that exercise improves CD8+ T cell efficacy by altering central carbon metabolism, thus optimizing their energy use and functional capacity (92).

In models of pancreatic cancer, exercise has been found to facilitate the mobilization and intra-tumoral clustering of IL15Rα+ CD8+ T cells, thereby amplifying the anti-tumor immune responses (93). Importantly, the augmentation in CD8+ T cells due to regular physical activity can significantly enhance the effectiveness of standard anti-cancer treatments, such as immunotherapy and radiotherapy (85, 91, 93).

Moreover, exercise impacts adaptive immunity by regulating various factors that not only increase the infiltration of CD8+ T cells into tumors but also boost their expression of functional molecules, crucial for their anti-tumor activity. Concurrently, exercise has been observed to decrease the presence of immunosuppressive regulatory T cells (Tregs), which can otherwise hinder effective immune responses against tumors. Additionally, physical activity appears to increase the number of memory CD8+ T cells, which are important for long-term immune surveillance and cancer control.

These findings indicate that regular exercise can potentiate the anti-cancer efficacy of treatments like radiotherapy and therapies targeting PD-(L)1, by modulating the immune landscape in favor of a more robust anti-tumor response. This highlights the potential of exercise as a strategic complement in cancer treatment protocols to leverage the body's own immune system against cancer.

3.5 Effects of exercise on cancer cells

Basic research has highlighted that exercise can impact cancer cells by hindering their proliferation, promoting apoptosis, and reducing their migration capabilities. For example, in a study involving colorectal cancer (CRC) patients, it was found that serum altered by exercise significantly suppressed the proliferation of LOVO cancer cells. Acute exercise leads to a rise in serum IL-6 levels (94), which in turn stimulates the release of anti-inflammatory cytokines, mobilizes immune cells, and helps mitigate DNA damage in early malignant cells, producing a range of beneficial biological effects. *In vitro* experiments further demonstrated that recombinant IL-6 at concentrations of 10 and 100 pg/mL could inhibit human CRC cell proliferation and reduce γ-H2AX expression, reflecting the anti-cancer properties associated with exercise. Additionally, recent research has shown that serum from metastatic castration-resistant prostate cancer (mCRPC) patients, who engaged in long-term regular exercise, exhibited delayed proliferation of human prostate cancer cells (95).

The mechanisms through which exercise influences cancer cells are complex and multifaceted. Firstly, exercise reduces levels of various nutrients and growth factors, such as glucose and insulin-like growth factors (IGFs), which are known to activate key pro-cancer signaling pathways like the PI3K/Akt/mTOR pathway (96, 97). At the same time, it activates anti-cancer signaling pathways, such as the AMPK pathway (98). Secondly, exercise affects cancer biology by altering the levels of critical growth factors and cytokines secreted by other organs. For instance, exercise-induced myokines like IL-10 and CCL4 have been shown to directly reduce tumor cell growth and migration in pancreatic cancer patients (99). Thirdly, exercise has been observed to suppress the Hippo/YAP signaling pathway in cancer cells, thereby inhibiting tumor formation and cell viability (100). Furthermore, moderate exercise increases dopamine levels in tumor tissues, which helps inhibit cancer cell growth and lung metastasis through mechanisms dependent on dopamine receptor 2 and TGF-β1 (101).

4 Conclusion and future perspectives

This review confirms the significant anti-cancer benefits of exercise, including reducing tumor incidence, suppressing tumor growth, mitigating treatment-related side effects, and enhancing overall survival rates. Such benefits underscore the necessity of integrating exercise as a standard component of cancer care protocols across all stages of the disease.

Future research should focus on elucidating the specific molecular and cellular mechanisms by which exercise impacts cancer, which will aid in developing targeted therapeutic strategies that leverage exercise's full potential (102). A deeper understanding of these mechanisms is essential for optimizing the design of exercise programs that can be tailored to individual needs based on cancer type, stage, and patient-specific characteristics such as genetic, metabolic, and immunological profiles.

There is a compelling need for personalized exercise prescriptions to maximize the therapeutic potential of exercise in oncology (103, 104). These prescriptions should be crafted by interdisciplinary teams, including oncologists, exercise physiologists, and data scientists, to ensure that exercise interventions are safe, effective, and specifically tailored to individual patient demographics. Additionally, it is crucial to address potential risks associated with exercise, particularly for patients with advanced cancer or significant comorbidities, by developing comprehensive guidelines that ensure exercise programs are implemented safely.

Enhancing cooperation among various healthcare professionals is vital for developing more effective exercise programs tailored to the specific needs of cancer patients (105). This collaborative approach can help overcome barriers to the implementation of exercise as a therapeutic strategy and pave the way for more inclusive, holistic cancer treatment plans.

Moreover, longitudinal studies are needed to better understand the long-term effects of exercise on cancer recurrence and survival. These studies will help establish robust, evidence-based guidelines for incorporating physical activity into cancer recovery and long-term survivorship plans (106, 107). Such research is essential for substantiating the benefits of exercise in the oncology setting and for encouraging its broader adoption in routine clinical practice.

Ultimately, these efforts will better harness the potential of exercise to complement traditional cancer therapies, potentially transforming the standard of care in oncology and markedly improving patient outcomes. By advancing our understanding and integration of exercise in cancer treatment, we can hope to significantly enhance the quality of life and survival rates for cancer patients worldwide.

Author contributions

YF: Writing – original draft, Writing – review & editing. XF: Writing – original draft, Writing – review & editing. RW: Writing – original draft, Writing – review & editing. ZL: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. LQ: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. QW: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported Suzhou Clinical Key Disease Diagnosis and Treatment Technology Special Project (LCZX202127), Kunshan High-level Medical Talent Program Project (Kunshan Health [2019] No. 6), Kunshan Chinese Medicine Science and Technology Development Special Project (KZYY2202).

Acknowledgments

We thank ChatGPT 4 for the language modification.

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.

Publisher's note

References

1. Koelwyn GJ, Quail DF, Zhang X, White RM, Jones LW. Exercise-dependent regulation of the tumour microenvironment. *Nat Rev Cancer*. (2017) 17:620–32. doi: 10.1038/nrc.2017.78

2. Sarich P, Canfell K, Egger S, Banks E, Joshy G, Grogan P, et al. Alcohol consumption, drinking patterns and cancer incidence in an Australian cohort of 226,162 participants aged 45 years and over. *Br J Cancer*. (2021) 124:513–23. doi: 10.1038/s41416-020-01101-2

3. Sohn W, Lee HW, Lee S, Lim JH, Lee MW, Park CH, et al. Obesity and the risk of primary liver cancer: A systematic review and meta-analysis. *Clin Mol Hepatol*. (2021) 27:157–74. doi: 10.3350/cmh.2020.0176

4. Matthews CE, Moore SC, Arem H, Cook MB, Trabert B, Håkansson N, et al. Amount and intensity of leisure-time physical activity and lower cancer risk. *J Clin Oncol*. (2020) 38:686–97. doi: 10.1200/JCO.19.02407

5. Hojman P, Gehl J, Christensen JF, Pedersen BK. Molecular mechanisms linking exercise to cancer prevention and treatment. *Cell Metab*. (2018) 27:10–21. doi: 10.1016/j.cmet.2017.09.015

6. Fiuza-Luces C, Valenzuela PL, Gálvez BG, Ramírez M, López-Soto A, Simpson RJ, et al. The effect of physical exercise on anticancer immunity. *Nat Rev Immunol*. (2024) 24:282–93. doi: 10.1038/s41577-023-00943-0

7. Schmitz KH, Campbell AM, Stuiver MM, Pinto BM, Schwartz AL, Morris GS, et al. Exercise is medicine in oncology: Engaging clinicians to help patients move through cancer. *CA Cancer J Clin*. (2019) 69:468–84. doi: 10.3322/caac.21579

8. Mctiernan A, Friedenreich CM, Katzmarzyk PT, Powell KE, Macko R, Buchner D, et al. Physical activity in cancer prevention and survival: A systematic review. *Med Sci Sports Exerc*. (2019) 51:1252–61. doi: 10.1249/MSS.0000000000001937

9. Cannioto RA, Hutson A, Dighe S, McCann W, McCann SE, Zirpoli GR, et al. Physical activity before, during, and after chemotherapy for high-risk breast cancer: relationships with survival. *JNCI J Natl Cancer Inst*. (2021) 113:54–63. doi: 10.1093/jnci/djaa046

10. Tohme S, Simmons RL, Tsung A. Surgery for cancer: A trigger for metastases. *Cancer Res*. (2017) 77:1548–52. doi: 10.1158/0008-5472.CAN-16-1536

11. Zhang H, Chen T, Ren J, Xia Y, Onuma A, Wang Y, et al. Pre-operative exercise therapy triggers anti-inflammatory trained immunity of Kupffer cells through metabolic reprogramming. *Nat Metab*. (2021) 3:843–58. doi: 10.1038/s42255-021-00402-x

12. Gulsin GS, Henson J, Brady EM, Sargeant JA, Wilmot EG, Athithan L, et al. Cardiovascular determinants of aerobic exercise capacity in adults with type 2 diabetes. *Diabetes Care*. (2020) 43:2248–56. doi: 10.2337/dc20-0706

13. McIsaac DI, Hladkowicz E, Bryson GL, Forster AJ, Gagne S, Huang A, et al. Home-based prehabilitation with exercise to improve postoperative recovery for older adults with frailty having cancer surgery: the PREHAB randomised clinical trial. *Br J Anaesth*. (2022) 129:41–8. doi: 10.1016/j.bja.2022.04.006

14. Roh JD, Houstis N, Yu A, Chang B, Yeri A, Li H, et al. Exercise training reverses cardiac aging phenotypes associated with heart failure with preserved ejection fraction in male mice. *Aging Cell*. (2020) 19:e13159. doi: 10.1111/acel.13159

15. Bausys A, Luksta M, Anglickiene G, Maneikiene VV, Kryzauskas M, Rybakovas A, et al. Effect of home-based prehabilitation on postoperative complications after surgery for gastric cancer: randomized clinical trial. *Br J Surg*. (2023) 110:1800–7. doi: 10.1093/bjs/znad312

16. Barnes K, Hladkowicz E, Dorrance K, Bryson GL, Forster AJ, Gagné S, et al. Barriers and facilitators to participation in exercise prehabilitation before cancer surgery for older adults with frailty: a qualitative study. *BMC Geriatr*. (2023) 23:356. doi: 10.1186/s12877-023-03990-3

17. Morris ZS, Harari PM. Interaction of radiation therapy with molecular targeted agents. *J Clin Oncol*. (2014) 32:2886–93. doi: 10.1200/JCO.2014.55.1366

18. Chen HHW, Kuo MT. Improving radiotherapy in cancer treatment: Promises and challenges. *Oncotarget*. (2017) 8:62742–58. doi: 10.18632/oncotarget.18409

19. Telarovic I, Wenger RH, Pruschy M. Interfering with tumor hypoxia for radiotherapy optimization. *J Exp Clin Cancer Res*. (2021) 40:197. doi: 10.1186/s13046-021-02000-x

20. Ashcraft KA, Warner AB, Jones LW, Dewhirst MW. Exercise as adjunct therapy in cancer. *Semin Radiat Oncol*. (2019) 29:16–24. doi: 10.1016/j.semradonc.2018.10.001

21. Dufresne S, Guéritat J, Chiavassa S, Noblet C, Assi M, Rioux-Leclercq N, et al. Exercise training improves radiotherapy efficiency in a murine model of prostate cancer. *FASEB J Off Publ Fed Am Soc Exp Biol*. (2020) 34:4984–96. doi: 10.1096/fj.201901728R

22. Vikmoen O, Strandberg E, Svindland KV, Henriksson A, Mazzoni A-S, Johansson B, et al. Effects of heavy-load strength training during (neo-)adjuvant chemotherapy on muscle strength, muscle fiber size, myonuclei, and satellite cells in women with breast cancer. *FASEB J Off Publ Fed Am Soc Exp Biol*. (2024) 38:e23784. doi: 10.1096/fj.202400634R

23. Zimmer P, Schmidt ME, Prentzell MT, Berdel B, Wiskemann J, Kellner KH, et al. Resistance exercise reduces kynurenine pathway metabolites in breast cancer patients undergoing radiotherapy. *Front Oncol*. (2019) 9:962. doi: 10.3389/fonc.2019.00962

24. Gouez M, Rébillard A, Thomas A, Beaumel S, Matera E-L, Gouraud E, et al. Combined effects of exercise and immuno-chemotherapy treatments on tumor growth in MC38 colorectal cancer-bearing mice. *Front Immunol*. (2024) 15:1368550. doi: 10.3389/fimmu.2024.1368550

25. Hiensch AE, Mijwel S, Bargiela D, Wengström Y, May AM, Rundqvist H. Inflammation mediates exercise effects on fatigue in patients with breast cancer. *Med Sci Sports Exerc*. (2021) 53:496–504. doi: 10.1249/MSS.0000000000002490

26. Neuzillet C, Bouché O, Tournigand C, Chibaudel B, Bauguion L, Bengrine-Lefevre L, et al. Effect of adapted physical activity in patients with advanced pancreatic cancer: the APACaP GERCOR randomized trial. *J Natl Compr Cancer Netw JNCCN*. (2023) 21:1234–1242.e17. doi: 10.6004/jnccn.2023.7065

27. Wakefield ZR, Tanaka M, Pampo C, Lepler S, Rice L, Guingab-Cagmat J, et al. Normal tissue and tumor microenvironment adaptations to aerobic exercise enhance doxorubicin anti-tumor efficacy and ameliorate its cardiotoxicity in retired breeder mice. *Oncotarget*. (2021) 12:1737–48. doi: 10.18632/oncotarget.28057

28. Florez Bedoya CA, Cardoso ACF, Parker N, Ngo-Huang A, Petzel MQ, Kim MP, et al. Exercise during preoperative therapy increases tumor vascularity in pancreatic tumor patients. *Sci Rep*. (2019) 9:13966. doi: 10.1038/s41598-019-49582-3

29. Schauer T, Mazzoni A-S, Henriksson A, Demmelmaier I, Berntsen S, Raastad T, et al. Exercise intensity and markers of inflammation during and after (neo-) adjuvant cancer treatment. *Endocr Relat Cancer*. (2021) 28:191–201. doi: 10.1530/ERC-20-0507

30. Campbell KL, Brown JC, Lee C, Weltzien E, Li J, Sternfeld B, et al. Advances in adherence reporting of resistance training in a clinical trial during adjuvant chemotherapy for colon cancer. *Med Sci Sports Exerc*. (2024) 56:1186–95. doi: 10.1249/MSS.0000000000003395

31. Kirkham AA, Gelmon KA, Van Patten CL, Bland KA, Wollmann H, McKenzie DC, et al. Impact of exercise on chemotherapy tolerance and survival in early-stage breast cancer: A nonrandomized controlled trial. *J Natl Compr Canc Netw*. (2020) 18:1670–7. doi: 10.6004/jnccn.2020.7603

32. Chiarotto JA, Akbarali R, Bellotti L, Dranitsaris G. A structured group exercise program for patients with metastatic cancer receiving chemotherapy and CTNNB1 (β-catenin) as a biomarker of exercise efficacy. *Cancer Manag Res*. (2017) 9:495–501. doi: 10.2147/CMAR.S147054

33. Groen WG, Naaktgeboren WR, van Harten WH, van Vulpen JK, Kool N, Sonke GS, et al. Physical fitness and chemotherapy tolerance in patients with early-stage breast cancer. *Med Sci Sports Exerc*. (2022) 54:537–42. doi: 10.1249/MSS.0000000000002828

34. Gustafson MP, Wheatley-Guy CM, Rosenthal AC, Gastineau DA, Katsanis E, Johnson BD, et al. Exercise and the immune system: taking steps to improve responses to cancer immunotherapy. *J Immunother Cancer*. (2021) 9:e001872. doi: 10.1136/jitc-2020-001872

35. Yan H, Jiang A, Huang Y, Zhang J, Yang W, Zhang W, et al. Exercise sensitizes PD-1/PD-L1 immunotherapy as a hypoxia modulator in the tumor microenvironment of melanoma. *Front Immunol*. (2023) 14:1265914. doi: 10.3389/fimmu.2023.1265914

36. Farley MJ, Bartlett DB, Skinner TL, Schaumberg MA, Jenkins DG. Immunomodulatory function of interleukin-15 and its role in exercise, immunotherapy, and cancer outcomes. *Med Sci Sports Exerc*. (2023) 55:558–68. doi: 10.1249/MSS.0000000000003067

37. Stout NL, Santa Mina D, Lyons KD, Robb K, Silver JK. A systematic review of rehabilitation and exercise recommendations in oncology guidelines. *CA Cancer J Clin*. (2021) 71:149–75. doi: 10.3322/caac.21639

38. Mustian KM, Alfano CM, Heckler C, Kleckner AS, Kleckner IR, Leach CR, et al. Comparison of pharmaceutical, psychological, and exercise treatments for cancer-related fatigue: A meta-analysis. *JAMA Oncol*. (2017) 3:961. doi: 10.1001/jamaoncol.2016.6914

39. Wagoner CW, Lee JT, Battaglini CL. Community-based exercise programs and cancer-related fatigue: a systematic review and meta-analysis. *Support Care Cancer Off J Multinatl Assoc Support Care Cancer*. (2021) 29:4921–9. doi: 10.1007/s00520-021-06135-7

40. Bower JE, Bak K, Berger A, Breitbart W, Escalante CP, Ganz PA, et al. Screening, assessment, and management of fatigue in adult survivors of cancer: an american society of clinical oncology clinical practice guideline adaptation. *J Clin Oncol*. (2014) 32:1840–50. doi: 10.1200/JCO.2013.53.4495

41. Ngo-Huang AT, Parker NH, Xiao L, SChadler KL, Petzel MQB, Prakash LR, et al. Effects of a pragmatic home-based exercise program concurrent with neoadjuvant therapy on physical function of patients with pancreatic cancer: the pancFit randomized clinical trial. *Ann Surg*. (2023) 278:22–30. doi: 10.1097/SLA.0000000000005878

42. Mavropalias G, Sim M, Taaffe DR, Galvão DA, Spry N, Kraemer WJ, et al. Exercise medicine for cancer cachexia: targeted exercise to counteract mechanisms and treatment side effects. *J Cancer Res Clin Oncol*. (2022) 148:1389–406. doi: 10.1007/s00432-022-03927-0

43. Hardee JP, Counts BR, Carson JA. Understanding the role of exercise in cancer cachexia therapy. *Am J Lifestyle Med*. (2019) 13:46–60. doi: 10.1177/1559827617725283

44. Franco FDO, Lopes MA, Henriques FDS, Neves RXD, Bianchi Filho C, Batista ML. Cancer cachexia differentially regulates visceral adipose tissue turnover. *J Endocrinol*. (2017) 232:493–500. doi: 10.1530/JOE-16-0305

45. Dodson S, Baracos VE, Jatoi A, Evans WJ, Cella D, Dalton JT, et al. Muscle wasting in cancer cachexia: clinical implications, diagnosis, and emerging treatment strategies. *Annu Rev Med*. (2011) 62:265–79. doi: 10.1146/annurev-med-061509-131248

46. Leal LG, Lopes MA, Peres SB, Batista ML. Exercise training as therapeutic approach in cancer cachexia: A review of potential anti-inflammatory effect on muscle wasting. *Front Physiol*. (2021) 11:570170. doi: 10.3389/fphys.2020.570170

47. Raun SH, Buch-Larsen K, Schwarz P, Sylow L. Exercise-A panacea of metabolic dysregulation in cancer: physiological and molecular insights. *Int J Mol Sci*. (2021) 22:3469. doi: 10.3390/ijms22073469

48. Pedersen L, Idorn M, Olofsson GH, Lauenborg B, Nookaew I, Hansen RH, et al. Voluntary running suppresses tumor growth through epinephrine- and IL-6-dependent NK cell mobilization and redistribution. *Cell Metab*. (2016) 23:554–62. doi: 10.1016/j.cmet.2016.01.011

49. Gallot YS, Durieux A-C, Castells J, Desgeorges MM, Vernus B, Plantureux L, et al. Myostatin gene inactivation prevents skeletal muscle wasting in cancer. *Cancer Res*. (2014) 74:7344–56. doi: 10.1158/0008-5472.CAN-14-0057

50. Kir S, White JP, Kleiner S, Kazak L, Cohen P, Baracos VE, et al. Tumour-derived PTH-related protein triggers adipose tissue browning and cancer cachexia. *Nature*. (2014) 513:100–4. doi: 10.1038/nature13528

51. Barnes O, Wilson RL, Gonzalo-Encabo P, Kang D-W, Christopher CN, Bentley T, et al. The effect of exercise and nutritional interventions on body composition in patients with advanced or metastatic cancer: A systematic review. *Nutrients*. (2022) 14:2110. doi: 10.3390/nu14102110

52. Solheim TS, Laird BJA, Balstad TR, Stene GB, Bye A, Johns N, et al. A randomized phase II feasibility trial of a multimodal intervention for the management of cachexia in lung and pancreatic cancer. *J Cachexia Sarcopenia Muscle*. (2017) 8:778–88. doi: 10.1002/jcsm.12201

53. Mandelblatt JS, Small BJ, Luta G, Hurria A, Jim H, McDonald BC, et al. Cancer-related cognitive outcomes among older breast cancer survivors in the thinking and living with cancer study. *J Clin Oncol*. (2018) 36:3211–22. doi: 10.1200/JCO.18.00140

54. Speidell AP, Demby T, Lee Y, Rodriguez O, Albanese C, Mandelblatt J, et al. Development of a human APOE knock-in mouse model for study of cognitive function after cancer chemotherapy. *Neurotox Res*. (2019) 35:291–303. doi: 10.1007/s12640-018-9954-7

55. Salerno EA, Culakova E, Kleckner AS, Heckler CE, Lin P-J, Matthews CE, et al. Physical activity patterns and relationships with cognitive function in patients with breast cancer before, during, and after chemotherapy in a prospective, nationwide study. *J Clin Oncol*. (2021) 39:3283–92. doi: 10.1200/JCO.20.03514

56. Naaktgeboren WR, Koevoets EW, Stuiver MM, van Harten WH, Aaronson NK, van der Wall E, et al. Effects of physical exercise during adjuvant chemotherapy for breast cancer on long-term tested and perceived cognition: results of a pragmatic follow-up study. *Breast Cancer Res Treat*. (2024) 205:75–86. doi: 10.1007/s10549-023-07220-7

57. Lugano R, Ramachandran M, Dimberg A. Tumor angiogenesis: causes, consequences, challenges and opportunities. *Cell Mol Life Sci*. (2020) 77:1745–70. doi: 10.1007/s00018-019-03351-7

58. Esteves M, Monteiro MP, Duarte JA. The effects of physical exercise on tumor vasculature: systematic review and meta-analysis. *Int J Sports Med*. (2021) 42:1237–49. doi: 10.1055/a-1533-1876

59. Esteves M, Monteiro MP, Duarte JA. Role of regular physical exercise in tumor vasculature: favorable modulator of tumor milieu. *Int J Sports Med*. (2021) 42:389–406. doi: 10.1055/a-1308-3476

60. Bennewith KL, Durand RE. Quantifying transient hypoxia in human tumor xenografts by flow cytometry. *Cancer Res*. (2004) 64:6183–9. doi: 10.1158/0008-5472.CAN-04-0289

61. Betof AS, Lascola CD, Weitzel D, Landon C, Scarbrough PM, Devi GR, et al. Modulation of murine breast tumor vascularity, hypoxia, and chemotherapeutic response by exercise. *JNCI J Natl Cancer Inst*. (2015) 107(5):djv040. doi: 10.1093/jnci/djv040

62. Fiorenza M, Gliemann L, Brandt N, Bangsbo J. Hormetic modulation of angiogenic factors by exercise-induced mechanical and metabolic stress in human skeletal muscle. *Am J Physiol-Heart Circ Physiol*. (2020) 319(4):H824–H834. doi: 10.1152/ajpheart.00432.2020

63. SChadler KL, Thomas NJ, Galie PA, Bhang DH, Roby KC, Addai P, et al. Tumor vessel normalization after aerobic exercise enhances chemotherapeutic efficacy. *Oncotarget*. (2016) 7:65429–40. doi: 10.18632/oncotarget.11748

64. Gallo O, Fini-Storchi I, Vergari WA, Masini E, Morbidelli L, Ziche M, et al. Role of nitric oxide in angiogenesis and tumor progression in head and neck cancer. *JNCI J Natl Cancer Inst*. (1998) 90:587–96. doi: 10.1093/jnci/90.8.587

65. Perego S, Sansoni V, Ziemann E, Lombardi G. Another weapon against cancer and metastasis: physical-activity-dependent effects on adiposity and adipokines. *Int J Mol Sci*. (2021) 22:2005. doi: 10.3390/ijms22042005

66. Kim J-S, Galvão DA, Newton RU, Gray E, Taaffe DR. Exercise-induced myokines and their effect on prostate cancer. *Nat Rev Urol*. (2021) 18:519–42. doi: 10.1038/s41585-021-00476-y

67. Hatfield SM, Kjaergaard J, Lukashev D, Schreiber TH, Belikoff B, Abbott R, et al. Immunological mechanisms of the antitumor effects of supplemental oxygenation. *Sci Transl Med*. (2015) 7(277):277ra30. doi: 10.1126/scitranslmed.aaa1260

68. Luo Z, Wan R, Liu S, Feng X, Peng Z, Wang Q, et al. Mechanisms of exercise in the treatment of lung cancer – a mini-review. *Front Immunol*. (2023) 14:1244764. doi: 10.3389/fimmu.2023.1244764

69. Huang C-W, Chang Y-H, Lee H-H, Wu J-Y, Huang J-X, Chung Y-H, et al. Irisin, an exercise myokine, potently suppresses tumor proliferation, invasion, and growth in glioma. *FASEB J Off Publ Fed Am Soc Exp Biol*. (2020) 34:9678–93. doi: 10.1096/fj.202000573RR

70. Mannelli M, Gamberi T, Magherini F, Fiaschi T. The adipokines in cancer cachexia. *Int J Mol Sci*. (2020) 21:4860. doi: 10.3390/ijms21144860

71. Pu X, Chen D. Targeting adipokines in obesity-related tumors. *Front Oncol*. (2021) 11:685923. doi: 10.3389/fonc.2021.685923

72. Vona-Davis L, Rose DP. Adipokines as endocrine, paracrine, and autocrine factors in breast cancer risk and progression. *Endocr Relat Cancer*. (2007) 14:189–206. doi: 10.1677/ERC-06-0068

73. Liang C, Li X, Song G, Schmidt SF, Sun L, Chen J, et al. Adipose Kiss1 controls aerobic exercise-related adaptive responses in adipose tissue energy homeostasis. *FASEB J Off Publ Fed Am Soc Exp Biol*. (2024) 38:e23743. doi: 10.1096/fj.202302598RR

74. Supriya R, Delfan M, Saeidi A, Samaie SS, Al Kiyumi MH, Escobar KA, et al. Spirulina supplementation with high-intensity interval training decreases adipokines levels and cardiovascular risk factors in men with obesity. *Nutrients*. (2023) 15:4891. doi: 10.3390/nu15234891

75. Golbidi S, Laher I. Exercise induced adipokine changes and the metabolic syndrome. *J Diabetes Res*. (2014) 2014:1–16. doi: 10.1155/2014/726861

76. Sturgeon K, Digiovanni L, Good J, Salvatore D, Fenderson D, Domchek S, et al. Exercise-induced dose-response alterations in adiponectin and leptin levels are dependent on body fat changes in women at risk for breast cancer. *Cancer Epidemiol Biomarkers Prev*. (2016) 25:1195–200. doi: 10.1158/1055-9965.EPI-15-1087

77. Wada A, Yamashita H, Togashi A, Ogawa S, Muroi A, Kido S, et al. Combination of parenteral amino acid infusion and intermittent loading exercise ameliorates progression of postoperative sarcopenia in rat model. *Nutrients*. (2024) 16:1218. doi: 10.3390/nu16081218

78. Raitakari OT, Porkka KVK, Räsänen L, Viikari JSA. Relations of life-style with lipids, blood pressure and insulin in adolescents and young adults. The Cardiovascular Risk in Young Finns Study. *Atherosclerosis*. (1994) 111:237–46. doi: 10.1016/0021-9150(94)90098-1

79. Stanford KI, Goodyear LJ. Exercise and type 2 diabetes: molecular mechanisms regulating glucose uptake in skeletal muscle. *Adv Physiol Educ*. (2014) 38:308–14. doi: 10.1152/advan.00080.2014

80. Kim J-S, Wilson RL, Taaffe DR, Galvão DA, Gray E, Newton RU. Myokine Expression and Tumor-Suppressive Effect of Serum after 12 wk of Exercise in Prostate Cancer Patients on ADT. *Med Sci Sports Exerc*. (2022) 54:197–205. doi: 10.1249/MSS.0000000000002783

81. Febvey-Combes O, Jobard E, Rossary A, Pialoux V, Foucaut A-M, Morelle M, et al. Effects of an exercise and nutritional intervention on circulating biomarkers and metabolomic profiling during adjuvant treatment for localized breast cancer: results from the PASAPAS feasibility randomized controlled trial. *Integr Cancer Ther*. (2021) 20:153473542097766. doi: 10.1177/1534735420977666

82. Dieli-Conwright CM, Courneya KS, Demark-Wahnefried W, Sami N, Lee K, Buchanan TA, et al. Effects of aerobic and resistance exercise on metabolic syndrome, sarcopenic obesity, and circulating biomarkers in overweight or obese survivors of breast cancer: A randomized controlled trial. *J Clin Oncol*. (2018) 36:875–83. doi: 10.1200/JCO.2017.75.7526

83. Orange ST, Leslie J, Ross M, Mann DA, Wackerhage H. The exercise IL-6 enigma in cancer. *Trends Endocrinol Metab TEM*. (2023) 34:749–63. doi: 10.1016/j.tem.2023.08.001

84. Zimmer P, Bloch W, Schenk A, Zopf E, Hildebrandt U, Streckmann F, et al. Exercise-induced natural killer cell activation is driven by epigenetic modifications. *Int J Sports Med*. (2015) 36:510–5. doi: 10.1055/s-0034-1398531

85. Wennerberg E, Lhuillier C, Rybstein MD, Dannenberg K, Rudqvist N-P, Koelwyn GJ, et al. Exercise reduces immune suppression and breast cancer progression in a preclinical model. *Oncotarget*. (2020) 11:452–61. doi: 10.18632/oncotarget.27464

86. Abdalla DR, Aleixo AAR, Murta EFC, Michelin MA. Innate immune response adaptation in mice subjected to administration of DMBA and physical activity. *Oncol Lett*. (2014) 7:886–90. doi: 10.3892/ol.2013.1774

87. Xiao Y, Yu D. Tumor microenvironment as a therapeutic target in cancer. *Pharmacol Ther*. (2021) 221:107753. doi: 10.1016/j.pharmthera.2020.107753

88. Almeida PWM, Gomes-Filho A, Ferreira AJ, Rodrigues CEM, Dias-Peixoto MF, Russo RC, et al. Swim training suppresses tumor growth in mice. *J Appl Physiol*. (2009) 107:261–5. doi: 10.1152/japplphysiol.00249.2009

89. Zielinski MR, Muenchow M, Wallig MA, Horn PL, Woods JA. Exercise delays allogeneic tumor growth and reduces intratumoral inflammation and vascularization. *J Appl Physiol*. (2004) 96:2249–56. doi: 10.1152/japplphysiol.01210.2003

90. Xiao Y, Cong M, Li J, He D, Wu Q, Tian P, et al. Cathepsin C promotes breast cancer lung metastasis by modulating neutrophil infiltration and neutrophil extracellular trap formation. *Cancer Cell*. (2021) 39:423–437.e7. doi: 10.1016/j.ccell.2020.12.012

91. Gomes-Santos IL, Amoozgar Z, Kumar AS, Ho WW, Roh K, Talele NP, et al. Exercise training improves tumor control by increasing CD8+ T-cell infiltration via CXCR3 signaling and sensitizes breast cancer to immune checkpoint blockade. *Cancer Immunol Res*. (2021) 9:765–78. doi: 10.1158/2326-6066.CIR-20-0499

92. Rundqvist H, Veliça P, Barbieri L, Gameiro PA, Bargiela D, Gojkovic M, et al. Cytotoxic T-cells mediate exercise-induced reductions in tumor growth. *eLife*. (2020) 9:e59996. doi: 10.7554/eLife.59996

93. Kurz E, Hirsch CA, Dalton T, Shadaloey SA, Khodadadi-Jamayran A, Miller G, et al. Exercise-induced engagement of the IL-15/IL-15Rα axis promotes anti-tumor immunity in pancreatic cancer. *Cancer Cell*. (2022) 40:720–737.e5. doi: 10.1016/j.ccell.2022.05.006

94. Runowicz CD, Leach CR, Henry NL, Henry KS, Mackey HT, Cowens-Alvarado RL, et al. American cancer society/american society of clinical oncology breast cancer survivorship care guideline. *CA Cancer J Clin*. (2016) 66:43–73. doi: 10.3322/caac.21319

95. Kim J-S, Taaffe DR, Galvão DA, Hart NH, Gray E, Ryan CJ, et al. Exercise in advanced prostate cancer elevates myokine levels and suppresses *in-vitro* cell growth. *Prostate Cancer Prostatic Dis*. (2022) 25:86–92. doi: 10.1038/s41391-022-00504-x

96. Thompson HJ, Jiang W, Zhu Z. Candidate mechanisms accounting for effects of physical activity on breast carcinogenesis. *IUBMB Life*. (2009) 61:895–901. doi: 10.1002/iub.233

97. Zhu Z, Jiang W, Sells JL, Neil ES, McGinley JN, Thompson HJ. Effect of nonmotorized wheel running on mammary carcinogenesis: circulating biomarkers, cellular processes, and molecular mechanisms in rats. *Cancer Epidemiol Biomarkers Prev*. (2008) 17:1920–9. doi: 10.1158/1055-9965.EPI-08-0175

98. Xia Y, Xu F, Xiong M, Yang H, Lin W, Xie Y, et al. Repurposing of antipsychotic trifluoperazine for treating brain metastasis, lung metastasis and bone metastasis of melanoma by disrupting autophagy flux. *Pharmacol Res*. (2021) 163:105295. doi: 10.1016/j.phrs.2020.105295

99. Schwappacher R, Dieterich W, Reljic D, Pilarsky C, Mukhopadhyay D, Chang DK, et al. Muscle-derived cytokines reduce growth, viability and migratory activity of pancreatic cancer cells. *Cancers*. (2021) 13:3820. doi: 10.3390/cancers13153820

100. Dethlefsen C, Hansen LS, Lillelund C, Andersen C, Gehl J, Christensen JF, et al. Exercise-induced catecholamines activate the hippo tumor suppressor pathway to

reduce risks of breast cancer development. *Cancer Res*. (2017) 77:4894–904. doi: 10.1158/0008-5472.CAN-16-3125

101. Zhang Q-B, Zhang B-H, Zhang K-Z, Meng X-T, Jia Q-A, Zhang Q-B, et al. Moderate swimming suppressed the growth and metastasis of the transplanted liver cancer in mice model: with reference to nervous system. *Oncogene*. (2016) 35:4122–31. doi: 10.1038/onc.2015.484

102. Luo Z, Mei J, Wang X, Wang R, He Z, Geffen Y, et al. Voluntary exercise sensitizes cancer immunotherapy via the collagen inhibition-orchestrated inflammatory tumor immune microenvironment. *Cell Rep*. (2024) 43(9):114697. doi: 10.1016/j.celrep.2024.114697

103. Luo Z, Zhu J, Xu R, Wan R, He Y, Chen Y, et al. Exercise-downregulated CD300E acted as a negative prognostic implication and tumor-promoted role in pan-cancer. *Front Immunol*. (2024) 15:1437068. doi: 10.3389/fimmu.2024.1437068

104. Luo Z, Zhu J, Fang Z, Xu R, Wan R, He Y, et al. Exercise-augmented THSD7B exhibited a positive prognostic implication and tumor-suppressed functionality in pan-cancer. *Front Immunol*. (2024) 15:1440226. doi: 10.3389/fimmu.2024.1440226

105. Luo Z, Zhang T, Chen S. Exercise Prescription: Pioneering the "Third Pole" for Clinical Health Management. *Research (Wash D C)*. (2023) 6:0284. doi: 10.34133/research.0284

106. Luo Z, Wan T, Liu S, Feng X, Peng Z, Wang Q, et al. Mechanisms of exercise in the treatment of lung cancer - a mini-review. *Front Immunol*. (2023) 14:1244764. doi: 10.3389/fimmu.2023.1244764

107. Luo ZW, Sun YY, Xia W, Xu JY, Xie DJ, Jiao CM, et al. Physical exercise reverses immuno-cold tumor microenvironment via inhibiting SQLE in non-small cell lung cancer. *Mil Med Res*. (2023) 10(1):39. doi: 10.1186/s40779-023-00474-8

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 23 September 2024
DOI 10.3389/fimmu.2024.1435334

OPEN ACCESS

EDITED BY
Raquel Alarcon Rodriguez,
University of Almeria, Spain

REVIEWED BY
Lola Rueda-Ruzafa,
University of Almeria, Spain
Xuehua Wan,
Nankai University, China

*CORRESPONDENCE
Jiali Zhang
zhangjiali@5thhospital.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 20 May 2024
ACCEPTED 05 September 2024
PUBLISHED 23 September 2024

CITATION
Qian C, Hui J, Peng Z, Sun X and Zhang J (2024) Mucosal microbiota characterization in gastric cancer identifies immune-activated–related transcripts relevant gastric microbiome signatures. *Front. Immunol.* 15:1435334. doi: 10.3389/fimmu.2024.1435334

Mucosal microbiota characterization in gastric cancer identifies immune-activated–related transcripts relevant gastric microbiome signatures

Chengjia Qian[1†], Jiang Hui[1†], Ziyao Peng[2†], Xiaoyan Sun[3] and Jiali Zhang[4*]

[1]Department of Gastrointestinal Surgery, Affiliated Hospital of Jiangnan University, Wuxi, China, [2]Department of Trauma-Emergency & Critical Care Medicine, Shanghai Fifth People's Hospital, Fudan University, Shanghai, China, [3]Department of Nuclear Medicine, Shanghai Fifth People's Hospital, Fudan University, Shanghai, China, [4]Central Laboratory, Shanghai Fifth People's Hospital, Fudan University, Shanghai, China

Tumor microenvironment (TME) immune cells and gastric mucosal microbiome constitute two vital elements of tumor tissue. Increasing evidence has elucidated their clinicopathological significance in predicting outcomes and therapeutic efficacy. However, comprehensive characterization of immune cell-associated microbiome signatures in the TME is still in the early stages of development. Here, we characterized the gastric mucosa microbiome and its associations with immune-activated related transcripts (IATs) in 170 GC tumor tissues and matched non-tumor tissues using 16s rRNA gene sequencing and quantitative reverse transcription-PCR. Microbial diversity and richness were significantly higher in GC tumor tissues than in non-tumor tissues. Differences in microbial composition between the groups were evident, with Firmicutes, Proteobacteria, Bacteroidota, Campilobacterota, Actinobacteria, Fusobacteriota, Verrucomicrobiota, Acidobacteriota, and Cyanobacteria being the dominant phyla in the gastric mucosal microbiota. Microbial interaction network analysis revealed distinctive centralities of oral bacteria (such as *Fusobacterium, Porphyromonas, Prevotella*, etc.) in both tumor and normal mucosae networks, suggesting their significant influence on GC microbial ecology. Furthermore, we analyzed the expression of IATs (CXCL9, CXCL10, GZMA, GZMB, PRF1, CD8A, IFNG, TBX2, and TNF) and characterized IAT-relevant gastric microbiome signatures in GC patients. Our results showed that the expression of CXCL9, CXCL10, GZMA, GZMB, PRF1 and IFNG was significantly higher in tumor tissues than in adjacent normal tissues in GC patients. Notably, high expression of IATs in tumor tissues was associated with improved survival in GC patients and could serve as a powerful predictor for disease-free survival. Additionally, analysis of IAT levels and mucosal microbiota diversity revealed a correlation between higher IAT expression and increased microbiota richness and evenness in the IATs ^{high} group, suggesting potential interactions between mucosal microbiota and tumor immunopathology. Spearman correlation analysis showed positive associations between IAT expression and specific mucosal bacterial species. Notably, *Akkermansia muciniphila* demonstrated potential involvement in modulating GZMB expression in the GC mucosal

10.3389/fimmu.2024.1435334

microenvironment. These findings underscore the importance of mucosal microbiota alterations in GC and suggest potential therapeutic targets focusing on modulating the tumor microbiota for improved clinical outcomes. The detailed characterization of these elements has profound implications for both treatment and survival prediction in GC. We observed that microbial diversity and richness were significantly higher in GC tumor tissues compared to non-tumor tissues. These differences highlight the unique microbial landscape of GC tumors and suggest that the microbiome could influence tumor development and progression. Importantly, our study demonstrated that high expression levels of IATs in GC tumor tissues were associated with improved patient survival. This suggests that IATs not only reflect immune activation but also serve as valuable biomarkers for predicting disease-free survival. The potential of IATs as predictive markers underscores their utility in guiding therapeutic strategies and personalizing treatment approaches. Moreover, the correlation between higher IAT expression and increased microbiota richness and evenness suggests that a diverse and balanced microbiome may enhance immune responses and contribute to better clinical outcomes. These findings highlight the critical need to consider mucosal microbiota alterations in GC management. Targeting the tumor microbiota could emerge as a promising therapeutic strategy, potentially leading to more effective treatments and improved patient outcomes. Understanding and modulating the microbiome's role in GC opens new avenues for innovative therapeutic interventions and personalized medicine.

KEYWORDS

mucosal microbiota, gastric cancer, immune-activated, transcripts, gastric microbiome signatures

Introduction

Gastric cancer (GC) is the fifth most common cancer worldwide and over 1 million new cases were diagnosed in 2020 (1). In China, GC was responsible for more than 509,421 new cases and 400,415 deaths in 2022 (2), making it the third most frequently diagnosed cancer and the third leading cause of cancer-related deaths. One of the primary risk factors for GC is infection with *Helicobacter pylori*, as the majority of GC cases are associated with this pathogen (3). Advances in sequencing technology have revealed that the stomach hosts a diverse microbiota beyond *H. pylori*. Notably, studies have found that the microbiota in GC patients was associated with decreased diversity and richness compared with intestinal metaplasia (4). Understanding how the microbiota composition in *H. pylori*-positive GC patients affects the local tumor microenvironment (TME) warrants further investigation.

To assess the immune contexture within the TME, numerous models (5–7) utilizing immunoscoring have been developed. These models provide robust statistical parameters for prognostic evaluation and therapeutic efficacy across various solid tumors, including GC (8). Traditionally, immunohistochemistry has been the predominant method for investigating cellular heterogeneity. However, immunohistochemistry has limitations, including a restricted set of phenotypic markers and the requirement for sizable biopsy specimens. Technical constraints in turn resulted in studies marked by small sample sizes, a scarcity of cell types, or both. Additionally, achieving standardized and reproducible staining intensity measurement, crucial for accurate protein expression quantification, remains inherently challenging in immunohistochemistry.

Recent innovations in prognostic tools aim to improve survival predictions post-GC diagnosis. These tools employ a novel computational algorithm to enumerate immune cell subsets from RNA specimens sourced from various tissue types, encompassing solid tumors (9–11). Furthermore, contemporary immune profiling studies have delved into the cytokine and chemokine milieu characterizing each gene cluster predictive of survival in patient cohorts sharing identical TNM stages. Analysis of the expression patterns of selected cytokine and chemokine mRNAs in 299 GC samples unveiled CXCL9, CXCL10, GZMA, GZMB, PRF1, CD8A, IFNG, TBX2, and TNF as immune activation-related transcripts (IATs), serving as robust statistical parameters for prognostic assessments in GC patients (10). Notably, CXCL9 and CXCL10 have been shown to cooperate in recruiting effector T cells into tumors. Newly strategies including plasmid-borne CXCL9 (12), intratumor injection of CXCL9 (13), recombinant CXCL10

protein with adoptive cell therapy (ACT) (14), intra-tumor injection of CXCL10 (15), retroviral transduction tumor cells with CXCL10 (16, 17) were effective in increasing T cell infiltration and reducing tumor growth in animal models (18). Despite their potential, these strategies have yet to be explored in clinical trials, partly due to challenges such as the limited bioavailability of injected proteins. Additionally, research has indicated that commensal bacteria colonizing could trigger activation of immune cells to express the chemokine CXCL10 which leaded to the formation of CXCL10–bacterial DNA complexes (19). The gastrointestinal mucosa is a well-studied interface for microbiota-IAT interactions. However, profiling of mucosal microbiota and IATs associated microbiome in the GC patients were lack. Therefore, further studies analyzing these interactions in gastric mucosa from cancer patients are urgently needed.

Recent advances in high-throughput sequencing based on conserved 16S ribosomal RNA and newly developed computational methods have uncovered a complex and distinct bacterial community that inhabits in the tumor mucosa compared with non-tumor mucosa, in addition to *H. pylori*. Species such as *Prevotella melaninogenica, Streptococcus anginosus* and *Propionibacterium acnes* have been identified (20, 21). It remains unclear whether the presence of *H. pylori* shapes the microbiota composition in gastric mucosa compared with non-tumor mucosa. Some studies suggest that *H. pylori* infection induces inflammation in the gastric mucosa, with changes in gastric acid and gastrin secretion, resulting in the gastric mucosa bacterial shifting (22, 23). However, the microbial profiling of GC mucosa with *H. pylori* infection and its association with IATs remain scarce.

To address this gap, we conducted this study employing 16s rRNA gene sequencing on tumor tissues and matched non-tumor tissues from 85 GC patients with *H. pylori* infection. This approach allowed us to characterize the mucosa-associated microbiota comprehensively. We also performed quantitative reverse transcription-PCR analysis of the paired GC tissue samples to quantify key IATs, including CXCL9, CXCL10, GZMA, GZMB, PRF1, CD8A, IFNG, TBX2, and TNF. By combining these analyses, we aimed to identify IATs relevant gastric microbiome signatures.

TABLE 1 Characteristics of Patients.

Characteristics	Patients (n = 85)
Age (means ± SD)	65.44± 21.56
Gender (Female/Male)	30/55
Weight (Kg, means ± SD)	66.1 ± 25.9
Height (cm, means ± SD)	166.5± 16.5
BMI (means ± SD)	23.73± 5.67
Complications, no	
Hypertension	35
Diabetes mellitus	10
Tumor localization, no	
Proximal stomach	21
Body/Fundus	25
Antrum	39
Tumor differentiation, no	
High differentiated	2
Moderately/poor differentiated	83
Lauren typing, no	
Intestinal type	16
Diffuse type	8
Mixed type	61
Tumor stage, no	
I (Ia, Ib)	12
II (IIa, IIb)	19
III (IIIa, IIIb, IIIc)	46
IV	8
HP infection,	
Positive	85
Negative	0
Antibiotics use, no	0
Pre-operative chemotherapy, no	0

BMI, Body mass index; HP, Helicobacter pylori; no, number; SD, standard deviation.

Results

Altered gastric mucosal microbiota in GC tumor tissues compared with matched non-tumor tissues

In this study, we investigated the microbial composition of gastric tumor tissues and compared it with matched non-tumor tissues from GC patients (Table 1), focusing on alterations in gastric mucosal microbiota. 16s rRNA gene sequencing yielded a median of 73,634 clean reads for 170 paired tumor and non-tumor tissues. To assess differences in microbial diversity, we analyzed alpha diversity measures. The observed OTUs, which reflects species richness, were significantly higher in tumor tissues than in non-tumor tissues (623.68 vs. 493.00; P = 0.01; Supplementary Figure S1). Additional alpha diversity indices, such as the Shannon, Simpson, and Pielou indices, also showed higher values in tumor tissues (P = 0.009; P = 0.033; P = 0.019; Figure 1A). Similarly, indices measuring species evenness, including ACE, Chao1, and Faith_PD, were significantly higher in tumor tissues (P = 0.004; P = 0.004; P = 0.006; Figure 1A). However, due to significant inter-individual variation, principal coordinate analysis (PCoA) could not separate the tumor and non-tumor mucosa microbiomes into distinct clusters (Supplementary Figure S1).

Given the observed differences in gastric microbiota between non-tumor and tumor tissues, which were dominated by Firmicutes, Proteobacteria, Bacteroidota, Campilobacterota, Actinobacteria, Fusobacteriota, Verrucomicrobiota, Acidobacteriota, and Cyanobacteria (Figures 1B, 2A), we hypothesized that there is a

10.3389/fimmu.2024.1435334

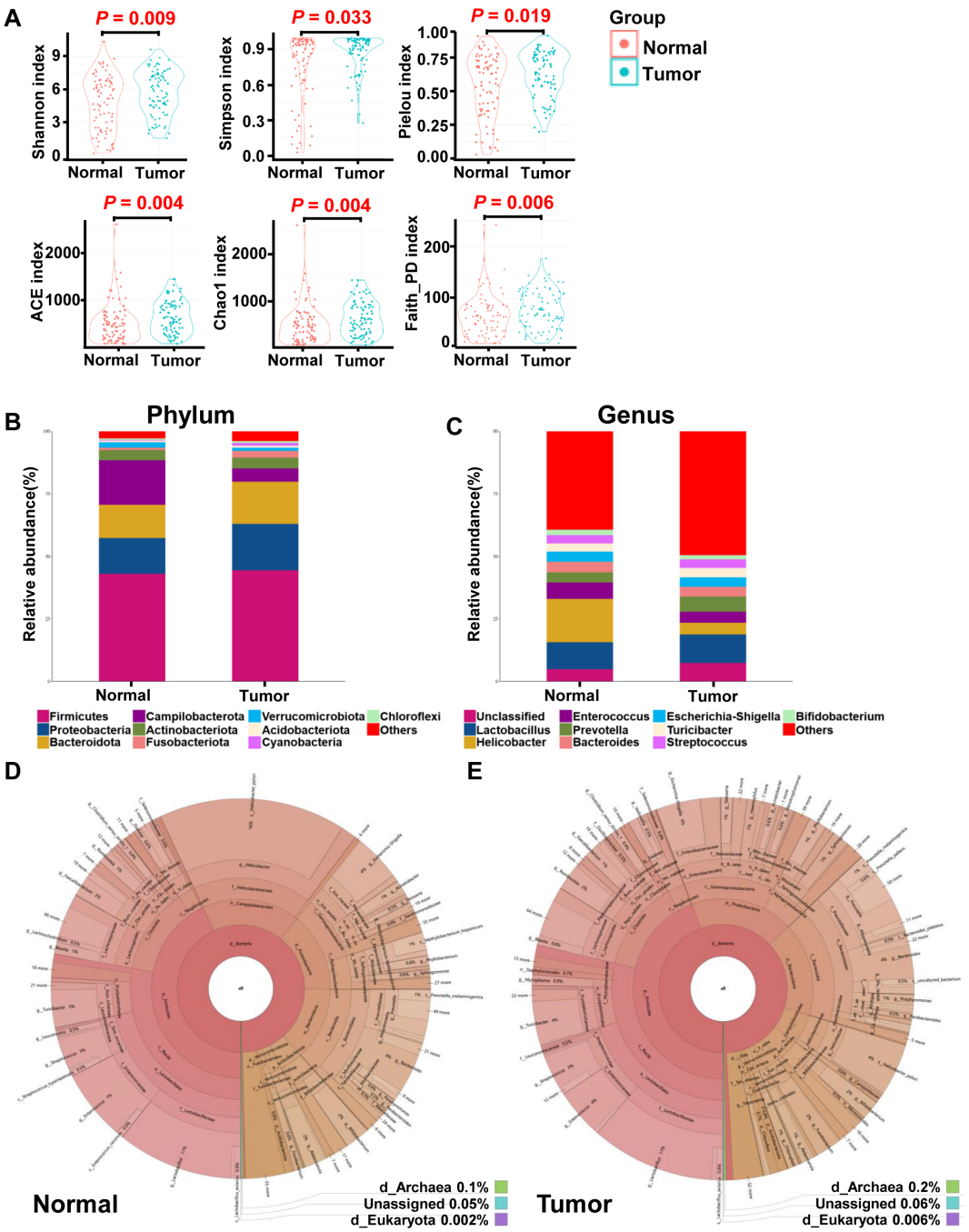


FIGURE 1
Overall structure and composition map of the gastric microbiota in the tumor and matched normal mucosae tissues. The α-diversity indices (Shannon, Simpson, Pielou, ACE, Choa and Faith_PD) were used to evaluate the overall structure of mucosae microbiota in GC patients **(A)**. The two-sided Wilcoxon signed rank test was utilized to analyze variations between tumor and adjacent normal mucosae. The microbiota structures at phylum **(B)** and genus **(C)** levels in both tumor and normal mucosae tissues are depicted, along with Krona species composition plots for normal **(D)** and tumor **(E)** tissues.

shift in mucosal microbiome profiles in GC patients. The top 10 genera showed in Figure 1C included *Lactobacillus*, *Helicobacter*, *Enterococcus*, *Prevotella*, *Bacteroides*, *Escherichia-Shigella*, *Turicibacter*, *Streptococcus*, and *Bifidobacterium*. Notably, the Proteobacteria/Campylobacterales ratio was significantly higher in the tumor mucosa group (p = .000; Figures 1D, E).

ANOVA was used to assess mean differences in species abundance at various taxonomic levels. Significant differences were observed at the phylum and genus levels (Figures 2A, B). At the phylum level, differences were noted inFirmicutes, Proteobacteria, Campilobacterota, Bacteroidota, Acidobacteriota, Actinobacteriota, Chloroflexi, Cyanobacteria, Fusobacteriota, and Verrucomicrobiota.

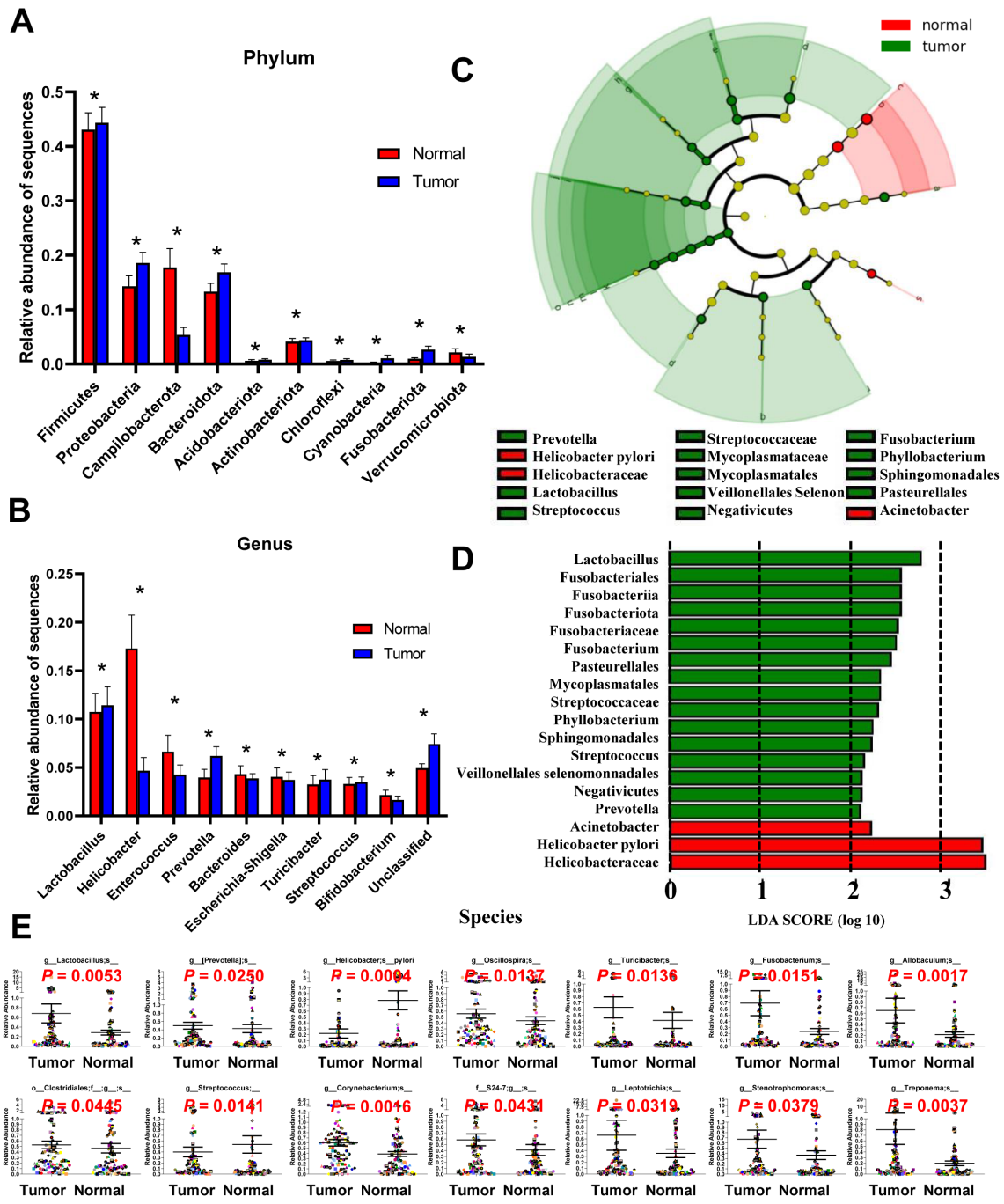


FIGURE 2
Different bacterial taxa between tumor and matched normal mucosae microbiomes. Relative abundance comparisons at the bacterial phylum **(A)** and genus **(B)** levels are presented; *P<0.05. The LEfSe analysis identifies taxa with significant differences in abundance between tumor and adjacent tissues **(C)**, with only those exceeding a significant LDA threshold value of >2 displayed **(D)**. Fourteen differentially abundant bacterial species were identified **(E)**, and representative dot plots illustrate their relative abundances, showing significant differences between tumor and adjacent tissues.

At the genus level, differences were observed in *Lactobacillus*, *Helicobacter*, *Enterococcus*, *Prevotella*, *Bacteroides*, *Escherichia-Shigella*, *Turicibacter*, *Streptococcus*, and *Bifidobacterium*. Discriminant analyses using LEfSe identified 18 bacterial phylotypes that were significantly different between GC tumor and normal mucosa microbiota (Figures 2C, D). The tumor microbiomes, present in over 90% of the patients, () exhibited increased abundances of several taxa, including *Lactobacillus*, *Prevotella*, *Clostridiales*, *Oscillospira*, *Turicibacter*, *Fusobacterium*, *Corynebacterium*, *Leptotrichia*, *Stenotrophomonas*, *Allobaculum*, *Treponema* and the family S24-7 (Figure 2E). Of note, taxa enriched in the normal mucosa microbiomes included *Helicobacter pylori* and *Streptococcus* genera, consistent with previous reports (19). A heatmap depicting the most abundant genera identified in GC mucosa microbiota

showed correlations between the mucosal microbiome and the abundance of selected genera (Supplementary Figure S2).

Significant centralities of oral bacteria in GC mucosae ecological network

SparCC algorithm-generated correlation-based microbial interaction networks identified co-occurrence and co-excluding interactions, highlighting the roles of oral bacteria such as *Fusobacterium*, *Porphyromonas*, *Prevotella*, *Leptotrichia*, *Aggregatibacter*, *Oribacterium*, *Parvimonas*, *Atopobium*, *Treponema*, and *Selenomonas* in both tumor and adjacent normal mucosae networks (Figures 3A, B; Supplementary Figures S4A, B).

To identify potential biomarkers, we focused on operational taxonomic units (OTUs) with significant weighted node connectivity (WNC) scores. This analysis identified *Prevotella*, *Porphyromonas*, *Fusobacterium*, *Aggregatibacter*, *Parvimonas*, *Oribacterium*, *Leptotrichia*, *Catonella*, *Atopobium*, *Allobaculum*, *Oscillospira*, *Lachnoanaerobaculum*, and *Selenomonas* as significant in the tumor mucosea network (Figure 3B). Furthermore, it highlighted differential microbial enrichment patterns between normal and tumor mucosae (Supplementary Figure S4B). OTU identification in the normal mucosea network included *Prevotella pallens*, *Eubacterium biforme* and *Helicobacter pylori*, which formed the backbone of the normal mucosae-specific network and likely exerted significant influence on normal microbial ecology (Supplementary Figure S4A). Given this, our

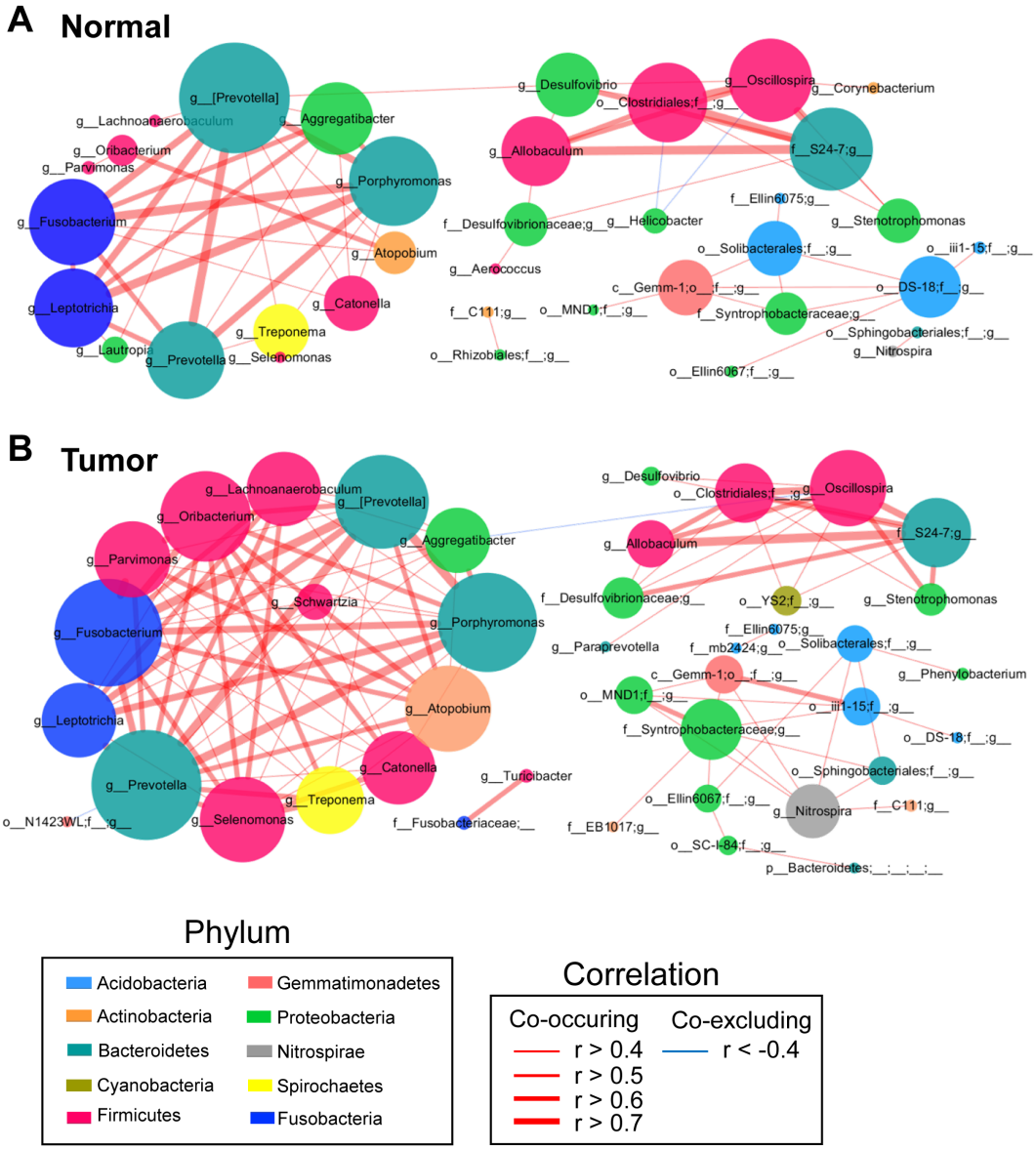


FIGURE 3
Co-occurrence network analysis of gastric bacterial genera with correlation coefficient >0.4 or < −0.4 in normal mucosae **(A)**, matched tumor mucosae **(B)**. SparCC algorithm was used for correlation coefficient calculation, and Cytoscape version 3.7.0 facilitated network construction. Red and blue lines represent positive and negative correlations, respectively.

results could suggest potential cooperation interactions among these species in the microenvironment of GC-associated gastric mucosae.

Mucosal microbiota richness and diversity is significantly higher in IATs^{high} group in GC tissues

Given the intrinsic gene expression signature closely linked to stromal activation and immune activation processes, we aimed to determine whether the IATs could accurately predict outcomes. This study focused on mainly IATs, namely CXCL9, CXCL10, GZMA, GZMB, PRF1, CD8A, IFNG, TBX2, and TNF. Quantitative PCR was performed to analyze the mRNA expression of these IATs in tumor and adjacent normal tissues from 85 GC patients. The levels of CXCL9, CXCL10, GZMA, GZMB, PRF1, CD8A, IFNG and TNF were significantly higher in tumor tissues compared to adjacent normal tissues whereas the TBX2 level was significantly lower in tumor tissues (Figure 4A). Relevance analyses revealed that several IATs were significantly associated with each other, indicating that different IATs might be co-regulated during tumor progression (Figure 4B). Moreover, Kaplan-Meier survival curves were plotted to investigate associations with survival. Positive correlations were observed between the expression levels of CXCL9, CXCL10, GZMA, GZMB, PRF1, CD8A, TNF and IATsand overall survival (OS) in tumor tissues (Figure 4C P = 0.0021; P = 0.0264; P = 0.0132; P = 0.0185; P = 0.026; P = 0.002; P = 0.0182; P = 0.0245; P < 0.0001). However, no significant correlation was found for IFNG in tumors or for these chemokines and cytokines in adjacent normal tissues (Figure 4C, data not showed). Multivariate Cox proportional hazards analysis was performed, and variables that were associated with survival by univariate analysis were adopted as covariates. In multivariate analysis, the expression level of TNF and IATs in tumor could emerge as an independent prognostic factor of either OS (HR, 0. 217; 95%CI, 0. 088-0. 536; P = 0.001; Table 2) or DFS (HR, 0.254; 95%CI, 0.105-0.618; P = 0.003; Table 2). These results suggested that IATs were significantly associated with GC progression and could serve as a powerful predictor of GC patient disease-free survival.

To identify mucosal microbiota signatures associated with IATs expressing patterns, we grouped patients into IATs ^{high} group/IATs ^{low} group. Stratification revealed that the high IATs group had greater mucosal microbiota richness and diversity, as indicated by alpha diversity measures (Figure 5A). Discriminant analyses using LEfSe identified 46 bacterial phylotypes significantly different between IATs^{high} group and IATs^{low} group (Figures 5B, C), with specific enrichment of *Akkermansia_muciniphila*, *Lactobacillus_intestinalis*, *Bacteroides_coprocola*, *MBNT15*, *uncultured_prokaryote*, and other bacteria in the IATs^{high} group. We further uncovered an enrichment of specific taxa, including *Proteobacteria*, *Bacteroides_stercoris*, *uncultured_gamma*, *Gammaproteobacteria*, *Oceanospirillales*, *Alcanivorax*, and *Alcanivoracaceae* within IATs^{low} group. This suggests a potential link between mucosal microbiota composition and IATs expression, with implications for targeted therapeutic interventions.

The association between discriminative species and distinctive selected IATs in GC mucosal tissues

To evaluate the effects of mucosal bacterial species on antitumor immune cell infiltration and the antitumor immune response, we used Spearman's correlation analysis to assess the association between 1483 discriminative species and four distinctive IATs (CXCL9, CXCL10, GZMB and IFNG) in the GC tissues. This analysis showed that the expression of these chemokines and cytokines was significantly correlated with the abundance of several OTUs (Figures 6A–D). Notably, all corelated OTUs were positively corelated with these four chemokines. Specifically, for the chemokine CXCL9, the mucosal bacterial species showing significant positive correlations included *Burkholderiales_bacterium*, *Desulfomicrobium_orale*, *Prevotella_genomosp*, *Treponema_vincentii*, *Verrucomicrobia_bacterium*, *Novosphingobium_rosa*, *Odoribacter_splanchnicus*, *Pyramidobacter_piscolens*, *Sulfuricaulis_limicola*, *bacterium_enrichment*, *Chlorobi_bacterium*, *Treponema_porcinum*, *Lactobacillus_mucosae*, and *Ileibacterium_valen*. For the chemokine CXCL10, the significantly positively correlated mucosal bacterial species encompassed *Firmicutes_bacterium*, *Lactobacillus_intestinalis*, *Spirochaeta_sp*, *Mesomycoplasma_moatsii*, *Nitrospira_japonica*, *Clostridium:spiroforme*, *Bacteroides_stercoris*, *Helicobacter_rodentium*, *Slackia_exigua*, *Prevotella_oris*, *bacterium_ROME 215Asa*, *Lactobacillus_aviarius*, *Streptococcus_anginosus*, and *Dialister_pneumosintes*. Regarding the cytokine GZMB, the significantly positively correlated mucosal bacterial species included *Alloprevotella_tannerae*, *Bacteroides_plebeius*, *Treponema_socranskii*, *Prevotella_salivae*, *Akkermansia_muciniphila*, *Bacteroides_coprocola*, *Prevotella_stercorea*, *Acidobacteria_bacterium*, *Treponema_medium*, *Prevotella_melaninogenica*, *Prevotella_nanceiensis*, *Prevotella_pallens*, *Prevotella_histicola*, *Prevotella_baroniae*, *Alloprevotella_rava*, *Actinomyces_graevenitzii*, *Helicobacter_typhlonius*, *Mucispirillum_schaedleri*, *Clostridiales_bacterium*, *Prevotella_shahii*, *Capnocytophaga_granulosa*, and *P*revotella_jejuni. Furthermore, with the cytokine IFNG, the significantly positively correlated mucosal bacterial species comprised *Nakamurella_multipartita*, *Lachnospiraceae*, *Lactobacillus_ingluviei*, *Campylobacter_canadensis*, and *bacterium_Ellin6543*. Finally, the bacterium *Akkermansia_muciniphila* may play a role in GZMB regulation, potentially influencing the tumor immune microenvironment. Its association with specific chemokine expression suggests its potential involvement in shaping the immune response within the tumor microenvironment, highlighting its significance in modulating the tumor's immune landscape.

10.3389/fimmu.2024.1435334

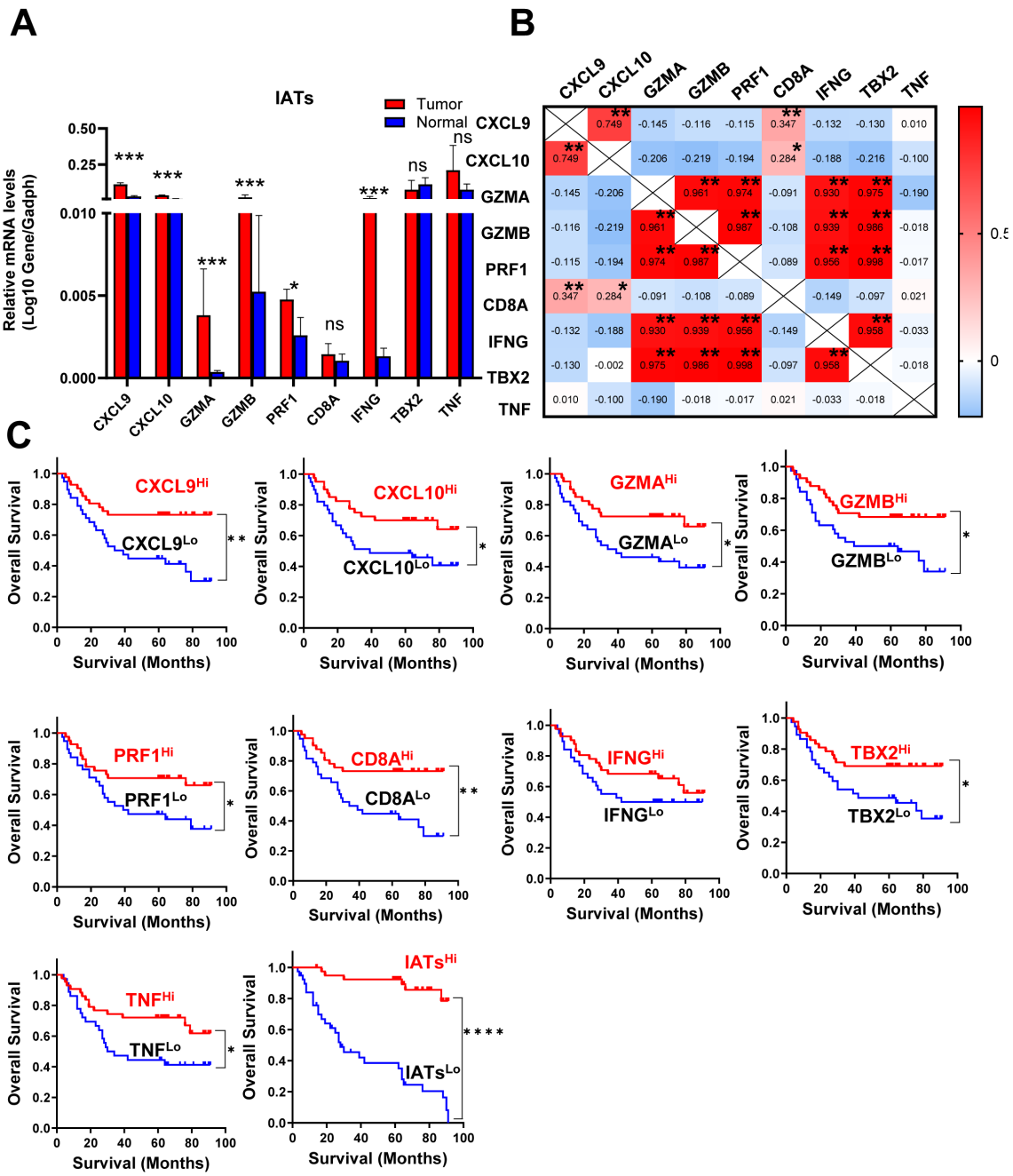


FIGURE 4
IATs were selectively regulated in tumor and the prognostic significance of IATs in GC patients. **(A)** Quantitative real-time polymeras chain reaction (qRT-PCR) was employed to detect the mRNA expression of each IATs in tumor and adjacent normal tissues (n=85). **(B)** Values denote the Pearson correlation coefficients; values closer to 1 indicate a better correlation. *P<0.05; **P<0.001; ***P<0.0001; ****P<0.00001; ns, no significant difference. **(C)** Cumulative OS times were calculated by the Kaplan-Meier method and analyzed by the log-rank test. The patients were divided into two groups according to the median value of CXCL9, CXCL10, GZMA, GZMB, PRF1, CD8A, IFNG, TBX2, TNF and IATs in tumor tissues.

Inferred functional changes in GC-associated gastric mucosal microbiota

The functional content of the gastric microbiota was predicted by PiCRUSt based on closed-reference OTU picking. In our present study, 25 Clusters of Orthologous Groups (COG) functional categories were tested, identifying 7 differentially abundant COGs with a QFDR < 0.05 between the GC tumor and normal mucosae microbiota (Figure 7A). These 7 COG categories, including cell motility, cell wall/membrane/envelope biogenesis, intracellular trafficking, secretion, and vesicular transport, extracellular structures, coenzyme transport and metabolism, transcription and general function prediction only, exhibited the most significant differences between the GC tumor and normal mucosae microbiota. Among these differential COGs, extracellular structures, transcription and general function prediction only were

10.3389/fimmu.2024.1435334

TABLE 2 Univariate and Multivariate Analyses of Factors Associated with Survival and Recurrence.

Variables	OS				DFS			
	Univariate *P*	Multivariate			Univariate *P*	Multivariate		
		HR	95% CI	*P*		HR	95% CI	*P*
Age, years (>64/≤64)	.241			NA	.348			NA
Gender (female/male)	.995			NA	. 230			NA
Tumor stage (pT4/pTis+pT1 +pT2+pT3)	**.0019**			NA	**.0034**			NA
Nodal status (pN1+pN2/pN0)	**.0017**	17.681	3.753-83.292	**.000**	**.0008**	5.790	2.158-15.534	**<.0001**
Distant metastases (Pos/Neg)	**.0019**	3.429	1.118-10.516	**.031**	**.0005**	3.415	1.410-8.269	**.006**
Differentiation(H+M/L)	.199			NA	.198			NA
$CXCL9_{high}/CXCL9_{low}$ tumor	**0.0021**			NA	**0.004**			NA
$CXCL10_{high}/CXCL10_{low}$ tumor	**0.0264**			NA	.123			NA
$GZMA_{high}/GZMA_{low}$ tumor	**0.0132**			NA	.106			NA
$GZMB_{high}/GZMB_{low}$ tumor	**0.0185**			NA	**.0034**			NA
$PRF1_{high}/PRF1_{low}$ tumor	**0.026**		.	NA	**.0017**			NA
$CD8A_{high}/CD8A_{low}$ tumor	**0.002**			NA	**< 0.0001**			NA
$IFNG_{high}/IFNG_{low}$ tumor	0.243			NA	.523			NA
$TBX2_{high}/TBX2_{low}$ tumor	**0.0182**			NA	.092			NA
TNF_{high}/TNF_{low} tumor	**0.0245**	.217	.088-.536	**.001**	.089			NA
$IATs_{high}/IATs_{low}$ tumor	**< 0.0001**			NA	**< 0.0001**	.254	.105-.618	**.003**

Cox proportional hazards regression model; Variables associated with survival by univariate analysis were adopted as covariates in multivariate analyses.
OS, overall survival; DFS, disease-free survival; HR, hazard ratio; CI, confidence interval. NA, not applicable. Pos, positive. Neg, negative. The bold values indicate that the P < 0.05.

significantly enriched in the tumoral microbiota. Additionally, we compared 40 Kyoto Encyclopedia of Genes and Genome (KEGG) pathways at level 2. At an FDR of 0.05, we identified 17 differentially abundant pathways between the GC tumor and normal mucosae microbiota (Figure 7B; Supplementary Figure S5). Consistent with the significant alterations in IATs-associated gastric microbiota, the KEGG pathways were changed between $IATs^{high}$ group and $IATs^{low}$ group in gastric mucosal tissues (Supplementary Figure S6). Together, these functional changes in the gastric microbiota may contribute to cytotoxic T cells infiltration and functional regulation.

Discussion

In this study, our analysis of 16S rRNA gene sequencing data reveals a noteworthy increase in richness and evenness within GC tumor tissues compared to their non-tumor counterparts. The elevated alpha diversity metrics emphasize heightened microbial diversity in tumor tissues, aligning with findings from a previous study (24). This may be due to the decreased diversity caused by the dominance of *Helicobacter pylori* in non-tumor counterparts. The other way, diverse microbial communities in tumor tissues may include bacteria that produce metabolites capable of influencing tumor growth. Moreover, diverse microbial populations could influence the local immune microenvironment by modulating the immune-subsets which could support tumor growth by creating an immunosuppressive environment. However, Liu et al. observed decreased diversity and richness in peritumoral and tumoral tissues compared to non-tumor tissues (6). The inconsistent findings underscore the lack of consensus regarding the relationship between microbial diversity and gastric mucosal tissues.

The dominant phyla in the gastric microbiota include Firmicutes, Proteobacteria, Bacteroidota, Campilobacterota, Actinobacteria, Fusobacteriota, Verrucomicrobiota, Acidobacteriota, and Cyanobacteria. Remarkably, the relative abundance of Helicobacter is reduced in GC tumor tissues compared to non-tumor tissues, consistent with previous studies (20, 24). This decrease may be attributed to the loss of specialized glandular tissues and decreased acid secretion (25). Discriminant analyses reveal significant shifts in microbial taxa between tumor and non-tumor mucosae, exemplified by an increased Proteobacteria/Campylobacterales ratio in tumor mucosae, indicating altered microbial profiles associated with GC. Additionally, network analysis highlights the central role of oral bacteria (*Fusobacterium*, *Porphyromonas*, *Prevotella*, etc.) in both tumor and normal mucosae networks, emphasizing their profound influence on GC microbial ecology. Notably, the top 10 genera of the gastric microbiota, including *Lactobacillus* and *Streptococcus*, are identified. *Lactobacillus* may produce metabolites serving as an

10.3389/fimmu.2024.1435334

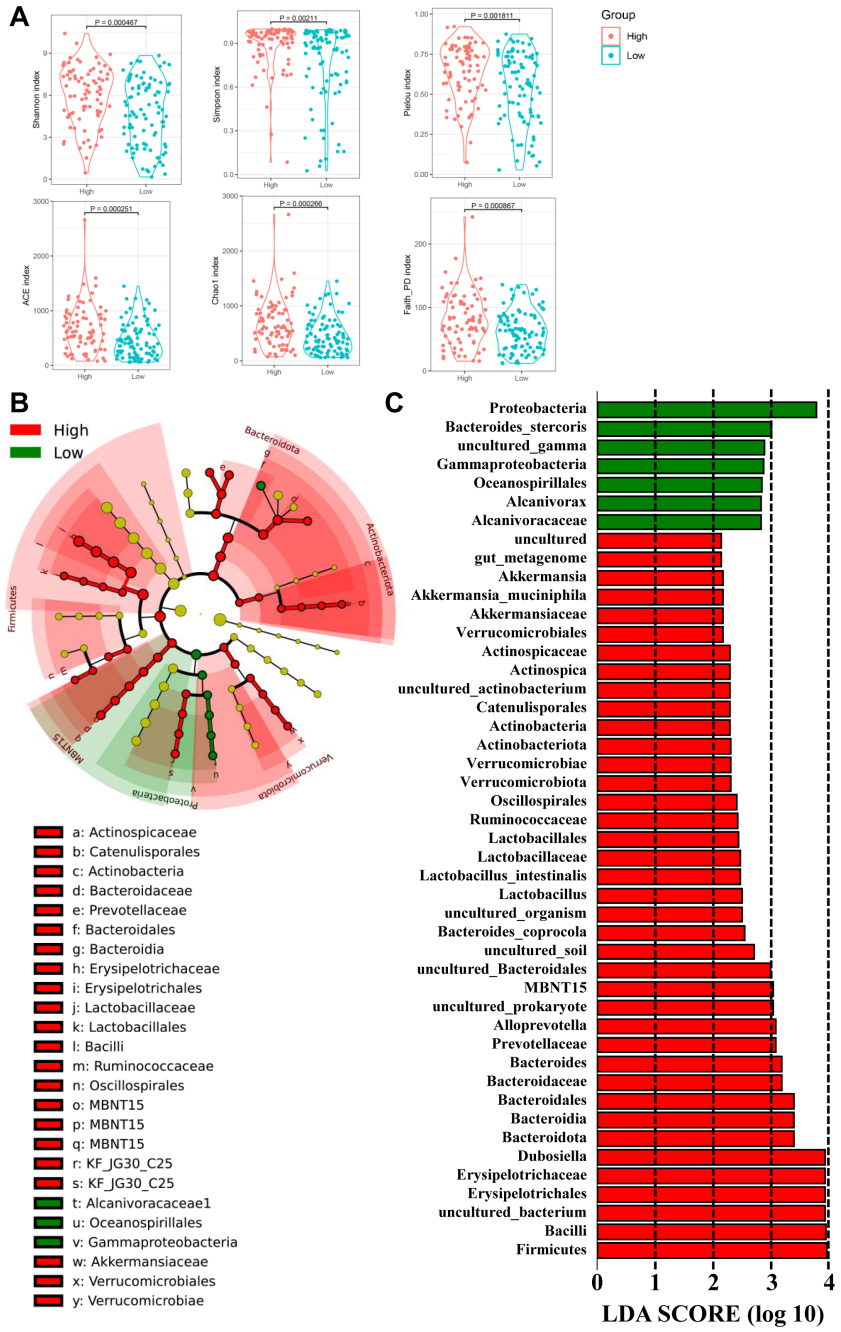


FIGURE 5
The diversity and richness of the gastric microbiota, and different bacterial taxa between the IATs^{high} group and IATs^{low} group. **(A)** The α-diversity indices (Shannon, Simpson, Pielou, ACE, Choa, and Faith_PD) evaluate the overall structure between the two groups, and **(B, C)** LEfSe identifies taxa with significant differences in abundance, shown if exceeding an LDA threshold value of >2.

energy source for tumor growth and angiogenesis (26), while the abundance of *Streptococcus* is increased in GC tumor tissues (24, 27).

Our assessment of IATs in GC tumor tissues and matched non-tumor tissues revealed a correlation between heightened chemokine expression and increased microbiota richness and evenness in the IATs^{high} group. Consistent with previous studies (10, 28, 29), the majority of IATs are upregulated in tumor tissues, mirroring the trend of mucosal microbiota diversity. Microbial enrichments within IATs high expression tissues suggest potential therapeutic interventions targeting the tumor microbiota for improved clinical outcomes. Moreover, Spearman's correlation analysis reveals significant associations between discriminative species and distinctive chemokines in GC tissues. Notably, the positive correlation of *Akkermansia_muciniphila* with specific chemokine expression underscores its potential role in modulating the GC tumor immune microenvironment. Recently studies have shown significant improvement in the pathological complete response rate among GC patients participating a randomized trial using

10.3389/fimmu.2024.1435334

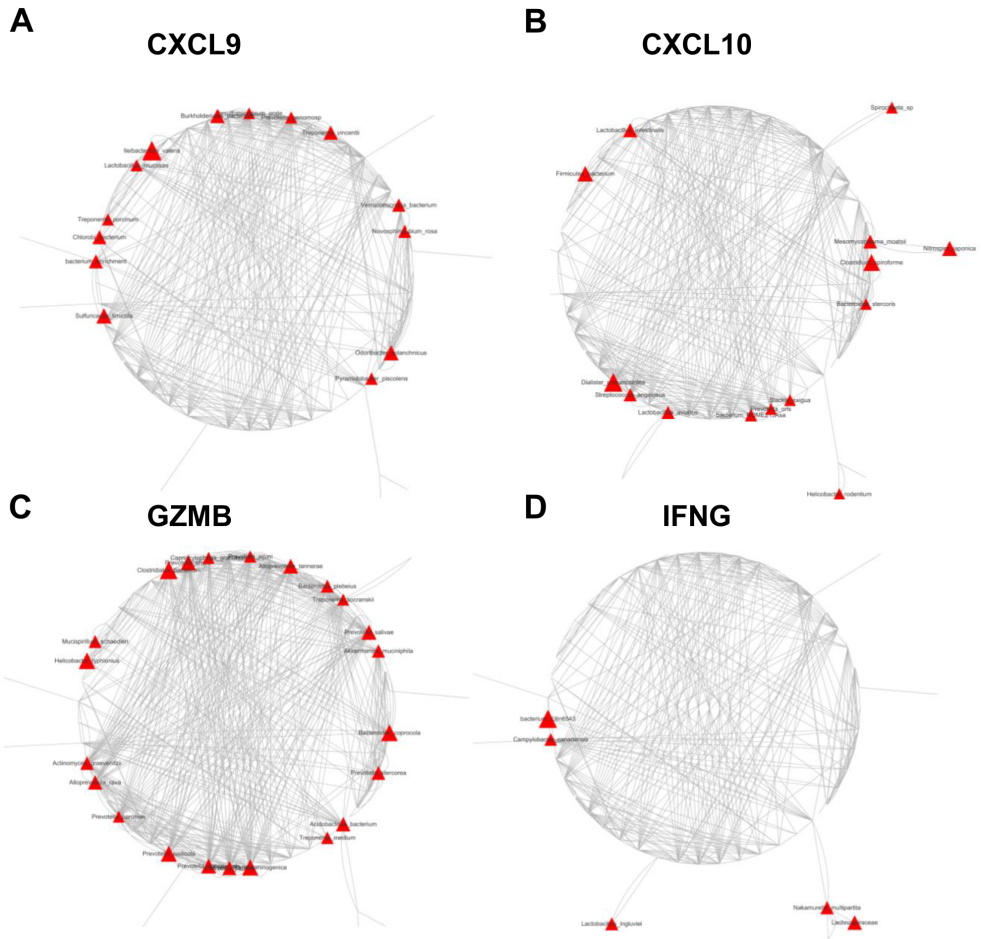


FIGURE 6
Network plots of operational taxonomic units (OTUs) based on mucosal abundance associated with each cytokine in the mucosal tissues of 85 GC patients. Nodes represent bacterial OTUs, and their abundance is significantly correlated with the expression of CXCL9 **(A)**, CXCL10 **(B)**, GZMB **(C)**, or IFNG **(D)**. The size of each node corresponds to the correlation coefficient.

perioperative PD-1 inhibitor plus chemotherapy for locally advanced GC (30). Studies have also demonstrated that intestinal microbiota composition significantly influences the effectiveness of anticancer immunosurveillance, impacting the therapeutic activity of immune-checkpoint inhibitors (ICIs) and immunogenic chemotherapies (31–33). Consistently, fecal microbiota transplantation (FMT) of microbiota from therapy-responding patients restored sensitivity to immunotherapy in germ-free environments or in mice treated with antibiotics and made insensitive to immunotherapy (34). Importantly, recent studies have further shown that response to ICIs therapy correlated with the percentage of *Akkermansia_muciniphila* in the intestinal microbiome of patients (35).

COG functional category tests showed that extracellular structures, transcription and general function prediction only were significantly enriched in the tumoral microbiota. Besides, KEGG enrichment analysis showed that pathways related to Amino acid metabolism including Isoflavonoid biosynthesis, Retinol metabolism, Lipoic acid metabolism, Adipocytokine signaling pathway, Fatty acid metabolism, and PPAR signaling pathway had significantly increased relative abundance in the tumoral microbiota. Liu et al. also reported that amino acid transport and metabolism exhibited the most significant differences among GC microhabitats (6). Studies have also shown that the PPAR signaling pathway is a crucial regulator in autocrine and paracrine signaling in the tumor microenvironment, modulating cancer-associated fibroblasts and tumor-associated macrophages/immune cells (36). Our observation of PPAR signaling pathway enrichment in GC tumor mucosal suggests pathway activation by the GC microbiome. Interestingly, we observed the enrichments of superpathway of L-lysine, L-threonine and L-methionine biosynthesis I, superpathway of arginine and polyamine biosynthesis, anhydromuropeptides recycling, superpathway of polyamine biosynthesis I, superpathway of L-methionine biosynthesis, reductive TCA cycle I, superpathway of S-adenosyl-L-methionine biosynthesis and tRNA processing in the IATs^{high} group mucosae microbiome. Increases in these pathways are predictive of bacterial involvement in amino acid metabolism by the gut microbiome, which has been linked to hyperproliferation of cells in the tumor microenvironment (37, 38). The enrichment of these pathway in IATs^{high} group mucosae microbiome highlights their potential contribution to the immune response. Furthermore, pathways involved in polyamine biosynthesis have been reported to

10.3389/fimmu.2024.1435334

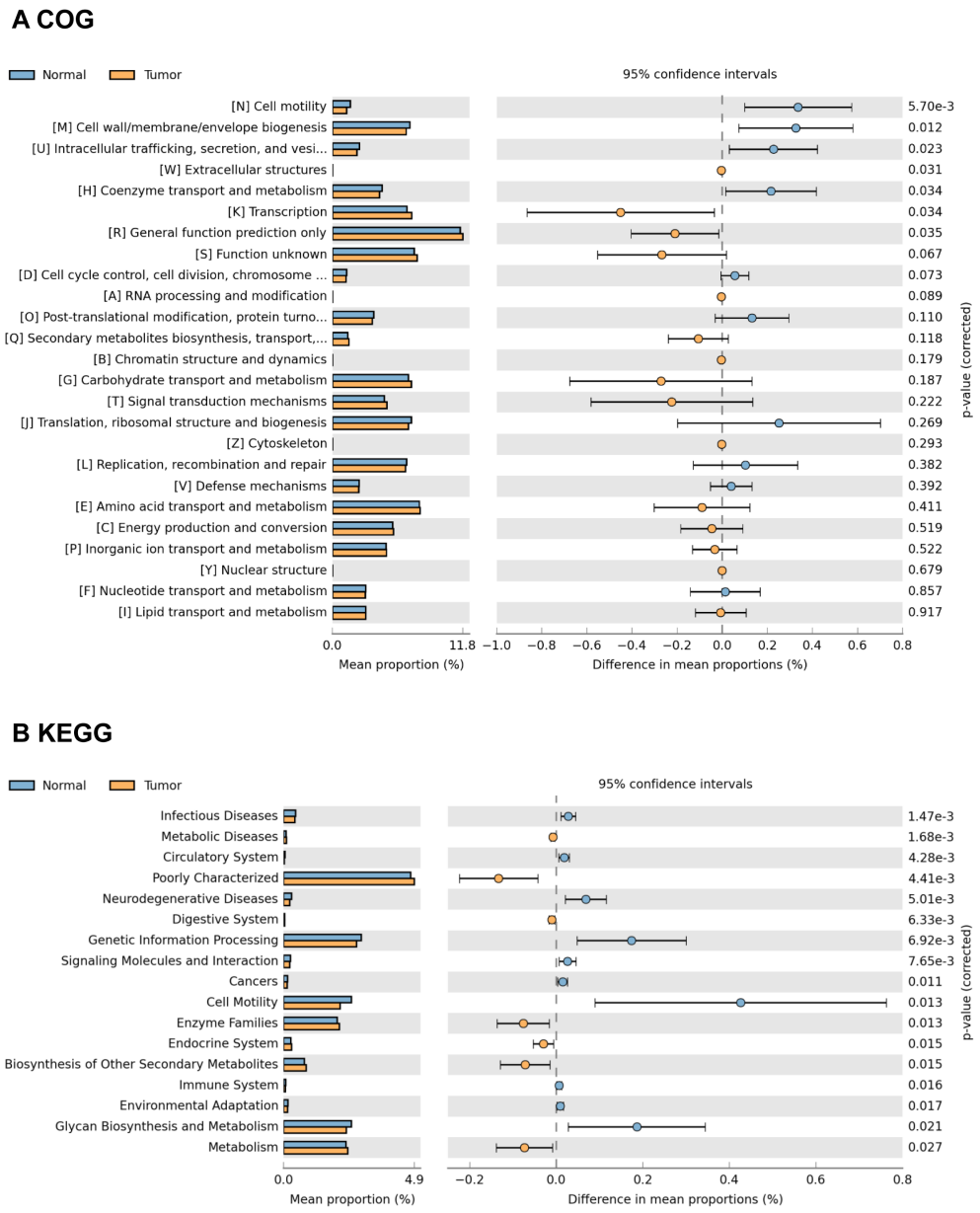


FIGURE 7
Representative COG functional category and KEGG functional pathways enriched in the tumor and adjacent normal mucosae microbiome. PiCRUSt-based analysis using Welch's t-test reveals differences between the two groups, and multiple testing correction by the Benjamini-Hochberg method is applied based on the false discovery rate (FDR) by STAMP. Comparisons for each COG functional category **(A)** and KEGG functional pathway **(B)** are shown as percentages.

remodel the tumor immune microenvironment by altering the activation and proliferation of $CD4^+$ and $CD8^+$ T lymphocytes (39, 40). The association of polyamine biosynthesis with IATs in this study supports the role that this pathway may play in $CD8^+$ T lymphocytes remodeling and supports previous observations in GC (41). Further investigations into the implications of microbiome functional dysbiosis in $IATs^{high}$ and $IATs^{low}$ groups are needed for a deeper understanding of the gastric immune microenvironment.

Our study had several limitations. First, the sample size is relatively small, resulting in the lack of significant correlation between clinical features and microbiome, and between clinical features and metabolome. Second, we did not perform longitudinal studies since we could not obtain serial tissue samples from the recruited patients. Third, we did not include gastric cancer patients from different regions, so our patient heterogeneity is insufficient. Fourth, the diet could heavily influence both the gastric microbiota and metabolites, but we could not obtain the diet information of patients to analyze the effect of diet on gastric microbiome and metabolome.

In conclusion, our study provides insights into the microbiome of GC tumor tissues and matched non-tumor tissues, unveiling IATs-associated bacteria and highlighting the pivotal role of mucosal microbiota alterations in GC. The identification of potential biomarkers and therapeutic targets, such as IATs-

associated bacteria, offers prospects for improving clinical outcomes in GC. Further research is warranted to delve into the functional implications of microbiome dysbiosis in IATs^{high} and IATs^{low} group mucosae, advancing our understanding of the gastric immune microenvironment.

Materials and methods

Patients and database

A total of 85 individuals, scheduled for primary tumor resection at the Affiliated Hospital of Jiangnan University between 2016 and 2019, were enrolled in the study. Exclusion criteria included prior chemo-radiotherapy. Patients received no antibiotics within a month before surgery but were administered intravenous antibiotics shortly before resection. Post-surgery, 85 paired fresh tissues, including gastric tumor and matched non-tumor tissues, were collected. Biopsies were snap-frozen in cryovial immediately with liquid nitrogen and then stored at -80°C until DNA extraction. Histopathological and clinical findings were scored according to the International Union Against Cancer (UICC)-TNM staging system.

DNA extraction and16S rRNA gene sequencing

A total of 170 tissue samples (one tumor and one adjacent normal sample per individual) were processed for DNA purification. The DNA extraction was carried out according to the AllPrep DNA/RNA extraction kit and total RNA were extracted using the Ultrapure RNA Kit (CWBIO, China). Total DNA was purified from tumor and paired normal adjacent mucosal tissue samples. Mucosa-associated microbiota was analyzed through 16S rRNA sequencing. 16S rRNA gene amplicon sequencing was carried out employing the 16S Meta-genomic Sequencing Library Preparation protocol developed by Illumina (San Diego, California, USA. Briefly, 200 ng of mucosal DNA was amplified from each sample using the primers 515F (5′ GTGCCAGCMGCCGCGGTAA 3′) with Titanium Adaptor B and 806R (5′ GGACTAC HVGGGTWTCTAAT 3′) with Titanium Adaptor A and a sample-specific barcode sequence consisting of twelve nucleotides targeting the V4 hypervariable region of the 16S rRNA gene using FastStart Taq DNA Polymerase (Roche). The resulting sequences were processed for bioinformatics analysis.

RNA isolation, chemokine mRNA expression and quantitative PCR

Total RNA from GC tumor and paired normal adjacent tissues was isolated using Trizol (Invitrogen, USA). cDNA was synthesized using Superscript III Reverse Transcriptase (Promega, USA). Real-time PCR reactions were conducted with SYBR Green (TaKaRa, Japan) and analyzed on the Step One Plus Real-time PCR System (Applied Biosystems, USA) with the following conditions: 95°C for 5 min, 95°C for 5 s, 60°C for 30 s, for 40 cycles. The relative mRNA expression value was calculated by 2 -△△T method. GAPDH was utilized as the internal control. The primers used were as follows: CXCL9, forward primer F(5′ AAGCAGCCAAGTCGGTTAGT 3′) and reverse primer R(5′ CAGCAGTGTGAGCAGTGATTC 3′); CXCL10, forward primer F(5′ AGCAGAGGAACCTCCAGTCT 3′) and reverse primer R(5′ AGGTACTCCTTGAATGCCACT 3′); GZMA, forward primer F(5′ GAAGAGACTCGTGCAAT GGAGA 3′) and reverse primer R(5′ AAGGCCAAAGGA AGTGACCC 3′); GZMB, F(5′ CCAGGGCAGATGCAGACTTT 3′) and reverse primer R(5′ CTCGTATCAGGAAGCCACCG 3′); PRF1, F(5′ GGGGCTGATGCCACCATT 3′) and reverse primer R (5′ GGCACTTGGGCTCTGGAAT 3′); CD8A, F(5′ CGGTTTCCT GGGGTAACAGT 3′) and reverse primer R(5′ TGCCTGAATCAG CCTTTCTGT 3′); IFNG, F(5′ GAGTGTGGAGACCATCAAGGA 3′) and reverse primer R(5′ TGGACATTCAAGTCAGTTA CCGAA 3′); TBX2, F(5′ TACGAGGAGCACTGCAAACC 3′) and reverse primer R(5′ CACGACTTCTCCTCAGCTCG 3′); TNF, F(5′ AGCCCATGTTGTAGCAAACC 3′) and reverse primer R(5′ ATGAGGTACAGGCCCTCTGA 3′); GAPDH, forward primer F(5′ TGACTTCAACAGCGACACCCA 3′) and reverse primer R(5′ CACCCTGTTGCTGTAGCCAAA 3′). Experiments were performed in triplicate.

Bioinformatics analysis

Microbiome bioinformatics were performed using QIIME 2 (2023.9) with slight modification according to the official tutorials. Briefly, raw sequence data were demultiplexed using the demux plugin following by primers cutting with cutadapt plugin (Martin, M., 2011). Sequences were then quality filtered, denoised, merged and chimera removed using the DADA2 plugin. Species annotation was performed using QIIME2 software. The annotation database is Silva Database. Alpha and beta diversity analyses were calculated with QIIME2 and displayed with R software (Version 3.6.2). Principal Coordinate Analysis (PCoA) was carried out to obtain principal coordinates and visualize differences of samples in complex multi-dimensional data. A matrix of weighted or unweighted unifrac distances among samples obtained previously was transformed into a new set of orthogonal axes, where the maximum variation factor was demonstrated by the first principal coordinate, and the second maximum variation factor was demonstrated by the second principal coordinate, and so on. The three-dimensional PCoA results were displayed using QIIME2 package, while the two-dimensional PCoA results were displayed using ade package and ggplot2 package in R software (Version 3.6.2).

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: NCBI SRA database under accession number PRJNA1032279.

Ethics statement

The Institutional Review Board of Affiliated Hospital of Jiangnan University reviewed and approved the study protocol and exempted the study from the obligation to obtain informed consent (the committee's reference number: JNU20210618IRB08). The studies were conducted in accordance with the local legislation and institutional requirements. The human samples used in this study were acquired from primarily isolated as part of your previous study for which ethical approval was obtained. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements. The manuscript presents research on animals that do not require ethical approval for their study.

Author contributions

CQ: Investigation, Methodology, Writing – original draft, Conceptualization, Data curation, Validation, Visualization. JH: Investigation, Methodology, Resources, Writing – review & editing. ZP: Data curation, Methodology, Software, Validation, Visualization, Writing – review & editing. XS: Data curation, Formal analysis, Methodology, Software, Writing – review & editing. JZ: Funding acquisition, Investigation, Methodology, Resources, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The Scientific Research Project funded by Shanghai Municipal Science and Technology Commission (No. 22ZR1449000) and Youth Foundation of Jiangsu Natural Science Foundation (No. BK20180617).

Acknowledgments

We thank Dr. Dongli Liang from Instrumental Analysis Center, Shanghai Jiao Tong University for recommendations on data analysis and helpful discussions and comments on 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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1435334/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) 71:209–49. doi: 10.3322/caac.21660

2. Xia C, Dong X, Li H, Cao M, Sun D, He S, et al. Cancer statistics in China and United States, 2022: profiles, trends, and determinants. *Chin Med J (Engl)*. (2022) 135:584–90. doi: 10.1097/CM9.0000000000002108

3. Peek RM, Crabtree JE. Helicobacter infection and gastric neoplasia. *J Pathol*. (2006) 208:233–48. doi: 10.1002/path.1868

4. Eun CS, Kim BK, Han DS, Kim SY, Kim KM, Choi BY, et al. Differences in gastric mucosal microbiota profiling in patients with chronic gastritis, intestinal metaplasia, and gastric cancer using pyrosequencing methods. *Helicobacter*. (2014) 19:407–16. doi: 10.1111/hel.12145

5. Galon J, Pagès F, Marincola FM, Thurin M, Trinchieri G, Fox BA, et al. The immune score as a new possible approach for the classification of cancer. *J Transl Med*. (2012) 10:1. doi: 10.1186/1479-5876-10-1

6. Angell H, Galon J. From the immune contexture to the immunoscore: the role of prognostic and predictive immune markers in cancer. *Curr Opin Immunol*. (2013) 25:261–7. doi: 10.1016/j.coi.2013.03.004

7. Galon J. From the immune contexture to the immunoscore in cancer. *Eur J Cancer*. (2014) 50:S8. doi: 10.1016/S0959-8049(14)50030-5

8. Lee K, Hwang H, Nam KT. Immune response and the tumor microenvironment: How they communicate to regulate gastric cancer. *Gut Liver*. (2014) 8:131–9. doi: 10.5009/gnl.2014.8.2.131

9. Zeng D, Zhou R, Yu Y, Luo Y, Zhang J, Sun H, et al. Gene expression profiles for a prognostic immunoscore in gastric cancer. *Br J Surg*. (2018) 105:1338–48. doi: 10.1002/bjs.10871

10. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, et al. Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. *Cancer Immunol Res*. (2019) 7:737–50. doi: 10.1158/2326-6066.CIR-18-0436

11. Jiang Y, Zhang Q, Hu Y, Li T, Yu J, Zhao L, et al. ImmunoScore signature: A prognostic and predictive tool in gastric cancer. *Ann Surg*. (2018) 267:504. doi: 10.1097/SLA.0000000000002116

12. Zhang R, Tian L, Chen L-J, Xiao F, Hou J-M, Zhao X, et al. Combination of MIG (CXCL9) chemokine gene therapy with low-dose cisplatin improves therapeutic efficacy against murine carcinoma. *Gene Ther*. (2006) 13:1263–71. doi: 10.1038/sj.gt.3302756

13. Pan J, Burdick MD, Belperio JA, Xue YY, Gerard C, Sharma S, et al. CXCR3/CXCR3 ligand biological axis impairs RENCA tumor growth by a mechanism of immunoangiostasis. *J Immunol*. (2006) 176:1456–64. doi: 10.4049/jimmunol.176.3.1456

14. Wang X, Lu X-L, Zhao H-Y, Zhang F-C, Jiang X-B. A novel recombinant protein of IP10-EGFRvIIIscFv and CD8(+) cytotoxic T lymphocytes synergistically inhibits the growth of implanted glioma in mice. *Cancer Immunol Immunother*. (2013) 62:1261–72. doi: 10.1007/s00262-013-1426-6

15. Wang X, Lu X-L, Zhao H-Y, Zhang F-C, Jiang X-B. A novel recombinant protein of IP10-EGFRvIIIscFv and CD8(+) cytotoxic T lymphocytes synergistically inhibits the

growth of implanted glioma in mice. *Cancer Immunol Immunother*. (2013) 62:1261–72. doi: 10.1007/s00262-013-1426-6

16. Feldman AL, Friedl J, Lans TE, Libutti SK, Lorang D, Miller MS, et al. Retroviral gene transfer of interferon-inducible protein 10 inhibits growth of human melanoma xenografts. *Int J Cancer*. (2002) 99:149–53. doi: 10.1002/ijc.10292

17. Sun Y, Finger C, Alvarez-Vallina L, Cichutek K, Buchholz CJ. Chronic gene delivery of interferon-inducible protein 10 through replication-competent retrovirus vectors suppresses tumor growth. *Cancer Gene Ther*. (2005) 12:900–12. doi: 10.1038/sj.cgt.7700854

18. Peng W, Liu C, Xu C, Lou Y, Chen J, Yang Y, et al. PD-1 blockade enhances T-cell migration to tumors by elevating IFN-γ inducible chemokines. *Cancer Res*. (2012) 72:5209–18. doi: 10.1158/0008-5472.CAN-12-1187

19. Nadella V, Nagao K. A bacteria-chemokine double act repairs the skin. *Nat Immunol*. (2020) 21:966–7. doi: 10.1038/s41590-020-0755-9

20. Chen X-H, Wang A, Chu A-N, Gong Y-H, Yuan Y. Mucosa-associated microbiota in gastric cancer tissues compared with non-cancer tissues. *Front Microbiol*. (2019) 10:1261. doi: 10.3389/fmicb.2019.01261

21. Ferreira RM, Pereira-Marques J, Pinto-Ribeiro I, Costa JL, Carneiro F, Machado JC, et al. Gastric microbial community profiling reveals a dysbiotic cancer-associated microbiota. *Gut*. (2018) 67:226–36. doi: 10.1136/gutjnl-2017-314205

22. Malfertheiner P, Camargo MC, El-Omar E, Liou J-M, Peek R, Schulz C, et al. Helicobacter pylori infection. *Nat Rev Dis Primers*. (2023) 9:1–24. doi: 10.1038/s41572-023-00431-8

23. Ozbey G, Sproston E, Hanafiah A. Helicobacter pylori infection and gastric microbiota. *Euroasian J Hepatogastroenterol*. (2020) 10:36–41. doi: 10.5005/jp-journals-10018-1310

24. Dai D, Yang Y, Yu J, Dang T, Qin W, Teng L, et al. Interactions between gastric microbiota and metabolites in gastric cancer. *Cell Death Dis*. (2021) 12:1–11. doi: 10.1038/s41419-021-04396-y

25. Yao X, Smolka AJ. Gastric parietal cell physiology and helicobacter pylori-induced disease. *Gastroenterology*. (2019) 156:2158–73. doi: 10.1053/j.gastro.2019.02.036

26. Sonveaux P, Copetti T, De Saedeleer CJ, Végran F, Verrax J, Kennedy KM, et al. Targeting the lactate transporter MCT1 in endothelial cells inhibits lactate-induced HIF-1 activation and tumor angiogenesis. *PloS One*. (2012) 7:e33418. doi: 10.1371/journal.pone.0033418

27. Shao D, Vogtmann E, Liu A, Qin J, Chen W, Abnet CC, et al. Microbial characterization of esophageal squamous cell carcinoma and gastric cardia adenocarcinoma from a high-risk region of China. *Cancer*. (2019) 125:3993–4002. doi: 10.1002/cncr.32403

28. Nakajima S, Mimura K, Matsumoto T, Thar Min AK, Ito M, Nakano H, et al. The effects of T-DXd on the expression of HLA class i and chemokines CXCL9/10/11 in HER2-overexpressing gastric cancer cells. *Sci Rep*. (2021) 11:16891. doi: 10.1038/s41598-021-96521-2

29. Ou J, Lei P, Yang Z, Yang M, Luo L, Mo H, et al. LINC00152 mediates CD8+ T-cell infiltration in gastric cancer through binding to EZH2 and regulating the CXCL9, 10/CXCR3 axis. *J Mol Histol*. (2021) 52:611–20. doi: 10.1007/s10735-021-09967-z

30. Yuan S-Q, Nie R-C, Jin Y, Liang C-C, Li Y-F, Jian R, et al. Perioperative toripalimab and chemotherapy in locally advanced gastric or gastro-esophageal junction cancer: a randomized phase 2 trial. *Nat Med*. (2024) 30:552–9. doi: 10.1038/s41591-023-02721-w

31. Luo B, Zhang Y, Zhang C, Liu X, Shi C. Intestinal microbiota: A potential target for enhancing the antitumor efficacy and reducing the toxicity of immune checkpoint inhibitors. *Cancer Lett*. (2021) 509:53–62. doi: 10.1016/j.canlet.2021.04.001

32. Li X, Zhang S, Guo G, Han J, Yu J. Gut microbiome in modulating immune checkpoint inhibitors. *EBioMedicine*. (2022) 82:104163. doi: 10.1016/j.ebiom.2022.104163

33. Simpson RC, Shanahan ER, Scolyer RA, Long GV. Towards modulating the gut microbiota to enhance the efficacy of immune-checkpoint inhibitors. *Nat Rev Clin Oncol*. (2023) 20:697–715. doi: 10.1038/s41571-023-00803-9

34. Zhang J, Wu K, Shi C, Li G. Cancer immunotherapy: Fecal microbiota transplantation brings light. *Curr Treat Options Oncol*. (2022) 23:1777–92. doi: 10.1007/s11864-022-01027-2

35. Derosa L, Routy B, Thomas AM, Iebba V, Zalcman G, Friard S, et al. Intestinal akkermansia muciniphila predicts clinical response to PD-1 blockade in patients with advanced non-small-cell lung cancer. *Nat Med*. (2022) 28:315–24. doi: 10.1038/s41591-021-01655-5

36. Cheng HS, Yip YS, Lim EKY, Wahli W, Tan NS. PPARs and tumor microenvironment: The emerging roles of the metabolic master regulators in tumor stromal-epithelial crosstalk and carcinogenesis. *Cancers (Basel)*. (2021) 13:2153. doi: 10.3390/cancers13092153

37. Zhang E, Ding C, Li S, Aikemu B, Zhou X, Fan X, et al. Polyamine metabolism patterns characterized tumor microenvironment, prognosis, and response to immunotherapy in colorectal cancer. *Cancer Cell Int*. (2023) 23:96. doi: 10.1186/s12935-023-02892-z

38. Yang Q, Wang B, Zheng Q, Li H, Meng X , Zhou F, et al. A review of gut microbiota-derived metabolites in tumor progression and cancer therapy. *Adv Sci*. (2023) 10. doi: 10.1002/advs.202207366

39. Elmarsafawi AG, Hesterberg RS, Fernandez MR, Yang C, Darville LN, Liu M, et al. Modulating the polyamine/hypusine axis controls generation of CD8+ tissue-resident memory T Cells. *JCI Insight*. (2023) 8:e169308. doi: 10.1172/jci.insight.169308

40. Bengsch B, Johnson AL, Kurachi M, Odorizzi PM, Pauken KE, Attanasio J, et al. Bioenergetic insufficiencies due to metabolic alterations regulated by the inhibitory receptor PD-1 are an early driver of CD8(+) T Cell exhaustion. *Immunity*. (2016) 45:358–73. doi: 10.1016/j.immuni.2016.07.008

41. McNamara KM, Gobert AP, Wilson KT. The role of polyamines in gastric cancer. *Oncogene*. (2021) 40:4399–412. doi: 10.1038/s41388-021-01862-x

frontiers | Frontiers in Immunology

TYPE Mini Review
PUBLISHED 25 September 2024
DOI 10.3389/fimmu.2024.1471409

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Qianhe Ren,
Jiangsu University, China

*CORRESPONDENCE
Ke Xu
cqghxuke@cqu.edu.cn
Xiang Shen
shenxiangdoctor@163.com
Shangke Huang
huangshangke001@swmu.edu.cn

†These authors have contributed equally to this work and share first authorship

RECEIVED 27 July 2024
ACCEPTED 30 August 2024
PUBLISHED 25 September 2024

CITATION
Yan Y, Shen S, Li J, Su L, Wang B, Zhang J, Lu J, Luo H, Han P, Xu K, Shen X and Huang S (2024) Cross-omics strategies and personalised options for lung cancer immunotherapy.
Front. Immunol. 15:1471409.
doi: 10.3389/fimmu.2024.1471409

Cross-omics strategies and personalised options for lung cancer immunotherapy

Yalan Yan[1†], Siyi Shen[1†], Jiamin Li[1†], Lanqian Su[1], Binbin Wang[2], Jinghan Zhang[3], Jiaan Lu[1], Huiyan Luo[4], Ping Han[4], Ke Xu[4*], Xiang Shen[5*] and Shangke Huang[6*]

[1]School of Clinical Medicine, The Affiliated Hospital, Southwest Medical University, Luzhou, China, [2]Intensive Care Unit, Xichong People's Hospital, Nanchong, China, [3]Department of Anaesthesiology, Southwest Medical University, Luzhou, China, [4]Department of Oncology, Chongqing General Hospital, Chongqing University, Chongqing, China, [5]Department of Respiratory and Critical Care Medicine, The Affiliated Hospital, Southwest Medical University, Luzhou, China, [6]Department of Oncology, The Affiliated Hospital, Southwest Medical University, Luzhou, China

Lung cancer is one of the most common malignant tumours worldwide and its high mortality rate makes it a leading cause of cancer-related deaths. To address this daunting challenge, we need a comprehensive understanding of the pathogenesis and progression of lung cancer in order to adopt more effective therapeutic strategies. In this regard, integrating multi-omics data of the lung provides a highly promising avenue. Multi-omics approaches such as genomics, transcriptomics, proteomics, and metabolomics have become key tools in the study of lung cancer. The application of these methods not only helps to resolve the immunotherapeutic mechanisms of lung cancer, but also provides a theoretical basis for the development of personalised treatment plans. By integrating multi-omics, we have gained a more comprehensive understanding of the process of lung cancer development and progression, and discovered potential immunotherapy targets. This review summarises the studies on multi-omics and immunology in lung cancer, and explores the application of these studies in early diagnosis, treatment selection and prognostic assessment of lung cancer, with the aim of providing more personalised and effective treatment options for lung cancer patients.

KEYWORDS

lung cancer, immunotherapy, precision medicine, multi-omics, individualised therapy, immune checkpoints

1 Introduction

Lung cancer has been one of the most common malignant tumours globally over the past decades. Despite the widespread use of conventional treatments such as surgery, radiotherapy, chemotherapy and targeted drug therapy, the five-year survival rate for lung cancer is usually less than 20% (1). Additionally, at all stages, less than 7% of patients

survive more than ten years after diagnosis (2). In recent years, the emergence of immunotherapy has marked a revolution in cancer treatments, which not only has an acceptable safety profile, but also produces durable therapeutic responses in a wide range of patient populations (3). Nonetheless, lung cancer exhibits significant histological heterogeneity, diverse genomic profiles, and differential responses to therapy (4), and still poses significant challenges for immunotherapy and prevention.

With the rapid development of multi-omics technology, covering genomics, transcriptomics, proteomics and metabolomics, our understanding of lung cancer is deepening (5, 6). Multi-omics technology has constructed a progressive analysis framework from the genetic basis to the effect of environmental exposure (7), and has deeply analysed the pathogenesis, pathophysiological process and molecular basis of immunotherapy of lung cancer, which has provided a strong support for the scientific formulation of precise treatment strategies.

The aim of this review is to explore recent advances in multi-omics studies of lung cancer and their potential applications in early diagnosis, treatment selection and prognostic assessment. By integrating immunotherapy and multi-omics data in order to better understand the complex molecular network of lung cancer, it provides new ideas and methods for individualised treatment and precision medicine of lung cancer.

2 Lung cancer immunotherapy and genomics

Lung cancer, as a highly heterogeneous disease, has been profoundly influenced by molecular biology in its pathogenesis and therapeutic strategies. In non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC), unique molecular features of different histological subtypes have been revealed through the identification of specific genetic variants and epigenetic modifications, thus providing new directions for individualised treatment of lung cancer.

In NSCLC, histological subtypes frequently dominated by lung adenocarcinoma (LUAD) and squamous cell carcinoma are common (8). The complexity of NSCLC is reflected in its variable genetic variants. Common target gene driver mutations include genes such as epidermal growth factor receptor (EGFR), KRAS, MET, BRAF, ALK, ROS proto-oncogene 1 (ROS1) and RET (9) (Figure 1A). Through combined whole exome sequencing (WES) technology, it was found that common mutations in LUAD include tumour suppressor genes TP53 (46%), STK11 (17%), KEAP1 (17%), NF1 (11%), RB1 (4%) and CDKN2A (4%), as well as chromatin modification genes SETD2 (9%), ARID1A (7%), SMARCA4 (6%) and RNA splicing genes RBM10 (8%) and U2AF1 (3%) (10). Mutations in the genes FGFR1, NRF2, AKT1 and DDR2 are particularly prominent in lung squamous cell carcinoma (10).

For SCLC, deep sequencing of key oncogenes by advanced integrated mutational profiling (MSK-IMPACT) technology (11) revealed inactivating mutations or deletion of tumour suppressor genes such as TP53, RB1, KMT2D, PTEN, NOTCH1, CREBBP, FAT1, NF1 and APC, and inactivating mutations in PIK3A, EGFR and KRAS activating mutations in oncogenes (12). Unlike NSCLC, SCLC is often accompanied by the expression of MYC oncogenes, which contribute to rapid cell proliferation and lead to DNA replication stress (13). In addition, epigenetic modifications play a key role in lung carcinogenesis, and heritable chromatin modifications such as DNA methylation, histone modifications, and non-coding RNA regulation regulate gene expression without altering the DNA sequence (9, 10, 14–16) (Figure 1A). Epigenetic mutations and disruptions are strongly associated with multiple tumour types, providing new ideas for targeted lung cancer therapy based on molecular subtype differences.

Genomics plays an important role in the classification, treatment and prognostic assessment of lung cancer. Traditionally, lung cancer classification was based on histological patterns, whereas advances in genomics have allowed lung cancer to be characterised also by tumour biomarkers and genetic alterations. For example, Stephen J Murphy et al. defined a common origin or lineage of lung cancer by analysing genomic rearrangements and somatic DNA linkages, and used these specific DNA linkages as precise tumour markers to differentiate between primary and metastatic lung cancer (17).

Genome sequencing technology has revealed key genetic variants in lung cancer, facilitating the development of personalised treatment strategies. The study noted that in non-small cell lung cancer (NSCLC), aberrant activation of the PI3K-AKT-mTOR pathway is closely associated with resistance to EGFR tyrosine kinase inhibitors (EGFR-TKIs), and that its activation is mainly caused by PIK3CA, AKT1 mutations and PTEN deletion. This discovery led to the development of drugs targeting mTOR (e.g., everolimus and temsirolimus) and EGFR-TKIs targeting EGFR and ALK (e.g., ositinib, gefitinib, ceritinib, and loratinib), which have demonstrated clinical efficacy in the treatment of lung cancer (18) (Table 1).

Immune checkpoint inhibitors (ICIs) have become key agents in tumour immunotherapy, especially in the treatment of lung cancer. However, their potential to cause immune-related adverse effects makes the search for biomarkers that predict response to ICI therapy crucial (9). Investigators assessed early predictors of anti-PD-L1 therapy by analysing circulating tumour DNA (ctDNA), and their study showed that a reduction in the frequency of the variant allele was associated with tumour shrinkage after 6 weeks of treatment, providing a valuable non-invasive method for predicting the effectiveness of treatment (44).

3 Lung cancer immunotherapy and transcriptomics

In the field of tumour immunotherapy, targeting the tumour microenvironment (TME) for precision medicine is one of the latest research directions. In this process, immune cells play a key role (45). Immune cell interactions are regulated by transcription factors and further contribute to the immune response. Wu et al. identified interactions between cancer cells and endothelial cells, fibroblasts,

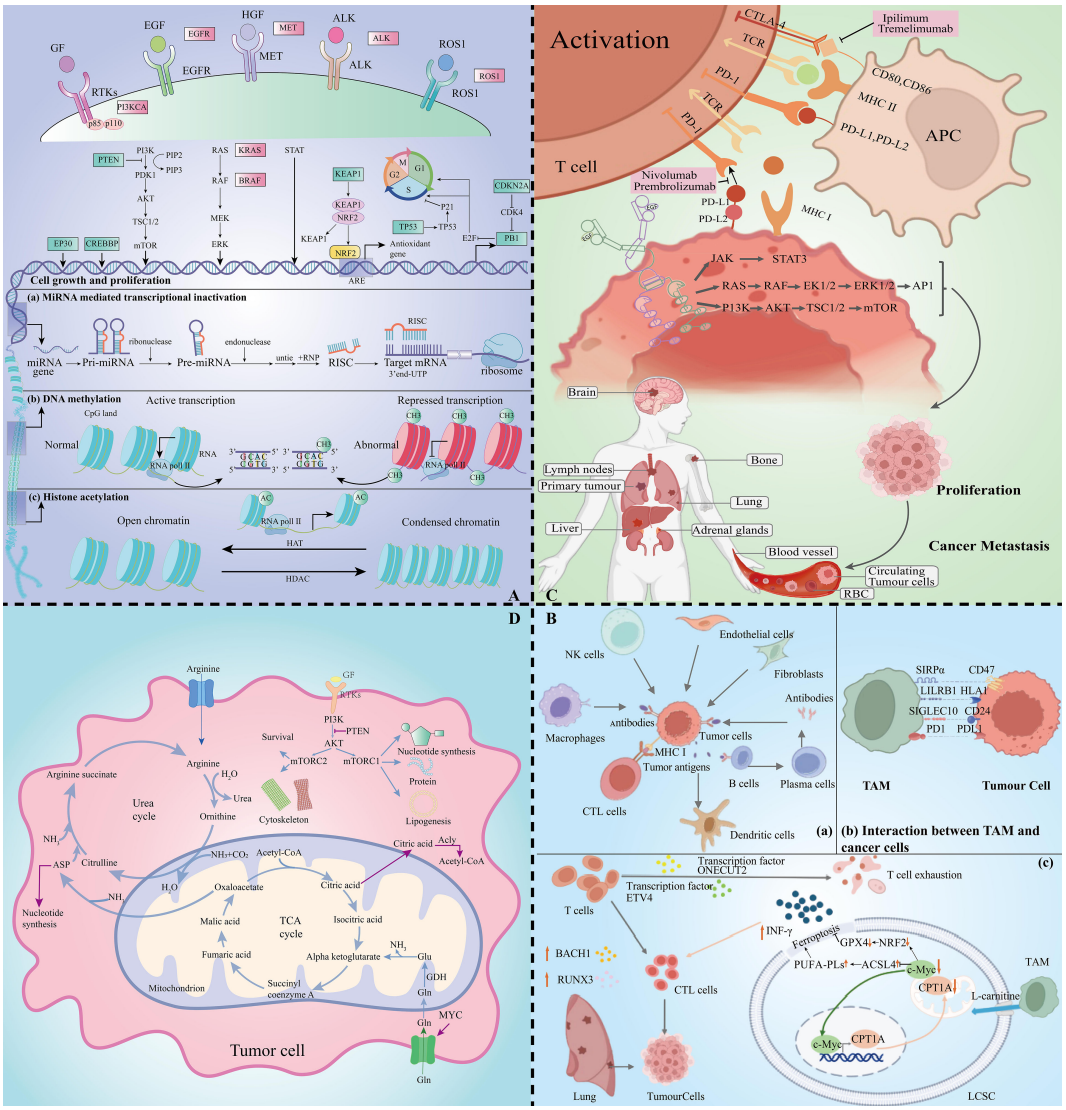


FIGURE 1
Lung cancer multidimensional histologic analysis map. **(A)** Exploratory mapping of changes in metabolomic profiles in lung cancer pathogenesis. **(B)** Immune cell profiles revealed by transcriptomics in the tumour microenvironment. **(C)** Critical mapping of proteomic changes during lung cancer progression. **(D)** Resolution of metabolic profiles in lung cancer.

and macrophages by single-cell RNA sequencing analysis of NSCLC patient samples, revealing multiple signalling pathways (e.g., EGFR, NOTCH, WNT, and PDGF, etc.) that are associated with carcinogenesis (Figure 1B). These findings shed more light on the molecular interactions and immunoregulatory mechanisms of NSCLC and provide a new perspective on the treatment of lung cancer (46).

In TME, individual cells are precisely regulated by transcription factors. Through the regulation of transcription factors, the killing of cells can be modulated. ONECUT2 and ETV4 were found to be likely potential regulators of CD8 T cell depletion in the blood of NSCLC patients, whereas the transcription factors BACH1 and RUNX3 were up-regulated in CD8 T cytotoxic subpopulations. Thus, regulation of these transcription factors may drive cytotoxic immune responses in NSCLC (47). Immune cell macrophages (TAM) in TME are among the most common immunosuppressive cells. Increased TAM in TME has been associated with immunotherapy resistance by transcriptomic analysis, and its expression is regulated by genes such as CD27, ITGAM and CCL5 (48). Recent studies have shown that TAM interacts with carnitine palmitoyltransferase 1A (CPT1A), increasing resistance to iron death and inactivation of CD8 T cells in lung cancer. Therefore, the use of CPT1 inhibitors enhances the killing of tumour cells by chemotherapy or immunotherapy (49) (Figure 1B).

In the field of lung cancer treatment, transcriptomics is often combined with other histological approaches to extend its application. By combining transcriptomics and metabolomics, the researchers analysed the effects of AZD-6482 (a PI3Kβ-targeted inhibitor) on 28 metabolite-related genes in LUAD. They found that the expression of three genes, LDHA, PPAT, and SMS, was increased in untreated LUAD samples; whereas after treatment with AZD-6482, the expression of these genes was significantly

TABLE 1 Targeted therapeutics drugs and targets in a genomic perspective.

Modifiable targets	Therapeutic drug	Reference
EGFR	Osimertinib	(19)
	Gefitinib	(20)
	Dacomitinib	(21)
	Erlotinib	(22)
	Afatinib	(23)
	Amivantamab	(24)
ALK	Crizotinib	(25)
	Ceritinib	(26)
	Alectinib	(27)
	Brigatinib	(28)
	Lorlatinib	(29)
RET	Cabozantinib	(30)
	Selpercatinib	(31)
	Pralsetinib	(32)
ROS1	Crizotinib	(33)
	Entrectinib	(34)
	Lorlatinib	(35)
MET	Glesatinib	(36)
BRAF V600E	Dabrafenib	(37)
	Trametinib	(38)
	Vemurafenib	(39)
KRAS	Adagrasib	(40)
	Sotorasib (AMG 510)	(41)
VEGFR	Bevacizumab	(42)
	Ramucirumab	(43)

decreased, suggesting that the inhibitor may improve the prognosis of LUAD patients (50). By joining forces with proteomics, Qing et al. used Cancer Cell Line Encyclopaedia (CCLE) RNA sequencing and proteomics profiles in human NSCLC cell lines to identify genes that are pan-sensitive and pan-resistant to drugs used in the treatment of NSCLC with systemic or targeted therapies (51).

4 Lung cancer immunotherapy and proteomics

With the rapid development of mass spectrometry (MS) technology, large-scale protein analysis has become a hotspot in scientific research, in which proteomics has achieved remarkable results in the study of protein phosphorylation, interaction, structure and function (52–55). In particular, proteomics has shown great potential for the discovery of new therapies and biomarkers. These biomarkers come from a wide range of sources, including body fluids and specific samples from lung cancer studies, such as breath condensate (56–59).

In the study of lung cancer proteomics, proteins in receptor tyrosine kinases such as EGFR and ALK and their downstream signalling pathways play a key role in the pathological process of lung cancer (60–62). EGFR, as a member of the ErbB family, promotes malignant cell survival, proliferation, etc. through a series of biochemical processes, making EGFR and its downstream signalling pathway an important target for lung cancer therapy.

In addition, the discovery of immune checkpoints (ICP) has led to a major breakthrough in the field of immunotherapy (63). In normal physiology, ICP maintains immune system homeostasis, but tumour cells evade immune attack by expressing ICP proteins (3). Among them, the interaction of programmed death receptor 1 (PD-1) with programmed cell death ligand-1 (PD-L1), cytotoxic T-lymphocyte-associated protein-4 (CTLA-4) and CD80/86 is the main mechanism of tumour cell escape (Figure 1C). Significant progress has been made in the development of targeted therapeutic agents for lung cancer against these immune checkpoint proteins (64). By blocking the function of these proteins, the immune system in the patient's body is activated to recognize and attack tumour cells more effectively, bringing new therapeutic hope to lung cancer patients. Over the past decade, tyrosine kinase inhibitors (TKIs) have made significant advances in the treatment of cancer, especially NSCLC.EGFR-TKI, as a potent agent for the treatment of over-activation of EGFR signalling, has been developed for multiple generations with remarkable efficacy (54, 65).

In the development of the field of immunotherapy, it is particularly important to achieve selective destruction of tumours by activating the immune response of T cells (64, 65). PD-1/PD-L1 inhibitors in combination with chemotherapy have become the standard of care in advanced NSCLC (62, 66). In clinical study finds,PD-1 antibodies such as Nivolumab and Pembrolizumab demonstrate durable efficacy in a variety of cancers (67, 68) (Figure 1C). In addition, anti-CTLA-4 antibodies such as Lpilimumab and Tremelimumab play an important role in immunotherapy.

Immunotherapy has great potential in the field of cancer treatment, including checkpoint inhibitors, monoclonal antibodies, and over-the-counter cell transplantation (69). Scholars such as Wang and Chiu emphasized that the combination of multiple therapies is the key to enhancing the effectiveness of cancer treatment and is expected to significantly improve patient survival rates (70, 71).

5 Lung cancer immunotherapy and metabolomics

Metabolomics delves into metabolite changes in organisms, providing new insights into the pathology and drug mechanisms of lung cancer. By analysing lung cancer samples and identifying metabolic markers closely related to lung cancer, it brings new

perspectives for early diagnosis, treatment planning and prognosis assessment (72–74).

In the oncogenic transformation of lung cancer cells, there are significant metabolic changes that are particularly dependent on energy sources such as ATP. Among these, the Warburg effect is particularly pronounced in lung cancer cells, where their metabolic needs are met by increased glucose uptake and support nucleotide and amino acid biosynthesis (75). SCLC is significantly dependent on exogenous arginine, which is associated with the deficiency or low expression of arginine succinyl synthase 1 (ASS1) (76, 77). The study found that up to 45% of SCLC samples and 50% of cell lines exhibited ASS1-negativity, highlighting the importance of arginine biosynthesis downregulation in the progression of lung carcinogenesis (78, 79).

The interaction between signalling and metabolism is critical in lung cancer research. mTOR kinases in the PI3K/Akt/mTOR pathway form the mTORC1 and mTORC2 complex, which influences protein, nucleotide, and lipid metabolism (80) (Figure 1D). MYC gene changes affect bioenergetic processes (81). Changes in metabolic state can also inversely regulate signalling pathway activity, e.g., mTORC1 activity is reduced during energy shortage (82).

Developing more effective lung cancer treatment regimens by modulating metabolic pathways or monitoring the disease using metabolic markers. In particular, lung cancer immunotherapy is closely linked to metabolomics, an important cornerstone of lung cancer treatment (83). The markers provided by metabolomics support personalized strategies for immunotherapy and improve treatment efficiency. In addition, the study by Ma et al. revealed the relationship between the regulation of amino acid metabolism, hypoxia-inducible factor-1 (HIF-1) and PI3K-Akt pathways and ositinib resistance, providing new perspectives for understanding the mechanism of drug resistance (84). These studies emphasize the critical role of metabolomics in monitoring marker changes after treatment (85).

For SCLC, polyethylene glycolated arginine deiminase (ADI-PEG20) and human recombinant polyethylene glycolated arginase (e.g., rhArgPEG, BCT-100, etc.) have been regarded as potential therapeutic targets due to arginine nutritional deficiency. And the combination of arginine with PD-1/PD-L1 inhibitors has also demonstrated efficacy in the clinic (76).

The combination of immunotherapy and metabolomics in lung cancer provides patients with more effective and personalized treatment options that are expected to improve their quality of life.

6 Discussion

In recent years, immunotherapies, particularly targeted therapies, have transformed the management and prognosis of lung cancer by providing personalized treatment options for lung cancer patients (86). And the integration of multi-omics data offers the unique advantage of aiming for a comprehensive assessment of each patient through extracted features, which promises a more complete picture of this complex immune ecosystem.

However, although ICI has been widely used in the treatment of lung cancer, there is still a lack of adequate understanding of prognostic biomarkers. We still need to increase research efforts on prognostic biomarkers for lung cancer at the multi-omics level such as proteomics and genomics. For example, genetic mutations in EGFR and KRAS have a key role in individualized therapy, but their specific impact and mechanisms in prognostic assessment need to be further explored (87, 88). In addition, proteomic markers such as CEA and CYFRA 21-1 show potential in disease surveillance, and their correlation with disease progression could provide additional information for disease management (89, 90).

Immune tolerance and therapeutic resistance are also current challenges in the combination of immunology and proteomics for the treatment of lung cancer. In particular, the EGFR T790M mutation leads to resistance to early drug (91–93). Specifically, although the third-generation EGFR-TKI ositinib has successfully treated patients with T790M mutations, new resistance mechanisms such as the EGFR T790M/C797S mutation are still emerging. Currently, there are investigators evaluating fourth-generation EGFR-TKIs clinically for new resistance issues (94, 95).

By integrating these multi-omics data, we can develop a more comprehensive understanding of the biological complexity of lung cancer, leading to the development of more effective therapeutic strategies and improved patient survival (96). Therefore, future research should focus on how to use these biomarkers to optimize treatment pathways, improve the accuracy of prognostic prediction, and ultimately achieve true precision medicine.

Author contributions

YY: Conceptualization, Data curation, Writing – original draft, Writing – review & editing. SS: Data curation, Writing – original draft. JML: Data curation, Writing – original draft. LS: Data curation, Writing – original draft. BW: Validation, Writing – original draft. JZ: Validation, Writing – original draft. JL: Validation, Writing – original draft. HL: Validation, Writing – original draft. PH: Validation, Writing – original draft. KX: Supervision, Writing – original draft. XS: Supervision, Writing – original draft. SH: Supervision, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This project is supported by the Doctoral Startup Fund of the Affiliated Hospital of Southwest Medical University (No. 19025), Innovation and Entrepreneurship Training Program of Southwest Medical University (No. 2023393), (Luzhou city- Southwest Medical University) Collaborative Application Foundation Project (No. 2023LZXNYDJ033).

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.

Publisher's note

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. Crinò L, Weder W, van Meerbeeck J, Felip E. Early stage and locally advanced (non-metastatic) non-small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. (2010) 21 Suppl 5:v103–15. doi: 10.1093/annonc/mdq207

3. Lahiri A, Maji A, Potdar PD, Singh N, Parikh P, Bisht B, et al. Lung cancer immunotherapy: progress, pitfalls, and promises. *Mol Cancer*. (2023) 22:40. doi: 10.1186/s12943-023-01740-y

4. de Sousa VML, Carvalho L. Heterogeneity in lung cancer. *Pathobiology*. (2018) 85:96–107. doi: 10.1159/000487440

5. Alyass A, Turcotte M, Meyre D. From big data analysis to personalized medicine for all: challenges and opportunities. *BMC Med Genomics*. (2015) 8:33. doi: 10.1186/s12920-015-0108-y

6. Ma A, McDermaid A, Xu J, Chang Y, Ma Q. Integrative methods and practical challenges for single-cell multi-omics. *Trends Biotechnol*. (2020) 38:1007–22. doi: 10.1016/j.tibtech.2020.02.013

7. Mohammadi-Shemirani P, Sood T, Paré G. From 'Omics to multi-omics technologies: the discovery of novel causal mediators. *Curr Atheroscler Rep*. (2023) 25:55–65. doi: 10.1007/s11883-022-01078-8

8. Chen P, Liu Y, Wen Y, Zhou C. Non-small cell lung cancer in China. *Cancer Commun (Lond)*. (2022) 42:937–70. doi: 10.1002/cac2.v42.10

9. Abbasian MH, Ardekani AM, Sobhani N, Roudi R. The role of genomics and proteomics in lung cancer early detection and treatment. *Cancers (Basel)*. (2022) 14:5144–60. doi: 10.3390/cancers14205144

10. Saller JJ, Boyle TA. Molecular pathology of lung cancer. *Cold Spring Harb Perspect Med*. (2022) 12:369–77. doi: 10.1101/cshperspect.a037812

11. MacGregor GA, Markandu ND, Smith SJ, Sagnella GA. Captopril: contrasting effects of adding hydrochlorothiazide, propranolol, or nifedipine. *J Cardiovasc Pharmacol*. (1985) 7 Suppl 1:S82–7. doi: 10.1097/00005344-198507001-00016

12. Rudin CM, Brambilla E, Faivre-Finn C, Sage J. Small-cell lung cancer. *Nat Rev Dis Primers*. (2021) 7:3. doi: 10.1038/s41572-020-00235-0

13. Yuan M, Zhao Y, Arkenau H-T, Lao T, Chu L, Xu Q. Signal pathways and precision therapy of small-cell lung cancer. *Signal Transduct Target Ther*. (2022) 7:187. doi: 10.1038/s41392-022-01013-y

14. Karpinets TV, Foy BD. Tumorigenesis: the adaptation of mammalian cells to sustained stress environment by epigenetic alterations and succeeding matched mutations. *Carcinogenesis*. (2005) 26:1323–34. doi: 10.1093/carcin/bgi079

15. Ansari J, Shackelford RE, El-Osta H. Epigenetics in non-small cell lung cancer: from basics to therapeutics. *Transl Lung Cancer Res*. (2016) 5:155–71. doi: 10.21037/tlcr.2016.02.02

16. Bajbouj K, Al-Ali A, Ramakrishnan RK, Saber-Ayad M, Hamid Q. Histone modification in NSCLC: molecular mechanisms and therapeutic targets. *Int J Mol Sci*. (2021) 22:11701. doi: 10.3390/ijms222111701

17. Murphy SJ, Harris FR, Kosari F, Barreto Siqueira Parrilha Terra S, Nasir A, Johnson SH, et al. Using genomics to differentiate multiple primaries from metastatic lung cancer. *J Thorac Oncol*. (2019) 14:1567–82. doi: 10.1016/j.jtho.2019.05.008

18. Hamouz M, Hammouz RY, Bajwa MA, Alsayed AW, Orzechowska M, Bednarek AK. A functional genomics review of non-small-cell lung cancer in never smokers. *Int J Mol Sci*. (2023) 24:13314. doi: 10.3390/ijms241713314

19. Roper N, Brown A-L, Wei JS, Pack S, Trindade C, Kim C, et al. Clonal evolution and heterogeneity of osimertinib acquired resistance mechanisms in EGFR mutant lung cancer. *Cell Rep Med*. (2020) 1:100007. doi: 10.1016/j.xcrm.2020.100007

20. Liu S-Y, Bao H, Wang Q, Mao W-M, Chen Y, Tong X, et al. Genomic signatures define three subtypes of EGFR-mutant stage II-III non-small-cell lung cancer with distinct adjuvant therapy outcomes. *Nat Commun*. (2021) 12:6450. doi: 10.1038/s41467-021-26806-7

21. Hou B, Lu X, Gao D-C, Liu Q-X, Zhou D, Zheng H, et al. REPORT- Clinical outcomes of using second - versus first-Generation EGFR-tkis for the First-Line treatment of advanced NSCLC patients with EGFR mutations: A meta-analysis. *Pak J Pharm Sci*. (2021) 34:1459–68.

22. Ma H, Tian X, Zeng X-T, Zhang Y, Wang Y, Wang F, et al. The efficacy of erlotinib versus conventional chemotherapy for advanced nonsmall-cell lung cancer: A PRISMA-compliant systematic review with meta-regression and meta-analysis. *Med (Baltimore)*. (2016) 95:e2495. doi: 10.1097/MD.0000000000002495

23. Yang JCH, Sequist LV, Geater SL, Tsai C-M, Mok TSK, Schuler M, et al. Clinical activity of afatinib in patients with advanced non-small-cell lung cancer harbouring uncommon EGFR mutations: a combined *post-hoc* analysis of LUX-Lung 2, LUX-Lung 3, and LUX-Lung 6. *Lancet Oncol*. (2015) 16:830–8. doi: 10.1016/S1470-2045(15)00026-1

24. Zhou C, Tang K-J, Cho BC, Liu B, Paz-Ares L, Cheng S, et al. Amivantamab plus chemotherapy in NSCLC with EGFR exon 20 insertions. *N Engl J Med*. (2023) 389:2039–51. doi: 10.1056/NEJMoa2306441

25. Solomon BJ, Bauer TM, Mok TSK, Liu G, Mazieres J, de Marinis F, et al. Efficacy and safety of first-line lorlatinib versus crizotinib in patients with advanced, ALK-positive non-small-cell lung cancer: updated analysis of data from the phase 3, randomised, open-label CROWN study. *Lancet Respir Med*. (2023) 11:354–66. doi: 10.1016/S2213-2600(22)00437-4

26. Soria J-C, Tan DSW, Chiari R, Wu Y-L, Paz-Ares L, Wolf J, et al. First-line ceritinib versus platinum-based chemotherapy in advanced ALK-rearranged non-small-cell lung cancer (ASCEND-4): a randomised, open-label, phase 3 study. *Lancet*. (2017) 389:917–29. doi: 10.1016/S0140-6736(17)30123-X

27. Wu Y-L, Dziadziuszko R, Ahn JS, Barlesi F, Nishio M, Lee DH, et al. Alectinib in resected ALK-positive non-small-cell lung cancer. *N Engl J Med*. (2024) 390:1265–76. doi: 10.1056/NEJMoa2310532

28. Camidge DR, Kim HR, Ahn M-J, Yang JC-H, Han J-Y, Lee J-S, et al. Brigatinib versus crizotinib in ALK-positive non-small-cell lung cancer. *N Engl J Med*. (2018) 379:2027–39. doi: 10.1056/NEJMoa1810171

29. Solomon BJ, Bauer TM, Ignatius Ou S-H, Liu G, Hayashi H, Bearz A, et al. *Post hoc* analysis of lorlatinib intracranial efficacy and safety in patients with ALK-positive advanced non-small-cell lung cancer from the phase III CROWN study. *J Clin Oncol*. (2022) 40:3593–602. doi: 10.1200/JCO.21.02278

30. Drilon A, Rekhtman N, Arcila M, Wang L, Ni A, Albano M, et al. Cabozantinib in patients with advanced RET-rearranged non-small-cell lung cancer: an open-label, single-centre, phase 2, single-arm trial. *Lancet Oncol*. (2016) 17:1653–60. doi: 10.1016/S1470-2045(16)30562-9

31. Drilon A, Oxnard GR, Tan DSW, Loong HHF, Johnson M, Gainor J, et al. Efficacy of selpercatinib in RET fusion-positive non-small-cell lung cancer. *N Engl J Med*. (2020) 383:813–24. doi: 10.1056/NEJMoa2005653

32. Gainor JF, Curigliano G, Kim D-W, Lee DH, Besse B, Baik CS, et al. Pralsetinib for RET fusion-positive non-small-cell lung cancer (ARROW): a multi-cohort, open-label, phase 1/2 study. *Lancet Oncol*. (2021) 22:959–69. doi: 10.1016/S1470-2045(21)00247-3

33. Shaw AT, Ou S-HI, Bang Y-J, Camidge DR, Solomon BJ, Salgia R, et al. Crizotinib in ROS1-rearranged non-small-cell lung cancer. *N Engl J Med*. (2014) 371:1963–71. doi: 10.1056/NEJMoa1406766

34. Desai AV, Robinson GW, Gauvain K, Basu EM, Macy ME, Maese L, et al. Entrectinib in children and young adults with solid or primary CNS tumors harboring NTRK, ROS1, or ALK aberrations (STARTRK-NG). *Neuro Oncol*. (2022) 24:1776–89. doi: 10.1093/neuonc/noac087

35. Shaw AT, Solomon BJ, Chiari R, Riely GJ, Besse B, Soo RA, et al. Lorlatinib in advanced ROS1-positive non-small-cell lung cancer: a multicentre, open-label, single-arm, phase 1-2 trial. *Lancet Oncol*. (2019) 20:1691–701. doi: 10.1016/S1470-2045(19)30655-2

36. Engstrom LD, Aranda R, Lee M, Tovar EA, Essenburg CJ, Madaj Z, et al. Glesatinib exhibits antitumor activity in lung cancer models and patients harboring MET exon 14 mutations and overcomes mutation-mediated resistance to type I MET inhibitors in nonclinical models. *Clin Cancer Res*. (2017) 23:6661–72. doi: 10.1158/1078-0432.CCR-17-1192

37. Planchard D, Kim TM, Mazieres J, Quoix E, Riely G, Barlesi F, et al. Dabrafenib in patients with BRAF(V600E)-positive advanced non-small-cell lung cancer: a single-arm, multicentre, open-label, phase 2 trial. *Lancet Oncol*. (2016) 17:642–50. doi: 10.1016/S1470-2045(16)00077-2

38. Bouffet E, Geoerger B, Moertel C, Whitlock JA, Aerts I, Hargrave D, et al. Efficacy and safety of trametinib monotherapy or in combination with dabrafenib in pediatric BRAF V600-mutant low-grade glioma. *J Clin Oncol*. (2023) 41:664–74. doi: 10.1200/JCO.22.01000

39. Hyman DM, Puzanov I, Subbiah V, Faris JE, Chau I, Blay J-Y, et al. Vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations. *N Engl J Med*. (2015) 373:726–36. doi: 10.1056/NEJMoa1502309

40. Skoulidis F, Goldberg ME, Greenawalt DM, Hellmann MD, Awad MM, Gainor JF, et al. STK11/LKB1 mutations and PD-1 inhibitor resistance in KRAS-mutant lung adenocarcinoma. *Cancer Discov*. (2018) 8:822–35. doi: 10.1158/2159-8290.CD-18-0099

41. Canon J, Rex K, Saiki AY, Mohr C, Cooke K, Bagal D, et al. The clinical KRAS (G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature*. (2019) 575:217–23. doi: 10.1038/s41586-019-1694-1

42. Patel JD, Socinski MA, Garon EB, Reynolds CH, Spigel DR, Olsen MR, et al. PointBreak: a randomized phase III study of pemetrexed plus carboplatin and bevacizumab followed by maintenance pemetrexed and bevacizumab versus paclitaxel plus carboplatin and bevacizumab followed by maintenance bevacizumab in patients with stage IIIB or IV nonsquamous non-small-cell lung cancer. *J Clin Oncol*. (2013) 31:4349–57. doi: 10.1200/JCO.2012.47.9626

43. Garon EB, Ciuleanu T-E, Arrieta O, Prabhash K, Syrigos KN, Goksel T, et al. Ramucirumab plus docetaxel versus placebo plus docetaxel for second-line treatment of stage IV non-small-cell lung cancer after disease progression on platinum-based therapy (REVEL): a multicentre, double-blind, randomised phase 3 trial. *Lancet*. (2014) 384:665–73. doi: 10.1016/S0140-6736(14)60845-X

44. Raja R, Kuziora M, Brohawn PZ, Higgs BW, Gupta A, Dennis PA, et al. Early reduction in ctDNA predicts survival in patients with lung and bladder cancer treated with durvalumab. *Clin Cancer Res*. (2018) 24:6212–22. doi: 10.1158/1078-0432.CCR-18-0386

45. Zhang P, Wu X, Wang D, Zhang M, Zhang B, Zhang Z. Unraveling the role of low-density lipoprotein-related genes in lung adenocarcinoma: Insights into tumor microenvironment and clinical prognosis. *Environ Toxicol*. (2024) 24:6212–22. doi: 10.1002/tox.24230

46. Wu F, Fan J, He Y, Xiong A, Yu J, Li Y, et al. Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. *Nat Commun*. (2021) 12:2540. doi: 10.1038/s41467-021-22801-0

47. Zhao Y, Zhang Q, Tu K, Chen Y, Peng Y, Ni Y, et al. Single-cell transcriptomics of immune cells reveal diversity and exhaustion signatures in non-small-cell lung cancer. *Front Immunol*. (2022) 13:854724. doi: 10.3389/fimmu.2022.854724

48. Larroquette M, Guegan J-P, Besse B, Cousin S, Brunet M, Le Moulec S, et al. Spatial transcriptomics of macrophage infiltration in non-small cell lung cancer reveals determinants of sensitivity and resistance to anti-PD1/PD-L1 antibodies. *J Immunother Cancer*. (2022) 10:e003890. doi: 10.1136/jitc-2021-003890

49. Ma L, Chen C, Zhao C, Li T, Ma L, Jiang J, et al. Targeting carnitine palmitoyl transferase 1A (CPT1A) induces ferroptosis and synergizes with immunotherapy in lung cancer. *Signal Transduct Target Ther*. (2024) 9:64. doi: 10.1038/s41392-024-01772-w

50. Thaiparambil J, Dong J, Grimm SL, Perera D, Ambati CSR, Putluri V, et al. Integrative metabolomics and transcriptomics analysis reveals novel therapeutic vulnerabilities in lung cancer. *Cancer Med*. (2023) 12:584–96. doi: 10.1002/cam4.v12.1

51. Ye Q, Hickey J, Summers K, Falatovich B, Gencheva M, Eubank TD, et al. Multi-omics immune interaction networks in lung cancer tumorigenesis, proliferation, and survival. *Int J Mol Sci*. (2022) 23:14978. doi: 10.3390/ijms232314978

52. Chen F, Chandrashekar DS, Varambally S, Creighton C.A.-O.X. Pan-cancer molecular subtypes revealed by mass-spectrometry-based proteomic characterization of more than 500 human cancers. *Nat Commun*. (2019) 10(1):5679. doi: 10.1038/s41467-019-13528-0

53. Hanash S, Taguchi A. Application of proteomics to cancer early detection. *Cancer J*. (2011) 17(6):423–8. doi: 10.1097/PPO.0b013e3182383cab

54. Moonmuang SA-O, Tantraworasin AA-O, Orrapin S, Udomruk S, Chewaskulyong B, Pruksakorn D, et al. The role of proteomics and phosphoproteomics in the discovery of therapeutic targets and biomarkers in acquired EGFR-TKI-resistant non-small cell lung cancer. *Int J Mol Sci*. (2023) 24 (5):4827. doi: 10.3390/ijms24054827

55. Zhang P, Zhang H, Tang J, Ren Q, Zhang J, Chi H, et al. The integrated single-cell analysis developed an immunogenic cell death signature to predict lung adenocarcinoma prognosis and immunotherapy. *Aging (Albany NY)*. (2023) 15:10305–29. doi: 10.18632/aging.205077

56. Tyers M, Mann M. From genomics to proteomics. *Nature*. (2003) 422 (6928):193–7. doi: 10.1038/nature01510

57. Zhou L, Beuerman RW. The power of tears: how tear proteomics research could revolutionize the clinic. *Expert Rev Proteomics*. (2017) 14(3):189–91. doi: 10.1080/14789450.2017.1285703

58. Geyer PA-O, Voytik E, Treit PV, Doll S, Kleinhempel A, Niu L, et al. Plasma Proteome Profiling to detect and avoid sample-related biases in biomarker studies. *EMBO Mol Med*. (2019) 11(11):e10427. doi: 10.15252/emmm.201910427

59. Youssef O, Sarhadi VK, Armengol G, Piirilä P, Knuuttila A, Knuutila S. Exhaled breath condensate as a source of biomarkers for lung carcinomas. A focus on genetic and epigenetic markers-A mini-review. *Genes Chromosomes Cancer*. (2016) 55 (12):905–14. doi: 10.1002/gcc.22399

60. Kumar R, George B, Campbell MR, Verma N, Paul AM, Melo-Alvim C, et al. HER family in cancer progression: From discovery to 2020 and beyond. *Adv Cancer Res*. (2020) 147:109–60. doi: 10.1016/bs.acr.2020.04.001

61. Holbro T, Hynes NE. ErbB receptors: directing key signaling networks throughout life. *Annu Rev Pharmacol Toxicol*. (2004) 44:195–217. doi: 10.1146/annurev.pharmtox.44.101802.121440

62. Sharma SV, Bell DW, Settleman J, Settleman DA, Haber DA. Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer*. (2007) 7(3):169–81. doi: 10.1038/nrc2088

63. Bagchi S, Yuan R, Engleman EG. Immune checkpoint inhibitors for the treatment of cancer: clinical impact and mechanisms of response and resistance. *Annu Rev Pathol*. (2021) 16:223–49. doi: 10.1146/annurev-pathol-042020-042741

64. Powles T, Durán I, van der Heijden MS, Loriot Y, Vogelzang NJ, De Giorgi U, et al. Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. *Lancet*. (2018) 391(10122):748–57. doi: 10.1016/S0140-6736(17)33297-X

65. Huang L, Jiang S, Shi YA-O. Tyrosine kinase inhibitors for solid tumors in the past 20 years (2001-2020). *J Hematol Oncol*. (2020) 13(1):143. doi: 10.1186/s13045-020-00977-0

66. Seegobin K, Majeed U, Wiest N, Manochakian R, Lou Y, Zhao Y. Immunotherapy in non-small cell lung cancer with actionable mutations other than EGFR. *Front Oncol*. (2021) 11:750657. doi: 10.3389/fonc.2021.750657

67. Antonia SJ, López-Martin JA, Bendell J, Ott PA, Taylor M, Eder JP, et al. Nivolumab alone and nivolumab plus ipilimumab in recurrent small-cell lung cancer (CheckMate 032): a multicentre, open-label, phase 1/2 trial. *Lancet Oncol*. (2016) 17 (7):883–95. doi: 10.1016/S1470-2045(16)30098-5

68. Patnaik A, Kang SP, Rasco D, Papadopoulos KP, Elassaiss-Schaap J, Beeram M, et al. Phase I study of pembrolizumab (MK-3475; anti-PD-1 monoclonal antibody) in patients with advanced solid tumors. *Clin Cancer Res*. (2015) 21(19):4286–93. doi: 10.1158/1078-0432.CCR-14-2607

69. Kelly RJ, Giaccone G. Lung cancer vaccines. *Cancer J*. (2011) 17(5):302–8. doi: 10.1097/PPO.0b013e318233e6b4

70. Wang J, Zou ZH, Xia HL, He JX, Zhong NS, Tao AL. Strengths and weaknesses of immunotherapy for advanced non-small-cell lung cancer: a meta-analysis of 12 randomized controlled trials. *PLOS ONE*. (2012) 7(3):e32695. doi: 10.1371/journal.pone.0032695

71. Chiu M, Armstrong EJL, Jennings V, Foo S, Crespo-Rodriguez E, Bozhanova G, et al. Combination therapy with oncolytic viruses and immune checkpoint inhibitors. *Expert Opin Biol Ther*. (2020) 20(6):635–52. doi: 10.1080/14712598.2020.1729351

72. Moreno P, Jiménez-Jiménez C, Garrido-Rodríguez M, Calderón-Santiago M, Molina S, Lara-Chica M, et al. Metabolomic profiling of human lung tumor tissues - nucleotide metabolism as a candidate for therapeutic interventions and biomarkers. *Mol Oncol*. (2018) 12(10):1778–96. doi: 10.1002/1878-0261.12369

73. Tonorezos ES, Jones LW. Energy balance and metabolism after cancer treatment. *Semin Oncol*. (2013) 40(6):745–56. doi: 10.1053/j.seminoncol.2013.09.011

74. Aoyagi T, Terracina KP, Raza A, Matsubara H, Takabe K. Cancer cachexia, mechanism and treatment. *World J Gastrointest Oncol*. (2015) 7(4):17–29. doi: 10.4251/wjgo.v7.i4.17

75. Nomura M, Morita M, Tanuma N. A metabolic vulnerability of small-cell lung cancer. *Oncotarget*. (2018) 9(64):32278–9. doi: 10.18632/oncotarget.25964

76. Carpentier JA-O, Pavlyk I, Mukherjee U, Hall PE, Szlosarek PW. Arginine deprivation in SCLC: mechanisms and perspectives for therapy. *Lung Cancer: Targets Ther*. (2022) 13:53–66. doi: 10.2147/LCTT.S335117

77. Hall PE, Ready N, Johnston A, Bomalaski JS, Venhaus RR, Sheaff M, et al. Phase II study of arginine deprivation therapy with pegarginase in patients with relapsed sensitive or refractory small-cell lung cancer. *Clin Lung Cancer*. (2020) 21(6):527–33. doi: 10.1016/j.cllc.2020.07.012

78. Szlosarek PW, Grimshaw MJ, Wilbanks GD, Hagemann T, Wilson JL, Burke F, et al. Aberrant regulation of argininosuccinate synthetase by TNF-alpha in human epithelial ovarian cancer. *Int J Cancer*. (2007) 121(1):6–11. doi: 10.1002/ijc.22666

79. Long Y, Tsai WB, Wangpaichitr M, Tsukamoto T, Savaraj N, Feun LG, et al. Arginine deiminase resistance in melanoma cells is associated with metabolic reprogramming, glucose dependence, and glutamine addiction. *Mol Cancer Ther*. (2013) 12(11):2581–90. doi: 10.1158/1535-7163.MCT-13-0302

80. Saxton RA, Sabatini DM. mTOR signaling in growth, metabolism, and disease. *Cell*. (2017) 169(2):361–71. doi: 10.1016/j.cell.2017.02.004

81. Stine ZE, Walton ZE, Altman BJ, Hsieh AL, Dang CV. MYC, metabolism, and cancer. *Cancer Discov*. (2015) 5(10):1024–39. doi: 10.1158/2159-8290.CD-15-0507

82. Krencz I, Sebestyen A, Khoor A. mTOR in lung neoplasms. *Pathol Oncol Res*. (2020) 26(1):35–48. doi: 10.1007/s12253-020-00796-1

83. Zhang PA-OX, Pei S, Zhou G, Zhang M, Zhang L, Zhang ZA-O. Purine metabolism in lung adenocarcinoma: A single-cell analysis revealing prognostic and immunotherapeutic insights. *J Cell Mol Med*. (2024) 28(8):e18284. doi: 10.1111/jcmm.18284

84. Ma QA-O, Wang J, Ren Y, Meng F, Zeng L. Pathological mechanistic studies of osimertinib resistance in non-small-cell lung cancer cells using an integrative metabolomics-proteomics analysis. *J Oncol*. (2020) 2020:6249829. doi: 10.1155/2020/6249829

85. Pan X, Chen W, Nie M, Liu Y, Xiao Z, Zhang Y, et al. A serum metabolomic study reveals changes in metabolites during the treatment of lung cancer-bearing mice with anlotinib. *Cancer Manag Res*. (2021) 13:6055–63. doi: 10.2147/CMAR.S300897

86. Zhang L, Cui Y, Mei J, Zhang Z, Zhang P.A.-O.X. Exploring cellular diversity in lung adenocarcinoma epithelium: Advancing prognostic methods and immunotherapeutic strategies. *Cell Prolif*. (2024) 30:e13703. doi: 10.1111/cpr.13703

87. Wang H, Yang X, Sun Y, Li Y, Dong Y, Shan D, et al. Synergistic effect of metformin and EGFR-TKI in the treatment of non-small cell lung cancer. *Transl Cancer Res*. (2020) 9(1):372–81. doi: 10.21037/tcr.2019.11.09

88. Reck M, Carbone DP, Garassino M, Barlesi F. Targeting KRAS in non-small-cell lung cancer: recent progress and new approaches. *Ann Oncol*. (2021) 32(9):1101–10. doi: 10.1016/j.annonc.2021.06.001

89. Cao WA-O, Tang Q, Zeng J, Jin X, Zu L, Xu S. A review of biomarkers and their clinical impact in resected early-stage non-small-cell lung cancer. *Cancers*. (2023) 15 (18):4561. doi: 10.3390/cancers15184561

90. Nijhawan HP, Prabhakar B, Misra A, Yadav KS. Fragmented antibodies in non-small cell lung cancer: A novel nano-engineered delivery system for detection and treatment of cancer. *Drug Discov Today*. (2023) 28(11):103701. doi: 10.1016/j.drudis.2023.103701

91. Oxnard GR, Arcila ME, Sima CS, Riely GJ, Chmielecki J, Kris MG, et al. Acquired resistance to EGFR tyrosine kinase inhibitors in EGFR-mutant lung cancer: distinct natural history of patients with tumors harboring the T790M mutation. *Clin Cancer Res*. (2011) 17(6):1616-. doi: 10.1158/1078-0432.CCR-10-2692

92. Wang ZF, Ren SX, Li W, Gao GH. Frequency of the acquired resistant mutation T790 M in non-small cell lung cancer patients with active exon 19Del and exon 21 L858R: a systematic review and meta-analysis. *BMC Cancer*. (2018) 18(1):148. doi: 10.1186/s12885-018-4075-5

93. Yu HA, Arcila ME, Rekhtman N, Sima CS, Zakowski MF, Pao W, et al. Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. *Clin Cancer Res*. (2013) 19(8):2240–7. doi: 10.1158/1078-0432.CCR-12-2246

94. Rangachari D, To C, Shpilsky JE, VanderLaan PA, Kobayashi SS, Mushajiang M, et al. EGFR-Mutated Lung Cancers Resistant to Osimertinib through EGFR C797S Respond to First-Generation Reversible EGFR Inhibitors but Eventually Acquire EGFR T790M/C797S in Preclinical Models and Clinical Samples. *J Thoracic Oncol*. (2019) 14 (11):1995–2002. doi: 10.1016/j.jtho.2019.07.016

95. Tripathi SK, Biswal BK. Allosteric mutant-selective fourth-generation EGFR inhibitors as an efficient combination therapeutic in the treatment of non-small cell lung carcinoma. *Drug Discov Today*. (2021) 26(6):1466–72. doi: 10.1016/j.drudis.2021.02.005

96. Ren Q, Zhang P, Lin H, Feng Y, Chi H, Zhang X, et al. A novel signature predicts prognosis and immunotherapy in lung adenocarcinoma based on cancer-associated fibroblasts. *Front Immunol*. (2023) 14:1201573. doi: 10.3389/fimmu.2023.1201573

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 30 September 2024
DOI 10.3389/fimmu.2024.1434300

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Jing Zhang,
University of South Dakota, United States
Yisheng Chen,
Fudan University, China
Guichuan Lai,
Chongqing Medical University, China

*CORRESPONDENCE
Jin Xing
xj980906260@163.com

†These authors have contributed equally to this work

RECEIVED 17 May 2024
ACCEPTED 09 September 2024
PUBLISHED 30 September 2024

CITATION
Zhao L, Shao W, Xiahou Z, Ren L, Liu C, Song Y, Xu H, Wang Z and Xing J (2024) IGFBP7+ subpopulation and IGFBP7 risk score in astrocytoma: insights from scRNA-Seq and bulk RNA-Seq.
Front. Immunol. 15:1434300.
doi: 10.3389/fimmu.2024.1434300

IGFBP7+ subpopulation and IGFBP7 risk score in astrocytoma: insights from scRNA-Seq and bulk RNA-Seq

Liang Zhao[1†], Wenwen Shao[2†], Zhikai Xiahou[3†], Li Ren[1], Chaobo Liu[1], Yanbing Song[1], Hao Xu[1], Zhihan Wang[1] and Jin Xing[1*]

[1]Department of Neurosurgery, Shanghai Pudong Hospital, Fudan University Pudong Medical Center, Pudong, Shanghai, China, [2]First Clinical Medical College, Shandong University of Traditional Chinese Medicine, Jinan, China, [3]China Institute of Sport and Health Science, Beijing Sport University, Beijing, China

Background: Glioma is the predominant malignant brain tumor that lacks effective treatment options due to its shielding by the blood-brain barrier (BBB). Astrocytes play a role in the development of glioma, yet the diverse cellular composition of astrocytoma has not been thoroughly researched.

Methods: We examined the internal diversity of seven distinct astrocytoma subgroups through single-cell RNA sequencing (scRNA-seq), pinpointed crucial subgroups using CytoTRACE, monocle2 pseudotime analysis, and slingshot pseudotime analysis, employed various techniques to identify critical subgroups, and delved into cellular communication analysis. Then, we combined the clinical information of GBM patients and used bulk RNA sequencing (bulk RNA-seq) to analyze the prognostic impact of the relevant molecules on GBM patients, and we performed *in vitro* experiments for validation.

Results: The analysis of the current study revealed that C0 IGFBP7+ Glioma cells were a noteworthy subpopulation of astrocytoma, influencing the differentiation and progression of astrocytoma. A predictive model was developed to categorize patients into high- and low-scoring groups based on the IGFBP7 Risk Score (IGRS), with survival analysis revealing a poorer prognosis for the high-IGRS group. Analysis of immune cell infiltration, identification of genes with differential expression, various enrichment analyses, assessment of copy number variations, and evaluation of drug susceptibility were conducted, all of which highlighted their significant influence on the prognosis of astrocytoma.

Conclusion: This research enhances comprehension of the diverse cell composition of astrocytoma, delves into the various factors impacting the prognosis of astrocytoma, and offers fresh perspectives on treating glioma.

KEYWORDS

astrocytoma, scRNA-seq, bulk RNA-seq, C0 IGFBP7+ glioma cells, prognosis

Introduction

Glioma is a tumor caused by glial cells or precursor cells (1). Gliomas are the predominant histological form of primary cancer in the central nervous system, including high-grade gliomas and low-grade gliomas (2, 3). As for the classification, WHO advocates dividing gliomas into I–IV grades (4). Glioblastoma multiforme (GBM) is the predominant malignant brain tumor, making up 60%–70% of malignant gliomas (2), and is classified as a highly invasive grade IV glioma (5). Glioblastoma, also known as malignant glioma, is the deadliest type of brain tumor, typically resulting in a median survival time of 15 months (6), glioblastoma is the most aggressive form of astrocytoma. Prior research has indicated that there are gender disparities in the occurrence of GBM in adults, with a higher prevalence among males (1).

Treating a brain tumor can be challenging due to the presence of the blood-brain barrier (BBB), which protects it (6). At present, surgery, radiotherapy, and chemotherapy are still the main treatment methods for glioma (4). GBM cannot be removed surgically because of its invasive nature and ability to infiltrate normal surrounding brain tissue (7). At present, the main drugs for GBM chemotherapy are temozolomide, or TMZ. TMZ slightly improved the survival rate of patients but caused many side effects (6). The GBM tumor has strong resistance to radiotherapy and cytotoxic chemotherapy (7). Hence, there is no superior remedy for GBM, necessitating a more profound comprehension of the illness and investigation into novel treatment approaches. Recent literature has indicated that the combination of temozolomide therapy and tumor-treating fields (TTFields) can enhance both progression-free survival and overall survival in patients with glioblastoma (8). TTFields represents a therapeutic modality that combats mitosis, although further investigation is needed to fully elucidate its experimental findings. Moreover, this treatment necessitates the utilization of a device, which entails head hair shaving and may impose an additional burden on patients. The adoption of a multimodal standard therapy still entails an inevitable recurrence rate, with a median survival exceedingly merely one year (9), so other therapeutic modalities still need to be explored.

Single-cell analysis has become an important tool for dissecting cellular heterogeneity (10, 11). This method has been extensively utilized for examining the internal diversity of different types of cancer, including non-small cell lung cancer (12), melanoma (13), cervical cancer (14), bladder cancer (15), prostate cancer (16) and clear cell renal cell carcinomas (ccRCCs) (17–20), among others. The characteristics and makeup of the tumor immune microenvironment (TIME) play a crucial role in the treatment and outlook of tumors. Research has shown that astrocytes play a role in the development of glioma, indicating that this relationship could be a potential focus for novel treatments (21). Research has extensively shown that astrocytes have the ability to control the attraction of tumor-associated macrophages (TAMs) to the tumor microenvironment (TME) through CCL2, leading to the progression of glioblastoma by encouraging a pro-tumor phenotype in TAMs (22). However, the tumor immune microenvironment of astrocytoma has not been fully explored

For this research, we utilized scRNA-seq to analyze single-cell data from GBM patients. We conducted dimensionality reduction clustering analysis on astrocyte subpopulations, followed by inferCNV analysis to identify astrocytoma. Our goal was to investigate the diverse heterogeneity of astrocytoma subpopulations and identify key subpopulations with the potential for high differentiation. Additionally, we explored the transcription factors associated with these subpopulations. Furthermore, a risk assessment model was developed, and the infiltration of immune cells in tumors was investigated along with clinical data from patients with glioma. Finally, we performed *in vitro* experimental validation. These studies could offer fresh insights for treating GBM.

Materials and methods

Get glioma data

The Glioma single-cell RNA-seq data utilized in this study were obtained from the NCBI Gene Expression Omnibus (GEO) database at https://www.ncbi.nlm.nih.gov/geo/. The identification code for logging in was GSE182109.

Data pertaining to bulk RNA-seq was acquired from the Cancer Genome Atlas (TCGA) website (https://portal.gdc.cancer.gov/), which included clinical details (age, gender, ethnicity) and somatic mutation information for glioma patients.

Raw data processing

The raw single-cell RNA data was analyzed using the "Seurat" package (version 4.3.0) (23, 24). To enhance data quality, the "DoubletFinder" R package (version 2.0.3) (17, 25) was utilized for eliminating doublet cells based on genetic data, followed by applying the "PercentageFeatureSet" function to filter out low-quality cells. High-quality cells meeting the criteria of (1) having 300 < nFeature < 7,500 genes detected in a single cell, (2) having 500 < nCount < 100,000 total transcriptomic count in a single cell, and (3) having the number of recognized genes in a single cell < 100,000 were retained. A single cell contains between 500 and 6,000 identifiable genes. Less than 20% of genes in a single cell were actively expressed by mitochondria.

Data clustering analysis with reduced dimensions

High-quality glioma cells were acquired and then normalized using the "NormalizeData"function, followed by the identification of the top 2000 variable genes using the "FindVariableFeatures" function. All genes were centered using "ScaleData" (26–29). To remove batch effects across various samples, the samples were processed and analyzed using the "harmony" R package (version 0.1.1) (14, 30).

The initial 30 primary components (PCs) were reduced in size with the "RunPCA" function, then the glioma cells were grouped and examined using "FindClusters" and "FindNeighbors" categorized based on the marker genes of cell subgroups mentioned in previous studies, and displayed through Uniform Manifold Approximation and Projection (UMAP) (31).

Detect astrocytoma utilizing InferCNV

By utilizing InferCNV (https://github.com/broadinstitute/inferCNV/wiki) (13), we were able to assess the astrocytes within the glioma cell subset and identify the differences in copy number within this subset. Taking EC (epithelial cell) as a control, the astrocytes with high-level copy number variation (CNV) were defined as astrocytoma.

Subgroup identification of astrocytoma

By clustering astrocytoma, we were able to identify various subgroups, revealing its internal heterogeneity. First of all, the top 2,000 highly mutated genes in astrocytoma were identified, then normalized, and the "harmony" R package was applied to reduce batch effects. Finally, the first 30 principal components (PC) were projected onto the two-dimensional map by using the UMAP map, and the different subsets of astrocytoma were marked according to the marker genes in previous literature (32, 33).

Furthermore, we investigated the origin of tissues and the cell cycle of various cell subgroups, computed staging scores like G2M.Score and S.Score, and compared the variations in G2M.Score, S.Score, nFeature, and nCount across different cell subgroups.

Identification and enrichment analysis of differentially expressed genes in astrocytoma subtypes

DEGs were identified for each astrocytoma subpopulation by screening with the "FindAllMarkers" function, detecting genes in a minimum of 25% of the cells with a false discovery rate (FDR) of less than 0.05 and an absolute log fold change (| logFCfilter |) greater than 1.

The "clusterProfiler" R package (version 0.1.1) (34, 35) was utilized for the analysis and enhancement of particular marker genes, with access to the Gene Ontology-Biological Processes (GOBP) database provided at http://www.geneontology.org (36, 37). During GO enrichment analysis, genes with p-values bel ow 0.05 were deemed to be statistically significant. Enriched entries were subjected to Gene Set Enrichment Analysis (GSEA) using gene sets obtained from the database (c2.cp.kegg.v7.5.1.symbols.gmt). Pathways that were significantly enriched were chosen using a false discovery rate (FDR) less than 0.05.

Trajectory analysis of astrocytoma

Stemness and developmental trajectories of astrocytoma subpopulations were comprehensively inferred using a variety of analytical methods, including CytoTRACE analysis, monocle2 analysis, and the Slingshot method.

CytoTRACE can re-establish the relative differentiation status of astrocytoma subpopulations based on gene expression profiles (38) and assess the stemness of different cellular subpopulations.

A proposed time-series analysis of astrocytoma subpopulations was performed using the Monocle R package (version 2.24.0). Monocle identified cellular alterations during astrocytoma differentiation as a means of inferring the developmental trajectory of the subpopulation.

Slingshot analysis (version 2.6.0) was used to detect and generate multiple differentiation trajectories for the astrocytoma subpopulation. The "getlineage" and "getCurves" functions were used to infer subpopulation differentiation trajectories and to assess changes in cell expression levels over time, respectively.

SCENIC analysis

To investigate the transcription factors (TFs) in the main subgroup, we utilized the pySCENIC algorithm to build a gene regulatory network, assessed the transcription factors' expression, and unveiled the general distribution of the main subgroup transcription factors.

Cell communication analysis

Astrocytoma subpopulations were analyzed for cellular communication using the 'CellChat' R package (version 1.6.1) (39), to examine and interpret inter-cellular communication networks derived from scRNA-seq data. The analysis was performed by integrating gene expression data from cells to establish the probability of communication through interactions between gene expression and signaling pathways, ligand-receptors, and their cofactors, which provided insights into the coordinated roles of signaling pathways in different cell types.

Construction of risk score and establishment of nomogram

Prognosis-related genes and corresponding risk scores for each sample were obtained through univariate COX risk regression analysis using the "survival" R package (40, 41), as well as Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression analysis (42–44) and multivariate COX risk regression. The risk score calculation formula: $Risk\ \ score = \sum_i^n Xi \times Yi$ (x: coefficient, y: gene expression level). According to the median risk score, the samples were divided into a high-risk group and a low-risk group. The prognostic features of various risk score categories

were assessed using Kaplan-Meier survival analysis and the "timeROC" R package (45–47).

We assessed the predictive precision of risk scores by merging patient clinical data with risk scores for multivariate COX risk regression analysis. We developed a nomogram model to predict 1-, 3-, and 5-year overall survival (OS) in glioma patients, visualized it using the "rms" R package, assessed the model's accuracy with c-index and ROC curves (48), and explored the relationship between model genes, risk scores, and OS.

Immune microenvironment analysis

In order to evaluate the correlation between risk characteristics and the immune microenvironment, we used a combination of the ESTIMATE, CIBERSORT, and Xcell algorithms to comprehensively evaluate the immune microenvironment of astrocytoma patients. Furthermore, the CIBERSORT algorithm (http://cibersort.stanford.edu/). was utilized to examine the distribution of 22 various immune cell types. We computed the ImmuneScore, StromalScore, ESTIMATEScore, and TumorPurity values, along with the TIDE (TumorImmune Dysfunction and Exclusion) scores. In addition, the relationship between model genes, risk score, and OS was explored to illustrate the important role of genes in immune-related functions.

Examining and enhancing the analysis of genes with varying expression levels in groups with high and low scores

The "DESeq2" was utilized to identify differentially expressed genes (DEGs) in groups with high and low risk scores, followed by enrichment analyses using the "clusterProfiler" R package (version 4.6.2) (49) for GO, Kyoto Encyclopedia of Genes and Genomes (KEGG) (50), and GSEA enrichment analyses.

Tumor mutation analysis

Glioma patient somatic mutation information was obtained from the TCGA database, and the Tumor Mutation Burden (TMB) was assessed in various scoring categories using the "maftools" R package (51), and the subjects were classified into high TMB and low TMB according to the median TMB. Participants were divided into high TMB and low TMB groups using the median TMB value, and survival differences were compared between the two groups using Kaplan-Meier curves. Pearson correlation coefficients were used to analyze the relationship between score and TMB. Furthermore, we analyzed the genetic variation in gene copies (CNV) of the modeled genes.

Drug sensitivity analysis

In order to better align with the clinical use of the drugs, we evaluated the sensitivity of the different drugs. The "pRRophetic" package (version 0.5) (52) was utilized to determine the IC50 value for each sample and assess the responsiveness of the groups with high and low risk scores.

Cell culture

The U87 MG and U251 MG cell lines were acquired from the American Type Culture Collection (ATCC). The two cell types were grown in DMEM medium with 10% fetal bovine serum and 1% streptomycin/penicillin (Gibco BRL, USA) at 37°C, 5% CO2, and 95% humidity as per usual conditions.

Cell transfection

Two small interfering RNAs (siRNAs) (GenePharma, Suzhou, China) were used to achieve FOSL2 knockdown, followed by inoculating cells in 6-well plates at 50% density. Transfection was performed with a negative control group (si-NC) and FOSL2 knockdown (si-FOSL2-1 and si-FOSL2-2). The transfection was carried out following the specific instructions provided by Lipofectamine 3000RNAiMAX (Invitrogen, USA).

Cell viability assay

The viability of U87 MG and U251 MG cells that were transfected was measured by utilizing the Cell Counting Kit-8 (CCK-8, A311-01, Vazyme). Cell suspensions were added to 96-well plates (5×10^3 cells per well) and left to incubate for 2 hours. The absorbance was then recorded at 450 nm on days 1, 2, 3, and 4. Mean optical density (OD) values were recorded, and the corresponding line graphs were plotted.

Quantitative real-time PCR

Cell lines were used to extract total RNA with TRIzol reagent (15596018, Thermo), followed by cDNA synthesis using PrimeScript™ RT Reagent Kit (R232-01, Vazyme). cDNA was isolated using the SYBR Green Real-Time PCR Kit from TaKaRa Biotechnology in Dalian, China, through real-time quantitative PCR (qRT-PCR). The primers and siRNAs utilized in this research are displayed in Supplementary Table 1.

Transwell

Cells (corning, USA) were either coated with or without matrix glue (BD Biosciences, USA) in a 24-well plate chamber. The cell suspension was then placed in the upper chamber with Costar and serum medium, while serum culture medium was added to the lower chamber. Put the cells in a cell incubator for 48 hours. Following incubation, the cells were treated with 4% paraformaldehyde and then stained with 0.1% crystal violet (Solarbio, China) to assess migration and invasion.

Plate-cloning experiment

Transfected cells were seeded in a 6-well plate at a density of 1×10^3 cells per well and incubated for 14 days. Next, the cells were rinsed with PBS and then treated with 4% paraformaldehyde (PFA) for a duration of 15 minutes. Finally, the cells were stained with 0.1% crystal violet (Solarbio, China) for 20 minutes and quantified.

Wound healing

After transfection, the cells were grown in 6-well dishes until they reached 95% confluence, then a 200-mL sterile pipette was used to wash away debris with PBS in a straight line through the cell layer. Next, the serum-free solution was exchanged to sustain cell culture, and images of the wounds at the identical spot at 0 hours and 48 hours were captured for assessing the breadth of the wounds.

5-Ethyl-2′-deoxyuridine proliferation assay

U87 MG and U251 MG cell lines that were transfected were plated in 6-well cell culture plates with 5×103 cells per well and left at room temperature for 24 hours. After that, a solution made by EdU was added to serum-free medium and incubated for 2 hours at 37°. Next, the cells were rinsed with PBS and then treated with 4% paraformaldehyde for a duration of 30 minutes. Afterward, the cells were exposed to glycine (2 mg/mL) and 0.5% Triton X-100 for 15 minutes, followed by incubation with 1 mL of 1× Apollo and 1 mL of 1× Hoechst 33342 for 30 minutes. The quantification of cell proliferation was ultimately determined using fluorescence microscopy.

Statistical analysis

The analysis of all the research was conducted using R software (version 4.3.0) and Python software (version 4.2.0). The Wilcoxon test, Pearson correlation coefficients, etc. Statistical tests were employed to evaluate the importance of variances among the groups (*P<0.05, **P<0.01, ***P<0.001).

Results

Main cell types of glioma

To comprehend the tumor microenvironment of glioma, we collected glioma cells from 18 patients following quality control of 234,148 high-quality cells. According to the marker genes, these high-quality cells were divided into 13 main cell types: microglia(49030), Myeloidcells (50565), Oligodendrocytes (29536), Astrocytes (46377), T_NK (28697), Excitatory_neuronal_cells (10997), Proliferating_cells (11346), Fibroblasts (1978), EndothelialCells(ECs)(1820), Muller_glia_cells (1580), B_Plasma(1245), Inhibitory_Neuronal _ Cells (814), Pericytes (163), and drawn into a 2D scatter plot by using Uniform Manifold Approximation and Projection (UMAP) technology (Supplementary Figure 1A). Additionally, we examined the tissue categories, cellular phases, and seurat groupings of each cell category, presenting them through UMAP visualizations paired with pie graphs (Supplementary Figures 1B–D). Bubble plots (Supplementary Figure 1E) displayed the top 5 marker genes for 13 cell types and 3 tissue types.

Subtype identification of astrocytoma

In order to distinguish malignant cells, we used the InferCNV algorithm to analyze the copy number variation (CNV) level of astrocytes, and the result was shown in Supplementary Figure 2. Based on the inferred CNV results, astrocytes with high levels of CNV were defined as tumor cells as astrocytoma. We classified the 40,650 astrocytomas obtained by Seurat and named the seven subclusters according to the marker genes as C0 IGFBP7+ Glioma cells, C1 OLIG2+ Glioma cells, C2 LINC02283+ Glioma cells, C3 LINC00632+ Glioma cells, C4 MX1+ Glioma cells, C5 FOSB+ Glioma cells, and C6 DLL3+ Glioma cells. The 2D map of UMAP dimensionality reduction combined with pie charts showed the distribution of subgroups and their proportion in different cell phases (G1, G2M, and S) and in different groups (II and IV) (Figure 1A). The results showed that most of the astrocytoma subclusters had a higher percentage of G1 cell cycle, in addition, C0 IGFBP7+ Glioma cells and C4 MX1+ Glioma cells had a higher percentage of Group IV, suggesting that the malignant degree of cells in these two subclusters might be higher. Figure 1B of the of the UMAP diagram showed the distribution of each subgroup and the proportion of cell cycle and group. Figure 1C UMAP faceted plots depicting the distribution of each subpopulation in detail.

Next, to dig deeper into the relevant features of astrocytoma, we calculated the Cell Stemness AUC (Area Under the Curve), nCount _RNA, nFeature _RNA, G2M.Calculated the Score, S. Score, and CNV Score for each subgroup and displayed them using UMAP plots (Figure 1D). The relevant features of different cellular phases and different groups were demonstrated with violin plots (Figures 1E–G). The results showed that C0 IGFBP7+ Glioma cells had the highest cell stemness among the seven subclusters, and C2 LINC02283+ Glioma cells had the highest G2M.Score (Figure 1E). In addition, compared with subgroup II, subgroup IV had higher cell stemness and had higher G2M.Score and CNV Score (Figure 1G).

Correlation enrichment analysis

To comprehend the biological mechanisms linked to each subgroup of astrocytoma, we conducted various enrichment analyses on the distinct genes within the seven subclusters of astrocytoma. Figure 2A violin plots demonstrated the expression levels of the named genes of the subclusters in each subcluster, and interestingly, IGFBP7, the named gene of the C0 subcluster, was

10.3389/fimmu.2024.1434300

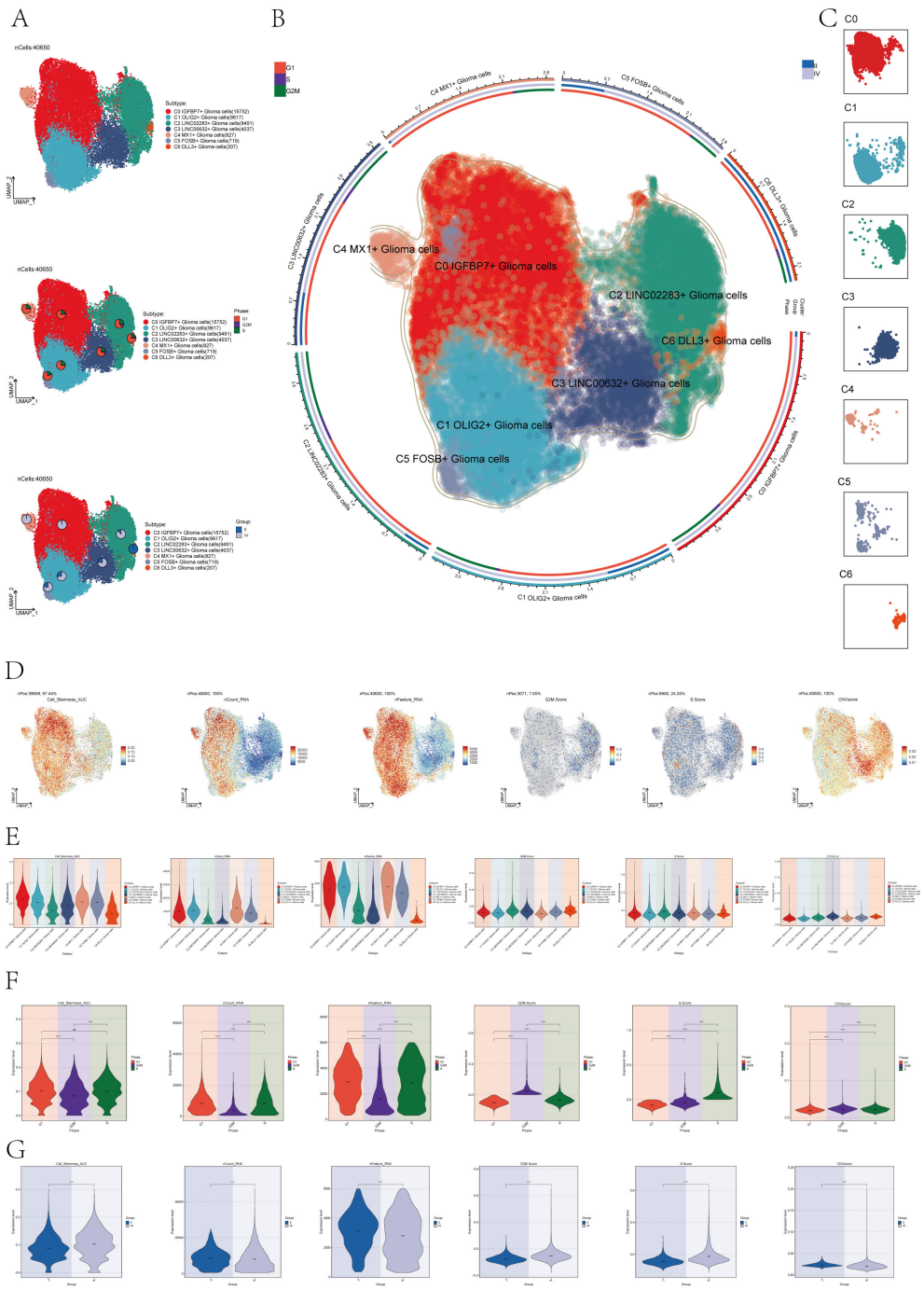


FIGURE 1
Subcluster identification of astrocytoma. **(A)** The UMAP plot revealed 7 subclusters of 40,650 astrocytoma (top). UMAP visualizations, along with pie graphs, illustrated the breakdown of individual subgroups based on Phases (G1, S, and G2M) (center) and Groups (II and IV) (lower section). **(B)** An integrated visualization demonstrated the distribution of astrocytoma subclusters, phases, and groups. **(C)** UMAP facet map exhibited the distribution of each astrocytoma subcluster. **(D)** UMAP plots individually showcased the Cell Stemness AUC, nCount RNA, nFeature RNA, G2M Score, S Score, and CNV Score of astrocytoma. **(E-G)** Violin plots respectively, displayed the levels of Cell Stemness AUC, nCount RNA, nFeature RNA, G2M Score, S Score, and CNV Score for each astrocytoma subcluster **(E)**, each cell phase **(F)**, and each group **(G)**. Significance levels were denoted as follows: ***P < 0.001, and ****P < 0.0001; NS was used to represent lack of significance.

also expressed in the C4 subcluster. We visualized the DEGs (differential expressed genes) in each subcluster of astrocytoma using volcano plots (Figure 2B).

Then, we plotted the gene cloud diagrams for each subgroup of astrocytoma and the cloud diagrams for enrichment analysis according to the number of gene occurrences and the level of enrichment scores of each subgroup, as shown in Figures 2C, D.

Furthermore, GSEA was conducted for every subgroup, revealing the pathways with the highest NES values displayed in Figure 2E. The top GSEA pathways for these seven subpopulations

10.3389/fimmu.2024.1434300

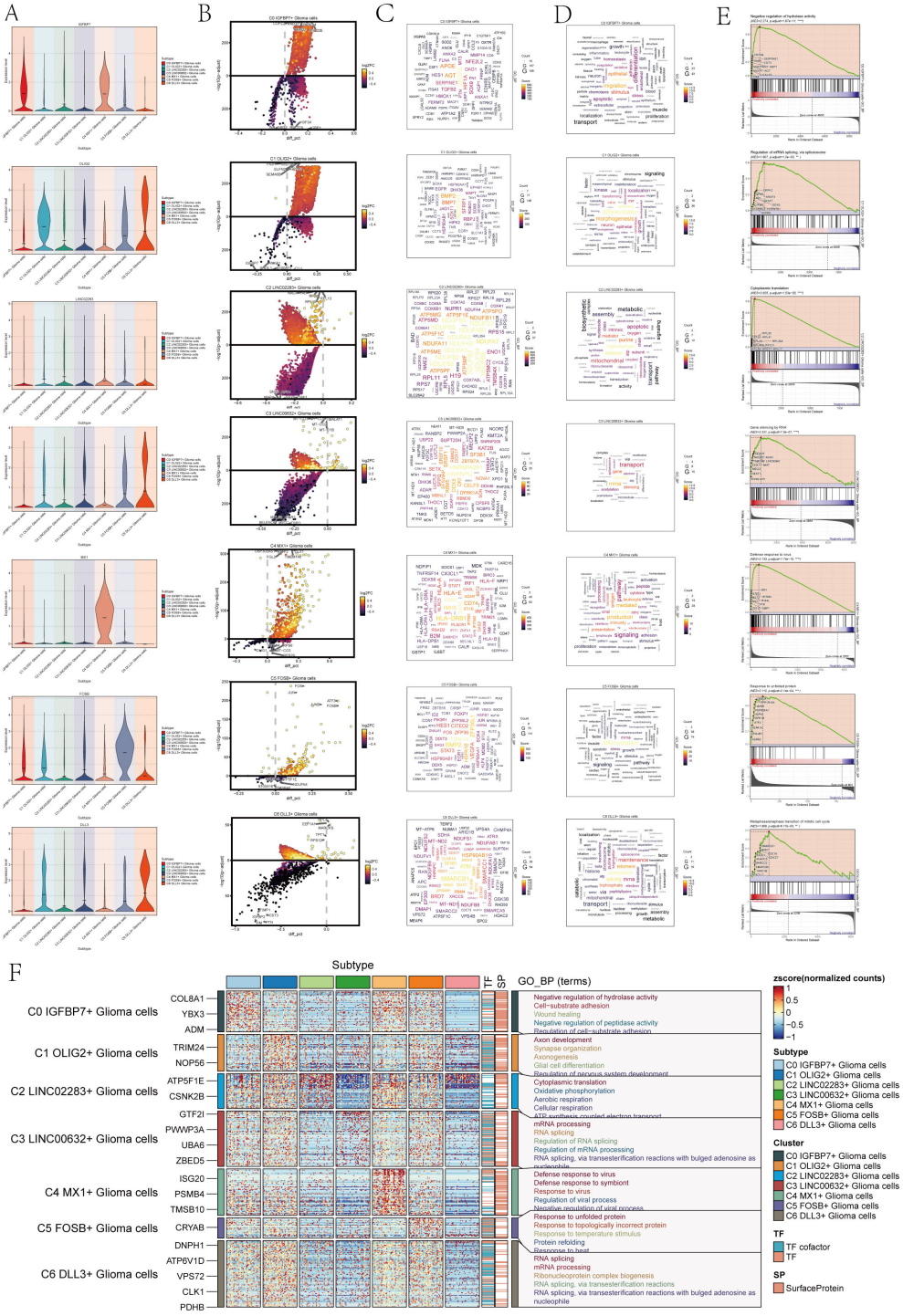


FIGURE 2
Enrichment analysis of astrocytoma subclusters. **(A)** Violin plots illustrated the distribution of named genes in each subcluster of the 7 astrocytoma subclusters. **(B)** Volcanic plots illustrated the genes with differential expression in the C0-C6 subgroups. **(C)** Cloud diagrams presented the expression patterns of highly-enriched genes in each astrocytoma subcluster. The font size indicated the quantity of genes, while the color indicated the enrichment score for each gene. **(D)** Cloud diagrams displayed the specific enriched pathways of highly-enriched genes in each astrocytoma subcluster. The font size indicated the quantity of genes, while the color indicated the enrichment score of genes within that pathway. **(E)** GSEA enrichment analysis results for each astrocytoma subcluster, showing only the pathway with the highest NES value. **(F)** Heatmap showed the gene expression and top 5 GO-BP enrichment analysis results for each astrocytoma subpopulation.

included negative regulation of hydrolase activity, regulation of mRNA splicing via the spliceosome, cytoplasmic translation, gene silencing by RNA, defense response to viruses, response to unfolded proteins, and metaphase/anaphase transition of the mitotic cell cycle based on the highest NES values.

In addition, in order to visualize the GOBP (Gene Ontology Biological Processes) enrichment analysis of each subpopulation of astrocytoma, we generated a heatmap to show the top 5 enriched terms of each subpopulation (Figure 2F).

The findings indicated that the enhanced pathways in C0 IGFBP7 + Glioma cells included inhibiting hydrolase activity, promoting cell-substrate adhesion, aiding in wound healing, inhibiting peptidase activity, and regulating cell-substrate adhesion. This result suggests that the C0 subpopulation may be associated with the adhesion movement of glioma cells. The enrichment pathways of C1 OLIG2+ Glioma cells for axis development, synapse organization, axionogenesis, glial cell differentiation, and regeneration of nervous system development suggest that this subpopulation may be involved in nervous system development and related tissue differentiation. Glioma cells with C2 LINC02283+ Glioma cells were found to have high levels of cytoplasmic translation, oxidative phosphorylation, aerobic respiration, cellular respiration, and ATP synthesis-linked electron transport, indicating a strong connection to cellular respiration and energy metabolism. On the other hand, glioma cells with C3 LINC00632+ Glioma cells showed enrichment in mRNA processing, RNA splicing, regulation of RNA splicing, regulation of mRNA processing, RNA splicing, and via transesterification reactions with bulged adenosine as a nucleophile, suggesting that this subpopulation may play a role in regulating RNA processing.

C4 MX1+ Glioma cells showed enrichment in immune responses to viruses and symbionts, as well as in regulating viral processes and negative regulation. On the other hand, C5 FOSB+ Glioma cells were enriched in responses to protein misfolding, temperature changes, and topologically incorrect proteins. Additionally, these cells also showed enrichment in responses to viruses, symbionts, viral processes, and negative regulation. Response to temperature stimulus, protein refolding, and resistance to heat pathways suggest that the C5 subpopulation may be involved in protein response-related biological processes. C6 DLL3+ Glioma cells are involved in RNA splicing, mRNA processing, and ribonucleoprotein complexes. Ribonucleoprotein complex formation occurs through RNA splicing, involving transesterification reactions and bulged adenosine. The C6 subpopulation may be involved in RNA splicing and other related biological processes through transesterification reactions involving bulged adenosine as a nucleophile.

Trajectory analysis of the astrocytoma subcluster

To investigate the differentiation status and developmental trajectory of seven subgroups of astrocytoma, we performed CytoTRACE analysis and monocle 2 pseudotime analysis on these subgroups. The related results of the CytoTRACE analysis were shown in Figures 3A, B. The CytoTRACE results showed that the CytoTRACE scores of subcluster C1, subcluster C4, and subcluster C0 were higher, indicating that the stemness was higher in these three subclusters. The gene correlations involved in the CytoTRACE analysis can be observed in the bar graph (Figure 3C). The findings from the pseudotime analysis of the astrocytoma subgroup were displayed in Figure 3D. The findings indicated that the pseudotime path deviated from the top right to the bottom left, encompassing six stages and three points of divergence. The pseudotime facets of the along-trajectory distribution of each subgroup of astrocytoma were shown in Figure 3E. In addition, we further demonstrated the pseudotime results of astrocytoma subgroups with Violin plots and ridge plots (Figures 3F–H). These results indicated that C0 IGFBP7+ Glioma cells might be at the end of differentiation and have high differentiation ability, and C6 DLL3+ Glioma cells might be at the initial stage of differentiation.

Transcription factors related to the C0 IGFBP7+ glioma cells subgroup

We analyzed the TOP1 transcription factor FOSL2 of the C0 IGFBP7+ Glioma cells subgroup, which may be at the end of differentiation. Initially, a UMAP visualization was created to display the distribution of the transcription factor FOSL2 (Figure 3I), revealing its limited presence in various subgroups. The specific differences in the distribution of transcription factor FOSL2 in each subgroup were shown in Figure 3J. The transcription factor FOSL2 was most distributed in the C0 IGFBP7+ Glioma cells subgroup, and the distribution in other subgroups was different, with statistical differences. In addition, the transcription factor FOSL2 was more distributed in highly differentiated tissues (Group IV) than in Group II, and the results were statistically different (Figure 3K).

Slingshot pseudotime analysis of the astrocytoma subcluster

In order to further confirm the differentiation relationship between different subgroups of astrocytoma, we conducted a slingshot pseudotime analysis on the astrocytoma subgroup. The findings indicated the presence of two lineages in the slingshot pseudotime assessment of seven subtypes of astrocytoma (Figure 4A). Lineage 1 originated from C2 and ends at CO. Lineage 2 originated from C2, passed through CO/C4→C1/C5→C3, and ended with C6. However, there was only one lineage in the slingshot pseudotime analysis of two Groups (II and IV) (Figure 4B). The expression of named genes with subpopulation slingshot pseudotime analysis lineage 1 was shown in Figure 4C scatter plots, and the expression of named genes with subpopulation slingshot pseudotime analysis lineage 2 was shown in Figure 4D. In addition, we also analyzed the trajectories of the slingshot pseudotime analysis of different groups (IV and II), and the slingshot pseudotime analysis of different groups only had lineage 1. The expression of string hot pseudotime analysis

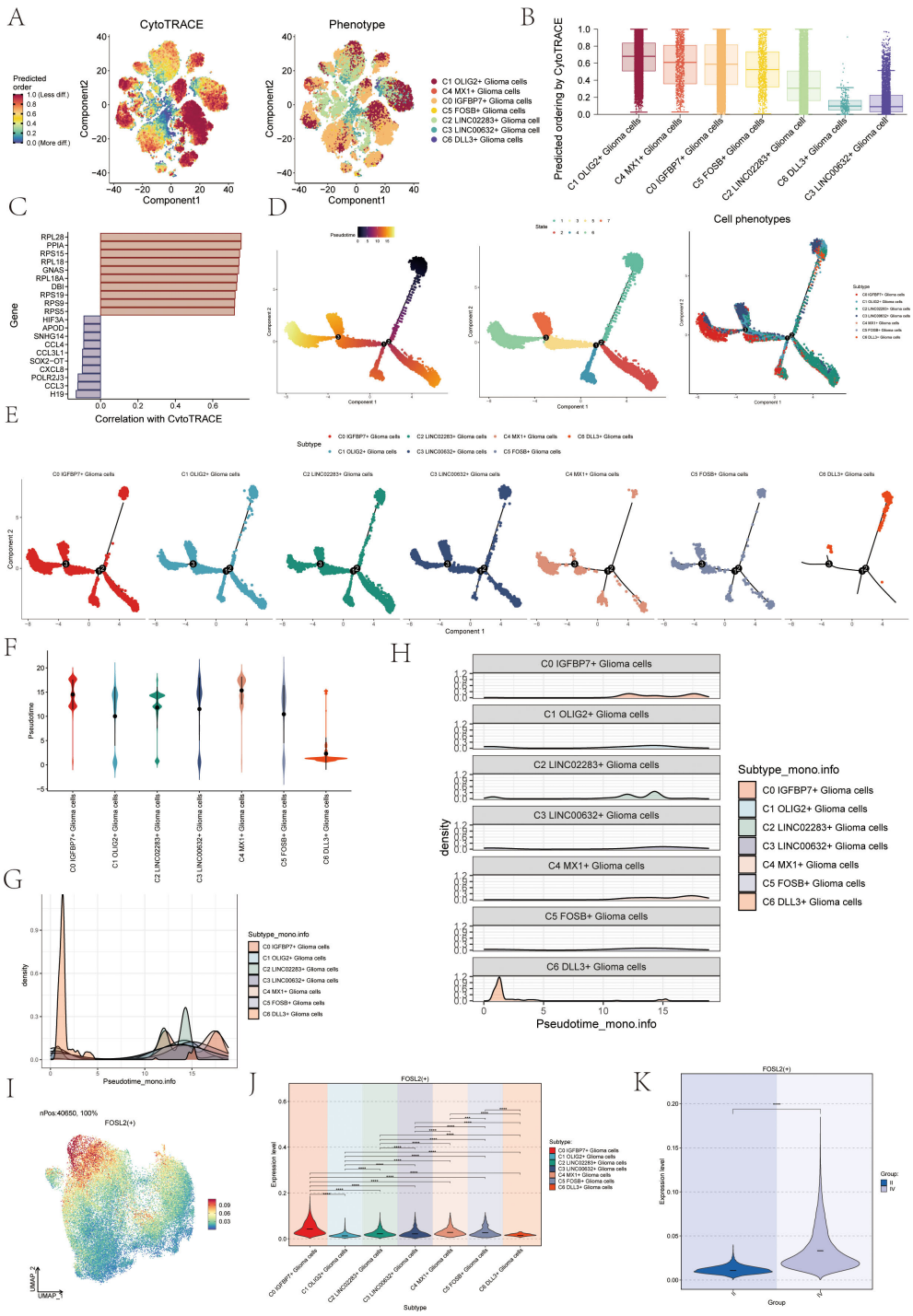


FIGURE 3
CytoTRACE and Monocle2 pseudotime analysis of astrocytoma subpopulations and related transcription factors. **(A)** CytoTRACE analysis and visualization of the differentiation degree for each astrocytoma subpopulation. In the left figure, dark-green indicated higher differences (low stemness), while dark-red indicated lower differences (high stemness). In the right figure, different colors represent different astrocytoma subpopulations. **(B)** Boxplot displayed the CytoTRACE analysis results, revealing that C1 OLIG2+ Glioma cells, C4 MX1+ Glioma cells, and C0 IGFBP7+ Glioma cells exhibited higher differentiation potential, while C3 LINC00632+ Glioma cells had the lowest differentiation potential. **(C)** Bar graph showed the gene correlations in the CytoTRACE analysis. **(D)** Trajectory analysis using Monocle2, with 3 branch points and 6 states. **(E)** Monocle2 pseudotime analysis facet map depicted the trajectories of the 7 astrocytoma subclusters. **(F)** Violin plots showed the distribution of the 7 astrocytoma subgroups along the pseudotime trajectory. **(G, H)** Ridge plots and their facet maps displayed the density distribution of the 7 astrocytoma subgroups along the pseudotime trajectory. **(I)** UMAP plot visualized the distribution of the top transcription factor (TF) FOSL2 in C0 IGFBP7+ Glioma cells. **(J)** Violin plot presented the distribution of FOSL2 in astrocytoma for each subcluster. **(K)** Violin plot illustrated the distribution of FOSL2 in different groups (II and IV). Significance levels were denoted as follows: **P < 0.01, ***P < 0.001, and ****P < 0.0001; NS was used to represent lack of significance.

10.3389/fimmu.2024.1434300

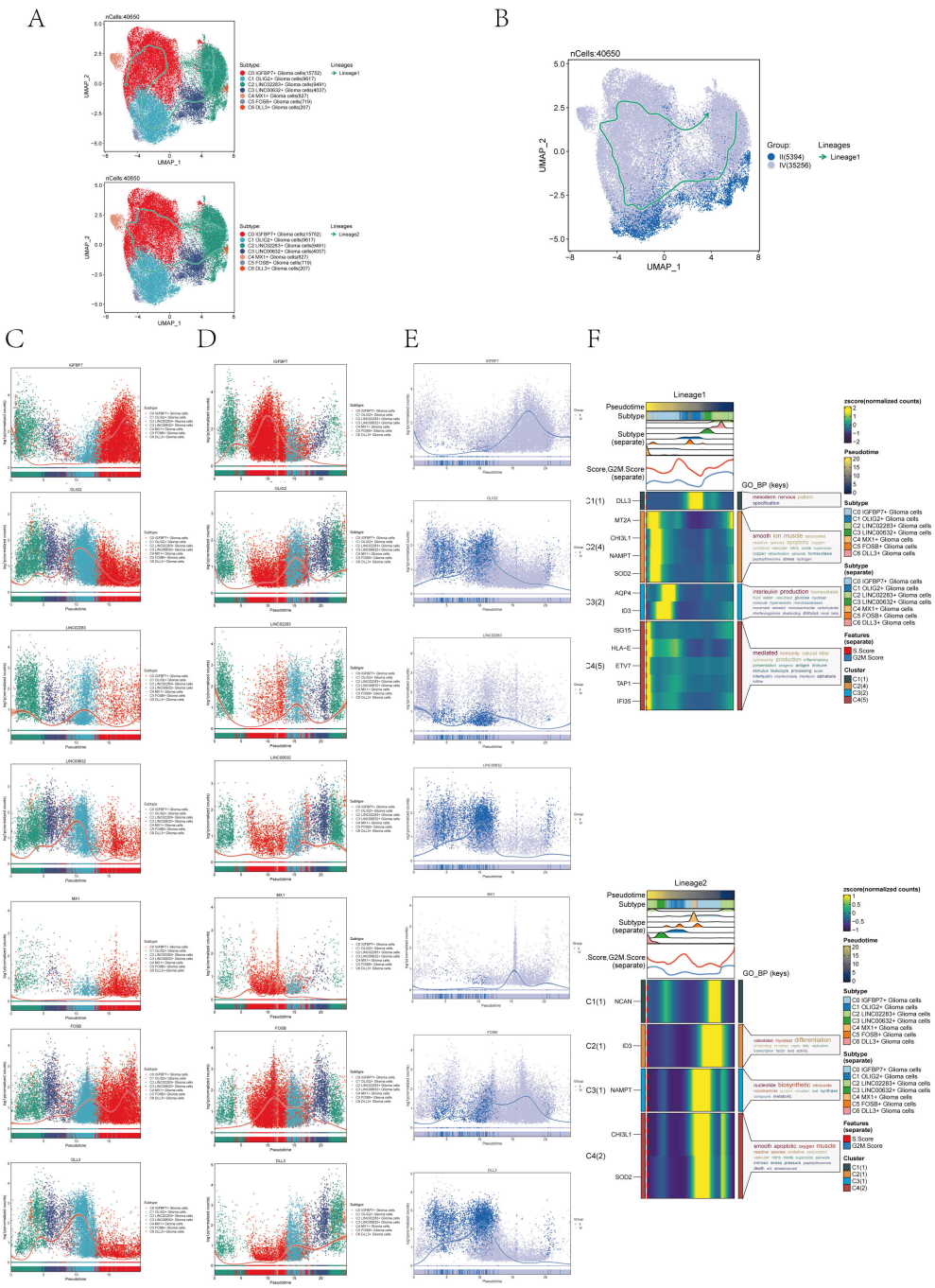


FIGURE 4
Slingshot pseudotime analysis of astrocytoma. **(A)** Slingshot pseudotime analysis results for the 7 subclusters of astrocytoma reveal 2 lineages. **(B)** Slingshot pseudotime analysis results for different Groups (II and IV) of astrocytoma, showing 1 lineage. **(C)** Scatter plots demonstrated the expression changes of lineage 1-associated genes in the astrocytoma subclusters. **(D)** Scatter plots illustrated the expression changes of lineage 2-associated genes in the astrocytoma subclusters. **(E)** Scatter plots displayed the expression changes of lineage 1-associated genes in the astrocytoma subclusters across the Groups. **(F)** Heatmaps exhibited the expression changes of differentially expressed genes along the trajectories of the 2 lineages of the astrocytoma subclusters, along with their GOBP enrichment analysis results.

lineage 1 with different groups of named genes was shown in Figure 4E. The findings indicated that the gene IGFBP7, belonging to the C0 subgroup, exhibited the highest expression levels in Group IV following the slingshot pseudotime analysis. This was consistent with the previous results of CytoTRACE analysis and monocle 2 pseudotime analysis that C0 IGFBP7+ Glioma cells were at the end of differentiation and had high cell stemness, with most of the C0 subclusters distributed in subgroup IV. We also analyzed the expression of two lineages of DEGs with the subgroup's slingshot pseudotime analysis, and the result was as shown in Figure 4F. We

also analyzed the enrichment of DEGs by GOBP and found that lineage 1 was mainly enriched with mesodem nervous, smooth ion muscle, DEGs of interleukin production, and mediated, while lineage 2 was enriched with osleoblast, nucleotide biosynthetic, smooth apoptotic, and other pathways.

Cellular communication network

In order to systematically explore the interaction of the tumor microenvironment in astrocytoma, we used Cellchat analysis to draw a cell communication network to show the intensity (Figure 5A) and quantity (Figure 5B) of ligand-receptor interaction between different cell groups. Then, we analyzed the signal patterns between astrocytoma and other cells and the interaction between cells and pathways. Three outgoing signal patterns and three incoming signal patterns were identified, and the results were shown in Figures 5C, D. Figures 5E, F displayed the communication patterns received by target cells and sent by secreting cells, respectively. The results showed that both C0-C6 subgroups were involved in the PTN signal network pathway.

The results of Sankey diagrams revealed that the cell groups in outgoing patterns were astrocytoma, Oligodendrocytes, Excitatory neuronal cells, Proliferating cells, Muller glia cells, Inhibitory neuronal cells and Pericytes send out signals in coordination with the signal paths PTN, MK and SPP1 belonging to pattern1 (Figure 5G).

The heatmap specifically showed (Figure 5H) the signal intensity of outgoing signaling patterns and incoming signaling patterns interacting with other cell types.

PTN signaling network pathway

Because the seven subpopulations of astrocytoma involved the PTN signaling network pathway in both Incoming communication and Outgoing communication, we initiated further studies on the PTN pathway. When glioma cells were used as receiver cells in PTN signaling network pathway, the interaction of various cell-like cells was shown in Figure 6A, and the expression levels of signal genes involved in this pathway (PTPRZ1, SDC2, SDC3, NCL) were shown in Figure 6B. The results showed that NCL has a certain expression level in various cell-like cells. We further explored the cellular communication networks involved in PTN signaling network. We further explored the PTN signaling network involved in the intercellular communication network, PTPRZ1 Glioma cells, SDC2 Glioma cells, SDC3 Glioma cells, NCL Glioma cells were used as receivers in the respective hierarchical plots shown in Figure 6C (PTPRZ1), Figure 6D (SDC2), Figure 6E (SDC3), and Figure 6F (NCL), respectively. The specific interaction between PTN signaling network cells was comprehensively demonstrated by the heatmap (Figure 6G). The heatmap (Figure 6H) displayed the varying significance of different cell types in the PTN signaling pathway network, acting as senders, receivers, mediators, and influencers. The results showed that in the PTN pathway, six subgroups of malignant asteroids (C0-C5) had high relative importance as sender, receiver, mediator and influencer.

Establishment and verification of a prognostic model

In order to better serve the clinic, we evaluated the prognostic characteristics of the C0 IGFBP7+ Glioma cell subgroup identified in this study.

Initially, we analyzed the leading 100 potential genes in this specific group through univariate Cox regression analysis, revealing that 29 genes were linked to patient prognosis (Figure 7A). In order to avoid the multiple contributions of the screened genes, we conducted LASSO regression analysis on these 29 genes (Figure 7B), and a total of 4 genes were determined to be significantly related to the prognosis of patients. After screening four genes (FAM20C, TIMP1, PMP22, and ID1), we performed a multivariate Cox regression analysis and identified three genes (FAM20C, TIMP1, and PMP22) as risk factors, with gene ID1 being a protective factor (Figure 7C). Using the Cox regression coefficient for each gene, we developed an IGFBP7 Risk Score (IGRS) and determined the IGRS for each sample based on gene expression and the associated coefficient. The specific formula was: IGFBP7 Risk Score (IGRS) = ID1 expression level * (-0.206) + TIMP1 expression level* 0.130 + FAM20C level* 0.192 + PMP22 level* 0.052. According to the score, we divided the C0 IGFBP7+ Glioma cell subgroup into High IGRS Group and Low IGRS Group, and further analyzed the high and low IGRS Groups. The IGFBP7 Risk Score of high and low IGRS Groups and the changes of their living state with time were shown on the left of Figure 7D. The expression of four construction model genes in High IGRS Group and Low IGRS Group was shown on the right of Figure 7D. The findings indicated that the genes FAM20C, TIMP1, and PMP22 exhibited high expression levels in the High IGRS Group, while the gene ID1 displayed high expression in the Low IGRS Group. Survival analysis comparing high and low IGRS groups indicated that the survival rate was lower in the high IGRS group compared to the low IGRS group (Figure 7E). AUC scores for 1 year and 3 years were shown in Figure 7F. We analyzed the survival of four modeling genes (FAM20C, TIMP1, PMP22, and ID1) (Figure 7G), and the results showed that three genes (FAM20C, PMP22, and ID1) had statistical differences. Among them, the high expression of FAM20C and PMP22 genes has a worse prognosis, while the high expression of gene ID1 has a better survival outcome. Further prove the previous conclusion: genes FAM20C and PMP22 were associated with adverse outcomes.

Nomogram creation

In order to further analyze whether IGFBP7 Risk Score can be an independent risk factor, we conducted multivariate Cox regression analysis on clinical factors (gender, age, and race) and IGFBP7 Risk Score (Figure 8A). The results of forest plot showed that IGRS Group and IGRS score can be independent prognostic factors.

In order to determine if the IGFBP7 Risk Score could act as a standalone risk factor, we conducted a multivariate Cox regression analysis that included clinical factors such as gender, age, and

ethnicity along with the IGFBP7 Risk Score (Figure 8B). The forest plot results suggested that both the IGRS Group and IGRS score may act as separate prognostic factors. Figure 8C displayed the AUCs for survival at 1-year and 3-year intervals, while Figures 8D, E illustrated the calibration curves for the nomograms at the same intervals, indicating that the nomograms accurately predicted the OS of the training group. Figure 8F displayed the pairwise correlation between the four modeling genes, OS, and IGFBP7 Risk Score. The two-by-two correlations between the four modeling genes, OS and IGFBP7 Risk Score were shown in Figure 9F. The

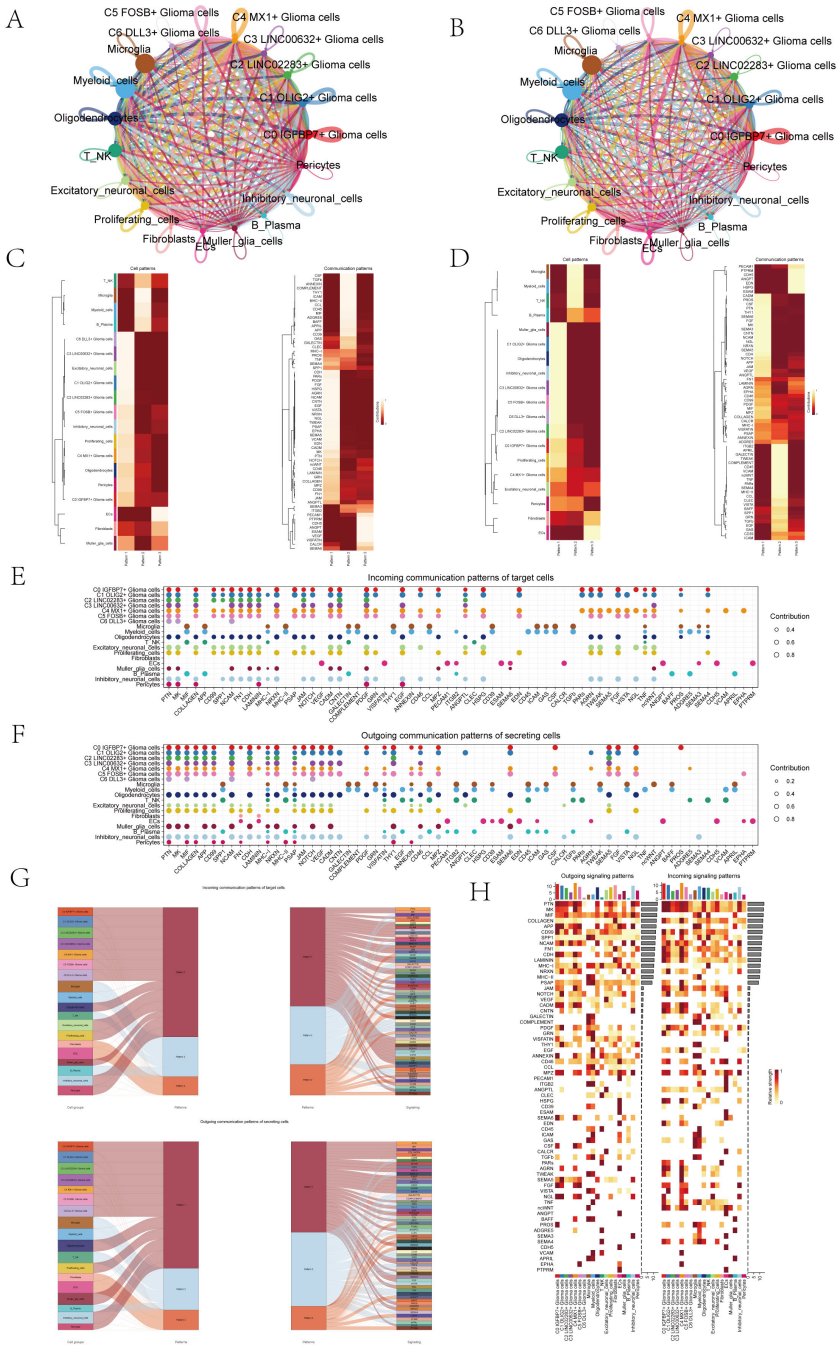


FIGURE 5
Overview of Cell Communication. **(A)** Weighted interaction network diagram of cellular interactions for all cell types. Thicker lines represented stronger interactions between the cell types. **(B)** Interaction count network diagram of cellular interactions for all cell types. Thicker lines indicated a higher count of interactions between the cell types. **(C, D)** Heatmaps respectively displayed the patterns identified in the incoming communication **(C)** and outgoing communication **(D)**. **(E, F)** Dot plots compared the communication patterns received by target cells **(E)** with the communication patterns sent out by secreting cells **(F)**. **(G)** Sankey charts illustrated the projected communication flow patterns of recipient cells, revealing the coordination between cells receiving signals and their interaction with specific signaling pathways in response (top). In addition, the secretion behaviors of cells were illustrated (bottom), demonstrating how cells interact as message transmitters and how they interact with specific signaling pathways to facilitate communication. **(H)** The heatmap displayed the communication patterns of all cells, showing both outgoing and incoming signals.

correlations between the four modeling genes and IGFBP7 Risk Score were visualized with scatter plots (Figure 8G), and the results showed that genes FAM20C, TIMP1, and PMP22 were positively correlated with Risk and gene ID1 was negatively correlated with Risk. The correlation analysis of the 4 modeled genes with OS was shown in Figure 8H, the results showed that FAM20C, TIMP1, and PMP22 were negatively correlated with OS, while gene ID1 was positively correlated with OS. Then, we further analyzed the specific expression of the four modeled genes in High IGRS Group and Low IGRS Group, and the results were demonstrated by ridge plots combined with box plots (Figure 8I). In the High IGRS Group, the genes FAM20C, TIMP1, and PMP22 exhibited increased expression

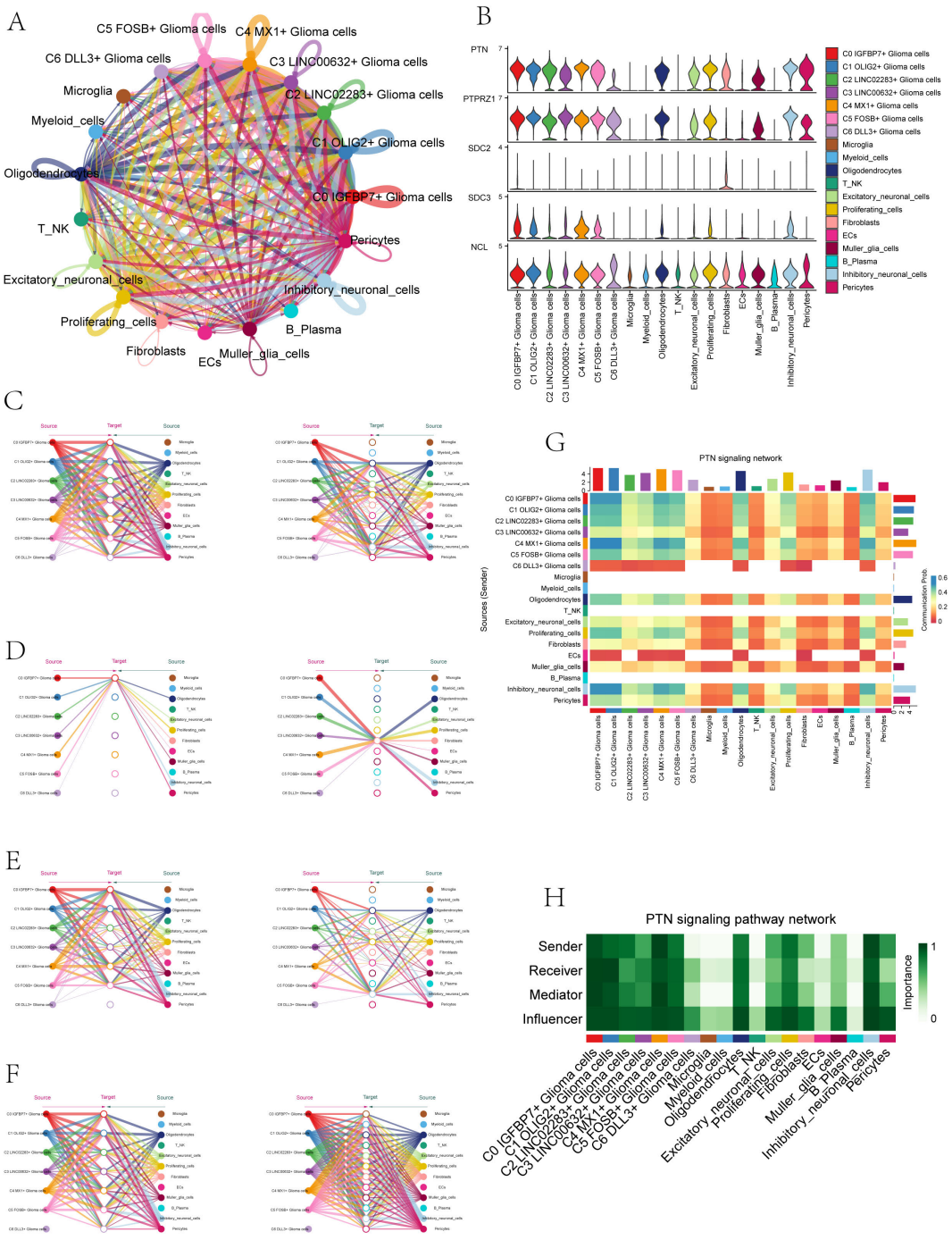


FIGURE 6
PTN signaling network. **(A)** Circle plot illustrated the interactions of astrocytoma in the PTN signaling network as receiver cells. **(B)** Violin plots displayed the levels of expression of signaling genes related to the PTN signaling network in astrocytoma subgroups and various cell types. **(C-F)** Hierarchical plots depicted the communication networks involving PTPRZ1 **(C)**, SDC2 **(D)**, SDC3 **(E)**, and NCL **(F)** in the inferred PTN signaling network. Source cells were represented by filled circles, while target cell types were represented by open circles. **(G)** The heatmap displayed the calculated four centrality metrics of the PTN signaling network, highlighting the significance of each cell type in terms of sending, receiving, mediating, and influencing.

10.3389/fimmu.2024.1434300

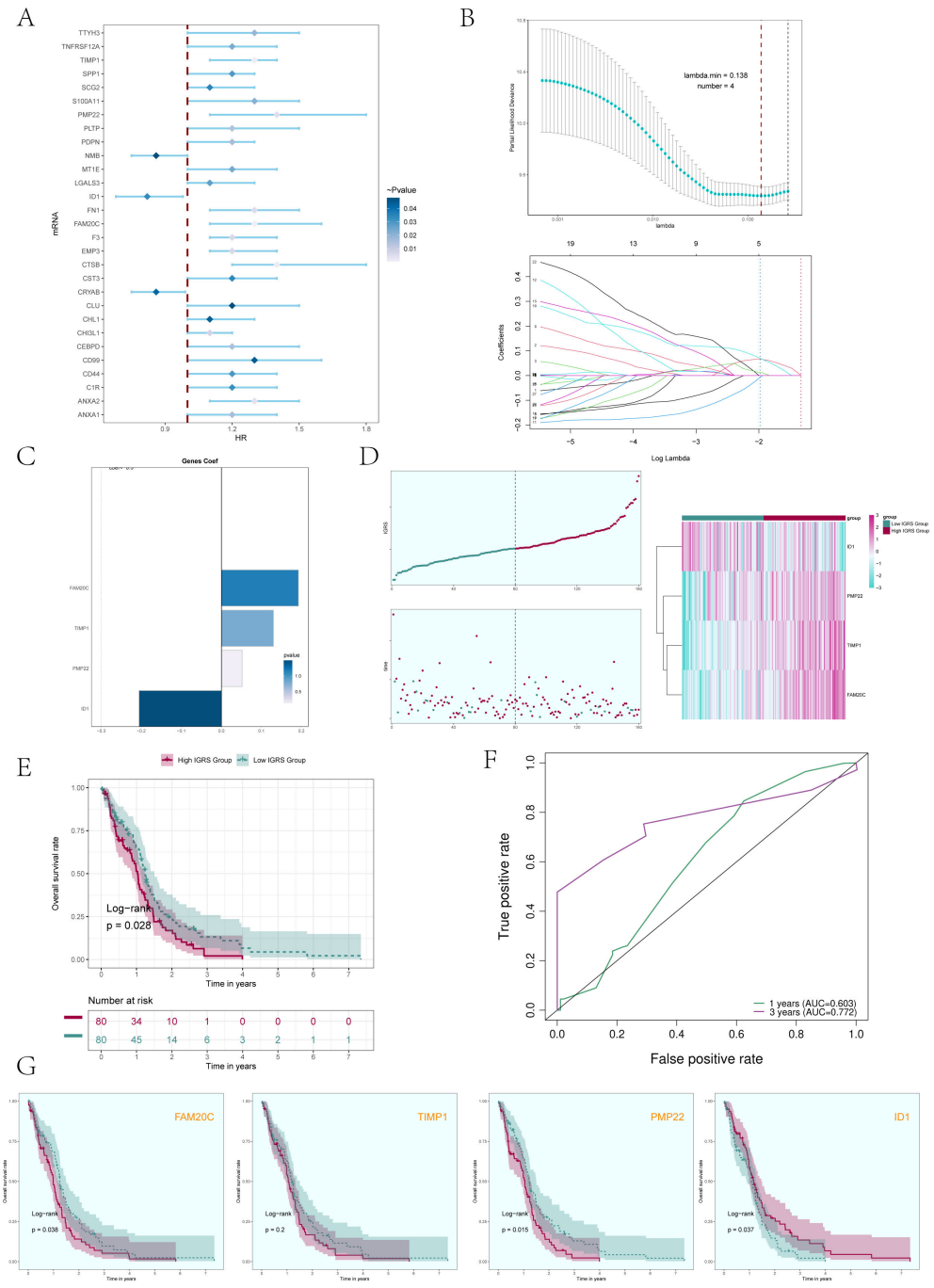


FIGURE 7
Construction and validation of a prognostic risk model. **(A)** Forest plot presented the results of the univariate Cox analysis (P < 0.05). A HR value less than 1 represented protective genes, whereas a HR value greater than 1 represented risk genes. The color depth represents the magnitude of the p-value. **(B)** The results of the LASSO regression analysis indicated that the optimum lambda value was 0.138, yielding the most favorable outcome. Four genes, namely FAM20C, TIMP1, PMP22, and ID1, had been incorporated into the construction of the risk model. **(C)** Bar graph displaying the Coef values and corresponding p-values for the 4 genes. **(D)** C0 subcluster was divided into High IGRS Group and Low IGRS Group based on the IGFBP7 Risk Score (IGRS). The scoring distribution of the C0 subcluster was displayed in the curve plot (top left), while the survival status of the High IGRS and Low IGRS Groups was shown in the scatter plot (bottom left), and the gene expression patterns contributing to the IGRS were visualized in the heatmap. The color green indicated the Low IGRS Group, while the color red indicated the High IGRS Group. **(E)** Kaplan-Meier analysis findings for the High IGRS Group and Low IGRS Group were presented. **(F)** ROC curves showed the AUC of the risk model for predicting survival at 1 and 3 years. **(G)** Survival plots for the four genes associated with prognosis that make up the IGFBP7 Risk Score.

10.3389/fimmu.2024.1434300

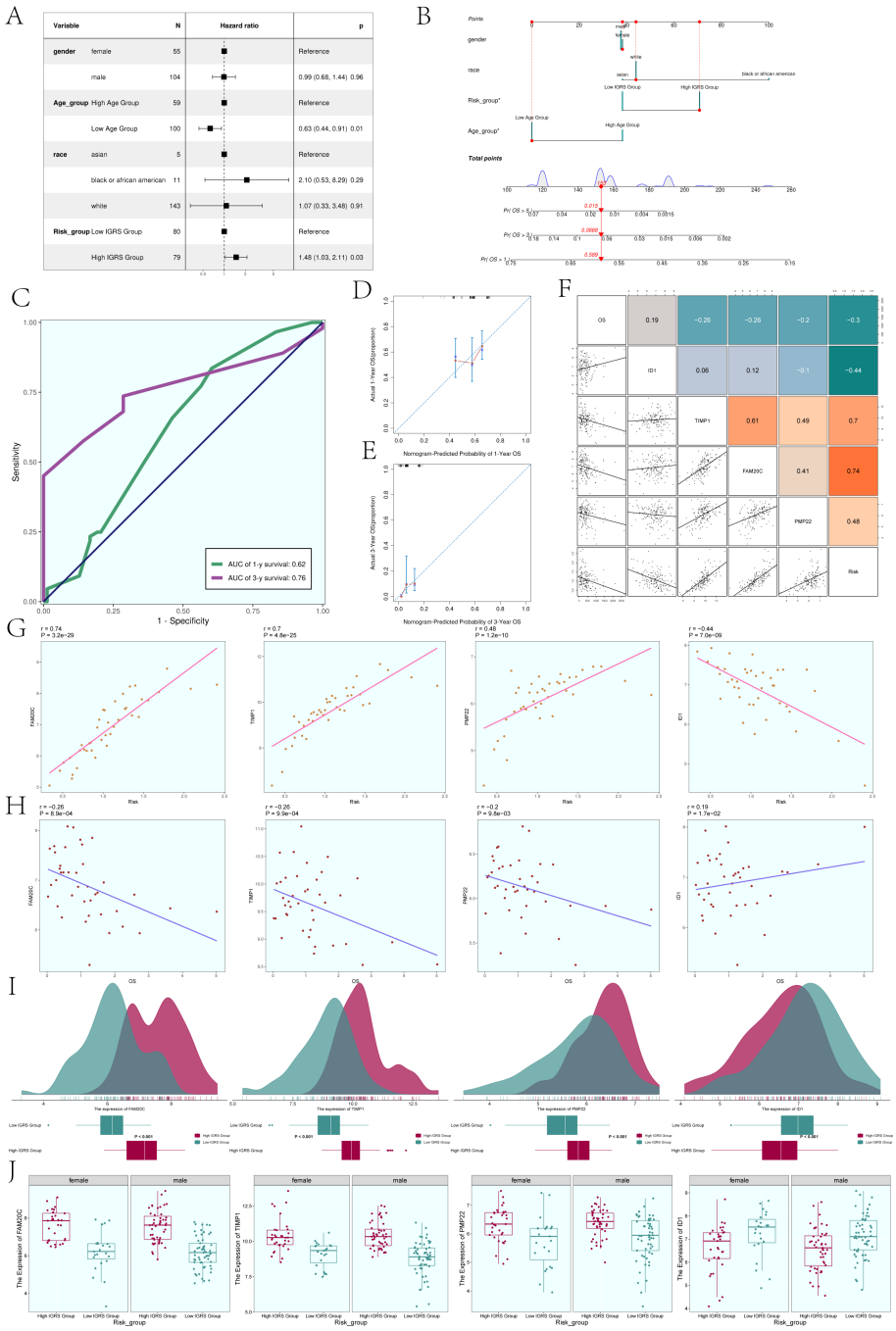


FIGURE 8
Construction of the Nomogram. **(A)** The forest plot displayed the findings from the multivariate Cox regression analysis, showing that age and IGRS score were identified as separate risk factors. **(B)** Nomogram constructed based on clinical factors (gender, race, age) and the IGFBP7 Risk Score. **(C)** The AUC values for 1-year and 3-year predictions were shown on the ROC curve for the nomogram. **(D, E)** Calibration curves were utilized to evaluate the predictive accuracy of the nomogram for both 1-year and 3-year overall survival (OS). **(F)** Scatter plots combined with a heatmap illustrating the correlations between OS, the four modeling genes, and the IGFBP7 Risk Score. **(G)** Scatter plots demonstrated the correlations between the four modeling genes and the IGFBP7 Risk Score. **(H)** Scatter plots showed the correlations between the four modeling genes and OS. **(I)** Ridge plots combined with box plots displaying the expression levels of the four modeling genes in the High IGRS Group and Low IGRS Group, with both groups sharing the same coordinate system. **(J)** Box plots compared the expression levels of the four modeling genes in the High IGRS Group and Low IGRS Group across different genders. Significance levels were denoted as follows: *P < 0.05; NS was used to represent lack of significance.

compared to the gene ID1, which was positively correlated with OS. The expression of the four modeled genes was higher in High IGRS Group and Low IGRS Group in different sexes (female and male), as shown in Figure 8J.

Immunoinfiltration analysis of high IGRS group and low IGRS group

To delve deeper into the tumor microenvironment of glioma, we examined the presence of immune cells infiltrating the tumor in both the High IGRS Group and Low IGRS Group of the training cohort, with the findings displayed in a heatmap (Figure 9A). The statistically different tumor-immune infiltrating cells were further visualized by box plot (Figure 9B), and the evaluation results showed that T cell regulatory (Tregs), Neutrophils, NK cells resting, and Macroghages M1 had higher expression in High IGRS Group, while NK cells activated had higher expression in Low IGRS Group than in High IGRS Group.

To validate the connection between immune cells and IGFBP7 Risk Score in the glioma tumor microenvironment, we assessed the correlation between immune cells and IGRS, presenting the findings through Lollipop plots depicted in Figure 9C. We thoroughly analyzed the relationship between immune cells and the four genes that make up IGRS, IGFBP7 Risk Score, and OS and displayed the findings using a heatmap (Figure 9D). The findings indicated an inverse relationship between IGRS Score and B cells naive, Eosinophils, Master cells Resting, and NK cells activated, while showing a positive correlation with Dendritic cells activated, Monocytes, NK cells Resting, and T cells CD4 memory Resting. It was worth noting that gene TIMP1 and gene FAM20C were negatively correlated with Eosinophils, Master Cells Resting and NK Cells Activated.

Next, we delved deeper into the variations in Stromal Score, Immune Score, Estmate Score, and Tumour Purity between the High IGRS Group and Low IGRS Group, finding statistically significant differences (Figures 9E, F). The Stromal Score, Immune Score, and Estmate Score were elevated in the High IGRS Group, whereas the Tumor Purity was increased in the Low IGRS Group. Nonetheless, there was no statistically significant difference in Tumor Immune Dysfunction and Exclusion (TIDE) between the two groups, suggesting that tumor immune dysfunction and exclusion were similar in both groups (Figure 9G). In the study, it was found that the gene TIMP1 exhibited a strong positive correlation with the majority of immune checkpoint-related genes, while the gene ID1 did not show any significant correlation with most immune checkpoint-related genes (Figure 9H).

Furthermore, we analyzed the variations in expression of immune checkpoint-associated genes between the High IGRS Group and Low IGRS Group, creating box plots to illustrate the genes exhibiting significant differences (Figure 9I). The results of the analysis indicated that the majority of genes associated with immune checkpoints exhibited increased levels of expression in the High IGRS Group, whereas VTCN1 and CD200 displayed higher expression levels in the Low IGRS Group. We used ESTIMATE, CIBERSORT, EPIC, and Xcell algorithms to analyze and display the variations in immune infiltrating cells, Stromal Score, Immune Score, and Tumor Purity between the High IGRS Group and Low IGRS Group in a heatmap. (Figure 9J)

Differentially expressed genes and their enrichment analysis in high and low IGRS groups

To compare the High IGRS Group and Low IGRS Group, we computed and studied the DEGs in both groups, presenting them using a volcano plot (Figure 10A) and showcasing the specific expression of these DEGs in the groups through a heatmap (Figure 10B). Immediately after that, we performed multiple enrichment analyses on these differentially expressed genes. Enrichment analyses were conducted on them, which included examining GOBP (Gene Ontology Biological Processes), GOCC (Gene Ontology Cellular Components), and GOMF (Gene Ontology Molecular Functions). The findings indicated that differentially expressed genes (DEGs) were highly concentrated in functions related to binding between receptors and ligands, signaling pathways mediated by cytokines, and activities involving chemokines (Figure 10C). The related genes of the enriched entries are shown in the chord plot (Figure 10D). The analysis of enriched pathways using KEGG for the identified DEGs (Figure 10E) indicated a significant enrichment in pathways related to viral protein interaction with cytokines and cytokine receptors, interactions between cytokines and cytokine receptors, signaling pathways for chemokines, the IL-17 signaling pathway, and more. According to the findings of GSEA (Gene Set Enrichment Analysis) (Figure 10F), the High IGRS Group exhibited increased activity in pathways related to Neutrophil Chemotaxis, Neutrophil Migration, Granulocyte Chemotaxis, and Granulocyte Migration, while showing decreased activity in pathways associated with Spinal Cord Development, Neurotransmitter Transport, Neuron Fate Specification, Neuron Migration, and Neuron Fate Commitment.

PCA was utilized to examine the diversity of gene expression patterns in the High IGRS Group and the Low IGRS Group, with PCA 1 and PCA 2 visualized through scatter plots. PCA 1 and PCA 2 exhibited variances of 13.2% and 8.1%, respectively, as shown in Figure 10G. Furthermore, we investigated the somatic gene mutations in both cohorts and highlighted the distinctions among the top 30 genes exhibiting the greatest mutation rates in each group. Variations among 12 genes across various groups indicated that the PTEN gene had the highest mutation frequency, as depicted in Figure 10H. Next, we assessed the gene model's chromosome copy number variation (CNV) and presented the findings using a bar graph (Figure 10I). The findings indicated that genes ID1 and TIMP1 did not exhibit any CNV loss or CNV gain, while gene FAM20C experienced both CNV loss and CNV gain events.

A comparison analysis was performed on the two groups' tumor mutation burden (TMB). The results revealed no statistically significant difference in TMB between the two groups (Figure 10J). The correlation analysis between TMB and Risk Score was shown in Figure 10K, with an R value of -0.12 and a corresponding p-value of

FIGURE 9
Immune infiltration analysis. **(A)** A heatmap was used to analyze the distribution of 22 immune infiltrating cells between the High IGRS Group and Low IGRS Group. Red represented the High IGRS Group, and green represented the Low IGRS Group. **(B)** Box plot illustrated the distribution of immune infiltrating cells with statistically significant differences between the High IGRS Group and Low IGRS Group. **(C)** Lollipop charts illustrated the relationship between immune infiltrating cells and IGRS. **(D)** The heatmap offered a comprehensive perspective on the relationships among immune infiltrating cells, the four modeling genes, IGRS, and overall survival. **(E)** The box plot illustrated variations in StromalScore, ImmuneScore, and ESTMATEScore between the High IGRS Group and Low IGRS Group. **(F)** Violin plot demonstrated the variations in Tumor Purity levels between the High IGRS Group and Low IGRS Group. **(G)** Violin plots compared the TIDE values and the differences between the ctla4-negative-pd1-negative and ctla4-positive-pd1-positive subgroups in the High IGRS Group and Low IGRS Group. **(H)** The dot plot illustrated the correlations between OS, the four modeling genes, IGFBP7 Risk Score, and immune checkpoint-associated genes. **(I)** Box plot displayed the expression levels of immune checkpoint-associated genes in the High IGRS Group and Low IGRS Group. **(J)** Heatmap provided a comprehensive display of the results from the ESTIMATE, CIBERSORT, EPIC, and Xcell algorithms. Significance levels were denoted as follows: *P < 0.05, **P < 0.01, ***P < 0.001, and ****P < 0.0001; NS was used to represent lack of significance.

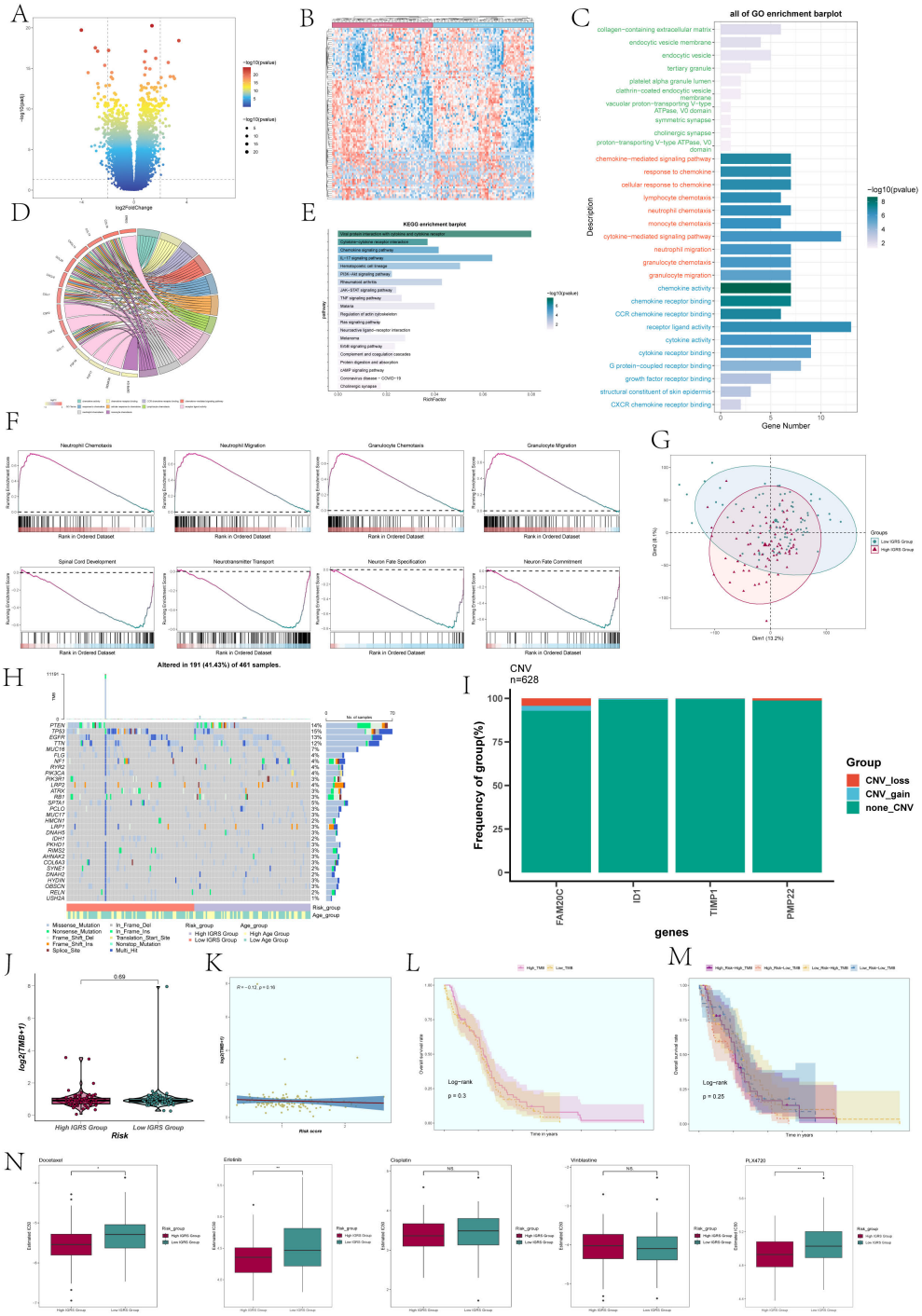


FIGURE 10
Differentially expressed genes and enrichment analysis in High and Low IGRS Groups. **(A)** The volcano plot displayed the genes that were expressed differently between the High IGRS Group and Low IGRS Group. **(B)** Heatmap depicted the distribution of DEGs in the High IGRS Group and Low IGRS Group. **(C)** Bar graph presented the results of the GOBP, GOCC, and GOMF enrichment analyses for the DEGs. **(D)** Chord plot displayed the relevant genes involved in the GO enrichment analysis items. **(E)** The bar graph displayed the findings of the KEGG examination for the differentially expressed genes. **(F)** GSEA enrichment analysis results for the DEGs, displaying the enrichment scores of genes on different pathways. **(G)** Principal Component Analysis (PCA) plot showing the gene expression clustering distribution differences between the High IGRS Group and Low IGRS Group. **(H)** A waterfall chart displayed the 30 most mutated genes in the High IGRS Group and Low IGRS Group. **(I)** Bar graph displayed the copy number variation status of the four modeling genes, with blue indicating chromosomal copy number increase, red indicating chromosomal copy number decrease, and green indicating no change in chromosomal copy number. **(J)** A box plot displayed the Tumor Mutation Burden (TMB) for both the High IGRS Group and Low IGRS Group. **(K)** The scatter plot displayed the relationship between Tumor Mutation Burden and IGFBP7 Risk Score. **(L)** Kaplan-Meier analysis demonstrated variations in prognosis between High TMB and Low TMB groups. **(M)** Kaplan-Meier survival analysis findings for the High Risk-High TMB, High Risk-Low TMB, Low Risk-High TMB, and Low Risk-Low TMB groups. **(N)** Box plots showed the findings of drug response analysis for the High IGRS Group and Low IGRS Group. Significance levels were denoted as follows: *P < 0.05, **P < 0.01, and NS was used to represent lack of significance.

0.16. Using the TMB as a basis, the participants were separated into two groups, High TMB and Low TMB, for examination of survival rates (Figure 10L). Furthermore, the participants were divided into four groups based on their risk level and tumor mutational burden (TMB): High Risk-High TMB, High Risk-Low TMB, Low Risk-High TMB, and Low Risk-Low TMB, which was then followed by an analysis of survival rates. Nevertheless, the findings indicated that there was no notable variation between the groups in terms of statistical significance (Figure 10M).

Drug sensitivity analysis

Analysis of drug sensitivity was performed on the High IGRS Group and Low IGRS Group, showing that Docetaxel had a lower IC50(semi-inhibitory concentration) in the High IGRS Group, as illustrated in Figure 10N.Conversely, PLX4720 demonstrated a lower IC50 value in the Low IGRS Group.

In vitro experimental validation

For further elucidation of the functionality of FOSL2, we conducted *in vitro* functional assessments. Two cell lines, U87 MG and U251 MG, were chosen for comparison with FOSL2 knockdown by establishing a negative control group. The cell activity test (Figures 11A, B) showed a notable reduction in cell viability after FOSL2 knockdown, as revealed by the results of the CCK-8 assay. For accuracy, we quantified the levels of FOSL2 mRNA expression in the U87 MG and U251 MG cell lines in both the control and FOSL2 knockdown groups (Figure 11C). The transwell test findings showed a significant decrease in the movement and infiltration of U87 MG and U251 MG cells following the suppression of FOSL2 in comparison to the control group (Figures 11D, E). Furthermore, the plate cloning results revealed a significant suppression in colony formation quantity after FOSL2 knockdown in both cell line models (Figure 11F).

A healing experiment was performed, revealing a notable increase in the width of the 48-hour scratch in both cell lines after FOSL2 knockdown compared to the negative control group. This suggests a reduction in cell migration rate, supported by statistically significant findings (Figures 12A, B). Additionally, EdU staining once again confirmed the decreased proliferative capacity of tumor cells after FOSL2 knockdown (Figures 12C, D). Thus, from the above tests, it was noted that reducing FOSL2 results in lower cell proliferation, migration, and invasion in U87 MG and U251 MG cell lines, indicating that FOSL2 could enhance glioma advancement.

Discussion

Astrocytoma tumors start in the glial cells called astrocytes. The most aggressive astrocytoma is a glioblastoma. Glioblastomas are the most aggressive and lethal brain tumors (53), being the most aggressive and deadly brain tumor with a high likelihood of recurrence and spreading to other areas of the brain (54). To further investigate the internal heterogeneity of glioma, we analyzed glioma single-cell RNA sequencing (scRNA-seq) data to identify the various cell types present, including microglia, oligodendrocytes, astrocytes, inhibitory neuronal cells, and pericytes, among others. Astrocytoma tumors originate in astrocytes. The most aggressive form of astrocytoma is glioblastoma. Additionally, astrocytes encompass the most abundant cellular entities within the central nervous system (55),so we used astrocytes as the main subpopulation of the study. By inferCNV analysis, we defined high levels of astrocytes as astrocytoma, analyzed them by dimensionality reduction clustering, and finally divided them into seven different cell subpopulations: C0 IGFBP7+ Glioma cells, C1 OLIG2+ Glioma cells, C2 LINC02283+ Glioma cells, C3 LINC00632+ Glioma cells, C4 MX1+ Glioma cells, C5 FOSB+ Glioma cells, and C6 DLL3+ Glioma cells. CytoTRACE and Monocle 2 analyses suggested that C0 IGFBP7+ glioma cells were likely at advanced stages of differentiation with high differentiation potential. Since astrocytomas often showed that higher malignancy could correlate with greater differentiation, identifying these cells might have been crucial. They could provide important insights into tumor progression and resistance, potentially guiding more effective treatments.

In order to delve deeper into the connections between the astrocytoma subcluster and various cell types, we employed CellChat analysis. This tool can deduce and examine intercellular communication networks based on single-cell sequencing data, forecasting the primary signals exchanged between cells and how they work together to carry out their functions (56). By analyzing afferent and efferent signals between subclusters of astrocytoma and other cells, it was found that all 7 subclusters of astrocytoma were involved in the PTN signaling network pathway in both Incoming communication and Outgoing communication. Previous research data has indicated that blocking the PTN pathway may serve as a means to combat glioblastoma (57). Disrupting the PTN receptor PTPRZ1 has been shown to inhibit the growth of glioblastoma stem cells (GSCs) (58). Therefore, we conducted further analysis of the PTN pathway and discovered that PTPRZ1 exhibits high expression in various subclusters of astrocytoma. When PTPRZ1 Glioma cells acted as receivers, the subclusters of astrocytoma showed a strong association with other cell types. Furthermore, in the PTN signaling pathway network, the C0 IGFBP7+ Glioma cells subcluster showed greater importance as a sender, receiver, mediator, and influencer when compared to other types of cells. Therefore, we hypothesized that the C0 subgroup was essential in the PTN pathway and impacted the advancement of glioblastoma via this pathway.

To assess the role of the C0 IGFBP7+ glioma cell subgroup in neuroglioma progression, we performed univariate Cox and LASSO regression analyses on candidate genes, identifying four genes strongly linked to prognosis. We developed a prognosis model based on these genes and established the IGFBP7 Risk Score (IGRS). This score classified the training cohort into High IGRS and Low IGRS groups, with survival analysis showing poorer outcomes for the High IGRS group. A nomogram incorporating clinical data and multivariate Cox regression confirmed the IGRS as a standalone predictor of patient outcomes. Analysis of the four genes revealed

their distribution and correlation with Risk Score and overall survival. In summary, the IGFBP7 Risk Score (IGRS) provided a robust prognostic tool for astrocytomas by categorizing patients into High and Low IGRS groups, with High IGRS correlating with worse outcomes. It integrated gene expression data to offer improved predictions of patient survival and highlighted key genes like FAM20C and PMP22 associated with poor prognosis. FAM20C has been proven to be a marker of glioma invasion and

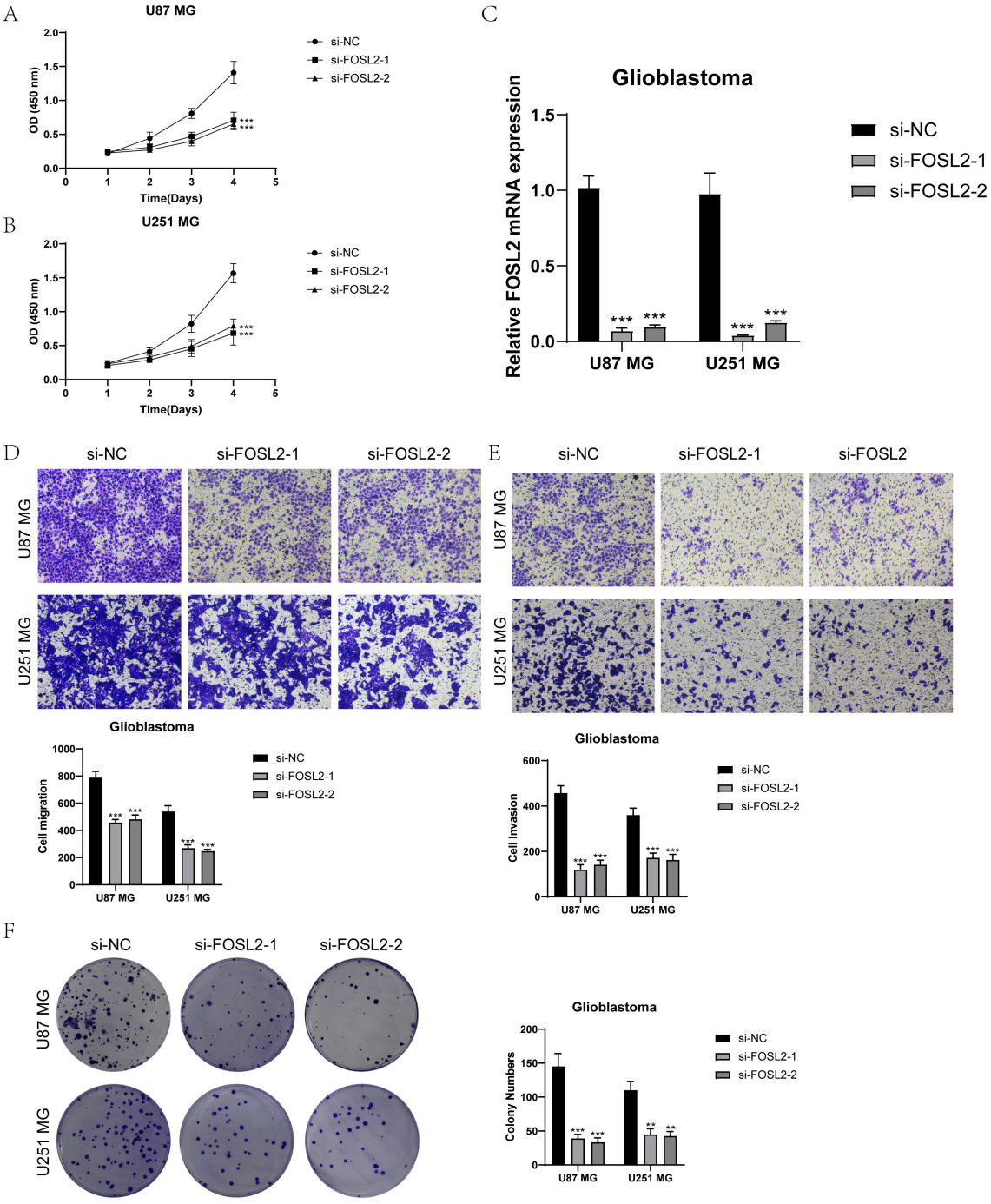


FIGURE 11
In vitro experimental validation. **(A, B)** The CCK-8 assay results showed a notable reduction in cell viability in the U87 MG and U251 MG cell lines following the knockdown of FOSL2. **(C)** The qPCR findings showed the initial levels of FOSL2 mRNA expression in the U87 MG and U251 MG cell lines, as well as the changes in FOSL2 mRNA expression following FOSL2 knockdown. **(D, E)** The transwell test showed that reducing FOSL2 expression greatly hinders the movement and infiltration capabilities of the U87 MG and U251 MG cell lines. **(F)** The plate cloning experiment showed a notable reduction in colony formation capacity in the U87 MG and U251 MG cell lines following the suppression of FOSL2. Significance levels were set at **P < 0.01, and ***P < 0.001.

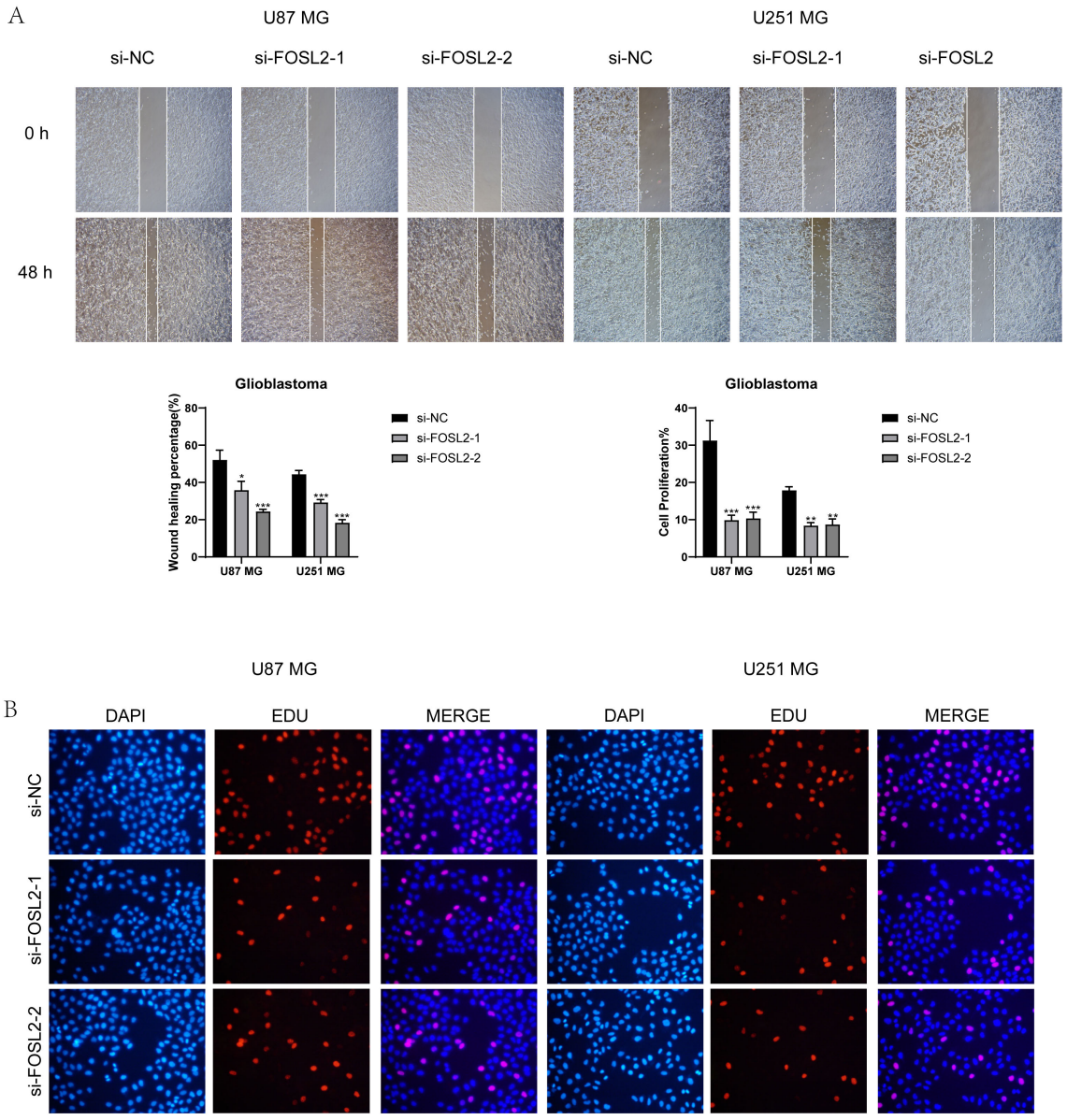


FIGURE 12
Scratch assay and EdU staining results. **(A, B)** The scratch assay showed that FOSL2 knockdown significantly decreased the movement and infiltration of the U87 MG and U251MG cells. **(C, D)** EdU staining demonstrated that FOSL2 knockdown was shown to inhibit the growth of U87 MG and U251MG cells. Significance levels were set at *P < 0.05, **P < 0.01, and ***P < 0.001.

can be used as a new therapeutic target for GBM (59). However, there are few studies on the relationship between PMP22 and glioma, which need to be further explored.

GBM is a highly immunosuppressive tumor. At present, there is no FDA-approved immunotherapy for glioblastoma (60). We further discussed the relationship between IGFBP7 Risk Score (IGRS) and the immune microenvironment of glioma and analyzed the tumor immune infiltration of the two groups based on High IGRS Group and Low IGRS Group. Compared with Low IGRS Group, T cell regulation (tregs), neutrophils, NK cells resting, and macroghages M1 in High IGRS Group have higher expression, while NK cells activated have higher expression in Low IGRS Group, which was related to the relationship between TME and immune cells in tumors (61).

ESTIMATEScore is calculated by adding ImmuneScore and StromalScore, which indicate the presence of immune or matrix components in the TME (62). The Stromal Score, Immune Score, and EstmateScore were higher in the High IGRS Group, which is intriguing. Furthermore, an examination of somatic cell mutation frequency revealed that the PTEN gene exhibited the highest mutation rate, and PTEN could suppress the activation of the PI3K/AKT/mTOR signaling pathway (63). When the functionality of PTEN is disrupted, such as through mutations in the PTEN gene, it leads to the loss of PTEN's tumor suppressor capabilities.

Finally, the IC50 (semi-inhibitory concentration) of PLX4720, Docetaxel, and Erlotinib in different groups was analyzed. Docetaxel, an FDA-approved medication, is now the primary therapy for various cancer forms, such as prostate cancer (64) and non-small cell lung cancer (NSCLC) (65). while Erlotinib, a tyrosine kinase inhibitor, is effective against lung cancer, head and neck squamous cell carcinoma (66, 67), and various other types of cancer. The analysis revealed that Docetaxel and Erlotinib had reduced IC50 values in the High IGRS Group, indicating improved efficacy of these drugs for this patient cohort. Consequently, Docetaxel and Erlotinib demonstrate greater therapeutic potential for patients in the High IGRS Group.

Analysis of the transcription factors in the C0 subgroup revealed that the distribution of the transcription factor FOSL2 of TOP1 in Group IV was greater than in Group II. Hence, we performed *in vitro* tests to support the role of crucial transcription regulators. The findings indicated that suppressing FOSL2 can decrease the growth, movement, and infiltration of U87 MG and U251 MG cells, aligning with the findings of Yiyun Chen and Ranhuo et al. (68). Thus, FOSL2 has the ability to enhance the invasion and advancement of gliomas.

However, there are some limitations to this study. First of all, the sample size is small, and the number of patients with glioma obtained in this study is limited. Secondly, we have only done scrna-seq and bulk RNA-seq analyses and *in vitro* experiments, and we need large sample and multi-center research to further explore the relationship between IGFBP7, FOSL2, the IGFBP7 Risk Score (IGRS), and glioma. Therefore, we plan to carry out various analytical methods, such as metabonomics and ATAC-seq, to demonstrate in many aspects.

Nevertheless, there are certain constraints to this research. First of all, the sample size is small, and the number of patients with glioma obtained in this study is limited. Additionally, our research has been limited to scRNA-Seq and bulk RNA-seq analyses along with *in vitro* experiments. To delve deeper into the connection between IGFBP7, FOSL2, the IGFBP7 Risk Score (IGRS), and glioma, we require extensive sample sizes and collaboration with multiple research centers. Therefore, we plan to carry out various analytical methods, such as metabonomics and ATAC-seq, to demonstrate this in many aspects.

Conclusion

Our exploration of the astrocyte tumor microenvironment highlighted the critical role of the C0 IGFBP7+ glioma subpopulation in astrocytoma progression. We developed the IGFBP7 Risk Score (IGRS) as an independent prognostic tool that effectively separates High and Low IGRS groups, with High IGRS indicating worse outcomes. The IGRS not only predicts patient survival but also identifies key genes like FAM20C and PMP22 linked to poor prognosis. Our study also pinpointed new therapeutic targets, showing that Docetaxel and Erlotinib are more effective in the High IGRS group. Additionally, *in vitro* tests confirmed that transcription regulators like FOSL2 enhance glioma invasion and progression. These insights improve our understanding of astrocytoma and offer promising avenues for future treatments.

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/s.

Ethics statement

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

Author contributions

LZ: Data curation, Methodology, Visualization, Writing – original draft. WS: Data curation, Methodology, Visualization, Writing – original draft. ZX: Validation, Writing – review & editing. LR: Supervision, Writing – review & editing. CL: Data curation, Writing – review & editing. YS: Methodology, Writing – review & editing. HX: Writing – review & editing, Supervision. ZW: Investigation, Writing – review & editing. JX: Funding acquisition, Resources, Writing – original draft, Writing – review & editing, Supervision.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research was financially supported by the Science and Technology Development Fund of Shanghai Pudong New Area under Grant No. PKJ2023-Y07.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1434300/full#supplementary-material.

SUPPLEMENTARY FIGURE 1
Visualization of large groups of Glioma. **(A)** The UMAP visualization displayed how the different cell types were distributed among 234,148 high-quality glioma cells. **(B-D)** UMAP visualizations, along with pie graphs, illustrated how 234,148 high-grade glioma cells were spread out among Categories (II, III, and IV) **(B)**, Phases (G1, S, and G2M) **(C)**, and 45 seurat clusters **(D)**. **(E)** The bubble chart displayed the top five genes for the 13 types of cells and illustrated how these genes were distributed among the various groups.

SUPPLEMENTARY FIGURE 2
Analysis of astrocytes through inferCNV. **(A)** The inferCNV analysis of astrocytes was visualized in a heatmap. Astrocytes with high levels of inferCNV were defined as astrocytoma. The red color represented high copy number variation (astrocytoma), while the blue color represented low copy number variation.

References

1. Ostrom QT, Cioffi G, Gittleman H, Patil N, Waite K, Kruchko C, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2012-2016. *Neuro-Oncology*. (2019) 21:v1–v100. doi: 10.1093/neuonc/noz150

2. Ghiaseddin AP, Shin D, Melnick K, Tran DD. Tumor treating fields in the management of patients with Malignant gliomas. *Curr Treat Option On*. (2020) 21:76. doi: 10.1007/s11864-020-00773-5

3. GBD 2016 Brain and Other CNS Cancer Collaborators. Global, regional, and national burden of brain and other CNS cancer 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol*. (2019) 18:376–93.

4. Weller M, Wick W, Aldape K, Brada M, Berger M, Pfister SM, et al. Glioma. *Nat Rev Dis Primers*. (2015) 1:15017. doi: 10.1038/nrdp.2015.17

5. Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, Burger PC, Jouvet A, et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol*. (2007) 114:97–109. doi: 10.1007/s00401-007-0243-4

6. Sengupta S, Marrinan J, Frishman C, Sampath P. Impact of temozolomide on immune response during Malignant glioma chemotherapy. *Clin Dev Immunol*. (2012) 2012:831090. doi: 10.1155/2012/831090

7. Hambardzumyan D, Bergers G. Glioblastoma: defining tumor niches. *Trends Cancer*. (2015) 1:252–65. doi: 10.1016/j.trecan.2015.10.009

8. Stupp R, Taillibert S, Kanner A, Read W, Steinberg D, Lhermitte B, et al. Effect of tumor-treating fields plus maintenance temozolomide vs maintenance temozolomide alone on survival in patients with glioblastoma: A randomized clinical trial. *JAMA-J Am Med Assoc*. (2017) 318:2306–16. doi: 10.1001/jama.2017.18718

9. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *New Engl J Med*. (2005) 352:987–96. doi: 10.1056/NEJMoa043330

10. Han X, Zhou Z, Fei L, Sun H, Wang R, Chen Y, et al. Construction of a human cell landscape at single-cell level. *Nature*. (2020) 581:303–9. doi: 10.1038/s41586-020-2157-4

11. Ma B, Qin L, Sun Z, Wang J, Tran LJ, Zhang J, et al. The single-cell evolution trajectory presented different hypoxia heterogeneity to reveal the carcinogenesis of genes in clear cell renal cell carcinoma: Based on multiple omics and real experimental verification. *Environ Toxicol*. (2024) 39:869–81. doi: 10.1002/tox.24009

12. Wu F, Fan J, He Y, Xiong A, Yu J, Li Y, et al. Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. *Nat Commun*. (2021) 12:2540. doi: 10.1038/s41467-021-22801-0

13. Zhao Z, Ding Y, Tran LJ, Chai G, Lin L. Innovative breakthroughs facilitated by single-cell multi-omics: manipulating natural killer cell functionality correlates with a novel subcategory of melanoma cells. *Front Immunol*. (2023) 14:1196892. doi: 10.3389/fimmu.2023.1196892

14. Lin Z, Li X, Shi H, Cao R, Zhu L, Dang C, et al. Decoding the tumor microenvironment and molecular mechanism: unraveling cervical cancer subpopulations and prognostic signatures through scRNA-Seq and bulk RNA-seq analyses. *Front Immunol*. (2024) 15:1351287. doi: 10.3389/fimmu.2024.1351287

15. Wang J, Zuo Z, Yu Z, Chen Z, Tran LJ, Zhang J, et al. Collaborating single-cell and bulk RNA sequencing for comprehensive characterization of the intratumor heterogeneity and prognostic model development for bladder cancer. *Aging (Albany Ny)*. (2023) 15:12104–19. doi: 10.18632/aging.205166

16. Sun Z, Wang J, Zhang Q, Meng X, Ma Z, Niu J, et al. Coordinating single-cell and bulk RNA-seq in deciphering the intratumoral immune landscape and prognostic stratification of prostate cancer patients. *Environ Toxicol*. (2024) 39:657–68. doi: 10.1002/tox.23928

17. Zhou W, Lin Z, Tan W. Deciphering the molecular landscape: integrating single-cell transcriptomics to unravel myofibroblast dynamics and therapeutic targets in clear cell renal cell carcinomas. *Front Immunol*. (2024) 15:1374931. doi: 10.3389/fimmu.2024.1374931

18. Wang J, Zuo Z, Yu Z, Chen Z, Meng X, Ma Z, et al. Single-cell transcriptome analysis revealing the intratumoral heterogeneity of ccRCC and validation of MT2A in pathogenesis. *Funct Integr Genomic*. (2023) 23:300. doi: 10.1007/s10142-023-01225-7

19. Zhu C, Sun Z, Wang J, Meng X, Ma Z, Guo R, et al. Exploring oncogenes for renal clear cell carcinoma based on G protein-coupled receptor-associated genes. *Discovery Oncol*. (2023) 14:182. doi: 10.1007/s12672-023-00795-z

20. Sun Z, Wang J, Fan Z, Yang Y, Meng X, Ma Z, et al. Investigating the prognostic role of lncRNAs associated with disulfidptosis-related genes in clear cell renal cell carcinoma. *J Gene Med*. (2024) 26:e3608. doi: 10.1002/jgm.v26.1

21. Mega A, Hartmark NM, Leiss LW, Tobin NP, Miletic H, Sleire L, et al. Astrocytes enhance glioblastoma growth. *Glia*. (2020) 68:316–27. doi: 10.1002/glia.23718

22. Perelroizen R, Philosof B, Budick-Harmelin N, Chernobylsky T, Ron A, Katzir R, et al. Astrocyte immunometabolic regulation of the tumour microenvironment drives glioblastoma pathogenicity. *Brain*. (2022) 145:3288–307. doi: 10.1093/brain/awac222

23. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WR, et al. Comprehensive integration of single-cell data. *Cell*. (2019) 177:1888–902. doi: 10.1016/j.cell.2019.05.031

24. Shao W, Lin Z, Xiahou Z, Zhao F, Xu J, Liu X, et al. Single-cell RNA sequencing reveals that MYBL2 in Malignant epithelial cells is involved in the development and progression of ovarian cancer. *Front Immunol*. (2024) 15:1438198. doi: 10.3389/fimmu.2024.1438198

25. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst*. (2019) 8:329–37. doi: 10.1016/j.cels.2019.03.003

26. Jin W, Zhang Y, Zhao Z, Gao M. Developing targeted therapies for neuroblastoma by dissecting the effects of metabolic reprogramming on tumor microenvironments and progression. *Theranostics*. (2024) 14:3439–69. doi: 10.7150/thno.93962

27. Nie W, Zhao Z, Liu Y, Wang Y, Zhang J, Hu Y, et al. Integrative single-cell analysis of cardiomyopathy identifies differences in cell stemness and transcriptional regulatory networks among fibroblast subpopulations. *Cardiol Res Pract*. (2024) 2024:3131633. doi: 10.1155/2024/3131633

28. Huang W, Kim BS, Zhang Y, Lin L, Chai G, Zhao Z. Regulatory T cells subgroups in the tumor microenvironment cannot be overlooked: Their involvement in prognosis and treatment strategy in melanoma. *Environ Toxicol*. (2024). doi: 10.1002/tox.v39.10

29. Ding Y, Zhao Z, Cai H, Zhou Y, Chen H, Bai Y, et al. Single-cell sequencing analysis related to sphingolipid metabolism guides immunotherapy and prognosis of skin cutaneous melanoma. *Front Immunol*. (2023) 14:1304466. doi: 10.3389/fimmu.2023.1304466

30. Liu P, Xing N, Xiahou Z, Yan J, Lin Z, Zhang J. Unraveling the intricacies of glioblastoma progression and recurrence: insights into the role of NFYB and oxidative phosphorylation at the single-cell level. *Front Immunol*. (2024) 15:1368685. doi: 10.3389/fimmu.2024.1368685

31. Ge Q, Zhao Z, Li X, Yang F, Zhang M, Hao Z, et al. Deciphering the suppressive immune microenvironment of prostate cancer based on CD4+ regulatory T cells: Implications for prognosis and therapy prediction. *Clin Transl Med*. (2024) 14:e1552. doi: 10.1002/ctm2.v14.1

32. Lin Z, Sui X, Jiao W, Chen C, Zhang X, Zhao J. Mechanism investigation and experiment validation of capsaicin on uterine corpus endometrial carcinoma. *Front Pharmacol*. (2022) 13:953874. doi: 10.3389/fphar.2022.953874

33. Xing J, Cai H, Lin Z, Zhao L, Xu H, Song Y, et al. Examining the function of macrophage oxidative stress response and immune system in glioblastoma multiforme through analysis of single-cell transcriptomics. *Front Immunol*. (2023) 14:1288137. doi: 10.3389/fimmu.2023.1288137

34. Zhao J, Jiao W, Sui X, Zou J, Wang J, Lin Z. Construction of a prognostic model of luteolin for endometrial carcinoma. *Am J Transl Res*. (2023) 15:2122–39.

35. Lin Z, Fan W, Yu X, Liu J, Liu P. Research into the mechanism of intervention of SanQi in endometriosis based on network pharmacology and molecular docking technology. *Medicine*. (2022) 101:e30021. doi: 10.1097/MD.0000000000030021

36. Zhao ZJ, Zheng RZ, Wang XJ, Li TQ, Dong XH, Zhao CY, et al. Integrating lipidomics and transcriptomics reveals the crosstalk between oxidative stress and neuroinflammation in central nervous system demyelination. *Front Aging Neurosci*. (2022) 14:870957. doi: 10.3389/fnagi.2022.870957

37. Li XY, Zhao ZJ, Wang JB, Shao YH, Hui-Liu, You JX, et al. m7G methylation-related genes as biomarkers for predicting overall survival outcomes for hepatocellular carcinoma. *Front Bioeng Biotech*. (2022) 10:849756. doi: 10.3389/fbioe.2022.849756

38. Gulati GS, Sikandar SS, Wesche DJ, Manjunath A, Bharadwaj A, Berger MJ, et al. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science*. (2020) 367:405–11. doi: 10.1126/science.aax0249

39. Zhang Y, Liu T, Hu X, Wang M, Wang J, Zou B, et al. CellCall: integrating paired ligand-receptor and transcription factor activities for cell-cell communication. *Nucleic Acids Res*. (2021) 49:8520–34. doi: 10.1093/nar/gkab638

40. Lin Z, Sui X, Jiao W, Wang Y, Zhao J. Exploring the mechanism and experimental verification of puerarin in the treatment of endometrial carcinoma based on network pharmacology and bioinformatics analysis. *BMC Complement Med*. (2022) 22:150. doi: 10.1186/s12906-022-03623-z

41. Zou J, Lin Z, Jiao W, Chen J, Lin L, Zhang F, et al. A multi-omics-based investigation of the prognostic and immunological impact of necroptosis-related mRNA in patients with cervical squamous carcinoma and adenocarcinoma. *Sci Rep-Uk*. (2022) 12:16773. doi: 10.1038/s41598-022-20566-0

42. Wang Y, Zhao ZJ, Kang XR, Bian T, Shen ZM, Jiang Y, et al. lncRNA DLEU2 acts as a miR-181a sponge to regulate SEPP1 and inhibit skeletal muscle differentiation and regeneration. *Aging (Albany Ny)*. (2020) 12:24033–56. doi: 10.18632/aging.104095

43. Zheng RZ, Zhao ZJ, Yang XT, Jiang SW, Li YD, Li WJ, et al. Initial CT-based radiomics nomogram for predicting in-hospital mortality in patients with traumatic brain injury: a multicenter development and validation study. *Neurol Sci*. (2022) 43:4363–72. doi: 10.1007/s10072-022-05954-8

44. Zheng R, Zhuang Z, Zhao C, Zhao Z, Yang X, Zhou Y, et al. Chinese admission warning strategy for predicting the hospital discharge outcome in patients with traumatic brain injury. *J Clin Med*. (2022) 11(4):974. doi: 10.3390/jcm11040974

45. Lin Z, Zou J, Sui X, Yao S, Lin L, Wang J, et al. Necroptosis-related lncRNA signature predicts prognosis and immune response for cervical squamous cell carcinoma and endocervical adenocarcinomas. *Sci Rep-Uk*. (2022) 12:16285. doi: 10.1038/s41598-022-20858-5

46. Zhao J, Zou J, Jiao W, Lin L, Wang J, Lin Z. Construction of N-7 methylguanine-related mRNA prognostic model in uterine corpus endometrial carcinoma based on multi-omics data and immune-related analysis. *Sci Rep-Uk*. (2022) 12:18813. doi: 10.1038/s41598-022-22879-6

47. Lin Z, Fan W, Sui X, Wang J, Zhao J. Necroptosis-related lncRNA signatures for prognostic prediction in uterine corpora endometrial cancer. *Reprod Sci*. (2023) 30:576–89. doi: 10.1007/s43032-022-01023-9

48. Zhao ZJ, Chen D, Zhou LY, Sun ZL, Wang BC, Feng DF. Prognostic value of different computed tomography scoring systems in patients with severe traumatic brain injury undergoing decompressive craniectomy. *J Comput Assist Tomo*. (2022) 46:800–7. doi: 10.1097/RCT.0000000000001343

49. 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

50. Zhao Z, Li T, Dong X, Wang X, Zhang Z, Zhao C, et al. Untargeted metabolomic profiling of cuprizone-induced demyelination in mouse corpus callosum by UPLC-orbitrap/MS reveals potential metabolic biomarkers of CNS demyelination disorders. *Oxid Med Cell Longev*. (2021) 2021:7093844. doi: 10.1155/2021/7093844

51. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. (2018) 28:1747–56. doi: 10.1101/gr.239244.118

52. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PloS One*. (2014) 9:e107468. doi: 10.1371/journal.pone.0107468

53. Majercikova Z, Dibdiakova K, Gala M, Horvath D, Murin R, Zoldak G, et al. Different approaches for the profiling of cancer pathway-related genes in glioblastoma cells. *Int J Mol Sci*. (2022) 23(18):10883. doi: 10.3390/ijms231810883

54. Velasquez C, Mansouri S, Mora C, Nassiri F, Suppiah S, Martino J, et al. Molecular and clinical insights into the invasive capacity of glioblastoma cells. *J Oncol*. (2019) 2019:1740763. doi: 10.1155/2019/1740763

55. Brix F, Jensen JM, Hebbinghaus D. Presentation of a new partial-field blocker for breast-preserving therapy of breast carcinoma. *Strahlenther Onkol*. (1987) 163:314–20.

56. 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

57. Wang Z, Zhong Z, Jiang Z, Chen Z, Chen Y, Xu Y. A novel prognostic 7-methylguanosine signature reflects immune microenvironment and alternative splicing in glioma based on multi-omics analysis. *Front Cell Dev Biol*. (2022) 10:902394. doi: 10.3389/fcell.2022.902394

58. Shi Y, Ping YF, Zhou W, He ZC, Chen C, Bian BS, et al. Tumour-associated macrophages secrete pleiotrophin to promote PTPRZ1 signalling in glioblastoma stem cells for tumour growth. *Nat Commun*. (2017) 8:15080. doi: 10.1038/ncomms15080

59. Du S, Guan S, Zhu C, Guo Q, Cao J, Guan G, et al. Secretory pathway kinase FAM20C, a marker for glioma invasion and Malignancy, predicts poor prognosis of glioma. *Oncotargets Ther*. (2020) 13:11755–68.

60. McGranahan T, Therkelsen KE, Ahmad S, Nagpal S. Current state of immunotherapy for treatment of glioblastoma. *Curr Treat Option On*. (2019) 20:24. doi: 10.1007/s11864-019-0619-4

61. Pena-Romero AC, Orenes-Pinero E. Dual effect of immune cells within tumour microenvironment: pro- and anti-tumour effects and their triggers. *Cancers*. (2022) 14(7):1681. doi: 10.3390/cancers14071681

62. Bi KW, Wei XG, Qin XX, Li B. BTK has potential to be a prognostic factor for lung adenocarcinoma and an indicator for tumor microenvironment remodeling: A study based on TCGA data mining. *Front Oncol*. (2020) 10:424. doi: 10.3389/fonc.2020.00424

63. Alvarez-Garcia V, Tawil Y, Wise HM, Leslie NR. Mechanisms of PTEN loss in cancer: It's all about diversity. *Semin Cancer Biol*. (2019) 59:66–79.

64. James ND, Sydes MR, Clarke NW, Mason MD, Dearnaley DP, Spears MR, et al. Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. *Lancet*. (2016) 387:1163–77. doi: 10.1016/S0140-6736(15)01037-5

65. Wu YL, Lu S, Cheng Y, Zhou C, Wang J, Mok T, et al. Nivolumab versus docetaxel in a predominantly chinese patient population with previously treated advanced NSCLC: checkMate 078 randomized phase III clinical trial. *J Thorac Oncol*. (2019) 14:867–75. doi: 10.1016/j.jtho.2019.01.006

66. Fatani WK, Aleanizy FS, Alqahtani FY, Alanazi MM, Aldossari AA, Shakeel F, et al. Erlotinib-loaded dendrimer nanocomposites as a targeted lung cancer chemotherapy. *Molecules*. (2023) 28(9):3974. doi: 10.3390/molecules28093974

67. Thomas F, Rochaix P, White-Koning M, Hennebelle I, Sarini J, Benlyazid A, et al. Population pharmacokinetics of erlotinib and its pharmacokinetic/pharmacodynamic relationships in head and neck squamous cell carcinoma. *Eur J Cancer*. (2009) 45:2316–23. doi: 10.1016/j.ejca.2009.05.007

68. Chen Y, Huo R, Kang W, Liu Y, Zhao Z, Fu W, et al. Tumor-associated monocytes promote mesenchymal transformation through EGFR signaling in glioma. *Cell Rep Med*. (2023) 4:101177. doi: 10.1016/j.xcrm.2023.101177

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 02 October 2024
DOI 10.3389/fonc.2024.1473515

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Chen Li,
Free University of Berlin, Germany
Muhammad Mansoor Majeed,
AIDM, Pakistan

*CORRESPONDENCE
Chenwei Li
nblcw2008@126.com

RECEIVED 31 July 2024
ACCEPTED 09 September 2024
PUBLISHED 02 October 2024

CITATION
Li K, Tan SC, Yang Z and Li C (2024) FAS gene expression, prognostic significance and molecular interactions in lung cancer.
Front. Oncol. 14:1473515.
doi: 10.3389/fonc.2024.1473515

FAS gene expression, prognostic significance and molecular interactions in lung cancer

Kaimin Li[1], Shing Cheng Tan[2], Zhihao Yang[1] and Chenwei Li[1]*

[1]Department of Thoracic Surgery, The First Affiliated Hospital of Ningbo University, Ningbo, Zhejiang, China, [2]UKM Medical Molecular Biology Institute, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia

Introduction: FAS has been implicated in the development of various cancers, but its involvement in lung cancer has not been systematically characterized. In this study, we performed data mining in online tumor databases to investigate the expression, methylation, alterations, protein interactions, co-expression and prognostic significance of FAS in lung cancer.

Method: The expression, prognostic significance and molecular interactions of FAS in lung cancer was mined and analyzed using GENT2, GEPIA2, UALCAN, cBioPortal, STRING, GeneMANIA, UCSC Xena, Enrichr, and OSluca databases. FAS expression was subsequently investigated at the protein level in samples from 578 lung cancer patients to understand its protein-level expression. *In vitro* validation of FAS gene expression was performed on H1299, H1993, A549 and HBE cell lines.

Result: We found that the expression of FAS was significantly downregulated in both lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) compared to normal lung tissue. In addition, we observed a higher level of FAS promoter methylation in LUSC tissue than in normal tissue. FAS alterations were rare (1.9%) in lung cancer samples, with deep deletions being more common than missense mutations, which occurred mainly in the TNFR-like cysteine-rich domain and the death domain. We also identified a list of proteins interacting with FAS and genes co-expressed with FAS, with LUAD having 11 co-expressed genes and LUSC having 90 co-expressed genes. Our results also showed that FAS expression has limited prognostic significance (HR=1.302, 95% CI=0.935-1.139, P=0.530). Protein level investigation revealed that FAS expression varied among individuals, with nTPM values ranging from 5.2 to 67.2.

Conclusion: This study provides valuable insights into the involvements and characteristics of FAS in lung cancer. Further studies are needed to investigate the clinical significance of FAS alterations in lung cancer and to explore the potential of targeting FAS for therapeutic intervention.

KEYWORDS

apoptosis, bioinformatics, data mining, FAS, *in silico*, lung carcinoma

1 Introduction

Lung cancer is the leading cause of cancer deaths worldwide. In 2020 alone, lung cancer was responsible for about 2.2 million new cases and nearly 1.8 million deaths (1). The cancer is more common in men than in women, accounting for 14.3% and 11.4% of all new cancer cases, respectively. Lung cancer also has a low 5-year survival rate of about 10-20%, due in part to the fact that the disease is often detected at an advanced stage, and accounts for nearly one-fifth of all cancer deaths (1, 2). The number of new lung cancer cases and deaths is expected to continue to increase over the next 15 years, continuing the upward trend in lung cancer incidence and mortality (3). An aging population and continued tobacco use in many regions of the world are predicted to be the main causes of this increase (4).

There are two main histologic types of lung cancer, namely small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) (5). SCLC is a rare but aggressive form of lung cancer. On the other hand, NSCLC is the most common form of lung cancer, accounting for about 80-85% of all cases (5). NSCLC includes more subtypes than SCLC, including lung adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC), which arises in glandular and squamous cells, respectively. However, the role of FAS in lung cancer subtypes, particularly LUAD and LUSC, remains poorly understood concerning its gene expression, mutational landscape, and clinical relevance.

Like other cancers, lung cancer exhibits several hallmarks that are commonly associated with the disease (6, 7). One of these hallmarks is the evasion of apoptosis, a mechanism in which the Fas receptor plays an important role. The transmembrane receptor belongs to the TNF receptor (TNFR) superfamily, consists of 319 amino acids and has a size of about 48 kDa. The protein consists of a TNFR-like domain at its N-terminus, which is rich in cysteines and necessary for interaction with Fas ligand (FasL) (8). It also contains a death domain near the carboxyl terminus that is essential for interaction with FasL. When FasL binds to Fas receptor homotrimers, the receptor is activated and recruits the adaptor protein, Fas-associated death domain (FADD), which in turn recruits procaspase-8 to form the death-inducing signaling complex (DISC) (9). Procaspase-8 is cleaved in the DISC into the active caspase-8, which then triggers activation of the caspase cascade, leading to cellular apoptosis.

In addition to apoptosis, the Fas/FasL pathway is also involved in the initiation of other cellular responses. These include maintenance of immune homeostasis, cell migration, and control of cancer cell invasiveness through regulation of mitogen-activated protein kinase and nuclear factor kappa B activation (10, 11). Apart from that, the Fas signaling pathway has been shown to drive cancer stemness through various mechanisms, such as activation of the ERK-JAG1 axis and the type I interferon/STAT1 axis (12, 13). While these studies have implicated *FAS* in cancer progression and response to therapy, studies exploring *FAS* gene expression and its potential role as a prognostic biomarker in lung cancer, particularly in NSCLC, are limited.

Given the important role of the Fas receptor in various aspects of oncogenesis, variations in the FAS gene have been shown to influence the risk and prognosis of many cancers (14–19). These effects are thought to be due to differential expression of *FAS* and its co-expressed genes in tumor cells compared with normal cells. However, the expression, prognostic significance and molecular interactions of *FAS* in lung cancer have not been systematically studied. We hypothesize that *FAS* gene expression is significantly altered in lung cancer compared to normal tissue and that genetic alterations, such as promoter methylation and mutations, could influence disease progression. However, the exact alterations are often not well-understood. For example, while FAS downregulation has been reported in lung cancer, there are also studies that show that can promote lung cancer growth *in vivo* (20, 21). The availability of genetic data in online tumor databases could provide useful information on the characteristics and potential role of *FAS* gene expression as a prognostic biomarker in cancers. Further, the potential modulation of *FAS* expression by patient characteristics such as age and smoking status in lung cancer context requires more detailed investigation. The clinical implications of understanding *FAS* expression in lung cancer are significant. If *FAS* expression and alterations are shown to have a prognostic impact, this could inform the development of new therapeutic approaches targeting the *FAS* pathway. Therefore, in this study, we performed data mining in online tumor databases to better understand the expression, prognostic significance and molecular interactions of *FAS* in lung cancer. This study aims to better understand the expression patterns, promoter methylation, genomic alterations, and potential protein-protein interactions of *FAS* in lung cancer, which could provide new insights into its role as a prognostic biomarker.

2 Materials and methods

2.1 Sample selection and preprocessing

All data used in this study were taken directly from publicly available databases and no additional pre-processing steps were performed. Sample selection criteria (i.e., inclusion of cancerous or normal tissue) were provided by the respective databases (e.g., TCGA, GEO, and Human Protein Atlas) and included data that passed the quality controls provided by the curators of the databases. We did not apply any specific inclusion or exclusion criteria beyond those specified in the databases. As all samples analyzed were from human subjects, biological replicates were not included in the analysis.

2.2 Gene expression analysis

The mRNA expression of FAS in human cancers was studied using GENT2 (http://gent2.appex.kr/gent2/), which extracts microarray data from the NCBI GEO database (22). GENT2 compiles gene expression profiles across a wide range of cancer types, allowing for the exploration of differential expression patterns in large datasets. Its strengths lie in its large sample size and robust statistical processing, but it is limited by the dependency on

microarray data, which can be subject to batch effects and platform-specific biases. Subsequently, differential expression of FAS between lung cancer and normal lung tissue was examined using GEPIA2 (http://gepia2.cancer-pku.cn/), which extracts data from The Cancer Genome Atlas (TCGA) (23). GEPIA2 is a web-based tool specifically designed for cancer gene expression profiling and survival analysis based on RNA-Seq data from TCGA and GTEx, offering statistical significance testing through analysis of variance (ANOVA) and t-tests with FDR correction to control for multiple comparisons. Subgroup analysis of the TCGA expression data in different clinicopathological features was performed using UALCAN (http://ualcan.path.uab.edu) (24). UALCAN is a comprehensive, user-friendly platform for investigating cancer omics data with a focus on subgroup analysis based on clinicopathological parameters, using data from TCGA. The Human Protein Atlas database (https://www.proteinatlas.org/) was used to examine gene expression at the protein level through immunohistochemistry images (25).

2.3 Promoter methylation analysis

To determine whether the differential gene expression was driven by promoter methylation, the methylation level of FAS in lung cancers and paired normal tissues were compared using the UALCAN database. Methylation analysis in UALCAN used TCGA level-3 data processed through beta-values ranging from 0 to 1, where values closer to 0 indicate unmethylated CpG sites and values closer to 1 indicate fully methylated sites. Statistical comparisons between tumor and normal tissues were performed using a two-sample t-test, with correction for multiple comparisons using FDR.

2.4 Mutation and copy number alteration analysis

The presence and characteristics of FAS gene alterations, including mutations and copy number alterations, were analyzed using cBioPortal (https://www.cbioportal.org/), which contains information on various types of cancer genomics data (26). cBioPortal aggregates data from multiple sources, including TCGA and other cancer genomics projects. Mutational data are derived from whole-exome and whole-genome sequencing, and copy number alterations (CNAs) are identified using GISTIC 2.0 algorithms. Statistical analysis was performed using Fisher's exact tests for comparing mutations.

2.5 Protein-protein interaction analysis

The protein-protein interaction networks of FAS were then reconstructed using the STRING (http://string.embl.de/) (27) and GeneMANIA (https://genemania.org/) (28) databases. STRING integrates known and predicted protein-protein interactions from multiple sources, including experimental data, computational prediction methods, and text mining, with interaction confidence scores based on the strength of evidence. GeneMANIA provides predictions using functional genomics data, including co-expression, colocalization, and physical interaction data. Both platforms employ machine-learning algorithms to predict novel interactions, but predictions can sometimes be prone to false positives or depend on incomplete datasets. STRING compiles data on protein-protein interactions from multiple sources and makes computational predictions to obtain a comprehensive global network of the interactions, whereas GeneMANIA uses extensive genomic and proteomic data to predict protein-protein interactions.

2.6 Co-expression analysis

Genes co-expressed with FAS are identified using GeneMANIA and UALCAN. GeneMANIA uses a combination of Pearson correlation coefficients and other statistical methods to identify genes that show similar expression patterns, which are then displayed in a network. UCSC Xena (https://xenabrowser.net/heatmap/) was then used to generate a correlation heat map with TCGA datasets to visualize the data (29). UCSC Xena applies Pearson correlation to measure the strength of co-expression between *FAS* and its associated genes, with statistical significance provided directly by the database.

2.7 Pathway analysis

Pathways involving FAS and the most frequently coexpressed genes were analyzed using Enrichr (https://maayanlab.cloud/Enrichr/), a gene set enrichment analysis tool, with default parameters (30). A Fisher's exact test was used to evaluate the enrichment of gene sets within biological pathways, adjusting for multiple testing using the Benjamini-Hochberg method. GO terms are categorized into biological processes, molecular functions, and cellular components, and significance is determined through odds ratios and combined scores, which take into account both the magnitude of enrichment and significance. Potential limitations include reliance on existing annotations, which may not fully capture the complexity of gene interactions. Based on the GO terms, the input genes were categorized into biological processes, molecular functions, and cellular components.

2.8 Survival analysis

The prognostic significance of FAS in lung cancer was assessed using the OSluca web server (http://bioinfo.henu.edu.cn/LUCA/LUCAList.jsp), which performs hazard ratio (HR) analysis of data from various datasets, such as TCGA and GEO (31). Kaplan-Meier survival analysis was performed, along with log-rank tests, to determine the statistical significance of survival differences between groups based on *FAS* expression levels. The HR and 95% confidence intervals are provided for each dataset, and multiple comparisons are controlled using FDR correction. Results are

pooled across datasets when appropriate to improve the statistical power of the analysis. HRs from the eligible datasets were then combined to estimate the impact of FAS gene expression on overall survival of lung cancer patients.

2.9 Protein expression

Data on FAS gene expression were retrieved from the Human Protein Atlas. This dataset comprised 578 samples from various tissues, with associated metadata including age, sex, and specific tissue type for each sample. The normalized transcripts per million (nTPM) values for FAS gene expression were extracted from the dataset. Immunohistochemistry images from the Human Protein Atlas were reviewed to confirm FAS protein expression across various tissues.

2.10 *In vitro* validation

Expression of *FAS* in lung cancer cell lines was examined using qRT-PCR. H1299, H1993, A549 and the normal bronchial epithelial cell line HBE were purchased from Shanghai Zhongqiao Xinzou Biotechnology Co., Ltd. and cultured in DMEM medium (Solarbio, Beijing, China) with 10% FBS and 1% penicillin-streptomycin. TRIzol reagent (from Invitrogen, Carlsbad, CA, USA) was used for total RNA extraction and RNA was transcribed into cDNA using ReverTra Ace qPCR RT Master Mix with gDNA Remover Kit. The qRT-PCR was performed using SYBR Premix Ex Taq II on the Mx3005P quantitative real-time fluorescence PCR system (from Stratagene, San Diego, CA, USA), and *GAPDH* was selected as the endogenous control for mRNA. The primer sequences are *FAS*, forward 5'-TCT GGT TCT TAC GTC TGT TGC-3', reverse 5'-CTG TGC AGT CCC TAG CTT TCC-3'; *GAPDH*, forward 5'-GGA GCG AGA TCC CTC CAA AAT-3', reverse 5'-GGC TGT TGT CAT ACT TCT CAT GG-3'. The reaction conditions were as follows: pre-denaturation at 95°C for 10 minutes, denaturation at 95°C for 5 seconds, annealing at 60°C for 30 seconds, for a total of 45 cycles. The target genes and the internal reference gene were amplified for each sample. Each sample group included three replicate wells. Data analysis was performed using the $2^{(-\Delta\Delta Ct)}$ method.

3 Results

3.1 Gene expression analysis

Using GENT2, data on FAS gene expression were available for the GPL570 and GPL96 platforms. For both platforms, gene expression of FAS was found to be significantly altered in several cancer types (see Supplementary Table 1). In lung cancer, the expression of FAS was found to be significantly downregulated ($P < 0.001$ and log2FC = -0.569 for GPL570; $P < 0.001$ and -0.263 for GPL96). Expression data in GEPIA also showed that the expression of FAS was lower in tumor tissues compared with normal lung tissues in both LUAD (TPM =12.17 in tumor tissues and 28.41 in normal tissues) and LUSC (TPM =10.16 in tumor tissues and 29.52 in normal tissues) (Figure 1A). A similar observation was also found in the UALCAN database (P=5.61×10-8 for LUAD, P=5.75×10-12 for LUSC; Figure 1B). At the protein level, data from The Human Protein Atlas showed that in a sample of 578 lung cancer samples, the average nTPM of FAS was 21.7 (range: 5.2-67.2, median: 19.1).

We also performed a subgroup analysis of the expression of FAS in TCGA samples using UALCAN. In all stages of LUAD, the expression of FAS was lower than in normal tissues (Figure 1C). However, no significant difference was found between the different stages of cancer ($P < 0.01$). Similarly, no significant difference in FAS expression was observed between men and women (Figure 1D), different nodal metastasis status (Figure 1E), and different TP53 mutants (Figure 1F). Interestingly, when stratified by patient age, it was found that the older the patients, the higher the expression of FAS in general, although the difference between the different age groups was not statistically significant (Figure 1G). Similarly, there was no significant difference in FAS gene expression between nonsmokers and smokers, but former smokers (who had quit smoking for less than 15 years) had higher FAS gene expression than current smokers (Figure 1H, $P = 0.019$).

In LUSC, a similar observation was found for different cancer stages (Figure 1I), sex (Figure 1J), nodal metastasis status (Figure 1K), and TP53 mutation (Figure 1L). There was also no difference in the expression of FAS among different age groups, but the expression was more constant (i.e., showing no trend of increasing expression with increasing age) (Figure 1M). Regarding smoking status, there was also no significant difference in FAS gene expression between nonsmokers and smokers (Figure 1N). However, former smokers who had quit smoking for less than 15 years had significantly higher FAS expression than those who had quit smoking for more than 15 years ($P < 0.01$).

3.2 Promoter methylation analysis

The extent of FAS promoter methylation in TGCA samples was observed using UALCAN. In LUAD, the median beta-value of FAS promoter methylation in normal tissue is 0.152 (range: 0.135-0.178), whereas the value in tumor tissue is 0.148 (range: 0.103-0.195). There was no statistically significant difference between normal and LUAD tissue (P=0.149; Figure 2A). For LUSC, the median beta-value in normal and tumor tissues was 0.112 (range: 0.093-0.134) and 0.118 (range: 0.064-0.197), respectively. Beta-value was significantly higher in LUSC tissues than in normal tissues ($P < 0.001$; Figure 2B).

3.3 Mutation and copy number alteration analysis

We used cBioportal to investigate the prevalence and types of genomic alterations of the FAS gene in lung cancer patients. Information was available from a total of six TCGA datasets (Firehose Legacy, Nature 2014 and PanCancer Atlas for LUAD; Firehose Legacy, Nature 2012 and PanCancer Atlas for LUSC). Overall, FAS alterations were found in 46 (1.9%) of the 2478 samples (Figure 2C). Specifically, 36 (1.45%) of the samples had

10.3389/fonc.2024.1473515

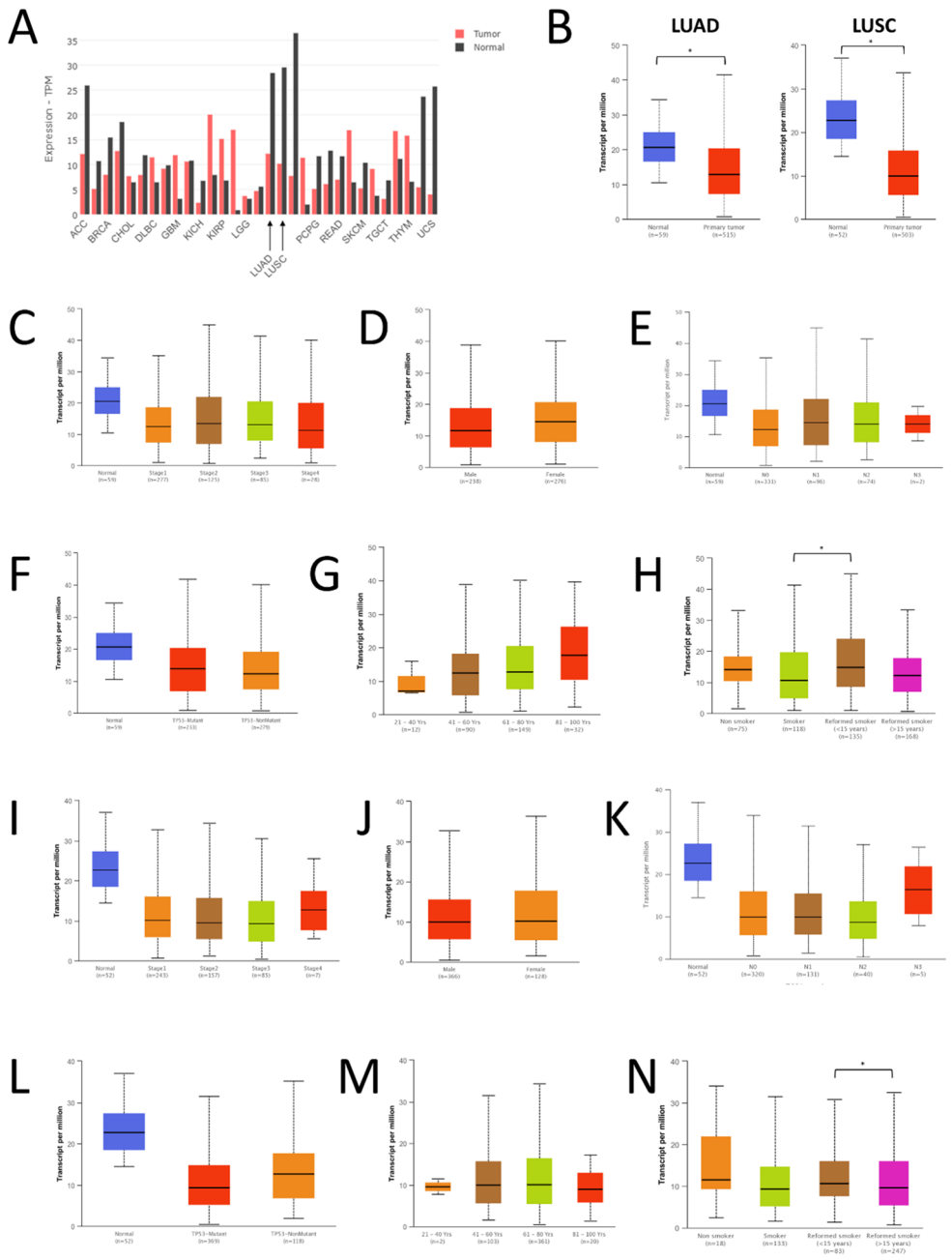


FIGURE 1
Expression of FAS in lung cancer and normal tissues. **(A)** Expression data from GEPIA. **(B)** Expression data from UALCAN. **(C)** Expression of FAS in different stages of LUAD. **(D)** Expression of FAS in men and women with LUAD. **(E)** Expression of FAS in LUAD patients with different nodal metastasis status. **(F)** Expression of FAS in LUAD patients with different TP53 mutations. **(G)** Expression of FAS in LUAD patients with different ages. **(H)** Expression of FAS in LUAD patients of different smoking status. **(I)** Expression of FAS in different stages of LUSC. **(J)** Expression of FAS in men and women with LUSC. **(K)** Expression of FAS in LUSC patients with different nodal metastasis status. **(L)** Expression of FAS in LUSC patients with different TP53 mutations. **(M)** Expression of FAS in LUSC patients with different ages. **(N)** Expression of FAS in LUSC patients of different smoking status. * Statistically significant (P<0.05).

deep deletions and 10 (0.40%) had missense mutations (Table 1; Figure 2D). In LUAD, the missense mutations included p.E114V (N=3) and p.C143F (N=1) in the TNFR-like cysteine-rich domain, p.E261K (N=1) and p.K300E (N=1) in the death domain. In contrast, for LUSC, mutations included p.I262M (N=1) in the death domain and p.S20F (N=2) and p.T219A (N=1) in the non-domain region of the protein product.

3.4 Protein-protein interaction analysis

STRING analysis scored the protein-protein interaction using a score from 0 to 1, where 1 represents the highest probability that the interaction is true based on current evidence. Several proteins were shown to interact with FAS, namely FASLG (score = 0.999), CASP8 (score = 0.999), FADD (score = 0.999), CASP10 (score = 0.997),

10.3389/fonc.2024.1473515

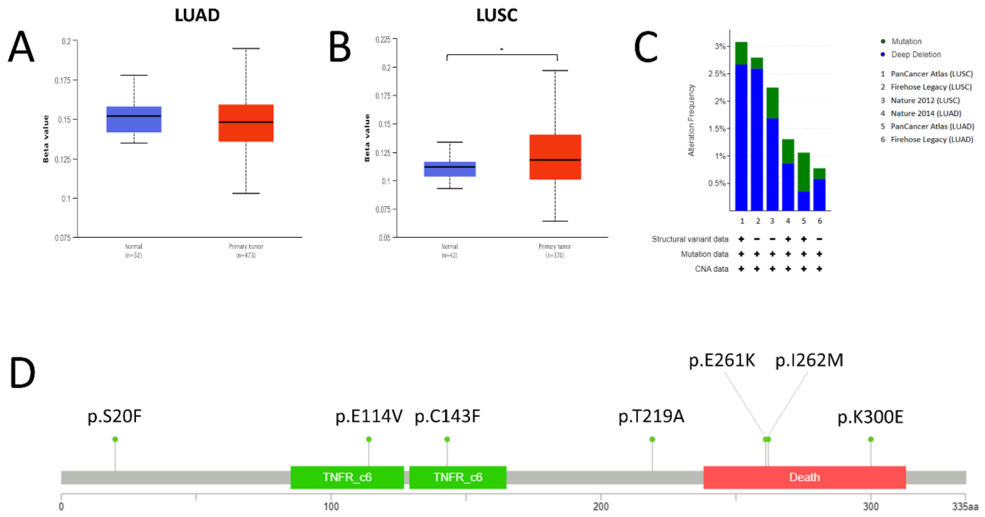


FIGURE 2
Promoter methylation, mutations and copy number status of FAS in lung cancer. **(A)** Promoter methylation level in LUAD. **(B)** Promoter methylation level in LUSC. **(C)** Prevalence of FAS alterations in different TCGA lung cancer datasets. **(D)** Lollipop diagram showing the location of FAS mutations in lung cancer. * Statistically significant (P<0.05).

CFLAR (score = 0.996), DAXX (score = 0.995), PTPN13 (score = 0.995), FAF1 (score = 0.992), RIPK1 (score = 0.985), TRADD (score = 0.983) (Figure 3A). GeneMANIA, on the other hand, categorizes related genes into several categories, namely (1) physical interaction (protein-protein interaction), (2) shared protein domains, (3) colocalization (when genes are expressed in the same tissue or proteins are found in the same location), (4) pathway (two proteins are related when they are involved in the same signaling pathway), and (5) predicted protein interactions. For physical interaction, FAS has been shown to interact with BID, CASP10, CASP8, CFLAR, DAXX, FADD, FAF1, FAIM2, FASLG, MAP3K5, NOL3, PLEC, PRKCA, RAP1A, RIPK1, TNFRSF10B, TNFSF10, TP63, and TRADD. In addition, FADD, CASP8, CFLAR, CASP10, NOL3, TNFRSF10B, RIPK1, and TRADD shared protein domains with FAS. Proteins colocalizing with FAS include FASLG, CASP8, TP63, FADD, PRKCA, BID, DAXX, MAP3K5, and RAP1A. Besides, proteins that participate in the same signaling pathway as FAS include FADD, FASLG, CASP8, CFLAR, CASP10, BID, DAXX, RIPK1, TRADD, FAF1, MAP3K5, FCMR, and FAIM2. Finally, FAIM2, NOL3, TNFRSF10B, TNFSF10, FADD, CASP8, DAXX, and FAF1 are predicted to interact with FAS. The overall interaction network of FAS, as generated by GeneMANIA, is shown in Figure 3B.

TABLE 1 Prevalence of FAS alterations in lung cancer.

Dataset	Prevalence	
	Missense mutations	Deep deletion
LUSC PanCancer Atlas	2/487 (0.41%)	13/487 (2.67%)
LUSC Firehose Legacy	1/501 (0.20%)	13/501 (2.59%)
LUSC Nature 2012	1/178 (0.56%)	3/178 (1.69%)
LUAD Nature 2014	1/230 (0.43%)	2/230 (0.87%)
LUAD PanCancer Atlas	4/566 (0.71%)	2/566 (0.35%)
LUAD Firehose Legacy	1/516 (0.19%)	3/516 (0.58%)
Combined	10/2478 (0.40%)	36/2478 (1.45%)

3.5 Co-expression analysis

Analysis with GeneMANIA showed that CASP10, CFLAR, PRKCA, TNFRSF10B, TNFSF10, and TRADD are frequently co-expressed with FAS. Specifically for LUAD, UALCAN revealed 11 genes whose expression correlates with that of FAS, namely ARL6IP5 (Pearson's coefficient, r = 0.61), RALB (r = 0.56), ELK3 (r = 0.53), CD44 (r = 0.52), DPYD (r = 0.52), GLIPR1 (r = 0.52), DAPP1 (r = 0.51), AIM1 (r = 0.51), LHFPL2 (r = 0.5), CFLAR (r = 0.5), and MDFIC (r = 0.5). On the other hand, in LUSC, 90 co-expressed genes were found, with the top 10 being VCAM1 (r = 0.72), TNFRSF9 (r = 0.69), RELB (r = 0.69), NFKB2 (r = 0.68), BTN2A2 (r = 0.67), BIRC3 (r = 0.67), SH2B3 (r = 0.67), PKDCC (r = 0.67), ZBTB46 (r = 0.66), and JAK2 (r = 0.66) (for the full list, please see Supplementary Table 2). Correlation heat maps generated by UCSC Xena showed for the top five co-expressed genes in LUAD and LUSC are shown in Figure 3C.

3.6 Pathway analysis

Gene set enrichment analysis confirmed that FAS and its most frequently co-expressed genes are involved in apoptosis-related pathways. Specifically, in the biological processes category, FAS and its co-expressed genes were found to be involved most predominantly in negative regulation of extrinsic apoptotic signaling pathway via death domain receptors (GO:1902042, P=2.53×10-18, adjusted P=7.15×10-16, OR=1725.2, combined score=69899.68). In terms of

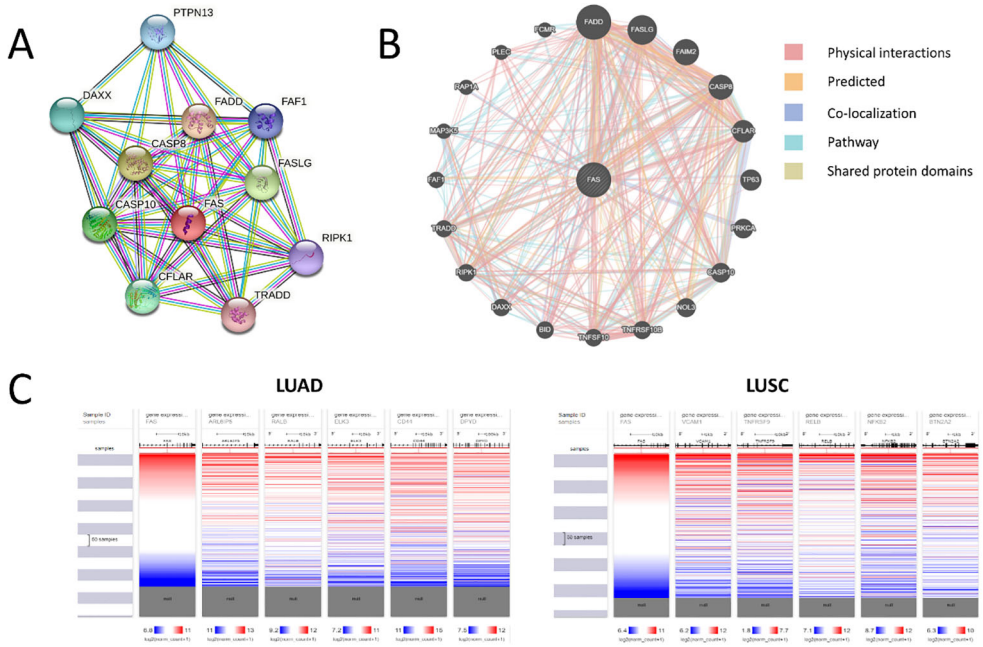


FIGURE 3
Protein-protein and genetic interactions of FAS. **(A)** Protein-protein interaction as predicted using STRING. **(B)** Overall protein-protein interaction network of FAS. **(C)** Correlation heat maps for top three co-expressed genes.

molecular functions, the genes were found to participate mainly in ubiquitin protein ligase binding (GO:0031625, P=5.937×10-6, adjusted P=1.049×10-4, odds ratio=50.39, combined score=606.4). The genes are also implicated in death-inducing signaling complex (GO:0031264, P=3.78×10-24, adjusted P=5.29×10-23, odds ratio=46641, combined score=2515479.09) in terms of cellular components.

3.7 Survival analysis

The OSluca web server contains survival data from 26 datasets. Eleven of the datasets had a hazard ratio (HR) value greater than 1.0 (indicating poor prognosis), although 10 of these were without statistical significance. Only the GSE68465 dataset showed borderline statistical significance at P=0.042 (HR =1.348, 95% CI=1.012-1.797). The remaining 15 datasets, which had an HR value of < 1.0, also did not reach statistical significance. Pooled results from all these datasets suggest that FAS expression has an HR of 1.302 (95% CI=0.935-1.139) (P=0.530), indicating a lack of prognostic significance.

3.8 Protein expression

The expression levels of the FAS protein across various lung cancer samples were investigated. From the dataset that comprised 578 samples, nTPM values of FAS protein in lung cancer samples ranged from 5.2 to 67.2. The average nTPM value across all samples was 21.7, with a median nTPM of 19.1. The standard deviation of the nTPM values was 9.5, indicating variability in FAS expression among the samples (Supplementary Table 3).

3.9 *In vitro* validation

The expression levels of the *FAS* were investigated in three lung cancer cell lines and a normal cell line. The relative expression of *FAS* in lung cancer cell lines (H1299, H1993, A549) was significantly lower than the normal bronchial epithelial cell line (HBE), with *GAPDH* serving as the internal control. The difference was statistically significant ($p < 0.001$ for H1299 and H1993, $p < 0.0001$ for A549; Figure 4; Supplementary Table 4).

4 Discussion

Apoptosis, a tightly regulated process of programmed cell death, plays a crucial role in the pathogenesis of lung cancer. However, the specific characteristics (in terms of expression, mutational and epigenetic profiles, and protein and gene interactions) of apoptosis-related genes such as FAS and their prognostic significance are not well understood. In this study, we sought to clarify these unknowns through an extensive search of online databases. As interpretation of the impact of a genetic variant may vary depending on the specific databases, it is important to consider multiple sources of information when investigating the clinical and biological characteristics of a gene. We therefore searched multiple databases (e.g., GENT2, GEPIA2, and UALCAN for gene expression; STRING and GeneMANIA for protein-protein interaction; etc.) to obtain a conclusive result about the role of FAS in lung cancer. These databases were chosen due to their extensive and curated datasets, wide adoption in cancer research, and their ability to provide different layers of information. For instance, GENT2 and GEPIA2 integrate data from large consortia

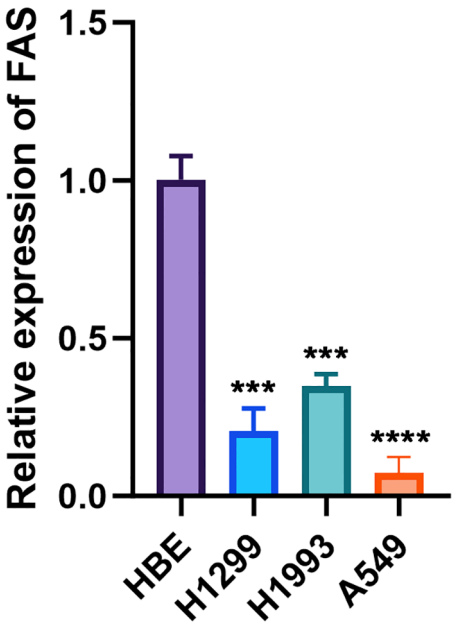


FIGURE 4
Relative expression of FAS in lung cancer cell lines and normal bronchial epithelial cells. The bar graph shows the relative expression levels of FAS in normal bronchial epithelial cell line (HBE) and lung cancer cell lines (H1299, H1993, A549). Data are presented as mean ± SD. Statistical significance is indicated by asterisks: ***$p < 0.001$, ****$p < 0.0001$ compared to HBE.

like TCGA and provide a detailed breakdown of gene expression across cancers, whereas STRING and GeneMANIA focus on elucidating protein and gene interaction networks, helping to contextualize gene function in a broader biological network. However, GENT2 and GEPIA2 are limited by their reliance on bulk RNA sequencing and microarray data, which may obscure cell-type-specific expression patterns, and they may not capture transcript variants or post-transcriptional modifications. UALCAN was selected because of its user-friendly interface and comprehensive analysis of clinicopathological features, which allows for meaningful subgroup analysis based on patient characteristics such as smoking status, age, and tumor stage. A limitation of UALCAN is its dependence on TCGA data, which, although extensive, may not be representative of all population demographics, and batch effects or data inconsistencies across studies can influence the outcomes. STRING and GeneMANIA focus on interaction networks, but STRING relies heavily on computational predictions and text mining, which can introduce false positives, while GeneMANIA's predictions are not always experimentally validated, and both tools may omit less well-characterized interactions. Nevertheless, the use of multiple databases allowed us to ensure robustness of our findings and reduce potential biases inherent in any single database (32). Nonetheless, it is important to note that the use of in silico data mining limits direct biological validation of our findings, and experimental confirmation is needed in future studies. Nevertheless, the findings of this study may provide valuable insight into the role of FAS in the pathogenesis of lung cancer.

One of the most important findings of this work is that FAS was significantly downregulated in lung cancer, both in LUAD and LUSC. The downregulation of FAS was also validated in our *in vitro* analysis in three lung cancer cell lines, when compared against the normal bronchial epithelial cell line HBE. However, it is important to clarify that while this downregulation was statistically significant, its clinical implications remain unclear, as no strong association with patient prognosis was observed. In LUAD, we also observed that former smokers (who had quit smoking for less than 15 years) had higher FAS gene expression than current smokers. This observation is not surprising, because smoking cessation is known to reverse smoking-induced DNA methylation changes (33). Thus, smoking cessation can restore the expression of FAS to levels that are close to those of nonsmokers. However, this finding should be interpreted with caution as it does not account for confounding variables such as the extent and duration of smoking exposure before quitting, which can significantly influence gene expression patterns. Future studies should aim to collect more detailed smoking history data to better understand these interactions. An interesting finding was noted in LUSC, where former smokers who had quit smoking for less than 15 years had significantly higher FAS expression than those who had quit smoking for more than 15 years. This observation is counterintuitive because, according to the logic above, the longer a person has quit smoking, the higher the FAS gene expression should be. However, this observation did not take into account the intensity and duration of smoking before quitting, which may also affect gene expression (34). In addition, gene expression may also be influenced by other factors such as age, sex, and other genetic and nongenetic factors that may interact with smoking status and affect the expression of FAS. Future studies are needed to clarify the underlying mechanisms and potential clinical implications of these findings.

At the protein level, we observed variability in the nTPM levels of FAS protein in the lung cancer samples, ranging from 5.2 to 67.2, suggesting considerable heterogeneity in FAS protein expression across tumor samples. While many samples have moderate FAS expression, some tumors have either significantly elevated or reduced levels of the protein. High FAS expression could indicate an increased capacity for apoptosis in some tumors, possibly serving as a mechanism for tumor suppression, while lower levels of FAS contribute to apoptosis evasion, facilitating tumor progression and resistance to cell death. This differential expression could also reflect differences in the molecular subtype of lung cancer, the tumor microenvironment or the influence of external factors such as smoking or previous treatments. Understanding the causes and consequences of this variability is critical as it may provide insight into tumor behavior, prognosis and therapeutic response, particularly in relation to therapies targeting apoptotic pathways. Further studies are needed to determine how FAS protein levels correlate with clinical outcomes and treatment efficacy in lung cancer.

Given the significantly different expression levels between lung tumor tissues and their normal counterparts, we next sought to determine whether there was a significant difference in the methylation levels of cancerous and noncancerous tissues of the lung. We did not find a statistically significant difference between

normal and LUAD tissues. However, beta-value was significantly higher in LUSC than in normal lung tissues, indicating greater methylation in cancerous tissues. Nevertheless, both normal and LUSC tissues have low levels of DNA methylation (beta-value < 0.3), suggesting that methylation is unlikely to play a dominant role in affecting the expression of FAS. Other factors, such as noncoding RNAs, silencers, enhancers, and transcription factors, might also affect FAS expression (35–37). Indeed, FAS is known to be transcriptionally regulated by members of the p53 family (38), and several silencer and enhancer sequences in the FAS gene have also been identified since the 1990s (39). More recently, antisense RNA of FAS, FAS-AS1 or Saf, has been identified, and is thought to affect the expression of FAS and shown to have functional effects (40, 41). Thus, further studies are needed to investigate the role of these factors in affecting FAS expression. It will also be important to examine whether the methylation status of FAS is tissue-specific and whether certain lung cancer subtypes exhibit unique epigenetic signatures that could provide therapeutic targets.

We have also shown that alterations in FAS are a rare event in lung cancer, occurring in 1.9% of all samples. Deep deletions represent the predominant form of FAS alterations. FAS deletions have been observed in many cancers, including prostate, colorectal, and gastric cancers, but the exact prevalence is not well known because previous studies have typically used small sample sizes (42–44). FAS deletions have been associated with impaired apoptosis, which may serve as an important mechanism of carcinogenesis (44). In addition to small sample sizes, discrepancies in the reported frequency of FAS deletions may also arise from differences in methodologies, such as the use of different sequencing platforms or variant calling algorithms, which can affect the detection of deletions. Furthermore, variations in patient populations, including differences in tumor stage, histological subtype, or demographic factors like age and smoking status, could contribute to variability in FAS deletion prevalence across studies. *In vivo* studies found that deletion of FAS can increase the size and number of intestinal adenomas in mice (45). Another study showed that deletion of FAS, when accompanied by deletion of PTEN, is associated with poor prognosis in hormone-refractory prostate cancer (43). However, deletion of FAS has been shown not to affect its expression (42). It is also difficult to determine whether these genomic changes have a significant clinical impact in lung cancer, as the frequency of these FAS alterations was low. Therefore, the significance of FAS deletion in carcinogenesis requires further research.

We also identified missense mutations in 0.40% of lung cancer samples. The mutations present in LUAD are different from those in LUSC. However, the small number of affected samples does not allow us to reliably determine whether the findings can be interpreted as different mechanisms of carcinogenesis in the two lung cancer subtypes. The functional consequences of these variants have not been thoroughly studied, so it is not known whether they play a driving role in lung cancer. Nevertheless, the p.E261K mutation found in LUAD and the p.I262M mutation found in LUSC, both of which result in amino acid changes in the death domain of the Fas protein, have been linked to autoimmune lymphoproliferative syndrome (46–48). *In vitro* studies revealed that the p.E261K mutation can impair the process of reorganization of Fas into large protein islands and also has a dominant-negative property that adversely affects the normal wild-type Fas during the formation of the Fas-FADD signaling complex (47, 48). While these findings suggest potential functional effects of these mutations, the lack of statistical significance and the small number of cases prevent us from making broad generalizations. Additional functional studies are needed to determine whether these mutations have oncogenic or tumor-suppressive roles in lung cancer.

Genes and proteins often engage in various forms of molecular interactions, such as gene-gene and protein-protein interactions, to perform their biological functions (49, 50). Therefore, it is important to understand these interactions to decipher the complexity of biological systems. We have performed protein-protein interactions and genetic co-expression analyses to identify proteins and genes that may interact with FAS. Perhaps not surprisingly, many of the identified genes/proteins, such as FASLG, CASP8, FADD, CASP10, BID, TRADD, and CFLAR, are involved in the apoptotic process. This finding supports the hypothesis that FAS and its associated genes and proteins play a key role in regulating apoptosis in lung cancer. However, the lack of mechanistic studies in lung cancer cells limits our ability to determine the functional importance of these interactions. Investigating these interactions *in vitro* or *in vivo* could provide deeper insights into their relevance in lung cancer progression (51). It is interesting to note that 90 genes were significantly co-expressed in LUSC, whereas only 11 genes were significantly co-expressed in LUAD, which may reflect the differences in molecular and cellular processes involved in the two lung cancer subtypes. The higher number of co-expressed genes in LUSC suggests that the molecular networks and signaling pathways in LUSC are more complex and interconnected than those in LUAD. Indeed, a recent study also demonstrated that many cancer-related signaling pathways, including Notch, Hedgehog, Wnt, and ErbB pathways, were significantly overrepresented in LUSC compared to LUAD (52). Another possible explanation for this observation is that compared to LUAD, LUSC is more frequently associated with tobacco smoking, which can cause extensive genomic damage and activate many cellular signaling pathways, including inflammation and oxidative stress, that can further drive cancer development (52, 53). The involvement (or lack thereof) of these signaling pathways in oncogenesis may also contribute to the observed differences in co-expression between the two types of lung cancer. The clinical significance of these differences in co-expressed genes between LUAD and LUSC is not well understood and represents a future research direction. Further exploration into the functional consequences of these differences could reveal important subtype-specific therapeutic targets and guide personalized treatment strategies (54).

The lack of significant prognostic value for FAS in this study, despite it being downregulated in lung cancer, highlights an important aspect of cancer research — the importance of negative findings. While FAS plays a role in several types of cancer, its role in lung cancer may be more context-dependent or its prognostic significance may be overshadowed by other factors. This negative result is valuable because it prompts future research to consider more complex, multifactorial prognostic models that include additional markers or signaling pathways. It also suggests that the role of FAS in lung cancer may not be clear, so its interactions with other apoptotic or non-apoptotic metabolic pathways and its behavior under different

conditions of the tumor microenvironment need to be further investigated (55). Negative results such as these help to refine research questions and focus on promising targets or combinations of biomarkers that could provide clinically meaningful prognostic information.

Despite successfully demonstrating the reduced expression of the FAS gene in lung cancer, survival analysis revealed no prognostic significance of the gene in lung cancer. This suggests that although downregulation of FAS is a common feature of lung cancer, it may not be a reliable predictor of disease outcome. This result was not consistent with previous findings in lung and other cancers, which showed a significant association between expression of FAS and cancer prognosis (56–60). However, it should be noted that previous studies on the prognostic significance of the FAS gene used small sample sizes, which may lead to inaccurate conclusions (61). Our in silico data mining combined survival data from 26 datasets, which greatly improved the statistical power required for accurate analysis. The lack of prognostic significance in our study may indicate that while FAS plays a role in tumor initiation, its downregulation may not be critical for disease progression or metastasis in lung cancer. One possible explanation for the limited prognostic significance of FAS is that although the gene plays a role in the early stages of lung cancer development, its expression may not be critical for tumor progression or metastasis. However, this postulation is not supported by several studies that showed that FAS can promote progression and metastasis in various cancers (10, 13, 62, 63). Nevertheless, none of these studies were conducted in lung cancer cells, and it remains unclear whether FAS plays a role in lung cancer cell progression and metastasis. Further *in vitro* and *in vivo* studies are needed to explore the role of FAS in lung cancer metastasis, particularly to assess whether its downregulation affects the invasive potential of lung cancer cells. It is also possible that there are other confounding factors or co-occurring genes that affect lung cancer progression and patient survival, which could limit the prognostic significance of FAS expression in the cancer. The role of FAS and other factors in influencing lung cancer progression and metastasis deserves further investigation. Future studies should aim to investigate whether FAS expression in combination with other apoptotic markers could provide a more accurate prognostic model for lung cancer.

5 Conclusions

In conclusion, we have successfully characterized the role of FAS in lung cancer. Specifically, we have shown that FAS is significantly downregulated in lung cancer and characterized its mutational and methylation profiles. We also identified its protein-protein interactions and co-expressed genes and reconfirmed the important role of FAS and its co-expressed genes in apoptosis-related pathways. Finally, we have shown that despite the above observations, the prognostic significance of FAS in lung cancer is limited. The clinical implications of FAS downregulation, alterations, and molecular interactions, as well as the differences between LUAD and LUSC in these features, remain to be investigated. Thus, there is a need for more comprehensive and integrative approaches to understand the molecular and cellular mechanisms of FAS that drive lung cancer progression. Future studies should focus on functional analyses of FAS and its mutations in lung cancer cells to better understand how downregulation of FAS contributes to apoptosis evasion. In addition, it will be important to investigate the role of non-coding RNAs, transcription factors and other regulatory elements that may influence FAS expression. Further research should also investigate the potential of FAS as part of a biomarker panel in combination with other apoptotic genes for a more accurate prognosis. Finally, *in vivo* studies are needed to assess whether modulation of FAS expression could have therapeutic potential in lung cancer, either as a direct target or in combination with existing treatments.

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

KL: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Visualization, Writing – original draft. SCT: Data curation, Formal analysis, Methodology, Writing – review & editing. ZY: Data curation, Formal analysis, Methodology, Writing – review & editing. CL: Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2024.1473515/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) 71:209–49. doi: 10.3322/caac.21660

2. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet*. (2018) 391:1023–75.doi: 10.1016/S0140-6736(17)33326-3

3. Didkowska J, Wojciechowska U, Mańczuk M, Łobaszewski J. Lung cancer epidemiology: contemporary and future challenges worldwide. *Ann Transl Med*. (2016) 4:150. doi: 10.21037/atm.2016.03.11

4. Jakobsen E, Olsen KE, Bliddal M, Hornbak M, Persson GF, Green A. Forecasting lung cancer incidence, mortality, and prevalence to year 2030. *BMC Cancer*. (2021) 21:985. doi: 10.1186/s12885-021-08696-6

5. Schabath MB, Cote ML. Cancer progress and priorities: lung cancer. *Cancer Epidemiol Biomarkers Prev*. (2019) 28:1563–79. doi: 10.1158/1055-9965.epi-19-0221

6. Hanahan D. Hallmarks of cancer: new dimensions. *Cancer Discov.*. (2022) 12:31–46. doi: 10.1158/2159-8290.cd-21-1059

7. Hanahan D, Weinberg Robert A. Hallmarks of cancer: the next generation. *Cell*. (2011) 144:646–74. doi: 10.1016/j.cell.2011.02.013

8. Levoin N, Jean M, Legembre P. CD95 structure, aggregation and cell signaling. *Front Cell Dev Biol*. (2020) 8:314. doi: 10.3389/fcell.2020.00314

9. Ivanisenko NV, Lavrik IN. Mechanisms of procaspase-8 activation in the extrinsic programmed cell death pathway. *Mol Biol*. (2019) 53:732–8. doi: 10.1134/s0026893319050091

10. Guégan JP, Ginestier C, Charafe-Jauffret E, Ducret T, Quignard J-F, Vacher P, et al. CD95/Fas and metastatic disease: What does not kill you makes you stronger. *Semin Cancer Biol*. (2020) 60:121–31. doi: 10.1016/j.semcancer.2019.06.004

11. Williams JW, Ferreira CM, Blaine KM, Rayon C, Velázquez F, Tong J, et al. Non-apoptotic fas (CD95) signaling on T cells regulates the resolution of th2-mediated inflammation. *Front Immunol*. (2018) 9:2521. doi: 10.3389/fimmu.2018.02521

12. Qadir AS, Stults AM, Murmann AE, Peter ME. The mechanism of how CD95/Fas activates the Type I IFN/STAT1 axis, driving cancer stemness in breast cancer. *Sci Rep*. (2020) 10:1310. doi: 10.1038/s41598-020-58211-3

13. Li L-J, Chang PM-H, Li C-H, Chang Y-C, Lai T-C, Su C-Y, et al. FAS receptor regulates NOTCH activity through ERK-JAG1 axis activation and controls oral cancer stemness ability and pulmonary metastasis. *Cell Death Discov*. (2022) 8:101. doi: 10.1038/s41420-022-00899-5

14. Tan SC. Low penetrance genetic polymorphisms as potential biomarkers for colorectal cancer predisposition. *J Gene Med*. (2018) 20:e3010. doi: 10.1002/jgm.3010

15. Tan SC, Ankathil R. Genetic susceptibility to cervical cancer: role of common polymorphisms in apoptosis-related genes. *Tumour biology: J Int Soc Oncodevelopmental Biol Med*. (2015) 36:6633–44. doi: 10.1007/s13277-015-3868-2

16. Tan SC, Ismail MP, Duski DR, Othman NH, Ankathil R. FAS c.-671A>G polymorphism and cervical cancer risk: a case-control study and meta-analysis. *Cancer Genet*. (2017) 211:18–25. doi: 10.1016/j.cancergen.2017.01.004

17. Daripally S, Peddi K. 5-year cumulative survival of oral cancer patients with FAS and FASL SNPs. *Meta Gene*. (2021) 27:100826. doi: 10.1016/j.mgene.2020.100826

18. Hoxhaj I, Vukovic V, Boccia S, Pastorino R. Single nucleotide polymorphisms and the risk of developing a second primary cancer among head and neck cancer patients: a systematic literature review and meta-analysis. *BMC Cancer*. (2021) 21:660. doi: 10.1186/s12885-021-08335-0

19. Chen Y, Wang H, Yan Y, Ren M, Yan C, Wang B. Correlation between FAS single nucleotide polymorphisms and breast carcinoma susceptibility in Asia. *Med (Baltimore)*. (2019) 98:e18240. doi: 10.1097/md.0000000000018240

20. Zhang Y, Liu Q, Zhang M, Yu Y, Liu X, Cao X. Fas signal promotes lung cancer growth by recruiting myeloid-derived suppressor cells via cancer cell-derived PGE21. *J Immunol*. (2009) 182:3801–8. doi: 10.4049/jimmunol.0801548

21. Huang X, Xiao S, Zhu X, Yu Y, Cao M, Zhang X, et al. miR-196b-5p-mediated downregulation of FAS promotes NSCLC progression by activating IL6-STAT3 signaling. *Cell Death Dis*. (2020) 11:785. doi: 10.1038/s41419-020-02997-7

22. Park S-J, Yoon B-H, Kim S-K, Kim S-Y. GENT2: an updated gene expression database for normal and tumor tissues. *BMC Med Genomics*. (2019) 12:101. doi: 10.1186/s12920-019-0514-7

23. Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res*. (2019) 47: W556–W60. doi: 10.1093/nar/gkz430

24. Chandrashekar DS, Karthikeyan SK, Korla PK, Patel H, Shovon AR, Athar M, et al. UALCAN: An update to the integrated cancer data analysis platform. *Neoplasia*. (2022) 25:18–27. doi: 10.1016/j.neo.2022.01.001

25. Sjöstedt E, Zhong W, Fagerberg L, Karlsson M, Mitsios N, Adori C, et al. An atlas of the protein-coding genes in the human, pig, and mouse brain. *Science*. (2020) 367: eaay5947. doi: 10.1126/science.aay5947

26. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. (2012) 2:401–4. doi: 10.1158/2159-8290.cd-12-0095

27. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. (2020) 49: D605–D12. doi: 10.1093/nar/gkab835

28. Franz M, Rodriguez H, Lopes C, Zuberi K, Montojo J, Bader GD, et al. GeneMANIA update 2018. *Nucleic Acids Res*. (2018) 46:W60–w4. doi: 10.1093/nar/gky311

29. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol*. (2020) 38:675–8. doi: 10.1038/s41587-020-0546-8

30. Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, et al. Gene set knowledge discovery with enrichr. *Curr Protoc*. (2021) 1:e90. doi: 10.1002/cpz1.90

31. Yan Z, Wang Q, Lu Z, Sun X, Song P, Dang Y, et al. OSluca: an interactive web server to evaluate prognostic biomarkers for lung cancer. *Front Genet*. (2020) 11:420. doi: 10.3389/fgene.2020.00420

32. Chen Y, Li C, Wang N, Wu Z, Zhang J, Yan J, et al. Identification of LINC00654-NINL regulatory axis in diffuse large B-cell lymphoma in silico analysis. *Front Oncol*. (2022) 12:883301. doi: 10.3389/fonc.2022.883301

33. McCartney DL, Stevenson AJ, Hillary RF, Walker RM, Bermingham ML, Morris SW, et al. Epigenetic signatures of starting and stopping smoking. *EBioMedicine*. (2018) 37:214–20. doi: 10.1016/j.ebiom.2018.10.051

34. Guida F, Sandanger TM, Castagné R, Campanella G, Polidoro S, Palli D, et al. Dynamics of smoking-induced genome-wide methylation changes with time since smoking cessation. *Hum Mol Genet*. (2015) 24:2349–59. doi: 10.1093/hmg/ddu751

35. Najafi S, Tan SC, Raee P, Rahmati Y, Asemani Y, Lee EHC, et al. Gene regulation by antisense transcription: A focus on neurological and cancer diseases. *Biomed Pharmacother*. (2022) 145:112265. doi: 10.1016/j.biopha.2021.112265

36. Panigrahi A, O'Malley BW. Mechanisms of enhancer action: the known and the unknown. *Genome Biol*. (2021) 22:108. doi: 10.1186/s13059-021-02322-1

37. Bushweller JH. Targeting transcription factors in cancer — from undruggable to reality. *Nat Rev Cancer*. (2019) 19:611–24. doi: 10.1038/s41568-019-0196-7

38. Schilling T, Schleithoff ES, Kairat A, Melino G, Stremmel W, Oren M, et al. Active transcription of the human FAS/CD95/TNFRSF6 gene involves the p53 family. *Biochem Biophys Res Commun*. (2009) 387:399–404. doi: 10.1016/j.bbrc.2009.07.063

39. Rudert F, Visser E, Forbes L, Lindridge E, Wang Y, Watson J. Identification of a silencer, enhancer, and basal promoter region in the human CD95 (Fas/APO-1) gene. *DNA Cell Biol*. (1995) 14:931–7. doi: 10.1089/dna.1995.14.931

40. Yuan Q, Zhang H, Pan Z, Ling X, Wu M, Gui Z, et al. Regulatory loop between lncRNA FAS-AS1 and DNMT3b controls FAS expression in hydroquinone-treated TK6 cells and benzene-exposed workers. *Environ pollut*. (2020) 261:114147. doi: 10.1016/j.envpol.2020.114147

41. Villamizar O, Chambers CB, Mo Y-Y, Torry DS, Hofstrand R, Riberdy JM, et al. Fas-antisense long noncoding RNA is differentially expressed during maturation of human erythrocytes and confers resistance to Fas-mediated cell death. *Blood Cells Molecules Dis*. (2016) 58:57–66. doi: 10.1016/j.bcmd.2016.03.002

42. Butler LM, Hewett PJ, Butler WJ, Cowled PA. Down-regulation of Fas gene expression in colon cancer is not a result of allelic loss or gene rearrangement. *Br J Cancer*. (1998) 77:1454–9. doi: 10.1038/bjc.1998.239

43. Williams J, Yoshimoto M, Sircar K, Zielenska M, Squire J. Abstract #4325: PTEN homozygous genomic deletion accompanied by deletion of FAS is associated with poor outcome in hormone refractory prostate cancer. *Cancer Res*. (2009) 69:4325.

44. Hui L, Wenjie W, Junhui Z, Qian L, Yuanfang Z eds. Expression and Clinical Significance of Claudin-7, Fas and Integrin ß1 in Gastric Carcinoma Tissues. In: *2015 7th International Conference on Information Technology in Medicine and Education (ITME)*. Huangshan, China: IEEEp. 13–5.

45. Guillen-Ahlers H, Suckow MA, Castellino FJ, Ploplis VA. Fas/CD95 deficiency in apcMin/+ Mice increases intestinal tumor burden. *PloS One*. (2010) 5:e9070. doi: 10.1371/journal.pone.0009070

46. Suspitsin EN, Guseva MN, Kostik MM, Sokolenko AP, Skripchenko NV, Levina AS, et al. Next generation sequencing analysis of consecutive Russian patients with clinical suspicion of inborn errors of immunity. *Clin Genet*. (2020) 98:231–9. doi: 10.1111/cge.13789

47. Frazzette N, Cruz AC, Wu X, Hammer JA, Lippincott-Schwartz J, Siegel RM, et al. Super-resolution imaging of fas/CD95 reorganization induced by membrane-bound fas ligand reveals nanoscale clustering upstream of FADD recruitment. *Cells*. (2022) 11:1908. doi: 10.3390/cells11121908

48. Wang L, Yang JK, Kabaleeswaran V, Rice AJ, Cruz AC, Park AY, et al. The Fas–FADD death domain complex structure reveals the basis of DISC assembly and disease mutations. *Nat Struct Mol Biol*. (2010) 17:1324–9. doi: 10.1038/nsmb.1920

49. Arshad Z, McDonald JF. Changes in gene-gene interactions associated with cancer onset and progression are largely independent of changes in gene expression. *iScience*. (2021) 24:103522. doi: 10.1016/j.isci.2021.103522

50. Gulfidan G, Turanli B, Beklen H, Sinha R, Arga KY. Pan-cancer mapping of differential protein-protein interactions. *Sci Rep*. (2020) 10:3272. doi: 10.1038/s41598-020-60127-x

51. Yu Y, Huang Y, Li C, Ou S, Xu C, Kang Z. Clinical value of M1 macrophage-related genes identification in bladder urothelial carcinoma and in *vitro* validation. *Front Genet*. (2022) 13:1047004. doi: 10.3389/fgene.2022.1047004

52. Anusewicz D, Orzechowska M, Bednarek AK. Lung squamous cell carcinoma and lung adenocarcinoma differential gene expression regulation through pathways of Notch, Hedgehog, Wnt, and ErbB signalling. *Sci Rep*. (2020) 10:21128. doi: 10.1038/s41598-020-77284-8

53. Caliri AW, Tommasi S, Besaratinia A. Relationships among smoking, oxidative stress, inflammation, macromolecular damage, and cancer. *Mutat Res Rev Mutat Res*. (2021) 787:108365. doi: 10.1016/j.mrrev.2021.108365

54. Chu Y, Xu Q, Zhou X, Qiuying N, Yao X, Fang J, et al. Pristimerin inhibits thioredoxin reductase in the treatment of non-small cell lung cancer. *Acta Materia Med*. (2024) 3:239-53. doi: 10.15212/amm-2024-0015

55. Wu X, Lu W, Xu C, Jiang C, Zhuo Z, Wang R, et al. Macrophages phenotype regulated by IL-6 are associated with the prognosis of platinum-resistant serous ovarian cancer: integrated analysis of clinical trial and omics. *J Immunol Res*. (2023) 2023:6455704. doi: 10.1155/2023/6455704

56. Koomägi R, Volm M. Expression of Fas (CD95/APO-1) and Fas ligand in lung cancer, its prognostic and predictive relevance. *Int J Cancer*. (1999) 84:239–43. doi: 10.1002/(sici)1097-0215(19990621)84:3<239::aid-ijc7>3.3.co;2-j

57. Ohno S, Tachibana M, Shibakita M, Dhar DK, Yoshimura H, Kinugasa S, et al. Prognostic significance of fas and fas ligand system-associated apoptosis in gastric cancer. *Ann Surg Oncol*. (2000) 7:750–7. doi: 10.1007/s10434-000-0750-1

58. de Carvalho-Neto PB, Santos MD, de Carvalho MB, Mercante AMDC, Santos VPPD, Severino P, et al. FAS/FASL expression profile as a prognostic marker in squamous cell carcinoma of the oral cavity. *PloS One*. (2013) 8:e69024. doi: 10.1371/journal.pone.0069024

59. Eser B, Sari I, Canoz O, Altuntas F, Cakmak E, Ozturk A, et al. Prognostic significance of Fas (CD95/APO-1) positivity in patients with primary nodal diffuse large B-cell lymphoma. *Am J Hematol*. (2006) 81:307–14. doi: 10.1002/ajh.20564

60. Shibakita M, Tachibana M, Dhar DK, Kotoh T, Kinugasa S, Kubota H, et al. Prognostic significance of Fas and Fas ligand expressions in human esophageal cancer. *Clin Cancer Res*. (1999) 5:2464–9.

61. Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ*. (2020) 368:m441. doi: 10.1136/bmj.m441

62. Koshkina N, Yang Y, Kleinerman ES. The Fas/FasL Signaling Pathway: Its Role in the Metastatic Process and as a Target for Treating Osteosarcoma Lung Metastases. In: Kleinerman ES, Gorlick R, editors. *Current Advances in the Science of Osteosarcoma: Research Perspectives: Tumor Biology, Organ Microenvironment, Potential New Therapeutic Targets, and Canine Models*. Springer International Publishing, Cham (2020). p. 177–87.

63. Li H, Fan X, Stoicov C, Liu JH, Zubair S, Tsai E, et al. Human and mouse colon cancer utilizes CD95 signaling for local growth and metastatic spread to liver. *Gastroenterology*. (2009) 137:934–44.e4. doi: 10.1053/j.gastro.2009.06.004

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 04 October 2024
DOI 10.3389/fimmu.2024.1470449

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Yuquan Chen,
Monash University, Australia
Yisheng Chen,
Fudan University, China
Biao Zhang,
Dalian Medical University, China
Min Miao,
Ningbo University, China

*CORRESPONDENCE
Yi Zhang
240294866@qq.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 25 July 2024
ACCEPTED 13 September 2024
PUBLISHED 04 October 2024

CITATION
Zhang S, Zhang X, Xiahou Z, Zuo S, Xue J and Zhang Y (2024) Unraveling the ecological landscape of mast cells in esophageal cancer through single-cell RNA sequencing.
Front. Immunol. 15:1470449.
doi: 10.3389/fimmu.2024.1470449

Unraveling the ecological landscape of mast cells in esophageal cancer through single-cell RNA sequencing

Shengyi Zhang[1†], Xinyi Zhang[2†], Zhikai Xiahou[3†], Shunqing Zuo[1], Jialong Xue[1] and Yi Zhang[1*]

[1]Department of Thoracic Surgery, Songjiang Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China, [2]Clinical Medical College, Southwest Medical University, Luzhou, China, [3]China Institute of Sport and Health Science, Beijing Sport University, Beijing, China

Background: Esophageal cancer (EC) is a major health issue, ranking seventh in incidence and sixth in mortality worldwide. Despite advancements in multidisciplinary treatment approaches, the 5-year survival rate for EC remains low at 21%. Challenges in EC treatment arise from late-stage diagnosis, high malignancy, and poor prognosis. Understanding the tumor microenvironment is critical, as it includes various cellular and extracellular components that influence tumor behavior and treatment response. Mast cells (MCs), as tissue-resident immune cells, play dual roles in tumor dynamics. High-throughput single-cell RNA sequencing offers a powerful tool for analyzing tumor heterogeneity and immune interactions, although its application in EC is limited.

Methods: In this study, we investigated the immune microenvironment of EC using single-cell RNA sequencing and established a comprehensive immune profile. We also performed analysis of upstream transcription factors and downstream pathway enrichment to further comprehensively decipher MCs in EC. Besides, we performed knockdown experiments to explore the role of epidermal growth factor receptor (*EGFR*) signaling pathway in MCs-tumor cell interactions, highlighting its potential as a prognostic marker. Finally, we constructed a prognostic model for EC, which provided valuable suggestions for the diagnosis and prognosis of EC.

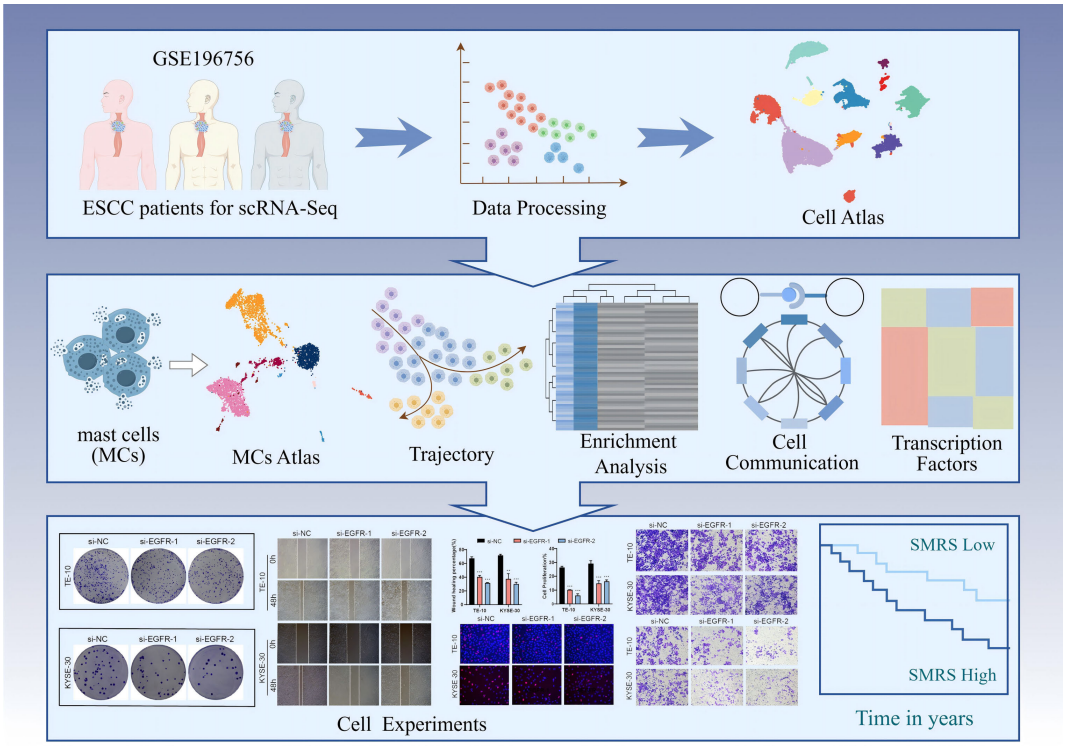
Results: Our analysis identified 11 major cell types, of which MCs were particularly present in pericarcinoma tissues. Further grouping of the 5,001 MCs identified 8 distinct subtypes, including *SRSF7*-highly expressed MCs, which showed strong tumor preference and potential tumor-promoting properties. Moreover, we identified the key signaling receptor *EGFR* and validated it by in vitro knockdown experiments, demonstrating its cancer-promoting effects. In addition, we established an independent prognostic indicator, *SRSF7*+ MCs risk score (SMRS), which showed a correlation between high SMRS group and poor prognosis.

10.3389/fimmu.2024.1470449

Conclusion: These findings illuminate the complex interactions within the tumor microenvironment of EC and suggest that targeting specific MCs subtypes, particularly via the *EGFR* signaling pathway, may present novel therapeutic strategies. This study establishes a comprehensive immune map of EC, offering insights for improved treatment approaches.

KEYWORDS

single-cell RNA sequencing, mast cells, EGFR signaling pathway, prognostic model, esophageal cancer



GRAPHICAL ABSTRACT
Article research flow chart. First, we downloaded the clinical data of three esophageal cancer patients from TCGA, and after a series of data processing, we obtained the UMAP map of esophageal cancer. After that, we focused on MCs to get its UMAP map and performed trajectory analysis, enrichment analysis, cellular communication analysis and transcription factor analysis sequentially. Finally, we performed *in vitro* experiments on the pathway derived from communication analysis, verified the effect of the pathway on esophageal cancer progression through a series of experiments, and analyzed the related prognosis, and the results obtained were consistent with our study.

1 Introduction

Esophageal cancer (EC) is a common malignant tumor of the gastrointestinal system, with the seventh highest incidence and sixth highest mortality rate in the world (1). In China, the incidence and mortality rates of EC rank third and fourth, respectively, among all malignant tumors (2). Despite the development of a multidisciplinary treatment approach, the prognosis remains unfavorable (3). The 5-year survival rate for EC is only 21%, after pancreatic and liver cancers (4). Therefore, EC has been a major malignant tumor threatening the health of Chinese residents. EC consists of two main subtypes, esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma, with ESCC accounting for about 90% of all EC

Abbreviations: EC, esophageal cancer; MCs, mast cells; TME, tumor microenvironment; scRNA-seq, single-cell RNA sequencing.

cases worldwide (5). EC is an aggressive cancer with rapid growth and a high rate of lymph node metastasis, usually involving the upper two thirds of the esophagus (6). In retrospective studies in EC, smoking, hot tea consumption, red meat consumption, poor oral health, low intake of fresh fruits and vegetables, and low socioeconomic status were associated with a higher risk of EC (7). Previous studies have shown that chronic inflammation plays a central role in progression from esophageal precancerous lesions (EPL) to esophageal squamous cell carcinoma, that dietary inflammatory potential has been linked to both EPL and ESCC, and that inflammatory imbalances promote tumorigenesis, and that the consumption of anti-inflammatory foods may be helpful in the prevention of EPL and ESCC (8–10). Difficulty swallowing and swollen lymph nodes in the neck do not appear until the cancer has progressed to an advanced stage (11), and the treatment of EC patients faces major challenges due to the lack of early symptoms, high malignancy, poor prognosis, and surgical complexity of EC. Although we have made great progress in the treatment of EC in recent years, especially through preoperative radiotherapy combined with immunotherapy, which shows a broad potential in the treatment of EC. However, due to the high rate of post-treatment recurrence and the limitations of drugs and treatment strategies after metastasis, only a small proportion of EC patients can benefit from the available treatments, while the majority of patients respond poorly to the treatments, and therefore, the overall survival rate of EC is still disappointing in China (3, 12).

In addition, due to the heterogeneity and complexity of tumors, the mechanisms of tumor proliferation, metastasis, drug resistance, and immunosuppression are unknown. Therefore, elucidating the molecular mechanisms of tumorigenesis and tumor progression is crucial for effective control and management of tumor development. Notably, the presence of non-tumor cells within the tumor tissue is also critical for tumor development (13). Therefore, shifting the therapeutic focus to other components of the tumor microenvironment (TME) may become an important strategy for future tumor therapy. The introduction of TME has played a very powerful role in advancing oncology research. TME has had an incredibly important role in the development and evolution of EC (14). The TME consists of multiple cellular components (e.g., fibroblasts, endothelial cells, and immune cells) and extracellular components (including cytokines, hormones, extracellular matrices, and growth factors), which form a complex network that encapsulates EC cells. These cells shape cancer biology and influence the response to treatment (15–17). In TME, mast cells (MCs) are tissue-resident immune cells that are important players in diseases associated with chronic inflammation such as cancer. Because MCs can infiltrate solid tumors and promote or limit tumor growth, MCs may polarize to either pro- or anti-tumor phenotypes and remain a challenging area of research (18). Previous articles have also hypothesized that *NRF2* in combination with AC-MCs may be a predictive marker for prognosis and may influence immunotherapy by modulating PD-L1 in EC (19).

High-throughput single-cell RNA sequencing (scRNA-seq), developed in recent years, is an effective method that has been shown to dissect heterogeneous tumors and decipher the interactions between cancer cells and their microenvironmental components, and to elucidate the transcriptomic profiles of both the cancer cells and the microenvironmental components (20–22), which is the basis and foundation for furthering the understanding of cancers and the development of effective early diagnostic and therapeutic strategies, previous studies have dissected the esophageal squamous cell carcinoma ecosystem by single-cell transcriptomic analysis (16), but its application in EC is still limited. At the same time, there is still a long way to go for early detection of esophageal cancer (23), and prognostic tools lack the necessary accuracy to facilitate individualized patient management strategies (24).

Therefore, in this study, scRNA-seq was used to sequence EC samples in order to decipher the immune microenvironment of EC, reveal the immune map of EC, and provide new insights for the treatment of EC. The functional role of MCs subtypes in EC and their association with tumor tissues are extensively discussed and summarized in this paper, and a prognostic model is established, which provides a valuable resource for deeper understanding of the causes and progression of EC and helps to improve its therapeutic strategies.

2 Materials and methods

2.1 Data source

The scRNA-seq data of EC were acquired from the GEO website (https://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE196756. Patient clinical sample information can be found at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi. Considering the utilization of publicly accessible data derived from databases, it was not required to secure an ethical endorsement for this investigation.

2.2 Single-cell sequencing

The gene expression data were imported into the R software and analyzed using the Seurat R package (25, 26). Cells of inferior quality were excluded based on the following criteria (1): nFeature between 300 and 7,500 (2); nCount between 500 and 100,000 (3); mitochondrial gene expression occupying no more than 25% of the total gene count within the cell (4); erythrocyte gene expression not surpassing 5% of the total gene count within the cell.

Subsequently, all gene expression data underwent normalization and scaling using the "NormalizeData" and "ScalData" functions within the Seurat R package (27). For the purpose of principal component analysis, the "FindVariableFeautres" function (28) was implemented to identify the top 2,000 most variable genes. These cells were then segregated into clusters based on the top 30 principal components (PCs) using the "FindClusters" function at a resolution

of 1.0. Finally, the top 30 significant PCs were selected to dimensionality reduction and visualization of gene expression through uniform manifold approximation and projection (UMAP) (29, 30). The harmony R package (31, 32) was employed to alleviate the influence of batch effects among the samples. The dim value was set to 30, while the resolution parameter was configured to 1.2.

2.3 Identification of cell subtypes

Cell clusters were initially discerned utilizing the "FindClusters" and "FindNeighbors" functions within Seurat (33–35), employing a default resolution of 0.8. Afterwards, these cell clusters were bestowed with annotations based on the average gene expression of representative markers. In order to evaluate differentially expressed genes (DEGs) across distinct cell clusters, a Wilcoxon rank sum test was employed utilizing Seurat's "FindAllMarkers" function (36, 37). The parameters min.pct and min.diff.pct were established at 0.25, while the LogFc threshold was set to 0.25.

2.4 Cancer preferences analysis

In order to evaluate the predilection of MCs subtypes for cancer, odds ratios were computed utilizing the calculation methodology (38).

2.5 Trajectory analysis of MCs subtypes

The slingshot R package was employed to deduce cellular lineages and pseudotimes. It delineated the structure of lineages through clustering-based minimum spanning trees and employed synchronized master curves to model branching trajectories for these lineages. The "getCurves" function was utilized to acquire refined trajectory curves. The association between gene expression and pseudotime was characterized by modeling the noise distribution of each gene through a generalized additive model with negative binomials. This approach allowed for the simulation of genes exhibiting a gradual alteration in expression throughout the pseudotime continuum (39).

2.6 Assessment of cell stemness

AUCell (40) represents a novel approach to discerning cells harboring active genes within single-cell RNA-seq datasets. Given a gene set as input, it provides an evaluation of the "activity" exhibited by that particular gene set in each individual cell. In the context of this study, AUCell was employed to quantitatively assess the level of stemness exhibited by various subtypes of MCs. To hypothesize the temporal trajectory of cell differentiation, the CytoTRACE R package was utilized (41).

2.7 Enrichment analysis of cellular subtypes

By leveraging the Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Genome Enrichment Analysis (GSEA) tools, available at http://software.broadinstitute.org/gsea/msigdb, within the Cluster Profiler R package (42–44), we carried out enrichment analysis on the DEGs. To discern the disparities among various risk groups within the bulk data, the DESeq2 R package was applied, employing a threshold of |logFC| > 2 and a p-value threshold below 0.05.

2.8 Cell communication analysis

The CellChat R package (45) was used to analyze complex cell-to-cell interactions and develop regulatory networks based on ligand-receptor expression. The "netVisual DiffInteraction" function was applied to depict differences in communication strength among cells, and the "IdentifyCommunicationPatterns" function was utilized to estimate the number of communication patterns. A significance threshold of 0.05 was set. Various visualizations, including circle plots, bubble plots, and violin plots, were used to represent the incoming and outgoing signals of all cells

2.9 Scenic analysis

In evaluating the transcriptional activity within diverse subtypes of tumor cells, we employed the SCENIC analysis with Python.

2.10 Cell culture

Cell lines TE-10 and KYSE-30 were acquired from the American Type Culture Collection. The TE-10 cell line was grown under standard conditions (37°C, 5% CO_2, 95% humidity) in RPMI1640 media with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin. KYSE-30 cell line was grown under standard conditions (37°C, 5% CO_2, 95% humidity) in RPMI1640 media with 10% FBS, 1% penicillin-streptomycin, and 1% sodium pyruvate.

2.11 Cell transfection

EGFR knockdown was accomplished through the use of GenePharma (Suzhou, China) small interfering RNA (siRNA) constructs. According to Lipofectamine 3000 RNAiMAX (Invitrogen, USA) manufacturer's instructions, transfection was carried out. Two knockdown constructs (Si-*EGFR*-1 and Si-*EGFR*-2) and a negative control (si-NC) were transfected into cells that had been plated at 50% confluency in six-well plates. Every transfection was carried out using Lipofectamine 3000 RNAiMAX (Invitrogen, USA).

2.12 Cell viability assay

Using the CCK-8 assay, the cell viability of transfected AGS and SGC-7901 cells was evaluated. After being cultivated for 24 hours, cells were planted at a density of 5×10^3 cells per well in 96-well plates. Following the addition of 10μL of CCK-8 reagent (A311-01, Vazyme) to each well, the plates were incubated for two hours at 37°C in the dark. On days 1, 2, 3, and 4 post-transfections, absorbance at 450 nm was measured using a microplate reader (A33978, Thermo). Plotting of the mean OD values was done.

2.13 5-Ethynyl-2′-deoxyuridine proliferation assay

In 6-well plates, 5×10^3 cells were planted per well with transfected CNE2 and HNE2 cells, and they were grown for an entire night. A 2x EdU working solution was then created by combining serum-free medium with 10 mM EdU. Following two hours of incubation at 37°C, the cells were rinsed with PBS, fixed for thirty minutes with 4% paraformaldehyde, permeabilized for fifteen minutes with a solution of 2 mg/mL glycine and 0.5% Triton X-100, then stained for thirty minutes at room temperature using a solution of 1X Apollo and 1X Hoechst 33342. The measurement of cell proliferation was done by fluorescence microscopy.

2.14 Wound-healing assay

In 6-well plates, stabilized transfected cells were plated and allowed to grow to confluence. Each well was scratched with a sterile 200μL pipette tip, and then it was cleaned with PBS to get rid of any remaining cell debris before being incubated in a medium without serum. Using Image-J software, the breadth of the scratches was measured after they were photographed at 0 and 48 hours.

2.15 Transwell assay

Before the experiment, cells were fasted for 24 hours in a serum-free medium. The upper chamber of Costar plates was filled with cell suspension after being treated with Matrigel (BD Biosciences, USA), while the lower chamber was filled with media containing serum. In a cell culture incubator, the cells were incubated for 48 hours. To evaluate the cells' ability to invade, they were fixed with 4% paraformaldehyde after incubation and stained with crystal violet.

2.16 Construction and validation of the prognostic model

We determined the most important predictive genes using LASSO regression analysis and univariate Cox analysis (46–48). The risk coefficients for each prognostic gene were then determined using multivariate Cox regression analysis, allowing for the creation of a risk score model:

$$\text{Risk score} = \sum_i^n \text{Xi} \times \text{Yi}$$

X stands for the coefficient and Y represents the gene expression level. Using the "surv-cutpoint" function to compute the best cutoff value, patients were divided into two groups: low-risk and high-risk. We also used the Survival R package for survival analysis of the created risk score model and the "ggsurvplot" function (27) to depict survival curves in order to observe the prognostic outcomes in various patient cohorts. ROC curves were plotted using the timeROC R package to assess the predictive model's accuracy and calibration (49, 50).

2.17 Kaplan-Meier survival curve of selected genes

We performed a survival analysis utilizing the R packages survminer and survival. The area under the ROC curve (AUC) was calculated after generating ROC curves for 1-year, 3-year, and 5-year survival rates using the Survive and Time ROC R packages. Model validation was conducted through survival analysis and time-dependent ROC analysis. To evaluate the model, we employed a heatmap, a scatter plot of survival status, and a distribution of risk scores.

3 Result

3.1 ScRNA sequencing revealed the main cell types in the EC

To identify the major cell types during the progression of EC, we collected pericarcinoma and tumor tissue samples from three EC patients for single-cell RNA sequencing (scRNA-seq). We also checked the quality and completeness of the raw data. This included checking for missing values, outliers, or any anomalies that might affect the analysis. We excluded genes in the sample that did not meet the minimum expression threshold. For example, genes with low counts or low variability were excluded as they may not provide meaningful insights. After performing initial quality control and removing batch effects, we retained a total of 29,719 cells. We categorized these 29,719 cells into 30 cell clusters by dimensionality reduction (Figure 1A). According to the cell gene map and typical markers, 30 cell clusters were finally identified into 11 cell types, including B-Plasma cells *(IGKC)*, T-NK cells (*IL32*), mast cells (MCs, *TPSB2*), neutrophils (*S100A8*), fibroblasts (*DCN*), myeloid cells (*LYZ*), epithelial cells (EPCs, *KRT5*), proliferating-cells (*MKI67*), endothelial cell (ECs, *AQP1*), smooth muscle cell (SMCs, *MYH11*), and neurons (*NRXN1*). From the pie charts and bar graphs, we could learn that for tissues, T-NK cells accounted for the largest proportion in

10.3389/fimmu.2024.1470449

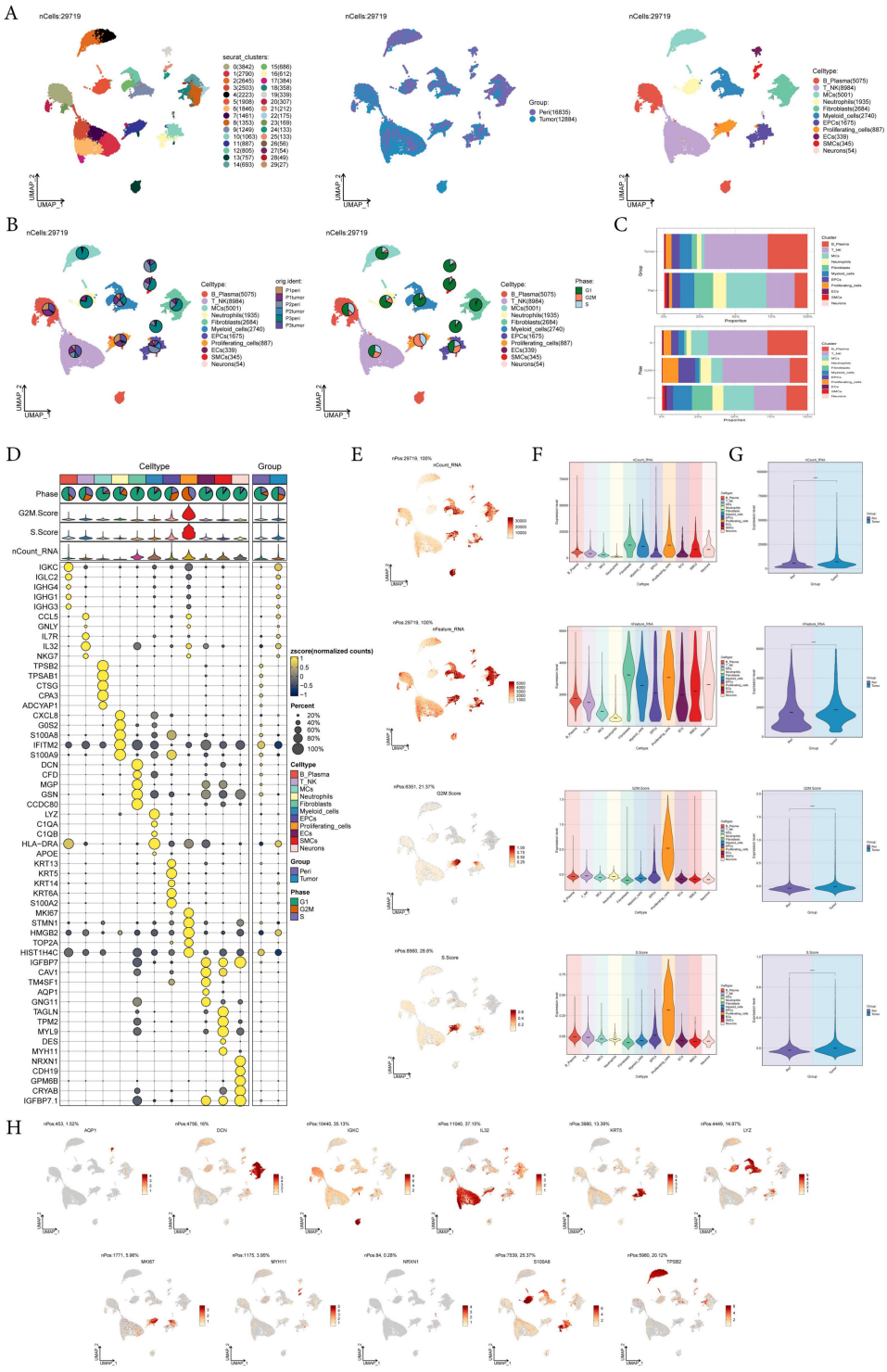


FIGURE 1
A single-cell profiling of EC, comprising 30 clusters and 11 cell types. **(A)** UMAP plot showed the 30 clusters of cells in EC patients and the number of cells in each cluster (left); UMAP plot showed the distribution of sample sources in the 11 cell types (middle); UMAP plot showed the 11 major cell types obtained by dimensionality reduction clustering of cells in EC (right). Each point corresponded to a single cell colored according to cell cluster or cell type. **(B)** The pie charts showed the proportion of different patient sources (left) and cell phases (right) in each cell type. **(C)** The bar graphs showed the proportion of different cell types in sample sources (top) and cell phase(bottom) respectively. **(D)** Bubble plot showed differential expression of top5 maker genes in EC cells across 11 cell types. Bubble colors were based on normalized data and sizes indicated the percentage of genes expressed in each cell type. **(E-G)** UMAP and violin plots revealed the expression levels of nCount-RNA, nFeature-RNA, G2M.Score, and S.Score in different cell types and sample sources. ****, $p < 0.0001$ indicated a significant difference. **(H)** UMAP plots visualized the differential genes of 11 cell types.

10.3389/fimmu.2024.1470449

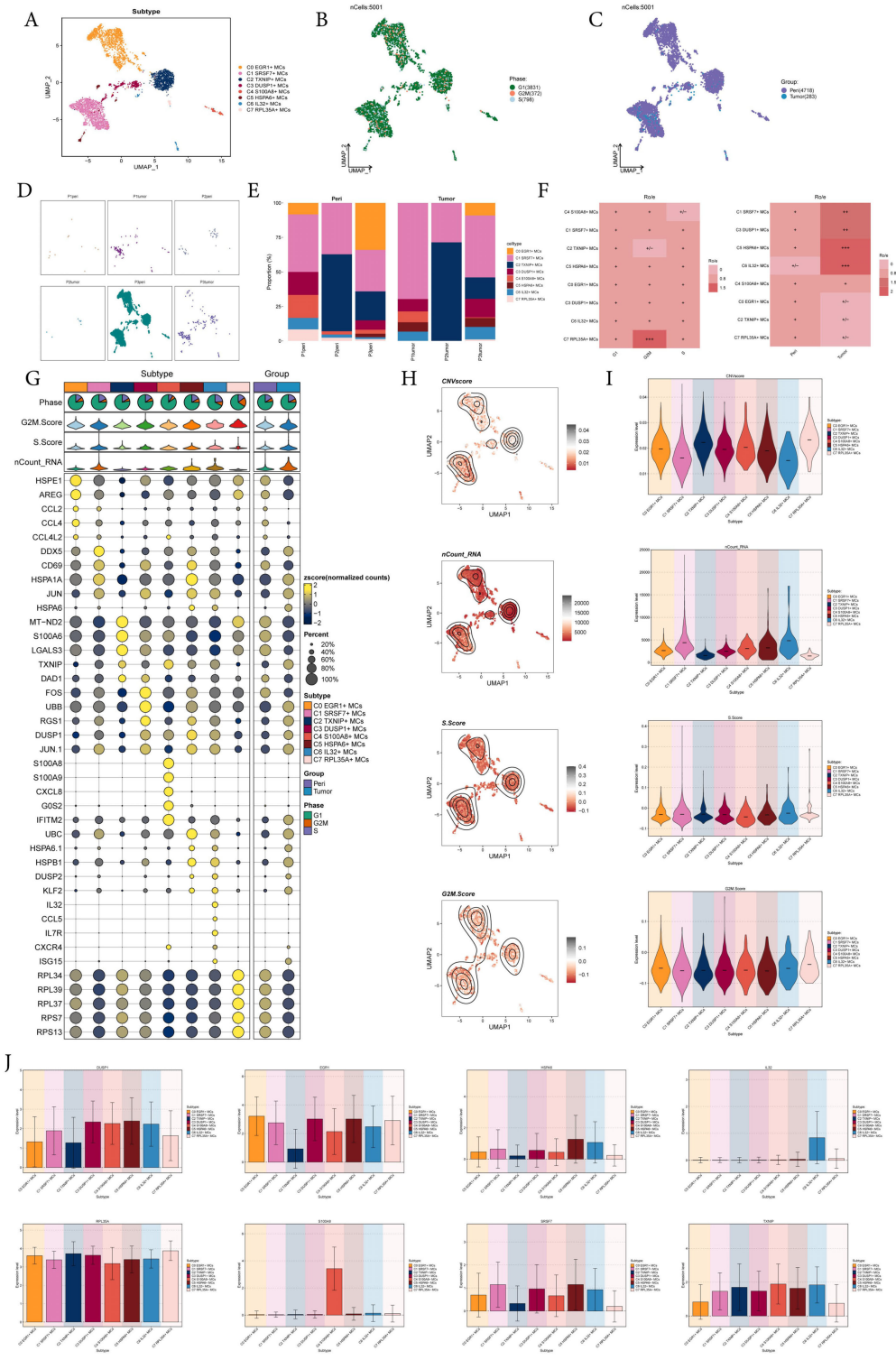


FIGURE 2
8 subtypes of MCs were identified with different markers. **(A-C)** UMAP plot demonstrated the 8 cell subtypes of MCs in EC patients and the number of cells in each cell subtype **(A)**; UMAP plot demonstrated the distribution of cell phases and sample sources in the 8 MCs subtypes respectively **(B, C)**. Each point corresponded to a single cell colored according to cell different groups. **(D)** UMAP plots showed the distribution of MCs in each patient source respectively. **(E)** The bar graphs showed the proportion of different MCs subtypes in each patient source. **(F)** Cell phases and sample sources preference of each MCs subtype estimated by Ro/e score. **(G)** Bubble plot showed differential expression of top 5 maker genes in 8 MCs subtypes. Bubble colors were based on normalized data and sizes indicated the percentage of genes expressed in each subtype. **(H, I)** UMAP and violin plots revealed the expression levels of CNVscore, nCount-RNA, S.Score and G2M.Score in different MCs subtypes. **(J)** Bar plots showed the expression levels of gene markers in each MCs subtype.

tumor tissues, followed by B-Plasma cells, while MCs were the most predominant cell type in pericarcinoma tissues; for phases, T-NK accounted for the largest proportion in both the G2M and the S phases, while on the contrary, most of MCs accounted for the largest proportion in the G1 phase (Figures 1B, C). Figure 1D showed the top 5 marker genes for 11 cell types. UMAP and violin plots were utilized to visualize nCount-RNA, nFeature-RNA, G2M.Score, and S.Score across all cells, demonstrating that proliferating cells exhibited the highest proliferative activity and vigorous division (Figures 1E-G). At the same time, the distribution of marker genes on UMAP for each cell type was presented (Figure 1H).

Among all cell types, MCs drew our attention. MCs play a crucial role in allergic reactions, pathogen immune responses during infections, angiogenesis, and the regulation of both innate and adaptive immunity. In addition to all these roles, MCs were increasingly recognized as regulators of the tumor microenvironment. Despite the accumulating evidence for MCs in tumors, their exact role in the tumor microenvironment remained incompletely understood (51). Therefore, we next performed a further analysis of mast cells.

3.2 Visualization of MCs subtypes in EC

Next, we analyzed the scRNA-seq data from tumor and pericarcinoma tissues, identified MCs, and performed further sub-clustering. This analysis resulted in eight distinct cell subtypes from a total of 5,001 mast cells: C0 *EGR1*+ MCs, C1 *SRSF7*+ MCs, C2 *TXNIP*+ MCs, C3 *DUSP1*+ MCs, C4 *S100A8*+ MCs, C5 *HSPA6*+ MCs, C6 *IL32*+ MCs, C7 *RPL35A*+ MCs (Figure 2A), and showed the distribution of phases and sample sources in the subtypes, while faceting gived a clearer picture of the distribution of MCs from different sample sources (Figures 2B-D). The bar graphs illustrated that the C1 *SRSF7*+ MCs had the highest proportion of tumor tissues of P1 and P3 origin and was enhanced over the pericarcinoma tissues share, and similarly, the Ro/e preference graph corroborated this, suggesting that the C1 *SRSF7*+ MCs was more preferred to tumor tissues (Figures 2E, F). In order to better explore the characteristics of different MCs subtypes, we visualized their typical genes. As shown in Figure 2G, C1 *SRSF7*+ MC highly express *DDX5*, which had been shown to be associated with a variety of key tumor promoting molecular interactions and was involved in tumorigenesis and tumor progression signaling pathways (52). This suggested that C1 *SRSF7*+ MCs in EC might be involved in tumor promoting effect. Several related features (CNVscore, ncount-RNA, S.Score and G2M.Score) of eight MCs subtypes were visualized (Figures 2H, I). From the Figures, we could learn that C7 *RPL35A*+ MCs had the highest expression level of CNVscore and G2M.Score, while C1 *SRSF7*+ MCs and C6 *IL32*+ MCs had the highest nCount-RNA expression level, and all subtypes had basically the same expression level of S.Score. In the end, bar plots showed the expression level of gene makers in each MCs subtype, validating the basis for delineating subtypes (Figure 2J).

3.3 Slingshot analysis of proposed temporal trajectories of MCs subtypes

To infer the lineage trajectory and pseudotime sequence of MCs, we employed slingshot analysis to assess the distribution of MCs differentiation trajectories across all MCs, visually represented through UMAP plots (Figure 3A). Then, we found 3 cell lineage trajectories of the MCs subtypes (Figures 3B-E). Including: lineage 1: C4 → C2 → C3 → C0; lineage 2: C4 → C2 → C3 → C1; lineage 3: C4 → C2 → C3 → C6. Slingshot analysis revealed that the differences among the three trajectories mainly reside in the late stages. Combined with Figures 3C-E to determine, lineage 1's endpoint was located in C0, which showed no preference for tumor tissue, lineage 3's endpoint was located in C6, which had a very small number of cells and a low percentage of tumor tissue, while lineage 2's endpoint was located in C1, which not only showed a preference for tumor tissue, but also had a high percentage of tumor tissue. Therefore, we concluded that lineage 2 represented the differentiation line of MCs associated with the tumor. In addition, we also noted that MCs are influenced by some cytokines or tumor cell-secreted proteins during development in TME, resulting in a possible transformation of the MCs phenotype to a tumor-associated or pro-tumorigenic phenotype (18), whereas C1 belonged to the terminal end and consisted predominantly of MCs originating from tumor tissues, and based on this observation, we hypothesized that C1 may play a crucial role in the differentiation of tumor-associated MCs(TAMCs) process. Subsequently, we confirmed the biological processes corresponding to the three cell lineage trajectories of MCs subtypes using GO-BP enrichment analysis (Figure 3F). It was found that C1 in lineage 2 was associated with biological processes such as endopeptidase and cysteinetype, C2 was linked to processes such as protein folding, C3 was related to leukocyte functions, and C4 was involved in processes such as lamellipodium formation, contraction, and production. Finally, the dynamic trends plot demonstrated the expression variation and distribution of marker genes for MCs subtypes along the three differentiation trajectories in pseudotime (Figure 3G).

3.4 Expression of stemness gene sets in subtypes of MCs

To investigate the expression of stem cell genes in MCs subtypes and to understand their differentiation potential, we used bubble plots to illustrate the different expression of stem cell genes in MCs subtypes. The results showed the corresponding expression of stem cell genes *KDM5B*, *EPAS1*, *CTNNB1*, *EZH2*, *KLF4*, *CD44*, *BMI1*, and *HIF1A* in MCs subtypes and different tissue types (Figure 4A). Subsequently, we visualized the cell stemness AUC scores of different MCs subtypes using a UMAP plot (Figure 4B). We then combined this with other analyses to assess the expression levels of stemness-related genes in different subtypes of MCs, and violin plots showed the different expression

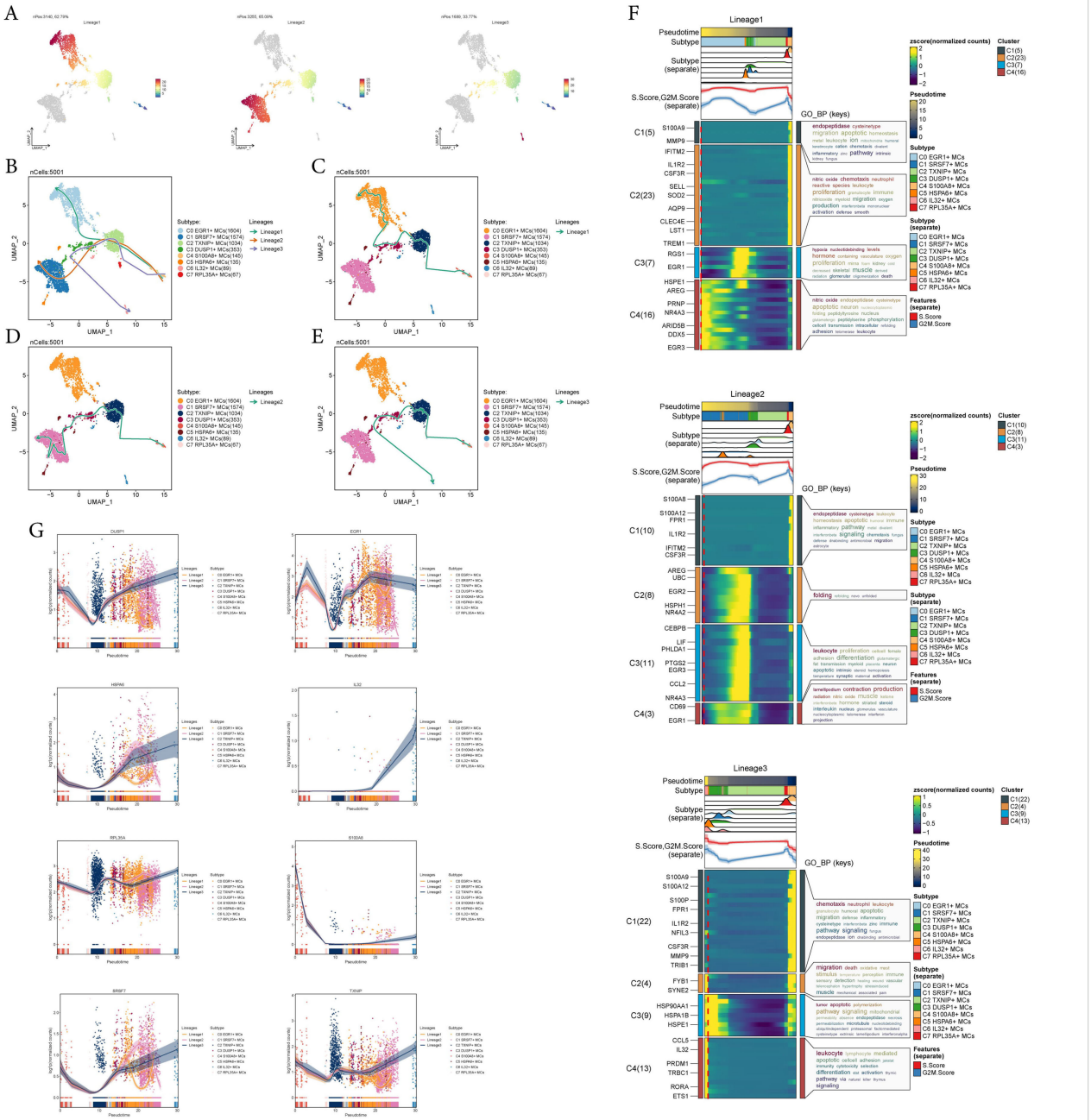


FIGURE 3
Slingshot analysis reveals three differentiation trajectories of MCs. **(A)** Demonstration of the distribution of slingshot-predicted MCs differentiation trajectories among all MCs by UMAP plot. Plotting each spectrum according to the pseudotime value to infer the result, the color from blue to red indicates the pseudotime from naïve to mature, and the grey part of the cells represent not belonging to the lineage. **(B-E)** The distribution of three differentiation trajectories of 8 MCs subtypes fitted by the pseudotime order in all mast cells **(B)**. Each trajectory was displayed respectively in UMAP **(C-E)**. **(F)** Heatmap demonstrated the related characteristics of 3 pseudotime trajectory lineages of MCs. The value of pseudotime correlated with differentiation, where 0 indicates the start point and 20 is the end point. **(G)** Scatter plots demonstrated the trajectories of named genes of 8 cell subtypes of mast cells changing on three lineages obtained after slingshot visualization.

levels of stemness genes in different sample sources, tissues, subtypes of MCs, and phases, respectively (Figures 4C-F). The results showed that C1 *SRSF7*+ MCs exhibited a higher level of cell stemness, indicating a lower degree of differentiation and higher differentiated potential; and it also showed that pericarcinoma tissues had the higher level of cell stemness. In addition, there was no significant difference in the expression levels of stemness genes in different cell phases. By CytoTRACE analysis, C1 *SRSF7*+ MCs showed the lowest degree of differentiation and the highest cell stemness among all subtypes, which we hypothesized might be related to the transformation of MCs to TAMCs (Figures 4G, H). Afterwards, the expression profiles of stemness genes with relatively elevated expression levels in Figure 4A were demonstrated in all MCs by UMAP plots (Figure 4I).

10.3389/fimmu.2024.1470449

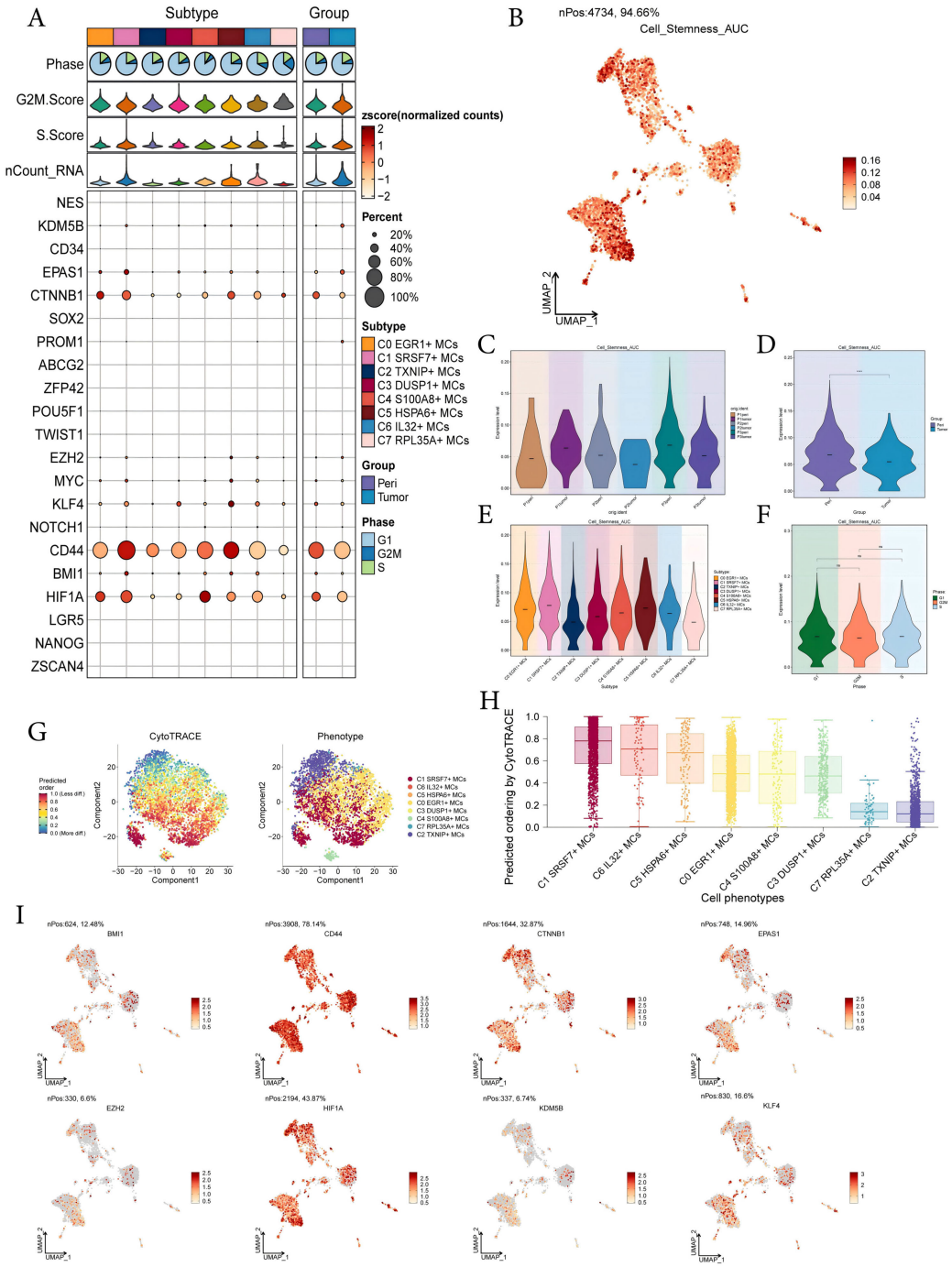


FIGURE 4
Analysis of cell stemness in mast cell subtypes. **(A)** Bubble plot showed expression of stemness genes in 8 MCs subtypes. Bubble colors were based on normalized data and sizes indicated the percentage of genes expressed in each subtype. **(B)** UMAP plot visualized the AUC values of cell stemness. **(C-F)** Violin plots revealed the expression levels of AUC values of MCs cell stemness in different patient sources **(C)**, sample sources **(D)**, subtypes **(E)** and phases **(F)**. ****,$p < 0.0001$ indicated a significant difference, ns indicated a non-significant difference. **(G)** The left panel demonstrated the distribution of MCs CytoTRACE scores. The color represented high or low cell stemness. The right panel indicated the distribution of MCs subtypes. The color represented different MCs subtypes. **(H)** Box line plot ranked the stemness of MCs subtypes according to CytoTRACE. **(I)** UMAP plots visualized the 8 stemness genes expressed in MCs subtypes.

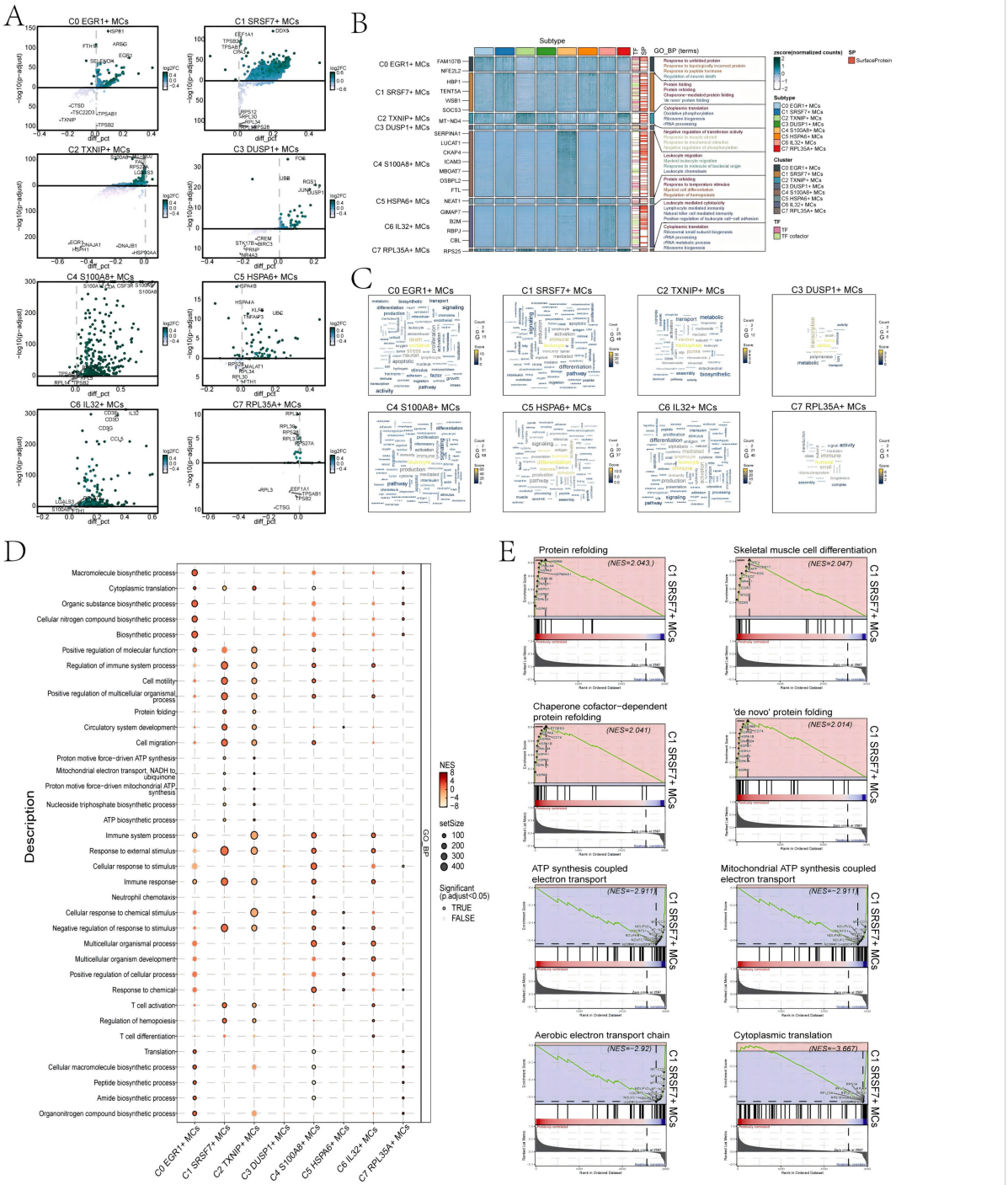


FIGURE 5
Results of functional enrichment analysis of differentially expressed genes in 8 MCs subtypes. **(A)** Volcano plots showed differentially expressed genes in 8 subtypes. **(B)** Heatmap showed the enrichment items of GO_BP scored. zscore > 0 was positive enrichment and < 0 was negative enrichment. **(C)** Word cloud diagrams demonstrated the activity of different pathways in MCs subtypes. **(D)** GSEA analysis diagram of different pathways in each MCs subtype. NES > 0 was positive enrichment and < 0 was negative enrichment. NES, N stands for standardization, and ES for enrichment scores. **(E)** GSEA enrichment analysis among C1 *SRSF7+* MCs.

3.5 Enrichment analysis of MCs subtypes in EC

First, we utilized volcano plots to represent the DEGs profiles between subtypes of MCs (Figure 5A). The results showed that the up-regulated DEGs in C1 *SRSF7+* MCs were mainly *DDX5*, *EEF1A1*, *TPSB2*, *TPSAB1*, and *CPA3*. In addition, we performed GO-BP enrichment analysis of the DEGs in the subtypes of MCs to reveal their enrichment in biological processes. The heatmap showed the results of the top four enrichment entries in the MCs

subtypes (Figure 5B). The C0 *EGR1*+ MCs subtype was mainly associated with pathways such as response to unfolded protein, response to topologically incorrect protein and regulation of neuron death; The C1 *SRSF7*+ MCs subtype was enriched in pathways such as protein folding, protein refolding, chaperone-mediated protein folding and '*de novo*' protein folding; The C2 *TXNIP*+ MCs subtype revealed their close association with cytoplasmic translation, oxidative phosphorylation, ribosome biogenesis and rRNA processing; The C3 *DUSP1*+ MCs subtype showed enrichment in pathways such as negative regulation of transferase activity, response to muscle stretch, response to mechanical stimulus and negative regulation of phosphorylation; The C4 *S100A8*+ MCs subtype was enriched in pathways related to leukocyte migration, myeloid leukocyte migration, response to molecule of bacterial origin and leukocyte chemotaxis; The C5 *HSPA6*+ MCs subtype mainly exhibited enrichment in pathways such as protein refolding, response to temperature stimulus, myeloid cell differentiation and regulation of hemopoiesis; The C6 *IL32*+ MCs subtype revealed pathways related to leukocyte mediated cytotoxicity, lymphocyte mediated immunity, natural killer cell mediated immunity and positive regulation of leukocyte cell-cell adhesion; The enrichment analysis conducted on the C7 *RPL35A*+ MCs subtype revealed their association with cytoplasmic translation, ribosomal small subunit biogenesis, rRNA processing and rRNA metabolic process. The word cloud plots illustrated the enrichment results of DEGs across various pathways in the eight MC subtypes (Figure 5C). The results showed that the C1 *SRSF7*+ MCs subtype was mainly enriched in leukocyte, immune and activation, and it was hypothesized that C1 *SRSF7*+ MCs subtype might be related to MCs activation and participation in immune regulation.

In addition, the results of GSEA enrichment analysis were also shown in the form of bubble plots (Figure 5D). It showed that C1 *SRSF7*+ MCs subtype was significantly expressed in regulation of immune system process, cell motility and migration, protein folding, response to immune and external stimulus pathways. All of the above pathways would suggest that MCs in the C1 *SRSF7*+ MCs subtype had likely transformed into TAMCs. Finally, we performed GSEA on the DEGs of the C1 *SRSF7*+ MCs subtype according to GO-BP terminology. The results were shown in Figure 5E. We observed that pathways associated with protein refolding, skeletal muscle cell differentiation, chaperone cofactor-dependent protein refolding and '*de novo*' protein folding were upregulated in the C1 subtype. In contrast, pathways associated with ATP synthesis coupled electron transport, mitochondrial ATP synthesis coupled electron transport, aerobic electron transport chain and cytoplasmic translation were downregulated in the C1 subtype. Combining the above up-regulated genes and enriched pathways with previous studies, we believed that the C1 subtype was affected by the endoplasmic reticulum stress state, which disrupts the homeostasis of the original proteins and generates aberrant protein folding, and that this stress state could control a variety of pro-tumorigenic attributes of cancer cells, dynamically re-programming the function of immune cells, transforming MCs into TAMCs, thus exerting pro-tumorigenic effects, and conferring a greater tumorigenic, metastatic, and drug-resistant capacity to the malignant cells (53).

3.6 Transcription factors regulate the carcinogenic mechanism of C1 SRSF7+ MCs

Transcription factors can directly act on the genome and regulate gene transcription and affect the biological function of cells by combining specific nucleotide sequences in the upstream of the gene. Therefore, we used scenic to analyze the gene regulatory network of C1 *SRSF7*+ MCs. First of all, we carried out cluster analysis of MCs according to regulator activity (Figure 6A). It was obvious that the discretization of UMAP diagram based on regulator activity was smaller, the interference factors were better excluded, and all MCs were clustered and distributed. Among them, C1 *SRSF7*+ MCs were mainly distributed on the right side of UMAP plot without significant discretization. By further analyzing the key regulators of different MCs subtypes, the five major regulators of C1 *SRSF7*+ MCs, ATF4, JUNB, NFκB2, MAFK and JUN, were identified (Figures 6B, C). After analyzing these five key regulators in depth in conjunction with previous studies and Figure 6D, ATF4 and JUNB caught our attention. ATF4, which was expressed at higher levels in C1 *SRSF7* + MCs than in other subtypes, was a major transcriptional regulator of the unfolded protein response to hypoxia, activated genes that promoted recovery of normal endoplasmic reticulum function and hypoxic survival (54), regulated mast cells through endoplasmic reticulum stress (55), and had been associated with programmed cell death in a variety of tumors, particularly ER stress-induced iron death (56, 57, 86). As for JUNB, its expression level was high in C1 *SRSF7*+ MCs, C4 *S100A8*+ MCs and C5 *HSPA6*+ MCs subtypes, and it is a potent inhibitor of endoplasmic reticulum stress and apoptosis, and, in particular, its modulation of endoplasmic reticulum stress is associated with ATF4 alterations (58).

3.7 CellChat analysis among all cells

In order to systematically elucidate complex cellular responses, we aimed to investigate cell-to-cell relationships and ligand-receptor communication networks to better understand interactions between cells. Using CellChat analysis, we initially established intercellular communication networks involving various cells such as tumor cells, fibroblasts, T-NK cells, and different subtypes of MCs, etc (Figure 7A). After establishing the intercellular communication networks using CellChat analysis, we calculated both the number of interactions (represented by the thickness of the connecting lines between two cell types) and the strength of interactions (indicated by the weight of the lines, where thicker lines denote stronger interaction strengths). This approach helped quantify the complexity and intensity of communication pathways between different cell types in the

10.3389/fimmu.2024.1470449

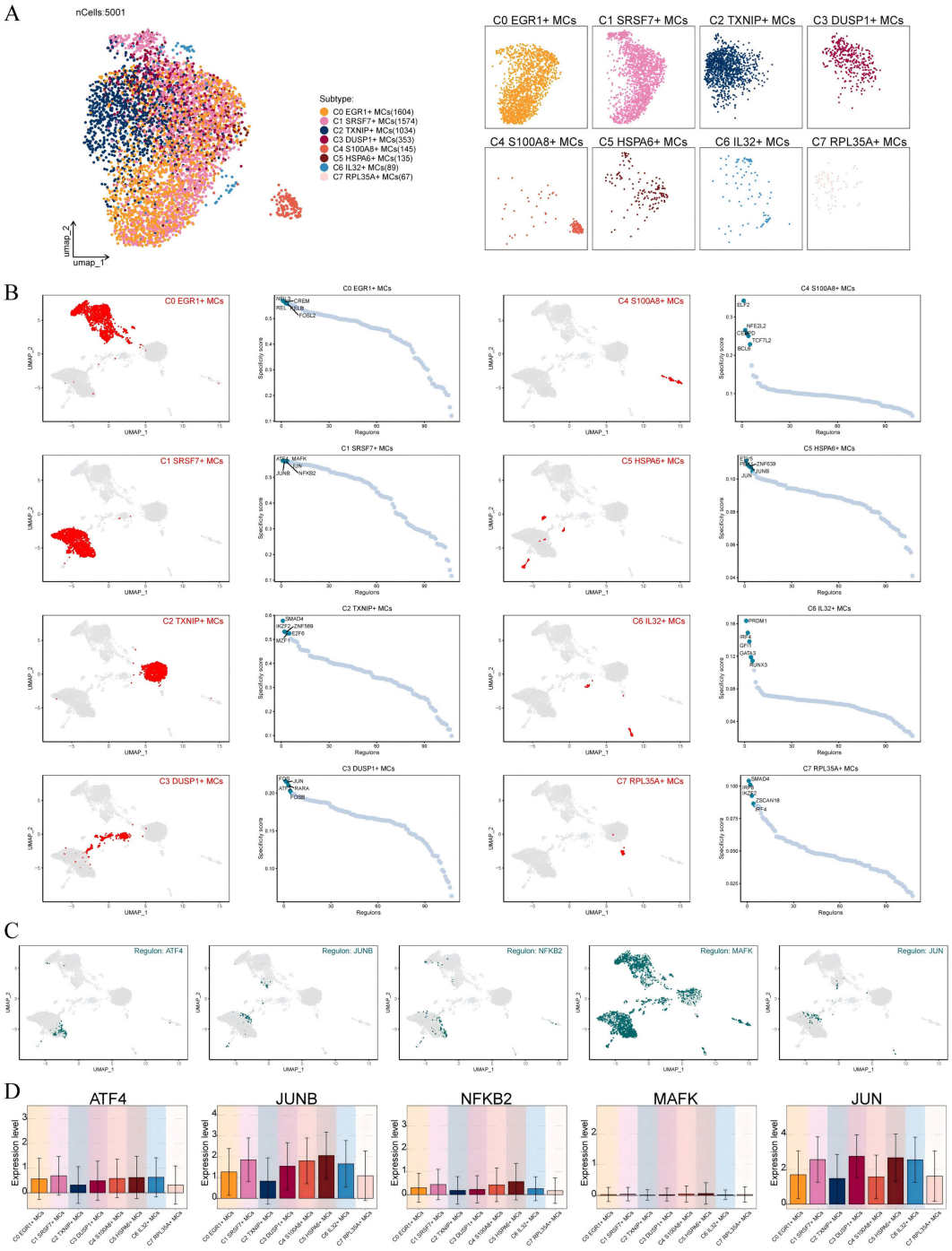


FIGURE 6
Identification of C1 SRSF7+ MCs Gene Regulatory Network. **(A)** UMAP visualized all MCs based on regulon activity score. Colored according to cell subtypes. **(B)** Different MCs subtypes were highlighted in the UMAP plots (red) (left); rank for regulators in different MCs subtypes based on regulon specificity score (RSS) (green) (right). **(C, D)** Expression of transcription factors ATF4, JUNB, NFKB2, MAFK and JUN of C1 *SRSF7+* MCs in different MCs subtypes.

network. We utilized gene expression pattern analysis methods available through CellChat to investigate how cells and signaling pathways interact. Initially, we assessed the relationship between inferred potential communication patterns and groups of cells that secrete signaling molecules to decipher outgoing communication patterns. Three distinct signaling patterns were identified through our analysis: pattern 1 (subtypes of MCs), pattern 2 (Neurons cells, fibroblasts, SMCs, tumor-cells and ECs) and pattern 3 (myeloid-cells, B-Plasma cells, neutrophils, proliferating-cells and T-NK cells) (Figure 7B). To identify the key incoming and outgoing signals associated with the eight MCs subtypes, we quantitatively analyzed the ligand-receptor network using CellChat. This approach allowed us to predict the primary incoming signals from secreting cells (signal senders) releasing

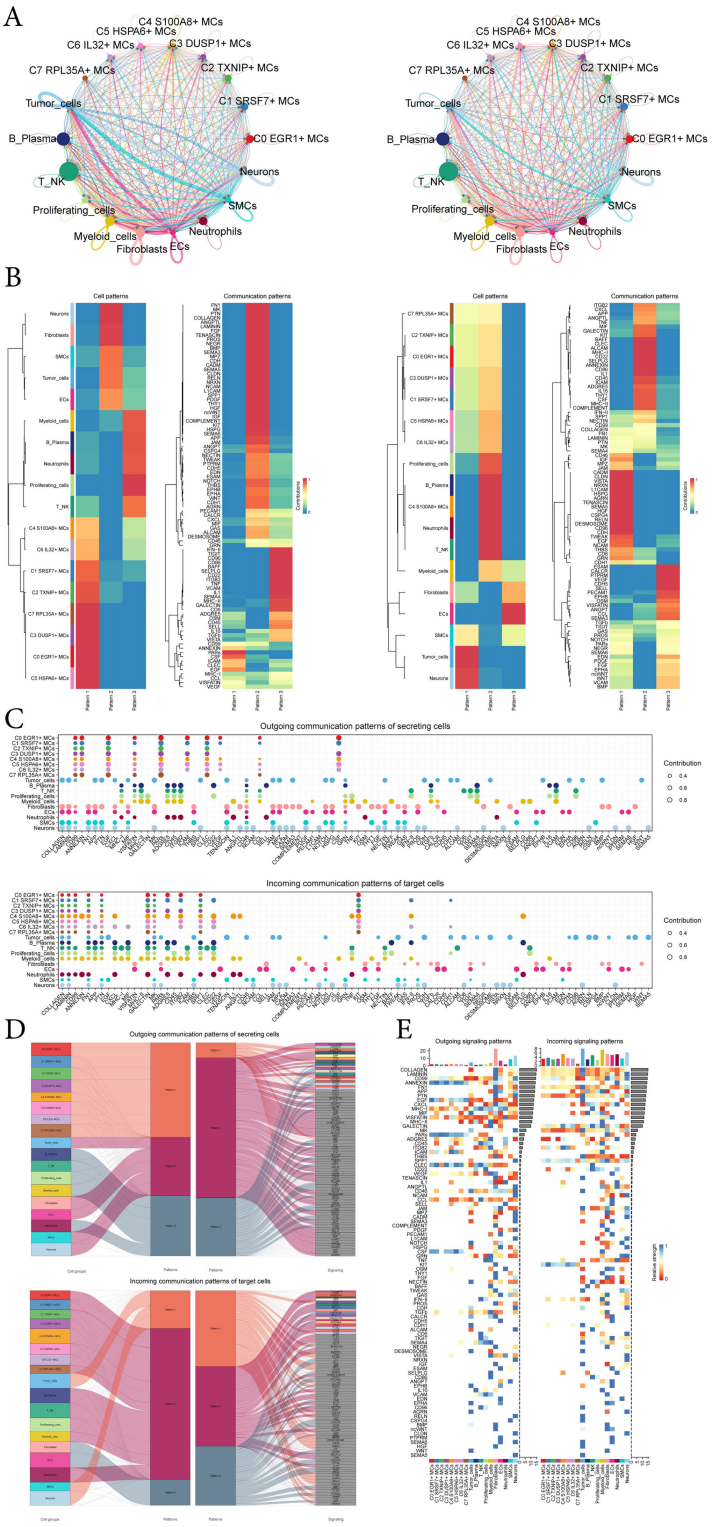


FIGURE 7
CellChat analysis among all cells. **(A)** Circle plots showed the number (left) and strength (right) of interactions between all cells. **(B)** Heatmap showed pattern recognition of outgoing cells (left), and incoming cells (right) among all cells. **(C)** Outgoing contribution bubble plot and incoming contribution bubble plot demonstrated the communication patterns between the secreting cells and target cells of EC, the color of the dots indicated different cells and the size of the dots indicated the contribution of cells. **(D)** Sankey diagrams showed inferred outgoing communication patterns of secreting cells and incoming communication patterns of target cells, as well as correspondence between inferred potential patterns, cell groups, and signaling pathways. The color and width of the branches represented the type and strength of the communication. **(E)** Heatmap showed ligands and receptors related to the incoming and outgoing signals of cell interactions.

10.3389/fimmu.2024.1470449

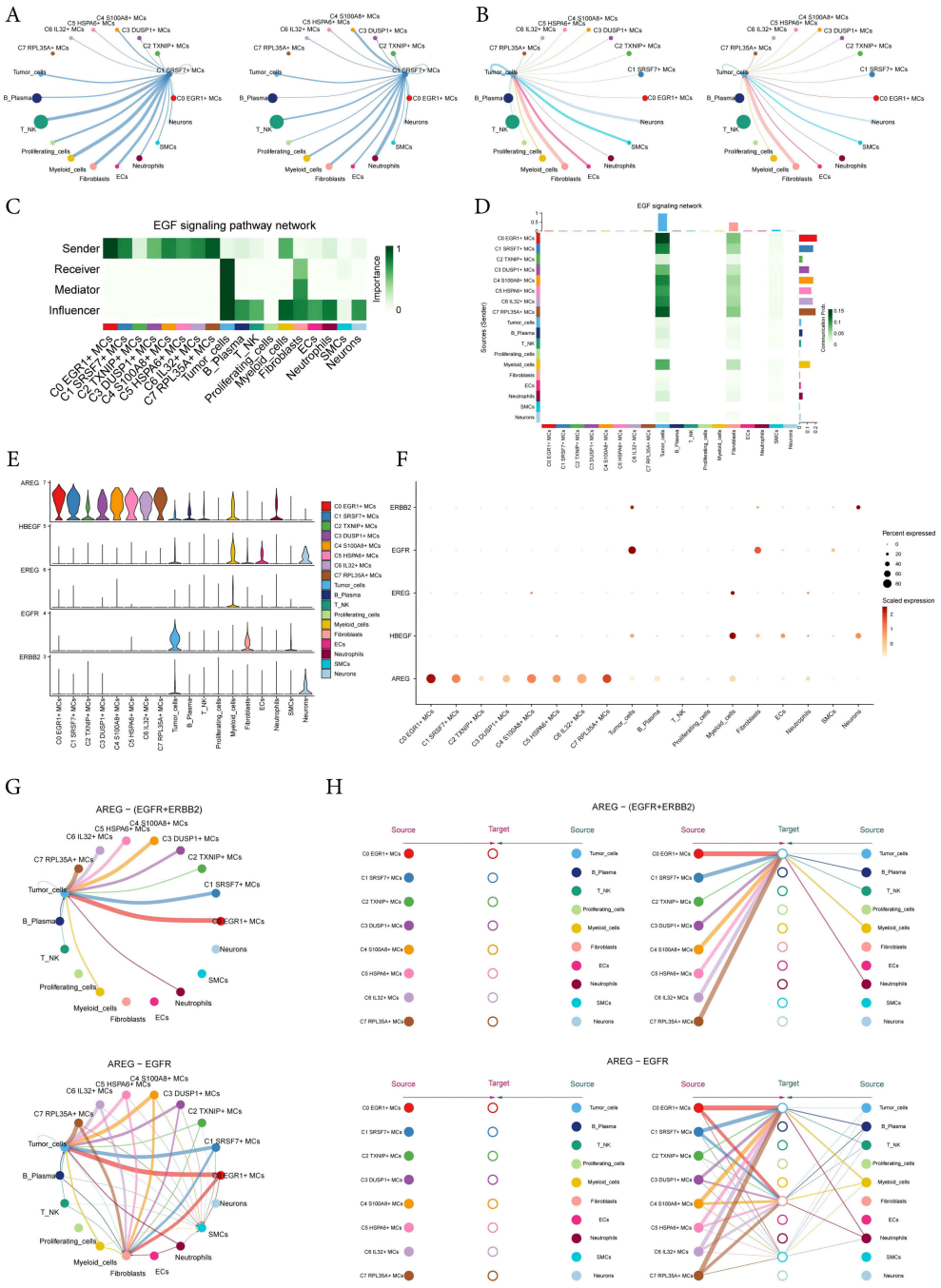


FIGURE 8
Visual analysis of AREG-EGFR/AREG-(EGFR+ERBB2) signaling pathway. **(A, B)** The number (left) and strength (right) of cellular interactions circled plots with C1 *SRSF7*+ MCs as source **(A)** and tumor as target **(B)**. **(C)** Heatmap demonstrated the centrality score of the EGF signaling pathway network, showing the relative importance of each cell group. **(D)** Heatmap showed the cell interactions of the EGF signaling pathway. **(E, F)** Violin and bubble plots demonstrated cellular interactions in the EGF signaling pathways. **(G–H)** Circle plot and hierarchical plots showed the inferred intercellular communication network for EGF signaling. Solid and hollow circles indicated source and target cell types in hierarchical plots, respectively. The edge color of the middle circle in hierarchical plots was consistent with the signal source.

various cytokines or ligands. Additionally, we assessed which cell types acted as targeting cells (signal receivers), and how ligand-receptor-mediated communications between different cell types contributed to the progression of EC. This analysis helped illustrate how receptors on these cells were targeted by ligands released either from the same type of cell or from other cell types (Figure 7C).

In addition to examining detailed communication within individual pathways, an important aspect was understanding how multiple cell populations and signaling pathways

coordinate their functions. To address this, CellChat employed a pattern recognition method based on nonnegative matrix decomposition. This method identified global communication patterns and key signals across different cell groups, shedding light on how various cells and pathways collaborate in their functions. The application of this analysis revealed three distinct incoming signaling patterns and three outgoing signaling patterns. For instance, this output indicated that the majority of outgoing MCs signaling was characterized by pattern 1, which represented multiple pathways, including but not limited to CD99, ANNFXIN, EGF, PARs, ICAM, CSF, etc. All output tumor-cells, fibroblasts, ECs, SMCs, neurons signalings were characterized by pattern 2, which represented pathways such as COLLAGEN, LAMININ, FN1, APP, PTN and so on. On the other hand, the analysis of communication patterns in target cells indicated that incoming signalings to tumor-cells, SMCs, and neurons were predominantly characterized by pattern 1. This pattern included signaling pathways such as EGF, TENASCIN, JAM, MPZ, CADM, and TWEAK. In contrast, the majority of incoming signalings to subtypes of MCs, B-plasma cells, T-NK cells, proliferating-cells, myeloid-cells, and neutrophils were characterized by pattern 2, which was driven by pathways such as CXCL and ANNEXIN (Figure 7D).

Combining the above analysis and the demonstration of all incoming and outgoing signal intensities in Figure 7E, the signaling molecule EGF caught our attention. EGF was present in the incoming pathway of tumor-cells, i.e., tumor-cells were the target cells, and EGF is again present in the outgoing pathway of C1 *SRSF7*+ MCs subtype, i.e. C1 *SRSF7*+ MCs subtype was the secreting cell, which links C1 *SRSF7*+ MCs subtype and tumor-cells, we speculated that this signaling pathway might be related to tumor progression, so we next focused on EGF.

3.8 Analysis of AREG-EGFR/AREG-(EGFR +ERBB2) signal pathway

The circular displayed the inferred cell-cell communication network between MCs and other cells (Figures 8A, B). The results showed that there was a strong crosstalk between C1 *SRSF7*+ MCs and tumor cells. We considered all identified cell types in ECEC as source cells for the EGF signaling pathway, and the results indicated that all subtypes of MCs could target tumor cells with released EGF. In addition to the senders and receivers of EGF signaling, based on the relative importance of each cell type in EGF signaling-mediated intercellular communication, we identified the cell types that act as mediators and influencers in this process, which is referred to as the "centrality measurement" algorithm. As can be seen from the Figure, C1 *SRSF7*+ MCs subtype had higher expression as a 'sender' in the EGF signaling pathway, whereas tumor-cells were acting as 'receiver', 'mediator' and 'influencer' in this signaling pathway (Figure 8C). Similarly, the heatmap corroborated this conclusion (Figure 8D). The violin plot showed the cell-cell interactions while giving the different ligands and receptors in the EGF signaling pathway, and the results showed that C1 *SRSF7*+ MCs subtype and tumor-cells were mainly contacted with *AREG* as a ligand and *EGFR* or *ERBB2* as receptors (Figure 8E). Bubble and circle plot as well as hierarchical plots likewise corroborated this conclusion (Figures 8F–H). Combined with the results of previous results in this paper, it can be concluded that the C1 *SRSF7*+ MCs and tumor cells crosstalk through the AREG-EGFR/AREG-(EGFR+ERBB2) signal pathway, thereby exerting a tumor-promoting effect.

3.9 *In vitro* experimental validation of EGFR

To further investigate the role of *EGFR* in EC, we conducted *in vitro* experiments using the TE-10 and KYSE-30 cell lines. Initially, we knocked down *EGFR* and measured the mRNA and protein expression levels before and after knockdown. We observed a significant reduction in both mRNA and protein expression levels in both cell lines compared to the control group (Figure 9A). Subsequently, the CCK-8 assay revealed a marked decrease in EC cell viability post-*EGFR* knockdown (Figure 9B). Colony formation assays and EDU experiments confirmed that *EGFR* knockdown inhibited EC cell proliferation (Figures 9C, E, F). Additionally, scratch and transwell assays were employed to assess the migration and invasion capabilities of EC cells post-*EGFR* knockdown, demonstrating a significant reduction in migration and invasion levels (Figures 9D, F-H). These results collectively indicate that *EGFR* knockdown suppresses the activity, proliferation, migration, and invasion of EC cells, thereby inhibiting tumor growth.

3.10 Enrichment analysis and construction of predictive models

To further investigate the impact of MCs with high *SRSF7* expression on EC patients, we divided the TCGA cohort patients into high and low SMRS (SMRS: *SRSF7*+MCs risk score) groups according to the gene expression levels of the *SRSF7*+ MCs subtype. A heatmap illustrated the expression profiles of the top 30 DEGs (Figure 10A), and a volcano plot depicted the up-regulation and down-regulation of DEGs (Figure 10B). Subsequently, we employed various enrichment methods to gain insights into the associated biological processes. KEGG enrichment analysis revealed that DEGs were primarily enriched in pathways such as cholesterol metabolism, PPAR signaling pathway, and Fat digestion and absorption (Figure 10C). In GO-BP analysis, enrichment was observed in the triglyceride metabolic process, acylglycerol metabolic process, and neutral lipid metabolic process (Figure 10D). In GO-CC analysis, enrichment included chylomicron and high- and low-density lipoproteins, and in GO-MF analysis, glycosaminoglycan binding and lipoprotein particle receptor binding were highlighted (Figures 10E, F). We then visualized the primary enrichment terms for each gene set and used t-SNE plots to graphically represent the risk score distribution of these enrichment terms (Figure 10G). GSEA results showed that the up-regulated genes were mainly enriched in processes such as intestinal absorption, peptidyl methionine modification, intestinal lipid absorption, and protein lipid complex assembly, while down-regulated genes were enriched in processes like regulation of release

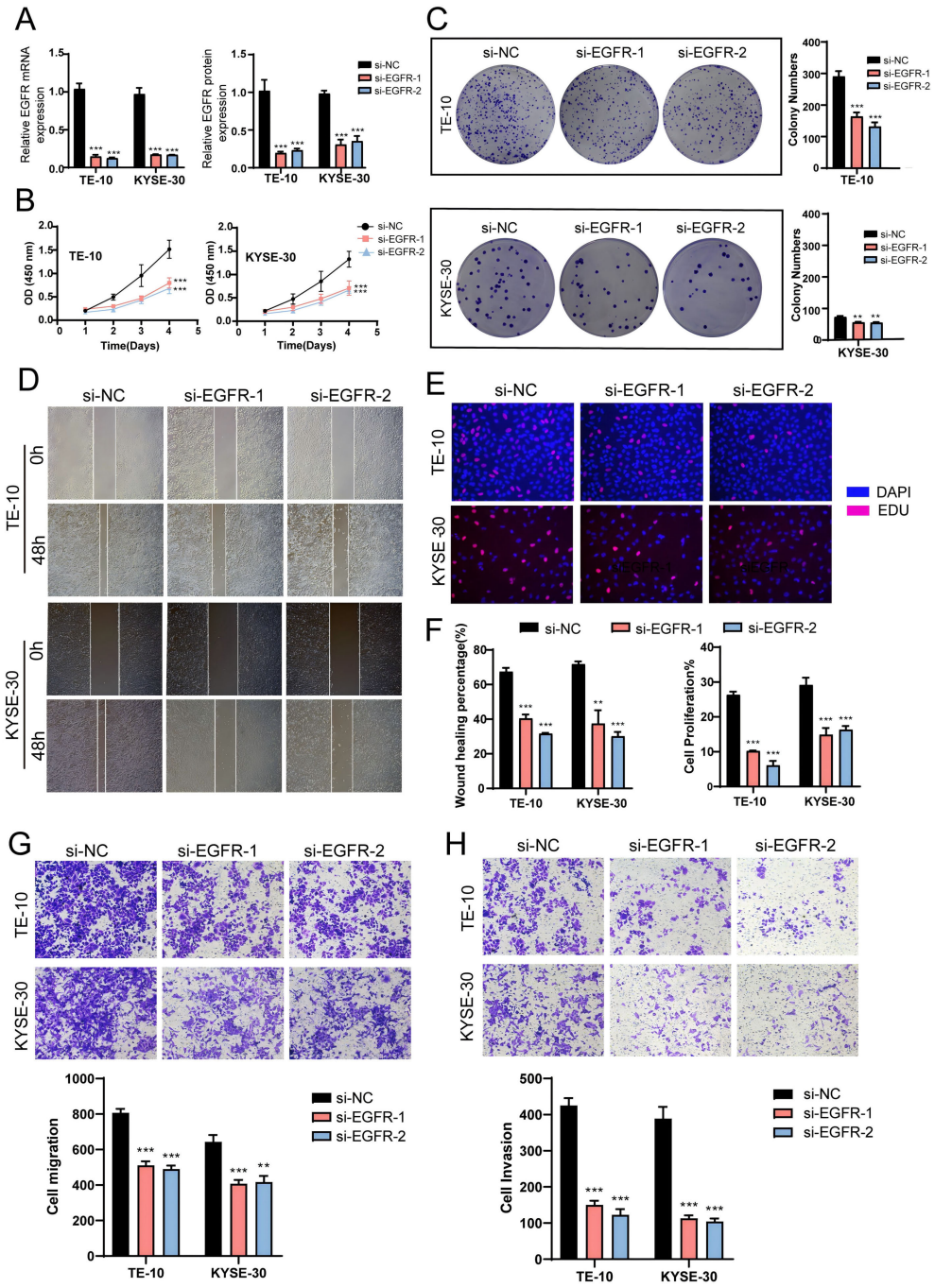


FIGURE 9
In vitro experiments confirmed the effects of *EGFR* knockdown. **(A)** Following *EGFR* knockdown, both mRNA and protein expression levels were significantly reduced. **(B)** The CCK-8 assay demonstrated a marked decrease in EC cell viability post-*EGFR* knockdown compared to the control group. **(C)** Colony formation assays revealed a significant reduction in colony numbers after *EGFR* knockdown. **(D)** The scratch assay indicated that *EGFR* knockdown inhibited EC cell migration. **(E)** The EDU staining assay confirmed that *EGFR* knockdown exerted an inhibitory effect on EC cell proliferation. **(F)** Bar graphs showed a significant reduction in both EC cell migration and proliferation capabilities post-*EGFR* knockdown ($P < 0.01$). **(G, H)** Transwell experiments indicated that *EGFR* knockdown inhibited the migration and invasion capabilities of tumor cells in the TE-10 and KYSE-30 cell lines. ***, $p < 0.001$; ****, $p < 0.0001$ indicates significant difference.

of sequestered calcium on into cytosol, external encapsulating structure organization, B cell receptor signaling pathway, and collagen fibril organization (Figure 10H). Additionally, we constructed a prognostic model to explore the clinical significance of MCs with high *SRSF7* expression. Univariate Cox regression analysis identified 11 genes significantly associated with prognosis (Figure 10I), with AHR as a protective factor (HR < 1) and the others as risk factors. To address the issue of multicollinearity among these genes, we further screened them using LASSO regression analysis, ultimately identifying eight prognostic-related

10.3389/fimmu.2024.1470449

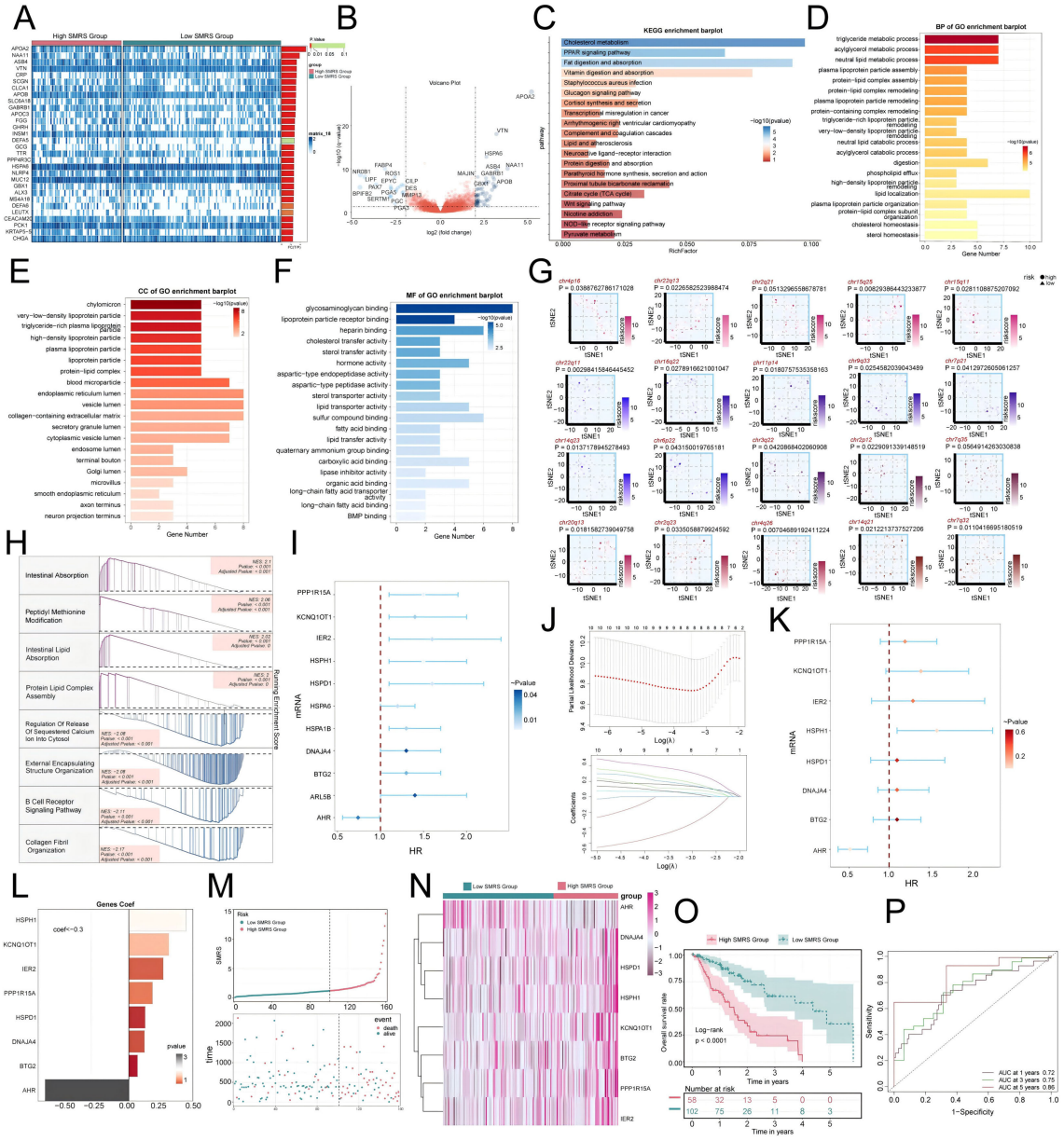


FIGURE 10
Enrichment analysis of differential genes and construction of the prognostic model. **(A)** Heatmap illustrated the expression profiles of differential genes in high and low SMRS groups. **(B)** Volcano plot depicted the distribution of differential genes in high and low SMRS groups. **(C-F)** Bar charts separately presented the enrichment analysis results of differential genes in KEGG, GO-BP, GO-CC, and GO-MF pathways for high and low SMRS groups. **(G)** t-SNE plot visualized the risk score distribution of the top-ranked GSVA enrichment term in high and low SMRS groups. **(H)** Detailed exposition of GSEA pathway enrichment results for differential genes across various pathways was provided. **(I)** Forest plot from univariate Cox regression analysis showcased statistically significant genes (P<0.05) with HR<1 indicating protective factors and HR>1 indicating risk factors. **(J)** Selection of eight prognostic-related genes (non-zero regression coefficients) was made via LASSO regression analysis, with optimal parameter (lambda) determined through cross-validation (top), and LASSO coefficient curve determined by optimal lambda (bottom). **(K)** Forest plot of eight prognosis-related genes. **(L)** Bar chart showed the Coef values of genes utilized for model construction. **(M)** Curve chart illustrated the risk scores of high and low SMRS groups, and scatter plot depicted survival/death events over time for both groups. **(N)** Heatmap displayed differential expression of model genes, with color scale based on normalized data. **(O)** Kaplan-Meier curves demonstrated survival disparities between high and low SMRS groups. **(P)** ROC curve and AUC value were used to evaluate the sensitivity and specificity of the prognostic model in predicting 1-year, 3-year and 5-year prognosis.

genes (Figure 10J). Cox regression analysis was then used to calculate the coefficient values of these genes (Figures 10K, L). Curve and scatter plots demonstrated the differences in risk scores and survival outcomes between the two groups, indicating that the high SMRS group was associated with poorer prognosis (Figure 10M). Moreover, a heatmap displayed the differential expression patterns of genes used in model construction (Figure 10N). Kaplan-Meier survival curves further confirmed the conclusion that the high SMRS group had a worse survival outcome (Figure 10O). ROC curves and AUC values for 1-year, 3-year, and

5-year outcomes indicated that the model had good predictive value (Figure 10P).

4 Discussion

In recent years, the rapid development and application of scRNA-seq in cancer research has revolutionized our understanding of the biological features and dynamics within cancer lesions, greatly facilitating the diagnosis, treatment, and prognosis prediction of a range of tumors (59–61). Overall, the present study focused on mast cells in esophageal cancer, and we validated the pro-carcinogenic role of this pathway by launching a comprehensive profiling of mast cell subtypes with an eye on the C1 *SRSF7*+ MCs and obtaining its reciprocal receptor, *EGFR*, using cellular communication analysis, and subsequently verifying the pro-carcinogenic role of this pathway through cellular knockdown experiments. In this study, we comprehensively characterized the cellular heterogeneity of EC using scRNA-seq technology. We identified immune cells including T-NK cells, MCs, and myeloid cells and so on, as well as non-immune cells such as smooth muscle cells and neuronal cells. In addition, we carefully analyzed the sample origin of these cell types and the distribution characteristics during the phase. Among them, MCs caught our attention. Until more than a hundred years ago, MCs were regarded as effectors of allergy, and it is only in the last two decades that MCs have gained recognition for their involvement in other physiological and pathological processes. MCs maturation, phenotype and function as a direct result of the local microenvironment (62), and by releasing a range of bioactive mediators has a significant effect on their ability to specifically recognize and respond to a variety of strategies (63–65). Therefore, depicting and analyzing the TME is important for MCs. And in previous studies, MCs have been shown to correlate with pro-tumorigenic effects (66–68). Despite the accumulating evidence for MCs in tumors, their exact role in the TME remains incompletely understood (51). We therefore focused our attention on the study of MCs. By further dimensionality reduction clustering, we obtained eight MCs subtypes, i.e., C0 *EGR1*+ MCs, C1 *SRSF7*+ MCs, C3 *TXNIP*+ MCs, C4 *S100A8*+ MCs, C5 *HSPA6*+ MCs, C6 *IL32*+ MCs, and C7 *RPL35A*+ MCs.

By integrating the proportions of MCs subtypes in sample sources and cell phases, Ro/e preference analyses, cell stemness analyses, and slingshot proposed pseudotime analyses, we identified the target subtype in this study: the C1 *SRSF7*+ MCs. C1 *SRSF7*+ MCs were significantly more abundant in tumor tissues than in pericancer tissues in P1 and P3 samples, and this was confirmed by Ro/e preference analysis. In slingshot proposed pseudotime analysis, Lineage 2 was considered to be representative of the differentiation trajectory of MCs associated with tumors. And the endpoint of Lineage 2 was a subtype of C1 *SRSF7*+ MCs, this result may prove that MCs are affected by some cytokines or tumor cell-secreted proteins during their development in the TME, leading to the transformation of MCs into a tumor-associated or pro-tumor phenotype, which is in line with the previous study (69). Meanwhile, cell stemness analysis by AUC value and CytoTRACE showed that the C1 *SRSF7*+ MCs subtype had the strongest cell stemness among all subtypes, with high differentiation potential, which did not contradict slingshot's results, and it is understandable that the transformation from normal phenotype to TAMCs phenotype would result in an increase in cell stemness. It can be seen that the C1 *SRSF7*+ MCs subtype is intricately linked to tumor progression.

To further investigate the tumor-promoting related roles of the C1 *SRSF7*+ MCs subtype, we performed enrichment analysis and obtained the upregulated genes *DDX5*, T*PSB2*, and *CPA3*, of which *DDX5* interacts with a variety of key pro-tumorigenic molecules and participates in tumorigenic and tumor progression signaling pathways, and when *DDX5* is expressed or its post-translational modifications are deregulated, the normal cellular signaling network collapses, leading to many pathological states, including tumorigenesis and tumor progression (52, 70). Moreover, the enriched pathways obtained by GO-BP and GSEA on the C1 *SRSF7*+ MCs subtype showed that the C1 *SRSF7*+ MCs subtype was extensively involved in protein folding and refolding, regulation of immune system processes, and response to external stimuli. All these pathways suggest that the C1 *SRSF7*+ MCs subtype has probably been transformed into TAMCs. Finally, combining the above up-regulated genes and enriched pathways, we suggest that the C1 *SRSF7*+ MCs subtype is affected by the endoplasmic reticulum stress state (71), which disrupts the original protein equilibrium (72) and produces aberrant protein folding (73, 74), and this stress state dynamically reprograms the function of MCs, transforming MCs into TAMCs, which exerts pro-tumorigenic effects (75) and confers cancer cells with enhanced tumorigenic, metastatic, and drug-resistant capabilities. In this regard, we can treat patients with esophageal cancer by targeting the abnormal protein folding to prevent MCs from entering the endoplasmic reticulum stress state in patients, thus preventing the conversion of MCs into TAMCs, and thus controlling the progression of the cancer.

In addition, gene regulatory network of C1 *SRSF7*+ MCs was revealed by scenic analysis, in which the most valuable key regulators were ATF4 and JUNB. ATF4 showed a dual role in iron death and cancer under endoplasmic reticulum stress (75), and under sustained stress conditions, ATF4 promotes apoptotic cell death induction. Characterizing the mechanisms that regulate ATF4-mediated transcription and its effects on cellular metabolism may identify novel targets for cancer therapy (56). As for JUNB, more and more studies have shown that it is involved in tumorigenesis by regulating cell proliferation, differentiation, senescence, and metastasis, and in particular, it affects the TME by transcriptionally promoting or repressing oncogenes in tumor cells or immune cells (76). Furthermore, previous mechanistic studies have shown that JUNB overexpression regulates the mitochondrial apoptosis pathway, mediating resistance to FasL and TRAIL-induced cell death, and thus tumor resistance to

immunotherapy (77). This study of ours provides a theoretical basis for subsequent analysis of drug sensitivity and provides new insights into the development of innovative targeted therapeutics.

To explore the interactions involving the C1 *SRSF7*+ mast cell subtype and other cell types, we employed CellChat communication pattern analysis. This approach helped reveal coordinated responses and interactions between different cell types in the context of their communication pathways. Different cell types can simultaneously activate common cell type-independent signaling pathways or different cell type-specific signaling transduction transduction pathways (77). Through CellChat analysis, we established the intercellular communication network between most cells, including tumor cells, fibroblasts, T-NK cells, and various subtypes of MCs, etc., as a way to characterize the relationship between the subtype of C1 *SRSF7*+ MCs and other cell types, and at the same time, we identified the three modes of outgoing, incoming and their corresponding signaling pathway expression. The C1 *SRSF7*+ MCs subtype belongs to mode 1 in the outgoing pathway, and its communication molecules, i.e., ligands, include ANNEXIN, PARs, CSF, ICAM, etc.; and it belongs to mode 2 in the incoming pathway, and its communication molecules, i.e., receptors, include BAFF, CLEC, ALCAM, SELPG, etc. It is also worth noting that tumor cells, which can be learned after our careful observation, belong to mode 2 on the outgoing and mode 1 on the incoming, echoing the subtype of C1 *SRSF7*+ MCs, which drew our attention.

By targeting tumor cells and the C1 *SRSF7*+ MCs subtype for interactions analysis, we have identified the secretion of AREG ligands by a subtype of C1 *SRSF7*+ MCs in the EGF signaling pathway that act on the protein receptor *EGFR* on the membrane of the tumor cells. In previous studies, the *EGFR* family has been validated to play a key role in *EGFR* signaling through the activation of many important cellular processes, including cell division, growth, and differentiation. Playing a key role in mediating cell growth factor signaling (78), overexpression of *EGFR* signaling widely promotes tumor progression and leads to promotion of proliferation and inhibition of apoptosis (79). And cancer immunotherapies, particularly immune checkpoint blockade (ICB), have transformed oncology care over the past decade and significantly improved survival in a wide range of metastatic tumors. Based on significant treatment benefits, ICB therapy is approved by the FDA as monotherapy or in combination with other cancer therapies for cancers such as melanoma, breast cancer, renal cell carcinoma, head and neck squamous cell carcinoma, and lung cancer (80–84). However, the MCs-mediated pro-tumor axis AREG-EGFR in EC has not yet been mentioned. Therefore, our study provides new EC target therapeutic approaches and provides a scientific basis for the treatment and prognosis of EC. Meanwhile, to further investigate the role of *EGFR* in EC, we performed *in vitro* experiments using TE-10 and KYSE-30 cell lines. We observed that *EGFR* knockdown inhibited tumor cell activity, migration and proliferation, thereby suppressing tumor growth. However, previous studies have shown that epidermal growth factor receptor inhibitors (EGFRIs) produce a variety of dermatologic side effects in the majority of patients, and this targeted therapeutic regimen needs to be further refined (85).

Given their role in promoting tumor growth and immune evasion, mast cells are considered potential therapeutic targets. Contemporary therapeutic strategies may include the use of mast cell stabilizers, mast cell mediator inhibitors, or blocking mast cell recruitment to tumor tissues and organs.

Finally, we constructed a prognostic model to indicate that the higher the SMRS score, the worse the prognosis.

Our study will direct attention to MCs in the progression of esophageal cancer, trigger attention to them, and promote researchers' understanding of the tumor microenvironment in esophageal cancer. At the same time, we discovered the communication pathway between the tumor and our target MCs subtype. Although EFGR antagonists are still proved to have certain side effects, we believe that the development of targeted therapy will be further advanced in the future. However, this study still has some limitations. The relatively small sample size chosen is one aspect, and secondly, we only performed transcriptomics studies and *in vitro* experiments. The analysis of mast cell in EC using SCENIC and AUCell in our article is well-founded though and provides a detailed understanding of the regulatory networks that drive mast cell behavior. However, to draw more reliable conclusions, these findings must be validated by further experiments and compared across different cancer types. Next, we will integrate *in vivo* and *in vitro* experiments to provide a more comprehensive validation.

In conclusion, the innovative features of our study lie in the use of high-resolution single-cell analysis technology, the construction of cell-cell interaction networks, the analysis of dynamic evolutionary trajectories, the identification of regulatory networks, and experimental verification, which provide new ideas for the targeted treatment of MCs in EC and new cell carriers for the development of *EGFR* targeted drugs. These will help to promote the in-depth development of the research on EC and provide new strategies for the disease.

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

SYZ: Conceptualization, Data curation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. XZ: Conceptualization, Data curation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. ZX: Data curation, Software, Validation, Visualization, Writing – review & editing. SQZ: Methodology, Software, Validation, Visualization, Writing – review & editing. JX: Software, Validation, Visualization, Writing – review & editing. YZ: Conceptualization, Data curation, Methodology, Software,

Supervision, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

Our Graphical Abstract was drawn using Figdraw. Image ID: WWRRY06006. We would like to thank Figdraw for its contribution to this article.

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.

Publisher's note

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. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. *CA Cancer J Clin*. (2016) 66:115–32. doi: 10.3322/caac.21338

3. Baba Y, Nomoto D, Okadome K, Ishimoto T, Iwatsuki M, Miyamoto Y, et al. Tumor immune microenvironment and immune checkpoint inhibitors in esophageal squamous cell carcinoma. *Cancer Sci*. (2020) 111:3132–41. doi: 10.1111/cas.14541

4. Siegel RL, Wagle NS, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2023. *CA Cancer J Clin*. (2023) 73:233–54. doi: 10.3322/caac.21772

5. Smyth EC, Lagergren J, Fitzgerald RC, Lordick F, Shah MA, Lagergren P, et al. Oesophageal cancer. *Nat Rev Dis Primers*. (2017) 3:17048. doi: 10.1038/nrdp.2017.48

6. Morgan E, Soerjomataram I, Rumgay H, Coleman HG, Thrift AP, Vignat J, et al. The global landscape of esophageal squamous cell carcinoma and esophageal adenocarcinoma incidence and mortality in 2020 and projections to 2040: new estimates from GLOBOCAN 2020. *Gastroenterology*. (2022) 163:649–58. doi: 10.1053/j.gastro.2022.05.054

7. Zhang Y. Epidemiology of esophageal cancer. *World J Gastroenterol*. (2013) 19:5598–606. doi: 10.3748/wjg.v19.i34.5598

8. Dong J, Gao M, Li L, Pan X, Chen SY, Li J, et al. Associations of dietary inflammatory potential with esophageal precancerous lesions and esophageal squamous-cell cancer: A cross-sectional study. *Nutrients*. (2023) 15(8):4078. doi: 10.3390/nu15184078

9. You Y, Chen Y, Wei M, Tang M, Lu Y, Zhang Q, et al. Mediation role of recreational physical activity in the relationship between the dietary intake of live microbes and the systemic immune-inflammation index: A real-world cross-sectional study. *Nutrients*. (2024) 16(6):777. doi: 10.3390/nu16060777

10. Dowling GP, Daly GR, Hegarty A, Hembrecht S, Bracken A, Toomey S, et al. Comment on: Predictive value of pretreatment circulating inflammatory response markers in the neoadjuvant treatment of breast cancer: meta-analysis. *Br J Surg*. (2024) 111(5):znae132. doi: 10.1093/bjs/znae187

11. Zhang R, Lau L, Wu P, Yip HC, Wong SH. Endoscopic diagnosis and treatment of esophageal squamous cell carcinoma. *Methods Mol Biol*. (2020) 2129:47–62. doi: 10.1007/978-1-0716-0377-2_5

12. Tang H, Wang H, Fang Y, Zhu JY, Yin J, Shen YX, et al. Neoadjuvant chemoradiotherapy versus neoadjuvant chemotherapy followed by minimally invasive esophagectomy for locally advanced esophageal squamous cell carcinoma: a prospective multicenter randomized clinical trial. *Ann Oncol*. (2023) 34:163–72. doi: 10.1016/j.annonc.2022.10.508

13. Davern M, Donlon NE, Power R, Hayes C, King R, Dunne MR, et al. The tumour immune microenvironment in oesophageal cancer. *Br J Cancer*. (2021) 125:479–94. doi: 10.1038/s41416-021-01331-y

14. Nallasamy P, Nimmakayala RK, Parte S, Are AC, Batra SK, Ponnusamy MP. Tumor microenvironment enriches the stemness features: the architectural event of therapy resistance and metastasis. *Mol Cancer*. (2022) 21:225. doi: 10.1186/s12943-022-01682-x

15. Dinh HQ, Pan F, Wang G, Huang QF, Olingy CE, Wu ZY, et al. Integrated single-cell transcriptome analysis reveals heterogeneity of esophageal squamous cell carcinoma microenvironment. *Nat Commun*. (2021) 12:7335. doi: 10.1038/s41467-021-27599-5

16. Zhang X, Peng L, Luo Y, Zhang S, Pu Y, Chen Y, et al. Dissecting esophageal squamous-cell carcinoma ecosystem by single-cell transcriptomic analysis. *Nat Commun*. (2021) 12:5291. doi: 10.1038/s41467-021-25539-x

17. Zheng Y, Chen Z, Han Y, Han L, Zou X, Zhou B, et al. Immune suppressive landscape in the human esophageal squamous cell carcinoma microenvironment. *Nat Commun*. (2020) 11:6268. doi: 10.1038/s41467-020-20019-0

18. Segura-Villalobos D, Ramirez-Moreno IG, Martinez-Aguilar M, Ibarra-Sanchez A, Munoz-Bello JO, Anaya-Rubio I, et al. Mast cell-tumor interactions: molecular mechanisms of recruitment, intratumoral communication and potential therapeutic targets for tumor growth. *Cells*. (2022) 11(3):349. doi: 10.3390/cells11030349

19. Guo X, Shen W, Sun M, Lv J, Liu R. Activated mast cells combined with NRF2 predict prognosis for esophageal cancer. *J Oncol*. (2023) 2023:4211885. doi: 10.1155/2023/4211885

20. Puram SV, Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S, et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell*. (2017) 171:1611–24. doi: 10.1016/j.cell.2017.10.044

21. Wagner J, Rapsomaniki MA, Chevrier S, Anzeneder T, Langwieder C, Dykgers A, et al. A single-cell atlas of the tumor and immune ecosystem of human breast cancer. *Cell*. (2019) 177:1330–45. doi: 10.1016/j.cell.2019.03.005

22. van Galen P, Hovestadt V, Wadsworth IM, Hughes TK, Griffin GK, Battaglia S, et al. Single-cell RNA-seq reveals AML hierarchies relevant to disease progression and immunity. *Cell*. (2019) 176:1265–81. doi: 10.1016/j.cell.2019.01.031

23. Liu S, Chen LX, Ye LS, Hu B. Challenges in early detection and endoscopic resection of esophageal cancer: There is a long way to go. *World J Gastrointest Oncol*. (2024) 16:3364–67. doi: 10.4251/wjgo.v16.i7.3364

24. Zhang K, Ye B, Wu L, Ni S, Li Y, Wang Q, et al. Machine learning–based prediction of survival prognosis in esophageal squamous cell carcinoma. *Sci Rep*. (2023) 13:13532. doi: 10.1038/s41598-023-40780-8

25. 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

26. Xing J, Cai H, Lin Z, Zhao L, Xu H, Song Y, et al. Examining the function of macrophage oxidative stress response and immune system in glioblastoma multiforme through analysis of single-cell transcriptomics. *Front Immunol*. (2023) 14:1288137. doi: 10.3389/fimmu.2023.1288137

27. Jiang H, Yu D, Yang P, Guo R, Kong M, Gao Y, et al. Revealing the transcriptional heterogeneity of organ-specific metastasis in human gastric cancer using single-cell RNA Sequencing. *Clin Transl Med*. (2022) 12:e730. doi: 10.1002/ctm2.730

28. Wu F, Fan J, He Y, Xiong A, Yu J, Li Y, et al. Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. *Nat Commun*. (2021) 12:2540. doi: 10.1038/s41467-021-22801-0

29. Ding Y, Zhao Z, Cai H, Zhou Y, Chen H, Bai Y, et al. Single-cell sequencing analysis related to sphingolipid metabolism guides immunotherapy and prognosis of skin cutaneous melanoma. *Front Immunol*. (2023) 14:1304466. doi: 10.3389/fimmu.2023.1304466

30. Ge Q, Zhao Z, Li X, Yang F, Zhang M, Hao Z, et al. Deciphering the suppressive immune microenvironment of prostate cancer based on CD4+ regulatory T cells: Implications for prognosis and therapy prediction. *Clin Transl Med*. (2024) 14:e1552. doi: 10.1002/ctm2.1552

31. Zhou Y, Yang D, Yang Q, Lv X, Huang W, Zhou Z, et al. Single-cell RNA landscape of intratumoral heterogeneity and immunosuppressive microenvironment in advanced osteosarcoma. *Nat Commun*. (2020) 11:6322. doi: 10.1038/s41467-020-20059-6

32. 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

33. Liu P, Xing N, Xiahou Z, Yan J, Lin Z, Zhang J. Unraveling the intricacies of glioblastoma progression and recurrence: insights into the role of NFYB and oxidative phosphorylation at the single-cell level. *Front Immunol*. (2024) 15:1368685. doi: 10.3389/fimmu.2024.1368685

34. Zhou W, Lin Z, Tan W. Deciphering the molecular landscape: integrating single-cell transcriptomics to unravel myofibroblast dynamics and therapeutic targets in clear cell renal cell carcinomas. *Front Immunol*. (2024) 15:1374931. doi: 10.3389/fimmu.2024.1374931

35. Shao W, Lin Z, Xiahou Z, Zhao F, Xu J, Liu X, et al. Single-cell RNA sequencing reveals that MYBL2 in Malignant epithelial cells is involved in the development and progression of ovarian cancer. *Front Immunol*. (2024) 15:1438198. doi: 10.3389/fimmu.2024.1438198

36. Lin Z, Sui X, Jiao W, Chen C, Zhang X, Zhao J. Mechanism investigation and experiment validation of capsaicin on uterine corpus endometrial carcinoma. *Front Pharmacol*. (2022) 13:953874. doi: 10.3389/fphar.2022.953874

37. Lin Z, Li X, Shi H, Cao R, Zhu L, Dang C, et al. Decoding the tumor microenvironment and molecular mechanism: unraveling cervical cancer subpopulations and prognostic signatures through scRNA-Seq and bulk RNA-seq analyses. *Front Immunol*. (2024) 15:1351287. doi: 10.3389/fimmu.2024.1351287

38. Zheng L, Qin S, Si W, Wang A, Xing B, Gao R, et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science*. (2021) 374:abe6474. doi: 10.1126/science.abe6474

39. Van den Berge K, Roux DBH, Street K, Saelens W, Cannoodt R, Saeys Y, et al. Trajectory-based differential expression analysis for single-cell sequencing data. *Nat Commun*. (2020) 11:1201. doi: 10.1038/s41467-020-14766-3

40. Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. (2017) 14:1083–86. doi: 10.1038/nmeth.4463

41. Gulati GS, Sikandar SS, Wesche DJ, Manjunath A, Bharadwaj A, Berger MJ, et al. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science*. (2020) 367:405–11. doi: 10.1126/science.aax0249

42. Alexa A, Rahnenfuhrer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*. (2006) 22:1600–07. doi: 10.1093/bioinformatics/btl140

43. Lin Z, Fan W, Yu X, Liu J, Liu P. Research into the mechanism of intervention of SanQi in endometriosis based on network pharmacology and molecular docking technology. *Med (Baltimore)*. (2022) 101:e30021. doi: 10.1097/MD.0000000000030021

44. Zhao J, Jiao W, Sui X, Zou J, Wang J, Lin Z. Construction of a prognostic model of luteolin for endometrial carcinoma. *Am J Transl Res*. (2023) 15:2122–39.

45. 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

46. Lin Z, Zou J, Sui X, Yao S, Lin L, Wang J, et al. Necroptosis-related lncRNA signature predicts prognosis and immune response for cervical squamous cell carcinoma and endocervical adenocarcinomas. *Sci Rep*. (2022) 12:16285. doi: 10.1038/s41598-022-20858-5

47. Zou J, Lin Z, Jiao W, Chen J, Lin L, Zhang F, et al. A multi-omics-based investigation of the prognostic and immunological impact of necroptosis-related mRNA in patients with cervical squamous carcinoma and adenocarcinoma. *Sci Rep*. (2022) 12:16773. doi: 10.1038/s41598-022-20566-0

48. Zhao J, Zou J, Jiao W, Lin L, Wang J, Lin Z. Construction of N-7 methylguanine-related mRNA prognostic model in uterine corpus endometrial carcinoma based on multi-omics data and immune-related analysis. *Sci Rep*. (2022) 12:18813. doi: 10.1038/s41598-022-22879-6

49. Lin Z, Sui X, Jiao W, Wang Y, Zhao J. Exploring the mechanism and experimental verification of puerarin in the treatment of endometrial carcinoma based on network pharmacology and bioinformatics analysis. *BMC Complement Med Ther*. (2022) 22:150. doi: 10.1186/s12906-022-03623-z

50. Lin Z, Fan W, Sui X, Wang J, Zhao J. Necroptosis-related lncRNA signatures for prognostic prediction in uterine corpora endometrial cancer. *Reprod Sci*. (2023) 30:576–89. doi: 10.1007/s43032-022-01023-9

51. Liu J, Zhang Y, Zhao J, Yang Z, Li D, Katirai F, et al. Mast cell: insight into remodeling a tumor microenvironment. *Cancer Metastasis Rev*. (2011) 30:177–84. doi: 10.1007/s10555-011-9276-1

52. Xu K, Sun S, Yan M, Cui J, Yang Y, Li W, et al. DDX5 and DDX17-multifaceted proteins in the regulation of tumorigenesis and tumor progression. *Front Oncol*. (2022) 12:943032. doi: 10.3389/fonc.2022.943032

53. Cubillos-Ruiz JR, Bettigole SE, Glimcher LH. Tumorigenic and immunosuppressive effects of endoplasmic reticulum stress in cancer. *Cell*. (2017) 168:692–706. doi: 10.1016/j.cell.2016.12.004

54. Mukherjee D, Bercz LS, Torok MA, Mace TA. Regulation of cellular immunity by activating transcription factor 4. *Immunol Lett*. (2020) 228:24–34. doi: 10.1016/j.imlet.2020.09.006

55. Wang HC, Zhou Y, Huang SK. SHP-2 phosphatase controls aryl hydrocarbon receptor-mediated ER stress response in mast cells. *Arch Toxicol*. (2017) 91:1739–48. doi: 10.1007/s00204-016-1861-1

56. Wortel I, van der Meer LT, Kilberg MS, van Leeuwen FN. Surviving stress: modulation of ATF4-mediated stress responses in normal and Malignant cells. *Trends Endocrinol Metab*. (2017) 28:794–806. doi: 10.1016/j.tem.2017.07.003

57. Han S, Zhu L, Zhu Y, Meng Y, Li J, Song P, et al. Targeting ATF4-dependent pro-survival autophagy to synergize glutaminolysis inhibition. *Theranostics*. (2021) 11:8464–79. doi: 10.7150/thno.60028

58. Beck A, Shatz-Azoulay H, Vinik Y, Isaac R, Boura-Halfon S, Zick Y. Nedd4 family interacting protein 1 (Ndfip1) promotes death of pancreatic beta cells. *Biochem Biophys Res Commun*. (2015) 465:851–56. doi: 10.1016/j.bbrc.2015.08.099

59. Gu X, Cai L, Luo Z, Shi L, Peng Z, Sun Y, et al. Identification and validation of a muscle failure index to predict prognosis and immunotherapy in lung adenocarcinoma through integrated analysis of bulk and single-cell RNA sequencing data. *Front Immunol*. (2022) 13:1057088. doi: 10.3389/fimmu.2022.1057088

60. Luo Z, He Z, Qin H, Chen Y, Qi B, Lin J, et al. Exercise-induced IL-15 acted as a positive prognostic implication and tumor-suppressed role in pan-cancer. *Front Pharmacol*. (2022) 13:1053137. doi: 10.3389/fphar.2022.1053137

61. 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) 14:91. doi: 10.1186/s13045-021-01105-2

62. Da SE, Jamur MC, Oliver C. Mast cell function: a new vision of an old cell. *J Histochem Cytochem*. (2014) 62:698–738. doi: 10.1369/0022155414545334

63. Dery RE, Lin TJ, Befus AD, Milne CD, Moqbel R, Menard G, et al. Redundancy or cell-type-specific regulation? Tumour necrosis factor in alveolar macrophages and mast cells. *Immunology*. (2000) 99:427–34. doi: 10.1046/j.1365-2567.2000.00982.x

64. Blair RJ, Meng H, Marchese MJ, Ren S, Schwartz LB, Tonnesen MG, et al. Human mast cells stimulate vascular tube formation. Tryptase is a novel, potent angiogenic factor. *J Clin Invest*. (1997) 99:2691–700. doi: 10.1172/JCI119458

65. Marech I, Ammendola M, Sacco R, Capriuolo GS, Patruno R, Rubini R, et al. Serum tryptase, mast cells positive to tryptase and microvascular density evaluation in early breast cancer patients: possible translational significance. *BMC Cancer*. (2014) 14:534. doi: 10.1186/1471-2407-14-534

66. Aller MA, Arias A, Arias JI, Arias J. Carcinogenesis: the cancer cell-mast cell connection. *Inflammation Res*. (2019) 68:103–16. doi: 10.1007/s00011-018-1201-4

67. Gorzalczany Y, Sagi-Eisenberg R. Role of mast cell-derived adenosine in cancer. *Int J Mol Sci*. (2019) (10):2603. doi: 10.3390/ijms20102603

68. Alda S, Ceausu RA, Gaje PN, Raica M, Cosoroaba RM. Mast cell: A mysterious character in skin cancer. *In Vivo*. (2024) 38:58–68. doi: 10.21873/invivo.13410

69. Lichterman JN, Reddy SM. Mast cells: A new frontier for cancer immunotherapy. *Cells*. (2021) 10(6):1270. doi: 10.3390/cells10061270

70. Nyamao RM, Wu J, Yu L, Xiao X, Zhang FM. Roles of DDX5 in the tumorigenesis, proliferation, differentiation, metastasis and pathway regulation of human Malignancies. *Biochim Biophys Acta Rev Cancer*. (2019) 1871:85–98. doi: 10.1016/j.bbcan.2018.11.003

71. Oakes SA, Papa FR. The role of endoplasmic reticulum stress in human pathology. *Annu Rev Pathol*. (2015) 10:173–94. doi: 10.1146/annurev-pathol-012513-104649

72. Mercier R, LaPointe P. The role of cellular proteostasis in antitumor immunity. *J Biol Chem*. (2022) 298:101930. doi: 10.1016/j.jbc.2022.101930

73. Van Drie JH. Protein folding, protein homeostasis, and cancer. *Chin J Cancer*. (2011) 30:124–37. doi: 10.5732/cjc.010.10162

74. Scott MD, Frydman J. Aberrant protein folding as the molecular basis of cancer. *Methods Mol Biol*. (2003) 232:67–76. doi: 10.1385/1-59259-394-1:67

75. Visciano C, Prevete N, Liotti F, Marone G. Tumor-associated mast cells in thyroid cancer. *Int J Endocrinol*. (2015) 2015:705169. doi: 10.1155/2015/705169

76. Ren FJ, Cai XY, Yao Y, Fang GY. JunB: a paradigm for Jun family in immune response and cancer. *Front Cell Infect Microbiol*. (2023) 13:1222265. doi: 10.3389/fcimb.2023.1222265

77. Joung J, Kirchgatterer PC, Singh A, Cho JH, Nety SP, Larson RC, et al. CRISPR activation screen identifies BCL-2 proteins and B3GNT2 as drivers of cancer resistance to T cell-mediated cytotoxicity. *Nat Commun*. (2022) 13:1606. doi: 10.1038/s41467-022-29205-8

78. Kyriakopoulou K, Kefali E, Piperigkou Z, Bassiony H, Karamanos NK. Advances in targeting epidermal growth factor receptor signaling pathway in mammary cancer. *Cell Signal*. (2018) 51:99–109. doi: 10.1016/j.cellsig.2018.07.010

79. Ayati A, Moghimi S, Salarinejad S, Safavi M, Pouramiri B, Foroumadi A. A review on progression of epidermal growth factor receptor (EGFR) inhibitors as an efficient approach in cancer targeted therapy. *Bioorg Chem*. (2020) 99:103811. doi: 10.1016/j.bioorg.2020.103811

80. Larkin J, Chiarion-Sileni V, Gonzalez R, Grob JJ, Cowey CL, Lao CD, et al. Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *N Engl J Med*. (2015) 373:23–34. doi: 10.1056/NEJMoa1504030

81. Schmid P, Adams S, Rugo HS, Schneeweiss A, Barrios CH, Iwata H, et al. Atezolizumab and nab-paclitaxel in advanced triple-negative breast cancer. *N Engl J Med*. (2018) 379:2108–21. doi: 10.1056/NEJMoa1809615

82. Motzer RJ, Rini BI, McDermott DF, Aren FO, Hammers HJ, Carducci MA, et al. Nivolumab plus ipilimumab versus sunitinib in first-line treatment for advanced renal cell carcinoma: extended follow-up of efficacy and safety results from a randomised, controlled, phase 3 trial. *Lancet Oncol.* (2019) 20:1370–85. doi: 10.1016/S1470-2045(19)30413-9

83. Burtness B, Harrington KJ, Greil R, Soulieres D, Tahara M, de Castro GJ, et al. Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-048): a randomised, open-label, phase 3 study. *Lancet*. (2019) 394:1915–28. doi: 10.1016/S0140-6736(19)32591-7

84. Gandhi L, Rodriguez-Abreu D, Gadgeel S, Esteban E, Felip E, De Angelis F, et al. Pembrolizumab plus chemotherapy in metastatic non-small-cell lung cancer. *N Engl J Med*. (2018) 378:2078–92. doi: 10.1056/NEJMoa1801005

85. Lacouture ME, Anadkat MJ, Bensadoun RJ, Bryce J, Chan A, Epstein JB, et al. Clinical practice guidelines for the prevention and treatment of EGFR inhibitor-associated dermatologic toxicities. *Support Care Cancer*. (2011) 19:1079–95. doi: 10.1007/s00520-011-1197-6

86. Tang H, Kang R, Liu J, Tang D. ATF4 in cellular stress, ferroptosis, and cancer. *Arch Toxicol.* (2024) 98:1025–41. doi: 10.1007/s00204-024-03681-x

frontiers | Frontiers in Immunology

TYPE Review
PUBLISHED 08 October 2024
DOI 10.3389/fimmu.2024.1483595

OPEN ACCESS

EDITED BY
Raquel Alarcon Rodriguez,
University of Almeria, Spain

REVIEWED BY
Yuquan Chen,
Monash University, Australia
Zhijia Xia,
Ludwig Maximilian University of Munich, Germany
Lanqian Su,
Southwest Medical University, China

*CORRESPONDENCE
Xiaofei Liu
drliuxf@126.com

†These authors have contributed equally to this work

RECEIVED 20 August 2024
ACCEPTED 23 September 2024
PUBLISHED 08 October 2024

CITATION
Zhang Y, Gong S and Liu X (2024) Spatial transcriptomics: a new frontier in accurate localization of breast cancer diagnosis and treatment.
Front. Immunol. 15:1483595.
doi: 10.3389/fimmu.2024.1483595

Spatial transcriptomics: a new frontier in accurate localization of breast cancer diagnosis and treatment

Yang Zhang[1†], Shuhua Gong[2†] and Xiaofei Liu[1*]

[1]Breast and Thyroid Surgery, Shandong University of Traditional Chinese Medicine Affiliated Hospital, Jinan, China, [2]Department of Student Affair, Shandong College of Traditional Chinese Medicine, Yantai, China

Breast cancer is one of the most prevalent cancers in women globally. Its treatment and prognosis are significantly influenced by the tumor microenvironment and tumor heterogeneity. Precision therapy enhances treatment efficacy, reduces unwanted side effects, and maximizes patients' survival duration while improving their quality of life. Spatial transcriptomics is of significant importance for the precise treatment of breast cancer, playing a critical role in revealing the internal structural differences of tumors and the composition of the tumor microenvironment. It offers a novel perspective in studying the spatial structure and cell interactions within tumors, facilitating more effective personalized treatments for breast cancer. This article will summarize the latest findings in the diagnosis and treatment of breast cancer from the perspective of spatial transcriptomics, focusing on the revelation of the tumor microenvironment, identification of new therapeutic targets, enhancement of disease diagnostic accuracy, comprehension of tumor progression and metastasis, assessment of drug responses, creation of high-resolution maps of tumor cells, representation of tumor heterogeneity, and support for clinical decision-making, particularly in elucidating the tumor microenvironment, tumor heterogeneity, immunotherapy and their correlation with clinical outcomes.

KEYWORDS

breast cancer, spatial transcriptomics, tumor heterogeneity, tumor microenvironment, immunotherapy

1 Introduction

The therapeutic approach to breast cancer necessitates a personalized strategy that accounts for the patient's unique profile, encompassing the tumor's biological attributes, genomic expression signatures, estrogen and progesterone receptor status, HER2 amplification, tumor microenvironmental dynamics, the patient's genetic predispositions,

and lifestyle factors (1). Breast cancer is stratified into distinct molecular subtypes, including Luminal A, Luminal B, HER2-positive, and triple-negative breast cancer (TNBC), each exhibiting divergent responses to therapeutic interventions, thereby underscoring the imperative for tailored treatment modalities such as chemotherapy, hormonal therapies, targeted therapies, and immunotherapies (Figure 1). Molecular subtyping also serves as a prognostic indicator, providing insights into the survival outcomes of breast cancer patients (2). Accurate molecular subtyping is pivotal for the formulation of effective treatment regimens. However, the intrinsic heterogeneity of breast cancer, characterized by variable gene expression and mutational landscapes among patients with ostensibly similar disease, may lead to therapeutic resistance, disease progression, prognostic variability, and clonal evolution, thereby advocating for individualized therapeutic approaches (3).

Precision oncology in breast cancer is witnessing a transformative evolution, propelled by the remarkable advancements in genomics, biomarker identification, and tailored therapeutic approaches. The meticulous genomic profiling of neoplastic tissues has facilitated the elucidation of pivotal molecular targets, thereby enabling the formulation of more precise and efficacious treatment regimens for affected individuals. Notably, immunotherapeutic interventions, particularly the utilization of immune checkpoint inhibitors (4), have emerged as a potent therapeutic modality for select breast cancer subtypes. Furthermore, the application of mitochondrial inhibitors (5) has demonstrated the potential to enhance the therapeutic efficacy by modulating the metabolic reprogramming of breast cancer cells. The incorporation of artificial intelligence (AI) and machine learning (ML) algorithms has revolutionized the analysis of extensive datasets, thereby unveiling novel predictive models that forecast treatment responsiveness and disease progression (6–8). Concurrently, the integration of lifestyle modifications and pharmacological interventions as preventive strategies (9) underscores the multifaceted and individualized nature of breast cancer management. Collectively, these developments herald an era of expanded therapeutic horizons and renewed optimism for breast cancer patients.

Advancements in oncology research have been significantly propelled by the advent of Spatial Transcriptomics (ST), a cutting-edge technology that enables the concurrent assessment of gene expression profiles in the context of cellular spatial architecture (10). ST has emerged as a valuable tool for dissecting tumor heterogeneity, thereby facilitating a more precise understanding of tumor progression and therapeutic responses. Its utility extends to the elucidation of tumor microenvironment (TME) intricacies, offering insights into the spatial distribution and reciprocal interactions between neoplastic and surrounding cells (11). ST has been instrumental in revealing the complex interplay between tumor cells and the immune milieu, as well as in delineating the gene expression patterns within the immune microenvironment, thereby enhancing our comprehension of immune function. For instance, ST has delineated five distinct immune microenvironmental subtypes in hepatocellular carcinoma, uncovering the heterogeneity of tumor-associated neutrophils (TAN) and identifying key subpopulations, such as $CCL4^+$ and $PD\text{-}L1^+$ TAN, which are implicated in tumor promotion, thus suggesting novel therapeutic targets for liver cancer (12). Utilizing ST, Ye et al. have delineated the tumor boundary region, which serves as a critical interface between malignant and non-malignant tissues, identifying specific cell subtypes, cellular interactions, and potential therapeutic targets enriched at this boundary, and have observed significant infiltration of FAP^+ fibroblasts and $SPP1^+$ macrophages in colorectal cancer, which correlates with adverse prognosis and resistance to immunotherapy (13, 14).

In the field of breast cancer research, the application of ST is gradually demonstrating its unique value and potential. Although

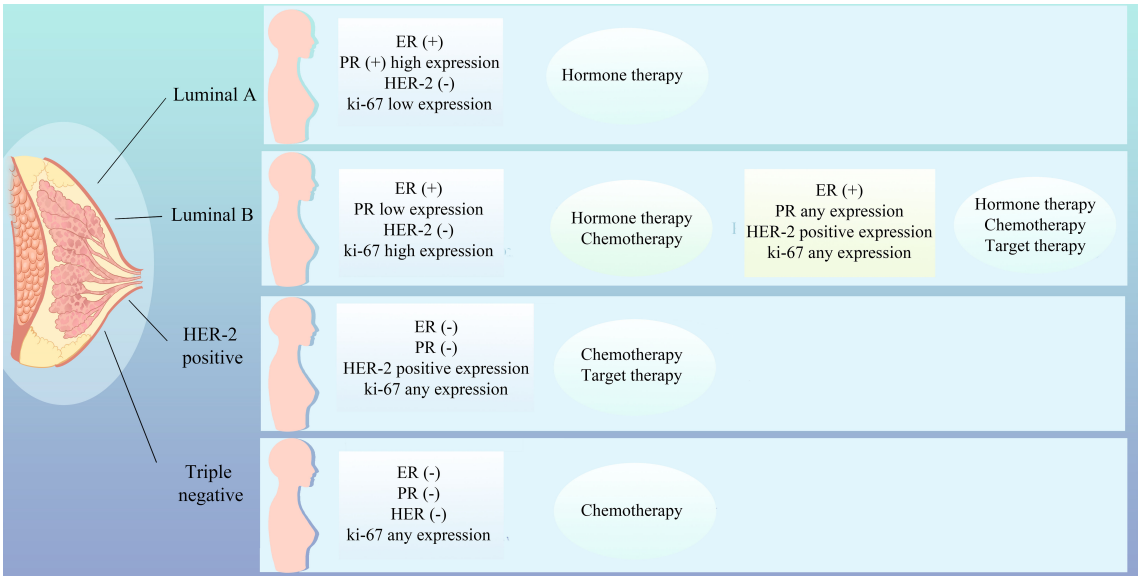


FIGURE 1
Molecular subtyping of breast cancer and therapeutic strategies.

the current research articles exhibit a certain degree of dispersion, each contains valuable information and insights that urgently await systematic organization and integration. This review aims to synthesize these research findings into a comprehensive knowledge map through meticulous literature review and in-depth analysis, with the expectation of revealing the integrated application and development trends of ST in breast cancer research.

Furthermore, this study is dedicated to exploring the potential application prospects of ST in breast cancer research, as well as the innovative and transformative impacts it may bring. Through this process, we anticipate providing new perspectives and ideas for understanding the biological mechanisms of breast cancer, improving diagnostic methods, and optimizing treatment strategies.

2 ST and the molecular characteristics of breast cancer

Breast cancer is marked by pronounced tumor heterogeneity, with neoplastic cells from disparate regions potentially manifesting divergent gene expression profiles and biological behaviors. A research consortium from the Australian Institute for Medical Research has harnessed ST in conjunction with single-cell RNA sequencing to delineate the most exhaustive cellular cartography of breast cancer to date. This investigation has not only unveiled the heterogeneity of neoplastic cells but also delineated nine tumor ecotypes correlated with the overall patient survival, with certain ecotypes being predictive of an adverse prognosis (15).

Invasive micropapillary carcinoma (IMPC), a distinctive histological variant of breast cancer, is distinguished by its high propensity for lymphovascular invasion and lymph node metastasis. A team led by Lv, J. has postulated the "IMPC tumor cell clump metastasis" hypothesis, pioneering the transcriptional profiling of IMPC and uncovering its profound heterogeneity, which is intricately linked to metabolic reprogramming. Metabolically aberrant IMPC subpopulations are spatially segregated and exhibit heightened lipid metabolism across all IMPC stratified clusters. Concurrently, elevated expression of the sterol regulatory element-binding transcription factor 1 (SREBF1) protein has been correlated with increased lymph node metastasis and diminished survival rates in IMPC patients, underscoring its potential as a diagnostic and therapeutic biomarker (16).

Yoshitake, R. and colleagues have elucidated that estrogen receptor-positive (ER+) breast cancer encompasses four spatially discrete populations with functional heterogeneity, including estrogen responsiveness, proliferation, hypoxia induction, and inflammation association. The "proliferative" subset is pivotal for estrogen-driven tumorigenesis, conferring a phenotype reminiscent of the luminal B subtype. Gene signatures emanating from proliferative, hypoxia-induced, and inflammation-associated populations are significantly associated with inferior clinical outcomes, whereas patients with estrogen-responsive signatures demonstrate a more favorable prognosis (17).

Sun H (18) associates have stratified the MDA-MB-231 tumor mass into necrotic, peripheral necrotic, hypoxic tumor, adaptive survival tumor, and invasive tumor compartments based on hypoxic status and transcriptomic profiling. Each compartment possesses a unique expression signature, with diverse gene networks activated under the influence of distinct hypoxic microenvironments, thereby dictating the fate of tumor cells across different regions. The spatial transcriptional distribution of 35 hypoxia-associated genes was mapped, revealing that disparate tumor regions with distinct hypoxia-related gene signatures exhibit unique characteristics. B lymphocytoma-2 gene-homology 3 (BINP3), implicated in the regulation of apoptosis, exhibits heightened expression in hypoxic regions, whereas prolyl-4-hydroxylase alpha polypeptide I (P4HA1), though broadly expressed, displays significant variation at the periphery of necrotic areas. Lactate dehydrogenase (LDHA), a marker of tumor metabolism, is ubiquitously upregulated across the tumor tissue, with the most pronounced differences in the invasive regions. Beyond hypoxia-inducible factor-1α (HIF-1α), alternative pathways are implicated in the modulation of these gene networks. These hypoxia-associated genes not only interact among themselves but also serve as key regulators within the gene regulatory networks of each compartment. The elucidation of these spatial heterogeneities is instrumental for advancing our understanding of breast cancer biology and for the development of novel therapeutic strategies targeting breast malignancies.

The architecture and functionality of mammary tissue exhibit intricate spatial complexity, with cellular constituents and states across regions potentially exerting distinct influences on the health of the breast and the evolution of pathologies. Precise stratification is paramount for the determination of appropriate therapeutic interventions. Research has posited that the heterogeneity of cellular populations within neoplastic tissue, encompassing both malignant and non-malignant entities, are organized into tumor regions or "niches" with distinct cellular compositions. These zones may reflect the recruitment of specific cellular subpopulations or the differentiation processes of cells (3). Kumar, T.'s investigative research has elucidated the molecular disparities between the ductal and alveolar compartments of the breast, as well as the ecosystem of resident immune cells within these tissues. Specifically, certain subtypes of immune and basal cells exhibit a higher prevalence in the ductal and alveolar regions, in contrast to their sparse distribution in the connective tissue areas. The ductal region has been correlated with an enhanced expression of genes associated with secretory luminal epithelial cells (LumSec), whereas the alveolar region is characterized by an upregulation of genes specific to hormone-responsive luminal epithelial cells (LumHR). Lymphocytes were predominantly detected in the connective tissue areas, whereas vascular cells showed a higher prevalence in the ductal and alveolar regions (19).

An additional study employing single-cell RNA sequencing (scRNA-seq) and ST has identified a cluster of disseminating cancer cells characterized by heightened oxidative phosphorylation (OXPHOS) activity. This investigation has discerned a metabolic shift between glycolysis and OXPHOS as the process of dissemination commences. Moreover, this distinctive cellular cluster is observed to be distributed along the tumor's leading edge (20). The heterogeneity of cellular positions at the tumor ductal periphery or core underscores the necessity of incorporating the spatial architecture

of the tumor in therapeutic strategies. Spatially resolved transcriptomics, genomics, and single-cell analyses have unveiled the intrinsic subtype heterogeneity within mixed infiltrating ductal and lobular carcinoma (MDLC). Compared to TNBC or basal ductal and estrogen receptor-positive (ER^+) phenotypes, MDLC exhibits a pronounced enrichment of luminal lobular region cells characterized by cell cycle arrest/senescence and oncogenic (ER and MYC) features, along with inactivation of E-cadherin 1 (CDH1) specific to the lobular rather than ductal regions. Furthermore, the identification of a unique oncogenic single-cell ductal and lobular subset accentuates the heterogeneity within the region. It has been substantiated that the tumor morphology and histological heterogeneity within MDLC are governed by intrinsic subtype and oncogenic heterogeneity, which may engender prognostic ambiguity and therapeutic challenges (21).

ST has been instrumental in elucidating the inter-regional interactions within human breast cancer tumors, as well as the regulatory mechanisms from receptor-ligand (LR) interactions to target gene expression. This approach has unveiled the intricate crosstalk between disparate cell types, which may exert substantial influence on the functional attributes of mammary tissue and the trajectory of disease progression. In a study conducted by Wang H et al., it was determined that multi-tiered signaling networks exist between any two tumor regions, with the affinity of LR interactions within these networks varying significantly between different regions (22). Certain LR pairs, such as Tumor Necrosis Factor (TNF)-Tumor Necrosis Factor Receptor Superfamily Member 21 (TNFRSF21), Retinol Binding Protein 4(RBP4)- Stimulated by Retinoic Acid Gene 6 (STRA6), Platelet-derived Growth Factor A (PDGFA)-Platelet Derived Growth Factor Receptor Beta (PDGFRB), Tenascin C (TNC)-Contactin 1 (CNTN1), and ALK and LTK ligand 2 (ALKAL2)-Anaplastic Lymphoma Kinase (ALK), have demonstrated enhanced interaction profiles. Prior research has underscored the significance of ALKAL2-ALK signaling (23, 24) and PDGFA-PDGFRB signaling (25) in the oncogenic processes of breast cancer, particularly in tumor growth and metastasis.

Employing ST analysis, investigators have discerned subpopulations of cells and molecular signatures that correlate with clinical outcomes, thereby introducing novel biomarkers pivotal for the stratification, treatment response prediction, and personalized therapeutics in breast cancer. Ductal carcinoma *in situ* (DCIS) represents an incipient phase of breast cancer with the potential to evolve into invasive ductal carcinoma. Dr. Satoi Nagasawa from the University of Tokyo performed ST analysis on DCIS, revealing that mutations in GATA Binding Protein 3 (GATA3) and Phosphoinositide-3-Kinase, Catalytic, Alpha Polypeptide (PIK3CA) are the most prevalent within this cohort. DCIS cells harboring GATA3 mutations have been observed to occasionally evolve into invasive cancers, implicating their role in epithelial-mesenchymal transition (EMT) and angiogenesis. In contrast, DCIS cells with PIK3CA mutations do not progress to malignancy (26), highlighting the cellular heterogeneity intrinsic to breast cancer tumors. ST sequencing data corroborate that DNA Damage-Inducible Transcript 3 (DDIT3) co-localizes with biomarkers of malignant epithelia (KRT19), myofibroblasts (ACTA2), and monocytic/macrophage populations (CD68), exhibiting heightened expression within cellular clusters. DDIT3's role in modulating the TME and intercellular communication is multifaceted, with positive correlations observed with pathways implicated in apoptosis, cell cycle regulation, DNA damage response, and the epithelial-mesenchymal transition (EMT) in breast cancer (27).

3 ST and the breast cancer TME

TME is a complex ecosystem comprising a diverse array of immune and stromal cells, vascular structures, extracellular matrix (ECM) components, and an array of soluble mediators (28). This multifaceted milieu is instrumental in modulating tumorigenesis and dictating the trajectory of cancer evolution. The cellular constituents of the TME exhibit significant heterogeneity, and their spatial architecture varies across distinct genomic subtypes of breast cancer. These variations are manifested in the interactions and topographical arrangements among cellular subsets, which in turn can significantly influence the neoplastic process and the tumor's responsiveness to therapeutic interventions. ST has elucidated the intricate distribution of these cellular elements within the TME and delineated the nuanced interactions between diverse cellular populations and malignant cells, thereby enhancing our comprehension of the molecular underpinnings of breast cancer and the intricacies of the TME, including its therapeutic resistance.

Investigators have meticulously mapped the spatial organization and context-specific landscape of breast cancer and its attendant microenvironment, profiling the expression of 37 proteins across a cohort of 483 tumor samples, these data helped the investigators to distinguish between different cellular phenotypes such as tumor cells, stromal cells and immune cells. For example, they were able to distinguish between different phenotypes of epithelial cells, fibroblasts, myofibroblasts, endothelial cells, T cells, B cells, and macrophages, among others, thereby unmasking the spatial heterogeneity of cellular constituents within the TME (29). Research endeavors have further characterized the spatial co-occurrence and interplay of various cellular phenotypes within breast cancer. Croizer H (30) has expounded on the malleability of FAP^+ cancer-associated fibroblasts (CAFs) and their intricate crosstalk with immune cells, identifying a spectrum of 10 spatially orchestrated FAP^+ CAF clusters associated with cellular modules, designated as EcoCellTypes (ECTs). These ECTs, which include immunosuppressive and immuno-permissive variants, encompass specific FAP^+ CAF clusters and immune cell populations that are situated at discrete distances from tumor conglomerates and vascular structures. Certain FAP^+ CAF clusters have been correlated with the invasive properties of breast cancer, suggesting that the heterogeneity among FAP^+ CAFs may play a pivotal role in the progression of DCIS. Another study focusing on CAFs has delineated the spatial organization of disparate CAF populations within breast cancer. Some specific CAF populations were found to coincide with heightened transforming growth factor-β (TGF-β) signaling, with elastin microfibril interface-derived protein 1 (EMILIN1) emerging as a paramount regulatory gene. Elevated EMILIN1 expression at the tumor periphery is associated with robust CD8 T-cell infiltration, and

such increased EMILIN1 expression correlates with an improved prognosis in breast cancer patients, underscoring its functional relevance in the recruitment of cytotoxic T cells to the TME (31).

3.1 ST and the tumor-immune microenvironment

The immunological components of the TME are under intense scrutiny due to their significant roles in oncological processes. Immune infiltration within tumor tissues surpasses that observed in normal tissues (32), with the potential to either combat or propagate tumorigenesis. For instance, TAN can stimulate the proliferation and invasiveness of malignant cells, yet they are also capable of exerting cytotoxic effects against tumor cells (33). Macrophages within the TME demonstrate diverse polarization states; the classically activated M1 macrophages are associated with anti-tumor activities, while the alternatively activated M2 macrophages are more inclined to support tumorigenesis (34). ST has delineated the distribution patterns of these cellular elements across tumor tissues, thereby providing a more refined representation of the spatial tumor-specificity inherent to each cell population. The composition of the TME is subject to variation predicated on spatial positioning, which corresponds to the distinct functionalities of immune cells. Research team reclustered immune cells to identify T cells and innate lymphoid cells, myeloid cells, B cells and plasmablasts. They identified 18 T-cell and innate lymphoid clusters, 13 clusters myeloid cells, three major cell types in the stromal compartment across patients, these cell clusters are mutually exclusive in their spatial arrangement, and the "ecotypes" composed of different cell clusters differ significantly from the tumor subtypes and prognosis (15). Research (19) has disclosed that immune cells are predominantly localized to the parenchymal compartments of mammary tissue, distinct from the intravascular locales, indicative of their tissue-residency. Furthermore, extensive ligand-receptor interactions between immune cells and other cellular constituents of the mammary tissue, such as epithelial and stromal fibroblasts, suggest an influential role for immune cells in the steady-state and pathological mechanisms of mammary tissue. In a cohort of 152 $HER2^+$ ductal breast carcinomas, cellular constituents of the TME displayed well-defined three-dimensional localization patterns; for instance, T lymphocytes exhibit a propensity to aggregate perivascularly and along vascular networks, whereas macrophage accumulations manifest distinct distributional configurations, ranging from uniform dispersion to localized aggregation (35). Within TNBC, there exists a pronounced variability in the spatial distribution of immune cell subpopulations. Intraepithelial T and B lymphocytes consistently exhibit a more clonal and less diverse immune repertoire compared to their stromal counterparts. Overamplification of T cell clones within the intraepithelial compartment is more pronounced than within the stromal T cells, indicative of an enriched accumulation of antigen-specific T cells at the tumor core (36).

Utilizing spatial information, we can more adeptly investigate the intricate interplay between immune cells and neoplastic cells. By amalgamating single-nucleus RNA sequencing (snRNA-seq) with ST datasets to elucidate the spatial heterogeneity of immune cells within the breast cancer TME, the ST datasets have been stratified into six principal zones: the luminal region, basal region, the interfacing area between the luminal region and basal region, stroma and infiltrating lymphocyte areas based on the principal component scores across all ST spots (37). Neutrophils were found to be enriched in the luminal region, whereas B cells were observed to be primarily infiltrating the basal region. Activated $CD8^+$ T cells display an enhanced tumor spatial specificity relative to their quiescent counterparts, aligning with their purported anti-tumor capabilities. Certain immunosuppressive cell populations, such as Regulatory T cells (Tregs) and Cancer-associated Fibroblasts (CAFs), exhibit diminished tumor spatial specificity. Furthermore, macrophage clusters expressing both M1 and M2 phenotypic markers, notably $Mac.FABP5^+$ cells, demonstrate heightened tumor spatial specificity in comparison to pro-inflammatory macrophages, implying a potential direct induction by neoplastic cells (32). Among HER2 positive patient cohorts, a shared spatial expression signature has been identified, with reciprocal interactions between Mø and T cell subsets evident within the context of type I interferon responses (38). Collectively, these insights underscore the qualitative disparities among immune cell clusters across distinct clinical subtypes, augmenting our comprehension of the intricate interactions between tumor cells and the TME's architectural intricacies, thereby highlighting the imperative for subtype-specific targeted therapeutic strategies in clinical practice.

3.2 ST and the tumor immune evasion

Recent research in the realm of tumor immune evasion has made significant strides. It is now understood that the intricate interplay among immune cells within TME, encompassing T lymphocytes, macrophages, and regulatory Treg, is crucially linked to the occurrence of immune evasion. Treg are known for their role in curbing the activity of immune effector cells, thus preventing unwarranted tissue damage and quelling inflammatory responses. However, within the inflammatory milieu of a tumor, Treg can undergo reprogramming that augments their suppressive capabilities, leading to a state that either facilitates tumor immune evasion or fosters tumor progression. Strategies that aim to diminish the Treg cell population or attenuate their activity within the tumor's inflammatory TME, while simultaneously impeding their reprogramming, have been shown to bolster the body's anti-tumor immune response (39). Notably, the identification of novel $PD\text{-}L1^+/PD\text{-}L2^+$ macrophage populations that correlate with clinical outcomes suggests that these macrophages might modulate immune responses in the TME through interactions with the T cell surface, playing a significant role in tumor progression and immune evasion (15). Furthermore, in basal-like tumors, epithelial cells under hypoxic conditions have been linked to the upregulation of CD274 and the downregulation of B2M, establishing a connection between hypoxia and the mechanisms underlying immune evasion (29). These findings underscore the complexity of the TME and highlight potential

targets for therapeutic intervention to counteract tumor immune evasion.

ST technology facilitates a sophisticated delineation of the genotype and phenotype of diverse immune cell populations and their states of activation or suppression, unveiling novel therapeutic targets for breast cancer intervention. Claudin-low breast cancer is characterized by a pronounced immune cell infiltration, with heightened presence of B cells, T cells, NK cells, macrophages, and neutrophils relative to other breast cancer subtypes. However, clinical investigations have demonstrated that despite abundant lymphocytic infiltration, a significant number of patients exhibit resistance to immune checkpoint therapies. Beyond CD274, a repertoire of additional immune checkpoint genes, including CD276 and Neuropilin-1 (NPR1), contribute to immunosuppressive mechanisms, thereby circumscribing the efficacy of PD-L1 inhibitory agents (40). In the context of metaplastic breast cancer (MBC), there is evidence of intratumoral permeation by Treg cells, M2-macrophages, and myeloid-derived suppressor cells (MDSCs), which orchestrate an immunosuppressive milieu replete with EMT and hypoxic elements. The interplay with Treg cells is shown to be mediated through signaling pathways involving fibroblast growth factor 2 (FGF2), fibroblast growth factor receptor 1 (FGFR1), and CD44, underscoring the potential therapeutic efficacy of interventions directed at Treg cells in MBC (41). Li CJ et al. (42), employing ST technology, identified an elevated expression of the mitochondrial calcium uniporter (MCU) within tumorigenic regions. Subsequent analyses revealed a positive correlation between MCU expression and the upregulation of pivotal T cell regulatory factors. Within the BRCA invasion cohort, a significant positive association was observed between T cell infiltration and MCU expression, suggesting that MCU not only offers prognostic insights into disease progression but also serves as an indicator of immune status.

4 ST in relation to breast cancer treatment response and guidance of clinical strategies

ST has elucidated subpopulations of cells correlated with tumor metastasis and chemotherapy resistance, enhancing our understanding of the molecular underpinnings of these intricate biological processes. By examining the TME-modulated pharmacological responses, it is possible to predict which patients are poised to garner clinical benefit from immunotherapeutic interventions. Employing single-cell transcriptomics in conjunction with spatial proteomics, the therapeutic efficacy of pembrolizumab in TNBC has been assessed. Tumors that were refractory to treatment demonstrated a dearth of immune cell infiltration both prior to and subsequent to therapy, alongside minimal alterations in immune profiles induced by treatment. In contrast, tumors that responded to therapy could be segregated into two distinct cohorts based on pre-treatment characteristics; one cohort was characterized by elevated expression of major histocompatibility complex molecules and the presence of tertiary lymphoid structures, indicative of pre-existing anti-tumor immune activity. The other cohort, akin to non-responders at baseline, exhibited a pronounced immune response following combined therapeutic intervention, marked by the interactive engagement of cytotoxic T lymphocytes and antigen-presenting myeloid cells (43). Following neoadjuvant therapy in patients with HER2-positive tumors, significant changes occur within the immunological landscape of the tumor. These changes are characterized by a substantial decrease in HER2 expression and its downstream Akt signaling, along with an increased expression of CD45 and CD8, which corresponds to the infiltration of leukocytes and cytotoxic T cells, respectively. Conversely, cases that did not achieve pathological complete response (pCR) are marked by an increase in CD56 expression, which may suggest the lysis of chemotherapy-stressed tumor cells by natural killer (NK) cells (44). ST has been harnessed to investigate the contribution of TNBC tumor cells to the response to neoadjuvant chemotherapy (45), revealing in pCR cases a spatial intermingling of tumor and lymphocytic infiltrates, underpinned by robust activation of interferon (IFN) signaling pathways. Conversely, non-responsive or progressive (pNR) lesions were typified by heightened angiogenic signaling and oxidative metabolism, likely ensuring the requisite energy provision to facilitate the proliferation and architectural reconfiguration necessary for tumor progression. ST offers multidimensional insights into the complexity of TNBC and enables the prognostication of tumor behavior with precision. Collectively, these findings substantiate the utility of ST in affording novel perspectives for the refinement and personalization of therapeutic strategies.

ST is instrumental in crafting therapeutic strategies that are precisely targeted to distinct TME and specific cellular subpopulations, thereby optimizing therapeutic outcomes and curtailing superfluous adverse effects. Trastuzumab serves as an efficacious therapeutic for HER2-positive breast cancer; however, the development of resistance within a year is a common clinical challenge. A novel bispecific antibody, IMM2902, directed against CD47 and HER2, has been engineered to address trastuzumab-resistant breast cancer. Utilizing ST analysis in conjunction with multiplex immunofluorescence (mIFC) and *in vitro* assays, it was determined that IMM2902 is capable of robustly inducing macrophages to secrete C-X-C motif chemokine ligands 9 and 10 (CXCL9 and CXCL10), which are pivotal for the recruitment of T lymphocytes and NK cells to the TME. The integration of IMM2902 into the current therapeutic regimens holds the potential to markedly alter the clinical management of HER2-positive breast cancer, offering a novel avenue of hope for patients, particularly those with limited therapeutic options due to acquired resistance to existing treatments (46). An additional study, employing both single-cell and whole-tissue analytical approaches, has demonstrated a correlation between high levels of inner mitochondrial membrane protein (IMMT) and the immunosuppressive tumor immune microenvironment (TIME). This research substantiates the role of IMMT in the immunosuppressive phenotype of TIME, the proliferation of cancer cells, and mitochondrial adaptive mechanisms, thereby nominating pyridostatin as a promising candidate for targeted therapeutic development in precision medicine (47).

5 ST and prognosis assessment in breast cancer

ST analysis is a cutting-edge technique that plays a pivotal role in discerning cell subpopulations and molecular signatures linked to patient prognostication. This technology affords novel biomarkers that are instrumental for the categorical stratification of breast cancer, the prognostication of therapeutic responses, and the tailoring of individualized treatment paradigms. ST sequencing of disparate regions within clinical breast cancer tissue specimens has unveiled a higher prevalence of follicular helper T cells, quiescent dendritic cells, and plasmacytes within regions abundant in tumor cells, in contrast to areas rich in immune cells where there is a diminished presence of resting $CD4^+$ memory T cells and T regulatory cells. The investigation has pinpointed activated leukocyte cell adhesion molecule (ALCAM), ADP-ribosylation factor-like protein 6-interacting protein 1 (ARL6IP1), and cyclin G2 (CCNG2) as potential immunoprotective agents in breast cancer pathology, while antizyme Inhibitor 1 (AZIN1), myoferlin (MYOF), and transforming acidic coiled-coil containing protein 2 (TACC2) are implicated as potential oncogenes. In locales of elevated tumor cell density, surfeit locus protein 4 (SURF4) and the lipid metabolic gene lysophospholipase I (LYPLA1) have been corroborated as biomarkers inversely related to favorable outcomes. Additionally, diacylglycerol o-acyltransferase 1 (DGAT1), LYPLA1, polymerase (RNA) II (DNA-directed) polypeptide K (POLR2K), and recombinant sphingomyelin phosphodiesterase 4 (SMPD4) are identified as influential factors that modulate patient survival outcomes (48). In the realm of TNBC, which lacks established biomarkers for outcome prediction, spatial profiling has uncovered that caspase 3 and cleaved poly (ADP-ribose) polymerase (cPARP), both indicators of cellular demise, are associated with inferior overall survival when in interaction with the epidermal growth factor receptor (EGFR). The absence of interplay between cells manifesting myoepithelial markers such as smooth muscle actin (SMA) and those indicative of cell cycle progression (mitotic figures) marked by phosphorylated histone H3 (pHH3) is correlated with diminished overall survival. Conversely, the interaction between stromal cells positive for vimentin and those exhibiting active receptor tyrosine kinase (RTK) signaling is associated with enhanced overall survival. These findings underscore the complexity and heterogeneity of the tumor microenvironment and highlight the potential of ST analysis in uncovering new therapeutic targets and prognostic biomarkers in breast cancer (49).

6 Potential challenges and future prospects for the clinical application of ST

These investigations underscore the transformative impact of ST on our apprehension of tumor heterogeneity, offering novel insights and methodological approaches for the molecular profiling of breast cancer, delineation of the TME, assessment of treatment responsiveness, and prognostic stratification, thereby holding the potential to catalyze the evolution of precision oncology in breast cancer. As technological advancements persist and applications are further explored, the anticipation is ripe for a cascade of transformative discoveries that are poised to innovate the diagnostic and therapeutic paradigms in breast cancer management.

Despite the progress made by ST profiling in discovering and identifying disease-specific and spatially specific factors, ST also has its limitations (Table 1). The spatial distribution within ST is subject to the positional integrity of tissue sections relative to the three-dimensional architecture of the organ or tissue and the fluctuating phases of disease progression. Given the considerable variability that can exist between tissue sections of different depths and orientations, there exists a legitimate query as to whether the spatial distribution captured by ST is comprehensively representative of the complete landscape and spectrum of variations within the multi-dimensional organ milieu (50, 51). The precision of pathological identification and selection is equally pivotal to the fidelity and congruence of ST maps and the inferred intercellular dynamics (52). The corpus of human specimens is constrained by the complexities surrounding sample acquisition, preservation, and transit, underscoring an urgent requirement for systematic, rigorously architected clinical research to elucidate the pathophysiological nuances of the disease (53). The translational journey of ST maps into clinical relevance is replete with challenges, stemming from discrepancies in molecular profiling and phenotypic manifestations between animal models and human subjects, as well as among various disease models (51). To encapsulate, the assembly and cartography of spatial profiles necessitate enhanced standardization and automation protocols. The

TABLE 1 Advantages, limitations, and application potential of spatial transcriptomics.

Feature	Description
Advantages	1. Cellular Spatial Localization: Retains spatial information of cells, providing gene expression characteristics *in situ* (10). 2. *In Situ* Tissue Research: Advances the study of genuine gene expression of cells in tissue sites (10). 3. Broad Application Potential: Demonstrates potential in various fields such as tumor, embryonic development, and pathology (51). 4. Biological Interactions: Clarifies interactions between cells and the influence of the microenvironment (11).
Limitations	1. ST Distribution Positional Influence: **The representativeness of ST distribution is influenced by the position of tissue sections and differences in disease stages** (50, 51). 2. Pathological Identification Accuracy**: The accuracy of pathological identification is crucial for the reliability of ST atlases** (52). 3. Human Sample Availability Limitation: **The limited availability of human samples restricts the conclusiveness of disease research** (53). 4. Animal Model-Human Disease Discrepancy: **Differences between animal models and human diseases pose challenges for the clinical application of ST atlases** (51). 5. Cost Issue: Higher cost compared to traditional RNA sequencing methods (54).
Application Potential	1. Revealing Cellular Heterogeneity: Identifying and locating different cell populations within tissues (10). 2. Drug Development: Identifying new biomarkers and drug targets (26, 27). 3. Spatiotemporal Dynamics Analysis: Revealing spatiotemporal dynamics within tissues (11). 4. Multi-Omics Integration: Combining with other omics data to provide a comprehensive perspective (57, 58).

clinical utility of ST is inextricably linked to the precision, reproducibility, and consistency of ST profiling, contingent upon the intricacies, severity, staging, pathological fidelity, and morphological exactitude of the disease context (54).

With the rapid development of new ST technology, data acquisition is continuously improving, and challenges in ST resolution, sensitivity, throughput, and accessibility are being overcome (55). ST is compatible with paraffin-embedded tissues, providing the possibility for retrospective analysis of samples collected in biobanks. This will potentially allow for the systematic detection of various tissues and the reconstruction of the three-dimensional spatial structure of gene expression in organisms (51).

ST has emerged as a transformative approach in breast cancer research, offering innovative vistas and analytical tools that have propelled our comprehension of the TME and the intricacies of tumor heterogeneity. It has also been instrumental in fostering substantial advancements in the realm of breast cancer immunotherapy.

7 Conclusions

The present review synthesizes the burgeoning role of spatial transcriptomics in elucidating the intricate landscapes of breast cancer. It underscores the technology's capacity to delineate the TME and identify cell subpopulations with unprecedented clarity. Despite its promise, spatial transcriptomics is still nascent and confronts hurdles such as cost-efficiency, data intricacy, and the need for analytical standardization.

Ongoing research must refine spatial transcriptomics to augment its resolution and scalability, while concurrently advancing bioinformatics methodologies to adeptly manage and interpret the voluminous datasets. Interdisciplinary cooperation, space multiple omics technology combines genomics, transcriptomics, proteomics and metabolomics, provides tissue space system in all or most of the gene expression level, has been used in neuroscience, development, cancer, plant biology and other fields, improve the biological insights on disease pathogenesis (56–58).

The burgeoning potential of spatial transcriptomics in breast cancer research is palpable, with the anticipation that it will engender transformative diagnostics and therapeutics. As the field matures, we anticipate its pivotal role in the realm of precision medicine, significantly impacting patient prognostication and treatment paradigms.

Author contributions

YZ: Writing – original draft, Writing – review & editing. SG: Data curation, Writing – review & editing. XL: Project administration, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

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.

Publisher's note

References

1. Harbeck N, Gnant M. Breast cancer. *Lancet*. (2017) 389:1134–50. doi: 10.1016/S0140-6736(16)31891-8

2. Burstein HJ, Curigliano G, Thürlimann B, Weber WP, Poortmans P, Reganet MM, et al. Customizing local and systemic therapies for women with early breast cancer: the St. Gallen International Consensus Guidelines for treatment of early breast cancer 2021. *Ann Oncol*. (2021) 32:1216–35. doi: 10.1016/j.annonc.2021.06.023

3. Nolan E, Lindeman GJ, Visvader JE. Deciphering breast cancer: from biology to the clinic. *Cell*. (2023) 186:1708–28. doi: 10.1016/j.cell.2023.01.040

4. Dvir K, Giordano S, Leone JP. Immunotherapy in breast cancer. *Int J Mol Sci*. (2024) 25:7517. doi: 10.3390/ijms25147517

5. Yan Y, Li S, Su L, Tang X, Chen X, Gu X, et al. Mitochondrial inhibitors: a new horizon in breast cancer therapy. *Front Pharmacol*. (2024) 15:1421905. doi: 10.3389/fphar.2024.1421905

6. Jiang C, Zhang S, Jiang L, Chen Z, Chen H, Huang J, et al. Precision unveiled: Synergistic genomic landscapes in breast cancer-Integrating single-cell analysis and decoding drug toxicity for elite prognostication and tailored therapeutics. *Environ Toxicol*. (2024) 39:3448–72. doi: 10.1002/tox.24205

7. Darbandi MR, Darbandi M, Darbandi S, Bado I, Hadizadeh M, Khorram Khorshid HR. Artificial intelligence breakthroughs in pioneering early diagnosis and precision treatment of breast cancer: A multimethod study. *Eur J Cancer*. (2024) 209:114227. doi: 10.1016/j.ejca.2024.114227

8. JiangCL,ZhangSK,JiangL,ChenZP,ChenHQ,HuangJB,etal.Precisionprognostication in breast cancer: unveiling a long noncoding RNA-based model linked to disulfidptosis for tailored immunotherapeutic strategies. *AGING*. (2024) 16. doi: 10.18632/aging.205946

9. Yan Y, Su L, Huang S, He Q, Lu J, Luo H, et al. Circadian rhythms and breast cancer: unraveling the biological clock's role in tumor microenvironment and ageing. *Front Immunol*. (2024) 15:1444426. doi: 10.3389/fimmu.2024.1444426

10. Wang Y, Liu B, Zhao G, Lee Y, Buzdin A, Mu X, et al. Spatial transcriptomics: Technologies, applications and experimental considerations. *Genomics*. (2023) 115:110671. doi: 10.1016/j.ygeno.2023.110671

11. Tian L, Chen F, Macosko EZ. The expanding vistas of spatial transcriptomics. *Nat Biotechnol*. (2023) 41:773–82. doi: 10.1038/s41587-022-01448-2

12. Xue R, Zhang Q, Cao Q, Kong R, Xiang X, Liu H, et al. Liver tumour immune microenvironment subtypes and neutrophil heterogeneity. *Nature*. (2022) 612:141–47. doi: 10.1038/s41586-022-05400-x

13. Xun Z, Ding X, Zhang Y, Zhang B, Lai S, Zou D, et al. Reconstruction of the tumor spatial microenvironment along the Malignant-boundary-nonmalignant axis. *Nat Commun*. (2023) 14:933. doi: 10.1038/s41467-023-36560-7

14. Qi J, Sun H, Zhang Y, Wang Z, Xun Z, Li Z, et al. Single-cell and spatial analysis reveal interaction of FAP+ fibroblasts and SPP1+ macrophages in colorectal cancer. *Nat Commun*. (2022) 13:1742. doi: 10.1038/s41467-022-29366-6

15. Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, et al. A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet*. (2021) 53:1334–47. doi: 10.1038/s41588-021-00911-1

16. Lv J, Shi Q, Han Y, Li W, Liu H, Zhang J, et al. Spatial transcriptomics reveals gene expression characteristics in invasive micropapillary carcinoma of the breast. *Cell Death Dis*. (2021) 12:1095. doi: 10.1038/s41419-021-04380-6

17. Yoshitake R, Mori H, Ha D, Wu X, Wang J, Wang X, et al. Molecular features of luminal breast cancer defined through spatial and single-cell transcriptomics. *Clin Transl Med*. (2024) 14:e1548. doi: 10.1002/ctm2.1548

18. Sun H, Li Y, Zhang Y, Zhao X, Dong X, Guo Y, et al. The relevance between hypoxia-dependent spatial transcriptomics and the prognosis and efficacy of immunotherapy in claudin-low breast cancer. *Front Immunol*. (2023) 13:1042835. doi: 10.3389/fimmu.2022.1042835

19. Kumar T, Nee K, Wei R, He S, Nguyen QH, Bai S, et al. A spatially resolved single-cell genomic atlas of the adult human breast. *Nature*. (2023) 620:181–91. doi: 10.1038/s41586-023-06252-9

20. Liu YM, Ge JY, Chen YF, Liu T, Chen L, Liu CC, et al. Combined single-cell and spatial transcriptomics reveal the metabolic evolvement of breast cancer during early dissemination. *Adv Sci (Weinh)*. (2023) 10:e2205395. doi: 10.1002/advs.202205395

21. Shah OS, Nasrazadani A, Foldi J, Atkinson JM, Kleer CG, McAuliffe PF, et al. Spatial molecular profiling of mixed invasive ductal and lobular breast cancers reveals heterogeneity in intrinsic molecular subtypes, oncogenic signatures, and mutations. *Proc Natl Acad Sci U S A*. (2024) 121:e2322068121. doi: 10.1073/pnas.2322068121

22. Wang H, Zhao J, Nie Q, Zheng C, Sun X. Dissecting spatiotemporal structures in spatial transcriptomics *via* diffusion-based adversarial learning. *Res (Wash D C)*. (2024) 7:390. doi: 10.34133/research.0390

23. Siraj AK, Beg S, Jehan Z, Prabhakaran S, Ahmed M, Hussain AR, et al. ALK alteration is a frequent event in aggressive breast cancers. *Breast Cancer Res*. (2015) 17:127. doi: 10.1186/s13058-015-0610-3

24. Zhao S, Li J, Xia Q, Liu K, Dong Z. New perspectives for targeting therapy in ALK-positive human cancers. *Oncogene*. (2023) 42:1959–69. doi: 10.1038/s41388-023-02712-8

25. Zishui F, Tian Y, Sui C, Guo Y, Hu X, Lai Y, et al. Single-cell transcriptomics of proliferative phase endometrium: systems analysis of cell–cell communication network using CellChat. *Front Cell Dev Biol*. (2022) 10:919731. doi: 10.3389/fcell.2022.919731

26. Nagasawa S, Kuze Y, Maeda I, Kojima Y, Motoyoshi A, Onishi T, et al. Genomic profiling reveals heterogeneous populations of ductal carcinoma in *situ* of the breast. *Commun Biol*. (2021) 4:438. doi: 10.1038/s42003-021-01959-9

27. Yu X, Li W, Sun S, Li J. DDIT3 is associated with breast cancer prognosis and immune microenvironment: an integrative bioinformatic and immunohistochemical analysis. *J Cancer*. (2024) 15:3873–89. doi: 10.7150/jca.96491

28. Roma-Rodrigues C, Mendes R, Baptista PV, Fernandes AR. Targeting tumor microenvironment for cancer therapy. *Int J Mol Sci*. (2019) 20:840. doi: 10.3390/ijms20040840

29. Ali HR, Jackson HW, Zanotelli VRT, Danenberg E, Fischer JR, Bardwell H, et al. Imaging mass cytometry and multiplatform genomics define the phenogenomic landscape of breast cancer. *Nat cancer*. (2020) 1:163–75. doi: 10.1038/s43018-020-0026-6

30. Croizer H, Mhaidly R, Kieffer Y, Gentric G, Djerroudi L, Leclere R, et al. Deciphering the spatial landscape and plasticity of immunosuppressive fibroblasts in breast cancer. *Nat Commun*. (2024) 15:2806. doi: 10.1038/s41467-024-47068-z

31. Honda CK, Kurozumi S, Fujii T, Pourquier D, Khellaf L, Boissiere F, et al. Cancer-associated fibroblast spatial heterogeneity and EMILIN1 expression in the tumor microenvironment modulate TGF-β activity and CD8 T-cell infiltration in breast cancer. *Theranostics*. (2024) 14:1873–85. doi: 10.7150/thno.90627

32. Mao X, Zhou D, Lin K, Zhang B, Gao J, Ling F, et al. Single-cell and spatial transcriptome analyses revealed cell heterogeneity and immune environment alternations in metastatic axillary lymph nodes in breast cancer. *Cancer Immunol Immunother*. (2023) 72:679–95. doi: 10.1007/s00262-022-03278-2

33. Peña-Romero AC, Orenes-Piñero E. Dual effect of immune cells within tumour microenvironment: pro- and anti-tumour effects and their triggers. *Cancers (Basel)*. (2022) 14:1681. doi: 10.3390/cancers14071681

34. Salmaninejad A, Layeghi SM, Falakian Z, Golestani S, Kobravi S, et al. An update to experimental and clinical aspects of tumor-associated macrophages in cancer development: hopes and pitfalls. *Clin Exp Med*. (2024) 24:156. doi: 10.1007/s10238-024-01417-w

35. Kuett L, Catena R, Özcan A, Plüss A, Cancer Grand Challenges IMAXT Consortium, , Schraml P, et al. Three-dimensional imaging mass cytometry for highly multiplexed molecular and cellular mapping of tissues and the tumor microenvironment. *Nat cancer*. (2022) 3:122–33. doi: 10.1038/s43018-021-00301-w

36. Romanens L, Chaskar P, Marcone R, Ryser S, Tille JC, Genolet R, et al. Clonal expansion of intra-epithelial T cells in breast cancer revealed by spatial transcriptomics. *Int J Cancer*. (2023) 153:1568–78. doi: 10.1002/ijc.34620

37. Liu SQ, Gao ZJ, Wu J, Zheng HM, Li B, Sun S, et al. Single-cell and spatially resolved analysis uncovers cell heterogeneity of breast cancer. *J Hematol Oncol*. (2022) 15:19. doi: 10.1186/s13045-022-01236-0

38. 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

39. Wu X, Zhou Z, Cao Q, Chen Y, Gong J, Zhang Q, et al. Reprogramming of Treg cells in the inflammatory microenvironment during immunotherapy: a literature review. *Front Immunol*. (2023) 14:1268188. doi: 10.3389/fimmu.2023.1268188

40. Sun H, Li Y, Zhang Y, Zhao X, Dong X, Guo Y, et al. The relevance between hypoxia-dependent spatial transcriptomics and the prognosis and efficacy of immunotherapy in claudin-low breast cancer. *Front Immunol*. (2023) 13:1042835. doi: 10.3389/fimmu.2022.1042835

41. He S, Jin Y, Nazaret A, Shi L, Chen X, Rampersaud S, et al. Starfysh integrates spatial transcriptomic and histologic data to reveal heterogeneous tumor-immune hubs. *Nat Biotechnol*. (2024). doi: 10.1038/s41587-024-02173-8

42. Li CJ, Tzeng YT, Hsiao JH, Tseng LM, Hsu TS, Chu PY. Spatial and single-cell explorations uncover prognostic significance and immunological functions of mitochondrial calcium uniporter in breast cancer. *Cancer Cell Int*. (2024) 24:140. doi: 10.1186/s12935-024-03327-z

43. Shiao SL, Gouin KH 3rd, Ing N, Ho A, Basho R, Shah A, et al. Single-cell and spatial profiling identify three response trajectories to pembrolizumab and radiation therapy in triple negative breast cancer. *Cancer Cell*. (2024) 42:70–84.e8. doi: 10.1016/j.ccell.2023.12.012

44. McNamara KL, Caswell-Jin JL, Joshi R, Ma Z, Kotler E, Bean GR, et al. Spatial proteomic characterization of HER2-positive breast tumors through neoadjuvant therapy predicts response. *Nat cancer*. (2021) 2:400–13. doi: 10.1038/s43018-021-00190-z

45. Donati B, Reggiani F, Torricelli F, Santandrea G, Rossi T, Bisagni A, et al. Spatial distribution of immune cells drives resistance to neoadjuvant chemotherapy in triple-negative breast cancer. *Cancer Immunol Res*. (2024) 12:120–34. doi: 10.1158/2326-6066.CIR-23-0076

46. Zhang B, Shi J, Shi X, Xu X, Gao L, Li S, et al. Development and evaluation of a human CD47/HER2 bispecific antibody for Trastuzumab-resistant breast cancer immunotherapy. *Drug Resist Updat*. (2024) 74:101068. doi: 10.1016/j.drup.2024.101068

47. Lin HY, Wu HJ, Chu PY. Multi-omics and experimental analysis unveil theragnostic value and immunological roles of inner membrane mitochondrial protein (IMMT) in breast cancer. *J Transl Med*. (2023) 21:189. doi: 10.1186/s12967-023-04035-4

48. Liu G, Wang L, Ji L, He D, Zeng L, Zhuo G, et al. Identifying prognostic markers in spatially heterogeneous breast cancer microenvironment. *J Transl Med*. (2023) 21:580. doi: 10.1186/s12967-023-04395-x

49. Tanevski J, Flores ROR, Gabor A, Schapiro D, Saez-Rodriguez J. Explainable multiview framework for dissecting spatial relationships from highly multiplexed data. *Genome Biol*. (2022) 23:97. doi: 10.1186/s13059-022-02663-5

50. You Y, Fu Y, Li L, Zhang Z, Jia S, Lu S, et al. Systematic comparison of sequencing-based spatial transcriptomic methods. *Nat Methods*. (2024) 21:1743–54. doi: 10.1038/s41592-024-02325-3

51. Zhang L, Chen D, Song D, Liu X, Zhang Y, Xu X, et al. Clinical and translational values of spatial transcriptomics. *Signal Transduct Target Ther*. (2022) 7:111. doi: 10.1038/s41392-022-00960-w

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

53. Saul D, Kosinsky RL. Spatial transcriptomics herald a new era of transcriptome research. *Clin Transl Med*. (2023) 13:e1264. doi: 10.1002/ctm2.1264

54. Bressan D, Battistoni G, Hannon GJ. The dawn of spatial omics. *Science*. (2023) 381:eabq4964. doi: 10.1126/science.abq4964

55. Tian L, Chen F, Macosko EZ. The expanding vistas of spatial transcriptomics. *Nat Biotechnol*. (2023) 41:773–82. doi: 10.1038/s41587-022-01448-2

56. Eisenstein M. Seven technologies to watch in 2022. *Nature*. (2022) 601:658–61. doi: 10.1038/d41586-022-00163-x

57. Vicari M, Mirzazadeh R, Nilsson A, Shariatgorji R, Bjärterot P, Larsson L, et al. Spatial multimodal analysis of transcriptomes and metabolomes in tissues. *Nat Biotechnol*. (2024) 42:1046–50. doi: 10.1038/s41587-023-01937-y

58. Li B, Zhang W, Guo C, Xu H, Li L, Fang M, et al. Benchmarking spatial and single-cell transcriptomics integration methods for transcript distribution prediction and cell type deconvolution. *Nat Methods*. (2022) 19:662–70. doi: 10.1038/s41592-022-01480-9

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 10 October 2024
DOI 10.3389/fimmu.2024.1470130

OPEN ACCESS

EDITED BY
Mar Requena Mullor,
University of Almeria, Spain

REVIEWED BY
Priyanka S. Rana,
Case Western Reserve University, United States
Ravi Kumar Sharma,
All India Institute of Medical Sciences Bilaspur, India

*CORRESPONDENCE
Sergey Sennikov
sennikov@niikim.ru

RECEIVED 25 July 2024
ACCEPTED 23 September 2024
PUBLISHED 10 October 2024

CITATION
Sennikov S, Volynets M, Alrhmoun S, Perik-Zavodskii R, Perik-Zavodskaia O, Fisher M, Lopatnikova J, Shevchenko J, Nazarov K, Philippova J, Alsalloum A, Kurilin V and Silkov A (2024) Modified Dendritic cell-based T-cell expansion protocol and single-cell multi-omics allow for the selection of the most expanded and *in vitro*-effective clonotype via profiling of thousands of MAGE-A3-specific T-cells. *Front. Immunol.* 15:1470130. doi: 10.3389/fimmu.2024.1470130

Modified Dendritic cell-based T-cell expansion protocol and single-cell multi-omics allow for the selection of the most expanded and *in vitro*-effective clonotype via profiling of thousands of MAGE-A3-specific T-cells

Sergey Sennikov*, Marina Volynets, Saleh Alrhmoun, Roman Perik-Zavodskii, Olga Perik-Zavodskaia, Marina Fisher, Julia Lopatnikova, Julia Shevchenko, Kirill Nazarov, Julia Philippova, Alaa Alsalloum, Vasily Kurilin and Alexander Silkov

Laboratory of Molecular Immunology, Research Institute of Fundamental and Clinical Immunology, Novosibirsk, Russia

Introduction: Adoptive cell therapy using TCR-engineered T-cells is one of the most effective strategies against tumor cells. The TCR T-cell approach has been well tested against a variety of blood neoplasms but is yet to be deeply tested against solid tumors. Among solid tumors, cancer-testis antigens are the most prominent targets for tumor-specific therapy, as they are usually found on cells that lie behind blood-tissue barriers.

Methods: We have employed a novel efficient protocol for MAGE-A3-specific T-cell clonal expansion, performed single-cell multi-omic analysis of the expanded T-cells via BD Rhapsody, engineered a selected T-cell receptor into a lentiviral construct, and tested it in an *in vitro* LDH-cytotoxicity test.

Results and discussion: We have observed a 191-fold increase in the MAGE-A3-specific T-cell abundance, obtained a dominant T-cell receptor via single-cell multi-omic BD Rhapsody data analysis in the TCRscape bioinformatics tool, and observed potent cytotoxicity of the dominant-clonotype transduced TCR T-cells against a MAGE-A3-positive tumor. We have demonstrated the efficiency of our T-cell enrichment protocol in obtaining potent anti-tumor T-cells and their T-cell receptors, especially when paired with the modern single-cell analysis methods.

KEYWORDS

MAGE-A3, ScRNA-seq, scTCR-seq, TCR, T-cell receptor repertoire, TCR T-cells, adoptive cell therapy, naturally-occurring T-cells

1 Introduction

Melanoma-associated gene (MAGE) protein family were the first proteins identified from the class of cancer-testis antigens whose expression pattern was restricted to germline cells and immune-privileged testes and placenta cells (1). In tumor cells, expression of cancer-testis antigens is associated with apoptosis avoidance, increased viability, migration with subsequent metastasis, and angiogenesis (2). More than sixty proteins belong to the MAGE family, but only type I proteins whose expression is restricted to the X chromosome belong to the class of cancer-testicular antigens (3). The first type includes the MAGE-A, -B, and -C subfamilies. MAGE family proteins have a MAGE homology domain (MHD) consisting of approximately one hundred and seventy amino acid residues (4).

The MHD is 46% conserved in the human population. Expression of MAGE-A subfamily proteins in patients with cancer correlates with poor clinical prognosis as well as increased recurrence after therapy (5). MAGE-A subfamily proteins are widely used as target epitopes for immunotherapy because they are present in a large number of tumor types (6). In 2009, the National Cancer Institute (NCI) ranked seventy-five tumor-associated antigens according to characteristics important for selecting an antigen as a target for immunotherapy (e.g. immunogenicity and oncogenicity). MAGE-A3 was ranked eighth (7).

One of the possible approaches to the treatment of MAGE-A3-positive tumors is TCR T-cells. Several clinical trials on the use of high-affinity TCR-T-cells with genetically modified antigen-recognition receptors specific to MAGE-A3 epitopes were prematurely terminated due to lethal outcomes, as the increase in T-cell receptor affinity may lead to off-target activity (8, 9). In the first case, some patients developed severe neurological toxicity due to possible cross-reactivity of the T-cell receptor with a highly homologous epitope of the MAGE-A12 protein (KMAELVHFL), which can be normally expressed in brain cells (8). In the second case, severe cardiac toxicity was observed due to additional recognition of the Titin protein epitope (ESDPIVAQY), which is specific for transverse striated muscles and is found in the myocardium (10).

T-cells during their natural encounters with antigen-presenting Dendritic cells (DCs) can undergo clonal expansion if they recognize the presented peptide (11). Clonal expansion is the process of selection of the most affine T-cell receptor (TCR), i.e. the best TCR clonotype (12). T-cell receptor is composed of two chains: alpha and beta, each of which is, in turn, composed of 4 framework (FR) and 3 complementarity-determining regions (CDR), among which beta-chain CDR3 is the most important as it is mostly responsible for the recognition of the antigen the TCR is specific to.

Previously TCR genomics were facilitated by bulk RNA-seq that had intrinsic inability to properly pair TCR alpha- and beta-chains (13). This insufficiency was later solved by the single-cell multi-omics that currently allow for either targeted or full transcriptome and surface proteome analysis (14), as well as full-length TCR sequencing, proper TCR chain pairing, along with the profiling of transcriptome and proteome of the studied T-cells (15–17).

In this work, we implemented a novel T-cell enrichment protocol (based on the peptide-loaded dendritic cell T-cell induction) to enrich for the MAGE-A3 antigen-specific T-cells (using KVAELVHFL peptide), performed scRNA-seq of the enriched MAGE-A3-specific T-cells, found the Dominant MAGE-A3-specific clonotype, and assessed its effectiveness in an *in vitro* LDH-based cytotoxicity test.

2 Materials and methods

2.1 Study population

The study population consisted of conditionally healthy adult donors (n = 7) who were preselected for the presence of the HLA-A02 allele via flow cytometry using PE antibodies against human HLA-A02 (343306, Biolegend, United States) and the Attune Nxt flow cytometer (A24858, Thermo Fisher Scientific, Waltham, Massachusetts, United States). The average age of the donors was 27.33 ± 6.34 years (mean ± SE) (male donors n = 3, female donors n = 4). All donors signed written informed consent to participate in the study.

2.2 PBMC isolation

We collected peripheral blood (n = 7) in vacuum tubes with EDTA and isolated PBMCs using the Ficoll™ (PanEco, Russia) density gradient centrifugation method. We performed the centrifugation at 400 g at room temperature for 40 min and then collected the buffy-coat cells.

2.3 T-cell clonal expansion protocol

We used peptide-loaded dendritic cell (DCs)-based T-cell clonal expansion protocol to induce the single-cell capture-compatible cell number of antigen-specific cytotoxic T lymphocytes *in vitro*. We utilized DCs as they have a unique ability to capture and present antigens in combination with MHC class I and class II molecules for the activation of naïve T-cells, which, in turn, leads to the clonal expansion and differentiation of the naïve T-cells into the effector T-cells (11, 18). In this paper, we developed a multi-stage, maximally optimized, and efficient protocol, based on a previous well-established in-lab protocol that has proven its effectiveness (19–21).

2.3.1 Obtaining DCs

The PBMCs (n = 7) were cultivated in RPMI-1640 medium, which was enriched with 10% fetal calf serum (FCS) (Biowest, Nuaillé, France), 2 mM L-glutamine (Biolot, Saint Petersburg, Russia), 5×10^{-4} M 2-mercaptoethanol (Sigma-Aldrich, St. Louis, MO, USA), 25 mM HEPES (Biolot, Saint Petersburg, Russia), 80 µg/mL gentamicin (KRKA, Novomesto, Slovenia), and 100 µg/mL benzylpenicillin (Biolot, Saint Petersburg, Russia). This

combination is referred to as the culture medium. The obtained PBMCs (80-100 million) were incubated in a culture vial with a surface area of 150 cm² (TPP, Switzerland) at a concentration of 1-2 million cells/mL for 30 minutes in a CO_2 incubator. The non-adherent cells were then used to isolate CD8+ cells, while the adherent cells were detached from the plastic surface using a cell scraper (TPP, Switzerland). The adherent fraction of PBMCs (20-25 million cells) was cultured in a 150 cm² culture vial (TPP, Switzerland) in the presence of recombinant human granulocyte-macrophage colony-stimulating factor (100 ng/mL, BioLegend, United States) and interleukin-4 (50 ng/mL, BioLegend, United States) for 4 days to generate immature dendritic cells (DCs). Partial media replacement was performed on the 3rd day of culture. On the 5th day, the cells were harvested using a cell scraper, counted, and transferred to 12-well plates (2 mL/well, 1 mL/well). The MAGE-A3 HLA-A02-binding peptide KVAELVHFL peptide (p112-120, Immunotex, Stavropol, Russia) was added to the cell culture at a concentration of 30 μg/mL, followed by the induction of DC maturation on the 6th day using TNF-alpha (25 ng/mL, BioLegend, United States). The HLA-A02 was chosen since it is the most frequent class I HLA genotype among almost all human populations (22).

2.3.2 Obtaining CD8+ T-cells

We used the non-adhesive fraction of cells (n = 7) from the previous paragraph for the isolation of CD8+ T-cells via MojoSort™ Human CD8 T-Cell Isolation Kit (480012, BioLegend, United States) according to the manufacturer's instructions. Immediately after the isolation of CD8+ T-cells, we added IL-7 (581906, BioLegend, United States), IL-15 (570306, BioLegend, United States), and IL-2 (589106, BioLegend, United States) each at a concentration of 10 ng/ml and cultivated the CD8+ cells, at a concentration of 2 million cells/mL, for 6 days. We performed total media replacement on the 3rd day of cultivation and the addition of another dose of IL-2/7/15, alongside anti-CD3 antibody (0.5 μg/ml) (830301, BioLegend, United States) and anti-CD28 antibody (1 μg/ml) (302902, BioLegend, United States). We saved the 6-day-cultivation conditioned media for later use.

2.3.3 DCs and CD8+ T-cells co-culture

Antigen-loaded mature DCs and T-cells were collected from the culture vessel surfaces using a cell scraper and subsequently co-cultured in a new 75 cm² vial (TPP, Switzerland) at a ratio of 1:10 (DCs: CD8+). The first 3 days of co-cultivation were carried out without additional stimulation to selectively eliminate cells that are not receiving stimulation through their T-cell receptor from the mature DCs. On the 4th day of the protocol we added anti-CD3 antibody (0.5 μg/ml) (830301, BioLegend, United States), anti-CD28 antibody (1 μg/ml) (302902, BioLegend, United States), IL-7 (581906, BioLegend, United States), IL-15 (570306, BioLegend, United States), and IL-2 (589106, BioLegend, United States) each at a concentration of 10 ng/ml the co-culture media to maintain cellular viability and promote proliferation of the obtained CD8+ T-cells.

2.3.4 Antigen-specific CD8+ T-cell isolation

On the 8th day of DCs and CD8+ T-cells co-culturing (n = 7), we isolated antigen-specific T-cells from the cell co-culture using a two-step isolation process. First, the DCs were eliminated using the MojoSort negative magnetic selection Human CD8 T-cell Isolation Kit (480012, BioLegend, United States). Second, the antigen-specific T-cells were sorted using Flex-T tetramers: we loaded MHC tetramers (HLA-A*02:01) with the MAGE-A3 HLA-A2-binding peptide KVAELVHFL (p112-120) (Immunotex, Stavropol, Russia) and labeled the Flex-T MHC tetramers (BioLegend, United States) with either phycoerythrin (PE) or allophycocyanin (APC) according to the manufacturer's instructions, and sorted the double-positive (APC-tetramer and PE-tetramer) lymphocytes on a BD FACS Aria I sorter (pressure: 20 psi, mode: "Purity", speed: 2500 events/sec) (BD Biosciences, Franklin Lakes, New Jersey, United States). The purity of the sorted cell populations ranged from 67% to 88%, yielding between 25,000 to 170,000 antigen-specific cells in total. Additional details can be found in the Supplementary Materials and Methods and in Supplementary Figure 1.

2.3.5 Stimulation of cell proliferation

Following the CD8 T-cell sorting, we transferred the antigen-specific T-cells (n = 7) to a flat-bottomed culture plate at a concentration of 2-4 million cells/mL, where we cultured them for 14 days in the presence of the following T-cell stimulating agents: anti-CD3 antibody (0.5 μg/ml) (830301, BioLegend, United States), anti-CD28 antibody (1 μg/ml) (302902, BioLegend, United States), IL-7 (581906, BioLegend, United States), IL-15 (570306, BioLegend, United States), and IL-2 (589106, BioLegend, United States) each at a concentration of 10 ng/ml. The culture media consisted of equal parts of RPMI-1640 culture media (Biolot, Russia) and the pre-saved CD8+ T-cell conditioned media. The total duration of cultivation of antigen-specific T-cells was 21 days. We visualized the enrichment results in GraphPad Prism 10.2.3 using bar plots.

2.4 Sample tag sample barcoding and cell counting for BD Rhapsody single-cell analysis

We incubated cells from different individuals (n = 3, other 4 donors were discarded due to the low number of antigen-specific T-cells enriched) with Sample Tag antibodies (BD Biosciences, Franklin Lakes, New Jersey, United States) for 20 minutes at room temperature according to the BD Rhapsody Single-Cell Analysis System User Guide Revision 5.0 (BD Biosciences, Franklin Lakes, New Jersey, United States). After three washing cycles, we stained the cells with Calcein, counted them using the Attune NxT flow cytometer (Thermo Fisher, United States), pooled the samples together, and resuspended them in a cold sample buffer to a final concentration of 60 cells/μl and a volume of 620 μl for loading onto a BD Rhapsody Cartridge. The quality of cell loading into the cartridge was assessed using the InCell Analyzer 6000 with the help of Calcein AM (GE Healthcare, United States).

2.5 cDNA library preparation and sequencing

We utilized the BD Rhapsody Express Single-Cell Analysis System (BD Biosciences, Franklin Lakes, New Jersey, United States) for single-cell capture and cDNA library preparation, following the manufacturer's "TCR/BCR Full Length, Targeted mRNA, and Sample Tag Library Preparation" Protocol. In summary, single cells were captured in the BD Rhapsody cartridge, and magnetic beads were introduced for poly-A mRNA capture. Cells were lysed, and reverse transcription was carried out on the magnetic beads with the captured poly-A mRNA. A template switch oligo was then added, followed by another round of reverse transcription. The Sample Tag cDNA was then denatured, Sample Tag PCR 1 was performed, and bead cDNA was extended using Klenow DNA polymerase fragment. Beads were treated with Exonuclease I, and the cDNA was amplified using TCR primers. The TCR amplicons were denatured and collected, followed by the Human Immune Response Primer Panel on the cDNA (targeting 397 genes with 399 primer pairs) to collect the mRNA panel amplicons. PCR1 products were purified using AMPure XP magnetic beads (A63880, Beckman Coulter, Brea, California, United States) and separated by amplicon size into TCR, mRNA panel, and Sample Tag products. Further amplification and size selection clean-up was carried out on the mRNA and Sample Tag PCR1 products yielding PCR2 mRNA and Sample Tag products. TCR amplicons were normalized to 1.5 ng/μL, followed by random primer extension (TCR RPE) with Klenow DNA polymerase fragment and TCR RPE library clean-up by double-sided selection. Concentrations of PCR2 of mRNA and Sample Tag products and TCR RPE products were measured using Qubit High-Sensitivity dsDNA Kit (Q33231, Thermo Fisher, Waltham, Massachusetts, United States). The final products were then normalized to 4.5 ng/μL for the mRNA panel library and 1.0 ng/μL for the Sample Tag library and the RPE TCR products were used undiluted for the TCR library, and final amplification was performed with Illumina indexes to prepare the libraries. The final libraries were quantified with Qubit 4 and Agilent BioAnalyzer 2100 (Agilent, Santa Clara, California, United States), then pooled (~83/11/5% TCR/mRNA/Sample Tag ratio, estimating 15000 (TCR), 2000 (mRNA), and 1000 (Sample Tag) reads per cell) to a final concentration of 2 nM. Sequencing was performed on a NovaSeq 6000 sequencer (Illumina, San Diego, California, United States) using an SP flow cell with (R1 = 85, R2 = 225, 600 million clusters).

2.6 Sequencing data processing

We processed the FASTQ files using the BD Rhapsody pipeline v1.12 (BD Biosciences, Franklin Lakes, New Jersey, United States). The pipeline first filtered out low-quality read pairs based on criteria such as read length, highest single-nucleotide frequency, and mean base quality score. It then analyzed the remaining high-quality R1 reads to identify cell label and unique molecular identifier (UMI) sequences. High-quality R2 reads were aligned to the reference panel sequences (mRNA) using Bowtie2. Reads with the same cell label, UMI sequence, and gene were collapsed into single molecules. UMI counts were adjusted using error correction algorithms — recursive substitution error correction (RSEC) and distribution-based error correction (DBEC) — to mitigate errors from both sequencing and PCR. Cell counts were estimated using second derivative analysis to filter out noise cell labels; only cell labels beyond a single observed inflection point were considered valid. The pipeline then used the sample tags (single-cell multiplexing kit; BD Biosciences) for sample demultiplexing and to exclude multiplets, identifying a total of 5.491 single cells. Following this, the pipeline aligned TCR RPE library reads on a per-cell basis to create TCR contigs, annotated these contigs, and generated gene expression (gene/cell) matrices (GEX matrices) for each biological sample. Additionally, a cumulative Adaptive Immune Receptor Repertoire (AIRR) matrix was created for the TCR contigs.

2.7 ERGO-II TCR-peptide-MHC affinity prediction

We downloaded the ERGO-II neural network repository (23) and initiated the tool from the terminal with the selection of the input file and database (McPAS-TCR) (24). McPAS-TCR is a manually curated database based on published literature containing information on over twenty thousand T-cell receptor sequences, the antigens they bind to, T-cell type (CD4+/CD8+), and MHC type (MHC-I/MHC-II). McPAS-TCR includes information about T lymphocytes that expand in various human or mouse pathological conditions (including viral infections, cancer, and autoimmune reactions). The CSV input file for ERGO-II (a pre-ERGO-II-generated input file, (25, pre ERGO-II.ipynb) contains information about TCR CDR3α and CDR3β sequence, peptide sequence, MHC type (MHC-I/MHC-II class), V and J genes, and T-cell type (CD4+/CD8+), the pre-ERGO-II-generated input file also contained single cell indices for later data frame merging. The output file contains a prediction value for the score of T-cell receptor binding to the peptide/MHC complex, which ranges from 0 to 1, where 0 is the minimum or no affinity and 1 is the maximum affinity.

2.8 TCRscape clonotype selection

We imported the GEX matrices of each biological sample, the multi-sample AIRR matrix, and the ERGO-II output file into TCRscape (25), merged the GEX matrices, performed "Counts Per Million" data normalization, replaced the zeroes in the data frame with the ones, log2-transformed the data, gated CD8+ T-cells, counted dominant full-length T-cell receptor clonotypes of the gated T-cells, created a merged data frame containing the gene expression of the key T-cell markers (*CD4, CD8, FOXP3*), clonotype information and ERGO-II-generated binding scores, performed Principal Component Analysis (PCA) on the merged

data frame, assessed the dimensionality of the merged data frame by Scree plot, and performed Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction using the first 4 principal components.

2.9 Plasmid construction and lentivirus preparation

We generated TCR-containing transfer plasmid, VSV-G encoding plasmid, and two third generation lentivirus packaging plasmids (containing Gag-Pol and Rev genes respectively) in E. coli NEB Stable strain (C3040I, New England Biolabs, Ipswich, Massachusetts, United States), and verified the resulting plasmids using restriction enzymes and gel electrophoresis. We then delivered the above plasmids using Lipofectamine 2000 (11668500, Thermo Fisher, Waltham, Massachusetts, United States) into HEK-293T packaging cells generously provided by Dr. Hiroshi Shiku (Mie University, Japan).

2.10 Lentivirus concentration and titration using qPCR

We concentrated the produced lentiviruses using the commercial TransLv™ Lentivirus Precipitation Solution (5×) (FV101-01, TransGen, China) and titrated them using quantitative (with dilution standards) PCR for proviral DNA (TransLv™ Lentivirus qPCR Titration Kit) (FV201-01, TransGen, China) in HEK-293T.

2.11 TCR T-cell manufacturing

To obtain TCR T-cells specific for MAGE-A3, Retronectin (Sci Store, Russia) at a concentration of 25 μg/mL and anti-CD3 antibodies (BioLegend) at 5 μg/mL in ACDA (citrate buffer with glucose) were adsorbed one day prior to the experiment. This solution was applied to the wells of a 12-well plate at 415 μL per well. CD3+ T-cells were isolated from the PBMCs (n = 7) of HLA A02-positive healthy donors using MojoSortTM Human CD3 negative magnetic Selection Kit (480134, Biolegend, United States). The isolated cells (0.75 million/mL) were then incubated in the plate with retronectin and anti-CD3 antibodies, supplemented with IL-2 (300 units/mL, Biotech LLC, Russia), for 48 hours.

The day before transduction, Retronectin (Sci Store, Russia) was adsorbed at a concentration of 25 μg/mL in ACDA. This solution was applied to the wells of a 24-well plate at 255 μL per well. On the day of transduction, cells were harvested, centrifuged, resuspended in a serum-free medium, and counted. Then, 200,000 cells were transferred into Retronectin-coated wells at a volume of 500 μL/well. After that, to each well we added lentivirus (200,000 particles/well, i.e. at a multiplicity of infection (MOI) of 1) and protamine sulfate (5-8 μg/mL) to enhance transduction (26). The plate was then centrifuged for 2 hours at 600 x g and 32°C. After centrifugation, 500 μL of warm serum-free medium containing IL-2 (final concentration 300 units/mL) was added to the cells that were allowed to incubate overnight. The following morning, cells were transferred into the wells of a 12-well plate with an equal volume of complete medium containing IL-2. Cell growth, conglomerate formation, and nutrient medium condition were monitored visually, and growth factors were refreshed every two days.

2.12 *In vitro* TCR T-cell assessment in a cytotoxicity test

To assess cytotoxic activity against tumor cells, transduced or vehicle-transduced (with a lentivirus without transfer plasmid) cells were harvested, centrifuged at 350 g for 10 minutes, and counted. Tumor cells, in the logarithmic phase of growth, were harvested using a 1:3 mixture of trypsin (0.25%) (PanEco, Russia) and Versen solution (Vector, Russia). Tumor cells were seeded into a 96-well flat-bottom plate at a concentration of 5,000 cells/well. T-cells (50,000 cells/well) were added 2-3 hours later, resulting in an effector-to-target ratio of 10:1. The co-culture was allowed to incubate for 16-18 hours in a medium containing 5% FCS. Forty-five minutes before the end of the incubation period, 10 μL of 10X lysing solution was added per 100 μL of cell suspension to control the maximum release of lactate dehydrogenase (LDH) from the cells.

After completing the lysis, the cell plate was centrifuged at 250 × g for 4 minutes to gently pellet the cells. Aliquots of 50 μL from each well were then transferred to a new 96-well flat-bottom plate for the immunoassay. To each well of cell culture supernatants, 50 μL of reconstituted lactate dehydrogenase (LDH) enzyme-substrate mixture was added. The plate was covered with foil or an opaque cover slip to protect it from light and was incubated for 30 minutes at room temperature. After the incubation, the reaction was stopped with 1M acetic acid solution. Optical density was measured at 490 or 492 nm immediately after stopping the reaction. Cytotoxic activity was calculated in doubles using the following formula:

$$\%\ Cytotoxicity = \frac{OD\ (T-cells\ +\ targets) - OD\ (T-cells)}{OD\ (maximal\ tumor\ lysis) - OD\ (spontaneous\ tumor\ lysis)} \times 100$$

We then analyzed the per-sample-averaged cytotoxicity data in GraphPad Prism 10.2.3 using one-way ANOVA with Tukey correction for multiple testing.

3 Results

3.1 T-cell clonality expansion

We performed a Dendritic cell-based antigen-specific T-cell induction protocol on PBMCs (n = 7) and observed an increase (mean fold change = 191.0, SD = 87.9) in the presence of MAGE-A3-specific T-cells after the protocol. The percentage of the MAGE-A3-specific T-cells isolated from healthy donors (n = 7) was 0.02% ± 0.015 (average ± SD) of the total lymphocytes, after applying the

protocol, this percentage increased to 3.33% ± 2.61 (average ± SD) (Figure 1). We have selected donors with the enrichment of the MAGE-A3-specific T-cells above 3% for single-cell multi-omics analysis that we performed on a BD Rhapsody platform using the "TCR/BCR Full Length, Targeted mRNA, and Sample Tag Library Preparation" Protocol.

3.2 TCR clonotype selection

We imported the BD Rhapsody-generated T-cell multi-omic data into TCRscape and identified 3000 T-cell receptor clonotypes, among which 191 clonotypes were present in 2 or more cells (Figure 2A). We were also able to perform post-sequencing quality control of the CD8+ T-cell sorting (99,9% of the T-cells were CD8+ T-cells) (Figure 2B). We used two main criteria for the Dominant clonotype selection: cell count per clonotype (main criterion) and predicted binding score towards the target peptide (secondary criterion). We identified a single dominant clonotype that was expressed by 14 cells (Figures 2A, E, F) and was predominantly represented by *CD8+ CD4−FOXP3−* T-cells (Figures 2B–D, G). We also observed that the dominant clonotype had a medium predicted binding score (i.e. affinity) towards the KVAELVHFL peptide, which is characteristic of the naturally occurring cytotoxic T-cells (27–29).

3.3 Plasmid construction

We then designed the insert for the lentiviral transfer plasmid. The insert included the TCRa and TCRb sequences in a single reading frame, separated by a signal to reset the polypeptide with the P2A polypeptide (See Supplementary Figure 2), the insert was then cloned into the pLenti hPGK GFP vector replacing the GFP gene (See Supplementary Figure 3).

3.4 Cytotoxicity assay

To evaluate the efficiency of the candidate TCR in targeting and eliminating MAGE-A3+ tumor cells, we engineered TCR T-cells via transduction of the anti-CD3-treated PBMCs with a lentivirus containing the aforementioned anti-MAGE-A3 construct (n=7), with the vehicle-transduced cells serving as a control. We then co-cultured the TCR T-cells and the non-transduced cells with the MAGE-A3-high SK-MEL-5, the MAGE-A3-low HCT-116, and the MAGE-A3-negative MDA-MB-231 cell lines. The LDH cytotoxicity assay results revealed high cytotoxicity against the SK-MEL-5 cell line (Figure 3, Table 1), indicating that the candidate TCR is effectively recognizing and targeting cells that express the MAGE-A3 antigen. Similar trends were observed in response to HCT-116 cells, which have

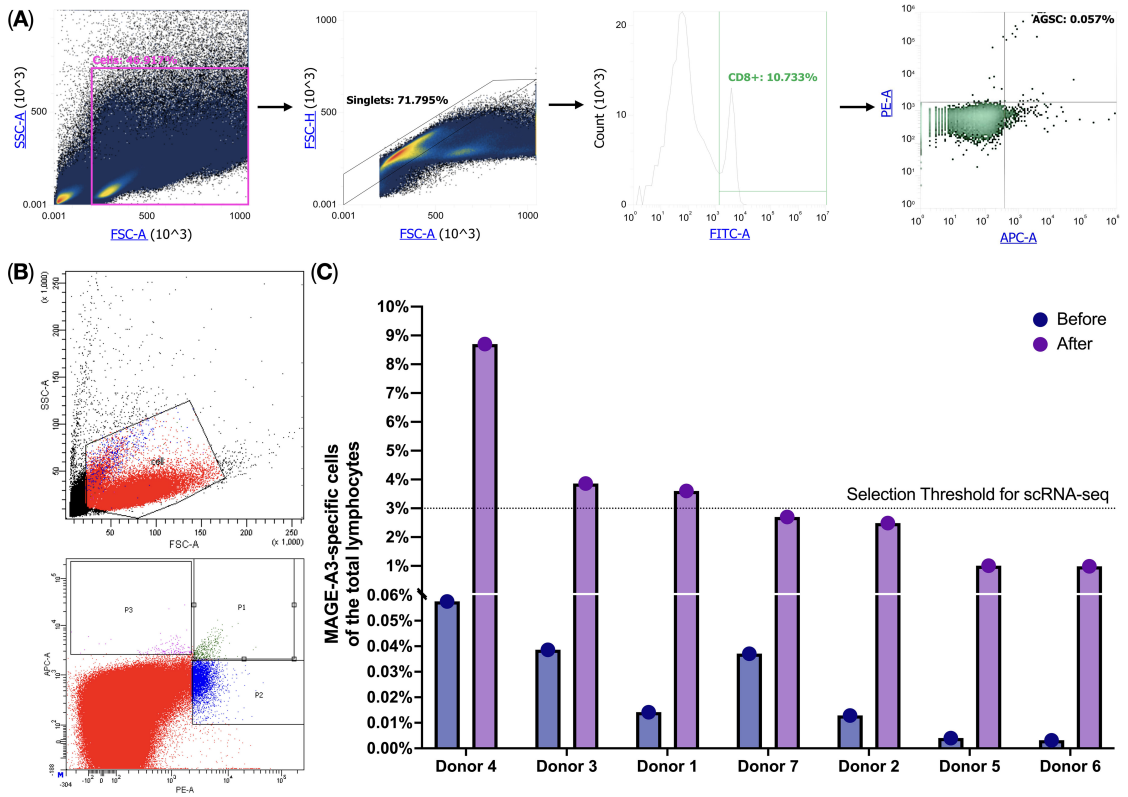


FIGURE 1
MAGE-A3-specific T-cell flow cytometry analysis. **(A)** HLA-A02-positive donor screening for MAGE-A3-specific T-cells; **(B)** MAGE-A3-specific T-cell enrichment after the cultivation protocol; **(C)** MAGE-A3-specific T-cell before and after the cultivation protocol.

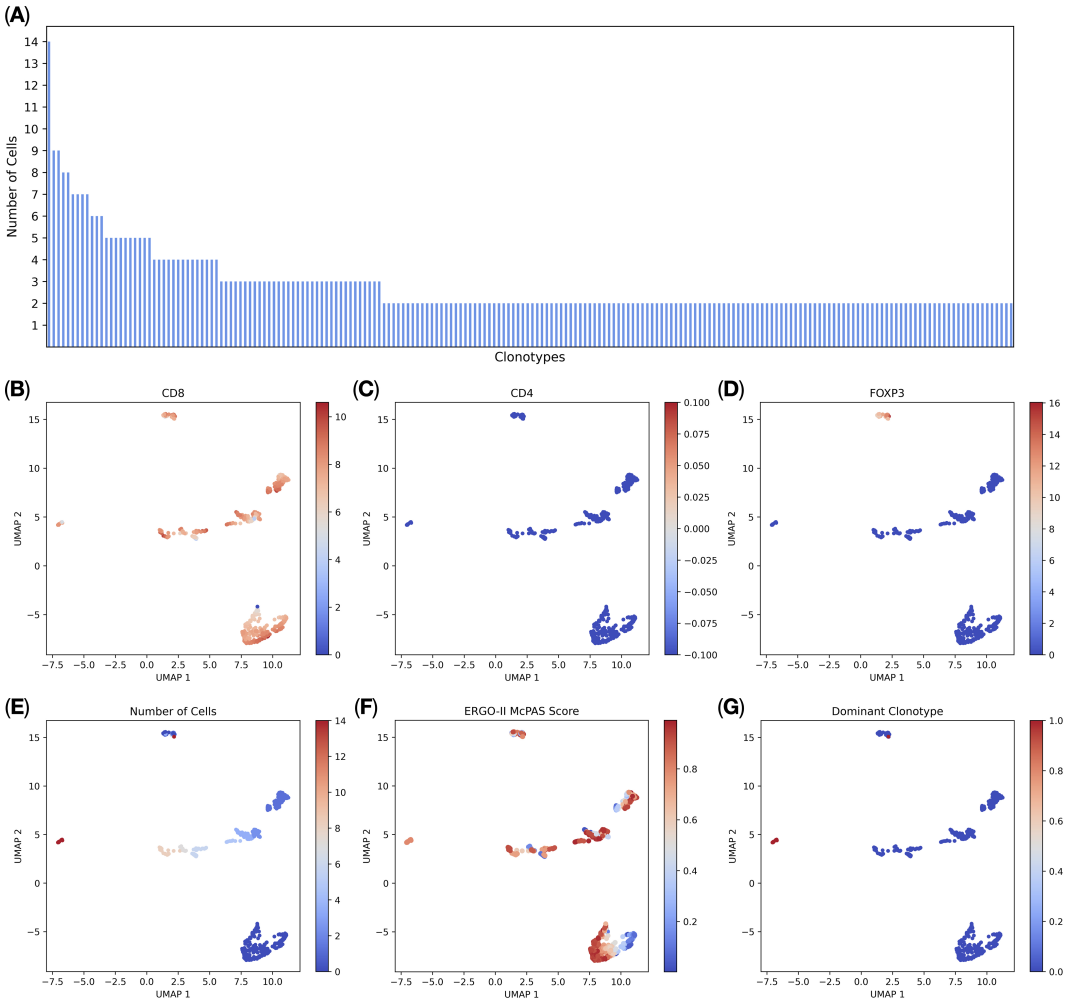


FIGURE 2
T-cell multi-omic analysis (*n* = 3, 5.491 single cells). **(A)** Distribution of complete T-cell receptor clonotype sequences. Each bar on the x-axis represents a clonotype, the y-axis shows the number of cells per clonotype, all clonotypes with 2 or more cells per clonotype are shown; **(B)** normalized *CD8* gene expression; **(C)** normalized *CD4* gene expression; **(D)** normalized *FOXP3* gene expression; **(E)** the number of cells per clonotype; **(F)** ERGO-II-predicted binding scores to the target peptide; **(G)** the Dominant clonotype (with 14 cells per clonotype) is shown in red.

lower levels of MAGE-A3 antigen expression, which aligns with the findings reported by Xiang Zhao et al. (30).

4 Discussion

In this paper, we performed a complex pipeline for the induction of MAGE-A3 antigen-specific T-cells: starting with the T-cell enrichment using peptide-loaded DCs, followed by single-cell RNA sequencing of the enriched T-cells on the BD Rhapsody platform, T-cell clonotype analysis was then performed via TCRscape, a clonotype discovery tool tailored to the BD Rhapsody data, which resulted in the identification of the dominant clonotype followed by the assessment of its effectiveness *in vitro* via an LDH cytotoxicity test.

We observed a successful T-cell clonality expansion after our protocol (mean fold change = 191.0, SD = 87.9), with sufficient T-cell numbers that allowed us to successfully perform scRNA-seq of such cells. This validates the described protocol as a suitable approach for the discovery of antigen-specific T-cells and their TCRs using modern single-cell analysis methods.

By utilizing single-cell sequencing technology, we obtained detailed information about the sequence of each TCR, enabling us to accurately construct the obtained TCR clones, as well as, information about the immune transcriptome of each T-cell, allowing us to assess the functionality state of the cells and enhance the selection process of candidate TCR clones. This provides a significant advantage over currently used methods like bulk RNA-seq, which do not provide any cell-of-origin information for sequenced TCRs making it impossible to determine which alpha and beta chains originated from the same T-cell, or multiplex PCR and 5'-RACE approaches, which exhibit lower accuracy and sensitivity (31).

Our analysis of the BD Rhapsody-generated single-cell multi-omic data in TCRscape revealed 191 unique clonotypes that were detected in 2 or more T-cells, thus also confirming the successful enrichment of the T-cells using our T-cell enrichment protocol.

10.3389/fimmu.2024.1470130

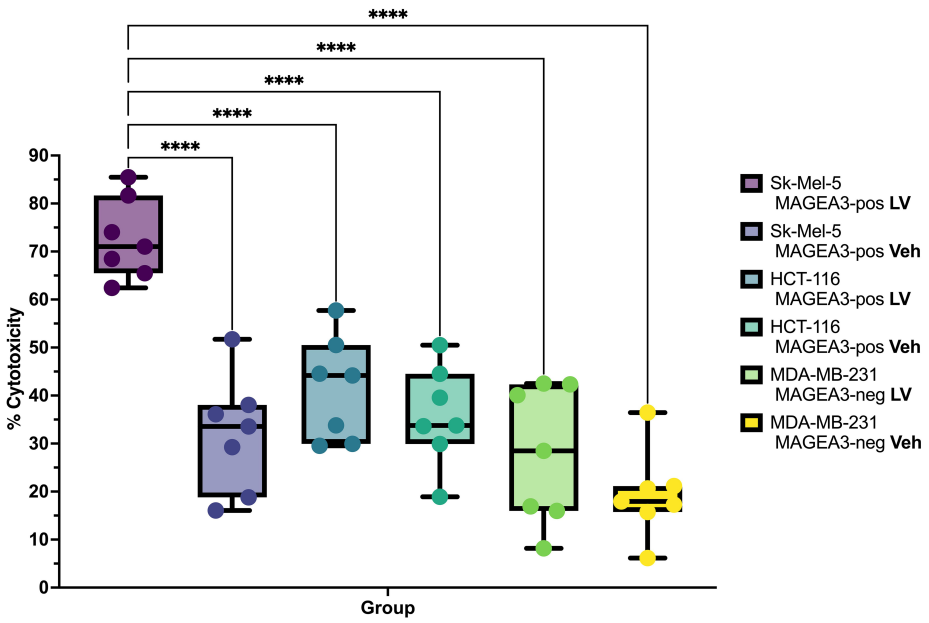


FIGURE 3
Cytotoxicity percentage and specificity of the candidate TCR against different tumor cell lines (n = 7), LV – lentiviral transduction (TCR-encoding), Veh– lentiviral transduction (Vehicle). **** - q-values < 0.00005.

Cytotoxicity results also validated our approach for clonotype selection, as TCR T-cells transduced with the dominant clonotype-lentivirus show potent cytotoxicity. This shows that the observed number of cells per clonotype could be a valid criterion for the *in-silico* search for the effective clonotype.

However, the small sample size constitutes a limitation of this study, as the data might not fully represent the overall population. Nevertheless, the findings remain highly relevant for TCR T-cell therapy, given that we identified a strong dominant clonotype with a 14-fold expansion compared with the baseline TCR occurrence, highlighting the potential efficacy of our approach even within a small cohort. Such results underscore the advantage of single-cell multi-omics technologies, which enable significant data acquisition and meaningful results even with limited sample sizes.

Previous efforts for the enrichment and isolation of antigen-specific T-cells have been extensively explored across various fields,

TABLE 1 Cytotoxicity percentage and specificity of the candidate TCR against different tumor cell lines (n = 7), LV, lentiviral transduction (TCR-encoding); Veh, lentiviral transduction (Vehicle).

	Donor 1	Donor 2	Donor 3	Donor 4	Donor 5	Donor 6	Donor 7	Mean	SD
Sk-Mel-5 MAGEA3-pos LV	68,46	71,02	65,48	62,43	74,02	85,48	81,67	72,65	8,41
Sk-Mel-5 MAGEA3-pos Veh	51,72	38,04	18,80	36,11	16,05	29,26	33,56	31,94	12,12
HCT-116 MAGEA3-pos LV	33,75	44,16	44,52	29,52	29,92	57,69	50,50	41,44	10,78
HCT-116 MAGEA3-pos Veh	33,75	44,52	29,92	50,50	18,93	33,62	39,55	35,83	10,28
MDA-MB-231 MAGEA3-neg LV	42,33	42,50	40,06	15,97	8,18	16,93	28,48	27,78	14,27
MDA-MB-231 MAGEA3-neg Veh	36,45	20,65	6,16	15,75	21,13	17,29	17,96	19,34	9,04

including infectious diseases, autoimmunity, and cancer. For example, Klinger et al. (32) pioneered a multiplex approach combining immune assays with receptor sequencing to identify antigen-specific TCRs, laying the groundwork for subsequent research. Similarly, Sharma et al. (33, 34) demonstrated the utility of TCR repertoire analysis monitoring transplant patients and autoimmune diseases, underscoring the broader applicability of such pipelines beyond oncology. Furthermore, Dziubianau et al. (35) provided further evidence of the importance of precise TCR selection processes through their work on enriching antigen-specific T-cells in viral infections. Collectively, these methodologies underscore the critical need to identify effective TCR candidates, a goal that aligns directly with the objectives of our current study, especially when integrated with single-cell sequencing.

In conclusion, we have modified a dendritic cell-bases protocol to be efficient for T-cell clonal expansion, obtained a potent TCR via single-cell sequencing, and successfully tested it in an *in vitro* cytotoxicity test against a MAGE-A3-positive tumor. Nevertheless, further investigation is required to determine the applicability of our results *in vivo*.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found via the Zenodo repository: accession number 12804940.

Ethics statement

The studies involving humans were approved by Local ethics committee of the Research Institute of Fundamental and Clinical Immunology. 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.

Author contributions

SS: Funding acquisition, Resources, Validation, Writing – review & editing, Conceptualization, Data curation, Methodology, Project administration, Supervision, Writing – original draft. MV: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Visualization, Writing – original draft. SA: Methodology, Software, Visualization, Writing – original draft, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation. RP-Z: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Visualization, Writing – original draft. OP-Z: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Visualization, Writing – original draft, Validation. MF: Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Writing – original draft. JL: Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Writing – original draft. JS: Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Writing – original draft. KN: Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Writing – original draft. JP: Data curation, Funding acquisition, Investigation, Methodology, Project administration, Validation, Writing – original draft. AA: Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Writing – original draft. VK: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – review & editing, Data curation. AS: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was carried out with the support of the Russian Science Foundation, project number 21-65-00004 (https://rscf.ru/project/21-65-00004).

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1470130/full#supplementary-material

References

1. Verma S, Swain D, Kushwaha PP, Brahmbhatt S, Gupta K, Sundi D, et al. Melanoma antigen family A (MAGE A) as promising biomarkers and therapeutic targets in bladder cancer. *Cancers (Basel)*. (2024) 16:246. doi: 10.3390/cancers16020246

2. Scanlan MJ, Gure AO, Jungbluth AA, Old LJ, Chen YT. Cancer/testis antigens: An expanding family of targets for cancer immunotherapy. *Immunol Rev*. (2002) 188:22–32. doi: 10.1034/J.1600-065X.2002.18803.X

3. Doyle JM, Gao J, Wang J, Yang M, Potts PR. MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases. *Mol Cell*. (2010) 39:963–74. doi: 10.1016/j.molcel.2010.08.029

4. Lee AK, Potts PR. A comprehensive guide to the MAGE family of ubiquitin ligases. *J Mol Biol*. (2017) 429:1114–42. doi: 10.1016/j.jmb.2017.03.005

5. Poojary M, Jishnu PV, Kabekkodu SP. Prognostic value of melanoma-associated antigen-A (MAGE-A) gene expression in various human cancers: A systematic review and meta-analysis of 7428 patients and 44 studies. *Mol Diagn Ther*. (2020) 24:537–55. doi: 10.1007/s40291-020-00476-5

6. Weon JL, Potts PR. The MAGE protein family And cancer. *Curr Opin Cell Biol*. (2015) 37:1–8. doi: 10.1016/j.ceb.2015.08.002

7. Cheever MA, Allison JP, Ferris AS, Finn OJ, Hastings BM, Hecht TT, et al. The prioritization of cancer antigens: A national cancer institute pilot project for the acceleration of translational research. *Clin Cancer Res*. (2009) 15:5323–37. doi: 10.1158/1078-0432.CCR-09-0737

8. Morgan RA, Chinnasamy N, Abate-Daga D, Gros A, Robbins PF, Zheng Z, et al. Cancer regression and neurological toxicity following anti-MAGE-A3 TCR gene therapy. *J Immunother*. (2013) 36:133–51. doi: 10.1097/CJI.0b013e3182829903

9. Home. ClinicalTrials.gov. Available at: https://clinicaltrials.gov/ (Accessed July 19, 2024).

10. Linette GP, Stadtmauer EA, Maus MV, Rapoport AP, Levine BL, Emery L, et al. Cardiovascular toxicity and titin cross-reactivity of affinity-enhanced T cells in myeloma and melanoma. *Blood*. (2013) 122:863–71. doi: 10.1182/blood-2013-03-490565

11. Arias CF, Herrero MA, Cuesta JA, Acosta FJ, Fernández-Arias C. The growth threshold conjecture: a theoretical framework for understanding T-cell tolerance. *R Soc Open Sci*. (2015) 2:150016. doi: 10.1098/rsos.150016

12. Kaech SM, Cui W. Transcriptional control of effector and memory CD8+ T cell differentiation. *Nat Rev Immunol*. (2012) 12:749–61. doi: 10.1038/nri3307

13. Rosati E, Dowds CM, Liaskou E, Henriksen EKK, Karlsen TH, Franke A. Overview of methodologies for T-cell receptor repertoire analysis. *BMC Biotechnol*. (2017) 17:61. doi: 10.1186/s12896-017-0379-9

14. Perik-Zavodskii R, Perik-Zavodskaia O, Alrhmoun S, Volynets M, Shevchenko J, Nazarov K, et al. Single-cell multi-omics reveal stage of differentiation and trajectory-dependent immunity-related gene expression patterns in human erythroid cells. *Front Immunol*. (2024) 15:1431303/BIBTEX. doi: 10.3389/FIMMU.2024.1431303/BIBTEX

15. Pai JA, Satpathy AT. High-throughput and single-cell T cell receptor sequencing technologies. *Nat Methods*. (2021) 18:881–92. doi: 10.1038/s41592-021-01201-8

16. Ulbrich J, Lopez-Salmeron V, Gerrard I. BD Rhapsody™ single-cell analysis system workflow: from sample to multimodal single-cell sequencing data. *Methods Mol Biol*. (2023) 2584:29–56. doi: 10.1007/978-1-0716-2756-3_2.

17. Jacobsen K, Inekci D, Brix L. 30 Single cell multiomic profiling of the antigen-specific immune response using antigen specific dCODE Dextramer® (RiO) reagents and BD® AbSeq Reagents on the BD Rhapsody™ single-cell analysis system. *J Immunother Cancer*. (2022) 10:A32–2. doi: 10.1136/jitc-2022-SITC2022.0030

18. Kurilin V, Alshevskaya A, Sennikov S. Development of cell technologies based on dendritic cells for immunotherapy of oncological diseases. *Biomedicines*. (2024) 12:699. doi: 10.3390/biomedicines12030699

19. Kuznetsova M, Lopatnikova J, Khantakova J, Maksyutov R, Maksyutov A, Sennikov S. Generation of populations of antigen-specific cytotoxic T cells using DCs transfected with DNA construct encoding HER2/neu tumor antigen epitopes. *BMC Immunol*. (2017) 18:31. doi: 10.1186/s12865-017-0219-7

20. Shevchenko J, Khristin A, Kurilin V, Kuznetsova M, Blinova D, Starostina N, et al. Autologous dendritic cells and activated cytotoxic T–cells as combination therapy for breast cancer. *Oncol Rep*. (2019) 43:671–80. doi: 10.3892/or.2019.7435

21. Obleukhova I, Kiryishina N, Falaleeva S, Lopatnikova J, Kurilin V, Kozlov V, et al. Use of antigen-primed dendritic cells for inducing antitumor immune responses in *vitro* in patients with non-small cell lung cancer. *Oncol Lett*. (2018) 15:1297–306. doi: 10.3892/OL.2017.7403

22. Arrieta-Bolaños E, Hernández-Zaragoza DI, Barquera R. An HLA map of the world: A comparison of HLA frequencies in 200 worldwide populations reveals diverse patterns for class I and class II. *Front Genet*. (2023) 14:866407. doi: 10.3389/FGENE.2023.866407

23. GitHub - IdoSpringer/ERGO-II: ERGO-II, an updated version of ERGO including more features for TCR-peptide binding prediction. Available online at: https://github.com/IdoSpringer/ERGO-II (Accessed July 19, 2024).

24. Tickotsky N, Sagiv T, Prilusky J, Shifrut E, Friedman N. McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences. *Bioinformatics*. (2017) 33:2924–9. doi: 10.1093/BIOINFORMATICS/BTX286

25. GitHub - Perik-Zavodskii/TCRscape: TCRscape - a tool for simultaneous multimodal gene expression and clonotype analysis of single T-cells profiled via the BD Rhapsody system . Available online at: https://github.com/Perik-Zavodskii/TCRscape (Accessed July 19, 2024).

26. Cornetta K, Anderson WF. Protamine sulfate as an effective alternative to polybrene in retroviral-mediated gene-transfer: implications for human gene therapy. *J Virol Methods*. (1989) 23:187–94. doi: 10.1016/0166-0934(89)90132-8

27. Moran AE, Hogquist KA. T-cell receptor affinity in thymic development. *Immunology*. (2012) 135:261–7. doi: 10.1111/j.1365-2567.2011.03547.x

28. Chao DL, Davenport MP, Forrest S, Perelson AS. The effects of thymic selection on the range of T cell cross-reactivity. *Eur J Immunol*. (2005) 35:3452–9. doi: 10.1002/eji.200535098

29. Palmer E, Naeher D. Affinity threshold for thymic selection through a T-cell receptor–co-receptor zipper. *Nat Rev Immunol*. (2009) 9:207–13. doi: 10.1038/nri2469

30. Zhao X, Kolawole EM, Chan W, Feng Y, Yang X, Gee MH, et al. Tuning T cell receptor sensitivity through catch bond engineering. *Science (80-)*. (2022) 376. doi: 10.1126/science.abl5282

31. Genolet R, Bobisse S, Chiffelle J, Speiser DE, Coukos G, Harari A. TCR sequencing and cloning methods for repertoire analysis and isolation of tumor-reactive TCRs. *Cell Rep Methods*. (2023) 3:100459. doi: 10.1016/j.crmeth.2023.100459

32. Klinger M, Pepin F, Wilkins J, Asbury T, Wittkop T, Zheng J, et al. Multiplex identification of antigen-specific T cell receptors using a combination of immune assays and immune receptor sequencing. *PloS One*. (2015) 10:e0141561. doi: 10.1371/journal.pone.0141561

33. Sharma RK, Yoosuf N, Afonso M, Scheffschick A, Avik A, Bartoletti A, et al. Identification of proteinase 3 autoreactive CD4+T cells and their T-cell receptor repertoires in antineutrophil cytoplasmic antibody–associated vasculitis. *Kidney Int*. (2023) 103:973–85. doi: 10.1016/j.kint.2023.01.023

34. Sharma RK, Boddul SV, Yoosuf N, Turcinov S, Dubnovitsky A, Kozhukh G, et al. Biased TCR gene usage in citrullinated Tenascin C specific T-cells in rheumatoid arthritis. *Sci Rep*. (2021) 11:24512. doi: 10.1038/s41598-021-04291-8

35. Dziubianau M, Hecht J, Kuchenbecker L, Sattler A, Stervbo U, Rödelsperger C, et al. TCR repertoire analysis by next generation sequencing allows complex differential diagnosis of T cell–related pathology. *Am J Transplant*. (2013) 13:2842–54. doi: 10.1111/ajt.12431vv

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 28 October 2024
DOI 10.3389/fimmu.2024.1471198

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Amita Joshi,
Biobay Solutions, India
Fangdie Ye,
Fudan University, China

*CORRESPONDENCE
Yongxin-Zhang
yoyozhang202204@163.com
Zhenjie-Chen
20180601@kmmu.edu.cn

†These authors share first authorship

RECEIVED 26 July 2024
ACCEPTED 08 October 2024
PUBLISHED 28 October 2024

CITATION
Lu Y, Wu J, Li X, Leng Q, Tan J, Huang H, Zhong R, Chen Z and Zhang Y (2024) Cuproptosis-related lncRNAs emerge as a novel signature for predicting prognosis in prostate carcinoma and functional experimental validation.
Front. Immunol. 15:1471198.
doi: 10.3389/fimmu.2024.1471198

Cuproptosis-related lncRNAs emerge as a novel signature for predicting prognosis in prostate carcinoma and functional experimental validation

Yangbai-Lu[1†], Jinfeng-Wu[2†], Xianzhe Li[3,4†], Qu-Leng[1], Jian-Tan[2], Hongxing-Huang[1], Rui-Zhong[1], Zhenjie-Chen[1*] and Yongxin-Zhang[5*]

[1]Department of Urology, Zhongshan City People's Hospital, Zhongshan, Guangdong, China, [2]Department of First Clinical Medical College, Guangdong Medical University, Zhanjiang, Guangdong, China, [3]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany, [4]Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany, [5]Department of MR, Zhongshan City People's Hospital, Zhongshan, Guangdong, China

Background: Prostate cancer (PCa) is one of the most common malignancies of the urinary system. Cuproptosis, a newly discovered form of cell death. The relationship between cuproptosis-related long non-coding RNAs (ClncRNAs) related to PCa and prognosis remains unclear. This study aimed to explore the clinical significance of novel ClncRNAs in the prognostic assessment of PCa.

Methods: ClncRNAs and differentially expressed mRNAs linked to these ClncRNAs were identified using Pearson's correlation and differential expression analyses. A prognostic signature (risk score) comprising three ClncRNAs was established based on multivariable Cox regression analysis. The predictive performance of this ClncRNAs signature was validated using receiver operating characteristic curves and nomograms. Finally, further *in vitro* cell experiments were conducted for validation, including quantitative polymerase chain reaction (qPCR), western blot (WB), cell proliferation assays, cell migration assays, cell invasion assays, apoptosis, and cell cycle analysis.

Results: We constructed a prognostic signature of ClncRNAs for PCa comprising three key differentially expressed ClncRNAs(AC010896-1, AC016394-2, and SNHG9). Multivariable Cox regression analysis indicated that clinical staging and risk scores of the ClncRNAs signature were independent prognostic factors for PCa. Compared to other clinical features, the ClncRNAs signature exhibited higher diagnostic efficiency and performed well in predicting the 1-, 3-, and 5-year progression-free intervals (PFIs) for PCa. Notably, in terms of immune activity, PCa patients with high-risk scores exhibited higher tumor mutational burden (TMB) levels, while their Tumor Immune Dysfunction and Exclusion (TIDE) scores were lower than those of PCa patients with low-risk scores. Additionally, *in vitro* cellular functional experiments, we knocked down SNHG9 that is the most significantly differentially expressed ClncRNA among the three key ClncRNAs. SNHG9 knockdown resulted in a significant increase in G1 phase cells and a decrease in S and G2 phases, indicating inhibition of DNA synthesis

and cell cycle progression. Colony formation assays showed reduced clonogenic ability, with fewer and smaller colonies. Western blot analysis revealed the upregulation of the key cuproptosis-related mRNAs FDX1 and DLST. These findings suggested that SNHG9 promotes PCa cell proliferation, migration, and invasion.

Conclusion: Building on the three ClncRNAs, we identified a novel prognostic signature of PCa. The ClncRNA SNHG9 can promote PCa cell proliferation, migration, and invasion.

KEYWORDS

cuproptosis, lncRNAs, prognosis signature, prostate carcinoma, SNHG9

1 Introduction

Estimates suggest that globally, there are close to 1.5 million new cases of prostate cancer (PCa), resulting in approximately 397,000 deaths. In 2022, it was the second most prevalent cancer among men and the fifth leading cause of cancer-related deaths among men (1). Although surgery, hormonal therapy, and radiation therapy show significant therapeutic efficacy in most PCa cases (2), the heterogeneity of PCa often renders these treatments ineffective. Therefore, novel prognostic methods and treatment strategies for PCa are urgently needed.

Biological features, primarily presented through gene expression and genomic profiling, can enhance the predictive capacity of traditional clinicopathological features (3). Understanding the genes associated with PCa may improve treatment selection and precision. Cuproptosis is a novel cell death mechanism that is distinct from other known mechanisms such as apoptosis, autophagy, and ferroptosis. Copper ions can induce cell death even when known cell death pathways are blocked. During mitochondrial respiration, copper ions directly bind to lipidated components of the tricarboxylic acid cycle, leading to protein aggregation. Additionally, copper ions can decrease the protein levels of Fe-S clusters, triggering proteinotoxic stress responses, and ultimately leading to cell death (4). By leveraging this novel approach to cell death, we are developing new strategies to advance the development of novel therapeutic options for patients with late-stage PCa.

Within the human genome, only a minority of genes encode proteins, with the majority encoding non-coding RNAs. Long non-coding RNAs (lncRNAs) are single-stranded RNAs that are more than 200 nucleotides long and lack protein-coding capabilities (5). lncRNAs regulate various physiological and biochemical cellular processes through chromatin modification, transcriptional activation, and interference (6, 7). lncRNAs can serve as non-invasive tumor markers for malignancies of the genitourinary system. Further elucidation of the molecular mechanisms by which lncRNAs function in normal and malignant cells will enhance our understanding of tumor biology and provide new therapeutic targets for genitourinary cancers (8). Recent studies have increasingly confirmed the crucial role of cuproptosis-related lncRNAs (ClncRNAs) in the prognosis and immunity of various genitourinary tumors. Zhang et al. (9) identified the ClncRNAs LINC01711 as a potential biomarker for early diagnosis and prognosis of clear cell renal carcinoma. Shen et al. (10) developed a risk signature comprising eight ClncRNAs with potential clinical applications in predicting outcomes and diagnosing and treating bladder cancer. However, research on the relationship between ClncRNAs and the prognosis of patients with PCa remains limited. Therefore, this study aimed to construct a novel prognostic signature related to ClncRNAs to predict the prognosis of PCa.

2 Materials and methods

2.1 Data download and pre–processing

Cuproptosis-related genes expression was standardized using the "limma" R package. Subsequently, the Perl programming language was used to distinguish between mRNAs and lncRNAs. Additionally, simple nucleotide variation data and masked somatic mutation data were downloaded from The Cancer Genome Atlas (TCGA) database and employed to compute the mutation burden of PCa. Nineteen cuproptosis-related mRNAs were sourced from previously published studies including NFE2L2, NLRP3, ATP7A, ATP7B, SLC31A1, FDX1, LIAS, LIPT1, LIPT2, DLD, DLAT, PDHA1, PDHB, MTF1, GLS, CDKN2A, DBT, GCSH, and DLST (4).

2.2 Screening the differentially expressed ClncRNAs

Pearson correlation analysis was employed to explore the association between ClncRNAs and PCa-related lncRNAs, using a screening criterion of $|R^2| > 0.5$ and $p < 0.05$ to ensure the significance and robustness of the discovered correlations.

Subsequently, further analysis was conducted using the "limma" R package, adjusting for p <0.05 and |log2 fold change (FC)| >1 as criteria to screen for differentially expressed ClncRNAs. These ClncRNAs then were visualized using Sankey plots and volcano plots to illustrate the aforementioned results.

2.3 Establishment and evaluation of a novel prognostic signature associated with ClncRNAs

Six candidate ClncRNAs, including two protective and four risk lncRNAs, were identified using univariate Cox regression analysis. Subsequently, to prevent overfitting, LASSO regression analysis and lambda spectra were employed to explore collinearity. Next, three key ClncRNAs were identified by multivariable Cox regression analysis. Finally, a prognostic ClncRNAs signature based on these three ClncRNAs was constructed and 420 patients were divided into high- and low-risk groups using the median score as the cutoff value. In this study, the risk score calculation formula of ClncRNAs signature was as follows: risk score = $f(x) = \sum_{i=1}^{n}(\mathrm{Coef}_i * \mathrm{Exp}_i)$, where Coef_i represents the corresponding coefficient for each ClncRNAs and Exp_i represents the gene expression level of the selected ClncRNAs. Subsequently, the Kaplan–Meier (K–M) method was used to plot the progression-free interval (PFI) and disease-free interval (DFI) of patients with PCa using the "survival" R package. Based on the risk score and clinical features, time-dependent receiver operating characteristic (ROC) curves and C-index curves were plotted using the "timeROC" R package to evaluate the accuracy and stability of the signature. Additionally, based on the results of multivariable Cox regression analysis, a nomogram was developed using the "rms" package to predict the 1-, 3-, and 5-year PFI and evaluate the long-term predictive accuracy of the signature. Furthermore, the correlation between the ClncRNAs signature and some clinical features was assessed, and their predictive ability in subgroups with different clinical features was evaluated using KM survival analysis.

2.4 Principal components analysis and functional enrichments analysis

Following the identification of differentially expressed mRNAs between the high-risk and low-risk groups, we conducted gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses using the "clusterProfiler" and "enrichment" packages to elucidate molecular functions and key signaling pathways (11, 12). The components of graphene oxide comprised biological process components, cellular component components, and molecular function components.

2.5 Calculation of tumor mutation burden score

The tumor mutational burden (TMB) reflects the number of mutations in a tumor. Mutation data from PCa samples were downloaded from TCGA, and analysis was conducted using the R package "maftools" (13). The TMB is correlated with the clinical efficacy of immunotherapy. Additionally, we explored the differences in somatic mutation characteristics between the high-risk and low-risk groups.

2.6 A comparative analysis of immune cell infiltration and immune-related functions in different risk groups

To evaluate the differences in immune infiltration between high-risk and low-risk groups, we performed Gene Set Variation Analysis (GSVA) using the "reshape2" and "GSEABase" packages and generated heatmaps to illustrate the variations in immune functions across different risk groups (14). Additionally, we obtained Tumor Immune Dysfunction and Exclusion (TIDE) scores from their website (http://tide.dfci.harvard.edu) and analyzed the differences in TIDE scores between the low-risk and high-risk groups. Subsequently, we examined other immune scores, including T cell dysfunction, T cell exclusion, PD-L1, CD8, IFNG, Merck18, CAF, TAM M2, and MDSC scores, to assess the differences in immune infiltration between the high-risk and low-risk groups. These scores are sourced from the TIDE website.

2.7 Cell lines, cell culture, and handling

The PCa cell lines RWPE-1, PC-3, U145, 22RV1, LNCaP, C4-2PC3, and DU145 were purchased from the ATCC (Manassas, VA, USA). PC-3 and DU145 cells were transfected using Lipofectamine™ RNAiMAX (Invitrogen, Cat. No. 13778075) using LncRNA-SNHG9-targeting small interfering RNA (siRNA) designed and synthesized by Beijing TsingKe Biological Technology Co., Ltd. The sequences for siSNHG9-1, siSNHG9-2, and siSNHG9-3 were as follows: siSNHG9-1 sequence: ACCCGAAGAGUGGCUAUAATT, siSNHG9-2 sequence: CCUCUUCACUUAGGACACUTT, and siSNHG9-3 sequence: CCACGUCUUUCAAAUAAAGTT. PC-3 and DU145 cells were cultured in the RPMI-1640 (Gibco, Cat. No. C11875500BT) and MEM (Gibco, Cat. No. C11095500BT) media, respectively, supplemented with 10% fetal bovine serum (FBS) and 1% dual antibiotics (penicillin-streptomycin) (HyClone, Cat. No. SH30010). Cultures were maintained at 37°C in a humidified atmosphere containing 5% CO_2.

2.8 RNA Extraction, reverse transcription, and quantitative real-time polymerase chain reaction

Total RNA was extracted from RWPE-1, PC-3, U145, 22RV1, LNCaP, C4-2PC3, and DU145 cells using the TRIzol reagent. The concentration and quality of each RNA sample were evaluated by measuring the OD values using a BioPhotometer Plus (Eppendorf Nucleic Acid and Protein Analyzer). Subsequently, the extracted RNA was reverse transcribed into complementary DNA (cDNA)

using the EasyScript First-Strand cDNA Synthesis SuperMix Kit. The resulting cDNA was subjected to quantitative real-time polymerase chain reaction (qRT-PCR) using SYBR Green Master Mix (Vazyme). Analysis was performed on the ABI PRISM® 7500 Sequence Detection System, with GAPDH serving as the internal reference gene. Relative gene expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method. PCR primers for the target gene (SNHG9) and the internal reference gene (GAPDH) are provided in Supplementary Table 2.

2.9 Cell proliferation experiment

The proliferative capacity of PC-3 and DU145 cells transfected with siRNA was assessed using the Cell Counting Kit-8 (CCK-8) reagent according to the manufacturer's instructions. PC-3 and DU145 cells were seeded in 96-well plates, and CCK-8 solution (Beyotime, Cat. No. C0039) was added at 0, 24, 48, and 72 h. After incubation for 4 h, optical density (OD450) was measured using a microplate reader (Thermo Fisher Scientific, Multiscan MK3) to determine the absorbance values for CCK-8 detection.

2.10 Cell apoptosis assay

Apoptosis was examined using an Annexin V-FITC Apoptosis Detection Kit (Keygen, Cat. No. KGA106) using flow cytometry. All procedures were performed according to manufacturer's instructions. Briefly, cells pretreated with phosphate-buffered saline (PBS) were washed twice to obtain a pellet. Subsequently, the cells at different stages of apoptosis were identified using fluorescein isothiocyanate and propidium iodide (PI) solutions. The cells were then incubated in the dark at room temperature (18–24°C) for 15 min. Finally, the apoptosis rate was measured using a BD FACS Calibur flow cytometer (Beckman Coulter, CA, USA).

2.11 Flow cytometric cell cycle analysis

A KGA511 cell cycle assay kit was used to assess the cellular proliferation status and cycle distribution under various conditions. Following transfection for 48 h, 1×10^6 cells were collected per sample, centrifuged to remove the supernatant, and washed twice with pre-chilled PBS. Subsequently, cells were treated with pre-chilled 70% ethanol and fixed overnight at 4°C or stored long-term at -20°C. Fixed cells were then collected by centrifugation, washed once with 1 mL of PBS, and resuspended in PBS solution containing 50 μg/mL PI, 100 μg/mL RNase A, and 0.2% Triton X-100. Following 30 min of incubation at 4°C in the dark to stain the cellular DNA, an appropriate volume of the stained cell suspension was subjected to standard flow cytometry analysis. The results were analyzed using the ModFit cell cycle analysis software, which allocates cells to different cell cycle phases based on DNA content and generates the corresponding cell cycle distribution plots.

2.12 Transwell assay for assessing cell migration and invasion capacity

Transwell cell culture inserts (BD, REF353097) were used for the transwell migration assays. Approximately 1×10^5 cells were seeded in 100 μL of serum-free culture medium in the upper chamber, while 600 μL of complete culture medium was added to the lower chamber. Following incubation at 37°C with 5% CO_2 for 12–48 h, the inserts were retrieved. The cells on the upper surface of the insert were gently removed using a cotton swab and fixed in 4% paraformaldehyde for 20 min. After a single wash with PBS, the cells were stained with crystal violet for 10 min, washed again with PBS, and subjected to cell counting and image analysis using an inverted microscope (OLYMPUS CKX41, U-CTR30-2) and Image J 1.44 software. For the transwell invasion assay, Matrigel (BD, 356234) was dissolved overnight at 4°C and diluted to a 1:3 ratio with pre-chilled serum-free culture medium. Forty microliters of diluted Matrigel was added to the pre-chilled transwell inserts and allowed to gel at 37°C for 2 h. Excess liquid was removed from the inserts, and 100 μL of serum-free culture medium was added to the upper chamber, while 600 μL was added to the lower chamber. After overnight equilibration at 37°C, cells transfected with siRNA were seeded at a density of 1×10^5 cells in 100 μL of serum-free DMEM-F12 or MEM medium in the upper chamber. Complete culture medium was added to the lower chamber, and the cells were incubated at 37°C with 5% CO_2 for 24 or 48 h. Following incubation, the cells on the upper surface of the insert were removed using a cotton swab and the remaining cells were fixed with 4% paraformaldehyde for 15 min. After a single wash with PBS, the cells were stained with crystal violet for 10 min, washed again with PBS, and subjected to observation and statistical analysis.

2.13 Cloning formation experiment

Using pancreatic enzymes, cells from different groups were initially digested and resuspended in 1 mL culture medium to achieve an appropriate cell count range. The cells were subsequently dissociated into single-cell suspensions and diluted to a concentration of 1×10^4 cells/ml. The diluted cells were then seeded into a 96-well plate, with 300 μL of suspension added to each well. After seeding, culture medium was added to each well to a final volume of 2000 μL to ensure sufficient growth medium for the cells. The plate was gently rocked horizontally and vertically to evenly distribute cells within each well. Cell growth was observed daily. When individual cells formed clustered clones, the culture medium was aspirated from each well. The wells were washed twice with PBS or saline, and then 200 μL of crystal violet staining solution was added to each well to ensure coverage of the well bottom, followed by a 20-min incubation period. The 96-well plate was rinsed with tap water, gently washed, and air-dried. Finally, the number of colonies was calculated and the stained 96-well plates were analyzed and scanned using an enzyme-linked immunospot (ELISPOT) AID iSpot system.

2.14 Protein extraction and western blotting

After washing the cells with PBS, protein extraction was performed using a radioimmunoprecipitation assay buffer containing 1% PMSF. The samples were then centrifuged at 14,000 rpm for 5 min at 4°C in appropriate centrifuge tubes. The protein concentration was determined using a BCA Protein Assay Kit (Thermo Scientific, USA). Total protein samples were separated on 15% or 10% SDS-PAGE gels and transferred onto polyvinylidene fluoride (PVDF) membranes. Membranes were blocked with TBST containing 5% skim milk powder. Antibodies against FDX1 (1:1000, 12592-1-AP, Proteintech, China), DLAT (1:2000, 13426-1-AP, Proteintech, China), and GAPDH (1:1000, KC-5G5, Shanghai Kangcheng Biotech, China) were incubated at 4°C. After overnight washing with TBST, the membranes were incubated with secondary antibodies at 37°C for 50 min to 3 h. The PVDF membranes were developed using ECL solution, followed by three washes and imaging.

2.15 Statistical analysis

Statistical analyses for visualization were performed using R software (version 4.2.0, https://www.r-project.org/). The Wilcoxon test was used to assess the differential expression of ClncRNAs and mRNAs, with significance set at $p < 0.05$.

3 Results

3.1 Clinical data of patients and identification of ClncRNAs

The flowchart in Figure 1 illustrates the process followed in this study. Gene expression profiles and clinical data of 52 normal and 501 PCa samples were obtained from the TCGA database. Detailed clinical characteristics of all the participants are presented in Table 1. To identify ClncRNAs, Pearson and differential expression analyses were performed, resulting in 47 differentially expressed ClncRNAs meeting the criteria of $|R^2| > 0.5$ and $p < 0.05$, and $|\log$ fold change (FC)$| > 1$, as shown in Supplementary Table 1. The Sankey and volcano plots in Figures 2A, B depict these findings. Subsequently, the TCGA cohort was divided into training and validation cohorts in a 1:1 ratio. The clinical characteristics of each cohort are listed in Table 1.

3.2 Construction and validation of a prognostic ClncRNAs signature for patients with PCa

Initially, univariate Cox regression analysis was used to preliminarily screen prognosis-related ClncRNAs, which identified six prognostic candidate ClncRNAs, including two protective lncRNAs (AC010896.1 and USP46-DT) and four high-risk lncRNAs (MHENCR, SNHG9, ACO11477.1, and AC016394.2), as shown in Figure 2C. Subsequently, LASSO regression analysis was conducted to refine and mitigate overfitting risks, as illustrated by the cvfit and lambda curves in Figures 2D, E. Multifactor Cox regression was employed to further screen for prognosis-related genes, resulting in three ClncRNAs (AC010896.1, AC016394.2 and SNHG9) with independent prognostic risk for PCa (Table 2). Additionally, using a heatmap, we demonstrated an association between these three ClncRNAs and 19 cuproptosis-related mRNAs (Figure 2F). Ultimately, based on these three ClncRNAs, a prognostic risk signature was constructed by dividing the 420 patients into high-risk and low-risk groups (1:1) using the median score as a cutoff point. The risk score for each patient was calculated with a multivariate Cox regression formula: expression level of AC010896.1*(-0.8161961) + expression level of AC016394.2*0.90231436 + expression level of SNHG9*0.44953283. Through survival status risk score curves and scatter plots for the total, training, and test cohorts, we observed a significant increase in mortality rates with increasing risk scores of the ClncRNAs signature, with the majority of deaths occurring in individuals identified as high risk (Figures 3A–I). Furthermore, the K–M survival curve analysis revealed a markedly poorer PFI in samples with high-risk scores in the total cohort (Figure 3J, $p < 0.001$), training cohort (Figure 3K, $p < 0.001$), and test cohort (Figure 3L, $p = 0.009$). However, when using DFI as the outcome measure, significant differences persisted between the total cohort (Figure 3M, $p = 0.002$) and the training cohort (Figure 3N, $p = 0.06$), while there was no statistical difference in the test cohort. Nonetheless, a trend of difference remained between the two groups in the test cohort (Figure 3O, $p = 0.154$).

3.3 Evaluation of the predictive accuracy of the ClncRNAs prognostic signature

Univariate and multivariate Cox regression analyses were conducted to validate whether the ClncRNAs signature risk score could independently predict PCa prognosis. Univariate analysis revealed significant differences in age, stage, and ClncRNAs risk score among the patients with PCa, whereas multivariate analysis excluded age as a significant factor (Figures 4A, B). The discriminative ability of the risk score was then evaluated against other clinical features using PFI as the outcome measure. The results showed that the ROC curve of ClncRNAs risk score had the highest AUC compared to the ROC curves of age and clinical stage, with AUCs of 0.730, 0.716, and 0.766 for the 1-, 3-, and 5-year ROCs, respectively (Figures 4C–E). Additionally, the 10-year C-index of the risk score was significantly higher than that of the other clinical features (Figure 4F). In summary, the ROC curve confirmed the significant prognostic predictive ability of the risk score of ClncRNAs signature compared with other clinical features. Furthermore, based on the results of the multivariable Cox regression analysis, a nomogram was constructed to accurately predict the 1-, 3-, and 5-year PFI (Figures 5A, B). To assess the

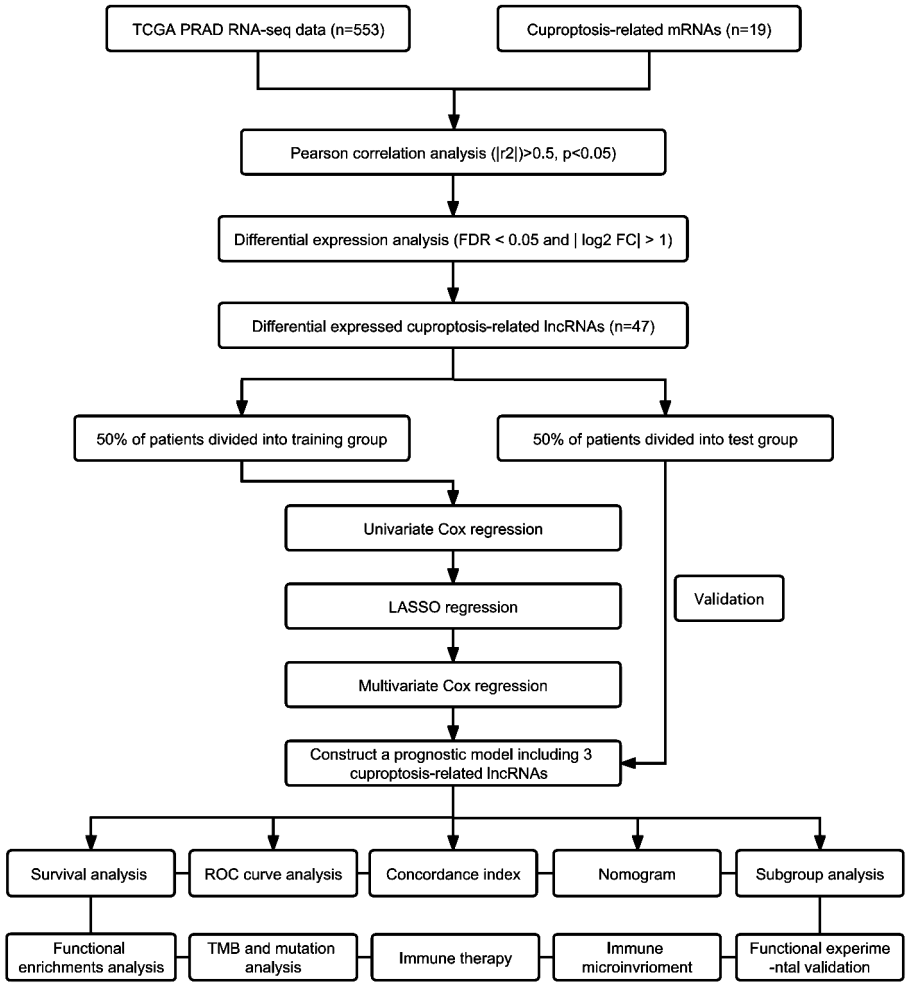


FIGURE 1
The schematic diagram of workflow.

TABLE 1 The clinical characteristics of patients in different cohorts.

Covariates	Total cohort	Training cohort	Test cohort	p value
Age				
<=65	296 (70.48%)	149 (70.95%)	147 (70%)	0.9148
>65	124 (29.52%)	61 (29.05%)	63 (30%)	
pN stage				
N0	341 (81.19%)	165 (78.57%)	176 (83.81%)	0.2118
N1	79 (18.81%)	45 (21.43%)	34 (16.19%)	
pT stage				
T2	143 (34.05%)	66 (31.43%)	77 (36.67%)	0.2685
T3	267 (63.57%)	137 (65.24%)	130 (61.90%)	
T4	10 (2.38%)	7 (3.33%)	3 (1.43%)	

accuracy of the prognostic prediction of the risk score in different stratified cohorts, a K–M survival analysis was conducted for each subgroup based on the clinicopathological characteristics including aged ≤65 years, >65 years, pT2, pT3-4, pN0, and pN1 (Figures 5C–H).

3.4 Principal components analysis and functional enrichments analysis

To elucidate the differences between the high-risk and low-risk groups, we conducted GO (Figures 6A, B) and KEGG (Figures 6C, D) enrichment analyses on the differentially expressed mRNAs between the two groups. GO analysis revealed potential associations of signaling receptor activation factor activity, contraction fibers, and muscle system processes with PCa. KEGG analysis indicated that enrichment was primarily in the adrenergic signaling pathway in cardiomyocytes, the neuroactive ligand-receptor interaction pathway, and the calcium signaling pathway.

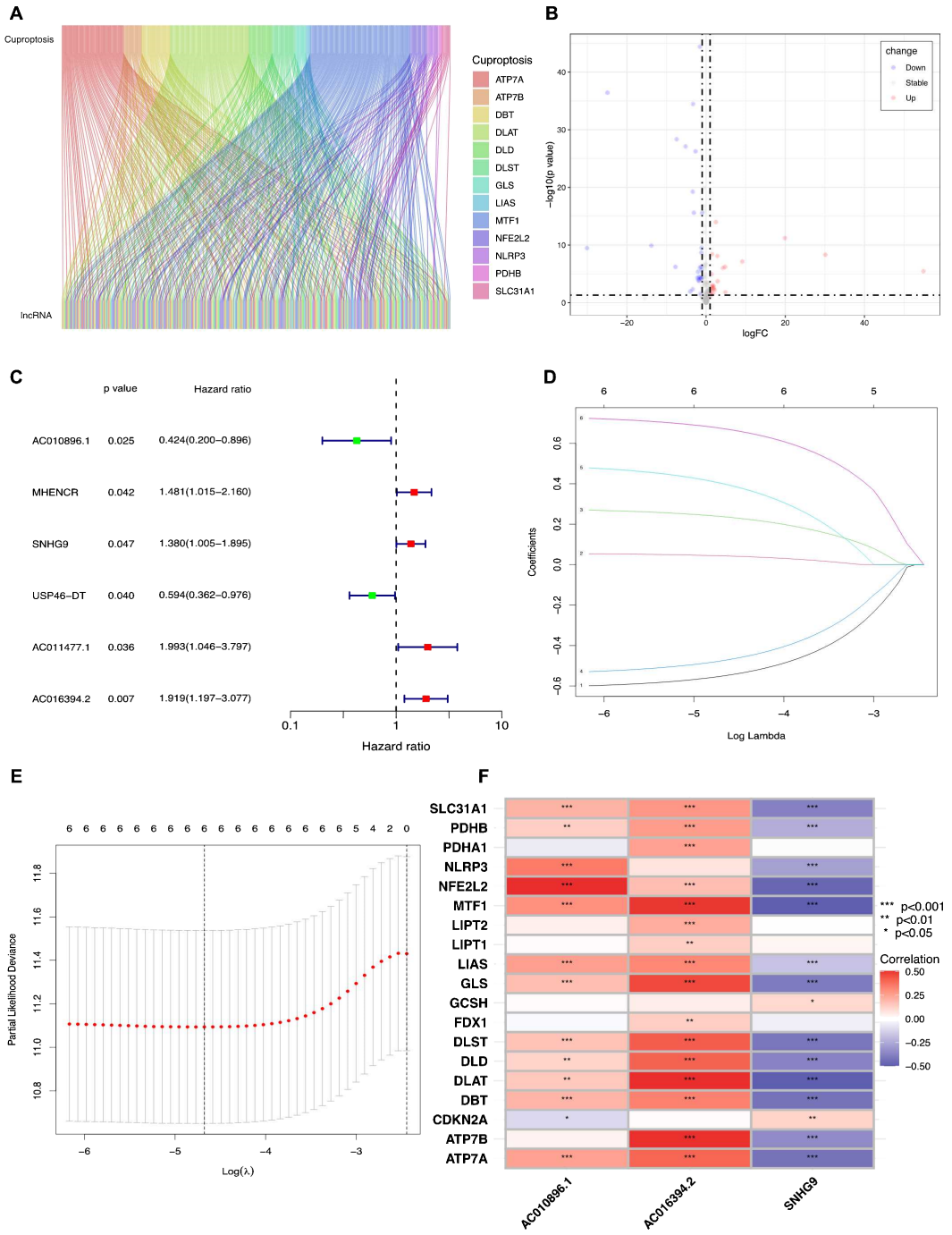


FIGURE 2
Identifying the ClncRNAs with prognostic significance in PCa. **(A)**The Sankey plot for the network of cuproptosis-related mRNAs and ClncRNAs. **(B)** The volcano diagram showed differentially expressed 47 ClncRNAs in PCa. **(C)** The forest plot displays six differentially expressed ClncRNAs with significant diagnostic value. **(D)** Lasso-Cox regression analysis is used to build a prognostic signature of ClncRNAs for PCa through 10-fold cross-validation for variable selection. **(E)** The LASSO penalty lambda (λ) of the six ClncRNAs. **(F)** The correlations of three prognostic ClncRNAs and cuproptosis-related mRNAs. *p <0.05, **p <0.01, and ***p <0.001.

3.5 Tumor mutation burden analysis

TMB differences between the high-risk and low-risk groups were also assessed, with the TMB of the low-risk group significantly lower than that of the high-risk group (p <0.001, Figure 6E). This indicates good consistency between the risk score and TMB for prognostic prediction. The waterfall plot depicted the top 15 genes with the highest mutation frequencies, highlighting the differences between the two groups. (Figures 6F, G). Among the top five genes with the highest mutation rates, only SPOP had a higher mutation rate in the low-risk group than in the high-risk group. The other genes showed lower mutation rates in the low-risk group.

10.3389/fimmu.2024.1471198

TABLE 2 Prognostic ClncRNAs signature via multivariate Cox regression analysis.

lncRNAs	coef	HR	HR.95L	HR.95H	p value
AC010896.1	-0.8161961	0.44211019	0.20397572	0.95825825	0.03864042
AC016394.2	0.90231436	2.46530212	1.55148936	3.91734207	0.00013408
SNHG9	0.44953283	1.56757968	1.07461132	2.28669289	0.01962196

3.6 Comparison of immune functions and TIDE scores in different risk groups

Through the analysis of immune-related functions, the immune statuses of the low-risk and high-risk groups were investigated. The results indicate significant differences in immune cell expression between the groups. Notably, expression levels of various immune cells, including antigen-presenting cells (APCs), co-stimulatory chemokine receptors (CCRs), pro-inflammatory markers, major histocompatibility complex (MHC) class I, sub-inflammation, and type II interferon (IFN) response, were notably lower in the high-risk group than in the low-risk group. Conversely, immune function

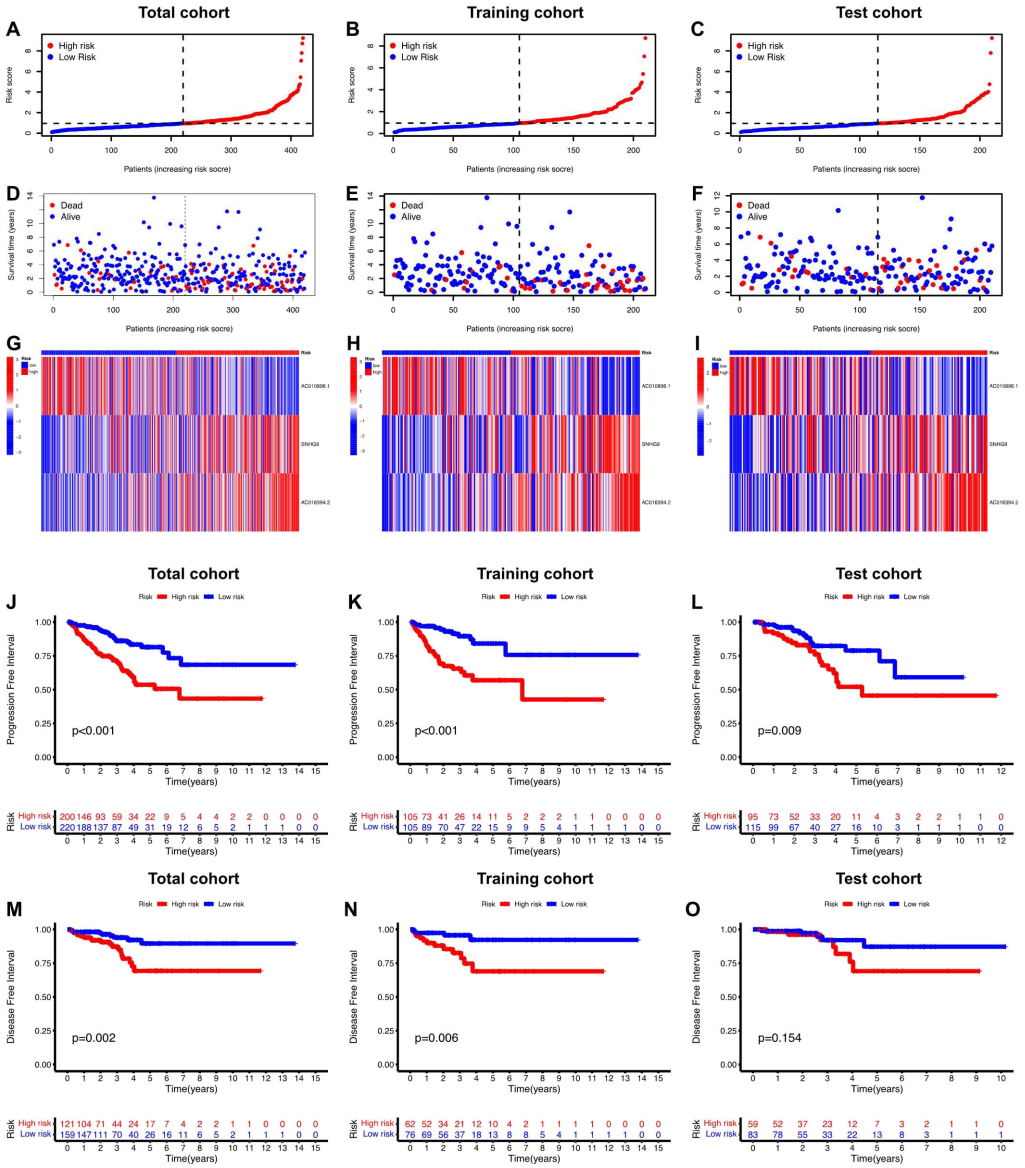


FIGURE 3
Prognosis value of the risk score of ClncRNAs signature in different cohorts. **(A-F)** The distribution of risk scores and statuses of patients with PCa in the total, training, and test cohorts, respectively. **(G-I)** Heatmaps of the three ClncRNAs expression in the total, training, and test cohorts, respectively. **(J-L)** Kaplan–Meier survival curves of PFI and **(M-O)** DFI in the high- and low-risk groups of PCa in the total, training, and test cohorts respectively.

10.3389/fimmu.2024.1471198

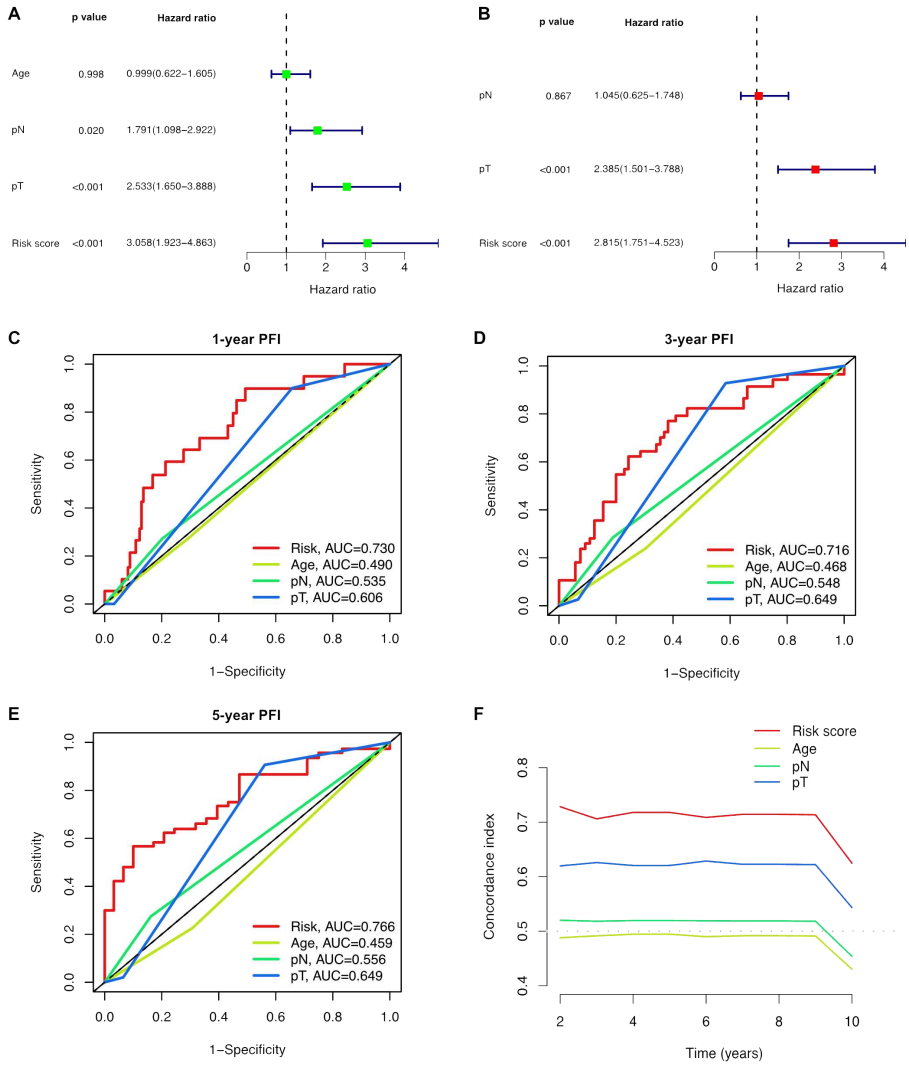


FIGURE 4
Independent prognostic analysis of the prognostic risk score of ClncRNAs signature in PCa. **(A, B)** Conducted both univariate and multivariate Cox regression analyses, taking into account the risk score and clinical characteristics. **(C-E)** TimeROC curve for 1-, 3-, and 5-year overall survival. **(F)** The C-index curve indicates that the predictive accuracy of the ClncRNAs signature risk score surpasses that of other clinical characteristics.

appeared to be more robust in the low-risk group (Figure 7A). TIDE scores obtained from the TIDE website further supported this observation, with the low-risk group showing higher scores than the high-risk group (Figure 7B). Additional investigations into scores for T-cell dysfunction, T-cell exhaustion, PD-L1, CD8, IFNG, Merck18, CAF, TAM M2, and MDSC also revealed significant differences between the high-risk and low-risk groups (Figures 7C–K).

3.7 Knockdown of SNHN9 affects cell viability, migration, and proliferation in PCa cells

Considering that among the ClncRNAs, SNHG9 exhibited the most significantly differential expression between tumor and normal samples (Supplementary Table 1), it was selected for further cellular experimental validation. The expression levels of SNHG9 were measured in human PCa cells, including RWPE-1, PC-3, DU145, 22RV1, LNCAP, and C4-2 cells, using qRT-PCR (Figure 8A). Subsequently, functional experiments targeting SNHG9 were conducted in highly expressed PC-3 and DU145 cells. Initially, siRNA targeting SNHG9 were designed and transfected into the cells and the transfection efficiency in DU145 cells was validated by qRT-PCR (Figure 8B). Cell viability was assessed using a CCK-8 assay. The results indicated a significant decrease in cell proliferation in PC3 and DU145 cells after SNHG9 knockdown compared to that in the control group (Figures 8C, D). Colony formation assays demonstrated a marked decrease in the cloning ability of PC3 and DU145 cells after SNHG9 knockdown, with significantly fewer and smaller colonies formed, further confirming the inhibitory effect of SNHG9 knockdown on cell proliferation and growth (Figures 8E–H). Transwell assays revealed a significant reduction in cell migration and invasion after SNHG9 knockdown (Figures 8I–P).

10.3389/fimmu.2024.1471198

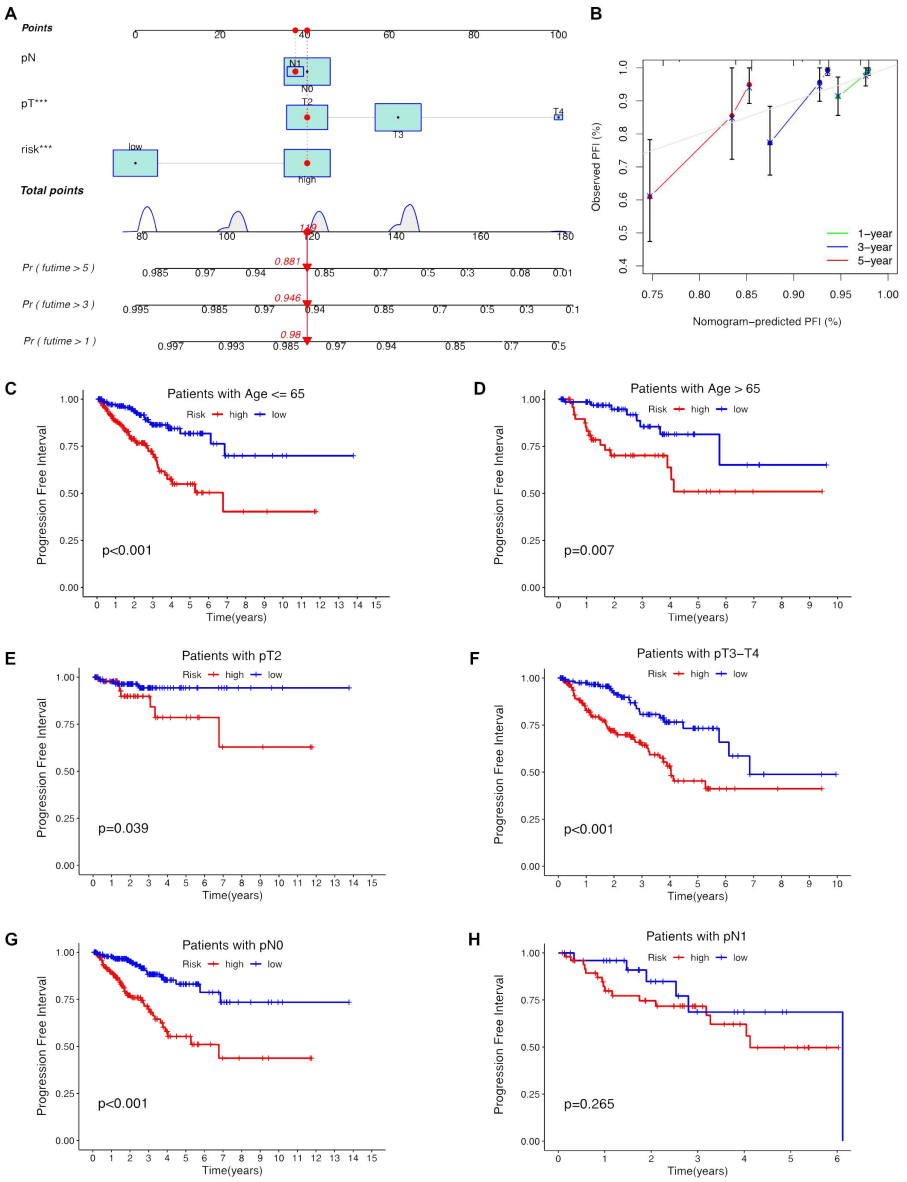


FIGURE 5
Construction the nomogram and subgroup analysis of the risk score of ClncRNAs signature. **(A, B)** The nomogram integrates the risk score and clinicopathological features to predict 1-, 3-, and 5-year DFS in PCa. **(C, D)** Kaplan–Meier curves for high-risk and low-risk groups across different age groups of patients with PCa. **(E, F)** Kaplan–Meier curves for high-risk and low-risk groups across different T stages of patients with PCa. **(G, H)** Kaplan–Meier curves for high-risk and low-risk groups across different N stages of patients with PCa.

3.8 Knockdown of SNHN9 affects apoptosis, cell cycle, and the expression of key cuproptosis-related mRNAs in PCa cells

Flow cytometry was used to assess apoptosis and cell cycle distribution, showing increased apoptosis in PC3 and DU145 cells after SNHG9 knockdown (Figures 9A–D). Additionally, cell cycle analysis showed a significant increase in the proportion of PC3 and DU145 cells in the G1 phase and a decrease in the proportion of cells in the S and G2 phases after SNHG9 knockdown, suggesting inhibition of DNA synthesis and cell cycle progression (Figures 9E–H). Moreover, western blot analysis revealed upregulation of the key cuproptosis-related mRNAs (FDX1 and DLST) upon SNHG9 knockdown (Figure 9I). Collectively, these findings suggest that SNHG9 activates the proliferation, migration, and invasion of PCa cells.

Discussion

In the study, gene expression profiles and clinical data of patients with PCa were obtained from TCGA database. Nineteen cuproptosis-related mRNAs were identified from published literature. Using Pearson's correlation, differential expression, LASSO regression, and Cox regression analyses, three lncRNAs

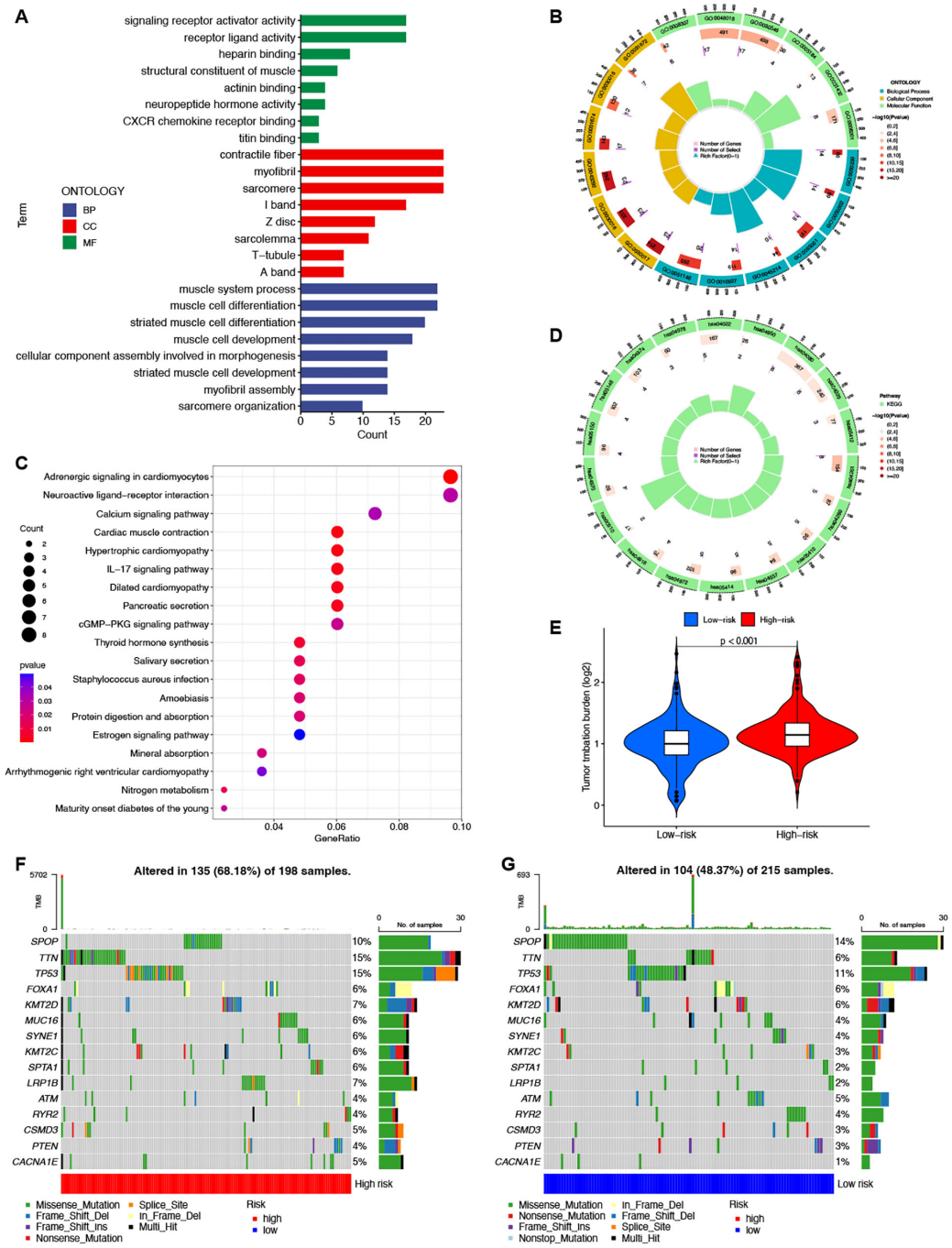


FIGURE 6
Enrichment and TMB analyses based on the prognostic ClncRNAs signature in high-risk and low-risk in patients with PCa. The GO **(A, B)** and KEGG **(C, D)** analysis shows significantly enriched biological processes and pathway between two different risk groups. **(E)** The differences of TMB scores between two different risk groups. **(F, G)** Waterfall plots showing the top 15 most frequently mutated genes in two different risk groups.

(AC010896.1, AC016394.2, and SNHG9) associated with prognosis were identified. These three lncRNAs are considered independent prognostic risk factors for PCa.

Previous studies have utilized these three ClncRNAs to predict the prognosis of other cancer types. For instance, the small nucleolar RNA host gene (SNHG) family comprises a group of lncRNAs that function as novel oncogenes in multiple cancers. Recent studies have shown that SNHG1 promotes immune evasion in breast cancer cells by regulating miRNA (15, 16), while SNHG6 facilitates colorectal cancer and glioma progression by modulating miR-101-3p expression (17, 18). Additionally, SNHG8 has been identified as a potential biomarker and therapeutic target for hepatocellular carcinoma and non-small cell lung cancer (19, 20). Elevated SNHG9 expression is correlated with poor prognosis, suggesting its involvement in PCa progression by influencing ribosomal function and immune cell infiltration (21).

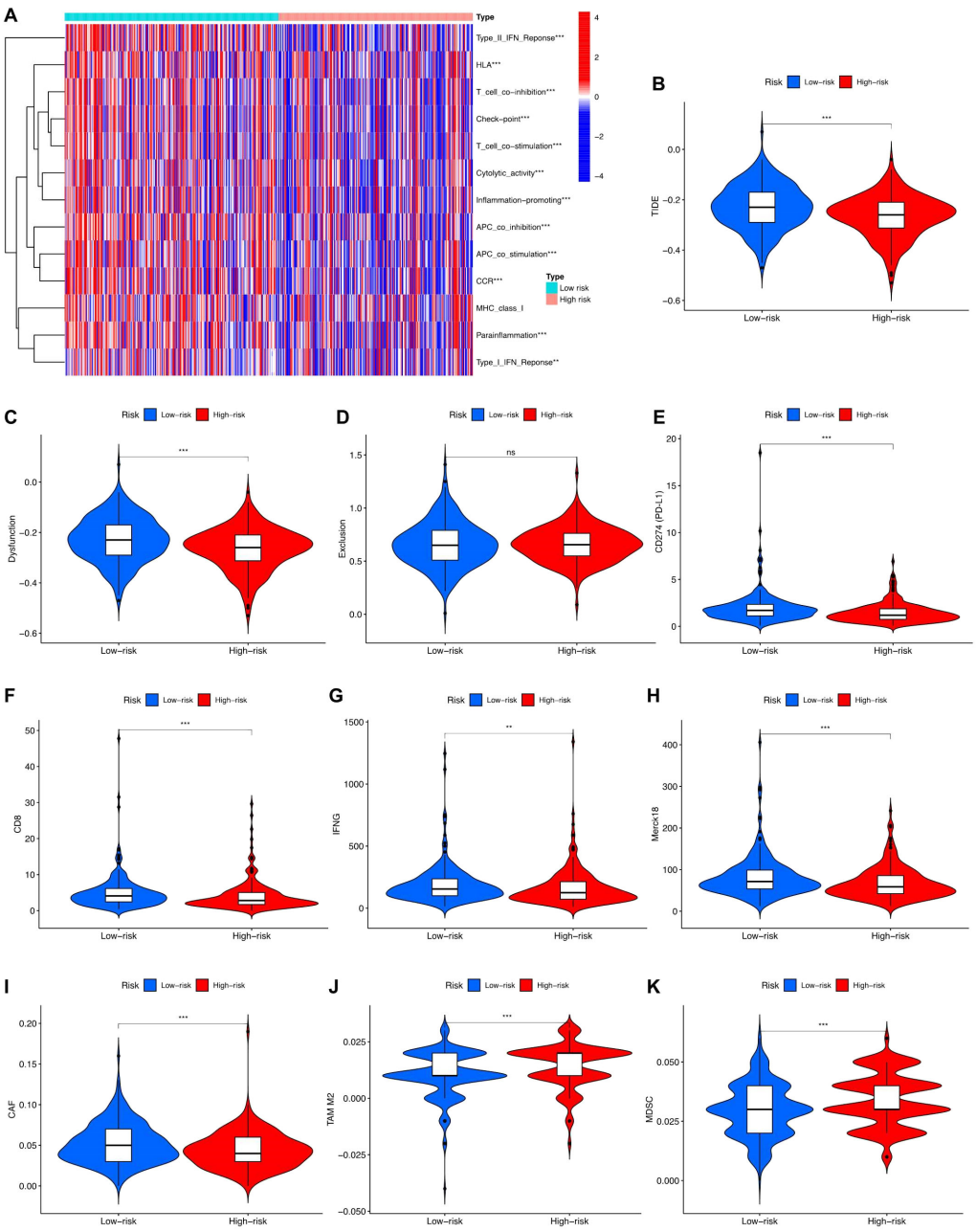


FIGURE 7
Immune function analysis in high-risk and low-risk groups. **(A)** Different proportions of immune function between the high-risk and the low-risk groups as shown in the heatmap. **(B-K)** TIDE, T cells dysfunction, T cells exclusion, CD274, CD8, INFG, Merck18, CAF, TAM M2, and MDS scores between the low-risk and high-risk groups of patients with PCa.

Consistently, SNHG9 was found to be the most highly expressed ClncRNAs observed in the PCa samples in our study. Functional experiments conducted on PC-3 and DU145 cell lines with elevated SNHG9 expression demonstrated significant inhibition of proliferation, migration, and invasion after SNHG9 knockdown. The expected outcomes were observed in cell cycle, clone formation, and western blotting experiments, highlighting the pivotal role of SNHG9 in PCa progression. Tu et al. (22) found that a cuproptosis-related prognostic gene signature involving AC016394.2 served as an independent risk indicator of gastric adenocarcinoma. Additionally, Kuo et al. (23) found that AC016394.2 can also serve as a lncRNA associated with disulfidptosis, distinct from other forms of cell death such as cuproptosis and ferroptosis, for prognostic prediction in gastric cancer. Therefore it is evident that these related lncRNAs play a crucial role in the regulation of tumor programmed cell death and warrant further investigation. However, the information about AC010896.1 in the published literature is limited. Therefore, further experiments are warranted in future studies to elucidate its role in PCa.

A prognostic signature was constructed based on the three ClncRNAs. To validate the predictive accuracy of risk score of the ClncRNAs signature, the survival analysis indicated the ability of

10.3389/fimmu.2024.1471198

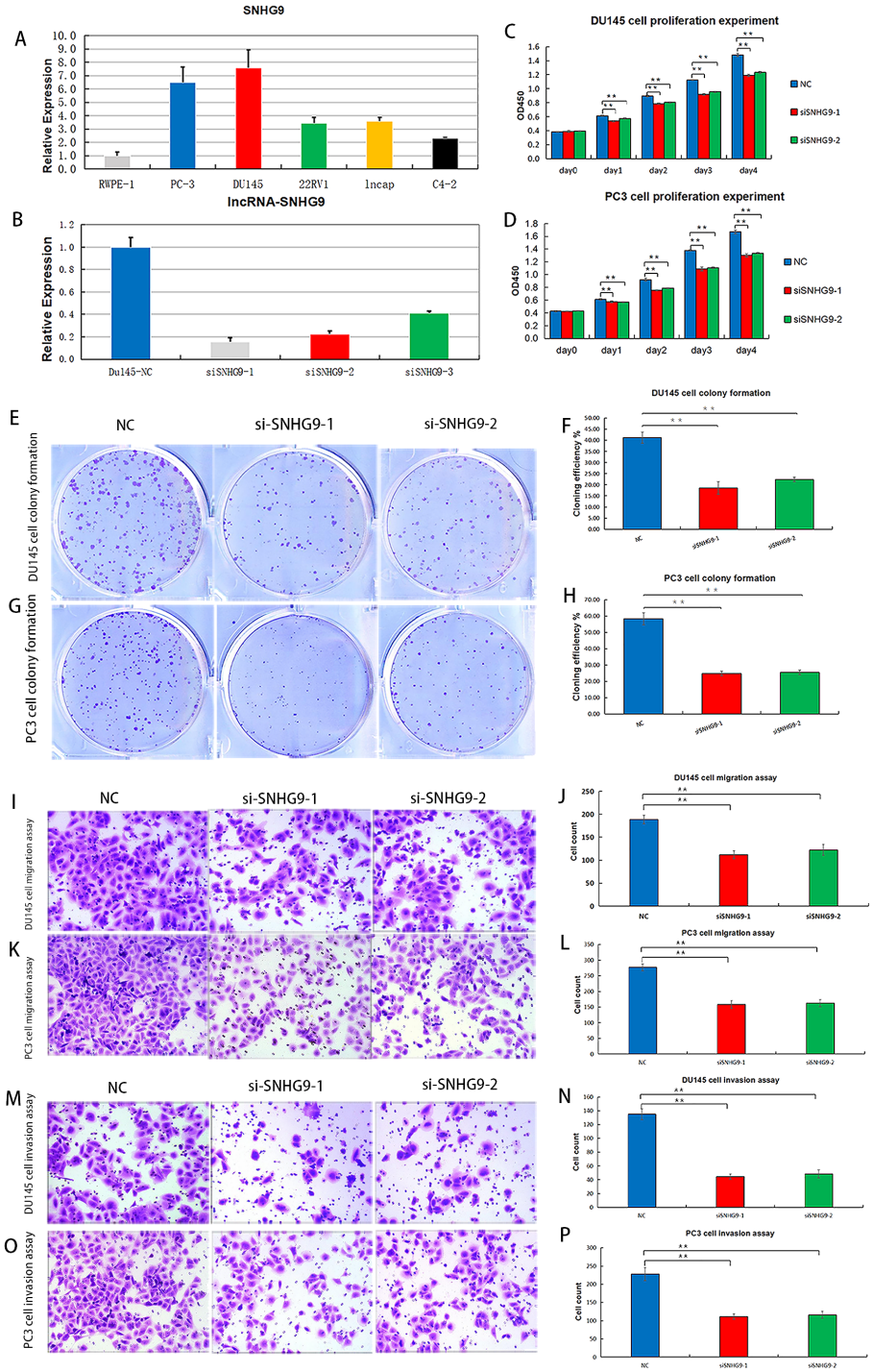


FIGURE 8
Proliferation, migration and invasion experiments after knockdown of ClncRNAs SNHG9 *in vitro*. **(A)** Expression levels of SNHG9 in human PCa cells. **(B)** Relative expression levels of DU145 cells after SNHG9 transfection with the corresponding siRNA. **(C, D)** CCK-8 assay was used to detect the effect of SNHG9 on PC3 and DU145 cells proliferation. **(E-H)** Evaluate the impact of SNHG9 on the proliferation and growth abilities of individual PC3 and DU145 cells through flat cloning experiments. **(I-L)** Use the transwell method to assess the effect of SNHG9 on the migration of PC3 and DU145 cells. **(M-P)** Use the transwell method to assess the effect of SNHG9 on the invasion of PC3 and DU145 cells. **p <0.01.

ClncRNAs signature to effectively identify and differentiate between high-risk and low-risk groups. Furthermore, the risk score of the ClncRNAs signature demonstrated superior predictive efficacy for predicting 1-, 3-, and 5-year PFI compared to other clinical characteristics, with AUC values exceeding 0.70. Compared to previously published prognostic signature related to ClncRNAs, Cheng et al. (24) constructed a signature based on differentially expressed ClncRNAs, including AC005790.1, AC011472.4, AC099791.2, AC144450.1, LIPE-AS1, and STPG3-AS1. This signature demonstrated AUC values exceeding 0.70 for predicting

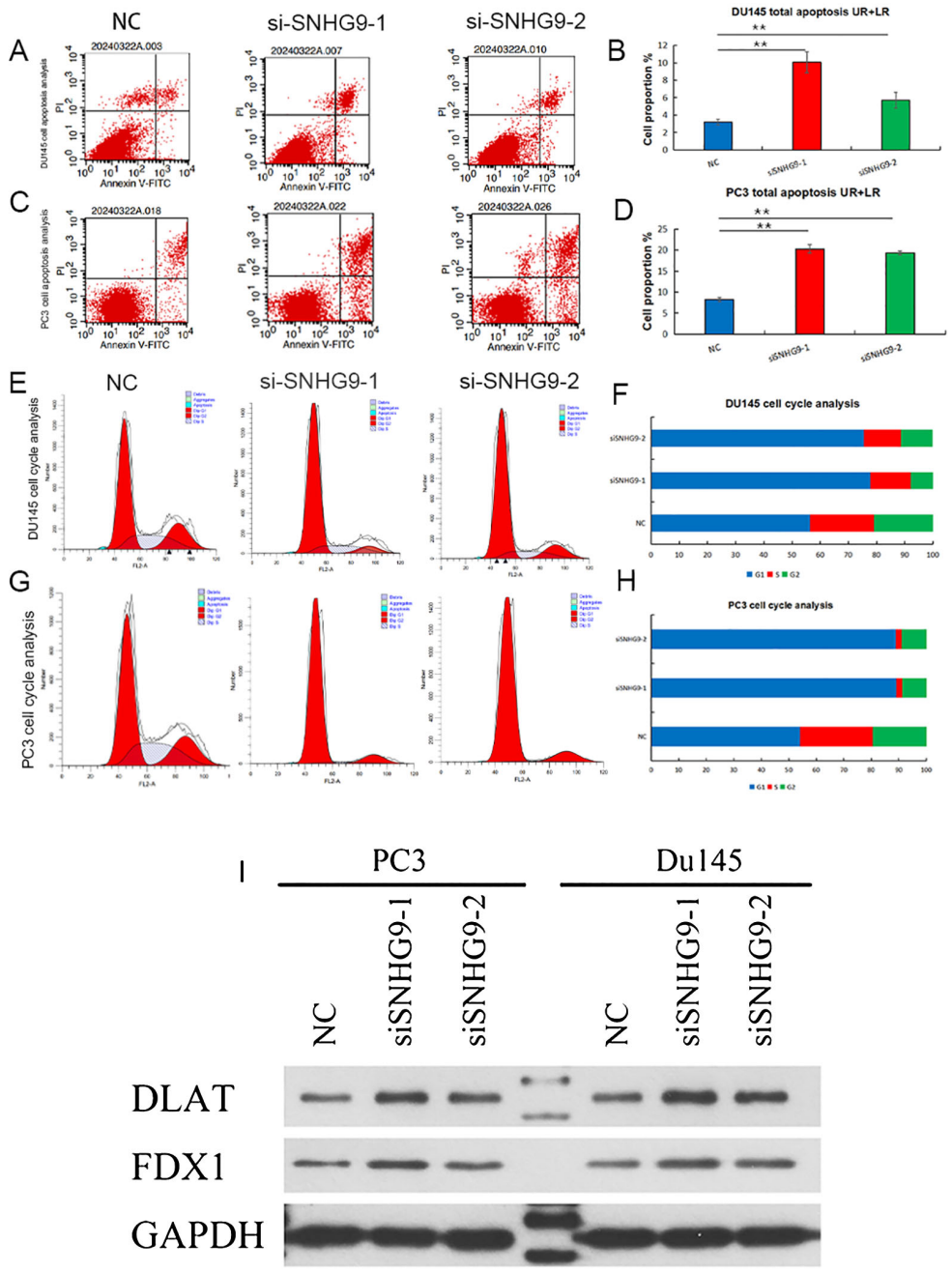


FIGURE 9
Apoptosis and cell cycle experiments, and key cuproptosis-related mRNAs validation after knockdown of ClncRNAs SNHG9 *in vitro*. **(A-H)** Examine the effect of SNHG9 on apoptosis and the cell cycle of PC3 and DU145 cells through flow cytometry. **(I)** Western blot indicated that key cuproptosis-related mRNAs (FDX1 and DLAT) were significantly upregulated in the two siSNHG9 groups relative to the control group. **p <0.01.

1-year, 3-year, and 5-year DFS. Specifically, the AUC for 5-year DFS was 0.703, while the AUC for 5-year PFI in our study reached 0.766. Additionally, Jiang et al. (25) developed a signature based on seven ClncRNAs (C1orf229, C9orf139, LIPE-AS1, MCPH1-AS1, PRR26, SGMS1-AS1, and SNHG1), which yielded an AUC of only 0.676 for 5-year DFS prediction. Although their research focused on DFS prediction, the ClncRNA signature validated through functional experiments in our study offered greater accuracy and clinical applicability. A novel nomogram was constructed to enhance the clinical applicability of the signature by combining the risk score with clinical characteristics, providing an intuitive and quantitative assessment method for predicting the 1-, 3-, and 5-year PFI in patients with PCa. Survival analysis in different subgroups revealed that the risk score of the ClncRNAs signature still exhibited highly accurate predictive capabilities for different subgroups while the predictive ability for pN1 patients was not ideal, which may be

attributed to the relatively small number of pN1 patients included. Furthermore, the results of GO and KEGG analyses between high-risk and low-risk groups indicated that the differentially expressed mRNAs were mainly enriched in processes related to the muscular system and the activation and transmission of signal receptors. Previous studies have shown significant metabolic changes in myocardial cells under an increased cardiac load, with acceleration of the tricarboxylic acid cycle, leading to heightened activity in shuttle systems and biosynthetic processes (26, 27). This further confirms the intricate relationship between copper ions and various aspects of mitochondrial respiration.

The accumulation of genetic mutations is widely recognized as a primary driver of tumorigenesis (28). By analyzing the genetic mutations between the two risk groups, we found that the high-risk group exhibited a higher TMB than the low-risk group. Previous studies have indicated that in various cancer types treated with immune checkpoint inhibitors, a higher TMB correlates with better survival rates, suggesting that a higher TMB implies the presence of more mutations in tumor cells, potentially leading to the generation of more neoantigens and indicating a potentially better response to immunotherapy in high-risk patients with PCa (29). Additionally, among the top 15 mutated genes in patients with PCa, the mutation rate of TTN showed the most significant difference between the high-risk and low-risk groups, followed by SPOP and P53. Yan et al. (30) reported the prognostic value of TTN in breast cancer. Many previous studies have confirmed that SPOP and P53 mutations are among the most common mutations in PCa, with TP53 mutations indicating a poor prognosis, whereas SPOP mutations suggest the opposite. This consistency with the genetic mutation profiles distinguished by our risk scoring between the high-risk and low-risk groups further validates the predictive accuracy of risk score of the ClncRNAs signature (31, 32). The TIDE scores between high-risk and low-risk patients were also evaluated. Consistent with the findings of Zhao et al. (33), we observed that PCa patients with lower TMB and low-risk scores tended to have higher TIDE scores. Higher TIDE scores correlated with lower responsiveness to anti-PD-1 and anti-CTLA-4 therapies (34). This further supports the notion that patients with high-risk scores may respond better to immunotherapy. However, TIDE cannot predict patient survival outcomes, so this finding did not contradict the prognostic results of risk score of the ClncRNAs signature.

Currently, immunotherapy has revolutionized treatment strategies for many cancers (35); however, treatment selection for castration-resistant prostate cancer (CRPC) remains a primary challenge in PCa management. Previous study has suggested immunotherapy as a promising treatment modality for CRPC (36). However, the role of androgens in the regulation of immune function and the effects of androgen deprivation on adaptive immune responses remain unclear. Androgens are conventionally believed to possess immunosuppressive effects, and androgen deprivation has been shown to enhance T-cell function in autoimmune disease models (37, 38). Additionally, research has identified the mutation status of SPOP as an important independent prognostic marker for metastatic PCa, with SPOP mutations rendering tumor cells susceptible to androgen deprivation therapy (39, 40). In the immunocellular infiltration analysis conducted in this study, the high-risk group exhibited downregulated APC co-stimulation, C-C CCR, pro-inflammatory markers, major MHC class I, anti-inflammatory, and type II IFN response states. Previous studies have indicated that APC co-stimulation can enhance T cell activation and sustain anti-tumor immunity to improve immunotherapeutic response rates in patients with PCa (41, 42). Therefore, further investigation of the differential expression of immune cells may hold significant promise for the future treatment of PCa.

The current study had some limitations. First, our data source was singular, relying solely on TCGA database for internal validation without an external validation from additional databases. Second, while the initial validation of the identified ClncRNA SNHG9 in PCa was conducted through cell experiments, further experimental confirmation *in vivo* is required. This will be our next plan for new in-depth research projects.

In conclusion, a prognostic signature of PCa based on three ClncRNAs was established, enabling accurate prognosis prediction for PCa. SNHG9 knockdown inhibited DNA synthesis, cell cycle progression, and clonogenic ability, while upregulating cancer-related genes, suggesting its role in promoting PCa cell proliferation, migration, and invasion. This study not only deepens our understanding of the interplay between ClncRNAs and PCa but also may provide fresh insights for devising advanced therapeutic strategies aimed at enhancing the management of PCa patients.

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/s.

Ethics statement

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used. Ethical approval was not required for the studies on animals in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

Author contributions

Y-L: Funding acquisition, Investigation, Methodology, Resources, Writing – review & editing. J-W: Data curation, Investigation, Validation, Writing – original draft. X-L: Methodology, Software,

Validation, Writing – review & editing, Formal analysis. Q-L: Data curation, Methodology, Validation, Writing – review & editing. J-T: Data curation, Formal analysis, Methodology, Validation, Writing – original draft. H-H: Funding acquisition, Resources, Supervision, Writing – review & editing. R-Z: Writing – review & editing, Software, Validation. Z-C: Data curation, Software, Validation, Writing – review & editing, Formal analysis. Y-Z: Data curation, Funding acquisition, Resources, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by Guangdong Basic and Applied Basic Research Foundation (2022A1515220032), Guangdong Medical Science and Technology Research Foundation(B2023195), Science and Technology Project of Zhongshan City (2020B1073), Zhongshan city people's hospital Major Project of Scientific Research Foundation (BG20228249) and Zhongshan City People's Hospital Outstanding Youth Project (SG2023106).

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1471198/full#supplementary-material

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I. 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. Teo MY, Rathkopf DE, Kantoff P. Treatment of advanced prostate cancer. *Annu Rev Med*. (2019) 70:479–99. doi: 10.1146/annurev-med-051517-011947

3. Zyner KG, Simeone A, Flynn SM, Doyle C, Marsico G, Adhikari S, et al. G-quadruplex DNA structures in human stem cells and differentiation. *Nat Commun*. (2022) 13:142. doi: 10.1038/s41467-021-27719-1

4. Tsvetkov P, Coy S, Petrova B, Dreishpoon M, Verma A, Abdusamad M, et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science*. (2022) 375:1254–61. doi: 10.1126/science.abf0529

5. Gao N, Li Y, Li J, Gao Z, Yang Z, Li Y, et al. Long non-coding RNAs: the regulatory mechanisms, research strategies, and future directions in cancers. *Front Oncol*. (2020) 10:598817. doi: 10.3389/fonc.2020.598817

6. Statello L, Guo C-J, Chen L-L, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol*. (2021) 22:96–118. doi: 10.1038/s41580-020-00315-9

7. Xiong Z, Li X, Yin S, Xie M, Mao C, Zhang F, et al. Prognostic value of N6-methyladenosine-related lncRNAs in early-stage colorectal cancer: association with immune cell infiltration and chemotherapeutic drug sensitivity. *Front Mol Biosci*. (2021) 8:724889. doi: 10.3389/fmolb.2021.724889

8. Martens-Uzunova ES, Böttcher R, Croce CM, Jenster G, Visakorpi T, Calin GA. Long noncoding RNA in prostate, bladder, and kidney cancer. *Eur Urol*. (2014) 65:1140–51. doi: 10.1016/j.eururo.2013.12.003

9. Zhang H, Zhang Y, Gao L, Song W, Zhang H, Yang G, et al. Cuproptosis-related LINC01711 promotes the progression of kidney renal clear cell carcinoma. *Am J Cancer Res*. (2023) 13:2617.

10. Shen J, Du M, Liang S, Wang L, Bi J. Construction of a cuproptosis-associated lncRNA prognostic signature for bladder cancer and experimental validation of cuproptosis-related lncRNA UBE2Q1-AS1. *Front Med*. (2023) 10. doi: 10.3389/fmed.2023.1222543

11. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: J Integr Biol*. (2012) 16:284–7. doi: 10.1089/omi.2011.0118

12. Bock C. Analysing and interpreting DNA methylation data. *Nat Rev Genet*. (2012) 13:705–19. doi: 10.1038/nrg3273

13. Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. (2018) 28:1747–56. doi: 10.1101/gr.239244.118

14. Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, et al. Pathway enrichment analysis and visualization of omics data using g: Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc*. (2019) 14:482–517. doi: 10.1038/s41596-018-0103-9

15. Pei X, Wang X, Li H. LncRNA SNHG1 regulates the differentiation of Treg cells and affects the immune escape of breast cancer via regulating miR-448/IDO. *Int J Biol Macromol*. (2018) 118:24–30. doi: 10.1016/j.ijbiomac.2018.06.033

16. Xiong X, Feng Y, Li L, Yao J, Zhou M, Zhao P, et al. Long non–coding RNA SNHG1 promotes breast cancer progression by regulation of LMO4. *Oncol Rep*. (2020) 43:1503–15. doi: 10.3892/or.2020.7530

17. Meng Q, Yang B-Y, Liu B, Yang J-X, Sun Y. Long non-coding RNA SNHG6 promotes glioma tumorigenesis by sponging miR-101-3p. *Int J Biol Markers*. (2018) 33:148–55. doi: 10.1177/1724600817747524

18. Shao Q, Xu J, Deng R, Wei W, Zhou B, Yue C, et al. SNHG 6 promotes the progression of Colon and Rectal adenocarcinoma via miR-101-3p and Wnt/β-catenin signaling pathway. *BMC Gastroenterol*. (2019) 19:1–8. doi: 10.1186/s12876-019-1080-3

19. Chen C, Zhang Z, Li J, Sun Y. SNHG8 is identified as a key regulator in non-small-cell lung cancer progression sponging to miR-542-3p by targeting CCND1/CDK6. *OncoTargets Ther*. (2018) 11:6081–90. doi: 10.2147/OTT.S170482

20. Dong J, Teng F, Guo W, Yang J, Ding G, Fu Z. lncRNA SNHG8 promotes the tumorigenesis and metastasis by sponging miR-149-5p and predicts tumor recurrence in hepatocellular carcinoma. *Cell Physiol Biochem*. (2018) 51:2262–74. doi: 10.1159/000495871

21. Li C, Hu J, Hu X, Zhao C, Mo M, Zu X, et al. LncRNA SNHG9 is a prognostic biomarker and correlated with immune infiltrates in prostate cancer. *Trans Androl Urol*. (2021) 10:215. doi: 10.21037/tau-20-1134

22. Tu H, Zhang Q, Xue L, Bao J. Cuproptosis-related lncRNA gene signature establishes a prognostic model of gastric adenocarcinoma and evaluate the effect of antineoplastic drugs. *Genes*. (2022) 13:2214. doi: 10.3390/genes13122214

23. Kang K, Li X, Peng Y, Zhou Y. Comprehensive analysis of disulfidptosis-related lncrnas in molecular classification, immune microenvironment characterization and prognosis of gastric cancer. *Biomedicines*. (2023) 11:3165. doi: 10.3390/biomedicines11123165

24. Cheng X, Zeng Z, Yang H, Chen Y, Liu Y, Zhou X, et al. Novel cuproptosis-related long non-coding RNA signature to predict prognosis in prostate carcinoma. *BMC Cancer*. (2023) 23:105. doi: 10.1186/s12885-023-10584-0

25. Jiang S, Li Z, Dou R, Lin Z, Zhang J, Zhang W, et al. Construction and validation of a novel cuproptosis-related long noncoding RNA signature for predicting the outcome of prostate cancer. *Front Genet*. (2022) 13:976850. doi: 10.3389/fgene.2022.976850

26. Tverskaya M, Sukhoparova V, Klyuchikov VY, Alipov N, Polivoda M, Bulgakov S. Histoenzymological characteristics of smooth muscle cells in myocardial vessels: a

comparative study under conditions of increased left or right ventricular afterload. *Bull Exp Biol Med.* (2016) 161:7–10. doi: 10.1007/s10517-016-3332-3

27. Williams GS, Boyman L, Lederer WJ. Mitochondrial calcium and the regulation of metabolism in the heart. *J Mol Cell Cardiol.* (2015) 78:35–45. doi: 10.1016/j.yjmcc.2014.10.019

28. Lulla RR, Saratsis AM, Hashizume R. Mutations in chromatin machinery and pediatric high-grade glioma. *Sci Adv.* (2016) 2:e1501354. doi: 10.1126/sciadv.1501354

29. Samstein RM, Lee C-H, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet.* (2019) 51:202–6. doi: 10.1038/s41588-018-0312-8

30. Yan Y, Liang Q, Liu Y, Zhou S, Xu Z. COL12A1 as a prognostic biomarker links immunotherapy response in breast cancer. *Endocr Rel Cancer.* (2023) 30:e230012. doi: 10.1530/ERC-23-0012

31. Nakazawa M, Fang M H, Marshall C, Lotan TL, Isaacsson Velho P, Antonarakis ES. Clinical and genomic features of SPOP-mutant prostate cancer. *Prostate.* (2022) 82:260–8. doi: 10.1002/pros.24269

32. Nientiedt C, Budczies J, Endris V, Kirchner M, Schwab C, Jurcic C, et al. Mutations in TP53 or DNA damage repair genes define poor prognostic subgroups in primary prostate cancer. *Urol Oncol.* (2022) 40:8.e11–8. doi: 10.1016/j.urolonc.2021.06.024

33. Zhao H, Wu T, Luo Z, Huang Q, Zhu S, Li C, et al. Construction and validation of a fatty acid metabolism-related gene signature for predicting prognosis and therapeutic response in patients with prostate cancer. *PeerJ.* (2023) 11:e14854. doi: 10.7717/peerj.14854

34. Woo S-R, Corrales L, Gajewski TF. The STING pathway and the T cell-inflamed tumor microenvironment. *Trends Immunol.* (2015) 36:250–6. doi: 10.1016/j.it.2015.02.003

35. Khalil DN, Smith EL, Brentjens RJ, Wolchok JD. The future of cancer treatment: immunomodulation, CARs and combination immunotherapy. *Nat Rev Clin Oncol.* (2016) 13:273–90. doi: 10.1038/nrclinonc.2016.25

36. Rosenberg SA, Yang JC, Restifo NP. Cancer immunotherapy: moving beyond current vaccines. *Nat Med.* (2004) 10:909–15. doi: 10.1038/nm1100

37. Fijak M, Schneider E, Klug J, Bhushan S, Hackstein H, Schuler G, et al. Testosterone replacement effectively inhibits the development of experimental autoimmune orchitis in rats: evidence for a direct role of testosterone on regulatory T cell expansion. *J Immunol.* (2011) 186:5162–72. doi: 10.4049/jimmunol.1001958

38. Radojevic K, Arsenovic-Ranin N, Kosec D, Pesic V, Pilipovic I, Perisic M, et al. Neonatal castration affects intrathymic kinetics of T-cell differentiation and the spleen T-cell level. *J Endocrinol.* (2007) 192:669. doi: 10.1677/joe.1.07019

39. Zhou J, Lai Y, Peng S, Tang C, Chen Y, Li L, et al. Comprehensive analysis of TP53 and SPOP mutations and their impact on survival in metastatic prostate cancer. *Front Oncol.* (2022) 12:957404. doi: 10.3389/fonc.2022.957404

40. Wang Z, Song Y, Ye M, Dai X, Zhu X, Wei W. The diverse roles of SPOP in prostate cancer and kidney cancer. *Nat Rev Urol.* (2020) 17:339–50. doi: 10.1038/s41585-020-0314-z

41. Reiner SL. Development in motion: helper T cells at work. *Cell.* (2007) 129:33–6. doi: 10.1016/j.cell.2007.03.019

42. Sturgill ER, Rolig AS, Linch SN, Mick C, Kasiewicz MJ, Sun Z, et al. Galectin-3 inhibition with belapectin combined with anti-OX40 therapy reprograms the tumor microenvironment to favor anti-tumor immunity. *Oncoimmunology.* (2021) 10:1892265. doi: 10.1080/2162402X.2021.1892265

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 04 November 2024
DOI 10.3389/fimmu.2024.1463258

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Lushan Xiao,
Southern Medical University, China
Dafina Fondaj,
University of Bari Aldo Moro, Italy

*CORRESPONDENCE
Hui Liu
thyroid2018@126.com
Yu Wu
429146964@qq.com

†These authors have contributed equally to this work

RECEIVED 11 July 2024
ACCEPTED 16 October 2024
PUBLISHED 04 November 2024

CITATION
Wu S, Liu Y, Zeng Y, Ruan X, Tao M, Lin W, Liu C, Chen H, Liu H and Wu Y (2024) Prognostic value of EMT-related genes and immune cell infiltration in thyroid carcinoma. *Front. Immunol.* 15:1463258. doi: 10.3389/fimmu.2024.1463258

Prognostic value of EMT-related genes and immune cell infiltration in thyroid carcinoma

Shuping Wu[1,2†], Yu Liu[2†], Yu Zeng[2†], Xianhui Ruan[2], Mei Tao[2], Wenrong Lin[3], Chang Liu[1], Hongbin Chen[1], Hui Liu[1*] and Yu Wu[1*]

[1]Department of Head and Neck Surgery, Clinical Oncology School of Fujian Medical University, Fujian Cancer Hospital, Fuzhou, Fujian, China, [2]Department of Thyroid and Neck Tumor, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin, China, [3]Department of Ultrasound, Clinical Oncology School of Fujian Medical University, Fujian Cancer Hospital, Fuzhou, Fujian, China

Background: The Epithelial–Mesenchymal Transition (EMT) is a very important process involved in cancer invasion and metastasis. Additionally, the Cathepsin K (CTSK) gene is closely related to the degradation of the extracellular matrix, which is a critical component of the EMT. The purpose of this study was to determine the relationships between EMT-related genes and immune cell infiltration and their prognostic value in Thyroid carcinoma (THCA). The effect of the CTSK gene on the aggressive biological features of THCA was assessed.

Methods: Within the framework of the present study, the THCA cohort was analyzed in detail based on data obtained from The TCGA database in the context of the EMT. The TCGA-THCA cohort was then divided into two groups, namely, high- and low-risk groups, based on the calculated EMT scores. Finally, based on the findings from the Weighted Gene Co-Expression Network Analysis (WGCNA) algorithm, LASSO regression analysis, and Kaplan–Meier plotter, we selected five genes (CTSK, C3ORF80, FBLN2, PRELP and SRPX2) associated with patient prognosis. Furthermore, this study examined the presence of various immune cells within the THCA samples using three distinct algorithms, namely ssGSEA, xCell, and MCPcounter. Additional studies have been conducted to establish the roles of CTSK in THCA cell proliferation and migration using various assays, such as CCK8, colony formation, EdU proliferation, Transwell migration and wound healing assays. Additionally, the involvement of CTSK in the regulation of various EMT-related markers was confirmed using Western blot analysis.

Results: Based on EMT scores, TCGA-THCA patients were further divided into two groups, and the study revealed that patients in the high-risk group had a worse prognosis than those in the low-risk group. Among the five genes linked to the prognostic value of EMT (CTSK, C3ORF80, FBLN2, PRELP, and SRPX2), CTSK exhibited notably elevated expression in the high-risk cohort. This group also exhibited pronounced immune cell infiltration, with a marked correlation observed between CTSK expression and the levels of macrophages, MDSCs, and various T-cell subtypes. Furthermore, *in vitro* studies demonstrated that reducing CTSK expression led to significant reductions in THCA cell viability; clonogenic, proliferative, motility and migratory capacities; and the expression of key EMT-related proteins, including N-cadherin, vimentin, slug, and snail.

Conclusion: Our results suggest that the expression of CTSK, a gene associated with the EMT, may be associated with THCA onset and progression and thus may serve as a promising prognostic biomarker.

KEYWORDS

THCA, EMT, CTSK, immune infiltration, biomarker

1 Introduction

Thyroid cancer (THCA), which originates from either follicular or parafollicular epithelial cells within the thyroid gland, is the predominant malignant endocrine tumor, accounting for approximately 90% of all malignancies within this system (1–3). Despite the generally favorable prognostic outlook for most cases of THCA, specific forms, such as anaplastic, medullary, and treatment-resistant thyroid cancers, demonstrate a propensity for recurrence and metastasis, ultimately leading to detrimental outcomes (4, 5). In response to these challenges, recent therapeutic advancements have included the adoption of sorafenib, a multitargeted small-molecule tyrosine kinase inhibitor that acts on VEGFR, BRAF, and RET. Although beneficial for managing advanced or metastatic forms of THCA, the application of sorafenib is frequently hampered by its potential to trigger a spectrum of adverse effects within patients (6). The limitations of current treatment options emphasize the pressing necessity of identifying and validating novel genetic markers. These markers not only gauge the aggressive nature of tumors but also facilitate the development of targeted treatments, thereby enhancing management strategies for patients suffering from THCA with an otherwise poor prognosis.

The epithelial–mesenchymal transition (EMT) is a process in which epithelial cells transform into mesenchymal cells, as described previously (7). This process is indispensable for embryonic development and wound healing and has a major impact on tumor growth and metastasis. This dual role of EMT in normal wound healing and pathophysiological processes such as cancer portrays the significance of EMT in living organisms. These interactions include promoting tumor cell motility and invasion, increasing tumor cell stemness, and increasing tumor cell chemoresistance and immunoresistance. The process known as EMT is intricately governed by an extensive range of factors that originate both internally within cells and externally from the cellular environment. These regulatory elements include numerous transcription factors, diverse mechanisms of posttranslational modification, comprehensive epigenetic changes, and various noncoding RNAs (8). Various studies have established that the EMT does not operate as a straightforward binary mechanism. In contrast, this process occurs in a step-by-step manner through several well-coordinated cellular stages (9). The connection between EMT and MSCs in THCA is crucial, and their interaction might be responsible for the poor outcome of patients with certain THCA subtypes through the stimulation of MSC-like cell proliferation for metastasis (10). Some papillary tumors are associated with metastatic and invasive behaviors despite the fact that most of the thyroid tumors are well differentiated because of dedifferentiation. This finding can be explained by the EMT, whereby thyroid epithelial cells undergo a transition, assume a fibroblastic morphology, become less cohesive and more motile and express mesenchymal markers (11). The tumor microenvironment (TME) has been described as a complex structure composed of both living and nonliving components. Other cell types that constitute this environment include endothelial cells, adipocytes, fibroblasts, epithelial cells and immune cells in addition to primary tumor cells. Moreover, the TME includes acellular components, including the extracellular matrix (ECM), cytokines, chemokines, growth factors and antibodies, which are involved in carcinogenesis and tumor advancement (12, 13). Several works have emphasized a strong correlation between a high level of immune cell infiltration in tumor tissue and patient survival. The relationships between the elements of an individual's antitumor defense and the features of the tumor, including the rates of tumor growth, invasion, and metastasis, influence the response to therapy and the predicted course of THCA (14). Cathepsin K (CTSK) is a ubiquitously expressed protease that plays enzymatic and nonenzymatic roles in numerous pathologies (15). Several recent studies have shown a strong correlation between increased CTSK levels and the onset and poor prognosis of pancreatic and hepatocellular carcinomas. Additionally, higher CTSK levels promote disease progression to the lymph nodes in patients with oral squamous cell carcinoma (16–18). It should be noted that CTSK is reportedly involved in the promotion of an M2-like macrophage phenotype in castration-resistant prostate cancer (19). However, it remains unclear how CTSK levels are associated with THCA patient prognosis or exactly how CTSK is involved in the development of THCA.

In the present study, we applied the WGCNA technique to systematically identify genes integral to the EMT and to formulate related coexpression networks. Patients within the TCGA-THCA dataset were stratified based on their EMT scores, yielding two distinct groups—those with high EMT and those with low EMT. Our analysis focused on exploring differences in prognosis and

immune cell infiltration between these categorizations. Notably, this study identified CTSK as a critical gene associated with the EMT, emphasizing its substantial role in modulating both tumor proliferation and the metastatic process in THCA. These insights could improve our understanding of the influence of the EMT on THCA pathophysiology and could significantly refine approaches to develop personalized treatment modalities for affected patients.

2 Methods

2.1 Data sources

Thus, the present study used an analytical dataset consisting of 501 THCA samples derived from the TCGA dataset. This cleaning of the initial data was performed using Perl programming to remove any duplicates or incomplete observations from the dataset. The next procedures were the normalization and annotation of the data to meet the requirements of the subsequent analysis. To determine the genes that were significantly differentially expressed between normal and THCA tissues, the 'limma' package within the R environment was used. This analysis used a stringent threshold for statistical analysis where only genes that had a log fold change of ±1 and a P value of less than 0.05 were considered significant. DEG visualization was informative; the chromosomal positions of the DEGs were presented in a circular form using the 'RCircos' package. Furthermore, to increase the applicability of this study on the EMT, genes connected to the EMT were selectively incorporated into the analysis. These genes were obtained from the MSigDB and complemented the study with a focused view on biological processes that could underlie the development of THCA and its metastatic spread.

2.2 Construction of WGCNA and identification of modules related to the EMT

In this comprehensive analysis, the 'WGCNA' package in R was used to identify the gene modules related to the EMT using the TCGA-THCA dataset. The flow of the study began with the identification of the first 1000 genes that showed the highest variability between samples split into two groups with low and high EMT scores. To make the data more suitable for analysis, two suspicious samples, which were determined using cluster analysis, were removed. This was followed by the analysis of scale independence as well as the mean connectivity across the modules at different power levels. This step was necessary for establishing the most appropriate soft threshold that would help in providing a stable analysis of the network with a signed R^2 value of 0. The 95% confidence interval is considered an adequate level of scale independence. After fixing the soft threshold, the next step was to examine the relationship between the gene expression modules and the EMT parameters. To this end, only those modules that had detectable correlations with the EMT were chosen for further assessment at the level of individual GO terms within the modules. This approach was proposed to uncover not only genes that are highly relevant to the EMT but also genes that are strongly associated with the selected modules based on eigengenes. The GS for each gene in these crucial modules was defined as the absolute value of the correlation with the clinical phenotypes. Moreover, MM was established by plotting the correlation of each gene's expression pattern to the module eigengene. To depict these findings, scatter plots were created to show the relationship between GS and MM for the genes in the highlighted modules and highlight the complexity of the relationship between these genes and their potential for clinical application.

In our WGCNA analysis, we meticulously selected genes for inclusion in the network based on a series of rigorous criteria aimed at enhancing the validity and interpretability of our findings. Initially, we filtered out genes with low expression levels by setting a threshold where only those genes exhibiting a mean expression value greater than 1 Transcripts Per Million (TPM) across all samples were retained. This step ensured that we focused on genes with sufficient expression for meaningful analysis. Subsequently, we calculated the coefficient of variation (CV) for each gene, which is defined as the ratio of the standard deviation to the mean expression. We included only the top 50% of genes exhibiting the highest CV values, thereby prioritizing those genes that demonstrated significant variability in expression across samples, indicative of their potential biological relevance. Furthermore, to align our analysis with existing biological knowledge, we cross-referenced our gene list with cancer-related genes obtained from well-established databases such as The Cancer Genome Atlas (TCGA) and GeneCards. This additional filtering step allowed us to focus specifically on genes that have documented associations with cancer pathways and processes. After these selection steps, the remaining genes were subjected to the standard WGCNA procedures to construct the co-expression network, wherein we employed a soft-thresholding power to achieve scale-free topology, followed by hierarchical clustering to identify modules of co-expressed genes. This comprehensive approach facilitated the identification of biologically relevant gene modules that may contribute to cancer pathology.

2.3 Enrichment analysis of key DEGs

Enrichment analysis of the 68 selected DEGs was performed using the R program's clusterProfiler package (20). This systematic review included assessments based on both GO and KEGG analyses. To ensure rigorous statistical evaluation, the study adhered to stringent criteria. Therefore, the significance levels were set at an adjusted P value and adjusted q value of less than 0.05. Genes or pathways for which the p value was less than 0.05 were considered to be significantly enriched; this defined the biological relevance of the gene or pathway. In this analysis, the FDR level was set at 0.05 or less to ensure the credibility of the identified gene relationships and pathway impacts.

2.4 Construction and validation of the EMT signature

The DEGs were then analyzed using univariate Cox regression analysis with the "tinyarray" package, and LASSO regression, which is a machine learning method, was used. The proposed approach facilitated improved evaluation of the probability effects of specific genes. Of the 68 DEGs, five (CTSK, C3ORF80, FBLN2, PRELP, and SRPX2) had clinical prognostic significance and were included in the prognostic model. In selecting CTSK, C3ORF80, FBLN2, PRELP, and SRPX2 for our EMT signature model, we based the decision on both their statistical significance and biological relevance to EMT and THCA. Each of these genes was identified through a rigorous screening process using WGCNA and LASSO regression, followed by functional enrichment analysis. These five genes stood out due to their significant association with poor patient prognosis and strong involvement in key EMT-related pathways. CTSK was selected because of its well-established role in extracellular matrix degradation, a critical component of EMT. C3ORF80 is involved in cellular processes that contribute to immune regulation and cancer progression. Its expression was correlated with immune infiltration, particularly macrophages and T cells, which are crucial to the tumor microenvironment in THCA. This made C3ORF80 a relevant marker for both immune-related and EMT-driven tumor progression. FBLN2 is part of the extracellular matrix, where it plays a role in stabilizing the structural integrity of tissues. Thus, FBLN2 contributes to our understanding of EMT by highlighting extracellular matrix remodeling in THCA. PRELP is involved in cell-matrix interactions and has been associated with the regulation of EMT through matrix reorganization. SRPX2 was chosen due to its role in angiogenesis and tumor cell invasion, two processes integral to EMT. These five genes together form a robust model that captures both the epithelial and mesenchymal aspects of EMT.

Additionally, to categorize the patients into low-risk and high-risk groups, a median risk score was used. This method highlighted the differences in prognosis between these groups. Risk score: CTSK * 0. 286 + C3ORF80 * 0. 478 - FBLN2 * 0. 636 - PRELP * (-0. 166) + SRPX2 * 0. 310. Subsequently, the TCGA-THCA cohort was split into training and validation sets based on a 2:1 ratio and an 8:1 ratio, respectively. The EMT prognostic model was developed using multiple regression analysis of the coefficients of five critical genes. This robust model facilitated the stratification of the TCGA-THCA cohort into two distinct groups, namely, the high-risk group and the low-risk group, depending on the likelihood of disease progression. To determine DEGs between these risk groups, the Wilcoxon rank-sum test was applied, which demonstrated the genetic differences that led to the different prognoses. To compare the discriminative ability of the EMT model for predicting the 1-, 2-, and 3-year PFIs, ROC curves were constructed. The AUC was calculated using the 'survivalROC' package to determine the efficiency of the model in predicting patient prognosis. Furthermore, a Kaplan–Meier estimator was used to compare PFIs among the various risk categories of patients. Statistical analysis of the differences in survival rates was performed using the log rank test at a significance level of $p < 0.05$, hence validating the model's ability to identify patients with higher and lower risks of disease progression.

2.5 Genomic mutation analysis

The step-by-step approach for obtaining CNV data for the THCA cohort was performed using the R-based 'TCGAbiolinks' package through which the GDC portal was accessed. All these analyses were performed using the Genome Reference Consortium Human Build 38 (GRCh38) to avoid variation in genomic alignment. CNV analysis was subsequently performed using the advanced GISTIC2.0 algorithm hosted on the GenePattern platform (21). Genetic analysis was performed on the website http://cloud.genepattern.org/gp/pages/index.jsf using default parameters, including a confidence level of 0.9 to provide statistically accurate results. To display the CNV data that were obtained in the study, the 'Maftools' package (22) in R was used to generate a clear map of genomic alterations within the patient population. Furthermore, to better visualize the distribution of highly mutated genes among the clinical subtypes within the THCA samples, waterfall plots were created. These plots were created with the most current version of 'maftools' (version 2.12). This process allowed for the compilation and depiction of precise mutation information alongside clinical categorizations, thus improving the understanding of the genomic environment in patients with this type of cancer.

2.6 Analysis of immune infiltration

We used three computational algorithms, namely, ssGSEA, xCell, and MCPcounter, to calculate immune infiltration scores, which were visualized using boxplots, stacked plots, correlation scatter plots, and heatmaps. xCell (https://xcell.ucsf.edu/) was used to quantify the infiltration abundance of 67 immune cell types based on transcriptomic data. xCell employs advanced machine learning techniques to derive gene signatures from thousands of diverse cell types, significantly reducing correlations among similar cell types. This approach has been validated using detailed computer simulations that analyze both features and cellular immunophenotyping, demonstrating the effectiveness of xCell in precisely delineating cellular heterogeneity across tissue expression profiles. Next, the ssGSEA method was applied to compute enrichment scores for individual samples and pairs of gene sets, enabling the assessment of the extent of immune infiltration within these samples. Furthermore, the MCPcounter tool was utilized to measure the presence of ten different immune cells within the transcriptomic data, providing a quantitative analysis of immune cell abundance.

2.7 Prognostic analysis using CTSK

Using transcriptome data from 513 patients with THCA obtained from the TCGA database, patients were categorized into groups based on high or low CTSK expression, with an established optimal threshold of 3.7326 for gene expression levels. Kaplan–Meier survival curves were then constructed to depict the survival outcomes for both the high-expression and low-expression groups, enabling a comparative analysis of their survival durations.

2.8 Cell culture and cell transfection

In this study, the THCA cancer cell lines CAL-62 and KTC-1 were acquired from the American Type Culture Collection (ATCC) and cultured under controlled conditions at 37°C in an atmosphere containing 5% CO_2. This study employed negative control(NC) and siRNAs specifically targeting CTSK, which were produced by the Tsingke Company (Beijing, China), and the sequences of siRNAs were as follows: siNC(5'-UUCUCCGAACGUGUCACGUTT-3'); siCTSK-1(5'-CAGCAAAGGTGTGTATTATGATGAA-3'); and siCTSK-2(5'-GGUUCAGAAGAUGACUGGA(dT)(dT)-3'). For gene silencing experiments, cells were transiently transfected with either the negative control or CTSK-targeted siRNAs utilizing Lipofectamine 2000 reagent following the protocols provided by Invitrogen.

2.9 Quantitative real-time PCR

RNA was isolated from THCA cells with TRIzol reagent (AC0101-B; SparkJade, China). Subsequently, 1 μg of extracted RNA was converted to cDNA utilizing a High-Capacity cDNA Reverse Transcription Kit (Vazyme, R223-01). This cDNA served as the template for subsequent exponential amplification, which was performed using 2 × HQ SYBR qPCR Mix (ZF501; ZOMANBIO; Beijing, China). ACTB served as the internal control for normalization. The forward sequence and reverse sequence of the primers for CTSK were 5'-ACACCCACTGGGAGCTATG-3' and 5'-GACAGGGGTACTTTGAGTCCA-3', respectively, and the forward sequence and reverse sequence of the primers for ACTB were 5'-CATGTACGTTGCTATCCAGGC-3' and 5'-CTCCTTAATGTCACGCACGAT-3', respectively.

2.10 Western blotting

For protein analysis, the collected cell samples were disrupted using RIPA lysis buffer (Catalog No. R0020; Solarbio, Shanghai, China), ensuring thorough cellular breakdown for protein extraction. After cell lysis, protein concentrations were accurately determined with a BCA protein assay kit, allowing for the quantification necessary for further analysis. The proteins were then resolved on SDS–PAGE gels to achieve separation based on molecular weight. After electrophoretic separation, the proteins were carefully transferred onto PVDF membranes obtained from Millipore. The membranes were then incubated with a 5% solution of nonfat milk from Solarbio to block nonspecific binding sites. Primary antibodies directed against CTSK(rabbit polyclonal, 1:1000, A1782, ABclonal), GAPDH(rabbit polyclonal, 1:4000, A19056, ABclonal), N-cadherin (rabbit polyclonal, 1:1000, A21308, ABclonal), vimentin(rabbit polyclonal, 1:4000, A19607, ABclonal), slug(rabbit polyclonal, 1:1000, 9585T, CST) and snail (rabbit polyclonal, 1:1000, 3879T, CST) were applied to the membranes, which were then incubated overnight at a steady temperature of 4°C. Following primary antibody binding, the membranes were exposed to appropriate secondary antibodies, and the appropriate settings were established for detection. The detection phase employed the chemiluminescent method using the Western blotting Detection Kit (ECL; Catalog No. ED0015-A, Sparkjade), ensuring sensitive visualization of the protein bands.

2.11 Cell proliferation assay

After transfection, the cells were allowed to adapt for 48 hours before cell activities were assessed. Assessment was conducted using the CCK-8 Cell Proliferation Assay Kit (catalog no. C6005M; US Everbright; Silicon Valley, CA, USA), which strictly adhered to the manufacturer's instructions. Simultaneously, to evaluate the proliferative responses, the EdU Cell Proliferation Assay Kit (Catalog No. C6015M; US Everbright) was used, which provides a parallel quantitative measure of cell division and growth. For the colony formation assays, an initial seeding density of 1000 cells per well was maintained in six-well plates, and the cultures were incubated for a period ranging between one and two weeks to allow for sufficient colony development. At the conclusion of the incubation period, the colonies were fixed in 4% paraformaldehyde solution for 20 minutes to ensure optimal preservation. Then, colonies were stained with a 0.5% crystal violet solution for 20 minutes to enhance visual contrast for subsequent analysis.

2.12 Transwell assay

This study employed Transwell migration assays using 24-well plates with polycarbonate membranes that had an 8-μm pore size (Corning, USA). In these experiments, we filled each lower chamber with 500 μl of RPMI 1640 medium enriched with 10% fetal bovine serum to facilitate cellular growth and migration. In parallel, 200 μl of a serum-free cell suspension, prepared at a density of 1×10^6 cells/ml, was gently pipetted into the upper chamber of the setup. This configuration was maintained in an incubator set at the optimal growth conditions of 37°C and an atmosphere containing 5% CO_2 for a 24-hour period to allow for effective cell migration. After incubation, the cells within the Transwell chambers were fixed in a 5% glutaraldehyde solution to preserve their structure and morphology. Staining was then performed using 0.1% crystal violet dye, allowing the visualization and subsequent analysis of cell migration patterns.

2.13 Wound healing

In the described experiment, six-well plates were seeded at a density of 1×10^6 cells per well. Following an overnight incubation period, a deliberate wound was introduced into the confluent cell monolayer utilizing the tip of a 10-μl pipette. Subsequently, the induced scratch was visualized using a high-resolution microscope equipped with options for 10× magnification.

2.14 Statistical analysis

Statistical evaluations in this study were conducted utilizing R software, version 4.1.1. Data analysis was performed by using GraphPad Prism 9.0 (San Diego, CA, USA). The two-sided Student's t-test was used to compare unpaired data. The Cox hazard regression model was used for univariate analysis, P value < 0.05 was considered statistically significant.

The analyses included one-way and multifactorial Cox regression using 'survival' and 'survminer' packages available within R, respectively. The criterion for statistical significance was set such that a p value less than 0.05 indicated statistical significance.

3 Results

3.1 Construction of the coexpression network in THCA

In this detailed study, we harnessed the EMT gene set from the MSigDB website to conduct gene pathway assessments for 501 patients diagnosed with THCA using the ssGSEA algorithm. Patients were divided into two groups according to a median EMT score of 0.67. The group with scores above this median, termed the EMT-High group, demonstrated significantly elevated EMT scores that were greater than those in the EMT-Low group, with statistical analyses confirming a significant difference ($p < 2.2e-16$), as detailed in Figure 1A. Furthermore, to explore the gene expression profiles across these patients, WGCNA was employed to scrutinize the expression data of 14,564 genes collected from the 501 THCA samples. Through meticulous determination, a soft-thresholding power of 15 was established based on achieving a scale-free topology criterion with an R^2 value of 0.9, as depicted in Figure 1B. The analytical process led to the identification of ten distinct gene modules after setting the dissolution threshold (DissThres) to 0.2 to merge dynamic modules. Notably, cluster dendrogram analysis revealed that the pink module had the most substantial correlation with the EMT scores, with a Pearson correlation coefficient of 0.58 and a statistically significant p value of 0, as illustrated in Figure 1C. Given the focus of our research on the EMT phenomenon within the TCGA-THCA dataset, the green module was identified as a hub module. This module's pivotal role is highlighted in Figure 1D, underscoring its relevance in our ongoing analysis. To further refine our study, thresholds for GS and MM were set at greater than 0.5 and 0.7, respectively. This stringent criterion facilitated the identification and selection of 68 key genes that exhibited strong associations with EMT characteristics, paving the way for subsequent detailed investigations. These pivotal genes are shown in Figure 1E, setting the stage for future exploratory and confirmatory studies.

3.2 Functional analyses of EMT-related genes

The functional enrichment analysis conducted in this study revealed a significant concentration of GO terms associated with components of the extracellular matrix. Notably, these included terms related to collagen-enriched extracellular matrices, the structural constituents of such matrices, and the organization of encapsulating structures external to cells. The analysis also highlighted significant enrichment in terms related to the broader organization of extracellular structures and matrices themselves (Figure 1F). KEGG pathway analysis revealed the enrichment of specific pathways that play pivotal roles in cellular interactions and signaling mechanisms. The identified pathways included the PI3K-Akt signaling pathway, protein digestion and absorption, ECM-receptor interactions, and focal adhesion, which are all essential for cellular communication and adhesion processes (Figure 1G). The enrichment of these pathways suggested that genes associated with EMT may actively contribute to the malignant progression of THCA by enhancing the activation of these critical signaling pathways. This activation potentially facilitates the invasive and metastatic behavior of cancer cells, underlining the importance of these pathways in the context of cancer progression and the potential for targeted therapeutic interventions.

3.3 Construction of the EMT-based signature

In this analysis, a LASSO regression approach was utilized to scrutinize the prognostic potential of 68 genes, and a critical minimum value of 5 was determined (Figure 2A). This analysis identified five genes with significant characteristics related to EMT: C3ORF80, CTSK, FBLN2, PRELP, and SRPX2. These genes were then used to construct a robust EMT risk score model. The model was formulated as follows: EMT risk score = (CTSK * 0.286) + (C3ORF80 * 0.478) + (FBLN2 * -0.636) + (PRELP * -0.166) + (SRPX2 * 0.310). Using this predictive model, patients with THCA were stratified into two distinct risk categories based on the median risk score of the cohort. The categorization placed 84 patients in the high-risk group, which corresponded with a markedly increased mortality rate. In contrast, the classification identified 83 patients as belonging to the low-risk group, which was associated with significantly enhanced survival rates, as depicted in Figures 2C, D. The disparity in survival probabilities between these groups was starkly illustrated in the Kaplan–Meier survival plots (Figure 2B), indicating a significantly shorter survival duration for patients in the high-risk group than for those in the low-risk group. Furthermore, the reliability of the EMT risk score was evaluated using receiver operating characteristic (ROC) curve analysis, yielding areas under the curve (AUCs) for 1-year, 2-year, and 3-year survival predictions of 0.87, 0.87, and 0.81, respectively, for the TCGA-THCA cohort (Figure 2E). This analysis underscores the prognostic accuracy of the EMT risk score model in predicting patient outcomes. Additionally, a comparative analysis of gene expression within these risk groups revealed that CTSK and SRPX2 were expressed at higher levels in the high-risk group, whereas FBLN2 and PRELP showed reduced expression levels in the same group compared to the low-risk group (Figure 2F). This differential expression pattern further corroborates the link between these genes and the aggressive clinical behavior associated with higher EMT risk scores.

10.3389/fimmu.2024.1463258

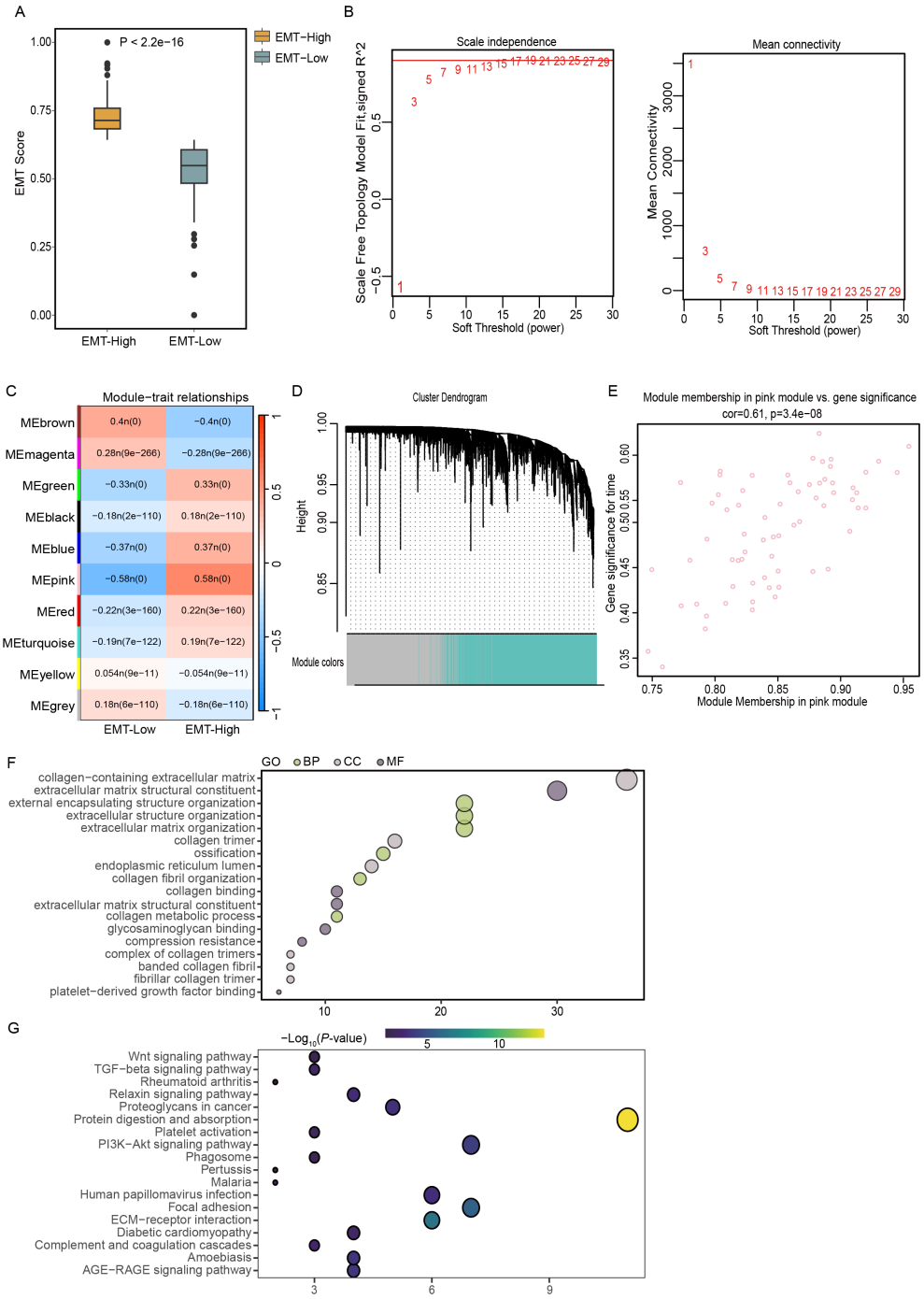


FIGURE 1
Elucidating EMT dynamics and key genetic players in THCA using TCGA data. **(A)** Classification of TCGA-THCA samples into EMT-High and EMT-Low groups using the ssGSEA algorithm. **(B)** Determination of the optimal soft-thresholding power at 3, illustrated using graphs depicting scale independence and mean connectivity for assessing scale-free network topology. **(C)** Correlation analysis between gene modules and EMT scores to identify relevant genetic interactions. **(D)** Construction of a coexpression network using WGCNA based on RNA-seq profiles from the TCGA-THCA dataset. **(E)** Scatter plot highlighting the pink module, where key genes with a GS greater than 0.5 and MM above 0.7 were identified, indicating significant topological overlap. **(F, G)** Functional enrichment analyses using GO and KEGG pathway analyses to explore the biological implications of genes within the EMT-based signature.

3.4 Validation of the training set and validation set

The dataset was divided into two comprehensive sections: approximately 80% were assigned to the training set, and the remaining 20% formed the validation set. Subsequent analyses of the training data indicated that the prognosis for patients identified as belonging to the high-risk group was significantly less favorable than that for those assigned to the low-risk group. This disparity was particularly evident in the elevated mortality rates observed

10.3389/fimmu.2024.1463258

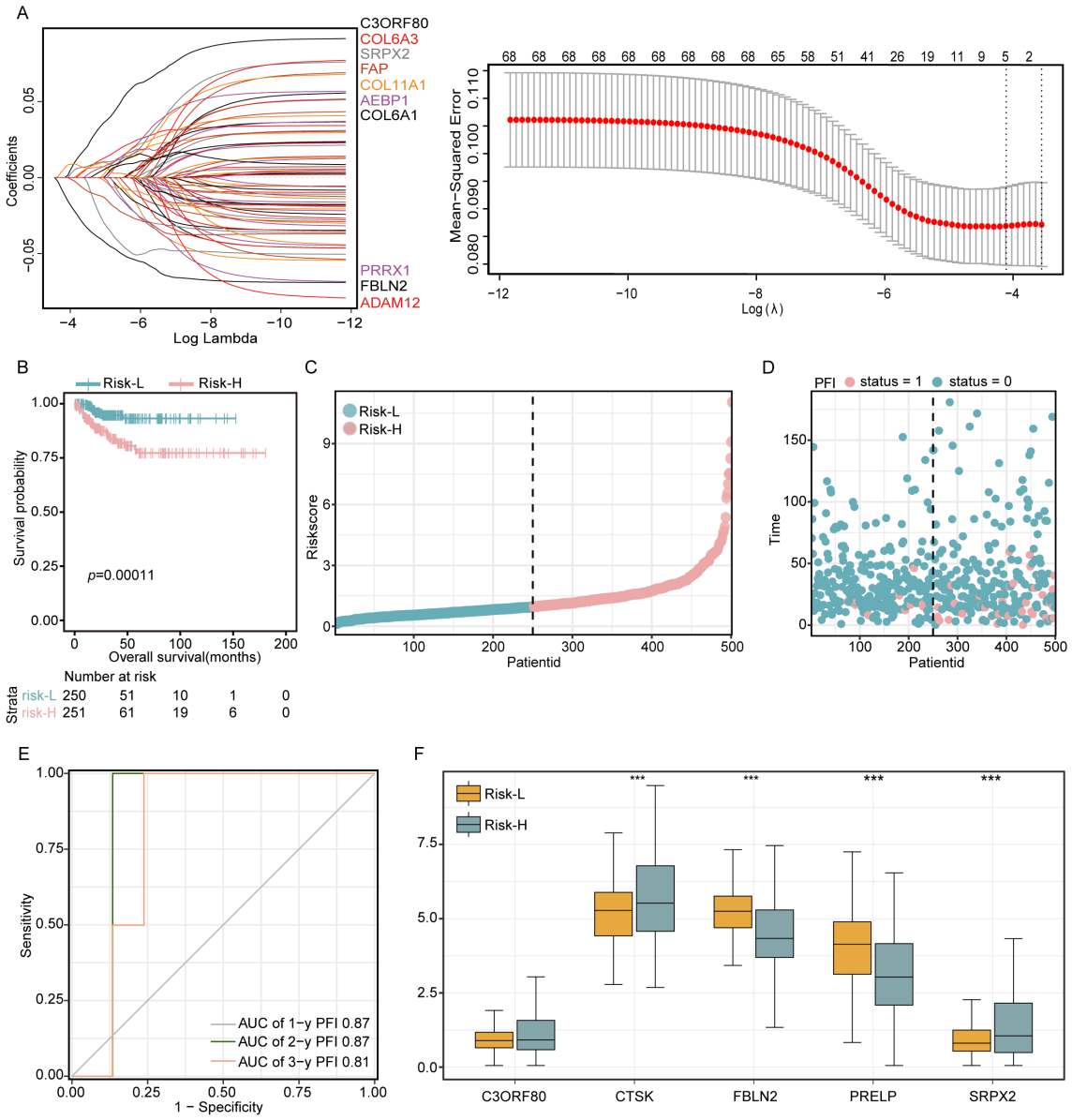


FIGURE 2
Development of a 68-gene prognostic signature based on differential expression analysis in two subtypes. **(A)** LASSO regression was applied to refine the gene selection for the prognostic model. **(B)** Kaplan–Meier survival curves delineating the outcomes of 509 patients stratified into high-risk and low-risk groups according to their EMT scores. **(C, D)** Presentation of risk curves illustrating the distribution of prognostic scores along with patient survival time and status. **(E)** Time-dependent ROC curves evaluating the predictive accuracy of survival probabilities based on DEGs. **(F)** Comparative analysis of the expression levels of five critical DEGs between patients in the low-risk and high-risk groups. Significance levels are denoted as ***P<0.001.

among individuals in the high-risk category (as detailed in Figure 3A). Similarly, evaluation of the validation set demonstrated consistent results with both the training set and the entire dataset, confirming the reproducibility and robustness of the findings across different subsets of data (Figure 3B). Based on the previously described LASSO linear regression, after removing redundant genes and constructing a risk model, we ultimately screened five DEGs (SRPX2, PRELP, FBLN2, CTSK, and C3ORF80). Of these, C3ORF80 expression showed a significant positive correlation with prognosis, whereas the expression of the other four genes exhibited no significant correlation with prognosis (Figure 3C). In addition, the chromosome circle plot illustrated the chromosomal locations of SRPX2, PRELP, FBLN2, CTSK and C3ORF80 (Figure 3D). Additionally, Spearman correlation analysis revealed significant negative correlations between the risk score and EDRNB and VEGFA, whereas positive correlations were observed with VTCN1, CD276, and TNFRSF4 (Figure 3E).

3.5 Construction of the nomogram and mutation analysis

Among the five DEGs screened, only FBLN2 exhibited a significant association with the hazard ratio (p=0.001), which was

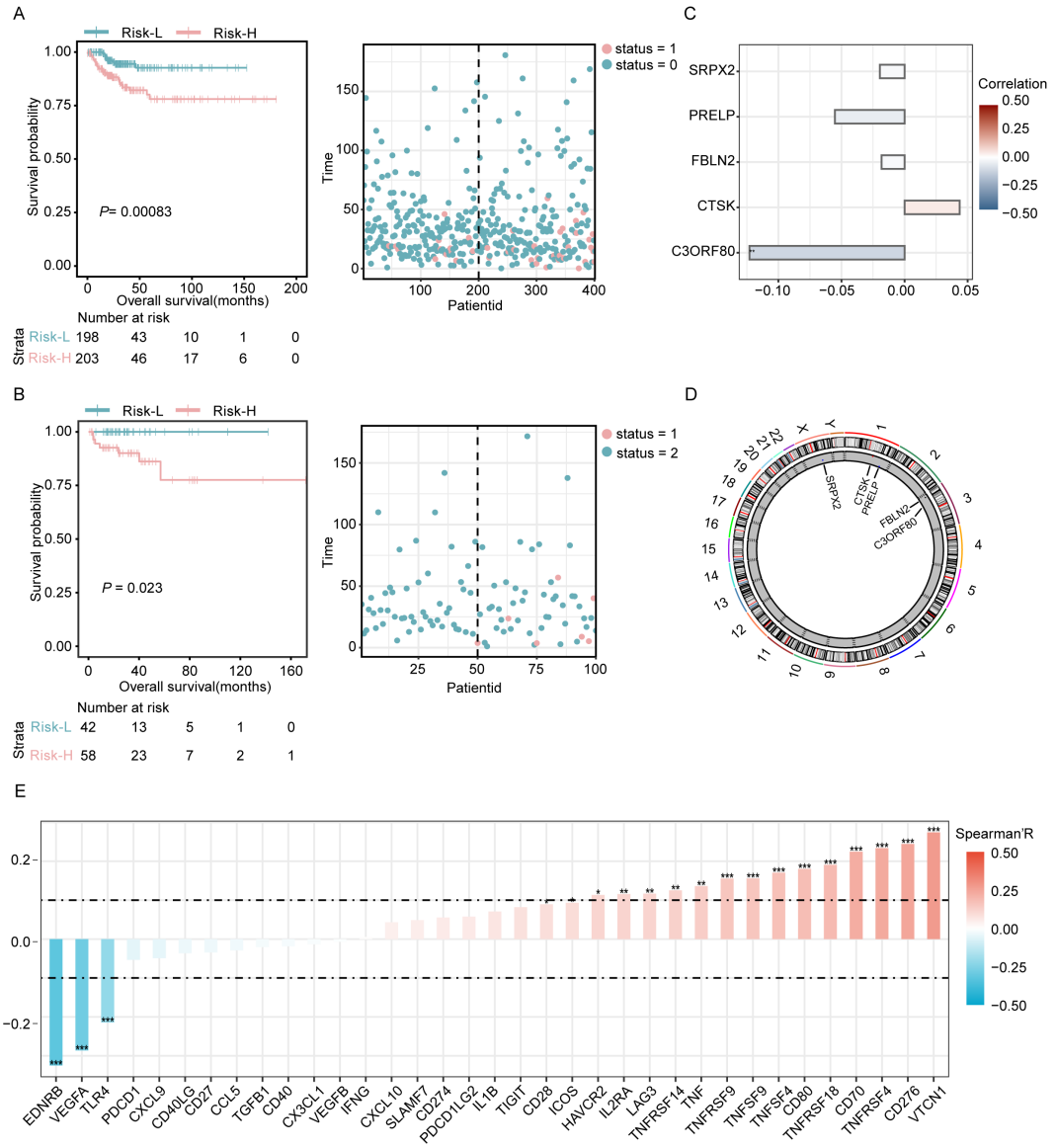


FIGURE 3
Prognostic evaluation in training and validation sets with examination of key genes. **(A)** Visualization of the risk curves showing the distribution of prognostic scores and survival statuses within the training cohort. **(B)** Risk curves depicting the prognostic scores and survival statuses across the validation cohort. **(C)** Analysis of the correlation between key genes and patient prognosis. **(D)** A circular chromosome plot illustrating the genomic positions of key genes relevant to the study. **(E)** Evaluation of the associations between the risk score model and 43 immune checkpoint genes conducted using Spearman's correlation coefficient. Significance levels are denoted as *P<0.05, **P<0.01, and ***P<0.001.

lower in patients in whom FBLN2 was highly expressed (Figure 4A). The interrelationships among these 5 genes are depicted in Figure 4B, where FBLN2 displayed the strongest correlation with PRELP. To evaluate the recurrence risk of individual THCA patients, a nomogram incorporating four predictive factors, namely, gender, age, risk score, and cancer stage, was developed. This tool indicates that for THCA patients with a high genetic risk score (GRS) and N0 stage disease, the probabilities of recurrence at 12, 36, and 60 months are estimated to be 0.127, 0.403, and 0.426, respectively (Figure 4C). Additionally, analysis of single nucleotide variants revealed that missense mutations were the predominant type of DNA mutation found within the five DEGs. Among these, single nucleotide polymorphisms (SNPs) are the most frequently occurring mutations, with transitions from cytosine to thymine representing the most common type of base substitution observed. In addition, BRAF, NRAS and HRAS were the most commonly mutated genes in THCA, and most of their mutations were missense mutations (Figure 4D).

3.6 Evaluation of the immune microenvironment

Using the ssGSEA algorithm to analyze the composition of tumor-infiltrating immune cells, we observed distinct profiles in different risk groups of tumor patients. This study revealed

10.3389/fimmu.2024.1463258

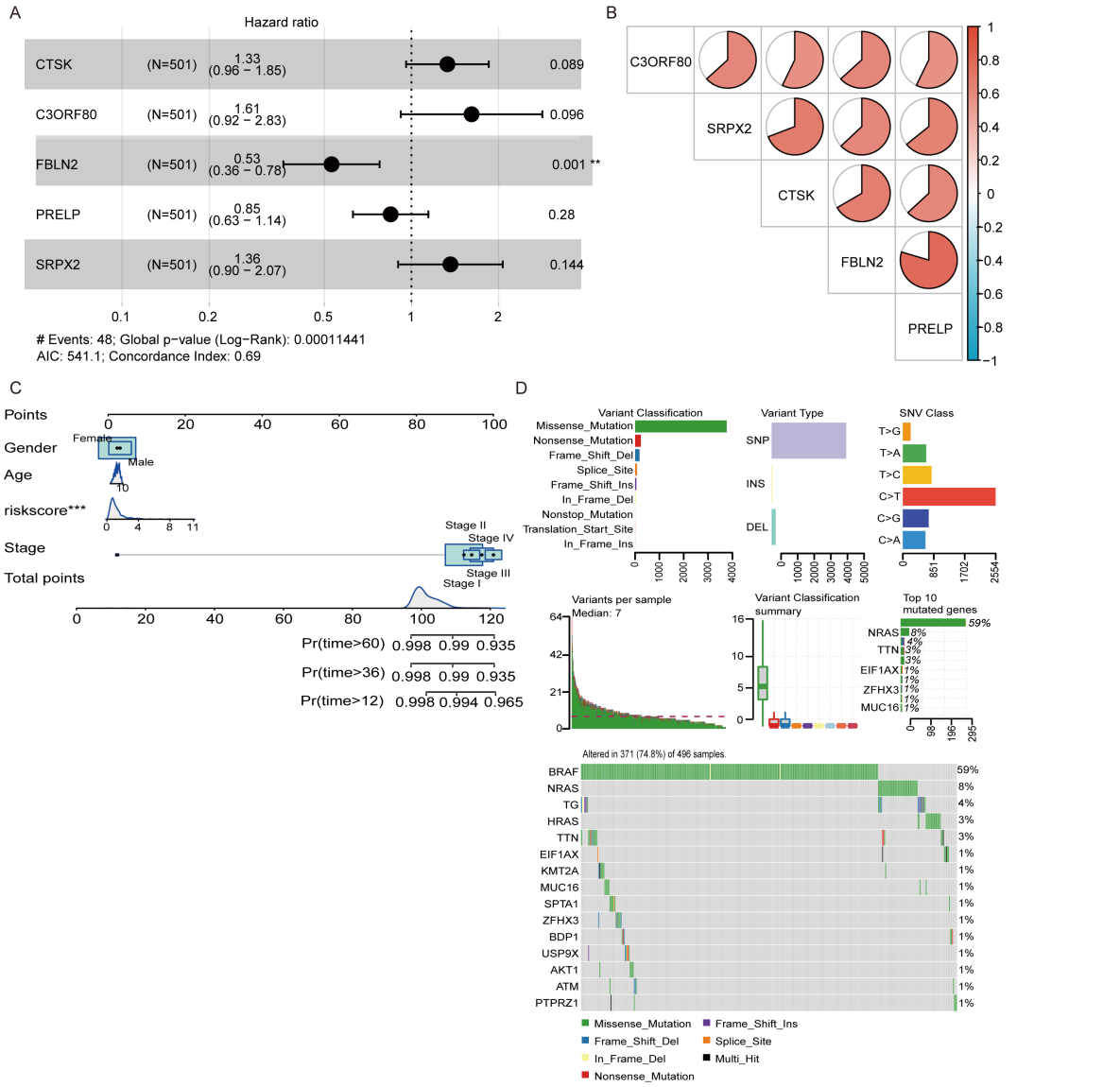


FIGURE 4
Development of a nomogram and analysis of genetic mutations. **(A)** Multivariate analysis was conducted to confirm the independent prognostic factors influencing patient outcomes. **(B)** Pie chart illustrating the interrelationships among the genes included in the model. **(C)** A nomogram was constructed that incorporates sex, age, risk score, and cancer stage to predict the risk of recurrence at 12, 36, and 60 months. **(D)** Waterfall plot displaying the spectrum of single nucleotide variant (SNV) mutations in the genes modeled, highlighting the genetic alterations within the study cohort. **P<0.01, ***P<0.001.

increased levels of various immune cells, including CD56dim natural killer cells, gamma delta T cells, CD56bright natural killer cells, immature dendritic cells, macrophages, MDSCs, activated dendritic cells, regulatory T cells, and type 17 T helper cells, in the high-risk group, all of which demonstrated statistically significant differences ($p < 0.05$). Conversely, the low-risk group exhibited significantly greater numbers of activated B cells, eosinophils, and type 2 T helper cells ($p < 0.05$), as shown in Figure 5A. Higher levels of macrophage infiltration, MDSC infiltration, and regulatory T-cells (Tregs) in the high-risk group suggest that the high-risk group may have more significant signs of immune evasion. The correlation scatter plot illustrates that the EMT risk score is positively associated with the infiltration of certain immune cells, including macrophages, activated dendritic cells, and gamma delta T cells. Conversely, this score showed a negative correlation with plasmacytoid dendritic cells, activated B cells, and monocytes, as depicted in Figure 5B. This analysis highlights the differential relationships between the EMT risk score and specific immune cell types, suggesting varying influences of these cells on EMT progression. In this analysis, we also selected 23 immune cells expressed in the TCGA cohort for analysis and calculated the correlation coefficients between the expression levels of the five genes and their infiltration levels. The results of the thermographic analysis are shown in Figure 5C. Among them, CTSK showed a significant positive correlation with regulatory T cells, macrophages, type I T helper cells, and

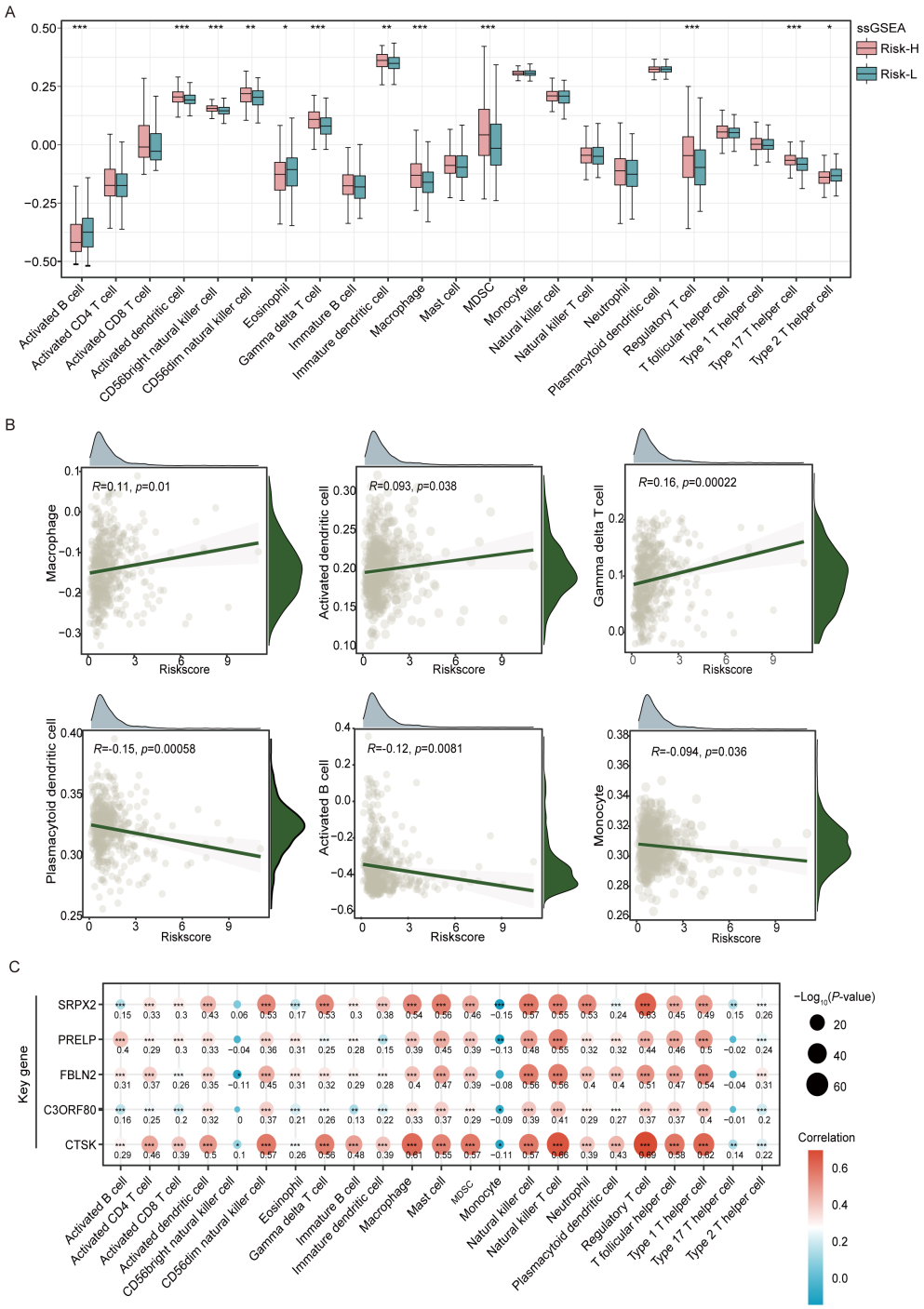


FIGURE 5
Analysis of tumor immune microenvironment variations in the high-risk and low-risk groups of the TCGA-THCA cohort. **(A)** Box plots reveal the variation in the levels of 23 different immune cell types between groups classified as high- and low-risk, as established by ssGSEA. **(B)** A scatterplot illustrates the correlation between the risk score and the distribution of different immune cell types within the tumor microenvironment. **(C)** The heatmap visualizes correlation coefficients linking crucial genes with immune cells, where red dots represent positive correlations, blue dots signify negative correlations, and the star symbol (*) highlights statistically significant findings. *P<0.05, **P<0.01, ***P<0.001.

natural killer cell (all R>0.6 and all p<0.001). SRPX2 also showed a positive correlation with regulatory T cells (R>0.6 and all p<0.001).

Using the xCell algorithm, our research investigated the correlation between tumor-infiltrating immune cells and the EMT risk score in THCA patients. This study revealed a robust positive correlation between the risk score and NK T cells, with a correlation coefficient exceeding 0.3 and a p value less than 0.001, confirming statistical significance (Figure 6A). Conversely, the most substantial negative correlation was observed with HSCs, where the correlation coefficient was less than -0.4, and the p value was less than 0.001.

Moreover, boxplot analyses highlighted that immune infiltration levels varied significantly between risk groups. Individuals categorized

10.3389/fimmu.2024.1463258

within the low-risk group displayed elevated levels of MSCs, microvascular endothelial cells, myocytes, and HSCs relative to their counterparts in the high-risk group. Conversely, the high-risk group was characterized by increased quantities of monocytes, NK T cells, sebocytes, Tregs, immature dendritic cells (iDCs), and macrophages. Additionally, high-risk patients had a greater overall immune score, whereas low-risk patients had an elevated stromal score, indicating a differential stromal contribution to the tumor microenvironment (Figure 6B). The stacked diagrams provided a detailed view of the immune cell infiltration landscape across individual patients, revealing notable differences in the proportions of infiltrating immune cell subsets among them (Figure 6C).

In the extended analysis of the THCA, we applied the MCPcounter algorithm to determine the associations between the

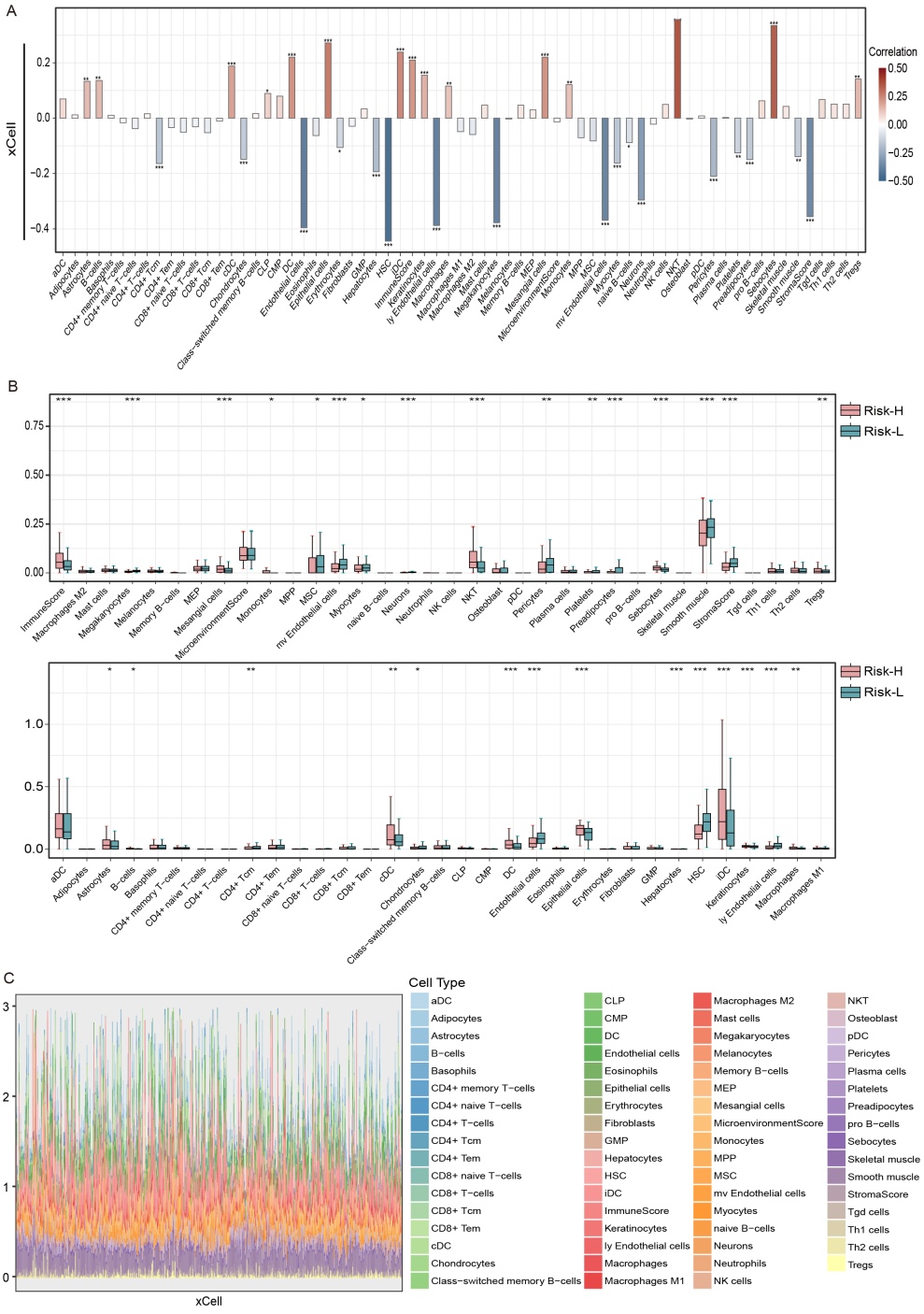


FIGURE 6
The correlation between risk score and immune cells types in THCA using the xCell algorithm and the TCGA dataset. **(A)** Bar graph of the risk score based on the xCell immune infiltration algorithm, **(B)** correlation boxplots of the risk score and 23 xCell immune cells, and **(C)** immune cell stacking plots of xCells from 501 thyroid cancer patients. ns = non-significant, *P<0.05, **P<0.01, ***P<0.001.

concentrations of immune cells infiltrating tumors and risk scores. Detailed boxplot assessments demonstrated significantly greater numbers of B lineage cells, endothelial cells, neutrophils, and NK cells in the high-risk group than in the low-risk group. These findings suggest a pronounced immunological signature that correlates with an increased risk of tumor development (Figure 7A). Additionally, using stacked diagrams, we observed that the proportions of infiltrated immune cell subsets varied significantly across patients, highlighting the diverse immune landscape present within the patient cohort (Figure 7B). The heatmap showed that all five EMT-related genes were positively associated with fibroblasts, whereas SRPX2 and CTSK were negatively associated with endothelial cells and neutrophils, as shown in Figures 7C, D.

3.7 CTSK potentially play an important oncogenic role in THCA

Further investigations have been conducted to explore the correlation between CTSK expression levels in overall THCA patient and individual patient outcomes. According to previously published data, the CTSK scores in the group at high risk were much greater than those in the group at low risk (Figure 2F). Our recent analysis builds on these findings, demonstrating that increased CTSK expression is significantly correlated with a decrease in patient prognosis (p=0.015, Figure 8A).

To investigate the impact of CTSK suppression on cellular dynamics, functional experiments were conducted using the KTC-1 and Cal-62 thyroid carcinoma cell lines. Initially, the effectiveness of CTSK knockdown was validated using RT–qPCR and Western blot analyses, demonstrating a significant reduction in CTSK expression ($p < 0.001$, Figures 8B, C). Subsequently, a series of assays were performed to assess cellular functions after knockdown. After CTSK elimination, the CCK-8 assay, colony formation assay, and EdU assay all revealed significant reductions in cell activity, colony development, and proliferation. This reduction was statistically significant ($p < 0.01$, $p < 0.001$; Figures 8D–F). Furthermore, Transwell and wound healing experiments demonstrated that cell motility and migration were substantially impaired in response to CTSK knockdown ($p < 0.01$, $p < 0.001$; Figures 9A, B). Following CTSK knockdown, Western blot analysis revealed a significant decrease in the expression of N-cadherin, Vimentin, Slug, and Snail (Figure 9C).

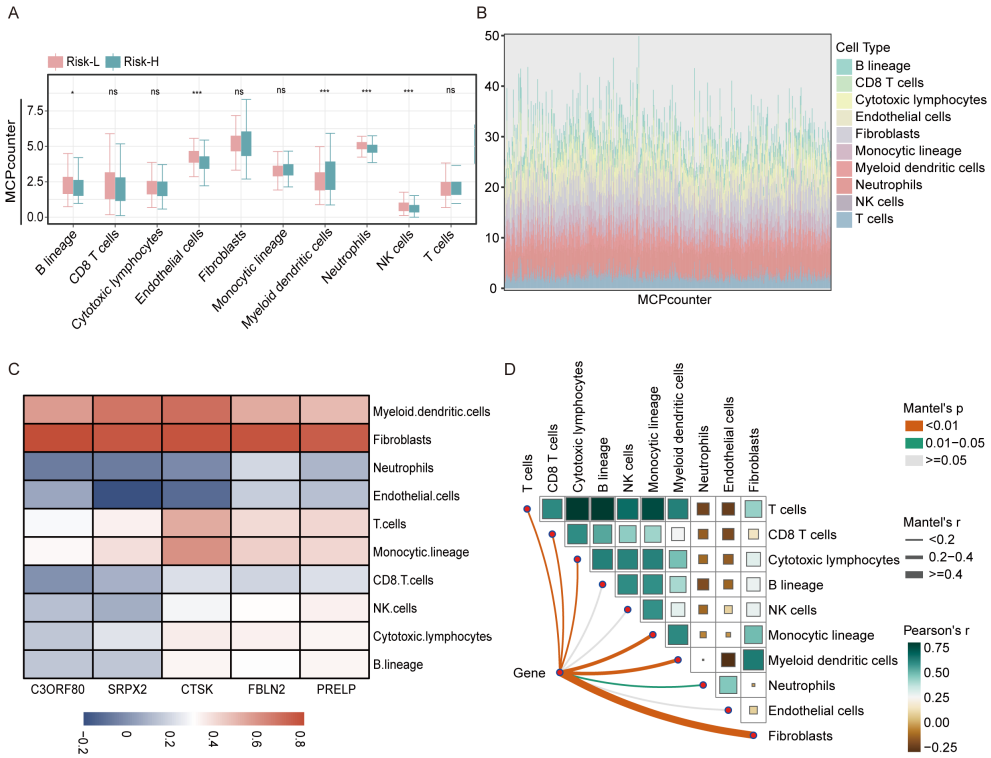


FIGURE 7
Analysis of the association between the risk score and immune cell types in the THCA using the MCPcounter approach. **(A)** A boxplot illustrates the variation in immune cell infiltration among the high-risk and low-risk categories as determined using the MCPcounter algorithm within the TCGA dataset. **(B)** Stacked bar chart illustrating the distribution of immune cells across 501 thyroid cancer patients as analyzed using MCPcounter. **(C)** Heatmap displaying the correlations between five key genes and the levels of various immune cell types, as quantified using MCPcounter. **(D)** LINKET map showing the relationships between the abundances of immune cells linked to ten specific immune cell genes and model genes within the immune microenvironment. ns = non-significant, *P<0.05, ***P<0.001.

10.3389/fimmu.2024.1463258

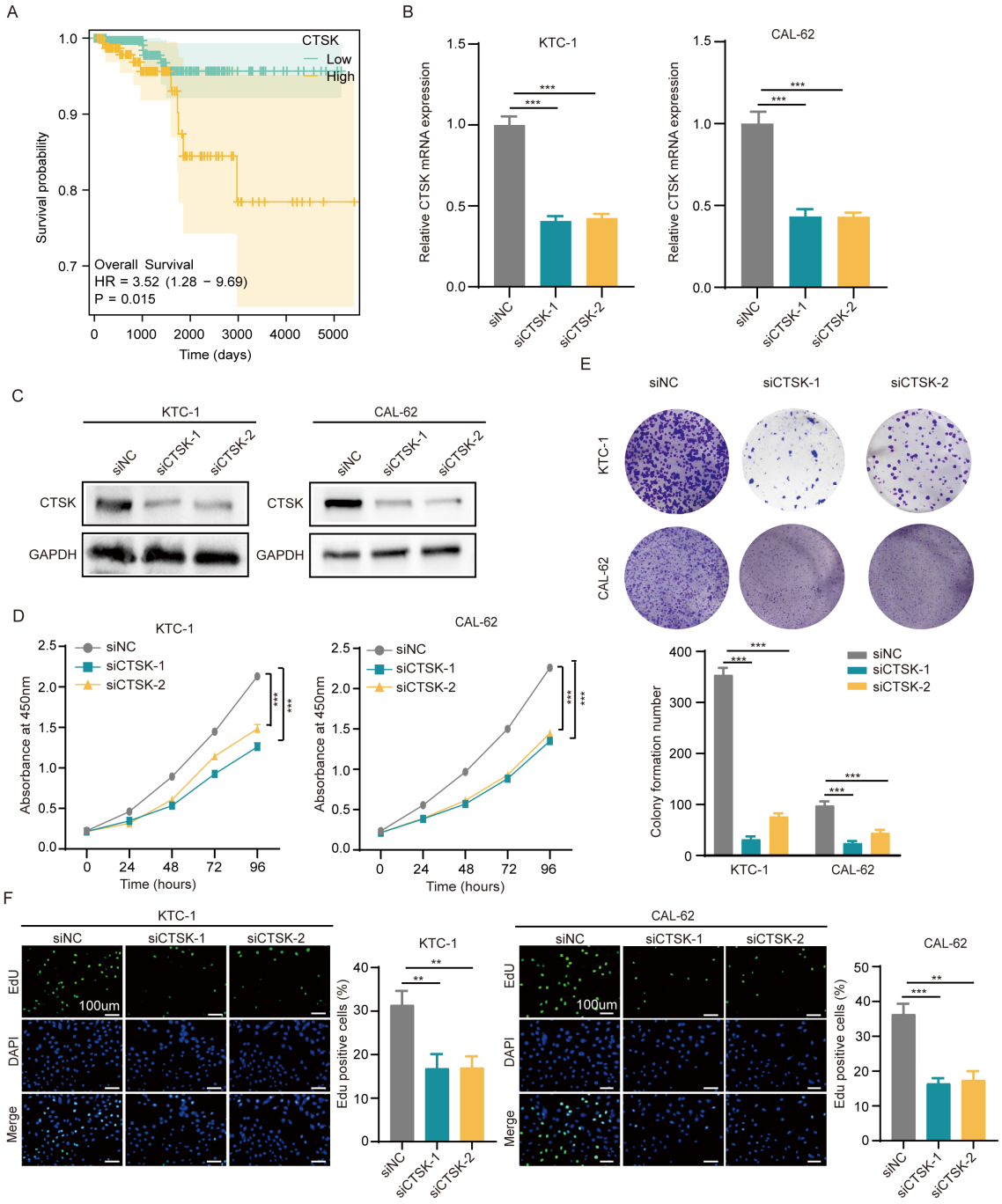


FIGURE 8
CTSK knockdown inhibits cell proliferation and metastasis. **(A)** Correlation between the CTSK expression level and overall survival of THCA patients. **(B, C)** The knockdown efficiency of CTSK at the gene level was verified using RT–qPCR and western blotting. **(D)** CCK-8 assays revealed that KTC-1 and cal-62 cells with CTSK knockdown exhibited significantly weaker cell activity than siNC cells. **(E)** The colony formation assay demonstrated that the colony formation ability of KTC-1 and cal-62 cells in which CTSK was knocked down was substantially lower than that of cells from the siNC control group. **(F)** The results of the EdU incorporation assay showed that the proliferation of CTSK-knockdown KTC-1 and CTSK-knockdown cal-62 cells was significantly lower than that of siNC-transfected cells. **P<0.01, ***P<0.001. Scale bar =100 μm.

4 Discussion

The majority of differentiated thyroid cancers exhibit a favorable prognosis. For these patients, the primary treatment modality is surgery, followed by subsequent radioactive iodine ablation (iodine-131) or thyroxine therapy. However, given that some patients with THCA are prone to tumor metastasis and recurrence or even progression to fatal THCA, systemic treatment is needed, and targeted therapies are preferred (23). The EMT is strongly associated with poor prognosis in THCA patients, and the EMT properties of THCA make therapy targeting EMT-related genes an attractive therapeutic option (24). Remarkably, the

10.3389/fimmu.2024.1463258

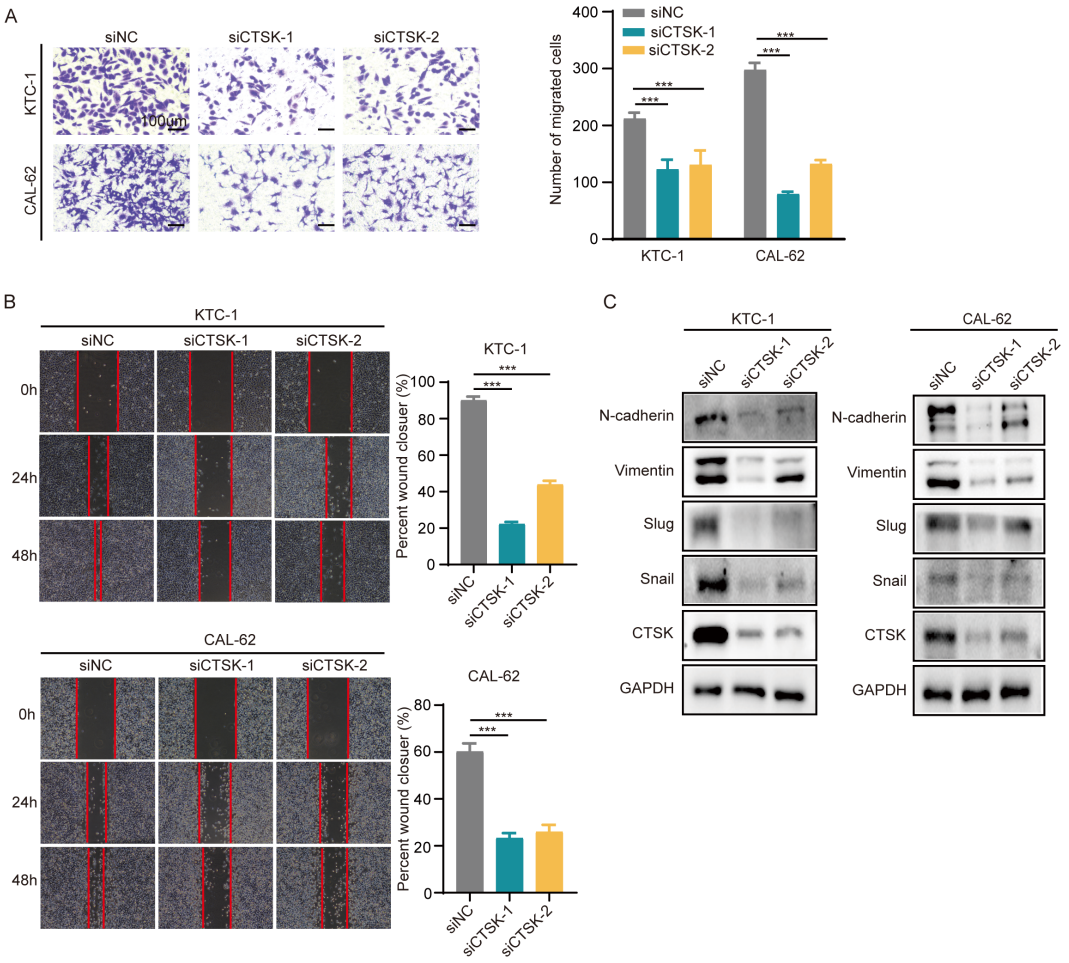


FIGURE 9
CTSK knockdown inhibits cell motility and migration. **(A)** A Transwell assay revealed reduced cellular mobility in CTSK-knockdown KTC-1 and cal-62 cells compared to control cells, demonstrating a significant reduction in the ability of these cells to traverse membrane pores. **(B)** Wound healing assays at 24 hours postwound creation revealed decreased motility in CTSK-knockdown KTC-1 and CTSK-knockdown cal-62 cells compared with that in the siNC group, as indicated by decreased closure rates. **(C)** Western blot analysis showing decreased levels of EMT markers, including Vimentin, N-cadherin, Snail, and Slug, in KTC-1 and CAL-62 cells following CTSK knockdown, with GAPDH serving as the loading control. The reduction in these proteins underscores significant suppression of EMT progression. ***P<0.001. Scale bar =100 μm.

potential and functions of EMT-related genes in THCA remain largely unexplored.

In recent years, precision medicine has revolutionized cancer treatment by aiming to personalize disease prevention and treatment strategies through the analysis of individual variations in genomics, the external environment, and lifestyle. An increasing number of researchers have already established subgroups based on the molecular profiles of patients, representing different phenotypes, prognoses and treatment responses. In the context of precision medicine, recent studies have illustrated the importance of gene expression profiling in various cancers. For example, in acute myeloid leukemia (AML), profiling based on the expression of genes linked to ferroptosis can identify a subset of patients with a poorer prognosis who may benefit from ferroptosis-inducing treatments (25). Patients with colorectal cancer (CRC) are stratified into high-risk and low-risk groups using patterns of autophagy-related gene expression, and this information facilitates decision making regarding more aggressive treatments (26). Moreover, in gastric cancer (GC), categorizing patients into subtypes based on RNA n6-methyladenosine-related regulator expression revealed that those patients in certain high-risk subtypes demonstrate significant resistance to immunotherapy (27).

During this investigation, we identified two distinct expression profiles linked to the EMT, designated as the EMT-high and EMT-low categories. These groups exhibited significant differences in terms of prognosis, with the EMT-high group demonstrating a poorer prognosis than the EMT-low group. ssGSEA showed more macrophage infiltration, MDSC infiltration and regulatory T cell (Treg) expression in the EMT-high group compared to the EMT-low group, all three of which modulate the immune response by inhibiting the activity of effector T cells and other immune cells, thereby suppressing the anti-tumor immune response and promoting tumor growth. The xCell algorithm revealed that Treg levels were generally greater in patients in the high-risk subgroup than in those in the low-risk subgroup, indicating a potential association between the EMT-high subgroup and immune evasion through Treg activation.

Numerous studies have reported that oncogenes induce malignant progression of tumors by activating both Treg cells and the EMT and that Treg cells can also induce the EMT in tumor cells (27, 28). The findings from this study indicate that Treg cells significantly influence THCA progression.

Based on the DEGs from the two groups, we identified five genes (SRPX2, PRELP, FBLN2, CTSK and C3ORF80) for the construction of prognostic models using one-way and LASSO Cox analyses. The five EMT-related genes identified here offer significant potential for clinical application, particularly in personalized medicine for THCA. These EMT biomarkers can categorize patients into distinct risk groups based on their EMT signatures. Patients with higher EMT scores, associated with elevated CTSK and SRPX2 expression, tend to have a worse prognosis. This stratification provides clinicians with valuable information on disease progression risk, enabling more intensive monitoring for high-risk individuals. For example, such patients could be prioritized for frequent imaging and biomarker assessments to ensure early detection of recurrence or metastasis. Integrating EMT-related biomarkers into clinical practice could significantly enhance precision medicine approaches. By considering the EMT signature in treatment planning, oncologists can customize therapies based on the tumor's molecular characteristics. Additionally, the EMT signature could identify patients who may be less responsive to conventional treatments like radioactive iodine, guiding them toward alternative therapies.Moreover, these EMT-related genes could be developed into a biomarker panel for early detection and regular screening of THCA patients. Detecting elevated levels of these genes in blood samples or biopsy tissues could help identify patients at higher risk of disease progression or recurrence before clinical symptoms arise. This early detection could improve survival outcomes by enabling prompt interventions.

Research has indicated an association between CTSK expression and the malignant advancement of various tumors. In prostate cancer, molecules downstream of CTSK act as control elements that regulate the expression of EMT-related genes and promote PC cell metastasis and hyperproliferation (19). CTSK has emerged as a crucial mediator linking gut microbiota dysbiosis to CRC metastasis, thereby contributing significantly to the invasive phenotype of CRC cells both *in vitro* and *in vivo* (29). Research on hepatocellular carcinoma (HCC) revealed that CTSK significantly influences cell proliferation. This action is accomplished through its interaction with the SIAH1/protein kinase B (AKT) signaling pathway, where CTSK enhances SIAH1 protein ubiquitination, thereby promoting HCC cell growth and proliferation (17). Although extensive research has been conducted on the biological functions of CTSK in various tumors, limited knowledge exists regarding its involvement in the biological processes of THCA. This study demonstrates that CTSK is linked to poor prognosis in thyroid cancer (THCA) and actively promotes the proliferation and migration of THCA cells. Additionally, it increases the expression of key epithelial-mesenchymal transition (EMT) markers, including N-cadherin, vimentin, slug, and snail, as shown by *in vitro* experiments. CTSK, a critical factor in extracellular matrix degradation and immune modulation, emerges as a promising therapeutic target in THCA. Targeting CTSK, especially in patients with high CTSK expression, may improve outcomes for those in the high-risk EMT group. Additionally, the relationship between EMT signatures and immune cell infiltration, particularly macrophages, myeloid-derived suppressor cells (MDSCs), and regulatory T cells, opens pathways for combining these biomarkers with immunotherapy. High-risk patients with increased immunosuppressive cell infiltration might benefit from treatments that reactivate the immune system, such as immune checkpoint inhibitors. Furthermore, EMT profiles could serve as predictive biomarkers for selecting suitable candidates for immunotherapy. We can also explore whether the combination of CTSK-targeted therapy with immunotherapy or chemotherapy can further improve the therapeutic efficacy of THCA, which can provide a basis for clinical personalized treatment.

Although we obtained the above analyses in this study and some of the results have been validated by *in vitro* experiments, there are still some shortcomings in this study. Firstly, there are inherent limitations of the data in the TCGA database, for example, the number of samples in the TCGA-THCA dataset is relatively small, which may lead to insufficient efficacy of statistical analyses to detect biomarkers or gene variants with small effect sizes. In addition, although the TCGA database provides a wealth of transcriptomic data, these data originate from multiple technology platforms, and technical differences between these platforms may also lead to inconsistencies in the data, as well as increasing the complexity of data integration and data analysis. Finally, we have only validated our analyses by *in vitro* cytological experiments and have not yet completed *in vivo* experiments; in the future, further refinement of the *in vivo* experiments as well as exploring the role of CTSK in immune cell infiltration will be the main focus of our research.

5 Conclusions

In conclusion, we identified and validated the key gene CTSK, which is closely related to the EMT in THCA, and we concluded that CTSK could serve as an important biomarker to assist in the diagnosis of THCA.

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

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used. Ethical approval was not required for the studies on animals in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

Author contributions

SW: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Software, Supervision, Writing – original draft, Writing – review & editing, Validation, Visualization. YL: Conceptualization, Data curation, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. YZ: Conceptualization, Data curation, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. XR: Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. MT: Formal Analysis, Project administration, Resources, Supervision, Validation, Writing – review & editing. WL: Methodology, Resources, Software, Validation, Writing – review & editing. CL: Methodology, Resources, Validation, Writing – review & editing. HC: Methodology, Resources, Software, Writing – review & editing. HL: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Project administration, Resources, Writing – review & editing. YW: Funding acquisition, Project administration, Resources, Software, Writing – original draft, Writing – review & editing, Conceptualization, Data curation, Formal Analysis.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was supported by the Startup Fund for Scientific Research, Fujian Medical University (Grant number: 2020QH1228), the Joint Funds for the innovation of science and Technology, Fujian province(Grant number: 2019Y91010062), the Natural Science Foundation of Fujian Province(Grant number: 2021J01446,2024J011102). Scientific Research Foundation of Fujian Cancer Hospital (grant number:2023YN21).

Acknowledgments

We sincerely thank the TCGA database, the molecular signature database, and the authors who uploaded the original data.

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.

Publisher's note

References

1. Pishkari S, Paryan M, Hashemi M, Baldini E, Mohammadi-Yeganeh S. The role of microRNAs in different types of thyroid carcinoma: a comprehensive analysis to find new miRNA supplementary therapies. *J Endocrinol Invest*. (2018) 41:269–83. doi: 10.1007/s40618-017-0735-6

2. Lim H, Devesa SS, Sosa JA, Check D, Kitahara CM. Trends in thyroid cancer incidence and mortality in the United States, 1974-2013. *JAMA*. (2017) 317:1338–48. doi: 10.1001/jama.2017.2719

3. 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

4. Berdelou A, Lamartina L, Klain M, Leboulleux S, Schlumberger M. Treatment of refractory thyroid cancer. *Endocr Relat Cancer*. (2018) 25:R209–R23. doi: 10.1530/ERC-17-0542

5. Liu Q, Sun W, Zhang H. Roles and new insights of macrophages in the tumor microenvironment of thyroid cancer. *Front Pharmacol*. (2022) 13:875384. doi: 10.3389/fphar.2022.875384

6. Puerto M, Borson-Chazot F, Tabarin A. Updates on therapy for medullary thyroid cancer in 2021. *Ann Endocrinol*. 83(2):114–8. doi: 10.1016/j.ando.2021.12.002

7. Brabletz T, Kalluri R, Nieto MA, Weinberg RA. EMT in cancer. *Nat Rev Cancer*. (2018) 18:128–34. doi: 10.1038/nrc.2017.118

8. Huang Y, Hong W, Wei X. The molecular mechanisms and therapeutic strategies of EMT in tumor progression and metastasis. *J Hematol Oncol*. (2022) 15:129. doi: 10.1186/s13045-022-01347-8

9. Pastushenko I, Blanpain C. EMT transition states during tumor progression and metastasis. *Trends Cell Biol*. (2019) 29:212–26. doi: 10.1016/j.tcb.2018.12.001

10. Hardin H, Zhang R, Helein H, Buehler D, Guo Z, Lloyd RV. The evolving concept of cancer stem-like cells in thyroid cancer and other solid tumors. *Lab Invest*. (2017) 97:1142–51. doi: 10.1038/labinvest.2017.41

11. Mardente S, Aventaggiato M, Splendiani E, Mari E, Zicari A, Catanzaro G, et al. Extra-cellular vesicles derived from thyroid cancer cells promote the epithelial to mesenchymal transition (EMT) and the transfer of Malignant phenotypes through immune mediated mechanisms. *Int J Mol Sci*. (2023) 24:2754. doi: 10.3390/ijms24032754

12. Dou A, Fang J. Heterogeneous myeloid cells in tumors. *Cancers*. (2021) 13:3772. doi: 10.3390/cancers13153772

13. Peña-Romero AC, Orenes-Piñero E. Dual effect of immune cells within tumour microenvironment: pro-and anti-tumour effects and their triggers. *Cancers*. (2022) 14:1681. doi: 10.3390/cancers14071681

14. Yin H, Tang Y, Guo Y, Wen S. Immune microenvironment of thyroid cancer. *J Cancer*. (2020) 11:4884. doi: 10.7150/jca.44506

15. Meng X, Huang Z, Inoue A, Wang H, Wan Y, Yue X, et al. Cathepsin K activity controls cachexia-induced muscle atrophy via the modulation of IRS1 ubiquitination. *J Cachexia Sarcopenia Muscle*. (2022) 13:1197–209. doi: 10.1002/jcsm.12919

16. Ding Y, Li Z, Wang H, Wang Q, Jiang H, Yu Z, et al. CTSK and PLAU as prognostic biomarker and related to immune infiltration in pancreatic cancer: Evidence from bioinformatics analysis and qPCR. *Int J Genomics*. (2023) 2023:3914687. doi: 10.1155/2023/3914687

17. Zhang C, Liu Z, Wang X, Zhang B, Cui L, Hu Q, et al. Cathepsin K promotes the proliferation of hepatocellular carcinoma cells through induction of SIAH1 ubiquitination and degradation. *Iscience*. (2023) 26(6). doi: 10.1016/j.isci.2023.106852

18. Leusink FK, Koudounarakis E, Frank MH, Koole R, van Diest PJ, Willems SM. Cathepsin K associates with lymph node metastasis and poor prognosis in oral squamous cell carcinoma. *BMC Cancer*. (2018) 18:1–9. doi: 10.1186/s12885-018-4315-8

19. Wu N, Wang Y, Wang K, Zhong B, Liao Y, Liang J, et al. Cathepsin K regulates the tumor growth and metastasis by IL-17/CTSK/EMT axis and mediates M2

macrophage polarization in castration-resistant prostate cancer. *Cell Death Dis*. (2022) 13:813. doi: 10.1038/s41419-022-05215-8

20. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS: J Integr Biol*. (2012) 16:284–7. doi: 10.1089/omi.2011.0118

21. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. *Nat Genet*. (2006) 38:500–1. doi: 10.1038/ng0506-500

22. Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. (2018) 28:1747–56. doi: 10.1101/gr.239244

23. Haddad RI, Bischoff L, Ball D, Bernet V, Blomain E, Busaidy NL, et al. Thyroid carcinoma, version 2.2022, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*. (2022) 20:925–51. doi: 10.6004/jnccn.2022.0040

24. Mitchell B, Dhingra JK, Mahalingam M. BRAF and epithelial-mesenchymal transition: lessons from papillary thyroid carcinoma and primary cutaneous melanoma. *Adv Anat Pathol*. (2016) 23:244–71. doi: 10.1097/PAP.0000000000000113

25. Cui Z, Fu Y, Yang Z, Gao Z, Feng H, Zhou M, et al. Comprehensive analysis of a ferroptosis pattern and associated prognostic signature in acute myeloid leukemia. *Front Pharmacol*. (2022) 13:866325. doi: 10.3389/fphar.2022.866325

26. Zhou Z, Mo S, Dai W, Ying Z, Zhang L, Xiang W, et al. Development and validation of an autophagy score signature for the prediction of post-operative survival in colorectal cancer. *Front Oncol*. (2019) 9:878. doi: 10.3389/fonc.2019.00878

27. Zhang B, Wu Q, Li B, Wang D, Wang L, Zhou YL. m 6 A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. *Mol Cancer*. (2020) 19:1–21. doi: 10.1186/s12943-020-01170-0

28. Song J, Lin Z, Liu Q, Huang S, Han L, Fang Y, et al. MiR-192-5p/RB1/NF-κBp65 signaling axis promotes IL-10 secretion during gastric cancer EMT to induce Treg cell differentiation in the tumour microenvironment. *Clin Transl Med*. (2022) 12:e992. doi: 10.1002/ctm2.992

29. Li R, Zhou R, Wang H, Li W, Pan M, Yao X, et al. Gut microbiota-stimulated cathepsin K secretion mediates TLR4-dependent M2 macrophage polarization and promotes tumor metastasis in colorectal cancer. *Cell Death Differ*. (2019) 26:2447–63. doi: 10.1038/s41418-019-0312-y

frontiers | Frontiers in Immunology

TYPE Review
PUBLISHED 08 November 2024
DOI 10.3389/fimmu.2024.1462697

OPEN ACCESS

EDITED BY
Mar Requena Mullor,
University of Almeria, Spain

REVIEWED BY
Lukman Olalekan Afolabi,
Indiana University School of Medicine, South Bend, United States
Giuseppe Schepisi,
Scientific Institute of Romagna for the Study and Treatment of Tumors (IRCCS), Italy

*CORRESPONDENCE
Alireza Nourazarian
noorazarian_a@khoyums.ac.ir

†These authors have contributed equally to this work and share first authorship

RECEIVED 10 July 2024
ACCEPTED 23 September 2024
PUBLISHED 08 November 2024

CITATION
Amiri M, Moaveni AK, Majidi Zolbin M, Shademan B and Nourazarian A (2024) Optimizing cancer treatment: the synergistic potential of CAR-T cell therapy and CRISPR/Cas9.
Front. Immunol. 15:1462697.
doi: 10.3389/fimmu.2024.1462697

Optimizing cancer treatment: the synergistic potential of CAR-T cell therapy and CRISPR/Cas9

Maryam Amiri[1†], Amir Kian Moaveni[1†], Masoumeh Majidi Zolbin[1], Behrouz Shademan[2] and Alireza Nourazarian[3*]

[1]Pediatric Urology and Regenerative Medicine Research Center, Children's Medical Center, Tehran University of Medical Sciences, Tehran, Iran, [2]Medical Journalism, School of Paramedical Sciences, Shiraz University of Medical Sciences, Shiraz, Iran, [3]Department of Basic Medical Sciences, Khoy University of Medical Sciences, Khoy, Iran

Optimizing cancer treatment has become a pivotal goal in modern oncology, with advancements in immunotherapy and genetic engineering offering promising avenues. CAR-T cell therapy, a revolutionary approach that harnesses the body's own immune cells to target and destroy cancer cells, has shown remarkable success, particularly in treating acute lymphoblastic leukemia (ALL), and in treating other hematologic malignancies. While CAR-T cell therapy has shown promise, challenges such as high cost and manufacturing complexity remain. However, its efficacy in solid tumors remains limited. The integration of CRISPR/Cas9 technology, a powerful and precise genome-editing tool, also raises safety concerns regarding unintended edits and off-target effects, offers a synergistic potential to overcome these limitations. CRISPR/Cas9 can enhance CAR-T cell therapy by improving the specificity and persistence of CAR-T cells, reducing off-target effects, and engineering resistance to tumor-induced immunosuppression. This combination can also facilitate the knockout of immune checkpoint inhibitors, boosting the anti-tumor activity of CAR-T cells. Recent studies have demonstrated that CRISPR/Cas9-edited CAR-T cells can target previously untreatable cancer types, offering new hope for patients with refractory cancers. This synergistic approach not only enhances the efficacy of cancer treatment but also paves the way for personalized therapies tailored to individual genetic profiles. This review highlights the ongoing research efforts to refine this approach and explores its potential to revolutionize cancer treatment across a broader range of malignancies. As research progresses, the integration of CAR-T cell therapy and CRISPR/Cas9 holds the promise of transforming cancer treatment, making it more effective and accessible. This review explores the current advancements, challenges, and future prospects of this innovative therapeutic strategy.

KEYWORDS

CAR-T cell therapy, CRISPR/Cas9 technology, immunotherapy, genetic engineering, personalized cancer treatment

1 Introduction

Cancer continues to be a major health challenge in modern times, with its burden increasing globally. The complexity of this disease and its high levels of genetic diversity have prompted scientists to consider more tailored approaches to treatment. Typical approaches in cancer treatment include surgery, radiation, and systemic drugs, which are usually used alone or in combination. Cytotoxic therapies for most people bring about significant suffering and do not provide long-term immunity against the disease (1).

Alongside the growing demand for cancer treatment, the cost per patient has consistently risen. Consequently, healthcare spending on cancer care has increased at a rate that surpasses the rise in cancer incidence (2). The pursuit of optimizing cancer treatment stands at the forefront of modern oncology, driven by the critical need to enhance patient outcomes and overcome the limitations of existing therapies. In response, advancements in immunotherapy and genetic engineering have emerged as promising avenues, offering more targeted and effective solutions.

What has been envisioned over the past decade is that immunotherapy has quickly transformed from a fancy concept to a practical and revolutionary method for treating cancer. One of the most promising and innovative approaches among these strategies is chimeric antigen receptor T (CAR-T) cell therapy. This marks a significant milestone in cancer treatment, departing from traditional methods that aimed to harness the body's immune system. More specifically, CAR T-cell therapy has shown tremendous potential for targeting and combating cancer cells in ways previously unattainable through conventional treatment methods (3). To achieve this, the patient's immune cells are genetically engineered to express chimeric receptors that specifically bind to antigens and activate cytotoxic T lymphocytes. This enhancement boosts the immune cells' efficiency in targeting and destroying cancer cells (4). In 2008, Malcolm Brenner and colleagues in Houston reached a pivotal achievement in the clinical application of CAR T cells (5). This method has demonstrated notable efficacy in managing blood cancers such as leukemia and lymphoma, resulting in substantial remission rates for patients with few prior treatment options (6). Although the potential of CAR T-cell therapy is promising, it still poses several challenges that researchers must overcome. The treatment has been noted for its success, although some patients are subject to severe side effects, toxicity of treatment, or inefficacy, in which case they face a recurrence of cancer (7–9). Scientists are thus probing different strategies to tackle these, such as optimization of the CAR structure, combination therapies with radiotherapy and chemotherapy, immune checkpoint inhibitors, and oncolytic viruses (10–13). In this regard, although these strategies have achieved improvement with respect to their efficacy and safety, they have not been able to completely alleviate all the concerns. Besides, such high treatment costs have also prevented the large-scale clinical application of CAR-T cell therapy. So, there is a continued need to further refine this technology for better effectiveness and safety, together with a lower manufacturing cost (14). These efforts would further make CAR-T cell therapy more feasible and accessible to a larger population of cancer patients.

The clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9 technology is a widely recognized genome editing tool that utilizes guide RNA (sgRNA) to target specific DNA sequences. It has garnered significant interest due to its ability to target multiple genes simultaneously, its ease of implementation, and its cost-effectiveness (15). A significant breakthrough was made by Zhang Feng and et al. (16) to demonstrate the effectiveness of CRISPR/Cas9 technology in human and mammalian cells. This advancement has expanded the capabilities of CAR-T cell therapy. Currently, researchers are utilizing the CRISPR/Cas9 system to edit and engineer CAR-T cells, enhancing their ability to target cancer. Specifically, these scientists are modifying the cells to improve specificity, target a greater range of antigens, prolong persistence, broaden the scope of action against cancer, and enhance safety. Of particular interest is the application of CRISPR/Cas9 in engineering CAR-T cells (17). This powerful gene-editing tool allows for precise genetic changes, thereby improving the specificity and efficacy of CAR-T cells. Furthermore, it enables the development of CAR-T cells that can overcome immunosuppression in the tumor microenvironment and potentially target a wider array of cancer types. CRISPR/Cas9 technologies will unlock these possibilities and more, presenting a flexible and highly promising option for CAR-T cell therapy for most cancer patients, thus mitigating current limitations. This review explores the synergistic potential of combining CAR-T cell therapy with CRISPR/Cas9 technology. By delving into the latest advancements, current challenges, and future prospects, we aim to shed light on how this innovative therapeutic strategy can transform cancer treatment, making it more effective, accessible, and personalized for patients worldwide.

2 CAR-T cell therapy

CAR-T cell therapy is a revolutionary approach in cancer treatment that leverages genetically engineered T cells to target and eliminate cancer cells. By introducing chimeric antigen receptors (CARs) into T cells, CAR-T therapy equips them with the ability to recognize and destroy specific cancer cells. CARs are artificial proteins composed of an extracellular antigen-binding domain, a hinge region, a transmembrane domain, and an intracellular signaling domain. The antigen-binding domain, typically derived from a single-chain variable fragment (scFv) of a monoclonal antibody, recognizes and binds to specific antigens on the surface of cancer cells. Upon binding to the target antigen, the CAR triggers a cascade of signaling pathways within the T cell, leading to its activation, proliferation, and release of cytotoxic molecules that destroy the cancer cell (6). However, CAR-T cell therapy also faces challenges, such as potential for severe side effects, high manufacturing costs, and limited efficacy in solid tumors.

2.1 Mechanism of action

CAR-T therapy is the culmination of extensive research across basic and clinical disciplines. CARs are artificial molecules displayed on cell surfaces, enabling T cells or other effector cells like natural killer (NK) cells to focus their cytotoxic activity on tumor cells expressing the CAR target antigen. These CAR

transgenes are introduced into T cells either temporarily or permanently (18, 19). Specifically, using CAR mRNA electroporation leads to temporary CAR expression, whereas employing lentiviral or gammaretroviral gene delivery methods results in the integration of CAR transgenes into T cell genomes, ensuring their stable expression (18, 19). CARs can target either tumor-specific antigens (TSAs) or tumor-associated antigens (TAAs) (20) (Figure 1). The capability of CARs to recognize and engage with these target antigens primarily relies on their extracellular domain, comprising a targeting domain and a hinge (or spacer) (21). CARs commonly utilize the scFv from a monoclonal antibody (mAb) as their targeting domain (22). Nevertheless, nanobodies (VHH) and toxins have also been employed for this purpose (23). The hinge serves as the bridge between the extracellular domain and the transmembrane domain of CARs (24). CARs also feature an intracellular domain comprising an activation domain, typically CD3ζ derived from the T-cell receptor (TCR) CD3 complex, and one or two co-stimulatory domains such as CD28, 4-1BB (CD137), ICOS, or OX40 (CD134) (24). The transmembrane domain connects the extracellular and intracellular domains of CARs, serving as an anchor to stabilize CAR molecules within the cell membrane. Upon encountering their target antigen, CAR molecules initiate downstream signaling pathways that activate T cells. This mechanism operates independently of the major histocompatibility complex (MHC) for activation (24). CARs possess the capability to identify target antigens directly, bypassing the need for antigen processing and presentation by MHC molecules on antigen-presenting cells. As a result, any surface antigen primarily expressed on malignant cells, not normal ones, and accessible to targeting with monoclonal antibodies (mAbs), is considered a viable candidate for CAR-T therapy (20, 21).

It hasn't been long since the first CAR-T product, named tisagenlecleucel, received approval from the US Food and Drug Administration (FDA) in 2017 for clinical use (25, 26). Today, CAR-T therapy stands as an effective treatment option available for patients with certain relapsed or refractory (R/R) hematologic malignancies, including B-cell acute lymphoblastic leukemia (B-ALL), diffuse large B-cell lymphoma (DLBCL), follicular lymphoma (FL), mantle cell lymphoma (MCL), and multiple myeloma (MM) (27).

2.2 Success in hematologic malignancies

The development of CAR T-cell therapy for select hematological malignancies represents one of the most remarkable therapeutic advances in the past decade. Currently, CD19-targeted CAR T-cell therapy is approved for relapsed/refractory diffuse large B-cell lymphoma and acute lymphoblastic leukemia (28). However, there

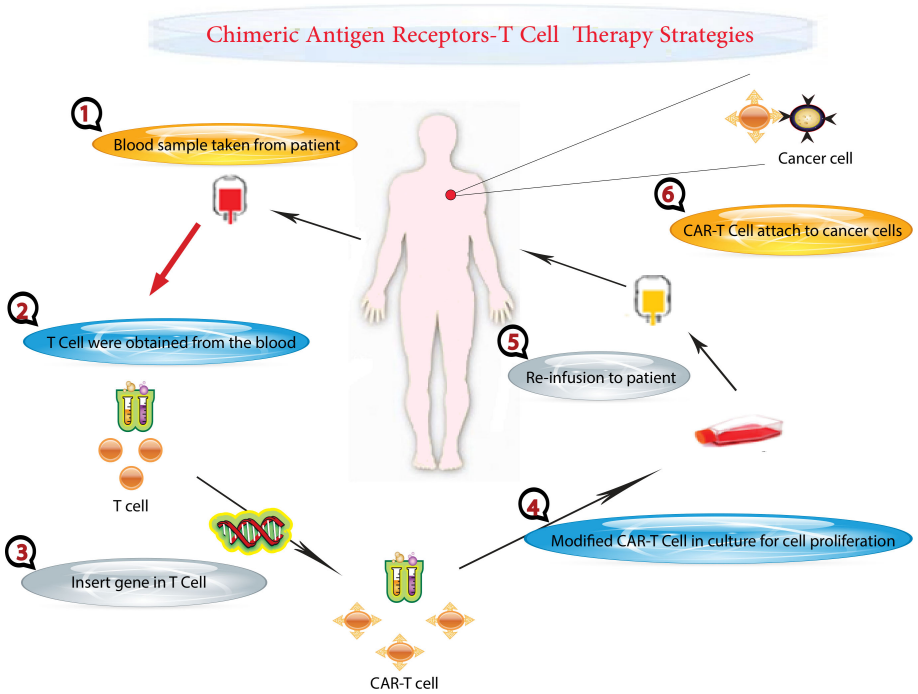


FIGURE 1
CAR T-Cell Therapy Process. This figure illustrates the step-by-step process of CAR T-cell therapy, a personalized treatment for cancer patients. The process begins at the hospital, where blood is drawn from a cancer patient. The collected blood undergoes a separation process to isolate T cells, a type of white blood cell critical for immune response. The remaining components of the blood are returned to the patient's body. The isolated T cells are then taken to a laboratory for genetic modification. Here, scientists introduce specific genetic changes to the T cells. These genetic modifications equip the T cells with chimeric antigen receptors (CARs) on their surface. CARs are specialized receptors that enhance the T cells' ability to recognize and target cancer cells accurately. The genetically modified T cells, now known as CAR T cells, are cultured in the laboratory to increase their number, ensuring there are sufficient cells for effective treatment. Finally, the expanded population of CAR T cells is administered back to the patient through an intravenous injection. These modified cells circulate in the patient's body, seeking out and destroying cancer cells.

is significant interest in the application of CAR T-cell therapy to other hematological malignancies, including multiple myeloma, where the current focus is on the development of B-cell maturation antigen-directed CAR T-cell therapy. Despite the successes achieved to date, there remain significant challenges associated with CAR T-cell therapy and substantial research efforts are underway to develop new targets and approaches.

Presently, CD19 and BCMA represent the predominant targets in CAR-T cell therapy. Despite remarkable success in treating B cell malignancies, relapse following anti-CD19 CAR-T cell therapy and anti-BCMA CAR-T cell therapy is common. Moreover, due to antigenic diversity in acute myeloid leukemia (AML) and the absence of CD19 expression in T cell malignancies, ongoing research is exploring various potential targets. CD19 is a crucial target antigen in B cell malignancies such as B-ALL and NHL. Recent advancements in anti-CD19 CAR-T cell therapy have led to rapid and long-lasting responses in patients with relapsed or refractory (R/R) B-ALL and NHL, fundamentally changing treatment approaches for these conditions. Currently, four anti-CD19 CAR-T cell products have received FDA approval for treating R/R B-ALL and NHL (29). Despite its clinical success, CD19 antigen loss is a common issue (30). To address this, combined therapy using anti-CD19 and anti-CD20 CAR-T cells has been explored for R/R DLBCL, demonstrating safety and feasibility (31). CD22 is prominently expressed on many B cell malignancies, including B-ALL and DLBCL (32, 33). Clinical trials have shown that anti-CD22 CAR-T cell therapy is highly effective in patients with R/R B-ALL and R/R DLBCL who have not responded to previous anti-CD19 CAR-T cell therapy (34–36). Moreover, humanized anti-CD22 CAR-T cells have exhibited potent activity against leukemia cells even with low CD22 expression (37). Emerging strategies in CAR T-cell therapy for B-cell malignancies are focusing on addressing challenges associated with autologous T-cell production, particularly for patients with insufficient healthy T cells. A promising approach involves the use of gene-editing technologies to create universal CAR T cells (UCART19) by modifying donor T cells to introduce a chimeric antigen receptor (CAR) and disrupt T-cell receptor (TCR) and CD52 genes. This process produces "off-the-shelf" CAR T cells capable of evading host immune responses, enabling their use in unmatched recipients. Qasim et al. successfully applied this strategy in two infants with relapsed refractory acute lymphocytic leukemia, achieving molecular remission and bridging them to successful allogeneic stem cell transplantation (38). This groundbreaking application of TALEN-mediated gene editing highlights the potential of universal CAR T cells in treating aggressive B-cell leukemias, offering a scalable and feasible alternative to patient-specific therapies.

Patients diagnosed with relapsed or refractory (R/R) T-cell acute lymphoblastic leukemia (T-ALL) and T-cell lymphomas typically face a grim prognosis. Unlike the notable clinical success of anti-CD19 CAR-T cell therapy in B cell malignancies, the effectiveness and safety of CAR-T cell therapy in T cell malignancies are largely under investigation. A general problem to consider in the production of CAR-T cells is on-target-off-tumor toxicity. This problem is associated with antigens of target expression on normal, non-malignant cells, leading to their destruction by CAR-Ts. Fratricide, i.e. killing of CAR-T cells by each other, is also a problem because the targets of this kind of T cell tumor antigen therapy, antigens expressed on T cells (such as CD5, CD7, etc.), lead to it (39). Other problems are T-cell aplasia and contamination of the CAR T-cell product with tumor cells (39).

CD7 is highly expressed in 95% of T-ALL patients, making it an attractive target for treating T-ALL (40). Two R/R T-ALL patients received allogeneic anti-CD7 CAR-T cell therapy in an open-label, single-arm clinical trial. One patient achieved remission lasting over a year, while the other relapsed 48 days after CAR-T cell infusion (40). Another phase I clinical trial involved 20 R/R T-ALL patients who received donor-derived anti-CD7 CAR-T cell therapy, with 90% achieving complete remission (CR) (41). Additionally, a case study reported the successful treatment of an 11-year-old T-ALL patient, who had not responded to initial treatment, with autologous anti-CD7 CAR-T cell therapy resulting in remission by day 17 and subsequent hematopoietic stem cell transplantation (HSCT) (42). CD5 is expressed in approximately 85% of T-cell malignancies, including T-cell lymphoblastic lymphoma (T-LBL) and peripheral T-cell lymphoma (PTCL). Recent studies have demonstrated the efficacy of anti-CD5 CAR-T cells in eliminating malignant T cells (43). In a phase I clinical trial, a refractory T-LBL patient with central nervous system (CNS) involvement achieved CR within four weeks of receiving anti-CD5 CAR-T cell therapy (44). Moreover, preclinical studies have shown promising activity of anti-CD4 CAR-T cells against T cell malignancies (45). However, targeting CD4, CD5, and CD7 may lead to depletion of normal T cells and fratricide of CAR-T cells, as these antigens are also expressed in normal T cells (46). The chemokine receptor CCR9 is expressed in over 70% of T-ALL patients but in less than 5% of normal T cells. It is associated with multidrug resistance and poor prognosis, making it an ideal target for CCR9-positive T-ALL. Preclinical studies have demonstrated the potent anti-leukemic activity of anti-CCR9 CAR-T cells, which are also resistant to fratricide (47). In addition, several recent clinical trials have shown excellent results for CAR-T cell therapy. In this regard, genome-editing technology can help overcome these problems (39).

AML is the most prevalent acute leukemia in adults. CAR T-cell therapy for AML has been elusive so far, mainly because of the lack of truly AML-specific surface antigens that make targeting AML very challenging. AML cells express various cell surface antigens such as CD123, CD34, and CD33. However, expression of these same antigens is also shared by healthy HSPCs and their myeloid and/or lymphoid progenitors. Besides, production of CAR T cells per se may also present difficulties in patients with active AML, possibly because of the inhibition of T-cell expansion by AML blasts, or previous exposure to chemotherapy damaging T cells (48). However, progress has been made toward the use of CAR T-cell therapy in this disease.

Recently, CLL-1, LILRB4, and Siglec-6 have emerged as potential targets. In preclinical studies, CLL-1, a myeloid cell surface marker overexpressed on leukemic stem cells, has shown specificity in eliminating CLL-1-positive leukemia (49–51). Notably, CLL-1 is absent in hematopoietic stem cells, enhancing its therapeutic potential. NPM1 mutations are present in 30%-35% of AML cases and are considered pivotal in leukemic cell initiation.

CAR-T cells targeting a nucleophosmin neoepitope, presented by HLA-A2, demonstrated potent anti-leukemia effects in preclinical models (52). CD70, expressed on AML blasts but not normal myeloid cells, is also being investigated as a promising CAR-T cell therapy target (53, 54). LILRB4, highly expressed in monocytic AML cells, presents another attractive target for monocytic AML (55). Siglec-6, found in approximately 60% of AML patients and absent on normal hematopoietic stem and progenitor cells, has effectively eliminated AML blasts in preclinical xenotransplantation models. These findings support Siglec-6 as a validated target for CAR-T cell therapy in AML (56). A novel approach to CAR T-cell therapy for AML aims to overcome the challenge of prolonged myeloablation while ensuring long-term persistence of therapeutic cells. This strategy involves gene editing to remove the CAR target antigen, such as CD33, from donor hematopoietic stem and progenitor cells (HSPCs). After these CD33−/− HSPCs are transplanted into the patient and engrafted, CD33-specific CAR T cells from the same donor can be administered, allowing normal hematopoiesis to continue without being targeted by the CAR T cells (48). Early studies using CRISPR/Cas9 technology have shown that CD33−/− HSPCs can resist CD33-directed CAR T cells while maintaining normal hematopoietic and immune functions (57). A clinical trial at the University of Pennsylvania is being developed to test this approach in patients with relapsed/refractory AML. This strategy not only highlights the potential of gene editing in improving CAR T-cell therapies but also opens new avenues for targeting other antigens, such as CD123, with careful consideration of their biological roles and potential impact on healthy tissues.

2.3 Challenges in treating solid tumors

CAR-T cell therapy, while highly effective in treating certain cancers, is associated with a range of adverse effects that can significantly impact patient outcomes. These include immune-related toxicities such as Cytokine Release Syndrome (CRS) and Immune-Effector-Cell-Associated Neurotoxicity Syndrome (ICANS), both of which can cause severe inflammatory and neurological complications. Metabolic toxicities like Tumor Lysis Syndrome (TLS) are also common, arising from the rapid destruction of cancer cells, which can lead to life-threatening metabolic imbalances. Additionally, on-target/off-tumor toxicity occurs when CAR-T cells attack healthy tissues that express the same antigen as the cancer cells, leading to potential organ damage. These toxicities are influenced by various factors, including the patient's health status, tumor burden, the dose of CAR-T cells, and the rate of infusion, necessitating careful monitoring and tailored management strategies to mitigate risks and enhance the safety and efficacy of CAR-T cell therapy.

2.3.1 Immune-related toxicities

The most serious and prevalent toxicity associated with CAR-T cell therapy is a systemic inflammatory reaction known as cytokine release syndrome (CRS) (58). Activation of CAR-T cells upon recognition of tumor antigens triggers CRS, characterized by severe systemic inflammation. Following CAR-T cell infusion, there is a significant increase in serum levels of cytokines such as interferon-gamma (IFN-γ), interleukin-6 (IL-6), tumor necrosis factor-alpha (TNF-α), granulocyte-macrophage colony-stimulating factor (GM-CSF), IL-2, IL-8, and IL-10, leading to a cytokine storm. This initial response is followed by a secondary inflammatory phase involving antigen-presenting cells (APCs) like dendritic cells (DCs), B cells, macrophages, and monocytes, which express the cell surface protein CD40. Activated CAR-T cells also express high levels of the CD40 ligand (CD40L) (59).

Patients should be closely monitored from the emergence of early symptoms of CRS and treated symptomatically with antipyretics and analgesics, although it is contraindicated for NSAIDs because they can alter kidney function. If a diagnosis of an infection is made, particularly in patients who are febrile and neutropenic, they should be ruled out and started on empiric antibiotics, taking into consideration the raised incidence of infection post a regimen that deploys lymphocytes to such a great extent. Studies show that almost one-quarter of patients under CD19-targeted CAR T-cell therapy experience infections, mostly bacteremias and respiratory viral infections, in the first four weeks post-infusion. The use of prophylactic antibiotics is not universally established but is done in some centers (60). Other supportive measures for CRS include antiemetics, oxygen, intravenous fluids, and low-dose vasopressors if needed; it generally avoids corticosteroids (60). Prompt recognition and treatment of severe CRS (sCRS) is important as it may lead to multiorgan failure akin to septic shock. First-line therapy of sCRS is tocilizumab, an IL-6 receptor antagonist, with a high response rate. Next steps in failure cases include the use of corticosteroids, whereas other inflammatory cytokine-targeted therapies, such as anti-TNFα or IL-1R inhibitors, should be offered in rare, resistant cases (60).

Immune-effector-cell-associated neurotoxicity Syndrome (ICANS), formerly known as CAR-T-cell-related encephalopathy syndrome, represents the second most common adverse effect of CAR-T cell therapy. ICANS is characterized by cytokine-mediated neurotoxicity rather than direct cytotoxic effects, although its exact pathophysiology remains unclear. Research suggests that endothelial activation plays a role in the development of ICANS (61). Endothelial activation associated with CRS can disrupt the blood-brain barrier (BBB), allowing inflammatory cytokines and immune cells to infiltrate the cerebrospinal fluid and reach the central nervous system (CNS). Once in the brain parenchyma, these cytokines and immune cells can cause inflammation, leading to neuronal impairment or damage. CAR-T cells, along with monocytes and macrophages, are attracted to the CNS and contribute to the release of cytokines, which are central to the development of ICANS (59). Pericytes, a type of mural cell surrounding capillary endothelium and expressing CD19, are crucial for maintaining the integrity of the BBB. In treatments targeting CD19, increased BBB permeability due to pericyte activation has been implicated in ICANS development (62). ICANS manifests with various neurological symptoms that often progress in a characteristic pattern. Symptoms can appear as early as the fourth or fifth day after CAR-T cell infusion and as late as the third or fourth week (63, 64). ICANS rarely occurs without preceding CRS, and when it does, it tends to be mild (65).

Common symptoms include headache, tremors, speech difficulties, confusion, delirium, impaired consciousness (such as obtundation, lethargy, and stupor), and occasionally, focal neurological deficits (66).

Prompt recognition and management of CNS toxicity in CAR T-cell therapy are crucial, similar to the approach for CRS. Patients experiencing neurotoxicity require close monitoring, with ICU transfer recommended for those with grade ≥3 toxicity and considered for those with grade 2 toxicity, depending on the center's policy. In severe cases, neurologic toxicity may necessitate intubation and mechanical ventilation for airway protection, even without respiratory failure. Notably, fever within the first 36 hours post-infusion has been associated with a high sensitivity for subsequent severe neurotoxicity. Management typically involves corticosteroids, tailored to the severity and specific CAR T-cell product used. Importantly, unlike CRS, neurotoxicity generally does not respond to tocilizumab, and in some cases, the drug may even exacerbate the condition. Additionally, neurotoxicity often resolves more slowly than CRS (60).

2.3.2 Metabolic toxicities

When cancer treatment effectively kills cells, it can release significant amounts of phosphate, potassium, and nucleic acids into the bloodstream, potentially leading to tumor lysis syndrome (TLS) (67). While TLS has traditionally been associated with chemotherapy, CAR-T cell therapy has also been linked to acute anaphylaxis and TLS, sometimes occurring even without prior conditioning chemotherapy (68, 69). The rapid death of lymphoma cells following CAR-T cell treatment can pose challenges if the kidneys are unable to process the byproducts of cell lysis quickly enough, resulting in conditions such as hyperuricemia, hyperkalemia, hyperphosphatemia, and hypocalcemia. Accumulation of uric acid, calcium phosphates, and ferritin can further exacerbate acute kidney injury, leading to systemic inflammation and iron overload.

Unlike other novel therapies for hematologic malignancies that have heightened the risk of TLS, TLS is relatively uncommon following CAR T-cell therapy, even in high-risk scenarios. Nevertheless, precautionary measures, such as intravenous hydration and prophylactic administration of allopurinol or febuxostat, should be implemented prior to starting conditioning lymphodepleting chemotherapy, particularly in patients with elevated uric acid levels or high tumor burden. It is essential to closely monitor for signs and symptoms of TLS and manage them according to established guidelines to prevent complications (60).

2.3.3 On-target/off-tumor toxicity

Ideally, CAR-T cell therapy targets antigens expressed exclusively on malignant cells, sparing healthy tissues. However, solid tumors have seen limited success with this approach. Many tumor antigens are tumor-associated antigens (TAAs), present on both healthy and tumor cells. Consequently, CAR-T cells often struggle to differentiate normal cells from cancerous ones, resulting in "on-target, off-tumor" toxicity (59). This phenomenon is more common in solid tumors, underscoring the need for extensive research into identifying tumor-specific antigens (TSAs). Several cases illustrate the challenges posed by TAAs expressed on normal tissues. For instance, the first patient treated with CAR-T cells targeting HER2 experienced respiratory distress and significant lung infiltrates within 15 minutes of infusion, leading to lung damage and death (70). Similar lung toxicity was observed in clinical trials testing CAR-T cells targeting CEA (71). It's important to recognize that the occurrence of adverse effects in CAR-T cell therapy can vary significantly depending on the specific CAR-T product (Figure 2), cancer type, and individual patient factors.

Thus, it is important to prevent "on-target/off-tumor" effects, which would result in collateral damage to normal tissues with CAR T-cell therapy. This can be achieved by targeting antigens that are more selective, such as κ and λ light chains of immunoglobulins to preserve humoral immunity, potentially maintaining antitumor activity in some B-cell malignancies (72, 73). Another approach is the extinction of the target antigen on the normal population of hematopoietic stem cells. This has been demonstrated in a study where CD33-deficient stem cells were genetically modified and co-infused with CD33+ CAR T cells. The infused T cells engrafted, allowing for normal myeloid function with no evidence of off-target effects. Furthermore, optimization of the CAR design, ranging from engineering dual CARs for multi-antigen recognition of tumor-specific antigens to the use of affinity-tuned CARs, may increase the precision of target recognition and minimize the risk of relapse from antigen deletion (74, 75). Lastly, the capacity for a targeted reduction of CAR activity in the event of severe toxicity is an active research area, which may be achieved using techniques such as inducible suicide genes, monoclonal antibodies, small molecule modulators, or CRISPR/Cas9 technology to remotely control or transiently turn off CAR T cells (61).

2.3.4 Other factors influencing toxicity

In vivo, the expansion of CAR-T cells and associated toxicity may be exacerbated by factors such as tumor burden, intensity of conditioning therapy, higher infusion doses, and CAR design. For example, pediatric patients with high baseline tumor burdens in B-ALL tend to experience more pronounced CAR-T cell proliferation and more severe cytokine release syndrome (CRS) (59). Clinical studies have consistently shown that patients with larger tumor burdens experience more severe and frequent CRS, likely due to heightened T cell activation levels (76, 77). Moreover, patients with higher initial burdens of ALL and those receiving higher doses of CD19 CAR-T cells have been found to have increased incidence rates of CRS (78, 79). Improving CAR-T cell efficacy and durability of response is the central goal of conditioning therapy aimed at enhancing clinical outcomes in cancer patients. Even without chemotherapy conditioning, adverse effects such as thrombocytopenia, anemia, and neutropenia have been documented (59). Effective strategies for prevention and mitigation of these toxicities include preemptive treatments, careful patient monitoring, and adjustments in CAR-T cell administration protocols. Understanding and managing these adverse effects is crucial for maximizing the therapeutic benefits of CAR-T cell therapy while minimizing its risks.

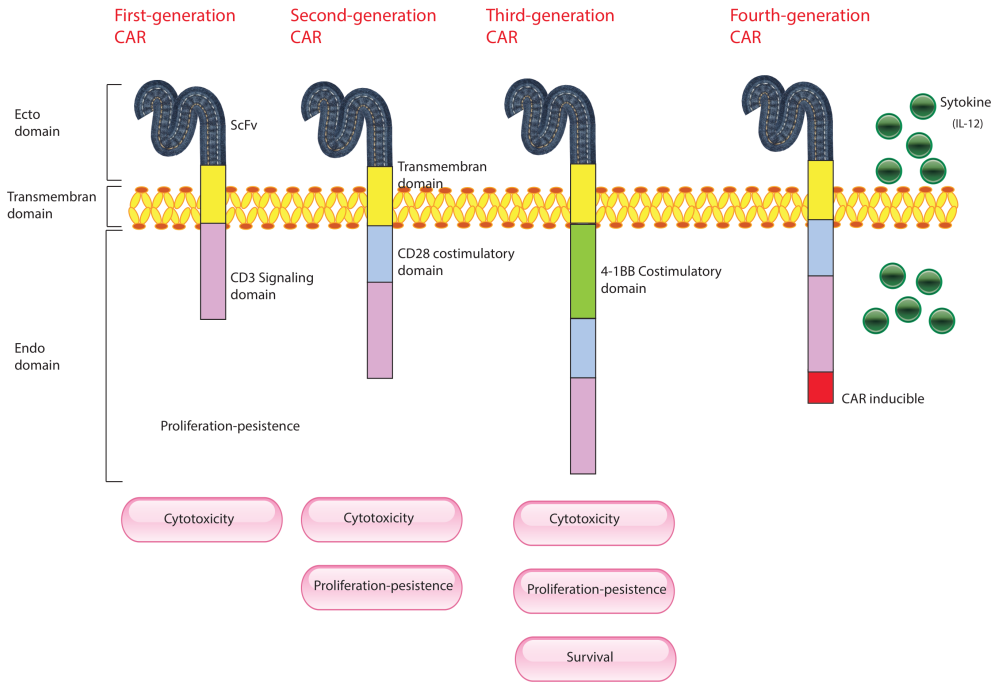


FIGURE 2
Evolution of CAR-T Cells. This figure illustrates the development of CAR-T cells through three generations, each incorporating advancements to improve their effectiveness against cancer. First Generation CAR-T Cells: These are composed of a single-chain variable fragment (scFv) derived from an antibody, which is responsible for targeting cancer cells, and the CD3 immunoglobulin, which is part of the T-cell receptor complex essential for initiating T-cell activation. Second Generation CAR-T Cells: Building on the first generation, these cells include additional co-stimulatory molecules, such as CD28. This enhancement provides stronger and more sustained activation signals to the CAR-T cells, improving their persistence and effectiveness in targeting and destroying cancer cells. Third Generation CAR-T Cells: These cells incorporate multiple co-stimulatory molecules, such as CD28, CD134 (OX40), and CD137 (4-1BB). The inclusion of multiple stimulatory signals further enhances the CAR-T cells' ability to proliferate, survive, and eliminate cancer cells, offering an even more robust therapeutic effect.

Recently, CAR-Natural Killer (CAR-NK) and CAR-macrophages (CAR-M) were introduced as a complement/alternative to CAR-T cell therapy for solid tumors. CAR Natural Killer (NK) cells have several advantages over CAR T cells as the NK cells can be manufactured from pre-existing cell lines or allogeneic NK cells with unmatched major histocompatibility complex (MHC); can kill cancer cells through both CAR-dependent and CAR-independent pathways; and have less toxicity, especially cytokine-release syndrome and neurotoxicity. At least one clinical trial showed the efficacy and tolerability of CAR NK cell therapy. Additionally, CAR-NK cells might be generated in large scale from several sources which would suggest them as promising off-the-shelf product. CAR-M immunotherapy with its capabilities of phagocytosis, tumor-antigen presentation, and broad tumor infiltration, is currently being investigated (80, 81).

In contrast, CAR-Natural Killer (CAR-NK) and CAR-Macrophages (CAR-M) have recently been introduced as alternatives or complements to CAR-T cell therapy for solid tumors. There are several advantages of CAR NK cells over CAR T cells: NK cells can be manufactured from pre-existing cell lines or allogeneic NK cells with unmatched major histocompatibility complex. They can kill cancer cells through both CAR-dependent and CAR-independent pathways, and have less toxicity, particularly when it comes to cytokine-release syndrome and neurotoxicity. At least one clinical trial has shown the effectiveness and safety of CAR NK-cell therapy. Furthermore, CAR-NK cells can be derived on a large scale from various sources, making them a potential off-the-shelf product. Currently, CAR-M immunotherapy, which involves phagocytosis, tumor antigen presentation, and extensive tumor infiltration, is under investigation. In this study, we designed an adenovirus-induced CAR-M using an anti-HER2 CAR and the CD3ζ intracellular domain. This CAR-M demonstrated *in vitro* specificity in terms of antigen-specific phagocytosis against HER2-positive tumor cells. A single injection of anti-HER2 CAR-M reduced tumor load and prolonged survival in mice. It also shifted M2 macrophages into M1 macrophages, stimulated an inflammatory tumor microenvironment (TME), and exhibited anti-tumor cytotoxicity. Importantly, HER2 CAR-M had the capability of inducing epitope spread, which could be an additional approach to prevent tumor immune escape. Another study combined anti-HER2 CAR with transduced primary human CD14+ peripheral blood monocyte-derived macrophages. These CAR-Ms stimulated phagocytosis of the HER2+ ovarian cancer cell line SKOV3 in a dose-dependent manner. The authors also demonstrated that the transduction of macrophages was not impaired by the antitumor activity, as their transduction with a control CAR did not exhibit any antitumor activity. *In vivo*, the SKOV3 tumor burden in NOD-SCID mice was significantly reduced in those cohorts that had been treated with primary human anti-HER2 CAR-Ms. The authors further showed that the CAR-Ms persisted and remained resistant to the

immunosuppressive cytokines secreted by the TME. In contrast, the CAR-Ms secreted pro-inflammatory cytokines, leading to the conversion of macrophages from the M2 to the M1 phenotype and subsequently transforming the TME into a proinflammatory environment. Additionally, the combination of donor-derived T cells with CAR-Ms enhanced the antitumor response *in vivo*. When murine-derived anti-HER2 CAR-Ms were infused, Pierini et al. reported inhibition of tumor growth, extended overall survival, increased levels of CD4+ and CD8+ T cells, NK cells, and dendritic cells in the TME. These researchers also found that CAR-Ms play a crucial role in regulating the TME through the upregulation of MHC I/II expression on cancer cells.

3 CRISPR/Cas9 technology

3.1 CRISPR/Cas9: advantages of CRISPR/Cas9 over traditional methods

Genome editing involves modifying genomic DNA to artificially alter genetic information, resulting in permanent changes to the function of the targeted gene (82). Several tools have been developed for precisely modifying specific regions of the genome. Genome editing nucleases, such as zinc finger nucleases (ZFNs), meganucleases, and transcription activator-like effector nucleases (TALENs), induce double-strand breaks at specific genomic sites (83). These nucleases facilitate targeted modifications by initiating endogenous DNA repair mechanisms, primarily non-homologous end joining (NHEJ), which repairs double-strand breaks without requiring a template (84). This approach can effectively replace or delete target genes, although designing and engineering these nucleases to target new sequences remains a significant challenge (83, 85).

The CRISPR/Cas9 system has emerged as a leading gene editing tool in recent years (Figure 3). Initially discovered in bacteria as a defense mechanism against viruses, this system consists of an endonuclease (Cas9) and a single-guide RNA (sgRNA) that directs Cas9 to specific locations in the genome through base pairing. Cas9 then cleaves the target DNA, prompting the host cell to repair the break. If a donor template with homologous arms is present, homology-directed repair (HDR) can occur, resulting in precise editing of the genome. Alternatively, non-homologous end joining (NHEJ) can resolve the break by inserting or deleting nucleotides (indels), often disrupting the gene's reading frame. This system is highly effective, simple to use, and widely applicable, making it a powerful tool for genome editing (86).

Compared to previous genome editing tools like TALEN and ZFN, CRISPR/Cas9 offers advantages such as rapid, cost-effective sgRNA production, in contrast to the synthesis of custom guide proteins for TALEN or ZFN. Additionally, CRISPR/Cas9 can simultaneously modify multiple genes by utilizing multiple sgRNAs targeting different genomic sites, surpassing the capabilities of ZFN and TALEN (87, 88). This system also excels in its ability to alter the epigenome, transcriptome, and genome of immune-related cells and cancer cells (89). The application of CRISPR/Cas in cancer treatment hinges on the careful selection of target genes (90), offering innovative solutions for clinical applications in cancer immunotherapy and gene therapy. Nonetheless, ongoing research is essential to further refine and optimize the CRISPR/Cas method for genome editing.

3.2 Applications in cancer research

In cancer research, the applications of CRISPR-Cas9 mainly involve the screening of oncogenic mutations and tumor suppressors, the construction of *in vivo* and *in vitro* cancer models, and cancer gene therapy (91). With a relatively high editing efficiency and few off-target effects, the CRISPR-Cas gene-editing system is able to change the biological behavior of tumor cells from the level of the genome, reduce the destruction of normal human tissue cells, and increase the survival time of patients.

Genome-wide CRISPR-Cas9 screening has produced numerous high-quality data (92, 93). Sidi Chen et al. obtained specific functionally defective mutations essential for tumor growth and metastasis using genome-wide CRISPR screening, such as Cdkn2a, Fga, and Cryba4 (94). Similarly, Ryan d. Chow et al. identified several functional suppressors and the cooccurrence and correlation of specific mutations in glioblastoma through *in vivo* CRISPR screening. They identified cooccurring driver combinations including B2m-Nf1, Mll3-Nf1 and Zc3h13-Rb1 by commutation analysis (95).

It is essential to develop models that accurately reflect the disease to better study cancer evolution and pathogenesis. For this reason, CRISPR-Cas9, with its precise gene-editing technology, is considered one of the game-changers employed to develop relevant models of cancer that best imitate human tumors. These models can be broadly divided into two categories: *in vitro* and *in vivo*. *In vitro* methods include organoid technology, which recapitulates tumor behavior via the introduction of loss-of-function mutations by knocking out or knocking down selected genes. Additionally, *in vivo* models would likely be established by introducing CRISPR-Cas9 specifically to edit oncogenic mutations or chromosomal rearrangements in tissues to further explain cancer biology thoroughly (96).

To effectively deliver CRISPR components into target cells, various systems have been developed, primarily categorized into viral-based and nonviral methods. In cancer research, viral-based delivery systems such as adeno-associated virus (AAV), lentivirus, and adenovirus are commonly used for the plasmid-based CRISPR-Cas9 system. AAV, in particular, stands out due to its small, non-enveloped single-stranded DNA structure derived from the non-pathogenic parvovirus family. It has gained attention for its minimal immunogenicity and ability to maintain gene expression in non-dividing cells, making it a promising tool for gene delivery. Additionally, AAV can serve as a donor template in homologous recombination, facilitating DNA strand exchange between similar sequences (97, 98). Although AAV vectors are not yet in clinical trials, they hold significant potential for future therapeutic applications. On the other hand, nonviral delivery systems, including hydrodynamic injection, electroporation, nanoparticles, and transposon carriers, offer greater safety despite being generally

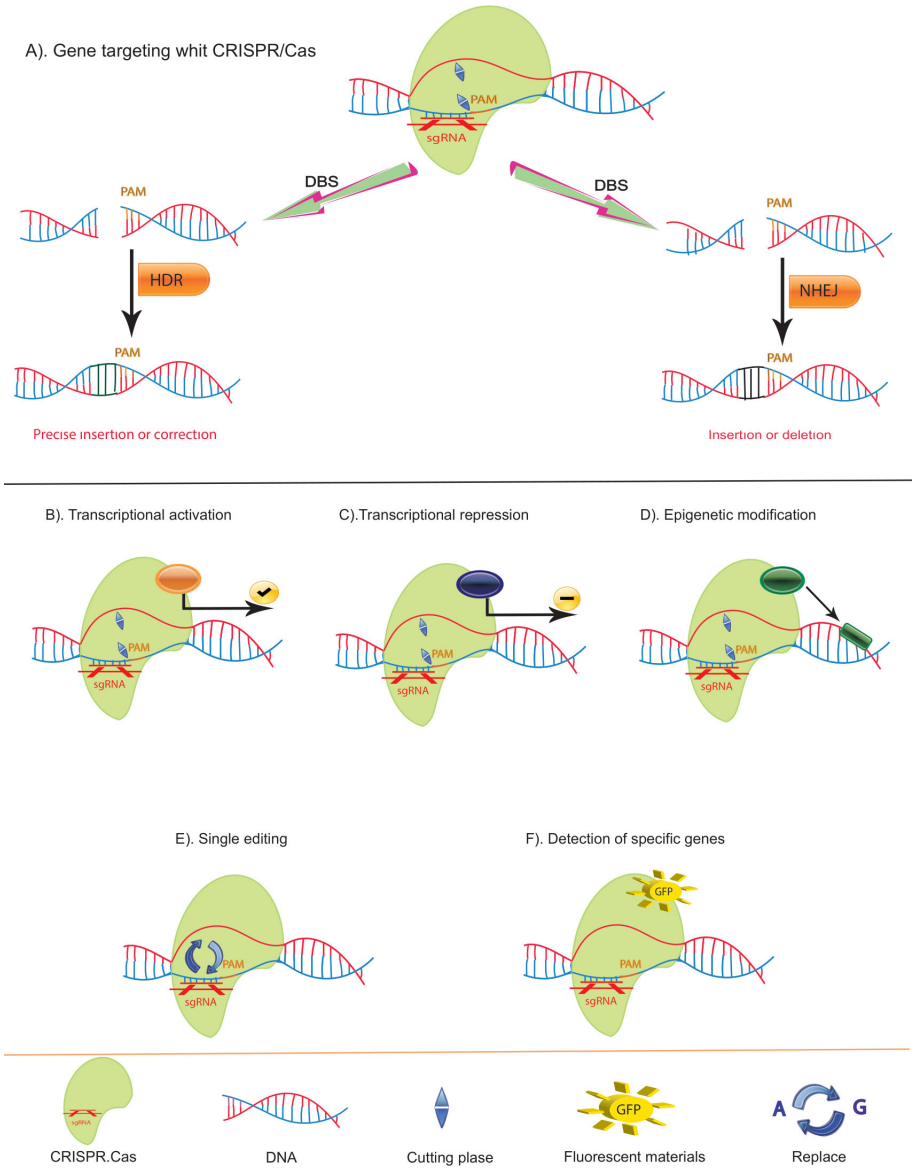


FIGURE 3
Function of CRISPR/Cas and Variants. **(A)** Double-Strand Break (DSB) Repair Mechanisms Using CRISPR/Cas: The CRISPR/Cas system can induce DSBs, which are repaired by two primary pathways: Non-Homologous End Joining (NHEJ) and Homology-Directed Repair (HDR). NHEJ often results in small insertions or deletions, leading to disruptive frameshift mutations and premature stop codons, making it ideal for gene knockouts or generating point mutants. In contrast, HDR enables precise mutations. **(B)** Cas9-VP64 Fusion: Cas9 can be fused with the VP64 transcriptional activator (yellow) to activate gene transcription by binding upstream of the transcription start site. **(C)** Cas9-KRAB Fusion: Cas9 can also be fused with the KRAB repressor (red) to downregulate gene transcription by binding to the transcription start site. **(D)** Cas9-Epigenetic Modifier Fusion: Cas9 can be linked with epigenetic modifiers (black) to alter local methylation patterns, thereby modifying gene expression epigenetically. **(E)** Cas9-Base Editors: By fusing Cas9 with base editors (purple), precise single nucleotide exchanges can be achieved without causing DSBs. **(F)** Fluorescent Cas9: If the cutting function of Cas9 is deactivated and it is equipped with a fluorescent marker, it can bind to specific DNA sequences, creating a green fluorescent signal. This enables the identification and visualization of specific sequences.

less efficient than viral vectors (96). A major challenge with nonviral methods is their lack of tissue specificity, an issue that may be mitigated by modifying the Cas9 protein accordingly.

CRISPR-based genetic and epigenetic manipulation of immune responses has emerged as a promising strategy in immunotherapy for combating cancer initiation and progression. This approach involves enhancing host immunity at specific genetic loci, boosting tumor immunogenicity, and overcoming tumor immune evasion mechanisms. Thus far, modifying immune cells ex vivo to suppress immune checkpoint expression or to introduce synthetic immune receptors, such as chimeric antigen receptors (CARs), has demonstrated efficacy in treating certain cancers like melanoma, lymphoma, liver, and lung cancer (99).Besides the success of the CRISPR-Cas9 system, the development of a nuclease-deactivated Cas9 (dCas9) variant has expanded CRISPR technologies to include epigenome engineering. By introducing two mutations—D10A and H840A—into Cas9, the wild-type system is converted into an inactivated cleavage capacity but retains RNA-guided DNA-

binding specificity (100). As was first shown with the engineered zinc finger proteins, dCas9 can be fused to a variety of effector domains that enable highly targeted and tunable transcriptional activation or repression, editing of epigenetic marks, or fluorescent tagging of endogenous genes without direct genomic modification (99).

In the area of oncology, these dCas-based tools have shown strong activation of tumor suppressor genes, such as PTEN, in breast cancer and melanoma; MASPIN, in breast and lung cancers; REPRIMO, in breast and gastric cancers; SARI, in colon cancer; and DKK3, in prostate cancer. Successful suppression of oncogenes was also attained using dCas9 in colon cancer, targeting BRAF, HER2, and MYC; pancreatic cancer, targeting KRAS; and liver cancer, targeting GRN (99). Also, several works describe that epigenome editing can be very efficient, as it has attained almost complete gene repression or robust (several-fold) gene activation with low off-target effects, which mostly depend on the effector domains' nature used (100, 101). Finally, in contrast to genome engineering by Cas9, which unavoidably leads to permanent changes, epigenetic approaches result in reversibility and thereby bypass the risk of inducing sequence changes in the target DNA—a most crucial factor when it comes to targeting tumors with high degrees of genetic instability (99). In addition, the durability of such epigenetic and transcriptional changes that are induced by dCas9 editing might depend on the specific combination of effectors or on targeted loci. Therefore, current research in epigenome engineering will have to focus on further fine-tuning the technology for the manipulation of different loci within diverse cell types with differing chromatin microenvironments.

4 Synergistic potential of CAR-T and CRISPR/Cas9

CRISPR/Cas9 technology has been extensively tested across various cell types and organisms. It has played a pivotal role in advancing CAR-T cell development and enhancing other genome editing tools (102). Notably, CRISPR/Cas9 modifications are currently under investigation in clinical trials aimed at improving CAR-T cell therapy (Table 1). Before the advent of CRISPR/Cas9, other genome editing methods like zinc finger nucleases (ZFNs) or TALENs were employed. However, CRISPR/Cas9 has surpassed these methods in terms of cost-effectiveness and practicality (103, 104). Moreover, it enables multiplex gene editing, a significant capability that facilitates the generation of universal CAR-T cells by long-term silencing of endogenous TCR, HLA class I molecules, and inhibitory checkpoints such as CTLA4 and PD1 (102–104). Research has proved that employing CRISPR/Cas9 to delete PD-1 resulted in enhanced long-term persistence and activity of CAR-T cells (preclinical study) (105), and similarly, the deletion of CTLA-4 using CRISPR/Cas9 improved the proliferation and activity of CAR-T cells (preclinical study) (106). Additionally, CRISPR/Cas9 provides a robust alternative to conventional lentiviral insertion of CARs by avoiding random genome integration and uncontrolled construct expression (107).

Some researchers have proven that, in terms of effectiveness, persistence, and reducing side effects, autologous CAR-T cells perform better than allogeneic CAR-T cells. This has prompted researchers to further optimize allogeneic CAR-T cells. Li et al. used CRISPR/Cas9 technology to knock out T-cell receptor (TCR) and HLA-I/II genes in CAR-T cells, introducing exogenous HLA-E expression to improve CAR-T cell persistence and prevent rejection in preclinical studies. A more promising approach being pursued is the development of a stable supply of universal allogeneic CAR-T cells for cancer therapy by generating iPSCs, which have virtually unlimited replicative potential and broad differentiation abilities. Wang et al. demonstrated that using CRISPR/Cas9 to integrate the CAR gene into the endogenous TCRα constant (TRAC) locus in iPSCs results in CAR-T cells with lower immunogenicity, enhanced tumor cytotoxicity, improved survival, and reduced allogeneic response risk. Furthermore, Ueda et al. enhanced these iPSC-derived CAR-T cells by editing genes with CRISPR/Cas9 to delete the diacylglycerol kinase gene and introduce IL-15 and its receptor subunit, leading to improved proliferation and increased longevity in preclinical studies.

Despite its advantages, the CRISPR/Cas9 system is not without limitations. Off-target effects can occur, potentially affecting cell fitness. However, several strategies, such as precise sgRNA design, truncated sgRNAs, chimeric DNA-RNA-based sgRNAs, and the use of different Cas9 variants, have been developed to mitigate these effects (108–111). Another limitation is the low frequency of homology-directed repair (HDR), which limits the efficiency of gene correction and addition. Enhancers of HDR and inhibitors of non-homologous end joining (NHEJ) can be employed to promote these processes (112–114). Despite these challenges, overcoming these limitations is crucial for generating effective CAR-T cells (Table 2).

To address issues such as targeting healthy cells and committing fratricide—an on-target off-tumor effect—rational selection of chimeric antigen receptors (CARs) is essential (121). In T-cell-derived malignancies, identifying suitable tumor-associated antigens (TAA) is particularly challenging, as many TAAs are shared between malignant T cells and CAR-T cells (122). To prevent self-destruction, researchers have disrupted the expression of widely expressed T-cell antigens in CAR-T cells. Specifically, clusters of differentiation 3, 5, and 7 (CD3, CD5, and CD7) have been successfully targeted using gene-editing technologies such as transcription activator-like effector nucleases (TALENs) and CRISPR/Cas9. The resulting CAR-T cells demonstrated significant antitumor efficacy while reducing the risk of fratricide, offering a promising approach to improving CAR-T cell therapies in T-cell malignancies (123–125). Pinz et al. (45). developed CD4-CAR-T cells to treat peripheral T-cell lymphomas but encountered significant fratricide, resulting in the enrichment of CD4+ CD8+ CD4-CAR-T cells. To maintain a stable CD4/CD8 CAR-T cell ratio, investigating the impact of a CD4 knockout could be beneficial. Similarly, to address the potential fratricide of anti-CD319 CAR-T cells in multiple myeloma treatment, Galetto et al. utilized TALENs technology to inactivate CD319, which is broadly expressed in activated T cells. This approach successfully prevented the loss of CD319-positive CAR-T cells during T-cell expansion, highlighting a promising strategy to mitigate fratricide in CAR-T cell therapies (126).

TABLE 1 Clinical trials of engineering CAR-T cells based on CRISPR/Cas9 technology.

NCT ID	Cancer	Target antigen	CART Cell	Patients (n)	Phase	Advantages
NCT04637763	Non-Hodgkin Lymphoma	CD19	CD19-CAR-T	72	1	Broaden applicability
NCT04502446	T cell malignancy, Diffuse Large B-Cell Lymphoma	TCR and MHC I KO via CRISPR/Cas9	CD70-CAR-T	26	1	–
NCT04035434	B-cell Lymphoma, B-cell ALL	CD19	CD19-CAR-T	227	1/2	Cost reduction
NCT04244656	Multiple Myeloma	–	BCMA-CAR-T	26	1	Enhance effector function
NCT03166878	Leukemia, Lymphoma	TCR and B2M (knock out)	CD19-CAR-T	80	1/2	Broaden applicability
NCT03398967	Leukemia, Lymphoma	TCR and B2M (knock out)	CD19/CD20/CD22-CAR-T	80	1/2	Cost reduction Broaden applicability
NCT04037566	ALL, Lymphoma	HPK1 (knock out)	CD19-CAR-T	40	1	Cost reduction Enhance effector function
NCT04438083	Renal Cell Carcinoma	–	CTX130	107	1	–
NCT03747965	Multiple Solid Tumors	PD-1 (knock out)	Mesothelin CAR-T	10	1	Enhance effector function
NCT05795595	Multiple Solid Tumors	–	CD70-CAR-T	250	1/2	–
NCT03545815	Solid Tumors	PD-1 and TCR (knock out)	Mesothelin CAR-T	10	1/2	Cost reduction Broaden applicability Enhance effector function
NCT05812326	Breast Cancer	PD-1 (knock out)	MUC1-CAR-T	15	1/2	

Data extracted from https://clinicaltrials.gov/. B2M, β-beta 2-microglobulin; CD, cluster of differentiation; HPK1, hematopoietic progenitor kinase 1; PD-1, programmed cell death protein 1; TCR, T- cell receptor.

CRISPR/Cas9 delivery into T cells can be achieved through various methods, including plasmid DNA, messenger RNA (mRNA), or delivery as ribonucleoproteins (RNPs). The efficiency of gene editing varies significantly depending on the cell type and donor, but both mRNA and RNP delivery methods have shown promise in achieving high rates of insertion or deletion mutations (indels) (127). This variability in efficiency underscores the importance of choosing the appropriate delivery method to ensure accurate editing of CAR-T cells.

Traditional CAR-T cell production relies on viral transduction of CARs, but concerns over potential side effects from integrative viruses have prompted the exploration of alternative methods. One preferred approach by several research groups involves targeted transgene integration into the TRAC gene using the CRISPR/Cas9 system. This method achieves high integration rates by delivering the repair template through AAV6 transduction or long single-stranded DNA (ssDNA) electroporation (128, 129). Another delivery method utilizes integrase-defective lentiviruses (IDLVs), known for their large genome capacity, low pathogenicity, and ability to transduce both dividing and non-dividing cells (130, 131). Despite these advantages, no preclinical trials utilizing IDLVs have been conducted thus far.

5 Integration of CAR-T and CRISPR/Cas9 in clinical settings: analysis of outcomes

CAR-T cell therapy stands out as a highly promising treatment option for refractory hematologic malignancies. Over the years, significant advancements have transformed the design of CARs (132). These include incorporating features such as coexpression of costimulatory molecules, cytokines, and suicide genes to enhance both efficacy and safety (133). Furthermore, the repertoire of tumor targets for CAR-T cells has expanded beyond CD19 to encompass a wide array of new targets. These include CD20, CD22, CD30, CD33, CD138, CD171, CEA, epidermal growth factor receptor (EGFR), EGFRvIII, ErbB, FAP, GD2, Glypican 3, Her 2, Mesothelin, and NKG2D, among others (134). However, current therapies face limitations, including off-target effects, fratricide, and challenges in identifying suitable TAA, especially in T-cell-derived cancers. These issues can lead to the destruction of healthy cells and diminished efficacy. CRISPR/Cas9 technology offers a solution by enabling precise gene editing to disrupt the expression of problematic antigens, such as CD3, CD5, and CD7, in CAR-T

10.3389/fimmu.2024.1462697

TABLE 2 Enhanced techniques and strategies for effective gene editing in CAR-T cells.

CAR-T cell	Target gene (s)	Cas9 delivery system	Gene-editing efficiency	Multiplex Efficiency	Antitumor Activity	Disadvantages	Ref
CD19-CAR-T cells PSCA-CAR-T cells	B2M PD1 TRAC TRBC	mRNA	50% NHEJ (multiplex) 70–90% NHEJ (single gene disruption)	Enhanced *in vivo* antitumor activity	Yes, *in vitro* and animal models	Multiplex gene editing led to reduced overall editing efficiencies.	(115)
CD19-CAR-T cells	LAG3	RNP	70% NHEJ (single gene Disruption)	Not applicable	Robust antigen-specific antitumor activity in cell culture and murine model	The efficiency of gene editing has room for further optimization.	(116)
GPC3-CAR-T cells	PD1	RNP	85% NHEJ (single gene disruption)	Not applicable	Enhanced *in vivo* antitumor activity, improved persistence and infiltration	Not applicable	(117)
139 CAR-T cells	DGKα DGKζ	RNP	60–70% NHEJ (single gene disruption)	Not applicable	Significant regression of tumors in a xenograft mouse model	The efficiency of gene editing has room for further optimization.	(118)
CD7-CAR-T cells CD19-CAR-T cells	CD7 TRAC	DNA plasmid	70 NHEJ (multiplex) 90% NHEJ (single gene disruption)	Not applicable	Efficacy *in vitro* and *in vivo* without induction of xenogeneic GvHD	Not applicable	(119)
CD19-CAR-T cells PSCA-CAR-T cells	TCR, HLA class I, Fas, PD1, CTLA4	lentivirus -based one-shot system/ mRNA/ RNP/mRNA	50% NHEJ (multiplex) and 90% NHEJ (single gene disruption) 82% NHEJ (single gene disruption) 76% NHEJ (single gene disruption)	Simultaneous gene editing of four loci attempted	Enhanced antitumor activity against multiple inhibitory pathways	Multiplex gene editing led to reduced overall editing efficiencies. The limited packaging capacity of lentiviral vectors led to reduced gene-editing efficiency.	(120)

B2M, β-2-microglobulin; CTLA4, cytotoxic T-lymphocyte-associated protein 4; DGKα/ζ, diacylglycerol kinase α/ζ subunit; GPC3, glypican 3; LAG3, lymphocyte activation gene 3; NHEJ, non-homologous end joining; PD1, programmed cell death protein 1; PSCA, prostate stem cell antigen; RNP, ribonucleoprotein; TRAC/TRBC, T-cell receptor α/β constant subunit.

cells. This approach enhances antitumor efficacy while reducing fratricide and off-target effects, paving the way for more effective and safer CAR-T cell therapies.

Currently, only a limited number of clinical trials are employing CRISPR/Cas9 technology in CAR-T cells. For instance, NCT04037566 represents a pioneering trial evaluating CD19 CAR-T cells with edited endogenous HPK1 in patients with relapsed/refractory leukemia or lymphoma. Another trial, NCT04637763, is a phase I study investigating the efficacy and safety of CRISPR-edited allogeneic CD19 CAR-T cells in patients with relapsed/refractory B cell non-Hodgkin lymphoma. Stadtmauer et al. recently reported on a phase 1 clinical trial focusing on the safety and feasibility of CRISPR-Cas9 gene editing in three patients with advanced cancer. In this trial, T lymphocytes were extracted from patients and genetically modified using CRISPR-Cas9 to disrupt three genes (TRAC, TRBC, and PDCD1) to enhance antitumor immunity. Additionally, a cancer-targeting transgene, NY-ESO-1, was introduced to specifically target tumors. The engineered cells were administered to patients and demonstrated good tolerance, with sustained engraftment observed throughout the study period. These promising findings lay the groundwork for future trials exploring CRISPR-engineered cancer immunotherapies (135). Finally, NCT03545815 represents a phase I clinical trial employing CRISPR/Cas9 to disable PD-1 and TCR in CAR-T cells before administering them to patients with mesothelin-positive multiple solid tumors. Research has demonstrated that CRISPR-Cas9 can effectively disrupt up to five genes simultaneously in mouse embryonic stem cells with high efficiency (136). In another study, protocols were developed to efficiently generate CAR-T cells with edits in two genes (TRAC and B2M) or three genes (TRAC, B2M, and PD-1), evaluating their antitumor activities *in vitro*. Results indicated that CRISPR-Cas9-mediated multiple gene editing is readily applicable to CAR-T cells (137). However, CRISPR-edited CAR-T cells are still in early stages, and further preclinical studies, potentially conducted under Good Laboratory Practice (GLP), are necessary to pave the way for clinical trials.

The translation of CRISPR/Cas9 technology beyond CAR-T cells into clinical applications faces significant challenges that currently hinder its successful therapeutic implementation (Additional details are given in section 4). These challenges include, but are not limited to, several key issues. One major obstacle is the occurrence of off-target modifications, where the sgRNA can sometimes match with regions similar to the target sequence, leading Cas9 to cleave unintended off-target sites. Efforts to enhance the specificity of CRISPR/Cas9 have been pursued through improved gRNA design, the development of more efficient delivery vehicles, and the creation of novel Cas9 nucleases (109, 138, 139). Notably, newly designed variants like

xCas9 and HypaCas9 appear to offer improved precision without compromising target activity (138, 140). Another concern associated with the use of CRISPR/Cas9 is its potential to introduce unintended deletions and complex genomic rearrangements into edited cells, which could pose irrecoverable genotoxicity risks in clinical applications (141). Addressing this challenge could involve strategies such as conducting whole-genome sequence analysis, employing in silico off-target prediction tools, assessing genotoxicity risks, and implementing long-term patient follow-up protocols (142, 143).

Additionally, the immunogenicity of the Cas9 protein presents another challenge that impedes the clinical adoption of the CRISPR/Cas9 system. Some individuals develop specific antibodies against the Cas9 protein, leading to T cell immune memory formation upon subsequent encounters. This immune response against Cas9 can diminish the editing efficiency and potentially lead to adverse effects. Moreover, CRISPR/Cas9 requires a specific PAM sequence (NGG) to exert its genome editing capability. The conventional Cas9 protein recognizes only a limited set of PAM sequences, restricting its versatility. However, the development of xCas9, an advanced variant, expands the range of PAM sequences recognized, thereby broadening the potential applications of the CRISPR/Cas9 system (138). Despite these advancements, the efficiency of HDR pathways for genomic insertion occasionally remains low (144). Strategies to address this issue include using single-stranded DNA templates instead of double-stranded DNA, inhibiting the NHEJ pathway, and employing advanced delivery methods like nucleofection (145, 146). Advancements in CAR-T cell design and the precision of CRISPR/Cas9 gene editing have paved the way for more effective and targeted therapies (Table 3).

Merging CRISPR/Cas9 with other emerging technologies, like mRNA vaccines and adoptive cell therapy, has the potential to dramatically impact cancer treatment. For instance, this could involve editing immune cells so they recognize and destroy cancerous cells, while mRNA vaccines prime the immune system against specific tumor-expressed antigens. Synergy in this direction may lead to more personalized, effective cancer therapies. Key findings suggest that CRISPR/Cas9 can greatly enhance the specificity and efficacy of CAR-T cells by deleting inhibitory genes or introducing new receptors that boost their function. However, challenges include ensuring the safety and precision of gene editing, preventing off-target effects, and optimizing delivery methods. Future integration of CAR-T cells with CRISPR/Cas9 may revolutionize cancer treatment by creating more potent, specific, and durable therapies. As these technologies advance, we may see next-generation CAR-T cells that are more effective against a broader range of cancers, safer, and tailored to individual patients. The future of cancer treatment may involve combining CRISPR-enhanced CAR-T cells, mRNA vaccines, and other innovative approaches to pave the way for breakthroughs in fighting cancer.

6 Personalized cancer therapies

The future of individualized therapy using combined approaches of CAR-T and CRISPR/Cas9 holds tremendous potential to revolutionize the treatment of cancer and genetic diseases. CAR-T therapy, which involves engineering a patient's T-cells to target cancer cells, has shown remarkable success in treating certain blood cancers. Meanwhile, CRISPR/Cas9 has transformed genetic research

TABLE 3 The integration of CAR-T and CRISPR/Cas9 in clinical settings.

Aspect	Description	Outcomes	Challenges	Ref
Advancements in CAR-T Design	Incorporation of costimulatory molecules, cytokines, and suicide genes to enhance efficacy and safety. Expanded tumor target repertoire.	Enhanced efficacy and safety in treating hematologic malignancies.	N/A	(132–134)
Clinical Trials of CRISPR in CAR-T	Trials like NCT04037566 and NCT04637763 are pioneering studies evaluating CRISPR-edited CAR-T cells. Stadtmauer et al.'s trial on CRISPR-Cas9 gene editing in advanced cancer patients.	Demonstrated feasibility and safety, sustained engraftment, and good tolerance.	Early stages of clinical trials; further studies needed for validation.	(135–137)
Off-target Modifications	Occurrence of unintended off-target modifications where sgRNA matches regions similar to the target sequence. Efforts to enhance specificity through improved gRNA design and novel Cas9 nucleases like xCas9 and HypaCas9.	Improved precision with xCas9 and HypaCas9 without compromising target activity.	Persistent risk of off-target effects; need for better prediction and assessment tools.	(109, 138–140)
Genotoxicity Risks	Potential for unintended deletions and complex genomic rearrangements posing genotoxicity risks. Strategies involve whole-genome sequence analysis, in silico off-target prediction, and long-term patient follow-up.	Genotoxicity risks need thorough assessment before clinical application	High genotoxicity risks require robust assessment and monitoring strategies.	(141–143)
Immunogenicity and HDR Efficiency	Cas9 protein immunogenicity leading to T cell immune memory formation. Limited PAM sequence recognition by conventional Cas9. xCas9 developed to recognize a broader range of PAM sequences. Efficiency of HDR pathways for genomic insertion is occasionally low.	Broadened potential applications with xCas9. Improved HDR efficiency with single-stranded DNA templates, NHEJ pathway inhibition, and advanced delivery methods.	Immunogenicity of Cas9 protein; limited HDR efficiency; strategies needed to overcome these challenges.	(99, 144–146)

Ref, References. "N/A" stands for "not available".

by allowing precise modifications to DNA, making it a powerful tool for correcting genetic defects. By merging these two cutting-edge technologies, we can create more effective and personalized treatments tailored to individual patients' genetic profiles. One of the most promising prospects of combining CAR-T and CRISPR/Cas9 is the enhancement of CAR-T cell therapy. CRISPR can be used to edit genes within T-cells to improve their ability to target and kill cancer cells. For example, CRISPR can knock out genes that inhibit T-cell activity or add genes that enhance their persistence and efficacy in the tumor microenvironment (147). Immune checkpoints, such as PD-1 and CTLA-4, are critical regulatory molecules in the immune system that prevent autoimmune reactions by dampening immune cell activity. However, this mechanism also allows cancer cells to evade immune responses. Immune checkpoint inhibitors work by blocking the interaction between these checkpoints and their ligands, thereby restoring T cell function. To counteract immune checkpoint inhibition in CAR-T cells, researchers have explored the combination of PD-1 inhibitors with CAR-T cell therapy. For instance, a research team engineered Mesothelin-CAR-T cells to treat pleural mesothelioma in mice and administered them alongside a PD-1 inhibitor. This combined approach effectively prolonged CAR-T cell activity, slowed tumor progression, and significantly extended median survival time (148). The feasibility of this combination therapy was further validated in a clinical trial, offering a promising strategy for enhancing cancer treatment outcomes (149). This synergy could potentially expand the success of CAR-T therapy beyond blood cancers to solid tumors, which have been challenging to treat with current CAR-T strategies. Moreover, the combined approach allows for the development of "off-the-shelf" CAR-T cells, which are derived from healthy donors rather than the patient. CRISPR/Cas9 can be used to edit these donor cells to prevent immune rejection and enhance their cancer-fighting properties (150).

This advancement would make CAR-T therapy more accessible, reducing the time and cost associated with generating personalized treatments from the patient's own cells (151). Clinical applications face limitations due to the high cost of CAR-T cell therapy, which is mainly driven by complex manufacturing processes and expensive raw materials (14). Cost reductions can be achieved by improving the efficiency of production and using strict quality control systems to ensure consistency and reduce waste. A typical production cycle for CAR-T cells takes about two weeks, but new platforms are emerging that significantly reduce this time. For example, the Novartis T-Charge platform reduces *in vitro* culture time and increases T-cell proliferation (152). Dickinson et al. (153) developed CD19-CAR-T YTB323 autoimmune cell therapy using this platform in just two days. This is different from tisagenlecleucel, as YTB323 maintains T cell regulation and improves *in vivo* expansion and anti-tumor efficacy at low doses. Similarly, Grasel Biotechnologies' FasTCAR platform has shortened the production time for CAR-T to one day. Clinical trials to evaluate the efficacy and safety of this process (NCT04638270, NCT05840107, and NCT04935580) are currently underway. In addition, CRISPR/Cas9 technology will accelerate more precise gene editing, allowing for rapid changes within a short period of time.

CRISPR can precisely correct genetic mutations in stem cells, which can then be differentiated into T-cells for CAR-T therapy. This approach not only targets cancer but also treats underlying genetic conditions, offering a dual therapeutic benefit. For example, patients with genetic immunodeficiencies could receive gene-corrected, cancer-targeting T-cells, addressing both their genetic disorder and cancer simultaneously (154). Despite these exciting prospects, challenges remain in ensuring the safety, efficacy, and ethical deployment of these therapies. Potential off-target effects of CRISPR and the long-term impacts of genetic modifications need thorough investigation. Regulatory frameworks must evolve to keep pace with technological advancements, ensuring that therapies are safe and ethically sound. Nonetheless, the combined use of CAR-T and CRISPR/Cas9 represents a promising frontier in personalized medicine, offering hope for more effective, targeted, and accessible treatments for cancer and genetic diseases.

7 Future directions

Editing CAR-T cells using CRISPR/Cas9 marks a pivotal advancement in the field of immunotherapy, addressing several critical challenges such as mitigating allogeneic reactions, overcoming tonic signaling and exhaustion within the tumor microenvironment (TME), and reducing potential toxicity. The scalability of CRISPR/Cas9 enables large-scale genetic screens, allowing researchers to efficiently and precisely investigate thousands of genes in T cells. This has led to groundbreaking preclinical studies that demonstrate the potential of CRISPR-edited CAR-T cells to significantly improve cancer treatment outcomes. As these technologies evolve, their integration into CAR-T cell therapy could revolutionize the approach to treating various forms of cancer, including those that have been resistant to conventional therapies.

Despite these promising advancements, there are still significant challenges that need to be addressed to fully realize the potential of CRISPR/Cas9 in CAR-T cell therapy. One of the primary concerns is the risk of off-target effects, where unintended genetic modifications could lead to adverse outcomes. Future research should focus on improving the precision of CRISPR/Cas9 through advancements in gRNA design, the development of novel Cas9 variants like xCas9 and HypaCas9, and the implementation of robust off-target detection methods. Additionally, optimizing the delivery systems for CRISPR/Cas9 components is crucial to ensure that the gene-editing machinery is accurately and efficiently introduced into target cells, minimizing the risk of unintended consequences.

Translating CRISPR/Cas9-engineered CAR-T cells into clinical practice also presents a set of challenges that must be carefully navigated. The transition from laboratory research to clinical application involves overcoming regulatory hurdles, designing clinical trials that accurately assess the safety and efficacy of these therapies, and scaling up production to meet the demands of a broader patient population. Establishing standardized manufacturing protocols and strict quality control measures will be essential to ensure the consistency and safety of CRISPR-edited CAR-T cells as they move toward widespread clinical use.

Looking forward, the future of CAR-T and CRISPR/Cas9 therapies could be further enhanced by integrating emerging technologies such as synthetic biology and artificial intelligence. Synthetic biology offers the potential to create programmable T cells

that can be fine-tuned for specific therapeutic purposes, while artificial intelligence could be used to optimize gene-editing strategies and predict patient responses. These innovations, combined with ongoing advancements in CRISPR/Cas9 technology, could lead to even more personalized and effective cancer treatments, ultimately transforming the landscape of cancer care and offering new hope for patients facing some of the most challenging and refractory cancers.

8 Conclusion

The integration of CAR-T cell therapy with CRISPR/Cas9 technology has revolutionized cancer treatment by enhancing efficacy and personalization. CAR-T has shown success in treating hematologic cancers, but CRISPR/Cas9 offers precision and versatility to address limitations. By engineering CAR-T cells, researchers can enhance persistence, reduce off-target effects, and even knockout immune checkpoint inhibitors. However, challenges remain, such as off-target effects, optimizing gene-editing precision, and developing robust clinical protocols. Scaling up production, ensuring patient safety, and navigating regulatory landscapes are crucial steps for translating these therapies from preclinical success to widespread clinical use. Collaboration across academic, clinical, and industry sectors is essential to overcome barriers and accelerate the development of transformative therapies. The potential of CRISPR-edited CAR-T cells to revolutionize cancer treatment remains immense.

Author contributions

MA: Conceptualization, Data curation, Validation, Visualization, Writing – original draft, Writing – review & editing. AM: Conceptualization, Validation, Visualization, Writing – original draft, Writing – review & editing. MM: Data curation, Writing – original draft, Writing – review & editing. BS: Conceptualization, Investigation, Writing – original draft, Writing – review & editing. AN: Conceptualization, Investigation, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We acknowledge the use of ChatGPT, version (v2), a large language model developed by OpenAI, for correcting grammatical structure and syntax in this article. All corrections and suggestions provided by the AI tool were reviewed and verified by the authors.

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.

Publisher's note

References

1. Jeschke J. Advances in cancer detection and diagnosis: from liquid biopsies and molecular biomarkers to opportunistic intratumoral bacteria. *Curr Opin Oncol.* (2023) 35:114. doi: 10.1097/CCO.0000000000000930

2. Hofmarcher T, Lindgren P, Wilking N, Jönsson B. The cost of cancer in Europe 2018. *Eur J Cancer.* (2020) 129:41–9. doi: 10.1016/j.ejca.2020.01.011

3. Spokeviciute B, Kholia S, Brizzi MF. Chimeric antigen receptor (CAR) T-cell therapy: harnessing extracellular vesicles for enhanced efficacy. *Pharmacol Res.* (2024) 13:107352. doi: 10.1016/j.phrs.2024.107352

4. Li X, Li W, Xu L, Song Y. Chimeric antigen receptor-immune cells against solid tumors: Structures, mechanisms, recent advances, and future developments. *Chin Med J.* (2024) 137:1285–302. doi: 10.1097/CM9.0000000000002818

5. Pule MA, Savoldo B, Myers GD, Rossig C, Russell HV, Dotti G, et al. Virus-specific T cells engineered to coexpress tumor-specific receptors: persistence and antitumor activity in individuals with neuroblastoma. *Nat Med.* (2008) 14:1264. doi: 10.1038/nm.1882

6. Shademan B, Karamad V, Nourazarian A, Avcı CB. CAR T cells: cancer cell surface receptors are the target for cancer therapy. *Advanced Pharm bulletin.* (2022) 12:476. doi: 10.34172/apb.2022.051

7. Yan W, Liu Z, Liu J, Xia Y, Hu K, Yu J. Application of chimeric antigen receptor T cells in the treatment of hematological Malignancies. *BioMed Res Int.* (2020) 2020:1–9. doi: 10.1155/2020/4241864

8. Goldsmith SR, Ghobadi A, DiPersio JF. Hematopoeitic cell transplantation and CAR T-cell therapy: complements or competitors? *Front Oncol.* (2020) 10:2904. doi: 10.3389/fonc.2020.608916

9. Yang X, Wang GX, Zhou JF. CAR T cell therapy for hematological Malignancies. *Curr Med Sci.* (2019) 39:874–82. doi: 10.1007/s11596-019-2118-z

10. Tao R, Han X, Bai X, Yu J, Ma Y, Chen W, et al. Revolutionizing cancer treatment: enhancing CAR-T cell therapy with CRISPR/Cas9 gene editing technology. *Front Immunol.* (2024) 15:1354825. doi: 10.3389/fimmu.2024.1354825

11. Fan J, Adams A, Sieg N, Heger JM, Gödel P, Kutsch N, et al. Potential synergy between radiotherapy and CAR T-cells-a multicentric analysis of the role of radiotherapy in the combination of CAR T cell therapy. *Radiotherapy Oncol.* (2023) 183:109580. doi: 10.1016/j.radonc.2023.109580

12. Wang AX, Ong XJ, D'Souza C, Neeson PJ, Zhu JJ. Combining chemotherapy with CAR-T cell therapy in treating solid tumors. *Front Immunol.* (2023) 14:1140541. doi: 10.3389/fimmu.2023.1140541

13. Wang A, Lv T, Song Y. Tandem CAR-T cells targeting MUC1 and PSCA combined with anti-PD-1 antibody exhibit potent preclinical activity against non-small cell lung cancer. *Cell Immunol.* (2023) 391:104760. doi: 10.1016/j.cellimm.2023.104760

14. Miao L, Zhang Z, Ren Z, Tang F, Li Y. Obstacles and coping strategies of CAR-T cell immunotherapy in solid tumors. *Front Immunol.* (2021) 12:687822. doi: 10.3389/fimmu.2021.687822

15. Shademan B, Nourazarian A, Hajazimian S, Isazadeh A, Biray Avci C, Oskouee MA. CRISPR technology in gene-editing-based detection and treatment of SARS-CoV-2. *Front Mol Biosciences*. (2022) 8:772788. doi: 10.3389/fmolb.2021.772788

16. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. (2013) 339:819–23. doi: 10.1126/science.1231143

17. Ghaffari S, Khalili N, Rezaei N. CRISPR/Cas9 revitalizes adoptive T-cell therapy for cancer immunotherapy. *J Exp Clin Cancer Res*. (2021) 40:269. doi: 10.1186/s13046-021-02076-5

18. Lana MG, Strauss BE. Production of lentivirus for the establishment of CAR-T cells. *Methods Mol Biol*. (2020) 2086:61–7. doi: 10.1007/978-1-0716-0146-4_4

19. Foster JB, Choudhari N, Perazzelli J, Storm J, Hofmann TJ, Jain P, et al. Purification of mRNA encoding chimeric antigen receptor is critical for generation of a robust T-cell response. *Hum Gene Ther*. (2019) 30:168–78. doi: 10.1089/hum.2018.145

20. Safarzadeh Kozani P, Safarzadeh Kozani P, Rahbarizadeh F. Novel antigens of CAR T cell therapy: new roads; old destination. *Transl Oncol*. (2021) 14:101079. doi: 10.1016/j.tranon.2021.101079

21. Safarzadeh Kozani P, Safarzadeh Kozani P, O'Connor RS. In like a lamb; out like a lion: marching CAR T cells toward enhanced efficacy in B-ALL. *Mol Cancer Ther*. (2021) 20:1223–33. doi: 10.1158/1535-7163.MCT-20-1089

22. Hashem Boroojerdi M, Rahbarizadeh F, Safarzadeh Kozani P, Kamali E, Safarzadeh KP. Strategies for having a more effective and less toxic CAR T-cell therapy for acute lymphoblastic leukemia. *Med Oncol*. (2020) 37:100. doi: 10.1007/s12032-020-01416-3

23. Safarzadeh Kozani P, Safarzadeh Kozani P, Rahbarizadeh F. CAR-T cell therapy in T-cell Malignancies: Is success a low-hanging fruit? *Stem Cell Res Ther*. (2021) 12:527. doi: 10.1186/s13287-021-02595-0

24. Safarzadeh Kozani P, Safarzadeh Kozani P, Rahbarizadeh F, Khoshtinat NS. Strategies for Dodging the obstacles in CAR T cell therapy. *Front Oncol*. (2021) 11:627549. doi: 10.3389/fonc.2021.627549

25. Maude SL, Laetsch TW, Buechner J, Rives S, Boyer M, Bittencourt H, et al. Tisagenlecleucel in children and young adults with B-cell lymphoblastic leukemia. *N Engl J Med*. (2018) 378:439–48. doi: 10.1056/NEJMoa1709866

26. Maude SL, Frey N, Shaw PA, Aplenc R, Barrett DM, Bunin NJ, et al. Chimeric antigen receptor T cells for sustained remissions in leukemia. *N Engl J Med*. (2014) 371:1507–17. doi: 10.1056/NEJMoa1407222

27. Safarzadeh Kozani P, Safarzadeh Kozani P, Ahmadi Najafabadi M, Yousefi F, Mirarefin SM, Rahbarizadeh F. Recent advances in solid tumor CAR-T cell therapy: driving tumor cells from hero to zero? *Front Immunol*. (2022) 13:795164. doi: 10.3389/fimmu.2022.795164

28. Holstein SA, Lunning MA. CAR T-cell therapy in hematologic Malignancies: a voyage in progress. *Clin Pharmacol Ther*. (2020) 107:112–22. doi: 10.1002/cpt.v107.1

29. Zhang X, Zhu L, Zhang H, Chen S, Xiao Y. CAR-T cell therapy in hematological Malignancies: current opportunities and challenges. *Front Immunol*. (2022) 13:927153. doi: 10.3389/fimmu.2022.927153

30. Pan J, Yang JF, Deng BP, Zhao XJ, Zhang X, Lin YH, et al. High efficacy and safety of low-dose CD19-directed CAR-T cell therapy in 51 refractory or relapsed B acute lymphoblastic leukemia patients. *Leukemia*. (2017) 31:2587–93. doi: 10.1038/leu.2017.145

31. Sang W, Shi M, Yang J, Cao J, Xu L, Yan D, et al. Phase II trial of co-administration of CD19-and CD20-targeted chimeric antigen receptor T cells for relapsed and refractory diffuse large B cell lymphoma. *Cancer Med*. (2020) 9:5827–38. doi: 10.1002/cam4.v9.16

32. Olejniczak SH, Stewart CC, Donohue K, Czuczman MS. A quantitative exploration of surface antigen expression in common B-cell Malignancies using flow cytometry. *Immunol investigations*. (2006) 35:93–114. doi: 10.1080/08820130500496878

33. Haso W, Lee DW, Shah NN, Stetler-Stevenson M, Yuan CM, Pastan IH, et al. Anti-CD22–chimeric antigen receptors targeting B-cell precursor acute lymphoblastic leukemia. *Blood J Am Soc Hematology*. (2013) 121:1165–74. doi: 10.1182/blood-2012-06-438002

34. Pan J, Niu Q, Deng B, Liu S, Wu T, Gao Z, et al. CD22 CAR T-cell therapy in refractory or relapsed B acute lymphoblastic leukemia. *Leukemia*. (2019) 33:2854–66. doi: 10.1038/s41375-019-0488-7

35. Fry TJ, Shah NN, Orentas RJ, Stetler-Stevenson M, Yuan CM, Ramakrishna S, et al. CD22-targeted CAR T cells induce remission in B-ALL that is naive or resistant to CD19-targeted CAR immunotherapy. *Nat Med*. (2018) 24:20–8. doi: 10.1038/nm.4441

36. Zhu H, Deng H, Mu J, Lyu C, Jiang Y, Deng Q. Anti-CD22 CAR-T cell therapy as a salvage treatment in B cell Malignancies refractory or relapsed after anti-CD19 CAR-T therapy. *OncoTargets Ther*. (2021) 2:4023–37. doi: 10.2147/OTT.S312904

37. Tan Y, Cai H, Li C, Deng B, Song W, Ling Z, et al. A novel full-human CD22-CAR T cell therapy with potent activity against CD22low B-ALL. *Blood Cancer J*. (2021) 11:71. doi: 10.1038/s41408-021-00465-9

38. Qasim W, Zhan H, Samarasinghe S, Adams S, Amrolia P, Stafford S, et al. S. Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Sci Trans Med*. (2017) 9:eaaj2013. doi: 10.1126/scitranslmed.aaj2013

39. Tashiro H. CAR T-cell therapy for T cell Malignancies: challenges and recent advances. *Japanese J Clin Hematology*. (2024) 65:644–51. doi: 10.11406/rinketsu.65.644

40. Li S, Wang X, Yuan Z, Liu L, Luo L, Li Y, et al. Eradication of T-ALL cells by CD7-targeted universal CAR-T cells and initial test of ruxolitinib-based CRS management. *Clin Cancer Res*. (2021) 27:1242–6. doi: 10.1158/1078-0432.CCR-20-1271

41. Pan J, Tan Y, Wang G, Deng B, Ling Z, Song W, et al. Donor-derived CD7 chimeric antigen receptor T cells for T-cell acute lymphoblastic leukemia: first-in-human, phase I trial. *J Clin Oncol*. (2021) 39:3340–51. doi: 10.1200/JCO.21.00389

42. Xie L, Ma L, Liu S, Chang L, Wen F. Chimeric antigen receptor T cells targeting CD7 in a child with high-risk T-cell acute lymphoblastic leukemia. *Int Immunopharmacology*. (2021) 96:107731. doi: 10.1016/j.intimp.2021.107731

43. Wada M, Zhang H, Fang L, Feng J, Tse CO, Zhang W, et al. Characterization of an anti-CD5 directed CAR T-cell against T-cell Malignancies. *Stem Cell Rev Rep*. (2020) 16:369–84. doi: 10.1007/s12015-019-09937-9

44. Feng J, Xu H, Cinquina A, Wu Z, Chen Q, Zhang P, et al. Treatment of aggressive T cell lymphoblastic lymphoma/leukemia using anti-CD5 CAR T cells. *Stem Cell Rev Rep*. (2021) 17:652–61. doi: 10.1007/s12015-020-10092-9

45. Pinz K, Liu H, Golightly M, Jares A, Lan F, Zieve GW, et al. Preclinical targeting of human T-cell Malignancies using CD4-specific chimeric antigen receptor (CAR)-engineered T cells. *Leukemia*. (2016) 30:701–7. doi: 10.1038/leu.2015.311

46. Alcantara M, Tesio M, June CH, Houot R. CAR T-cells for T-cell Malignancies: challenges in distinguishing between therapeutic, normal, and neoplastic T-cells. *Leukemia*. (2018) 32:2307–15. doi: 10.1038/s41375-018-0285-8

47. Maciocia PM, Wawrzyniecka PA, Maciocia NC, Burley A, Karpanasamy T, Devereaux S, et al. Anti-CCR9 chimeric antigen receptor T cells for T-cell acute lymphoblastic leukemia. *Blood J Am Soc Hematology*. (2022) 140:25–37. doi: 10.1182/blood.2021013648

48. Mardiana S, Gill S. CAR T cells for acute myeloid leukemia: state of the art and future directions. *Front Oncol*. (2020) 10:697. doi: 10.3389/fonc.2020.00697

49. Bakker AB, van den Oudenrijn S, Bakker AQ, Feller N, van Meijer M, Bia JA, et al. C-type lectin-like molecule-1: a novel myeloid cell surface marker associated with acute myeloid leukemia. *Cancer Res*. (2004) 64:8443–50. doi: 10.1158/0008-5472.CAN-04-1659

50. Wang J, Chen S, Xiao W, Li W, Wang L, Yang S, et al. CAR-T cells targeting CLL-1 as an approach to treat acute myeloid leukemia. *J Hematol Oncol*. (2018) 11:1–3. doi: 10.1186/s13045-017-0553-5

51. Lin G, Zhang Y, Yu L, Wu D. Cytotoxic effect of CLL–1 CAR–T cell immunotherapy with PD–1 silencing on relapsed/refractory acute myeloid leukemia. *Mol Med Rep*. (2021) 23:1–. doi: 10.3892/mmr.2021.11847

52. Xie G, Ivica NA, Jia B, Li Y, Dong H, Liang Y, et al. CAR-T cells targeting a nucleophosmin neoepitope exhibit potent specific activity in mouse models of acute myeloid leukaemia. *Nat Biomed Engineering*. (2021) 5:399–413. doi: 10.1038/s41551-020-00625-5

53. Sauer T, Parikh K, Sharma S, Omer B, Sedloev D, Chen Q, et al. CD70-specific CAR T cells have potent activity against acute myeloid leukemia without HSC toxicity. *Blood J Am Soc Hematology*. (2021) 138:318–30. doi: 10.1182/blood.2020008221

54. Riether C, Pabst T, Höpner S, Bacher U, Hinterbrandner M, Banz Y, et al. Targeting CD70 with cusatuzumab eliminates acute myeloid leukemia stem cells in patients treated with hypomethylating agents. *Nat Med*. (2020) 26:1459–67. doi: 10.1038/s41591-020-0910-8

55. John S, Chen H, Deng M, Gui X, Wu G, Chen W, et al. A novel anti-LILRB4 CAR-T cell for the treatment of monocytic AML. *Mol Ther*. (2018) 26:2487–95. doi: 10.1016/j.ymthe.2018.08.001

56. Jetani H, Navarro-Bailón A, Maucher M, Frenz S, Verbruggen C, Yeguas A, et al. Siglec-6 is a novel target for CAR T-cell therapy in acute myeloid leukemia. *Blood J Am Soc Hematology*. (2021) 138:1830–42. doi: 10.1182/blood.2020009192

57. Kim MY, Yu K-R, Kenderian SS, Ruella M, Chen S, Shin TH, et al. Genetic Inactivation of CD33 in hematopoietic stem cells to enable CAR T cell immunotherapy for acute myeloid leukemia. *Cell*. (2018) 173:1439–1453.e19. doi: 10.1016/j.cell.2018.05.013

58. Ferreri CJ, Bhutani M. Mechanisms and management of CAR T toxicity. *Front Oncol*. (2024) 14:1396490. doi: 10.3389/fonc.2024.1396490

59. Zhang Y, Qin D, Shou AC, Liu Y, Wang Y, Zhou L. Exploring CAR-T cell therapy side effects: mechanisms and management strategies. *J Clin Med*. (2023) 12:6124. doi: 10.3390/jcm12196124

60. Yáñez L, Sánchez-Escamilla M, Perales MA. CAR T cell toxicity: current management and future directions. *Hemasphere*. (2019) 3:e186. doi: 10.1097/HS9.0000000000000186

61. Gust J, Hay KA, Hanafi L-A, Li D, Myerson D, Gonzalez-Cuyar LF, et al. Endothelial activation and blood-brain barrier disruption in neurotoxicity after adoptive immunotherapy with CD19 CAR-T cells. *Cancer Discovery*. (2017) 7:1404–19. doi: 10.1158/2159-8290.CD-17-0698

62. Parker KR, Migliorini D, Perkey E, Yost KE, Bhaduri A, Bagga P, et al. Single-cell analyses identify brain mural cells expressing CD19 as potential off-tumor targets for CAR-T immunotherapies. *Cell*. (2020) 183:126–142.e17. doi: 10.1016/j.cell.2020.08.022

63. Neelapu SS, Tummala S, Kebriaei P, Wierda W, Gutierrez C, Locke FL, et al. Chimeric antigen receptor T-cell therapy—Assessment and management of toxicities. *Nat Rev Clin Oncol*. (2018) 15:47–62. doi: 10.1038/nrclinonc.2017.148

64. Hunter BD, Jacobson CA. CAR T-cell associated neurotoxicity: mechanisms, clinicopathologic correlates, and future directions. *J Natl Cancer Inst*. (2019) 111:646–54. doi: 10.1093/jnci/djz017

65. Freyer CW, Porter DL. Cytokine release syndrome and neurotoxicity following CAR T-cell therapy for hematologic Malignancies. *J Allergy Clin Immunol*. (2020) 146:940–8. doi: 10.1016/j.jaci.2020.07.025

66. Sheth VS, Gauthier J. Taming the beast: CRS and ICANS after CAR T-cell therapy for ALL. *Bone Marrow Transplant*. (2021) 56:552–66. doi: 10.1038/s41409-020-01134-4

67. Barbar T, Jaffer Sathick I. Tumor lysis syndrome. *Adv Chronic Kidney Dis*. (2021) 28:438–446.e1. doi: 10.1053/j.ackd.2021.09.007

68. Maus MV, Haas AR, Beatty GL, Albelda SM, Levine BL, Liu X, et al. T cells expressing chimeric antigen receptors can cause anaphylaxis in humans. *Cancer Immunol Res*. (2013) 1:26–31. doi: 10.1158/2326-6066.CIR-13-0006

69. Kochenderfer JN, Dudley ME, Carpenter RO, Kassim SH, Rose JJ, Telford WG, et al. Donor-derived CD19-targeted T cells cause regression of Malignancy persisting after allogeneic hematopoietic stem cell transplantation. *Blood*. (2013) 122:4129–39. doi: 10.1182/blood-2013-08-519413

70. Morgan RA, Yang JC, Kitano M, Dudley ME, Laurencot CM, Rosenberg SA. Case report of a serious adverse event following the administration of T cells transduced with a chimeric antigen receptor recognizing ERBB2. *Mol Ther*. (2010) 18:843–51. doi: 10.1038/mt.2010.24

71. Thistlethwaite FC, Gilham DE, Guest RD, Rothwell DG, Pillai M, Burt DJ, et al. The clinical efficacy of first-generation carcinoembryonic antigen (CEACAM5)-specific CAR T cells is limited by poor persistence and transient pre-conditioning-dependent respiratory toxicity. *Cancer Immunol Immunother*. (2017) 66:1425–36. doi: 10.1007/s00262-017-2034-7

72. Ramos CA, Savoldo B, Torrano V, Ballard B, Zhang H, Dakhova O, et al. Clinical responses with T lymphocytes targeting Malignancy-associated κ light chains. *J Clin Invest*. (2016) 126:2588–96. doi: 10.1172/JCI86000

73. Ranganathan R. Chimeric antigen receptor T cells targeting the lambda light chain of human immunoglobulin as a viable target for B cell non-Hodgkin lymphoma. *J Clin Oncol*. (2018) 36:12079. doi: 10.1200/JCO.2018.36.15_suppl.12079

74. Li H, Zhao Y. Increasing the safety and efficacy of chimeric antigen receptor T cell therapy. *Protein Cell*. (2017) 8:573–89. doi: 10.1007/s13238-017-0411-9

75. Chu F, Cao J, Neelalpu SS. Versatile CAR T-cells for cancer immunotherapy. *Contemp Oncol (Poznan Poland)*. (2018) 22:73–80. doi: 10.5114/wo.2018.73892

76. Kochenderfer JN, Dudley ME, Feldman SA, Wilson WH, Spaner DE, Maric I, et al. B-cell depletion and remissions of Malignancy along with cytokine-associated toxicity in a clinical trial of anti-CD19 chimeric-antigen-receptor-transduced T cells. *Blood*. (2012) 119:2709–20. doi: 10.1182/blood-2011-10-384388

77. Brentjens RJ, Davila ML, Riviere I, Park J, Wang X, Cowell LG, et al. CD19-targeted T cells rapidly induce molecular remissions in adults with chemotherapy-refractory acute lymphoblastic leukemia. *Sci Transl Med*. (2013) 5:177ra38. doi: 10.1126/scitranslmed.3005930

78. Finney OC, Brakke HM, Rawlings-Rhea S, Hicks R, Doolittle D, Lopez M, et al. CD19 CAR T cell product and disease attributes predict leukemia remission durability. *J Clin Investig*. (2019) 129:2123–32. doi: 10.1172/JCI125423

79. Davila ML, Riviere I, Wang X, Bartido S, Park J, Curran K, et al. Efficacy and toxicity management of 19-28z CAR T cell therapy in B cell acute lymphoblastic leukemia. *Sci Transl Med*. (2014) 6:224ra25. doi: 10.1126/scitranslmed.3008226

80. Maalej KM, Merhi M, Inchakalody VP, Mestiri S, Alam M, Maccalli C, et al. CAR-cell therapy in the era of solid tumor treatment: current challenges and emerging therapeutic advances. *Mol Cancer*. (2023) 22:20. doi: 10.1186/s12943-023-01723-z

81. Pan K, Farrukh H, Chittepu VC, Xu H, Pan CX, Zhu Z. CAR race to cancer immunotherapy: from CAR T, CAR NK to CAR macrophage therapy. *J Exp Clin Cancer Res*. (2022) 41:119. doi: 10.1186/s13046-022-02327-z

82. Furtado RN. Gene editing: the risks and benefits of modifying human DNA. *Rev Bioetica*. (2019) 27:223–33. doi: 10.1590/1983-80422019272304

83. Novellasdemunt L, Foglizzo V, Cuadrado L, Antas P, Kucharska A, Encheva V, et al. USP7 is a tumor-specific WNT activator for APC-mutated colorectal cancer by mediating β-catenin deubiquitination. *Cell Rep*. (2017) 21:612–27. doi: 10.1016/j.celrep.2017.09.072

84. Brunet E, Jasin M. Induction of chromosomal translocations with CRISPR-Cas9 and other nucleases: understanding the repair mechanisms that give rise to translocations. *Chromosome Translocation*. (2018) 1044:15–25.

85. Costa JR, Bejcek BE, McGee JE, Fogel AI, Brimacombe KR, Ketteler R. Genome editing using engineered nucleases and their use in genomic screening. In: *Assay Guidance Manual* (2017).

86. Shademan B, Masjedi S, Karamad V, Isazadeh A, Sogutlu F, Rad MH, et al. CRISPR technology in cancer diagnosis and treatment: opportunities and challenges. *Biochem Genet*. (2022) 60:1446–70. doi: 10.1007/s10528-022-10193-9

87. Gupta RM, Musunuru K. Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. *J Clin Invest*. (2014) 124:4154–61. doi: 10.1172/JCI72992

88. Sekine R, Kawata T, Muramoto T. CRISPR/Cas9 mediated targeting of multiple genes in Dictyostelium. *Sci Rep*. (2018) 8:8471. doi: 10.1038/s41598-018-26756-z

89. Zhan T, Rindtorff N, Betge J, Ebert MP, Boutros M. CRISPR/Cas9 for cancer research and therapy. In: *Seminars in Cancer Biology*, vol. 55. Academic Press (2019). p. 106–19.

90. Martinez-Lage M, Torres-Ruiz R, Puig-Serra P, Moreno-Gaona P, Martin MC, Moya FJ, et al. *In vivo* CRISPR/Cas9 targeting of fusion oncogenes for selective elimination of cancer cells. *Nat Commun*. (2020) 11:5060. doi: 10.1038/s41467-020-18875-x

91. Fellmann C, Gowen BG, Lin PC, Doudna JA, Corn JE. Cornerstones of CRISPR-Cas in drug discovery and therapy. *Nat Rev Drug Discovery*. (2017) 16:89–100. doi: 10.1038/nrd.2016.238

92. Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet*. (2017) 49:1779–84. doi: 10.1038/ng.3984

93. Behan FM, Iorio F, Picco G, Gonçalves E, Beaver CM, Migliardi G, et al. Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature*. (2019) 568:511–6. doi: 10.1038/s41586-019-1103-9

94. Chen S, Sanjana NE, Zheng K, Shalem O, Lee K, Shi X, et al. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell*. (2015) 160:1246–60. doi: 10.1016/j.cell.2015.02.038

95. Chow RD, Guzman CD, Wang G, Schmidt F, Youngblood MW, Ye L, et al. AAV-mediated direct *in vivo* CRISPR screen identifies functional suppressors in glioblastoma. *Nat Neurosci*. (2017) 20:1329–41. doi: 10.1038/nn.4620

96. Zhao Z, Li C, Tong F, Deng J, Huang G, Sang Y. Review of applications of CRISPR-Cas9 gene-editing technology in cancer research. *Biol Procedures Online*. (2021) 23:1–3. doi: 10.1186/s12575-021-00151-x

97. Asad AS, Moreno Ayala MA, Gottardo MF, Zuccato C, Nicola Candia AJ, Zanetti FA, et al. Viral gene therapy for breast cancer: progress and challenges. *Expert Opin Biol Ther*. (2017) 17:945–59. doi: 10.1080/14712598.2017.1338684

98. Gaj T, Epstein BE, Schaffer DV. Genome engineering using adeno-associated virus: basic and clinical research applications. *Mol Ther*. (2016) 24:458–64. doi: 10.1038/mt.2015.151

99. Alves E, Taifour S, Dolcetti R, Chee J, Nowak AK, Gaudieri S, et al. Reprogramming the anti-tumor immune response via CRISPR genetic and epigenetic editing. *Mol Therapy-Methods Clin Dev*. (2021) 21:592–606. doi: 10.1016/j.omtm.2021.04.009

100. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. (2013) 152:1173–83. doi: 10.1016/j.cell.2013.02.022

101. Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, PR Iyer E, et al. Highly efficient Cas9-mediated transcriptional programming. *Nat Methods*. (2015) 12:326–8. doi: 10.1038/nmeth.3312

102. Ureña-Bailén G, Lamsfus-Calle A, Daniel-Moreno A, Raju J, Schlegel P, Seitz C, et al. CRISPR/Cas9 technology: towards a new generation of improved CAR-T cells for anticancer therapies. *Briefings Funct Genomics*. (2020) 19:191–200. doi: 10.1093/bfgp/elz039

103. Mali P, Esvelt KM, Church GM. Cas9 as a versatile tool for engineering biology. *Nat Methods*. (2013) 10:957–63. doi: 10.1038/nmeth.2649

104. Teimourian S, Abdollahzadeh R. Technology developments in biological tools for targeted genome surgery. *Biotechnol Lett*. (2015) 37:29–39. doi: 10.1007/s10529-014-1656-5

105. Dötsch S, Svec M, Schober K, Hammel M, Wanisch A, Gökmen F, et al. Long-term persistence and functionality of adoptively transferred antigen-specific T cells with genetically ablated PD-1 expression. *Proc Natl Acad Sci*. (2023) 120:e2200626120. doi: 10.1073/pnas.2200626120

106. Agarwal S, Aznar MA, Rech AJ, Good CR, Kuramitsu S, Da T, et al. Deletion of the inhibitory co-receptor CTLA-4 enhances and invigorates chimeric antigen receptor T cells. *Immunity*. (2023) 56:2388–407. doi: 10.1016/j.immuni.2023.09.001

107. Eyquem J, Mansilla-Soto J, Giavridis T, van der Stegen SJ, Hamieh M, Cunanan KM, et al. Targeting a CAR to the TRAC locus with CRISPR/Cas9 enhances tumour rejection. *Nature*. (2017) 543:113. doi: 10.1038/nature21405

108. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*. (2014) 32:279. doi: 10.1038/nbt.2808

109. Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, et al. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*. (2016) 529:490–5. doi: 10.1038/nature16526

110. Yin H, Song C-Q, Suresh S, Kwan SY, Wu Q, Walsh S, et al. Partial DNA-guided Cas9 enables genome editing with reduced off-target activity. *Nat Chem Biol*. (2018) 14:311. doi: 10.1038/nchembio.2559

111. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. *Science*. (2016) 351:84–8. doi: 10.1126/science.aad5227

112. Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol*. (2015) 33:538. doi: 10.1038/nbt.3190

113. Paulsen S, Mandal BP, Frock R, Boyraz B, Yadav R, Upadhyayula S, et al. Ectopic expression of RAD52 and dn53BP1 improves homology-directed repair during

CRISPR–Cas9 genome editing. *Nat BioMed Eng*. (2017) 1:878–88. doi: 10.1038/s41551-017-0145-2

114. Song J, Yang D, Xu J, Zhu T, Chen YE, Zhang J. RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat Commun*. (2016) 7:10548. doi: 10.1038/ncomms10548

115. Ren J, Liu X, Fang C, Jiang S, June CH, Zhao Y. Multiplex genome editing to generate universal CAR T cells resistant to PD1 inhibition. *Clin Cancer Res*. (2017) 23:2255–66. doi: 10.1158/1078-0432.CCR-16-1300

116. Zhang Y, Zhang X, Cheng C, Mu W, Liu X, Li N, et al. CRISPR-Cas9 mediated LAG-3 disruption in CAR-T cells. *Front Med*. (2017) 11:554–62. doi: 10.1007/s11684-017-0543-6

117. Guo X, Jiang H, Shi B, Zhou M, Zhang H, Shi Z, et al. Disruption of PD-1 enhanced the anti-tumor activity of chimeric antigen receptor T cells against hepatocellular carcinoma. *Front Pharmacol*. (2018) 9:1118. doi: 10.3389/fphar.2018.01118

118. Jung IY, Kim YY, Yu HS, Lee M, Kim S, Lee J. CRISPR/Cas9-mediated knockout of DGK improves antitumor activities of human T cells. *Cancer Res*. (2018) 78:4692–703. doi: 10.1158/0008-5472.CAN-18-0030

119. Cooper ML, Choi J, Staser K, Ritchey JK, Devenport JM, Eckardt K, et al. An "off-the-shelf" fratricide-resistant CAR-T for the treatment of T cell hematologic Malignancies. *Leukemia*. (2018) 32:1970–83. doi: 10.1038/s41375-018-0065-5

120. Ren J, Zhang X, Liu X, Fang C, Jiang S, June CH, et al. A versatile system for rapid multiplex genome-edited CAR T cell generation. *Oncotarget*. (2017) 8:17002. doi: 10.18632/oncotarget.15218

121. Jung I-Y, Lee J. Unleashing the therapeutic potential of CAR-T cell therapy using gene-editing technologies. *Mol Cells*. (2018) 41:717–23. doi: 10.14348/molcells.2018.0242

122. Scherer LD, Brenner MK, Mamonkin M. Chimeric antigen receptors for T-cell Malignancies. *Front Oncol*. (2019) 9:126. doi: 10.3389/fonc.2019.00126

123. Cooper ML, Choi J, Staser K, Ritchey JK, Devenport JM, Eckardt K, et al. An "off-the-shelf" fratricide-resistant CAR-T for the treatment of T cell hematologic Malignancies. *Leukemia*. (2018) 32:1970–83. doi: 10.1038/s41375-018-0065-5

124. Fleischer LC, Raikar SS, Moot R, Knight KA, Doering CB, Spencer HT. Engineering CD5-targeted chimeric antigen receptors and edited T cells for the treatment of T-cell leukemia. *Blood*. (2017) 130:1914. doi: 10.1182/blood.V130.Suppl_1.1914.1914

125. Rasaiyaah J, Georgiadis C, Preece R, Preece R, Mock U, Qasim W. TCRαβ/CD3 disruption enables CD3-specific antileukemic T cell immunotherapy. *JCI Insight*. (2018) 3:e99442. doi: 10.1172/jci.insight.99442

126. Galetto R, Chion-Sotinel I, Gouble A, Smith J. Bypassing the constraint for chimeric antigen receptor (CAR) development in T-cells expressing the targeted antigen: improvement of anti-CS1 CAR activity in allogenic TCRa/CS1 double knockout T-cells for the treatment of multiple myeloma (MM). *Blood*. (2015) 126:116. doi: 10.1182/blood.V126.23.116.116

127. Antony JS, Haque AKMA, Lamsfus-Calle A, Daniel-Moreno A, Mezger M, Kormann MS. CRISPR/Cas9 system: a promising technology for the treatment of inherited and neoplastic hematological diseases. *Adv Cell Gene Ther*. (2018) 1:e10. doi: 10.1002/acg2.v1.1

128. Cavazzana-Calvo M, Payen E, Negre O, Wang G, Hehir K, Fusil F, et al. Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. *Nature*. (2010) 467:318–22. doi: 10.1038/nature09328

129. Roth TL, Puig-Saus C, Yu R, Shifrut E, Carnevale J, Li PJ, et al. Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature*. (2018) 559:405–9. doi: 10.1038/s41586-018-0326-5

130. Banasik MB, McCray PB Jr. Integrase-defective lentiviral vectors: progress and applications. *Gene Ther*. (2009) 17:150. doi: 10.1038/gt.2009.135

131. Genovese P, Schiroli G, Escobar G, Di Tomaso T, Firrito C, Calabria A, et al. Targeted genome editing in human repopulating haematopoietic stem cells. *Nature*. (2014) 510:235. doi: 10.1038/nature13420

132. Sadelain M, Brentjens R, Rivière I. The basic principles of chimeric antigen receptor design. *Cancer discovery*. (2013) 3:388–98. doi: 10.1158/2159-8290.CD-12-0548

133. Wang X, Rivière I. Clinical manufacturing of CAR T cells: foundation of a promising therapy. *Mol Therapy-Oncolytics*. (2016) 3:16015. doi: 10.1038/mto.2016.15

134. Gill S, Maus MV, Porter DL. Chimeric antigen receptor T cell therapy: 25 years in the making. *Blood Rev*. (2016) 30:157–67. doi: 10.1016/j.blre.2015.10.003

135. Stadtmauer EA, Fraietta JA, Davis MM, Cohen AD, Weber KL, Lancaster E, et al. CRISPR-engineered T cells in patients with refractory cancer. *Science*. (2020) 367:eaba7365. doi: 10.1126/science.aba7365

136. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*. (2013) 153:910–8. doi: 10.1016/j.cell.2013.04.025

137. Liu X, Zhang Y, Cheng C, Cheng AW, Zhang X, Li N, et al. CRISPR-Cas9-mediated multiplex gene editing in CAR-T cells. *Cell Res*. (2017) 27:154–7. doi: 10.1038/cr.2016.142

138. Hu JH, Miller SM, Geurts MH, Tang W, Chen L, Sun N, et al. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature*. (2018) 556:57–63. doi: 10.1038/nature26155

139. Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, et al. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nat Methods*. (2014) 11:399–402. doi: 10.1038/nmeth.2857

140. Chen JS, Dagdas YS, Kleinstiver BP, Welch MM, Sousa AA, Harrington LB, et al. Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. *Nature*. (2017) 550:407–10. doi: 10.1038/nature24268

141. Shin HY, Wang C, Lee HK, Yoo KH, Zeng X, Kuhns T, et al. CRISPR/Cas9 targeting events cause complex deletions and insertions at 17 sites in the mouse genome. *Nat Commun*. (2017) 8:15464. doi: 10.1038/ncomms15464

142. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. (2013) 31:827–32. doi: 10.1038/nbt.2647

143. Zhang XH, Tee LY, Wang XG, Huang QS, Yang SH. Off-target effects in CRISPR/Cas9-mediated genome engineering. *Mol Therapy-Nucleic Acids*. (2015) 4:e264. doi: 10.1038/mtna.2015.37

144. Komor AC, Badran AH, Liu DR. CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell*. (2017) 168:20–36. doi: 10.1016/j.cell.2016.10.044

145. Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol*. (2015) 33:543–8. doi: 10.1038/nbt.3198

146. Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol*. (2016) 34:339–44. doi: 10.1038/nbt.3481

147. Wei W, Chen ZN, Wang K. CRISPR/Cas9: A powerful strategy to improve CAR-T cell persistence. *Int J Mol Sci*. (2023) 24:12317. doi: 10.3390/ijms241512317

148. Cherkassky L, Morello A, Villena-Vargas J, Feng Y, Dimitrov DS, Jones DR, et al. Human CAR T cells with cell-intrinsic PD-1 checkpoint blockade resist tumor-mediated inhibition. *J Clin Invest*. (2016) 126:3130–44. doi: 10.1172/JCI83092

149. Adusumilli PS, Zauderer MG, Rivière I, Solomon SB, Rusch VW, O'Cearbhaill RE, et al. A phase I trial of regional mesothelin-targeted CAR T-cell therapy in patients with Malignant pleural disease, in combination with the anti–PD-1 agent pembrolizumab. *Cancer Discovery*. (2021) 11:2748–63. doi: 10.1158/2159-8290.CD-21-0407

150. Aparicio C, Acebal C, González-Vallinas M. Current approaches to develop "off-the-shelf" chimeric antigen receptor (CAR)-T cells for cancer treatment: a systematic review. *Exp Hematol Oncol*. (2023) 12:73. doi: 10.1186/s40164-023-00435-w

151. Cutmore LC, Marshall JF. Current perspectives on the use of off the shelf CAR-T/NK cells for the treatment of cancer. *Cancers*. (2021) 13:1926. doi: 10.3390/cancers13081926

152. Ghassemi S, Durgin JS, Nunez-Cruz S, Patel J, Leferovich J, Pinzone M, et al. Rapid manufacturing of non-activated potent CAR T cells. *Nat Biomed Engineering*. (2022) 6:118–28. doi: 10.1038/s41551-021-00842-6

153. Dickinson MJ, Barba P, Jäger U, Shah NN, Blaise D, Briones J, et al. A novel autologous CAR-T therapy, YTB323, with preserved T-cell stemness shows enhanced CAR T-cell efficacy in preclinical and early clinical development. *Cancer Discovery*. (2023) 13:1982–97. doi: 10.1158/2159-8290.CD-22-1276

154. Gao Q, Dong X, Xu Q, Zhu L, Wang F, Hou Y, et al. Therapeutic potential of CRISPR/Cas9 gene editing in engineered T-cell therapy. *Cancer Med*. (2019) 8:4254–64. doi: 10.1002/cam4.2257

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 12 November 2024
DOI 10.3389/fimmu.2024.1452294

OPEN ACCESS

EDITED BY
Mar Requena Mullor,
University of Almeria, Spain

REVIEWED BY
Dmitry Aleksandrovich Zinovkin,
Gomel State Medical University, Belarus
Yanpeng Tian,
First Affiliated Hospital of Zhengzhou University, China

*CORRESPONDENCE
Wen Qiang
wenqiang@zjcc.org.cn
Xu Chao
xuchao@zjcc.org.cn

RECEIVED 20 June 2024
ACCEPTED 17 October 2024
PUBLISHED 12 November 2024

CITATION
Xiaorong Y, Lu X, Fangyue X, Chao X, Jun G and Qiang W (2024) Integrated multiomics characterization reveals cuproptosis-related hub genes for predicting the prognosis and clinical efficacy of ovarian cancer. *Front. Immunol.* 15:1452294. doi: 10.3389/fimmu.2024.1452294

Integrated multiomics characterization reveals cuproptosis-related hub genes for predicting the prognosis and clinical efficacy of ovarian cancer

Yang Xiaorong[1], Xu Lu[2,3], Xu Fangyue[2,3], Xu Chao[2*], Gao Jun[1] and Wen Qiang[4*]

[1]Department of Gynecologic Oncology, Jiangxi Cancer Hospital, The Second Affiliated Hospital of Nanchang Medical College, Jiangxi Clinical Research Center for Cancer, Nanchang, China, [2]Department of Integrated Chinese and Western Medicine, Zhejiang Cancer Hospital, Hangzhou, China, [3]The Second Clinical Medical College, Zhejiang Chinese Medical University, Hangzhou, China, [4]Department of Gynecologic Oncology, Zhejiang Cancer Hospital, Hangzhou, China

Background: As a prevalent malignancy in women, ovarian cancer (OC) presents a challenge in clinical practice because of its poor prognosis and poor therapeutic efficacy. The mechanism by which cuproptosis activity is accompanied by immune infiltration in OC remains unknown. Here, we investigated cuproptosis-related OC subtypes and relevant immune landscapes to develop a risk score (RS) model for survival prediction.

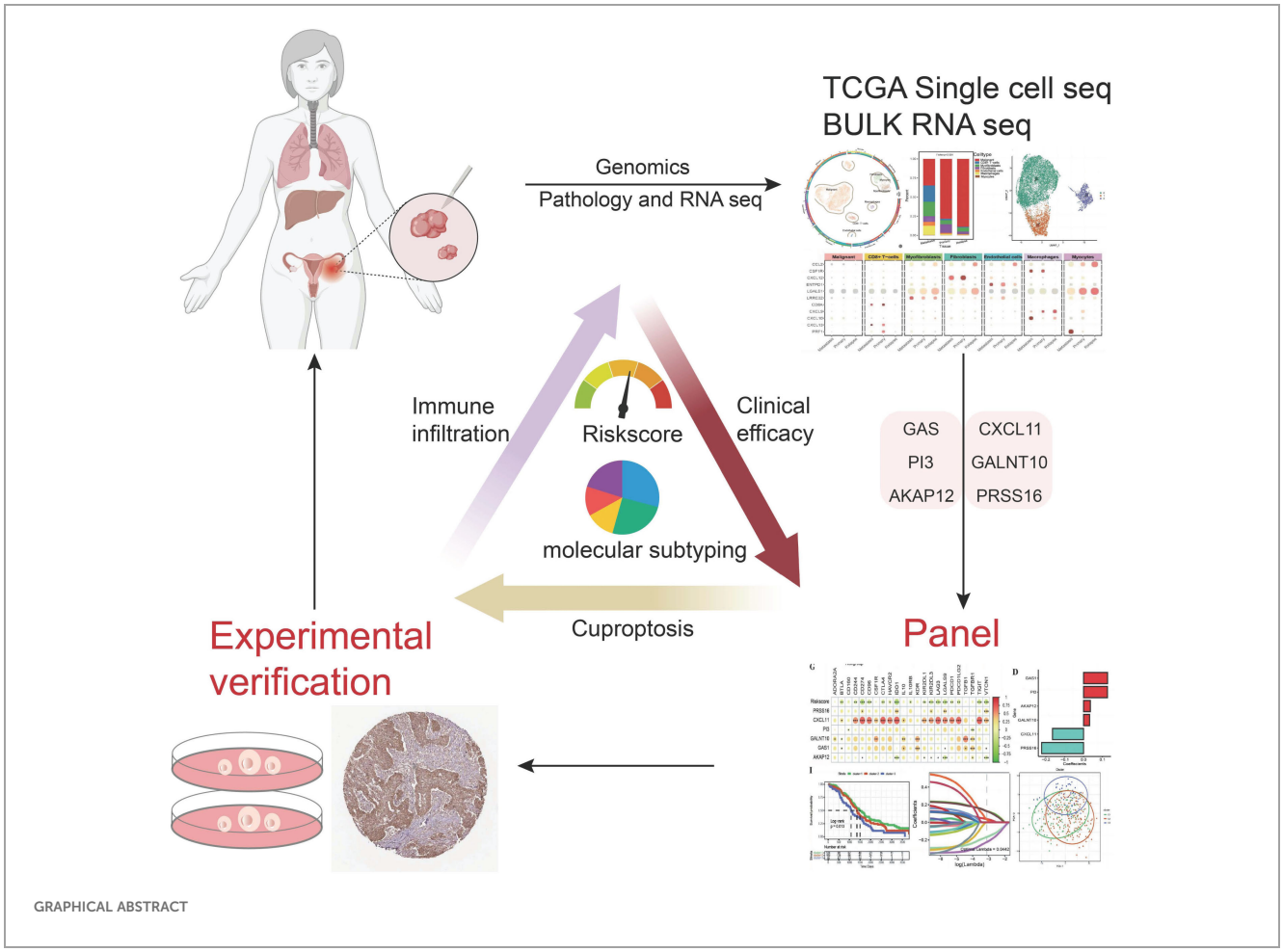
Methods: Cuproptosis-related genes (CRGs) were identified to construct molecular subtypes via an unsupervised clustering algorithm based on the expression profiles of survival-related CRGs in the GEO database. Single-cell datasets were used to estimate immune infiltration among subtypes. The RS oriented from molecular subtypes was developed via LASSO Cox regression in the TCGA OC dataset and independently validated in the GEO and TCGA datasets. Hub markers from RS were identified in tissues and cell lines. The function of the key gene from RS was identified *in vitro*.

Results: We investigated cuproptosis activity and immune infiltration to establish three clinical subtypes of OC based the differentially expressed genes (DEGs) from CRGs to create an RS model validated for clinical efficacy and prognosis. Six hub genes from the RS served as ongenic markers in OC tissues and cell lines. The function of GAS1 in the RS model revealed that it exerts oncogenic effects.

Conclusions: Our study provides a novel RS model including 6 hub genes associated with cuproptosis and immune infiltration to predict OC prognosis as well as clinical efficacy.

KEYWORDS

cuproptosis, RiskScore, ovarian cancer, single-cell sequencing, bulk RNA sequencing



GRAPHICAL ABSTRACT

1 Introduction

As the leading cause of cancer death in the female reproductive system, OC is defined as a "silent killer" because of its insidious symptoms at an early stage and advanced disease at the time of diagnosis. According to the World Health Organization, there were approximately 313,959 new cases worldwide and approximately 207,252 deaths due to OC in 2022 (1). Given the high morbidity and mortality of OC, current diagnostic tools, such as the International Federation of Gynecology and Obstetrics (FIGO) stage system (2) and several common serum biomarkers, such as carbohydrate antigen 125 (CA125) (3) and human epididymis protein 4 (HE4) (4), are far from ideal models for precisely estimating the prognosis and curative effect of each patient. A reliable prognostic model is needed to accurately evaluate the prognosis of patients, which is crucial for optimal individualized management and treatment.

The tumor microenvironment (TME) is a highly complex and heterogeneous ecosystem consisting of tumor cells, infiltrating immune cells, and stromal cells intertwined with noncellular components, in which immune-related genes and immune infiltrating cells play indispensable roles (5). Despite its high morbidity and mortality, OC is a recognized immunogenic tumor, and immunotherapies have attracted substantial attention because of their promising potential in OC therapy. However, cancer immunotherapy is not effective for everyone, and distinct immune cell infiltration patterns result in different responses to cancer immunotherapies (6). In addition, the variability in therapy response and the determinants underpinning these tumor immune phenotypes remain elusive (7). Thus, there is an urgent need to discover key molecular determinants involved in immune

Abbreviations: AUC, area under the curve; BP, biological process; CRGs, cuproptosis-related genes; CMGrisk, cuproptosis marker gene risk score; CIBERSORT, cell type identification by estimating relative subsets of RNA transcripts; CC, cell composition; DEGs, differentially expressed genes; GSVA, gene set variation analysis; GSEA, gene set enrichment analysis; GO, Gene Ontology; GEO, Gene Expression Omnibus; ICI, immune checkpoint inhibitor; IPS, immunophenoscore; IHC, immunohistochemistry; KM, Kaplan–Meier; KEGG, Kyoto Encyclopedia of Genes and Genomes; LASSO, least absolute shrinkage and selection operator; MT, metastasis tissue; MSigDB, Molecular Signature Database; MF, molecular function; OS, overall survival; OC, ovarian cancer; PCs, principal components; PCA, principal component analysis; qRT –PCR, quantitative real-time polymerase chain reaction; RS, risk score; ROC curve, receiver operating characteristic curve; ScRNA-seq, single-cell RNA sequencing; SNVs, single nucleotide variants; TCGA, The Cancer Genome Atlas; TIDE, tumor immune dysfunction and exclusion; UMAP, uniform manifold approximation and projection.

infiltration signatures and construct prognostic signatures based on these signatures.

Recently, a novel form of cell death pathway triggered by copper, named "cuproptosis", which differs from apoptosis, pyroptosis, necroptosis, and ferroptosis, was discovered by Peter Tsvetkov and colleagues in 2022 (8). Copper is an extremely essential element involved in all types of biological procedures in the human body, especially in tumor growth and metastasis (9). Additionally, cuproptosis-related genes (CRGs) have been reported to be associated with immune cell infiltration in melanoma (10), esophageal carcinoma (11), and hepatocellular carcinoma (12), with an increase in protumor or antitumor immune components in tumors; however, few CRGs have been reported in OC (9). Due to the promising future of immunotherapy in the treatment of OC and the crucial role of cuproptosis in immune cell infiltration, evaluating cuproptosis may be an effective way to predict the prognosis and therapeutic benefit for patients.

In this study, molecular subtypes and prognostic models of CRGs in OC were established, and their relationships with the characteristics of immune infiltrating cells in OC were elucidated at the single-cell level. The results were further validated using four cohorts from major public databases. In addition, we aimed to illustrate the potential of the risk model to predict the efficacy of immunotherapy and chemotherapy. Our study revealed a potential association between cuproptosis, prognosis, and the TME in OC patients. These findings may provide a new method to predict outcomes in OC patients and ameliorate them.

2 Materials and methods

2.1 Data download and preprocessing

The FPKM expression profile data of OCs in the TCGA database were downloaded via the R package (TCGAbiolinks), and log2(FPKM+1) transformations were performed to calculate their abundance. Corrected TCGA survival data (13) were used for prognostic survival analysis without samples whose survival time was less than 30 days. Mutation data were also downloaded for genomic variation analysis. The clinical information of the patients in the TCGA_OC cohort used for the analysis is presented in Table 1. The expression and metadata files from the GSE130000 dataset were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) as the input files to construct Seurat single-cell objects for the analysis of the dataset. The expression and survival data of the GSE26712 dataset were downloaded from the GEO database as the GEO external validation set for the risk model. Moreover, the expression and clinical annotation data of the uroepithelial carcinoma dataset IMvigor210 were downloaded from http://research-pub.gene.com/IMvigor210CoreBiologies as the immunotherapy dataset without samples whose survival time was less than 30 days. Sample information for each of the above external datasets is shown in Table 2.

TABLE 1 TCGA OC cohort sample clinical information sheet.

TCGA cohort	Information	Number
Status	Alive	131
	Dead	211
Age	Age>=55	221
	Age<55	121
Grade	G1/2	41
	G3/4	292
Stage	Stage I/II	19
	Stage III/IV	320
venous_invasion	YES	60
	NO	32
lymphatic_invasion	YES	92
	NO	40

2.2 Gene set acquisition

The HALLMARK and GO gene sets were obtained from the MsigDB database (Table 3), as were the immune function-related gene sets from a previous article (14) (Supplementary Table 1: Immune_function_geneset). A total of 43 copper death-related genes were also obtained from previous work (15) (Supplementary Table 1: cuproptosis_related_gene).

2.3 Differential expression analysis

Molecular subtypes of the TCGA_OC cohort were analyzed for differential expression via the limma R package, and differentially expressed genes (DEGs) were subsequently screened with Benjamini–Hochberg (FDR)-corrected thresholds of p values<0.01 and |log2FC|> 0.5.

2.4 Single cell identification and profiling

The single-cell dataset was analyzed via the specialized single-cell transcriptome analysis tool R package Seurat, which includes

TABLE 2 Sample information table for GEO dataset.

Cohort	Subgroup Category	Information	Number
GSE26712	**Status**	Alive	56
		Dead	129
IMvigor210	**Response**	CR/PR	68
		SD/PD	227
	Status	Alive	108
		Dead	187

TABLE 3 Table for immune infiltration.

Cohort	Information
GSE130000	immune cell infiltration
MsigDB	immune function

the steps of constructing objects, data normalization, data downscaling and clustering, identifying marker genes, and visualizing the analysis results. First, the single-cell sequencing results were selected for data merging, and the Seurat object was then constructed via CreateSeuratObject() of Seurat. After harmony, the data were further downscaled via uniform manifold approximation and projection (UMAP) to classify the cell types in the low-dimensional space. The cell subpopulations were identified via FindClusters(), and all the marker genes of all the subpopulations were identified via the FindAllMarkers() function. The function DimPlot() was used to visualize the results of dimensionality reduction of single-cell expression data and the distribution of active cell populations. Moreover, the R package plot1 cell was used for cell type proportion analysis and visualization of marker gene expression.

2.5 Single cell annotation

The R package SingleR was used to annotate the subpopulation results on the basis of Seurat 0.6 resolution, and BlueprintEncodeData were selected as the cell type reference database.

2.6 Identification of cuproptosis-active cell populations

The cuproptosis-associated marker genes of our malignancy cell subpopulation were used to calculate the activity score of each malignant cell via the R package AUCell and to determine the threshold for delineating active cells in the current gene set via the AUCell_exploreThresholds() function. The cell clustering UMAP embedding was then colored based on the AUC score of each cell to show which subpopulation-specific cuproptosis-associated factors were active in which cell subpopulations.

2.7 Construction of the prognostic risk model

One-way Cox analysis was first performed to screen the genes related to prognosis for intersubtype differences, and a prognostic risk score model for ovarian cancer was then constructed on the basis of this gene via least absolute shrinkage and selection operator (lasso). The Tibshirani (1996) method was used to screen the variables to reduce the number of genes in the risk model. The final multifactorial Cox regression model was established to construct the OC risk score (Riskscore) model.

2.8 Survival curves of gene expression

The TCGA OC tumor samples were divided into high- and low-expression groups, with the median gene expression as the threshold point. Survival curves for prognostic analysis were generated via the Kaplan–Meier method, and the significance of the differences was determined via the log-rank test.

2.9 Estimation of the proportion of immune infiltrating cells and the immune score

Based on the expression profile of the TCGA_OC dataset, four algorithms, CIBERSORT, TIMER, ssGSEA, and ESTIMATE, of the R package IOBR were applied to calculate the proportion of infiltrating immune cells.

2.10 Gene set enrichment analysis

Gene set (immune function gene set/HALLMARK) enrichment scores were calculated for each cancer sample on the basis of gene expression in TCGA OC samples via the ssGSEA algorithm of the R package GSVA, which first performs a kernel estimate of the cumulative distribution density function for each gene in all samples. The enrichment score differences between subgroups were then calculated via statistical tests, and the enrichment score heatmap was plotted via the R package pheatmap combined with the clinical characteristics of the samples. The correlations among the expression of model genes, risk scores and enrichment scores were also calculated via the cor() function and visualized via the R package corrplot.

Gene set enrichment analysis (GSEA) uses a predefined set of genes to rank genes according to their differential expression in two types of samples and then tests whether the predefined set of genes is enriched at the top or bottom of this ranking table. Enrichment analysis was performed through the R package clusterProfiler on the basis of GO functional gene sets and KEGG pathways, and the top 8 gene sets with significant enrichment results were then selected to generate bubble plots showing the enrichment results.

2.11 Genomic SNV analysis

Based on the maf file of somatic mutation detection results of the TCGA_OC cohort, the oncoplot() function of the R package mafTools was used to draw waterfall plots to show the differences in SNV mutations between different model groupings. Finally, the maf data of the high- and low-risk groups were analyzed via the mafCompare() function to obtain genes with significant differences in mutations between the two groups.

2.12 Immunotherapy response prediction

The algorithm is based on expression profiles prior to tumor treatment and scores multiple published transcriptomic biomarkers to predict patients' immunotherapeutic response. The TIDE score

(http://tide.dfci.harvard.edu) integrates T-cell dysfunction and exclusion features, simulates tumor immune escape with different levels of tumor-infiltrating cytotoxic T cells, and appears to be highly advantageous compared with other biomarkers. The immunophenoscore (IPS) can be used to identify immunogenicity and predict the response to immunotherapy in multiple tumor types. We obtained IPS scores for tumor samples in the TCGA_OC dataset via an online website (https://tcia.at/home) and performed between-group difference analysis via statistical tests.

2.13 Samples and cell collection

All tumor tissues were obtained from Jiangxi Cancer Hospital and stored in liquid nitrogen at -80°C until use. This research was approved by the ethics committee of Jiangxi Cancer Hospital (Approval number: 2022ky305). Every patient provided informed consent prior to the collection and usage of these clinical materials. The OC cell lines used in this study were obtained from the ATCC cell bank.

2.14 Quantitative real-time polymerase chain reaction (qRT–PCR)

mRNA was extracted from tissues with TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. The SYBR Green PCR Master One-Mix Kit (TransGen, Beijing, China) was used for qPCR to determine mRNA expression. Detailed information about all the primers used is listed in Supplementary Table 12.

2.15 Wound healing assay

Approximately 3×10^5 A2780 and SKOV3 OC cells were seeded in a 6-well plate. After the cells filled the entire area, the culture inserts were removed. The cells were treated with PB (1.0 or 2.0 μM) or DMSO for 48 h. The cells were then rinsed twice with PBS to remove floating cells. Images were obtained under an optical microscope (ix71, Olympus, Japan) at 0 h, 24 h and 48 h after wound induction.

2.16 Western blot analysis

The details of the assay were as described in our previous study (16). The antibodies used are listed in Supplementary Table 12.

2.17 Immunohistochemical analysis

The details of the assay were as described in our previous study (17). The antibodies used are listed in Supplementary Table 12. When tissue slices are observed under an optical microscope, they are graded based on the degree of staining and the extent of positivity. The degree of staining can range from 0 to 3, representing varying depths of color (negative staining, pale yellow, light brown, and dark brown). The extent of positivity can range from 1 to 4, representing different percentages of positive cells (0-25%, 26-50%, 51-75%, 76-100%). By grading the intensity of cellular staining and the percentage of positive cells, these two scores are multiplied to obtain the final score.

2.18 Plasmid construction and transfection

Stably transfected small interfering RNAs (siRNAs) were obtained from ElifeBio (Hangzhou, China) and transfected into cells via Lipofectamine iMax (Invitrogen, AL, USA) following the manufacturer's instructions. The transfection efficiency was verified via qRT–PCR.

2.19 Transwell assays

The premixed matrix gel was added to the Transwell chamber, which was then placed in a 24-well deep-well plate and incubated in a cell culture incubator for 2 hours. Adherent cells were digested with trypsin solution, resuspended by pipetting, and counted via a cell counter. OC cells transfected with GAS1 were digested and seeded in the upper chamber, which was supplemented with 0.2 ml of serum-free medium. Then, 700 μl of complete medium containing serum was added to the 24-well deep-well plate housing the Transwell chamber. After 24 hours, the Transwell chamber was removed, and the cells were processed for subsequent fixation, staining, and counting.

2.20 Cell proliferation analysis

We evaluated the proliferation of OC cells transfected with GAS1 via colony formation and EdU assays. The specific experimental procedures were performed as described previously (18).

2.21 Description of the statistical analysis

For statistical mapping, the Wilcoxon test was used to compare the differences between two groups of samples, and the Kruskal–Wallis test was used to compare the differences between multiple groups of samples.

3 Results

3.1 Single-cell landscape and cuproptosis activity in OC

3.1.1 Expression of the TME and CRGs in OC

A total of 13,511 cells from OC tissue in GSE130000 (Table 3) were obtained. For initial data dimensionality reduction analysis (PCA), we selected the top 15 PCs for subsequent PCA and then drew a clusterree (clustree.res.pdf) to annotate the cell subtypes with SingleR (Supplementary Table 1: cell_info). We classified all epithelial cells, including seven cell subtypes, in tumor tissue as a malignant cell group (Figure 1A). To investigate the differences in the immune function of different cell types in different tissues, we first plotted an expression

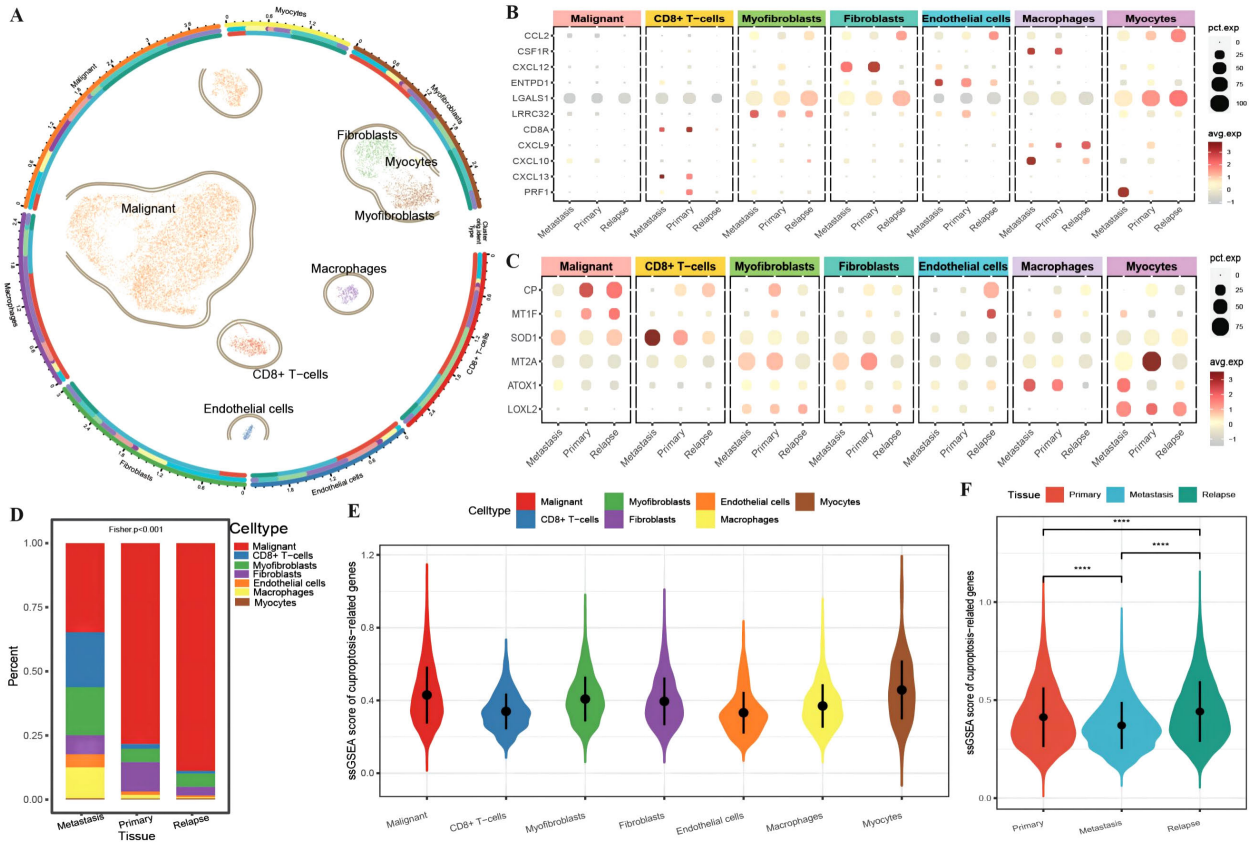


FIGURE 1
Differences in the tumor microenvironment at the single-cell level. **(A)** UMAP distribution map of cell clustering and annotation results, where the circles represent cell annotation results. From outer to inner, they respectively indicate cell type, sample source, and tissue type; **(B)** Bubble plot of the expression of anti-tumor immune and anti-tumor response genes combined with different cell types from various tissue sources, with the size of the dots representing the number of cells expressing the genes and the color indicating the level of expression; **(C)** Bubble plot of the expression of ferroptosis-related marker genes combined with different cell types from various tissue sources, with the size of the dots representing the number of cells expressing the genes and the color indicating the level of expression; **(D)** Bar chart depicting the distribution of cell types in different tissues, with different colors representing different cell types; **(E)** Violin plot showing the enrichment scores of ferroptosis-related genes in various cell clusters; **(F)** Violin plot illustrating the differences in ferroptosis enrichment scores among primary tissue, metastatic tissue, and recurrent tissue. The blue color represents cells from metastatic tissue, red indicates cells from primary cancer tissue, and green represents cells from recurrent tissue. ****p < 0.0001.

bubble chart of antitumor response factors and antitumor immune genes, which revealed low expression levels of these genes in malignant cells, whereas the antitumor response factor in CD8+ T cells was active in both metastatic and primary cancer tissues. Moreover, antitumor immune factors from macrophages, fibroblasts, and endothelial cells exhibited much greater activity in primary and metastatic tissues than in recurrent tissue (Figure 1B). We also plotted an expression bubble chart of the CRGs, which revealed that CP was expressed at higher levels in the malignant cells of primary and recurrent tissues and that MT2A was expressed at higher levels in the primary myocardial cells. However, SOD1 preferred CD8+ T cells in metastatic tissue (Figure 1C). Next, we explored the distribution of different cell types across different tissue sources, revealing the diverse proportions of malignant cells, CD8+ T cells, myofibroblasts and macrophages in different tissues (Figure 1D). Finally, according to the expression of each gene in single cells, we calculated the ssGSEA scores of the CRGs (Supplementary Table 1: cuproptosis_ssGSEA), which revealed that the enrichment level of CRGs in malignant cells was relatively high (Figure 1E). Moreover, the enrichment score of CRGs in recurrent tissue was greater than that in primary and metastatic tissue (Figure 1F).

3.1.2 CRGs associated with cellular immunity in OC

We subdivided the 10,358 malignant cells into three malignant subgroups (Figure 2A) and then generated an expression bubble map of all the CRGs by identifying the marker genes of each cluster, which revealed that the expression of CRGs was significantly greater in subgroup 2 than in subgroups 1 and 0 (Figure 2B). Moreover, the enrichment scores of the CRGs were highest in subgroup 2, followed by subgroups 1 and 0 (Figure 2C). We also performed immune functional analysis of the three malignant subtypes, which revealed that the enrichment scores for tumor cell immune response ability and interferon-γ response genes were significantly greater in subgroup 2 than in subgroups 1 and 0 (Figures 2D, E).

3.1.3 Identification of cuproptosis activity in malignant cells

To further investigate the expression and functional characteristics of CRGs at the single-cell level, we identified 192

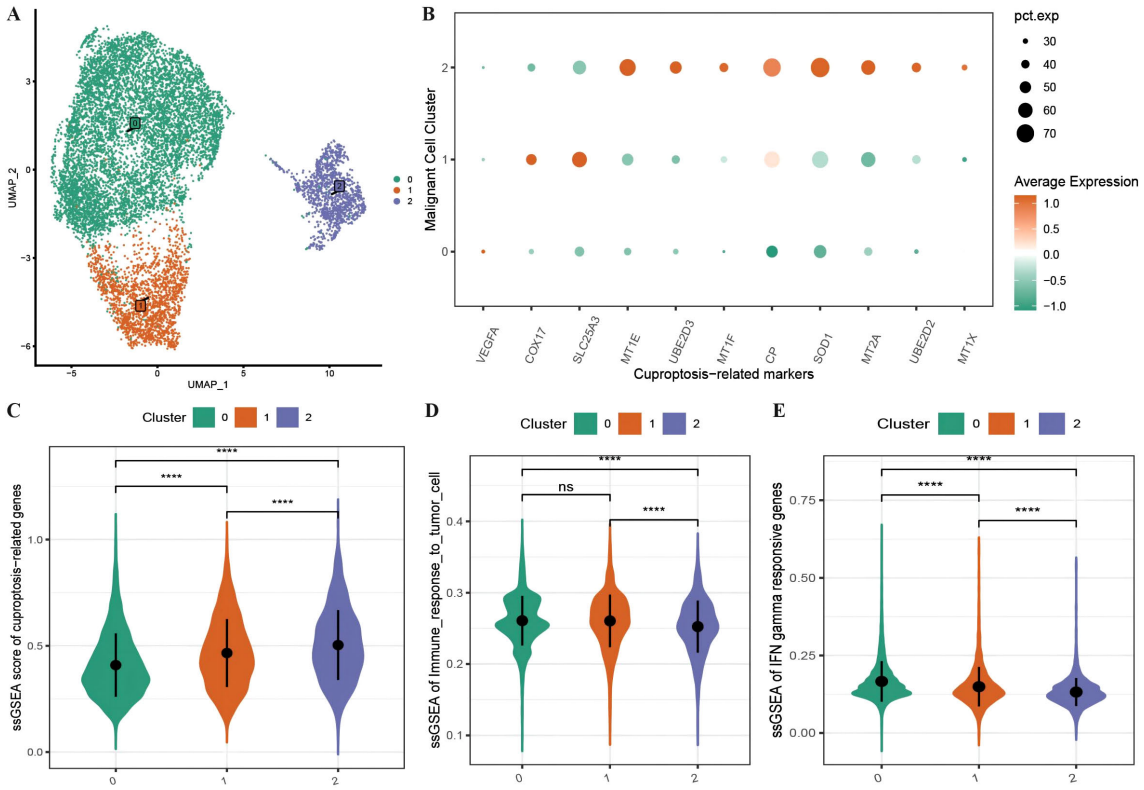


FIGURE 2
Characteristics of cuproptosis in malignant cells. **(A)** UMAP distribution map of the re-clustering results of malignant cells; **(B)** Bubble plot showing the expression of all ferroptosis-related marker genes in malignant subtypes, where the size of the dots represents the number of cells expressing the genes and the color represents the level of expression; **(C)** Violin plot displaying the differences in enrichment scores of ferroptosis-related gene sets among malignant subtypes; **(D, E)** Violin plots illustrating the differences in enrichment scores of tumor cell immune response and interferon response gene sets among malignant subtypes. ****p < 0.0001.

cells as an active malignant population related to active cuproptosis in malignancies (Figure 3A) and used the optimal threshold to determine cell viability (Supplementary Table 2: cell_info). A cumulative distribution histogram was subsequently plotted to display the distribution of active and inactive cuproptotic malignant cells in each tissue type, which revealed that the number of active cuproptotic malignant cells was highest in the recurrent tissue (Figure 3B). In addition, a bubble chart of CRGs in three subtypes and different active types revealed that CRGs were enriched in the active population of all malignant subtypes, especially in subtype 2 (Figure 3C). To identify functional differences between the active and inactive populations, we compared their immune functions and identified the active population with a significantly lower level of immune response to tumor cells and interferon-γ response gene enrichment scores compared with the inactive population (Figures 3D, E). Finally, differential expression analysis was conducted on the two cell populations, with 157 DEGs identified (Supplementary Table 2: Active_cells_DEGS), which were enriched mainly in response to metal ions such as copper, zinc, and cadmium, as well as in the detoxification of organic compounds, according to GOBP enrichment analysis (Figure 3F).

3.2 Role of the identification of molecular subtypes associated with active necroptosis DEGs

3.2.1 Immune infiltration between molecular subtypes in OC

Based on the identification of DEGs between active and inactive cuproptotic malignant cell populations via single-cell analysis, we identified OC molecular subtypes with the best clustering performance achieved via the PAM clustering algorithm, Spearman distance, and K=3 (Figures 4A–D; Supplementary Table 3). The consistency clustering cumulative distribution function (CDF) with the KM curves revealed clear boundaries between the three subtypes, indicating good clustering results, and K=3 was the result of our molecular subtype identification (Figures 4E–K; Supplementary Table 3: cc_group). We also explored the different cuproptosis expression patterns of the three subtypes (Supplementary Figure 1).

With the results of the immunoreactive cell proportion analysis of the TCGA OC samples (Supplementary Table 4), we analyzed the proportion of immunoreactive cells among the three subtypes, which revealed that the stromal, immune, and

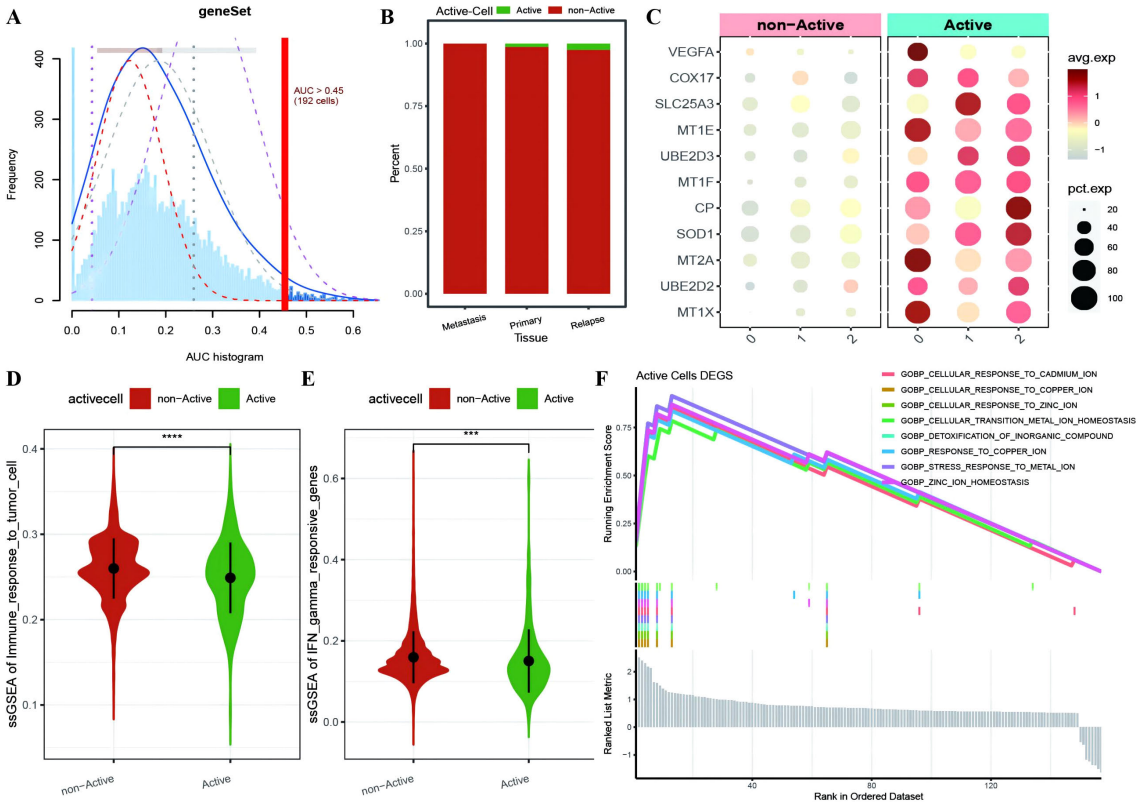


FIGURE 3
Identification of active cuproptosis malignant cell population. **(A)** Single-cell AUC scores for ferroptosis genes, with the optimal threshold being 0.45; **(B)** Bar chart displaying the distribution of active populations in different tissue types, where red represents inactive populations and green represents active populations; **(C)** Bubble plot depicting the expression of ferroptosis marker genes in various malignant cell subtypes and different activity types, with a redder color indicating higher average expression values and the size of the dots representing the number of expressing cells; **(D, E)** Violin plots showing differences in immune function between active and non-active malignant cell populations with respect to ferroptosis activity, where green represents the active population and red represents the non-active population; **(F)** Line plot illustrating the enrichment results of differentially expressed genes between the active and non-active ferroptosis groups based on GOBP-GSEA. ***p < 0.001, ****p < 0.0001.

ESTIMATE scores of the C3 subtype were significantly greater than those of the other two subtypes, whereas the tumor purity was lower (Figure 5A). Moreover, the B-cell infiltration level was lower in the C3 subtype than in the other subtypes, whereas the infiltration of T cells, macrophages, DCs, and neutrophils was greater according to the TIMER algorithm (Figure 5B). Finally, we explored the infiltration proportions of 28 immune cells among the three subtypes via the ssGSEA algorithm and found that there were significant differences in the infiltration proportions of 24 cell types and that the infiltration proportion of the C3 subtype was significantly greater than those of the other two subtypes (Figure 5C). Furthermore, we also detected differences in the enrichment scores of the six immune-related gene sets among the three MSs, from which we found that the immune factor enrichment score in C3 was higher than that in C1 and C2 and that C2 had a higher score than C1 (Figures 6A–F). Additionally, the enrichment of four immune function gene sets among the subtypes revealed that T-cell activation and innate immunity in C3 were significantly stronger than those in C2 and C1 (Figures 6G–J).

3.2.2 Differential expression of immune markers across molecular subtypes

The expression levels of immune checkpoints from the TISIDB database in the C3 subtype were generally significantly greater than those in the C1 and C2 subtypes (Supplementary Figure 2A). Second, to explore the differences in other biological functions between molecular subtypes, we identified the upregulated genes in each subtype relative to those in the other two subtypes and then combined the three upregulated gene sets as the DEGs between subtypes. A total of 1033 subtype-specific DEGs were obtained (Supplementary Table 3: All_diff_gene). The KEGG pathway results revealed that the subtype-specific DEGs were enriched mainly in ECM–receptor interactions, proteoglycans in cancer, the PI3K–Akt signaling pathway, protein digestion and absorption, complement and coagulation cascades, phagosomes, and *Staphylococcus aureus* infection (Supplementary Figure 2B). Moreover, the GO functional enrichment results revealed that the DEGs were enriched mainly in biological processes related to cell tissue, migration, adhesion, regulation of vasculature development regulation, and wound healing; molecular functions such as extracellular matrix

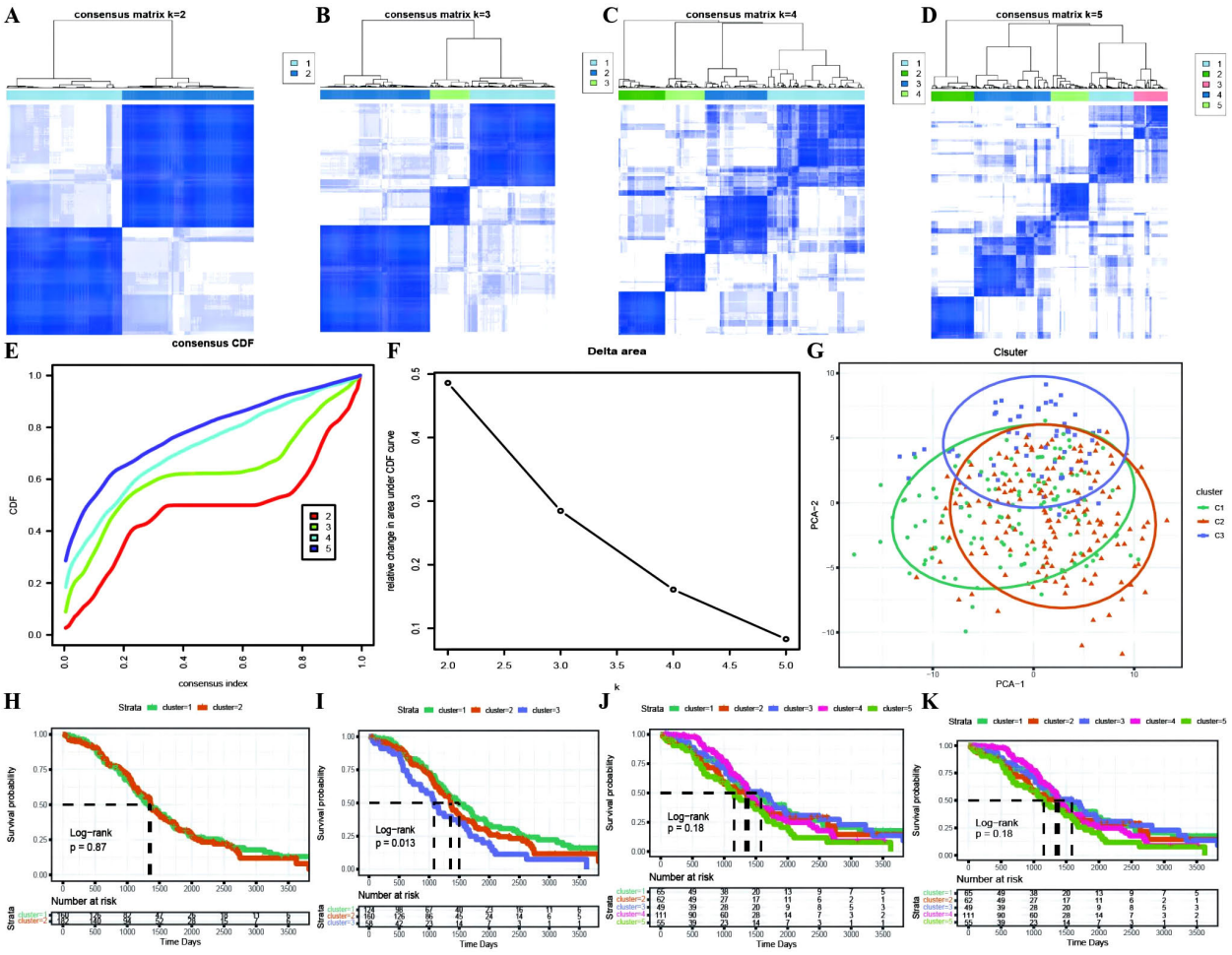


FIGURE 4
Identification of TCGA molecular subtypes in OC. **(A–D)** Clustering results for different numbers of clusters, specifically k=2, k=3, k=4, and k=5; **(E)** Distribution of the cumulative distribution function (CDF) curve for consensus clustering; **(F)** Distribution of the area under the cumulative distribution function (CDF) curve for consensus clustering; **(G)** Scatter plot showing the results of clustering using PCA dimensionality reduction algorithm, where green represents C1, orange represents C2, and purple represents C3; **(H–K)** Survival curves for different numbers of clusters (k=2, k=3, k=4, k=5), where different colored curves represent different clusters.

structural constituents, glycosaminoglycan binding, and growth factor binding; and cellular component gene sets such as the collagen-containing extracellular matrix, endoplasmic reticulum lumen, and collagen triple helix complex (Supplementary Figures 2C–E).

3.3 Construction and validation of a prognostic risk model for OC

We identified 24 genes significantly associated with OC prognosis from the DEGs associated with OC among the cuproptosis subtypes (Figure 7A; Supplementary Table 5: cox_res). We subsequently applied Lasso linear regression to eliminate redundant genes according to these 24 genes, resulting in six prognosis-related signatures (Figures 7B–D; Supplementary Table 5: lasso_res). Then, we established Kaplan–Meier survival curves for these genes in the overall TCGA cohort. We found that there were significant differences between the KM curves for four genes: PRSS16 and CXCL11 had better prognostic values in the high-expression group, whereas PI3 and GAS1 had better prognostic values in the low-expression group (Figures 7E–J). Furthermore, we calculated the RS for each sample via the formula RiskScore = PRSS16*-0.223 + CXCL11*-0.166 + PI3*0.127 + GALNT10*0.032 + GAS1*0.127 + AKAP12*0.036, which led to high-risk and low-risk groups via the median risk score of 0.4525444 as the threshold (Supplementary Table 5: TCGA_Train). These six model genes were diverse in the high- and low-risk groups of the training set and were related to prognosis (Supplementary Figures 3A–D). The receiver operating characteristic (ROC) curve of the prognostic signature, with area under the ROC curve (AUC) values of 0.709, 0.711, and 0.773 at 3, 5, and 8 years, revealed good predictive performance of the model score (Supplementary Figure 3E). Similar results were also validated in the TCGA test set, TCGA_OC dataset, and GEO dataset GSE26712 (Supplementary Figures 4–6).

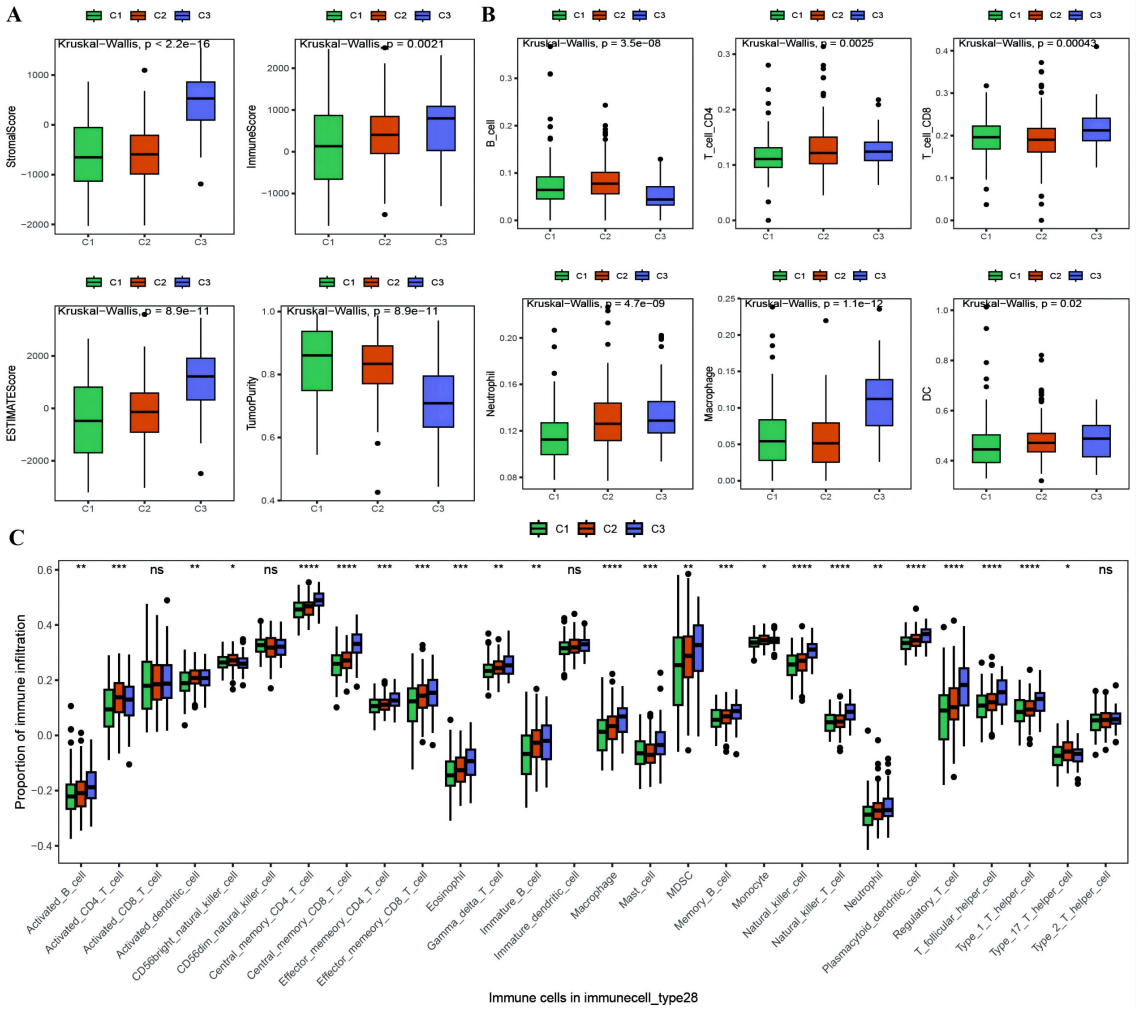


FIGURE 5
Differences in proportion of immune-infiltrating cells among TCGA molecular subtypes. **(A)** Box plot showing the differences in ESTIMATE scores between molecular subtypes of ovarian cancer, where green represents C1, orange represents C2, and purple represents C3; **(B)** Box plot illustrating the differences in the proportions of immune infiltrating cells between molecular subtypes calculated using the TIMER algorithm; **(C)** Box plot displaying the differences in the proportions of 28 types of immune infiltrating cells between molecular subtypes. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

3.4 Association between RS and multiple features in OC

3.4.1 Association between risk score and immune microenvironment in OC

The RS was confirmed to be an independent prognostic factor for age (>55), stage and venous invasion in OC patients (Figure 8A; Supplementary Table 5: Clinical_stat). The RS is a clinical factor that contributes to survival time and survival status in combination with the clinical indicator age (Figure 8B). The risk score was significantly different between both the age subgroups and the molecular subtypes (Figures 8C, D), indicating that the age and molecular subtypes of the ovarian cancer samples were significantly associated with the RS. To investigate the differences in the tumor immune microenvironment between the high- and low-risk groups in the RS, in light of the estimation results of the proportion of infiltrating immune cells, we discovered that 5 of the 28 immune cells, including activated CD4+ T cells, activated CD8+ T cells, effector memory CD8+ T cells, immature B cells, and type 2 T helper cells, were significantly greater in the low-RS group than in the high-RS group according to the ssGSEA algorithm (Figures 9A–E). We subsequently explored the expression of 23 immunosuppressive checkpoints (Supplementary Table 6: check_model_data) in the high- and low-risk groups, 16 of which were significantly more highly expressed in the low-risk group than in the high-risk group (Figure 9F). In addition, the associations between the expression of 6 genes from the RS and that of 23 immunosuppression checkpoint genes (Supplementary Table 6: corrdata & corrp) revealed that the expression of CXCL11 was significantly positively correlated with that of multiple immune checkpoint genes; however, the expression of the RS was generally negatively correlated with that of other genes (Figure 9G). Finally, the stromal score (Figure 9H) from the ESTIMATE algorithm was higher in

10.3389/fimmu.2024.1452294

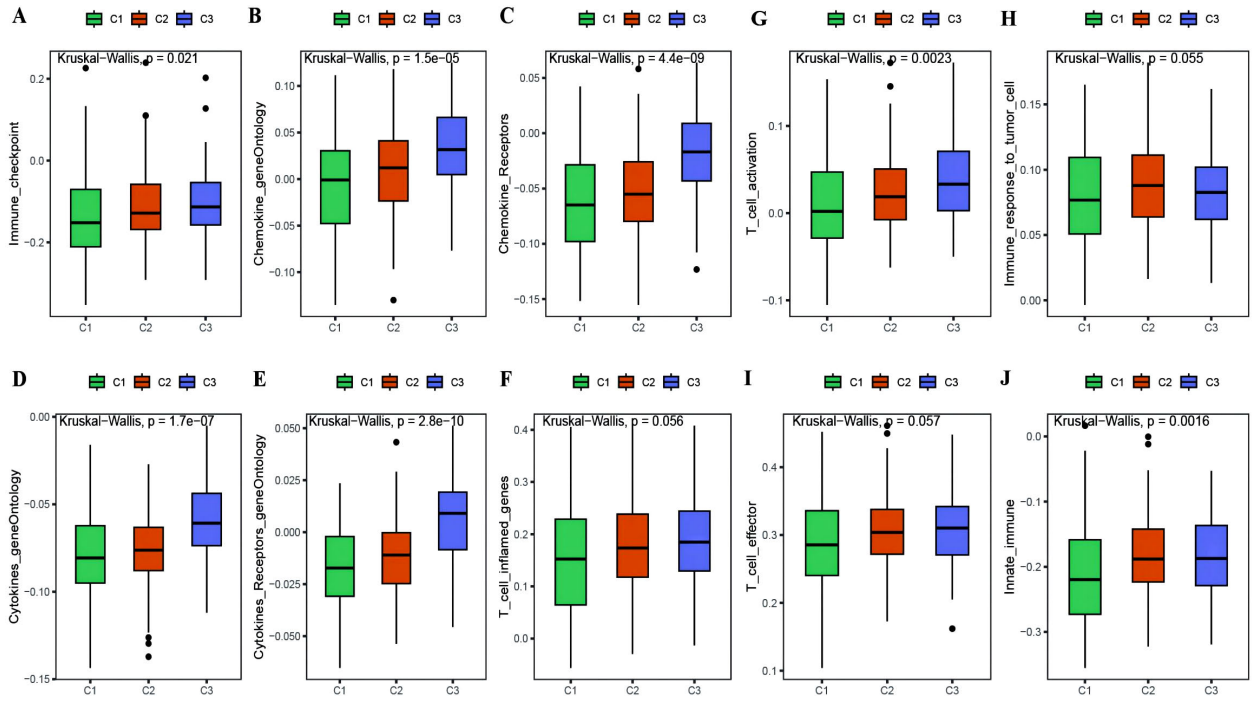


FIGURE 6
Differences in immune enrichment among subtypes. **(A–F)** Box plots showing the differences in enrichment of immune factor gene sets between subtypes, including immune checkpoints, chemokine gene ontology, chemokine receptor gene ontology, cytokine gene ontology, cytokine receptor gene ontology, and T cell inflammatory genes; **(G–J)** Box plots illustrating the differences in immune functions between subtypes, including T cell activation, immune response to tumor cells, effector T cells, and innate immunity.

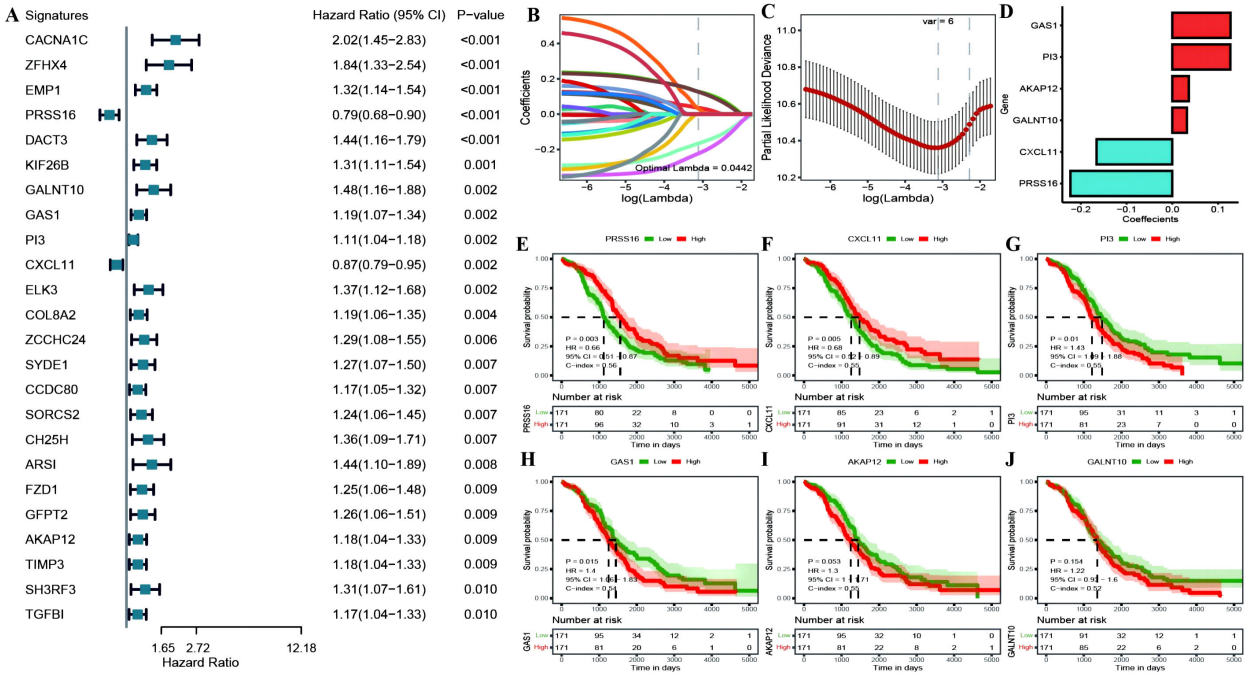


FIGURE 7
(A) construction of the prognostic risk model. **(A)** Forest plot of the results of single-factor Cox analysis of differentially expressed genes between subtypes; **(B)** Trajectory of changes in the independent variables of LASSO regression, where the x-axis represents the logarithm of the independent variable Lambda and the y-axis represents the coefficient of the independent variable; **(C)** Confidence interval for each Lambda in LASSO regression; **(D)** LASSO regression coefficients for 8 key prognostic factors; **(E–J)** KM curves for model genes, where red represents high expression group and green represents low expression group.

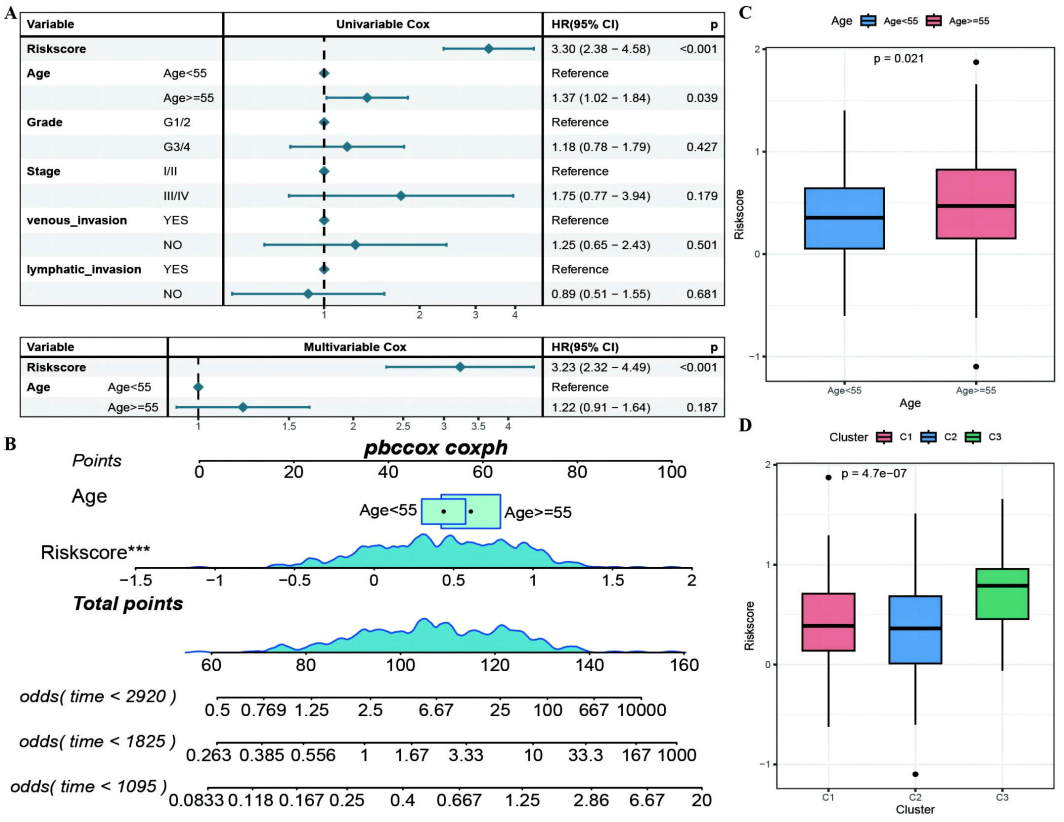


FIGURE 8
The RiskScore as an independent prognostic factor. **(A)** Forest plot of the results of single- and multi-factor Cox analysis for clinical factors; **(B)** Nomogram of the predictive model, where the square plus line segment represents the contribution of each clinical factor to the outcome event, Total Points represents the total score obtained by adding up the scores for all variables, and the three lines at the bottom represent the 8/5/3-year survival probabilities corresponding to each value point; **(C, D)** Box plots showing the differences in risk scores distribution between different Age and molecular subtype groups, with different colors representing different groups, and the p-value indicating the significance of the difference. ***p < 0.001.

the high-RS group than in the low-RS group, and the immune score (Figure 9I) results were the opposite.

Analysis of the differences in the expression of antitumor immune and antitumor response genes in the RS groups revealed that the expression of antitumor response-related genes was significantly lower in the high-risk group than in the low-risk group. We then selected 197 DEGs between the two groups (Supplementary Table 7) to explore their functions via Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) enrichment analyses, which revealed that the DEGs were enriched mainly in antigen processing and presentation, Epstein–Barr virus (EBV) infection, allograft rejection, autoimmune thyroid disease and other Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, as well as in gene ontology (GO) functional gene sets related to antigen processing, processing and presentation, response to viruses, and major histocompatibility complex (MHC) protein complexes.

3.4.2 Correlation of the RS model with the HALLMARK

Based on the HALLMARK enrichment score results of the OC samples and the RS information, we aimed to explore the pathway enrichment differences between the high- and low-risk groups (Supplementary Table 8), and 28 of the pathways presented significant differences in enrichment scores between the high- and low-risk groups (Supplementary Figure 6A). Moreover, the expression of CXCL11 and GALNT10 was significantly positively correlated with multiple pathway enrichment scores; however, the expression of RS was significantly negatively correlated with the interferon α and γ response (Supplementary Figure 6B). RS was found to be significantly negatively correlated with multiple immune function gene sets, where we selected four functional gene sets with strong correlations, including Co inhibition, interferon γ response, immune response to tumor cells and antigen processing and presentation (Supplementary Figures 6C–F).

3.4.3 Association between RS and tumor genome mutations

Waterfall plots of the top 20 genes with mutation frequencies were generated separately for the high- and low-risk groups combined with other clinical information, demonstrating the distribution of gene mutations between the two groups and samples with different clinical characteristics, with TP53, TTN,

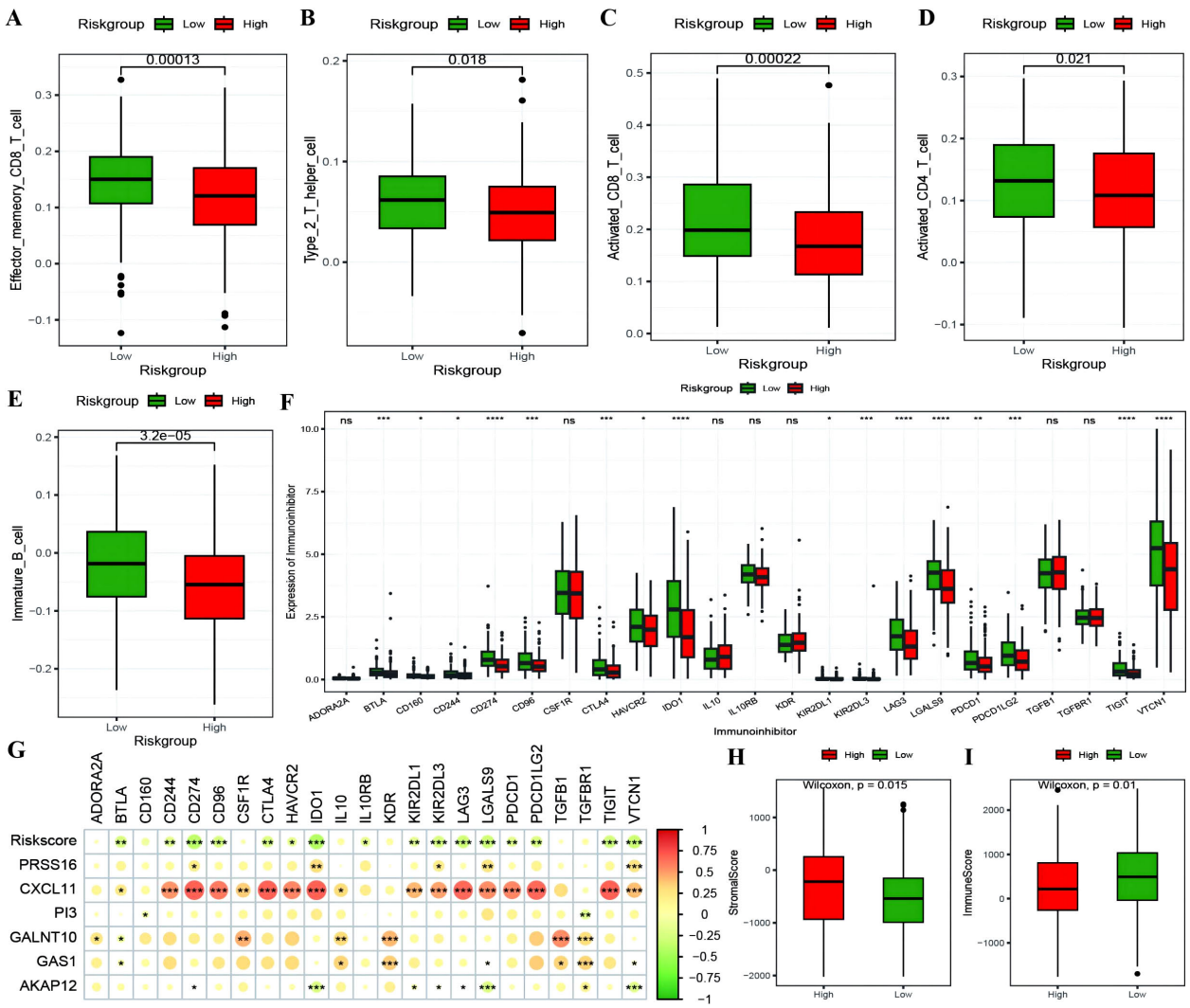


FIGURE 9
The correlation between RiskScore and immune checkpoints. **(A–E)** Box plots showing the differences in proportions of immune-infiltrating cells between high and low-risk groups calculated using the ssGSEA algorithm, where red represents the high-risk group and green represents the low-risk group; **(F)** Box plot illustrating the expression differences of 23 immune checkpoint inhibitors between high and low-risk groups, where red represents the high-risk group and green represents the low-risk group; **(G)** Heatmap of the correlation coefficients between the expression of model genes, risk scores, and the expression of immune checkpoint inhibitors, where the color of the dots represents the strength of the correlation, and "*" denotes significance; **(H, I)** Box plots displaying the differences in stromal score and immune score between high and low-risk groups. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

and CSMD3 having the highest mutation frequencies in OC (Figures 10A, B). Due to differences in gene mutation frequencies between the high- and low-risk groups (Supplementary Table 9: mafCompare_High_VS_Low), we divided the samples into MT and WT groups to analyze the difference in RS between them. We found that the risk scores of the TICRR, CACNA1S and C7 gene groups were significantly different from those of the wild-type group, and the mutation frequency of these genes was greater (Figures 10C–E). Using the same methodology, we also scrutinized the correlation between tumor mutational burden (TMB) (Supplementary Table 9: Riskscore_tmb_res) or RS and prognosis, revealing a more favorable prognosis associated with high TMB (Figure 10F).

3.5 RS model estimation for predicting patient clinical efficacy

To investigate whether genes from the Riskscore model can serve as markers for immunotherapy response, we explored the ability of tumor risk scores to predict patient benefit from immunotherapy. Initially, within the immunotherapy dataset, risk scores were computed for each sample and then categorized into high- and low-risk groups (Supplementary Table 10: IMvigor210_res). The low-risk group had a better prognosis (Figure 11A), with more pronounced benefits following immunotherapy (Figures 11B, C). TIDE (Supplementary Table 10: TIDE_res) was subsequently used to predict the

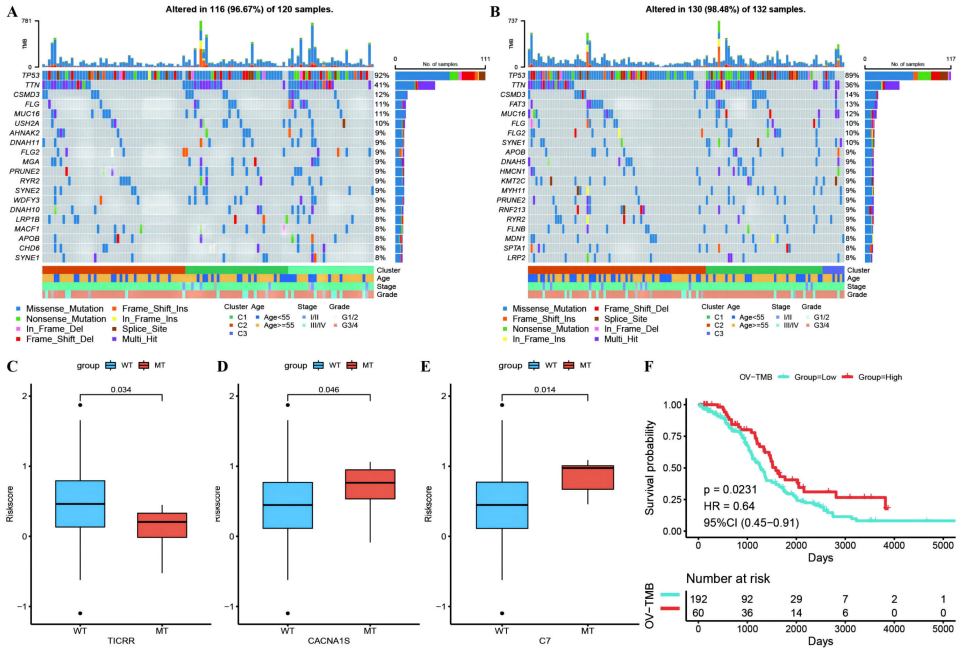


FIGURE 10
Differences of gene mutant enrichment in Riskscore model. **(A)** Waterfall plot of the top 20 genes' SNV mutation frequencies in the high-risk group; **(B)** Waterfall plot of the top 20 genes' SNV mutation frequencies in the low-risk group; **(C–E)** Box plots illustrating the difference in Riskscore between mutated and wild-type groups for genes with significantly different mutation frequencies between high and low-risk groups, where red represents the mutated group and blue represents the wild-type group; **(F)** Survival curve comparison between high and low TMB groups, where red represents the high TMB group and blue represents the low TMB group.

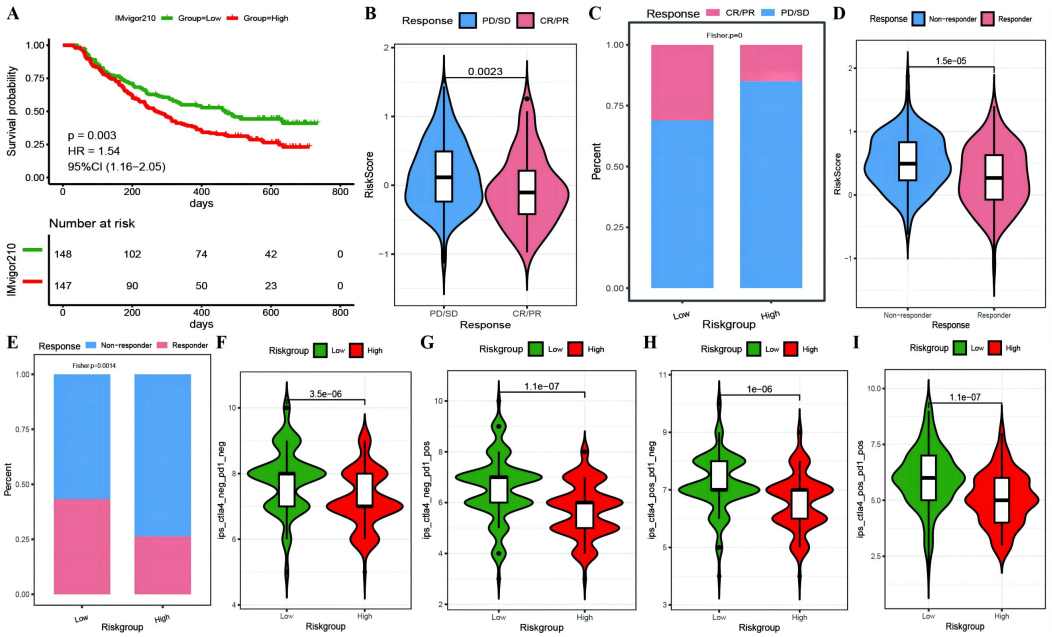


FIGURE 11
Predictive modeling of immunotherapy efficacy via risk stratification. **(A)** Kaplan-Meier curves for high-risk versus low-risk groups within the IMvigor210 cohort; **(B)** Violin plots illustrating differences in risk scores between responder and non-responder groups within the IMvigor210 cohort, with responders depicted in red and non-responders in blue; **(C)** Bar graphs showing the cumulative distribution differences between responder and non-responder groups among high-risk and low-risk categories within the IMvigor210 cohort; **(D)** Violin plots of risk scores for responder versus non-responder groups as predicted by TIDE; **(E)** Bar graphs depicting the cumulative distribution of responder versus non-responder groups among model-based stratifications as forecasted by TIDE; **(F–I)** Violin plots representing the differences in IPS scores between high-risk and low-risk groups, with high-risk groups colored in red and low-risk groups in green.

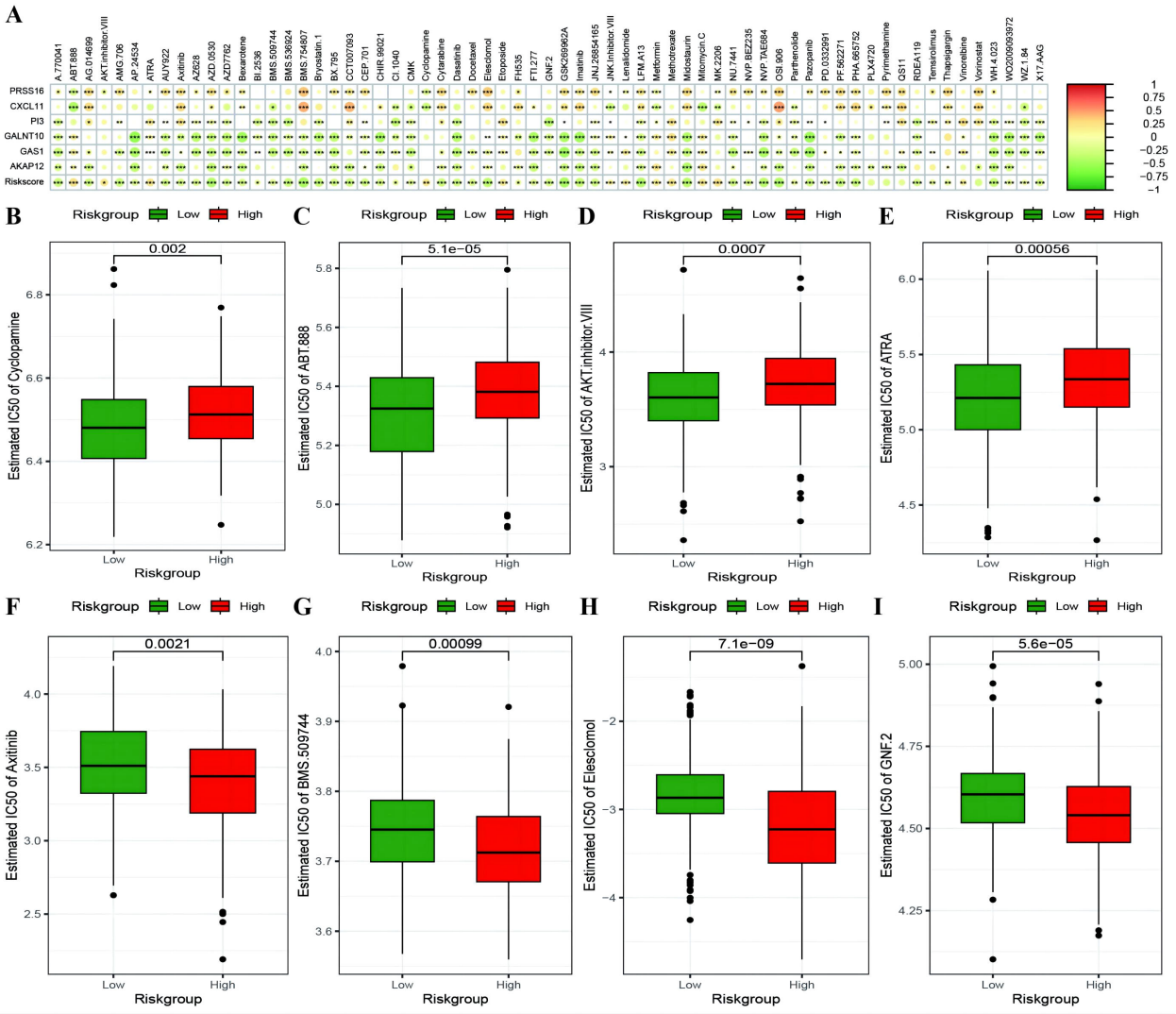


FIGURE 12
Risk model predicts chemotherapy drug resistance. **(A)** Heatmap illustrating the correlation between IC50 values of drugs showing significant differences in sensitivity between high-risk and low-risk groups, model gene expression, and RiskScore. The intensity of the color indicates the level of correlation, with * denoting significance; **(B–E)** Box plots depicting differences in IC50 values for drugs to which high-risk groups are more sensitive, with high-risk groups in red and low-risk groups in green; **(F–I)** Box plots showing differences in IC50 values for drugs to which low-risk groups are more sensitive. *p < 0.05, **p < 0.001, ***p < 0.001.

immune response status of samples in the TCGA_OC dataset to evaluate the model's ability to predict the immune response. The risk score of the responsive group was significantly lower than that of the nonresponsive group (Figures 11D, E). Additionally, the proportion of responders in the low-RS group was significantly greater than that in the high-RS group. The analysis from TIDE indicated a significant correlation between risk scores, RS grouping, and immune response, suggesting a strong predictive ability of the model for the immune response. Furthermore, we employed the immunophenoscore (IPS) (Supplementary Table 10: IPS_res) to explore clinical efficacy from the perspective of tumor immunogenicity and found that all four IPS scores in the low-RS group were significantly higher than those in the high-RS group, further revealing that patients in the low-RS group are more likely to benefit from immunotherapy (Figures 11F–I). Furthermore, we analyzed the relationships between the risk score model and the IC50 values of targeted and chemotherapeutic drugs via the GDS and TCGA databases (Supplementary Table 11: drug_res&sign_stat; Supplementary Table 11: corr_IC50_ModelGene_corr). Our findings confirmed a significant negative correlation between the expression of GALNT10, GAS1, and AKAP12 and the risk score and IC50 values of the drugs (Figure 12A). Additionally, we observed a strong positive correlation between the expression of the genes CXCL11 and PRSS16 and the IC50 values of the drugs. The different treatment responses to various targeted drugs' IC50 values were also evident between the high- and low-RS groups (Figures 12B–I).

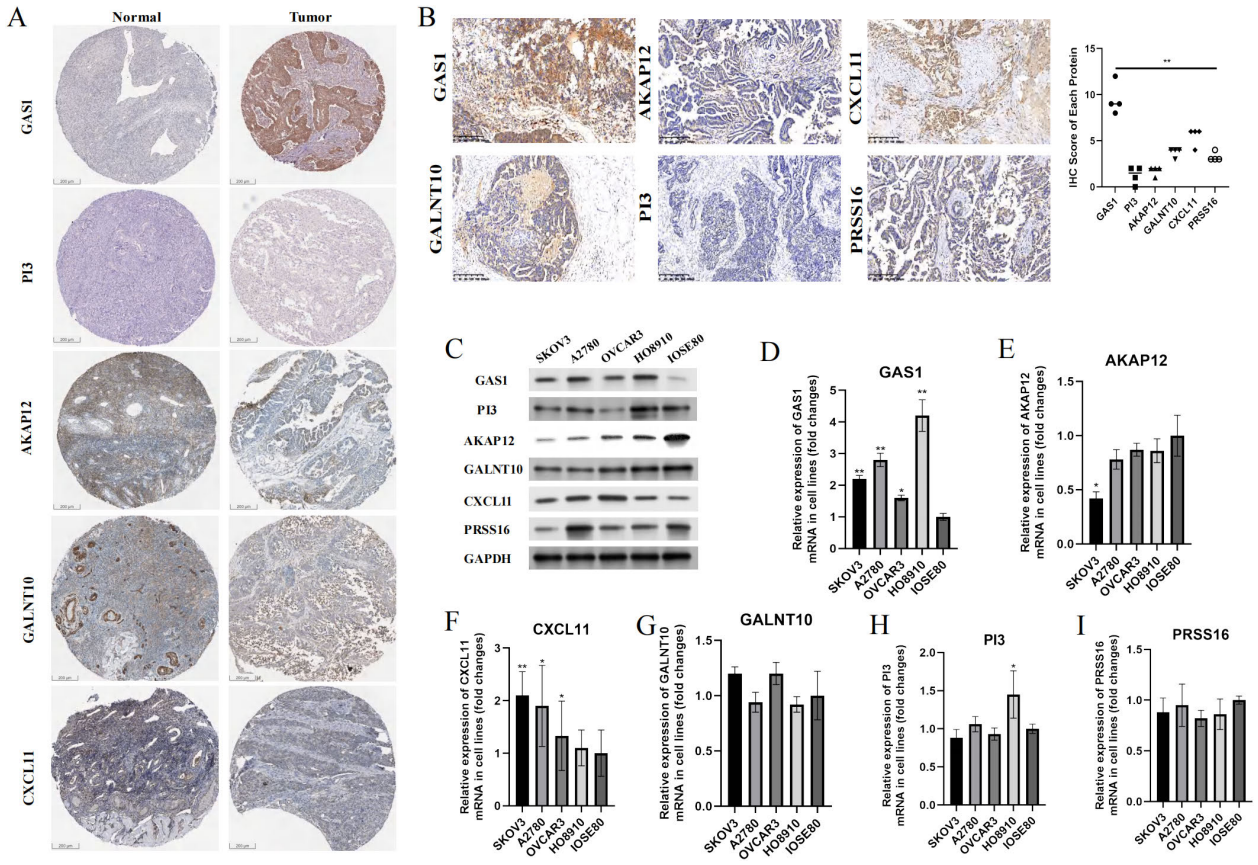


FIGURE 13
Validation of the expression levels of hub genes in samples. **(A)** IHC results of proteins in RS model from HPA database. **(B)** IHC results of proteins from RS model in our clinical samples. **(C)** Western-blot assay results of proteins from RS model in cell lines. **(D-I)** qPCR results of markers from RS model in cell lines. **p < 0.001.

3.6 Validation of the expression levels of hub genes in samples

In addition to PRSS16, we explored the protein expression of the other hub genes in the HPA database (https://www.proteinatlas.org/). Among them, GAS1 was significantly overexpressed in OC tissues, whereas AKAP12 and GALNT10 were markedly upregulated in normal tissues. PI3 exhibited low expression in both normal and OC tissues. The remaining proteins were moderately expressed in normal and OC tissues (Figure 13A). To further corroborate the results obtained from the HPA database, we conducted validation in OC cell lines and specimens at the protein and mRNA levels through Western blotting and qPCR assays (Figures 13B–I), and we discovered that the outcomes aligned with those derived from the HPA database. Due to the coefficient scores, we identified GAS1 as a crucial oncogene from the RS model. Knocking down GAS1 in A2780 and SKOV3 cells resulted in a notable decrease in the proliferative and invasive capacities of OC cells (Figures 14A–D). Most notably, diminished EDU fluorescence and weakened reparative abilities were observed following GAS1 knockdown (Figures 14E, F).

4 Discussion

OC is the most aggressive reproductive system cancer in females worldwide, with a low survival of less than 35% by five years after diagnosis, despite various improved treatment strategies, such as cytoreductive surgery, modified chemotherapy and hormonal therapy (19). Thus, the discovery and development of individualized diagnostic and therapeutic strategies are urgently needed. Fortunately, as medicine has evolved from empirical to evidence-based, personalized, precision medicine has become increasingly achievable. Against the backdrop of the emerging era of bioinformatics, many genetic signatures and corresponding risk models have been mined from internationally available genomic databases and increasingly accepted by the scientific community (20).

Due to the tumor complexity/heterogeneity of neoplasms and their surrounding tumor microenvironment (TME) (5, 21), most OC patients experience recurrence after first-line treatment. Several cells involved in both innate and adaptive immunity, including tumor-associated macrophages (TAMs), tumor-associated neutrophils (TANs), myeloid-derived suppressor cells (MDSCs), γδ T cells, and natural killer (NK) cells, directly or indirectly shape

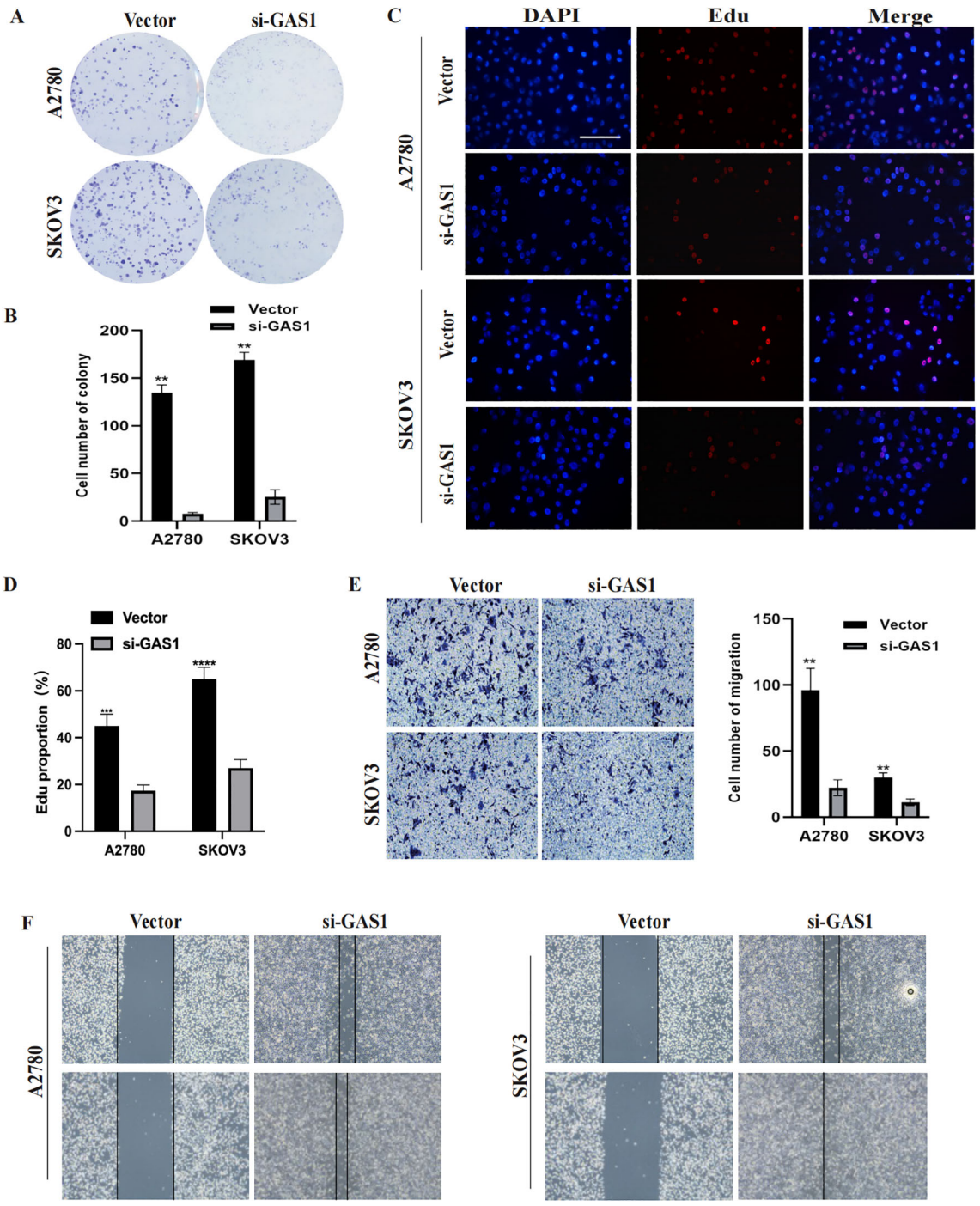


FIGURE 14
GAS1 regulates OC cell progression. **(A, B)** Colony formation assay results of A2780 and SKOV3 cells transfected with GAS1 siRNA. **C, D)** EdU assay results for A2780 and SKOV3 cells transfected with GAS1 siRNA. **(E)** Transwell assay results of A2780 and SKOV3 cells transfected with GAS1 siRNA.. **(F)** Wound healing assay for A2780 and SKOV3 cells transfected with GAS1 siRNA. **p < 0.001, ****p < 0.0001.

the OC TME (22, 23), which displays unique features leading to immune suppression and tolerance and the impairment of immune system components, including TAMs (24), TANs (25), γδ T cells, and NK cells (26). The immune context of OC acts as a crucial orchestrator of OC progression, playing an indispensable role in predicting patient prognosis.

Cuproptosis, an unconventional cell death mechanism, is involved in numerous pathological conditions, including cancers, albeit mainly through undefined underlying mechanisms. The process of cuproptosis has an unambiguous relationship with mitochondrial respiration (27). Excess copper within cells can be transported to the mitochondria by ionophores and directly bind to

lipoylated components of the tricarboxylic acid (TCA) cycle, triggering an accumulation of lipoylated proteins and loss of iron–sulfur cluster proteins, leading to proteotoxic stress and, eventually, cell death (9). However, little work has been performed on optimizing the prognostic model in OC utilizing cuproptosis combined with immune infiltration.

Taken together, in this study, the differential expression of CRGs and TME-related genes was analyzed at the single-cell level, and the correlations between these two types of genes were described, with the active cooperative death of malignant cells further identified. We subsequently identified three molecular subtypes on the basis of differential gene expression between active and inactive cuproptosis cells and demonstrated differential enrichment via multiple analyses. In particular, a prognostic risk model integrating 6 genes was constructed, which was validated with four datasets from major public databases. In addition, the patients were divided into high- and low-risk groups based on the risk model, and the associations between the risk score and multiple features were depicted. Finally, the ability of the risk model to predict the efficacy of immunotherapy and chemotherapy drug resistance was delineated.

Six genes, including PRSS16, CXCL11, PI3, GALNT10, GAS1, and AKAP12, constitute a prognostic risk model, and the RS was calculated. The AUC value of the risk score in the diagnostic ROC analysis using four datasets covering the TCGA and GEO datasets further verified the experimental phenomenon. Among the core genes, PRSS16, encoding a thymus-specific serine protease (TSSP), which is involved in CD4+ T-cell maturation in the thymus, has tumor suppressor activity (28). CXCL11, known as interferon-inducible T-cell alpha chemoattractant (I-TAC), has been reported to be the ligand of specific atypical chemokine receptors, including CXCR7 (29) and GPR182 (30). PI3, encoding elafin, which is a serine protease inhibitor critical for host defense, is reportedly associated with an unfavorable OS and a better immunotherapy response in OC (31). Moreover, GALNT10, an enzyme that mediates protein and lipid modifications, was found to be an independent predictor of prognosis in OC patients with immunosuppression (32). AKAP12, a scaffolding protein, anchors PKA to compartmentalize cycle AMP signaling and was found to be a promoter in tumors (33). Intriguingly, GAS1, which has been reported to play a role in growth suppression, blocks entry into the S phase, prevents the cycling of normal and transformed cells, and functions as a putative tumor suppressor, whereas it had the opposite effect in our study (34). This interplay highlights the complex gene profile heterogeneity of OC, which promotes changes in the tumor microenvironment.

Notably, we investigated the relationship between the risk score and tumor immunology and the differences in tumor immunology among patients with different risk scores. We found that the risk score acts as an independent prognostic factor and is associated with the condition of the tumor microenvironment and the efficacy of immunotherapy. With respect to immunotherapy, a lower risk score was more likely to be beneficial, as it indicated a higher ESTIMATE immune score, fewer mutations in tumor suppressor genes and a greater immune response. In addition, the half-maximal inhibitory concentration (IC50) curves of 138 chemotherapy drugs in the GDSC database were examined to determine their predicted chemotherapeutic effectiveness. Among them, the IC50 values of cyclopamine, ABT-888, and AKT inhibitor were determined. VIII, and ATRA were relatively more common in the high-risk score group, whereas axitinib, BMS.509744, Elesclomol and GNF.2 were relatively more common in the high-risk score groups. Both results were statistically significant after log-rank testing. Accordingly, our prognostic risk model not only predicts the OS rate but is also conducive to more precise therapy choices.

Despite these limitations, including the lack of real-world clinical cohorts and IC50 data from benches, the present study highlights the outstanding ability of the risk model to predict the prognosis of OC and its association with tumor immunology. These findings may contribute to the development of immunotherapy- and chemotherapy-based interventions in the future.

Three studies have confirmed the role of cuproptosis-related genes in OC from the perspectives of molecular subtyping and the risk score. Compared with Li's results (35), our research focused more on the relationship between the risk score and evaluation indicators such as chemoresistance, genomic mutations, and the efficacy of immunotherapies. We focused primarily on cuproptosis without incorporating ferroptosis, aiming for a more direct exploration of the roles of CRGs in OC. Our study also focused on the role of key genes within the risk score, delving into their oncogenic functions. Unlike Wang's study (36), our research closely examines the relationships between the risk score and the tumor immune microenvironment as well as immunotherapeutic responses. Additionally, via single-cell analysis, we identified genes that were differentially expressed between the active copper-depleted malignant cell population and the inactive cell population to characterize the molecular subtypes in the TCGA ovarian cancer cohort. Finally, compared with Zhang's research (8), our strength lies in uncovering the oncogenic functions of key genes within the risk score and validating these findings in samples. Furthermore, our molecular subtyping method is more specific and precise.

In conclusion, our study provides a novel risk score model including 6 hub genes associated with cuproptosis and immune infiltration to predict OC prognosis as well as clinical efficacy.

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 authors.

Ethics statement

This research was approved by the ethics committee of Jiangxi Cancer Hospital (Approval number: 2022ky305). 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.

Author contributions

YX: Data curation, Formal analysis, Investigation, Methodology, Supervision, Writing – original draft. XL: Data curation, Formal analysis, Investigation, Validation, Writing – review & editing. XF: Data curation, Formal analysis, Project administration, Writing – review & editing. XC: Formal analysis, Investigation, Methodology, Project administration, Validation, Writing – review & editing. GJ: Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Writing – review & editing. WQ: Conceptualization, Formal analysis, Investigation, Methodology, Validation, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This project is supported by the National Natural Science Foundation of China (82060477) and the Talent Training Support Plan of Jiangxi Cancer Hospital (the Second Affiliated Hospital of Nanchang Medical College).

Acknowledgments

The authors would like to thank the participants.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1452294/full#supplementary-material

References

1. Dalmartello M, La Vecchia C, Bertuccio P, Boffetta P, Levi F, Negri E, et al. European cancer mortality predictions for the year 2022 with focus on ovarian cancer. *Ann Oncol*. (2022) 33:330–9. doi: 10.1016/j.annonc.2021.12.007

2. Duska LR, Kohn EC. The new classifications of ovarian, fallopian tube, and primary peritoneal cancer and their clinical implications. *Ann Oncol*. (2017) 28:viii8–viii12. doi: 10.1093/annonc/mdx445

3. Zhang M, Cheng S, Jin Y, Zhao Y, Wang Y. Roles of CA125 in diagnosis, prediction, and oncogenesis of ovarian cancer. *Biochim Biophys Acta Rev Cancer*. (2021) 1875:188503. doi: 10.1016/j.bbcan.2021.188503

4. Zhuang H, Tan M, Liu J, Hu Z, Liu D, Gao J, et al. Human epididymis protein 4 in association with Annexin II promotes invasion and metastasis of ovarian cancer cells. *Mol Cancer*. (2014) 13:243. doi: 10.1186/1476-4598-13-243

5. Jiang Y, Wang C, Zhou S. Targeting tumor microenvironment in ovarian cancer: Premise and promise. *Biochim Biophys Acta Rev Cancer*. (2020) 1873:188361. doi: 10.1016/j.bbcan.2020.188361

6. Xu J, Fang Y, Chen K, Li S, Tang S, Ren Y, et al. Single-cell RNA sequencing reveals the tissue architecture in human high-grade serous ovarian cancer. *Clin Cancer Res*. (2022) 28:3590–602. doi: 10.1158/1078-0432.CCR-22-0296

7. Ferri-Borgogno S, Zhu Y, Sheng J, Burks JK, Gomez JA, Wong KK, et al. Spatial transcriptomics depict ligand-receptor cross-talk heterogeneity at the tumor-stroma interface in long-term ovarian cancer survivors. *Cancer Res*. (2023) 83:1503–16. doi: 10.1158/0008-5472.CAN-22-1821

8. Zhang J, Lu M, Xu H, Ren F, Zhu L. Molecular subtypes based on cuproptosis-related genes and tumor microenvironment infiltration characterization in ovarian cancer. *Cancer Cell Int*. (2022) 22:328. doi: 10.1186/s12935-022-02756-y

9. Ge EJ, Bush AI, Casini A, Cobine PA, Cross JR, DeNicola GM, et al. Connecting copper and cancer: from transition metal signalling to metalloplasia. *Nat Rev Cancer*. (2022) 22:102–13. doi: 10.1038/s41568-021-00417-2

10. Liu D, Yang F, Zhang T, Mao R. Leveraging a cuproptosis-based signature to predict the prognosis and drug sensitivity of cutaneous melanoma. *J Transl Med*. (2023) 21:57. doi: 10.1186/s12967-023-03891-4

11. Jiang R, Huan Y, Li Y, Gao X, Sun Q, Zhang F, et al. Transcriptional and genetic alterations of cuproptosis-related genes correlated to Malignancy and immune-infiltrate of esophageal carcinoma. *Cell Death Discovery*. (2022) 8:370. doi: 10.1038/s41420-022-01164-5

12. Yan C, Niu Y, Ma L, Tian L, Ma J. System analysis based on the cuproptosis-related genes identifies LIPT1 as a novel therapy target for liver hepatocellular carcinoma. *J Transl Med*. (2022) 20:452. doi: 10.1186/s12967-022-03630-1

13. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell*. (2018) 173:400–16.e11. doi: 10.1016/j.cell.2018.02.052

14. Yang K, Shi Y, Luo M, Mao M, Zhang X, Chen C, et al. Identification of a unique tumor cell subset employing myeloid transcriptional circuits to create an immunomodulatory microenvironment in glioblastoma. *Oncoimmunology*. (2022) 11:2030020. doi: 10.1080/2162402X.2022.2030020

15. Cai Z, He Y, Yu Z, Hu J, Xiao Z, Zu X, et al. Cuproptosis-related modification patterns depict the tumor microenvironment, precision immunotherapy, and prognosis of kidney renal clear cell carcinoma. *Front Immunol*. (2022) 13:933241. doi: 10.3389/fimmu.2022.933241

16. Chen H, Ye C, Wu C, Zhang J, Xu L, Wang X, et al. Berberine inhibits high fat diet-associated colorectal cancer through modulation of the gut microbiota-mediated lysophosphatidylcholine. *Int J Biol Sci*. (2023) 19:2097–113. doi: 10.7150/ijbs.81824

17. Xu C, Zheng L, Li D, Chen G, Gu J, Chen J, et al. CXCR4 overexpression is correlated with poor prognosis in colorectal cancer. *Life Sci*. (2018) 208:333–40. doi: 10.1016/j.lfs.2018.04.050

18. Zhu X, Zhang Y, Wang Y, Zhang H, Wang X, Tang H, et al. Agrimoniin sensitizes pancreatic cancer to apoptosis through ROS-mediated energy metabolism dysfunction. *Phytomedicine*. (2022) 96:153807. doi: 10.1016/j.phymed.2021.153807

19. Kurian AW, Abrahamse P, Furgal A, Ward KC, Hamilton AS, Hodan R, et al. Germline genetic testing after cancer diagnosis. *JAMA*. (2023) 330:43–51. doi: 10.1001/jama.2023.9526

20. Rebbeck TR, Mitra N, Wan F, Sinilnikova OM, Healey S, McGuffog L, et al. Association of type and location of BRCA1 and BRCA2 mutations with risk of breast and ovarian cancer. *JAMA*. (2015) 313:1347–61. doi: 10.1001/jama.2014.5985

21. Schoutrop E, Moyano-Galceran L, Lheureux S, Mattsson J, Lehti K, Dahlstrand H, et al. Molecular, cellular and systemic aspects of epithelial ovarian cancer and its

tumor microenvironment. *Semin Cancer Biol*. (2022) 86:207–23. doi: 10.1016/j.semcancer.2022.03.027

22. Mukherjee A, Bilecz AJ, Lengyel E. The adipocyte microenvironment and cancer. *Cancer Metastasis Rev*. (2022) 41:575–87. doi: 10.1007/s10555-022-10059-x

23. Maniati E, Berlato C, Gopinathan G, Heath O, Kotantaki P, Lakhani A, et al. Mouse ovarian cancer models recapitulate the human tumor microenvironment and patient response to treatment. *Cell Rep*. (2020) 30:525–40.e7. doi: 10.1016/j.celrep.2019.12.034

24. Miyamoto T, Murakami R, Hamanishi J, Tanigaki K, Hosoe Y, Mise N, et al. B7-H3 suppresses antitumor immunity via the CCL2-CCR2-M2 macrophage axis and contributes to ovarian cancer progression. *Cancer Immunol Res*. (2022) 10:56–69. doi: 10.1158/2326-6066.CIR-21-0407

25. Yang M, Zhang G, Wang Y, He M, Xu Q, Lu J, et al. Tumour-associated neutrophils orchestrate intratumoural IL-8-driven immune evasion through Jagged2 activation in ovarian cancer. *Br J Cancer*. (2020) 123:1404–16. doi: 10.1038/s41416-020-1026-0

26. Hornburg M, Desbois M, Lu S, Guan Y, Lo AA, Kaufman S, et al. Single-cell dissection of cellular components and interactions shaping the tumor immune phenotypes in ovarian cancer. *Cancer Cell*. (2021) 39:928–44.e6. doi: 10.1016/j.ccell.2021.04.004

27. Tong X, Tang R, Xiao M, Xu J, Wang W, Zhang B, et al. Targeting cell death pathways for cancer therapy: recent developments in necroptosis, pyroptosis, ferroptosis, and cuproptosis research. *J Hematol Oncol*. (2022) 15:174. doi: 10.1186/s13045-022-01392-3

28. Brisson L, Pouyet L, N'guessan P, Garcia S, Lopes N, Warcollier G, et al. The thymus-specific serine protease TSSP/PRSS16 is crucial for the antitumoral role of CD4 (+) T cells. *Cell Rep*. (2015) 10:39–46. doi: 10.1016/j.celrep.2014.12.009

29. Singh AK, Arya RK, Trivedi AK, Sanyal S, Baral R, Dormond O, et al. Chemokine receptor trio: CXCR3, CXCR4 and CXCR7 crosstalk via CXCL11 and CXCL12. *Cytokine Growth Factor Rev*. (2013) 24:41–9. doi: 10.1016/j.cytogfr.2012.08.007

30. Melgrati S, Radice E, Ameti R, Hub E, Thelen S, Pelczar P, et al. Atlas of the anatomical localization of atypical chemokine receptors in healthy mice. *PloS Biol*. (2023) 21:e3002111. doi: 10.1371/journal.pbio.3002111

31. Al Bakir M, Gabra H. The molecular genetics of hereditary and sporadic ovarian cancer: implications for the future. *Br Med Bull*. (2014) 112:57–69. doi: 10.1093/bmb/ldu034

32. Zhang G, Lu J, Yang M, Wang Y, Liu H, Xu C. Elevated GALNT10 expression identifies immunosuppressive microenvironment and dismal prognosis of patients with high grade serous ovarian cancer. *Cancer Immunol Immunother*. (2020) 69:175–87. doi: 10.1007/s00262-019-02454-1

33. Deng Y, Gao J, Xu G, Yao Y, Sun Y, Shi Y, et al. HDAC6-dependent deacetylation of AKAP12 dictates its ubiquitination and promotes colon cancer metastasis. *Cancer Lett*. (2022) 549:215911. doi: 10.1016/j.canlet.2022.215911

34. Li Q, Qin Y, Wei P, Lian P, Li Y, Xu Y, et al. Gas1 inhibits metastatic and metabolic phenotypes in colorectal carcinoma. *Mol Cancer Res*. (2016) 14:830–40. doi: 10.1158/1541-7786.MCR-16-0032

35. Li Y, Fang T, Shan W, Gao Q. Identification of a novel model for predicting the prognosis and immune response based on genes related to cuproptosis and ferroptosis in ovarian cancer. *Cancers (Basel)*. (2023) 15:579. doi: 10.3390/cancers15030579

36. Wang K, Zhang Y, Ao M, Luo H, Mao W, Li B. Multi-omics analysis defines a cuproptosis-related prognostic model for ovarian cancer: Implication of WASF2 in cuproptosis resistance. *Life Sci*. (2023) 332:122081. doi: 10.1016/j.lfs.2023.122081

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 25 November 2024
DOI 10.3389/fimmu.2024.1496886

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong, Hong Kong SAR, China

REVIEWED BY
Yuquan Chen,
Monash University, Australia
Xin Yu,
Baylor College of Medicine, United States
Kaige Chen,
Wake Forest University, United States

*CORRESPONDENCE
Xiaoyuan Qiao
qiaoxiaoyuan@sxmu.edu.cn

RECEIVED 15 September 2024
ACCEPTED 06 November 2024
PUBLISHED 25 November 2024

CITATION
Li Y, Chen Y, Zhang Y, Fang Y, Wu L, Zhao Y, Wang D and Qiao X (2024) Integrating multi-omics techniques and *in vitro* experiments reveals that GLRX3 regulates the immune microenvironment and promotes hepatocellular carcinoma cell proliferation and invasion through iron metabolism pathways.
Front. Immunol. 15:1496886.
doi: 10.3389/fimmu.2024.1496886

Integrating multi-omics techniques and *in vitro* experiments reveals that GLRX3 regulates the immune microenvironment and promotes hepatocellular carcinoma cell proliferation and invasion through iron metabolism pathways

Yang Li[1], Yuan Chen[1,2], Yang Zhang[3], Yunsheng Fang[4,5], Ling Wu[6], Ying Zhao[1], Danqiong Wang[1] and Xiaoyuan Qiao[7]*

[1]Department of General Medicine, Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Third Hospital of Shanxi Medical University, Tongji Shanxi Hospital, Taiyuan, China, [2]Department of Geriatrics, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [3]School of Mechanical Engineering, Taiyuan University of Science and Technology, Taiyuan, China, [4]The Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, China, [5]Bioinspired Engineering & Biomechanics Center (BEBC), Xi'an Jiaotong University, Xi'an, China, [6]Tumor Center, Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Third Hospital of Shanxi Medical University, Tongji Shanxi Hospital, Taiyuan, China, [7]Department of Comprehensive Medicine, Shanxi Province Cancer Hospital/Shanxi Hospital Affiliated to Cancer Hospital, Chinese Academy of Medical Sciences/Cancer Hospital Affiliated to Shanxi Medical University, Taiyuan, China

Background: Hepatocellular carcinoma (HCC) is a common malignancy worldwide, and its development is closely related to abnormalities in iron metabolism. This study aims to systematically analyze changes in iron metabolism in the tumor microenvironment of HCC using single-cell sequencing technology, and investigate the potential mechanisms by which iron metabolism regulation affects the survival of liver cancer patients.

Materials and methods: Single-cell sequencing data from hepatocellular carcinoma patients were obtained from the GEO database. By iron metabolism genomic scoring, we assessed differences in iron metabolism levels in hepatocellular carcinoma samples. By cell communication analysis as well as GO and KEGG enrichment analysis, we determined the functional role of iron metabolism in different cell types. We used survival analysis and Kaplan-Meier curves to assess the impact of iron metabolism levels on patient prognosis. In addition, we identified and analyzed the expression profile of the GLRX3 gene, investigated its key regulatory role in iron metabolism, and validated its clinical value as a prognostic marker. Finally, we explored the effect of GLRX3 on hepatocellular carcinoma phenotype by *in vitro* experiments such as PCR, transwell, CCK8, and wound healing assay.

Results: Bioinformatics results and experimental validation confirmed the dysregulation of iron metabolism in the development of hepatocellular carcinoma, revealing iron's regulatory influence across various cell types. Additionally, GLRX3 was identified as a key regulatory factor in iron metabolism, and the mechanism by which GLRX3 regulates tumor cell proliferation and immune evasion was determined. Furthermore, experiments verified GLRX3's role in facilitating tumor cell proliferation and invasion.

Conclusion: This study highlights the critical role of iron metabolism in the progression of hepatocellular carcinoma, particularly the regulatory mechanism of the GLRX3 gene in tumor cell proliferation and immune evasion. Iron metabolism abnormalities are not only drivers of liver cancer development but also key indicators of patient prognosis.

KEYWORDS

iron metabolism, GLRX3, immunotherapy, precision medicine, multi-omics analysis, immune signatures, personalized therapy

1 Introduction

Primary liver cancer (PLC) is the fifth most common cancer globally and the second leading cause of cancer-related deaths (1–4). Its incidence and mortality rates are rising rapidly, particularly in Western countries. Hepatocellular carcinoma (HCC), the most prevalent form of PLC, makes up 90% of all primary liver tumors and around 5% of all cancers (5–8). The malignant transformation of hepatocytes results in HCC (9), with known risk factors including excessive alcohol consumption, hepatitis B virus (HBV) infection, fat accumulation in the liver, and autoimmune liver diseases (10–12). While liver transplantation, surgery, and local therapies can be curative at early stages (13, 14), most liver cancer patients are diagnosed late, where treatment options are extremely limited and the prognosis is poor (15–17). Thus, understanding changes in the tumor microenvironment during liver cancer progression and gaining deeper insights into its pathogenesis are critical for developing effective treatments (18).

Iron metabolism encompasses the comprehensive processes of iron absorption, transport, storage, and utilization within a biological system. Although the body's requirement for iron is relatively modest, it is an essential trace element that plays a critical role in numerous physiological processes (19). The key components involved in maintaining cellular iron homeostasis include transferrin receptor 1 (TfR1), which internalizes transferrin-bound iron; ferroportin (Fpn), the sole iron export protein; and ferritin, which stores excess iron (20). Iron levels are tightly regulated at both systemic and cellular levels to remain within an optimal range. However, excessive iron can promote the production of highly reactive and toxic oxidants via the Fenton reaction, impairing immune function and disrupting various physiological processes (21).

Hepatocytes play a crucial role in maintaining stable plasma glucose and lipoprotein levels in humans (22). Under normal conditions, hepatocytes remain quiescent; however, when liver tissue is excessively exposed to viruses, toxic substances, or metabolites, significant physiological changes occur. Given that the liver is a primary organ for excess iron accumulation, it plays a crucial role in maintaining iron homeostasis (19, 23). Dysregulation of iron metabolism significantly increases the risk of liver cancer. Research has shown that iron overload is not only associated with cancer development but also actively contributes to carcinogenesis. Excess iron induces oxidative stress-mediated DNA damage in hepatocytes and promotes the rapid proliferation of tumor cells (24, 25). Thus, understanding the intrinsic link between abnormal iron metabolism and changes in the tumor microenvironment of liver cancer is critical for developing precise treatment strategies and for uncovering the broader impact of metal ions on cancer progression (22, 26).

2 Materials and methods

2.1 Cell culture

The human hepatocellular carcinoma cell lines, Hep3B and Huh7, were cultured in RPMI/1640 medium (Gibco) supplemented with 10% fetal bovine serum (FBS) (Hyclone), along with 100 U/L of penicillin and 100 mg/L of streptomycin (Thermo Fisher). Cells were maintained at 37°C in a humidified atmosphere of 5% CO2. The culture medium was changed every 2-3 days to ensure optimal growth conditions. When cells reached 80-90% confluence, they were passaged using trypsin-EDTA for further experiments.

2.2 shRNA knockdown

Plasmids expressing shRNA, specifically designed to target GLRX3, were carefully constructed with the assistance of GenePharma. During cultivation, the cells were treated with viral supernatants and polybrene (Sigma Aldrich) in the culture medium. After 24 hours of incubation, the cells were transferred to fresh medium containing 2.0 μg/ml of puromycin. The efficiency of GLRX3 knockdown was confirmed two days later using qRT-PCR analysis.

2.3 qPCR assay

Total RNA extraction was carried out utilizing the RNA Eazy Fast Tissue/Cell Kit (TIANGEN Biotech) in accordance with the manufacturer's guidelines. Subsequently, cDNA synthesis was performed using the FastKing RT Kit (TIANGEN Biotech), adhering to the provided protocol. Real-time PCR analysis was conducted with the application of the SuperReal PreMix Plus (TIANGEN Biotech) reagent, implemented on the StepOnePlus Real-Time PCR System. The PCR reaction encompassed an initial pre-denaturation phase at 95°C for 15 minutes, followed by 40 amplification cycles, comprising denaturation at 95°C for 10 seconds, annealing at 72°C for 20 seconds, and extension at 60°C for 20 seconds. Primer sequences utilized were procured from Sangon Biotech. (Species of Human Origin) GLRX3 Forward Primer: GGGCGGCTGAGGCAGCT,reverse primer GCAGG GGGCAGCATGAGTC;(Species of Human Origin) IL10 Forward Primer: GACTTTAAGGGTTACCTG GGTTG,Reverse Primer: TCACATGCGCCTTGATGTCTG; At last, PCR signals 2-44Ct was used to calculate the expression of genes mRNA levels. The following sequences were used: 5′-GTGGAAATTCTTCA CAAACAT-3′ for human GLRX3 shRNA and 5′-GGAATC TCATTCGATGCATAC-3′ for the control shRNA.

2.4 Transwell assay

A seeding density of 1×10^5 cells was allocated to either Matrigel-coated chambers (BD Biosciences, San Jose, CA) for the invasion assay or uncoated chambers designated for the migration assay. The upper chamber was filled with serum-free medium, while the lower chamber was supplied with complete RPMI/1640 medium. Following a 24-hour incubation period, cells that had traversed the membrane were meticulously fixed with a 4% paraformaldehyde solution and subsequently subjected to staining with 0.1% crystal violet. Cell quantification was carried out using a light microscope, specifically the Thermo Fisher instrument based in Waltham, MA, USA.

2.5 CCK-8 assay

Cell viability was ascertained via the Cell Counting Kit-8 (CCK-8) assay. Cells were appropriately seeded at a density of 1500 cells per well, and each well contained 200 μl of complete medium within 96-well plates. Subsequent to seeding, the cells were diligently cultured under standard conditions at 37°C. Following each experimental procedure, 20 μl of CCK-8 reagent (Beyotime) was introduced into every well. A further incubation period of 2 hours ensued, after which the optical density value (OD450nm) was meticulously determined utilizing a microplate reader.

2.6 Wound healing assay

A wound healing assay was conducted to evaluate the migratory capacity of hepatocellular carcinoma cells. Transfected cells in six-well plates were incubated at 37°C until they reached around 80% confluence. Then, a 200 μL sterile pipette tip was used to create uniform wounds in the cell monolayer. Cells were washed twice with phosphate-buffered saline to remove any debris, and the medium was replaced with serum-free medium. Cell migration into the wound area was carefully monitored under an Olympus inverted microscope at 0 and 24 hours.

2.7 Protein expression and immunohistochemistry

We used the CTPAC database to validate the difference in the expression of GLRX3 protein in hepatocellular carcinoma tissues and normal liver tissues. The expression levels of GLRX3 in hepatocellular carcinoma tissues and normal tissues were verified by immunohistochemical sections from the HPA database.

2.8 Data sources

The single-cell sequencing data used in this study was obtained from the GEO database, specifically from dataset GSE149614, which includes sequencing data from 10 HCC patients. We selected two types of samples, primary tumors and non-tumorous liver tissues, for analysis. Spatial transcriptomics sequencing data from one HCC tumor tissue sample was sourced from GSM6177612, with tissue sections derived from primary hepatocellular carcinoma regions. Additionally, RNA-seq data for HCC was downloaded from the UCSC Xena platform (https://xena.ucsc.edu/), originating from the TCGA (The Cancer Genome Atlas) cohort. This dataset contains sequencing information from 424 samples along with corresponding survival data, which was used for survival analysis. External validation sets utilized in this study included GSE144269, GSE76427, and ICGC_LIRI.

2.9 Single-cell sequencing data processing

After processing single-cell sequencing data from 10 tumor and 8 normal liver samples, we obtained a total of 63,101 cells. Preliminary data analysis was conducted using the Seurat package, which included quality control, dimensionality reduction, clustering, and visualization. To ensure the reliability of the sequencing data, stringent quality control measures were implemented. Specifically, cells with fewer than 500 or more than 6,000 detected genes, as well as those with over 20% mitochondrial gene content, were excluded. This step minimized the presence of empty droplets, doublets, and senescent cells. Following data normalization and scaling, PCA-based dimensionality reduction was performed, and batch effects were mitigated using the Harmony package. We then selected the top 20 principal components for clustering with a resolution of 0.3, resulting in the identification of 17 cell clusters, which were visualized using UMAP (27–29).

2.10 Cell type identification and subpopulation segmentation

We employed common cell marker genes and the "FindAllMarkers" function to conduct preliminary cell type identification. Based on the expression patterns of marker genes in each cluster and the upregulation of specific genes, we assigned cell type labels. Subpopulations within larger groups, such as myeloid cells, B cells, and T/NK cells, were further subdivided. Using a resolution of 0.1, we identified distinct cell types, including plasma cells, cytotoxic T lymphocytes (CTLs), epithelial-mesenchymal transition cells (EMTs), regulatory T cells (Tregs), and macrophages (30, 31).

2.11 Tumor cell identification and stemness assessment

To identify tumor cells, we utilized the "copykat" package for copy number variation (CNV) analysis. CopyKAT (Copy-number Karyotyping of Tumors) is a computational tool that employs an integrative Bayesian approach to detect whole-genome aneuploidy in single cells at a 5MB resolution, allowing us to distinguish tumor cells from normal cells. Cells displaying extensive whole-genome CNV (aneuploidy) were classified as tumor cells, while stromal and immune cells typically exhibited 2N diploid or near-diploid CNV profiles. To assess the differentiation status of tumor cells and support pseudotime analysis of T cell subpopulations, we applied the "cytotrace" package for cell stemness scoring. Cytotrace provides a continuous measure of developmental potential, ranging from 0 (fully differentiated) to 1 (pluripotent). Pseudotime inference for T cell subpopulations was performed using the "monocle" package (32).

2.12 Iron metabolism level assessment

To quantify iron metabolism levels across different cell types using 73 iron metabolism-related genes, we applied several gene set scoring methods, including AddModuleScore, ssGSEA, AUCell, UCell, and singscore. Each method generated a score for each cell, and after centering and standardizing these scores, the final score for each cell was obtained by summing the five scores. The use of multiple scoring methods helps reduce errors and biases in gene set scoring, providing more comprehensive information, increased robustness, and better biological interpretation. For cell types that showed significant changes in iron metabolism levels between groups, cells were classified into high- and low-score groups based on the average score, representing different levels of iron metabolism (27, 33, 34).

2.13 Enrichment and cell communication analysis

To investigate the biological functional differences among cells with varying iron metabolism levels, we conducted Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses. Genes for enrichment analysis were identified using the "FindMarkers" function, focusing on those upregulated in the high-score group cells. The "clusterProfiler" package was employed to retrieve gene sets from the GO, KEGG, and GSEA databases and to visualize the results. Additionally, we utilized the "GSVA" package, which employs the "HALLMARK" gene set to identify tumor-associated biological processes. To compare differences in cell communication between high- and low-score cells, we conducted cell communication network analysis using the "CellChat" package. CellChat simulates and analyzes intercellular communication by integrating gene expression data with known interactions between signaling ligands, receptors, and cofactors (35–37).

2.14 Infiltration and prognostic analysis of high- and low-score cells

Using the marker genes of high- and low-score cells, we performed ssGSEA scoring on TCGA data to classify patients into high- and low-infiltration groups. Survival data from these groups were then used to plot Kaplan-Meier (K-M) curves, allowing us to compare prognostic differences. The "survival" and "survminer" packages were employed to plot K-M curves for both overall survival and progression-free survival (38).

2.15 Spatial transcriptomics data deconvolution analysis

For the initial processing of spatial transcriptomics data, we utilized the "Seurat" package. During quality control, only mitochondrial and ribosomal genes were excluded, while data for each spot were retained. After normalization and centering using the "SCTransform" function, we performed PCA-based dimensionality reduction and clustering. We clustered the data using the top 20 principal components, resulting in the identification of 7 cell clusters. The "scMetabolism" package was employed to infer metabolic activity in each cell cluster from the spatial

transcriptomics data. This package includes human-specific metabolic gene sets covering 85 KEGG pathways and 82 REACTOME entries, employing the VISION algorithm to score each cell. To address the resolution limitations of spatial transcriptomics and leverage spatial location information, deconvolution analysis was performed using the "spacexr" package, specifically the RCTD deconvolution analysis. Annotated single-cell data were used to deconvolute spatial transcriptomics data, inferring the probability of each cell type at each sequencing spot. Cells from high- and low-score groups were also included in the analysis to compare iron metabolism levels across different locations.

2.16 Expression and prognostic analysis of key iron metabolism genes in tumors

For key iron metabolism genes, we performed differential gene expression analysis using TCGA data and validated the results with GEO data. Six significantly differentially expressed iron metabolism genes were then used to score bulk data, categorizing patients into high and low groups for comparison of prognostic differences, reflecting the impact of key iron metabolism genes on HCC prognosis. The ssGSEA method was employed to score and plot K-M curves using various survival datasets. Two additional datasets from GEO and ICGC were used as external validation sets to assess the impact of key iron metabolism genes on HCC patient prognosis. Furthermore, we examined the expression of key iron metabolism genes in spatial transcriptomics data, comparing gene expression in normal cells, mixed cells, and malignant cells, and their correlation with various cell types.

2.17 Prognostic and clinical analysis of GLRX3, a key iron metabolism gene

For GLRX3, a key gene in iron metabolism, we conducted subgroup differential expression analysis using clinical information from TCGA. The prognostic value of GLRX3 was evaluated using TCGA and multiple external validation sets. Enrichment analysis and spatial transcriptomics data were also employed in the study of GLRX3.

2.18 Statistical analysis

Statistical analyses were performed using R 4.2.2 64-bit version and its supported packages. The non-parametric Wilcoxon rank-sum test was used to assess relationships between groups for continuous variables. Spearman correlation analysis was used to test correlation coefficients. All statistical analyses were considered significant at $P<0.05$.

3 Results

3.1 Quality control of liver cancer samples

In this study, we obtained single-cell transcriptomic data from the GEO database (dataset GSE149614), which includes 18 liver cancer tumor tissue samples and adjacent normal liver tissue samples from ten patients. To ensure high-quality single-cell data analysis, we first performed quality control on all samples. To minimize the impact of aging cells, red blood cells, and a high percentage of mitochondrial reads, we evaluated key quality metrics, such as UMI counts and the percentages of mitochondrial and hemoglobin gene expression (Figure 1A). Additionally, we employed the Harmony package to correct for potential batch effects in sequencing, ensuring that observed differences were due to biological variation between samples (Figure 1B). After dimensionality reduction and clustering, we visualized 61,776 cells that passed quality control, which were grouped into 16 distinct clusters via UMAP (Figure 1C). Furthermore, we analyzed differences in data distribution across samples (Figure 1D), between tumor and normal groups (Figure 1E), and in mRNA density (Figure 1F).

Next, using common cell marker genes and the "FindAllMarkers" function, we performed preliminary cell type identification. Based on the expression patterns of marker genes and upregulated genes, we named the cell types (Figure 1G). Figure 1H displays the distribution of different cell types across the tumor and normal groups, while the heatmap in Figure 1I shows the marker genes for each cell cluster. For the large groups of myeloid cells, B cells, and T/NK cells, we further subdivided the populations, using a resolution of 0.1, identifying plasma cells, CTLs, EMTs, Tregs, macrophages, and more (Figures 1J–M). Lastly, we displayed the distribution differences of all cell types across the tumor and normal groups (Figure 1N).

3.2 Tumor cell identification

To identify tumor cells, we used the "copykat" package for copy number variation (CNV) analysis, which distinguishes tumor from normal cells by identifying aneuploidy. Cells exhibiting extensive genome-wide CNV were classified as tumor cells. Figure 2A shows a group of cells with high levels of CNV abnormalities detected by copykat. Additionally, we performed stemness scoring using the "cytotrace" package (Figure 2B). By combining these results with those from copykat, we confirmed that hepatocytes constituted a highly malignant tumor cell population. Next, we applied five scoring methods (AddModuleScore, ssGSEA, AUCell, UCell, and singscore) to assess the expression of iron metabolism genes across different cell populations (Figures 2C, D). We also compared the iron metabolism scores of each cell type between tumor and normal groups (Figure 2E).

10.3389/fimmu.2024.1496886

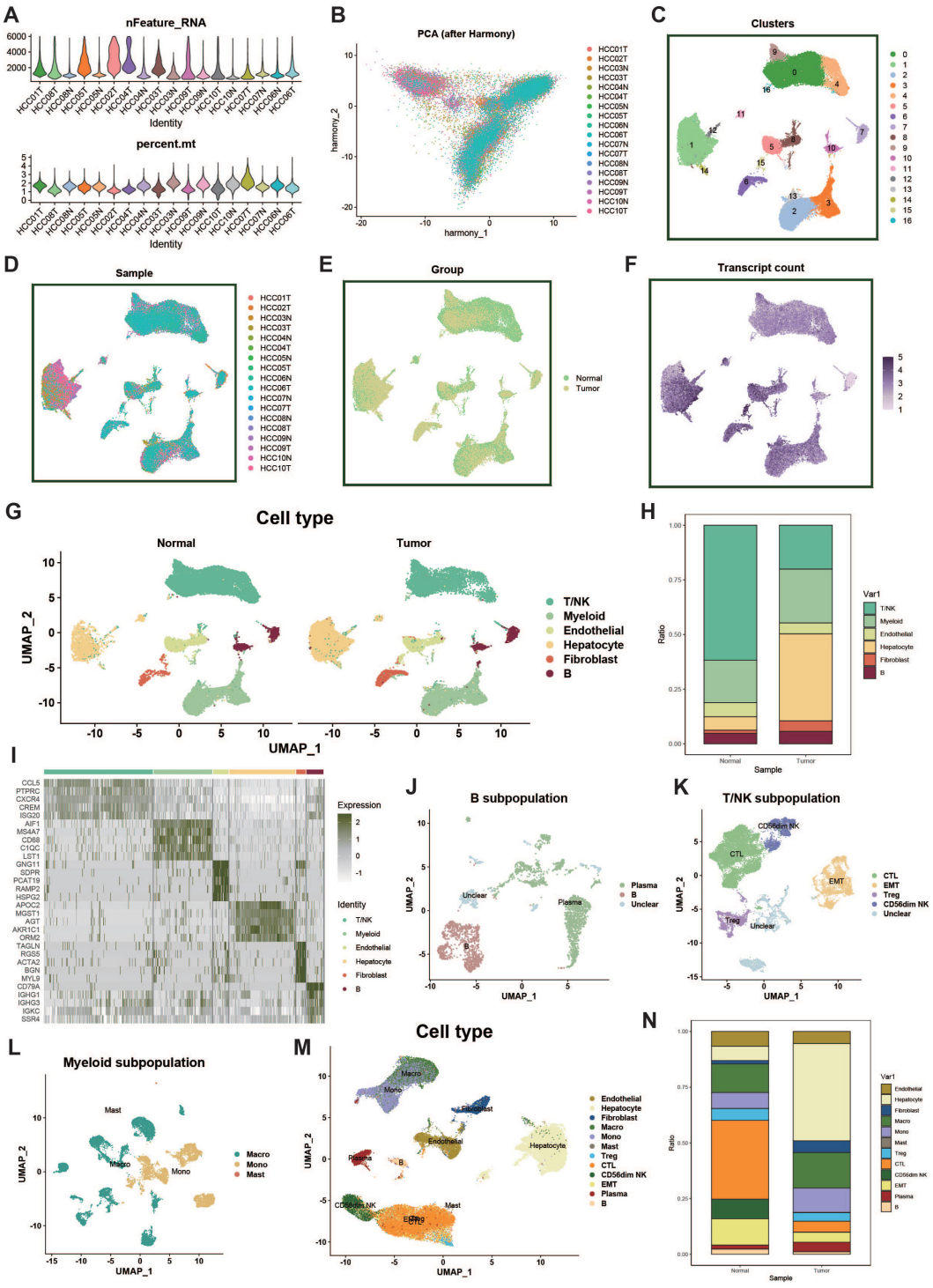


FIGURE 1
Single-cell data processing and cell type identification. **(A)** Violin plot showing sample characteristics after single-cell data quality control. The upper plot displays the number of detected genes, while the lower plot shows the proportion of mitochondrial genes. **(B)** PCA plot of cell distribution across samples after batch effect removal. **(C)** UMAP of dimensionality-reduced clustered cell populations, with a total of 17 clusters. **(D)** UMAP showing cell distribution across different samples. **(E)** UMAP showing cell distribution across different groups. **(F)** UMAP of cell counts. **(G)** Results of cell type identification, displaying the distribution and number of each cell type in different groups. **(H)** Bar plot of cell proportions. **(I)** Heatmap of cell marker gene expression. **(J)** UMAP of B cell subpopulations. **(K)** UMAP of T/NK cell subpopulations. **(L)** UMAP of myeloid cell subpopulations. **(M)** UMAP of overall cell types. **(N)** Bar plot showing proportions of overall cell types.

10.3389/fimmu.2024.1496886

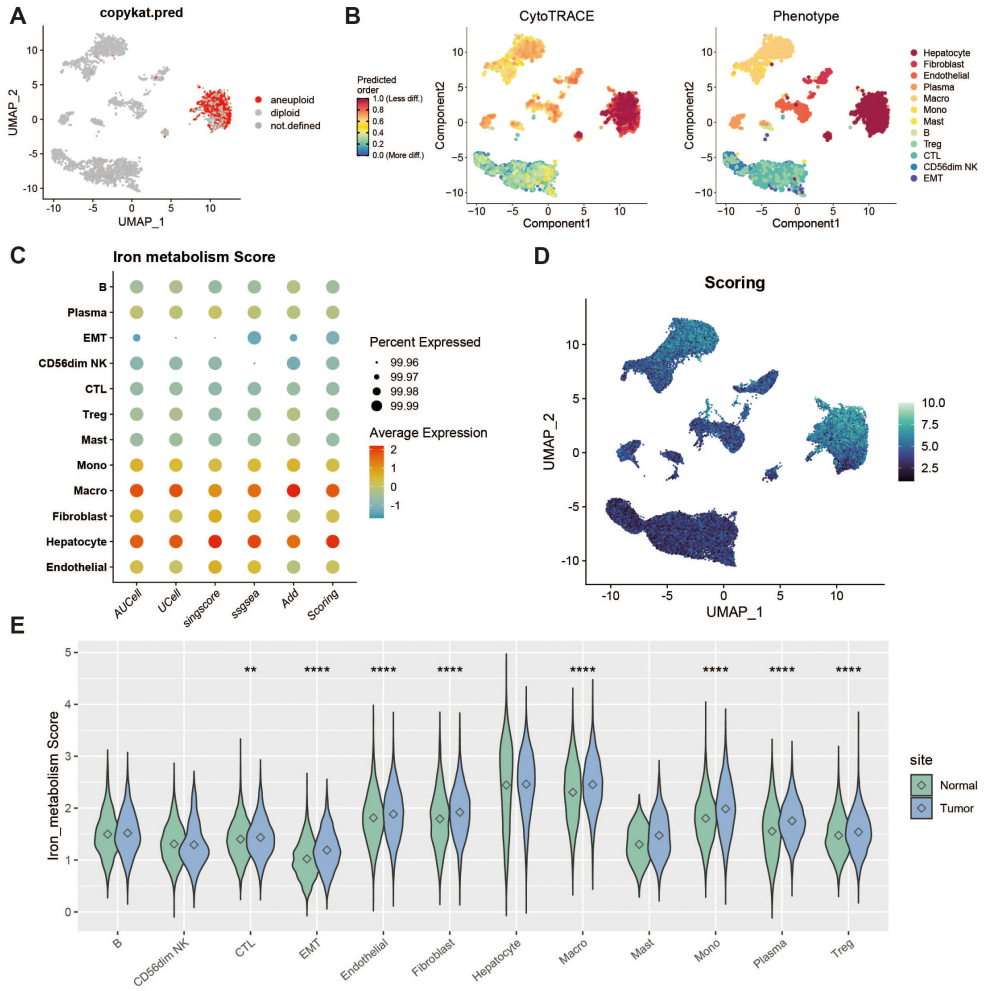


FIGURE 2
Tumor microenvironment analysis and iron metabolism level assessment. **(A)** UMAP showing copy number variation, where red indicates polyploid cells (tumor cells). **(B)** Heatmap of stemness score, ranging from 0 (differentiated) to 1 (pluripotent). **(C)** Bubble plot of gene set scoring results. **(D)** Heatmap of gene set scores. **(E)** Violin plot of score differences between tumor and normal groups. ** represents a p-value < 0.01, **** represents a p-value < 0.0001.

3.3 Plasma cell iron metabolism analysis

To investigate the characteristics of iron metabolism in plasma cells within the liver cancer tumor microenvironment, we categorized plasma cells into high and low expression groups based on their iron metabolism scores (Figure 3A). Notably, plasma cells in the tumor group exhibited significantly higher iron metabolism scores compared to those in the normal group (Figure 3B). To assess the heterogeneity between the two groups, we performed Gene Set Variation Analysis (GSVA), which revealed functional differences between plasma cells with high and low iron metabolism scores (Figure 3C). Plasma cells with elevated iron metabolism scores demonstrated enhanced lipogenesis, metabolic activity, and oxidative phosphorylation. Furthermore, we evaluated the expression of antibody secretion-related genes and observed a reduction in antibody secretion functionality in the high iron metabolism group (Figure 3D). Cell communication analysis indicated that plasma cells with high iron metabolism scores exhibited stronger communication and signaling output (Figures 3E, F). GO enrichment analysis indicated that plasma cells with high iron metabolism scores exhibited increased iron ion transport and oxidative response capabilities (Figure 3G). KEGG pathway analysis suggested that these cells were more active in ferroptosis and pyrimidine/nucleotide metabolism (Figure 3H). Finally, survival curve analysis revealed that patients with high iron metabolism had shorter overall survival (OS) and progression-free survival (PFS) compared to those with low iron metabolism (Figures 3I–K).

3.4 Cytotoxic T cell iron metabolism analysis

To assess the impact of iron metabolism on immune cytotoxicity, we categorized cytotoxic T cells (CTLs) into high

10.3389/fimmu.2024.1496886

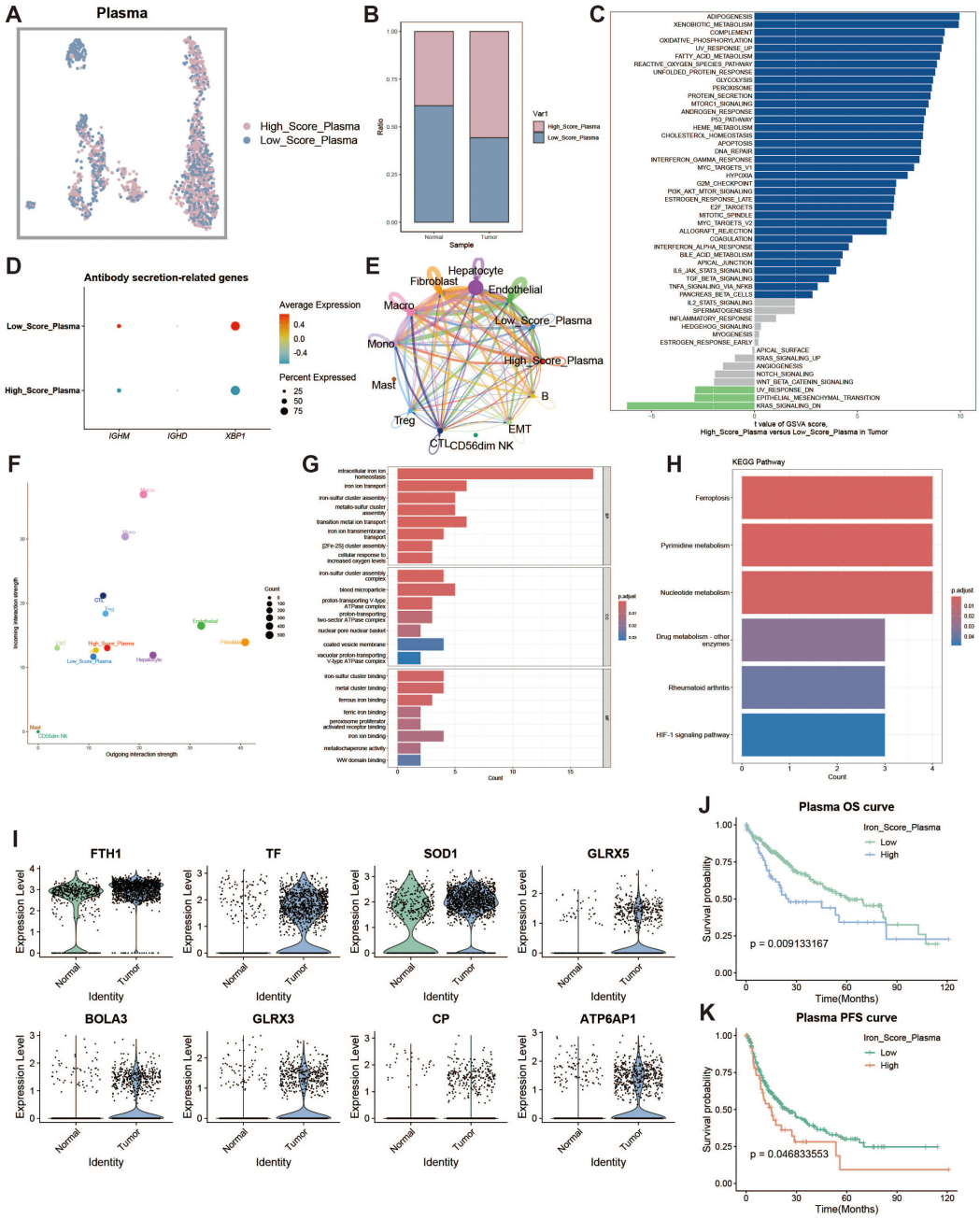


FIGURE 3
Plasma cell analysis. **(A)** UMAP of high and low iron metabolism score cells, where pink indicates high iron metabolism score cells and blue indicates low score cells. **(B)** Bar plot of cell proportions. **(C)** Diverging bar plot of GSVA enrichment results for Hallmark gene sets, showing pathways enriched in low-score and high-score cells. **(D)** Bubble plot of antibody secretion-related gene expression. **(E)** Circle plot showing cell communication frequency, where line thickness represents the number of communications. **(F)** Scatter plot of signal transmission and reception strength between cells. **(G)** Bar plot of GO enrichment analysis results. **(H)** Bar plot of KEGG enrichment analysis results. **(I)** Violin plot of differential expression of iron metabolism-related genes in plasma cells. **(J)** KM survival curve for high and low infiltration groups (overall survival). **(K)** KM survival curve for high and low infiltration groups (progression-free survival).

and low expression groups based on their iron metabolism scores (Figure 4A). Notably, CTLs in the tumor group exhibited significantly higher iron metabolism scores compared to those in the normal group (Figure 4B). GSVA analysis of the functional differences between these groups revealed a heightened oxidative profile in the high iron metabolism group (Figure 4C). Evaluation of cytotoxicity-related genes showed reduced cytotoxic function in CTLs with elevated iron metabolism (Figure 4D). Cell communication analysis further demonstrated enhanced intercellular communication and signal output in these cells (Figures 4E, F). GO and KEGG enrichment analyses indicated increased iron ion transport and oxidative responses in CTLs

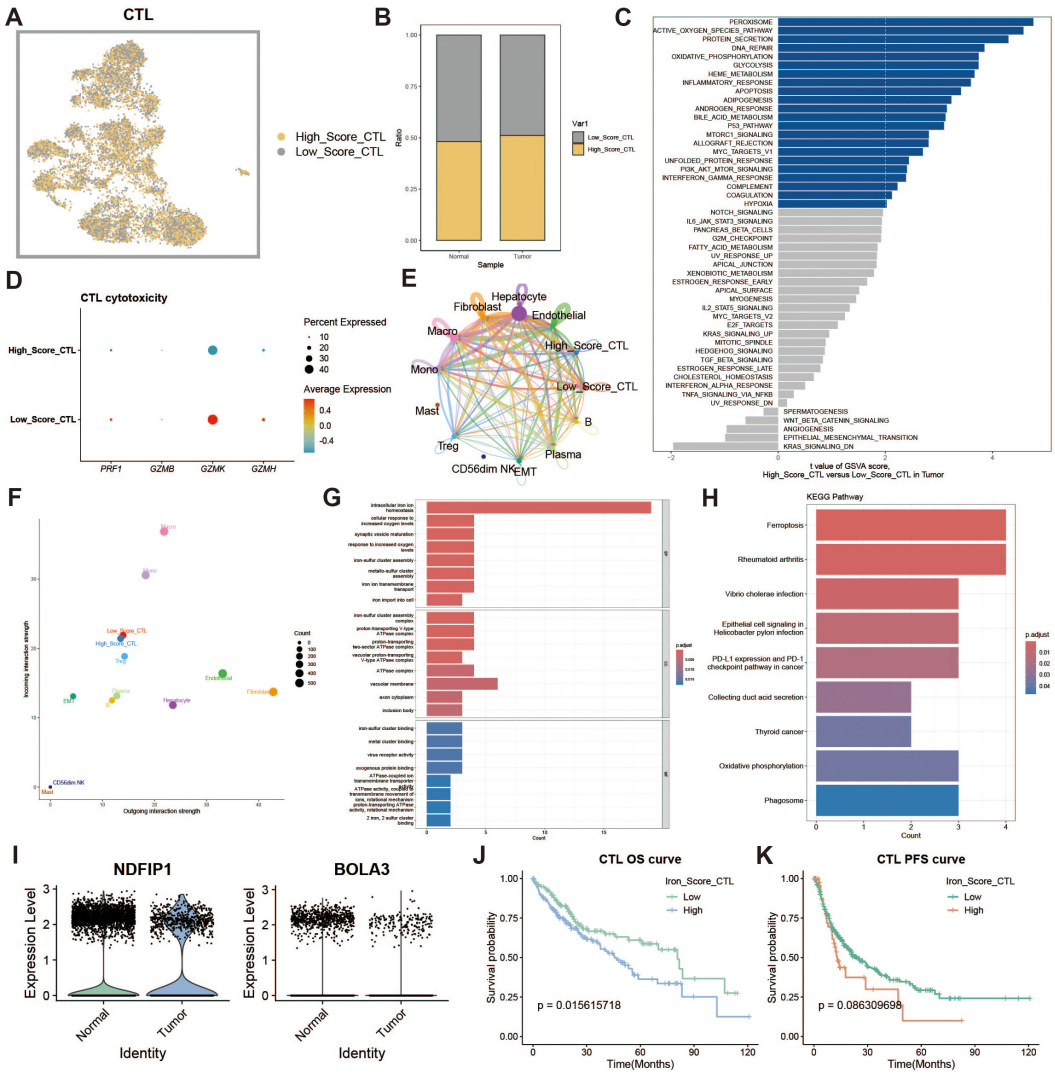


FIGURE 4
Cytotoxic T cell analysis. **(A)** UMAP of high and low iron metabolism score cells, where yellow indicates high-score cells and gray indicates low-score cells. **(B)** Bar plot of cell proportions. **(C)** Diverging bar plot of GSVA enrichment results for Hallmark gene sets. **(D)** Bubble plot of cytotoxic-related gene expression. **(E)** Circle plot of cell communication frequency. **(F)** Scatter plot of signal transmission and reception strength. **(G)** Bar plot of GO enrichment analysis results. **(H)** Bar plot of KEGG enrichment analysis results. **(I)** Violin plot of iron metabolism-related gene expression in CTLs. **(J)** KM survival curve for overall survival in high and low infiltration groups. **(K)** KM survival curve for progression-free survival in high and low infiltration groups.

with high iron metabolism (Figures 4G, H). Additionally, analysis of iron metabolism-related gene expression revealed a downregulation of NDFIP1 and BOLA3 in tumor tissues (Figure 4I). Survival curve analysis showed that patients with high iron metabolism had shorter overall survival (OS) and progression-free survival (PFS) compared to those with low iron metabolism (Figures 4J, K).

3.5 Effector memory T cell iron metabolism analysis

Effector memory T cells (TEMs) can rapidly produce effector cytokines to provide immune protection. We investigated the effects of iron metabolism on their immune function. TEMs were divided into high and low iron metabolism score groups (Figure 5A), and TEMs in the tumor group exhibited significantly higher iron metabolism scores than those in the normal group (Figure 5B). GSVA analysis revealed increased protein synthesis in high iron metabolism TEMs (Figure 5C). However, these cells also exhibited reduced cell proliferation and migration capacities (Figure 5D). Cell communication analysis showed enhanced intercellular communication and signal output in TEMs with high iron metabolism (Figures 5E, F). GO and KEGG analyses indicated elevated iron ion transport, protein localization, and cell differentiation in high iron metabolism TEMs (Figures 5G, H). Additionally, most iron metabolism-related genes were upregulated in tumor tissues (Figure 5I). Survival analysis demonstrated shorter overall survival (OS) and progression-free survival (PFS) in patients with high iron metabolism (Figures 5J, K).

10.3389/fimmu.2024.1496886

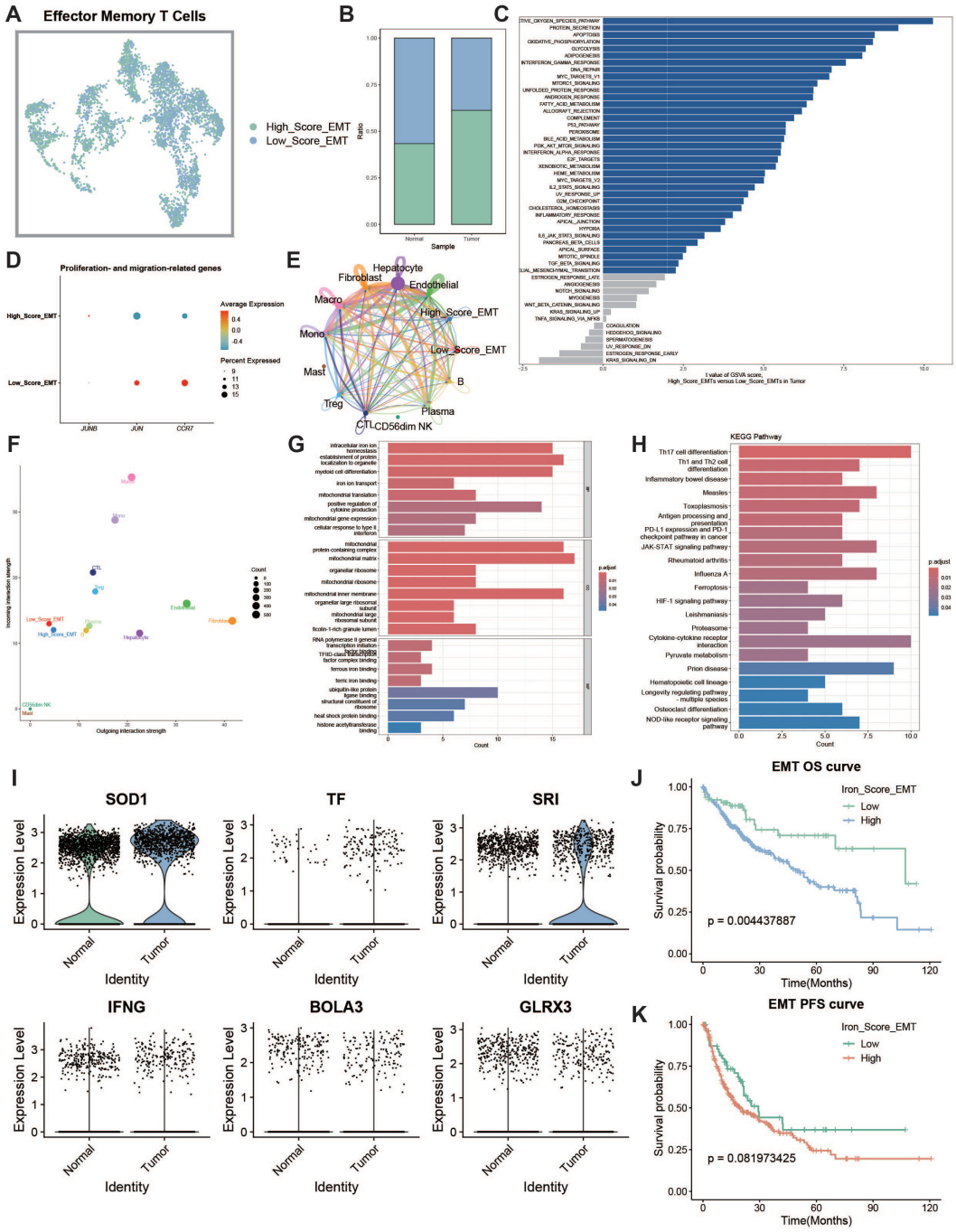


FIGURE 5
Effector memory T cell analysis. **(A)** UMAP of high and low iron metabolism score cells, with green indicating high-score cells and blue indicating low-score cells. **(B)** Bar plot of cell proportions. **(C)** Diverging bar plot of GSVA enrichment results for Hallmark gene sets. **(D)** Bubble plot of proliferation and migration-related gene expression. **(E)** Circle plot showing cell communication frequency. **(F)** Scatter plot of signal transmission and reception strength. **(G)** Bar plot of GO enrichment analysis results. **(H)** Bar plot of KEGG enrichment analysis results. **(I)** Violin plot of iron metabolism-related gene expression in EMTs. **(J)** KM survival curve for overall survival in high and low infiltration groups. **(K)** KM survival curve for progression-free survival in high and low infiltration groups.

3.6 Regulatory T cell iron metabolism analysis

Regulatory T cells (Tregs) are responsible for modulating immune responses and maintaining self-tolerance. We explored the impact of iron metabolism on their function. Tregs were divided into high and low expression groups based on their iron metabolism scores (Figure 6A), and Tregs in the tumor group exhibited higher iron metabolism scores compared to those in the normal group (Figure 6B). GSVA analysis revealed that oxidative phosphorylation was a dominant feature in high iron metabolism Tregs (Figure 6C). Immune suppression-related genes displayed distinct expression

10.3389/fimmu.2024.1496886

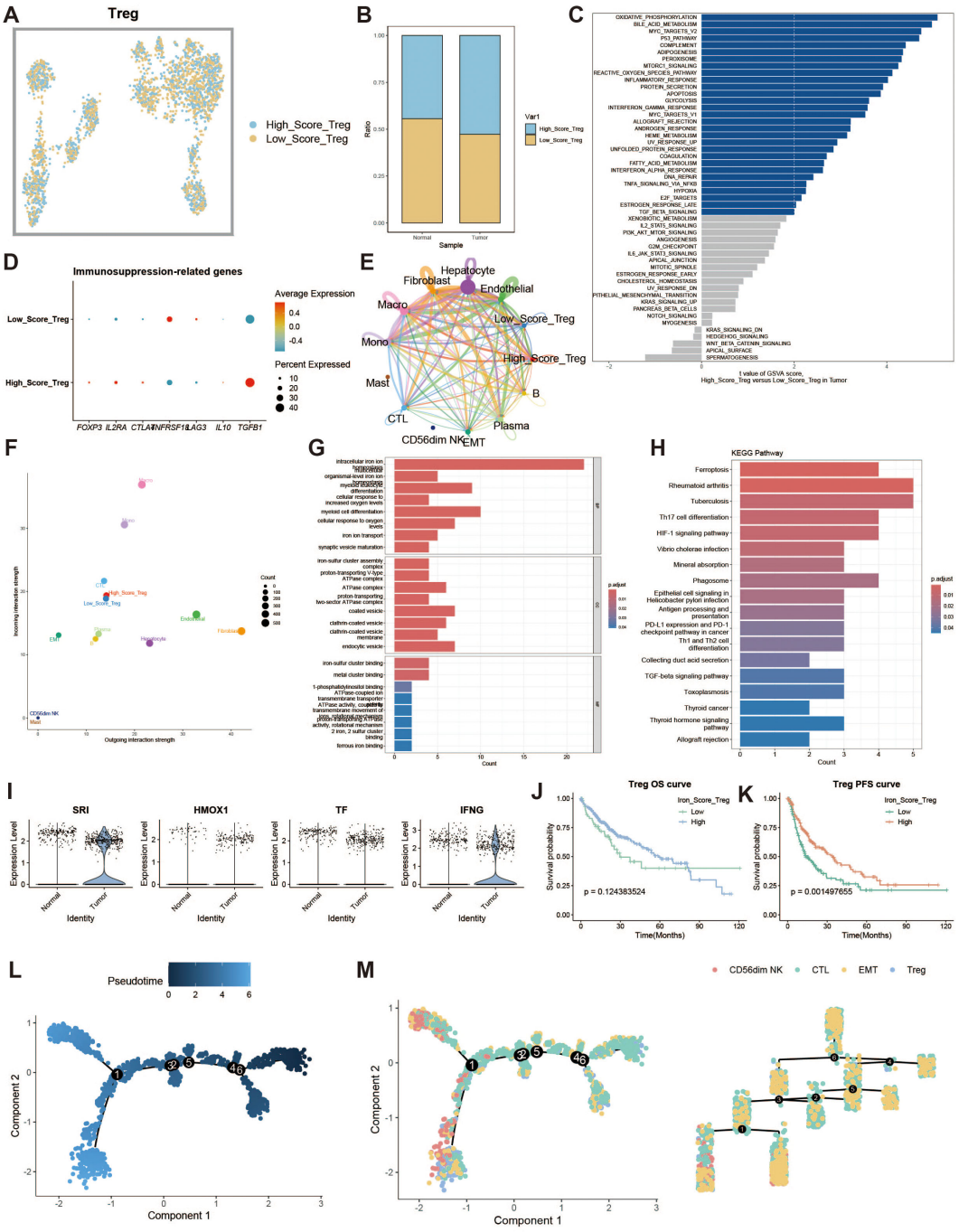


FIGURE 6
Regulatory T cell analysis. **(A)** UMAP of high and low iron metabolism score cells, with blue indicating high-score cells and yellow indicating low-score cells. **(B)** Bar plot of cell proportions. **(C)** Diverging bar plot of GSVA enrichment results for Hallmark gene sets. **(D)** Bubble plot of immunosuppressive-related gene expression. **(E)** Circle plot of cell communication frequency. **(F)** Scatter plot of signal transmission and reception strength. **(G)** Bar plot of GO enrichment analysis results. **(H)** Bar plot of KEGG enrichment analysis results. **(I)** Violin plot of iron metabolism-related gene expression in Tregs. **(J)** KM survival curve for overall survival in high and low infiltration groups. **(K)** KM survival curve for progression-free survival in high and low infiltration groups. **(L)** Pseudotime trajectory plot of T cell subpopulations, with color representing pseudotime. **(M)** Left panel shows the distribution of T cell subtypes on the trajectory plot, and the right panel shows the trajectory tree diagram.

patterns in the high iron metabolism group (Figure 6D). Cell communication analysis showed enhanced intercellular communication and signal output in Tregs with high iron metabolism scores (Figures 6E, F). GO and KEGG analyses revealed increased iron ion transport and oxidative responses in these cells (Figures 6G, H). Interestingly, high iron metabolism was associated with longer overall survival (OS) and progression-free survival (PFS) (Figures 6I–K). Pseudotime analysis revealed that CD56dim NK cells and Tregs appeared in the later stages of T cell development (Figures 6L, M).

10.3389/fimmu.2024.1496886

3.7 Endothelial cell iron metabolism analysis

To explore iron metabolism in endothelial cells within the liver cancer microenvironment, we divided endothelial cells into high and low expression groups based on iron metabolism scores (Figure 7A). Endothelial cells in the tumor group exhibited higher iron metabolism scores compared to those in the normal group (Figure 7B). Endothelial cells with high iron metabolism showed upregulation of pro-angiogenic genes, such as VEGFA (Figure 7C).

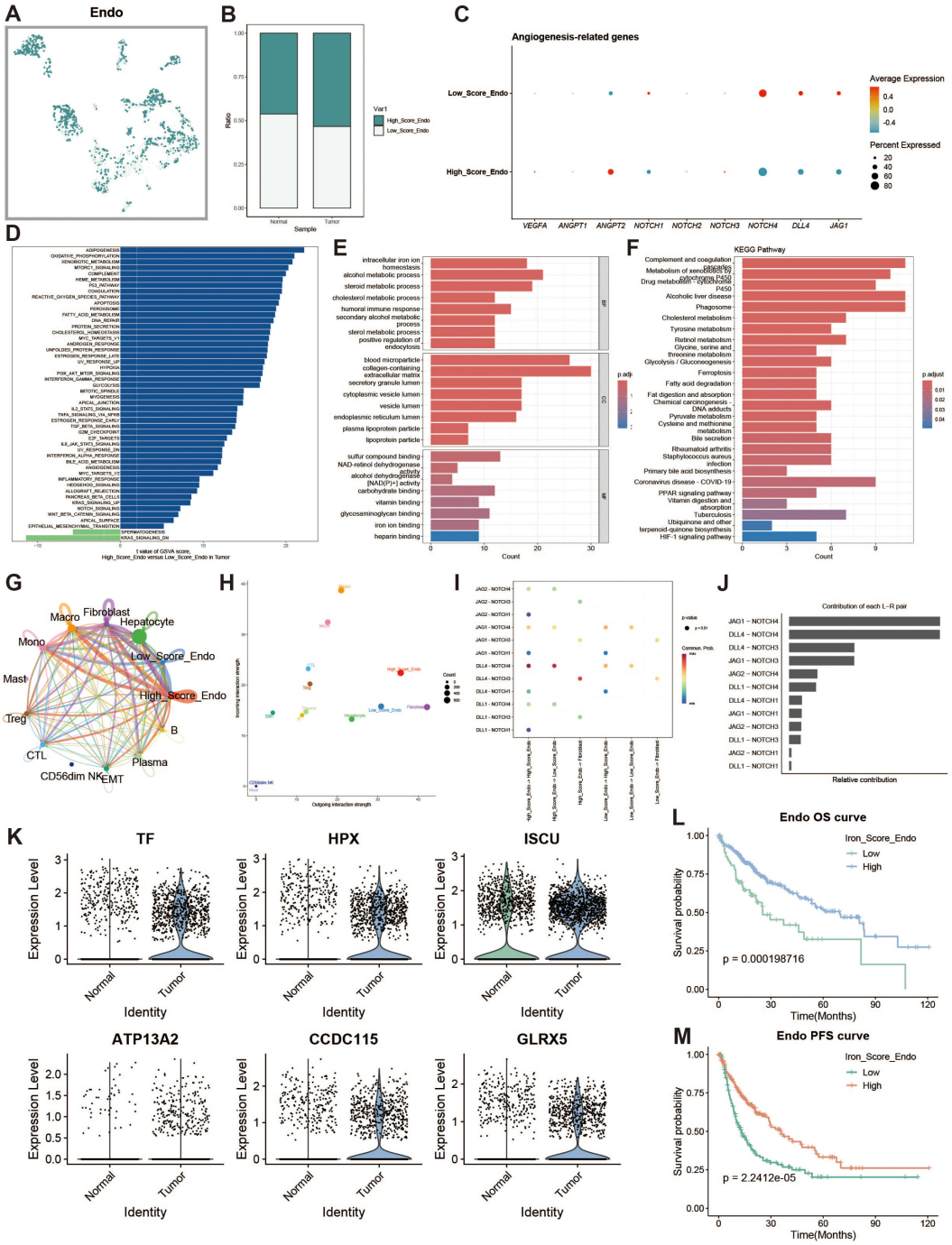


FIGURE 7
Endothelial cell analysis. **(A)** UMAP of high and low iron metabolism score cells, with green indicating high-score cells and white indicating low-score cells. **(B)** Bar plot of cell proportions. **(C)** Bubble plot of angiogenesis-related gene expression. **(D)** Diverging bar plot of GSVA enrichment results for Hallmark gene sets. **(E)** Bar plot of GO enrichment analysis results. **(F)** Bar plot of KEGG enrichment analysis results. **(G)** Circle plot of cell communication frequency. **(H)** Scatter plot of signal transmission and reception strength. **(I)** Bubble plot showing NOTCH signaling pathway communication between high and low iron metabolism endothelial cells. **(J)** Contribution of ligand-receptor pairs to communication in the NOTCH pathway. **(K)** Violin plot of iron metabolism-related gene expression in endothelial cells. **(J)** KM survival curve for overall survival in high and low infiltration groups. **(K)** KM survival curve for progression-free survival in high and low infiltration groups.

10.3389/fimmu.2024.1496886

GSVA analysis indicated increased oxidative phosphorylation and metabolic activity in high iron metabolism endothelial cells (Figure 7D). GO and KEGG enrichment analyses revealed elevated metabolic activity in these cells (Figures 7E, F). Cell communication analysis showed enhanced intercellular communication and signal output (Figures 7G, H). The JAG1-NOTCH1 pathway was identified as a key mediator of active communication (Figures 7I, J). Survival analysis revealed that patients with high iron metabolism had a longer survival period (Figures 7L, M).

3.8 Characterization of iron metabolism in fibroblasts

Fibroblasts are an important cellular component of the tumor microenvironment. Tumor-associated fibroblasts play a key role at all stages of tumor development, promoting tumor proliferation and migration, enhancing tumor angiogenesis, regulating tumor immunity, and increasing tumor drug resistance. We divided fibroblasts into high- and low-expression groups based on iron metabolism scores (Figures 8A, B). To explore the heterogeneity between the two groups, we analyzed the functional differences between fibroblasts with high and low iron metabolism scores through GSVA and found that the high iron metabolism score group exhibited higher oxidative phosphorylation, reactive oxygen species response, and fat generation-related characteristics (Figure 8C). We constructed a bubble chart to visualize the expression levels of tumor-associated fibroblast marker genes, aiming to explore the impact of different metabolic scores on the generation of tumor-associated fibroblasts (Figure 8D). The results indicated that the expression of tumor-associated fibroblast marker genes was elevated in fibroblasts with low iron metabolism scores.

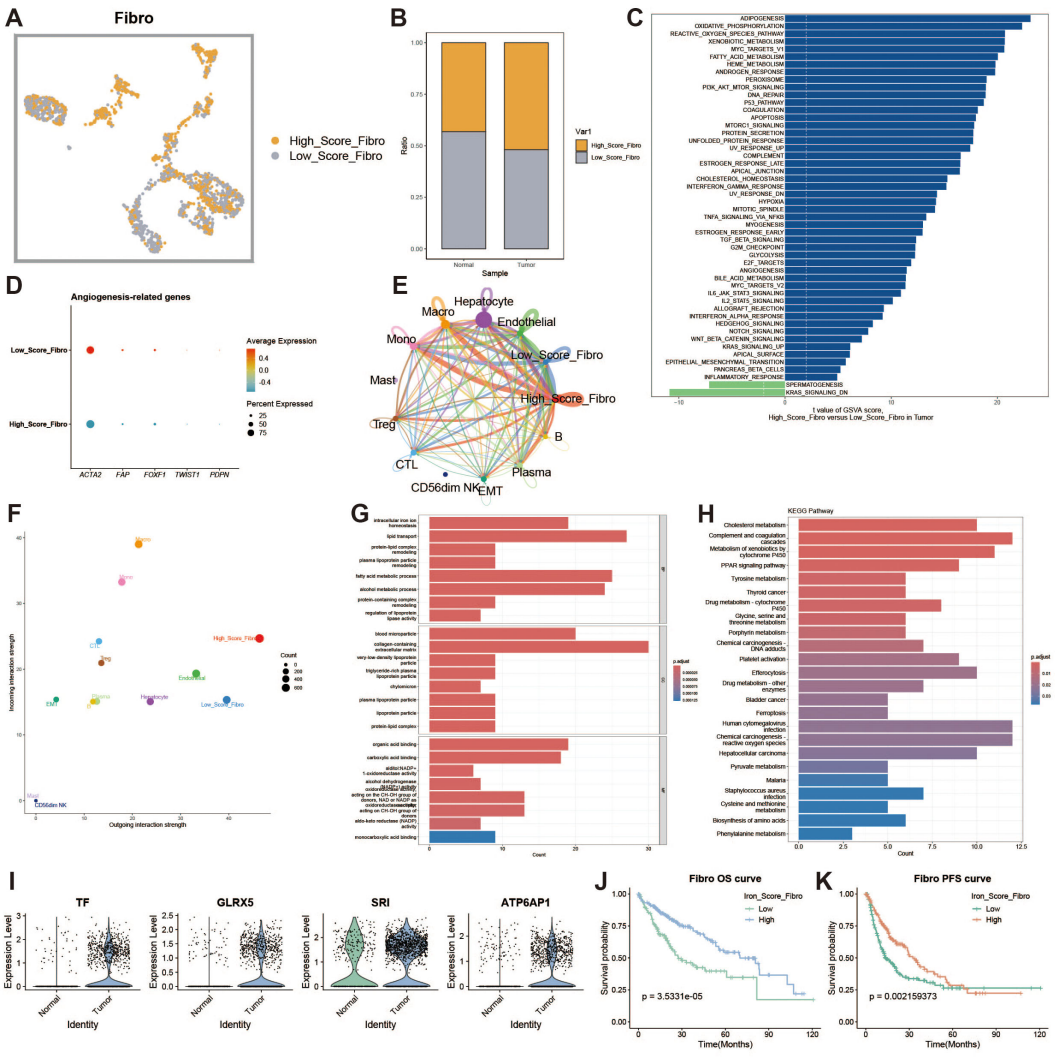


FIGURE 8
Fibroblast analysis. **(A)** UMAP of high and low iron metabolism score cells, with yellow indicating high-score cells and gray indicating low-score cells. **(B)** Bar plot of cell proportions. **(C)** Diverging bar plot of GSVA enrichment results for Hallmark gene sets. **(D)** Bubble plot of tumor-associated fibroblast marker gene expression. **(E)** Circle plot of cell communication frequency. **(F)** Scatter plot of signal transmission and reception strength. **(G)** Bar plot of GO enrichment analysis results. **(H)** Bar plot of KEGG enrichment analysis results. **(I)** Violin plot of iron metabolism-related gene expression in fibroblasts. **(J)** KM survival curve for overall survival in high and low infiltration groups. **(K)** KM survival curve for progression-free survival in high and low infiltration groups.

In a cell communication analysis, we discovered that fibroblasts with high iron metabolism scores exhibited enhanced levels of cell communication and stronger signal output characteristics (Figures 8E, F). Gene Ontology (GO) enrichment analysis revealed that fibroblasts with high iron metabolism scores demonstrated increased iron ion transport characteristics and a heightened oxidative response (Figure 8G).KEGG pathway analysis found that fibroblasts with high iron metabolism scores displayed more active metabolism (Figure 8H). We further analyzed the expression levels of iron metabolism genes and found that most of these genes were downregulated in liver cancer tissues (Figure 8I). Survival curves indicated that patients with high iron metabolism had a longer survival period compared to those with low iron metabolism (Figures 8J, K).

3.9 Spatial distribution of iron metabolism

To further investigate the features of iron metabolism in liver cancer, we conducted a deconvolution analysis of spatial transcriptomic data. We obtained spatial transcriptomic sequencing data from hepatocellular carcinoma (HCC) tumor tissue (GSM6177612), specifically from the tumor region of primary hepatocellular carcinoma. Following dimensionality reduction and clustering of the spatial transcriptomic data, we visualized the results using UMAP, which revealed seven distinct cell clusters (Figures 9A, B). The spatial distribution of these cell clusters is illustrated in Figure 9C. We assessed the iron metabolism-related gene scores for each cell cluster (Figure 9D) and analyzed the metabolic differences among the clusters, highlighting elevated metabolic activity in clusters 0, 1, and 2 (Figure 9E). Additionally, we examined the activity levels of glycolytic and oxidative phosphorylation metabolic pathways across different spatial regions (Figures 9F, G). High metabolic activity generally indicates that these cells play a more active role in tumor growth and progression, particularly in scenarios where energy demands are elevated. The spatial differences in the glycolytic and oxidative phosphorylation pathways suggest that cells in different regions may employ unique metabolic strategies to adapt to changes in the microenvironment. To further clarify the metabolic characteristics of cells in each spot and reveal the spatial distribution of their iron metabolism levels, we displayed the single-cell annotation results at the spatial level through deconvolution analysis (Figures 9H, I), showing the primary and secondary probabilities of cells with different iron metabolism levels in each spot (Figures 9J, K).

3.10 Survival analysis of iron metabolism-related genes

We analyzed the differential expression of key iron metabolism genes in tumor versus normal samples using TCGA and GEO data (Figure 10A), and performed ssGSEA scoring of iron metabolism levels in both groups (Figures 10B, C). The results revealed that iron metabolism levels were markedly elevated in the tumor group compared to the normal group. To further investigate the role of iron metabolism across different tumor stages, we examined and illustrated the differences in iron metabolism scores among patients at various clinical stages (Figures 10D, E). In the spatial transcriptomics data, we identified malignant, mixed, and normal cells, and depicted the iron metabolism score intensities for these three cell types (Figures 10F–H). The elevated iron metabolism scores in malignant cells compared to normal cells reflect the heterogeneity of iron metabolism within the tumor microenvironment. The correlation between AUC scores of key iron metabolism genes and microenvironment components further indicated that iron metabolism plays a significant role in regulating the tumor microenvironment, potentially influencing intercellular metabolic communication and tumor growth.

Through Spearman correlation analysis of AUC scores for key iron metabolism genes and microenvironment components, we further validated the pivotal role of iron metabolism in liver cancer (Figure 10I). We extracted overall survival (OS), disease-free survival (DFS), progression-free interval (PFI), and disease-free interval (DFI) data from liver cancer samples and examined the survival durations of patients with varying iron metabolism levels. The results indicated that higher iron metabolism levels were associated with poorer prognoses across these survival metrics (Figures 10J–M). This suggests that elevated iron metabolism levels may serve as a potential biomarker for increased tumor malignancy and adverse prognosis. Additionally, we utilized ICGC-LIRI and GSE76427 datasets to generate prognostic curves, thereby corroborating our findings (Figures 10N, O). Furthermore, conducting Gene Ontology (GO) enrichment analysis on patients with high and low iron metabolism scores revealed potential mechanisms through which iron metabolism influences prognosis (Figure 10P). Overall, these results suggest that iron metabolism represents a critical target for diagnosis and treatment in liver cancer.

3.11 GLRX3 expression and prognostic analysis

GLRX3, a key iron metabolism gene, was found to be highly expressed in HCC. Glutaredoxin 3 (GLRX3) is a type II monothiol glutaredoxin involved in iron balance, redox reactions, and antioxidant responses. In the TCGA cohort, GLRX3 expression was higher in advanced tumor grades (Figure 11A) and higher-stage tumors (Figure 11B). M1-stage tumors also showed increased GLRX3 expression compared to M0-stage tumors (Figure 11C). These differences indicate that high GLRX3 expression correlates with more advanced tumors and poorer prognosis (Figure 11D). We validated these findings by analyzing OS and DSS in patients with different GLRX3 expression levels (Figures 11E–H). Meta-analysis confirmed our conclusions (Figure 11I). GO enrichment analysis of high and low GLRX3 score patients revealed potential

mechanisms influencing prognosis (Figure 11J). Spatial transcriptomics data showed that GLRX3 was highly expressed in malignant regions (Figure 11K). Spearman correlation analysis further confirmed the role of GLRX3 in HCC (Figure 11L).

3.12 Knocking down the expression level of GLRX3 significantly inhibited the proliferation, invasion and migration of hepatocellular carcinoma cells

Knocking down the expression level of GLRX3 significantly inhibited the proliferation, invasion, and migration of hepatocellular carcinoma cells. Considering the importance of GLRX3, we validated its role in hepatocellular carcinoma through a series of *in vitro* experiments. First, we reduced the expression of GLRX3 and confirmed via PCR that its level was significantly decreased compared to the control group (Figures 12A, B). Subsequently, CCK8 assays demonstrated that the knockdown of GLRX3 markedly inhibited the activity of hepatocellular carcinoma cells (Figure 12C, D). To investigate the relationship between GLRX3 and the invasive migration of hepatocellular carcinoma, we conducted transwell and wound healing assays, revealing that GLRX3 knockdown significantly inhibited the invasive migration of these cells (Figures 12E, F). Immunohistochemistry experiments indicated that GLRX3 was highly expressed in hepatocellular carcinoma tissues (Figure 12G), and Western blot analysis confirmed the elevated protein expression of GLRX3 in these tissues (Figure 12H). In summary, GLRX3 enhances the invasive migration of hepatocellular carcinoma cells, correlating with the malignant characteristics of the disease.

4 Discussion

Iron serves dual roles in cancer biology: it acts as an initiator in the early stages of tumor development and functions as a promoter during malignancy, allowing transformed cells to maximize their potential for uncontrolled proliferation. Concurrently, cancer cells exhibit an increased demand for iron to support cellular growth, leading to alterations in iron metabolism-related gene expression that facilitate enhanced iron acquisition. Previous studies have demonstrated that tumor cells often upregulate transferrin receptor 1 (TFR1) while downregulating ferroportin (FNP), thereby limiting iron release (4, 19, 21). In our study, by scoring iron metabolism-related genes, we classified all cells into high and low iron metabolism groups and found that iron metabolism activity was consistently higher in tumor samples.

Disruption of the cellular iron homeostasis mechanism can lead to abnormal iron accumulation or depletion within cells. Under normal conditions, cells finely regulate iron levels to maintain a balance between demand and supply (39, 40). Heightened iron metabolism can disrupt homeostasis, resulting in abnormal iron levels that adversely impact cellular function and overall health. This dysregulation may alter the cellular redox balance, potentially inducing oxidative damage and ultimately resulting in iron-dependent programmed cell death, known as ferroptosis. These factors can significantly impact the prognosis of liver cancer patients (41). In our study, we generated multiple prognostic curves based on iron metabolism, showing that patients with higher iron metabolism scores experienced worse outcomes to varying degrees (42). Similarly, elevated iron metabolism was correlated with advanced tumor grades and stages. These findings underscore the potential of abnormal iron metabolism as a predictive biomarker and therapeutic target in cancer. They provide single-cell-level evidence to support the clinical investigation of iron chelators in cancer therapy (43). For example, oral iron chelators, such as deferasirox, have shown efficacy in leukemia patients, while the thiosemicarbazone Dp44mT has inhibited cancer cell proliferation *in vitro* by inducing the expression of p21, a cyclin-dependent kinase inhibitor involved in cell cycle arrest (44–46).

Our study also identified GLRX3 (Glutaredoxin 3) as a key iron metabolism target gene that significantly influences liver cancer progression (47). GLRX3 is a critical iron-sulfur cluster protein primarily involved in regulating iron metabolism. As a member of the oxidoreductase family, it performs multiple biological roles within cells, particularly in maintaining iron homeostasis and facilitating the assembly and transport of iron-sulfur clusters. In our study, we observed that GLRX3 was abnormally expressed in liver cancer patients (48). Iron-sulfur clusters serve as essential cofactors for many enzymes and proteins. Overexpression of GLRX3 can enhance the assembly and transport of these clusters in the cytoplasm, leading to excessive production and distribution (49). This overactivation may disrupt the metabolic balance in certain cells by over activating iron-sulfur cluster-dependent proteins. This finding is consistent with previous studies showing that tumor cells increase their demand for iron to sustain proliferation, with alterations in iron metabolism gene expression facilitating iron acquisition (50). Consequently, the expression level of GLRX3 may serve as a significant biomarker for liver cancer prognosis and as a potential indicator for assessing the efficacy of immunotherapy. Future research should prioritize elucidating the specific mechanisms through which GLRX3 contributes to tumor progression and developing targeted therapeutic strategies to enhance prognosis and treatment outcomes for liver cancer patients (51).

Although our study reveals the critical impact of iron metabolism on liver cancer, several limitations should be noted. First, the limited sample size may affect the generalizability of our findings. We hope future studies will analyze larger datasets, incorporating single-cell data from liver cancer patients across different databases, to fully explore the effects of iron metabolism dysregulation on the tumor microenvironment. Second, future studies should integrate proteomics and metabolomics approaches to provide multi-omics analyses that better elucidate the functional

10.3389/fimmu.2024.1496886

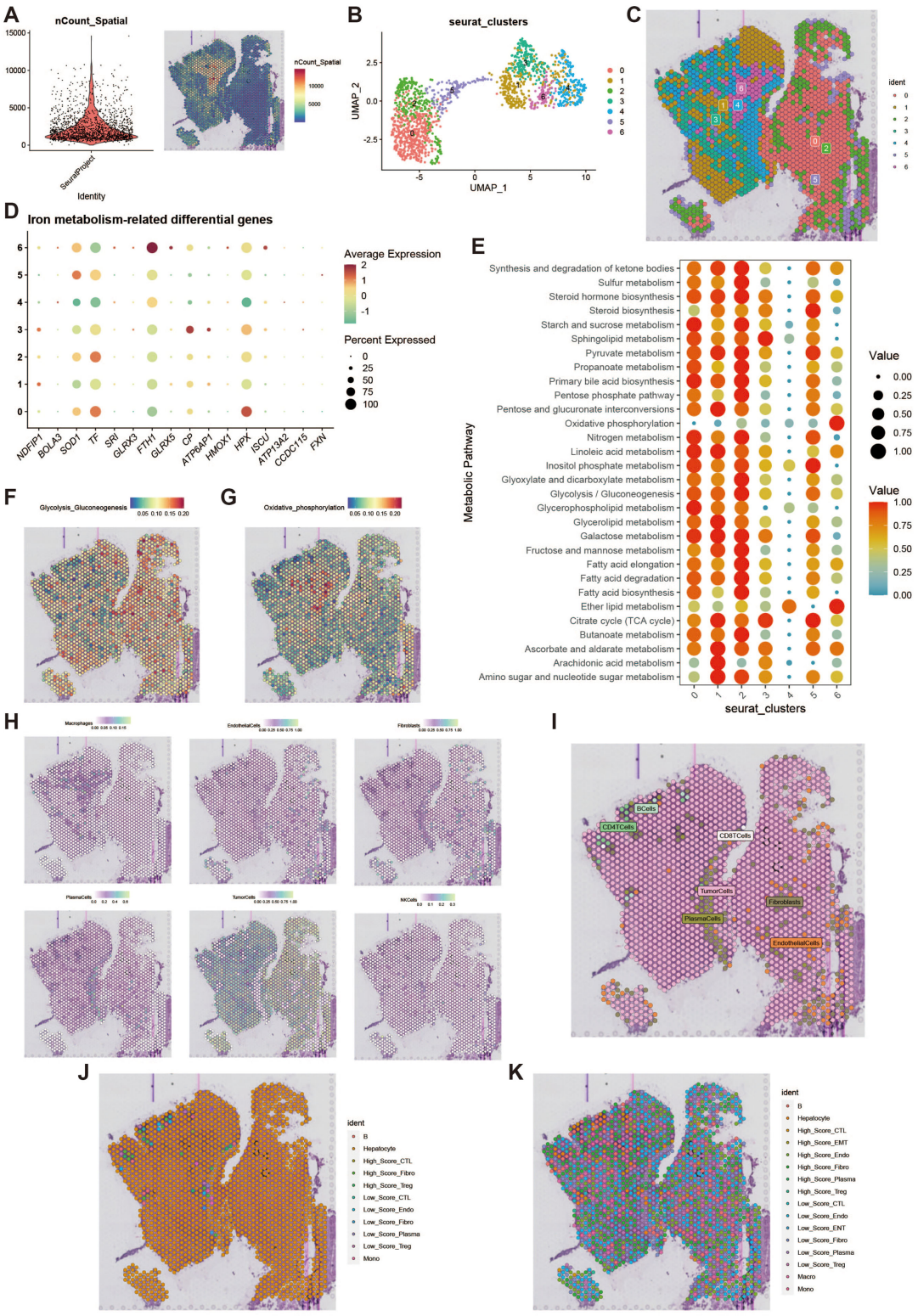


FIGURE 9
Spatial transcriptomics deconvolution analysis. **(A)** Heatmap of count values on spatial transcriptomic slices. **(B)** UMAP of dimensionality-reduced clustering results. **(C)** Plot of reduced dimensional clustering on spatial transcriptomic slices. **(D)** Bubble plot showing expression of key iron metabolism-related genes in spatial transcriptomics data. **(E)** Bubble plot of metabolic pathway activity scores. **(F)** Heatmap of glycolysis activity. **(G)** Heatmap of oxidative phosphorylation activity. **(H)** Deconvolution analysis results, including tumor cells, macrophages, fibroblasts, etc. **(I)** Plot of the most likely cell type for each spot. **(J)** Deconvolution analysis incorporating high and low iron metabolism levels, showing primary cell type results. **(K)** Deconvolution analysis incorporating high and low iron metabolism levels, showing secondary cell type results.

10.3389/fimmu.2024.1496886

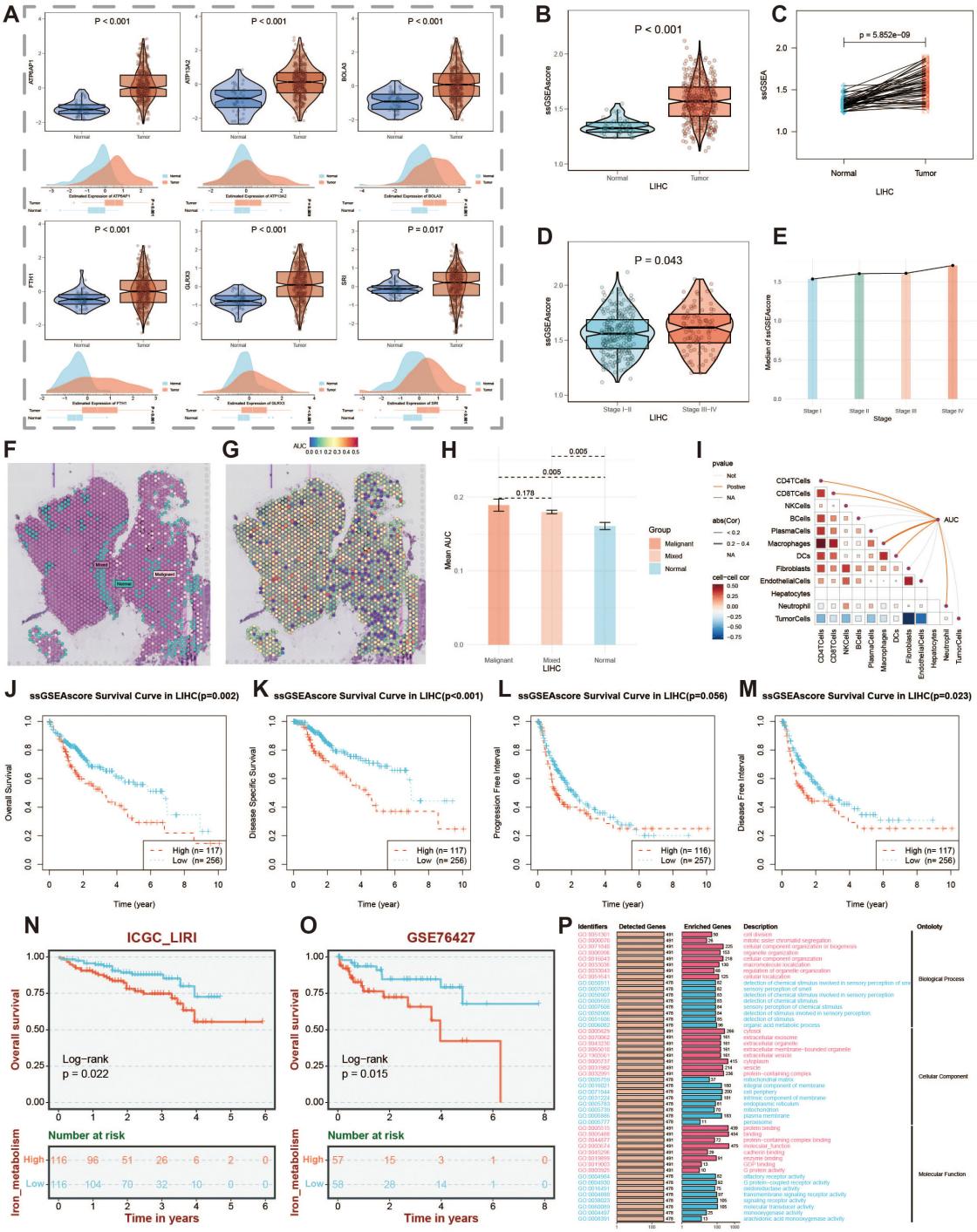


FIGURE 10
Expression of key iron metabolism-related genes and prognosis analysis. **(A)** Violin plot showing differential expression of key iron metabolism-related genes in TCGA and GEO data. **(B)** Violin plot of ssGSEA score results. **(C)** Differential expression of ssGSEA scores in paired samples. **(D)** Violin plot of score differences among patients at different clinical stages. **(E)** Line chart showing changes in scores across different clinical stages. **(F)** Identification of malignant, mixed, and normal cells in spatial transcriptomics data. **(G)** Active landscape of key iron metabolism-related genes in microregions. **(H)** Differences in the AUC scores of key iron metabolism-related genes between malignant, mixed malignant, and normal microregions at spatial transcriptomics resolution. **(I)** Spearman correlation between the AUC scores of key iron metabolism genes and microenvironment components. **(J-M)** Prognostic curves of patients with high and low iron metabolism scores. **(N)** Prognostic curve in ICGC-LIRI data. **(O)** Prognostic curve in GSE76427 data. **(P)** GO enrichment analysis of patients with high and low iron metabolism scores.

10.3389/fimmu.2024.1496886

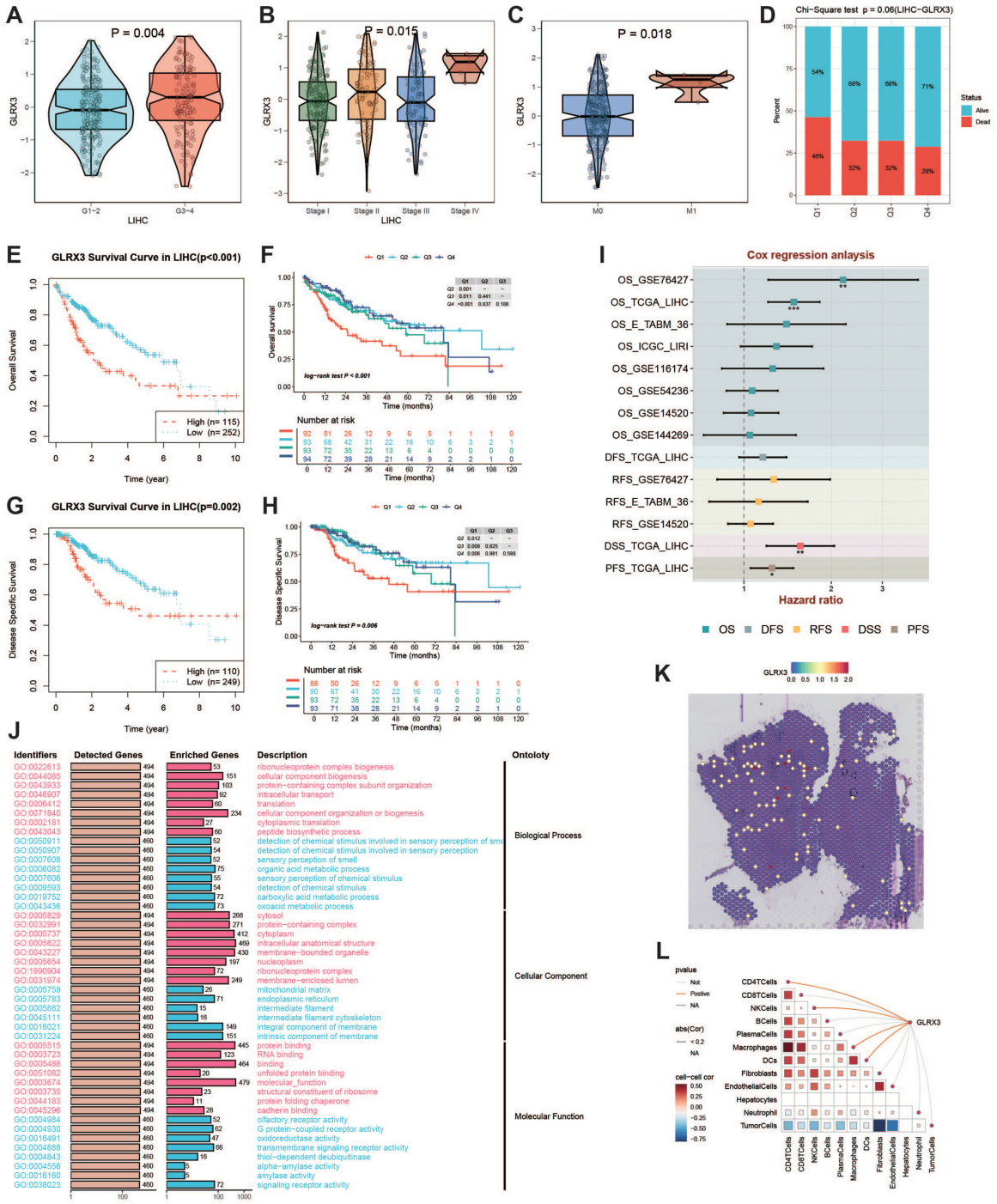


FIGURE 11
Clinical subgroup expression and K-M curve of GLRX3. **(A)** Differential expression of GLRX3 in high/low tumor grades in the TCGA cohort. **(B)** Differential expression of GLRX3 in high and low stages in the TCGA cohort. **(C)** Differential expression of GLRX3 in M1 and M0 stages in the TCGA cohort. **(D)** Bar chart of chi-square test showing the number of survival and death samples with different expression levels. The X-axis represents patients with different GLRX3 expression levels, and the Y-axis represents the proportion of deaths (red) and survivors (blue). **(E)** Kaplan-Meier survival analysis of OS. **(F)** Kaplan-Meier survival analysis of OS dividing patients into four groups (Q1, Q2, Q3, and Q4) based on GLRX3 expression levels. **(G)** Kaplan-Meier survival analysis of DSS. **(H)** Kaplan-Meier survival analysis of DSS dividing patients into four groups (Q1, Q2, Q3, and Q4) based on GLRX3 expression levels. **(I)** Meta-analysis of survival risk ratios. **(J)** GO enrichment analysis of high and low expression groups. **(K)** Each dot represents a microregion (spot) from spatial transcriptomics sequencing. The darker the color (red), the higher the expression level of the gene in the spot. **(L)** Correlation between cell content and GLRX3 expression levels in all spots, and correlation between cell content and GLRX3 gene expression.

10.3389/fimmu.2024.1496886

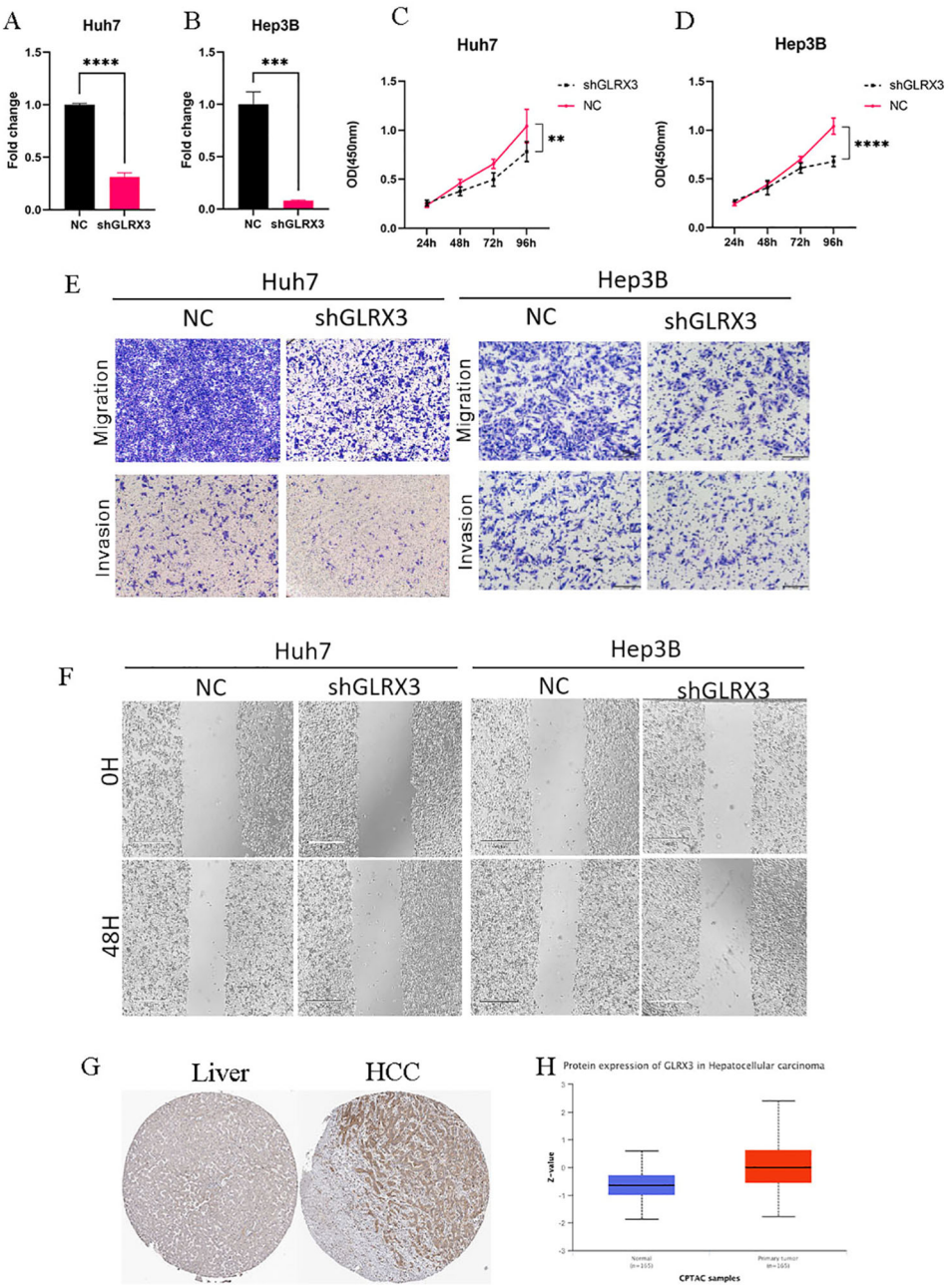


FIGURE 12
In vitro experiments to validate the role of GLRX3 in hepatocellular carcinoma. **(A, B)** PCR assay to detect GLRX3 knockdown efficiency. **(C, D)** CCK8 assay to detect cell viability. **(E)** transwell assay. **(F)** Wound healing assay. **(G)** IHC assay. **(H)** CTPAC database to verify GLRX3 expression. ** represents a p-value < 0.01, *** represents a p-value < 0.001, **** represents a p-value < 0.0001.

role of GLRX3. Third, due to technical and funding constraints, while we investigated the impact of GLRX3 on liver cancer prognosis, we did not conduct a comprehensive analysis of other key iron metabolism genes. Future research should broaden the scope of iron metabolism studies in liver cancer (4). Despite these limitations, our understanding of many processes remains incomplete, but growing recognition of the importance of iron metabolism in cancer biology offers new opportunities to uncover the mechanisms driving tumorigenesis. This, in turn, could lead to the development of more effective iron-targeted therapies for liver cancer.

5 Conclusion

This study revealed that iron metabolism plays a critical role in the progression of liver cancer, focusing on the role of GLRX3 (Glutaredoxin 3) in modulating iron homeostasis and driving

tumor progression. The study showed that disruptions in iron metabolism lead to abnormal iron accumulation or deficiency within liver cancer cells, inducing oxidative damage and ferroptosis. GLRX3, a key regulatory protein for iron-sulfur clusters, is abnormally overexpressed in liver cancer patients, and its overexpression facilitates iron-sulfur cluster assembly and transport, thereby disrupting metabolic balance and promoting tumor cell growth and metastasis. Survival analysis and experimental validation demonstrated that high GLRX3 expression correlates with poor patient prognosis, highlighting its potential as a prognostic biomarker and an indicator for assessing immune therapy response.

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/s.

Ethics statement

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

Author contributions

YL: Conceptualization, Data curation, Validation, Visualization, Writing – original draft, Writing – review & editing. YC: Writing – original draft. YaZ: Writing – original draft. YF: Writing – original draft. LW: Writing – original draft. YiZ: Writing – original draft. DW: Writing – original draft. XQ: Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The study was approved by Youth Foundation of Science and Technology Commission of Shanxi Province (No.202203021212066), General Programs of Shanxi Provincal Health Commission (No.2024024).

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.

Publisher's note

References

1. Ilyas SI, Wang J, El-Khoueiry AB. Liver cancer immunity. *Hepatology*. (2021) 73 Suppl 1:86–103.

2. Su L, Luo H, Yan Y, Yang Z, Lu J, Xu D, et al. Exploiting gender-based biomarkers and drug targets: advancing personalized therapeutic strategies in hepatocellular carcinoma. *Front Pharmacol*. (2024) 15:1433540.

3. Tang J, Zhang S, Jiang L, Liu J, Xu J, Jiang C, et al. Causal relationship between immune cells and hepatocellular carcinoma: a Mendelian randomisation study. *J Cancer*. (2024) 15:4219–31.

4. Molina-Sánchez P, Lujambio A. Iron overload and liver cancer. *J Exp Med*. (2019) 216:723–4.

5. Zhao C, Wu X, Chen J, Qian G. The therapeutic effect of IL-21 combined with IFN-γ inducing CD4(+)CXCR5(+)CD57(+)T cells differentiation on hepatocellular carcinoma. *J Adv Res*. (2022) 36:89–99.

6. Su K, Wang F, Li X, Chi H, Zhang J, He K, et al. Effect of external beam radiation therapy versus transcatheter arterial chemoembolization for non-diffuse hepatocellular carcinoma (≥ 5 cm): a multicenter experience over a ten-year period. *Front Immunol*. (2023) 14:1265959.

7. Li H, Wu Z, Chen J, Su K, Guo L, Xu K, et al. External radiotherapy combined with sorafenib has better efficacy in unresectable hepatocellular carcinoma: a systematic review and meta-analysis. *Clin Exp Med*. (2023) 23:1537–49.

8. Chen X, Chi H, Zhao X, Pan R, Wei Y, Han Y. Role of exosomes in immune microenvironment of hepatocellular carcinoma. *J Oncol*. (2022) 2022:2521025.

9. Chi H, Zhao S, Yang J, Gao X, Peng G, Zhang J, et al. T-cell exhaustion signatures characterize the immune landscape and predict HCC prognosis via integrating single-cell RNA-seq and bulk RNA-sequencing. *Front Immunol*. (2023) 14:1137025.

10. Fu Y, Maccioni L, Wang XW, Greten TF, Gao B. Alcohol-associated liver cancer. *Hepatology*. (2024).

11. Zhang S, Jiang C, Jiang L, Chen H, Huang J, Gao X, et al. Construction of a diagnostic model for hepatitis B-related hepatocellular carcinoma using machine learning and artificial neural networks and revealing the correlation by immunoassay. *Tumour Virus Res*. (2023) 16:200271.

12. Zhang S, Jiang C, Jiang L, Chen H, Huang J, Zhang J, et al. Uncovering the immune microenvironment and molecular subtypes of hepatitis B-related liver cirrhosis and developing stable a diagnostic differential model by machine learning and artificial neural networks. *Front Mol Biosci*. (2023) 10:1275897.

13. Su K, Liu Y, Wang P, He K, Wang F, Chi H, et al. Heat-shock protein 90α is a potential prognostic and predictive biomarker in hepatocellular carcinoma: a large-scale and multicenter study. *Hepatol Int*. (2022) 16:1208–19.

14. Su K, Guo L, Ma W, Wang J, Xie Y, Rao M, et al. PD-1 inhibitors plus anti-angiogenic therapy with or without intensity-modulated radiotherapy for advanced hepatocellular carcinoma: A propensity score matching study. *Front Immunol*. (2022) 13:972503.

15. Paradis V, Zucman-Rossi J. Pathogenesis of primary liver carcinomas. *J Hepatol*. (2023) 78:448–9.

16. Su K, Shen Q, Tong J, Gu T, Xu K, Li H, et al. Construction and validation of a nomogram for HBV-related hepatocellular carcinoma: A large, multicenter study. *Ann Hepatol*. (2023) 28:101109.

17. Li H, Guo L, Su K, Li C, Jiang Y, Wang P, et al. Construction and validation of TACE therapeutic efficacy by ALR score and nomogram: A large, multicenter study. *J Hepatocell Carcinoma*. (2023) 10:1009–17.

18. Ladd AD, Duarte S, Sahin I, Zarrinpar A. Mechanisms of drug resistance in HCC. *Hepatology*. (2024) 79:926–40.

19. Martins AC, Almeida JI, Lima IS, Kapitão AS, Gozzelino R. Iron metabolism and the inflammatory response. *IUBMB Life*. (2017) 69:442–50.

20. Capelletti MM, Manceau H, Puy H, Peoc'h K. Ferroptosis in liver diseases: an overview. *Int J Mol Sci*. (2020) 21.

21. Recalcati S, Correnti M, Gammella E, Raggi C, Invernizzi P, Cairo G. Iron metabolism in liver cancer stem cells. *Front Oncol*. (2019) 9:149.

22. Zeng K, Huang N, Liu N, Deng X, Mu Y, Zhang X, et al. LACTB suppresses liver cancer progression through regulation of ferroptosis. *Redox Biol*. (2024) 75:103270.

23. He S, Su L, Hu H, Liu H, Xiong J, Gong X, et al. Immunoregulatory functions and therapeutic potential of natural killer cell-derived extracellular vesicles in chronic diseases. *Front Immunol*. (2023) 14:1328094.

24. Satriano L, Lewinska M, Rodrigues PM, Banales JM, Andersen JB. Metabolic rearrangements in primary liver cancers: cause and consequences. *Nat Rev Gastroenterol Hepatol*. (2019) 16:748–66.

25. Zhou H, Chen J, Fan M, Cai H, Dong Y, Qiu Y, et al. KLF14 regulates the growth of hepatocellular carcinoma cells via its modulation of iron homeostasis through the repression of iron-responsive element-binding protein 2. *J Exp Clin Cancer Res*. (2023) 42:5.

26. Mossmann D, Müller C, Park S, Ryback B, Colombi M, Ritter N, et al. Arginine reprograms metabolism in liver cancer via RBM39. *Cell*. (2023) 186:5068–5083.e23.

27. Jiang L, Ren X, Yang J, Chen H, Zhang S, Zhou X, et al. Mitophagy and clear cell renal cell carcinoma: insights from single-cell and spatial transcriptomics analysis. *Front Immunol*. (2024) 15:1400431.

28. Jiang L, Liu J, Zhang S, Jiang C, Huang J, Chen H, et al. Role of glycosylation-related gene MGAT1 in pancreatic ductal adenocarcinoma. *Front Immunol*. (2024) 15:1438935.

29. Xu Q, Liu C, Wang H, Li S, Yan H, Liu Z, et al. Deciphering the impact of aggregated autophagy-related genes TUBA1B and HSP90AA1 on colorectal cancer evolution: a single-cell sequencing study of the tumor microenvironment. *Discovery Oncol*. (2024) 15:431.

30. Chen H, Zuo H, Huang J, Liu J, Jiang L, Jiang C, et al. Unravelling infiltrating T-cell heterogeneity in kidney renal clear cell carcinoma: Integrative single-cell and spatial transcriptomic profiling. *J Cell Mol Med*. (2024) 28:e18403.

31. Wang L, Zhou X, Yan H, Miao Y, Wang B, Gu Y, et al. Deciphering the role of tryptophan metabolism-associated genes ECHS1 and ALDH2 in gastric cancer: implications for tumor immunity and personalized therapy. *Front Immunol*. (2024) 15:1460308.

32. Gao R, Bai S, Henderson YC, Lin Y, Schalck A, Yan Y, et al. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. *Nat Biotechnol*. (2021) 39:599–608.

33. Andreatta M, Carmona SJ. UCell: Robust and scalable single-cell gene signature scoring. *Comput Struct Biotechnol J*. (2021) 19:3796–8.

34. Jiang C, Zhang S, Jiang L, Chen Z, Chen H, Huang J, et al. Precision unveiled: Synergistic genomic landscapes in breast cancer-Integrating single-cell analysis and decoding drug toxicity for elite prognostication and tailored therapeutics. *Environ Toxicol*. (2024) 39:3448–72.

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.

36. Huang J, Liu M, Chen H, Zhang J, Xie X, Jiang L, et al. Elucidating the Influence of MPT-driven necrosis-linked LncRNAs on immunotherapy outcomes, sensitivity to chemotherapy, and mechanisms of cell death in clear cell renal carcinoma. *Front Oncol*. (2023) 13:1276715.

37. Chi H, Huang J, Yan Y, Jiang C, Zhang S, Chen H, et al. Unraveling the role of disulfidptosis-related LncRNAs in colon cancer: a prognostic indicator for immunotherapy response, chemotherapy sensitivity, and insights into cell death mechanisms. *Front Mol Biosci*. (2023) 10:1254232.

38. Tu H, Hu Q, Ma Y, Huang J, Luo H, Jiang L, et al. Deciphering the tumour microenvironment of clear cell renal cell carcinoma: Prognostic insights from programmed death genes using machine learning. *J Cell Mol Med*. (2024) 28:e18524.

39. Sun Y, Li Q, Guo H, He Q. Ferroptosis and iron metabolism after intracerebral hemorrhage. *Cells*. (2022) 12.

40. Tang J, Wei W, Xu Y, Chen K, Miao Y, Fan W, et al. CXC chemokine receptor 4 - mediated immune modulation and tumor microenvironment heterogeneity in gastric cancer: Utilizing multi-omics approaches to identify potential therapeutic targets. *Biofactors*. (2024).

41. Rais Y, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, et al. Deterministic direct reprogramming of somatic cells to pluripotency. *Nature*. (2013) 502:65–70.

42. Kulaszyńska M, Kwiatkowski S, Skonieczna-Żydecka K. The iron metabolism with a specific focus on the functioning of the nervous system. *Biomedicines*. (2024) 12.

43. Kazan HH, Urfali-Mamatoglu C, Gunduz U. Iron metabolism and drug resistance in cancer. *Biometals*. (2017) 30:629–41.

44. Kwiatkowski J, Duffner U, Abdel-Mageed A. Deferasirox decreases busulfan clearance. *Ann Pharmacother*. (2018) 52:497–8.

45. Enko D, Moro T, Holasek S, Baranyi A, Schnedl WJ, Zelzer S, et al. Branched-chain amino acids are linked with iron metabolism. *Ann Transl Med*. (2020) 8:1569.

46. Chaffer CL, Brueckmann I, Scheel C, Kaestli AJ, Wiggins PA, Rodrigues LO, et al. Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proc Natl Acad Sci U.S.A*. (2011) 108:7950–5.

47. Pandya P, Braiman A, Isakov N. PICOT (GLRX3) is a positive regulator of stress-induced DNA-damage response. *Cell Signal*. (2019) 62:109340.

48. Mechoud MA, Pujol-Carrion N, Montella-Manuel S, de la Torre-Ruiz MA. Interactions of GMP with human glrx3 and with saccharomyces cerevisiae grx3 and grx4 converge in the regulation of the gcn2 pathway. *Appl Environ Microbiol*. (2020) 86.

49. Li B, Chen M, Lu M, Xin-Xiang J, Meng-Xiong P, Jun-Wu M. Glutaredoxin 3 promotes migration and invasion via the Notch signalling pathway in oral squamous cell carcinoma. *Free Radic Res*. (2018) 52:390–401.

50. Camponeschi F, Prusty NR, Heider SAE, Ciofi-Baffoni S, Banci L. GLRX3 acts as a [2Fe-2S] cluster chaperone in the cytosolic iron-sulfur assembly machinery transferring [2Fe-2S] clusters to NUBP1. *J Am Chem Soc*. (2020) 142:10794–805.

51. Jo JH, Kim SA, Lee JH, Park YR, Kim C, Park SB, et al. GLRX3, a novel cancer stem cell-related secretory biomarker of pancreatic ductal adenocarcinoma. *BMC Cancer*. (2021) 21:1241.

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 29 November 2024
DOI 10.3389/fimmu.2024.1493673

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong, Hong Kong SAR, China

REVIEWED BY
Chen Li,
Free University of Berlin, Germany
Zhijia Xia,
Ludwig Maximilian University of Munich, Germany

*CORRESPONDENCE
Xiang Zhang
cq_zhangxiang@163.com
Daqiang Song
cqmusdq@163.com

†These authors have contributed equally to this work

RECEIVED 09 September 2024
ACCEPTED 13 November 2024
PUBLISHED 29 November 2024

CITATION
Li X, Pan Z, Luan T, Xiao Q, Li L, Wu Q, Yao G, Zhang X and Song D (2024) Fibroblast growth factor receptor risk signature predicts patient prognosis and immunotherapy resistance in colorectal cancer.
Front. Immunol. 15:1493673.
doi: 10.3389/fimmu.2024.1493673

Fibroblast growth factor receptor risk signature predicts patient prognosis and immunotherapy resistance in colorectal cancer

Xiaofang Li[1†], Zhiling Pan[2,3†], Tiankuo Luan[4†], Qian Xiao[4], Liuying Li[4], Qianxue Wu[4], Guoqing Yao[5], Xiang Zhang[6*] and Daqiang Song[4*]

[1]Department of Pharmacy, The Affiliated Hospital of Southwest Medical University, Luzhou, Sichuan, China, [2]Department of Operating Room, Affiliated Hospital of Youjiang Medical University for Nationalities, Baise, China, [3]Baise Key Laboratory of Molecular Pathology in Tumors, Baise, China, [4]Chongqing Key Laboratory of Molecular Oncology and Epigenetics, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China, [5]Department of Vascular Surgery, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China, [6]Department of Breast and Thyroid Surgery, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China

Background: Fibroblast Growth Factor Receptor (FGFR) signaling is linked with tumor progression and tumor immunoevasion, yet the potential effect of FGFR signature on the prognosis of patient with colorectal cancer (CRC) and response to immune therapy remains elusive.

Methods: The fibroblast growth factor receptor risk signature (FRS) was identified through single-cell RNA sequencing, bulk RNA sequencing, and machine learning techniques. Signaling enrichment analyses were conducted using Gene Set Enrichment Analysis (GSEA) and the Kyoto Encyclopedia of Genes and Genomes (KEGG). Drugs targeting the FRS were predicted using the Cancer Therapeutics Response Portal (CTRP) and PRISM databases. The analysis of T cell function and the tumor microenvironment (TME) was performed using flow cytometry.

Results: In this study, we characterized the FRS in cancer patients with CRC. By integrating advanced techniques, we identified the FRS and revealed the intricate molecular landscape and diversity of the FRS within the TME. Notably, the FRS effectively predicted unfavorable prognosis and resistance to immunotherapy in CRC patients. Furthermore, PHA-793887, identified as a potential FRS inhibitor by the CTRP and PRISM databases, significantly restructured the immunosuppressive TME and enhanced the antitumor immune response, resulting in a reduced tumor burden in the MC38 murine tumor model.

Conclusion: Together, these data support FRS positively correlates with poor prognosis and therapy resistance. The PHA-793887 could be a potential FRS inhibitor to improving the effectiveness of CRC management via bolstering antitumor immunity.

KEYWORDS

fibroblast, tumor immunity, prognosis, therapy resistance, colorectal cancer

Introduction

Colorectal cancer (CRC) stands as a significant contributor to cancer-related deaths globally, featuring an intricate interaction between genetic and environmental elements that propel its onset and advancement (1, 2). The disease's diversity and inconsistent reaction to therapies highlight the necessity for a more profound comprehension of its molecular foundations. Despite advancements in early detection and treatment, the precise mechanisms driving CRC progression and the genes influencing patient prognosis remain insufficiently understood. Elucidating these complexities is essential for developing new prognostic indicators and treatment strategies.

Fibroblast growth factor receptor (FGFR) signaling plays a critical role in multiple aspects of tumor biology, including cell proliferation, survival, angiogenesis, and invasion. The activation of FGFR signaling is observed in various cancers, contributing to aggressive disease phenotypes and poor outcomes (3–5). Nevertheless, the correlation between FGFR signaling and the tumor microenvironment (TME) in CRC remains inadequately defined. With an increasing understanding of the TME's influence on cancer progression and response to immunotherapy, clarifying the association between FGFR signaling and the immunological landscape of CRC is highly relevant (6, 7). Examining FGFR risk signature (FRS) holds promise for shedding light on CRC pathogenesis and guiding prognostic assessments of patients.

The emergence of cancer immunotherapy has revolutionized the management of various malignancies, including CRC. Despite its potential, the clinical effectiveness of immunotherapy is frequently impeded by the development of treatment resistance, benefiting only a subset of patients (8, 9). The mechanisms driving this resistance are complex, involving aberrant activation of specific intracellular transcriptional pathways in tumor cells, among other factors (10, 11). In the context of CRC, our investigation has revealed a significant association between the FRS and immunotherapy resistance. Given the established role of FRS in influencing cellular behavior and the TME, targeting this signature presents a promising strategy to overcome this resistance.

In this study, we utilized bioinformatic tools and experimental models to identify and forecast the effectiveness of inhibitors targeting the FRS. Our goal was to devise a therapeutic strategy capable of selectively disrupting FRS, thereby enhancing the anti-tumor immune response and impeding tumor progression. Our results indicate that FRS-targeting inhibitors have the potential to reshape the TME, enhance T cell function, and elicit antitumor responses. This discovery carries substantial translational significance, suggesting that the integration of FRS inhibitors with existing immunotherapies may provide a novel approach to improving treatment efficacy and clinical outcomes for CRC patients.

Materials and methods

Animal experiments

In the present investigation, male C57BL/6J mice, aged 6-8 weeks, were obtained from Ensiwer Corporation and utilized in the ICB-resistant MC38 tumor model. The colorectal cancer MC38 cell line, adjusted to a concentration of 1 × 10^5 cells, was subcutaneously inoculated into these mice. On the seventh day post-inoculation, the animals were administered either IgG or anti-CTLA4 antibodies (Bio X cell, 9H10) at a dosage of 5 mg/kg for a continuous period of five days. At the culmination of the designated experimental timelines, the tumor tissues were meticulously excised for further analysis. In a subsequent experimental paradigm, mice bearing MC38 tumors were subjected to treatment with either a control vehicle or PHA-793887 (MedChemExpress, HY-11001), administered at a dosage of 10 mg/kg. Post-treatment, the excised tumor tissues underwent flow cytometric analysis to elucidate the impact of PHA-793887 on the TME. Concurrently, tumor dimensions were meticulously monitored biweekly throughout the experimental period.

Flow cytometry analysis

Live cells were assessed using the Fixable Viability Dye eFluor 450. To evaluate cytokine production, the cells were stimulated with the Cell Stimulation Cocktail and subsequently labeled with anti-IFN-γ (BioLegend, XMG1.2) and anti-TNF-α (BioLegend, MP6-XT22) antibodies. Other antibodies included an-CD45 (Biolegend, 2D1), anti-CD11b (BioLegend, M1/70), anti-CD8a (BioLegend, 53-6.7), anti-CD4 (BioLegend, GK1.5), anti-FOXP3 (BioLegend, MF-14), anti-F4/80 (BioLegend, BM8).

Analysis of stained cells was conducted using a BD FACSCanto II Flow Cytometer in conjunction with BD FACSDiva software (BD Biosciences), and the resulting data were processed using FlowJo software (version 10.5.3).

RNA sequencing analysis

Total RNA was isolated from MC38 tumor tissues using the Trizol reagent (Invitrogen, catalog number 15596026). RNA samples were forwarded to ANNOROAD for construction of sequencing libraries and subsequent sequencing on the NovaSeq platform (Illumina). The resultant raw fastq files were processed to quantify gene expression as transcripts per million (TPM) using the htseq-count tool, facilitating the downstream analysis. Differential expression analysis of genes (DEGs) was performed using the "DESeq2" R package, applying stringent filtering criteria: a fold change threshold greater than 2, an adjusted *P*-value of less than 0.05, and a mean log2-TPM in the high-expression cohort exceeding 0.

Data acquisition

In the present investigation, we procured five distinct public datasets from the NCBI Gene Expression Omnibus (GEO) repository. Our approach entailed the application of scRNA-seq datasets, namely GSE231559 and GSE166555, to dissect the heterogeneity of fibroblast populations within both normal and neoplastic colorectal tissues. Additionally, we leveraged the COAD

cohort from the TCGA database and an aggregate of three bulk RNA sequencing datasets, GSE17536, GSE29621, and GSE38832, for the development and substantiation of our prognostic model. It is important to note that for analyses based on publicly available datasets, neither patient consent nor ethical board approval is required.

scRNA-seq data analysis

The processing of scRNA-seq data was executed utilizing the 'Seurat' R package (version 5.0.2), following established protocols (12). Initially, a stringent filtration criterion was applied, excluding cells with gene expression levels below the 200-gene threshold or exceeding the 6000-gene ceiling, as well as those with mitochondrial gene expression surpassing the 5% mark. This exclusionary process was pivotal for the retention of a substantial cell population representative of the datasets in question. Subsequently, the SCTransform function was utilized to standardize and normalize the raw count data, which was then subjected to principal component analysis (PCA) to identify the underlying patterns. To address and neutralize batch effects inherent in the dissociated scRNA-seq datasets, the "Harmony" R package, was strategically implemented. Clustering was conducted by assessing the edge weights connecting pairs of cells, culminating in the construction of a shared nearest-neighbor graph. This graph was adeptly derived using the Louvain algorithm, facilitated by the FindNeighbors and FindClusters functions. The outcome of this process was a visual representation of the cells, rendered through the UMAP algorithm, providing a comprehensive overview of cellular distribution. The "FindMarkers" function was employed to pinpoint genes that were preferentially expressed within specific clusters, in addition to identifying DEGs. Each resultant cell cluster was annotated with reference to established cell-type marker genes, enhancing the interpretability of the data. To elucidate the distinct expression profiles of the identified genes at the single-cell level, the "scRNAtoolVis" package was utilized, providing a graphical interface that enabled precise and clear visualization of gene expression patterns.

High dimensional weighted gene co-expression network analysis

To explore genes associated with FGFRS-positive fibroblasts, we conducted a hdWGCNA using the "hdWGCNA" package. We created metacells for each sample and cell cluster, with 50 cells per metacell, and applied a standard pipeline of functions to analyze gene expression patterns and visualize module relationships in a reduced-dimensional space.

Machine learning-based construction of an FGFRS-fibroblast-related prognosis model

As previous published study (13, 14), we identified FGFRS-fibroblast-related genes using hdWGCNA and validated their predictive potential in tumor development with three transcriptome datasets, GSE17536 GSE29621, and GSE38832. Using the "mlr3" package, we developed a predictive model incorporating seven machine learning algorithms. Following cross-validation, we selected the most accurate model and assessed its predictive ability on an independent test set.

Trajectory and cell-cell communication analysis

Employing an unsupervised approach, pseudotemporal analysis was performed using the "Monocle" package, applying the DDR-Tree algorithm with its default parameters to delineate the developmental trajectory of fibroblasts. Following the pseudotemporal trajectory mapping, the "plot_pseudotime_heatmap" function was engaged to craft a heatmap. This visual tool effectively depicted the fluctuating expression patterns of a cohort of genes, illustrating their dynamic behavior along the pseudotime trajectories of fibroblasts. Furthermore, to uncover potential cellular interactions, both intracellular and extracellular, the "CellChat" package was deployed using its default settings and recommended pipeline configurations. This application facilitated the identification of communication networks among fibroblasts and other cellular components within the TME.

Enrichment analysis

The Seurat package's "FindMarkers" function was deployed to discern DEGs within each delineated cell subcluster. The selection criteria for these genes were stringent, requiring a fold change surpassing a threshold of 2 and an adjusted p-value below the significance level of 0.05. Subsequently, leveraging the identified DEGs, a comprehensive GSEA and KEGG enrichment analyses were conducted to explore the functional profiles of the cell subgroups. These analyses were executed utilizing the "clusterProfiler" package, which provided a robust framework for assessing the overrepresentation of specific gene sets and biological processes. To visually represent the functional enrichment results, the "GseaVis" package was employed. This tool facilitated the creation of intuitive and informative visualizations that encapsulated the enriched biological themes and pathways associated with the cell subclusters under investigation.

Non–negative matrix factorization analysis

To explore the diversity of FGFRS subtypes, we applied the NMF algorithm from the "NMF" package (15). The objective was to identify distinct subtypes characterized by unique gene expression patterns. To assess the prognostic value of these genes, we conducted survival analysis using the "survival" package. Additionally, we used the "ggrisk" package to analyze the survival and risk profiles of cancer patients, categorizing them into high- and low-risk groups. The analyses were considered statistically significant for P values below 0.05.

Statistical analysis

All computational analyses and graphical representations were executed utilizing the R software (v4.3.2). The strength and direction of the linear relationships between pairs of continuous variables were evaluated through Pearson's correlation coefficients. In the case of quantitative datasets, statistical comparisons among subgroups were made using either a two-tailed, unpaired Student's t-test for two-group comparisons or a one-way analysis of variance (ANOVA) complemented with Tukey's *post hoc* test for multiple group analyses. The threshold for statistical significance was set at a *P*-value of less than 0.05.

Results

Single-cell RNA sequencing analysis reveals FGFRS-positive fibroblast subsets

To elucidate the role of the fibroblast growth factor receptor (FGFR) signature in tumors, we analyzed single-cell RNA sequencing (scRNA-seq) data from colorectal cancer contained in the Gene Expression Omnibus (GEO) database (GSE231559, GSE166555) to identify potential fibroblast subsets that are positive for FGFRS ($FGFRS^+$). In the GSE231559 dataset, we annotated cell types using established markers, including T cells, malignant cells, neutrophils, myeloid cells, fibroblasts, B cells, epithelial cells, and plasma cells, each characterized by their unique gene expression profiles (Figure 1A). Notably, the analysis highlights genes like GZMK, GZMM, CD3G, which are expressed in T cells, and genes such as CD79A and CD19, which are characteristic of B cells. The myeloid and plasma cells are associated with genes like C1QC and IGHA1, respectively (Figure 1B). In the GSE166555 dataset, we identified 22 clusters and further annotated these clusters to 10 main clusters based classic markers, including B cells, DC cells, endothelial cells, fibroblasts, malignant cells, mast cells, monocyte/macrophage, plasma and T cells (Supplementary Figure S1A). These established markers including of MZB1, MS4A1, EPCAM, CD3D, CPA3, CD163, COL1A2, VWF, and IDO1 (Supplementary Figure S1C). Significantly, each cluster exhibited different gene expression profiles (Supplementary Figure S1D). Furthermore, to analyze the effect of FGFR signaling, we constructed an FGFR signature (FGFRS) using 86 genes involved in the FGFR signaling pathway from Molecular Signatures Database (MSigDB). This signature was then applied to score the identified cell groups using the "AddModuleScore" function in the Seurat package. Our findings reveal that fibroblast cell groups exhibit the highest FGFRS score, indicating a significant involvement of FGFR signaling in these cells (Figures 1C, Supplementary Figure S1B). To further dissect the role of FGFRS in fibroblast cell cluster, we isolated these cells and identified nine main fibroblast subgroups. Notably, in tumor tissues, a distinct fibroblast subgroup was found to have an increased proportion relative to normal tissues and showed highest FGFRS score (here after as "$FGFRS^+$ fibroblast") (Figures 1D, Supplementary Figure S1E). This observation suggested a potential role of FGFRS in the transformation and proliferation of fibroblasts in the TME. Furthermore, Gene Set Enrichment Analysis (GSEA) was performed on the transcriptome of this specific fibroblast subsets. The enrichment analysis identified several key cells signaling pathways that are potentially modulated by FGFRS. In the GSE231559 dataset, these pathways include receptor-mediated endocytosis, positive regulation of phagocytosis, and the regulation of miRNA metabolic processes, among others. In the $FGFRS^+$ fibroblast, some signaling pathways were significantly enriched, including epithelial tube morphogenesis and collagen fibril organization (Figure 1E). In the GSE16655 dataset, $FGFRS^+$ fibroblasts exhibited upregulated signaling pathways, including cytokine-cytokine receptor interactions, while downregulated pathways included the Wnt and Hippo signaling pathways (Supplementary Figure S1F). In summary, the integration of scRNA-seq data analysis highlighted that FGFRS is enriched in the fibroblast cluster within the TME, and these relevant signaling pathways may be critical for the role of $FGFRS^+$ fibroblasts in tumor progression.

Predictive model to identify core genes correlated to $FGFRS^+$ fibroblasts

To identify potential core genes associated with $FGFRS^+$ fibroblasts, we conducted high-dimensional Weighted Gene Co-expression Network Analysis (hdWGCNA), a comprehensive methodology for analyzing co-expression networks in the scRNA-seq data. This analysis aimed to detect co-expressed gene modules and unravel their functional roles within $FGFRS^+$ fibroblasts. Subsequently, we constructed a scale-free co-expression network by applying an optimal soft thresholding power of 12 (Supplementary Figure S2A). From this analysis, we distinguished a total of 19 distinct gene co-expression modules, identified the top 10 hub genes from these modules, and constructed protein-protein interaction (PPI) networks for the identified hub genes in each module (Supplementary Figure S2B, C, Figure 2B). Additionally, we investigated the correlation between each module (Supplementary Figure S2D), where modules 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 15, 18, and 19 displayed significant activation primarily in $FGFRS^+$ fibroblasts (Figure 2C, Supplementary Figure S2E). Subsequently, 325 genes from these modules underwent univariate Cox regression analysis in the TCGA COAD cohort, leading to the identification of 13 genes significantly associated with overall survival in CRC patients (Figure 2D). Additionally, by integrating data from the TCGA COAD and GEO database cohorts, we constructed robust models using 101 algorithmic combinations and calculated the area under the curve (AUC) for each model across all cohorts to assess their predictive capacity. (Figure 2A). Among the 101 models, the Step Cox (direction = both) algorithm in conjunction with a Random Survival Forest (RSF) demonstrated the highest AUC, serving as the basis for the final model creation. We further utilized RSF analysis to assess the prognostic relevance of various genes in predicting patient survival outcomes systematically. Notably, genes such as JDP2, HEYL, NRG1, RPS17, and MANF exhibited substantial predictive value, as indicated by their low Minimal Depth and high Variable Importance scores, thus influencing the accuracy of the survival model significantly (Figures 2E, F, Supplementary Figure S2F). The gene signature, comprising JDP2, HEYL, NRG1, RPS17, and MANF, known as the FRS, has demonstrated substantial predictive capabilities for patient survival at the 1-year, 3-year, and 5-year

10.3389/fimmu.2024.1493673

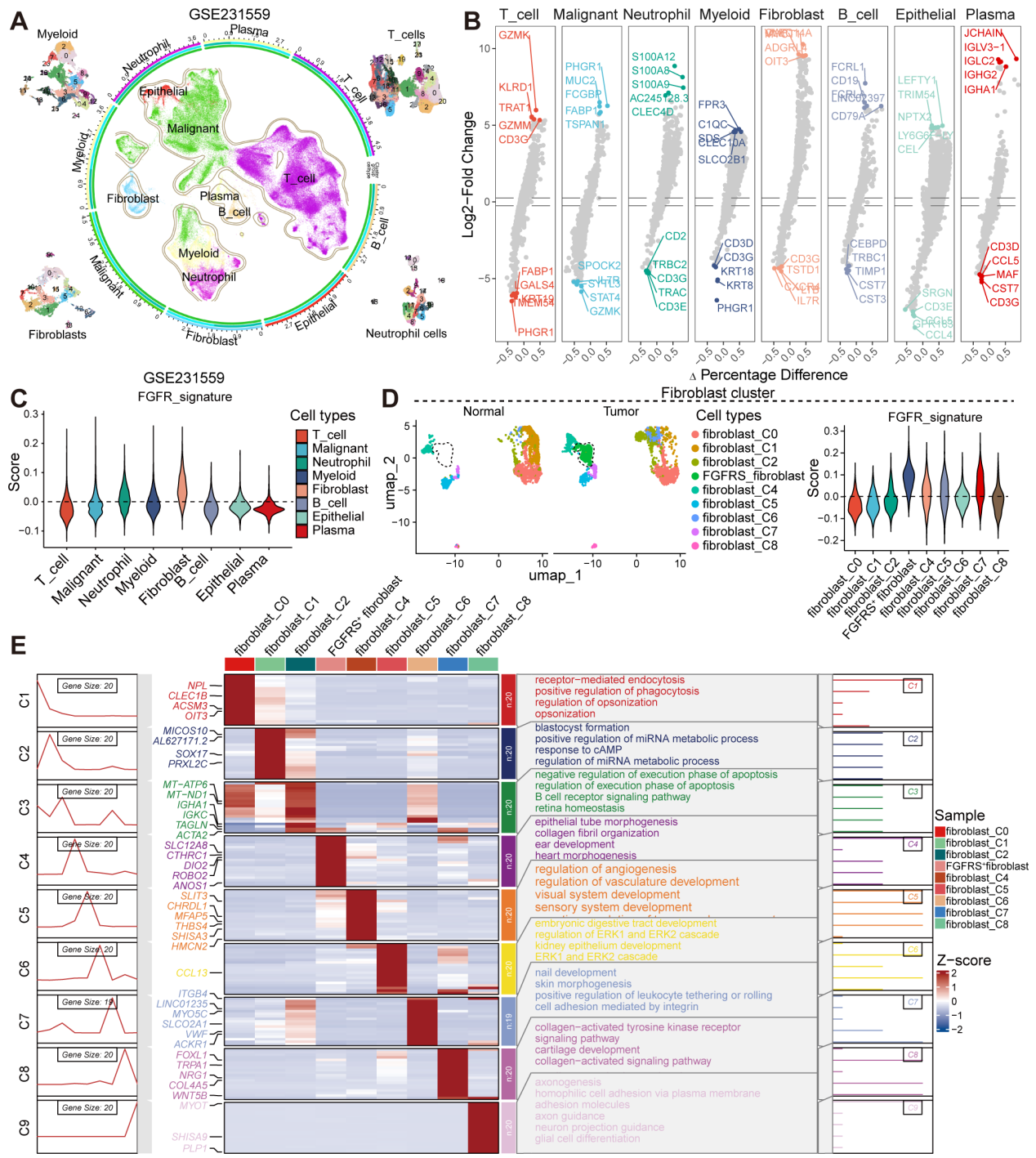


FIGURE 1
The scRNA-seq analysis uncovers the diversity of fibroblast subsets within CRC. **(A)** UMAP plot representing 8 distinct groups. The color coding corresponds to predominant cell types, with contours delineating cluster boundaries. The four corner insets provide a detailed view of myeloid, T cells, fibroblasts, and neutrophils. The peripheral axis displays the log-transformed total cell counts per class. The concentric colored tracks (exterior to interior) signify class identity (aligned with the central UMAP), cluster, group, and cell types. **(B)** Volcano plot showing the differential expression of markers in the distinct cell types. **(C)** Violin plot showing the score of FGFRS in the distinct cell types. **(D)** UMAP plot showing the fibroblast subsets in the normal or tumor tissues (left) and the score of FGFRS in the distinct fibroblast subsets (right). **(E)** The left panels illustrate the dynamic patterns of DEGs specific to each subset, while the central heatmap compares the expression profiles of these DEGs across populations. The right panels summarize the GO terms, providing insights into the biological functions associated with each cluster.

intervals. This integrated genetic profile serves as a valuable prognostic tool, providing insights into anticipated clinical outcomes for patients over the specified time frames (Figure 2G). Based on the FRS, cancer patients were categorized into high- and low-risk groups, with high-FRS patients exhibiting a poor prognosis (Figure 2H). These results were further validated using cohorts from the TCGA and GEO databases. In the TCGA cohort, Kaplan-Meier survival curve analysis revealed adverse outcomes for high-FRS patients, with an Area Under the Receiver Operating Characteristic curve (AUROC) of 0.7334 serving as a critical metric for model evaluation. These comprehensive results highlight the predictive capacity of the model and align with the analysis of COAD cohorts from the GEO database (Figures 2I-L).

FRS is correlated to the tumor immunity

To further analyze the correlation between the FRS and tumor immunity, we employed Non-negative Matrix Factorization (NMF), a robust technique for clustering analysis, to investigate the FRS of cancer patients based on the FGFRS. Our analysis specifically aimed to identify the optimal number of clusters (K) that would best capture the heterogeneity within the patient cohort. Following the NMF-based clustering, we observed that when K was set to 2, the consensus matrix revealed a clear distinction between two distinct patient groups. This finding suggests that dividing the patients into two groups at K=2 provides the most meaningful separation in terms of their FRS (Figure 3A, Supplementary Figure S3A). Notably, patients in cluster 1 exhibited a poorer prognosis compared to those in cluster 2 (Figure 3B). Upon TME score analysis, cluster 1 displayed lower immune scores, including stromal score, immune score, and ESTIMATE score, in contrast to cluster 2 (Supplementary Figure S3B). Additionally, CIBERSORT analysis was conducted for both clusters. Notably, tumors from patients in cluster 1 demonstrated reduced $CD8^+$ T cell infiltration but increased M0-like macrophages compared to those in cluster 2 (Figure 3C). Moreover, tumors from patients in cluster 1 exhibited decreased expression levels of immunostimulatory molecules (Supplementary Figure S3C). These findings imply that tumors from patients in cluster 2 might possess heightened antitumoral immunity within an immune-stimulating TME. The cell-cell interactions within the TME play a crucial role in tumor progression. Therefore, to further understand the interactions between $FGFRS^+$ fibroblasts and T cells, we conducted CellChat analysis. While the total interaction strength was notably increased in tumor tissues compared to normal tissues, the number and weight of interactions showed no significant differences (Supplementary Figures S3D, E). Subsequently, we investigated whether specific signaling interactions between $FGFRS^+$ fibroblasts and T cells were altered. Utilizing signal flow analysis, a quantitative method for assessing information transmission in biological systems, revealed substantial changes in signaling pathways between tumor and normal tissues. These results highlight a significant divergence in the flow of information within distinct tissue states (Figure 3D). In our subsequent analysis, we delved into the ligand-receptor interactions within the TME, focusing on the $FGFRS^+$ fibroblasts and their potential binding partners. Utilizing a comprehensive approach, we identified a significant interaction between collagen (COL)-related ligands and the CD44 receptor on T cells. Notably, this interaction was found to be markedly enhanced in the TME compared to normal tissues (Figure 3E). Our findings underscore the complex interplay between the extracellular matrix components and immune cells. Specifically, the pairs COL4A2-CD44, COL6A1-CD44, and COL1A2-CD44 demonstrated a pronounced increase in their relative contribution to the signal transduction within the tumor context (Figures 3F, Supplementary Figure S3F). Moreover, there was a significant increase in COLLAGEN signaling between FGFRS-positive fibroblasts and T cells in the tumor tissue compared to normal tissues (Figures 3G, H). This heightened interaction between collagen ligands and the CD44 receptor on T cells implies a potential role in influencing the immune response within the TME.

FRS were negatively correlated with T cell function

In our investigation of the influence of the FRS on T cell dynamics, we conducted a comprehensive analysis of scRNA-seq data from the GEO database. Using unsupervised clustering techniques, we analyzed T cell populations and identified seven primary clusters. By annotating these clusters with characteristic markers, we delineated six major T cell subsets: naïve T cells, regulatory T cells (Tregs), effector T cells (Teffs), central memory T cells (Tcms), natural killer T cells (NKTs), and exhausted T cells (Tex). Each subset exhibited distinct expression patterns of specific markers (Figures 4A, D, Supplementary Figure S4A, B). Our results revealed a notable shift in the proportion of T cell subsets between tumor tissues with high FRS scores and those with low FRS scores. Specifically, there was a significant increase in the proportion of Teff and NKT in tumors with low FRS scores, whereas the proportion of Tex was markedly reduced in comparison to tissues with high FRS scores (Figures 4B, C). Given the pivotal role of T cell function in determining the success of tumor immunity, we further analyzed the impact of FRS on the functionality of Teff cells. We observed a significant upregulation of effector molecules in Teff cells within tumors with low FRS scores, including PRF1, GZMH, GZMB, and IFNG (Figure 4E). Additionally, we utilized pseudotime analysis to elucidate the temporal evolution of T cell subset differentiation. Our findings indicated that in the early stages of differentiation, the T cell population was predominantly composed of Treg cells. As pseudotime progressed, there was a transition towards a predominance of naive T cells. In the late stages of differentiation, the population was characterized by an increase in Tex cells and NKT cells (Figures 4F, Supplementary Figure S4C). Importantly, the expression levels of CD8, GZMB, and IFNG and other markers associated with T cell function were significantly increased in the late stages of differentiation (Supplementary Figures S4D, E).

Utilizing the transcriptomic sequencing data from the TCGA COAD cohort, we embarked on a comparative analysis of gene expression differences between tumor tissues with high and low levels of FRS score. This analysis unveiled a spectrum of

10.3389/fimmu.2024.1493673

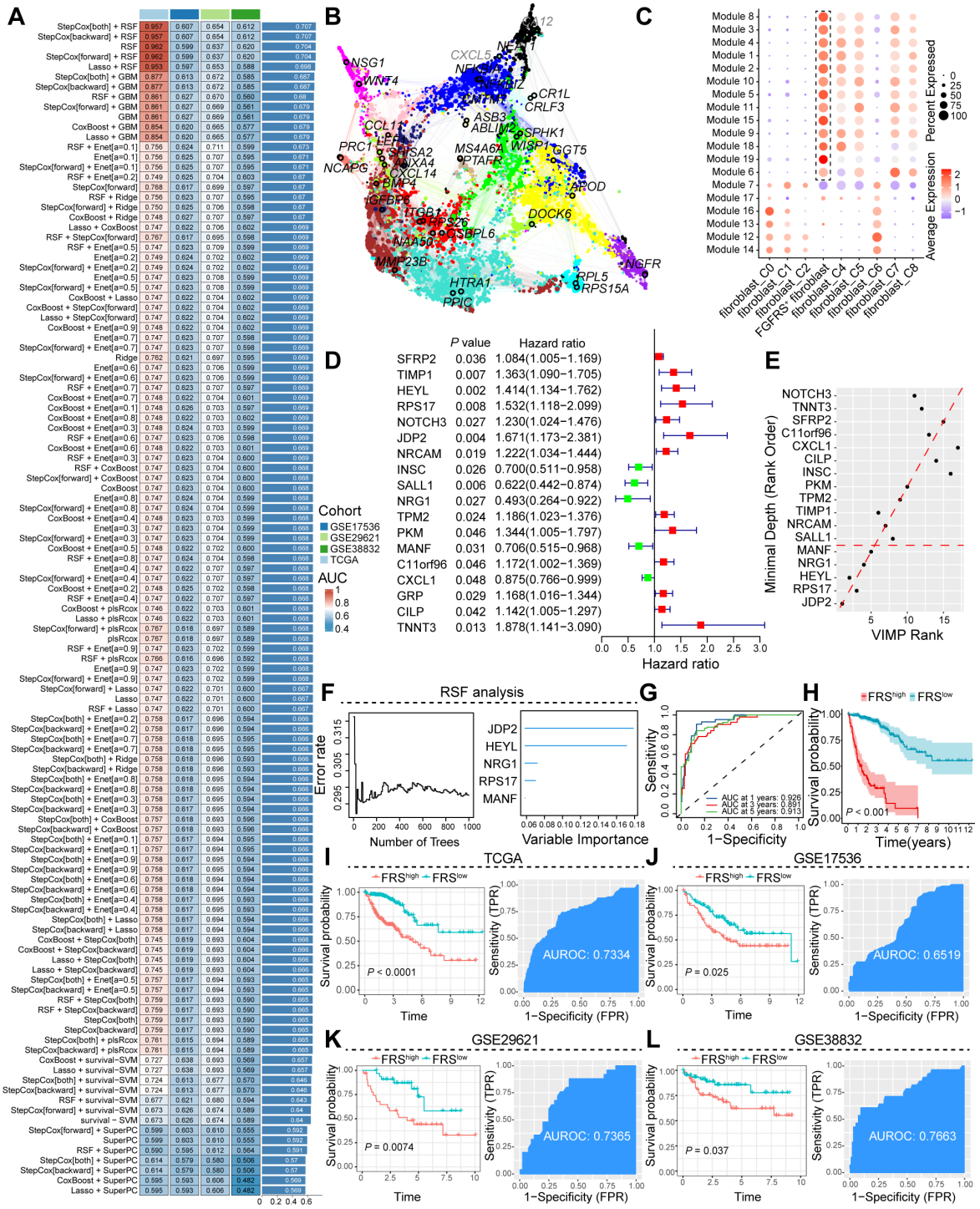


FIGURE 2
Construction and validation of the artificial intelligence-derived prognostic model. **(A)** Assessment of the area under the receiver operating characteristic curve (AUC) for multiple models generated from diverse combinations of machine learning algorithms across four cohorts. **(B)** Illustration of the interaction network among core prognostic genes, highlighting their interconnectivity. **(C)** Dot plot representation of the expression levels of module genes across distinct fibroblast subpopulations. **(D)** Forest plot delineating the hazard ratios associated with prognostic genes, quantifying their impact on patient outcomes. **(E, F)** RSF analysis, emphasizing the relative importance of prognostic genes in predicting survival. **(G)** The AUC curve at various time points, demonstrating the predictive accuracy of the FGFRS for survival. **(H)** Kaplan-Meier survival curves comparing the overall survival of cancer patients stratified by high and low FGFRS scores, FRS, FGFR risk score. **(I-L)** A comprehensive analysis was conducted to evaluate the survival benefits and predictive efficacy of the FSR score in cancer patients using data from TCGA and the GEO. The analysis was stratified according to high and low FRS expression.

10.3389/fimmu.2024.1493673

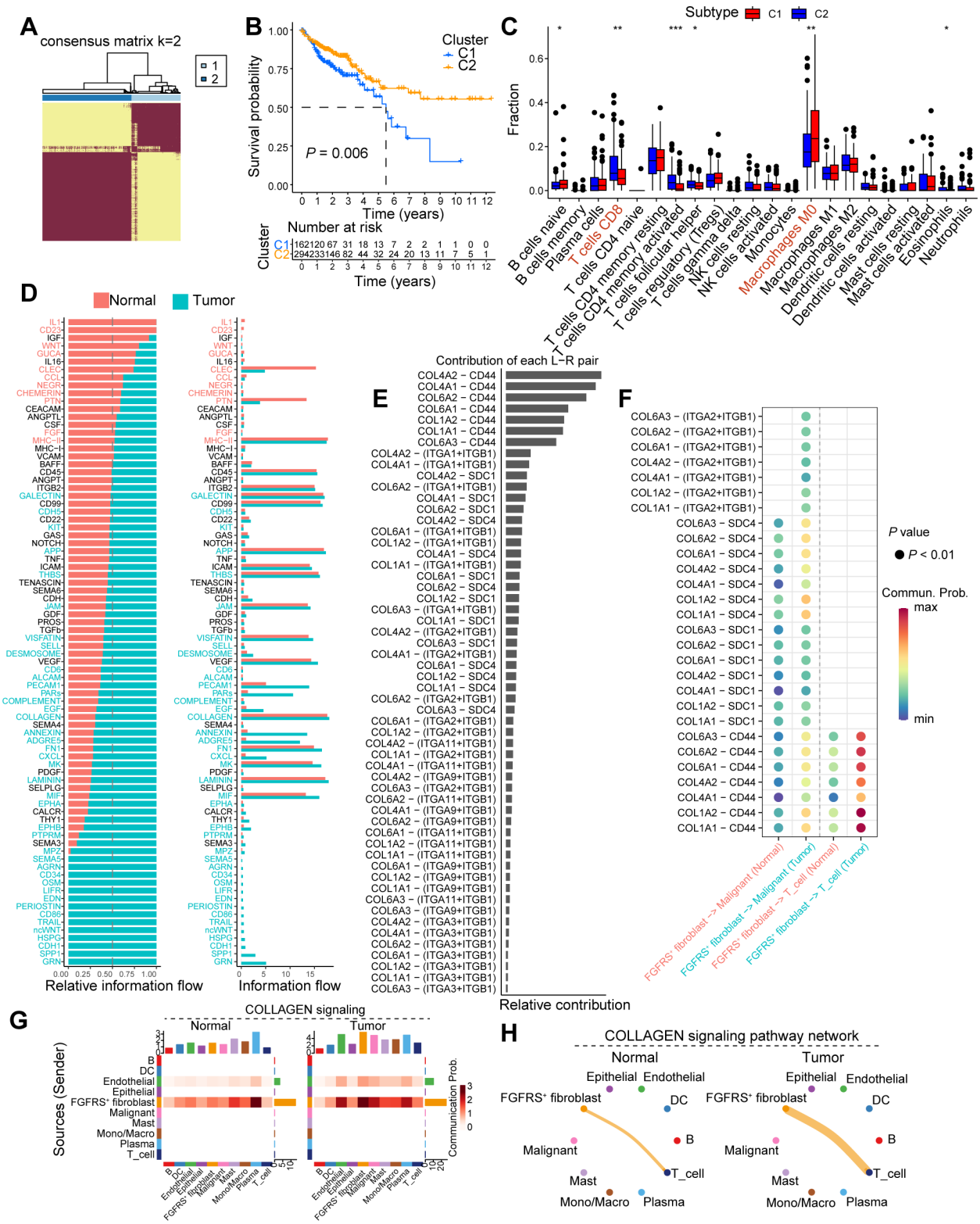


FIGURE 3
FGFRS-related molecule subtype correlates to tumor immunity. **(A)** Heatmap showing the clustering outcomes derived from NMF analysis, which has been applied to categorize patients into two distinct molecular subtypes based on the FGFRS signature. **(B)** Survival analysis of cancer patients in the cluster 1 and cluster 2. **(C)** Box plot showing the fractions of immune cells in the COAD tissues between cluster 1 and cluster 2. **(D)** The relative flow among cell clusters in the TME of normal and COAD tumors. **(E)** The contribution of ligand-receptor pair in the TME of COAD tissues. **(F)** The dot plot showing the interaction of ligand-receptor between FGFRS-positive fibroblasts and T cells. **(G)** Heatmap showing the changes of collagen signaling among cell clusters between normal and tumor tissues. **(H)** A shell plot illustrating the interaction of collagen signaling between FGFR score (FGFRS)-positive fibroblasts and T cells is presented. *P* values are from log-rank test **(B)** and two-way ANNOVA **(C)**. $^{*}P < 0.05$, $^{*}P < 0.01$, $^{***}P < 0.0001$.

10.3389/fimmu.2024.1493673

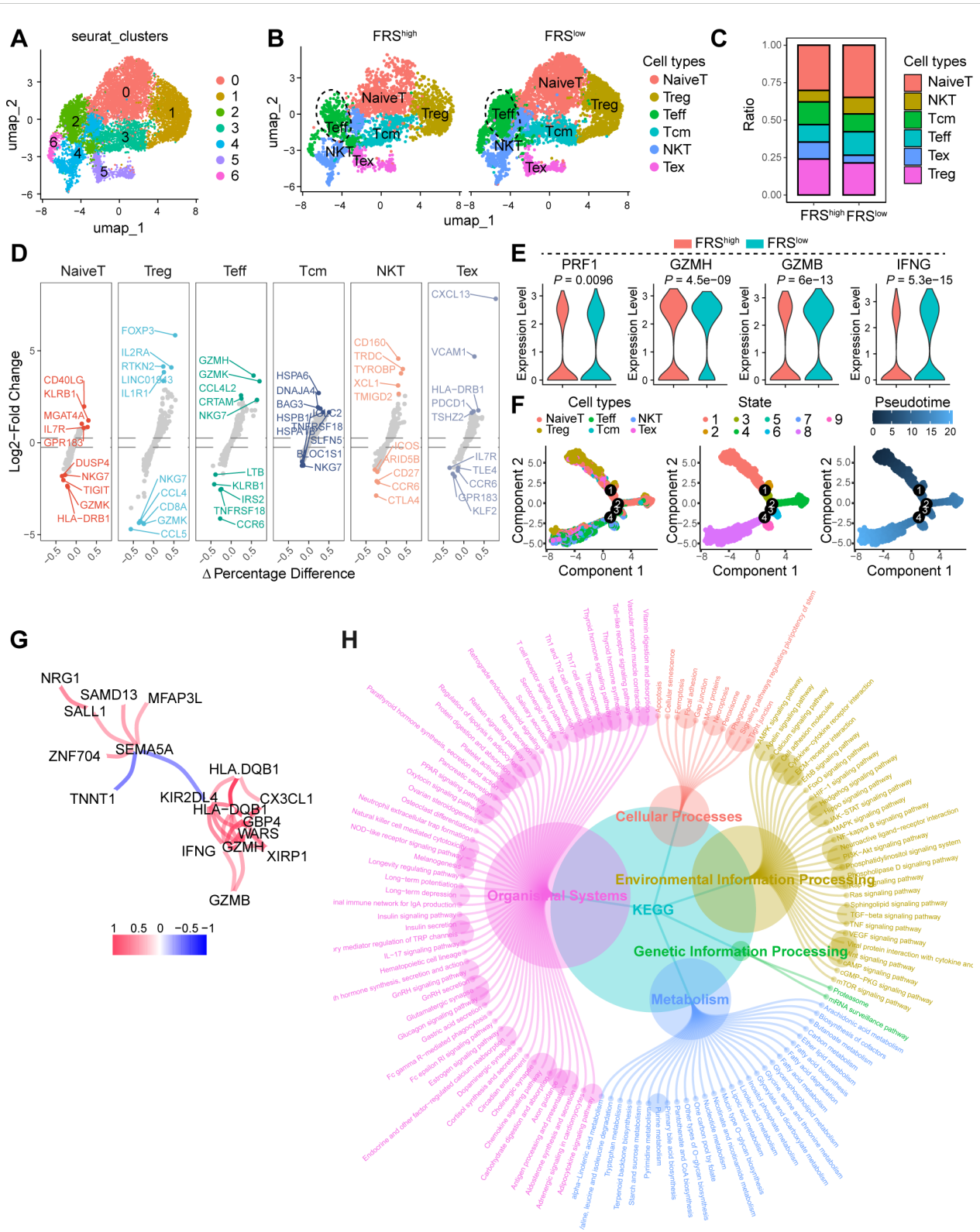


FIGURE 4
FRS correlates to the T cell function in the TME. **(A)** The UMAP depicting the cell cluster of T cells within COAD tumors, with each distinct color indicating a different cell type. **(B)** The UMAP visualizing the distribution of cell clusters in both FRS high and FAMS low COAD tumors, Treg, regulator T cells, Teff, effector T cells, Tcm, central memory T cells, NKT, natural killer cells, Tex, exhausted T cells. **(C)** A bar plot illustrating the ratio of cell clusters in the FRS high and FRS low COAD tumors. **(D)** Multiple volcano exhibiting the changes in markers across various cell types. **(E)** A violin plot displaying the expression levels of cytotoxic molecules in T effector cells among specified groups. **(F)** Pseudotime analysis demonstrating the differentiation trajectory of T cell subsets. **(G)** The network diagram illustrating the interactions of markers associated with T cell function. **(H)** KEGG analysis revealing the enriched signaling pathways.

differentially expressed genes (DEGs), which we subsequently subjected to a co-expression network analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. Notably, genes such as NRG1, SAMD13, MFAP3L, SALL1, ZNF704, SEMA5A, and HLA-DQB1 were identified as having altered expression patterns, potentially reflecting the diverse roles of FGFRS in tumor biology. The presence of immune-related genes such as CX3CL1, HLA-DQBY, and KIR2DL4, along with cytotoxic markers like GZMH and GZMB, suggests an intricate interplay between FRS signaling and immune response modulation within the TME (Figure 4G). Furthermore, the KEGG analysis elucidated the biological pathways and cellular processes significantly enriched among the differentially expressed genes (DEGs). Pathways such as the NF-kappa B signaling pathway, the STAT signaling pathway, and the TNF signaling pathway were prominently highlighted, indicating the roles of immune response and cell signaling in tumors altered by the FRS. Additionally, the identification of processes, including neutrophil extracellular trap formation and natural killer cell-mediated cytotoxicity, underscores the influence of the FRS on immune cell functions (Figure 4H). Collectively, FRS exhibited a negative correlation with T cell function, thereby impairing antitumor immunity and facilitating tumor progression.

FRS correlates to ICB resistance in murine MC38 tumor model

The expression of FRS is inversely associated with T cell function, but it is unclear whether this correlation contributes to immunotherapy resistance. In order to elucidate the role of FRS in colon tumors following immune checkpoint blockade (ICB) therapy, we developed an ICB-resistant MC38 tumor model and conducted RNA sequencing analysis. Within our experimental framework, mice engrafted with MC38 tumors were treated with either IgG or anti-CTLA4 antibodies, and we meticulously monitored changes in tumor volume. Upon completion of the treatment, we classified tumors that exhibited a significant reduction in volume following anti-CTLA4 treatment as "responder" phenotypes, in stark contrast to those that showed minimal change with IgG treatment, which we labeled as "nonresponder" phenotypes (Figure 5A). Of particular interest, the expression levels of HEYL, RPS17, and JDP2 were significantly elevated in tumors from nonresponders compared to those from responders (Figure 5B). To investigate the potential correlation between FRS expression and immune checkpoint blockade (ICB) resistance, we stratified nonresponders into two groups based on the gene set variation analysis (GSVA) score of the FRS and identified differentially expressed genes (DEGs) between high and low FRS-expressing tumors (Figures 5C, D). Utilizing these DEGs, we performed GSEA, revealing that several signaling pathways were significantly enriched in low FRS -expressing tumors, including those involved in extracellular matrix organization, collagen formation, and oxidative stress response. Notably, key genes associated with collagen and oxidative stress were found to be significantly downregulated in low FRS -expressing tumors (Supplementary Figures S5A-F). Further investigation into the impact of FRS on the TME of nonresponsive MC38 tumors using MCPcounter analysis indicated elevated levels of endothelial and fibroblast infiltration, alongside a diminished presence of T cells in the TME of tumors with high FRS expression (Figure 5E). Additionally, the expression of GZMA, a marker for cytotoxic T lymphocytes, was significantly higher in tumors with low FRS expression (Figure 5F). Our data suggest that reduced FRS may enhance T cell function within the TME. To comprehensively assess the influence of FRS expression on ICB-resistant tumor development, we conducted pseudotime analysis on RNA sequencing data from nonresponsive tumors. The findings revealed that the expression patterns of three key genes closely mirrored one another and were significantly elevated during the advanced stages of tumor progression (Figures 5G, H). Collectively, these results suggest that the FRS is positively associated with ICB resistance and may facilitate tumor progression.

FRS inhibitor significantly rewires the TME to promote tumor regression

As our analysis, the correlation between FRS expression and tumor immunity has been established, yet the extent to which FRS inhibitors can curtail tumor progression through the enhancement of antitumor immunity is not fully understood. To address this, we utilized the CTRP and PRISM databases to prognosticate potential inhibitors that target the FRS and subsequently validated the antitumoral efficacy of these inhibitors in murine tumor models. Through the amalgamation of data from both databases, we identified PHA-793887, recognized as a quintessential cyclin-dependent kinase (CDK) inhibitor, as a promising candidate with potential inhibitory effects on the FRS (Figures 6A, B). Then, we developed a mouse xenograft model of colorectal cancer by subcutaneously inoculating MC38 tumor cells and evaluated the impact of PHA-793887 on tumor progression and the TME via oral administration (Figure 6C). Notably, the treatment did not induce any discernible harm to the vital organs of the mice, encompassing the heart, liver, intestine, stomach, and lungs (Figure 6D). Importantly, the PHA-793887 significantly suppressed tumor growth and reduced the tumor burden relative to the control cohort (Figure 6E). Further examination of the TME utilizing flow cytometry disclosed that PHA-793887 markedly augmented the secretion of cytotoxic T cell-associated cytokines, including interferon-gamma (IFN-γ), tumor necrosis factor-alpha (TNF-α), and granzyme B (GZMB), within the tumor milieu (Figure 6F). This suggests an enhancement of T cell functionality. Moreover, in tumors treated with PHA-793887, there was a significant increase in the infiltration of $CD8^+$ T cells and effector $CD4^+$ T cells, along with a decrease in tumor-associated macrophages (TAMs) and regulatory T cells (Tregs) (Figures 6G–J). These findings underscore that PHA-793887, acting as an FRS inhibitor, can profoundly rewire the immunosuppressive TME, bolster anti-tumor immune responses, and reduce tumor growth.

10.3389/fimmu.2024.1493673

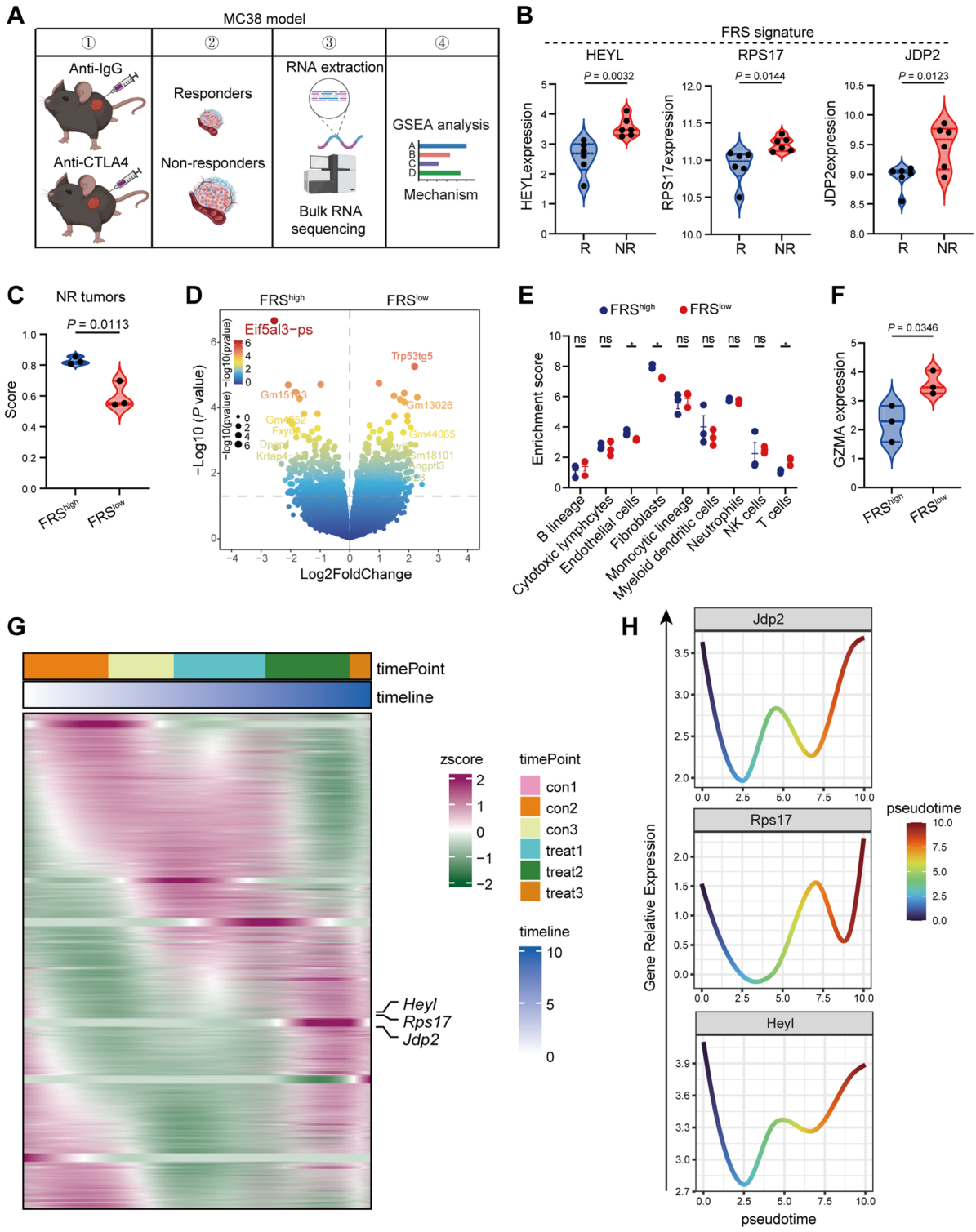


FIGURE 5
High FRS score correlates positively with ICB resistance in murine MC38 tumors. **(A)** The schematic representation of the experimental timeline for the RNA-seq analysis of ICB resistant murine MC38 tumors. The image was illustrated by the Biorender.com. **(B)** Expression profiles of pivotal genes within the FRS in responsive and nonresponsive MC38 tumors, highlighting variations in gene activity, R, responders, NR, non-responders. **(C)** Violin plot depicting the distribution of scores for tumors with high and low expression of FRS, specifically within the nonresponsive MC38 tumor group. **(D)** A volcano plot illustrating the differential gene expression between tumors exhibiting high versus low FRS score in the nonresponsive category. **(E)** MCPcounter analysis revealing the enrichment of distinct cell clusters associated with high and low FRS score in nonresponsive tumors. **(F)** A violin plot displaying the expression levels of GZMB in tumors with high and low FRS score among the nonresponsive group. **(G)** Heatmap representing the expression patterns of key genes derived from pseudotime analysis of bulk RNA-seq data from nonresponsive tumors. **(H)** Line graphs illustrating the trend of key gene expression changes within the FRS across the nonresponsive tumors. P values are from unpaired t-tests **(B, C, F)** and two-way ANOVA **(E)**.

10.3389/fimmu.2024.1493673

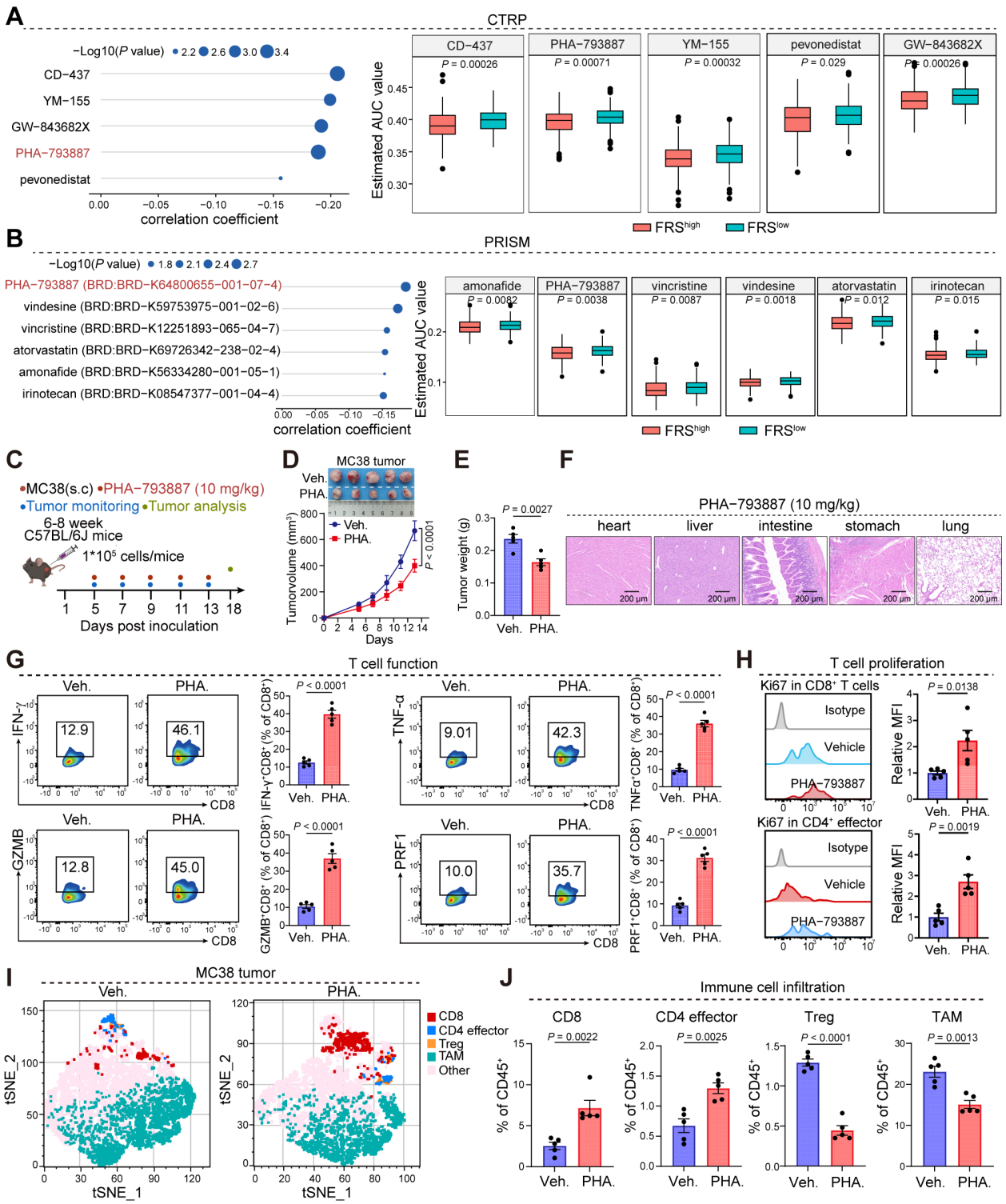


FIGURE 6
PHA-793887 enhances antitumor immunity in MC38 tumor model. **(A, B)** The prediction of drugs targeting FRS via CTRP **(A)** and PRISM **(B)** database. **(C)** The schedule process of animal experiments. **(D, E)** Tumor growth **(D)** and tumor weight **(E)** of MC38 tumors bearing mice treated with Veh. or PHA. (10 mg/kg). Veh., vehicle; PHA., PHA-793887. **(F)** Representative images of hematoxylin and eosin (HE) staining were obtained to evaluate organ toxicity in mice treated with either vehicle or PHA-793887. **(G)** The expression levels of cytotoxic molecules in the $CD8^+$ T cells within MC38 tumors treated with Veh. or PHA. **(H)** The expression of ki67 in the $CD8^+$ T cells (top) and CD4 effector cells (bottom). **(I)** tSNE plot showing the distribution of immune cell clusters. Distinct colors represent the different cell clusters. **(J)** The box plot showing the percentages of $CD8^+$ T cells, CD4 effector cells, Treg cells, and TAMs in the MC38 tumors treated with Veh. or PHA, Treg, regulator T cell, TAM, tumor-associated macrophage, Veh., Vehicle, PHA., PHA-793887. *P* values are from a two-tailed unpaired Student's t-test **(A, B, E, G, H, J)** and two-way ANOVA **(D)**.

10.3389/fimmu.2024.1493673

FRS is positively correlated ICB therapy resistance in cancer patients

To further elucidate the association between FRS and the therapeutic efficacy of ICB in cancer patients, we conducted an integrative analysis using TCGA and the Tumor Immune Dysfunction and Exclusion (TIDE) database. Our analysis revealed that patients with tumors exhibiting high FRS score had a significantly lower survival rate compared to those with low FRS score within their tumors (Figure 7A). Additionally, by examining patients within the TIDE database who underwent ICB treatment, we observed a notable increase FRS among patients with stable disease (SD) or progressive disease (PD) compared to those with partial response (PR) (Figure 7B). Utilizing the TIDE scoring system, we identified a higher prevalence of responders among patients with low FRS expression (Figure 7C). Furthermore, the SubMap analysis revealed that the low FRS group exhibited a high likelihood of response to anti-PD-1 in the immunotherapy cohorts (IMvigor210) (Figure 7D). In the context of patients undergoing ICB therapy, we conducted a detailed analysis of the prognostic impact of key genes within the FRS. Our findings indicate that tumors with elevated expression levels of HEYL, RPS17, and JDP2 are associated with a poorer prognosis for cancer patients (Figures 7E-G). These results suggest that the FRS may serve as a

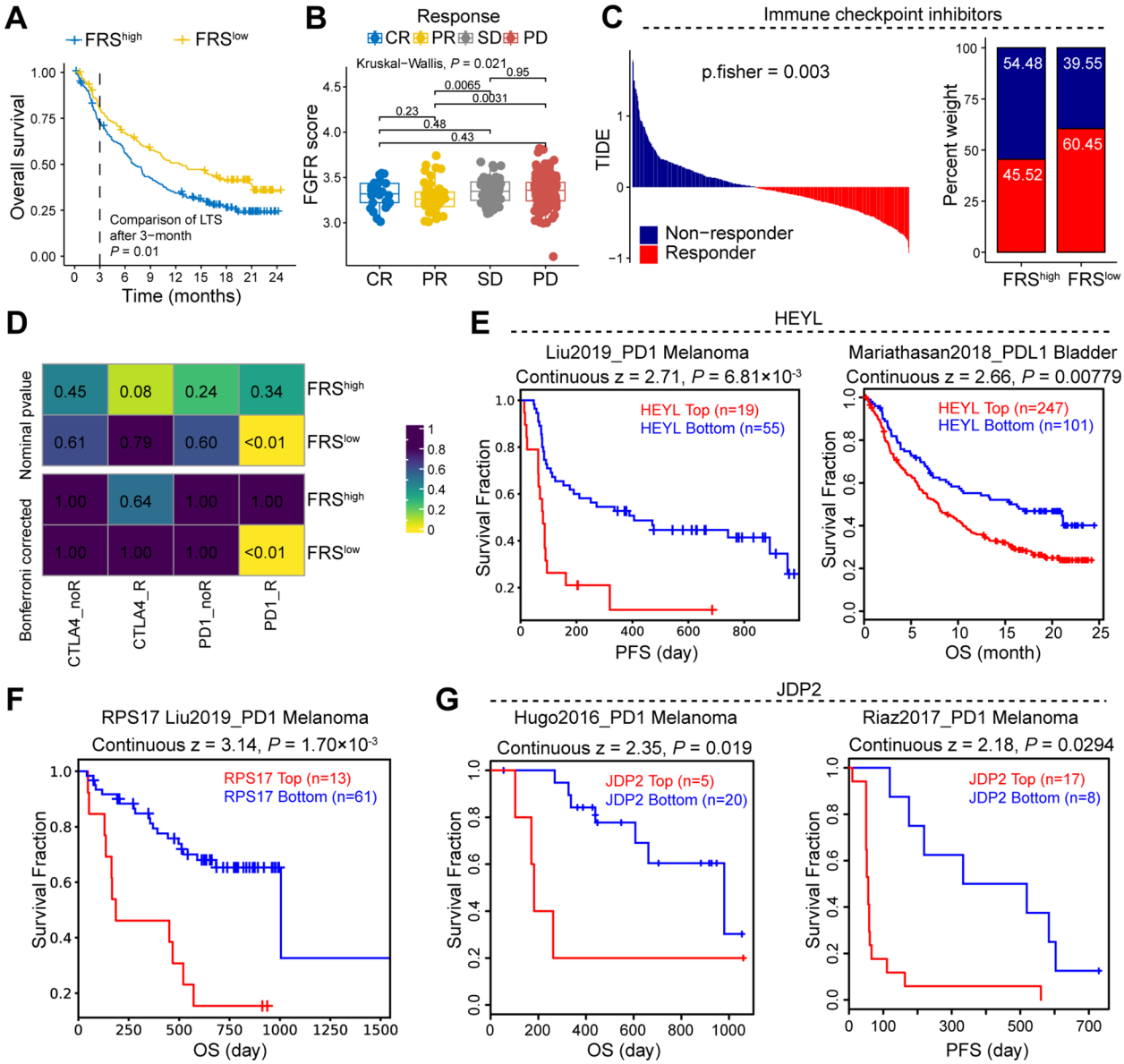


FIGURE 7
FRS is negatively correlated to the response of immune therapy in cancer patients. **(A)** Survival curve of cancer patients with high or low FRS score. **(B)** The FRS score of cancer patients in the indicated groups, CR, complete response, PR, partial response, SD, stable disease, PD, progressive disease. **(C)** TIDE analysis showing the immune therapy response of cancer patients. **(D)** Submap showing the correlation between FRS score and therapy response of ICB in the cancer patients. **(E-G)** The survival analysis of cancer patients received ICB therapy in indicated groups, PFS, progression-free survival, OS, overall survival. P values are from the log-rank test **(A, E-G)** and one-way ANOVA **(B)**. Pearson's correlation coefficient is calculated **(D)**.

prognostic indicator for ICB treatment outcomes, with a lower FRS score potentially correlating with a more favorable response to ICB therapy.

Discussion

Our comprehensive analysis leveraging scRNA-seq data from CRC cohorts has unveiled the intricate relationship between FRS and tumor biology, particularly within the TME. The identification of $FGFRS^+$ fibroblast subsets through scRNA-seq analysis provides a foundation for understanding the heterogeneity of cancer and the potential for targeted therapeutic intervention. The significant association between FRS and fibroblasts suggests a pivotal role for these cells in modulating the TME, which could be a critical factor in tumor progression and response to therapy. The enrichment of FRS in fibroblasts, as evidenced by the highest module scores, indicates that these cells may be key players in the tumor's resistance to ICB therapy. This finding is supported by the observation that $FGFRS^+$ fibroblasts display distinct gene expression profiles linked to pathways, including epithelial tube morphogenesis and collagen fibril organization, which are known to influence tumor growth and invasion.

Previous studies have reported the relevant signature of immune cells, such as macrophages, predicted patient prognosis and therapy resistance (16–19). The development of a predictive model using machine learning algorithms has been instrumental in identifying a gene signature that strongly predicts patient survival outcomes (20, 21). The FRS, composed of genes like JDP2, HEYL, NRG1, RPS17, and MANF, has demonstrated robust predictive capabilities across various cohorts. This underscores the potential of using FRS as a prognostic tool in clinical settings to stratify patients into high and low-risk groups, thereby personalizing treatment strategies. Furthermore, our analysis elucidates the correlation between the FRS and tumor immunity. The identification of distinct patient clusters based on FRS, along with the associated differences in immune scores and cell infiltration patterns, suggests a complex interplay between the FRS and immune cell dynamics within the TME. Notably, the reduced infiltration of $CD8^+$ T cells and the increased presence of M0-like macrophages in tumors with higher FRS scores indicate a possible mechanism through which FRS enrichment may suppress antitumor immunity.

Many drugs harbor significant anti-tumor, whereas the mechanism remains elusive (22). Previous research has established PHA-793887 as a potent, ATP-competitive cyclin-dependent kinase (CDK) inhibitor capable of inhibiting key cell cycle regulators such as CDK2, CDK1, CDK4, AND CDK9 (23, 24). While its efficacy in disrupting cell cycle progression is well-documented, the exploration of additional potential targets and its impact on tumor immunity remains less explored (25). Our study takes a significant step toward addressing this gap by employing in silico drug prediction methods to identify FRS as a potential target for PHA-793887. The application of PHA-793887 in our murine model of colorectal cancer has yielded promising results, demonstrating a significant inhibitory effect on tumor growth without any detectable organ toxicity. In the TME, various factors could contribute to T cell dysfunction, subsequently promoting immunotherapy resistance and tumor progression (26, 27). In this study, we found PHA-793887 treatment induced an enhanced secretion of cytotoxic T cell-associated cytokines and facilitated an increase in the infiltration of $CD8^+$ T cells and effector $CD4^+$ T cells within the TME. These observations suggest an improvement in T cell functionality, indicative of the compound's potential to bolster antitumor immunity. Together, these findings underscore the PHA-793887 could be a potential inhibitor enhancing antitumor immunotherapy for the cancer patients with CRC. Future research should focus on clinical trials to evaluate the efficacy of FRS inhibitors in combination with ICB therapies and on further elucidating the mechanisms by which FRS modulates the TME and immune responses in CRC patients.

Conclusion

Our study presents a multi-faceted perspective on the role of FRS in colorectal cancer, highlighting its enrichment in fibroblasts, its correlation with ICB resistance, and its impact on tumor immunity. The findings have implications for the development of novel therapeutic strategies targeting the FRS and for the refinement of prognostic models to better predict patient outcomes. Importantly, PHA-793887 could be a potential inhibitor targeting the FRS in the immunotherapy of CRC patients.

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. The public databases presented in the study are deposited in the Gene Expression Omnibus repository (http://www.ncbi.nlm.nih.gov/geo/), accession number: GSE231559, GSE166555, GSE17536, GSE29621, and GSE38832.

Ethics statement

The animal studies were approved by Ethics Committee of the First Affiliated Hospital of Chongqing Medical University (2023-6S). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent was obtained from the owners for the participation of their animals in this study.

Author contributions

XL: Data curation, Investigation, Software, Writing – original draft. ZP: Data curation, Formal Analysis, Software, Writing – review & editing. TL: Investigation, Methodology, Software, Writing – review & editing. QX: Methodology, Writing – review & editing. LL: Methodology, Writing – review & editing. QW: Writing – review & editing. GY: Software, Writing – review & editing. XZ: Funding acquisition, Supervision, Writing – review & editing. DS: Software, Supervision, Validation, Visualization, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was funded by the Sponsored by Natural Science Foundation of Chongqing, China (CSTB2024NSCQ-MSX0036).

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.

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.2024.1493673/full#supplementary-material

References

1. Biller LH, Schrag D. Diagnosis and treatment of metastatic colorectal cancer: A review. *Jama*. (2021) 325:669–85. doi: 10.1001/jama.2021.0106

2. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *Lancet (London England)*. (2019) 394:1467–80. doi: 10.1016/s0140-6736(19)32319-0

3. Xie Y, Su N, Yang J, Tan Q, Huang S, Jin M, et al. FGF/FGFR signaling in health and disease. *Signal Transduction Targeted Ther*. (2020) 5:181. doi: 10.1038/s41392-020-00222-7

4. Katoh M, Loriot Y, Brandi G, Tavolari S, Wainberg ZA, Katoh M. FGFR-targeted therapeutics: clinical activity, mechanisms of resistance and new directions. *Nat Rev Clin Oncol*. (2024) 21:312–29. doi: 10.1038/s41571-024-00869-z

5. Ruan R, Li L, Li X, Huang C, Zhang Z, Zhong H, et al. Unleashing the potential of combining FGFR inhibitor and immune checkpoint blockade for FGF/FGFR signaling in tumor microenvironment. *Mol Cancer*. (2023) 22:60. doi: 10.1186/s12943-023-01761-7

6. Janani B, Vijayakumar M, Priya K, Kim JH, Prabakaran DS, Shahid M, et al. EGFR-based targeted therapy for colorectal cancer-promises and challenges. *Vaccines (Basel)*. (2022) 10:499. doi: 10.3390/vaccines10040499

7. Guler I, Askan G, Klostergaard J, Sahin IH. Precision medicine for metastatic colorectal cancer: an evolving era. *Expert Rev Gastroenterol Hepatol*. (2019) 13:919–31. doi: 10.1080/17474124.2019.1663174

8. Zhao W, Jin L, Chen P, Li D, Gao W, Dong G. Colorectal cancer immunotherapy-Recent progress and future directions. *Cancer Lett*. (2022) 545:215816. doi: 10.1016/j.canlet.2022.215816

9. Fan A, Wang B, Wang X, Nie Y, Fan D, Zhao X, et al. Immunotherapy in colorectal cancer: current achievements and future perspective. *Int J Biol Sci*. (2021) 17:3837–49. doi: 10.7150/ijbs.64077

10. Kennedy LB, Salama AKS. A review of cancer immunotherapy toxicity. *CA Cancer J Clin*. (2020) 70:86–104. doi: 10.3322/caac.21596

11. Rui R, Zhou L, He S. Cancer immunotherapies: advances and bottlenecks. *Front Immunol*. (2023) 14:1212476. doi: 10.3389/fimmu.2023.1212476

12. Mann JE, Lucca L, Austin MR, Merkin RD, Robert ME, Al Bawardy B, et al. ScRNA-seq defines dynamic T-cell subsets in longitudinal colon and peripheral blood samples in immune checkpoint inhibitor-induced colitis. *J Immunother Cancer*. (2023) 11:e007358. doi: 10.1136/jitc-2023-007358

13. Sun Y, Wu J, Zhang Q, Wang P, Zhang J, Yuan Y. Single-cell hdWGCNA reveals metastatic protective macrophages and development of deep learning model in uveal melanoma. *J Transl Med*. (2024) 22:695. doi: 10.1186/s12967-024-05421-2

14. 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

15. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature*. (1999) 401:788–91. doi: 10.1038/44565

16. Wu X, Lu W, Xu C, Jiang C, Zhuo Z, Wang R, et al. Macrophages phenotype regulated by IL-6 are associated with the prognosis of platinum-resistant serous ovarian cancer: integrated analysis of clinical trial and omics. *J Immunol Res*. (2023) 2023:6455704. doi: 10.1155/2023/6455704

17. Yu Y, Huang Y, Li C, Ou S, Xu C, Kang Z. Clinical value of M1 macrophage-related genes identification in bladder urothelial carcinoma and in *vitro* validation. *Front Genet*. (2022) 13:1047004. doi: 10.3389/fgene.2022.1047004

18. Zhang S, Jiang C, Jiang L, Chen H, Huang J, Gao X, et al. Construction of a diagnostic model for hepatitis B-related hepatocellular carcinoma using machine learning and artificial neural networks and revealing the correlation by immunoassay. *Tumor Virus Res*. (2023) 16:200271. doi: 10.1016/j.tvr.2023.200271

19. Zhang X, Zhang P, Cong A, Feng Y, Chi H, Xia Z, et al. Unraveling molecular networks in thymic epithelial tumors: deciphering the unique signatures. *Front Immunol*. (2023) 14:1264325. doi: 10.3389/fimmu.2023.1264325

20. Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. *N Engl J Med*. (2023) 388:1201–8. doi: 10.1056/NEJMra2302038

21. Patel L, Shukla T, Huang X, Ussery DW, Wang S. Machine learning methods in drug discovery. *Molecules (Basel Switzerland)*. (2020) 25:5277. doi: 10.3390/molecules25225277

22. Noorbakhsh Varnosfaderani SM, Ebrahimzadeh F, Akbari Oryani M, Khalili S, Almasi F, Mosaddeghi Heris R, et al. Potential promising anticancer applications of β-glucans: a review. *Biosci Rep*. (2024) 44:BSR20231686. doi: 10.1042/bsr20231686

23. Locatelli G, Bosotti R, Ciomei M, Brasca MG, Calogero R, Mercurio C, et al. Transcriptional analysis of an E2F gene signature as a biomarker of activity of the cyclin-dependent kinase inhibitor PHA-793887 in tumor and skin biopsies from a phase I clinical study. *Mol Cancer Ther*. (2010) 9:1265–73. doi: 10.1158/1535-7163.Mct-09-1163

24. Wu B, Yang W, Fu Z, Xie H, Guo Z, Liu D, et al. Selected using bioinformatics and molecular docking analyses, PHA-793887 is effective against osteosarcoma. *Aging (Albany NY)*. (2021) 13:16425–44. doi: 10.18632/aging.203165

25. Massard C, Soria JC, Anthoney DA, Proctor A, Scaburri A, Pacciarini MA, et al. A first in man, phase I dose-escalation study of PHA-793887, an inhibitor of multiple cyclin-dependent kinases (CDK2, 1 and 4) reveals unexpected hepatotoxicity in patients with solid tumors. *Cell Cycle*. (2011) 10:963–70. doi: 10.4161/cc.10.6.15075

26. Deng Y, Shi M, Yi L, Naveed Khan M, Xia Z, Li X. Eliminating a barrier: Aiming at VISTA, reversing MDSC-mediated T cell suppression in the tumor microenvironment. *Heliyon*. (2024) 10:e37060. doi: 10.1016/j.heliyon.2024.e37060

27. Xia Z, Chen S, He M, Li B, Deng Y, Yi L, et al. Editorial: Targeting metabolism to activate T cells and enhance the efficacy of checkpoint blockade immunotherapy in solid tumors. *Front Immunol*. (2023) 14:1247178. doi: 10.3389/fimmu.2023.1247178

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 03 December 2024
DOI 10.3389/fimmu.2024.1493377

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Cheng Wang,
Sun Yat-sen University, China
Lin-Lin Bu,
Wuhan University, China
Chunjie Li,
Sichuan University, China

*CORRESPONDENCE
Jiannan Liu
liujiannan@sh9hospital.org.cn

†These authors have contributed equally to this work and share first authorship

RECEIVED 09 September 2024
ACCEPTED 18 November 2024
PUBLISHED 03 December 2024

CITATION
Sun Y, Cheng G, Wei D, Luo J and Liu J (2024) Integrating omics data and machine learning techniques for precision detection of oral squamous cell carcinoma: evaluating single biomarkers. *Front. Immunol.* 15:1493377. doi: 10.3389/fimmu.2024.1493377

Integrating omics data and machine learning techniques for precision detection of oral squamous cell carcinoma: evaluating single biomarkers

Yilan Sun[1,2,3,4,5,6†], Guozhen Cheng[7†], Dongliang Wei[1,2,3,4,5,6], Jiacheng Luo[2] and Jiannan Liu[1,2,3,4,5,6*]

[1]Department of Oral and Maxillofacial Head and Neck Oncology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [2]College of Stomatology, Shanghai Jiao Tong University, Shanghai, China, [3]National Center for Stomatology, Shanghai, China, [4]National Clinical Research Center for Oral Diseases, Shanghai, China, [5]Shanghai Key Laboratory of Stomatology, Shanghai, China, [6]Shanghai Research Institute of Stomatology, Shanghai, China, [7]College of Mechanical and Electrical Engineering, Fujian Agriculture and Forestry University, Fuzhou, China

Introduction: Early detection of oral squamous cell carcinoma (OSCC) is critical for improving clinical outcomes. Precision diagnostics integrating metabolomics and machine learning offer promising non-invasive solutions for identifying tumor-derived biomarkers.

Methods: We analyzed a multicenter public dataset comprising 61 OSCC patients and 61 healthy controls. Plasma metabolomics data were processed to extract 29 numerical and 47 ratio features. The Extra Trees (ET) algorithm was applied for feature selection, and the TabPFN model was used for classification and prediction.

Results: The model achieved an area under the curve (AUC) of 93% and an overall accuracy of 76.6% when using top-ranked individual biomarkers. Key metabolic features significantly differentiated OSCC patients from healthy controls, providing a detailed metabolic fingerprint of the disease.

Discussion: Our findings demonstrate the utility of integrating omics data with advanced machine learning techniques to develop accurate, non-invasive diagnostic tools for OSCC. The study highlights actionable metabolic signatures that have potential applications in personalized therapeutics and early intervention strategies.

KEYWORDS

machine learning, oral squamous cell carcinoma, precision metabolomics, feature selection, personalized therapy

1 Introduction

Oral squamous cell carcinoma (OSCC) is the most common malignancy affecting the oral cavity, with a mortality rate exceeding 50% (1). Postoperative OSCC can severely impact patients' speech, chewing, and swallowing functions, significantly affecting their quality of life (2). Many OSCC patients are diagnosed at advanced stages, missing the window for optimal treatment. Early diagnosis is crucial, as it significantly improves survival rates and treatment outcomes and reduces costs. This highlights the need for more sensitive and specific diagnostic methods for early OSCC detection.

Currently, imaging combined with histopathology remains the gold standard for OSCC screening (3). Detecting early asymptomatic cases of OSCC is challenging despite straightforward oral imaging and sampling. Incisional biopsies cause physical trauma and suffer from sampling accuracy issues due to tumor heterogeneity. Molecular diagnostics, which detect subtle phenotypic changes that occur prior to malignancy or metastasis, have become crucial tools for early detection. Therefore, developing effective multianalyte detection methods for biofluids is urgently needed (4).

Tumors, including OSCC, are rich in blood vessels, facilitating the shedding of tumor cells and molecules into the bloodstream, making blood-based tests an effective screening tool for early detection. Thus, plasma is an ideal diagnostic fluid for the molecular diagnosis and early screening of OSCC (5). It can be sampled alongside routine blood tests, making it convenient to collect samples during outpatient visits or regular check-ups. Owing to its diverse components, including the genome, transcriptome, proteome, microbiome, and metabolome, plasma is a potential source of biomarkers. Its diversity makes blood a promising medium for OSCC metabolite marker screening, which can offer insights into metabolic pathways (5). Previous studies have reported various blood metabolites associated with early OSCC screening, demonstrating its potential as a noninvasive diagnostic tool.

Research has shown that the lipid content in the plasma of OSCC patients is significantly lower than that in the plasma of healthy controls (HC), with certain types of lipids being reduced by at least twofold (4). Disparities in sphingolipid levels between OSCC patients and healthy individuals have led to diagnostic methods with high accuracy, sensitivity, and specificity. Lower levels of certain amino acids and phosphatidylcholines in OSCC patients are associated with poorer survival rates, suggesting their roles in tumor progression and potential as predictive biomarkers (6). An integrated analysis of plasma metabolomics data revealed distinct profiles indicative of disrupted metabolic pathways, particularly in advanced disease stages, potentially fostering tumor growth and suppressing immune responses (7).

While the genome consists of approximately 20,000 protein-coding genes, the metabolome presents a smaller yet more dynamic landscape with approximately 220,000 metabolites noted in the HMDB (4). The metabolome's precise nature and direct reflection of the physiological state make metabolites ideal candidates for prognostic, diagnostic, and therapeutic monitoring applications (5, 8). However, the diversity among cancer patients requires a deeper understanding of specific tumor metabolisms, including those involved in OSCC, to tailor effective treatments and screening strategies.

To be clinically applicable, molecular screening must consider several factors: 1) the inclusion of highly specific and sensitive measurable markers; 2) convenient sampling with high patient acceptance; 3) affordable and accessible analytical technology platforms; and 4) rapid feedback for clinical diagnostic decision-making. Balancing these factors is essential for developing effective analytical methods (9, 10).

Mass spectrometry (MS) is a widely utilized technology in metabolomics that is capable of qualitative and quantitative analysis of small molecules and is widely used in biomedical fields (4, 11). The advantages of the MS platform include high specificity and sensitivity for biomarker screening, mature detection techniques, clear detection processes, and controlled costs. Additionally, MS can provide rapid feedback, enabling quick molecular screening. The combination of MS and machine learning (ML) successfully translates metabolomics analysis into clinical diagnostic decisions (12, 13). The application of MS/ML methods can achieve routine blood diagnostics, enabling rapid, accurate, cost-effective, and sustainable early screening and intelligent diagnosis of OSCC, thereby offering new strategies for early detection, diagnosis, and treatment (14).

In this study, we used a publicly available dataset with plasma samples from OSCC patients and healthy controls. This dataset, which was chosen for its comprehensive coverage and validated data, forms a robust foundation for developing our diagnostic model. By integrating metabolomic profiles with advanced ML algorithms, we aimed to identify key metabolic biomarkers associated with OSCC. This approach bridges the gap between molecular data and clinical applicability, ensuring scientifically rigorous and clinically relevant findings, ultimately contributing to the development of reliable, noninvasive diagnostic tools for early OSCC detection.

2 Methods

2.1 Chou's 5-step rule

The methods section of this study is organized essentially by following Chou's 5-step rule (15), outlined as follows: 1) Build a benchmark dataset: We utilized a publicly available metabolomics dataset that includes plasma samples from 61 OSCC patients and 61 HC. The dataset comprises 131 numerical features and 104 ratio features, ensuring a comprehensive foundation for training and testing the predictor. 2) Dataset Representation: To effectively represent the dataset, we employed the extra trees (ET) classifier for feature selection. These selected features were then standardized to ensure uniformity across the dataset. This step includes data preprocessing such as cleaning, handling missing values, and normalizing data to prepare it for feature selection. 3) Introducing a powerful algorithm: We evaluate multiple machine learning models, including support vector machines (SVMs), random forests (RFs), neural networks (NNs), XGBoost, TabNet,

logistic regression (LR), and TabPFN. Each model's parameters were initially tuned via Bayesian optimization to maximize accuracy, except for the TabPFN model, which does not require hyperparameter tuning and is straightforward to use. 4) Statistical analysis: To evaluate the prediction accuracy, we performed statistical analysis via cross-validation methods. For the ET model, out-of-bag (OOB) estimates were used to observe the accuracy of the top-ranked features. This step also incorporates model evaluation, where metrics like accuracy, precision, recall, F1 score and the area under the ROC curve (AUC) are used to assess the performance of the predictive models. 5) The development of a user-friendly webserver for the predictor is left for future work. This future development aims to provide a practical tool for clinicians and researchers, enhancing the clinical applicability and impact of our findings.

2.2 Public dataset collection

2.2.1 Data acquisition and preparation

We utilized a publicly available metabolomics dataset, initially detailed in the study "Plasma metabolomics of oral squamous cell carcinomas on the basis of NMR and MS approaches provides biomarker identification and survival prediction" published in *Scientific Reports* (4). Although the original dataset includes both NMR and MS analyses, for this study, we focused solely on the MS data for our analysis, as it aligns better with the objectives of our research. This dataset includes plasma samples from 61 OSCC patients and 61 healthy controls, which were meticulously collected by four institutions: the Faculty of Medicine at the University of São Paulo, Heliopolis Hospital, Arnaldo Vieira de Carvalho Cancer Institute, and Barretos Cancer Hospital. These samples were sourced from diverse demographics within São Paulo State, Brazil, with the OSCC patients having not received any prior radiotherapy or chemotherapy to ensure unaltered metabolic profiles. The dataset specifics are cataloged in Supplementary Table S1.

2.3 Mass spectrometry analysis

The metabolic profiling of the dataset was conducted via the AbsoluteIDQ® p180 Kit by BIOCRATES Life Sciences AG, Innsbruck, Austria. This comprehensive platform facilitates the quantification of up to 188 distinct metabolites spanning various classes, such as 21 amino acids, 21 biogenic amines, one hexose (total hexose), 40 acylcarnitines, 90 glycerophospholipids (including 76 phosphatidylcholines and 14 lysophosphatidylcholines), and 15 sphingolipids along with their derivatives. For detailed categorization, metabolites are systematically labeled on the basis of chain length and type of linkage—e.g., Cx:y, where 'x' denotes the number of carbon atoms and 'y' denotes double bonds in the lipid side chains (4).

To ensure high precision in metabolite quantification, sample derivatization was performed using phenyl isothiocyanate (PITC) with internal standards. Subsequent analyses employed flow injection analysis-tandem mass spectrometry (FIA-MS/MS) for acylcarnitines, lipids, and hexoses and liquid chromatography-mass spectrometry (LC-MS/MS) for amino acids and biogenic amines. These procedures were executed via advanced mass spectrometry equipment, namely, the SCIEX 4000 QTrap® and Waters XEVO TQMS® systems with electrospray ionization. The specific methodologies are detailed in the patent US 2007/0004044. To increase data reliability, only metabolites above the detection threshold and with identifiable peaks were considered (16), as detailed in the analysis of samples in Supplementary Table S2.

2.4 Machine learning analysis

The machine learning analysis for this study followed a structured approach encompassing data preprocessing, model construction, model optimization, feature selection and model evaluation. This comprehensive approach ensures the reliability and accuracy of the models used to diagnose OSCC. The overall process is illustrated in Figure 1.

2.4.1 Dataset preprocessing

Data preprocessing was a crucial step in preparing the dataset for effective modeling. The preprocessing involved several steps. Initially, irrelevant variables and extreme outliers, defined as values beyond the mean ± 2 standard deviations, were removed (17). Missing values were handled by imputing with median values or removing features with substantial missing data (18). The features were then standardized so that each had a mean of zero and a standard deviation of one, ensuring consistency throughout the dataset. After preprocessing, the dataset consisted of 253 characteristic values, including 131 numerical features and 104 ratio features. The labels were divided into two categories: OSCC (61 samples) and HC (61 samples).

2.4.2 Model construction and optimization

To distinguish OSCC patients from HC via plasma metabolite profiles, multiple machine learning algorithms have been evaluated. The dataset was split into training (70%), validation (10%), and testing (20%) sets. The models assessed included support vector machines (SVMs), extra trees (ET), XGBoost, TabNet, logistic regression (LR), TabPFN, multilayer perceptron (MLP) and voting method. Bayesian optimization leverages Bayes' theorem to guide the search for optimal solutions by using prior knowledge from previous iterations. It avoids poor-performing areas and focuses on regions with better results, improving the efficiency of finding the optimal solution. Thus Bayesian optimization was employed to fine-tune the hyperparameters of most models, focusing on optimizing validation accuracy (ACC) (19). Unlike other models, the TabPFN does not require hyperparameter tuning, offering a straightforward implementation (20). This optimization method systematically explores the hyperparameter space, using a probabilistic approach to identify the best configuration for each model.

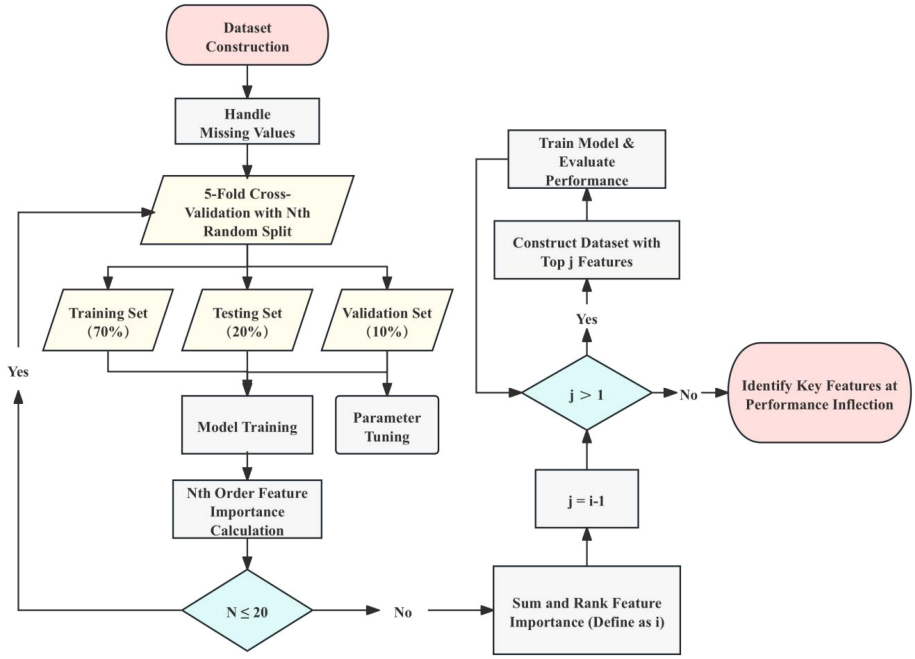


FIGURE 1
Workflow for dataset construction and model training. This figure outlines the workflow for constructing the dataset and training the machine learning models. The process starts with dataset preparation, handling missing values, and performing 5-fold cross-validation with random splits of the dataset. This cross-validation process is repeated 20 times, totaling 100 model training and validation iterations. For each iteration, models are trained with the top n important features, and the change in accuracy (ACC) is monitored to identify the inflection point, representing the most critical features for classification. The importance of each feature is determined by summing the feature importance scores calculated over 100 iterations using the Extra Trees (ET) model. The training set (70%) and validation set (10%) are used for parameter tuning and feature ranking, while the test set (20%) is reserved for final performance evaluation. The process iteratively narrows down the feature set until the top-ranked features are determined based on performance stabilization at the inflection point.

2.4.3 Feature selection

Feature selection was performed via the ET algorithm (21), which was chosen for its effectiveness in handling high-dimensional data and robustness in identifying the most informative features. The ET algorithm constructs multiple decision trees with random splits at each node, increasing the variance among trees and reducing overfitting. This method offers several advantages: 1) Handling high-dimensional data: ET is particularly effective in datasets with a large number of features, reducing dimensionality while retaining significant predictive power. 2) Robustness: By averaging over many trees, ET reduces the variance of the model, making it less sensitive to noise in the training data. 3) Feature importance evaluation: The ET algorithm evaluates the importance of each feature on the basis of the mean decrease in impurity, which measures the effectiveness of a feature in reducing uncertainty in predictions (22). The importance scores derived from the ET algorithm were used to identify the most significant features.

The ET model parameters were fine-tuned via Bayesian optimization, and the accuracy of the top-ranked features was observed via out-of-bag (OOB) estimates (23). This combined approach ensures that the most relevant features are selected and that the model parameters are optimized for the best performance.

2.4.4 Model evaluation

Following feature selection, the identified significant features were used to train and evaluate the previously selected best models, ensuring that the models were built using the most informative and relevant data. Model performance was assessed via metrics such as accuracy, precision, recall, F1 score and ROC/AUC (24). The specific formula is as follows:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$AUC = \int_0^1 TPR\left(FPR^{-1}(x)\right)dx$$

where TP is the number of true positive cases, TN represents the number of true negative cases, FP represents the number of false positive cases, and FN represents the number of false negative cases. where $FPR^{-1}(x)$ denotes the inverse function of FPR with respect to x. This integral essentially computes the area under the ROC curve, which plots the TPR against the FPR as the discrimination threshold varies. The ROC curve illustrates the diagnostic ability of the classifier

system by plotting the true positive rate against the false positive rate at various thresholds. The AUC measures the overall performance, with values closer to 1 indicating better model discrimination capability. These metrics provide a comprehensive evaluation of model performance, ensuring that the selected model not only achieves high accuracy but also maintains a balance between precision and recall, which is crucial for effective OSCC diagnosis.

In this study, we used Python 3.8 and several Python packages, including sklearn, xgboost, pytorch, pytorch_tabnet, matplotlib, and TabPFN, to implement and evaluate our machine learning models. The entire process was run on a system with an AMD Ryzen 7 5800H CPU and an NVIDIA GeForce RTX 3070 Laptop GPU. These tools and hardware allowed for efficient training and optimization of the models. The details and algorithms of the machine learning models can be found in their respective documentation and publications.

In all the statistical P value calculations, the significance levels are indicated as follows: *P ≤ 0.05 (significant), **P ≤ 0.01 (highly significant), and ***P ≤ 0.001 (extremely significant).

3 Results

The machine learning analysis for this study followed a structured approach encompassing data preprocessing, model construction, feature selection, model optimization, parameter tuning and evaluation. This comprehensive approach ensures the reliability and accuracy of the models used to diagnose OSCC.

3.1 Modeling performance and comparisons

Figure 2 illustrates the Bayesian optimization procedure used to fine-tune the parameters of various machine learning models and the before and afterwards accuracy of all models. The Bayesian optimization procedure (Figure 2A) fine-tunes parameters such as the number of estimators, and maximum depth (25) as detailed in Supplementary Table S3. This process highlights the importance of parameter optimization in improving model performance, as evidenced by the notable differences in accuracy scores.

The comparative accuracy of different models after Bayesian optimization is shown in Figures 2B, C, with the TabPFN model outperforming others without optimization in terms of accuracy, precision, recall, and the F1 score. The detailed performance metrics and the differences before and after optimization are presented in Supplementary Tables S4 and S5. Both before and after hyperparameter tuning, the TabPFN and ET models consistently performed well, indicating their effectiveness (20). Although TabNet is highly dependent on hyperparameters, Bayesian optimization has significantly improved its performance (26). The results indicate that the TabPFN model achieved an accuracy of

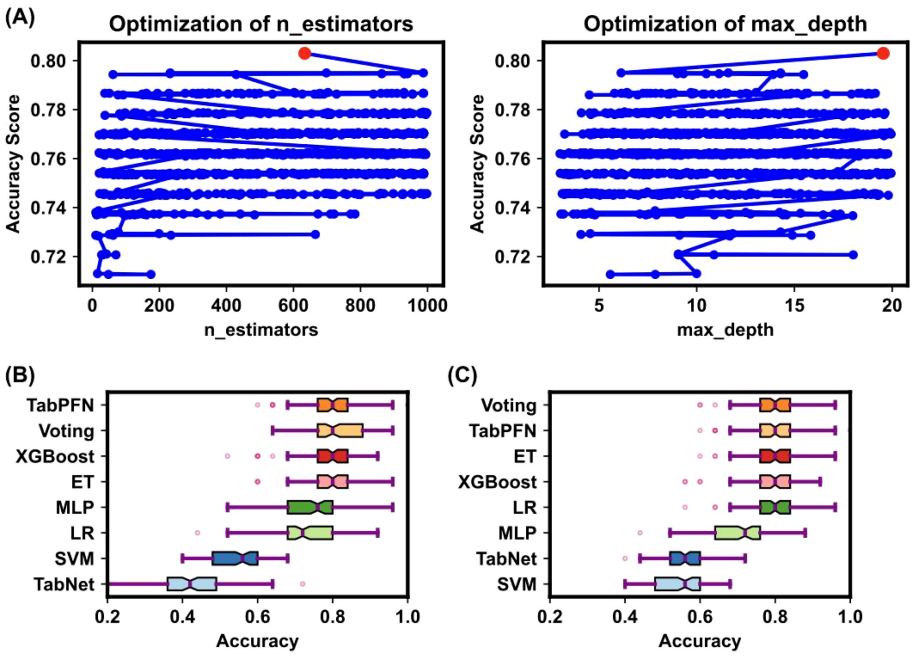


FIGURE 2
Bayesian optimization procedure for model parameters. **(A)** The Bayesian optimization procedure for tuning parameters such as the number of estimators, and maximum depth. The x-axis represents parameter values, and the y-axis represents accuracy changes. Blue dots indicate parameter values attempted by Bayesian optimization, and the red dot indicates the optimal parameter value. This process highlighted the importance of parameter optimization in improving model performance. **(B, C)** Box plots of model accuracy before and after Bayesian optimization. The box's central red line represents the median, the outer red lines represent the maximum and minimum values, and the box edges represent the first and third quartiles. Outliers are shown as individual points around the box. **(B)** shows the accuracy before parameter tuning, while **(C)** shows the accuracy after parameter tuning. The comparison demonstrates that TabPFN outperformed others in terms of accuracy.

80% with a comparatively short running time in distinguishing OSCC and premalignant lesions from healthy conditions on a person-by-person basis. The selected plasma metabolites are significantly dysregulated in OSCC patients, highlighting their potential as biomarkers for early diagnosis (27). Unlike traditional supervised learning methods, the TabPFN is a single transformer model pretrained on a large amount of generated data, making it particularly suitable for small-sample table classification (20). This model can approximate the calculation of a posterior prediction distribution (PPD) on the basis of the likelihood of given data and prior probability, providing a generic model applicable to various small tabular classification tasks without retraining or model selection (20). The breakthrough of this method lies in its ability to quickly and accurately solve small table classification problems (20).

The results demonstrate that the integration of metabolomics analysis with advanced machine learning techniques, particularly the TabPFN model, provides a powerful tool for the early detection and clinical management of OSCC. This combined approach offers high accuracy and reliability, underscoring the potential for practical implementation in clinical settings (20).

3.2 Important features

Feature selection via the ET algorithm identified 29 numerical features and 47 ratio features as crucial for the model's predictive power. The importance of these features is depicted in Figure 3, which shows the value of feature importance for each selected feature. These features were identified using the ET algorithm's feature importance calculation, where the importance of a feature increases each time it is used to effectively split the data and improve purity. Subsequently, significance analysis was performed on the selected important features to further validate their impact on the model's classification accuracy.

The trend of accuracy changes when models are built using different numbers of top-ranked important features is illustrated in Figure 3A. The figure shows the optimization process, highlighting an inflection point where 76 features yielded the highest accuracy (ACC = 0.8057). This inflection point, calculated using the kneed algorithm (28), indicates the optimal number of features needed to achieve the best model performance without overfitting.

The specific important features identified are detailed in Tables 1 and 2, which shows a mix of individual metabolites and metabolite ratios (Figure 3B), and the top 13 features are all ratios (standard deviation, mean, and significance information for all the features are detailed in Supplementary Tables S6, S7). These features play critical roles in differentiating OSCC patients from HC, underscoring their potential as biomarkers for early diagnosis (29). The important features include various sphingomyelins (SMs), phosphatidylcholines (PCs), and amino acid ratios, each contributing uniquely to the model's predictive ability (4).

A heatmap of the Pearson correlation coefficients for the top-ranked features is shown in Figure 3C. This heatmap illustrates the correlation between each pair of selected features, highlighting the relationships and dependencies among them. High correlation coefficients indicate strong relationships, which can provide insights into the underlying metabolic pathways affected in OSCC (14, 30). In the analysis of all metabolite correlations, the top 10 feature pairs exhibit strong positive correlations (0.89 to 0.9995), for example C2/C0 and (C2+C3)/C0 (0.9995), indicating substantial redundancy between these features (Supplementary Table S8). In practice, when two features are highly correlated, detecting both may not add significant value to the diagnostic model, as they convey similar information about the underlying metabolic changes.

The significance of these selected features is further validated by their impact on model performance metrics. As shown in Figure 4, the evaluation of the TabPFN model with all features versus only the important features demonstrated significant improvements in accuracy (0.851 ± 0.066), precision (0.858 ± 0.065), recall (0.851 ± 0.066), and F1 score (0.85 ± 0.067) when the important features were SMs such as used (Figure 4A). This evaluation underscores the efficacy of the feature selection process in enhancing model performance.

The ROC curve and AUC value for the TabPFN model, depicted in Figure 4B, further confirmed the model's high diagnostic capability. The ROC curve shows a high true positive rate against the false positive rate, with an AUC of 0.93, indicating excellent model performance. The performance of the machine learning model was further evaluated via a confusion matrix (Figure 4C). The model correctly identified 83.6% of the HC and 86.6% of the OSCC patients. This finding indicates a high level of accuracy in distinguishing between healthy individuals and those with OSCC, with only a small percentage of misclassifications in each group. This robust classification performance underscores the model's potential for reliable early screening and diagnosis of OSCC on the basis of plasma metabolite profiles.

To assess the predictive power of individual features and the feasibility of using single features for practical screening in clinical settings, all important features were used independently to predict OSCC status. The accuracy of these predictions is presented in Figure 5. Each feature's ability to distinguish between OSCC patients and HC was evaluated, with the highest accuracy observed for the top-ranked feature. These findings demonstrate that even single features can provide substantial predictive power for early OSCC screening (5, 31). These results indicate that the top-ranked feature alone can achieve an accuracy of 76.6%, highlighting its potential for use in rapid early screening of OSCC.

These results highlight the potential of using machine learning models combined with plasma metabolite profiling for accurate and automated diagnosis of OSCC. The integration of these techniques offers a robust and reliable approach for early detection and improved patient outcomes (32). Feature selection via Bayesian-optimized ET classifier and model construction with TabPFN yielded the highest accuracy, demonstrating the suitability of these methods (20, 33). Notably, the top-ranked features, primarily ratios, were found to be particularly useful for rapid early screening.

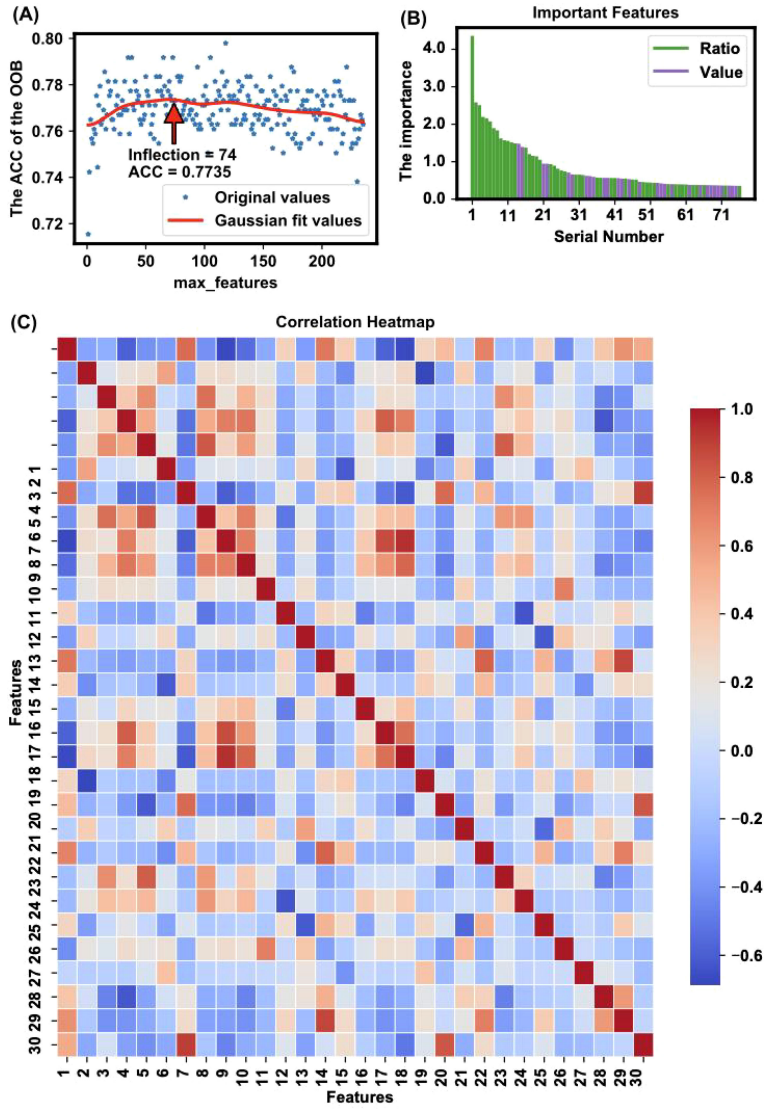


FIGURE 3
Feature selection and importance analysis. **(A)** The trend of accuracy changes when applying different features for modeling is illustrated, highlighting an inflection point where 76 features yielded the highest accuracy (ACC = 0.8057) with the OOB method. The x-axis represents the number of top important features used for modeling, and the y-axis represents the corresponding OOB accuracy. The red line is a Gaussian fit curve indicating the trend. **(B)** The importance of features is depicted, showing the value of feature importance for each selected feature. The identified important features include a mix of individual metabolites and metabolite ratios, which together capture key metabolic changes linked to OSCC. This combination improves the model's ability to distinguish between healthy and cancerous states The feature importance is calculated based on the sum of importance scores from 100 random splits and model trainings. **(C)** The heatmap presents Pearson correlation coefficients for the top 30 features ranked by importance in the model, as listed in Supplementary Table S8. The color intensity indicates the strength of the correlation: red represents a strong positive correlation, blue indicates a strong negative correlation, and white shows little to no correlation. This visualization helps identify relationships and dependencies among the selected features, providing insights into potential interactions that could influence the model's performance.

4 Discussion

Predicting clinical outcomes can significantly optimize diagnostic and treatment strategies for OSCC. In this study, we developed a prediction model using a combination of metabolomic profiles and machine learning techniques. The data were processed and analyzed to identify significant attributes, which were then utilized in a TabPFN model to make predictions (20).

The important step in creating our prediction model involved validating and collecting key metabolic biomarkers associated with OSCC. Through the application of the ET algorithm for feature selection, we identified 29 numerical features and 47 ratio features as crucial for the model's predictive power (Figures 3A, B). These biomarkers provide a detailed metabolic fingerprint of OSCC, highlighting the significant metabolic alterations that occur in this disease (4).

The results show that these selected features correlate strongly with OSCC diagnosis, as validated by high accuracy, precision, recall, and F1 scores across different ML models. The TabPFN model, which leverages pretrained data for small-sample table

10.3389/fimmu.2024.1493377

TABLE 1 List of important features (ratio).

Number	Important Features (ratio)	HC	OSCC	p-value	Significant
1	SM C24:1/(Met/PC aa C40:3)	0.746 ± 0.348	1.352 ± 0.596	0.000	***
2	C3/C4	2.034 ± 0.649	1.412 ± 0.659	0.000	***
3	(Ala/Gln)/Orn	0.008 ± 0.004	0.006 ± 0.003	0.000	***
4	Phe/PC aa C42:4	405.921 ± 131.665	300.782 ± 71.689	0.000	***
5	(Ala/Gln)/(Tyr/Phe)	0.688 ± 0.256	0.48 ± 0.192	0.000	***
6	Val/C5	1610.717 ± 561.589	1121.013 ± 832.575	0.000	***
7	(Tyr/Phe)/(Met/PC aa C40:3)	0.018 ± 0.007	0.031 ± 0.017	0.000	***
8	Ala/Gln	0.706 ± 0.241	0.517 ± 0.135	0.000	***
9	Met/PC aa C40:3	67.062 ± 35.062	43.083 ± 15.703	0.000	***
10	Ala/PC aa C40:2	1892.746 ± 910.595	1269.65 ± 382.683	0.000	***
11	SM (OH) C24:1/SM C16:0	0.013 ± 0.002	0.011 ± 0.003	0.000	***
12	Gln/Thr	5.248 ± 1.299	6.569 ± 1.791	0.000	***
13	Total PC ae/Total SM	0.534 ± 0.089	0.476 ± 0.081	0.000	***
14	Thr/Ser	1.276 ± 0.379	1.035 ± 0.321	0.000	***
15	Phe/PC aa C40:3	170.475 ± 75.462	125.024 ± 34.121	0.000	***
16	Met/PC aa C40:2	119.134 ± 55.558	78.964 ± 26.552	0.000	***
17	C4/C0	0.006 ± 0.002	0.008 ± 0.005	0.001	***
18	(Tyr/Phe)/Ala	0.003 ± 0.001	0.004 ± 0.002	0.000	***
19	(Ala/Gln)/Tyr	0.01 ± 0.004	0.008 ± 0.004	0.000	***
20	Asn/Gln	0.074 ± 0.025	0.063 ± 0.019	0.000	***
21	Total SM/Total Lipids	0.135 ± 0.02	0.148 ± 0.019	0.000	***
22	Total SMOH/Total SM nonOH	0.16 ± 0.022	0.142 ± 0.027	0.000	***
23	C4/C5	1.339 ± 0.524	1.173 ± 1.101	0.005	**
24	(Tyr/Phe)/Met	0.042 ± 0.013	0.063 ± 0.04	0.001	**
25	Total acylcarnitines/C0	0.169 ± 0.046	0.225 ± 0.098	0.001	***
26	Glutaminolysis: (Ala+Asp+Glu)/Gln	0.86 ± 0.307	0.716 ± 0.383	0.000	***
27	PC_ae_C32:1/PC_ae_C34:1	0.297 ± 0.041	0.284 ± 0.049	0.125	
28	Pro/Orn	2.68 ± 0.907	2.235 ± 0.753	0.017	*
29	PUFA(PC)/MUFA(PC)	5.621 ± 0.883	5.298 ± 1.193	0.211	
30	Ala/lysoPC a C18:1	37.608 ± 22.76	26.298 ± 10.968	0.000	***
31	PC_aa_C40:3/PC_aa_C42:5	1.451 ± 0.301	1.606 ± 0.288	0.005	**
32	Leu/Gln	0.259 ± 0.087	0.211 ± 0.063	0.004	**
33	PC ae C44:5/PC ae C42:5	0.798 ± 0.169	0.821 ± 0.132	0.074	
34	CPT1: (C16+C18)/C0	0.003 ± 0.001	0.003 ± 0.001	0.041	*
35	Met/lysoPC a C18:1	2.41 ± 1.777	1.642 ± 0.816	0.000	***
36	(Asn/Asp)/Glu	0.279 ± 0.379	0.176 ± 0.253	0.103	
37	CPT2: (C16+C18.1)/C2	0.031 ± 0.008	0.031 ± 0.012	0.579	
38	Met-SO/Met	0.025 ± 0.017	0.043 ± 0.048	0.000	***
39	Asn/Orn	0.537 ± 0.164	0.462 ± 0.159	0.017	*

(Continued)

TABLE 1 Continued

Number	Important Features (ratio)	HC	OSCC	p-value	Significant
40	(Glnlysis)/(Asp/Gln)	180.706 ± 189.567	126.211 ± 154.392	0.090	
41	lysoPC_a_C20:4/lysoPC_a_C20:3	2.856 ± 0.876	2.837 ± 1.38	0.214	
42	SDMA/Arg	0.006 ± 0.004	0.008 ± 0.006	0.081	
43	Total lyso(PC)/Total(PC)	0.107 ± 0.027	0.104 ± 0.031	0.558	
44	(Ala/Gln)/Ile	0.008 ± 0.003	0.007 ± 0.004	0.004	**
45	MUFA/SFA	10.77 ± 1.646	11.309 ± 2.812	0.669	
46	C2/C0	0.139 ± 0.044	0.182 ± 0.093	0.016	*
47	(C2+C3)/C0	0.149 ± 0.044	0.191 ± 0.092	0.022	*

This table lists the important ratio features identified through feature selection, including their mean values in healthy controls (HC) and OSCC patients, and the significance of the difference between the two groups. p-values are represented as follows: $p < 0.05$ *, $p < 0.01$ **, and $p < 0.001$ ***.

TABLE 2 List of important features (value).

Number	Important Features (value)	HC	OSCC	p-value	Significant
1	SM C24:1	41.628 ± 7.955	51.259 ± 12.199	0.000	***
2	C5	0.201 ± 0.18	0.499 ± 0.647	0.000	***
3	PC aa C36:6	0.624 ± 0.225	0.471 ± 0.205	0.000	***
4	SM C26:1	0.285 ± 0.079	0.361 ± 0.114	0.000	***
5	PC aa C42:4	0.181 ± 0.045	0.215 ± 0.053	0.000	***
6	SM C16:0	103.144 ± 18.64	121.256 ± 27.081	0.000	***
7	PC aa C36:0	1.584 ± 0.474	1.336 ± 0.521	0.031	*
8	C4	0.228 ± 0.09	0.314 ± 0.211	0.008	**
9	C14:1	0.091 ± 0.021	0.116 ± 0.055	0.002	**
10	PC aa C36:5	16.314 ± 6.622	13.535 ± 6.794	0.006	**
11	C3	0.429 ± 0.137	0.358 ± 0.147	0.003	**
12	PC ae C44:3	0.093 ± 0.021	0.108 ± 0.03	0.010	**
13	PC ae C38:0	1.428 ± 0.394	1.234 ± 0.392	0.012	*
14	C14:2	0.025 ± 0.011	0.037 ± 0.025	0.000	***
15	PC aa C34:4	1.644 ± 0.738	1.31 ± 0.554	0.007	**
16	PC aa C40:2	0.244 ± 0.057	0.283 ± 0.065	0.003	**
17	Ala	431.684 ± 125.636	347.628 ± 94.083	0.000	***
18	SDMA	0.484 ± 0.353	0.549 ± 0.333	0.066	
19	PC aa C32:2	3.13 ± 1.303	2.663 ± 1.183	0.034	*
20	Ser	101.627 ± 23.284	113.342 ± 40.304	0.191	
21	Gln	632.779 ± 142.604	687.898 ± 165.148	0.031	*
22	C0	41.368 ± 9.438	39.007 ± 10.676	0.123	
23	PC ae C38:6	5.379 ± 1.328	4.966 ± 1.4	0.121	
24	PC aa C42:1	0.245 ± 0.088	0.285 ± 0.083	0.004	**
25	PC aa C40:3	0.44 ± 0.11	0.518 ± 0.122	0.001	***
26	Lys	246.288 ± 63.078	212.745 ± 56.475	0.003	**

(Continued)

TABLE 2 Continued

Number	Important Features (value)	HC	OSCC	p-value	Significant
27	PC ae C44:4	0.273 ± 0.077	0.328 ± 0.104	0.002	**
28	Met	27.623 ± 10.139	21.974 ± 8.371	0.002	**
29	PC aa C38:6	51.035 ± 12.787	46.733 ± 14.33	0.110	

This table lists the important metabolic value features identified through feature selection,including their mean values in healthy controls (HC) and OSCC patients, and the significance of the difference between the two groups. p-values are represented as follows: $p < 0.05$ *, $p < 0.01$ **, and $p < 0.001$ ***.

classification, demonstrated superior performance in distinguishing OSCC patients from HC (Figure 4). This approach underscores the model's ability to handle complex, high-dimensional and small sample data efficiently (20).

Among the identified metabolites (Table 2), SMs such as SM C24:1 and SM C16:0 play crucial roles in cell membrane integrity and signaling pathways that regulate cell proliferation and apoptosis (34). Altered levels of these sphingomyelins have been linked to cancer cell survival and resistance to apoptosis, which are characteristic of OSCC progression (31, 35). PCs, like PC aa C36:6 and PC aa C42:4, are also central to membrane structure and cellular signaling, with abnormal PC metabolism being a common feature in many cancers (36). Both SMs and PCs are primarily formed via the Kennedy pathway, which is significant for OSCC progression and can be targeted for therapeutic interventions (37). Elevated choline kinase activity, crucial for PC synthesis, has been linked to poor prognosis and could play a similar role in OSCC (38). The consistent detection of SM and PC features among the top-ranked markers underscores their relevance as both potential biomarkers and therapeutic targets in OSCC.

Within the candidate biomarkers identified in this study, short-chain acylcarnitine (ACar) like C3/C4, C4, and C5, along with medium-chain ACars such as C8 and C10, demonstrate significant potential for OSCC diagnosis (39). These ACars are crucial intermediates in fatty acid oxidation (FAO), a metabolic pathway reprogrammed in OSCC cells to meet the high energy demands and adapt to the harsh tumor microenvironment characterized by hypoxia and acidosis (40). The observed upregulation of short-

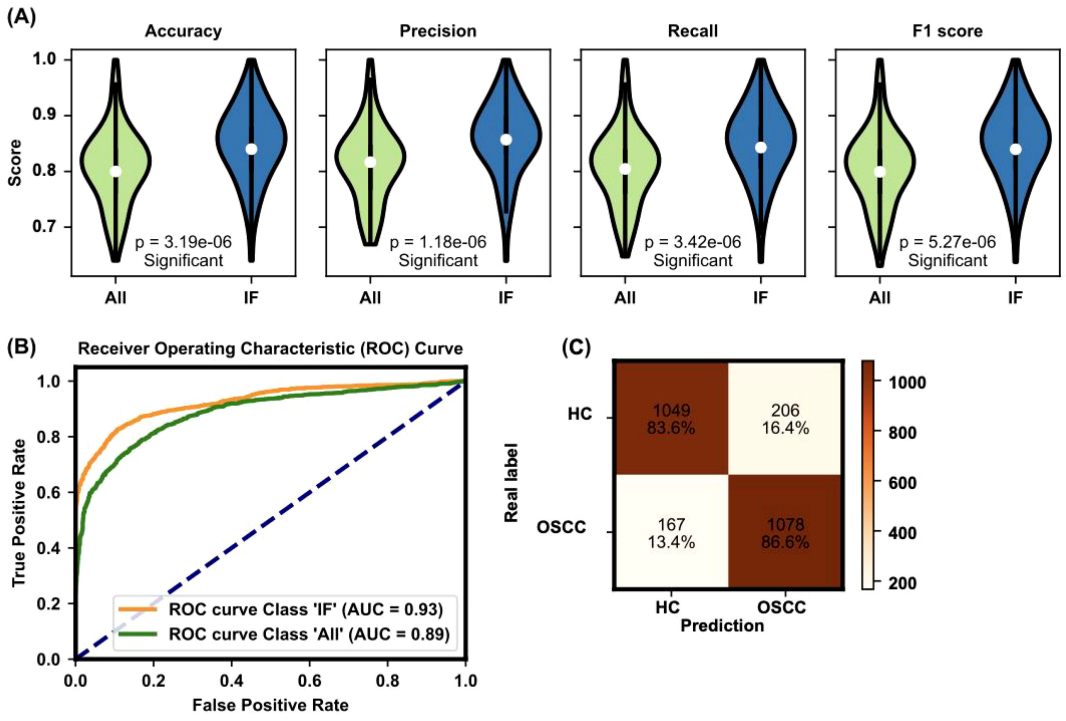


FIGURE 4
Evaluation of model performance with all features and important features. **(A)** The evaluation of the TabPFN model with all features versus only the important features demonstrates significant improvements in accuracy (0.851 ± 0.066), precision (0.858 ± 0.065), recall (0.851 ± 0.066), and F1 score (0.851 ± 0.067) when the important features are used. The figure presents violin plots with embedded box plots. The box plots' central red line represents the median, with the edges of the box denoting the first and third quartiles, and whiskers extending to the minimum and maximum values. Outliers are shown as individual points. **(B)** The ROC curve and AUC value for the TabPFN model further confirm the model's high diagnostic capability, with an AUC of 0.93 indicating excellent performance. The violin plot, a variant of the box plot, shows the density of accuracy values, highlighting the distribution of accuracy scores. **(C)** Confusion matrix analysis of model predictions: The confusion matrix compares real labels (HC for HC and OSCC for oral squamous cell carcinoma patients) against predicted labels. The numbers represent the count and percentage of correctly and incorrectly classified samples in each category. The model accurately classified 83.6% of HC and 86.6% of OSCC patients.

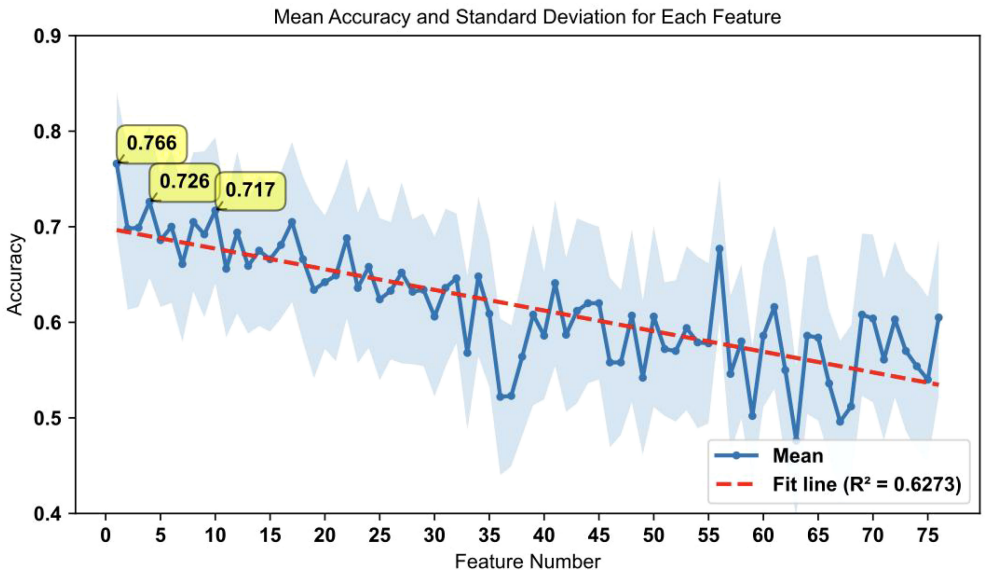


FIGURE 5
Accuracy performance for individual features ranked by importance. This figure shows the accuracy results for individual feature predictions, with the accuracy trend generally following the feature importance ranking. The shaded area represents the standard deviation of accuracy for each feature. The top three features with the highest accuracy are labeled on the graph (ACC= 0.766, 0.698, 0.699, respectively). The red dashed line represents the fit line with an R^2 value of 0.6273, indicating the overall trend.

chain ACars suggests an increased reliance on FAO for energy production, while the downregulation of medium-chain ACars may indicate selective consumption by OSCC cells. Such metabolic alterations not only reflect the underlying pathophysiology of OSCC but also highlight the value of ACars as biomarkers (41). Their ease of detection in plasma makes them particularly suitable for non-invasive early screening, offering a promising avenue for early diagnosis and personalized intervention in OSCC.

The key biomarker ratios, (Ala/Gln)/Orn, Phe/PC aa C42:4, and (Ala/Gln)/(Tyr/Phe), reflect significant metabolic reprogramming in OSCC. The (Ala/Gln)/Orn ratio highlights disruptions in nitrogen metabolism, as glutamine and alanine are crucial for tumor growth, while ornithine links to altered urea cycle activity (42). The Phe/PC aa C42:4 ratio connects amino acid metabolism with lipid synthesis, underscoring the interplay between phenylalanine uptake and phosphatidylcholine pathways, both critical in cancer progression (43). Meanwhile, (Ala/Gln)/(Tyr/Phe) captures the balance of nitrogen and aromatic amino acid metabolism, further emphasizing OSCC's reliance on reprogrammed amino acid pathways (43). These ratios offer potential as diagnostic markers and therapeutic targets in OSCC. These ratios are indicative of the extensive metabolic reprogramming that occurs in cancer cells to support their rapid growth and proliferation (15).

Furthermore, our analysis of the top-ranked individual features demonstrated substantial predictive power even when used independently, and the highest accuracy achieved with a single feature (SM C24:1/(Met/PC aa C40:3)) was 76.6% (Figure 5). These findings indicate the potential for the use of top-ranked features in rapid screening protocols for OSCC. The ability of individual biomarkers to predict disease status underscores their importance and utility in clinical settings (44).

The correlation analysis reinforces the feasibility of simplifying OSCC diagnostic protocols by focusing on individual metabolites (Figure 3C). The strong correlations observed among the top 10 feature pairs (Supplementary Table S8), ranging from 0.89 to 0.9995, suggest redundancy, where detecting a single feature in each pair could be sufficient for accurate diagnosis. For instance, the near-perfect correlation between C2/C0 and (C2+C3)/C0 (0.9995) implies that either could be selected based on practical considerations, such as ease of detection. Similarly, highly correlated pairs like Total acylcarnitines/C0 with (C2+C3)/C0 (0.9834), and C4/C0 with C4 (0.9061), indicate that prioritizing the more detectable metabolite is a viable strategy.

Despite the promising performance of our machine learning model, which achieved an accuracy of 85% (Figure 4), several factors may have contributed to it not reaching 100%. One significant factor is the inherent biological variability among patients (44, 45). Variations in age, sex, ethnicity, diet, and lifestyle can influence metabolic profiles, potentially introducing noise into the data and affecting the model's ability to generalize across diverse populations (44). Additionally, metabolic alterations due to factors other than cancer, such as chronic diseases or medication, can confound the data and reduce predictive accuracy (44, 45). The list of biomarkers used in our model, although comprehensive, may still be incomplete. The molecular mechanisms underlying OSCC are complex and not fully understood. There may be other relevant metabolites or metabolic pathways that were not included in our analysis, potentially limiting the model's ability to capture all aspects of the disease (45). Recognizing and accounting for these confounding factors can increase the accuracy and reliability of metabolomic studies and associated diagnostic models for OSCC.

5 Limitations

Although our machine learning model demonstrated significant predictive power via plasma metabolite profiles, several limitations need to be addressed. First, the list of biomarkers identified and used in this study may be incomplete because the molecular mechanisms underlying OSCC are complex and not fully understood. The biomarkers included in our model may not encompass all relevant metabolic changes associated with OSCC, potentially affecting the accuracy and generalizability of the model (27). Additionally, the variability in drug dosage, treatment regimens, and patient responses in clinical settings could influence metabolic profiles and their diagnostic utility, which our model does not account for (29).

Furthermore, our analysis was based on publicly available datasets, which may not fully represent the diverse populations affected by OSCC (29). Access to more extensive and diverse datasets, including patient-level data, would likely enhance the model's predictive capability and robustness (29). The reliance on a single dataset and the exclusion of patients who had undergone radiotherapy or chemotherapy to ensure unaltered metabolic profiles may limit the applicability of our findings to the broader OSCC patient population (27). Future studies should aim to validate these findings across multiple datasets and consider the inclusion of treated patients to better understand the impact of various treatments on metabolic profiles.

6 Conclusion

This study highlights the effectiveness of integrating advanced machine learning techniques with plasma metabolomics for the early diagnosis of OSCC. By leveraging biomarkers identified through metabolomic profiling and applying sophisticated algorithms such as TabPFN, we achieved high diagnostic accuracy, underscoring the potential of this approach for precision medicine. Additionally, we explored the potential of using individual features for early screening, with the advantage of avoiding accuracy inflation through multiple k-fold cross-validations. The results demonstrate that combining multiple disease features, including specific metabolite levels and ratios, significantly enhances the predictive power of the models (27).

Future research should incorporate multi-omics data, such as proteomics and transcriptomics, to enrich biomarker discovery and explore the immune landscape associated with OSCC (29). Integrating these multi-omics approaches with immunotherapy-related biomarkers could offer novel insights into personalized therapeutic strategies. Additionally, expanding patient-level data across diverse cohorts and developing a publicly accessible web platform for interactive biomarker analysis could enhance clinical utility. Such a platform could enable personalized diagnostics and immune-based treatment planning, ultimately improving patient outcomes in OSCC.

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

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. 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. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

YS: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Writing – original draft. GC: Data curation, Investigation, Methodology, Visualization, Writing – review & editing. DW: Writing – review & editing. JCL: Visualization, Writing – review & editing. JNL: Conceptualization, Funding acquisition, Methodology, 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 project was supported by the National Natural Science Foundation of China Outstanding Youth Fund (Grant No. 62322114) and the Medical Engineering Cross Foundation of Shanghai Jiao Tong University (Grant No. YG2023LC06).

Acknowledgments

We extend our sincere thanks to Prof. Eloiza H Tajara and her research team for developing the publicly available plasma metabolomics dataset titled "*Plasma metabolomics of oral squamous cell carcinomas based on NMR and MS approaches provides biomarker identification and survival prediction.*" This dataset was fundamental to our study and enabled our exploration of biomarkers in OSCC diagnosis. We also express our appreciation for their commitment to open science, as this work is licensed under the Creative Commons Attribution 4.0 International License. Additionally, we thank ChatGPT for its support in English language editing.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1493377/full#supplementary-material

SUPPLEMENTARY TABLE 1
Clinicopathological data of OSCC patients and HC. Most patients are male, over 40 years, current smokers and alcoholics, with large tumors from the tongue (C02) and floor of mouth (C04) subsites, often with nodal metastases.

SUPPLEMENTARY TABLE 2
MS analysis of all samples. Concentration (uM) of 131 metabolites and 104 metabolite ratios/sums in plasma samples from 61 OSCC patients and 61 controls.

SUPPLEMENTARY TABLE 3
Bayesian optimization parameter list. This table lists the parameters used in Bayesian optimization for various machine learning models, including their start values, end values, and optimal values.

SUPPLEMENTARY TABLE 4
Performance metrics for all machine learning models before Bayesian optimization. This table presents the accuracy, precision, recall, and F1 scores for various machine learning models, along with their running times, measured over 100 runs (20 iterations of 5-fold cross-validation) before Bayesian optimization.

SUPPLEMENTARY TABLE 5
Differences in model performance metrics before and after Bayesian optimization. This table displays the differences in model performance metrics before and after Bayesian optimization for various machine learning models. Metrics include Accuracy, Precision, Recall, and F1 scores. Statistical significance is denoted as follows: $p * \leq 0.05$, $p ** \leq 0.01$, $p *** \leq 0.001$.

SUPPLEMENTARY TABLE 6
Standard deviation, mean, and significance information for ratio features. This table provides the standard deviation, mean values, and significance information for the ratio features in HC and OSCC patients.

SUPPLEMENTARY TABLE 7
Standard deviation, mean, and significance information for value features. This table provides the standard deviation, mean values, and significance information for the numerical features in HC and OSCC patients.

SUPPLEMENTARY TABLE 8
Correlation matrix of important metabolic features used in classification. The table presents the Pearson correlation coefficients among the important features identified in the study.

References

1. Radaic A, Kamarajan P, Cho A, Wang S, Hung GC, Najarzadegan F, et al. Biological biomarkers of oral cancer. *Periodontol 2000*. (2023) 96:250–80. doi: 10.1111/prd.12542

2. Hasegawa T, Yatagai N, Furukawa T, Wakui E, Saito I, Takeda D, et al. The prospective evaluation and risk factors of dysphagia after surgery in patients with oral cancer. *J Otolaryngol Head Neck Surg*. (2021) 50:4. doi: 10.1186/s40463-020-00479-6

3. Faedo RR, Da SG, Da SR, Ushida TR, Da SR, Lacchini R, et al. Sphingolipids signature in plasma and tissue as diagnostic and prognostic tools in oral squamous cell carcinoma. *Biochim Biophys Acta Mol Cell Biol Lipids*. (2022) 1867:159057. doi: 10.1016/j.bbalip.2021.159057

4. Polachini GM, de Castro TB, Smarra L, Henrique T, de Paula C, Severino P, et al. Plasma metabolomics of oral squamous cell carcinomas based on NMR and MS approaches provides biomarker identification and survival prediction. *Sci Rep*. (2023) 13:8588. doi: 10.1038/s41598-023-34808-2

5. Wang S, Yang M, Li R, Bai J. Current advances in noninvasive methods for the diagnosis of oral squamous cell carcinoma: a review. *Eur J Med Res*. (2023) 28:53. doi: 10.1186/s40001-022-00916-4

6. Balonov I, Mattis M, Jarmusch S, Koletzko B, Heinrich K, Neumann J, et al. Metabolomic profiling of upper GI Malignancies in blood and tissue: a systematic review and meta-analysis. *J Cancer Res Clin Oncol*. (2024) 150:331. doi: 10.1007/s00432-024-05857-5

7. An R, Yu H, Wang Y, Lu J, Gao Y, Xie X, et al. Integrative analysis of plasma metabolomics and proteomics reveals the metabolic landscape of breast cancer. *Cancer Metab*. (2022) 10:13. doi: 10.1186/s40170-022-00289-6

8. Pekarek L, Garrido-Gil MJ, Sanchez-Cendra A, Cassinello J, Pekarek T, Fraile-Martinez O, et al. Emerging histological and serological biomarkers in oral squamous cell carcinoma: Applications in diagnosis, prognosis evaluation and personalized therapeutics (Review). *Oncol Rep*. (2023) 50(6):213. doi: 10.3892/or.2023.8650

9. Graf EH, Pancholi P. Appropriate use and future directions of molecular diagnostic testing. *Curr Infect Dis Rep*. (2020) 22:5. doi: 10.1007/s11908-020-0714-5

10. Kurzrock R, Chaudhuri AA, Feller-Kopman D, Florez N, Gorden J, Wistuba II. Healthcare disparities, screening, and molecular testing in the changing landscape of non-small cell lung cancer in the United States: a review. *Cancer Metastasis Rev*. (2024) 43:1217–31. doi: 10.1007/s10555-024-10187-6

11. Zhang XW, Li QH, Xu ZD, Dou JJ. Mass spectrometry-based metabolomics in health and medical science: a systematic review. *Rsc Adv*. (2020) 10:3092–104. doi: 10.1039/c9ra08985c

12. Zhang L, Ma F, Qi A, Liu L, Zhang J, Xu S, et al. Integration of ultra-high-pressure liquid chromatography-tandem mass spectrometry with machine learning for identifying fatty acid metabolite biomarkers of ischemic stroke. *Chem Commun (Camb)*. (2020) 56:6656–59. doi: 10.1039/d0cc02329a

13. Galal A, Talal M, Moustafa A. Applications of machine learning in metabolomics: Disease modeling and classification. *Front Genet*. (2022) 13:1017340. doi: 10.3389/fgene.2022.1017340

14. Mumtaz M, Bijnsdorp IV, Bottger F, Piersma SR, Pham TV, Mumtaz S, et al. Secreted protein markers in oral squamous cell carcinoma (OSCC). *Clin Proteomics*. (2022) 19:4. doi: 10.1186/s12014-022-09341-5

15. Plans-Beriso E, Babb-de-Villiers C, Petrova D, Barahona-Lopez C, Diez-Echave P, Hernandez OR, et al. Biomarkers for personalised prevention of chronic diseases: a common protocol for three rapid scoping reviews. *Syst Rev*. (2024) 13:147. doi: 10.1186/s13643-024-02554-9

16. Song X, Yang X, Narayanan R, Shankar V, Ethiraj S, Wang X, et al. Oral squamous cell carcinoma diagnosed from saliva metabolic profiling. *Proc Natl Acad Sci U.S.A*. (2020) 117:16167–73. doi: 10.1073/pnas.2001395117

17. Alapati S, Fortuna G, Ramage G, Delaney C. Evaluation of metabolomics as diagnostic targets in oral squamous cell carcinoma: A systematic review. *Metabolites*. (2023) 13(8):890. doi: 10.3390/metabo13080890

18. Khan SI, Hoque A. SICE: an improved missing data imputation technique. *J Big Data*. (2020) 7:37. doi: 10.1186/s40537-020-00313-w

19. Ganapathy S, Harichandrakumar KT, Penumadu P, Tamilarasu K, Nair NS. Comparison of Bayesian, Frequentist and Machine learning models for predicting the two-year mortality of patients diagnosed with squamous cell carcinoma of the oral cavity. *Clin Epidemiol Glob Health*. (2022) 17:101145. doi: 10.1016/j.cegh.2022.101145

20. Hollmann N, Müller S, Eggensperger K, Hutter F. Data from: TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Second (2023). Available online at: https://go.exlibris.link/HqMS7xW0. doi: 10.48550/arXiv.2207.01848

21. Talukder MSH, Sulaiman RB, Angon MBP. Data from: Unleashing the Power of Extra-Tree Feature Selection and Random Forest Classifier for Improved Survival Prediction in Heart Failure Patients (2023). Available online at: https://go.exlibris.link/DyP7vpkZ. doi: 10.48550/arXiv.2308.05765

22. Dalleau K, Couceiro M, Smail-Tabbone M. Unsupervised extra trees: a stochastic approach to compute similarities in heterogeneous data. *Int J Data Sci Anal*. (2020) 9:447–59. doi: 10.1007/s41060-020-00214-4

23. Goldstein BA, Polley EC, Briggs FB. Random forests for genetic association studies. *Stat Appl Genet Mol Biol*. (2011) 10:32. doi: 10.2202/1544-6115.1691

24. Meysam V, Mohammad G, Masoumeh R. Data from: Performance Analysis and Comparison of Machine and Deep Learning Algorithms for IoT Data Classification (2020). Available online at: https://arxiv.org/abs/2001.09636. doi: 10.48550/arXiv.2001.09636

25. Rimal Y, Sharma N, Alsadoon A. The accuracy of machine learning models relies on hyperparameter tuning: student result classification using random forest, randomized search, grid search, bayesian, genetic, and optuna algorithms. *Multimed Tools Appl*. (2024) 83:74349–64. doi: 10.1007/s11042-024-18426-2

26. Zhang H, Wu Y, Zhang W, Zhang Y. FFNN–tabNet: an enhanced stellar age determination method based on tabNet. *Appl Sci*. (2024) 14:1203. doi: 10.3390/app14031203

27. Wang Y, Zhang X, Wang S, Li Z, Hu X, Yang X, et al. Identification of metabolism-associated biomarkers for early and precise diagnosis of oral squamous cell carcinoma. *Biomolecules*. (2022) 12(3):400. doi: 10.3390/biom12030400

28. Satopaa V, Albrecht J, Irwin D, Raghavan B. *Data from: finding a "Kneedle" in a haystack: detecting knee points in system behavior*. IEEE. (2011). doi: 10.1109/ICDCSW.2011.20.

29. Yang W, Zhou W, Zhao X, Wang X, Duan L, Li Y, et al. Prognostic biomarkers and therapeutic targets in oral squamous cell carcinoma: a study based on cross-database analysis. *Hereditas*. (2021) 158:15. doi: 10.1186/s41065-021-00181-1

30. Jiang W, Zhang T, Zhang H, Han T, Ji P, Ou Z. Metabolic patterns of high-invasive and low-invasive oral squamous cell carcinoma cells using quantitative metabolomics and 13C-glucose tracing. *Biomolecules*. (2023) 13:1806. doi: 10.3390/biom13121806

31. Tan Y, Wang Z, Xu M, Li B, Huang Z, Qin S, et al. Oral squamous cell carcinomas: state of the field and emerging directions. *Int J Oral Sci*. (2023) 15:44. doi: 10.1038/s41368-023-00249-w

32. Chen Z, Huang X, Gao Y, Zeng S, Mao W. Plasma-metabolite-based machine learning is a promising diagnostic approach for esophageal squamous cell carcinoma investigation. *J Pharm Anal*. (2021) 11:505–14. doi: 10.1016/j.jpha.2020.11.009

33. Alfian G, Syafrudin M, Fahrurrozi I, Fitriyani NL, Atmaji FTD, Widodo T, et al. Predicting breast cancer from risk factors using SVM and extra-trees-based feature selection method. *Computers*. (2022) 11:136. doi: 10.3390/computers11090136

34. Hirano K, Kinoshita M, Matsumori N. Impact of sphingomyelin acyl chain heterogeneity upon properties of raft-like membranes. *Biochim Biophys Acta Biomembr*. (2022) 1864:184036. doi: 10.1016/j.bbamem.2022.184036

35. Tallima H, Azzazy H, El RR. Cell surface sphingomyelin: key role in cancer initiation, progression, and immune evasion. *Lipids Health Dis*. (2021) 20:150. doi: 10.1186/s12944-021-01581-y

36. Dickinson A, Saraswat M, Joenvaara S, Agarwal R, Jyllikoski D, Wilkman T, et al. Mass spectrometry-based lipidomics of oral squamous cell carcinoma tissue reveals aberrant cholesterol and glycerophospholipid metabolism - A Pilot study. *Transl Oncol*. (2020) 13:100807. doi: 10.1016/j.tranon.2020.100807

37. Gibellini F, Smith TK. The Kennedy pathway–*De novo* synthesis of phosphatidylethanolamine and phosphatidylcholine. *IUBMB Life*. (2010) 62:414–28. doi: 10.1002/iub.337

38. Santos CR, Schulze A. Lipid metabolism in cancer. *FEBS J*. (2012) 279:2610–23. doi: 10.1111/j.1742-4658.2012.08644.x

39. Indiveri C, Iacobazzi V, Tonazzi A, Giangregorio N, Infantino V, Convertini P, et al. The mitochondrial carnitine/acylcarnitine carrier: Function, structure and physiopathology. *Mol Aspects Med*. (2011) 32:223–33. doi: 10.1016/j.mam.2011.10.008

40. Wu L, Ye C, Yao Q, Li Q, Zhang C, Li Y. The role of serum acylcarnitine profiling for the detection of multiple solid tumors in humans. *Heliyon*. (2024) 10:e23867. doi: 10.1016/j.heliyon.2023.e23867

41. Xu J, Chen Y, Zhang R, Song Y, Cao J, Bi N, et al. Global and targeted metabolomics of esophageal squamous cell carcinoma discovers potential diagnostic and therapeutic biomarkers. *Mol Cell Proteomics*. (2013) 12:1306–18. doi: 10.1074/mcp.M112.022830

42. Wu SL, Zha GY, Tian KB, Xu J, Cao MG. The metabolic reprogramming of gamma-aminobutyrate in oral squamous cell carcinoma. *BMC Oral Health*. (2024) 24:418. doi: 10.1186/s12903-024-04174-0

43. Cao J, Balluff B, Arts M, Dubois LJ, van Loon L, Hackeng TM, et al. Mass spectrometry imaging of L-[ring-(13)C(6)]-labeled phenylalanine and tyrosine kinetics in non-small cell lung carcinoma. *Cancer Metab*. (2021) 9:26. doi: 10.1186/s40170-021-00262-9

44. Ran R, Zhong X, Yang Y, Tang X, Shi M, Jiang X, et al. Metabolomic profiling identifies hair as a robust biological sample for identifying women with cervical cancer. *Med Oncol*. (2023) 40:75. doi: 10.1007/s12032-022-01848-z

45. Qiang YX, You J, He XY, Guo Y, Deng YT, Gao PY, et al. Plasma metabolic profiles predict future dementia and dementia subtypes: a prospective analysis of 274,160 participants. *Alzheimers Res Ther*. (2024) 16:16. doi: 10.1186/s13195-023-01379-3

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 03 December 2024
DOI 10.3389/fimmu.2024.1485109

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong, China

REVIEWED BY
Yuquan Chen,
Monash University, Australia
Luo Qiang,
Children's Hospital of Chongqing Medical University, China

*CORRESPONDENCE
Lingyun Liu
tmxlly@jlu.edu.cn
Yumeng Li
yingzhilifeng1995@163.com
Zhou Sun
sunzhou19@mails.jlu.edu.cn

†The authors have contributed equally to this work

RECEIVED 23 August 2024
ACCEPTED 04 November 2024
PUBLISHED 03 December 2024

CITATION
Zhao Z, Jia H, Sun Z, Li Y and Liu L (2024) A new perspective on macrophage-targeted drug research: the potential of *KDELR2* in bladder cancer immunotherapy. *Front. Immunol.* 15:1485109. doi: 10.3389/fimmu.2024.1485109

A new perspective on macrophage-targeted drug research: the potential of *KDELR2* in bladder cancer immunotherapy

Zhiyi Zhao[1], Hongling Jia[2], Zhou Sun[3*†], Yumeng Li[2*†] and Lingyun Liu[1*†]

[1]Department of Andrology, The First Hospital of Jilin University, Changchun, Jilin, China, [2]The First Clinical Medical College, Shandong University of Traditional Chinese Medicine, Jinan, Shandong, China, [3]Department of Urology, China-Japan Union Hospital of Jilin University, Changchun, Jilin, China

Introduction: Bladder cancer was recognized as one of the most common malignant tumors in the urinary system, and treatment options remained largely limited to conventional surgery, radiotherapy, and chemotherapy, which limited patient benefits.

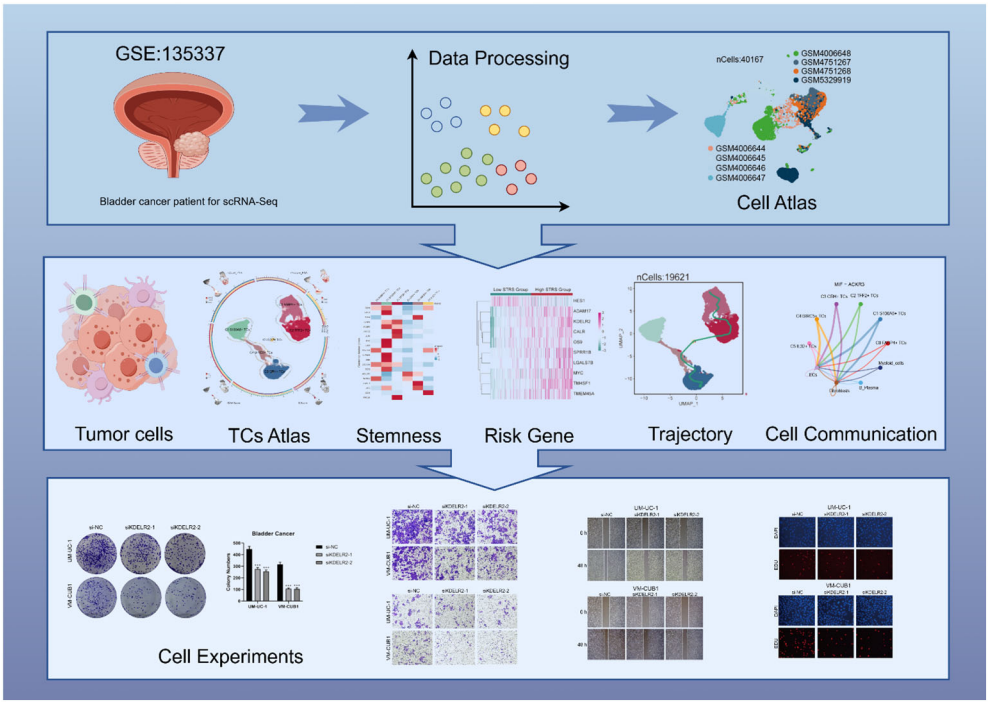
Methods: Researchers constructed an RNA transcriptome map of bladder cancer by integrating single-cell RNA sequencing and clinical data, identifying potential molecular targets for diagnosis and treatment. We also verified the antitumor activity of the target through in vitro experiment.

Results: A distinct tumor cell subpopulation characterized by elevated *S100A8* expression exhibited high copy number variation, high stemness, and low differentiation. It interacted with myeloid cells via the MIF-(CD74+CD44) and MIF-(CD74+CXCR4) signaling pathways. This study underscored *KDELR2*'s role in promoting cell proliferation, invasion, and migration, providing new therapeutic insights. Prognostic analysis revealed that *KDELR2* correlated with poor survival, higher immune scores, and increased macrophage infiltration.

Discussion: The findings suggested that patients with high *KDELR2* expression might benefit from immune checkpoint therapy. *KDELR2* was also shown to enhance bladder cancer cell proliferation, invasion, and migration, highlighting it as a promising target for macrophage-focused drug development.

KEYWORDS

bladder cancer, *S100A8*, TCs, *KDELR2*, macrophage



GRAPHICAL ABSTRACT
Workflow diagram of this study. Workflow showed bladder cancer single-cell RNA sequencing data were used to perform cell stemness, pseudo-time, and cell communication analyses, leading to the identification of a cell subpopulation and the creation of a prognostic model. The *KDELR2* gene was validated by in vitro experiments.

1 Introduction

On a global scale, bladder cancer holds the second position among urological malignancies in terms of prevalence, while it ranks ninth among all malignancies and stands as the 13th leading contributor to cancer-related deaths worldwide (1). The main risk factors include old age, smoking, pelvic radiation therapy, the use of cyclophosphamide, and an enlarged prostate and urinary retention in men, which may also increase the risk of cancer due to the presence of carcinogens in the urine (2). Clinically recognized bladder tumors typically present with symptoms such as significant hemorrhage, urinary system irritative symptoms (e.g., urgency, frequency, and burning), and may or may not be accompanied by massive hematuria. This is especially true for patients with diffuse *in situ* carcinoma, muscle-invasive tumors, or secondary infection-related lesions. While some bladder tumors may be asymptomatic, they can be detected during the evaluation of asymptomatic hematuria (3).

Current modalities for bladder cancer management encompass surgery, chemotherapy, and radiation therapy (4). Immunotherapy for bladder cancer had made strides in recent years, but the benefits were limited, particularly in certain types of bladder cancer. Therefore, researchers aimed to develop more precise targeted therapies and combination strategies. The recently advanced single-cell RNA sequencing(scRNA-seq) technology facilitates precise profiling of diverse gene modules utilizing minimal cell numbers, enabling high-resolution characterization.

KDELR2 is an ER-resident protein involved in the ER stress response and proper protein folding. Its expression in cancer cells could be associated with tumor proliferation, migration, and invasion. For instance, in breast cancer, *KDELR2* was identified as a novel target of HDAC3, with aberrant expression predicting poor outcomes in patients (5). Though *KDELR2*'s role in bladder cancer had not yet been fully explored, its involvement in other cancers suggested it could also be crucial in bladder cancer development.

Hence, *KDELR2* was considered a promising therapeutic target, and the development of drugs aimed at *KDELR2* could provide new treatment options for bladder cancer patients. With our endeavors, it is anticipated that this research will offer personalized diagnostic and treatment strategies as well as immunotherapy guidance for bladder cancer patients for patients with bladder cancer to enhance prognosis and diminish mortality rates.

2 Materials and methods

2.1 Data source

The scRNA-seq data for bladder cancer, sourced from Gene Expression Omnibus (https://ww.ncbinlm.nih.gov/geo/) (GSE135337), was utilized in this study. Data pertaining to bulk RNA-seq was acquired from the Cancer Genome Atlas (TCGA) website (https://portal.gdc.cancer.gov/), which included clinical details (age, gender, ethnicity) and somatic mutation information

for bladder cancer patients. By leveraging publicly accessible data, the quirement for ethical approval was bypassed.

2.2 Single-cell sequencing

The gene expression data was processed in R (v4.2.0) with the Seurat package (v4.3.0). Low-quality cells were filtered out based on stringent criteria, including nFeature range (300–5000), nCount range (500–50,000), and limits on mitochondrial ≤(1) and erythrocyte gene expression ≤(5 contributions to the total gene count. A total of 40,167 cells were obtained after quality control.

Seurat's NormalizeData prepared the gene expression data. The FindVariableFeatures function pinpointed the 2,000 most variable genes (6–10). ScaleData functions (v3.1.4) prepared the gene expression data (11), guiding subsequent principal component analysis. The harmony R package (v.0.1.1) (12, 13) mitigated batch effects between sample.CellCycleScoring is used to calculate cell cycle phases (14). Clusters were formed based on the top 30 principal components (PCs). For UMAP visualization, the same 30 PCs were selectively used to depict gene expression patterns (15–18). A dimensionality reduction of 30 and a resolution of 1.2 were used.

2.3 Identification of cell subpopulations

Cell clusters were initially identified with Seurat's FindClusters and FindNeighbors functions (18–21). Clusters were annotated based on marker genes expression averages.

2.4 Trajectory analysis of TCs subpopulations

Slingshot (v2.6.0) inferred cell lineages and pseudotimes, leveraging clustering-based MSTs to define lineage architecture. Synchronized master curves and branching curve fittings were applied, with getCurves used to derive smooth trajectory curves.

2.5 Assessment of cell stemness

The AUCell method (22) is utilized for the identification of cells exhibiting active gene single-cell RNA-seq data underwent analysis to assess gene expression profiles. It takes gene sets as input and provides an assessment of the 'activity' of each gene set within every individual cell. In this particular study, it was employed to evaluate the stemness characteristics associated with different tumor cell subpopulations.

2.6 Enrichment analysis of cellular subpopulations

Differentially expressed genes (DEGs) were identified using FindAllMarkers, applying a Wilcoxon test with thresholds of min.pct and min.diff.pct at 0.25, and a logfc threshold of 0.25 (23). ClusterProfiler (v4.6.0) facilitated DEGs enrichment and analysis using GO (24), KEGG, and GSEA (25–27). Significant GO terms were identified with an adjusted P-value < 0.05.

2.7 Cell communication analysis

CellChat (v1.6.1) enables quantitative inference and analysis of scRNA-seq data to decipher cellular interactions. CellChat's netVisual diffInteraction function assesses alterations in cell-cell communication strength, while IdentifyCommunicationPatterns determines the count of distinct communication patterns. Scatter plots, heatmaps, and various visualization techniques are used to visually analyze the signals entering and exiting each cell. The CellChatDB database (http://www.cellchat.org/) is then used to identify signaling pathways and receptor pairs associated with specific types of TCs in cancer. A P-value threshold of 0.05 was set to identify statistically significant cell-cell interactions across various cell types.

2.8 Scenic analysis

Using pySCENIC (v0.10.0) in Python 3.7, we inferred single-cell regulatory networks and clustered tumor cell subpopulations. GRNBoost identified TF target genes, which were then refined by DNA-motif analysis. AUCell scored regulon activities, revealing the top 5 transcription factors (TFs) with most significant changes in expression, based on human gene motif rankings from https://resources.aertslab.org/cistarget/.

2.9 Development and Verification of a Prognostic Prediction Model

The purpose of this research was to evaluate the prognostic potential of specific prognostic genes associated with diverse bladder cancer subpopulations in predicting patient survival outcomes. Through a rigorous process involving univariate and multivariate Cox proportional hazards analysis (28, 29), combined with Least Absolute Shrinkage and Selection Operator (LASSO) regression, we pinpointed the most influential prognostic genes as key predictors for developing a robust prognostic model. Subsequently, we formulated a risk scoring system, where the risk score is calculated as the sum of the products between individual gene expression levels and their respective coefficients.

$$\text{Risk Score} = \sum_{i}^{n} Xi \times Yi$$

Employing an optimized cutoff threshold derived from the "surv_cutpoint" function, patients were stratified into distinct low risk and high STRS groups, allowing for a comparative analysis of prognostic differences across patient subpopulations. To visualize and statistically validate the predictive capabilities of our risk score model, we leveraged the "Survival" package (v3.3.1) in R for survival analysis and utilized the "ggsurvplot" function to generate survival

curves (30–32). Moreover, we ensured the reliability of our model by assessing its accuracy and calibration through the generation of Receiver Operating Characteristic (ROC) curves (33–37) using the "timeROC" package (v0.4.0), providing a comprehensive evaluation of the model's predictive performance.

2.10 Immune microenvironment analysis

We utilized the CIBERSORT R package (version 0.1.0) to compute immune-related scores for immune cells, providing a comprehensive evaluation of the patients' immune milieu. Furthermore, we scrutinized the levels of immune cell infiltration and differential gene expression associated with immune checkpoints, and conducted investigations into the correlation between risk scores, immune cells, and model genes (38). Concurrently, we employed the TumorImmune Dysfunction program to appraise the response to tumor immunotherapy.

2.11 Identification of malignant cells by inferCNV

To distinguish between cancerous and non-cancerous cellular populations, we initially estimated the baseline copy number variation (CNV) across various genomic regions by analyzing disruptions in chromosome gene expression patterns. This was accomplished utilizing the inferCNV R package (accessible at https://github.com/broadinstitute/inferCNV/wiki), a tool specifically designed for CNV inference. Using endothelial cells (ECs) as a benchmark, we leveraged the inferCNV algorithm to characterize the CNV landscape within distinct cellular subpopulations. Subsequently, those EPCs subpopulations that exhibited marked alterations in their CNV profiles, indicative of significant genomic instability, were identified and classified as malignant cells, thereby facilitating the differentiation between cancerous and non-cancerous cell types.

2.12 Cell culture

The UM-UC-1 tumor-derived cell line was propagated in MEM medium under standardized conditions (37°C, 5% CO2 atmosphere, and 95% humidity) supplemented with 10% fetal bovine serum (FBS) and 1% antibiotics. Similarly, the VM-CUB1 cell line was maintained in DMEM medium under identical conditions, also enriched with 10% FBS and 1% antibiotics to ensure optimal growth and health.

2.13 Cell transfection

RNA constructs sourced from GenePharma (Suzhou, China) facilitated the downregulation of *KDELR2* expression. The cells were seeded onto a 6-well plate at a moderate density of 50%, subsequently subjected to transfection procedures involving *KDELR2*-specific knockdown constructs (si-*KDELR2*-1 and si-*KDELR2*-2), as well as a negative control construct (si-NC) for comparison. Lipofectamine 3000RNAiMAX (Invitrogen, USA) was used for transfection under manufacturer directions. Every si-RNA (RIbbio, China) was transfected into cells. siRNA sequences: si-1: GUAGUCCAGACCAUCCUAU; si-2: UCGUGCUUUGUAUCUUGUC. qRT-PCR Primers: F: TGGATCTGGCGCTTCTACTT; R: GCTGGCAAACTGAGCTTCTT.

2.14 Western blotting

After the transfected cells attained a 70% confluency level, they were lysed in RIPA buffer. The resulting lysates were clarified through centrifugation at 12,000 rpm for 15 minutes, preparing them for SDS-PAGE separation. The separated proteins were then transferred onto PVDF membranes, which were subsequently blocked with 5% BSA for 1.5 hours at ambient temperature. Following an overnight incubation with an Anti-*KDELR2* antibody at 4°C, the membranes were further incubated with a secondary antibody for an hour. Ultimately, the presence of *KDELR2* protein bands was detected using an ECL Western Blot substrate for visualization.

2.15 Quantitative real-time polymerase chain reaction

The RNA extraction process involved the utilization of Trizol reagent, Trizol reagent was used to lyse cells and release RNA, with chloroform and isopropanol employed to precipitate the RNA while suppressing RNase activity. Throughout the extraction process, it was essential to confirm that the workbench, tools, and water utilized were RNase-free to avoid RNA degradation. Followed by a reverse transcription step facilitated by the PrimeScript™ Kit. Subsequently, the quantitative Real-time Polymerase Chain Reaction was performed using SYBR Green premix as the fluorescent dye for amplification detection.

2.16 Cell viability assay

To assess the viability of UM-UC-1 and VM-CUB1 cells post-transfection, the Cell Counting Kit-8(CCK-8) assay was employed. Cells were seeded in 96-well plates at a density of 5×10^3 cells per well and allowed to incubate for 24 hours. Subsequently, 10μL of CCK-8 labeling reagent (A311-01, Vazyme) was added to each well, and the plates were incubated in the dark at 37°C for two hours. From day one to day four, cell viability was quantitatively determined by measuring the absorbance at 450nm using an enzymatic marker (A33978, Thermo). The average optical density values were computed and plotted on a line graph to visually represent the cellular viability trends over time.

2.17 Transwell assay

Prior to the experiment, the cells underwent a 24-hour serum starvation period in medium devoid of serum. Afterward, the cell suspension was mixed with Matrigel (BD Biosciences, USA) and

seeded into the upper chamber of Costar plates, while the lower chamber was filled with serum-rich medium to create a chemoattractant gradient. The cells were then incubated for 48 hours in a cell culture incubator to allow for migration and invasion. Following incubation, the cells were fixed with 4% paraformaldehyde and stained with crystal violet to visually assess their invasive potential.

2.18 Wound healing assay

The stably transfected cells were plated in 6-well dishes placed in the incubator and cultured at 37°C with 5% CO_2 until the cells reached confluence. Using a sterile 200μL pipette tip, uniform scratches were generated across the cell monolayer in each well. Subsequently, the wells were gently washed with PBS to remove any dislodged cells or debris. The scratched areas were then subjected to incubation in serum-free medium to monitor cell migration. Images of the scratch wounds were documented at 0 hours and again after 48 hours of incubation, with the widths of the scratches measured utilizing Image-J software for quantitative analysis.

2.19 5-Ethynyl-2′-deoxyuridine proliferation experiments

The UM-UC-1 and VM-CUB1 cell lines, post-transfection, were seeded at a concentration of 5×10^3 cells per well in 6-well plates. Following a 24-hour incubation period at ambient temperature, the EdU working solution was introduced into the culture medium and allowed to incubate for 2 hours. Subsequently, the cells underwent a double wash with PBS and were fixed using a 4% paraformaldehyde solution for 15 minutes to stabilize them. After fixation, the cells were permeabilized and quenched with a mixture of 2 mg/ml glycine and 0.5% Triton X-100 for 15 minutes. Finally, the cells were stained with a combination of 1X Apollo solution (1 ml) and Hoechst staining reaction solution (1 ml), followed by a 30-minute incubation period. Fluorescence microscopy was then employed to assess cell proliferation by capturing images of the stained cells.

2.20 Statistical analysis

R and Python software packages are utilized for analyzing data from databases, while GraphPad Prism, specifically version 8.0.1, serves as the tool of choice for experimental data analysis. Throughout the analyses, two-tailed p-values are employed, and statistical significance is determined based on values falling below the threshold of 0.05.

3 Results

3.1 Single-cell sequencing analysis revealed major transcriptomic features of the TME in bladder cancer

Based on our research objectives and requirements, we conducted a comprehensive review of the scRNA-seq data related to bladder cancer in public database. To explore the cellular transcriptome characteristics of the bladder cancer TME, we performed a systematic and comprehensive analysis of the collected scRNA-seq data. The dataset was obtained from seven primary tumor tissue samples and one paracancerous sample from seven bladder cancer patients that included diverse demographic backgrounds and clinical stages of bladder cancer. This diversity enhanced the relevance of our findings to a broader patient population. After removing batch effects and primary quality control, we obtained 40,167 high-quality cells (Figure 1A). Dimension reduction cluster analysis was performed on the screened cells, and we found that they could be divided into 5 cell types, including B plasma cells, ECs, epithelial cells (EPCs), myeloid cells, and fibroblasts (Figure 1B). The bubble plot depicted the levels of expression for the most significant genes in distinct cellular subpopulations, categorized by cell types and tissue samples (Figure 1C). The results found that ECs highly expressed *PLVAP, GNG11, SPARC, RGCC, IGFBP7*; fibroblasts highly expressed *CFD, LUM, DCN, GSN, MT2A*; EPCs showed high expression of *SPINK1, KRT19, LY6D, FXYD3* and *S100A2;* B plasma cells showed high expression of *IGHG1, IGLC3, IGHA1, IGKC, IGLC2*; myeloid cells showed high expression of *HLA-DRA, CCL3, HLA-DPB1, C1QB* and *SPP1*; bladder cancer tissue cells showed high expression of *SPINK1, LY6D, KRT19, FXYD3, S100A2;* paracancerous cells highly expressed genes including *CFD, LUM, DCN, GSN, MT2A,RGCC* and so on. In conclusion, we could find that EPCs marker genes had high consistency with the highly expressed genes in cancer tissues. At the same time, we could observe that the cell phases are different in these 5 cell types, and most of the cells were in G1 phase (Figure 1D). Moreover, most ECs and EPCs were mainly derived from bladder cancer tissue, while fibroblasts were mainly derived from paracancerous cells (Figure 1E). The above study were consistent with Ro/e analysis, that EPCs tended to be derived from bladder cancer tissues (Figure 1H). Through cell cycle analysis of cell subpopulations, each cell subpopulation tended to be distributed in different cycle phases (Figures 1C, D, F). In addition, the proportion of EPCs in bladder cancer tissue cells and paracancerous cells was much greater than that of other cells, especially in bladder cancer tissue cells, as high as 97.7% (Figures 1E, G).

Volcano plots results showed that the different genes among ECs, fibroblasts, EPCs, B plasma cells, and myeloid cells (Figure 1I). To investigate the biological processes of different cell types, by using GO-BP enrichment analysis (Figure 1J), we identified that EPCs were mainly enriched in ATP synthesis coupled electron transport, mitochondrial ATP synthesis coupled electron transport, cytoplasmic translation, proton motive force-driven mitochondrial ATP synthesis and mitochondrial respiratory chain complex assembly. Further GSEA (Figure 1K) of EPCs was performed according to GO-BP terms. The results showed that pathways such as mitochondrial electron transport, NADH to ubiquinone; oxidative phosphorylation and ATP synthesis coupled electron transport showed a positive enrichment trend, by comparison, pathways such as adaptive immune response and extracellular structure organization showed a trend of negative enrichment in this genome.

10.3389/fimmu.2024.1485109

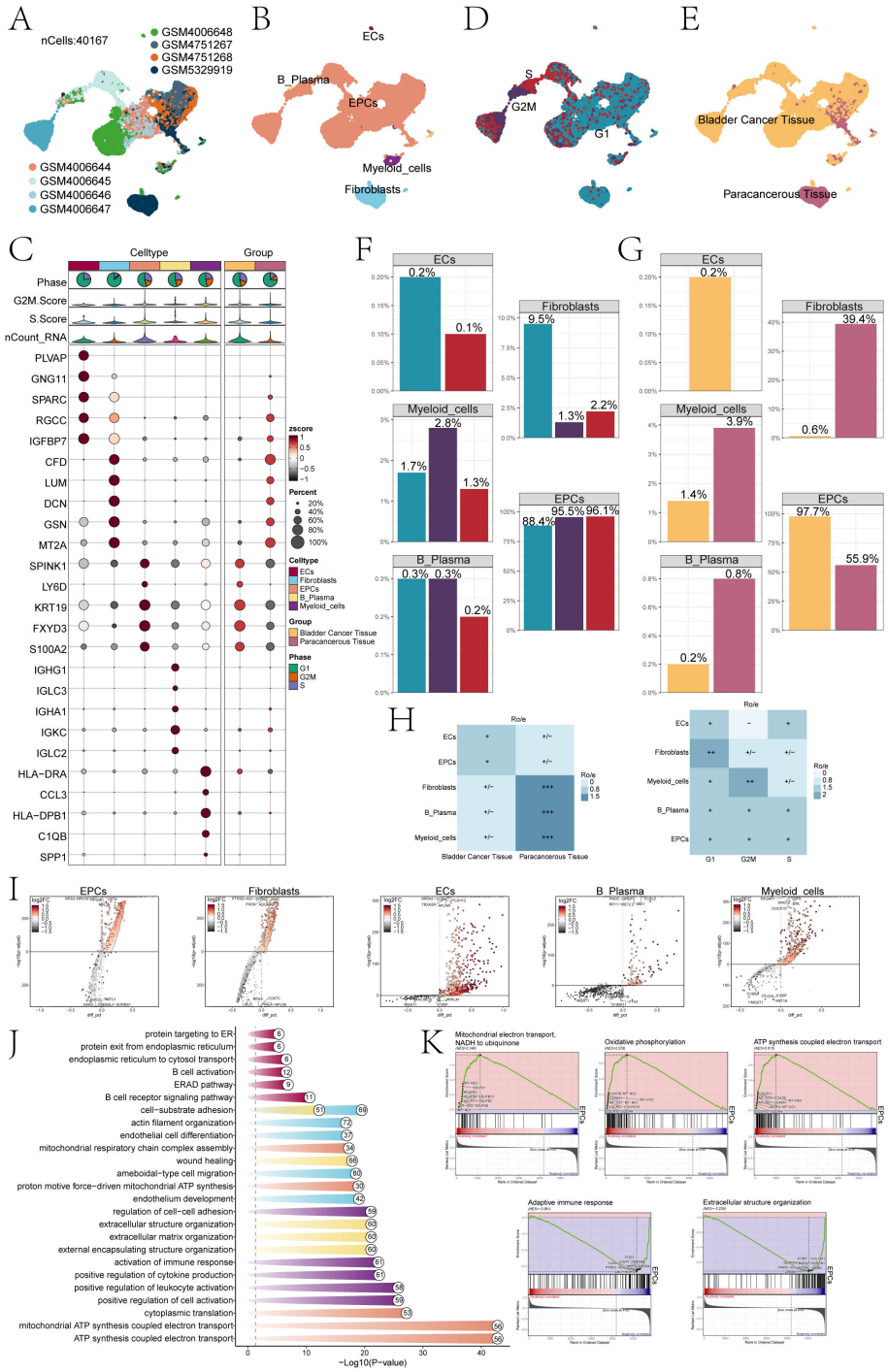


FIGURE 1
Single-cell profiling of bladder cancer identified 5 cell types. **(A)** The UMAP plot visualized the analysis encompassing all cells from eight bladder cancer samples sourced from seven patients,by using single-cell RNA sequencing method(n=40,167). **(B)** UMAP plot showed clusters of five different types of cells (ECs; EPCs; B plasma cells; myeloid cells; fibroblasts). **(C)** The bubble chart displayed the top 5 marker genes associated with each individual cell cluster. The bar charts were color-coded by cell subpopulation, and the pie charts illustrated the proportion of each phase. The violin plots visualized the expression levels of G2M.Score, S.Score, and nCount-RNA, with bubble size representing the percentage of gene expression and color indicating z-score. **(D)** UMAP plot illustrated the cellular distribution at three different cell cycle phases (Phase: G1, G2M, S). **(E)** UMAP plot showed the distribution of cell samples from bladder cancer tissue and paracancerous tissue. **(F)** The bar charts represented the proportions of three cell phases in five different types of cells. **(G)** The bar charts represented the proportions of bladder cancer tissue and paracancerous tissue in 5 different types of cells. **(H)** The heatmap revealed the distribution preferences of different cell subpopulations in terms of sample origin and cell cycle phase. **(I)** Volcano plots showed different expressed genes in 5 types of cells. **(J)** Enrichment analyses of DEGs across all cells unveiled their key biological roles and functions. **(K)** GSEA identified both positively and negatively enriched biological pathways in EPCs, including mitochondrial electron transport, NADH-ubiquinone reduction, oxidative phosphorylation coupled ATP synthesis, adaptive immune response, and extracellular structure organization.

3.2 Visualization analysis of bladder cancer tissue cells

The findings of various studies have demonstrated that the presence of TCs can trigger significant alterations at a molecular, cellular, and physical level within the surrounding tissue, leading to the formation of a specialized environment known as the TME. During the initial stages of tumor development, there is an intricate interplay between cancer cells and various components of this microenvironment, facilitating tumor survival, local infiltration, and metastatic dissemination (39). Therefore, TCs are the real culprit of tumor development. We first used the infer CNV algorithm to successfully separate 19,621 bladder cancer cells from EPCs based on CNV (Supplementary Figure S1). Next, we identified six subpopulations of bladder cancer cells based on the level of marker genes expression. The circle diagram showed six cell subpopulations. These were C0 *FABP4*+ TCs, C1 *S100A8*+ TCs, C2 *TFF2*+ TCs, C3 *CRH*+ TCs, C4 *BIRC5*+ TCs and C5 *IL32*+ TCs. In addition, we used UMAP plots to show the nCount-RNA, nFeature-RNA, G2M.Score and S.Score of each subpopulation (Figure 2A). Subsequently, the bubble plots showed the marker genes of different tissue cells and subpopulations (Figure 2B). According to the comparison of named gene expression of each subpopulation of tumor cell in bar charts, C0 was characterized by high expression of *FABP4*. The high expression of *S100A8* characterizes C1, while C2 was characterized by the elevated expression of *TFF2*. Similarly, C3 was distinguished by its heightened expression of *CRH*, and C4 exhibits a notable increase in *BIRC5* expression. Lastly, the high expression of *IL32* served as a defining feature for C5. UMAP plots showed the named gene expression of each subpopulation, which was consistent with the results of bar charts (Figure 2C).

For further analysis of the subpopulations of TCs, we conducted deeper visualization study. First, we performed sample sources and cell cycle phases analysis of TCs, and combined the UMAP plots to observe the main distribution of different TCs (Figure 2D). The results showed that most of the C1 subpopulation of TCs in G2M stage, compared to the other subpopulations, the highest proportion, as high as 50.90%,at the same time, the tissue classification results show that the C1 subpopulation is different from other subpopulations, its mainly come from bladder cancer tissue cells (Figures 2E, F), and the CNV score was relatively high (Figure 2G). In addition, we conducted CNV score analysis of bladder cancer tissue cells and paracancerous cells, the cancer tissue cell expression level of CNV score was higher than the paracancerous cells. There were significant differences between them. UMAP plot showed the CNV score distribution characteristics. Additional examination of the stemness characteristics of specific subpopulations of cells in bladder cancer demonstrated that the C1 subpopulation had the highest Area Under the Curve (AUC) score for stemness. This meant that C1 had the characteristics of low differentiation degree and strong differentiation potential. The stemness score of bladder cancer cells exhibited a slightly higher level compared to that of paracancerous cells; however, the difference did not reach statistical significance (Figure 2H).

3.3 Heterogeneity of stemness and development in bladder cancer cell subpopulations

Subsequently, we conducted an analysis of gene expression related to cell stemness in the subpopulations of TCs. And we found that the stemness gene expression level was more significant in the C1 subpopulation compared with the other cell subpopulations (Figure 3A). The bar charts confirmed the above conclusion that the C1 subpopulation is higher in *CTNNB1, MYC, HIF1A* and *BMI1* expression levels than other subpopulations cells (Figure 3B). Research had indicated that the atypical activation of *CTNNB1* was linked to the development of various types of tumors, including but not limited to colorectal cancer, ovarian cancer, prostate cancer, hepatoblastoma, and hepatocellular carcinoma (40). *HIF1A*, hypoxia-inducing factor, promotes angiogenesis and is important for the vascular system in the embryo and for cancer tumors (41, 42). As a proto-oncogene, *BMI1* was an important component of the polycomb gene family. High levels of *BMI1* expression have been found to be significantly associated with the onset, progression, and prognosis of diverse malignancies. *BMI1* could participate in tumorigenesis by inhibiting multiple gene loci, synergizing with other proto-oncogenes, and enhancing telomerase activity (43). The synergistic effect of these genes precisely demonstrateed the pro-tumor effect of C1 subpopulations. The UMAP plots showed the expression distribution of *CTNNB1, MYC, HIF1A, BMI1* genes (Figure 3C).

Cell stemness was closely related to cell differentiation and development. To explore the differentiation trajectory of tumor cell subpopulations, we performed a pseudotime analysis and presented it using UMAP plots. First, we showed the lineage trajectory of bladder cancer tissue cells and paracancerous cells (Figure 3D), and the cell cycle lineage trajectory (Figure 3E). The initial cell differentiation trajectory began in the G1 phase and gradually transitioned to the S and G2M phases according to the cell cycle sequence. We then analyzed six different TCs subpopulations, and the UMAP plots showed two main cell lineage tracks (Figure 3F), including lineage 1: C0→C2→C5→C4→C3; lineage 2:C0→C2→C4→C1. The difference between the two trajectories mainly existed in the later stage. Combined with Figure 3D, we could find that the ended of lineage 1 ended mainly C3 subpopulation, which contained cells from both tumor tissues and cells from paracancerous cells. The end of lineage 2 ended mainly C1 subpopulation, which was a tumor cell subpopulation derived entirely from a bladder cancer tissue sample. Combined with the evolution of TCs and based on the tissue origin, CNV score and cell stemness of C1 subpopulation, the conclusion that C1 subpopulation has a high degree of malignancy and is closely related to the progression of bladder cancer was reaffirmed.

Furthermore, we conducted an analysis on the temporal expression patterns of marker genes belonging to 6 distinct subpopulations. The findings indicated that. different from other cell subpopulations. The high expression of marker genes for the C1 subpopulation was predominantly observed during the later stage (Figure 3G). To confirm the above findings, we also conducted a temporal analysis of the expression of cell stemness genes. The

10.3389/fimmu.2024.1485109

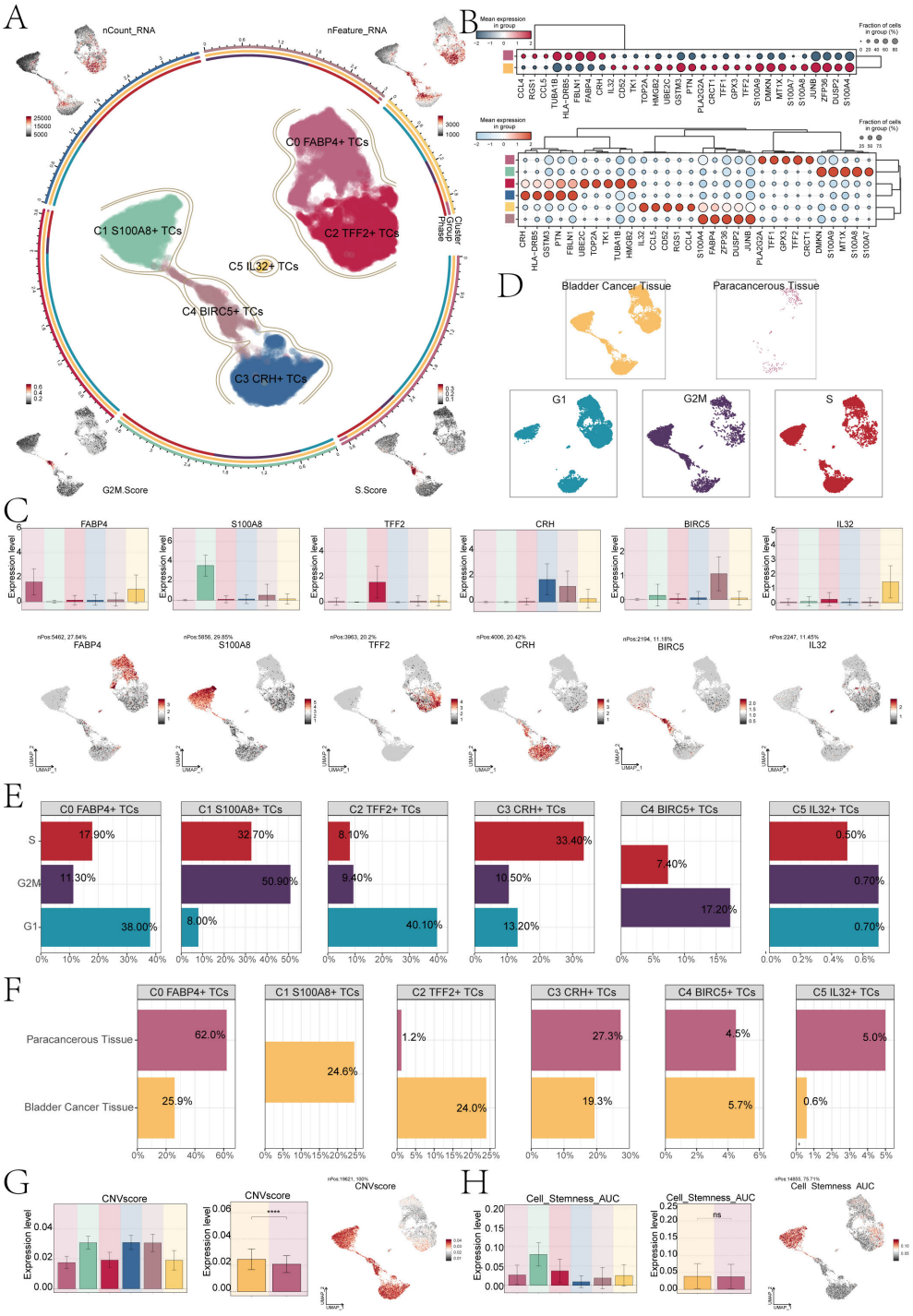


FIGURE 2
S100A8+ TCs specifically expressed in malignant EPCs and are associated with cell stemness. **(A)** The circle plot represented the clustering of the six tumor cell subpopulations identified in bladder cancer. and the contour curve outlines the boundaries of each cell subpopulation. The outer axis of the circle plot represents the logarithmic scale of the entire cell count in each cell category. The three-color tracks representative the ratio of each tumor cell subpopulation in cell types, cell sample types, and cell phases, respectively, and are colored according to cell categories. The UMAP graphs in the four corners start from the upper left corner and go clockwise to show the expression distribution of nCount-RNA, nFeature-RNA, S.score, and G2M.score across all TCs was shown. **(B)** The bubble charts showed the manifestation of marker genes in two sample tissues (top) and in six tumor cell clusters (bottom). **(C)** The bar charts and UMAP plots collectively presented the expression profiles of six marker genes *FABP4*, *S100A8*, *TFF2*, *CRH*, *BIRC5*, and *IL32* across six tumor cell clusters. **(D)** UMAP visualized the distribution of TCs in bladder cancer tissue, paracancerous tissue, and G1, G2M, and S phases. **(E)** The bar charts showed the percentage of G1, G2M, and S phases in six tumor cell clusters. **(F)** The bar charts illustrated the percentage of bladder cancer tissue and paracancerous tissue across six tumor cell clusters. **(G)** The bar plots illustrated the CNV score expression levels across six tumor cell clusters, bladder cancer tissue, and paracancerous tissue. Meanwhile, the UMAP plot visualized the distribution pattern of CNV scores. ****$P < 0.0001$. **(H)** The bar plots showed the AUC score of cell stemness for six tumor cell clusters and bladder cancer tissue and paracancerous tissue. The UMAP plot showed the distribution of cell stemness AUC score. "ns" was used to say that there was no significant difference.

10.3389/fimmu.2024.1485109

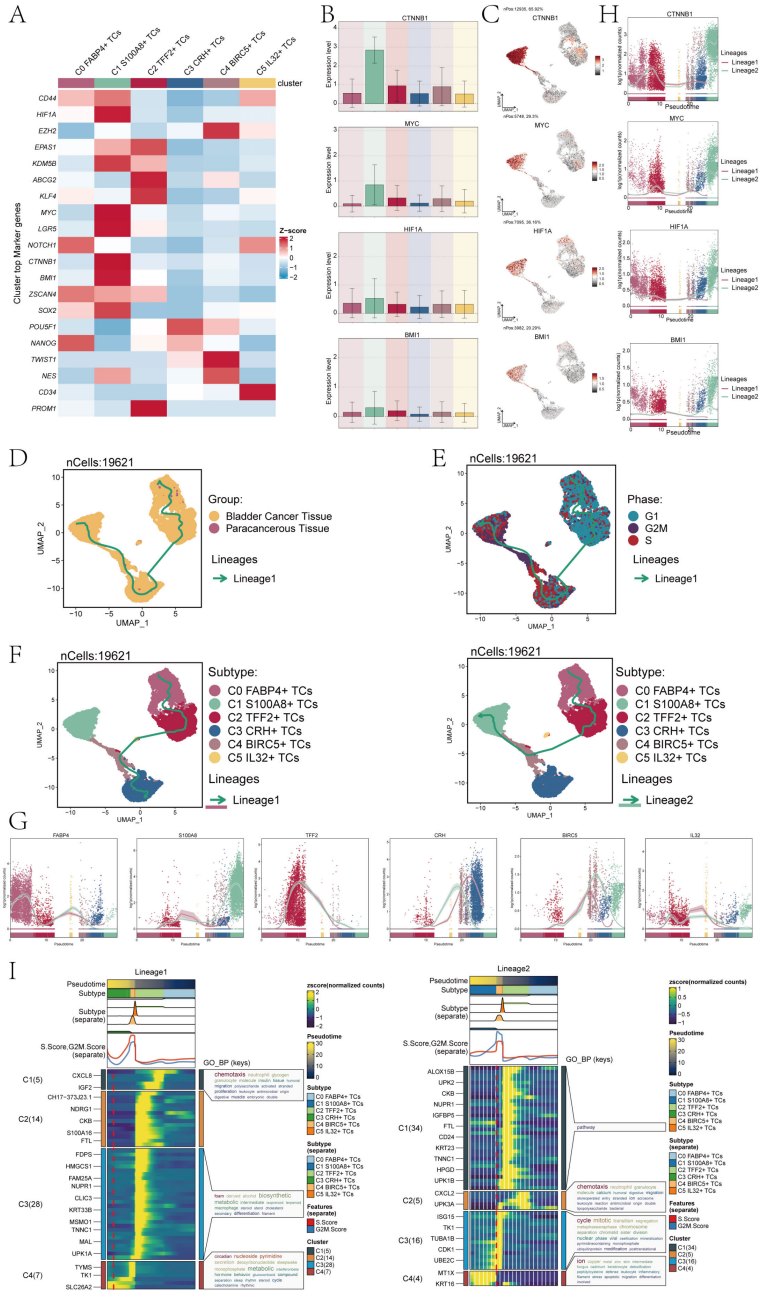


FIGURE 3
Analysis of tumor cell clusters for cell pluripotency and analysis of the developmental trajectory of cells. **(A)** The heatmap showed the z-scores of the marker genes related to cellular stemness in six cell clusters. **(B)** The bar plots displayed the expression levels of four key genes associated with cell stemness across six cellular clusters. **(C)** The UMAP plots showed the dispersion of the four genes related to cellular stemness within all TCs. **(D)** The UMAP plot illustrated the tumor cellular trajectory changes inferred based on bladder cancer tissue and paracancerous tissue. **(E)** The UMAP plot illustrated the temporal trajectories of tumor cellular differentiation, depicted based on three cell cycle phases: G1, G2M, and S. The lineage showed the trajectory from G1 and S to G2M. **(F)** The UMAP plots showed the two lineages of cellular differentiation over time for the six tumor cell clusters discussed. The lineage showed the trajectory from C0 *FABP4*+ TCs to C3 *CRH*+ TCs(left), another lineage showed the trajectory from C0 *FABP4*+ TCs to C1 *S100A8*+ TCs (right). **(G)** The dynamic trend graphs showed the expression of six marker genes over time at different differentiation stages. **(H)** The dynamic trend graphs showed the expression of four stemness genes over time at different stages. **(I)** The heatmaps showed GO enrichment pathways during the differentiation process of TCs. The top bar chart represents pseudo-time and six different types of cells. The faceted mountain plot showed the distribution density of six tumor cell subpopulations spanning various pseudo-time stages. The trajectory plot showed the expression of S.Score and G2M. Score (red represented S.Score, blue represented G2M. Score) as they changed with pseudotime.

primary distinction between lineage1 and lineage2 was observed in the later stage, with a notable increase in stemness gene expression detected in lineage2 compared to lineage1 (Figure 3H). This was consistent with previous conclusions. Subsequently, we analyzed the subpopulations using GO-BP enrichment analysis to verify the related biological processes of the two lineages (Figure 3I), and the dynamic timing showed the expression changes of the different genes of TCs along the two trajectories within the pseudotiming.

3.4 Heterogeneity of biological functions of bladder cancer TCs

To explore the biological functions of TCs in bladder cancer, we first analyzed the different expression genes of six tumor cell subpopulations. The results showed that the main upregulated genes in C1 *S100A8*+ TCs were *RPL32, RPS8, BEST3, IGF2* and *NUPR1* (Figure 4A). Among them, *RPL32* played a crucial role in modulating cellular signaling pathways. It had found that *RPL32* could participate in multiple signaling pathways during tumorigenesis and significantly contribute to the initiation and progression of neoplastic growth (44, 45). *RPS8* was mainly involved in protein folding and stability, and in tumor studies, *RPS8* has the potential to serve as a biomarker specific to tumors due to its tendency for elevated expression levels in tumor tissues and cells compared to normal tissues (46). *BEST3* was a coding gene that encoded proteins with good histocompatibility and low expression levels. *BEST3* had been observed to be present in multiple types of cancerous tissues, and its expression level appears to be linked with the tumor's capacity for invasion. Furthermore, the correlation between the expression level of *BEST3* and the efficacy of tumor therapy has prompted its investigation as a potential therapeutic target or biomarker (47). The role of *IGF2* was pivotal in regulating cellular proliferation, development, motility, differentiation, and viability. Furthermore, *IGF2* played a role in a variety of cancer development (48). *NUPR1* had been linked to the onset and progression of cancer, and it was commonly observed to be upregulated in different cancer forms. Moreover, *NUPR1* could act as a transcriptional regulator affecting the expression of genes involved in cell cycle regulation, apoptosis, and stress response. These genes had a significant correlation with the onset and progression of neoplasms (49, 50).

Furthermore, we conducted enrichment analysis using GO-BP and KEGG to identify enriched biological processes and pathways associated with the DEGs in the subpopulations of TCs. And the heatmap showed the results of the genes enrichment in the TCs subpopulations (Figure 4B). C0 *FABP4*+ TCs was mainly associated with cytoplasmic translation, translational initiation, intrinsic apoptotic signaling pathway, ATP synthesis coupled electron transport and mitochondrial ATP synthesis coupled electron transport biological processes related in GO-BP, and was associated with the ribosome, coronavirus disease-COVID-19, chemical carcinogenesis-reactive oxygen species, oxidative phosphorylation and parkinson disease pathways in KEGG.C1 *S100A8*+ TCs was mainly associated with cytoplasmic translation, protein folding, chaperone-mediated protein folding, regulation of apoptotic signaling pathway and the biological process of ribonucleoprotein complex biogenesis in GO-BP, and was associated with the fluid shear stress and atherosclerosis, ribosome, protein processing in endoplasmic reticulum, prion disease and parkinson disease pathways in KEGG. The association of C2 *TFF2*+ TCs primarily pertained to the biological processes related to hypoxia response, decreased oxygen levels response, oxygen levels response, cellular response to decreased oxygen levels, and cellular response to hypoxia in GO-BP, and was associated with the estrogen signaling pathway, HIF-1signaling pathway, legionellosis, lipid and atherosclerosis and the glycolysis pathways correlation in KEGG. The association of C3 *CRH*+ TCs primarily pertained to the process of ATP production through electron transport, mitochondrial ATP synthesis coupled with electron transport, aerobic electron transport chain, and oxidative phosphorylation in GO-BP, and was associated with the parkinson disease, ribosome, huntington disease, oxidative phosphorylation and the amyotrophic lateral sclerosis pathways correlation in KEGG.C4 *BIRC5*+ TCs was mainly associated with chromosome segregation, ATP synthesis coupled electron transport, mitochondrial ATP synthesis coupled electron transport, nuclear chromosome segregation, respiratory electron transport chain in GO-BP, and was associated with the parkinson disease, huntington disease, amyotrophic lateral sclerosis, prion disease and alzheimer disease pathways correlation in KEGG. C5 *IL32*+ TCs was mainly associated with leukocyte mediated cytotoxicity, leukocyte mediated immunity, cell killing, lymphocyte mediated immunity and regulation of T cell activation biological processes in GO-BP, and was associated with the natural killer cell mediated cytotoxicity, graft-versus-host disease, antigen processing and presentation, primary immunodeficiency and the Epstein-Barr virus infection pathways in KEGG.

In addition, we performed a further step GSEA of the six tumor cell subpopulations, and we could observe that C1 *S100A8*+ TCs were mainly enriched in intracellular zinc ion homeostasis (Figure 4C). GSEA results (Figure 4D) also showed that the C1 subpopulation was positively enriched in the regulation of extent of cell growth, regulation of axon extension, positive regulation of mononuclear cell migration, the intracellular zinc ion homeostasis, while it was negatively regulated in positive regulation of peptide hormone secretion, antigen processing and presentation of exogenous antigen, glycolytic process through glucose-6-phosphate and glycolytic process through fructose-6-phosphate.

In summary, C1 subpopulation was mainly associated with biological pathways such as protein folding and ribosomes, which promoted and accelerated tumor cell migration and invasion. Cells were very sensitive to physiological conditions, and when the physiological conditions changed, the cells would change accordingly. When the loaded proteins exceed the folding capacity of the endoplasmic reticulum stress (ERS), ERS will be caused. The core of ERS lied in the occurrence of protein misfolding, which significantly influenced the growth and viability of cancerous cells and may facilitate the progression of bladder carcinoma (51, 52). Ribosomes, as a kind of nutrient, can escort the smooth and rapid translation of viruses and provide conditions for tumor proliferation, thus affecting the progression and development of bladder cancer (53).

3.5 Analysis of the cell interactions in the TME

To gain a thorough and organized comprehension of intricate cellular reactions, our aim was to conduct an examination on intercellular connections and networks involved in ligand-receptor communication, with the intention of visualizing interactions between

10.3389/fimmu.2024.1485109

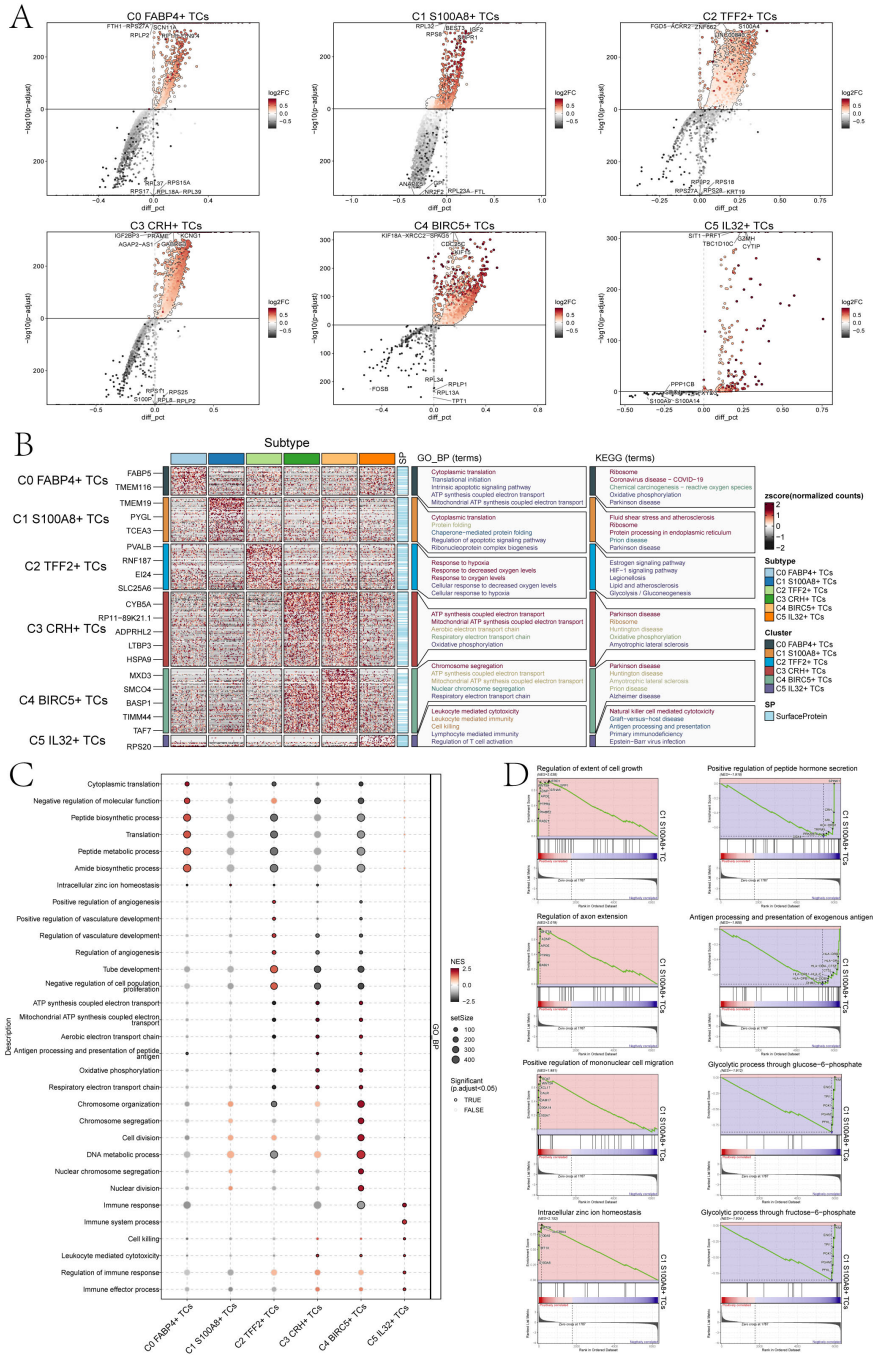


FIGURE 4
Perform gene and pathway enrichment analysis for each tumor cluster set. **(A)** The volcano plots showed the differential gene expression signatures across the six clusters. **(B)** The heatmap displayed the top five enrichment pathways among the six clusters identified through GO-BP and KEGG enrichment analysis. **(C)** The bubble plot showed the GSEA results of the six tumor cell clusters. **(D)** GSEA analyzed eight positively or negatively enriched pathways in C1 *S100A8*+ TCs.

cells. By means of CellChat analysis, we successfully established a network for cellular communication encompassing the majority of cells, including ECs, fibroblasts, B plasma cells, myeloid cells and six tumor cell subpopulations (Figure 5A). Then, we determined the quantity of interactions through the 'line' connection linking two cell types, with thicker lines indicating a higher number of interaction paths. Additionally, we represented the strength of interactions using line weight, where thicker lines indicated stronger interactions. Among different cell subpopulations. Next, we combined the CellChat maps to explore how different cell subpopulations interact with each other through these pathways. First, we categorized cell communication patterns into three main types. In addition, the displayed heatmaps illustrated the expression of cell interaction proteins in the three modes. The two heatmaps on the left showed the outgoing signal patterns of different cell subpopulations and the active interacting proteins in the signal patterns, and the two heatmaps on the right showed the incoming

10.3389/fimmu.2024.1485109

signal patterns and interacting proteins of cell subpopulations (Figure 5B). We could observe that in the outgoing signal pattern diagrams, C1 *S100A8*+ TCs were mainly in pattern 3. Correspondingly, ncWNT, CDH, MPZ, MK, and OCLN showed high expression levels in pattern 3. Additionally, the incoming signal patterns of C1 *S100A8*+ TCs exhibited pattern 1 including but not limited to the pathways represented by such as EPHA, CDH, MPZ, IGF, and OCLN. The Sankey diagrams and heatmaps (Figure 5C, D) showed the communication pattern and target signal of each cell subpopulations, which was consistent with the results in Figures 5B.

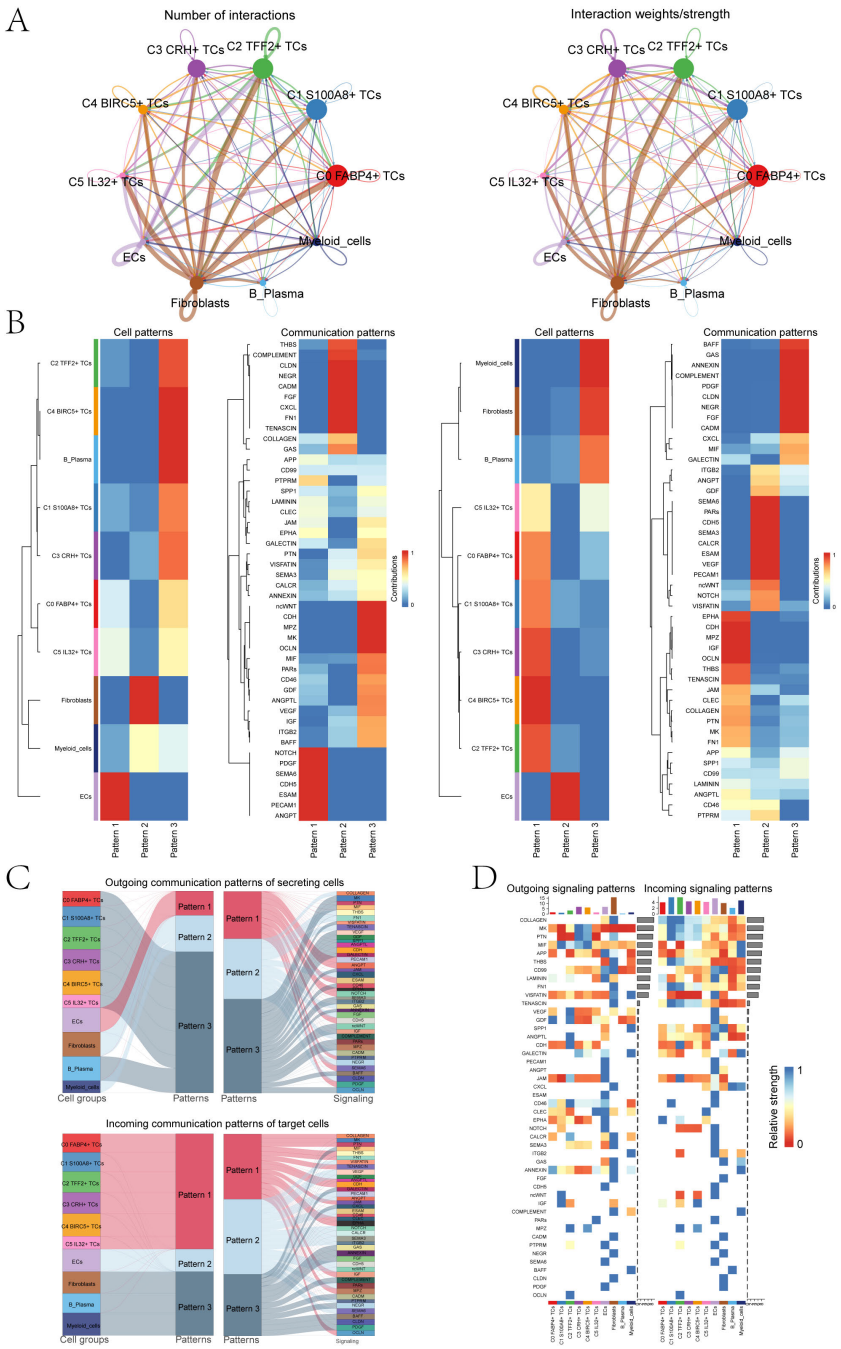


FIGURE 5
Bladder cancer cells was characterized by cell-to-cell signaling networks. **(A)** The circle charts summarized the quantity and intensity of interactions between six tumor cell clusters and four distinct cell types, providing insights into their interconnectedness. **(B)** The heatmaps separately showed the contributions of the six tumor cell clusters and four cell types in the outgoing (left) and incoming (right) signaling under three cell communication patterns, as well as the contributions of various proteins in the three communication patterns. **(C)** The Sankey diagrams illuminated the outbound communication pattern of secretory cells and the inbound communication pattern received by target cells. **(D)** The bar charts compared the relative signaling strengths of six tumor cell clusters and four cell types in both incoming and outgoing patterns. Complementarily, the heatmaps visualized the reception intensities of various proteins within these communication patterns across the same cell groups.

3.6 Explored the cell signaling pathways in C1*S100A8*+ TCs

Subsequently, we analyzed the number of cell interaction strength between the C1 *S100A8*+ TCs and other cell subsets in depth (Figures 6A, B). We could find that both in the incoming signal or the outgoing signal, the number and intensity of the interaction between myeloid cells and C1 *S100A8*+ TCs were more significant than other cells, so we could infer that C1 *S100A8*+ TCs cells and myeloid cells had a strong interaction. Subsequently, we performed signaling network analysis and showed that C1 *S100A8*+ TCs and myeloid cells mainly communicate with MIF signals pathway (Figure 6C). Then we analyzed the role of the two in the signaling pathway (Figure 6D), which showed that C1 *S100A8*+ TCs cells mainly played the role of signal sender, while myeloid cells mainly played the role of signal sender and influencer.

By comparison, we could infer that the MIF ligand of C1 *S100A8*+ TCs cells acted on the CD74-CXCR4 and CD74-CD44 receptor of myeloid cells (Figures 6E, F). The cell interaction circle diagrams (Figures 6G, H) showed the interaction relationship between C1 *S100A8*+ TCs cells and myeloid cells in MIF-(CD74 +CXCR4) and MIF-(CD74+CD44) cell signaling pathway, verifying the above statement. In order to make the results more visual, we used the interaction hierarchy diagrams (Figures 6I, J) to show the relationship between C1 *S100A8*+ TCs cells and myeloid cells (the thickness of the line represents the interaction strength, the thicker of the line, the relationship is more significant). The findings indicated that the intercellular communication between C1 *S100A8*+ TCs and myeloid cells primarily took place through paracrine, leading to signal crosstalk. This could disrupt normal signal transduction, leading to alterations in the tumor immune microenvironment, thereby facilitating the proliferation, metastasis, and invasion of TCs (54).

3.7 TFs regulate the oncogenic mechanism of C1 *S100A8*+ TCs

TFs act on genes to regulate gene transcription by binding specific nucleotide sequences upstream of the gene, thereby affecting cell biological functions. Firstly, we conducted cluster analysis of bladder cancer tissue cells according to genes expression (Figure 7A). In addition, the UMAP plots displayed the expression of TFs in six tumor cell clusters (Figure 7B). Subsequently, based on the heatmap of TFs correlation (Figure 7C), we divided the TFs with similar functional and expression into three main modules, namely M1, M2 and M3. Next, we visualized the analysis of the divided modules, and observed the expression levels of each TCs subpopulation in different modules through the UMAP plots. The results showed that C1 *S100A8*+ TCs cells had significant expression in the M1 module (Figure 7D). To further verify the above statement, we performed transcription factor regulatory activity score analysis (Figure 7E), and we can observe that in the M1 module, C5 *IL32*+ TCs cells have the highest transcription factor regulatory activity score, followed by C1 *S100A8*+ TCs cell. The bar charts also verify the above statement that C1 *S100A8*+ TCs cells have high expression levels in the M1 module, second only to C5 *IL32*+ TCs cells (Figure 7F). Subsequently, we conducted an analysis on the top 5 TFs in the subpopulations of TCs (Figure 7G). The specificity score identified HES1, AHR, TBL1XR1, IRF4, and SPI1 as the top 5 TFs in C1 *S100A8*+ TCs. Finally, we used the UMAP plots to show the TFs distribution characteristics of C1 *S100A8*+ TCs cells (Figure 7H).

3.8 Constructed the prognostic model of bladder cancer

To investigate the prognostic factors in patients, we employed univariate Cox regression analysis to discover 21 genes that exhibited an association with prognosis (Figure 8A). We could observe *HES1* HR<1, while HR values of other genes were>1. Therefore, *HES1* was a protective factor that favors patient prognosis, while others were risk factors and unfavorable to patient prognosis. To mitigate the issue of multicollinearity among genes, we employed LASSO regression analysis to identify 10 genes that were associated with prognosis (Figure 8B). Multivariate Cox regression analysis was then performed on the above genes and used to calculate the genes risk coefficient for these genes (Figures 8C, D). The curve graph and scatter plot revealed variability in risk scores and survival outcomes between low STRS (*S100A8*+ tumor risk score) group and high STRS group, while high STRS group was associated with worse outcome (Figure 8E). The heatmap showcased the distinctive expression patterns of the prognostic genes in both cohorts, setting them apart from the rest of the gene pool (Figure 8F), *HES1* was more significantly expressed in low STRS group and *HES1* favors the prognosis of the patient, consistent with previous conclusions. The ROC curves and AUC values (Figure 8G) for 1-year, 3-years, and 5-years periods exceeded the threshold of 0.6, indicating that the prediction model was specific and valuable. The Kaplan-Meier survival curve further confirmed the low survival profile of high STRS group (Figure 8H). Combined with Figure 8D, the *KDELR2* risk score was the highest. Hence, we conducted an analysis using the Kaplan-Meier survival curve (Figure 8I) and observed a notable decline in survival rates among patients exhibiting elevated *KDELR2* gene expression.

3.9 Prognostic model enrichment analysis

To provide clarity on the distinction between the two scoring cohorts, an analysis was conducted on the genes that exhibited differential expression (Figure 9A). The volcano plot showed the up-regulation and downregulation trends of DEGs (Figure 9B). Subsequently, to understand the biological processes for the above genes, we performed various enrichment analyses. The first was GO enrichment analysis, which reveals the main biological process, cell composition and molecular function (Figure 9C). In GO-BP, genes were mainly enriched in biological processes like epidermis development, keratinization, keratinocyte differentiation, intermediate filament organization, skin development. In GO-CC, its predominant enrichment were observed in the cornified envelope,

10.3389/fimmu.2024.1485109

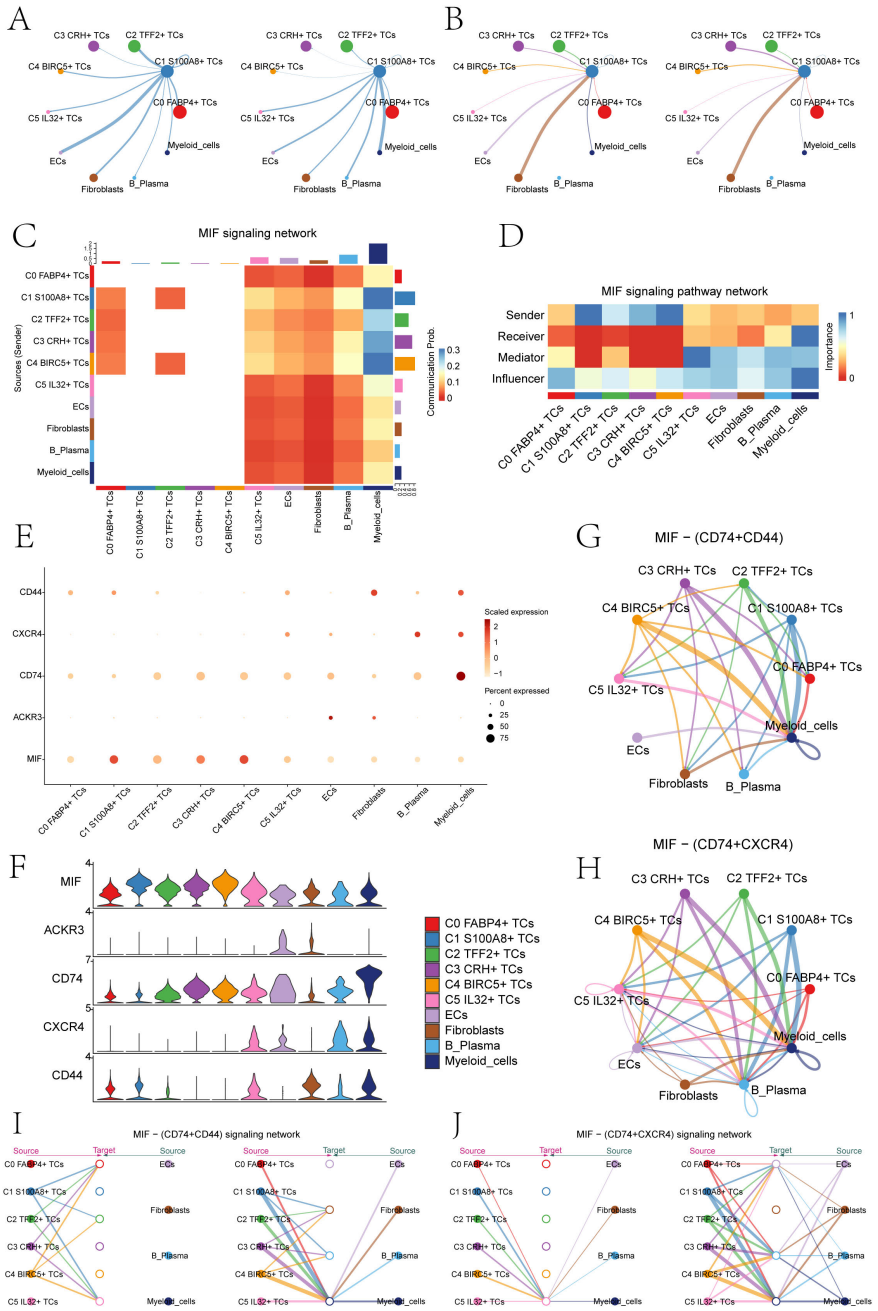


FIGURE 6
The MIF signaling network was the main communication method between *S100A8*+ TCs and Myeloid cells. **(A)** The circle diagrams showed TCs as the signal emitter and other cells as signal receiver The left side represented the communication strength, while the right side represented the number of communications. **(B)** The circle diagrams showed other cells as the signal emitter and TCs as signal receiver. The left side represented the communication strength, while the right side represented the number of communications. **(C)** The heatmap showed the communication probability of various cell clusters based on the MIF signaling pathway. There was a high probability of communication between C1 *S100A8*+ TCs and Myeloid cells. **(D)** The heatmap showed that C1 *S100A8*+ TCs mainly acted as signal senders, while myeloid cells mainly played a signal receiver and influencer in the MIF pathway. **(E, F)** The bubble chart and violin plots displayed that in the MIF pathway, the communication crosstalk between C1 *S100A8*+ TCs and myeloid cells through the MIF- (CD74+CD44) and MIF-(CD74+CXCR4) ligand receptor pair. **(G, H)** The circle diagrams showed the interactions between C1 *S100A8*+ TCs and myeloid cells in the MIF- (CD74+CD44) and MIF- (CD74+ CXCR4) signaling pathways. **(I, J)** The hierarchy diagrams illustrated the autocrine and paracrine interactions between the six tumor cell clusters and ECs, fibroblasts, B plasma cells, and myeloid cells on the MIF- (CD74+CD44) and MIF- (CD74+ CXCR4) signaling pathway.

intermediate filament, and cytoskeletal components of intermediate filaments as well as keratin filament. In GO-MF, mainly enriched in structural constituent of skin epidermis, serine-type endopeptidase activity, serine hydrolase activity, endopeptidase activity and receptor ligand activity. In addition, we performed KEGG pathway enrichment analysis of differential genes (Figure 9D), and showed that the related pathways were staphylococcus aureus infection, retinol metabolism, ECM-receptor interaction, steroid hormone

10.3389/fimmu.2024.1485109

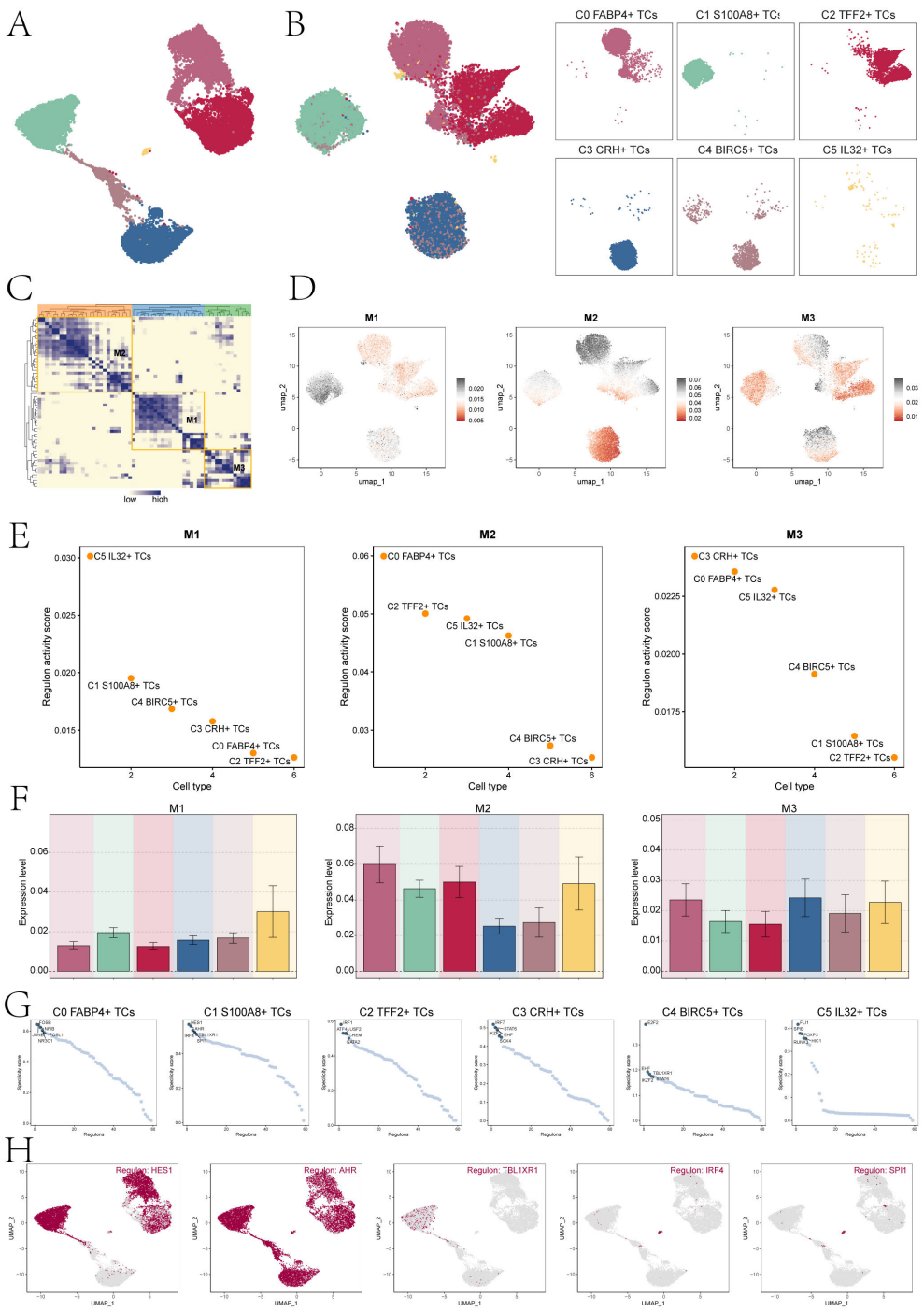


FIGURE 7
Cluster analysis of TFs and the top five TFs in C1 *S100A8*+ TCs. **(A)** The UMAP plot displayed tumor cell clustering based on gene expression levels. **(B)** The UMAP plots visualizations highlighted distinct clustering patterns among TCs, grouped according to the activation levels of various TFs. **(C)** The heatmap displayed three modules M1, M2, and M3 of transcription factor hierarchical clustering. **(D)** The UMAP plots depicted the distinct expression patterns of TFs across the three tumor cell modules. **(E)** The dot plots displayed the ranking of transcription factor regulatory activity scores for different tumor cell clusters in three modules. **(F)** The bar charts showed the expression levels of six cell clusters in three modules. **(G)** Ranking of the top 5 transcription factor activity scores of different cell types. **(H)** The UMAP plots displayed the expression of the top five TFs in C1 *S100A8*+ TCs.

biosynthesis and cytokine-cytokine receptor interaction. Finally, we performed a GSEA of the enriched pathways (Figure 9E), the results showed that keratinization, keratinocyte differentiation, skin development and the epidermis development related pathways showed a positive enrichment trend, mitochondrial respiratory chain complex assembly, uronic acid metabolic process, the flavonoid metabolic process and ATP synthesis coupled electron transport related pathways showed a negative enrichment trend.

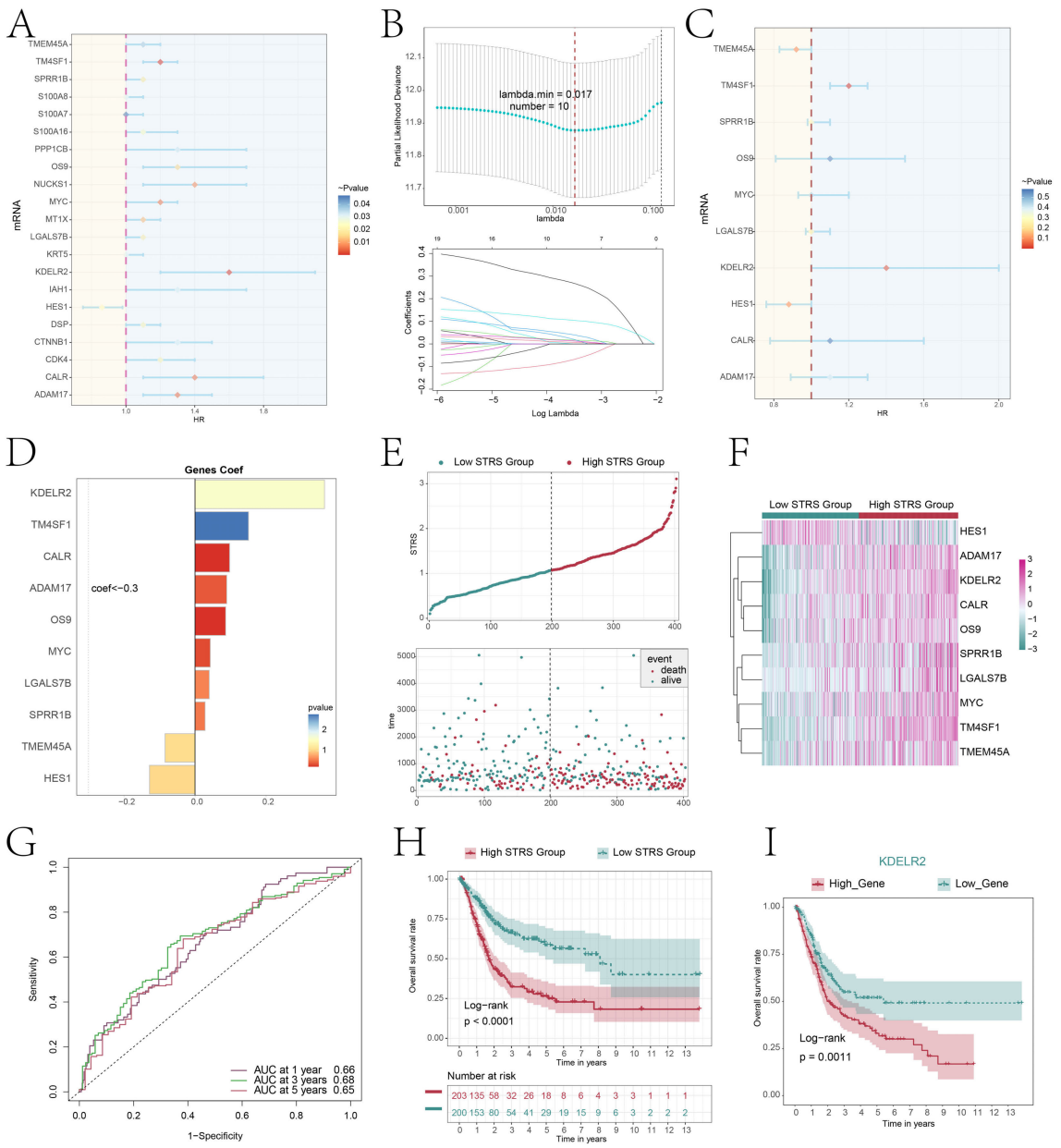


FIGURE 8
Constructing a risk prediction model through a combined approach of univariate Cox proportional hazards analysis and Lasso regression. **(A)** The forest plot displayed the top 21 genes obtained from univariate Cox analysis that were associated with prognosis. (aHR > 1 indicated poor prognosis). **(B)** By setting the lambda.min = 0.017 for LASSO regression curve, we obtained 10 prognostic-related genes(up). Each line depicted the coefficient assigned to a distinct variable, selected for its significant prognostic value. (bottom). **(C)** The forest plot displayed the top 10 genes obtained from multivariate Cox analysis that were associated with prognosis. (aHR > 1 indicated poor prognosis). **(D)** The bar plot showed gene coefficients about those 10 prognostic-related genes. **(E)** A curve graph compared the risk scores between patients in the low and high STRS groups, while a scatter plot visualized survival outcomes, with blue dots indicating survival events and red dots signifying death events. **(F)** The heatmap contrasted the expression levels of ten risk genes between the high and low STRS groups, providing insights into their differential activation patterns. **(G)** The ROC curve analysis, along with its corresponding AUC value, offered a quantitative assessment of the predictive performance of the model in estimating patient survival cycles. **(H)** A Kaplan-Meier survival analysis was conducted to compare the survival outcomes between patients in the high STRS group and those in the low STRS group. **(I)** A Kaplan-Meier survival analysis distinguished survival trends between patients stratified into high *KDELR2* group and low *KDELR2* group.

3.10 Disparities in immune infiltration observed between high and low STRS groups

We performed immune infiltration analysis on high STRS group and low STRS group. The box plot revealed higher expression levels of macrophages M0 and M1, and neutrophils in the high STRS group (Figure 10A). Those cell types belong to myeloid cells. It was noteworthy that macrophages exhibited a higher abundance of immune infiltration and exerted a more pronounced effect on the high STRS group. Subsequently, we further investigated the correlation between risk score and immune infiltration (Figure 10B).

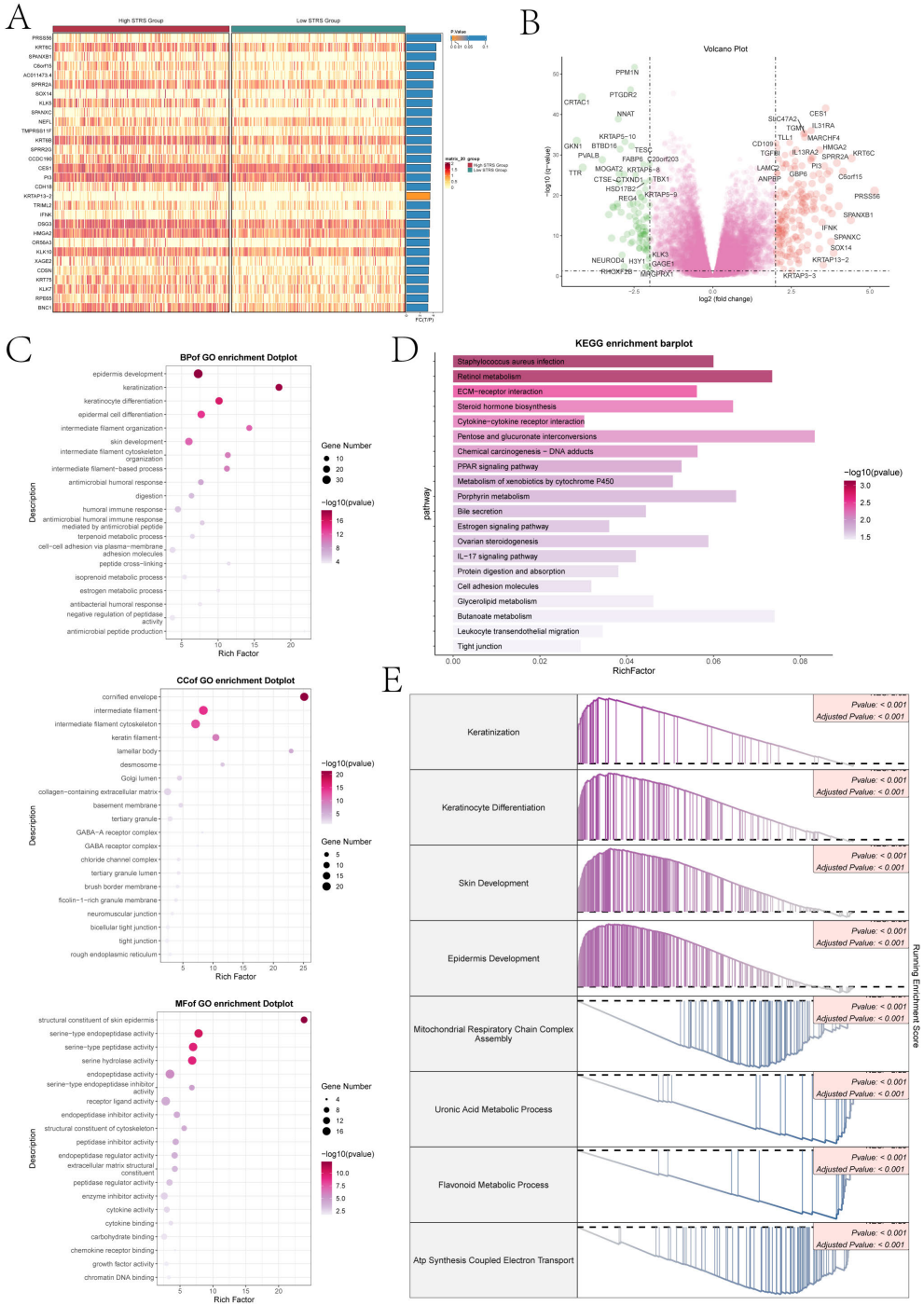


FIGURE 9
Differential gene expression and enrichment analysis. **(A)** The heatmap illustrated distinct patterns of gene expression between the high and low STRS groups. **(B)** The volcano plot visually displayed the variation in expression levels among genes that exhibited differential expression. **(C)** The dot plots sequentially displayed Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) categories from the GO enrichment analysis. **(D)** KEGG enrichment bar plot showed top 20 enrichment pathways. **(E)** Eight GSEA pathways that were positively and negatively enriched.

In the correlation lollipop chart, macrophages also showed a close positive correlation with risk score. In addition, the immune-score, stromal-score, and ESTIMATE-score of the high STRS group were significantly higher than those of the low STRS group, while the tumor-purity was significantly lower than that of the low STRS group (Figure 10C). Scatter plots demonstrated positive associations between *KDELR2* expression levels with immune-score, stromal-score, and ESTIMATE-score as well as macrophages M0 (Figure 10D). Elevated *KDELR2* expression was linked to increased infiltration of macrophages and a more prominent immunosuppressive microenvironment. Therefore, target mining of macrophages might be a potential strategy for bladder cancer immunotherapy. The

heatmap provided visual comparisons that supported our previous findings (Figure 10E).

Furthermore, the bubble plot (Figure 10F) showed the correlation between immune checkpoints and prognostic genes, OS, Risk. Among them, CD274, CD276, CD80, CD86, HAVCR2, LAIR1, NRP1, PDCD1LG2, TNFRSF8, TNFRSF9, TNFRSF4, and VTCN1 were significantly positively correlated with *KDELR2*. This suggested that patients with high *KDELR2* expression were potential candidates for these immune checkpoint inhibitors (ICIs). In order to optimize ICIs utilization and improve therapeutic efficacy further analysis was conducted on gene expression levels of different immune checkpoints between high STRS group and low STRS group using a box plot (Figure 10G). The results indicated significantly higher expression levels for most genes in the high STRS group compared to those in the low STRS group. These findings suggested that application of these ICIs to patients classified as high STRS group might result in improved treatment outcomes.

3.11 *KDELR2* inhibits the proliferation and invasion of bladder cancer cells

In an effort to delve deeper into the impact of *KDELR2* on the patient outcome in cases of bladder cancer, we conducted *in vitro* experiments. Firstly, we divided tumor cell lines of different sources into three groups: si-NC, si-*KDELR2*-1 and si-*KDELR2*-2. The findings indicated notable variances in the expression levels of protein and mRNA for UM-UC-1 and VM-CUB1. Specifically, the si-*KDELR2* group exhibited lower expression levels compared to the si-NC group (Figure 11A). In addition, we investigated the effect of *KDELR2* on tumor cell activity. First, tumor cell activity was tested by CCK-8, and fluorescence staining showed that the tumor cell activity in si-NC was higher than that in si-*KDELR2*-1 and si-*KDELR2*-2. Subsequently, we used line plots to quantify indicators. The OD values of si-*KDELR2*-1 and si-*KDELR2*-2 were found to be significantly reduced compared to the si-NC group, aligning with the outcomes obtained from CCK-8 staining (Figure 11B). Next, we studied the proliferation ability of TCs. The colony formation assay (Figure 11C) similarly corroborated the aforementioned statements. Through EDU experiments, the microbiota density was lower after *KDELR2* was knocked out. Subsequent data analysis showed that si-*KDELR2* inhibited cell proliferation and cloning (Figures 11D, E). In order to investigate the invasion and migration ability of the cells, we conducted wound healing assay and transwell assay (Figures 11F, G). After 48 hours of assay, si-*KDELR2* TCs' distance was greater than that of si-NC, while the density was lower than that of si-NC, which proved that si-*KDELR2* TCs had low invasion and migration ability. The bar plots were consistent with the experimental conclusion that UM-UC-1 and VM-CUB1 cells inhibited the invasion, migration and wound healing of TCs after the knockout of si-*KDELR2* (Figure 11H). In summary, *KDELR2* had obvious tumor promoting effect, which played a key role in the development of TCs. In conclusion, our research provided a more comprehensive understanding of the *KDELR2* gene. Through a series of studies, we discovered that *KDELR2* promoted tumor cell development. It could be inferred that *KDELR2* played a key role in driving tumor cell proliferation, migration, and other related processes. Accordingly, we could achieve the purpose of improving prognostic survival by inhibiting *KDELR2*, making it possible focus for therapeutic intervention in cases of bladder cancer.

4 Discussion

In today's society, bladder cancer is still a global problem, posing a threat to human health, and its incidence has significant population differences, the incidence of men is generally higher than that of women, but most female patients with bladder cancer are more serious and have a poor prognosis (55). In most cases, the prognosis of bladder cancer is determined by the distinct attributes and differences observed among individual patients (56). Hence, it is imperative to conduct a thorough investigation into the diversity of bladder cancer in order to improve patient prognosis and identify possible targets for therapeutic interventions in this disease. Using scRNA-seq, the cellular and molecular features of bladder cancer tissues and identified five known cell types. For different cell subpopulations, we performed temporal phase, sample source as well as pathway enrichment analysis. We found that unlike other cells, EPCs were mainly derived from bladder cancer tissue cells.

Enrichment analysis showed that ATP synthesisation-related pathways were significantly expressed in EPCs, and Mitochondrial electron transport and NADH to ubiquinone were positively enriched in EPCs. ATP played a very important role in cells and was one of the important links of energy information, which could control the life activity of tumors (57)and make the adverse process, namely tumor proliferation, possible, thus promoting the development of TCs (58). Mitochondria could provide an energy source for cells, and in cells they could influence signature functions, including avoiding cell death, bioenergy dysregulation, genome mutation, and promoting tumor inflammation and metastasis (59). Cancer cells would derive a kind of tumor stem cells with the ability of self-renewal, metastasis and spread and treatment resistance during their self-evolution, which played a key role in tumor occurrence and development (60). Based on this, we studied TCs of bladder cancer. First, tumor tissue was divided into six subpopulations based on marker genes. Among them, we discovered that TCs exhibiting elevated levels of *S100A8* were exclusively derived from tumor tissue specimens. Bubble chart showed the C1 subpopulation high expressed *DMKN MT1X, S100A7, S100A9, S100A8* genes, etc. Among them, *S100A9* and *S100A8* had strong pro-inflammatory functions. According to their high expression in cancer, they showed abundant expression in TCs and infiltrate immune cells. The involvement of these factors was crucial in the progression of cancer. Furthermore, it was observed that C1 subpopulation had higher expression levels of CNV score and Cell-stemness. Therefore, we could conclude that C1 subpopulation cells have higher malignant degree and higher differentiation potential. Overall, C1 *S100A8*+ TCs had complex relationships with further development of bladder cancer.

To clarify the relationship between C1 *S100A8*+ TCs subpopulation and tumor, we performed various enrichment analyses on cell subpopulations. C1 *S100A8*+ TCs were

10.3389/fimmu.2024.1485109

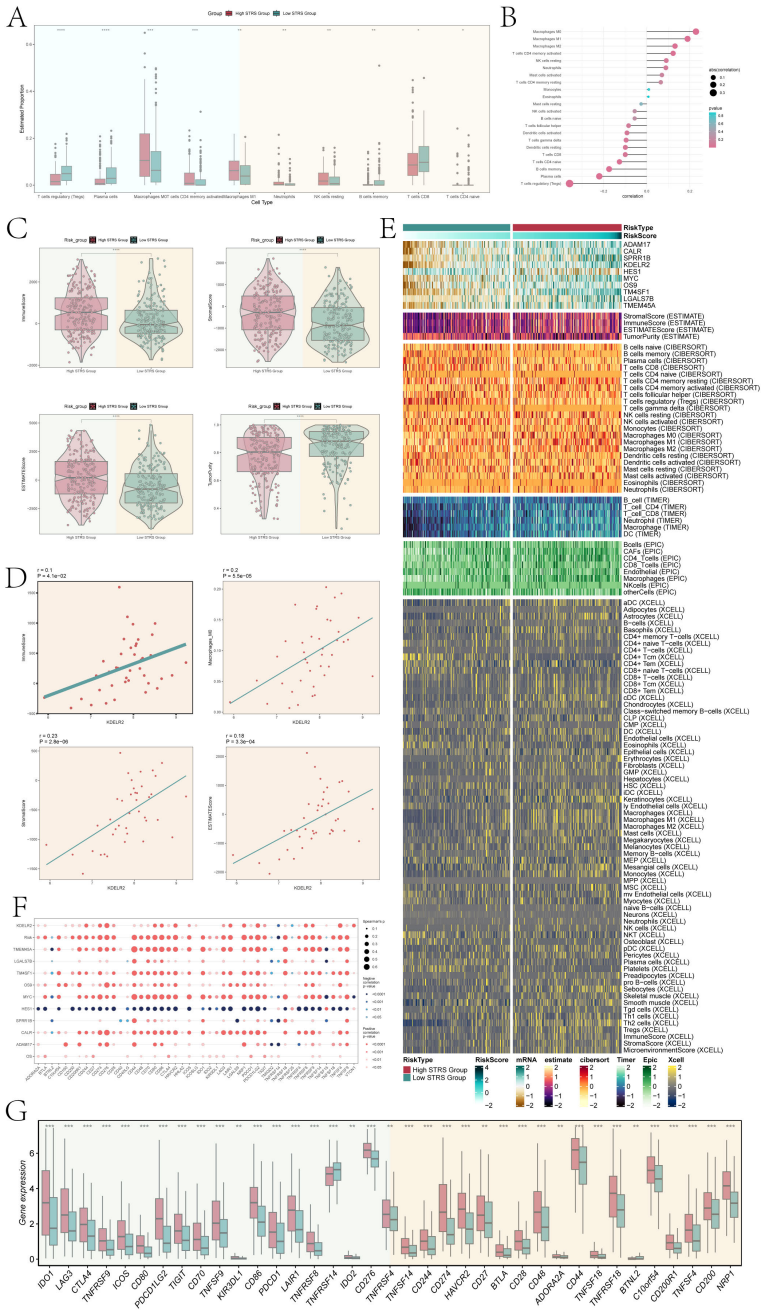


FIGURE 10
Infiltration of immune cells in high STRS group and low STRS group. **(A)** The box plot showed estimated proportion of immune cells that were statistically different between high STRS group and low STRS group. **(B)** Lollipop chart showed the correlation between different immune pathways and risk scores, with bubble size representing the abs(correlation) and color indicating p-value. **(C)** Immune-score, stromal-score, ESTIMATE-score, tumor-purity between high STRS group and low STRS group. **(D)** The scatter diagram showed the correlation among *KDELR2* and immune-score, macrophages-M0, stromal-score, ESTIMATE-score. **(E)** The heatmap showed risk scores for different immune cells in the high STRS group and the low STRS group. **(F)** The bubble plot showed the degree of association between risk genes and immune checkpoints. **(G)** The box plot showed immune checkpoints with statistical differences in the high STRS group and the low STRS group. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ and ****$P < 0.0001$.

mainly involvedin the cytoplasmic translation, protein folding, chaperone-mediated protein folding and signaling pathways like ribonucleoprotein complex biogenesis. Studies had shown that there was a special class of ribosomes in cancer cells, which could change cell progression and metabolic reprogramming, accelerate carcinogenic translation, and change cell function (61). In other words, bladder cancer cells might develop more rapidly under the support of ribosomes, and correspondingly, their malignancy might be higher. In addition, protein folding tends to have a relationship with the endoplasmic reticulum. When the body's ability to tolerate ERS was improved, the survival ability of cancer cells will become stronger. Similarly, the ability of immunosuppression, angiogenesis and drug resistance would be improved accordingly (62). To further enhanced our understanding of ribosome function and protein

10.3389/fimmu.2024.1485109

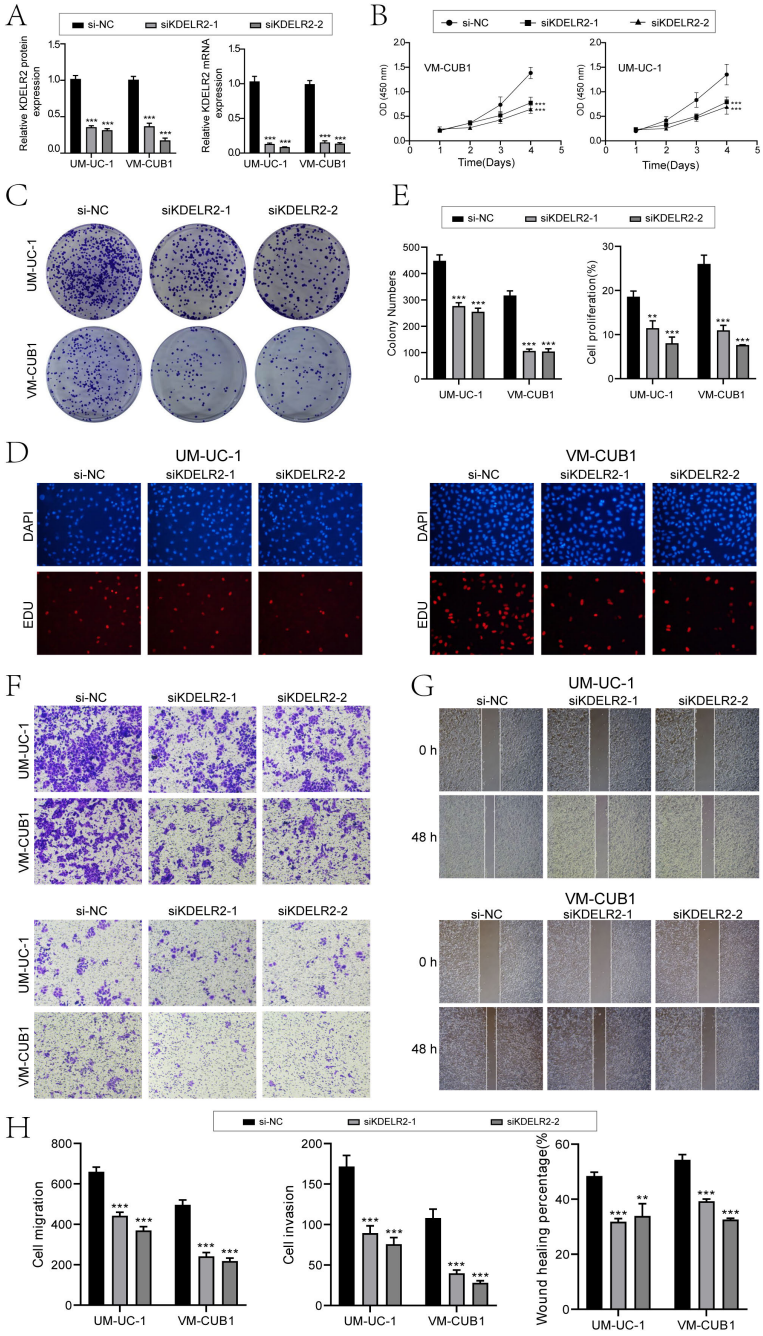


FIGURE 11
In vitro experiments confirmed the effects of *KDELR2* knockdown. **(A)** The bar charts depicted the altered patterns of gene-encoded protein (left) and gene RNA (right) expression in UM-UC-1 and VM-CUB1 cell lines, comparing three groups: si-NC, si*KDELR2*-1, and si*KDELR2*-2. Following targeted *KDELR2* knockdown, notable reductions in both mRNA and protein abundance levels were evident. **(B)** The line plot showed the longitudinal growth of three distinct groups across two cell lines. **(C)** Colony-formation assay revealed a significant reduction in cell viability subsequent to *KDELR2* knockdown, in contrast to the unaltered control group. **(D)** The EDU staining assay confirmed that *KDELR2* knockdown exerted an inhibitory effect on cell proliferation. **(E)** The bar plots showed the colony numbers and cell proliferation of three groups in two cell lines. **(F)** The transwell assay assessed the migratory and invasive potential of three distinct groups across two cellular lines, offering quantitative insights into their motility and aggressiveness. **(G)** Post-treatment migration capacity of TCs was quantitatively assessed using wound healing assays. **(H)** *KDELR2* knockdown led to a statistically significant decrease in cell migration, invasion, and wound healing capacities, as evident from bar graph analysed. $^{**}P < 0.01$, $^{***}P < 0.001$.

folding, we focused on their critical roles in cancer drug resistance and tumor evolution. Ribosomes played a vital role in maintaining cellular proteostasis, and alterations in ribosomal biogenesis enhanced protein synthesis capacity, allowing cancer cells to adapt to therapeutic pressure (63). Additionally, the involvement of chaperone proteins and the accumulation of misfolded proteins triggered the unfolded protein response (64), activating autophagy or heat shock protein-mediated survival pathways, thereby promoting resistance to chemotherapy and targeted therapies. Defects in protein folding also led to the accumulation of

mutations, exacerbating tumor heterogeneity and evolution (65), and affected tumor antigen presentation, potentially resulting in immune evasion. Moreover, ribosomal alterations impacted translational fidelity, increasing genetic and phenotypic diversity, further driving tumor adaptation to the microenvironment and therapeutic interventions (66). In summary, ribosome and protein folding and other related biological processes played an important role in tumor development, which can accelerate the proliferation and metastasis of TCs. In simpler terms, the progression of bladder cancer could be hindered by targeting specific factors like ribosomes and the process of protein folding.

To further explore the interaction between C1 *S100A8*+ TCs and other cells, CellChat found that C1 *S100A8*+ TCs acted on myeloid cells through the MIF-(CD74+CXCR4) and MIF-(CD74 +CD44) signaling pathways. MIF was a representative pro-inflammatory factor that played a role in regulating immune responses. Studies had shown that MIF expression is significantly increased in a variety of tumors (67). High expression of MIF had been observed to promote tumor progression and metastasis, stimulate angiogenesis, and create an immunosuppressive microenvironment conducive to tumor development. C1 *S100A8*+ TCs acted on myeloid cells CD74-CXCR4 and CD74-CD44 receptors through MIF ligands, which might promote the transformation of normal myeloid cells into cancer-related myeloid cells and inhibit the body's normal anti-tumor immune effect. Therefore, targeting MIF in TCs could be used as a potential strategy to treat tumors (67).

To further elucidate the mechanism of carcinogenesis in the subpopulation of C1 *S100A8*+ TCs, we analyzed the TFs and selected the top five TFs according to the expression active, namely HES1, AHR, TBL1XR1, IRF4 and SPI1.We could find that HES1 and AHR were significantly expressed in C1 subpopulation. HES1 had a complex relationship with various pathways, which had the potential to trigger cellular metamorphosis and enhance its invasive capabilities, while also playing a significant role in the differentiation, proliferation, and immune suppression of cancer cells (68). Furthermore, research had indicated a strong correlation between HES1 and the stem-like properties of TCs, as well as their ability to spread and develop resistance against drugs (69). AHR activated a specific gene, *CYP1A1*, which caused many oncogenic genes to combine with DNA to form cancer-promoting combinations, thus promoting the development of cancer (70). The over-expression or abnormal activation of AHR or its endogenous agonists in the TME, and these regulatory abnormalities promoted the immune escape of tumors. In summary, the above studies provided innovative prospects for future immunotherapeutic interventions in bladder cancer.

In addition, to enhance the chances of survival and extend the lifespan of individuals and improve the quality of life of patients, our research group constructed a bladder cancer prognostic risk model based on the top 10 marker genes of C1 *S100A8*+ TCs, and evaluated the prognosis of patients characterized by high expression of *S100A8* in the TCGA cohort. We observed that a higher high STRS group risk score, often symbolizing higher mortality, had a worse prognosis than low STRS group. According to research findings, *KDELR2* had been found to contribute to the advancement of bladder cancer and was often associated with unfavorable outcomes in individuals who had been diagnosed with cancer (71). In vitro experiments had also proved that when *KDELR2* was highly expressed, the proliferation, migration and invasion ability of TCs were significantly improved, while when *KDELR2* was knocked down, it had the opposite effect. This was consistent with previous studies (5, 72). Therefore, the potential of targeting *KDELR2* should be acknowledged as a viable strategy in the management of patients diagnosed with bladder cancer.

Additionally, the results of immune cell infiltration analysis showed that the high STRS group had higher macrophage infiltration levels, and the immune-score and stromal-score were significantly higher than those of the low STRS group. Macrophages have two phenotypes: "anti-tumor M1" and "pro-tumor M2", which respectively performed two opposite biological effects (73, 74). Therefore, reshaping the macrophage phenotype and promoting the transformation of the M0 phenotype to the M1 direction was crucial to resist tumor immunosuppression. The higher macrophage infiltration level in the high STRS group was obviously beneficial for macrophage-targeted therapy in bladder cancer patients to help promote the normal anti-tumor effect of macrophages. In addition, with the increase of *KDELR2* expression level, the macrophage M0 immune infiltration status became more obvious, and both immune and stromal scores increased. *KDELR2* could enhance tumor immune infiltration, inhibit anti-tumor immune response, and lead to low response to immunotherapy in patients (71). *KDELR2* was instrumental in promoting macrophage infiltration into the TME of bladder cancer, where it influenced the polarization of macrophages towards the M2 phenotype (75). This M2 polarization was associated with immunosuppression, increased angiogenesis, and enhanced tumor survival, creating a favorable environment for tumor progression (76). By modulating cytokine and chemokine production, *KDELR2* helped recruit macrophages that secreted factors supporting tumor growth. Additionally, the presence of M2 macrophages facilitated immune evasion, allowing cancer cells to escape detection by the immune system. Thus, targeting *KDELR2* presented a promising therapeutic strategy to disrupt these pro-tumorigenic processes and improve treatment outcomes in bladder cancer.

To further elucidate the modulation of macrophage phenotypes for the purpose of anti-tumor therapy, particularly focusing on the conversion of M2 macrophages to the pro-inflammatory M1 phenotype, previous studies showed promise in reprogramming macrophages toward an anti-tumor state using Toll-Like Receptor agonists, such as MPLA (77), and CSF1R inhibitors, such as pexidartinib (78, 79). Additionally, researchers explored targeting the STAT3 (80) and PI3K/AKT signaling pathways (81) to suppress M2 polarization while promoting M1 activity. Cytokine-based therapies, including IFN-γ and IL-12 (82), were also employed to enhance the M1 phenotype. Our findings on the role of *KDELR2* in macrophage polarization potentially complemented these therapeutic strategies, as *KDELR2* might have represented a novel target for further enhancing macrophage reprogramming. Future studies could have explored combining *KDELR2* targeting with these existing therapies to improve the efficacy of macrophage-based immunotherapies in bladder cancer.

In our study, although we made every effort to conduct a thorough analysis and carefully selected the study cohort, we acknowledged several limitations. First, the small sample size might have led to an erroneous association between the investigated samples and the target genes, potentially affecting the accuracy of our assessment. Second, bladder cancer was composed of various subpopulations. Due to the relative rarity of bladder squamous cell carcinoma and adenocarcinoma, we focused primarily on urothelial carcinoma in our study. This focus might have limited the generalizability of our findings, as different subpopulations of bladder cancer might exhibit distinct biological behaviors and treatment responses. Lastly, our experimental investigation into the interactions between tumor cells and macrophages was not sufficiently in-depth. The complexity of the TME, especially the interactions between tumor cells and immune cells, was a current research hotspot. In bladder cancer, the cellular components of the TME, such as tumor-associated macrophages and cancer-associated fibroblasts, co-evolve with tumor cells, contributing to tumor heterogeneity and promoting tumor progression and drug resistance. Therefore, future studies need to focus more on these interactions and how they influence tumor progression and treatment response.

In future research, we plan to start with *KDELR2* and further investigate macrophage-targeted therapeutic strategies for bladder cancer. By exploring *KDELR2*'s mechanisms in bladder cancer, we aim to develop novel therapeutic strategies to improve patient prognosis. Additionally, we will explore combining different TME-targeting strategies into a rational approach that offers better efficacy with fewer side effects, providing more treatment options for bladder cancer patients.

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 authors.

Author contributions

ZZ: Data curation, Methodology, Visualization, Writing – original draft, Writing – review & editing. HJ: Data curation, Methodology, Visualization, Writing – original draft. ZS: Data curation, Methodology, Validation, Visualization, Writing – review & editing. YL: Data curation, Methodology, Validation, Visualization, Writing – review & editing. LL: Data curation, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

Thanks to everyone especially Figdraw (ID: PIRIU6b24e) who have contributed to this article.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1485109/full#supplementary-material

SUPPLEMENTARY FIGURE 1
The heatmaps showed the scores of CNV for ECs, EPCs, and six tumor cell clusters. **(A)** The heatmaps showed the scores of CNV for ECs and EPCs. **(B)** The heatmaps showed the scores of CNV for ECs and six tumor cell clusters.

References

1. Dobruch J, Daneshmand S, Fisch M, Lotan Y, Noon AP, Resnick MJ, et al. Gender and Bladder Cancer: A Collaborative Review of Etiology, Biology, and Outcomes. *Eur Urol.* (2016) 69:300–10. doi: 10.1016/j.eururo.2015.08.037

2. Lenis AT, Lec PM, Chamie K, Mshs MD. Bladder Cancer: A Review. *Jama.* (2020) 324:1980–91. doi: 10.1001/jama.2020.17598

3. Whitmore WJ. Bladder cancer: an overview. *CA Cancer J Clin.* (1988) 38:213–23. doi: 10.3322/canjclin.38.4.213

4. Chang SS, Bochner BH, Chou R, Dreicer R, Kamat AM, Lerner SP, et al. Treatment of Non-Metastatic Muscle-Invasive Bladder Cancer: AUA/ASCO/ASTRO/SUO Guideline. *J Urol.* (2017) 198:552–59. doi: 10.1016/j.juro.2017.04.086

5. Wei H, Ma W, Lu X, Liu H, Lin K, Wang Y, et al. KDELR2 promotes breast cancer proliferation via HDAC3-mediated cell cycle progression. *Cancer Commun (Lond).* (2021) 41:904–20. doi: 10.1002/cac2.12180

6. Sun L, Shao W, Lin Z, Lin J, Zhao F, Yu J. Single-cell RNA sequencing explored potential therapeutic targets by revealing the tumor microenvironment of neuroblastoma and its expression in cell death. *Discovery Oncol.* (2024) 15:409. doi: 10.1007/s12672-024-01286-5

7. Zhao F, Hong J, Zhou G, Huang T, Lin Z, Zhang Y, et al. Elucidating the role of tumor-associated ALOX5+ mast cells with transformative function in cervical cancer progression via single-cell RNA sequencing. *Front Immunol.* (2024) 15:1434450. doi: 10.3389/fimmu.2024.1434450

8. Shao W, Lin Z, Xiahou Z, Zhao F, Xu J, Liu X, et al. Single-cell RNA sequencing reveals that MYBL2 in malignant epithelial cells is involved in the development and progression of ovarian cancer. *Front Immunol*. (2024) 15:1438198. doi: 10.3389/fimmu.2024.1438198

9. Zhang Y, Zhao Z, Huang W, Kim BS, Lin L, Li X, et al. Pan-Cancer Single-Cell Analysis Revealing the Heterogeneity of Cancer-Associated Fibroblasts in Skin Tumors. *Curr Gene Ther*. (2024). doi: 10.2174/0115665232331353240911080642

10. Nie W, Zhao Z, Liu Y, Wang Y, Zhang J, Hu Y, et al. Integrative Single-Cell Analysis of Cardiomyopathy Identifies Differences in Cell Stemness and Transcriptional Regulatory Networks among Fibroblast Subpopulations. *Cardiol Res Pract*. (2024) 2024:3131633. doi: 10.1155/2024/3131633

11. Ding Y, Zhao Z, Cai H, Zhou Y, Chen H, Bai Y, et al. Single-cell sequencing analysis related to sphingolipid metabolism guides immunotherapy and prognosis of skin cutaneous melanoma. *Front Immunol*. (2023) 14:1304466. doi: 10.3389/fimmu.2023.1304466

12. Chen N, Fan B, He Z, Yu X, Wang J. Identification of HBEGF+ fibroblasts in the remission of rheumatoid arthritis by integrating single-cell RNA sequencing datasets and bulk RNA sequencing datasets. *Arthritis Res Ther*. (2022) 24:215. doi: 10.1186/s13075-022-02902-x

13. 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

14. Ge Q, Zhao Z, Li X, Yang F, Zhang M, Hao Z, et al. Deciphering the suppressive immune microenvironment of prostate cancer based on CD4+ regulatory T cells: Implications for prognosis and therapy prediction. *Clin Transl Med*. (2024) 14:e1552. doi: 10.1002/ctm2.1552

15. Li X, Lin Z, Zhao F, Huang T, Fan W, Cen L, et al. Unveiling the cellular landscape: insights from single-cell RNA sequencing in multiple myeloma. *Front Immunol*. (2024) 15:1458638. doi: 10.3389/fimmu.2024.1458638

16. Zhou W, Lin Z, Tan W. Deciphering the molecular landscape: integrating single-cell transcriptomics to unravel myofibroblast dynamics and therapeutic targets in clear cell renal cell carcinomas. *Front Immunol*. (2024) 15:1374931. doi: 10.3389/fimmu.2024.1374931

17. Liu P, Xing N, Xiahou Z, Yan J, Lin Z, Zhang J. Unraveling the intricacies of glioblastoma progression and recurrence: insights into the role of NFYB and oxidative phosphorylation at the single-cell level. *Front Immunol*. (2024) 15:1368685. doi: 10.3389/fimmu.2024.1368685

18. Lin Z, Sui X, Jiao W, Chen C, Zhang X, Zhao J. Mechanism investigation and experiment validation of capsaicin on uterine corpus endometrial carcinoma. *Front Pharmacol*. (2022) 13:953874. doi: 10.3389/fphar.2022.953874

19. Lin Z, Li X, Shi H, Cao R, Zhu L, Dang C, et al. Decoding the tumor microenvironment and molecular mechanism: unraveling cervical cancer subpopulations and prognostic signatures through scRNA-Seq and bulk RNA-seq analyses. *Front Immunol*. (2024) 15:1351287. doi: 10.3389/fimmu.2024.1351287

20. Xing J, Cai H, Lin Z, Zhao L, Xu H, Song Y, et al. Examining the function of macrophage oxidative stress response and immune system in glioblastoma multiforme through analysis of single-cell transcriptomics. *Front Immunol*. (2023) 14:1288137. doi: 10.3389/fimmu.2023.1288137

21. Jin W, Zhang Y, Zhao Z, Gao M. Developing targeted therapies for neuroblastoma by dissecting the effects of metabolic reprogramming on tumor microenvironments and progression. *Theranostics*. (2024) 14:3439–69. doi: 10.7150/thno.93962

22. Zhao Z, Ding Y, Tran LJ, Chai G, Lin L. Innovative breakthroughs facilitated by single-cell multi-omics: manipulating natural killer cell functionality correlates with a novel subcategory of melanoma cells. *Front Immunol*. (2023) 14:1196892. doi: 10.3389/fimmu.2023.1196892

23. Huang W, Kim BS, Zhang Y, Lin L, Chai G, Zhao Z. Regulatory T cells subgroups in the tumor microenvironment cannot be overlooked: Their involvement in prognosis and treatment strategy in melanoma. *Environ Toxicol*. (2024) 39:4512–30. doi: 10.1002/tox.24247

24. Zhao ZJ, Zheng RZ, Wang XJ, Li TQ, Dong XH, Zhao CY, et al. Integrating Lipidomics and Transcriptomics Reveals the Crosstalk Between Oxidative Stress and Neuroinflammation in Central Nervous System Demyelination. *Front Aging Neurosci*. (2022) 14:870957. doi: 10.3389/fnagi.2022.870957

25. Tang C, Deng L, Luo Q, He G. Identification of oxidative stress-related genes and potential mechanisms in atherosclerosis. *Front Genet*. (2022) 13:998954. doi: 10.3389/fgene.2022.998954

26. Zhao Z, Luo Q, Liu Y, Jiang K, Zhou L, Dai R, et al. Multi-level integrative analysis of the roles of lncRNAs and differential mRNAs in the progression of chronic pancreatitis to pancreatic ductal adenocarcinoma. *BMC Genomics*. (2023) 24:101. doi: 10.1186/s12864-023-09209-4

27. Lin Z, Fan W, Yu X, Liu J, Liu P. Research into the mechanism of intervention of SanQi in endometriosis based on network pharmacology and molecular docking technology. *Med (Baltimore)*. (2022) 101:e30021. doi: 10.1097/MD.0000000000030021

28. Lin Z, Sui X, Jiao W, Wang Y, Zhao J. Exploring the mechanism and experimental verification of puerarin in the treatment of endometrial carcinoma based on network pharmacology and bioinformatics analysis. *BMC Complement Med Ther*. (2022) 22:150. doi: 10.1186/s12906-022-03623-z

29. Zou J, Lin Z, Jiao W, Chen J, Lin L, Zhang F, et al. A multi-omics-based investigation of the prognostic and immunological impact of necroptosis-related mRNA in patients with cervical squamous carcinoma and adenocarcinoma. *Sci Rep*. (2022) 12:16773. doi: 10.1038/s41598-022-20566-0

30. Lin Z, Fan W, Sui X, Wang J, Zhao J. Necroptosis-Related LncRNA Signatures for Prognostic Prediction in Uterine Corpora Endometrial Cancer. *Reprod Sci*. (2023) 30:576–89. doi: 10.1007/s43032-022-01023-9

31. Lin Z, Zou J, Sui X, Yao S, Lin L, Wang J, et al. Necroptosis-related lncRNA signature predicts prognosis and immune response for cervical squamous cell carcinoma and endocervical adenocarcinomas. *Sci Rep*. (2022) 12:16285. doi: 10.1038/s41598-022-20858-5

32. Zhao J, Zou J, Jiao W, Lin L, Wang J, Lin Z. Construction of N-7 methylguanine-related mRNA prognostic model in uterine corpus endometrial carcinoma based on multi-omics data and immune-related analysis. *Sci Rep*. (2022) 12:18813. doi: 10.1038/s41598-022-22879-6

33. Zhao Z, Li T, Dong X, Wang X, Zhang Z, Zhao C, et al. Untargeted Metabolomic Profiling of Cuprizone-Induced Demyelination in Mouse Corpus Callosum by UPLC-Orbitrap/MS Reveals Potential Metabolic Biomarkers of CNS Demyelination Disorders. *Oxid Med Cell Longev*. (2021) 2021:7093844. doi: 10.1155/2021/7093844

34. Wang Y, Zhao ZJ, Kang XR, Bian T, Shen ZM, Jiang Y, et al. lncRNA DLEU2 acts as a miR-181a sponge to regulate SEPP1 and inhibit skeletal muscle differentiation and regeneration. *Aging (Albany NY)*. (2020) 12:24033–56. doi: 10.18632/aging.104095

35. Zhao ZJ, Chen D, Zhou LY, Sun ZL, Wang BC, Feng DF. Prognostic Value of Different Computed Tomography Scoring Systems in Patients With Severe Traumatic Brain Injury Undergoing Decompressive Craniectomy. *J Comput Assist Tomogr*. (2022) 46:800–07. doi: 10.1097/RCT.0000000000001343

36. Zhao ZJ, Wei DP, Zheng RZ, Peng T, Xiao X, Li FS. The Gene Coexpression Analysis Identifies Functional Modules Dynamically Changed After Traumatic Brain Injury. *Comput Math Methods Med*. (2021) 2021:5511598. doi: 10.1155/2021/5511598

37. Zheng R, Zhuang Z, Zhao C, Zhao Z, Yang X, Zhou Y, et al. Chinese Admission Warning Strategy for Predicting the Hospital Discharge Outcome in Patients with Traumatic Brain Injury. *J Clin Med*. (2022) 11974,1-13. doi: 10.3390/jcm11040974

38. Yuan J, Zhang J, Luo Q, Peng L. Effects of nonalcoholic fatty liver disease on sarcopenia: evidence from genetic methods. *Sci Rep*. (2024) 14:2709. doi: 10.1038/s41598-024-53112-1

39. Tang L, Xu H, Wu T, Wu W, Lu Y, Gu J, et al. Advances in tumor microenvironment and underlying molecular mechanisms of bladder cancer: a systematic review. *Discovery Oncol*. (2024) 15:111. doi: 10.1007/s12672-024-00902-8

40. Yang F, Fang E, Mei H, Chen Y, Li H, Li D, et al. Cis-Acting circ-CTNNB1 Promotes beta-Catenin Signaling and Cancer Progression via DDX3-Mediated Transactivation of YY1. *Cancer Res*. (2019) 79:557–71. doi: 10.1158/0008-5472.CAN-18-1559

41. Paredes F, Williams HC, San MA. Metabolic adaptation in hypoxia and cancer. *Cancer Lett*. (2021) 502:133–42. doi: 10.1016/j.canlet.2020.12.020

42. Wu Q, You L, Nepovimova E, Heger Z, Wu W, Kuca K, et al. Hypoxia-inducible factors: master regulators of hypoxic tumor immune escape. *J Hematol Oncol*. (2022) 15:77. doi: 10.1186/s13045-022-01292-6

43. Zhao Y, Yang W, Zheng K, Chen J, Jin X. The role of BMI1 in endometrial cancer and other cancers. *Gene*. (2023) 856:147129. doi: 10.1016/j.gene.2022.147129

44. Xie J, Zhang W, Liang X, Shuai C, Zhou Y, Pan H, et al. RPL32 Promotes Lung Cancer Progression by Facilitating p53 Degradation. *Mol Ther Nucleic Acids*. (2020) 21:75–85. doi: 10.1016/j.omtn.2020.05.019

45. Hou G, Lu Z, Jiang J, Yang X. Ribosomal protein L32 enhances hepatocellular carcinoma progression. *Cancer Med*. (2023) 12:10791–803. doi: 10.1002/cam4.5811

46. Bi N, Sun Y, Lei S, Zeng Z, Zhang Y, Sun C, et al. Identification of 40S ribosomal protein S8 as a novel biomarker for alcohol–associated hepatocellular carcinoma using weighted gene co–expression network analysis. *Oncol Rep*. (2020) 44:611–27. doi: 10.3892/or.2020.7634

47. Tan WK, Maroni R, Offman J, Zamani SA, Sasieni PD, Fitzgerald RC. Targeted Screening for Barrett's Esophagus and Esophageal Cancer: Post Hoc Analysis From the Randomized BEST3 Trial. *Gastroenterology*. (2024) 167:798–800. doi: 10.1053/j.gastro.2024.04.030

48. Brouwer-Visser J, Huang GS. IGF2 signaling and regulation in cancer. *Cytokine Growth Factor Rev*. (2015) 26:371–77. doi: 10.1016/j.cytogfr.2015.01.002

49. Zhang L, Gao S, Shi X, Chen Y, Wei S, Mi Y, et al. NUPR1 imparts oncogenic potential in bladder cancer. *Cancer Med*. (2023) 12:7149–63. doi: 10.1002/cam4.5518

50. Dong Y, Ma WM, Shi ZD, Zhang ZG, Zhou JH, Li Y, et al. Role of NRP1 in Bladder Cancer Pathogenesis and Progression. *Front Oncol*. (2021) 11:685980. doi: 10.3389/fonc.2021.685980

51. Brewer JW. Regulatory crosstalk within the mammalian unfolded protein response. *Cell Mol Life Sci*. (2014) 71:1067–79. doi: 10.1007/s00018-013-1490-2

52. Clarke HJ, Chambers JE, Liniker E, Marciniak SJ. Endoplasmic reticulum stress in malignancy. *Cancer Cell*. (2014) 25:563–73. doi: 10.1016/j.ccr.2014.03.015

53. Dai X, Zhu M. Coupling of Ribosome Synthesis and Translational Capacity with Cell Growth. *Trends Biochem Sci*. (2020) 45:681–92. doi: 10.1016/j.tibs.2020.04.010

54. Xiong S, Dong L, Cheng L. Neutrophils in cancer carcinogenesis and metastasis. *J Hematol Oncol*. (2021) 14:173. doi: 10.1186/s13045-021-01187-y

55. Dyrskjot L, Hansel DE, Efstathiou JA, Knowles MA, Galsky MD, Teoh J, et al. Bladder cancer. *Nat Rev Dis Primers*. (2023) 9:58. doi: 10.1038/s41572-023-00468-9

56. Meeks JJ, Al-Ahmadie H, Faltas BM, Taylor JR, Flaig TW, DeGraff DJ, et al. Genomic heterogeneity in bladder cancer: challenges and possible solutions to improve outcomes. *Nat Rev Urol*. (2020) 17:259–70. doi: 10.1038/s41585-020-0304-1

57. Shukla S, Dalai P, Agrawal-Rajput R. Metabolic crosstalk: Extracellular ATP and the tumor microenvironment in cancer progression and therapy. *Cell Signal*. (2024) 121:111281. doi: 10.1016/j.cellsig.2024.111281

58. Vander HM, DeBerardinis RJ. Understanding the Intersections between Metabolism and Cancer Biology. *Cell*. (2017) 168:657–69. doi: 10.1016/j.cell.2016.12.039

59. Giampazolias E, Tait SW. Mitochondria and the hallmarks of cancer. *FEBS J*. (2016) 283:803–14. doi: 10.1111/febs.13603

60. Najafi M, Mortezaee K, Ahadi R. Cancer stem cell (a)symmetry & plasticity: Tumorigenesis and therapy relevance. *Life Sci*. (2019) 231:116520. doi: 10.1016/j.lfs.2019.05.076

61. Elhamamsy AR, Metge BJ, Alsheikh HA, Shevde LA, Samant RS. Ribosome Biogenesis: A Central Player in Cancer Metastasis and Therapeutic Resistance. *Cancer Res*. (2022) 82:2344–53. doi: 10.1158/0008-5472.CAN-21-4087

62. Cubillos-Ruiz JR, Bettigole SE, Glimcher LH. Tumorigenic and Immunosuppressive Effects of Endoplasmic Reticulum Stress in Cancer. *Cell*. (2017) 168:692–706. doi: 10.1016/j.cell.2016.12.004

63. Mills EW, Green R. Ribosomopathies: There's strength in numbers. *Science*. (2017) 358 608,1-8. doi: 10.1126/science.aan2755

64. Walter P, Ron D. The unfolded protein response: from stress pathway to homeostatic regulation. *Science*. (2011) 334:1081–86. doi: 10.1126/science.1209038

65. Hartl FU, Hayer-Hartl M. Converging concepts of protein folding *in vitro* and *in vivo*. *Nat Struct Mol Biol*. (2009) 16:574–81. doi: 10.1038/nsmb.1591

66. Truitt ML, Ruggero D. New frontiers in translational control of the cancer genome. *Nat Rev Cancer*. (2016) 16:288–304. doi: 10.1038/nrc.2016.27

67. Liang J, Lei K, Liang R, Huang J, Tan B, Lin H, et al. Single-cell RNA sequencing reveals the MIF-ACKR3 receptor-ligand interaction between iCAFs and tumor cells in esophageal squamous cell carcinoma. *Cell Signal*. (2024) 117:111093. doi: 10.1016/j.cellsig.2024.111093

68. Rani A, Greenlaw R, Smith RA, Galustian C. HES1 in immunity and cancer. *Cytokine Growth Factor Rev*. (2016) 30:113–17. doi: 10.1016/j.cytogfr.2016.03.010

69. Liu ZH, Dai XM, Du B. Hes1: a key role in stemness, metastasis and multidrug resistance. *Cancer Biol Ther*. (2015) 16:353–59. doi: 10.1080/15384047.2015.1016662

70. Murray IA, Patterson AD, Perdew GH. Aryl hydrocarbon receptor ligands in cancer: friend and foe. *Nat Rev Cancer*. (2014) 14:801–14. doi: 10.1038/nrc3846

71. Ma S, Sa L, Zhang J, Jiang K, Mi B, Shan L. KDELR2 as a diagnostic and prognostic biomarker of bladder urothelial carcinoma and its correlation with immune infiltration. *Genet Mol Biol*. (2023) 46:e20230002. doi: 10.1590/1678-4685-GMB-2023-0002

72. Mao H, Nian J, Wang Z, Li X, Huang C. KDELR2 is an unfavorable prognostic biomarker and regulates CCND1 to promote tumor progression in glioma. *Pathol Res Pract*. (2020) 216:152996. doi: 10.1016/j.prp.2020.152996

73. Oh MH, Sun IH, Zhao L, Leone RD, Sun IM, Xu W, et al. Targeting glutamine metabolism enhances tumor-specific immunity by modulating suppressive myeloid cells. *J Clin Invest*. (2020) 130:3865–84. doi: 10.1172/JCI131859

74. 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

75. Mantovani A, Biswas SK, Galdiero MR, Sica A, Locati M. Macrophage plasticity and polarization in tissue repair and remodelling. *J Pathol*. (2013) 229:176–85. doi: 10.1002/path.4133

76. Qian BZ, Pollard JW. Macrophage diversity enhances tumor progression and metastasis. *Cell*. (2010) 141:39–51. doi: 10.1016/j.cell.2010.03.014

77. Mata-Haro V, Cekic C, Martin M, Chilton PM, Casella CR, Mitchell TC. The vaccine adjuvant monophosphoryl lipid A as a TRIF-biased agonist of TLR4. *Science*. (2007) 316:1628–32. doi: 10.1126/science.1138963

78. Ries CH, Cannarile MA, Hoves S, Benz J, Wartha K, Runza V, et al. Targeting tumor-associated macrophages with anti-CSF-1R antibody reveals a strategy for cancer therapy. *Cancer Cell*. (2014) 25:846–59. doi: 10.1016/j.ccr.2014.05.016

79. Pyonteck SM, Akkari L, Schuhmacher AJ, Bowman RL, Sevenich L, Quail DF, et al. CSF-1R inhibition alters macrophage polarization and blocks glioma progression. *Nat Med*. (2013) 19:1264–72. doi: 10.1038/nm.3337

80. Kortylewski M, Kujawski M, Wang T, Wei S, Zhang S, Pilon-Thomas S, et al. Inhibiting Stat3 signaling in the hematopoietic system elicits multicomponent antitumor immunity. *Nat Med*. (2005) 11:1314–21. doi: 10.1038/nm1325

81. Kaneda MM, Messer KS, Ralainirina N, Li H, Leem CJ, Gorjestani S, et al. PI3Kgamma is a molecular switch that controls immune suppression. *Nature*. (2016) 539:437–42. doi: 10.1038/nature19834

82. Trinchieri G. Interleukin-12 and the regulation of innate resistance and adaptive immunity. *Nat Rev Immunol*. (2003) 3:133–46. doi: 10.1038/nri1001

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 11 December 2024
DOI 10.3389/fimmu.2024.1486652

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong, China

REVIEWED BY
Ming Yi,
Zhejiang University, China
Zhiling Yang,
Clinical Medical College and Affiliated Hospital of Chengdu University, China
Guang Lei,
Central South University, China

*CORRESPONDENCE
Xiuwen Wang
xiuwenwang12@sdu.edu.cn
Fangling Ning
ningfangling@126.com†These authors have contributed equally to this work

RECEIVED 26 August 2024
ACCEPTED 25 November 2024
PUBLISHED 11 December 2024

CITATION
Liu Z, Han L, Ji X, Wang X, Jian J, Zhai Y, Xu Y, Wang F, Wang X and Ning F (2024) Multi-omics analysis and experiments uncover the function of cancer stemness in ovarian cancer and establish a machine learning-based model for predicting immunotherapy responses.
Front. Immunol. 15:1486652.
doi: 10.3389/fimmu.2024.1486652

Multi-omics analysis and experiments uncover the function of cancer stemness in ovarian cancer and establish a machine learning-based model for predicting immunotherapy responses

Zhibing Liu[1,2†], Lei Han[3†], Xiaoyu Ji[4†], Xiaole Wang[1], Jinbo Jian[1], Yujie Zhai[1], Yingjiang Xu[5], Feng Wang[1], Xiuwen Wang[2*] and Fangling Ning[1*]

[1]Department of Oncology, Binzhou Medical University Hospital, Binzhou, Shandong, China, [2]Department of Oncology, Qilu Hospital of Shandong University, Jinan, Shandong, China, [3]Department of Reproductive Medicine, Binzhou Medical University Hospital, Binzhou, Shandong, China, [4]Department of Oncology, Huashan Hospital Fudan University, Shanghai, China, [5]Department of Interventional Vascular Surgery, Binzhou Medical University Hospital, Binzhou, Shandong, China

Background: The heterogeneity of cancer makes it challenging to predict its response to immunotherapy, highlighting the need to find reliable biomarkers for assessment. The sophisticated role of cancer stemness in mediating resistance to immune checkpoint inhibitors (ICIs) is still inadequately comprehended.

Methods: Genome-scale CRISPR screening of RNA sequencing data from Project Achilles was utilized to pinpoint crucial genes unique to Ovarian Cancer (OV). Thirteen publicly accessible OV transcriptomic datasets, seven pan-cancer ICI transcriptomic cohorts, and one single-cell RNA dataset from melanoma patients treated with PD-1 were utilized to scale a novel cancer stemness index (CSI). An OV single-cell RNA dataset was amassed and scrutinized to uncover the role of Small Nuclear Ribonucleoprotein Polypeptide E (SNRPE) in the tumor microenvironment (TME). Vitro experiments were performed to validate the function of SNRPE in promoting proliferation and migration of ovarian cancer.

Results: Through the analysis of extensive datasets on ovarian cancer, a specific gene set that impacts the stemness characteristics of tumors has been identified and we unveiled a negative correlation between cancer stemness, and benefits of ICI treatment in single cell ICI cohorts. This identified gene set underpinned the development of the CSI, a groundbreaking tool leveraging advanced machine learning to predict prognosis and immunotherapy responses in ovarian cancer patients. The accuracy of the CSI was further confirmed by applying PD1/PD-L1 ICI transcriptomic cohorts, with a mean AUC exceeding 0.8 for predicting tumor progression and immunotherapy benefits. Remarkably, when compared to existing

immunotherapy and prognosis markers, CSI exhibited superior predictive capabilities across various datasets. Interestingly, our research unveiled that the amplification of SNRPE contribute to remodeling the TME and promoting the evasion of malignant cells from immune system recognition and SNRPE can server as a novel biomarker for predicting immunotherapy response.

Conclusions: A strong relationship between cancer stemness and the response to immunotherapy has been identified in our study. This finding provides valuable insights for devising efficient strategies to address immune evasion by targeting the regulation of genes associated with cellular stemness.

KEYWORDS

cancer stemness, immunotherapy, ovarian cancer, TME, snrpe

Introduction

Continuous progress in immuno-oncology, including the creation of checkpoint inhibitors and CAR-T cell therapy, provide hopeful approaches for fighting tumors through the stimulation of the body's natural defense system (1). However, a significant number of patients do not experience the expected benefits from immunotherapy, highlighting the urgent need to identify the the population eligible for immunotherapy at this stage.

The efficacy of immunotherapy relies on a complex network involving multiple modulators, including the tumor immune microenvironment and genetic heterogeneity. Previous research has primarily focused on analyzing these factors through RNA sequencing of intact tumor tissue (2, 3). However, the variability in the tumor immune microenvironment across different cancer types and individuals, as well as the genetic heterogeneity of tumors, present challenges in accurately predicting patient responses to immunotherapy (4). While certain biomarkers like tumor mutation burden (TMB) have been linked to immunotherapy responses, they may not always accurately predict the effectiveness of specific immunotherapy treatments (5). This underscores the importance of developing robust markers and optimizing combinations of biomarkers to better stratify patients for optimal therapeutic outcomes.

Cancer stem cells contribute to the initiation, progression, and spread of tumors (6). Recently, research has shown a strong link between stem cell characteristics and the ability of cancer cells to evade the immune system and resist treatment (7). Previous study demonstrated a negative correlation between cancer stemness and immune cell infiltration in 21 solid cancers and indicated that high level of stemness have a negative impact on the effectiveness of ICI treatment across various cancer types (8–11). Nevertheless, the connection between tumor stemness and ICI response in ovarian cancer has been disregarded.

This study utilized integrative analyses of transcriptome and CRISPR cell line datasets to identify specific cancer stemness-related mRNAs of ovarian cancer. We also found a negative relationship between cancer-intrinsic variability, cancer stemness, and outcomes of ICI treatment in single-cell SKCM ICI cohorts (12). Subsequently, a CSI was developed by analyzing 13 ovarian cancer cohorts with 2407 patients. The accuracy of CSI in predicting immunotherapy response was assessed using 7 independent anti-PD-1/PD-L1 ICI cohorts with 929 patients and the submap algorithm. We observed a significant inverse correlation between CSI and intrinsic variations, including TMB, mutations, copy number variations, and Homologous Recombination Defects (HRD). Furthermore, combining CSI with TMB was found to improve the predictive accuracy of immunotherapeutic efficacy. Of note, a pivotal gene, SNRPE, was identified as having a promoting effect on tumor growth. This finding suggests that SNRPE could be a potential novel target for immunotherapy in the future. Collectively, our detailed analysis offers valuable insights into the role of cancer stemness in immunotherapy for ovarian cancer.

Methods

Acquisition and preprocessing of extensive ovarian cancer datasets

The Cancer Genome Atlas (TCGA) dataset on ovarian cancer RNA sequencing and survival data was retrieved from the UCSC Xena database (13). Additionally, 12 GEO cohorts focusing on ovarian cancer (GSE13876, GSE138866, GSE140082, GSE14764, GSE17260, GSE18520, GSE19829, GSE26712, GSE31245, GSE49997, GSE63885, GSE9891) were acquired, each containing detailed survival information.

Collection of immunotherapy-associated datasets

The research gathered various sets of data from groups of patients treated with anti-PD-L1/PD-1 medications to investigate the correlation between cancer stem cell characteristics and the effectiveness of immunotherapy. The cohorts included the following: Rose TL cohort (14) (GSE176307: ICB treated metastatic urothelial cancer), Jung H cohort (15) (GSE135222: anti-PD-1/PD-L1 treated non-small cell lung carcinoma), Riaz N cohort (16) (GSE91061: anti-CTLA4 and PD-1 treated melanoma), Liu/VanAllen cohort (phs000452.v3: anti-PD1/CTLA4-treated metastatic melanoma) from the dbGaP database, Necchi cohort (17) (IMvigor210: Atezolizumab treated advanced or metastatic urothelial carcinoma) obtained using the "IMvigor210CoreBiologies" R package, Wang GY cohort (anti-PD-1/PD-L1 treated melanoma), and Braun DA cohort (anti-PD-1 treated advanced clear cell renal cell carcinoma). Gene expression and clinical data were also gathered for these immunotherapy-treated datasets. The details of all cohorts used in this study can be found in Supplementary Table S1.

Collection of single cell datasets for OV and ICI-treated SKCM

Gene expression profiles of single cell OV dataset were preprocessed and retrieved from the GEO database with accession number GSE184880 (18). The dataset consisted of five non-malignant tissues and seven high-grade serous ovarian cancer tissues. Moreover, an examination was conducted on a melanoma cohort to explore the correlation between cancer cell stemness and the efficacy of immunotherapy. This cohort comprised data on both ICI response and single-cell RNA sequencing, sourced from GEO under accession number GSE115978 (12).

Identifying essential genes for OV

The CRISPR screening of OV cells at a genome-wide level was acquired through the DepMap portal (https://depmap.org/portal/download). Utilizing the CERES algorithm, dependency scores were computed for about 17,000 potential genes (19). Genes deemed essential demonstrated a CERES score below -1 in 75% of the OV cell lines (n = 73).

Development and validation of CSC prediction model

An innovative pipeline was developed to construct a predictive model for Cancer stem cells (CSC), illustrated in Figure 1A. Initially, by utilizing the CERES algorithm with cell line data, we pinpointed 687 mRNAs that displayed an association with the survival and progression of ovarian cancer cells. Subsequently, we computed mRNA stemness indices (10) across 12 GEO datasets, the TCGA-OV dataset and evaluated the relationship between total mRNA and mRNA expression-based stemness index (mRNAsi). We then identified the mRNAs showing significantly positive correlations in at least half of the cohorts (7 out of 13) as mRNAsi-associated mRNAs (Cor>0.2 and P<0.01), resulting in the discovery of 60 mRNAsi-associated mRNAs (Supplementary Tables S2, S3).

Finally, 8 ovarian cancer (OV) cohorts were used in the creation of a predictive model for CSCs. To accomplish this, a variety of machine learning techniques were employed, including random forest (RSF), elastic net (Enet), gradient boosting (GBM), ridge regression, Stepcox, plsRcox, CoxBoost, and SuperPC.

Prediction of immunotherapy outcomes using TIDE webserver

To evaluate the effectiveness of PD-1/CTLA4 immunotherapy, our first step involved the calculation of scores for tumor immune dysfunction and exclusion (TIDE). This analysis was performed using the adjusted expression data collected from patients with ovarian cancer. The resultant matrix of expression profiles was then submitted to the TIDE database website (http://tide.dfci.harvard.edu/) to assess the response of the patients (20). Next, we employed the submap algorithm available on the GenePattern website to determine the differences in response likelihood between the low- and high-CSI groups.

Identification of optimal therapeutics for low and high CSI groups and drug sensitivity analysis

By analyzing gene expression profiles, drug sensitivity prediction in cell lines was achieved using the 'oncoPredict' R package and the calcPhenotype method. To estimate drug IC50, a ridge regression model was developed utilizing gene expression profiles of cell lines acquired from GDSC through the pRRophetic algorithm (21).

Cell lines

Human ovarian cancer cell lines, specifically OVCAR-3, A2780, and SK-OV-3, were supplied by the Cell Bank of the Committee for Conservation of Typical Cultures, which is part of the Chinese Academy of Sciences. These cell lines were cultured using Dulbecco's Modified Eagle Medium (DMEM) from Gibco (New York, USA) and enriched with 10% fetal bovine serum. Furthermore, the culture medium was supplemented with 100 IU/mL penicillin and streptomycin, both of which were also procured from Gibco (New York, USA).

IHC

Following the removal of paraffin, the sections embedded in paraffin were subjected to a treatment with 3% hydrogen peroxide at

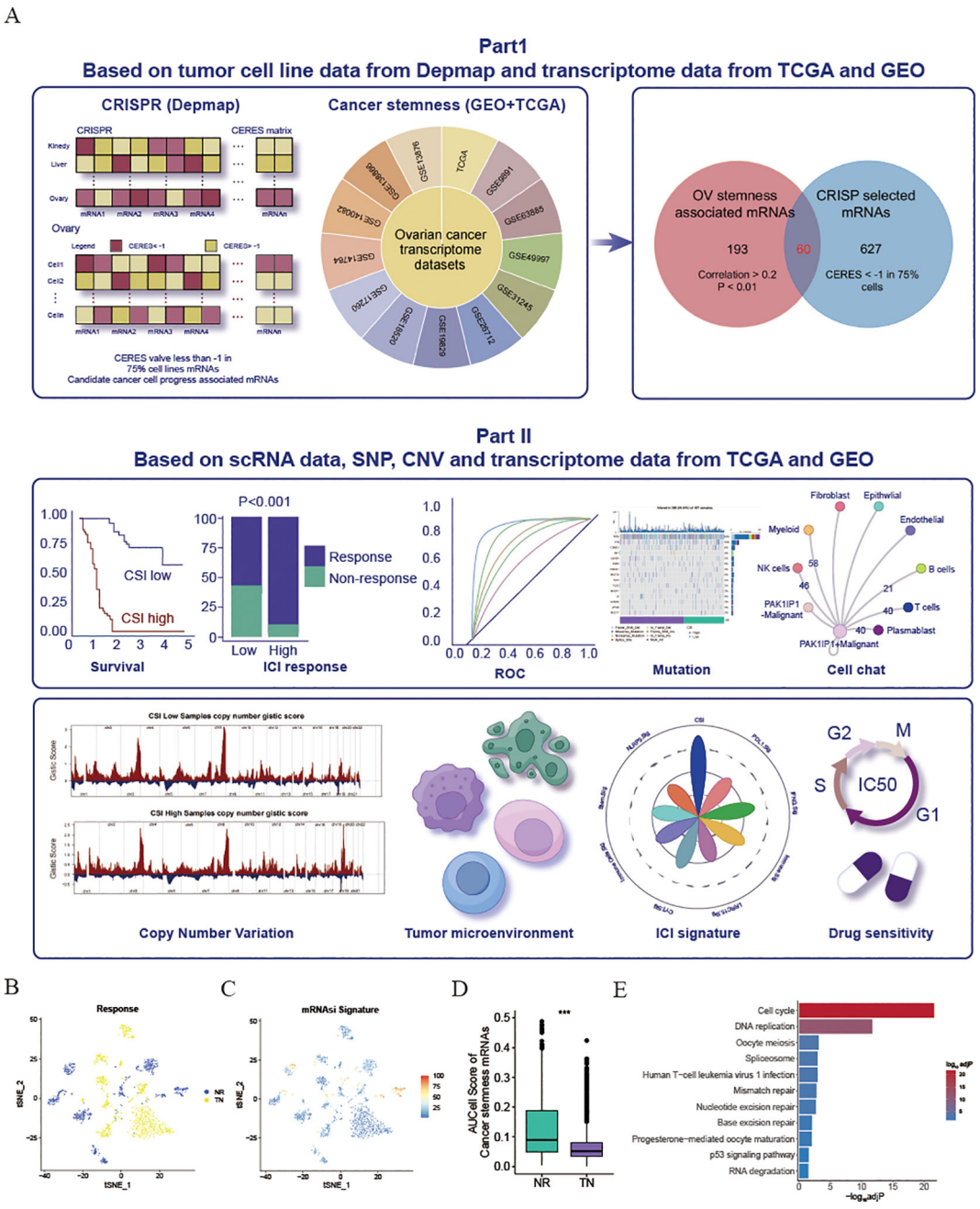


FIGURE 1
Exploration and validation of an inverse relationship between ovarian cancer stemness and ICI response. **(A)** Schematic representation of the process for identifying mRNA markers associated with cancer's intrinsic heterogeneity and stemness, and the development of predictive models using various machine learning techniques. **(B–D)** Visualization of t-Distributed Stochastic Neighbor Embedding (tSNE) plots for malignant cells in the SKCM dataset. **(B)** Malignant cells categorized by response phenotype in the tSNE plot. **(C)** tSNE plot depicting the AUCell scores of cancer stem cell (CSC)-related gene sets in malignant cells, with red indicating higher scores (indicative of high stemness) and blue representing lower scores (indicative of low stemness). **(D)** Box plot illustrating the distribution of AUCell scores across response phenotypes (non-responders vs. treatment-naïve patients) in the SKCM cohort. The median values are marked at the center of the box plot, with the box boundaries representing the 25% and 75% quantiles (Wilcoxon test; *** P < 0.001). Abbreviations: NR, non-responders; TN, treatment-naïve patients. **(E)** KEGG enrichment of cancer stemness associated mRNAs.

26°C for 10 minutes to suppress the activity of endogenous peroxidase. Next, the sections were blocked using 10% goat serum to avoid non-specific binding. Afterward, the sections were incubated overnight at 4°C with primary antibodies. Subsequently, rabbit secondary antibodies were applied to the sections, which were then stained using DAB.

Knockout and overexpression in ATC cell lines

Lentiviral vectors designed for the overexpression of SNRPE were sourced from Genechem in Shanghai, China. Stable transfection of cells with these SNRPE-overexpressing lentiviruses, along with the

corresponding control plasmids, was performed to induce puromycin resistance. Following the instructions provided by the manufacturer, selection of stable transfectants was carried out using 2 μg/mL puromycin over a period of 7 days to establish stable SNRPE-overexpressing cell lines. Biotend Co., Ltd. synthesized the siRNA targeting SNRPE. The siRNA, at a concentration of 50 nM, was transfected into cells using the Lipofectamine 3000 transfection kit provided by Thermo Fisher Scientific in Waltham, Massachusetts, USA, with a 24-hour incubation period.

Western blotting

To conduct western blot analysis, cells in culture were washed with ice-cold PBS before extracting total cell protein lysates at 4°C with RIPA lysis buffer (Beyotime, Shanghai, China) supplemented with 1% protease inhibitor cocktail (MedChemExpress, New Jersey, USA). Following centrifugation at 12,000 g for 20 minutes at 4°C, the supernatant containing proteins was gathered and mixed with loading buffer. The samples were then subjected to separation by 10% SDS-PAGE and transfer onto a PVDF membrane. The membrane was then blocked for 2 hours at room temperature with 5% skim milk before incubating overnight at 4°C with primary antibodies. After rinsing with Tris Buffered Saline, the membrane was exposed to secondary antibodies for detection of protein bands using enhanced chemiluminescence reagents (Beyotime, Shanghai, China). Antibodies used in the analysis included SNRPE (20407-1-AP, Proteintech, Wuhan, China) and GAPDH (60004-1-Ig, Proteintech, Wuhan, China).

Assessment of cell proliferation, colony formation, and migration abilities

To evaluate cell proliferation, 2×10^3 cells were introduced into each well of a 96-well plate and maintained for the required period. Afterward, each well was treated with 10 μl of CCK-8 reagent (Dojindo Molecular Technologies, Kumamoto, Japan) and left to incubate for one hour. The absorbance was then recorded at a wavelength of 450 nm (OD450) for further analysis.

To assess the ability of colonies to form, a range of 500 to 2000 cells were placed in each well of a 6-well plate and left to incubate for around one week. Upon detection of colonies with over 50 cells, they were treated with 0.2% crystal violet for a duration of 30 minutes. Following three rounds of washing with PBS, the colonies were both captured in pictures and tallied for measurement.

To evaluate the migratory potential of cells, a total of 40,000 cells were suspended in 200 μL of culture medium and placed in the upper compartment of Transwell plates from BD Biosciences. At the same time, 600 μL of culture medium with 10% FBS was introduced into the lower compartment. After an overnight incubation at 37°C, the cells located beyond the Transwell membrane were fixed using 4% paraformaldehyde for half an hour and subsequently stained with 0.25% crystal violet for an additional 30-minute period. Following the removal of cells from the interior of the chamber, the migrated cells outside the membrane were visualized and quantified.

Statistical analysis

We utilized the Wilcoxon test to assess various attributes of the high- SNRPE and low- SNRPE groups. The Chisq test was employed to scrutinize the variability in immunotherapy response among the high-CSI and low-CSI groups. The correlation between mRNA and mRNAsi was investigated through the calculation of Pearson's correlation coefficient. Kaplan-Meier survival analysis was performed to explore the connection between CSI, SNRPE, and survival, utilizing the log-rank test to determine the significance of observed distinctions. To assess the prognostic and immunotherapy advantages of CSI, time-dependent receiver operating characteristic (ROC) curves were generated with the assistance of the 'pROC' R package (22) being utilized. Key factors influencing immunotherapy efficacy were identified using xGboost, a scalable tree boosting system. Patients were grouped based on the optimal threshold established by the 'survminer' R package. Statistical significance was defined by a significance level of P or adjP < 0.05.

Results

Revelation of the link between cancer stemness and immunotherapy resistance through scRNA ICI cohort

Considering the potential influence of cancer stemness on the resistance to ICIs, a comprehensive analysis was conducted on 13 transcriptome datasets related to OV obtained from the GEO and TCGA databases. The mRNAsi was calculated for each patient (10). By utilizing the Pearson correlation coefficient, mRNAs that exhibited a significant relationship with mRNAsi across multiple samples (Cor>0.2 & P<0.01) were identified. Subsequently, 253 mRNAs that detected in more than 50% of the datasets (7 out of 13) were considered as tumor stemness-associated mRNAs. Moreover, to pinpoint crucial candidate genes involved in OV malignancy, an in-depth examination of CRISPR-based loss-of-function screens was undertaken on a global scale based on DepMap database. As a result, a total of 687 genes essential for the survival of 73 OV cell lines (CERE score < -1 in 75% OV malignant cells) were identified. Of these, 60 mRNAs were selected through the overlap of OV mRNAsi-associated mRNAs with the mRNAs highlighted by CRISP (Supplementary Table S3). All these genes associated with cancer stemness were chosen for further study. To validate the influence of cancer stemness associated genes on immunotherapy effectiveness, a previously published scRNA-seq dataset of PD1 ICI-treated patients with melanoma (SKCM) was initially employed to investigate the correlation between cancer stemness and ICI responses. After excluding individuals lacking data on malignant cells, a total of 23 patients from this cohort were included, comprising 10 non-responders (NR) and 13 treatment-naïve (TN) patients. Ideally, a

comparison of cancer stemness between responders (R) and NR to ICI treatment would have been preferred. However, the dataset available did not contain specific data on responders. Given that treatment-naïve patients may consist of both potential responders and non-responders, our next step was to compare the stemness levels between the NR group and the TN group, as previously described. As depicted in Figures 1B, C, the NR subcategory displayed a higher frequency of cancer cells with increased stemness rankings. Further investigation revealed that individuals from the NR subgroup had significantly higher levels of stemness ($P < 0.001$, Figure 1D), indicating an inverse relationship between cancer intrinsic driver and stemness with immune checkpoint inhibitor outcomes. Furthermore, we also found that these cancer stemness related genes significantly enriched Cell proliferation-related pathways, including Cell cycle and DNA replication (Figure 1E), indicating that tumor stemness-related genes may stimulate tumor cell proliferation.

Establishing the cancer stemness index through machine learning methodologies

To further develop a prediction model for CSI, nine machine learning algorithms were used with a combination of six GEO OV datasets and the OV TCGA dataset. Subsequently, we calculated the risk score for each sample in the eight cohorts, which included survival data, using these predictors. The performance was evaluated by determining the average C-index for each algorithm. Interestingly, most of these predictors showed a considerably high average C-index (Figure 2A). This finding can be partly attributed to the exceptional quality of our cancer stemness markers. Among all the models, random forest (RSF) demonstrated the highest level of precision (average C-index = 0.922, Figure 2A) and was chosen as the definitive CSI. Additionally, through univariate cox analysis, a significant correlation between high CSI in the seven cohorts and poor survival outcomes was established ($P<0.05$, Figure 2B).

The progress in next-generation sequencing and large-scale data mining technologies has facilitated the thorough investigation and advancement of gene expression-derived markers that are able to forecast prognosis results. To thoroughly assess the effectiveness of the CSI in comparison to alternative markers, we methodically compiled previously published markers from the past decade. A total of 79 markers were analyzed in this research (Supplementary Table S4). Notably, the reliability of the CSI in predicting survival outcomes exceeded that of all other models across eight different OV cohorts, achieving an average AUC > 0.9 in the mentioned cohorts (Figure 2C).

Given the importance of cancer stemness mRNAs in predicting the success of tumor immunotherapy, we utilized the submap algorithm available on GenePattern website to predict the

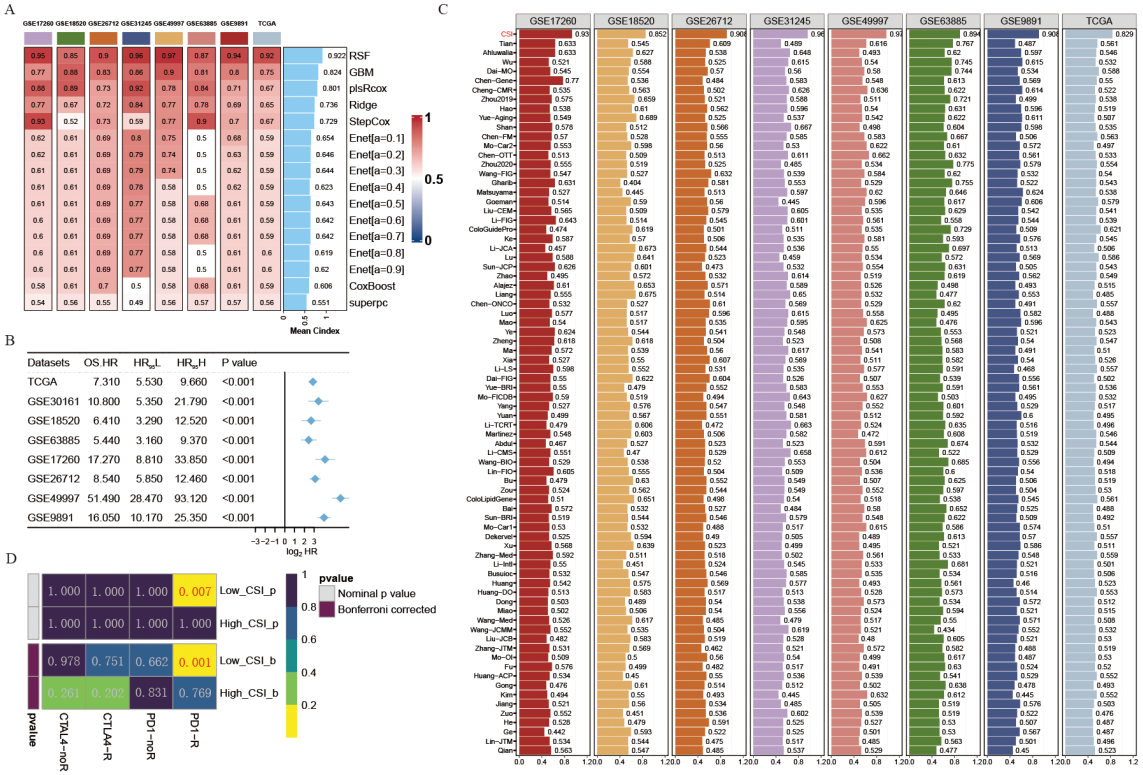


FIGURE 2
Development of a cancer stemness index utilizing extensive scRNA-seq and bulk RNA-seq datasets. **(A)** C-index of eight algorithms across eight validation cohorts. **(B)** Univariate Cox regression analysis of the RSF score in eight OV cohorts. **(C)** Estimated response rates to immune checkpoint inhibitors (PD-1/PD-L1) across different CSI groups (R: Response, NR: No Response). **(D)** ROC values comparing the predictive accuracy of the CSI and 79 other models for clinical status in eight OV cohorts.

probability of immune therapy response based on high- and low-CSI groups. A noteworthy finding was revealed when individuals from the low CSI group exhibited a significant reaction to PD-1 immunotherapy ($P = 0.007$, Bonferroni corrected $P < 0.001$, Figure 2D), demonstrating the remarkable predictive power of the CSI model in the context of PD-1 immunotherapy efficacy.

Collectively, these results suggest that CSI could function as predictive markers for the prognosis of ovarian cancer, with higher robust than other models. In addition, CSI can also be used as a prognostic indicator of PD1 immunotherapy response.

Cancer stemness index demonstrates predictive capabilities for immunotherapy outcomes

To further verify the predictive performance of CSI on the therapeutic effect of PD-1 ICIs, we collected different datasets associated with PD-1/PD-L1 immune checkpoint inhibitors. Our results consistently indicated that patients diagnosed with specific cancers (such as SKCM, UC, KIRC, or metastatic urothelial carcinoma) who had lower CSI scores experienced notably enhanced overall survival (OS) or progression-free survival (PFS) following immunotherapy compared to those with higher CSI scores (Figures 3A, B). This suggests that higher CSI scores may impede the benefits of PD-1 immunotherapy. Furthermore, the response to PD-1/PD-L1 ICI therapy varied between patients with high and low CSI scores. Individuals with higher CSI scores exhibited suboptimal response to the treatment, whereas over half of those with lower CSI scores responded positively (Figure 3C). More specifically, the group with higher CSI scores predominantly displayed no response (progressive disease or stable disease), whereas the group with lower CSI scores mainly demonstrated a response (complete response or partial response). Significantly, our analysis indicated that CSI serves as a reliable predictor of PD-1/PD-L1 ICI immunotherapy response, as demonstrated by the area under the curve (AUC) values. The AUC curve portrayed outstanding predictive performance, with an average AUC > 0.8 across the six cohorts examined (Figure 3D). Additionally, we performed further analysis using the IMvigor210 dataset and observed that even upon excluding samples with incomplete clinical data, CSI remained a robust predictor of immunotherapy outcomes. Intriguingly, it held greater significance compared to parameters such as PD-L1 expression in tumor cells (TC), immune phenotype, ECOG score, Stage, or tumor mutation burden (TMB), as indicated by a multivariate Cox regression analysis (refer to Figure 3E). To extend the clinical utility of our model, we explored the potential benefits of combining CSI with other commonly utilized markers of immunotherapy response. Specifically, we investigated the synergistic effects of CSI and TMB, a well-known indicator of immunotherapy effectiveness. Our findings revealed that patients exhibiting low CSI scores and high TMB levels experienced the most favorable outcomes with immunotherapy treatment, whereas those with elevated CSI scores demonstrated the least benefits from such therapies (Figure 3F). Additionally, we conducted a comparative analysis of CSI with established signatures for predicting immunotherapy response. Notably, CSI outperformed various signatures, including IFNG.Sig (23), Immune.Sig (23), ImmuneCells.Sig (16), PDL1.Sig (24), LRRC15.CAF.Sig (25), NLRP3.Sig (26), Stem.Sig (11), and CYT.Sig (27) in six PD1/PD-L1 immunotherapy cohorts, while the majority of these signatures exhibited optimal performance in only one or two cohorts (Figure 3G).

In conclusion, our research offers important insights into the predictive significance of CSI on immunotherapy results. Elevated CSI levels could potentially hinder the advantages of immunotherapy, whereas lower CSI levels have been linked to better survival rates and treatment responses. Integrating CSI with TMB could potentially improve the classification of patients for immunotherapy.

Intrinsic somatic mutations and copy number variation patterns of different CSI group patients

Examining somatic mutations and copy number variations (CNVs) as factors influencing both antitumor immunity and tumor advancement (28), we analyzed the most commonly mutated genes in various CSI categories (see Supplementary Figure S1A). Among OV patients, TP53 exhibits the greatest mutation rate, trailed by TTN and CSMD3. TMB stands for the tally of somatic non-synonymous mutations in a specific genetic area, usually expressed as mutations per megabase (mut/Mb). Prior research has demonstrated a negative correlation between TMB and the efficacy of immunotherapy (29). Notably, our investigation revealed that the TMB levels were greater in patients from the low-CSI category compared to those in the high-CSI group (Supplementary Figure S1B, $P<0.05$). Consistent with this, the SNV neoantigens and rate of nonsilent mutations were notably elevated in the low-CSI group compared to the high-CSI cohort. (Supplementary Figures S1C, D, $P<0.05$). In addition, Homologous Recombination Repair (HRR), a key mechanism for repairing DNA double strand breaks in cells, plays a critical role in maintaining the stability and integrity of the genome. HRD refers to conditions that occur when this repair mechanism is impaired, which may be due to genetic mutations in key repair proteins (such as BRCA1 and BRCA2) or dysfunction of other regulators. Prior research has indicated that targeting HRD defects can be an effective strategy for combating cancer. This includes not only conventional treatments like chemotherapy and radiotherapy, which cause DNA damage, but also newer approaches such as targeted therapies and immunotherapies (30, 31). Consistently, our research revealed that the Homologous Recombination Defects rating was notably elevated in the low CSI group compared to the high CSI group (Supplementary Figure S1E, $P <0.001$). Furthermore, to assess the prevalence of CNV across various CSI groups, we subsequently utilized the Genomic Identification of Significant Targets in Cancer (GISTIC) algorithm. Notably, we observed that the amplification GISTIC score was greater in patients from the low CSI group than in those from the high CSI group (Supplementary Figure S1F).

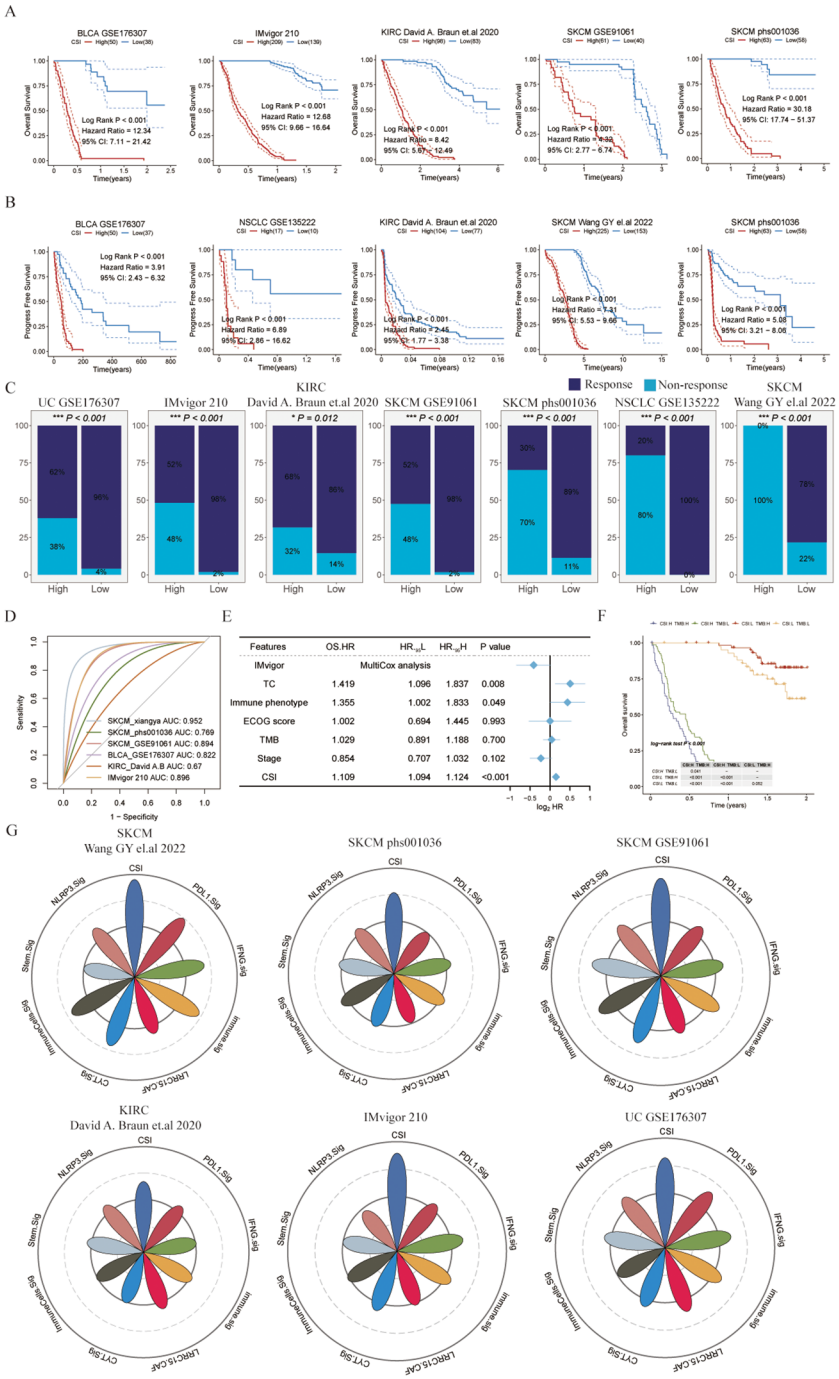


FIGURE 3
Evaluating the cancer stemness index as a potent prognostic tool for immunotherapy outcomes across various cancer types. **(A)** Kaplan-Meier survival curves depicting overall survival of patients undergoing immunotherapy in bladder cancer (UC, GSE176307), metastatic urothelial carcinoma (IMvigor210), kidney renal clear cell carcinoma (KIRC, David **(A)** Braun et al., (26)), and melanoma (SKCM, GSE91061 and phs000452.v3.p1). **(B)** Kaplan-Meier curves for progression-free survival of patients receiving immunotherapy in bladder cancer (UC, GSE176307), non-small cell lung cancer (NSCLC, GSE135222), KIRC (David **(A)** Braun et al.), and melanoma (SKCM, Wang GY et al., 2022 and phs000452.v3.p1). **(C)** Immunotherapy response rates in patients categorized by different CSI groups, with response defined as complete response (CR) or partial response (PR), and non-response as progressive disease (PD) or stable disease (SD). **(D)** ROC curves of the CSI for predicting response status in various immune checkpoint inhibitor (ICI) cohorts. **(E)** Multivariate Cox regression analysis of the CSI and clinical features in metastatic urothelial carcinoma (IMvigor210). **(F)** Kaplan-Meier survival curves for overall survival in different patient groups within the IMvigor210, with log-rank P values comparing each pair of groups displayed in the table. **(G)** Radar plot comparing the ROC values of eight ICI response prediction models and the CSI.

Collectively, these findings further elucidate the rationale behind the improved immunotherapeutic efficacy observed in patients with lower tumor stemness and demonstrate that CSI could serve as a prognostic indicator for predicting the therapeutic benefits of ICIs in ovarian cancer.

The amplification of SNRPE promotes the progress of ovarian cancer

To explore the clonal architecture and cell origins of ovarian malignant cells, we initially obtained a single-cell RNA profile from ovarian carcinoma. After filtering cells with a minimum expression of 200 genes and excluded those with over 20% expression of mitochondrial genes, we grouped the residual cells into eight major cell types based on traditional biomarkers. As described in Figures 4A, B all cells divided into diverse cell populations including B cells (MS4A1, and CD79A), Endothelial cells (PLVAP, and VWF), Epithelial cells (KRT18, and EPCAM), Fibroblasts (COL1A1, and ACTA2), Myeloid cells (LYZ, and CD14), NK cells (KLRD1, and TRDC), Plasmablast cells (JCHAIN, and MZB1), as well as T cells (CD3D, and IL7R).

Subsequently, the inferCNV algorithm was utilized to assess copy number variations (CNV) and clonality in ovarian malignant cells derived from epithelial cells (ECs). Among the 2807 ECs from THCA tissues, 6230 displayed high CNV scores, indicating malignancy (Figure 4C). Notably, amplifications in chromosomal 8q were identified as specific driving variations in ovarian cancer,

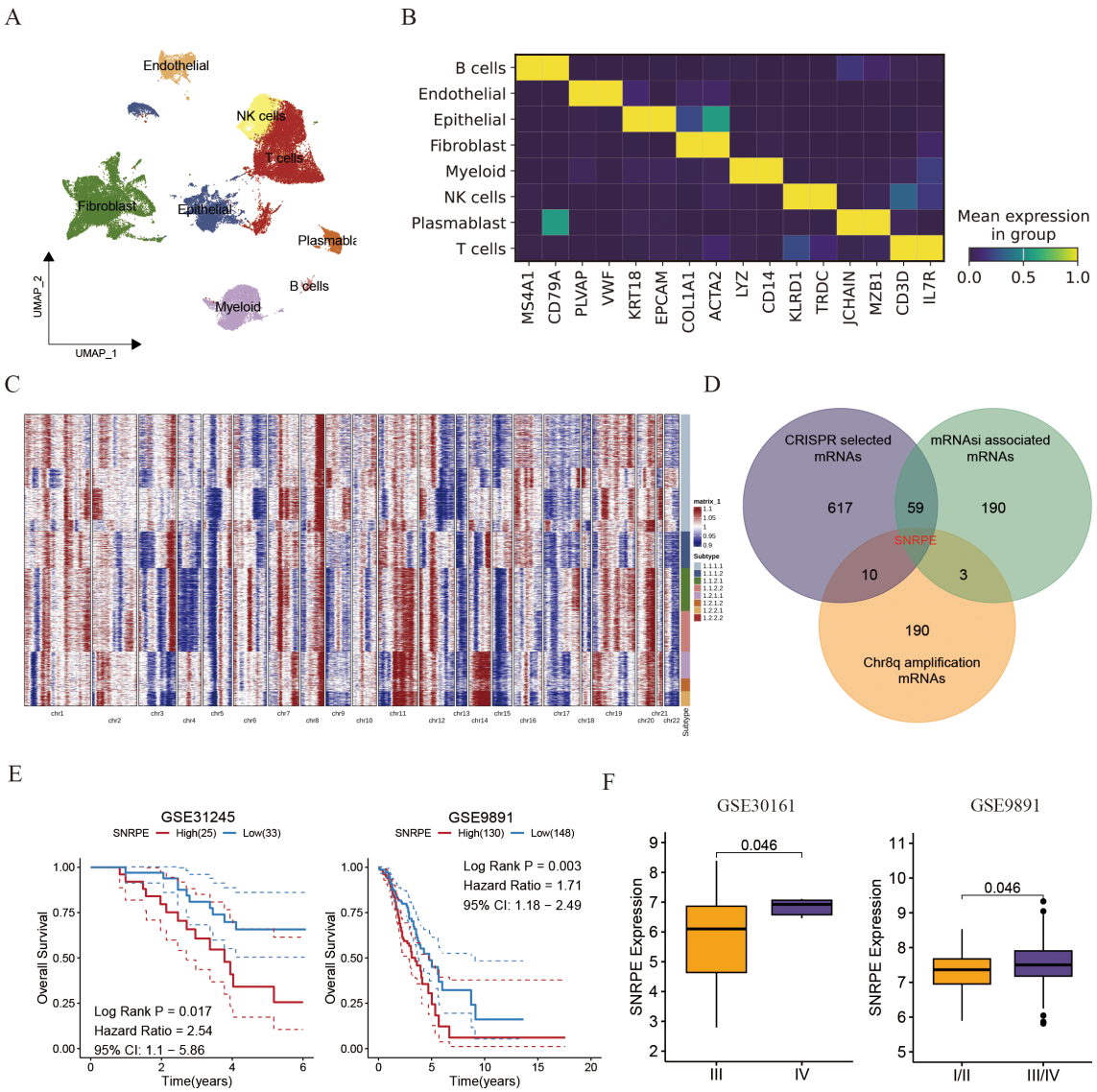


FIGURE 4
Single cell analysis uncovered the amplification of SNRPE promotes the progress of ovarian cancer. **(A)** UMAP plot showing the composition of 8 main subtypes derived from OV malignant cells. **(B)** Heatmap showing expression of each cell markers in each cell type. **(C)** Heatmap showing copy number variation of reference cells and epithelial cells. **(D)** Venn diagram showing the overlap between intratumoral heterogeneity (8q amplification) driven genes, cancer cell develop associated genes and cancer stemness associated genes. **(E)** Kaplan-Meier curves for overall survival of ovarian patients between SNRPE -high and -low groups (GSE31245 and GSE9891). **(F)** Box plots comparing SNRPE expression between early and advanced OV groups.

with SNRPE amplification in chromosomal 8q being linked to cancer stemness and essential for the survival of ovarian cell (Figure 4D). Prior studies have indicated SNRPE's role in promoting cell growth and advancing high-grade prostate cancer by controlling the expression of the androgen receptor [22740892]. Consistently, patients exhibiting high SNRPE expression consistently demonstrated poorer prognoses compared to those with low SNRPE expression (Figure 4E). Additionally, we observed that elevated SNRPE expression was specifically linked to tumor stage progression (Figure 4F).

Given that tumor heterogeneity and stemness play a crucial role in immune evasion and response to immunotherapy, we analyzed the TME among high and low SNRPE patients categorized based on the optimal threshold established by the 'survminer' R package. Tertiary lymphoid structures (TLS), which serve as germinal centers for immune cells, were analyzed for the expression of various interleukins associated with the formation of TLS and the activation of immune responses. Our findings indicate that the majority of these interleukins exhibited elevated expression levels in the group with high SNRPE. Specifically, we found that patients with low SNRPE expression showcased increased expression of IL10RA, IL10RB, IL18, IL21R, IL2RA, IL2RB, IL2RG and IL9R (Figure 5A). Additionally, numerous interferons along with their receptors (for example, IFNE, IFNG, IFNAR2, IFNGR2) as well as the majority of interleukins and their corresponding receptors were linked to immune-activating transcripts. Our discovery revealed that the levels of these interferons and receptors were elevated in the low SNRPE group, a pattern that aligns with the inverse relationship of interleukins within the tumor microenvironment (Figure 5B). Furthermore, recognizing the importance of immune checkpoint presence as a critical element in immunotherapy with ICIs, we carried out an additional investigation into the levels of immune checkpoints within two distinct groups. It is worth mentioning that the expression levels of several checkpoints (such as HAVCR2/TIM-3, ICOS, LAG3, LGALS9, PDCD1/PD-1, and PDCD1LG2/PD-L2) were significantly higher in the low SNRPE group compared to the high SNRPE group, indicating that higher expression SNRPE patients may benefit from immunotherapy benefit (Figure 5C). We also analyzed classical immune signatures in each sample, and we found that the most immune signatures were lower in high SNRPE group, suggesting that these immune cell and immune function were suppressed (Figure 5D). These results indicated that SNRPE might impact the effectiveness of immunotherapy by regulating the expression of immune checkpoints and immune microenvironment factors.

To further explore the mechanism of SNRPE on ovarian malignant cells, we then divided malignant tumor cells into SNRPE + malignant cells and SNRPE- malignant cells according to whether they expressed SNRPE. Through cell-cell interaction analysis, after eliminating common cell communication pairs, we found that SNRPE+ malignant cells had significantly higher specific cell communication with endothelial cells and fibroblasts than SNRPE- malignant cells (Figure 5E). Hence, we speculated that SNRPE+ malignant may promote tumor invasion and migration by promoting endothelial mesenchymal transformation. Interestingly, we found that NOTCH signaling pathway (DLL1_NOTCH1, JAG1_NOTCH2, NOTCH1_JAG1, and DLL4_NOTCH3), TGFβ1 signaling pathway (TGFB1_TGFBR3, TGFB1_TGFbeta receptor1, and TGFB2_TGFBR3) and VEGF signaling pathway (NRP1_VEGFA, and VEGFA_KDR) were significantly activated in cell communication pairs between SNRPE+ malignant cells and endothelial cells (Figure 5F). Then, we reclustered the tumor cells and found high expression of SNRPE in the C3 and C4 subpopulations (Supplementary Figures S2A, B). Using the HALLMARK pathway scoring, we identified that C3 can be defined as the EMT subpopulation, characterized by high activity in EMT signaling pathways. On the other hand, C4 can be defined as the proliferative subpopulation, characterized by the activation of proliferation-related pathways, including E2F_TARGETS, G2M_CHECKPOINT, MYC_TARGETS_V1, and MYC_TARGETS_V2 (Supplementary Figure S2C). Additionally, through analysis of cell-cell communications, we further discovered that the interaction intensity between SNRPE+ malignant cells and endothelial cells was significantly higher than that between SNRPE- malignant cells and endothelial cells (Supplementary Figure S2D). These findings further demonstrate how SNRPE enhances tumor cell proliferation and invasion. Furthermore, a univariate Cox regression analysis of the pan-cancer cohorts revealed that SNRPE expression was negatively correlated with prognosis across multiple cancer types, including Adrenocortical Carcinoma (ACC), Head and Neck squamous cell carcinoma (HNSC), Kidney Chromophobe (KICH), Kidney renal clear cell carcinoma (KIRC), liver hepatocellular carcinoma (LIHC), Brain Lower Grade Glioma (LGG), Lung adenocarcinoma (LUAD), and Pheochromocytoma and Paraganglioma (PCPG) (Supplementary Figure S3A).

Immunohistochemistry analysis revealed a marked increase in CSE1L expression in tumor tissues relative to the adjacent non-cancerous tissues (Figure 6A). To further validate SNRPE's oncogenic role in ovarian cancer, SNRPE was knocked down in OVCAR-3 and A2780 cell lines, effectiveness confirmed at the protein-level through western blot analyses (Figure 6B). Significantly, SNRPE knockdown markedly suppressed cell proliferation in both OVCAR-3 and A2780 cell lines (Figure 6C), underscoring SNRPE's contribution to promoting ovarian cancer cell growth. This was further supported by reduced clonogenic capacity in SNRPE knockdown cells compared to controls (Figure 6D), highlighting SNRPE's involvement in fostering growth in ovarian cancer cells. Notably, transwell migration assays revealed decreased cell migration upon SNRPE depletion in OVCAR-3 and A2780 cell lines (Figure 6E).

Concurrently, SNRPE overexpression in SK-OV-3 cells was validated through western blot analyses (Figure 6F), demonstrating its significant enhancement of proliferation (Figure 6G). Furthermore, SNRPE overexpression notably boosted clonogenic potential and significantly increased migration capability in SK-OV-3 cells (Figures 6H, I).

In summary, the amplification of SNRPE can drive the progression of ovarian malignant cells and it may serve as an efficient biomarker in foreseeing the prognosis and immunotherapy response in ovarian cancer.

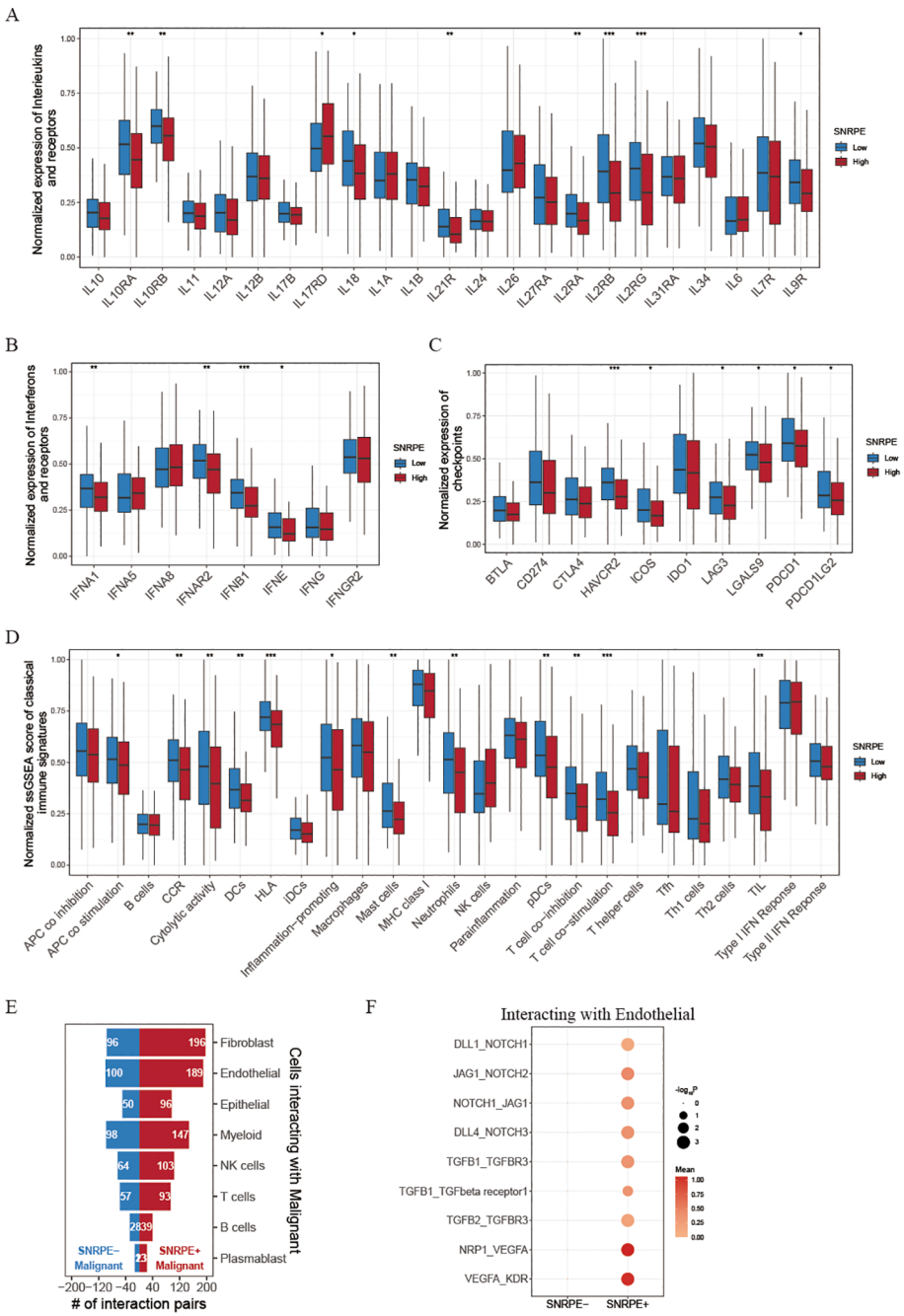


FIGURE 5
Investigating immune landscapes related to SNRPE expression. **(A, B)** Box plots comparing expression of interleukins, interferons and their receptors between low and high SNRPE groups. **(C)** Box plots for checkpoint expression comparison. **(D)** Normalized ssGSEA scores of classical immune signatures in the two groups. **(E)** Bar plots showing cell communication numbers between SNRPE- or SNRPE+ malignant cells and other cell types. **(F)** Dot plot depicting cell communication pairs of malignant cells (stratified by SNRPE status) with endothelial cells. Wilcoxon test; *P<0.05, **P<0.01, ***P<0.001.

Discussion

Tumor heterogeneity and stemness play key roles in influencing tumor immune evasion and the efficacy of immunotherapy. Numerous studies have explored the association between cancer stemness and the ICI response against tumors (6, 7). However, currently, there is no direct evidence linking tumor stemness to the response to ICI in OV. Furthermore, previous research has failed to acknowledge the predictive ability of tumor stemness in determining the response to ICI in OV (11).

In this research, we initially used a stemness index from a previous publication to identify mRNAs linked to tumor stemness by analyzing various omics data, such as transcriptome and CRISPR cell line data (10). Subsequently, to investigate the correlation between tumor stemness and immunotherapy, we examined a single-cell dataset of PD1/PD-L1 ICI-treated cells (12). It should

10.3389/fimmu.2024.1486652

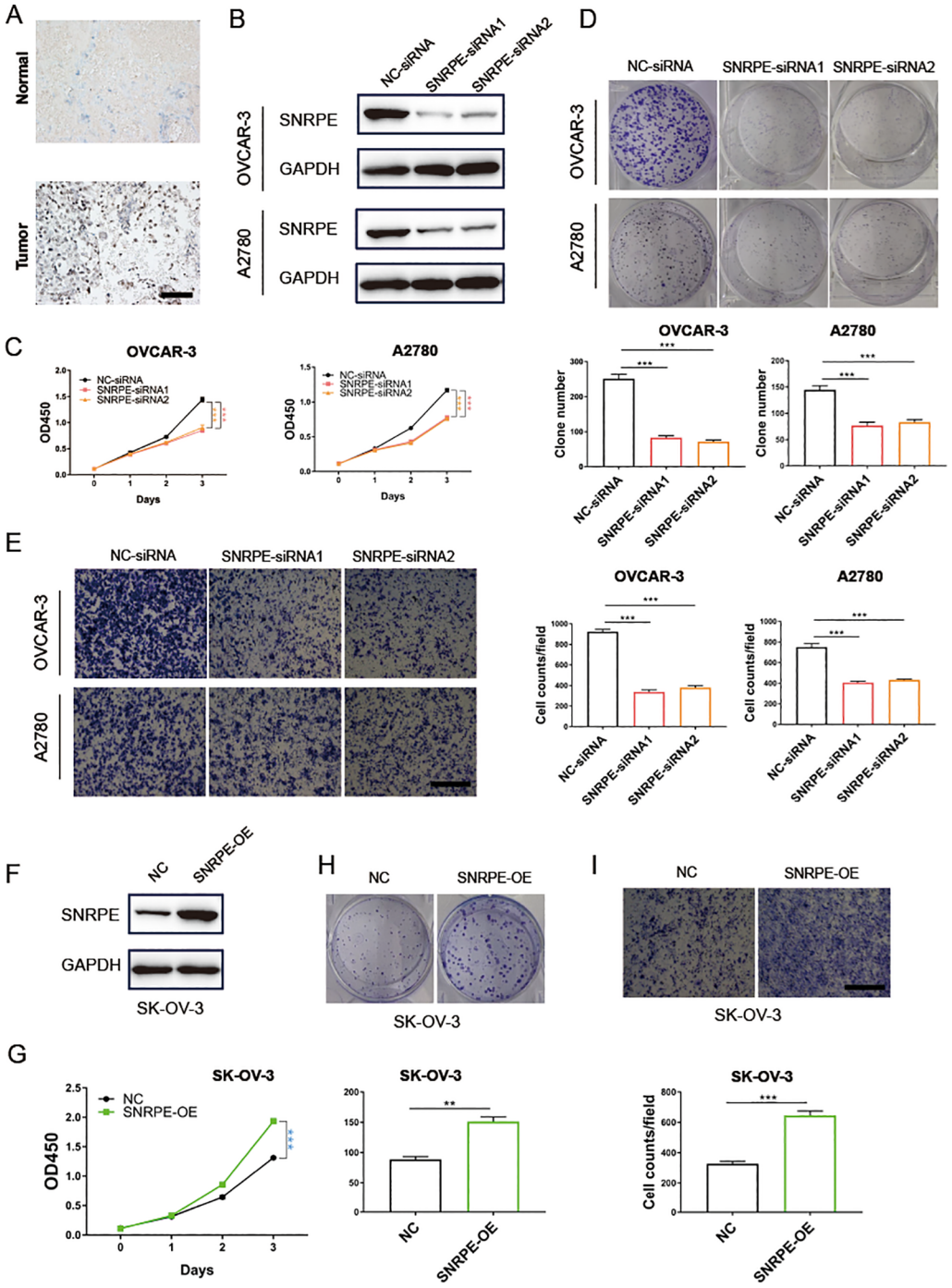


FIGURE 6
SNRPE promotes proliferation and migration of ovarian cancer *in vitro*. **(A)** Immunohistochemistry analysis revealed a marked increase in CSE1L expression in tumor tissues relative to the adjacent non-cancerous tissues. **(B)** Knockout of SNRPE in OVCAR-3 and A2780 cell lines validated by Western Blotting analysis. **(C)** The proliferative abilities of SNRPE knockout ovarian cancer cells detected with CCK8 assay. **(D)** The clone formation abilities of SNRPE knockout OVCAR-3 and A2780 cell lines. **(E)** The migrating abilities of SNRPE knockout OVCAR-3 and A2780 cell lines evaluated with transwell assay. Scale bar, 100 μm. **(F)** Overexpression of SNRPE validated by Western Blotting analysis in SK-OV-3 cells. **(G)** The proliferative abilities of SNRPE overexpressing SK-OV-3 cells detected with CCK8 assay. **(H)** The clone formation abilities assessed in SK-OV-3 cells upon SNRPE overexpression. **(I)** The migrating abilities of SK-OV-3 cells detected by transwell assay upon SNRPE overexpression. Scale bar, 100 μm. T test; **P < 0.01; ***P < 0.001.

be emphasized that we observed an inverse relationship between cancer stemness and the responses to ICI. CSCs are specialized cells that play a critical role in tumor initiation, progression, and spreading (6). Additionally, our KEGG enrichment analysis of genes associated with cancer stemness showed a significant enrichment in pathways such as Cell cycle, DNA replication, Mismatch repair, Nucleotide excision repair, and Base excision repair. A previous investigation highlighted that the abnormal

activation of the cell cycle pathway can lead to an increase in the expression of transcription factors like CDK, MKI67, and p53, which may support the occurrence and sustenance of tumor stemness (32). Besides, improved DNA repair ability readied cancerous cells for harsh surroundings (33). Our findings aligned with prior research and proposed that the gene set linked with cancer stemness which we revealed could strongly and uniquely correspond with cancer stemness.

Then, to assess the impact of tumor stemness genes on the prognosis of ovarian cancer patients, we employed various machine learning techniques to create a predictive model for CSI. This model's performance was then validated in eight separate datasets using a variety of assessment measures. Ultimately, the RSF model was chosen as the optimal CSI due to its increased stability and accuracy compared to the 79 previously established models. Notably, our research highlighted the CSI's effectiveness in predicting the response to PD-1 immune checkpoint inhibitors in ovarian cancer (11). Prior investigations have indicated a correlation between tumor stem cells and immune checkpoint inhibitor effectiveness. Building on these findings, we hypothesized that the CSI could be widely applicable for forecasting immunotherapy responses in various cancer types. As a result, we conducted an extensive analysis to evaluate the CSI's precision in predicting immunotherapy responses in other cancer types. Impressively, the CSI displayed exceptional accuracy in predicting ICI responses across diverse datasets utilizing bulk RNA-Seq data, with an average AUC exceeding 0.8. Additionally, our CSI demonstrated superior predictive capabilities compared to eight existing ICI response prognostic models. Notably, leveraging the IMvigor 210 dataset, we found that the CSI had better prognostic accuracy for post-immunotherapy patients compared to TMB. Our analysis also identified significantly different survival rates between low and high TMB patients. These results highlight the CSI's impressive predictive ability for both prognosis and immunotherapy outcomes in cases of ovarian cancer.

The quantity of neoantigens on tumor cells is determined by intrinsic variations within the tumor, which in turn impacts the immune system's ability to recognize and combat the tumor (34, 35). TMB serves as a crucial biomarker for predicting the effectiveness of immune checkpoint inhibitors. Clinical research has consistently shown that patients with high TMB tumors have a higher rate of clinical benefit when treated with these inhibitors (34, 36). Our research revealed a significant negative correlation between TMB levels, SNV neoantigens, nonsilent mutation rates, and CSI. Previous studies have indicated that HRD defects can be targeted by various anti-cancer treatments, including chemotherapy, radiotherapy, targeted therapies, and immunotherapies (30, 31). Notably, the level of HRD was found to be higher in the low CSI group as opposed to the high CSI group. Additionally, the GISTIC score was also observed to be higher in patients from the low CSI group compared to those in the high CSI group. Overall, CSI may provide valuable insights into the immune resistance mechanisms of high TMB tumors, underscoring its significance as a predictive biomarker for immune checkpoint inhibitors.

The tumor microenvironment has been established as a vital factor in the progression of various types of tumors. Tumor immune cell subpopulations vary among different tumor types and even among patients with the same pathological type (37). Through our selection process, we identified that SNRPE, correlated with the amplification of the long arm of chromosome 8, is linked to tumor stemness and encourages tumor cell proliferation. Previous research has demonstrated that SNRPE facilitates HCC tumorigenesis by regulating FGFR4 expression via alternative splicing mechanisms (38). In addition, we observed that patients with high SNRPE levels exhibited suppressed APC co-stimulation, Cytolytic activity, and HLA signatures, indicating that SNRPE can impede the activation and cytotoxic function of immune cells.

Tertiary lymphoid structures function as germinal centers for immune cells within the tumor microenvironment. In our study, we evaluated the expression levels of various interferons, interleukins, and their corresponding receptors that play roles in the formation of TLS (39). Our analysis demonstrated a considerable negative correlation between SNRPE expression and the levels of interleukins and interferons. Immune checkpoint inhibitors have emerged as a promising treatment strategy for advanced cancer. Higher levels of immune checkpoints facilitate tumor immune evasion and indicate a greater likelihood of response to these inhibitors. Additionally, we identified that several key immune checkpoints, such as TIM-3/HAVCR2, LAG3, PD-1/PDCD1, and PD-L2/PDCD1LG2, were significantly upregulated in the low SNRPE group. High expression of PD-L1 on tumor cells can bind to PD-L1 receptors on immune cells, initiating negative regulatory signals that impair T cell recognition of cancer cells, thereby allowing the tumor cells to evade the immune response (40). These results imply that patients with low SNRPE expression show an enhanced response to ICIs, likely due to the inhibition of TME components that support tumor progression, including the NOTCH1 signaling cascade, tumor necrosis factor (TGFB), and VEGF pathways. Furthermore, our findings confirmed that the overexpression of SNRPE notably boosted the proliferation and invasion abilities of ovarian cancer cells, indicating its potential as a therapeutic target for this type of cancer. In summary, these results suggest that SNRPE could affect the efficacy of immunotherapy by modifying the composition of the tumor microenvironment and influencing the recruitment of immune cells through its effects on chemokines and immune checkpoints.

While it is important to highlight the impressive accuracy of the CSI in predicting the success of immunotherapy, it is essential to acknowledge certain limitations in this study. The ability of the OV model to forecast outcomes of immunotherapy for ovarian cancer is based on projections generated by the submap algorithm, and the reliability of the CSI requires validation using real OV ovarian cancer immunotherapy groups.

Conclusion

In summary, we have developed a reliable and consistent signature of CSCs by conducting an integrated analysis of CRISPR OV cell lines, large-scale OV tissues, and single-cell cohorts. This signature allows for the classification of OV patients and the prediction of outcomes for immunotherapy. Our research

represents a groundbreaking exploration into the association between cancer stemness and immunotherapy in OV. It establishes a solid framework for understanding the importance of cancer stemness in immuno-oncology, clinical benefits, and practical implications. Based on our discoveries, this study enhances our comprehension of the link between cancer stemness and immunotherapy in OV, presenting new possibilities for treatment strategies.

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 authors.

Ethics statement

The studies involving humans were approved by Binzhou Medical University Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The human samples used in this study were acquired from primarily isolated as part of your previous study for which ethical approval was obtained. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements. Ethical approval was not required for the studies on animals in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

Author contributions

ZL: Conceptualization, Formal analysis, Funding acquisition, Investigation, Validation, Writing – original draft, Writing – review & editing. LH: Formal analysis, Validation, Writing – original draft, Writing – review & editing. XJ: Validation, Writing – original draft, Writing – review & editing. XLW: Writing – review & editing. JJ: Writing – review & editing. YZ: Writing – review & editing. YX: Writing – review & editing. FW: Writing – review & editing. XWW: Conceptualization, Project administration, Supervision, Writing – review & editing. FN: Conceptualization, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (NO.82200981); The Special Funds of Taishan Scholars Project of Shandong Province (grant number tsqn202312384); Natural Science Foundation of Shandong Province (ZR2022QH358, ZR2023MH222); Scientific Research Fund of Binzhou Medical University (BY2021KYQD34).

Acknowledgments

The authors would like to thank TCGA, GEO, DepMap and dbGaP projects for the data access.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1486652/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Analysis of intrinsic somatic mutations and copy number variation patterns in ovarian cancer (OV) patients with differing CSI levels. **(A)** Waterfall plot depicting the mutation frequency of the top 15 genes in OV. **(B-E)** Box plots comparing tumor mutation burden, single nucleotide variant neoantigens, nonsilent mutation rate, and homologous recombination defects between low- and high-CSI groups. **(F)** GISTIC scores for low- and high-CSI groups in OV patients. (Wilcoxon test; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$).

SUPPLEMENTARY FIGURE 2
SNRPE is highly expressed in EMT and proliferative tumor cells. **(A)** TSNE plot showing the composition of 5 main subtypes derived from OV malignant cells. **(B)** Dotplot showing the expression of SNRPE in malignant cell subtypes. **(C)** Dotplot showing the score of HALLMARK pathways in malignant cell subtypes. **(D)** There is a significant difference in cell-cell communication strength between SNRPE+ malignant and SNRPE- malignant cells with endothelial cells.

SUPPLEMENTARY FIGURE 3
Pan-cancer validation of the association between SNRPE and prognosis. **(A)** Univariate Cox regression analysis reveals a significant association between SNRPE and poor prognosis in various cancers.

References

1. Waldman AD, Fritz JM, Lenardo MJ. A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nat Rev Immunol*. (2020) 20:651–68. doi: 10.1038/s41577-020-0306-5

2. Yang Z, Wei X, Pan Y, Xu J, Si Y, Min Z, et al. A new risk factor indicator for papillary thyroid cancer based on immune infiltration. *Cell Death Dis*. (2021) 12:51. doi: 10.1038/s41419-020-03294-z

3. Pan Y, Sha Y, Wang H, Zhuang H, Ren X, Zhu X, et al. Comprehensive analysis of the association between tumor-infiltrating immune cells and the prognosis of lung adenocarcinoma. *J Cancer Res Ther*. (2020) 16:320–6. doi: 10.4103/jcrt.JCRT_954_19

4. Ge R, Wang Z, Cheng L. Tumor microenvironment heterogeneity an important mediator of prostate cancer progression and therapeutic resistance. *NPJ Precis Oncol*. (2022) 6:31. doi: 10.1038/s41698-022-00272-w

5. Hegde PS, Chen DS. Top 10 challenges in cancer immunotherapy. *Immunity*. (2020) 52:17–35. doi: 10.1016/j.immuni.2019.12.011

6. Chen P, Hsu W-H, Han J, Xia Y, DePinho RA. Cancer stemness meets immunity: from mechanism to therapy. *Cell Rep*. (2021) 34:1. doi: 10.1016/j.celrep.2020.108597

7. Bayik D, Lathia JD. Cancer stem cell–immune cell crosstalk in tumour progression. *Nat Rev Cancer*. (2021) 21:526–36. doi: 10.1038/s41568-021-00366-w

8. Shang S, Yang C, Chen F, Xiang RS, Zhang H, Dai SY, et al. ID1 expressing macrophages support cancer cell stemness and limit CD8+ T cell infiltration in colorectal cancer. *Nat Commun*. (2023) 14:766. doi: 10.1038/s41467-023-43548-w

9. Miranda A, Hamilton PT, Zhang AW, Pattnaik S, Becht E, Mezheyeuski A, et al. Cancer stemness, intratumoral heterogeneity, and immune response across cancers. *Proc Natl Acad Sci*. (2019) 116:9020–9. doi: 10.1073/pnas.1818210116

10. Lian H, Han YP, Zhang YC, Zhao Y, Yan S, Li QF, et al. Integrative analysis of gene expression and DNA methylation through one-class logistic regression machine learning identifies stemness features in medulloblastoma. *Mol Oncol*. (2019) 13:2227–45. doi: 10.1002/1878-0261.12557

11. Zhang Z, Wang Z-X, Chen Y-X, Wu H-X, Yin L, Zhao Q, et al. Integrated analysis of single-cell and bulk RNA sequencing data reveals a pan-cancer stemness signature predicting immunotherapy response. *Genome Med*. (2022) 14:45. doi: 10.1186/s13073-022-01050-w

12. Jerby-Arnon L, Shah P, Cuoco MS, Rodman C, Su M-J, Melms JC, et al. Lin: A cancer cell program promotes T cell exclusion and resistance to checkpoint blockade. *Cell*. (2018) 175:984–97. doi: 10.1016/j.cell.2018.09.006

13. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol*. (2020) 38:675–8. doi: 10.1038/s41587-020-0546-8

14. Rose TL, Weir WH, Mayhew GM, Shibata Y, Eulitt P, Uronis JM, et al. Fibroblast growth factor receptor 3 alterations and response to immune checkpoint inhibition in metastatic urothelial cancer: a real world experience. *Br J Cancer*. (2021) 125:1251–60. doi: 10.1038/s41416-021-01488-6

15. Jung H, Kim HS, Kim JY, Sun J-M, Ahn JS, Ahn M-J, et al. DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. *Nat Commun*. (2019) 10:4278. doi: 10.1038/s41467-019-12159-9

16. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, et al. Tumor and microenvironment evolution during immunotherapy with nivolumab. *Cell*. (2017) 171:934–949. e16. doi: 10.1016/j.cell.2017.09.028

17. Necchi A, Joseph R, Loriot Y, Hoffman-Censits J, Perez-Gracia J, Petrylak D, et al. Atezolizumab in platinum-treated locally advanced or metastatic urothelial carcinoma: post-progression outcomes from the phase II IMvigor210 study. *Ann Oncol*. (2017) 28:3044–50. doi: 10.1093/annonc/mdx518

18. Xu J, Fang Y, Chen K, Li S, Tang S, Ren Y, et al. Single-cell RNA sequencing reveals the tissue architecture in human high-grade serous ovarian cancer. *Clin Cancer Res*. (2022) 28:3590–602. doi: 10.1158/1078-0432.CCR-22-0296

19. Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, et al. Computational correction of copy number effect improves specificity of CRISPR–Cas9 essentiality screens in cancer cells. *Nat Genet*. (2017) 49:1779–84. doi: 10.1038/ng.3984

20. 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

21. Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting *in vivo* or cancer patient drug response and biomarkers from cell line screening data. *Briefings Bioinf*. (2021) 22:bbab260. doi: 10.1093/bib/bbab260

22. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf*. (2011) 12:1–8. doi: 10.1186/1471-2105-12-77

23. Vredevoogd DW, Kuilman T, Ligtenberg MA, Boshuizen J, Stecker KE, de Bruijn B, et al. Augmenting immunotherapy impact by lowering tumor TNF cytotoxicity threshold. *Cell*. (2020) 180:404–5. doi: 10.1016/j.cell.2020.01.005

24. Lim SA, Wei J, Nguyen T-LM, Shi H, Su W, Palacios G, et al. Lipid signalling enforces functional specialization of Treg cells in tumours. *Nature*. (2021) 591:306–11. doi: 10.1038/s41586-021-03235-6

25. Pan D, Kobayashi A, Jiang P, Ferrari de Andrade L, Tay RE, Luoma AM, et al. A major chromatin regulator determines resistance of tumor cells to T cell–mediated killing. *Science*. (2018) 359:770–5. doi: 10.1126/science.aao1710

26. Braun DA, Hou Y, Bakouny Z, Ficial M, Sant'Angelo M, Forman J, et al. Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. *Nat Med*. (2020) 26:909–18. doi: 10.1038/s41591-020-0839-y

27. Manguso RT, Pope HW, Zimmer MD, Brown FD, Yates KB, Miller BC, et al. *In vivo* CRISPR screening identifies Ptpn2 as a cancer immunotherapy target. *Nature*. (2017) 547:413–8. doi: 10.1038/nature23270

28. Zhou Y, Bian S, Zhou X, Cui Y, Wang W, Wen L, et al. Single-cell multiomics sequencing reveals prevalent genomic alterations in tumor stromal cells of human colorectal cancer. *Cancer Cell*. (2020) 38:818–828. e5. doi: 10.1016/j.ccell.2020.09.015

29. Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 inhibition. *New Engl J Med*. (2017) 377:2500–1. doi: 10.1056/NEJMc1713444

30. Rempel E, Kluck K, Beck S, Ourailidis I, Kazdal D, Neumann O, et al. Pan-cancer analysis of genomic scar patterns caused by homologous repair deficiency (HRD). *NPJ Precis Oncol*. (2022) 6:36. doi: 10.1038/s41698-022-00276-6

31. Herzog TJ, Vergote I, Gomella LG, Milenkova T, French T, Tonikian R, et al. Testing for homologous recombination repair or homologous recombination deficiency for poly (ADP-ribose) polymerase inhibitors: A current perspective. *Eur J Cancer*. (2023) 179:136–46. doi: 10.1016/j.ejca.2022.10.021

32. Melzer C, von der Ohe J, Lehnert H, H. Ungefroren and R. Hass: Cancer stem cell niche models and contribution by mesenchymal stroma/stem cells. *Mol Cancer*. (2017) 16:1–15. doi: 10.1186/s12943-017-0595-x

33. Maugeri-Saccà M, Bartucci M, De Maria R. DNA damage repair pathways in cancer stem cells. *Mol Cancer Ther*. (2012) 11:1627–36. doi: 10.1158/1535-7163.MCT-11-1040

34. Anagnostou V, Bardelli A, Chan TA, Turajlic S. The status of tumor mutational burden and immunotherapy. *Nat Cancer*. (2022) 3:652–6. doi: 10.1038/s43018-022-00382-1

35. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Mutational landscape determines sensitivity to PD-1 blockade in non–small cell lung cancer. *Science*. (2015) 348:124–8. doi: 10.1126/science.aaa1348

36. Niu X, Martinez L. Harnessing p53 to improve immunotherapy for lung cancer treatment. *Cancer Res*. (2024) 84:179–80. doi: 10.1158/0008-5472.CAN-23-3929

37. Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, Prabhakaran S, et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell*. (2018) 174:1293–1308. e36. doi: 10.1016/j.cell.2018.05.060

38. Wu Q, Liao R, Miao C, Hasnat M, Li L, Sun L, et al. Oncofetal SNRPE promotes HCC tumorigenesis by regulating the FGFR4 expression through alternative splicing. *Br J Cancer*. (2024) 131:77–89. doi: 10.1038/s41416-024-02689-5

39. Yang J, Xiong X, Zheng W, Xu H, Liao X, Wei Q, et al. The roles of tertiary lymphoid structures in genitourinary cancers: a comprehensive review of current advances and clinical applications. *Int J Surg*. (2024) 10:1097. doi: 10.1097/JS9.0000000000001939

40. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer*. (2012) 12:252–64. doi: 10.1038/nrc3239

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 16 December 2024
DOI 10.3389/fimmu.2024.1494106

OPEN ACCESS

EDITED BY
Raquel Alarcon Rodriguez,
University of Almeria, Spain

REVIEWED BY
Pengpeng Zhang,
Nanjing Medical University, China
Kaige Chen,
Wake Forest University, United States

*CORRESPONDENCE
Xinrui Gao
gaoxr1899@163.com
Ke Xu
cqghxuke@cqu.edu.cn
Yewei Xiao
ywx78@swmu.edu.cn

†These authors have contributed equally to this work

RECEIVED 10 September 2024
ACCEPTED 29 October 2024
PUBLISHED 16 December 2024

CITATION
He G, Jiang L, Zhou X, Gu Y, Tang J, Zhang Q, Hu Q, Huang G, Zhuang Z, Gao X, Xu K and Xiao Y (2024) Single-cell transcriptomics reveals heterogeneity and prognostic markers of myeloid precursor cells in acute myeloid leukemia. *Front. Immunol.* 15:1494106. doi: 10.3389/fimmu.2024.1494106

Single-cell transcriptomics reveals heterogeneity and prognostic markers of myeloid precursor cells in acute myeloid leukemia

Guangfeng He[1†], Lai Jiang[2†], Xuancheng Zhou[2†], Yuheng Gu[2†], Jingyi Tang[2], Qiang Zhang[3], Qingwen Hu[2], Gang Huang[2], Ziye Zhuang[4], Xinrui Gao[5*], Ke Xu[6*] and Yewei Xiao[7*]

[1]Department of Hematology, Affiliated Hospital of Southwest Medical University, Luzhou, China, [2]Department of Clinical Medicine, Southwest Medical University, Luzhou, China, [3]Department of Laboratory Medicine, Southwest Medical University, Luzhou, China, [4]First Clinical Medical College, Guangdong Medical University, Zhanjiang, China, [5]Department of Oncology, Affiliated Hospital of Southwest Medical University, Luzhou, China, [6]Department of Oncology, Chongqing General Hospital, Chongqing University, Chongqing, China, [7]Department of Physiology, School of Basic Medical Sciences, Southwest Medical University, Luzhou, China

Background: Acute myeloid leukemia (AML) is a hematologic tumor with poor prognosis and significant clinical heterogeneity. By integrating transcriptomic data, single-cell RNA sequencing data and independently collected RNA sequencing data this study aims to identify key genes in AML and establish a prognostic assessment model to improve the accuracy of prognostic prediction.

Materials and methods: We analyzed RNA-seq data from AML patients and combined it with single-cell RNA sequencing data to identify genes associated with AML prognosis. Key genes were screened by bioinformatics methods, and a prognostic assessment model was established based on these genes to validate their accuracy.

Results: The study identified eight key genes significantly associated with AML prognosis: SPATS2L, SPINK2, AREG, CLEC11A, HGF, IRF8, ARHGAP5, and CD34. The prognostic model constructed on the basis of these genes effectively differentiated between high-risk and low-risk patients and revealed differences in immune function and metabolic pathways of AML cells.

Conclusion: This study provides a new approach to AML prognostic assessment and reveals the role of key genes in AML. These genes may become new biomarkers and therapeutic targets that can help improve prognostic prediction and personalized treatment of AML.

KEYWORDS

acute myeloid leukemia, prognostic biomarkers, immune escape, personalized treatment, immunotherapy

1 Introduction

AML is a highly heterogeneous hematologic malignancy characterized by clonal proliferation of myeloid precursor cells leading to impaired differentiation and accumulation of immature primitive cells in the bone marrow and peripheral blood (1–3). AML accounts for approximately 80% of adult acute leukemia cases and carries a poor prognosis, especially in elderly patients (4, 5). Despite advances in therapeutic strategies, including chemotherapy, hematopoietic stem cell transplantation, and targeted therapies, overall survival in AML remains poor, with a 5-year survival rate of only 25% to 30% (6). An important reason for this poor prognosis is the high degree of heterogeneity in the biological and clinical manifestations of AML, which highlights the importance of searching for reliable prognostic biomarkers in order to predict the patient's prognosis and develop a personalized treatment plan (7, 8).

Over the past decade, many studies have been devoted to unraveling the molecular features of AML in an attempt to improve AML therapeutic approaches by identifying gene mutations, chromosomal abnormalities, and gene expression profiles that are associated with disease progression and prognosis (9). For example, AML with NPM1 and CEBPA mutations usually has a better prognosis, whereas AML with FLT3 mutations has a worse prognosis. These differences are critical for the choice of treatment strategies. High-throughput sequencing technologies, particularly RNA sequencing (RNA-seq), have revealed a variety of genes and abnormal signaling pathways that are frequently mutated in AML, providing valuable clues for understanding the pathogenesis of the disease. However, despite these advances, a comprehensive understanding of the molecular mechanisms affecting AML prognosis is still lacking, and the search for reliable prognostic biomarkers remains challenging (10, 11).

In this study, we adopted an integrated multi-omics approach to systematically identify and screen for possible prognostic biomarkers using transcriptomic and epigenetic data from public AML datasets. We performed survival analysis, gene expression analysis, and pathway enrichment analysis to mine genes that may be associated with AML prognosis. In addition, to validate our findings, we performed RNA-seq sequencing from bone marrow samples of 10 AML patients and 10 healthy donors to assess the expression levels of the screened candidate genes.

The main goal of this study was to provide a comprehensive prognostic biomarker analysis of AML, providing insight into the molecular mechanisms behind AML progression. By integrating RNA-seq data from public databases and our own patient cohort, we aim to identify gene signatures that are not only associated with clinical prognosis, but also hope to provide potential targets for the development of personalized treatment strategies. Ultimately, our findings are expected to improve risk stratification in AML and provide a basis for future therapeutic development targeting these prognostic genes.

2 Materials and methods

2.1 Sample source and collection

A total of 20 bone marrow samples were collected for this study, consisting of 10 bone marrow samples from patients with acute myeloid leukemia (AML) (AML group) and 10 bone marrow samples from healthy individuals (control group). All AML patients were diagnosed by bone marrow smear morphology and cytogenetic testing, and individuals in the healthy control group underwent a thorough physical examination to exclude any history of blood disorders and tumors. The samples were collected from 2023 to 2024 at Zhongshan Campus of Southwest Medical University Hospital and stored in liquid nitrogen for further processing. All subjects signed an informed consent form, and the study was approved by the Ethics Committee of the Affiliated Hospital of Southwest Medical University (Ethics Approval No. KY2024070).

2.2 RNA extraction and sequencing

Total RNA was extracted from frozen bone marrow samples using TRIzol reagent (Thermo Fisher Scientific, USA) for RNA extraction according to the manufacturer's instructions. The quality of extracted RNA was assessed by Agilent 2100 Bioanalyzer to ensure that the RNA Integrity Index (RIN) was greater than 7.0, and the RNA concentration was quantified using Qubit 2.0 (Thermo Fisher Scientific) to ensure that it met the requirements for sequencing. The RNA libraries were constructed using the Illumina TruSeq RNA Library Prep Kit (Illumina, USA) and library quality control was performed by Qubit and Bioanalyzer. All samples were bipartite sequenced on the Illumina NovaSeq 6000 platform with a read length of 150 bp and a target sequencing depth of 50M reads per sample.

2.3 Public data sources

The single cell sequencing data used in this study were obtained from the dataset GSE116256 in the GEO database (12). This dataset includes 16 patients at the time of AML diagnosis, 19 patients during treatment, and 5 healthy control donors. Considering our primary focus on the disease pathogenesis of AML, we selected three high-quality samples from each of the patients at the time of AML diagnosis and healthy control donors, excluding the interference of drug treatment. The RNA-seq data for AML were obtained from datasets GSE12417 and GSE71014 in the GEO database, containing 405 and 104 samples, respectively (13). The GSE12417 data was used as a training set during the construction of the prognostic model, while the GSE71014 data was used as an external validation set to evaluate the performance of the model.

2.4 Single-cell sequencing data processing and cell type identification

After reading single-cell sequencing data from three AML patients and three normal bone marrow samples, we used the Seurat package to perform initial processing of the data, including quality control, dimensionality reduction clustering, and visualization (14). To ensure that subsequent analyses were based

on high-quality sequencing data, we performed stringent quality control criteria on the cells, removing those with fewer than 500 or more than 5000 genes measured, as well as those with a proportion of mitochondrial genes greater than 15%, and avoiding the interference of empty droplets, doublets, and senescent cells. After data standardization and normalization, we performed PCA downscaling analysis. Based on the PCA downscaling results, a batch effect correction was performed using the "harmony" package. Next, we selected the top 20 principal components for cluster analysis, and the cluster resolution was set to 0.3, resulting in 10 cell clusters, which were visualized by UMAP. We performed preliminary cell type annotation for each population with the help of common cell marker genes and "FindAllMarkers" function, and finally identified myeloid precursor cells, monocytes, T cells, erythrocytes, NK cells and B cells.

2.5 Cellular communication analysis and pseudo-time analysis

In analyzing cell-cell interactions in the AML tumor microenvironment, we used the "CellChat" package for cellular communication network analysis, which simulates and analyzes cell-cell communication patterns by combining gene expression data with information on known signaling pathways, including ligand, receptor, and cofactor interactions. communication patterns (15). To further explore the developmental trajectories of different cell types in the tumor microenvironment and their dynamics during tumor progression, we used the Monocle R package to perform pseudo-temporal analyses of single-cell RNA sequencing data to reveal key transitions during cell development (16–18). We also used the Slingshot package for pseudo-temporal analysis to further explore the developmental trajectories of cells in single-cell RNA sequencing data. Slingshot is a powerful tool that can efficiently handle data with complex branching structures, helping us to gain a deeper understanding of the changes in the cellular state at the single-cell level (19). With these two analyses, we were able to paint a comprehensive picture of cellular developmental pathways and their dynamic behaviors in the AML tumor microenvironment.

2.6 Enrichment analysis

To explore the biological characteristics of myeloid precursor cells, we performed GSEA (Gene Set Enrichment Analysis). The gene sets used for enrichment analysis are differentially expressed genes identified by the "FindAllMarkers" function. We performed enrichment analysis of myeloid precursor cells using the KEGG gene set to identify pathways related to their biological functions. In addition, we performed GO and KEGG enrichment analyses on patients in the high-risk and low-risk groups. The set of genes used for enrichment analysis was derived from genes up regulated for expression in the high-risk group. To facilitate the retrieval of gene sets in the GO and KEGG databases, we used the "clusterProfiler" package and visualized the results of the analyses using this package (20). These analyses helped us to gain a deeper understanding of the biology of myeloid precursor cells and patients in different risk groups.

2.7 Construction of myeloid precursor cell marker genes and prognostic models

We screened myeloid precursor cells for marker genes using the "FindMarkers" function with a log2FC threshold of 0.25. Then, univariate Cox regression analyses were performed to initially screen out genes with prognostic value from the marker genes. Subsequently, we built multiple prognostic models using the GSE12417 and GSE71014 datasets in combination with 101 algorithm combinations and calculated the average C-index of each model across all cohorts to assess its predictive power. The analysis showed that the model combining the GBM (Gradient Booster) and Lasso (Least Absolute Shrinkage and Selection Operator) algorithms had the highest average C-index and was selected as the final model. The Lasso algorithm was used to identify the most prognostic genes while the GBM algorithm was used to build the final prognostic model, which consisted of 8 genes (21, 22). Finally, we plotted the Kaplan-Meier (K-M) survival curves for each gene as well as the K-M survival curves for the high- and low-risk groups using the "survival" and "survminer" packages to assess the model's prognostic performance.

2.8 Analysis of immune infiltration in high and low risk groups

To gain insight into the relationship between the prognostic impact of myeloid precursor cells and the immune microenvironment, we used the CIBERSORT tool to analyze samples for immune infiltration. Specifically, we performed quantitative assessment of immune cell composition for single gene grouped and risk grouped samples. CIBERSORT utilizes transcriptomic data to make inferences about the relative abundance of immune cell subpopulations, thereby revealing the infiltration status and characterization of immune cells in different risk groups. We also performed immune checkpoint analysis to assess the expression levels of immune checkpoint molecules and their differences in high and low risk groups. The expression of immune checkpoint molecules plays a key role in the suppression and activation of the immune system, and changes in them may affect the immune escape mechanisms of tumors. We analyzed the expression patterns of common immune checkpoints, including PD-1, PD-L1, and CTLA-4, and explored their correlation with high and low risk groups. In addition, we performed an immune function analysis to further reveal the functional status of the immune system in different risk groups by assessing the enrichment of immune-related pathways and functions. This analysis helped us to understand the dynamics of immune function in the tumor microenvironment and its potential impact on disease progression.

2.9 Statistical analysis

Statistical analyses were performed using R 4.2.2 64-bit version and its supporting software packages. For continuous variables, the nonparametric Wilcoxon rank sum test was used to assess the relationship between the two groups. The Sperman correlation analysis was used to test the correlation coefficients. All statistical analyses were performed at a level of significance of $P<0.05$.

3 Results

3.1 Single-cell data dimensionality reduction clustering and cell type identification

We obtained bone marrow aspiration single cell sequencing data from the GSE116256 dataset for three acute myeloid leukemia patients and three healthy individuals, with samples in the AML group coming from unmedicated patients. We processed the single-cell data using the Seurat package. To remove senescent and low-quality cells, we performed quality control on the cells (Figure 1A). After going through a series of steps of normalization, finding highly variable genes, and normalizing the expression matrix, a dimensionality reduction clustering step was performed. After PCA dimensionality reduction, we briefly observed the distribution of sample cells and the contribution of dimensions (Figures 1B, C) and selected the top 20 PCs for further dimensionality reduction clustering, and we clustered the cells into a total of 10 cell clusters (Figure 1D). We obtained relevant cell marker genes from the CellMarker website (http://xteam.xbio.top/CellMarker/index.jsp) and utilized the expression of these marker genes in the cell clusters for cell type identification (Figures 1E, F). Eventually we identified myeloid precursor cells, monocytes, erythrocytes, T cells, NK cells, and B cells (Figure 1G). Based on cell types, we observed the general spectrum of cells using the proposed temporal trajectory analysis (Figures 1H, I), which verified the accuracy of the myeloid precursor cells we identified. Meanwhile, we also utilized violin plots and heat maps to demonstrate the expression of marker genes in each cell type (Figures 1J, M). By comparing the number and distribution of cells in the normal and AML groups, we found that myeloid precursor cells were significantly increased in the AML group (Figures 1K, L).

3.2 Analysis of cellular communication in the AML tumor microenvironment

To understand the cellular communication between various cell types in the AML tumor microenvironment, we performed cellular communication inference using the CellChat package. Interaction between myeloid precursor cells and NK cells was relatively significant in the tumor microenvironment, and additionally, communication between monocytes in immune cells was also very active (Figures 2A, F, G). Because both myeloid precursor cells and monocytes belong to the myeloid lineage, their patterns of autocrine and paracrine communication were similar (Figure 2B). However, myeloid precursor cells send signals with higher intensity than monocytes and are the most active presence in the tumor microenvironment. In contrast, myeloid precursor cells received signals at a lower intensity, which may suggest their unregulated presence in the context of AML disease (Figure 2C). In addition we observed that the macrophage migration inhibitory factor (MIF) signaling pathway played an important role in the cellular communication process of myeloid precursor cells, which may be related to the large proliferation of myeloid precursor cells (Figures 2D, E) (23–25). We also resolved the communication patterns in the tumor microenvironment by NMF analysis, which also helped us to identify the specificity of myeloid precursor cells in communication (Figures 2H, I).The MIF pathway played multiple roles in the communication process of myeloid precursor cells, through which myeloid cells mainly communicated with monocytes and B-cells, and the cellular expression of MIF, CD74 and CD44 genes MIF, CD74 and CD44 genes were expressed at high levels in the cells and were the main ligand receptors mediating the communication through this pathway (Figures 2J–L).

3.3 Pseudo-time analysis and enrichment of myeloid precursor cells

To further understand the developmental trajectories and lineages in the tumor microenvironment, we performed a proposed-time analysis using the monocle2 package, and the trajectory results presented three distinct nodes and three branches (Figures 3A–C). Among them, monocytes, a cell type already present in the tumor microenvironment, underwent extensive proliferative development of myeloid precursor cells on the proposed temporal trajectory, and their peak numbers appeared in the middle and late stages of the proposed time (Figures 3D, E). We also observed a number of genes that were significantly differentially expressed at the proposed time, including genes such as AHSP and CA1 that functioned at a late stage, but also genes such as THBS1, CD14, S100A9, and FCN1 that functioned at an early stage of development (Figure 3F). By enrichment analysis of myeloid precursor cells, we observed apical enrichment for acute myeloid leukemia (Figures 3G–J).

3.4 Screening of myeloid precursor cell-related prognostic genes

Abnormal proliferation and dysregulated differentiation of myeloid precursor cells contribute to the onset and progression of AML, and we wish to screen genes with prognostic value based on myeloid precursor cells for AML patients (26, 27). We used the marker genes of myeloid precursor cells from single-cell data and modeled them by a combination of ten machine learning algorithms, and the results showed that the combination of Lasso and GBM was the most effective (Figure 4A). The Lasso and GBM

10.3389/fimmu.2024.1494106

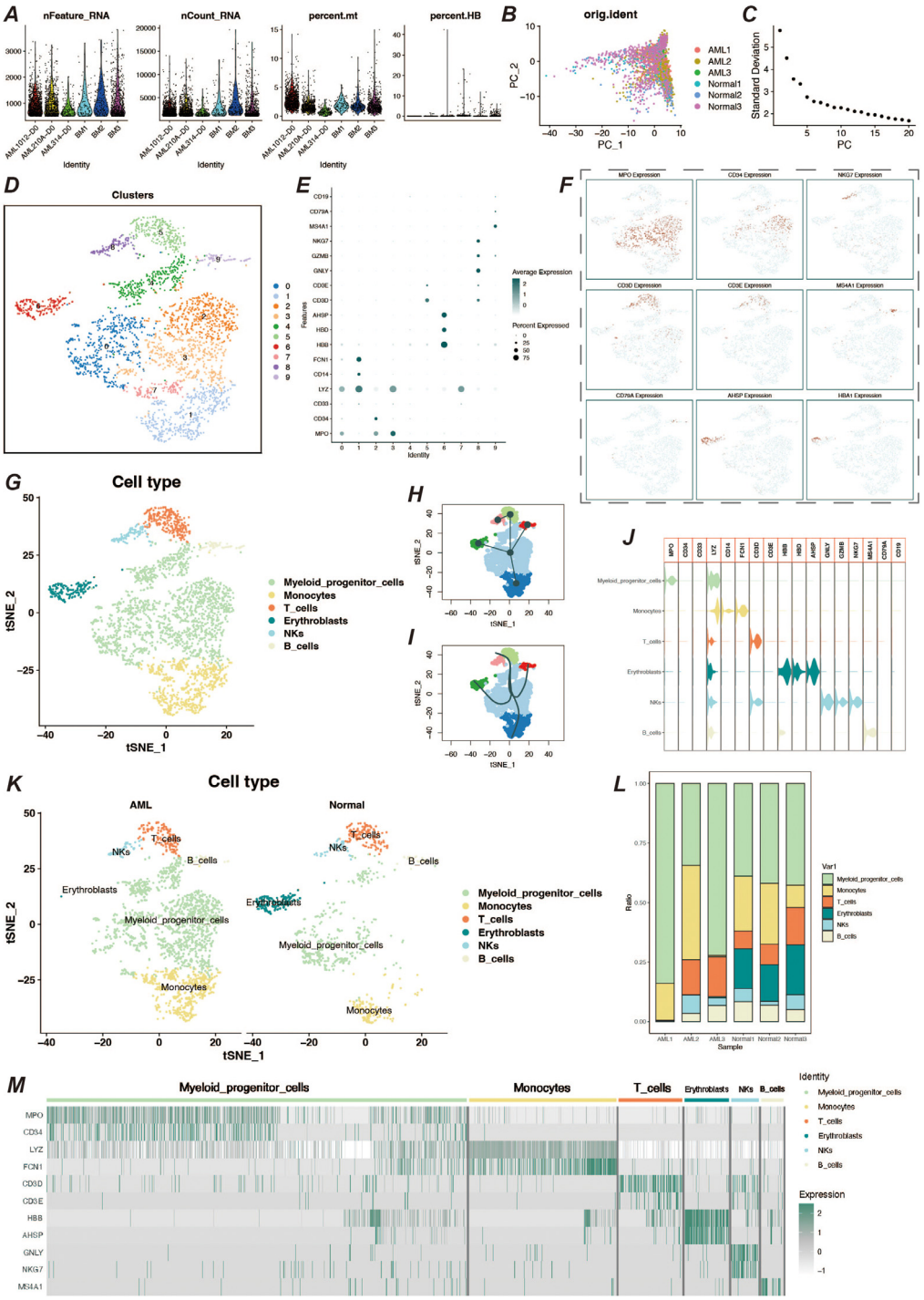


FIGURE 1
Single-cell data processing and cell type identification. **(A)** Violin plots of sample characteristics after quality control of single-cell data, showing the distribution of gene counts, total counts, mitochondrial ratio, and red blood cell ratio. **(B)** Visualization of cell distribution from different samples in the principal component analysis (PCA) space. **(C)** Elbow plot for dimension selection, used to determine the optimal number of dimensions in the dimensionality reduction process. **(D)** t-SNE dimensionality reduction clustering plot, showing the result of cells being divided into 10 clusters. **(E)** Bubble chart of marker gene expression, displaying the expression levels of marker genes across different cell populations. **(F)** Feature plot of marker gene expression, presenting the distribution of specific marker genes within cell populations. **(G)** t-SNE plot showing the results of cell type identification, displaying the distribution of cell types in two-dimensional space. **(H, I)** Pseudo-time trajectory plots, respectively showing the developmental trajectory of cells in pseudo-time analysis. **(J)** Violin plot of marker gene expression, describing the expression distribution of specific marker genes across different cell populations. **(K)** Cell type identification result display, showing the distribution and quantity of cell types by different groups. **(L)** Bar chart of cell proportions, displaying the proportion of different cell types within the total cell population. **(M)** Heatmap of cell marker gene expression, showing the expression levels of marker genes within various cell populations.

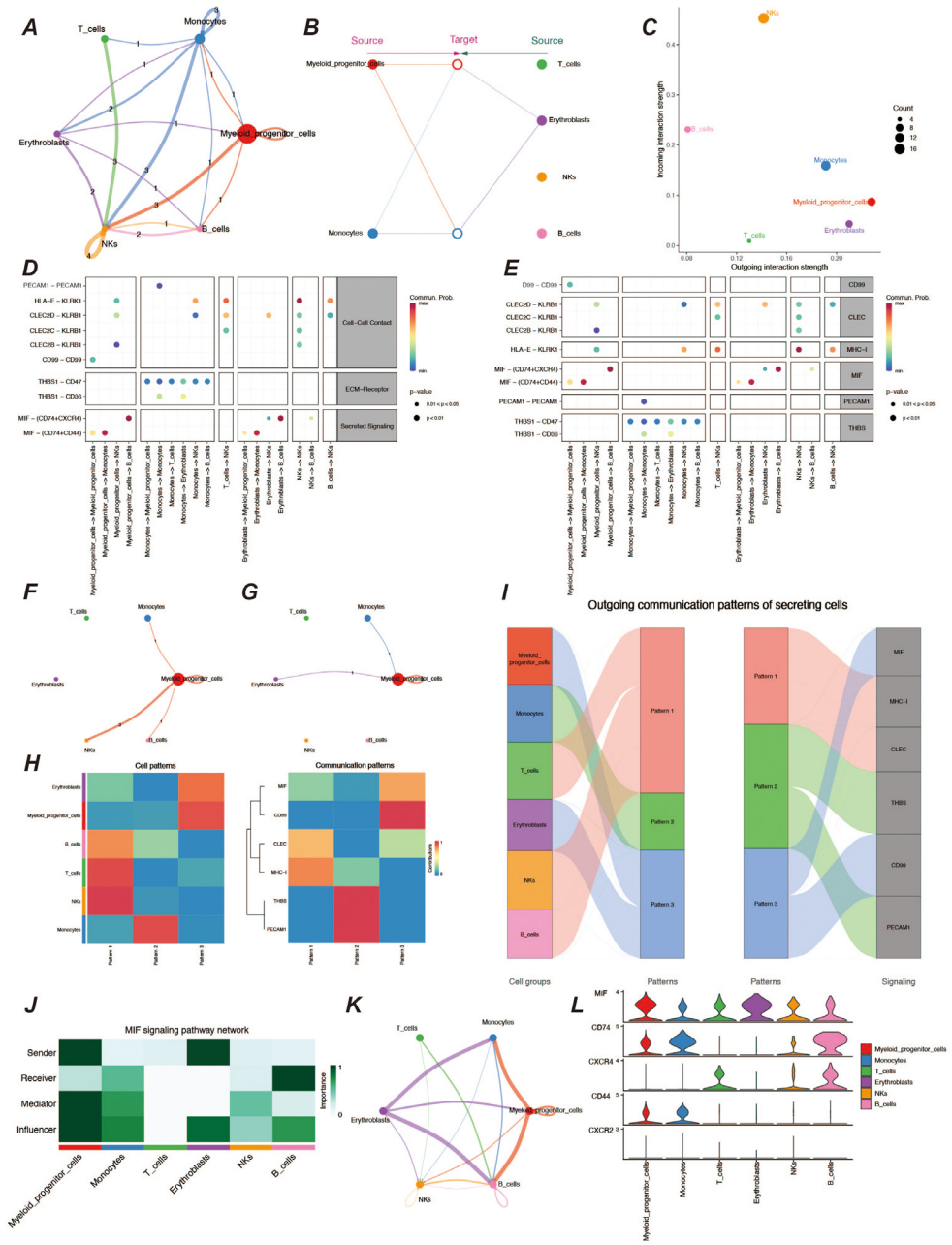


FIGURE 2
Analysis of cellular communication in the tumor microenvironment. **(A)** Cellular communication chord diagram, where the thickness of the lines represents the frequency of communication between different cells, reflecting the strength of cell interactions. **(B)** Hierarchical diagram of cellular communication, showing the communication relationships and hierarchical structure among cells. **(C)** Scatter plot of signal emission and reception intensity, displaying the distribution of signal emission and reception strength among different cells. **(D, E)** Bubble charts of ligand-receptor communication intensity, respectively showing the communication strength between ligands and receptors. The size and color of the bubbles indicate the strength and frequency of communication. **(F, G)** Chord diagrams of signal emission and reception by myeloid precursor cells, respectively showing the communication patterns of myeloid precursor cells when emitting and receiving signals. **(H)** Heatmap of cellular communication patterns, displaying the communication patterns and intensities between different cell types. **(I)** Sankey diagram of cellular communication patterns, showing the distribution of communication patterns across various cell types. **(J)** Heatmap of role preferences during cellular communication via MIF communication family ligand receptors in different cell types. **(K)** Chordal plot of cellular communication via MIF family ligand receptors. **(L)** Violin plot of MIF family ligand receptor gene expression in various cell types.

algorithms resulted in eight signature genes, SPATS2L, SPINK2, AREG, CLEC11A, HGF, IRF8, ARHGAP5, CD34, and their correlation was demonstrated by heatmap (Figure 4B). We plotted the K-M curves for each characterized gene using the corresponding overall survival data (Figure 4C).

3.5 Construction of prognostic models

We screened the characterized genes with p-value less than 0.05 as myeloid precursor cell-associated prognostic genes, which were SPATS2L, SPINK2, AREG,CLEC11A, HGF, IRF8, ARHGAP5. we

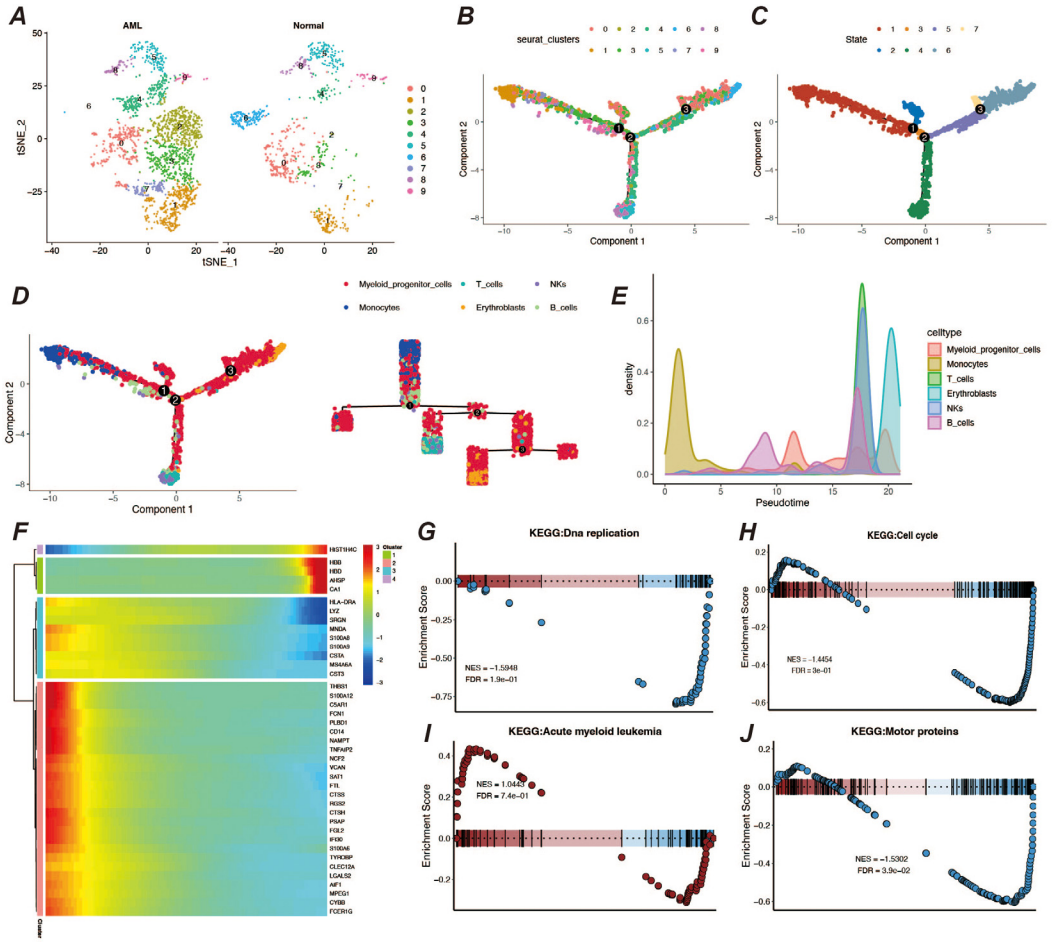


FIGURE 3
Pseudo-time analysis and enrichment analysis. **(A)** Display of cell clustering results, showing the distribution and quantity of cell types in different groups. **(B, C)** Pseudo-time developmental trajectory plots. **(D)** Distribution of different cell types on the pseudo-time trajectory, with a developmental trajectory dendrogram on the right. **(E)** Cell count peak plot. **(F)** Heatmap of differentially expressed genes in pseudo-time. **(G–J)** Results of GSEA enrichment analysis for myeloid precursor cells.

demonstrated the p-value and HR value of each gene by forest plot (Figure 5A). We categorized patients into high and low risk groups based on the expression of myeloid precursor cells (Figure 5B). The prognostic model demonstrated good discrimination and prognostic value in both the training group data and the validation group data (Figures 5C, D). We constructed prognostic column-line plots by combining the risk score with age and gender, thereby predicting the likelihood of patient survival at 1, 3, and 5 years from the composite score (Figure 5E). The calibration curves demonstrated the accuracy of the model in predicting the likelihood of survival at 1, 3, and 5 years (Figure 6A).The AUC scores of the ROC demonstrated the reliability of our risk model and the column-line diagrams (Figure 6B).

3.6 High and low risk group enrichment analysis

To explore the characteristics of patients in the high- and low-risk groups under the myeloid precursor cell prognostic model, we performed KEGG and GO enrichment analyses. the KEGG results showed that immune-related pathways were less enriched in the high-risk group, while metabolic and signaling pathways were higher (Figure 6C). For example, Natural Killer Cell Mediated Cytotoxicity, the pathway showed low enrichment in the high-risk group. Natural killer cells are an important part of the innate immune system, responsible for recognizing and destroying cancerous or virally infected cells. reduced NK cell activity may imply that immune surveillance is impaired in the high-risk group, making it easier for leukemia cells to evade clearance by the immune system, further exacerbating the disease (28–30). The T Cell Receptor Signaling Pathway, a pathway that is also higher in the high-risk group, was also found to be less active in the high-risk group (Figure 6C). T cells are at the core of the adaptive immune response, recognizing antigens and activating the immune response through the T cell receptor (TCR), and the low enrichment of the T cell signaling pathway suggests that patients in the high-risk group may be immunosuppressed and unable to effectively initiate an immune response against leukemia cells, which may be related to the immune escape mechanism of the leukemia cells (31). The results

10.3389/fimmu.2024.1494106

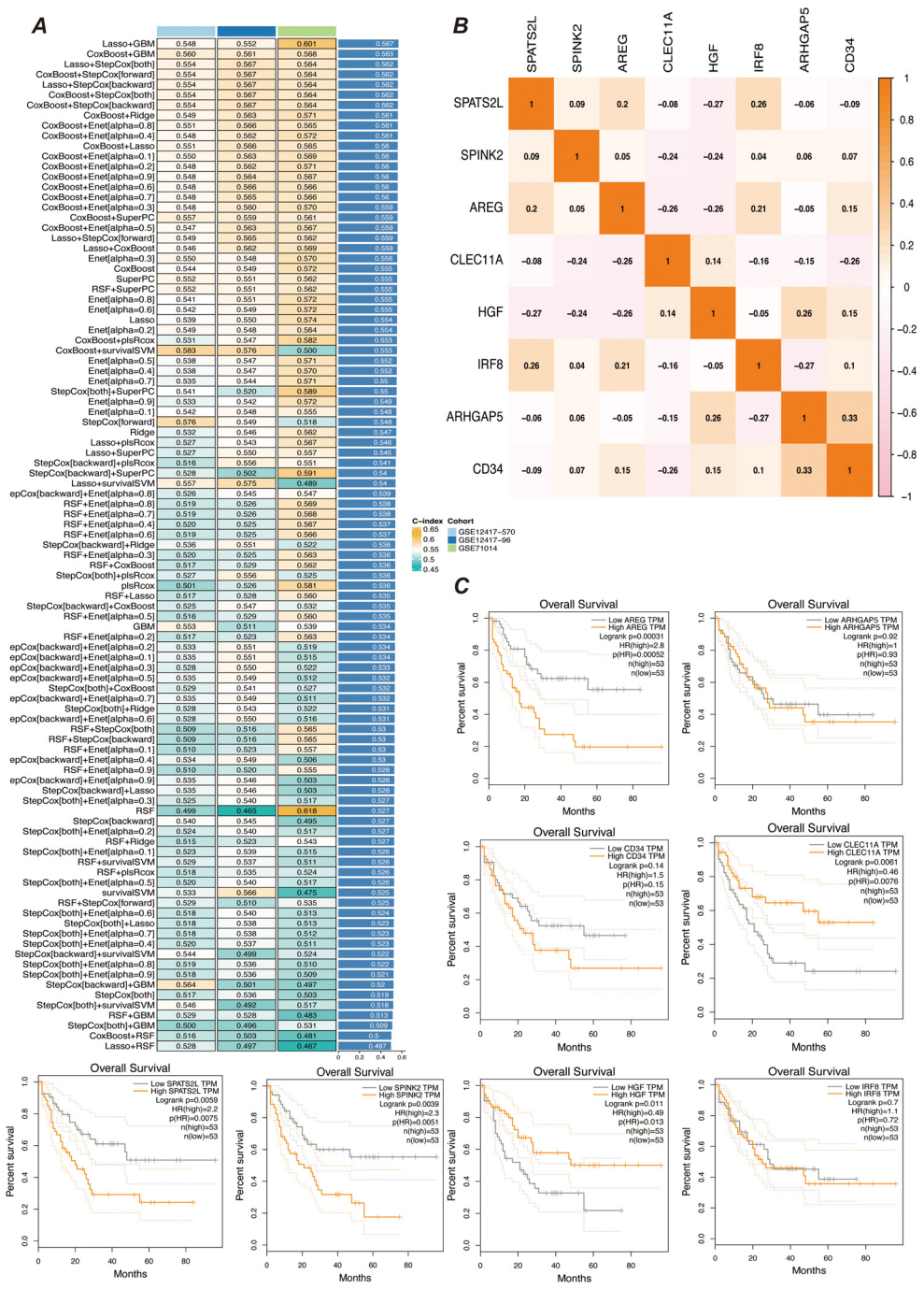


FIGURE 4
Prognostic gene selection. **(A)** Consensus model construction diagram. An illustration of the consensus model built using 101 different algorithm combinations, showing the different algorithm combinations and their consistency assessment results during the model construction process. **(B)** Heatmap of prognostic gene correlations. It displays the correlations between the selected prognostic genes. The color intensity in the heatmap indicates the strength of gene correlations, helping to identify key prognostic genes and their interrelationships. **(C)** K-M survival curves for gene expression high and low groups. Kaplan-Meier survival curves are plotted after grouping samples into high and low expression groups based on gene expression levels. The curves show differences in survival rates between gene expression level groups to assess their role in prognosis.

of the enrichment of GO function showed that patients in the high-risk group had impaired immune surveillance function, making it easier for leukemia cells to evade the immune system and further exacerbate the disease. results showed that the high-risk group was enriched for more functions related to transcriptional regulation and metabolism, while the low-risk group exhibited more functions related to development and differentiation (Figures 6D, E). From the results of these enrichment analyses, there were significant differences in immune system function, cellular differentiation, and metabolic regulation between patients in the AML high- and low-risk groups. Patients in the high-risk group exhibited suppressed immune function, activation of metabolic pathways (e.g., GPI-anchored

10.3389/fimmu.2024.1494106

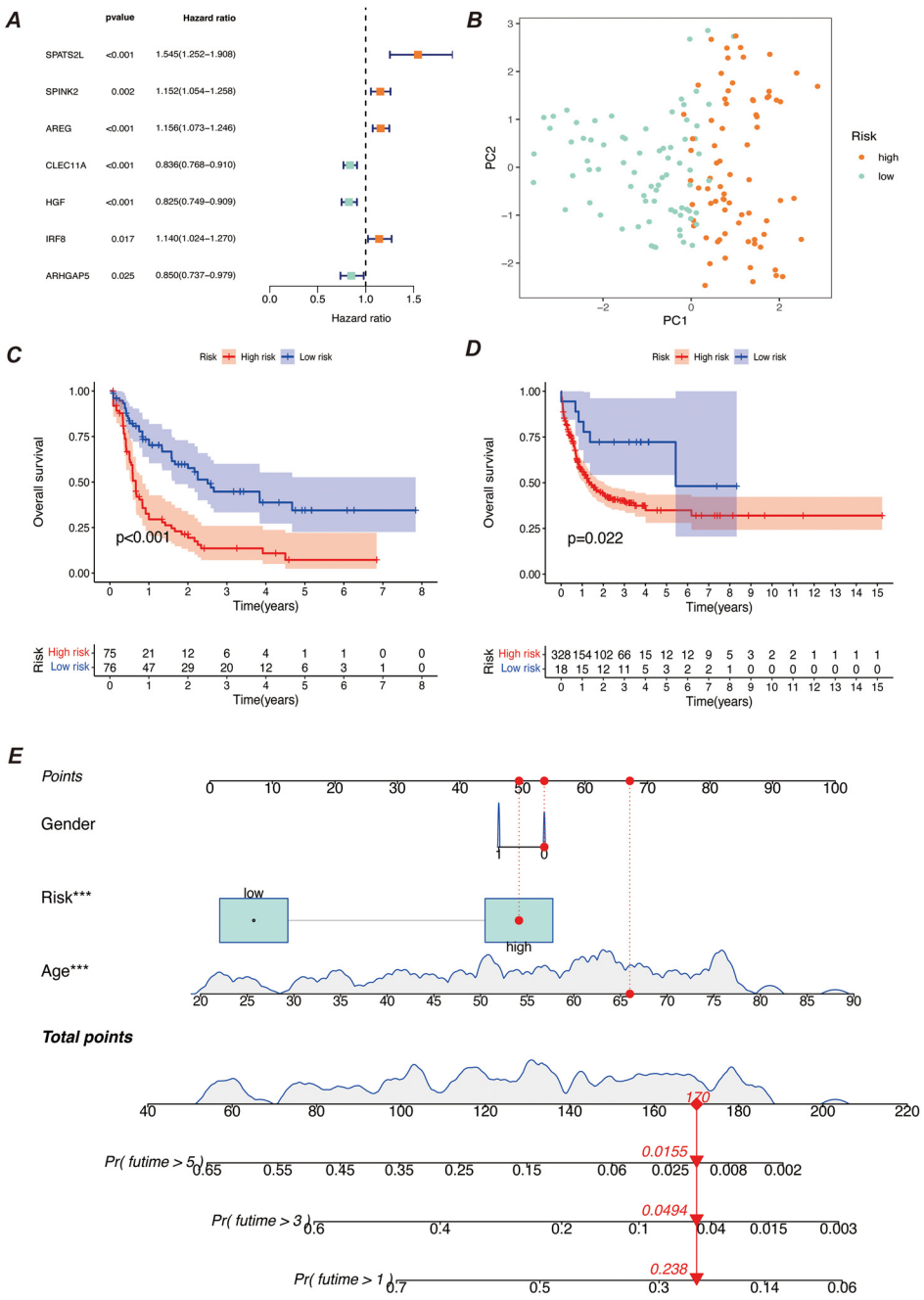


FIGURE 5
Construction of the prognostic model. **(A)** Meta-analysis diagram of univariate Cox survival analysis. It shows the results of meta-analysis of univariate Cox regression analysis for different prognostic factors to assess the correlation of each factor with survival time. The diagram includes the HR and 95% confidence interval for each factor. **(B)** PCA diagram of high-risk and low-risk groups. It compares the data distribution after dividing samples into high-risk and low-risk groups based on risk scores. The PCA diagram shows the distribution difference of the two groups on the principal components. **(C, D)** K-M survival curves. **(C)** shows the survival curves of high-risk and low-risk groups in the training set; **(D)** shows the survival curves of the corresponding groups in the validation set. Both diagrams are used to compare survival differences between different risk groups and to assess the prognostic predictive ability of the model. **(E)** Bar chart of comprehensive risk scores and clinical characteristics. It shows the association between comprehensive risk scores and clinical features, with the bar chart indicating the weight of each clinical characteristic in the comprehensive risk score and its predictive value for prognosis. *** indicates that the factor is statistically significant for disease prognosis with a p-value of less than 0.001.

synthesis, proteasome pathway), and abnormalities in gene transcriptional activation (e.g., increased activity of transcription factors, increased activity of histone methylation), which may contribute to the proliferation and survival of leukemic cells. In contrast, patients in the low-risk group were more enriched for functions related to cell differentiation, immune regulation, and developmental processes, suggesting that these patients may have more normal hematopoietic functions and immune responses.

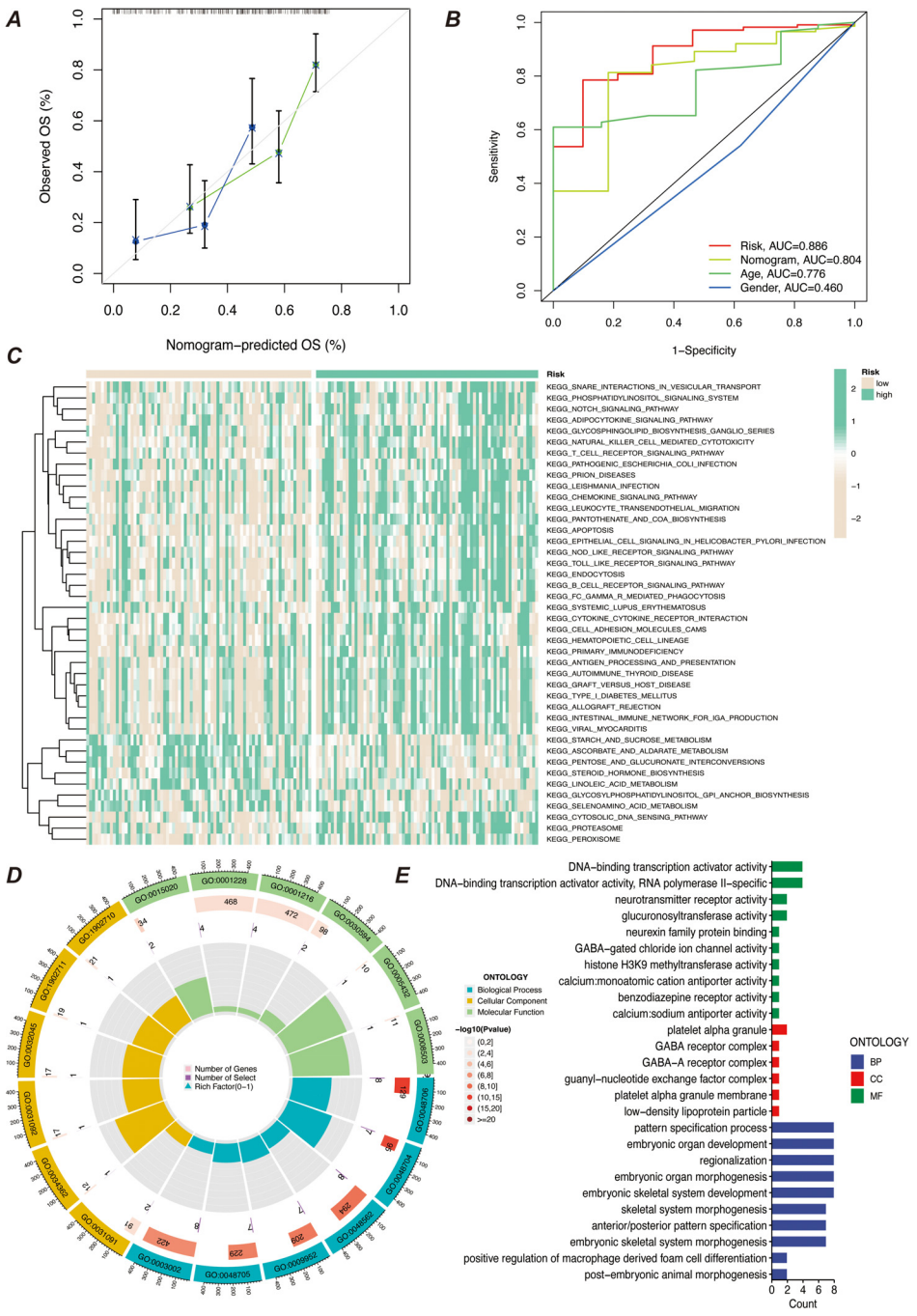


FIGURE 6
Predictive efficacy and functional enrichment analysis of the prognostic model. **(A)** Calibration curve of the prognostic model's predictive efficacy. It shows the match between the model's predicted risk scores and actual observed results, assessing the model's predictive accuracy. **(B)** ROC curve of the prognostic model's efficacy. It displays the receiver operating characteristic (ROC) curve of the model and the AUC, used to assess the model's classification ability. **(C)** KEGG pathway enrichment analysis heatmap. It shows the enrichment of prognostic genes in KEGG pathways, helping to identify key biological pathways. **(D)** GO enrichment analysis bubble chart. It shows the enrichment of prognostic genes in GO categories, with the bubble chart displaying the relationships and enrichment levels of different GO terms. **(E)** GO enrichment analysis bar chart. It shows the enrichment levels of prognostic genes in different GO categories, with the bar chart reflecting the significance of each GO category.

3.7 Immune infiltration analysis

We first performed immune infiltration analyses on subgroups of patients with the eight myeloid precursor cell signature genes, thus observing the consistency and differences in their effects on the tumor microenvironment (Figures 7A–H). Immune infiltration analyses were also performed on the high- and low-risk groups, and the high-risk group had higher levels of monocyte and macrophage M0-type infiltration. This may imply that macrophage polarization in the tumor microenvironment is

10.3389/fimmu.2024.1494106

associated with the progression of AML, and that M0-type macrophages are in an unpolarized state, where they may not yet have fully exerted their antitumor effects in the tumor microenvironment, and may even contribute to the growth of leukemic cells. The low-risk group showed a higher proportion of activated natural killer cells and activated mast cells, which may suggest that anti-tumor immune surveillance mechanisms are still more active in low-risk patients. Regulatory T cells were in higher proportion in the high-risk group. By suppressing the immune response, regulatory T cells may help leukemia cells to evade the attack of the immune system and further promote disease progression. The results of immune checkpoint analysis showed

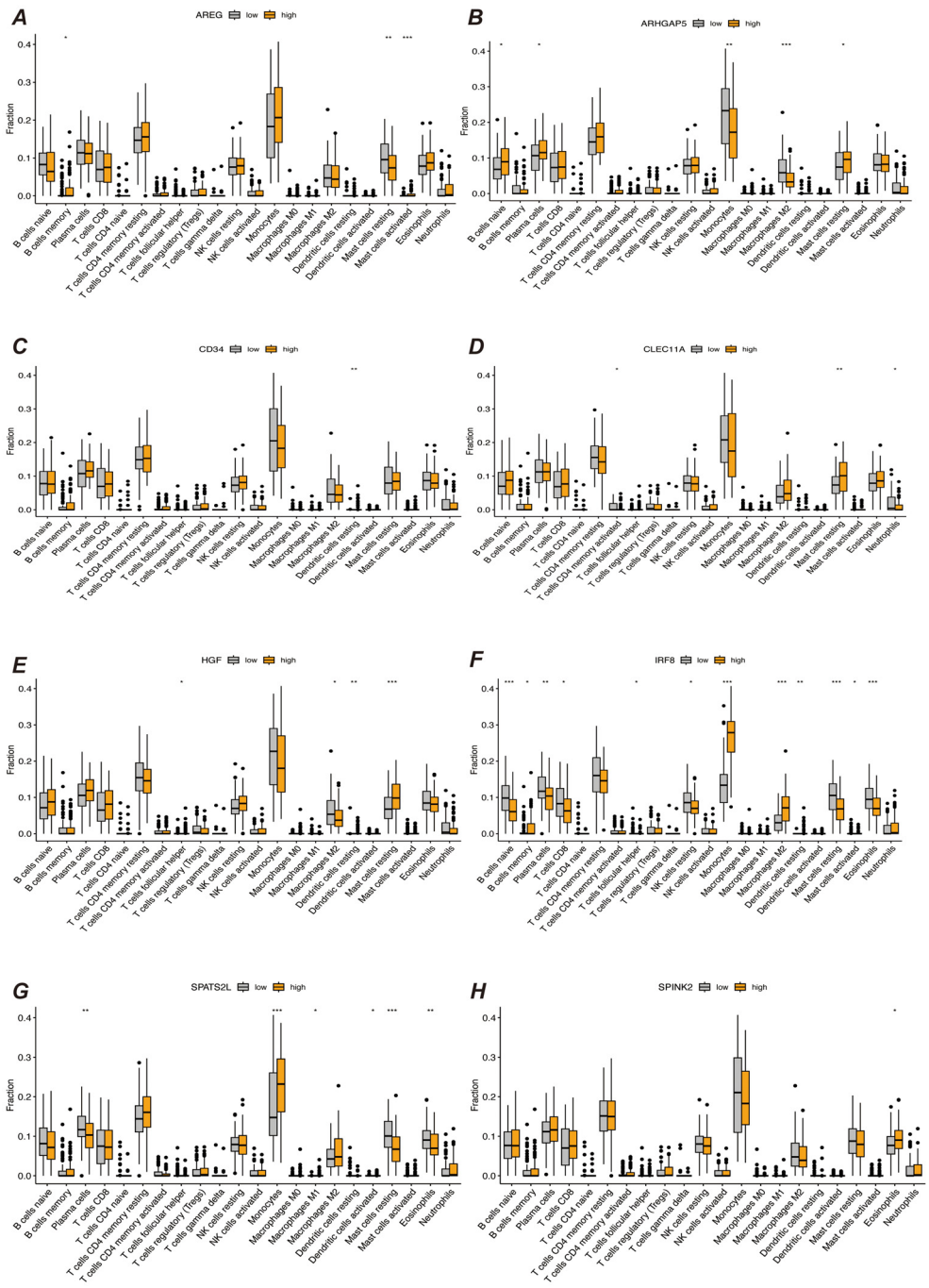


FIGURE 7
Immune infiltration analysis. **(A–H)** Results of immune infiltration analysis based on the expression levels of prognostic marker genes, grouped into different groups. Each subfigure shows the immune cell infiltration between different groups, comparing the immune infiltration levels between high and low expression groups of prognostic marker genes. "*" represents a p-value less than 0.05, "**" represent a p-value less than 0.01, and "***" represent a p-value less than 0.001.

that the expression of classical immune checkpoint molecules, such as PDCD1, CD274, and CTLA4, was significantly elevated in the high-risk group. This suggests that leukemia cells in high-risk patients may inhibit the activity of the immune system through these checkpoint molecules, allowing tumor immune escape and thus promoting disease progression. The expression of the novel immune checkpoints, such as LAG3, TIM3, and TIGIT, was likewise higher in the high-risk group, suggesting that these molecules may play a key role in immunosuppression in high-risk AML patients. LAG3 and TIM3, in particular, have been recognized as potential therapeutic targets in recent years and may be closely related to immune escape mechanisms. The lower expression levels of these immune checkpoints in the low-risk group may reflect the relatively more active immune system of these patients, which is capable of recognizing and clearing leukemia cells more effectively. The results of the immune function score analysis showed that APC co-inhibition and Check-point scored higher in the high-risk group, suggesting that the activity of antigen-presenting cells was suppressed in high-risk patients and that the immune system may have difficulty in effectively initiating an anti-leukemia response. In addition, increased immune checkpoint activity implies activation of immunosuppressive pathways, further suppressing T-cell function. Inflammation-promoting scored higher in the high-risk group, which may be related to the chronic inflammatory state in the tumor microenvironment. Chronic inflammation may promote cancer cell survival and proliferation in some cases. Type I IFN response and Type II IFN response scored higher in the low-risk group, suggesting that low-risk patients may have stronger antiviral and antitumor immune responses. Together, these analyses reveal that there are significant differences in the immune environments of high- and low-risk AML patients, that immunosuppression is an important mechanism of disease progression in the high-risk group, and that immunotherapies (e.g., checkpoint inhibitors) may have a positive therapeutic effect in these patients (Figures 8A–C).

3.8 Prognostic marker gene expression validation and enrichment analysis

We validated the expression of prognostic marker genes using the AML-BM RNA-seq Cohort with the aim of confirming the differences in the expression of these genes between acute myeloid leukemia (AML) patients and normal controls. By comparing the data from the AML group with that of the normal group, we used an independent samples t-test to statistically analyze the gene expression. The results showed that seven prognostic marker genes showed significant expression differences between the AML and normal groups (Figure 9A). We performed gene differential expression analysis using normal and disease samples from the AML-BM RNA-seq cohort and demonstrated the top thirty and bottom thirty differentially expressed genes by heatmap (Figure 9B). Then based on the differentially expressed genes, we performed GO enrichment analysis and demonstrated the enriched active pathways by bar graph (Figure 9C).

4 Discussion

In this study, we comprehensively analyzed the tumor microenvironment and prognosis-related genes in acute myeloid leukemia (AML) and constructed a prognostic model for AML patients based on single-cell RNA sequencing (scRNA-seq) and RNA-seq data. Our study revealed the abnormal proliferation and dysregulated differentiation of myeloid precursor cells in AML, elucidating the important role of this cell type in the tumor microenvironment and its potential as a potential therapeutic target for AML. Through systematic cell communication analysis, mimetic timing analysis and enrichment analysis, we not only deeply explored the function and developmental trajectory of myeloid precursor cells, but also screened out genes that are closely related to patient prognosis.

Firstly, we successfully identified myeloid precursor cells and other major cell types through dimensionality reduction clustering and cell type identification of single-cell data. Myeloid precursor cells are significantly more prevalent in AML patients, a finding that suggests their critical role in the pathological process of AML. Abnormal proliferation of these cells may be closely related to the development of AML, especially their abnormal activity in the tumor microenvironment (32, 33). Mimetic time-series analysis further validated the developmental trajectory of myeloid precursor cells, showing their extensive proliferation in AML and peaking in the late stage of the disease. Combined with the results of enrichment analysis, the gene expression profile of myeloid precursor cells revealed their specific functions in AML, especially the activation of the MIF signaling pathway, which may be related to the proliferation and autocrine regulation of myeloid precursor cells (34, 35).

Cell communication analysis also revealed for us the complex interactions between myeloid precursor cells and other cell types in the AML tumor microenvironment. We found that myeloid precursor cells interacted with NK cells and monocytes particularly in the AML microenvironment. the role of NK cells was significantly reduced in the high-risk group, possibly suggesting that the immune system of these patients was significantly suppressed. This was further verified in our analysis of immune infiltration in the high- and low-risk groups, which demonstrated suppression of immune function with activation of immune escape mechanisms in the high-risk group (36). These findings emphasize the association between immune microenvironment characteristics and disease progression in AML patients and provide a theoretical basis for targeted immunotherapy (37).

Based on the marker gene screening of myeloid precursor cells, we constructed a prognostic model for AML and selected a combined model of Lasso and GBM by comparing multiple machine learning algorithms. The final 8 characterized genes (SPATS2L, SPINK2, AREG, CLEC11A, HGF, IRF8, ARHGAP5, CD34) showed significant prognostic value in patients' survival analysis. These genes may not only help predict the prognosis of AML patients but may also serve as potential targets for future therapy (38–40). For example, high expression of SPINK2 and

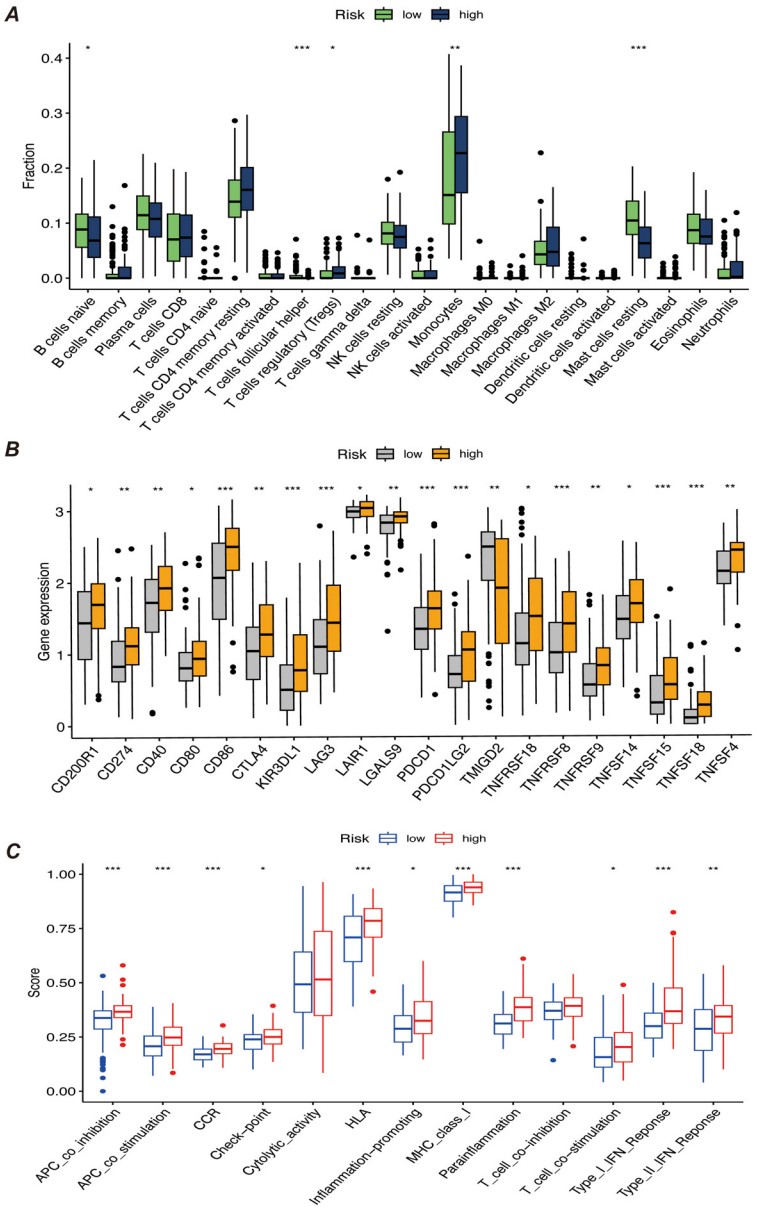


FIGURE 8
(A) Bar chart of immune infiltration analysis results for high-risk and low-risk patients, showing the infiltration levels of immune cells in different risk groups. **(B)** Bar chart comparing the expression levels of immune checkpoint genes, showing differences in the expression of major immune checkpoint genes between high-risk and low-risk groups. **(C)** Bar chart of immune function score results, comparing the immune function scores of high-risk and low-risk patients. "*" represents a p-value less than 0.05, "**" represent a p-value less than 0.01, and "***" represent a p-value less than 0.001, indicating the statistical significance of differences between groups.

AREG is closely associated with malignant progression of AML, suggesting their role in disease regulation. By K-M curves and Cox regression analysis, we further validated the expression differences of these genes in patients of high and low risk groups, and the good performance of the prognostic model was further supported by the data from the training and validation groups (41).

To further validate the expression patterns of these marker genes and their actual roles in AML patients, we collected bone marrow samples from 10 AML patients and 10 healthy individuals and performed RNA-seq sequencing analysis. The results showed that some of the prognosis-related genes screened in this paper were significantly highly expressed in bone marrow samples from AML patients and were statistically significant compared with normal individuals (42). This result is consistent with our findings obtained from publicly available databases and bioinformatics analyses, further enhancing the reliability of these genes as potential prognostic markers for AML. In addition, the gene expression data revealed that certain genes may play key regulatory roles in the development of AML, providing direction for subsequent mechanistic studies (43).

In addition, our immunoassays revealed significant features of high-risk AML patients in terms of immune escape, especially the high

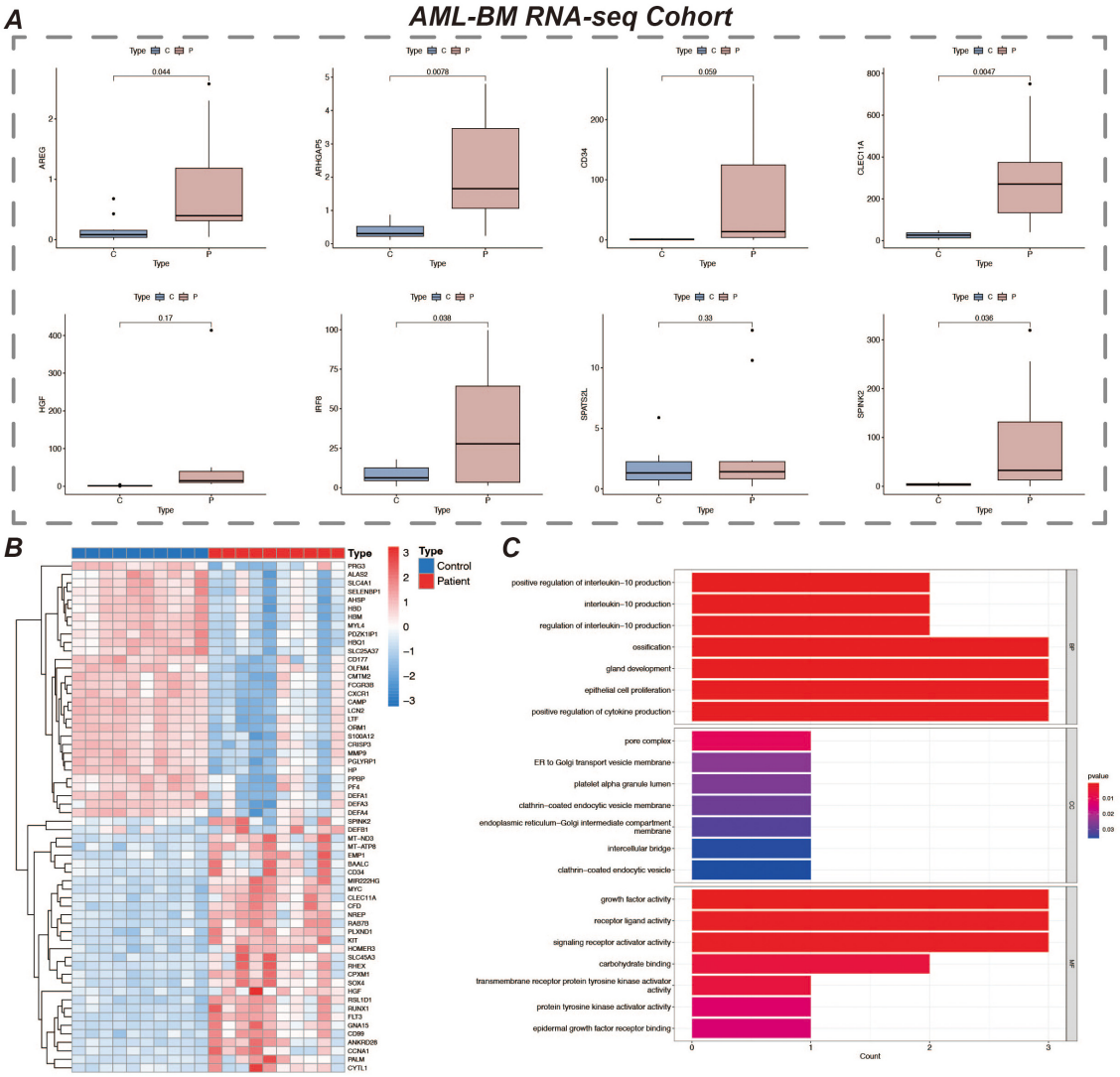


FIGURE 9
Expression of prognostic marker genes and GO enrichment analysis. **(A)** It shows the expression of 8 prognostic marker genes in the AML-BM RNA-seq Cohort. The figure displays the expression levels of these genes in samples, revealing their expression characteristics in AML bone marrow samples. **(B)** Heatmap of differentially expressed genes in the normal and disease groups, showing 30 genes each that are up- and down-regulated. Red represents up-regulated expression and blue represents down-regulated expression. **(C)** Bar chart of GO enrichment analysis results.

expression of classical immune checkpoint molecules (e.g., PDCD1, CTLA4), which further suggests that these patients may escape from the attack of the immune system through immune checkpoint inhibition mechanisms (44–46). Patients in the high-risk group had higher levels of M0-type macrophage infiltration, whereas the low-risk group showed greater immune surveillance. This suggests that the role of the immune system in the progression of AML patients is crucial, and future treatment in combination with immune checkpoint inhibitors or other immunotherapies may be needed to improve the prognosis of patients in the high-risk group (47, 48).

In summary, this study constructed a prognostic model of AML by systematic single-cell analysis and machine learning modelling and revealed the critical role of myeloid precursor cells in the pathological process of AML. Future studies should further validate the prognostic value of these genes in independent cohorts and explore myeloid precursor cell-based therapeutic interventions, thus providing new ideas for individualized treatment of AML patients.

5 Conclusion

In this study, the cellular heterogeneity and potential molecular mechanisms in the tumor microenvironment of acute myeloid leukemia (AML) were deeply resolved by integrating single-scRNA-seq and bulk RNA-seq data. The aberrant proliferation of myeloid precursor cells and their critical role in AML development were revealed by cellular communication, mimetic time-series analysis, and screening of prognosis-related genes. The AML prognostic model constructed based on survival analysis identified a variety of

prognostic-related genes and demonstrated their potential application value in survival prediction of AML patients. In addition, this study revealed features related to immunosuppression and tumor immune escape, which provided new ideas and potential targets for personalized treatment of AML.

Data availability statement

The data presented in the study are deposited in the Dryad repository, accession link: https://doi.org/10.5061/dryad.h9w0vt4t2.

Ethics statement

The studies involving humans were approved by Clinical Trial Ethics Committee, Affiliated Hospital of Southwest Medical University Affiliated Hospital of Southwest 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.

Author contributions

GFH: Conceptualization, Formal analysis, Writing – original draft. LJ: Conceptualization, Data curation, Visualization, Writing – original draft. XZ: Conceptualization, Formal analysis, Visualization, Writing – original draft. YG: Conceptualization, Formal analysis, Visualization, Writing – original draft. JT: Data curation, Formal analysis, Visualization, Writing – original draft. QZ: Formal analysis, Writing – original draft. QH: Formal analysis, Visualization, Writing – original draft. GH: Formal analysis, Visualization, Writing – original draft. ZZ: Formal analysis, Writing – original draft. XG: Conceptualization, Writing – review & editing. KX: Conceptualization, Writing – review & editing. YX: Conceptualization, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We gratefully acknowledge the Gene Expression Omnibus (GEO) for their public datasets. We also appreciate the R programming language community for the essential tools that facilitated our 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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1494106/full#supplementary-material

References

1. Newell LF, Cook RJ. Advances in acute myeloid leukemia. *Bmj*. (2021) 375:n2026. doi: 10.1136/bmj.n2026

2. Chopra M, Bohlander SK. The cell of origin and the leukemia stem cell in acute myeloid leukemia. *Genes Chromosomes Cancer*. (2019) 58:850–8. doi: 10.1002/gcc.22805

3. Long NA, Golla U, Sharma A, Claxton DF. Acute myeloid leukemia stem cells: origin, characteristics, and clinical implications. *Stem Cell Rev Rep*. (2022) 18:1211–26. doi: 10.1007/s12015-021-10308-6

4. De Kouchkovsky I, Abdul-Hay M. Acute myeloid leukemia: a comprehensive review and 2016 update. *Blood Cancer J*. (2016) 6:e441. doi: 10.1038/bcj.2016.50

5. Shimony S, Stahl M, Stone RM. Acute myeloid leukemia: 2023 update on diagnosis, risk-stratification, and management. *Am J Hematol*. (2023) 98:502–26. doi: 10.1002/ajh.26822

6. Hansrivijit P, Gale RP, Barrett J, Ciurea SO. Cellular therapy for acute myeloid Leukemia - Current status and future prospects. *Blood Rev*. (2019) 37:100578. doi: 10.1016/j.blre.2019.05.002

7. Yang X, Wang J. Precision therapy for acute myeloid leukemia. *J Hematol Oncol*. (2018) 11:3. doi: 10.1186/s13045-017-0543-7

8. Wysota M, Konopleva M, Mitchell S. Novel therapeutic targets in acute myeloid leukemia (AML). *Curr Oncol Rep*. (2024) 26:409–20. doi: 10.1007/s11912-024-01503-y

9. Jayavelu AK, Wolf S, Buettner F, Alexe G, Häupl B, Comoglio F, et al. The proteogenomic subtypes of acute myeloid leukemia. *Cancer Cell*. (2022) 40:301–317.e312. doi: 10.1016/j.ccell.2022.02.006

10. Prada-Arismendy J, Arroyave JC, Röthlisberger S. Molecular biomarkers in acute myeloid leukemia. *Blood Rev*. (2017) 31:63–76. doi: 10.1016/j.blre.2016.08.005

11. Shahrajabian MH, Sun W. Survey on multi-omics, and multi-omics data analysis, integration and application. *Curr Pharm Anal*. (2023) 19:267–81. doi: 10.2174/1573412919666230406100948

12. van Galen P, Hovestadt V, Wadsworth Ii MH, Hughes TK, Griffin GK, Battaglia S, et al. Single-cell RNA-seq reveals AML hierarchies relevant to disease progression and immunity. *Cell*. (2019) 176:1265–1281.e1224. doi: 10.1016/j.cell.2019.01.031

13. Metzeler KH, Hummel M, Bloomfield CD, Spiekermann K, Braess J, Sauerland MC, et al. An 86-probe-set gene-expression signature predicts survival in cytogenetically normal acute myeloid leukemia. *Blood*. (2008) 112:4193–201. doi: 10.1182/blood-2008-02-134411

14. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell*. (2021) 184:3573–3587.e3529. doi: 10.1016/j.cell.2021.04.048

15. 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

16. 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

17. Qiu X, Hill A, Packer J, Lin D, Ma YA, Trapnell C, et al. Single-cell mRNA quantification and differential analysis with Census. *Nat Methods*. (2017) 14:309–15. doi: 10.1038/nmeth.4150

18. Zulibiya A, Wen J, Yu H, Chen X, Xu L, Ma X, et al. Single-cell RNA sequencing reveals potential for endothelial-to-mesenchymal transition in tetralogy of fallot. *Congenital Heart Dis*. (2023) 18:611–25. doi: 10.32604/chd.2023.047689

19. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics*. (2018) 19:477. doi: 10.1186/s12864-018-4772-0

20. 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

21. Friedman JH, 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

22. Song Z, Yu J, Wang M, Shen W, Wang C, Lu T, et al. CHDTEPDB: transcriptome expression profile database and interactive analysis platform for congenital heart disease. *Congenital Heart Dis*. (2023) 18:693–701. doi: 10.32604/chd.2024.048081

23. Abdul-Aziz AM, Shafat MS, Mehta TK, Di Palma F, Lawes MJ, Rushworth SA, et al. MIF-induced stromal PKCβ/IL8 is essential in human acute myeloid leukemia. *Cancer Res*. (2017) 77:303–11. doi: 10.1158/0008-5472.Can-16-1095

24. Spertini C, Bénéchet AP, Birch F, Bellotti A, Román-Trufero M, Arber C, et al. Macrophage migration inhibitory factor blockade reprograms macrophages and disrupts prosurvival signaling in acute myeloid leukemia. *Cell Death Discov*. (2024) 10:157. doi: 10.1038/s41420-024-01924-5

25. Falantes JF, Trujillo P, Piruat JI, Calderón C, Márquez-Malaver FJ, Martín-Antonio B, et al. Overexpression of GYS1, MIF, and MYC is associated with adverse outcome and poor response to azacitidine in myelodysplastic syndromes and acute myeloid leukemia. *Clin Lymphoma Myeloma Leuk*. (2015) 15:236–44. doi: 10.1016/j.clml.2014.10.003

26. Yamashita M, Dellorusso PV, Olson OC, Passegué E. Dysregulated haematopoietic stem cell behaviour in myeloid leukaemogenesis. *Nat Rev Cancer*. (2020) 20:365–82. doi: 10.1038/s41568-020-0260-3

27. Pabon CM, Abbas HA, Konopleva M. Acute myeloid leukemia: therapeutic targeting of stem cells. *Expert Opin Ther Targets*. (2022) 26:547–56. doi: 10.1080/14728222.2022.2083957

28. Crinier A, Dumas PY, Escalière B, Piperoglou C, Gil L, Villacreces A, et al. Single-cell profiling reveals the trajectories of natural killer cell differentiation in bone marrow and a stress signature induced by acute myeloid leukemia. *Cell Mol Immunol*. (2021) 18:1290–304. doi: 10.1038/s41423-020-00574-8

29. D'Silva SZ, Singh M, Pinto AS. NK cell defects: implication in acute myeloid leukemia. *Front Immunol*. (2023) 14:1112059. doi: 10.3389/fimmu.2023.1112059

30. Romee R, Rosario M, Berrien-Elliott MM, Wagner JA, Jewell BA, Schappe T, et al. Cytokine-induced memory-like natural killer cells exhibit enhanced responses against myeloid leukemia. *Sci Transl Med*. (2016) 8:357ra123. doi: 10.1126/scitranslmed.aaf2341

31. Vadakekolathu J, Rutella S. Escape from T-cell-targeting immunotherapies in acute myeloid leukemia. *Blood*. (2024) 143:2689–700. doi: 10.1182/blood.2023019961

32. Chakraborty S, Park CY. Pathogenic mechanisms in acute myeloid leukemia. *Curr Treat Options Oncol*. (2022) 23:1522–34. doi: 10.1007/s11864-022-01021-8

33. Xu X, Xiong M, Ye H, Qi Y, Zhao Y. Analysis of single-cell sequencing results of an elderly patient with myeloid leukemia reveals high expression of multiple oncogenes in monocytes and hematopoietic stem cells. *Hematology*. (2023) 28:2240129. doi: 10.1080/16078454.2023.2240129

34. Simpson KD, Templeton DJ, Cross JV. Macrophage migration inhibitory factor promotes tumor growth and metastasis by inducing myeloid-derived suppressor cells in the tumor microenvironment. *J Immunol*. (2012) 189:5533–40. doi: 10.4049/jimmunol.1201161

35. Jiang L, Liu J, Zhang S, Jiang C, Huang J, Chen H, et al. Role of glycosylation-related gene MGAT1 in pancreatic ductal adenocarcinoma. *Front Immunol*. (2024) 15:1438935. doi: 10.3389/fimmu.2024.1438935

36. Shao R, Li Z, Xin H, Jiang S, Zhu Y, Liu J, et al. Biomarkers as targets for CAR-T/NK cell therapy in AML. *biomark Res*. (2023) 11:65. doi: 10.1186/s40364-023-00501-9

37. Jiang L, Ren X, Yang J, Chen H, Zhang S, Zhou X, et al. Mitophagy and clear cell renal cell carcinoma: insights from single-cell and spatial transcriptomics analysis. *Front Immunol*. (2024) 15:1400431. doi: 10.3389/fimmu.2024.1400431

38. Jiang C, Zhang S, Jiang L, Chen Z, Chen H, Huang J, et al. Precision unveiled: Synergistic genomic landscapes in breast cancer-Integrating single-cell analysis and decoding drug toxicity for elite prognostication and tailored therapeutics. *Environ Toxicol*. (2024) 39:3448–72. doi: 10.1002/tox.24205

39. Tu H, Hu Q, Ma Y, Huang J, Luo H, Jiang L, et al. Deciphering the tumour microenvironment of clear cell renal cell carcinoma: Prognostic insights from programmed death genes using machine learning. *J Cell Mol Med*. (2024) 28:e18524. doi: 10.1111/jcmm.18524

40. Xu Y, Cheng D, Hu L, Dong X, Lv L, Zhang C, et al. Single-cell sequencing analysis reveals the molecular mechanism of promotion of SCAP proliferation upon AZD2858 treatment. *Biocell*. (2023) 47:825–36. doi: 10.32604/biocell.2023.026122

41. Wu L, Lin Q, Chatla S, Amarachintha S, Wilson AF, Atale N, et al. LepR+ niche cell-derived AREG compromises hematopoietic stem cell maintenance under conditions of DNA repair deficiency and aging. *Blood*. (2023) 142:1529–42. doi: 10.1182/blood.2022018212

42. Zhang S, Jiang C, Jiang L, Chen H, Huang J, Gao X, et al. Construction of a diagnostic model for hepatitis B-related hepatocellular carcinoma using machine learning and artificial neural networks and revealing the correlation by immunoassay. *Tumour Virus Res*. (2023) 16:200271. doi: 10.1016/j.tvr.2023.200271

43. Zhang S, Jiang C, Jiang L, Chen H, Huang J, Zhang J, et al. Uncovering the immune microenvironment and molecular subtypes of hepatitis B-related liver cirrhosis and developing stable a diagnostic differential model by machine learning and artificial neural networks. *Front Mol Biosci*. (2023) 10:1275897. doi: 10.3389/fmolb.2023.1275897

44. Long H, Jia Q, Wang L, Fang W, Wang Z, Jiang T, et al. Tumor-induced erythroid precursor-differentiated myeloid cells mediate immunosuppression and curtail anti-PD-1/PD-L1 treatment efficacy. *Cancer Cell*. (2022) 40:674–693.e677. doi: 10.1016/j.ccell.2022.04.018

45. Minnie SA, Hill GR. Immunotherapy of multiple myeloma. *J Clin Invest*. (2020) 130:1565–75. doi: 10.1172/jci129205

46. Sauerer T, Velázquez GF, Schmid C. Relapse of acute myeloid leukemia after allogeneic stem cell transplantation: immune escape mechanisms and current implications for therapy. *Mol Cancer*. (2023) 22:180. doi: 10.1186/s12943-023-01889-6

47. Toffalori C, Zito L, Gambacorta V, Riba M, Oliveira G, Bucci G, et al. Immune signature drives leukemia escape and relapse after hematopoietic cell transplantation. *Nat Med*. (2019) 25:603–11. doi: 10.1038/s41591-019-0400-z

48. Chen H, Zuo H, Huang J, Liu J, Jiang L, Jiang C, et al. Unravelling infiltrating T-cell heterogeneity in kidney renal clear cell carcinoma: Integrative single-cell and spatial transcriptomic profiling. *J Cell Mol Med*. (2024) 28:e18403. doi: 10.1111/jcmm.18403

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 24 January 2025
DOI 10.3389/fimmu.2024.1463747

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Hongyun Shi,
Affiliated Hospital of Hebei University, China
Longbin Wang,
Deutsches Zentrum für Neurodegenerative Erkrankungen, Germany

*CORRESPONDENCE
Wen-Ming Hong
hongwenming@ahmu.edu.cn
Fang Zhang
2009500029@ahmu.edu.cn
Dong-Hui Chen
20202508@qq.com

†These authors have contributed equally to this work

RECEIVED 12 July 2024
ACCEPTED 31 December 2024
PUBLISHED 24 January 2025

CITATION
Yu T-H, Ding Y-Y, Zhao S-G, Zhao J-H, Gu Y, Chen D-H, Zhang F and Hong W-M (2025) Single-cell sequencing uncovers the mechanistic role of DAPK1 in glioma and its diagnostic and prognostic implications. *Front. Immunol.* 15:1463747. doi: 10.3389/fimmu.2024.1463747

Single-cell sequencing uncovers the mechanistic role of DAPK1 in glioma and its diagnostic and prognostic implications

Tian-Hang Yu[1†], Yan-Yu Ding[2,3†], Si-Guo Zhao[1†], Jie-Hui Zhao[4], Yu Gu[4], Dong-Hui Chen[5*], Fang Zhang[4*] and Wen-Ming Hong[1,6*]

[1]Department of Neurosurgery, First Affiliated Hospital of Anhui Medical University, Hefei, China, [2]Department of Immunology, School of Basic Medicine, Anhui Medical University, Hefei, China, [3]Institute of Health and Medicine, Hefei Comprehensive National Center, Hefei, China, [4]School of Nursing, Anhui Medical University, Hefei, China, [5]Department of Neurosurgery, Lu'an People's Hospital, Luan, China, [6]Open Project of Key Laboratory of Dermatology, Ministry of Education, Anhui Medical University, Hefei, China

Background: We conducted an investigation into the characteristics of single-cell differentiation data in gliomas, with a focus on developing DAPK1-based prognostic markers to predict patient outcomes. Dysregulated expression of DAPK1 has been associated with the invasive behavior of various malignancies, including gliomas. However, the precise role and underlying mechanisms of DAPK1 in gliomas remain inadequately understood.

Methods: We performed analyses on RNA-seq and microarray datasets from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO), in addition to single-cell RNA sequencing (scRNA-seq) data from glioma patients available in GEO. Utilizing the Seurat R package, we identified gene clusters associated with survival from the scRNA-seq data. Prognostic models were developed using LASSO and stepwise regression algorithms. Furthermore, we assessed the predictive potential of these genes within the immune microenvironment and their relevance in immunotherapy contexts.

Results: Our scRNA-seq data analysis revealed 32 distinct cell clusters corresponding to 10 cell types. Through dimensionality reduction and clustering, we identified three glial cell subpopulations based on their differentiation trajectories. DAPK1, serving as a marker gene for the terminal subpopulation, exhibited an association with poor prognosis.

Conclusions: DAPK1-based prognostic models show promise for accurately predicting outcomes in glioblastoma and glioma. An in-depth examination of DAPK1's specific mechanisms in glioblastoma could elucidate its role in immunotherapy response. Targeting the DAPK1 gene may offer therapeutic benefits for glioma patients.

KEYWORDS

glioma, DAPK1, diagnosis, prognosis, immune infiltration, tumor immune microenvironment

1 Introduction

Gliomas are among the most prevalent and lethal intracranial tumors in adults, presenting a significant challenge to global public health due to their high incidence and poor prognosis (1, 2). The diagnosis and treatment of gliomas are particularly challenging because of the complex etiology of neurological diseases (3, 4), unclear pathophysiological processes, and limited advances in pharmacological treatments and therapeutic interventions (5, 6). Treatment options for glioma patients remain limited; while emerging immunotherapies show promise in other tumors, their efficacy in gliomas is hindered by the tumor's immune-suppressive environment (7, 8). Traditional approaches such as surgery, radiotherapy, and chemotherapy also provide limited success (9–11). Early diagnosis, effective intervention, and accurate prognosis are critical for improving outcomes in glioma patients, as those diagnosed early and treated with targeted therapies tend to have better prognoses (12, 13). Thus, there is an urgent need for research to discover new and more effective strategies to combat glioma (14). Despite significant progress in understanding tumor cell heterogeneity and the tumor microenvironment (TME) in recent years (15, 16), translating these findings into enhanced diagnostics and immunotherapy advancements for glioma patients has been challenging due to the complex interactions among various cells within the TME (17, 18). Therefore, elucidating the interactions among diverse cells within the TME related to glioma initiation and progression, and identifying potential pathological mechanisms and therapeutic targets, is essential for achieving significant clinical advancements (19, 20).

Macrophages play a crucial role in gliomas (21, 22). These immune cells, primarily found in tissues including the brain, are responsible for clearing foreign substances, dead cells, and debris while participating in immune responses (23, 24). In gliomas, macrophages can exhibit dual roles. They can identify and destroy tumor cells, thereby inhibiting tumor growth and spread through the release of cytotoxic factors such as oxidants and nitric oxide (25, 26), and by promoting the activation of other immune cells (27). However, macrophages can also be exploited by tumor cells to facilitate tumor growth and metastasis (28). Tumor cells can release signaling molecules that induce macrophages to polarize into tumor-associated macrophages (TAMs) (29, 30), which promote tumor cell proliferation, angiogenesis, and metastasis through the release of growth factors, cytokines, and proteases (31, 32). Thus, understanding the dual role of macrophages in gliomas and exploring strategies to modulate macrophage activity is crucial for enhancing therapeutic efficacy.

The protein encoded by the DAPK1 (Death-associated protein kinase 1) gene is a serine/threonine kinase involved in various biological processes, including apoptosis, cell cycle regulation, cell motility, and cell adhesion (33, 34). DAPK1 is considered a significant tumor suppressor gene, capable of inhibiting tumor development by regulating apoptosis and the cell cycle (35). Recently, DAPK1 has gained interest in research on neurological disorders (36, 37). Elevated DAPK1 expression has been observed in Alzheimer's disease (AD), where it may contribute to pathogenesis by modulating apoptosis, inflammatory responses, and pathways associated with neuronal synaptic damage (38, 39). In Parkinson's disease (PD) (40), DAPK1 hyperactivation is linked to neuronal apoptosis and synaptic injury, exacerbating disease progression (41). In stroke, DAPK1 may play a role in neuroprotection and regeneration by regulating apoptosis, inflammation, and synaptic reconstruction (42). Conversely, reduced DAPK1 expression is reported in various cancers, such as lung, breast, colon, and gastric cancers, highlighting its role in tumorigenesis and progression (43). DAPK1 can regulate multiple apoptosis pathways, including phosphorylation-mediated regulation of Bcl-2 family members, modulation of p53 activity, and activation of the mitochondrial pathway. Dysregulation of these pathways can lead to tumor cell evasion of apoptosis and tumor progression (44). Additionally, DAPK1 influences cell cycle progression by phosphorylating proteins such as p53, p21, and Cdc25, thereby inhibiting tumor cell proliferation (45). DAPK1 also affects signaling pathways involved in tumor initiation and progression, such as Wnt/β-catenin, NF-κB, and MAPK pathways, impacting tumor proliferation, invasion, metastasis, and drug resistance. Although limited, some research suggests a relationship between DAPK1 and macrophages. DAPK1 may modulate macrophage function by influencing pathways related to apoptosis, inflammatory responses, and autophagy, potentially affecting macrophage roles in inflammation, infection, and the TME. DAPK1 might also regulate macrophage polarization into TAMs, which are associated with tumor growth, invasion, and metastasis. Thus, investigating DAPK1's specific mechanisms in gliomas and its interaction with TAMs is crucial for advancing our understanding and therapeutic approaches.

2 Methods

2.1 Data source

Single-nucleus RNA sequencing (snRNA-seq) data were retrieved from the Gene Expression Omnibus (GEO) database (accession number GSE141383), with a particular focus on dataset GSE138794. The samples analyzed included GSM4119521 to GSM4119530. Furthermore, bulk RNA sequencing data were obtained from The Cancer Genome Atlas (TCGA) via the TCGA portal.

2.2 Data filtering and the standard process

The initial processing of the snRNA-seq data was carried out using the Seurat package (version 4.3.0) in R (version 4.2.0). The DoubletFinder package (version 2.0.3) was used to detect and remove potential doublets according to its standard procedure. Cells were filtered based on predefined quality control criteria: 300 < nFeature < 7,500, 500 < nCount < 100,000, mitochondrial

gene expression < 20% of total cellular genes, and erythroid gene expression < 5% of total cellular genes (46, 47). The data were normalized, and the top 2,000 highly variable genes were identified for subsequent analysis. To address batch effects, the Harmony R package (version 0.1.1) was employed. Dimensionality reduction was carried out using 30 principal components (PCs), followed by clustering and visualization of cellular heterogeneity through the UMAP method (48). Cell clusters were annotated using established cell markers from the literature and the CellMarker database (CellMarker database) (49–52), allowing for the determination of cell type proportions within the dataset (53, 54).

2.3 Differentiation and enrichment analysis

Differential gene expression analysis within each cell cluster was performed using the "FindAllMarkers" function in the Seurat package, with the Wilcoxon rank sum test applied to identify differentially expressed genes (DEGs). Genes with a log fold change (logFC) > 0.25 and expression in more than 25% of the cells within the cluster were considered significant. Functional enrichment analysis of the DEGs was conducted using the "clusterProfiler" R package (version 0.1.1), which enabled Gene Ontology (GO) biological process (BP) and pathway enrichment analysis. Enriched terms were visualized to interpret the biological significance of the identified DEGs.

2.4 Subpopulation analysis of macrophage

To investigate macrophage heterogeneity, we isolated macrophage cells and identified the top 2000 variable genes. Data normalization and batch effect correction were carried out using the Harmony method prior to principal component analysis (PCA). The top 30 PCs were used for clustering and downsampling, and macrophage subpopulations were visualized using the UMAP method.

2.5 InferCNV identifies malignant cells

Copy number variation (CNV) analysis was conducted using the inferCNV algorithm to identify malignant cells among non-tumor cells. Vascular endothelial cells were used as a reference, and subpopulations exhibiting significant CNV variability were classified as glioma cells.

2.6 Difference analysis and enrichment analysis of cell subpopulations

Differential expression analysis within each cell subpopulation was performed using the "FindAllMarkers" function in conjunction with the Wilcoxon rank sum test. Enrichment analysis of Gene Ontology Biological Processes (GO-BP) was conducted using the clusterProfiler package.

2.7 Trajectory analysis

Cell differentiation trajectories in oligodendrocyte subpopulations were analyzed using the cytoTRACE, Monocle R (version 2.24.0), and Slingshot (version 2.6.0) software tools. Trajectories were reconstructed using the DDRTree algorithm, and cell lineage was inferred through minimum spanning trees (MSTs) to track developmental progression along the identified paths.

2.8 Cell communication

Intercellular communication within GBM tumor subpopulations and microenvironmental cells was analyzed using the CellChat R package (version 1.6.1) and the CellChatDB.human database for ligand-receptor interactions. Signaling pathways and receptor-ligand interactions were assessed to understand coordinated signaling across different cell types.

2.9 Prognostic modeling of glioma-associated cells

Prognostic gene signatures for glioma were identified using univariate Cox analysis and Lasso regression (55–58). A multivariate Cox model integrated key genes (59, 60) to compute risk scores, which were validated using survival analysis and receiver operating characteristic (ROC) curves at 1, 3, and 5 years (61).

2.10 Assessment of tumor-infiltrating immune cells

Immune-related scores were calculated using the CIBERSORT, ESTIMATE, and xCell algorithms to assess the immune microenvironment. CIBERSORT analysis identified 22 immune cell types across various groups, and correlated these with risk scores, model genes, and overall survival (OS). Additionally, TIDE scores and AODRA2A expression were evaluated.

2.11 Differential and enrichment analysis of bulk data

Differential expression analysis in high- and low-risk groups was performed using the DESeq2 R package, with enrichment analysis conducted using GO, KEGG, and GSEA for identified DEGs with $|logFC| > 2$ and p-value < 0.05 (62, 63).

2.12 Somatic mutation analysis

TCGA somatic mutation data were used to identify highly mutated genes and compare mutational patterns with control genes. TMB-based stratification and Kaplan-Meier analysis were

employed to assess survival differences, while CNV patterns of target genes were also evaluated (64, 65).

2.13 Drug sensitivity analysis

The pRRophetic R software was used to project IC50 values for chemotherapeutic agents and evaluate sensitivity across categories.

2.14 Cell culture

The LN229 and U-251 cell lines were obtained from American Type Culture Collection (ATCC, Manassas, VA, USA). Cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Gibco, Cat. No. 11965092), supplemented with 10% fetal bovine serum (FBS, Gibco, Cat. No. 10099141) and 1% penicillin-streptomycin (Gibco, Cat. No. 15140122). Cultures were maintained in a 37°C incubator with 5% CO2. For passaging, cells were detached using 0.25% trypsin-EDTA (Gibco, Cat. No. 25200056), counted, and reseeded at a density of 1 × 10^5 cells per flask. Cultures were monitored regularly for mycoplasma contamination and subcultured when they reached 80-90% confluency.

2.15 Cell transfection

For siRNA transfection targeting DAPK1, Lipofectamine 3000 was utilized. The siRNA sequences, synthesized by Ribobio, were used for subsequent experimental procedures. Detailed sequences targeting NUSAP1 are provided in Supplementary Table 1.

2.16 RT-qPCR analysis

RNA was isolated using TRIzol reagent, and cDNA synthesis was conducted using the PrimeScript™ RT kit. Gene expression was analyzed with SYBR qPCR Master Mix on a Roche LightCycler 480 system. GAPDH served as the internal control. Gene expression levels were quantified using the 2−ΔΔCt method. Primer sequences, sourced from Tsingke Biotech (Beijing, China), are listed in Supplementary Table 1.

2.17 The experiment of cell-cunting-kit-8 assay

Cell proliferation was assessed using the Cell Counting Kit-8 (CCK-8) assay (Dojindo Molecular Technologies, Cat. No. CK04) following the manufacturer's instructions. LN229 and U251 cells were seeded in 96-well plates at a density of 5 × 10^3 cells per well in 100 µL of complete culture medium and incubated at 37°C with 5% CO_2 for 24, 48, 72, and 96 hours. At each time point, 10 µL of CCK-8 solution was added to each well, and the plates were incubated for 2 hours at 37°C. The absorbance at 450 nm was measured using a microplate reader (BioTek, Synergy H1) to evaluate cell viability. The relative cell proliferation rate was calculated as the absorbance value of the experimental group divided by the absorbance value of the control group. All experiments were performed in triplicate.

2.18 The experiment of colony formation

The colony formation assay was performed to evaluate the proliferative capacity of LN229 and U251 cells. Cells were seeded in 6-well plates at a density of 500 cells per well in complete culture medium and incubated at 37°C with 5% CO_2 for 10–14 days to allow colony formation. The culture medium was changed every 3 days. After the incubation period, the cells were fixed with 4% paraformaldehyde (Sigma-Aldrich, Cat. No. P6148) for 15 minutes, followed by staining with 0.1% crystal violet (Sigma-Aldrich, Cat. No. C3886) for 30 minutes. Colonies were counted under a light microscope (Leica, DM3000). Colonies consisting of more than 50 cells were considered positive. The colony formation efficiency was calculated as the number of colonies formed divided by the number of seeded cells, expressed as a percentage. All experiments were conducted in triplicate.

2.19 The experiment of wound healing

To assess cell proliferation and migration capacity, a wound healing assay was performed using LN229 and U251 cells. Cells were seeded in 6-well plates (Corning, Cat. No. 353046) and cultured in complete medium until 90% confluence. A uniform "scratch" was made using a 200 µL pipette tip, creating a cell-free gap. The medium was replaced with serum-free medium (Gibco, Cat. No. 31603-029) to prevent cell proliferation during the migration phase. The cells were then cultured in a 37°C incubator with 5% CO_2. Images were captured at 0 and 24 hours post-scratch using a light microscope (Leica, DM3000). Migration ability was quantified by measuring the gap distance using ImageJ software (National Institutes of Health). The results were expressed as the percentage of wound closure, calculated as:

$$\text{Wound closure percentage} = \text{Gap at 0 hours} - \text{Gap at 24 hours} \times 100$$

Experiments were performed in triplicate to ensure reproducibility.

2.20 The experiment of transwell

The migration and invasion abilities of LN229 and U251 cells were assessed using a Transwell assay. For the migration assay, cells were suspended in serum-free medium (Gibco, Cat. No. 31603-029) at a density of 1×10^5 cells/mL and added to the upper chamber of a Transwell insert (Corning, Cat. No. 3422) with a pore size of 8 µm. The lower chamber was filled with complete medium containing 10% fetal bovine serum (FBS, Gibco, Cat. No. 10099-141) to serve as a chemoattractant. After incubating for 24 hours at 37°C and 5%

CO_2, non-migrated cells were removed from the upper surface using a cotton swab. Migrated cells were fixed with 4% paraformaldehyde (Sigma-Aldrich, Cat. No. 158127) for 15 minutes and stained with 0.1% crystal violet (Sigma-Aldrich, Cat. No. C3886) for 30 minutes. The number of migrated cells was counted in five random fields per membrane using a light microscope (Leica, DM3000).

2.21 Apoptotic rate assessed through flow cytometric analysis

Apoptosis in LN229 and U251 cells was assessed using an Annexin V-FITC Apoptosis Detection Kit (BD Biosciences, Cat. No. 556547) and flow cytometry. Briefly, cells were cultured in 6-well plates (Corning, Cat. No. 353046) and treated with the appropriate experimental conditions. After treatment, cells were harvested and washed twice with phosphate-buffered saline (PBS, Gibco, Cat. No. 10010-023). The harvested cells were resuspended in binding buffer (BD Biosciences, Cat. No. 556454), and 5 μL of Annexin V-FITC reagent and 5 μL of propidium iodide (PI, BD Biosciences, Cat. No. 556463) were added to each sample. The cells were incubated at room temperature for 15 minutes in the dark. After incubation, the cells were analyzed using a flow cytometer (BD FACSCalibur, BD Biosciences). The Annexin V-positive/PI-negative population was considered to be early apoptotic, while the Annexin V-positive/PI-positive population was considered to be late apoptotic. Data were analyzed using FlowJo software (FlowJo, LLC, Version 10).

The experiment was performed in triplicate, and apoptosis rates were calculated as the percentage of apoptotic cells (early and late apoptosis) relative to the total cell population.

2.22 Statistical analysis

Statistical analysis was performed using SPSS software (version 26.0, IBM Corp., Armonk, NY) and R software (version 4.0.3, R Foundation for Statistical Computing, Vienna, Austria). Data are presented as mean ± standard deviation (SD) for normally distributed variables and median with interquartile range (IQR) for non-normally distributed variables. The normality of data was assessed using the Shapiro-Wilk test. Comparisons between two groups were made using the Student's t-test for normally distributed data or the Mann-Whitney U test for non-normally distributed data. For comparisons among multiple groups, one-way analysis of variance (ANOVA) followed by *post hoc* Tukey's test was used. The correlation between variables was assessed using Pearson's or Spearman's correlation coefficient, as appropriate. Survival analysis was performed using the Kaplan-Meier method, and differences in survival rates were compared with the log-rank test. Univariate and multivariate Cox regression analyses were used to identify independent prognostic factors. A p-value < 0.05 was considered statistically significant.

3 Results

3.1 snRNA sequencing reveals major cell types during GBM progression

Single-nucleus RNA sequencing (snRNA-seq) was performed on tumor samples from ten glioma patients to profile cellular populations. Following rigorous quality control and filtering, 22,392 cells were analyzed using dimensionality reduction clustering, resulting in the identification of 32 distinct cell clusters (Figure 1A, top left). These clusters were annotated into ten cell types: T cells (1234 cells), neutrophils (8467 cells), microglia (54,289 cells), endothelial cells (5025 cells), B/plasma cells (1308 cells), glia/neuroglia (903 cells), mural cells (1342 cells), macrophages (38,746 cells), proliferating macrophages (2442 cells), and microglial/macrophages (12,833 cells) (Figure 1A, upper right). Cells were further categorized based on their origin within the tumor core (63,833 cells) versus the tumor periphery (62,756 cells) (Figure 1A, bottom left). Analysis of cell cycle phases showed distribution ratios of S phase (41,662 cells), G1 phase (46,150 cells), and G2M phase (38,777 cells) (Figure 1A, bottom right). Marker gene analysis identified the top ten markers for each cell type, which were visualized through bubble plots (Figure 1B). Proportional distributions of cell types between the tumor core and peripheral tissues across four cases were illustrated using histograms, highlighting inter-patient variability (Figure 1C). Differential expression patterns across experimental groups were depicted with box plots (Figure 1D). Uniform Manifold Approximation and Projection (UMAP) plots were employed to visualize distributions of key parameters (nCount_RNA, nFeature_RNA, G2M score, and S score) across all cells (Figure 1E). Gene Ontology Biological Process (GO-BP) enrichment analysis highlighted specific biological processes associated with each cell type, represented by heatmaps (Figure 1F). Volcano plots illustrated differential gene expression among cell types (Figure 1G).

3.2 Displaying the intracellular heterogeneity of macrophages

Dimensionality reduction clustering identified four distinct macrophage subpopulations. Using the inferCNV algorithm, cells with high genomic copy number variation (CNV) were classified as GBM cells (Supplementary Figure 1). The macrophage subpopulations were C0 RGS16+ Macrophages (15,867 cells), C1 DAPK1+ Macrophages (13,796 cells), C2 VCAN+ Macrophages (6,258 cells), and C3 CXCL3+ Macrophages (2,825 cells) (Figure 2A, upper left). UMAP plots visualized the distributions and proportions of these subpopulations based on cell cycle staging, subpopulation identity, and patient sample sources (Figure 2A). Further UMAP plots detailed G2M scores, nCount_RNA, S scores, and CNV scores across subpopulations (Figure 2B). Proportional distributions of subpopulations between tumor core and peripheral tissues were compared (Figure 2D). While a higher proportion of

10.3389/fimmu.2024.1463747

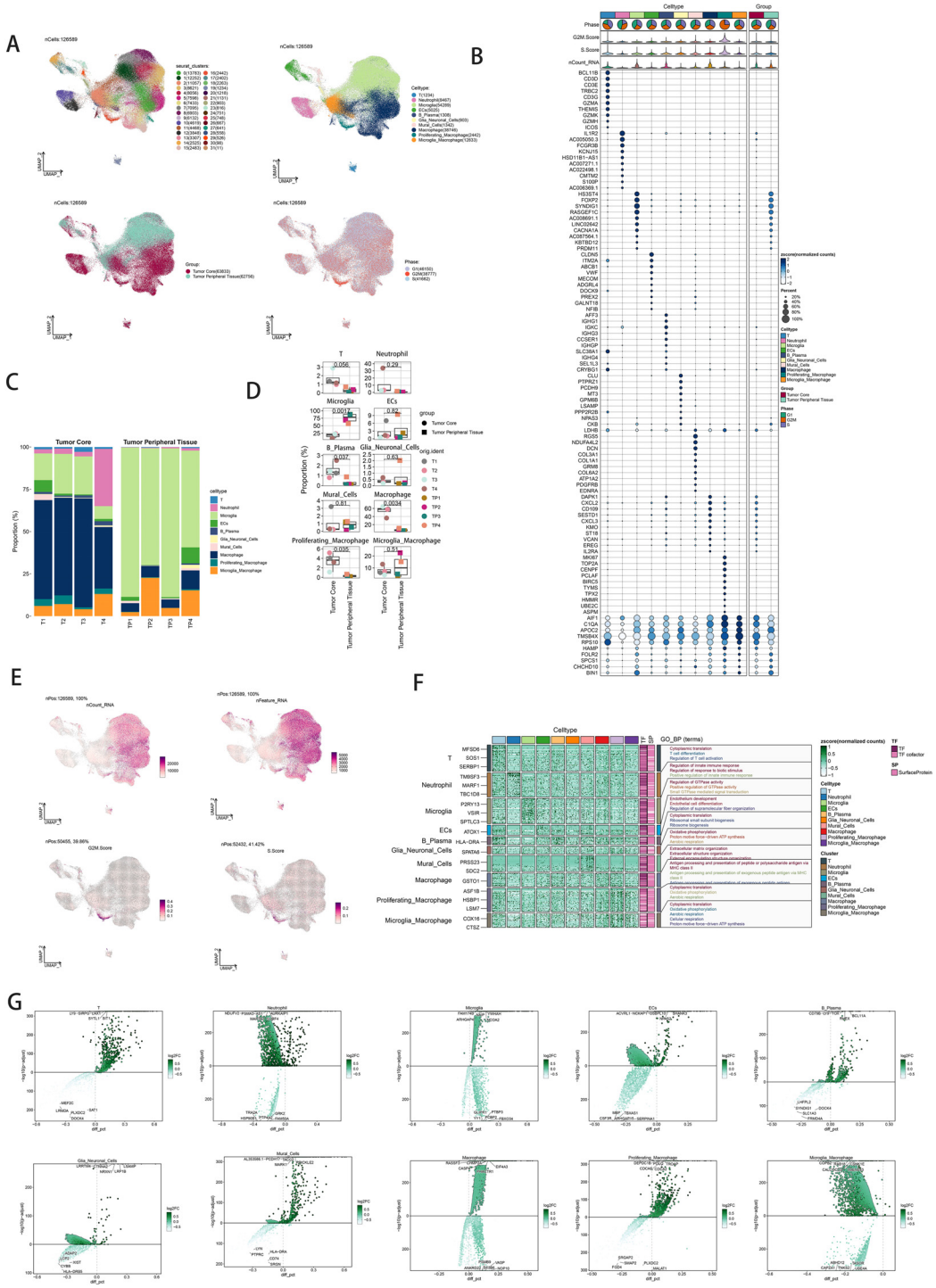


FIGURE 1
snRNA sequencing reveals major cell types in neuroblastic tumor. **(A)** UMAP plot showing the 32 clusters of cells in glioma patients and the number of cells in each cluster (top left); UMAP plot showing the 10 major cell types (top right); UMAP plot showing the distribution of the tumor core and tumor peripheral tissue for the 10 cell types (bottom left); and UMAP plot showing the distribution of different cell cycle phases (lower right). **(B)** Bubble plot showing differential expression of Top10maker genes in glioma cells across cell types. The color of the bubbles is based on the normalized data and the size indicates the percentage of genes expressed in the subpopulation. **(C)** Bar graph showing the percentage of the 10 cell types in the tumor core group versus tumor peripheral tissue group. **(D)** Box line plot depicting the percentage of the 10 cell types in the tumor core group versus tumor peripheral tissue group. **(E)** The UMAP plot showcases the distribution of the following relevant features: nCount_RNA, nFeature_RNA,S.score, and G2M.score. **(F)** GO-BP enrichment analysis demonstrating biological processes associated with the 10 cell types. **(G)** Volcano plot demonstrating differential gene expression in 10 cell types.

10.3389/fimmu.2024.1463747

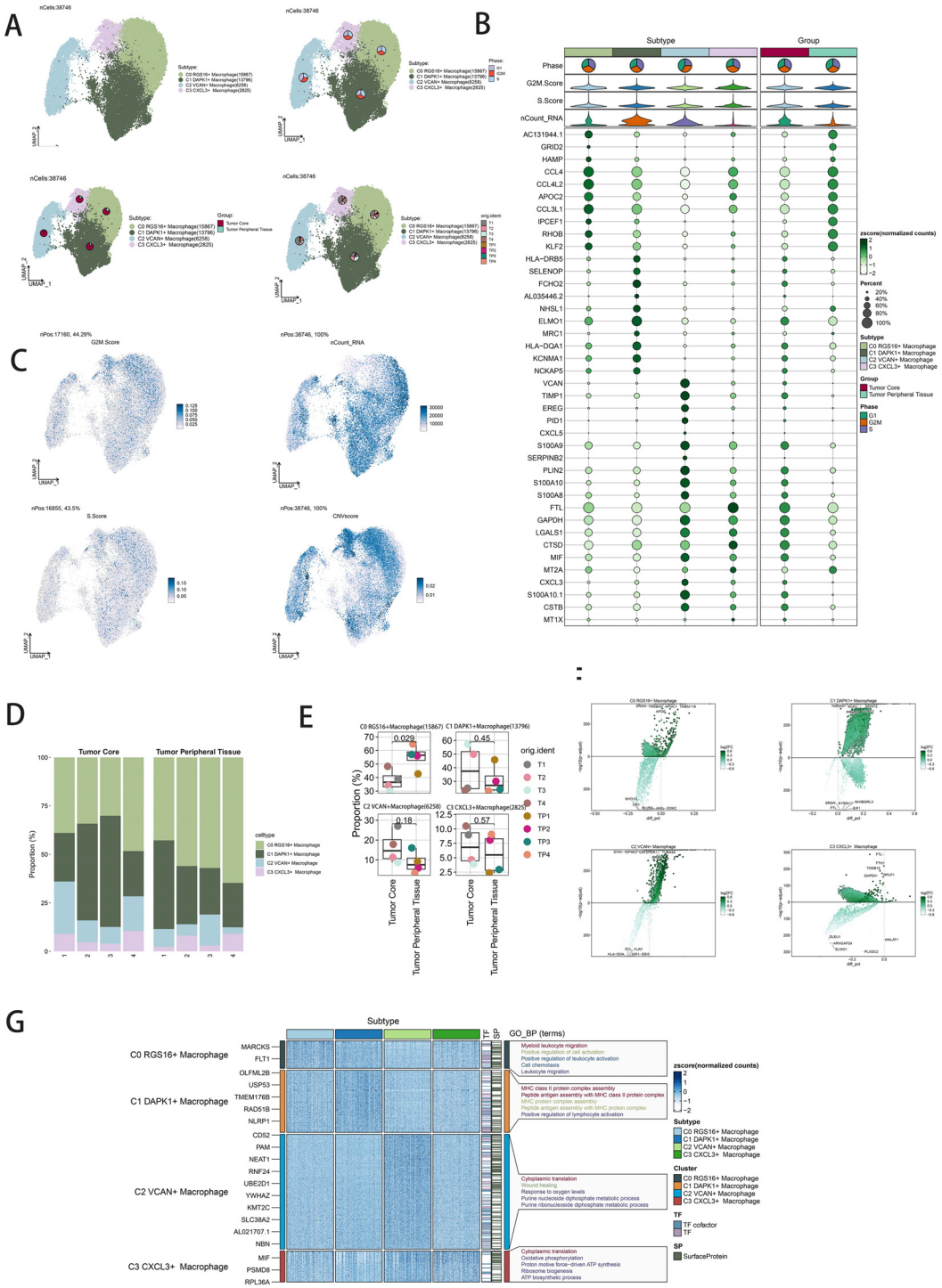


FIGURE 2
Visualization of macrophage subpopulations. **(A)** UMAP diagram demonstrating the 4 cell subpopulations of tumor cells in glioma patients and the number of cells in each subpopulation (top left); UMAP diagram demonstrating the percentage of different cell cycles in the 4 cell subpopulations (top right); UMAP diagram demonstrating the distribution of the tumor core group versus tumor peripheral tissue group in the 4 cell subpopulations (bottom left); and UMAP diagram demonstrating the patient origin of the 4 cell subpopulations (lower right). **(B)** Bubble plot showing differential expression of Top10maker genes in 4 cell subpopulations. The color of the bubbles is based on the normalized data and the size indicates the percentage of genes expressed in the subpopulation. **(C)** UMAP plot visualizing the relevant features of the 4 cell subpopulations: G2M.score, nCount_RNA,S.score,CNVscore. **(D)** Bar graph demonstrating the percentage of the 4 cell subpopulations in the tumor core group versus tumor peripheral tissue group. **(E)** Box line graph depicting the percentage of the 4 cell subpopulations in the tumor core group versus tumor peripheral tissue group. **(F)** Volcano plot demonstrating the expression of differential genes in the 4 cellular subpopulations. **(G)** GO-BP enrichment analysis demonstrating biological processes associated with the 4 cell subpopulations.

C1 DAPK1+ Macrophages was observed in the tumor core, statistical analysis did not reveal significant differences (Figure 2E). Volcano plots depicted differential gene expression patterns (Figure 2F). Bubble plots highlighted the top ten marker genes distinguishing macrophages from GBM cells (Figure 2C). GO-BP enrichment analysis of differentially expressed genes within subpopulations was visualized using heatmaps (Figure 2G).

3.3 Visualization of macrophage and GBM cell subpopulations for time-series analysis

Trajectory analysis using CytoTRACE revealed differentiation relationships among macrophage and GBM cell subpopulations (Figures 3A, B). Bar graphs compared proportions of subpopulations between tumor core and peripheral tissues, highlighting distinct distributions across cell cycle stages (Figure 3C). Distribution percentages across trajectory differentiation states were depicted in Figure 3D. Detailed cell percentage distributions within cell cycle stages and trajectory states were provided (Figure 3E). Trajectory analysis indicated continuous differentiation among subpopulations, revealing branching points and developmental stages (Figures 3F, G). Scatter plots visualized pseudotime series of selected genes specific to subpopulations (Figure 3H).

3.4 Pseudotemporal trajectories of macrophage and GBM cell subpopulations

Slingshot analysis delineated pseudotemporal trajectories of macrophage and GBM cell subpopulations, revealing continuous distribution and differentiation into distinct lineages (Figure 4A). Pseudotemporal sequences were estimated at the cellular level (Figures 4B, C). GO-BP enrichment analysis highlighted biological processes associated with pseudotemporal trajectories (Figure 4D). Scatter plots visualized differentiation trajectories across pseudotime series (Figure 4E).

3.5 Intercellular cell interactions and IL-10 signaling pathway visualization

Quantification of cellular interactions involved measuring interaction frequency between different cell types (represented by line thickness) and the strength of these interactions (indicated by line weight) (Figure 5B). CellChat analysis elucidated intercellular communication networks and signaling pathways, emphasizing IL-10 signaling pathways mediated by C1 DAPK1+ Macrophages (Figure 5A). Heatmaps depicted signal intensity of intercellular interactions (Figure 5C). Pathway visualization revealed mechanisms of IL-10 signaling (Figure 5D). Violin plots illustrated cell-cell interactions, emphasizing IL-10 signaling in C1 DAPK1+ Macrophages (Figure 5E). The IL-10 signaling pathway's interactions were visualized, showing IL-10 release and reception by different cell types (Figure 5F left). Cell interaction plots detailed the receptor-ligand profiles of the C1 DAPK1+ Macrophage subpopulation with other intercellular receptors (Figure 5F right). Scatter plots detailed cellular communication patterns in IL-10 signaling pathways (Figure 5G).

3.6 Prognostic modeling and validation

Univariate Cox analysis identified 19 genes associated with patient prognosis, with 8 genes constituting the DAPK1+ Macrophage risk score through lasso regression (Figures 6A, B). Survival analysis stratified patients based on DAPK1+ Macrophage risk scores, revealing significant prognostic implications (Figure 6C, D). TCGA-GBM dataset analysis validated negative prognostic impacts associated with higher DAPK1+ Macrophage risk scores (Figure 6E). Correlation analysis demonstrated a negative association between overall survival (OS) and DAPK1+ Macrophage risk scores, with TCF12 exhibiting significant negative correlation. Scatter plots illustrated the relationship between the model genes, risk scores, and OS (Figure 6F). ROC curves assessed predictive accuracy for survival outcomes (Figure 6G). Scatter plots visualized the relationship between model genes and risk scores (Figure 6H), while expression levels of the 8 genes varied between high and low-risk groups (Figure 6I). Multifactorial Cox regression confirmed the DAPK1+ Macrophage risk score as an independent prognostic factor (Figure 6J). Nomogram diagrams predicted survival probability integrating clinical and pathological factors (Figure 6K). Internal validation via cross-validation demonstrated the nomogram's accuracy (Figure 6L).

3.7 Immune infiltration patterns between high and low DAPK1+ macrophage risk score groups

Heatmaps visualized distinct immune infiltration patterns between high and low DAPK1+ Macrophage risk score groups (Figure 7A). The CIBERSORT algorithm estimated proportions of immune cell types, highlighting differences in immune cell composition (Figure 7B). Bar graphs correlated immune infiltrating cells with macrophage subpopulation scores (Figure 7C). Heatmaps identified correlations between model genes and immune cells (Figure 7D). Stromal scores, immune scores, and ESTIMATE scores were elevated in high-risk groups (Figure 7E). Tumor purity was lower in high DAPK1+ Macrophage risk score groups (Figure 7F).

3.8 Analysis of variance and enrichment

Volcano plots depicted differentially expressed genes between high and low DAPK1+ Macrophage risk score groups (Figures 8A, B). GO and KEGG enrichment analyses highlighted biological processes and pathways associated with these differentially expressed genes

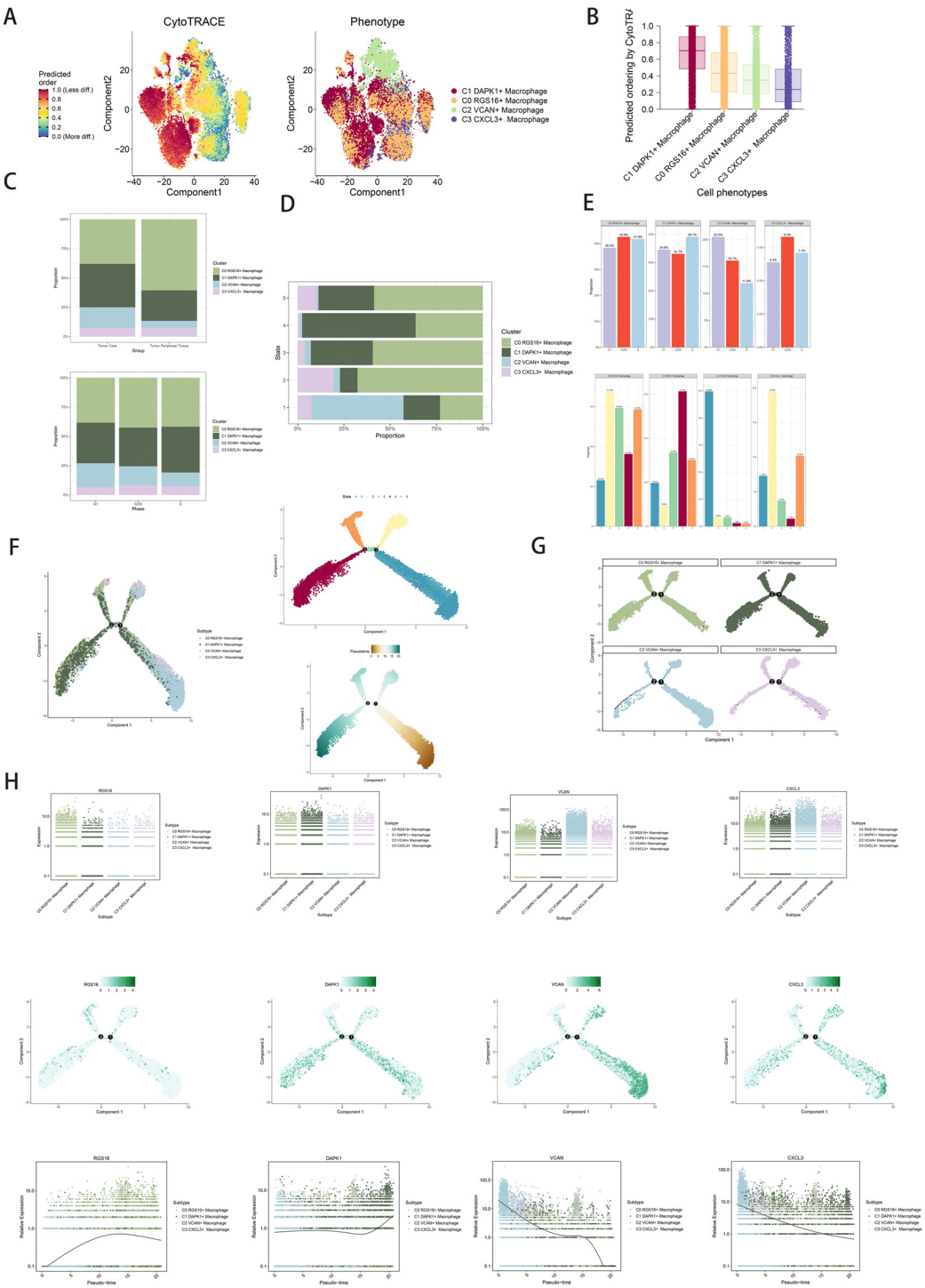


FIGURE 3
Visualization of proposed time series analysis of macrophage subpopulations by cytotrace and monocle. **(A)** The differentiation of macrophage subpopulations is analyzed using cytotrace and displayed in a 2D graph. The color can represent the level of differentiation. The figure on the right represents the cytotrace results displayed according to different macrophage subpopulations. The colors represent different cell subpopulations. **(B)** Box line plot showing the visualization results of cytotrace analysis of macrophage subpopulations. **(C)** The occupancy of relevant features in different pseudotime stages of 4 cell subpopulations was visualized: the occupancy of 4 cell subpopulations in the tumor core group versus tumor peripheral tissue group(top); the occupancy of 4 cell subpopulations in different cell cycles (bottom). **(D)** Bar charts illustrating the proportions of different pseudotime stages (state1-state6) within the four cell subgroups. **(E)** Bar graph demonstrating the expression of the 4 cell subpopulations in different phases (top) *vs*. different states(bottom), respectively. **(F)** Demonstrating the derivation process of macrophage subpopulations. Left: UAMP plot of the proposed temporal trajectory showing the 4 cell subpopulations; Middle: UMAP plot showing the pseudotime trajectory of macrophage subpopulation, starting from the lower right, are divided into two branches, one of which differentiates upward to the right and the other to the left followed by two branches, one of which differentiates upward to the left, and the other down and left;Right: UMAP plot showing the distribution of 5 states on the proposed temporal trajectory plot. **(G)** Split-plane plots of the proposed temporal trajectories of macrophage subpopulations showing the distribution of different cell subpopulations on the proposed temporal trajectories, respectively. **(H)** Scatter plot showing the changes of 4 cell subpopulations of macrophage subpopulations with the proposed chronological sequence; proposed chronological sequence UMAP plot showing the changes of the cell subpopulations corresponding to the 4 named genes with the proposed chronological sequence; and the expression of the 4 named genes of cell subpopulations (RGS16, DAPK1, VCAN, CXCL3) on the pseudotime trajectory.

10.3389/fimmu.2024.1463747

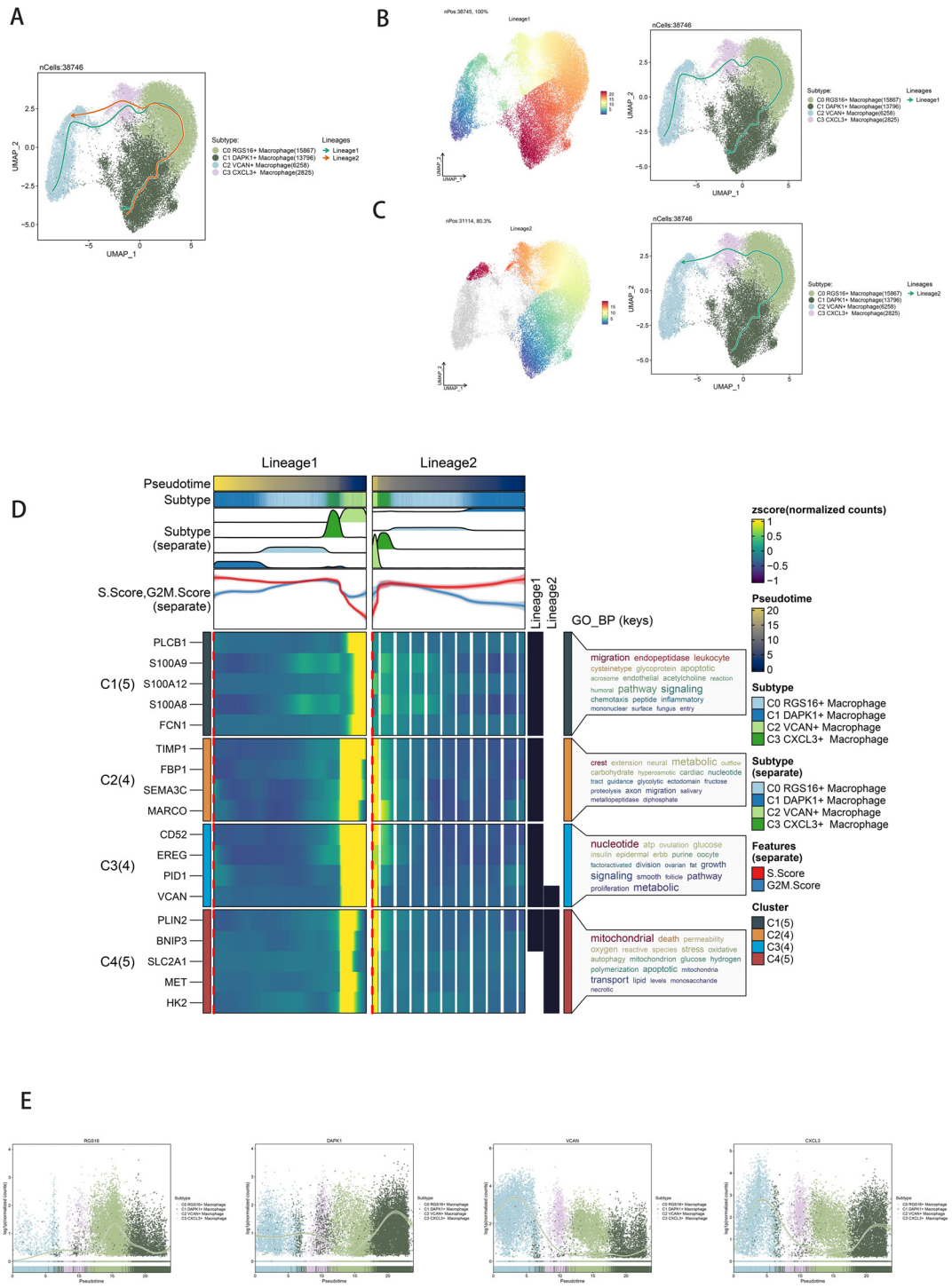


FIGURE 4
Slingshot analysis of the pseudotime trajectories of macrophage subpopulations. **(A)** UMAP plot showing the distribution of two differentiation trajectories of macrophage subpopulation fitted by the pseudotime order in all cells. **(B)** UMAP plot demonstrating the change of Lineage1 with the fitted temporal order (left); UMAP plot demonstrating the differentiation trajectory of Lineage1 on the fitted temporal order (right). **(C)** UMAP plot demonstrating the change of Lineage2 with the fitted temporal order (left); UMAP plot demonstrating the differentiation trajectory of Lineage2 on the fitted temporal order (right). **(D)** GO-BP enrichment analysis demonstrating the biological processes corresponding to the two proposed chronological trajectories of macrophage subpopulation. **(E)** Scatterplot demonstrating the trajectories of the named genes of the four cell subpopulations of macrophage subpopulation obtained after slingshot visualization.

10.3389/fimmu.2024.1463747

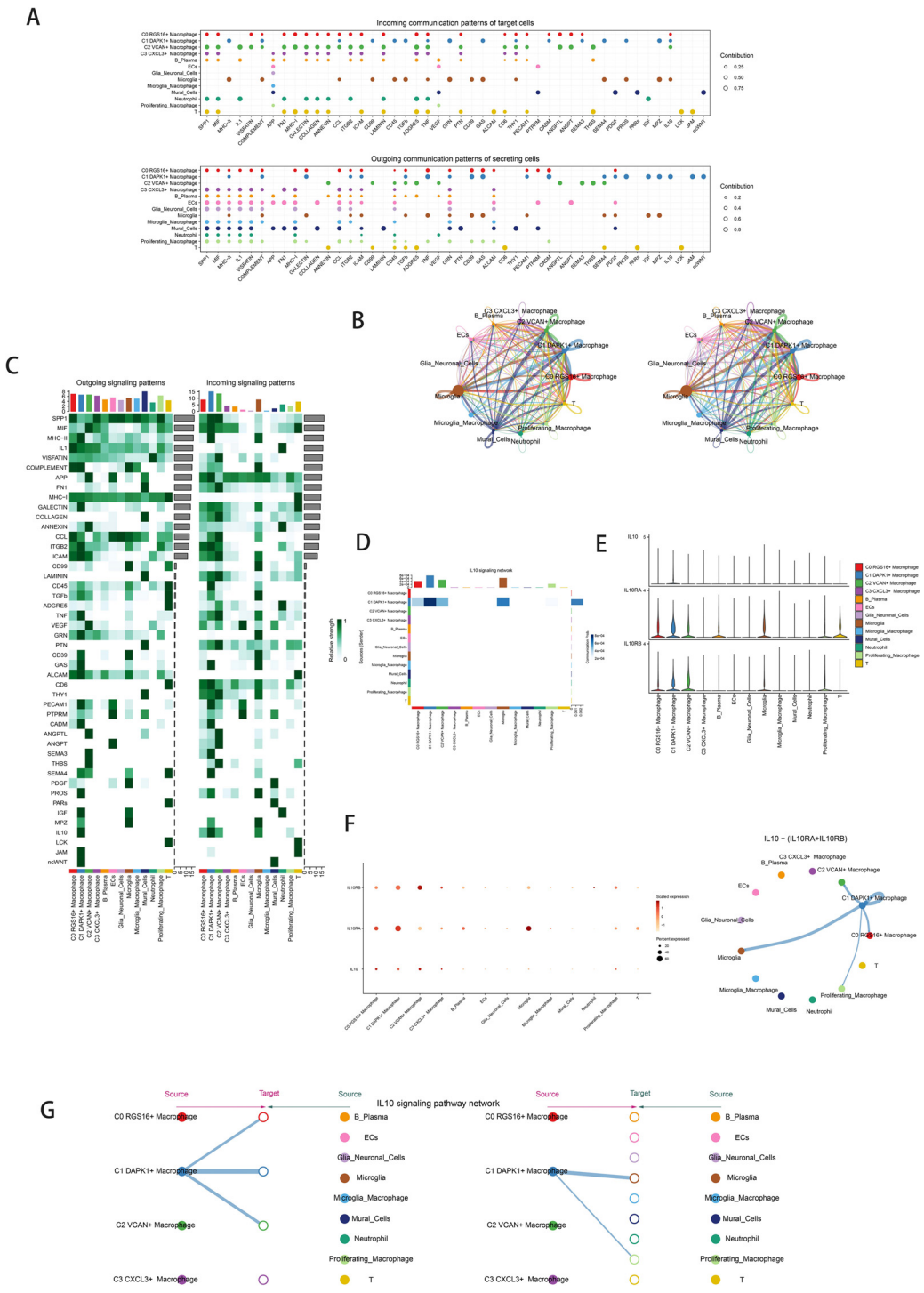


FIGURE 5
Cellchat results presentation. **(A)** Outcoming contribution bubble plots and incoming contribution bubble plots showing the expression of cellular communication patterns between each cell subpopulation and other cells in the macrophage subpopulation. **(B)** Circle plot showing the number (left) and strength (right) of interactions between all cells. **(C)** Heatmap showing afferent and efferent signal intensities of all cell interactions. **(D)** Interaction of cells in the IL-10 signaling pathway shown by heatmap. **(E)** Violin plot of cellular interactions in the IL-10 signaling. **(F)** Interactive bubble diagram of IL-10 macrophage.The color of the dots indicates varying degrees of functional strength and the size of the dots indicates the number of cells. p-value < 0.01, statistically different. **(G)** IL10 Macrophage as a receiver interaction ligand diagram.Hierarchical diagram of macrophage subpopulations interacting with other cells in the IL-10 signaling pathway.IL10 Macrophage as a receiver interaction ligand diagram.

10.3389/fimmu.2024.1463747

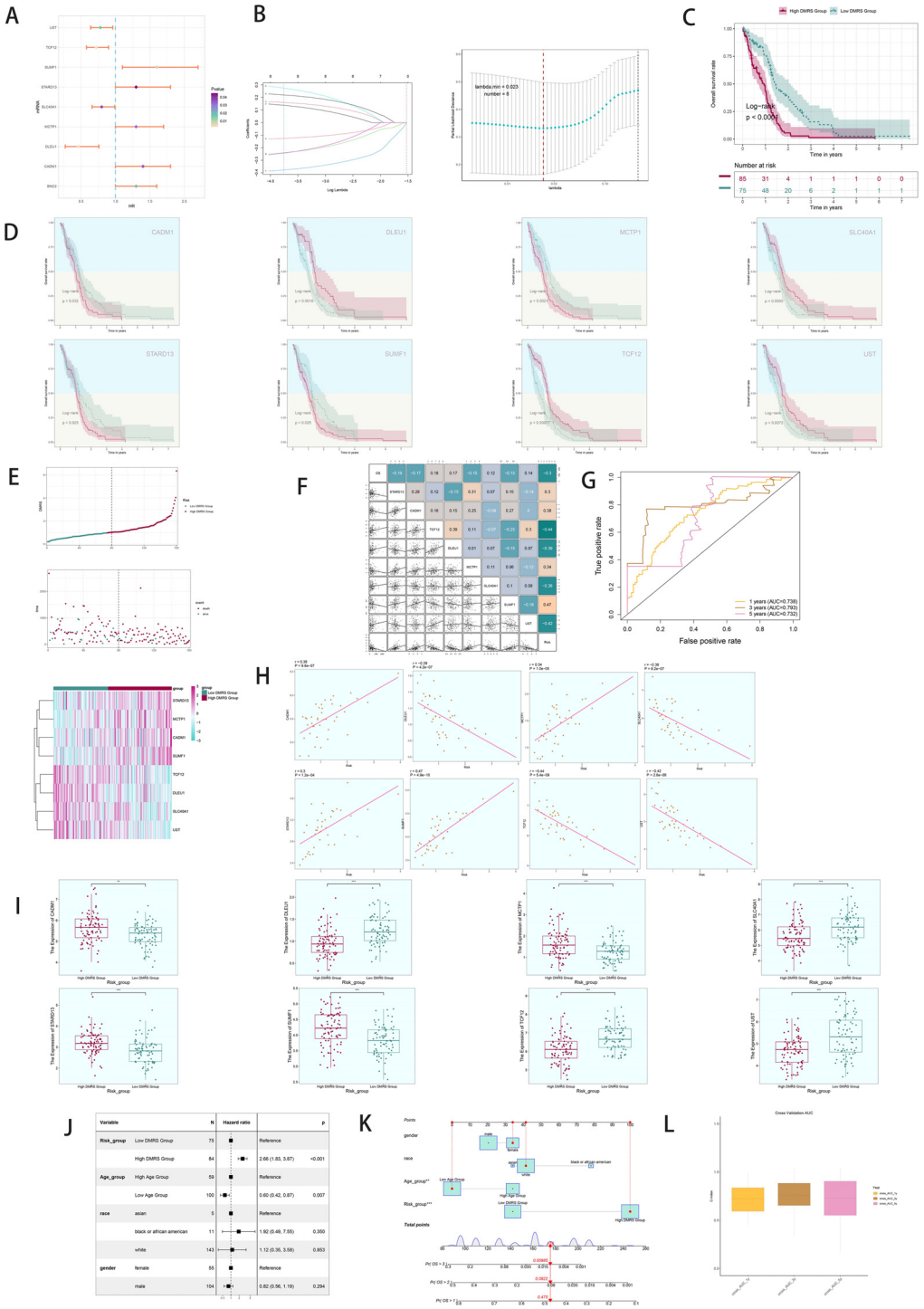


FIGURE 6
Development of a prognostic model associated with DAPK1+ Macrophage scores. **(A)** Forest plot showing univariate cox analysis of genes constituting C1 DAPK1+ Macrophage risk score. Null line HR=1, HR<1 protective factor, HR>1 risk factor. **(B)** 8 genes that constitute DAPK1+ Macrophage scores. screened by lasso regression (top); Lambda plot of genes that constitute DAPK1+ Macrophage scores (right). **(C)** Survival analysis status of the screened 8 genes constituting DAPK1+ Macrophage scores. in both high and low DAPK1+ Macrophage scores. groups. **(D)** Survival analysis plot of the 8 genes constituting the high and low DAPK1+ Macrophage scores. groups. **(E)** Curve plots showing hazard scores in the high and low DAPK1+GBM score groups (top); scatter plot illustrates survival status variations between high and low DAPK1+ Macrophage scores. groups(middle);heatmaps showing gene expression of genes constituting the high and low DAPK1+ Macrophage scores. groups, with color scales based on normalized data (bottom). Green indicates the low DAPK1+ Macrophage scores. group and red indicates the high DAPK1+ Macrophage scores. group. **(F)** Correlation analysis between DAPK1+ Macrophage scores.s, overall survival (OS), and genes used in model establishment. Red indicates positive correlation, blue indicates negative correlation, and color shades indicate high or low correlation. **(G)** AUC scores for 1, 3, and 5 years are shown by ROC plot. AUC(1-year): 0.687, AUC(3-year):0.703, AUC (5-year):0.599. **(H)** Scatter plot showing the correlation analysis of the genes constituting DAPK1+ Macrophage scores. with DAPK1+ Macrophage scores. **(I)** Peak and box plot showing the difference in expression of the eight genes constituting DAPK1+ Macrophage scores. in the high and low DAPK1+GBM score groups. **(J)** Forest plot showing multivariable Cox regression analysis of DAPK1+ Macrophage scores. in conjunction with other clinical factors. Null line HR=1, HR<1 protective factor, HR>1 risk factor. **(K)** Nomogram plots predicting OS (overall survival) at 1, 3, and 5 years based on age, high and low DAPK1+ Macrophage scores. subgroups, and stage. **(L)** Box line plot for internal cross validation of AUC scores at 1, 3, and 5 years. **:p<0.01, ***:p<0.001.

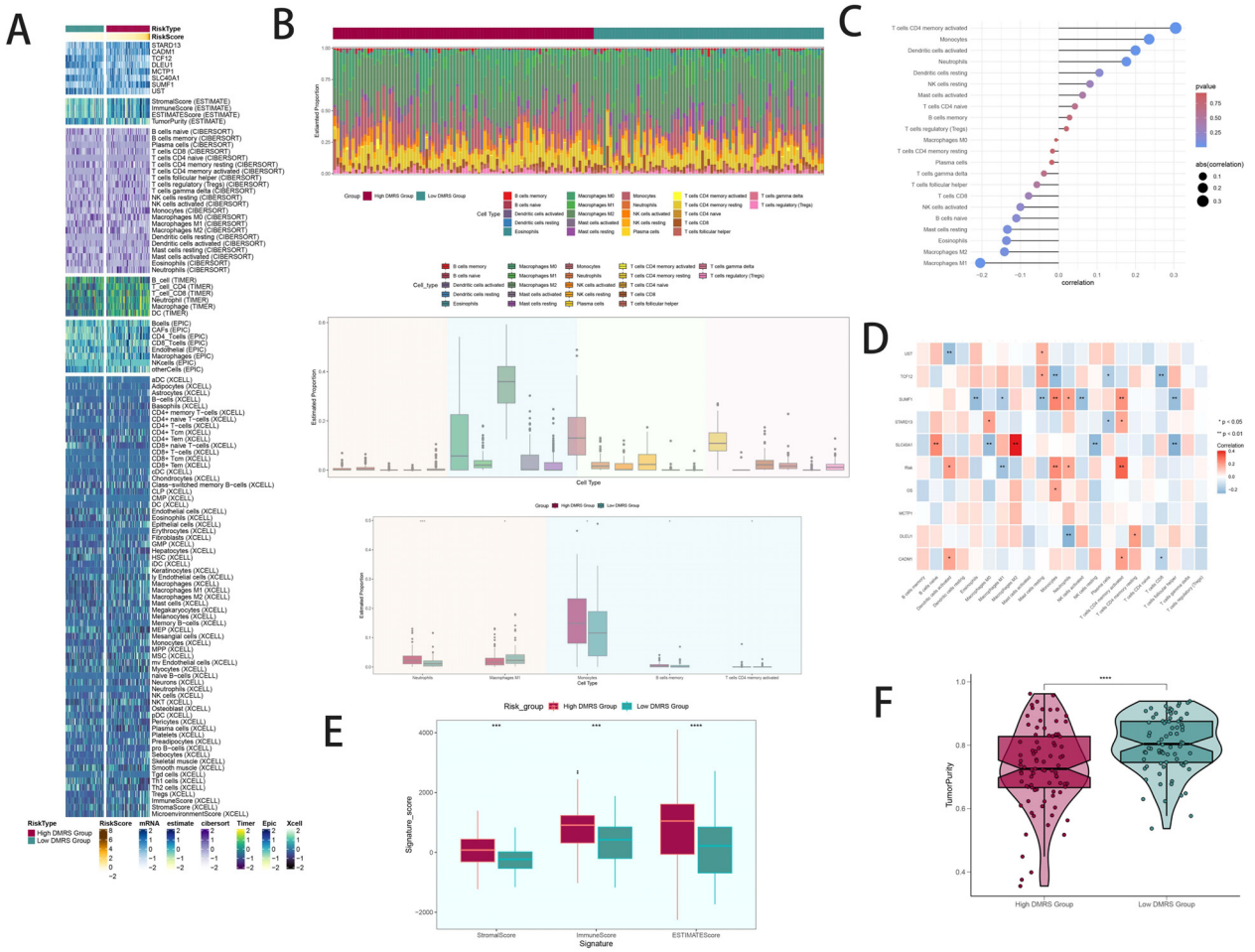


FIGURE 7
Patterns and differences in immune infiltration between high and low DAPK1+ macrophage risk score groups in GBM patients. **(A)** Heatmap visualization of distinct expression patterns of immune infiltration between patients with high and low DAPK1+ Macrophage risk scores. **(B)** Estimation of immune cell proportions using the CIBERSORT algorithm in GBM patients from the TCGA database. Top: Predicted composition of various immune cell subpopulations in the high and low DAPK1+ Macrophage risk score groups. Bottom: Differences in the predicted abundance of five immune cell types between the two groups. **(C)** Bar graph illustrating the correlation between immune infiltrating cells and macrophage subpopulation labeling scores, indicating positive correlations in red and negative correlations in blue. **(D)** Heatmap presentation of the relationship between the eight model genes and immune cells, showing positive correlations in red and negative correlations in blue. **(E)** Evaluation of stromal scores, immune scores, and ESTIMATE scores in the high versus low DAPK1+ Macrophage risk score groups, demonstrating elevated levels in the high-risk group compared to the low-risk group. **(F)** Assessment of tumor purity in the high versus low DAPK1+ Macrophage risk score groups, revealing lower tumor purity in the high-risk group compared to the low-risk group. ***:p<0.001.

(Figures 8C, D). GSEA scoring visualized enrichment scores across pathways (Figure 8E).

3.9 Mutation analysis

Comprehensive mutation analysis illustrated gene mutations and their associations within the tumor microenvironment (TME), revealing differences in mutation load and copy number variation (CNV) between high and low DAPK1+ Macrophage risk score groups (Figures 8F–H). Heatmaps displayed the correlation of mutation profiles among the genes comprising the DAPK1+ Macrophage risk score (Figure 8I). Violin plots depicted mutation loads and their correlation with DAPK1+ Macrophage risk scores (Figure 8J). A scatter plot demonstrated the statistically significant correlation ($p < 0.05$) between mutation load and DAPK1+ Macrophage risk score (Figure 8K). Survival analysis correlated mutation load and DAPK1+ Macrophage risk scores with patient outcomes (Figure 8L).

3.10 Drug sensitivity analysis

Violin plots compared drug sensitivity between high and low DAPK1+ Macrophage risk score groups (Figure 8M), highlighting potential differences in therapeutic responses.

To explore the functional implications of DAPK1 in glioma, we conducted DAPK1 gene knockdown experiments, validating

10.3389/fimmu.2024.1463747

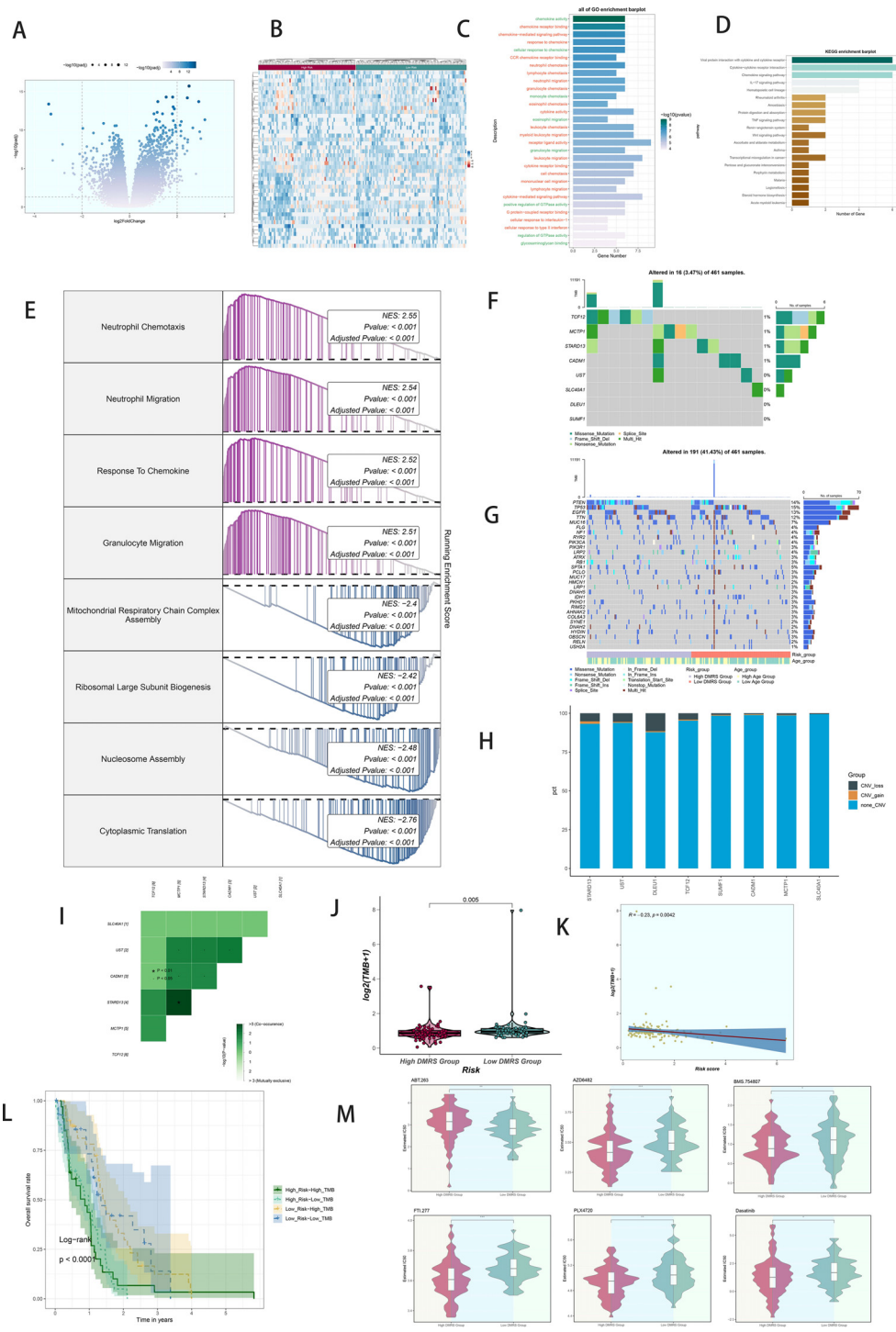


FIGURE 8
Analysis of variance, enrichment, mutation analysis, and drug sensitivity analysis in GBM patients stratified by DAPK1+ macrophage risk score. **(A)** Volcano plot illustrating the expression of differentially expressed genes between the high and low DAPK1+ Macrophage risk score groups. **(B)** Heatmap visualization of the expression patterns of differentially expressed genes between the high and low DAPK1+ Macrophage risk score groups. **(C)** Bar graphs depicting the results of Gene Ontology (GO) enrichment analysis, highlighting associations with chemokine activity, chemokine receptor binding, chemokine-mediated signaling pathway, and response to chemokine. **(D)** Bar graphs presenting the results of Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis, showing significant enrichment of pathways including Viral protein interaction with cytokine and cytokine receptor, Cytokine-cytokine receptor interaction, Chemokine signaling pathway, and IL-17 signaling pathway. **(E)** GSEA scoring of GO-BP-enriched entries of differentially expressed genes, demonstrating the enrichment scores of genes on different pathways. **(F)** Visualization of cellular mutation data depicting mutations in eight genes. **(G)** Comparison of the top 30 genes with the highest mutation frequency in two cohorts of mesenchymal cells. **(H)** Bar graphs illustrating chromosomal copy number variation (CNV) gain and loss. **(I)** Heatmaps displaying the correlation of mutation profiles among genes comprising the DAPK1+ Macrophage risk score. **(J)** Violin plot showing the disparity in mutation load between the high and low DAPK1+ Macrophage risk score groups. **(K)** Scatter plot demonstrating the statistically significant correlation between mutation load and DAPK1+ Macrophage risk score. **(L)** Survival analysis curves depicting outcomes based on mutational load and DAPK1+ Macrophage risk score. **(M)** Violin plots depicting variations in drug sensitivity between the high and low DAPK1+ Macrophage risk score groups, with distinct responses to specific drugs observed. *:p<0.05, **:p<0.01, ***:p<0.001.

transfection efficiency through RT-qPCR (Supplementary Figure 2). Knockdown of DAPK1 significantly inhibited glioma cell proliferation, migration, and metastatic potential, as confirmed by colony formation assays, migration assays, CCK-8 assays, and apoptosis assays (Figure 9).

3.11 Knockdown of DAPK1 inhibits cell proliferation, migration, and induces apoptosis in glioma

To investigate the role of DAPK1 in glioma cell proliferation, migration, and apoptosis, we performed a series of functional assays following DAPK1 knockdown in U251 and LN229 glioma cell lines. Colony formation assays revealed a significant reduction in colony size in the si-DAPK1 groups compared to the negative control (NC), suggesting that DAPK1 knockdown impairs the proliferative capacity of glioma cells (Figures 9A, B). The CCK-8 assay further confirmed these findings, showing a notable decrease in cell viability in the DAPK1 knockdown groups (Figures 9F, G). Scratch and transwell assays were conducted to assess the migratory potential of glioma cells. Representative images from both assays indicated that DAPK1 silencing resulted in impaired migration (Figure 9C). Quantification of the migration assays revealed a significant reduction in the migratory ability of both U251 and LN229 glioma cells upon DAPK1 knockdown, supporting the notion that DAPK1 is essential for glioma cell migration (Figures 9D, E). To examine the effect of DAPK1 knockdown on apoptosis, flow cytometry was performed. The results demonstrated that DAPK1 silencing significantly enhanced apoptosis in both U251 and LN229 glioma cell lines, indicating that the loss of DAPK1 promotes cell death in these tumor cells (Figures 9H, I). Together, these data suggest that DAPK1 knockdown inhibits glioma cell proliferation, migration, and enhances apoptosis, highlighting its potential role as a modulator of glioma cell survival and progression.

4 Discussion

The treatment of glioma remains a significant challenge due to its aggressive nature, high recurrence rate, and poor prognosis (66). Despite advancements in surgical techniques, radiotherapy, and chemotherapy, the therapeutic options for glioma patients are still limited, underscoring the need for novel therapeutic strategies (67). In recent years, immunotherapy has emerged as a promising avenue for glioma treatment, with growing interest in understanding how immune mechanisms contribute to glioma progression (68). However, the pace of research in this area has been relatively slow, particularly in elucidating the complex interactions between glioma cells and the immune microenvironment (69). Studies have shown that the tumor microenvironment, including immune cells, plays a pivotal role in shaping glioma behavior, yet the precise immune-modulatory mechanisms remain poorly understood (70). The role of DAPK1 (Death-associated protein kinase 1) in glioma has garnered attention for its involvement in regulating apoptosis, cell migration, and invasion. DAPK1's function in glioma development is increasingly recognized as being linked to its modulation of immune responses (71). Recent studies suggest that DAPK1 can influence glioma progression through immune mechanisms, particularly by affecting immune cell infiltration and cytokine production. However, there is still a gap in our understanding of how DAPK1 interacts with the immune system to modulate glioma growth and treatment responses. To accelerate progress in this area, single-cell technologies offer a powerful tool to unravel the complex cellular heterogeneity of glioma and its immune microenvironment. Single-cell RNA sequencing (scRNA-seq) enables the high-resolution analysis of individual cell types within tumors, allowing for a more precise mapping of immune cell interactions and the identification of novel therapeutic targets. By integrating single-cell data with immune profiling, researchers can gain deeper insights into the dynamic interplay between glioma cells and the immune system, ultimately advancing the development of more effective immunotherapies for glioma patients (72).

In this study, we employed single-nucleus RNA sequencing (snRNA-seq) to comprehensively characterize the cellular heterogeneity within the glioblastoma (GBM) microenvironment. Our analysis revealed 32 distinct cellular clusters, categorized into 10 major cell types including T cells, neutrophils, microglia, endothelial cells, B/plasma cells, astrocytes, pericytes, macrophages, proliferative macrophages, and microglia/macrophages. We observed significant variations in the distribution of these cell types between the tumor core and peripheral tissues, as well as their differential representation across the cell cycle. Marker gene analyses provided insights into the distinct functional roles and heterogeneous distributions of these cell types within GBM tumors. Notably, our findings underscored the dynamic interplay and functional diversity among different cell populations in the context of GBM pathogenesis. Furthermore, we conducted detailed intratumoral heterogeneity analyses focusing on macrophages, leveraging copy number variation (CNV) analysis to distinguish between normal and cancerous cells within GBM tissues. This approach identified four distinct macrophage subgroups, with one subgroup exhibiting elevated DAPK1 gene expression levels. Utilizing advanced computational methods including UMAP plots, we visualized the distribution patterns of these macrophage subgroups across various parameters such as cell cycle stages, cellular subtypes, and patient samples. Our analysis further highlighted differential gene expression profiles and enriched biological processes within these macrophage subgroups, providing deeper insights into their functional specialization and potential roles in GBM progression. To elucidate the developmental trajectories and differentiation patterns of macrophage and GBM cell subgroups, we employed pseudotime analysis, revealing continuous differentiation trajectories and developmental stages across pseudotime sequences. These findings shed light on the temporal dynamics of cellular differentiation within the GBM microenvironment, suggesting potential developmental milestones and critical points in tumor evolution. Importantly, our study investigated the clinical relevance of identified cell types, particularly focusing on the high DAPK1+ macrophage subgroup,

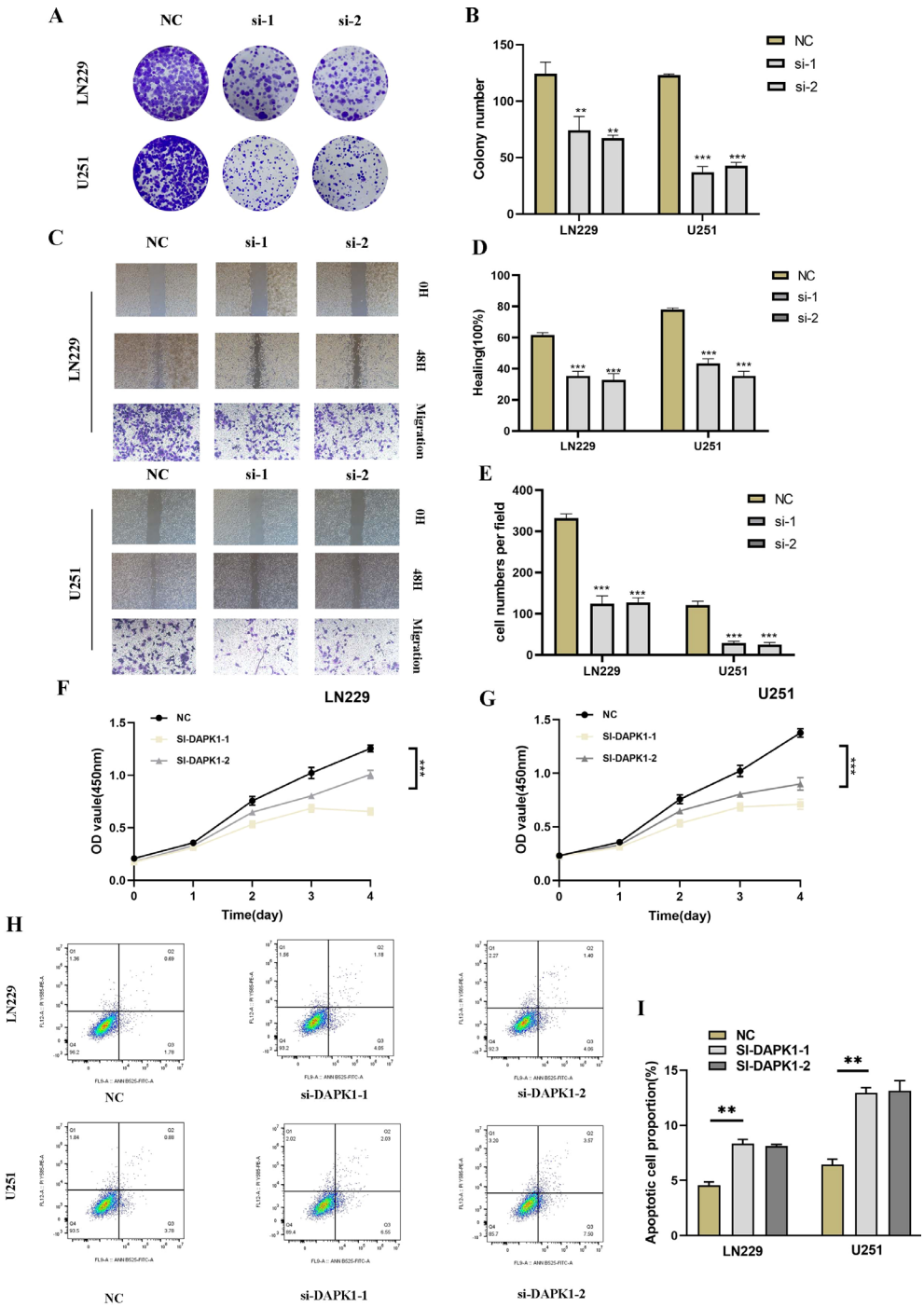


FIGURE 9
Impact of DAPK1 knockdown on glioma cell proliferation, migration, and apoptosis. **(A)** Representative images of colony formation assays conducted on U251 and LN229 glioma cells in the negative control (NC) and si-DAPK1 groups. **(B)** Quantification of colony size indicating the inhibitory effect of DAPK1 knockdown on glioma cell proliferation. **(C)** Representative images of scratch and transwell assays performed to assess the migration capability of U251 and LN229 glioma cells upon DAPK1 knockdown. **(D, E)** Quantification of scratch and transwell assay results demonstrating reduced migration capability following DAPK1 knockdown. **(F, G)** Cell viability assessed by CCK-8 assay indicating the inhibitory effect of DAPK1 knockdown on glioma cell proliferation. **(H, I)** Apoptosis assay results showing enhanced apoptosis in U251 and LN229 glioma cell lines upon DAPK1 downregulation. **:p<0.01, ***:p<0.001.

through comprehensive survival analysis and prognostic modeling. We established a robust multivariable Cox regression model that identified the DAPK1+ macrophage risk score as an independent prognostic factor, correlating significantly with patient survival outcomes. Our findings underscored the prognostic implications of high DAPK1+ macrophage risk scores in GBM, emphasizing their potential as predictive biomarkers in clinical settings. Furthermore, we explored immune infiltration patterns between high and low DAPK1+

macrophage risk score groups, revealing distinct immune cell compositions and interactions within the tumor microenvironment. This analysis highlighted significant correlations between immune cell subtypes and macrophage subgroup markers, implicating intricate immune-modulatory roles of DAPK1+ macrophages in GBM pathophysiology.

Finally, our drug sensitivity analysis demonstrated differential responses to therapeutic agents between high and low DAPK1+ macrophage risk score groups, suggesting potential implications for personalized treatment strategies targeting DAPK1-associated pathways in GBM (73, 74). In conclusion, our study provides a comprehensive characterization of cellular heterogeneity within the GBM microenvironment, highlighting the critical roles of macrophage subpopulations, particularly those with elevated DAPK1 expression. These findings not only deepen our understanding of GBM pathogenesis but also offer novel insights into potential therapeutic targets and prognostic markers for improved patient management. Future investigations should further elucidate the mechanistic underpinnings of DAPK1-mediated pathways in GBM and explore their clinical implications in therapeutic interventions.

The findings from our functional assays underscore the critical role of DAPK1 in regulating glioma cell proliferation, migration, and apoptosis, consistent with previous studies implicating DAPK1 as a key modulator of cancer cell behavior. DAPK1, a pro-apoptotic kinase, has been shown to influence various cellular processes, including cell survival, motility, and death, through its involvement in signaling pathways such as the MAPK and PI3K/Akt pathways. In line with our results, a number of studies have reported that the downregulation of DAPK1 leads to reduced cell proliferation and migration, along with increased apoptosis, in several cancer types, including glioma. For instance, loss of DAPK1 expression in glioma cells has been associated with enhanced tumorigenicity and resistance to apoptosis, further supporting its role as a tumor suppressor. Additionally, DAPK1's influence on mitochondrial dynamics and autophagic processes has been suggested to be a key mechanism underlying its ability to regulate glioma progression. The impaired migratory capacity observed in our study is also consistent with findings that DAPK1 promotes cell adhesion and migration, potentially through its effects on actin cytoskeleton remodeling. Collectively, these results reinforce the notion that DAPK1 serves as an important regulator of glioma cell survival and migration, and its downregulation may contribute to tumor aggressiveness. Therefore, targeting DAPK1 or its downstream signaling pathways may offer promising therapeutic strategies for glioma treatment.

5 Conclusions

In summary, models incorporating DAPK1-related parameters offer robust patient stratification for prognostic evaluation and immunological profiling in glioblastoma. Our findings contribute significant insights into the diagnosis, therapeutic strategies, and mechanistic investigations of gliomas.

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 authors.

Author contributions

T-HY: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. Y-YD: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. S-GZ: Conceptualization, Investigation, Methodology, Project administration, Software, Supervision, Writing – original draft. J-HZ: Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing – original draft. YG: Funding acquisition, Investigation, Project administration, Resources, Software, Validation, Visualization, Writing – original draft. D-HC: Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing. FZ: Conceptualization, Investigation, Project administration, Software, Supervision, Validation, Writing – original draft. W-MH: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Postgraduate Innovation Research and Practice Program of Anhui Medical University (Nos.YJS20230074, Nos.YJS20230184), Youth Foundation of National Natural Science Foundation of China (Nos. 82003795), Youth Project of Anhui Natural Science Foundation (Nos. 2108085QH330), University Natural Science Research Project of Anhui Province (2023AH053318), Scientific Research Fund, Scientific Research and Cultivation Project of Department of Nursing, and Doctor Research Project of the First Affiliated Hospital of Anhui Medical University (Nos. 2021xkj020, hlpy20210014), Quality Engineering Teaching Research Project of Anhui Medical University(No. 2021xjjyxm12 and 2022zyxwjxalk057), Early contact scientific research training plan for clinical medicine (No. 2021-ZQKY-123), Open Project of Key Laboratory of Anti-Inflammatory and Immune Medicine, Ministry of Education (No. KFJJ-2020-01), Clinical Science Foundation project (No. 2023xkj193) and Open Project of Key lab of Dermatology, Ministry of Education, Anhui Medical University(No. AYPYS2022-5).

Acknowledgments

We are very grateful for the data provided by databases such as TCGA, and GEO. Thanks to the reviewers and editors for their sincere comments.

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.

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.2024.1463747/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Discrimination of glioblastoma (GBM) cells from normal cells based on genomic copy number variation (CNV) analysis. The intercnv algorithm was utilized to analyze single-cell data and identify cells with high CNV levels, classifying them as GBM cells.

SUPPLEMENTARY FIGURE 2
DAPK1 gene transfection knock-down low efficiency verification. Compared with untransfected cells, the mRNA level of DAPK1 gene was significantly decreased in the transfected knockdown group.

References

1. Gusyatiner O, Hegi ME. Glioma epigenetics: from subclassification to novel treatment options. *Semin Cancer Biol*. (2018) 51:50–8. doi: 10.1016/j.semcancer.2017.11.010

2. Malta TM, de Souza CF, Sabedot TS, Silva TC, Mosella MS, Kalkanis SN, et al. Glioma cpg island methylator phenotype (G-cimp): biological and clinical implications. *Neuro Oncol*. (2018) 20:608–20. doi: 10.1093/neuonc/nox183

3. van den Bent MJ, Geurts M, French PJ, Smits M, Capper D, Bromberg JEC, et al. Primary brain tumours in adults. *Lancet*. (2023) 402:1564–79. doi: 10.1016/S0140-6736(23)01054-1

4. Nicholson JG, Fine HA. Diffuse glioma heterogeneity and its therapeutic implications. *Cancer Discovery*. (2021) 11:575–90. doi: 10.1158/2159-8290.CD-20-1474

5. Yu K, Lin CJ, Hatcher A, Lozzi B, Kong K, Huang-Hobbs E, et al. Pik3ca variants selectively initiate brain hyperactivity during gliomagenesis. *Nature*. (2020) 578:166–71. doi: 10.1038/s41586-020-1952-2

6. Curry RN, Aiba I, Meyer J, Lozzi B, Ko Y, McDonald MF, et al. Glioma epileptiform activity and progression are driven by igsf3-mediated potassium dysregulation. *Neuron*. (2023) 111:682–95 e9. doi: 10.1016/j.neuron.2023.01.013

7. Wen PY, van den Bent M, Youssef G, Cloughesy TF, Ellingson BM, Weller M, et al. Rano 2.0: update to the response assessment in neuro-oncology criteria for high- and low-grade gliomas in adults. *J Clin Oncol*. (2023) 41:5187–99. doi: 10.1200/JCO.23.01059

8. Thompson EM, Landi D, Brown MC, Friedman HS, McLendon R, Herndon JE 2nd, et al. Recombinant polio-rhinovirus immunotherapy for recurrent paediatric high-grade glioma: A phase 1b trial. *Lancet Child Adolesc Health*. (2023) 7:471–8. doi: 10.1016/S2352-4642(23)00031-7

9. Weller M, Wick W, Aldape K, Brada M, Berger M, Pfister SM, et al. Glioma. *Nat Rev Dis Primers*. (2015) 1:15017. doi: 10.1038/nrdp.2015.17

10. Xie X, Shen C, Zhang X, Wu G, Yang B, Qi Z, et al. Rapid intraoperative multi-molecular diagnosis of glioma with ultrasound radio frequency signals and deep learning. *EBioMedicine*. (2023) 98:104899. doi: 10.1016/j.ebiom.2023.104899

11. Hu H, Mu Q, Bao Z, Chen Y, Liu Y, Chen J, et al. Mutational landscape of secondary glioblastoma guides met-targeted trial in brain tumor. *Cell*. (2018) 175:1665–78 e18. doi: 10.1016/j.cell.2018.09.038

12. Loding S, Andersson U, Kaaks R, Schulze MB, Pala V, Urbarova I, et al. Altered plasma metabolite levels can be detected years before a glioma diagnosis. *JCI Insight*. (2023) 8. doi: 10.1172/jci.insight.171225

13. Mellinghoff IK, van den Bent MJ, Blumenthal DT, Touat M, Peters KB, Clarke J, et al. Vorasidenib in idh1- or idh2-mutant low-grade glioma. *N Engl J Med*. (2023) 389:589–601. doi: 10.1056/NEJMoa2304194

14. Xu S, Tang L, Li X, Fan F, Liu Z. Immunotherapy for glioma: current management and future application. *Cancer Lett*. (2020) 476:1–12. doi: 10.1016/j.canlet.2020.02.002

15. Friebel E, Kapolou K, Unger S, Nunez NG, Utz S, Rushing EJ, et al. Single-cell mapping of human brain cancer reveals tumor-specific instruction of tissue-invading leukocytes. *Cell*. (2020) 181:1626–42 e20. doi: 10.1016/j.cell.2020.04.055

16. Bikfalvi A, da Costa CA, Avril T, Barnier JV, Bauchet L, Brisson L, et al. Challenges in glioblastoma research: focus on the tumor microenvironment. *Trends Cancer*. (2023) 9:9–27. doi: 10.1016/j.trecan.2022.09.005

17. Khan F, Pang L, Dunterman M, Lesniak MS, Heimberger AB, Chen P. Macrophages and microglia in glioblastoma: heterogeneity, plasticity, and therapy. *J Clin Invest*. (2023) 133. doi: 10.1172/JCI163446

18. Zhao T, Zeng J, Xu Y, Su Z, Chong Y, Ling T, et al. Chitinase-3 like-protein-1 promotes glioma progression via the nf-kappab signaling pathway and tumor microenvironment reprogramming. *Theranostics*. (2022) 12:6989–7008. doi: 10.7150/thno.75069

19. Singh SK, Hawkins C, Clarke ID, Squire JA, Bayani J, Hide T, et al. Identification of human brain tumour initiating cells. *Nature*. (2004) 432:396–401. doi: 10.1038/nature03128

20. Sturm D, Capper D, Andreiuolo F, Gessi M, Kolsche C, Reinhardt A, et al. Multiomic neuropathology improves diagnostic accuracy in pediatric neuro-oncology. *Nat Med*. (2023) 29:917–26. doi: 10.1038/s41591-023-02255-1

21. Hambardzumyan D, Gutmann DH, Kettenmann H. The role of microglia and macrophages in glioma maintenance and progression. *Nat Neurosci*. (2016) 19:20–7. doi: 10.1038/nn.4185

22. Chen C, Jing W, Chen Y, Wang G, Abdalla M, Gao L, et al. Intracavity generation of glioma stem cell-specific car macrophages primes locoregional immunity for postoperative glioblastoma therapy. *Sci Transl Med*. (2022) 14:eabn1128. doi: 10.1126/scitranslmed.abn1128

23. Wynn TA, Vannella KM. Macrophages in tissue repair, regeneration, and fibrosis. *Immunity*. (2016) 44:450–62. doi: 10.1016/j.immuni.2016.02.015

24. Guilliams M, Scott CL. Liver macrophages in health and disease. *Immunity*. (2022) 55:1515–29. doi: 10.1016/j.immuni.2022.08.002

25. Ochocka N, Segit P, Walentynowicz KA, Wojnicki K, Cyranowski S, Swatler J, et al. Single-cell rna sequencing reveals functional heterogeneity of glioma-associated brain macrophages. *Nat Commun*. (2021) 12:1151. doi: 10.1038/s41467-021-21407-w

26. Wu M, Wu L, Wu W, Zhu M, Li J, Wang Z, et al. Phagocytosis of glioma cells enhances the immunosuppressive phenotype of bone marrow-derived macrophages. *Cancer Res*. (2023) 83:771–85. doi: 10.1158/0008-5472.CAN-22-1570

27. Ausejo-Mauleon I, Labiano S, de la Nava D, Laspidea V, Zalacain M, Marrodan L, et al. Tim-3 blockade in diffuse intrinsic pontine glioma models promotes tumor regression and antitumor immune memory. *Cancer Cell*. (2023) 41:1911–26 e8. doi: 10.1016/j.ccell.2023.09.001

28. Mei Y, Wang X, Zhang J, Liu D, He J, Huang C, et al. Siglec-9 acts as an immune-checkpoint molecule on macrophages in glioblastoma, restricting T-cell priming and immunotherapy response. *Nat Cancer*. (2023) 4:1273–91. doi: 10.1038/s43018-023-00598-9

29. Pyonteck SM, Akkari L, Schuhmacher AJ, Bowman RL, Sevenich L, Quail DF, et al. Csf-1r inhibition alters macrophage polarization and blocks glioma progression. *Nat Med*. (2013) 19:1264–72. doi: 10.1038/nm.3337

30. Pan Z, Zhao R, Li B, Qi Y, Qiu W, Guo Q, et al. Ewsr1-induced circneil3 promotes glioma progression and exosome-mediated macrophage immunosuppressive

polarization via stabilizing igf2bp3. *Mol Cancer*. (2022) 21:16. doi: 10.1186/s12943-021-01485-6

31. Mantovani A, Marchesi F, Malesci A, Laghi L, Allavena P. Tumour-associated macrophages as treatment targets in oncology. *Nat Rev Clin Oncol*. (2017) 14:399–416. doi: 10.1038/nrclinonc.2016.217

32. Liu J, Cao X. Glucose metabolism of tams in tumor chemoresistance and metastasis. *Trends Cell Biol*. (2023) 33:967–78. doi: 10.1016/j.tcb.2023.03.008

33. Brenner C, Galluzzi L, Kepp O, Kroemer G. Decoding cell death signals in liver inflammation. *J Hepatol*. (2013) 59:583–94. doi: 10.1016/j.jhep.2013.03.033

34. Ghafouri-Fard S, Noroozi R, Abak A, Taheri M, Salimi A. Emerging role of lncrnas in the regulation of rho gtpase pathway. *BioMed Pharmacother*. (2021) 140:111731. doi: 10.1016/j.biopha.2021.111731

35. Gade P, Roy SK, Li H, Nallar SC, Kalvakolanu DV. Critical role for transcription factor C/ebp-beta in regulating the expression of death-associated protein kinase 1. *Mol Cell Biol*. (2008) 28:2528–48. doi: 10.1128/MCB.00784-07

36. Khan ZA, Sumsuzzman DM, Choi J, Hong Y. Neurodegenerative effect of dapk1 after cerebral hypoxia-ischemia is associated with its post-transcriptional and signal transduction regulations: A systematic review and meta-analysis. *Ageing Res Rev*. (2022) 76:101593. doi: 10.1016/j.arr.2022.101593

37. Zhang T, Kim BM, Lee TH. Death-associated protein kinase 1 as a therapeutic target for Alzheimer's disease. *Transl Neurodegener*. (2024) 13:4. doi: 10.1186/s40035-023-00395-5

38. Wang L, Shui X, Mei Y, Xia Y, Lan G, Hu L, et al. Mir-143-3p inhibits aberrant tau phosphorylation and amyloidogenic processing of app by directly targeting dapk1 in Alzheimer's disease. *Int J Mol Sci*. (2022) 23. doi: 10.3390/ijms23147992

39. Shu S, Zhu H, Tang N, Chen W, Li X, Li H, et al. Selective degeneration of entorhinal-ca1 synapses in Alzheimer's disease via activation of dapk1. *J Neurosci*. (2016) 36:10843–52. doi: 10.1523/JNEUROSCI.2258-16.2016

40. Su Y, Deng MF, Xiong W, Xie AJ, Guo J, Liang ZH, et al. Microrna-26a/death-associated protein kinase 1 signaling induces synucleinopathy and dopaminergic neuron degeneration in Parkinson's disease. *Biol Psychiatry*. (2019) 85:769–81. doi: 10.1016/j.biopsych.2018.12.008

41. Li R, Zhi S, Lan G, Chen X, Zheng X, Hu L, et al. Ablation of death-associated protein kinase 1 changes the transcriptomic profile and alters neural-related pathways in the brain. *Int J Mol Sci*. (2023) 24. doi: 10.3390/ijms24076542

42. Lai TW, Zhang S, Wang YT. Excitotoxicity and stroke: identifying novel targets for neuroprotection. *Prog Neurobiol*. (2014) 115:157–88. doi: 10.1016/j.pneurobio.2013.11.006

43. Ding J, Wang C, Sun Y, Guo J, Liu S, Cheng Z. Identification of an autophagy-related signature for prognosis and immunotherapy response prediction in ovarian cancer. *Biomolecules*. (2023) 13. doi: 10.3390/biom13020339

44. Du Y, Kong C. Stat3 regulates mir93-mediated apoptosis through inhibiting dapk1 in renal cell carcinoma. *Cancer Gene Ther*. (2021) 28:502–13. doi: 10.1038/s41417-020-00235-y

45. Pietrocola F, Izzo V, Niso-Santano M, Vacchelli E, Galluzzi L, Maiuri MC, et al. Regulation of autophagy by stress-responsive transcription factors. *Semin Cancer Biol*. (2013) 23:310–22. doi: 10.1016/j.semcancer.2013.05.008

46. Xing J, Cai H, Lin Z, Zhao L, Xu H, Song Y, et al. Examining the function of macrophage oxidative stress response and immune system in glioblastoma multiforme through analysis of single-cell transcriptomics. *Front Immunol*. (2023) 14:1288137. doi: 10.3389/fimmu.2023.1288137

47. Lin Z, Li X, Shi H, Cao R, Zhu L, Dang C, et al. Decoding the tumor microenvironment and molecular mechanism: unraveling cervical cancer subpopulations and prognostic signatures through scrna-seq and bulk rna-seq analyses. *Front Immunol*. (2024) 15:1351287. doi: 10.3389/fimmu.2024.1351287

48. Lin Z, Sui X, Jiao W, Chen C, Zhang X, Zhao J. Mechanism investigation and experiment validation of capsaicin on uterine corpus endometrial carcinoma. *Front Pharmacol*. (2022) 13:953874. doi: 10.3389/fphar.2022.953874

49. Huang W, Kim BS, Zhang Y, Lin L, Chai G, Zhao Z. Regulatory T cells subgroups in the tumor microenvironment cannot be overlooked: their involvement in prognosis and treatment strategy in melanoma. *Environ Toxicol*. (2024). doi: 10.1002/tox.24247

50. Ge Q, Zhao Z, Li X, Yang F, Zhang M, Hao Z, et al. Deciphering the suppressive immune microenvironment of prostate cancer based on cd4+ Regulatory T cells: implications for prognosis and therapy prediction. *Clin Transl Med*. (2024) 14:e1552. doi: 10.1002/ctm2.1552

51. Ding Y, Zhao Z, Cai H, Zhou Y, Chen H, Bai Y, et al. Single-cell sequencing analysis related to sphingolipid metabolism guides immunotherapy and prognosis of skin cutaneous melanoma. *Front Immunol*. (2023) 14:1304466. doi: 10.3389/fimmu.2023.1304466

52. Zhao Z, Ding Y, Tran LJ, Chai G, Lin L. Innovative breakthroughs facilitated by single-cell multi-omics: manipulating natural killer cell functionality correlates with a novel subcategory of melanoma cells. *Front Immunol*. (2023) 14:1196892. doi: 10.3389/fimmu.2023.1196892

53. Liu P, Xing N, Xiahou Z, Yan J, Lin Z, Zhang J. Unraveling the Intricacies of Glioblastoma Progression and Recurrence: Insights into the Role of Nfyb and Oxidative Phosphorylation at the Single-Cell Level. *Front Immunol*. (2024) 15:1368685. doi: 10.3389/fimmu.2024.1368685

54. Zhou W, Lin Z, Tan W. Deciphering the molecular landscape: integrating single-cell transcriptomics to unravel myofibroblast dynamics and therapeutic targets in clear cell renal cell carcinomas. *Front Immunol*. (2024) 15:1374931. doi: 10.3389/fimmu.2024.1374931

55. Wang Y, Zhao ZJ, Kang XR, Bian T, Shen ZM, Jiang Y, et al. Lncrna dleu2 acts as a mir-181a sponge to regulate sepp1 and inhibit skeletal muscle differentiation and regeneration. *Aging (Albany NY)*. (2020) 12:24033–56. doi: 10.18632/aging.104095

56. Zheng RZ, Zhao ZJ, Yang XT, Jiang SW, Li YD, Li WJ, et al. Initial ct-based radiomics nomogram for predicting in-hospital mortality in patients with traumatic brain injury: A multicenter development and validation study. *Neurol Sci*. (2022) 43:4363–72. doi: 10.1007/s10072-022-05954-8

57. Li XY, Zhao ZJ, Wang JB, Shao YH, Hui L, You JX, et al. M7g methylation-related genes as biomarkers for predicting overall survival outcomes for hepatocellular carcinoma. *Front Bioeng Biotechnol*. (2022) 10:849756. doi: 10.3389/fbioe.2022.849756

58. Zheng R, Zhuang Z, Zhao C, Zhao Z, Yang X, Zhou Y, et al. Chinese admission warning strategy for predicting the hospital discharge outcome in patients with traumatic brain injury. *J Clin Med*. (2022) 11. doi: 10.3390/jcm11040974

59. Lin Z, Sui X, Jiao W, Wang Y, Zhao J. Exploring the mechanism and experimental verification of puerarin in the treatment of endometrial carcinoma based on network pharmacology and bioinformatics analysis. *BMC Complement Med Ther*. (2022) 22:150. doi: 10.1186/s12906-022-03623-z

60. Zhao ZJ, Chen D, Zhou LY, Sun ZL, Wang BC, Feng DF. Prognostic value of different computed tomography scoring systems in patients with severe traumatic brain injury undergoing decompressive craniectomy. *J Comput Assist Tomogr*. (2022) 46:800–7. doi: 10.1097/RCT.0000000000001343

61. Zhao ZJ, Wei DP, Zheng RZ, Peng T, Xiao X, Li FS. The gene coexpression analysis identifies functional modules dynamically changed after traumatic brain injury. *Comput Math Methods Med*. (2021) 2021:5511598. doi: 10.1155/2021/5511598

62. Zhao Z, Li T, Dong X, Wang X, Zhang Z, Zhao C, et al. Untargeted metabolomic profiling of cuprizone-induced demyelination in mouse corpus callosum by uplc-orbitrap/ms reveals potential metabolic biomarkers of cns demyelination disorders. *Oxid Med Cell Longev*. (2021) 2021:7093844. doi: 10.1155/2021/7093844

63. Zhao ZJ, Zheng RZ, Wang XJ, Li TQ, Dong XH, Zhao CY, et al. Integrating lipidomics and transcriptomics reveals the crosstalk between oxidative stress and neuroinflammation in central nervous system demyelination. *Front Aging Neurosci*. (2022) 14:870957. doi: 10.3389/fnagi.2022.870957

64. Lin Z, Fan W, Yu X, Liu J, Liu P. Research into the mechanism of intervention of sanqi in endometriosis based on network pharmacology and molecular docking technology. *Med (Baltimore)*. (2022) 101:e30021. doi: 10.1097/MD.0000000000030021

65. Zhao J, Jiao W, Sui X, Zou J, Wang J, Lin Z. Construction of a prognostic model of luteolin for endometrial carcinoma. *Am J Transl Res*. (2023) 15:2122–39.

66. Lucas CG, Sloan EA, Gupta R, Wu J, Pratt D, Vasudevan HN, et al. Multiplatform molecular analyses refine classification of gliomas arising in patients with neurofibromatosis type 1. *Acta Neuropathol*. (2022) 144:747–65. doi: 10.1007/s00401-022-02478-5

67. Li Z, Li J, Bai X, Huang X, Wang Q. Tumor microenvironment as a complex milieu driving cancer progression: A mini review. *Clin Transl Oncol*. (2024). doi: 10.1007/s12094-024-03697-w

68. Shen G, Wang Q, Li Z, Xie J, Han X, Wei Z, et al. Bridging chronic inflammation and digestive cancer: the critical role of innate lymphoid cells in tumor microenvironments. *Int J Biol Sci*. (2024) 20:4799–818. doi: 10.7150/ijbs.96338

69. Wang Q, Wang H, Ding Y, Wan M, Xu M. The role of adipokines in pancreatic cancer. *Front Oncol*. (2022) 12:926230. doi: 10.3389/fonc.2022.926230

70. Chen P, Zhao D, Li J, Liang X, Li J, Chang A, et al. Symbiotic macrophage-glioma cell interactions reveal synthetic lethality in pten-null glioma. *Cancer Cell*. (2019) 35:868–84 e6. doi: 10.1016/j.ccell.2019.05.003

71. Saha D, Martuza RL, Rabkin SD. Macrophage polarization contributes to glioblastoma eradication by combination immunovirotherapy and immune checkpoint blockade. *Cancer Cell*. (2017) 32:253–67 e5. doi: 10.1016/j.ccell.2017.07.006

72. Xuan F, Huang M, Liu W, Ding H, Yang L, Cui H. Homeobox C9 suppresses beclin1-mediated autophagy in glioblastoma by directly inhibiting the transcription of death-associated protein kinase 1. *Neuro Oncol*. (2016) 18:819–29. doi: 10.1093/neuonc/nov281

73. Zhao Q, Wang F, Chen YX, Chen S, Yao YC, Zeng ZL, et al. Comprehensive profiling of 1015 patients' exomes reveals genomic-clinical associations in colorectal cancer. *Nat Commun*. (2022) 13:2342. doi: 10.1038/s41467-022-30062-8

74. Fu D, Zhang B, Wu S, Zhang Y, Xie J, Ning W, et al. Prognosis and characterization of immune microenvironment in acute myeloid leukemia through identification of an autophagy-related signature. *Front Immunol*. (2021) 12:695865. doi: 10.3389/fimmu.2021.695865

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 24 January 2025
DOI 10.3389/fimmu.2024.1466806

OPEN ACCESS

EDITED BY
Wenyi Jin,
City University of Hong Kong,
Hong Kong SAR, China

REVIEWED BY
Yuquan Chen,
Monash University, Australia
Jing Zhang,
University of South Dakota, United States
Xin Yu,
Baylor College of Medicine, United States

*CORRESPONDENCE
Shijian Zhao
zhaoshijian1025@163.com
Yinteng Wu
wuyinteng@stu.gxmu.edu.cn

RECEIVED 18 July 2024
ACCEPTED 20 December 2024
PUBLISHED 24 January 2025

CITATION
Huang D, Li J, He Z, Liang W, Zhong L, Huang J, Wu Y and Zhao S (2025) Pan-cancer and experimental analyses reveal the immunotherapeutic significance of CST2 and its association with stomach adenocarcinoma proliferation and metastasis. *Front. Immunol.* 15:1466806. doi: 10.3389/fimmu.2024.1466806

Pan-cancer and experimental analyses reveal the immunotherapeutic significance of CST2 and its association with stomach adenocarcinoma proliferation and metastasis

Dan Huang[1], Jing Li[1], Zhijun He[1], Wenjing Liang[1], Likun Zhong[1], Jun Huang[1], Yinteng Wu[2]* and Shijian Zhao[3]*

[1]Department of Gastroenterology, The People's Hospital of Guangxi Zhuang Autonomous Region (Guangxi Academy of Medical Sciences), Nanning, Guangxi, China, [2]Department of Orthopedic and Trauma Surgery, The First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, China, [3]Department of Cardiology, The Affiliated Cardiovascular Hospital of Kunming Medical University (Fuwai Yunnan Cardiovascular Hospital), Kunming, Yunnan, China

Purpose: Cystatin 2 (CST2) is a cysteine protease inhibitor, and recent research suggests its potential involvement in cancer development. However, its role in the occurrence, progression, and prognosis of pan-cancer has not been systematically investigated.

Materials and methods: This study comprehensively analyzes the differential expression of CST2 in pan-cancer. The expression distribution patterns of CST2 were examined using single-cell datasets. Furthermore, we conducted a comprehensive evaluation of the correlation between CST2 expression and various factors. These factors include prognosis, immune cell infiltration, immune-related genes, mutations, methylation, tumor mutation burden (TMB), and microsatellite instability (MSI). In addition, we analyzed the sensitivity of drugs dependent on CST2 expression. We utilized gene set enrichment analysis (GSEA) analysis to explore the biological functions of CST2 across different cancer types. Finally, in gastric cancer cell lines, we will investigate the impact of CST2 knockout on expression levels, clonal proliferation, cell apoptosis, and cell migration.

Results: CST2 exhibits abnormal overexpression in multiple tumors. Single-cell analysis reveals high expression of CST2 in fibroblasts. CST2 is closely associated with prognosis, immune cell infiltration, immune-related genes, mutations, methylation, TMB, and MSI. Enrichment analysis demonstrated a significant correlation between CST2 and immune-related pathways. In stomach adenocarcinoma (STAD), CST2-related risk models are associated with prognosis and demonstrate strong predictive capabilities, while also being closely linked to

the immune microenvironment. Drug sensitivity analysis indicates the correlation between CST2 and 21 chemotherapy drugs. Finally, experimental validation revealed significantly elevated expression of CST2 in STAD, indicating its role as a driver gene in regulating malignant cell proliferation and migration.

Conclusion: CST2 serves as a potential tumor immune biomarker, playing a critical facilitating role in the proliferation and migration processes of STAD.

KEYWORDS

cystatin 2 (CST2), pan-cancer, stomach adenocarcinoma (STAD), the tumor microenvironment (TME), immunotherapy, single-cell

Introduction

Cancer remains a leading cause of global human mortality (1). Given its high incidence and mortality rates, the pursuit of more valuable and targeted biomarkers for early diagnosis, treatment, and prevention is of paramount importance and urgency. The involvement of the tumor immune microenvironment in cancer initiation and progression has been gradually elucidated. Currently, molecular targeted therapies are progressively being applied in clinical practice. The discovery of biomarkers has significantly accelerated the development of anti-cancer drugs.

With the widespread adoption of high-throughput sequencing technologies and the improvement of tumor data sharing platforms, pan-cancer research is receiving increasing attention. By combining and analyzing cancers originating from different organs, pan-cancer studies can provide a deeper and broader understanding of common oncogenic signaling pathway characteristics, allowing researchers to focus on datasets with relatively larger sample sizes. Larger sample sizes enhance the statistical power of the data and make it easier to identify cancer-associated genomic alterations, potentially leading to the discovery of previously unidentified drug targets. Additionally, new tumor classification methods based on an understanding of common signaling pathway features can help certain cancer patients receive more personalized treatments, increasing the likelihood of disease relief (2).

Cystatin 2 (CST2) is a gene encoding a protein that belongs to the cysteine protease inhibitor superfamily (3). Previous studies have indicated that CST2 can predict disease progression in certain non-tumor conditions (4, 5). A recent study discovered that CST2 is overexpressed in pancreatic cancer, functioning as an oncogene. Knockdown of CST2 in pancreatic cancer inhibits tumor cell proliferation, migration, and invasion, while also suppressing the activation of the PI3K/AKT signaling pathway (6). Elevated levels of CST2 in colorectal cancer are associated with shortened overall survival in patients (7). The upregulation of CST2 has been linked to breast cancer development (8). In gastric cancer samples, CST2 is upregulated, enhancing tumor cell growth, migration, and invasion by regulating epithelial-mesenchymal transition (EMT) and the TGF-β1 signaling pathway, leading to poor prognosis in patients (9). CST2 may participate in prostate cancer metastasis by modulating the EMT signaling pathway (10). CST2 is associated with overall survival rate (OS) in hepatocellular carcinoma (11). High expression of CST2 promotes bone metastasis occurrence (12), which is common in solid tumors. Therefore, dysregulation of CST2 is implicated in human cancer. Nevertheless, there remains a need for a comprehensive understanding of CST2's role in pan-cancer. Hence, further exploration of the mechanisms underlying CST2's involvement in tumors holds significant importance in providing new directions and strategies for clinical cancer treatment.

By integrating multiple databases, we conducted an analysis of CST2 expression levels in pan-cancer and its relationship with tumor-infiltrating immune cells, immune-related genes, mutations, DNA methylation, tumor mutation burden (TMB), microsatellite instability (MSI), and their impact on patient prognosis. The results demonstrate that CST2 overexpression in pan-cancer contributes to carcinogenesis and is closely associated with the tumor immune microenvironment (TIM). Furthermore, we performed molecular biology validation in gastric cancer to further confirm the oncogenic role of CST2. In summary, CST2 represents a promising therapeutic target in cancer treatment and serves as a potential biomarker for predicting immunotherapy efficacy and prognosis.

Materials and methods

CST2 expression analysis in human pan-cancer

CST2 expression data in 35 normal tissues were obtained and downloaded from the Genotype-Tissue Expression (GTEx, https://commonfund.nih.gov/GTEx) database. Additionally, CST2 expression data from 31 tumor cell lines were obtained from the Cancer Cell Line Encyclopedia (CCLE, https://portals.broadinstitute.org/ccle/) database. By combining the data from normal tissues in the GTEx database and the cancer genomic atlas from The Cancer Genome Atlas (TCGA, https://

www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga), the differential expression of CST2 between cancer tissues and normal tissues was analyzed. Data from 15 different types of cancer were retrieved from the TCGA database. We also investigated the expression levels of CST2 in different clinical stages.

Single-cell analysis

To further analyze the single-cell expression distribution pattern of CST2 in pan-cancer, we analyzed various single-cell datasets, including BRCA, BTCC, CAC, CCRCC, CRPC, ESCC, NNSC, ICC, NPC, NSCLC, OV, PDAC, STAD, UCEC. These data single-cell data were obtained from the Single-cell and Spatially-resolved Cancer Resources (SCAR) database (http://scaratlas.com).

Prognostic assessment in pan-cancer

Utilize the R package "survival" to establish a univariate Cox regression model. Evaluate the prognostic value of CST2 in different tumor types based on Overall Survival (OS), Progression Free Interval (PFI), Disease Specific Survival (DSS), and Disease Free Interval (DFI). Plot Kaplan-Meier survival analysis curves to depict the relationship between CST2 expression and OS, DSS, disease-free interval (DFS), and progression-free interval (PFS) specifically in STAD. The Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) database was used to validate the prognostic analysis of CST2 in STAD. We also analyzed the prognostic significance of CST2 in the immunotherapy cohort. Prognostic evaluation criteria include Hazard Ratio (HR), 95% confidence intervals, and p-values considered statistically significant when $p<0.05$.

Immunological analyses

The TIMER2.0 database (https://cistrome.shinyapps.io/timer/) and Biomarker Exploration for Solid Tumors database (https://rookieutopia.hiplot.com.cn/app_direct/BEST/) can be employed to analyze tumor immune cell infiltration., which can be utilized for correlation analysis between CST2 and various types of immune cells. Visualization of the analysis results can be done using the "ggplot2" R package. Furthermore, the ESTIMATE algorithm can be applied to calculate the ESTIMATEScore, ImmuneScore, and StromalScore for different tumor types. Spearman algorithm can be used to determine the correlation coefficients between CST2 expression and these three scoring systems.

Immunological correlation analysis

In each sample, the expression data for 130 immune-related genes were extracted, including 38 chemokines, 43 immunostimulators, 18 receptors, 8 immune checkpoint genes, and 23 immunoinhibitors. Subsequently, the Spearman algorithm was utilized to calculate the correlations between CST2 and individual immune-related genes.

The evaluation of CST2 in pan-cancer regarding mutations, methylation, TMB, and MSI

The CBioPortal database (https://www.cbioportal.org/) was used to explore the mutation characteristics and locations of CST2 in tumors. To analyze the relationship between CST2 expression levels and the methylation status of its promoter region, we utilized the TCGA database and visualized the results using the "ggplot2" R package. We employed Spearman's test in the TCGA database to assess the correlation between CST2 expression and TMB as well as MSI across different tumor types. The correlation results were then visualized using the "fmsb" R package.

Enrichment analysis of CST2

To gain further insights into the biological functions and molecular mechanisms of CST2, we conducted Gene Set Enrichment Analysis (GSEA) using the R package ClusterProfiler. In this analysis, we employed the hallmark gene set from the Molecular Signatures Database (MSigDB), which consists of 50 gene sets associated with key cancer pathways. For each pathway, we calculated the Normalized Enrichment Score (NES) and False Discovery Rate (FDR) to assess the enrichment of CST2. Statistical significance was determined at a p-value < 0.05.

Drug sensitivity analysis

To investigate the correlation between CST2 expression levels and drugs, we employed two databases. Firstly, the CellMiner database (https://ngdc.cncb.ac.cn/databasecommons/database/id/6092) allowed us to analyze gene expression profiles and their association with drug response. Furthermore, we explored the expression levels of CST2 in gastric cancer and its correlation with drugs using the Genomics of Drug Sensitivity in Cancer (GDSC, https://www.cancerrxgene.org) database and Cancer Therapeutics Response Portal (CTRP, https://portals.broadinstitute.org/ctrp.v2.1/) database.

Cell culture and transfection

Human STAD cell lines MKN-45 and SGC-7901 were obtained from X-Y Biotechnology and maintained in DMEM medium supplemented with 10% fetal bovine serum at 37ru in a cell incubator with 5% CO2. For transfection experiments, 3 × 10^5 thyroid cancer cells were seeded into 6-well dishes and cultured for 24 hours to allow for cell attachment. Transfections were performed using Lipofectamine 3000 reagent (Thermo Fisher Scientific) according to the manufacturer's protocol. Biological experiments were conducted following the appropriate transfection period. The siRNA sequences used in this study are as follows: si-CST2-1: 5'-GC UCCUCGAGACAUGUAAU-3' (targeting 70-90bp downstream of the start codon, the Antisense strand: 5'-AUUACAUGACU

CGAGAAGC-3'), si-CST2-2: 5'-GGACGAGGUUCUUGUAAAU-3' (targeting 150-170bp downstream of the start codon, the Antisense strand: 5'-AUUUCCAAGAACCACGUCC-3').

Plate clone formation assay

Cells were collected and resuspended to obtain a single-cell suspension. The cells were seeded in 35 mm cell culture dishes at a density of 200 cells per dish. The medium was replaced every 2–3 days, with half of the medium being replaced each time. After approximately 14 days, the colonies were analyzed. Clones with more than 20 cells were counted under a microscope. Subsequently, the colonies were fixed with methanol and stained using a 1% crystal violet staining solution (G1063, Solarbio, China).

Cell migration assay

During the transwell invasion experiment, first add 150 μL of DMEM culture medium with an additional 10% fetal bovine serum (FBS) to the lower chamber. Then, seed a total of 2 × 10^4 MKN-45 and SGC-7901 cells in the upper chamber of the transwell device. After 24 hours of incubation, carefully remove the non-invasive cells on the membrane surface. Next, fix the cells that have successfully penetrated the membrane and invaded the lower surface, followed by staining with crystal violet dye (G1063, Solarbio, China). Finally, use a microscope to capture images of the stained invasive cells and quantify the number of invasive cells by counting the cells in three randomly selected fields per well.

Flow-cytometric analysis

To evaluate cell apoptosis, first seed MKN-45 and SGC-7901 cells into a 6-well plate. Then, transfect the cells with si-NC and si-CST2 separately for 24 hours. After transfection, digest the cells with trypsin and wash them twice with pre-chilled PBS (4S-c to remove residual culture medium and enzymes. The washed cells are resuspended in binding buffer, and according to the instructions provided by Absin company (China) using the Annexin V-FITC/PI Cell Apoptosis Detection Kit (abs50001-25T), stain the cells with Annexin V-FITC and propidium iodide (PI). The stained cells are transferred to tubes specifically designed for flow cytometry analysis and cell apoptosis is detected using a BD FACSCalibur flow cytometer. Data collection and analysis are performed using FlowJo software. The percentage of apoptotic cells is accurately calculated by distinguishing Annexin V-positive and PI-negative cells (early apoptosis) from Annexin V-positive and PI-positive cells (late apoptosis).

Western blot

The cells are first washed with ice-cold PBS, followed by lysis using a protein extraction kit (20127ES60, Yeasen). The collected cell samples are centrifuged at 700g for 10 minutes at 4°C. After centrifugation, the supernatant is transferred to a new Eppendorf tube to avoid particle contamination. Subsequently, at 4tbs the supernatant is centrifuged again at 14,000g for 30 minutes to pellet cell membrane debris. The cell pellet is then resuspended in 200 μL of BCA protein assay kit B solution (containing PMSF), vortexed for 5 seconds, and placed on ice for 5 minutes. This step is repeated twice. After the initial step, centrifugation is performed at 14,000g for 5 minutes at 4tn to collect the supernatant containing the membrane proteins. The protein concentration is determined using the BCA protein assay kit (20201ES76, Yeasen) following standard operating procedures. Subsequently, total protein is separated by electrophoresis on a 4-20% Bis-Tris gel (Genscript China) and transferred to a PVDF membrane (ISEQ00010, Millipore). The membrane is initially incubated at room temperature in TBST buffer (Tris-buffered saline with Tween) containing 5% skim milk for 1 hour to block non-specific binding. Following the blocking step, it is further incubated overnight at 4te with primary antibodies (including Anti-CST2, 1:1000, abs115674, Absin; and Anti-GAPDH, 1:2000, ab8245, Abcam) diluted in TBST with 0.5% skim milk. After the primary antibody incubation, the membrane is washed three times with TBST and subsequently incubated with HRP-conjugated secondary antibodies (34201ES60, 34101ES60, Yeasen) diluted in TBST at room temperature for 1 hour. Immunoreactive bands are detected using an ECL Western blotting substrate (36208ES60, Yeasen). Finally, the density of the bands is quantitatively analyzed using ImageJ software.

Results

Aberrant expression of CST2 in human pan-cancer

Initially, we analyzed the expression of CST2 across 35 normal tissues utilizing the GTEx database. As depicted in Figure 1A, we observed relatively high expression levels of CST2 in Fibroblast, Skin, and Vagina tissues. Additionally, we accessed data from the CCLE database to investigate CST2 expression in 31 tumor cell lines. As shown in Figure 1B, CST2 was expressed in all 26 types of tumor cells examined. To determine the differential expression of CST2 between tumor and normal tissues, we conducted an analysis using the TCGA database for 33 different cancer types. The results revealed significantly higher expression levels of CST2 in BLCA, BRCA, CHOL, COAD, ESCA, KIRC, KIRP, LIHC, LUAD, LUSC, PRAD, READ, STAD, THCA, and UCEC tissues when compared to normal tissues (Figure 1C). Considering the limited number of normal samples available in TCGA, we integrated data from the GTEx and TCGA databases to analyze CST2 expression differences in 15 tumor types. We found that CST2 was significantly upregulated in 13 tumor types, including BLCA, BRCA, COAD, ESCA, KIRC, KIRP, LIHC, LUAD, LUSC, PRAD, STAD, THCA, and UCEC, as compared to their respective normal tissues (Figure 1D). These findings indicate the aberrant overexpression of CST2 in human pan-cancer.

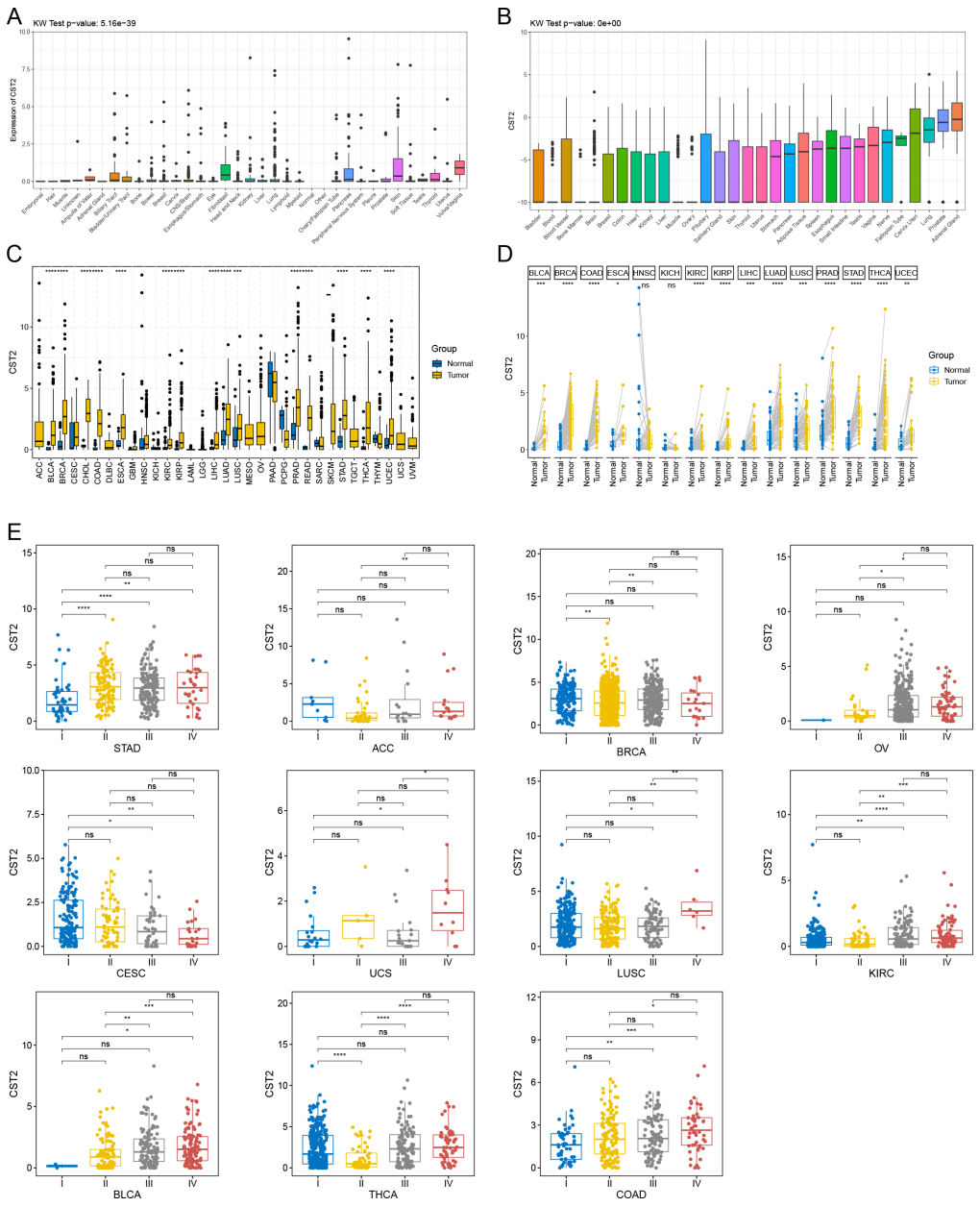


FIGURE 1
Cystatin 2 (CST2) aberrant expression and its correlation with clinical stages in human pan-cancer **(A)** Expression of CST2 across 35 tissues in the Genotype-Tissue Expression (GTEx) database. **(B)** Expression of CST2 in 31 tumor cell lines in the Cancer Cell Line Encyclopedia (CCLE) database. **(C)** Differential expression of CST2 between cancerous and normal tissues in The Cancer Genome Atlas (TCGA) database. **(D)** Aberrant overexpression of CST2 in 13 types of cancer based on analysis of the GTEx and TCGA databases. **(E)** Variations in CST2 expression levels among distinct clinical stages in STAD, ACC, BRCA, OV, CESC, UCS, LUSC, KIRC, BLCA, THCA, and COAD. (*p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001).

Different clinical stages

Through comprehensive analysis and evaluation, we have discovered distinct expression patterns of CST2 across different clinical stages in several tumor types. Specifically, higher levels of CST2 expression were observed in the advanced stages of STAD, ACC, BRCA, OV, UCS, LUSC, KIRC, BLCA, THCA, and COAD (Figure 1E). However, in the case of CESC, CST2 exhibited lower expression levels in the advanced stages.

Single-cell analysis

As depicted in Figure 2 and Supplementary Figure S1, the single-cell analysis results reveal distinct expression patterns of CST2 across various cancer types. Specifically, in BRCA, STAD, BTCC, NSCLC, and PDAC, CST2 exhibits predominantly high expression levels in the Fibroblast cell population. In CRPC, CST2 is primarily highly expressed in the Luminal cell population. On the other hand, in OV and ICC, CST2 shows elevated expression levels mainly in the Malignant cell population.

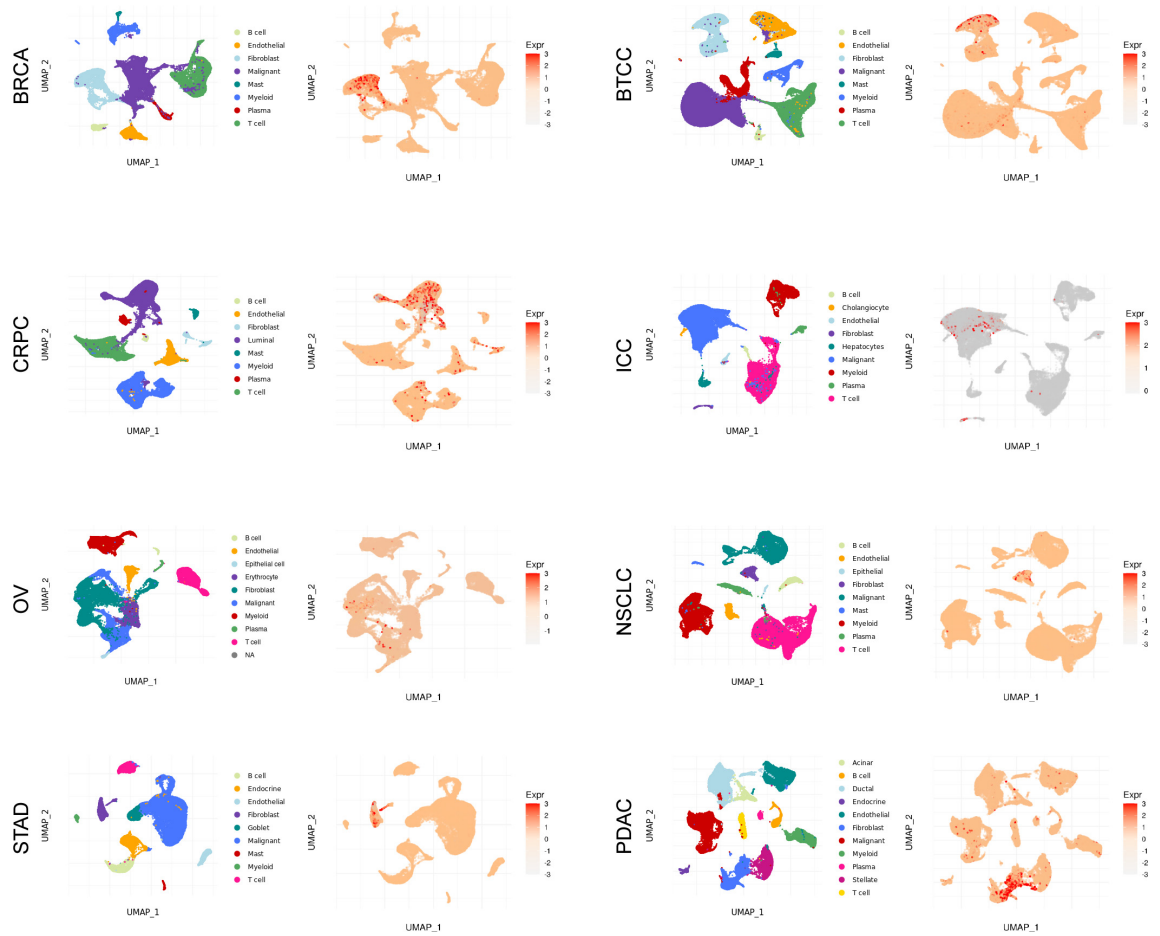


FIGURE 2
Analysis of single-cell expression distribution patterns of CST2 in pan-cancer.

Survival analysis

COX regression analysis revealed that high expression of CST2 is a risk factor for OS in 10 tumor types, including STAD, SKCM, READ, PAAD, LGG, KIRP, KIRC, HNSC, GBM, and COAD (Figure 3A). Further investigations demonstrated a significant correlation between CST2 expression and PFI in several cancer categories, including STAD, SKCM, READ, PAAD, LGG, KIRC, GBM, and COAD (Figure 3B). Moreover, elevated expression of CST2 in STAD, SKCM, PAAD, LGG, KIRP, KIRC, GBM, and COAD was associated with improved DSS (Figure 3C). Additionally, the univariate Cox regression model established a link between CST2 expression and adverse prognosis in DFI for STAD and PAAD (Figure 3D). The survival analysis of CST2 in OS, PFI, DSS, and DFI highlighted its prognostic value across STAD (Figure 3E). Furthermore, Kaplan-Meier survival analysis curves were employed to explore the relationship between CST2 expression and OS, DSS, DFS, and PFS specifically in STAD. It was observed that patients with low CST2 expression had better outcomes compared to those with high CST2 expression (Figure 3F). The GEO database validated the relationship between CST2 and STAD prognosis in OS (Figure 4A), RFS (Figure 4B) and PFS (Figure 4C).

Immune cell infiltration

The correlation heatmap reveals that CST2 is closely associated with various immune cells in pan-cancer, including cancer-associated fibroblasts (CAF), endothelial cells, hematopoietic stem cells, regulatory T cells (Tregs), B cells, macrophages, monocytes, myeloid dendritic cells, and CD8+ T cells (Figures 5A, B). According to the results obtained from the ESTIMATE algorithm, CST2 exhibits a significant positive correlation with the ESTIMATEScore, ImmuneScore, and StromalScore across different tumor types such as GBM, UCEC, KIRP, LUSC, KIRC, SARC, COAD, READ, BLCA, LIHC, HNSC, STAD, ESCA, THYM, PAAD, PCPG, THCA, and LGG (Figure 6A).

Correlation analysis of CST2 and immune-related genes

The study findings indicate that CST2 exhibits significant positive correlations with several chemokines, such as CCL19, CCL21, CCL11, CCL26, CXCL14, CX3CL1, and CXCL12, in most tumors (Figure 6B). Furthermore, CST2 shows significant positive

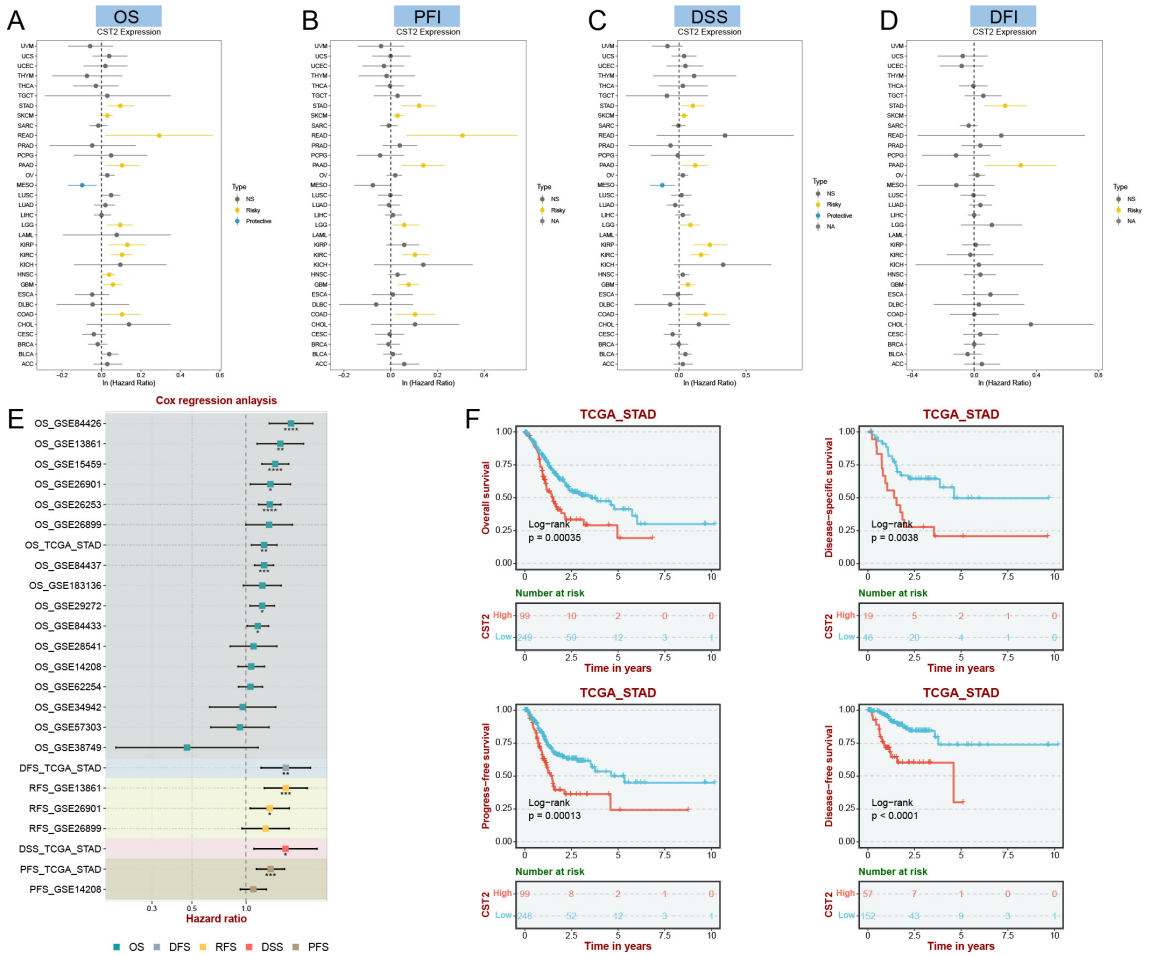


FIGURE 3
Univariate COX regression analysis was performed to examine the association of CST2 with Overall Survival (OS), Progression Free Interval (PFI), Disease Specific Survival (DSS), and Disease Free Interval (DFI) in pan-cancer. The correlation between CST2 expression and OS **(A)**, PFI **(B)**, DSS **(C)**, and DFI **(D)**. **(E)** Cox regression analysis in STAD from TGCA and GEO database. **(F)** Survival analysis using Kaplan-Meier (KM) curves was conducted to investigate the expression of CST2 and its relationship with OS, DSS, progression-free interval (PFS), and disease-free interval (DFS) in STAD.

correlations with immunostimulators, including CD48, LTA, CD27, TNFRSF4, IL2RA, TNFRSF9, CD80, and CD86 (Figure 6C). Additionally, CST2 demonstrates significant positive correlations with receptors such as CXCR4, CCR7, CXCR3, XCR1, CCR2, CCR4, CCR8, CXCR5, CCR3, CXCR6, CCR1, and CCR5 (Figure 6D). Moreover, CST2 exhibits significant positive correlations with immune checkpoint genes, including CD274, CTLA4, HAVCR2, LAG3, PDCD1, PDCD1LG2, SIGLEC15, and TIGIT (Figure 6E). Lastly, CST2 shows significant positive correlations with immunoinhibitors like CSF1R, HAVCR2, PDCD1LG2, LAG3, CD96, PDCD1, CTLA4, TIGIT, BTLA, and IL10 (Figure 6F).

The immunotherapeutic potential of CST2

Prognostic analysis within the immunotherapy cohorts reveals that high expression of CST2 improves patient prognosis in the Lauss cohort 2017 (CAR-T) (Figure 7A). Conversely, low expression of CST2 enhances patient prognosis in the Kim cohort 2019 (Anti-PD-1/PD-L1), Nathanson cohort 2017 (Anti-CTLA-4), and IMvigor210 cohort 2018 (Anti-PD-L1) (Figure 7A). Low expression of CST2 enhances prognosis in Anti-PD-1/PD-L1 treatment. Furthermore, drug sensitivity analysis demonstrates a significant association between CST2 and 21 chemotherapy drugs. Examples include Pyrazoloacridine, XL-147, Lificguat, Ethinyl estrdiol, Curcumin, Vincristine, Floxuridine, Fenretinide, Entinostat, RH1, Fluorouracil, (+)-JQ1, Axitinib, Temsirolimus, Batracylin, Lapatinib, AT-13387, Cordycepin, Benzimate, 5-fluoro deoxy uridine, and Triapine (Figure 7B). GDSC database (Figure 7C) and CTRP database (Figure 7D) validate this result.

Correlation analysis of CST2 with mutations, methylation, TMB, and MSI

The predominant mutation type observed is “Mutation,” with the highest mutation frequency of CST2 found in Uterine corpus

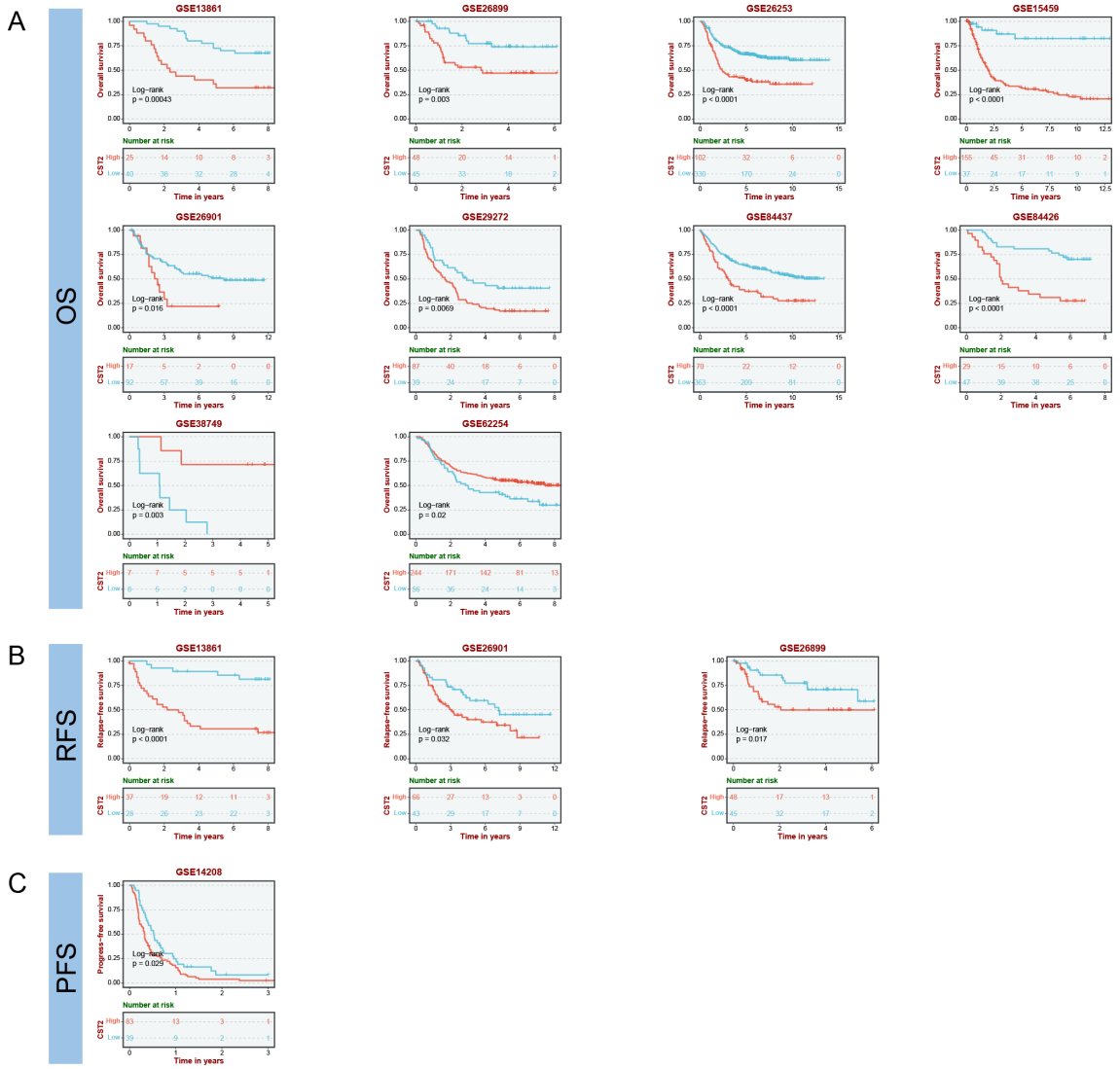


FIGURE 4
Prognostic analysis of CST2 in the immunotherapy cohort, including OS **(A)**, RFS **(B)** and PFS **(C)**.

endometrioid carcinoma (Figure 8A). In pan-cancer analysis, a significant correlation is observed between CST2 and methylation (Figure 8B). The specific mutation sites of CST2 are illustrated in Figure 8C. Furthermore, STAD patients were categorized into two groups based on the median expression level of CST2, and a comparison of gene mutations was conducted between these two groups. The results indicate that patients with low CST2 expression in STAD exhibit a higher frequency of gene mutations compared to those with high CST2 expression (Figure 8D). Additionally, Figure 8E depicts a significant positive correlation between CST2 and TMB in TGCT and THCA. However, in BRCA, LUAD, and STAD, CST2 exhibits a significant negative correlation with TMB. Regarding MSI, CST2 shows a significant positive correlation with MSI in SKCM and TGCT. Conversely, in BRCA, KIRC, LGG, LUAD, LUSC, MESO, PCPG, STAD, and UCEC, CST2 demonstrates a significant negative correlation with MSI (Figure 8F).

The results of GSEA

The GSEA results depicted in Figure 9A demonstrate that CST2 exhibits a significant positive correlation with various biological processes and signaling pathways in the majority of tumors. Specifically, CST2 shows significant positive correlations with myogenesis, KRAS signaling up, interferon gamma response, interferon alpha response, inflammatory response, IL6 JAK STAT3 signaling, IL2 STAT5 signaling, epithelial-mesenchymal transition, complement, coagulation, apical junction, angiogenesis, allograft rejection, and PI3K AKT mtor signaling (Figure 9A). In STAD, GO analysis reveals that CST2 is significantly enriched in External encapsulating structure organization, Collagen fibril organization, Collagen metabolic process (Figure 9B). Furthermore, KEGG analysis demonstrates significant enrichment of CST2 in Hedgehog signaling pathway, Glycosphingolipid biosynthesis globo series, Gap junction, and Wnt signaling pathway in STAD (Figure 9C).

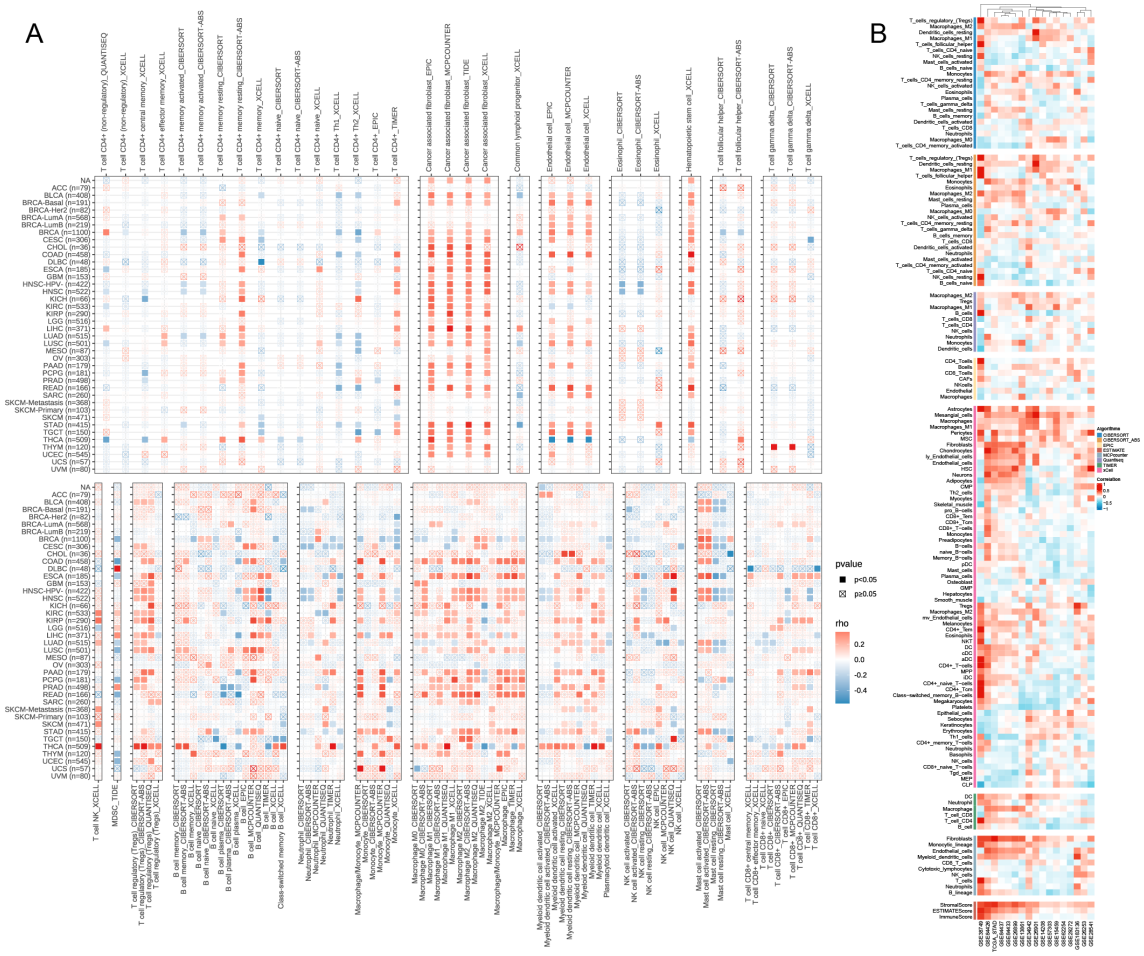


FIGURE 5
The Relationship of CST2 expression with immune cell infiltration analysis. Immunoinfiltration analysis from Timer2.0 database **(A)** and best database **(B)**.

Enrichment analysis of CST2 positively and negatively co-expressed genes

Applying the Spearman algorithm, we identified the top 20 genes that showed positive and negative co-expression with CST2 in STAD (Figure 10A). The heatmap (Figure 10B) depicts the expression levels of these genes. Positive co-expressed genes with CST2 were notably enriched in extracellular matrix organization and extracellular structure organization (Figure 10C). Conversely, negative co-expressed genes with CST2 were significantly enriched in the regulation of mitotic cell cycle phase transition and regulation of cell cycle phase transition (Figure 10D).

Validating the key role of CST2 in STAD development

We selected two STAD cell lines, MKN-45 and SGC-7901, and used gene knockdown techniques to decrease the expression level of CST2. Western blot analysis confirmed a significant reduction in CST2 expression (Figures 11A, B). Furthermore, to comprehensively evaluate the effect of CST2 in STAD, we performed transwell migration assays to investigate its impact on cell migration ability. Under conditions of CST2 silencing, cell migration was significantly reduced, which was confirmed by crystal violet staining (Figure 11C). To further explore the impact of CST2 on cancer cell proliferation, we conducted colony formation assays. The results demonstrated a significant decrease in the proliferation ability of both cell lines following CST2 knockdown (Figures 11D–F), highlighting the crucial role of CST2 in maintaining malignant cell proliferation. Additionally, the levels of apoptosis were significantly increased in both cell lines following CST2 knockdown (Figures 11G, H), further supporting the influence of CST2 on cancer cell survival. These findings underscore the central role of CST2 in STAD pathogenesis, influencing disease progression by regulating cancer cell proliferation and migration.

Discussion

In recent years, mounting evidence has shown the overexpression of CST2 in gastric cancer, colorectal cancer,

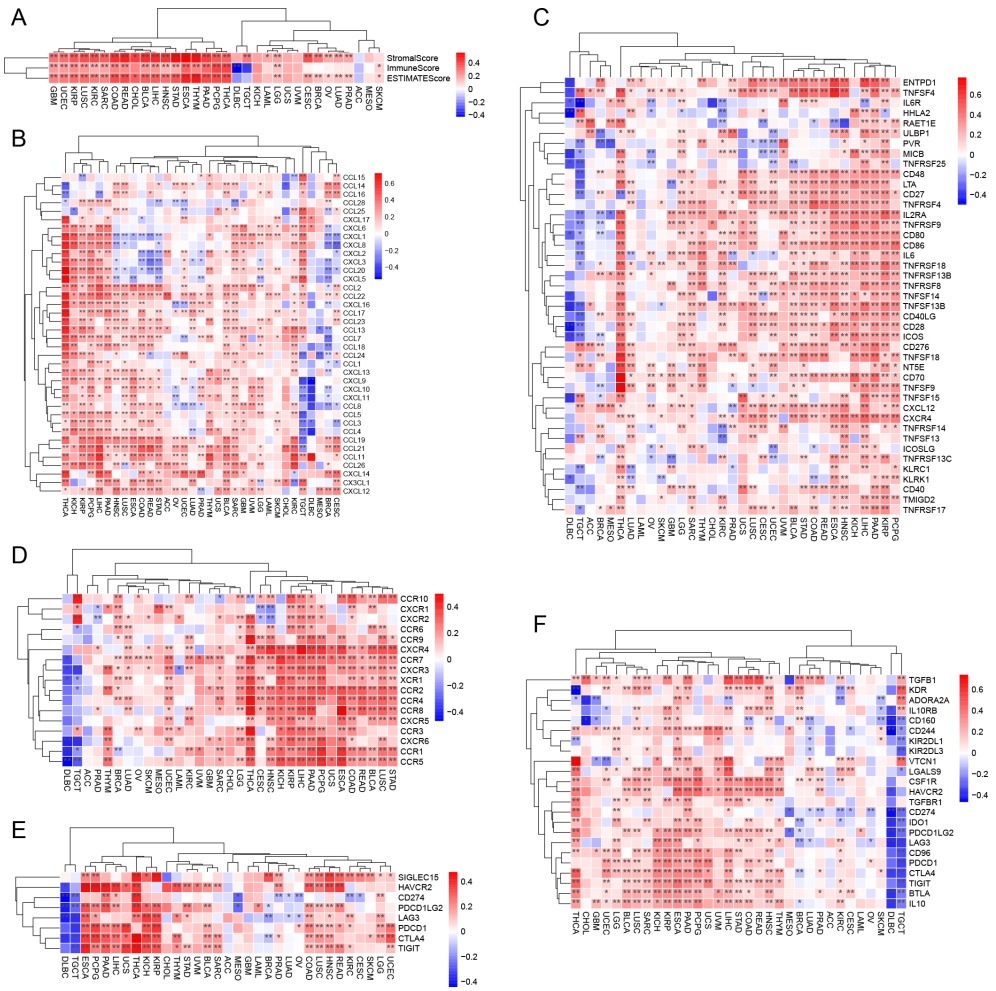


FIGURE 6
CST2 correlates with immune-related genes. **(A)** Correlation of CST2 with three scores including ESTIMATEScore, ImmuneScore and StromalScore. The Relationship of CST2 expression with **(B)** Chemokine, **(C)** Immunostimulator, **(D)** receptor, **(E)** immune checkpoint genes, and **(F)** Immunoinhibitor (*p < 0.05 and **p < 0.01).

prostate cancer, hepatocellular carcinoma, and breast cancer. In these cancers, CST2 functions as an oncogene. However, its role in other malignancies remains largely unknown.

Previous studies have demonstrated that overexpression of CST2 contributes to the progression of pancreatic cancer. Analysis of the TCGA and GTEx databases has confirmed abnormally high expression of CST2 in 15 types of tumors. COX regression analysis demonstrates significant correlations between CST2 and OS in STAD, SKCM, READ, PAAD, LGG, KIRP, KIRC, HNSC, GBM, and COAD. Particularly in STAD and PAAD, CST2 exhibits associations with OS, DSS, DFS, and PFS. Further analysis using Kaplan-Meier survival curves demonstrates the prognostic value of CST2 in STAD, where lower expression is associated with better patient outcomes. CST2 shows potential as a reliable biomarker, supported by COX regression analysis and KM survival curves. Additionally, overexpression of CST2 at both mRNA and protein levels is correlated with poor prognosis in late-stage cancer across multiple types, emphasizing its significance in cancer progression. These findings provide direct evidence of CST2's involvement in cancer initiation and development, solidifying its potential as a target for cancer diagnosis, treatment, and prognosis. However, more research is still needed to confirm the overexpression of CST2 and its oncogenic role in cancer.

Recently, a study has suggested that CST2 may promote the malignant progression of pancreatic cancer through the activation of the PI3K AKT signaling pathway (6). EMT is known to be one of the major factors contributing to tumor cell proliferation, invasion, and metastasis (13). Furthermore, EMT is a crucial factor in the development of drug resistance in cancer treatment (14–16). In gastric cancer, CST2 promotes tumor cell growth, migration, and invasion by modulating EMT and the TGF-β1 signaling pathway (17). Additionally, CST2 may be involved in prostate cancer cell migration through the regulation of the EMT signaling pathway (10). Approximately 1/5 of cancer patients have RAS mutations, which play a significant role in tumorigenesis and progression (18). KRAS is the most common subtype among RAS mutations and is correlated with poor prognosis in cancer patients (19). Interferon gamma levels significantly increase upon stimulation by CST2 (20). In the tumor microenvironment, the IL2 STAT5 signaling pathway can induce CD8+ T cell exhaustion (21).

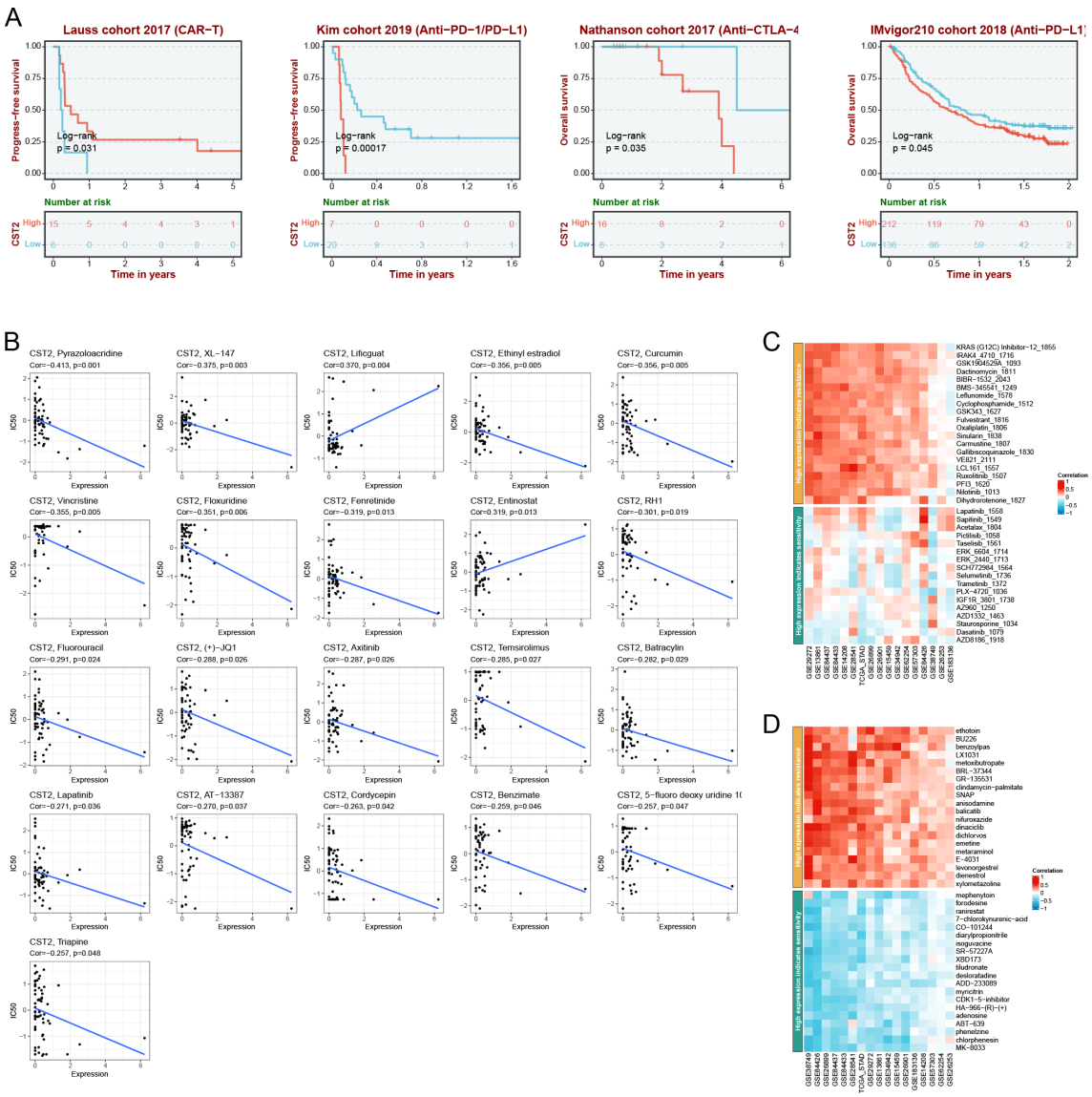


FIGURE 7
Immunotherapy analysis. **(A)** CST2 expression and prognostic analysis of chimeric antigen receptor-modified T (CAR-T), anti-programmed cell death protein 1/programmed cell death ligand 1 (PD-1/PD-L1), anti-cytotoxic T lymphocyte-associated antigen 4 (CTLA-4), and anti-PD-L1 immunotherapy cohorts. Drug sensitivity analysis of CST2 from cellMiner **(B)**, Genomics of Drug Sensitivity in Cancer (GDSC) database **(C)** and Cancer Therapeutics Response Portal (CTRP) database **(D)**.

The tumor microenvironment (TME) is a complex ecosystem consisting of various immune cell types, CAFs, and endothelial cells that surround cancer cells, all embedded within the extracellular matrix (ECM) (22). These cells have been demonstrated to play critical roles in cancer pathogenesis. In our study, we discovered that CST2 is closely associated with multiple immune cells, including CAFs, endothelial cells, hematopoietic stem cells, regulatory T cells (Tregs), B cells, macrophages, monocytes, myeloid dendritic cells, and CD8+ T cells, in pan-cancer. CAFs are an essential cell population in the TME and have been demonstrated to promote tumorigenesis and lead to poorer survival outcomes (23). TAMs are another significant component of the TME, playing roles in coordinating angiogenesis, ECM remodeling, cancer cell proliferation, metastasis, immunosuppression, and resistance to chemotherapy and immune checkpoint blockade therapy (24). CD8 + T cells are potent effector cells that play a crucial role in anti-tumor immune responses, including ICB and adoptive T-cell therapy (25). In cancer, CD4+ T cells exhibit a dual role. Th1 subtype CD4+ T cells contribute to anti-tumor activity by assisting cytotoxic CD8+ T cells and B cells and directly killing cancer cells through interferon and tumor necrosis factor-alpha production. Conversely, Th2 subtype CD4+ T cells secrete anti-inflammatory mediators, promoting tumor growth (26). Tregs, on the other hand, are essential for regulating immune homeostasis and can inhibit effective anti-tumor immunity through various mechanisms (27). Additionally, B cells in cancer exert anti-tumor effects through antibody-dependent cellular cytotoxicity and complement activation (28). These observations highlight the potential role of CST2 in tumor progression by

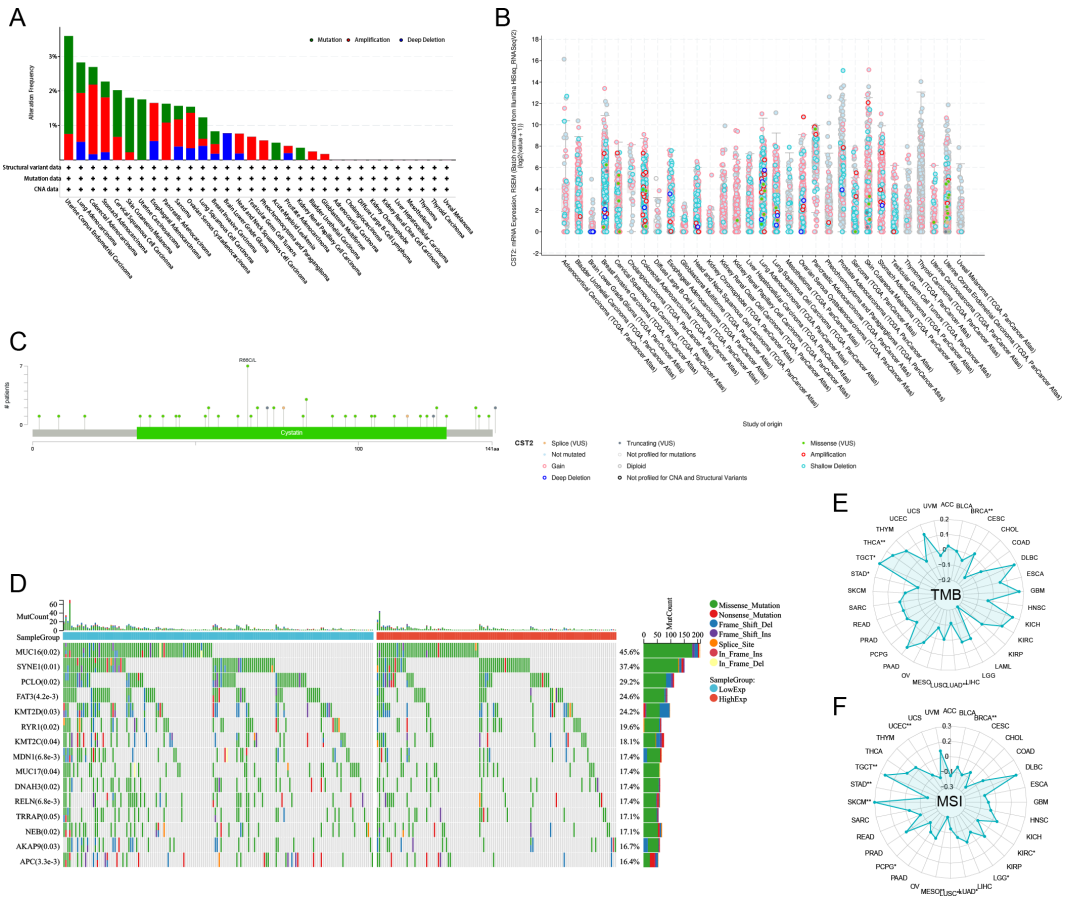


FIGURE 8
The correlation of CST2 expression with mutation, DNA methylation, tumor mutation burden (TMB), and microsatellite instability (MSI). **(A)** The mutation frequency and corresponding mutation types of CST2 in different cancers. **(B)** The correlation between CST2 expression and DNA methylation. **(C)** Mutation sites of CST2. **(D)** The R package "mafTools" was used to calculate the top 15 genes with the highest mutation frequencies in the low-CST2 (left) and high-CST2 (right) groups of STAD, respectively. Radar plots represent the correlation of CST2 expression with TMB **(E)** and MSI **(F)** in pan-cancer (*p < 0.05 and **p < 0.01).

modulating the functions of various immune cells mentioned earlier. Therefore, CST2 represents an attractive therapeutic target in cancer treatment.

Survival analysis demonstrated that high CST2 expression is associated with improved PFS in patients undergoing chimeric antigen receptor-modified T (CAR-T) cell therapy. These findings suggest that patients exhibiting elevated CST2 levels may experience enhanced survival rates following CAR-T treatment. In contrast, low CST2 expression correlates with increased PFS in patients receiving anti-programmed cell death protein 1/programmed cell death ligand 1 (PD-1/PD-L1) therapy, as well as improved OS in those treated with anti-cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) therapy. CAR-T cell therapy represents a groundbreaking approach in adoptive immunotherapy, significantly broadening the horizons for cancer treatment (29). Immune checkpoint proteins, such as PD-1 and PD-L1, are frequently overexpressed in cancer cells and tumor-associated myeloid cells, leading to the suppression of immune surveillance by adaptive immune cells within the TME. Consequently, targeting the PD-1/PD-L1 axis through immune checkpoint blockade (ICB) has emerged as a standard therapeutic strategy for various malignancies (30). Additionally, CTLA-4, another inhibitory immune checkpoint expressed on activated T cells, serves as an effective target for cancer therapy. Agents that inhibit both the PD-1/PD-L1 and CTLA-4 pathways have received approval for the treatment of multiple cancer types. Our research further indicates that among tumor patients receiving CAR-T cell therapy, those with CST2 overexpression are at a lower risk of mortality and demonstrate a better prognosis. Conversely, in patients undergoing anti-PD-1/PD-L1 and anti-CTLA-4 treatments, those with low CST2 expression exhibit more favorable prognoses.

Immunotherapy has become a crucial approach in human anti-tumor treatment, but it has certain limitations (31). Due to the heterogeneity of the TME, immunotherapy is not effective for all cancer patients. The prediction of biomarkers enables more accurate identification of individuals who are likely to benefit from immunotherapy. This precise guidance for treatment helps in determining the appropriate course of action. In tumors, CST2 primarily exhibits mutations at the DNA level. Methylation at the DNA, RNA, and protein levels and their associated downstream signaling pathways participate in various biological processes (32). The use of methylation for cancer diagnosis and treatment is an intriguing research direction. We have observed a significant

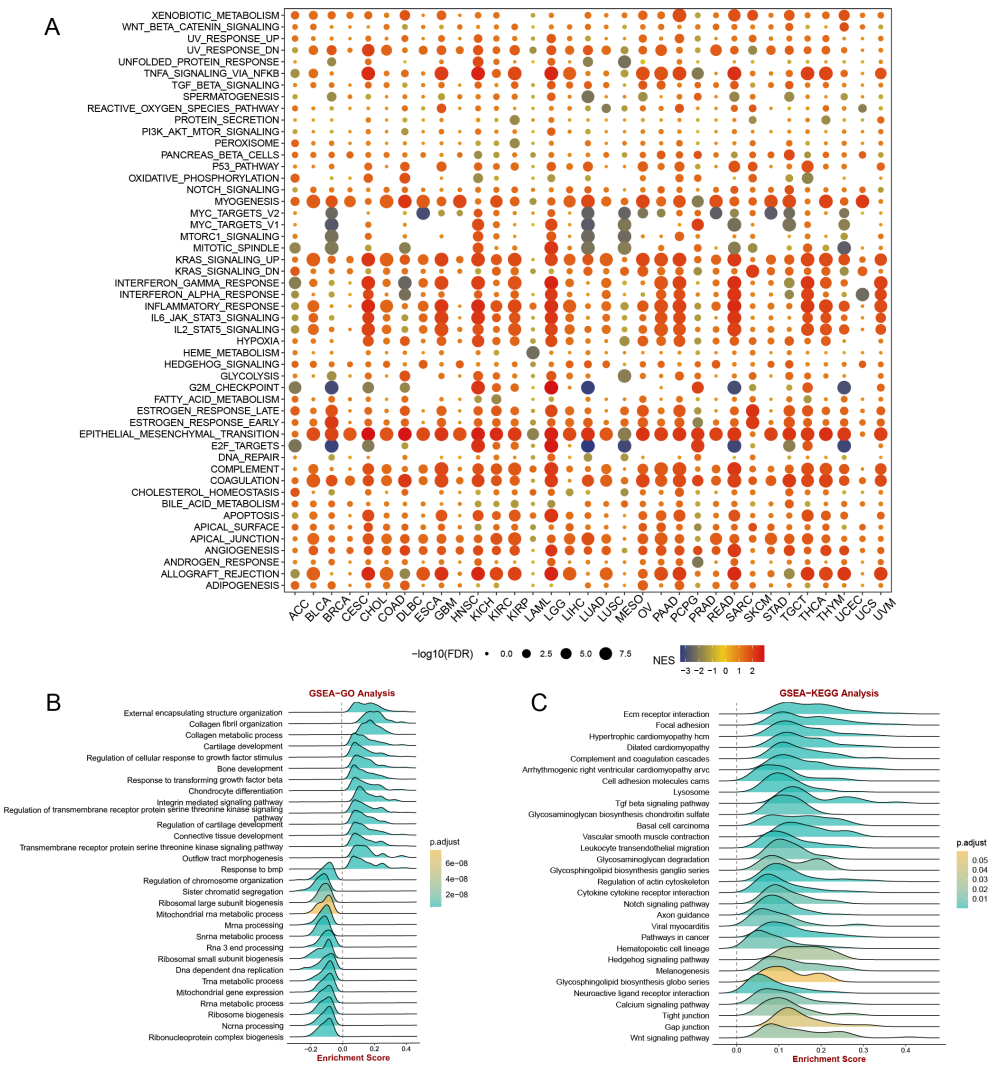


FIGURE 9
(A) gene set enrichment analysis (GSEA) of CST2 in hallmarks gene set. CST2 in STAD for **(B)** Gene Ontology (GO), and **(C)** Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis.

correlation between CST2 and methylation in pan-cancer. Therefore, CST2 holds promise as a diagnostic marker for detecting mutations and epigenetic alterations across various types of cancer. TMB and MSI are two closely related biomarkers that play important roles in tumor diagnosis, treatment, and prognosis assessment (33, 34). TMB has been used as a predictive biomarker for the efficacy of various immunotherapies, particularly PD-1/PD-L1 inhibitors. Currently, in clinical practice, MSI and TMB are detected to determine if a tumor patient is suitable for immunotherapy and to predict the response and outcome of immunotherapy (35). MSI and TMB, as tumor biomarkers, play significant roles in precision medicine, guiding treatment decisions and improving treatment outcomes. Our study found a significant correlation between CST2 and TMB/MSI in various tumors such as TGCT, THCA, BRCA, LUAD, STAD, SKCM, KIRC, LGG, LUSC, MESO, PCPG, STAD, and UCEC. Thus, CST2 can serve as a predictive biomarker for immunotherapy efficacy in these specific cancers.

In addition, we also analyzed the sensitivity of CST2-related drugs. We identified 21 chemotherapy drugs, including Pyrazoloacridine, XL-147, Lificguat, Ethinyl estradiol, Curcumin, Vincristine, Floxuridine, Fenretinide, Entinostat, RH1, Fluorouracil, (+)-JQ1, Axitinib, Temsirolimus, Batracylin, Lapatinib, AT-13387, Cordycepin, Benzimate, 5-fluoro deoxy uridine, and Triapine, that are associated with CST2 expression. CST2 has the potential to serve as a predictive marker for the efficacy of chemotherapy drugs.

In summary, CST2 is upregulated in various tumor types and is associated with unfavorable prognosis in stomach adenocarcinoma. It is linked to gene mutations, methylation patterns, tumor mutational burden (TMB), microsatellite instability (MSI), immune regulatory genes, immune checkpoint genes, immune cell infiltration, and sensitivity to chemotherapy drugs. While potential molecular mechanisms and related signaling pathways of CST2 have been identified, it is important to note that these findings are primarily based on gastric cancer. Further investigation is necessary to establish

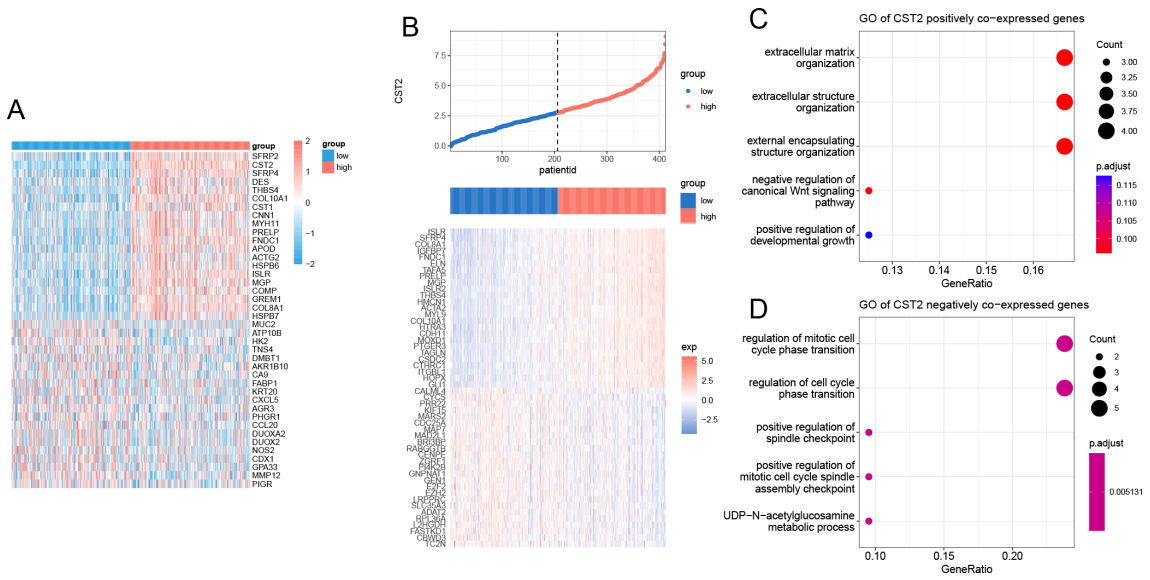


FIGURE 10
Enrichment analysis of CST2. **(A)** positive and negative co-expression with CST2 in STAD. **(B)** Heatmap. Enrichment analysis of positive **(C)** and negative **(D)** co-expression with CST2.

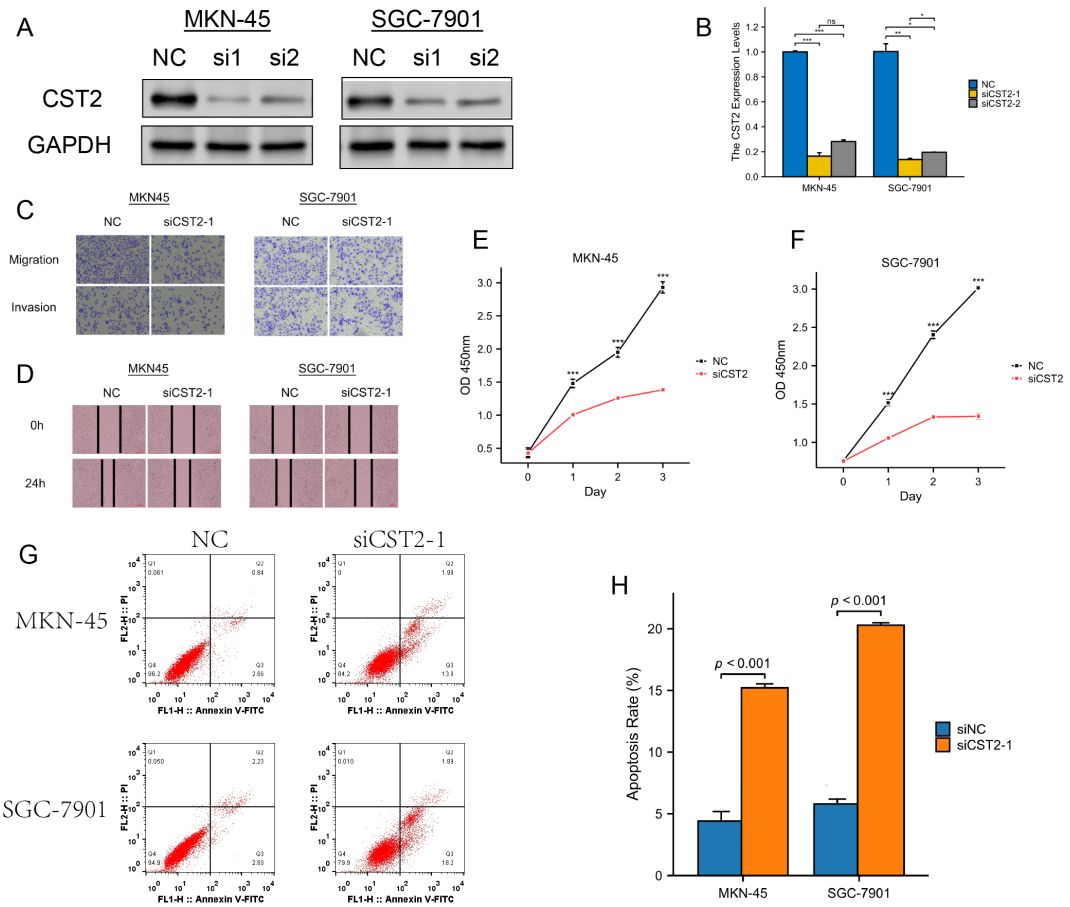


FIGURE 11
Experimental validation analysis of CST2 in STAD. **(A)** WB showing decreased CST2 expression after CST2 knockout. **(B)** Differential expression analysis. **(C)** Cell migration. **(D-F)** Significant impact on cell cloning afterCST2 knockout. **(G, H)** Apoptosis analysis. Increased levels of apoptosis in both cell lines after CST2 knockdown. *p < 0.05, **p < 0.01, ***p < 0.001.

whether CST2 can serve as a novel target for cancer diagnosis, treatment, and prognosis across different cancer types, as well as its potential value in predicting the efficacy of anti-tumor immune responses. These insights contribute to a better understanding of the molecular mechanisms underlying CST2's involvement in tumor initiation and progression, laying the groundwork for future research into targeted therapies and precision medicine.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Materials, further inquiries can be directed to the corresponding authors.

Ethics statement

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

Author contributions

DH: Writing – original draft, Data curation, Methodology, Validation, Visualization. JL: Methodology, Validation, Writing – original draft. ZH: Methodology, Writing – original draft. WL: Methodology, Writing – original draft. LZ: Methodology, Writing – original draft. JH: Validation, Visualization, Methodology, Writing – original draft. YW: Writing – review & editing, Methodology, Supervision. SZ: Writing – review & editing, Funding acquisition, Methodology, Resources, Supervision.

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 Nature and Science Foundation of Guangxi (2024GXNSFAA999352).

Acknowledgments

We thank the public database TCGA for its assistance in our 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.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1466806/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Analysis of single-cell expression distribution patterns of CST2 in CAC, NNSC, NPC, CCRCC, ESCC and UCEC.

References

1. Zhao S, Chi H, Yang Q, Chen S, Wu C, Lai G, et al. Identification and validation of neurotrophic factor-related gene signatures in glioblastoma and Parkinson's disease. *Front Immunol.* (2023) 14:1090040. doi: 10.3389/fimmu.2023.1090040

2. 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

3. Brzin J, Popovic T, Turk V, Borchart U, Machleidt W. Human cystatin, a new protein inhibitor of cysteine proteinases. *Biochem Biophys Res Commun.* (1984) 118:103–9. doi: 10.1016/0006-291X(84)91073-8

4. Mueller SK, Wendler O, Nocera A, Grundtner P, Schlegel P, Agaimy A, et al. Escalation in mucus cystatin 2, pappalysin-A, and periostin levels over time predict need for recurrent surgery in chronic rhinosinusitis with nasal polyps. *Int Forum Allergy rhinology.* (2019) 9:1212–9. doi: 10.1002/alr.22407

5. Techatanawat S, Surarit R, Chairatvit K, Roytrakul S, Khovidhunkit W, Thanakun S, et al. Salivary and serum cystatin SA levels in patients with type 2 diabetes mellitus or diabetic nephropathy. *Arch Oral Biol.* (2019) 104:67–75. doi: 10.1016/j.archoralbio.2019.05.020

6. Ou R, Lin C, Chen Y. CST2 is activated by RUNX1 and promotes pancreatic cancer progression by activating PI3K/AKT pathway. *Arch Biochem biophysics.* (2023) 747:109760. doi: 10.1016/j.abb.2023.109760

7. Xie Q, Liu L, Chen X, Cheng Y, Li J, Zhang X, et al. Identification of cysteine protease inhibitor CST2 as a potential biomarker for colorectal cancer. *J Cancer.* (2021) 12:5144–52. doi: 10.7150/jca.53983

8. Bao Y, Wang L, Shi L, Yun F, Liu X, Chen Y, et al. Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer. *Cell Mol Biol letters.* (2019) 24:38. doi: 10.1186/s11658-019-0162-0

9. Zhang WP, Wang Y, Tan D, Xing CG. Cystatin 2 leads to a worse prognosis in patients with gastric cancer. *J Biol regulators homeostatic agents*. (2020) 34:2059–67. doi: 10.23812/20-293-A

10. Song F, Zhang Y, Pan Z, Hu X, Yi Y, Zheng X, et al. Identification of novel key genes associated with the metastasis of prostate cancer based on bioinformatics prediction and validation. *Cancer Cell Int*. (2021) 21:559. doi: 10.1186/s12935-021-02258-3

11. Song ZB, Yu Y, Zhang GP, Li SQ. Genomic instability of mutation-derived gene prognostic signatures for hepatocellular carcinoma. *Front Cell Dev Biol*. (2021) 9:728574. doi: 10.3389/fcell.2021.728574

12. Blanco MA, LeRoy G, Khan Z, Alečković M, Zee BM, Garcia BA, et al. Global secretome analysis identifies novel mediators of bone metastasis. *Cell Res*. (2012) 22:1339–55. doi: 10.1038/cr.2012.89

13. Zhang Y, Weinberg RA. Epithelial-to-mesenchymal transition in cancer: complexity and opportunities. *Front Med*. (2018) 12:361–73. doi: 10.1007/s11684-018-0656-6

14. Brabletz T, Kalluri R, Nieto MA, Weinberg RA. EMT in cancer. *Nat Rev Cancer*. (2018) 18:128–34. doi: 10.1038/nrc.2017.118

15. Debaugnies M, Rodríguez-Acebes S, Blondeau J, Parent MA, Zocco M, Song Y, et al. RHOJ controls EMT-associated resistance to chemotherapy. *Nature*. (2023) 616:168–75. doi: 10.1038/s41586-023-05838-7

16. Shibue T, Weinberg RA. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat Rev Clin Oncol*. (2017) 14:611–29. doi: 10.1038/nrclinonc.2017.44

17. Xie KQ, Zhang LM, Cao Y, Zhu J, Feng LY. Adenosine A(1) receptor-mediated transactivation of the EGF receptor produces a neuroprotective effect on cortical neurons *in vitro*. *Acta pharmacologica Sin*. (2009) 30:889–98. doi: 10.1038/aps.2009.80

18. Prior IA, Hood FE, Hartley JL. The frequency of ras mutations in cancer. *Cancer Res*. (2020) 80:2969–74. doi: 10.1158/0008-5472.CAN-19-3682

19. Guo TA, Wu YC, Tan C, Jin YT, Sheng WQ, Cai SJ, et al. Clinicopathologic features and prognostic value of KRAS, NRAS and BRAF mutations and DNA mismatch repair status: A single-center retrospective study of 1,834 Chinese patients with Stage I-IV colorectal cancer. *Int J cancer*. (2019) 145:1625–34. doi: 10.1002/ijc.v145.6

20. Kato T, Ito T, Imatani T, Minaguchi K, Saitoh E, Okuda K. Cystatin SA, a cysteine proteinase inhibitor, induces interferon-gamma expression in CD4-positive T cells. *Biol Chem*. (2004) 385:419–22. doi: 10.1515/BC.2004.047

21. Liu Y, Zhou N, Zhou L, Wang J, Zhou Y, Zhang T, et al. IL-2 regulates tumor-reactive CD8(+) T cell exhaustion by activating the aryl hydrocarbon receptor. *Nat Immunol*. (2021) 22:358–69. doi: 10.1038/s41590-020-00850-9

22. de Visser KE, Joyce JA. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. *Cancer Cell*. (2023) 41:374–403. doi: 10.1016/j.ccell.2023.02.016

23. Chen Y, McAndrews KM, Kalluri R. Clinical and therapeutic relevance of cancer-associated fibroblasts. *Nat Rev Clin Oncol*. (2021) 18:792–804. doi: 10.1038/s41571-021-00546-5

24. Mantovani A, Allavena P, Marchesi F, Garlanda C. Macrophages as tools and targets in cancer therapy. *Nat Rev Drug discovery*. (2022) 21:799–820. doi: 10.1038/s41573-022-00520-5

25. Philip M, Schietinger A. CD8(+) T cell differentiation and dysfunction in cancer. *Nat Rev Immunol*. (2022) 22:209–23. doi: 10.1038/s41577-021-00574-3

26. Borst J, Ahrends T, Bąbała N, Melief CJM, Kastenmüller W. CD4(+) T cell help in cancer immunology and immunotherapy. *Nat Rev Immunol*. (2018) 18:635–47. doi: 10.1038/s41577-018-0044-0

27. Togashi Y, Shitara K, Nishikawa H. Regulatory T cells in cancer immunosuppression - implications for anticancer therapy. *Nat Rev Clin Oncol*. (2019) 16:356–71. doi: 10.1038/s41571-019-0175-7

28. Laumont CM, Banville AC, Gilardi M, Hollern DP, Nelson BH. Tumour-infiltrating B cells: immunological mechanisms, clinical impact and therapeutic opportunities. *Nat Rev Cancer*. (2022) 22:414–30. doi: 10.1038/s41568-022-00466-1

29. Sterner RC, Sterner RM. CAR-T cell therapy: current limitations and potential strategies. *Blood Cancer J*. (2021) 11:69. doi: 10.1038/s41408-021-00459-7

30. Zhang H, Dai Z, Wu W, Wang Z, Zhang N, Zhang L, et al. Regulatory mechanisms of immune checkpoints PD-L1 and CTLA-4 in cancer. *J Exp Clin Cancer research: CR*. (2021) 40:184. doi: 10.1186/s13046-021-01987-7

31. Zhang Y, Zhang Z. The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. *Cell Mol Immunol*. (2020) 17:807–21. doi: 10.1038/s41423-020-0488-6

32. Dai X, Ren T, Zhang Y, Nan N. Methylation multiplicity and its clinical values in cancer. *Expert Rev Mol Med*. (2021) 23:e2. doi: 10.1017/erm.2021.4

33. Jardim DL, Goodman A, de Melo Gagliato D, Kurzrock R. The challenges of tumor mutational burden as an immunotherapy biomarker. *Cancer Cell*. (2021) 39:154–73. doi: 10.1016/j.ccell.2020.10.001

34. Yamamoto H, Watanabe Y, Maehata T, Imai K, Itoh F. Microsatellite instability in cancer: a novel landscape for diagnostic and therapeutic approach. *Arch toxicology*. (2020) 94:3349–57. doi: 10.1007/s00204-020-02833-z

35. Palmeri M, Mehnert J, Silk AW, Jabbour SK, Ganesan S, Popli P, et al. Real-world application of tumor mutational burden-high (TMB-high) and microsatellite instability (MSI) confirms their utility as immunotherapy biomarkers. *ESMO Open*. (2022) 7:100336. doi: 10.1016/j.esmoop.2021.100336

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