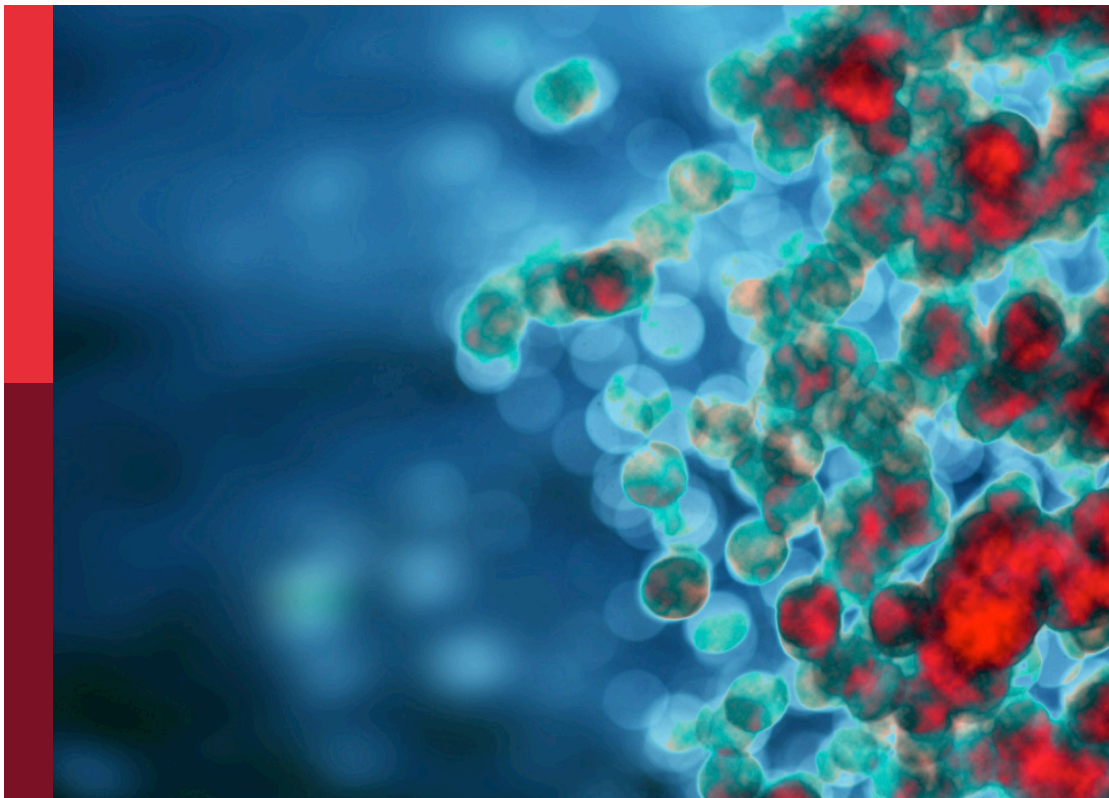
frontiers | Research Topics

Community series in unveiling the tumor microenvironment by machine learning to develop new immunotherapeutic strategies, volume II

Edited by
Ping Zheng, Meng Zhou, Jun Liu
and Nan Zhang

Published in
Frontiers in Immunology
Frontiers in Oncology



ISSN 1664-8714
ISBN 978-2-8325-4123-4
DOI 10.3389/978-2-8325-4123-4

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

Community series in unveiling the tumor microenvironment by machine learning to develop new immunotherapeutic strategies, volume II

Topic editors

Ping Zheng — The University of Melbourne, Australia

Meng Zhou — Wenzhou Medical University, China

Jun Liu — Dongguan Hospital of Guangzhou University of Chinese Medicine, China

Nan Zhang — Harbin Medical University, China

Citation

Zheng, P., Zhou, M., Liu, J., Zhang, N., eds. (2024). *Community series in unveiling the tumor microenvironment by machine learning to develop new immunotherapeutic strategies, volume II*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-4123-4

Table of contents

frontiers | Frontiers in Immunology

TYPE Editorial
PUBLISHED 24 January 2024
DOI 10.3389/fimmu.2024.1351597

OPEN ACCESS

EDITED AND REVIEWED BY
Fuhai Li,
Washington University in St. Louis,
United States

*CORRESPONDENCE
Ping Zheng
jojo_ras@126.com

RECEIVED 06 December 2023
ACCEPTED 11 January 2024
PUBLISHED 24 January 2024

CITATION
Zheng P and Liu J (2024) Editorial: Community series in unveiling the tumor microenvironment by machine learning to develop new immunotherapeutic strategies, volume II. *Front. Immunol.* 15:1351597. doi: 10.3389/fimmu.2024.1351597

Editorial: Community series in unveiling the tumor microenvironment by machine learning to develop new immunotherapeutic strategies, volume II

Ping Zheng[1]* and Jun Liu[2]

[1]Department of Neurosurgery, Pudong New Area People's Hospital, Shanghai, China, [2]Department of Medicine, Yuebei People's Hospital, Shaoguan, China

KEYWORDS

tumor microenvironment, machine learning, tumor, disease, immune infiltrations

Editorial on the Research Topic
Community series in unveiling the tumor microenvironment by machine learning to develop new immunotherapeutic strategies, volume II

A total of 25 papers are included in this series. We selected seven as representative:

Yang et al. discussed the role of lipid metabolism in gastric cancer. Their study identified 78 genes related to fatty acid metabolism that are differentially expressed between normal and gastric cancer tissues. The ConsensusClusterPlus R package was used to divide the genes into two gastric cancer subtypes, cluster 1 and cluster 2. Patients in cluster 2 were found to have a poorer prognosis than those in cluster 1. The study used machine learning to select 8 differentially expressed genes between the subtypes to construct a fatty acid prognostic risk score (FARS) model, which displayed good prognostic efficacy. Certain anticancer drugs, such as bortezomib, elesclomol, GW843682X, and nilotinib, showed significant sensitivity in the high FARS score group. RGS2 was identified as the core gene in gastric cancer single-cell analysis, and Western blotting and immunofluorescence staining results revealed high levels of expression of this gene in gastric cancer cells. The results of immunohistochemical staining showed that a large amount of RGS2 was deposited in the stroma in gastric cancer. The pan-cancer analysis also revealed a significant association of RGS2 with TMB, TIDE, and CD8+ T-cell infiltration in other cancer types as well. RGS2 may thus be further studied as a new target for immunotherapy in future studies on gastric cancer. The FARS model developed here enhances our understanding of lipid metabolism in the TME in gastric cancer, and provides a theoretical basis for predicting tumor prognosis and clinical treatment.

Peng et al. investigated the development of a prognostic model based on oxidative stress for lung adenocarcinoma (LUAD). The study extracted oxidative stress-related genes (ORGs) from Genecards and performed machine learning algorithms to build the OS-score and OS-signature. The study identified ten ORGs with prognostic value and the OS-signature containing three prognostic ORGs. The efficiency and accuracy of the OS-signature in predicting the prognosis for LUAD patients was confirmed by survival ROC

curves and two external validation data sets. Patients with high OS-scores were found to have lower levels of immunomodulators, stromal score, immune score, ESTIMATE score, and infiltrating immune cell populations. Conversely, patients with higher OS-scores were more likely to have higher tumor purity. PCR assays showed that MRPL44 and CYCS were significantly upregulated in LUAD cell lines, while CAT was significantly downregulated.

Wang et al. discussed the role of MYBL1 in clear cell renal cell carcinoma (ccRCC). The study comprehensively investigated the role of MYBL1 in ccRCC and found that MYBL1 was correlated with progressive clinical characteristics and worse prognosis performance. The study also found that MYBL1 can activate multiple oncogenic pathways in ccRCC and can remodel the immune microenvironment of ccRCC and affect the immunotherapy response. *In vitro* and *in vivo* assays indicated that MYBL1 is upregulated in ccRCC cells and can promote the cellular malignant behavior of ccRCC. Finally, a machine learning algorithm - LASSO logistic regression was utilized to identify a prognostic signature based on the MYBL1-derived molecules, which showed satisfactory ability to predict patient prognosis in both training and validation cohorts. The study concluded that MYBL1 is a novel biomarker of ccRCC that can remodel the tumor microenvironment, influence immunotherapy responses, and guide precision medicine in ccRCC.

Ke et al. discussed the potential and significance of immune-related diagnostic biomarkers in differentiating Uterine leiomyosarcoma (ULMS) from Uterine leiomyoma (ULM). The study downloaded two public gene expression profiles containing ULMS and ULM samples and identified differentially expressed genes (DEGs) among 37 ULMS and 25 ULM control samples. The DEGs were used for Gene Ontology (GO), Kyoto Encyclopaedia of Genes and Genomes (KEGG), and Disease Ontology (DO) enrichment analysis in addition to gene set enrichment analysis (GSEA). The study identified DPP6 and MFAP5 as diagnostic biomarkers for ULMS, which were verified in the GSE9511 and GSE68295 datasets. Low expression of DPP6 and MFAP5 was associated with ULMS. The study concluded that DPP6 and MFAP5 are potential diagnostic biomarkers for ULMS.

Zhang et al. investigated the development of a mitochondria-related signature in osteosarcoma patients. Transcriptomic data and clinical information of osteosarcoma samples were collected from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) and Gene Expression Omnibus (GEO) databases. Comprehensive bioinformatics analysis was performed on the samples at the bulk RNA sequencing level and single-cell RNA sequencing (scRNA-seq). The study constructed a mitochondria-related signature in osteosarcoma patients and explored its prognostic value, predictive value in the immune microenvironment and chemotherapeutic agents. The study also investigated the association between mitochondria and immunity in the tumor microenvironment of osteosarcoma at the scRNA-seq level. The tumorigenic role of the critical mitochondria-related gene, PCCB, was verified by *in vitro* validation. The study concluded that a mitochondria-related signature was developed in osteosarcoma with solid predictive values for the immune microenvironment, chemotherapeutic agents, and prognosis.

Xu et al. discussed the identification of a glycolysis and cholesterol synthesis-related genes (GCSRGs) signature for effective prognostic assessment of osteosarcoma patients. Gene expression data and clinical information were obtained from the GSE21257 and TARGET-OS datasets. Patients diagnosed with osteosarcoma were classified into one of 4 subtypes (quiescent, glycolysis, cholesterol, and mixed subtypes), which differed significantly in terms of prognosis and tumor microenvironment. Both univariate and LASSO Cox regression analyses were conducted on the screened module genes to identify 5 GCSRGs (RPS28, MCAM, EN1, TRAM2, and VEGFA) that constituted a prognostic signature for osteosarcoma patients. The signature was an effective prognostic predictor, independent of clinical characteristics, as further verified via Kaplan-Meier analysis, ROC curve analysis, and univariate and multivariate Cox regression analysis. Additionally, the GCSRG signature had a strong correlation with drug sensitivity, immune checkpoints and immune cell infiltration.

Cholangiocarcinoma (CHOL) is a prevalent type of malignancy and the second most common form of primary liver cancer, resulting in high rates of morbidity and mortality. Necroptosis is a type of regulated cell death that appears to be involved in the regulation of several aspects of cancer biology, including tumorigenesis, metastasis, and cancer immunity. Xu et al. aimed to construct a necroptosis-related gene (NRG) signature to investigate the prognosis of CHOL patients using an integrated bioinformatics analysis. CHOL patient data were obtained from the GEO (GSE89748, GSE107943) and TCGA databases, and NRG data from necroptosis were obtained from the KEGG database. A total of 65 differentially expressed (DE) NRGs were screened, of which five were selected to establish the prognostic signature of NRGs based on multivariate Cox regression analysis. Low-risk patients survived significantly longer than high-risk patients. Patients with high-risk scores experienced higher immune cell infiltration, drug resistance, and more somatic mutations than patients with low-risk scores. Sensitivities to GW843682X, mitomycin C, rapamycin, and S-trityl-L-cysteine were significantly higher in the low-risk group than in the high-risk group. Finally, the expression of five NRGs was validated.

All of the above studies select a specific geneset and identify prognosis-related or stage-related genes using machine learning methods, some of them further associate these genes with immune status and finally validate them with *in vitro* or *in vivo* experimental methods.

Author contributions

PZ: Conceptualization, Data curation, Writing – original draft, Writing – review & editing. JL: Data curation, Investigation, Methodology, 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.

Publisher's note

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 30 November 2022
DOI 10.3389/fonc.2022.1084192

OPEN ACCESS

EDITED BY
Nan Zhang,
Harbin Medical University, China

REVIEWED BY
Pan Li,
Southern Medical University, China
Xia Chen,
Southern Medical University, China

*CORRESPONDENCE
Zhuna Wu
wuzhuna@aliyun.com

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

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Oncology

RECEIVED 30 October 2022
ACCEPTED 16 November 2022
PUBLISHED 30 November 2022

CITATION
Ke Y, You L, Xu Y, Wu D, Lin Q and Wu Z (2022) DPP6 and MFAP5 are associated with immune infiltration as diagnostic biomarkers in distinguishing uterine leiomyosarcoma from leiomyoma.
Front. Oncol. 12:1084192.
doi: 10.3389/fonc.2022.1084192

DPP6 and MFAP5 are associated with immune infiltration as diagnostic biomarkers in distinguishing uterine leiomyosarcoma from leiomyoma

Yumin Ke[1†], LiuXia You[2†], YanJuan Xu[3], Dandan Wu[4], Qiuya Lin[1] and Zhuna Wu[1*]

[1]Department of Gynecology and Obstetrics, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian, China, [2]Department of Clinical Laboratory, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian, China, [3]Department of Pathology, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian, China, [4]Department of Gynecology, The First Hospital of Quanzhou Affiliated to Fujian Medical University, Quanzhou, Fujian, China

Objective: Uterine leiomyosarcoma (ULMS) is the most common subtype of uterine sarcoma and is difficult to discern from uterine leiomyoma (ULM) preoperatively. The aim of the study was to determine the potential and significance of immune-related diagnostic biomarkers in distinguishing ULMS from ULM.

Methods: Two public gene expression profiles (GSE36610 and GSE64763) from the GEO datasets containing ULMS and ULM samples were downloaded. Differentially expressed genes (DEGs) were selected and determined among 37 ULMS and 25 ULM control samples. The DEGs were used for Gene Ontology (GO), Kyoto Encyclopaedia of Genes and Genomes (KEGG) and Disease Ontology (DO) enrichment analyses as well as gene set enrichment analysis (GSEA). The candidate biomarkers were identified by least absolute shrinkage and selection operator (LASSO) and support vector machine recursive feature elimination (SVM-RFE) analyses. The receiver operating characteristic curve (ROC) was applied to evaluate diagnostic ability. For further confirmation, the biomarker expression levels and diagnostic value in ULMS were verified in the GSE9511 and GSE68295 datasets (12 ULMS and 10 ULM), and validated by immunohistochemistry (IHC). The CIBERSORT algorithm was used to calculate the compositional patterns of 22 types of immune cells in ULMS.

Result: In total, 55 DEGs were recognized *via* GO analysis, and KEGG analyses revealed that the DEGs were enriched in nuclear division, and cell cycle. The recognized DEGs were primarily implicated in non–small cell lung carcinoma and breast carcinoma. Gene sets related to the cell cycle and DNA replication were activated in ULMS. *DPP6* and *MFAP5* were distinguished as diagnostic

biomarkers of ULMS (AUC = 0.957, AUC = 0.899, respectively), and they were verified in the GSE9511 and GSE68295 datasets (AUC = 0.983, AUC = 0.942, respectively). The low expression of *DPP6* and *MFAP5* were associated with ULMS. In addition, the analysis of the immune microenvironment indicated that resting mast cells were positively correlated with *DPP6* and *MFAP5* expression and that eosinophils and M0 macrophages were negatively correlated with *DPP6* expression (P<0.05).

Conclusion: These findings indicated that *DPP6* and *MFAP5* are diagnostic biomarkers of ULMS, thereby offering a novel perspective for future studies on the occurrence, function and molecular mechanisms of ULMS.

KEYWORDS

diagnostic biomarkers, machine-learning, DPP6, MFAP5, immune infiltration, uterine leiomyosarcoma

Introduction

Uterine leiomyosarcoma (ULMS) is a rare but aggressive tumor subtype, accounting for approximately 1% of all uterine malignancies (1). ULMS is the most common subtype of uterine sarcoma and originates from the smooth muscles of the myometrium. In the past several decades, the prognosis of ULMS patients has not changed with an overall 5-year survival rate of only 15%-25% (2). Currently, complete surgical resection is the primary treatment for early-stage ULMS (3), and chemotherapy is regarded as the standard therapy for advanced or metastatic ULMS (4, 5), but with an estimated recurrence rate of approximately 50 to 70% (6). ULMS constitutes a sizable proportion of uterine cancer deaths (7). Additionally, compared to other gynecological malignancies, ULMS etiology, pathogenesis and earlier diagnosis are poorly understood. Considering that ULM can currently be treated with minimally invasive surgery, it is important to discern ULMS from ULM preoperatively to avoid disseminated spread by laparoscopic morcellation or delayed diagnosis with conservative treatment (8, 9). Considering that ULMS has a high trend towards local recurrence, metastasis and poor prognosis, the misdiagnosis of a ULMS for a leiomyoma may lead to therapy delays and higher morbidity (10, 11).

ULMS patients generally present with abnormal vaginal bleeding, pelvic pain and palpable pelvic mass. Because these symptoms resemble ULM, particularly degenerated ULM, it is difficult to discern ULMS and ULM by pelvic ultrasound and MRI preoperatively (12). Postoperative pathological diagnosis is currently the only available method to distinguish the two tumor conditions. A meta-analysis containing 133 studies has indicated that undiagnosed ULMS estimated to be approximately 1 in 2000 surgeries for presumed ULM (13). It is well known that tumor-associated immunity plays a vital role in the occurrence, development and metastasis of tumors (14). The recent development of integrated microarray technology with bioinformatics analysis may allow identification of novel genes that might act as diagnostic and prognostic biomarkers in cancers (15, 16). Definitive molecular diagnosis added to histopathological diagnosis should be considered to decrease the risk of misdiagnosis. Verification of highly novel diagnostic biomarkers for ULMS related to immune cell infiltration will further improve the diagnostic accuracy of ULMS.

Herein, the aim of this study was to identify novel diagnostic immune-related genes for ULMS. Machine-learning algorithms and logistic regression were used to verify diagnostic biomarkers of ULMS. Furthermore, CIBERSORT was applied to compute the quotas of infiltrating immune cells between ULMS and ULM samples. Finally, the correlation among the recognized diagnostic biomarkers and infiltrating immune cells was explored to offer a foundation for further research.

Materials and methods

Microarray data processing and identification of DEGs

First, we obtained datasets (GSE36610 and GSE64763 as the training group; and GSE9511 and GSE68295 as the testing group) from the GEO database (https://www.ncbi.nlm.nih.gov/gds) (Table 1). The background correction and normalization of raw data were processed by the limma package of R software (http://www.bioconductor.org/). Two datasets were merged into

TABLE 1 GEO database data of preeclampsia mRNA expression profile.

Dataset ID	Platform	leiomyosarcoma	leiomyoma
Train group			
GSE36610	GPL7363-11635	12	0
GSE64763	GPL571-17391	25	25
Test group			
GSE9511	GPL80-30376	9	7
GSE68295	GPL6480-9577	3	3

a metadata cohort, and the batch effect was removed with the SVA package of R software (17). Genes with |log fold change (FC)| > 2 and adjusted P < 0.05 were defined as DEGs.

Functional enrichment of DEGs

The DEGs were analyzed using the clusterProfiler, org.Hs.eg.db, enrichplot and ggplot2 packages of R software for GO and KEGG analyses. The clusterProfiler and DOSE packages of R software were used to perform DO analyses on DEGs. GSEA was conducted to recognize the most important feature between the ULMS and ULM groups. "c2.cp.kegg.v7.4.symbols.gmt" was applied as the reference gene set from the Molecular Signatures Database (MSigDB). P <0.05 was considered a significant enrichment.

Screening candidate biomarker for diagnosis of ULMS

We applied two machine-learning algorithms to increase the prediction accuracy. LASSO is a regression-based analysis that scrutinizes variable selection and regularization in ULMS models. The glmnet package of R software was applied to perform LASSO regression analysis on the identification of DEGs correlated with the discernment between ULMS and ULM. The support vector machine (SVM) is an efficient and widely applied supervised machine-learning algorithm for disease classification and regression tasks (18). Consequently, we screened the overlapping genes by conjugating LASSO and SVM-RFE followed by verification using the GSE9511 and GSE68295 datasets.

Significance of diagnostic biomarkers in ULMS

We obtained mRNA expression data from 37 ULMS and 25 ULM samples, which were applied to create ROC curves to verify the biomarker predictive ability. The area under the ROC curve (AUC) was utilized to determine the ability of diagnosis in distinguishing ULMS from ULM samples followed by verification using the GSE9511 and GSE68295 datasets.

Evaluating the level of immune infiltration

We downloaded a gene signature matrix with interpretation, known as the 22 immune cell (LM22) matrix with 1,000 permutations from CIBERSORT (http://cibersort.stanford.edu/) (19). The CIBERSORT algorithm was applied to quantify the proportion of 22 infiltrating immune cells in the tissue using the expression of 547 immune-related genes. The corrplot package in R software was applied to conduct the correlation and visualization of 22 types of infiltrating immune cells. The vioplot package in R software was used to study the infiltration of immune cells between the ULMS and ULM groups. Pearson correlation analysis was applied to explore the selected diagnostic biomarkers correlated with the levels of infiltrating immune cells.

Patient and tissue samples

Twenty-six paraffin-embedded ULMS and twenty-three ULM specimens were diagnosed at The Second Affiliated Hospital of Fujian Medical University (Fujian, China) from September 2010 to February 2022. The main treatment of all patients underwent hysterectomy with bilateral adnexal resection. The research was approved by the Research Ethics Committee of The Second Affiliated Hospital of Fujian Medical University prior to the study.

Immunohistochemistry

IHC staining was operated as previously described (20). The primary antibodies included anti-DPP6 (Bioss, Beijing), anti-MFAP5 (Proteintech, USA). The proportion of DPP6 and MFAP5 staining intensity was scored as follows: negative = 0; light yellow = 1; brownish yellow = 2; or tan=3. The staining was scored as follows: less than 1/3 = 1; between 1/3 and 2/3 = 2; or more than 2/3 = 3. The final score for DPP6 and MFAP5 expression was calculated by multiplying the 2 scores. The slides were classified to low and high expression group, corresponding to scores of <3 or ≥3, respectively. The histopathological diagnosis of the patients included in our study was established by two pathologists specialized in Gynecologic Oncology.

Statistical analysis

We utilized R software to conduct (v.4.1.1) all statistical analyses. We used the Mann–Whitney U test to compare the different groups. LASSO regression, SVM algorithm, ROC curve, Pearson's correlation and unpaired t test were used as described above. Differences with $P < 0.05$ were considered statistically significant for all statistical analyses.

Result

Study procedure

The analysis procedure of the present is shown in Figure 1. The transcriptome RNA-seq data were downloaded from the GEO database. We identified the DEGs between the ULMS and ULM group. DEGs were anlayzed using the GO, KEGG and DO analyses as well as GSEA. LASSO and SVM-RFE were used to select the candidate overlapping genes, and ROC curves were applied to check the predictive ability of biomarkers, which was further verified in the GSE9511 and GSE68295 datasets. The compositional patterns of 22 immune cells were calculated using CIBERSORT in ULMS. Finally, correlation analysis among the diagnostic markers and infiltrating immune cells was performed.

Identification of DEGs in ULMS

The present study utilized two datasets (GSE36610, GSE64763) and included 37 ULMS and 25 ULM samples. We identified 55 DEGs by comparing ULMS and ULM (Figure 2A). Among these DEGs, 21 genes were significantly downregulated, and 34 genes were significantly upregulated. The volcano plot in Figure 2B shows the distribution of the top 50 DEGs in ULMS and ULM.

Correlation and functional enrichment analysis

The GO analysis indicated that the DEGs mainly participated in chromosome segregation and the cell cycle

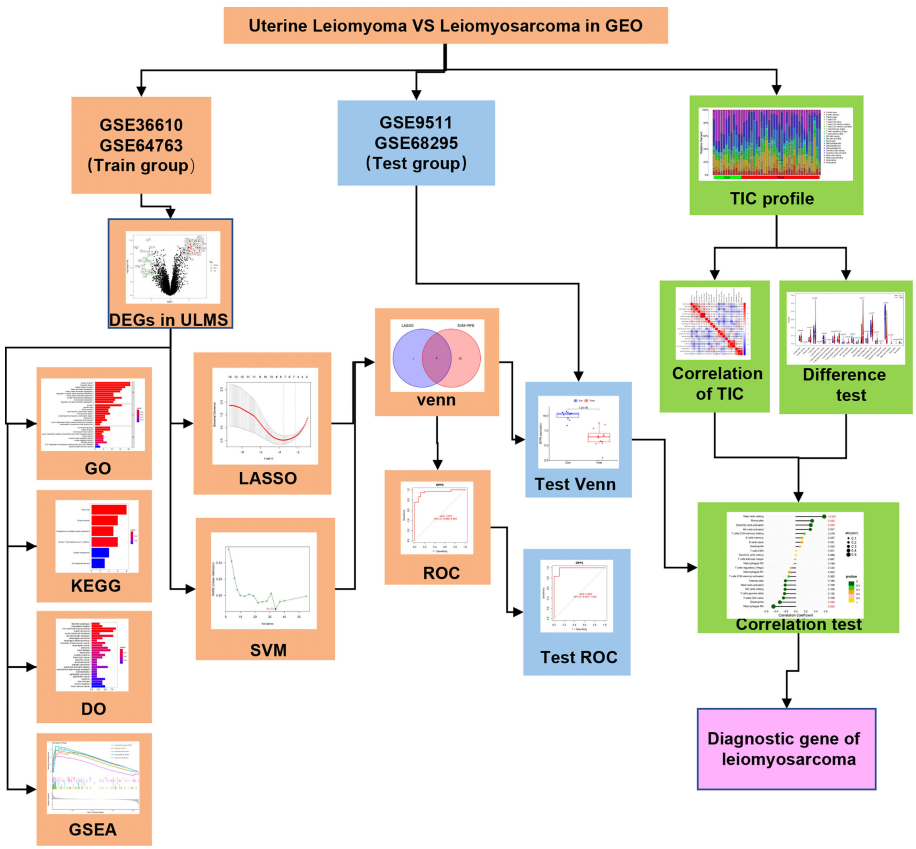


FIGURE 1
Analysis flow diagram of this study.

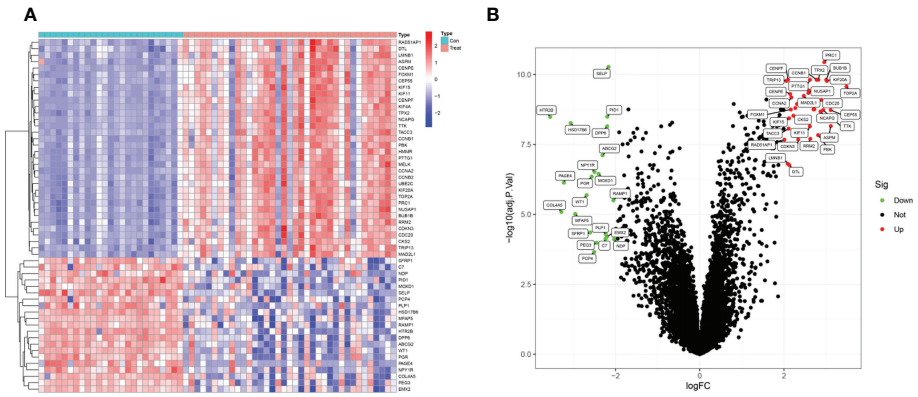


FIGURE 2
Identifcation of DEGs. **(A)** Heatmap plots of 55 DEGs between ULMS and ULM samples from GEO database. Row name of heatmap is the gene name, and column name is the ID of samples which not shown in plot.The colors from red toblue represent expression level from high to low in the heatmaps. **(B)** Volcano plots of top 50 DEGs between ULMS and ULM samples. The red dots in the volcano plots represent up-regulation, the green dots represent down-regulation and black dots represent genes without differential expression.

(Figure 3A). In addition, KEGG analysis showed enrichment of the cell cycle and immune-related pathways, such as HTLV-1 infection way (Figure 3B). The DO enrichment showed that DEGs were mostly related to solid malignant tumors and haematological malignancies (Figure 3C). The GSEA results revealed negative enrichment in cell adhesion and the Wnt signalling pathway in ULM (Figure 3D, Table S1) as well as positive enrichment in the cell cycle, DNA replication and mismatch repair in ULMS (Figure 3E, Table S2). These results indicated that mismatch repair, related immunity and the cell cycle play vital roles in ULMS.

Verification and validation of diagnostic biomarkers

We employed the LASSO and SVM-RFE algorithm methods to select potential biomarkers. We identified 7 DEGs as diagnostic biomarkers using LASSO regression for ULMS (Figure 4A), and we verified 34 DEGs using the SVM-RFE (Figure 4B). When integrating both algorithms, six overlapping candidate genes (*PRC1*, *SELP*, *PID1*, *DPP6*, *MFAP5* and *HSD17B6*) were selected (Figure 4C). In addition, with the purpose of producing more dependable and exact DEGs, we verified the expression levels of six DEGs using the GSE9511 and GSE68295 datasets. The *DPP6* and *MFAP5* expression levels in ULMS samples were significantly lower than those in the ULM group (Figures 5A, B; $P < 0.05$). However, *SLEP* gene expression was not significantly different between the two groups (Figure 5C). Subsequently, we investigated the latent ability of the two identified DEGs as diagnostic biomarkers utilizing a logistic regression algorithm.

Effectiveness of diagnostic biomarkers in ULMS

The ability of the two diagnostic biomarkers indicated good diagnostic value in early discernment of ULMS as the AUC values of the *DPP6* and *MFAP5* genes were 0.957 and 0.899, respectively (Figure 6A). Subsequently, a persuasive screening capacity was verified in the GSE9511 and GSE68295 datasets with AUC values of 0.983 in *DPP6* and 0.942 in *MFAP5* (Figure 6B). We assessed the expression of *DPP6* and *MFAP5* across ULMS and ULM tissues *via* immunohistochemistry and found that low expression of *DPP6* and *MFAP5* were associated with ULMS. DPP6 was expressed in the cytoplasm, MFAP5 was expressed in the stroma. (Figures 6C, D; $P < 0.05$). The above results indicating that the *DPP6* gene and *MFAP5* had a higher diagnostic capacity.

DPP6 and *MFAP5* genes correlate with the percentage of immune cell infiltration

Next, we verified the correlation of the *DPP6* and *MFAP5* genes with immune cell infiltration. We determined the proportions of 22 immune cells in the ULMS and ULM samples using the CIBERSORT algorithm (Figures 7A, B). The components of immune cells in the ULMA vs. ULM group were explored. The ratios of resting CD4+ memory T cells ($P = 0.023$), activated NK cells ($P = 0.031$) and resting mast cells ($P < 0.001$) in the ULMS group were markedly lower than those in the ULM group. However, the ratio of M0 macrophages ($P = 0.011$) was significantly higher in ULMS compared to ULM (Figure 7C).

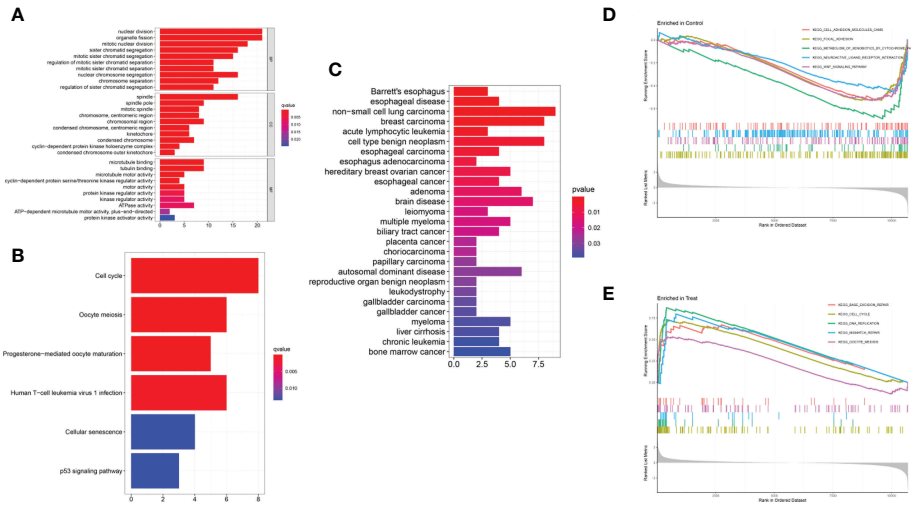


FIGURE 3
Functional enrichment analyses to identify potential biological processes. **(A)** GO analysis. GO analysis divided DEGs into three functional groups: molecular function (MF), biological processes (BP), and cell composition (CC). **(B)** KEGG analysis of DEGs. **(C)** Disease ontology enrichment analysis of DEGs between ULMS and ULM samples. **(D, E)** Enrichment analyses between ULMS and ULM samples *via* gene set enrichment analysis.

Furthermore, we studied the relationship between the *DPP6* and *MFAP5* genes and infiltrating immune cells. *DPP6* was positively related to resting mast cells (r = 0.570, p < 0.001), monocytes (r = 0.328, P = 0.032) and activated dendritic cells (r = 0.301, p =0.0495) but negatively related to eosinophils (r = −0.321, P = 0.036) and M0 macrophages (r = -0.450, P =0.003) (Figure 7D). Moreover, *MFAP5* was positively related to resting mast cells (r = 0.413, p = 0.006) (Figure 7E). These findings supported that *DPP6* and *MFAP5* are related to immune activity.

Discussion

ULMS is one of the most common subtypes in mesenchymal neoplasms, but research on ULMS is limited. Because the incidence rate is low, different clinical features and histopathological appearances result in a lack of molecular biomarkers, offering no superior treatment regimen (21). The biological behaviour of ULMS is difficult to predict. Although the tumor is often restricted to the uterus, recurrence and metastasis are highly common (22). An increasing number of studies have employed immune cells as a new bioinformatic approach to investigate the diagnosis and prognosis of various diseases, including gastric cancer (23), breast cancer (24) and osteosarcoma (25). However, there are few studies on the immune cell infiltration association with DEGs in ULMS. Thus, we focused on the identification of significant diagnostic DEGs for ULMS and determined the correlation of these DEGs with infiltrating immune cells in ULMS.

To our knowledge, our study is the first to apply multiple GEO datasets for knowledge mining using a machine learning approach in ULMS to identify significant diagnostic biomarkers related to immune cells. In the present study, which utilized the GSE36610, GSE64763, GSE9511 and GSE68295 datasets from the GEO database, 55 DEGs were identified by comparing ULMS and ULM samples. DO enrichment showed that the 55 DEGs were mainly related to solid malignant tumors and haematological malignancies. KEGG analysis and GSEA indicated that the DEGs were involved in regulating immune-related pathways and the cell cycle. Risinger et al. reported that defective postreplication mismatch repair resulting in microsatellite instability is present in considerable portions of sarcomas in gynecology (26). Similarly, mismatch repair (MMR) protein has been screened in uterine carcinosarcomas and leiomyosarcomas by immunohistochemical assays but has not been identified in other types of uterine mixed epithelial/mesenchymal or mesenchymal malignancies (27). Anderson et al. found that p53 expression may act as a prognostic biomarker for ULMS (28). Abnormal p53 staining (null or strong/diffuse) has been observed in ULMS with 70% sensitivity and 100% specificity against inflammatory myofibroblastic tumors (IMTs) and is related to genomic alterations (29). Relevant study has demonstrated that HTLV-1 infection correlates with the occurrence of ULMS. However, HTLV-1 has been thoroughly

10.3389/fonc.2022.1084192

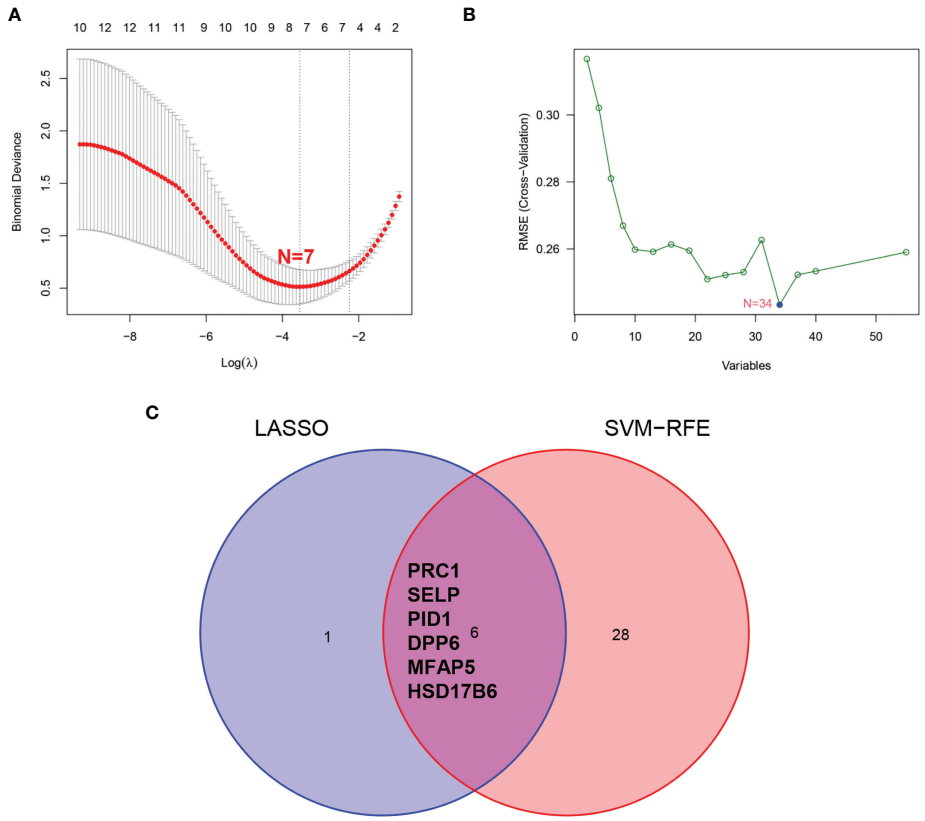


FIGURE 4
Screening process of diagnostic biomarker candidates for ULMS diagnosis. **(A)** Tuning feature selection in the LASSO model. **(B)** A plot of biomarkers selection *via* SVM-RFE algorithm. **(C)** Venn diagram demonstrating 6 diagnostic biomarkers shared by the LASSO and SVM-RFE algorithms.

studied in adult T-cell leukaemia/lymphoma (ATL) (30–32), an aggressive CD4+ T-cell malignancy. HTLV-1 increases genomic instability by directly altering the expression of host genes; conversely, abnormal gene expression may influence the longevity of infected CD+4 T cell clones and profileration rate, allowing further mutations to accumulate and the host genome structure to vary, ultimately leading to malignant transformation (33). Because HTLV-1 mediates immune-related pathways, it possible that regulation of the immune response is strongly associated with the occurrence of ULMS.

We identified two diagnostic biomarkers based on integrating two machine-learning algorithms and diagnostic ability analysis, and we verified these markers using the GSE9511 and GSE68295 datasets. The dipeptidyl peptidase 6

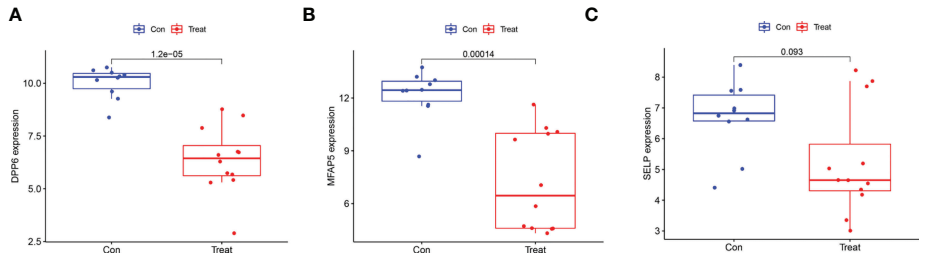


FIGURE 5
Validation of the expression of diagnostic biomarkers in the GSE9511 and GSE68295. **(A)** *DPP6*; **(B)** *MFAP5*; **(C)** *SLEP*.

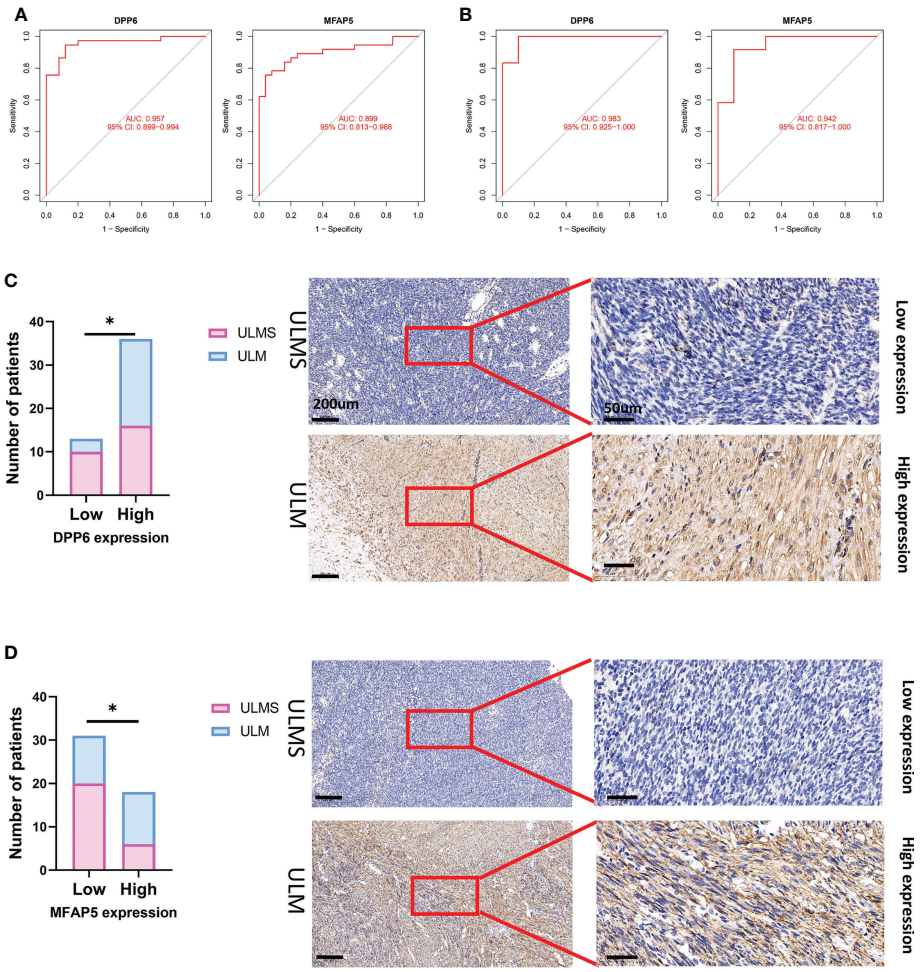


FIGURE 6
The receiver operating characteristic (ROC) curve of the diagnostic effectiveness of the six diagnostic markers. **(A)** ROC curve of *DPP6* and *MFAP5* after fitting to one variable in the metadata cohort; **(B)** ROC curve of *DPP6* and *MFAP5* after fitting to one variable in the GSE9511 and GSE68295 dataset. **(C)** Significantly low *DPP6* expression was observed in ULMS tissues compared with ULM specimens (ULMS=26, ULM=23). Representative images (×50 and ×400) of IHC staining for *DPP6* in 26 ULMS and 23 ULM patients (high expression vs. low expression). **(D)** Significantly low *MFAP5* expression was observed in ULMS tissues compared with ULM specimens (ULMS=26, ULM=23). Representative images (×50 and ×400) of IHC staining for *MFAP5* in 26 ULMS and 23 ULM patients (high expression vs. low expression). Scale bars are shown. *P < 0.05. P values were calculated by chi-square tests.

(*DPP6*) gene encodes a single transmembrane peptidase without activity. Most likely, *DPP6* enhances its expression and regulates its gating feature by combining at the permeation and gating modules of the potassium channel (34). In breast cancer tissues, *DPP6* has low expression at the transcription and protein levels, and in breast cancer patients, low expression of *DPP6* indicates poor prognoses, suggesting that *DPP6* may serve as a tumour suppressor in tumour development (35), which agreed with our study. However, in surgically treated clear cell renal cell carcinoma (ccRCC) patients, the promoter methylation of *DPP6* genes is related to an aggressive phenotype and early progression of distant metastasis (36). Similarly, in pancreatic ductal adenocarcinoma tissues, the promoter methylation of *DPP6* genes is significantly higher than that in normal tissues (37). Microfibril-associated protein 5 (*MFAP5*) is a 25 kDa glycoprotein present in the extracellular matrix and stroma in all tissues (38), and it is crucial for elastic microfibril assembly. Using a microarray to investigate prostate tumors, researchers have detected 3800 significant expression alterations between the tumor stroma and benign stroma, and they reported that the downregulation of *MFAP5* expression is the most significant alteration in the prostate cancer stroma among all genes examined (39). Significant loss of *MFAP5* expression in colon cancer stroma may facilitate the difference between pseudoinvasive and true invasive tumors with a specificity of 75% and a sensitivity of 80% in colonic adenomatous polyps

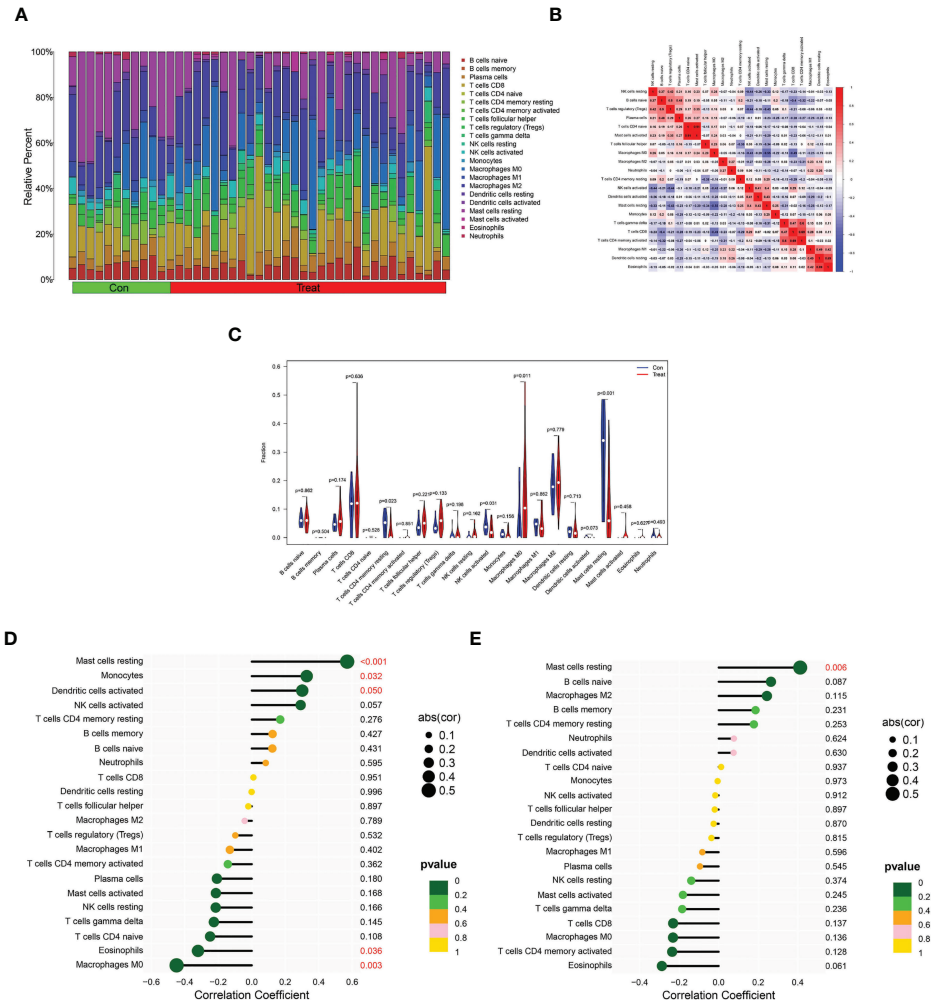


FIGURE 7
Distribution and visualization of immune cell infiltration and correlation analysis. **(A)** Barplot showing the proportion of 22 immune cell subtypes between ULMS and ULM samples. **(B)** Heatmap showing the correlation matrix of all 22 immune cell subtype compositions. Both horizontal and vertical axes demonstrate immune cell subtypes. Immune cell subtype compositions (higher, lower, and same correlation levels are displayed in red, blue, and white, respectively), and Pearson coefficient was used for significance test. **(C)** Violin plot showed the the total distribution of immune cells in ULMS and ULM samples. Correlation between *DPP6* **(D)**, *MFAP5* **(E)** and infiltrating immune cells in ULMS.

(40). However, high expression levels of *MFAP5* are associated with a worse prognosis in ovarian cancer (both in epithelium and stroma). In the present study, we observed significantly low expression of *MFAP5* in the stromal component of ULMS specimens, similar to the above study.

We applied CIBERSORT to assess the types of immune cell infiltration in ULMS and ULM. We discovered that decreased infiltration of resting CD4+ memory T cells, activated NK cells and resting mast cells in addition to increased infiltration of M0 macrophages were potentially correlated with the occurrence and development of ULMS. Xiaoqing et al. found that the infiltration of two types of immune cells (resting mast cells resting and activated NK cells) is lower in ULMS tissues, while the infiltration of five types of immune cells (memory B cells, M0 macrophages, activated mast cells, M1 macrophages and follicular helper T cells) is higher in ULMS tissues than in normal myometrium (NL) tissues (41). Similarly, our study demonstrated that the infiltration of immune cell types was lower due to the selection of the control group. Additionally, we found that the *DPP6* gene was positively correlated with resting mast cells, monocytes and activated dendritic cells. However, M0 macrophages and eosinophils had a negative correlation with the *DPP6* gene. Together, these findings indicated that the *DPP6* gene is associated with several types of immune cell infiltration and plays an important role in ULMS, suggesting that should be a focus in future experimental work.

The present study had limitations. First, due to the low incidence rate of ULMS, the number of cases was not enough in

the GSE36610 and GSE64763 datasets. Second, the function and reproducibility of the *DPP6* and *MFAP5* genes as well as the related immune cell infiltration should be further validated by prospective studies with larger sample sizes in ULMS.

Conclusion

Based on the GEO database, the two hub genes and the infiltration of five types of immune cells were related to ULMS occurrence. *DPP6* and *MFAP5* genes may affect the occurrence of ULMS through immune-related pathways. Thus, these findings provided molecular evidence for the treatment of ULMS in the future.

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 human participants were reviewed and approved by Ethics committee of The Second Affiliated Hospital of Fujian Medical University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Fujian Provincial Health Technology Project (No. 2019-1-15).

Acknowledgments

The authors acknowledge the GEO database for providing data of ULMS and ULM available.

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.2022.1084192/full#supplementary-material

References

1. Kobayashi H, Uekuri C, Akasaka J, Ito F, Shigemitsu A, Koike N, et al. The biology of uterine sarcomas: A review and update. *Mol Clin Oncol* (2013) 1:599–609. doi: 10.3892/mco.2013.124

2. Ricci S, Stone RL, Fader AN. Uterine leiomyosarcoma: Epidemiology, contemporary treatment strategies and the impact of uterine morcellation. *Gynecol Oncol* (2017) 145:208–16. doi: 10.1016/j.ygyno.2017.02.019

3. George S, Serrano C, Hensley ML, Ray-Coquard I. Soft tissue and uterine leiomyosarcoma. *J Clin Oncol* (2018) 36:144–50. doi: 10.1200/JCO.2017.75.9845

4. D'Ambrosio L, Touati N, Blay JY, Grignani G, Flippot R, Czarnecka AM, et al. Doxorubicin plus dacarbazine, doxorubicin plus ifosfamide, or doxorubicin alone as a first-line treatment for advanced leiomyosarcoma: A propensity score matching analysis from the European organization for research and treatment of cancer soft tissue and bone sarcoma group. *Cancer* (2020) 126:2637–47. doi: 10.1002/cncr.32795.European Organization for Research and Treatment of Cancer Soft Tissue and Bone Sarcoma Group.

5. Casali PG, Abecassis N, Aro HT, Bauer S, Biagini R, Bielack S, et al. Soft tissue and visceral sarcomas: ESMO-EURACAN clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* (2018) 29:iv268–9. doi: 10.1093/annonc/mdy321

6. Kapp DS, Shin JY, Chan JK. Prognostic factors and survival in 1396 patients with uterine leiomyosarcomas: emphasis on impact of lymphadenectomy and oophorectomy. *Cancer* (2008) 112:820–30. doi: 10.1002/cncr.23245

7. Roberts ME, Aynardi JT, Chu CS. Uterine leiomyosarcoma: A review of the literature and update on management options. *Gynecol Oncol* (2018) 151:562–72. doi: 10.1016/j.ygyno.2018.09.010

8. Mori KM, Abaid LN, Mendivil AA, Brown JV3rd, Beck TL, Micha JP, et al. The incidence of occult malignancy following uterine morcellation: A ten-year single institution experience retrospective cohort study. *Int J Surg* (2018) 53:239–42. doi: 10.1016/j.ijsu.2018.03.075

9. Kaganov H, Ades A, Fraser DS. Preoperative magnetic resonance imaging diagnostic features of uterine leiomyosarcomas: A systematic review. *Int J Technol Assess Health Care* (2018) 34:172–9. doi: 10.1017/S0266462318000168

10. Major FJ, Blessing JA, Silverberg SG, Morrow CP, Creasman WT, Currie JL, et al. Prognostic factors in early-stage uterine sarcoma. *cancer* (1993) 71:1702–9. doi: 10.1002/cncr.2820710440

11. Marret H, Fritel X, Ouldamer L, Bendifallah S, Brun JL, De Jesus I, et al. And cngof, therapeutic management of uterine fibroid tumors: updated French

guidelines. *Eur J Obstet Gynecol Reprod Biol* (2012) 165:156–64. doi: 10.1016/j.ejogrb.2012.07.030

12. D'Angelo E, Prat J. Uterine sarcomas: A review. *Gynecol Oncol* (2010) 116:131–9. doi: 10.1016/j.ygyno.2009.09.023

13. Pritts EA, Vanness DJ, Berek JS, Parker W, Feinberg R, Feinberg J, et al. The prevalence of occult leiomyosarcoma at surgery for presumed uterine fibroids: a meta-analysis. *Gynecol Surg* (2015) 12:165–77. doi: 10.1007/s10397-015-0894-4

14. Giraldo NA, Becht E, Remark R, Damotte D, Sautes-Fridman C, Fridman WH. The immune contexture of primary and metastatic human tumours. *Curr Opin Immunol* (2014) 27:8–15. doi: 10.1016/j.coi.2014.01.001

15. Wu J, Zhao Y, Zhang J, Wu Q, Wang W. Development and validation of an immune-related gene pairs signature in colorectal cancer. *Oncoimmunology* (2019) 8:1596715. doi: 10.1080/2162402X.2019.1596715

16. Zhao E, Zhou C, Chen S. A signature of 14 immune-related gene pairs predicts overall survival in gastric cancer. *Clin Transl Oncol* (2021) 23:265–74. doi: 10.1007/s12094-020-02414-7

17. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* (2012) 28:882–3. doi: 10.1093/bioinformatics/bts034

18. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. *Machine Learn.* (2002) 46:389–422. doi: 10.1023/A:1012487302797

19. 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

20. Chen H, Wang J, Yang H, Chen D, Li P. Association between FOXM1 and hedgehog signaling pathway in human cervical carcinoma by tissue microarray analysis. *Oncol Lett* (2016) 12:2664–73. doi: 10.3892/ol.2016.4932

21. Tsuyoshi H, Yoshida Y. Molecular biomarkers for uterine leiomyosarcoma and endometrial stromal sarcoma. *Cancer Sci* (2018) 109:1743–52. doi: 10.1111/cas.13613

22. Desar IME, Ottevanger PB, Benson C, van der Graaf WTA. Systemic treatment in adult uterine sarcomas. *Crit Rev Oncol Hematol* (2018) 122:10–20. doi: 10.1016/j.critrevonc.2017.12.009

23. Xiang A, Lin X, Xu L, Chen H, Guo J, Zhou F. PCOLCE is potent prognostic biomarker and associates with immune infiltration in gastric cancer. *Front Mol Biosci* (2020) 7:544895. doi: 10.3389/fmolb.2020.544895

24. Yao S, Cheng TD, Elkhanany A, Yan L, Omilian A, Abrams SI, et al. Breast tumor microenvironment in black women: A distinct signature of CD8+ T-cell exhaustion. *J Natl Cancer Inst* (2021) 113:1036–43. doi: 10.1093/jnci/djaa215

25. Zheng J-H, Zhang C, Lin Z-H, Lv H-Y, Ye Z-M, Chen Y-P, et al. Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of osteosarcoma. *Aging* (2020) 12:3486–501. doi: 10.18632/aging.102824

26. Risinger JI, Boyer JC, Evans AC, Berchuck A, Kunkel TA, Barrett JC. Microsatellite instability in gynecological sarcomas and in hMSH2 mutant uterine sarcoma cell lines defective in mismatch repair activity. *Cancer Res* (1995) 55:5664–9.

27. Hoang LN, Ali RH, Lau S, Gilks CB, Lee CH. Immunohistochemical survey of mismatch repair protein expression in uterine sarcomas and carcinosarcomas. *Int J Gynecol Pathol* (2014) 33:483–91. doi: 10.1097/PGP.0b013e31829ff239

28. Anderson SE, Chuai S, Olshen AB, D. CHI, Sabbatini P, Soslow RA. p53, epidermal growth factor, and platelet-derived growth factor in uterine leiomyosarcoma and leiomyomas. *Int J Gynecol Cancer* (2006) 16:849–53. doi: 10.1111/j.1525-1438.2006.00542.x

29. Schaefer IM, Hornick JL, Sholl LM, Quade BJ, Nucci MR, Parra-Herran C. Abnormal p53 and p16 staining patterns distinguish uterine leiomyosarcoma from inflammatory myofibroblastic tumour. *Histopathology* (2017) 70:1138–46. doi: 10.1111/his.13176

30. Kannagi M, Hasegawa A, Nagano Y, Kimpara S, Suehiro Y. Impact of host immunity on HTLV-1 pathogenesis: potential of tax-targeted immunotherapy against ATL. *Retrovirology* (2019) 16:23. doi: 10.1186/s12977-019-0484-z

31. Kannagi M, Ohashi T, Harashima N, Hanabuchi S, Hasegawa A. Immunological risks of adult T-cell leukemia at primary HTLV-I infection. *Trends Microbiol* (2004) 12:346–52. doi: 10.1016/j.tim.2004.05.005

32. Matsuoka M. Human T-cell leukemia virus type I (HTLV-I) infection and the onset of adult T-cell leukemia (ATL). *Retrovirology* (2005) 2:27. doi: 10.1186/1742-4690-2-27

33. Cook L, Melamed A, Yaguchi H, Bangham CR. The impact of HTLV-1 on the cellular genome. *Curr Opin Virol* (2017) 26:125–31. doi: 10.1016/j.coviro.2017.07.013

34. Kaulin YA, De Santiago-Castillo JA, Rocha CA, Nadal MS, Rudy B, Covarrubias M. The dipeptidyl-peptidase-like protein DPP6 determines the unitary conductance of neuronal Kv4.2 channels. *J Neurosci* (2009) 29:3242–51. doi: 10.1523/JNEUROSCI.4767-08.2009

35. Choy TK, Wang CY, Phan NN, Khoa Ta HD, Anuraga G, Liu YH, et al. Identification of dipeptidyl peptidase (DPP) family genes in clinical breast cancer patients via an integrated bioinformatics approach. *Diagnostics (Basel)* (2021) 11 (7):1204. doi: 10.3390/diagnostics11071204

36. Kang HW, Park H, Seo SP, Byun YJ, Piao XM, Kim SM, et al. Methylation signature for prediction of progression free survival in surgically treated clear cell renal cell carcinoma. *J Korean Med Sci* (2019) 34:e144. doi: 10.3346/jkms.2019.34.e144

37. Zhao X, Cao D, Ren Z, Liu Z, Lv S, Zhu J, et al. Dipeptidyl peptidase like 6 promoter methylation is a potential prognostic biomarker for pancreatic ductal adenocarcinoma. *Biosci Rep* (2020) 40(7):BSR20200214. doi: 10.1042/BSR20200214

38. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. tissue-based map of the human proteome. *Science* (2015) 347:1260419. doi: 10.1126/science.1260419

39. Jia Z, Wang Y, Sawyers A, Yao H, Rahmatpanah F, Xia XQ, et al. Diagnosis of prostate cancer using differentially expressed genes in stroma. *Cancer Res* (2011) 71:2476–87. doi: 10.1158/0008-5472.CAN-10-2585

40. Zhao L, Westerhoff M, Hornick JL, Krausz T, Antic T, Xiao SY, et al. Loss of microfibril-associated protein 5 (MFAP5) expression in colon cancer stroma. *Virchows Arch* (2020) 476:383–90. doi: 10.1007/s00428-019-02649-y

41. Shen X, Yang Z, Feng S, Li Y. Identification of uterine leiomyosarcoma-associated hub genes and immune cell infiltration pattern using weighted co-expression network analysis and CIBERSORT algorithm. *World J Surg Oncol* (2021) 19:223. doi: 10.1186/s12957-021-02333-z

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 01 December 2022
DOI 10.3389/fonc.2022.1085065

OPEN ACCESS

EDITED BY
Hao Zhang,
Chongqing Medical University, China

REVIEWED BY
Lingping Yang,
Chongqing Medical University, China
Tiezhu Chen,
Hunan Provincial People's Hospital, China

*CORRESPONDENCE
Xu Cao
hughcaoxu@hotmail.com

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Oncology

RECEIVED 31 October 2022
ACCEPTED 15 November 2022
PUBLISHED 01 December 2022

CITATION
Zhang L, Wu S, Huang J, Shi Y, Yin Y and Cao X (2022) A mitochondria-related signature for predicting immune microenvironment and therapeutic response in osteosarcoma. *Front. Oncol.* 12:1085065. doi: 10.3389/fonc.2022.1085065

A mitochondria-related signature for predicting immune microenvironment and therapeutic response in osteosarcoma

Lina Zhang, Song Wu, Junjie Huang, Yanbin Shi, Yuesong Yin and Xu Cao*

Department of Orthopaedics, The Third Xiangya Hospital, Central South University, Changsha, China

Background: Osteosarcoma remains to be the most devastating malignant tumor in children and teenagers. Mitochondria have also been proven to play critical roles in osteosarcoma. However, a mitochondria-related signature has been established in osteosarcoma to comprehensively evaluate the pathogenic roles and regulatory roles of mitochondria in osteosarcoma.

Methods: In this study, osteosarcoma samples' transcriptome data and clinical information were collected from Therapeutically Applicable Research to Generate Effective Treatments (TARGET) and Gene Expression Omnibus (GEO) databases. A comprehensive bioinformatics analysis was performed on the samples at the bulk RNA sequencing level and single-cell RNA sequencing (scRNA-seq) level. EdU, Transwell, and immunohistochemistry (IHC) were performed on PCCB.

Results: A mitochondria-related signature was constructed in osteosarcoma patients. The prognostic value of the mitochondria-related signature was explored. The predictive value of the mitochondria-related signature in the immune microenvironment and chemotherapy agents was explored. The association between mitochondria and immunity in the tumor microenvironment of osteosarcoma at the scRNA-seq level was investigated. The tumorigenic role of the critical mitochondria-related gene, PCCB, was verified by *in vitro* validation.

Conclusion: In conclusion, a mitochondria-related signature was developed in osteosarcoma with solid predictive values in the immune microenvironment, chemotherapy agents, and prognosis.

KEYWORDS

osteosarcoma, immune, mitochondria-related signature, therapeutic response, single-cell analysis

Introduction

Osteosarcoma remains the most devastating malignant tumor in children and teenagers (1). Osteosarcoma develops from the mesenchymal cell line, and the rapid growth of the cancer is due to the direct or indirect formation of tumor osteoid tissue and bone tissue during the chondral stage (1). The closer the tumor site is to the trunk, the higher the mortality. The key factors affecting the prognosis are early diagnosis, complete tumor resection, chemotherapy, and radiotherapy before and after surgery. In addition, it is also related to the tissue type and size of tumor cells, the increase of serum alkaline phosphatase before and after surgery, and whether local lymph nodes are involved (2). The primary treatment for osteosarcoma is radical surgical resection. Consolidation of chemical or radiation therapy after tumor resection is significant for controlling tumor metastasis and improving survival rate (3). Immunotherapy involves the intravenous infusion of lymphocytes or interferon and transfer factors, but the efficacy is uncertain (4).

Mitochondria have been well recognized as a critical mediator for oncogenesis. Based on their function as major bioenergy promoters, mitochondria are actively involved in regulating tumor anabolism, controlling REDOX and calcium homeostasis, participating in transcriptional regulation, and controlling cell death. Mitochondrial dysfunction leads to the release of cytochrome C, the production of mitochondrial reactive oxygen species (mtROS), and the generation of metabolites, further initiating signaling cascades that affect gene expression, cell activation, cell proliferation, and cell differentiation (5, 6). Mitochondria may promote malignant transformation through three main mechanisms: (1) Reactive oxygen species (ROS), mainly derived from the mitochondrial respiratory chain, contribute to the accumulation of potential oncogenic DNA defects, and the activation of potential oncogenic signaling pathways (7); (2) Abnormal accumulation of some mitochondrial metabolites, including fumaric acid, succinic acid, and 2-hydroxyglutaric acid (2-Hg) (8); (3) Defective mitochondrial permeability transition (MPT) function promotes the formation of malignant precursors (9).

Mitochondria have also been proven to play critical roles in osteosarcoma. AICAR was reported to induce mitochondrial apoptosis in osteosarcoma through an AMPK-dependent pathway (10). The mitochondrial BIG3-PHB2 complex formation was reported to promote the survival and proliferation of osteosarcoma (11). Besides, targeting autophagy was reported to enhance atezolizumab-induced mitochondria-related apoptosis in osteosarcoma (12). Mitochondria-regulated cell death and energetic metabolism are intimately connected in osteosarcoma (13). However, a mitochondria-related signature has never been established in osteosarcoma to comprehensively evaluate the pathogenic roles and regulatory roles of mitochondria in osteosarcoma. More importantly, the interconnection between mitochondria and the tumor microenvironment of osteosarcoma remains to be deciphered.

In this study, osteosarcoma samples' transcriptome data and clinical information were collected from Therapeutically Applicable Research to Generate Effective Treatments (TARGET) and Gene Expression Omnibus (GEO) databases. A mitochondria-related signature was constructed in osteosarcoma patients. The prognostic value of the mitochondria-related signature was explored. The predictive value of the mitochondria-related signature in the immune microenvironment was explored. The predictive value of the mitochondria-related signature in chemotherapy agents was explored. The association between mitochondria and immunity in the tumor microenvironment of osteosarcoma at the single-cell RNA sequencing (scRNA-seq) level was investigated. The tumorigenic role of the critical mitochondria-related gene, PCCB, was verified by *in vitro* validation. To the best of our knowledge, this is the first study assessing the effect of mitochondria on the prognosis, immune microenvironment, and therapeutic efficacy in osteosarcoma.

Materials and methods

This study was ethically approved by the institutional review board (IRB) of the Third Xiangya Hospital, Central South University (No: 2020-S221). All experiments involving human tissues were performed based on guidelines approved by the IRB. A signed informed consent form was obtained from each patient.

Data collection and procession

84 osteosarcoma samples with transcriptome data and clinical information were accessed from the TARGET database

Abbreviations: TARGET, Therapeutically Applicable Research to Generate Effective Treatments; GEO, Gene Expression Omnibus; scRNA-seq, single-cell RNA sequencing; mtROS, mitochondrial reactive oxygen species; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic; AUC, area under the curve; ssGSEA, single-sample gene-set enrichment analysis; ESTIMATE, Estimated Stromal and Immune cells in Malignant Tumor tissues using Expression data; DEGs, differentially expressed genes; GSVA, gene set variation analysis; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; IHC, Immunohistochemistry; RT-qPCR, Real-time quantitative polymerase chain reaction.

((https://xenabrowser.net/) and were designed as the training cohort. 53 osteosarcoma samples with transcriptome data and clinical information were accessed from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and were designed as the validating cohort (GSE21257). Samples with less than one month follow-up time and a lack of overall survival information were excluded. The count data were normalized using the R package 'DEseq2'. The scRNA-seq data of osteosarcoma samples (primary osteosarcoma lesions, 'BC2' and 'BC3') were accessed from GSE152048 in the GEO database. The mitochondria-related genes were accessed from the MitoCarta3.0 database (http://www.broadinstitute.org/mitocarta).

Construction of the mitochondria-related signature

The univariate Cox regression analysis was performed on the mitochondria-related genes to determine their prognostic values. The least absolute shrinkage and selection operator (LASSO) regression analysis was performed on the prognostic mitochondria-related genes for dimension reduction. The stepwise multivariate Cox regression analysis was further performed on the prognostic mitochondria-related genes for dimension reduction. A mitochondria-related signature was developed based on the following formula: Risk Score=$\sum Expression(Gene)\times Coefficient$.

Prognostic value of the mitochondria-related signature

The survival curves were generated using the R package 'survival'. The survival curves regarding different clinical factors were developed independently. The receiver operating characteristic (ROC) curve was generated using the R package 'timeROC', and the area under the curve (AUC) value was calculated.

Predictive value of the mitochondria-related signature in the immune microenvironment

The single-sample gene-set enrichment analysis (ssGSEA) algorithm was used to quantify the abundance of 28 specific immune cell types using the R package 'GSVA' (14). The Estimated Stromal and Immune cells in Malignant Tumor tissues using Expression data (ESTIMATE) algorithm was used to determine the microenvironment scores using the R package 'estimate' (15).

Functional annotation of the mitochondria-related signature

The differentially expressed genes (DEGs) between two mitochondria-related signature score groups were determined. The DEGs were visualized by volcano plot using the R package 'EnhancedVolcano'. The DEGs were visualized by heatmap using the R package 'pheatmap'. The gene sets of 'c2.cp.kegg.v7.4.symbols' and 'c5.go.bp.v7.4.symbols' were obtained from MSigDB database for performing the gene set variation analysis (GSVA). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was conducted with the R package 'clusterProfiler' and 'org.Hs.eg.db'.

Prediction value of the mitochondria-related signature in chemotherapy agents

The transcriptome data and drug response information of over 1,000 cancer cell lines were collected from the Genomics of Drug Sensitivity in Cancer (GDSC, http://www.cancerrxgene.org/) database. The mitochondria-related signature was developed in each cancer cell line. The Spearman method was used to evaluate the correlation between risk score and half-maximal inhibitory concentration (IC50) of each cancer cell line.

scRNA-seq analysis of the mitochondria-related signature

The scRNA-seq matrix of primary osteosarcoma samples from GSE152048 was processed using the R package 'Seurat'. The function 'NormalizedData' was used to normalize the scRNA-seq data. The function 'FindVariableFeatures' was used to identify the 1,000 most variable genes. The function 'RunPCA' was used for dimension reduction. A K-nearest neighbor was analyzed using the function 'FindNeighbors', and the cells were combined with the function 'FindClusters'. The function Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) was used for visualization. All cells were annotated using the R package 'Single R'. The function 'FindMarkers' was used to find DEGs between two mitochondria-related signature score groups of osteosarcoma cells. The pseudotime trajectory analysis was performed using the R package 'monocle' (16). GO and KEGG enrichment analysis was conducted with the R package 'clusterProfiler' and 'org.Hs.eg.db'. The cell communication pattern was explored using the R package 'iTalk'.

Immunohistochemistry

Three pairs of formalin-fixed paraffin-embedded osteosarcoma tissue and para-carcinoma tissue blocks from three osteosarcoma patients (post-chemotherapy) were collected and used for 5 μm paraffin sections. IHC was performed following the manufacturer's protocol of the Rabbit-enhanced polymer method detection system (ZSGB-BIO, PV-9000, China). The slides were deparaffinized and rehydrated using xylene and gradient-concentration ethyl alcohol. The antigen retrieval was performed with sodium citrate at 95°C. The slides were blocked using an endogenous peroxidase blocker for 10 min at room temperature. Samples were incubated with primary antibody against PCCB (127549, Zenbio, China) overnight at 4°C, reaction enhancer for 20 min at 37°C, and enhanced enzyme-conjugated sheep anti-rabbit IgG polymer for 20 min at 37°C. The slides were stained with 3, 30-diaminobenzidine tetrahydrochloride (DAB) and counterstained with hematoxylin.

Cell culture

Two osteosarcoma cell lines (U2OS and MNNG/HOS) were obtained from the Procell Life Science & Technology Co., Ltd. U2OS and MNNG/HOS were correspondingly cultured in McCoy's 5A (Procell, China) and MEM (Procell, China) supplemented with 10% fetal bovine serum (FBS, Gibco, USA) and 1% penicillin-streptomycin solution (Biosharp, China) at 37°C with saturated humidity and 5% CO2. The average time of culture medium exchange was 24-48h. The cells were digested with trypsin-EDTA (Gibco, USA) and passaged when cell adhesion exceeded 80% confluency.

Small interfering RNA transfection

The PCCB siRNA (si-PCCB) and the nonspecific control siRNA (si-NC) were synthesized by JTSBio (Wuhan, China). The siRNAs sequences are as follows: PCCB-1 (F: CCCUAACAGACUUCACGUUTT R: AACGUGAAGUCUGUUAGGGTT), PCCB-2 (F: CCAAGCUUCUCUACGCAUUTT R: AAUGCGUAGAGAAGCUUGGTT), PCCB-3 (F: CCGCAGAGAUUGCAGUCAUTT R: AUGACUGCAAUCUCUGCGGTT). The siRNAs were transfected into U2OS and MNNG/HOS cells using a jetPRIME transfection reagent (Polyplus, France). RNA extraction was performed 48h after transfection.

Real-time quantitative polymerase chain reaction

The primer sequences are as follows: PCCB (F: TGTCTTCAGTCAGGATTTTACAGTT R: GGCCTGGTCCATGATTTTGC), GAPDH (F: AATGGGCAGCCGTTAGGAAA R: GCCCAATACGACCAAATCAGAG). Total RNA from cultured cells was extracted using Rnafast200 (Fastagen, Japan), and cDNA was synthesized using HiScript II Q RT SuperMix for qPCR (Vazyme, China). ChamQ Universal SYBR qPCR Master Mix (Vazyme, China) was used to conduct RT-qPCR based on the manufacturer's protocol. All steps for the RT-qPCR reaction were conducted as follows: initial denaturation at 95°C for 30s, one cycle; denaturation at 95°C for 10s, 40 cycles; dissolution curve at 95°C for 15s, 60°C for 60s, 95°C for 15s, one cycle. Gene expression levels were normalized to those of GAPDH and calculated using lg2–△△Ct method.

Western blot

A mixture of RIPA (Beyotime, China) and PMSF (Beyotime, China) was used to lyse U2OS and MNNG/HOS cells for protein extraction. Loading Buffer (Biosharp, China) was added to the protein supernatant, and then the sample was boiled to denature the protein. Then proteins were separated using SDS–PAGE gel (Biosharp, China), transferred to PVDF membranes (Millipore, USA), and blocked in 5% skimmed milk for 1 h. Then membranes were incubated overnight at 4°C with primary antibodies, including PCCB (127549, Zenbio, China) and GAPDH (10494-1-AP, Proteintech, China). The membranes were incubated with HRP-conjugated secondary antibody (SA00001-2, Proteintech, China) the following day. Protein bands were captured with a UVP Chem studio PLUS 815 (Analytik Jena, Germany).

EdU assay

Proliferating U2OS and MNNG/HOS cells were identified using the EdU Imaging Kits (APExBIO, USA), and cell nuclei were stained using Hoechst (Invitrogen, USA). Image Pro-Plus version 6.0 (Media Cybernetics, USA) was used for counting EdU-positive cells.

Statistical analysis

All bioinformatics statistical analyses were performed using the R project (version 4.0.3, https://www.r-project.org/). The

Wilcoxon and One-way ANOVA tests were used to compare the difference among groups. All statistical analyses in the cell experiment are based on mean ± SD using Graphpad Prism (version 8.0.2.263). The Benjamini-Hochberg method was used to obtain adjusted p values. The adjusted p value< 0.05 was considered statistically significant.

Results

Construction of the mitochondria-related signature

The univariate Cox regression analysis was performed on the mitochondria-related genes to determine their prognostic values (Figure 1A). The LASSO regression analysis was performed on the prognostic mitochondria-related genes for dimension reduction (Figure 1B). The stepwise multivariate Cox regression analysis was further performed on the prognostic mitochondria-related genes for dimension reduction (Figure 1C). Survival analysis was performed on the prognostic mitochondria-related genes, among which nine genes predicted worse survival and eight genes predicted better survival (Figure S1). The mitochondria-related signature was developed based on the sum of the expression values of the prognostic mitochondria-related genes and their corresponding coefficients. The formula is as follows: OGDH x (1.299)+GUF1 x (-1.34)+FDX1 x (1.115)+ACADVL x (1.335)+PCCB x (1.635) +PDK1 x (0.658)+STOML2 x (0.727)+LACTB x (-0.852) +UQCRB x (1.145)+MFN2 x (-2.086)+CKMT2 x (0.368) +ALDH7A1 x (-0.688)+TRMT1 x (1.189)+EPHX2 x (0.841) +BAK1 x (-1.113)+SPATA20 x (-0.958).

Prognostic value of the mitochondria-related signature

Survival analysis was performed on the two mitochondria-related signature score groups in TARGET, and the high score group was associated with shortened survival time (Figure 1D). Survival analysis was also performed on the two mitochondria-related signature score groups in GSE21257, and the high score group was associated with shortened survival time (Figure 1E). The 1-year, 3-year, and 5-year ROC curves had values of 0.97, 0.942, and 0.951 in TARGET, while the 1-year, 3-year, and 5-year ROC curves had values of 0.689, 0.71, and 0.656 in GSE21257 (Figures 1F, G). The mitochondria-related signature score was not significantly different between the two age groups (Figure 2A). The mitochondria-related signature score was not significantly different between the two gender groups (Figure 2B). Notably, the mitochondria-related signature score was significantly different between the metastatic group and the non-metastatic group (Figure 2C). In osteosarcoma patients with age< 18, the high score group was associated with shortened survival time (Figure 2D). Likewise, in osteosarcoma patients with age > 18, the high score group was associated with shortened survival time (Figure 2D). In male osteosarcoma patients, the high score group was associated with shortened survival time (Figure 2E). Likewise, in osteosarcoma patients, the high score group was associated with shortened survival time (Figure 2E). In metastatic osteosarcoma patients, the high score group was associated with shortened survival time (Figure 2F). Likewise, in non-metastatic osteosarcoma patients, the high score group was associated with shortened survival time (Figure 2F).

Predictive value of the mitochondria-related signature in the immune microenvironment

The mitochondria-related signature was negatively associated with multiple immune cells, including T cell, B cell, natural killer T cell, macrophage, mast cell, and neutrophil (Figure 3A). Notably, central memory CD8 T cell, natural killer cell, CD56bright natural killer cell, macrophage, and activated B cell were the top five cells highly correlated with the mitochondria-related signature (Figure 3B). The high score group was associated with lower levels of microenvironment scores, including stromal score, immune score, and ESTIMATE score (Figure 3C). Besides, the mitochondria-related signature was negatively associated with stromal score (Figure 3D), immune score (Figure 3E), and ESTIMATE score (Figure 3F).

Functional annotation of the mitochondria-related signature

The DEGs between the two mitochondria-related signature score groups were identified (Figure 4A). The distribution of the DEGs between the two mitochondria-related signature score groups is shown in Figure 4B. GO enrichment analysis was performed on the DEGs (Figure 4C). Ossification, embryonic skeletal system development, and pattern specification process were highly enriched in the high score group. T cell activation, extracellular matrix organization, and leukocyte cell-cell adhesion were highly enriched in the low score group. KEGG enrichment analysis was performed on the DEGs (Figure 4D). TGF-β signaling pathway, hippo signaling pathway, and wnt signaling pathway were highly enriched in the high score group. Cytokine-cytokine receptor interaction, ECM-receptor interaction, and focal adhesion were highly enriched in the low score group. Besides, GSVA of GO and KEGG pathways confirmed that autophagosome-lysosome fusion, recognition of apoptotic cell, T cell receptor signaling pathway, and B cell receptor signaling pathway were negatively associated with the mitochondria-related signature (Figure 4E).

10.3389/fonc.2022.1085065

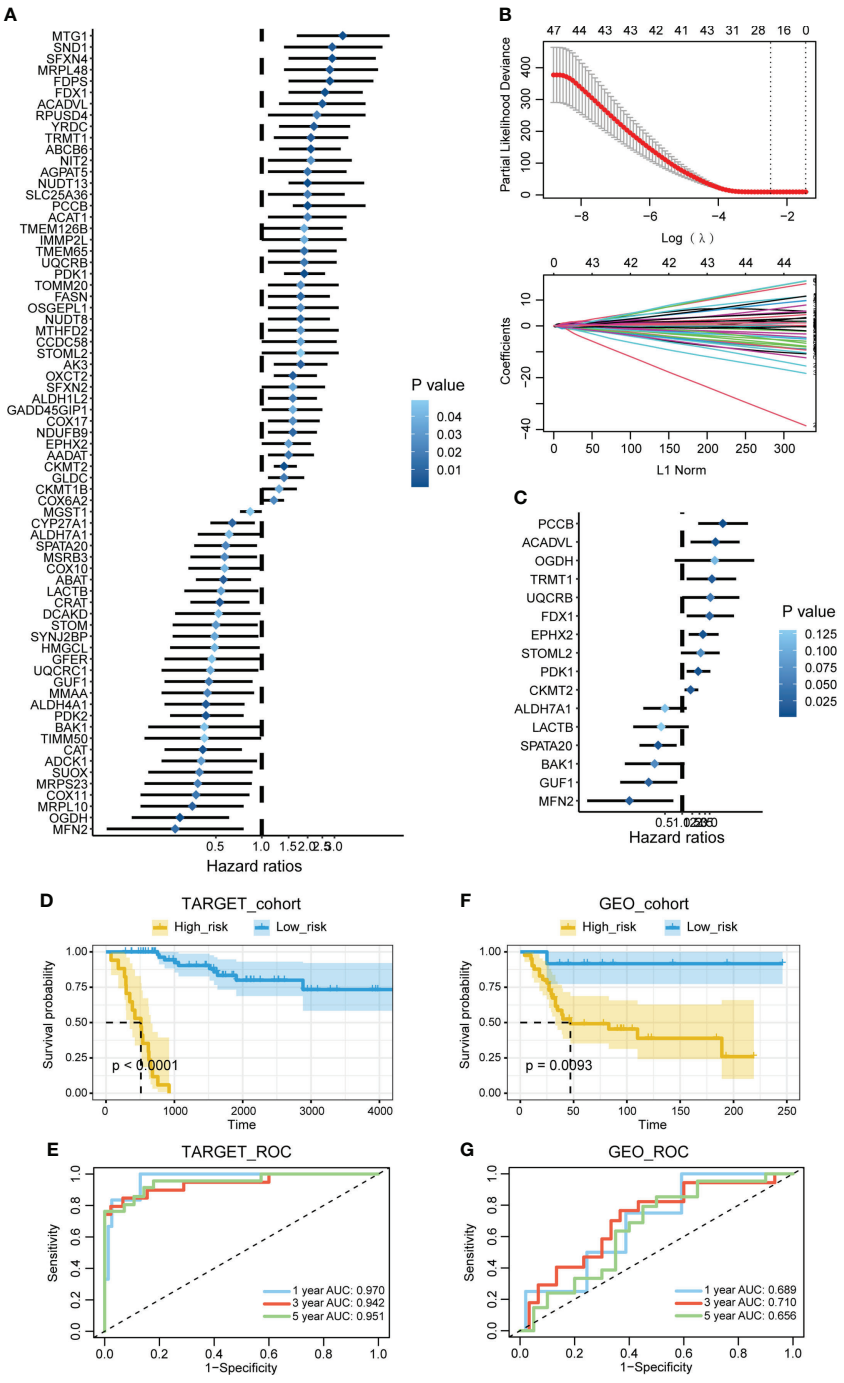


FIGURE 1
Construction of the mitochondria-related signature. **(A)** The univariate Cox regression analysis was performed on the mitochondria-related genes. **(B)** The LASSO regression analysis was performed on the prognostic mitochondria-related genes. **(C)** The stepwise multivariate Cox regression analysis was performed on the prognostic mitochondria-related genes. **(D)** Survival analysis was performed on the two mitochondria-related signature score groups in TARGET. **(E)** Survival analysis was performed on the two mitochondria-related signature score groups in GSE21257. **(F)** The 1-year, 3-year, and 5-year ROC curves of the mitochondria-related signature in TARGET. **(G)** The 1-year, 3-year, and 5-year ROC curves of the mitochondria-related signature in GSE21257.

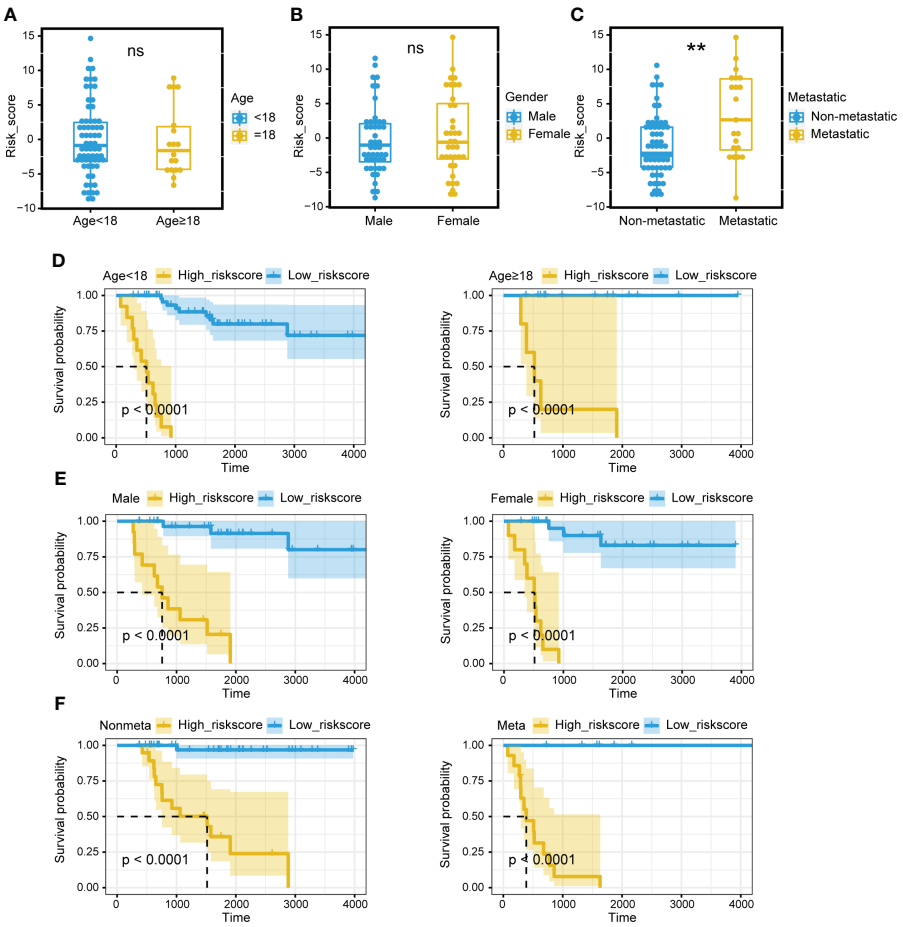


FIGURE 2
Prognostic value of the mitochondria-related signature. **(A)** The different levels of the mitochondria-related signature score regarding age. **(B)** The different levels of the mitochondria-related signature score regarding gender. **(C)** The different levels of the mitochondria-related signature score regarding metastasis. **(D)** Survival analysis was performed on the two mitochondria-related signature score groups regarding age in TARGET. **(E)** Survival analysis was performed on the two mitochondria-related signature score groups regarding gender in TARGET. **(F)** Survival analysis was performed on the two mitochondria-related signature score groups regarding metastasis in TARGET. ns, no significance; **p < 0.01.

Prediction value of the mitochondria-related signature in chemotherapy agents

The potential value of the mitochondria-related signature in predicting chemotherapy agents was explored based on the GSDC database. The correlation between IC50 of drugs and the mitochondria-related signature in cancer cell lines was explored. The drug sensitivity of 30 drugs was identified to be significantly associated with the mitochondria-related signature (Figure S2A). Besides, the targeted signaling pathways of these drugs were exhibited (Figure S2B). 24 drugs were negatively associated with the mitochondria-related signature, including apoptosis regulation inhibitor AZD5991, protein stability and degradation inhibitor ML323, and kinases inhibitor BMS-345541. In addition, six drugs were positively associated with the mitochondria-related signature, including ERK MAPK signaling inhibitor Refametinib, RTK signaling inhibitor NVP-TAE684, and kinases inhibitor A-770041. The overall predicted drug sensitivity and drug resistance in targeted signaling pathways are shown in Figure S2C.

scRNA-seq analysis for the mitochondria-related signature

The identified cells in the tumor microenvironment of osteosarcoma are shown in Figure 5A. The levels of the mitochondria-related signature score in identified cells are shown in Figure 5B. The DEGs between the two mitochondria-related signature score groups of osteosarcoma cells were identified. GO enrichment analysis was performed on the DEGs (Figure 5C). ATP

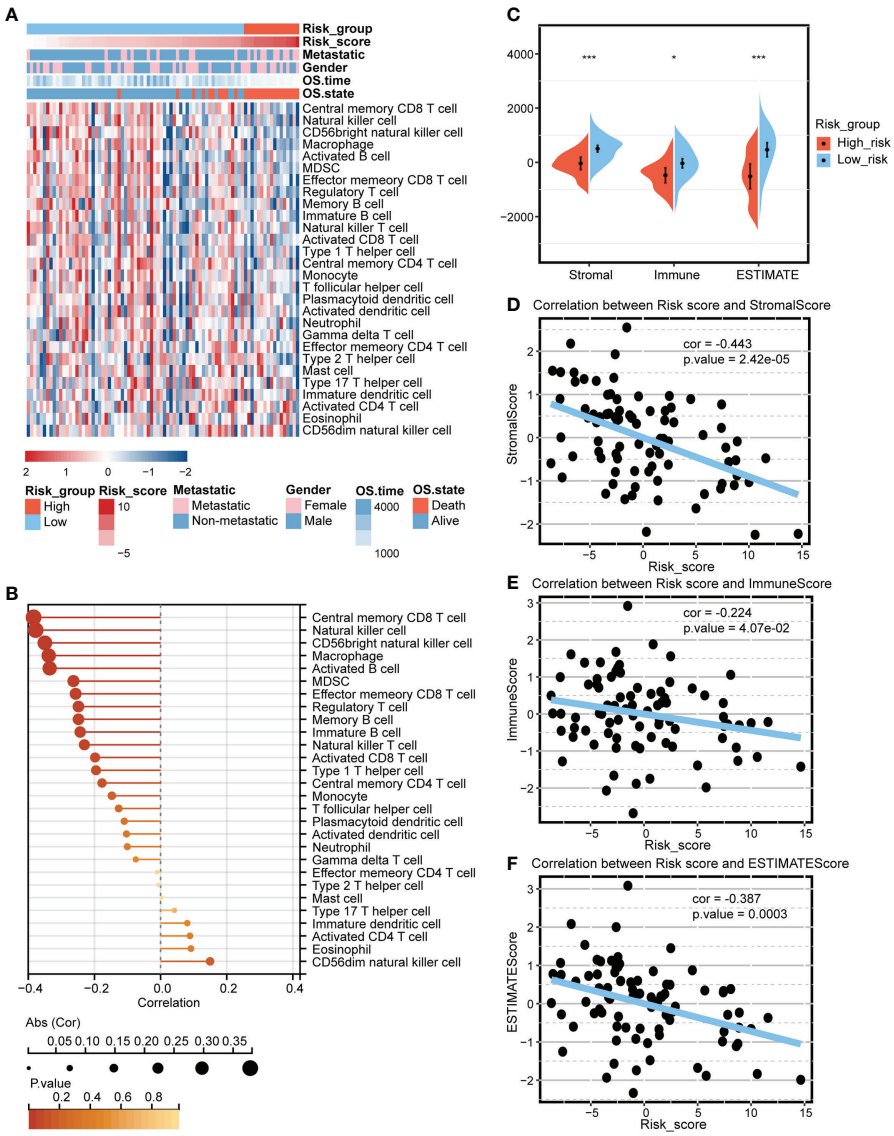


FIGURE 3
Predictive value of the mitochondria-related signature in the immune microenvironment. **(A)** Heatmap depicting the association between the mitochondria-related signature and immune cells. **(B)** Dot plot depicting the association between the mitochondria-related signature and immune cells. **(C)** The different levels of the mitochondria-related signature score regarding microenvironment scores. **(D)** The association between the mitochondria-related signature and stromal score. **(E)** The association between the mitochondria-related signature and immune score. **(F)** The association between the mitochondria-related signature and ESTIMATE score. *p < 0.05; ***p < 0.001.

metabolic process, ossification, and energy derivation by oxidation of organic compounds were highly enriched in the high score group. Immune response-regulating signaling pathway, mononuclear cell proliferation, and positive regulation of T cell activation were highly enriched in the low score group. KEGG enrichment analysis was performed on the DEGs (Figure 5D). Oxidative phosphorylation, chemical carcinogenesis-reactive oxygen species, and glycolysis/gluconeogenesis were highly enriched in the high score group. Ferroptosis, Th1 and Th2 cell differentiation, and natural killer cell mediated cytotoxicity were highly enriched in the low score group. The pseudotime trajectory analysis was performed on the osteosarcoma cells, and five cell states were determined (Figure 5E). As pseudotime increased (Figure 5F), osteosarcoma cells tended to have increased mitochondria-related signature scores (Figure 5G). The DEGs between osteosarcoma cells around branch point 1 were identified and clustered into four types (Figure S3A). GO enrichment analysis was performed on the DEGs in four clusters (Figure S3B-S3E). The expression pattern of the mitochondria-related genes in different cell states is shown in Figure S4.

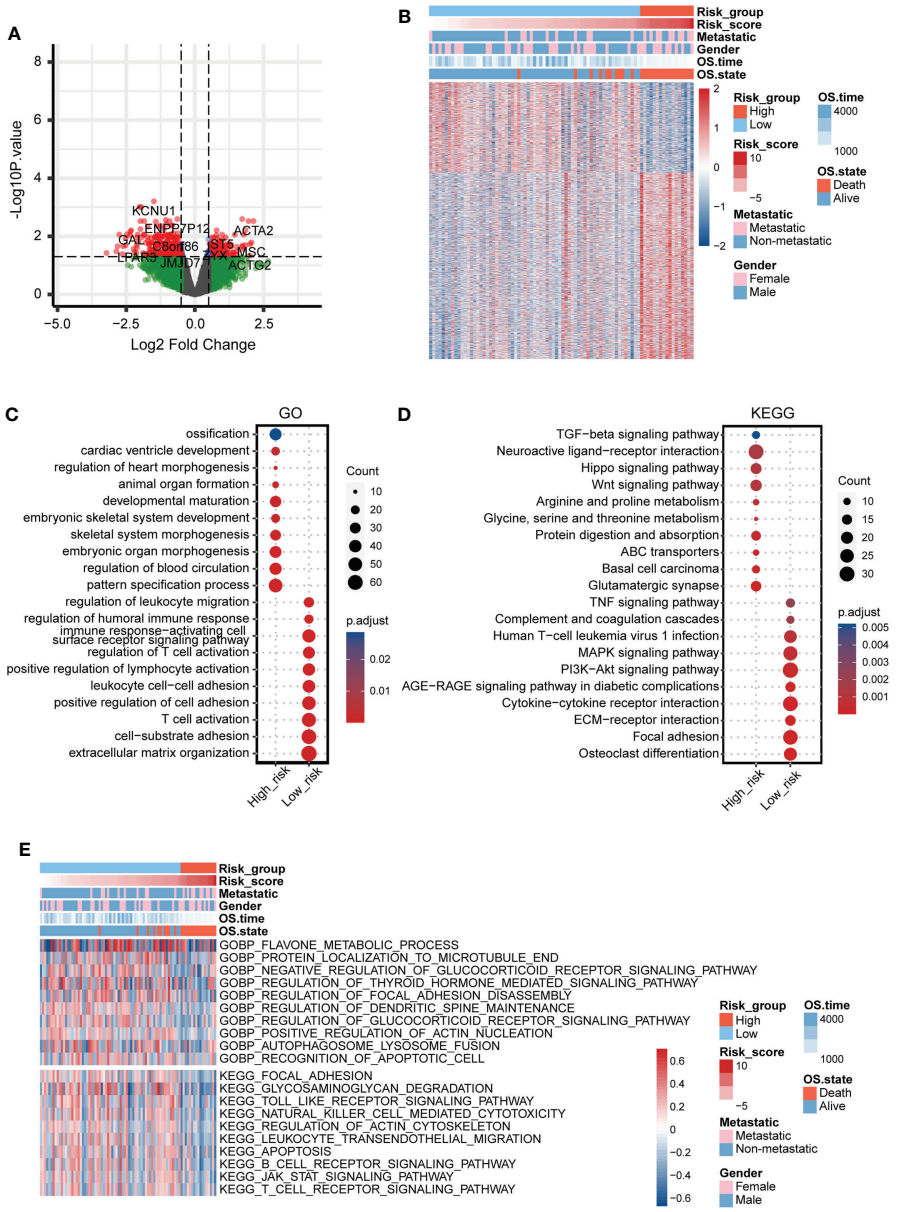


FIGURE 4
Functional annotation of the mitochondria-related signature. **(A)** Volcano plot for the DEGs between the two mitochondria-related signature score groups. **(B)** Heatmap for the DEGs between the two mitochondria-related signature score groups. **(C)** GO enrichment analysis for the DEGs between the two mitochondria-related signature score groups. **(D)** KEGG enrichment analysis for the DEGs between the two mitochondria-related signature score groups. **(E)** GSVA for the DEGs between the two mitochondria-related signature score groups.

Cell communication pattern of the mitochondria-related signature

Different cellular signaling pathways regarding checkpoints between two mitochondria-related signature score groups of osteosarcoma cells and microenvironment cells are shown in Figures 6A, B, in which ITGB2, HAVCR2, and LGALS9 were the most active signaling pathways in osteosarcoma cells with the high mitochondria-related signature score. Different cellular signaling pathways regarding cytokine between two mitochondria-related signature score groups of osteosarcoma cells and microenvironment cells are shown in Figures 6C, D, in which ITGB1 was the most active signaling pathway in osteosarcoma cells with the high mitochondria-related signature score. Different

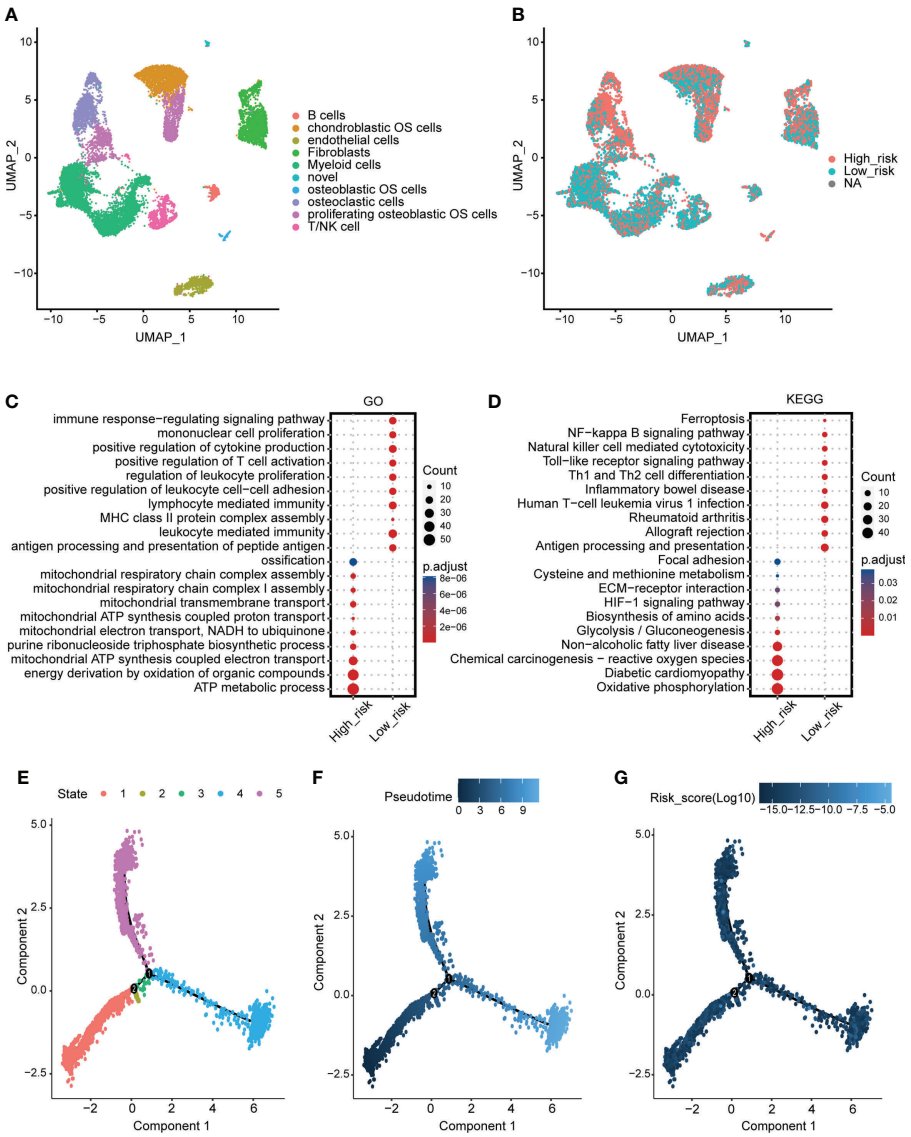


FIGURE 5
scRNA-seq analysis for the mitochondria-related signature. **(A)** The identified cells in the tumor microenvironment of osteosarcoma. **(B)** The levels of the mitochondria-related signature score in identified cells. **(C)** GO enrichment analysis for the DEGs between the two mitochondria-related signature score groups of osteosarcoma cells. **(D)** KEGG enrichment analysis for the DEGs between the two mitochondria-related signature score groups of osteosarcoma cells. **(E)** Different cell states of the pseudotime trajectory analysis on the osteosarcoma cells. **(F)** Pseudotime pattern of the pseudotime trajectory analysis on the osteosarcoma cells. **(G)** The mitochondria-related signature score of the pseudotime trajectory analysis on the osteosarcoma cells.

cellular signaling pathways regarding growth factor between two mitochondria-related signature score groups of osteosarcoma cells and microenvironment cells are shown in Figures 6E, F, in which ITGB2, SDC2, PGF, and TGFB1 were the most active signaling pathways in osteosarcoma cells with the high mitochondria-related signature score. Different cellular signaling pathways regarding other between two mitochondria-related signature score groups of osteosarcoma cells and microenvironment cells are shown in Figures 6G, H, in which CD63, COL1A1, COL1A2, and TIMP1 were the most active signaling pathways in osteosarcoma cells with the high mitochondria-related signature score.

In vitro validation on PCCB

The expression pattern of PCCB in the identified cells in the tumor microenvironment of osteosarcoma is shown in Figures S5A-S5C, in which PCCB was highly expressed by osteosarcoma

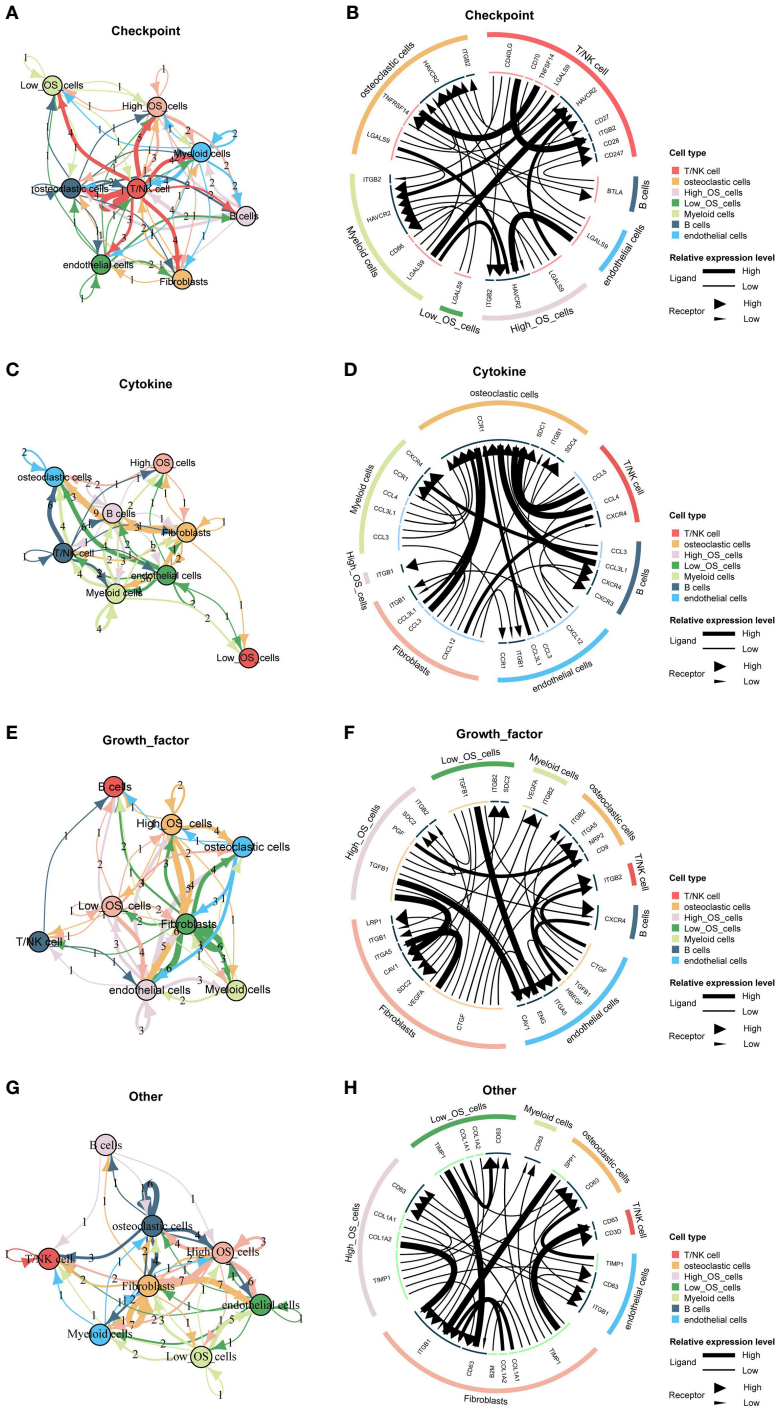


FIGURE 6
Cell communication pattern of the mitochondria-related signature. **(A)** Different cellular signaling pathways regarding checkpoints between two mitochondria-related signature score groups of osteosarcoma cells and microenvironment cells. **(B)** Cell communication pattern regarding checkpoints between two mitochondria-related signature score groups of osteosarcoma cells and microenvironment cells. **(C)** Different cellular signaling pathways regarding cytokine between two mitochondria-related signature score groups of osteosarcoma cells and microenvironment cells. **(D)** Cell communication pattern regarding cytokine between two mitochondria-related signature score groups of osteosarcoma cells and microenvironment cells. **(E)** Different cellular signaling pathways regarding growth factor between two mitochondria-related signature score groups of osteosarcoma cells and microenvironment cells. **(F)** Cell communication pattern regarding growth factor between two mitochondria-related signature score groups of osteosarcoma cells and microenvironment cells. **(G)** Different cellular signaling pathways regarding other between two mitochondria-related signature score groups of osteosarcoma cells and microenvironment cells. **(H)** Cell communication pattern regarding other between two mitochondria-related signature score groups of osteosarcoma cells and microenvironment cells.

cells. GO enrichment analysis revealed that DNA replication, sterol biosynthetic process, and cholesterol biosynthetic process were highly enriched in osteosarcoma patients with high expression of PCCB. In contrast, T cell activation, B cell activation, and lymphocyte proliferation were highly enriched in osteosarcoma patients with low expression of PCCB (Figure S5D). KEGG enrichment analysis revealed that cell cycle, carbon metabolism, and biosynthesis of amino acids were highly enriched in osteosarcoma patients with high expression of PCCB. In contrast, Th1 and Th2 cell differentiation, T cell receptor signaling pathway, and B cell receptor signaling pathway were highly enriched in osteosarcoma patients with low expression of PCCB (Figure S5E). As the most hazardous gene based on the stepwise multivariate Cox regression analysis, the biological function of PCCB in osteosarcoma was explored. RT-qPCR (Figure 7A) and western blot (Figure 7B) results showed that the expression of PCCB was significantly inhibited in three si-PCCB groups compared to NC and si-NC groups in U2OS and MNNG/HOS cells. The si-PCCB, with the most vital ability to suppress the expression of PCCB, was used for the follow-up experiment. EdU assay revealed that the proliferation ability of U2OS (Figure 7D) and MNNG/HOS (Figure 7C) cells was significantly inhibited after transfection with si-PCCB. The IHC results further confirmed that the expression of PCCB was considerably higher in osteosarcoma tumor tissues than in normal tissues (Figure 8).

Immunotherapy prediction of PCCB

The expression of PCCB in responders and non-responders in immunotherapy cohorts is shown in Figure 9A, in which non-responders had higher expression of PCCB in the Dizier cohort 2013 and Riaz cohort 2018 while responders had higher expression of PCCB in the Hugo cohort 2016 and IMvigor210 cohort 2018. Survival analysis was performed on the two groups regarding PCCB expression in immunotherapy cohorts (Figure 9B). PCCB was associated with better survival in the Hugo cohort 2016 and IMvigor210 cohort 2018, while PCCB was associated with worse survival in the Cho cohort 2020 and Kim cohort 2019. PCCB showed potent efficacy in predicting immunotherapy response in ten immunotherapy cohorts (Figure 9C).

Discussion

In the big data era, mining potential diagnostic, prognostic, and predictive markers in cancer based on large-scale bioinformatics analysis has been increasingly important. Many established markers showed great potential in clinical application (17, 18). Mitochondrial dysfunction is known as a hallmark of cancer. Briefly, mitochondrial dysfunction can be caused by mtDNA mutation, oxidative stress, defective mitochondrial electron transport chain, defective mitochondrial TCA cycle enzyme, tumor-promoting signals, etc (19). The result is the change of cell metabolic pathway, the destruction of intracellular REDOX homeostasis, and the generation of apoptosis and drug resistance. Finally, mitochondrial dysfunction would lead to genomic instability, the aging process, and the occurrence and development of cancer (19). Mitochondria dysfunction in CD8+ T cells has also been demonstrated to be an essential contributing factor for cancer development and a potential target for cancer treatment (20). Therefore, an in-depth understanding of mitochondrial dysfunction in cancer is essential for developing novel effective therapeutic strategies. Most studies have focused on the pathogenic molecular mechanisms of individual mitochondria-related genes. Although several studies have comprehensively explored the potential values of mitochondria-related genes in hepatocellular carcinoma, clear cell renal cell carcinoma, etc., a comprehensive evaluation of the mitochondria-related genes in osteosarcoma has never been conducted.

The current study developed a mitochondria-related signature in osteosarcoma patients based on 16 mitochondria-related genes. ACADVL was reported to be associated with the loss of heterozygosity of 17p13 in the pathogenesis of adrenocortical tumors (21). OGDH was a critical tumor promoter in cancer (22). TRMT1 was found to serve as a promising biomarker in clear cell renal cell carcinoma (23). Downregulation of mitochondrial UQCRB was reported to inhibit cancer stem cell-like properties in glioblastoma (24). FDX1 was revealed to impact the prognosis and mediate the metabolism of lung adenocarcinoma (25). EPHX2 could inhibit colon cancer progression by promoting fatty acid degradation (26). STOML2 was reported to potentiate metastasis of hepatocellular carcinoma by promoting PINK1-mediated mitophagy and regulating sensitivity to Lenvatinib (27). Glycolysis gatekeeper PDK1 could reprogram breast cancer stem cells under hypoxia (28). ALDH7A1 knockdown significantly reduces tumor formation in pancreatic ductal adenocarcinoma (29). LACTB could suppress melanoma progression by attenuating PP1A and YAP interaction (30). miR-125b/BAK1 pathway was essential in promoting tumorigenesis and inhibiting apoptosis of cervical cancer (31). MFN2 could suppress cancer progression by inhibiting mTORC2/Akt signaling (32).

As the most hazardous gene in the mitochondria-related signature, PCCB was found to mediate the proliferation of proliferation and migration of osteosarcoma cells. Besides, PCCB was found with significantly higher expression in osteosarcoma tumor tissues than in normal tissues. Therefore, PCCB was a potential oncogene in osteosarcoma.

Despite different clinical factors (age, gender, metastasis), Osteosarcoma patients with high mitochondria-related signature scores presented decreased survival time. Besides, the mitochondria-related signature was associated with tumor metastasis. Therefore, the mitochondria-related signature was a potential prognostic marker in osteosarcoma patients. Besides, osteosarcoma patients with high mitochondria-

10.3389/fonc.2022.1085065

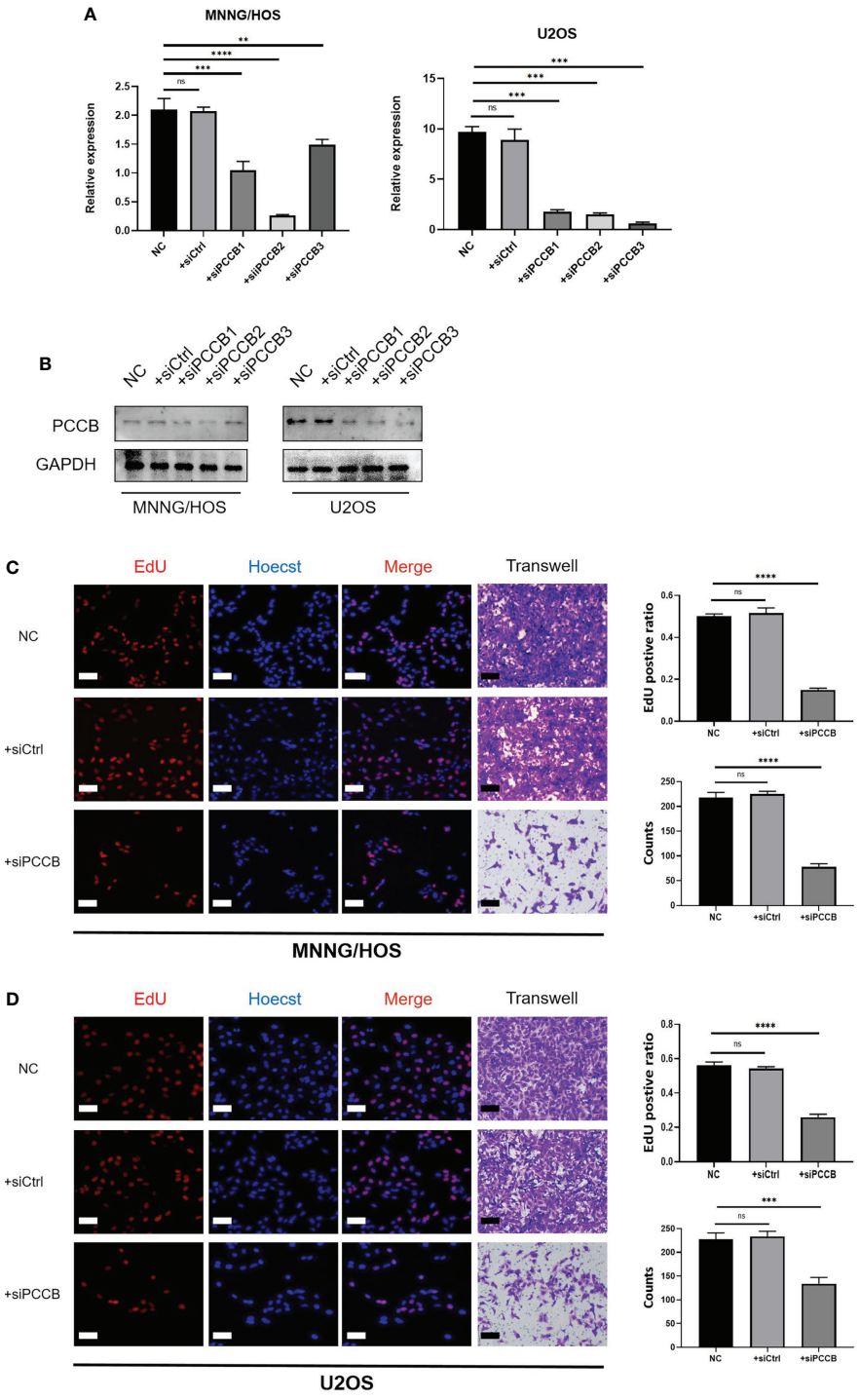


FIGURE 7
In vitro validation on PCCB. **(A)** RT-qPCR results of the expression of PCCB in five groups (NC, si-NC, si-PCCB-1, si-PCCB-2, si-PCCB-3) in two cell lines. **(B)** Western blot results of the expression of PCCB in five groups (NC, si-NC, si-PCCB-1, si-PCCB-2, si-PCCB-3) in two cell lines. **(C)** EdU assay in three groups (NC, si-NC, si-PCCB) in the MNNG/HOS cell line. Statistical analysis was based on mean ± SD. **(D)** EdU assay in three groups (NC, si-NC, si-PCCB) in the U2OS cell line. Statistical analysis was based on mean ± SD. The si-PCCB refers to siRNA transfection of PCCB. The si-NC refers to siRNA transfection of nonspecific control. ns, no significance; **p < 0.01; ***p < 0.001;****p < 0.0001

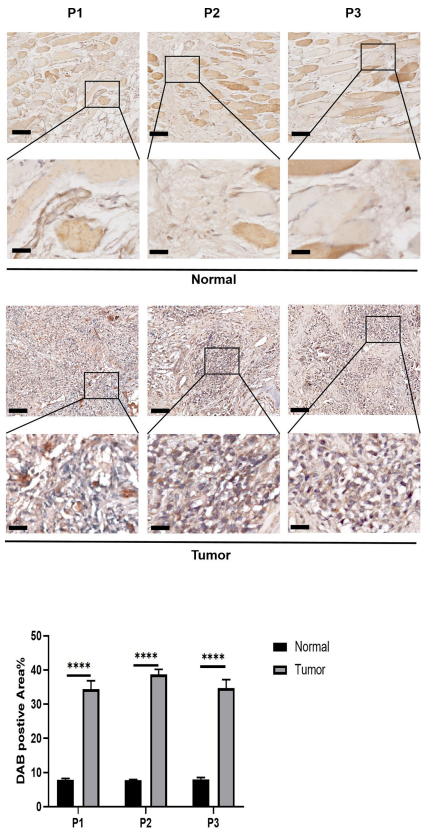


FIGURE 8
In vitro validation on PCCB. IHC results of the expression of PCCB in osteosarcoma tumor tissues and normal tissues. The upper row is 10x images of the sections, and the lower row is 40x images of the sections. Statistical analysis was based on mean ± SD. ****p < 0.0001.

related signature scores were found with a relatively immune cold microenvironment, indicating the un-suppressed malignancy of the tumor. The functional annotation of the mitochondria-related signature further confirmed that the tumorigenic pathways were more active in osteosarcoma patients with high mitochondria-related signature scores.

In contrast, the immunogenic pathways were more involved in osteosarcoma patients with low mitochondria-related signature scores. The potential value of the mitochondria-related signature in predicting chemotherapy agents was also confirmed. 24 drugs were negatively associated with the mitochondria-related signature, including apoptosis regulation inhibitor AZD5991, protein stability and degradation inhibitor ML323, and kinases inhibitor BMS-345541. In addition, six drugs were positively associated with the mitochondria-related signature, including ERK MAPK signaling inhibitor Refametinib, RTK signaling inhibitor NVP-TAE684, and kinases inhibitor A-770041.

At the scRNA-seq level, osteosarcoma cells gradually evolved into tumors with high mitochondria-related signature scores. The functional annotation of the mitochondria-related signature also confirmed the active tumorigenic pathways and inactive immunogenic pathways in osteosarcoma cells with high mitochondria-related signature scores. The tumor microenvironment has already been proven to essentially influence the proliferation, migration, and invasion of cancer (33, 34). Checkpoints (ITGB2, HAVCR2, and LGALS9), cytokine (ITGB1), growth factor (ITGB2, SDC2, PGF, and TGFB1), and other (CD63, COL1A1, COL1A2, and TIMP1) were the most active signaling pathways involved in the cell communication between osteosarcoma cells with a high mitochondria-related signature score and microenvironment cells, indicating the potential immune evasion and tumor progression in osteosarcoma cells with a high mitochondria-related signature score.

Conclusion

Taken together, a mitochondria-related signature was developed in osteosarcoma with robust predictive values in the immune microenvironment, chemotherapy agents, and prognosis.

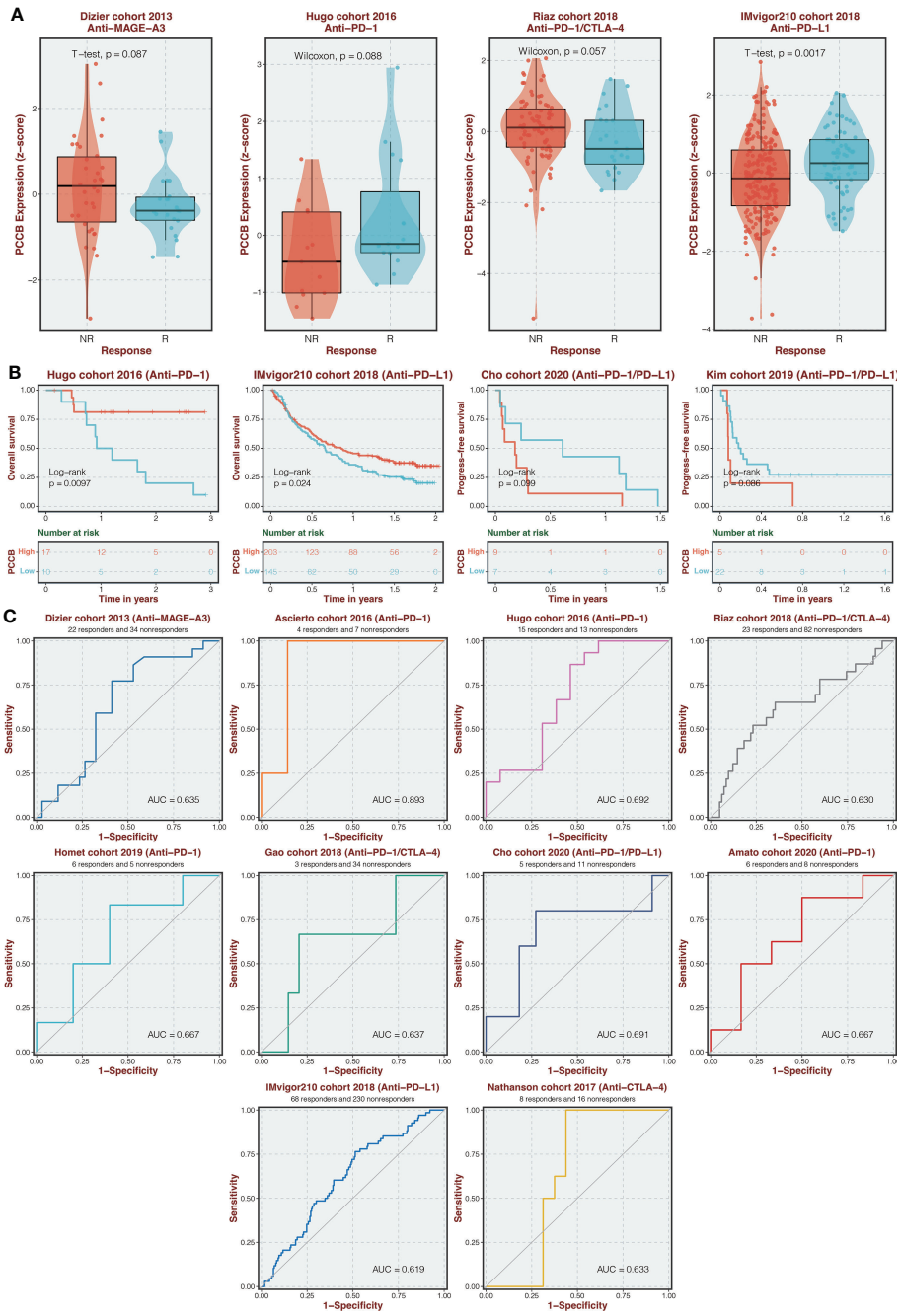


FIGURE 9
Immunotherapy prediction of PCCB. **(A)** The expression of PCCB in responders and non-responders in immunotherapy cohorts. **(B)** Survival analysis was performed on the two groups regarding PCCB expression in immunotherapy cohorts. **(C)** The ROC curve of PCCB in predicting immunotherapy response in immunotherapy cohorts.

The potential clinical application of the mitochondria-related signature is expected to be further validated by real-world cohorts. PCCB was a potential oncogene in osteosarcoma, and the related complex regulatory mechanisms remain to be further explored.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: TARGET-OS(https://xenabrowser.net/) and GEO (https://www.ncbi.nlm.nih.gov/geo).

Ethics statement

The studies involving human participants were reviewed and approved by the institutional review board (IRB) of the Third Xiangya Hospital, Central South University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

LZ conceived and performed most of the experiments. JH wrote the manuscript. YS and YY collected and analyzed the data. XC provided experimental advice and supervised the study; SW provided funding support. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (82072501), Science and Technology Innovation Leading Plan of High Tech Industry in Hunan Province (2020SK2011).

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/fonc.2022.1085065/full#supplementary-material

References

1. Ritter J, Bielack SS. Osteosarcoma. *Ann Oncol* (2010) 21 Suppl 7:vii320–325. doi: 10.1093/annonc/mdq276

2. Kansara M, Teng MW, Smyth MJ, Thomas DM. Translational biology of osteosarcoma. *Nat Rev Cancer* (2014) 14(11):722–35. doi: 10.1038/nrc3838

3. Jafari F, Javdansirat S, Sanaie S, Naseri A, Shamekh A, Rostamzadeh D, et al. Osteosarcoma: A comprehensive review of management and treatment strategies. *Ann Diagn Pathol* (2020) 49:151654. doi: 10.1016/j.anndiagpath.2020.151654

4. Chen C, Xie L, Ren T, Huang Y, Xu J, Guo W. Immunotherapy for osteosarcoma: Fundamental mechanism, rationale, and recent breakthroughs. *Cancer Lett* (2021) 500:1–10. doi: 10.1016/j.canlet.2020.12.024

5. Mehta MM, Weinberg SE, Chandel NS. Mitochondrial control of immunity: Beyond ATP. *Nat Rev Immunol* (2017) 17(10):608–20. doi: 10.1038/nri.2017.66

6. Rambold AS, Pearce EL. Mitochondrial dynamics at the interface of immune cell metabolism and function. *Trends Immunol* (2018) 39(1):6–18. doi: 10.1016/j.it.2017.08.006

7. Sabharwal SS, Schumacker PT. Mitochondrial ROS in cancer: Initiators, amplifiers or an achilles' heel? *Nat Rev Cancer* (2014) 14(11):709–21. doi: 10.1038/nrc3803

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

9. Czabotar PE, Lessene G, Strasser A, Adams JM. Control of apoptosis by the BCL-2 protein family: Implications for physiology and therapy. *Nat Rev Mol Cell Biol* (2014) 15(1):49–63. doi: 10.1038/nrm3722

10. Morishita M, Kawamoto T, Hara H, Onishi Y, Ueha T, Minoda M, et al. AICAR induces mitochondrial apoptosis in human osteosarcoma cells through an AMPK-dependent pathway. *Int J Oncol* (2017) 50(1):23–30. doi: 10.3892/ijo.2016.3775

11. Toki S, Yoshimaru T, Matsushita Y, Aihara H, Ono M, Tsuneyama K, et al. The survival and proliferation of osteosarcoma cells are dependent on the mitochondrial BIG3-PHB2 complex formation. *Cancer Sci* (2021) 112(10):4208–19. doi: 10.1111/cas.15099

12. Liu Z, Wang H, Hu C, Wu C, Wang J, Hu F, et al. Targeting autophagy enhances atezolizumab-induced mitochondria-related apoptosis in osteosarcoma. *Cell Death Dis* (2021) 12(2):164. doi: 10.1038/s41419-021-03449-6

13. Lai HT, Naumova N, Marchais A, Gaspar N, Geoerger B, Brenner C. Insight into the interplay between mitochondria-regulated cell death and energetic metabolism in osteosarcoma. *Front Cell Dev Biol* (2022) 10:948097. doi: 10.3389/fcell.2022.948097

14. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep* (2017) 18(1):248–62. doi: 10.1016/j.celrep.2016.12.019

15. 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

16. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* (2017) 14(10):979–82. doi: 10.1038/nmeth.4402

17. Zhang H, Zhang N, Wu W, Zhou R, Li S, Wang Z, et al. Machine learning-based tumor-infiltrating immune cell-associated lncRNAs for predicting prognosis and immunotherapy response in patients with glioblastoma. *Brief Bioinform* (2022) 23(6):bbac386. doi: 10.1093/bib/bbac386

18. Zhang N, Zhang H, Wu W, Zhou R, Li S, Wang Z, et al. Machine learning-based identification of tumor-infiltrating immune cell-associated lncRNAs for

improving outcomes and immunotherapy responses in patients with low-grade glioma. *Theranostics* (2022) 12(13):5931–48. doi: 10.7150/thno.74281

19. Luo Y, Ma J, Lu W. The significance of mitochondrial dysfunction in cancer. *Int J Mol Sci* (2020) 21(16):5598. doi: 10.3390/ijms21165598

20. Zhang L, Zhang W, Li Z, Lin S, Zheng T, Hao B, et al. Mitochondria dysfunction in CD8+ T cells as an important contributing factor for cancer development and a potential target for cancer treatment: A review. *J Exp Clin Cancer Res* (2022) 41(1):227. doi: 10.1186/s13046-022-02439-6

21. Soon PS, Libe R, Benn DE, Gill A, Shaw J, Sywak MS, et al. Loss of heterozygosity of 17p13, with possible involvement of ACADVL and ALOX15B, in the pathogenesis of adrenocortical tumors. *Ann Surg* (2008) 247(1):157–64. doi: 10.1097/SLA.0b013e318153ff55

22. Allen EL, Ulanet DB, Pirman D, Mahoney CE, Coco J, Si Y, et al. Differential aspartate usage identifies a subset of cancer cells particularly dependent on OGDH. *Cell Rep* (2016) 17(3):876–90. doi: 10.1016/j.celrep.2016.09.052

23. Li C, Yao Y, Long D, Lin X. KDELC1 and TRMT1 serve as prognosis-related SARS-CoV-2 proteins binding human mRNAs and promising biomarkers in clear cell renal cell carcinoma. *Int J Gen Med* (2021) 14:2475–90. doi: 10.2147/IJGM.S312416

24. Jung N, Kwon HJ, Jung HJ. Downregulation of mitochondrial UQCRB inhibits cancer stem cell-like properties in glioblastoma. *Int J Oncol* (2018) 52 (1):241–51. doi: 10.3892/ijo.2017.4191

25. Zhang Z, Ma Y, Guo X, Du Y, Zhu Q, Wang X, et al. FDX1 can impact the prognosis and mediate the metabolism of lung adenocarcinoma. *Front Pharmacol* (2021) 12:749134. doi: 10.3389/fphar.2021.749134

26. Zhou Y, Li X, Guan A, Zhou H, Zhu Y, Wang R, et al. EPHX2 inhibits colon cancer progression by promoting fatty acid degradation. *Front Oncol* (2022) 12:870721. doi: 10.3389/fonc.2022.870721

27. Zheng Y, Huang C, Lu L, Yu K, Zhao J, Chen M, et al. STOML2 potentiates metastasis of hepatocellular carcinoma by promoting PINK1-mediated mitophagy and regulates sensitivity to lenvatinib. *J Hematol Oncol* (2021) 14(1):16. doi: 10.1186/s13045-020-01029-3

28. Peng F, Wang JH, Fan WJ, Meng YT, Li MM, Li TT, et al. Glycolysis gatekeeper PDK1 reprograms breast cancer stem cells under hypoxia. *Oncogene* (2018) 37(8):1062–74. doi: 10.1038/onc.2017.368

29. Lee JS, Lee H, Woo SM, Jang H, Jeon Y, Kim HY, et al. Overall survival of pancreatic ductal adenocarcinoma is doubled by Aldh7a1 deletion in the KPC mouse. *Theranostics* (2021) 11(7):3472–88. doi: 10.7150/thno.53935

30. Ma Y, Wang L, He F, Yang J, Ding Y, Ge S, et al. LACTB suppresses melanoma progression by attenuating PP1A and YAP interaction. *Cancer Lett* (2021) 506:67–82. doi: 10.1016/j.canlet.2021.02.022

31. Wang YD, Cai N, Wu XL, Cao HZ, Xie LL, Zheng PS. OCT4 promotes tumorigenesis and inhibits apoptosis of cervical cancer cells by miR-125b/BAK1 pathway. *Cell Death Dis* (2013) 4:e760. doi: 10.1038/cddis.2013.272

32. Xu K, Chen G, Li X, Wu X, Chang Z, Xu J, et al. MFN2 suppresses cancer progression through inhibition of mTORC2/Akt signaling. *Sci Rep* (2017) 7:41718. doi: 10.1038/srep41718

33. Hinshaw DC, Shevde LA. The tumor microenvironment innately modulates cancer progression. *Cancer Res* (2019) 79(18):4557–66. doi: 10.1158/0008-5472.CAN-18-3962

34. Wang Y, Zhang H, Liu C, Wang Z, Wu W, Zhang N, et al. Immune checkpoint modulators in cancer immunotherapy: Recent advances and emerging concepts. *J Hematol Oncol* (2022) 15(1):111. doi: 10.1186/s13045-022-01325-0

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 14 December 2022
DOI 10.3389/fimmu.2022.1080403

OPEN ACCESS

EDITED BY
Hao Zhang,
Second Affiliated Hospital, Chongqing Medical University, China

REVIEWED BY
Longxin Lin,
Hokkaido University, Japan
Qiang Peng,
The Chinese University of Hong Kong, China

*CORRESPONDENCE
Cheng Zhang
doctorcheng77@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Immunology

RECEIVED 26 October 2022
ACCEPTED 30 November 2022
PUBLISHED 14 December 2022

CITATION
Wang T, Jian W, Xue W, Meng Y, Xia Z, Li Q, Xu S, Dong Y, Mao A and Zhang C (2022) Integration analysis identifies MYBL1 as a novel immunotherapy biomarker affecting the immune microenvironment in clear cell renal cell carcinoma: Evidence based on machine learning and experiments. *Front. Immunol.* 13:1080403. doi: 10.3389/fimmu.2022.1080403

Integration analysis identifies MYBL1 as a novel immunotherapy biomarker affecting the immune microenvironment in clear cell renal cell carcinoma: Evidence based on machine learning and experiments

Tengda Wang[1†], Wengang Jian[1†], Wei Xue[1], Yuyang Meng[1], Zhinan Xia[1], Qinchen Li[2], Shenhao Xu[2], Yu Dong[2], Anli Mao[2] and Cheng Zhang[1,2*]

[1]Urology Surgery Department, The First Affiliated Hospital of Harbin Medical University, Harbin, China, [2]The Fourth Affiliated Hospital, International Institutes of Medicine, Zhejiang University School of Medicine, Yiwu, Zhejiang, China

Background: Previous studies have identified MYBL1 as a cancer-promoting molecule in numerous types of cancer. Nevertheless, the role of MYBL in renal cancer remains unclear.

Methods: Genomic and clinical data of clear cell renal cell carcinoma (ccRCC) was get from the Cancer Genome Atlas (TCGA) database. CCK8, colony formation, and 5-ethynyl-2'-deoxyuridine assay were utilized to evaluate the performance of cell proliferation. Cell apoptosis was detected using the flow cytometric analysis. The protein level of MYBL1 in different tissues was evaluated using immunohistochemistry. A machine learning algorithm was utilized to identify the prognosis signature based on MYBL1-derived molecules.

Results: Here, we comprehensively investigated the role of MYBL1 in ccRCC. Here, we noticed a higher level of MYBL1 in ccRCC patients in both RNA and protein levels. Further analysis showed that MYBL1 was correlated with progressive clinical characteristics and worse prognosis performance. Biological enrichment analysis showed that MYBL1 can activate multiple oncogenic pathways in ccRCC. Moreover, we found that MYBL1 can remodel the immune microenvironment of ccRCC and affect the immunotherapy response. *In vitro* and *in vivo* assays indicated that MYBL1 was upregulated in

ccRCC cells and can promote cellular malignant behaviors of ccRCC. Ultimately, an machine learning algorithm – LASSO logistics regression was utilized to identify a prognosis signature based on the MYBL1-derived molecules, which showed satisfactory prediction ability on patient prognosis in both training and validation cohorts.

Conclusions: Our result indicated that MYBL1 is a novel biomarker of ccRCC, which can remodel the tumor microenvironment, affect immunotherapy response and guide precision medicine in ccRCC.

KEYWORDS

MYBL1, immunotherapy, machine learning, prognosis, ccRCC

Introduction

Renal cancer represents a frequent malignancy globally with approximately 4 and 1.9 million new cases and death each year, respectively (1). Among all cases of renal cancer, clear cell renal cell carcinoma (ccRCC) is the leading pathological subtype (2). As a multifactorial multi-factorial disorder, ccRCC is associated with obesity, smoking, dietary habits, environmental exposure, genetic susceptibility, and so on (3). For ccRCC patients at the local stage, surgical resection combined with adjuvant agents is still the mainstay therapy choice (4). Nonetheless, the prognosis is still unsatisfactory for patients in the advanced stage or with distant metastasis (4). Consequently, identifying a novel target for ccRCC diagnosis and treatment is meaningful for clinical application.

The rapid development of bioinformatics analysis has brought researchers great convenience in deeply understanding specific molecules in diseases (5). For instance, Wei et al. revealed that MX2 might be a biomarker indicating sunitinib resistance (6). The MYBL1 has been implicated in multiple diseases. Zhu et al. indicated that MYBL1 was tightly associated with higher endothelial vessel density by inducing the transcriptional activation of ANGPT2, further affecting sorafenib resistance in liver cancer (7). Guo et al. indicated that the O-GlcNAc can regulate MYBL1 expression in an epigenetic modification manner, leading to an aberrant cancer stem cell compartment and cancer progression (8). Brayer et al. demonstrated that the fusion of MYB and MYBL1 contributes to the oncogenic pathway in salivary gland adenoid cystic carcinoma (9). Ramkissoon et al. indicated that MYBL1 also acted as an oncogene factor in glioma (10). Nikolaus et al. revealed that the MYBL1 might be a trigger for autoimmune encephalitis, indicating its role in the disease immune (11). However, there is no previous study focused on MYBL1 in ccRCC.

In this study, we comprehensively the effects MYBL1 produces in ccRCC through bioinformatics analysis and experiments. We noticed a higher level of MYBL1 in ccRCC patients in both RNA and protein levels. Further analysis showed that MYBL1 was correlated with progressive clinical characteristics and worse prognosis performance. Biological enrichment analysis was conducted to identify the biological role of MYBL1 in ccRCC. Moreover, we noticed that MYBL1 can remodel the immune microenvironment of ccRCC and affect the immunotherapy response. *In vitro* and *in vivo* assays indicated that MYBL1 was upregulated in ccRCC cells and can promote cellular malignant behaviors of ccRCC. Meanwhile, the prognosis signature based on the MYBL1-derived molecules showed great prediction ability on patient prognosis.

Methods

Public data collection

The open-accessed data on transcriptional profiles and clinical features were all download from The Cancer Genome Atlas (TCGA) database, the KIRC project and the Arrayexpress database, E-MTAB-1980 project. Initial expression profile of ccRCC patients (STAR-Counts form) was collated to a combined matrix (TPM" form) for subsequent analysis. We extracted the survival and clinical information based on the bcr-xml file in TCGA-KIRC. Pan-cancer expression data was obtained from the UCSC Xena website (https://xenabrowser.net/). The limma and affay packages in the R environment were utilized for data preprocessing, consisting of background correction, probe ID annotation, missing value completion and normalization. Extra gene expression information of normal tissue was get from the GETx database. The open-accessed immunohistochemical images of renal cancer were get from

The Human Protein Atlas (HPA) database. Baseline features of ccRCC patients enrolled in the analysis were presented in Tables 1, 2. The limma package in the R environment was utilized for DEGs analysis under specific parameters.

Biological enrichment

The clueGO app built in Cytoscape software provides the biological terms for the input molecules identified, as well as an intuitive representation (12). Clusterprofiler in R was used to enrich Gene Ontology (GO) terms (13). Gene Set Enrichment Analysis (GSEA) was utilized to investigate the potential biological differences between two selected groups based on set reference gene sets, including Hallmark gene set (14). The single sample GSEA (ssGSEA) was utilized to quantify the enrichment score based on a specific reference file (15).

Immune-related analysis

The XCELL, MCPCOUNTER, CIBERSORT, TIMER, EPIC and QUANTISEQ algorithms were utilized for immune microenvironment quantification (16–19). The Immunophenoscore (IPS), an machine learning algorithm from The Cancer Immunome Database (TCIA), was utilized to quantify the IPS score of ccRCC patients based on their transcriptional profile, indicating the response to immunotherapy (20). Meanwhile, the Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was also utilized to assess patients immunotherapy response (21).

Genomic characterization

Two important immunotherapy markers, tumor mutation burden (TMB) and microsatellite instability (MSI) score were get from the TCGA database. The gene mutation characteristic of MYBL1 in the TCGA database was obtained and visualized based on an online website, https://www.home-for-researchers.com/. The score of the tumor stemness index (mRNAsi) in the TCGA-KIRC project was obtained from the previous study, which was completed using the OCLR machine-learning algorithm (22).

Nomogram

With the survival and rms packages, a nomogram plot was created based on multiple factors. The calibration and Decision Curve Analysis (DCA) plot were utilized to evaluate nomogram performance.

TABLE 1 Baseline information of patients in TCGA-KIRC.

Clinical features		Numbers	Percentage (%)
Age	<=65	352	65.5
	>65	185	34.5
Gender	Female	191	35.6
	Male	346	64.4
Grade	G1	14	2.6
	G2	230	42.8
	G3	207	38.5
	G4	78	14.5
	Unknown	8	1.5
Stage	Stage I	269	50.1
	Stage II	57	10.6
	Stage III	125	23.3
	Stage IV	83	15.5
	Unknown	3	0.6
T-stage	T1	275	51.2
	T2	69	12.8
	T3	182	33.9
	T4	11	2.0
M-stage	M0	426	79.3
	M1	79	14.7
	Unknown	32	6.0
N-stage	N0	240	44.7
	N1	17	3.2
	Unknown	280	52.1

TABLE 2 Baseline information of patients in E-MTAB-1980.

Clinical features		Numbers	Percentage (%)
Age	<=65	145	60.4
	>65	95	39.6
Gender	Female	56	23.3
	Male	184	76.7
Grade	G1	42	17.5
	G2	141	58.8
	G3	49	20.4
	G4	6	2.5
	Unknown	2	0.8
T-stage	T1	187	77.9
	T2	18	7.5
	T3	33	13.8
	T4	2	0.8
M-stage	M0	215	89.6
	M1	25	10.4
N-stage	N0	238	99.2
	N1	2	0.8

Identification of prognosis signature based on machine learning

For TCGA database, all patients were randomly divided into training and internal validation cohorts according to the ratio of 1 to 1. A univariate Cox regression analysis was conducted on the DEGs identified between high and low MYBL1 expression to identify the molecules significantly correlated with patient survival. Then, an machine learning algorithm, LASSO logistic regression, was utilized for variable optimization (23). The optimized variables were then set as the input file for multivariate Cox regression analysis. Ultimately, a prognosis signature was identified with the formula of "Risk score = Expression of A * Coef A + Expression of B * Coef B + … + Expression of X * Coef X". The E-MTAB-1980 project was used as the external validation cohort.

Cell culture

Four renal cancer cell lines (A498, ACHN, 786-O and OS-RC-2) and normal cell line HK-2 were laboratory stocks. All cells were routinely cultured using the 10% heat-inactivated fetal bovine serum (37°C with 5% CO_2) and passaged for three days a time (24).

RNA isolation and quantitative RT-PCR

Total RNA extraction and steps for qRT-PCR were conducted following our previous study (24). The sequence of primers was as follows: MYBL1, forward primer, 5'-TAG CACTCCACCAGCCATCCTC-3', reverse primer, 5'-ACCAC CATCGTTCAATGAGCCATC-3'.

Retroviral infection, and transfection

We purchased HBLV-h-MYBL1 shRNA#1-PURO, HBLV-h-MYBL1 shRNA#2-PRUO and HBLV-h-MYBL1-Ctl-PURO from Hanbio. Cell transfection was performed by jetPRIME (Polyplus, NY, USA) referring to the manufacturer's protocol. We constructed the stably lentivirus-transinfected cells with puromycin (MCE.NJ) to collect the MYBL1 stable-knockdown cells. The sequence used for shRNA were as follow: MYBL1 sh#1, top strand, 5'-GATCCGGACGAGGATGATAAA TTACTCGAGTAATTTATCATCCTCGTCCTTTTTTG-3', b ottom strand, 5'-AATTCAAAAAAGGACGAGGATGAT AAATTACTCGAGTAATTTATCATCCTCGTCCG-3'; sh#2, top strand, 5'-GATCCGCCATGGAATGCCAATT TACTCGAGTAAATTGGCATTCCATGGCTTTTTTG-3', bottom strand, 5'-AATTCAAAAAAGCCATGGAATGCCAA TTTACTCGAGTAAATTGGCATTCCATGGCG-3'.

CCK-8 assay

Steps for **CCK-8** were completed following our previous study based on the sh-MYBL1 and control cells (24).

Colony formation assay

Steps for colony formation assay were completed following our previous study based on the sh-MYBL1 and control cells (24).

5-ethynyl-2'-deoxyuridine assay

Steps for EdU assay were completed following our previous study based on the sh-MYBL1 and control cells (24).

Cell apoptosis assays

Steps for cell apoptosis detection were completed following our previous study based on the sh-MYBL1 and control cells (24). The results were analyzed through FlowJo 6.2 software.

Xenograft models

Steps for xenograft model assay were completed following our previous study based on the sh-MYBL1 and control cells (Approximately 6×10^6 786-O cells, MYBL1 sh#Ctrl and sh#1) (24). Animal procedures were performed in line with the Association for Assessment and Accreditation of Laboratory Animal Care and approved by the Animal Care and Use Committee of the First Affiliated Hospital of Harbin Medical University.

Patient and clinical samples

The study was admitted by the First Affiliated Hospital of Harbin Medical University. ccRCC and adjacent tissue samples were obtained from patients who were aware of the purpose of the study and signed informed consent at the Medical Ethics Committee of First Affiliated Hospital of Harbin Medical University. After the operations of radical nephrectomy, Half of the samples were frozen in liquid nitrogen, and half were embedded in paraffin after being fixed with 4%

paraformaldehyde overnight at room temperature until further use.

Immunohistochemistry

The clinical samples and tumor tissues embedded in paraffin were cut into 5-μm-thick sections. The sections were blocked by 5% goat serum and incubated overnight at 4°C with antibodies against MYBL1 (1:400, Boster, Wuhan, China). We used the DAB (Beyotime, Shanghai, China) system for detection. We chose the three fields in each section and took photos and analyzed the images by ImageJ software.

Statistical analysis

R and GraphPad Prism 8 software were utilized for all statistical analysis. The statistical threshold of the P value in comparison was 0.05. All the experiments were repeated at least three times.

Results

Pan-cancer analysis of MYBL1 and its clinical role in ccRCC

The flow chart of the whole study was shown in Figure S1. The expression landscape of MYBL1 was illustrated in Figure 1A, in which MYBL1 showed an abnormal expression pattern in most of the cancers, indicating its important role in cancer development. According to the GTEx and TCGA data, MYBL1 all showed a higher expression level in ccRCC tumor tissue compared with the control normal tissue (Figures 1B, C). Moreover, based on the immunohistochemical result from the HPA database, a higher protein level of MYBL1 in renal cancer tissue was observed (Figure 1D). Furthermore, we tried to investigate the prognosis role of MYBL1 in ccRCC. Results indicated that MYBL1 might be correlated with worse prognosis performance of ccRCC patients, including overall survival (OS), disease-free survival (DSS) and progression-free survival (PFI) in both TCGA and E-MTAB-1980 cohorts (Figures 1E, F). Then, we explored the clinical correlation of MYBL1 in ccRCC patients. No remarkable difference in MYBL1 expression was found in patients with different T-stage and grades (Figures 1G, H). However, a significantly higher level of MYBL1 was noticed in patients with worse M- and N-stage, indicating its promoting effect in cancer metastasis (Figures 1I, J). Based on univariate and multivariate analyses, MYBL1 was an independent risk factor for ccRCC (Figures 1K, L).

MYBL1 exerts a wide biological regulatory effect in ccRCC

A total of 154 downregulated and 136 upregulated DEGs were identified in patients with high and low MYBL1 expression (Figure 2A). These DEGs were involved in L-alpha-amino acid transmembrane transport, inorganic anion transport, monovalent inorganic cation homeostasis, retinoid metabolic process, active ion transmembrane transporter activity, metanephros development, embryonic pattern specification, excretion, negative regulation of chemotaxis, vasodilation, negative regulation of coagulation based on clueGO analysis (Figure 2B). Moreover, the results of the ssGSEA algorithm indicated that MYBL1 was positively correlated with most immune terms, especially base excision repair, fanconi anemia pathway, and homologous recombination; MYBL1 was negatively correlated with most metabolism terms (Figure 2C). GO analysis indicated that for the Biological Process (BP), the top enriched terms were excretion, metanephros development, chloride transmembrane transport (Figure 2D). For the Cellular Component (CC) and Molecular Function (MF), the terms were mainly enriched in the transport complex (Figures 2E, F). The GSEA analysis of Hallmark gene set was mainly enriched in the terms of the inflammatory response, G2M checkpoint and E2F targets (Figures 2G–I).

MYBL1 can remodel the ccRCC immune microenvironment

Based on the XCELL, MCPCOUNTER, CIBERSORT, TIMER, EPIC and QUANTISEQ algorithms, we quantified the immune microenvironment of ccRCC samples. A different immune infiltration pattern was observed in patients with high and low MYBL1 expression (Figure 3A). Immune correlation indicated that MYBL1 can increase Tregs, M2 macrophages, neutrophils, B cells, monocytes, CD8+ T cells, yet decrease endothelial cells level in the ccRCC microenvironment (Figures 3B–I). Moreover, we found that MYBL1 was positively correlated with immune score, stromal score and estimate score (Figures 3J–L). Interestingly, we found that the key immune checkpoints PD-1, CTLA4, PD-L1 and PD-L2 all present a high level in patients with higher MYBL1 expression, indicating that MYBL1 might affect the immunotherapy response of ccRCC patients (Figures 3M–P).

Role of MYBL in ccRCC genomic characteristics

TMB and MSI are important markers for cancer immunotherapy and can also indicate genomic instability. A

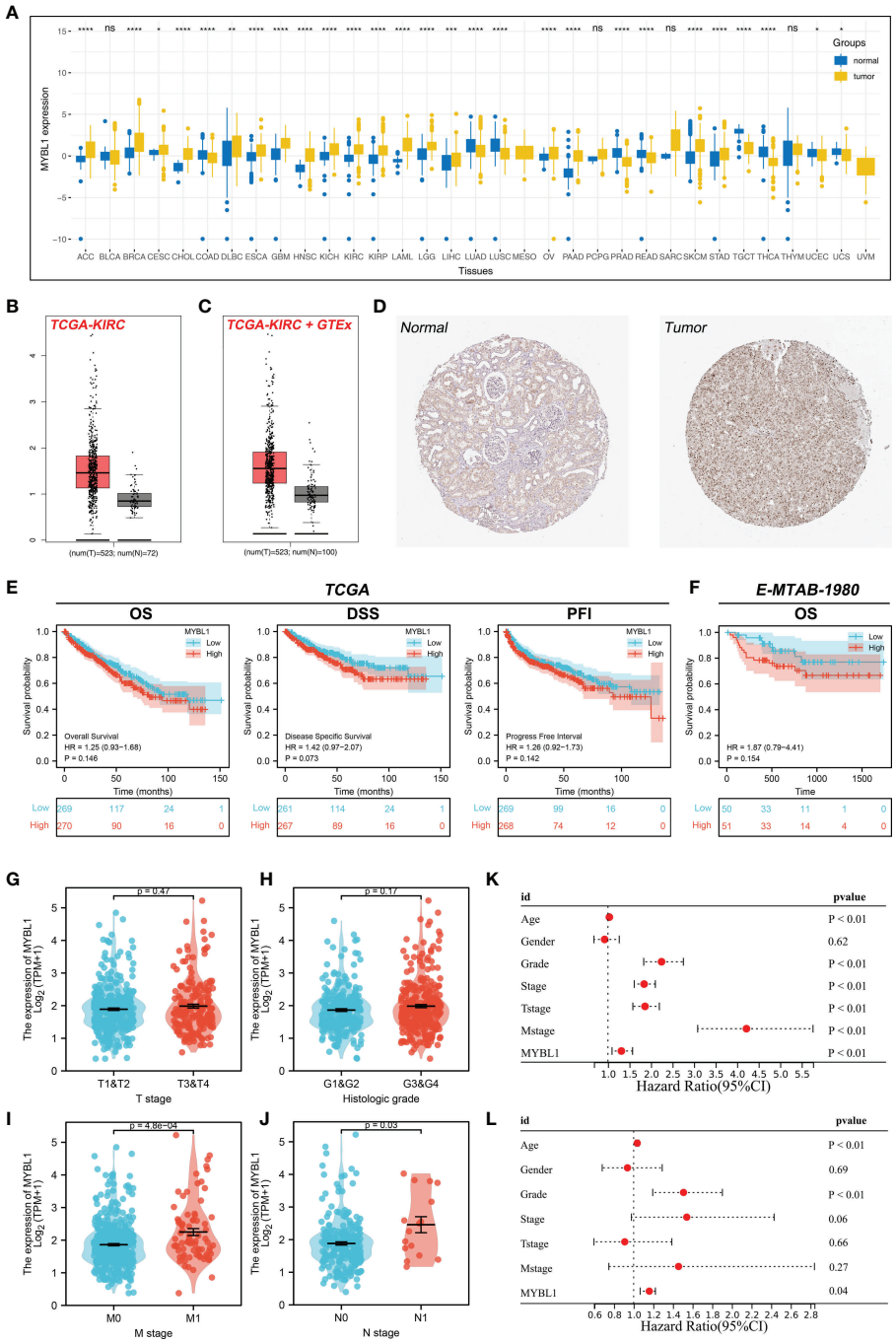


FIGURE 1
Expression pattern and clinical role of MYBL in ccRCC. **(A)** The landscape of MYBL1 in pan-cancer; **(B)** Expression of MYBL1 in ccRCC tumor and normal tissue based on the TCGA database; **(C)** Expression of MYBL1 in ccRCC tumor and normal tissue based on the TCGA+GTEx database; **(D)** Representative immunohistochemical images of MYBL1 in HPA database; **(E)** Prognosis performance of MYBL1 in TCGA database; **(F)** Prognosis performance of MYBL1 in E-MTAB-1980 database; **(G-J)** Clinical correlation of MYBL1; **(K, L)** Univariate and multivariate analysis of MYBL1. *P < 0.05, **P < 0.01, ***P <0.001, ****P <0.0001, "ns" means P > 0.05.

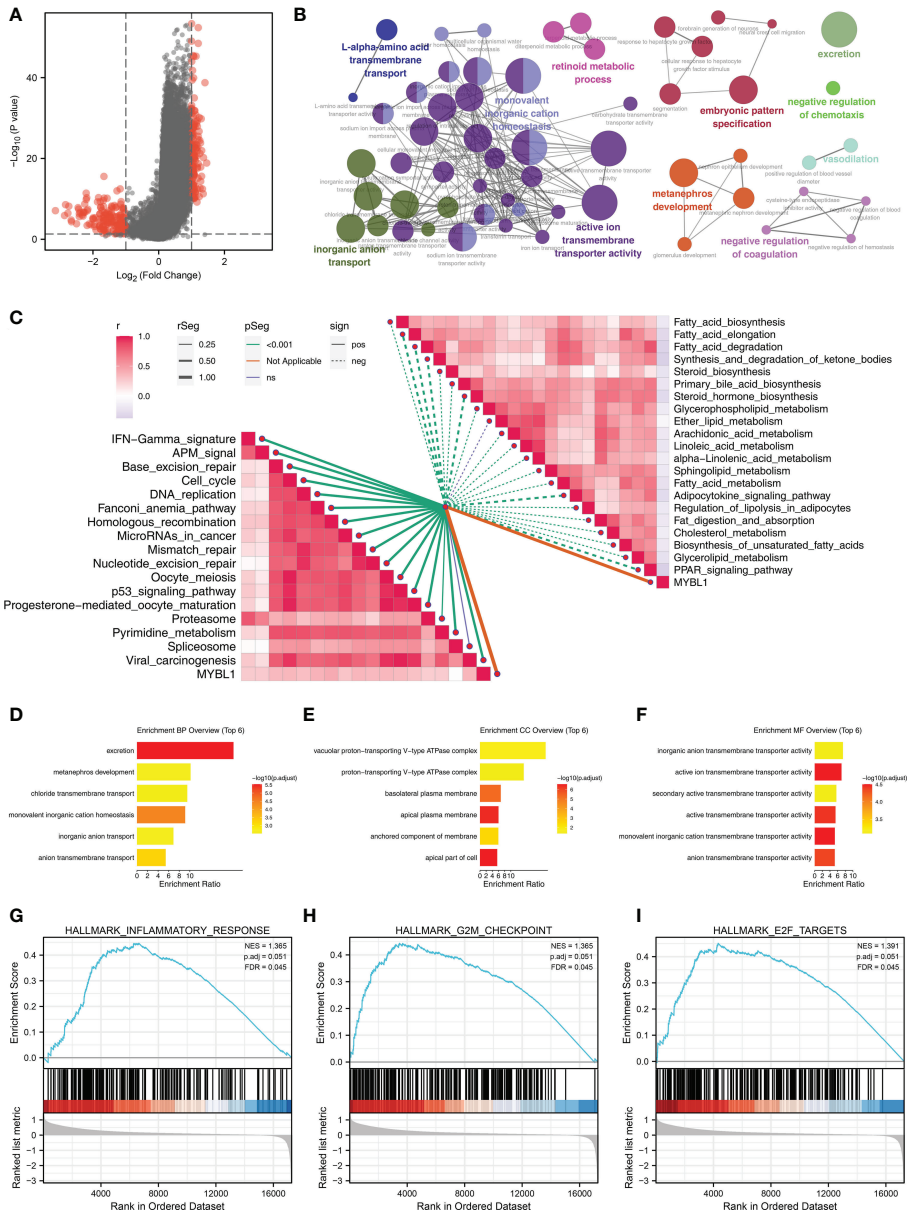


FIGURE 2
The biological role of MYBL1. **(A)** DEGs between the high and low MYBL1 expression; **(B)** clueGO analysis of DEGs; **(C)** Correlation of the pathways quantified by ssGSEA algorithm; **(D–F)** GO analysis of the DEGs; **(G–I)** GSEA analysis of Hallmark gene set.

positive correlation was observed in MYBL1 with TMB and MSI (Figures 4A, B; TMB: R = 0.13; MSI: R = 0.22). Nevertheless, MYBL1 might has no significant effect on mRNAsi (Figure 4C). The genomic mutation characteristics of MYBL1 was shown in Figure 4D (0.6% somatic mutation rate). The top five most differentially mutated genes in patients with high and low MYBL1 expression were VHL, PBRM1, TTN, SETD2, and BAP1 (Figures 4E, F).

Immunotherapy response, drug sensitivity and nomogram plot of MYBL1 in ccRCC

We next calculate the TIDE score of each ccRCC patient based on the TIDE analysis. It seems that MYBL1 had no significant influence on the TIDE score, immune exclusion, immune dysfunction quantified by the TIDE analysis

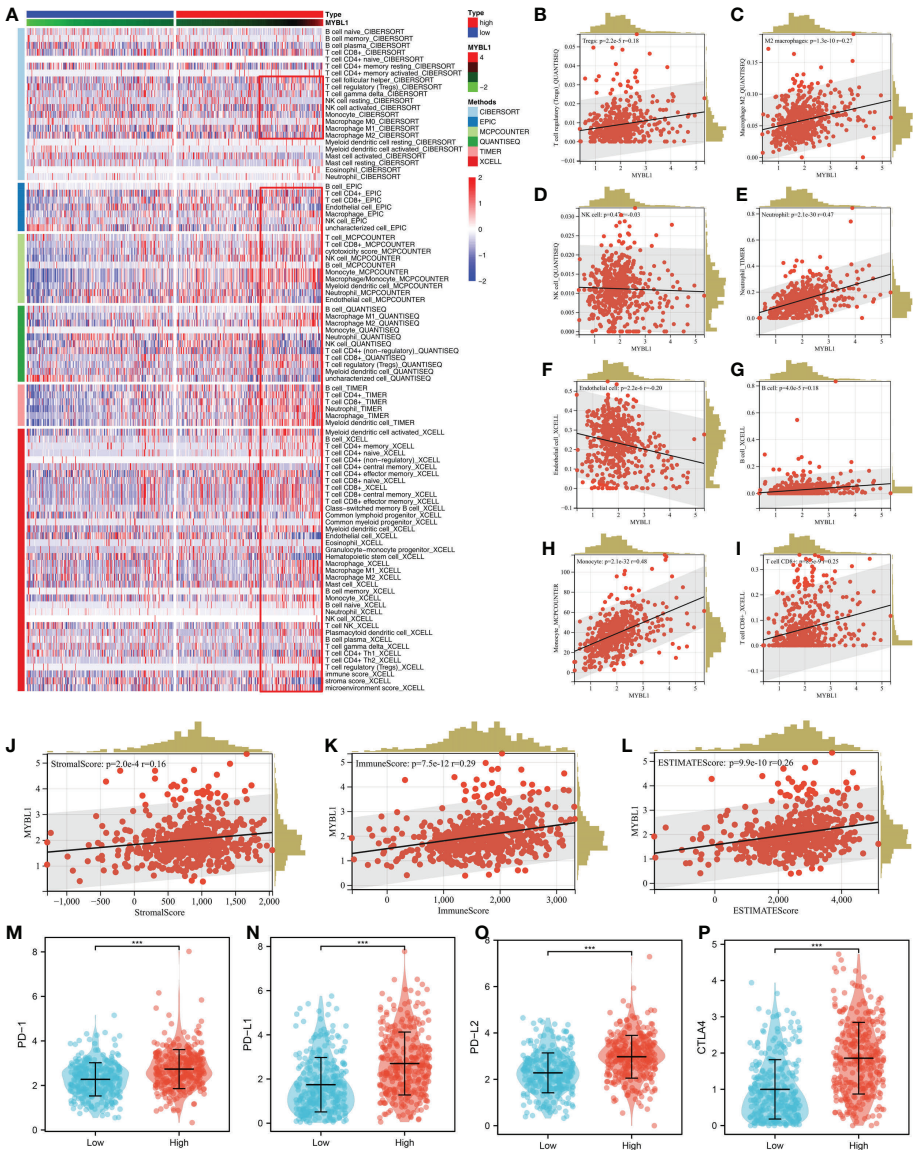


FIGURE 3
The immune landscape of MYBL1. **(A)** Immune infiltration of MYBL1; **(B–I)** Correlation between MYBL1 and multiple immune cells; **(J–L)** Correlation of MYBL1 and immune score, stromal score and estimate score; **(M–P)** The expression level of key immune checkpoints in patients with high and low MYBL1 expression.

(Figures 5A–D). Another aspect, a negative correlation was found in MYBL1 with ips_CTLA4_pos_PD1_neg and ips_CTLA4_neg_PD1_pos, indicating that MYBL1 could affect the immunotherapy response of ccRCC patients (Figures 5E–H). For the common target drugs for ccRCC, we found that MYBL1 can increase the sensitivity of vinblastine and pazopanib (Figures 5I–L). Then, a nomogram plot was constructed by combining the clinical features and MYBL1 expression (Figure 5M). The calibration curve indicated that a good fit between actual and nomogram predicted survival (Figure 5N). Also, the DCA curves showed that the clinical features can improve the performance of MYBL1 on prognosis prediction (Figure 5O).

MYBL1 enhances the malignant biological behaviors of ccRCC

The qRT-PCR result of cell lines indicated that the MYBL1 was overexpressed in ccRCC cells compared to the control cells

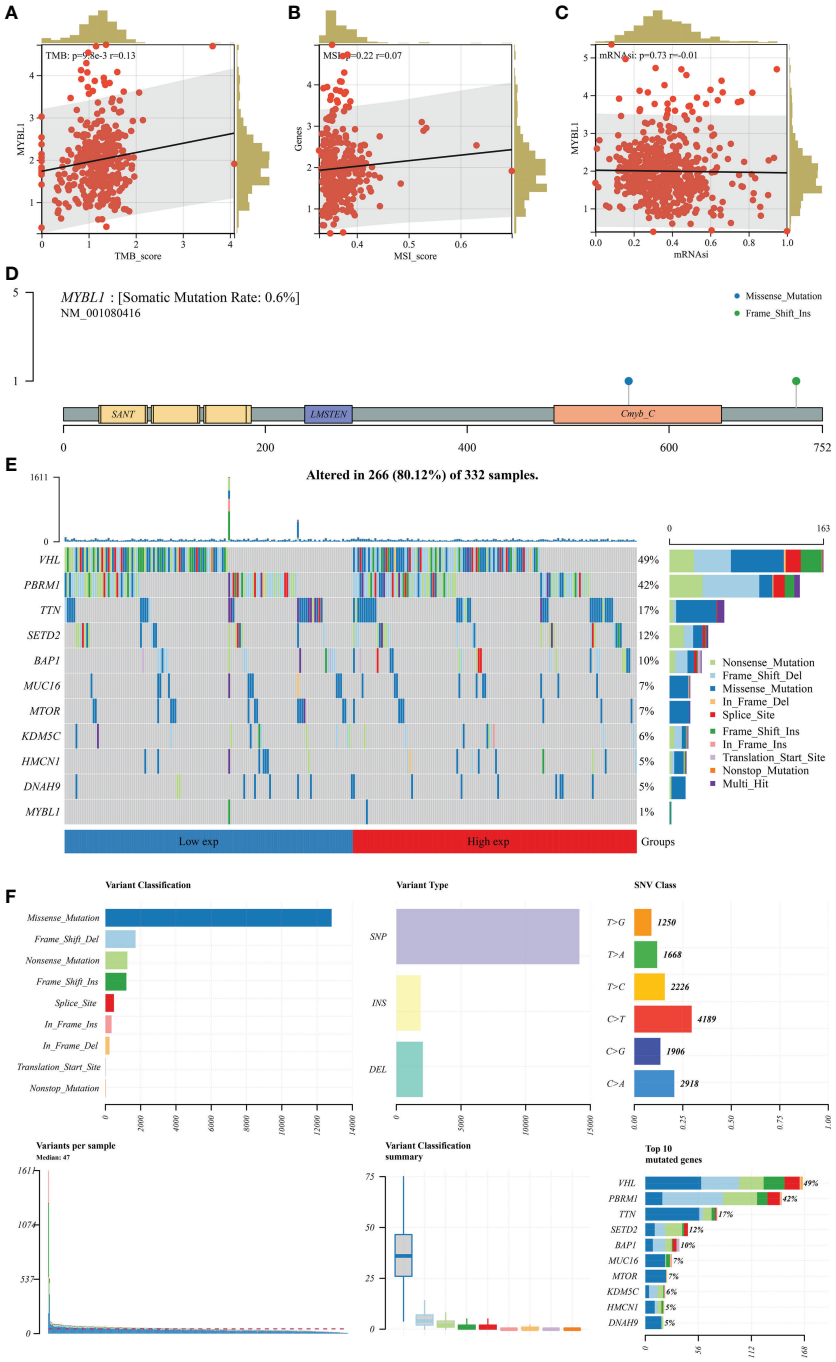


FIGURE 4
Genomic characteristics of MYBL1. **(A)** Correlation of MYBL1 with TMB; **(B)** Correlation of MYBL1 with MSI; **(C)** Correlation of MYBL1 and mRNAsi; **(D–F)** The mutation characteristics of MYBL1.

(Figure 6A). Knockdown efficacy of MYBL1 was shown in Figure 6B. CCK8 assay indicated that the inhibition of MYBL1 might remarkably increase the proliferation ability of ccRCC cells (Figures 6C, D). The same trend was validated by colony formation assay (Figure 6E). Meanwhile, we observed a lower percentage of EdU-positive cells in MYBL1 knockdown cells (Figure 6F). Flow cytometry results indicated that the knockdown of MYBL1 could remarkably increase the apoptosis rate of ccRCC cells (Figure 6G). *In vivo* assay showed that the inhibition of MYBL1 could also hamper tumor growth in mice (Figures 6H–J). IHC showed that MYBL1 was overexpressed in ccRCC cancer tissue compared with the normal tissue obtained from four patients (Figure 7).

Machine learning identified the prognosis signature based on MYBL1-derived molecules

Based on the MYBL1-derived DEGs identified above, univariate Cox regression analysis was conducted to identify the prognosis-related molecules with $P < 0.05$. The top 50 significant prognosis-related genes were shown in Figure 8A. The machine learning algorithm – LASSO logistics regression was utilized to identify the best variable (Figures 8B, C). Multivariate Cox regression analysis was utilized to identify a prognosis signature with the formula of "Risk score = CASR * -0.492 + F11 * -0.167 + IGF2BP3 * 0.262 + TAGLN3 * 0.327 +

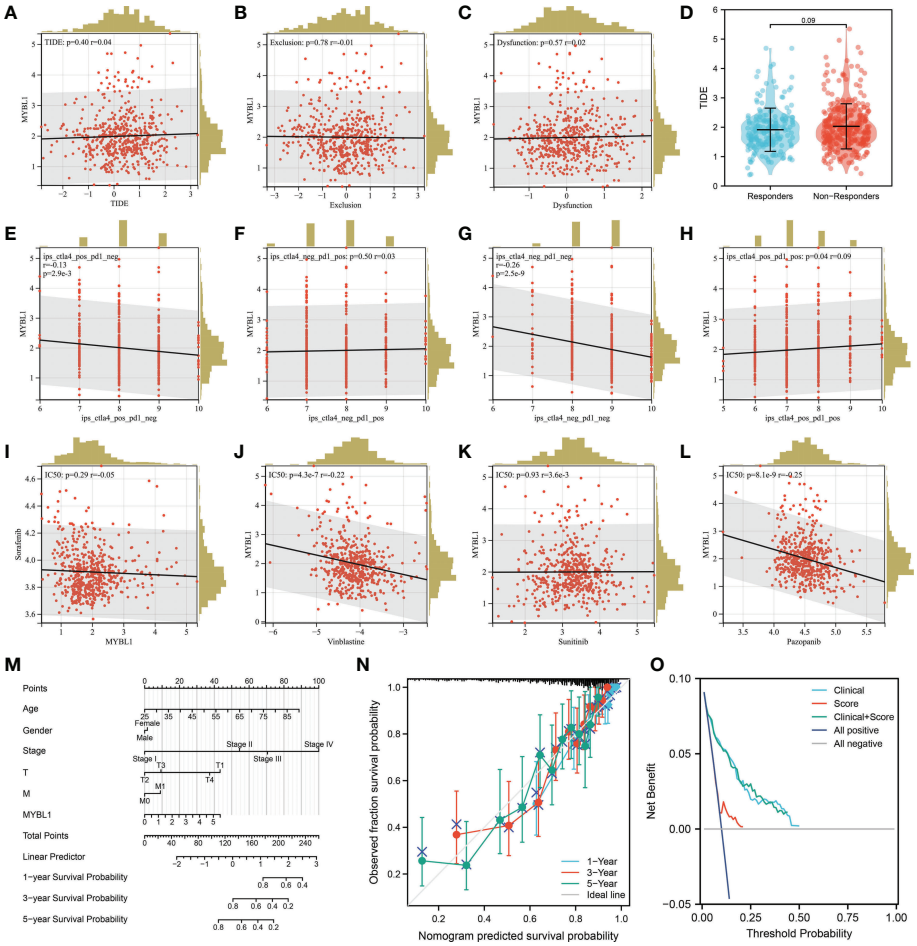


FIGURE 5
Immunotherapy and drug sensitivity of MYBL1. **(A)** Correlation of MYBL1 with TIDE; **(B)** Correlation of MYBL1 with immune exclusion; **(C)** Correlation of MYBL1 with immune dysfunction; **(D)** The MYBL1 expression in immunotherapy responders and non-responders; **(E–H)** Correlation of MYBL1 with IPS score; **(I–L)** Correlation of MYBL1 and drug sensitivity; **(M)** The nomogram plot based on MYBL and nomogram plot; **(N)** The calibration curve; **(O)** The DCA curve.

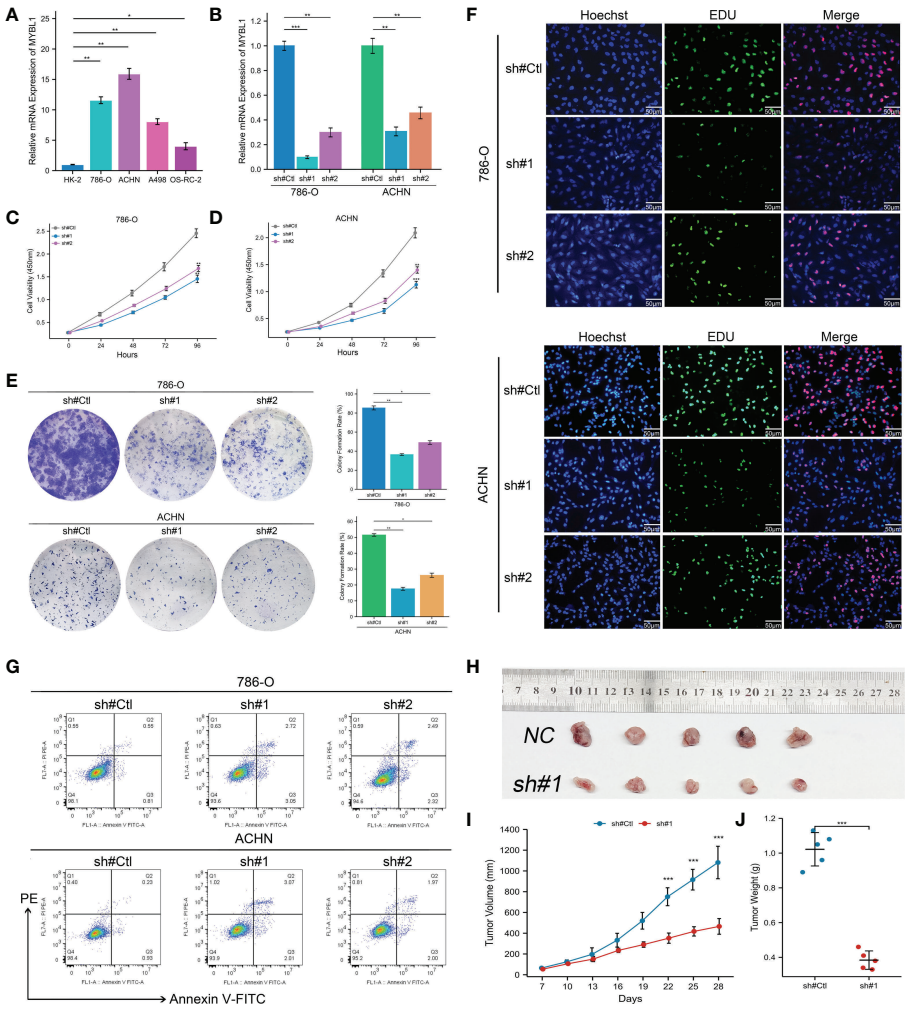


FIGURE 6
MYBL1 enhances the malignant biological behaviors of ccRCC cells. **(A)** The expression level of MYBL1 in renal cancer cells; **(B)** Knockdown efficiency of MYBL1 in ccRCC cells; **(C, D)** CCK8 assay between sh-MYBL1 and control cells; **(E)** Colony formation assay between sh-MYBL1 and control cells; **(F)** EdU assay between sh-MYBL1 and control cells; **(G)** Flow cytometry detecting cell apoptosis between sh-MYBL1 and control cells; **(H–J)** *In vivo* assay showed that the inhibition of MYBL1 could also hamper tumor growth in mice. *P < 0.05, **P < 0.01, ***P < 0.001.

PLPPR1 * -0.245 + SIM2 * 0.334 + RALYL * 0.601 + RUFY4 * 0.381" (Figure 8D). In the training cohort, the patients in the high risk group might have a worse OS (Figure 9A). Also, our signature showed a good prediction ability in patients survival (Figure 9A; 1-year AUC = 0.77; 3-year AUC = 0.74, 5-year AUC = 0.71). Meanwhile, the satisfactory performance of our signature was also observed in the internal validation and external validation cohort (Figures 9B, C). Next, we noticed a positive correlation between the risk score and TIDE score (Figure 9D, R = 0.17, P < 0.001). We found that the immunotherapy responders defined by TIDE analysis tend to have a higher risk score level (Figure 9E). Also, the percentage of immunotherapy responders in high risk patients was 26.7%, greatly lower than 38.9% in low risk patients (Figure 9F). GSEA analysis indicated that the in high risk patients, the pathway of pancreas beta cells, allograft rejection, KRAS signaling, IL6/JAK/STAT3 signaling, spermatogenesis, E2F targets, G2M checkpoints, angiogenesis were significantly activated (Figure S2).

Discussion

Renal cancer is still a threatening disease globally and responsible for almost 2 million deaths per year, resulting in a great public burden (25). Although surgical treatment can provide reliable prognosis benefits for early patients, the survival performance of advanced patients is still limited.

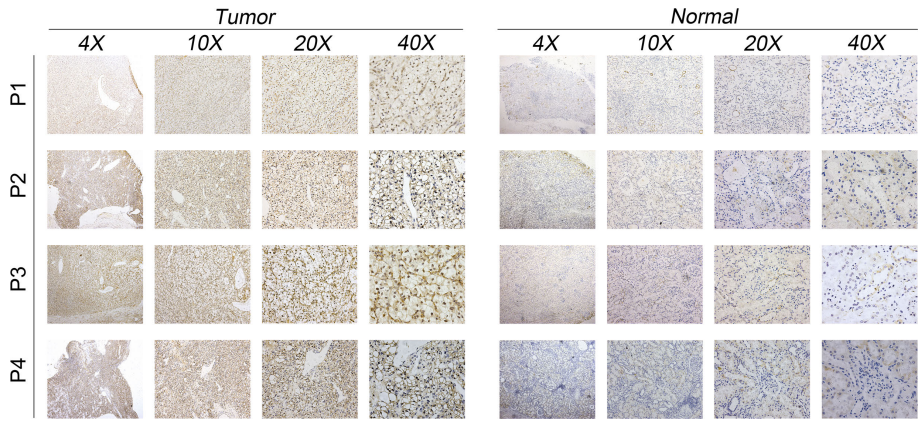


FIGURE 7
IHC result of MYBL1 between ccRCC tumor and control tissue.

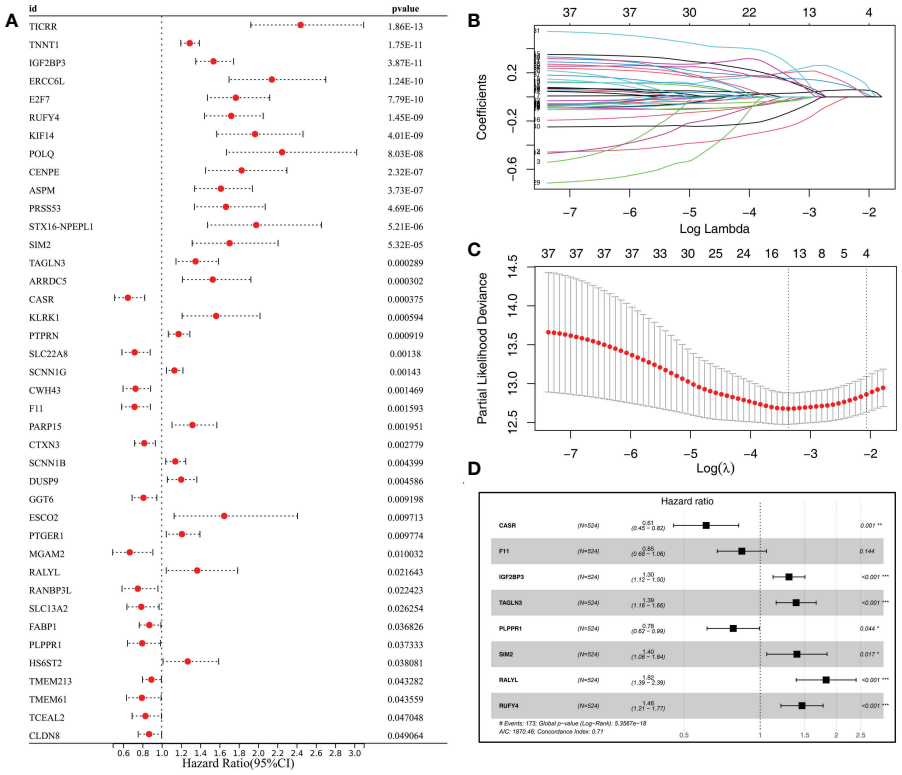


FIGURE 8
Identification of the prognosis signature based on the machine learning algorithm from MYBL1-derived molecules. **(A)** The top 50 molecules significantly correlated with patients prognosis from MYBL1-derived molecules; **(B, C)** LASSO logistics regression was utilized for data dimension reduction; **(D)** Multivariate cox regression analysis.

Consequently, it is extremely meaningful to identify novel biomarkers for ccRCC diagnosis and therapy options.

Our study examined the role of MYBL1 in ccRCC in depth. According to our knowledge, this is the first study investigating MYBL1 in ccRCC. Here, through a series of bioinformatics analysis, we found that MYBL1 was highly expressed in ccRCC patients in both RNA and protein levels. Prognosis analysis revealed that MYBL1 was correlated with poor prognosis performance, including OS, DSS and PFI. Clinical correlation analysis showed that MYBL1 was higher in the worse N- and M-stage. Furthermore, biological enrichment analysis was conducted to explore the biological role of MYBL1 in ccRCC. Moreover, we found that MYBL1 can remodel the immune microenvironment of ccRCC and affect the immunotherapy response. *In vitro* and *in vivo* assays indicated that MYBL1 was upregulated in ccRCC cells and can promote cellular malignant behaviors of ccRCC. Meanwhile, the prognosis signature based on the MYBL1-derived molecules

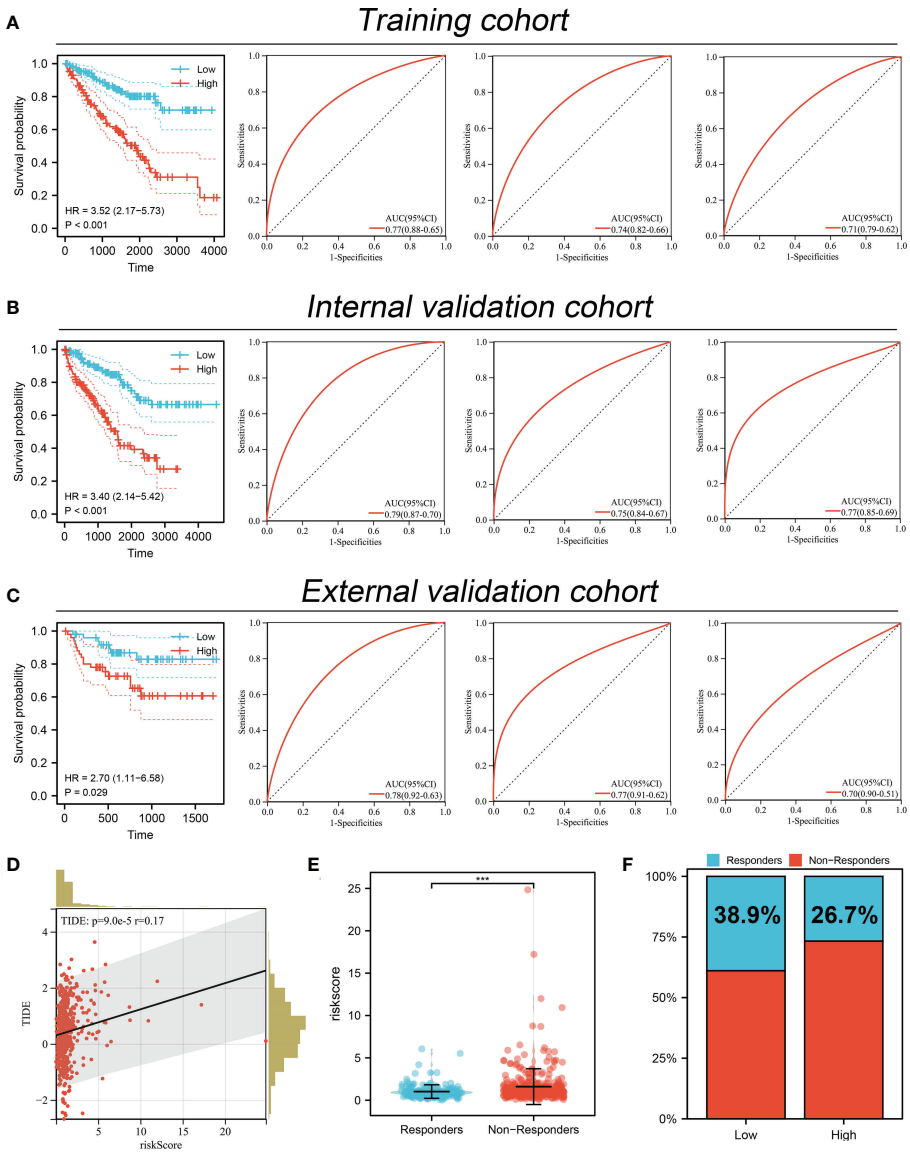


FIGURE 9
Performance of prognosis signature in predicting patients survival and immunotherapy. **(A)** The performance of our signature in predicting patients survival in the training cohort; **(B)** The performance of our signature in predicting patients survival in the internal validation cohort; **(C)** The performance of our signature in predicting patients survival in external validation cohort; **(D)** Correlation of risk score and TIDE; **(E)** The risk score level in immunotherapy responders and non-responders; **(F)** Percentage of immunotherapy responder in patients with high and low MYBL1 expression. ***P <0.001.

showed great prediction ability on patient prognosis in both training and validation cohorts.

We found that MYBL1 is a risk factor for ccRCC based on bioinformatic analysis and experiments. Also, biological enrichment analysis indicated that MYBL1 was mainly enriched in the terms of the inflammatory response, G2M checkpoint and E2F targets. Batova et al. indicated that the acute inflammatory response could be regulated by Englerin A, therefore changing the cell metabolism level and affecting renal cancer progression (26). During the cell cycle, the G2/M checkpoint is an important step. Ding et al. indicated that the dioscin could hamper cell proliferation of osteosarcoma cells based on a G2/M checkpoint-dependent manner (27). Meanwhile, Kent et al. revealed that the dysfunction of E2F in cancers tends to induce the carcinogenic cascade reaction (28). Our result indicated that the MYBL1 might enhance ccRCC progression by affecting the activity of the above pathways.

The immune microenvironment can influence cancer development through complex biological interactions. In our study, we found that the MYBL1 MYBL1 can increase Tregs, M2 macrophages, neutrophils, B cells, monocytes, CD8+ T cells, yet decrease endothelial cells level in the ccRCC microenvironment. Tregs can result in an inhibitory immune microenvironment. In ccRCC, Ji et al. found that the hamper of Tregs in the colon cancer microenvironment can improve the anti-tumor effect and inhibit cancer metastasis (29). Li et al. indicated that aiduqing formula can inhibit Treg infiltration induced by TAM/CXCL1, further hampering breast cancer metastasis (30). In solid tumors, M2 macrophages are generally cancer promoters (31). Chen et al. revealed that gastric and breast cancer metastasis can be facilitated by M2 macrophages recruited by the local tumor microenvironment based on secreted CHI3L1 (32). Xie et al. revealed that the CXCL13 secreted by M2 macrophages facilitated the metastatic potential of ccRCC (33). Based on a comprehensive review conducted by Xiong et al., carcinogenesis and metastasis of cancer can be facilitated by neutrophils (34). These results indicate that MYBL1 might be immune-related molecules that can remodel the immune microenvironment of ccRCC patients.

Moreover, we established a prognosis signature based on the MYBL1-derived molecules. Our signature presents a good prediction ability on patient survival performance. Moreover, the ccRCC patients in different groups might have different responses to immunotherapy. These results indicated the clinical application value of MYBL1 in the clinical.

Although our research is based on high-quality analysis and rigorous experiments, some limitations still need to be noted. Firstly, the population used for analysis was mainly Western. The biological difference between different races can decrease the reliability of our conclusions. Secondly, the in-deep mechanism of MYBL1 to enhance the cellular malignant behaviors of ccRCC is still unclear. In the future, more basic studies focused on MYBL1 in cancers, especially in ccRCC, are needed.

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 human participants were reviewed and approved by Medical Ethics Committee of First Affiliated Hospital of Harbin Medical University. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by Animal Care and Use Committee of the First Affiliated Hospital of Harbin Medical University.

Author contributions

TW and WJ performed the public data analysis. TW and WX wrote the manuscript. TW and WJ performed the experiments. CZ designed this work. All authors contributed to the article and approved the submitted version.

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.2022.1080403/full#supplementary-material

SUPPLEMENTARY FIGURE 1
The flow chart of our study

SUPPLEMENTARY FIGURE 2
GSEA analysis based on Hallmark gene set between high and low risk patients.

References

1. Capitanio U, Montorsi F. Renal cancer. *Lancet (London England)* (2016) 387:894–906. doi: 10.1016/s0140-6736(15)00046-x

2. Wolf MM, Kimryn Rathmell W, Beckermann KE. Modeling clear cell renal cell carcinoma and therapeutic implications. *Oncogene* (2020) 39:3413–26. doi: 10.1038/s41388-020-1234-3

3. Sato Y, Yoshizato T, Shiraishi Y, Maekawa S, Okuno Y, Kamura T, et al. Integrated molecular analysis of clear-cell renal cell carcinoma. *Nat Genet* (2013) 45:860–7. doi: 10.1038/ng.2699

4. Jonasch E, Gao J, Rathmell WK. Renal cell carcinoma. *BMJ (Clinical Res ed.)* (2014) 349:g4797. doi: 10.1136/bmj.g4797

5. Yin Z, Wu D, Shi J, Wei X, Jin N, Lu X, et al. Identification of ALDH3A2 as a novel prognostic biomarker in gastric adenocarcinoma using integrated bioinformatics analysis. *BMC Cancer* (2020) 20:1062. doi: 10.1186/s12885-020-07493-x

6. Wei Y, Chen X, Ren X, Wang B, Zhang Q, Bu H, et al. Identification of MX2 as a novel prognostic biomarker for sunitinib resistance in clear cell renal cell carcinoma. *Front Genet* (2021) 12:680369. doi: 10.3389/fgene.2021.680369

7. Zhu J, Wu Y, Yu Y, Li Y, Shen J, Zhang R. MYBL1 induces transcriptional activation of ANGPT2 to promote tumor angiogenesis and confer sorafenib resistance in human hepatocellular carcinoma. *Cell Death Dis* (2022) 13:727. doi: 10.1038/s41419-022-05180-2

8. Guo H, Zhang B, Nairn AV, Nagy T, Moremen KW, Buckhaults P, et al. O-Linked n-acetylglucosamine (O-GlcNAc) expression levels epigenetically regulate colon cancer tumorigenesis by affecting the cancer stem cell compartment *via* modulating expression of transcriptional factor MYBL1. *J Biol Chem* (2017) 292:4123–37. doi: 10.1074/jbc.M116.763201

9. Brayer KJ, Frerich CA, Kang H, Ness SA. Recurrent fusions in MYB and MYBL1 define a common, transcription factor-driven oncogenic pathway in salivary gland adenoid cystic carcinoma. *Cancer Discov* (2016) 6:176–87. doi: 10.1158/2159-8290.Cd-15-0859

10. Ramkissoon LA, Horowitz PM, Craig JM, Ramkissoon SH, Rich BE, Schumacher SE, et al. Genomic analysis of diffuse pediatric low-grade gliomas identifies recurrent oncogenic truncating rearrangements in the transcription factor MYBL1. *Proc Natl Acad Sci United States America* (2013) 110:8188–93. doi: 10.1073/pnas.1300252110

11. Nikolaus M, Koch A, Stenzel W, Elezkurtaj S, Sahm F, Tietze A, et al. Atypical NMDA receptor expression in a diffuse astrocytoma, MYB- or MYBL1-altered as a trigger for autoimmune encephalitis. *Acta Neuropathol* (2022) 144:385–9. doi: 10.1007/s00401-022-02447-y

12. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinf (Oxford England)* (2009) 25:1091–3. doi: 10.1093/bioinformatics/btp101

13. 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

14. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci United States America* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102

15. Ren X, Chen X, Zhang X, Jiang S, Zhang T, Li G, et al. Immune microenvironment and response in prostate cancer using Large population cohorts. *Front Immunol* (2021) 12:686809. doi: 10.3389/fimmu.2021.686809

16. Chen B, Khodadoust MS, Liu CL, Newman AM. & alizadeh, a. a. profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol (Clifton NJ)* (2018) 1711:243–59. doi: 10.1007/978-1-4939-7493-1_12

17. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res* (2017) 77:e108–10. doi: 10.1158/0008-5472.Can-17-0307

18. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* (2016) 17:218. doi: 10.1186/s13059-016-1070-5

19. Plattner C, Finotello F. & rieder, d. deconvoluting tumor-infiltrating immune cells from RNA-seq data using quanTIseq. *Methods Enzymol* (2020) 636:261–85. doi: 10.1016/bs.mie.2019.05.056

20. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Sci (New York NY)* (2015) 350:207–11. doi: 10.1126/science.aad0095

21. 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

22. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* (2018) 173:338–354.e315. doi: 10.1016/j.cell.2018.03.034

23. McNeish DM. Using lasso for predictor selection and to assuage overfitting: A method long overlooked in behavioral sciences. *Multivariate Behav Res* (2015) 50:471–84. doi: 10.1080/00273171.2015.1036965

24. Yu Y, Yao W, Wang T, Xue W, Meng Y, Cai L, et al. FBXL6 depletion restrains clear cell renal cell carcinoma progression. *Trans Oncol* (2022) 26:101550. doi: 10.1016/j.tranon.2022.101550

25. Morris MR, Latif F. The epigenetic landscape of renal cancer. *Nat Rev Nephrol* (2017) 13:47–60. doi: 10.1038/nrneph.2016.168

26. Batova A, Altomare D, Creek KE, Naviaux RK, Wang L, Li K, et al. Englerin a induces an acute inflammatory response and reveals lipid metabolism and ER stress as targetable vulnerabilities in renal cell carcinoma. *PloS One* (2017) 12:e0172632. doi: 10.1371/journal.pone.0172632

27. Ding Q, Zhang W, Cheng C, Mo F, Chen L, Peng G, et al. Dioscin inhibits the growth of human osteosarcoma by inducing G2/M-phase arrest, apoptosis, and GSDME-dependent cell death *in vitro* and *in vivo*. *J Cell Physiol* (2020) 235:2911–24. doi: 10.1002/jcp.29197

28. Kent LN, Leone G. The broken cycle: E2F dysfunction in cancer. *Nat Rev Cancer* (2019) 19:326–38. doi: 10.1038/s41568-019-0143-7

29. Ji D, Song C, Li Y, Xia J, Wu Y, Jia J, et al. Combination of radiotherapy and suppression of tregs enhances abscopal antitumor effect and inhibits metastasis in rectal cancer. *J Immunother Cancer* (2020) 8(2):e000826. doi: 10.1136/jitc-2020-000826

30. Li J, Wang S, Wang N, Zheng Y, Yang B, Wang X, et al. Aiduqing formula inhibits breast cancer metastasis by suppressing TAM/CXCL1-induced treg differentiation and infiltration. *Cell Communication Signaling CCS* (2021) 19:89. doi: 10.1186/s12964-021-00775-2

31. Mantovani A, Sozzani S, Locati M, Allavena P, Sica A. Macrophage polarization: tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes. *Trends Immunol* (2002) 23:549–55. doi: 10.1016/s1471-4906(02)02302-5

32. Chen Y, Zhang S, Wang Q, Zhang X. Tumor-recruited M2 macrophages promote gastric and breast cancer metastasis *via* M2 macrophage-secreted CHI3L1 protein. *J Hematol Oncol* (2017) 10:36. doi: 10.1186/s13045-017-0408-0

33. Xie Y, Chen Z, Zhong Q, Zheng Z, Chen Y, Shangguan W, et al. M2 macrophages secrete CXCL13 to promote renal cell carcinoma migration, invasion, and EMT. *Cancer Cell Int* (2021) 21:677. doi: 10.1186/s12935-021-02381-1

34. 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

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 14 December 2022
DOI 10.3389/fimmu.2022.1065927

OPEN ACCESS

EDITED BY
Hao Zhang,
Chongqing Medical University, China

REVIEWED BY
Guangquan Zhang,
The Eighth Affilliated Hospital, Sun Yat-sen University, China
Xiao Wang,
China Medical University, China

*CORRESPONDENCE
Xinyu Zhang
xinyuzhang3052000@163.com

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

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Immunology

RECEIVED 10 October 2022
ACCEPTED 30 November 2022
PUBLISHED 14 December 2022

CITATION
Yang S, Sun B, Li W, Yang H, Li N and Zhang X (2022) Fatty acid metabolism is related to the immune microenvironment changes of gastric cancer and RGS2 is a new tumor biomarker. *Front. Immunol.* 13:1065927. doi: 10.3389/fimmu.2022.1065927

Fatty acid metabolism is related to the immune microenvironment changes of gastric cancer and RGS2 is a new tumor biomarker

Shifeng Yang[1,2†], Boshi Sun[1†], Wenjing Li[3†], Hao Yang[1], Nana Li[1] and Xinyu Zhang[1*]

[1]Department of General Surgery, The Second Affiliated Hospital of Harbin Medical University, Harbin, China, [2]The Key Laboratory of Myocardial Ischemia, Ministry of Education, Harbin, China, [3]Department of Otorhinolaryngology, Head and Neck Surgery, Second Affiliated Hospital of Harbin Medical University, Harbin, China

Background: Alterations in lipid metabolism promote tumor progression. However, the role of lipid metabolism in the occurrence and development of gastric cancer have not been fully clarified

Method: Here, genes that are related to fatty acid metabolism and differentially-expressed between normal and gastric cancer tissues were identified in the TCGA-STAD cohort. The intersection of identified differentially-expressed genes with Geneset was determined to obtain 78 fatty acid metabolism-related genes. The ConsensusClusterPlus R package was used to perform differentially-expressed genes, which yielded divided two gastric cancer subtypes termed cluster 1 and cluster 2.

Results: Patients in cluster 2 was found to display poorer prognosis than patients in cluster 1. Using machine learning method to select 8 differentially expressed genes among subtypes to construct fatty acid prognostic risk score model (FARS), which was found to display good prognostic efficacy. We also identified that certain anticancer drugs, such as bortezomib, elesclomol, GW843682X, and nilotinib, showed significant sensitivity in the high FARS score group. RGS2 was selected as the core gene upon an analysis of the gastric cancer single-cell, and Western blotting and immunofluorescence staining results revealed high level of expression of this gene in gastric cancer cells. The results of immunohistochemical staining showed that a large amount of RGS2 was deposited in the stroma in gastric cancer. A pan-cancer analysis also revealed a significant association of RGS2 with TMB, TIDE, and CD8+ T-cell infiltration in other cancer types as well. RGS2 may thus be studied further as a new target for immunotherapy in future studies on gastric cancer.

10.3389/fimmu.2022.1065927

Conclusion: In summary, the FARS model developed here enhances our understanding of lipid metabolism in the TME in gastric cancer, and provides a theoretical basis for predicting tumor prognosis and clinical treatment.

KEYWORDS

gastric cancer, fatty acid, immunotherapy, tumor microenvironment, machine learning

1 Introduction

Gastric cancer (GC) is one of the most prevalent malignant digestive system tumors, characterized by a high degree of heterogeneity, difficulty of treatment, and a poor prognosis (1, 2). The liver is the most frequently affected organ by hematogenous metastases of gastric cancer tumors, after liver metastasis, the survival rate was only 20% (3) The development of neoadjuvant chemotherapy and immunotherapy for gastric cancer treatment in recent years has led to improvements in the diagnosis and prognosis of gastric cancer to a certain extent, yet further improvement is still necessary (4). To this end, new tumor markers, therapeutic targets, and treatment strategies need to be developed (5). Previous studies have shown that the occurrence, proliferation, and metastasis of tumors are closely related to their microenvironment. Various tumor cell metabolites can affect the activation of surrounding immune cells in various ways, and suppress their antitumor activity. Alterations in the tumor microenvironment promotes proliferation and development of tumor cells (6). Growing evidence suggests that reprogramming of energy metabolism towards e.g. lactic acid production and acetylation enzymes contributes to the progression of gastric cancer (7). An in-depth investigation of metabolic changes in the tumor microenvironment of gastric cancer may thus provide with a new marker or therapeutic target to improve gastric cancer prognosis and treatment.

In lipid metabolism and especially fatty acid (FA) synthesis, nutrients are converted into metabolic intermediates for membrane biosynthesis, energy storage, and signal molecule production (8). Alterations in lipid metabolism is a hallmark and metabolic phenotype of cancer cells. Blocking the supply of lipids to cancer cells has a significant impact on cancer cell bioenergetics, membrane biosynthesis, and intracellular signal transduction (9). Most tumors were previously shown to display an abnormal lipid metabolism (10). Polymorphonuclear myelogenous suppressor cells (PMN-MDSCs) are pathologically-activated neutrophils that play an important role in the regulation of cancer immune response.The selective pharmacological inhibition of FATP2 was also found to eliminate the activity of PMN-MDSCs, and significantly delay tumor progression in mice. Inhibition of PMN-MDSCs thus improves the efficiency of cancer treatment (11, 12). Therefore, targeted fatty acid-induced oxidative stress can prevent cancer-induced cachexia.

In recent years, inhibition of FA synthesis has attracted attention as a potential strategy for cancer treatment, yet it is not yet implemented in clinical practice (13). The role of lipid metabolism in gastric cancer has also not been widely studied previously. Therefore, we conducted an in-depth study here on the expression and significance of fatty acid disorder-related genes in gastric cancer. We identified differentially-expressed fatty acid metabolism-related genes in gastric cancer, and determined two subtypes based on consistency clustering analysis. A prognostic signature (FRAS) model was constructed by performing a univariate Cox regression analysis of differentially-expressed genes in different subtypes, and used as a potential molecular marker of gastric cancer to identify immune infiltration and genomic instability patterns. FeaturePlot visualization was performed to display the expression and distribution of model genes in the cell population and to verify the accuracy of the model. A "core gene", *RGS2* was selected for subsequent experiments, and the relationship between the expression level of RGS2 protein and the prognosis of patients with gastric cancer was evaluated. Finally, we also discussed the biological significance of the *RGS2* gene in multiple cancer types to fully understand the role of fatty acid metabolism in gastric cancer, and to provide a theoretical basis for effective treatment.

2 Materials and methods

2.1 Patients and tissues samples

All patients were admitted to the Second Affiliated Hospital of Harbin Medical University between May 2020 and June 2022, and diagnosed by pathological examination.Pathological diagnosis was based on the 8th edition of the American Joint Commission on Cancer (14). All participants have informed consent. The study design was approved by the Internal Audit and Ethics Committee of the Second Affiliated Hospital of Harbin Medical University (No : KY2021-075).

2.2 Western blotting

The protein content of the cells was extracted, and the expression of *RGS2* protein was analyzed by Western blotting after the cell density of cultures of AGS, HGC27, MKN-45, MKN-1, and the GES-1 cell lines reached 90%.

2.3 Data preparation and processing

STAD clinical information and expression data were obtained from the American Cancer Genome Map Database (TCGA, https://cancergenome.nih.gov/) using the TCGA R package biollinks. Tumor samples with both expression and survival information were retained for follow-up analysis, which included 373 cancer and 32 paracancerous samples. Fatty acid-related genes (Geneset) are derived from fatty acid-related factors (fattyacid) in the MsigDB database (HALLMARK, KEGG, REACTOME). A total of 14 pathways and 342 related genes were identified.

2.4 Clustering analysis

An intersection between the identified differentially-expressed genes with Geneset yielded 78 differentially-expressed fatty acid-related genes. Using the ConsensusClusterPlus R package, differentially-expressed genes related to fatty acid disorder were clustered based on Euclidean distance. The maximum number of clusters was set to five, and the clustering method to K-means, in order to find a stable and reliable subgroup classification. The results yielded two subtypes, and the differential gene expression between two subtypes was analyzed (screening condition of the difference was: absolute value of log2FC > 1, P< 0.05).

2.5 Construction of prognostic risk model

The genes differentially-expressed between the subtypes were analyzed using univariate Cox regression analysis to identify genes related to the prognosis of subtypes. For this purpose, LASSO penalty Cox regression analysis was used *via* the Rglmnet package to construct a prognostic model to minimize the risk of overfitting. Patient scores were calculated according to the expression levels of the pathway genes and their corresponding regression coefficients.

$$\text{Score} = \sum_{i=0}^{n} \beta i * xi$$

βi: weight coefficient of each gene; χi: expression of each gene (FPKM). Patients were divided into high and low score groups based on the median score, and the survminer R package was used for survival analysis of OS based on high and low scores.

2.6 Evaluation of immune cell filtration

The CIBERSORT algorithm provided by the IOBR R package was used to calculate the scores of immune cells in 22 types of tumor microenvironments using the default parameters. Based on the gene expression profile in the TCGA-STAD data, the proportion of immune cell infiltration was calculated.

2.7 Single-cell dataset analysis

The Seurat R package, which is single-cell transcriptome analysis tool, was used to analyze the single-cell dataset. The analysis workflow mainly included the steps of constructing objects, data standardization, data dimensionality reduction, clustering, and searching for marker genes. Then, the SingleR R packagewas used to annotate the clustering results obtained from Seurat.

2.8 Drug sensitivity

Using the pRRophetic R package and the expression data of model genes, the sensitivity (IC50 value) of 138 drugs in the GDSC database was predicted, and the sensitivity of STAD patients to drug therapy was evaluated based on the predicted IC50 values.

2.9 Statistical analysis

The R program (version 4.1.2) was used for statistical analysis. The survival curve was generated using the Kaplan-Meier method, and the differences between groups were compared using the log-rank test. A Cox regression model was used for univariate and multivariate analyses combined with other clinical features to determine the independent prognostic value of the risk score. The R package timeROC was used for time-dependent ROC curve analysis to evaluate the predictive value of prognostic characteristics. ROC analysis was used to evaluate the sensitivity and specificity of the score in predicting prognosis, and the area under the ROC curve (AUC) was considered to judge prognosis. Statistical significance was set at p< 0.05. The same formula is used to calculate verification scores.

3 Results

3.1 Differential expression of fatty acid related genes in tumor tissues and their biological functions

The study population included 373 STAD and 32 paracancerous tissue samples obtained from the TCGA-STAD cohort.|Using log2FC | > 0.585, BH-corrected, and P< 0.05 as differential expression criteria, 3857 genes were found to be differentially-expressed in gastric cancer and paracancerous tissues, with 2801 and 1056 up- and down-regulated genes, respectively. A total of 78 fatty acid-related differentially-expressed genes were identified by determining the intersection of these genes with the Geneset, A volcano map and a differentially-expressed fatty acid metabolism-related gene thermogram is shown in Figures 1A, B. The PPI network showed that *HSP90AA1*, *EPHX2*, *ACOX2*, *ACADM*, *ACLY*, and other genes had high connectivity in the network (Figure 1C). The correlation between the expression levels of differentially-expressed fatty acid metabolism-related genes was also calculated. Fatty acid metabolism-related genes were found to be classified into three groups (Figure 1D). A functional GO enrichment t was found for oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor, acting on paired donors and binding or reducing oxygen molecules, CH-CH group acting on donor, and easy to bind iron ions. These enzymes participate in long-chain fatty acid metabolism, fatty acid biosynthesis, eicosane-like metabolism, olefin metabolism, and unsaturated fatty acid metabolism (Figures 1E–H). The clinical feature analysis revealed that there were significant differences in the expression of some fatty acid metabolism-related genes between different age, sex, stage, and grade groups (Figures S1A-D).

3.2 Determination of molecular subtypes based on fatty acid metabolism related genes

Subtyping can be used to reveal distinct states of the tumor, and thus help implement personalized treatment strategies. Cancer samples from the TCGA gastric cancer data were subjected to consistency clustering based on expression patterns of 78 different fatty acid metabolism-related genes to identify groups of samples with similar expression patterns. According to the cumulative distribution function and incremental region map of consistent clustering, the change in the CDF curve for the case of two clusters (k = 2, clusters 1 and 2) was found to be close to smooth. Hence, the samples were divided into two subtypes (Figures 2A-C). We found that there were significant differences in survival time between patients with different fatty acid metabolism subtypes, and the prognosis of patients in cluster 2 was worse than that of cluster 1 patients (Figure 2D). In addition, the scores of angiogenesis-related pathways in the HALLMARK and GOBP gene sets in the MSigDB database were calculated using SSGSEA. The results showed significant differences between the scores of all pathways related to angiogenesis between the fatty acid metabolism related molecular subtypes (Figure 2E). A large number of blood vessels (Figure 2F) were found in gastric cancer tissues by HE staining. Immune cell infiltration was calculated using CIBERSORT, and the immune score, matrix score, and tumor purity (Figure 2G) were calculated using ESTIMATE algorithms. The heat map of immune cell infiltration in subtypes showed that there were significant differences in Mast_cells_activated, Dendritic_cells_resting, Macrophages_M0, etc. Inter-subtype immune scores and matrix scores (Figure 2H).

3.3 Construction of a fatty acid metabolism-related prognostic signature

We have identified 515 genes differentially-expressed between the two subtypes under the screening condition of | log2FC | > 1 dBH correction p< 0.05. A total of 454 and 61 genes were up- and down-regulated, respectively. Univariate Cox regression analysis showed that 146 genes were associated with OS. KM analysis revealed eight genes (eight genes screened after LASSO-Cox regression analysis) (Figure S2). The signature (Figures 3A-C) composed of eight genes, and was determined based on the optimal value of λ. The regression coefficient of each gene is shown in Table S1.

3.4 Verification of prognostic efficacy of FARS based on an analysis of training and external independent verification sets

The score of each patient was calculated according to the formula and the patients were divided into high score group and low score group by the median score. KM curve showed that the survival probability of patients with high score was significantly lower than that of patients with low score (Figures 3D, E). To evaluate the predictive efficiency of prognostic models in 1 -, 2 -, and 3-year survival rates, we performed a time-related ROC analysis. The area under the ROC curve (AUC) is 0.627 at 1 year, 0.643 at 2 years and 0.631 at 3 years, indicating that the prediction effect of the model is good (Figure 3F). Univariate and multivariate Cox analysis were used to determine whether Score was an independent prognostic factor for OS. In univariate Cox analysis, Score obtained from TCGA data queue was significantly correlated with OS. After correcting other confounding factors, multivariate Cox analysis showed that Score was still an independent predictor of OS (Figure 3G).

In order to verify the stability of the model, the Score of each sample is also calculated in GSE13861 dataset and GSE26899 dataset based on the same algorithm. According to the median of Score, gastric cancer samples were divided into high score group and low score group. Consistent with the results obtained by the TCGA

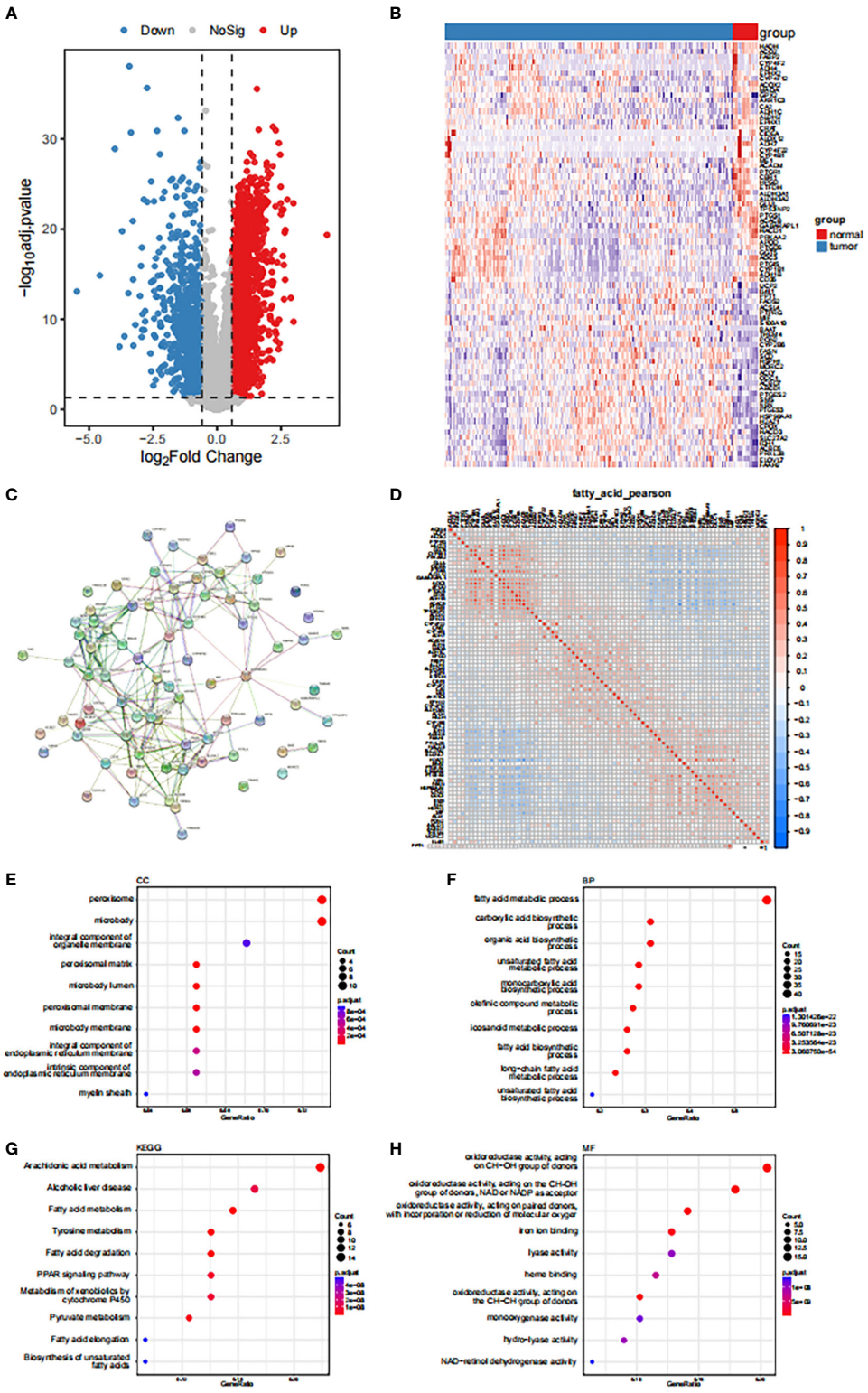


FIGURE 1
Expression disturbance of fatty acid-related genes in tumors **(A)** Volcanogram of fatty acid related genes differentially expressed in gastric cancer and paracancerous tissues. **(B)** Heat map of differentially expressed fatty acid-related genes. **(C)** Based on the differentially expressed fatty acid related genes, the PPI network was constructed by using STRING database. **(D)** Expression correlation analysis of genes related to differential fatty acids. **(E–H)** GO functional enrichment analysis of genes related to differential fatty acids.

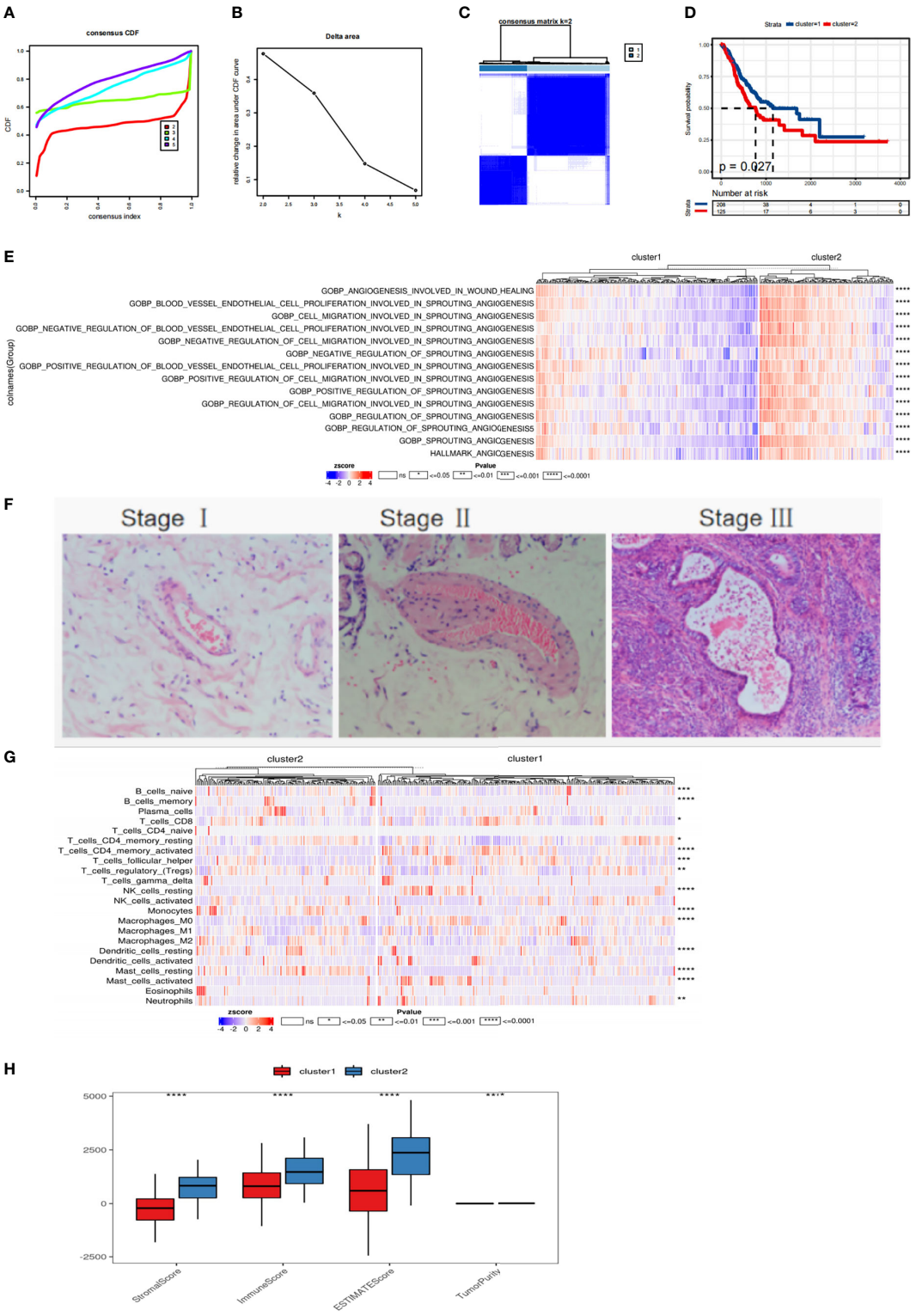


FIGURE 2
Molecular subtype recognition of fatty acid related genes **(A, B)** TCGA gastric cancer samples were analyzed by consistent cluster analysis based on 78 fatty acid related genes. **(C)** The consistency matrix heat map when the number of clusters is 2. **(D)** Survival curve of patients with fatty acid subtypes. **(E)** The scores of ANGIOGENESIS-related pathways in HALLMARK,GOBP gene set in MSigDB database were calculated by SSGSEA. **(F)** HE staining of gastric cancer tissue. **(G)** Using CIBERSORT to calculate the heat map of immune cell infiltration. **(H)** ESTIMATE was used to calculate immune score, matrix score and tumor purity. ns means p > 0.05, *p<=0.05, **p<=0.01, ***p<=0.001 and ****p<=0.0001.

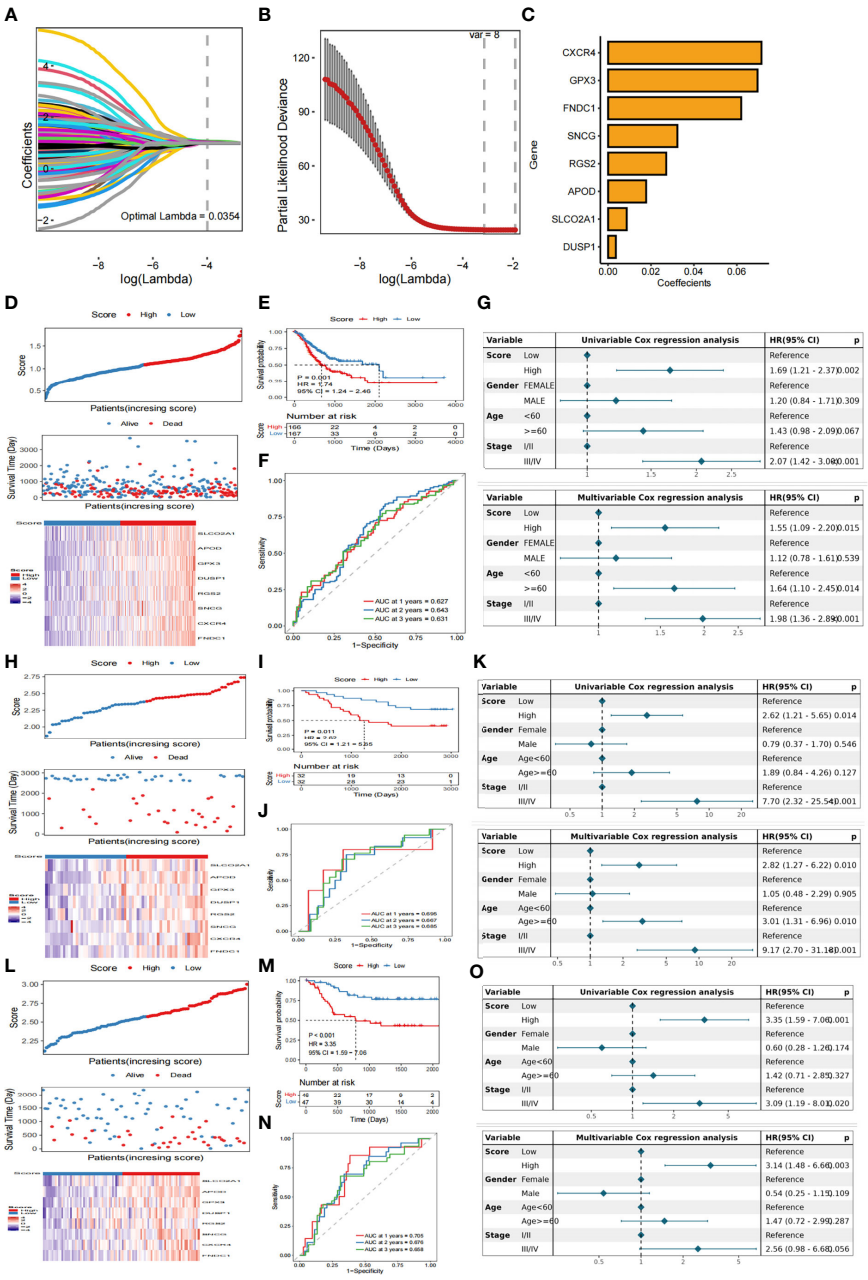


FIGURE 3
Construction of hierarchical scoring system for Fatty Acid related prognosis **(A)** LASSO coefficient profiles of eight genes. **(B)** Tuning parameter lambda (λ) selected by cross-validation error curve. **(C)** Eight genes determined based on the optimal value of λ **(D)** The relationship between survival status/risk score of TCGA cohort, mRNA expression heat map of 8 genes and survival time (days)/risk score. **(E)** Kaplan-Meier OS analysis of gastric cancer patients in low-risk and high-risk groups **(F)** The time-dependent ROC curve of TCGA training data set OS. AUC was evaluated at 1 year, 3 years and 5 years, respectively. p value was calculated using the log-rank test. p< 0.001. **(G)** Univariate and multivariate Cox analysis were used to determine whether Score was an independent prognostic factor for OS. **(H)** GSE13861 dataset was used to analyze the relationship between survival status/risk score, mRNA expression heat map of 8 genes and survival time (days)/risk score. **(I)** Kaplan-Meier OS analysis of gastric cancer patients in low-risk and high-risk groups based on GSE13861 data set **(J)** Time-dependent ROC curve of GSE13861 dataset OS. AUC was evaluated at 1 year, 3 years and 5 years, respectively. P value was calculated using the log-rank test. P< 0.001. **(K)** The GSE13861 dataset uses univariate and multivariate Cox analysis to determine whether Score is an independent prognostic factor for OS. **(L)** The relationship between survival status/risk score, mRNA expression heat map of 8 genes and survival time (days)/risk score was analyzed by GSE26899 dataset. **(M)** Kaplan-Meier OS analysis of gastric cancer patients in low-risk and high-risk groups based on GSE26899 data set **(N)** Time-dependent ROC curve of GSE26899 dataset OS. AUC was evaluated at 1 year, 3 years and 5 years, respectively. P value was calculated using the log-rank test. P< 0.001. **(O)** The GSE26899 dataset uses univariate and multivariate Cox analysis to determine whether Score is an independent prognostic factor for OS.

cohort, patients with high scores had a lower probability of survival than patients with low scores (Figures 3H, I, L, M). In addition ,the prognostic model revealed that the 1- year AUC was 0.695, 2-year AUC was 0.667, 3-year AUC was 0.685 in the GSE13861 dataset (Figure 3J),and 1-year AUC was 0.705, 2-year AUC was 0.676, and3-year AUC was 0.658 in the GSE26899 dataset (Figure 3N). In the validation set, univariate and multivariate Cox analyses were also used to determine whether Score was an independent prognostic factor for OS. The results show that in univariate Cox analysis, there is a significant correlation between Score and OS. After correcting other confounding factors, multivariate Cox analysis shows that Score is still an independent predictor of OS (Figures 3K, O).

3.5 FARS is related to the clinical characteristics of tumor

We found that the score of patients with *Helicobacter pylori* infection was significantly higher than that of patients without infection and significant differences were also detected between patients with first-, second-, and third-grade cancer: higher grades corresponded to higher scores and poorer prognosis (Figure 4A). Immune cell infiltration as calculated by the CIBERSORT algorithm revealed that many immune cell types, such as Mast_cells_activated, Dendritic_cells_resting, and Macrophages_M0, are significantly correlated with the FARS score (Figure 4B). Figure 4C shows the difference in gene expression of immune checkpoints in the high- and low-risk groups of scores, in which the expression leves of *CD276*, *CTLA4*, *PDCD1*, and *PDCD1LG2* were significantly higher in the high score group. This high expression level helps gastric cancer cells escape immune surveillance and promote immune escape. Based on the calculation of the Pearson correlation between the fatty acid risk score (Score) and the identified gene signature score, we detected several gene sets related to immunity and EMT from the literature, and then performed mapping between the SSGSEA score and the fatty acid risk score (Score) of these samples. We found a significant correlation between the FRAS score and EMT2, EMT3, and PanFTBRS, which promote the EMT process in gastric cancer cells (Figure 4D). We further evaluated the relationship between fatty acid risk score and chemotherapeutic drug resistance, and also calculated the difference in chemotherapeutic drug resistance between highFARS and lowFARS using the pRRophetic package. The IC50 values of bortezomib, elesclomol, and nilotinib were found to be significantly different between highFARS and lowFARS, and with stronger chemotherapeutic effects (Figure 4E) in the low-score group.

3.6 Single-cell dataset analysis

Using the STAD samples in the single-cell data set downloaded from the GEO database (GSE142750), the cells were grouped and annotated based on an t-SNE analysis. A total of 107597 cells (33694 features) were grouped into 13 clusters, and finally annotated as two large cell groups (Figure 5A). Then, the union of the top 5 marker genes of in each cluster was used to draw a heat map to show the differential expression of each marker gene in each subtype. No genes included in the constructed model (model genes) was detected (Figure 5B) among these top 5 marker genes.

Feature plot visualization was used to show the expression and distribution of model genes in the cell population. The results showed that *RGS2* and *DUSP1* were significantly expressed in the cell population, and three model genes, *CXCR4*, *SLCO2A1*, and *FNDC1*, were not in the cluster, indicating that the two model genes that were significantly expressed could be used as marker genes (Figure 5C) of cancer.

3.7 Biological significance of *RGS2* in gastric cancer

We found that the high expression of *RGS2* in gastric cancer was significantly correlated with a shorter survival time (Figure 6A). The TME score showed that the high expression of *RGS2* was positively correlated with the stomalscore, Immunescore, and Estimatescore, which indicated a worse immune response (Figure 6B). Correlation analysis of immune cell infiltration showed that the expression of 10 types of immune cells in 22 types of immune cells was correlated with *RGS2* expression (Figure 6C). We also analyzed the clinicopathological features of patients with high and low *RGS2* expression, including age, sex, survival, grade grade, T stage, and N stage. The figure shows that there is no statistical difference in age and sex between the high and low *RGS2* expression groups. High *RGS2* expression was found to be closely related to poor prognosis. This finding shows that high expression of *RGS2* represents a higher degree of malignancy based on clinicopathological features (Figures 6D, E). We also analyzed the relationship between expression levels of *RGS2* and immune checkpoints (Figure 6F). We found that the lower tumor mutation load in the group with high expression of *RGS2* increased the difficulty of receiving the benefit of immune checkpoint inhibitors for patients (Figure 6G). We found that the TIDE score of the *RGS2* high expression group was significantly higher than that of low expression group (Figure 6H). This also indicates that high *RGS2* expression is more likely to lead to immune dysfunction and immune rejection. We have determined the mRNA and protein expression levels of *RGS2* in GES-1 gastric mucosal cells and AGS, HGC-27, MKN-1, and MKN-45 gastric cancer cell lines. Accordingly, the expression of RGS2 in gastric cancer cell line was found to be higher than that of GES-1 (Figure 6I) at both mRNA level and protein level. Immunofluorescence staining

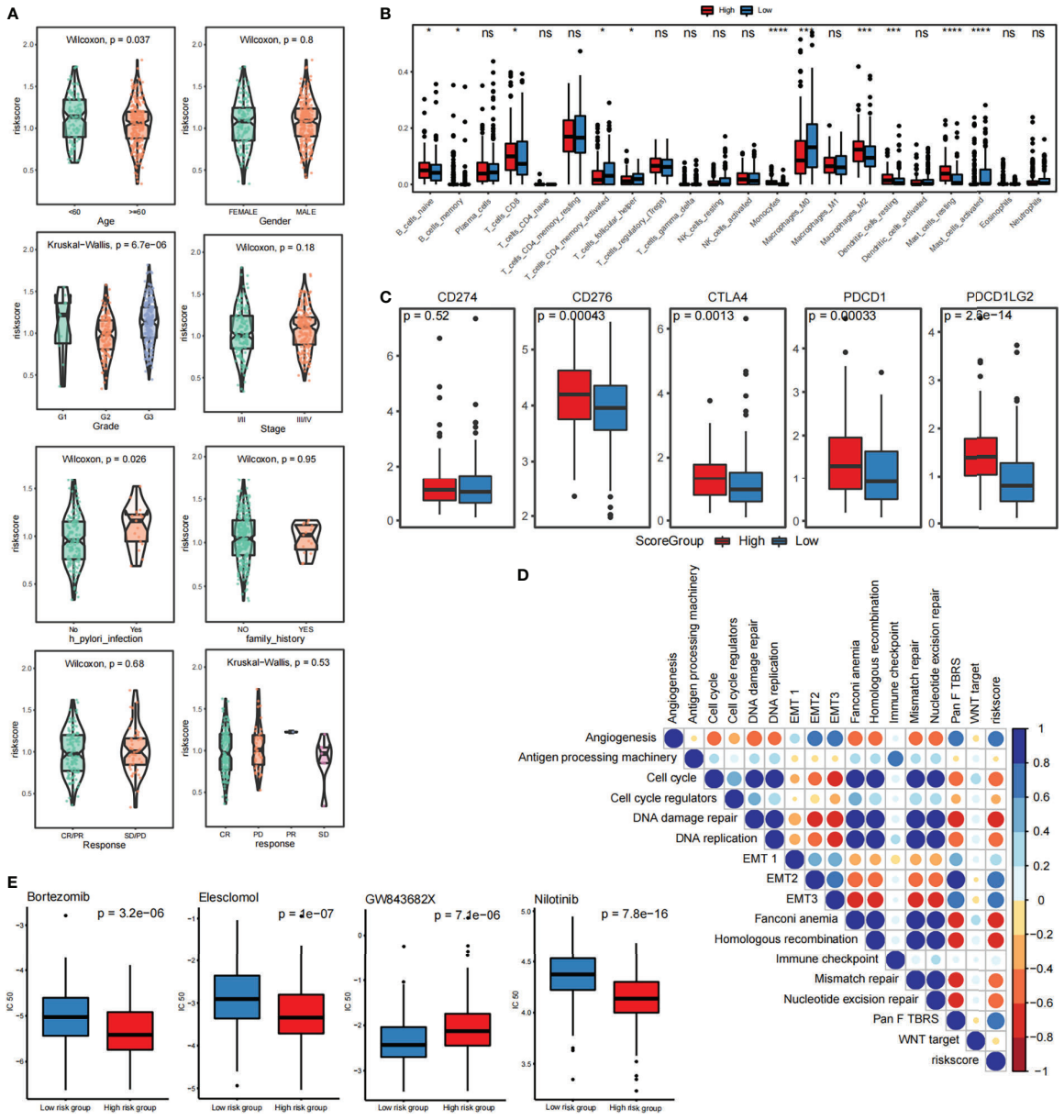


FIGURE 4
Relationship between risk scoring system and immune infiltration and immunotherapy **(A)** Analysis of correlation between risk score and clinical characteristics of patients with STAD **(B)** CIBERSORT calculated the relationship between immune cell infiltration and risk score **(C)** The difference of gene expression in immune checkpoint between high and low score groups. **(D)** Pearson analysis of SSGSEA score and fatty acid risk score. **(E)** Correlation analysis between fatty acid risk score and chemotherapy resistance. *p<=0.05, ***p<=0.001, ****p<=0.0001, ns means p > 0.05.

showed that RGS2 was highly expressed in gastric cancer cell lines AGS and MKN45, and most of them were located in the cytoplasm (Figure 6J). In order to verify the expression of RGS2 in gastric cancer, we found that RGS2 was expressed to varying degrees in different clinical stages of gastric cancer by immunohistochemical staining, and with the increase of staging, the more RGS2 deposition (Figure 6K).

3.8 Biological significance of RGS2 in other cancer types

A pan-cancer analysis has shown that *RGS2* is expressed in many tumor types (Figures 7A, B). We found that the expression of *RGS2* in the overall survival time (OS) was significantly correlated with the survival rates of BLAC, KIRC, LIHC, SKCM, STAD,

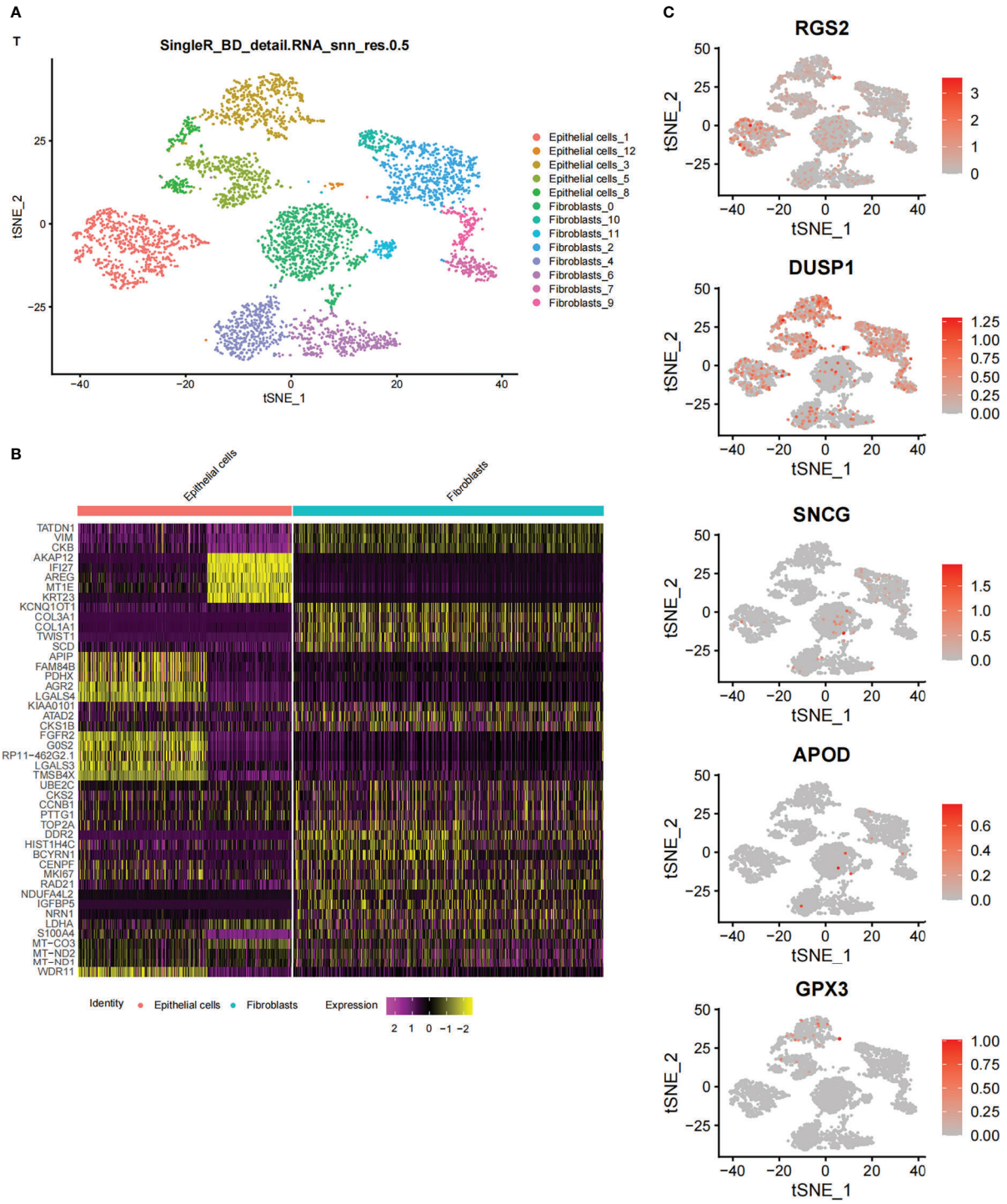


FIGURE 5
Single cell data set analysis **(A)** T-SNE clustering and cell group annotation based on STAD samples from single-cell data sets. **(B)** Draw a heat map of the Marker gene of TOP5 in each Cluster. **(C)** Display the expression and distribution of model genes in the cell population by FeaturePlot visualization.

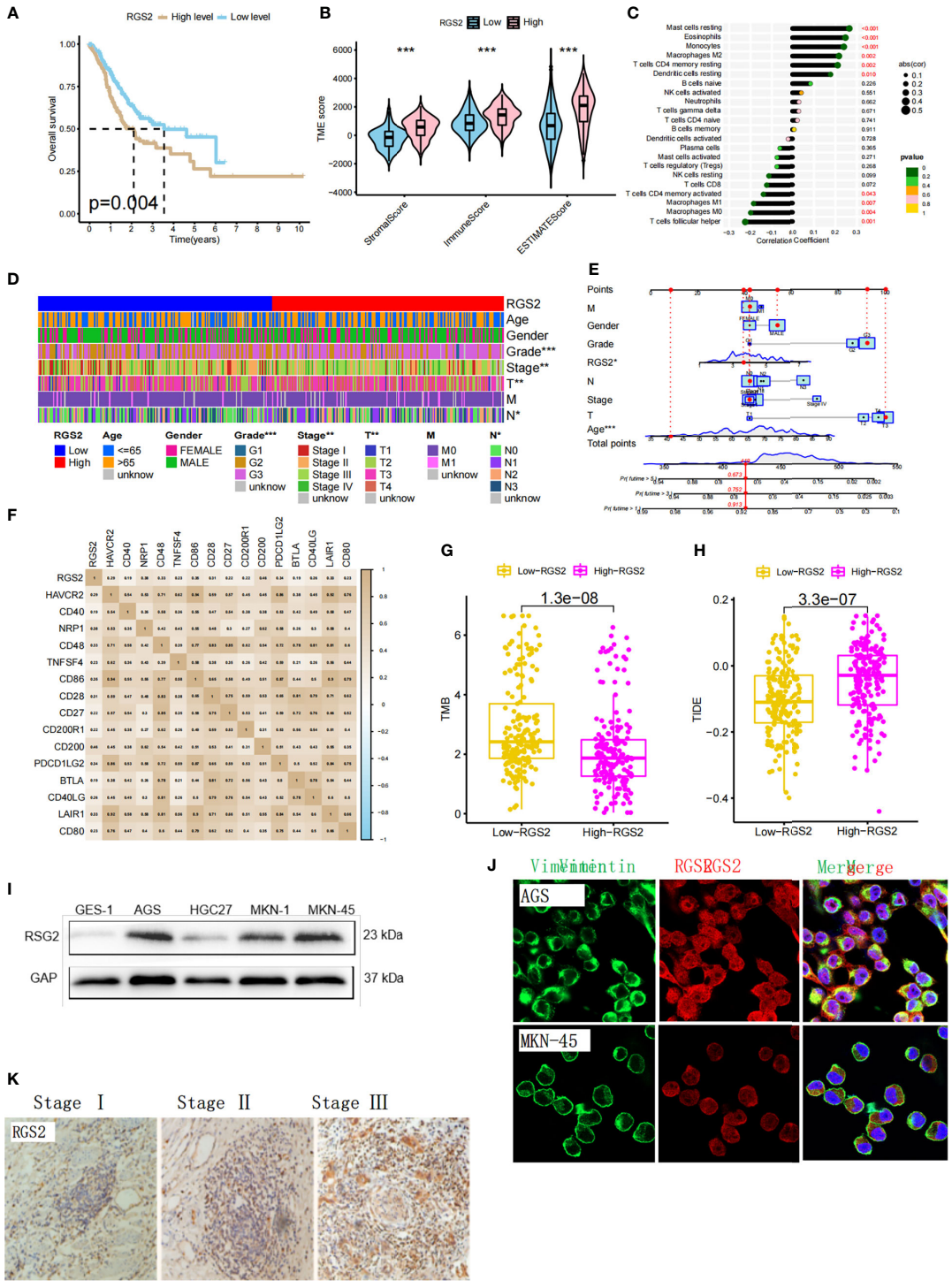


FIGURE 6
Biological significance of RGS2 in gastric cancer and expression level of gastric cancer cell line **(A)** The relationship between the high and low expression of RGS2 in gastric cancer and prognosis **(B)** The relationship between the high and low expression of RGS2 and the score of TME **(C)** The correlation between the high and low expression of RGS2 and immune cell infiltration. **(D, E)** Analysis of the correlation between the high and low expression of RGS2 and the clinicopathological features of patients **(F)** The relationship between the high and low expression of RGS2 and immune checkpoints **(G)** The relationship between the high and low expression of RGS2 and TMB score **(H)** The relationship between high and low expression of RGS2 and TIDE score **(I)** Expression at mRNA level of RGS2 in gastric cancer cell lines AGS, HGC27, MKN-1, MKN45 and normal gastric mucosal cells GES-1 **(J)** Expression and localization of RGS2 in gastric cancer cell lines AGS and MKN45 **(K)** Expression of RGS2 in different clinical stages of gastric cancer. ***p<=0.001.

THCA, and THYM (Figure 7C). There was also no significant difference in the expression of *RGS2* between cancer and disease-free survival (DFS) groups (Figure 7D). There was a correlation between disease-specific survival and ACC, BLCA, KIRC, PRAD, SKCM, STAD, and THYM (Figure 7E), and also a significant correlation between progression-free survival and ACC, KIRC, and THYM (Figure 7F). We analyzed the correlation between RGS2, TMB, and MSI, and found that it was significantly correlated with TGCT, STAD, PAAD, COAD, and CESC, suggesting that it can be used as a basis of detection for immunotherapy of the above tumors (Figures 7G, H). Finally, we found that *RGS2* was closely related to the level of immune cell infiltration in most tumors, suggesting that *RGS2* participates in the regulation of the tumor immune response in the tumor microenvironment (Figure 7I).

4 Discussion

Rapid proliferation and insufficient angiogenesis of tumor cells lead to hypoxia, low pH levels, and depletion of nutrients in

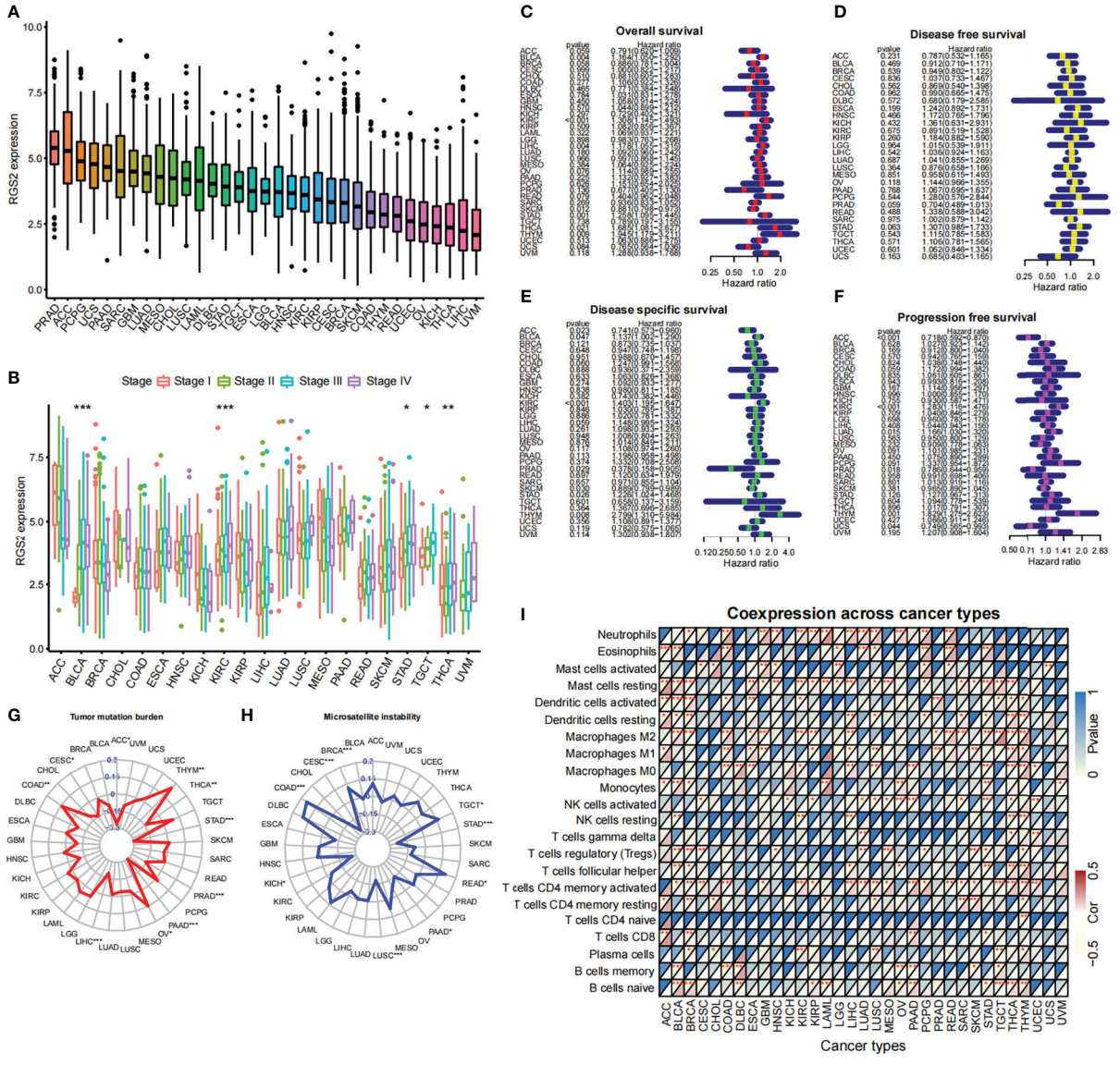


FIGURE 7
Biological significance of RGS2 in pan-cancer **(A)** RGS2 expression level in multiple tumors **(B)** Correlation analysis between RGS2 expression level and tumor stage **(C)** Relationship between RGS2 expression level and OS **(D)** Relationship between RGS2 expression level and disease-free survival **(E)** Relationship between RGS2 expression level and Disease specific survival **(F)** Relationship between RGS2 expression level and Progression free survival **(G)** Correlation analysis between RGS2 and Tumor mutation burden (TMB) **(H)** Analysis of the correlation between RGS2 and Microsatellite instability (MSI) **(I)** The relationship between RGS2 and immune cell infiltration. *p<=0.05, **p<=0.01,***p<=0.001.

the tumor microenvironment (15, 16). Therefore, tumor cells show unique metabolic characteristics that are distinct from those of normal cells. Tumor cells deal with a variety of adverse microenvironments by reprogramming their metabolism, and blocking carcinogenic signals to maintain their proliferating state and survival. Abnormal energy metabolism is thus a hallmark of cancer, which indicates that the metabolism of carbohydrates, lipids, and amino acids in tumor cells is significantly different from that in normal cells. Fatty acid metabolism was previous shown to maintain tumorigenesis, disease progression, and therapeutic resistance by enhancing lipid synthesis, storage, and decomposition (17, 18). Recently, increasing attention has been paid to the role of membrane fatty acids (with respect to e.g. the ratio of saturated fatty acids, monounsaturated fatty acids, and polyunsaturated fatty acids) in promoting cell survival, limiting lipotoxicity, and iron-dependent cell death (19–21). Here, eight fatty acid metabolism-related genes related to gastric cancer prognosis were identified based on an analysis of genomic information of 373 STAD samples and 32 paracancerous tissue samples using univariate COX regression, clustering, and principal component analyses. A model called "FRAS" was constructed, and the score calculated using this model (FRAS score) was found to be closely related to increased immune cell infiltration, genomic instability, immune escape and sensitivity of immune checkpoint inhibitor (ICIs). This fatty acid metabolism-related model was comprehensively evaluated as well. The fatty acid prognostic risk score model was found to be able to independently predict the prognosis of patients with gastric cancer, and effectively distinguish the sensitivity of patients to chemotherapeutic drugs. In addition, the relationship between the prognostic risk score model and characteristics of TME cell infiltration was studied. The prognostic risk score model was found to identify patients with gastric cancer who are suitable for anti-CTLA4 antibody immunotherapy sucessfully, and thereby also indicated that fatty acid metabolism is crucial for shaping individual TME characteristics. These findings may provide a new perspective for exploring the mechanisms of fatty acid metabolism and treatment of gastric cancer.

Rapidly proliferating tumor cells show a high affinity for lipids and cholesterol by increasing exogenous lipid uptake, or by overactivating their biosynthetic pathways (22). Therefore, fatty acid synthesis (FAS) inhibitors, especially fatty acid synthase (FASN), have been the focus of cancer treatment studies (23–25). *RGS2, DUSP1, CXCR4, FNDC1, SNCG, SLCO2A1, APOD*, and *GPX38* were selected to construct this risk model. This model can predict the prognosis of patients with gastric cancer more accurately that a single clinical variable, which may be helpful for clinicians in making clinical decisions. The model was used to classify patients with stage G2/G3, patients aged > 65 years and< 65 years, and patients with *Helicobacter pylori* infection into two groups. This was found to have a significant impact on prognosis, as it confers the advantage of using genetic characteristics in predicting clinical grouping and prognosis.

Gastric cancer patients also develop drug resistance eventually, even though 60% of them are sensitive to chemotherapy. This leads to a 5-year survival rate of less than 10% (26–28). Therefore, understanding the mechanism of chemotherapy resistance in gastric cancer cells is important for improving the prognosis and survival rate. Previous studies have revealed that some cancer cells require fatty acid oxidation to provide energy that is required to maintain the stem cell state. Studies on resistance of breast tumor stem cells (BCSCs) to chemotherapy have found that JAK/STAT3 signaling systems help breast cancer cells maintain their stem cell status, and resistance to chemotherapy by promoting fatty acid oxidation (29, 30). Animal experiments have further confirmed that drugs that inhibit the JAK/STAT3 signaling system can greatly reduce the population of stem cells in breast cancer, and improve the efficiency of chemotherapy (31, 32). Here, we further analyzed the relationship between the develop fatty acid metabolism-related risk score and chemotherapy resistance in gastric cancer cells, and identified significant differences in sensitivity to chemotherapeutic drugs between the high- and low-score groups. Specifically, bortezomib, elesclomol, and nilotinib showed better therapeutic effects in the low-score groups. Targeting of the fatty acid metabolism may thus be a new strategy for reversing drug resistance in gastric cancer cells.

The G protein signal transduction regulatory factor (RGS) gene family, which includes negative regulators of G protein-coupled receptors, are potential drug targets for the treatment of malignant tumors (33, 34). RGS is a large family of genes with multiple functions (35–37). These proteins share an RGS domain with a conserved core that includes 130 amino acid residues, which can directly bind to the activated G-α subunit to inactivate GTP, and thus help negatively regulate GPCR-related signaling pathways (38–40). RGS gene has been proved to be closely related to the occurrence and development of many systemic diseases and cancers (41–43). Here, we analyzed the role of RGS2, in the tumor microenvironment in gastric cancer, and also in other cancer types for the first time. The results showed that the expression of *RGS2* was correlated with interstitial and immune scores. Therefore, we speculate that RGS2 participates in the occurrence and development of gastric cancer by affecting the migration of immune cells. Moreover, we also found that the TMB score of the RGS2 high-expression group was lower than that of the low-expression group, and the TIDE score was higher than that of the low-expression group. This indicates that it is more difficult for gastric cancer patients to benefit from immunotherapy, and have a worse prognosis. High expression levels of *RGS2* were detected by Western blot analysis, which indicates a role of RGS2 in the progression of gastric cancer.In gastric cancer, the deposition of RGS2 increased with the increase of clinical stage. Therefore, in the microenvironment of gastric cancer, RGS2 may predict a poor

prognosis. *RGS2* expression in various tumor types was also found to be significantly correlated with survival, clinical stage, immune score, TMB score, and MSI. Therefore, RGS2 could be used as a new tumor marker as well.

However, our study has suffered from some limitations as well. For example, further research is still needed to reveal how fatty acid-related genes affect immune cell infiltration and genomic instability in gastric cancer. In addition, as this study mainly used online datasets for analysis, more clinical data supplement is necessary.

5 Conclusion

In conclusion, we analyzed here the expression of fatty acid metabolism-related genes in gastric cancer, and constructed a model based on fatty acidification to calculate a disease risk score for gastric cancer. Our analysis revealed that FARS score in gastric cancer is closely related to tumor mutation load, genomic instability, ICIs treatment response, immune cell infiltration, and immune escape. This score provides with a new tool for the diagnosis and treatment of gastric cancer, and the genes related to FARS may become new tumor markers or therapeutic targets. In general, the FARS score developed in this study can be used as a potential molecular classification tool for gastric cancer, and thus help identify immune infiltration and genomic instability patterns in gastric cancer. FARS can also be used to evaluate response of patients to ICIs treatment.

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 study design was approved by the Internal Audit and Ethics Committee of the Second Affiliated Hospital of Harbin Medical University (No: KY2021-075). The patients/participants provided their written informed consent to participate in this study.

Author contributions

We would like to thank all the participants and staff for their valuable contributions. SY, BS and WL contributed equally as co-first authors. XZ conceived and designed the study; HY and NL drafted the manuscript. All authors contributed to the article and approved the submitted version.

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.2022.1065927/full#supplementary-material

SUPPLEMENTARY TABLE 1
The regression coefficient of each gene

SUPPLEMENTARY FIGURE 1
Differential expression of fatty acid related genes in different clinical characteristics **(A)** Difference of fatty acid related gene expression among Age groups of gastric cancer **(B)** Difference of fatty acid related gene expression among gastric cancer Grade groups **(C)** Difference of fatty acid related gene expression among gastric cancer Gender groups **(D)** Difference of fatty acid related gene expression among gastric cancer Stage groups

SUPPLEMENTARY FIGURE 2
Distribution of different groups among subtypes in clinical characteristics. **(A)** Correlation analysis between fatty acid subtypes and different clinical features. **(B)** KM curve analysis of 8 genes (8 genes screened after lasso-cox)

References

1. Lin Y, Zheng Y, Wang H-L, Wu J. Global patterns and trends in gastric cancer incidence rates (1988-2012) and predictions to 2030. *Gastroenterology* (2021) 161(1):116–127.e8. doi: 10.1053/j.gastro.2021.03.023

2. Wong MCS, Huang J, Chan PSF, Choi P, Lao XQ, Chan SM, et al. Global incidence and mortality of gastric cancer, 1980-2018. *JAMA Netw Open* (2021) 4 (7):e2118457. doi: 10.1001/jamanetworkopen.2021.18457

3. Tanaka H, Kanda M, Miwa T, Umeda S, Sawaki K, Tanaka C, et al. G-Protein subunit gamma-4 expression has potential for detection, prediction and therapeutic targeting in liver metastasis of gastric cancer. *Br J Cancer* (2021) 125(2):220–8. doi: 10.1038/s41416-021-01366-1

4. Zhang X, Chen Xu., Liu J, Li Y, Wu J, Chen M, et al. A novel metabolism-related prognostic gene development and validation in gastric cancer. *Clin Transl Oncol* (2022). doi: 10.1007/s12094-022-02958-w

5. Zhang W, Auguste A, Liao X, Walterskirchen C, Bauer K, Lin Y-H, et al. A novel B7-H6-targeted IgG-like T-cell engaging antibody for the treatment of gastrointestinal tumors. *Clin Cancer Res* (2022) 28(23). doi: 10.1158/1078-0432.CCR-22-2108

6. Yu H, Wang M, Zhang T, Cao L, Li Z, Du Y, et al. Dual roles of β-arrestin 1 in mediating cell metabolism and proliferation in gastric cancer. *Proc Natl Acad Sci U.S.A.* (2022) 119(40):e2123231119. doi: 10.1073/pnas.2123231119

7. Zhang L, Li S. Lactic acid promotes macrophage polarization through MCT-HIF1α signaling in gastric cancer. *Exp Cell Res* (2020) 388(2):111846. doi: 10.1016/j.yexcr.2020.111846

8. Luo Xu., Li R-R, Li Y-Q, Yu H-P, Yu H-N, Jiang W-G, et al. Reducing VEGFB expression regulates the balance of glucose and lipid metabolism in mice *via* VEGFR1. *Mol Med Rep* (2022) 26(3):285. doi: 10.3892/mmr.2022.12801

9. Zhang C, Zhu N, Li H, Gong Y, Gu J, Shi Y, et al. New dawn for cancer cell death: Emerging role of lipid metabolism. *Mol Metab* (2022) 63:101529. doi: 10.1016/j.molmet.2022.101529

10. Vriens K, Christen S, Parik S, Broekaert D, Yoshinaga K, Talebi A, et al. Evidence for an alternative fatty acid desaturation pathway increasing cancer plasticity. *Nature* (2019) 566:403–6. doi: 10.1038/s41586-019-0904-1

11. Veglia F, Tyurin VA, Blasi M, De Leo A, Kossenkov AV, Donthireddy L, et al. Fatty acid transport protein 2 reprograms neutrophils in cancer. *Nature* (2019) 569:73–8. doi: 10.1038/s41586-019-1118-2

12. Fukawa T, Yan-Jiang BC, Min-Wen JC, Jun-Hao ET, Huang D, Qian C-N, et al. Excessive fatty acid oxidation induces muscle atrophy in cancer cachexia. *Nat Med* (2016) 22(6):666–71. doi: 10.1038/nm.4093

13. Röhrig F, Schulze A. The multifaceted roles of fatty acid synthesis in cancer. *Nat Rev Cancer* (2016) 16(11):732–49. doi: 10.1038/nrc.2016.89

14. Rice TW, Ishwaran H, Hofstetter WL, Kelsen DP, Apperson-Hansen C, Blackstone EH, et al. Recommendations for pathologic staging (pTNM) of cancer of the esophagus and esophagogastric junction for the 8th edition AJCC/UICC staging manuals. *Dis Esophagus* (2016) 29(8):897–905. doi: 10.1111/dote.12533

15. Infantino V, Santarsiero A, Convertini P, Todisco S, Iacobazzi V. Cancer cell metabolism in hypoxia: Role of HIF-1 as key regulator and therapeutic target. *Int J Mol Sci* (2021) 22(11):5703. doi: 10.3390/ijms22115703

16. Ma Z, Xiang X, Li S, Xie P, Gong Q, Goh B-C, et al. Targeting hypoxia-inducible factor-1, for cancer treatment: Recent advances in developing small-molecule inhibitors from natural compounds. *Semin Cancer Biol* (2022) 80:379–90. doi: 10.1016/j.semcancer.2020.09.011

17. Junior RP, Chuffa LG deA, Simão VA, Sonehara NatháliaM, Chammas R, Reiter RJ, et al. Melatonin regulates the daily levels of plasma amino acids, acylcarnitines, biogenic amines, sphingomyelins, and hexoses in a xenograft model of triple negative breast cancer. *Int J Mol Sci* (2022) 23(16):9105. doi: 10.3390/ijms23169105

18. Hu C, Pang Bo., Lin G, Zhen Yu., Yi H. Energy metabolism manipulates the fate and function of tumour myeloid-derived suppressor cells. *Br J Cancer* (2020) 122(1):23–9. doi: 10.1038/s41416-019-0644-x

19. Zhang M, Wei T, Zhang X, Guo D. Targeting lipid metabolism reprogramming of immunocytes in response to the tumor microenvironment stressor: A potential approach for tumor therapy. *Front Immunol* (2022) 13:937. doi: 10.3389/fimmu.2022.937

20. Cormerais Y, Vučetić M, Parks SK, Pouyssegur J. Amino acid transporters are a vital focal point in the control of mTORC1 signaling and cancer. *Int J Mol Sci* (2020) 22(1):23. doi: 10.3390/ijms22010023

21. Liu L, Wang Y, Zheng T, Cao B, Li M, Shi J, et al. Metabolic shifts induced by human H460 cells in tumor-bearing mice. *BioMed Chromatogr* (2016) 30(3):337–42. doi: 10.1002/bmc.3553

22. Salisbury TB, Arthur S. The regulation and function of the l-type amino acid transporter 1 (LAT1) in cancer. *Int J Mol Sci* (2018) 19(8):2373. doi: 10.3390/ijms19082373

23. Muthusamy T, Cordes T, Handzlik MK, You Le., Lim EW, Gengatharan J, et al. Serine restriction alters sphingolipid diversity to constrain tumour growth. *Nature* (2020) 586(7831):790–5. doi: 10.1038/s41586-020-2609-x

24. Jones CL, Stevens BM, D'Alessandro A, Reisz JA, Culp-Hill R, Nemkov T, et al. Inhibition of amino acid metabolism selectively targets human leukemia stem cells. *Cancer Cell* (2018) 34(5):724–740.e4. doi: 10.1016/j.ccell.2018.10.005

25. Beloribi-Djefaflia S, Vasseur S, Guillaumond F. Lipid metabolic reprogramming in cancer cells. *Oncogenesis* (2016) 5:e189. doi: 10.1038/oncsis.2015.49

26. McClellan B, Gries P, Harlow B, Tiziani S, Jolly C, deGraffenried L. An IGF-1R-mTORC1-SRPK2 signaling axis contributes to FASN regulation in breast cancer. *BMC Cancer* (2022) 22(1):976. doi: 10.1186/s12885-022-10062-z

27. Jiang W, Xing X-L, Zhang C, Yi L, Xu W, Ou J, et al. MET and FASN as prognostic biomarkers of triple negative breast cancer: A systematic evidence landscape of clinical study. *Front Oncol* (2021) 11:604801. doi: 10.3389/fonc.2021.604801

28. Chen Q, Yang Z, Ding H, Li H, Wang W, Pan Z. CircWHSC1 promotes breast cancer progression by regulating the FASN/AMPK/mTOR axis through sponging miR-195-5p. *Front Oncol* (2021) 11:649242. doi: 10.3389/fonc.2021.649242

29. Narita Y, Shoji H, Kawai S, Mizukami T, Nakamura M, Moriwaki T, et al. REVIVE study: A prospective observational study in chemotherapy after nivolumab therapy for advanced gastric cancer. *Future Oncol* (2021) 17(8):869–75. doi: 10.2217/fon-2020-0621

30. Kim JW, Cho S-Y, Chae J, Kim J-W, Kim T-Y, Lee K-W, et al. Adjuvant chemotherapy in microsatellite instability-high gastric cancer. *Cancer Res Treat* (2020) 52(4):1178–87. doi: 10.4143/crt.2020.313

31. Tanaka H, Kanda M, Shimizu D, Tanaka C, Inokawa Y, Hattori N, et al. Transcriptomic profiling on localized gastric cancer identified CPLX1 as a gene promoting malignant phenotype of gastric cancer and a predictor of recurrence after surgery and subsequent chemotherapy. *J Gastroenterol* (2022) 57(9):640–53. doi: 10.1007/s00535-022-01884-6

32. Wang T, Fahrmann JF, Lee H, Li YJ, Tripathi SC, Yue C, et al. JAK/STAT3-regulated fatty acid β-oxidation is critical for breast cancer stem cell self-renewal and chemoresistance. *Cell Metab* (2018) 27(6):1357. doi: 10.1016/j.cmet.2017.11.001

33. Hu Y, Zheng M, Wang S, Gao L, Gou R, Liu O, et al. Identification of a five-gene signature of the RGS gene family with prognostic value in ovarian cancer. *Genomics* (2021) 113(4):2134–44. doi: 10.1016/j.ygeno.2021.04.012

34. Hurst JH, Hooks SB. Regulator of G-protein signaling (RGS) proteins in cancer biology. *Biochem Pharmacol* (2009) 78(10):1289–97. doi: 10.1016/j.bcp.2009.06.028

35. DiGiacomo V, Maziarz M, Luebbers A, Norris JM, Laksono P, Garcia-Marcos M. Probing the mutational landscape of regulators of G protein signaling proteins in cancer. *Sci Signal* (2020) 13(617):eaax8620. doi: 10.1126/scisignal.aax8620

36. Lin C, Koval A, Tishchenko S, Gabdulkhakov A, Tin U, Solis GP, et al. Double suppression of the gα protein activity by RGS proteins. *Mol Cell* (2014) 53(4):663–71. doi: 10.1016/j.molcel.2014.01.014

37. Hurst Jillian H, Henkel Paul A, Brown April L, Hooks SB. Endogenous RGS proteins attenuate galpha(i)-mediated lysophosphatidic acid signaling pathways in ovarian cancer cells. *Cell Signal* (2008) 20(2):381–9. doi: 10.1016/j.cellsig.2007.10.026

38. Cha P-H, Cho Y-H, Lee S-K, Lee J, Jeong W-J, Moon B-S, et al. Small-molecule binding of the axin RGS domain promotes β-catenin and ras degradation. *Nat Chem Biol* (2016) 12(8):593–600. doi: 10.1038/nchembio.2103

39. Ahlers KE, Chakravarti B, Fisher RA. RGS6 as a novel therapeutic target in CNS diseases and cancer. *AAPS J* (2016) 18(3):560–72. doi: 10.1208/s12248-016-9899-9

40. Huang Di., Chen X, Zeng X, Lao L, Li J, Xing Y, et al. Targeting regulator of G protein signaling 1 in tumor-specific T cells enhances their trafficking to breast cancer. *Nat Immunol* (2021) 22(7):865–79. doi: 10.1038/s41590-021-00939-9

41. Wang L, Lau JS, Patra CR, Cao Y, Bhattacharya S, Dutta S, et al. RGS-GAIP-interacting protein controls breast cancer progression. *Mol Cancer Res* (2010) 8(12):1591–600. doi: 10.1158/1541-7786.MCR-10-0209

42. Asli A, Higazy-Mreih S, Avital-Shacham M, Kosloff M. Residue-level determinants of RGS R4 subfamily GAP activity and specificity towards the G subfamily. *Cell Mol Life Sci* (2021) 78(null):6305–18. doi: 10.1007/s00018-021-03898-4

43. Caballero-Franco C, Kissler S. The autoimmunity-associated gene RGS1 affects the frequency of T follicular helper cells. *Genes Immun* (2016) 17(4):228–38. doi: 10.1038/gene.2016.16

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 23 December 2022
DOI 10.3389/fimmu.2022.1096009

OPEN ACCESS

EDITED BY
Hao Zhang,
Chongqing Medical University, China

REVIEWED BY
Meng Hao,
Fudan University, China
Hui Wang,
University of Calgary, Canada

*CORRESPONDENCE
Jinglong Yan
yanjinglong2020@126.com

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Immunology

RECEIVED 11 November 2022
ACCEPTED 09 December 2022
PUBLISHED 23 December 2022

CITATION
Xu F, Yan J, Peng Z, Liu J and Li Z (2022) Comprehensive analysis of a glycolysis and cholesterol synthesis-related genes signature for predicting prognosis and immune landscape in osteosarcoma.
Front. Immunol. 13:1096009.
doi: 10.3389/fimmu.2022.1096009

Comprehensive analysis of a glycolysis and cholesterol synthesis-related genes signature for predicting prognosis and immune landscape in osteosarcoma

Fangxing Xu[1], Jinglong Yan[1]*, Zhibin Peng[2], Jingsong Liu[2] and Zecheng Li[1]

[1]Department of Orthopedics, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China, [2]Department of Orthopedics, The First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China

Background: Glycolysis and cholesterol synthesis are crucial in cancer metabolic reprogramming. The aim of this study was to identify a glycolysis and cholesterol synthesis-related genes (GCSRGs) signature for effective prognostic assessments of osteosarcoma patients.

Methods: Gene expression data and clinical information were obtained from GSE21257 and TARGET-OS datasets. Consistent clustering method was used to identify the GCSRGs-related subtypes. Univariate Cox regression and LASSO Cox regression analyses were used to construct the GCSRGs signature. The ssGSEA method was used to analyze the differences in immune cells infiltration. The pRRophetic R package was utilized to assess the drug sensitivity of different groups. Western blotting, cell viability assay, scratch assay and Transwell assay were used to perform cytological validation.

Results: Through bioinformatics analysis, patients diagnosed with osteosarcoma were classified into one of 4 subtypes (quiescent, glycolysis, cholesterol, and mixed subtypes), which differed significantly in terms of prognosis and tumor microenvironment. Weighted gene co-expression network analysis revealed that the modules strongly correlated with glycolysis and cholesterol synthesis were the midnight blue and the yellow modules, respectively. Both univariate and LASSO Cox regression analyses were conducted on screened module genes to identify 5 GCSRGs (RPS28, MCAM, EN1, TRAM2, and VEGFA) constituting a prognostic signature for osteosarcoma patients. The signature was an effective prognostic predictor, independent of clinical characteristics, as verified further via Kaplan-Meier analysis, ROC curve analysis, univariate and multivariate Cox regression analysis. Additionally, GCSRGs signature had strong correlation with drug sensitivity, immune checkpoints and immune cells infiltration. In cytological

experiments, we selected TRAM2 as a representative gene to validate the validity of GCSRGs signature, which found that TRAM2 promoted the progression of osteosarcoma cells. Finally, at the pan-cancer level, TRAM2 had been correlated with overall survival, progression free survival, disease specific survival, tumor mutational burden, microsatellite instability, immune checkpoints and immune cells infiltration.

Conclusion: Therefore, we constructed a GCSRGs signature that efficiently predicted osteosarcoma patient prognosis and guided therapy.

KEYWORDS

glycolysis, cholesterol, osteosarcoma, prognosis, signature, immune, TRAM2

1 Introduction

Osteosarcoma mostly occurs in the metaphysis of long bone and is the second leading factor of cancer deaths in children and adolescents (1, 2). Currently, surgical resection, chemotherapy, radiation therapy, hormone therapy, and small molecule targeted therapy are the mainstays in osteosarcoma treatment (3). Although the survival rate of osteosarcoma patients has been drastically increased with the combined chemotherapy, the 5-year survival rate is still not ideal for patients with distant metastasis, even with the use of large doses of adjuvant chemotherapy combined with radical resection (4). In addition, the psychological trauma caused by radical resection and the side effects of chemotherapy drugs are also problems that need to be addressed in the current treatment of osteosarcoma. To aid in improving osteosarcoma treatment, identifying novel therapeutic targets and biomarkers is crucial.

Unlike normal cells, cancerous cells often experience metabolic reprogramming. Metabolic reprogramming refers to the modifications to the tumor cells metabolic mode in the starvation state that allow adaption to the nutritional microenvironment; that is, to accommodate the requirements of their own quick growth through sufficient nutrients intake, metabolic reprogramming is a vital hallmark of malignant tumors (5). Glycolysis produces a small amount of energy during the entire glucose metabolism process. Normal cells mainly obtain energy through aerobic respiration. However, cancerous cells deviate from normal cells in various aspects. Even in an aerobic condition, cancerous cells favor the consumption of extra glucose for aerobic glycolysis in order for lactate production, a phenomenon referred to as Warburg effect (6). Calcium-binding protein A10 can accelerate glycolysis by mediating the AKT/mTOR signaling pathway in osteosarcoma, thereby enhancing malignancy of osteosarcoma cells (7). In addition, the novel lncRNA HCG18 enhances aerobic glycolysis in osteosarcoma cells *via* miR-365a-3p/PGK1 signaling pathway regulation, which accelerating the development of osteosarcoma cells (8). HIF-1α oncogene is present in numerous malignancies, including ovarian, breast, and bladder cancers, and can induce the glycolytic pathway in malignant tumors (9–11).

In recent years, the reprogramming of lipid synthesis has been considered to be another significant metabolic abnormality required for tumor growth, in which changes within the cholesterol biosynthetic pathway are vital (12). Cholesterol accumulation within cancerous cells can influence cell proliferation and metastasis, and enhance tumor microenvironmental adaptability, hence reinforcing tumor incidence and progression (13). Studies have demonstrated that several genes involved in cholesterol production are overactive in malignant tissue, such as squalene monooxygenase and the cholesterol biosynthesis rate-limiting enzyme 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase (HMGCR), which is upregulated within several types of malignancies, comprising glioma and prostate cancer (14, 15). HMGCR overexpression enhances cancer progression and metastasis, while its inhibition can suppress tumors; therefore, HMGCR has been used to treat solid cancers, hematological cancers, and tumors with drug resistance (16–18). In addition, the copy number of the SQLE locus encoding squalene monooxygenase is also increased in a variety of tumors. This copy number increase has been related to pancreatic cancer radiation tolerance and the development of several cancers within breast, prostate and colorectal cancer, or a poor patient prognosis (19, 20). However, similar to gene heterogeneity, tumor cell metabolism is also highly heterogeneous. In other words, no single universal change occurs within cancer metabolism. Tumorous metabolic changes are mainly characterized by changes in lipid and glucose metabolism. Recently, relevant research has discovered that changes in the

combined effects on lipid and glucose metabolism have become vital in pancreatic cancer, breast cancer, and skin malignant melanoma (21–23). High-throughput sequencing technologies are developing rapidly, and researchers possess the better understanding of pathogenic genes for various diseases, which is helpful for the discovery of novel biomarkers and pathogenic mechanisms (24). In recent years, differentially expressed genes have been screened through bioinformatics analysis to construct a prognostic signature for predicting osteosarcoma patient prognosis. For example, Zheng et al. constructed a prognostic signature and a nomogram relied on characteristics and clinical variables, which are used to screen out the tumor suppressor gene FHIT in osteosarcoma (25). However, to our knowledge, no gene signature related to glycolysis and cholesterol synthesis has been established to predict osteosarcoma patient prognosis.

During this research, relying on glycolysis and cholesterol synthesis-related genes (GCSRGs), osteosarcoma patients were categorized into one of 4 subtypes, and the differences in patient prognosis and tumor microenvironment between subtypes were also studied. A GCSRGs signature and an efficient nomogram were constructed by screening gene modules and their core genes for associations with glycolysis and cholesterol synthesis. In addition, the relationship of GCSRGs signature with drug sensitivity, immune infiltration and immune checkpoints was investigated, thereby expanding the genes signature's prognostic values for patients with osteosarcoma. Finally, we performed *in vitro* functional experiments and pan-cancer analysis to validate the genes of interest among the GCSRGs.

2 Materials and methods

2.1 Data download

GSE21257 dataset (n=53) was downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) and the Therapeutically Applicable Research to Generate Effective Treatment-Osteosarcoma (TARGET-OS) dataset (n=95) was obtained from the TARGET database (https://ocg.cancer.gov/programs/target). Both osteosarcoma datasets contain RNA sequences and clinical information. To obtain the total cohort dataset for subsequent mining, we combined TARGET-OS normalized by log2 of the transcript count per million (TPM) and GSE21257 with the batch effect removed by the ComBat function. Supplementary Table 1 illustrates all patients' clinical information in the total cohort. GCSRGs were obtained from the "REACTOME_GLYCOLYSIS" (n=72) and "REACTOME_CHOLESTEROL_BIOSYNTHESIS" (n=25) datasets in the Molecular Signatures Database (MSigDB) (https://www.gsea-msigdb.org/gsea/msigdb/). In addition, we downloaded the original pan-cancer mRNA matrix data, clinical data and copy number data from the University of California, Santa Cruz (UCSC) database (https://xenabrowser.net/).

2.2 Identification of the GCSRGs-related subtypes

Based on the expression of GCSRGs, the total cohort excluded metabolic genes with a standard deviation ≤ 0.5 and then used the genes as the main objects to perform consistent clustering using ConsensusClusterPlus R package to remove co-expressed metabolic genes and obtain co-expressed GCSRGs. The median expression level classified the metabolic subtypes, which were the quiescent type (glycolysis ≤ 0, cholesterol synthesis ≤ 0), glycolysis type (glycolysis > 0, cholesterol synthesis ≤ 0), cholesterol type (glycolysis > 0, cholesterol synthesis > 0), and mixed type (glycolysis > 0, cholesterol synthesis > 0). The prcomp function was used for principal component analysis (PCA) between subtypes, and survival R package and survminer R package analyzed survival differences between subtypes. The ESTIMATE algorithm calculated tumor purity, immune, stromal, and ESTIMATE scores in different subtypes.

2.3 Construction of weighted gene co-expression network and enrichment analysis

Weighted gene co-expression network analysis (WGCNA) employs gene expression data for scale-free network construction. For the top 25% of expression profiles in terms of variation coefficients, we built a network using the WGCNA R package. The modules strongly correlated with glycolysis and cholesterol subtype were screened, and the genes in the modules were pooled as key metabolic genes. Enrichment analysis of GO and KEGG pathway was conducted using clusterProfiler package.

2.4 Establishment and validation of a GCSRGs prognostic signature

To screen prognosis-related genes, in a random manner we categorized the total cohort into training and verification cohort, and utilized survival R package to do univariate Cox regression analysis upon the key modules' genes in training cohort. In order to further minimize the dimensionality and build the risk signature, least absolute shrinkage and selection operator (LASSO) Cox regression analysis has been conducted *via* glmnet R package and survminer R package, and patients' risk scores were then determined. The training, verification, and total cohorts were categorized into high- and low-risk groups based on risk score's median value. Survminer R package and

survivalROC R package generated survival and receiver operating characteristic (ROC) curves for the high- and low-risk groups. Area under curve (AUC) determined the signature's predictive ability. Once AUC > 0.6, signature became reliably predictive. We then performed univariate and multivariate Cox regression analyses to see if the risk score was an independent prognostic factor for osteosarcoma patients.

2.5 Nomogram construction and validation

The rms R package plotted the clinical nomogram. Performance of nomogram in predicting overall survival (OS) of osteosarcoma patients was evaluated using independent risk factors such as sex, age, metastatic status, and risk score. The calibration curve then proved the nomogram's efficacy.

2.6 Analysis of immune landscape and drug sensitivity

The single-sample gene set enrichment analysis (ssGSEA) method analyzed immune cells infiltration differences across the high- and low-risk groups. Differential expression analysis of immune checkpoints was used to assess the difference in the efficacy of immunotherapy. The pRRophetic R package was utilized to assess the drug sensitivity of different groups.

2.7 Pan-cancer analysis of TRAM2

To perform additional research into the role of TRAM2 in tumors, TRAM2 differential expression was assessed in pan-cancer, and we performed a correlation analysis of TRAM2 with patient prognosis, tumor mutational burden (TMB), and microsatellite instability (MSI). Furthermore, we performed a co-expression analysis of TRMA2 with immune cells and immune checkpoints.

2.8 Cell culture and transfection

All cell lines had been obtained from Procell (Wuhan, China). These cell lines were cultivated into DMEM/F12 medium containing 10% fetal bovine serum. TRAM2 siRNA and the corresponding si-control had been bought from GenePharma (Shanghai, China). Lipofectamine 3000 reagent (Invitrogen, California, USA) transfected cells as per the guidelines. After 48h of transfection, cells were utilized for protein quantification. The following sequences were utilized for the targeting of TRAM2: 5'-GCGUCCUCAUCGGGCUUAUTT-3' (si-TRAM2-1); 5'-CCUCGGUGAUUUGGUGCUUTT-3' (si-TRAM2-2); 5'-GCACGCACUUCCUGAGCUATT-3' (si-TRAM2-3).

2.9 Western blotting

In a nutshell, the protein samples were first isolated using SDS-PAGE. Later, proteins on the gel were moved to PVDF membrane and blocked. Primary antibodies were incubated overnight at a temperature of 4 °C, including anti-TRAM2 (Proteintech, 13311-1-AP, Wuhan, China), anti-E-cadherin (Proteintech, 20874-1-AP), anti-N-cadherin (Proteintech, 22018-1-AP), anti-Vimentin (Proteintech, 10366-1-AP), and anti-GAPDH (Zhongshanjinqiao, TA-08, Beijing, China). On day 2, the membrane underwent secondary antibody incubation. Next, enhanced chemiluminescence (ECL) color developing solution was utilized to develop the membrane after it had been rinsed with TBST three times.

2.10 Cell viability assay

The transfected cells have been cultured within 96-well plates at 5000 cells/well. Prior to Detection, Cell Counting Kit 8 (CCK8) reagent (Dojindo, Kumamoto, Japan) was added and incubated at 37 °C. A microplate reader took 450 nm absorbance readings once every 24 h up until 72 h.

In order to evaluate the osteosarcoma cells' capabilities for colony formation, a plate cloning assay was carried out. The transfected cells were evenly seeded in 6-well plate, and then cultured for 12 days with periodic replacements of the medium. Fixation and staining were accomplished with paraformaldehyde and crystal violet staining solution. A digital camera was used to snap photographs of the cells and recorded data.

2.11 Migration and invasion assays

To determine if osteosarcoma cells underwent migratory changes, a scratch assay was performed. 6-well plate was seeded with the transfected cells. After reaching 80% - 90% cell density, the cells were scratched using a pipettor tip oriented perpendicular to the plate's base. Results were photographed and recorded at 0 h and 48 h.

The invasive potential of osteosarcoma cells was measured using the Transwell assay. After pre-plating the Transwell chamber with Matrigel, the transfected cells were resuspended in fresh basal medium and added to the upper chamber. In the

lower chamber, we put in full medium. The upper chamber's cells were completely removed following 48 h. The remaining cells were stained after fixation, and photographed under a microscope.

2.12 Statistical analysis

GraphPad Prism 7 and R (version 3.6.3) were utilized throughout this investigation for all statistical testing and analysis. We used ClusterProfiler R package for consistent clustering. The Kaplan-Meier (KM) method was utilized for the survival analysis, and survival R package performed the log-rank test. In order to conduct LASSO analysis with cross-validation, the glmnet R package was used. The survminer R package and survival R package were used to create the ROC curve. Features selection was performed *via* univariate and multivariate Cox regression analyses. Wilcoxon test compared the continuous variables. Spearman correlation test was used for correlation analysis. $P < 0.05$ was considered statistically significant unless otherwise stated.

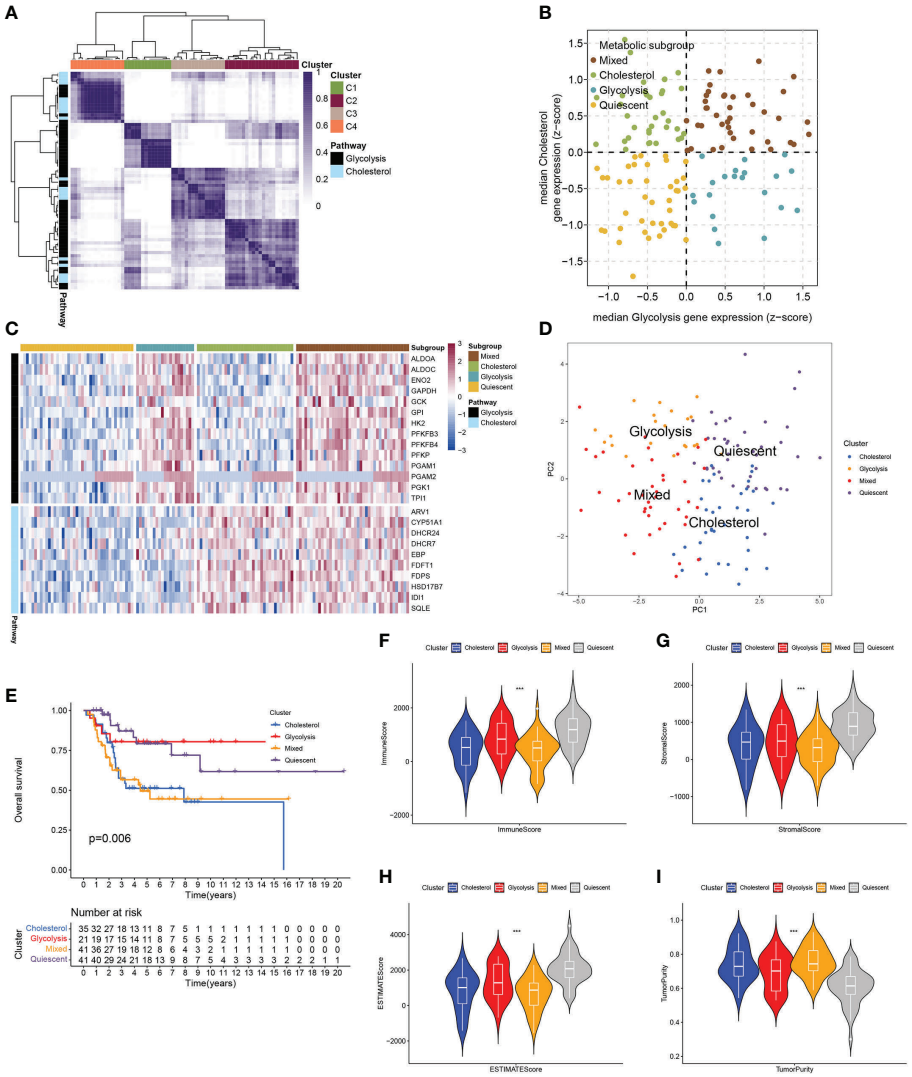


FIGURE 1
Classification of osteosarcoma patients based on expression of GCSRGs. **(A)** Heatmap showing consensus clustering solution for GCSRGs in osteosarcoma sample **(B)** Scatter plot depicting classification of samples based on GCSRGs expression. **(C)** Heatmap showing expression levels of co-expressed GCSRGs across each subgroup. **(D)** PCA showing significant differentiation between different subgroups of patients. **(E)** Kaplan-Meier survival curves of patients in the different subgroups. Log-rank test P values are displayed. **(F–I)** Violin plots showing the immune score, stromal score, ESTIMATE score and tumor purity across different metabolic subgroups. ***P < 0.001.

3 Results

3.1 Identification of the 4 subtypes of osteosarcoma patients by analysis of the expression of GCSRGs

The RNA-seq data and clinical information in the GSE21257 dataset and the TARGET-OS dataset were integrated after the batch effect was removed. The total cohort was obtained for subsequent analysis. Based on the gene sets of GCSRGs, metabolic-related genes with a standard deviation ≤0.5 were excluded from the total cohort. Then, consistent clustering was performed using the genes as the main body, thereby removing the co-expressed mixed metabolic genes C2 and C3, and the respective co-expressed metabolic genes were obtained including co-expressed glycolysis genes C1 and co-expressed cholesterol genes C4 (Figure 1A). We classified the total cohort into 4 metabolic subtypes based on the median expression levels of GCSRGs. Glycolysis ≤ 0 and cholesterol synthesis ≤ 0 was the quiescent subtype, glycolysis > 0 and cholesterol synthesis ≤ 0 was the glycolysis subtype, glycolysis ≤ 0 and cholesterol synthesis > 0 was the cholesterol subtype, and glycolysis > 0 and cholesterol synthesis > 0 was the mixed subtype (Figure 1B). Figure 1C

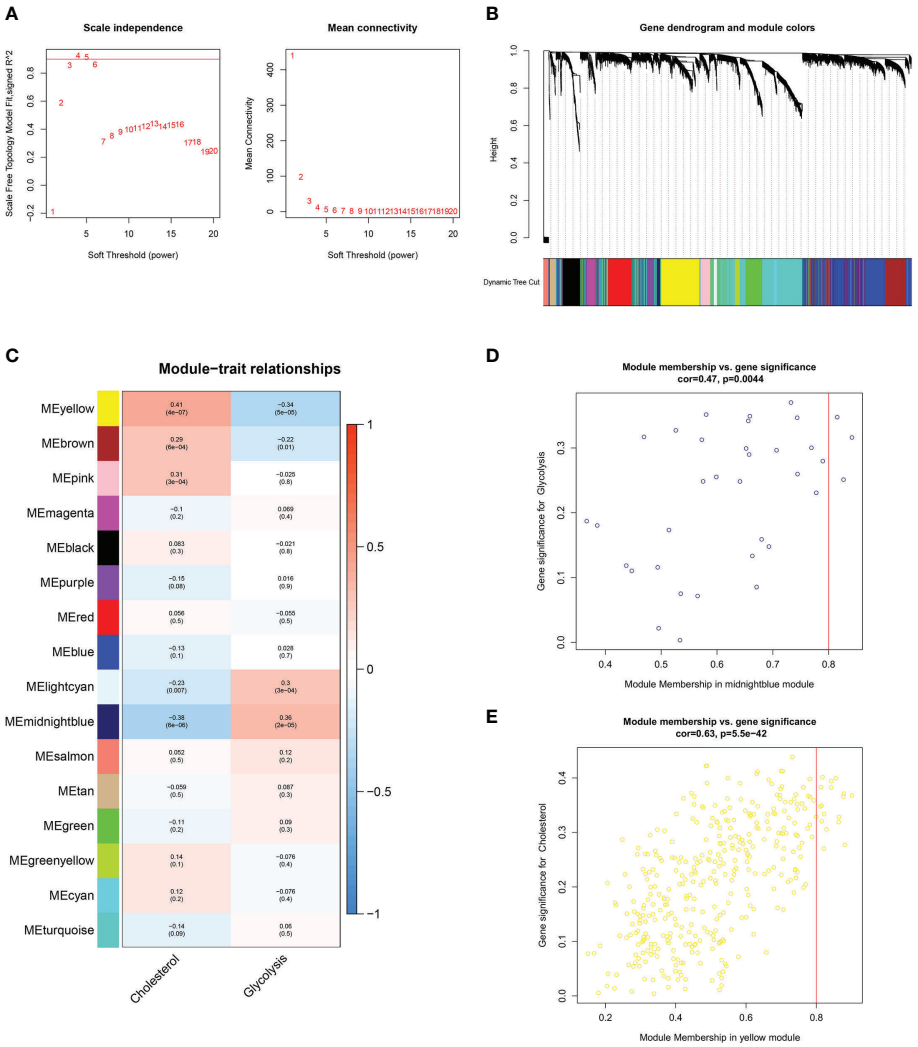


FIGURE 2
WGCNA to identify similar genes networks of GCSRGs. **(A)** The scale independence (left) and mean connectivity (right) of WGCNA analysis.**(B)** Color coding of co-expression network modules for genes. **(C)** Heatmap showing the correlation of gene modules and glycolysis-cholesterol synthesis. **(D)** Scatter plot displaying the correlation between module membership and gene significance in midnight blue network. **(E)** Scatter plot displaying the correlation between module membership and gene significance in yellow network.

illustrates the expression levels of GCSRGs in the 4 subtypes. According to the PCA of the 4 subtypes, the principal components of the 4 subtypes had a good degree of discrimination (Figure 1D). Further analysis of the differences in the prognosis between subtypes revealed the significant differences in the prognosis of different subtypes. Among them, prognosis for the glycolysis subtype was significantly better than the cholesterol subtype, and the quiescent subtype's prognosis was significantly better than the mixed subtype, and the mixed subtype's prognosis was similar to the cholesterol subtype (Figure 1E). In addition, to further analyze the differences in tumor microenvironment between different subtypes, ESTIMATE algorithm ranked the immune, stromal, and ESTIMATE scores as quiescent subtype > glycolysis subtype > cholesterol subtype > mixed subtype, but the reverse trend was noted for the tumor purity (Figures 1F–I).

3.2 GCSRGs co-expression network and biological activity

WGCNA was used to discover additional GCSRGs for further studies. The gene network achieved both high internal connectivity and gene similarity when the soft threshold was 4 (Figure 2A). Using hybrid dynamic shear tree, with a minimum of 25 genes per gene network module, 16 networks were found to be different from one another and were assigned distinct colors to represent them (Figure 2B). Then, the modules with strong correlations with glycolysis and cholesterol synthesis were screened, namely, the midnight blue and the yellow modules (Figure 2C). Among them, the glycolysis-related midnight blue module ($P = 0.0044$) contained 35 genes, and the cholesterol synthesis-related yellow module ($P < 0.001$) contained 367 genes. Figures 2D, E illustrates gene significance and module membership of the 2 modules. A robust positive relationship was identified between these variables' values.

A total of 402 genes within the midnight blue and yellow modules were pooled and used as key metabolic genes. The ClusterProfiler R package was conducted for GO and KEGG pathway enrichment analysis. The bubble plots showed the top 10 in GO-BP, GO-CC, and GO-MF and the top 7 in KEGG. GO functional annotation indicated that GCSRGs were mainly associated with hypoxia response, decreased oxygen response, focal adhesion, cell–substrate junction, ribosome, ribosome structural constituent, and monosaccharide binding (Figure 3A). KEGG functional annotation showed that GCSRGs were mainly associated with pathways including ribosome, HIF-1 signaling pathway, glycolysis/gluconeogenesis, and central carbon metabolism in cancer (Figure 3B).

3.3 Identification and construction of the GCSRGs signature to predict OS in osteosarcoma patients

The total cohort was categorized in a random manner into training and verification cohorts. Univariate Cox analysis was carried out on key metabolic genes (genes in the midnight blue and yellow modules) in the training cohort to screen prognosis-related genes *via* the survival R package, and 12 genes were related to patient prognosis ($P < 0.05$) (Figure 4A). As Figures 4B–G indicates, the Kaplan-Meier (KM) survival curves of the top 6 genes from low to high in terms of the P value were listed. Then, LASSO Cox regression analysis further reduced dimensionality and constructed genes signature. In the Cox regression based on the LASSO penalty, as log λ changed, the corresponding coefficient of the determined gene also decreased to 0, and in the cross-validation, 12 genes reached the partial likelihood estimation bias minimum value (Figures 4H, I). 5 genes were identified as independent predictors by LASSO Cox regression analysis in training cohort, namely, RPS28, MCAM, EN1, TRAM2, and VEGFA. We determined the risk scores *via* following formula: Risk score = RPS28 × 0.513 + MCAM × 0.701 - EN1 × 0.718 + TRAM2 × 0.575 + VEGFA × 0.467. The training, verification, and total cohorts were all categorized into high- and low-risk groups based on their median risk score. In each of the three cohorts, it was discovered that the low-risk group's survival probability was significantly greater than the other group ($P < 0.005$) (Figures 5A–C). Then, ROC curve analysis evaluated whether the GCSRGs signature is an efficient prognosis predictor of osteosarcoma patients. The 1-, 3-, and 5-year AUC predicted by the genes signature in training cohort were, 0.873, 0.889, and 0.856, respectively; in verification cohort, were 0.673, 0.810, and 0.823, respectively; in total cohort, were 0.747, 0.835, and 0.820, respectively (Figures 5D–F). In the low-risk group, the expression of 4 high-risk genes (RPS28, MCAM, TRAM2, and VEGFA) was low, while the low-risk gene EN1 expression was high (Figures 5G–I). Finally, we compared the survival status between the two groups in the three cohorts (Figures 5J–L) and plotted an expression heatmap of the risk genes (Figures 5M–O).

3.4 Independent prognostic analysis of the GCSRGs signature

To determine if the risk score and the other clinical characteristics are independent prognostic factors for osteosarcoma patients, univariate and multivariate Cox regression analyses were conducted. Univariate Cox regression analysis revealed the risk score ($P = 0.019$) and the clinical pathological parameters of metastasis ($P = 0.001$) were independent prognostic factors for osteosarcoma patients

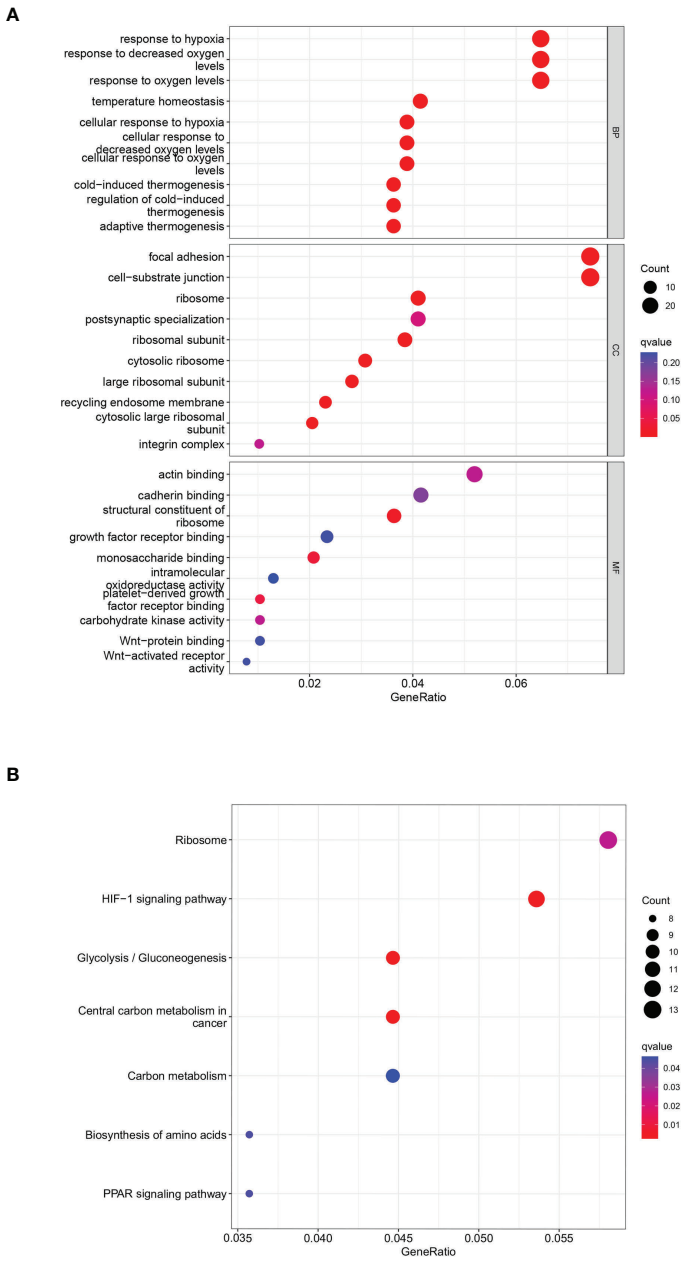


FIGURE 3
Functional enrichment analysis of genes in the screened modules. **(A)** The results of GO functional enrichment in GCSRGs. **(B)** The results of KEGG pathways enrichment in GCSRGs.

(Figure 6A), and multivariate Cox regression analysis showed the same results (Figure 6B). Furthermore, we developed a prognostic nomogram for estimating the osteosarcoma patients' survival likelihood (Figure 6C). This prognostic nomogram could systematically anticipate the 1-, 3-, and 5-year OS of osteosarcoma patients. The calibration curve showed that actual results were consistent with predicted results (Figure 6D).

3.5 Immune landscape and drug sensitivity analysis of the GCSRGs signature

For confirming if the GCSRGs signature was associated with tumor immunity, we used the ssGSEA method for evaluating differences in immune cells infiltration between the two groups. As Figure 7A indicates, the expression of

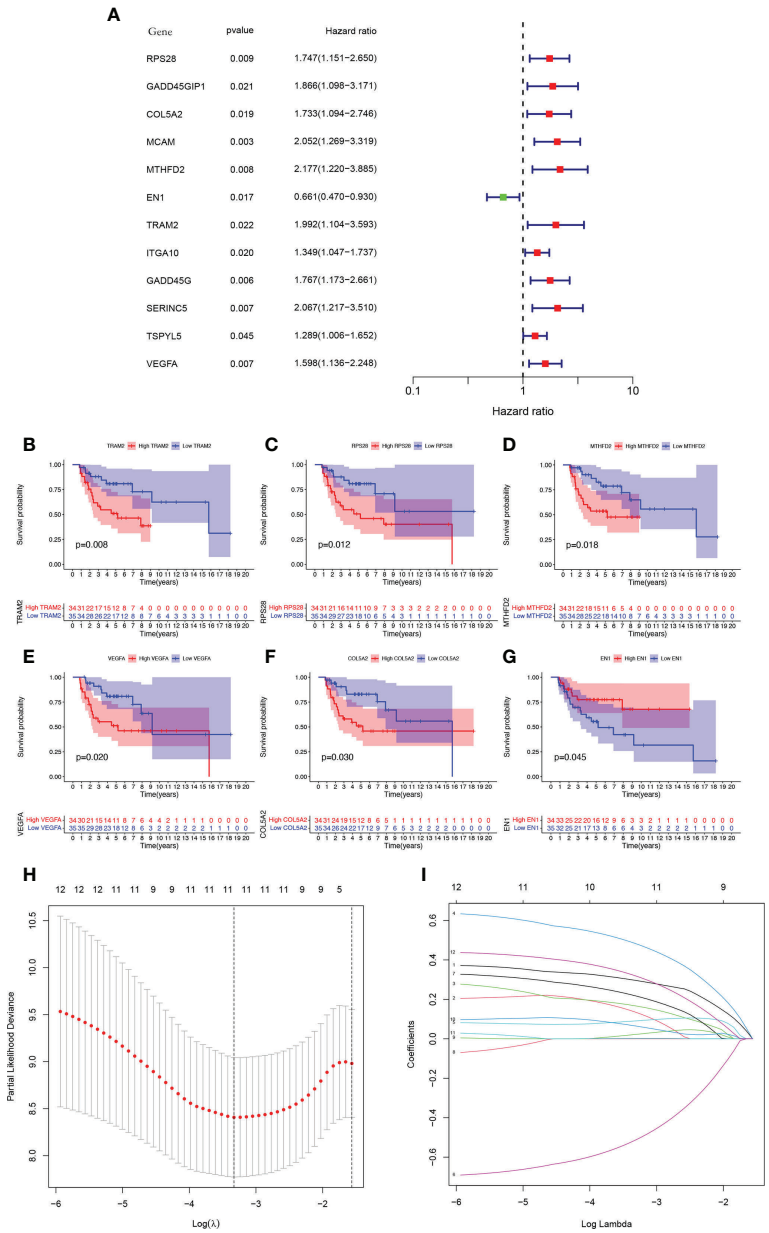


FIGURE 4
Construction of a GCSRGs prognostic signature in training cohort. **(A)** Forest plot of univariate cox regression analysis of the survival-related 12 differentially expressed genes. **(B–G)** Kaplan-Meier survival curves of patients with differential expression of prognosis-related genes. **(H)** Obtainment of the optimal λ value. **(I)** The LASSO Cox analysis identified 5 genes associated with prognosis.

eosinophils, macrophages, and natural killer cells had significant difference between the two groups. Among them, within the high-risk group, eosinophils proportion was significantly increased, while the opposite results occurred in macrophages and natural killer cell proportions. Additionally, as Figure 7B indicates, significant differences were found in immune checkpoints expression, including LGALS9, HAVCR2, LAIR1, TNFSF4, PDCD1LG2, TNFSF15, ICOS, CD200R1, TNFSF14, and BTLA between the two groups, with higher expression within the low-risk group than the other, pointing to the fact that there may be limited differences in the efficacy of immunotherapy. Drug sensitivity analysis indicated that 11 drugs were sensitive to patients in the high-risk group (Figure 7C), and 13 drugs were sensitive to patients in the low-risk group (Figure S1).

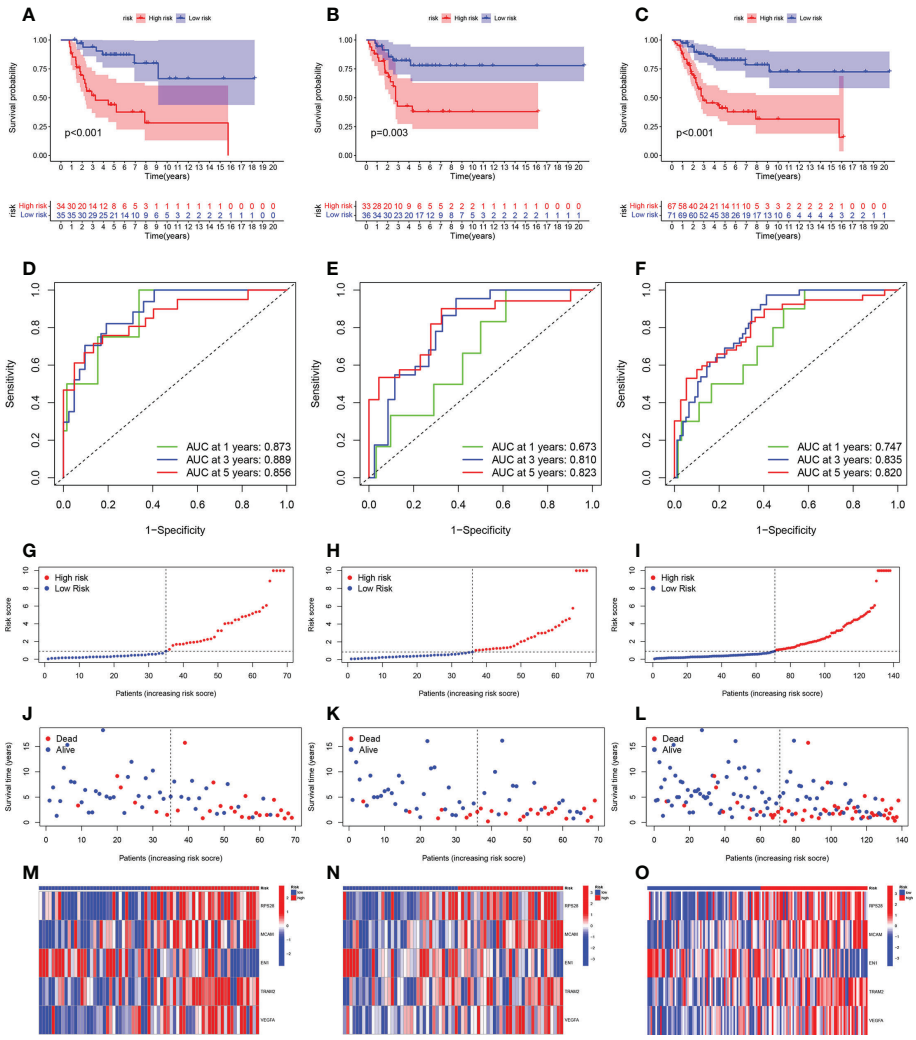


FIGURE 5
Prognostic value of the GCSRGs signature in training cohort, verification cohort and total cohort. **(A–C)** Kaplan-Meier survival curves according to risk score in the training cohort **(A)**, verification cohort **(B)**, and total cohort **(C)**. **(D–F)** ROC curves for predicting overall survival in the training cohort **(D)**, verification cohort **(E)**, and total cohort **(F)**. **(G–I)** Distribution of risk score in the high-risk group and the low-risk group in the training cohort **(G)**, verification cohort **(H)**, and total cohort **(I–L)** Survival status between the high-risk group and the low-risk group in the training cohort **(J)**, verification cohort **(K)**, and total cohort **(L-O)**.Heatmap of the expression profile of the included glycolysis-cholesterol synthesis related genes in the training cohort **(M)**, verification cohort **(N)**, and total cohort **(O)**.

3.6 Functional verification of TRAM2 *in vitro*

We searched the relevant literature of the aforementioned GCSRGs and found that TRAM2 was crucial in some malignancies. However, studies on the mechanism of TRAM2 action in osteosarcoma are scarce. Therefore, TRAM2 is expected to emerge as a promising new biological target in osteosarcoma treatment. Our study first revealed that TRAM2 expression in osteosarcoma cell lines was higher than the human osteoblast cell line according to Western blot results (Figure 8A). Then, si-TRAM2 was transferred to HOS and U2OS cell lines to discover the effect of TRAM2 on the osteosarcoma cell progression. Western blot results confirmed transfection efficiency (Figure 8B), and si-TRAM2-2 was chosen for further experiments. Based on CCK8 experiment results, TRAM2 downregulation inhibited HOS and U2OS cell lines viability (Figure 8C). According to the results of the plate cloning assay, downregulation of TRAM2 expression inhibited the colony-forming ability of the HOS and U2OS cell lines (Figure 8D). Furthermore, we conducted cell scratch and Transwell cell invasion assays. Experimental results indicated TRAM2 downregulation inhibited HOS and U2OS cell migration ability (Figure 8E) and invasion (Figure 8F). Prior studies have

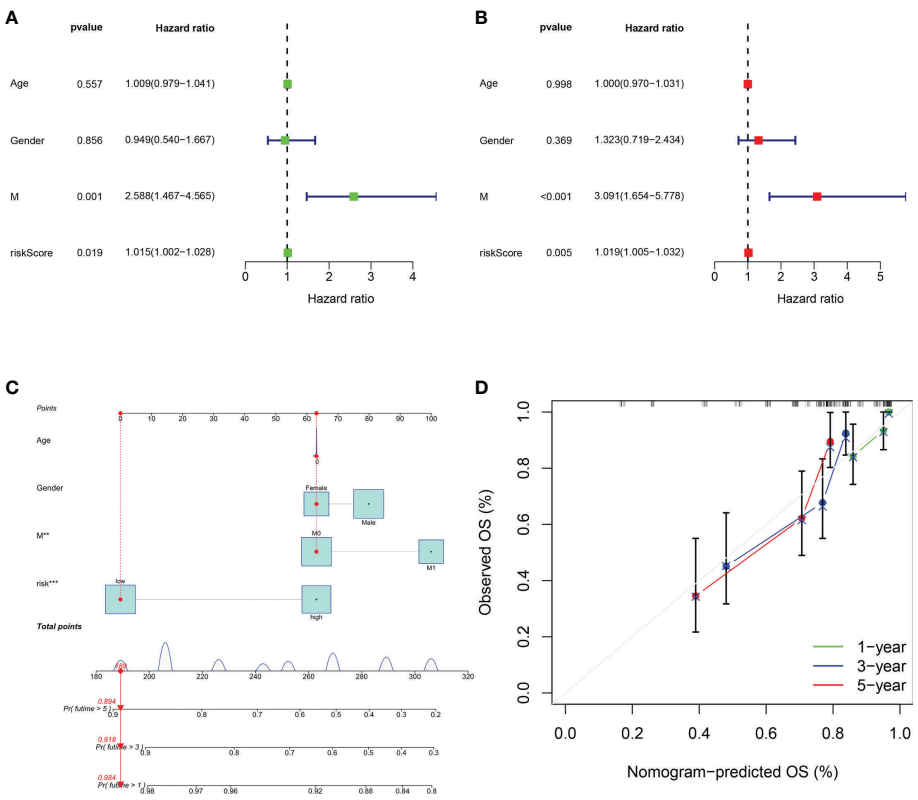


FIGURE 6
Assessment of the independent prognostic value and construction of the nomogram based on risk score and clinical factors. **(A)** Forest plot of univariate cox regression analysis of various clinical feature and risk score in osteosarcoma. **(B)** Forest plot of multivariate cox regression analysis of various clinical feature and risk score in osteosarcoma. **(C)** The nomogram to predict the 1-, 3- and 5-year survival risk of osteosarcoma patients. **(D)** Calibration curve for the 1-, 3-, and 5-year predicted survival nomogram. **P < 0.01, ***P < 0.001.

revealed that epithelial-mesenchymal transition (EMT) was vital in tumor progression and metastasis (26). So, we examined TRAM2 downregulation effect on EMT-related proteins expression. TRAM2 downregulation promoted E-cadherin expression while suppressing N-cadherin and vimentin expression in the HOS and U2OS cell lines, according to Western blot results (Figure 8G).

3.7 Pan-cancer analysis of TRAM2

To further analyze the important role of TRAM2 in other malignant tumors, we performed pan-cancer analysis of TRAM2. Figure 9A shows the expression of TRAM2 in 33 types of cancers, where TRAM2 had the highest expression in SARC. In addition, TRAM2 expression differed significantly between tumor tissues and normal paracancerous tissues in several types of cancer (Figure 9B). As shown in Figures 9C–E, TRAM2 was relevant to OS, progression free survival (PFS) and disease specific survival (DSS) in a range of cancers. Further analysis of the above data obtained KM survival curves (Figure S2). Moreover, TRAM2 was relevant to TMB and MSI in a range of cancers (Figures 9F, G). To elucidate the relationship of TRAM2 with immune-related genes and immune checkpoints, we conducted gene co-expression analysis. As Figures 9H, I illustrates, TRAM2 can affect immune cell infiltration and immune checkpoint expression in pan-cancer.

4 Discussion

Osteosarcoma is a highly invasive cancer. Its poor prognosis is related to problems with current treatments (27). Therefore, there is a need to develop and study prognostic models of osteosarcoma to guide targeted therapy. With the development of bioinformatics and sequencing technology, many scholars have constructed different prognostic models of osteosarcoma to analyze the characteristics of the disease (28–30). However, most of the parameters used to construct prognostic models consider only the genome or transcriptome and do not consider biological processes. As a result, osteosarcoma features cannot be represented accurately within these models. Recently, tumor

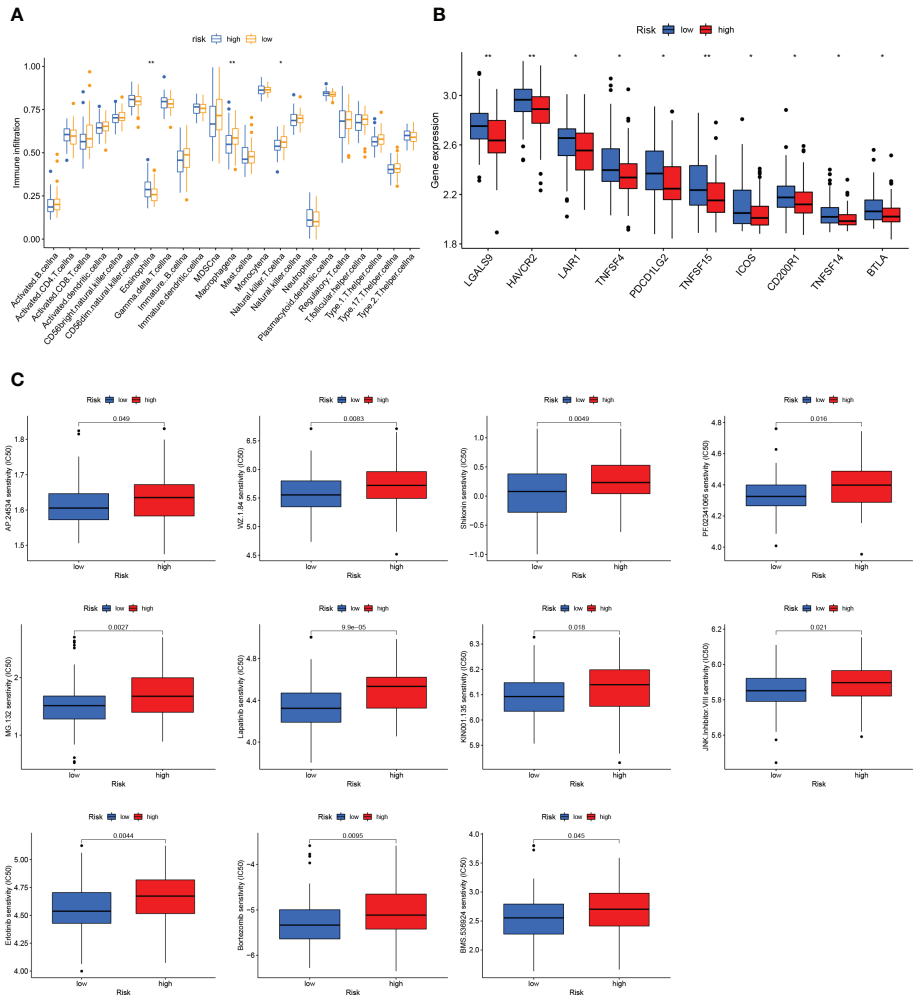


FIGURE 7
Immune status and drug sensitivity differences between high- and low-risk groups. **(A)** Comparison of immune cell infiltration between the high-risk group and low-risk group. **(B)** Comparison of the expression of immune checkpoints between the high-risk group and low-risk group. **(C)** Drug sensitivity in the high-risk group and low-risk group. *P < 0.05, **P < 0.01.

energy metabolism has attracted increasing interest. Glycolysis and cholesterol synthesis pathways are involved in the metabolic reprogramming of tumors and are crucial in tumor progression (31, 32). In our work, for the first time, we constructed a prognostic signature with glycolysis and cholesterol synthesis as the main characteristics, which can effectively predict osteosarcoma patient prognosis.

We first utilized consensus clustering to confirm the 2 groups of stable independent metabolic genes of glycolysis and cholesterol synthesis and then divided osteosarcoma patients into 4 subtypes (glycolysis subtype, cholesterol subtype, quiescent subtype, and mixed subtype) on basis of median gene expression. Survival across the subtypes showed significant differences based on the prognostic analysis, with the cholesterol subtype and the mixed subtype having the worst prognosis. Additionally, significant differences were observed in tumor purity, scores of immune, stroma, and ESTIMATE, which also confirmed prognosis differences in the 4 subtypes. We used WGCNA to screen out the modules related to glycolysis and cholesterol synthesis and conducted GO and KEGG enrichment analysis. Hypoxia is strongly correlated with poor prognosis, with its pathway activated throughout cancer advancement (33). The HIF-1 protein is heterodimeric with two different subunits, HIF-1α and HIF-1β. This protein activates several genes transcription that encode proteins engaged with angiogenesis, extracellular mesenchymal remodeling, migration, invasion, and metastasis (34). Consistent with the above conclusions, the results of enrichment analysis, such as response to hypoxia and decreased oxygen, and HIF-1 signaling pathway, indicated that this module's key metabolic genes had tight association with

10.3389/fimmu.2022.1096009

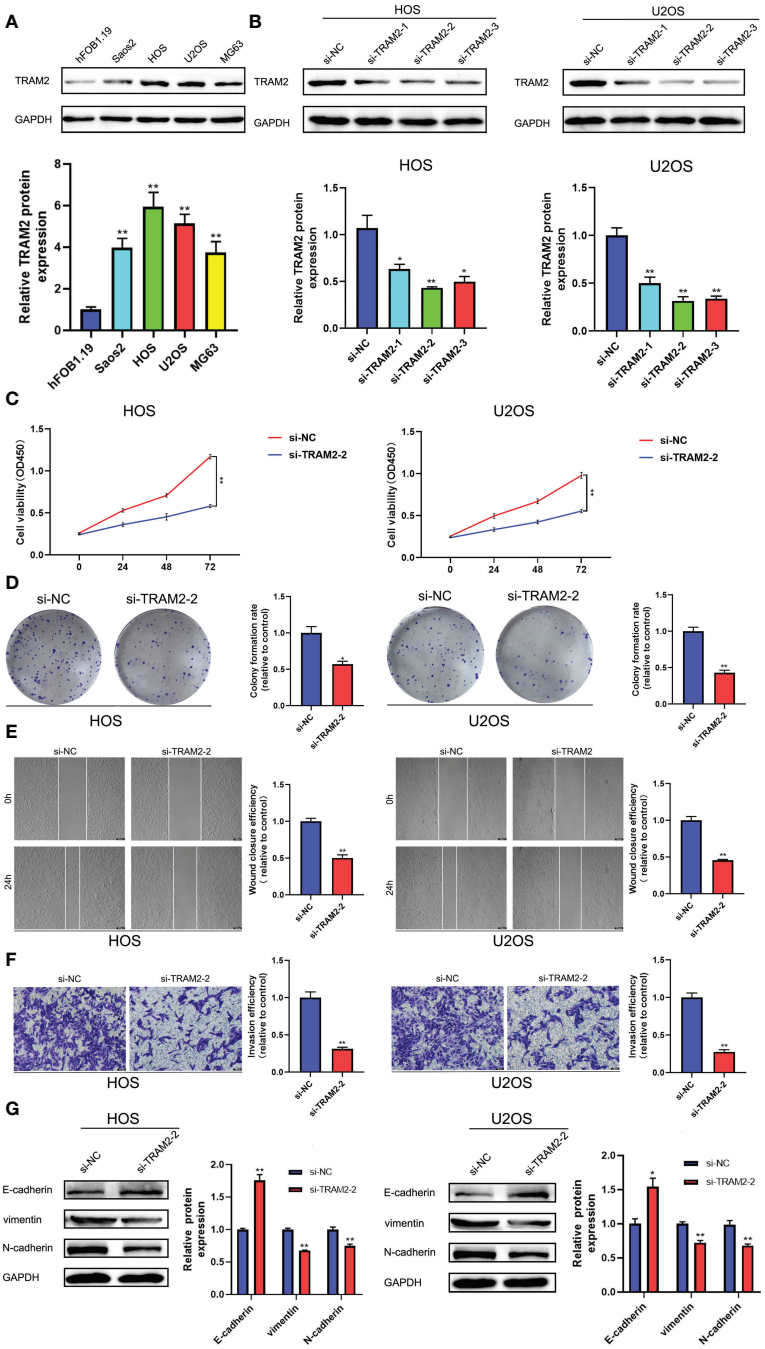


FIGURE 8
TRAM2 suppressed the progression of osteosarcoma *in vitro*. **(A)** The protein levels of TRAM2 in Saos2, HOS, U2OS, MG63 cells and normal hFOB1.19 cells. **(B)** The protein levels of TRAM2 in HOS and U2OS cells after transfection of si-NC, si-TRAM2-1, si-TRAM2-2 and si-TRAM2-3. **(C)** CCK-8 proliferation assay in HOS and U2OS cells after transfection of si-NC and si-TRAM2-2. **(D)** Plate cloning assay in HOS and U2OS cells after transfection of si-NC and si-TRAM2-2. **(E)** Scratch assay in HOS and U2OS cells after transfection of si-NC and si-TRAM2-2. **(F)** Transwell assay in HOS and U2OS cells after transfection of si-NC and si-TRAM2-2. **(G)** The protein levels of EMT-related proteins including E-cadherin, vimentin and N-cadherin in HOS and U2OS cells after transfection of si-NC and si-TRAM2-2. All results are presented as mean ± SEM. *P < 0.05, **P < 0.01.

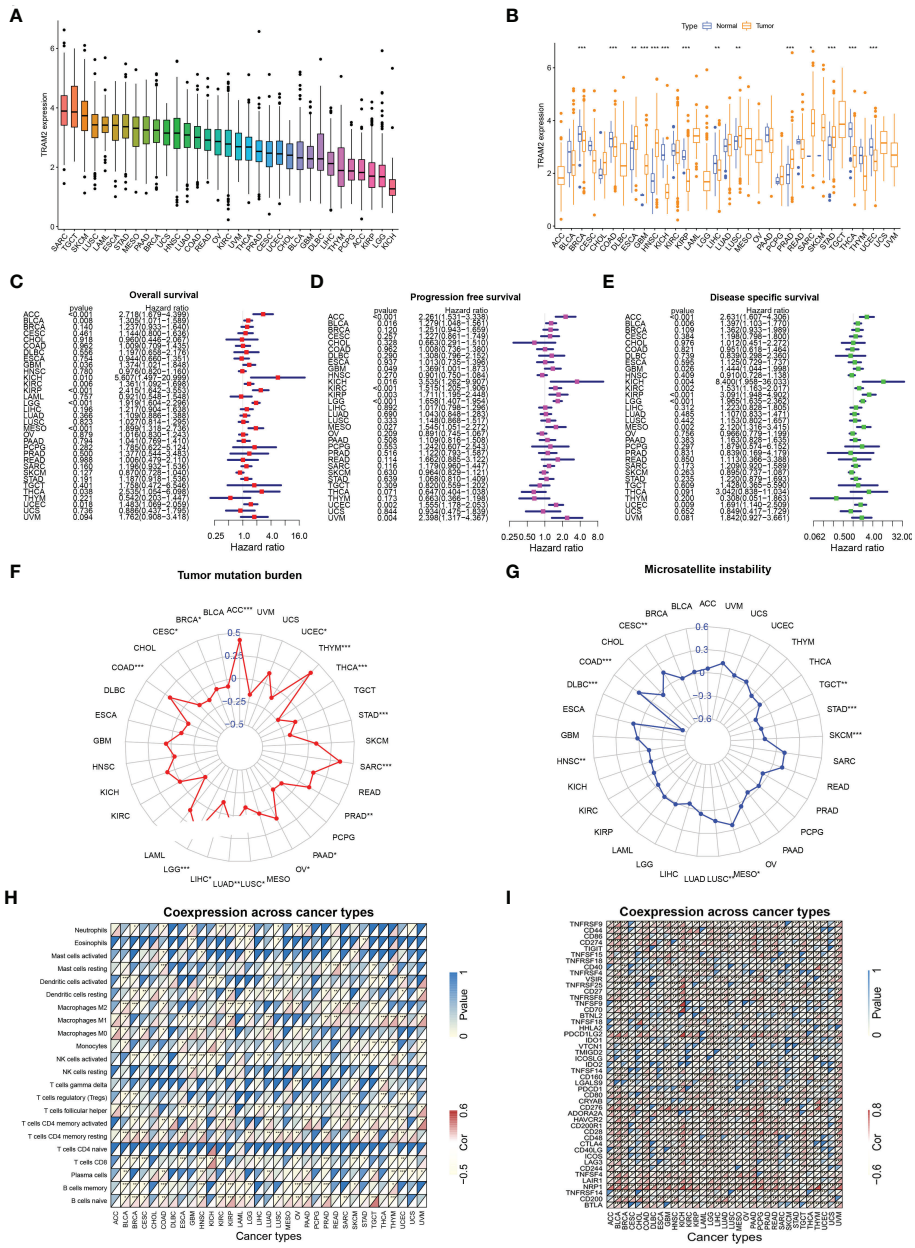


FIGURE 9
Analysis of TRAM2 in pan-cancer. **(A)** Expression of TRAM2 in 33 cancers. **(B)** Expression of TRAM2 in tumor and normal tissue in pan-cancer. **(C)** Overall survival of TRAM2 in pan-cancer. **(D)** Progression free survival of TRAM2 in pan-cancer. **(E)** Disease specific survival of TRAM2 in pan-cancer. **(F)** Tumor mutation burden of TRAM2 in pan-cancer. **(G)** Microsatellite instability of TRAM2 in pan-cancer. **(H)** Co-expression analysis of TRAM2 and immune cells in pan-cancer. **(I)** Co-expression analysis of TRAM2 and immune checkpoints in pan-cancer. *P < 0.05, **P < 0.01, ***P < 0.001.

hypoxia process. Previous studies have demonstrated that focal adhesions, as mediators of tumor cells and the extracellular matrix, are vital in various ways within tumor migration, invasion, and drug resistance (35). The results of GO enrichment analysis, such as focal adhesion and cell-substrate junction, indicated that the key metabolic genes in the module may be closely associated with metastasis. Subsequently, univariate Cox and LASSO Cox regression analysis had been conducted on key metabolic genes in the selected modules, and 5 genes (RPS28, MCAM, EN1, TRAM2, and VEGFA) were screened as relevant genes for the GCSRGs signature construction. The GCSRGs signature had good predictive ability in all cohorts and can be utilized as an independent prognostic factor for osteosarcoma patients. Several researchers

have investigated the relationship among glycolysis, cholesterol synthesis and immune responses. Regulating cholesterol metabolism can improve CD8 (+) T cells' anticancer effect (36). Additionally, Li et al. indicated that the glycolysis process of tumor tissues within breast cancer had association with low natural killer T (NKT) cells infiltration (37). In our study, macrophages and NKT cells expression levels within the low-risk group were significantly higher than the other group. According to our knowledge, NKT cells are crucial for controlling tumor cell progression and affecting cancer patient prognosis (38). For macrophages, high infiltration of tumor-associated macrophages in some malignant tumors has a strong correlation with better prognosis (39, 40). That's consistent with our study findings and helps explain, to a certain extent, why patients who were classified as low-risk group had superior survival outcomes. In addition, our study found that the total 10 immune checkpoint genes expression showed different levels between the two groups, with low-risk group showing higher expression than the other group, indicating that there may be limited differences in the efficacy of immunotherapy.

In our analysis, we selected 5 GCSRGs (RPS28, MCAM, EN1, TRAM2, and VEGFA) as the relevant genes for constructing the risk genes signature. RPS28 is a 40S ribosome component and is critical for 18S rRNA biosynthesis (41). There are few studies on the effect of RPS28 on cancer, and most research results are only predictions generated by bioinformatics and have not been confirmed by corresponding biological experiments (42, 43). However, some researchers have found that reducing the expression of RPS28 protein can reduce the cell viability of HeLa cells and induce tumor cell apoptosis (44), indicating that RPS28 has a major regulatory function in cancer. Additionally, RPS28 can influence tumor immunosurveillance and regulate T cell killing (45). MCAM is highly expressed in various malignancies and has tight association with their growth and metastasis, such as melanoma (46), prostate cancer (47), gastric cancer (48), and lung cancer (49). Prior investigations revealed that MCAM was associated with poor prognosis of osteosarcoma patients and can improve the migration ability of osteosarcoma cells (50). For immunotherapy, MCAM deficiency significantly impairs T cell-mediated antitumor effect (51). Solid tumor progression and metastasis are accompanied by angiogenesis stimulation, with VEGFA as the main factor driving tumor vascular bed expansion (52). VEGFA is involved in angiogenesis, progression, and metastasis in various malignancies, including osteosarcoma, and has a strong association with a poor prognosis (53–55). Moreover, the expression of co-inhibitory receptor and regulatory T cell expansion are both influence by VEGFA signaling (56). Hence, targeted VEGFA therapy is a key area for improving the osteosarcoma prognosis (57). TRAM2 is a translocon component and can transport proteins synthesized by ribosomes to the endoplasmic reticulum (ER), acting as ER channels for calcium concentration regulation within it (58). In glioma, through its PI3K/AKT/mTOR signaling pathway regulation, TRAM2 is able to enhance tumor cells migration, invasion, proliferation, and EMT (59). In addition, TRAM2 and YAP activity in various cancers shows a very strong expression correlation, demonstrating that TRAM2 acts a significant role in malignant proliferation and invasion caused by YAP (60). However, no relevant studies have shown the relationship between TRAM2 and osteosarcoma. Therefore, to ensure the validity of the GCSRGs signature, we chose to use TRAM2 for cell function validation and pan-cancer analysis.

TRAM2 protein expression was demonstrated to be significantly different across the osteosarcoma and the human osteoblast cell lines during experimental validation. In addition, inhibiting of EMT-related protein expression, cell viability, colony formation, migration, and invasion were achieved by downregulating TRAM2 protein expression in osteosarcoma cells. These findings provide further support for validity of genes signature based on glycolysis and cholesterol synthesis and suggest that TRAM2 is involved in osteosarcoma cells progression. In addition, TRAM2 was not only involved in osteosarcoma progression but also closely related to OS, PFS, DSS, TMD, MSI, immune cell infiltration and immune checkpoints in pan-cancer, suggesting that the GCSRGs signature and the target genes in the signature have the potential to serve as the prognostic indicators for a wide range of cancers.

Although we confirmed the effective role of the GCSRGs signature in predicting the prognosis of osteosarcoma patients and confirmed the tumor-promoting effect of TRAM2 in osteosarcoma cells in cytological experiments *in vitro*, this study still has certain drawbacks that require further research. First, the patient sample size was small within the datasets used, and their clinical characteristics were not sufficiently detailed. Therefore, a larger sample size with more detailed clinical characteristics is needed. In addition, besides TRAM2, other signature-related genes should also be verified at the cytological level.

During this research, osteosarcoma patients were categorized into 4 subtypes according to GCSRGs expression matrix, and these subtypes differed significantly from one another in terms of prognosis and tumor microenvironment. Through WGCNA, the gene modules most closely associated with glycolysis and cholesterol synthesis were screened, and a risk signature of osteosarcoma consisting of 5 GCSRGs was constructed for the first time. In addition, we found that this signature was closely related to immune cells infiltration and immune checkpoint expression in osteosarcoma patients. These findings not only provide a new method to predict the prognosis of osteosarcoma patients but also provide novel therapeutic targets.

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 and FX designed the research and drafted the manuscript. FX and ZP completed the bioinformatics analysis. JL and ZL completed cytology experiments. The final manuscript has been read by all of the authors, and they have all given their approval for it to be submitted.

Funding

The National Natural Science Foundation of China (No.82072472) provided financial backing for this investigation.

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.2022.1096009/full#supplementary-material

SUPPLEMENTARY FIGURE 1
The 13 drugs were selected for osteosarcoma patients with the low-risk group

SUPPLEMENTARY FIGURE 2
Correlation between the expression of TRAM2 and survival prognosis in pan-cancer **(A)** Kaplan-Meier survival curves of the relationship between TRAM2 expression and overall survival in pan-cancer. **(B)** Kaplan-Meier survival curves of the relationship between TRAM2 expression and progression free survival in pan-cancer. **(C)** Kaplan-Meier survival curves of the relationship between TRAM2 expression and Disease specific survival in pan-cancer.

References

1. Xu F, Sun G, Peng Z, Liu J, Li Z, Yan J. Cannabidiol promotes apoptosis of osteosarcoma cells and by activating the Sp1-Cbx2 axis. *Am J Transl Res* (2022) 14 (2):1188–203.

2. Liu Y, Qiu G, Luo Y, Li S, Xu Y, Zhang Y, et al. Circular rna Rock1, a novel circrna, suppresses osteosarcoma proliferation and migration *Via* altering the mir-532-5p/Pten axis. *Exp Mol Med* (2022) 54(7):1024–37. doi: 10.1038/s12276-022-00806-z

3. Pan F, Zhang J, Tang B, Jing L, Qiu B, Zha Z. The novel Circ_0028171/Mir-218-5p/Ikbkb axis promotes osteosarcoma cancer progression. *Cancer Cell Int* (2020) 20:484. doi: 10.1186/s12935-020-01562-8

4. Huang K, Chen Y, Zhang R, Wu Y, Ma Y, Fang X, et al. Honokiol induces apoptosis and autophagy *Via* the Ros/Erk1/2 signaling pathway in human osteosarcoma cells in vitro and in vivo. *Cell Death Dis* (2018) 9(2):157. doi: 10.1038/s41419-017-0166-5

5. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013

6. Yang Y, Zhang Y, Miao L, Liao W, Liao W. Lncrna Ppp1r14b-As1 promotes tumor cell proliferation and migration *Via* the enhancement of mitochondrial respiration. *Front Genet* (2020) 11:557614. doi: 10.3389/fgene.2020.557614

7. Ling F, Lu Q. S100 calcium-binding protein A10 contributes to malignant traits in osteosarcoma cells by regulating glycolytic metabolism *Via* the Akt/Mtor pathway. *Bioengineered* (2022) 13(5):12298–308. doi: 10.1080/21655979.2022.2071022

8. Pan X, Guo J, Liu C, Pan Z, Yang Z, Yao X, et al. Lncrna Hcg18 promotes osteosarcoma growth by enhanced aerobic glycolysis *Via* the mir-365a-3p/Pgk1 axis. *Cell Mol Biol Lett* (2022) 27(1):5. doi: 10.1186/s11658-021-00304-6

9. Lyu Y, Zhang Y, Wang Y, Luo Y, Ding H, Li P, et al. Hif-1 regulated wtap overexpression promoting the warburg effect of ovarian cancer by M6a-dependent manner. *J Immunol Res* (2022) 2022:6130806. doi: 10.1155/2022/6130806

10. Wang J-Z, Zhu W, Han J, Yang X, Zhou R, Lu H-C, et al. The role of the hif-1α/Alyref/Pkm2 axis in glycolysis and tumorigenesis of bladder cancer. *Cancer Commun (Lond)* (2021) 41(7):560–75. doi: 10.1002/cac2.12158

11. Du Y, Wei N, Ma R, Jiang S, Song D. A mir-210-3p regulon that controls the warburg effect by modulating hif-1α and P53 activity in triple-negative breast cancer. *Cell Death Dis* (2020) 11(9):731. doi: 10.1038/s41419-020-02952-6

12. Cai D, Wang J, Gao B, Li J, Wu F, Zou JX, et al. Rorγ is a targetable master regulator of cholesterol biosynthesis in a cancer subtype. *Nat Commun* (2019) 10 (1):4621. doi: 10.1038/s41467-019-12529-3

13. Yue S, Li J, Lee S-Y, Lee HJ, Shao T, Song B, et al. Cholesteryl ester accumulation induced by pten loss and Pi3k/Akt activation underlies human prostate cancer aggressiveness. *Cell Metab* (2014) 19(3):393–406. doi: 10.1016/j.cmet.2014.01.019

14. Ashida S, Kawada C, Inoue K. Stromal regulation of prostate cancer cell growth by mevalonate pathway enzymes Hmgcs1 and hmgcr. *Oncol Lett* (2017) 14 (6):6533–42. doi: 10.3892/ol.2017.7025

15. Qiu Z, Yuan W, Chen T, Zhou C, Liu C, Huang Y, et al. Hmgcr positively regulated the growth and migration of glioblastoma cells. *Gene* (2016) 576(1 Pt 1):22–7. doi: 10.1016/j.gene.2015.09.067

16. Kong Y, Cheng L, Mao F, Zhang Z, Zhang Y, Farah E, et al. Inhibition of cholesterol biosynthesis overcomes enzalutamide resistance in castration-resistant prostate cancer (Crpc). *J Biol Chem* (2018) 293(37):14328–41. doi: 10.1074/jbc.RA118.004442

17. Lee JS, Roberts A, Juarez D, Vo T-TT, Bhatt S, Herzog L-O, et al. Statins enhance efficacy of venetoclax in blood cancers. *Sci Transl Med* (2018) 10(445): eaaq1240. doi: 10.1126/scitranslmed.aaq1240

18. Yang J, Wang L, Jia R. Role of cholesterol synthesis enzymes in cancer. *J Cancer* (2020) 11(7):1761–7. doi: 10.7150/jca.38598

19. Brown DN, Caffa I, Cirmena G, Piras D, Garuti A, Gallo M, et al. Squalene epoxidase is a bona fide oncogene by amplification with clinical relevance in breast cancer. *Sci Rep* (2016) 6:19435. doi: 10.1038/srep19435

20. Cirmena G, Franceschelli P, Isnaldi E, Ferrando L, De Mariano M, Ballestrero A, et al. Squalene epoxidase as a promising metabolic target in cancer treatment. *Cancer Lett* (2018) 425:13–20. doi: 10.1016/j.canlet.2018.03.034

21. Karasinska JM, Topham JT, Kalloger SE, Jang GH, Denroche RE, Culibrk L, et al. Altered gene expression along the glycolysis-cholesterol synthesis axis is associated with outcome in pancreatic cancer. *Clin Cancer Res* (2020) 26(1):135–46. doi: 10.1158/1078-0432.CCR-19-1543

22. Zhong P-C, Shu R, Wu H-W, Liu Z-W, Shen X-L, Hu Y-J. Altered gene expression in glycolysis-cholesterol synthesis axis correlates with outcome of triple-negative breast cancer. *Exp Biol Med (Maywood)* (2021) 246(5):560–71. doi: 10.1177/1535370220975206

23. Zhang E, Chen Y, Bao S, Hou X, Hu J, Mu OYN, et al. Identification of subgroups along the glycolysis-cholesterol synthesis axis and the development of an associated prognostic risk model. *Hum Genomics* (2021) 15(1):53. doi: 10.1186/s40246-021-00350-3

24. Jia Y, Liu Y, Han Z, Tian R. Identification of potential gene signatures associated with osteosarcoma by integrated bioinformatics analysis. *PeerJ* (2021) 9: e11496. doi: 10.7717/peerj.11496

25. Zheng D, Xia K, Yu L, Gong C, Shi Y, Li W, et al. A novel six metastasis-related prognostic gene signature for patients with osteosarcoma. *Front Cell Dev Biol* (2021) 9:699212. doi: 10.3389/fcell.2021.699212

26. Dudas J, Ladanyi A, Ingruber J, Steinbichler TB, Riechelmann H. Epithelial to mesenchymal transition: A mechanism that fuels cancer Radio/Chemoresistance. *Cells* (2020) 9(2):428. doi: 10.3390/cells9020428

27. Somarelli JA, Rupprecht G, Altunel E, Flamant EM, Rao S, Sivaraj D, et al. A comparative oncology drug discovery pipeline to identify and validate new treatments for osteosarcoma. *Cancers (Basel)* (2020) 12(11):3335. doi: 10.3390/cancers12113335

28. Luo X, Tang J, Xuan H, Liu J, Li X. Identification and validation of a potent multi-mirna signature for prediction of prognosis of osteosarcoma patients. *Med Sci Monit* (2020) 26:e919272. doi: 10.12659/MSM.919272

29. Yang J, Zhang A, Luo H, Ma C. Construction and validation of a novel gene signature for predicting the prognosis of osteosarcoma. *Sci Rep* (2022) 12(1):1279. doi: 10.1038/s41598-022-05341-5

30. Zhang Y, Lei X, He R, Mao L, Jiang P, Ni C, et al. Identification and preliminary validation of a four-gene signature to predict metastasis and survival in osteosarcoma. *Am J Transl Res* (2021) 13(11):12264–84.

31. Lu J, Tan M, Cai Q. The warburg effect in tumor progression: Mitochondrial oxidative metabolism as an anti-metastasis mechanism. *Cancer Lett* (2015) 356(2 Pt A):156–64. doi: 10.1016/j.canlet.2014.04.001

32. Xu H, Zhou S, Tang Q, Xia H, Bi F. Cholesterol metabolism: New functions and therapeutic approaches in cancer. *Biochim Biophys Acta Rev Cancer* (2020) 1874(1):188394. doi: 10.1016/j.bbcan.2020.188394

33. Boutilier AJ, Elsawa SF. Macrophage polarization states in the tumor microenvironment. *Int J Mol Sci* (2021) 22(13):6995. doi: 10.3390/ijms22136995

34. Schito L, Semenza GL. Hypoxia-inducible factors: Master regulators of cancer progression. *Trends Cancer* (2016) 2(12):758–70. doi: 10.1016/j.trecan.2016.10.016

35. Mao D, Xu R, Chen H, Chen X, Li D, Song S, et al. Cross-talk of focal adhesion-related gene defines prognosis and the immune microenvironment in gastric cancer. *Front Cell Dev Biol* (2021) 9:716461. doi: 10.3389/fcell.2021.716461

36. Yang W, Bai Y, Xiong Y, Zhang J, Chen S, Zheng X, et al. Potentiating the antitumour response of Cd8(+) T cells by modulating cholesterol metabolism. *Nature* (2016) 531(7596):651–5. doi: 10.1038/nature17412

37. Li W, Xu M, Li Y, Huang Z, Zhou J, Zhao Q, et al. Comprehensive analysis of the association between tumor glycolysis and Immune/Inflammation function in breast cancer. *J Transl Med* (2020) 18(1):92. doi: 10.1186/s12967-020-02267-2

38. Webb TJ, Yuan W, Meyer E, Dellabona P. Editorial: Nkt cells in cancer immunotherapy. *Front Immunol* (2020) 11:1314. doi: 10.3389/fimmu.2020.01314

39. Forssell J, Oberg A, Henriksson ML, Stenling R, Jung A, Palmqvist R. High macrophage infiltration along the tumor front correlates with improved survival in colon cancer. *Clin Cancer Res* (2007) 13(5):1472–9. doi: 10.1158/1078-0432.CCR-06-2073

40. Di Caro G, Cortese N, Castino GF, Grizzi F, Gavazzi F, Ridolfi C, et al. Dual prognostic significance of tumour-associated macrophages in human pancreatic adenocarcinoma treated or untreated with chemotherapy. *Gut* (2016) 65(10):1710–20. doi: 10.1136/gutjnl-2015-309193

41. Robledo S, Idol RA, Crimmins DL, Ladenson JH, Mason PJ, Bessler M. The role of human ribosomal proteins in the maturation of rrna and ribosome production. *RNA* (2008) 14(9):1918–29. doi: 10.1261/rna.1132008

42. Yau C, Esserman L, Moore DH, Waldman F, Sninsky J, Benz CC. A multigene predictor of metastatic outcome in early stage hormone receptor-negative and triple-negative breast cancer. *Breast Cancer Res* (2010) 12(5):R85. doi: 10.1186/bcr2753

43. Huang Z, Lan T, Wang J, Chen Z, Zhang X. Identification and validation of seven rna binding protein genes as a prognostic signature in oral cavity squamous cell carcinoma. *Bioengineered* (2021) 12(1):7248–62. doi: 10.1080/21655979.2021.1974328

44. Kim HK, Fuchs G, Wang S, Wei W, Zhang Y, Park H, et al. A transfer-Rna-Derived small rna regulates ribosome biogenesis. *Nature* (2017) 552(7683):57–62. doi: 10.1038/nature25005

45. Wei J, Kishton RJ, Angel M, Conn CS, Dalla-Venezia N, Marcel V, et al. Ribosomal proteins regulate mhc class I peptide generation for immunosurveillance. *Mol Cell* (2019) 73(6):1162–1173.e5. doi: 10.1016/j.molcel.2018.12.020

46. Ruma IMW, Putranto EW, Kondo E, Murata H, Watanabe M, Huang P, et al. Mcam, as a novel receptor for S100a8/A9, mediates progression of malignant melanoma through prominent activation of nf-κb and ros formation upon ligand binding. *Clin Exp Metastasis* (2016) 33(6):609–27. doi: 10.1007/s10585-016-9801-2

47. Zoni E, Astrologo L, Ng CKY, Piscuoglio S, Melsen J, Grosjean J, et al. Therapeutic targeting of Cd146/Mcam reduces bone metastasis in prostate cancer. *Mol Cancer Res* (2019) 17(5):1049–62. doi: 10.1158/1541-7786.MCR-18-1220

48. Chen J, Dang Y, Feng W, Qiao C, Liu D, Zhang T, et al. Sox18 promotes gastric cancer metastasis through transactivating mcam and Ccl7. *Oncogene* (2020) 39(33):5536–52. doi: 10.1038/s41388-020-1378-1

49. Zhang X, Wang Z, Kang Y, Li X, Ma X, Ma L. Mcam expression is associated with poor prognosis in non-small cell lung cancer. *Clin Transl Oncol* (2014) 16 (2):178–83. doi: 10.1007/s12094-013-1057-6

50. Du X, Zhang Q, Wang S, Chen X, Wang Y. Mcam is associated with metastasis and poor prognosis in osteosarcoma by modulating tumor cell migration. *J Clin Lab Anal* (2022) 36(2):e24214. doi: 10.1002/jcla.24214

51. Duan H, Jing L, Jiang X, Ma Y, Wang D, Xiang J, et al. Cd146 bound to lck promotes T cell receptor signaling and antitumor immune responses in mice. *J Clin Invest* (2021) 131(21):e148568. doi: 10.1172/JCI148568

52. Claesson-Welsh L, Welsh M. Vegfa and tumour angiogenesis. *J Intern Med* (2013) 273(2):114–27. doi: 10.1111/joim.12019

53. Zhang J, Zhang J, Pang X, Chen Z, Zhang Z, Lei L, et al. Mir-205-5p suppresses angiogenesis in gastric cancer by downregulating the expression of vegfa and Fgf1. *Exp Cell Res* (2021) 404(2):112579. doi: 10.1016/j.yexcr.2021.112579

54. Cao W, Zhao Y, Wang L, Huang X. Circ0001429 regulates progression of bladder cancer through binding mir-205-5p and promoting vegfa expression. *Cancer biomark* (2019) 25(1):101–13. doi: 10.3233/CBM-182380

55. Situ Y, Xu Q, Deng L, Zhu Y, Gao R, Lei L, et al. System analysis of in renal cell carcinoma: The expression, prognosis, gene regulation network and regulation targets. *Int J Biol Markers* (2022) 37(1):90–101. doi: 10.1177/17246008211063501

56. Petroni G, Buqué A, Zitvogel L, Kroemer G, Galluzzi L. Immunomodulation by targeted anticancer agents. *Cancer Cell* (2021) 39(3):310–45. doi: 10.1016/j.ccell.2020.11.009

57. Xie L, Ji T, Guo W. Anti-angiogenesis target therapy for advanced osteosarcoma (Review). *Oncol Rep* (2017) 38(2):625–36. doi: 10.3892/or.2017.5735

58. Stefanovic B, Stefanovic L, Schnabl B, Bataller R, Brenner DA. Tram2 protein interacts with endoplasmic reticulum Ca2+ pump Serca2b and is necessary for collagen type I synthesis. *Mol Cell Biol* (2004) 24(4):1758–68. doi: 10.1128/MCB.24.4.1758-1768.2004

59. Gao X, Jiang W, Ke Z, Huang Q, Chen L, Zhang G, et al. Tram2 promotes the malignant progression of glioma through Pi3k/Akt/Mtor pathway. *Biochem Biophys Res Commun* (2022) 586:34–41. doi: 10.1016/j.bbrc.2021.11.061

60. Li L, Ugalde AP, Scheele CLGJ, Dieter SM, Nagel R, Ma J, et al. A comprehensive enhancer screen identifies Tram2 as a key and novel mediator of yap oncogenesis. *Genome Biol* (2021) 22(1):54. doi: 10.1186/s13059-021-02272-8

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 26 January 2023
DOI 10.3389/fonc.2022.1087762

OPEN ACCESS

EDITED BY
Nan Zhang,
Harbin Medical University, China

REVIEWED BY
Gong YiJia,
Chongqing Medical University, China
Jingwei Zhang,
Xiangya Hospital, Central South University, China

*CORRESPONDENCE
Xiaolei Shu
1466540113@qq.com

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

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Oncology

RECEIVED 02 November 2022
ACCEPTED 28 December 2022
PUBLISHED 26 January 2023

CITATION
Wen J, Zhao W and Shu X (2023) A novel cuproptosis-related LncRNA signature: Prognostic and therapeutic value for low grade glioma.
Front. Oncol. 12:1087762.
doi: 10.3389/fonc.2022.1087762

A novel cuproptosis-related LncRNA signature: Prognostic and therapeutic value for low grade glioma

Jun Wen[1†], Wenting Zhao[2†] and Xiaolei Shu[1*]

[1]Chongqing Cancer Multi-Omics Big Data Application Engineering Research Center, Chongqing University Cancer Hospital, Chongqing, China, [2]The First Clinical College, Hubei University of Chinese Medicine, Wuhan, China

Background: As a common primary intracranial tumor, the diagnosis and therapy of low-grade glioma (LGG) remains a pivotal barrier. Cuproptosis, a new way induces cell death, has attracted worldwide attention. However, the relationship between cuproptosis and LGG remains unknown. Our study is all about finding out if there are any genes related to coproptosis that can be used to predict the outcome of LGG.

Methods: RNA data and clinical information were selected from Cancer Genome Atlas (TCGA) datasets and the Genotype-Tissue Expression (GTEx), 5 lncRNAs (GAS5.AS1, MYLK.AS1, AC142472.1, AC011346.1, AL359643.3) were identified by Cox univariate and multivariate regression, as well as LASSO Cox regression. In the training and test sets, a dual validation of the predictive signature comprised of these 5 lncRNAs was undertaken. The findings demonstrate that the risk model is able to predict the survival regression of LGG patients and has a good performance in either the KM curve approach or the ROC curve. GO, GSEA and KEGG were carried out to explore the possible molecular processes that affecting the prognosis of LGG. The characteristics of immune microenvironment were investigated by using CIBERSORT, ESTIMATE and ssGSEA.

Results: We identified five lncRNAs related with cuproptosis that were closely associated with the prognosis of LGG and used these five lncRNAs to develop a risk model. Using this risk model, LGG patients were then divided into high-risk and low-risk groups. The two patient groups had significantly distinct survival characteristics. Analyses of Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) revealed that the differential genes of the two patient groups were primarily concentrated in neural active ligand-receptor interaction and cytokine-cytokine receptor interaction. The ssGSEA score determined the information related to immune infiltration, and the two groups were differentially expressed in immune subpopulations such as T cells and B cells as well.

Conclusion: Our study discovered 5 cuproptosis-related lncRNAs which contribute to predicting patients' survival of LGG and provide ideas for the exploration of new targets for LGG in the future.

KEYWORDS

cuproptosis, lncRNA, low grade glioma, The Cancer Genome Atlas, prognosis

1 Introduction

According to the classification of the World Health Organization, gliomas can be divided into I-IV grades based on the malignant degree of tumor cells, of which grades II-III belong to LGG and grade IV to glioblastoma (1). Glioblastoma is the most frequent malignant intracerebral tumor, accounting for about 57% of all gliomas and 48% of all primary malignant central nervous system tumors (2). Its prognosis is poor, and the median survival time is less than two years (3, 4). With better prognosis, the life expectancy of patients with LGG is often more than 10 years. However, The natural history of these tumors is marked by frequent recurrences, despite the fact that the clinical course of the majority of tumors is initially benign (5). Some patients will ultimately worsen, posing grave risks to human life and health (4).

Since 2016, the World Health Organization (WHO) has added molecular characteristics, such as 1p19q co-deletion, ATRX, TP53, and IDH mutations, in the diagnostic categorization of LGG, offering a more thorough and accurate diagnosis (6, 7). High frequencies of epidermal growth factor receptor (EGFR) amplification (8), TERT promoter mutation (9), and PTEN loss are characteristic in idh wild-type glioblastomas (10). Because the presence of these distinctions impacts the prognosis of LGG, the current therapeutic strategy is deeply influenced by these molecular markers.

Cuproptosis is a unique type of cell death recently discovered (11, 12). Specifically, copper binds directly to the fatty components of the tricarboxylic acid (TCA) cycle, resulting in the accumulation of lipoproteins and the subsequent loss of Fe-S cluster proteins, resulting in protein toxic stress and eventually cell death (13). Recent studies have showed higher levels of copper in lots of malignant tumors compared with normal tissues, such as breast (14), lung (15), colorectal (16), oral (17) and bladder cancers (18). Change of the copper protein levels may contributes to the growth or invasion of tumor (19). Its specific mechanism includes stabilizing the nuclear hypoxia-inducible factor-1 (HIF-1) (19, 20), which provides help to subsequent angiogenesis, and ultimately leads to tumor progression and metastasis.

Long non-coding RNA (LncRNA) have a significant role in the control of gene expression and are also implicated in the regulation of programmed cell death (PCD), including autophagy, apoptosis, necrotizing apoptosis, and iron death, which impact the growth of cancer cells in cancer patients (21). In recent years, the lncRNA-constructed LGG prognostic model has demonstrated a degree of success. Shengchao Xu and coworkers developed a model consisting of 19 hypoxia-related lncRNAs that accurately predicts the prognosis and treatment response of LGG patients (22). We developed a model of cuproptosis-related lncRNAs with the purpose of better predicting patient prognosis. Figure 1 depicts the workflow for this research.

2 Materials and methods

2.1 Data and resources

The transcriptome profiles and clinical characteristics of LGG patients were retrieved from The Cancer Genome Atlas (TCGA, https://www.tcga.org/) (23), and the transcriptional profiles of normal brain tissues were collected from the Genotype-Tissue Expression Project Database (GTEx, https://commonfund.nih.gov/GTEx). The Counts type data are downloaded from the UCSC xena database (http://xena.ucsc.edu/) (24). Data from patients without complete clinical information were excluded from the study.

2.2 Identification of cuproptosis-related lncRNAs

Firstly, the 'LIMMA' package (25) in R language (Version 4.1.0) is utilized to pre-process. Then, using the limma software, adjust adj.P values <0.05 and $|logFC| > 1$ condition, identify lncRNAs with differential expression. Using "cuproptosis" as the key word, 13 related genes were selected from PubMed (https://pubmed.ncbi.nlm.nih.gov/). Finally, by the Pearson correlation analysis (26) (with a Correlation coefficient >0.7 and adjust P values <0.001), the cuproptosis-related lncRNA is obtained. Protein-Protein Interaction Networks (PPI, https://cn.string-db.org/) (27) was used to investigate the interaction between these genes and lncRNAs.

2.3 Construction of a prognostic cuproptosis-related lncRNA signature

523 LGG patients were randomly selected and divided into training set and test set, in which the training set accounted for 70% and the test set accounted for 30%. According to their median lncRNA expression, patients in the training set were separated into two groups: high and low expressing individuals. When comparing the median survival times of the two groups of patients, we drew KM curves to see whether high or low lncRNA expression had an impact on outcome (28). Univariate COX regression and LASSO regression were used to the KM-curve-selected lncRNAs. The R packages 'survminer' and 'glmnet' (29) performed the aforementioned tasks. We indicated that univariate and lasso Cox regression analyses were useful in identifying candidate lncRNAs with prognostic significance and reducing the impact of overfitting. Risk signatures were built after a preliminary round of multivariate Cox regression analysis. Risk score=$\sum_1^n coef_i * x_i$ (Coefi indicates the correlation coefficient of each ferroptosisrelated signature, and X indicates the level of gene

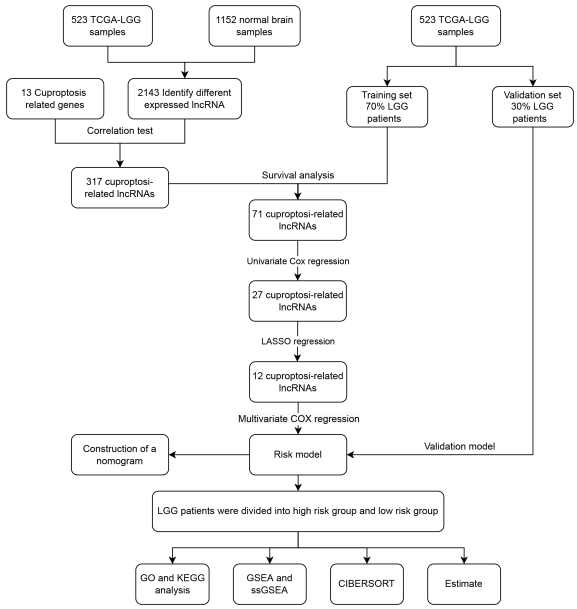


FIGURE 1
The flow chart of data analysis.

expression) was the formula used to determine the level of danger. The median risk score was used to classify the training and testing sets into high-risk and low-risk groups.

2.5 Independent prognostic value of the signature

Then, we analyzed the signature's predictive power by running univariate and multivariate Cox regressions. The patients' chances of survival were also estimated using the risk score's predictive nomogram. The R package "survival" was used to calculate risk scores and determine OS. The model's accuracy was then assessed with the use of the ROC (constructed using the 'survminer' package) and Kaplan-Meier curves (generated by the 'survivalROC' package). Dual validation was performed on the training set and the test set to further assess the model's prediction ability.

2.6 Enrichment analysis

Differentially expressed genes (DEGs) were identified between low risk and high risk groups using the limma program in R (with criteria of FDR< 0.05 and | log2 fold change (FC) | ≥1 or greater). Among the many analytical tools available for functional annotation, gene set enrichment analysis (GSEA) stands out as particularly potent. It may be used to decode the expression profile of the whole genome and investigate the connections between various cancer-related, metabolic, transcriptional, and stress-related pathways and activities. Get the HALLMARK genes set from the MSigDB database (https://www.gsea-msigdb.org/GSEA/msigdb) (30), and then run a GSEA analysis using the 'GSVA' program (P<0.05 and FDR<0.25) to compare the high-risk and low-risk groups.

To compare the DEGs of high-risk and low-risk groups, we used the R tool 'ClusterProfiler' in conjunction with the KEGG and GO databases (31). And infer its purpose from studies of gene sets. Several biological activities and pathways are overrepresented in differentially expressed genes between these two groups; we explore here whether these could contribute to disparities in survival.

2.7 Landscape of immune cells infiltration

The "gsva" R package was used for single-sample gene set enrichment analysis (ssGSEA) to assess the immune infiltration status of LGG patients in various risk categories. Using the CIBERSORT software (http://cibersort.stanford.edu/), estimate the cell subgroup abundance by analyzing whole gene expression profiles (32). Scores are produced using the ESTIMATE algorithm (https://bioinformatics.mdanderson.org/public-software/estimate/) to forecast the amount of infiltrating immune and stromal cells, which serve as the foundation for inferring tumor immunity.

2.8 RNA extraction and rt-PCR

The U251 glioma cell line and human astrocyte cell line NHA were purchased from Beyotime (Shanghai, China) and cultured in Dulbecco's Modified Eagle Medium (DMEM; Gibco, NY, USA) containing 10% fetal bovine serum (FBS; Gibco, NY, USA), penicillin (100 units/ml), and streptomycin (100 μg/ml) in a humidified incubator maintained at 5% CO2 and 37° C. Extracted total RNA from cell lines by using Universal RNA Extraction Kit (Takara; Dalian, China). PrimeScript RT-PCR Kit and TB Green were used for reverse transcription and relative lncRNA expression assessment, respectively. Primer information is shown in Table 1.

TABLE 1 Primer sequences.

lncRNA	Primer
GAS5-AS1	Forward: 5'-TGTGCCCTTTATACCCACTTT-3' Reverse: 5'-GCCCAACTAGTGATAGGCATTA-3'
MYLK-AS1	Forward: 5'-TTGCAGTGTTCAGCACTGGCAC-3' Reverse: 5'-ATTCGACGACCAGTGTTTCAGT-3'
GAPDH	Forward: 5'-GGTGTGAACCATGAGAAGTATGA-3' Reverse: 5'-GAGTCCTTCCACGATACCAAAG-3'

2.9 Statistical analysis

R software version 4.0.4 was used for data analysis. Unpaired Student's t test and Wilcoxon test were used to compare data conforming to normal distribution and non-normal distribution, respectively. $p<0.05$ was considered as the threshold for statistical significance.

3 Results

3.1 Construction of a cuproptosis-related lncRNAs prognostic model signature

Brain tissues from LGG patients and controls showed differential expression for 2143 lncRNAs in TCGA (Figures 2A, B). From a search of PubMed, we know that there are 13 genes involved in cuproptosis-related genes: DLST, FDX1, LIAS, SLC31A1, LIPT1, ATP7A, DLD, ATP7B, PDHB, and DBT (33, 34). The chosen genes were used to create a correlation network map with differential expression lncRNAs (Figure S1), from which 317 lncRNAs with cuproptosis-related differential expression were extracted. Further confirming the usefulness of these lncRNAs, KM curves were generated for 70% of patients chosen from the TCGA database, and lncRNAs with minor survival significance were omitted. Finally, 71 lncRNAs were successfully extracted.

The lncRNAs identified in the preceding phase were subjected to univariate and lasso regression analysis. In the univariate regression analysis, 27 lncRNAs were discovered to be substantially related to OS (Figure 3A). In addition, LASSO regression analysis of these 27 lncRNAs removed 15 lncRNAs and yielded 12 lncRNAs associated with cuproptosis (Figures 3B, C). These 12 lncRNAs underwent multifactorial regression analysis, and a risk model for 5-cuproptosis-related lncRNAs signature was developed (Figures 3D, E).

3.2 Validation of the prognostic model

Based on the risk scores, we plotted KM curves and time-dependent ROC curves (Figures 4A–D) for the high-risk and low-risk groups of patients in the training set and the test set (Figures 4A–D). As seen in the graph, our model had a high predictive value at 1, 3, and 5 years for both the training and validation sets (AUC were greater than 0.75).

Subsequently, a predictive Nomogram was created, by this 5-cuproptosis-related lncRNA signature (Figures 4E). This line graph includes clinical characteristics such as age, gender, and grade. The calibration curve showed that the Nomogram could accurately predict the overall survival at 1, 3, and 5 years (Figure S2).

3.3 Functional enrichment analysis

GO enrichment and KEGG pathway were carried out to analysis the possible molecular processes. Results showed that the differential genes were mainly involved in signal pathways such as neuroreceptor-ligand interaction, cytokine-cytokine interaction, and tumor proteoglycan (Figures 5A, B). Subsequently GSEA analysis also showed that the differential gene pathway was mainly concentrated in MTORC1 signal pathway and apoptosis, KRAS signal pathway (Figure 5C; Table S1).

3.4 Immune-related analysis of LGG patients

We employed the CIBERSORT and Estimate method to identify immune cell infiltration in LGG patients, since the enrichment analysis revealed that the association between cuproptosis and LGG

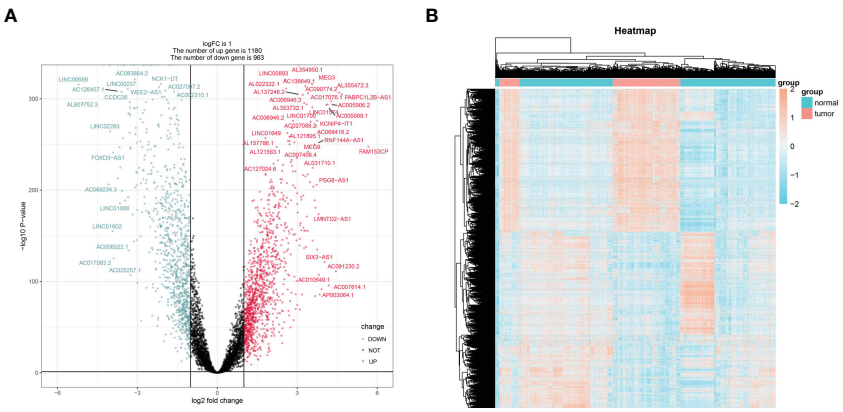


FIGURE 2
The screening of differentially expressed lncRNAs lncRNAs. The volcano graph **(A)** and heatmap **(B)** showed that 1180 lncRNAs were down-regulated and that 963 lncRNAs were up-regulated in tissues of LGG compared to normal tissues.

10.3389/fonc.2022.1087762

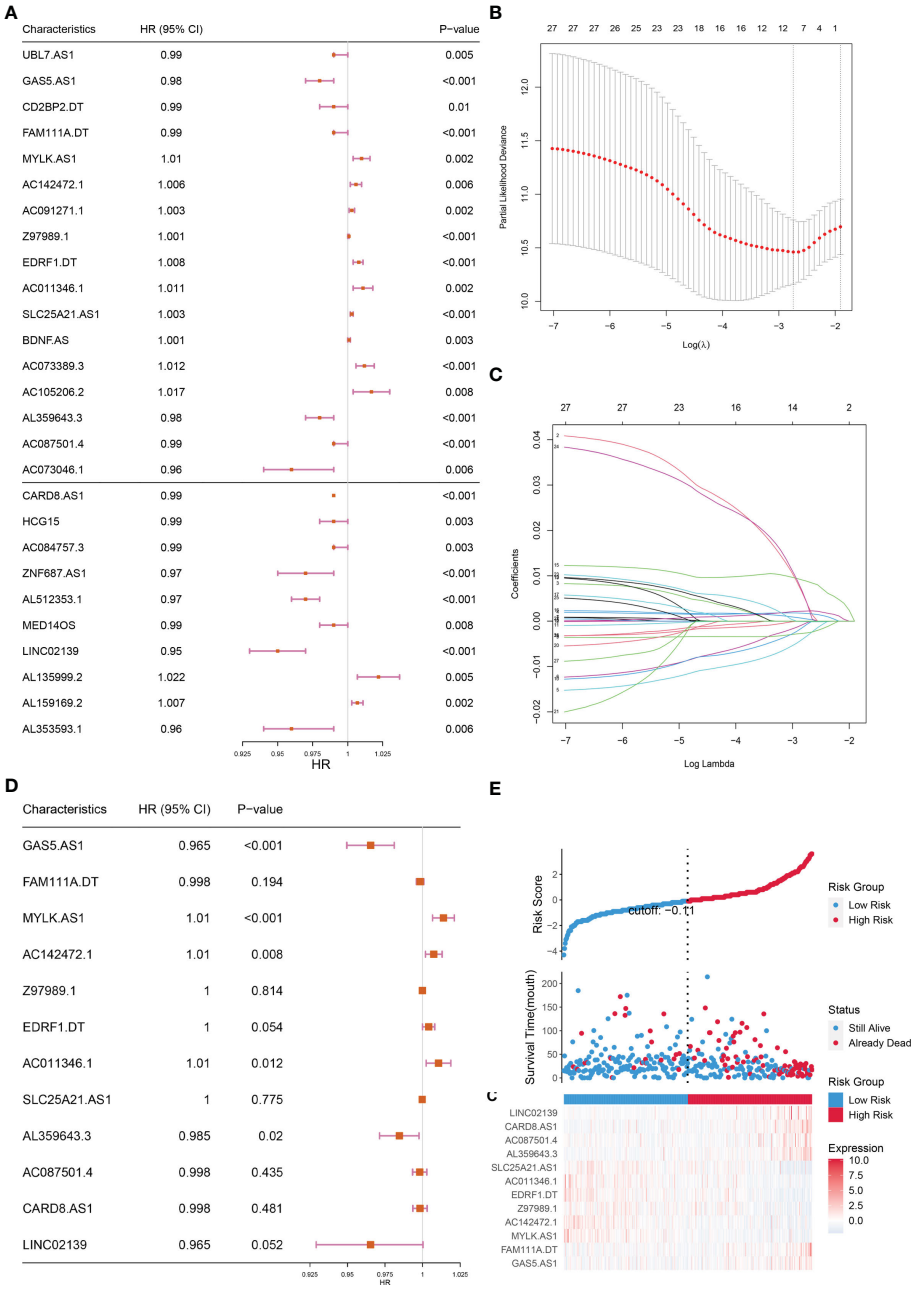


FIGURE 3
Construction of the prognostic cuproptosis-related lncRNAs signature for in the training set. **(A)** Based on univariate Cox regression analysis, 21 of the 71 cuproptosis-related lncRNAs were screened, as shown by the forest map. **(B, C)** Lasso regression analysis was used to further screen out 12 lncRNAs based on 10-fold cross-validation. **(D)** Forest plot of 12 cuproptosis-related lncRNAs based on Multivariate Cox regression. **(E)** The riskscore distribution, OS, and the Heat map of five lncRNAs of patients in the training set.

is mostly reliant on the tumor inflammatory pathway. We used CIBERSOR and Estimate algorithms to calculate the relative proportion of 22 immune cells in each LGG patient. The correlation analysis between risk score and the level of immune cell infiltration showed that the infiltration degree of many immune cells was different among subgroups ($P< 0.05$). The results showed that the scores of monocytes and M1 macrophages and mast cells decreased in the high-risk group (Figure 6).

3.5 rt-PCR was used to verify the expression of lncRNAs in glioma cell line

Among the 5 lncRNAs, AC142472.1, AC011346.1, and AL359643.3 lacked relevant studies, therefore MYLK.AS1 and GAS5.AS1, which have been shown to be strongly associated with tumors in prior research, were chosen. In the U251 cell line, the expressions of GAS5.AS1 and MYLK.AS1 were up-regulated and down-regulated, respectively (Figure 7).

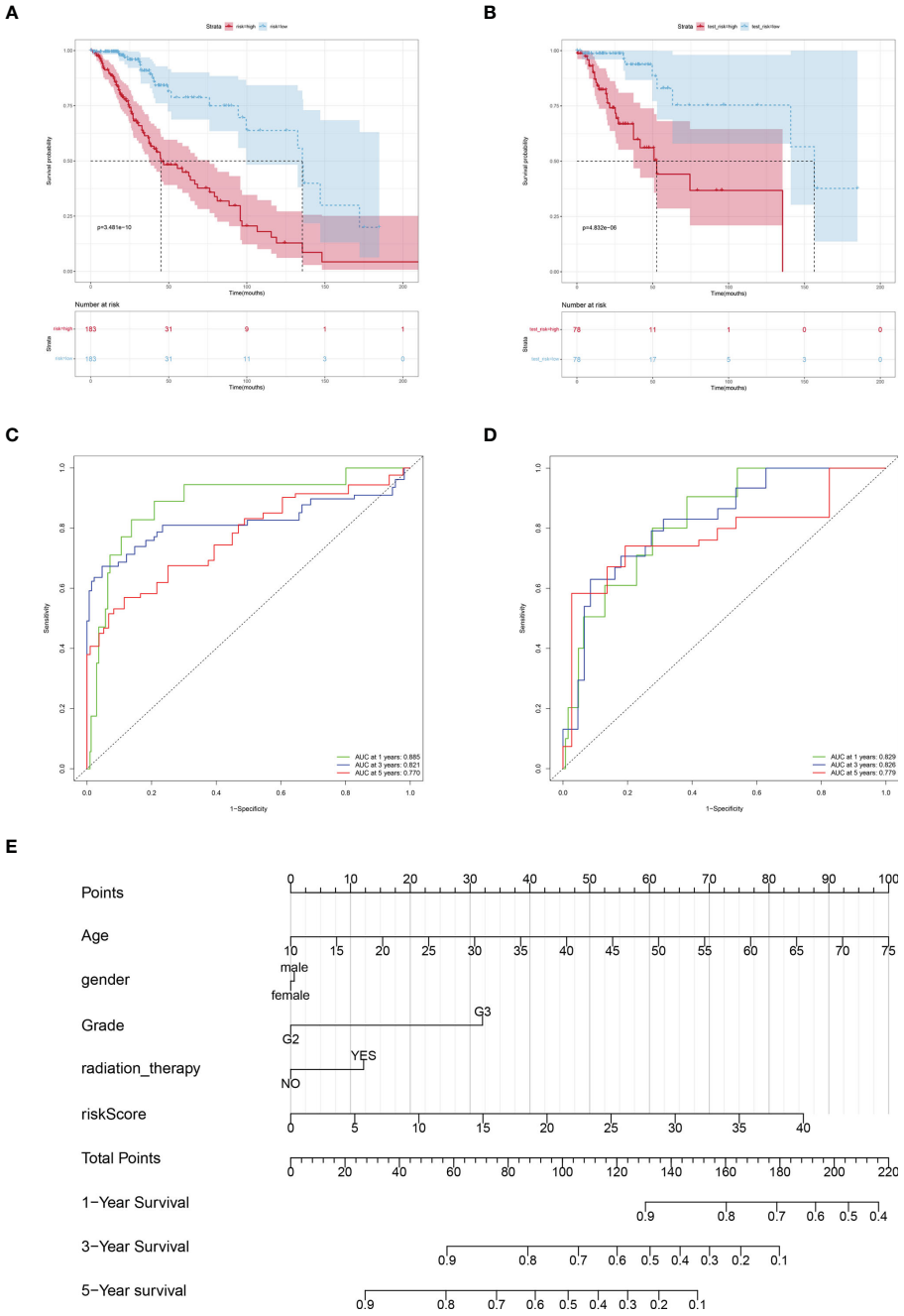


FIGURE 4
Verification the prognostic value of risk score. Kaplan-Meier curves of LGGs patients in the TCGA training cohort **(A)** and testing cohort **(B)**. AUC values at 1, 3, and 5 years in the TCGA training cohort **(C)** and testing cohort **(D)**. **(E)** Nomogram integrating risk score and clinical variables predicts 1-, 3-, and 5-year OS probabilities.

4 Discussion

Low-grade gliomas are primary brain tumors that tend to occur in young people. Common treatments include surgery and chemotherapy, accompany with good prognosis and long survival (35). But with our timely treatment, it will seriously affect the quality of life. Therefore, new approaches to LGG diagnosis and treatment are urgently needed.

Has a fundamental effect on biological processes (36), copper can regulates several biological pathways based on external stimulation (37). The copper accumulation is closely related to tumor proliferation and growth, angiogenesis, and metastasis (19, 37).

In this research, by analyzing the clinical data of LGG patients in TCGA and combining it with coproptosis, we constructed 5 (GAS5.AS1, MYLK.AS1, AC142472.1, AC011346.1, AL359643.3) cuproptosis-related lncRNAs prognostic models, analyzed and predicted their clinical prognosis, and found the relationship between them and tumor immunity.The discriminability and precision of the developed lncRNA signatures were validated using Kaplan-Meier survival analysis and area under the curve (AUC). The

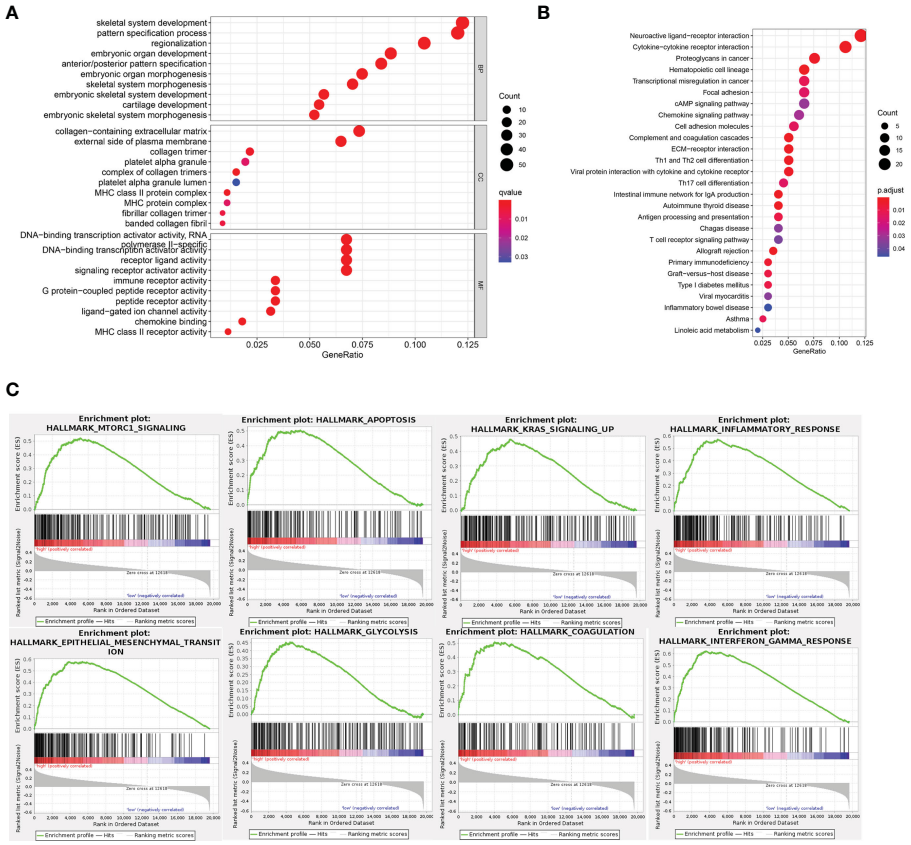


FIGURE 5
Functional analysis of DEGs. **(A, B)** GO and KEGG pathway enrichment analyses of DEGs in low-risk and high-risk groups. **(C)** In GSEA, the top 8 pathways or biological processes were sorted by P value.

test set was then used to validate the risk model's predictive value. The prognostic model performed well in ROC curve analysis, with auc values between 0.88 and 0.77. In addition, the risk score was determined to be an independent risk factor. Thus, the model demonstrated high clinical predictive value. In addition, the signature consists of just five lncRNAs, making it more applicable to clinical applications than previous signatures.

GAS5-AS1 is a down-regulated gene found in glioma tissues and cells. Its high expression can inhibit the proliferation, migration, and invasion of glioma cells. The expression of GAS5-AS1 is related to the tumor grade of glioma and can be used as a new target for the treatment and prognosis prediction of glioma (38). In glioma tissues and cells, lncRNA GAS5-AS1 was repressed, whereas miR-106b-5p was increased. Through the sponge effect, lncRNA GAS5-AS1 may bind miR-106b-5p, therefore promoting the expression of its target gene TUSC2 and inhibiting the growth and spread of glioma (38). In addition, MYLK-AS1 has been found to promote the growth and invasion of hepatocellular carcinoma cells through EGFR/HER2-ERK1/2 signal pathway (39), At the same time, it can also target miR-424-5p/E2F7 axis, activate VEGFR-2 signal pathway, and promote tumor progression and angiogenesis (39), And promote the invasion of nephroblastoma (40). Combined with our research, it may help us to better understand the molecular mechanism of glioma progression. Our research expands the field and provides a reference and direction for their application in cuproptosis and LGG.

Furthermore, based on the 5 lncRNA risk models developed, we estimated the risk scores of LGG patients in the TCGA database and categorized them into high-risk and low-risk groups. Then, Enrichment Analysis was performed. GSEA analysis revealed that the differences between the high-risk and low-risk groups were primarily enriched in the mTORC1 signal, the KARS signal, and the apoptosis. The mTOR pathway is an important regulator of cell survival or proliferation and plays a central role in regulating many basic cellular processes from protein synthesis to autophagy (41). It has been reported that the expression of mTOR pathway is up-regulated in GBM (42). At the same time, mTOR can promote the differentiation and expansion of CD4+ FoxP3+ regulatory T cells and CD8+ memory T cells, and inhibit CD8+ and CD4+ effector T cells (43, 44). This is consistent with our findings in GSEA, but its specific mechanism remains to be further studied, which provides a reference research direction for cuproptosis-related genes to predict the prognosis of LGG gliomas. Complex signaling cascades stimulate RAS, which then activates downstream signaling pathways to regulate a wide variety of cellular functions (45). The KRAS gene, which is part of the RAS gene family, is tied with glioma development and progression (46, 47). KRAS influences the inflammatory milieu of cancer by activating the MAPK and PI3K signaling pathways, which results in the release of additional IL-6/IL-8 cytokines and cancer cell proliferation (48, 49). As for apoptosis, it is inseparable with tumor and almost participates in the whole process of tumor.

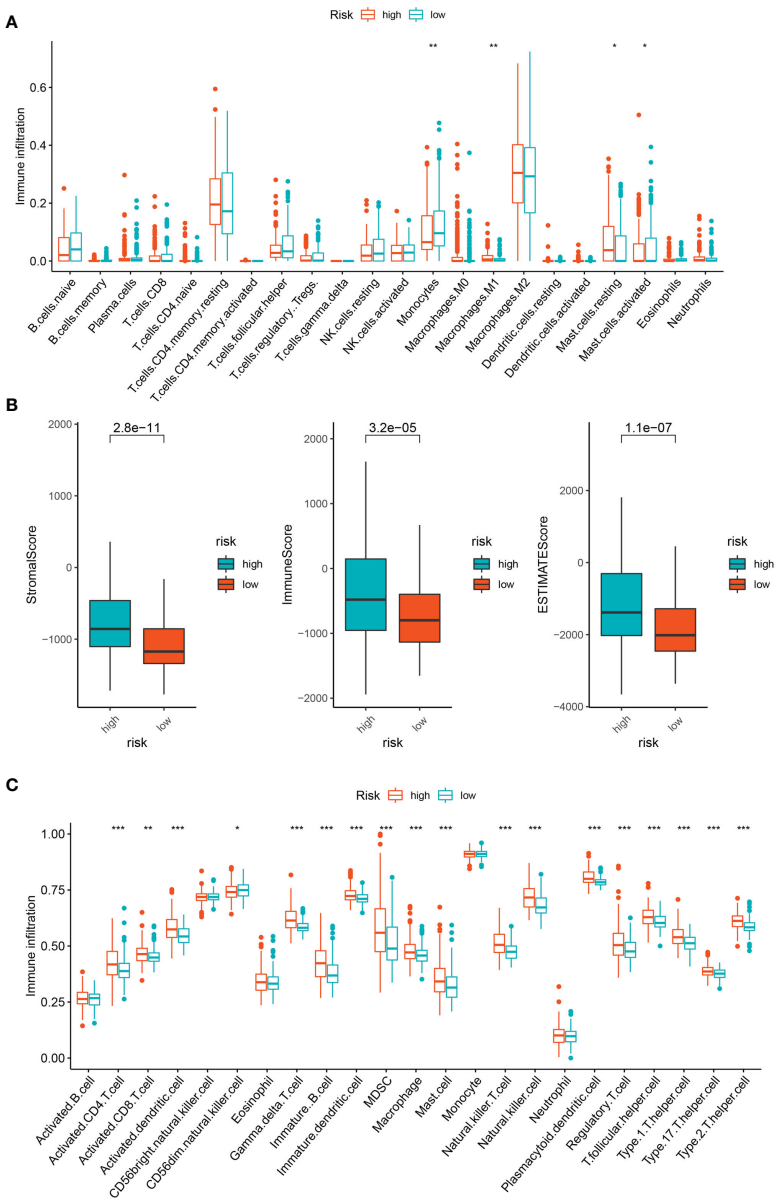


FIGURE 6
Immune infiltration analysis of DEGs. **(A)** Immune cell subpopulations in ssGSEA. **(B)** Different socres in high- and low-risk group. **(C)** Immune cell subpopulations in CIBERSORT. *p < 0.05, **p < 0.01, and ***p < 0.001.

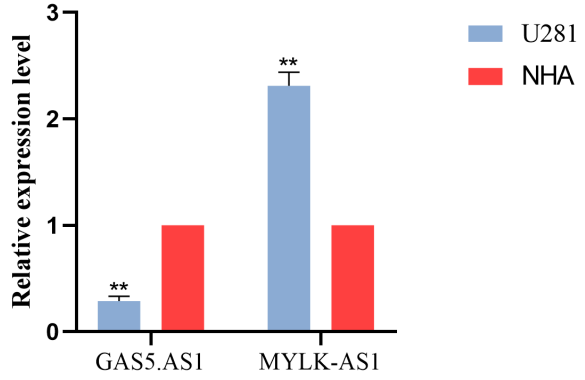


FIGURE 7
Validation of the expression level of GAS5.AS1 and MYLK.AS1 in cell lines and tissues.

Together, these studies support our findings, while there are still many important questions remain unanswered. The specific mechanism of coproptosis-related lncRNAs in LGG, and how they affect tumor development by affecting immunity need more details.

Our data provides a direction and a certain possibility for the treatment of LGG. But there are still certain limitations. Our sample was based entirely on public databases with limited clinical evidence. The prognostic model established in this study needs further experimental verification.

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

JW conceived the research and prepared the publication, whilst WZ conducted the bioinformatics-related analysis. WZ carried out the experimental validation phase of the research. The study was evaluated and arranged by XS. All authors contributed to editorial changes in the manuscript. All authors read and approved the final 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

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/fonc.2022.1087762/full#supplementary-material

SUPPLEMENTARY FIGURE 1
PPI network between cuproptosis-related genes and LGG lncRNAs.

SUPPLEMENTARY FIGURE 2
Kaplan-Meier curves of OS in 12 DEGs.

SUPPLEMENTARY TABLE 1
GSEA analysis of DEGs based on HALLMARK genes set.

References

1. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO classification of tumors of the central nervous system: a summary. *Neuro Oncol* (2021) 23:1231–51. doi: 10.1093/neuonc/noab106

2. Ostrom QT, Gittleman H, Liao P, Vecchione-Koval T, Wolinsky Y, Kruchko C, et al. CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the united states in 2010-2014. *Neuro Oncol* (2017) 19:v1–v88. doi: 10.1093/neuonc/nox158

3. Ohgaki H, Kleihues P. Genetic alterations and signaling pathways in the evolution of gliomas. *Cancer Sci* (2009) 100:2235–41. doi: 10.1111/j.1349-7006.2009.01308.x

4. Tan AC, Ashley DM, López GY, Malinzak M, Friedman HS, Khasraw M. Management of glioblastoma: State of the art and future directions. *CA Cancer J Clin* (2020) 70:299–312. doi: 10.3322/caac.21613

5. Yu Y, Villanueva-Meyer J, Grimmer MR, Hilz S, Solomon DA, Choi S, et al. Temozolomide-induced hypermutation is associated with distant recurrence and reduced survival after high-grade transformation of low-grade IDH-mutant gliomas. *Neuro Oncol* (2021) 23:1872–84. doi: 10.1093/neuonc/noab081

6. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 world health organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol* (2016) 131:803–20. doi: 10.1007/s00401-016-1545-1

7. Ohgaki H, Kleihues P. The definition of primary and secondary glioblastoma. *Clin Cancer Res* (2013) 19:764–72. doi: 10.1158/1078-0432.CCR-12-3002

8. Verhaak RGW, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* (2010) 17:98–110. doi: 10.1016/j.ccr.2009.12.020

9. Simon M, Hosen I, Gousias K, Rachakonda S, Heidenreich B, Gessi M, et al. TERT promoter mutations: a novel independent prognostic factor in primary glioblastomas. *Neuro Oncol* (2015) 17:45–52. doi: 10.1093/neuonc/nou158

10. Li J, Yen C, Liaw D, Podsypanina K, Bose S, Wang SI, et al. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science* (1997) 275:1943–7. doi: 10.1126/science.275.5308.1943

11. Kahlson MA, Dixon SJ. Copper-induced cell death. *Science* (2022) 375:1231–2. doi: 10.1126/science.abo3959

12. Tang D, Chen X, Kroemer G. Cuproptosis: a copper-triggered modality of mitochondrial cell death. *Cell Res* (2022) 32:417–8. doi: 10.1038/s41422-022-00653-7

13. 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

14. Jouybari L, Kiani F, Islami F, Sanagoo A, Sayehmiri F, Hosnedlova B, et al. Copper concentrations in breast cancer: A systematic review and meta-analysis. *Curr Med Chem* (2020) 27:6373–83. doi: 10.2174/0929867326666190918120209

15. Zhang X, Yang Q. Association between serum copper levels and lung cancer risk: A meta-analysis. *J Int Med Res* (2018) 46:4863–73. doi: 10.1177/0300060518798507

16. Stepien M, Jenab M, Freisling H, Becker N-P, Czuban M, Tjønneland A, et al. Pre-diagnostic copper and zinc biomarkers and colorectal cancer risk in the European prospective investigation into cancer and nutrition cohort. *Carcinogenesis* (2017) 38:699–707. doi: 10.1093/carcin/bgx051

17. Chen F, Wang J, Chen J, Yan L, Hu Z, Wu J, et al. Serum copper and zinc levels and the risk of oral cancer: A new insight based on large-scale case-control study. *Oral Dis* (2019) 25:80–6. doi: 10.1111/odi.12957

18. Basu S, Singh MK, Singh TB, Bhartiya SK, Singh SP, Shukla VK. Heavy and trace metals in carcinoma of the gallbladder. *World J Surg* (2013) 37:2641–6. doi: 10.1007/s00268-013-2164-9

19. Lelièvre P, Sancey L, Coll J-L, Deniaud A, Busser B. The multifaceted roles of copper in cancer: A trace metal element with dysregulated metabolism, but also a target or a bullet for therapy. *Cancers (Basel)* (2020) 12:E3594. doi: 10.3390/cancers12123594

20. Agostinelli E, Vianello F, Magliulo G, Thomas T, Thomas TJ. Nanoparticle strategies for cancer therapeutics: Nucleic acids, polyamines, bovine serum amine oxidase and iron oxide nanoparticles (Review). *Int J Oncol* (2015) 46:5–16. doi: 10.3892/ijo.2014.2706

21. Jiang N, Zhang X, Gu X, Li X, Shang L. Progress in understanding the role of lncRNA in programmed cell death. *Cell Death Discovery* (2021) 7:30. doi: 10.1038/s41420-021-00407-1

22. Xu S, Tang L, Liu Z, Luo C, Cheng Q. Hypoxia-related lncRNA correlates with prognosis and immune microenvironment in lower-grade glioma. *Front Immunol* (2021) 12:731048. doi: 10.3389/fimmu.2021.731048

23. Ganini C, Amelio I, Bertolo R, Bove P, Buonomo OC, Candi E, et al. Global mapping of cancers: The cancer genome atlas and beyond. *Mol Oncol* (2021) 15:2823–40. doi: 10.1002/1878-0261.13056

24. Wang S, Xiong Y, Zhao L, Gu K, Li Y, Zhao F, et al. UCSCXenaShiny: An R/CRAN package for interactive analysis of UCSC xena data. *Bioinformatics* (2021) 38(2):527–9. doi: 10.1093/bioinformatics/btab561

25. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43:e47. doi: 10.1093/nar/gkv007

26. Bishara AJ, Hittner JB. Testing the significance of a correlation with nonnormal data: comparison of Pearson, spearman, transformation, and resampling approaches. *Psychol Methods* (2012) 17:399–417. doi: 10.1037/a0028087

27. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* (2019) 47:D607–13. doi: 10.1093/nar/gky1131

28. Schober P, Vetter TR. Kaplan-Meier Curves, log-rank tests, and cox regression for time-to-Event data. *Anesth Analg* (2021) 132:969–70. doi: 10.1213/ANE.0000000000005358

29. Engebretsen S, Bohlin J. Statistical predictions with glmnet. *Clin Epigenet* (2019) 11:123. doi: 10.1186/s13148-019-0730-1

30. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst* (2015) 1:417–25. doi: 10.1016/j.cels.2015.12.004

31. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an r package for comparing biological themes among gene clusters. *OMICS* (2012) 16:284–7. doi: 10.1089/omi.2011.0118

32. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med* (2015) 21:938–45. doi: 10.1038/nm.3909

33. Zhou C, Li C, Zheng Y, Huang X. Regulation, genomics, and clinical characteristics of cuproptosis regulators in pan-cancer. *Front Oncol* (2022) 12:934076. doi: 10.3389/fonc.2022.934076

34. Liu H. Pan-cancer profiles of the cuproptosis gene set. *Am J Cancer Res* (2022) 12:4074–81. doi: 10.21203/rs.3.rs-1716214/v1

35. Wang TJC, Mehta MP. Low-grade glioma radiotherapy treatment and trials. *Neurosurg Clin N Am* (2019) 30:111–8. doi: 10.1016/j.nec.2018.08.008

36. Lm R AL, Aa E. Role of copper on mitochondrial function and metabolism. *Front Mol Biosci* (2021) 8:711227. doi: 10.3389/fmolb.2021.711227

37. Li Y. Copper homeostasis: Emerging target for cancer treatment. *IUBMB Life* (2020) 72:1900–8. doi: 10.1002/iub.2341

38. Huang W, Shi Y, Han B, Wang Q, Zhang B, Qi C, et al. LncRNA GAS5-AS1 inhibits glioma proliferation, migration, and invasion *via* miR-106b-5p/TUSC2 axis. *Hum Cell* (2020) 33:416–26. doi: 10.1007/s13577-020-00331-z

39. Teng F, Zhang J-X, Chang Q-M, Wu X-B, Tang W-G, Wang J-F, et al. LncRNA MYLK-AS1 facilitates tumor progression and angiogenesis by targeting miR-424-5p/E2F7 axis and activating VEGFR-2 signaling pathway in hepatocellular carcinoma. *J Exp Clin Cancer Res* (2020) 39:235. doi: 10.1186/s13046-020-01739-z

40. Zhu S, Zhang J, Gao X, Tang X, Cui Y, Li D, et al. Silencing of long noncoding RNA MYLK-AS1 suppresses nephroblastoma *via* down-regulation of CCNE1 through transcription factor TCF7L2. *J Cell Physiol* (2021) 236:5757–70. doi: 10.1002/jcp.30259

41. Saxton RA, Sabatini DM. mTOR signaling in growth, metabolism, and disease. *Cell* (2017) 169:361–71. doi: 10.1016/j.cell.2017.03.035

42. Dumas AA, Pomella N, Rosser G, Guglielmi L, Vinel C, Millner TO, et al. Microglia promote glioblastoma *via* mTOR-mediated immunosuppression of the tumour microenvironment. *EMBO J* (2020) 39:e103790. doi: 10.15252/embj.2019103790

43. Araki K, Turner AP, Shaffer VO, Gangappa S, Keller SA, Bachmann MF, et al. mTOR regulates memory CD8 T-cell differentiation. *Nature* (2009) 460:108–12. doi: 10.1038/nature08155

44. Haxhinasto S, Mathis D, Benoist C. The AKT-mTOR axis regulates *de novo* differentiation of CD4+Foxp3+ cells. *J Exp Med* (2008) 205:565–74. doi: 10.1084/jem.20071477

45. Mukhopadhyay S, Vander Heiden MG, McCormick F. The metabolic landscape of RAS-driven cancers from biology to therapy. *Nat Cancer* (2021) 2:271–83. doi: 10.1038/s43018-021-00184-x

46. Guan Q, Yuan L, Lin A, Lin H, Huang X, Ruan J, et al. KRAS gene polymorphisms are associated with the risk of glioma: a two-center case-control study. *Transl Pediatr* (2021) 10:579–86. doi: 10.21037/tp-20-359

47. Chiang J, Li X, Liu APY, Qaddoumi I, Acharya S, Ellison DW. Tectal glioma harbors high rates of KRAS G12R and concomitant KRAS and BRAF alterations. *Acta Neuropathol* (2020) 139:601–2. doi: 10.1007/s00401-019-02112-x

48. Ryall S, Zapotocky M, Fukuoka K, Nobre L, Guerreiro Stucklin A, Bennett J, et al. Integrated molecular and clinical analysis of 1,000 pediatric low-grade gliomas. *Cancer Cell* (2020) 37:569–83.e5. doi: 10.1016/j.ccell.2020.03.011

49. Munoz L, Yeung YT, Grewal T. Oncogenic ras modulates p38 MAPK-mediated inflammatory cytokine production in glioblastoma cells. *Cancer Biol Ther* (2016) 17:355–63. doi: 10.1080/15384047.2016.1139249

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 31 January 2023
DOI 10.3389/fonc.2023.1078697

OPEN ACCESS

EDITED BY
Nan Zhang,
Harbin Medical University, China

REVIEWED BY
Lingge Yang,
Zhejiang University School of Medicine, China
Qihang Yuan,
Dalian Medical University, China

*CORRESPONDENCE
Wei Wang
769816780@qq.com

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Oncology

RECEIVED 24 October 2022
ACCEPTED 05 January 2023
PUBLISHED 31 January 2023

CITATION
Peng H, Li X, Luan Y, Wang C and Wang W (2023) A novel prognostic model related to oxidative stress for treatment prediction in lung adenocarcinoma.
Front. Oncol. 13:1078697.
doi: 10.3389/fonc.2023.1078697

A novel prognostic model related to oxidative stress for treatment prediction in lung adenocarcinoma

Haijun Peng, Xiaoqing Li, Yanchao Luan, Changjing Wang and Wei Wang*

Department of Thoracic Surgery, Hebei Chest Hospital, Hebei Provincial Key Laboratory of Lung Disease, Shijiazhuang, Hebei, China

Background: The prognostic model based on oxidative stress for lung adenocarcinoma (LUAD) remains unclear.

Methods: The information of LUAD patients were acquired from TCGA dataset. We also collected two external datasets from GEO for verification. Oxidative stress-related genes (ORGs) were extracted from Genecards. We performed machine learning algorithms, including Univariate Cox regression, Random Survival Forest, and Least Absolute Shrinkage and Selection Operator (Lasso) analyses on the ORGs to build the OS-score and OS-signature. We drew the Kaplan-Meier and time-dependent receiver operating characteristic curve (ROC) to evaluate the efficacy of the OS-signature in predicting the prognosis of LUAD. We used GISTIC 2.0 and maftool algorithms to explore Genomic mutation of OS-signature. To analyze characteristic of tumor infiltrating immune cells, ESTIMATE, TIMER2.0, MCPcounter and ssGSEA algorithms were applied, thus evaluating the immunotherapeutic strategies. Chemotherapeutics sensitivity analysis was based on pRRophetic package. Finally, PCR assays was also used to detect the expression values of related genes in the OS-signature in cell lines.

Results: Ten ORGs with prognostic value and the OS-signature containing three prognostic ORGs were identified. The significantly better prognosis of LUAD patients was observed in LUAD patients. The efficiency and accuracy of OS-signature in predicting prognosis for LUAD patients was confirmed by survival ROC curves and two external validation data sets. It was clearly observed that patients with high OS-scores had lower immunomodulators levels (with a few exceptions), stromal score, immune score, ESTIMATE score and infiltrating immune cell populations. On the contrary, patients with higher OS-scores were

more likely to have higher tumor purity. PCR assays showed that, MRPL44 and CYCS were significantly higher expressed in LUAD cell lines, while CAT was significantly lower expressed.

Conclusion: The novel oxidative stress-related model we identified could be used for prognosis and treatment prediction in lung adenocarcinoma.

KEYWORDS

lung adenocarcinoma, oxidative stress, prognostic model, machine learning, tumor microenvironment

Introduction

According to the global cancer statistics analysis in 2020, the incidence of lung cancer ranks second only to breast cancer in the world, accounting for about 18% of all cancer deaths, and being the leading cause of cancer death in the world (1). The causes of lung cancer are very complex, including history of exposure to smoking and secondhand smoke, air pollution, history of pulmonary diseases, family history of cancer, occupational exposure to silica and asbestos, poor diet, mental and psychological factors (2–4). The early symptoms of lung cancer are not obvious. Generally, there are corresponding clinical symptoms in the middle and late stage, such as chest pain, hemoptysis, etc. According to relevant studies, 75% of lung cancer patients have been diagnosed at stage III or IV, at which time they have lost the opportunity for surgery, and the treatment means are relatively limited. Conventional radiotherapy and chemotherapy have no obvious effect, and the survival and prognosis are very poor (5). The overall 5-year survival rate for patients with lung cancer is 19%, which drops to 5% if distant metastasis is present at the time of diagnosis, and approximately 57% for patients in the localized stage (6). The diagnosis and treatment of lung cancer are still the focus of current research.

According to pathological types, lung cancer can be divided into non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC accounts for about 85% of all cases diagnosed with lung cancer, which mainly includes lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LSCC), and large cell lung cancer (LCLC) (7). LUAD is the most common pathological type of lung cancer, accounting for approximately 50-70% of surgically resected lung cancers (8) and almost 50% of all lung cancers (9). Precision medicine for disease requires accurate prognostic prediction, such as the risk of future recurrence after the initial treatment and responsiveness to different treatments (10, 11). At present, TNM staging is still the main basis for the treatment of LUAD and has been used clinically for many years as a prognostic predictor of LUAD (12). However, the reproducibility and discrimination ability of TNM staging for prognosis prediction are still not satisfied, and the prognosis is also different among LUAD patients with the same pathological type and stage. At the same time, although the emerging diagnosis and treatment technologies such as gene testing, targeted therapy and immunotherapy have been applied in the clinical diagnosis and treatment of lung cancer, the overall survival rate of lung cancer has only slightly improved compared with other malignant tumors (6, 13). Therefore, there are individual differences in LUAD, and prognosis prediction needs individual predictors.

Tumors often have oxidative stress (OXS), which is an imbalance between oxidation and anti-oxidation in the body that causes aberrant oxidative signal regulation and macromolecular oxidative damage (14). Cellular OXS is caused by ROS accumulation (15). OXS is the principal cause of cell damage, targeting intracellular macromolecules and promoting and suppressing tumor growth (14, 16–18). Tumor cell redox homeostasis control may improve tumor therapy. OXS regulates tumor cell fate in various ways that depend on tumor type and etiology. Future study will focus on controlling OXS's anti-tumor and tumor-promoting effects. We can evaluate OXS heterogeneity in cancers and find new therapeutic targets using bioinformatics and other big data analysis methods.

With the emergence of public biomedical databases such as TCGA (The Cancer Genome Atlas) database, the use of bioinformatics to mine disease gene data has become an important direction of disease research (19). TCGA aims to focus on acquired changes of cancer genes. Up to now, a total of 33 types of cancers have been included in TCGA database (19). Clinical sample information and sequencing data (including transcriptome data, epigenetic data, genomic mutation data, etc.) of more than 20,000 patients can be accessed openly, which has become an important database for cancer research (19, 20). The gene expression data and clinical information of LUAD patients needed in this study were obtained from public databases. In this study, we obtained transcriptome and corresponding clinical data from TCGA, Genecards, and GEO databases. Firstly, Univariate Cox regression analysis was performed and oxidative stress-related genes (ORGs) affecting overall survival of

Abbreviations: NSCLC, non-small cell lung cancer; SCLC, small cell lung cancer; LUAD, lung adenocarcinoma; LSCC, lung squamous cell carcinoma; LCLC, large cell lung cancer; TCGA, The Cancer Genome Atlas; ORG, oxidative stress-related gene; ROC, receiver operating characteristic; NCI, National Cancer Institute; FPKM, fragments per kilobase of transcript per million fragments mapped; TPM, transcripts per kilobase million; LASSO, least absolute shrinkage and selection operator; CNV, copy number variation; GISTIC, Genomic Identification of Significant Targets in Cancer; ssGSEA, single sample gene set enrichment analysis; GSVA, Gene Set Variation Analysis; ESTIMATE, Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression; GDSC, Genomics of Drug Sensitivity in Cancer.

LUAD were selected. Random Survival Forest and Least absolute shrinkage and selection operator (LASSO) analyses were used to screen and construct the OS-signature. We carried out efficacy evaluation for the OS-signature of LUAD using Kaplan-Meier and receiver operating characteristic (ROC) curves and the LUAD-cohort from GEO was used to validate the OS-signature. In addition, we evaluated the somatic mutation, genomic mutation, immunological characteristics, and sensitivity to chemotherapy for OS-signature. Finally, Quantitative Real-time PCR assays were used to detect the expression of the three genes establishing the OS-signature in LUAD cell lines.

Materials and methods

Collection and preprocessing the data of lung adenocarcinoma

The Cancer Genome Atlas (TCGA) is a major government-funded research initiative funded by the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) (21). Transcripts and clinical information of lung adenocarcinoma (LUAD) were extracted from TCGA (https://xenabrowser.OS/) (19, 22). We excluded LUAD patients without information of OS (Overall Survival), thus obtaining the clinical information and expression profiles of 502 LUAD patients. The data form of fragments per kilobase of transcript per million fragments mapped (FPKM) was transformed into transcripts per kilobase million (TPM) (22). We also used GEO data, including GSE37745 and GSE31210, generated from Affymetrix Human Genome U133 Plus 2.0 chip based on GPL570 platform as external validation groups (23). Genecards (https://www.genecards.org) is a comprehensive searchable gene database, where we can obtain information about almost all known human genes (24, 25). In order to obtain genes related to oxidative stress (oxidative stress related genes, ORGs), we set the screening threshold as relevance score>20 (26).

Establishment of the OS-signature for LUAD

After collection and preprocessing the data of LUAD, the Univariate Cox regression analysis was performed on the ORGs collected to identify ORGs with prognostic value (prognostic ORGs, $P<0.05$) (27). We used randomForestSRC package in R to execute Random Survival Forest (RSF) analysis, thus filtrating prognostic ORGs with greater value (variable importance>0.25) (28). Least absolute shrinkage and selection operator (LASSO) analysis is a compression estimation method for linear model (29). The regression coefficients can be compressed by minimizing the sum of residual squares under the constraint that the sum of absolute values of various coefficients is less than a constant, thus getting a sparse model (29). This model can effectively select variables for high dimensional and collinearity data (30). The Cox regression model for LASSO analysis provided by glmnet package in R software (31) was used to calculate the OS-scores and construct the prognostic OS-signature for LUAD.

Efficacy evaluation for the OS-signature of LUAD

The survminer package in R software was used to select the best separated value (cutoff value) of gene expression or OS-scores. The survival curves (Kaplan-Meier curves) of the high- or low-risk groups were drawn, and the survival differences between the two groups were compared (32). The receiver operating characteristic curve (ROC) is also known as the sensitivity curve (33). The research method is to analyze the Area Under the ROC Curve (AUC) of the research objects to judge the recognition ability of different diagnostic test objects for diseases. The timeROC package of R software was used to draw time-dependent (1-, 3-, and 5-year) ROC curves to evaluate the diagnostic efficacy and predictive effect of OS-signature for LUAD.

Genomic mutation analysis for OS-signature in LUAD

Somatic mutation and copy number variation (CNV) data of LUAD patients were downloaded from cBioPortal (http://www.cbioportal.org/datasets) (34) and FireBrowse (http://firebrowse.org/) (35) respectively. To determine the mutational burden in LUAD patients, the total number of non-synonymous mutations in LUAD was calculated. Somatic alterations of driver genes in LUAD were evaluated by OS-signature grouping. The R software package maftool was used to identify the driver genes of LUAD and analyze the top 20 driver genes with the highest frequency of change. We assessed genomic characteristics by Genomic Identification of Significant Targets in Cancer 2.0 (GISTIC 2.0, https://gatk.broadinstitute.org) analysis (36).

Characteristic analysis of tumor infiltrating immune cells

According to the transcriptome expression data from TCGA-LUAD cohort, the single sample gene set enrichment analysis (ssGSEA) algorithm in R package GSVA (Gene Set Variation Analysis) was used to rank the genes contained in the sample according to their expression level from high to low, and the rank of all genes was obtained (37). Each type of immune cell is characterized by a separate subset of genes. In this study, 783 genes were used to characterize 28 common immune infiltrating cell types. According to the background gene sets generated by each sample and arranged according to the expression situation, the enrichment scores of all samples for 28 types of immune infiltrating cells in each subset could be obtained by systematic calculation (38, 39). The advantages of this method are that it uses gene sets instead of single genes to annotate immune cell subsets and combined with multiple validation methods to improve the annotation accuracy of enrichment scores. The ESTIMATE ((The Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression) method was used to evaluate the ESTIMATE score, immune score, and stromal score of each LUAD patient (40). Besides, we assessed the levels of six kinds immune infiltrating cells (B cell, T cell CD4, T cell CD8, Neutrophil, Macrophage, and DC) *via* Tumor Immune Estimation Resource 2.0

(TIMER 2.0; http://timer.cistrome.org/) (41). We also used the MCPcounter algorithm to estimate the relative proportions of ten immune cells (T cells, CD8 T cells, Cytotoxic lymphocytes, B lineage, NK cells, Monocytic lineage, Myeloid dendritic cells, Neutrophils, Endothelial cells, and Fibroblasts) in LUAD (42). We extracted seven kinds of immunomodulators (Antigen presentation, Cell adhesion, Co-inhibitor, Co-stimulator, Ligand, Other, and Receptor) from previous study to explore the association between OS-scores and immune processes (43).

Chemotherapeutics sensitivity analysis for OS-signature in LUAD

The Genomics of Drug Sensitivity in Cancer (GDSC) database was used to screen the wide range of chemotherapeutics for LUAD (44). The prediction model was constructed based on Ridge's regression between drug sensitivity and expression profile of cell lines using pRRophetic algorithm (45, 46). Subsequently, we calculated the IC50 value of corresponding chemotherapeutics for each LUAD patients.

Quantitative real-time PCR assays detecting gene expression in cell lines

The human normal lung epithelial cells named BEAS-2B was supplied by Beyotime Biotechnology (Hangzhou, China). The LUAD cell lines, including A-549 and NCI-H1299, were purchased from National Collection of Authenticated Cell Cultures (Shanghai, China). BEAS-2B and NCI-H1299 were cultured in 90% RPMI (Roswell Park Memorial Institute)-1640 with 10% FBS (fetal bovine serum). A-549 was cultured in 89% F-12K + 10% FBS + 1% Glutamax. We extracted the total RNA of the cell lines by RNAsimple Total RNA Kit (Tiangen, China). Whereafter, to acquire cDNA, we reverse transcribed the cell RNA that we have obtained applying PrimeScript RT reagent Kit (Takara, Otsu, Japan). Finally, based on the premixed system of 2 μL cDNA with SYBR Premix Ex Taq (Takara, Otsu, Japan) and primers, we detected the expression values of related genes in cell lines by Applied Biosystems StepOne Plus Real-Time PCR system (Life Technologies, Grand Island, NY, USA). The primers of the target gene were supplied by Sangon Biotech (Shanghai, China). The sequences of the primers used were listed in Table 1.

TABLE 1 The primer sequences in PCR analysis.

Symbol	Sequences (5′-3′)
MRPL44-F	TTGAAGACGAGTACCCAGACA
MRPL44-R	GGGCTCCAATAACTGCAAAGAA
CYCS-F	CTTTGGGCGGAAGACAGGTC
CYCS-R	TTATTGGCGGCTGTGTAAGAG
CAT-F	TGGGATCTCGTTGGAAATAACAC
CAT-R	TCAGGACGTAGGCTCCAGAAG
GAPDH-F	GGAGCGAGATCCCTCCAAAAT
GAPDH-R	GGCTGTTGTCATACTTCTCATGG

Results

Establishment of OS-signature for patients with LUAD

For LUAD, we carried out the Univariate Cox regression analysis on a total of 80 OXRGs matched (relevance score>20). We identified a total of ten OXRGs with prognostic value (Figure 1A), including eight prognostic genes with HR>1 (MRPL44, CYCS, G3BP1, GFM1, SOD1, TXN, OSGIN2, and CRP) and two prognostic genes with HR<1 (CAT and XBP1). Hence, we observed eight malignant factors and two protective factors for patients with LUAD (Figure 1A). Whereafter, we conducted Lasso (Figures 1–C) and RSF (Figures 1D–E) analyses on the ten prognostic ORGs gained. The OS-signature ended up containing three genes: CAT, CYCS, and MRPL44 (Figures 1B–E). The three prognostic ORGs selected above were weighted by the regression coefficients of Lasso regression model, and finally the calculation formula of OS-signature for prognosis assessment of LUAD was obtained: OS-score =1.0002*CYCS - 0.9272*CAT + 1.7096*MRPL44. Figures 1B, C displayed the lambda selection diagram of the three genes in the OS-signature. The distribution of error rates generated by RSF analysis was shown in Figures 1D, E.

Evaluating the efficacy of OS-signature for LUAD

After establishing the OS-signature based on three prognostic ORGs (CAT, CYCS, and MRPL44) for LUAD, we computed the OS-score for each LUAD patient based on the LASSO coefficients and expression value for each ORG. We compared the OS-score of LUAD patients in TCGA database among clinical features (Stage, Gender, Age and Survival Status) and the expression values of the three ORGs included in the OS-signature, which was shown in the heatmap (Figure 2A). Overall, patients with high OS-scores were more likely to have high expression of MRPL44 and CYCS, whereas patients with high OS-scores were strongly associated with low expression of CAT (Figure 2A). Kaplan-Meier analysis was used to analyze the survival and prognosis of LUAD patients in TCGA. As shown in the Figure 2B, patients with low OS-score had a better prognosis, while patients with high OS-score had a worse prognosis (Figure 2B). The AUCs of 1-year (AUC=0.688), 3-year (AUC=0.668), and 5-year (AUC=0.660) survival ROC curves predicted by the OS-signature were all larger than 0.66, suggesting the efficiency of OS-signature in predicting prognosis for LUAD to a certain extent (Figure 2C). To further verify the conclusion, two independent external datasets (GSE37745 and GSE31210) were included in our study, and the significantly better clinical outcomes of LUAD patients with lower OS-scores were observed (Figures 3A, B). Therefore, OS-signature may serve as a malignancy factor for LUAD.

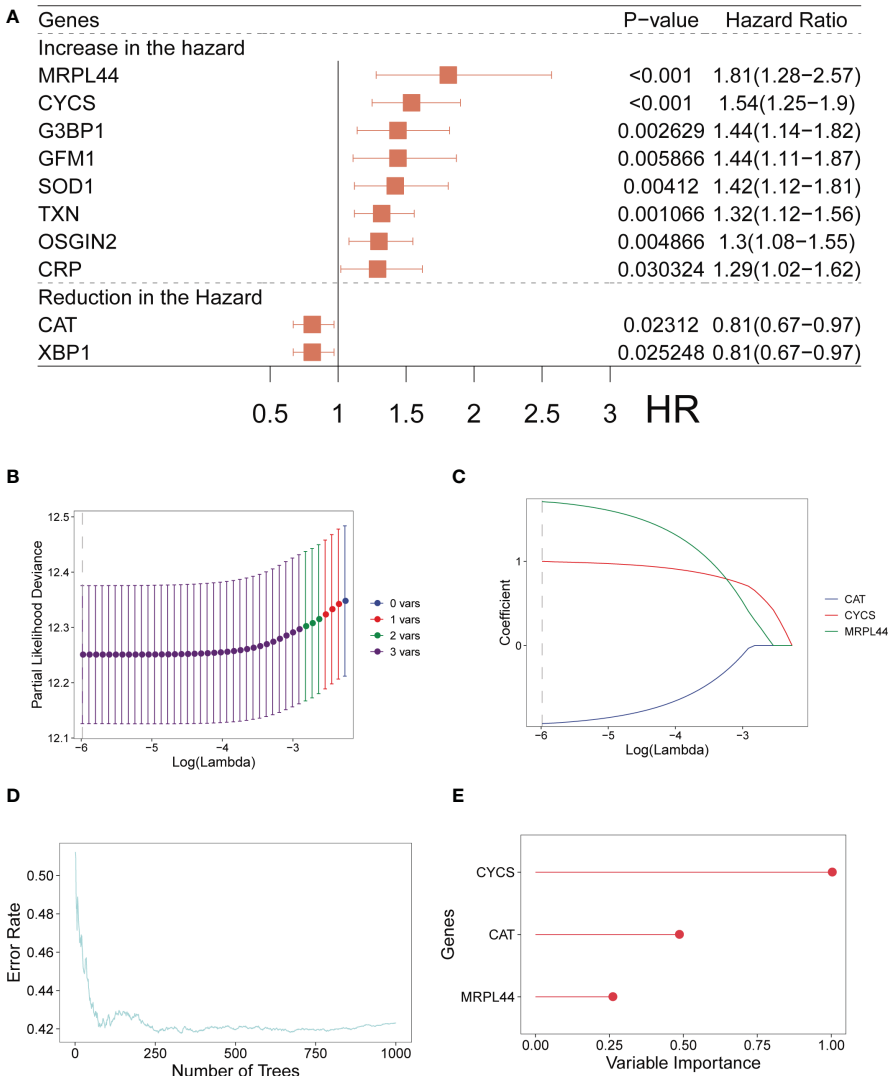


FIGURE 1
Establishment of OS-signature for patients with LUAD. **(A)** Forest plot for Univariate Cox regression analysis identifying ten oxidative stress related genes (MRPL44, CYCS, G3BP1, GEM1, SOD1, TXN, OSGIN2, CRP, CAT, and XBP1). **(B, C)** Lambda selection diagram for Least Absolute Shrinkage and Selection Operator (Lasso) analysis identifying three oxidative stress related genes (CAT, CYCS, and MRPL44) in the OS-signature. **(D)** The distribution of error rates in Random Survival Forest model. **(E)** The distribution of the variable relative importance of 12 TRP-related genes (variable importance>0.25).

Genomic mutation analysis for OS-signature in LUAD

We carried out genomic mutation analysis for OS-signature in LUAD. From the waterfall diagram (Figures 4A, B), we could find that TP53, TTN, CSMD3, MUC16, RYR2, ZFHX4, LRP1B, USH2A, SPTA1, XIRP2, KEAP1, KRAS, FLG, CSMD1, MUC17, ADAMTS12, APOB, PAPPA2, COL11A1, and FAT3 were the top 20 genes with the highest mutation rate in LUAD patients with high OS-scores (Figure 4A). TP53, TTN, MUC16, RYR2, CSMD3, LRP1B, USH2A, KRAS, FLG, ZFHX4, ANK2, SPTA1, XIRP2, ZNF536, NAV3, COL11A1, FAT3, PCDH15, PCLO, and TNR were the top 20 genes with the highest mutation rate in LUAD patients with low OS-scores (Figure 4B). Thus, the mutation rates of TP53, TTN, MUC16, RYR2, ZFHX4, LRP1B, USH2A, SPTA1, XIRP2, KRAS, FLG, COL11A1, and FAT3 in the two subgroups were both relatively high. We performed Pair-wise Fisher's Exact test to detect mutually exclusive or co-occurrence events (Figures 4C, D). We also Draw forest plot for mutation differences between cohorts.

Genomic characterization landscapes of LUAD patients with high OS-scores or patients with low OS-scores were analyzed by GISTIC algorithm and shown in Figure 5A. Figure 5B showed the plots significantly altered cytobands as a function of number samples in which it is altered and number genes it contains. Figure 5C showed a genomic plot with segments highlighting significant Amplifications and Deletion regions. Further, we drew the detailed amplificated or deleted CNV onco-plots of high OS-score and low OS-score subgroups (Figure 5D).

Characteristic analysis of tumor infiltrating immune cells

Since immunomodulators (IMs) play a critical role in tumor immunotherapy, we assessed the correlation between the IMs

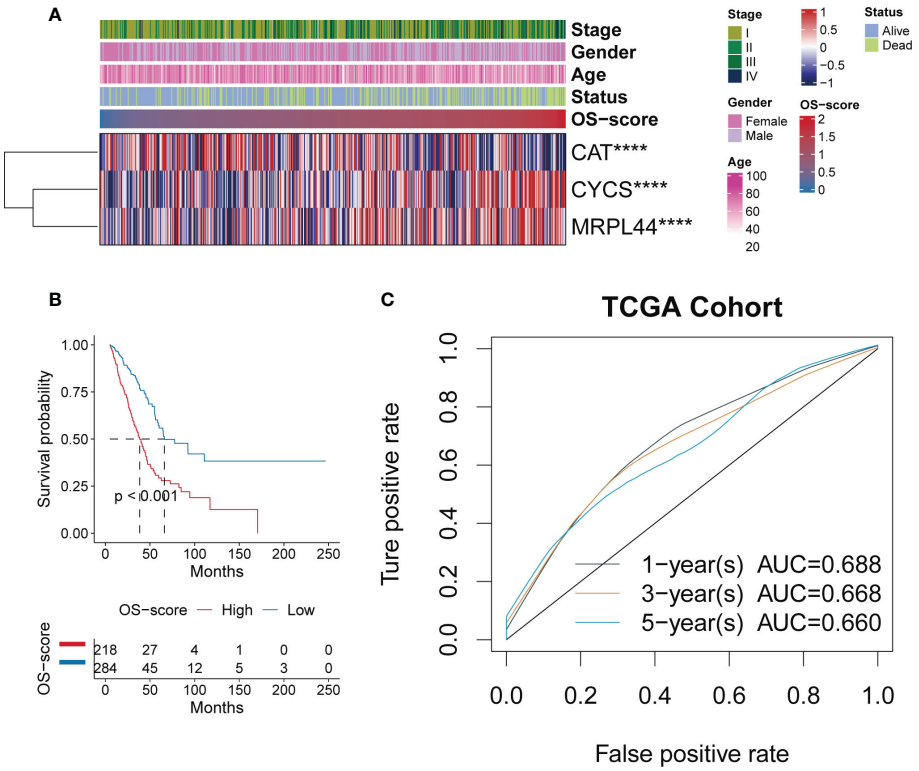


FIGURE 2
Evaluating the efficacy of OS-signature in TCGA for LUAD. **(A)** The heatmap displaying the distribution of the three oxidative stress related genes (CAT, CYCS, and MRPL44) in the OS-signature, clinicopathological characteristics (Stage, Gender, Age, Survival Status), and OS-score. Red represents high gene expression and blue represents low gene expression. **(B)** Kaplan-Meier curves displaying the correlation between the OS-score and LUAD patients. The blue curve represents the patients with lower OS-score, and the red curve represents patients with higher OS-score. **(C)** The 1-year (0.688), 3-year (0.668), 5-year (0.660) survival ROC curves predicted by the OS-signature. Different colored curves represent different years. ****p<0.0001.

levels (Antigen presentation, Cell adhesion, Co-inhibitor, Co-stimulator, Ligand, Other, and Receptor). It was clearly observed that patients with high OS-scores had lower IMs levels, with a few exceptions, such as CD276, TNFSF9, and HMGB1 (Figure 6A). From a general view, the level of stromal score, immune score, ESTIMATE score and infiltrating immune cell populations decreased as the OS-scores increased (Figure 6B). It was worth mentioning that patients with higher OS-scores were more likely to have higher tumor purity (Figure 6B).

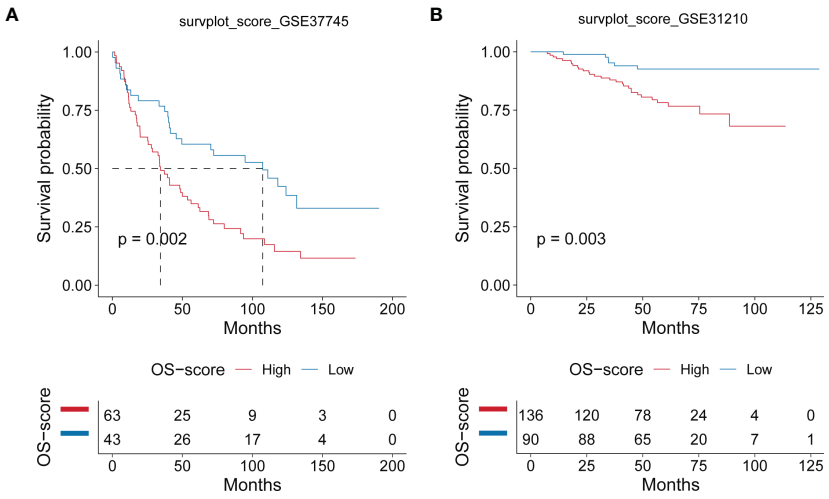


FIGURE 3
Evaluating the efficacy of OS-signature in GEO for LUAD. **(A, B)** Kaplan-Meier curves displaying the correlation between the OS-score and LUAD patients in GSE37745 **(A)** and GSE31210 **(B)**. The blue curve represents the patients with lower OS-score, and the red curve represents patients with higher OS-score.

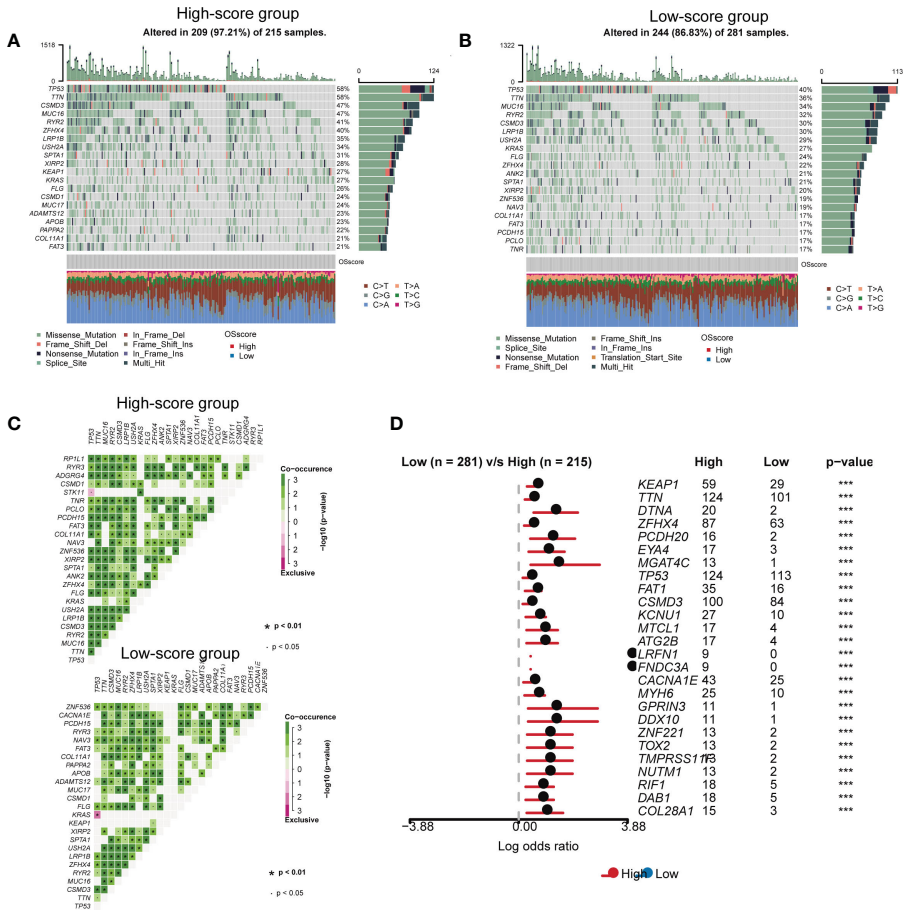


FIGURE 4
The somatic mutation features of the established OS-signature for LUAD. **(A, B)** The waterfall plot of somatic mutation features established with high **(A)** and low **(B)** OS-score. **(C)** We performed Pair-wise Fisher's Exact test to detect mutually exclusive or co-occurrence events. **(D)** Forest plot for mutation differences between cohorts.

Chemotherapeutics sensitivity analysis for OS-signature in LUAD

In order to find more effective chemotherapeutics drugs for LUAD patients with high OS-scores, we evaluated the differences in chemotherapeutics sensitivity between subgroups with high OS-score or low OS-score as described in the MATERIALS AND METHODS. The IC50 levels of nine chemotherapy drugs (Osimertinib_1919, Sapitinib_1549, Acetalax_1804, Ibrutinib_1799, Erlotinib_1168, Gefitinib_1010, AZD3759_1915, Afatinib_1032, and Lapatinib_1558) were compared between subgroups with high OS-score or low OS-score. We found that the IC50 values of the nine chemotherapy drugs were lower in LUAD patients with high OS-scores than that of LUAD patients with low OS-scores, suggesting LUAD patients with high OS-scores may be more sensitive to these nine chemotherapeutics drugs (Figure 7).

Quantitative real-time PCR

We selected the three genes in the OS-signature to detect their expression in cell lines.

As could be seen from the survival curves, the higher the expression of MRPL44 (Figure 8A) and CYCS (Figure 8B), the worse the prognosis, while the opposite was true for CAT (Figure 8C). Compared with control cell lines (BEAS-2B), MRPL44 (Figure 8D) and CYCS (Figure 8E) were significantly higher expressed in cancer cell lines (A549 and H1299), while CAT (Figure 8F) was significantly lower expressed.

Discussion

Lung cancer is a malignant tumor originating from the bronchial epithelium. According to histopathological classification, lung cancer is divided into non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC is the main pathological type of lung cancer, and lung adenocarcinoma (LUAD) accounts for the vast majority of NSCLC. Lung cancer ranks second only to breast cancer in incidence and is the most important cause of cancer-related deaths. Late diagnosis, poor sensitivity to chemoradiotherapy, acquired resistance to targeted therapy and other related factors can lead to poor prognosis of patients with lung cancer (47, 48). At present, histopathological diagnosis and

10.3389/fonc.2023.1078697

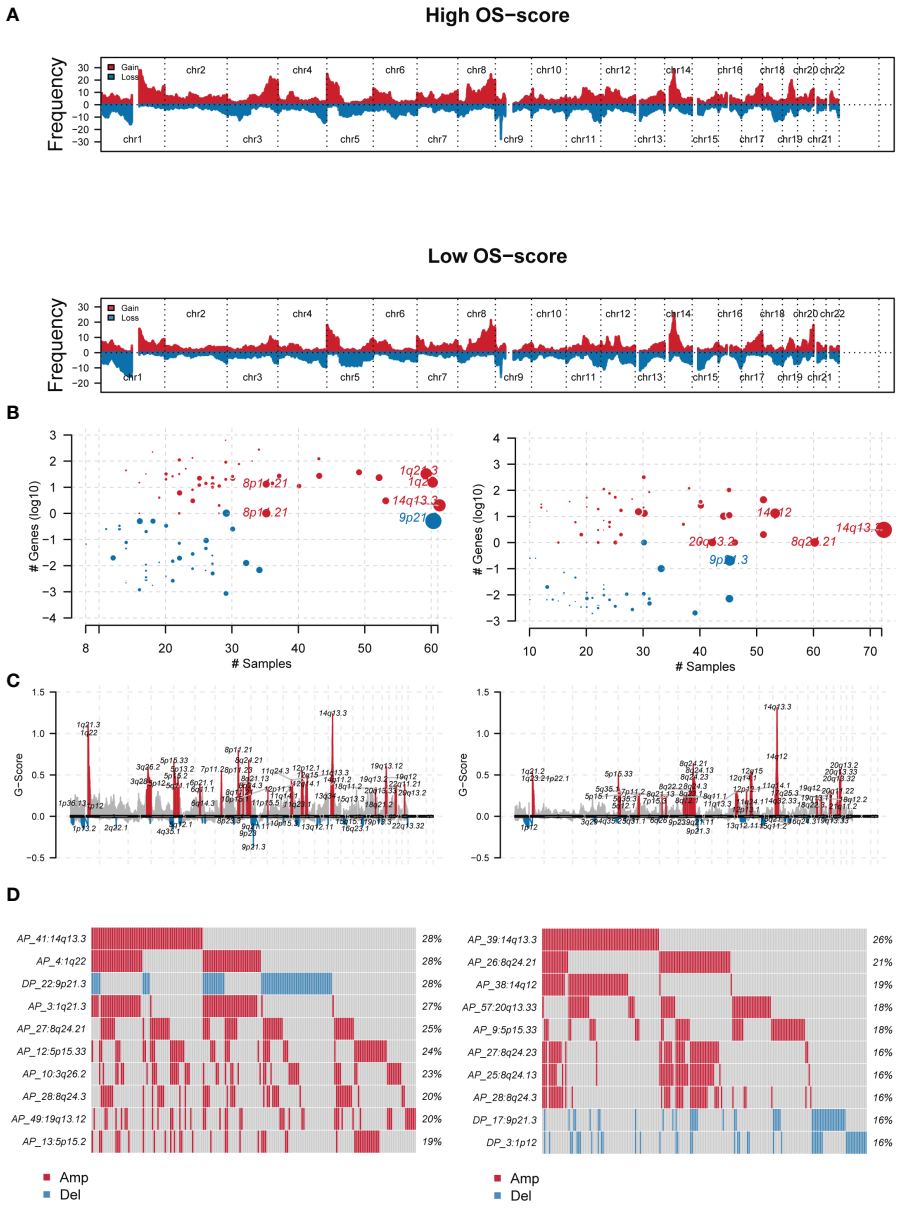


FIGURE 5
The genomic mutation analysis of the established OS-signature for LUAD. **(A)** Genomic characterization landscape of groups with high OS-scores or low OS-scores. **(B)** Plots significantly altered cytobands as a function of number samples in which it is altered and number genes it contains. Size of each bubble is according to -log10 transformed q values. **(C)** A genomic plot with segments highlighting signififcant Amplifications and Deletion regions. **(D)** The detailed amplificated or deleted CNV onco-plots of groups with high OS-scores or low OS-scores. #p<0.05.

tumor staging system are still the main basis for predicting the prognosis and survival of lung cancer patients. However, traditional methods cannot accurately assess the prognosis of patients with LUAD. In addition, Computed Tomography (CT) and serum tumor markers such as carcinoembryonic antigen (CEA) are often used to determine the prognosis of lung cancer. However, traditional methods are limited by cumulative radiation damage, low sensitivity and specificity (49, 50). Therefore, clinicians need an accurate prognostic prediction model to help optimize the treatment strategy of LUAD patients. Bioinformatics is one of the emerging fields of biological research. It uses mathematics, statistics and computer technology to process and analyze biological data. In our study, extracting data from public database, we identified eight prognostic genes with HR>1 (MRPL44, CYCS, G3BP1, GFM1, SOD1, TXN, OSGIN2, and CRP) and two prognostic genes with HR<1 (CAT and XBP1). And the OS-signature could be used for prognosis and treatment prediction in LUAD.

Because of the functional correlation between genes in a cell, diseases are rarely the result of abnormalities in a single gene, but rather result from abnormalities in a complex intracellular gene network (51–53). Like most diseases, the occurrence and development of LUAD is a complex process involving multiple genes and multiple pathogenic mechanisms, involving the activation of proto-oncogenes and the inactivation or mutation of tumor suppressor genes (54, 55). Therefore, the application of network for gene interaction in LAC research can simplify and visualize complex and high-throughput data. Compared with the focus on local gene function in single gene and single molecule

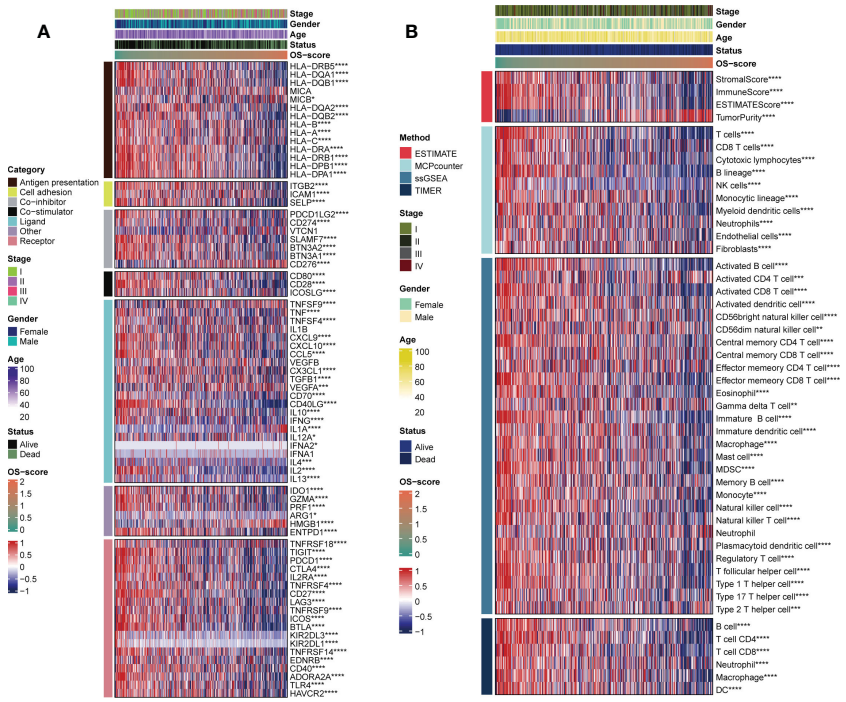


FIGURE 6
Evaluation of immunological characteristics for OS-signature. **(A)** Correlation of OS-scores with seven immunomodulators in LUAD. Red represents high enrichment score, and blue represents low enrichment score. **(B)** Heatmap displaying the abundance of infiltrating immune cell populations with different OS-scores. *p<0.05, ***p<0.001, ****p<0.0001.

biological research methods, network analysis focuses more on the integrity and systematization of biological processes (51). In the OS-signature, there were three ORGs (CAT, CYCS, and MRPL44), forming a network to predict the prognosis of LUAD. It is more reliable to explore the occurrence and development of LUAD from the perspective of multiple genes.

For early and mid-stage NSCLC that cannot be completely resected by surgery, and for some locally evolved or metastatic NSCLC that is advanced or advanced (stage IIIA-IV), comprehensive systemic and local combination therapy can be used, including surgical resection, chemotherapy, radiotherapy, targeted therapy and immunotherapy. At present, the 3rd generation chemotherapy drugs, including Docetaxel, Vinorelbine, Gemcitabine and Paclitaxel, have been widely used in clinical practice, combining platinum drugs to develop personalized treatment plans for patients. Radiotherapy is an effective means of local treatment of lung cancer, which plays a positive role in slowing down the clinical symptoms, prolonging the survival time and improving the quality of life of patients with advanced lung cancer. These treatment methods have been widely studied and applied at home and abroad. Genomic studies have shown that adenocarcinoma and squamous cell carcinoma have significantly different gene mutation types, and tyrosine kinase inhibitor (TKI) can be used to inhibit the catalytic phosphorylation of the corresponding kinases in the treatment of NSCLC patients with significant clinical benefits. Genomic studies have shown that adenocarcinoma and squamous cell carcinoma have significantly different gene mutation types, and tyrosine kinase inhibitor (TKI) can be used to inhibit the catalytic phosphorylation of the corresponding kinases in the treatment of NSCLC patients with significant clinical benefits (56, 57). A variety of effective and well-tolerated TKIs targets, including EGFR, ALK, ROSI, HER2, etc., have emerged continuously, and promoted significant progress in cancer treatment. For example, EGFR driver gene mutations have a high incidence in various subtypes of NSCLC. The most common EGFR mutations include exon19 deletion (delE746-750, etc.) or exon 21 arginine substitution leucine (L858R) mutation. EGFR inhibitors such as Gefitinib, Erlotinib, Afatinib, or Osimertinib play an important role in the treatment of NSCLC patients (58). However, some studies have shown that the proportion of NSCLC patients carrying EGFR mutations is about 30-40%, and there are still a large number of patients who cannot benefit directly from targeted therapy (59). With the development of Crizotinib and next-generation ALK-TKIs, considerable progress has been made in the treatment of patients with ALK recombinant NSCLC (60). Crizotinib, a first-generation ALK inhibitor originally approved for patients with ALK-positive NSCLC, was found to have a median progression-free survival of 8-10 months in treated patients (61). Subsequent randomized controlled trials compared Crizotinib with chemotherapy in patients undergoing treatment with a significant improvement in progression-free survival. Subsequently, second-generation ALK inhibitors Ceritinib, Alectinib and Brigatinib were developed to overcome Crizotinib resistance in patients (62). So far, other treatments, including third-generation ALK inhibitors Lorlatinib, Entrectinib and Ensartinib, have shown better results (60). For the above mentioned chemotherapeutic drugs and small molecule targeted therapy drugs, the Genomics of Drug Sensitivity in Cancer (GDSC) database was created. The immediate goal is to identify potential therapeutic biomarkers that may predict drug response (chemotherapeutic

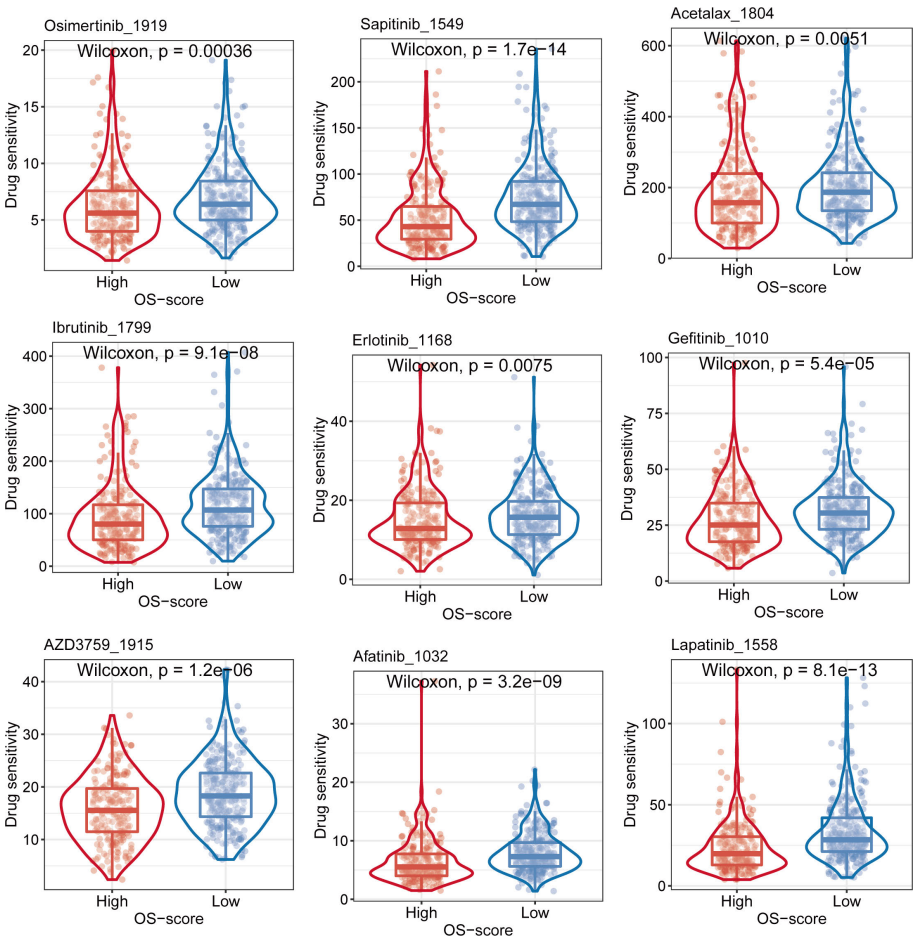


FIGURE 7
Evaluation of sensitivity to chemotherapy for OS-signature.

drugs, small molecule targeted drugs, and other drugs), while the ultimate goal is to improve the current status of cancer treatment based on biomarkers (44, 63). It has been shown that changes in the tumor genome directly affect the therapeutic effect of the tumor (64). With the emergence of novel compounds, the screening of predictive biomarkers in their early development process will have a profound impact on the entire process of new cancer drug development, including its design, development cost and final outcome (64). Based on the clinical and basic research background, researchers present the results of large-scale drug screening in human cancer cell lines in GDSC, a database that combines detailed genomic profiles and gene expression analysis to systematically provide biomarker identification patterns for drug sensitivity science for a variety of cancer drugs. In our study, we compared the IC50 levels of Osimertinib_1919, Sapitinib_1549, Acetalax_1804, Ibrutinib_1799, Erlotinib_1168, Gefitinib_1010, AZD3759_1915, Afatinib_1032, and Lapatinib_1558 between subgroups with high OS-score or low OS-score based on GDSC database. We found that LUAD patients with high OS-scores may be more sensitive to these nine chemotherapeutics drugs. Our study will provide reference for the treatment of LUAD.

Recent studies have shown that tumor microenvironment (TME) plays an important role in the development and treatment of tumors (65). TME refers to the microenvironment surrounding the occurrence, growth and metastasis of tumor cells, including not only the tumor cells themselves, but also the immune cells, inflammatory cells, fibroblasts, various signaling molecules, extracellular matrix and blood vessels (66). To fully understand and overcome the complexity of TME is helpful for clinicians to provide more feasible and precise individualized treatment plan for tumor treatment. The rapid development of single-cell sequencing, second-generation sequencing and other technologies has gradually deepened researchers' understanding of the relationship between T cells and other immune cell populations and immunotherapy. Tumor-associated immune cells play an important role in tumor spread, recurrence, metastasis and influencing immunotherapy treatment (67). They can be used as biomarkers to predict the efficacy of immunotherapy drugs or predict the prognosis of patients (67). Increased levels of tumor-infiltrating lymphocytes (TILs), such as CD4+T cells and CD8+T cells, are associated with immunotherapy response and longer survival (68). Immune checkpoint inhibition activates existing TILs, which recognize and eliminate abnormal and tumor cells, and TILs play a key role in immunotherapy response. Studies have shown that increased T-cell infiltration and increased IFN-γ-related mRNA expression can increase ICIs (immune checkpoint inhibitors) benefit and significantly improve patient prognosis in a variety of tumor types (69, 70). In advanced NSCLC patients, increased expression of CD8+ TILs detected by IHC or CD8A mRNA transcripts was associated with prolonged PFS treatment

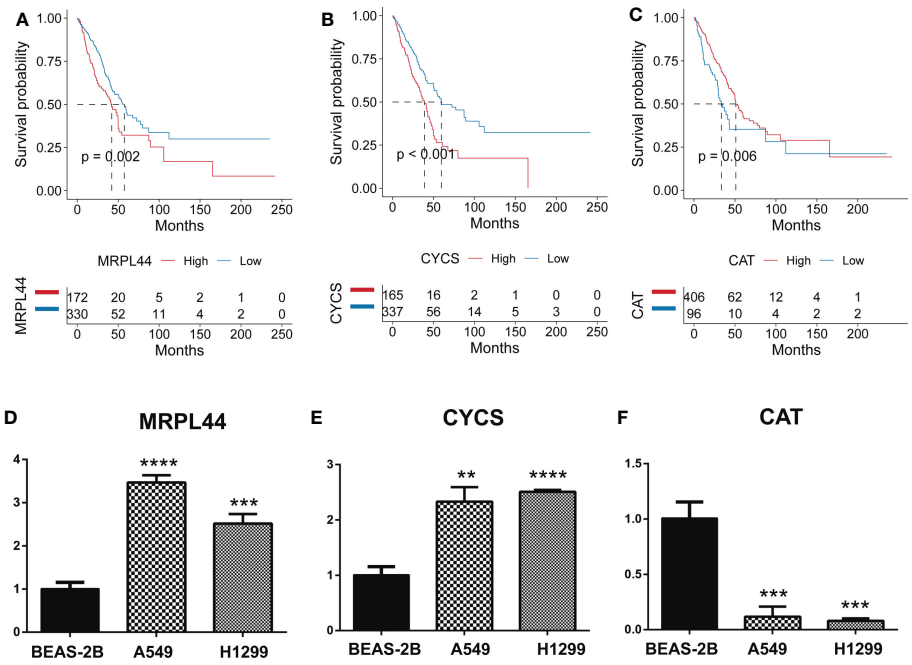


FIGURE 8
Quantitative Real-time PCR. **(A–C)** Kaplan-Meier curves displaying the correlation between the expression of the signature genes, including MRPL44 **(A)**, CYCS **(B)**, and CAT **(C)**, and the survival status of LUAD patients. The blue curve represents the patients with lower gene expression, and the red curve represents patients with higher gene expression. **(D–F)** Quantitative Real-time PCR assays using cell lines for MRPL44 **(D)**, CYCS **(E)**, and CAT **(F)**. **p<0.01, ***p<0.001, ****p<0.0001.

with PD-L1 inhibitors, especially in combination with PD-L1 mRNA and protein expression, suggesting that integrated biomarkers may provide higher predictive value (71). Another study using multiple quantitative immunofluorescences to detect TIL in paraffin tumor specimens found that ICI treatment resulted in lasting clinical benefits and longer OS in NSCLC patients with increased CD3+ T-cell infiltration (72). In addition, studies have found that tumor-associated macrophages (TAMs) secrete interleukin-10 (interleukin-10), Il-10, Transforming growth factor-B (TGFb) and other immunosuppressive cytokines play a variety of tumor-promoting effects, which increase the density of TAM and inhibit other related immune cells (73).

The limitations this study remain. The OS-signature we constructed and validated by retrospectively using the public database hence, more prospective studies are needed for clinical practicability. We selected the three genes in the OS-signature to detect their expression in cell lines. Biological experiments in this study are lacking, and more wet experiments are needed to explore the function of related genes.

In conclusion, immunotherapy by regulating the immune microenvironment may become a promising new strategy for cancer treatment. The precise regulation of immune gene expression is the key to generate strong immunity and intervene the development of cancer. In our study, we found that patients with high OS-scores had lower immunomodulators levels except CD276, TNFSF9, and HMGB1. From a general view, the level of infiltrating immune cell populations decreased as the OS-scores increased. It is necessary to further study the tumor microenvironment (TME) of lung cancer.

Data availability statement

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

Author contributions

WW conceived, designed, and supervised the study. HP performed data analysis and drafted the manuscript. XL performed PCR experiments. YL and CW assisted with the analysis. All authors contributed to the article and approved the submitted version.

Acknowledgments

We greatly thank the Research on application of public welfare technology of Zhejiang Province (GF22H085719) and Health Science and Technology Plan of Zhejiang Provincial (2021KY400) for supporting the research and work.

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. Toumazis I, Bastani M, Han SS, Plevritis SK. Risk-based lung cancer screening: A systematic review. *Lung Cancer* (2020) 147:154–86. doi: 10.1016/j.lungcan.2020.07.007

3. Nasim F, Sabath BF, Eapen GA. Lung cancer. *Med Clin North Am* (2019) 103:463–73. doi: 10.1016/j.mcna.2018.12.006

4. Pallis AG, Syrigos KN. Lung cancer in never smokers: disease characteristics and risk factors. *Crit Rev Oncol Hematol* (2013) 88:494–503. doi: 10.1016/j.critrevonc.2013.06.011

5. Blandin Knight S, Crosbie PA, Balata H, Chudziak J, Hussell T, Dive C, et al. Progress and prospects of early detection in lung cancer. *Open Biol* (2017) 7. doi: 10.1098/rsob.170070

6. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* (2020) 70:7–30. doi: 10.3322/caac.21590

7. Travis WD. Pathology of lung cancer. *Clin Chest Med* (2011) 32:669–92. doi: 10.1016/j.ccm.2011.08.005

8. Asamura H, Goya T, Koshiishi Y, Sohara Y, Eguchi K, Mori M, et al. A Japanese lung cancer registry study: prognosis of 13,010 resected lung cancers. *J Thorac Oncol* (2008) 3:46–52. doi: 10.1097/JTO.0b013e31815e8577

9. Wu K, House L, Liu W, Cho WCS. Personalized targeted therapy for lung cancer. *Int J Mol Sci* (2012) 13:11471–96. doi: 10.3390/ijms130911471

10. Chin L, Andersen JN, Futreal PA. Cancer genomics: from discovery science to personalized medicine. *Nat Med* (2011) 17:297–303. doi: 10.1038/nm.2323

11. Hernández Patiño CE, Jaime-Muñoz G, Resendis-Antonio O. Systems biology of cancer: moving toward the integrative study of the metabolic alterations in cancer cells. *Front Physiol* (2012) 3:481. doi: 10.3389/fphys.2012.00481

12. Chansky K, Sculier JP, Crowley JJ, Giroux D, Van Meerbeeck J, Goldstraw P, et al. The international association for the study of lung cancer staging project: prognostic factors and pathologic TNM stage in surgically managed non-small cell lung cancer. *J Thorac Oncol* (2009) 4:792–801. doi: 10.1097/JTO.0b013e3181a7716e

13. 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

14. Klaunig JE. Oxidative stress and cancer. *Curr Pharm Des* (2018) 24:4771–8. doi: 10.2174/1381612825666190215121712

15. Apel K, Hirt H. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu Rev Plant Biol* (2004) 55:373–99. doi: 10.1146/annurev.arplant.55.031903.141701

16. Udensi UK, Tchounwou PB. Dual effect of oxidative stress on leukemia cancer induction and treatment. *J Exp Clin Cancer Res* (2014) 33:106. doi: 10.1186/s13046-014-0106-5

17. Jelic MD, Mandic AD, Maricic SM, Srdjenovic BU. Oxidative stress and its role in cancer. *J Cancer Res Ther* (2021) 17:22–8. doi: 10.4103/jcrt.JCRT_862_16

18. Sosa V, Moliné T, Somoza R, Paciucci R, Kondoh H, LLeonart ME, et al. Oxidative stress and cancer: an overview. *Ageing Res Rev* (2013) 12:376–90. doi: 10.1016/j.arr.2012.10.004

19. Wang Z, Jensen MA, Zenklusen JC. A practical guide to the cancer genome atlas (TCGA). *Methods Mol Biol* (2016) 1418:111–41. doi: 10.1007/978-1-4939-3578-9_6

20. Yuan Q, Deng D, Pan C, Ren J, Wei T, Wu Z, et al. Integration of transcriptomics, proteomics, and metabolomics data to reveal HER2-associated metabolic heterogeneity in gastric cancer with response to immunotherapy and neoadjuvant chemotherapy. *Front Immunol* (2022) 13:951137. doi: 10.3389/fimmu.2022.951137

21. Tomczak K, Czerwińska P, Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)* (2015) 19:A68–77. doi: 10.5114/wo.2014.47136

22. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, et al. Visualizing and interpreting cancer genomics data. *via Xena platform. Nat Biotechnol* (2020) 38:675–8. doi: 10.1038/s41587-020-0546-8

23. He R, Zuo S. A robust 8-gene prognostic signature for early-stage non-small cell lung cancer. *Front Oncol* (2019) 9:693. doi: 10.3389/fonc.2019.00693

24. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards suite: From gene data mining to disease genome sequence analyses. *Curr Protoc Bioinf* (2016) 54:1.30.31–31.30.33. doi: 10.1002/cpbi.5

25. Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)* (2017) 2017). doi: 10.1093/database/bax028

26. Wu Z, Wang L, Wen Z, Yao J. Integrated analysis identifies oxidative stress genes associated with progression and prognosis in gastric cancer. *Sci Rep* (2021) 11:3292. doi: 10.1038/s41598-021-82976-w

27. van Dijk PC, Jager KJ, Zwinderman AH, Zoccali C, Dekker FW. The analysis of survival data in nephrology: basic concepts and methods of cox regression. *Kidney Int* (2008) 74:705–9. doi: 10.1038/ki.2008.294

28. Taylor JM. Random survival forests. *J Thorac Oncol* (2011) 6:1974–5. doi: 10.1097/JTO.0b013e318233d835

29. Tibshirani R. The lasso method for variable selection in the cox model. *Stat Med* (1997) 16:385–95. doi: 10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3

30. Gui J, Li H. Penalized cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data. *Bioinformatics* (2005) 21:3001–8. doi: 10.1093/bioinformatics/bti422

31. Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med* (2007) 26:5512–28. doi: 10.1002/sim.3148

32. Ranstam J, Cook JA. Kaplan-Meier Curve. *Br J Surg* (2017) 104:442. doi: 10.1002/bjs.10238

33. Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. *Cjem* (2006) 8:19–20. doi: 10.1017/s1481803500013336

34. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* (2013) 6:l1. doi: 10.1126/scisignal.2004088

35. Chabanais J, Labrousse F, Chaunavel A, Germot A, Maftah A. POFUT1 as a promising novel biomarker of colorectal cancer. *Cancers (Basel)* (2018) 10. doi: 10.3390/cancers10110411

36. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G, et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* (2011) 12:R41. doi: 10.1186/gb-2011-12-4-r41

37. Shen S, Wang G, Zhang R, Zhao Y, Yu H, Wei Y, et al. Development and validation of an immune gene-set based prognostic signature in ovarian cancer. *EBioMedicine* (2019) 40:318–26. doi: 10.1016/j.ebiom.2018.12.054

38. Xiao B, Liu L, Li A, Xiang C, Wang P, Li H, et al. Identification and verification of immune-related gene prognostic signature based on ssGSEA for osteosarcoma. *Front Oncol* (2020) 10:607622. doi: 10.3389/fonc.2020.607622

39. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf* (2013) 14:7. doi: 10.1186/1471-2105-14-7

40. Yoshihara K, Shahmoradgoli M, Martínez 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

41. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res* (2017) 77:e108–10. doi: 10.1158/0008-5472.Can-17-0307

42. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* (2016) 17:218. doi: 10.1186/s13059-016-1070-5

43. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. *Immunity* (2018) 48:812–830.e814. doi: 10.1016/j.immuni.2018.03.023

44. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* (2013) 41:D955–961. doi: 10.1093/nar/gks1111

45. Geeleher P, Cox NJ, Huang RS. Clinical drug response can be predicted using baseline gene expression levels and *in vitro* drug sensitivity in cell lines. *Genome Biol* (2014) 15:R47. doi: 10.1186/gb-2014-15-3-r47

46. 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

47. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* (2019) 69:7–34. doi: 10.3322/caac.21551

48. Dela Cruz CS, Tanoue LT, Matthay RA. Lung cancer: epidemiology, etiology, and prevention. *Clin Chest Med* (2011) 32:605–44. doi: 10.1016/j.ccm.2011.09.001

49. Villalobos P, Wistuba II. Lung cancer biomarkers. *Hematol Oncol Clin North Am* (2017) 31:13–29. doi: 10.1016/j.hoc.2016.08.006

50. Welch HG. Cancer screening, overdiagnosis, and regulatory capture. *JAMA Intern Med* (2017) 177:915–6. doi: 10.1001/jamainternmed.2017.1198

51. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet* (2011) 12:56–68. doi: 10.1038/nrg2918

52. Chen X, Yuan Q, Liu J, Xia S, Shi X, Su Y, et al. Comprehensive characterization of extracellular matrix-related genes in PAAD identified a novel prognostic panel related to clinical outcomes and immune microenvironment: A silico analysis with *in vivo* and vitro validation. *Front Immunol* (2022) 13:985911. doi: 10.3389/fimmu.2022.985911

53. Yuan Q, Ren J, Wang Z, Ji L, Deng D, Shang D, et al. Identification of the real hub gene and construction of a novel prognostic signature for pancreatic adenocarcinoma based on the weighted gene Co-expression network analysis and least absolute shrinkage and selection operator algorithms. *Front Genet* (2021) 12:692953. doi: 10.3389/fgene.2021.692953

54. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. *Cell* (2012) 150:12–27. doi: 10.1016/j.cell.2012.06.013

55. Ren J, Wang A, Liu J, Yuan Q. Identification and validation of a novel redox-related lncRNA prognostic signature in lung adenocarcinoma. *Bioengineered* (2021) 12:4331–48. doi: 10.1080/21655979.2021.1951522

56. Peter SH, Michael SL, Douglas V, Rui J, Kristian C, Andrey S, et al. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* (2012) 489:519–25. doi: 10.1038/nature11404

57. Eric AC, Joshua DC, Angela NB, Alice HB, William L, Juliann C, et alComprehensive molecular profiling of lung adenocarcinoma. *Nature* (2014) 511:543–50. doi: 10.1038/nature13385

58. Li X, Wang F, Jia H, Lian Z, Ren K, Yuan Z, et al. Efficacy and safety of EGFR inhibitors and radiotherapy in locally advanced non-small-cell lung cancer: a meta-analysis. *Future Oncol* (2022). doi: 10.2217/fon-2022-0491

59. Zhang YL, Yuan JQ, Wang KF, Fu XH, Han XR, Threapleton D, et al. The prevalence of EGFR mutation in patients with non-small cell lung cancer: a systematic review and meta-analysis. *Oncotarget* (2016) 7:78985–93. doi: 10.18632/oncotarget.12587

60. Huang L, Jiang S, Shi Y. Tyrosine kinase inhibitors for solid tumors in the past 20 years (2001-2020). *J Hematol Oncol* (2020) 13:143. doi: 10.1186/s13045-020-00977-0

61. Shaw AT, Bauer TM, de Marinis F, Felip E, Goto Y, Liu G, et al. First-line lorlatinib or crizotinib in advanced ALK-positive lung cancer. *N Engl J Med* (2020) 383:2018–29. doi: 10.1056/NEJMoa2027187

62. Camidge DR, Kim HR, Ahn MJ, Yang JC, Han JY, Lee JS, 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

63. Lu X, Jiang L, Zhang L, Zhu Y, Hu W, Wang J, et al. Immune signature-based subtypes of cervical squamous cell carcinoma tightly associated with human papillomavirus type 16 expression, molecular features, and clinical outcome. *Neoplasia* (2019) 21:591–601. doi: 10.1016/j.neo.2019.04.003

64. Shoemaker RH. The NCI60 human tumour cell line anticancer drug screen. *Nat Rev Cancer* (2006) 6:813–23. doi: 10.1038/nrc1951

65. Murciano-Goroff YR, Warner AB, Wolchok JD. The future of cancer immunotherapy: microenvironment-targeting combinations. *Cell Res* (2020) 30:507–19. doi: 10.1038/s41422-020-0337-2

66. Huang TX, Fu L. The immune landscape of esophageal cancer. *Cancer Commun (Lond)* (2019) 39:79. doi: 10.1186/s40880-019-0427-z

67. Schalper KA, Brown J, Carvajal-Hausdorf D, McLaughlin J, Velcheti V, Syrigos KN, et al. Objective measurement and clinical significance of TILs in non-small cell lung cancer. *J Natl Cancer Inst* (2015) 107. doi: 10.1093/jnci/dju435

68. Vassilakopoulou M, Avgeris M, Velcheti V, Kotoula V, Rampias T, Chatzopoulos K, et al. Evaluation of PD-L1 expression and associated tumor-infiltrating lymphocytes in laryngeal squamous cell carcinoma. *Clin Cancer Res* (2016) 22:704–13. doi: 10.1158/1078-0432.Ccr-15-1543

69. 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

70. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest* (2017) 127:2930–40. doi: 10.1172/jci91190

71. Fumet JD, Richard C, Ledys F, Klopfenstein Q, Joubert P, Routy B, et al. Prognostic and predictive role of CD8 and PD-L1 determination in lung tumor tissue of patients under anti-PD-1 therapy. *Br J Cancer* (2018) 119:950–60. doi: 10.1038/s41416-018-0220-9

72. Gettinger SN, Choi J, Mani N, Sanmamed MF, Datar I, Sowell R, et al. A dormant TIL phenotype defines non-small cell lung carcinomas sensitive to immune checkpoint blockers. *Nat Commun* (2018) 9:3196. doi: 10.1038/s41467-018-05032-8

73. Chen YP, Wang YQ, Lv JW, Li YQ, Chua MLK, Le QT, et al. Identification and validation of novel microenvironment-based immune molecular subgroups of head and neck squamous cell carcinoma: implications for immunotherapy. *Ann Oncol* (2019) 30:68–75. doi: 10.1093/annonc/mdy47

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 02 March 2023
DOI 10.3389/fimmu.2023.1118816

OPEN ACCESS

EDITED BY
Ping Zheng,
The University of Melbourne, Australia

REVIEWED BY
Juan Lu,
Zhejiang University, China
Zhu Hongwen,
Shanghai Institute of Materia Medica (CAS), China

*CORRESPONDENCE
Zhixian Guo
zx_g2764@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Immunology

RECEIVED 08 December 2022
ACCEPTED 17 February 2023
PUBLISHED 02 March 2023

CITATION
Xu L, Gao X, Xing J and Guo Z (2023) Identification of a necroptosis-related gene signature as a novel prognostic biomarker of cholangiocarcinoma.
Front. Immunol. 14:1118816.
doi: 10.3389/fimmu.2023.1118816

Identification of a necroptosis-related gene signature as a novel prognostic biomarker of cholangiocarcinoma

Lixia Xu[1†], Xueping Gao[2†], Jiyuan Xing[1] and Zhixian Guo[1*]

[1]Department of Infectious Diseases, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China, [2]Department of Clinical Laboratory Medicine, Southwest Hospital, Third Military Medical University (Army Medical University), Gaotanyan, Chongqing, China

Background: Cholangiocarcinoma (CHOL) is the most prevalent type of malignancy and the second most common form of primary liver cancer, resulting in high rates of morbidity and mortality. Necroptosis is a type of regulated cell death that appears to be involved in the regulation of several aspects of cancer biology, including tumorigenesis, metastasis, and cancer immunity. This study aimed to construct a necroptosis-related gene (NRG) signature to investigate the prognosis of CHOL patients using an integrated bioinformatics analysis.

Methods: CHOL patient data were acquired from the Gene Expression Omnibus (GEO) (GSE89748, GSE107943) and The Cancer Genome Atlas (TCGA) databases, with NRGs data from the necroptosis pathway in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Univariate and multivariate regression analyses were performed to establish the NRG signatures. Kaplan–Meier (KM) curves were used to evaluate the prognosis of patients with CHOL. Functional enrichment analysis was performed to identify key NRG-associated biological signaling pathways. We also applied integrative multi-omics analysis to the high- and low-risk score groups. Spearman's rank correlation was used to clarify the relationship between the NRG signature and immune infiltration.

Results: 65 differentially expressed (DE) NRGs were screened, five of which were selected to establish the prognostic signature of NRG_S based on multivariate Cox regression analysis. We observed that low-risk patients survived significantly longer than high-risk patients. We found that patients with high-risk scores experienced higher immune cell infiltration, drug resistance, and more somatic mutations than patients with low-risk scores. We further found that sensitivities to GW843682X, mitomycin C, rapamycin, and S-trityl-L-cysteine were significantly higher in the low-risk group than in the high-risk group. Finally, we validated the expression of five NRGs in CHOL tissues using the TCGA database, HPA database and our clinical data.

Conclusion: These findings demonstrate that the five-NRG prognostic signature for CHOL patients is reasonably accurate and valid, and it may prove to be of considerable value for the treatment and prognosis of CHOL patients in the future.

KEYWORDS

necroptosis, cholangiocarcinoma, prognostic signature, immune microenvironment, biomarker

1 Introduction

Cholangiocarcinoma (CHOL) is a highly heterogeneous malignancy stemming from biliary epithelia. CHOL is the most prevalent type of malignancy and the second most common form of primary liver cancer, accounting for approximately 20% of all primary liver cancers (1, 2). Surgical treatment, immunotherapy, chemotherapy, and other comprehensive tumor treatment methods have changed the prognosis of many patients with CHOL. Patients with CHOL nonetheless still tend to have unfavorable prognoses, with only 10% of patients surviving for five years (3). The main factors contributing to poor prognosis are the heterogeneity, infiltrative nature, and rapid drug resistance of CHOL, making it difficult to completely remove the tumor by surgical procedures and identify the therapeutic target of CHOL (1, 4, 5). There is, therefore, a pressing need to further explore the occurrence and progression of CHOL to improve the treatment and survival rates of CHOL patients.

Necroptosis is a self-destruction cellular process that is regulated *via* a complex signaling cascade (6), and it is closely related to key aspects of cancer biology regulation, including tumorigenesis, metastasis, and cancer immunity (7, 8). There is increasing evidence that overcoming apoptosis resistance by induction of cancer cell necroptosis may be an attractive therapeutic approach for patients with CHOL (9–11). For instance, the application of both TNFα and gemcitabine has been shown to induce RIPK1/RIPK3/MLKL-dependent necrosis when apoptosis-inhibitory proteins and caspases are blocked, as evidenced by increased expression of RIPK3 and MLKL in CHOL cell lines (9, 12). In addition, Xu et al. found that the alkaloid matrine can induce necroptosis in CHOL by enhancing the expression of RIP3 and the RIP3/MLKL/ROS signaling pathway, thus providing a new individualized strategy for overcoming chemoresistance in CHOL therapy based on the expression of RIP3 (12). Hence, exploring the role of necroptosis in tumorigenesis and the progression of CHOL has great potential for the diagnosis and treatment of CHOL patients. The rapid development of high-throughput sequencing and multi-omics studies has allowed a substantial body of reliable information to be obtained regarding the treatment and prognosis of patients with CHOL (13–15).

In this study, we first profiled the necroptosis-related genes in CHOL and developed a risk prediction model based on five genes to explore their functional enrichment and ability to predict outcomes. The performance of the prediction models was validated in three independent cohorts (TCGA, GSE89748, and GSE107943). Additionally, we examined the differences in drug resistance, somatic mutations, and immune infiltration between the low- and high-risk groups. In brief, our prognostic signature provides a reliable method for predicting the prognosis of patients with CHOL, and it offers clinicians a reference for early diagnosis and treatment of CHOL.

Abbreviations: CHOL, Cholangiocarcinoma; TCGA, The Cancer Genome Atlas; OS, Overall survival; ROC, Receiver operating characteristic; AUC, Area under the receiver operating characteristic; PCA, Principal component analysis; DEGs, Differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSEA, Gene set enrichment analysis; HPA, Human Protein Atlas.

2 Materials and methods

2.1 Data collection and preprocessing

TCGA biolinks was used to extract RNA-Seq data from 36 CHOL and 9 normal samples, as well as relevant clinical information from TCGA database (http://portal.gdc.cancer.gov) (16). Additionally, the University of California Santa Cruz (UCSC) provided FPKM, somatic mutation, and clinical data on CHOL. In the present study, CHOL datasets GSE89748 and GSE107943 (17, 18) from the GEO database (https://www.ncbi.nlm.nih.gov/geo) were downloaded using the GEO query R package, which was used as the external validation set, including available expression profile data and clinical information of bile duct cancer samples. In total, 72 CHOL samples from the GSE89748 dataset and 30 CHOL samples from the GSE107943 dataset were acquired. A total of 159 necroptosis-associated genes (NRGs) were obtained from the necroptosis pathway (hsa04217) in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.

2.2 Identification of the expression patterns and biological functions of DENRGs in CHOL

First, we extracted the NRG_S expression matrix from TCGA and then screened for differentially expressed necroptosis-related genes (DENRGs) between the CHOL and normal groups using the

limma package (19). Significant DENRGs were visualized using volcano plots constructed using the ggplot2 package. The criteria for differentially expressed genes (DEGs) were FDR < 0.05 and $|\log_2 FC| > 1$. Furthermore, differences in DENRGs between the CHOL and normal groups were visualized using boxplots. DENRGs were also analyzed based on a protein-protein interaction (PPI) network using the STRING database (20), and correlations between them were visualized using heatmaps. To investigate the biological role of DENRGs, we examined biological processes (BP), cellular components (CC), and molecular functions (MF) according to the Gene Ontology (GO) database and KEGG signaling pathways using the R tool cluster Profile (21). The enrichment significance thresholds were set at an adjusted p-value of < 0.05.

2.3 Development and validation of DENRGs-based prognostic models

DENRGs were first identified for their prognostic values in the TCGA cohort by univariate Cox proportional hazards regression analysis, and the genes with p-values < 0.05 were then entered into the multivariate Cox regression analysis. A risk score model was built based on the expression levels of the prognosis-associated genes and the contribution coefficient (and beta) of the multivariate Cox proportional hazard regression model. Based on the above risk score model, we calculated the prognostic risk value for each patient sample in TCGA (training cohort), GSE89748 (validation cohort 1), and GSE107943 (validation cohort 2). All CHOL samples were divided into high- and low-risk groups, with the median risk score as the cutoff value. Kaplan–Meier survival analyses were performed using the 'survival' and 'survminer' (22) packages between the high- and low-risk groups. To further assess the clinical diagnostic value of the risk score, time-dependent receiver operating characteristic (ROC) curves for overall survival (OS) and area under the ROC curves (AUCs) at 1, 3, and 5 years in TCGA (training cohort), GSE89748 (validation cohort 1), and GSE107943 (validation cohort 2) were generated using the R package "survivalROC" (23). OS is defined as the time from randomization to death. Furthermore, we constructed a risk plot to explore the relationship between the risk score and the prognosis status.

2.4 Process of the screening signature for the Cox regression model and building of the nomogram models

Univariate Cox regression was performed to examine the relationship between patient clinical characteristics (age, sex, stage, pathology, weight, height, and BMI), risk score, and OS. Significant prognostic factors (p < 0.05) in the univariate analyses were selected for multivariate Cox regression analysis. Forest plots were used to present the results of the univariate and multivariate Cox analyses, including all of the above variables. A nomogram was built based on the identified variables in the multivariate Cox regression analysis to facilitate clinical application.

2.5 Exploration of differences in biological functions between CHOL subgroups

To determine the differences in biological functions between the high- and low-risk groups, DEGs between the two groups were screened using the limma R package with FDR thresholds of < 0.05, and absolute $\log_2 FC > 1$. A volcano plot was then used to illustrate the DEGs using ggplot2. To visualize the expression patterns of DEGs between the low- and high-risk groups, we used R package (pheatmap) to generate a heatmap. All DEGs were subjected to GO and KEGG pathway enrichment analyses using Metascape (http://metascape.org) (24). A p-value < 0.01 and a minimum of three counts were set as the cutoff criteria for selecting significant enrichment results. GO and KEGG analyses were also performed using the R package "cluster Profiler" to explore the underlying biological roles of the DEGs (21). The enrichment results were visualized using bar and dot plots. Gene set enrichment analysis (GSEA) (25) was performed using cluster Profiler, with a p-value of < 0.05 as the threshold for significantly enriched KEGG pathways. The top 20 significantly enriched pathways ranked by normalized enrichment scores were visualized using a ridgeline plot.

2.6 Applying integrative multi-omics analysis between the high- and low-risk score groups

The R package "Rcircos" (26) was used to map the chromosomal locations of clinically significant NRGs. The Friends tool was then used to functionally annotate these genes, which were subsequently estimated by semantic analysis using the R package GOSemSim (27). By building a ridgeline regression model based on the Genomics of Drug Sensitivity in Cancer (GDSC) database (www.cancerrxgene.org/), we predicted the half-maximal inhibitory concentration (IC_{50}) for chemotherapy drugs in the high- and low-risk groups and we inferred the sensitivity of the patients (28). To detect somatic mutations in CHOL patients between the high-risk and low-risk subgroups, we used the mutation annotation format (MAF) in TCGA database. The results were visualized using a waterfall plot (oncoplot). Using the online tool Network Analyst (29), we explored the transcriptional regulators and chemical targets of hub necroptosis genes based on the JASPAR Tarbase and mir-Tarbase databases.

2.7 Correlation analysis between the prognostic DENRGs and immune cell infiltration

Immune infiltration is a significant factor in tumor progression, treatment, and prognosis. We used the "ESTIMATE" R package to estimate the stromal score, immune score, and tumor purity in the high- and low-risk subgroups (30). The R package "ggplot2" was then applied to generate boxplots to visualize differences between the two groups for the above-mentioned immune scores and tumor

purity. CIBERSORT is a deconvolution algorithm that can calculate the infiltration abundance of 22 immune cell types in all tumor samples (31). Heatmaps were drawn using the R package pheatmap to illustrate the fractions of immune cell types for each sample, and a correlation analysis between 22 immune cell types and prognostic necroptosis genes was performed using the corrplot package. The results were visualized using the 'pheatmap' package. Immune infiltration differences between the high- and low-risk groups of CHOL patients were determined using the ggplot2 package. Additionally, the most positively and negatively correlated gene-immune cell pairs were displayed using a scatter plot.

2.8 Immunohistochemical analysis of five NRGs in HPA

The protein expression of the five NRGs between CHOL and normal tissues was measured by immunohistochemistry from the Human Protein Atlas (HPA) (https://www.proteinatlas.org/), which is a valuable database providing the data of immunohistochemistry expression for specific human tissues and cells (32).

2.9 Tumor samples collection and qRT-PCR

A total of 12 CHOL tissue samples and 10 corresponding normal hepatobiliary duct tissues were obtained from patients who underwent surgical resection between March 2021 and October 2022 at the First Affiliated Hospital of Zhengzhou University, Henan, China. The samples were immediately frozen in liquid nitrogen after tissue resection. The total RNA of the tissue samples was extracted using TRIzol reagent (Invitrogen) according to the manufacturer's protocol. The RNA samples were reverse-transcribed into cDNA by using iScriptTM cDNA Synthesis Kit. RT-qPCR was performed using a thermal cycler (Roche LightCycler 480) using IQTM SYBR® Green Supermixes for Real-Time PCR. The mRNA expression was normalized to the expression of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA and counted by the $2-\Delta\Delta Ct$ method. The PCR primer sequences are shown in Table 1. This study conforms to the guidelines issued in the Declaration of Helsinki and was approved by the Ethics Committee of the First Affiliated Hospital of Zhengzhou University (Approval Number: SS-2019-018).

2.10 Statistical analysis

All data processing and statistical analyses were performed using R software (version 4.2.1). A detailed description of the bioinformatics analyses is provided in the corresponding subsections. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. *A* p-value < 0.05 was taken as representing statistical significance.

3 Results

3.1 Identification of DENRGs

According to the filter criteria, a total of 67 DENRGs were screened, including 64 upregulated genes and 3 downregulated genes. The expression distribution of the DENRGs was visualized using volcano plots (Figure 1A). Based on the boxplot and heatmap, it was clear that *H2AW*, *PYGB*, *PYCARD*, *CAPN2*, *BIRC3*, *H2AX*, *CHMP4C*, *STAT1*, *CHMP3*, *CHMP4B*, *CAPN1*, *H2AZ1*, and *BAX* were highly expressed in the CHOL group, whereas *FTL*, *GLUD1*, and *PYGL* were expressed at very low levels compared with the normal group (Figure S1; Figure 1B). Principal component analysis (PCA) of these DENRGs clearly distinguished the CHOL group from the control group (Figure 1C). Mutation analysis indicated that missense mutations were the most common, and *TYK2* had the highest mutation rate, which was a missense mutation with a frameshift deletion (Figure 1D). The heat map showed that *FTL*, *GLUD1*, and *PYGL* were positively correlated with each other and negatively correlated with the other DENRGs (Figure 1E). Furthermore, the PPI network diagram suggested that *CASP8*, *MLKL*, and *RIPK3* exhibited the strongest interactions with the other DENRGs (Figure 1F).

3.2 GO and KEGG functional analysis of the DENRGs

The results show that the DENGs were mainly related to cell death processes, such as programmed necrotic cell death, midbody abscission, necrotic cell death, mitotic cytokinetic process, necrotic process and virtual budding, and ESCRT complex,

TABLE 1 Primer list of PCR.

Gene Name	Forward primer	Reverse primer
GAPDH	GGAGCGAGATCCCTCCAAAAT	GGCTGTTGTCATACTTCTCATGG
PYGB	AGGTGCGGAAGAGCTTCAAC	TCGCGCTCGTAGTAGTGCT
IFNGR2	CTCCTCAGCACCCGAAGATTC	GCCGTGAACCATTTACTGTCG
TICAM1	GCCAGCAACTTGGAAATCAGC	GGGGTCGTCACAGAGCTTG
STAT6	GTTCCGCCACTTGCCAATG	TGGATCTCCCCTACTCGGTG
VPS4B	ATGTCATCCACTTCGCCCAAC	TTGCTTGGCTTTATCACCCTG

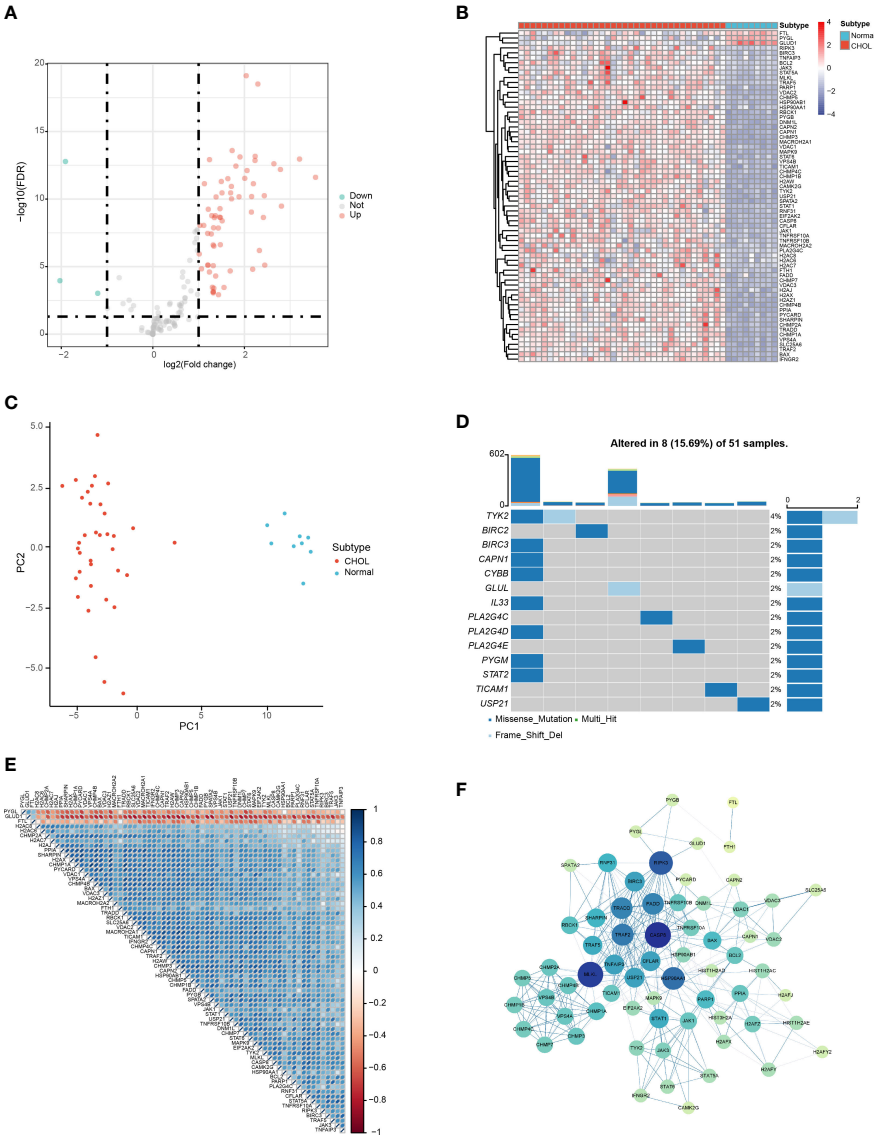


FIGURE 1
Identification of DENRGs in the CHOL group. **(A)** Volcano plot of the DENRGs. Genes indicated in red, blue, and gray colors were significantly upregulated (Up), downregulated (Down), or not significantly different (Not), respectively. **(B)** Heatmap showing the expression of 65 DENRGs in the normal and CHOL samples. Red, CHOL group; Blue, normal group **(C)** Principal components analysis (PCA) indicating the expression patterns of DENRGs. **(D)** Oncoplot of the DENRG mutations. **(E)** Heat map of the correlation between the DENRGs. Red colors indicate positive correlations and blue colors represent negative correlations. The darker the color, the stronger the correlation. **(F)** PPI network of the DENRGs. The larger the node, the higher the number of interactions with other genes, and the thicker the line, the higher the correlation coefficient.

nucleosome, DNA packaging complex, protein DNA complex, nuclear chromatin, tumor necrotic factor receiver superfamily binding, tumor necrotic factor receiver binding, cytokine receiver binding, ubiquitin-like protein ligase binding, and protein binding (Figure 2A; Table S1). The KEGG results suggest that the DENRGs were mainly involved in multiple functional pathways (e.g., Necroptosis, NOD-like receptor signaling pathway, Apoptosis, Influenza A, TNF signaling pathway, Th17 cell differentiation, IL-17 signaling pathway, and Neutrophil extracellular trap formation pathway) (Figure 2B; Table S1). A panoramic view of the necroptosis pathway in KEGG was generated (Figure 2C).

3.3 Construction of a prognostic model within necroptosis-associated genes

The 67 DENRGs were subjected to univariate Cox proportional hazard regression analysis. Five prognostic genes (*PYGB*, *IFNGR2*, *TICAM1*, *STAT6*, and *VPS4B*) were selected and further analyzed using multivariate Cox proportional hazards regression analysis. The coefficients from the multivariate Cox proportional hazards regression model were used to evaluate the potential prognostic factors. Risk scores were also calculated in TCGA (training cohort), GSE89748 (validation cohort 1), and GSE107943 (validation cohort 2) according to the prognostic gene expression values and their

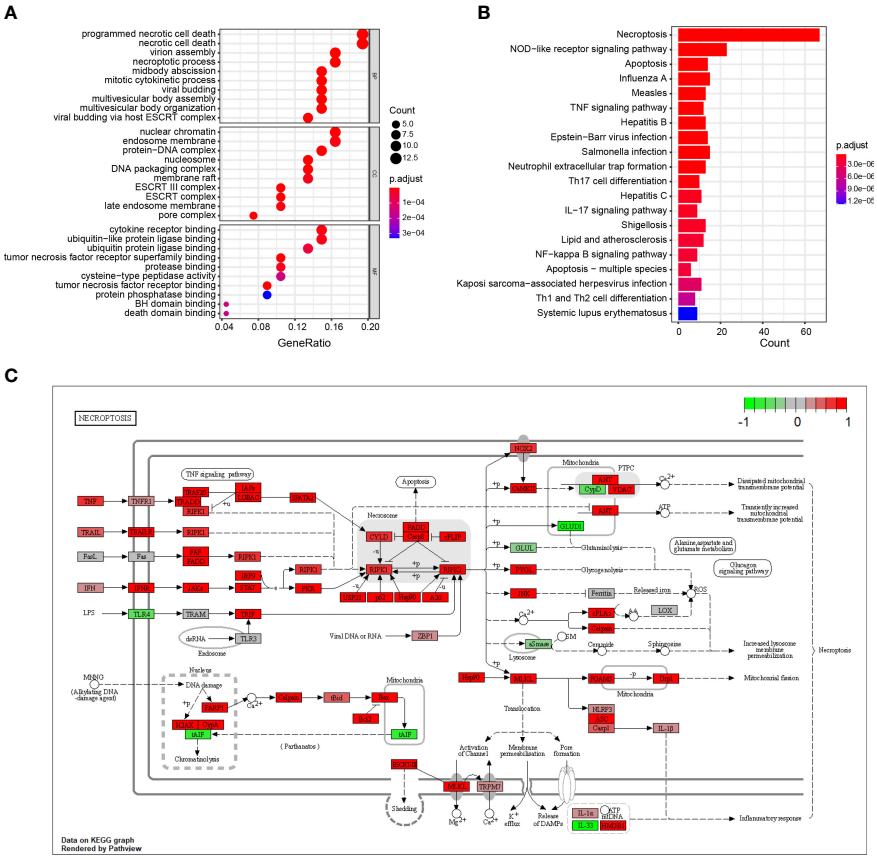


FIGURE 2
GO and KEGG enrichment analysis of DENRGs. **(A)** Dot plot showing the top 10 biological functions enriched in Gene Ontology (GO) terms. **(B)** Bar plot showing the top 10 signaling pathways enriched in KEGG terms. **(C)** Diagrammatic outline of the necroptosis pathway.

regression coefficients. Taking the median risk score of the samples as the cutoff value, CHOL patients were divided into high- and low-risk groups. Survival analysis showed that the low-risk group exhibited a better outcome in TCGA (log-rank test p-value < 0.05) (Figure 3A), GSE89748 (log-rank test p-value < 0.001) (Figure 3B), and GSE107943 (log-rank test p-value < 0.001) (Figure 3C). Next, we performed 1-, 3-, and five-year time-dependent ROC analyses in three independent datasets (TCGA, GSE89748, and GSE107943). The results show that the AUC of time-dependent ROC curves was greater than 0.6 in all datasets (Figures 3D–F). Notably, the AUC of the 1-year time-dependent ROC exceeded 0.7, indicating that the prognostic risk score had good prediction abilities. A risk plot also illustrated the distributions of the risk scores and the OS status in the three dependent datasets (Figures 3G–I). It is worth mentioning that the increase in the prognostic risk score and the number of death events in patients increased.

3.4 Construction and evaluation of the nomogram model

Univariate and multivariate Cox regression analyses were performed on the clinical characteristics and risk scores in TCGA to explore the prognostic factors of patients. The results show that two factors, the risk score and pathologic N, were significantly associated with patient prognosis ($p < 0.05$) (Figures 4A, B). Subsequently, a nomogram model for predicting 1-, 3-, and 5-year OS was constructed, which integrated the two factors that were significantly correlated with prognosis: pathologic N and the prognostic risk score (Figure 4C). Besides, we established calibration curves to verify the effectiveness of nomogram model for predicting the rates of OS for CHOL patients at 1, 3, and 5 years. The results showed that the calibration curves displayed a suitable agreement between the prediction by nomogram and actual survival (Figure S2).

A risk classification system was then constructed based on the risk scores calculated from the nomogram model for each CHOL patient. Using this system, the enrolled patients were divided into low- and high-risk groups. The outcomes show that the low-risk group had the best prognosis, and the high-risk group had the worst prognosis (Figures 4D, E). Time-dependent ROC analysis showed that the 1-, 3-, and 5-year nomogram models exhibited AUC > 0.7, and even the 1- and 3-year time-dependent ROC exhibited AUC > 0.8 (Figures 4F, G). We further used decision curve analysis (DCA) to evaluate the clinical predictive models. The results showed that the DCA curves at 1, 3, and 5 years remained above the gray and black lines between 0 and 1.0, in TCGA CHOL and GSE89748

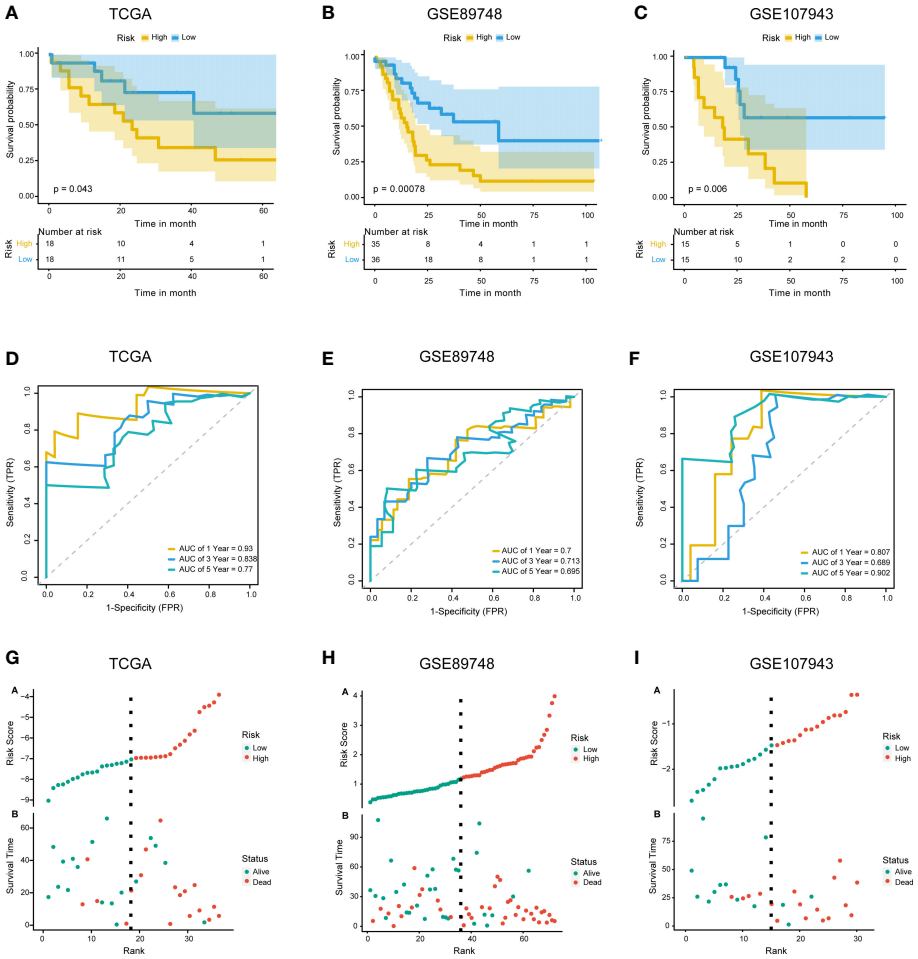


FIGURE 3
Construction and validation of the prognostic model. **(A–C)**. KM survival curves for overall survival in TCGA training cohort **(A)**, GSE89748 validation cohort **(B)**, and GSE107943 validation cohort **(C)**. **(D–F)** Time-dependent ROC curve of TCGA cohort **(D)**, GSE89748 cohort **(E)**, and GSE107943 cohort **(F)**. Sensitivity (TRP) = TP/(TP+FN) and false positive prediction rate (FPR) (1-specificity = FP/(FP+TN)) were used as the y-axis and x-axis variables, where TPs (true positives) are positive predictions which belong to gold standard positives (GSPs), FNs (false negatives) are negative predictions which belong to GSPs.TP, true positive; FP, false positive; TN, true negative. **(G–I)** Distributions of risk scores and OS status are shown for TCGA cohort **(G)**, GSE89748 cohort **(H)**, and GSE107943 cohort **(I)**.

datasets, suggesting that CHOL patients may benefit from decisions based on the prognostic model (Figures 4H, I).

3.5 Identification of DEGs and functional enrichment analysis

Next, we performed differential expression analysis on TCGA CHOL datasets of the high- and low-risk groups to obtain DEGs. According to the screening thresholds ($|\log2FC| > 0.5$ and $p < 0.05$), 179 DEGs were identified in the high- and low-risk groups, including 96 upregulated genes and 83 downregulated genes (Figure 5A). In addition, the heatmaps revealed that the expression patterns of genes were also classified into two categories, along with the high- and low-risk groups (Figure 5B).

GO and KEGG functional enrichment analyses of the DEGs were performed using Metascape. The top 20 enriched biological function terms were displayed in the network diagrams according to their enrichment scores (Figures 5C, D). The GO analysis results show that the DEGs were mainly associated with mitotic cell cycle, mitotic spindle organization, mitotic spindle assembly, intercellular bridge, polymeric cytoskeletal fiber, hexosyltransferase activity, DNA, Binding transcription activator activity, and protein kinase binding (Figure 5E). According to the KEGG analysis results, pathways in cancer, viral carcinogenesis, TNF signaling pathway, *Salmonella* infection, pathogenic *Escherichia coli* infection, IL−17 signaling pathway, hepatitis B, chemical carcinogenesis-receptor activation, and apoptosis were significantly enriched (Figure 5F). The detailed results are summarized in Table S2.

To further analyze the functional implications of the five necroptosis gene signatures in CHOL, we performed GSEA of TCGA CHOL expression profiles according to low- and high-risk groups. As shown in Figure 6A, the ridgeline plot reveals the top 20 enriched KEGG terms in the low- and high-risk groups. These results show that cytokine-cytokine receptor interaction, alcoholism, neutrophilic extracellular trap formation, influenza A,

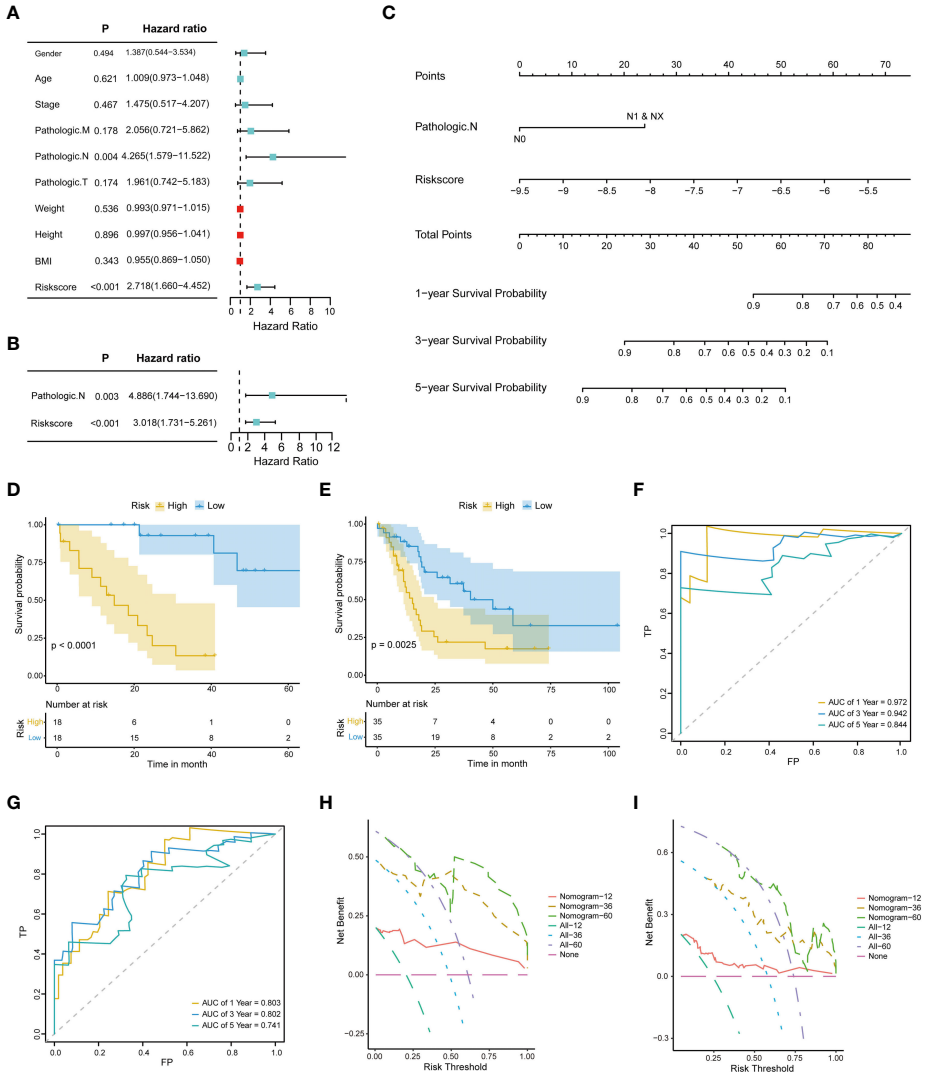


FIGURE 4
Construction and evaluation of the nomogram model. **(A)** Univariate Cox proportional hazard regression analysis of the clinical characteristics. **(B)** Multivariate Cox proportional hazard regression analysis of selected clinical characteristics. **(C)** Prediction of 1-, 3-, and 5-year survival probabilities for CHOL patients using the nomogram model. **(D, E)**. Survival curve for the low-risk and high-risk subgroups in the training dataset and the validation dataset. **(F, G)**. Time-dependent ROC curves of the training cohort and the validation cohort. Sensitivity (TRP) = TP/(TP+FN) and false positive prediction rate (FPR) (1-specificity = FP/(FP+TN)) were used as the y-axis and x-axis variables, where TPs (true positives) are positive predictions which belong to gold standard positives (GSPs), FNs (false negatives) are negative predictions which belong to GSPs.TP, true positive; FP, false positive; TN, true negative **(H, I)**. DCA curves of the training cohort and the validation cohort.

JAK-STAT signaling pathway, and cell adhesion molecules were significantly enriched in the low-risk group (Figures 6B–G). Detailed GSEA results are presented in Table S3.

3.6 Multi-omics analysis based on prognostic risk scores

We then used the R package "Rcircos" to map the chromosomal locations of the above five NRGs. The gene chromosome location diagram revealed that *PYGB, IFNGR2, TICAM1, STAT6*, and *VPS4B* are located on chr20, chr21, chr19, chr12, and chr18, respectively (Figure 7A). Friends analyses of the necroptosis-associated prognostic genes revealed that *TICAM1* was the most important term (Figure 7B). In the low-risk group, the ESTIMATE, immune, and stromal scores were all higher than those in the high-risk group, according to violin plots (Figure 7C). The therapeutic effects of the four drugs on CHOL are shown as boxplots. The results show that the sensitivity to GW843682X, mitomycin C (MMC), rapamycin, and S-trityl-L-cysteine (STLC) was significantly higher in the low-risk group than in the high-risk group (Figure 7D). The oncoplot demonstrated different mutation patterns between the high- and low-risk groups (Figures 7E, F).

We further used Network-Analyst to obtain network diagrams of the interaction between the five NRGs and miRNAs, transcription factors (TFs), and potential chemicals. The results show that 124 miRNAs targeting the five necroptosis prognosis genes fit a network diagram (Figure 8A). In the TF-necroptosis

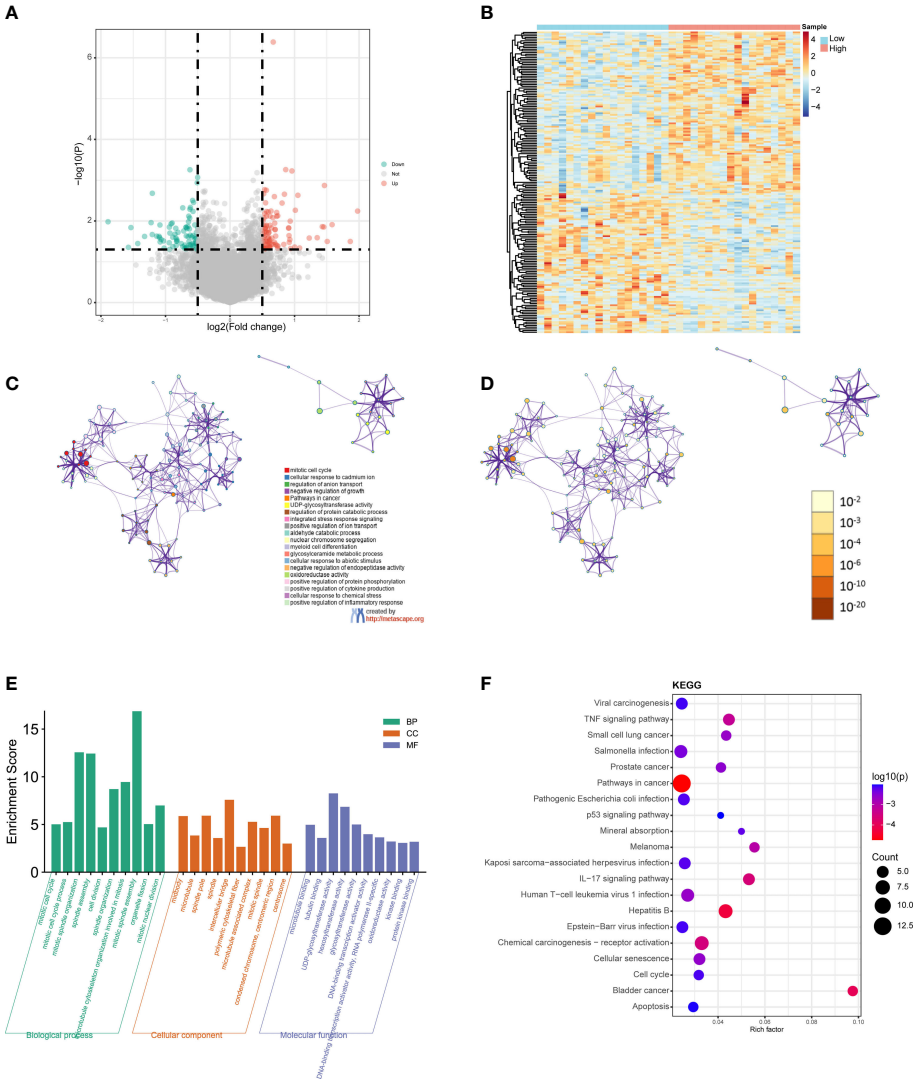


FIGURE 5
GO and KEGG functional enrichment analyses between the low- and high-risk groups. **(A)** Volcano plot of the DEGs. Red represents upregulated genes (Up), blue represents downregulated genes (Down), and gray represents not significantly different genes (Not). **(B)** Heatmap of the DEGs between the high-risk group and the low-risk group. Red indicates the high-risk group (High) and blue indicates the low-risk group (Low). **(C)** A network diagram of the top 20 enriched biological functions. Cluster IDs are represented using different colors, while enriched terms are indicated by nodes. **(D)** Twenty enriched biological functions are shown in this network diagram, and the p-values are displayed as different colors, while the enriched terms are indicated as nodes. **(E)** Bar-plot of GO terms, with the height of the column indicating the enrichment score. **(F)** Dot plot of the KEGG enrichment analyses results. The dot scale represents the number of genes in each KEGG term; the depth of the dot color represents the p-value.

prognosis gene network diagram, 114 TFs were observed (Figure 8B). A total of 117 potential chemical targets were identified using Network Analyst (Figure 8C).

3.7 Analysis of immune cell infiltration and its correlation with the five NRGs

Immune cell infiltration is a critical factor in the progression of CHOL, and it significantly affects the survival rate of patients with CHOL (9, 33). We analyzed the relationship between the expression of the five NRGs and infiltration of 22 immune cell types in CHOL. The results show that *IFNGR2* and *STAT6* were negatively correlated with resting natural killer (NK) cells, whereas *PYGB* was significantly positively correlated with $CD8^+$ T cells, M0 macrophages, Tregs, and eosinophils. *TICAM1* was positively correlated with resting central memory $CD4^+$ T cells and activated NK cells, and *VPS4B* was positively correlated with plasma cells and T follicular helper cells. *STAT6* expression positively correlated with monocytes and Tregs (Figure 9A). A heatmap of the correlation between the 22 different immune cell types indicates that M2 macrophages had a clear positive correlation with monocytes; naive B cells had a clear positive correlation with activated mast cells and naive $CD4^+$ T cells; memory B cells had a clear positive correlation with naive $CD4^+$ T cells, while activated mast cells exhibited obvious inverse

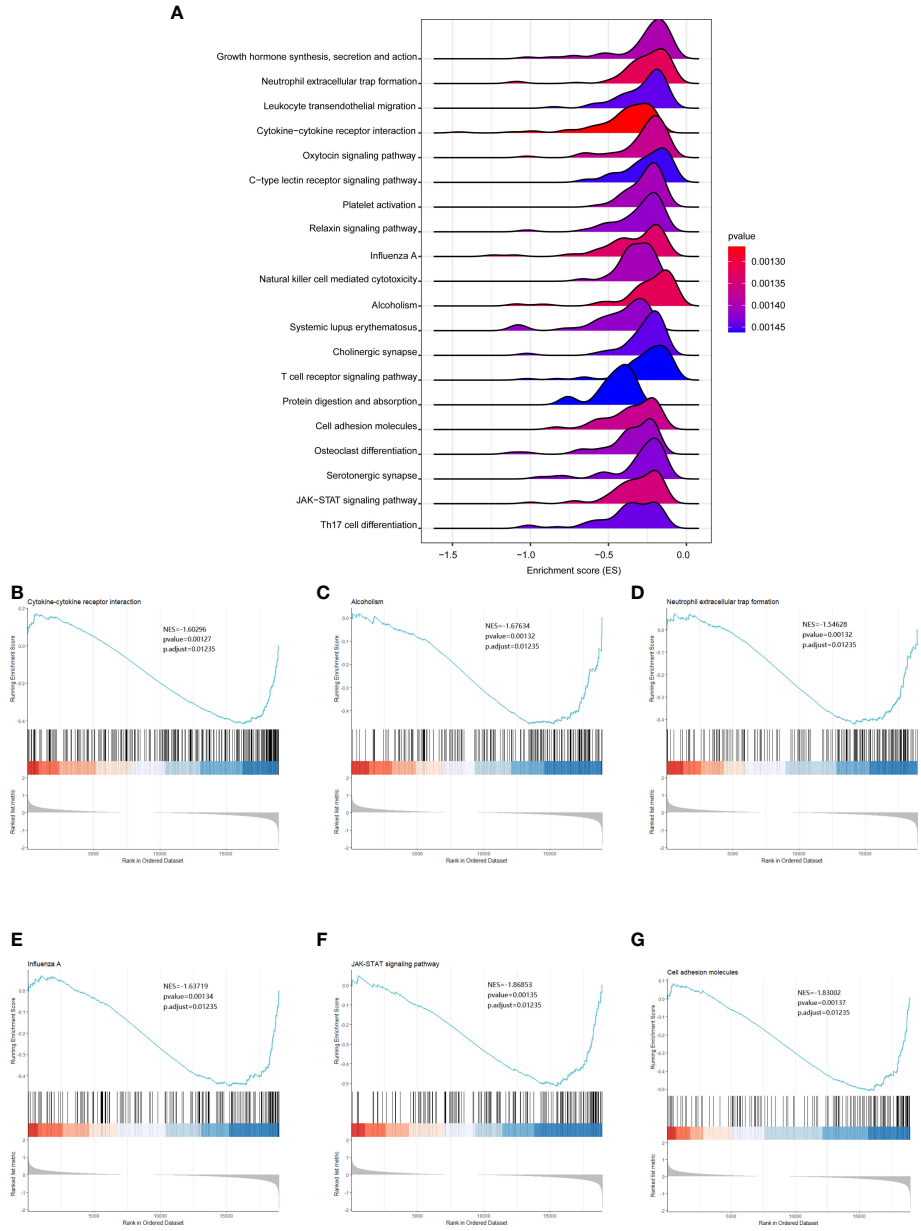


FIGURE 6
GSEA analysis results between the low- and high-risk groups. **(A)** Ridgeline plots showing the top 20 enriched KEGG terms in the low- and high-risk groups. ES (enrichment score) reflected the correlation between the gene set and the sample. B-G. Cytokine-cytokine receptor interaction **(B)**, alcoholism **(C)**, neutrophilic extracellular trap formation **(D)**, influenza A **(E)**, JAK-STAT signaling pathway **(F)**, and cell adhesion molecules **(G)** were significantly enriched in the low-risk group.

correlations with resting mast cells and M2 macrophages; activated NK cells had an obvious inverse correlation with monocytes, M2 macrophages, and neutrophils (Figure 9B). The strongest positive correlation was observed between *IFNGR2* and eosinophils (Figure 9C). In contrast, *STAT6* exhibited the strongest negative correlation with resting NK cells (Figure 9D). The high-risk and low-risk groups exhibited significantly different levels of immune cell infiltration in the heatmap (Figure 9E). The boxplot indicates that there was a significant difference in the proportion of immune cells between the high- and low-risk groups. B cells accounted for a higher proportion in the low-risk group, whereas T cells accounted for a higher proportion in the high-risk group (Figure 9F).

3.8 Validation of the five NRGs expressions in CHOL tissue samples

We further validated the expression of five NRGs using the TCGA database, HPA database and our clinical data. TCGA database results showed that *PYGB* (Figure 10A), *IFNGR2* (Figure 10D), *TICAM1* (Figure 10G), *STAT6* (Figure 10J) and *VPS4B* (Figure 10M) were expressed at high levels in CHOL tissues. Based on the protein expression data from the HPA, the immunohistochemistry results confirmed that the protein expression levels of *PYGB* (Figure 10B), *IFNGR2* (Figure 10E), *TICAM1* (Figure 10H), *STAT6* (Figure 10K) and *VPS4B*

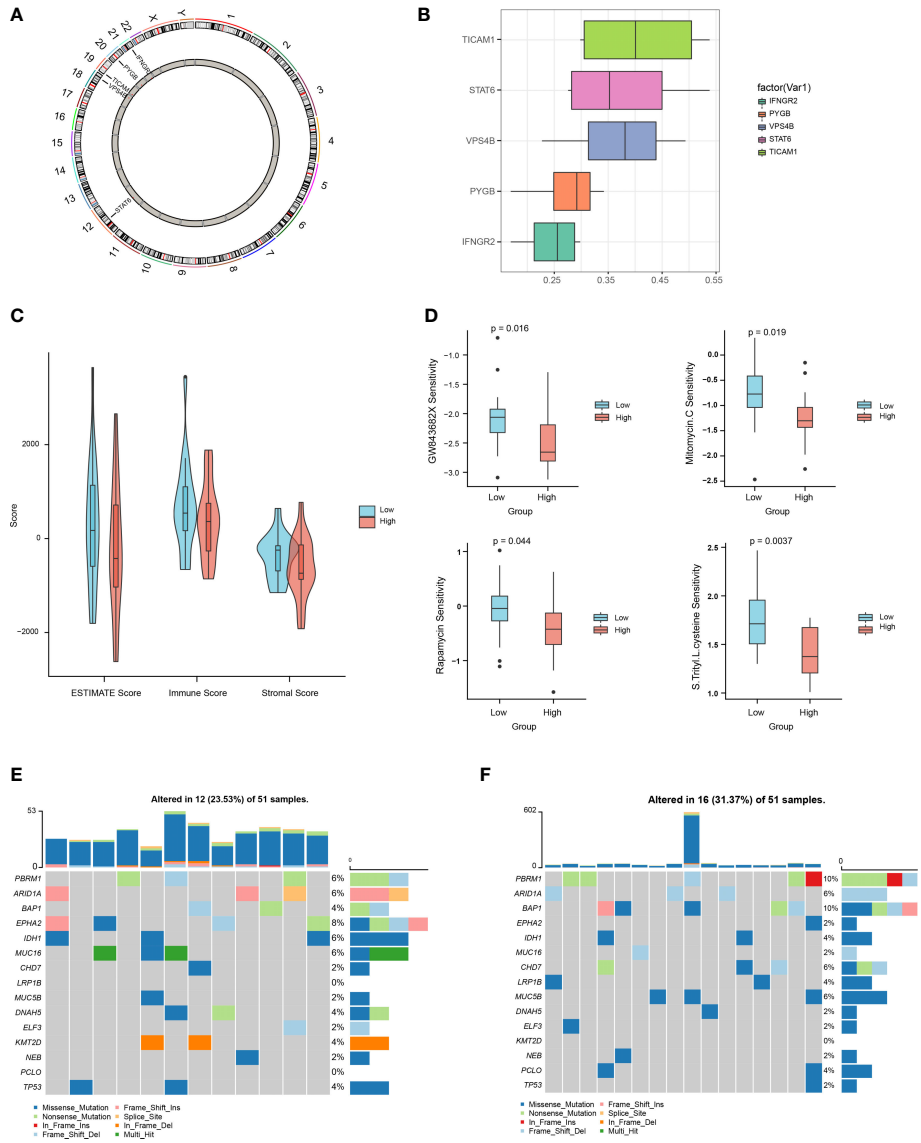


FIGURE 7
Multi-omics analysis based on the prognostic risk scores. **(A)** Chromosome localization map of necroptosis prognosis genes. **(B)** Friends analysis of necroptosis prognosis genes. **(C)** Differences in ESTIMATE, immune, and stromal scores between the high- and low-risk groups. **(D)** Differences between the high- and low-risk groups in terms of drug sensitivity. **(E, F)**. Oncoplot mutations in the low- and high-risk groups.

(Figure 10N) were higher in CHOL tissues than normal hepatobiliary duct tissues. Finally, we detected their expression levels in 10 non-tumor hepatobiliary duct tissues and 12 CHOL tissues by using RT-qPCR assay. The results showed that the expression levels of *PYGB* (Figure 10C), *IFNGR2* (Figure 10F), *TICAM1* (Figure 10I), *STAT6* (Figure 10L) and *VPS4B* (Figure 10O) in CHOL tissues showed an overall upward trend compared with non-tumor hepatobiliary duct tissues.

4 Discussion

CHOL is the second most common primary malignancy of the liver after hepatocellular carcinoma, with a steady increase in its incidence and mortality rate (1). When hepatocytes die due to necroptosis, the necroptosis-dominated microenvironment leads to the development of CHOL. Recent studies have also found that necroptosis plays a pivotal role in regulating carcinogenesis, cancer subtypes, immunity, metastasis, and anticancer treatments (2, 3). The molecular mechanism by which necroptosis is involved in the genesis and development of CHOL remains unclear, however.

In this study, we focused on developing and validating a prognostic signature for CHOL using necroptosis-related genes. First, 65 DENRGs were identified between the CHOL and control groups. Secondly, five genes (*PYGB*, *IFNGR2*, *TICAM1*, *STAT6*, and *VPS4B*) were identified as prognostic signatures based on multivariate Cox regression analysis. The Kaplan–Meier survival curves in TCGA also indicate that the low-risk group had

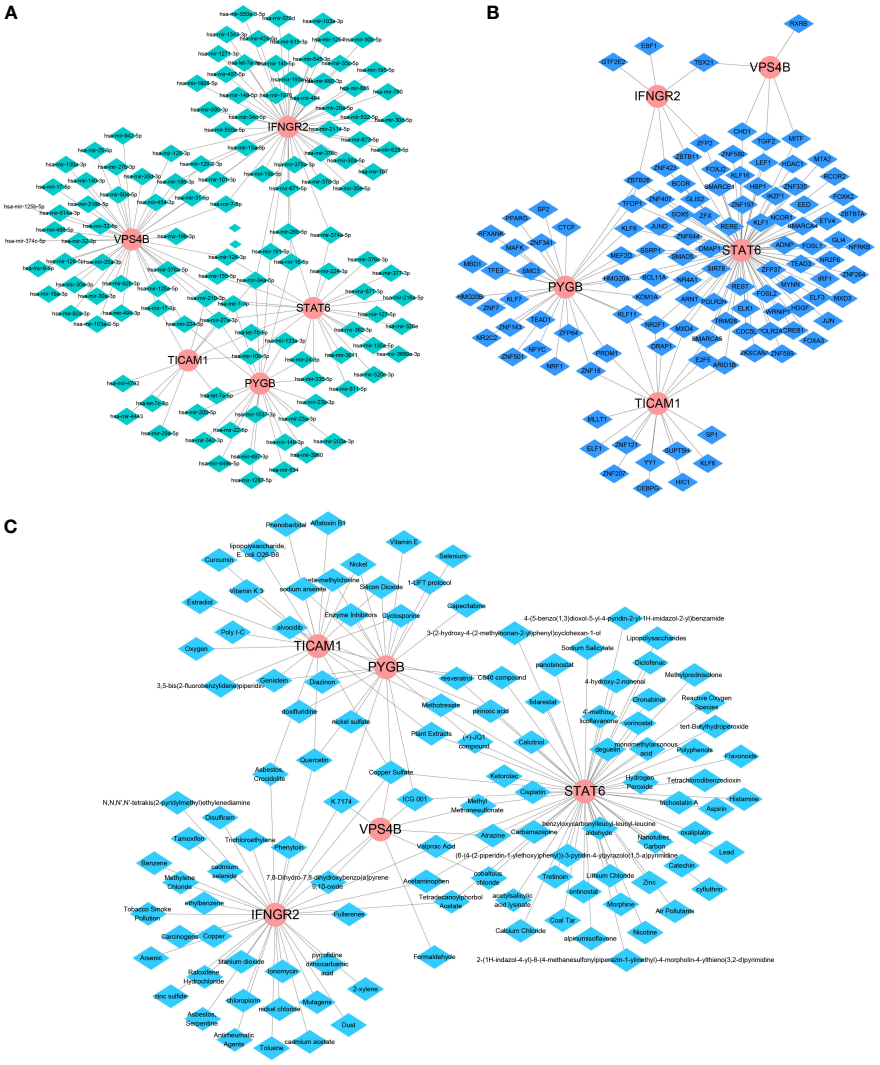


FIGURE 8
An integrated network of TFs, miRNAs, and chemicals target the necroptosis prognosis genes. **(A)** The integrated network diagrams between the five NRGs and miRNAs. **(B)** The integrated network diagrams between the five NRGs and TFs. **(C)** The integrated network diagrams between the five NRGs and potential chemical modulators.

significantly longer patient survival than the high-risk group. The survival results were also validated independently using the GSE89748 and GSE107943 datasets. In addition, the nomogram model was highly discriminatory for OS based on the pathologic N and risk score. Moreover, patients with high-risk scores experienced higher immune cell infiltration, drug resistance, and more somatic mutations. In summary, these results suggest that the five genes related to necroptosis play prominent roles in modulating drug resistance, somatic mutations, and the tumor microenvironment, indicating that these risk signatures were highly robust and accurate in predicting the prognosis of patients with CHOL.

Our prognostic signature consists of five genes, *PYGB*, *IFNGR2*, *TICAM1*, *STAT6*, and *VPS4B*, each of which plays a critical role in necroptosis and tumor progression. *PYGB* codes for the protein glycogen phosphorylase B, which is found predominantly in the brain (34). *PYGB* has been reported to be involved in the progression of gastric and liver cancers (35, 36). *IFNGR2* codes for the IFN-γ receptor, which has been found to mediate a non-immunogenic tumor phenotype associated with checkpoint inhibitor resistance in renal carcinoma (37, 38). *TICAM1* codes for an essential necrosome adaptor protein that functions as an essential signal transducer in Toll-like receptor (TLR) 3 and TLR4 signaling pathways (39). It has been reported that *TLR3/TICAM1* signaling is involved in tumor cell RIP3-dependent necroptosis, which contributes to immune effector-mediated tumor elimination (38). In our study, *TICAM1* was highly expressed in the CHOL group and was positively correlated with resting central memory $CD4^+$ T cells and NK cell activation, suggesting that the *TICAM1* gene product is involved in the tumor microenvironment. STAT6 is highly expressed in a variety of human cancers and has been suggested to induce apoptosis and growth inhibition of hepatocellular carcinoma-derived cells by lowering RANKL expression (40). *VPS4B* codes for a protein that is involved in autophagy that can reduce the sensitivity of T cell-mediated tumor

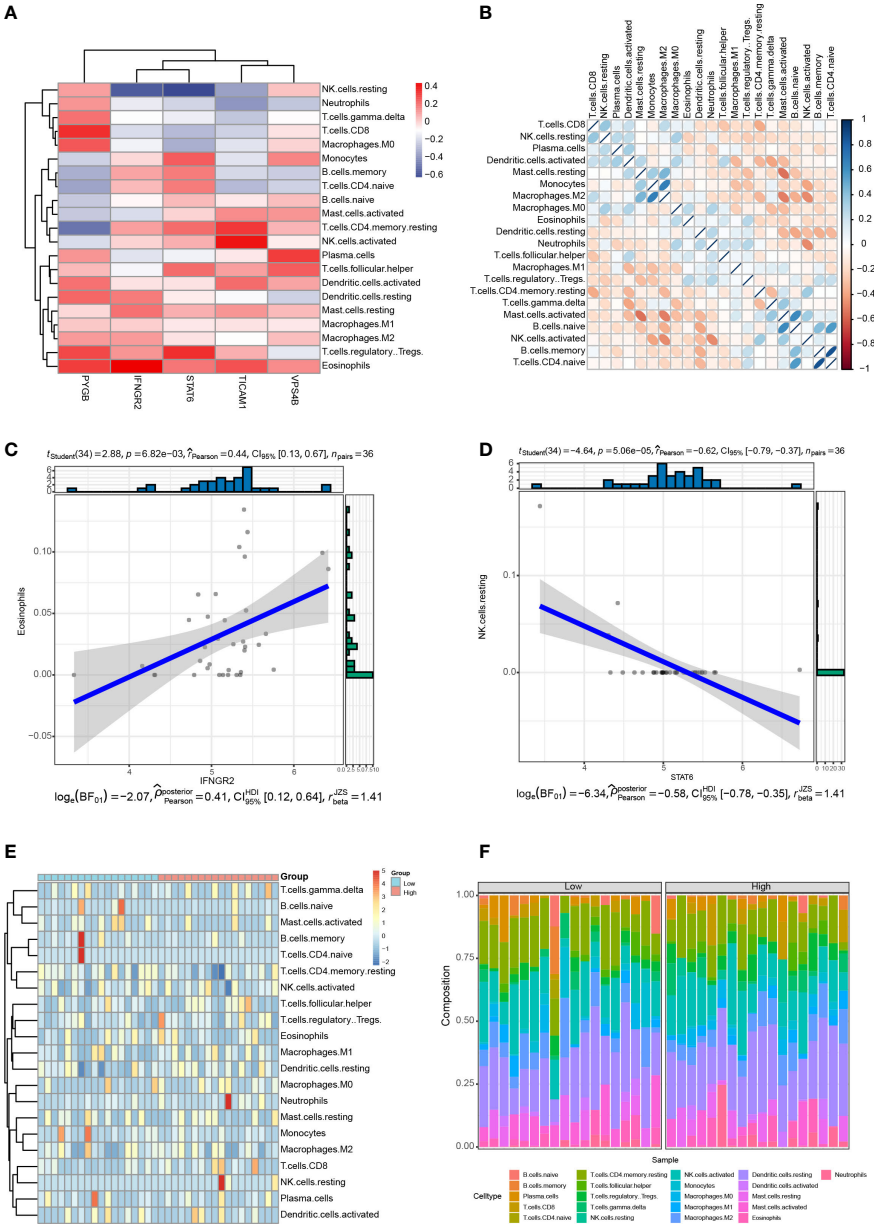


FIGURE 9
Correlation between the five NRGs and immune cell infiltration of CHOL. **(A)** Correlation analyses between 22 different immune cell types and the five NRGs in the CHOL group. Red color represents positive correlation whereas blue color indicates negative correlation. **(B)** Heatmap of the correlation between 22 different immune cell types. Positive correlations are in red and negative correlations are in blue. The darker the color, the stronger the correlation. **(C)** Correlation analysis between *IFNGR2* and Eosinophils. **(D)** Correlation analysis between *STAT6* and resting NK cells. **(E)** A heatmap showing the difference in immune cell infiltration between the high-risk and low-risk groups. **(F)** Box plot of the proportion of immune cell infiltration between the high-risk and low-risk groups.

cell lysis by lowering granzyme B content, and it is an essential factor required for escaping $CD8^+$ T cell-mediated killing in tumors (41, 42). In keeping with this, *VPS4B* was negatively correlated with follicular helper T cells and was found to be highly expressed in CHOL in our study. Overall, our study investigated the prognostic value of five necroptosis-related markers in CHOL. Further in-depth experimental research is needed to explore the potential regulatory effects of this gene set on necroptosis.

In recent years, regulation of the tumor immune microenvironment through immunotherapy has revolutionized cancer treatment (43, 44). Numerous studies have confirmed that immunotherapy based on alteration of the tumor immune microenvironment can affect tumor metastasis, immune escape, and immunotherapy resistance by modifying the immune response (45–47). For instance, a study has suggested that increasing the number or function of NK cells may be a promising approach for the treatment of CHOL (48). Our study found a negative correlation of *STAT6* with resting NK cells, thus suggesting that STAT6 is a potential immunotherapy target. Higher infiltration of M1 and M2 macrophages is related to a poor prognosis by accelerating tumor

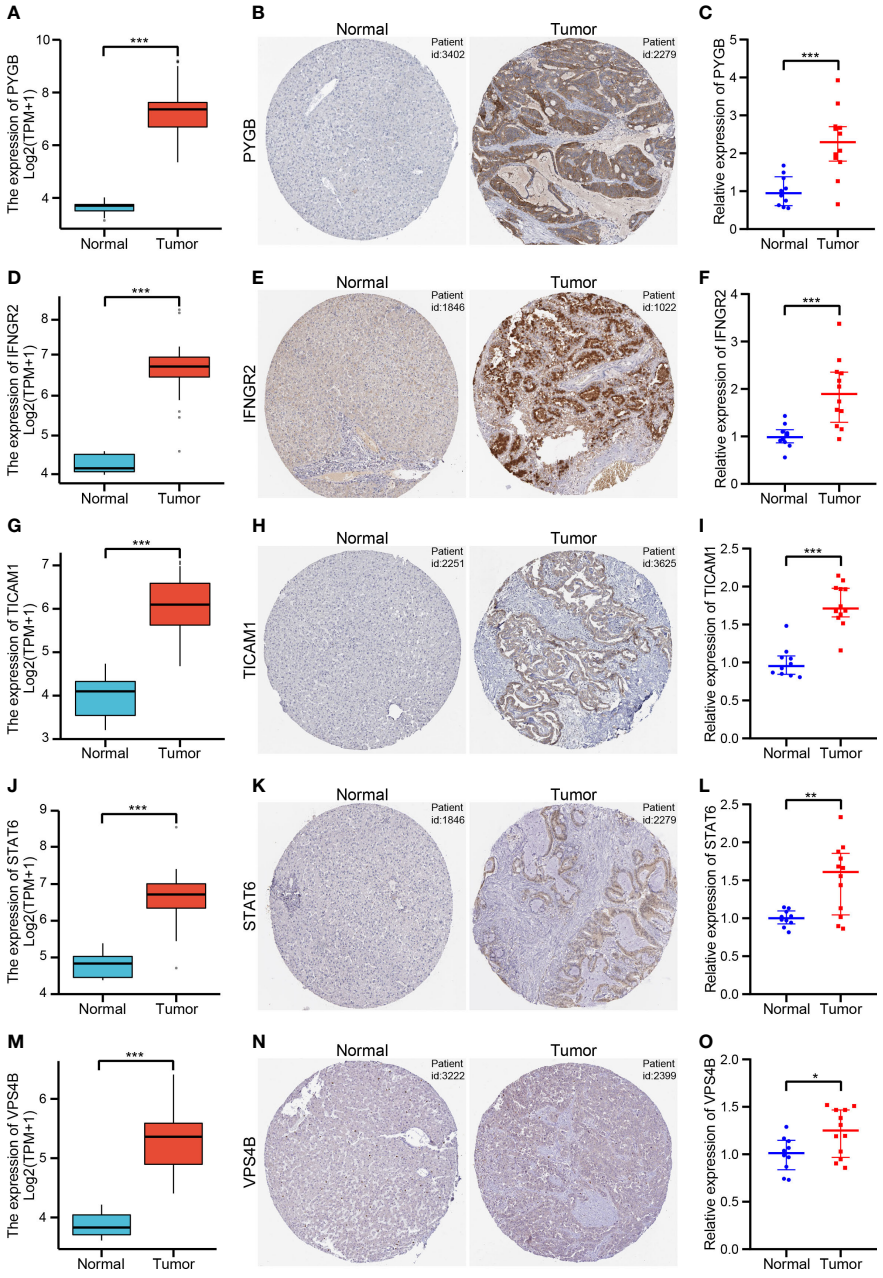


FIGURE 10
Validation of the five NRGs expressions in CHOL tissue samples **(A, D, G, J, M)**. The expression levels of *PYGB* **(A)**, *IFNGR2* **(D)**, *TICAM1* **(G)**, *STAT6* **(J)** and *VPS4B* **(M)** between CHOL and normal samples using the TCGA database. **(B, E, H, K, N)**. Immunohistochemistry of *PYGB* **(B)**, *IFNGR2* **(E)**, *TICAM1* **(H)**, *STAT6* **(K)** and *VPS4B* **(N)** in CHOL and normal samples from the HPA database. **(C, F, I, L, O)**. Relative expression of *PYGB* **(C)**, *IFNGR2* **(F)**, *TICAM1* **(I)**, *STAT6* **(L)** and *VPS4B* **(O)** was detected by qRT-PCR in CHOL and normal samples. *p < 0.05, **p < 0.01, ***p < 0.001.

progression through the secretion of pro-angiogenic factors, activation of the Wnt/β-catenin pathway, and suppression of the antitumor functions of T cells (49). In our study, the high-risk group, which had a poor prognosis, had a higher level of M0 macrophage infiltration, indicating that a greater number of non-activated macrophages were present.

The DEGs between the high- and low-risk groups were enriched in immune-related biological processes and pathways. The five genes involved in our prognostic signature correlated with different levels of immune cell infiltration, such as NK cells, T cells, monocytes, M0 macrophages, and plasma cells. Our results show that, based on the gene signature, there were clear differences in the degree of immune cell infiltration between the high-risk and low-risk groups. The high-risk group tended to exhibit a higher proportion of multiple types of T cells, whereas the low-risk group exhibited a higher proportion of multiple types of B cells. In addition, the low-risk group had higher stromal, immune, and ESTIMATE scores than the high-risk group. In summary, our prognostic signature for CHOL based on necroptosis-related genes could reflect the tumor immune microenvironment of CHOL, which could potentially contribute to personalized immunotherapy and targeted therapy for patients with CHOL.

According to previous studies examining genomic alterations, gene mutations in CHOL usually result in poor outcomes (50). Our study also demonstrated that necroptosis-related genes were positively correlated with genomic alterations, and the high-risk group (mutation rate: 31.37%) exhibited more somatic mutations than the low-risk group (mutation rate: 23.53%). In particular, missense mutations were by far the most predominant mutation type found in CHOL. Moreover, *PBRM1* and *BAP1* exhibited significantly increased mutation rates and multiple mutation types in the high-risk group. In addition, the high-risk group exhibited higher levels of resistance to treatment with GW843682X, mitomycin C, rapamycin, and S-trityl-L-cysteine. These results show that our prognostic signature could be used as a potential predictor of the efficacy of medical treatment for CHOL. Moreover, the occurrence of drug resistance may be reduced by regulation of this signature, which could potentially lead to new breakthroughs in the choice of individual therapeutic strategies.

However, the current study has some limitations. First, the data gathered were from public databases, which were limited in sample size. Future research with a larger sample size is needed to overcome these limitations. Secondly, the identified genes have complex functions and molecular mechanisms that need to be further verified in cellular and animal models. Finally, more detailed clinical follow-up data are required to confirm the value of our prognostic model.

5 Conclusion

In this study, we shed further light on the role of necroptosis in the prognosis of CHOL. Our results indicate that the prognostic model derived from the five NRGs can accurately predict the prognosis of patients with CHOL. Furthermore, the risk score derived from the necroptosis model is associated with important biological functions and is clinically significant. Therefore, the predictive signature of the five NRGs may help devise individualized treatments for patients in the future.

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 within the article/Supplementary Material.

Ethics statement

The studies involving human participants were reviewed and approved by The Ethics Committee of the First Affiliated Hospital of Zhengzhou University (Approval Number: SS-2019-018). The patients/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

LX and ZG designed the project and analyzed the data. XG downloaded the data and drafted the manuscript. JX collected and analyzed the data. All authors contributed to the article and approved the submitted version.

Funding

This research is supported by Henan Medical Science and Technology Joint Building Program (no. LHGJ20190255).

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.2023.1118816/full#supplementary-material

References

1. Raggi C, Taddei ML, Rae C, Braconi C, Marra F. Metabolic reprogramming in cholangiocarcinoma. *J Hepatol* (2022) 77(3):849–64. doi: 10.1016/j.jhep.2022.04.038

2. Lee H-Y, Hong I-S. Targeting liver cancer stem cells: An alternative therapeutic approach for liver cancer. *Cancers* (2020) 12(10):2746. doi: 10.3390/cancers12102746

3. Xu L, Wang L, Zhou L, Dorfman RG, Pan Y, Tang D, et al. The SIRT2/cMYC pathway inhibit peroxidation-related apoptosis in cholangiocarcinoma through metabolic reprogramming. *Neoplasia* (2019) 21(5):429–41. doi: 10.1016/j.neo.2019.03.002

4. Sun Q, Wang H, Xiao B, Xue D, Wang G. Development and validation of a 6-gene hypoxia-related prognostic signature for cholangiocarcinoma. *Front Oncol* (2022) 12. doi: 10.3389/fonc.2022.954366

5. Edge SB, Compton CC. The American joint committee on cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol* (2010) 17(6):1471–4. doi: 10.1245/s10434-010-0985-4

6. Kim EY, Yu JS, Yang M, Kim AK. Sub-Toxic dose of apigenin sensitizes HepG2 cells to TRAIL through ERK-dependent up-regulation of TRAIL receptor DR5. *Molecules Cells* (2013) 35(1):32–40. doi: 10.1007/s10059-013-2175-2

7. Baik JY, Liu Z, Jiao D, Kwon H-J, Yan J, Kadigamuwa C, et al. ZBP1 not RIPK1 mediates tumor necroptosis in breast cancer. *Nat Commun* (2021) 12(1):2666. doi: 10.1038/s41467-021-23004-3

8. Bai L, Kong M, Duan Z, Liu S, Zheng S, Chen Y. M2-like macrophages exert hepatoprotection in acute-on-chronic liver failure through inhibiting necroptosis-S100A9-necroinflammation axis. *Cell Death disease* (2021) 12(1):93. doi: 10.1038/s41419-020-03378-w

9. Sarcognato S, de Jong IEM, Fabris L, Cadamuro M, Guido M. Necroptosis in cholangiocarcinoma. *Cells* (2020) 9(4):982. doi: 10.3390/cells9040982

10. Cadamuro M, Brivio S, Spirli C, Joplin RE, Strazzabosco M, Fabris L. Autocrine and paracrine mechanisms promoting chemoresistance in cholangiocarcinoma. *Int J Mol Sci* (2017) 18(1):149. doi: 10.3390/ijms18010149

11. Morton SD, Cadamuro M, Brivio S, Vismara M, Stecca T, Massani M, et al. Leukemia inhibitory factor protects cholangiocarcinoma cells from drug-induced apoptosis *via* a PI3K/AKT-dependent mcl-1 activation. *Oncotarget* (2015) 6 (28):26052–64. doi: 10.18632/oncotarget.4482

12. Xu B, Xu M, Tian Y, Yu Q, Zhao Y, Chen X, et al. Matrine induces RIP3-dependent necroptosis in cholangiocarcinoma cells. *Cell Death Discov* (2017) 3:16096. doi: 10.1038/cddiscovery.2016.96

13. Chen Y, Li Z-Y, Zhou G-Q, Sun Y. An immune-related gene prognostic index for head and neck squamous cell carcinoma. *Clin Cancer Res* (2021) 27(1):330–41. doi: 10.1158/1078-0432.CCR-20-2166

14. Shen Y, Peng X, Shen C. Identification and validation of immune-related lncRNA prognostic signature for breast cancer. *Genomics* (2020) 112(3):2640–6. doi: 10.1016/j.ygeno.2020.02.015

15. Zhang J, Zhang S, Zuo L, Yue W, Li S, Xin S, et al. Differential expression profiling of lncRNAs related to Epstein-Barr virus infection in the epithelial cells. *J Med Virol* (2019) 91(10):1845–55. doi: 10.1002/jmv.25516

16. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* (2016) 44(8):e71. doi: 10.1093/nar/gkv1507

17. Ahn KS, O'Brien D, Kang YN, Mounajjed T, Kim YH, Kim T-S, et al. Prognostic subclass of intrahepatic cholangiocarcinoma by integrative molecular-clinical analysis and potential targeted approach. *Hepatol Int* (2019) 13(4):490–500. doi: 10.1007/s12072-019-09954-3

18. Jusakul A, Yong CH, Cutcutache I, Lim JQ, Huang MN, Padmanabhan N, et al. Whole-genome and epigenomic landscapes of etiologically distinct subtypes of cholangiocarcinoma. *Cancer Discov* (2017) 7(10):1116–35. doi: 10.1158/2159-8290.cd-17-0368

19. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43(7):e47. doi: 10.1093/nar/gkv007

20. von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res* (2003) 31(1):258–61. doi: 10.1093/nar/gkg034

21. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an r package for comparing biological themes among gene clusters. *Omics-a J Integr Biol* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118

22. Collarino A, Garganese G, Fragomeni SM, Arias-Bouda LMP, Ieria FP, Boellaard R, et al. Radiomics in vulvar cancer: First clinical experience using f-18-FDG PET/CT images. *J Nucl Med* (2019) 60(2):199–206. doi: 10.2967/jnumed.118.215889

23. Ogura A, Konishi T, Beets GL, Cunningham C, Garcia-Aguilar J, Iversen H, et al. Lateral nodal features on restaging magnetic resonance imaging associated with lateral local recurrence in low rectal cancer after neoadjuvant chemoradiotherapy or radiotherapy. *JAMA Surg* (2019) 154(9):e192172. doi: 10.1001/jamasurg.2019.2172

24. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* (2019) 10(1):1523. doi: 10.1038/s41467-019-09234-6

25. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci United States America* (2005) 102 (43):15545–50. doi: 10.1073/pnas.0506580102

26. Zhang H, Meltzer P, Davis S. RCircos: an r package for circos 2D track plots. *BMC Bioinf* (2013) 14:244. doi: 10.1186/1471-2105-14-244

27. Li A, Zhang Z, Ru X, Yi Y, Li X, Qian J, et al. Identification of SLAMF1 as an immune-related key gene associated with rheumatoid arthritis and verified in mice collagen-induced arthritis model. *Front Immunol* (2022) 13:961129. doi: 10.3389/fimmu.2022.961129

28. 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(9):e107468. doi: 10.1371/journal.pone.0107468

29. Xia J, Gill EE, Hancock REW. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. *Nat Protoc* (2015) 10(6):823–44. doi: 10.1038/nprot.2015.052

30. 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

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 (5):453–7. doi: 10.1038/nmeth.3337

32. Uhlen M, Zhang C, Lee S, Sjostedt E, Fagerberg L, Bidkhori G, et al. A pathology atlas of the human cancer transcriptome. *Science* (2017) 357(6352):eaan2507. doi: 10.1126/science.aan2507

33. Kang YJ, Bang B-R, Han KH, Hong L, Shim E-J, Ma J, et al. Regulation of NKT cell-mediated immune responses to tumours and liver inflammation by mitochondrial PGAM5-Drp1 signalling. *Nat Commun* (2015) 6:8371. doi: 10.1038/ncomms9371

34. Migocka-Patrzalek M, Elias M. Muscle glycogen phosphorylase and its functional partners in health and disease. *Cells* (2021) 10(4):883. doi: 10.3390/cells10040883

35. Xia B, Zhang K, Liu C. PYGB promoted tumor progression by regulating wnt/beta-catenin pathway in gastric cancer. *Technol Cancer Res Treat* (2020) 19:1533033820926592. doi: 10.1177/1533033820926592

36. Cui G, Wang H, Liu W, Xing J, Song W, Zeng Z, et al. Glycogen phosphorylase b is regulated by miR101-3p and promotes hepatocellular carcinoma tumorigenesis. *Front Cell Dev Biol* (2020) 8. doi: 10.3389/fcell.2020.566494

37. Williams JB, Li S, Higgs EF, Cabanov A, Wang X, Huang H, et al. Tumor heterogeneity and clonal cooperation influence the immune selection of IFN-gamma-signaling mutant cancer cells. *Nat Commun* (2020) 11(1):602. doi: 10.1038/s41467-020-14290-4

38. Zhuang X, Veltri DP, Long EO. Genome-wide CRISPR screen reveals cancer cell resistance to NK cells induced by NK-derived IFN-gamma. *Front Immunol* (2019) 10:2879. doi: 10.3389/fimmu.2019.02879

39. Yang Q, Liu T-T, Lin H, Zhang M, Wei J, Luo W-W, et al. TRIM32-TAX1BP1-dependent selective autophagic degradation of TRIF negatively regulates TLR3/4-mediated innate immune responses. *PloS Pathog* (2017) 13(9):e1006600. doi: 10.1371/journal.ppat.1006600

40. Tian Q, Zhang Y, Wang G, Jin Y, Shen Z. STAT6 silencing induces hepatocellular carcinoma-derived cell apoptosis and growth inhibition by decreasing the RANKL expression. *Biomed Pharmacother* (2017) 92:1–6. doi: 10.1016/j.biopha.2017.05.029

41. Frey N, Tortola L, Egli D, Janjuha S, Rothgangl T, Marquart KF, et al. Loss of Rnf31 and Vps4b sensitizes pancreatic cancer to T cell-mediated killing. *Nat Commun* (2022) 13(1):1804. doi: 10.1038/s41467-022-29412-3

42. Szymanska E, Nowak P, Kolmus K, Cybulska M, Goryca K, Derezinska-Wolek E, et al. Synthetic lethality between VPS4A and VPS4B triggers an inflammatory response in colorectal cancer. *EMBO Mol Med* (2020) 12(2):e10812. doi: 10.15252/emmm.201910812

43. Schwabe RF, Luedde T. Apoptosis and necroptosis in the liver: a matter of life and death. *Nat Rev Gastroenterol Hepatol* (2018) 15(12):738–52. doi: 10.1038/s41575-018-0065-y

44. Yang C, Sun P, Deng M, Loughran P, Li W, Yi Z, et al. Gasdermin d protects against noninfectious liver injury by regulating apoptosis and necroptosis. *Cell Death Dis* (2019) 10(7):481. doi: 10.1038/s41419-019-1719-6

45. Zeng D, Wu J, Luo H, Li Y, Xiao J, Peng J, et al. Tumor microenvironment evaluation promotes precise checkpoint immunotherapy of advanced gastric cancer. *J Immunother Cancer* (2021) 9(8):e002467. doi: 10.1136/jitc-2021-002467

46. Danaher P, Warren S, Lu R, Samayoa J, Sullivan A, Pekker I, et al. Pan-cancer adaptive immune resistance as defined by the tumor inflammation signature (TIS): results from the cancer genome atlas (TCGA). *J Immunother Cancer* (2018) 6(1):63. doi: 10.1186/s40425-018-0367-1

47. Dong Y, Wan Z, Gao X, Yang G, Liu L. Reprogramming immune cells for enhanced cancer immunotherapy: Targets and strategies. *Front Immunol* (2021) 12. doi: 10.3389/fimmu.2021.609762

48. Jung IH, Kim DH, Yoo DK, Baek SY, Jeong SH, Jung DE, et al. *In vivo* study of natural killer (NK) cell cytotoxicity against cholangiocarcinoma in a nude mouse model. *In Vivo* (2018) 32(4):771–81. doi: 10.21873/invivo.11307

49. Fabris L, Sato K, Alpini G, Strazzabosco M. The tumor microenvironment in cholangiocarcinoma progression. *Hepatology* (2021) 73:75–85. doi: 10.1002/hep.31410

50. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* (2013) 499(7457):214–8. doi: 10.1038/nature12213

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 03 March 2023
DOI 10.3389/fimmu.2023.1149989

OPEN ACCESS

EDITED BY
Jun Liu,
Yuebei People's Hospital, China

REVIEWED BY
Boxuan Zhou,
Third Affiliated Hospital of Sun Yat-sen University, China
Tong Jiao,
Karolinska Institutet (KI), Sweden

*CORRESPONDENCE
Xiaoming Zou
zou4930@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Immunology

RECEIVED 23 January 2023
ACCEPTED 23 February 2023
PUBLISHED 03 March 2023

CITATION
Yang H, Zou X, Yang S, Zhang A, Li N and Ma Z (2023) Identification of lactylation related model to predict prognostic, tumor infiltrating immunocytes and response of immunotherapy in gastric cancer. *Front. Immunol.* 14:1149989. doi: 10.3389/fimmu.2023.1149989

Identification of lactylation related model to predict prognostic, tumor infiltrating immunocytes and response of immunotherapy in gastric cancer

Hao Yang†, Xiaoming Zou*, Shifeng Yang†, Ange Zhang†, Nana Li and Zhen Ma

Department of General Surgery, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China

Background: The epigenetic regulatory chemical lactate is a product of glycolysis. It can regulate gene expression through histone lactylation, thereby promoting tumor proliferation, metastasis, and immunosuppression.

Methods: In this study, a lactylation-related model for gastric cancer (GC) was constructed, and its relationships to prognosis, immune cell infiltration, and immunotherapy were investigated. By contrasting normal tissues and tumor tissues, four lactylation-related pathways that were substantially expressed in GC tissues were found in the GSEA database. Six lactylation-related genes were screened for bioinformatic analysis. The GC data sets from the TCGA and GEO databases were downloaded and integrated to perform cluster analysis, and the lactylation related model was constructed by secondary clustering.

Results: The fingding demonstrated that the lactylation score has a strong correlation with the overall survival rate from GC and the progression of GC. Mechanistic experiments showed that abundant immune cell infiltration (macrophages showed the highest degree of infiltration) and increased genetic instability are traits of high lactylation scores. Immune checkpoint inhibitors (ICIs) demonstrated a reduced response rate in GC with high lactylation scores. At the same time, tumors with high lactylation scores had high Tumor Immune Dysfunction and Exclusion scores, which means that they had a higher risk of immune evasion and dysfunction.

Discussion: These findings indicate that the lactylation score can be used to predict the malignant progression and immune evasion of GC. This model also can guide the treatment response to ICIs of GC. The constructed model of the lactate gene is also expected to become a potential therapeutic target for GC and diagnostic marker.

KEYWORDS

lactylation, prognosis, infiltrating immunocytes, immuntherapy, gastric cancer

Introduction

One of the most prevalent malignant tumors in the world is gastric cancer (GC), which poses a severe threat to human health (1). The evolution of GC is a complicated pathological process involving a number of variables and phases, which is the result of the interaction between dietary factors, host genes, *Helicobacter pylori* infection, and environmental factors. Due to the atypical symptoms of early-stage GC, majority of cases are already advanced when they are diagnosed. GC patients typically have a poor prognosis, a high risk of local recurrence, and distant metastasis (2–4). As GC research has advanced, it has been found that GC is not caused only by specific gene mutations, but cellular metabolic dysfunction is also a key marker of GC progression (5–8). There is mounting evidence that tumor metabolism is crucial to the initiation and development of malignancies as well as affecting immune cells through the release of metabolites (such as lactate and arginine). There is metabolic competition in the tumor environment as a outcome of this energetic transition between tumors and immunocytes, which limits effective supply of nutrients and leads to microenvironmental acidosis, thereby hindering the function of immune cells (9).

Lactate was once considered only a metabolite of glycolysis and the final product of the Warburg effect. A rising number of research have revealed that lactate is not just an essential energy source., signaling molecule, and immunomodulatory molecule but can also control body's metabolism, immunological response, and intercellular communication (10). When tumor cells undergo abnormal glycolysis, they continue to intake a lot of glucose and make a lot of lactate even when there is plenty of oxygen available to them (11). Lactate accumulates in the cells and is exported to the extracellular environment by activating the transport proteins on the cell membrane, eventually forming the acidic tumor microenvironment. In addition to providing the energy required by cells as a fuel substrate, lactate can be utilized as a precursor material to modify histone lysines through lactylation (12). Histone lactylation is a crucial mechanism through which lactate performs its duties and takes part in crucial biological processes, such as glycolysis-related cellular functions (13), macrophage polarization (14), neurodevelopment (15), and regulation of tumor proliferation (16).

Although the lactylation modification has received extensive attention, the relevant articles are still limited. In particular, there are few articles on lactylation modification in GC. This detailed investigation of lactylation-related gene expression and relevance in GC was carried out by us. First, by comparing normal gastric tissues and GC tissues, four lactylation-related pathways with significantly elevated gene expression in GC tissues were identified by GSEA. We hypothesize that these pathways directly or indirectly contribute to the development of GC, leading to a poorer prognosis. After differential analysis and univariate Cox analysis of the above pathways, we obtained six prognostic lactylation genes. By secondary clustering based on these six lactation-related genes, we constructed a model ("lactylation score") that classified as potential screening molecules for GC, which helped to discover various immunocellular infiltration and genetic instability patterns.

Through further analysis, the results revealed that GC patients with high lactylation scores possessed greater potential for immune evasion and lower rates of immunotherapy response, which also means that the lactylation score could become a method for forecasting patients' reactions to immune checkpoint inhibitor (ICI) therapy. The scoring model's PPI network was built, and the hub gene PLOD2 and its downstream lactylation target gene GLUT3 were chosen for experimental validation. Both GC cell lines and GC tissues have significantly higher levels of PLOD2 and GLUT3 expression. After treating GC cell lines with lactate dehydrogenase inhibitors, their expression was decreased, demonstrating the strong relationship between lactylation and the expression of these two genes. Flowchart of this study shows in Figure 1.

Materials and methods

Data retrieval and processing

Through the GSEA database (https://www.gsea-msigdb.org, December 2021), lactylation-related pathways and their genes were downloaded. Through the TCGA database (https://portal.gdc.cancer.gov/, December 2021), we obtained the raw mRNA matrix data in fragments per kilobase million (FPKM) format and copy number data of GC tissues. From the TCGA database, we also obtained relevant clinical data of the GC patients. Baseline characteristics of patients was summarized in Table 1. We downloaded the GSE84437 dataset from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/, December 2021) to obtain the mRNA matrix and the clinical data of GC patients. The batch effect was removed using ComBat function of the SVA package in R for subsequent analysis. The STRING website (https://string-db.org/cgi/input.pl, December 2021) was used to construct the PPI network of lactylation-related genes.

Screening of prognostic lactylation-related genes in GC

By performing differential analysis and univariate Cox analysis of lactylation-related pathways in all GC samples, we obtained six prognostic lactylation-related genes. The copy number variation frequency of lactylation-related genes was calculated from the increase and decrease in gene copy numbers in GC samples from the TCGA database. The number of gene mutations in GC samples from the TCGA database was calculated to draw a waterfall plot. The "RCircos" package of the R language was used to plot the gene copy number circle diagram. Through Cox analysis and coexpression analysis, the prognostic network of lactylation-related genes was plotted. We used the Kaplan–Meier method to calculate the survival curves of GC patients and plotted them using the "survminer" package (17). The clinical data and lactylation scores of all GC patients were analyzed to calculate survival time, survival status, and risk division and to construct nomograms and receiver operating characteristic (ROC) curves.

TABLE 1 Baseline characteristics of patients from TCGA and GEO database.

Clinical features	Total patients (803)		TCGA (377)		GSE84437 (426)	
	Number	Percentage (%)	Number	Percentage (%)	Number	Percentage (%)
Age						
<=65	456	56.79	176	46.68	280	65.73
>65	347	43.21	201	53.32	146	34.27
Gender						
Female	265	33	130	34.48	135	31.69
Male	538	67	247	65.52	291	68.31
Fustat						
Alive	451	56.16	227	60.21	224	55.58
Dead	352	43.84	150	39.79	202	47.42
Grade						
1-2	135	16.81	135	35.81	0	0
3	242	30.14	242	64.19	0	0
Unkown	426	53.05	0	0	426	100
Stage T						
1-2	147	18.31	98	25.99	49	11.5
3-4	656	81.69	279	74.01	377	88.5
Stage N						
N0	196	24.41	116	30.77	80	18.78
N1-3	607	75.59	261	69.23	346	81.22

Cluster analysis

We categorized the GC cohort by "ConsensusClusterPlus" package to determine whether the expression of genes related to lactylation was connected with GC. The intragroup correlation is strong and the intergroup correlation is modest for k=2. Using heatmaps, we connected the amounts of lactylation-related gene expression in several types of GC with patient clinical information. We measured the immune cell expression in various GC types using the ssGSEA technique, and then plotted the data using box plots. Through the GSEA website (https://www.gsea-msigdb.org, December 2021), the files of the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were downloaded, and the functional pathways enriched in GC types were plotted as a heatmap using the "GSEABase" and "GSVA" packages of R.

Construction of the lactylation score model

We developed an lactylation score scheme to quatify the lactylation modification level of individual patients by using PCA. Specifically,the overlapping genes(3112) identified from two clusters of GC were selected and employed to perform prognostic analysis for each gene using a univariate COX regression model. The genes with a significant prognostic were extracted for further feature selection by using recursive feature elimination(RFE) with random forest and the 10-fold cross-validation method in the caret package. We then curated the expression profile of the final determined genes to perform PCA analysis, and principal components 1 and 2 were extracted and served as the signature score. We then adopted a formula similar to previous studies to define the lactylation score: lactylation score =∑(PC1i+PC2i).

We performed correlation analysis between lactylation score and immune cells present in the tissue using the ssGSEA algorithm. With the "ggpubr" and "reshape2" packages of R, we analyzed the relationship between clustering classification, lactylation score, and tumor mutation burden. With the "survival" package and the "survminer" package, we conducted joint survival analysis of the high-tumor-mutation-load group, low-tumor-mutation-load group, high-lactylation-score group, and low-lactylation-score group. With the "plyr" package and the "ggpubr" package, we plotted the different clinicopathological features in the lactylation score group as bar graphs and box plots. From the Cancer Immunome Database (TCIA) (https://tcia.at/, December 2021),

the scoring data of ICI treatment and MSI status for GC were downloaded. According to lactylation score, we analyzed the ICI treatments CTLA-4 and PD-1 for pancreatic cancer (18). All gastric cancer patients were divided into MSS, MSI-L and MSI-H groups, and lactylation scores among all groups were calculated. The Tumor Immune Dysfunction and Exclusion (TIDE) score, exclusion score, and dysfunction data of GC were downloaded from the TIDE database (http://tide.dfci.harvard.edu/, December 2021), and the immune evasion and immune dysfunction in high- and low-lactylation-score groups were analyzed. We used "pRRophetic" package of R to predict drug sensitivity of high- and low-lactylation score groups ($p<0.001$).

Collection of tissue samples from GC patients

Tissue samples were collected from GC patients who underwent surgical resection at the Second Affiliated Hospital of Harbin Medical University. These GC patients had not received any other treatments, such as radiotherapy, chemotherapy, or biological treatment, before surgery. All specimens were histopathologically diagnosed by two pathologists according to the diagnostic criteria for GC. All patients provided informed consent.

Cell culture and transfection

Normal gastric mucosal epithelial cells (GES-1) and GCcells (AGS, HGC-27, KATO3, MKN-45) were purchased from Procell Life Science & Technology (Wuhan, China), and the cells were cultured according to the manual instructions. A lactate dehydrogenase A (LDHA) inhibitor (GSK2837808A, MCE) was used to treat HGC cell lines. The HGC cell lines were cultured in six-well plates, and plasmids were transfected using Lipofectamine 3000 according to the instructions. The target sequences of the short hairpin RNA (shRNA) were as follows: PLOD2-RNAi (8407-1), caGCAAGTGTCCTTAAGTCAA; PLOD2-RNAi (8408-1), ggA AATGGACCCACCAAGATT; PLOD2-RNAi (8409-1), c tTTGCCGAAATGCTAGAGAA.

Western blotting

WB assay was done in accordance with prior literature descriptions (19). We used RIPA buffer to extract the proteins needed for western blotting. The subsequent antibodies were utilized: PLOD2 (Proteintech, China), GLUT3 (Abcam, UK) and GAPDH (ImmunoWay, USA).

Immunohistochemistry

The tissue samples were sectioned at 5 µm after being submerged in 4% paraformaldehyde for an entire night. Following an overnight incubation at 4°C with anti-PLOD2 and anti-SLC2A3 antibodies, tissue slices were treated with secondary antibodies conjugated to horseradish peroxidase (20).

RNA extraction and RT-PCR

TRIzol was applied to extract cellular RNA, and a cDNA synthesis kit was applied to create single-stranded cDNA from the recovered cellular RNA (21). The FastStart Universal SYBR Green Master Mix's instructions were followed to perform the qPCR analysis. A 10-µl reaction system was prepared.

Colony formation assays

After 10 days of incubation, the GC cell lines were seeded in a six-well plate, and colony formation was apparent to the naked eye.

Ethylenediurea experiment

EdU experiment was performed according to the previous literature (22, 23).

Wound-healing assay

Cells were cultured in six-well plates until full confluence, then starved with serum-free medium. We scratched a 10-µl pipette tip across the plate, removing a line of cells. Under a microscope, pictures were obtained at 0, 12, and 24 hours to record the extent of wound healing (24).

Transwell assays

The GC cell lines were inoculated into a Transwell chamber containing 200 µl of serum-free medium. Matrigel mix was coated on the upper chamber surface of the Transwell chamber to detect the invasion ability of the cells. When testing the cell migration ability, the bottom of the chamber does not need to be coated with Matrigel. Medium containing 10% FBS was added to the lower culture plate. After 24 h of incubation, the chamber was removed and stained with crystal violet for 30 min. Five randomly selected fields of view were photographed, and their cells were counted under a microscope.

Data analysis

The GraphPad Prism 8.0 software was used to illustrate the results of the data analysis, which was carried out using the SPSS 18.0 program. In particular, the means of the two groups were compared using Student's t tests, and one-way ANOVA analysis was used to determine the statistical significance of the means from multiple groups (>2). The one-way ANOVA analysis for the corrective test was followed by the *post hoc* testing (Tukey test).

Results

Expression of lactylation-related genes in GC

Through the GSEA database, we identified four lactylation-related pathways that were significantly upregulated in GC tissues (Figure 2A). We speculate that these four lactylation-related pathways are directly linked to the incidence and evolvemetn of GC.

After differential and Cox analysis of the above lactylation pathways, we obtained six prognostic lactylation genes and plotted a forest diagram (Figure 2B). Box plot showed the expression of lactylation-related genes in normal gastric samples and GC samples (Figure 2C). Additionally, we examined the prevalence of copy number variation in target genes across all GC samples (Figure 2D). Except for *EFNA3* and *PLOD2*, which showed increased copy number, the other four genes all showed copy number reductions. We also plotted single-gene mutation frequency waterfall plots (Figure 2E) and gene copy number circle plots (Figure 2F) of the six lactylation-related genes.

Next, we combined the TCGA and GEO data to analyze the data of a total of 804 GC patients and performed Kaplan–Meier survival analysis based on lactylation-related gene groups (Figure 2G). The results showed that the survival curves of the six lactylation-related genes all had statistically significant differences. A prognostic network map was constructed through coexpression analysis (Figure 2H). It can be seen from the network map that *NUP50* and *EFNA3* were favorable factors, and the other four lactylation genes were risk factors. They regulate each other within a network, which can form a functional whole and together affect the progression of GC.

Cluster analysis

We separated all GC samples into two clusters using cluster analysis of lactylation-related genes (Figures 3A–C). Significant variations between the two clusters of GC were shown by the survival analysis (Figure 3D). Compared to cluster 1, cluster 2 had a much lower survival rate. We determined the ratios of 23 different immune cell types in the two clusters of GC using the ssGSEA method. 18 types of immunocytes were significantly different between the two clusters of GC (Figure 3E). We also plotted the clinicopathological characteristics and lactylation genes of GC samples into a heatmap (Figure 3F). Using the GSVA algorithm, we plotted the GO and KEGG pathways enriched in the two clusters of GC into a heatmap. The GO-enriched pathway in cluster 2 was mostly focused on angiogenesis and epithelial cell proliferation (Figure 3G), such as ENDOTHELIAL_CELL_PROLIFERATION and REGULATION_OF_VASCULATURE_DEVELOPMENT. The KEGG-enriched pathways in cluster 2 was mostly focused on metastasis-related pathways (Figure 3H), such as FOCAL_ADHENSION, and TGF_BETA_SIGNALING_PATHWAY.

Constructing lactylation score model

The results of PCA on all GC samples (Figure 4A) revealed that there was little overlap between the two GC clusters and that the components within each GC cluster were well correlated. Figure 4B demonstrates that cluster 2's lactylation score was substantially greater than cluster 1's. The survival rate of high lactylation score group was considerably lower than low lactylation score group (Figure 4C), demonstrating that the prognosis is worse the higher the lactylation score. The TMB analysis was also examined in both groups, and the findings revealed a negative correlation between the two (Figures 4D, E). We also constructed a Sankey plot of the cluster

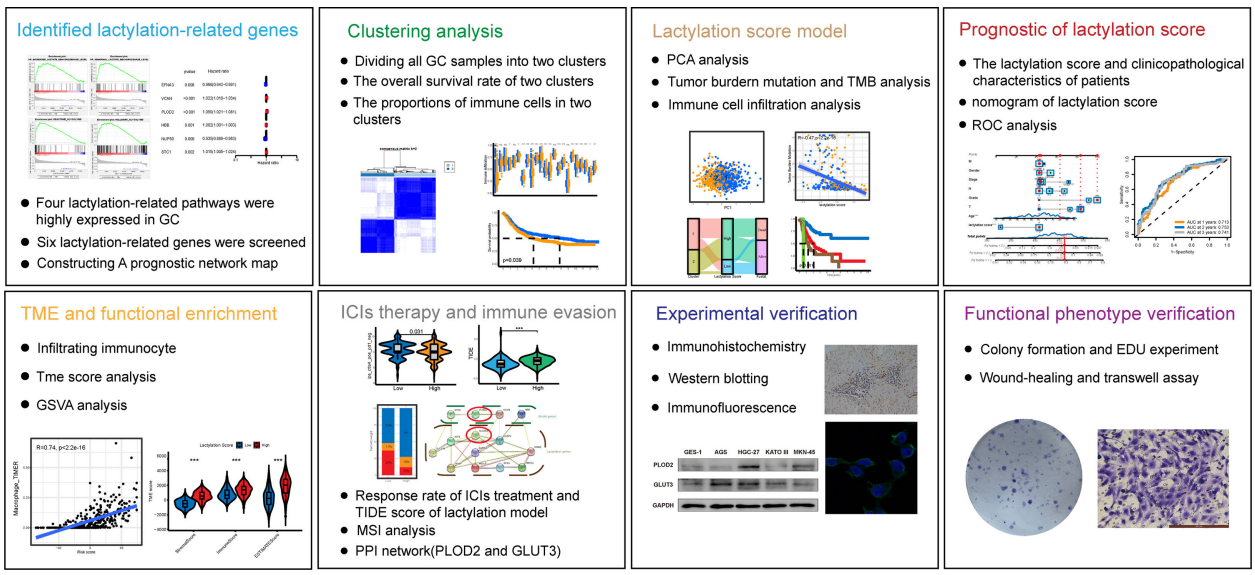


FIGURE 1
Flowchart of the study.

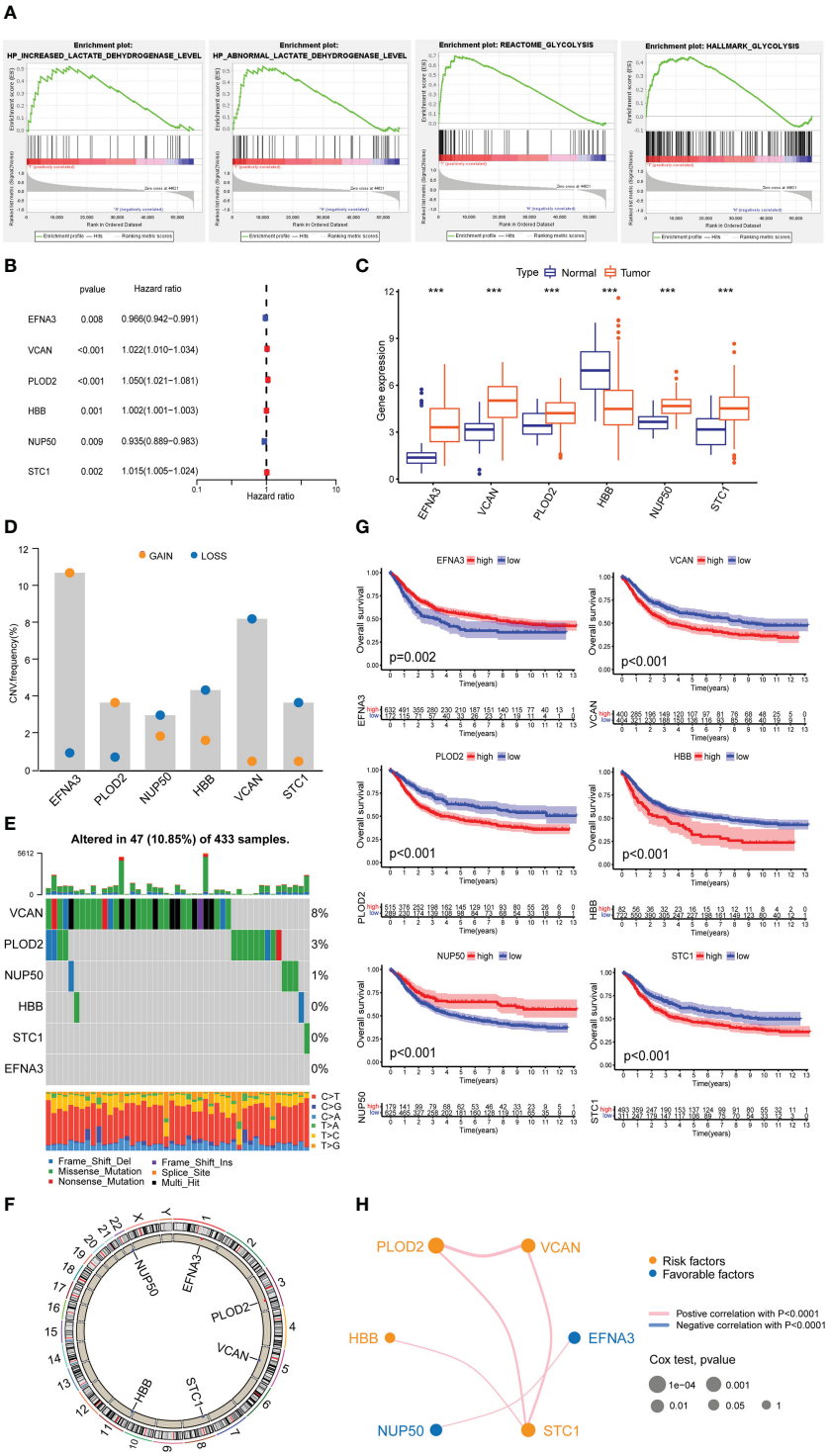


FIGURE 2
Expression of lactylation-related genes in GC. **(A)** Highly expressed GSEA pathway in GC tissues. **(B)** Univariate Cox analysis of lactylation-related genes. **(C)** Expression of lactylation-related genes in GC tissues and normal gastric tissues. ***p<0.001. **(D)** Diagram of the copy number frequencies of lactylation-related genes. **(E)** Waterfall plot of the mutation frequencies of lactylation-related genes. **(F)** Circle plot of the copy numbers of lactylation-related genes. **(G)** Overall survival rate of lactylation-related gene patient groups. **(H)** Prognostic network of lactylation-related genes.

of GC, lactylation score, and patient survival. Figure 4F shows that cluster 2 GC had a high correlation with a high lactylation score, while cluster 1 was associated with a low lactylation score. While the most part of samples in the low-lactylation-score group were in the survival condition, the majority of the patients who died belonged to the high-lactylation-score group. We created a waterfall plot to compare the prevalence of single-gene mutations in two groups (Figures 4G, H). The low-lactylation-score group had a higher

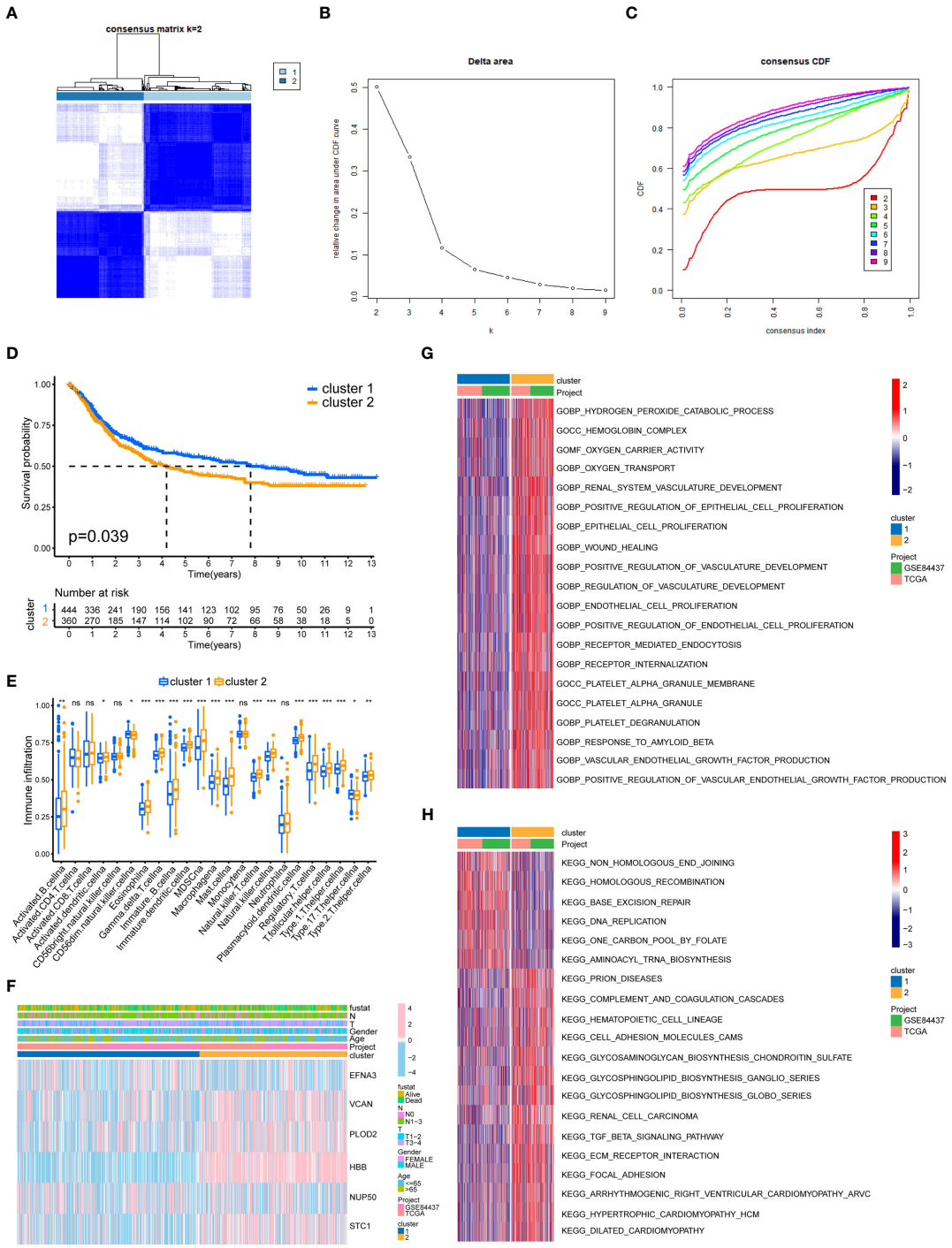


FIGURE 3
GC classification of lactylation-related genes. **(A)** Changes in the length and inclination of the cumulative distribution function (CDF) curve when k=2-9. **(B)** Area under the CDF curve when k=2-9. **(C)** GC samples were divided into two tumor clusters. **(D)** Kaplan–Meier survival curves of the two clusters. **(E)** Immune cell infiltration of the two clusters. *p<0.05; **p<0.01; ***p<0.001; ns, no signifiificance. **(F)** Heatmap of GC classification, lactylation-related genes, and clinicopathological characteristics. **(G)** GO enrichment analysis of the two GC clusters. **(H)** KEGG enrichment analysis of the two clusters.

prevalence of single-gene mutations than the high-lactylation-score group. Figure 4I demonstrates that the high-TMB group's survival rate was significantly greater than the low-TMB group's. Additionally, we performed conjoint analysis between tumor mutation burdens groups and lactylation score groups (Figure 4J). The result revealed statistical differences, demonstrating that the prognosis of GC patients was influenced by both TMB and lactylation scores.

10.3389/fimmu.2023.1149989

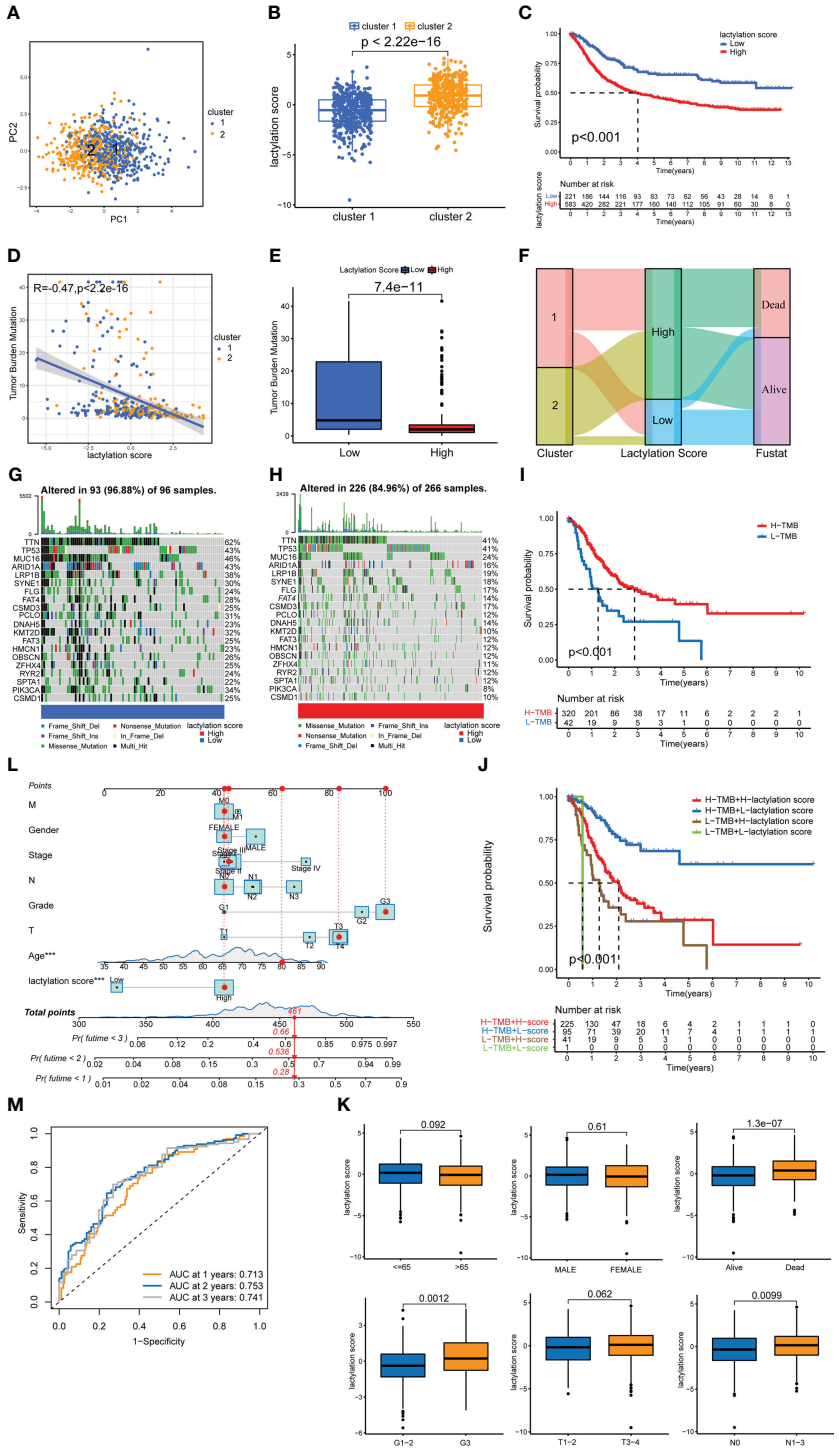


FIGURE 4
lactylation score model. **(A)** PCA of lactylation-related genes. **(B)** Lactylation scores of GC types. **(C)** Kaplan–Meier survival curve of the high- and low-lactylation-score groups. **(D)** Spearman correlation analysis of the lactylation score and immune cells. **(E)** *UBQLN4* expression in the high- and low-lactylation-score groups. **(F)** Sankey plot of GC clusters, lactylation score, and patient survival. **(G)** Waterfall plot of mutation frequencies in the low-lactylation-score group. **(H)** Waterfall plot of mutation frequencies in the high-lactylation-score group. **(I)** Kaplan–Meier survival curve of the high-tumor-mutation-burden group and the low-tumor-mutation-burden group. **(J)** Joint survival analysis was performed in the high and low-tumor-mutation-burden groups and the high- and low-lactylation-score groups. **(K)** The age, sex, survival, grade, T stage, and N stage of patients in the high- and low-lactylation-score groups. **(L)** Nomogram of the lactylation score and clinical information. ***p<0.001 **(M)** AUC of the lactylation score model.

Prognosis and clinicopathological characteristics in different lactylation score groups

Additionally, we examined the patients' clinicopathological traits, as well as their lactylation score. Age, sex, and T stage showed no significantly different between two lactylation scores groups, as shown in Figure 4K. Additionally, a poor prognosis, a high tumor grade, and lymph node metastases were all closely related to a high lactylation score. This shows that a high lactylation score frequently suggests a higher degree of malignancy from the perspective of clinicopathological features. We also constructed the nomogram (Figure 4L), in which we can classify risks of the patients and predict patient survival by statistically scoring the clinicopathological characteristics and lactylation scores. The results show that if the total score of the patients reached 448 points, the 1-, 2-, and 3-year mortality rates of the patients were 28%, 53.6%, and 66%, respectively. We also draweed the ROC curve of the nomogram (Figure 4M). The area of AUCs at 1, 2, and 3 years were all greater than 0.65.

Immune cells infiltration statement and functional enrichment analysis

According to the ESTIMATE algorithm, the stromal and immune scores were statistically higher in high lactylation score group than low lactylation score group. This demonstrated that the group with a high lactylation score had a larger proportion of stromal cell and immune cell infiltration (25). The ESTIMATE score was considerably greater in the group with high lactylation scores than in the group with low lactylation scores, demonstrating a negative correlation between lactylation score level and tumor purity (Figure 5C). By combining various immunocytes analysis methods, we analyzed the correlation between lactylation score and immunocytes infiltration (Figure 5A). The findings revealed that the degree of macrophage and M2-type macrophage infiltration was positively connected with lactylation score (Figure 5B). KEGG and Hallmark enrichment analysis showed that lactylation model was closely relevanted to multiple oncogenic pathways (including WNT, TGF_BETA, MTOR, P53_SIGNALING) (Figures 5D, E). At the same time, Hallmark enrichment analysis revealed that the lactylation model was closely relevanted with HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION and HALLMARK_ANGIOGENESIS (Figure 5E). This indicated that the higher lactylation score, the stronger proliferation, metastasis and invasion ability of gastric cancer.

Immunotherapy response and immune evasion

Our study also revealed that the lactylation score was relevanted to the efficacy of immune checkpoint treatment. Figure 6A shows that patients with lower lactylation scores possessed a higher response rate to ICIs(ctla4_pos_pd1_neg, $p<0.05$). Additionally, high lactylation score group are more liabled to acquire ctla4 immunotherapy resistance. We also predicted the drug sensitivity of the high- and low-lactylation score groups, and found that most drugs expressed higher sensitivity in the low-lactylation score group, while only two drugs (Gefitinib and Metformin) showed higher sensitivity in the high-lactylation score group (Figure S1). Immune evasion can result in resistance to immunotherapy. An algorithm called the TIDE score is used to determine T-cell malfunction and rejection in different tumor types. The scores can be used to forecast the impact of immune checkpoint therapy in patients with tumors in addition to being consistent with immune evasion features (26). Figure 6B demonstrates that the high-lactylation-score group's TIDE, Exclusion, and Dysfunction score were all considerably greater than those in the low-lactylation-score group. Further resulting in immune evasion and immunotherapy resistance was the high-lactylation-score group's increased susceptibility to immunological dysfunction and immune rejection. We also looked at the lactylation score and the instability of the microsatellites. The low-lactylation-score group showed increased microsatellite instability, as shown in Figures 6C, D.

Expression of lactylation-related genes in GC

Figure 6E shows the correlation between the lactylation score and lactylation genes. We constructed a PPI network of lactylation-related genes/proteins and lactylation target genes reported in the literature (Figure 6F). At the center of the PPI network, PLOD2 and SLC2A3 (GLUT3) had a greater priority. Additional, there was an interaction between PLOD2 and GLUT3. We speculate that PLOD2 may affect the lactylation level of GC through GLUT3, thereby leading to the development and progression of GC. Figure 6G shows the correlation between GLUT3 and lactylation genes, where it can be seen that GLUT3 and PLOD2 showed a strong connection. Moreover, the high-lactylation-score group's expression level of GLUT3 was noticeably higher than the low-lactylation-score group's (Figure 6H).

The relation between PLOD2, GLUT3 and lactylation in GC

We performed immunohistochemistry in normal gastric tissues and GC samples (Figure 7A). The figure shows that the expression degrees of PLOD2 and GLUT3 in GC samples were all considerably higher than those in normal gastric samples.Figures 7B, C shows that PLOD2 and GLUT3 were expressed more highly in GC cell lines than GES-1 cells. The PLOD2 expression was significantly high in HGC-27 and MKN-45, while GLUT3 expression was the highest in HGC-27.

To observe whether PLOD2 and GLUT3 were associated with lactylation, we selected different concentrations of lactate dehydrogenase inhibitors to treat HGC-27 cells for 48 h. Figure 7D shows that PLOD2 and GLUT3 were both decreased after HGC-27 were treated by the lactate dehydrogenase inhibitor. When the lactate dehydrogenase concentration was set as 1 μM, the

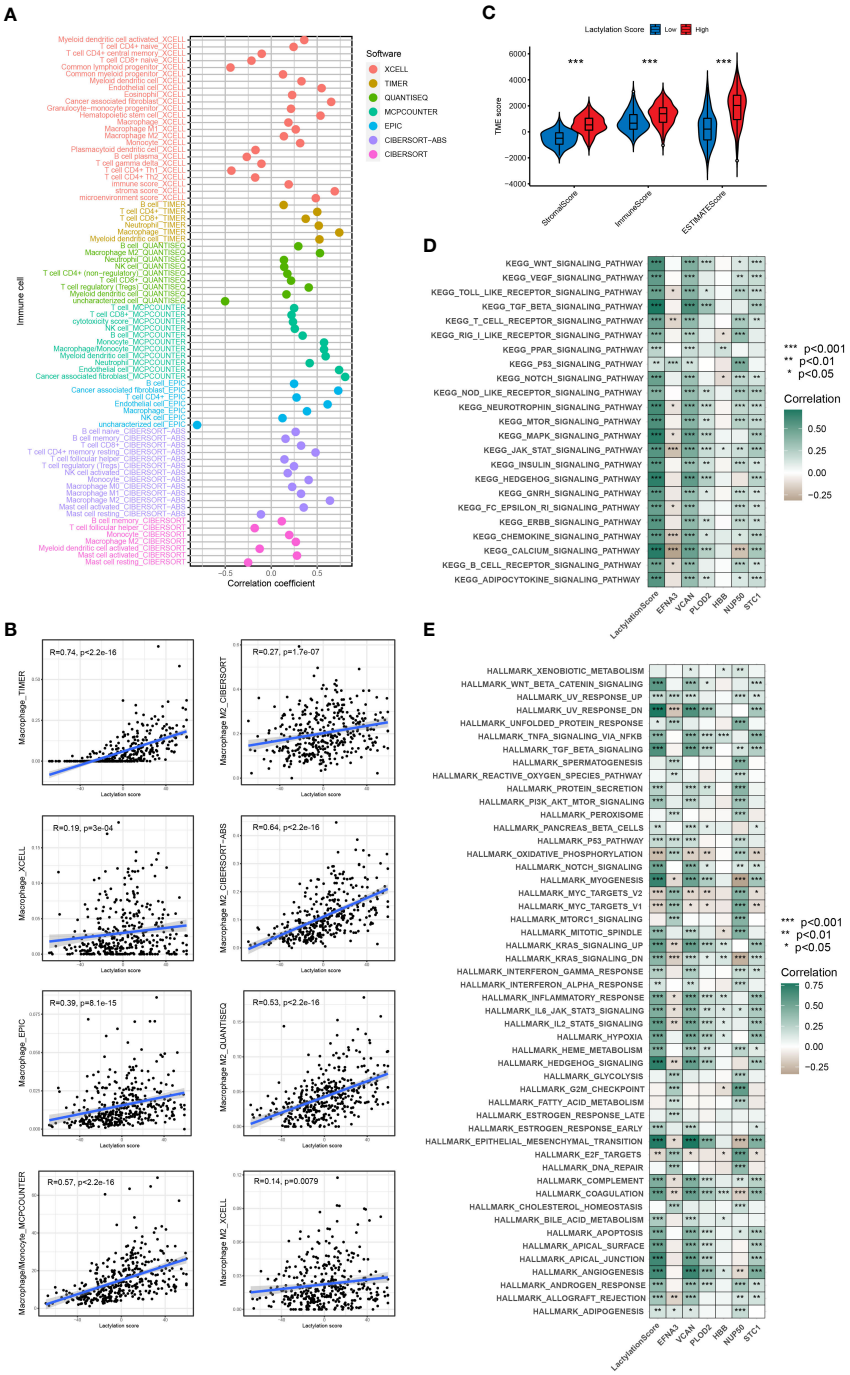


FIGURE 5
Immune cells infiltration and function enrichment analysis. *p<0.05; **p<0.01; ***p<0.001. **(A, B)** The correlation between lactylation score and immune cell infiltration by various immunocytes analysis methods. **(C)** Correlation between lactylation score and the tumor microenvironment of gastric cancer assessed using the ESTIMATE algorithm. **(D, E)** GSVA analysis of lactylation score and lactylation-related genes.

inhibitory effect on PLOD2 and GLUT3 was the most significant. We also constructed three PLOD2-knockdown plasmids expressing heterogeneous nuclear RNAs, namely, sh-8407, sh-8408, and sh-8409. After transfection of each of these three plasmids individually into HGC cells, the expression of PLOD2 was knocked down at the protein level, and the knockdown effect of sh-8409 was the most significant (Figure 7E). Therefore, we chose the sh-8409 plasmid to create a PLOD2-KD group of cells. Knockdown of PLOD2 by sh-8409 also downregulated the expression of GLUT3 (Figure 7E), further supporting a PPI relationship between PLOD2 and GLUT3. L-Lactyl is a pan-antibody of lactylation modification that reflects the level of lactylation in tissue samples. Immunofluorescence assay results showed that GLUT3 and L-Lactyl expression were significantly declined after PLOD2 knockdown (Figure 7F). This indicated that PLOD2 could regulate lactylation modification in GC cell line.

10.3389/fimmu.2023.1149989

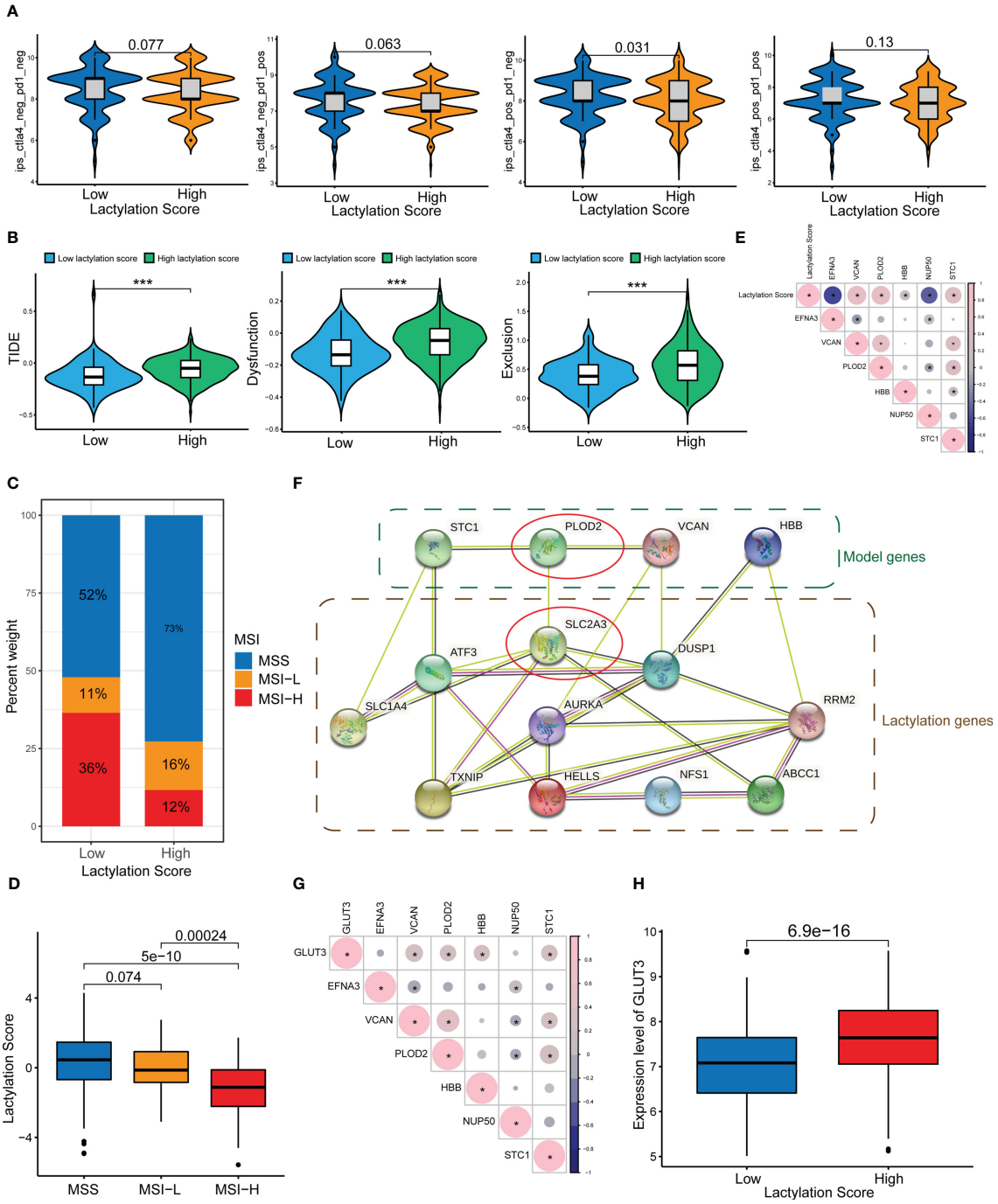


FIGURE 6
ICIs and immune evasion. *p<0.05; ***p<0.001. **(A)** Sensitivity analysis of the high- and low-lactylation-score groups to immunotherapy. **(B)** TIDE score, Dysfunction score, and Exclusion score of the high- and low-lactylation-score groups. **(C, D)** Microsatellite instability in the high- and low-lactylation-score groups. **(E)** Correlations between the lactylation score and lactylation-related gene expression. **(F)** PPI network between lactylation-related proteins and lactylation target genes. **(G)** Correlation between GLUT3 and lactylation-related genes. **(H)** GLUT3 expression in the high- and low-lactylation-score groups.

Functional phenotype of PLOD2 in GC cells

After knocking down PLOD2, we examined the functional phenotyoe change in GC cells. Figure 8A shows that there were considerably fewer cell clones in PLOD2-KD than in PLOD2-NC. EdU experiment also show that the cell viability of the PLOD2-KD group was significantly reduced (Figure 8B). In the scratch assay, the cells in the PLOD2-KD group showed considerably broader scratches than PLOD2-NC group (Figure 8C). Similarly, the Transwell assay demonstrated that PLOD2 gene knockdown prevented HGC-27 and MKN45 cells from migrating and invading (Figure 8D).

10.3389/fimmu.2023.1149989

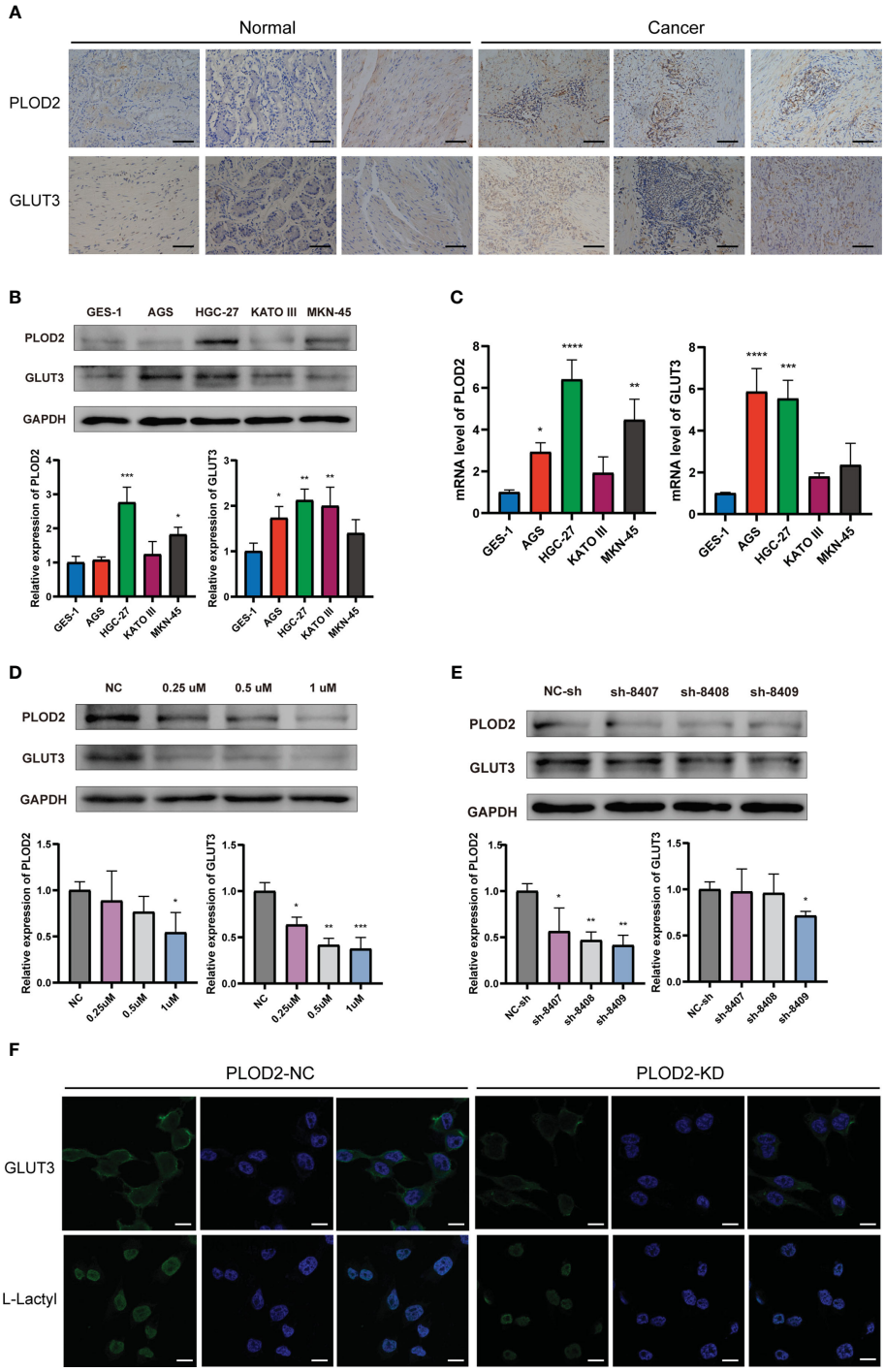


FIGURE 7
Expression of PLOD2 and GLUT3 in GC cell lines and GC tissues. *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001. **(A)** Immunohistochemistry for PLOD2 and GLUT3 in normal gastric tissues and GC tissues. Scale bar, 100 μm. **(B)** Western blots for PLOD2 and GLUT3. **(C)** PCR results of PLOD2 and GLUT3. **(D)** Western blots of HGC cells treated with different concentrations of lactate dehydrogenase inhibitors (0.25 μM, 0.5 μM, and 1 μM). **(E)** Western blots after transfection with the PLOD2-KD plasmid. **(F)** Expression levels of GLUT3 and L-Lactyl in PLOD-NC group and PLOD-KD group were compared by immunofluorescence. Scale bar, 200 μm.

Discussion

Recently, growing evidence has shown that lactate is not only the most important direct source of nutrition for tumor cells, it can also promote the growth, proliferation, metastasis, drug resistance, and immunosuppression of tumors, such as by acidifying the immune microenvironment and increasing the expression of tumor resistance proteins (27, 28). More importantly, researchers at the University of Chicago demonstrated that lactate is an important epigenetic modification molecule that can affect degree

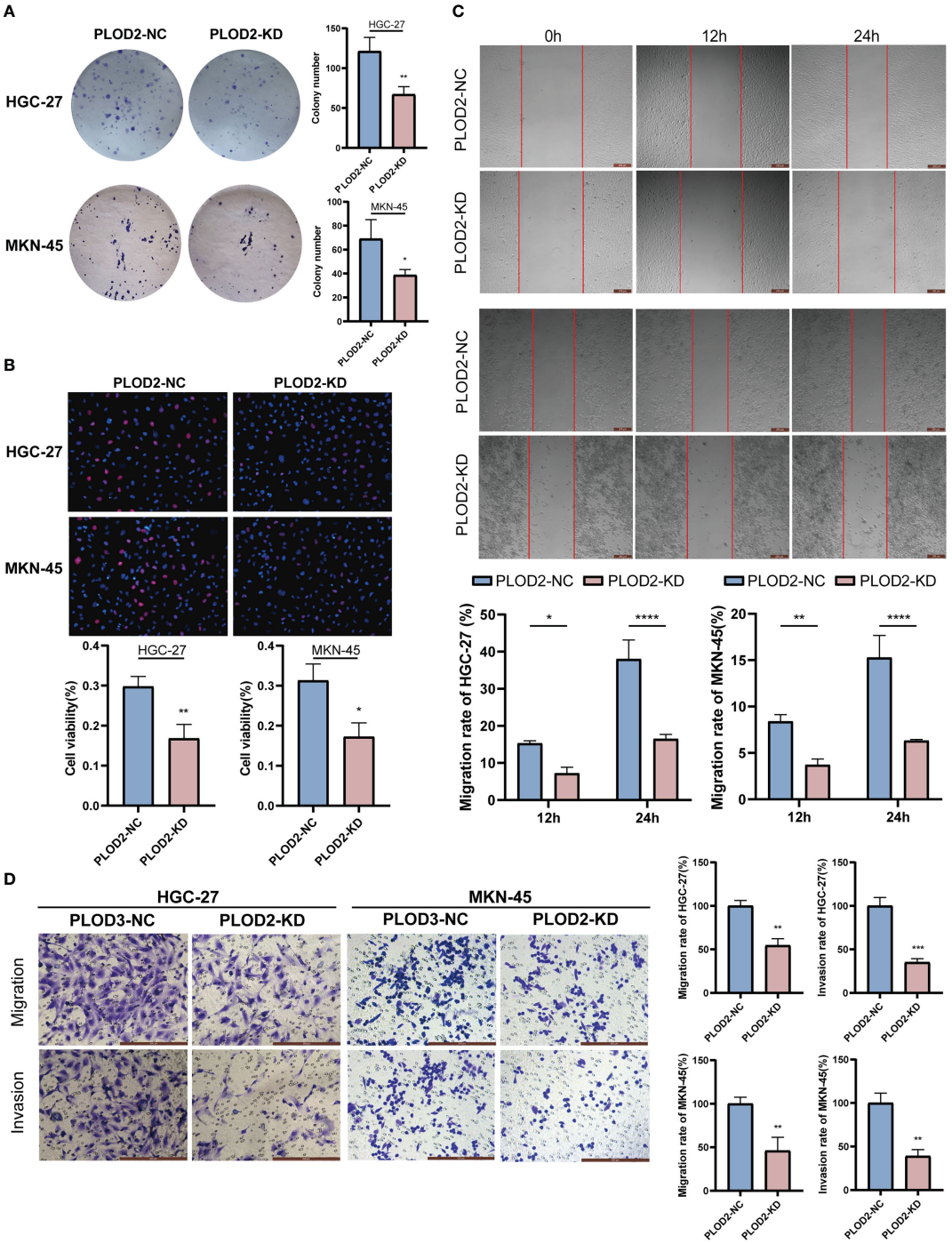


FIGURE 8
Vitro assay results of PLOD2 in gastric cancer cell.*p<0.05; **p<0.01; ***p<0.001; ****p<0.0001. Scale bar, 200 μm. **(A)** Plate cloning assay results of *PLOD2*-NC and PLOD2-KD in HGC-27 and MKN-45 cells. **(B)** EdU assay results of PLOD2-NC and PLOD2-KD in HGC-27 and MKN-45 cells. **(C)** Wound-healing assay results of PLOD2-NC and PLOD2-KD in HGC-27 cells. **(D)** Transwell assay results of PLOD2-WT and PLOD2-KD in HGC-27 and MKN-45 cells.

of macrophage polarization through histone lactylation (12). The lactylation modification of histone lysines is indeed widespread in human and mouse cells and is regulated by glycolysis. In addition, the lysine lactylation of histones is highly sensitive to lactate produced by glycolysis, which can change with the intensity of glycolysis or the level of lactate. Protein lactylation modification is the farthest known downstream molecular mechanism of glycolysis, and lactate in the adjustment of cellular functions.

To investigate the function of GC lactylation modification, we downloaded four lactylation pathways with significantly elevated expression in GC tissues from the GSEA database. Six lactylation-related genes correlatived with GC development were screened by

cluster analysis, and PCA analysis. A lactylation score model was constructed, which was closely associated with increased tumor infiltrating immunocytes, genetic instability and ICI treatment. The six lactylation-related genes in the constructed model wee also closely correlated with the prognosis of GC.

Tumor growth relies heavily on glycolysis. Tumor cells can produce lactate through aerobic glycolysis and maintain a high-lactate environment, thereby inhibiting T cells that attack tumor cells. Moreover, lactate enhances the expression of regulatory T cells to contribute to the defense of malignant cells, thereby evading the attack of the immune system (29). In this study, the TIDE score of high-lactylation-score group was considerably higher than low-lactylation-score group, which means that the higher lactylation score, tumor cell more likely achieve immune evasion and immune dysfunction. Experiments have shown that the key factor in tumor drug resistance is lactate, and the lactate concentration and glycolysis rate can reflect the sensitivity of tumor drugs to a certain extent (30). Meanwhile, higher glycolysis rate of cancer cells is closely related to lower response rate to ICI treatment (31). We also performed drug sensitivity analysis on the lactylation score model. The findings revealed that high-lactylation-score group's sensitivity to ctla4 immunotherapy was lower than low-lactylation-score group's. It proves that patients with high lactylation scores are more likely to develop resistance to immunotherapy, which agrees with the findings of the majority of recent investigations. The above results showed that lactate and lactylation are closely related to immune evasion and sensitivity to ICI treatment. Moreover, lactate and lactylation have received extensive attention as novel target for tumor immunotherapy (32).

LDHA is one of the key enzymes in the reprogramming of glucose metabolism in the TME. It is also the hub protein that connects various cellular metabolic pathways. It directly or indirectly activates signal transduction pathways and regulates immune responses to participate in development and progression of tumors (33). The increase in LDHA level is mainly caused by the increase in tumor glycolytic activity and tumor hypoxic necrosis, which are important drivers of the immunosuppressive microenvironment. LDHA can promote the conversion of pyruvate to lactate, and its activity is positively correlated with the Warburg effect (34). The enzymatic activity of LDHA is regulated by posttranslational modifications, including lactylation, acetylation, and phosphorylation (35). Some studies have shown that when LDH is knocked out, the lactylation level is also inhibited (12). Wang et al. reported that by directly binding to LDHA and PKM2, HULC increases their phosphorylation levels to regulate their activity and ultimately accelerate glycolysis and promote cell proliferation (36). Chen et al. found that CENP-N affects tumor progression by participating in the aerobic glycolysis process of nasopharyngeal carcinoma cells (37). These studies fully demonstrate that aerobic glycolysis and lactylation modification are involved in the regulation of posttranslational modification and that LDHA inhibitors can inhibit the lactylation process by inhibiting the activity of the LDHA enzyme. Current clinical studies are testing antitumor drugs targeting LDHA (38). In our study, after the use of LDHA inhibitors to block lactylation in GC cells, the expression of the lactylation-related gene PLOD2 and the lactylation target gene GLUT3 were inhibited to varying degrees, and the degree of inhibition was positively correlated with the concentration of LDHA inhibitor. This also indicates that the expression levels of the PLOD2 and GLUT3 proteins are regulated by lactylation modification.

PLOD2 is a functional enzyme located in the rough endoplasmic reticulum of the cytoplasm. It participates in the posttranslational modification of collagen and promotes the synthesis of collagen fibers. Fibrotic collagen plays a key role in promoting tumor invasiveness (39). PLOD2 can participate in extracellular matrix remodeling by promoting the deposition of collagen fibers, thereby improving the invasiveness of tumor cells (40–44). In a variety of tumors, clinical data indicate that increased PLOD2 expression can be used as an independent poor-prognostic factor and is associated with poor patient survival (45). This is consistent with the results of our study, in which the overall survival rate of GC patients in the high-PLOD2-expression group was significantly lower than that in the low-PLOD2-expression group, and PLOD2 was the highest-ranked hub gene of the lactylation score model, which was also validated by immunohistochemistry. We found that the expression of PLOD2 in GC tissues, especially in GC foci, was significantly higher than that in normal gastric tissues. At the same time, after PLOD2 knockdown, the proliferation and invasiveness of GC HGC-27 and MKN-45 cells were significantly reduced. TGF-β1 seems to be an important factor in the regulation of PLOD2 (46–48). Through clustering analysis of lactylation-related genes, including PLOD2, we found that cluster 2 GC was enriched in the TGF_BETA_SIGNALING_PATHWAY and ECM_RECEPTOR_INTERACTION pathways, which is consistent with our other results.

The glucose transporter is one of the most important transmembrane proteins in the human body, with a total of 14 subtypes (GLUT1-GLUT14). Among them, GLUT3 encoded by the SLC2A family is a tissue-specific glucose transporter with high affinity for glucose in the GLUT family. GLUT3 is highly expressed in tumor cells and promotes the transport and uptake of glucose by tumor cells (49–52). The affinity of GLUT3 for glucose is five times that of the well-known GLUT1 transporter (53). Taekyu Ha et al. reported that Caveolin-1 factor upregulated glucose transport and uptake and aerobic glycolysis by promoting direct binding of HMGA1 to the GLUT3 promoter region (54). Liu reported that the transcriptional repressor family member ZBTB7A could inhibit the transcription and expression of GLUT3, and the knockdown of ZBTB7A could increase glucose transport and uptake as well as the synthesis of lactate (55). Wang et al. demonstrated that the enhancement of the insulin-stimulated PI3K/Akt phosphorylation pathway can promote the expression of GLUT3, LDHA and monocarboxylic acid transporter 1 (MCT1) (56). Through the above studies, we found that GLUT3 is closely related to glycolysis and lactylation. Therefore, we selected *GLUT3* as the lactylation target gene. We found that the expression level of *GLUT3* was increased in GC cell lines and GC tissues. The results show that lactate dehydrogenase inhibitors could affect the expression of GLUT3 by inhibiting the lactylation level.

In this study, by analyzing the lactylation pathways with elevated component expression in GC tissues, we found

lactylation-related genes that play a role in the occurrence and development of GC and constructed a lactylation score model. Lactylation score in GC was closely associated with tumor mutational load, genomic instability, response to ICI treatment, immune cell infiltration, and immune evasion. These findings provide novel ideas for the diagnosis and treatment of GC, and lactylation-related genes may become novel tumor markers or therapeutic targets. Although our paper has many strengths, it also has limitations. For example, further studies are needed to reveal the pathways through which lactylation-related genes affect immune cell infiltration and genomic instability in GC. The accuracy of the lactylation score in predicting the response to ICI treatment by GC still needs to be verified by large-scale clinical trials.

In summary, the lactylation score might be useful in the molecular classification of GC, as it could help to identify different patterns of immune infiltration and genomic instability. The lactylation score can also be used as a method to assess the response of patients to ICI treatment.

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 within the article/Supplementary Materials.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of the second affiliated hospital of Harbin Medical University(sydwgzr2021-161). The patients/participants provided their written informed consent to participate in this study.

Author contributions

HY: Bioinformatics analysis, manuscript writing, vitro assays. XZ: Conceptual design, manuscript revision. SY: Prepare tissue and conduct Immunohistochemistry, graphical visualization. AZ: Experimental design and supervision. NL: Experimetal design, manuscript editing. ZM: Conceptual design, bioinformatics analysis, manuscript revision. All authors contributed to the article and approved the submitted version.

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.2023.1149989/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Drug sensitivity in the high and low lacylation score groups.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel R, Torre L, 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(6):394–424. doi: 10.3322/caac.21492

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* (2016) 66 (1):7–30. doi: 10.3322/caac.21332

3. 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(10125):1023–75. doi: 10.1016/S0140-6736(17)33326-3

4. Feng F, Tian Y, Xu G, Liu Z, Liu S, Zheng G, et al. Diagnostic and prognostic value of CEA, CA19-9, AFP and CA125 for early gastric cancer. *BMC cancer.* (2017) 17 (1):737. doi: 10.1186/s12885-017-3738-y

5. Xiao S, Zhou L. Gastric cancer: Metabolic and metabolomics perspectives (Review). *Int J Oncol* (2017) 51(1):5–17. doi: 10.3892/ijo.2017.4000

6. Abbassi-Ghadi N, Kumar S, Huang J, Goldin R, Takats Z, Hanna G. Metabolomic profiling of oesophago-gastric cancer: A systematic review. *Eur J Cancer* (2013) 49 (17):3625–37. doi: 10.1016/j.ejca.2013.07.004

7. Hanahan D, Weinberg R. Hallmarks of cancer: The next generation. *Cell.* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013

8. Nagarajan A, Malvi P, Wajapeyee N. Oncogene-directed alterations in cancer cell metabolism. *Trends cancer.* (2016) 2(7):365–77. doi: 10.1016/j.trecan.2016.06.002

9. Zhang L, Li S. Lactic acid promotes macrophage polarization through MCT-HIF1α signaling in gastric cancer. *Exp Cell Res* (2020) 388(2):111846. doi: 10.1016/j.yexcr.2020.111846

10. Ippolito L, Morandi A, Giannoni E, Chiarugi P. Lactate: A metabolic driver in the tumour landscape. *Trends Biochem Sci* (2019) 44(2):153–66. doi: 10.1016/j.tibs.2018.10.011

11. Cai Z, Li C, Han F, Liu C, Zhang A, Hsu C, et al. Phosphorylation of PDHA by AMPK drives TCA cycle to promote cancer metastasis. *Mol Cell* (2020) 80(2):263–78.e7. doi: 10.1016/j.molcel.2020.09.018

12. Zhang D, Tang Z, Huang H, Zhou G, Cui C, Weng Y, et al. Metabolic regulation of gene expression by histone lactylation. *Nature.* (2019) 574(7779):575–80. doi: 10.1038/s41586-019-1678-1

13. Li L, Chen K, Wang T, Wu Y, Xing G, Chen M, et al. Glis1 facilitates induction of pluripotency *via* an epigenome-metabolome-epigenome signalling cascade. *Nat Metab* (2020) 2(9):882–92. doi: 10.1038/s42255-020-0267-9

14. Irizarry-Caro R, McDaniel M, Overcast G, Jain V, Troutman T, Pasare C. TLR signaling adapter BCAP regulates inflammatory to reparatory macrophage transition by promoting histone lactylation. *Proc Natl Acad Sci United States America.* (2020) 117 (48):30628–38. doi: 10.1073/pnas.2009778117

15. Hagihara H, Shoji H, Otabi H, Toyoda A, Katoh K, Namihira M, et al. Protein lactylation induced by neural excitation. *Cell Rep* (2021) 37(2):109820. doi: 10.1016/j.celrep.2021.109820

16. Yu J, Chai P, Xie M, Ge S, Ruan J, Fan X, et al. Histone lactylation drives oncogenesis by facilitating mA reader protein YTHDF2 expression in ocular melanoma. *Genome Biol* (2021) 22(1):85. doi: 10.1186/s13059-021-02308-z

17. Wang Z, Wang X, Zhang N, Zhang H, Dai Z, Zhang M, et al. Pentraxin 3 promotes glioblastoma progression by negative regulating cells autophagy. *Front Cell Dev Biol* (2020) 8:795. doi: 10.3389/fcell.2020.00795

18. Wu X, Su D, Mei Y, Xu M, Zhang H, Wang Z, et al. Identified lung adenocarcinoma metabolic phenotypes and their association with tumor immune microenvironment. *Cancer immunology immunotherapy CII.* (2021) 70(10):2835–50. doi: 10.1007/s00262-021-02896-6

19. Luo Z, Rong Z, Zhang J, Zhu Z, Yu Z, Li T, et al. Circular RNA circCCDC9 acts as a miR-6792-3p sponge to suppress the progression of gastric cancer through regulating CAV1 expression. *Mol cancer.* (2020) 19(1):86. doi: 10.1186/s12943-020-01203-8

20. Ding L, Yin Y, Han L, Li Y, Zhao J, Zhang W. TSC1-mTOR signaling determines the differentiation of islet cells. *J endocrinology.* (2017) 232(1):59–70. doi: 10.1530/JOE-16-0276

21. Burgermeister E, Xing X, Röcken C, Juhasz M, Chen J, Hiber M, et al. Differential expression and function of caveolin-1 in human gastric cancer progression. *Cancer Res* (2007) 67(18):8519–26. doi: 10.1158/0008-5472.CAN-07-1125

22. Zhan W, Liao X, Li L, Chen Z, Tian T, Yu L, et al. *In vitro* mitochondrial-targeted antioxidant peptide induces apoptosis in cancer cells. *OncoTargets Ther* (2019) 12:7297–306. doi: 10.2147/OTT.S207640

23. Zhang N, Zhang H, Wu W, Zhou R, Li S, Wang Z, et al. Machine learning-based identification of tumor-infiltrating immune cell-associated lncRNAs for improving outcomes and immunotherapy responses in patients with low-grade glioma. *Theranostics.* (2022) 12(13):5931–48. doi: 10.7150/thno.74281

24. Cai A, Landman K, Hughes B. Multi-scale modeling of a wound-healing cell migration assay. *J Theor Biol* (2007) 245(3):576–94. doi: 10.1016/j.jtbi.2006.10.024

25. Zhang H, Zhang N, Wu W, Zhou R, Li S, Wang Z, et al. Machine learning-based tumor-infiltrating immune cell-associated lncRNAs for predicting prognosis and immunotherapy response in patients with glioblastoma. *Briefings Bioinf* (2022) 23(6): bbac368. doi: 10.1093/bib/bbac386

26. 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(10):1550–8. doi: 10.1038/s41591-018-0136-1

27. Rabinowitz J, Enerbäck S. Lactate: the ugly duckling of energy metabolism. *Nat Metab* (2020) 2(7):566–71. doi: 10.1038/s42255-020-0243-4

28. de la Cruz-López K, Castro-Muñoz L, Reyes-Hernández D, García-Carrancá A, Manzo-Merino J. Lactate in the regulation of tumor microenvironment and therapeutic approaches. *Front Oncol* (2019) 9:1143. doi: 10.3389/fonc.2019.01143

29. Kumagai S, Koyama S, Itahashi K, Tanegashima T, Lin Y, Togashi Y, et al. Lactic acid promotes PD-1 expression in regulatory T cells in highly glycolytic tumor microenvironments. *Cancer Cell* (2022) 40(2):201–18.e9. doi: 10.1016/j.ccell.2022.01.001

30. Sun H, Zhu A, Zhou X, Wang F. Suppression of pyruvate dehydrogenase kinase-2 re-sensitizes paclitaxel-resistant human lung cancer cells to paclitaxel. *Oncotarget.* (2017) 8(32):52642–50. doi: 10.18632/oncotarget.16991

31. Renner K, Bruss C, Schnell A, Koehl G, Becker H, Fante M, et al. Restricting glycolysis preserves T cell effector functions and augments checkpoint therapy. *Cell Rep* (2019) 29(1):135–50.e9. doi: 10.1016/j.celrep.2019.08.068

32. Siska P, Singer K, Evert K, Renner K, Kreutz M. The immunological warburg effect: Can a metabolic-tumor-stroma score (MeTS) guide cancer immunotherapy? *Immunol Rev* (2020) 295(1):187–202. doi: 10.1111/imr.12846

33. Cheng C, Tan H, Wang N, Chen L, Meng Z, Chen Z, et al. Functional inhibition of lactate dehydrogenase suppresses pancreatic adenocarcinoma progression. *Clin Trans Med* (2021) 11(6):e467. doi: 10.1002/ctm2.467

34. Le A, Cooper C, Gouw A, Dinavahi R, Maitra A, Deck L, et al. Inhibition of lactate dehydrogenase a induces oxidative stress and inhibits tumor progression. *Proc Natl Acad Sci United States America.* (2010) 107(5):2037–42. doi: 10.1073/pnas.0914433107

35. Fan J, Hitosugi T, Chung T, Xie J, Ge Q, Gu T, et al. Tyrosine phosphorylation of lactate dehydrogenase a is important for NADH/NAD(+) redox homeostasis in cancer cells. *Mol Cell Biol* (2011) 31(24):4938–50. doi: 10.1128/MCB.06120-11

36. Wang C, Li Y, Yan S, Wang H, Shao X, Xiao M, et al. Interactome analysis reveals that lncRNA HULC promotes aerobic glycolysis through LDHA and PKM2. *Nat Commun* (2020) 11(1):3162. doi: 10.1038/s41467-020-16966-3

37. Qi C, Huang M, Zou Y, Yang R, Jiang Y, Sheng J, et al. The IRF2/CENP-N/AKT signaling axis promotes proliferation, cell cycling and apoptosis resistance in nasopharyngeal carcinoma cells by increasing aerobic glycolysis. *J Exp Clin Cancer Res CR.* (2021) 40(1):390. doi: 10.1186/s13046-021-02191-3

38. Feichtinger R, Lang R. Targeting l-lactate metabolism to overcome resistance to immune therapy of melanoma and other tumor entities. *J Oncol* (2019) 2019:2084195. doi: 10.1155/2019/2084195

39. Chen Y, Guo H, Terajima M, Banerjee P, Liu X, Yu J, et al. Lysyl hydroxylase 2 is secreted by tumor cells and can modify collagen in the extracellular space. *J Biol Chem* (2016) 291(50):25799–808. doi: 10.1074/jbc.M116.759803

40. Du H, Chen Y, Hou X, Huang Y, Wei X, Yu X, et al. PLOD2 regulated by transcription factor FOXA1 promotes metastasis in NSCLC. *Cell Death disease.* (2017) 8(10):e3143. doi: 10.1038/cddis.2017.553

41. Xu F, Zhang J, Hu G, Liu L, Liang W. Hypoxia and TGF-β1 induced PLOD2 expression improve the migration and invasion of cervical cancer cells by promoting epithelial-to-mesenchymal transition (EMT) and focal adhesion formation. *Cancer Cell Int* (2017) 17:54. doi: 10.1186/s12935-017-0420-z

42. Cheriyamundath S, Kumar A, Gavert N, Brabletz T, Ben-Ze'ev A. The collagen-modifying enzyme PLOD2 is induced and required during L1-mediated colon cancer progression. *Int J Mol Sci* (2021) 22(7):3552. doi: 10.3390/ijms22073552

43. Gilkes D, Bajpai S, Chaturvedi P, Wirtz D, Semenza G. Hypoxia-inducible factor 1 (HIF-1) promotes extracellular matrix remodeling under hypoxic conditions by inducing P4HA1, P4HA2, and PLOD2 expression in fibroblasts. *J Biol Chem* (2013) 288 (15):10819–29. doi: 10.1074/jbc.M112.442939

44. Kurozumi A, Kato M, Goto Y, Matsushita R, Nishikawa R, Okato A, et al. Regulation of the collagen cross-linking enzymes LOXL2 and PLOD2 by tumor-suppressive microRNA-26a/b in renal cell carcinoma. *Int J Oncol* (2016) 48(5):1837–46. doi: 10.3892/ijo.2016.3440

45. Du H, Pang M, Hou X, Yuan S, Sun L. PLOD2 in cancer research. *Biomedicine pharmacotherapy = Biomedecine pharmacotherapie.* (2017) 90:670–6. doi: 10.1016/j.biopha.2017.04.023

46. Gjaltema R, de Rond S, Rots M, Bank R. Procollagen lysyl hydroxylase 2 expression is regulated by an alternative downstream transforming growth factor β-1 activation mechanism. *J Biol Chem* (2015) 290(47):28465–76. doi: 10.1074/jbc.M114.634311

47. Remst D, Blaney Davidson E, Vitters E, Bank R, van den Berg W, van der Kraan P. TGF-ß induces lysyl hydroxylase 2b in human synovial osteoarthritic fibroblasts through ALK5 signaling. *Cell Tissue Res* (2014) 355(1):163–71. doi: 10.1007/s00441-013-1740-5

48. Mia M, Bank R. The IκB kinase inhibitor ACHP strongly attenuates TGFβ1-induced myofibroblast formation and collagen synthesis. *J Cell Mol Med* (2015) 19 (12):2780–92. doi: 10.1111/jcmm.12661

49. Watanabe M, Abe N, Oshikiri Y, Stanbridge E, Kitagawa T. Selective growth inhibition by glycogen synthase kinase-3 inhibitors in tumorigenic HeLa hybrid cells is mediated through NF-κB-dependent GLUT3 expression. *Oncogenesis.* (2012) 1:e21. doi: 10.1038/oncsis.2012.21

50. Masin M, Vazquez J, Rossi S, Groeneveld S, Samson N, Schwalie P, et al. GLUT3 is induced during epithelial-mesenchymal transition and promotes tumor cell proliferation in non-small cell lung cancer. *Cancer Metab* (2014) 2:11. doi: 10.1186/2049-3002-2-11

51. Deng D, Sun P, Yan C, Ke M, Jiang X, Xiong L, et al. Molecular basis of ligand recognition and transport by glucose transporters. *Nature.* (2015) 526(7573):391–6. doi: 10.1038/nature14655

52. Kawauchi K, Araki K, Tobiume K, Tanaka N. p53 regulates glucose metabolism through an IKK-NF-kappaB pathway and inhibits cell transformation. *Nat Cell Biol* (2008) 10(5):611–8. doi: 10.1038/ncb1724

53. Flavahan W, Wu Q, Hitomi M, Rahim N, Kim Y, Sloan A, et al. Brain tumor initiating cells adapt to restricted nutrition through preferential glucose uptake. *Nat Neurosci* (2013) 16(10):1373–82. doi: 10.1038/nn.3510

54. Ha T, Her N, Lee M, Ryu B, Lee J, Han J, et al. Caveolin-1 increases aerobic glycolysis in colorectal cancers by stimulating HMGA1-mediated GLUT3 transcription. *Cancer Res* (2012) 72(16):4097–109. doi: 10.1158/0008-5472.CAN-12-0448

55. Liu X, Haines J, Mehanna E, Genet M, Ben-Sahra I, Asara J, et al. ZBTB7A acts as a tumor suppressor through the transcriptional repression of glycolysis. *Genes Dev* (2014) 28(17):1917–28. doi: 10.1101/gad.245910.114

56. Gan L, Huang S, Hu Y, Zhang J, Wang X. Heat treatment reduced the expression of miR-7-5p to facilitate insulin-stimulated lactate secretion by targeting IRS2 in boar sertoli cells. *Theriogenology.* (2022) 180:161–70. doi: 10.1016/j.theriogenology.2021.12.029

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 16 March 2023
DOI 10.3389/fonc.2023.1156455

OPEN ACCESS

EDITED BY
Nan Zhang,
Harbin Medical University, China

REVIEWED BY
Zeyu Wang,
University of Edinburgh, United Kingdom
Zhijia Xia,
Ludwig Maximilian University of Munich, Germany

*CORRESPONDENCE
Ting Cai
ct0702@csu.edu.cn
Song Wu
xy3ws1969@hotmail.com

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Oncology

RECEIVED 01 February 2023
ACCEPTED 28 February 2023
PUBLISHED 16 March 2023

CITATION
Hu H, Yin Y, Jiang B, Feng Z, Cai T and Wu S (2023) Cuproptosis signature and PLCD3 predicts immune infiltration and drug responses in osteosarcoma. *Front. Oncol.* 13:1156455. doi: 10.3389/fonc.2023.1156455

Cuproptosis signature and PLCD3 predicts immune infiltration and drug responses in osteosarcoma

Hai Hu[1], Yuesong Yin[1], Binbin Jiang[1], Zhennan Feng[1], Ting Cai[2]* and Song Wu[1]*

[1]Department of Orthopedics, The Third Xiangya Hospital, Central South University, Changsha, China, [2]Department of Gastroenterology, The Third Xiangya Hospital, Central South University, Changsha, China

Osteosarcoma (OS) is a cancer that is frequently found in children and adolescents and has made little improvement in terms of prognosis in recent years. A recently discovered type of programmed cell death called cuproptosis is mediated by copper ions and the tricarboxylic acid (TCA) cycle. The expression patterns, roles, and prognostic and predictive capabilities of the cuproptosis regulating genes were investigated in this work. TARGET and GEO provided transcriptional profiling of OS. To find different cuproptosis gene expression patterns, consensus clustering was used. To identify hub genes linked to cuproptosis, differential expression (DE) and weighted gene co-expression network analysis (WGCNA) were used. Cox regression and Random Survival Forest were used to build an evaluation model for prognosis. For various clusters/subgroups, GSVA, mRNAsi, and other immune infiltration experiments were carried out. The drug-responsive study was carried out by the Oncopredict algorithm. Cuproptosis genes displayed two unique patterns of expression, and high expression of FDX1 was associated with a poor outcome in OS patients. The TCA cycle and other tumor-promoting pathways were validated by the functional study, and activation of the cuproptosis genes may also be connected with immunosuppressive state. The robust survival prediction ability of a five-gene prognostic model was verified. This rating method also took stemness and immunosuppressive characteristics into account. Additionally, it can be associated with a higher sensitivity to medications that block PI3K/AKT/mTOR signaling as well as numerous chemoresistances. U2OS cell migration and proliferation may be encouraged by PLCD3. The relevance of PLCD3 in immunotherapy prediction was verified. The prognostic significance, expressing patterns, and functions of cuproptosis in OS were revealed in this work on a preliminary basis. The cuproptosis-related scoring model worked well for predicting prognosis and chemoresistance.

KEYWORDS

osteosarcoma, cuproptosis, immune infiltration, tumor microenvironment, drug response, machine learning

1 Introduction

Osteosarcoma (OS) continues to be the most prevalent primary bone cancer in children and adolescents, although being uncommon globally (1). With 4.4 instances per million people in the US, this tumor reaches its peak incidence in adolescence, which is consistent with a pubertal growth surge (2). Patients with OS have a >60% five-year survival rate thanks to the present conventional therapeutic approach of surgery and chemotherapy, but since 1980, little has been learned about the pathophysiology and targeted therapy of OS. Patients with metastatic disease and relapse cannot benefit from additional surgery or chemoradiotherapy (3). In-depth research on novel etiology and treatment targets for OS is urgently needed given the non-negligible severe socioeconomic burden on young people.

The tailored treatment of OS may greatly benefit from further study into programmed cell death (PCD), which is still a hot topic in oncology. For instance, cisplatin, a traditional first-line chemotherapeutic treatment for OS, induces apoptosis (4). By inducing oxidative stress dependent on GSH depletion and ROS overproduction, ferroptosis promoters such as phenethyl isothiocyanate (PEITC), baldachin, and ursolic acid have been identified as potential adjuvant chemotherapy treatments (5–7). Similar to this, inhibiting RIP1- and RIP3-dependent necroptosis effectively reduced lung metastasis and tumor growth in an OS mouse model (8).

A new PCD variant called cuproptosis was published in March 2022 by Peter T et al. (9). The buildup of monovalent copper ions may interact directly with proteins that have been lipoylated, which are mostly found in the mitochondria that power the TCA cycle. The loss of proteins containing the Fe-S cluster and the production of acute proteotoxic stress as a result of copper chelating lipoylated protein aggregation led to an independent type of cell death. For oncology researchers interested in copper toxicity in the treatment of cancer, this result is encouraging. A significant anti-tumor effect in patients with low plasma lactate dehydrogenase (LDH) was revealed in the phase 3 clinical trial to apply copper ionophores for melanoma, suggesting malignancies with a high dependence on mitochondrial metabolism were likely to benefit from cuproptosis-related molecular therapies (10).

The metabolic reprogramming in OS (11) is characterized by abnormally suppressed TCA cycle and high levels of oxidized glutathione (GSH), and GSH regulates copper ion cytotoxicity by inhibiting the oxidation of divalent copper ions to monovalent copper ions (9). It is important to talk about the activities of the lipoylation and cuproptosis pathways. In this investigation, we seek to identify functional pathways and genetic targets closely associated with cuproptosis, investigate the expression patterns of cuproptosis regulatory genes, and assess the influence of these targets on the prognosis of OS patients. Additionally, immune infiltrates landscapes, chemotherapeutic responsiveness, and cancer stem-like cellular features are also implicated in identifying their distinctions in patients with various cuproptosis patterns. This study might offer initial recommendations and a feasibility analysis for treatment plans that aim to treat copper toxicity in OS patients.

2 Materials and methods

2.1 Dataset obtaining and processing

In this investigation, public transcriptional profiling datasets from OS patients were used, including the TARGET OS dataset and the GSE21257 dataset from GEO. For the TARGET OS dataset, the GDC portal (https://portal.gdc.cancer.gov/), along with complete clinical information and RNA expression data in raw count and TPM format, were downloaded by GDC client. Expression data in TPM format was then converted into a log2(TPM+1) matrix for further analysis, and 85 samples with full RNA expression and clinical data were finally included. The URL for GSE21257 was https://www.ncbi.nlm.nih.gov/geo/. 53 samples with complete information were eventually included after starting with raw data and moving on to obtain a normalized expression matrix and clinical data using the R package beadarray and illuminaHumanv2.db. R (version 4.1.3) and Bioconductor programs for data cleaning and gene analysis were used to analyze all the aforementioned data signature annotation.

2.2 Cuproptosis regulatory gene set and unsupervised consensus clustering

The cuproptosis regulatory gene set was obtained from the latest literature by Peter T et al. (9), including *FDX1*, *LIAS*, *LIPT1*, *DLD*, *DLAT*, *PDHA1*, *PDHB*, *MTF1*, *GLS* and *CDKN2A*. The sweep() function in R was used to normalize the expression matrix for these genes in log2(TPM+1) format before package ConsensusClusterPlus was used for unsupervised clustering. The study's parameters were maxK = 4, reps = 500, pItem = 0,8, pFeature = 1, title = title, clusterAlg = hc, and distance = canberra. Each clustering was evaluated using the consensus CDF value and CDF curve delta area.

2.3 Differential expressing analysis

The TARGET OS dataset's expression data in raw count format and the R package DEseq2 were used for the DE analysis. Briefly, grouping information was first established using results from previous clustering; next, the entire expression matrix in TARGET OS was pre-screened to remove genes with zero expression in more than 20% of samples; finally, a DEseqDataSet object was built; the DESeq() function was used to calculate DE fold change and perform a significance test. FDR 0.05 was the cutoff for identifying genes as significantly differentially expressed (DE), and these genes were referred to as cuproptosis-related DE genes (CRDEGs).

2.4 Weight gene correlation network analysis and identification of cuproptosis-related hub genes

To find additional genes connected to cuproptosis clustering, WGCNA was carried out using DE genes. Hierarchical clustering

analysis was first performed using the hclust tool. Then, the pickSoftThreshold duty during module construction screened the soft thresholding power setting (6 in this study). Various modules' average connectivity degrees and independence were tested using candidate power (1 to 30). A suitable power value was chosen if the autonomy level was greater than 0.8. Co-expression networks (modules) were built using the WGCNA R package (The R package WGCNA is a collection of functions for calculating various weighted association analyses, which can be used for network construction, gene screening, gene cluster identification, topological feature calculation, data simulation, and visualization). The minimum module size was set to 30, giving each module a distinct color label. On the basis of its correlation with clusters, the core module was chosen. Genes in the core module with GS values greater than 0.8 and Module Membership (MM)>0.5 was defined as hub genes, termed cuproptosis-related hub genes (CRHGs).

2.5 Construction and validation of the cuproptosis-related prognostic 0 0predicting model

Based on the aforementioned CRHGs, a Random Survival Forest (RSF) plus Cox regression algorithmic technique was used for the selection of predictive features, model development, and internal and external validation. Details are as follows:

2.5.1 Univariate Cox regression for preliminary feature screening

TARGET_OS dataset was first randomly divided into the train (70%) and internal test (30%) datasets by createDataPartition() function in the R package caret. Univariate Cox regression analysis was then applied for all CRHGs by R package survival and survminer. Given the low sample volume for the TARGET_OS dataset, a bootstrap (12) sampling strategy was adopted: in 1,000 replicates of sampling with replacement, a gene was proved as prognosis-related only when the univariate cox regression results showed FDR< 0.05 for more than 900 times; this step was accomplished by sample() function in R.

2.5.2 RSF model for prognostic genes selection

The randomforestSRC (13) R package's rfsrc() function was used to access the remaining genes in order to build an RSF model. The optimal values were ntree=1000, block.size=1, mtry=2, nodesize=13, splitrule="logrank" after adjustments. The var.select () function was used to choose features based on minimal depth in order to build the final prognostic model. As cuproptosis-related prognostic genes, these genes (CRPGs).

2.5.3 Prognostic model construction by multivariate Cox regression

A multivariate Cox regression model was built based on CRPGs. Coefficients in this regression were applied for a final cuproptosis-related prognostic scoring (CRP score) model calculated as follows:

$$CRP \quad score = \sum_{i=1}^{n} \mathrm{Coef}_i * \mathrm{x}_i$$

Where Coefi was the coefficient of multivariate Cox regression and xi was the log2(TPM+1) expression value corresponding to the No.i CRPG.

2.5.4 Model validation

Then, for patients in TARGET OS for train and internal tests and GSE21257 for external validation, the CRP score was determined. The timeROC package represented the Time-dependent ROC curve, and the area under the curve (AUC) was used as the foundation for evaluating the CRP score model's ability to predict outcomes.

2.6 Gene set variation analysis

GSVA was used by the R package GSVA and GSEAbase to investigate various enrichment statuses in gene function for distinct clusters and subgroups. Two gene sets, c2.cp.kegg.v7.4.symbols and h.all.v7.4.symbols, were used for functional annotation from MsigDB (http://www.gsea-msigdb.org/gsea/msigdb/). After that, the LIMMA package was used to identify the enrichment variations between various subgroups (13–15).

2.7 Calculation of the stemness index (mRNAsi)

Based on the mean-centered gene expression profiles of PSCs in the PCBC database (syn2701943), the stemness signature was derived *via* the one-class logistic regression (OCLR) machine learning algorithm (16), which was also verified by leave-one-out cross-validation. Then, we calculated the Spearman correlations between the normalized expression matrix of OS samples and the stemness signature. Eventually, the stemness index was identified by scaling the Spearman correlation coefficients to be between 0 and 1. The higher the mRNAsi, the greater the tumor dedifferentiation and higher stemness (17).

2.8 Compound resistance and sensitivity analysis

Genomics of Drug Sensitivity in Cancer (GDSC, http://www.cancerrxgene.org/downloads) database (18), which contained drug sensitivity data (IC50) of 1,000 cell lines, was accessed to get drug sensitivity and resistance information for osteosarcoma cell lines. Then R package Oncopredict (19) based on the Ridge Regression algorithm was applied to predict the drug response of samples in the TARGET_OS cohort. Spearman correlation analysis was performed to calculate the correlation between drug sensitivity and CRP_Score. The absolute value of correlation coefficient > 0.4 and FDR< 0.05 were regarded as significant.

2.9 Analysis for immune cell infiltration and immune signatures

Following the usual analysis procedure, we first used the ESTIMATE program in R to evaluate the stromal purity and general immune infiltration of tumor samples. For the investigation of tumor immune cell infiltration, we used the algorithms CIBERSORT and ssGSEA (13, 20, 21). The original publications' archives with the defining gene signatures for each type of immune cell were obtained.

2.10 q-PCR experiment

The primers used for q-PCR are as follows: β-actin (https://www.ncbi.nlm.nih.gov/gene/60; F ACCCTGAAGTACCCCATCGAG; R AGCACAGCCTGGATAGCAAC). PLCD3 (https://www.ncbi.nlm.nih.gov/gene/113026; F CTCATTCGGGAGGCAGGGAA; R CTGGGGACTGTAGTTGGCTG). The cell groups are as follows: NC, si-PLCD3-1, si-PLCD3-2, and si-PLCD3-3.

2.11 Transwell experiment

In DMEM with 10% FBS and 1% double antibody, U2OS cells were grown. Pancreatic enzymes were used to digest the U2OS cells at the logarithmic growth stage before being counted and distributed uniformly in six-well plates with roughly 1x105 cells per well. The cells were transfected with NC and si-PLCD3 on the second day. The cells were switched to a full medium for 48 hours after 6 hours, and they were then cultured in an EDU37°C incubator overnight. Paraformaldehyde was used to fix the samples after collection. To defrost on ice, remove the necessary si-PLCD3 and NC. They took four sterile tubes. A total of 95 uL of serum-free MEM/DMEM media was added to two tubes. The tubes were then filled with 5 uL of NC and 5 uL of Lip2000, respectively. The equivalent centrifuge tubes received the addition of si-PLCD3 in the same manner. Mix gently, then set aside for five minutes at room temperature. Following a 20-minute rest period at room temperature, combine the two tubes. Finally, the mixture was blended and added uniformly to the transfection hole. Replace with fresh and full culture medium six hours after starting the culture in incubators at 37°C. The following are the cell groups: Si-PLCD3-1 and Si-PLCD3-2, NC. Corning sold the Transwell chamber (3428), which was acquired.

2.12 EdU experiment

In DMEM with 10% FBS and 1% double antibody, U2OS cells were grown. Pancreatic enzymes were used to digest the U2OS cells at the logarithmic growth stage before being counted and distributed uniformly in six-well plates with roughly 1x105 cells per well. The cells were transfected with NC and si-PLCD3 on the second day. The cells were switched to a full medium for 48 hours after 6 hours, and they were then cultured in an EDU37°C incubator overnight. Paraformaldehyde was used to fix the samples after collection. To defrost on ice, remove the necessary si-PLCD3 and NC. They took four sterile tubes. A total of 95 uL of serum-free MEM/DMEM media was added to two tubes. The tubes were then filled with 5 uL of NC and 5 uL of Lip2000, respectively. The equivalent centrifuge tubes received the addition of si-PLCD3 in the same manner. Mix gently, then set aside for five minutes at room temperature. Following a 20-minute rest period at room temperature, combine the two tubes. Finally, the mixture was blended and added uniformly to the transfection hole. Replace with fresh and full culture medium six hours after starting the culture in incubators at 37°C. The following are the cell groups: Si-PLCD3-1 and Si-PLCD3-2, NC. Ribo supplied the EdU kit (RN: R11078.2).

2.13 Statistical analysis

All statistical calculations were done in R. (version 4.1.3). The comparison of count data was assessed using Fisher's test and the Chi-square test. The Student-t test was used for measurement data with a normal distribution, whereas the Wilcox test was used for data with an abnormal distribution. All correlation investigations must be completed using Spearman analysis. The Kaplan-Meier survival curve was represented using the R package survival and survminer.

3 Results

The workflow chart of the study is shown in Supplement Figure 1.

3.1 Distinct expression patterns for cuproptosis regulatory genes were identified in osteosarcoma patients

We initially examined the expression pattern of the genes that regulate cuprotosis based on the log2(TMP+1) expression matrix. All 10 genes were expressed in the TARGET OS and GSE21257 datasets, as seen in Figures 1A, B, and their expression followed a normal distribution, which gave us the foundation for further investigation. By combining the Kaplan-Meier survival curve and univariate Cox regression with survival data, we found that elevated FDX1 expression is linked to both lower overall survival (OvS) and disease-free survival (DFS) in OS patients (Figures 1C–E). These findings suggested that cuproptosis might contribute to the malignant biological activity of OS since FDX1 was shown to be a key regulator in cuproptosis by linking the cytotoxicity of cooper ions and protein lipoylation in the TCA cycle.

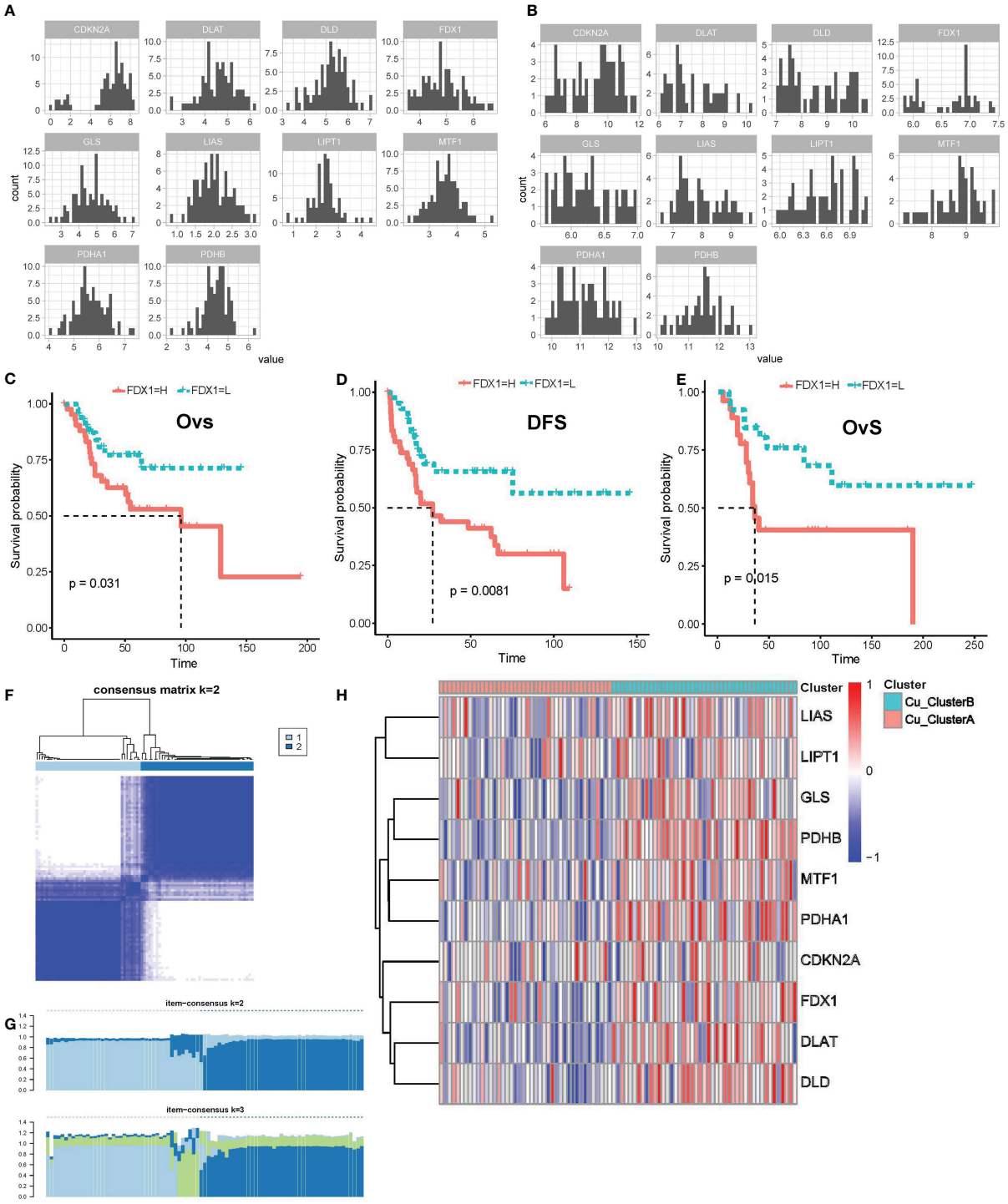


FIGURE 1
Cuproptosis regulatory genes were expressed in distinct patterns in OS samples. **(A, B)** Expression distributions of cuproptosis regulatory genes in TARGET_OS **(A)** and GSE21257 **(B)** datasets. **(C–E)** K-M survival curve for FDX1 high- and low- expression subgroups in TARGET_OS **(C, D)** and GSE21257 datasets **(E)**, Ovs, overall survival; DFS, disease-free survival. **(F, G)** Results of consensus clustering based on the expression of cuproptosis regulatory genes, **(F)** Consensus heatmap, **(G)** Item-Consensus plot. **(H)** Heatmap shows the expression of cuproptosis genes in distinct patterns.

Following that, unsupervised consensus clustering was carried out using the expression matrix of 10 genes involved in cuproptosis. Eighty-five samples in TARGET OS were best grouped into two clusters, referred to as Cu ClusterA (42 samples) and Cu ClusterB (43 samples), as shown in Figures 1F, G. The samples in Cu ClusterB tended to overexpress all cuproptosis genes, as seen in Figure 1H, whereas CDKN2A appeared to be indistinguishable. These results showed that the cuproptosis pathway genes' activity in OS patients showed two different patterns.

3.2 Cuproptosis clusters in OS patients represented differences in immune infiltration and stemness properties

Then, we pondered how the two clusters' malignant biological characteristics varied from one another. So, we carried out a number of functional investigations. The expression of cuproptosis genes was typically active in Cu ClusterB samples, according to GSVA analysis, which first revealed a number of pathways that were sparked. TCA cycle-related pathways (such as citrate metabolism, oxoglutarate metabolism, and pantothenic acid biosynthesis) and traditional cancer-promoting pathways (such as TGF-, WNT/-catenin, p53, and IL./STAT5 signalings) are two categories of important findings (Figure 2A). We developed the mRNAsi index to show the difference in cellular stemness between the two clusters because the majority of these enriched pathways were involved in the destiny control of cancer stem-like cells (CSLCs). Cu ClusterB displayed a substantially higher mRNAsi than Cu ClusterA, as illustrated in Figure 2B, indicating that enhanced cuproptosis gene expression may function as an initiating factor in immortal proliferation, quick metastasis, and chemo-resistance linked to CSLC activities.

In addition, we carried out a number of researches on immunological infiltration between two clusters. Figures 2C–E illustrates how the ESTIMATE approach revealed that samples in Cu ClusterB had lower stromal scores than Cu ClusterA, indicating that Cu ClusterB had fewer stromal components. Samples in Cu ClusterA tended to enhance activated immune cells, according to

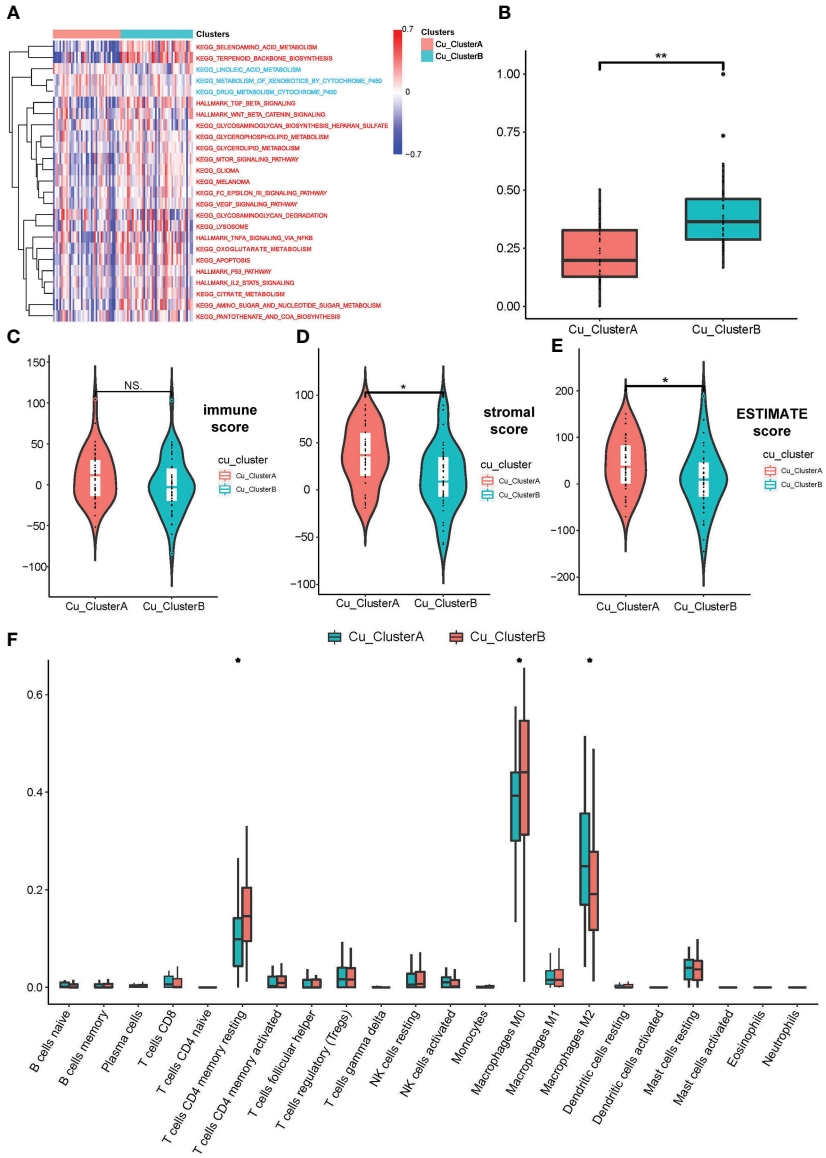


FIGURE 2
OS samples in different Cu_Clusters exhibited distinct tumor biological characteristics. **(A)** GSVA analysis showed diverse enriched pathways in different Cu_Clusters. **(B)** Divergence in the mRNAsi index showed differences in stemness properties between Cu_ClusterA and Cu_ClusteB. **(C–E)** ESTIMATE analysis for the overall status of immune cell infiltration and stromal component samples in the TARGET_OS dataset. **(F)** ssGSEA for the infiltration analysis of 29 types of immune cells in different Cu_Clusters *P < 0.05; **P < 0.01.

CIBERSORT and ssGSEA for a study of just one type of immune cell (e.g., Activated CD4 T cell, dendritic cell, and Macrophages M2). In contrast, models in Cu ClusterB (such as Regulatory T cell, MDSC, and Macrophages M0) may show signs of a dormant immunological state (Figure 2F and Supplement Figure 2A).

3.3 Screening of cuproptosis-related genes revealed a functional connection between cuproptosis and other biological processes in OS

We first performed a differential expression (DE) study to find CRGs associated with cuproptosis clusters. A total of 6537 genes, including 3565 up-regulated genes in Cu ClusterA and 2972 up-regulated genes in Cu ClusterB, matched the criteria for DE, as shown in Figures 3A, B. WGCNA analysis was used, using DEGs as input objects, to further narrow down the potential genes highly connected with cuproptosis clusters, and 16 modules were ultimately discovered (Figures 3C, D). Notably, the cuproptosis clusters had the strongest correlation with Module turquoise (MEturquoise), which had 1762 genes and had a R = 0.73 with Cu ClusterB, p = 1e-15, in Figure 3E. Further verification showed that genes in MEturquoise had strong consistency in principal component representation (shown by Module Membership, MM) and external connection with cuproptosis clusters (indicated by Module Membership, MM) (Figure 3F).

We eventually discovered 331 hub genes in MEturquoise based on the selection criteria of GS>0.8 and MM>0.5 described above. For further examination, these signatures were classified as CRGs. For CRGs, enrichment analysis was used to investigate the co-regulated pathways and biological processes. Notably, as shown in Figures 3G, H, the KEGG analysis revealed a high enrichment of the TCA cycle and NAD(P)+ activity pathways, further demonstrating the close relationship between cuproptosis and the TCA cycle. The terms RNA synthesis, metabolism & splicing, and ubiquitin-proteasome pathway also commonly appeared in search results. Traditional methods of controlling cuproptosis are suggested by AMPK and Hedgehog signaling. The emergence of the PD-1 checkpoint pathway suggested that cuproptosis might contribute to the responsiveness of tumor treatment.

3.4 Selection of cuproptosis-related prognostic genes and construction of cuproptosis-related prognostic score model

Patients in the TARGET OS cohort were 7:3 randomly split between the training and testing groups. 16 prognostic CRGs were left after performing univariate Cox regression with bootstrap sampling to reduce redundancy based on CRGs. Then, using a 1,000-tree random survival forest model, the minimum depth values selected the five gene signatures that would ultimately be used as CRPGs: BTBD10, DLX1, MRTFA, PLCD3, and RFX3 (Figures 4A, B). The scatter plot revealed no obvious association between the expression of these five genes, ruling out model redundancy in the process (Figure 4C). The CRP score model was then created by performing multivariate Cox regression using CRPGs:

$$CRP\quad score = (-1.8626130)*Exp_{BTBD10} + 0.2978399*Exp_{DLX1} - 0.9252084*Exp_{MRTFA} + 0.1514946*Exp_{PLCD3} + 1.0547832*Exp_{RFX3}$$

For each patient in a train, test, and external validation dataset, we computed a CRP score. In these datasets, we used a time-dependent ROC curve to find the prediction power for overall survival. The area under the curve (AUC) was convincingly confirmed in the TARGET testing set GSE21257 validation set and reached 0.809 at three years, 0.816 at five years, and 0.769 at eight years (Figures 4D–F). Based on the median score, OS patients were divided into CRP score high and CRP score low subgroups. The K-M curve further demonstrated that OS patients with higher CRP scores had considerably worse OvS times (Figures 4G, H).

3.5 Correlation analysis between CRP score and malignant biological behaviors

We also carried out a number of functional studies. First, GSVA analysis indicated that the CRP high subgroup was enriched for various cancer-promoting pathways, including Wnt/-catenin, TGF-, and JAK/STAT signaling, which overlapped with Cu ClusterB. Improvements were made to the TCA cycle-related pathways, demonstrating the coherence between Cu Clusters and CRP subgroups. Notably, the CRP high fraction also had activation of the epithelial-mesenchymal transition (EMT) pathway, suggesting that samples with poorer prognoses were more likely to develop distant metastases (Figure 5A).

Additionally, the mRNAsi index was used in connection studies with CRP results. In contrast to the CRP low subgroup, samples in the CRP high subgroup showed a considerably higher mRNAsi index (Figure 5B). Additionally, the TARGET OS dataset revealed a strong association between CRP score and mRNAsi in every person (R=0.32, p=0.031, Figure 5C), suggesting that OS samples with higher CRP values may have more pronounced stemness features.

According to the immune infiltration analysis, a higher CRP score was linked to immunosuppression (R = - 0.25, p = 0.023; Figure 5D); while a lower CRP score was linked to a greater stromal score (R = -0.31, p = 0.017; Figure 5E). The CRP high subgroup was related with higher infiltration of Macrophages M0, Type 17 T helper cells, and T cells, according to an examination of infiltration for various immune cells. The CRP low subgroup, on the other hand, was connected to enhanced infiltration of Macrophages M2, Regulatory T cell, Central memory CD8 T cell, and Activated B cell, showing different immune infiltration patterns in OS patients (Figure 5F and Supplement Figure 2B). It should be highlighted that the CRP low subgroup showed increased expression of PDL1, TIM3, and TIGIT (Supplement Figure 2C). Because immunosuppression and CRP score are correlated, anti-PD-1/PD-L1 immunotherapeutic medicines may be more effective for OS patients with lower CRP values.

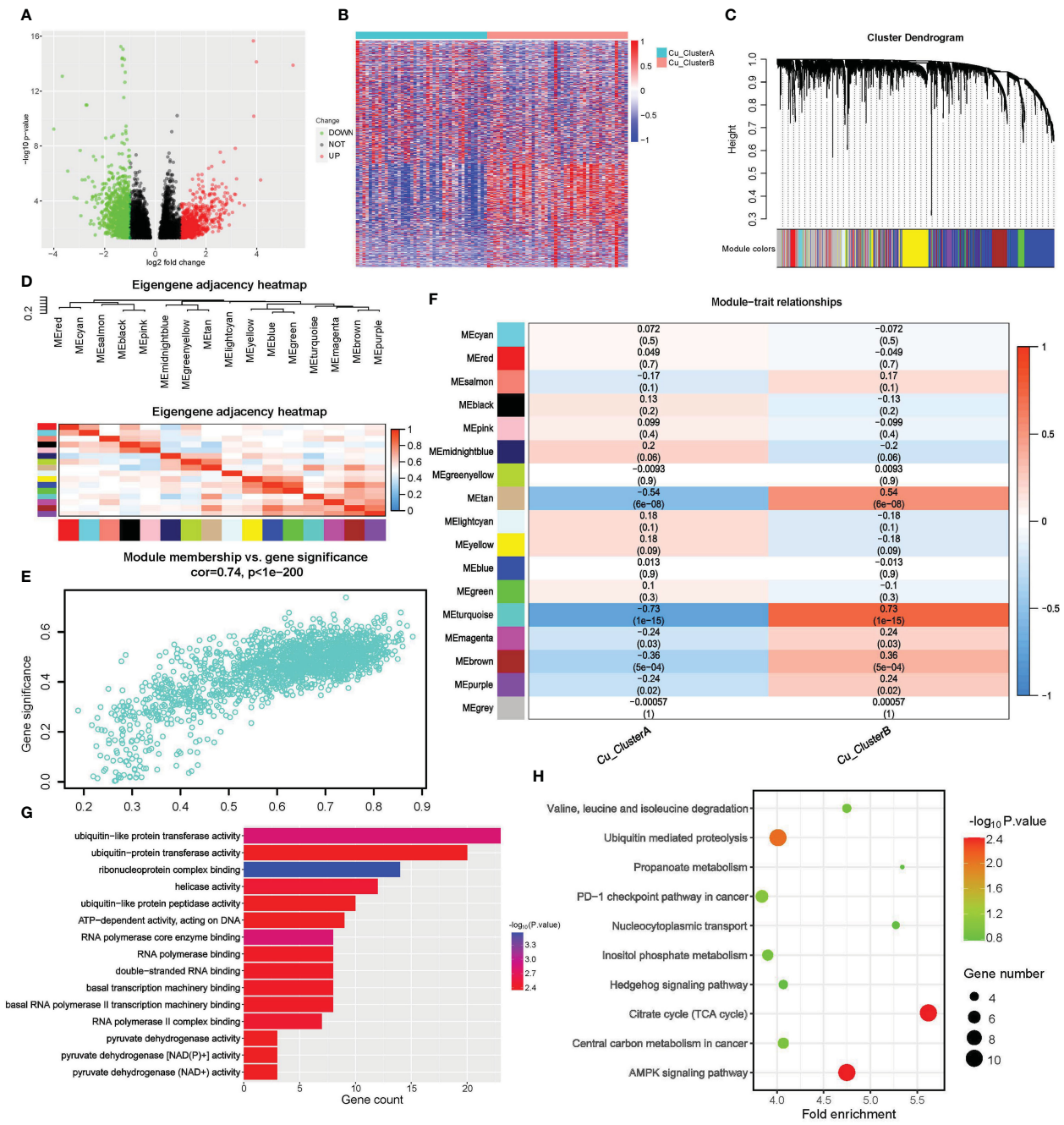


FIGURE 3
Screening of cuproptosis-related genes by DE analysis and WGCNA. **(A, B)** Results of DE analysis in different Cu_Cluseters; **(A)** Volcano plot of DEGs; **(B)** Heatmap of DEGs. **(C–F)** WGCNA analysis for DEGs to identify gene module that was most correlated with Cu_Clusters; **(G, H)** Pathway enrichment analysis for hub genes obtained from WGCNA; **(G)** Molecular function analysis of WGCNA hub genes in GO **(H)** Pathway enrichment analysis of WGCNA hub genes in KEGG.

3.6 Correlation analysis between CRP score and malignant biological behaviors

From CCLE, 10 OS cell lines and their expression matrix were taken. In order to determine the link between the CRP score and the IC50 for each molecule contained in the GDSC v2 database, we first calculated the CRP score for these cell lines. The elevated CRP score was linked to greater resistance to a number of medications, particularly those that target the ERK/MAPK pathway and cell cycle, as shown in Supplement Figure 3A. Unexpectedly, cell lines with higher CRP ratings appeared to be more responsive to AT13148, a medication that blocks PI3K/Akt/mTOR signaling. To further forecast the pharmacological reactions of samples in the TARGET OS dataset, we utilized a machine learning system. Two medications that target the PI3K/mTOR pathway, AZD6482 and AZD8055, were probably more sensitive in OS samples with

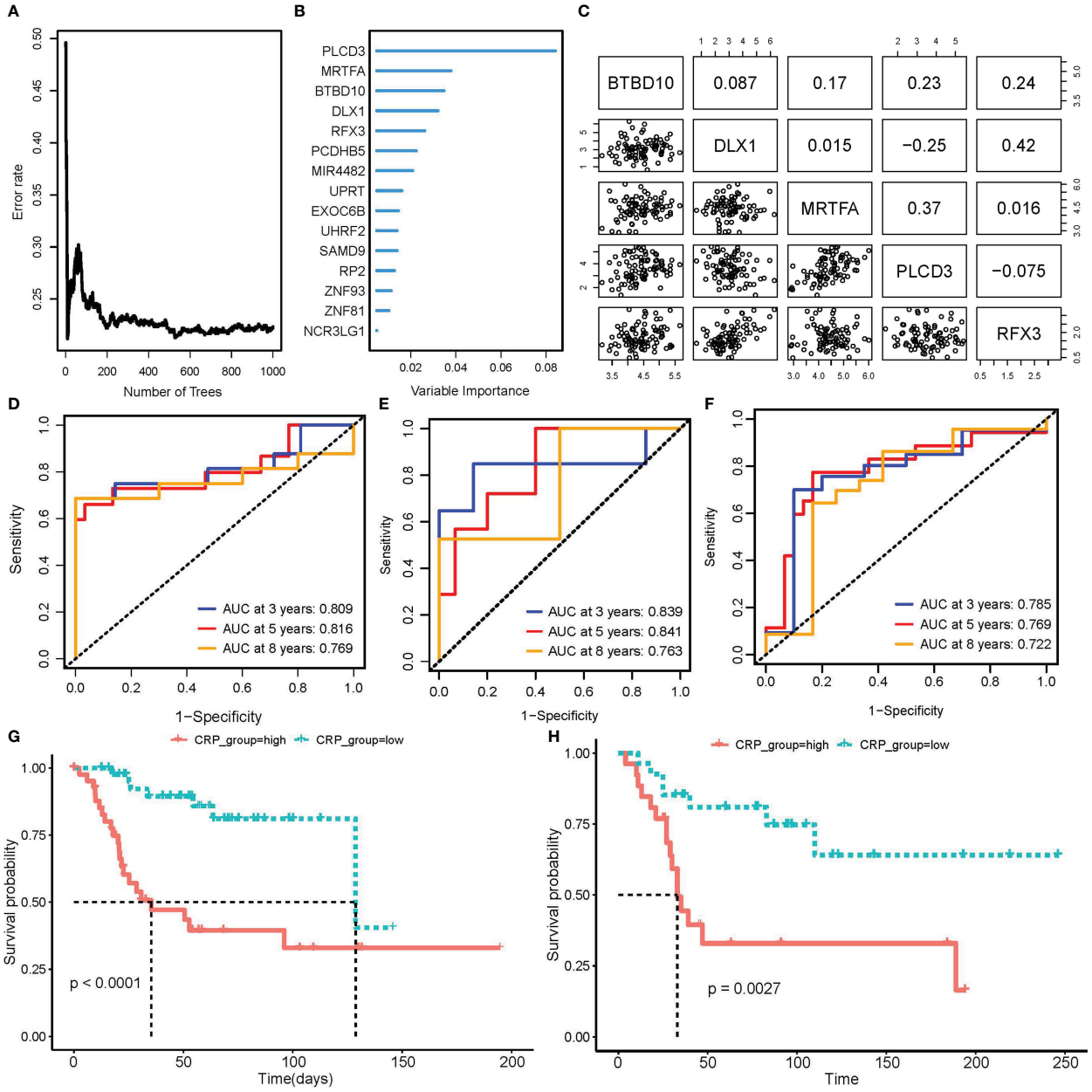


FIGURE 4
Construction and validation of CRP score model. **(A, B)** RSF model training and variable selection; **(A)** error rate trends as the number of trees increased when training RSF model; **(B)** Variable importance of selected features. **(C)** Correlation of expression in 5 genes that RSF selected to train CRP score model. **(D–F)** Time-dependent ROC curve to test the predictive ability of CRP score for OS patients in TARGET_OS train set **(D)**, test set **(E)**, as well as GSE21573 external validation set **(F)**. **(G, H)** K-M curve of CRP scores high and low subgroups for patients' overall survival in TARGET_OS **(G)** and GSE21573 **(H)** datasets.

higher CRP scores, as demonstrated in Supplement Figure 3B and Supplement Figure 2D. Additionally, samples with high CRP values responded more favorably to linsitinib targeting IGF1R.

In light of the aforementioned findings, treating patients with high CRP scores who were thought to have bad prognoses may involve targeting PI3K/Akt/mTOR signaling. Contrarily, drugs that target the cell cycle and Wnt signaling pathways are frequently ineffective against patients with high CRP values. Given that most first-line chemotherapeutics for OS used cell cycle inhibition as their primary mechanism of action, the CRP score model may also be able to predict clinical chemoresistance in OS patients.

3.7 Pan-cancer analysis on model genes

The expression pattern of model genes in pan-cancer is shown in Figure 6A. BTBD10, DLX1, MRTFA, PLCD3, and RFX3 were highly expressed in PRAD, COAD, LUSC, HNSC, and KIRC. The

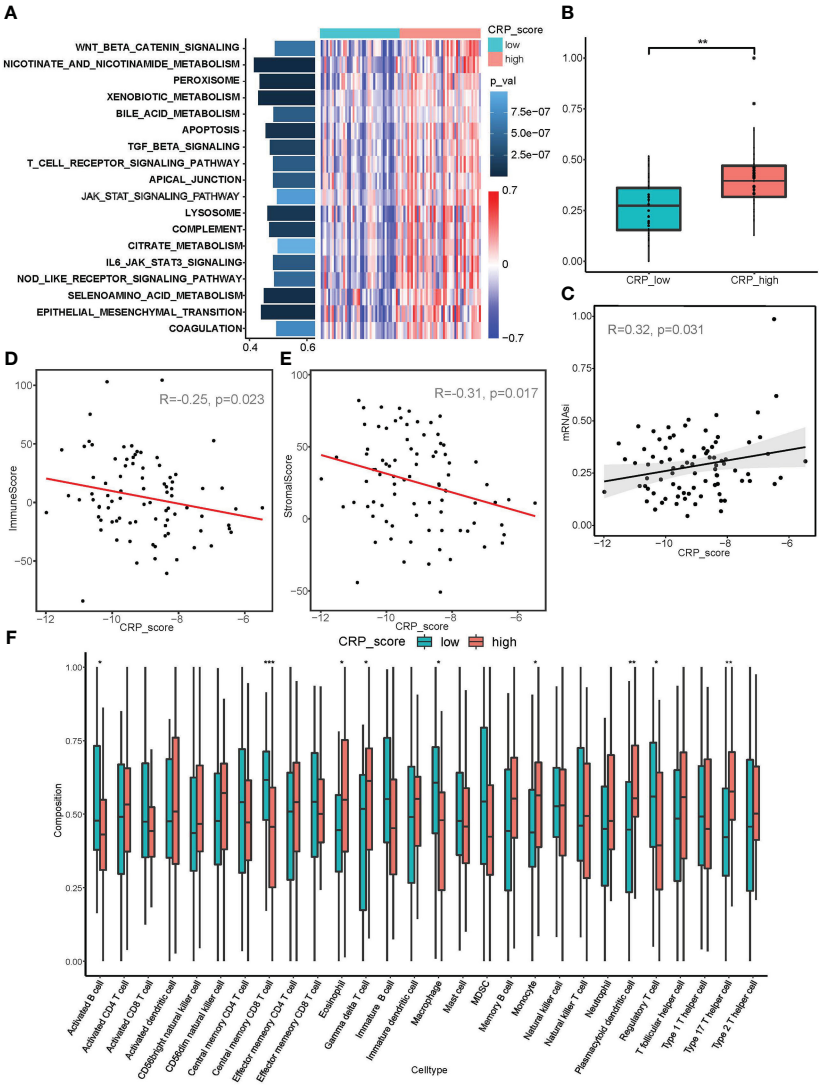


FIGURE 5
Correlation analysis between CRP score and malignant biological behaviors. **(A)** GSVA analysis showed diverse enriched pathways between CRP_high and CRP_low subgroups. **(B, C)** mRNAsi index analysis revealed differences in stemness properties between CRP_high and CRP_low subgroups and a significant correlation between CRP_score and mRNAsi. **(D, E)** ESTIMATE analysis for the correlation between CRP_score and immune cell infiltration as well as a stromal component in samples of the TARGET_OS dataset. **(F)** ssGSEA for the infiltration analysis of 29 types of immune cells in CRP_high and CRP_low subgroups *P < 0.05; **P < 0.01; ***P < 0.001.

somatic mutation frequency of model genes is shown in Figure 6B. BTBD10, DLX1, MRTFA, PLCD3, and RFX3 had relatively high mutation rates in UCEC and SKCM. The somatic mutation landscape of model genes is shown in Figure 6C. BTBD10 (43%), PLCD3 (27%), and RFX3 (24%) were frequently mutated in pan-cancer.

The heterozygous CNV profiles (amplification and depletion) of model genes are shown in Figure 7A. The homozygous CNV profiles (amplification and depletion) of model genes are shown in Figure 7B. Pathway analysis revealed that PLCD3 was related to activated apoptosis, EMT, hormone AR, hormone ER, PI3K/Akt, RAS/MAPK, RTK, and TSC/mTOR (Figure 7C). The miRNA regulation network of model genes is shown in Figure 7D.

3.8 *In vitro* validation on PLCD3

The tumor-promoting activity of PLCD3 was investigated by *in vitro* tests since it is a crucial gene in the CRP score. Three si-RNA significantly reduced the relative RNA expression of PLCD3 in the NC and three si-RNA groups, according to a q-PCR experiment (Figure 8A). Figure 8B displays the statistical analysis of the cell counts in the NC and two si-RNA groups using the Transwell test. Figure 8C illustrates the statistical analysis of the proliferation rate (EdU/DAPI) in the NC and two si-RNA groups. Transwell assay representative photos of the cell counts in the NC and two si-RNA groups (Figure 8D), showing that the number of migrated cells was dramatically decreased in the two si-RNA groups. Typical pictures

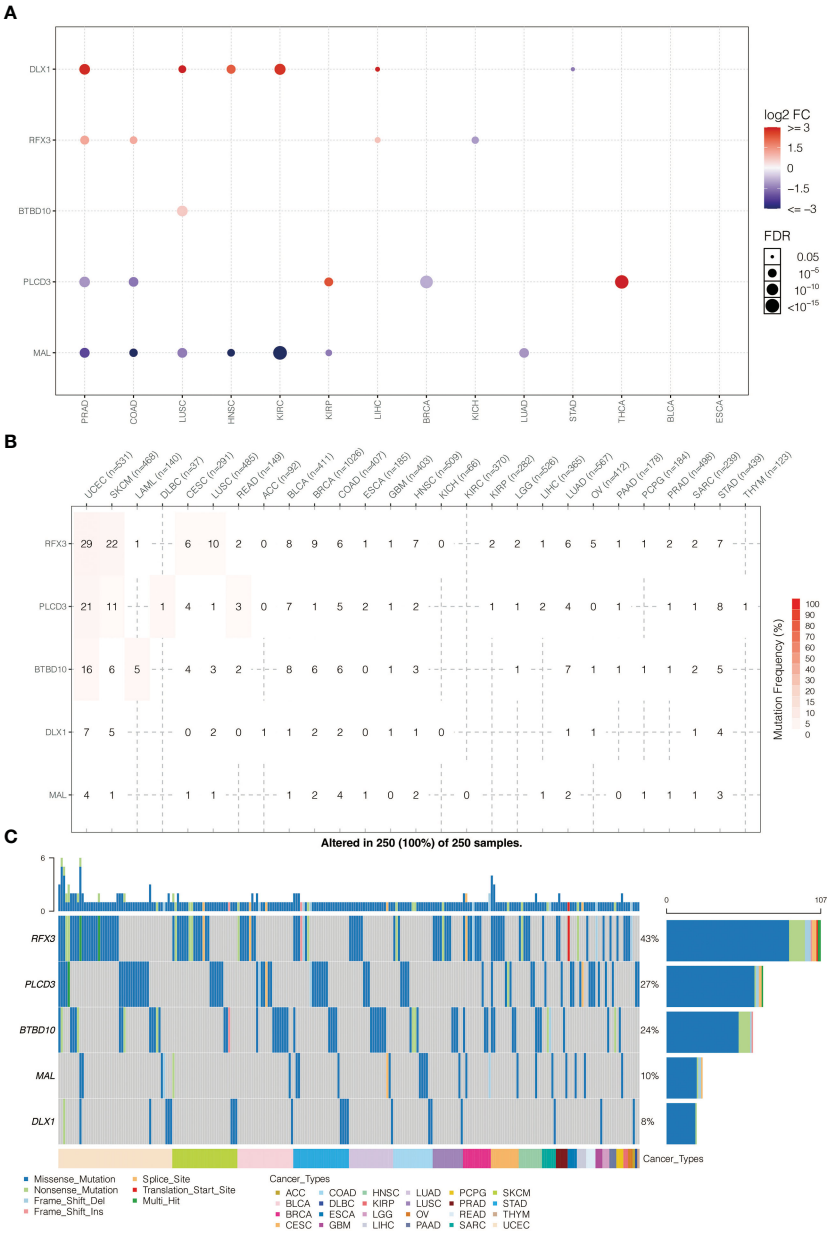


FIGURE 6
(A) Pan-cancer expression pattern of model genes. **(B)** Pan-cancer SNP analysis on model genes. **(C)** Pan-cancer SNP landscape on model genes.

of the proliferation rate in NC (EdU/DAPI). Examples of the proliferation rate (EdU/DAPI) in the NC and two si-RNA groups by EdU test are shown in Figure 8E, where the positively stained cells in the two si-RNA groups were dramatically decreased.

3.9 Immunotherapy prediction of PLCD3

Figure 9A depicts the expression of PLCD3 in immunotherapy cohorts of responders and non-responders, with responders exhibiting higher expression of PLCD3 in the Lauss cohort of 2017 and Kim cohort of 2019. Regarding the two groups' PLCD3 expression in immunotherapy cohorts, a survival analysis was carried out (Figure 9B). In the VanAllen cohort of 2015 and the Cho cohort of 2020, PLCD3 was linked to improved survival; in the Kim cohort of 2019, the Nathanson cohort of 2017, and the Lauss cohort of 2019, PLCD3 was linked to worse survival. In eight immunotherapy cohorts, PLCD3 demonstrated strong predictive power for immunotherapy response (Figure 9C).

Figure 10A illustrates the relationship between PLCD3 and T dysfunction value (core dataset), normalized Z score calling from Cox-PH regression (immunotherapy datasets), normalized Z score calling from selection log2FC (CRISPR screening datasets), and normalized expression value from immune-suppressive cell types. PLCD3 had an AUC greater than 0.5 in ten immunotherapy cohorts with regard to its predictive value (Figure 10B). In seven mouse cohorts, the cytokine treatment prediction revealed that PLCD3 could strongly predict the treatment with cytokines (Figure 10C). In

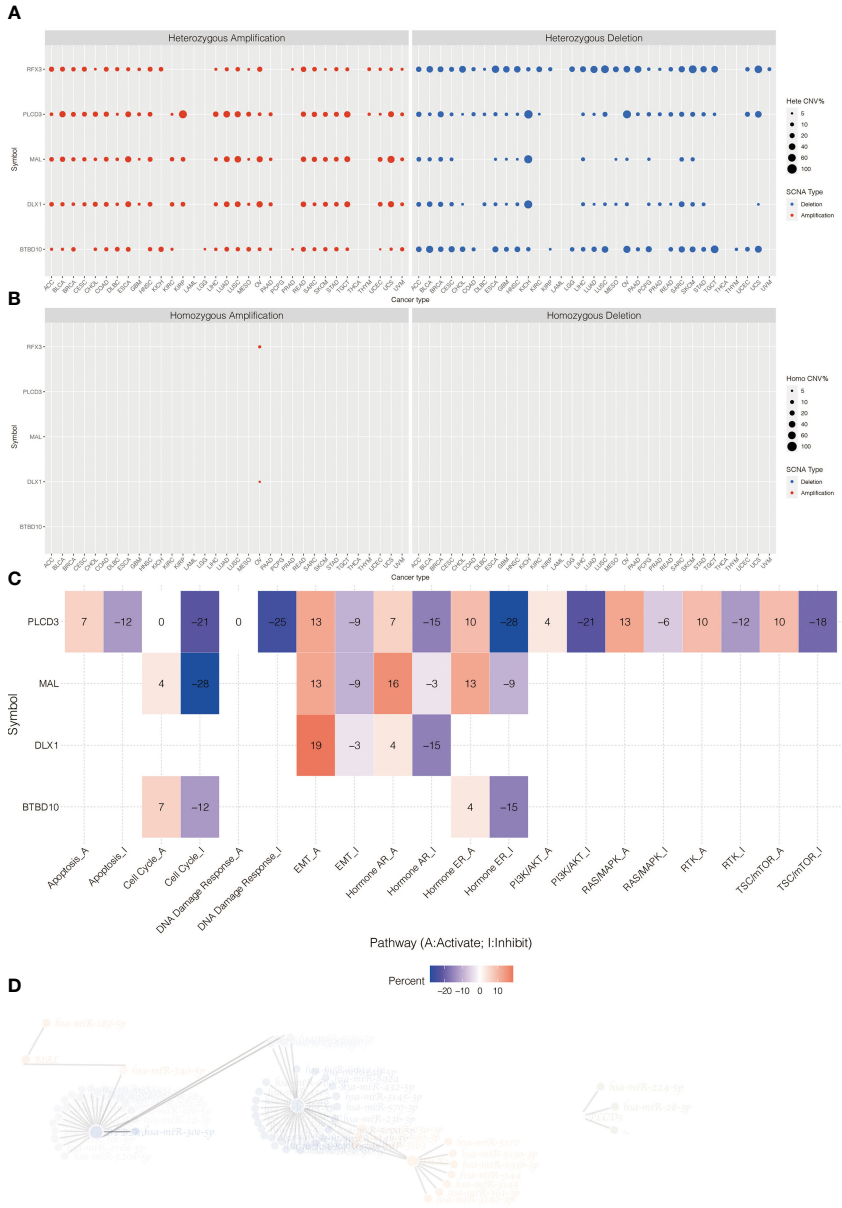


FIGURE 7
(A) The heterozygous CNV profiles (amplification and depletion) of model genes. **(B)** The homozygous CNV profiles (amplification and depletion) of model genes. **(C)** Pathway analysis related to PLCD3. **(D)** The miRNA regulation network of model genes.

two mouse cohorts, the immunotherapy prediction demonstrated that PLCD3 could accurately predict immunotherapy (Figure 10D).

3.10 Protein interaction network, illness network, and pan-cancer immune infiltration pattern of PLCD3

PLCD3 was found to interact with ITRP3, ITPR1, PRKCA, and PIP4K families by STRING (Figure 11A). PLCD3 was involved in hypertension, cutaneous melanoma, and breast adenocarcinoma by Open Targets Platform (Figure 11B). PLCD3 positively correlated with macrophages and negatively correlated with T cells in most cancers by TIMER (Figure 11C).

4 Discussion

Cuproptosis is a recently identified type of programmed cell death. Little is currently known about this unique, mitochondrial-dependent mechanism, however Peter T. et al. This work provides a preliminary description of the regulatory environment of cuproptosis-related pathways in osteosarcoma based on the available information. Some key findings may serve as an inspiration for work on OS and many other pathological conditions. First, it is found that OS patients have a poor prognosis and high FDX1 expression. FDX1 was first discovered as a mitochondrial electron transporter for cytochrome P450 metabolism (22, 23). Some sporadic studies identified FDX1 as a

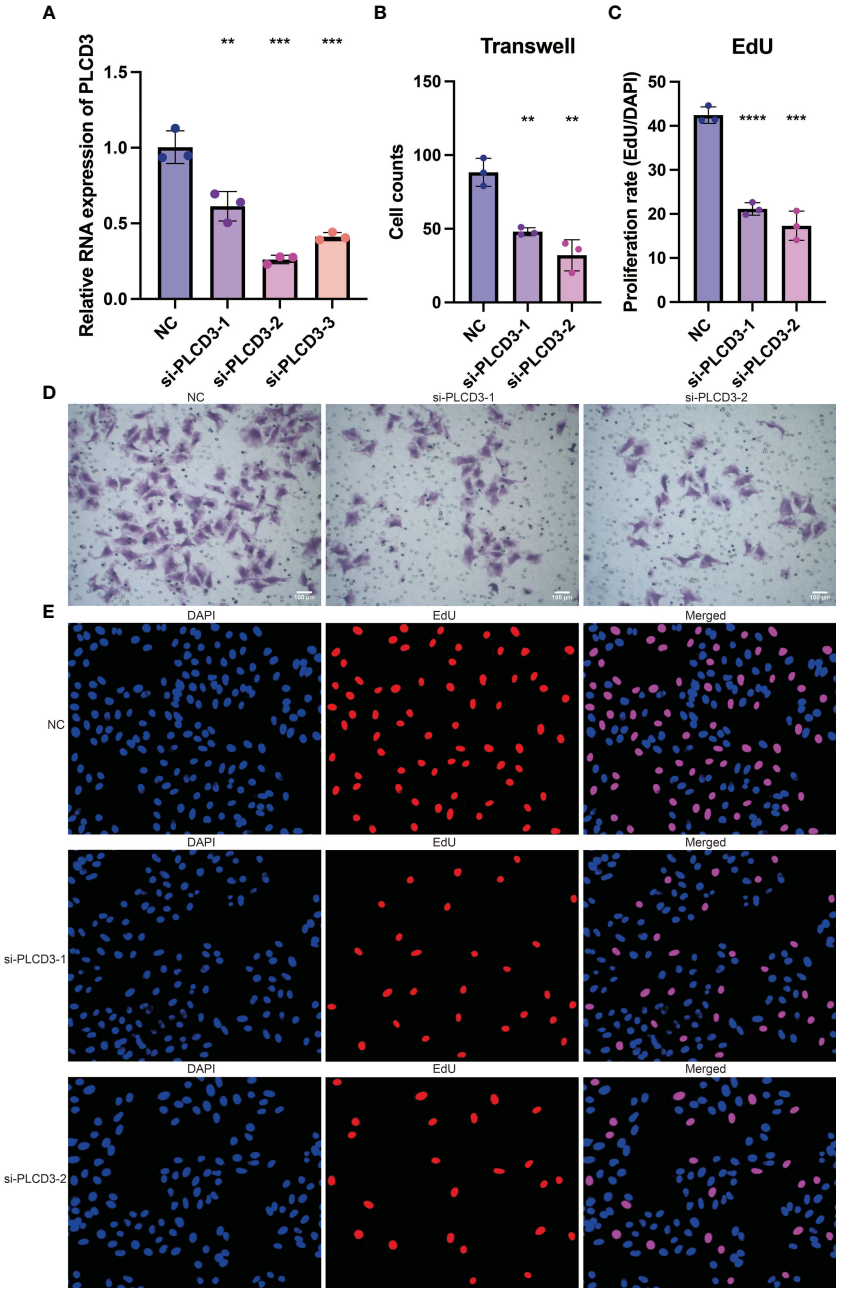


FIGURE 8
The tumor-promoting role of PLCD3. **(A)** The relative RNA expression of PLCD3 in NC and three si-RNA groups by q-PCR assay. **(B)** Statistical analysis of the cell counts in NC and two si-RNA groups by Transwell assay. **(C)** Statistical analysis of the proliferation rate (EdU/DAPI) in NC and two si-RNA groups by EdU assay. **(D)** The cell counts in NC and two si-RNA groups by Transwell assay. **(E)** The proliferation rate (EdU/DAPI) in NC and two si-RNA groups by EdU assay. **, P<0.01; ***, P<0.001; ****, P<0.0001.

tumorigenesis regulator. In xenograft models of multi-tumors, Tsvepkov P et al. (24) proved that FDX1 worked as an oncogene rescuing elesclomol-induced cell death. Zhang Y et al. (25) found that FDX1 could regulate iron metabolism and mitochondrial homeostasis in tumor cells through the p53 pathway. Our work may inspire more research on FDX1 as a key element in cuproptosis and an oncogene to control the pathogenesis of OS because no studies on the association between FDX1 and OS have been located.

Our study has also thoroughly examined the regulatory pathways connected to cuproptosis and its potential roles in OS. Cuproptosis gene up-regulation resulted in the enrichment of a few well-known cancer-promoting pathways, including TGF-, Wnt/-catenin, and p53 signaling. These results might offer suggestions for further experiments on cuproptosis regulation pathways. Additionally, we discovered that cuproptosis may also generate an immunosuppressive state and CSLC characteristics. Ferroptosis and cuproptosis have certain molecular commonalities in several types of programmed cell death. Both were correlated with the reduction of metal ions and redox metabolic pathway mediated by GSH/NADPH in mitochondria (26, 27). Some recent studies have

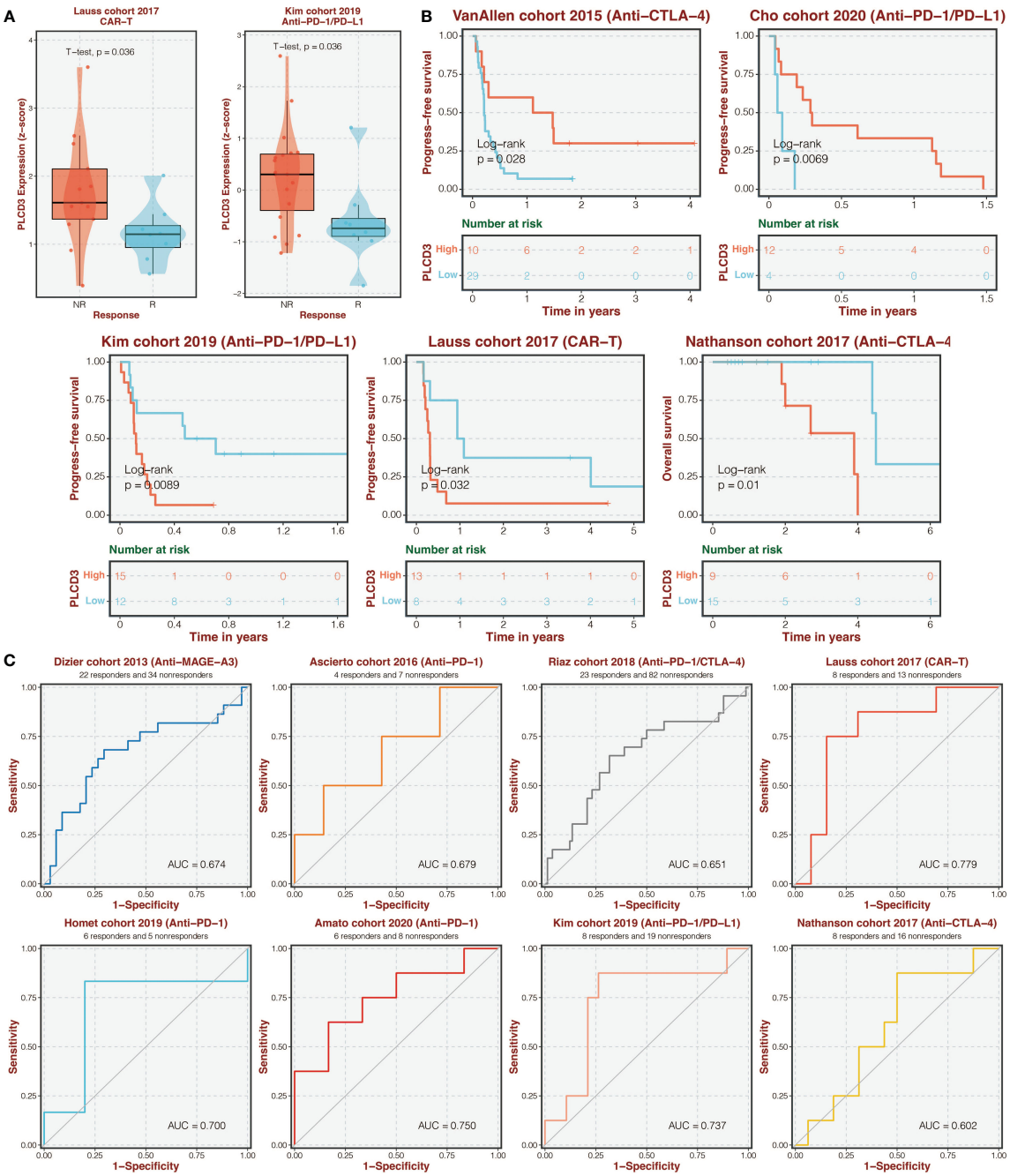


FIGURE 9
Immunotherapy prediction of PLCD3. **(A)** The expression of PLCD3 in responders and non-responders in immunotherapy cohorts. **(B)** Survival analysis was performed on the two groups regarding PLCD3 expression in immunotherapy cohorts. **(C)** The ROC curve of PLCD3 in predicting immunotherapy response in immunotherapy cohorts.

suggested that CSLCs might be sensitive to ferroptosis due to their relatively strong dependency on nutrition intake and higher intracellular levels of metal trace elements to maintain their self-renewal (28, 29). For tumor immunology, ferroptosis might also play a crucial role in regulating T cells. Ferroptosis induction in CD8+ and CD4+ T cells could lead to phospholipid hydroperoxide and impair its antitumor function (30, 31). These results are in line with our research, which show that cuproptosis-regulated gene activation is positively correlated with a higher mRNAsi index and an increase in the infiltration of immunosuppressive cells. Therefore, it is encouraging that future studies will concentrate on controlling cuproptosis in CSLCs and tumor microenvironments.

Our study identified five cuproptosis-related prognostic genes and built a reliable prognostic predicting model (CRP score model) based on them using a number of bioinformatic and machine learning methods. Our search revealed that studies on the role of these five genes, except MRTFA and RFX, in the etiology of OS had

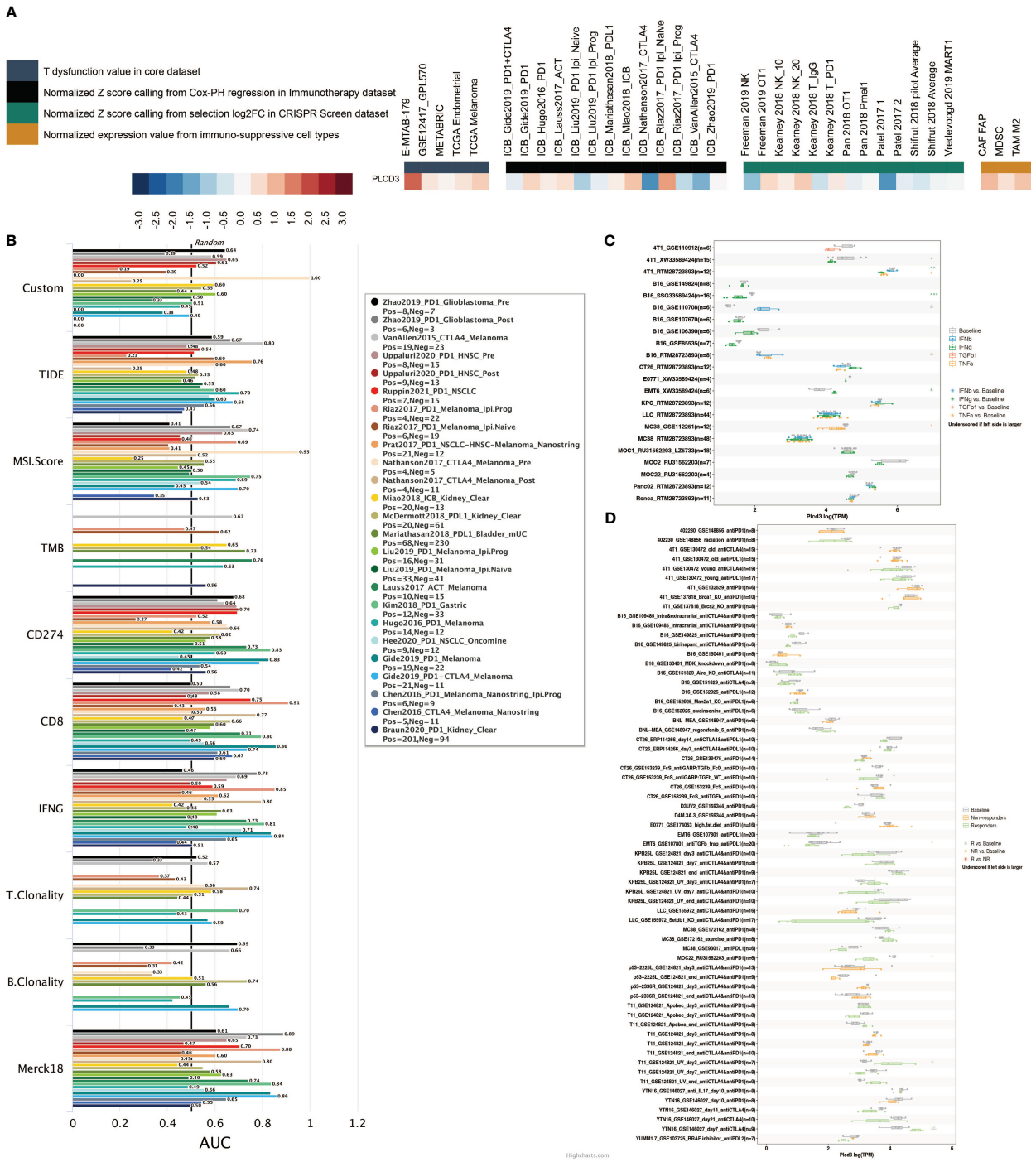


FIGURE 10
Immunotherapy prediction of PLCD3. **(A)** Regulator prioritization performed by TIDE. **(B)** Biomarker evaluation by TIDE. **(C)** Cytokine treatment prediction by TISMO. **(D)** Immunotherapy prediction by TISMO.

yet to be published. Matrix stiffness regulates EMT *via* cytoskeletal remodeling and MRTFA translocation in osteosarcoma (32). MRTFA is strongly associated with cell viability of its correlation with cytoskeleton and actin (33). It has been identified as an EMT and metastasis regulator in NPC (34) and NSCLC (35). BTBD10 functions as an activator of AKT family members by inhibiting PPP2CA-mediated dephosphorylation, and a few studies have identified it as a prognostic risk factor in hepatocellular carcinoma (36) and glioma (37). DLX1 serves as a two-sided transcriptional regulator of the TGF-β superfamily that may be either an oncogene or a suppressor in different types of tumors (38, 39). PLCD3 is a member of the phospholipase C family, which catalyzes the hydrolysis of phosphatidylinositol 4,5-bisphosphate to generate the second messenger diacylglycerol and inositol 1,4,5-

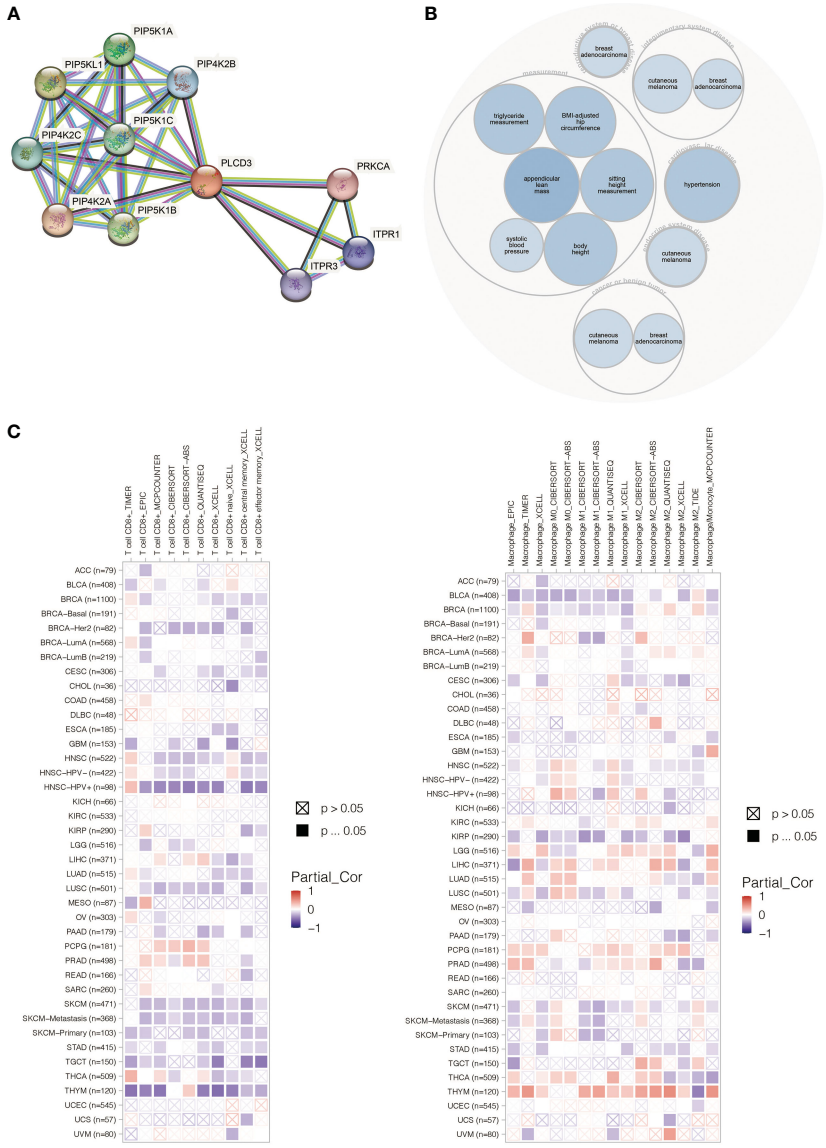


FIGURE 11
(A) The PLCD3 protein interaction network. **(B)** The PLCD3 illness network. **(C)** PLCD3's pan-cancer immune infiltration pattern.

trisphosphate (IP3) (40). PLCD3 is involved in the proliferation, migration, and invasion of nasopharyngeal carcinoma (41). PLCD3 inhibits apoptosis and promotes thyroid cancer's proliferation, migration, and invasion *via* the Hippo pathway (42). As PLCD3 was not studied in osteosarcoma, the *in vitro* validation was performed on PLCD3. PLCD3 could facilitate the proliferation and migration of osteosarcoma. p53 could directly regulate target genes, including MDM2, TP53I3, and RRM2B, or indirectly regulate numerous further genes through several hub genes, including EHF and RFX, through various drug treatments in osteosarcoma (43). RFX3 is a transcription factor that is essential for the differentiation of nodal monocilia (44). It has been reported that these two genes may also be involved in malignant biological behaviors of cancers (42, 45), but the mechanisms are poorly understood. Given that drug responses were predicted for OS patients with varying CRP scores, this 5-gene prognostic model is not only deserving of exploration of their mechanism in cuproptosis regulation and OS tumorigenesis/progression but also potential for translational medical outcomes, particularly for the future targeted therapy targeting PI3K/AKT/mTOR signaling, as compounds targeting this pathway could remain highly sensitive in patients with high CRP scores (poor prognosis).

We had to acknowledge that this study has limitations as researchers in the fields of bioinformatics and machine learning. Since OS is a relatively uncommon tumor, it is challenging to gather WGS data, and the sample size is modest when compared to other cancer types. The TARGET database provided by GDC has to have certain types of data, including SNP, copy number variation, and protein expression profiling, completed or accessible. These flaws likely decreased the power of statistical tests throughout the study,

particularly for machine learning-related studies like ridge regression and random survival forests. In a summary, our study demonstrated the distinctive cuproptosis regulatory gene expression profiles in osteosarcoma patients. It revealed some fresh information on the connections between this recently discovered kind of PCD and cancer-related pathways, stemness features, and immune infiltration traits. A scoring model based on cuproptosis-related clustering may have a significant impact on OS patient prognosis prediction and may influence clinical chemotherapy regimen selection and the creation of novel targeted medications.

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

SW and TC designed the concept of this research. HH conducted all the studies in this paper. YY validated and proofread the data. HH wrote this paper. BJ and ZF revised this paper. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (82072501), Science and Technology Innovation Leading Plan of High-Tech Industry in Hunan Province (2020SK2011).

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.2023.1156455/full#supplementary-material

FIGURE S1
Workflow diagram of this study.

FIGURE S2
Figure S2 **(A)** The expression of immune cells in two Cu Clusters. **(B)** The expression of immune cells in two CRP score groups. **(C)** The expression of checkpoints in two CRP score groups. **(D)** The correlation of drug response and CRP score.

FIGURE S3
Figure S3 Pharmacogenomics Analysis of CRP score. **(A)** Signaling pathways targeted by drugs resistant or sensitive to CRP_score in OS cell lines; **(B)** Molecular targets and signaling pathways targeted by drugs resistant or susceptible to CRP_score in OS samples predicted by the Oncopredict algorithm.

References

1. Mirabello L, Troisi R, Savage S. Osteosarcoma incidence and survival rates from 1973 to 2004: Data from the surveillance, epidemiology, and end results program. *Cancer* (2009) 115(7):1531–43. doi: 10.1002/cncr.24121

2. Siegel R, Miller K, Fuchs H, Jemal A. Cancer statistics, 2021. *CA: Cancer J Clin* (2021) 71(1):7–33. doi: 10.3322/caac.21654

3. Isakoff M, Bielack S, Meltzer P, Gorlick R. Osteosarcoma: Current treatment and a collaborative pathway to success. *J Clin Oncol* (2015) 33(27):3029–35. doi: 10.1200/JCO.2014.59.4895

4. Li S, Liu F, Zheng K, Wang W, Qiu E, Pei Y, et al. CircDOCK1 promotes the tumorigenesis and cisplatin resistance of osteogenic sarcoma *via* the miR-339-3p/IGF1R axis. *Mol cancer* (2021) 20(1):161. doi: 10.1186/s12943-021-01453-0

5. Tang Z, Dong H, Li T, Wang N, Wei X, Wu H, et al. The synergistic reducing drug resistance effect of cisplatin and ursolic acid on osteosarcoma through a multistep mechanism involving ferritinophagy. *Oxid Med Cell longevity* (2021) 2021:5192271. doi: 10.1155/2021/5192271

6. Luo Y, Gao X, Zou L, Lei M, Feng J, Hu Z. Bavachin induces ferroptosis through the STAT3/P53/SLC7A11 axis in osteosarcoma cells. *Oxid Med Cell longevity* (2021) 2021:1783485. doi: 10.1155/2021/1783485

7. Lv H, Zhen C, Liu J, Shang P. PEITC triggers multiple forms of cell death by GSH-iron-ROS regulation in K7M2 murine osteosarcoma cells. *Acta pharmacologica Sinica* (2020) 41(8):1119–32. doi: 10.1038/s41401-020-0376-8

8. Li S, Zhang T, Xu W, Ding J, Yin F, Xu J, et al. Sarcoma-targeting peptide-decorated polypeptide nanogel intracellularly delivers shikonin for upregulated osteosarcoma necroptosis and diminished pulmonary metastasis. *Theranostics* (2018) 8(5):1361–75. doi: 10.7150/thno.18299

9. Tsvetkov P, Coy S, Petrova B, Dreishpoon M, Verma A, Abdusamad M, et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. *Sci (New York NY)* (2022) 375(6586):1254–61. doi: 10.1126/science.abf0529

10. O'Day S, Eggermont A, Chiarion-Sileni V, Kefford R, Grob J, Mortier L, et al. Final results of phase III SYMMETRY study: Randomized, double-blind trial of elesclomol plus paclitaxel versus paclitaxel alone as treatment for chemotherapy-naive patients with advanced melanoma. *J Clin Oncol* (2013) 31(9):1211–8. doi: 10.1200/JCO.2012.44.5585

11. Zhong Z, Mao S, Lin H, Li H, Lin J, Lin J. Alteration of intracellular metabolome in osteosarcoma stem cells revealed by liquid chromatography-tandem mass spectrometry. *Talanta* (2019) 204:6–12. doi: 10.1016/j.talanta.2019.05.088

12. Hesterberg T. *Bootstrap* Vol. 3. Wiley Interdisciplinary Reviews: Computational Statistics (2011) p. 497–526.

13. Chi H, Xie X, Yan Y, Peng G, Strohmer DF, Lai G, et al. Natural killer cell-related prognosis signature characterizes immune landscape and predicts prognosis of HNSCC. *Front Immunol* (2022) 13:1018685. doi: 10.3389/fimmu.2022.1018685

14. Chi H, Jiang P, Xu K, Zhao Y, Song B, Peng G, et al. A novel anoikis-related gene signature predicts prognosis in patients with head and neck squamous cell carcinoma

and reveals immune infiltration. *Front Genet* (2022) 13:984273. doi: 10.3389/fgene.2022.984273

15. Chi H, Peng G, Wang R, Yang F, Xie X, Zhang J, et al. Cuprotosis programmed-Cell-Death-Related lncRNA signature predicts prognosis and immune landscape in PAAD patients. *Cells* (2022) 11(21). doi: 10.3390/cells11213436

16. Malta T, Sokolov A, Gentles A, Burzykowski T, Poisson L, Weinstein J, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* (2018) 173(2):338–54.e15. doi: 10.1016/j.cell.2018.03.034

17. Lian H, Han Y, Zhang Y, Zhao Y, Yan S, Li Q, 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 (10):2227–45. doi: 10.1002/1878-0261.12557

18. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of drug sensitivity in cancer (GDSC). *Nucleic Acids Res* (2015) 41:955–61. doi: 10.1093/nar/gks1111

19. 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) 11). doi: 10.1093/bib/bbab260

20. Chi H, Peng G, Yang J, Zhang J, Song G, Xie X, et al. Machine learning to construct sphingolipid metabolism genes signature to characterize the immune landscape and prognosis of patients with uveal melanoma. *Front Endocrinol (Lausanne)* (2022) 13:1056310. doi: 10.3389/fendo.2022.1056310

21. Peng G, Chi H, Gao X, Zhang J, Song G, Xie X, et al. Identification and validation of neurotrophic factor-related genes signature in HNSCC to predict survival and immune landscapes. *Front Genet* (2022) 13:1010044. doi: 10.3389/fgene.2022.1010044

22. Sheftel A, Stehling O, Pierik A, Elsässer H, Mühlenhoff U, Webert H, et al. Humans possess two mitochondrial ferredoxins, Fdx1 and Fdx2, with distinct roles in steroidogenesis, heme, and Fe/S cluster biosynthesis. *Proc Natl Acad Sci United States America* (2010) 107(26):11775–80. doi: 10.1073/pnas.1004250107

23. Strushkevich N, MacKenzie F, Cherkesova T, Grabovec I, Usanov S, Park H. Structural basis for pregnenolone biosynthesis by the mitochondrial monooxygenase system. *Proc Natl Acad Sci United States America* (2011) 108(25):10139–43. doi: 10.1073/pnas.1019441108

24. Tsvetkov P, Detappe A, Cai K, Keys H, Brune Z, Ying W, et al. Mitochondrial metabolism promotes adaptation to proteotoxic stress. *Nat Chem Biol* (2019) 15 (7):681–9. doi: 10.1038/s41589-019-0291-9

25. Zhang Y, Qian Y, Zhang J, Yan W, Jung Y, Chen M, et al. Ferredoxin reductase is critical for p53-dependent tumor suppression *via* iron regulatory protein 2. *Genes Dev* (2017) 31(12):1243–56. doi: 10.1101/gad.299388.117

26. Zheng J, Conrad M. The metabolic underpinnings of ferroptosis. *Cell Metab* (2020) 32(6):920–37. doi: 10.1016/j.cmet.2020.10.011

27. Chen X, Li J, Kang R, Klionsky D, Tang D. Ferroptosis: Machinery and regulation. *Autophagy* (2021) 17(9):2054–81. doi: 10.1080/15548627.2020.1810918

28. Raz N, Daugherty A. Pathways to brain aging and their modifiers: Free-Radical-Induced energetic and neural decline in senescence (FRIENDS) model - a mini-review. *Gerontology* (2018) 64(1):49–57. doi: 10.1159/000479508

29. Wu S, Li T, Liu W, Huang Y. Ferroptosis and cancer: Complex relationship and potential application of exosomes. *Front Cell Dev Biol* (2021) 9:733751. doi: 10.3389/fcell.2021.733751

30. Xu S, Chaudhary O, Rodríguez-Morales P, Sun X, Chen D, Zappasodi R, et al. Uptake of oxidized lipids by the scavenger receptor CD36 promotes lipid peroxidation and dysfunction in CD8 T cells in tumors. *Immunity* (2021) 54(7):1561–77.e7. doi: 10.1016/j.immuni.2021.05.003

31. Wang W, Green M, Choi J, Gijón M, Kennedy P, Johnson J, et al. CD8 T cells regulate tumour ferroptosis during cancer immunotherapy. *Nature* (2019) 569 (7755):270–4. doi: 10.1038/s41586-019-1170-y

32. Dai J, Qin L, Chen Y, Wang H, Lin G, Li X, et al. Matrix stiffness regulates epithelial-mesenchymal transition *via* cytoskeletal remodeling and MRTF-a translocation in osteosarcoma cells. *J Mech Behav BioMed Mater* (2019) 90:226–38. doi: 10.1016/j.jmbbm.2018.10.012

33. Chatzifrangkeskou M, Pefani D, Eyres M, Vendrell I, Fischer R, Pankova D, et al. RASSF1A is required for the maintenance of nuclear actin levels. *EMBO J* (2019) 38 (16):e101168. doi: 10.1101/559310

34. Wang J, Zhong Q, Zhang H, Liu S, Li S, Xia T, et al. Nogo-b promotes invasion and metastasis of nasopharyngeal carcinoma *via* RhoA-SRF-MRTFA pathway. *Cell Death disease* (2022) 13(1):76. doi: 10.1038/s41419-022-04518-0

35. Du F, Qi X, Zhang A, Sui F, Wang X, Proud C, et al. MRTF-A-NF-κB/p65 axis-mediated PDL1 transcription and expression contributes to immune evasion of non-small-cell lung cancer *via* TGF-β. *Exp Mol Med* (2021) 53(9):1366–78. doi: 10.1038/s12276-021-00670-3

36. Li J, Tian X, Nie Y, He Y, Wu W, Lei X, et al. BTBD10 is a prognostic biomarker correlated with immune infiltration in hepatocellular carcinoma. *Front Mol biosci* (2021) 8:762541. doi: 10.3389/fmolb.2021.762541

37. Chen J, Xu J, Ying K, Cao G, Hu G, Wang L, et al. Molecular cloning and characterization of a novel human BTB domain-containing gene, BTBD10, which is down-regulated in glioma. *Gene* (2004) 340(1):61–9. doi: 10.1016/j.gene.2004.05.028

38. Leyten G, Hessels D, Smit F, Jannink S, de Jong H, Melchers W, et al. Identification of a candidate gene panel for the early diagnosis of prostate cancer. *Clin Cancer Res* (2015) 21(13):3061–70. doi: 10.1158/1078-0432.CCR-14-3334

39. Chan D, Hui W, Wang J, Yung M, Hui L, Qin Y, et al. DLX1 acts as a crucial target of FOXM1 to promote ovarian cancer aggressiveness by enhancing TGF-β/SMAD4 signaling. *Oncogene* (2017) 36(10):1404–16. doi: 10.1038/onc.2016.307

40. Pawelczyk T, Matecki A. Phospholipase c-delta3 binds with high specificity to phosphatidylinositol 4,5-bisphosphate and phosphatidic acid in bilayer membranes. *Eur J Biochem* (1999) 262(2):291–8. doi: 10.1046/j.1432-1327.1999.00388.x

41. Liu W, Liu X, Wang L, Zhu B, Zhang C, Jia W, et al. PLCD3, a flotillin2-interacting protein, is involved in proliferation, migration and invasion of nasopharyngeal carcinoma cells. *Oncol Rep* (2018) 39(1):45–52. doi: 10.3892/or.2017.6080

42. Lin L, Wen J, Lin B, Chen H, Bhandari A, Qi Y, et al. Phospholipase c delta 3 inhibits apoptosis and promotes proliferation, migration, and invasion of thyroid cancer cells *via* hippo pathway. *Acta Biochim Biophys Sin (Shanghai)* (2021) 53(4):481–91. doi: 10.1093/abbs/gmab016

43. Zhou H, Cui X, Yuan H, Zhang B, Meng C, Zhao D. Effects of distinct drugs on gene transcription in an osteosarcoma cell line. *Oncol Lett* (2017) 14(4):4694–700. doi: 10.3892/ol.2017.6767

44. Smith S, Qu H, Taleb N, Kishimoto N, Scheel D, Lu Y, et al. Rfx6 directs islet formation and insulin production in mice and humans. *Nature* (2010) 463(7282):775–80. doi: 10.1038/nature08748

45. Qian Y, Cheng B, Luo J, Hu Y, Gao L, Cheng H. CircRFX3 up-regulates its host gene RFX3 to facilitate tumorigenesis and progression of glioma. *J Mol Neurosci* (2022) 72(6):1195–1207. doi: 10.1007/s12031-022-02005-x

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 21 March 2023
DOI 10.3389/fonc.2023.1161666

OPEN ACCESS

EDITED BY
Nan Zhang,
Harbin Medical University, China

REVIEWED BY
Wenfeng Lin,
The Sixth Affiliated Hospital of Guangzhou Medical University, China
Gen Wu,
Southern Medical University, China

*CORRESPONDENCE
Xiaoyu Cui
cuixiaoyu126@163.com

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Oncology

RECEIVED 08 February 2023
ACCEPTED 07 March 2023
PUBLISHED 21 March 2023

CITATION
Xia G, Wu S and Cui X (2023) An immune infiltration-related prognostic model of kidney renal clear cell carcinoma with two valuable markers: CAPN12 and MSC. *Front. Oncol.* 13:1161666. doi: 10.3389/fonc.2023.1161666

An immune infiltration-related prognostic model of kidney renal clear cell carcinoma with two valuable markers: CAPN12 and MSC

Guang Xia[1], Song Wu[1] and Xiaoyu Cui[2]*

[1]Department of Orthopaedics of the 3rd Xiangya Hospital, Central South University, Changsha, China, [2]Department of Anesthesiology of the 3rd Xiangya Hospital, Central South University, Changsha, China

Background: Since its discovery, clear cell renal cell carcinoma (ccRCC) has been the most prevalent and lethal kidney malignancy. Our research aims to identify possible prognostic genes of ccRCC and to develop efficient prognostic models for ccRCC patients based on multi-omics investigations to shed light on the treatment and prognosis of ccRCC.

Methods: To determine a risk score for each patient, we screened out differentially expressed genes using data from tumor samples, and control samples mined from The Cancer Genome Atlas (TCGA) and GTEx datasets. Somatic mutation and copy number variation profiles were analyzed to look for specific genomic changes connected to risk scores. To investigate potential functional relationships of prognostic genes, gene set variation analysis (GSVA) and gene set enrichment analysis (GSEA) were carried out. We created a prognostic model by fusing risk ratings with other clinical variables. For validation, the 786-O cell line was used to carry out the dual-gRNA approach to knock down CAPN12 and MSC. This was followed by qRT-PCR to verify the knockdown of CAPN12 and MSC.

Results: For ccRCC, seven predictive genes were discovered: PVT1, MSC, ALDH6A1, TRIB3, QRFPR, CYS1, and CAPN12. The most enriched pathways in the GSVA study and GSEA analysis promote tumorigenesis and immune system modulation. The risk score derived from prognostic genes corresponds with immune infiltration cells and helps predict how well a medicine will work. The mutation of numerous oncogenes was also linked to a high-risk score. A prognostic model with a high ROC value was created for the risk score. An *in vitro* study demonstrates that the suppression of CAPN12 and MSC dramatically reduced the ability of 786-O cells to proliferate in the CCK-8 proliferation assay and plate clonality assays.

Conclusions: A thorough prognostic model with good performance has been developed for ccRCC patients using seven prognostic genes that were discovered to be related to ccRCC prognosis. In ccRCC, CAPN12 and MSC were significant indicators and would make good therapeutic targets.

KEYWORDS

ccRCC, prognostic model, immune infiltration, somatic mutation, cell proliferation

Background

Kidney cancer has long been a common malignant tumor in the urinary system, with an increasing incidence rate worldwide. In the USA, 65,000 individuals are newly diagnosed with kidney cancers per year (1). Among all kinds, clear cell renal cell carcinoma (ccRCC) accounts for approximately 80% of kidney cancers, which also correlates with worse survival outcomes (2). Although the 5-year overall survival (OS) of patients with early diagnosis of ccRCC is about 90%, the 5-year OS for patients diagnosed at an advanced stage is down to 12% (3). Unfortunately, almost 20% of cases are in advanced malignant stages when diagnosed (4). Regarding treatment, nephrectomy continues to be the optimal approach for localized ccRCC. A phase 3 clinical trial has proved that nephrectomy with adjuvant chemotherapy increased the progressive free survival (PFS) of ccRCC patients to 6.8 years compared with nephrectomy alone (5.8 years) (5). Although chemotherapy is a good option for multiple cancer types, ccRCC shows resistance to chemotherapy *via* secreting vascular endothelial cell growth factor (VEGF) (6). Other molecules, such as the mammalian target of rapamycin (mTOR) and the mitogen-activated protein kinase (MAPK), have also been demonstrated to be involved in the carcinogenesis of ccRCC and dampen the effectiveness of chemotherapy (7, 8).

Recently, immunotherapies combined with conventional surgical resection and radiotherapy have gradually improved the clinical management of ccRCC (9). However, the mortality rate of ccRCC remains high due to diagnostic difficulty at the early stage of the disease. Thus 30% of patients inevitably would suffer from tumor recurrence and progression (9). Combining ccRCC prognostic genes, researchers have built some predictive models for ccRCC patients based on online databases, such as The Cancer Genome Atlas (TCGA), with many genetic ccRCC samples. However, no prognostic model of ccRCC has been widely accepted. Thus, a risk stratification model identifying ccRCC-related biomarkers and assessing the prognosis of ccRCC patients is urgently needed. In this study, we present a ccRCC prognostic model after mining and screening multiple predictive genes from the TCGA dataset, aiming to shed light on optimizing the clinical management of ccRCC patients.

Abbreviations: ccRCC, Clear cell renal cell carcinoma; DEGs, Differentially expressed genes; GSVA, Geneset variation analysis; GSEA, Geneset enrichment analysis; CCK-8, Cell Counting Kit-8; mTOR, mammalian target of rapamycin; VEGF, vascular endothelial cell growth factor; MAPK, mitogen-activated protein kinase; TCGA, The Cancer Genome Atlas; KIRC, Kidney Renal Clear Cell Carcinoma; ROC, Receiver operating characteristic; AUC, area under the curve; CNV, Copy number variation; TIDE, Tumor Immune Dysfunction and Exclusion.

Materials and methods

Datasets and preprocessing

We gathered two cohorts of patients with ccRCC for this study: GSE29609 (microarray) from the platform (GPL1708) and TCGA Kidney Renal Clear Cell Carcinoma (KIRC) (RNA-seq) cohort. Raw data from the microarray dataset generated by Agilent was downloaded from the Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/). Gene expression profile induced by Illumina and corresponding clinical information were downloaded from The Cancer Genome Atlas (TCGA) data source (https://xena.ucsc.edu). Raw data for the dataset from Agilent were processed using the RMA algorithm for background adjustment in the *limma* software package. The raw data from Illumina was processed using the lumi software package (10). For the TCGA cohort, RNA-sequencing data (FPKM values) were transformed into transcripts per kilobase million (TPM) values that are more similar to the values from the microarray. Samples without survival information were eliminated, 528 KIRC samples in TCGA were screened out for the risk score construction, and 39 KIRC samples in GEO were screened out for external validation of the risk score. One hundred standard pieces were downloaded from https://xenabrowser.net/datapages/?cohort=TCGA%20TARGET%20GTEx&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443, among which 28 regular renal models were from the GTEx database (https://xena.ucsc.edu), and 72 normal renal samples were from the TCGA database (https://xena.ucsc.edu). These 100 normal samples were already combined, so removing the batch effect was unnecessary. The TCGA KIRC cohort was randomly divided into two equal parts: the train set (set 1) and the validation set (set 2). The total TCGA KIRC data were used as another verification set (set 3), while the GEO cohort was used as the external validation set in the following studies (set 4).

Identification of differentially expressed genes in KIRC

Probes without corresponding gene symbols were filtered out, and the average value of gene symbols with multiple searches was calculated. Between the two groups, the Linear Models for Microarray Data Analysis (limma) package (10) was used to screen the differentially expressed genes (DEGs). Threshold values were set as adjusted P<0.05 and the absolute value of logFC> 2. A principal component analysis was also applied to categorize the data further to assess the DEGs' accuracy.

Screening and confirmation of the prognostic value of the genes

By intersecting the obtained differential expressed gene with the genes of TCGA, genes for further analysis were obtained. In the training set (set 1), univariate Cox proportional hazard regression analysis was performed using the survival package in R to investigate the relationship between patients' overall survival (OS) and gene expression level. Genes were considered significant with prognostic potential at a P-value<0.05. Next, we applied an L1-penalized (Lasso) regression to identify the differentially expressed genes with independent predictive values. Lasso regression is a valuable method to determine interpretable prediction rules in high dimension data (11). We obtained a set of prognostic genes and their corresponding LASSO coefficients based on the highest lambda value selected through 1,000 cross-validations in the Lasso method (lambda.1se). To evaluate whether the selected genes were related to the prognosis of KIRC patients, patients of set 1 were assigned into two groups based on the median expression value of each gene. Kaplan-Meier plots were used to determine their prognostic value, and P<0.05 was considered statistically significant. A genes-based survival risk assessment model was established using the LASSO coefficients. Then, patients were divided into low-risk and high-risk groups using the median risk scores in the other three sets. Kaplan-Meier plots and Log-rank tests were used to estimate and compare the OS of patients between the two risk groups; P<0.05 was set as the cutoff. The time-dependent receiver operating characteristic (ROC) curve and the area under the curve (AUC) were applied to evaluate the prediction accuracy of the risk model and the selected genes. Furthermore, stratified survival analyses were also conducted to explore whether the gene-based risk assessment model has predictive value among different age groups (older or younger than 60), primary tumor lesions (T1, T2, T3, T4), and stage (stage I, stage ii, stage iii, stage iv).

Consensus clustering of prognostic genes

To investigate the function of seven prognostic genes in KIRC, we clustered the KIRCs into different groups with "ConsensusClusterPlus" (50 iterations, resample rate of 80%, and Pearson correlation). PCA with the R package for R v3.4.1 was adopted to study the gene expression patterns in different KIRC groups.

Genomic alterations of samples clustered by risk scores

To determine whether risk score levels are associated with specific genomic characteristics in ccRCC, we performed copy number variation (CNV) and somatic mutation analysis using the TCGA dataset. GSITIC analysis was adopted to determine the genomic event enrichment.

Prognostic model based on clinical features and risk score

Univariate Cox proportional hazard regression analysis was performed using the survival package for the risk score and clinical features (Age, Tumor primary lesion, Stage) with a P value <0.05 as the cutoff. Then we built a Multivariate Cox model based on the selected features, and the Nomogram chart was drawn using the replot package. The Calibration curve and the AUC assessed the risk model.

Gene set variation analysis and geneset enrichment analysis

The gene set *variation analysis (GSVA) and geneset enrichment analysis (GSEA)* packages were used to calculate the enrichment status in Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) terms of TCGA samples. Correlation analysis was performed by expression values of risk score, GO terms, and KEGG terms. The items with p<0.05 and a high correlation coefficient were selected (12).

Immunological function analyses

A single sample gene set enrichment analysis (ssGSEA) was performed using R software to quantify 28 tumor-infiltrating immune cells (Foroutan et al., 2018). Correlation analysis between risk score and tumor-infiltrating immune cell expressions was performed using gene expression profiles from the TCGA datasets.

Prediction of chemotherapeutic and immunotherapy response

The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was performed to infer individual responses to immunotherapy, such as immune checkpoint blockade (e.g., anti-PD-1 therapy). The submap analysis was applied to show the difference in response to anti-PD-1 and CTAL-4 therapy (13). The chemotherapeutic response for each ccRCC patient was predicted according to the public pharmacogenomic database, Genomics of

Drug Sensitivity in Cancer (GDSC, www.cancerrxgene.org). The prediction of drug sensitivity (IC50) values was conducted using the R package "prophetic" (14).

CAPN12 and MSC knockdown

Knockdown plasmids were constructed by the dual-gRNA method (15), targeting CAPN12 and MSC. Vectors without specific gRNAs were used as control. All PCR products were verified by DNA sequencing. Transfection of plasmids was carried out using Lipofectamine 2000 (Invitrogen, USA) according to the manufacturer's instructions. After the transfection, cells were seeded and grown in the RPMI-1640 supplemented with 5% FBS. Then 786-O cell clones were picked, and the expression of CAPN12 and MSC were validated by qRT-PCR. Plate clonality assays were also used to measure the impact of knockdown on cell clonality and cell cycle in the 786-O cell line after silencing CAPN12 and MSC.

Quantitative real-time polymerase chain reaction

Three biological replicates were analyzed, with technical replicates for each triplicate biological sample. Total RNAs were extracted, reversed, and transcribed into cDNA by HiScript Q RT SuperMix for qRT-PCR. ChamQ SYBR qRT-PCR Master Mix was used for qRT-PCR experiments, and its protocol was as follows: 95° C 30 s, 95°C 10 s, 60°C 30 s, for a total of 40 cycles reactions. The expression level of target genes was quantified using the 2-ΔΔCT method. GADPH was used as the internal standard. The primers are as follows: CAPN12, 5'-CTCCATTTCGACACCGTGCAG-3', 5'-GAGTTGAAGCCACGCACCCA-3'; MSC, 5'-CAACTCG TAGTCCACGCTCC-'3, 5'-TAAAAACCCAGGCCGGGAAG-3'.

Cell proliferation assay

Cell Counting Kit-8 (CCK-8) proliferation assay was conducted to assess the proliferation ability of cells according to the manufacturer's instructions. After cell counting, 1×104 cells were seeded into 96-well plates and incubated at 37°C for 24 h, 48 h, and 72 h. ten μL CCK-8 reagent was added into each well, and the absorbance at 450 nm was tested one h later.

Colony forming assay

Cells were digested and plated in 6-well plates (300 cells per well) and cultured with 5% CO2 at 37C for two weeks. The colonies were then fixed with 4% methanol (1 ml per well) for 15 minutes and stained with crystal violet for 30 minutes at room temperature. After the photograph, discoloration was performed with 10% acetic acid, and cells were measured absorbance at 550 nm.

Statistical analysis

All statistical analyses were performed using R software. A two-tailed t-test and one-way ANOVA determined significant quantitative differences between and among groups. The chi-square test was used to analyze the correlation of the classified data. The Kaplan–Meier method calculated the overall survival difference. Cox regression analysis was performed using the survival package in R. Spearman correlation to measure the strength of the association between two ranked variables. The *gene sets enrichment analysis* (*GSVA*) box was used to calculate the enrichment status in GO (Biological Process) (12). The R package survival ROC was used to plot and visualize receiver operating characteristic (ROC) curves to calculate the area under the curve (AUC) (16). All figures and statistical analyses were performed based on R language for Windows, version 3.5.1(http://www.r-project.org). Somatic mutations and somatic copy number alternations (CNAs) data were downloaded from the TCGA database. Copy number alternations associated with risk scores were analyzed using GISTIC 2.0 (https://gatkforums.broadinstitute.org). Adjusted P values were obtained by False Discovery Rate (FDR) correction. P values and adjusted P values of less than 0.05 were considered statistically significant.

Results

Data preprocessing and DEGs screening

The flow chart of this study is shown in Supplementary Figure S1A. After mining the data in the GTEx and TCGA databases, 528 KIRC and 100 normal samples were gathered and clustered to screen for differentially expressed genes between cancer and normal tissue. With a threshold of logFC>2 and adjust P ≦0.05, 594 genes (Table S1) were found to be differentially expressed, among which 227 genes were up-regulated, and 367 genes were down-regulated (Figure 1A). Those DEGs in KIRC and normal tissue can be separated by PCA (Figure 1B). The heatmap shows that the DEGs effectively separate KIRC and normal tissue (Figure 1C).

Development of the risk score with TCGA train set

To calculate the risk score, five hundred twenty-eight samples from TCGA were randomly separated into 264 and 264. In the train set containing 264 patients, lasso regression was adopted to analyze the data. After multiplying gene expression with LASSO coefficients, we came to seven prognostic genes: PVT1, MSC, ALDH6A1, TRIB3, QRFPR, CYS1, and CAPN12 (Table S2). The risk score was then

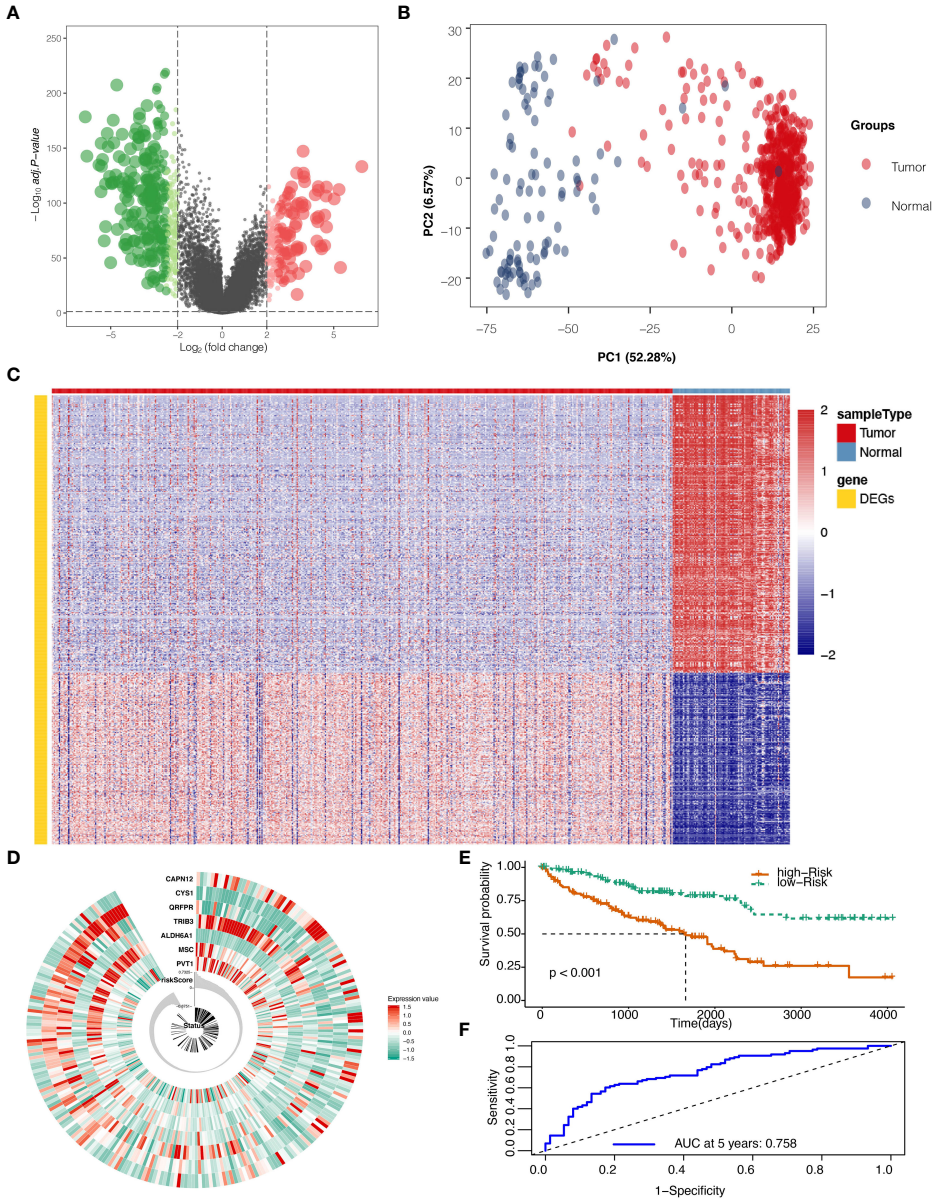


FIGURE 1
Differentially expressed genes (DEGs) screening and localization. **(A)** Volcano plot for DEGs with adjusted P values (FDR correction) less than 0.05. **(B)** Principal component analysis (PCA) to validate screening results. **(C)** Heat map result for DEGs screening. **(D)** Risk scores in the Cancer Genome Atlas (TCGA) train set, patient survival, and expression of 7 DEGs in the train set. **(E)** Risk score and patients' survival probabilities in the TCGA train set. **(F)** Receiver operating characteristic curve (ROC) of the risk score in the TCGA train set.

calculated for patients using seven prognostic genes between high and low-risk groups set at the median value (Figure 1D). Risk score=0.0009* PVT1 (gene expression level) + 0.0015*MSC + -0.0029*ALDH6A1 + 0.0022*TRIB3 + -0.0003*QRFPR +-0.0038*CYS1 + 0.0011* CAPN12. The calculated risk score ranged from -0.875 to 0.733 and had a median value of -0.007, in which the patients were grouped into a high-risk group and a low-risk group based on the median value of the risk score. In the train set, the high-risk and low-risk groups presented significantly different survival probabilities (Figure 1E) with an AUC of 0.758 in the time-dependent ROC curve at five years (Figure 1F).

Validation of the risk score with TCGA and GEO data

The Risk Score was calculated in the test set (Figure 2A). With the cutoff of risk score, survival probability between the high and low-risk score groups is statistically significant (P<0.001) with an AUC value of 0.716 (Figures 2B, C). When summed up, the risk score was further calculated with a P value of less than 0.001 between high and low-risk score groups and an AUC of 0.833 (Figures 2D–F). The model was then tested using GEO data in microarray GSE29609 from platform GPL1708. The difference

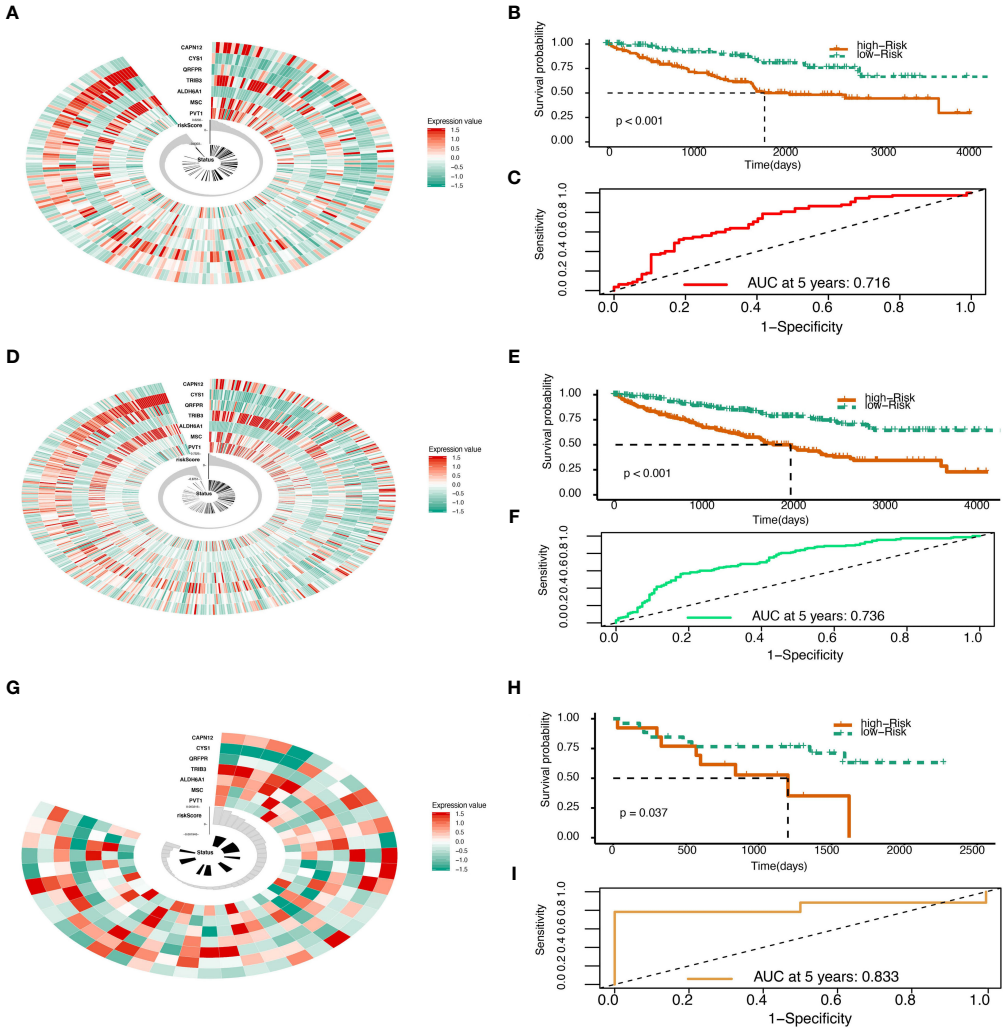


FIGURE 2
Risk score validation. **(A)** Risk scores in the TCGA test set, patient survival, and expression of 7 DEGs in the test set. **(B)** Risk score and patients' survival probabilities in the TCGA test set. **(C)** The ROC of risk scores in the TCGA test set. **(D)** Risk scores in the TCGA sum set, patient survival, and expression of 7 DEGs in the sum set. **(E)** Risk score and patients' survival probabilities in the TCGA sum set. **(F)** The ROC of the risk scores in the TCGA sum set. **(G)** Risk scores in the Gene Expression Omnibus (GEO) validation set, patient survival, and expression of 7 DEGs in the validation set. **(H)** Risk scores and patients' survival probabilities in the GEO validation set. **(I)** The ROC of risk scores in the GEO set.

between high and low-risk score groups in GEO data analysis was also statistically significant (P=0.037) with an AUC of 0.833 (Figures 2G–I). All model evaluation was based on time-dependent ROC at five years.

Genomic alterations and gene set enrichment analyses

To determine whether risk score levels were associated with specific genomic characteristics, we performed CNV and somatic mutation analysis using the TCGA dataset (Table S13). In high-score samples, frequently amplified genomic regions included oncogenic driver genes such as RSRC1 (3q25.32, p<0.001), SLC2A9 (4p16.1, p<0.001), EXOC2 (6p25.3, p<0.001), EGFR (7p11.2, p<0.001), and ERC1 (12p13.33, p<0.001) (Figure 3A). In contrast, deleted regions contained tumor suppressor genes including PTENP1 (9p13.3, p<0.001), FAM138C (9p24.3, p<0.001), and OR4K15 (14q11.2, p<0.001) (Figure 3A). In low-score samples, most amplified and deleted genomic regions were similar to those in high-score models. Analysis of somatic mutation profiles based on risk score levels revealed a high frequency of mutations in SETD2 (19%, p < 0.001), BAP1 (17%, p < 0.001), and KDM5C (10%, p < 0.01) in the high-score group (n = 166) (Figure 3B; Table S14). Genomic event enrichments were identified in either the low-score or high-score groups, respectively (Figure 3B).

Consensus clustering of seven prognostic genes

Consensus clustering of the seven prognostic genes identified three clusters of KIRCs in the TCGA dataset with distinct clinical

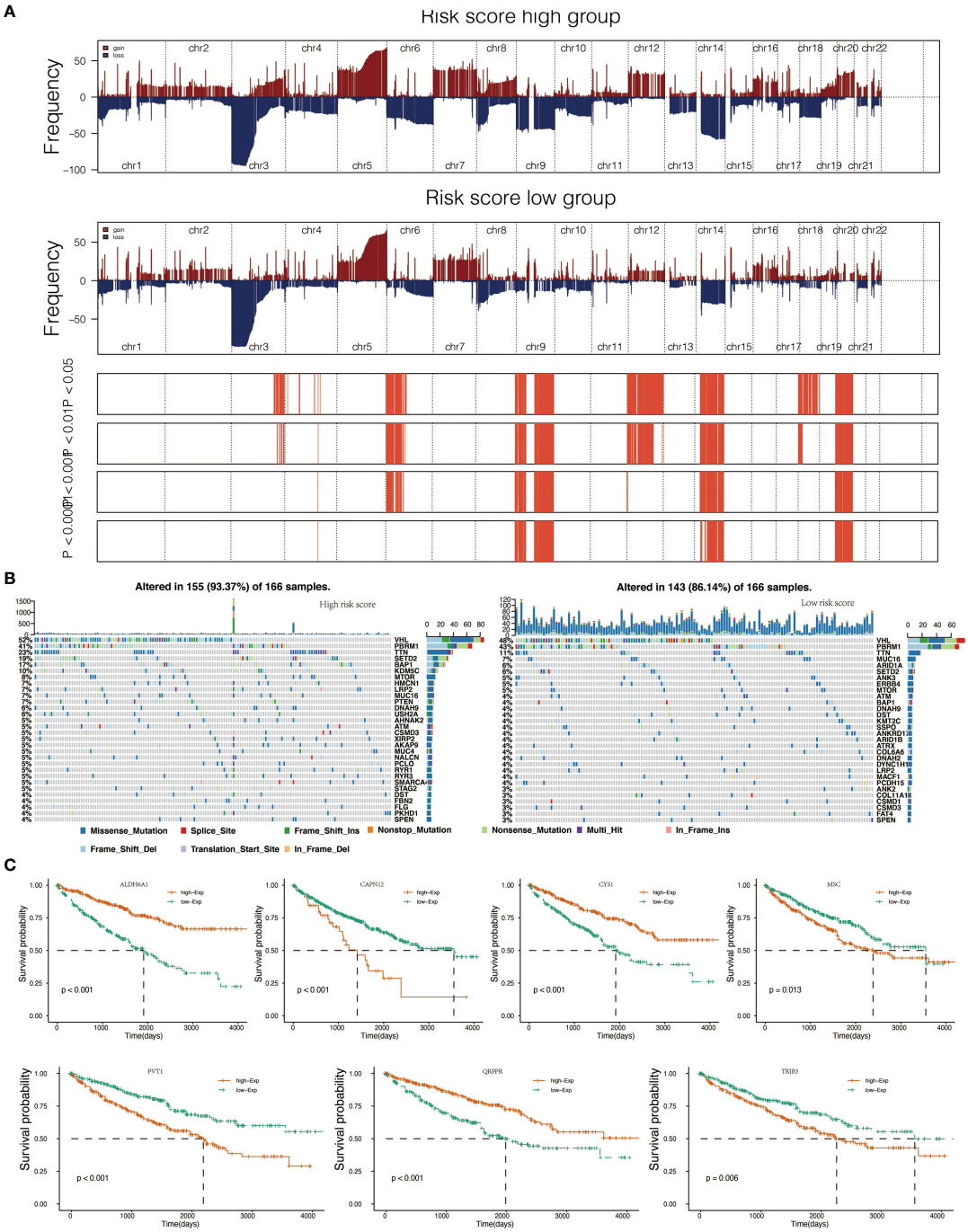


FIGURE 3
Genomic alterations in score low vs. high clusters and DEGs expression in cell lines. **(A)** Copy number variation (CNV) profile in the low score group and CNV profile in the high score group. **(B)** Genomic event enrichment in the low score cluster and genomic event enrichment in the high score cluster. **(C)** Kaplan-Meier overall survival (OS) of patients and expression level of seven prognostic genes.

outcomes, clinical features, and pathological features (Figures 4A, B). In the TCGA dataset, according to the expression similarity, k=3 was selected with clustering stability rising from k=2 to 10 in the TCGA dataset since the consensus cumulative distribution function (CDF) curve was flattest at k=3. Thus, consensus and cluster confidence are also maximal at this k (Figure 4C). The Venn diagram further showed the DEGs among three clusters (Figure 4D). Among the three groups, survival probability is distinctively separated (Figure 4E), which was also confirmed by PCA (Figure 4F).

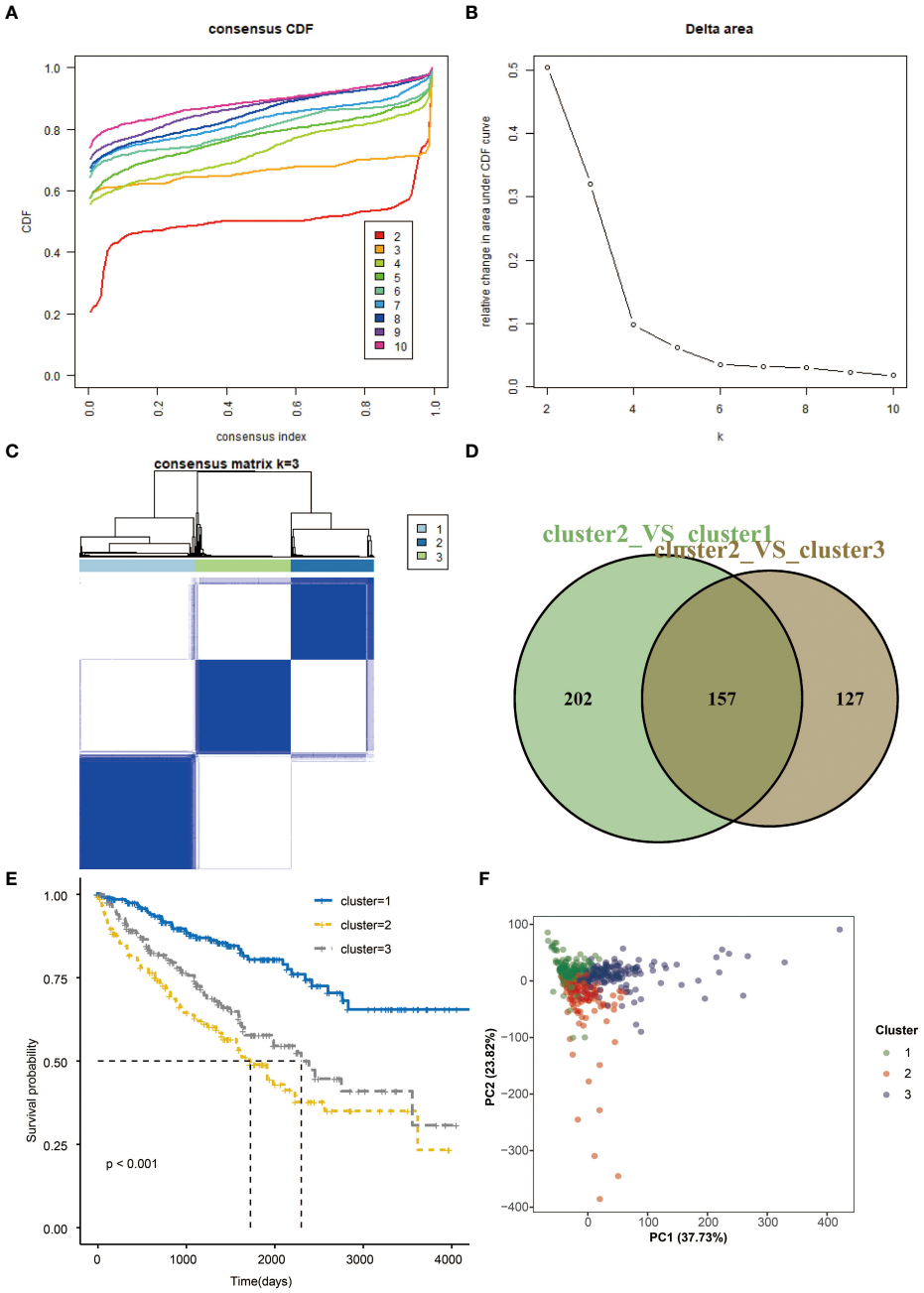


FIGURE 4
Consensus clustering and overall survival in three subgroups. **(A)** Consensus clustering cumulative distribution function (CDF) for k=2 to 10 in TCGA data. **(B)** Relative change in area under CDF curve for k=2 to 10. **(C)** Consensus matrixes of TCGA for each k=3. **(D)** Venn plots the two DEGs groups **(E)** Kaplan-Meier overall survival (OS) curves using TCGA data. **(F)** PCA results for two groups of patients.

Gene set variation analysis and geneset enrichment analysis

To further explore the function of seven prognostic genes, GSVA was conducted using TCGA data (Tables S3, S4). The most enriched GO functions are the regulation of the Wnt signaling pathway, regulation of MAPK cascade, regulation of apoptotic signaling pathway, base excision repair gap filling, positive regulation of T cell apoptotic process, etc. (Figure 5A). Analyses in KEGG pathways revealed that systemic lupus erythematosus, linoleic acid metabolism, regulation of autophagy, Notch signaling pathway, MAPK signaling pathway, Wnt signaling pathway, apoptosis, ERBB signaling pathway, and mTOR signaling pathway were correlated with the seven prognostic genes (Figure 5B). GSEA (Tables S5, S6) further confirmed that the seven predictive genes were enriched in GO pathways such as cytokine activity, humoral immune response, regulation of apoptotic signaling pathway, regulation of Wnt signaling pathway, regulation of signal transduction by p53 class mediator, ERBB signaling pathway, and regulation of Notch signaling

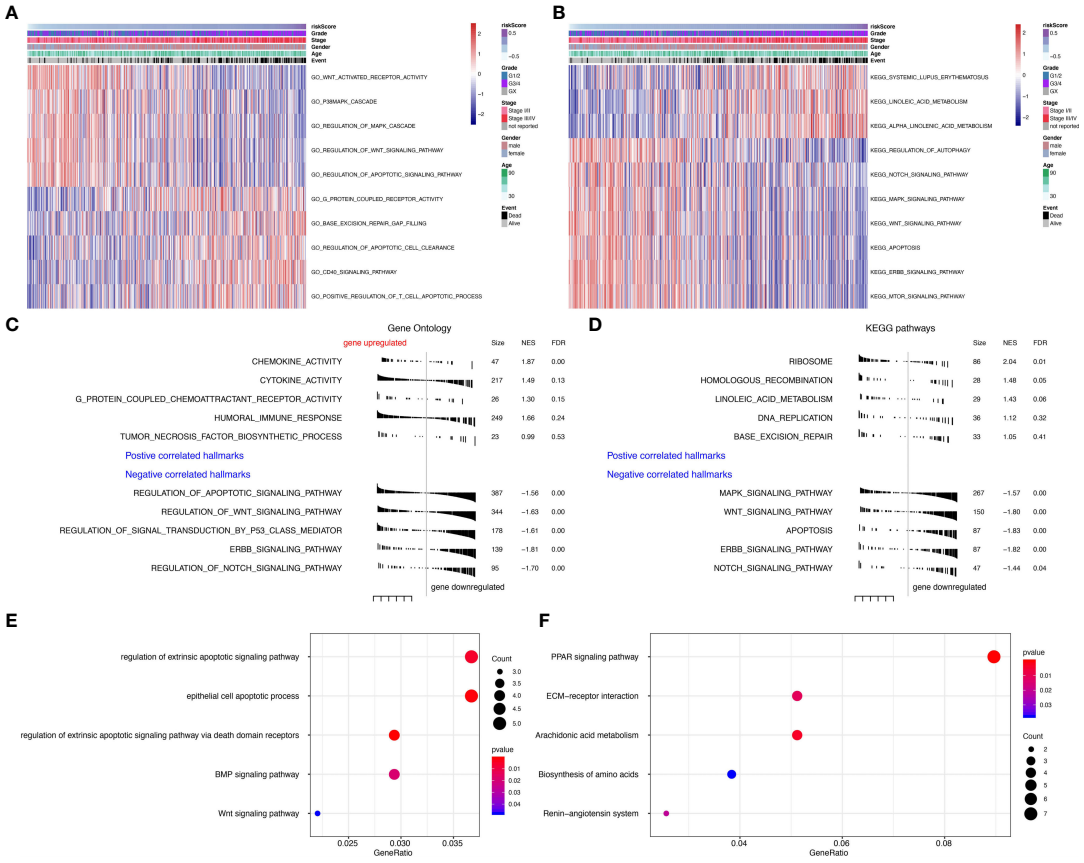


FIGURE 5
Gene set variation analysis (GSVA) and Geneset enrichment analysis (GSEA) in the TCGA dataset. **(A)** Gene Ontology (GO) results based on GSVA in TCGA dataset. **(B)** Kyoto Encyclopedia of Genes and Genomes (KEGG) results based on GSVA in TCGA dataset. **(C)** GO results based on GSEA in TCGA dataset. **(D)** KEGG results based on GSEA in TCGA dataset. **(E)** GO enrichment analyses in TCGA dataset. **(F)** KEGG enrichment analyses in TCGA dataset.

pathway (Figure 5C). As for KEGG pathways, seven prognostic genes were enriched in the ribosome, MAPK signaling pathway, Wnt signaling pathway, apoptosis, ERBB signaling pathway, and Notch signaling pathway (Figure 5D). The dot plot of GO and KEGG enrichment analysis (Tables S7, S8) further revealed that high risk scores were associated with regulation of extrinsic apoptotic signaling pathway, epithelial cell apoptotic process, BMP signaling pathway, and Wnt signaling pathway in GO pathways (Figure 5E), while the risk score was enriched in PRAR signaling pathway, ECM-receptor interaction, arachidonic acid metabolism, biosynthesis of amino acids and the renin-angiotensin system in KEGG pathways (Figure 5F). The correlation between seven prognostic genes and GO pathways was shown in Supplementary Figure S2B, while the correlation between seven predictive genes and KEGG pathways was shown in Supplementary Figure S2C.

Immunological function analyses

The risk scores calculated from prognostic genes are correlated with immune infiltrating cells in the tumor microenvironment (TME). High-risk scores were significantly associated with the relative expression levels of macrophage, MDSC, activated CD4 T cell, activated CD8 T cell, and type 1 T helper cell. In contrast, low-risk scores were correlated with the relative expression levels of immature dendritic cells and neutrophils (Figures 6A, B, correlation > 0.2, P<0.001). Three clusters identified by the seven prognostic genes were also significantly correlated with regulating immune cells in TME (Figure S4A). The correlation between seven predictive genes and immune infiltrating cells was shown in Supplementary Figure S2A, in which seven genes are highly correlated with multiple immune infiltrating cells.

Survival impact of prognostic genes

When comparing survival probabilities between patients with different expression levels of the seven prognostic genes, we found that high ALDH6A1, CYS1, and QRFPR were associated with worse overall survival (OS). In contrast, increased expression of CAPN12, PVT1, MSC, and TRIB3 indicated a better prognosis (Figure 3C). The time-dependent ROC curve at five years of these seven prognostic genes was shown in Figure S1B. We next conducted

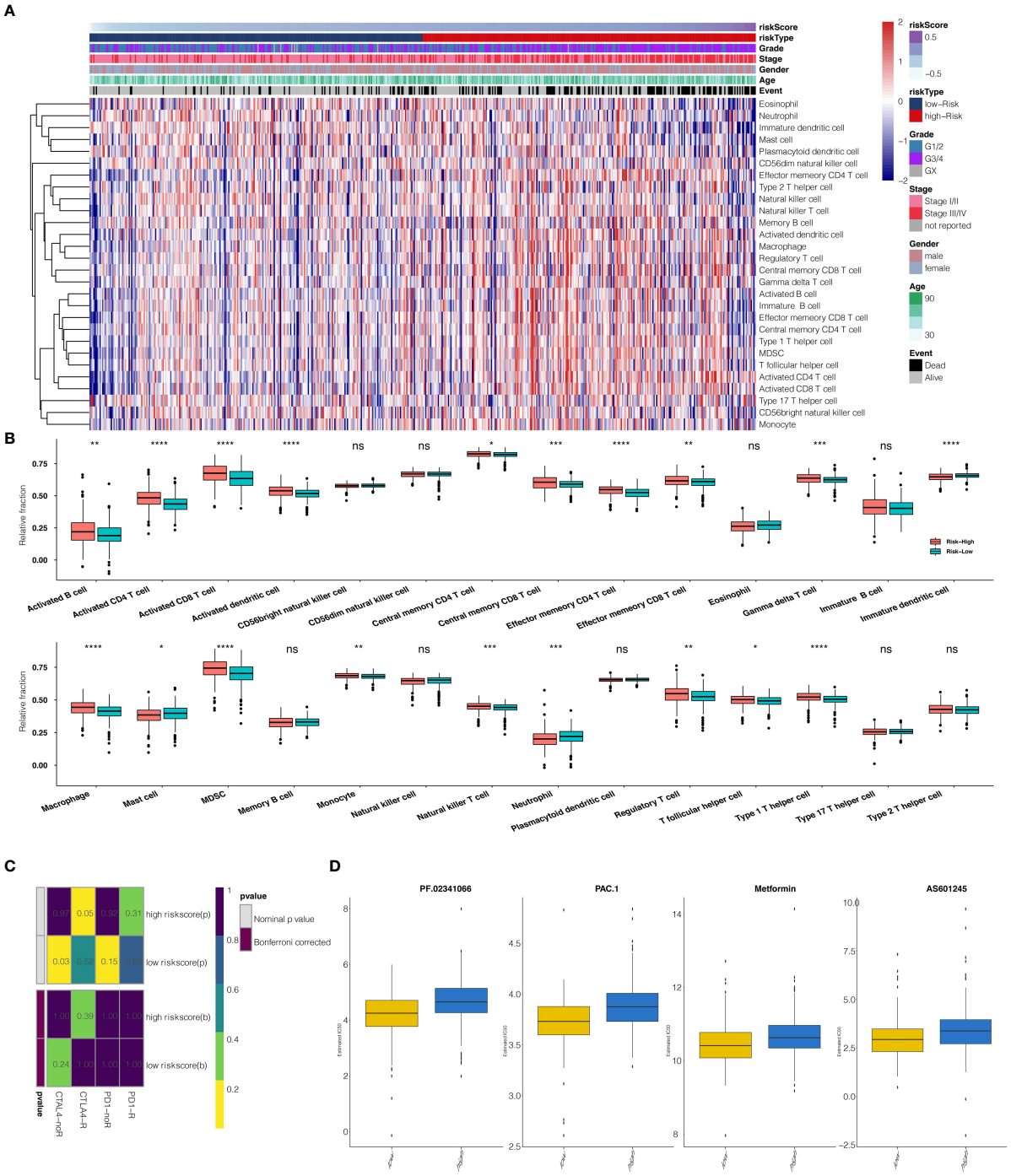


FIGURE 6
(A) The heatmap illustrates the association between risk scores and immune infiltrating cells. **(B)** Correlation between risk score and immune infiltrating cells. **(C)** Submap analysis showed that a high-risk score could be more sensitive to the CTLA-4 inhibitor (Nominal p-value = .05 *P<0.05; **P<0.01; ***P<0.001; ****P<0.0001, ns, not statistically significant). **(D)** The box plots show the estimated IC50 for PF.02341066, PAC.1, Metformin, and AS601245 for high-risk scores and low-risk scores.

the survival analysis of the risk score. High-risk scores were associated with worse OS in different age groups, sex, grade, and stage (Figure S3A). The expression pattern of risk scores in various prognostic factors was shown in Figure 6B, in which high-risk scores were significantly correlated with older patients, male patients, KIRC at grade 4, and KIRC at stage iv. We also revealed that the high-risk scores connected with T4N1M1 KIRC based on the TNM location (Figure S3B). High-risk scores were also related to worse disease-specific survival (DSS) and progressive-free survival (PFS) in the KIRC cohort (Figure S4B). We next verified

the seven prognostic genes in kidney renal papillary cell carcinoma (KIRP), in which high-risk scores indicated worse OS, DSS, and PFS in the KIRP cohort (Figure S4C).

Prediction of risk scores for immunotherapy and chemotherapy

The potential response to immunotherapy in TCGA based on the TIDE algorithm was evaluated, in which our results showed that patients with high-risk scores had a better answer to anti-Cytotoxic T-Lymphocyte Associated Protein (CTLA4) immunotherapy than those with low-risk scores (Nominal p-value = .05) (Figure 6C). Considering that chemotherapy is the standard way to treat ccRCC, we tried to assess the response of patients with different risk scores to various chemo drugs. We could observe a significant difference in the estimated IC50 between high-risk scores and low-risk scores for PF.02341066, PAC.1, Metformin, and AS601245, which low-risk scores could be more sensitive to commonly administered chemotherapies (P <.001 for PF.02341066, PAC.1, Metformin, and AS601245, respectively) (Figure 6D).

Development of the prognostic model with TCGA data

The risk score was subsequently validated as an independent prognostic marker after adjusting for several risk factors, including age group, primary tumor lesion, and stage in univariate and multivariate Cox regression analysis concerning OS, DSS, and PFS (Tables S8, S10, S11. respectively) in the TCGA dataset. The predictive model we built includes risk score, age group, primary tumor lesion, and stage (Figure 7A). At both the three-year and five-year survival, the model had satisfying results in the evaluation nomogram (Figure 7B). Survival difference between high and low-risk patients was statistically significant (Figure 7C). In TCGA data, the AUC at three years is 0.800 and the AUC at five years is 0.788 in the sensitivity test (Figure 7D).

CAPN12 and MSC suppress cell proliferation in ccRCC cells

According to the endogenous CAPN12 and MSC expression level, two independent siRNAs targeting CAPN12 and MSC were

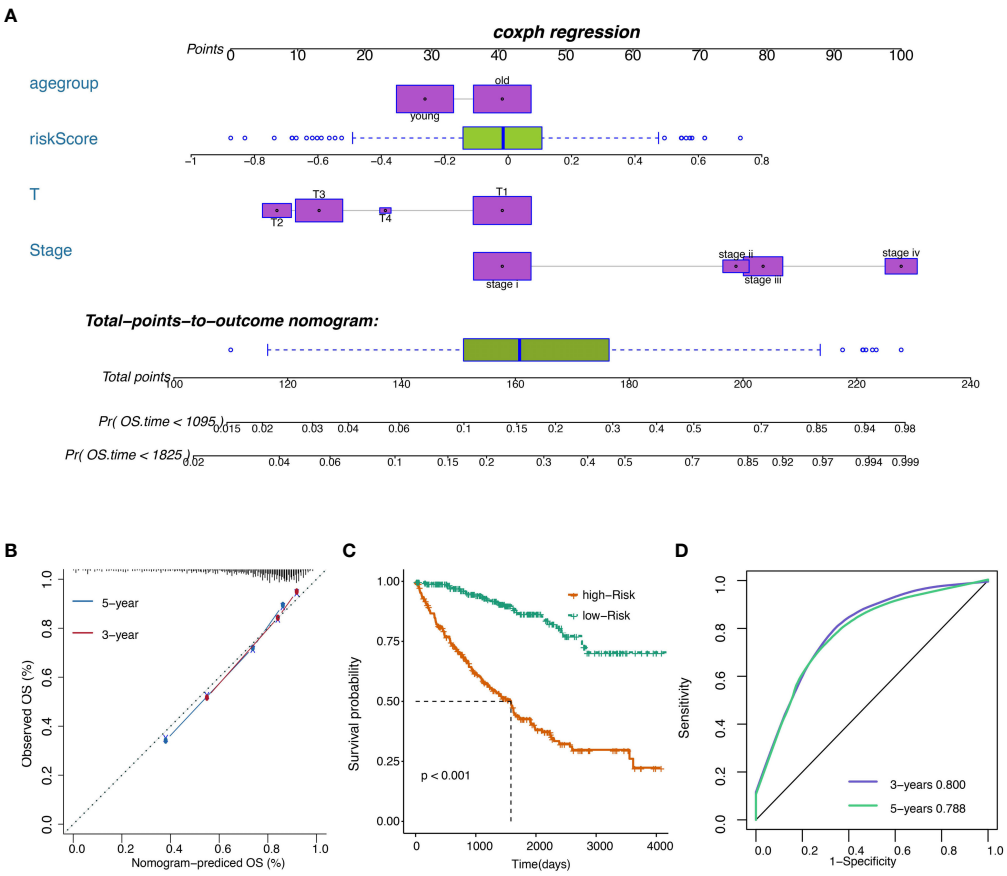


FIGURE 7
(A) Nomogram of the prognostic model. **(B)** Model evaluation results. **(C)** OS of patients with high or low overall risks. **(D)** ROC of the model in TCGA datasets.

transfected into the 786-O cell line with relatively high expression of CAPN12 and MSC. The efficiency of the knockdown of CAPN12 and MSC expression was validated by qRT-PCR (Figure 8A, p<0.001). It was demonstrated that the proliferative capacity of 786-O cells was significantly repressed by CAPN12 and MSC knockdown (Figure 8B). Plate clonality assays revealed the remarkable suppression of cell clonality and cell cycle in the 786-O cell line after silencing CAPN12 and MSC (Figures 8C, D).

Pan-cancer analysis on CAPN12 and MSC

To further explore the prognostic value and immune infiltration pattern of CAPN12 and MSC, pan-cancer samples from TCGA were used for analysis. CAPN12 (Figure 9A) and MSC (Figure 9B) were hazardous markers in most cancer types. Besides, CAPN12 (Figure 10A) and MSC (Figure 10B) correlated with the infiltration of multiple immune cells in most cancer types. These results

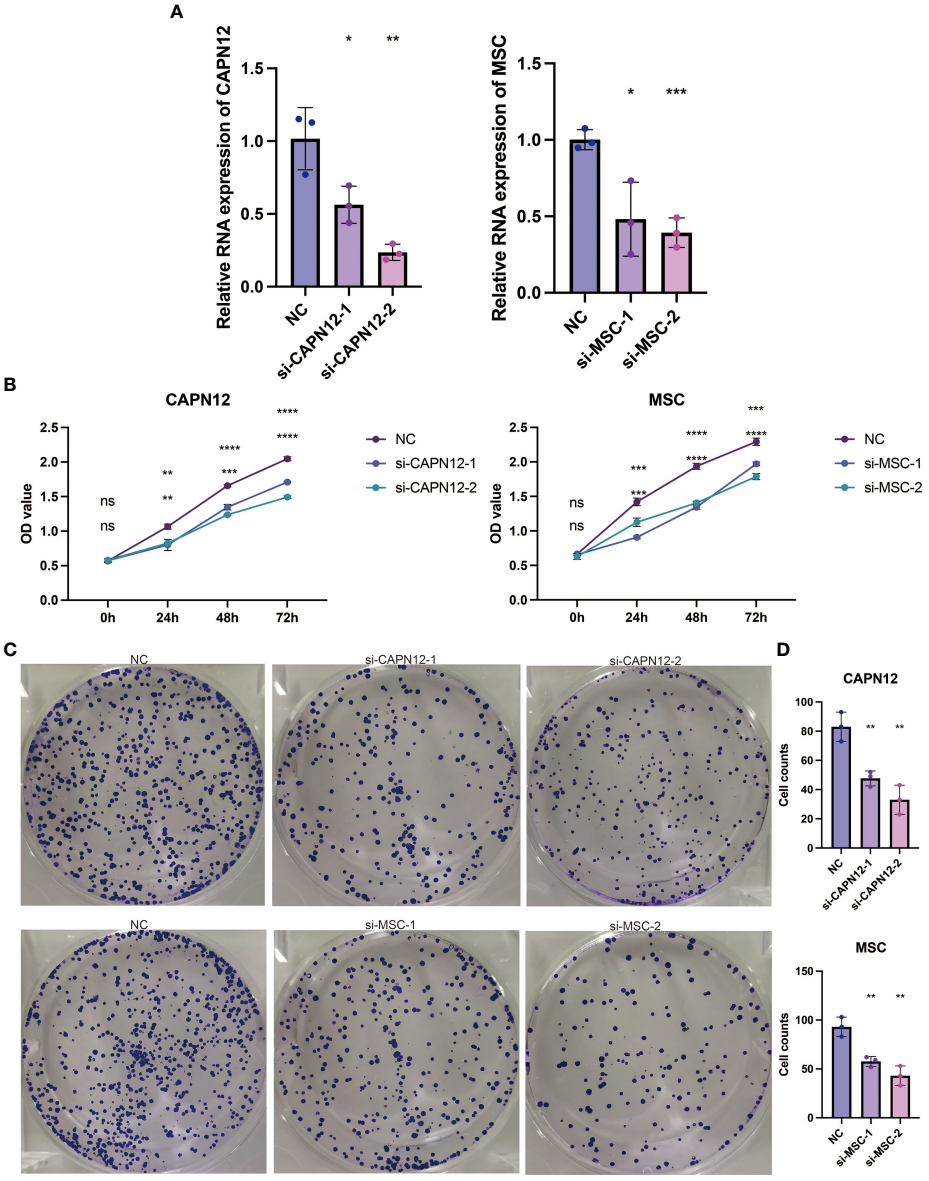


FIGURE 8
(A) qRT-PCR assays for the CAPN12 and MSC levels in 786-O cells transfected with two different siRNAs targeting CAPN12 and MSC (si#1 and is #2), respectively. Tukey HSD test. *P <.05, **P <.01, ***P <.001, ****P <.0001, ns, not statistically significant. **(B)** CAPN12 and MSC knockdown cell proliferation were measured using CCK-8 assay. **(C)** Plate clonality assays measuring the impact on cell clonality and cell cycle in 786-O cell line after silencing CAPN12 and MSC. **(D)** Statistical analysis in plate clonality assay.

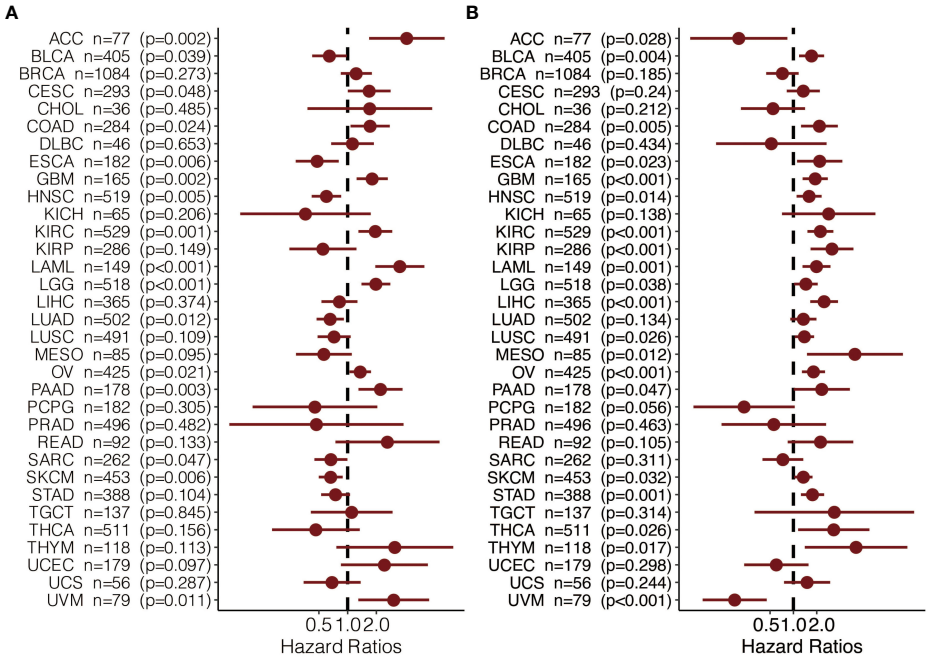


FIGURE 9
The prognostic value of **(A)** CAPN12 and **(B)** MSC in pan-cancer.

suggested that CAPN12 and MSC could be predictive markers of prognosis and immune infiltration in cancer.

Discussion

High mortality and recurrence rates have made ccRCC the most devastating tumor in the urinary system. Previous tic stratification and treatment strategies studies have focused on investigating single potential prognostic biomarkers for ccRCC (17–19). However, none has been immensely satisfying. As high-throughput sequencing and bioinformatics quickly develop, mining the large volume of genetic data has been increasingly appealing to researchers. After data mining, a prognostic model built on genetic profiles of ccRCCs poses significance in developing a prong.

In this study, specifically, after comparing global gene expression in ccRCC samples and controls, 594 DEGs were identified. After univariate and lasso regression analyses, 7 out of 594 DEGs were considered prognostic value: PVT1, MSC, ALDH6A1, TRIB3, QRFPR, CYS1, and CAPN12. Notably, high ALDH6A1, CYS1, and QRFPR were associated with worse OS, while high expressions of CAPN12, PVT1, MSC, and TRIB3 showed statistically significant survival benefits.

Calpains (CAPNs), a family of cysteine proteases, have been demonstrated to play a critical role in cancer development and progression and the insufficient response to cancer therapiesStarsky (20). CAPN12, a gene involved in apoptosis and suppressed by p53, is the critical determinant of anti-tumor response in medulloblastoma (21). Long non-coding RNA plasmacytoma variant translocation 1 (PVT1), up-regulated in various human cancers, inhibits renal cancer cell apoptosis *via* up-regulating Mcl-1 (22) and downregulating miR-16-5p (23). The knockdown of PVT1 induces apoptosis and cell cycle arrest through the epidermal growth factor receptor pathway (24). Multiple studies have also proved that PVT1 predicts unfavorable prognosis in patients with ccRCC (25, 26). MSC, also belonging to the lncRNA family, activates the Wnt/β-catenin signaling pathway to modulate cell proliferation and migration in ccRCC *via* miR-3924/WNT5A (27).

ALDH6A1, regulated by transcription factor HNF4A, has already been verified in other bioinformatics analyses to suppress tumorigenic capability in ccRCC and to be a prognostic biomarker (28, 29).

Tribbles pseudokinase 3 (TRIB3), a member of the mammalian pseudokinase tribbles family, is involved in multiple biological processes, including tumor progression. The previous study has revealed that TRIB3 promoted the proliferation and invasion of ccRCC *via* activating MAPK signaling pathway (30).

QRFPR, also named GPR103, activates glutamine RF–amide peptide (QRFP), is over-expressed in human prostate cancer, and stimulates the neuroendocrine differentiation and the migration of androgen-independent prostate cancer cells (31, 32).

CYS1 mutation on chromosome 2p25 has been proven to be a candidate for recessive cystic kidney disease (33). CAPN12 and MSC were selected for *in vitro* gene silencing among the seven prognostic genes. The cell proliferation assay demonstrated that the proliferative capacity of 786-O cells was significantly repressed by CAPN12 and MSC knockdown, revealing the tumorigenic role of CAPN12 and MSC.

Further geneset variation analysis was conducted in these seven prognostic genes to explore involved signaling pathways. GO analysis revealed that predictive genes are primarily enriched in the Wnt signaling pathway, MAPK cascade, regulation of apoptotic

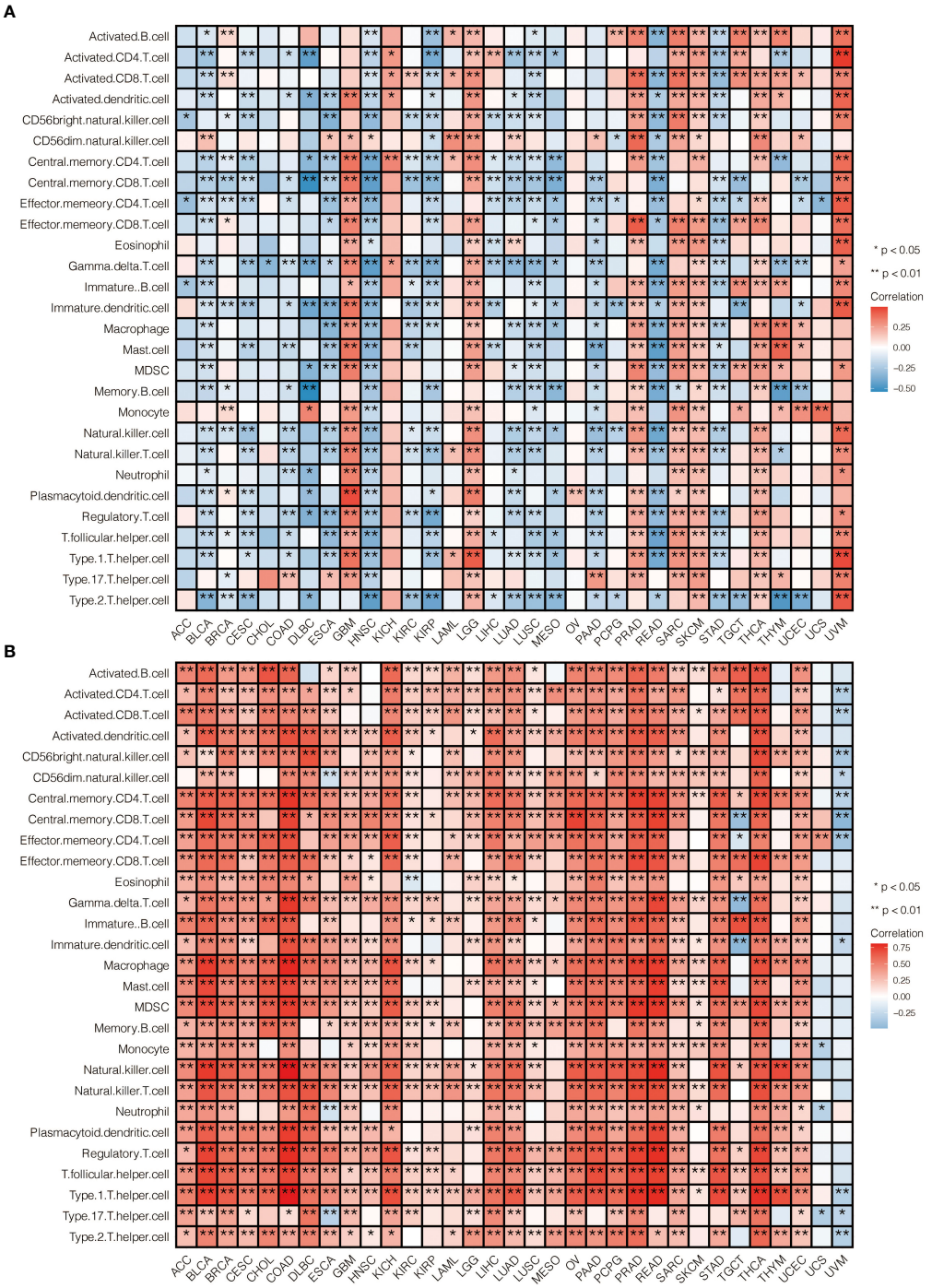


FIGURE 10
The immune infiltration pattern of **(A)** CAPN12 and **(B)** MSC in pan-cancer.

signaling pathway, base excision repair gap filling, positive regulation of T cell apoptotic process, etc. KEGG pathway revealed systemic lupus erythematosus, linoleic acid metabolism, regulation of autophagy, Notch signaling pathway, MAPK signaling pathway, Wnt signaling pathway, apoptosis, ERBB signaling pathway, and mTOR signaling pathway were correlated with the seven prognostic genes. GSEA further confirmed that these seven predictive genes were involved in the tumor-genic process. All these results support the significance of predictive genes in ccRCC.

Hence, risk scores were calculated for each patient based on the seven prognostic genes. When The high-risk group showed a significant survival disadvantage when we separated patients according to the median risk scores contrast, patients with low-risk scores had better responses to chemotherapy. The risk score was further validated in the TCGA test set, TCGA sum set, and GEO data set. High-risk patients showed significantly worse survival in all data sets than low-risk patients. Before we developed a prognostic model, consensus clustering was adopted

to evaluate the predictive genes, and the clustering findings suggested that predictive genes are closely related to survival probability. Patients with high-risk scores also had infiltrating immune cell levels similar to those in cluster 2. Given that the increased risk score group and cluster 2 predicted worse survival, the validity of these genes was supported from another aspect.

Moreover, the risk score was correlated with immune cell expression. High-risk scores were significantly associated with the macrophage, MDSC, activated CD4 T cell, activated CD8 T cell, and type 1 T helper cell expression.

In contrast, low-risk scores were associated with immature dendritic cells and neutrophils, which implicates a suppression in both the innate and acquired immune response system. This finding would open a gate to targeting the immune system to fight ccRCC. Though the activated CD4 T cell and activated CD8 T cell expression increased under such a situation, it could represent positive feedback from a tumor attack.

Interestingly, when examining the genomic alteration profiles of low- and high-risk groups, we found that VHL expression was much higher in the low-risk group. Since VHL plays a tumor-suppressing role, this connection validates the value of the calculated risk score from another perspective. However, further studies are in need to explore the causality in between.

Next, a prognostic model containing risk score, age group, primary tumor lesion, and stage was developed satisfyingly. The remarkable ROC value indicates that the predictive model could be an essential predicting tool. Although ccRCC patient overall survival is influenced by age group, primary tumor lesion, and stage, our risk score adds value to disease prognosis independently by categorizing patients into groups with distinct survival probability. Notably, the high-risk score was significantly correlated with male patients, KIRC at grade 4, KIRC at stage iv, and T4N1M1 based on the TNM location.

Conclusion

Some prognostic models(Wang et al., 2019; Zhang et al., 2019) focus on various ccRCC prognostic factors, such as DNA methylation-driven genes and metastasis-associated predictive genes. However, only some models are acknowledged as a golden standard due to the complex nature of ccRCC, which leaves much space for further research. Seven prognostic genes were eliminated from this study's analysis of data from two databases (TCGA and GEO), all of which were most likely to be highly linked with the onset and progression of ccRCC. We subsequently conducted extensive investigations to create a full prognostic model for ccRCC patients, offering a reliable signature for prognosis prediction and supporting data for drug discovery against these predictive genes. Although knockdown cell RNA-seq was not performed to examine expression profiles and the knockout specificity, qRT-PCR was used to confirm the knockout of CAPN12 and MSC. Additionally, further investigation based on a large cohort is required to fully understand the exciting findings that T2, T3, and T4 were linked to better outcomes (HR 1). Due to this limitation, there is still space for additional validation.

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

GX conducted the data analysis. SW provided funding support, and XC supervised and revised the study. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Science and Technology Innovation Leading Plan of High Tech Industry in Hunan Province (2020SK2011).

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.2023.1161666/full#supplementary-material

SUPPLEMENTARY TABLE 1
The list of differentially expressed genes (DEGs).

SUPPLEMENTARY TABLE 2
LASSO results of prognostic genes.The list of differentially expressed genes (DEGs).

SUPPLEMENTARY TABLE 3
Gene set variation analysis (GSVA) of the risk score in GO.

SUPPLEMENTARY TABLE 4
Gene set variation analysis (GSVA) of the risk score in KEGG.

SUPPLEMENTARY TABLE 5
Gene set enrichment analysis (GSEA) of the risk score in GO

SUPPLEMENTARY TABLE 6
Gene set enrichment analysis (GSEA) of the risk score in KEGG

SUPPLEMENTARY TABLE 7
GO enrichment analysis of risk score.

SUPPLEMENTARY TABLE 8
KEGG enrichment analysis of risk score

SUPPLEMENTARY TABLE 9
Univariate and multivariate analysis regarding OS in the prognostic model.

SUPPLEMENTARY TABLE 10
Univariate and multivariate analysis regarding DSS in the prognostic model.

SUPPLEMENTARY TABLE 11
Univariate and multivariate analysis regarding PFS in the prognostic model.

SUPPLEMENTARY TABLE 12
The relative abundance of 28 immune cell types in each sample.

SUPPLEMENTARY TABLE 13
Significantly different CNV regions of risk score.

SUPPLEMENTARY TABLE 14
Significantly different mutated genes of high and low-risk score groups.

References

1. Sanchez DJ, Simon MC. Genetic and metabolic hallmarks of clear cell renal cell carcinoma. *Biochim Biophys Acta Rev Cancer.* (2018) 1870:23–31. doi: 10.1016/j.bbcan.2018.06.003

2. Hakimi AA, Voss MH, Kuo F, Sanchez A, Liu M, Nixon BG, et al. Transcriptomic profiling of the tumor microenvironment reveals distinct subgroups of clear cell renal cell cancer: Data from a randomized phase III trial. *Cancer Discovery* (2019) 9:510–25. doi: 10.1158/2159-8290.CD-18-0957

3. Atkins MB, Tannir NM. Current and emerging therapies for first-line treatment of metastatic clear cell renal cell carcinoma. *Cancer Treat Rev* (2018) 70:127–37. doi: 10.1016/j.ctrv.2018.07.009

4. Mitchell TJ, Turajlic S, Rowan A, Nicol D, Farmery JHR, O'Brien T, et al. Timing the landmark events in the evolution of clear cell renal cell cancer: TRACERx renal. *Cell* (2018) 173:611–23 e17. doi: 10.1016/j.cell.2018.02.020

5. Ravaud A, Motzer RJ, Pandha HS, George DJ, Pantuck AJ, Patel A, et al. Adjuvant sunitinib in high-risk renal-cell carcinoma after nephrectomy. *N Engl J Med* (2016) 375:2246–54. doi: 10.1056/NEJMoa1611406

6. Qian CN, Huang D, Wondergem B, Teh BT. Complexity of tumor vasculature in clear cell renal cell carcinoma. *Cancer* (2009) 115:2282–9. doi: 10.1002/cncr.24238

7. Li JK, Chen C, Liu JY, Shi JZ, Liu SP, Liu B, et al. Long noncoding RNA MRCCAT1 promotes metastasis of clear cell renal cell carcinoma *via* inhibiting NPR3 and activating p38-MAPK signaling. *Mol Cancer.* (2017) 16:111. doi: 10.1186/s12943-017-0681-0

8. Syafruddin SE, Rodrigues P, Vojtasova E, Patel SA, Zaini MN, Burge J, et al. A KLF6-driven transcriptional network links lipid homeostasis and tumour growth in renal carcinoma. *Nat Commun* (2019) 10:1152. doi: 10.1038/s41467-019-09116-x

9. Miao D, Margolis CA, Gao W, Voss MH, Li W, Martini DJ, et al. Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science* (2018) 359:801–6. doi: 10.1126/science.aan5951

10. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43:e47. doi: 10.1093/nar/gkv007

11. Goeman JJ. L1 penalized estimation in the cox proportional hazards model. *Biom J* (2010) 52:70–84. doi: 10.1002/bimj.200900028

12. Hanzelmann S, Castelo R, Guinney J. GSVA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf* (2013) 14:7. doi: 10.1186/1471-2105-14-7

13. 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

14. 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

15. Chen B, Liu J, Ho TT, Ding X, Mo YY. ERK-mediated NF-kappaB activation through ASIC1 in response to acidosis. *Oncogenesis* (2016) 5:e279. doi: 10.1038/oncsis.2016.81

16. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: An open-source package for r and s+ to analyze and compare ROC curves. *BMC Bioinf* (2011) 12:77. doi: 10.1186/1471-2105-12-77

17. Hansson J, Lindgren D, Nilsson H, Johansson E, Johansson M, Gustavsson L, et al. Overexpression of functional SLC6A3 in clear cell renal cell carcinoma. *Clin Cancer Res* (2017) 23:2105–15. doi: 10.1158/1078-0432.CCR-16-0496

18. Liu Y, Cheng G, Song Z, Xu T, Ruan H, Cao Q, et al. RAC2 acts as a prognostic biomarker and promotes the progression of clear cell renal cell carcinoma. *Int J Oncol* (2019) 55:645–56. doi: 10.3892/ijo.2019.4849

19. Martignoni G, Brunelli M, Segala D, Munari E, Gobbo S, Cima L, et al. Validation of 34betaE12 immunoexpression in clear cell papillary renal cell carcinoma as a sensitive biomarker. *Pathology* (2017) 49:10–8. doi: 10.1016/j.pathol.2016.05.014

20. Starska K, Forma E, Jozwiak P, Lewy-Trenda I, Danilewicz M, Stasikowska-Kanicka O, et al. Gene/protein expression of CAPN1/2-CAST system members is associated with ERK1/2 kinases activity as well as progression and clinical outcome in human laryngeal cancer. *Tumour Biol* (2016) 37:13185–203. doi: 10.1007/s13277-016-5178-8

21. Waye S, Naeem A, Choudhry MU, Parasido E, Tricoli L, Sivakumar A, et al. The p53 tumor suppressor protein protects against chemotherapeutic stress and apoptosis in human medulloblastoma cells. *Aging (Albany NY).* (2015) 7:854–68. doi: 10.18632/aging.100831

22. Wu Q, Yang F, Yang Z, Fang Z, Fu W, Chen W, et al. Long noncoding RNA PVT1 inhibits renal cancer cell apoptosis by up-regulating mcl-1. *Oncotarget* (2017) 8:101865–75. doi: 10.18632/oncotarget.21706

23. Ren Y, Huang W, Weng G, Cui P, Liang H, Li Y. LncRNA PVT1 promotes proliferation, invasion and epithelial-mesenchymal transition of renal cell carcinoma cells through downregulation of miR-16-5p. *Onco Targets Ther* (2019) 12:2563–75. doi: 10.2147/OTT.S190239

24. Li W, Zheng Z, Chen H, Cai Y, Xie W. Knockdown of long non-coding RNA PVT1 induces apoptosis and cell cycle arrest in clear cell renal cell carcinoma through the epidermal growth factor receptor pathway. *Oncol Lett* (2018) 15:7855–63. doi: 10.3892/ol.2018.8315

25. Bao X, Duan J, Yan Y, Ma X, Zhang Y, Wang H, et al. Upregulation of long noncoding RNA PVT1 predicts unfavorable prognosis in patients with clear cell renal cell carcinoma. *Cancer biomark* (2017) 21:55–63. doi: 10.3233/CBM-170251

26. Posa I, Carvalho S, Tavares J, Grosso AR. A pan-cancer analysis of MYC-PVT1 reveals CNV-unmediated deregulation and poor prognosis in renal carcinoma. *Oncotarget* (2016) 7:47033–41. doi: 10.18632/oncotarget.9487

27. Hu Z, Li L, Cheng P, Liu Q, Zheng X, Peng F, et al. lncRNA MSC-AS1 activates wnt/beta-catenin signaling pathway to modulate cell proliferation and migration in kidney renal clear cell carcinoma *via* miR-3924/WNT5A. *J Cell Biochem* (2020). 121 (10):4085–93 doi: 10.1002/jcb.29594

28. Chen B, Jiao Z, Yin X, Qian Z, Gu J, Sun H. Novel insights into biomarkers associated with renal cell carcinoma. *Oncol Lett* (2018) 16:83–90. doi: 10.3892/ol.2018.8665

29. Lu J, Chen Z, Zhao H, Dong H, Zhu L, Zhang Y, et al. ABAT and ALDH6A1, regulated by transcription factor HNF4A, suppress tumorigenic capability in clear cell renal cell carcinoma. *J Transl Med* (2020) 18:101. doi: 10.1186/s12967-020-02268-1

30. Hong B, Zhou J, Ma K, Zhang J, Xie H, Zhang K, et al. TRIB3 promotes the proliferation and invasion of renal cell carcinoma cells *via* activating MAPK signaling pathway. *Int J Biol Sci* (2019) 15:587–97. doi: 10.7150/ijbs.29737

31. Alonzeau J, Alexandre D, Jeandel L, Courel M, Hautot C, El Yamani FZ, et al. The neuropeptide 26RFa is expressed in human prostate cancer and stimulates the neuroendocrine differentiation and the migration of androgeno-independent prostate cancer cells. *Eur J Cancer.* (2013) 49:511–9. doi: 10.1016/j.ejca.2012.05.028

32. Kawan MA, Kyrou I, Ramanjaneya M, Williams K, Jeyaneethi J, Randeva HS, et al. Involvement of the glutamine RFamide peptide and its cognate receptor GPR103 in prostate cancer. *Oncol Rep* (2019) 41:1140–50. doi: 10.3892/or.2018.6893

33. Fliegauf M, Frohlich C, Horvath J, Olbrich H, Hildebrandt F, Omran H. Identification of the human CYS1 gene and candidate gene analysis in boichis disease. *Pediatr Nephrol.* (2003) 18:498–505. doi: 10.1007/s00467-003-1141-1

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 23 March 2023
DOI 10.3389/fimmu.2023.1161960

OPEN ACCESS

EDITED BY
Meng Zhou,
Wenzhou Medical University, China

REVIEWED BY
Quan Zhu,
Xiangya Hospital, Central South University, China
Haoyu Chen,
The Chinese University of Hong Kong, China

*CORRESPONDENCE
Bo Qin
qinbozf@163.com
Chang Zou
zouchang@cuhk.edu.cn

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

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Immunology

RECEIVED 09 February 2023
ACCEPTED 13 March 2023
PUBLISHED 23 March 2023

CITATION
Tan Y, Pan J, Deng Z, Chen T, Xia J, Liu Z, Zou C and Qin B (2023) Monoacylglycerol lipase regulates macrophage polarization and cancer progression in uveal melanoma and pan-cancer. *Front. Immunol.* 14:1161960. doi: 10.3389/fimmu.2023.1161960

Monoacylglycerol lipase regulates macrophage polarization and cancer progression in uveal melanoma and pan-cancer

Yao Tan[1†], Juan Pan[2,3†], Zhenjun Deng[4,5†], Tao Chen[6], Jinquan Xia[3], Ziling Liu[1,7], Chang Zou[8*] and Bo Qin[1,9*]

[1]Shenzhen Aier Eye Hospital, Aier Eye Hospital, Jinan University, Shenzhen, China, [2]National Center for International Research of Bio-targeting Theranostics, Guangxi Key Laboratory of Bio-targeting Theranostics, Collaborative Innovation Center for Targeting Tumor Diagnosis and Therapy, Guangxi Talent Highland of Bio-targeting Theranostics, Guangxi Medical University, Nanning, Guangxi, China, [3]Department of Clinical Medical Research Center, The Second Clinical Medical College, The First Affiliated Hospital of Southern University of Science and Technology, Jinan University (Shenzhen People's Hospital), Shenzhen, Guangdong, China, [4]Department of Dermatology, The Second Clinical Medical College, Jinan University (Shenzhen People's Hospital), Shenzhen, China, [5]The First Affiliated Hospital, Jinan University, Guangzhou, China, [6]School of Medicine, The First Affiliated Hospital, Southern University of Science and Technology, Shenzhen, China, [7]Institute of Biopharmaceutical and Health Engineering, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, China, [8]School of Life and Health Sciences, The Chinese University of Kong Hong, Shenzhen, China, [9]Shenzhen Aier Ophthalmic Technology Institute, Shenzhen, China

Background: Although lipid metabolism has been proven to play a key role in the development of cancer, its significance in uveal melanoma (UM) has not yet been elucidated in the available literature.

Methods: To identify the expression patterns of lipid metabolism in 80 UM patients from the TCGA database, 47 genes involved in lipid metabolism were analyzed. Consensus clustering revealed two distinct molecular groups. ESTIMATE, TIMER, and ssGSEA analyses were done to identify the differences between the two subgroups in tumor microenvironment (TME) and immune state. Using Cox regression and Lasso regression analysis, a risk model based on differentially expressed genes (DEGs) was developed. To validate the expression of monoacylglycerol lipase (MGLL) and immune infiltration in diverse malignancies, a pan-cancer cohort from the UCSC database was utilized. Next, a single-cell sequencing analysis on UM patients from the GEO data was used to characterize the lipid metabolism in TME and the role of MGLL in UM. Finally, *in vitro* investigations were utilized to study the involvement of MGLL in UM.

Results: Two molecular subgroups of UM patients have considerably varied survival rates. The majority of DEGs between the two subgroups were associated with immune-related pathways. Low immune scores, high tumor purity, a low number of immune infiltrating cells, and a comparatively low immunological state were associated with a more favorable prognosis. An examination of GO and KEGG data demonstrated that the risk model based on genes involved with lipid metabolism can accurately predict survival in patients with UM. It has been demonstrated that MGLL, a crucial gene in this paradigm, promotes the

proliferation, invasion, and migration of UM cells. In addition, we discovered that MGLL is strongly expressed in macrophages, specifically M2 macrophages, which may play a function in the M2 polarization of macrophages and M2 macrophage activation in cancer cells.

Conclusion: This study demonstrates that the risk model based on lipid metabolism may be useful for predicting the prognosis of patients with UM. By promoting macrophage M2 polarization, MGLL contributes to the evolution of malignancy in UM, suggesting that it may be a therapeutic target for UM.

KEYWORDS

uveal melanoma (UM), lipid metabolism, cancer prognosis, tumor microenvironment (TME), macrophage polarization, monoacylglycerol lipase (MGLL)

1 Introduction

Uveal melanoma (UM) is the most prevalent primary intraocular malignant tumor and the second most prevalent kind of malignant melanoma (1), originating in the iris, choroid, and ciliary body (2, 3). Although UM and cutaneous melanoma are melanocyte-derived malignant tumors, UM has distinct clinical and biological characteristics (4). UM, a rare malignancy is most common in non-Hispanic whites with lighter skin and blue eyes (5). A recent meta-analysis revealed that the incidence rates in North America, Europe, and Asia were 5.74 (95% CI: 4.37-7.11), 7.3 (95% CI: 6.36-8.24), and 0.53 (95% CI: 0.31-0.74) respectively (6). The onset of UM is associated with some risk factors including fair skin color, light eye color, ability to tan, oculodermal melanocytosis, nevi, and BRCA1-associated protein 1 (BAP1) mutation (7). Nowadays, enucleation and radiotherapy—plaque and proton beam—are the most widely used treatments for UM (8, 9). A gene expression profile study divided UM patients into two kinds (low metastatic risk and high metastatic risk) (10). Only 15% of advanced (metastatic) UM patients have a one-year survival rate, and median survival varies from 4 to 15 months (11). A psychological test found that nearly all UM patients desire to know their survival prognosis at the time the tumor was diagnosed (12). An increasing number of research have been focused to elucidate the genetic and pathological mechanisms involved in UM prognosis, however precise prognostication for patients is far from unattainable (13, 14). Currently, there is a huge need to investigate important indicators that can provide reassurance to patients with a high chance of better survival or provide counseling, screening, and systemic adjuvant therapy to patients at high risk (15–17). Consequently, it is imperative to identify a risk classification strategy and prognostic genes for the development of personalized therapy for UM patients.

Due to the rapid proliferation of tumor cells and inadequate blood vessels formation, the tumor microenvironment (TME) is characterized by hypoxia, high oxidation, acidity, and malnutrition, therefore tumor cells reshape their microenvironment *via* multiple processes including lipid metabolic reprogramming (18, 19) to sustain unrestricted cell proliferation and survival. Metabolic reprogramming has been considered a hallmark of cancer for its ability to adapt TME, and dysregulation of lipid metabolism has been a focal point of recent research (20), contributing to the progression of various cancers including glioblastomas, prostate cancer, breast cancer, hepatocellular carcinoma, pancreatic cancer (21–25). Numerous studies have demonstrated that alterations in tumor lipid metabolism led to tumor formation and immunosuppression in the TME (26). Increasing evidence suggests a significant role of lipid metabolism in melanoma pathogenesis (27). However, the function of lipid metabolism-related genes (LMRGs) in determining the outcome of UM is not well understood.

In this study, we sought to identify key LMRGs associated with TME in UM and to construct a predictive model for UM. This project seeks to find novel prognostic indicators and therapeutic targets as well as clarify the condition of the tumor immune microenvironment in UM in order to build a molecularly-based technique for predicting survival and treatment advantages for UM patients.

2 Methods

2.1 Data collection

The Cancer Genome Atlas (TCGA) provided UM patients' clinicopathological characteristics and gene expression matrices. The training cohort included 80 UM patients (28). The validation cohort (containing 28 cases from GSE84976) was derived from the Omnibus (GEO) datasets (29). Clinical information included survival time, survival status, gender, age, tumor grade, and stage. Missing clinical information samples were eliminated.

LMRGs were chosen from Gene Set Enrichment Analysis (GSEA) and Kyoto Encyclopedia of Genes and Genomes (KEGG)

databases. GSEA has 17 lipid metabolism-related gene sets, while KEGG has 16. Table S1 lists GSEA and KEGG gene sets. 1168 LMRGs remained for research after eliminating duplication genes.

2.2 Consensus clustering

The R package "ConsensusClusterPlus" (version 1.54.0) was used to classify UM patients into two subgroups based on the 47 LMAG expression matrix (30, 31). LMRGs consensus clustering analysis found the optimal number of clusters, the lowest fraction of ambiguous clusterings, and the best CDF value for k = 2.

2.3 Calculation of microenvironment cell abundance

The "ESTIMATE" R package (version 1.0.13) calculated ESTIMATE scores, immune scores, stromal scores, and tumor purity (32). microenvironment cell populations (MCPs) and immune cells were quantified using transcriptomic data from the "MCPcounter" package in R (version 1.1) (33). Single sample gene set enrichment analysis (ssGSEA) was performed using the R package "GSVA" (version 1.24.0) to evaluate 28 immune infiltrating cell types (34, 35).

2.4 Differential gene expression and functional enrichment analysis

"limma" R package (version 3.40.6) performed differential gene expression analysis. We selected differentially expressed genes (DEGs) using |log2(fold change)| >1 and a false discovery rate (FDR) adjusted $p < 0.05$. The DEGs list was then analyzed using Gene Ontology (GO) and KEGG *via* "clusterProfiler" R package (version 4.4.4). Enrichment analysis and protein-protein interaction (PPI) analyses were done with "Metascape" (36). Using the molecular signature database's "GO biological process" gene set, the "GSVA" R program (version 1.24.0) investigated the two clusters' signaling pathways. GSEA analyzed the significant pathways.

2.5 Construction of the immune-related risk model

The R package "glmnet" (version 4.1-2) and absolute shrinkage and selection operator algorithm (LASSO) analysis generated a risk model based on univariable regression analysis of prognostic genes. The smallest lambda value was ideal. Multivariate Cox regression analysis determined the gene risk model. The following formula calculated the risk score: risk score = - 0.154969330859525 * expression value of ectonucleotide pyrophosphatase 2 (ENPP2) + 0.168756185717411 * expression value of MGLL - 0.491974590575217 * expression value of phospholipase C delta 1 (PLCD1) - 0.329592494818697 * expression value of solute carrier family 44 member 3 (SLC44A3). The lipid-related gene signature risk score evenly divided patients into low-risk and high-risk groups.

Kaplan-Meier analysis in R package "survival" (version 3.2-7) assessed the survival difference between two subgroups. Time-dependent receiver operating characteristic (ROC) curve analysis using the "survivalROC" R package (version 1.0.3) was also used to verify the risk model's prediction accuracy. Figure 1 illustrates the data analysis procedure.

2.6 Pan-cancer analysis of MGLL expression

The pan-cancer data set was obtained from the UCSC database and retrieved monoacylglycerol lipase (MGLL) gene expression data from each sample. Using the TIMER2 database (http://timer.cistrome.org/), pan-cancer macrophage infiltration and MGLL mRNA expression were correlated (37). Reassessed pan-cancer patient immune cell infiltration score *via* the R package "IOBR" (version 0.99.9) QUANTISEQ (38). Survival data were integrated by sample barcode to examine MGLL mRNA expression in pan-cancers. The median value of MGLL expression was utilized to distinguish high- and low-expressing tumor samples. Using the R package "survival" (version 3.2-7), survival time and status were fitted within the two groups. Cox proportional hazards models and log-rank tests examined the correlation between MGLL mRNA expression and overall survival (OS), disease-specific survival (DSS), disease-free survival (DFS), and progression-free survival (PFS). All cancer types are abbreviated in Table S2.

2.7 Single-cell RNA-seq online analysis

The scTIME Portal (http://sctime.sklehabc.com/unicellular/home) is a database with single-cell time-specific analytic tools for

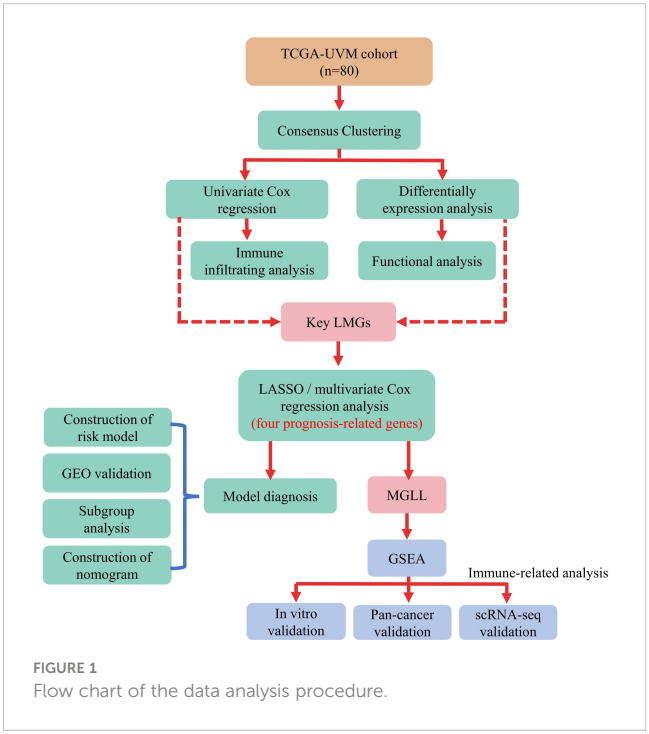


FIGURE 1
Flow chart of the data analysis procedure.

exploring and analyzing TIMEs (39). The "NormalizeData" function, the LogNormalize procedure, and a scale factor of 10,000 were used to normalize the GSE139829 dataset before nonlinear dimension reduction with the "RunUMAP" function and a dims parameter of "1:30". UM patients' cell type proportions and connections were then determined.

2.8 Cell lines and cultures

The human UM cell line MuM-2B (iCELL-h148; Shanghai, China) (40) and the human monocytic leukemia THP-1 cells (#TIB202; ATCC, USA) were grown in RPMI-1640 media containing 10% FBS and 1% penicillin-streptomycin (Gibco, USA) at 37°C and 5% CO2. The adult retinal pigment epithelial cell line (ARPE-19) cells from American Type Culture Collection (ATCC, Manassas, VA, USA) were cultured in DMEM/F12 (Gibco, USA). Shanghai Baoyi Applied Biotechnology Co., Ltd. did a short tandem repeat (STR) analysis. Experiments were done during logarithmic cell growth.

2.9 Monoacylglycerol lipase small interfering RNA construction and transfection

MGLL-knockdown small interfering RNA (siRNA) and negative control siRNA were purchased from Guangzhou RiboBio Co., Ltd. MGLL siRNA was transfected into MuM-2B cells by Lipofectamine RNAi Max (Invitrogen, CA, USA) to transient knockdown MGLL. Western blotting confirmed siRNA inhibition after 48 h of transfection. The siRNA sequence is CCAGGACAAGACTCTCAAGAT (41).

2.10 Cell proliferation assay

Cell proliferation experiment was performed using Cell Counting Kit-8 (CCK-8) (MedChem Express, Monmouth Junction, NJ, USA). 2000 MuM-2B cells per well were seeded into 96-well plates transfected with MGLL or negative control siRNA. After seeding for 24 h, 48 h, 72 h, 96 h, and 120 h, each well received 10 µL CCK-8 solution and was incubated at 37 °C for 1.5 h in the dark. A microtiter plate reader (BIO-TEK Instruments, Winooski, VT, USA) measured live cells at 450 nm.

2.11 Migration and invasion assays

The cell invasion assay used Matrigel (BD Biosciences, Mississauga, Canada), while the cell migration assay did not. $4 \times 10\text{^}4$ cells per well suspended in 200 µL serum-free medium were added to the cell culture insert (24-well insert, 8-µm pore size), and 500 µL 10% FBS-supplemented RPMI-1640 were added to the well to stimulate cell migration or invasion. After 24 hours, the cells in the inserts were removed and the cells that penetrated and attached to the bottom membrane were fixed with 4% paraformaldehyde (PFA) and stained with crystal violet (0.05% [w/v]). A photomicroscope took images in three randomly selected fields in each well.

2.12 Scratch wound healing assay

MuM-2B cells were seeded at $5 \times 10\text{^}5$ cells per well in a 6-well microplate. Scratches were made in the middle of the well with a sterile 200 µL pipette tip when cells reached 95% confluence. Serum-free medium replaced 10% FBS-supplemented RPMI-1640. Photographs were taken at 0 h, 24 h, and 48 h to estimate gap closing. 2.13 RNA extraction and real-time quantitative PCR (RT-qPCR) assays

TRIzol Reagent (Invitrogen, USA) extracted total RNA from cells, and cDNA was generated from 1 µg of RNA using the M-MLV Reverse Transcriptase Kit (Promega, USA) according to the manufacturer's instructions. A Bio-Rad iQ5 RT-qPCR System performed RT-qPCR. GAPDH normalized transcript expression. Table S3 lists the primer sequences used in this study.

2.14 Western blot

The total cell protein was isolated using radio immunoprecipitation assay buffer (RIPA; Beyotime, China) and quantified using BCA Protein Assay Kits (Pierce, Rockford, IL, USA). SDS-PAGE separated the identical protein samples, which were electro-transferred into PVDF membranes (Millipore Corp, Atlanta, GA, USA). After 1 hour of blocking in 5% non-fat milk, anti-MGLL (1:1000 dilution, #A6654, ABclonal, China) and anti-β-actin (#4970, Cell Signaling Technology, USA) primary antibodies were incubated overnight at 4°C. HRP-conjugated secondary antibodies (Cell Signaling Technology, USA) were incubated for 1 h at room temperature. Thermo Fisher ECL reagents detected band signals. It is recommended that strips be washed with stripping buffer (P0025, Beyotime, China) to ensure that the previous antibody has been removed and imaging can be repeated if necessary.

2.15 THP-1 polarization

THP-1 cells were seeded at 1 x 10^6 per well in 6-well plates and treated with PMA (100 nmol; Sigma-Aldrich, St. Louis, MO, USA) for 48 h. M1 macrophages were polarized by incubation with IFN-gamma (20 ng/mL; R&D System, USA) and LPS (100 ng/mL; Sigma, USA) for 48 h. IL-4 (100 ng/ml; PeproTech) was added for 48 h to elicit M2-phenotype polarization.

2.16 Statistical analysis

The data were analyzed using R (version 4.0.3) and GraphPad Prism (version 8.0.1). Univariate and multivariate Cox proportional hazards regression identified independent prognostic factors. After multivariate Cox regression analysis, a nomogram was created to predict 3-year RFS and validated with C-index. Following the regrouping of patients by age, sex, and metastasis, a subgroup analysis was conducted. Student's t-test was used for statistical analysis between two groups, however one-way ANOVA was applied when there were three or more groups. $P < 0.05$ was statistically significant.

3 Results

3.1 Two LMRGs-based molecular subtypes and their prognostic significance

Gene expression profiles and clinical data for 80 UM patients were gathered from the TCGA data portal. First, univariate Cox analysis screened gene sets linked with UM patient survival, then 47 LMRG-associated genes were selected (Table S4). Unsupervised consensus clustering established the optimal number of groups (k = 2) based on the expression patterns of 47 survival-related LMRGs genes (Figures 2A–C). We found two kinds of UM patients: cluster 1 (45, 56.25%) and cluster 2 (35, 43.75%). These survival-related LMRGs genes in the two clusters differed, as shown by the heatmap (Figure 2D). Cluster 2 had a substantially lower OS rate than cluster 1 (Figure 2E, $P < 0.001$). These data showed that LMRGs are greatly linked to UM patients' overall survival.

3.2 Identification of differentially expressed lipid metabolism-related genes and functional annotation

We used the "limma" R tool to compare gene expression between groups. The two clusters had 647 DEGs, 170 upregulated and 477 downregulated (Figure 3A). To elucidate these DEGs' immunity-related functions, GO (Figures 3B, C) and KEGG (Figure 3D) analyses were conducted. Most of these DEGs were involved in antigen processing and presentation, Th1 and Th2 cell differentiation, and other immune-related functions, according to signaling pathway analyses. The PPI analysis identified 14 sub-models, most of which (MCODE1, 2, 5, 7, 8, 10, and 12) were closely related to tumor formation and immunity, suggesting that immunity may contribute to UM through lipid metabolism (Figure 3E; Table S5). GSEA was performed in the two clusters to determine the relationship between enriched pathways and immune cell infiltration in UM patients, which found that T cell receptor signaling pathway, natural killer cell-mediated cytotoxicity, antigen processing and presentation, cytokine-cytokine receptor interaction, and chemokine signaling pathways expressed highly in cluster 1 patients (Figure 3F). These findings suggest that LMRG expression is crucial in the UM TME.

3.3 UM patients in two molecular subtypes exhibited significant differences in TME and immune status

The scoring signature of tumor-infiltrated immune cells can predict immunological treatment response and UM prognosis (42). Thus, we identify the relationship between lipid metabolism and TME in UM. We calculated each subgroup's immune score, ESTIMATE score, and stromal score using the ESTIMATE technique to see if there was an immunological difference. Cluster 2 showed much higher immune scores than cluster 1 (Figure 4A). Furthermore, we estimated immune infiltration in the UM microenvironment using the TIMER database. Cluster 1 had more fibroblasts, but cluster 2 had more T cells, CD8 T cells,

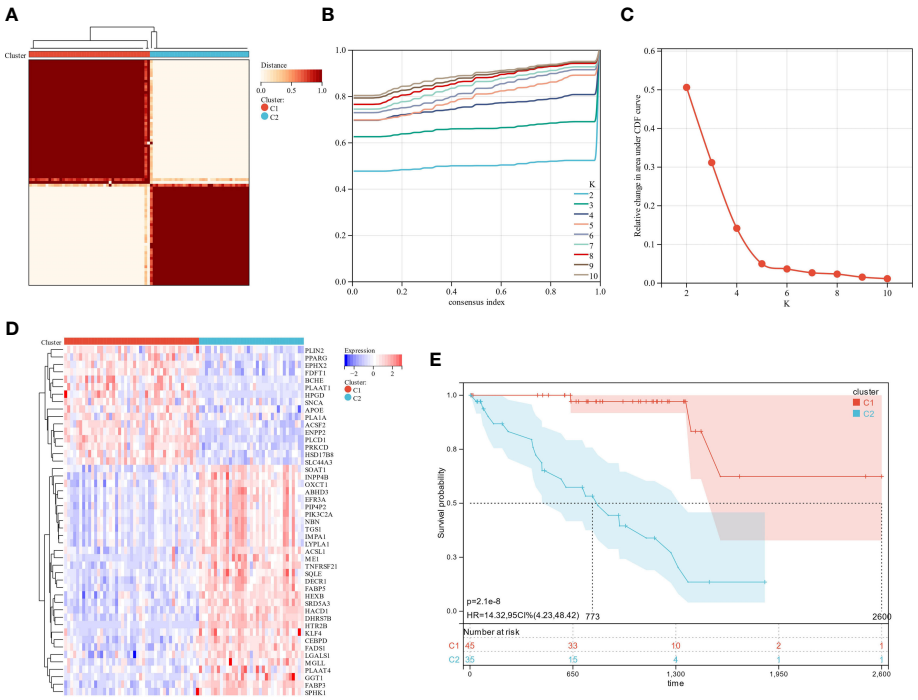


FIGURE 2
The analysis of consensus clustering. **(A–C)** Consensus clustering is best performed with K = 2. **(D)** A heatmap illustrating the differential expression of lipid metabolism genes between the two groups. **(E)** Two subgroups of patients displayed different survival curves.

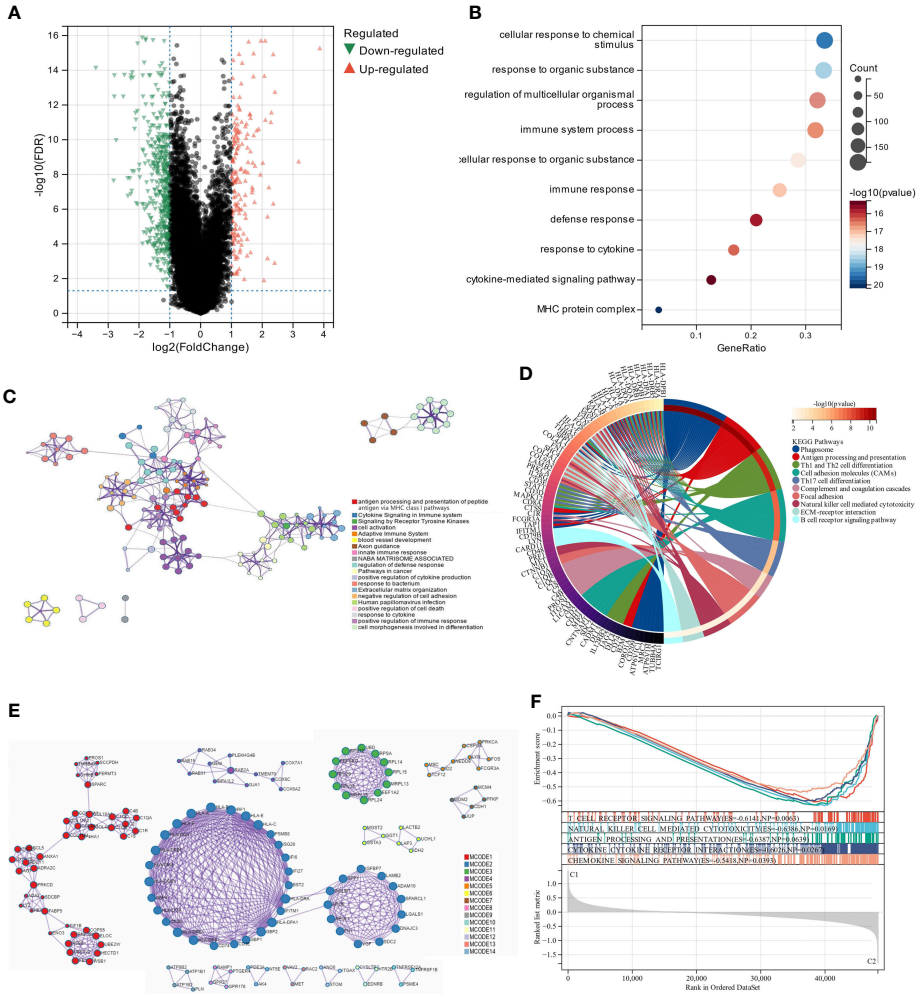


FIGURE 3
The analysis of differentially expressed genes and the evaluation of their functions. **(A)** Volcano plot illustrates the expression of DEGs between the two subgroups. **(B, C)** Visualization of biological processes that have been enhanced by GO analysis using bubble diagrams and networks. **(D)** A circle plot displaying the signaling pathways that KEGG analysis has enriched. **(E)** An analysis of DEGs based on PPI. **(F)** A GSEA plot depicts the signaling pathway analysis.

cytotoxic lymphocytes, NK cells, monocytic lineage, and myeloid dendritic cells (Figure 4B). B lineage, neutrophils, and endothelial cells were not statistically different between these two clusters (Figure 4B). Cluster 1 had a relatively low immune status, while cluster 2 was high (Figures 4C, D). These findings suggest that two molecular subtypes have distinct TME.

3.4 Construction of a risk model

A prognosis prediction model was used to test LMRGs' ability to predict UM prognosis. We extracted four LMRGs in UM] at the minimum likelihood of a deviative pattern ($\lambda_{min} = 0.09$) using Lasso regression on these DEGs (Figures 5A, B). Four genes (ENPP2, MGLL, PLCD1, and SLC44A3) were discovered through multivariate Cox regression. These four genes were utilized to create a risk regression model that categorized UM patients into low- and high-risk groups (Figure 5C). Patients with high risk exhibited significantly shorter survival periods than those with low risk (P = 4.7e-13; Figure 5D). The model's robustness was assessed by plotting ROC curves for 1-year, 3-year, and 5-year OS, with areas under curves (AUCs) of 0.86, 0.95, and 0.93, respectively (Figure 5D), indicating a positive accuracy rate. We used the ESTIMATE algorithm to evaluate the two groups' TME to better understand TME involvement in the UM risk model. The high-risk group had significantly higher stromal, immune, and ESTIMATE scores than the low-risk group (Figure 5E). Generally, high immune cell infiltration in the TME is associated with a positive prognosis, but in the UM it is associated with poor outcomes (42), which is consistent with our findings.

Additionally, the verification cohort was used to validate the created predictive risk score model. The heatmap showed the four candidate genes expression in the verification cohort stratified by risk level (Figure 5F). Survival analysis showed that the high-risk group had a worse prognosis than the low-risk group (P = 1.9e-7; Figure 5G). ROC analysis showed that the risk model predicted 5-year survival best (Figure 5H). Our findings show that the risk score model we created can correctly forecast the prognosis of UM patients.

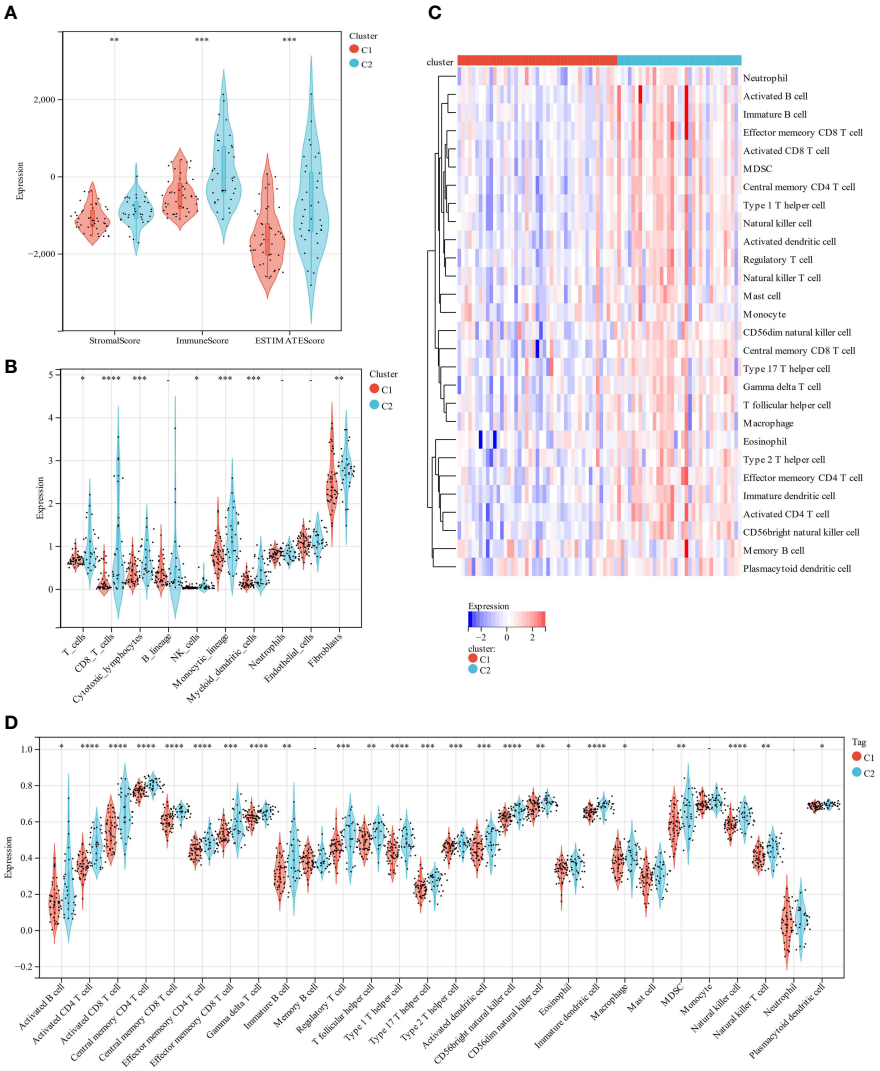


FIGURE 4
A comparison of immune infiltration between the two clustered subgroups. **(A)** The Violin diagram presented the calculation of stromal score, immune score, and estimate score in two subgroups. **(B)** MCP evaluated the abundance of ten immune filtrating cells. **(C)** A heatmap displays the level of enrichment for 28 immune-related cells based on a ssGSEA algorithm. **(D)** Statistical analysis of ssGSEA. *p < 0.05; **p < 0.01; ***p < 0.001 ****p < 0.0001.

3.5 Testing for independence in the constructed risk model

Subgroup analysis and regression analysis were used to evaluate the risk model's independence and the risk score's relevance to clinical features. Risk scores did not differ between patients of different sexes (Figure S1A), ages (Figure S1B), or distant metastases (Figure S1C), indicating no correlation between risk scores and clinical characteristics. The risk model was highly predictive when patients were classified by sex (Figures S1D, E), age (Figures S1F, G), or distant metastasis (Figures S1H, I). The risk model was also an independent predictor of patient prognosis in the univariate Cox regression study (Table 1). These findings demonstrate that the risk model is extremely independent in predicting UM patients' prognoses.

3.6 Construction and calibration of an integrated nomogram combining clinicopathological features and risk signature

Clinical parameters should be considered while predicting UM patients' prognoses, therefore risk score, age, gender, and distant metastases were used to create a prognosis nomogram for UM patients (Figure S2A). The nomogram was validated in the training and verification cohorts using the concordance index (C-index) and calibration curve. The nomogram's C-index in the training group was 0.919 (95% CI, 0.887–0.951, Figure S2B), which matched the verification cohort (Figure S2C), indicating its predictive power. Overall, the nomogram was more accurate in predicting UM patients' prognosis.

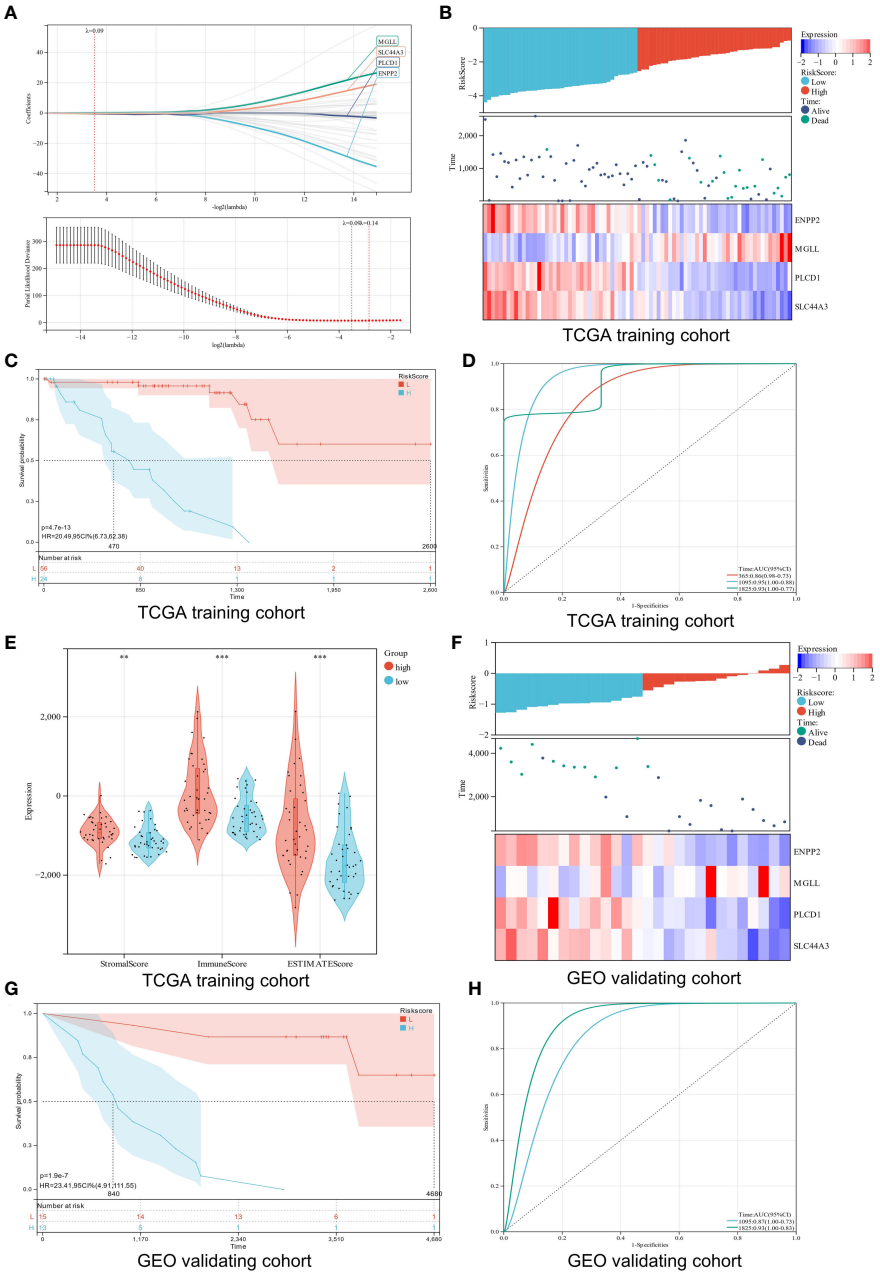


FIGURE 5
Construction and verification of a risk model. **(A)** Analysis using LASSO with a minimal lambda. **(B)** A heatmap displays the survival status and risk score of UM patients based on the expression of four potential genes in two groups. **(C)** Survival curves of UM patients in two different groups. **(D)** A risk model with ROC curves that are time-dependent. **(E)** Stomal score, immune score, and ESTIMATE score are calculated using the ESTIMATE algorithm. **(F)** Four candidate genes were expressed in the verification cohort with survival status and risk score indicated. **(G)** An analysis of the survival curves for high-risk and low-risk patients in the verification cohort. **(H)** The ROC curve for the risk model's verification cohort. **p < 0.01; ***p < 0.001

3.7 MGLL affects the proliferation, migration, and invasion of UM cells *in vitro*

When ENPP2, MGLL, PLCD1, and SLC44A3 mRNA expression levels in ARPE-19 cells and MuM-2B cells were compared, we found that MGLL had significantly higher expression while PLCD1 and SLC44A3 had significantly lower expression (Figure 6A). ENPP2 mRNA expression was not detectable in this study. MuM-2B cells were found to have much higher expression levels of MGLL than ARPE-19 cells (Figure 6B, Figure S3). siRNA-targeted MGLL and siRNA control were transfected into MuM-2B cells to explore the functional role of MGLL in UM. Transfection effectiveness was assessed using Western blotting (Figures 6C, S4). The impact of MGLL on cell proliferation was then assessed using the CCK-8 test, which revealed that MGLL knockdown reduced the proliferation of MuM-2B cells (Figure 6D). Additionally, after reducing the expression of MGLL, MuM-2B cells' capacity for migration and invasion was suppressed (Figures 6E–G, S5–S7). According to these

TABLE 1 Univariate analysis of risk score and characteristics in training cohort.

Characteristics	Total (N)	HR (95% CI) Univariate analysis	P value Univariate analysis	HR (95% CI) Multivariate analysis	P value Multivariate analysis
Age	80	1.046 (1.008 - 1.085)	< 0.05	1.092 (1.019 - 1.171)	< 0.05
Gender	80		0.316		
Male	45	Reference			
Female	35	0.649 (0.274 - 1.536)	0.325		
Clinical T stage	80		0.169		
T2	5	Reference			
T3	36	0.956 (0.116 - 7.864)	0.966		
T4	39	2.163 (0.278 - 16.817)	0.461		
Clinical N stage	80		0.058		
N0	76	Reference		Reference	
NX	4	6.177 (1.302 - 29.304)	< 0.05	8.059 (0.560 - 115.918)	0.125
Clinical M stage	80		< 0.01		
M0	73	Reference		Reference	
M1	3	35.072 (4.689 - 262.335)	< 0.001	0.241 (0.006 - 9.795)	0.452
MX	4	2.226 (0.507 - 9.778)	0.289	0.694 (0.051 - 9.367)	0.783
Clinical stage	80		< 0.001		
Stage II	39	Reference		Reference	
Stage III	37	1.235 (0.504 - 3.028)	0.644	1.074 (0.318 - 3.623)	0.909
Stage IV	4	72.950 (7.056 - 754.174)	< 0.001	20.228 (0.383 - 1066.974)	0.137
Metastasis	80		< 0.001		
Yes	26	Reference		Reference	
No	54	0.044 (0.010 - 0.189)	< 0.001	0.118 (0.015 - 0.956)	< 0.05
RiskScore	80	36.695 (8.799 - 153.035)	< 0.001	12.017 (1.637 - 88.233)	< 0.05

HR, hazard ratio; CI, confidence interval.

findings, MGLL inhibition reduced cell proliferation and migration in UM.

3.8 Pan-cancer MGLL expression and prognosis

Differential expression analysis of pan-cancer samples revealed that MGLL was generally underexpressed in cancers, including bladder urothelial carcinoma (BLCA, $p < 0.001$), breast invasive carcinoma (BRCA, $p < 0.001$), colon adenocarcinoma (COAD, $p < 0.001$), glioblastoma multiforme (GBM, $p < 0.01$), head and neck squamous cell carcinoma (HNSC, $p < 0.001$), kidney renal papillary cell carcinoma (KIRP, $p < 0.01$), liver hepatocellular carcinoma (LIHC, $p < 0.05$), lung adenocarcinoma (LUAD, $p < 0.001$), lung squamous cell carcinoma (LUSC, $p < 0.05$), prostate adenocarcinoma (PRAD, $p < 0.001$), rectum adenocarcinoma (READ, $p < 0.05$), stomach adenocarcinoma (STAD, $p < 0.05$), and uterine corpus endometrial carcinoma (UCEC, $p < 0.001$), with the exception of kidney renal clear cell carcinoma (KIRC, $p < 0.001$) (Figure 7A).

MGLL expression was then evaluated in relation to OS, DSS, and PFS. An analysis of 44 tumors using cox regression revealed that MGLL expression was significantly positively related to the OS of UM patients ($p = 1.5e-6$), acute myeloid leukemia patients (LAML, $p = 9.8e-4$ in TCGA, $p = 5.9e-3$ in TARGET), and pancreatic adenocarcinoma patients (PAAD, $p = 1.2e-3$), but negatively related to the OS of KIRC patients ($p = 1.9e-3$) and sarcoma patients (SARC, $p = 0.03$) patients (Figure 7B). Further cox regression analysis of 38 tumors indicated that MGLL expression significantly correlated with PFS in 5 cancers and was a risk factor for UM ($p = 6.0e-7$), PAAD ($p = 1.7e-3$), STAD ($p = 4.0e-3$), and ACC ($p = 0.03$), but a protective factor for ovarian serous cystadenocarcinoma (OV, p=0.03; Figure 7C). MGLL expression was significantly correlated with DSS in four tumors. For UM ($p = 3.2e-7$) and PAAD patients ($p = 1.1e-3$), MGLL was a protective factor, whereas for KIRC ($p = 2.5e-4$) and SARC patients ($p = 0.01$), it was a risk factor (Figure 7D).

10.3389/fimmu.2023.1161960

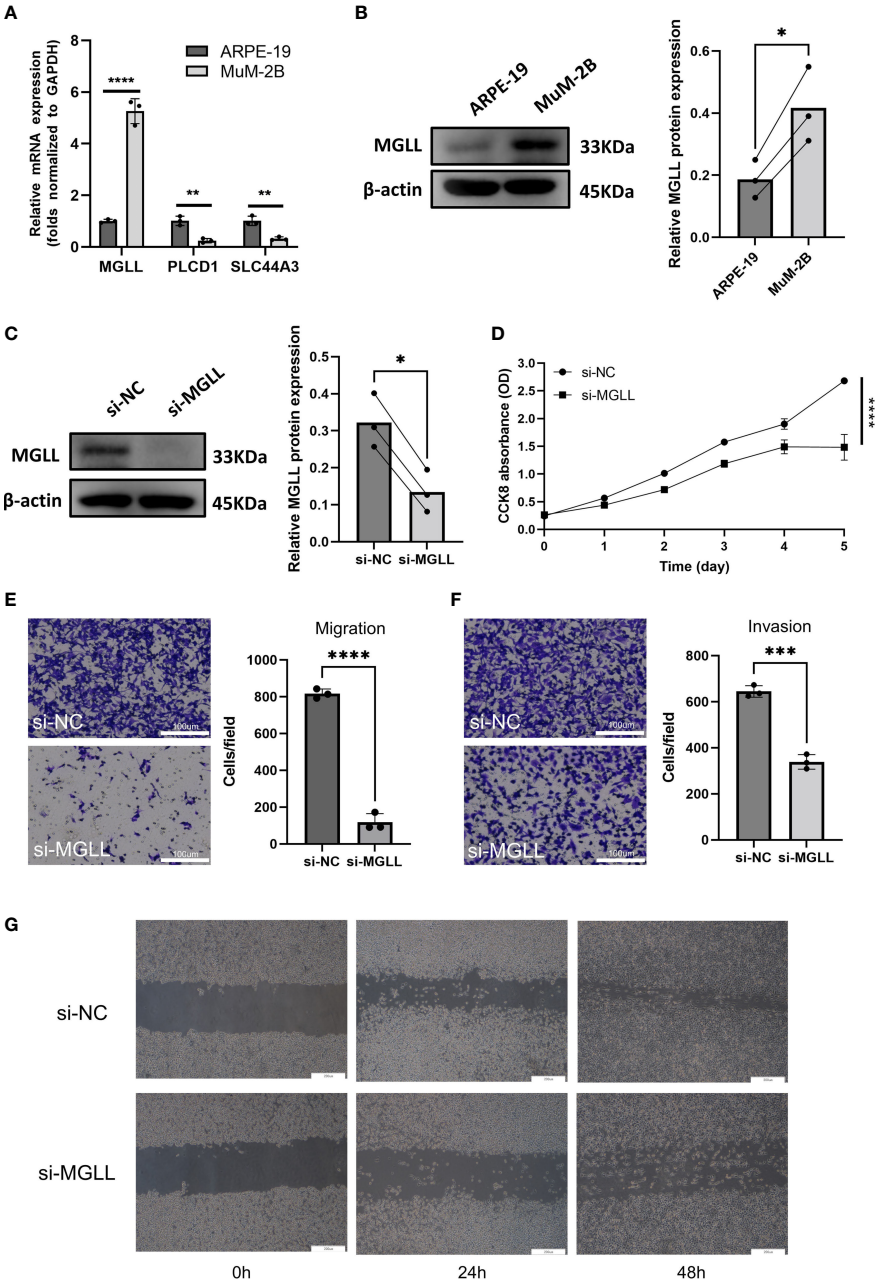


FIGURE 6
MGLL knockdown reduces UM cell migration, invasion, and activity *in vitro*. **(A)** Using qRT-PCR, the expression levels of ENPP2, MGLL, PLCD1, and SLC44A3 in ARPE-19 and MuM-2B cells were determined. ENPP2 mRNA expression was undetectable. Normalization of Ct values to GAPDH was performed. (Student's t-test) **(B)** Using WB, MGLL levels in ARPE-19 and MuM-2B cells were examined, and relative protein expression was adjusted using -actin levels. (left: images indicative of three independent experiments; right: quantitative analysis, n = 3, paired Student's t-test) **(C)** Western blotting revealed MGLL protein level in response to MGLL-siRNA treatment; relative protein expression levels were normalized based on -actin levels. (left: photos indicative of three independent experiments; right: quantitative analysis, n = 3, paired Student's t-test) **(D)** Using the CCK-8 test, growth curves for MuM-2B cells treated with MGLL knockdown were determined. (Student's t-test) **(E)** To determine the migration of MuM-2B cells, a Transwell test was carried out. (left: typical images from three independent experiments; right: quantitative analysis, n = 3, unpaired Student's t-test, scale bar represents 100 μm) **(F)** A Transwell experiment was conducted to identify the invasion of MuM-2B cells following MGLL knockdown treatment. (left: typical images from three independent experiments; right: quantitative analysis, n = 3, unpaired Student's t-test, scale bar represents 100 μm) **(G)** MuM-2B cells were used in a wound healing experiment to identify MGLL knockdown-induced migration. (typical images of three independent experiments, scale bar represents 200 μm) *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001.

Tumor tissues contain many non-tumor cells, including immune cells, stromal cells, and interstitial cells, which help tumor formation and growth (32). Tumor purity correlated with clinical characteristics, genome expression, and characteristics (43). Thus, MGLL expression and tumor purity should be assessed in tumor samples (Figure 7E). 22 tumors had significant Pearson correlations, 1 of which was positive and 21 of which were negative (Table S6).

3.9 Immune infiltration analysis

Since this work indicated a relationship between MGLL expression and immune cell infiltration in various malignancies (Figure 7E), we explored immune cell infiltration in pan-cancers. QUANTISEQ analyzed the immune cell infiltration of a uniformly normalized dataset from the UCSC database. MGLL and immune cell infiltration scores were strongly associated with 10,180 tumor samples from 44 tumor types. Besides, MGLL expression and immune infiltration were correlated in 43 cancer types (Figure 8A). Interestingly, MGLL gene expression was highly positively related to M1-type macrophages in acute lymphoblastic leukemia (ALL, R = 0.58) and lymphoid neoplasm diffuse large B-cell lymphoma (DLBC, R =0.60; Figure 8A) and considerably positively associated with M2-type macrophages in UM (R = 0.60; Figure 8A).

Using TIMER2, we examined the connection between pan-cancer macrophage infiltration and MGLL expression and found a

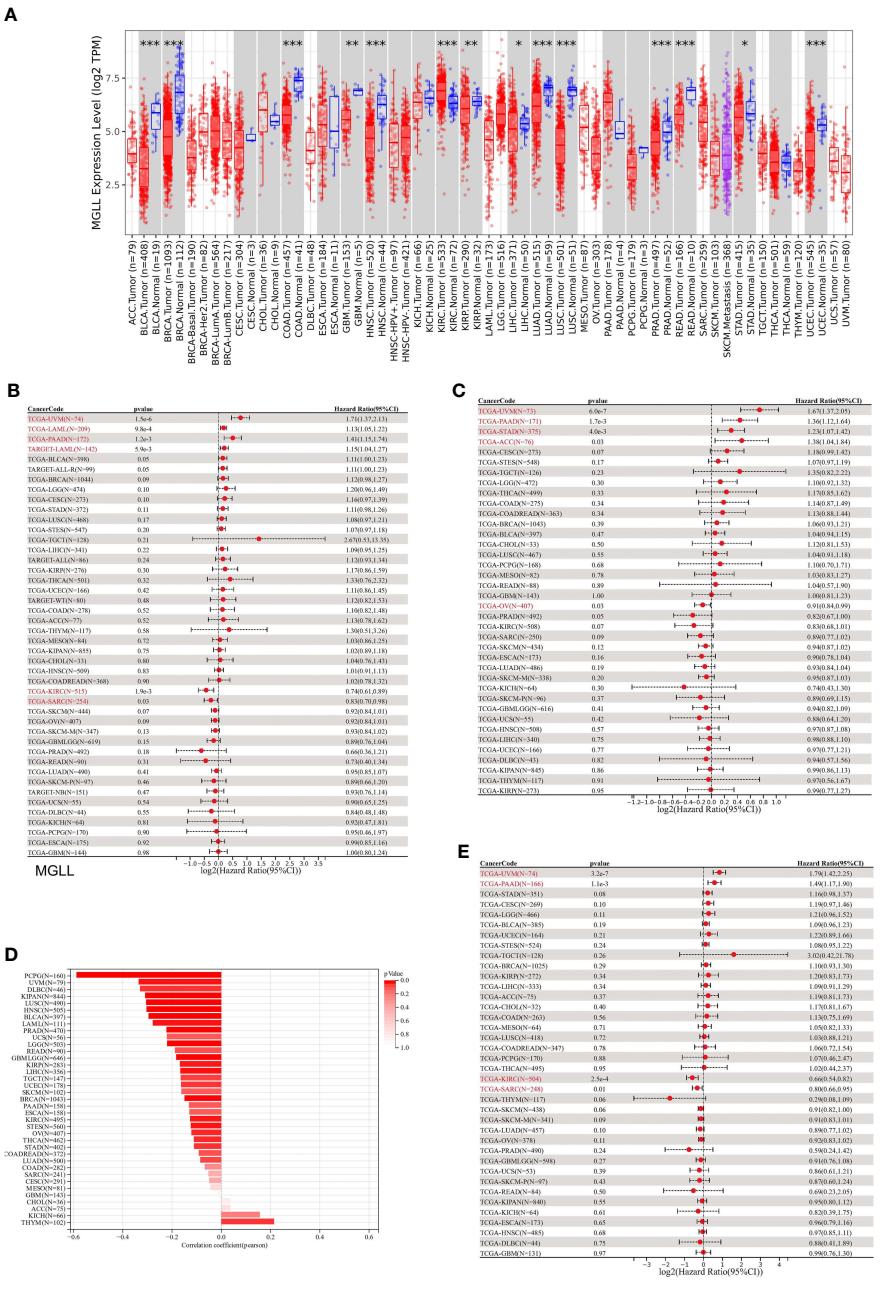


FIGURE 7
A correlation between MGLL and prognosis in pan-cancer samples. **(A)** Expression levels of MGLL across pan-cancer samples in the TCGA dataset. (http://timer.cistrome.org/) **(B)** Cox regression model study of MGLL expression and OS in the UCSC dataset. **(C)** Cox regression model-based analysis of PFS and MGLL expression. **(D)** Cox regression analysis of MGLL expression with DSS in different types of tumor. **(E)** MGLL expression and tumor purity were correlated using the ESTIMATE algorithm in the UCSC dataset. DSS, disease-specific survival; PFS, progression-free survival. *p < 0.05; **p < 0.01; ***p < 0.001.

positive correlation between them in DLBC, testicular germ cell tumor (TGCT), thymoma (THYM), and UM (Figure 8B). CIBERSORT algorithm also revealed a favorable association between UM MGLL expression and M2 macrophages (R = 0.59; Figure 8B). Therefore, MGLL affects immune cell infiltration especially macrophage polarization in various cancers.

3.10 Effect of MGLL on macrophage infiltration in TME

We used Uniform Manifold Approximation and Projection (UMAP) to cluster and designate 59,916 cells into 40 categories utilizing UM single-cell sequencing data (GSE139829) to study MGLL expression and function in TME at single-cell resolution (Figure 9A). Each sample's cell type fraction, with cancer cells, makes up practically the entire part (Figure 9B). Cancer cells and macrophages predominately expressed MGLL (Figure 9C). We then examined MGLL expression in cancer cells and macrophage makeup invading them (Figures 9D, S8A–F) and found that MGLL expression on cancer cells was positively correlated with SPP1-ACP5 macrophage infiltration (Figure 9D). SPP1-ACP5 macrophages were confirmed using the gene markers TLR2 of macrophage M1 and CD36 of M2. SPP1-ACP5 and IL1B macrophages express high levels of CD36 (Figure S8G), while ARG and IL1B macrophages express high levels of TLR2 (Figure S8H). Interestingly, MGLL may polarize M2 macrophages (SPP1-ACP5) by eliminating IL1B macrophages (Figure 9D).

To test if tumor cell-expressed MGLL can polarize macrophages to the M2 type, ARPE-19 or MuM-2B cells were co-cultured with THP-1 cells. Co-culturing MuM-2B and THP-1 cells elevated M2 macrophage markers including CD36 (Figure 9F), but not M1 markers like TLR2 (Figure 9E). In contrast, ARPE-19 cells co-cultured with THP-1 cells showed opposing macrophage polarization. These data suggest that UM-produced MGLL regulates macrophage polarization.

4 Discussion

Numerous lines of evidence imply that lipid metabolism is reprogrammed in cancers (44), which contributes to tumor

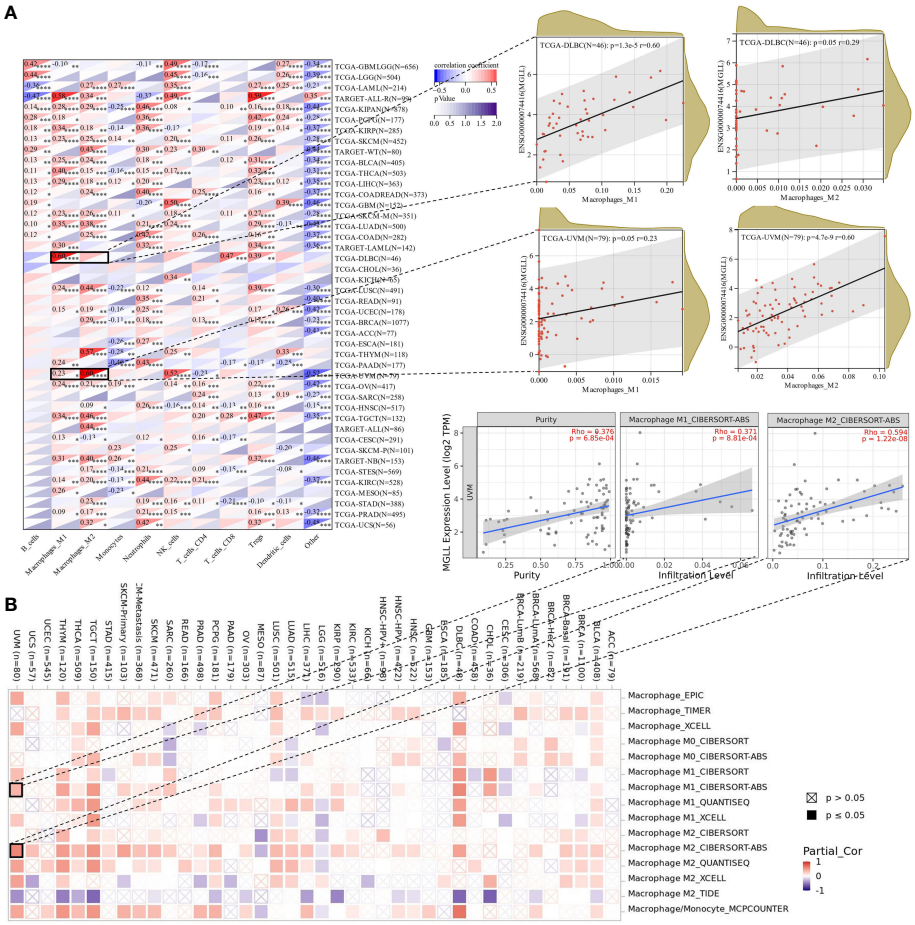


FIGURE 8
Assessment of immune infiltration. **(A)** Using the QUANTISEQ method, MGLL expression and immune cell infiltration were associated in the UCSC dataset. **(B)** Various algorithms identifies a relationship between MGLL expression and macrophages in different types of cancer from the TCGA database. *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001.

progression and local immunosuppression in the TME (26). Lipids, cell membrane components and second messengers that transduce signals within cells serve as vital energy storage sources under nutrient scarcity (45). The function of lipid metabolic abnormalities in cancer cells has been a major topic of study in recent years. The key processes of lipid metabolism are synthesis, storage, and breakdown. There is evidence that aberrant lipid metabolism plays a significant role in the development, progression, invasion, and treatment response of numerous cancers (46).

UM, the most common primary malignant eye tumor has been a major public health issue. Although UM will be diagnosed earlier as diagnostic technology advances, a fraction of early-stage patients are still diagnosed at an advanced stage, and the 5-year survival rate is still dismal, with the median survival of metastatic UM patients being less than 1 year due to high metastasis rates and restricted therapy options (3, 47, 48). Thus, better risk stratification strategies are needed to identify high-risk cancer patients to improve their prognosis.

Using consensus clustering, we classified samples into two categories based on the mRNA expression patterns of 47 prognostic genes derived from univariable Cox analysis. We found that lipid metabolism abnormalities may affect patient outcomes, as the two molecular categories had significant differences in overall survival, which we speculated may be linked to immune activity. Therefore, the ESTIMATE algorithm was then used to give additional insight into the immunological landscapes of UM, revealing that UM patients with bad prognoses had higher immune scores and ESTIMATE scores than those with better prognoses. Based on the aforementioned data, it can be assumed that immune variants may have a significant role in UM survival.

Furthermore, functional investigations were conducted to investigate the underlying mechanisms. In this study, we demonstrated that the prognosis of UM is significantly influenced by LMAGs through immune-associated signaling pathways, per the GO analysis and KEGG analysis. Subsequently, GSEA was used to explore the association between lipid metabolism and aberrant immunity. The findings showed that cluster 2 had lower immune cell differentiation expression. These findings provided a preliminary explanation for the prognostic differences between the two groupings, showing that immunological activity and the associated-LMRGs were responsible.

Additionally, we constructed a predictive risk model based on LMRGs and verified it in a validation cohort to establish that lipid metabolic disorders affect TME in UM patients. We also constructed a prognostic risk model using the four LMRGs signatures, including ENPP2, MGLL, PLCD1, and SLC44A3, and we found that most of them were correlated with tumor progression. For example, ENPP2, which encodes autotaxin, is overexpressed in chronic inflammatory diseases and cancer and synthesizes lysophosphatidic acid (49, 50). PLCD1 is known to convert phosphatidylinositol bisphosphate into diacylglycerol and inositol triphosphate, which serve as scaffolds and signaling molecules (51). PLCD1 is also identified as a new tumor suppressor gene, which is suppressed by promoter methylation in various cancer types (51–53). However, SLC44A3's role in cancer is unknown. Notably, only MGLL was highly expressed in the risk

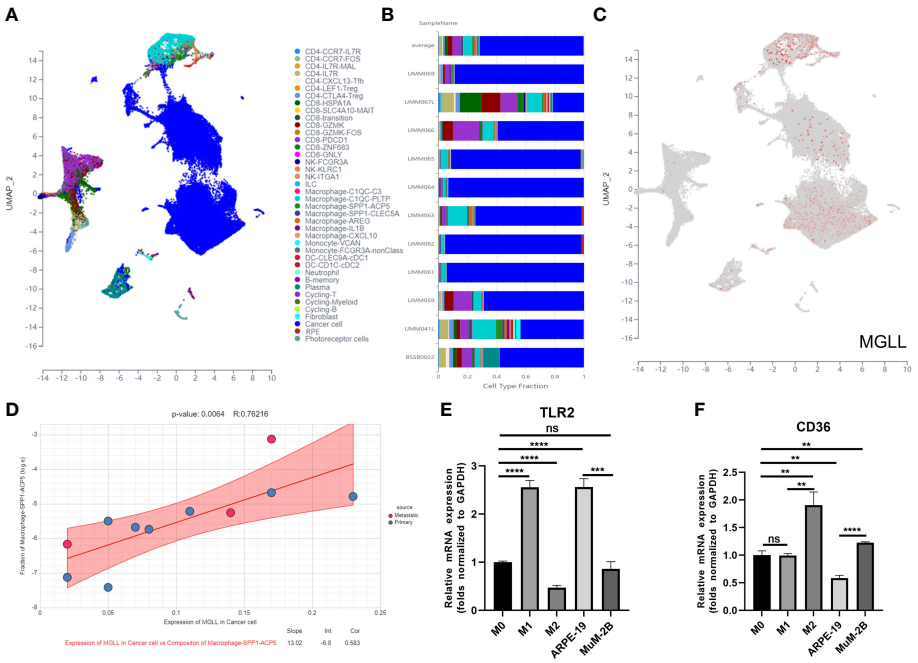


FIGURE 9
MGLL affects macrophage polarization in the TME. **(A)** Distribution of different cell clusters in the UMAP plot. **(B)** The fraction of each sample depends on the type of cell. **(C)** MGLL expression in distinct clusters of cells. **(D)** Correlation analysis between expression of MGLL in cancer cells and composition of infiltrating macrophages (SPP1-ACP5). **(E)** THP-1 macrophages were polarized towards an M1-like phenotype by the conditioned media of ARPE-19 cells with reduced MGLL expression. **(F)** THP-1 macrophages were polarized to an M2-like phenotype by the conditioned medium of MuM-2B cells with highly expressed MGLL. (n = 3, paired Student's t-test) **p < 0.01; ***p < 0.001; ****p < 0.0001 ns, no significance.

model. MGLL is a metabolic enzyme that transforms triglycerides into free fatty acids and is involved in tumor signaling (54). MGLL has been implicated to play a pathophysiological role in various cancers (54, 55). MALL is found to be involved in multiple cellular processes in cancer cells and is highly elevated in multiple aggressive cancer types (56). For instance, in endometrial adenocarcinoma, MGLL promoted tumor proliferation, metastasis, and the occurrence of progestogen resistance (57). In lung cancer, MGLL inhibition led to a decrease in cell proliferation, invasion, and metastasis (58, 59). In addition to inhibiting cell proliferation, migration, invasion, and tumor growth, MAGL inhibition also induced apoptosis in cervical cancer (60). According to a study on TNBC suggested that inhibiting MGLL can suppress inflammation, tumor growth, and brain colonization (61). Interestingly, MGLL in cancer cells promoted tumor progression by releasing special fatty acids whereas MGLL in TAMs suppressed cancer development by attenuating endogenous cannabinoid receptor 2 signaling (62). Additionally, MGLL has been reported to play a role in melanoma. Baba et al. observed that melanoma samples with lymphovascular invasion tended to be expressed more MGLL than samples without invasion, suggesting that the expression of MGLL in tumor cells may serve as a marker of tumor invasion and progression in malignant melanoma (63). In this study, we found that MGLL is upregulated in UM, and inhibition of MGLL suppressed the cell proliferation, migration, and invasion of UM cells, suggesting an oncogenic role of MGLL in UM.

Lipid metabolism has a substantial impact on macrophage regulatory functions. For example, lipids not only supply energy but also provide precursors to bioactive lipids and cell membrane components to macrophage (64). Besides, lipids regulate gene expression and signal transduction during macrophage activation (65). Several studies revealed that MGLL can induce the accumulation of 2-arachidonoylglycerol in the TME, which promotes the shift of tumor-associated macrophages into a tumor-promoting M2-like state by activating CB-2 (56, 62). Similarly, we discovered that MGLL was abundantly expressed in UM cells, which prompted a phenotypic shift in macrophages to a pro-tumor M2 state. Therefore, inhibiting MGLL in tumor cells may be a curative treatment for UM.

The 4-LMRGs risk model assigned a risk score to each UM patient, and the examination of survival in both the training and validation cohorts demonstrated powerful prediction ability. Using univariate/multivariate Cox regression analysis, the built risk model was an independent predictor of prognosis for UM patients irrespective of age, sex, or metastatic status. Additionally, we created and validated a nomogram that combines risk ratings and clinical characteristics to predict survival. The results indicated that aberrant lipid metabolism and TME may have an effect on therapy and survival, particularly in patients with metastatic UM.

Although our established LMRGs-based risk score for the prognosis of UM patients showed potential, there were disadvantages to being notified in our study. An initial disadvantage of the study is that differences in demographic variables, such as race, lifestyle, and living conditions were not taken into account. Second, our result was derived from open databases and not our cohorts' data. Thirdly, *in vivo* experiments are needed to further validate our results.

In conclusion, we discovered two genetic subgroups based on LMRGs in UM and assessed the significance of LMAGs in patients' prognosis and immune microenvironment. In addition, the molecular mechanisms may include the deregulation of lipid metabolism, which impedes the immune system and contributes to a bad prognosis. Additionally, we found that the actions of LMAGs in UM may be mediated by immune-related signaling pathways. We also discovered that cancer cells had elevated levels of MGLL expression, which switched macrophages to the pro-tumor M2 phenotype. Our work may shed light on the creation of new targeted medications and gives a potential direction for future UM research and personalized therapy.

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

YT and BQ conceived the study. YT and ZD collected and cleaned the data. YT and JP wrote the original draft. YT, JP, and ZL participated in completing the experiment. TC and JX participated in the formal analysis and data curation. ZD, CZ, and BQ revised the article. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Science Research Foundation of Shenzhen Aier Eye Hospital (SZAE2020B04), the Science Research Grant of Aier Eye Hospital Group (AM2001D2, AF2001D9, AM2101D1), the Natural Science Foundation of Guangdong Province, China (2022A1515010742), and Hunan Provincial Natural Science Foundation of China (2021JJ30045).

Acknowledgments

The authors are appreciative to the TCGA working groups for contributing their data and to T Huibertus van Essen and colleagues for uploading their data to GEO. Additionally, the authors would like to thank the patients and their families for donating tissue samples to TCGA and GEO, as well as the R development team and the R community for creating statistical packages.

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.

The reviewer HC declared a shared parent affiliation with the author CZ to the handling editor at the time of review.

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.2023.1161960/full#supplementary-material

References

1. Patel DR, Patel BC. Ocular melanoma. (2019). doi: 10.12968/opti.2019.11.221606

2. Wang Y, Mo L, Wei W, Shi X. Efficacy and safety of dendrimer nanoparticles with coexpression of tumor necrosis factor-A and herpes simplex virus thymidine kinase in gene radiotherapy of the human uveal melanoma ocm-1 cell line. *Int J Nanomedicine* (2013) 8:3805–16. doi: 10.2147/IJN.S48950

3. Wei C, Zhao X, Wang L, Zhang H. Trip suppresses cell proliferation and invasion in choroidal melanoma *Via* promoting the proteasomal degradation of Twist1. *FEBS Lett* (2020) 594(19):3170–81. doi: 10.1002/1873-3468.13882

4. Jager MJ, Shields CL, Cebulla CM, Abdel-Rahman MH, Grossniklaus HE, Stern M-H, et al. Uveal melanoma. *Nat Rev Dis Primers* (2020) 6(1):1–25. doi: 10.1038/s41572-020-0158-0

5. Hu DN, Yu GP, McCormick SA, Schneider S, Finger PT. Population-based incidence of uveal melanoma in various races and ethnic groups. *Am J Ophthalmol* (2005) 140(4):612–7. doi: 10.1016/j.ajo.2005.05.034

6. Naseripoor M, Azimi F, Mirshahi R, Khakpoor G, Poorhosseingholi A, Chaibakhsh S. Global incidence and trend of uveal melanoma from 1943-2015: A meta-analysis. *Asian Pac J Cancer Prev* (2022) 23(5):1791–801. doi: 10.31557/apjcp.2022.23.5.1791

7. Kaliki S, Shields CL. Uveal melanoma: Relatively rare but deadly cancer. *Eye (Lond)* (2017) 31(2):241–57. doi: 10.1038/eye.2016.275

8. Shields CL, Manalac J, Das C, Ferguson K, Shields JA. Choroidal melanoma: Clinical features, classification, and top 10 pseudomelanomas. *Curr Opin Ophthalmol* (2014) 25(3):177–85. doi: 10.1097/ICU.0000000000000041

9. Chen Y, Lu X, Gao L, Dean DC, Liu Y. Spheroid-induced heterogeneity and plasticity of uveal melanoma cells. *Cell Oncol (Dordr)* (2022) 45(2):309–21. doi: 10.1007/s13402-022-00671-y

10. Onken MD, Worley LA, Ehlers JP, Harbour JW. Gene expression profiling in uveal melanoma reveals two molecular classes and predicts metastatic death. *Cancer Res* (2004) 64(20):7205–9. doi: 10.1158/0008-5472.Can-04-1750

11. Hawkins BS. The coms randomized trial of iodine 125 brachytherapy for choroidal melanoma: v. twelve-year mortality rates and prognostic factors: coms report no. 28. *Arch Ophthalmol* (2006) 124(12):1684–93.

12. Cook SA, Damato B, Marshall E, Salmon P. Psychological aspects of cytogenetic testing of uveal melanoma: Preliminary findings and directions for future research. *Eye (Lond)* (2009) 23(3):581–5. doi: 10.1038/eye.2008.54

13. Rantala ES, Hernberg MM, Piperno-Neumann S, Grossniklaus HE, Kivelä TT. Metastatic uveal melanoma: The final frontier. *Prog Retin Eye Res* (2022) 90:101041. doi: 10.1016/j.preteyeres.2022.101041

14. Smit KN, Jager MJ, de Klein A, Kiliç E. Uveal melanoma: Towards a molecular understanding. *Prog Retin Eye Res* (2020) 75:100800. doi: 10.1016/j.preteyeres.2019.100800

15. Damato B. Progress in the management of patients with uveal melanoma. *2012 Ashton Lecture. Eye (Lond)* (2012) 26(9):1157–72. doi: 10.1038/eye.2012.126

16. Yan C, Hu X, Liu X, Zhao J, Le Z, Feng J, et al. Upregulation of Slc12a3 and Slc12a9 mediated by the Hcp5/Mir-140-5p axis confers aggressiveness and unfavorable prognosis in uveal melanoma. *Lab Invest* (2023), 100022. doi: 10.1016/j.labinv.2022.100022

17. Ren Y, Yan C, Wu L, Zhao J, Chen M, Zhou M, et al. Iumrg: Multi-layered network-guided propagation modeling for the inference of susceptibility genes and potential drugs against uveal melanoma. *NPJ Syst Biol Appl* (2022) 8(1):18. doi: 10.1038/s41540-022-00227-8

18. Yi M, Li J, Chen S, Cai J, Ban Y, Peng Q, et al. Emerging role of lipid metabolism alterations in cancer stem cells. *J Exp Clin Cancer Res* (2018) 37(1):118. doi: 10.1186/s13046-018-0784-5

19. Zhao J, Yi Q, Li K, Chen L, Dai L, Feng J, et al. A multi-omics deep learning model for hypoxia phenotype to predict tumor aggressiveness and prognosis in uveal melanoma for rationalized hypoxia-targeted therapy. *Comput Struct Biotechnol J* (2022) 20:3182–94. doi: 10.1016/j.csbj.2022.06.034

20. Hu B, Lin J-Z, Yang X-B, Sang X-T. Aberrant lipid metabolism in hepatocellular carcinoma cells as well as immune microenvironment: A review. *Cell Prolif* (2020) 53(3):e12772. doi: 10.1111/cpr.12772

21. Guo D, Prins RM, Dang J, Kuga D, Iwanami A, Soto H, et al. Egfr signaling through an akt-Srebp-1-Dependent, rapamycin-resistant pathway sensitizes glioblastomas to antilipogenic therapy. *Sci Signal* (2009) 2(101):ra82. doi: 10.1126/scisignal.2000446

22. Ettinger SL, Sobel R, Whitmore TG, Akbari M, Bradley DR, Gleave ME, et al. Dysregulation of sterol response element-binding proteins and downstream effectors in prostate cancer during progression to androgen independence. *Cancer Res* (2004) 64(6):2212–21. doi: 10.1158/0008-5472.CAN-2148-2

23. Yang Y-A, Morin PJ, Han WF, Chen T, Bornman DM, Gabrielson EW, et al. Regulation of fatty acid synthase expression in breast cancer by sterol regulatory element binding protein-1c. *Exp Cell Res* (2003) 282(2):132–7. doi: 10.1016/S0014-4827(02)00023-X

24. Yin F, Sharen G, Yuan F, Peng Y, Chen R, Zhou X, et al. Tip30 regulates lipid metabolism in hepatocellular carcinoma by regulating Srebp1 through the Akt/Mtor signaling pathway. *Oncogenesis* (2017) 6(6):e347. doi: 10.1038/oncsis.2017.49

25. Sun Y, He W, Luo M, Zhou Y, Chang G, Ren W, et al. Srebp1 regulates tumorigenesis and prognosis of pancreatic cancer through targeting lipid metabolism. *Tumour Biol* (2015) 36(6):4133–41. doi: 10.1007/s13277-015-3047-5

26. Hao Y, Li D, Xu Y, Ouyang J, Wang Y, Zhang Y, et al. Investigation of lipid metabolism dysregulation and the effects on immune microenvironments in pan-cancer using multiple omics data. *BMC Bioinf* (2019) 20(Suppl 7):195. doi: 10.1186/s12859-019-2734-4

27. Pellerin L, Carrié L, Dufau C, Nieto L, Ségui B, Levade T, et al. Lipid metabolic reprogramming: Role in melanoma progression and therapeutic perspectives. *Cancers (Basel)* (2020) 12(11): 3147. doi: 10.3390/cancers12113147

28. Robertson AG, Shih J, Yau C, Gibb EA, Oba J, Mungall KL, et al. Integrative analysis identifies four molecular and clinical subsets in uveal melanoma. *Cancer Cell* (2017) 32(2):204–20 e15. doi: 10.1016/j.ccell.2017.07.003

29. van Essen TH, van Pelt SI, Bronkhorst IH, Versluis M, Némati F, Laurent C, et al. Upregulation of hla expression in primary uveal melanoma by infiltrating leukocytes. *PloS One* (2016) 11(10):e0164292. doi: 10.1371/journal.pone.0164292

30. Wilkerson MD, Hayes DN. Consensusclusterplus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* (2010) 26(12):1572–3. doi: 10.1093/bioinformatics/btq170

31. Shen W, Song Z, Zhong X, Huang M, Shen D, Gao P, et al. Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta* (2022) 1(3):e36. doi: 10.1002/imt2.36

32. Yoshihara K, Shahmoradgoli M, Martínez 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

33. Becht E, de Reyniès A, Giraldo NA, Pilati C, Buttard B, Lacroix L, et al. Immune and stromal classification of colorectal cancer is associated with molecular subtypes and relevant for precision immunotherapydistinct immune phenotypes of colorectal cancer molecular subtypes. *Clin Cancer Res* (2016) 22(16):4057–66. doi: 10.1158/1078-0432.CCR-15-2879

34. Hänzelmann S, Castelo R, Guinney J. Gsva: Gene set variation analysis for microarray and rna-seq data. *BMC Bioinf* (2013) 14(1):1–15.

35. 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(5):453–7. doi: 10.1038/nmeth.3337

36. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* (2019) 10(1):1–10. doi: 10.1038/s41467-019-09234-6

37. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, et al. Timer2. 0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res* (2020) 48(W1):W509–W14.

38. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of rna-seq data. *Genome Med* (2019) 11(1):1–20.

39. Hong F, Meng Q, Zhang W, Zheng R, Li X, Cheng T, et al. Single-cell analysis of the pan-cancer immune microenvironment and sctime portalpan-cancer profile of single-cell times and sctime database. *Cancer Immunol Res* (2021) 9(8):939–51. doi: 10.1158/2326-6066.CIR-20-1026

40. Folberg R, Kadkol SS, Frenkel S, Valyi-Nagy K, Jager MJ, Pe'er J, et al. Authenticating cell lines in ophthalmic research laboratories. *Invest Ophthalmol Visual Sci* (2008) 49(11):4697–701. doi: 10.1167/iovs.08-2324

41. Liu R, Wang X, Curtiss C, Landas S, Rong R, Sheikh MS, et al. Monoglyceride lipase gene knockout in mice leads to increased incidence of lung adenocarcinoma. *Cell Death Dis* (2018) 9(2):36. doi: 10.1038/s41419-017-0188-z

42. Lv X, Ding M, Liu Y. Landscape of infiltrated immune cell characterization in uveal melanoma to improve immune checkpoint blockade therapy. *Front Immunol* (2022) 13:848455. doi: 10.3389/fimmu.2022.848455

43. Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. *Nat Commun* (2015) 6(1):1–12. doi: 10.1038/ncomms9971

44. Cheng C, Geng F, Cheng X, Guo D. Lipid metabolism reprogramming and its potential targets in cancer. *Cancer Commun (Lond)* (2018) 38(1):27. doi: 10.1186/s40880-018-0301-4

45. Currie E, Schulze A, Zechner R, Walther TC, Farese RV. Cellular fatty acid metabolism and cancer. *Cell Metab* (2013) 18(2):153–61. doi: 10.1016/j.cmet.2013.05.017

46. Pavlova NN, Thompson CB. The emerging hallmarks of cancer metabolism. *Cell Metab* (2016) 23(1):27–47. doi: 10.1016/j.cmet.2015.12.006

47. Zhao X, Kong F, Wang L, Zhang H. C-flip and the Noxa/Mcl-1 axis participate in the synergistic effect of pemetrexed plus cisplatin in human choroidal melanoma cells. *PloS One* (2017) 12(9):e0184135. doi: 10.1371/journal.pone.0184135

48. Hou P, Bao S, Fan D, Yan C, Su J, Qu J, et al. Machine learning-based integrative analysis of methylome and transcriptome identifies novel prognostic DNA methylation signature in uveal melanoma. *Briefings Bioinf* (2021) 22(4):bbaa371. doi: 10.1093/bib/bbaa371

49. Argaud D, Boulanger MC, Chignon A, Mkannez G, Mathieu P. Enhancer-mediated enrichment of interacting Jmjd3-Ddx21 to Enpp2 locus prevents r-loop formation and promotes transcription. *Nucleic Acids Res* (2019) 47(16):8424–38. doi: 10.1093/nar/gkz560

50. Panagopoulou M, Fanidis D, Aidinis V, Chatzaki E. Enpp2 methylation in health and cancer. *Int J Mol Sci* (2021) 22(21): 11958. doi: 10.3390/ijms222111958

51. Sanchez-Solana B, Wang D, Qian X, Velayoudame P, Simanshu DK, Acharya JK, et al. The tumor suppressor activity of Dlc1 requires the interaction of its start domain with phosphatidylserine, Plcd1, and caveolin-1. *Mol Cancer* (2021) 20(1):141. doi: 10.1186/s12943-021-01439-y

52. Hu X-T, Zhang F-B, Fan Y-C, Shu X-S, Wong A, Zhou W, et al. Phospholipase c delta 1 is a novel 3p22. 3 tumor suppressor involved in cytoskeleton organization, with its epigenetic silencing correlated with high-stage gastric cancer. *Oncogene* (2009) 28 (26):2466–75.

53. Xiang T, Li L, Fan Y, Jiang Y, Ying Y, Putti TC, et al. Plcd1 is a functional tumor suppressor inducing G2/M arrest and frequently methylated in breast cancer. *Cancer Biol Ther* (2010) 10(5):520–7. doi: 10.4161/cbt.10.5.12726

54. Zhang J, Song Y, Shi Q, Fu L. Research progress on fasn and mgll in the regulation of abnormal lipid metabolism and the relationship between tumor invasion and metastasis. *Front Med* (2021) 15(5):649–56. doi: 10.1007/s11684-021-0830-0

55. Lum CT, Liu X, Sun RW-Y, Li X-P, Peng Y, He M-L, et al. Gold (Iii) porphyrin 1a inhibited nasopharyngeal carcinoma metastasis in vivo and inhibited cell migration and invasion in vitro. *Cancer Lett* (2010) 294(2):159–66. doi: 10.1016/j.canlet.2010.01.033

56. Nomura DK, Long JZ, Niessen S, Hoover HS, Ng SW, Cravatt BF. Monoacylglycerol lipase regulates a fatty acid network that promotes cancer pathogenesis. *Cell* (2010) 140(1):49–61. doi: 10.1016/j.cell.2009.11.027

57. Ma X, Xia M, Wei L, Guo K, Sun R, Liu Y, et al. Abx-1431 inhibits the development of endometrial adenocarcinoma and reverses progesterone resistance by targeting mgll. *Cell Death Dis* (2022) 13(12):1067. doi: 10.1038/s41419-022-05507-z

58. Zhang H, Guo W, Zhang F, Li R, Zhou Y, Shao F, et al. Monoacylglycerol lipase knockdown inhibits cell proliferation and metastasis in lung adenocarcinoma. *Front Oncol* (2020) 10:559568. doi: 10.3389/fonc.2020.559568

59. Prüser JL, Ramer R, Wittig F, Ivanov I, Merkord J, Hinz B. The monoacylglycerol lipase inhibitor Jzl184 inhibits lung cancer cell invasion and metastasis *Via* the Cb(1) cannabinoid receptor. *Mol Cancer Ther* (2021) 20(5):787–802. doi: 10.1158/1535-7163.Mct-20-0589

60. Wang C, Li Z, Zhong L, Chen Y. Inhibition of monoacylglycerol lipase restrains proliferation, migration, invasion, tumor growth and induces apoptosis in cervical cancer. *J Obstet Gynaecol Res* (2022) 48(2):456–66. doi: 10.1111/jog.15110

61. Benchama O, Tyukhtenko S, Malamas MS, Williams MK, Makriyannis A, Avraham HK. Inhibition of triple negative breast cancer-associated inflammation, tumor growth and brain colonization by targeting monoacylglycerol lipase. *Sci Rep* (2022) 12(1):5328. doi: 10.1038/s41598-022-09358-8

62. Xiang W, Shi R, Kang X, Zhang X, Chen P, Zhang L, et al. Monoacylglycerol lipase regulates cannabinoid receptor 2-dependent macrophage activation and cancer progression. *Nat Commun* (2018) 9(1):2574. doi: 10.1038/s41467-018-04999-8

63. Baba Y, Funakoshi T, Mori M, Emoto K, Masugi Y, Ekmekcioglu S, et al. Expression of monoacylglycerol lipase as a marker of tumour invasion and progression in malignant melanoma. *J Eur Acad Dermatol Venereol* (2017) 31(12):2038–45. doi: 10.1111/jdv.14455

64. Yan J, Horng T. Lipid metabolism in regulation of macrophage functions. *Trends Cell Biol* (2020) 30(12):979–89. doi: 10.1016/j.tcb.2020.09.006

65. Van den Bossche J, O'Neill LA, Menon D. Macrophage immunometabolism: Where are we (Going)? *Trends Immunol* (2017) 38(6):395–406. doi: 10.1016/j.it.2017.03.001

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 23 March 2023
DOI 10.3389/fimmu.2023.1086907

OPEN ACCESS

EDITED BY
Meng Zhou,
Wenzhou Medical University, China

REVIEWED BY
Xinpei Deng,
Sun Yat-sen University Cancer Center (SYSUCC), China
Huimei Wang,
Fudan University, China

*CORRESPONDENCE
Long Wang
wanglong@csu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Immunology

RECEIVED 01 November 2022
ACCEPTED 09 March 2023
PUBLISHED 23 March 2023

CITATION
Li C, Peng D, Gan Y, Zhou L, Hou W, Wang B, Yuan P, Xiong W and Wang L (2023) The m^{6}A methylation landscape, molecular characterization and clinical relevance in prostate adenocarcinoma. *Front. Immunol.* 14:1086907. doi: 10.3389/fimmu.2023.1086907

The m^{6}A methylation landscape, molecular characterization and clinical relevance in prostate adenocarcinoma

Chao Li[1†], Dongyi Peng[1†], Yu Gan[2], Lei Zhou[1], Weibin Hou[1], Bingzhi Wang[1], Peng Yuan[1], Wei Xiong[1] and Long Wang[1*]

[1]Department of Urology, Third Xiangya Hospital, Central South University, Changsha, China, [2]Department of Urology, Xiangya Hospital, Central South University, Changsha, China

Background: Despite the recent progress of therapeutic strategies in treating prostate cancer (PCa), the majority of patients still eventually relapse, experiencing dismal outcomes. Therefore, it is of utmost importance to identify novel viable targets to increase the effectiveness of treatment. The present study aimed to investigate the potential relationship between N6-methyladenosine (m6A) RNA modification and PCa development and determine its clinical relevance.

Methods: Through systematic analysis of the TCGA database and other datasets, we analyzed the gene expression correlation and mutation profiles of m6A-related genes between PCa and normal tissues. Patient samples were divided into high- and low-risk groups based on the results of Least Absolute Shrinkage and Selection Operator (LASSO) Cox analysis. Subsequently, differences in biological processes and genomic characteristics of the two risk groups were determined, followed by functional enrichment analysis and gene set enrichment (GSEA) analysis. Next, we constructed the protein-protein interaction (PPI) network of differentially expressed genes between patients in high- and low-risk groups, along with the mRNA-miRNA-lncRNA network. The correlation analysis of tumor-infiltrating immune cells was further conducted to reveal the differences in immune characteristics between the two groups.

Results: A variety of m6A-related genes were identified to be differentially expressed in PCa tissues as compared with normal tissues. In addition, the PPI network contained 278 interaction relationships and 34 m6A-related genes, and the mRNA-miRNA-lncRNA network contained 17 relationships, including 91 miRNAs. Finally, the immune characteristics analysis showed that compared with the low-risk group, the levels of M1 and M2 macrophages in the high-risk

group significantly increased, while the levels of mast cells resting and T cells CD4 memory resting significantly decreased.

Conclusions: This study provides novel findings that can further the understanding of the role of m6A methylation during the progression of PCa, which may facilitate the invention of targeted therapeutic drugs.

KEYWORDS

prostate adenocarcinoma, RNA N6-methyladenosine, prognosis, molecular characterization, immune infiltration

Introduction

According to the statistics of the American Cancer Society, prostate cancer (PCa) is the second leading cause of cancer-related death in men in the United States, with an estimated 288,300 new cases and 34,700 deaths per year, accounting for 28.5% and 10.8% of all cancers, respectively (1). With the substantial increase in the aging population in China, the incidence of PCa has also increased year by year, and PCa has become the most common urogenital tumor in elderly men (2). Despite recent advances in surgical and drug treatments, the mortality rates of patients with recurrent or metastatic PCa remain close to 100% (1). Therefore, in-depth study of molecular markers related to treatment and prognosis of PCa and searching for more effective therapeutic targets are of significant importance for the clinical benefit of PCa patients.

To date, more than 150 RNA post-transcriptional modifications have been identified in eukaryotes (3). N6-methyladenosine (m^6A) is the most common RNA modification in mammalian cells that has important roles in different biological processes (4, 5). Abnormalities in regulatory mechanisms of m^6A have been identified as involved in a variety of human diseases including cancer (6). m^6A, as the methylation at the sixth N position of adenylate in RNA, is the most common modification of RNA in eukaryotes, accounting for about 80% of RNA methylation modifications, and each mRNA contains 3 to 5 m^6A residues on average (3). This process is dynamically and reversibly regulated by methyl transfer-related proteins (METTL3, METTL14, and WTAP, etc.) and demethylases (FTO, ALKBH3, and ALKBH5, etc.), and affects various steps of mRNA metabolism reader, including mRNA processing, nuclear export, translation and degradation, by binding to the m^6A (7). Several studies have established the model for m^6A risk-related prognosis to evaluate the treatment effect and prognosis of metastatic PCa, finding that in patients with metastatic PCa, a higher m^6A risk score indicates a worse prognosis, which is significantly associated with biological functions such as DNA mismatch repair. Therefore, patients with high m^6A risk scores may be a more suitable population for DNA repair-targeted drug therapy (8, 9). In addition, several studies have reported the potential tumor-promoting or tumor-suppressing effects of m^6A methylation-related factors such as METTL3, METTL14 and FTO in PCa (10–14). However, there is still a lack of integrative analysis of the expression of m^6A RNA methylation regulator, clinicopathological features, malignant progression, and prognosis in PCa.

In this study, we used published sequencing data to investigate the possible role of m^6A methylation in the progression of PCa, and to establish relevant clinical prediction model to analyze the predictive power of prognosis in PCa.

Materials and methods

Data acquirement and processing

The gene expression data of gene sequencing of patients with prostate adenocarcinoma (PRAD) was downloaded from the TCGA GDC (https://portal.gdc.cancer.gov/). The clinical characteristics of the corresponding patients, including age, gender, and survival prognosis, were also downloaded. After deleting the PRAD patients with missing clinical information, 481 tumor tissues and 51 normal tissues were ultimately included in the analysis. The somatic mutation data of PRAD patients were downloaded and maftools package of R software was used to visualize the somatic mutation (15). The tumor mutation burden (TMB) of each patient was collected. Besides, datasets including GSE46602 and GSE69223 were downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) (16, 17). Moreover, GSE46602 contains 36 tumor tissues and 14 normal tissues, and GSE69223 contains 15 tumor tissues and 15 normal tissues. Both datasets came from the GPL570 sequencing platform, where the species origin was Homo sapiens.

Construction of a risk model for PCa

To analyze the expression of m^6A-related genes in PRAD, we first analyzed the differential expression and gene expression correlation of m^6A-related genes in PRAD and normal tissues. The risk genes associated with PCa prognosis were obtained through univariate cox regression analysis of the expression and survival of PRAD patients from TCGA. The risk genes associated with PCa prognosis were subsequently incorporated into the model, and the Least Absolute Shrinkage and Selection Operator (LASSO)

was used to reduce the data dimensionality and obtain prognostic-related signature genes. The normalized values of expression of each gene were weighted by the penalty coefficient by LASSO Cox analysis, a risk score formula was established, and the patients were divided into high-risk group and low-risk group according to median value of the risk score, as follows:

$$\text{riskScore} = \sum_i \text{Coefficient (risk gene}_i) * \text{mRNA Expression (risk gene}_i)$$

Differentially expressed genes analysis

To analyze the effect of risk score on DEGs analysis of PRAD, the R package "DESeq2" was used to perform DEGs analysis on samples in high-risk and low-risk groups of the dataset from TCGA-PRAD to screen for significant differential genes (18). The absolute value of log2 fold change (logFC) > 1.5 and Padj< 0.05 were set as the thresholds of differential genes. Genes with logFC > 1.5 and Padj< 0.05 were up-regulated DEGs, and genes with logFC< -1.5 and Padj< 0.05 were down-regulated DEGs (19).

Genomic characteristics and biological characteristics of patients in high-risk group and low-risk group

Following the development of tumor genomics, the Mutation Annotation Format (MAF) has become widely accepted and used to store detected somatic variants. In order to evaluate the variation of gene copy number variation in risk-grouping, the GISTIC2.0 in the Genepattern (https://cloud.genepattern.org/) analysis platform was used to analyze the copy number variation in the risk groups of TCGA database (20).

In this study, the MSIpred method was used to analyze the relationship between risk-grouping and TMB or microsatellite instability (MSI), respectively (21). In addition, in order to investigate the variation of biological process of samples in high-risk group compared with that in low-risk group, we performed gene set variation analysis using the R package "GSVA" based on the gene expression profiling dataset of PRAD patients from TCGA (22).

The reference gene set "h.all.v7.4.symbols.gmt" was downloaded from the MSigDB database to calculate the enrichment score of each sample in each pathway in the dataset (23), and evaluate the relationship between the enrichment score and the risk score. P< 0.05 was considered statistically significant.

Functional enrichment analysis and gene set enrichment analysis

GO analysis is a method commonly used for large-scale functional enrichment studies, including biological process (BP), molecular function (MF) and cellular component (CC) (24). KEGG is a widely used database for storing data about genomes, biological pathways, diseases, and drugs (25). GO annotation analysis and KEGG pathway enrichment analysis of differentially expressed genes were performed using the clusterProfiler package of R and a cutoff value of FDR< 0.05 was considered statistically significant (26).

To investigate differences in biological processes between two groups, based on the gene expression profiling dataset of PRAD patients, gene set enrichment analysis was performed using GSEA, which is a computational method to analyze the potential existence of significant differences in a specific gene set between two biological states (27). Also, GSEA is often used to estimate changes in pathway and biological process activity in samples of expression dataset. The "c2.cp.kegg.v7.4.symbols.gmt" gene set and the "c5.go.v7.2.symbols.gmt" gene set were downloaded from the MSigDB database for GSEA analysis. P< 0.05 was considered statistically significant.

Identification and correlation analysis of tumor infiltrating immune cells

CIBERSORT is an algorithm that deconvolves the expression matrix of immune cell subtypes based on the principle of linear support vector regression, which utilizes RNA-Seq data to estimate the abundance of immune cells in tissues (28). The CIBERSORT in R software was used to estimate the abundance of 22 kinds of immune cells in high-risk and low-risk groups in the dataset, and boxplots were performed to visualize the immune cell composition of disease samples and normal samples. The Wilcoxon test calculated differences in the proportion of immune cells between disease samples and normal samples, and P< 0.05 was considered statistically significant. The dataset on the interaction of PRAD cell lines with drugs was obtained from the GDSC database (29), and the R package oncoPredict was used for drug sensitivity analysis of the expression data of patients in the high-risk group and the low-risk group from TCGA-PRAD so as to compare the sensitivity differences in anti-tumor drugs between patients in high-risk group and low-risk group (30).

Construction of protein-protein interaction network and key gene-miRNA network

The PPI network includes interactions of individual protein with each other that participate in all aspects of life processes such as biological signal transmission, gene expression regulation, energy and material metabolism, and cell cycle regulation. Therefore, systematic analysis of the interaction of a large number of proteins in biological systems is useful for elucidating the working principle of proteins in biological systems, understanding the mechanism of biological signals and energy metabolism under special physiological conditions such as diseases, as well as the functional connections between proteins.

The STRING database is used for searching for interactions between known protein and predicted protein (31). In this study, we used the STRING database and selected genes with a combined score > 400 to construct a protein-protein interaction network related to DEGs. Besides, Cytoscape (v3.7.2) was used to visualize

the PPI network model. Genes in the PPI network were functionally annotated using clueGO (32, 33).

In order to analyze the relationship between key genes and miRNAs in the post-transcriptional stage, miRNAs related to differentially expressed genes from the miRNet database were obtained to construct an mRNA-miRNA regulatory network (34). The mRNA-miRNA regulatory network was visualized using Cytoscape software. lncRNA is a class of RNA molecules with transcripts longer than 200 nt, which are generally considered to not encode proteins, but participate in the regulation of protein-coding genes in the form of RNA in epigenetic regulation, transcriptional regulation and post-transcriptional regulation (35).

To analyze the relationship among DEGs and miRNAs and lncRNAs in the post-transcriptional stage, we obtained miRNAs and lncRNAs related to DEGs from the miRNet database to construct an mRNA-miRNA-lncRNA regulatory network (34), which was visualized by the Cytoscape software.

Construction of clinical prediction model based on risk model

To demonstrate the individualized assessment of prognosis of patients by risk scores combined with clinicopathologic characteristics, univariate and multivariate Cox analyses were subsequently performed to analyze the predictive power of risk scores combined with clinicopathologic characteristics of patients for overall survival (OS). Subsequently, the risk score model with clinicopathologic characteristics was selected to construct a clinical predictive nomogram. To quantify discriminative performance, a calibration curve was generated to assess the performance of the nomogram by comparing the predicted value of the nomogram with the observed actual survival.

Cell culture

Human prostate normal cell line RWPE-1, PCa cell line 22Rv1 and PC3 were purchased from American type culture collection (ATCC). All cells were cultured in RPMI-1640 cell culture medium containing 10% FBS in a 5% CO_2 humidified atmosphere at 37°C. When used in experiments, these cell lines were cultured within 20 passages, and regular routine testing was employed to confirm them as negative for mycoplasma.

Real-time-qPCR analysis

In order to detect the mRNA levels of each m^{6}A-related factor, total RNA was extracted from cells using the RNAsimple Total RNA Kit (TIANGEN), after which the obtained RNA was reverse transcribed into cDNA using the RevertAid First Strand cDNA Synthesis Kit (ThermoFisher). Each cDNA sample was amplified using SuperReal PreMix Plus SYBR Green Supermix (TIANGEN) in the LightCycler 480 Real-Time PCR System (Roche) following the manufacturer's instructions. Primers used for RT-qPCR analysis are shown in Supplementary Table 1. Relative RNA levels were calculated using the $2^{-\Delta\Delta Ct}$ method, and normalized to β-actin as an internal control.

Western blot

To denature proteins, cell lysates were added to 5× loading buffer (Beijing TDY Biotech) and heated to 95°C for 5 min. Protein samples were separated by SDS-PAGE electrophoresis, transferred semi-dry onto NC membranes (Millipore), and blocked in Tris-buffered saline-Tween 20 (TBST) containing 5% nonfat milk for 30 min, after which the immunoblotting was performed by incubating with the primary antibody for 10 min at room temperature, and then overnight at 4°C. After being subjected to 5 washes, the membranes were incubated with goat anti-mouse/rabbit IgG (H+L)-HRP secondary antibody (Beijing TDY Biotech, 1:10000 dilution) for 40 min and were subsequently exposed to light using western ECL Substrate (Millipore). The relative expression levels of each protein were assessed using ImageJ software. Primary antibodies used in this study are listed in Supplementary Table 2.

Statistical analysis

All data processing and analysis were performed by R software (version 4.1.1). The student's t-test was used to estimate the statistical significance of normally distributed variables for the comparison of measurement data between two groups. The Wilcoxon rank-sum test was used to calculate the statistical significance of non-normally distributed variables between two groups. The Chi-square test or Fisher's exact test was used to compare and analyze the statistical significance of categorical data between two groups. Correlation coefficients between different genes were calculated by Pearson correlation analysis. The Kaplan-Meier survival curve was used to show the difference in survival, and the log-rank test was used to evaluate the significant difference in survival between the two groups. All statistical P values were two-sided, and $P< 0.05$ was considered statistically significant.

Results

Expression and mutation of m^{6}A-related genes in PRAD patients

The baseline data of patients with PRAD are shown in Supplementary Table 3. To analyze the expression levels of m^{6}A-related genes in PRAD patients, we analyzed genomic mutations and mRNA expression, respectively. First, a comprehensive analysis of expression profiles in PCa tissues and normal tissues from TCGA data and GEO data was performed with de-batch effects (Figure 1). Principal Component Analysis (PCA) showed significant differences in m^{6}A-related gene signatures between PRAD tissues and normal tissues.

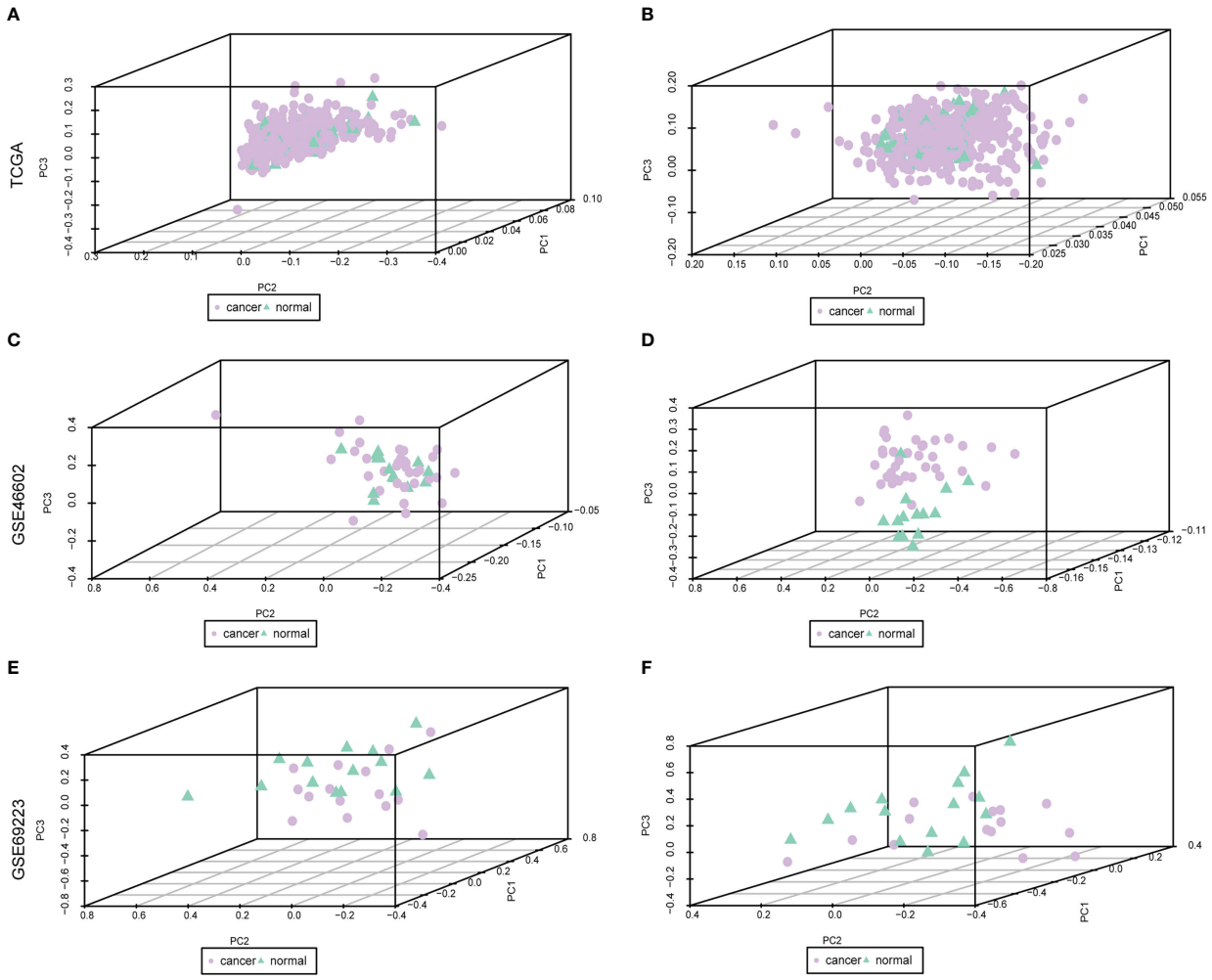


FIGURE 1
Dataset on PRAD after correction. Purple nodes indicate tumor samples, and green nodes indicate normal samples. **(A, C, E)** are the data before correction, and **(B, D, F)** are the data after correction.

Subsequently, the differential analysis showed that a variety of m^{6}A-related genes were significantly differentially expressed between PCa tissues and normal tissues, including FTO, METTL14, METTL16, ZC3H13, YTHDC1, YTHDF3, RBM15B, etc. (Figure 2).

Mutation analysis showed that most of the mutations were missense mutations, and most of the mutation types were SNPs (Figure 3A). There were 22 patients with PRAD and single nucleotide mutations in m^{6}A-related genes, among which the ZC3HI3 had the highest mutation rate (Figure 3B). The correlation analysis of the heat map showed a positive correlation of m^{6}A-related genes in PRAD tissues (Figure 3C).

The total number of mutations was obtained to calculate the TMB of the high-risk group of PRAD patients and low-risk PRAD patients. TMB was higher in PRAD patients in the high-risk group (Figure 3D), suggesting that PRAD patients in the high-risk group may be more likely to respond to immunotherapy. MSI is also an important treatment for predicting the effect of immunotherapy. Thus, we predicted the status distribution of MSI-H and MSI of PRAD patients in the high-risk group and low-risk group based on mutation data (Figure 3E). Our results showed that patients with MSI-H were all PRAD patients in the high-risk group and that MSI-H samples may be more sensitive to immunotherapy and more benefit from immunotherapeutic drugs.

Construction of risk model and prognostic analysis

In order to analyze the impact of genes on the prognosis of PRAD patient, 278 risk genes associated with PCa prognosis were identified by univariate cox regression analysis, and enrolled in LASSO-Cox analysis to select and obtain 18 genes with the best prognostic value (Figures 4A, B). Subsequently, the correlation among the expression levels of these genes was analyzed, which showed that the signature genes were broadly represented (Figure 4C). At the same time, based on penalty coefficients of important signature genes calculated by LASSO-Cox analysis, the gene expression was multiplied by the corresponding coefficients and added to establish a risk score. Besides, the final risk score of each sample was calculated. Next, patients were divided into high-risk group and low-risk group based

10.3389/fimmu.2023.1086907

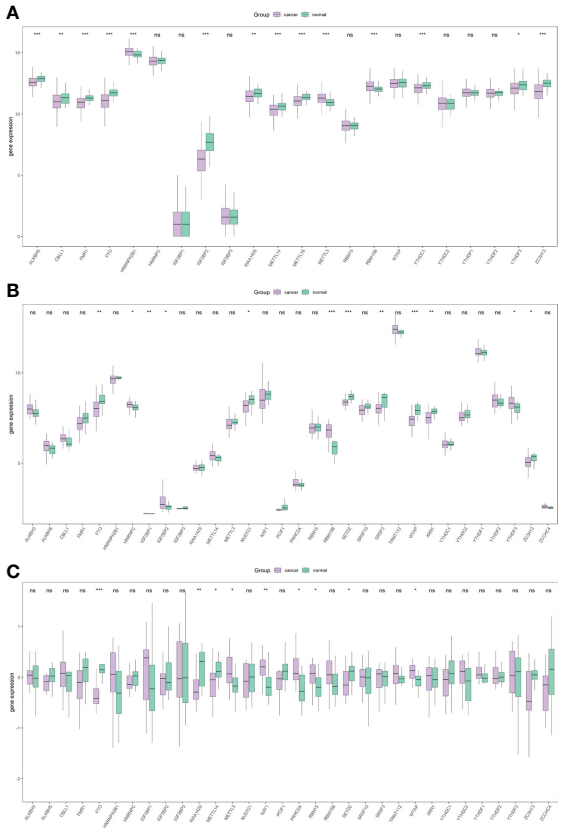


FIGURE 2
Overall expression of m^{6}A-related genes in PRAD patients. Purple indicates the tumor sample, and green indicates the normal sample. Three images indicate TCGA **(A)**, GSE46602 **(B)**, GSE69223 **(C)**. *P < 0.05, **P < 0.01, ***P < 0.001, ns, not significant.

on the mean value of PRAD patients' risk scores. Kaplan-Meier analysis showed that patients in high-risk group had relatively poor OS (Log-rank *P*< 0.0001, Figure 4D). Moreover, a significant correlation was found between the expression levels of m^{6}A-related genes and the risk score of patients (Figure 4E).

Next, we analyzed the differences in m^{6}A-related gene expression levels of patients between the high-risk group and low-risk group, finding 27 m^{6}A-related genes with significantly differential expression between patients in high-risk group and low-risk group (all *P*<0.05, Figure 5).

Differences in biological processes and genomic characteristics of risk-groups

The mutation types of mutated genes in PRAD patients in the high-risk group and low-risk group were analyzed, and more gene mutations were found in PRAD patients in the high-risk group (Figures 6A, B). Subsequently, we analyzed the high-frequency mutation genes of patients in the two groups, finding that the gene with the highest mutation frequency of patients in the high-risk group was TP53 (Figure 6C), while the gene with the highest mutation frequency among patients in the low-risk group was SPOP (Figure 6D). The relationship between mutated genes of patients in the two groups was compared, showing significant co-mutation between MACF1 and PCLO in PRAD patients in the high-risk group (Figure 6E), and significant co-mutation between SPOP and ASH1L in PRAD patients in the low-risk group (Figure 6F).

Finally, GISTIC 2.0 was used to identify genes with significant amplification or deletion in the copy number variation data of patients in two groups, respectively. The results showed more gene copy number amplifications on chromosomes 2, 12, 13, 20, and 21 in PRAD patients in the high-risk group (Figures 6G, H).

To identify the underlying biological features of the different risk models, we calculated the correlation between the enrichment score and the risk score at the hallmark for each sample, and the results showed that the risk score had a significant negative association with DNA repair, MYC targets V1, G2M checkpoint, unfolded protein response, MYC targets v2, E2F targets and oxidative phosphorylation, and significant positive association with an apical surface and myogenesis (all *P*<0.05, Figure 7).

Difference analysis between high-risk group and low-risk group

As the level of risk has a significant impact on the survival rate of patients, we conducted a differential analysis on the gene expression of patients in the high-risk group and the low-risk group, taking the genes with Padj< 0.01 and |logFC|> 1.5 as the differentially expressed

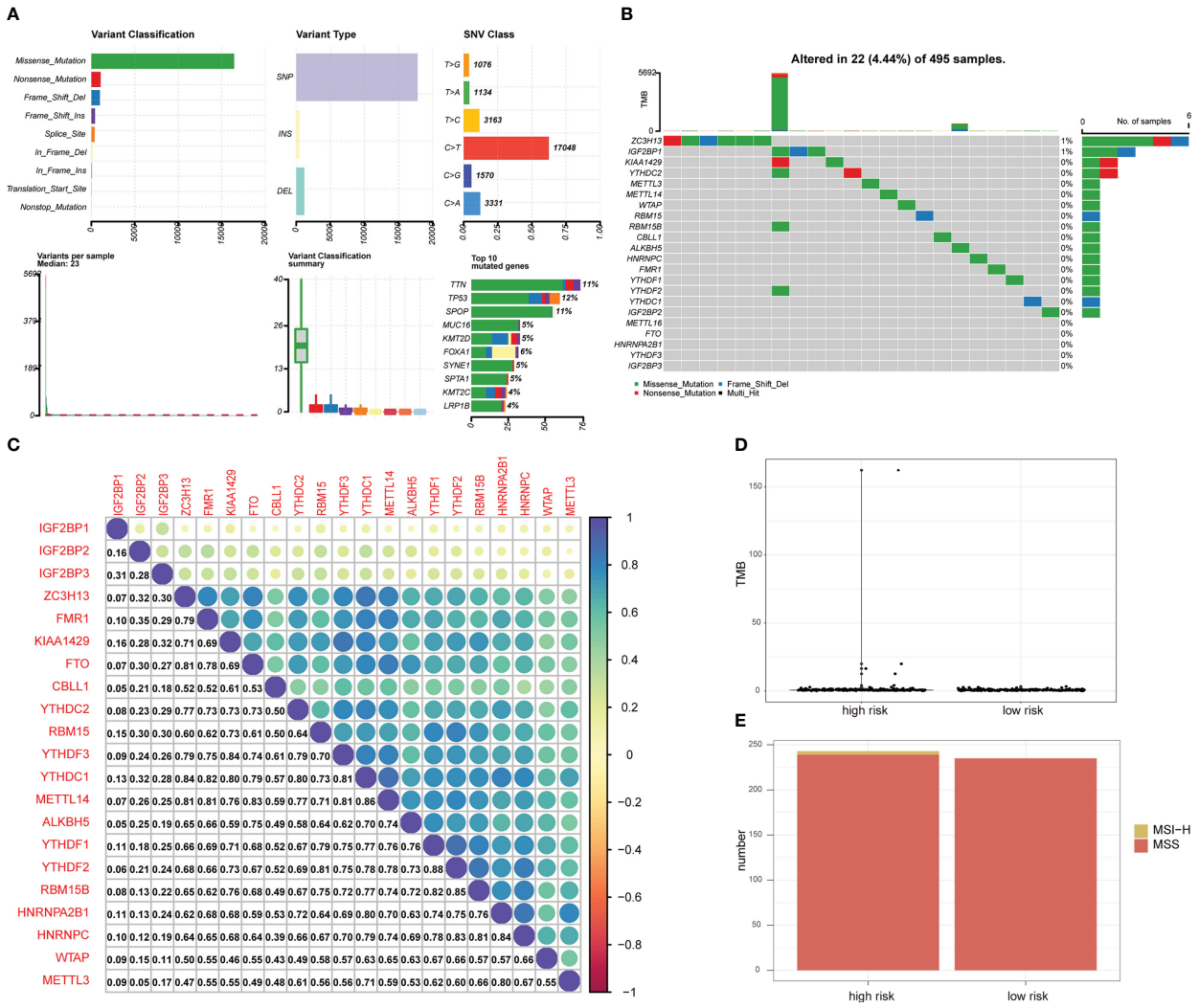


FIGURE 3
Mutation status of m^{6}A-related genes in PRAD patients. **(A)** Summary of PRAD patients mutation data from TCGA. **(B)** Mutation map of m^{6}A-related genes in PRAD patients from TCGA. Samples are ordered according to somatic nonsynonymous mutational burden and genes are ordered by mutation frequency, with various colors indicating different mutation types. Subsection above legend shows mutational burden. **(C)** The expression level correlation of m^{6}A-related genes in the gene expression profile of PRAD patients from TCGA. The numbers in the figure and the annotation bar on the right indicate the magnitude of the correlation. **(D)** Differences in TMB between PRAD patients in high-risk group and low-risk group. **(E)** Differences in MSI status between PRAD patients in high-risk group and low-risk group.

genes. We identified 284 differentially expressed genes, including 207 up-regulated genes and 77 down-regulated genes (Figure 8A). At the same time, the differentially expressed genes were divided into differentially expressed mRNAs and differentially expressed lncRNAs. There were 164 up-regulated miRNAs and 71 down-regulated miRNAs (Figure 8B) identified, and 43 up-regulated lncRNAs and 6 down-regulated lncRNAs (Figure 8C).

Subsequently, we analyzed the impact of differentially expressed mRNAs between the high-risk group and low-risk group on biologically relevant functions of patients. First, GO functional annotation was performed on the differentially expressed genes (Figure 9A; Supplementary Table 4), revealing that these differentially expressed genes were mainly enriched in biological processes including muscle filament sliding, actin-myosin filament sliding, striated muscle cell development, myofibril assembly, thyroid hormone metabolic process, cellular component assembly involved in morphogenesis and thyroid hormone generation (Figure 9B); in cellular components including sarcomere, myofibril, contractile fiber, muscle myosin complex, and myosin II complex (Figure 9C), and in molecular functions including lipase inhibitor activity, endopeptidase Inhibitor activity, peptidase inhibitor activity, microfilament motor activity, endopeptidase regulator activity, enzyme inhibitor activity (Figure 9D). At the same time, these differentially expressed genes were enriched in KEGG pathways such as Thyroid hormone synthesis, Chemical carcinogenesis-DNA adducts, Pancreatic secretion, Drug metabolism-cytochrome P450 (Figure 9E; Supplementary Table 5). The enrichment of the expression levels of differentially expressed genes in pathways hsa00982, hsa04918, and hsa04972 is shown in detail in Figures 9F–H.

Next, GSEA was performed on all gene expressions between the high-risk group and the low-risk group, showing significant differences in the following biological processes between groups (Supplementary

10.3389/fimmu.2023.1086907

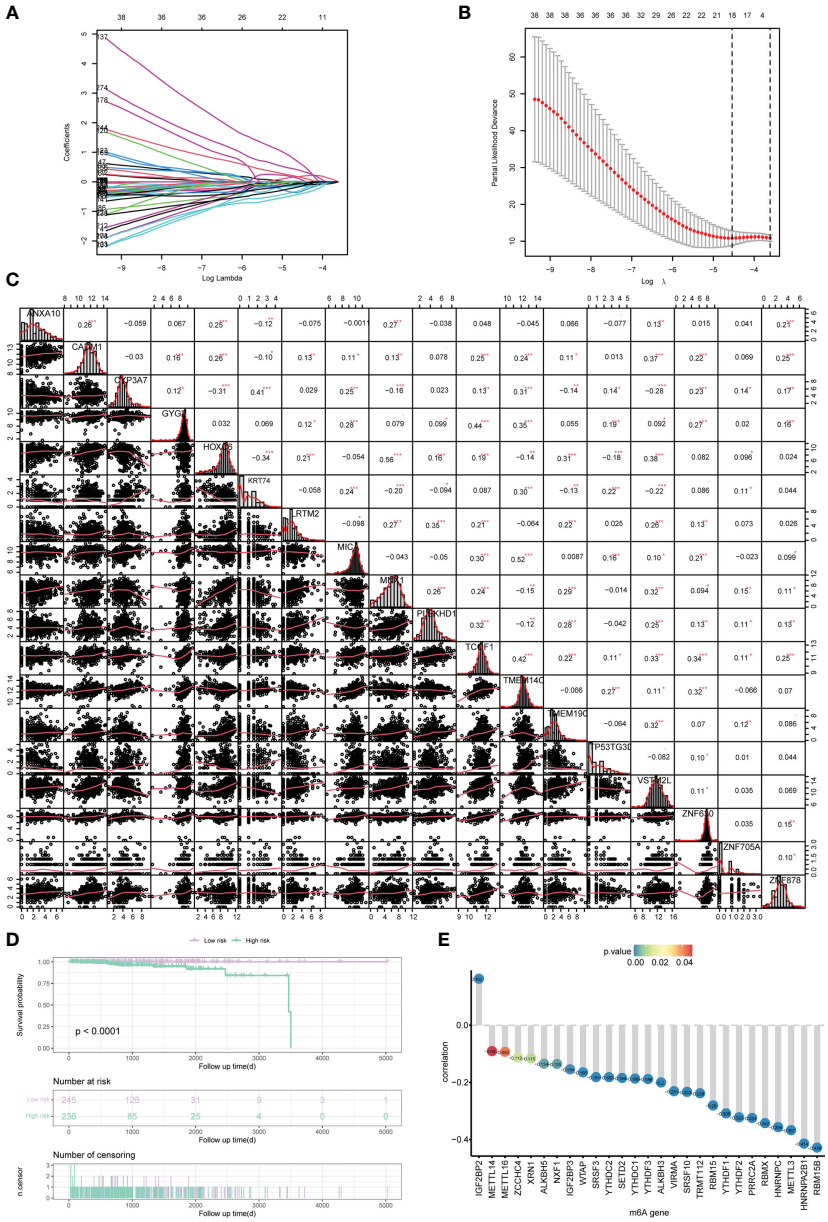


FIGURE 4
Construction of the risk scoring model. **(A, B)** LASSO Cox analysis identified 18 signature genes most associated with OS in the dataset of PRAD patients from TCGA. **(C)** Expression correlation analysis of signature genes in PRAD. **(D)** Kaplan-Meier curve assessed the effect of risk score on overall survival in PRAD patients, with patients with low risk in purple and patients with high risk in green. **(E)** The correlation analysis of m^6A-related genes and risk scores. The horizontal axis shows m^6A-related genes, the vertical axis shows the size of correlation, and the node color indicates the significance level. *P < 0.05, **P < 0.01, *** P < 0.001.

Table 6). Among them, biological processes such as centromere complex assembly, mitotic sister chromatid segregation, DNA replication independent nucleosome organization, kinetochore, and axoneme assembly were inhibited, while biological processes such as myofibril assembly, contractile fiber, muscle filament sliding, sarcomere organization, and structural constituent of muscle were activated (Figures 10A, B). Meanwhile, it was found that pathways involved in hypertrophic cardiomyopathy, dilated cardiomyopathy, arrhythmogenic right ventricular cardiomyopathy, glutathione metabolism, cytokine-cytokine receptor interaction were activated, while pathways involved in cell cycle, maturity onset diabetes of the young, aminoacyl tRNA biosynthesis, mismatch repair, ribosome were inhibited (Figures 10C, D).

PPI network of differentially expressed genes between patients in high-risk group and low-risk group

In order to explore the mechanism affecting the difference between high-risk and low-risk groups, the PPI network of differentially expressed genes in a high-risk group and low-risk group was obtained from the

10.3389/fimmu.2023.1086907

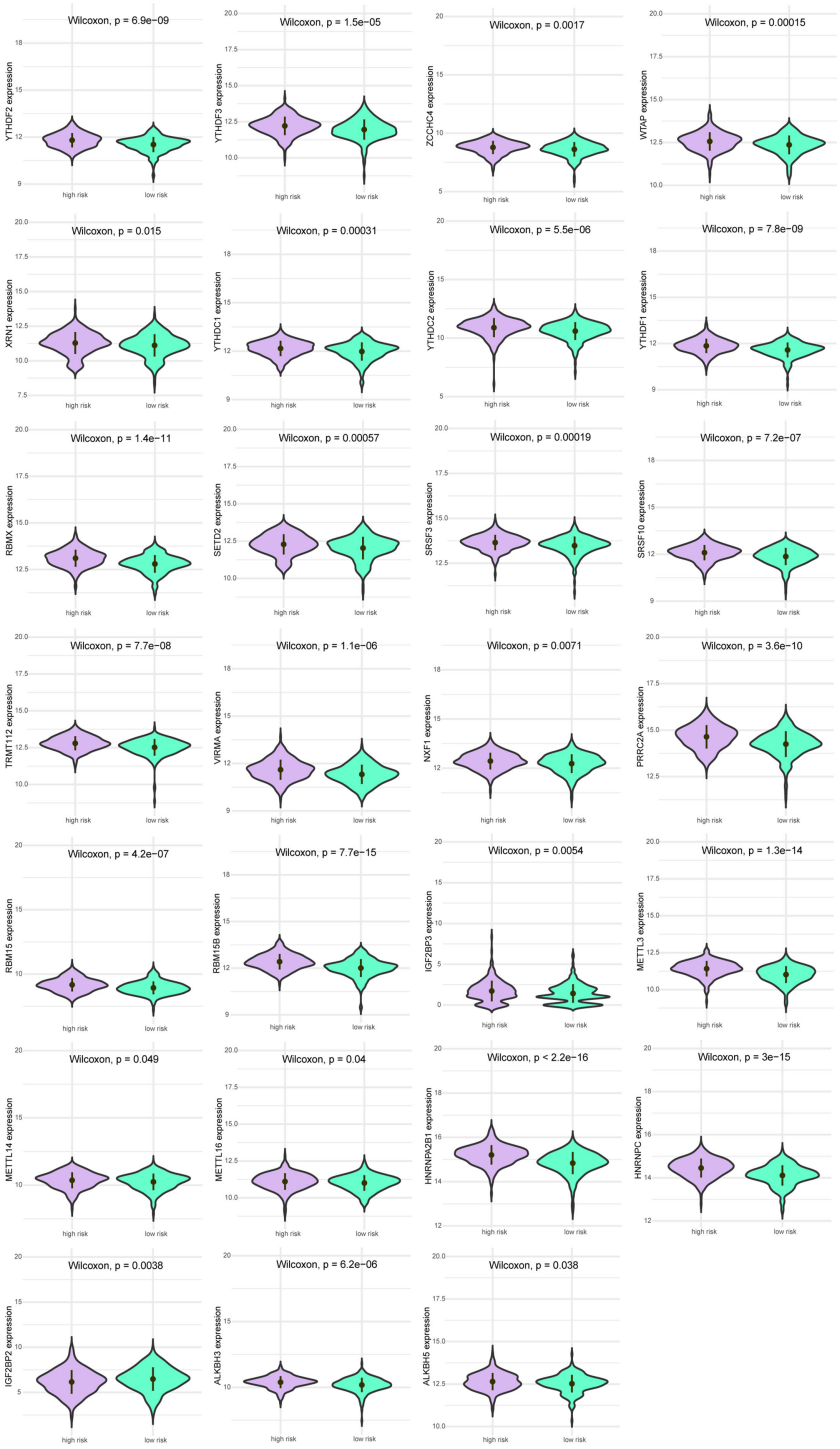


FIGURE 5
Expression levels of m^{6}A gene between patients in a high-risk group and low-risk group. Purple indicates patients with low-risk, and green indicates patients with high-risk.

String database, which was visualized by cytoscape (Figure 11A). The network contained 170 genes, where INS was also closely linked with 32 differentially expressed genes, while both MYH6 and MYH7 were linked with 18 differentially expressed genes. The functional interaction subnet was extracted by MCODE (Figure 11B). The ACTA1, ACTC1, and MYH4 in the subnet were all linked to multiple DEGs in PPI. To verify the functions of genes in the PPI, ClueGO functional enrichment analysis was performed, which showed that genes in PPI were significantly enriched in biological functions including ion transmembrane transporter activity, phosphorylative, regulation of serine-type endopeptidase activity mechanism, endopeptidase inhibitor activity, and glucuronosyltransferase activity (Figure 11C).

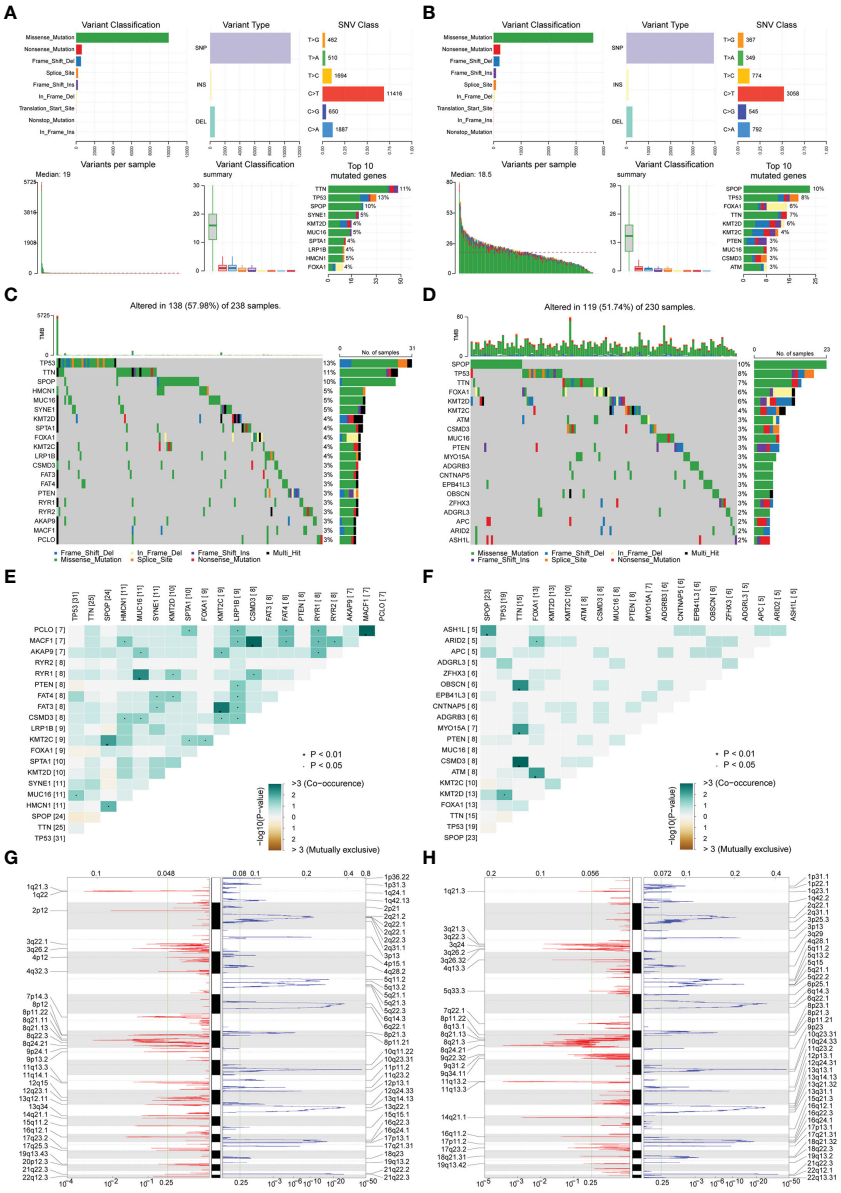


FIGURE 6
Correlation analysis of risk scores and genomic characteristics. **(A, B)** Summary data on mutation for patients with low-risk and patients with high-risk. **(C, D)** Statistics of top 20 mutant genes in patients with high-risk and patients with low-risk. Samples are ordered according to somatic nonsynonymous mutational burden and genes are ordered by mutation frequency, with various colors indicating different mutation types. The subsection above the legend shows mutational burden. **(E, F)** Demonstration of synergy and mutational relationships between mutated genes in patients with high-risk and patients with low-risk. **(G, H)** Identified genes with significant amplifications and deletions in patients with high-risk and patients with low-risk. Q-value and change score of GISTIC2.0 (x-axis) versus genomic location (y-axis). Dashed lines indicated centromeres. The green line represents the 0.25 Q-value cut-off point for determining significance. *P < 0.05.

The differentially expressed mRNA and differentially expressed lncRNA were used to construct the mRNA-miRNA network and lncRNA-miRNA network, respectively. The intersection of the miRNAs in the two networks was taken to obtain the mRNA-miRNA-lncRNA network associated with patients in the high-risk group and the low-risk group (Figure 11D). The network contained 17 mRNA-miRNA-lncRNA relationships, including 91 miRNAs.

At the same time, the PPI network between m^{6}A-related genes was constructed (Figure 11E). The network contained 278 interaction relationships and 34 m^{6}A-related genes, among which METTL3, YTHDF1, and YTHDF3 were the three nodes with the highest degree.

Similarly, the mRNA-miRNA network of m^{6}A-related genes was constructed (Figure 11F), and the network contained 34 m^{6}A-related genes and 1121 miRNAs. The top 5 m^{6}A-related genes were IGF2BP1 regulated by 241 miRNAs, HNRNPA2B1 regulated by 207 miRNAs, YTHDF1 regulated by 155 miRNAs, PRRC2A regulated by 144 miRNAs, and YTHDF3 regulated by 143 miRNAs. The top 4 of miRNAs that controlled multiple m^{6}A-related genes simultaneously were hsa-mir-1-3p controlling 24 m^{6}A-related genes, hsa-let-7b-5p controlling 20 m^{6}A-related genes, hsa-mir-124-3p controlling 19 m^{6}A-related genes, and hsa-mir-16-5p controlling 17 m^{6}A-related genes. Moreover, the heatmap of m6A-related genes, risk scores

10.3389/fimmu.2023.1086907

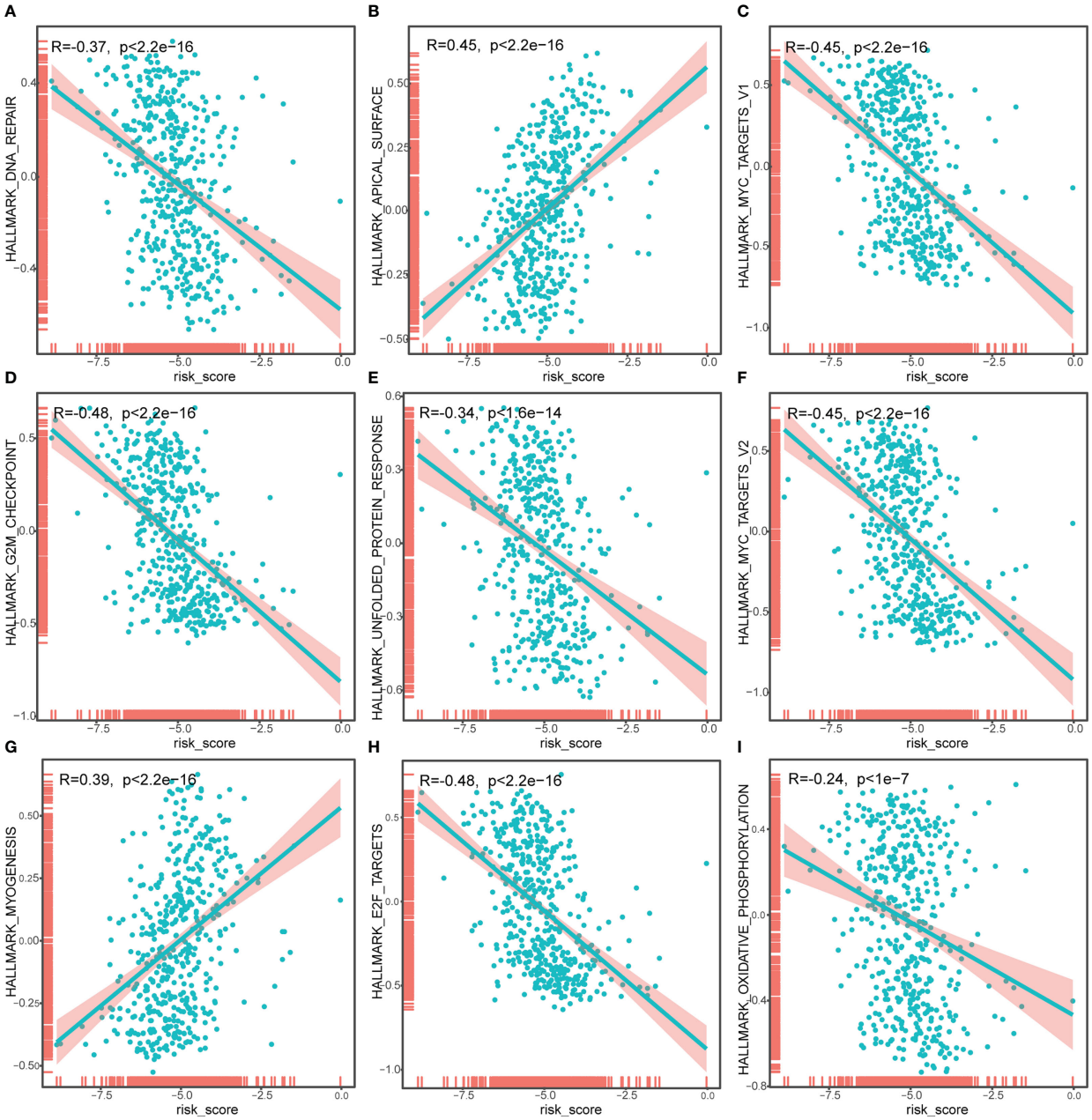


FIGURE 7
Correlation analysis of risk score and Hallmark_DNA_repair **(A)**, Hallmark_APICAL_surface **(B)**, Hallmark_myc_targets_v1 **(C)**, Hallmark_G2M_checkpoint **(D)**, Hallmark_unfolded_protein_response **(E)**, Hallmark_myc_targets_v2 **(F)**, Hallmark_myogenesis **(G)**, Hallmark_E2F_targets **(H)**, Hallmark_oxidative_phosphorylation **(I)**. The horizontal axis represents the risk score, and the vertical axis represents the enrichment score of the hallmark.

combined with clinicopathological characteristics was shown to further explore the relationship among risk scores, m6A-related genes and clinicopathological characteristics (Figure 11G).

Differences in immune characteristics and drug sensitivity prediction of patients in high-risk group and low-risk group

Next, the effect of risk score on the overall immune profile and different infiltration levels of immune cell in PRAD patients was assessed, revealing that compared with the low-risk group, the levels of M1 macrophages and M2 macrophages in the high-risk group significantly increased, while the levels of mast cells resting and T cells CD4 memory resting significantly decreased ($P< 0.05$, Figure 12A). We further calculated the correlation between the level of immune cell and the expression level of m^6A-related gene (Figure 12B), finding that resting memory CD4+ T cells and regulatory T cells (Tregs) were strongly correlated with multiple m^6A-related genes ($P< 0.05$).

We also predicted the drug sensitivity of PRAD patients in the high-risk group and low-risk group, finding that patients in the low-

10.3389/fimmu.2023.1086907

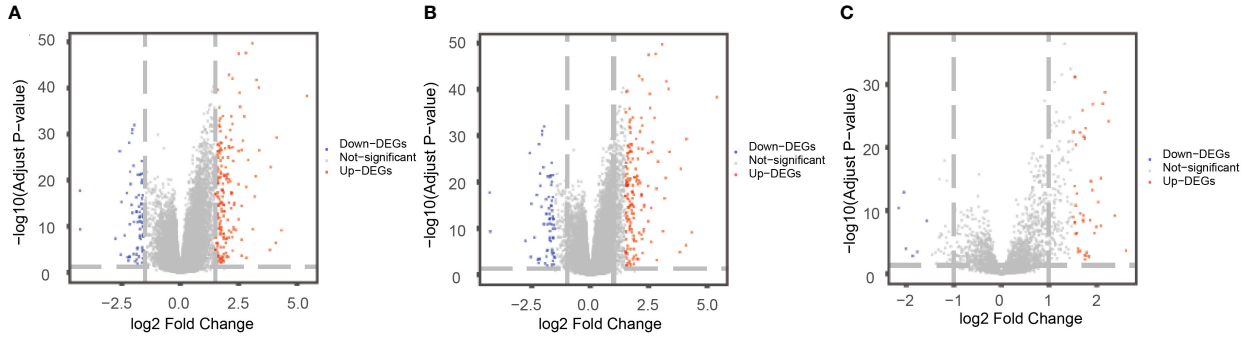


FIGURE 8
Differentially expressed mRNAs **(A)**, miRNAs **(B)**, lncRNAs **(C)** between patients in a high-risk group and a low-risk group. The horizontal axis was logFC; the vertical axis was -log10 (Adjust P-value). Red nodes represent up-regulated differentially expressed genes, blue nodes represent down-regulated differentially expressed genes, and gray nodes represent genes that were not significantly differentially expressed.

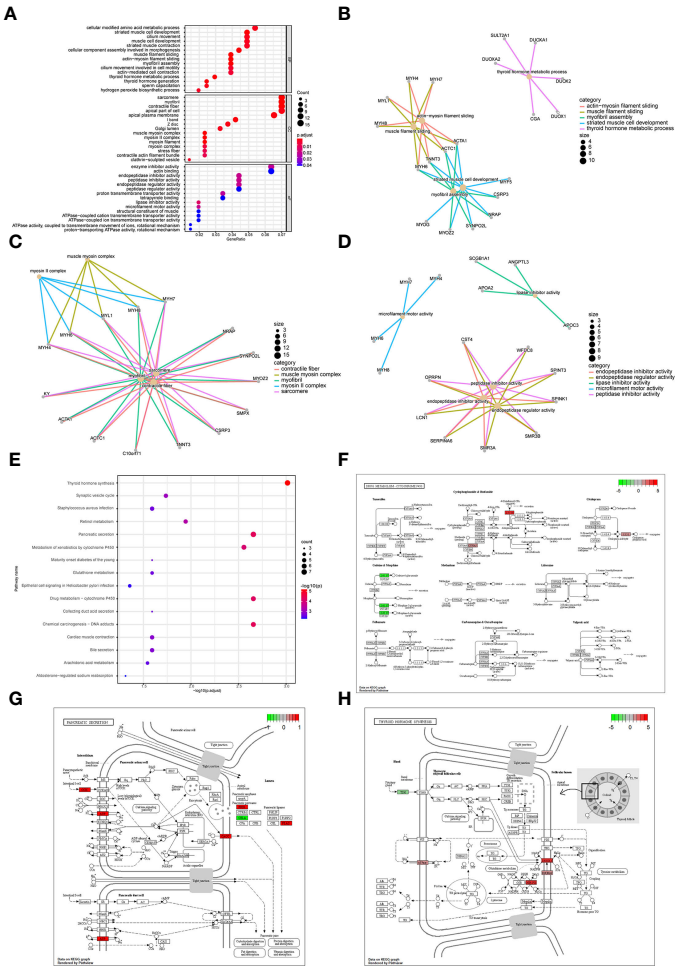


FIGURE 9
Enrichment analysis of differentially expressed genes between patients in a high-risk group and low-risk group. **(A)** GO functional enrichment analysis, the vertical axis is gene ratio, the horizontal axis is GO terms, the node color indicates -log10 (p value), and the node size indicates the number of genes contained in the current GO Term. **(B)** The first 5 items of BP are listed, the node's size indicates the number of genes contained in the current GO Term, and the different colors indicate different GO Term. **(C)** The first 5 items of CC are listed, the node size indicates the number of genes contained in the current GO Term, and the different colors indicate different GO Term. **(D)** The first 5 items of MF are listed, the node size indicates the number of genes contained in the current GO Term, and the different colors indicate different GO Term. **(E)** KEGG pathway enrichment analysis, the horizontal axis was -log10 (p value), the vertical axis is the Pathway name, the node size indicates the number of genes enriched in the pathway, and the node color indicated -log10 (p value). **(F)** KEGG pathway with significant enrichment. hsa00982: Drug metabolism - cytochrome P450. **(G)** KEGG pathway with significant enrichment, hsa04972: Pancreatic secretion. **(H)** KEGG pathway with significant enrichment, hsa04918: Thyroid hormone synthesis.

10.3389/fimmu.2023.1086907

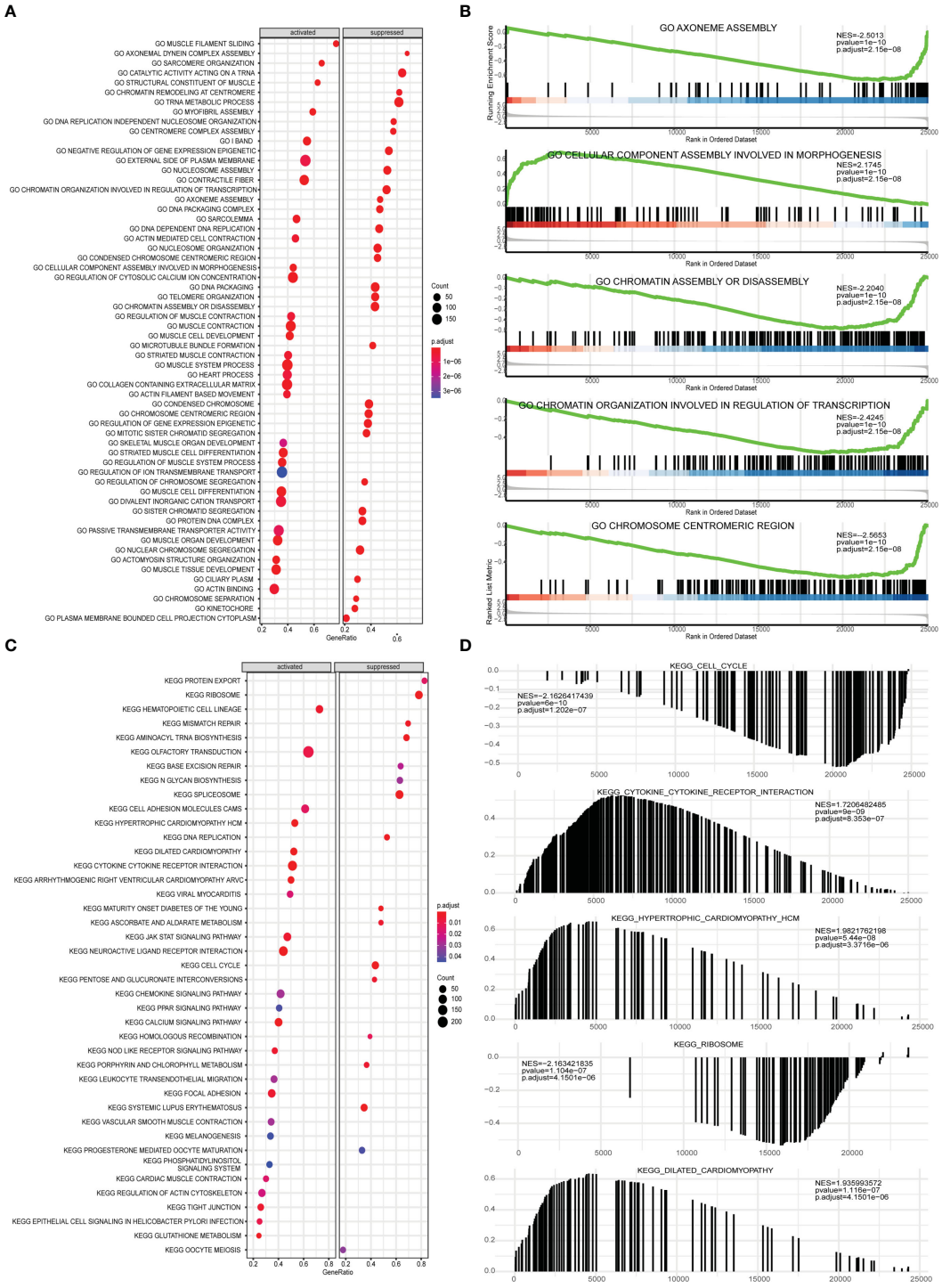


FIGURE 10
GSEA analysis of high-risk group and low-risk group. **(A)** GSEA-GO analysis of a dataset of PRAD patients from TCGA, the horizontal axis is the gene ratio, the vertical axis is the GO terms, and the color represents -log10 (p value). **(B)** The first 5 items of the GSEA-GO analysis of the entire dataset of PRAD patients from TCGA are shown. **(C)** GSEA-KEGG analysis of dataset of PRAD patients from TCGA, the horizontal axis is the gene ratio, the vertical axis is the GO terms, the node size represents the number of genes enriched in GO terms, and the node color represents log10 (p value). **(D)** The first 5 items of the GSEA-KEGG analysis of dataset of PRAD patients from TCGA.

risk group were more sensitive to PD0325901, trametinib, GSK1059615, dasatinib, PARP_0108 and Z-LLNle-CHO, while patients in the high-risk group were more sensitive to WZ3105, WYE-125132, CD532, pevonedistat, and other drugs (Figure 12C).

Subsequently, risk scores were combined with different clinicopathological characteristics to construct a predictive nomogram to predict OS in PRAD patients (Figure 13A). Moreover, the calibration curves showed good agreement between

10.3389/fimmu.2023.1086907

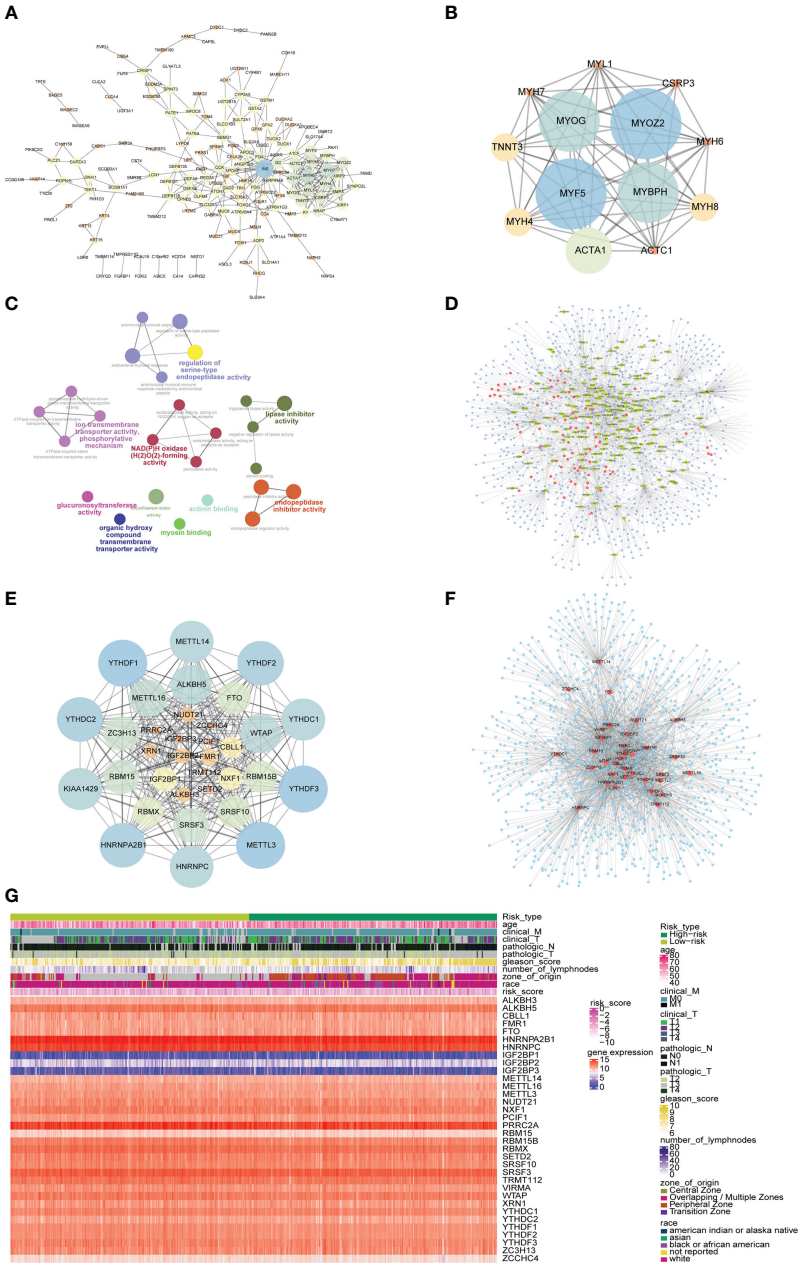


FIGURE 11
PPI network and mRNA-miRNA-lncRNA network of differentially expressed genes. **(A)** PPI network of differentially expressed genes. The node size represents the degree of the node. **(B)** The first subnet in the PPI network of differentially expressed gene. The node size represents the score of mcode. **(C)** Graph of enrichment analysis of PPI network of differentially expressed gene. **(D)** mRNA-miRNA-lncRNA network of differentially expressed genes. Blue nodes represent miRNAs, red nodes represent differentially expressed lncRNAs, and yellow nodes represent differentially expressed mRNAs. **(E)** PPI network of m^6A-related gene. The node size indicates the degree of the node. **(F)** mRNA-miRNA network of m^6A-related gene. Blue nodes represent miRNAs, and red nodes represent m^6A-related genes. **(G)** The heat map of m^6A-related genes, risk scores combined with clinicopathological characteristics.

the 2-, 3-, and 5-year OS estimates by comparing the nomogram and actual value of OS (Figures 13B–D). We also assessed the effect of risk scores on the prognosis of PRAD patients. Dot plot of risk score showed that all death samples belonged to the high-risk group, and as the risk score increased, while the survival time of the patients was shorter (Figure 13E). Univariate and multivariate Cox analysis revealed that risk score was an independent risk factor for predicting the prognosis of PRAD patients (Figures 13F, G; Supplementary Table 6). By analyzing the correlation between m^6A-related genes and risk scores or clinicopathological characteristics, it was found that the patients in the high-risk group were more in the middle and late stages. Patients in the high-risk group were older, and the cancerous sites were mostly in the central area with multiple points. m^6A-related genes were significantly differentially expressed between patients in high-risk group and the low-risk group (Figure 13G). Besides, the time-ROC

10.3389/fimmu.2023.1086907

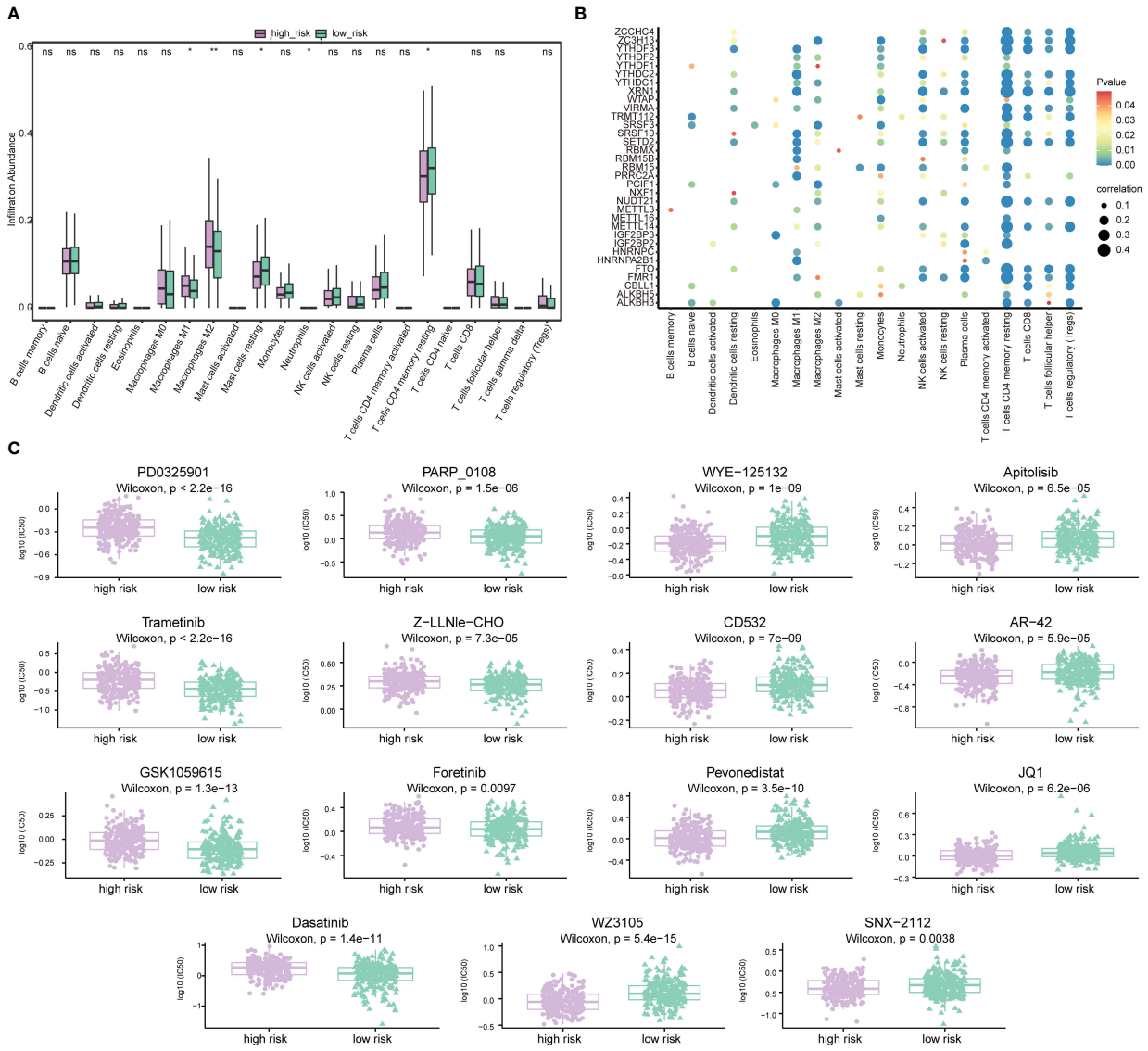


FIGURE 12
Association of risk score-m^{6}A-related gene-immune cell infiltration and drug sensitivity. **(A)** Histogram of the level of immune cells infiltration between patients in a high-risk group and low-risk group. Light green represents the high-risk group, dark green represents the low-risk group, the horizontal axis represents immune cell subtypes, and the vertical axis represents the infiltration level of cells. **(B)** Correlation diagram between m^{6}A-related genes and immune cells. The horizontal axis represents immune cell subtypes, the vertical axis represents m^{6}A-related genes, the node size represents the absolute value of the correlation size, and the node color represents the significance level. **(C)** Differences in drug sensitivity between patients in the high-risk group and low-risk group. The horizontal axis indicates grouping, and the vertical axis indicates -log0 (IC50). *P < 0.05, **P < 0.01, ns, not significant.

also showed that the predictive performance of the prognostic model was 100% for one-year survival, 96.9% for three-year survival, and 97.9% for five-year survival (Figure 13H).

Expression validation of m^{6}A-related gene in PCa cells

Based on the comprehensive analysis of TCGA data and GEO data above, significant differences were found in expression of multiple m^{6}A-related genes between PCa tissues and normal tissues, which were further verified at the cellular level. By comparing the expression of m^{6}A-related genes in prostate normal cell line (RWPE-1) and 2 PCa cell lines (22Rv1 and PC3), 8 significantly DEGs were screened out by RT-qPCR, among which METTL3, ALKBH5 and hnRNPA2B1 were highly expressed in PCa cells, while METTL5, YTHDF1, IGF2BP2, IGF2BP3 and hnRNPC were lowly expressed in PCa cells (Figure 14A). Moreover, three m^{6}A-related genes with the same expression trend as RT-qPCR results were screened out by Western blot, including METTL3, METTL5 and YTHDF1 (Figure 14B).

10.3389/fimmu.2023.1086907

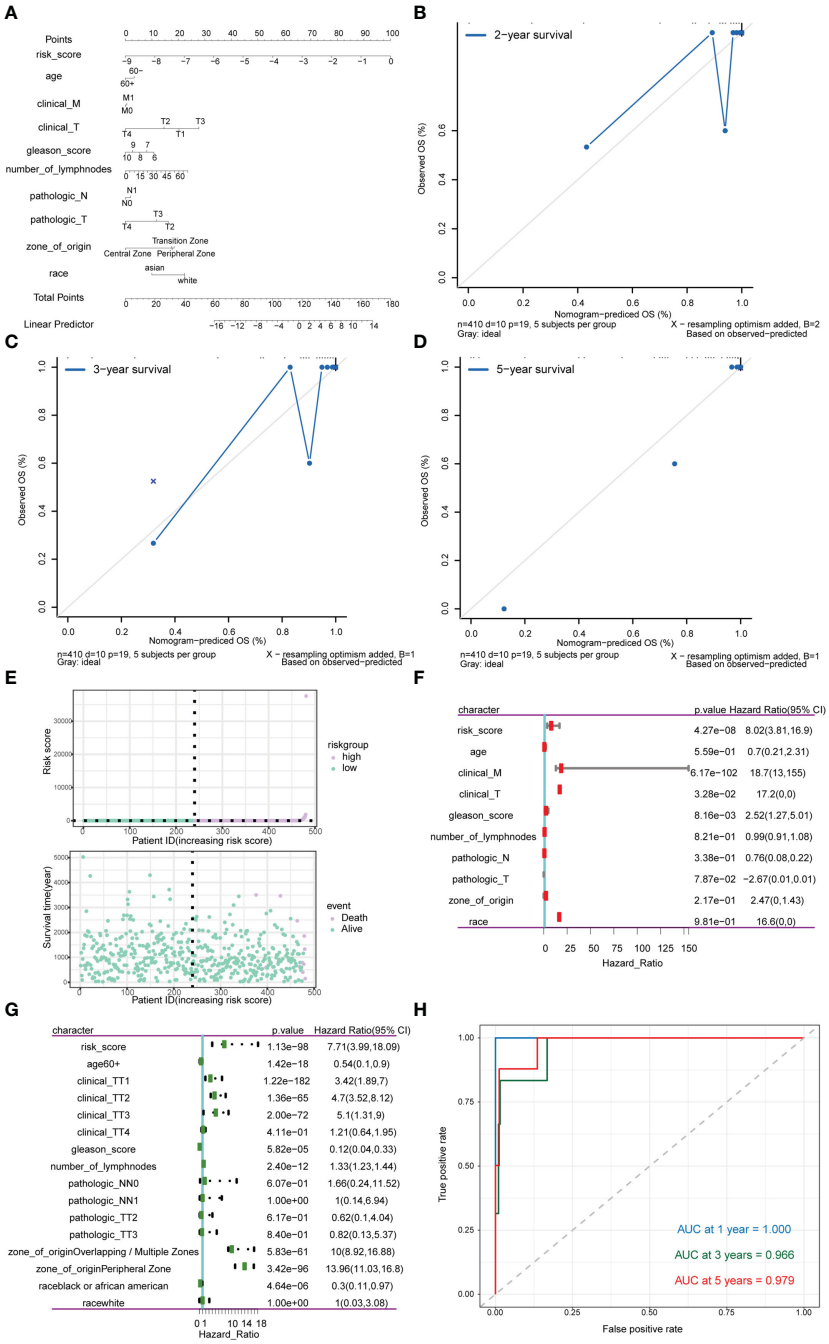


FIGURE 13
Analysis of the predictive power of risk scores for prognosis in PRAD patients. **(A–D)** Calibration curves of the nomogram. The horizontal axis is the survival predicted by the nomogram, and the vertical axis is the actual survival with repeated 1000 times each time. The curve shows the model had good predictive value of prognosis of patients for 2 years, 3 years and 5 years. **(E)** The risk group of the risk model. The horizontal axis shows the order of patient risk gradually increasing; the purple nodes represent patients with high-risk, the green nodes represent patients with low-risk, the vertical axis of the upper graph indicates the patient's transformed risk score, and the vertical axis of the lower graph indicate survival time of patients. **(F)** HR and P values for risk scores by Univariate Cox regression analysis combined with clinicopathological features. **(G)** Multivariate Cox regression analysis of risk score combined with HR and P values of clinicopathological characteristics. The analysis showed that score of m^6A group was an independent risk factor for the prognosis of PRAD patients. **(H)** Time-ROC curve of nomogram model for predicting 1-year survival, 3-year survival and 5-year survival of PRAD patients.

Discussion

Cumulative evidence over the two decades suggested that various types of RNA modifications, such as 5-methylcytosine (m^5C), m^6A, inosine (I), and 2′-O-methylation (2′-O-Me) are implicated in PCa (6, 36–38). Among them, m^6A has attracted the most attention due to the wide distribution of this modification across the human transcriptome. Yet, the interplay between m^6A and PCa development is still not clearly understood. In this study, we systematically examined the relationship between expression of

10.3389/fimmu.2023.1086907

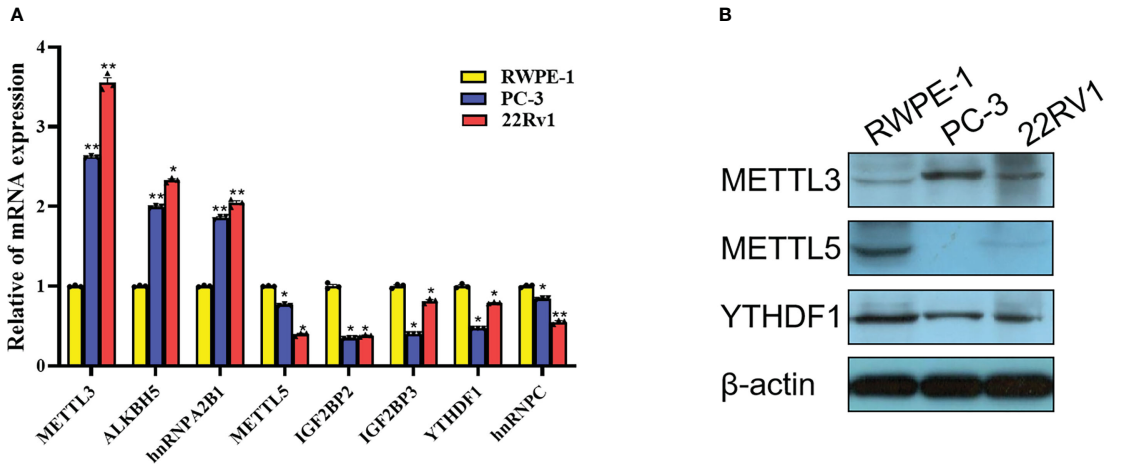


FIGURE 14
Expression validation of m^{6}A-related gene in PCa cells. **(A)** Differences in mRNA expression of 8 m^{6}A-related genes in 22Rv1, PC3 and RWPE-1 by RT-qPCR. **(B)** Differences in protein expression of METTL3, METTL5 and YTHDF1 in 22Rv1, PC3 and RWPE-1 by western blot. *P<0.05, **P<0.01.

m^{6}A regulators and progression/prognosis of PCa with the help of multiple bioinformatic tools. In addition, expression patterns of three candidates, i.e., METTL3, METTL5 and YTHDF1, have been successfully validated by experimental approaches.

As an important enzyme catalyzing the formation of m^{6}A, METTL3 forms an m^{6}A methyltransferase complex with METTL14, WTAP, and VIRMA to confer m^{6}A marks to its binding RNA transcripts (39). One study revealed that METTL3 inhibits apoptosis of PCa cells *via* Sonic Hedgehog (SHH)-GLI pathway, indicating an oncogenic role of METTL3 during PCa progression (40). Another study demonstrated that METTL3 regulates the expression of Integrin β1 (ITGB1) through m^{6}A-HuR-dependent mechanism, which subsequently promotes the bone metastasis of PCa (41). Notably, MYC, a well-known oncogene in PCa, was recently identified as a functional target of METTL3-mediated m^{6}A modification. As a result, over-expression of MYC was sufficient to rescue the inhibitory effect of METTL3 knockdown on the tumorigenic activities of PCa cells (42). Consistent with these previous studies, we re-confirmed the elevated expression of METTL3 in PCa cells, identifying it as the key node of the PPI network and further unveiling its potential in the prognosis of advanced PCa.

Other than METTL3, which is responsible for more than 100,000 methylation events in humans, methyltransferase of METTL5 can only catalyze m^{6}A in human 18S rRNA at position A1832 site, thus participating in translational control (43). Dysregulation of METTL5 has been revealed in breast cancer, pancreatic cancer and gastric cancer (44–46). To the best of our knowledge, this is the first study that reported METTL5 being downregulated in PCa samples compared to normal control. Considering the fact that METTL5 is mostly found to be upregulated in other cancer types and gas oncogenic functions, it will be interesting to investigate the reason for the downregulation of METTL5 in PCa and uncover its clinical relevance.

As an m^{6}A reader, YTHDF1 interacts with several translation initiation factors to mediate the translation of m^{6}A-modified transcripts (47). A recent study suggested that YTHDF1 is highly expressed in both PCa tissues and promotes the proliferation of PCa cells by regulating TRIM44 (48). Surprisingly, although we also identified YTHDF1 as a key node of both PPI and mRNA-miRNA networks, both RT-qPCR and western blot results showed a significant decrease of YTHDF1 in PCa cells compared to normal RWPE-1 cell line. This discrepancy may reflect the complexity of m^{6}A-related regulation in PCa, which should be further investigated.

Increasing studies have revealed the m^{6}A regulatory patterns of PCa and correlated these modification patterns with the tumor immune cell infiltration microenvironment characteristics (49–51). In addition, a recent paper found that m^{6}A reader HNRNPC can regulate Treg cell abundance as a possible mechanism for m^{6}A methylation-mediated response against CTLA-4, indicating that activation of the immune microenvironment by targeting m^{6}A regulators may serve as a potential therapeutic approach for advanced PCa(52).Our study synthetically analyzed the relationship between the expression of m^{6}A regulators and immune characteristics and drug sensitivity of PCa patients. In accordance with the previous reports, we confirmed that resting memory CD4+ T cells and Tregs are highly correlated with m^{6}A-related genes (53, 54), while both high- and low-risk groups are sensitive to a number of therapeutic drugs. Some of these drugs are known to be effective in the treatment of PCa and have even been approved for clinical use (55–57). Thus, it will be informative to determine whether combinational treatment of m^{6}A inhibitors and conventional PCa drugs could achieve a synergistic effect.

In the current study, four miRNAs, including hsa-mir-1-3p, hsa-let-7b-5p, hsa-mir-124-3p, and hsa-mir-16-5p were ranked as the top miRNAs, which dedicate the expression of m^{6}A regulators. As expected, most of them have been validated to be closely associated with PCa progression and metastasis (58–61), which further confirmed our observations.

It still remains some limitations in this study. For instance, although the dysregulation m^{6}A-related genes have been validated

in a number of PCa cell lines, additional studies are needed to investigate the change of global m^6A level in PCa specimen as compared with normal control. More importantly, the underlying mechanism by which the m^6A modification is modulated in response to oncogenic signals during PCa development is yet to be discovered. Future efforts should be made to systematically deconstruct how the m^6A-targeting axis promotes PCa tumorigenesis and unveil its clinical relevance.

Conclusions

The present study systematically evaluated the expression pattern, functional network, and potential prognostic value of m^6A regulators in PCa, which may provide novel insights into the understanding of PCa molecular pathology and facilitate the risk surveillance and clinical decision-making for patients diagnosed with PCa.

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

CL: Conception/design, collection and/or assembly of data, data analysis and interpretation, and manuscript writing. DP: Provision of study material, collection and/or assembly of data, and manuscript writing. YG, LZ, WH, BW, PY, and WX: Provision of study material and collection and/or assembly of data. LW: Conception/design and final approval of manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by The National Natural Science Foundation of China (82101694).

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.2023.1086907/full#supplementary-material

References

1. Siegel R, Miller K, Wagle N, Jemal A. Cancer statistic. *CA: Cancer J Clin* (2023) 73:17–48. doi: 10.3322/caac.21763

2. He HR, Liang L, Han DD, Xu FS, Lyu J. Different trends in the incidence and mortality rates of prostate cancer between China and the USA: A joinpoint and age-Period-Cohort analysis. *Front Med* (2022) 9. doi: 10.3389/fmed.2022.824464

3. Boccaletto P, Stefaniak F, Ray A, Cappannini A, Mukherjee S, Purta E, et al. MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Res* (2022) 50:D231–5. doi: 10.1093/nar/gkab1083

4. Yue Y, Liu J, He C. RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes Dev* (2015) 29:1343–55. doi: 10.1101/gad.262766.115

5. Jiang X, Liu B, Nie Z, Duan L, Xiong Q, Jin Z, et al. The role of m6A modification in the biological functions and diseases. *Signal Transduct Target Ther* (2021) 6:74. doi: 10.1038/s41392-020-00450-x

6. Sun T, Wu R, Ming L. The role of m6A RNA methylation in cancer. *BioMed Pharmacother* (2019) 112:108613. doi: 10.1016/j.biopha.2019.108613

7. Zaccara S, Ries RJ, Jaffrey SR. Reading, writing and erasing mRNA methylation. *Nat Rev Mol Cell Biol* (2019) 20:608–24. doi: 10.1038/s41580-019-0168-5

8. Xu J, Liu Y, Liu J, Xu T, Cheng G, Shou Y, et al. The identification of critical mA RNA methylation regulators as malignant prognosis factors in prostate adenocarcinoma. *Front Genet* (2020) 11:602485. doi: 10.3389/fgene.2020.602485

9. Liu Q, Li Z, He L, Li K, Hu C, Chen J, et al. Molecular characterization and clinical relevance of n-methyladenosine regulators in metastatic prostate cancer. *Front Oncol* (2022) 12:914692. doi: 10.3389/fonc.2022.914692

10. Chen Y, Pan C, Wang X, Xu D, Ma Y, Hu J, et al. Silencing of METTL3 effectively hinders invasion and metastasis of prostate cancer cells. *Theranostics* (2021) 11:7640–57. doi: 10.7150/thno.61178

11. Wang Y, Chen J, Gao W, Yang R. METTL14 promotes prostate tumorigenesis by inhibiting THBS1 *via* an m6A-YTHDF2-dependent mechanism. *Cell Death Discov* (2022) 8:143. doi: 10.1038/s41420-022-00939-0

12. Zou L, Chen W, Zhou X, Yang T, Luo J, Long Z, et al. N6-methyladenosine demethylase FTO suppressed prostate cancer progression by maintaining CLIC4 mRNA stability. *Cell Death Discov* (2022) 8:184. doi: 10.1038/s41420-022-01003-7

13. Zhang S, Lv C, Niu Y, Li C, Li X, Shang Y, et al. RBM3 suppresses stemness remodeling of prostate cancer in bone microenvironment by modulating N6-methyladenosine on CTNNB1 mRNA. *Cell Death Dis* (2023) 14:91. doi: 10.1038/s41419-023-05627-0

14. Zheng Y, Qi F, Li L, Yu B, Cheng Y, Ge M, et al. LncNAP1L6 activates MMP pathway by stabilizing the m6A-modified NAP1L2 to promote malignant progression in prostate cancer. *Cancer Gene Ther* (2023) 30:209–18. doi: 10.1038/s41417-022-00537-3

15. 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

16. Mortensen MM, Hoyer S, Lynnerup AS, Orntoft TF, Sorensen KD, Borre M, et al. Expression profiling of prostate cancer tissue delineates genes associated with recurrence after prostatectomy. *Sci Rep* (2015) 5:16018. doi: 10.1038/srep16018

17. Meller S, Meyer HA, Bethan B, Dietrich D, Maldonado SG, Lein M, et al. Integration of tissue metabolomics, transcriptomics and immunohistochemistry reveals ERG- and gleason score-specific metabolomic alterations in prostate cancer. *Oncotarget* (2016) 7:1421–38. doi: 10.18632/oncotarget.6370

18. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43:e47. doi: 10.1093/nar/gkv007

19. Dalman M, Deeter A, Nimishakavi G, Duan Z. Fold change and p-value cutoffs significantly alter microarray interpretations. *BMC Bioinf* (2012) S11(Suppl 2):13. doi: 10.1186/1471-2105-13-S2-S11

20. 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

21. Wang C, Liang C. MSIpred: A python package for tumor microsatellite instability classification from tumor mutation annotation data using a support vector machine. *Sci Rep* (2018) 8:17546. doi: 10.1038/s41598-018-35682-z

22. Hanzelmann S, Castelo R, Guinney J. GSVA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf* (2013) 14:7. doi: 10.1186/1471-2105-14-7

23. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst* (2015) 1:417–25. doi: 10.1016/j.cels.2015.12.004

24. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. the gene ontology consortium. *Nat Genet* (2000) 25:25–9. doi: 10.1038/75556

25. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* (2000) 28:27–30. doi: 10.1093/nar/28.1.27

26. Yu G, Wang L, Han Y, He Q. 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

27. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102

28. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* (2019) 37:773–82. doi: 10.1038/s41587-019-0114-2

29. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* (2013) 41:D955–961. doi: 10.1093/nar/gks1111

30. 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. *Brief Bioinform* (2021) 22(6):bbab260. doi: 10.1093/bib/bbab260

31. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* (2019) 47:D607–13. doi: 10.1093/nar/gky1131

32. 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

33. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: A cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* (2009) 25:1091–3. doi: 10.1093/bioinformatics/btp101

34. Chang L, Zhou G, Soufan O, Xia J. miRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology. *Nucleic Acids Res* (2020) 48:W244–51. doi: 10.1093/nar/gkaa467

35. Zhang X, Hong R, Chen W, Xu M, Wang L. The role of long noncoding RNA in major human disease. *Bioorg Chem* (2019) 92:103214. doi: 10.1016/j.bioorg.2019.103214

36. Salameh A, Lee AK, Cardo-Vila M, Nunes DN, Efstathiou E, Staquicini FI, et al. PRUNE2 is a human prostate cancer suppressor regulated by the intronic long noncoding RNA PCA3. *Proc Natl Acad Sci USA* (2015) 112:8403–8. doi: 10.1073/pnas.1507882112

37. Wang K, Zhong W, Long Z, Guo Y, Zhong C, Yang T, et al. 5-methylcytosine RNA methyltransferases-related long non-coding RNA to develop and validate biochemical recurrence signature in prostate cancer. *Front Mol Biosci* (2021) 8:775304. doi: 10.3389/fmolb.2021.775304

38. Yi Y, Li Y, Meng Q, Li Q, Li F, Lu B, et al. A PRC2-independent function for EZH2 in regulating rRNA 2'-O methylation and IRES-dependent translation. *Nat Cell Biol* (2021) 23:341–54. doi: 10.1038/s41556-021-00653-6

39. Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol* (2014) 10:93–5. doi: 10.1038/nchembio.1432

40. Cai J, Yang F, Zhan H, Situ J, Li W, Mao Y, et al. RNA m(6)A methyltransferase METTL3 promotes the growth of prostate cancer by regulating hedgehog pathway. *Onco Targets Ther* (2019) 12:9143–52. doi: 10.2147/OTT.S226796

41. Li E, Wei B, Wang X, Kang R. METTL3 enhances cell adhesion through stabilizing integrin beta1 mRNA *via* an m6A-HuR-dependent mechanism in prostatic carcinoma. *Am J Cancer Res* (2020) 10:1012–25.

42. Yuan Y, Du Y, Wang L, Liu X. The M6A methyltransferase METTL3 promotes the development and progression of prostate carcinoma *via* mediating MYC methylation. *J Cancer* (2020) 11:3588–95. doi: 10.7150/jca.42338

43. Xing M, Liu Q, Mao C, Zeng H, Zhang X, Zhao S, et al. The 18S rRNA m(6) a methyltransferase METTL5 promotes mouse embryonic stem cell differentiation. *EMBO Rep* (2020) 21:e49863. doi: 10.15252/embr.201949863

44. Rong B, Zhang Q, Wan J, Xing S, Dai R, Li Y, et al. Ribosome 18S m(6)A methyltransferase METTL5 promotes translation initiation and breast cancer cell growth. *Cell Rep* (2020) 33:108544. doi: 10.1016/j.celrep.2020.108544

45. Wang Z, Liu J, Yang Y, Xing C, Jing J, Yuan Y. Expression and prognostic potential of ribosome 18S RNA m(6)A methyltransferase METTL5 in gastric cancer. *Cancer Cell Int* (2021) 21:569. doi: 10.1186/s12935-021-02274-3

46. Huang H, Li H, Pan R, Wang S, Khan AA, Zhao Y, et al. Ribosome 18S m(6)A methyltransferase METTL5 promotes pancreatic cancer progression by modulating cMyc translation. *Int J Oncol* (2022) 60(1):9.

47. Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, et al. N(6)-methyladenosine modulates messenger RNA translation efficiency. *Cell* (2015) 161:1388–99. doi: 10.1016/j.cell.2015.05.014

48. Li W, Chen G, Feng Z, Zhu B, Zhou L, Zhang Y, et al. YTHDF1 promotes the proliferation, migration, and invasion of prostate cancer cells by regulating TRIM44. *Genes Genomics* (2021) 43:1413–21. doi: 10.1007/s13258-021-01175-z

49. Zhao Y, Sun H, Zheng J, Shao C. Analysis of RNA m(6)A methylation regulators and tumour immune cell infiltration characterization in prostate cancer. *Artif Cells Nanomed Biotechnol* (2021) 49:407–35. doi: 10.1080/21691401.2021.1912759

50. Ye X, Wang R, Yu X, Wang Z, Hu H, Zhang H. m(6)A/ m(1)A /m(5)C/m(7)G-related methylation modification patterns and immune characterization in prostate cancer. *Front Pharmacol* (2022) 13:1030766. doi: 10.3389/fphar.2022.1030766

51. Zou C, He Q, Feng Y, Chen M, Zhang D. A m(6)Avalue predictive of prostate cancer stemness, tumor immune landscape and immunotherapy response. *NAR Cancer* (2022) 4:zcac010.

52. Cheng Y, Li L, Wei X, Xu F, Huang X, Qi F, et al. HNRNPC suppresses tumor immune microenvironment by activating treg cells promoting the progression of prostate cancer. *Cancer Sci* (2023). doi: 10.1111/cas.15745

53. Liu Z, Zhong J, Zeng J, Duan X, Lu J, Sun X, et al. Characterization of the m6A-associated tumor immune microenvironment in prostate cancer to aid immunotherapy. *Front Immunol* (2021) 12:735170. doi: 10.3389/fimmu.2021.735170

54. Liu J, Zhang W, Wang J, Lv Z, Xia H, Zhang Z, et al. Construction and validation of N6-methyladenosine long non-coding RNAs signature of prognostic value for early biochemical recurrence of prostate cancer. *J Cancer Res Clin Oncol* (2022). doi: 10.1007/s00432-022-04040-y

55. Araujo JC, Trudel GC, Saad F, Armstrong AJ, Yu EY, Bellmunt J, et al. Docetaxel and dasatinib or placebo in men with metastatic castration-resistant prostate cancer (READY): a randomised, double-blind phase 3 trial. *Lancet Oncol* (2013) 14:1307–16. doi: 10.1016/S1470-2045(13)70479-0

56. Lee JK, Phillips JW, Smith BA, Park JW, Stoyanova T, Mccaffrey EF, et al. N-myc drives neuroendocrine prostate cancer initiated from human prostate epithelial cells. *Cancer Cell* (2016) 29:536–47. doi: 10.1016/j.ccell.2016.03.001

57. Thiery-Vuillemin A, De Bono J, Hussain M, Roubaud G, Procopio G, Shore N, et al. Pain and health-related quality of life with olaparib versus physician's choice of next-generation hormonal drug in patients with metastatic castration-resistant prostate cancer with homologous recombination repair gene alterations (PROfound): An open-label, randomised, phase 3 trial. *Lancet Oncol* (2022) 23:393–405. doi: 10.1016/S1470-2045(22)00017-1

58. Shi XB, Ma AH, Xue L, Li M, Nguyen HG, Yang JC, et al. miR-124 and androgen receptor signaling inhibitors repress prostate cancer growth by downregulating androgen receptor splice variants, EZH2, and src. *Cancer Res* (2015) 75:5309–17. doi: 10.1158/0008-5472.CAN-14-0795

59. Jin W, Chen F, Wang K, Song Y, Fei X, Wu B. miR-15a/miR-16 cluster inhibits invasion of prostate cancer cells by suppressing TGF-beta signaling pathway. *BioMed Pharmacother* (2018) 104:637–44. doi: 10.1016/j.biopha.2018.05.041

60. Li SM, Wu HL, Yu X, Tang K, Wang SG, Ye ZQ, et al. The putative tumour suppressor miR-1-3p modulates prostate cancer cell aggressiveness by repressing E2F5 and PFTK1. *J Exp Clin Cancer Res* (2018) 37:219. doi: 10.1186/s13046-018-0895-z

61. Yi Y, Li Y, Li C, Wu L, Zhao D, Li F, et al. Methylation-dependent and -independent roles of EZH2 synergize in CDCA8 activation in prostate cancer. *Oncogene* (2022) 41:1610–21. doi: 10.1038/s41388-022-02208-x

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 05 April 2023
DOI 10.3389/fonc.2023.1180191

OPEN ACCESS

EDITED BY
Meng Zhou,
Wenzhou Medical University, China

REVIEWED BY
Gen Wu,
Southern Medical University, China
Liu Yang,
First Affiliated Hospital of Army Medical University, China

*CORRESPONDENCE
Xiaoyu Cui
cuixiaoyu126@163.com

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Oncology

RECEIVED 05 March 2023
ACCEPTED 27 March 2023
PUBLISHED 05 April 2023

CITATION
Xia G, Wu S, Luo K and Cui X (2023) By using machine learning and *in vitro* testing, SERPINH1 functions as a novel tumorigenic and immunogenic gene and predicts immunotherapy response in osteosarcoma. *Front. Oncol.* 13:1180191. doi: 10.3389/fonc.2023.1180191

By using machine learning and *in vitro* testing, SERPINH1 functions as a novel tumorigenic and immunogenic gene and predicts immunotherapy response in osteosarcoma

Guang Xia[1], Song Wu[1], Ke Luo[2] and Xiaoyu Cui[3]*

[1]Department of Orthopaedics, Third Xiangya Hospital, Central South University, Changsha, China, [2]Department of Anesthesiology, Hunan Cancer Hospital, Central South University, Changsha, Hunan, China, [3]Department of Anesthesiology, Third Xiangya Hospital, Central South University, Changsha, China

Introduction: The most prevalent bone tumor with a relatively high level of aggressiveness and malignancy is osteosarcoma. The characteristics of the serpin family in osteosarcoma have not been defined.

Methods: In this study, the predictive significance of the serpin superfamily was investigated in the osteosarcoma and Gene Expression Omnibus (GEO) databases from The Cancer Genome Atlas (TCGA).

Results: It was discovered that SERPINH1 is a significant biological marker in osteosarcoma. According to the CCK-8, EdU, and Transwell assays as well as the IHC assay, SERPINH1 may promote osteosarcoma proliferation and migration. It is also more expressed in tumor samples than in healthy samples. SERPINH1 might forecast the effects of immunotherapy. Additionally, immune cells are interacted with through checkpoint, cytokine, and growth factor pathways in osteosarcomas with high SERPINH1 levels. The biological function, immunological characteristics, and treatment response (immunotherapy and chemotherapy responses) of patients with osteosarcoma were successfully predicted using a model related to SERPINH1. SERPINH1 and the SERPINH1-related score predict ferroptosis/pyroptosis/apoptosis/necroptosis in osteosarcoma.

Discussion: The SERPINH1-related score was an effective method for identifying osteosarcoma patients who would respond to immunotherapy and chemotherapy, as well as for predicting the survival outcomes of such patients.

KEYWORDS

SERPINH1, tumor microenvironment, immunotherapy, osteosarcoma, machine learning

Introduction

The most prevalent bone tumor with a relatively high level of aggressiveness and malignancy is osteosarcoma. Children and teenagers are frequently affected by osteosarcoma, which has an annual incidence of 1 to 3 incidences per million persons globally (1). Osteosarcoma typically involves lung metastasis and mostly develops from primitive mesenchymal cells found in bone. Surgery to remove pathogenic lesions and adjuvant chemotherapy are still the go-to treatments for people with osteosarcoma (2). Sometimes, the cure rates are underwhelming. The scientific community has faced a formidable challenge in enhancing the chances of survival for people with osteosarcoma (3). A thorough understanding of osteosarcoma biology has been made possible by the rapid advancement of high-throughput sequencing, growing viability of molecular profiling, and reliable model systems based on extensive bioinformatics research.

A popular family of protease inhibitors called serpin uses conformational changes to block the activity of the target enzyme. The coagulation route, inflammation, immunology, and cancer are just a few of the crucial proteolytic cascades that the serpin superfamily regulates (4). It has been demonstrated that the brain metastasis Serpin superfamily promotes tumor growth and vascular co-option (5). A biomarker for colorectal cancer that interacts with CEA was identified as SERPINB5 (6). The motility and invasiveness of oral carcinoma cells may be aided by overexpressed SERPINB1 (7). Notably, SERPINE2 could encourage osteosarcoma tumor cell proliferation and medication resistance (8). The serpin superfamily in osteosarcoma has not yet been thoroughly examined, yet. It is still unknown if the serpin superfamily has a significant predictive significance in osteosarcoma. As SERPINH1 has been widely studied in pan-cancer except for osteosarcoma (9, 10), the prognostic value of SERPINH1 in osteosarcoma was reasonably expected. We, therefore, paid special attention to SERPINH1 and performed a comprehensive analysis on its role in osteosarcoma.

The prognostic utility of the serpin superfamily was investigated in this study. It was discovered that SERPINH1 is a significant biological marker in osteosarcoma. The physical function, immunological characteristics, and medication response of patients with osteosarcoma were successfully predicted using a SERPINH1-related model.

Abbreviations: TARGET, Therapeutically Applicable Research to Generate Effective Treatments; GEO, Gene Expression Omnibus; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic; AUC, area under the curve; ssGSEA, single-sample gene-set enrichment analysis; ESTIMATE, Estimated Stromal and Immune cells in Malignant Tumor tissues using Expression data; DEGs, differentially expressed genes; GSVA, gene set variation analysis; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; IHC, Immunohistochemistry; RT-qPCR, Real-time quantitative polymerase chain reaction.

Method

Data gathering

The TCGA osteosarcoma and GSE21257 databases were used to gather the osteosarcoma samples. The data were transformed to TPM value. The R package SVA was used for data standardization and batch-to-batch difference removal.

Value for prognosis of SERPINH1

Based on the cutpoint value of SERPINH1 determined using the R package survminer, the osteosarcoma patients were split into two groups. The R package pROC was used to create the time-dependent receiver operating characteristic (ROC) curve for SERPINH1. On SERPINH1, the gene set variation analysis (GSVA) of the Kyoto Encyclopedia of Genes and Genomes (KEGG) and gene ontology (GO) pathways was carried out. The TISCH2 database was used to investigate the SERPINH1 expression pattern in the osteosarcoma tumor microenvironment (TME).

Immunotherapy prediction of SERPINH1

The TIDE was used for the immunotherapy prediction of SERPINH1 (11). The TISMO was used for immunotherapy and cytokine treatment prediction of SERPINH1 (12).

Immunohistochemistry (IHC)

Detailed methods were provided in the Supplementary File following the previous study (13).

Cell culture

Detailed methods were provided in the Supplementary File following the previous study (13).

Small interfering RNA (siRNA) transfection

The siRNAs sequences are as follows: SERPINH1-1 (F: GCAGCAAGCAGCACUACAATT R: UUGUAGUGC UGCUUGCUGCTT), SERPINH1-2 (F: CCAGCCUCAUCAU CCUCAUTT R: AUGAGGAUGAUGAGGCUGGTT), SER PINH1-3 (F: GGCCUAAGGGUGACAAGAUTT R: AUCUUG UCACCCUUAGGCCTT). The exact method is provided in the Supplementary Material.

Real-time quantitative polymerase chain reaction (RT-qPCR)

The primer sequences are SERPINH1 (F: ATATTTAT AGCCAGGTACCTTCTCACC R: TTTTATAGTTGGGAGAGGT TGGGATAG), GAPDH (F: AATGGGCAGCCGTTAGGAAA R: GCCCAATACGACCAAATCAGAG). The exact method is provided in the Supplementary Material.

Cell counting Kit-8 (CCK-8) assay

The U2OS and MNNG/HOS cells were seeded into 96-well plates at 5,000 cells/well density. After 24h, 1/10 volume of CCK-8 reagent (Proteintech, USA) was added to the wells, and the absorbance value was detected at 450nm after one h incubation at 37°C. The exact method is provided in the Supplementary Material.

EdU assay

Detailed methods were provided in the Supplementary File following the previous study (13).

Transwell assay

The migration of U2OS and MNNG/HOS cells was assessed using a Transwell chamber (Corning, USA) with polycarbonic membranes (6.5 mm in diameter and eight μm pore size). Cells in serum-free medium were added into the upper chamber at the density of 5×10^5 cells/ml (200 μl/well), and the culture medium with 10% FBS was added to the lower chamber. After incubating for 48h at 37°C, U2OS and MNNG/HOS cells that penetrated the lower surface were stained with 0.1% crystal violet and counted. The exact method is provided in the Supplementary Material.

SERPINH1 was the subject of a weighted correlation network analysis (WGCNA) in the TCGA dataset. The input was the matrix from the TCGA dataset. For further investigation, the genes from the turquoise module were taken out.

Building the SERPINH1-related score

Between two SERPINH1-related groups, the differentially expressed genes (DEGs) were identified. The prognostic DEGs were discovered using a single-variable Cox regression analysis. Dimension reduction and the creation of the SERPINH1-related score employed the Random Survival Forest method and the least absolute shrinkage and selection operator (LASSO) technique. The expression value of the gene's * coefficient was used to construct the SERPINH1-related score. The R package survival was used to create the Kaplan-Meier survival curve for the SERPINH1-related score. The R package pROC was used to create the time-dependent ROC curve for the SERPINH1-related score. The predicted clinical variables were identified using univariate and multivariate Cox regression analysis.

Immune characteristics of the SERPINH1-related score

The correlation between the SERPINH1-related score and immune infiltrating cells (10 cells from MCPcounter algorithm (14), 28 cells from the ssGSEA algorithm (15), six cells from TIMER algorithm (16)) were analyzed. The association between the SERPINH1-related and microenvironment scores (17) (ESTIMATE, Immune, and Stromal) was investigated.

Single-cell RNA sequencing (scRNA-seq) analysis on SERPINH1

For the scRNA-seq investigation, GSE152048 was employed. The R package "Seurat" was used to annotate the cell type. The R package "iTalk" was used to evaluate the cell communication pattern. The R package "monocle" was used to diagnose the pseudotime trajectory analysis.

Immunotherapy prediction of the SERPINH1-related score

The correlation between the SERPINH1-related score and classical immune checkpoints was analyzed. The association between the SERPINH1-related score and APM (antigen processing and presenting machinery) score, Cytotoxic activity (CYT), T cell-inflamed gene expression profile (GEP), and interferon γ (IFN-γ) were analyzed (18–20). The TIDE was used for immunotherapy prediction of the SERPINH1-related score (11).

Drug prediction using the SERPINH1-related score

The chemotherapy medications associated with the SERPINH1-related score were predicted using the R package OncoPredict (21).

Results

Prognostic value of serpin superfamily and SERPINH1

Figure 1A depicts the location of the serpin superfamily on human chromosomes. In patients with osteosarcoma, SERPINH1 and SERPINA7 were found to be two independent prognostic indicators by univariate Cox regression analysis on the serpin

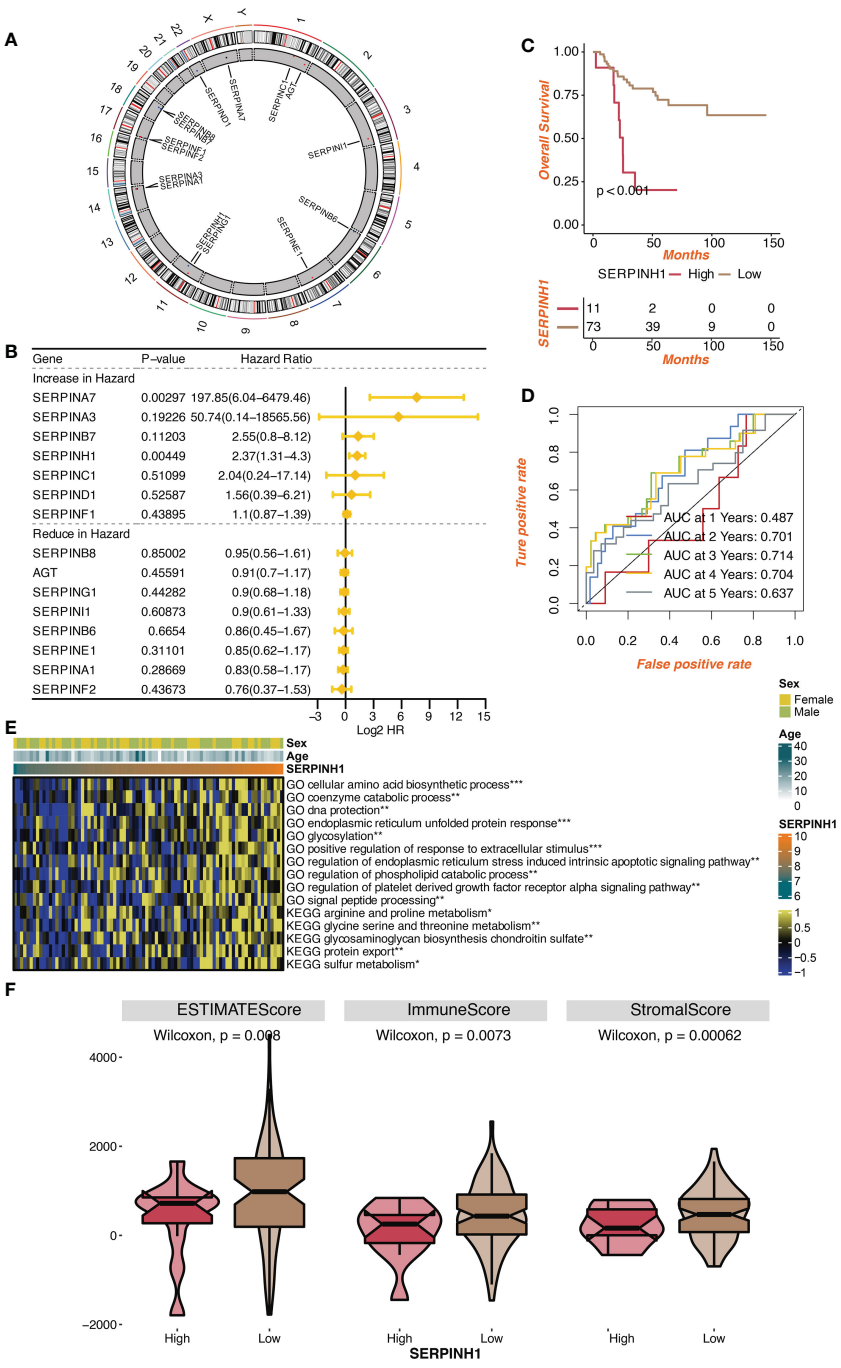


FIGURE 1
Prognostic value of serpin superfamily and SERPINH1. **(A)** Circplot of the serpin superfamily in human chromosomes. **(B)** Univariate Cox regression analysis on the serpin superfamily. **(C)** Survival curves of the two SERPINH1-stratified groups. **(D)** The 1-year, 2-year, 3-year, 4-year, and 5-year ROC regarding SERPINH1. **(E)** The correlation between SERPINH1 and GO, KEGG terms quantified by GSVA. **(F)** The levels of microenvironment scores in two SERPINH1 groups. *P<0.05; **P<0.01; ***P<0.001.

superfamily (Figure 1B). We focused specifically on SERPINH1, which has drawn more attention from the scientific community since it is more likely to contribute to the development of osteosarcoma as a malignancy. Patients with osteosarcoma were divided into two groups according to the expression level of SERPINH1, and it was shown that the two groups had dramatically different survival rates (Figure 1C). The prognostic value of SERPINH1 was validated by the 1-year, 2-year, 3-year, 4-year, and 5-year ROCs, which have the respective values of 0.487, 0.701, 0.714, 0.704, and 0.637. (Figure 1D). SERPINH1 was favorably related with biological processes, including cellular amino acid biosynthetic process, glycosylation, phospholipid carbolic process control, and protein export, according to GSVA of GO and KEGG keywords (Figure 1E). In the low SERPINH1

10.3389/fonc.2023.1180191

group, microenvironment scores such as ESTIMATE, Immune, and Stromal scores were significantly higher (Figure 1F).

Immunotherapy prediction of SERPINH1

SERPINH1 was given the highest priority possible for mechanistic follow-up research using TIDE's regulator prioritization module (Figure 2A). Based on the Normalized Z score calling using Cox-PH regression in the Immunotherapy dataset, high expression of SERPINH1 was discovered in the ICB Mariathasan2018 PDL1 and ICB VanAllen2015 CTLA4 datasets. Based on the Normalized Z score calling from selection log2FC in the CRISPR Screen dataset, it was discovered that SERPINH1 had a high expression level in the Pech 2019 NK E/T=2.5 dataset. Based on normalized expression values from immunosuppressive cell types, increased expression of SERPINH1 was discovered in CAF FAP and MDSC. Eight human immunotherapy datasets that included SERPINH1 achieved AUC values greater than 0.5. (Figure 2B). SERPINH1 exhibited a higher predictive value than TMB and B. Clonality, which respectively gave AUC values of > 0.5 in seven and seven immunotherapy cohorts. The predictive value of SERPINH1 was, however, lower than the MSI score

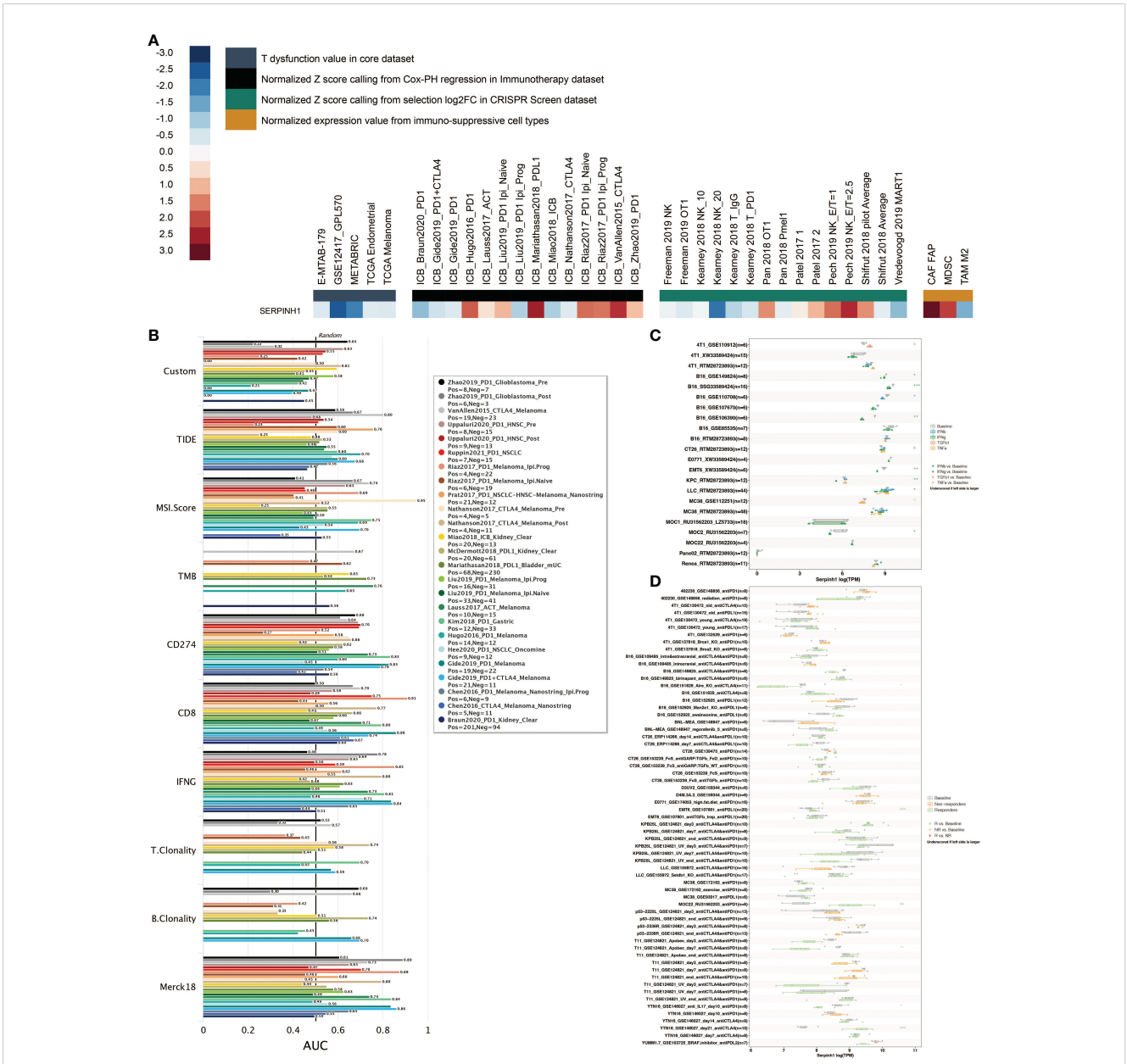


FIGURE 2
Immunotherapy prediction of SERPINH1. **(A)** The regulator prioritization module in TIDE. **(B)** Comparison of SERPINH1 and immunotherapy determinants in human immunotherapy datasets. **(C)** The prediction of SERPINH1 in cytokine treatment (IFN-β, IFN-γ, TGF-β, and TNF-α) in murine datasets. **(D)** The prediction of SERPINH1 in immunotherapy in murine datasets. *P<0.05; **P<0.01; ***P<0.001.

(AUC > 0.5 in 13 immunotherapy cohorts), T.Clonality (AUC > 0.5 in 9 immunotherapy cohorts), CD274 (AUC > 0.5 in 21 immunotherapy cohorts), TIDE (AUC > 0.5 in 18 immunotherapy cohorts), IFNG (AUC > 0.5 in 17 immunotherapy cohorts), and CD8 (AUC > 0.5 in 18 immunotherapy cohorts). In 11 mouse datasets, SERPINH1 could accurately predict the effects of cytokine treatment (IFN, IFN, TGF, and TNF) (Figure 2C). SERPINH1 might accurately forecast immunotherapy in 5 murine datasets (Figure 2D).

Additionally, in six immunotherapy cohorts, including Dizier (AUC = 0.651), Ascierto (AUC = 0.786), Riaz (AUC = 0.608), Homet (AUC = 0.733), Amato (AUC = 0.729), and Kim (AUC = 0.678), SERPINH1 shown excellent sensitivity in predicting immunotherapy response (Figure 3A). Patients in Hugo and Nathanson cohorts had more remarkable survival outcomes with high SERPINH1 expression, while patients in Lauss, Kim, IMvigor210, and Van Allen cohorts had more remarkable survival outcomes with low SERPINH1 expression (Figure 3B).

Pan-cancer analysis on the serpin superfamily

Pan-cancer mutation analysis on the serpin superfamily was performed. The expression pattern of the serpin superfamily between tumor and normal samples is shown in Figure S1A. The methylation difference of the serpin superfamily between tumor and normal samples is shown in Figure S1B. The heterozygous CNV of the serpin superfamily is shown in Figure S1C. The homozygous CNV of the serpin superfamily is shown in Figure S1D. Pan-cancer function analysis on the serpin superfamily was performed. Pathway annotation

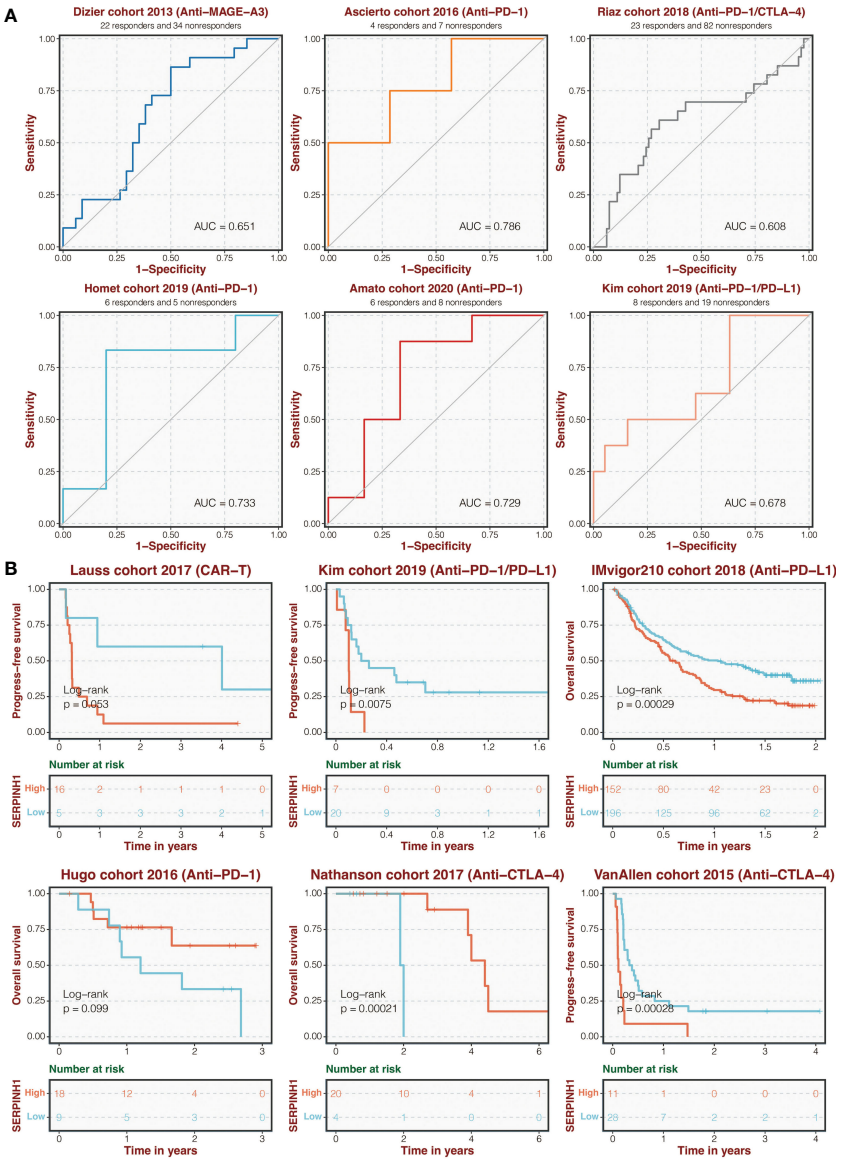


FIGURE 3
Immunotherapy prediction of SERPINH1. **(A)** ROC regarding the sensitivity of SERPINH1 in immunotherapy prediction in six immunotherapy cohorts. **(B)** Survival curves of the two SERPINH1-stratified groups in six immunotherapy cohorts.

on the serpin superfamily is shown in Figure S2A, in which EMT activation showed a strong positive correlation with the serpin superfamily. miRNA network annotation on the serpin superfamily is shown in Figure S2B. Pan-cancer drug prediction analysis on the serpin superfamily was performed based on the CTRP database (Figure S3A) and the GDSC database (Figure S3B), in which SERPINB6, SERPINE1, and SERPINH1 could predict most of the commonly used chemotherapy drugs. We further explored the disease network of SERPINH1, in which SERPINH1 was highly enriched in osteogenesis processes such as osteogenesis imperfecta (Figure S4A). Besides, the protein interaction network of SERPINH1 showed that HSPA8, MIA3, COL1A1, COL1A2, COL1A4, COL26A1, FKBP10, PPIB, LEPRE1, and CRTAP highly connected with SERPINH1 (Figure S4B).

In vitro validation on SERPINH1

It was investigated how SERPINH1 functions biologically in osteosarcoma. The expression of SERPINH1 was considerably reduced in three si-SERPINH1 groups compared to the si-NC group in U2OS and MNNG/HOS cells, according to the results of RT-qPCR (Figure 4A). For the follow-up experiment, si-SERPINH1, which has the most critical ability to suppress the expression of SERPINH1, was utilized. CCK-8 (Figure 4B) and EdU tests (Figure 4C) demonstrated that infection with si-SERPINH1 greatly reduced the ability of U2OS and MNNG/HOS cells to proliferate. Transwell experiment revealed that after exposure to si-SERPINH1, U2OS and MNNG/HOS cells' ability to migrate was drastically reduced (Figure 5A). The IHC results further demonstrated that osteosarcoma tumor tissues had much greater levels of SERPINH1 expression than did normal tissues (Figure 5B).

Construction of the SERPINH1-related score

It was determined that 17 DEGs were predictive genes, of which 15 were risky indicators and two were beneficial markers (Figure 6A). Using machine learning Random Survival Forest, the most effective prognostic genes were chosen (Figure 6B). Additionally, the most potent predictor genes were identified using machine learning LASSO, and the SERPINH1-related score was created (Figure 6C). The

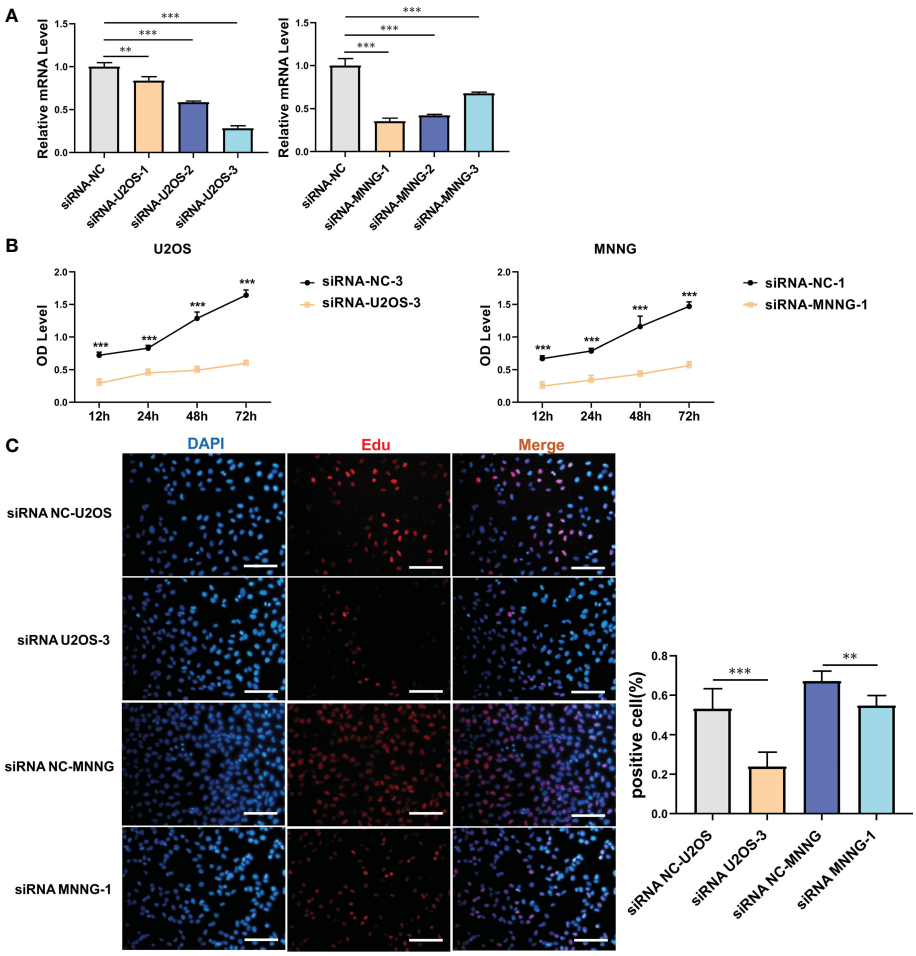


FIGURE 4
In vitro validation on SERPINH1. **(A)** RT-qPCR results of the expression of SERPINH1 in four groups (si-NC, si-SERPINH1-1, si-SERPINH1-2, si-SERPINH1-3) in two cell lines. **(B)** CCK-8 assay in two groups (si-NC, si-SERPINH1). In two cell lines, **(C)** EdU assay in two groups (si-NC, si-SERPINH1). **P<0.01; ***P<0.001.

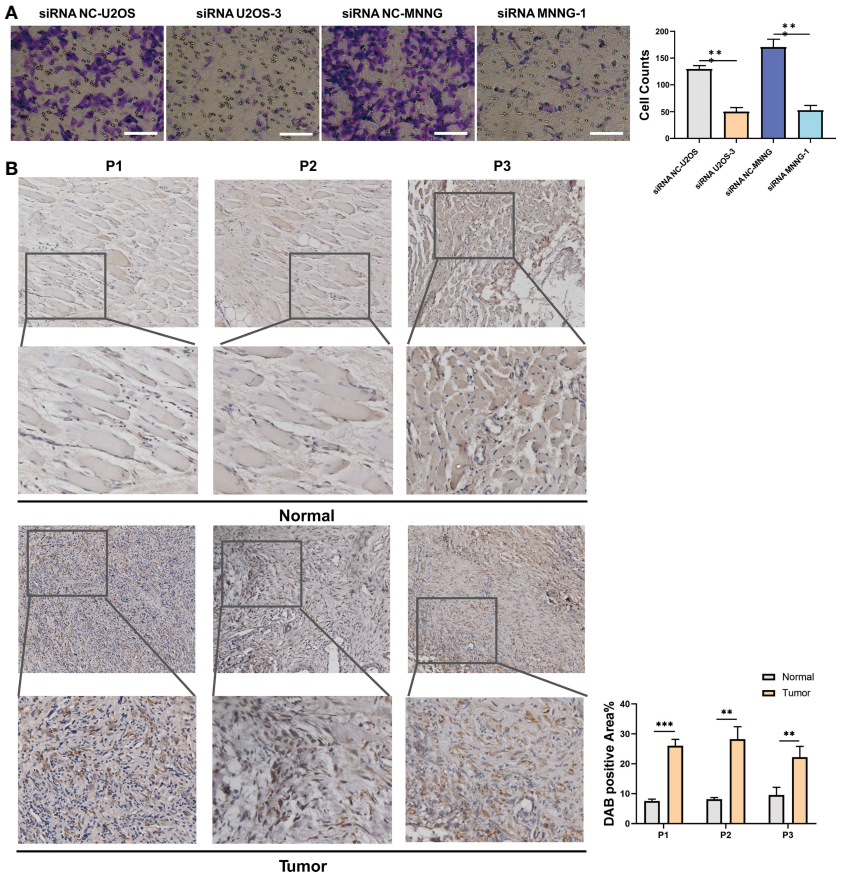


FIGURE 5
In vitro validation on SERPINH1. **(A)** Transwell assay in two groups (si-NC, si-SERPINH1) in two cell lines. **(B)** IHC results of the expression of SERPINH1 in osteosarcoma tumor tissues and normal tissues. **P<0.01; ***P<0.001.

following formula was used to determine the SERPINH1 score: 1.1077*CGREF1, 1.7743*TAC4, 1.4629*PROSER2, and -1.2053*PCDHB7. SERPINH1-related score was used to classify osteosarcoma patients. Figure 7A displays the expression profiles of PCDHB7, TAC4, PROSER2, and CGREF1 in two score-stratified groups associated to SERPINH1. Osteosarcoma patients with higher PCDHB7 expression exhibited better survival rates (Figure 7B), while those with lower CGREF1, PROSER2, and TAC4 expression had worse survival rates (Figures 7C, D, E). The two SERPINH1-related score-stratified groups in the TCGA and GSE21257 datasets had significantly different survival outcomes (Figures 6D, E). In the TCGA dataset, the 1-year, 3-year, and 5-year ROC values for the SERPINH1-related score are 0.822, 0.876, and 0.851, respectively (Figure 6F). In contrast, the GSE21257 dataset's 1-year, 3-year, and 5-year ROCs for the SERPINH1-related score had the respective values of 0.679, 0.674, and 0.683 (Figure 6G), confirming the predictive usefulness of the SERPINH1-related score. The SERPINH1-related score was shown to be an independent predictive factor by univariate and multivariate Cox regression analysis on the clinical variables (Figure 6H).

Immune characteristics of the SERPINH1-related score

Multiple immune infiltrating cells, including T cells, B cells, NK cells, macrophages, mast cells, Type 1 T helper cells, and Type 2 T helper cells, were observed to strongly negatively correlate with the SERPINH1-related score (Figure 8A). According to the aforementioned research, immune-cold microenvironment was present in osteosarcoma patients with high SERPINH1-related scores. In addition, the group with a high SERPINH1-related score exhibited considerably lower levels of the crucial immunotherapy determinants APM, CYT, GEP, and IFN- (Figure 8B). ESTIMATE, Immune, and Stromal scores were considerably lower in the group with high SERPINH1-related scores (Figure 8C). Positive correlations between the SERPINH1-related score and the traditional immunological checkpoints CD274, PD-1, PDCD1LG2, and CTLA-4 were observed (Figure 8D). Most importantly, patients with osteosarcoma who had low SERPINH1-related scores had a higher likelihood of responding to anti-PD-1 immunotherapy (Figure 8E).

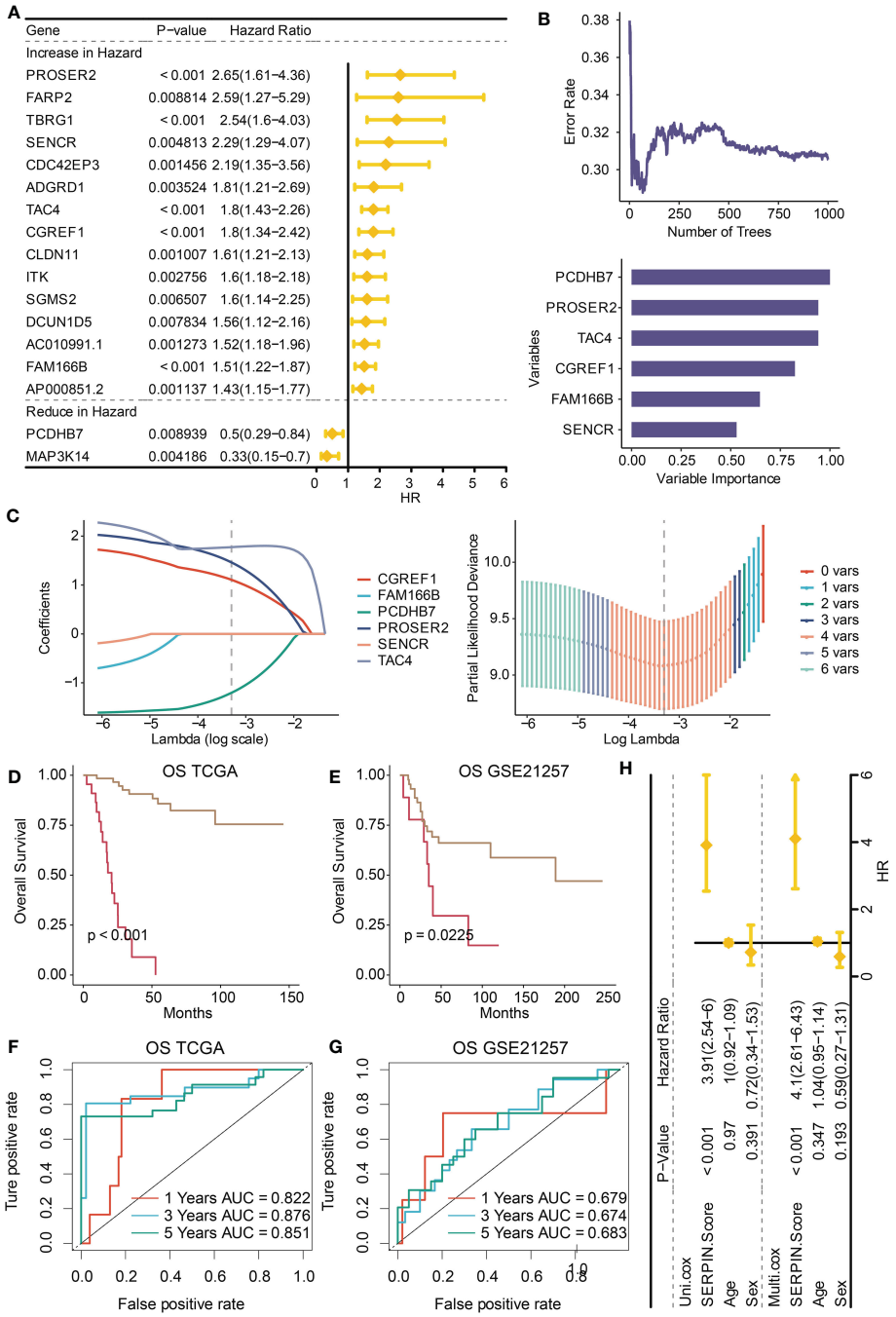


FIGURE 6
Construction of the SERPINH1-related score. **(A)** Univariate Cox regression analysis on the prognostic DEGs. **(B)** Machine learning Random Survival Forest for dimension reduction. **(C)** Machine learning LASSO for dimension reduction and constructing the SERPINH1-related score. **(D)** Survival curves of the two SERPINH1-related score-stratified groups in the TCGA dataset. **(E)** Survival curves of the two SERPINH1-related score-stratified groups in the GSE21257 dataset. **(F)** The 1-year, 3-year, and 5-year ROC regarding the SERPINH1-related score in the TCGA dataset. **(G)** The 1-year, 3-year, and 5-year ROC regarding the SERPINH1-related score in the GSE21257 dataset. **(H)** Univariate and multivariate Cox regression analysis on the clinical factors.

Biological functions of the SERPINH1-related score

SERPINH1 was a hazardous marker in most of the cancer types (Figure 9A). Besides, SERPINH1 positively correlated with the SERPINH1-related score (Figure 9B). SERPINH1 positively correlated with ferroptosis, pyroptosis, apoptosis, and necroptosis (Figure 9C). Besides, the SERPINH1-related score positively correlated with ferroptosis, pyroptosis, apoptosis, and necroptosis (Figure 9C).

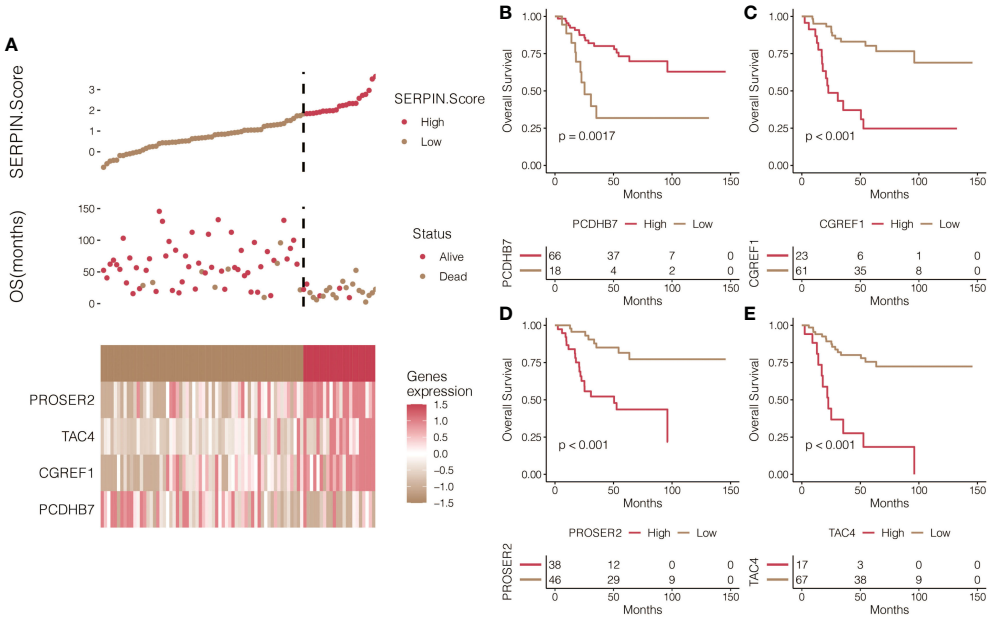


FIGURE 7
Prognostic value of the SERPINH1-related score. **(A)** The expression patterns of PCDHB7, TAC4, PROSER2, and CGREF1 in two SERPINH1-related score-stratified groups. **(B)** Survival curves of the two PCDHB7-stratified, **(C)** TAC4-stratified, **(D)** PROSER2-stratified, and **(E)** CGREF1-stratified groups.

Drug prediction of the SERPINH1-related score

Patients with osteosarcoma who had high SERPINH1-related scores had significantly lower drug sensitivity to AMG-319 2045, AZD3759 1915, AZD8186 1918, Cisplatin 1005, CZC24832 1615, Dactinomycin 1811, Dactolisib 1057, Entospletinib 1630, Foretinib 2040, GSK343 (Figure 10).

Discussion

The most prevalent malignant bone tumor with a significant capacity for invasion and metastasis is osteosarcoma. Patients with osteosarcoma are currently treated mostly with surgery, radiation, chemotherapy, and neoadjuvant chemotherapy. But overall survival rates for osteosarcoma patients are still woefully inadequate. Additionally, some osteosarcoma patients are resistant to conventional chemotherapy medicines. Clinically speaking, treating people with osteosarcoma is quite difficult. Some modern treatments, including as antiangiogenic drugs, immunotherapies, and tumor apoptotic promoters, have been used to treat osteosarcoma, however it is still unknown how well they work. In order to screen the essential molecules or biomarkers for early diagnosis, targeted therapy, and prognosis analysis of osteosarcoma, it is crucial to have a complete understanding of the molecular pathological pathways relating to the occurrence and development of osteosarcoma.

Serpins are homologous proteins with different functions, including tumor development, blood coagulation, fibrinolysis, programmed cell death, and inflammation (22). The relationship between the serpin superfamily and the etiology of various malignancies has been the subject of a plethora of research. The specific roles of the serpin superfamily in osteosarcoma haven't been explored in any studies, nevertheless. SERPINA7, SERPINA3, SERPINB7, SERPINH1, SERPINC1, SERPIND1, SERPINF1, SERPINB8, AGT, SERPING1, SERPINI1, SERPINB6, SERPINE1, SERPINA1, and SERPINF2 are the primary family members of the serpin superfamily. In this study, SERPINH1 outperformed other members of the serpin superfamily in terms of prognostic value. Specifically, SERPINH1 could regulate EMT and metastasis of gastric cancer *via* the Wnt/β-catenin signaling pathway (23). Splicing factor-derived circular RNA circCAMSAP1 was reported to accelerate the tumorigenic process of nasopharyngeal carcinoma *via* a SERPINH1/c-Myc positive feedback loop (24). SERPINH1 was found to be a potential prognostic biomarker based on a pan-cancer analysis (25). Circular RNA circ-TNPO3 could inhibit metastasis of clear cell renal cell carcinoma by binding to IGF2BP2 and destabilizing SERPINH1 (26). CyPA could interact with SERPINH1 to promote extracellular matrix production and inhibit the EMT of trophoblast (27). It has been established that SERPINH1 is crucial in controlling how biological processes in osteosarcoma are carried out. A potential target in the study of immunosuppressive TME and immunotherapy may also be SERPINH1. In human and mouse immunotherapy datasets, SERPINH1 displayed impressive performance in predicting cytokine treatment and immunotherapy.

Given the potent prognostic and predictive value of SERPINH1, we constructed the SERPINH1-related score based on the SERPINH1-related prognostic genes using two machine learning algorithms. PCDHB7, TAC4, PROSER2, and CGREF1 were finally included in the score. PCDHB7 was a potential marker in colorectal

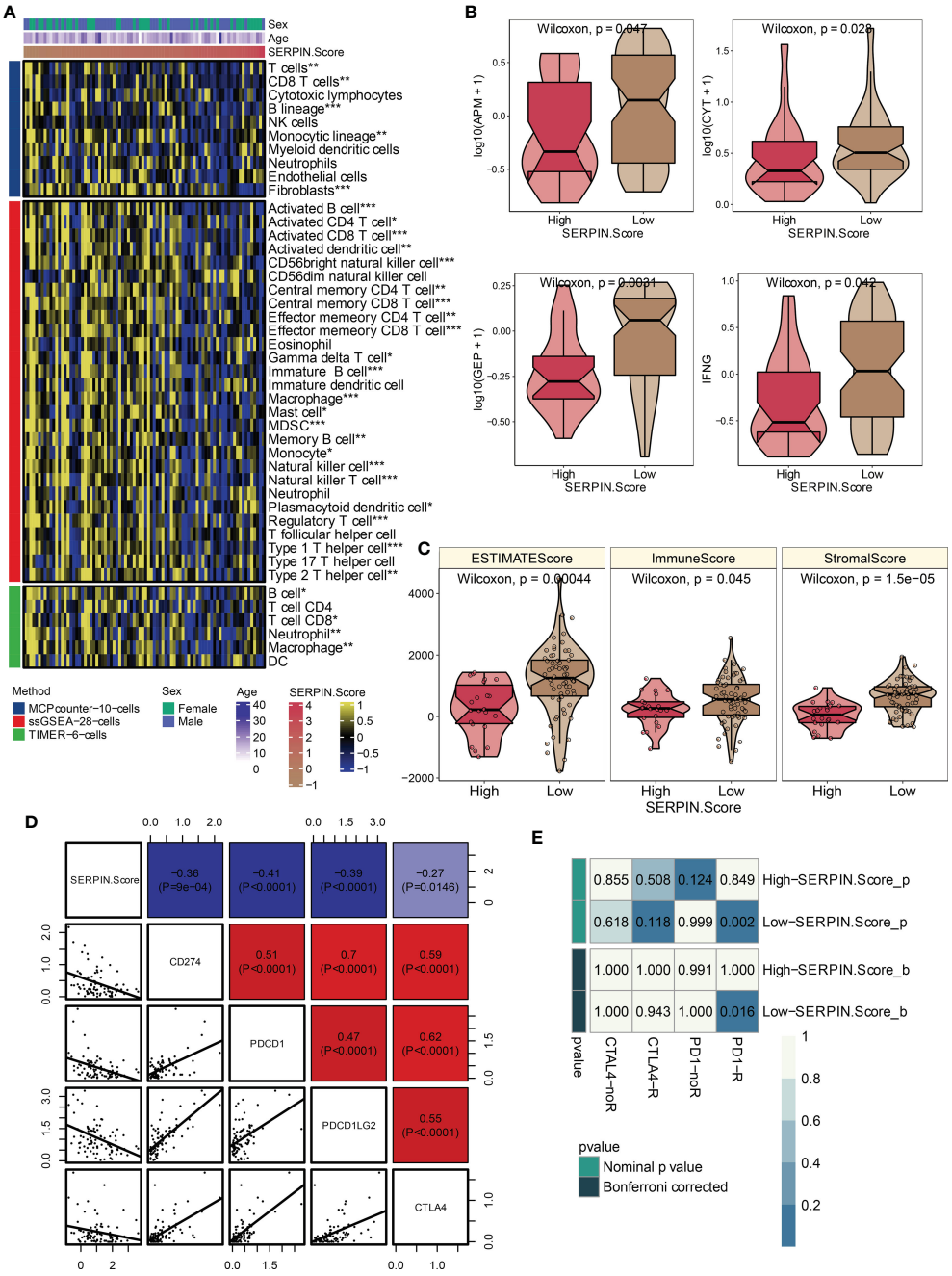


FIGURE 8
Immune characteristics of the SERPINH1-related score. **(A)** The correlation between the SERPINH1-related score and immune infiltrating cells. **(B)** APM, CYT, GEP, and IFN-γ in the two SERPINH1-related score-stratified groups. **(C)** ESTIMATE, Immune, and Stromal scores in the two SERPINH1-related score-stratified groups. **(D)** The correlation between the SERPINH1-related score and immune checkpoints. **(E)** TIDE-based immunotherapy prediction of the SERPINH1-related score. *P<0.05; **P<0.01; ***P<0.001.

cancer (28). TAC4 was a potential marker in osteosarcoma (29). PROSER2 was a potential marker in melanoma (30). CGREF1 was a potential marker in prostate cancer (31). As expected, the SERPINH1-related score efficiently stratified the survival outcomes of osteosarcoma patients and served as an independent prognostic factor. Notably, SERPINH1 and the SERPINH1-related score predict ferroptosis/pyroptosis/apoptosis/necroptosis in osteosarcoma. In combating medication resistance, numerous preclinical and clinical trials have been conducted. Interesting correlations have been found between ferroptosis and cancer therapeutic resistance, and it has been shown that activating ferroptosis can overcome medication resistance (32). Recently, some studies found that pyroptosis can influence tumors' proliferation, invasion, and metastasis, which is regulated by some non-coding RNAs and other molecules (33). Apoptosis is a coordinated and organized cellular process that takes place under both healthy and unhealthy circumstances. Moreover, it is one of the subjects that cell biologists study the most. One condition where

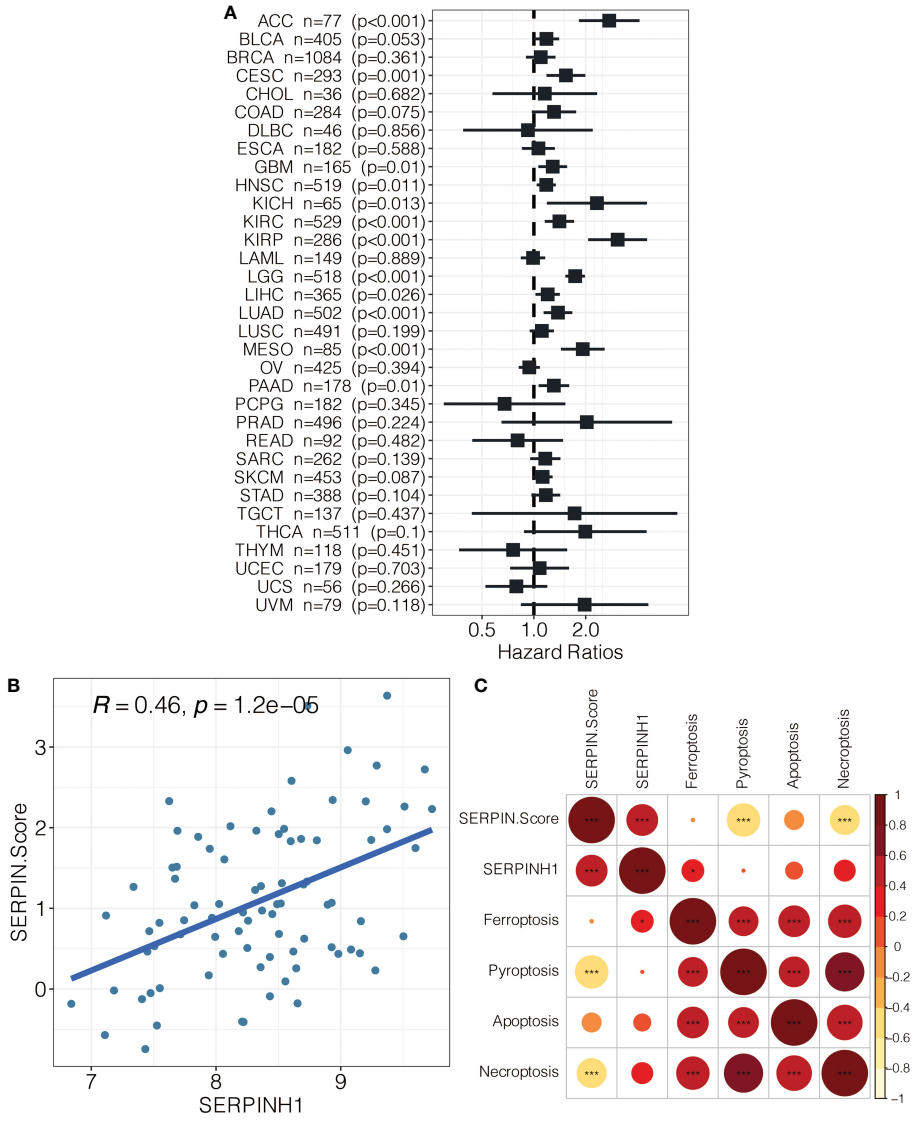


FIGURE 9
(A) Pan-cancer analysis on SERPINH1. **(B)** The correlation between SERPINH1 and SERPINH1-related score. **(C)** The correlation between SERPINH1, SERPINH1-related score, ferroptosis, pyroptosis, apoptosis, and necroptosis.

insufficient apoptosis occurs is cancer, which results in malignant cells that resist death. Apoptosis has a complicated mechanism that encompasses numerous routes. Anywhere along these pathways, flaws might develop, resulting in the malignant transformation of the afflicted cells, tumor spread, and medication resistance. Despite being the root of the issue, apoptosis is a crucial component of many therapeutic regimens for cancer and plays a significant part in its management. The wealth of literature implies that it is possible to target apoptosis in cancer (34). Receptor-Interacting Protein 1 (RIP1), RIP3, and Mixed Lineage Kinase Domain-Like are the major mediators of necroptosis, a controlled necrotic cell death mode that is caspase-independent (MLKL). Necroptosis is a type of programmed cell death that can be used to treat cancer patients resistant to apoptosis. It can also activate and intensify antitumor immunity (35).

In recent years, researchers have learned more and more about the tumor microenvironment (36, 37). The tumor microenvironment consists of cellular components, including immune cells, endothelial cells, and fibroblasts, and non-cellular components, including extracellular matrix, cytokines, and hormones. Immune infiltrating cells play an essential role in influencing tumor progression and therapeutic response (38). In this study, osteosarcoma patients with high SERPINH1-related scores had lower immune infiltrating cells, including as T cells, B cells, NK cells, macrophages, mast cells, Type 1 T helper cells, and Type 2 T helper cells. In osteosarcoma patients with high SERPINH1-related scores, an immune-cold microenvironment that promotes tumor development may exist.

Immunotherapy has emerged as a promising treatment option for several cancers. Cancer immunotherapy tries to provide people

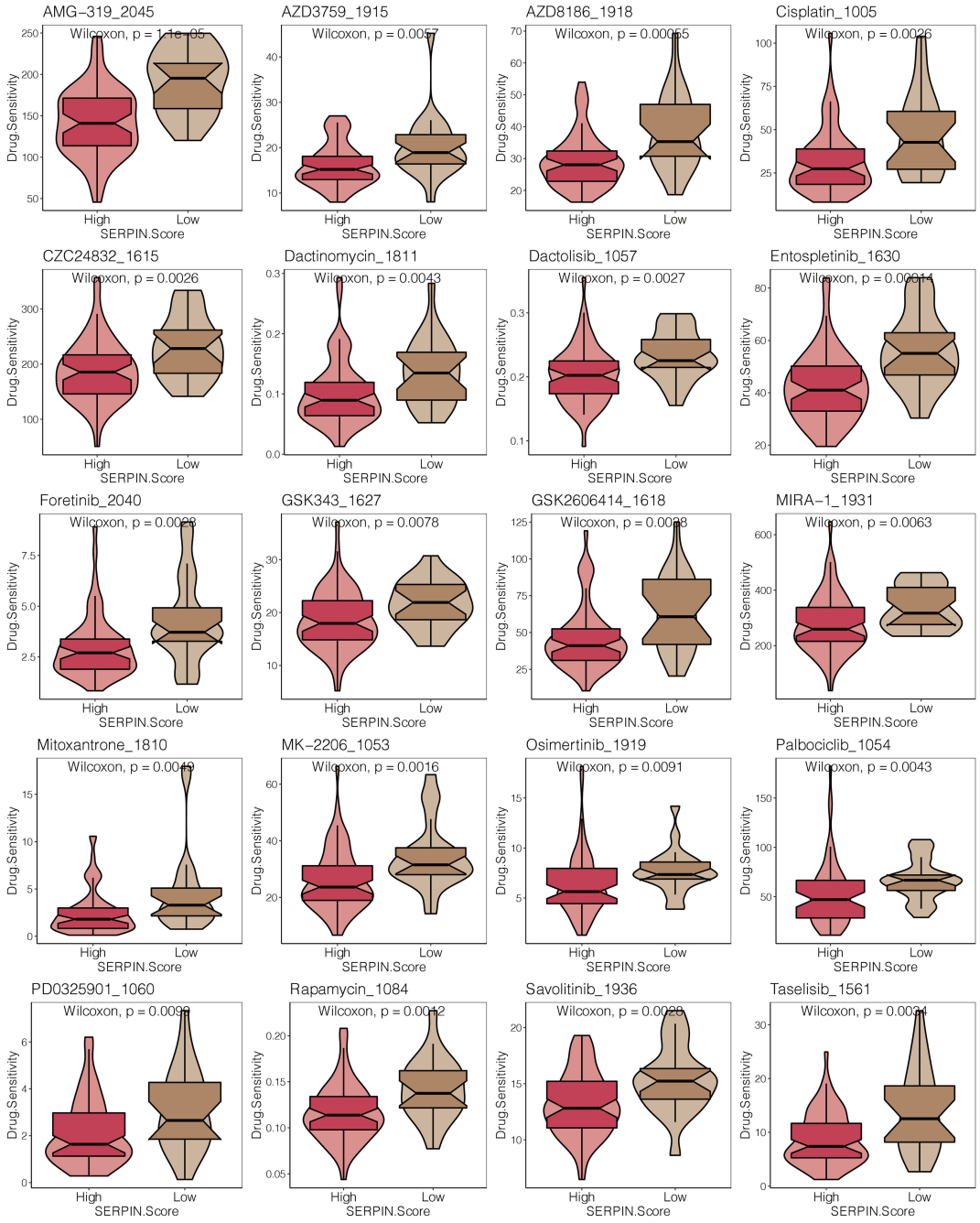


FIGURE 10
Drug prediction of the SERPINH1-related score. The drug sensitivity of 23 chemotherapy drugs in the two SERPINH1-related score-stratified groups.

immunity to fight cancer. In recent years, a large number of novel cancer-specific immunotherapeutic medicines have been approved, highlighting the effectiveness and promise of immunotherapy as an anticancer strategy (39). Among the multiple immunotherapy approaches, immune checkpoint inhibitor based on T cells has received the most attention due to their outstanding performance (40). APM, CYT, GEP, and IFN- levels were all statistically lower in the group with a high SERPINH1-related score, which were all important factors in immunotherapy. Additionally, the SERPINH1-related score was strongly adversely correlated with the CD274, PD-1, PDCD1LG2, and CTLA-4 immunological checkpoints. The likelihood of patients with osteosarcoma responding to anti-PD-1 immunotherapy was significantly higher in those with low SERPINH1-related scores. The SERPINH1-related score was accurate at identifying osteosarcoma patients who could respond to immunotherapy. Additionally, 23 chemotherapeutic agents had significantly decreased drug sensitivity in osteosarcoma patients with high SERPINH1-related scores. AZD3759 could inhibit the proliferation and progression of osteosarcoma through the blockade of the EGFR and JAK pathways (41). Inhibiting PI3Kβ with AZD8186 could regulate key metabolic pathways in PTEN-null tumors (42). Cisplatin is a well-known chemotherapeutic drug used

to treat numerous human cancers, including bladder, head and neck, lung, ovarian, and testicular cancers (43). Therefore, the SERPINH1-related score was accurate in identifying osteosarcoma patients who could respond to chemotherapy.

Conclusion

In conclusion, the biological role of SERPINH1 in osteosarcoma was investigated using *in vitro* confirmation. SERPINH1 and the SERPINH1-related score predict ferroptosis/pyroptosis/apoptosis/necroptosis in osteosarcoma. Further research is still needed to determine the precise processes underlying the osteosarcoma pathogenesis mediated by SERPINH1. The SERPINH1-related score was an effective method for identifying osteosarcoma patients who would respond to immunotherapy and chemotherapy, as well as for predicting the survival outcomes of such patients. A real-world cohort must be used to further validate the SERPINH1-related score's universality.

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 human participants were reviewed and approved by institutional review board (IRB) of the Third Xiangya Hospital, Central South University. The patients/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

GX conceived and performed most of the experiments. GX wrote the manuscript. SW and KL collected and analyzed the data. SW acquisition of the financial support for the project. XC provided experimental advice and supervised the study. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (82,072,501).

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.2023.1180191/full#supplementary-material

References

1. Kansara M, Teng MW, Smyth MJ, Thomas DM. Translational biology of osteosarcoma. *Nat Rev Cancer* (2014) 14(11):722–35. doi: 10.1038/nrc3838

2. Gill J, Gorlick R. Advancing therapy for osteosarcoma. *Nat Rev Clin Oncol* (2021) 18(10):609–24. doi: 10.1038/s41571-021-00519-8

3. Ritter J, Bielack SS. Osteosarcoma. *Ann Oncol* (2010) 21 Suppl 7:vii320–325. doi: 10.1093/annonc/mdq276

4. Heit C, Jackson BC, McAndrews M, Wright MW, Thompson DC, Silverman GA, et al. Update of the human and mouse SERPIN gene superfamily. *Hum Genomics* (2013) 7:22. doi: 10.1186/1479-7364-7-22

5. Valiente M, Obenauf AC, Jin X, Chen Q, Zhang XH, Lee DJ, et al. Serpins promote cancer cell survival and vascular co-option in brain metastasis. *Cell* (2014) 156 (5):1002–16. doi: 10.1016/j.cell.2014.01.040

6. Baek JY, Yeo HY, Chang HJ, Kim KH, Kim SY, Park JW, et al. Serpin B5 is a CEA-interacting biomarker for colorectal cancer. *Int J Cancer* (2014) 134(7):1595–604. doi: 10.1002/ijc.28494

7. Tseng MY, Liu SY, Chen HR, Wu YJ, Chiu CC, Chan PT, et al. Serine protease inhibitor (SERPIN) B1 promotes oral cancer cell motility and is over-expressed in invasive oral squamous cell carcinoma. *Oral Oncol* (2009) 45(9):771–6. doi: 10.1016/j.oraloncology.2008.11.013

8. Mao M, Wang W. SerpinE2 promotes multiple cell proliferation and drug resistance in osteosarcoma. *Mol Med Rep* (2016) 14(1):881–7. doi: 10.3892/mmr.2016.5316

9. Zhong H, Wang Z, Wei X, Liu Y, Huang X, Mo X, et al. Prognostic and immunological role of SERPINH1 in pan-cancer. *Front Genet* (2022) 13:900495. doi: 10.3389/fgene.2022.900495

10. Zhang H, Yan X, Gu H, Xue Q, Liu X. High SERPINH1 expression predicts poor prognosis in lung adenocarcinoma. *J Thorac Dis* (2022) 14(12):4785–802. doi: 10.21037/jtd-22-1518

11. 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(10):1550–8. doi: 10.1038/s41591-018-0136-1

12. Zeng Z, Wong CJ, Yang L, Ouardaoui N, Li D, Zhang W, et al. TISMO: syngeneic mouse tumor database to model tumor immunity and immunotherapy response. *Nucleic Acids Res* (2022) 50(D1):D1391–7. doi: 10.1093/nar/gkab804

13. Zhang L, Wu S, Huang J, Shi Y, Yin Y, Cao X. A mitochondria-related signature for predicting immune microenvironment and therapeutic response in osteosarcoma. *Front Oncol* (2022) 12:1085065. doi: 10.3389/fonc.2022.1085065

14. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* (2016) 17(1):218. doi: 10.1186/s13059-016-1070-5

15. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep* (2017) 18 (1):248–62. doi: 10.1016/j.celrep.2016.12.019

16. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res* (2017) 77(21): e108–10. doi: 10.1158/0008-5472.CAN-17-0307

17. 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

18. Senbabaoglu Y, Gejman RS, Winer AG, Liu M, Van Allen EM, de Velasco G, et al. Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol* (2016) 17(1):231. doi: 10.1186/s13059-016-1092-z

19. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest* (2017) 127(8):2930–40. doi: 10.1172/JCI91190

20. Roh W, Chen PL, Reuben A, Spencer CN, Prieto PA, Miller JP, et al. Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci Transl Med* (2017) 9(379):eaah3560. doi: 10.1126/scitranslmed.aah3560

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. *Brief Bioinform* (2021) 22(6):bbab260. doi: 10.1093/bib/bbab260

22. van Gent D, Sharp P, Morgan K, Kalsheker N. Serpins: structure, function and molecular evolution. *Int J Biochem Cell Biol* (2003) 35(11):1536–47. doi: 10.1016/s1357-2725 (03)00134-1

23. Tian S, Peng P, Li J, Deng H, Zhan N, Zeng Z, et al. SERPINH1 regulates EMT and gastric cancer metastasis *via* the wnt/beta-catenin signaling pathway. *Aging (Albany NY)* (2020) 12(4):3574–93. doi: 10.18632/aging.102831

24. Wang Y, Yan Q, Mo Y, Liu Y, Wang Y, Zhang S, et al. Splicing factor derived circular RNA circCAMSAP1 accelerates nasopharyngeal carcinoma tumorigenesis *via* a SERPINH1/c-myc positive feedback loop. *Mol Cancer* (2022) 21(1):62. doi: 10.1186/s12943-022-01502-2

25. Wang Y, Gu W, Wen W, Zhang X. SERPINH1 is a potential prognostic biomarker and correlated with immune infiltration: A pan-cancer analysis. *Front Genet* (2021) 12:756094. doi: 10.3389/fgene.2021.756094

26. Pan X, Huang B, Ma Q, Ren J, Liu Y, Wang C, et al. Circular RNA circ-TNPO3 inhibits clear cell renal cell carcinoma metastasis by binding to IGF2BP2 and destabilizing SERPINH1 mRNA. *Clin Transl Med* (2022) 12(7):e994. doi: 10.1002/ctm2.994

27. Hu H, Ma J, Li Z, Ding Z, Chen W, Peng Y, et al. CyPA interacts with SERPINH1 to promote extracellular matrix production and inhibit epithelial-mesenchymal transition of trophoblast *via* enhancing TGF-beta/Smad3 pathway in preeclampsia. *Mol Cell Endocrinol* (2022) 548:111614. doi: 10.1016/j.mce.2022.111614

28. Wong CC, Xu J, Bian X, Wu JL, Kang W, Qian Y, et al. In colorectal cancer cells with mutant KRAS, SLC25A22-mediated glutaminolysis reduces DNA demethylation to increase WNT signaling, stemness, and drug resistance. *Gastroenterology* (2020) 159 (6):2163–2180 e2166. doi: 10.1053/j.gastro.2020.08.016

29. Berger A, Paige CJ. Hemokinin-1 has substance p-like function in U-251 MG astrocytoma cells: A pharmacological and functional study. *J Neuroimmunol* (2005) 164(1-2):48–56. doi: 10.1016/j.jneuroim.2005.03.016

30. Yepes S, Tucker MA, Koka H, Xiao Y, Jones K, Vogt A, et al. Using whole-exome sequencing and protein interaction networks to prioritize candidate genes for germline cutaneous melanoma susceptibility. *Sci Rep* (2020) 10(1):17198. doi: 10.1038/s41598-020-74293-5

31. Diaz de la Guardia-Bolivar E, Barrios-Rodriguez R, Zwir I, Jimenez-Moleon JJ, Del Val C. Identification of novel prostate cancer genes in patients stratified by Gleason classification: Role of antitumoral genes. *Int J Cancer* (2022) 151(2):255–64. doi: 10.1002/ijc.33988

32. 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(1):47. doi: 10.1186/s12943-022-01530-y

33. Fang Y, Tian S, Pan Y, Li W, Wang Q, Tang Y, et al. Pyroptosis: A new frontier in cancer. *BioMed Pharmacother* (2020) 121:109595. doi: 10.1016/j.biopha.2019.109595

34. Wong RS. Apoptosis in cancer: from pathogenesis to treatment. *J Exp Clin Cancer Res* (2011) 30(1):87. doi: 10.1186/1756-9966-30-87

35. Gong Y, Fan Z, Luo G, Yang C, Huang Q, Fan K, et al. The role of necroptosis in cancer biology and therapy. *Mol Cancer* (2019) 18(1):100. doi: 10.1186/s12943-019-1029-8

36. Zhang H, Zhang N, Wu W, Zhou R, Li S, Wang Z, et al. Machine learning-based tumor-infiltrating immune cell-associated lncRNAs for predicting prognosis and immunotherapy response in patients with glioblastoma. *Brief Bioinform* (2022) 23 (6):bbac386. doi: 10.1093/bib/bbac386

37. Zhang N, Zhang H, Wu W, Zhou R, Li S, Wang Z, et al. Machine learning-based identification of tumor-infiltrating immune cell-associated lncRNAs for improving outcomes and immunotherapy responses in patients with low-grade glioma. *Theranostics* (2022) 12(13):5931–48. doi: 10.7150/thno.74281

38. Wu T, Dai Y. Tumor microenvironment and therapeutic response. *Cancer Lett* (2017) 387:61–8. doi: 10.1016/j.canlet.2016.01.043

39. Szeto GL, Finley SD. Integrative approaches to cancer immunotherapy. *Trends Cancer* (2019) 5(7):400–10. doi: 10.1016/j.trecan.2019.05.010

40. Zhang H, Wang Y, Zhao Y, Liu T, Wang Z, Zhang N, et al. PTX3 mediates the infiltration, migration, and inflammation-resolving-polarization of macrophages in glioblastoma. *CNS Neurosci Ther* (2022) 28(11):1748–66. doi: 10.1111/cns.13913

41. Yin W, Zhang K, Deng Q, Yu Q, Mao Y, Zhao R, et al. AZD3759 inhibits glioma through the blockade of the epidermal growth factor receptor and janus kinase pathways. *Bioengineered* (2021) 12(1):8679–89. doi: 10.1080/21655979.2021.1991160

42. Lynch JT, Polanska UM, Delpuech O, Hancox U, Trinidad AG, Michopoulos F, et al. Inhibiting PI3Kbeta with AZD8186 regulates key metabolic pathways in PTEN-null tumors. *Clin Cancer Res* (2017) 23(24):7584–95. doi: 10.1158/1078-0432.CCR-17-0676

43. Dasari S, Tchounwou PB. Cisplatin in cancer therapy: Molecular mechanisms of action. *Eur J Pharmacol* (2014) 740:364–78. doi: 10.1016/j.ejphar.2014.07.025

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 05 April 2023
DOI 10.3389/fonc.2023.1169395

OPEN ACCESS

EDITED BY
Nan Zhang,
Harbin Medical University, China

REVIEWED BY
Aimin Jiang,
Second Military Medical University, China
Jinxiao Li,
Huazhong University of Science and Technology, China

*CORRESPONDENCE
Rongshan Li
rongshanli13@163.com

SPECIALTY SECTION
This article was submitted to Cancer Immunity and Immunotherapy, a section of the journal Frontiers in Oncology

RECEIVED 19 February 2023
ACCEPTED 17 March 2023
PUBLISHED 05 April 2023

CITATION
Wang B, Li M and Li R (2023) Identification and verification of prognostic cancer subtype based on multi-omics analysis for kidney renal papillary cell carcinoma. *Front. Oncol.* 13:1169395. doi: 10.3389/fonc.2023.1169395

Identification and verification of prognostic cancer subtype based on multi-omics analysis for kidney renal papillary cell carcinoma

Baodong Wang[1], Mei Li[2] and Rongshan Li[1]*

[1]Department of Nephrology, Fifth Hospital of Shanxi Medical University (Shanxi Provincial People's Hospital), Taiyuan, China, [2]Department of Laboratory Medicine, Shanxi Provincial Hospital of Integrated Traditional Chinese and Western Medicine, Taiyuan, China

Background: Identifying Kidney Renal Papillary Cell Carcinoma (KIRP) patients with high-risk, guiding individualized diagnosis and treatment of patients, and identifying effective prognostic targets are urgent problems to be solved in current research on KIRP.

Methods: In this study, data of multi omics for patients with KIRP were collected from TCGA database, including mRNAs, lncRNAs, miRNAs, data of methylation, and data of gene mutations. Data of multi-omics related to prognosis of patients with KIRP were selected for each omics level. Further, multi omics data related to prognosis were integrated into cluster analysis based on ten clustering algorithms using MOVICS package. The multi omics-based cancer subtype (MOCS) were compared on biological characteristics, immune microenvironmental cell abundance, immune checkpoint, genomic mutation, drug sensitivity using R packages, including GSVA, clusterProfiler, TIMER, CIBERSORT, CIBERSORT-ABS, quanTIseq, MCPcounter, xCell, EPIC, GISTIC, and pRRophetic algorithms.

Results: The top ten OS-related factors for KIRP patients were annotated. Patients with KIRP were divided into MOCS1, MOCS2, and MOCS3. Patients in the MOCS3 subtype were observed with shorter overall survival time than patients in the MOCS1 and MOCS2 subtypes. MOCS1 was negatively correlated with immune-related pathways, and we found global dysfunction of cancer-related pathways among the three MOCS subtypes. We evaluated the activity profiles of regulons among the three MOCSs. Most of the metabolism-related pathways were activated in MOCS2. Several immune microenvironmental cells were highly infiltrated in specific MOCS subtype. MOCS3 showed a significantly lower tumor mutation burden. The CNV occurrence frequency was higher in MOCS1. As for treatment, we found that these MOCSs were sensitive to different drugs and treatments. We also analyzed single-cell data for KIRP.

Conclusion: Based on a variety of algorithms, this study determined the risk classifier based on multi-omics data, which could guide the risk stratification and medication selection of patients with KIRP.

KEYWORDS

kidney renal papillary cell carcinoma, prognosis, immune microenvironment, drug response, multi-omics

1 Introduction

Renal cell carcinoma (RCC) is a common malignancy of urinary system (1). Clear cell carcinoma (ccRCC) is the most common pathological type of RCC, accounting for about 70% of RCC, followed by Kidney Renal Papillary Cell Carcinoma (KIRP) and chromophobe renal cell carcinoma (1). Although surgical resection is a good treatment for renal cell carcinoma at early stage, 30% of patients are diagnosed in advanced stage due to delayed diagnosis, and 10% to 20% of patients develop metastatic kidney cancer due to postoperative recurrence, which eventually leads to death (2–4). At present, due to the lack of understanding of the pathogenesis of renal cancer, there is a lack of effective treatment for metastatic renal cancer. Therefore, to explore the pathogenesis of kidney cancer is of great significance for the treatment of kidney cancer. The search for biomarkers related to kidney cancer can help clinicians personalize patient treatment strategies and increase patient benefits.

Kidney cancer is a heterogeneous disease with multiple subtypes, multiple genes, different biochemical characteristics and multiple forms (5). KIRP, the second most common type of RCC, is a heterogeneous disease originating in the tubular epithelium of the kidney (6). The histological features of KIRP are the papillary arrangement of tumor cells, and the axis of the papilla is fibrous vascular tissue (7). In 1997, Delahunt and Eble divided KIRP into type 1 and type 2 according to histopathological characteristics and prognostic differences (8). In many studies, histological subtypes have been shown to be important prognostic predictors, with type 1 KIRP having a better prognosis than type 2 KIRP (9). Previous studies have reported that type 2 KIRP have higher nuclear grading, later staging, and poorer prognosis than type 1 KIRP (10). In contrast, Bigot et al. showed in a study of 486 patients with KIRP who underwent nephron-sparing surgery that the histological subtype of KIRP had no effect on postoperative tumor outcome (11). In conclusion, whether the histological subtype involved in type 1 or type 2 can be used as an independent prognostic factor is controversial, and correct histological phenotype and prognostic prediction are essential for the formulation of medical protocols.

Advances in sequencing technology and machine learning of all kinds have led to significant advances in the acquisition and analysis of omics data, which have deepened the understanding of tumors at the molecular level (12). Compared with a single type of data, omics data reflect the characteristics of biological individuals at multiple levels, which provides the possibility to delineate cellular molecular mechanisms in detail. Different levels of omics data reflect different relationships between genomic distribution, cancer occurrence, progression, and prognosis (13). At the same time, each omics data has its own advantages. For example, methylation chip data and lncRNA expression matrix have good tissue conservation, which can be used as efficient markers for the early diagnosis of specific tumor tissues (14). miRNA data are characterized by dissociation and can be used for non-invasive diagnosis and dynamic detection of disease (15). Common transcriptome, or mRNA sequencing, is the cheapest and most readily available, and is suitable for use in a wide range of cohort studies to explore general patterns in patient populations (16–25).

RCCs with different pathological types have different therapeutic methods and prognosis. In addition, existing targeted drugs are mainly used for ccRCCs, with unclear clinical efficacy in non- ccRCCs (26, 27). It is important to note that there is currently a lack of multi-omics prognostic molecular typing based on KIRP to guide the diagnosis and treatment of KIRP. In this study, the risk stratification of KIRP was studied by integrating multiple omics, and the differences of subgroups were analyzed in each single omics data to characterize the key events in the development of KIRP. The study provides a reference for precision medicine of KIRP.

2 Materials and methods

2.1 Extraction and preprocessing of multi-omics data for KIRP

The dataset for KIRP was downloaded from The Cancer Genome Atlas (TCGA) (28) and TCGA database had the multi-omics data for our analysis in this study. We acquired gene expression profile for transcriptomics (including mRNAs encoding protein, long noncoding RNAs as known as lncRNAs, microRNA known as miRNAs, data of methylation, and data of gene mutations). We applied TCGAbiolinks package of R application to acquire clinicopathologic information and multi omics-based data. We downloaded the gene expression profiles of 34 cases with KIRP from GSE2748 as the external validation cohort (29). The patients with KIRP in the GSE2748 dataset had the prognostic information (29). In addition, we searched and downloaded the single-cell RNA sequencing for KIRP from GSE152938 (30). There was a total of four KIRP samples and one normal kidney sample included in GSE152938 (30). The matrix for single-cell RNA sequencing was generated by R package Seurat (31).

2.2 Identification of multi omics-based cancer subtypes by integrative analysis

The MOVICS package aimed to show the multi-omics comprehensive clustering and visualization of cancer typing studies (32). There were ten algorithms included in the MOVICS package: CIMLR, iClusterBayes, MoCluster, COCA, ConsensusClustering, IntNMF, LRAcluster, NEMO, PINSPlus and SNF (32). For the multi omics-based data, we focused on the characteristic related to prognosis (OS). The OS-related features, including mRNAs, lncRNAs, miRNAs, methylation, and gene mutations were analyzed by Univariate Cox regression analysis, and we screened out features with the threshold P-value<0.05. Due to the small amount of mutation matrix and miRNA expression data, only the top 30 mutations and 200 miRNA data were extracted. We carried out analysis for Clustering Prediction Index (CPI) (33) and Gaps-statistics (33) to filtrate out the optimal number of cancer subtypes. We finally identified the multi omics-based cancer subtype (MOCS) based on consensus ensembles and high robustness, thus separating the patients with KIRP into different MOCSs.

2.3 Nearest template prediction validation

Nearest template prediction (NTP) algorithm could also be applied to cross-platform, cross-species and multi-class predictions without any optimization of analysis parameters (34). In this study, we also used NTP algorithm of CMScaller package to test the dependability and stability of MOCS subtypes *via* the external GSE2748 cohort.

2.4 Biological characteristics for MOCS subtypes

The gene sets (including immune-related pathways) were analyzed, and enrichment scores were calculated using gene set variation analysis (GSVA) from R package GSVA (35). The differentially expressed genes (DEGs) among the three MOCS subtypes were assessed using limma package (36). Pathway enrichment analysis was performed by clusterProfiler package with the employment of Biological Processes in Gene Ontology (GO) (37).

2.5 Calculation of immune microenvironmental cell abundance and immune checkpoint

Tumor Immune Estimation Resource (TIMER) is a website from which researchers can use RNA-Seq expression profile data to detect the infiltration of immune cells in tumor tissue (38). The TIMER provides the infiltrations of six kinds of immune cells (B cells, $CD4^+$ T cells, $CD8^+$ T cells, Neutrophil, Macrophages and Myeloid dendritic cells) (38). CIBERSORT (39) and CIBERSORT-ABS (40) algorithms were used to acquire the infiltrations of 22 kinds of immune cells. quanTIseq is a deconvolution tool developed specifically for RNA-seq data, enabling accurate quantification of unknown tumor content, as well as the immune cell component of the overall tissue (41). quanTIseq implemented a complete deconvolution process for analyzing RNA-seq data based on constrained least squares regression and a new eigenmatrix from 51 purified or enriched RNA-seq data sets, avoiding inconsistencies between mixtures and eigenmatrices (41). MCPcounter (42), xCell (43), and EPIC (44) algorithms (Estimate the Proportion of Immune and Cancer cells) were also used to assess the immune microenvironmental cell abundance. We estimated the infiltrating level of immune or stromal scores using ESTIMATE R package (45). In addition, DNA methylation of tumor-infiltrating lymphocyte (MeTIL) for TCGA- KIRP cohort was also calculated (46).

2.6 Evaluation of genomic mutation for MOCS subtypes

Mutation profiles of KIRP were acquired and we compared and visualized the difference of mutation among the MOCS subtypes utilizing Maftools package of R (47). We applied the Maftools function to analyze the oncogenic pathway and mutually exclusive or coexisting mutations (48). The loss and gain in genomic level was evaluated by GISTIC 2.0 algorithm (49).

2.7 Drug sensitivity profiles for MOCS subtypes

R package pRRophetic was employed to predict the drug sensitivity profiles for MOCS subtypes (50, 51). Subclass mapping was used to explore the immunotherapy of KIRP based on the literature published (52, 53).

2.8 Statistical analyses

R was used to conduct statistical analyses (v4.0.2). We also provide the codes of all methods used in this paper in Supplementary Code. P values or adjusted P values less than 0.05 were considered significant for all statistical comparisons.

3 Results

3.1 Three MOCSs were categorized for KIRP patients by MOVICS package

We discovered three MOCS subtypes for KIRP patients based on CPI analysis and Gaps-statistics, due to the optimal average statistic value with the number of MOCSs was found to be three

(Figure 1A). Hence, patients with KIRP were divided into MOCS1, MOCS2, and MOCS3, indicating the robustness of the classification system (Figure 1B). The silhouette plot indicated that the silhouette score of MOCS1 was 0.70, the silhouette score of MOCS2 was 0.42, while the silhouette score of MOCS3 was 0.71, which substantiated that the MOCS subtypes were distinguishable and separated well from each other (Supplementary Figure S1A). From the Figure 1C, the consistency of the classification system for MOCSs were observed in consideration of four statistics (54) (Figure 1C), including Rand Index (RI), Adjusted Mutual Information (AMI), Jaccard Index (JI), and Fowlkes-Mallows (FM) (54). Furthermore, we observed that all the patients in the MOCS1 and MOCS1 were at AJCC Stage I and Pstage I (Figure 1C). Among the three MOCSs, we displayed the distribution of the multi-omics data for mRNA, lncRNA, miRNA, DNA methylation, and gene mutations as shown in the heatmap (Figure 1D). In the distribution diagram (Figure 1D), RBP4, MSLN, VSTM2L, FTCD, AC147651.5, RP11-23P13.6, RP11-326C3.2, RP11-124N19.3, CHL1-AS2, and RP11-807H17.1 were the top ten OS-related factors of transcriptome (mRNAs and lncRNAs). As for miRNA, hsa-mir-127, hsa-mir-1247, hsa-mir-1-1, hsa-mir-1-2, hsa-mir-1180, hsa-mir-1269a, hsa-mir-10b, hsa-mir-126, hsa-mir-105-1, and hsa-mir-105-2 were the top ten OS-related factors of miRNAs (Figure 1D). As for DNA methylation, cg16434331, cg06775420, cg25244238, cg06282596, cg02239902, cg22688012, cg23591302, cg03994717, cg06223834, and cg06234051 were the top ten OS-related factors (Figure 1D). SETD2, PBRM1, SYNE2, NF2, MET, LRP2, CUL3, PKHD1, TTN, and PCF11 were the top ten OS-related factors (Figure 1D). Further, we compared the outcome of clinical prognosis of patients with KIRP among MOCS1, MOCS2, and MOCS3. Patients in the MOCS3 subtype were observed with shorter overall survival time than patients in the MOCS1 and MOCS2 subtypes (Figure 1E), which was also observed for progression free survival time (Figure 1E). Using NTP algorithm, three MOCSs were also identified as predicted by the external GSE2748 cohort (Supplementary Figure S1B). Patients in the MOCS3 subtype were observed with shorter overall survival time (Supplementary Figure S1C).

3.2 Biological characteristics for MOCS subtypes

Further, we depicted the molecular features characterization for MOCS subtypes. We computed the enrichment score of immune-related pathways (including Cell Functions, B Cell Functions, T Cell Functions, Leukocyte Functions, Pathogen Defense, Interleukins, TNF Superfamily, Chemokines, Cytokines, Regulation NK Cell Functions Complement, Antigen Processing, Cytotoxicity, Microglial Functions, TLR, Adhesion, Transporter Functions, Cell Cycle, Macrophage Functions and Senescence) based on GSVA analysis. We could find that MOCS1was negatively correlated with immune-related pathways (Figure 2A). As for other pathways, we found global dysfunction of cancer-related pathways among the three MOCS subtypes (Figure 2B). Generally, MOCS1 showed relatively lower enrichment level of Nature metabolism Hypoxia, Hu hypoxia signature, Exosomal secretion, Ferroptosis, MT exosome and exosome assembly (Figure 2B), suggesting the three MOCS subtypes were association with exosomes strongly. Biological processes of AXONEMAL DYNEIN COMPLEX ASSEMBLY, CILIUM MOVEMENT, AXONEME ASSEMBLY, INNER DYNEIN ARM ASSEMBLY, INTRACILIARY TRANSPORT, MICROTUBULE BUNDLE FORMATION, PROTEIN LOCALIZATION TO CILIUM, MRNA SPLICE SITE SELECTION, INTRACILIARY TRANSPORT INVOLVED IN CILIUM ASSEMBLY, and EXTRACELLULAR TRANSPORT were overactivated in MOCS1 (Figure 2C). Biological processes of RESPIRATORY ELECTRON TRANSPORT CHAIN, RESPIRATORY ELECTRON TRANSPORT CHAIN, ELECTRON TRANSPORT CHAIN, ORGANIC ACID CATABOLIC PROCESS, OXIDATIVE PHOSPHORYLATTON, GOTATP SYNTHESIS COUPLED ELECTRON TRANSPORT, COFACTOR METABOLIC PROCESS, COENZYME METABOLIC PROCESS, AEROBIC RESPIRATION, SMALL MOLECULE CATABOLIC PROCESS, and ALPHA AMINO ACID METABOLIC PROCESS (Figure 2C). Biological processes of CORNIFICATION, NEURON FATE SPECIFICATION, TONGUE DEVELOPMENT, AUTONOMIC NERVOUS SYSTEM DEVELOPMENT, INNERVATION, FORELIMB MORPHOGENESIS, APPENDAGE DEVELOPMENT, ENDOCARDIAL CUSHION MORPHOGENESIS, APPENDAGE MORPHOGENESIS, EYELID DEVELOPMENT IN CAMERA TYPE EYE (Figure 2C). In addition, we evaluated the activity profiles of regulons among the three MOCSs, thus highlighting the additional potential regulatory differences. The higher level of several regulon, such as ZNF683, IRF4, CEBPB, EPAS1, and TFE3 was observed in MOCS2 and MOCS3 (Figure 2D), indicating the important differentiators of epigenetically driven transcriptional networks among the three MOCS subtypes. GSVA analysis was carried out regarding metabolism-related pathways, we found that most of the metabolism-related pathways were activated in MOCS2 (Supplementary Figure S2A). Consistently, most of immune-associated signatures were enriched in MOCS2 (Supplementary Figure S2B).

3.3 Calculation of immune microenvironmental cell abundance and immune checkpoint

In consideration of the critical role of immunity in KIRP progression, we investigated the immune microenvironmental cell abundance and immune checkpoint among the three MOCS subtypes. Several immune microenvironmental cells were highly infiltrated in specific MOCS subtype. For instance, B cell in MOCS3, Macrophage M2 in MOCS2, NK cell in MOCS1 and so on (Figure 3A). As for the immune checkpoint genes, on the whole, MOCS3 was associated with higher levels of immune checkpoint genes (Figure 3B). MOCS3 was also associated with higher levels of MeTIL (Figure 3B). MOCS1 was found to be associated with lower

levels of immune, stromal and ESTIMATE scores (Figure 3C). Additionally, we found that MOCS3 showed a significantly lower tumor mutation burden (TMB, Figure 3D). MOCS1 was found to be associated with lower signature score of CD8$^+$ T effector, Immune checkpoint, APM, TME score A, Pan F TBRs, EMT2, EMT3, and TME score B (Supplementary Figure S3A). The level of RNAss, DMPss, ENHss, EREG.EXPss and HRD was found to be lower in MOCS1 (Supplementary Figure S3B).

3.4 Evaluation of genomic mutation for MOCS subtypes

The differences in copy number variations (CNV) among the three MOCS subtypes were compared, and the result revealed that the CNV occurrence frequency was higher in MOCS1 (Figures 4A–C). In detail, amplification in chr 2p, 2q, 3p, 3q, 4p, 7p, 7q, 12p, 12q, 16p, 16q, 17p, 17q, 18p, 18q, 20p, 20q, and 21q were higher in MOCS1

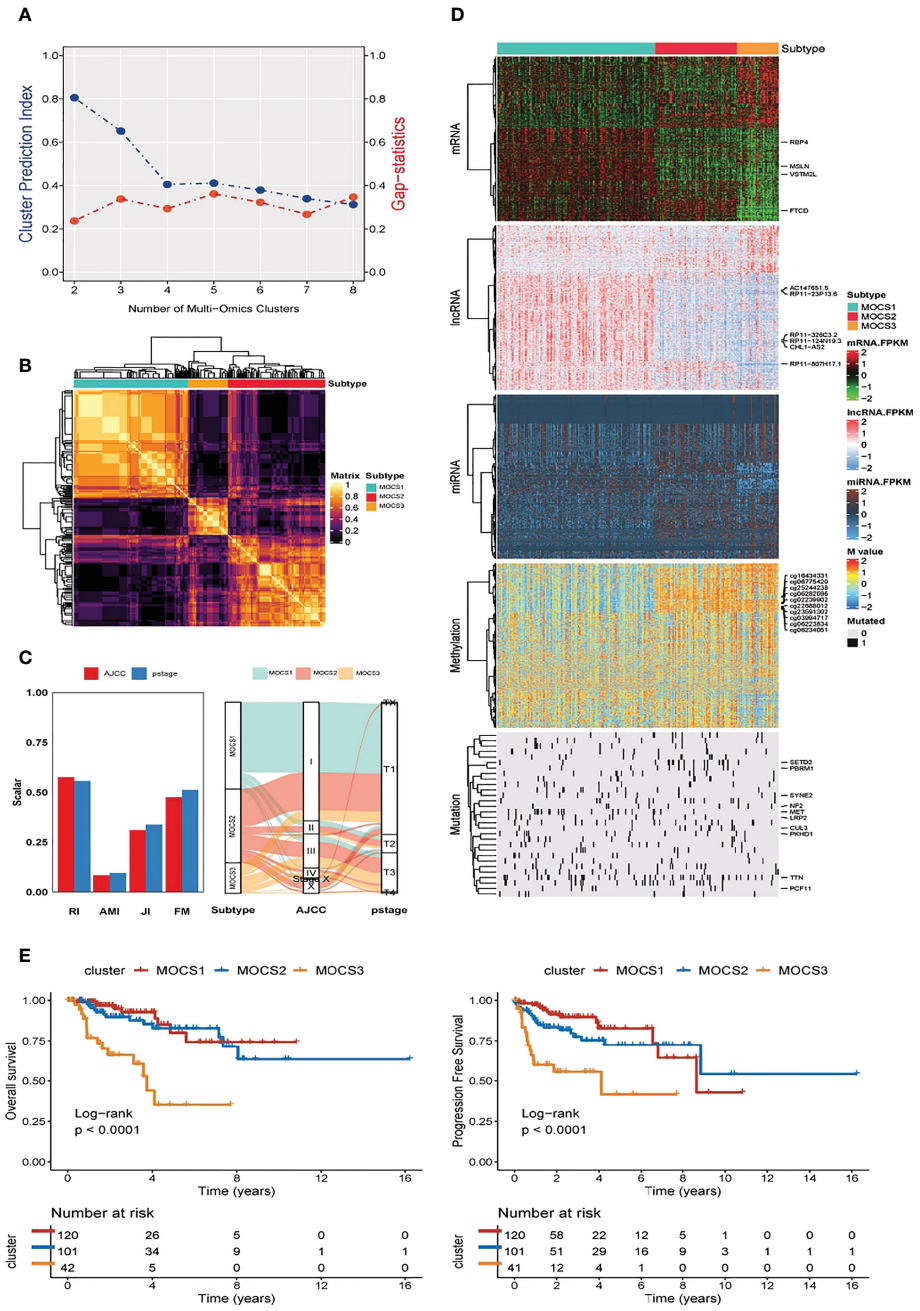


FIGURE 1
Three MOCSs were categorized for KIRP patients by MOVICS package. **(A)** Determination of optimal cluster number through calculating CPI (blue line) and Gaps-statistics (red line) in TCGA- KIRP cohort. **(B)** Consensus heatmap based on outcomes from 10 multi-omics integrative clustering approaches with subtype number of three showing perfect diagonal rectangle. **(C)** Quantification of sample similarity using silhouette score based on the consensus ensembles result and alluvial diagram presenting the flow distribution among different multi omics-based cancer subtypes (MOCSs). **(D)** Comprehensive heatmap showing the detailed molecular landscape multi-omics data for mRNA, lncRNA, miRNA, DNA methylation, and gene mutations among the three MOCSs. **(E)** log-rank test for overall survival time and progression free survival time for patients with KIRP.

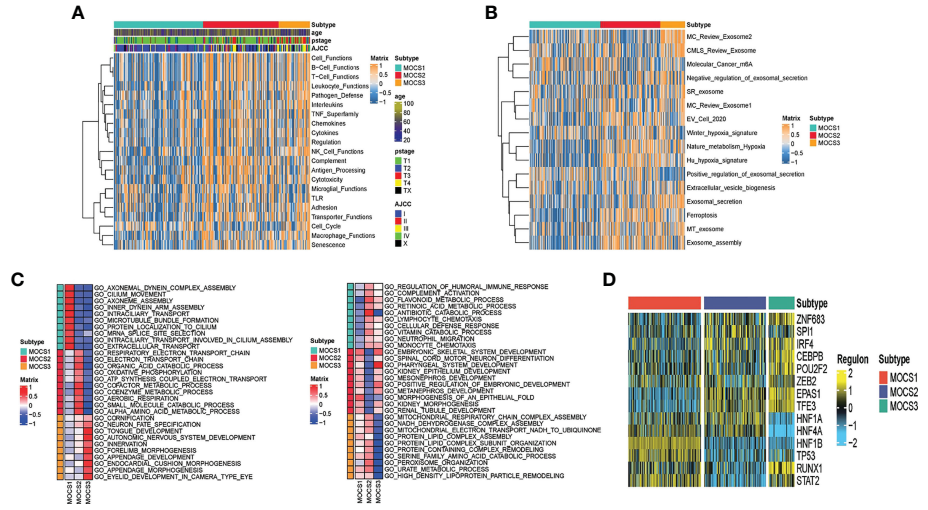


FIGURE 2
Biological characteristics for MOCS subtypes. **(A)** Heatmap showing GSVA enrichment score of immune-related pathways among the three MOCS subtypes. **(B)** Heatmap showing GSVA enrichment score of cancer-related pathways among the three MOCS subtypes. **(C)** GO enrichment analysis showing the upregulated pathways and the downregulated pathways. **(D)** Heatmap showing the regulon distribution among the three MOCS subtypes.

(Figure 4B). The above results were also proved by the total copy number alteration rate as shown in Figure 4C. MOCS1 displayed a higher rate in focal and arm-level mutation level gain (Figure 4D). Mutation patterns of the top 20 most frequently mutated genes among the three MOCSs were displayed in the waterfall plot (Supplementary Figure S4A), from which we could see that, TTN, MET, CUBN, SYNE1, HERC2, KIAA1109, MUC16, PKHD1, WDFY3, DNAH8, KMT2C, LRP2, MACF1, NEB, PCLO, SMARCA4, ANK3, COL18A1, DDX5, DYNC2H1 were the top 20 mutated genes for MOCS1; TTN, SETD2, MUC16, CUL3, KIAA1109, KMT2C, PBRM1, PCF11, BAP1, FAT1, KMT2D, PKHD1, KDM6A, LRBA, SRRM2, ARID1A, ASAP2, BIRC6, CENPE and CNOT1 were the top 20 mutated genes for MOCS2; NF2, TTN, TXNIP, BAP1, CAMK1D, CDH8, CMYA5, CREBBP, EBF2, HECTD4, ITGAL, KRAS, MAP1B, TAS1R2, TG, EIF4G3, FAT1, HELZ2, KDM6A, and SYNE1 were the top 20 mutated genes for MOCS3 (Supplementary Figure S4A). The synthetic lethal mutations in MOCS1, MOCS2, and MOCS3 were displayed in Supplementary Figure S4B. The potential druggable gene categories from the mutation dataset for MOCS1, MOCS2, and MOCS3 were shown in Supplementary Figure S4C, we found that ANK3, CUBN, LRP2, MET, PKHD1 and so on were the potential therapeutic targets for MOCS1; ARID1A, BAP1, CUL3, FAT1, KDM6A and so on were the potential therapeutic targets for MOCS2; BAP1, CAMK1D, CMYA5, CREBBP, FAT1 and so on were the potential therapeutic targets for MOCS3. The fraction of pathways and samples affected were the minimum among the three MOCSs (Supplementary Figure S4C).

3.5 Drug sensitivity profiles for MOCS subtypes

We collected drug response data reflected by the IC50 value *via* GDSC database. We observed that patients in MOCS3 were more sensitive to Crizotinib, Erlotinib, Pazopanib, Saracatinib, Sunitinib, and Temsirolimus (Figure 5A). We found that patients in MOCS1 were more sensitive to AS601245, Bosutinib, PAC.1, ABT.888, and Bleomycin (Figure 5B). Whereafter, we carried out subclass mapping and the results revealed that patients in MOCS2 were more likely to respond to anti-PD1 blockades (Figure 5C).

3.6 Single-cell analysis

A total of 16 cell clusters were identified after gene filtering, normalization and principal component analysis, as shown in Figure 6A. There were nine specific cell types, including B cell, $CD8^+$ T cell, Endothelial cell, Plasma cell, TAM cell, CAF cell, Dendritic cell, Fibroblast cell, pRCC cell (Figure 6B). In addition, a total of three cell clusters (C0, C1, and C2) were predicted by Scissor tool, as shown in Figure 6C. The bar graph displayed the fraction of specific cell types in each cell cluster predicted by Scissor tool (Figure 6D). C0 cluster was rich in TAM cell, CAF cell, Fibroblast cell, CD8+ T cell, Endothelial cell, and pRCC cell (Figure 6D). C1 cluster was rich in Dendirtic cell, Plasma cell, and B cell (Figure 6D). The correlation networks were generated to show the interactions among different cells (Figure 6E). The ligand–receptor pairs among cells were displayed in Figure 6F.

4 Discussion

Global cancer data show that RCC accounts for about 3%~5% of adult malignant tumors, and its incidence is higher in males than in females (4, 55). RCC is the 9th most common male cancer and the 14th most common female cancer worldwide (56). As for etiology, tobacco exposure of any kind is thought to be associated with the development of kidney cancer (57). In addition, diets high in fat, high in protein, low in fruits and vegetables, and increased intake of dairy

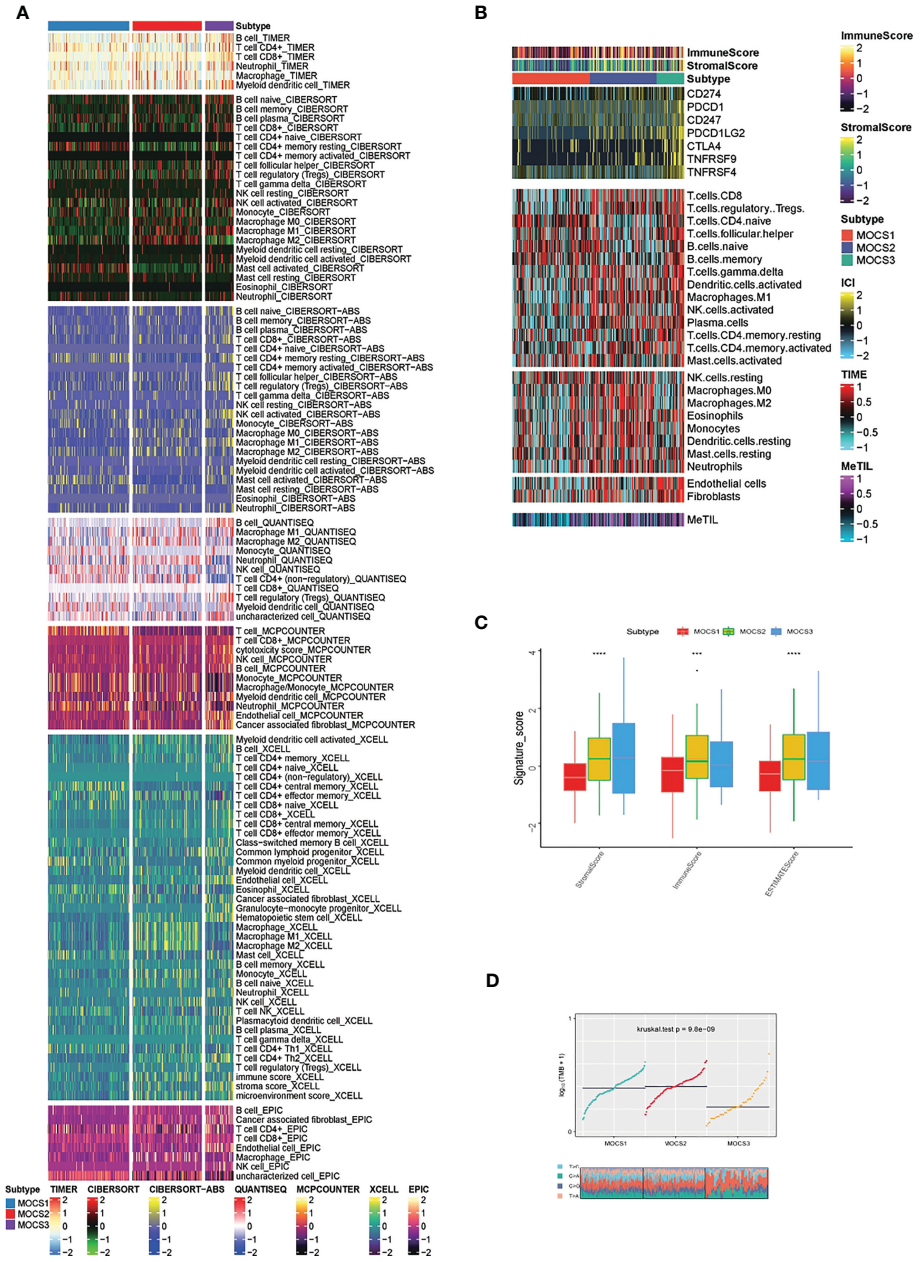


FIGURE 3
Calculation of immune microenvironmental cell abundance and immune checkpoint. **(A)** Heatmap showing the immune microenvironmental cell profile for TCGA- KIRP cohort based on TIMER, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, MCPCOUNTER, XCELL, EPIC algorithms. **(B)** Heatmap showing the profile for immune checkpoint genes, and DNA methylation of tumor-infiltrating lymphocytes (MeTILs). **(C)** Boxplot showing the distribution of immune, stromal and ESTIMATE scores. **(D)** Distribution of TMB and TiTv (transition to transversion) among the three MOCS subtypes. *** means P < 0.001 and **** means P < 0.0001.

products are associated with kidney cancer, but the relative risk is not very high, and many scholars have different views (57, 58). The current diagnosis and treatment problem of renal cancer is the heterogeneity of the tumor, which often leads to different prognosis of patients with the same stage and grade (59). In addition, the incidence of tumor resistance and metastasis is high in renal cancer, and the treatment options for these patients are extremely limited, resulting in a low 5-year survival rate (60). In view of the above problems, it is necessary to determine new diagnosis and treatment strategies to improve the survival rate of patients with kidney cancer.

TCGA is an oncology research initiative of The Cancer Genome Atlas and the National Human Genome Research Institute (28). The plan includes multifactorial data on common tumor tissues and prognostic information for patients. The data included pathological sections, cancer and para-cancer transcriptome, methylation chip data and genome data (28). The development of multi-omics has made it easy for researchers to deepen their understanding of cancer at the molecular level. At the same time, a large number of omics data also brings new challenges to analysts (61). It is particularly critical to reduce data noise and obtain key characteristics of tumor occurrence

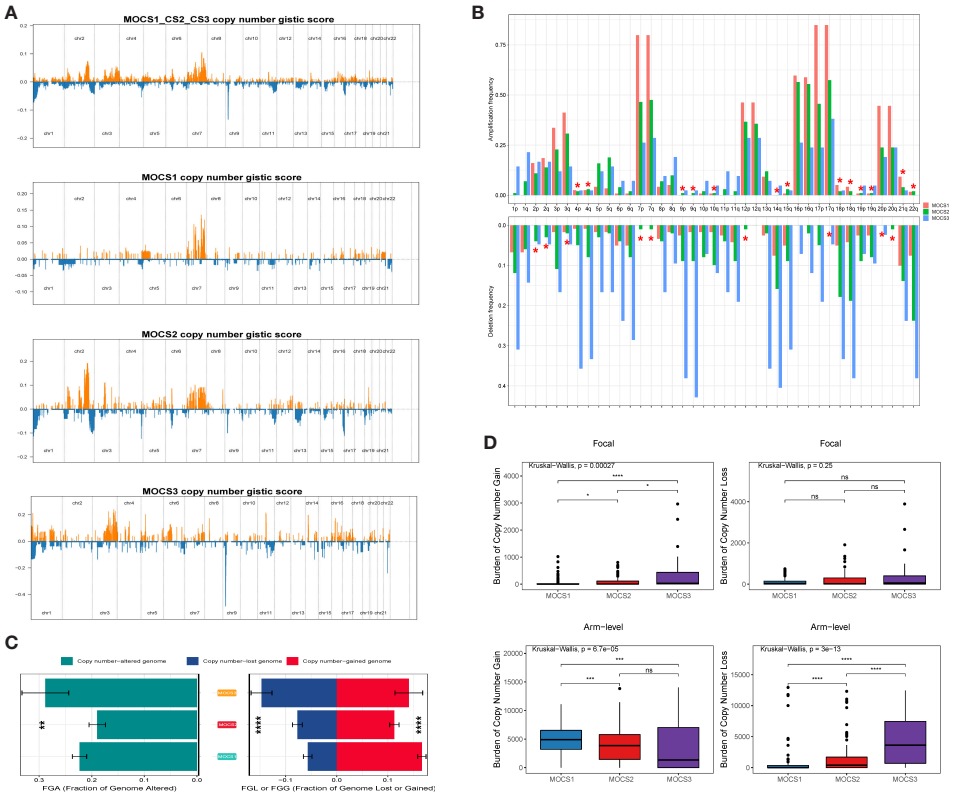


FIGURE 4
Landscapes of copy number variations. **(A)** Comparison of overall copy number among all patients with KIRP, MOCS1, MOCS2, and MOCS3. **(B)** The amplification or deletion frequency in chromosome among the three MOCSs. **(C)** Bar-plot indicating the total alteration frequency among the three MOCSs. **(D)** Different burden of copy number gain at focal and arm-level among the three MOCSs. * means P < 0.05, ** means P < 0.01, *** means P < 0.001, and **** means P < 0.0001, ns, no significance.

and development while preserving tumor characteristics (61). Few studies have attempted to establish a comprehensive model based on multiple omics data to predict prognosis and personalized drug selection in patients with KIRP. Therefore, it is particularly important to develop a comprehensive and robust prognostic and drug selection model for patients with KIRP to assist in prognostic prediction and guide personalized treatment. In this study, we conducted a comprehensive integrated analysis of multiple omics data, including mRNA, lncRNA, miRNA, DNA methylation profile and somatic mutation data, and constructed a classifier to evaluate the prognosis of patients with KIRP and assist drug selection. Omics data are complex, multi-layered, and high weaves, so a key goal of analyzing multi-omics data is to screen for valid predictors to predict phenotypic characteristics and thus elucidate the biological significance behind them. Another major difficulty in omics data processing is dimensionality reduction, omics noise elimination and overfitting avoidance. In this study, R package Survival was first used to screen the molecular features associated with patient prognosis in each omics for subsequent analysis. The classification of cancer patients into different molecular subgroups based on multi-omics data is an important problem in the context of precision medicine. MOVICS provides a unified interface to 10 state-of-the-art multiomics ensemble clustering algorithms and integrates the downstream analyses most commonly used in cancer typing studies, including characterization and comparison of identified subtypes from multiple perspectives and validation of subtypes in external corporations of multi-class predictions using two model-free methods. Patients with KIRP were divided into three multi omics-based cancer subtypes (MOCS1, MOCS2, and MOCS3). Patients in the MOCS3 subtype were observed with shorter overall survival time than patients in the MOCS1 and MOCS2 subtypes, therefore, the classification system can be used as an important prognostic tool. Similar prognostic outcomes were observed in independent external datasets. Therefore, the classification system established by us is reliable in prognostic assessment.

In recent years, molecular typing of kidney cancer has been emerging. Molecular typing of renal cancer from genomic changes, DNA methylation profiles, RNA and protein levels has revealed repeated mutations in the PI3K/AKT pathway, suggesting that this pathway is a potential therapeutic target (62). A large number of molecular typing studies of renal cancer have emerged based on single omics or specific gene sets, recently. Chen et al. integrated multi-omics data of all kidney cancer patients based on a single algorithm but did not include data of lncRNA data in the analysis (63). Ricketts et al. integrated the multi-omics data of kidney cancer for reclassification, while this study only conducted classification from the level of each omics, without realizing the real sense of integrated multi-omics data for classification (64). Although these studies provide new directions for the diagnosis and treatment of kidney cancer to some extent, they also have certain shortcomings.

10.3389/fonc.2023.1169395

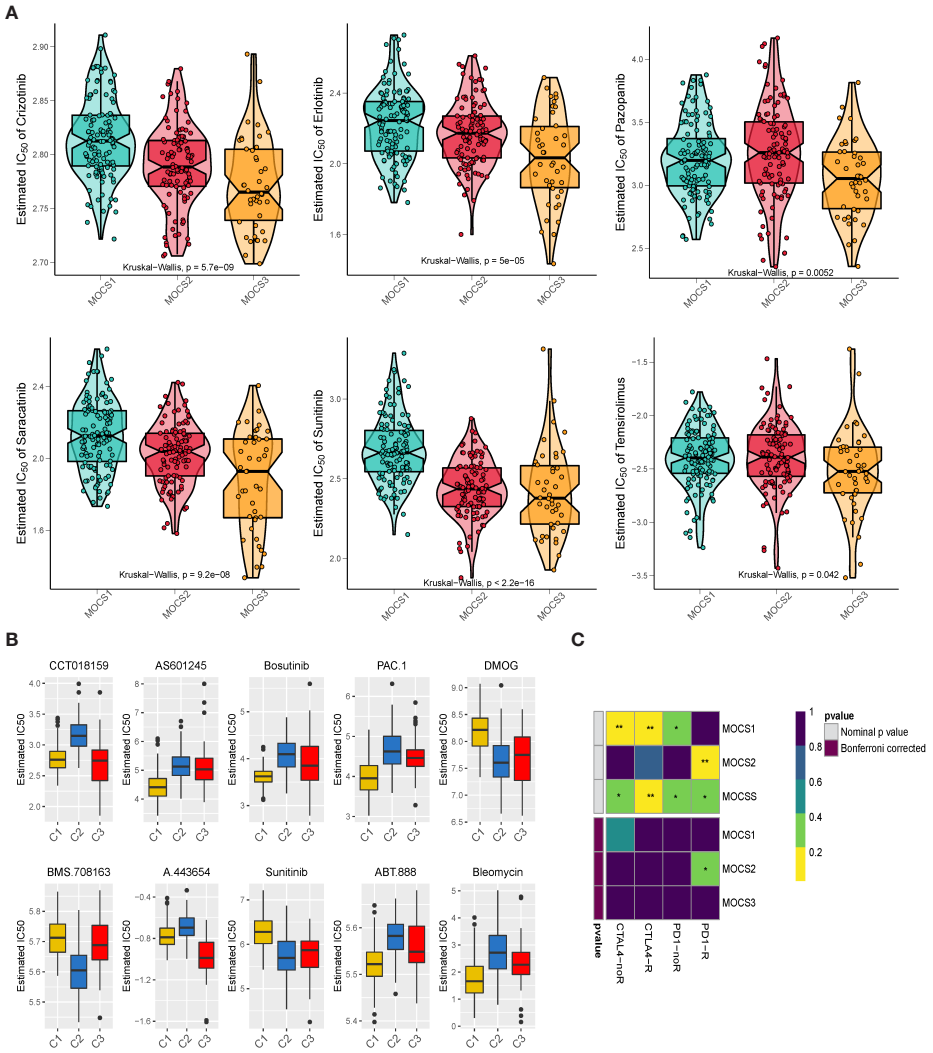


FIGURE 5
Drug Sensitivity Profiles for MOCS subtypes. **(A)** Estimated IC50 of Crizotinib, Erlotinib, Pazopanib, Saracatinib, Sunitinib, and Temsirolimus among the three MOCSs. **(B)** Estimated IC50 of CCT018159, AS601245, Bosutinib, PAC.1, DMOG, BMS.708163, A.443654, Sunitinib, ABT.888, and Bleomycin. **(C)** Subclass analysis manifested that MOCS2 were more likely to respond to anti-PD1 blockades. * means P < 0.05 and ** means P < 0.01.

The classification methods used in most typing studies are relatively simple. These shortcomings make it difficult to apply these classification studies to clinical practice. In this study, ten robust clustering algorithms based on MOVICS package were used, combined with multiple omics information, to conduct multi-omics cross-validation for patients with KIRP. Further, intra-omics heterogeneity analysis was conducted at each omics level to crack the omics differences among patients with different prognostic characteristics. Specifically, patients in the MOCS3 subtype were observed with shorter overall survival time than patients in the MOCS1 and MOCS2 subtypes. Compared to the other two subtypes, MOCS1was negatively correlated with immune-related pathways. Global dysfunction of cancer-related pathways among the three MOCS subtypes were also observed. We also evaluated the immune microenvironmental cell abundance and immune checkpoint and compared the discrepancy among the MOCSs.

Our study unexpectedly found that these three MOCSs also have significant differences in sensitivity to molecularly targeted drugs. We observed that patients in MOCS3 were more sensitive to Crizotinib, Erlotinib, Pazopanib, Saracatinib, Sunitinib, and Temsirolimus; while patients in MOCS1 were more sensitive to AS601245, Bosutinib, PAC.1, ABT.888, and Bleomycin. Whereafter, the results of subclass mapping revealed that patients in MOCS2 were more likely to respond to anti-PD1 blockades. In recent years, the treatment of kidney cancer has evolved from non-specific immune approaches to targeted therapy of vascular endothelial growth factor (VEGF), and now to novel immunotherapies. Our study assessed therapeutic differences among different subtypes and therefore can be a potential therapeutic direction for patients with KIRP.

In summary, our study provides a new reference for molecular subtypes of KIRP risk. In this study, a robust prognostic and drug selection subtype system was constructed by integrating multiple omics data using multiple algorithms. However, there are still some

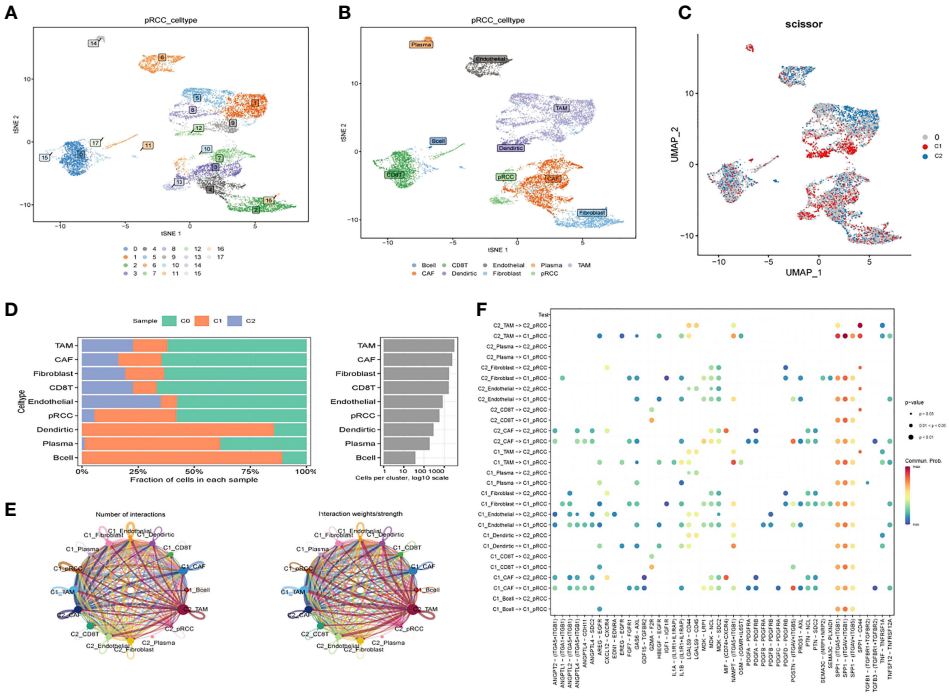


FIGURE 6
Cell–cell interactions for KIRP. **(A)** tSNE plot of the distribution of 16 samples; **(B)** tSNE plot of the distribution of nine cell clusters after clustering. **(C)** UMAP plot showing three subclusters (C0, C1, C2) of the KIRP. **(D)** The fraction of specific cell types in each cell cluster predicted by Scissor tool. **(E)** Circle plot showing the intercellular communication among major cell types in KIRP. **(F)** The ligand–receptor pairs among cells.

limitations in our study. Firstly, multi-omics data used for molecular subtypes is difficult to be applied in clinical practice. Second, although we compared the enrichment pathway and drug sensitivity between subgroups, further experiments and external data sets are still needed for verification.

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

RL designed the study. BW and ML performed data analysis. BW drafted the manuscript. RL revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the Provincial Special Supporting Fund Scientific Research Project of Shanxi Provincial People's Hospital (No. SZ2019022 and No. XY2018011), the Shanxi Fundamental Research Program (No. 201801D221382), and the Project of Administration of Traditional Chinese Medicine of Shanxi Province (2023ZYYA019).

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/fonc.2023.1169395/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Identifying and verifying of the multi omics-based cancer subtype. **(A)** Quantification of sample similarity using silhouette score based on the consensus ensembles result. **(B)** Three MOCSs were also identified as predicted by the external GSE2748 cohort. **(C)** Comparison of the overall survival time for the three MOCSs.

SUPPLEMENTARY FIGURE 2
Functional enrichment analysis of MOCS1, MOCS2 and MOCS3 subgroups. **(A)** Heatmap of metabolism-related enrichment scores among the three MOCSs. **(B)** Heatmap of immune-related enrichment scores among the three MOCSs.

SUPPLEMENTARY FIGURE 3
The landscapes of specific immune scores among the three MOCSs. **(A)** Comparison of the signature score among the three MOCSs. **(B)** Comparison of the RNAss, DNAss, DMPss, ENHss, EREG.EXPss, and HRD among the three MOCSs.

SUPPLEMENTARY FIGURE 4
Landscapes of somatic mutations and potential targets in the two subtypes. **(A)** Waterfall plot showing the mutation patterns of the top 20 most frequently mutated genes among the three MOCSs. **(B)** The synthetic lethal mutations in MOCS1, MOCS2, and MOCS3. **(C)** Potential druggable gene categories from the mutation dataset for MOCS1, MOCS2, and MOCS3. **(D)** The fraction of pathways or samples of oncogenic signaling pathways for MOCS1, MOCS2, and MOCS3.

References

1. Motzer RJ, Jonasch E, Michaelson MD, Nandagopal L, Gore JL, George S, et al. NCCN guidelines insights: Kidney cancer, version 2.2020. *J Natl Compr Canc Netw* (2019) 17:1278–85. doi: 10.6004/jnccn.2019.0054

2. Shinder BM, Rhee K, Farrell D, Farber NJ, Stein MN, Jang TL, et al. Surgical management of advanced and metastatic renal cell carcinoma: A multidisciplinary approach. *Front Oncol* (2017) 7:107. doi: 10.3389/fonc.2017.00107

3. McGuire S. World cancer report 2014. In: *Adv nutr*, vol. 7. . Geneva, Switzerland: World Health Organization, International Agency for Research on Cancer, WHO Press, 2015 (2016). p. 418–9. doi: 10.3945/an.116.012211

4. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin* (2022) 72:7–33. doi: 10.3322/caac.21708

5. Owens B. Kidney cancer. *Nature* (2016) 537:S97. doi: 10.1038/537S97a

6. Thoenes W, Störkel S, Rumpelt HJ. Histopathology and classification of renal cell tumors (adenomas, oncocytomas and carcinomas). the basic cytological and histopathological elements and their use for diagnostics. *Pathol Res Pract* (1986) 181:125–43. doi: 10.1016/s0344-0338(86)80001-2

7. Troxell ML, Higgins JP. Renal cell carcinoma in kidney allografts: Histologic types, including biphasic papillary carcinoma. *Hum Pathol* (2016) 57:28–36. doi: 10.1016/j.humpath.2016.06.018

8. Delahunt B, Eble JN. Papillary renal cell carcinoma: A clinicopathologic and immunohistochemical study of 105 tumors. *Mod Pathol* (1997) 10:537–44.

9. Allory Y, Ouazana D, Boucher E, Thiounn N, Vieillefond A. Papillary renal cell carcinoma. prognostic value of morphological subtypes in a clinicopathologic study of 43 cases. *Virchows Arch* (2003) 442:336–42. doi: 10.1007/s00428-003-0787-1

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

11. Bigot P, Bernhard JC, Gill IS, Vuong NS, Verhoest G, Flamand V, et al. The subclassification of papillary renal cell carcinoma does not affect oncological outcomes after nephron sparing surgery. *World J Urol* (2016) 34:347–52. doi: 10.1007/s00345-015-1634-0

12. Sánchez-Gastaldo A, Kempf E, González Del Alba A, Duran I. Systemic treatment of renal cell cancer: A comprehensive review. *Cancer Treat Rev* (2017) 60:77–89. doi: 10.1016/j.ctrv.2017.08.010

13. Dimitrieva S, Schlapbach R, Rehrauer H. Prognostic value of cross-omics screening for kidney clear cell renal cancer survival. *Biol Direct* (2016) 11:68. doi: 10.1186/s13062-016-0170-1

14. Mariani MP, Chen JA, Zhang Z, Pike SC, Salas LA. MethylMasteR: A comparison and customization of methylation-based copy number variation calling software in cancers harboring Large scale chromosomal deletions. *Front Bioinform* (2022) 2:859828. doi: 10.3389/fbinf.2022.859828

15. Miranda-Poma J, Trilla-Fuertes L, López-Camacho E, Zapater-Moros A, López-Vacas R, Lumbreras-Herrera MI, et al. MiRNAs in renal cell carcinoma. *Clin Transl Oncol* (2022) 24:2055–63. doi: 10.1007/s12094-022-02866-z

16. Mishra SK, Thakran P. Intron specificity in pre-mRNA splicing. *Curr Genet* (2018) 64:777–84. doi: 10.1007/s00294-017-0802-8

17. Wahl MC, Will CL, Lührmann R. The spliceosome: Design principles of a dynamic RNP machine. *Cell* (2009) 136:701–18. doi: 10.1016/j.cell.2009.02.009

18. Sillekens PT, Beijer RP, Habets WJ, van Verooij WJ. Molecular cloning of the cDNA for the human U2 snRNA-specific a' protein. *Nucleic Acids Res* (1989) 17:1893–906. doi: 10.1093/nar/17.5.1893

19. Kim YD, Lee J, Kim HS, Lee MO, Son MY, Yoo CH, et al. The unique spliceosome signature of human pluripotent stem cells is mediated by SNRPA1, SNRPD1, and PNN. *Stem Cell Res* (2017) 22:43–53. doi: 10.1016/j.scr.2017.05.010

20. Liu Y, Zhu X, Zhu J, Liao S, Tang Q, Liu K, et al. Identification of differential expression of genes in hepatocellular carcinoma by suppression subtractive hybridization combined cDNA microarray. *Oncol Rep* (2007) 18:943–51. doi: 10.3892/or.18.4.943

21. Dou N, Yang D, Yu S, Wu B, Gao Y, Li Y, et al. SNRPA enhances tumour cell growth in gastric cancer through modulating NGF expression. *Cell Prolif* (2018) 51: e12484. doi: 10.1111/cpr.12484

22. Fish L, Khoroshkin M, Navickas A, Garcia K, Culbertson B, Hänisch B, et al. A prometastatic splicing program regulated by SNRPA1 interactions with structured RNA elements. *Science* (2021) 372(6543):eabc7531. doi: 10.1126/science.abc7531

23. Bolduc F, Turcotte MA, Perreault JP. The small nuclear ribonucleoprotein polypeptide a (SNRPA) binds to the G-quadruplex of the BAG-1 5'UTR. *Biochimie* (2020) 176:122–7. doi: 10.1016/j.biochi.2020.06.013

24. Zeng Q, Lei F, Chang Y, Gao Z, Wang Y, Gao Q, et al. An oncogenic gene, SNRPA1, regulates PIK3R1, VEGFC, MKI67, CDK1 and other genes in colorectal cancer. *BioMed Pharmacother* (2019) 117:109076. doi: 10.1016/j.biopha.2019.109076

25. Yuan P, Ling L, Gao X, Sun T, Miao J, Yuan X, et al. Identification of RNA-binding protein SNRPA1 for prognosis in prostate cancer. *Aging (Albany NY)* (2021) 13:2895–911. doi: 10.18632/aging.202387

26. Négrier S, Gravis G, Pérol D, Chevreau C, Delva R, Bay JO, et al. Temsirolimus and bevacizumab, or sunitinib, or interferon alfa and bevacizumab for patients with advanced renal cell carcinoma (TORAVA): A randomised phase 2 trial. *Lancet Oncol* (2011) 12:673–80. doi: 10.1016/s1470-2045(11)70124-3

27. Ljungberg B, Bensalah K, Canfield S, Dabestani S, Hofmann F, Hora M, et al. EAU guidelines on renal cell carcinoma: 2014 update. *Eur Urol* (2015) 67:913–24. doi: 10.1016/j.eururo.2015.01.005

28. Wang Z, Jensen MA. & zenklusen, j. c. a practical guide to the cancer genome atlas (TCGA). *Methods Mol Biol* (2016) 1418:111–41. doi: 10.1007/978-1-4939-3578-9_6

29. Yang XJ, Tan MH, Kim HL, Ditlev JA, Betten MW, Png CE, et al. A molecular classification of papillary renal cell carcinoma. *Cancer Res* (2005) 65:5628–37. doi: 10.1158/0008-5472.Can-05-0533

30. Su C, Lv Y, Lu W, Yu Z, Ye Y, Guo B, et al. Single-cell RNA sequencing in multiple pathologic types of renal cell carcinoma revealed novel potential tumor-specific markers. *Front Oncol* (2021) 11:719564. doi: 10.3389/fonc.2021.719564

31. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* (2015) 33:495–502. doi: 10.1038/nbt.3192

32. Lu X, Meng J, Zhou Y, Jiang L, Yan F. MOVICS: An r package for multi-omics integration and visualization in cancer subtyping. *Bioinformatics* (2020) 36(22–23):5539–41. doi: 10.1093/bioinformatics/btaa1018

33. Hu X, Wang Z, Wang Q, Chen K, Han Q, Bai S, et al. Molecular classification reveals the diverse genetic and prognostic features of gastric cancer: A multi-omics consensus ensemble clustering. *BioMed Pharmacother* (2021) 144:112222. doi: 10.1016/j.biopha.2021.112222

34. Hoshida Y. Nearest template prediction: A single-sample-based flexible class prediction with confidence assessment. *PloS One* (2010) 5:e15543. doi: 10.1371/journal.pone.0015543

35. Hu J, Tzeng JY. Integrative gene set analysis of multi-platform data with sample heterogeneity. *Bioinformatics* (2014) 30:1501–7. doi: 10.1093/bioinformatics/btu060

36. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43:e47. doi: 10.1093/nar/gkv007

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

38. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res* (2017) 77:e108–10. doi: 10.1158/0008-5472.Can-17-0307

39. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol* (2018) 1711:243–59. doi: 10.1007/978-1-4939-7493-1_12

40. Zhu K, Xiaoqiang L, Deng W, Wang G, Fu B. Development and validation of a novel lipid metabolism-related gene prognostic signature and candidate drugs for patients with bladder cancer. *Lipids Health Dis* (2021) 20:146. doi: 10.1186/s12944-021-01554-1

41. Plattner C, Finotello F, Rieder D. Deconvoluting tumor-infiltrating immune cells from RNA-seq data using quanTIseq. *Methods Enzymol* (2020) 636:261–85. doi: 10.1016/bs.mie.2019.05.056

42. Sturm G, Finotello F, List M. Immunedeconv: An r package for unified access to computational methods for estimating immune cell fractions from bulk RNA-sequencing data. *Methods Mol Biol* (2020) 2120:223–32. doi: 10.1007/978-1-0716-0327-7_16

43. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol* (2017) 18:220. doi: 10.1186/s13059-017-1349-1

44. Racle J, Gfeller D. EPIC: A tool to estimate the proportions of different cell types from bulk gene expression data. *Methods Mol Biol* (2020) 2120:233–48. doi: 10.1007/978-1-0716-0327-7_17

45. Yoshihara K, Shahmoradgoli M, Martínez 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

46. Jeschke J, Bizet M, Desmedt C, Calonne E, Dedeurwaerder S, Garaud S, et al. DNA Methylation-based immune response signature improves patient diagnosis in multiple cancers. *J Clin Invest* (2017) 127:3090–102. doi: 10.1172/jci91095

47. 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

48. Jiang A, Song J, Fang X, Fang Y, Wang Z, Liu B, et al. A novel thinking: DDR axis refines the classification of ccRCC with distinctive prognosis, multi omics landscape and management strategy. *Front Public Health* (2022) 10:1029509. doi: 10.3389/fpubh.2022.1029509

49. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* (2014) 513:202–9. doi: 10.1038/nature13480

50. Geeleher P, Cox NJ, Huang RS. Clinical drug response can be predicted using baseline gene expression levels and. *Vitro Drug sensitivity Cell lines. Genome Biol* (2014) 15:R47. doi: 10.1186/gb-2014-15-3-r47

51. 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

52. McGranahan N, Furness AJ, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* (2016) 351:1463–9. doi: 10.1126/science.aaf1490

53. Lu X, Jiang L, Zhang L, Zhu Y, Hu W, Wang J, et al. Immune signature-based subtypes of cervical squamous cell carcinoma tightly associated with human papillomavirus type 16 expression, molecular features, and clinical outcome. *Neoplasia* (2019) 21:591–601. doi: 10.1016/j.neo.2019.04.003

54. Zhang G, Liu Z, Deng J, Liu L, Li Y, Weng S, et al. Smooth muscle cell fate decisions decipher a high-resolution heterogeneity within atherosclerosis molecular subtypes. *J Transl Med* (2022) 20:568. doi: 10.1186/s12967-022-03795-9

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

56. Fernando A. Time to focus on the rare-encouraging progress in the management of non-clear cell renal cell carcinoma. *Eur Urol* (2017) 72:984–5. doi: 10.1016/j.eururo.2017.07.028

57. Scelo G, Larose TL. Epidemiology and risk factors for kidney cancer. *J Clin Oncol* (2018) 36:Jco2018791905. doi: 10.1200/jco.2018.79.1905

58. Gray RE, Harris GT. Renal cell carcinoma: Diagnosis and management. *Am Fam Physician* (2019) 99:179–84.

59. Ohashi R, Schraml P, Batavia A, Angori S, Simmler P, Rupp N, et al. Allele loss and reduced expression of CYCLOPS genes is a characteristic feature of chromophobe renal cell carcinoma. *Transl Oncol* (2019) 12:1131–7. doi: 10.1016/j.tranon.2019.05.005

60. Lalani AA, McGregor BA, Albiges L, Choueiri TK, Motzer R, Powles T, et al. Systemic treatment of metastatic clear cell renal cell carcinoma in 2018: Current paradigms, use of immunotherapy, and future directions. *Eur Urol* (2019) 75:100–10. doi: 10.1016/j.eururo.2018.10.010

61. Subramanian I, Verma S, Kumar S, Jere A, Anamika K. Multi-omics data integration, interpretation, and its application. *Bioinform Biol Insights* (2020) 14:1177932219899051. doi: 10.1177/1177932219899051

62. Chowdhury N, Drake CG. Kidney cancer: An overview of current therapeutic approaches. *Urol Clin North Am* (2020) 47:419–31. doi: 10.1016/j.ucl.2020.07.009

63. Nargund AM, Pham CG, Dong Y, Wang PI, Osmangeyoglu HU, Xie Y, et al. The SWI/SNF protein PBRM1 restrains VHL-Loss-Driven clear cell renal cell carcinoma. *Cell Rep* (2017) 18:2893–906. doi: 10.1016/j.celrep.2017.02.074

64. Peña-Llopis S, Vega-Rubín-de-Celis S, Liao A, Leng N, Pavía-Jiménez A, Wang S, et al. BAP1 loss defines a new class of renal cell carcinoma. *Nat Genet* (2012) 44:751–9. doi: 10.1038/ng.2323

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 03 May 2023
DOI 10.3389/fimmu.2023.1162843

OPEN ACCESS

EDITED BY
Nan Zhang,
Harbin Medical University, China

REVIEWED BY
Huashan Shi,
Sichuan University, China
Renato Brito Baleeiro,
Queen Mary University of London, United Kingdom
Yang Wang,
Chinese Academy of Medical Sciences and Peking Union Medical College, China

*CORRESPONDENCE
Bin Luo
vibrio@163.com
Lin Wang
wanglinron@126.com

†These authors share first authorship

RECEIVED 10 February 2023
ACCEPTED 13 April 2023
PUBLISHED 03 May 2023

CITATION
Shen X, Zuo X, Liang L, Wang L and Luo B (2023) Integrating machine learning and single-cell trajectories to analyze T-cell exhaustion to predict prognosis and immunotherapy in colon cancer patients. *Front. Immunol.* 14:1162843. doi: 10.3389/fimmu.2023.1162843

Integrating machine learning and single-cell trajectories to analyze T-cell exhaustion to predict prognosis and immunotherapy in colon cancer patients

Xiaogang Shen[1,2†], Xiaofei Zuo[1,2†], Liang Liang[2,3], Lin Wang[1,2*] and Bin Luo[1,2*]

[1]Department of Gastrointestinal Surgery, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, China, [2]Chinese Academy of Sciences Sichuan Translational Medicine Research Hospital, Chengdu, China, [3]Cancer Center, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, China

Introduction: The incidence of colon adenocarcinoma (COAD) has recently increased, and patients with advanced COAD have a poor prognosis due to treatment resistance. Combining conventional treatment with targeted therapy and immunotherapy has shown unexpectedly positive results in improving the prognosis of patients with COAD. More study is needed to determine the prognosis for patients with COAD and establish the appropriate course of treatment.

Methods: This study aimed to explore the trajectory of T-cell exhaustion in COAD to predict the overall survival and treatment outcome of COAD patients. Clinical data were derived from the TCGA-COAD cohort through "UCSC", as well as the whole genome data. Prognostic genes driving T-cell trajectory differentiation were identified on the basis of single-cell trajectories and univariate Cox regression. Subsequently, T-cell exhaustion score (TES) was created by iterative LASSO regression. The potential biological logic associated with TES was explored through functional analysis, immune microenvironment assessment, immunotherapy response prediction, and in vitro experiments.

Results: Data showed that patients with significant TES had fewer favorable outcomes. Expression, proliferation, and invasion of COAD cells treated with TXK siRNA were also examined by cellular experiments. Both univariate and multivariate Cox regression indicated that TES was an independent prognostic factor in patients with COAD; in addition, subgroup analysis supported this finding. Functional assay revealed that immune response and cytotoxicity pathways are associated with TES, as the subgroup with low TES has an active immune microenvironment. Furthermore, patients with low TES responded better to chemotherapy and immunotherapy.

Conclusion: In this study, we systematically explored the T-cell exhaustion trajectory in COAD and developed a TES model to assess prognosis and provide guidelines for the treatment decision. This discovery gave rise to a fresh concept for novel therapeutic procedures for the clinical treatment of COAD.

KEYWORDS

colon adenocarcinoma (COAD), t-cell exhaustion, tumor microenvironment, prognosis, immunotherapy

Introduction

Colon adenocarcinoma (COAD) is a malignant gastrointestinal tumor that originates in the colon and is the third most prevalent tumor and the second leading cause of cancer-related death worldwide (1). Early-stage COAD is often difficult to detect, and despite rapid advances in early cancer screening technology, most COAD patients are diagnosed with advanced COAD only when they have obvious symptoms such as hematochezia and colonic obstruction, etc (2). Patients with advanced COAD are often not eligible for surgical resection and can only be treated with traditional chemotherapy and targeted therapy (Monoclonal antibody drugs such as bevacizumab and cetuximab) (3, 4). However, treatment resistance often occurs in advanced COAD, leading to a worse prognosis for patients (5). As cancer treatment has entered the field of immunotherapy, most cancer patients have achieved clinical success (6). However, only a small percentage of patients respond positively to immunotherapy.

Most patients with advanced cancer have T cells in a state of exhaustion, and T-cell exhaustion is an important factor in the efficacy of immunotherapy (7). T-cell exhaustion is a common feature of the cancer process and immune dysfunction, resulting from sustained antigenic stimulation and immune response (8). Exhausted T cells have a progressive loss of immune effector function, persistent high expression of suppressor receptors (such as ENTPD1, LAYN, LAG3, and HAVCR2) (9), and loss of self-renewal capacity, along with a unique transcriptional signature (8, 10). Recent studies suggest that interventions to alleviate T-cell exhaustion may lead to superior clinical outcomes and dramatic advances in cancer immunotherapy (11). Encouragingly, related studies have made some progress in lung cancer (12, 13). Thus, finding pre-depleted T cells in COAD could lead to a larger clinical window. Single-cell sequencing technology provides a new perspective for analyzing T-cell exhaustion trajectories (14), and by integrating scRNA-Seq data and bulk RNA-seq data, we may be able to gain a preliminary understanding of the T-cell trajectories and core regulatory targets that are pre-exhausted in COAD. This will facilitate initial protocol development and the development of novel targeted therapies.

Tumor processes are complex dynamic systems, and although it is commonly assumed that exhausted T cells result from sustained antigenic stimulation, the phenotype and transcriptional characteristics of exhausted T cells are also shaped by multiple factors in the immune microenvironment (15). Including the expression of suppressive receptors and ligands, the regulation of suppressor cells such as Tregs, suppressor cytokines such as IL-10 and TGFb, and some metabolic products also suppress T-cell function (15). Identification of the transcriptional pathways mediating T cell dysfunction is complex because the genetic profile of exhausted T cells overlaps to a considerable extent with that of activated T cells (16, 17). It has been suggested that failing T cells cannot be accurately defined by suppressor molecules alone (8), however, the driver genes that regulate T cell failure in COAD are currently unknown. These genes may be a breakthrough in targeting T-cell exhaustion and may also serve as important clinical prognostic treatment targets.

In this study, we aimed to characterize the dynamic trajectory of T-cell exhaustion in COAD and identify prognostic markers associated with exhausted CD8+ T cells. We first depicted the evolutionary trajectory of CD8+ T cells in a single-cell dataset and identified core regulatory genes of the exhausted CD8+ T trajectory based on the pseudo-time trajectory. Based on these genes we constructed a T-cell exhaustion score (TES) to assist in prognosis and quantify the degree of exhaustion. We then assessed the heterogeneity of different TES subgroups in terms of function, immune infiltration, and genomic alterations, and evaluated the predictive efficacy of the TES for immunotherapy. Finally, we preliminarily validated the core genes of TES by qRT-PCR, CCK8 assay, and transwell invasion experiment. In conclusion, this study not only identified the trajectory of exhausted CD8+ T cells in COAD and provided a tool to quantify T-cell exhaustion. Moreover, it confirmed the reliable efficacy of T-cell exhaustion in predicting COAD prognosis and immunotherapy. We hope that this study will provide novel prognostic markers and immunotherapeutic targets for COAD patients.

Methods

Data collection

For the TCGA-COAD cohort, we obtained copy number variant (CNV) data from the UCSC Xena (https://xena.ucsc.edu) database, somatic variant data from maf files on the Muctect 2

platform, and transcriptome RNA-seq data. Additionally, relevant clinical follow-up data was gathered. A TCGA-COAD cohort of 432 COAD patients was collected and utilized as a training cohort after patients with pathological typing as COAD were included and patients with missing follow-up information were excluded. Additionally, data from three large COAD cohorts: GSE14333 (18), GSE17536 (19), and GSE41258 (20) from the GEO database was gathered. After data combination and removing those with insufficient follow-up data, 654 patients with COAD were included after collecting the patient follow-up data from the original Supplementary Material. The external validation was done using the meta-GEO COAD cohort. Finally, a single-cell transcriptome dataset GSE146771 of 10 primary tumor sections, was obtained and processing use "Seurat" R packages. We explore T-cell exhaustion trajectories in colon cancer through the single-cell data. The specific data processing and standardization pipeline can be obtained from the original article (21).

Exploring T-cell depletion trajectories in colon cancer

First, we evaluated the cytotoxicity score and cell exhaustion score of each cell in the scRNA-seq dataset by the "AUCell" package based on previously reported genetic markers (9, 14). Subsequently, the R package "monocle" was used to calculate and map the pseudo-time trajectories of T cells. The differentialGeneTest() function was used to calculate the characteristic genes in different trajectories. Finally, we identified the signature genes in the T cell depletion trajectories as T cell depletion markers in COAD.

Construction of the T-cell exhaustion scoring model

In order to find independent predictive markers for COAD, we first conducted a univariate cox regression analysis for T-cell exhaustion markers. The T-cell exhaustion score (TES) was then created using LASSO regularization through 300 random iterations. After selecting a penalty factor λ, the regularization model will remove insignificant markers and generate coefficients for each TES model gene. To prevent overfitting, we set up a 5-fold cross-validation and determine the final stable TES model based on the number of builds in 300 random iterations. The final TES model was generated GAS according to the following equation:

$$TES = \sum iCoefficient(mRNA_i) \times Expression(mRNA_i)$$

In order to evaluate prognostic effectiveness, the "survcomp" program computed the C-index of the TES (22). A prediction made by the model that is more optimum and stable has a higher C-index. The independent prognostic significance of TES was thoroughly investigated using Kaplan-Meier survival analysis, univariate and multifactorial Cox regression, and time-dependent ROC (tROC) curves. High TES and low TES groups were separated by the median value of TES. Finally, to measure the chance of survival more accurately for specific patients, we created a nomogram based on TES and other clinical characteristics.

Cell culture

The normal human colonic epithelial cell line NCM460 and the human colon cancer cell lines SW460 and SW48 were bought from Bioss, China. All cells were grown in DMEM media with 10% FBS in a 37°C cell incubator with 5% CO2.

qRT-PCR

We then used qRT-PCR to assay patient tissues and COAD cell lines to assess TXK level. ChamQ Universal SYBR qPCR Master Mix was used to run each real-time PCR experiment (Vazyme, China). Using GAPDH as a control, the amplified PCR products were measured and standardized.

Cell proliferation detection

For the transfection of siRNA in this work, LipofectamineTM 2000 Transfection Reagent (Invitrogen, USA) was used. Cell Counting Kit-8 kit was used to measure the proliferation rate of COAD cells (Bioss, China). At a density of around 1500 cells per well, the digested single-cell solution was injected in 96-well plates. Three wells from each group were chosen at random at 0, 12, 24, 48, and 72 hours. Then, 10 mL of the Cell Counting Kit-8 reagent was added, and the wells were incubated at 37°C for two hours. identification of 450 nm absorbance values.

Transwell cell invasion analysis

We used a Transwell kit (Merck Millipore, USA) with a pore size of 8um to detect the degree of invasion of different SW460 cells. Briefly, SW460 cells were inoculated in the upper chamber of a 24-well plate, and DMEM medium containing 20% FBS was added dropwise in the lower chamber. Cells in the upper layer of transwell chambers were wiped off with a cotton swab after incubation for 32 h at 37°C in an incubator. The invading cells were stained with 0.1% crystal violet staining solution (Solarbio, China) and counted using ImageJ software after microscopic visualization.

Assessment of immune heterogeneity between TES subgroups

In each COAD sample, we calculated the relative abundance of 22 different immune cell types using the "CIBERSORT" program (23). The "ESTIMATE" system evaluated the samples' immunological score and tumor purity (24). The ssGSEA algorithm of the "GSVA" software was then used to evaluate the activity of the relevant immunological

pathways. The changes in the expression of 6 classical immunological checkpoints between subgroups were then examined.

Dissecting genomic alterations between subgroups

The "maftools" package was used to handle the maf files, and it computed the amount of nonsynonymous mutations for each patient (25). The variations in top 20 mutated genes across subgroups were next examined using OncodriveCLUST algorithm. We also used the "Sigminer" package to extract significant mutation signatures from the maf files for different subgroups and compared the mutation signatures with the COSMIC database (26). Finally, Gistic2.0 was used to process CNV data and count amplicons and deletions according to a threshold of 0.2. The "ggplot2" package was used for visualization.

Assessment of chemotherapy applications for TES

Three medications routinely used in COAD (5-FU, Cisplatin, and Camptothecin) were initially predicted using "pRRophetic," which was built on the GDSC database (27). IC50 values were estimated by ridge regression, with lower IC50 values indicating higher sensitivity. Since the differentially expressed genes between the high and low TES subgroups were thought to represent potential therapeutic targets, we uploaded the Top150 up- and down-regulated genes to the CMap database (https://clue.io/) to investigate prospective small molecule compounds. Additionally, to revealing the biomolecular pathways that medications target, it may infer pharmaceuticals based on gene expression patterns.

Predicting immunotherapy response

We calculated the Immunophenoscore (IPS) of patients based on the genetic profile of different immune cell phenotypes (28). A higher IPS indicates an active immune response and a higher response to immunotherapy. We used the TIDE method to mimic the tumor immune escape mechanism in order to forecast how each patient would respond therapeutically to immune checkpoint inhibitors (29). In addition, we collected two well-established immunotherapy cohort, Imvigor210, which contained 298 patients with complete follow-up information who received anti-PD-L1 immunotherapy for uroepithelial cancer (30). And Liu David. cohort, which contained 121 patients who received anti-PD-1 immunotherapy for melanoma (31). In order to evaluate the TES's immunotherapy prediction capability, the transcriptome data from the Imvigor210 cohort and Liu David. cohort were utilized to build the TES based on the same methodology.

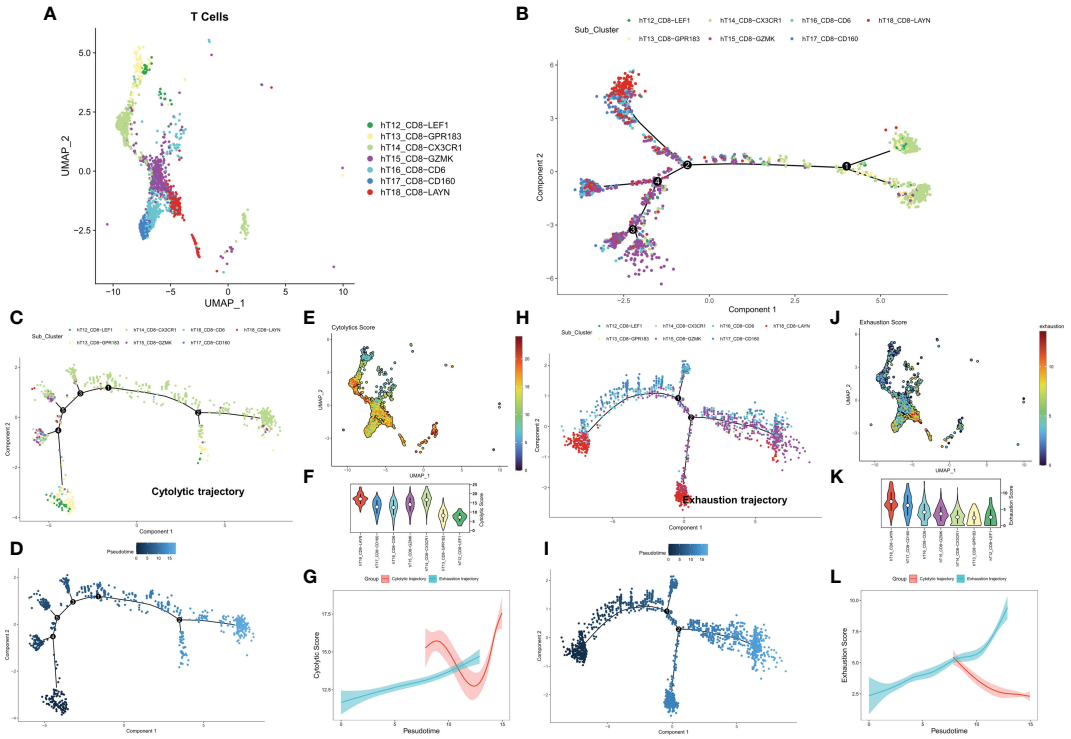


FIGURE 1
Exploring the trajectory of T-cell exhaustion in COAD. **(A)** The landscape of 7 CD8+ T cell subtypes in GSE146771. **(B)** Overall pseudo-time differentiation trajectory of seven T cell subtypes. **(C, D)** The pseudo-time differentiation trajectory of cytolytic T cell subtypes. **(E)** Density of cytolytic scores of different T cell subtypes. **(F)** Comparison of cytolytic scores of different T cell subtypes. **(G)** Cytolytic scores for different differentiation trajectories. **(H-J)** The pseudo-time differentiation trajectory of exhaustion T cell subtypes. **(J)** Density of exhaustion scores of different T cell subtypes. **(K)** Comparison of exhaustion scores of different T cell subtypes. **(L)** Exhaustion scores for different differentiation trajectories.

Bioinformatics and statistical analysis

Fisher's exact test was used to find proportional differences, the Wilcoxon test or T-test to determine group differences, the Kaplan-Meier plotter to produce survival curves, and the log-rank test to detect differences in survival. The R package "survivalROC" was used to plot time-dependent ROC curves (tROC). Using the R package "survival," univariate and multivariate Cox regressions were carried out. The nomogram and calibration curves were plotted using the R package "rms". The prediction power of several variables on the outcome of immunotherapy was evaluated using the R package "pROC". If not mentioned differently, two-tailed p-values 0.05 were regarded as significant. Every analysis was done using the R software (Version 4.1.0).

Results

Exploring T-cell exhaustion trajectories in colon cancer

We first identified seven CD8+ T cell subtypes based on the cell annotation in the original article (Figure 1A). The overall differentiation trajectory of CD8+ T cells was then identified by monocle algorithm (Figure 1B). The results showed that CD8-CX3CR1 was at the beginning of the trajectory, while CD8+ T cells with high expression of the exhaustion marker LAYN were distributed at the end of the trajectory. Subsequently, we differentiated two different developmental trajectories based on the cell exhaustion fraction and cytolytic fraction. Among them, the cytolytic trajectory was mainly composed of CD8-CX3CR1 cells with CD8-LEF1 and CD8-GPR183 at the beginning (Figures 1C, D). We found the strongest cytolytic activity of the CD8-CX3CR1 (Figures 1E, F). Over time, the exhaustion score showed an increasing trend while the cytolytic score showed a decreasing and then increasing trend (Figure 1G). The cell exhaustion trajectory was mainly composed of CD8-LAYN and CD8-GZMK, with CD8-LAYN distributed at the beginning and the end of the trajectory and CD8-GZMK mainly distributed at the end of the trajectory (Figures 1H, I). Subsequently, we found the highest exhaustion score of CD8-LAYN, which confirmed the reliability of the exhaustion trajectory (Figures 1J, K). Over time, the cell exhaustion score increased while the cytolytic activity decreased (Figure 1L). Finally, we identified 477 ordered genes in the cell exhaustion trajectory as the T-cell exhaustion markers in COAD.

Dissecting key T-cell exhaustion genes in COAD

Based on a P<0.05 criterion, 27 T-cell exhaustion genes were discovered (Figure 2A). A correlation network for these 27 genes was constructed and the results indicated that most of them were positively correlated (20/27) (Figure 2B). In the TCGA-COAD cohort, the mutation landscape of these 27 genes was shown in Figure 2C. The gene with the greatest frequency of mutations is RASGRP2, and missense mutations are the most common form of mutation (Figure 2D). Finally, we summarized the CNV events of the 27 key genes (Figure 2E). The results showed that prevalent CNV events occurred in most genes, the highest amplification frequency was LIME1, and the highest deletion frequency was RUNX3.

Construction of T-cell exhaustion score

To construct a more robust TES model, we enrolled 27 T-cell genes with independent prognostic efficacy and performed 300 iterations of LASSO regression to retrieve the most robust model. The results showed that the model containing 13 genes was the most robust TES model (215/300) (Figure 3A). Good predictive efficacy was demonstrated in both TCGA and meta-GEO cohorts (C index: 0.666 for TCGA; 0.635 for GEO) (Figure 3A). Compared to the commonly used clinical indicators, TES was slightly weaker than stage but better than age and gender (Figure 3B). In many COAD cohorts, survival analysis revealed that patients in the high TES group had worse results than its rival (Figures 3C, D). Roc analysis showed that TES had acceptable performance in the TCGA cohort (ROC>0.65, Figure 3E), while a favorable effect was also observed in the GEO cohort (ROC>0.65, Figure 3F). The tROC results showed that TES had an effective performance in predicting survival within five years in both cohort (Figures 3G, H).

Analysis of the predictive efficacy and independence of TES

First, we used univariate and multifactorial Cox regression to investigate the relationship between TES and patients' clinical variables (e.g., age, gender, and stage). In the training and validation cohorts, univariate Cox regression indicated that TES was an independent predictor (p 0.05) (Figure 4A). In both the training and validation cohorts, multifactorial Cox regression showed that TES remained an unfavorable predictor of OS (P<0.05) (Figure 4B). In addition, subgroup analysis showed that TES performed best in predicting prognosis in all subgroups of patients, especially those with advanced tumors in all age groups (Figure 4C). Therefore, TES could be a trustworthy prognostic indicator of OS in COAD patients. We then created a Nomogram to more accurately measure the risk assessment of COAD patients (Figure 4D). The calibration curve of Nomogram showed high stability and accuracy after 1, 3 and 5 years (Figure 4E). tROC study showed that Nomogram model performed better than TES alone (Figure 4F). Finally, DCA analysis showed that the nomogram model had the best decision effectiveness after 1, 3 and 5 years (Figure 4G).

Cellular experimental validation of key TES model indicator

We extracted risk coefficients for each indicator in the final TES model, and the results showed that TXK was the most potent risk factor (Figure 5A). We sought to explore whether TXK affects the malignant

10.3389/fimmu.2023.1162843

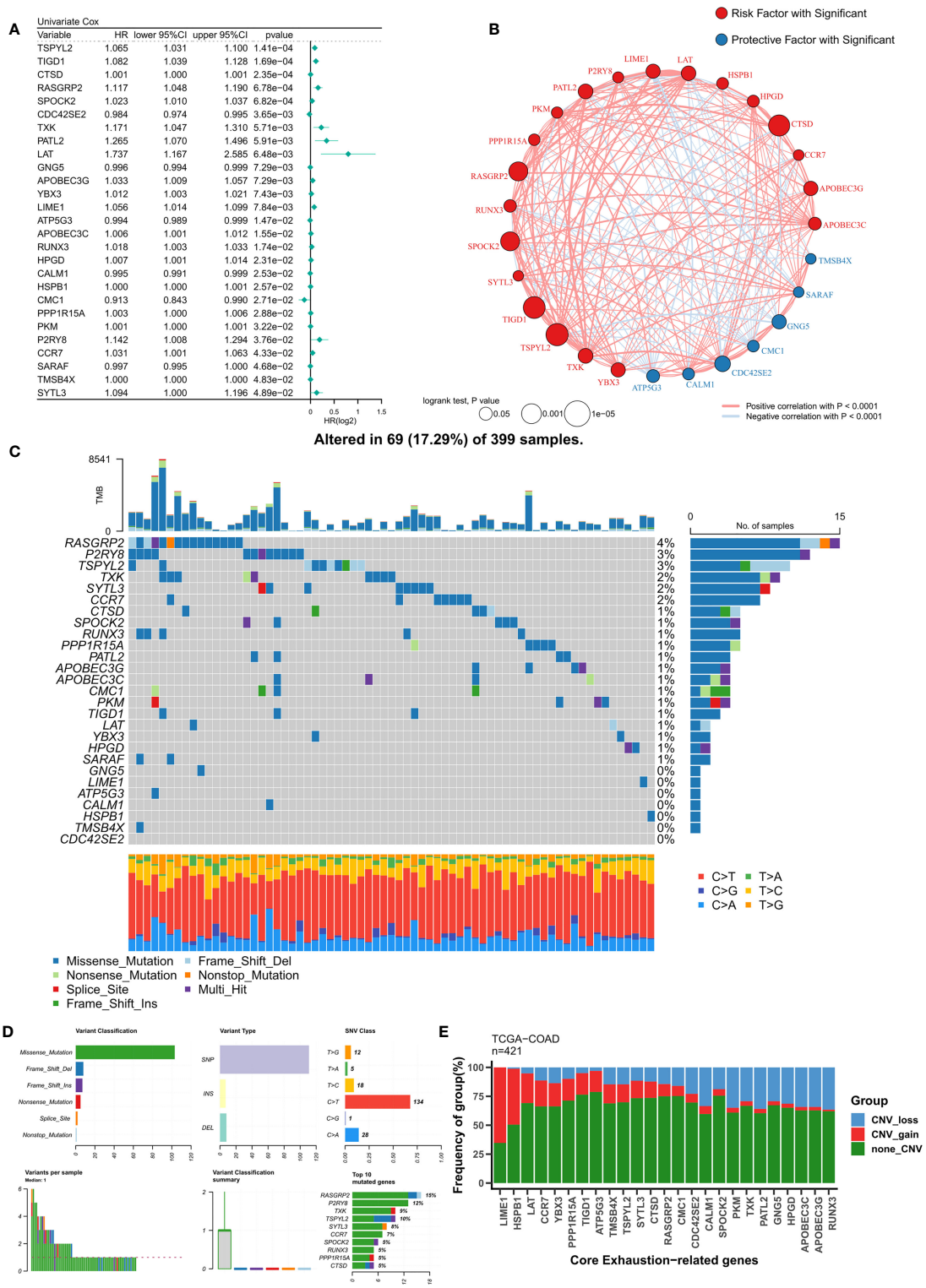


FIGURE 2
Identification of ordered T-cell exhaustion indicators in TCGA-COAD. **(A)** Univariate Cox regression identified 27 key T-cell exhaustion indicators with prognostic efficacy. **(B)** The correlation network of 27 key T-cell exhaustion indicators. **(C)** The landscape of somatic mutation of 27 key T-cell exhaustion indicators. **(D)** The summary of somatic mutation of 27 key T-cell exhaustion indicators. **(E)** The summary of CNV status of 27 key T-cell exhaustion indicators.

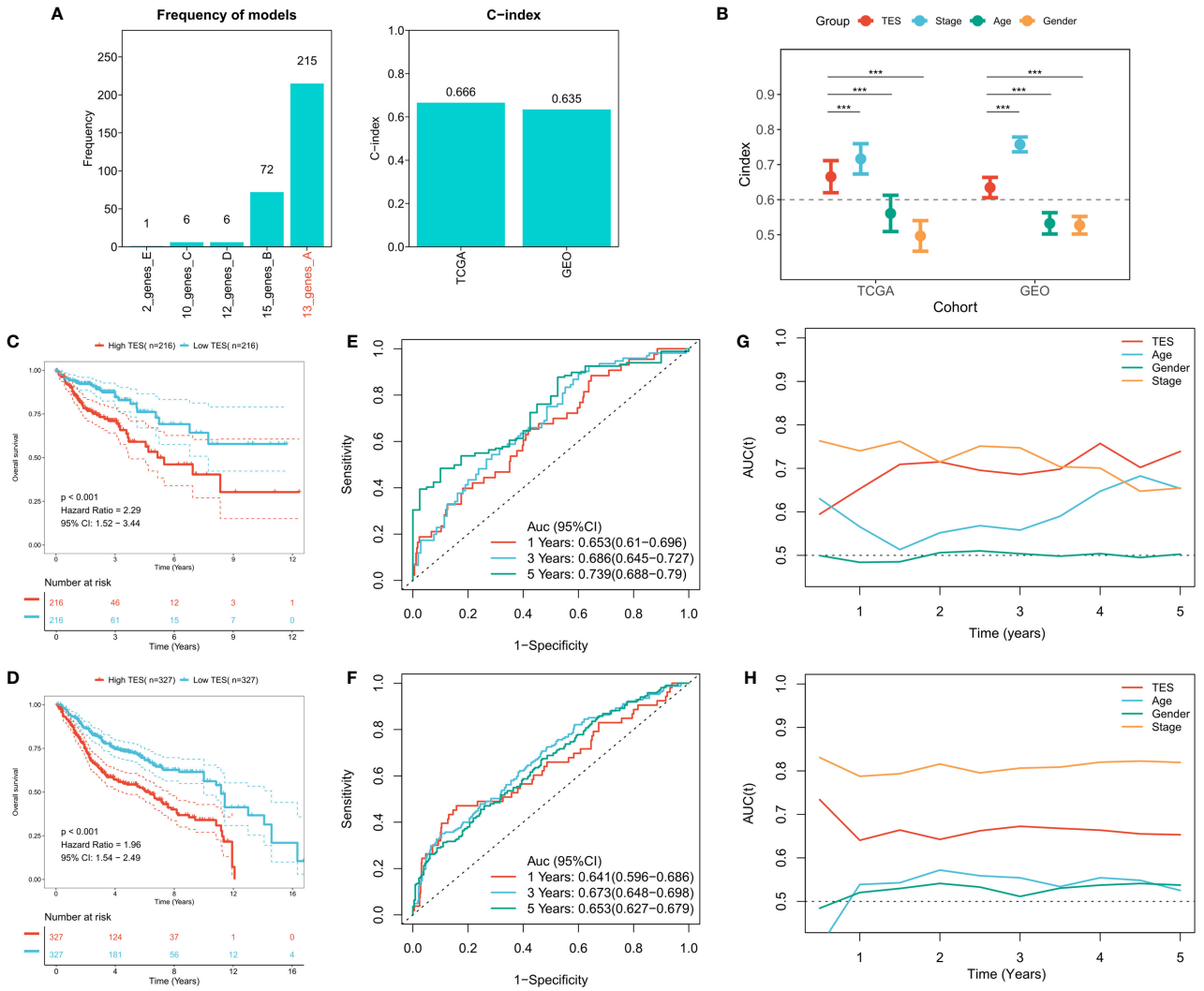


FIGURE 3
Construction of the T-cell exhaustion scoring model. **(A)** Iterative LASSO regression to select the most stable prognostic model. left: frequency of different gene pairs in LASSO models; right: C-index of the best combination in TCGA and GEO datasets. **(B)** Comparison of C-index differences between TES model and clinical characteristics. ***: P<0.001. **(C)** Kaplan-Meier survival curve of patients with high and low TES in TCGA dataset. **(D)** Kaplan-Meier survival curve of patients with high and low TES in the meta-GEO dataset. **(E)** ROC curves of TES at 1, 3, and 5 years in the TCGA dataset. **(F)** ROC curves of TES in the meta-GEO dataset at 1, 3, and 5 years. **(G)** tROC curves of TES in 5 years in the TCGA dataset. **(H)** tROC curves of TES in 5 years in the GEO dataset.

activity of tumor cells to influence prognosis through cellular experiments. We first found that the mRNA expression level of TXK was increased in COAD cell lines compared to normal colonic epithelial cell lines (Figure 5B). We then found the reduced proliferative activity of cells after knockdown of TXK in SW480 cell line by CCK8 kit (Figure 5C). After knockdown of TXK in SW460, invasive cells in transwell cells were reduced (Figure 5D). By counting the invading cells, we found that the degree of invasion of SW460 cells was significantly reduced after knockdown of TXK (Figure 5E).

Low TES is associated with abundant immune infiltration

We then dissected the tumor immune microenvironment of TES. Estimate results revealed more tumor purity in the high TES group, while the low TES group had better estimate and immunological scores (Figure 6A). Further we found elevated expression of six typical immune checkpoints (PD-1, CTLA-4, LAG-3, TIM-3, PD-L2, and PD-L1) in low TES (Figure 6A). We also found increased enrichment of CD8+ and CD4+ T cells in the low TES group and increased infiltration of Tregs and M0 macrophages in the high TES group (Figure 6A). Subsequently, we examined the differences in immune recycling cycles between the two subgroups, and the results showed an increased recruiting of CD4+ T, DC, and macrophages in the low TES group, accompanied by an increased promotion of antitumor immunity in step V (Figure 6B). We then assessed immune-related pathway activity using ssGSEA. The findings indicated that the low TES group had a considerable enrichment of most immune-related pathways (Figure 6C). Finally, GSEA results showed significant enrichment of cell adhesion, MAPK, NOTCH and VEGF signaling

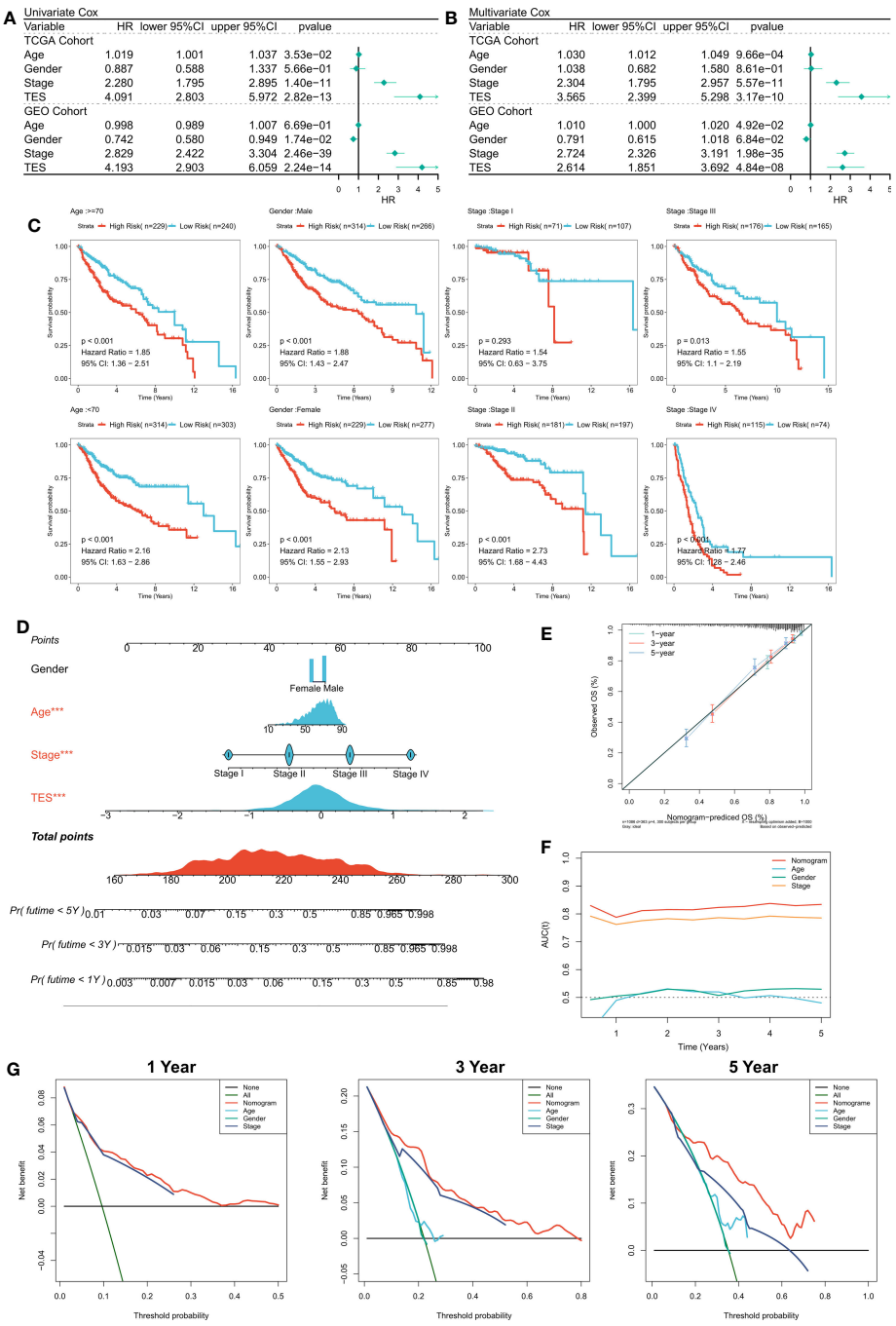


FIGURE 4
Verifying the independence and robustness of TES. **(A)** Univariate COX regression analysis of OS in TCGA and GEO datasets. **(B)** Multivariate COX regression analysis of OS in TCGA and GEO datasets. **(C)** The subgroup analysis of TES in the whole cohort. **(D)** Nomogram based on TES and clinical characteristics. **(E)** Calibration curve of Nomogram. **(F)** tROC curve of Nomogram and clinical characteristics. **(G)** The DCA curves of Nomogram and clinical characteristics at 1, 3, and 5 years.

pathways in the high TES group (Figure 6D). In contrast, the TCA cycle, fatty acid metabolism, protein export, and oxidative phosphorylation pathways were significantly enriched in the low TES group (Figure 6E). Therefore, our hypothesis was that CD8+ and CD4+ T cells enhanced anti-tumor immunity in the low TES group, but Tregs reduced anti-tumor immune responses in the high TES group.

Correlation of TES with genomic alterations

We then analyzed genome-wide data of the TCGA-COAD to decipher the genomic alteration status of different TES groups. The overall mutation profiles among the TES subgroups (including TMB, mutation signatures, SNP, and CNV) are shown in

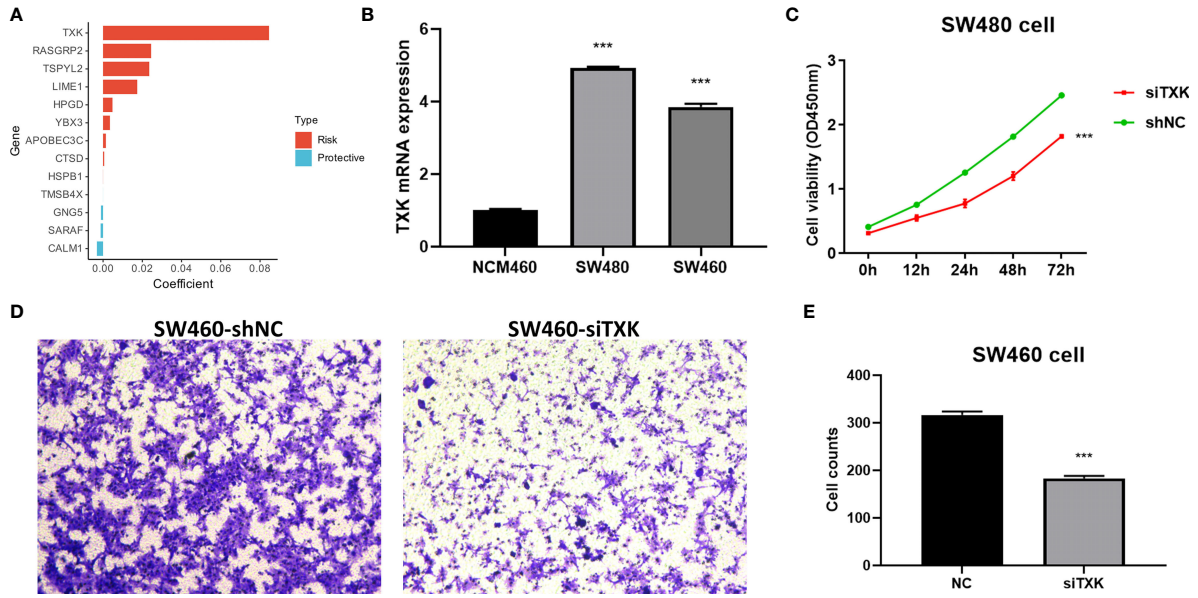


FIGURE 5
Cellular experiments to verify the malignancy of TXK. **(A)** Gene coefficients in the TES model showed that TXK was the most potent risk factor. **(B)** Differential mRNA expression levels of TXK in NCM460, SW480, and SW460 cell lines by qPCR. **(C)** Cell proliferation of SW480 cells transfected with TXK siRNA or siNC. **(D)** Transwell assay of invasive ability of SW460 cells transfected with TXK siRNA or siNC. **(E)** Cell counting of SW460 transfected with TXK siRNA or siNC in transwell cells. *** P<0.001.

Figure 7A. Between the two groups, we didn't detect any discernible differences in TMB (Figure 7B). In addition, we discovered no discernible variations in high frequency mutations across groupings except for TP53 (Figure 7A). We then detected no change in the total chromosomal amplification and deletion number between the two categories (Figures 7C, D). However, we discovered that the high TES group had greater gain and loss events in both arm and gene level (Figure 7A), while the low TES group had more deletions on 2p and 2q arms (Figure 7E).

Patients with low TES are more sensitive to chemotherapy

We proposed the hypothesis that TES might predict the response to chemotherapy in COAD patients given the disparities in biological function and CNV across different TES patients. On the basis of the GDSC database, we first assessed the IC50 of frequently used chemotherapeutic agents for COAD in various TES groups. The findings revealed that patients with low TES were more responsive to 5-Fluorouracil (Figures 8A, B). Contrarily, patients in the validation group with low TES were more responsive to Cisplatin and Camptothecin (Figure 8B). We examined the response of patients with different TES to chemotherapy in the TCGA dataset. The results showed that patients in the low TES group had a greater chance of complete and partial remission. In contrast, the proportion of patients with disease progression was higher in the high TES group (Figure 8C). Survival analysis showed better survival in the low-TES group of COAD patients receiving 5-FU and oxaliplatin, especially in those receiving Oxaliplatin (Figures 8D, F). In contrast, the survival difference between the different groups of COAD patients receiving Irinotecan was not significant, which may be due to the small sample size (Figure 8E). In conclusion, we speculate that patients with low TES are more suitable for treatment with 5-FU and platinum-based chemotherapeutic agents. Finally, we retrieved 20 small molecule compounds that may target TES through the Cmap database (Figure 8G)

Inferring immunotherapy response

We hypothesized that the low TES group would respond more strongly to immunotherapy because they have a more powerful antitumor immune response. First, we determined the unique IPS of each patient and found that individuals in the low TES group had greater IPS in both cohorts (Figures 9A, B). The TIDE algorithm was then used to predict the response of patients in the TCGA and GEO cohorts to immune checkpoint inhibitors, and the results showed that patients with low TES in both cohorts had a higher response rate to immunotherapy (Figures 9C, D). The efficacy of TES in the TCGA and GEO cohorts was only lower than MSI with reliable predictive efficacy compared to other indicators of immune efficacy (Figures 9E, F). Subsequently, we worked in two real-world immunotherapy cohorts (Imvigor210 and Liu David). The results showed significantly better survival in the low-TES group in both cohorts (Figures 9G, H). Moreover, patients in the low-TES group had higher remission rates in both cohorts (Figures 9I, J). Subsequently, we found no significant association between TES and neoantigens in both cohorts (Figures 9K, L). Although there was also no significant association between TES and TMB in both cohorts (Figures 9M, N), TMB was significantly higher in the low-TES group in the Liu David cohort (Figure 9N).

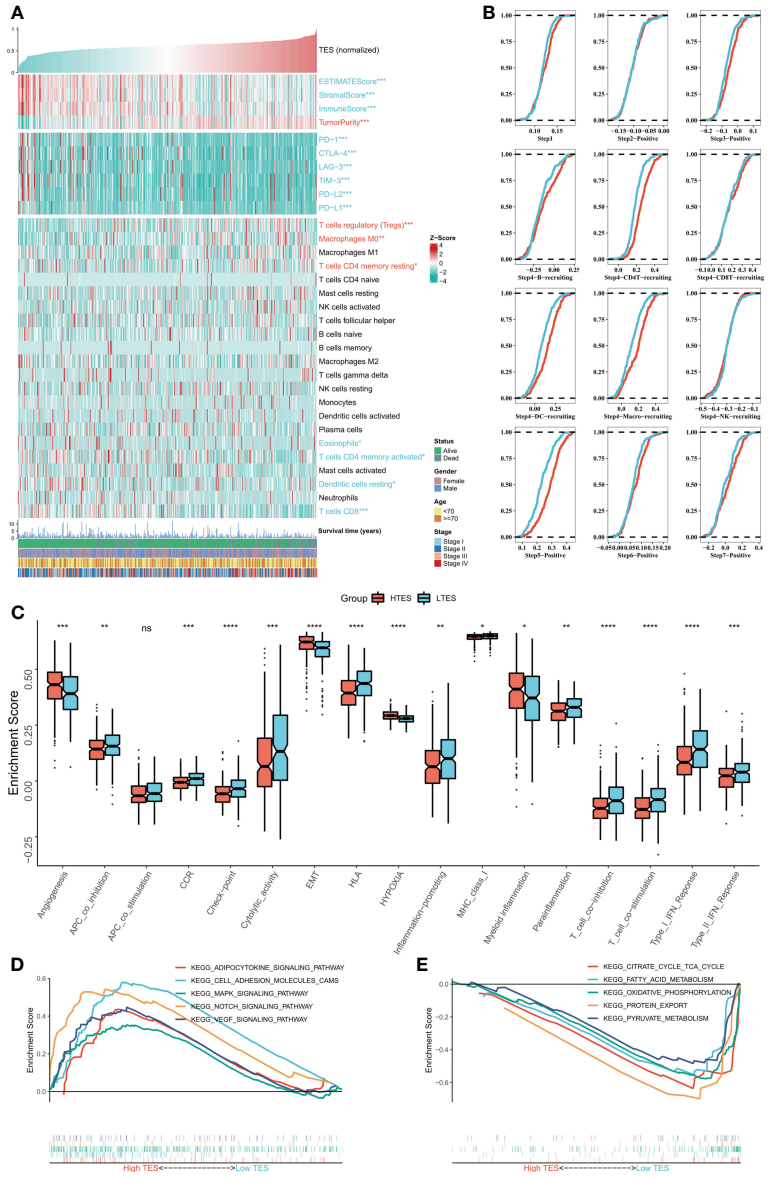


FIGURE 6
Dissecting the immune microenvironment of different TES groups. **(A)** The heat map shows the distribution of Estimate score, immune checkpoint expression and immune cell abundance among different TES groups in TCGA-COAD cohort. **(B)** Differences in tumor immune cycle among different TES groups in the TCGA-COAD cohort. **(C)** Box plots showing the immune-related pathway activity between different TES groups. *: P<0.05; **: P<0.01; ***: P<0.001; ****: P<0.0001. **(D)** GSEA analysis revealed 5 enriched pathways in the high TES group. **(E)** GSEA analysis revealed 5 enriched pathways in the low TES group. ns, not significant.

Discussion

Advanced COAD is extremely malignant and often has a poor prognosis due to treatment resistance (32). The immune system is involved in resisting the proliferation and invasion of malignant cells in COAD process, and T-cell depletion is one of the main causes of diminished antitumor immunity (33). In addition, targeted T-cell depletion is emerging in the field of cancer immunotherapy (11). To provide a new approach to the treatment and prognosis of advanced COAD, we aimed in this study to analyze the trajectory of T-cell depletion in COAD patients.

In this work, by analyzing scRNA-Seq data of TCGA-COAD, we differentiated cytolytic trajectories and exhaustion trajectories of CD8 T cells to comprehensively identify the major T-cell exhaustion indicators. We identified a total of 477 ordered genes for exhaustion trajectories and subsequently identified 27 effective T-cell exhaustion markers by one-way Cox regression. We observed a significant positive correlation between them, suggesting a potential mutual regulation between them. The primary transcriptome regulator for all core genes was CNV. we built a 13-gene T-cell depletion score (TES) model based on these 27 core genes using iterative LASSO regression. We first report the

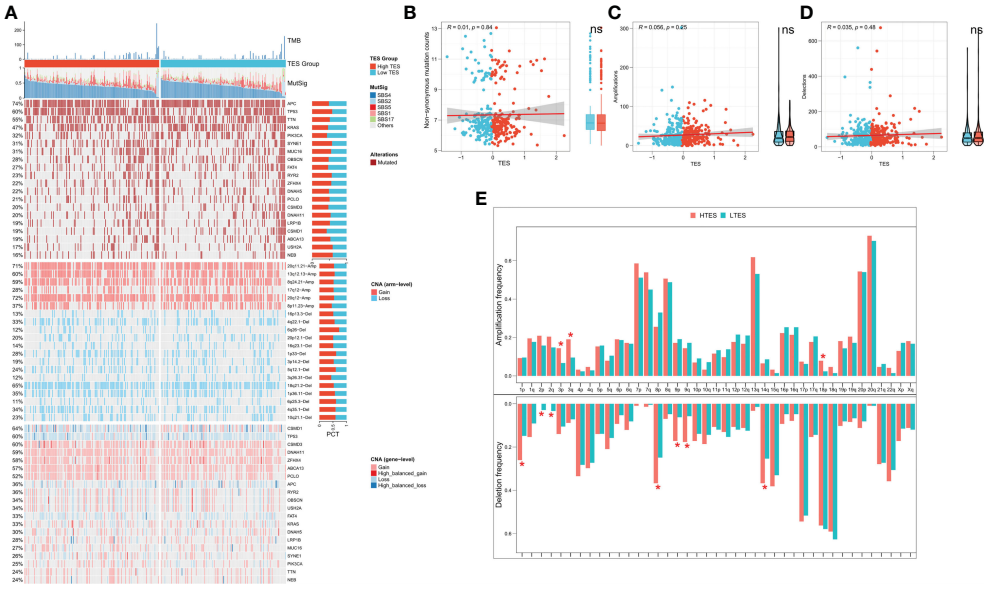


FIGURE 7
TES distinguishing genomic alteration patterns in COAD patients. **(A)** Genomic alterations landscape between different TES groups in the TCGA-COAD cohort, from top to bottom: TMB, mutational signatures, single-nucleotide mutations of top 20 driver mutated genes, CNVs of chromosomal segments, and CNVs in top 20 driver mutated genes. **(B)** Correlation of TES with Non-synonymous mutation counts. **(C)** Correlation between TES and total amplification number. **(D)** Correlation between TES and total deletion number. **(E)** CNV differences between different TES subgroups on the chromosome arms. *: P<0.05. ns, not significant.

prognostic efficacy of a systematic TES model for COAD patients compared to previously presented T-cell exhaustion markers (HAVCR2, ENTPD1, LAYN, and LAG3). We confirmed that TES is a strong prognostic indicator of OS in COAD patients and that TES works well in different COAD cohorts.

We next attempted to comprehend the molecular rationale behind TES, which is the key path of the dynamic processes of T cell differentiation in malignancies and is involved in the proliferation and spread of tumor cells. By comparing the two groups from multiple perspectives, including immune cell infiltration, immunological pathways, and immune checkpoints, we were able to explore in more detail the differences in the immune environment between the different TES groups. the results of ESTIMATE showed that the immune rating and Estimate composite score were higher in the low TES group while the tumor purity was higher in the high TES group. In addition, in the low TES group, we later found enhanced expression of six typical immune checkpoints, suggesting that patients with low TES may benefit more from treatment with immune checkpoint inhibitors (34). The tumor immune cycle system is characterized by most of the processes of antitumor immunity (35), and we found increased activity of the recruiting process of active immune cells (including T cells, DCs, and macrophages) as well as the positive regulatory processes of antitumor immunity in the low-TES group. Most immune-related pathways were also considerably elevated in the low TES group, supporting lower T-cell exhaustion in the low TES group, and pointing to a more potent and aggressive antitumor immune response in this group (35). Notably, there was a statistically significant increase in the number of Treg cells infiltrating the high TES group. This finding may have stifled the immune environment and antitumor immune response in patients with high TES, leading to a poorer prognosis (36).

Considering the significant relevance of genomic mutations for the course of tumor progression and treatment response, especially immunotherapy response, we then analyzed genome-wide data to explore the differentiation of genomic variation patterns of TES for individual patients. We were not able to detect significant TMB differences between TES subgroups. However, we found that the TP53 gene was significantly more frequently mutated in the high TES group. Previous studies have demonstrated that TP53 is the most frequently mutated tumor suppressor gene during tumor progression (37). Loss of function or dominant inactivation of wild type p53 is also frequently detected in patients with colon cancer (38, 39), which is consistent with our results. Our study indicates that increased mutations in TP53 may lead to a higher malignancy of tumor cells in the high TES group, resulting in a worse prognosis. Finally, we discovered more CNV occurrences in the group with high TES. Additionally, it has been shown that CNV has a significant role in the regulation of genes that affect drug response and metabolism, which in turn speeds up the development of anticancer drug resistance and results in treatment failure and disease recurrence (40, 41). As a result, we deduce that patients with low TES are suited for chemotherapy, whereas those with high TES are resistant to it. We confirmed the resistance to chemotherapy in patients with high TES through drug sensitivity data provided by the GDSC database. We found that low TES patients were more susceptible to 5-FU. In addition, survival analysis in the TCGA cohort also demonstrated an elevated remission rate for chemotherapy in patients with low TES, especially for oxaliplatin and 5-FU treatment. For high-risk COAD patients based on TES,

10.3389/fimmu.2023.1162843

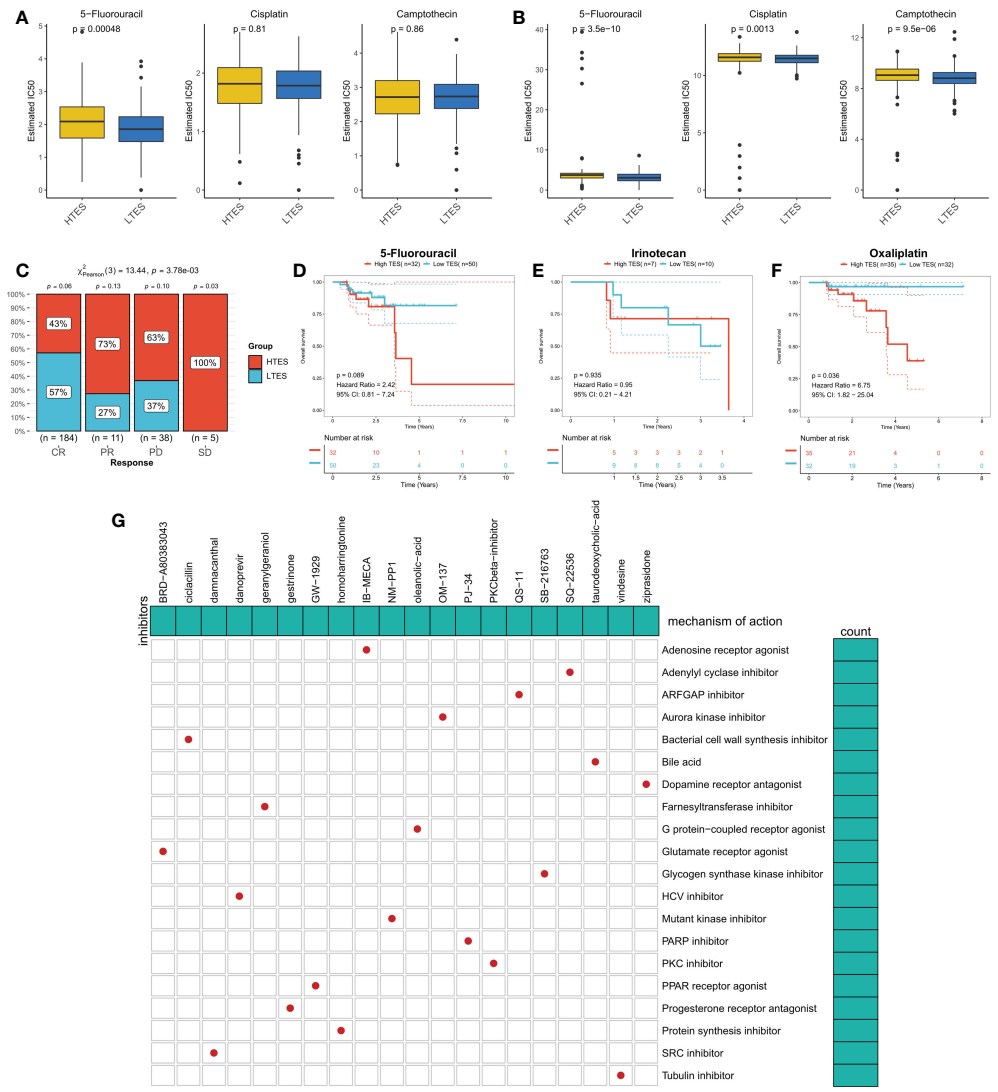


FIGURE 8
TES can predict chemotherapy. The IC50 values of the three commonly used drugs (5-Fluorouracil, Cisplatin, and Camptothecin) in the **(A)** TCGA cohort and **(B)** meta-GEO cohort were predicted based on the GDSC database. **(C)** Remission rates of different TES patients after receiving chemotherapy. **(D)** Kaplan-Meier survival curves for patients treated with 5-Fluorouracil in different TES groups. **(E)** Kaplan-Meier survival curves for patients treated with Irinotecan in different TES groups. **(F)** Kaplan-Meier survival curves for patients treated with Oxaliplatin in different TES groups. **(G)** Prediction of TES-related small molecule compounds as well as target pathway from the CMap database.

we also screened for potential therapeutic targets and found comparable small-molecule drugs. Finally, we identified the 20 most probable small molecule compounds.

Finnaly, we made a prediction that, from a variety of angles, people with low TES are more susceptible to immunotherapy. Moreover, IPS was greater among COAD patients with low TES, indicating that these patients would respond to immunotherapy better (28). The TIDE algorithm also demonstrated that individuals with low TES had greater rates of immune checkpoint inhibitor response (e.g., anti-PD-1, anti-PD-L1, and anti-CTLA-4) (29). Additionally, TES was more reliable than traditional predictors in predicting immunotherapy response. It is worth noting that the predictive efficacy of TES is not higher than that of MSI, which has been shown in numerous studies to be a reliable predictor of immunotherapy in colon cancer and is now being tested in clinical practice to assist in the prognosis of patients with COAD. Therefore, although TES does not show a leading advantage, it can be used as a complement to MSI in clinical practice (42–44).

In addition, two real-world cohorts used for validation confirmed our predicted results for immunotherapy sensitivity. In both the Imvigor210 and Liu David cohorts we found that the low TES group exhibited better survival rates. However, we did not find a significant association between TES and the number of detected neoantigens and TMB in these two cohorts. Previous studies have shown that TMB and neoantigens are indicators of a strong relationship with immunotherapy efficacy and can be used to predict benefits for patients. However, TES in our results exhibited independent predictive accuracy for immunotherapy. More insight into the specific regulatory mechanisms is needed in future studies (45, 46).

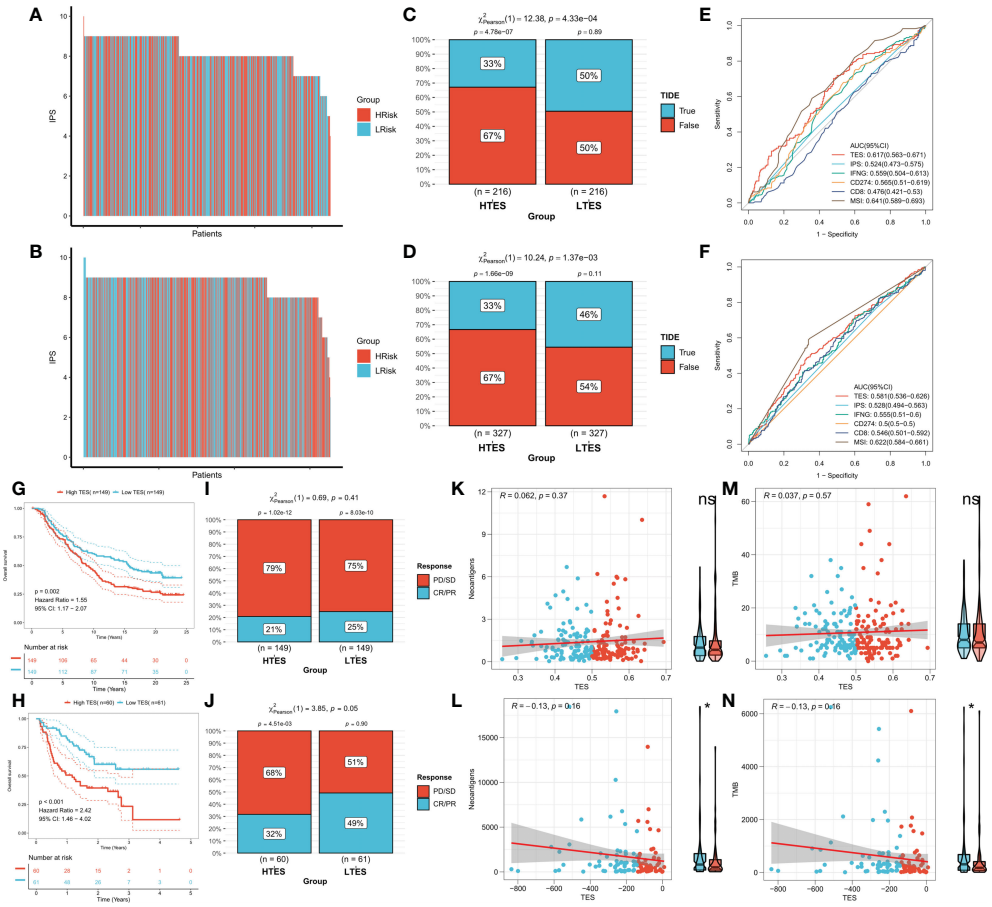


FIGURE 9
TES can predict immunotherapy. The IPS of individual COAD patients in the **(A)** TCGA cohort and **(B)** meta-GEO cohort. TIDE algorithm predicts response rates to immune checkpoint inhibitors for patients in the **(C)** TCGA cohort and **(D)** meta-GEO cohort. ROC curve shows the predictive efficiency for the response rate to immunotherapy by TES and other indicators in the **(E)** TCGA cohort and **(F)** meta-GEO cohort. **(G)** KM survival curves for patients in the high- and low-TES groups in IMvigor210 cohort. **(H)** KM survival curves for patients in the high- and low-TES groups in Liu David cohort. Remission rates of different TES patients after immunotherapy in the **(I)** IMvigor210 cohort and **(J)** Liu David cohort. Scatter plot and box plot show the correlation of TES with **(K)** neoantigens; **(M)** TMB in IMvigor210 cohort. Scatter plot and box plot show the correlation of TES with **(L)** neoantigens; **(N)** TMB in Liu David cohort. *: P<0.05. ns, not significant.

This study still contains some limitations. The study only contains two non-COAD immunotherapy RNA-seq data, which is due to the scarcity of data in this field. We hope to collect more immunotherapy sequences or platform data for COAD in the future to further validate the predictive accuracy of TES for immunotherapy. In addition, genomic regulation is a large field, and we have only focused on a portion of mRNAs and may have neglected data from some other regulatory genomes. Finally, the mechanism of how TES affects biological function as well as the phenotype is unclear. However, we synthesized the results of functional enrichment analysis to make reasonable speculations, which is an inspiration for future mechanistic studies.

Conclusions

In this study, we identified possible depleted CD8+ T cell differentiation trajectories in COAD patients and developed a TES model to quantify the level of T cell depletion in the tumor microenvironment. Patients with lower TES responded more strongly to chemotherapy and immunotherapy and had a better prognosis. This finding not only advances the development of cancer genetics and immunotherapy but also provides new perspectives on the clinical treatment of colon cancer and innovative immunotherapy strategies.

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 within the article/supplementary materials.

Author contributions

XS performed the data analyses and wrote the manuscript. XZ contributed significantly to analysis and manuscript preparation. LL

helped perform the analysis with constructive discussions. BL and LW contributed to the conception of the study. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Sichuan Provincial Science and Technology Department Project (Foundation of Science and Technology Innovation Seedling Project) (No.2022YFS0169).

Acknowledgments

We greatly appreciate the analytical data provided by the TCGA and GEO databases.

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: Cancer J Clin* (2021) 71(1):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(3):209–49. doi: 10.3322/caac.21660

3. Jung F, Lee M, Doshi S, Zhao G, Cheung KLT, Chesney T, et al. Neoadjuvant therapy versus direct to surgery for T4 colon cancer: meta-analysis. *Br J surg* (2021) 109 (1):30–6. doi: 10.1093/bjs/znab382

4. Chau I, Cunningham D. Adjuvant therapy in colon cancer–what, when and how? *Ann Oncol* (2006) 17(9):1347–59. doi: 10.1093/annonc/mdl029

5. Aldea M, Andre F, Marabelle A, Dogan S, Barlesi F, Soria JC. Overcoming resistance to tumor-targeted and immune-targeted therapies. *Cancer discov* (2021) 11 (4):874–99. doi: 10.1158/2159-8290.Cd-20-1638

6. Kanani A, Veen T, Søreide K. Neoadjuvant immunotherapy in primary and metastatic colorectal cancer. *Br J surg* (2021) 108(12):1417–25. doi: 10.1093/bjs/znab342

7. Pauken KE, Wherry EJ. Overcoming T cell exhaustion in infection and cancer. *Trends Immunol* (2015) 36(4):265–76. doi: 10.1016/j.it.2015.02.008

8. Blank CU, Haining WN, Held W, Hogan PG, Kallies A, Lugli E, et al. Defining 'T cell exhaustion'. *Nat Rev Immunol* (2019) 19(11):665–74. doi: 10.1038/s41577-019-0221-9

9. Liu B, Hu X, Feng K, Gao R, Xue Z, Zhang S, et al. Temporal single-cell tracing reveals clonal revival and expansion of precursor exhausted T cells during anti-PD-1 therapy in lung cancer. *Nat cancer* (2022) 3(1):108–21. doi: 10.1038/s43018-021-00292-8

10. McLane LM, Abdel-Hakeem MS, Wherry EJ. CD8 T cell exhaustion during chronic viral infection and cancer. *Annu Rev Immunol* (2019) 37:457–95. doi: 10.1146/annurev-immunol-041015-055318

11. Chow A, Perica K, Klebanoff CA, Wolchok JD. Clinical implications of T cell exhaustion for cancer immunotherapy. *Nat Rev Clin Oncol* (2022) 27(4):450–61. doi: 10.1038/s41571-022-00689-z

12. Peng DH, Rodriguez BL, Diao L, Chen L, Wang J, Byers LA, et al. Collagen promotes anti-PD-1/PD-L1 resistance in cancer through LAIR1-dependent CD8(+) T cell exhaustion. *Nat Commun* (2020) 11(1):4520. doi: 10.1038/s41467-020-18298-8

13. Guo Y, Xie YQ, Gao M, Zhao Y, Franco F, Wenes M, et al. Metabolic reprogramming of terminally exhausted CD8(+) T cells by IL-10 enhances anti-tumor immunity. *Nat Immunol* (2021) 22(6):746–56. doi: 10.1038/s41590-021-00940-2

14. Sun K, Xu R, Ma F, Yang N, Li Y, Sun X, et al. scRNA-seq of gastric tumor shows complex intercellular interaction with an alternative T cell exhaustion trajectory. *Nat Commun* (2022) 13(1):4943. doi: 10.1038/s41467-022-32627-z

15. Thommen DS, Schumacher TN. T Cell dysfunction in cancer. *Cancer Cell* (2018) 33(4):547–62. doi: 10.1016/j.ccell.2018.03.012

16. Tirosh I, Izar B, Prakadan SM, Wadsworth MH2nd, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Sci (New York NY).* (2016) 352(6282):189–96. doi: 10.1126/science.aad0501

17. Doering TA, Crawford A, Angelosanto JM, Paley MA, Ziegler CG, Wherry EJ. Network analysis reveals centrally connected genes and pathways involved in CD8+ T cell exhaustion versus memory. *Immunity* (2012) 37(6):1130–44. doi: 10.1016/j.immuni.2012.08.021

18. Jorissen RN, Gibbs P, Christie M, Prakash S, Lipton L, Desai J, et al. Metastasis-associated gene expression changes predict poor outcomes in patients with dukes stage b and c colorectal cancer. *Clin Cancer Res an Off J Am Assoc Cancer Res* (2009) 15 (24):7642–51. doi: 10.1158/1078-0432.Ccr-09-1431

19. Chen MS, Lo YH, Chen X, Williams CS, Donnelly JM, Criss ZK2nd, et al. Growth factor-independent 1 is a tumor suppressor gene in colorectal cancer. *Mol Cancer Res MCR.* (2019) 17(3):697–708. doi: 10.1158/1541-7786.Mcr-18-0666

20. Martin ML, Zeng Z, Adileh M, Jacobo A, Li C, Vakiani E, et al. Logarithmic expansion of LGR5(+) cells in human colorectal cancer. *Cell signal* (2018) 42:97–105. doi: 10.1016/j.cellsig.2017.09.018

21. Zhang L, Li Z, Skrzypczynska KM, Fang Q, Zhang W, O'Brien SA, et al. Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. *Cell* (2020) 181(2):442–59.e29. doi: 10.1016/j.cell.2020.03.048

22. Schröder MS, Culhane AC, Quackenbush J, Haibe-Kains B. Survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinf (Oxford England).* (2011) 27(22):3206–8. doi: 10.1093/bioinformatics/btr511

23. 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 (5):453–7. doi: 10.1038/nmeth.3337

24. Yoshihara K, Shahmoradgoli M, Martínez 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

25. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* (2018) 28(11):1747–56. doi: 10.1101/gr.239244.118

26. Wang S, Li H, Song M, Tao Z, Wu T, He Z, et al. Copy number signature analysis tool and its application in prostate cancer reveals distinct mutational processes and clinical outcomes. *PloS Genet* (2021) 17(5):e1009557. doi: 10.1371/journal.pgen.1009557

27. 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(9):e107468. doi: 10.1371/journal.pone.0107468

28. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep* (2017) 18 (1):248–62. doi: 10.1016/j.celrep.2016.12.019

29. 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(10):1550–8. doi: 10.1038/s41591-018-0136-1

30. 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 (London England).* (2018) 391(10122):748–57. doi: 10.1016/s0140-6736(17)33297-x

31. Liu D, Schilling B, Liu D, Sucker A, Livingstone E, Jerby-Arnon L, et al. Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nat Med* (2019) 25(12):1916–27. doi: 10.1038/s41591-019-0654-5

32. Hu T, Li Z, Gao CY, Cho CH. Mechanisms of drug resistance in colon cancer and its therapeutic strategies. *World J gastroenterol* (2016) 22(30):6876–89. doi: 10.3748/wjg.v22.i30.6876

33. Jaaks P, Coker EA, Vis DJ, Edwards O, Carpenter EF, Leto SM, et al. Effective drug combinations in breast, colon and pancreatic cancer cells. *Nature* (2022) 603 (7899):166–73. doi: 10.1038/s41586-022-04437-2

34. Topalian SL, Drake CG, Pardoll DM. Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell* (2015) 27(4):450–61. doi: 10.1016/j.ccell.2015.03.001

35. Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* (2013) 39(1):1–10. doi: 10.1016/j.immuni.2013.07.012

36. Nishikawa H, Koyama S. Mechanisms of regulatory T cell infiltration in tumors: implications for innovative immune precision therapies. *J immunother Cancer* (2021) 9 (7):e002591. doi: 10.1136/jitc-2021-002591

37. Giacomelli AO, Yang X, Lintner RE, McFarland JM, Duby M, Kim J, et al. Mutational processes shape the landscape of TP53 mutations in human cancer. *Nat Genet* (2018) 50(10):1381–7. doi: 10.1038/s41588-018-0204-y

38. Warren RS, Atreya CE, Niedzwiecki D, Weinberg VK, Donner DB, Mayer RJ, et al. Association of TP53 mutational status and gender with survival after adjuvant treatment for stage III colon cancer: results of CALGB 89803. *Clin Cancer Res* (2013) 19 (20):5777–87. doi: 10.1158/1078-0432.Ccr-13-0351

39. Kandioler D, Mittlböck M, Kappel S, Puhalla H, Herbst F, Langner C, et al. TP53 mutational status and prediction of benefit from adjuvant 5-fluorouracil in stage III colon cancer patients. *EBioMedicine* (2015) 2(8):825–30. doi: 10.1016/j.ebiom.2015.06.003

40. He Y, Hoskins JM, McLeod HL. Copy number variants in pharmacogenetic genes. *Trends Mol Med* (2011) 17(5):244–51. doi: 10.1016/j.molmed.2011.01.007

41. Sansregret L, Vanhaesebroeck B, Swanton C. Determinants and clinical implications of chromosomal instability in cancer. *Nat Rev Clin Oncol* (2018) 15 (3):139–50. doi: 10.1038/nrclinonc.2017.198

42. Llosa NJ, Cruise M, Tam A, Wicks EC, Hechenbleikner EM, Taube JM, et al. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. *Cancer discov* (2015) 5(1):43–51. doi: 10.1158/2159-8290.Cd-14-0863

43. Hou W, Yi C, Zhu H. Predictive biomarkers of colon cancer immunotherapy: present and future. *Front Immunol* (2022) 13:1032314. doi: 10.3389/fimmu.2022.1032314

44. Bao X, Zhang H, Wu W, Cheng S, Dai X, Zhu X, et al. Analysis of the molecular nature associated with microsatellite status in colon cancer identifies clinical implications for immunotherapy. *J immunother Cancer* (2020) 8(2):e001437. doi: 10.1136/jitc-2020-001437

45. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Sci (New York NY).* (2015) 348(6230):69–74. doi: 10.1126/science.aaa4971

46. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol* (2019) 30(1):44–56. doi: 10.1093/annonc/mdy495

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 05 May 2023
DOI 10.3389/fimmu.2023.1138203

OPEN ACCESS

EDITED BY
Nan Zhang,
Harbin Medical University, China

REVIEWED BY
Ju Wang,
Tianjin Medical University, China
Haijun Han,
Zhejiang University, China
Kun Zhao,
The First Affiliated Hospital of Sun Yat-sen University, China

*CORRESPONDENCE
Hui Liu
liuhui@wmu.edu.cn
Quanwei Zhou
1091040044@qq.com

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

RECEIVED 05 January 2023
ACCEPTED 24 April 2023
PUBLISHED 05 May 2023

CITATION
Li S, Guo Y, Hu H, Gao N, Yan X, Zhou Q and Liu H (2023) TANK shapes an immunosuppressive microenvironment and predicts prognosis and therapeutic response in glioma.
Front. Immunol. 14:1138203.
doi: 10.3389/fimmu.2023.1138203

TANK shapes an immunosuppressive microenvironment and predicts prognosis and therapeutic response in glioma

Shasha Li[1,2†], Youwei Guo[3†], Huijuan Hu[1,4], Na Gao[1,4], Xuejun Yan[5], Quanwei Zhou[6*] and Hui Liu[1,4*]

[1]National Engineering Research Center of Ophthalmology and Optometry, Eye Hospital, Wenzhou Medical University, Wenzhou, China, [2]Oujiang Laboratory (Zhejiang Lab for Regenerative Medicine, Vision and Brain Health), Wenzhou, China, [3]Department of Neurosurgery, Xiangya Hospital, Central South University, Changsha, China, [4]State Key Laboratory of Ophthalmology, Optometry and Visual Science, Eye Hospital, Wenzhou Medical University, Wenzhou, China, [5]Department of Geriatrics, National Key Clinical Specialty, Guangzhou First People's Hospital, Guangzhou Medical University, Guangzhou, China, [6]The National Key Clinical Specialty, Department of Neurosurgery, Zhujiang Hospital, Southern Medical University, Guangzhou, China

Background: Glioma, the most prevalent malignant intracranial tumor, poses a significant threat to patients due to its high morbidity and mortality rates, but its prognostic indicators remain inaccurate. Although TRAF-associated NF-kB activator (TANK) interacts and cross-regulates with cytokines and microenvironmental immune cells, it is unclear whether TANK plays a role in the immunologically heterogeneous gliomas.

Methods: TANK mRNA expression patterns in public databases were analyzed, and qPCR and IHC were performed in an in-house cohort to confirm the clinical significance of TANK. Then, we systematically evaluated the relationship between TANK expression and immune characteristics in the glioma microenvironment. Additionally, we evaluated the ability of TANK to predict treatment response in glioma. TANK-associated risk scores were developed by LASSO-Cox regression and machine learning, and their prognostic ability was tested.

Results: TANK was specifically overexpressed in glioma and enriched in the malignant phenotype, and its overexpression was related to poor prognosis. The presence of a tumor microenvironment that is immunosuppressive was evident by the negative correlations between TANK expression and immunomodulators, steps in the cancer immunity cycle, and immune checkpoints. Notably, treatment for cancer may be more effective when immunotherapy is combined with anti-TANK therapy. Prognosis could be accurately predicted by the TANK-related risk score.

Conclusions: High expression of TANK is associated with the malignant phenotype of glioma, as it shapes an immunosuppressive tumor microenvironment. Additionally, TANK can be used as a predictive biomarker for responses to various treatments and prognosis.

KEYWORDS

glioma, immunosuppressive microenvironment, immune infiltration, TANK, prognosis

Introduction

Glioma is the most common malignant intracranial tumor, and more than 60% of primary central nervous system (CNS) tumors (1) with a diffusely invasive nature (2). The median overall survival time for patients with gliomas ranges from ~14 months (3) to less than 10 years (4), giving rise to severe morbidity and mortality of patients. Over the past decade, remarkable advances in molecular profiling studies (5, 6) have deepened the understanding of the classification, diagnosis, treatment, and prognosis of glioma. As such, the 2016 WHO classification of glioma incorporated morphological and molecular features for division of gliomas into distinct subgroups (7) for precision diagnosis and treatment. In 2021, the latest version of the WHO classification subdivided diffuse glioma into adult-type and pediatric-type (8), with adult-type comprising three subtypes characterized by histological features and genetic mutation status. The current standard therapeutic strategies are surgical resection, radiotherapy, and chemotherapy, whereas combined therapy demonstrates modest efficacy. Physiological barriers, chemo- and radioresistance, and the paucity of clear targeted pathways contribute to the limitations of these treatments (9, 10). Consequently, novel treatment modalities focused on improving the life expectancy of glioma patients are urgently needed.

Accompanied by a deeper understanding of glioma biology, numerous preclinical and clinical trials have explored immunotherapies, such as immune checkpoint inhibitors (ICIs) (11), oncolytic viral therapies (12), adoptive cellular therapies (13), cytokine therapies (14) and vaccinations (15). The immune checkpoint pathway, integral to modulating self-tolerance and immune responses, serves as a major mechanism by which glioma escapes immunosurveillance and maintains immune resistance (16). ICIs restore tumoricidal activities by targeting coinhibitory molecules such as CTLA-4 and PD-1and have demonstrated encouraging efficacy in clinical trials of metastatic melanoma (17), Hodgkin's lymphoma (18), HCC (19), and NSCLC (20). Various preclinical trials of ICIs or ICIs combined with other strategies in patients with glioma have been explored and found to have potential clinical value (21–24). Considering the paucity of strong clinical evidence and the exact mechanisms of action, the current state of knowledge emphasizes the urgency of exploring primary therapeutic approaches in combination with novel therapeutic targets to prolong the survival of glioma patients.

TRAF-associated NF-kB activator (TANK) is a protein with dual functions in activating NF-kB (25–27) that is indispensable for immune responses and inflammatory processes, as well as for activating survival and proinflammatory genes within the tumor microenvironment (28–30). Subsequent studies identified TANK as an adaptor protein that interacts with canonical IKKs (NEMO and IKKγ) (31) and IKK-related kinases (TBK1 and IKKε) (32) to modulate NF-κB and TLR-induced antiviral pathways and prevent autoimmunity (32, 33). Consequently, the immunological functions that TANK perform in the pathophysiological process of these diseases, especially cancers, deserve in-depth discussion.

Downregulated genes, including TANK/I-TRAF, were analyzed in HPV-16 E6-transfected carcinoma cells (34), whereas treatment with the antiproliferative agent cisplatin reversed the condition and downregulated the TRAF2-mediated NF-κB activity (35), indicating the tumorigenic properties of TANK and novel therapeutic targets for cervical cancer. Deregulated expression of TANK not only orchestrates the signaling network of the ERK1/2, AKT and IRF3 pathways in controlling the survival, proliferation, migration and invasion of glioblastoma (GBM) cells but also mediates the relative expression of genes in inflammatory signaling cascades (36). Moreover, TANK indirectly phosphorylates the transcription factor STAT3 to increase the release of interleukin-6 (IL-6) and ultimately accelerate the progression of glioma in terms of enhanced angiogenesis and proliferation. TANK interacts with and cross-regulates cytokines and microenvironmental immune cells, but its exact biological function remains uncertain and needs to be further investigated.

Materials and methods

Obtaining and processing data

The methods used for obtaining and processing data are the same as those described in previous literature (37). All data were downloaded from Chinese Glioma Genome Atlas (CGGA) datasets (CGGA-693, CGGA-325, CGGA-301), TCGA, GSE16011 and Rembrandt datasets.

Human specimens

We retrospectively defined two cohorts from Xiangya Hospital, Central South University. Cohort 1 included 29 normal tissues and 200 glioma tissues for examining the mRNA expression of TANK

by qPCR. Cohort 2 included 23 normal tissues and 203 glioma tissues for examining the protein expression of TANK by IHC. The relevant information can be seen in Table 1. Informed consent was obtained from all patients. Ethical approval was obtained for this study.

RNA extraction and quantitative real-time PCR

We carried out these processes by referring to the previous study (37). The relative expression of TANK was calculated to ACTB by $2^{-\Delta\Delta Ct}$ method. Primer sequences are given below:

TANK (F) 5′- CCACTTCTGGACCCATCTGATG-3′,
TANK (R) 5′- GCAGTTCTGAGTCTGTGCCACT-3′,
ACTB (F) 5′-ACAGAGCCTCGCCTTTGCCGAT-3′,
ACTB (R) 5′- CTTGCACATGCCGGAGCCGTT-3′.

Immunohistochemistry (IHC)

A TMA was constructed from 23 normal tissues and 203 glioma tissues. With reference to the previous literature (37), we conducted experiments. IHC assay was performed with primary antibodies against PD-1 (Proteintech, China), HIF1A (CST, United States), CD11b (AiFang, China), CD40 (Proteintech, China), PD-L1 (CST, United States), CD163 (Proteintech, China), STAT3 (Proteintech, China), and TANK (Bioss, China).

Immunological characteristics of the glioma microenvironment

The immunological characteristics of the glioma microenvironment were evaluated by considering the expression levels of immunomodulators and infiltration levels of TIICs, the activity of the tumor immunity cycle, and inhibitory immune checkpoints. Date for a total of 122 immunomodulators, such as MHC, chemokines and immune stimulators, were obtained from previous studies (38). By single sample gene set enrichment analysis (ssGSEA), the seven steps of the tumor immunity cycle were assessed based on the gene expression profiles (39). Seven algorithms, including MCP-counter, CIBERSORT, quanTIseq, ssGSEA, xCELL, TIMER and TIP, were used to calculate the abundances of TIICs in the tumor microenvironment. Based on Auslander's study, we identified 22 immunosuppressive checkpoints with potential for therapeutic intervention (40). The T-cell inflammation score was calculated based on the mRNA expression of 18 genes. Immune and stromal scores were evaluated with the ESTIMATE R package (41).

Identification and functional enrichment analysis of differentially expressed genes (DEGs)

The median expression level of TANK was used as the cutoff for dividing all patients into the high and low TANK expression groups. With the limma R package, TANK-related DEGs between the two groups (42) in the TCGA and CGGA-693 cohorts were identified. Adjusted $P < 0.05$ and |log (fold change) |>1 were considered the criteria for identifying DEGs. Gene set enrichment analysis (GSEA) was performed with GSEA software (vision 3.0) (http://www.broadinstitute.org/gsea) to explore the potential mechanism of TANK.

Development and validation of a TANK-associated risk score by LASSO and machine learning

In the TCGA and CGGA-693 cohorts, univariate Cox regression analysis of the DEGs was performed using the survival R package. The TANK-associated prognostic model was established by using the LASSO procedure to identify 13 prognostic markers from among the 347 TANK-related DEGs significantly associated with prognosis by the R package "glmnet" in the TCGA cohort.

The root mean squared error (RMSE) was a measure of how well the machine learns the model, and was calculated by taking the square root of the average of the residuals (errors not explained by the regression equation) over the total sample size. The support vector machine (SVM) model selection of hyper-parameters was made based on lowest RMSE values (43). Using SVM regression, 10 genes were selected from 13 prognostic markers.

Individual risk scores were calculated based on the Cox regression coefficient (β) and mRNA expression levels. The training and validation sets were divided in the TCGA cohort at a ratio of 7:3. An R package was used to assess the statistical performance of the prognostic model. Additionally, the TANK-associated risk score was validated as a prognostic indicator and performer in the TCGA internal validation set, TCGA-all set, CGGA-693 set, CGGA-301 set, CGGA-325 set, GSE16011 dataset, and Rembrandt dataset.

Statistical analysis

Pearson or Spearman correlation analysis was performed to investigate correlations between variables. The *t* test was used to compare continuous variables fitting a normal distribution between binary groups. Kaplan-Meier curves were generated for prognostic analyses based on categorical variables, and the log-rank test was used to estimate statistical significance. $P < 0.05$ was used as the criterion for significance, and all tests were two-sided.

Results

The expression pattern of TANK

By integrating GTEx data with TCGA data, we were able to increase the number of normal tissue samples. TANK levels in various tumor tissues, including LGG and GBM tissues, were markedly higher than those in nontumor tissues ($P < 0.05$,

10.3389/fimmu.2023.1138203

TABLE 1 Clinical characteristics of Cohort 1 and Cohort 2.

		Cohort 1 (n=200)	Cohort 2 (n=203)
Age (Mean ± SD)		43.98 ± 15.78	45.02 ± 15.90
Gender, n (%)	Female	83 (41.5%)	87 (42.9%)
	Male	112 (56%)	116 (57.1%)
	Unknown	5 (2.5%)	0 (0%)
WHO grade, n (%)	2	57 (28.5%)	52 (25.6%)
	3	36 (18%)	42 (20.7%)
	4	84 (42%)	95 (46.8%)
	Unknown	23 (11.5%)	14 (6.9%)
IDH1 status, n (%)	Mutant	67 (33.5%)	71 (35%)
	Unknown	22 (11%)	23 (11.3%)
	Wild-type	111 (55.5%)	109 (53.7%)
Histology, n (%)	Astrocytoma	74 (37%)	85 (41.9%)
	Gangliocytoma	5 (2.5%)	5 (2.5%)
	GBM	82 (41%)	96 (47.3%)
	Oligodendroglioma	24 (12%)	17 (8.4%)
	Unknown	15 (7.5%)	0 (0%)
Radiotherapy, n (%)	No	50 (25%)	55 (27.1%)
	Unknown	54 (27%)	30 (14.8%)
	Yes	96 (48%)	118 (58.1%)
Chemotherapy, n (%)	No	44 (22%)	55 (27.1%)
	Unknown	54 (27%)	30 (14.8%)
	Yes	102 (51%)	118 (58.1%)
Laterality, n (%)	Both	5 (2.5%)	6 (3%)
	Left	86 (43%)	89 (43.8%)
	Middle	10 (5%)	7 (3.4%)
	Right	96 (48%)	101 (49.8%)
	Unknown	3 (1.5%)	0 (0%)
Tumor location, n (%)	Brainstem	3 (1.5%)	3 (1.5%)
	Cerebellar	10 (5%)	7 (3.4%)
	Frontal	87 (43.5%)	93 (45.8%)
	Insular	6 (3%)	5 (2.5%)
	Occipital	11 (5.5%)	14 (6.9%)
	Parietal	18 (9%)	22 (10.8%)
	Sellar	3 (1.5%)	2 (1%)
	Temporal	55 (27.5%)	54 (26.6%)
	Thalamus	4 (2%)	3 (1.5%)
	Unknown	3 (1.5%)	0 (0%)

Figure 1A). Analysis of the TCGA cohort showed higher TANK expression in tumor tissues than in normal brain tissues, and analyses of the GSE16011 and Rembrandt cohorts validated this observation ($P < 0.001$, Figures 1B–D). Additionally, we examined the expression pattern of TANK in four cohorts of patients with glioma. In the TCGA cohort, higher grade glioma tissues expressed significantly higher levels of TANK than lower grade glioma tissues ($P < 0.001$, Figure S1A). Results similar to those observed in the CGGA-693, Rembrandt and GSE16011 datasets were also observed ($P < 0.05$, Figures S1B–D). TANK expression was higher in gliomas with wild-type IDH than in those with mutant IDH in the three cohorts ($P < 0.05$, Figures S1E–G). TANK was generally highly expressed in GBM ($P < 0.05$, Figures S1H–K). We examined TANK expression in 29 normal tissues and 200 fresh tumor tissues, as well as in 27 paired tumor samples and peritumor tissues by qPCR. In the in-house cohort, TANK expression was high in glioma tissues ($P < 0.001$, Figure 1E). Furthermore, the tumor tissues exhibited significantly higher levels of TANK expression than the matched peritumor tissues ($P < 0.001$, Figure 1F). High expression of TANK was found at a significantly higher rate in WHO grade IV gliomas, wild-type IDH1 gliomas, and GBM ($P < 0.001$, Table 2), consistent with the above results. Immunohistochemical analysis of the tissue microarrays

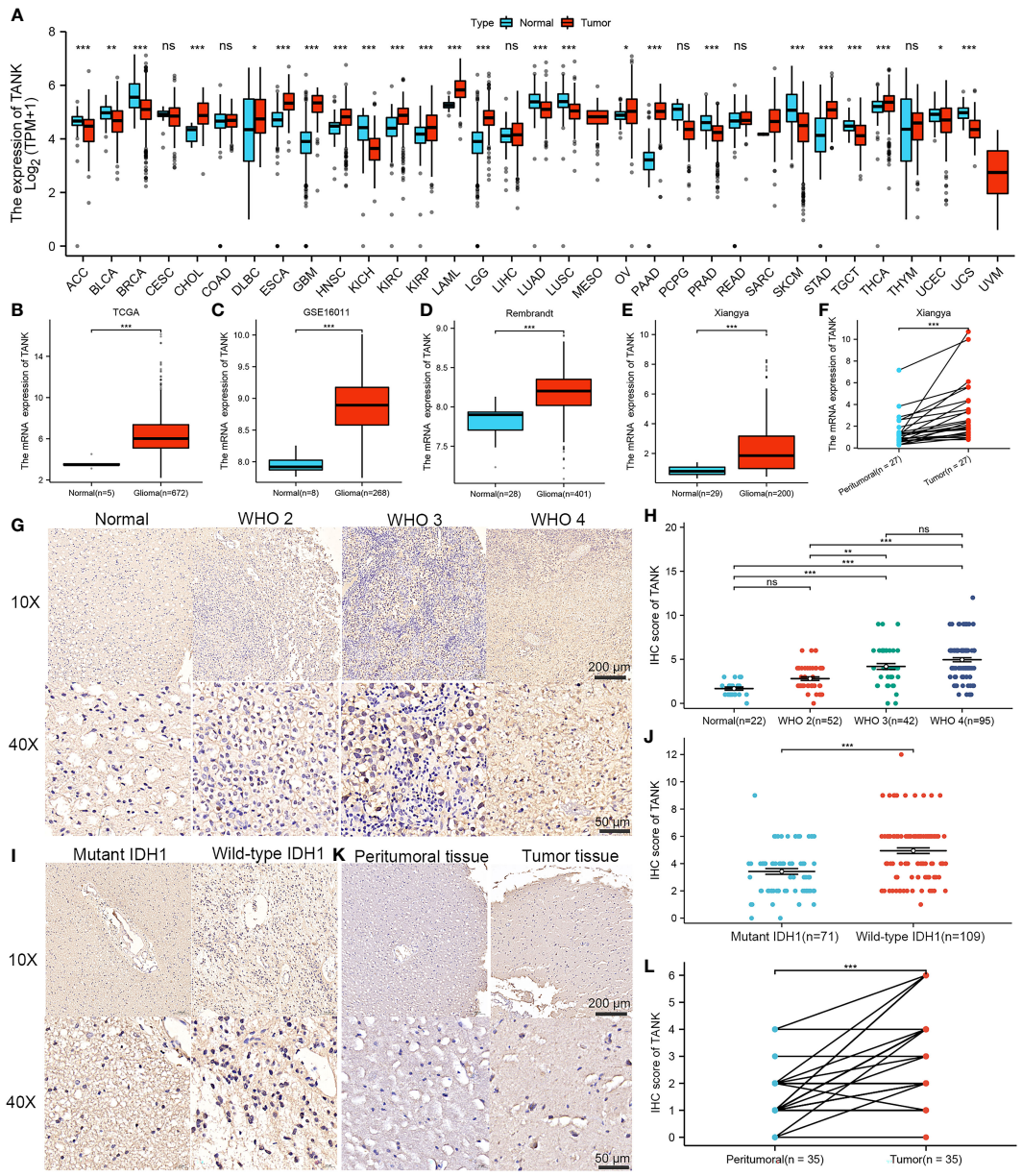


FIGURE 1
Elevated expression of TANK in glioma. **(A)** Differential expression of TANK between tumor and normal tissues from the TCGA dataset; **(B-E)** The expression level of TANK in the normal tissues and glioma tissues in the TCGA cohort **(B)**, GSE16011 **(C)**, Rembrandt cohort **(D)** and Xiangya cohort **(E)**; **(F)** The expression of TANK in glioma and peritumor tissues was analyzed by RT-qPCR; **(G-J)** The expression level of TANK in gliomas with different WHO grades **(G, H)**, and wild-type and mutant IDH1 **(I, J)** was analyzed by immunohistochemical staining; **(K, L)** The expression of TANK in glioma and peritumor tissues was analyzed by IHC (-no significance, *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$).

revealed that TANK was upregulated in gliomas ($P < 0.05$, Figures 1G, H). TANK was significantly enriched in glioma with WHO grade IV, wild-type IDH1, and GBM ($P < 0.05$, Figures 1G–J) ($P < 0.05$, Table 2). Furthermore, TANK was significantly overexpressed in tumor tissues in 35 tumor-peritumor tissue pairs ($P < 0.001$, Figures 1K, L). Thus, higher TANK expression is associated with more malignant glioma phenotypes.

TANK is an indicator of poor prognosis in glioma

Survival curves were used to explore the prognostic implications of TANK expression in glioma in the six cohorts, and the results consistently revealed that TANK is an indicator of poor prognosis in glioma (TCGA, HR=2.85 (2.16-3.75); CGGA-693, HR=1.57

TABLE 2 Association of TANK expression with clinical parameters in two cohorts.

Characteristic	Cohort1 (n=200)		*P*	Cohort2 (n=203)		*P*
	Low (n=100)	High (n=100)		Low (n=128)	High (n=75)	
Age, Mean ± SD	42.95 ± 15.20	45.02 ± 16.34	0.212	43.91 ± 15.19	46.91 ± 16.98	0.107
Gender, n (%)			0.789			0.062
Female	43 (21.5%)	40 (20%)		48 (23.6%)	39 (19.2%)	
Male	54 (27%)	58 (29%)		80 (39.4%)	36 (17.7%)	
Unknown	3 (1.5%)	2 (1%)		0 (0%)	0 (0%)	
WHO grade, n (%)			0.010			< 0.001
2	37 (18.5%)	20 (10%)		48 (23.6%)	4 (2%)	
3	20 (10%)	16 (8%)		28 (13.8%)	14 (6.9%)	
4	31 (15.5%)	53 (26.5%)		42 (20.7%)	53 (26.1%)	
Unknown	12 (6%)	11 (5.5%)		10 (4.9%)	4 (2%)	
IDH1 status, n (%)			0.005			< 0.001
Mutant	42 (21%)	25 (12.5%)		57 (28.1%)	14 (6.9%)	
Unknown	14 (7%)	8 (4%)		20 (9.9%)	3 (1.5%)	
Wild-type	44 (22%)	67 (33.5%)		51 (25.1%)	58 (28.6%)	
Histology, n (%)			0.018			< 0.001
Astrocytoma	45 (22.5%)	29 (14.5%)		66 (32.5%)	19 (9.4%)	
Gangliocytoma	3 (1.5%)	2 (1%)		4 (2%)	1 (0.5%)	
GBM	29 (14.5%)	53 (26.5%)		43 (21.2%)	53 (26.1%)	
Oligodendroglioma	14 (7%)	10 (5%)		15 (7.4%)	2 (1%)	
Unknown	9 (4.5%)	6 (3%)				
Radiotherapy, n (%)			0.205			0.093
No	21 (10.5%)	29 (14.5%)		41 (20.2%)	14 (6.9%)	
Unknown	32 (16%)	22 (11%)		16 (7.9%)	14 (6.9%)	
Yes	47 (23.5%)	49 (24.5%)		71 (35%)	47 (23.2%)	
Chemotherapy, n (%)			0.243			0.438
No	19 (9.5%)	25 (12.5%)		37 (18.2%)	18 (8.9%)	
Unknown	32 (16%)	22 (11%)		16 (7.9%)	14 (6.9%)	
Yes	49 (24.5%)	53 (26.5%)		75 (36.9%)	43 (21.2%)	
Laterality, n (%)			0.295			0.238
Both	3 (1.5%)	2 (1%)		2 (1%)	4 (2%)	
Center	38 (19%)	48 (24%)		60 (29.6%)	29 (14.3%)	

(Continued)

TABLE 2 Continued

Characteristic	Cohort1 (n=200)		P	Cohort2 (n=203)		P
	Low (n=100)	High (n=100)		Low (n=128)	High (n=75)	
Middle	6 (3%)	4 (2%)		3 (1.5%)	4 (2%)	
Right	50 (25%)	46 (23%)		63 (31%)	38 (18.7%)	
Unknown	3 (1.5%)	0 (0%)		0 (0%)	0 (0%)	
Tumor location, n (%)			0.440			0.903
Brainstem	2 (1%)	1 (0.5%)		2 (1%)	1 (0.5%)	
Cerebellar	6 (3%)	4 (2%)		6 (3%)	1 (0.5%)	
Frontal	46 (23%)	41 (20.5%)		56 (27.6%)	37 (18.2%)	
Insular	2 (1%)	4 (2%)		3 (1.5%)	2 (1%)	
Occipital	3 (1.5%)	8 (4%)		9 (4.4%)	5 (2.5%)	
Parietal	8 (4%)	10 (5%)		12 (5.9%)	10 (4.9%)	
Sellar	2 (1%)	1 (0.5%)		1 (0.5%)	1 (0.5%)	
Temporal	25 (12.5%)	30 (15%)		37 (18.2%)	17 (8.4%)	
Thalamus	3 (1.5%)	1 (0.5%)		2 (1%)	1 (0.5%)	
Unknown	3 (1.5%)	0 (0%)		0 (0%)	0 (0%)	

(1.28-1.91); GSE16011, HR=1.78 (1.37-2.32); Rembrandt, HR=1.45 (1.16-1.8); CGGA-301, HR=1.54 (1.15-2.07); CGGA-325, HR=2.11 (1.60-2.78), log-rank test $P < 0.05$, Figures 2A–F). When further exploring the relationship of TANK expression with DSS and PFS, patients with high expression of TANK were found to have shorter survival times (DSS, HR=3.02 (2.32-3.92); PFS, HR=2.35 (1.89-2.92), log-rank test $P < 0.05$, Figures 2G, H). In our in-house cohort of 158 glioma patients, we found that glioma patients with high TANK expression generally had shorter OS and PFS times than patients with low TANK expression by qPCR (PFS, HR=2.37 (143-3.92); OS, HR=2.34 (1.27-4.29); log-rank test $P < 0.05$, Figures 2I, J). Similarly, significant prognostic differences were observed in other in-house cohorts using IHC, and the survival time for patients with glioma with high TANK expression was shorter than that of patients with low TANK expression (PFS, HR=2.96 (1.48-5.94); OS, HR=2.45 (1.37-4.39); log-rank test $P < 0.05$, Figures 2K, L). Thus, TANK is a marker of unfavorable prognosis in glioma.

The correlations of TANK with immunological parameters

Considering that TANK expression is correlated with glioma malignancy, we inferred that abnormal expression of TANK might promote the progression of glioma. Among the DEGs, 2902 were significantly upregulated and 1370 were significantly downregulated in the TCGA cohort (Figure 3A). TANK's underlying pathways were further clarified using GSEA. Gliomas with high TANK levels exhibited enrichment in immunomodulatory pathways, including "hypoxia", "angiogenesis", "inflammatory response", "NF-kappaB signaling", and "IL6/STAT3 signaling" in the TCGA cohort (Figure 3B). Next, the complex microenvironment of glioma was assessed by using the ESTIMATE algorithm (41). Furthermore, we found that gliomas with high levels of TANK consistently exhibited higher immune and stromal scores than those with low levels of TANK in four cohorts ($P < 0.05$, Figure 3C), indicating that TANK may regulate immune and stromal cells. We assessed cell infiltration in 33 cancers using seven algorithms. TANK expression was negatively correlated with infiltration of antitumor immune cells such as $CD8^+$ T cells, follicular helper T cells and NK cells ($P < 0.05$, Figures 3D–J). In summary, TANK plays a vital role in the tumor microenvironment.

TANK shapes an immunosuppressive microenvironment in glioma

An increasing number of studies have shown that glioma is a brain tumor characterized by an immunosuppressive microenvironment formed by immunosuppressive cells, which limits the prognosis of tumor therapy (44, 45). Given that TANK may remodel the tumor microenvironment through immunobiological processes, the distribution of 35 immune cell types in gliomas with high and low expression of TANK was examined in the TCGA cohort (Figure 4A). Correlation analysis between TANK expression and infiltration of protumor immune cells in the TCGA cohort revealed that gliomas with high TANK expression contained more immunosuppressive cells, except for CD56dim NK cells ($P < 0.05$, Figure 4B). Other cohorts showed similar results (Figures 4C-E, $P < 0.05$). However, there was no difference in the abundance of Th2 cells in gliomas in other cohorts (Figures 4C-E, $P > 0.05$). Additionally, though the difference in neutrophil infiltration was not observed in the Rembrandt cohort (Figure 4E, $P > 0.05$), gliomas with high-expression TANK had higher neutrophilic infiltration than those with low-expression TANK in other cohorts (Figures 4B-D, $P < 0.05$).

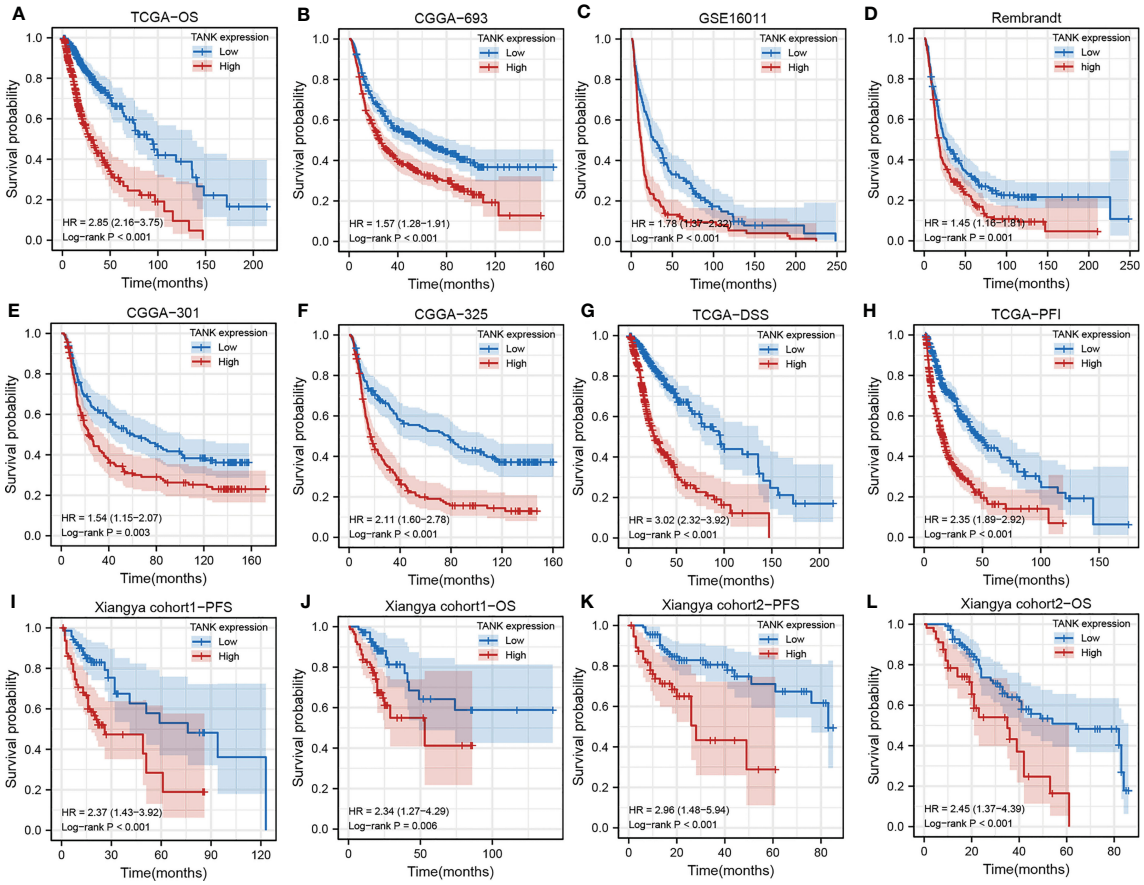


FIGURE 2
TANK is an unfavorable prognostic marker in glioma. **(A-F)** Kaplan-Meier curves displaying the correlations between TANK expression and OS in glioma patients in TCGA **(A)**, CGGA-693 **(B)**, GSE16011 **(C)**, Rembrandt **(D)**, CGGA-301 **(E)** and CGGA-325 **(F)** datasets; **(G, H)** Kaplan-Meier curves showing the correlations between TANK expression and DSS **(G)** and PFI **(H)** in the TCGA cohort; **(I, J)** Kaplan-Meier curves showing the correlations between TANK expression and PFS **(I)** and OS **(J)** in in-house cohort 1 based on qPCR data; **(K, L)** Kaplan-Meier curves showing the correlations between TANK expression and PFS **(K)** and OS **(L)** in in-house cohort 2 based on immunohistochemical data; *P* values were calculated by the log-rank test, and $P < 0.05$ was considered significant.

In the high-TANK group, most MHC molecules were overexpressed, indicating an enhanced ability to present and process antigens. In addition, the levels of CXCL9, CXCL10, and CCR3, which increase the recruitment of CD8+ T cells into the microenvironment of glioma, were increased in gliomas with high TANK expression (46, 47). Chemokines and paired receptors, including CCL2 and CCR2, were upregulated in TANK-expressing gliomas (Figure 4F). The recruitment of effector TIICs is promoted by these chemokines and receptors. Due to the complex and diverse functions of the chemokine system, studies on the relationship between TANK and individual chemokines are insufficient to elucidate the overall immune effect of TANK in the microenvironment.

The cancer immunity cycle includes seven steps: release of cancer cell antigens (Step 1), cancer antigen presentation (Step 2), priming and activation (Step 3), trafficking of immune cells to tumors (Step 4), infiltration of immune cells into tumors (Step 5), recognition of cancer cells by T cells (Step 6), and killing of cancer cells (Step 7). The activity of the tumor immune cycle is a direct result of the function of the chemokine system and immunomodulators (48). In the high-TANK group, the activities of most steps were downregulated, including Step 1, Step 3, and Step 4 (macrophage recruitment, Th1 cell recruitment, NK cell recruitment, and Th17 recruitment), was downregulated (Figure 4G). Consequently, these reduced activities may reduce the level of effector TIIC infiltration into the microenvironment. Interestingly, the activity of cancer cell recognition by T cells was downregulated in the low-TANK group. The activity of Step 7 (killing of cancer cells) was downregulated in the high-TANK group. Immune cell markers were upregulated in the high-TANK group compared with the low-TANK group (Figure 4H).

TANK predicts clinical response and therapeutic opportunities

Pan-cancer analyses showed that the immunological role of TANK is critical for determining the types of cancers that may benefit from anti-TANK immunotherapy. We found that expression of TANK was mutually exclusive with that of several immune checkpoints, including PD-L1, PD-1, CD44, CTLA-4, and PD-L2 (Figure 5A; Table S1). TANK expression was positively correlated with

10.3389/fimmu.2023.1138203

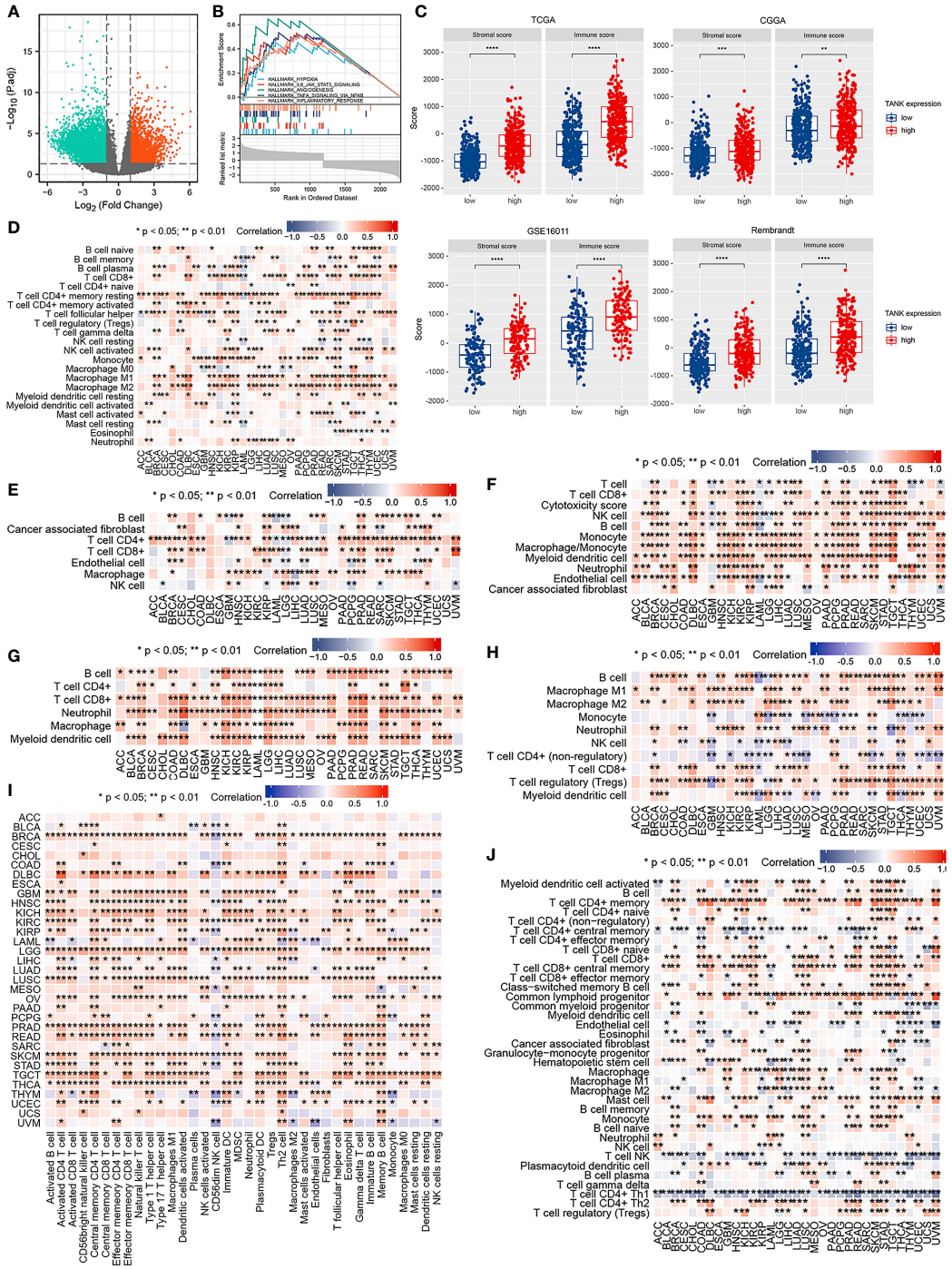


FIGURE 3
Immune relevance of TANK. **(A)** The volcano plot shows differentially expressed genes between the low- and high- TANK groups in the TCGA cohort; **(B)** GSEA of gliomas with low and high expression of TANK in the TCGA cohort; thresholds of a nominal $P < 0.05$ and an FDR < 25% were used to determine the significance of the enrichment score (ES); **(C)** The associations between the stromal and immune scores and TANK expression in the TCGA cohort, CGGA-693 cohort, GSE16011, and Rembrandt cohort; differences between the two groups were compared by Student's t test, and the *P* values are labeled above each boxplot with asterisks (**$P < 0.01$, ***$P < 0.001$, and ****$P < 0.0001$); **(D-J)** Correlation of TANK expression with immune cell infiltration, as evaluated using seven algorithms (TIMER, EPIC, xCELL, CIBERSORT, QUANTISEQ, MCP-counter, and ssGSEA).

the ssGSEA scores of most immunotherapy-associated signatures (Figure 5B). In addition, genetic abnormalities are classical biomarkers of the anti-PD-1/PD-L1 therapeutic response (49). Mutations in the high-TANK group were shown using a waterfall plot (Figure 5C). IDH1 and ATRX were not frequently mutated in gliomas with high TANK expression (IDH1, 43% and ATRX, 26%) compared with those with low TANK expression (IDH1, 78% and ATRX, 35%), while PTEN, TTN and EGFR were more frequently mutated in gliomas with high TANK expression (PTEN, 17%, TTN, 16% and EGFR, 15%) than in those with low-level TANK (PTEN, 4%, TTN, 10% and EGFR, 5%). Several oncogenic pathways cooperatively form the immunosuppressive microenvironment of glioma.

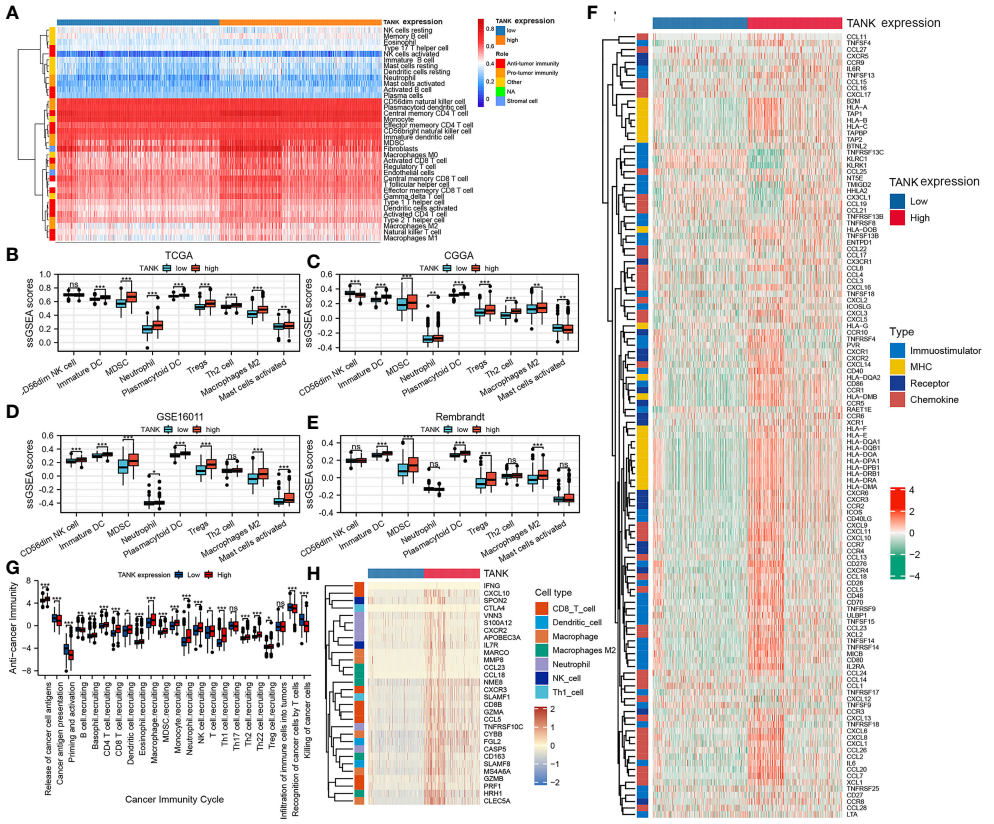


FIGURE 4
TANK shapes an immunosuppressive TME in glioma. **(A)** The landscape of immune cells and stromal cells in the low- and high- TANK groups in the TCGA cohort; **(B–E)** The association between TANK expression and the abundances of seven types of protumor immune cells in the TCGA cohort **(B)**, CGGA-693 cohort **(C)**, GSE16011 **(D)**, and Rembrandt cohort **(E)**; **(F)** Differences in the expression of 122 immunomodulators (chemokines, receptors, MHC, and immunostimulators) between the high- and low-TANK groups in glioma; **(G)** Differences in the various steps of the cancer immunity cycle between the high- and low-TANK groups; **(H)** Differences in the markers of immune cells between the high- and low-TANK groups in the TCGA cohort (- no significance, *P < 0.05, **P < 0.01, and ***P < 0.001).

Therefore, blocking these pathways suppresses the formation of an immunosuppressive microenvironment. We found that immunosuppressive oncogenic pathways were significantly enriched in gliomas with high expression of TANK (P < 0.05, Figures 5D, E).

TANK expression is positively correlated with CD163, CD11b, PD-1, PD-L1, CD40, STAT3 and HIF1A expression in glioma

As mentioned above, the expression of TANK was correlated with the abundances of immune cells, including neutrophils and macrophages, and the expression of immune checkpoints. TANK was also found to be involved in a variety of signaling pathways, including "hypoxia" and "IL6/STAT3 signaling" (Figure 3B). Therefore, we further analyzed the correlation between TANK expression and the expression of CD163 and CD11b, surface markers of M2 macrophages and neutrophils, respectively by IHC. M2 macrophages and neutrophils are important components of the glioma microenvironment and have been reported to be closely related to the prognosis of patients (50). TANK expression was positively correlated with the expression of CD163 and CD11b (CD163, Spearman r = 0.342, P < 0.05; CD11b, Spearman r = 0.360, P < 0.05, Figure 6). Considering that PD-1/PD-L1 and CD40 are important immunosuppressive molecules, we investigated the relationship between TANK expression and the expression of these molecules at the protein level. IHC showed that PD-1, PD-L1 and CD40 expression increased with increasing TANK expression (PD-1, Spearman r = 0.293, P < 0.05; PD-L1, Spearman r = 0.316, P < 0.05; CD40, Spearman r = 0.338, P < 0.05, Figure 6). Finally, TANK expression was found to be positively related to the expression of HIF1A, a core molecule of the hypoxia-induced signaling pathway, and STAT3, a key molecule of the IL6/STAT3 signaling pathway (HIF1A, Spearman r = 0.450, P < 0.05; Spearman r = 0.503, P < 0.05, Figure 6). These observations confirm that TANK may be involved in regulating the complex tumor microenvironment.

Development, validation, and evaluation of the TANK-associated risk score

In this study, we identified 892 overlapping TANK-associated DEGs in the CGGA-693 and TCGA cohorts (Table S2). Among these DEGs, we also identified 299 TANK-associated DEGs significantly

10.3389/fimmu.2023.1138203

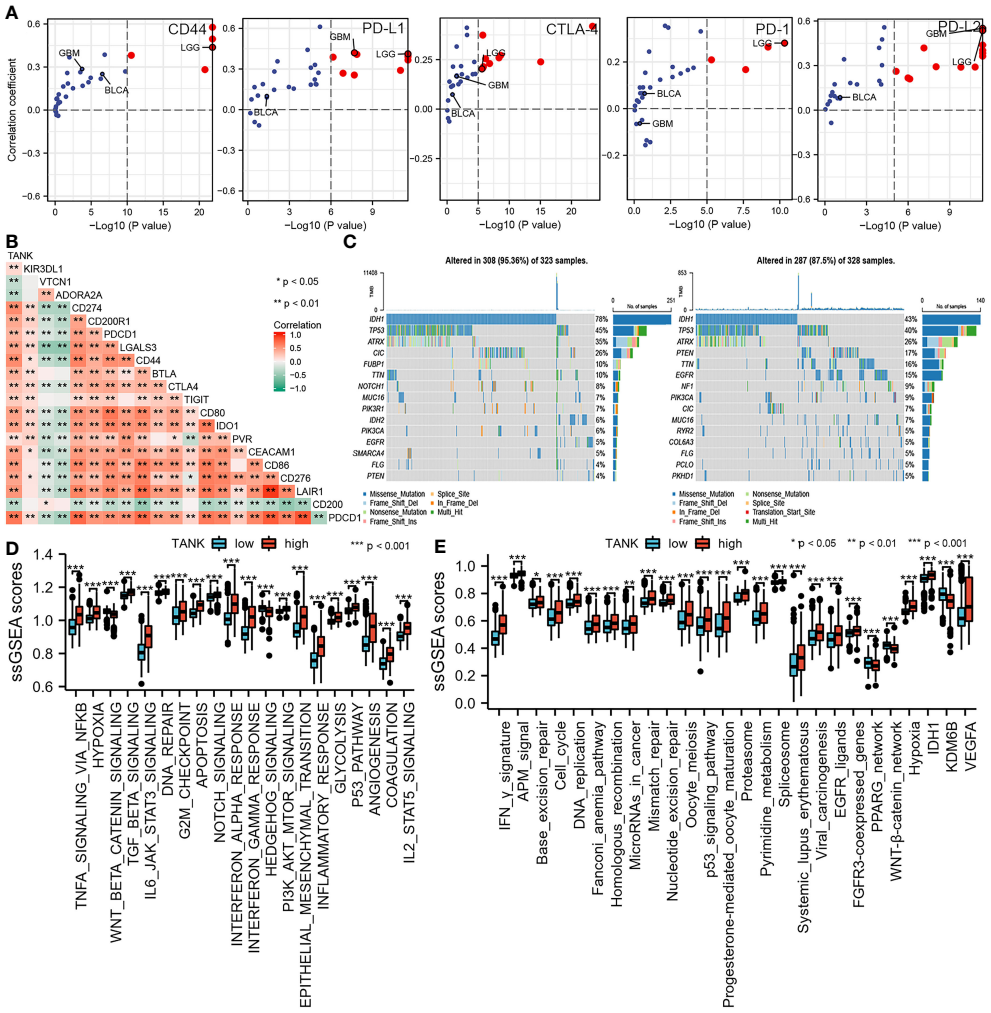


FIGURE 5
TANK predicts clinical response and therapeutic opportunities. **(A)** The associations between TANK expression and the mRNA expression of several immune checkpoints across pan-cancer; **(B)** The association between TANK expression and the mRNA expression of several immune checkpoints in glioma; **(C)** Mutation profiles in the low- and high-TANK groups in the TCGA cohort; **(D)** Correlations between TANK expression and the enrichment scores of several oncogenic pathways; **(E)** Differences in the enrichment scores of immunotherapy-related pathways between the high- and low-TANK groups in the TCGA cohort. *p<0.05, **p<0.01, and ***p<0.001.

associated with prognosis (Table S3). Then, LASSO Cox regression model and SVM were applied to select the 10 most useful factors for developing a prognostic model in the TCGA training set (Figures 7A–C) and obtain a TANK-associated risk score for each patient based on the mRNA expression of thirteen genes and the corresponding LASSO Cox coefficients (Figure 7D). Patients with low-risk scores had significantly longer overall survival times than those with high-risk scores in the TCGA training set, TCGA internal validation set and TCGA set (TCGA training set, HR = 7.80 (5.55-10.97), $P < 0.001$; TCGA validation set, HR = 5.66 (3.50-9.14), $P < 0.001$; TCGA set, HR = 7.28 (5.51-9.61), $P < 0.001$, Figures 7E–G). The CGGA-693 cohort, CGGA-301, CGGA-325, GSE16011, and Rembrandt cohorts were used as external validation sets. The results revealed that the risk score could effectively divide patients into two distinct groups in these external validation cohorts. Patients in the high-risk group had a significantly poorer prognosis than those in the low-risk group (CGGA-693, HR = 3.18 (2.59-3.90), $P < 0.001$; CGGA-301, HR = 3.22 (2.37-4.37), $P < 0.001$; CGGA-325, HR = 4.10 (3.07-5.49), $P < 0.001$; GSE16011, HR = 3.12 (2.34-4.17), $P < 0.001$; Rembrandt, HR = 2.97 (2.34-3.76), $P < 0.001$, Figures 7H–L). The predictive accuracy of the risk score was well validated in the TCGA training set, TCGA internal validation set and TCGA set. The AUC of the risk score was more than 0.80 for survival at 12, 36, and 60 months in the TCGA training set, TCGA internal validation set and TCGA set (Figures 7M–O). Similarly, the AUC of the risk score was more than 0.80 for survival at 12, 36, and 60 months in the CGGA-693 cohort, CGGA-301, CGGA-325, GSE16011 and Rembrandt cohort (Figures 7P–T).

Discussion

In this study, TANK was highly expressed in glioma ($P < 0.05$, Figure 1), as confirmed by qPCR and IHC ($P < 0.05$, Figure 1). In

10.3389/fimmu.2023.1138203

addition, the qPCR and IHC results indicated that higher TANK expression was associated with more malignant phenotypes in glioma ($P < 0.05$, Figure 1, Table 2). Furthermore, TANK was identified as a marker of poor prognosis in glioma. In two in-house cohorts, glioma patients with high expression of TANK generally had shorter OS and PFS times than those with low expression of TANK, as determined by qPCR and IHC (log-rank test $P < 0.05$, Figures 2I–L). Previous studies have shown that downregulation of TANK can arrest cells in S-phase and prevent tumor cell migration (36). Our results are consistent with previous results, indicating that TANK could play a protumorigenic role in glioma, and consistent with previous results. However, the clinical significance and expression pattern of TANK in glioma have not been reported. The possibility that TANK is a potential immunotherapeutic target for glioma needs further exploration.

TANK has been identified as a TRAF-interacting protein and can activate the NF-κB signaling pathway (26). The NF-κB pathway is indispensable for immune responses and inflammatory processes, as well as in activating survival and proinflammatory genes within the tumor microenvironment (51). Therefore, we hypothesized that

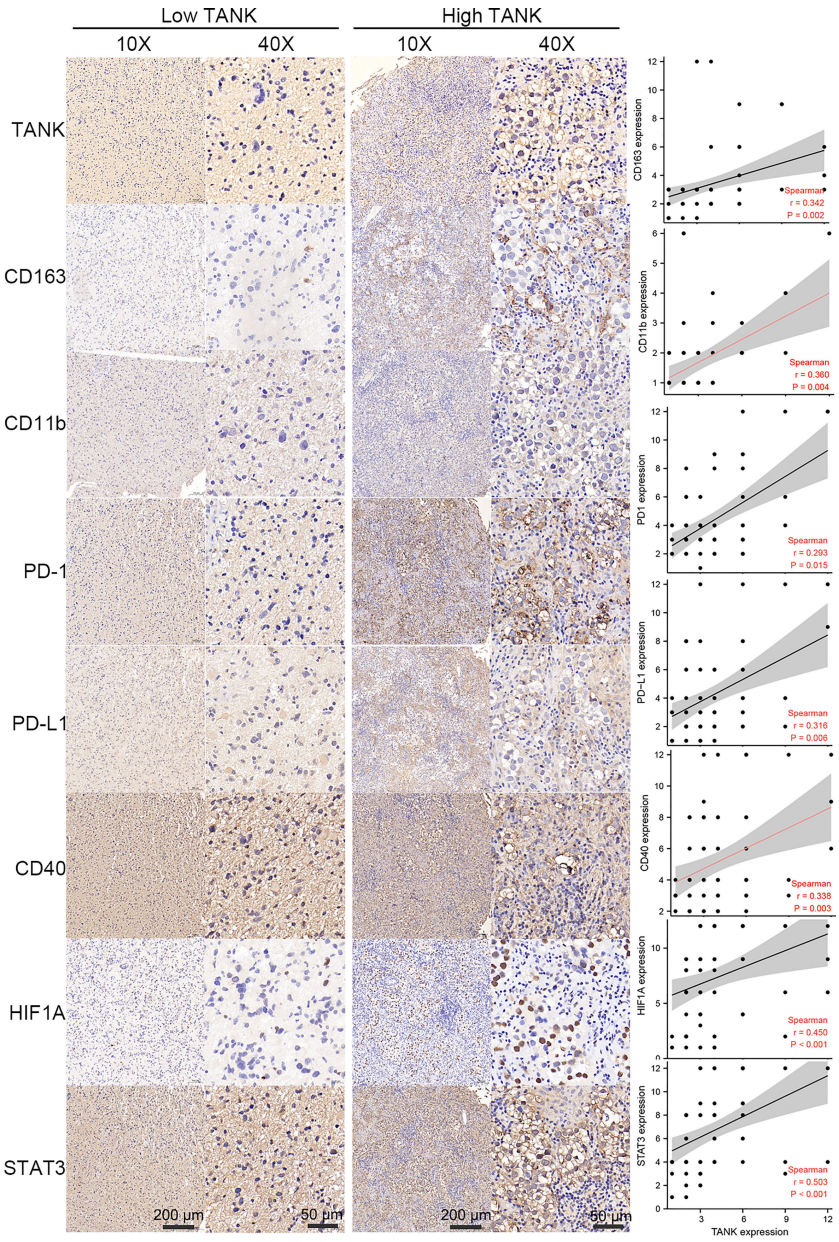


FIGURE 6
TANK expression is positively correlated with CD163, CD11b, PD-1, PD-L1, CD40, STAT3 and HIF1A expression in glioma. As surface markers of M2 macrophages and neutrophils, IHC shows that CD163 and CD11b are positively correlated with the expression of TANK (CD163, Spearman r = 0.342, $P < 0.05$; CD11b, Spearman r = 0.360, $P < 0.05$). As important immunosuppressive molecules, PD-1, PD-L1 and CD40 expression increased with the increase of TANK expression (PD-1, Spearman r = 0.293, $P < 0.05$; PD-L1, Spearman r = 0.316, $P < 0.05$; CD40, Spearman r = 0.338, $P < 0.05$). TANK was found to be positively related to HIF1A expression (Spearman r = 0.450, $P < 0.05$), a core molecule of the hypoxia-induced signaling pathway, and STAT3 (Spearman r = 0.503, $P < 0.05$), a key molecule of IL6/STAT3 signaling pathway.

TANK might be involved in the remodeling of the tumor microenvironment. GSEA showed that TANK was involved in immunoregulatory pathways in the TCGA cohort (Figure 3B). TANK was found to be closely and positively associated with the expression of HIF1A and STAT3 (HIF1A, Spearman r = 0.450, $P < 0.05$; Spearman r = 0.503, $P < 0.05$, Figure 6), a result that validated the above findings. In four cohorts, the immune and stromal scores were higher in gliomas with high TANK expression than in gliomas with low TANK expression ($P < 0.05$, Figure 3C). Pan-cancer immune infiltration analysis based on seven algorithms showed that TANK expression was closely correlated with infiltration of immunosuppressive cells ($P < 0.05$, Figures 3D–J). Importantly, the

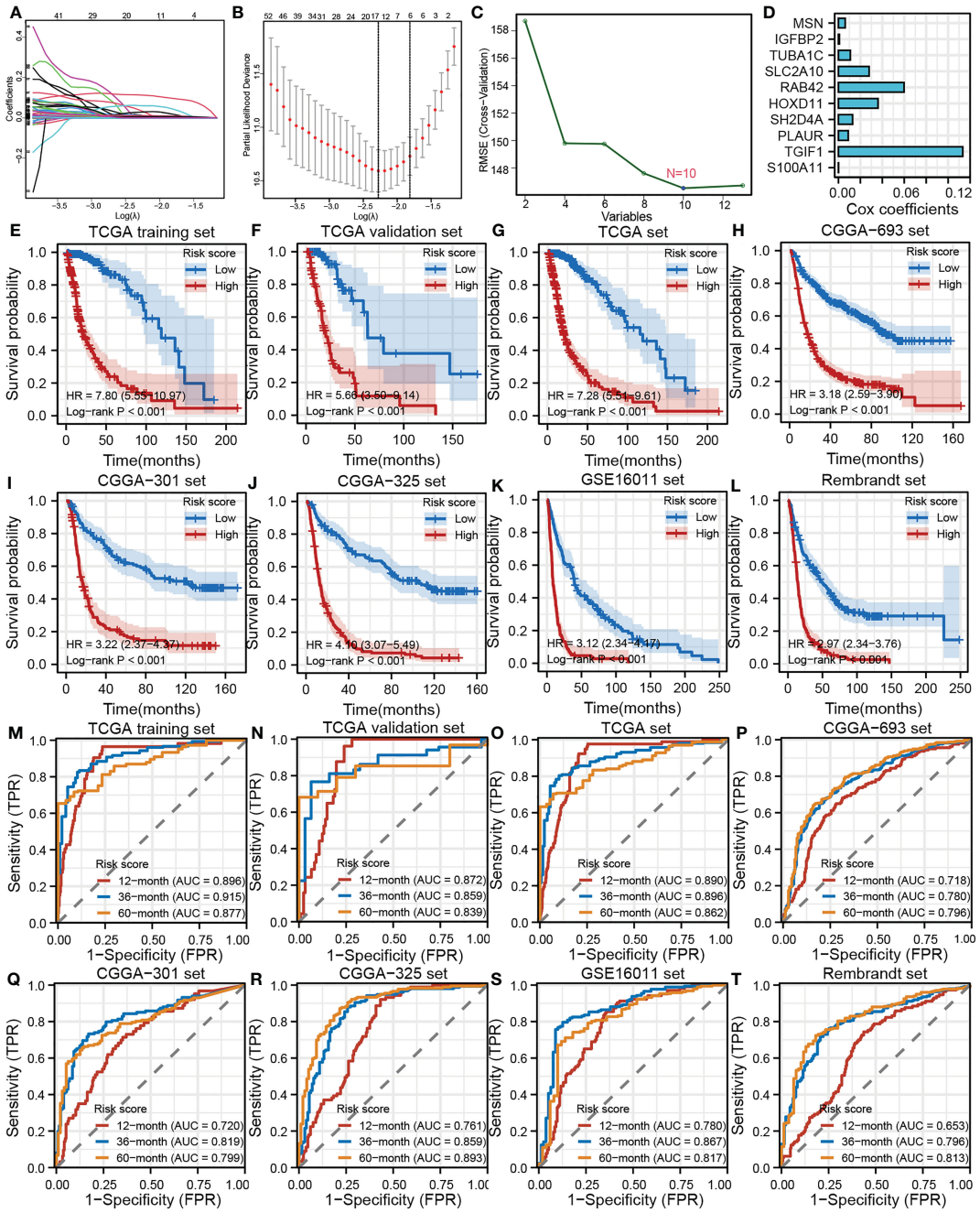


FIGURE 7
Development, validation and evaluation of the TANK-associated risk score. **(A)** The partial likelihood deviance distribution of the LASSO coefficient; **(B)** Partial likelihood deviance determined by the LASSO regression model; **(C)** Identification of hub genes by SVM; **(D)** An ensemble of 10 TANK-associated signatures with the Cox regression coefficients; **(E–L)** Kaplan-Meier curves show the correlation between the risk scores and overall survival of patients in the TCGA training set **(E)**, TCGA internal validation set **(F)**, whole TCGA set **(G)**, CGGA-693 cohort **(H)**, CGGA-301 **(I)**, CGGA-325 **(J)**, GSE16011 **(K)**, and Rembrandt cohorts **(L)**; *P* values were calculated by the log-rank test, and $P < 0.05$ was defined as the cutoff criterion; **(M-T)** Time-dependent ROC analysis of survival at 12 months, 36 months, and 60 months showed the predictive accuracy of the TANK-associated prognostic model in the TCGA training set **(M)**, TCGA internal validation set **(N)**, whole TCGA set **(O)**, CGGA-693 cohort **(P)**, CGGA-301 **(Q)**, CGGA-325 **(R)**, GSE16011 **(S)**, and Rembrandt cohorts **(T)**. AUC, area under the curve; ROC, receiver operating characteristic.

abundances of most immunosuppressive cells were higher in gliomas with high TANK expression in the four cohorts ($P < 0.05$, Figures 4B–E). TANK was negatively related to various immunomodulators (Figure 4F). In the high-TANK group, the activities of most of the steps were downregulated (Figure 4G). Subsequently, inactivity of these steps may weaken the infiltration of immune cells into the microenvironment. Therefore, we inferred that TANK shapes an inflamed TME in glioma. Although the role of TANK in regulating the tumor microenvironment of glioma has not been reported in previous studies, its important role in other tumors has been reported. Moreover, TANK expression was positively correlated with PD-L1, PD-1, and CD44 expression in various cancers (Figure 5A). The immunohistochemical results showed that PD-1, PD-L1 and CD40 expression increased with the increasing TANK expression in glioma (Figure 6).

Inhibiting oncogenic pathways blocks the formation of an immunosuppressive microenvironment, thereby reactivating cancer immunity. Most pathways were observably upregulated in the high-TANK group ($P < 0.05$, Figures 5D, E). Our observations provide insight for subsequent research on the mechanism by which TANK expression regulates immunity and lay a foundation for developing new treatment options. Similarly, for gliomas with high expression of TANK, one of the previous treatment methods was to transform the immunosuppressive microenvironment into an immune-activated state, thus triggering the anticancer immune response. The expression of inhibitory immune checkpoints may be upregulated by negative feedback regulation. Therefore, subsequent ICB therapy may reactivate suppressed anticancer immunity. This approach may enhance the efficacy of anti-TANK therapy and help trigger anticancer immunity. The combination of different ICB drugs with anti-TANK therapy is more effective than single therapy. Current therapeutic targets of ICB therapy, such as PD-L1and PD-1, are associated with each other in other tumors. Therefore, the combination of these drugs seems to have a synergistic effect. In contrast, TANK expression was significantly positively correlated with that of some ICB targets, suggesting complementary effects of anti-TANK and ICB therapy.

Finally, we established a TANK-related risk model to predict prognosis based on the expression of TANK-related genes. The TANK-associated risk score could effectively divide patients into two distinct groups, and patients in the high-risk group had a significantly worse prognosis than those in the low-risk group (log-rank test $P < 0.001$, Figures 7E–L). The risk score showed good predictive accuracy (Figures 7M–T).

However, there are limitations to our study. First, some contradictory findings need to be confirmed by carrying out reliable experiments. Second, the role of TANK in the tumor microenvironment and its underlying regulatory mechanisms need to be further explored. Third, more functional experiments are needed to validate the role of TANK in the glioma microenvironment, especially the immune microenvironment.

In conclusion, high expression of TANK indicates a malignant phenotype of glioma, predicts a poor prognosis and shapes an immunosuppressive tumor microenvironment. Combined anti-tank cancer immunotherapy may be a more effective strategy than monotherapy.

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 within the article/Supplementary Materials.

Ethics statement

Informed consent was obtained from all patients, and this study was approved by the ethics committee of Xiangya Hospital, Central South University.

Author contributions

HL, QZ, SL, and YG conceived and designed the study and analyzed the results. Other authors performed analysis procedures. QZ, SL, and YG, wrote the manuscript. HL provided materials and funding support. All authors contributed to the editing of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (82101143), Zhejiang Provincial Natural Science Foundation of China (LQ20C120001), Wenzhou Medical University Basic Scientific Research Operating Expenses (KYYW202105). We are sincerely grateful to those who created and maintained these public databases.

Acknowledgments

We are sincerely grateful to those who created and maintained these public databases such as TCGA, CGGA and GEO.

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.2023.1138203/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Expression pattern of TANK in glioma. **(A–D)** The expression level of TANK in glioma with different WHO grades in the TCGA cohort **(A)**, CGGA-693 cohort **(B)**, GSE16011 **(C)**, and Rembrandt cohort **(D)**; **(E–G)** The expression level of TANK in glioma with wild-type and mutant IDH or IDH1 in the TCGA cohort **(E)**, CGGA-693 cohort **(F)**, and GSE16011 **(G)**; **(H–K)** The expression level of TANK in glioma with different histologies in the TCGA cohort **(H)**, CGGA-693 cohort **(I)**, GSE16011 **(J)**, and Rembrandt cohort **(K)**. $P < 0.05$ was considered significant. (- no significance, *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$).

SUPPLEMENTARY TABLE 1
The relationship between the expression of TANK and several immune checkpoints, including PD-L1, PD-1, CD44, CTLA-4, and PD-L2.

SUPPLEMENTARY TABLE 2
TANK-associated DEGs in the CGGA-693 and TCGA cohorts.

SUPPLEMENTARY TABLE 3
299 TANK-associated DEGs significantly associated with prognosis.

References

1. Zhou Q, Yan X, Liu W, Yin W, Xu H, Cheng D, et al. Three immune-associated subtypes of diffuse glioma differ in immune infiltration, immune checkpoint molecules, and prognosis. *Front Oncol* (2020) 10:586019. doi: 10.3389/fonc.2020.586019

2. Masui K, Kato Y, Sawada T, Mischel PS, Shibata N. Molecular and genetic determinants of glioma cell invasion. *Int J Mol Sci* (2017) 18(12):2609. doi: 10.3390/ijms18122609

3. Brennan CW, Verhaak RG, McKenna A, Campos B, Noushmehr H, Salama SR, et al. The somatic genomic landscape of glioblastoma. *Cell* (2013) 155(2):462–77. doi: 10.1016/j.cell.2013.09.034

4. Duffau H, Taillandier L. New concepts in the management of diffuse low-grade glioma: proposal of a multistage and individualized therapeutic approach. *Neuro Oncol* (2015) 17(3):332–42. doi: 10.1093/neuonc/nou153

5. Reifenberger G, Wirsching HG, Knobbe-Thomsen CB, Weller M. Advances in the molecular genetics of gliomas - implications for classification and therapy. *Nat Rev Clin Oncol* (2017) 14(7):434–52. doi: 10.1038/nrclinonc.2016.204

6. Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, et al. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell* (2016) 164(3):550–63. doi: 10.1016/j.cell.2015.12.028

7. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 world health organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol* (2016) 131(6):803–20. doi: 10.1007/s00401-016-1545-1

8. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO classification of tumors of the central nervous system: a summary. *Neuro-Oncology* (2021) 23(8):1231–51. doi: 10.1093/neuonc/noab106

9. van Tellingen O, Yetkin-Arik B, de Gooijer MC, Wesseling P, Wurdinger T, de Vries HE. Overcoming the blood-brain tumor barrier for effective glioblastoma treatment. *Drug Resist Update* (2015) 19:1–12. doi: 10.1016/j.drup.2015.02.002

10. Lim M, Xia Y, Bettegowda C, Weller M. Current state of immunotherapy for glioblastoma. *Nat Rev Clin Oncol* (2018) 15(7):422–42. doi: 10.1038/s41571-018-0003-5

11. Ghouzlani A, Kandoussi S, Tall M, Reddy KP, Rafii S, Badou A. Immune checkpoint inhibitors in human glioma microenvironment. *Front Immunol* (2021) 12:679425. doi: 10.3389/fimmu.2021.679425

12. Foreman PM, Friedman GK, Cassady KA, Markert JM. Oncolytic virotherapy for the treatment of malignant glioma. *Neurotherapeutics* (2017) 14(2):333–44. doi: 10.1007/s13311-017-0516-0

13. Yip A, Webster RM. The market for chimeric antigen receptor T cell therapies. *Nat Rev Drug Discov* (2018) 17(3):161–2. doi: 10.1038/nrd.2017.266

14. Iwami K, Natsume A, Wakabayashi T. Cytokine therapy of gliomas. *Prog Neurol Surg* (2018) 32:79–89. doi: 10.1159/000469682

15. Li M, Han S and Shi X. *In situ* dendritic cell vaccination for the treatment of glioma and literature review. *Tumour Biol* (2016) 37(2):1797–801. doi: 10.1007/s13277-015-3958-1

16. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* (2012) 12(4):252–64. doi: 10.1038/nrc3239

17. Prieto PA, Yang JC, Sherry RM, Hughes MS, Kammula US, White DE, et al. CTLA-4 blockade with ipilimumab: long-term follow-up of 177 patients with metastatic melanoma. *Clin Cancer Res* (2012) 18(7):2039–47. doi: 10.1158/1078-0432.CCR-11-1823

18. Ansell SM, Lesokhin AM, Borrello I, Halwani A, Scott EC, Gutierrez M, et al. PD-1 blockade with nivolumab in relapsed or refractory hodgkin's lymphoma. *N Engl J Med* (2015) 372(4):311–9. doi: 10.1056/NEJMoa1411087

19. Sangro B, Gomez-Martin C, de la Mata M, Iñarrairaegui M, Garralda E, Barrera P, et al. A clinical trial of CTLA-4 blockade with tremelimumab in patients with hepatocellular carcinoma and chronic hepatitis c. *J Hepatol* (2013) 59(1):81–8. doi: 10.1016/j.jhep.2013.02.022

20. Suresh K, Naidoo J, Lin CT, Danoff S. Immune checkpoint immunotherapy for non-small cell lung cancer: benefits and pulmonary toxicities. *Chest* (2018) 154 (6):1416–23. doi: 10.1016/j.chest.2018.08.1048

21. Ruan S, Xie R, Qin L, Yu M, Xiao W, Hu C, et al. Aggregable nanoparticles-enabled chemotherapy and autophagy inhibition combined with anti-PD-L1 antibody for improved glioma treatment. *Nano Lett* (2019) 19(11):8318–32. doi: 10.1021/acs.nanolett.9b03968

22. Stessin AM, Clausi MG, Zhao Z, Lin H, Hou W, Jiang Z, et al. Repolarized macrophages, induced by intermediate stereotactic dose radiotherapy and immune checkpoint blockade, contribute to long-term survival in glioma-bearing mice. *J Neurooncol* (2020) 147(3):547–55. doi: 10.1007/s11060-020-03459-y

23. Fecci PE, Ochiai H, Mitchell DA, Grossi PM, Sweeney AE, Archer GE, et al. Systemic CTLA-4 blockade ameliorates glioma-induced changes to the CD4+ T cell compartment without affecting regulatory T-cell function. *Clin Cancer Res* (2007) 13 (7):2158–67. doi: 10.1158/1078-0432.CCR-06-2070

24. Kim JE, Patel MA, Mangraviti A, Kim ES, Theodros D, Velarde E, et al. Combination therapy with anti-PD-1, anti-TIM-3, and focal radiation results in regression of murine gliomas. *Clin Cancer Res* (2015) 23(suppl 5):124–36. doi: 10.1158/1078-0432.CCR-15-1535

25. Rothe M, Xiong J, Shu HB. I-TRAF is a novel TRAF-interacting protein that regulates TRAF-mediated signal transduction. *Proc Natl Acad Sci* (1996) 93(16):8241–6. doi: 10.1073/pnas.93.16.8241

26. Cheng G, Baltimore D. TANK, a co-inducer with TRAF2 of TNF- and CD 40L-mediated NF-kappaB activation. *Genes Dev* (1996) 10(8):963–73. doi: 10.1101/gad.10.8.963

27. Kaye KM, Devergne O, Harada JN, Izumi KM, Yalamanchili R, Kieff E, et al. Tumor necrosis factor receptor associated factor 2 is a mediator of NF-kappa b activation by latent infection membrane protein 1, the Epstein-Barr virus transforming protein. *Proc Natl Acad Sci USA* (1996) 93(20):11085–90. doi: 10.1073/pnas.93.20.11085

28. Didonato JA, Mercurio F, Karin M. NF-κB and the link between inflammation and cancer. *Immunol Rev* (2012) 246(1):379–400. doi: 10.1111/j.1600-065X.2012.01099.x

29. Nomura F, Kawai T, Nakanishi K, Akira S. NF-κB activation through IKK-i-dependent I-TRAF/TANK phosphorylation - nomura - 2001 - genes to cells - Wiley online library. *Genes to Cells* (2000) 5(3):191–202. doi: 10.1046/j.1365-2443.2000.00315.x

30. Lawrence T. The nuclear factor NF-kappaB pathway in inflammation. *Cold Spring Harb Perspect Biol* (2009) 1(6):a001651. doi: 10.1101/cshperspect.a001651

31. Chariot A, Leonardi A, Muller J, Bonif M, Brown K, Siebenlist U. Association of the adaptor TANK with the I kappa b kinase (IKK) regulator NEMO connects IKK complexes with IKK epsilon and TBK1 kinases. *J Biol Chem* (2002) 277(40):37029–36. doi: 10.1074/jbc.M205069200

32. Guo B, Cheng G. Modulation of the interferon antiviral response by the TBK1/IKKi adaptor protein TANK. *J Biol Chem* (2007) 282(16):11817–26. doi: 10.1074/jbc.M700017200

33. Clark K, Takeuchi O, Akira S, Cohen P. The TRAF-associated protein TANK facilitates cross-talk within the IkappaB kinase family during toll-like receptor signaling. *Proc Natl Acad Sci USA* (2011) 108(41):17093–8. doi: 10.1073/pnas.1114194108

34. Yim EK, Meoyng J, Namakoong SE, Um SJ, Park JS. Genomic and proteomic expression patterns in HPV-16 E6 gene transfected stable human carcinoma cell lines. *DNA Cell Biol* (2004) 23(12):826. doi: 10.1089/dna.2004.23.826

35. Yim EK, Lee KH, Kim CJ, Park JS. Analysis of differential protein expression by cisplatin treatment in cervical carcinoma cells. *Int J Gynecol Cancer* (2010) 16(2):690–7. doi: 10.1111/j.1525-1438.2006.00586.x

36. Stellzig J, Chariot A, Shostak K, Ismail Göktuna S, Renner F, Acker T, et al. Deregulated expression of TANK in glioblastomas triggers pro-tumorigenic ERK1/2 and AKT signaling pathways. *Oncogenesis* (2013) 2:e79. doi: 10.1038/oncsis.2013.42

37. Zhou Q, Wei M, Shen W, Huang S, Fan J, Huang H. SYK is associated with malignant phenotype and immune checkpoints in diffuse glioma. *Front Genet* (2022) 13:899883. doi: 10.3389/fgene.2022.899883

38. Hu J, Yu A, Othmane B, Qiu D, Li H, Li C, et al. Siglec15 shapes a non-inflamed tumor microenvironment and predicts the molecular subtype in bladder cancer. *Theranostics* (2021) 11(7):3089–108. doi: 10.7150/thno.53649

39. Xu L, Deng C, Pang B, Zhang X, Liu W, Liao G, et al. TIP: a web server for resolving tumor immunophenotype profiling. *Cancer Res* (2018) 78(23):6575–80. doi: 10.1158/0008-5472.CAN-18-0689

40. Auslander N, Zhang G, Lee JS, Frederick DT, Miao B, Moll T, et al. Robust prediction of response to immune checkpoint blockade therapy in metastatic melanoma. *Nat Med* (2018) 24(10):1545–9. doi: 10.1038/s41591-018-0157-9

41. 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

42. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43(7):e47. doi: 10.1093/nar/gkv007

43. Blaschek M, Roudier P, Poggio M, Hedley CB. Prediction of soil available water-holding capacity from visible near-infrared reflectance spectra. *Sci Rep* (2019) 9 (1):12833. doi: 10.1038/s41598-019-49226-6

44. van Hooren L, Vaccaro A, Ramachandran M, Vazaios K, Libard S, van de Walle T, et al. Agonistic CD40 therapy induces tertiary lymphoid structures but impairs responses to checkpoint blockade in glioma. *Nat Commun* (2021) 12(1):4127. doi: 10.1038/s41467-021-24347-7

45. Yang R, Zhang Z, Fu S, Hou T, Mu W, Liang S, et al. Charge and size dual switchable nanocage for novel triple-interlocked combination therapy pattern. *Adv Sci (Weinh)* (2020) 7(18):2000906. doi: 10.1002/advs.202000906

46. Richmond JM, Strassner JP, Essien KI, Harris JE. T-Cell positioning by chemokines in autoimmune skin diseases. *Immunol Rev* (2019) 289(1):186–204. doi: 10.1111/imr.12762

47. Tulic MK, Cavazza E, Cheli Y, Jacquel A, Luci C, Cardot-Leccia N. Innate lymphocyte-induced CXCR3B-mediated melanocyte apoptosis is a potential initiator of T-cell autoreactivity in vitiligo. *Nat Commun* (2019) 10(1):2178. doi: 10.1038/s41467-019-09963-8

48. Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* (2013) 39(1):1–10. doi: 10.1016/j.immuni.2013.07.012

49. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* (2015) 348(6230):124–8. doi: 10.1126/science.aaa1348

50. Zhang C, Cheng W, Ren X, Wang Z, Liu X, Li G, et al. Tumor purity as an underlying key factor in glioma. *Clin Cancer Res* (2017) 23(20):6279–91. doi: 10.1158/1078-0432.CCR-16-2598

51. Vlantis K, Wullaert A, Polykratis A, Kondylis V, Dannappel M, Schwarzer R, et al. NEMO prevents RIP kinase 1-mediated epithelial cell death and chronic intestinal inflammation by NF-kappaB-Dependent and -independent functions. *Immunity* (2016) 44(3):553–67. doi: 10.1016/j.immuni.2016.02.020

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 09 May 2023
DOI 10.3389/fonc.2023.1163695

OPEN ACCESS

EDITED BY
Nan Zhang,
Harbin Medical University, China

REVIEWED BY
Cailian Wang,
Southeast University, China
Yiqiao Luo,
Sichuan University, China

*CORRESPONDENCE
Jiao Yang
Jessieyang321@163.com

†These authors have contributed equally to this work

RECEIVED 11 February 2023
ACCEPTED 18 April 2023
PUBLISHED 09 May 2023

CITATION
Huang Y, Lei X, Sun L, Liu Y and Yang J (2023) Leveraging various extracellular matrix levels to assess prognosis and sensitivity to immunotherapy in patients with ovarian cancer.
Front. Oncol. 13:1163695.
doi: 10.3389/fonc.2023.1163695

Leveraging various extracellular matrix levels to assess prognosis and sensitivity to immunotherapy in patients with ovarian cancer

Youqun Huang[1†], Xingxing Lei[2†], Lisha Sun[2†], Yu Liu[3] and Jiao Yang[2*]

[1]Department of Nephrology-2, Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China, [2]Department of Endocrinology, Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China, [3]Department of Nephrology, South China Hospital, Health Science Center, Shenzhen University, Shenzhen, China

Background: Ovarian cancer (OC) is the fifth leading cause of cancer-related deaths among women. Late diagnosis and heterogeneous treatment result in a poor prognosis for patients with OC. Therefore, we aimed to develop new biomarkers to predict accurate prognoses and provide references for individualized treatment strategies.

Methods: We constructed a co-expression network applying the "WGCNA" package and identified the extracellular matrix-associated gene modules. We figured out the best model and generated the extracellular matrix score (ECMS). The ECMS' ability to predict accurate OC patients' prognoses and responses to immunotherapy was evaluated.

Results: The ECMS was an independent prognostic factor in the training [hazard ratio (HR) = 3.132 (2.068–4.744), p< 0.001] and testing sets [HR = 5.514 (2.084–14.586), p< 0.001]. The receiver operating characteristic curve (ROC) analysis showed that the AUC values for 1, 3, and 5 years were 0.528, 0.594, and 0.67 for the training set, respectively, and 0.571, 0.635, and 0.684 for the testing set, respectively. It was found that the high ECMS group had shorter overall survival than the low ECMS group [HR = 2 (1.53–2.61), p< 0.001 in the training set; HR = 1.62 (1.06–2.47), p = 0.021 in the testing set; HR = 1.39 (1.05–1.86), p = 0.022 in the training set]. The ROC values of the ECMS model for predicting immune response were 0.566 (training set) and 0.572 (testing set). The response rate to immunotherapy was higher in patients with low ECMS.

Conclusion: We created an ECMS model to predict the prognosis and immunotherapeutic benefits in OC patients and provided references for individualized treatment of OC patients.

KEYWORDS

ovarian cancer, ECMS, immune landscape, prognosis, immunotherapy

Introduction

In 2022, approximately 19,880 patients in the United States were diagnosed with ovarian cancer (OC) and 12,810 patients died from OC. It is the 11th most prevalent cancer and the fifth leading cause of cancer-related deaths among women (1). Indeed, OC includes a variety of pathological types, and epithelial OC is the most frequent pathological type, accounting for approximately 80% (2). The 5-year overall survival (OS) rate after OC diagnosis is only 47% due to failure to diagnose early, metastasis, relapse, and drug resistance (3). The first-line treatment for OC includes surgery and the administration of platinum drugs combined with paclitaxel, and maintenance therapies include bevacizumab and poly(ADP-ribose) polymerase inhibitors. In addition, the idea that immunotherapy has potential effects on various cancers, including OC, has been demonstrated. Therapeutic targeting of the programmed cell death protein 1 (PD-1) and cytotoxic T-lymphocyte antigen 4 (CTLA-4) is effective in many cancers, which can improve the survival rate (4). Previous studies have constructed several models that could predict chemotherapy's prognosis and efficacy in OC patients (5–8). However, these models do not consider the role of the extracellular matrix (ECM).

The ECM comprises different macromolecules, including glycoprotein, collagens, and proteoglycans, assembled into a three-dimensional supramolecular network to regulate cell growth, survival, motility, and differentiation (9). In addition, ECM is related to the formation of a tumor microenvironment (TME) and its dysregulation can promote tumor progression (10). Deposition of ECM is related to poor outcomes in multiple tumors. For example, in patients with uroepithelial carcinoma of the bladder, inflammatory cancer-associated fibroblasts were significantly associated with poor outcomes (11). In addition, in a study about pancreatic ductal adenocarcinoma, stromal-derived fibroblast growth factor 10 could activate fibroblast growth factor receptor 2 expressed on cancer cells to induce migration and invasion, which was correlated with poor prognosis (12). Similarly, the matrix remodeling gene expression correlated with poor prognosis in breast cancer (BC) patients (13). Abnormal ECM deposition may reduce the effects of chemotherapy and immunotherapy. In preclinical mouse tumor models, inhibition of collagen crosslinking decreased ECM content and tumor stiffness, thereby increasing the efficacy of PD-1 blockade treatment (14). In addition, inhibition of ECM deposition could inhibit colorectal cancer metastasis and enhance the effects of bevacizumab (15). On the contrary, an analysis of pancreatic cancer confirmed that the TGF-b signaling pathway could induce ECM deposition, resulting in the inability to block PD-1 (16). Since ECM is linked to the efficacy and prognosis of many tumor patients, exploring ECM-based prognostic and efficacy prediction models for OC may help the prognostic assessment and individualized treatment strategies to benefit more OC patients.

In this study, we constructed a co-expression network applying the "WGCNA" package and identified the extracellular matrix-associated gene modules. Independent prognostic factors in candidate ECM genes were then screened. We determined the best model utilizing the Cox proportional hazard model with the LASSO penalty. Therefore, a new ECM score (ECMS) model was developed, and its ability to predict accurate OC patients' prognoses and responses to immunotherapy was evaluated.

Methods

Data extraction and data processing

The transcriptome RNA-seq data and the corresponding information of OC patients were downloaded from the Cancer Genome Atlas (TCGA) database (https://cancergenome.nih.gov/) by the Genomic Data Commons platform. We obtained 349 OV samples after excluding participants with lost visits and missing information. We standardized the original fragments per kilobase per million (FPKM) expression data to transcripts per kilobase per million (TPM) and which served as a training set. In addition, from the University of California Santa Cruz (UCSC) Xena platform (https://xena.ucsc.edu/), we downloaded transcriptome RNA-seq data and the corresponding information of 111 OC patients in ICGC database and used them as a testing set. We collected publicly available immunotherapy cohorts to predict immunotherapy response and used them as a validation set. Finally, the IMvigor210 dataset was collected from http://research-pub.gene.com/IMvigor210CoreBiologie. The IMvigor210 cohort contained 298 urothelial carcinoma patients receiving anti-PD-L1 therapies. A total of 1,028 ECM genes were collected from the hallmark dataset on the MSIDGB website (https://www.gsea-msigdb.org/gsea/msigdb/). We included eligible OC samples based on the following criteria: (a) primary diagnosis of ovarian cancer; (b) having a complete gene expression matrix; (c) having well-established clinical follow-up information (including prognosis, stage, and age).

Screening of candidate ECM genes

The tumor purity and immune activity were assessed by the ESTIMATE algorithm. Then, we built a co-expression network based on transcriptomic data and ESTIMATE results by the "WGCNA" package and identified ECM-associated gene modules. Parameter settings: unsigned network architecture was adopted with a minimum module gene of 30, deepSplit = 2, cutNet = 0.02, and a correlation threshold of 0.9 used to identify genes of the same module. The intersection of the most relevant ECM-associated gene modules with ECM genes was considered candidate ECM genes.

Construction of the ECMS model

We determined the independent prognostic factors in candidate ECM genes by univariate COX regression. We selected the best predictive model applying the Cox proportional hazard model with the LASSO penalty and set fivefold cross-validation to prevent overfitting. To achieve cross-validated random sampling, we carried out 500 iterations to figure out the most robust model.

After 500 iterations, the model with the highest frequency was regarded as the final model and generated the ECMS:

$$ECMS = \sum iCoefficient(mRNA_i) \times Expression(mRNA_i)$$

We calculated the concordance index (C-index) utilizing the R package "survcomp."

Then, we calculated the ECMS of all patients and divided them into the high and low ECMS groups (also called high- and low-risk groups) according to the median ECMS. To assess the model's prognostic utility, Kaplan–Meier (KM) curves, time-dependent receiver operating characteristic curves (tROC), and univariate and multivariate Cox regression analyses were applied.

Functional enrichment and immune infiltration analyses

We carried out a single-sample gene set enrichment analysis (ssGSEA) by applying the R package "gsva" based on the molecular markers mentioned in previous studies (17–20). The detailed molecular markers are provided in Table S1. In addition, we applied the GSEA to compare two ECMS groups and used the $p< 0.05$ criterion to discover the significant KEGG pathway. The R package "limma" had been proposed to identify differentially expressed genes (DEGs) between two ECMS groups at a significance threshold of fdr<0.05, FC >2. In addition, we applied the Metascape (http://www.metascape.org) database to carry out functional enrichment analysis. The evaluation of the immune cell infiltration was performed through the R package "CIBERSORT" (21). Applying the ESTIMATE algorithm, we evaluated the tumor purity and immune activity (22). Finally, we collected SNV neoantigens and indel neoantigens samples from Thorsson et al. (23).

Prediction of immunotherapy response

We calculated the patients' immunophenoscore (IPS) based on the genetic characteristics of different immune cell phenotypes. A higher IPS indicates a more active immune response and a higher response to immunotherapy. We applied the TIDE algorithm to simulate the mechanism of tumor immune escape to predict the therapeutic effect of patients for immune checkpoint blockers. Finally, we tested the predictive effectiveness of ECMS through the Imvigor210 cohort.

Cell lines

The OC cell lines A2780 and SKOV3 and the normal ovarian epithelial cell line IOSE-80 were purchased from iCell Bioscience Inc. All the cells were cultured in DMEM with 10% FBS (Biological, Israel).

RT-qPCR

RNA was extracted using the RNeasy Mini Kit (QIAGEN). The HiScript II Q RT SuperMix for qPCR Kit (Vazyme, China) was used for reverse transcription. ChamQ Universal SYBR qPCR Master Mix (Vazyme, China) was used for RT-qPCR. GAPDH was the housekeeper gene. The results were calculated using the $2^{-\Delta\Delta CT}$ method.

Statistical analysis

We utilized R software (version 4.04) to conduct all statistical analyses and graphs. The Wilcoxon test was utilized to measure the differences between the two ECMS groups. Moreover, the chi-square test was applied to compare the differences in proportions. We used a KM plotter to generate survival curves and assessed the differences by log-rank test. We applied the R package "survivalROC" to plot tROC and evaluated the predictive power utilizing the area under the curve (AUC). We applied the R package "survival" to conduct the univariate and multivariate Cox regression analyses and "rms" to plot the nomogram and calibration curves. All tests were two-tailed, and $p< 0.05$ was considered statistically significant if not otherwise stated.

Results

Identification of the candidate ECM genes

A total of three OC cohorts (TCGA-OV, ICGC-OV, and Imvigor210) were considered suitable for this study. We collected 1,028 ECM genes from the hallmark dataset on the MSIDGB website. The WGCNA algorithm was applied to determine ECM-associated genes. The scale-free network was constructed with the scale-free topology fitting index set to 0.9, and the corresponding optimal soft threshold value was 8 (Figure 1A). We used a clustering dendrogram to identify 46 modules (Figure 1B). The correlation coefficient between the Darkorange2 module and ImmuneScore was 0.79, and the correlation coefficient between the Darkorange2 module and ESTIMATEScore was 0.8, suggesting that the Darkorange2 module was selectively expressed in samples with high immune cell infiltration (Figure 1C). The 1,028 ECM genes and 669 genes from the most relevant gene modules were intersected to obtain 61 candidate ECM genes (Figure 1D). These 61 candidate ECM genes were screened for independent prognostic factors by doing univariate Cox regression analysis, and we identified 10 genes (Figure 1E). To comprehensively analyze these genes, we used Metascape for functional enrichment analysis. We listed the top 20 enrichment terms in which candidate ECM genes were mostly enriched in NABA MATRISOME-ASSOCIATED signaling pathways (Figure 1F).

Construction and validation of the ECMS model

It was found that the risk model containing eight genes was the best one (Figure 2A). Detailed information on the eight genes is shown in Table S2. As we know, C-index is used to assess prediction capacity and

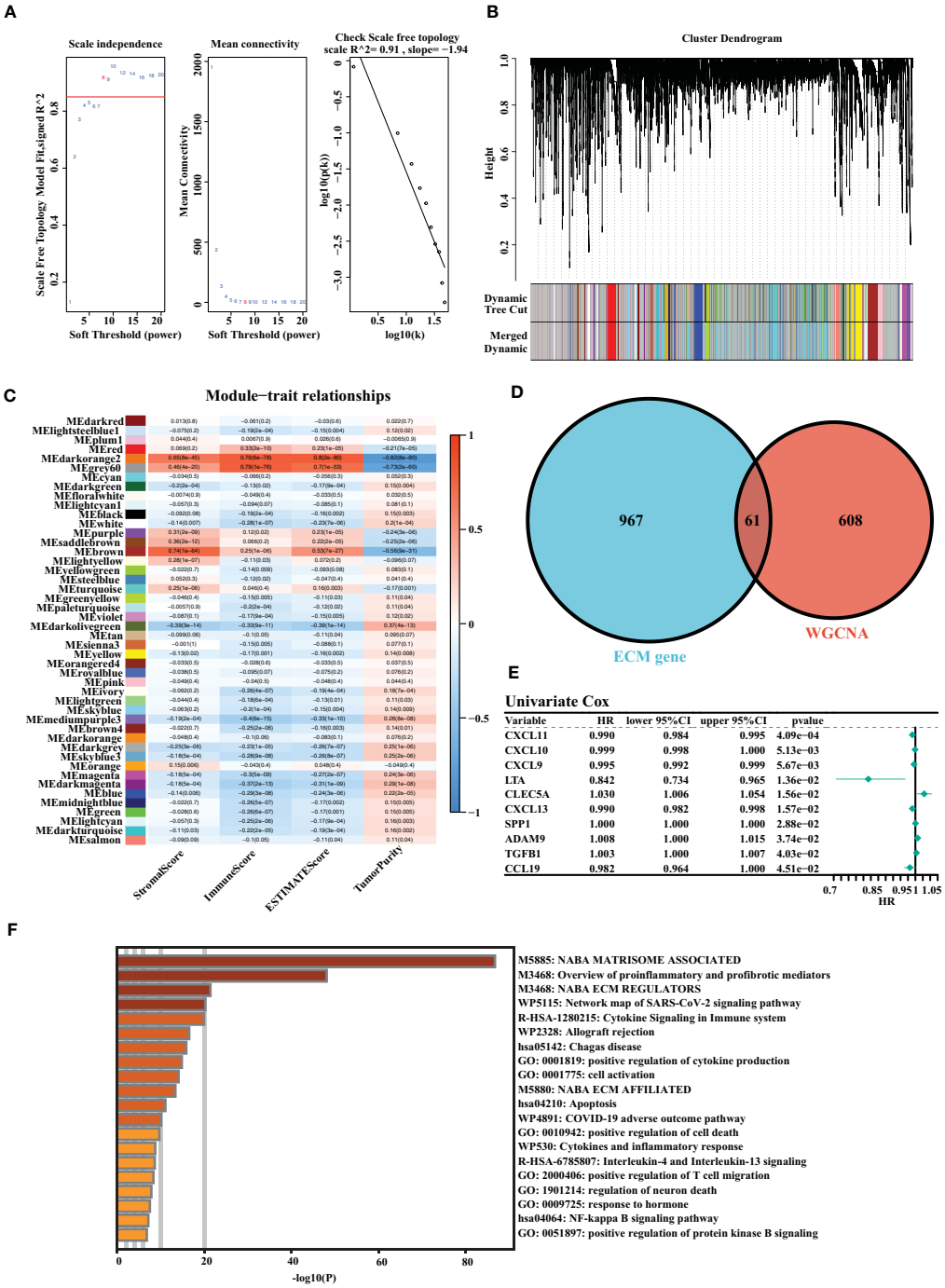


FIGURE 1
Identification of the highly valuable ECM-related genes. **(A)** The relation between the scale-free topology fit index and soft threshold. **(B)** Gene modules identified by cluster dendrogram. **(C)** Correlation analysis of modules with StromalScore, ImmuneScore, ESTIMATEScore, and TumorPurity. **(D)** Venn diagram of the 1,028 ECM genes and 661 ECM-associated genes from WGCNA. **(E)** Forest plots presenting univariate Cox regression analyses of candidate ECM genes as independent prognostic factors. **(F)** Enrichment analysis of the candidate ECM genes.

reliability. The C-indexes were 0.603 (training set) and 0.597 (testing set). The details are shown in Figure 2A. We constructed the risk model containing eight genes based on the optimal λ value of 0.01339134 (Figure 2B). The survival analysis demonstrated that the high-risk group had shorter OS than the low-risk group in the training set [hazard ratio (HR) = 2 (1.53–2.61), p< 0.001, Figure 2C]. Moreover, the testing set showed similar results [HR = 1.62 (1.06–2.47), p = 0.021, Figure 2D]. To further test the validity of the ECMS, we performed ROC analysis on the training and testing sets. We used the AUC analysis to assess the reliability of our signature. The AUC values of 0.528, 0.594, and 0.67 for the training set (Figure 2E) at 1, 3, and 5 years, respectively, and 0.571, 0.635, and 0.684 for the testing set (Figure 2F), respectively. The tROC analysis indicated that ECMS was a reliable predictor for OC patients (Figures 2G, H).

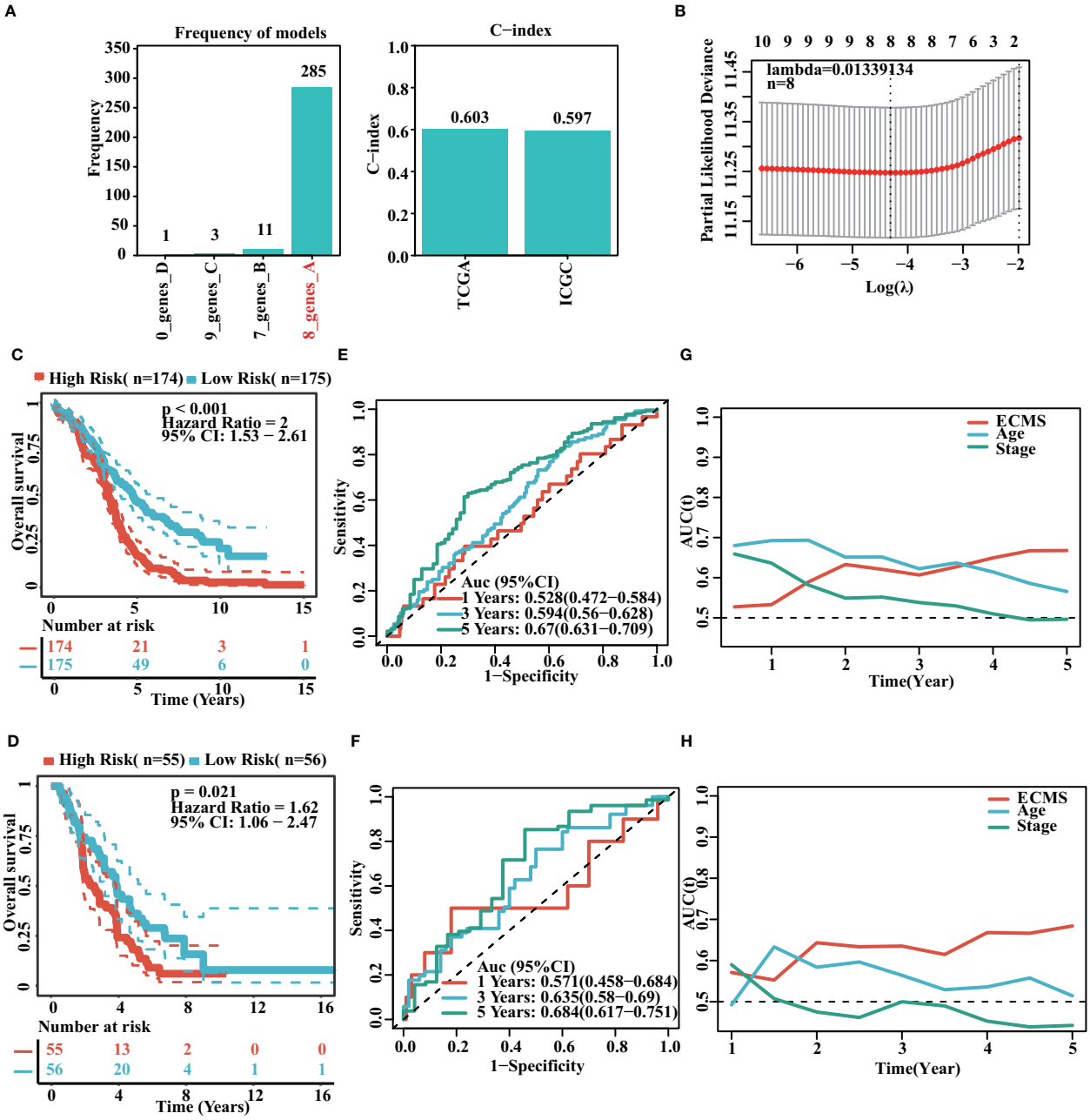


FIGURE 2
Construction of risk model. **(A)** Frequency of models and C-indices of training and testing sets. **(B)** The selection of the tuning parameter (λ) in the LASSO model. **(C, D)** Survival curves of the training **(C)** and testing sets **(D)**. **(E, F)** ROC curves of ECMS in the training **(E)** and testing sets **(F–H)**. Time-dependent AUC values for the ECMS and clinical characteristics in the training **(G)** and testing sets **(H)**.

We analyzed the association between age, stage, ECMS, and prognosis. The ECMS was an independent risk factor according to the univariate Cox regression analysis [HR = 3.243 (2.141–4.913), p< 0.001 in the training set; HR = 5.410 (2.031–14.413), p< 0.001 in the testing set, Figure 3A]. In multivariate Cox regression analysis, ECMS also exhibited an excellent prognostic performance [HR = 3.132 (2.068–4.744), p< 0.001 in the training set; HR = 5.514 (2.084–14.586), p< 0.001 in the testing set, Figure 3B]. We constructed a nomogram to assess the survival probability for OC patients (Figure 3C). The calibration curve analysis indicated this nomogram was accurate (Figure 3D). In addition, the tROC analysis revealed that the nomogram outperformed other variables (Figure 3E).

Enrichment analysis

We obtained DEGs and input these genes into Metascape. It was observed that the genes elevated in the high ECMS group were significantly related to trans-synaptic signaling, heart development, regulation of synaptic plasticity, presynapse assembly, and intermediate filament organization (Figure 4A). Moreover, the

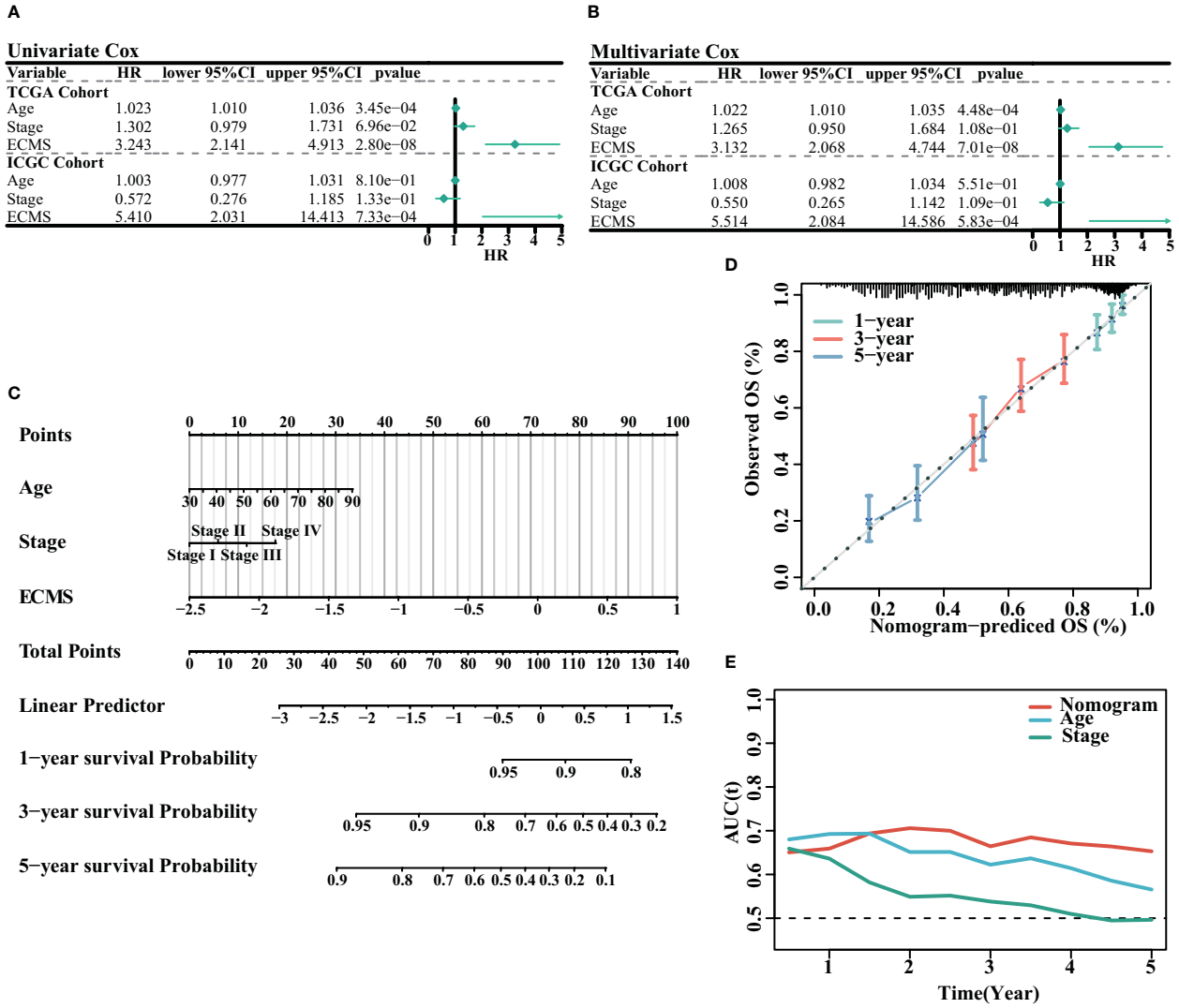


FIGURE 3
Assessment of the prognostic value of ECMS. **(A, B)** Forest plots of univariate regression analyses **(A)** and multivariable regression analyses **(B)** involving the ECMS, age, and stage. **(C)** Nomogram showing the risk assessment of patients. **(D)** Nomogram calibration curves predicting the OS at 1, 3, and 5 years. **(E)** Time-dependent AUC values of nomogram, age, and stage.

genes elevated in the low ECMS group were notably connected with allograft rejection, regulation of lymphocyte activation, positive regulation of immune response, extrafollicular B-cell activation by SARS-CoV-2, and adaptive immune response (Figure 4B). In addition, functional enrichment analyses in the high ECMS group showed that adherens junction, ECM receptor interactions, mitogen-activated protein kinase (MAPK) signaling pathway, pathways in cancer, and vascular endothelial growth factor (VEGF) signaling pathway were enriched (Figure 4C). In contrast, in the low ECMS group, antigen processing and presentation, asthma, natural killer cell-mediated cytotoxicity, oxidative phosphorylation, and primary immunodeficiency were mainly enriched (Figure 4D). We also performed an enrichment analysis on the validation set. The results revealed that elevated genes were mainly associated with spliceosome in the high ECMS group (Figure S1A) and were mainly related to inflammatory response, neutrophil degranulation, positive regulation of cytokine production, phagosome, and osteoclast differentiation in the low ECMS group (Figure S1B). The GSEA revealed that in the high ECMS group, adherens junction, gap junction, MAPK signaling pathway, o-glycan biosynthesis, and pathways in cancer were mostly enriched (Figure S1C), whereas oxidative phosphorylation, primary immunodeficiency, protein export, and ribosome were enriched in the low ECMS group (Figure S1D).

Immune landscape

ssGSEA showed a significant difference in ImmuneScore between the two ECMS groups, with the low ECMS group exhibiting higher immune activity (Figure 5A). In addition, we selected CTLA-4, T-cell immunoglobulin and mucin domain 3 (TIM-3), PD-1, PD-L1, PD-L2, and lymphocyte-activation gene 3 (LAG3) as biomarkers of immune checkpoint activity. We analyzed the differences of their expression

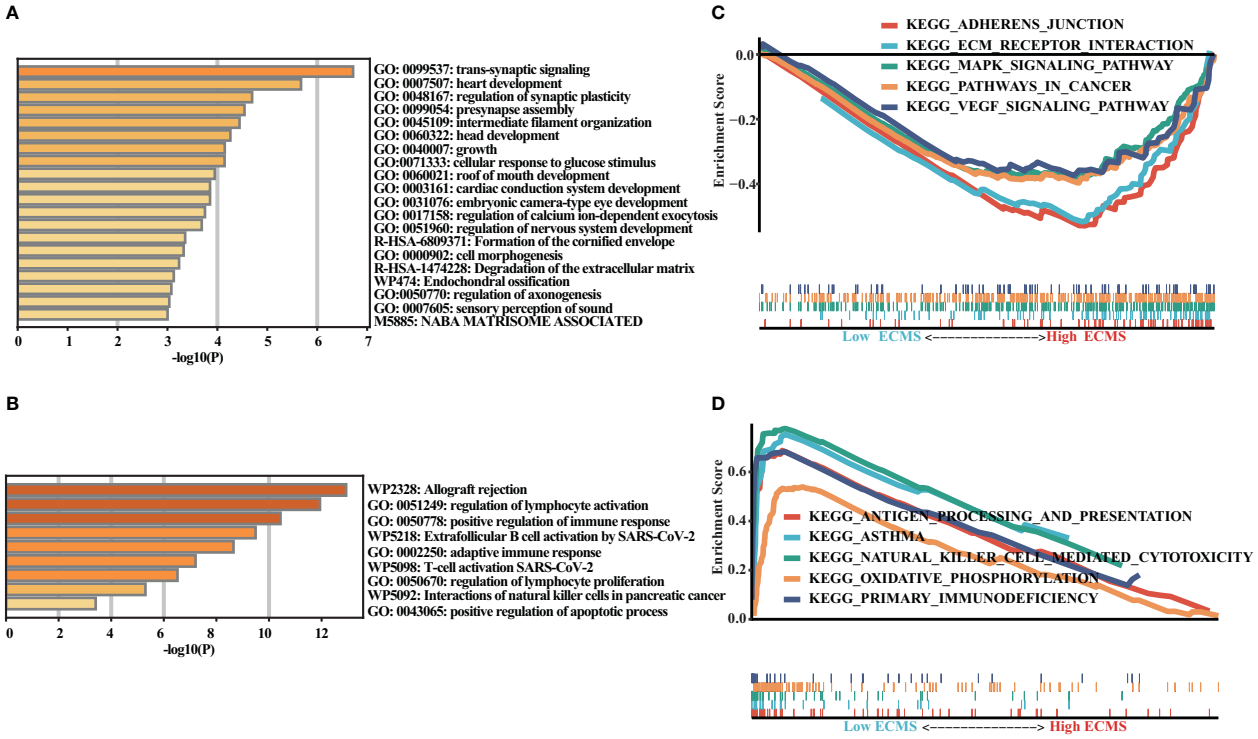


FIGURE 4
Functional enrichment analysis. **(A, B)** Enrichment analysis of the DEGs in the high **(A)** and low **(B)** ECMS groups; **(C, D)** Enrichment plots of GSEA in the high **(C)** and low **(D)** ECMS group.

between two ECMS groups. We found that the expression was notably more active in the low ECMS group (Figure 5B). Then, we analyzed the correlation between ECMS and enrichment scores and the relationship between ECMS and differences in immune checkpoint expression, as shown in Figure 5C. Subsequently, we assessed immune cell infiltration fraction and pathway activity in two groups. The activity of most immune pathways was notably lower in the high ECMS group (Figure 5D). A significant difference was seen between the two ECMS groups in the infiltration degree of most immune cells (e.g., T cells, macrophages, mast cells), as shown in Figure 5E. We also performed ssGSEA on the validation set to assess the immune-related pathways' activity. The enrichment score between both groups was not significantly different (Figure S2A), but there were differences in the expression of LAG3 and TIM-3 (Figure S2B). The heat map revealed the correlation between ECMS and enrichment score and the expression difference of immune checkpoint (Figure S2C).

Prediction of immunotherapy response

Neoantigen is one of the biomarkers of immunotherapy, which can guide the application of immunotherapy. We analyzed the correlation between indel neoantigens, SNV neoantigens, neoantigens, and ECMS, and the results are presented in Figures 6A, B. There were significant negative correlations between SNV neoantigens and ECMS (R = -0.46, p< 0.0001, Figure 6B), whereas no correlation was observed between indel neoantigens and ECMS (p = 0.23, Figure 6A). IPS can be used to assess the response to immunotherapy. The IPS of patients in TCGA and ICGC cohorts are shown in Figure 6C and Figure S3A. In addition, the response rate to anti-PD-L1 immunotherapy in the training set was higher in the low ECMS group (p = 0.03) (Figure 6D). The testing set presented similar results (p = 0.01) (Figure S3B). Then, we performed a ROC analysis on TCGA set, and the AUC value was 0.566 (Figure 6E). In contrast, the AUC value in the IGCA set was 0.572 (Figure S3C), indicating that ECMS was a more reliable predictor than other commonly used indicators. We also found that the high ECMS group had shorter OS than the low ECMS group in the Imvigor210 cohort (hazard ratio = 1.39, p = 0.022, Figure 6F). Moreover, we also found a significant negative correlation between neoantigens and ECMS (Figure 6G).

Verify the expression of ECMS genes in ovarian cancer cell lines

We evaluated the risk coefficients of genes in the ECMS model. Among them, CLEC5A is the strongest risk factor whereas LTA is the strongest protective factor (Figure 7A). Then, we performed RT-qPCR to verify our result (Figure 7B). *CLEC5A*, *ADAM9*, and *TGFB1* were highly expressed in OC cell lines compared with the normal ovarian epithelial cell line, whereas *LTA*, *CCL19*, *CXCL11*, and *CXCL9* were downregulated in OC cell lines. However, the expression level of SPP1 showed no difference between normal and malignant ovarian epithelial cell lines (Figure 7B).

10.3389/fonc.2023.1163695

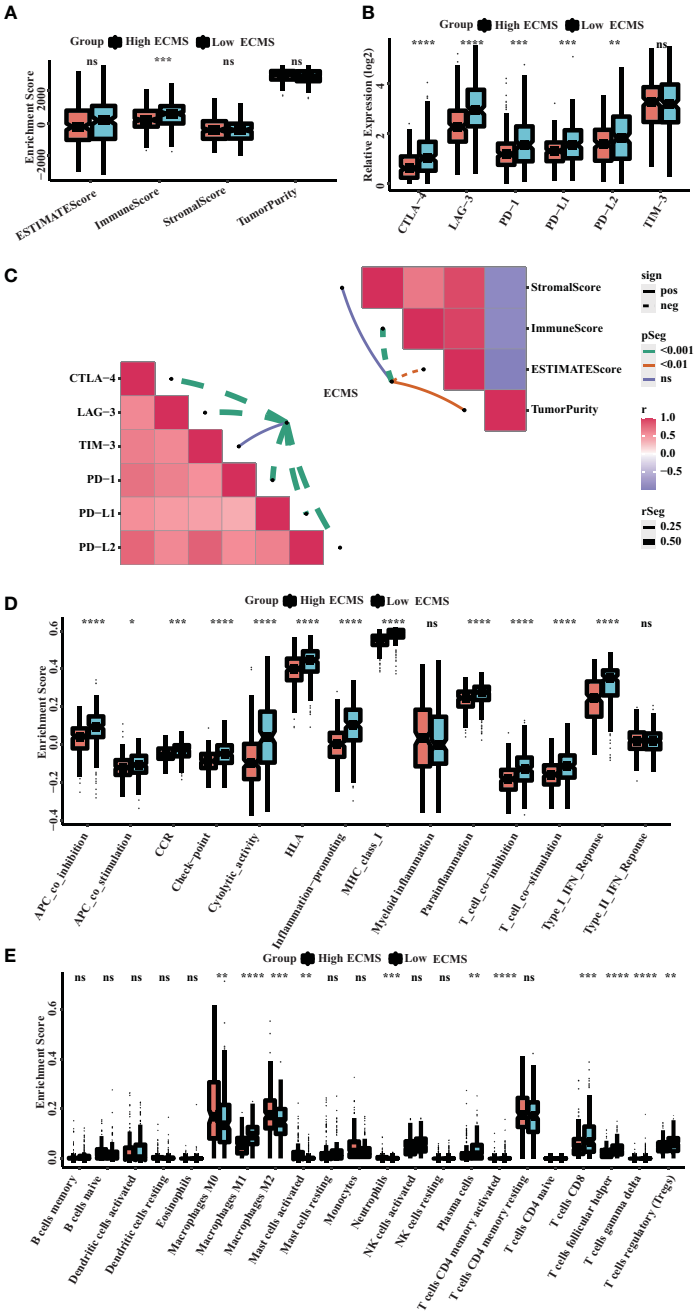


FIGURE 5
Immune landscape in the high and low ECMS group. **(A)** Box plot of the enrichment score. **(B)** Box plot of the immune checkpoints expression. **(C)** Heatmap of the correlation between the ECMS and enrichment score, as well as immune checkpoint expression, respectively. **(D)** Box plots of the immune pathway activity. **(E)** Box plots of the proportion of immune cell infiltration. *p< 0.05; **p< 0.01; ***p< 0.001; ****p< 0.0001; ns, no significance.

Discussion

In the present research, we explored the role of ECM in OC patients and linked them for the first time with the prognosis and effectiveness of immunotherapy. Our results suggested that the ECMS model performed well. The AUC values for the training set at 1, 3, and 5 years were 0.528, 0.594, and 0.67, respectively. We also found that the high ECMS group had shorter OS than the low ECMS group. In addition, the immune landscape demonstrated that the immune checkpoints' expression was more active in the low ECMS group, and the response rate to anti-PD-L1 treatment was lower in the high ECMS group. The ROC values of the ECMS model for predicting immune response were 0.566 (validation set) and 0.572 (testing set), indicating that the model could predict, to some extent, the response rate to immunotherapy.

The ECM plays a role in regulating cell growth, motility, and differentiation (6). The most widely known ECM alteration in

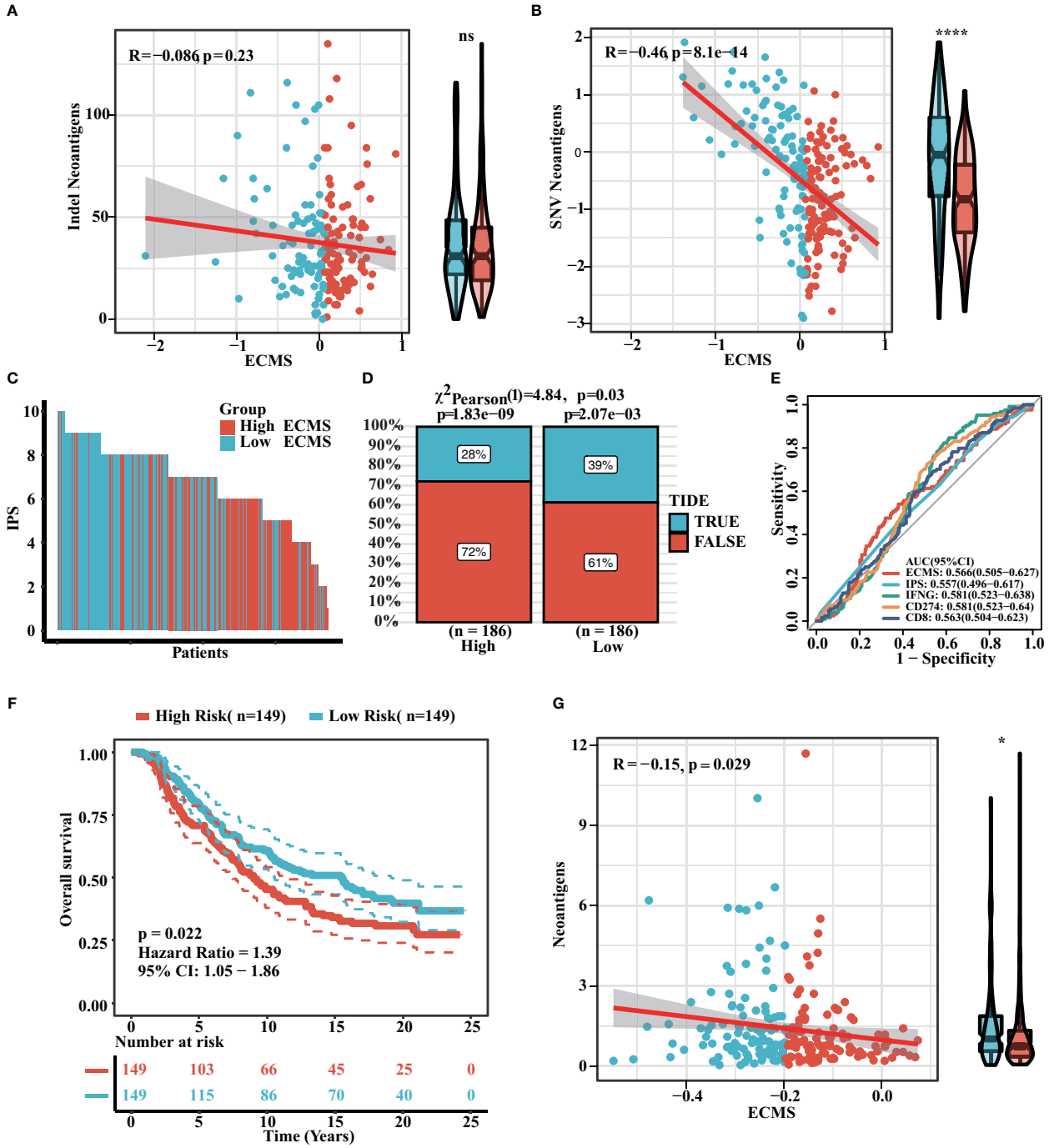


FIGURE 6
Prediction of immunotherapy response. **(A)** Correlation between indel neoantigens and ECMS. **(B)** Correlation between SNV neoantigens and ECMS. **(C)** IPS of the patients in TCGA cohort. **(D)** Distribution of the response rate to anti-PD-L1 immunotherapy in the high- and low-ECMS group. **(E)** ROC curves of ECMS, IPS, IFNG, CD274, and CD8 in TCGA cohort. **(F)** Survival curves of the Imvigor210 cohort. **(G)** Correlation between the neoantigens and ECMS. *p< 0.05; ****p< 0.0001; ns, no significance.

tumor tissue is increased collagen deposition. The increased collagen deposition affects the properties of the TME, thereby modulating cancer cell polarity, migration, and signaling transduction (24–27). Previous studies have shown that increased expression of proteins mediating ECM remodeling can increase mortality in patients with BC, lung cancer, or gastric cancer (GC) (28, 29). In addition, histological studies have observed excessive ECM deposition and remodeling in OC. The fibrosis rich in COL6/collagen VI and fibronectin is already present around the micro metastases, which develops into an extensive connective tissue proliferative TME as the disease progress. COL6 is involved in tumor growth and apoptosis escape in early metastases of OC (30). These findings confirm that ECM is closely associated with the clinical manifestations and prognosis of OC.

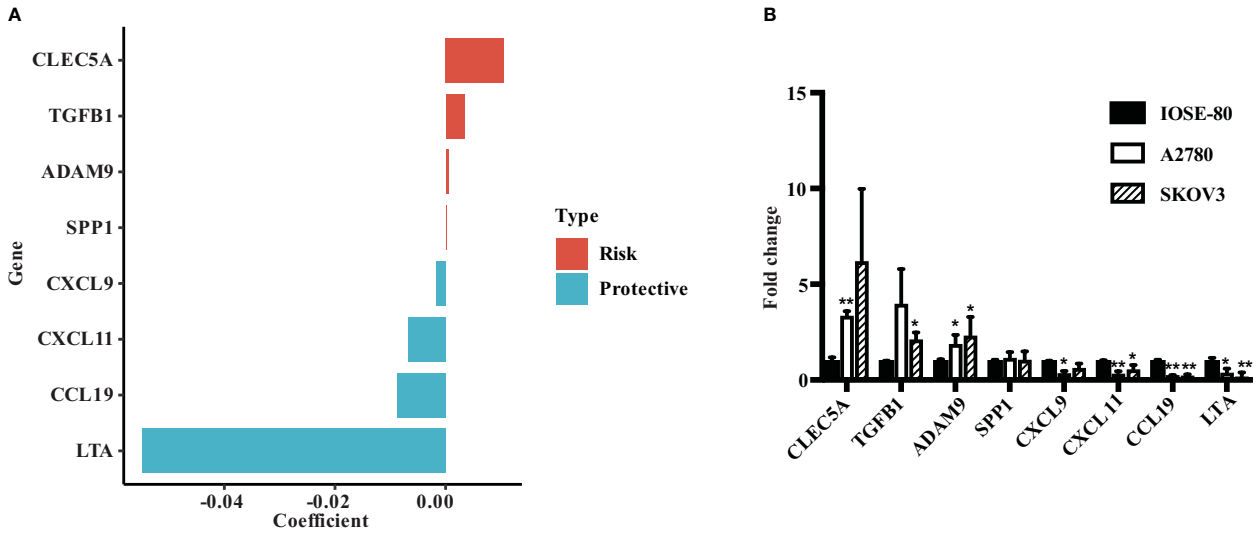


FIGURE 7
Verify the expression of ECMS genes in ovarian cancer cell lines. **(A)** The coefficient index of genes in the ECMS model. **(B)** The expression levels of ECMS genes in OC and ovarian epithelial cell lines detected by qRT-PCR. *p < 0.05; **p < 0.01.

It was observed that candidate ECM genes were mostly enriched in the NABA MATRISOME-ASSOCIATED signaling pathway, which had been mentioned to be associated with tumor development (31). The GSEA was performed, and the results indicated that the high ECMS was enriched in adherens junction, ECM receptor interactions, MAPK signaling pathway, pathway in cancer, and VEGF signaling pathway. Adherens junction is cell–cell adhesion complexes that take part in embryogenesis and tissue homeostasis (32). ECM receptor interactions regulate cell behavior and are vital in cell proliferation, adhesion, and migration (33). The MAPK signaling pathway regulates various cellular processes, such as cell proliferation and differentiation (34). The VEGF signaling pathway is a major regulator of angiogenesis and vascular permeability (35). These pathways are involved in tumorigenesis, progression, invasion, and metastasis (36–39). The results indicated that the tumors were developing and metastasizing. In the low ECMS group, enrichment of immune-related pathways such as antigen processing and presentation, asthma, oxidative phosphorylation, natural killer cell-mediated cytotoxicity, and primary immunodeficiency were observed. This suggested that low ECMS patients presented powerful immune function.

We also found that the high ECMS group had shorter OS than the low ECMS group. It could be explained by the enrichment of tumor-related signaling pathways and active tumor growth in the high ECMS group. On the other hand, the low ECMS group was enriched in antigen processing and signal presentation pathways, which showed stronger immune function, thus helping the body to clear tumor cells. Furthermore, the ECMS model was based on candidate ECM genes mainly enriched in NABA MATRISOME-ASSOCIATED signaling pathways associated with tumor development. Thus, this could also explain why the high ECMS group had shorter OS time. Similar results were also seen in the study by Liu et al., who classified BC patients into two groups based on the ECM index (ECMI), which was based on ECM-associated immunogens, and assessed their clinical, biological, and genomic characteristics. The researchers believed the low ECMI group had significantly improved OS (40). In addition, Yang et al. used a gene set variation analysis algorithm to establish ECM scores, and higher ECM scores predicted poor prognosis in GC (41). Similarly, Ding et al. established a new immune-related signature to stratify the risk of OC patients and then predict the prognosis (7). Considering the role of ECM in OC, our study constructed a risk model on the basis of ECM, and the results suggested that ECMS can well predict the prognosis of OC patients.

Tumor-associated ECM may have immunomodulatory effects, influencing antitumor immunity by controlling the localization and migration of immune cells (42). Thus, ECM may influence the effect of immunotherapy. Indeed, previous studies have proved that combination therapy targeting the immune and stromal microenvironment had better therapeutic effects (43, 44). Therefore, we developed an ECMS model to predict the patients' responses to immunotherapy. The high ECMS group was observed to have lower immune pathway activity. This indicated that ECM might affect the immune regulation of OC. Then, we selected CTLA-4, LAG3, PD-1, PD-L1, PD-L2, and TIM-3 as markers of the immune checkpoint. A significant difference was seen between the two ECMS groups, with more active expression in the low ECMS group. We also found that the response rate to anti-PD-L1 immunotherapy was lower in the high ECMS group in both the validation and testing sets. This could be due to the higher immune pathway activity and more active expression of immune checkpoints in the low ECMS group, so it has a higher response rate to the immune checkpoint inhibitors. Similarly, Mao et al. established a stromal score and investigated the relationship between immunotherapy-related markers or immune cell types and the stromal score in GC. The results of this study also confirmed that the stroma was related to immunotherapy-related markers (45). This is the first study that proposes the role of ECM in predicting immunotherapeutic response in OC. Our study confirmed the association between ECM and the immune-related pathway of OC

and the immunotherapeutic response. Improving the OS of OC patients is a common problem in advanced OC. Precision therapy is a good entry point. Our study revealed that the low ECMS group had higher immune pathway activity, the more active expression of the immune checkpoint, and a higher response rate to immune checkpoint inhibitors. Therefore, we can identify OC patients who may benefit from immunotherapy based on the established ECMS model and help clinicians and patients to make individualized treatment decisions.

Nevertheless, this research has several limitations. First, our data are from TCGA, ICGC, and publicly available immunotherapy cohorts, which need to be verified with large samples in reality. Second, the immunotherapy cohort is an advanced uroepithelial cancer cohort with PD-L1 immunotherapy (Imvigor210), and further validation in an OC immunotherapy cohort is needed in the future. Third, the ECMS model needs to be authenticated in reality before application.

In conclusion, we created an ECMS model to predict the prognosis and immunotherapeutic benefits in OC patients and provided references for individualized treatment of OC 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.

Author contributions

YH, LS: designed the study, analyzed the data and wrote the manuscript. YL: established the methods. JY: reviewed and revised the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

We thank the founders who built and developed the R package and the software.

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.2023.1163695/full#supplementary-material

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin* (2022) 72(1):7–33. doi: 10.3322/caac.21708

2. Bast RC Jr., Hennessy B, Mills GB. The biology of ovarian cancer: new opportunities for translation. *Nat Rev Cancer* (2009) 9(6):415–28. doi: 10.1038/nrc2644

3. Peres LC, Cushing-Haugen KL, Köbel M, Harris HR, Berchuck A, Rossing MA, et al. Invasive epithelial ovarian cancer survival by histotype and disease stage. *J Natl Cancer Inst* (2019) 111(1):60–8. doi: 10.1093/jnci/djy071

4. Egen JG, Ouyang W, Wu LC. Human anti-tumor immunity: insights from immunotherapy clinical trials. *Immunity* (2020) 52(1):36–54. doi: 10.1016/j.immuni.2019.12.010

5. Hua T, Wang RM, Zhang XC, Zhao BB, Fan SB, Liu DX, et al. ZNF76 predicts prognosis and response to platinum chemotherapy in human ovarian cancer. *Biosci Rep* (2021) 41(12). doi: 10.1042/BSR20212026

6. Bao M, Zhang L, Hu Y. Novel gene signatures for prognosis prediction in ovarian cancer. *J Cell Mol Med* (2020) 24(17):9972–84. doi: 10.1111/jcmm.15601

7. Ding J, Zhang Q, Chen S, Huang H, He L. Construction of a new tumor immunity-related signature to assess and classify the prognostic risk of ovarian cancer. *Aging (Albany NY)* (2020) 12(21):21316–28. doi: 10.18632/aging.103868

8. Ding Q, Dong S, Wang R, Zhang K, Wang H, Zhou X, et al. A nine-gene signature related to tumor microenvironment predicts overall survival with ovarian cancer. *Aging (Albany NY)* (2020) 12(6):4879–95. doi: 10.18632/aging.102914

9. Hynes RO. The extracellular matrix: not just pretty fibrils. *Science* (2009) 326 (5957):1216–9. doi: 10.1126/science.1176009

10. Lu P, Weaver VM, Werb Z. The extracellular matrix: a dynamic niche in cancer progression. *J Cell Biol* (2012) 196(4):395–406. doi: 10.1083/jcb.201102147

11. 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(1):5077. doi: 10.1038/s41467-020-18916-5

12. Tian C, Clauser KR, Öhlund D, Rickelt S, Huang Y, Gupta M, et al. Proteomic analyses of ECM during pancreatic ductal adenocarcinoma progression reveal different contributions by tumor and stromal cells. *Proc Natl Acad Sci U.S.A.* (2019) 116(39):19609–18. doi: 10.1073/pnas.1908626116

13. Pickup MW, Mouw JK, Weaver VM. The extracellular matrix modulates the hallmarks of cancer. *EMBO Rep* (2014) 15(12):1243–53. doi: 10.15252/embr.201439246

14. Nicolas-Boluda A, Vaquero J, Vimeux L, Guilbert T, Barrin S, Kantari-Mimoun C, et al. Tumor stiffening reversion through collagen crosslinking inhibition improves T cell migration and anti-PD-1 treatment. *Elife* (2021) 10. doi: 10.7554/eLife.58688.sa2

15. Shen Y, Wang X, Lu J, Salfenmoser M, Wirsik NM, Schleussner N, et al. Reduction of liver metastasis stiffness improves response to bevacizumab in metastatic colorectal cancer. *Cancer Cell* (2020) 37(6):800–817.e7. doi: 10.1016/j.ccell.2020.05.005

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

17. McDermott DF, Huseni MA, Atkins MB, Motzer RJ, Rini BI, Escudier B, et al. Clinical activity and molecular correlates of response to atezolizumab alone or in combination with bevacizumab versus sunitinib in renal cell carcinoma. *Nat Med* (2018) 24(6):749–57. doi: 10.1038/s41591-018-0053-3

18. Gibbons DL, Creighton CJ. Pan-cancer survey of epithelial-mesenchymal transition markers across the cancer genome atlas. *Dev Dyn* (2018) 247(3):555–64. doi: 10.1002/dvdy.24485

19. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest* (2017) 127(8):2930–40. doi: 10.1172/JCI91190

20. Liang JY, Wang DS, Lin HC, Chen XX, Yang H, Zheng Y, et al. A novel ferroptosis-related gene signature for overall survival prediction in patients with

hepatocellular carcinoma. *Int J Biol Sci* (2020) 16(13):2430–41. doi: 10.7150/ijbs.45050

21. 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 (5):453–7. doi: 10.1038/nmeth.3337

22. Yoshihara K, Shahmoradgoli M, Martínez 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

23. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. *Immunity* (2018) 48(4):812–830.e14. doi: 10.1016/j.immuni.2018.03.023

24. Fraley SI, Feng Y, Krishnamurthy R, Kim DH, Celedon A, Longmore GD, et al. A distinctive role for focal adhesion proteins in three-dimensional cell motility. *Nat Cell Biol* (2010) 12(6):598–604. doi: 10.1038/ncb2062

25. Fraley SI, Feng Y, Giri A, Longmore GD, Wirtz D. Dimensional and temporal controls of three-dimensional cell migration by zyxin and binding partners. *Nat Commun* (2012) 3:719. doi: 10.1038/ncomms1711

26. Paszek MJ, Zahir N, Johnson KR, Lakins JN, Rozenberg GI, Gefen A, et al. Tensional homeostasis and the malignant phenotype. *Cancer Cell* (2005) 8(3):241–54. doi: 10.1016/j.ccr.2005.08.010

27. Levental KR, Yu H, Kass L, Lakins JN, Egeblad M, Erler JT, et al. Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell* (2009) 139 (5):891–906. doi: 10.1016/j.cell.2009.10.027

28. Chang HY, Nuyten DS, Sneddon JB, Hastie T, Tibshirani R, Sørlie T, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc Natl Acad Sci U.S.A.* (2005) 102(10):3738–43. doi: 10.1073/pnas.0409462102

29. Chang HY, Sneddon JB, Alizadeh AA, Sood R, West RB, Montgomery K, et al. Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. *PloS Biol* (2004) 2(2):E7. doi: 10.1371/journal.pbio.0020007

30. Pietilä EA, Gonzalez-Molina J, Moyano-Galceran L, Jamalzadeh S, Zhang K, Lehtinen L, et al. Co-Evolution of matrisome and adaptive adhesion dynamics drives ovarian cancer chemoresistance. *Nat Commun* (2021) 12(1):3904. doi: 10.1038/s41467-021-24009-8

31. Naba A, Clauser KR, Hoersch S, Liu H, Carr SA, Hynes RO. The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Mol Cell Proteomics* (2012) 11(4):M111.014647. doi: 10.1074/mcp.M111.014647

32. Wei M, Ma Y, Shen L, Xu Y, Liu L, Bu X, et al. NDRG2 regulates adherens junction integrity to restrict colitis and tumourigenesis. *EBioMedicine* (2020) 61:103068. doi: 10.1016/j.ebiom.2020.103068

33. Nersisyan S, Novosad V, Engibaryan N, Ushkaryov Y, Nikulin S, Tonevitsky A, et al. ECM-receptor regulatory network and its prognostic role in colorectal cancer. *Front Genet* (2021) 12:782699. doi: 10.3389/fgene.2021.782699

34. Park HB, Baek KH. E3 ligases and deubiquitinating enzymes regulating the MAPK signaling pathway in cancers. *Biochim Biophys Acta Rev Cancer* (2022) 1877 (3):188736. doi: 10.1016/j.bbcan.2022.188736

35. Eremina V, Baelde HJ, Quaggin SE. Role of the VEGF–a signaling pathway in the glomerulus: evidence for crosstalk between components of the glomerular filtration barrier. *Nephron Physiol* (2007) 106(2):p32–7. doi: 10.1159/000101798

36. Liu SR, Yang X, Qi L, Zhu Z, Ji YZ. SMARCA4 promotes benign skin malignant transformation into melanoma through adherens junction signal transduction. *Clin Transl Oncol* (2021) 23(3):591–600. doi: 10.1007/s12094-020-02453-0

37. 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 Lett* (2019) 24:38. doi: 10.1186/s11658-019-0162-0

38. Yang M, Huang CZ. Mitogen-activated protein kinase signaling pathway and invasion and metastasis of gastric cancer. *World J Gastroenterol* (2015) 21(41):11673–9. doi: 10.3748/wjg.v21.i41.11673

39. Apte RS, Chen DS, Ferrara N. VEGF in signaling and disease: beyond discovery and development. *Cell* (2019) 176(6):1248–64. doi: 10.1016/j.cell.2019.01.021

40. Liu J, Lei B, Yu X, Li Y, Deng Y, Yang G, et al. Combining immune-related genes for delineating the extracellular matrix and predicting hormone therapy and neoadjuvant chemotherapy benefits in breast cancer. *Front Immunol* (2022) 13:888339. doi: 10.3389/fimmu.2022.888339

41. Yang Z, Xue F, Li M, Zhu X, Lu X, Wang C, et al. Extracellular matrix characterization in gastric cancer helps to predict prognosis and chemotherapy response. *Front Oncol* (2021) 11:753330. doi: 10.3389/fonc.2021.753330

42. Salmon H, Franciszkiewicz K, Damotte D, Dieu-Nosjean MC, Validire P, Trautmann A, et al. Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *J Clin Invest* (2012) 122 (3):899–910. doi: 10.1172/JCI45817

43. Ravi R, Noonan KA, Pham V, Bedi R, Zhavoronkov A, Ozerov IV, et al. Bifunctional immune checkpoint-targeted antibody-ligand traps that simultaneously disable TGFβ enhance the efficacy of cancer immunotherapy. *Nat Commun* (2018) 9 (1):741. doi: 10.1038/s41467-017-02696-6

44. Lan Y, Zhang D, Xu C, Hance KW, Marelli B, Qi J, et al. Enhanced preclinical antitumor activity of M7824, a bifunctional fusion protein simultaneously targeting PD-L1 and TGF-β. *Sci Transl Med* (2018) 10(424). doi: 10.1126/scitranslmed.aan5488

45. Mao M, Yu Q, Huang R, Lu Y, Wang Z, Liao L, et al. Stromal score as a prognostic factor in primary gastric cancer and close association with tumor immune microenvironment. *Cancer Med* (2020) 9(14):4980–90. doi: 10.1002/cam4.2801

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 09 May 2023
DOI 10.3389/fonc.2023.1177120

OPEN ACCESS

EDITED BY
Nan Zhang,
Harbin Medical University, China

REVIEWED BY
Yi Zhuo,
Hunan Normal University, China
Lei Wang,
Hunan Cancer Hospital, Central South University, China

*CORRESPONDENCE
Zhengtao Yu
yuzhengtao22@163.com

RECEIVED 01 March 2023
ACCEPTED 11 April 2023
PUBLISHED 09 May 2023

CITATION
Li Y, Li X and Yu Z (2023) Novel methylation-related long non-coding RNA clinical outcome prediction method: the clinical phenotype and immune infiltration research in low-grade gliomas. *Front. Oncol.* 13:1177120. doi: 10.3389/fonc.2023.1177120

Novel methylation-related long non-coding RNA clinical outcome prediction method: the clinical phenotype and immune infiltration research in low-grade gliomas

Youjun Li, Xiaobo Li and Zhengtao Yu*

Department of Neurosurgery, Affiliated Haikou Hospital of Xiangya School of South University, Haikou, Hainan, China

Background: Recent studies have suggested that long non-coding RNAs (lncRNAs) may play crucial role in low-grade glioma; however, the underlying mechanisms linking them to epigenetic methylation remain unclear.

Methods: We downloaded expression level data for regulators associated with N1 methyladenosine (m1A), 5-methyladenine (m5C), and N6 methyladenosine (m6A) (M1A/M5C/M6A) methylation from the Cancer Genome Atlas-low-grade glioma (TCGA-LGG) database. We identified the expression patterns of lncRNAs, and selected methylation-related lncRNAs using Pearson correlation coefficient>0.4. Non-negative matrix dimensionality reduction was then used to determine the expression patterns of the methylation-associated lncRNAs. We constructed a weighted gene co-expression network analysis (WGCNA) network to explore the co-expression networks between the two expression patterns. Functional enrichment of the co-expression network was performed to identify biological differences between the expression patterns of different lncRNAs. We also constructed prognostic networks based on the methylation presence in lncRNAs in low-grade gliomas.

Results: We identified 44 regulators by literature review. Using a correlation coefficient greater than 0.4, we identified 2330 lncRNAs, among which 108 lncRNAs with independent prognostic values were further screened using univariate Cox regression at P< 0.05. Functional enrichment of the co-expression networks revealed that regulation of trans-synaptic signaling, modulation of chemical synaptic transmission, calmodulin binding, and SNARE binding were mostly enriched in the blue module. The calcium and CA2 signaling pathways were associated with different methylation-related long non-coding chains. Using the Least Absolute Shrinkage Selector Operator (LASSO) regression analysis, we analyzed a prognostic model containing four lncRNAs. The model's risk score was 1.12 *AC012063 + 0.74 * AC022382 + 0.32 * AL049712 + 0.16 * GSEC. Gene set variation analysis (GSVA) revealed significant differences in mismatch repair, cell cycle, WNT signaling pathway, NOTCH signaling pathway, Complement and Cascades, and cancer pathways at different GSEC

expression levels. Thus, these results suggest that GSEC may be involved in the proliferation and invasion of low-grade glioma, making it a prognostic risk factor for low-grade glioma.

Conclusion: Our analysis identified methylation-related lncRNAs in low-grade gliomas, providing a foundation for further research on lncRNA methylation. We found that GSEC could serve as a candidate methylation marker and a prognostic risk factor for overall survival in low-grade glioma patients. These findings shed light on the underlying mechanisms of low-grade glioma development and may facilitate the development of new treatment strategies.

KEYWORDS

low-grade glioma, prognostic model, LncRNA, M1A/M5C/M6A, immune infiltration

Introduction

Gliomas are the most common primary tumors in the human brain and spinal cord. The World Health Organization (WHO) classified the primary central nervous system (CNS) tumors in 2007 using histopathological diagnostic analysis. Gliomas can be classified by cell type into astrocytomas, oligodendrogliomas, neuronal and mixed neuron gliomas, ependymomas, or oligodendrogliomas. Gliomas can also be graded from the least to most aggressive (Grades I to IV), with grades I and II indicating low-grade gliomas and grades III and IV showing high-grade gliomas (1–3). The median survival time is 11.6 years for low-grade gliomas, about three years for patients with grade III gliomas, and 15 months for patients with grade IV gliomas (4). Therefore, it is essential to study the mechanisms mediating the progression and prognosis of glioma.

RNA post-transcriptional modifications, including N6 methyladenosine (m6A), 5-methyladenine (m5C), N1 methyladenosine (m1A), and 7-methyladenosine m7G methylation (5), have recently gained attention in epigenetic research. The m6A, m1A, and 5-m5C modifications are the most common in eukaryotic messenger RNA (mRNA) regulation. Current studies have proved that m6A, m1A, and m5C regulators play essential roles in methylation, which is related to tumor progression (6–9). M6A regulatory genes methyltransferase 3 (METTL3), METTL14, and WTAP reportedly initiated m6A modification (7). MRTTL3 is usually overexpressed in endometrial epithelial ovarian cancer (EEOC) and can be used as a risk factor for the overall survival of EEOC patients. Similarly, M5C methyltransferase NSUN2 is overexpressed in gastric cancer and can be used as a risk factor for the overall survival of gastric cancer patients. Cell experiments demonstrated that NSUN2 promoted gastric cancer cell proliferation, migration, and invasion (9). Several studies have recently developed genetic risk models to evaluate the prognostic status of cancer patients and demonstrated the independent roles of the predictive variables (10–13).

Researchers have found that although long non-coding RNA (lncRNA) cannot be converted to protein, it impacts many biological processes, such as tumorigenesis and progression (14, 15). Methylation-related lncRNAs are involved in various biological processes associated with cancer progression (15) and have recently been found to influence cell proliferation, migration, and metastasis of many tumors (16–18). Meanwhile, the relationship between methylation and lncRNAs is being extensively studied, but their interaction mechanism is still unclear.

The role of m6A, m1A, and m5C regulatory genes in the progression of low-grade gliomas needs to be better understood. Therefore, this study aimed to evaluate the biological roles of m6A/m1A/m5C regulatory genes in the progression of low-grade gliomas using data from the Cancer Genome Atlas (TCGA) database and identify the lncRNAs associated with their regulatory networks.

Currently, computational biology and high-throughput sequencing data have been widely used in the research of the biomedicine field by Yutao (10, 19, 20). Wang et al. used computational biology methods such as WGCNA to identify biomarkers in different tumors, which provided us with a reliable methodological basis for studying the mechanism of tumorigenesis (21, 22). Weighted gene co-expression network analysis (WGCNA) and a prognostic risk model were used to calculate the prognosis signature score for the low-grade gliomas with methylation-associated lncRNAs.

Method

Data collection

We accessed the TCGA database (https://portal.gdc.Cancer.gov/) to obtain the gene matrix profiles and the relevant clinical information of the low-grade glioma patients, including age, sex, survival time, survival rate, and tissue or organ sample availability. We obtained 514 low-grade glioma tumor samples from patients with primary tumors and metastatic gliomas (23), and 44 m6A/m5C/m1A regulators were determined based on the existing research on methylation (Supplementary Table 1). To ensure the accuracy and

feasibility of analysis, we merged all data and converted them into TPM data format after downloading.

Determination of methylation-related lncRNAs

We determined the lncRNA expression levels of the TCGA-low grade glioma (LGG) cohort and used Pearson's correlation to identify 44 m6A/m5C/m1A methylation regulators associated with lncRNAs.

LncRNA univariate COX regression analysis

We downloaded the clinical follow-up data, including disease status, of the TCGA-LGG cohort from the TCGA database and individually matched the gene expression data to the clinical information. We eliminated the samples with no match (20, 21, 24) and used univariate Cox proportional-hazards regression analysis to determine the lncRNAs highly associated with overall survival. The $P < 0.05$ indicated a significant prognostic statistical significance. These prognostic lncRNAs were used for non-negative matrix factorization and predictive model construction.

Determination of the different lncRNA expression patterns related to methylation regulators

The prognostically significant lncRNAs were first clustered using non-negative matrix dimensionality reduction with 50 iterations. We obtained 9 clusters with the k-mer of 2-10, and the minimum sample size of each group was set to 10 by the 'non-negative matrix factorization (NMF)' R package. The number of our most desirable cluster groups was selected based on the Cophenetic, Dispersion, and Silhouette parameters. After that, survival analysis was used to determine the survival differences between the expression patterns, and $P < 0.05$ was considered significant.

WGCNA analysis

To investigate the biological differences among the different expression patterns of methylation-associated lncRNAs, we constructed protein-coding gene co-expression networks using the WGCNA method. We performed the functional enrichment of the co-expression networks. The TCGA-LGG co-expression network was created using the WGCNA R package, and optimum weighting parameters of the adjacent functions were obtained using the pickSoftThreshold function, which served as a soft threshold for subsequent network construction., Furthermore, the weighted adjacency matrix and the related gene modules were constructed based on the hierarchical clustering of the topological overlap matrix (25). To determine the biological significance of the co-expression modules, we calculated the correlation between the characteristic genes of each module and the NMF cluster analysis groups. Consequently, we identified the most relevant co-expression networks of methylation-associated lncRNAs.

Intersection function analysis

The Database for The Annotation, Visualization, and Integrated Discovery (DAVID, v6.8) was used to annotate the protein-coding genes enriched in co-expression biology, biological processes, and cellular composition (26). Moreover, the Kyoto Encyclopedia of Genes and Genomes (KEGG) (https://www.genome.jp/kegg/) (27) and Gene Ontology (28) (http://geneontology.org/) analyses were applied to identify the biological function of the genes.

Least absolute shrinkage and selection operator regression

The LASSO (29) regression algorithm was used to identify the prognostic survival of low-grade glioma patients and construct a predictive gene model. We used the single factor data of methylation-related lncRNAs to build the model, with the random number seed set as 27. After that, the time-dependent receiver operating characteristic curve (ROC) was used to evaluate the model's predictive performance. The different survival outcomes between the two groups were compared using the Kaplan-Meier survival curve and the log-rank test.

Immune microenvironment analysis

We assessed the proportion of immune cells in the immune microenvironment of TCGA-LGG using several methods. These methods included CIBERSORT (30, 31), EPIC (32), quanTIseq (33), MCPcounter (34), XCELL (35), and TIMER (36). After that, tumor purity of the tumor immune microenvironment was assessed using ESTIMATE, which estimated the proportion of stroma and immune cells in malignant tumor tissues using expression data to generate the purity score. The gene sets associated with multiple confirmatory responses were evaluated to explore the relationship between the model and the confirmatory responses in the immune microenvironment. These gene sets included major histocompatibility complex class II(MHC-II), lymphocyte-specific kinase (LCK), hematopoietic cell kinase (HCK), immunoglobulin G (IgG), signal transducer and activator of transcription 1(STAT1), costimulatory molecule (B7-CD28), interferon, and tumor necrosis factor (TNF) (37). Genes in these gene sets are presented in Supplementary Table 2.

Results

The research routine

Multiple methylation regulatory genes were obtained through a literature review using the analysis process shown in Figure 1. The

Pearson correlation analysis identified the methylation-associated lncRNAs, and we subjected the lncRNAs with independent aftereffects to a prognostic analysis. Thus, two cohorts of low-grade gliomas with different expression patterns of lncRNA were obtained. The survival analysis revealed a significant difference in the overall survival between the two groups of patients with varying expression patterns. After functional enrichment, WGCNA was used to analyze the co-expression networks between the two groups and determine the differences in their biological functions. We also constructed a prognostic survival model for low-grade glioma using the lncRNAs to demonstrate the involvement of lncRNAs in cell proliferation and invasion through cell experiments.

We screened 2330 lncRNAs to identify those with correlation coefficients greater than 0.4 based on the 42 methylation-related protein-coding genes (Figure 2A). Using univariate COX regression, we analyzed the association between these lncRNAs and overall survival. The risk ratios of lncRNAs and the corresponding statistical parameters are shown in Figure 2B. AP005482 was a prognostic protective factor with a risk ratio of 0.710, and AC020910 was a prognostic protective factor for low-grade gliomas.

Survival differences associated with the long non-coding RNA expression patterns

Cluster analysis of methylated lncRNAs was performed after the univariate COX regression analysis at $P < 0.05$. We used 108

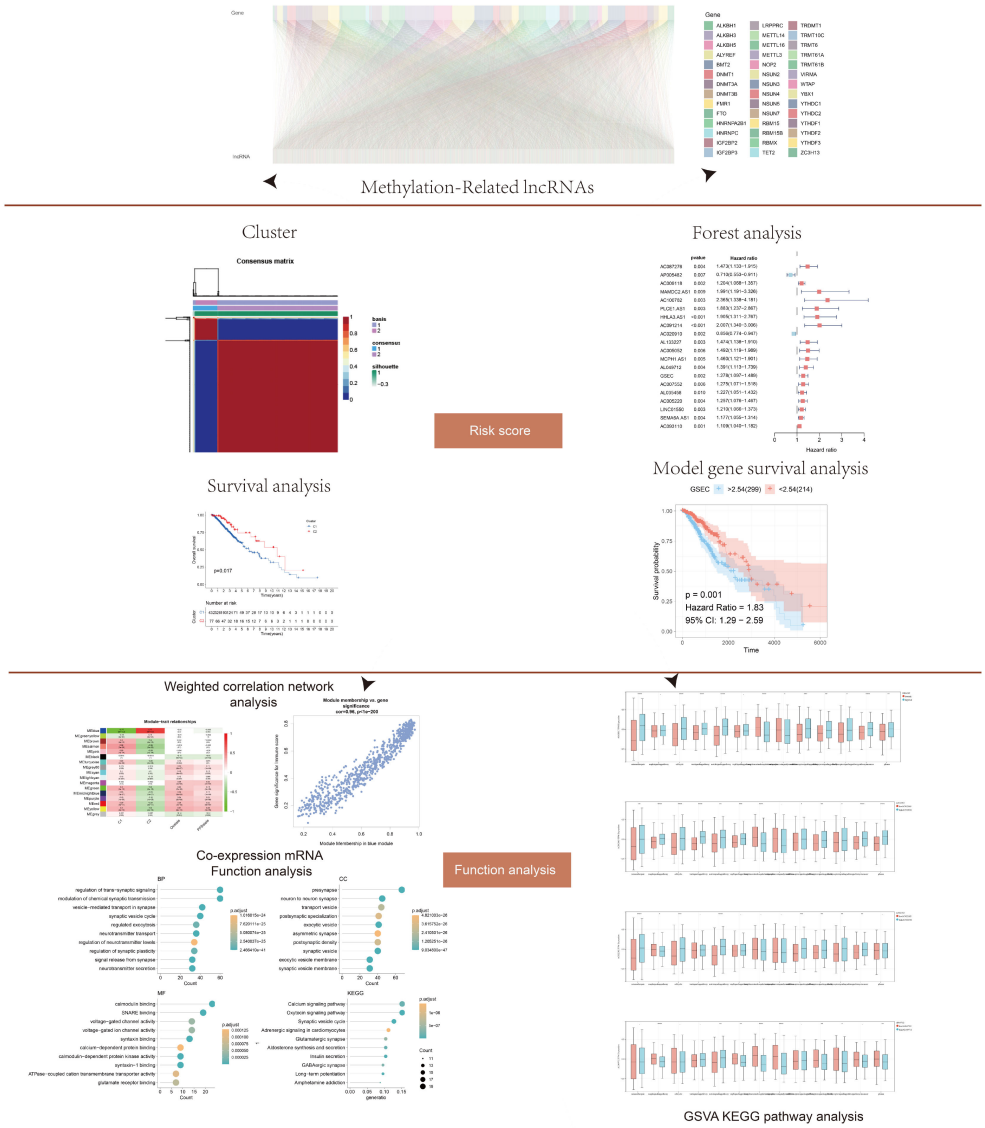


FIGURE 1
Flow chart showing the methodology of the study. Long non-coding RNAs (lncRNAs) associated with DNA methylation were screened from the literature. Cluster analysis, forest map, and survival analysis were used to determine the risk score. Functional analysis was performed *via* weighted gene co-expression network analysis (WGCNA), co-expression enrichment, and gene set variation analysis (GSVA) analyses. The model was verified by cell experiment.

10.3389/fonc.2023.1177120

lncRNAs for the non-negative matrix dimensionality reduction, whose results are shown in Figures 2C, D. There was a strong correlation between the two groups, as indicated by the red coloration; however, blue coloration indicated a weak correlation between the two groups. The clustering between the two groups was excellent, and there were no significant differences between the two groups. We also evaluated the overall and disease-free survival and found that the overall survival and relapse-free rates were lower in group C1 than in group red (Figure 2E).

Identification of biological function differences between two different methylation patterns

We identified two different methylation-related expression patterns of the lncRNAs. WGCNA was used to analyze the protein-coding gene network of low-grade gliomas using the optimal soft threshold of 5 (Figures 3A, B). We obtained 17 co-expression modules which were then used to calculate the correlation of the different methylation

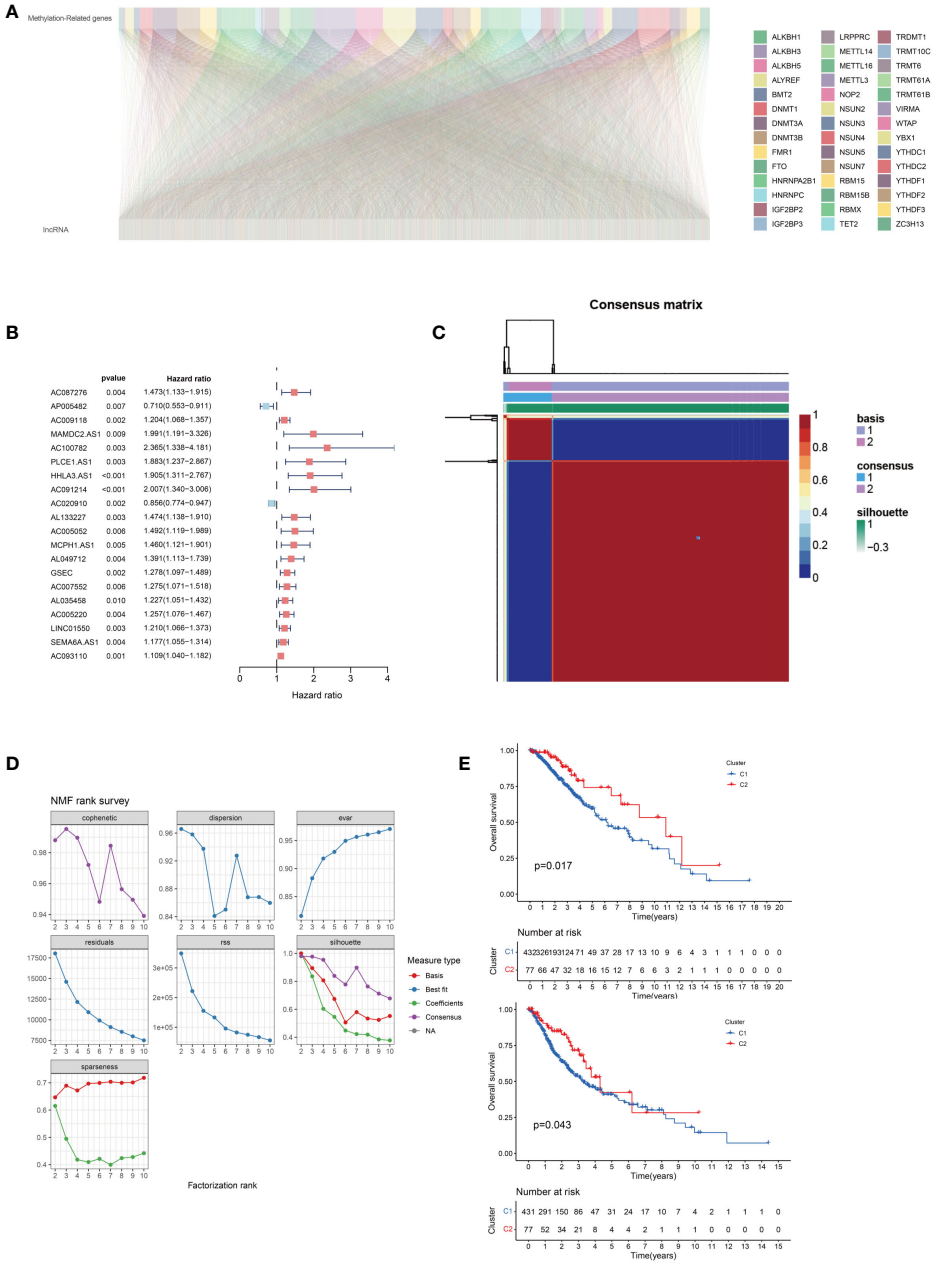


FIGURE 2
(A) Correlation between 5-methyladenine (m5C)-related genes and long non-coding RNAs (lncRNAs) in lower-grade gliomas. **(B)** Univariate Cox regression analysis of prognostic lncRNAs associated with m5C. **(C)** Negative matrix factorization clustering of m5C-relatedlncRNAs gene sets. **(D)** The parameters of negative matrix factorization clustering. **(E)** Overall and disease-free survival prognosis curve of the different subtypes.

10.3389/fonc.2023.1177120

expression patterns (Figure 3C). The correlation between the blue module and the C2 group was 0.77, while the correlation within the module was 0.96 (Figure 3D). We functionally enriched the blue module to determine the biological function differences between the different lncRNAs. We found that the co-expressed genes in the blue module were mostly related to trans-synaptic signaling in biological processes, modulation of chemical synaptic transmission, calmodulin binding, SNARE binding, calcium signaling pathway, and oxytocin signaling pathway (Figure 4).

Screening and construction of long non-coding RNA prognostic models using machine learning methods

Using the random forest method, we first screened prognostic lncRNAs and identified 102 typical lncRNAs based on their importance ranking. LASSO regression analysis was then performed on these lncRNAs to construct a methylation-related prognostic model using four prognostically significant genes. The risk score of the model was 1.12 *AC012063 + 0.74 * AC022382 + 0.32 * AL049712 + 0.16 * GSEC. Furthermore, we analyzed the independent prognostic value of four lncRNAs in the prognostic model and found that the four lncRNAs encoded AC012063, AC022382, AL04971, and GSEC. All four lncRNAs were independent prognostic factors for low-grade gliomas (Figure 5).

Result evaluation of the model

We summarized the risk factor (gene) expression of each sample and the clinical follow-up information for generating a heatmap (Figures 6A, B). The samples were presented in ascending order in the heatmap based on their risk scores. Since patients with higher risk scores had a poor prognosis, we marked the actual survival status of the patients with red and blue plot points and determined the corresponding points on the ordinate survival time. The number of patients with red plot points increased, but their survival time reduced as the risk score increased. These plots were concentrated in the lower right corner of the heat map, demonstrating that patients with low-grade gliomas exhibit poor prognoses with increasing risk scores. This also indicated the possible prognostic roles of lncRNAs such as AC012063, AC022382, AL04971, and GSEC. Significant impact, suggesting important research value. The expression levels of the prognostic risk factors of each patient were annotated on the x-axis of the heatmap. The results showed that the expression levels of AC012063, AC022382, AL04971, and GSEC gradually increased with the risk score progression, but survival time reduced (Figures 6A, B). The survival curve and the ROC analysis results of the different risk groups were shown in Figures 6C-F, which indicated the patients with high risk score might lead the worse clinical outcome.

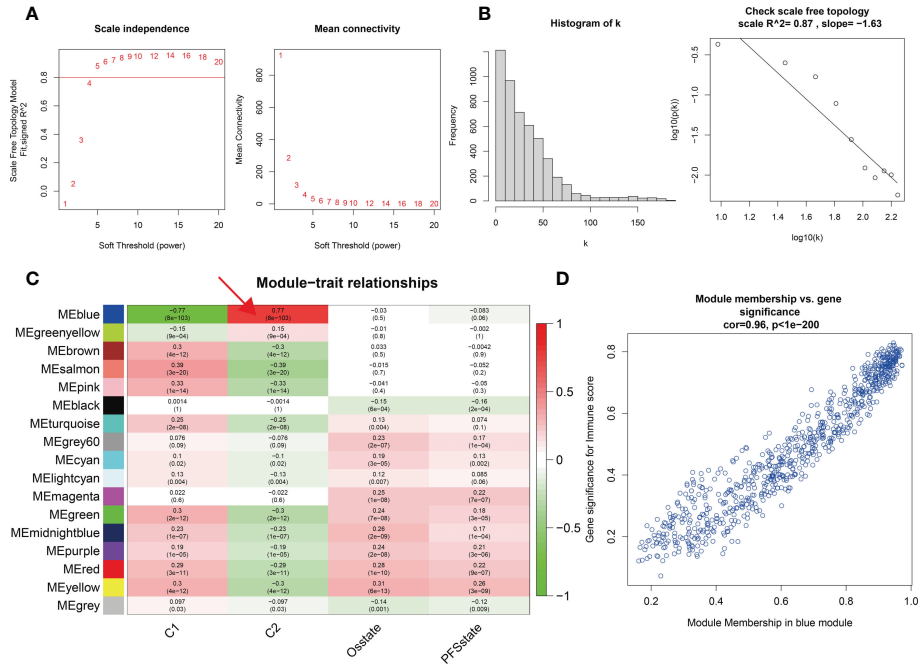


FIGURE 3
(A, B) Soft threshold and scale-free topology of weighted correlation network analysis. **(C)** Module-trait relationships of different modules and different molecular typing. The relationship between the blue module and C2 is mostly connective. **(D)** The Pearson correlation coefficient between the significance of the immune score and module membership in the blue module is 0.96.

LNCRNAs prognostic model and immune microenvironment and response

The enrichment analysis showed that antigen binding, B cell-mediated immunity, complement activation, and immunoglobulin receptor binding were highly enriched in the high-risk score group. However, the low-risk group had significantly enriched exocytic vesicle membrane, neurotransmitter transport, and positive synaptic transmission regulation (Figures 7A, B). Based on these results, we further investigated how the risk scores related to the immune microenvironment and immune validation response. We assessed tumor purity and the immune and stromal scores in low-grade gliomas using the ESTIMATE method, and the analysis included 8 immune-validated response gene sets. The gene sets included virulent T lymphocyte-related biomarkers representing the strength of the cellular immune response. The immune-validation response gene sets, such as the tumor necrosis family, were also included. The results showed that IgG, HCK, MHC-II, LCK, STAT1 interferon, B7-CD28, and TNF-related tumor immune responses were significantly enhanced with the increasing risk scores, indicating that the immune microenvironment in high-risk glioma patients regulates response changes (Figure 7C).

Effects of risk score-independent prognostic variables on biological pathways

We divided patients into two groups based on the median expression of the risk score variables. We then assessed the gene set variation analysis (GSVA) score of the C2 KEGG pathway using the GSVA method and conducted t-tests with completely randomized data. The results showed that mismatch repair, cell cycle, wnt signaling pathway, NOTCH signaling pathway, complement, coagulation cascades, cancer pathways, and other pathways significantly differed in GSEC expression levels of the two groups. Groups with high GSEC expression were associated with poorer prognoses, and cell proliferation-related pathways, such as mismatch repair and cell cycle, were highly expressed in the group with high GSEC expression. Two classic biological pathways, WNT and NOTCH signaling, were also significantly upregulated. This suggests that GSEC may affect the prognosis of low-grade glioma cells by enhancing their proliferation, thus providing a reference for future research (Figure 8). In addition, we found higher levels of PDCD1 expression in groups with high risk scores (Figure 9).

Discussion

Researchers are committed to developing prognostic assessment risk scores to evaluate cancer patients' prognoses. However, the role of lncRNAs associated with methylation-related regulators in the prognosis of patients with low-grade gliomas and the immune microenvironment of malignant tumors is unclear. Given the heterogeneity of m6A/M1A/M5C methylation modifications in low-grade gliomas, it is essential to quantify the long-chain non-coding modification group in low-grade glioma patients. In low-grade gliomas, we identified many methylation-related regulators for screening and identifying methylation-associated lncRNAs. We identified methylation-related regulators that could be modified by lncRNAs from the literature and screened 2330 lncRNAs. Furthermore, four lncRNAs in LGG were used to establish a

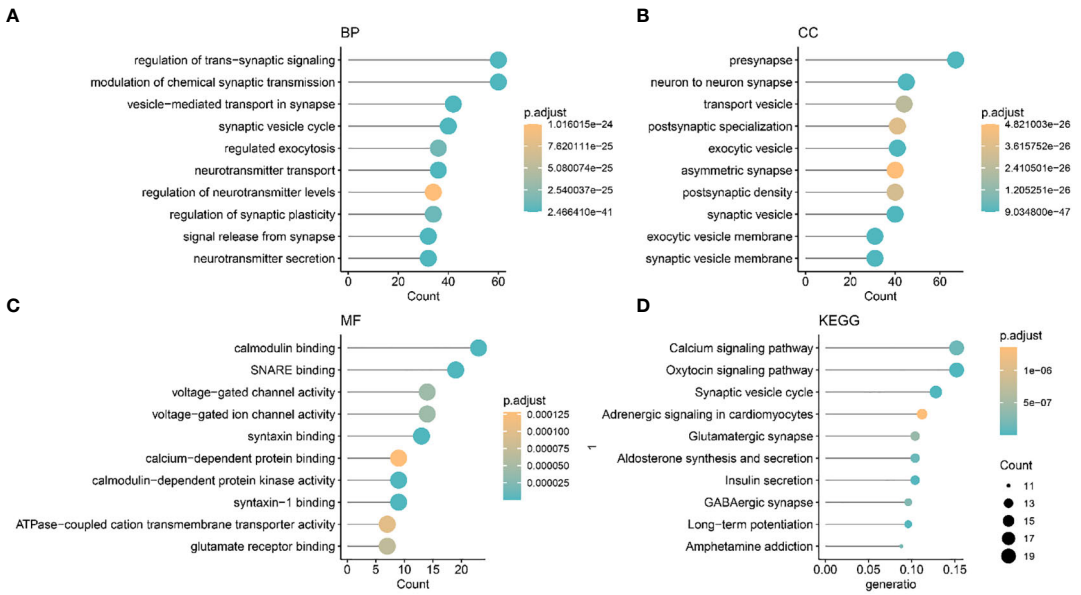


FIGURE 4
The Gene Ontology and Kyoto Encyclopedia of Genes and Genomes biological pathways showing the mainly enriched co-expression modules. **(A)** BP: biological process. **(B)** CC: Cellular component. **(C)** MF: Molecular function **(D)** KEGG.

lncRNAs-based prognostic model to determine the overall survival and prognosis. Patients were divided into low-risk and high-risk groups according to their risk scores. The data showed that AC012063, AC022382, AL04971, and GSEC were prognostic lncRNAs associated with methylation regulators in LGG. Moreover, the AUC of the ROC curve showed that the methylation-associated lncRNAs prognostic model was more accurate than the ones reported in other studies due to its 5-year specific survival and specificity.

Methylation, a common epigenetic modification, plays a crucial role in gene expression regulation. Recent studies have revealed a complex interplay between DNA methylation and long noncoding RNA (lncRNA) in various biological processes. Specifically, lncRNAs have been shown to recruit DNA methyltransferases to specific genomic regions, leading to site-specific DNA methylation. Moreover, some lncRNAs have been found to function as "decoys" that prevent DNA methylation by sequestering DNA methyltransferases away from their target genes. In addition, lncRNAs themselves can also be subject to methylation, which affects their stability and expression levels. Thus, the relationship between methylation and lncRNA is intricate and multifaceted, and further research is needed to fully elucidate its mechanisms and biological implications.

Long non-coding RNA (lncRNA) has a variety of biological functions in glioma, including promoting or inhibiting tumor growth, metastasis, angiogenesis and drug resistance. Among them, lncRNA H19 has been extensively studied. H19, a hepatocyte growth factor (HGF) -induced lncRNA, has been shown to be highly expressed in a variety of tumors, including gliomas. H19 can promote the proliferation and self-renewal of glioma stem cells, and promote the migration and metastasis of tumor cells through different mechanisms, including the regulatory relationship with miRNA, EZH2-mediated epigenetic regulation, etc. Therefore, H19 may serve as a therapeutic target for glioma stem cells and tumor migration. Therefore, the use of computational biology in this study to analyze methylation-related long non-coding RNA is very important for the study of the genesis and development mechanism of glioma

In addition, lncRNA MALAT1 was also up-regulated in glioma. MALAT1 can regulate the proliferation, migration and invasion of tumor cells and participate in the malignant transformation of tumors. In addition, lncrnas such as CCAT1 and TUG1 also play a role in promoting tumor growth and metastasis in glioma. In general, lncrnas play an important role in the occurrence and development of glioma. Understanding their mechanism of action will help to discover new therapeutic targets and develop more effective therapeutic strategies. Overall survival in LGG could also be predicted by the methylation-associated lncRNAs prognostic model, suggesting its potential application in future clinical cohort studies on glioma. The role of DNA and epigenetic histone modifications in cancer progression has led to the development of various drugs, including histone deacetylase inhibitors and

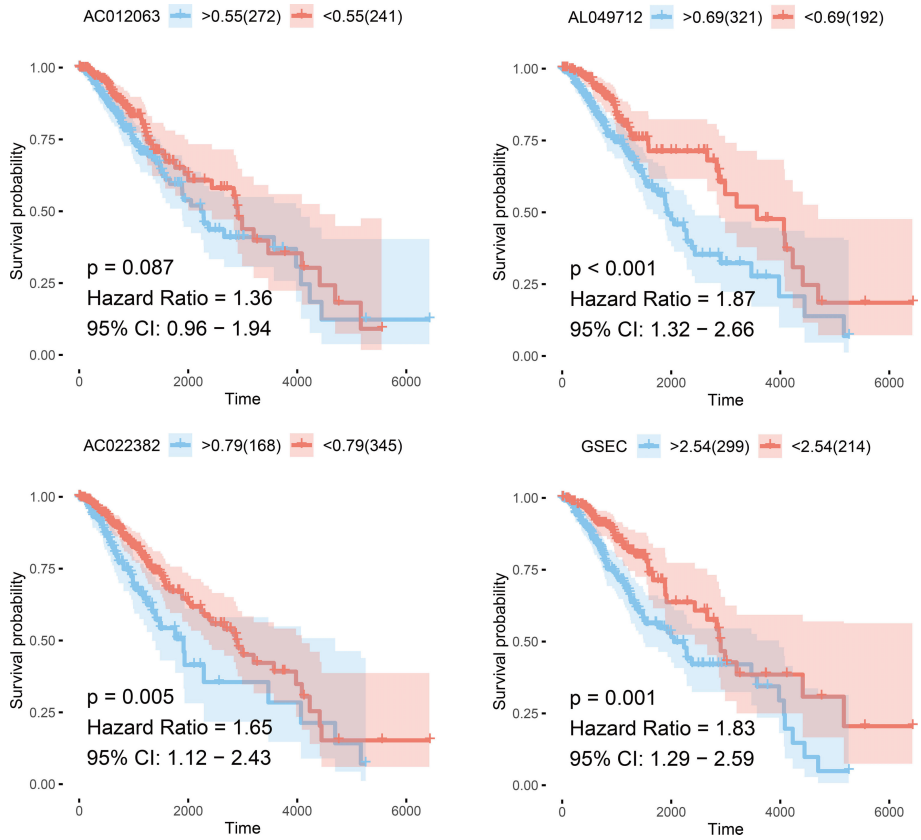


FIGURE 5
Survival curves of the four selected genes.

10.3389/fonc.2023.1177120

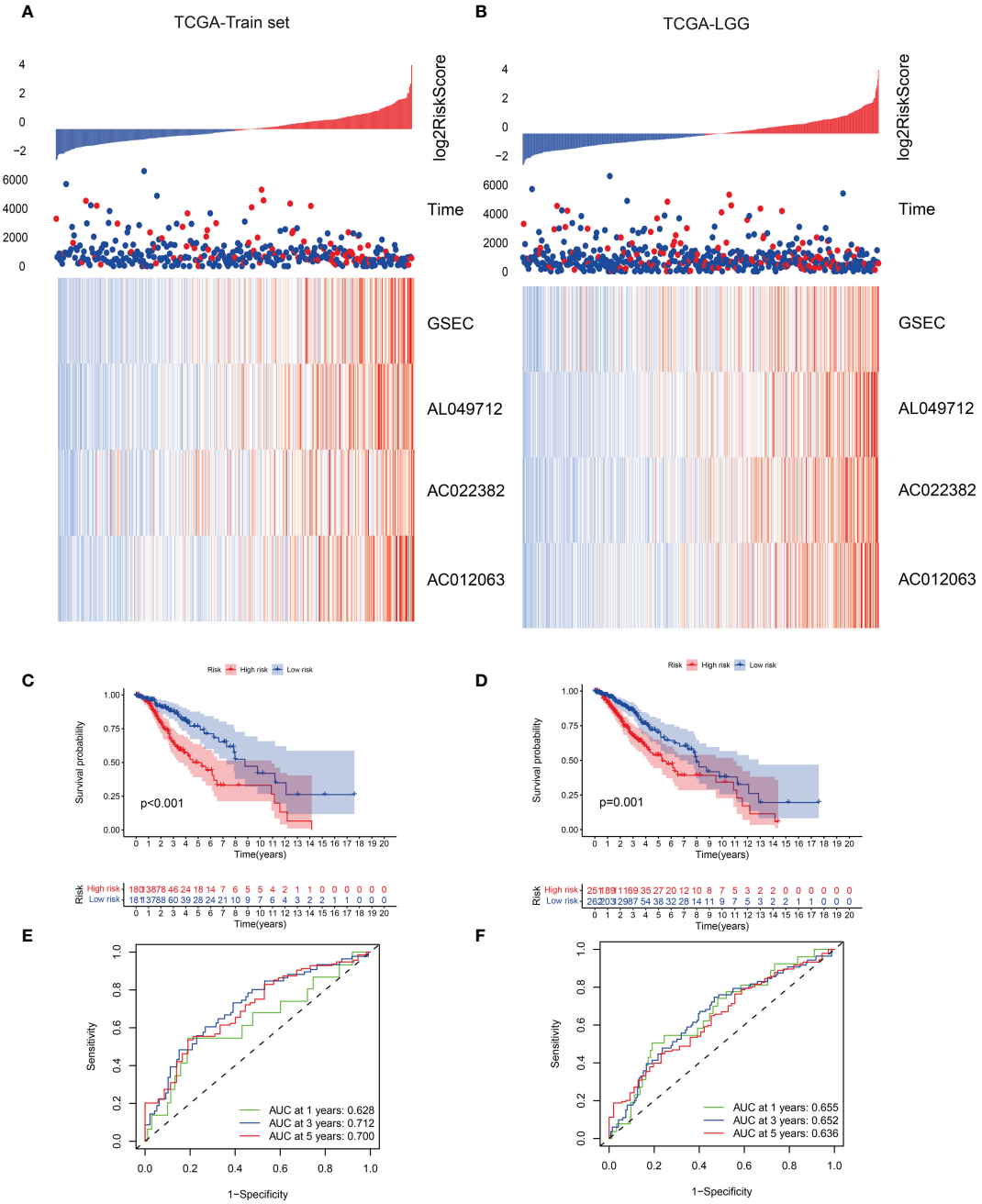


FIGURE 6
(A, B) Risk scores and survival status of gene signatures in the training and validation data set. **(C, D)** Survival curves of the two risk groups with different subtypes. **(E, F)** The receiver operating characteristic (ROC) curve of the two risk groups based on the gene signature classification.

hypoxia-targeting drugs. However, studying the different methylation mechanisms in cells has recently gained attention.

m6A is one of the crucial post-transcriptional modifications of the protein-coding genes in cancer pathogenesis. However, the biological function of lncRNA methylation remains unclear. Several studies showed that m6A might be crucial in cancer pathogenesis, but the mechanisms by which lncRNAs influence cancer progression and metastasis are unclear. M6A modulators extensively modify lncRNAs to control gene expression and cell biology at the transcriptional and post-transcriptional levels. Zhang Jun et al. predicted the interaction between lncRNAs and alkylation repair homolog protein 5 (ALKBH5), a demethylase that reverses methylation. Furthermore, nuclear paraspeckle assembly transcript 1 (NEAT1) was evaluated by gene silencing, RT-PCR, nuclear and cytoplasmic separation, scraping test, and transwell migration test (38). Yewen Shi evaluated the biological function of hepatic nuclear factor 1α antisense RNA 1 (HNF1A–AS1) and its regulatory mechanism in laryngeal squamous cell carcinoma. The study found that HNF1A-AS1 may act as a tumor suppressor lncRNA in LSCC by regulating the epithelial-mesenchymal transition (EMT) process. As a result, new therapeutic targets and strategies were discovered for treating patients with nasopharyngeal

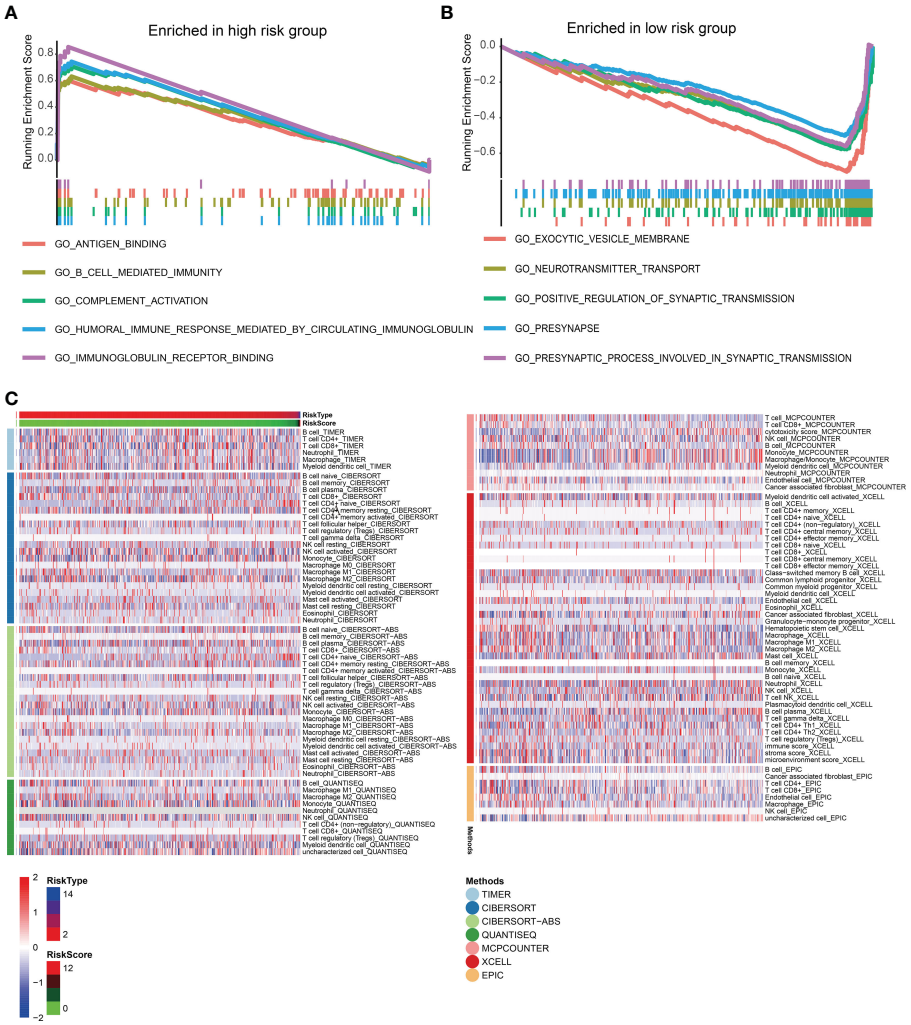


FIGURE 7
(A, B) The gene set enrichment analysis of high-risk and low-risk groups. **(C)** Heat maps of immune responses in the high-risk and low-risk groups based on TIMER, CIBERSORT, Cibersort-ABS, QUANTISEQ, MCPCOUNTER, XCELL, and EPIC algorithms.

carcinoma (NPC) (39). Overexpressing APCDD1L-AS1, a novel lncRNA, inhibited the growth and metastasis of ccRCC cells *in vitro* and *in vivo*. Dysregulation of histone expression caused by APCDD1L-AS1 overexpression may also inhibit ccRCC progression (40). However, APCDD1L-AS1 expression was decreased by DNA hypermethylation and inactivation of von Hippel Lindau (VHL) protein expression. METTL3-mediated modification upregulated LINC00958 by stabilizing its RNA transcript, and the LINC00958 activated miR-3619-5p to upregulate hepatoma-derived growth factor (HDGF) expression. This facilitated tumor lipogenesis and progression (41), indicating the importance of studying the methylation of lncRNAs.

We identified several lncRNAs that may be involved in methylation modification by analyzing the methylation-related regulator RNAs. G-quadruplex forming sequence containing (GSEC)-lncRNA is mostly associated with methylation in low-grade glioma but has also been extensively studied in various cancers. Jianhua Zhang et al. found that GSEC was significantly upregulated in TNBC tissues and cancer cell lines. Moreover, high GSEC levels were associated with tumor staging, positive lymph node metastasis, and poor prognosis in TNBC patients. The study also found that downregulating Mir-202-5p attenuated the inhibitory effect of GSEC knockdown on TNBC cell proliferation, invasion, and migration *in vitro*. Meanwhile, AXL overexpression reversed the *in vitro* mimicry inhibitory effect of Mir-202-5p on TNBC progression (42).

Shangshang Hu et al. constructed a GSEC/Mir-101-3p/SNX16/PAPOLG network to predict the prognosis of hepatocellular carcinoma (43). Xiulin Jiang et al. also found that ferroptosis-related GSEC-lncRNAs, mirNA-101-3p, and CISD1 axis play a functional role in lung adenocarcinoma (LUAD) and may serve as useful diagnostic and therapeutic biomarkers for the disease. The study reported that the ferroptosis-related GSEC- lncRNA/mirNA-101-3P/CISD1 axis could be an independent prognostic marker for lung adenocarcinoma (44).

10.3389/fonc.2023.1177120

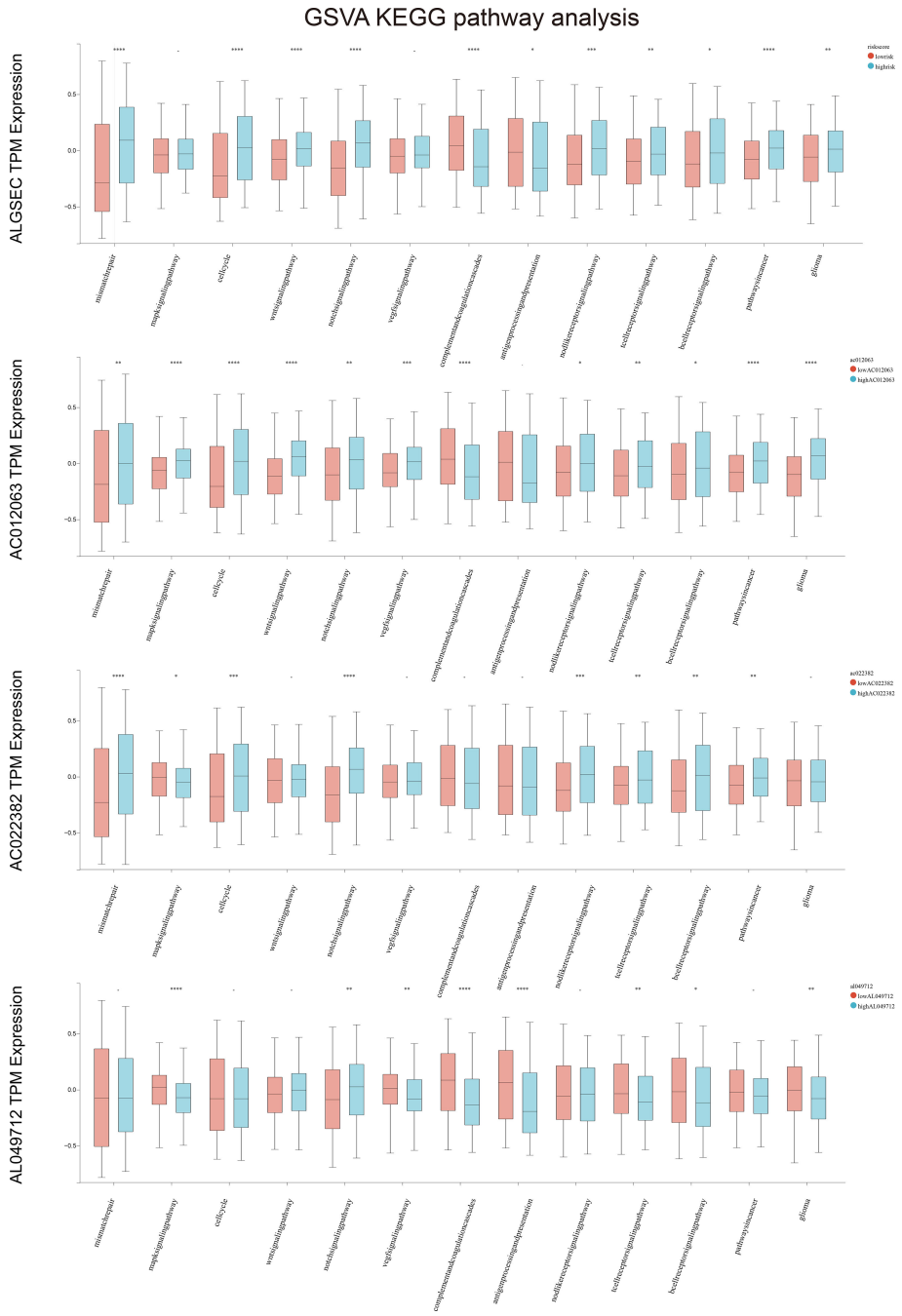


FIGURE 8
Gene set variation analysis (GSVA) in high and low-risk groups among four key genes. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

The present study constructed a methylation-associated lncRNAs prognostic model using computational biology and public databases. The model proved accurate and reliable in training and validating data sets. This suggested the importance of the four long non-coding RNAs, and GSEC has been identified as a potential methylation-related lncRNA. Despite these findings, there were several limitations to this study. This study utilized information from public databases for the *in silico* analysis. Although we have proved the significance of GSEC in cancer progression through literature review, there is a need to verify these findings through more external cohorts and *in vivo* experiments.

Conclusion

This study identified methylation-related lncRNAs in glioma and determined their expression patterns. We found two expression

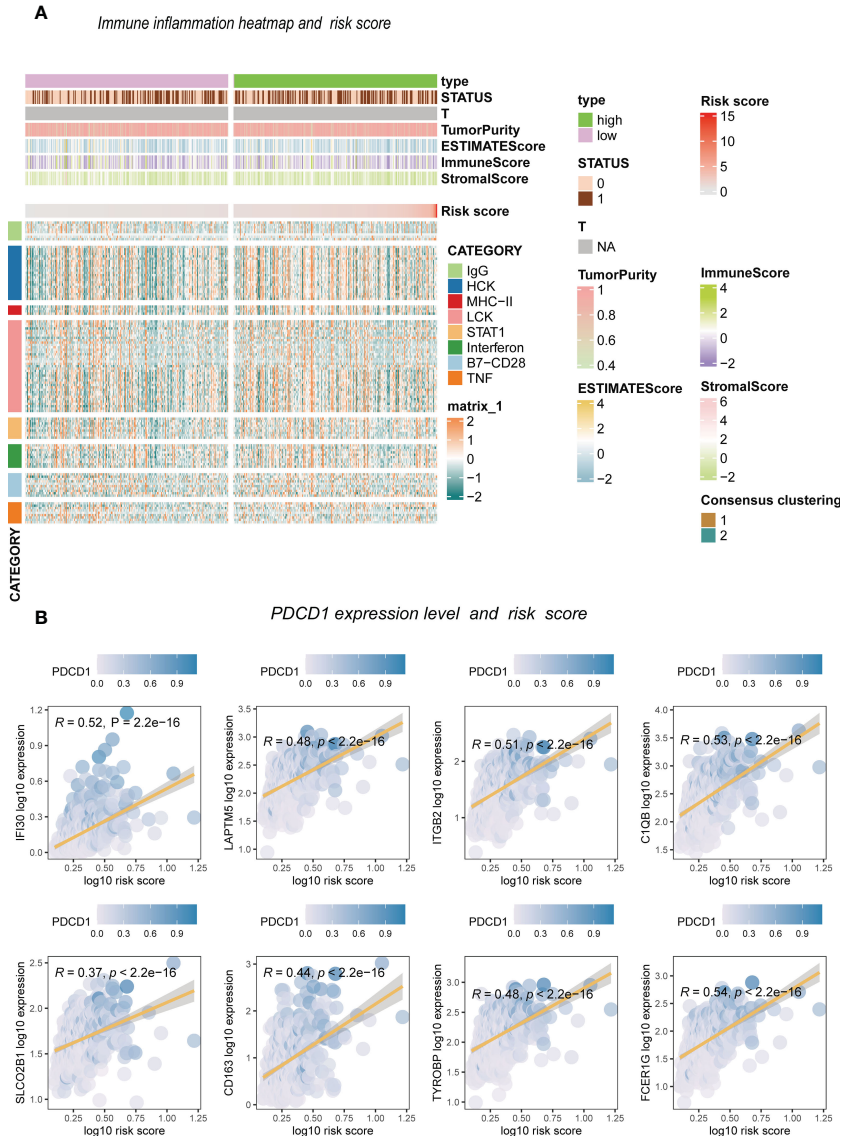


FIGURE 9
The inflammation response with risk score. **(A)** Heatmap. **(B)** Risk score and inflammation genes.

patterns of the methylation-related lncRNAs, and there was a significant difference between the two expression patterns. A prognostic model was also constructed based on these lncRNAs. GSEC was considered a lncRNA with a significant value in cancer progression, thus providing a basis for studying epigenetic methylation. Therefore, this study provides new strategies and research directions in the prognosis and treatment of glioma.

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 human participants were reviewed and approved by Affiliated Haikou Hospital of Xiangya School of South University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YL: data analysis, methodology, figures construction, and article writing. YL, XL, and ZY: methodology and validation. YL, XL, and ZY: supervision. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the Hainan Provincial Natural Science Foundation of China (No. 820MS163) and the Hainan Cerebrovascular Disease Clinical Research Center of China (No. LCYX202206).

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.2023.1177120/full#supplementary-material

References

1. 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(2):97–109. doi: 10.1007/s00401-007-0243-4

2. van den Bent MJ. Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective. *Acta Neuropathol* (2010) 120(3):297–304. doi: 10.1007/s00401-010-0725-7

3. Chen R, Smith-Cohn M, Cohen AL, Colman H. Glioma subclassifications and their clinical significance. *Neurotherapeutics* (2017) 14(2):284–97. doi: 10.1007/s13311-017-0519-x

4. 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

5. Zhou W, Wang X, Chang J, Cheng C, Miao C. The molecular structure and biological functions of RNA methylation, with special emphasis on the roles of RNA methylation in autoimmune diseases. *Crit Rev Clin Lab Sci* (2022) 59(3):203–18. doi: 10.1080/10408363.2021.2002256

6. Zhao Y, Zhao Q, Kaboli PJ, Shen J, Li M, Wu X, et al. m1A regulated genes modulate PI3K/AKT/mTOR and ErbB pathways in gastrointestinal cancer. *Transl Oncol* (2019) 12(10):1323–33. doi: 10.1016/j.tranon.2019.06.007

7. Ma Z, Li Q, Liu P, Dong W, Zuo Y. METTL3 regulates m6A in endometrioid epithelial ovarian cancer independently of METTl14 and WTAP. *Cell Biol Int* (2020) 44 (12):2524–31. doi: 10.1002/cbin.11459

8. Yang Z, Wang T, Wu D, Min Z, Tan J, Yu B. RNA N6-methyladenosine reader IGF2BP3 regulates cell cycle and angiogenesis in colon cancer. *J Exp Clin Cancer Res* (2020) 39(1):203. doi: 10.1186/s13046-020-01714-8

9. Hu Y, Chen C, Tong X, Chen S, Hu X, Pan B, et al. NSUN2 modified by SUMO-2/3 promotes gastric cancer progression and regulates mRNA m5C methylation. *Cell Death Dis* (2021) 12(9):842. doi: 10.1038/s41419-021-04127-3

10. Wang Y, Yan K, Wang L, Bi J. Genome instability-related long non-coding RNA in clear renal cell carcinoma determined using computational biology. *BMC Cancer* (2021) 21(1):727. doi: 10.1186/s12885-021-08356-9

11. Yan K, Lu Y, Yan Z, Wang Y. 9-gene signature correlated with CD8(+) T cell infiltration activated by IFN-γ: a biomarker of immune checkpoint therapy response in melanoma. *Front Immunol* (2021) 12:622563. doi: 10.3389/fimmu.2021.622563

12. Guo Q, Xiao XY, Wu CY, Li D, Chen J-L, Ding X-C, et al. Clinical roles of risk model based on differentially expressed genes in mesenchymal stem cells in prognosis and immunity of non-small cell lung cancer. *Front Genet* (2022) 13:823075. doi: 10.3389/fgene.2022.823075

13. Zhang N, Zhang H, Wu W, Zhou R, Li S, Wang Z, et al. Machine learning-based identification of tumor-infiltrating immune cell-associated lncRNAs for improving outcomes and immunotherapy responses in patients with low-grade glioma. *Theranostics* (2022) 12(13):5931–48. doi: 10.7150/thno.74281

14. Peng WX, Koirala P, Mo YY. LncRNA-mediated regulation of cell signaling in cancer. *Oncogene* (2017) 36(41):5661–7. doi: 10.1038/onc.2017.184

15. Yi YC, Chen XY, Zhang J, Zhu JS. Novel insights into the interplay between m(6) A modification and noncoding RNAs in cancer. *Mol Cancer* (2020) 19(1):121. doi: 10.1186/s12943-020-01233-2

16. Liu Z, Li S, Huang S, Wang T, Liu Z. N(6)-methyladenosine regulators and related LncRNAs are potential to be prognostic markers for uveal melanoma and indicators of tumor microenvironment remodeling. *Front Oncol* (2021) 11:704543. doi: 10.3389/fonc.2021.704543

17. Wang Y, Li N, Tian D, Zhou C-W, Wang Y-H, Yang C, et al. Analysis of m6A-related lncRNAs for prognosis value and response to immune checkpoint inhibitors therapy in hepatocellular carcinoma. *Cancer Manag Res* (2021) 13:6451–71. doi: 10.2147/CMAR.S322179

18. Zhang P, Liu G, Lu L. N6-Methylandenosine-Related lncRNA signature is a novel biomarkers of prognosis and immune response in colon adenocarcinoma patients. *Front Cell Dev Biol* (2021) 9:703629. doi: 10.3389/fcell.2021.703629

19. Wang Y, Yan K, Lin J, Li J, Bi J. Macrophage M2 Co-expression factors correlate with the immune microenvironment and predict outcome of renal clear cell carcinoma. *Front Genet* (2021) 12:615655. doi: 10.3389/fgene.2021.615655

20. Wang Y, Yan K, Wang J, Lin J, Bi J. M2 macrophage Co-expression factors correlate with immune phenotype and predict prognosis of bladder cancer. *Front Oncol* (2021) 11:609334. doi: 10.3389/fonc.2021.609334

21. Wang Y, Wang J, Yan K, Lin J, Zheng Z, Bi J. Identification of core genes associated with prostate cancer progression and outcome *via* bioinformatics analysis in multiple databases. *PeerJ* (2020) 8:e8786. doi: 10.7717/peerj.8786

22. Wang Y, Yan K, Lin J, Liu Y, Wang J, Li X, et al. CD8+ T cell Co-expressed genes correlate with clinical phenotype and microenvironments of urothelial cancer. *Front Oncol* (2020) 10:553399. doi: 10.3389/fonc.2020.553399

23. Blum A, Wang P, Zenklusen JC. SnapShot: TCGA-analyzed tumors. *Cell* (2018) 173(2):530. doi: 10.1016/j.cell.2018.03.059

24. Huang Z, Yang Q, Huang Z. Identification of critical genes and five prognostic biomarkers associated with colorectal cancer. *Med Sci Monit* (2018) 24:4625–33. doi: 10.12659/MSM.907224

25. Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabási AL. Hierarchical organization of modularity in metabolic networks. *Science* (2002) 297(5586):1551–5. doi: 10.1126/science.1073374

26. Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, et al. The DAVID gene functional classification tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol* (2007) 8(9):R183. doi: 10.1186/gb-2007-8-9-r183

27. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* (2017) 45 (D1):D353–61. doi: 10.1093/nar/gkw1092

28. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. the gene ontology consortium. *Nat Genet* (2000) 25(1):25–9. doi: 10.1038/75556

29. Tibshirani R. The lasso method for variable selection in the cox model. *Stat Med* (1997) 16(4):385–95. doi: 10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3

30. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* (2019) 37(7):773–82. doi: 10.1038/s41587-019-0114-2

31. Steen CB, Liu CL, Alizadeh AA, Newman AM. Profiling cell type abundance and expression in bulk tissues with CIBERSORTx. *Methods Mol Biol* (2020) 2117:135–57. doi: 10.1007/978-1-0716-0301-7_7

32. Racle J, de Jonge K, Baumgaertner P, Speiser DE, Gfeller D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *Elife* (2017) 6. doi: 10.7554/eLife.26476

33. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Med* (2019) 11(1):34.

34. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* (2016) 17(1):218.

35. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol* (2017) 18(1):220. doi: 10.1186/s13059-017-1349-1

36. Li B, Severson E, Pignon JC, Zhao H, Li T, Novak J, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biol* (2016) 17(1):174. doi: 10.1186/s13059-016-1028-7

37. Rody A, Holtrich U, Pusztai L, Liedtke C, Gaetje R, Ruckhaeberle E, et al. T-Cell metagene predicts a favorable prognosis in estrogen receptor-negative and HER2-positive breast cancers. *Breast Cancer Res* (2009) 11(2):R15. doi: 10.1186/bcr2234

38. Zhang J, Guo S, Piao HY, Wang Y, Wu Y, Meng X-Y, et al. ALKBH5 promotes invasion and metastasis of gastric cancer by decreasing methylation of the lncRNA NEAT1. *J Physiol Biochem* (2019) 75(3):379–89. doi: 10.1007/s13105-019-00690-8

39. Shi Y, Zhang Q, Xie M, Feng Y, Ma S, Yi C, et al. Aberrant methylation–mediated decrease of lncRNA HNF1A–AS1 contributes to malignant progression of laryngeal squamous cell carcinoma *via* EMT. *Oncol Rep* (2020) 44(6):2503–16. doi: 10.3892/or.2020.7823

40. Yang W, Zhou J, Zhang Z, Zhang K, Xu Y, Li L, et al. Downregulation of lncRNA APCDD1L-AS1 due to DNA hypermethylation and loss of VHL protein expression promotes the progression of clear cell renal cell carcinoma. *Int J Biol Sci* (2022) 18(6):2583–96. doi: 10.7150/ijbs.71519

41. Zuo X, Chen Z, Gao W, Zhang Y, Wang J, Wang J, et al. M6A-mediated upregulation of LINC00958 increases lipogenesis and acts as a nanotherapeutic target in hepatocellular carcinoma. *J Hematol Oncol* (2020) 13(1):5. doi: 10.1186/s13045-019-0839-x

42. Zhang J, Du C, Zhang L, Wang Y, Zhang Y, Li J. lncRNA GSEC promotes the progression of triple negative breast cancer (TNBC) by targeting the miR-202-5p/AXL axis. *Onco Targets Ther* (2021) 14:2747–59. doi: 10.2147/OTT.S293832

43. Hu S, Zhang J, Guo G, Zhang L, Dai J, Gao Y. Comprehensive analysis of GSEC/miR-101-3p/SNX16/PAPOLG axis in hepatocellular carcinoma. *PloS One* (2022) 17(4):e0267117. doi: 10.1371/journal.pone.0267117

44. Poliak MG. The angiotensin-forming system of the brain and the regulation of water-electrolyte metabolism and hemodynamics. *Probl Endokrinol (Mosk)* (1986) 32(4):84–90.

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 16 May 2023
DOI 10.3389/fonc.2023.1180722

OPEN ACCESS

EDITED BY
Nan Zhang,
Harbin Medical University, China

REVIEWED BY
Wantao Wu,
Central South University, China
Siqing Sun,
The First Affiliated Hospital of China Medical University, China

*CORRESPONDENCE
Yan Wu
cmuwuyan@163.com

RECEIVED 06 March 2023
ACCEPTED 20 April 2023
PUBLISHED 16 May 2023

CITATION
Bai Y, He J, Ma Y, Liang H, Li M and Wu Y (2023) Identification of DNA repair gene signature and potential molecular subtypes in hepatocellular carcinoma.
Front. Oncol. 13:1180722.
doi: 10.3389/fonc.2023.1180722

Identification of DNA repair gene signature and potential molecular subtypes in hepatocellular carcinoma

Yi Bai[1], Jinyun He[2], Yanquan Ma[1], He Liang[3], Ming Li[4] and Yan Wu[5]*

[1]Department of Critical Care Medicine, Panjin Liaoyou Baoshihua Hospital, Liaoning, China, [2]Department of hepatobiliary surgery, Panjin Liaoyou Baoshihua Hospital, Liaoning, China, [3]Department of integrated Chinese and Western medicine, Panjin Liaoyou Baoshihua Hospital, Liaoning, China, [4]Fuxin Municipal Discipline Inspection Commission, Liaoning, China, [5]Department of rheumatology and immunology, Panjin Liaoyou Baoshihua Hospital, Liaoning, China

DNA repair is a critical factor in tumor progression as it impacts tumor mutational burden, genome stability, PD-L1 expression, immunotherapy response, and tumor-infiltrating lymphocytes (TILs). In this study, we present a prognostic model for hepatocellular carcinoma (HCC) that utilizes genes related to the DNA damage response (DDR). Patients were stratified based on their risk score, and groups with lower risk scores demonstrated better survival rates compared to those with higher risk scores. The prognostic model's accuracy in predicting 1-, 3-, and 5-year survival rates for HCC patients was analyzed using receiver operator curve analysis (ROC). Results showed good accuracy in predicting survival rates. Additionally, we evaluated the prognostic model's potential as an independent factor for HCC prognosis, along with tumor stage. Furthermore, nomogram was employed to determine the overall survival year of patients with HCC based on this independent factor. Gene set enrichment analysis (GSEA) revealed that in the high-risk group, apoptosis, cell cycle, MAPK, mTOR, and WNT cascades were highly enriched. We used training and validation datasets to identify potential molecular subtypes of HCC based on the expression of DDR genes. The two subtypes differed in terms of checkpoint receptors for immunity and immune cell filtration capacity.Collectively, our study identified potential biomarkers of HCC prognosis, providing novel insights into the molecular mechanisms underlying HCC.

KEYWORDS

hepatocellular carcinoma, DNA repair gene, prognosis, molecular subtypes of HCC, immune infiltrition

1 Introduction

According to the 2018 Global Cancer Statistics report, Hepatocellular Carcinoma (HCC) ranks 6th among malignancies and is the 4th leading cause of cancer-related mortality (1, 2). Despitesignificant advancements in HCC treatment, the outcomes are still unsatisfactory (3, 4). Therefore,identifying novel therapeutic targets and diagnostic biomarkers for HCC is crucial to improve patientprognosis. The DNA damage response (DDR) pathway is considered a potential source of therapeutictargets as damaged DNA is a hallmark of cancerous cells (5).

Research studies have reported that genes involved in DDR pathways, such as nucleotide and base excision and mismatch repair, are aberrantly expressed during cancer development and progression (6–10). Dysregulated DDR is associated with increased genome instability in HCC cells and has asignificant impact on patient prognosis (11).

The use of high-throughput sequencing technology has become increasingly prevalent in recentyears, and sequencing data and clinical follow-up information can be downloaded from many cancerdatabases. In this study, we downloaded the hepatocellular carcinoma dataset from TCGA and GEOdatabases to explore the prognostic potential of DNA damage response (DDR)-linked genes inhepatocellular carcinoma (HCC) and develop a risk model.

We identified 150 DDR-related genes from the MSigDB database and constructed an 11-gene HCC prognostic signature using univariate Cox regression and random forest analyses. The robustness ofthe model was validated through internal and external validation. Additionally, we used Gene SetEnrichment Analysis (GSEA) to identify potential pathways associated with the risk model in HCCand analyzed the correlation between clinical traits and the risk score. Finally, we identified andvalidated two molecular subtypes of HCC using DDR gene expression. Our findings provide novelinsights into the molecular mechanisms of HCC and establish an independent DDR gene-basedprognostic signature. The use of this signature could aid in personalized therapy and improve clinicaldecision-making for HCC patients. With the increasing availability of sequencing data, this studyprovides a useful example of how these data can be utilized to better understand the underlyingbiology of cancer and improve patient outcomes.

2 Methods

2.1 Data collection

We obtained clinical data and gene expression information for HCC samples from the ICGC-LIRI (https://dcc.icgc.org/) and TCGA-LIHC (https://portal.gdc.cancer.gov/) datasets. Genes linked toDNA damage response (DDR) were collected from MSigDB, V7.1 (https://www.gseamsigdb.org/gsea/msigdb), and only those genes present in both datasets were retained for furtheranalysis.

2.2 Risk signature construction

We utilized univariate and multivariate Cox regression analyses to identify DDR-linked genes in the

LIRI-JP and LIHC datasets. To calculate the risk score for each patient, we used the equation: (Exp i * β i), where Expi represents the expression level of prognostic genes and β i represents the coefficient of cox regression for each prognostic gene. The median score was used to classify patients into high and low-risk groups, and survival differences were calculated using the "survival" and "survminer" packages. To determine the accuracy of the risk model for 1-, 3-, and 5-year survival, we utilized the "SurvivalROC" package (https://cran.rproject.org/web/packages/survivalROC/index.html). We also employed univariate and multivariate cox regression analyses to determine the prognostic independence of clinical features and the risk score. Potential pathways linked to low and high-risk groups were identified by GSEA, using c2.cgp.v7.1.symbols.gmt as the reference gene set.

2.3 Nomogram and DCA curve construction

We constructed a nomogram utilizing independent prognostic factors, and analyzed the benefit of the prognostic factor using decision curve analysis. The discriminative ability of the nomogram was assessed using a calibration plot with the bootstrap approach and 1,000 replications (12). Furthermore, we evaluated the benefit of the prognostic factor using decision curve analysis (13).

2.4 Consensus clustering

We utilized the "ConsensusClusterPlus" R package (with 50 iterations and 80% resampling samples) to group patients into distinct clusters based on DDR-related genes, with the aim of determining molecular subtypes of HCC (14). Principal components analysis (PCA) was employed to distinguish between various LIHC subgroups, and all analyses were validated using the LIRI-JP dataset.

2.5 Immune infiltration analysis

We evaluated the enrich score of immune cells and infiltration levels of 28 immune cells for each sample in both high- and low-risk groups using the ssGSEA algorithm, which was implemented using the "GSVA" R package (15–17). Furthermore, we analyzed the expression of immune checkpoint genes in both groups.

2.6 Cell culture and transfection

Human HCC cell lines Hep G2 and MHCC-97H were purchased from (National collection of authenticated cell culture,

Shanghai, CN), and incubated at 37 °C C with 5% CO2 in a humidity saturated environment. Cells were cultured in DMED (Hyclone, LA, USA) and supplied with 10% fetal bovine serum (BI, Israel), anti biotics (0.1 U/l penicillin and 100 g/l streptomycin). DGUOK siRNA were obtained from RiboBio Co., Ltd. (Guangzhou, China). The siRNA was dissolved in DEPC-treated water. Lipofectamine 2000 reagent (Invitrogen, CA, USA) were used for transfection according to the manufacturer's protocol. The solutions were mixed together and incubated at room temperature for 30 minutes. 30 nM siRNA was added into each well andincubated at 37 °C

2.7 Hoechst staining

Cell apoptosis was observed by the morphological changes of the cell nucleus (chromatin agglutination or DNA fragmentation). Cells were treated with si-NC or si-DGUOK, and washed with

PBS twice, Hoechst 33258 (1 μg/ml) was added for 20 min at room temperature avoiding light.Images were gathered by fluorescence microscope (Nikon, Japan)

2.8 Cell viability assays

Cells were seeded in 96-well plates at 10,000 cells per well, and cultured for 24h. They were treated with si-NC or si-DGUOK. Then CCK-8 were added to each plate, absorbance was measured at 450 nm using a FLUOstar Omega microplate reader (BMG Labtech). Cell viability of samples wascalculated according to the manufacturer's instructions

2.9 Statistical analyses

The statistical analysis was performed using R (https://www.r-project.org/). Kaplan-Meier (KM)method was employed to analyze the survival data and a p-value less than 0.05 was consideredstatistically significant.

3 Results

3.1 Identification of survival-related DDR risk model

150 DDR-related genes data were obtained from the TCGA dataset, which consisted of 343 HCCsamples. Using univariate Cox regression analysis, 37 prognostic genes that affect the survival ofpatients with HCC were identified. For developing a risk model, stepwise multivariate Coxregression analysis was conducted and 11 genes (AAAS, CANT1, CLP1, DGUOK, GTF2B,GTF2H1, NCBP2, POLA1, POLE4, POLR2D, and POLR2E) were selected. The risk score for eachpatient was calculated using the following method and computation: AAAS * -0.022 + CANT1 *0.016 + CLP1 * -0.098 + DGUOK * -0.016 + GTF2B*0.018 + GTF2H1 * 0.034 + NCBP2 * 0.042 +POLA1 * 0.089 + POLE4 * 0.015 + POLR2D * 0.047 + POLR2E * 0.007.

The patients were then divided into low and high-risk groups based on the median risk score. Asshown in Figure 1A, patients in the low-risk group had a longer expected survival rate compared tothose in the high-risk group. Furthermore, KM analysis confirmed better prognosis in the low-riskgroup than in the high-risk group ($p<0.001$) (Figure 1B). The predictive performance of the riskmodel was assessed using ROC analysis, and the area under curve (AUC) values for 1- and 3-yearsurvival were 0.76 and 0.66, respectively (Figure 1C), indicating good accuracy.

3.2 External validation of the DDR-gene prognostic signature

To evaluate the reliability and robustness of the 11-gene signature, we obtained a dataset of 231 HCC samples from ICGC (https://dcc.icgc.org/). Risk scores were computed for each patient, and thecohort was divided into high- and low-risk groups. Consistent with the previous findings, themajority of surviving cases were classified into the low-risk group, while a smaller proportion ofsurviving patients were classified into the high-risk group with higher mortality rate (Figure 2A).KM analysis confirmed better survival outcomes for individuals in the low-risk group compared totheir high-risk counterparts (Figure 2B). The AUC values for 1- and 3-year survival were 0.77 and0.73, respectively, indicating a good prognostic performance of the risk model in HCC (Figure 2C).

3.3 The risk model is an independent prognostic predictor of HCC

Univariate and multivariate Cox regression analyses showed that the risk model and tumor stagewere independent risk factors for HCC, as reported in Figures 3A, B. Furthermore, ROC analysisdemonstrated that the risk model performed better than tumor stage in predicting 1-year prognosis,with AUC values of 0.746 and 0.700, respectively (Figure 4A). Subsequently, we constructed anomogram that integrated the risk model and tumor stage to predict overall survival (OS) at 1-, 3-,and 5-year timepoints (Figure 5A). The nomogram exhibited good prognostic performance, asindicated by AUC values at 1-, 3-, and 5-year timepoints (Figures 4B, C). The stability of thenomogram was further validated by calibration curve plots (Figures 5B–D). Overall, the DDR genebased risk score and tumor stage-based nomogram can robustly predict the prognosis of HCCpatients and thus, can be useful in clinical decision-making.

3.4 Gene set enrichment analysis

To identify enriched pathways in HCC, GSEA was conducted for both high- and low-risk groups.

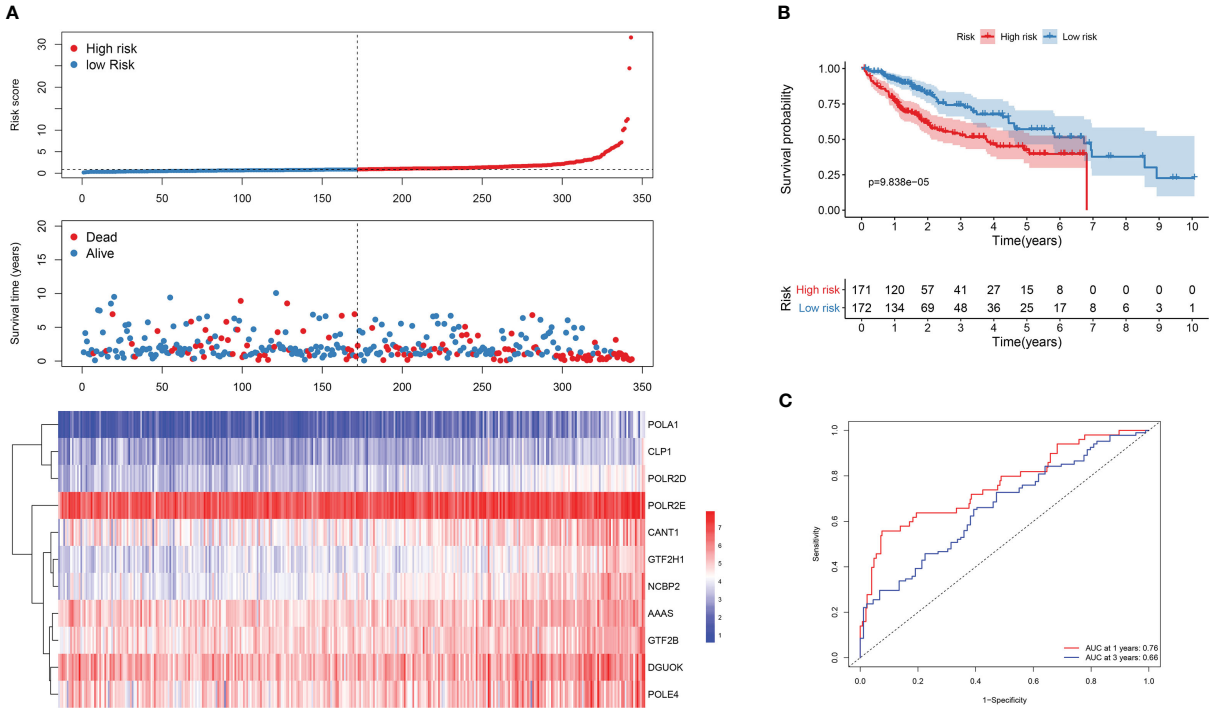


FIGURE 1
Survival analysis of DDR genes in TCGA dataset. **(A)** Risk score plot for the DDR signature. Upper panel is demonstrating the risk score distribution, middle panel is for case distribution, and the lower panel indicates the level of expression of 11 DDR genes. **(B)** KM survival curves of bothgroups. **(C)** ROC curve analysis of the risk gene signature.

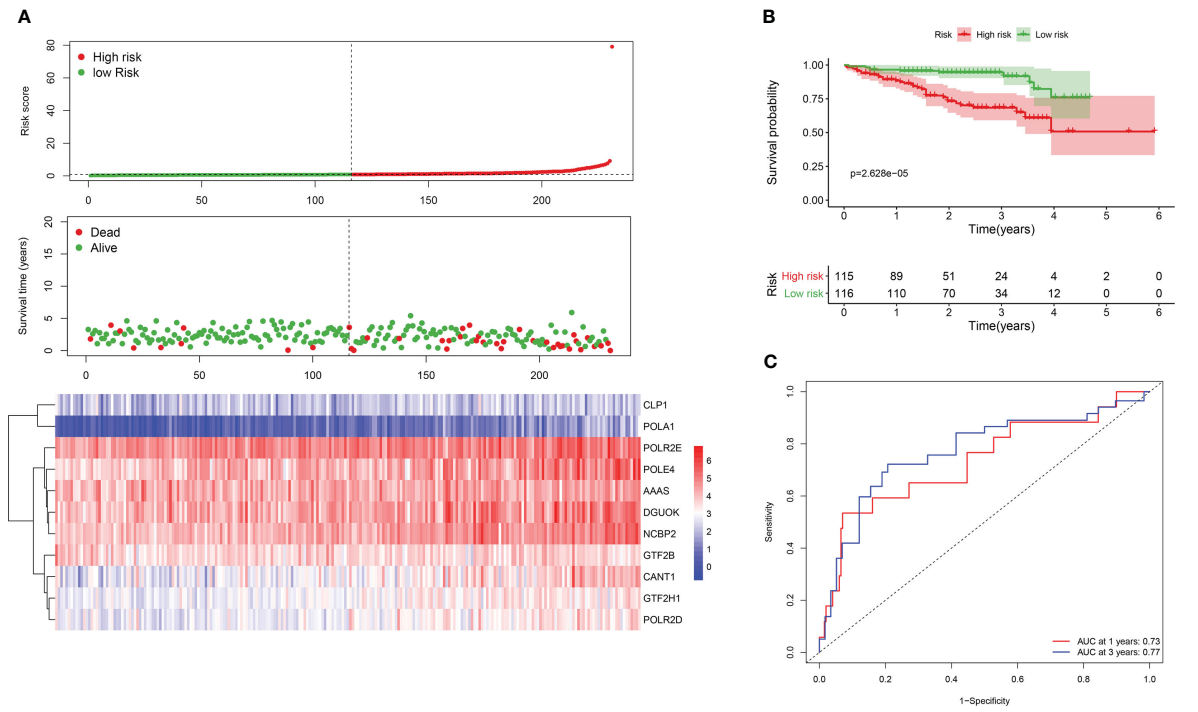


FIGURE 2
Survival analysis of DDR genes in ICGC dataset. **(A)** Risk score plot for the DDR signature. Upper panel is demonstrating the risk score distribution. Lower panel shows the level of expression of the 11 DDR genes and middle is for case distribution **(B)** KM survival curves for both groups. **(C)** ROC curve for risk gene signature.

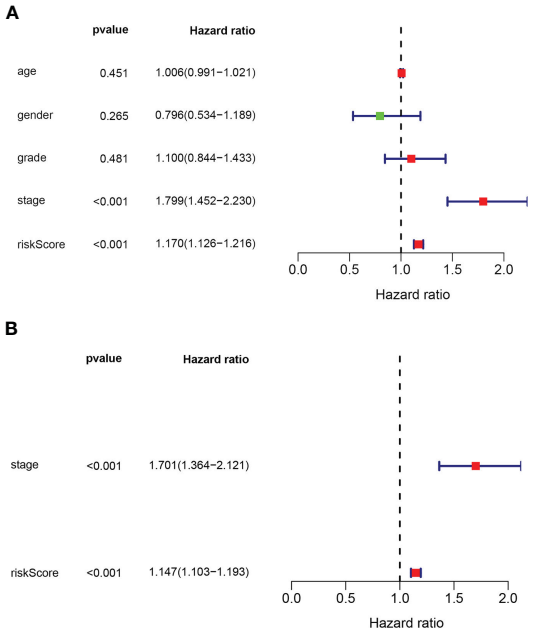


FIGURE 3
(A) Univariate and **(B)** multivariate cox regression analyses used to get prognostic value of the gene biosignature and clinical traits.

Results showed that pathways associated with apoptosis, cell cycle, and MAPK, mTOR, NOCTH, UBIQUITIN, and WNT signaling were enriched in the high-risk group, while the low-risk groupexhibited enrichment of pathways related to metabolism of fatty acid and retinol (Figures 6A, B).These findings suggest that favorable prognosis and low-risk scores are correlated with metabolismlinked pathways, whereas cancer-related pathways coincide with high-risk scores and poor prognosis.

3.5 Identification of molecular subtypes of HCC

Consensus clustering algorithm based on the DDR genes was employed to identify the molecularsubtypes of HCC. The optimal cluster was determined to be K=2 using the cumulative distributionfunction curve and the consensus heatmap (Figures 7A–C). PCA further differentiated patients intotwo distinct subgroups (Figure 7D), with subgroup 1 exhibiting better overall survival compared tosubgroup 2 (Figure 7E). To validate the robustness of the classification, we also evaluated thesubgroups in the ICGC dataset (Figure 8). Further analysis of the correlation and clinicalcharacteristics in both TCGA and ICGC datasets for the two subgroups showed that the group withbetter survival outcomes included more early-stage cases (Figures 9A, B).

3.6 Correlation of the immune infiltration with HCC subclasses

The ssGSEA algorithm was used to analyze the infiltration of 24 immune cells in both the high- and low-risk groups. The high-risk

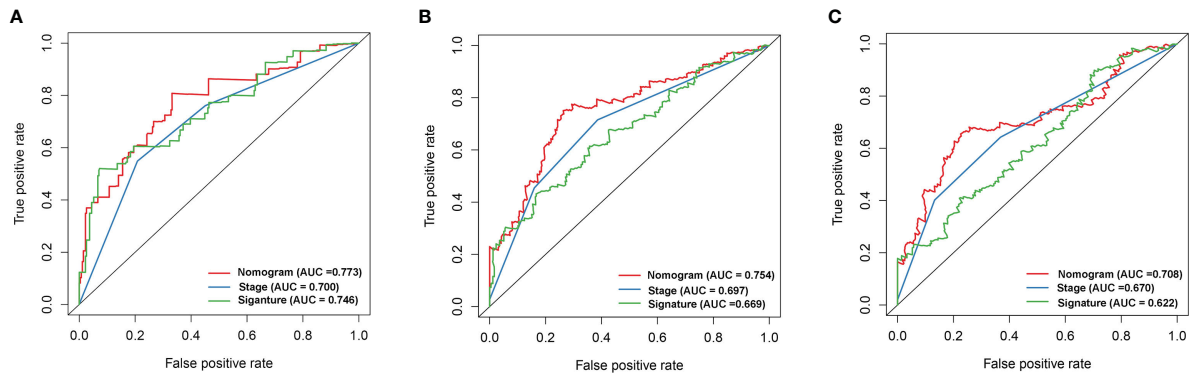


FIGURE 4
The gene signature, ROC curve analysis of the nomogram, and disease stage for 1- **(A)**, 3- **(B)**, and 5-year **(C)** survival.

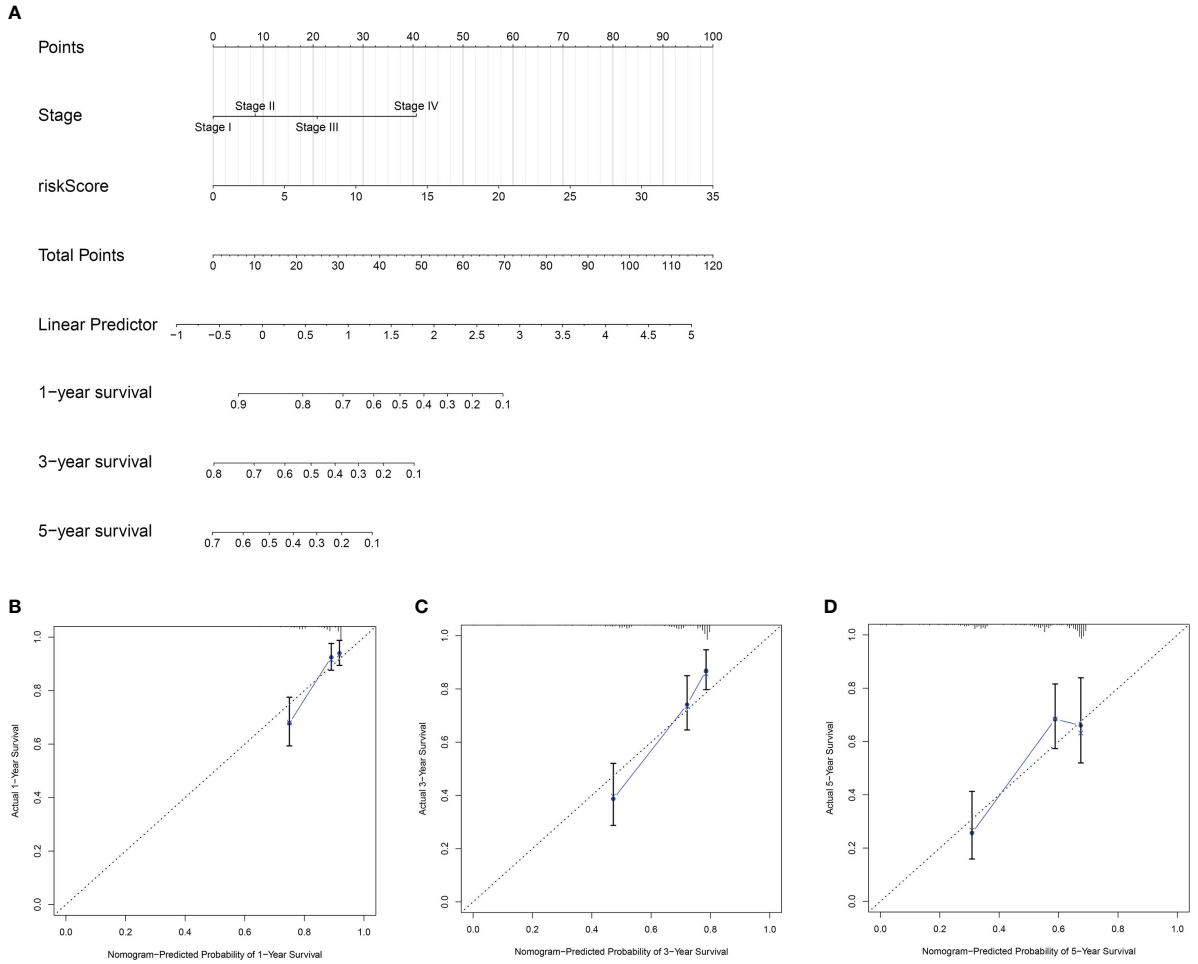


FIGURE 5
Construction and validation of a nomogram(prognostic). **(A)** A nomogram based on the risk model and tumor stage to estimate overall survival of HCC patients. The estimating of 1- **(B)**, 3- **(C)**and 5-year **(D)** survival of HCC patients using calibration curve plot of the nomogram.

group showed high levels of infiltration of activated CD4+ T cells, CD4+ T (central memory) CD8+ T cells (central memory), CD4+ T cells (effector memory), B cells(memory), regulatory T cells, T follicular helper cells, Th17 cells, Th2 cells, activated CD8+ T cells,immature dendritic cells (DCs), and plasmacytoid DCs, whereas the low-risk group had greaterinfiltration of activated CD8+ T cells and eosinophils (Figure 10A). Additionally, the high-risk group exhibited comparatively higher expression of all inhibitory immune receptors compared to the low-risk group (Figure 10B). These findings suggest that the anti-tumor

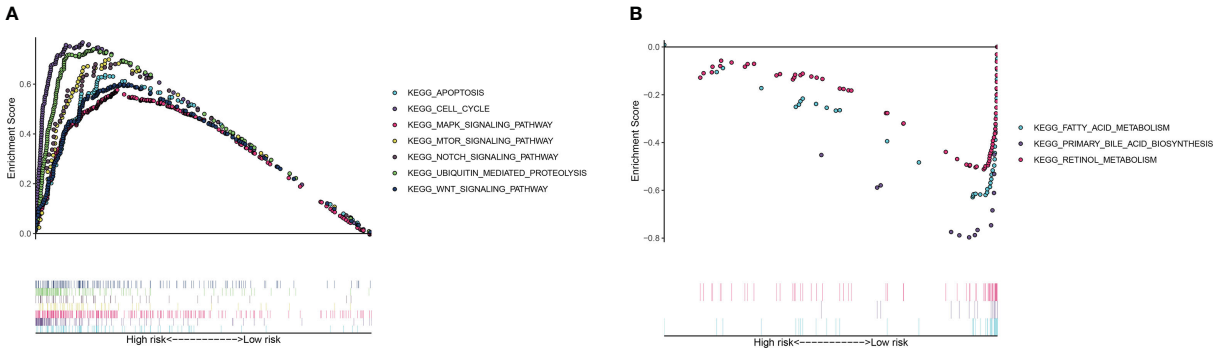


FIGURE 6
Gene set enrichment analysis results showing enriched pathways in **(A)** high-risk and **(B)** low-risk groups.

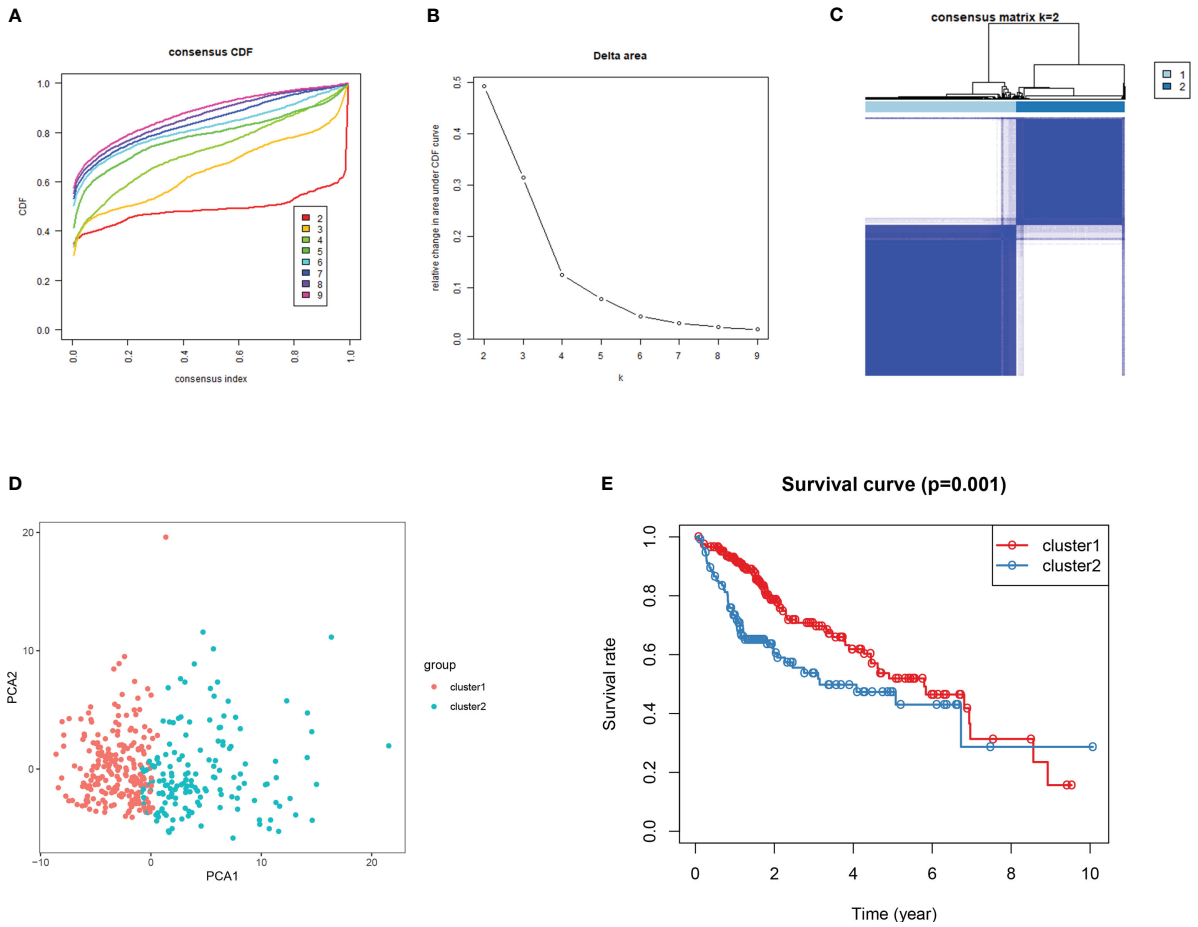


FIGURE 7
Consensus clustering for DDR genes in HCC patients from TCGA dataset. **(A)** Thecumulative distribution function (CDF) curve plot for k = 2 to k = 9 **(B)**. The change in the areaunder the CDF curve when k = 2 to k = 9. **(C)** Consensus heatmap at k =2. **(D)** Principal componentsanalysis for the DDR gene expression. **(E)** For the 2 subgroups, KM survival curve analysis.

properties of high T cell infiltration were offset by a strong immunosuppressive tumor microenvironment due to the overexpression of immune checkpoint proteins (18).

3.7 Cell assays

In vitro validation on DGUOK. HepG2 and MHCC-97H cells were treated with CCK8 and were performed to detect the cell viability. Hoechst 33258 fluorescent dye staining was used to show nuclear morphological changes and to assess apoptosis. Data in A are presented as means ± SD. **, p< 0.01. magnification: 200×. The results showed the proliferation ability is higher in the si-NC group compared by the si-DGUOK group (Figure 11).

3.8 MiRNA-mRNA Network

Based on the targetscan database data, we performed a relevant microRNA analysis of DNA repair genes in Siganture using the Cytoscape software. and conducted DNA repair related miRNA-mRNA Network (Figure 12).

4 Discussion

HCC is a highly heterogeneous cancer with multiple risk factors, including alcohol consumption, hepatitis B/C infection, and obesity (19). The initiation of HCC is associated with DNA damage and chromosomal abnormalities, which triggers a DNA damage response (DDR) in affected cells. DNA lesions can be repaired through various mechanisms, including homologous recombination, mismatch repair, and double-strand break repair. Dysfunctional DDR pathways can result in genomic instability, mutations, and eventually lead to HCC development and progression (20). Many DNA repair proteins, such as sphingolipid signaling, TP53, hOGG1, XRCC1, PARP-1, MRE11-Rad50NBS1 (MRN) complex, and ataxia-telangiectasia mutant (ATM) kinase, are frequently mutated in HCC (21). Furthermore, ionizing radiation (IR)-induced DDR pathways can create an

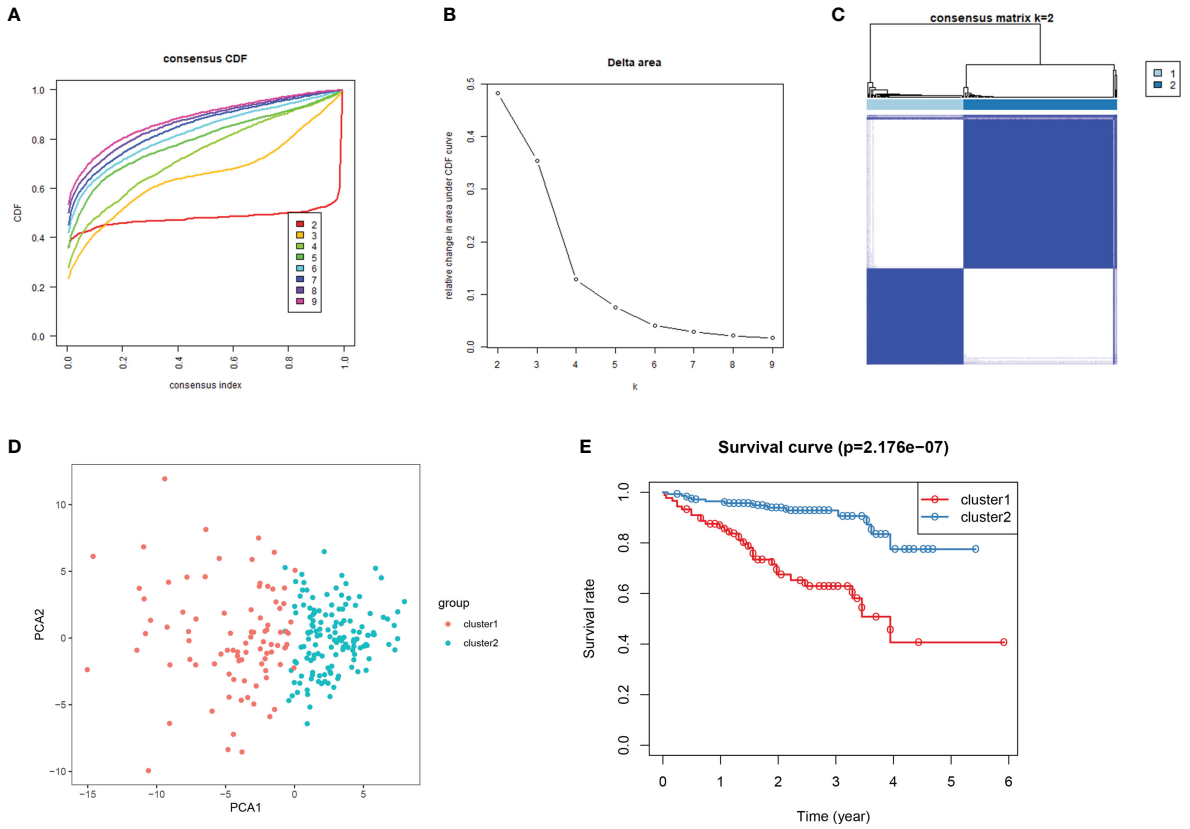


FIGURE 8
Consensus clustering for DDR genes in HCC patients from the ICGC cohort. **(A)** At k = 2 to k = 9, the cumulative distribution function (CDF) curve plot. **(B)**. The change in area under CDFcurve at two different values of k viz: k = 2 to k = 9. **(C)** At k =2, the consensus heatmap **(D)** Principal components analysis for the expression of DDR gene **(E)** KM survival curve analysis forthe 2 subgroups.

immunosuppressive tumor microenvironment, thereby weakening the anti-tumor effect of radioimmunotherapy. DDR inhibitors can reverse the immunosuppressive state of HCC and inhibit tumor progression, providing a potential therapeutic strategy (22).

There is a substantial body of evidence suggesting that DNA damage response (DDR) genes play a crucial role in the development of cancer (23). DDR genes are often expressed abnormally in mucosal or tumor tissues and are closely associated with patient prognosis (24, 25). However, the ability of individual genes to serve as prognostic markers is limited (26, 27), and multi-gene signatures may be better suited for predicting the prognosis of hepatocellular carcinoma (HCC). Despite this, no study has yet investigated the prognostic value of DDR genes in HCC. To address this gap, we developed an 11-gene signature based on DDR gene expression data and clinical data obtained from the ICGC and TCGA databases. The risk score generated by the 11-gene signature enabled the classification of patients into low- and high-risk groups, with the latter group exhibiting poorer survival outcomes. The risk model demonstrated good predictive performance in both TCGA and ICGC datasets. Additionally, the risk model was found to be an independent prognostic factor for HCC. A nomogram constructed using the risk score and tumor stage allowed the clear differentiation of two prognostic groups, which may be helpful in guiding preoperative management of HCC patients. The DDR gene signature identified in this study was found to be linked with several cancer related pathways including cell cycle, WNT signaling, mTOR signaling and apoptosis in the high risk group, which may be indicative of the potential mechanisms underlying HCC progression. On the other hand, the low-risk group was enriched in metabolism-related pathways. Most of the genes in the DDR-based risk signature have been implicated in tumorigenesis. For instance, CANT1 is known to regulate pyrimidine metabolism in melanoma cells and is associated with tumor progression (28). High expression of CANT1 in prostate cancer cells has been associated with better prognosis, while its silencing significantly suppressed cell proliferation and DNA synthesis (29). CLP1, on the other hand, plays an important role in motor neuron function (30). Mitochondrial deoxyguanosine kinase (DGUOK) is an enzyme that controls the rate of deoxy nucleoside salvage pathway in the mitochondria. Overexpression of DGUOK has been associated with worse prognosis in lung cancer, and its depletion suppressed lung adenocarcinoma growth, CSC self-renewal and metastasis (31). GTF2B has been identified as a prognostic marker for colorectal cancer and neuroblastoma, while GTF2H1 is a p62 subunit of complex transcription factor IIH (TFIIH) that regulates nucleotide excision repair and transcription (32, 33). Certain polymorphisms/haplotypes of GTF2H1 have been associated with increased susceptibility to

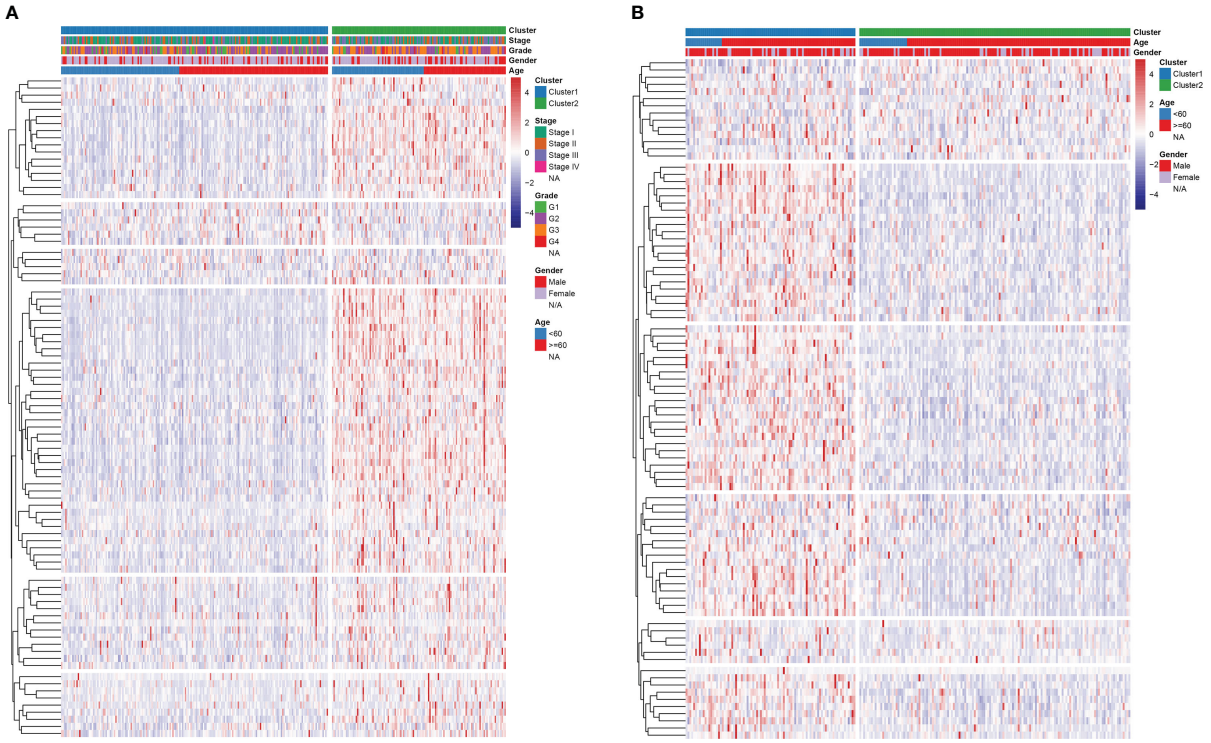


FIGURE 9
Heatmap analysis of the relationship between subgroup and clinical traits in **(A)** TCGA and **(B)** ICGC cohorts.

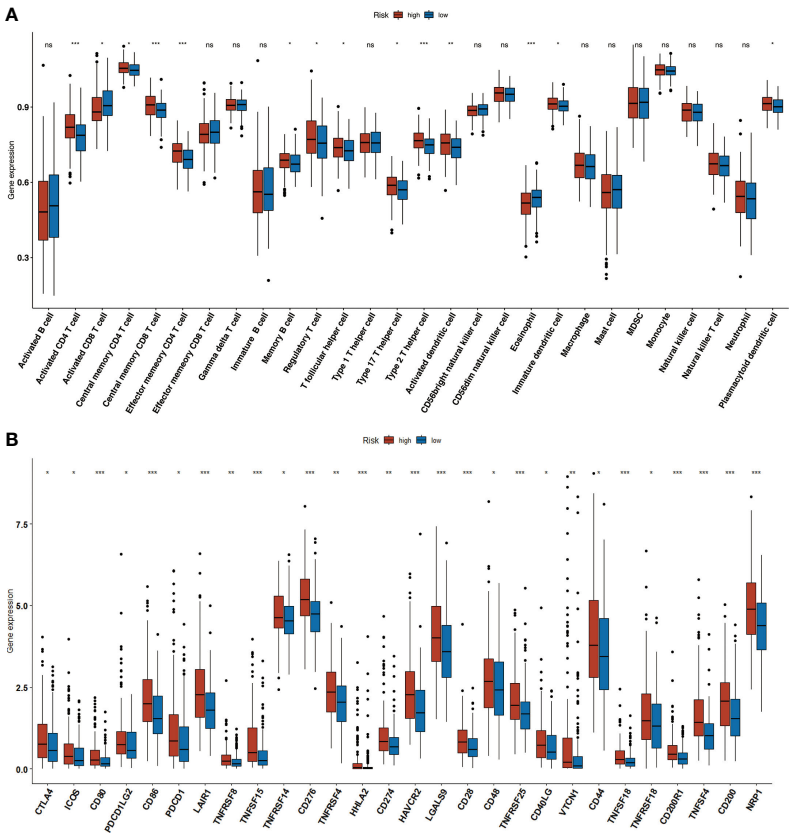


FIGURE 10
Immune infiltration. **(A)** The increased number of tumor infiltrating immune cells in high-and low-risk groups. **(B)** Boxplot showing the immune-checkpoint genes expression in low- andhigh-risk groups. * p<0.05, ** p <0.01, *** p<0.001.

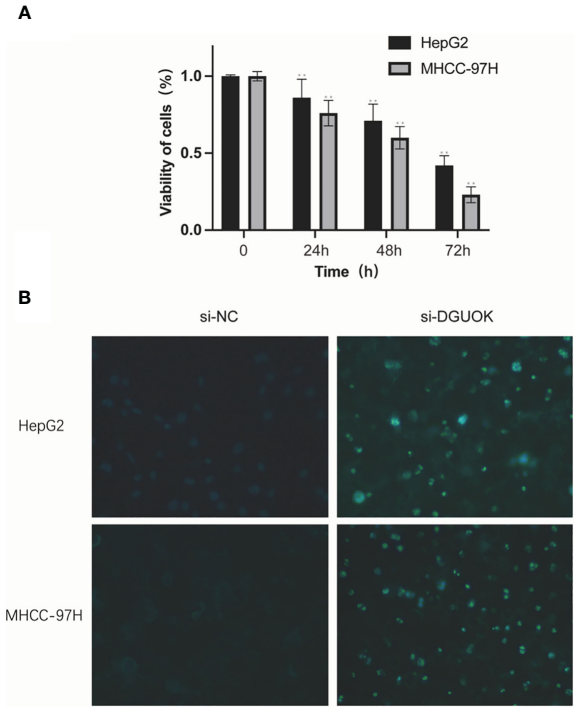


FIGURE 11
In vitro validation on DGUOK. HepG2 and MHCC-97H cells were treated with siDGUOK for indicated time. **(A)**: CCK8 was performed to detect the cell viability. **(B)**: Hoechst 33258 fluorescent dye staining was used to show nuclear morphological changes and to assess apoptosis. Data in **(A)** are presented as means ± SD. **p< 0.01. magnification: 200×.

lung cancer (34). Additionally, the budding yeast orthologs of POLE4 have been shown to enhance Polε processivity *in vitro*, but have the opposite effect *in vivo*, leading to accelerated tumorigenesis (35).

Moreover, the POLR2E rs3787016 polymorphism may enhance the risk of developing the prostatecancer, liver cancer esophageal cancer, papillary thyroid carcinoma, and breast cancer (36–38). However, the role of AAAS, NCBP2, POLA1 and POLR2D in HCC

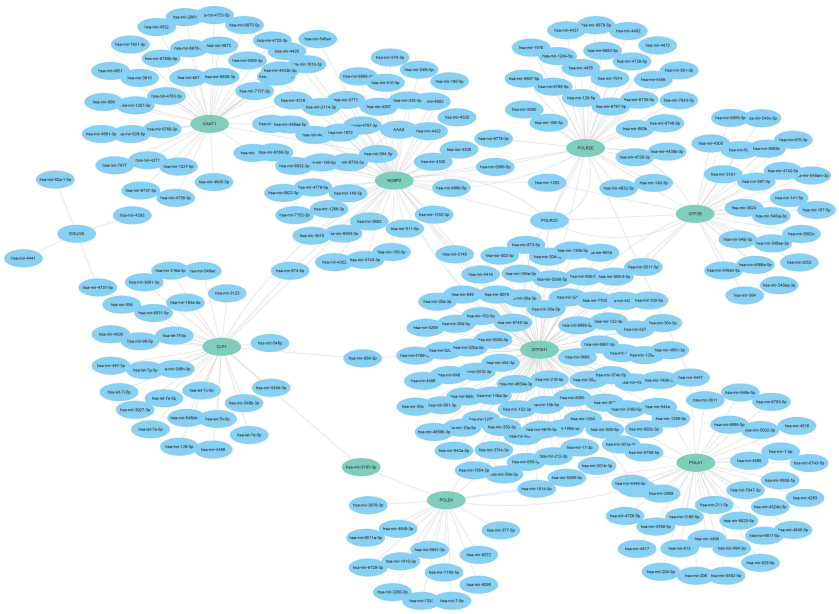


FIGURE 12
DNA Repair Gene miRNA-mRNA Network.

is unknown, and will have to beexperimentally verified. Immunotherapy has achieved encouraging results in various malignancies (39), including HCC. For instance, the "T+A" scheme is increasingly becoming the first-line optionfor advanced HCC (40). Despite achieving good outcomes in multiple cancers, a significantpercentage of the patients do not benefit from immunotherapy. Hence it is necessary to recognize thebiomarkers that can reveal the outcomes of immunotherapy, and screen for patients that can respondto immunotherapeutic regimens. Galon et al. (41) had proposed the concept of "cold" and "hot" tumors to evaluate their sensitivity to immunotherapies. In this study, we detected increased infiltration of immunosuppressive cells and overexpression of receptors responsible for immune checkpoint in the high-risk group, which indicates that the high-risk group patients are likely unresponsive to immunotherapy.

In summary, we identified biomarkers of HCC based on computational biology in oncologymethods (42, 43), and constructed prognostic models using machine learning methods (44–46). Wehave established an 11-DDR gene signature that can accurately forecast the prognosis ofhepatocellular carcinoma (HCC). The utilization of this prognostic signature not only advances our comprehension of the underlying molecular mechanisms that contribute to HCC progression but also provides a practical guide for clinical decision-making.

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 human participants were reviewed and approved by Intensive Care Medicine, Liaoyou Baoshihua Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YB contributed to data analysis, methodology, figures construction, and article writing. YWcontributed to investigation and validation. JH, YM, HL and ML contributed to methodology and validation. All authors contributed to the article and approved the submitted version.

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. 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(6):394–424. doi: 10.3322/caac.21492

2. Davis GL, Dempster J, Meler JD, Orr DW, Walberg MW, Brown B, et al. Hepatocellular carcinoma: management of an increasingly common problem. *Proc (Bayl Univ Med Cent)* (2008) 21(3):266–80. doi: 10.1080/08998280.2008.11928410

3. Marengo A, Rosso C, Bugianesi E. Liver cancer: connections with obesity, fatty liver, and cirrhosis. *Annu Rev Med* (2016) 67:103–17. doi: 10.1146/annurev-med-090514-013832

4. McGlynn KA, Petrick JL, London WT. Global epidemiology of hepatocellular carcinoma: an emphasis on demographic and regional variability. *Clin Liver Dis* (2015) 19(2):223–38. doi: 10.1016/j.cld.2015.01.001

5. Lans H, Hoeijmakers JHJ, Vermeulen W, Marteijn JA. The DNA damage response to transcription stress. *Nat Rev Mol Cell Biol* (2019) 20(12):766–84. doi: 10.1038/s41580-019-0169-4

6. Nickoloff JA, Jones D, Lee SH, Williamson EA, Hromas R. Drugging the cancers addicted to DNA repair. *J Natl Cancer Inst* (2017) 109(11). doi: 10.1093/jnci/djx059

7. Laurini E, Marson D, Fermeglia A, Aulic S, Fermeglia M, Pricl S. Role of Rad51 and DNA repair in cancer: a molecular perspective. *Pharmacol Ther* (2020) 208:107492. doi: 10.1016/j.pharmthera.2020.107492

8. Baretti M, Le DT. DNA Mismatch repair in cancer. *Pharmacol Ther* (2018) 189:45–62. doi: 10.1016/j.pharmthera.2018.04.004

9. Mateo J, Boysen G, Barbieri CE, Bryant HE, Castro E, Nelson PS, et al. DNA Repair in prostate cancer: biology and clinical implications. *Eur Urol* (2017) 71(3):417–25. doi: 10.1016/j.eururo.2016.08.037

10. Wang X, Tan C, Ye M, Wang X, Weng W, Zhang M, et al. Development and validation of a DNA repair gene signature for prognosis prediction in colon 334 cancer. *J Cancer* (2020) 11(20):5918–28. doi: 10.7150/jca.46328

11. Lin J, Shi J, Guo H, Yang X, Jiang Y, Long J, et al. Alterations in DNA damage repair genes in primary liver cancer. *Clin Cancer Res* (2019) 337 25(15):4701–11. doi: 10.1158/1078-0432.CCR-19-0127

12. Iasonos A, Schrag D, GV R. Panageas KS: how to build and interpret a nomogram for cancer prognosis. *J Clin Oncol* (2008) 26(8):1364–70. doi: 10.1200/JCO.2007.12.9791

13. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* (2006) 26(6):565–74. doi: 10.1177/0272989X06295361

14. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* (2010) 26(12):1572–3. doi: 10.1093/bioinformatics/btq170

15. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf* (2013) 14:7. doi: 10.1186/1471-2105-14-7

16. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep* (2017) 18(1):248–62. doi: 10.1016/j.celrep.2016.12.019

17. Jia Q, Wu W, Wang Y, Alexander PB, Sun C, Gong Z, et al. Local mutational diversity drives intratumoral immune heterogeneity in non-small cell lung cancer. *Nat Commun* (2018) 9(1):5361. doi: 10.1038/s41467-018-07767-w

18. Matsushita H, Sato Y, Karasaki T, Nakagawa T, Kume H, Ogawa S, et al. Neoantigen load, antigen presentation machinery, and immune signatures determine

prognosis in clear cell renal cell carcinoma. *Cancer Immunol Res* (2016) 4(5):463–71. doi: 10.1158/2326-6066.CIR-15-0225

19. Villanueva A. Hepatocellular carcinoma. *N Engl J Med* (2019) 380(15):1450–62. doi: 10.1056/NEJMra1713263

20. O'Connor MJ. Targeting the DNA damage response in cancer. *Mol Cell* (2015) 60(4):547–357 560. doi: 10.1016/j.molcel.2015.10.040

21. Yang SF, Chang CW, Wei RJ, Shiue YL, Wang SN, Yeh YT. Involvement of DNA damage response pathways in hepatocellular carcinoma. *BioMed Res Int* (2014) 2014:153867. doi: 10.1155/2014/153867

22. Sheng H, Huang Y, Xiao Y, Zhu Z, Shen M, Zhou P, et al. ATR inhibitor AZD6738 enhances the antitumor activity of radiotherapy and immune checkpoint inhibitors by potentiating the tumor immune microenvironment in hepatocellular carcinoma. *J Immunother Cancer* (2020) 8(1). doi: 10.1136/jitc-2019-000340

23. Ali R, Rakha EA, Madhusudan S, Bryant HE. DNA Damage repair in breast cancer and its therapeutic implications. *Pathology* (2017) 49(2):156–65. doi: 10.1016/j.pathol.2016.11.002

24. Jinjia C, Xiaoyu W, Hui S, Wenhua L, Zhe Z, Xiaodong Z, et al. The use of DNA repair genes as prognostic indicators of gastric cancer. *J Cancer* (2019) 10(20):4866–75. doi: 10.7150/jca.31062

25. Kun S, Duan Q, Liu G, Lu JM. Prognostic value of DNA repair genes based on stratification of glioblastomas. *Oncotarget* (2017) 8(35):58222–30. doi: 10.18632/oncotarget.17452

26. Deng CX, Wang RH. Roles of BRCA1 in DNA damage repair: a link between development and cancer. *Hum Mol Genet* (2003) 12 Spec No 1:R113–123. doi: 10.1093/hmg/ddg082

27. Fang M, Xia F, Mahalingam M, Virbasius CM, Wajapeyee N, Green MR. MEN1 is a melanoma tumor suppressor that preserves genomic integrity by stimulating transcription of genes that promote homologous recombination-directed DNA repair. *Mol Cell Biol* (2013) 33(13):2635–375 2647. doi: 10.1128/MCB.00167-13

28. Edwards L, Gupta R, Filipp FV. Hypermutation of DPYD deregulates pyrimidine metabolism and promotes malignant progression. *Mol Cancer Res* (2016) 14(2):196–206. doi: 10.1158/1541-7786.MCR-15-0403

29. Gerhardt J, Steinbrech C, Buchi O, Behnke S, Bohnert A, Fritzsche F, et al. The androgen-regulated calcium-activated nucleotidase 1 (CANT1) is commonly overexpressed in prostate cancer and is tumor-biologically relevant *in vitro*. *Am J Pathol* (2011) 178(4):1847–60. doi: 10.1016/j.ajpath.2010.12.046

30. Hanada T, Weitzer S, Mair B, Bernreuther C, Wainger BJ, Ichida J, et al. Komnenovic V et al: CLP1 links tRNA metabolism to progressive motor-neuron loss. *Nature* (2013) 495(7442):474–80. doi: 10.1038/nature11923

31. Lin S, Huang C, Sun J, Bollt O, Wang X, Martine E, et al. The mitochondrial deoxyguanosine kinase is required for cancer cell stemness in lung adenocarcinoma. *EMBO Mol Med* (2019) 11(12):e10849. doi: 10.15252/emmm.201910849

32. Di Pietro C, Ragusa M, Barbagallo D, Duro LR, Guglielmino MR, Majorana A, et al. Involvement of GTA protein NC2beta in neuroblastoma pathogenesis suggests that it physiologically participates in the regulation of cell proliferation. *Mol Cancer* (2008) 7:52. doi: 10.1186/1476-4598-7-52

33. Ribeiro-Silva C, Aydin OZ, Mesquita-Ribeiro R, Slyskova J, Helfricht A, Marteijn JA, et al. DNA Damage sensitivity of SWI/SNF-deficient cells depends on TFIIH subunit p62/GTF2H1. *Nat Commun* (2018) 9(1):4067. doi: 10.1038/s41467-018-06402-y

34. Wu W, Liu H, Lei R, Chen D, Zhang S, Lv J, et al. Genetic variants in GTF2H1 and risk of lung cancer: a case-control analysis in a Chinese population. *Lung Cancer* (2009) 63(2):180–6. doi: 10.1016/j.lungcan.2008.05.014

35. Bellelli R, Borel V, Logan C, Svendsen J, Cox DE, Nye E, et al. Polepsilon instability drives replication stress, abnormal development, and tumorigenesis. *Mol Cell* (2018) 70(4):707–721 e707.

36. Chen B, Li J, Yi C, Jiao Y, Gu X, Feng X. Long non-coding RNA POLR2E rs3787016 is associated with the risk of papillary thyroid carcinoma in Chinese population. *Pathol Res Pract* (2018) 214(7):1040–4.

37. Chen B, Jiao Y, Yaolong F, Li T, Liu Y, Wang M, et al. The POLR2E rs3787016 polymorphism is strongly associated with the risk of female breast and cervical cancer. *Pathol Res Pract* (2019) 215(5):1061–5.

38. Chen B, Wang S, Ma G, Han J, Zhang J, Gu X, et al. The association of POLR2E rs3787016 polymorphism and cancer risk: a Chinese case-control study and meta-analysis. *Biosci Rep* (2018) 38(6).

39. Yang Y. Cancer immunotherapy: harnessing the immune system to battle cancer. *J Clin Invest* (2015) 125(9):3335–7.

40. Finn RS, Qin S, Ikeda M, PR G, Ducreux M, TY K, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *N Engl J Med* (2020) 382 (20):1894–905.

41. Galon J, Bruni D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat Rev Drug Discovery* (2019) 18(3):197–218. doi: 10.1038/s41573-018-0007-y

42. Wang Y, Yan K, Lin J, Li J, Bi J. Macrophage M2 Co-expression factors correlate with the immune microenvironment and predict outcome of renal clear cell carcinoma. *[J] .Front Genet* (2021) 12:615655. doi: 10.3389/fgene.2021.615655

43. Wang Y, Yan K, Wang J, Lin J, Bi J. M2 macrophage Co-expression factors correlate with immune phenotype and predict prognosis of bladder cancer. *Front Oncol* (2021) 11:609334(undefined). doi: 10.3389/fonc.2021.609334

44. Wang Y, Yan K, Wang L, Bi J. Genome instability-related long non-coding RNA in clear renal cell carcinoma determined using computational biology. *[J] .BMC Cancer* (2021) 21:727. doi: 10.1186/s12885-021-08356-9

45. Wang L, Shen J, Wang Y, Bi J. Identification of fatty acid metabolism-based molecular subtypes and prognostic signature to predict immune landscape and guide clinical drug treatment in renal clear cell carcinoma. *[J] .Int Immunopharmacol* (2023) 116:109735. doi: 10.1016/j.intimp.2023.109735

46. Hu J, Wang L, Li L, Wang Y, Bi J. A novel focal adhesion-related risk model predicts prognosis of bladder cancer – a bioinformatic study based on TCGA and GEO database. *[J] .BMC Cancer* (2022) 22:1158. doi: 10.1186/s12885-022-10264-5

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 23 May 2023
DOI 10.3389/fimmu.2023.1185916

OPEN ACCESS

EDITED BY
Jun Liu,
Yuebei People's Hospital, China

REVIEWED BY
Bufu Tang,
Fudan University, China
Jiahao Gao,
Fudan University, China

*CORRESPONDENCE
Yihang Yuan
surgeryuan@outlook.com
Peng Sun
sp2082@shtrhospital.com

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

RECEIVED 14 March 2023
ACCEPTED 10 May 2023
PUBLISHED 23 May 2023

CITATION
Sun W, Wang J, Wang Z, Xu M, Lin Q, Sun P and Yuan Y (2023) Combining WGCNA and machine learning to construct basement membrane-related gene index helps to predict the prognosis and tumor microenvironment of HCC patients and verifies the carcinogenesis of key gene CTSA.
Front. Immunol. 14:1185916.
doi: 10.3389/fimmu.2023.1185916

Combining WGCNA and machine learning to construct basement membrane-related gene index helps to predict the prognosis and tumor microenvironment of HCC patients and verifies the carcinogenesis of key gene CTSA

Weijie Sun[1,2†], Jue Wang[1†], Zhiqiang Wang[1], Ming Xu[1], Quanjun Lin[1], Peng Sun[1*] and Yihang Yuan[1*]

[1]Department of General Surgery, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [2]Department of Infectious Diseases, The First Affiliated Hospital of Anhui Medical University, Hefei, China

Hepatocellular carcinoma (HCC) is a malignant tumor with high recurrence and metastasis rates and poor prognosis. Basement membrane is a ubiquitous extracellular matrix and is a key physical factor in cancer metastasis. Therefore, basement membrane-related genes may be new targets for the diagnosis and treatment of HCC. We systematically analyzed the expression pattern and prognostic value of basement membrane-related genes in HCC using the TCGA-HCC dataset, and constructed a new BMRGI based on WGCNA and machine learning. We used the HCC single-cell RNA-sequencing data in GSE146115 to describe the single-cell map of HCC, analyzed the interaction between different cell types, and explored the expression of model genes in different cell types. BMRGI can accurately predict the prognosis of HCC patients and was validated in the ICGC cohort. In addition, we also explored the underlying molecular mechanisms and tumor immune infiltration in different BMRGI subgroups, and confirmed the differences in response to immunotherapy in different BMRGI subgroups based on the TIDE algorithm. Then, we assessed the sensitivity of HCC patients to common drugs. In conclusion, our study provides a theoretical basis for the selection of immunotherapy and sensitive drugs in HCC patients. Finally, we also considered CTSA as the most critical basement membrane-related gene affecting HCC progression. *In vitro* experiments showed that the proliferation, migration and invasion abilities of HCC cells were significantly impaired when CTSA was knocked down.

KEYWORDS

hepatocellular carcinoma, basement membranes, prognosis, immunotherapy, machine learning, ScRNA-seq, CTSA, *vitro* experiment

1 Introduction

Hepatocellular carcinoma (HCC) is responsible for about 90% of primary liver cancers (1). It is also one of the most fatal malignant tumors worldwide, with high morbidity and mortality rates (2, 3). The frequent occurrence of metastasis and recurrence is a major contributing factor to the poor prognosis of HCC patients (4). Despite the development of numerous drug combination strategies for the treatment of HCC, the current level of patient survival time has not yet met satisfactory standards. Consequently, there is an urgent need to identify new biomarkers that can more accurately predict the prognosis of HCC.

Basement membranes (BM) are a ubiquitous and unique type of extracellular matrix that plays a key role in cancer cell metastasis (5). In the case of HCC and the surrounding uninvolved liver tissue, the BM is primarily made up of three components: fibronectin (FN), laminin (LAM), and collagen IV (Coll IV) (6). BM is known to affect numerous physiological and pathological activities of cells including cell proliferation, adhesion, migration, and vascular remodeling (7, 8). As a result, in most cancers, BM plays a crucial role in driving cell metastasis (5, 9, 10). Due to the significant role of BM in cancer metastasis, it is an ideal target for anticancer drugs. Previous studies have found that stable markers can be created using different gene sets such as cuproptosis and necroptosis to predict the prognosis of HCC patients (11, 12). Recently, Jayadev et al. have redefined 222 BM-related genes (BMRG) and proteins (13), but a robust prognostic model based on BMRG in HCC is yet to be developed.

In this study, we screened 222 BMRG and identified 4 that were used to construct the basement membrane-related gene prognostic index (BMRGI). This index helps to more accurately predict the prognosis of patients with hepatocellular carcinoma (HCC). Furthermore, we evaluated the clinical relevance and impact of BMRGI on the tumor microenvironment. More importantly, we identified CTSA as a key BMRG in HCC, and comprehensively analyzed the expression differences of CTSA in HCC, and have confirmed that the expression of CTSA has a significant impact on the proliferation, migration, and metastasis of HCC cells.

2 Methods and materials

2.1 Data download and processing

The mRNA expression data of HCC patients with corresponding clinical information and somatic mutation data were downloaded from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) database, and the mRNA expression data and clinical information from the Japan-HCC cohort were downloaded at International Cancer Genome Consortium (ICGC, https://dcc.icgc.org/). When performing correlation data analysis, we excluded cases with missing data. Finally, when performing prognostic analysis, we chose to exclude cases with a survival time ≤30. The single-cell RNA seq (scRNA-seq) data of 4 HCC patients were obtained from GSE146115 in the GEO database (https://www.ncbi.nlm.nih.gov/geo/), with a total of 27227 genes and 3200 cells obtained.

2.2 Screening of WGCNA and differential BMRG

The "WGCNA" package was used to construct the gene co-expression network of BMRG in the TCGA-HCC dataset (14). The core module was considered the one with the highest Pearson coefficient and also the one most associated with clinical traits. Furthermore, we analyzed differentially expressed genes ($|\log_2 FC| > 0.585$, False Discovery Rate (FDR)< 0.05) in the TCGA-HCC dataset using the "Limma" package. Finally, we further investigated the 47 common genes.

2.3 Construction and verification of BMRGI

47 common genes were analyzed by univariate Cox analysis based on the "survival" package, and potential BMRG affecting the overall survival of HCC patients were screened out ($p<0.05$). Then, these candidate genes were analyzed by using the least absolute shrinkage and selection operator (LASSO). Based on the analysis results, we established a four-gene optimal prognostic model. The calculation formula of BMRGI for each HCC patient is as follows:

$$BMRGI = \sum_{i=1}^{n} Expression(i) * Coefficient(i)$$

Where X refers to the expression level of the selected gene, and Coef is the coefficient of the selected gene. In addition, the same calculation method is applied to the verification queue ICGC. According to the median BMRGI, HCC patients were divided into high BMRGI group and low BMRGI group. Kaplan–Meier curves were used to assess differences in OS between different BMRGI groups.

2.4 scRNA-seq data processing and analysis

The Seurat package is used for preprocessing and filtering of scRNA-seq (15). The PercentageFeatureSet function is used to calculate the mitochondrial gene content in cells. We further analyzed the cells in which the number of genes was >200 and the proportion of mitochondrial genes was<10%. We set the number of principal components (PC) to 20, the resolution to 0.4, and the 1500 genes with the largest variation between cells to cluster the cells.

2.5 CellChat analysis

We use CellChat to quantify and infer the communication links between different cell types from scRNA-seq data, and identify the

signal input and output among them. In this study, we filtered for cell communication of less than 10 cells.

2.6 Gene ontology analysis, kyoto encyclopedia of genes and genomes analysis and gene set enrichment analysis

The "Limma" package was used to analyze differentially expressed genes (DEGs) between high and low BMRGI groups ($|\log_2FC| > 1$, FDR< 0.05). GO analysis was performed based on the " clusterProfiler" package. KEGG analysis was performed based on the "clusterProfiler", "org.Hs.eg.db", "enrichplot" package. In addition, we also performed GSEA using the "clusterProfiler" package to explore biological differences among different BMRGI groups.

2.7 Analysis of immunological properties

The enrichment scores of 16 immune-related cells and 13 immune-related terms in HCC samples were calculated using the ssGSEA algorithm based on the R packages "GSVA" and "GSEABase".

We also summarized common immune checkpoint molecules and HLA family genes, and analyzed the correlation between BMRGI and the expression of each gene, and displayed it with a radar map. Furthermore, we explored the somatic mutation profile of TCGA-HCC samples and listed the top10 mutation-prone genes in different BMRGI subgroups. In addition, we also compared the difference of TMB in the high BMRGI group and the low BMRGI group. Finally, the TIDE (http://tide.dfci.harvard.edu/) algorithm was used to predict and evaluate the response of HCC patients to immunotherapy.

2.8 Sensitivity analysis of common drugs

We use the R package "oncoPredict" (16) for the evaluation of common drug sensitivities.

2.9 Identification of core BMRG

SVM-REF (17), LASSO (18)and RandomForest (19) are commonly used machine learning methods with excellent classification performance. In biology-related research, it is often used for the screening of characteristic genes (20). In this study, we use these three types of machine learning to filter out characteristic BMRG, and use intersection to filter out the most critical BMRG.

2.10 Multilevel expression verification of CTSA

We analyzed the differential expression of CTSA at the mRNA level of HCC tissues online from the GEPIA2 database (http://gepia2.cancer-pku.cn/#index) (combined samples from TCGA and GTEx databases) (21). In addition, we analyzed the expression differences of CTSA at the protein level in the CPTAC database online using the UCLCAN database (http://ualcan.path.uab.edu/index.html) (22). Finally, the HPA database obtained the immunohistochemical images of CTSA in normal liver tissues and HCC tissues (23), and obtained the basic information of the corresponding tissue samples.

2.11 RNA extraction, and real-time quantitative PCR

Cell total RNA was extracted using Trizol reagent (Invitrogen, USA)). RNA extraction and RT-qPCR as previously described (24). Briefly, RNA was reversed to cDNA using PrimeScript™ RT Master Mix (Takara Bio, JAPAN). Fluorescence quantification was performed by TB-Green qPCR (Takara Bio, JAPAN) and normalized to β-actin. The information of all designed primers is listed in Supplementary Table 1.

2.12 Cell culture, transient transfection

All cell lines used in this study (including normal liver cell line LO2 and HCC cell lines HEPG2, BEL7402 and HCCLM3) were donated by Dr. Dai (25). All cell lines were cultured in complete DMEM medium (DMEM medium with 10% fetal bovine serum and 1% penicillin-streptomycin). Transient transfections were performed using jetPRIME Transfection Reagent (Polyplus, China) and followed the manufacturer's instructions. siRNA sequences were designed by Tsingke Biotechnology Co., Ltd. The SiCTSA sequence is as follows, SiCTSA-1: sense-GCCUCUUUC CGGAGUACAA; antisense-UUGUACUCCGGAAAGAGGC. SiCTSA-2: sense-CUGCUUAGCUCACAGAAAU; antisense-AUUUCUGUGAGCUAAGCAG.

2.13 Cell counting kit-8 (CCK8) experiment

We planted 2×10^3 cells in a 96-well plate, and set 5 replicate wells in each group, cultured them for 0 hour, 24 hours and 48 hours, respectively, and then added 10ul CCK8 reagent (Targetmol, USA) and incubated at 37°C for 2 hours. Absorbance was then measured at 450 nm using a microplate reader (TECAN, Switzerland).

2.14 Transwell experiment

We planted 5×10^4 cells in the upper chamber (Corning, USA) containing 250ul serum-free medium. The upper chamber was without Matrigel (Corning, USA) for migration experiments, with Matrigel for invasion experiments, and the lower chamber add 800ul complete medium. After 24 hours of incubation, the cells were fixed with 4% paraformaldehyde and stained with 0.1% crystal violet. The cells on the upper surface of the upper chamber were wiped with a cotton swab, photographed under a microscope (Leica, Germany) at 100 times, and then counted.

2.15 Statistical analysis

All bioinformatics analyzes were performed on R software (version 4.1.2). Continuous variables that were not normally distributed were tested using the Wilcoxon test. Correlation analysis between BMRGI and drug IC50 was performed using the spearman method. The Kaplan-Meier method was used to draw the survival curves of different subgroups. All experimental data were analyzed for variance using Student's T-test. p or FDR< 0.05 represents a statistical difference.

3 Result

3.1 WGCNA identified BM key module genes

According to the Materials and methods section, we identified 222 BMRG. First, we conducted WGCNA on 222 BMRG. By setting a minimum of 25 genes within a module, module connectivity (Figure 1A), 4 modules were finally identified (Figure 1B). According to the correlation thermograms of the modules, we found that the blue modules had the highest correlations with clinical traits (Figure 1C). Therefore, we choose the blue module for further analysis. Second, we performed differential analysis on 222 genes, and the results showed a total of 131 DEGs, of which 122 BMRG were up-regulated and 9 BMRG were down-regulated (Figure 1D). Furthermore, we showed the correlation of blue module genes with DEGs by Venn diagram, and finally obtained 47 common genes (Figure 1E).

3.2 Construction of BMRGI for HCC patients

We first performed univariate Cox analysis on 47 common genes, and the results showed that 18 BMRG were risk factors affecting OS in HCC patients (Figure 2A). LASSO analysis was further performed on 18 prognostic genes, and finally we identified 4 BMRG (Figure 2B). In addition, we also analyzed the expression differences and prognostic value of the 4 BMRG. Differential analysis showed that CTSA, ADAM9, LAMB3, and SPON2 were highly expressed in HCC (Figure 2C), and kaplan-meier analysis showed that high expression usually means poor prognosis (Figure 2D). Finally, we constructed the basement membrane-related gene prognostic index BMRGI based on the results of LASSO analysis.

According to the median BMRGI, HCC patients were divided into high BMRGI group and low BMRGI group. We use TCGA-HCC as the training cohort and ICGC as the validation cohort. First, we showed the risk scores of the training cohort and validation

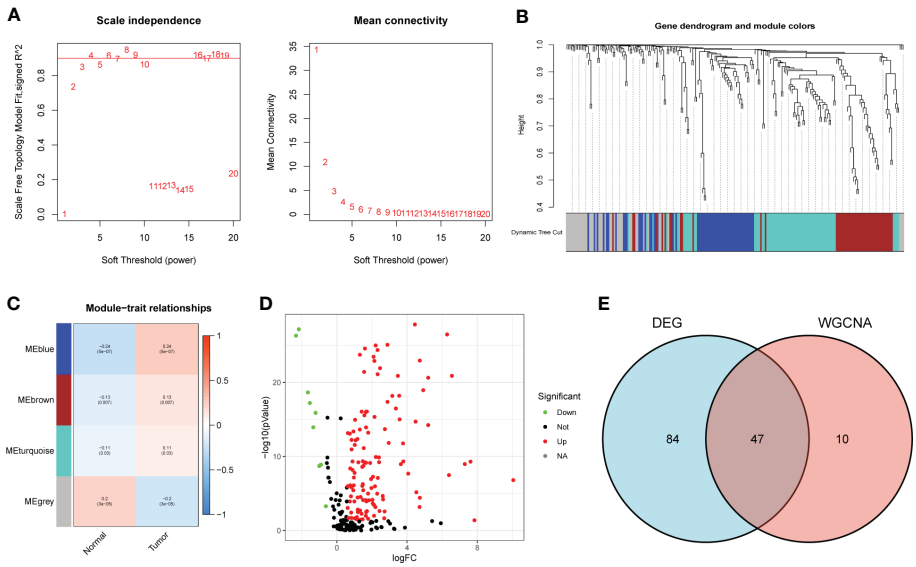


FIGURE 1
DEG screening of key module genes. **(A)** Scale independence and mean connectivity. **(B)** Gene clustering dendrogram, a total of 4 modules were identified. **(C)** Correlation heatmap of modules and clinical traits. **(D)** Volcano plot for BMRG differential analysis. **(E)** Venn diagram showing common genes of key module genes and differential genes.

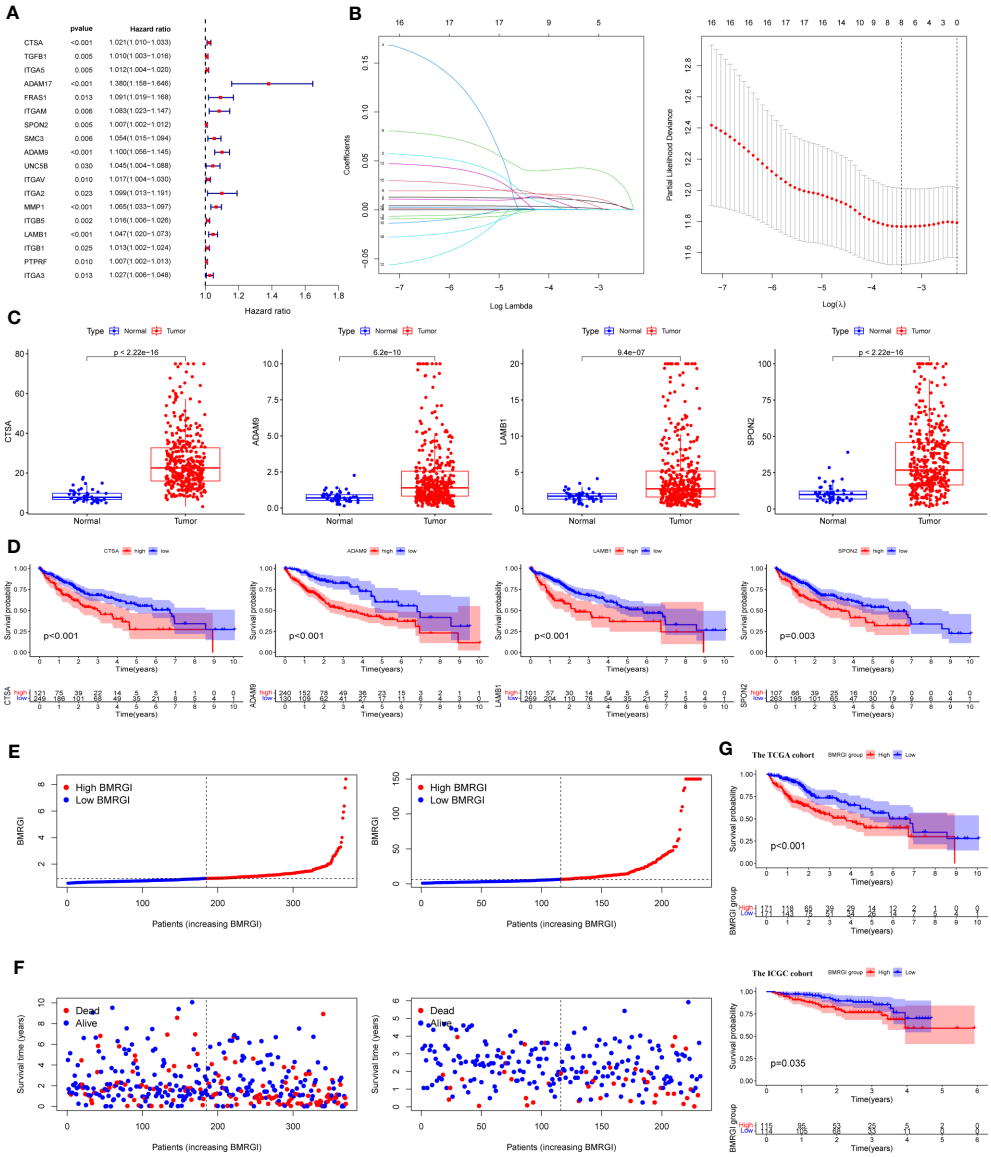


FIGURE 2
Construction of BMRGI and verification of its prognostic value. **(A)** Univariate Cox regression analysis of common genes. **(B)** LASSO analysis. **(C)** Expression difference analysis of CTSA, ADAM9, LAMB3 and SPON2. **(D)** Kaplan-Meier survival curves of CTSA, ADAM9, LAMB3 and SPON2. **(E)** Visual distribution of risk scores for TCGA cohort and ICGC cohort. **(F)** Survival status and time of TCGA and ICGC cohort. **(G)** Kaplan-Meier survival curves of BMRGI in TCGA and ICGC cohort.

cohort more intuitively (Figure 2E). Furthermore, in both the TCGA-HCC cohort and the ICGC cohort, the higher the BMRGI, the shorter the survival time of HCC patients (Figure 2F). Finally, we performed a kaplan-meier analysis, and the results showed that the OS of the high BMRGI group was significantly lower than that of the low BMRGI group in both the TCGA cohort and the ICGC cohort (Figure 2G).

3.3 Single-cell transcriptional profiling and cell-cell interactions in HCC tissue

We used tSNE to perform dimensionality reduction and clustering on the preprocessed scRNA-seq data, and finally obtained 12 clusters (Figure 3A). In addition, we also displayed the most significantly expressed genes in the 12 clusters using a heatmap (Figure 3B). Cell types were automatically annotated by the SingleR package, and these 12 clusters were clustered into 5 cell types, including Hepatocytes, T cells, Macrophage, B cell and NK cell (Figure 3C).

In addition, we further evaluated the interactions between different cells using the "CellChat" package. Figure 3D shows the number and weight of interactions among the five cell types. Furthermore, we present these results separately for a clearer picture of the strength of cell-cell interactions (Figure 3E). Overall, Hepatocytes rarely act as receptors for signals from the other four types of immune cells, but they can communicate with immune cells by emitting signals. Immune cells interact and

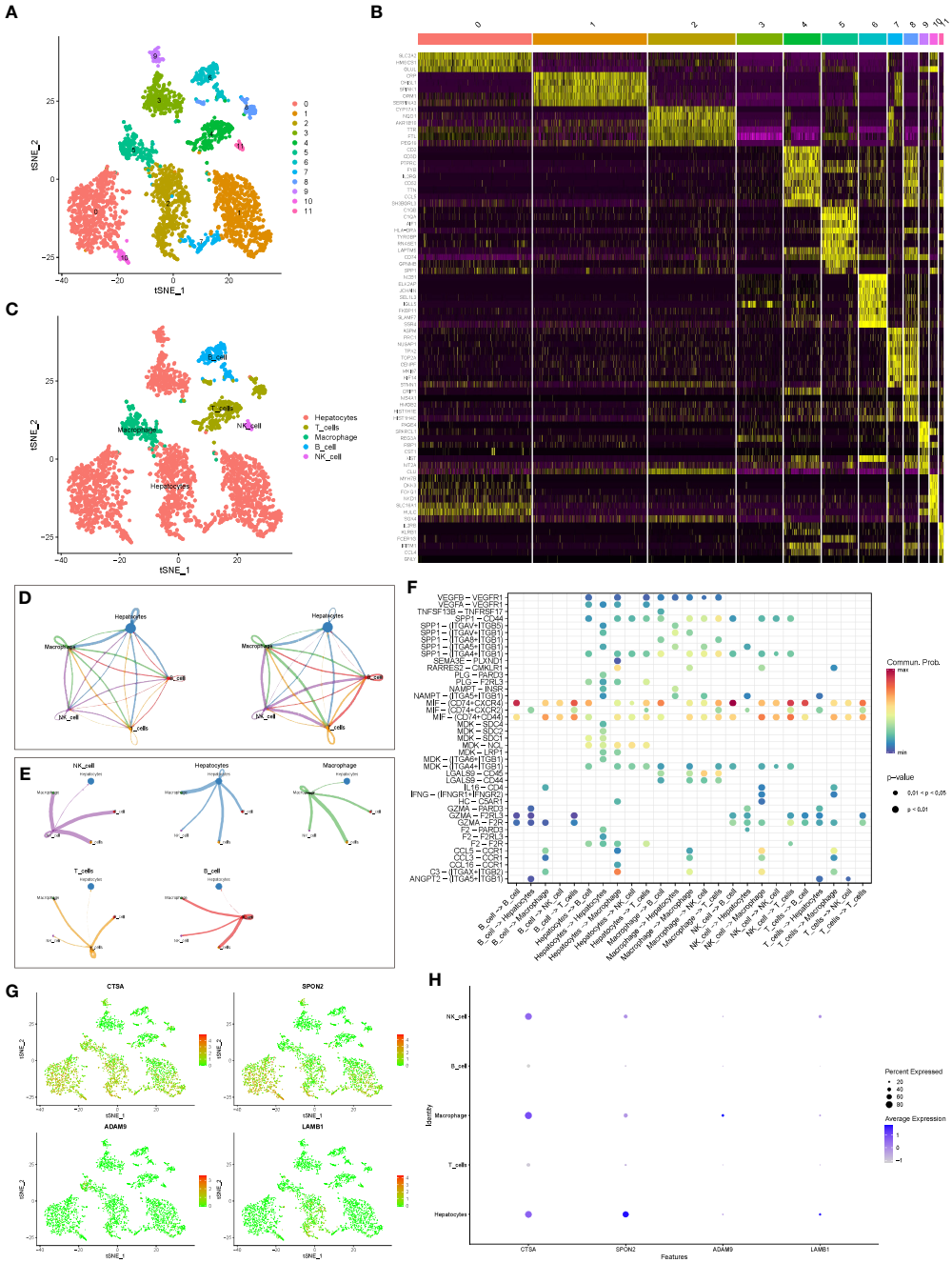


FIGURE 3
scRNA-seq analysis and CellChat Analysis **(A)** tSNE analysis to classify cell clusters. **(B)** Heatmap showing highly expressed genes in cell clusters. **(C)** The "SingleR" package annotates cell clusters into 5 cell types. **(D)** Network diagram of the number and weight of connections between different cell types. **(E)** Diagram of the communication network between cells and other cells. **(F)** Bubble diagram of receptor ligand molecules involved in cell communication. **(G, H)** Distribution and expression levels of model genes in 5 cell types.

receive signals frequently. In addition, receptor ligand molecules when mediating cell-cell interactions are shown in Figure 3F.

Finally, we explored the distribution of model genes in different cell types and showed the expression levels of the model in different cell types in bubble plots (Figures 3G, H). In brief, CTSA was highly expressed in hepatocytes, NK cells, and macrophages, SPON2 was highly expressed only in hepatocytes, whereas ADAM9 and LAMB1 were expressed at low levels in all cell types.

3.4 Comprehensive analysis of clinical parameters in HCC patients

As shown in Figure 4A, the heat map of BMRGI and common clinicopathological parameters, the results showed that the tumor stages of HCC patients with different BMRGI groups had statistical differences. In addition, we further determined the prognostic value of BMRGI in patients with different pathological features. The results showed that the high BMRGI group had significantly

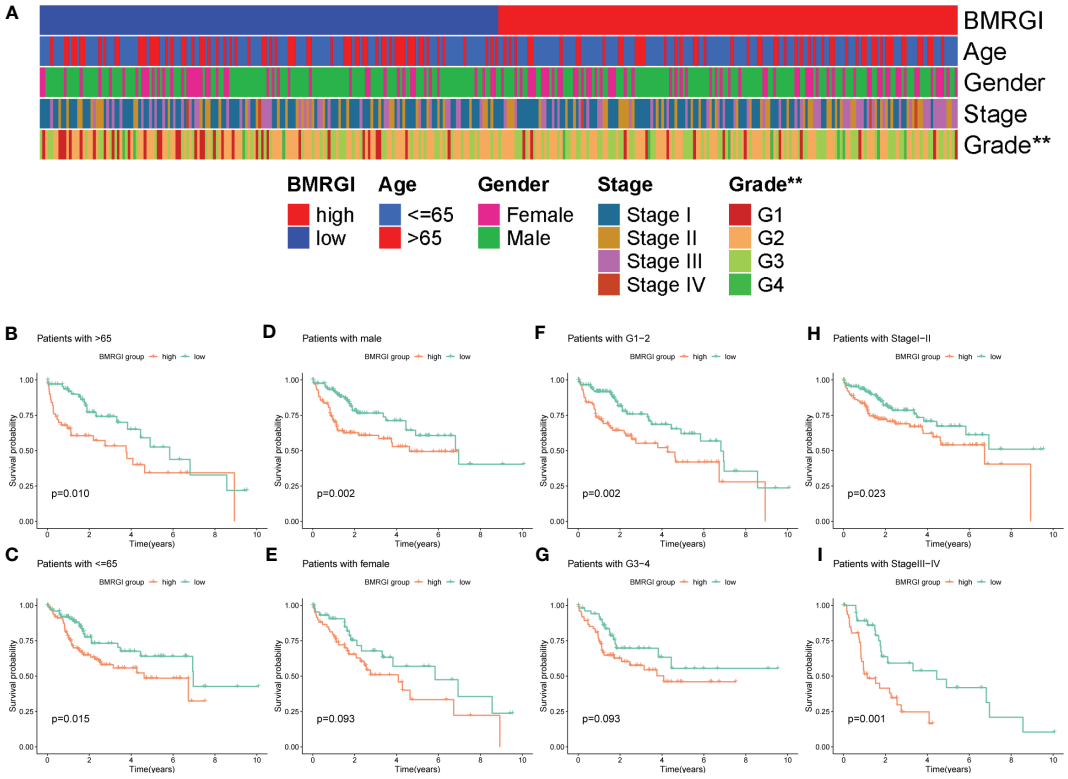


FIGURE 4
Correlation analysis between BMRGI and clinicopathological features. **(A)** Heatmap of the distribution of clinical case characteristics of patients in the high BMRGI group and low BMRGI group. **(B, C)** Kaplan-meier survival curves of high BMRGI group and low BMRGI group in different age groups. **(D, E)** Kaplan-meier survival curves of high BMRGI group and low BMRGI group in different gender groups. **(F, G)** Kaplan-meier survival curves of high BMRGI group and low BMRGI group under different grading groups. **(H, I)** Kaplan-meier survival curves of high BMRGI group and low BMRGI group in different stages.

worse OS than the low BMRGI group in HCC patients with different clinicopathological parameters (age, gender, tumor grade and stage) ($p < 0.1$, Figures 4B–I). These results suggest that our BMRGI can effectively predict the prognosis of HCC patients with different clinicopathological features.

3.5 Construction and evaluation of clinical nomogram based on BMRGI

In order to construct a more practical and stable nomogram, we incorporated several common clinicopathological parameters (age, gender, tumor grade and stage). Univariate Cox analysis showed that tumor stage and BMRGI were risk factors affecting the prognosis of HCC patients (Figure 5A). Multivariate Cox analysis confirmed that tumor stage and BMRGI were independent risk factors affecting the prognosis of HCC patients after adjusting for other clinicopathological parameters (Figure 5B). Given the high correlation between BMRGI and prognosis of HCC patients. We constructed a new nomogram combining common clinicopathological parameters and BMRGI (Figure 5C). We first evaluated the AUC value of various indicators to predict the prognosis of HCC patients using the ROC curve, and the results showed that the ability of BMRGI to predict the prognosis of HCC patients was significantly better than other clinicopathological features (including the classic indicator tumor stage), and the nomogram constructed based on this further improved the accuracy of predicting the prognosis of HCC patients (Figure 5D). In addition, the excellent accuracy and robustness of the nomogram in assessing the 1-year, 3-year, and 5-year survival of patients was further illustrated by ROC curves and calibration curves (Figures 5E, F).

3.6 GO, KEGG and GSEA analysis related to BMRGI

First, we analyzed the genetic differences between the high BMRGI group and the low BMRGI group ($|\log_2 FC| > 1$, FDR< 0.05). Based on these differential genes, we further performed GO analysis and KEGG analysis to explore their biological characteristics. GO analysis results showed that, in terms of biological process, DEGs were mainly enriched in " membrane invagination, phagocytosis, engulfment, phagocytosis, recognition, plasma membrane invagination, phagocytosis, humoral immune response mediated by circulating immunoglobulin, humoral immune response, B cell receptor signaling pathway, cell chemotaxis, leukocyte migration". In terms of cellular

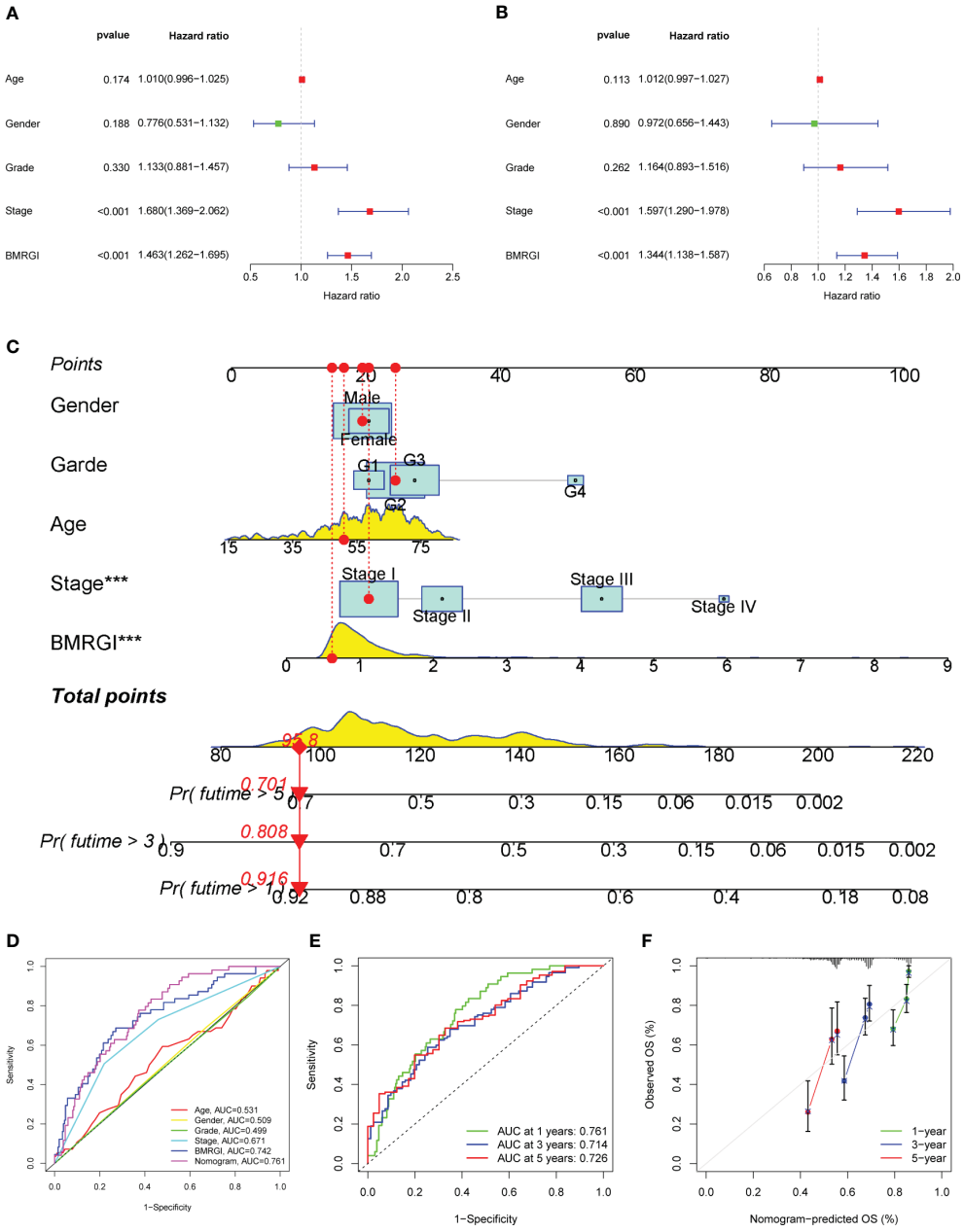


FIGURE 5
Construction of clinical nomogram. **(A, B)** Forest plots for univariate and multivariate Cox regression analysis. **(C)** Nomogram combining common clinical parameters and BMRGI. **(D)** ROC curves for clinical parameters, BMRGI and nomogram. **(E)** ROC curves of nomograms predicting 1-year, 3-year and 5-year survival rates. **(F)** Calibration curves for nomograms.

composition, DEGs were mainly enriched in "immunoglobulin complex, collagen−containing extracellular matrix, immunoglobulin complex, circulating, external side of plasma membrane, basal plasma membrane, Golgi lumen, basal part of cell, apical plasma membrane, basolateral plasma membrane, apical part of cell". and in terms of molecular functions, DEGs were mainly enriched in "antigen binding, immunoglobulin receptor binding, extracellular matrix structural constituent, collagen binding, glycosaminoglycan binding, fibronectin binding, sulfur compound binding, heparin binding, growth factor binding, insulin−like growth factor binding" (Figure 6A). The results of KEGG analysis showed that DEGs were only enriched in the Focal adhesion signaling pathway (Figure 6B).

In addition, GSEA analysis was further carried out in this study. The results showed that the signal pathways affected by the high BMRGI group were mainly enriched in " KEGG CELL ADHESION MOLECULES CAMS, KEGG CYTOKINE CYTOKINE RECEPTOR INTERACTION, KEGG ECM RECEPTOR INTERACTION, KEGG FOCAL ADHESION, KEGG NEUROACTIVE LIGAND RECEPTOR INTERACTION". The signal pathways affected by the low BMRGI group were mainly enriched in "KEGG DRUG METABOLISM CYTOCHROME P450,

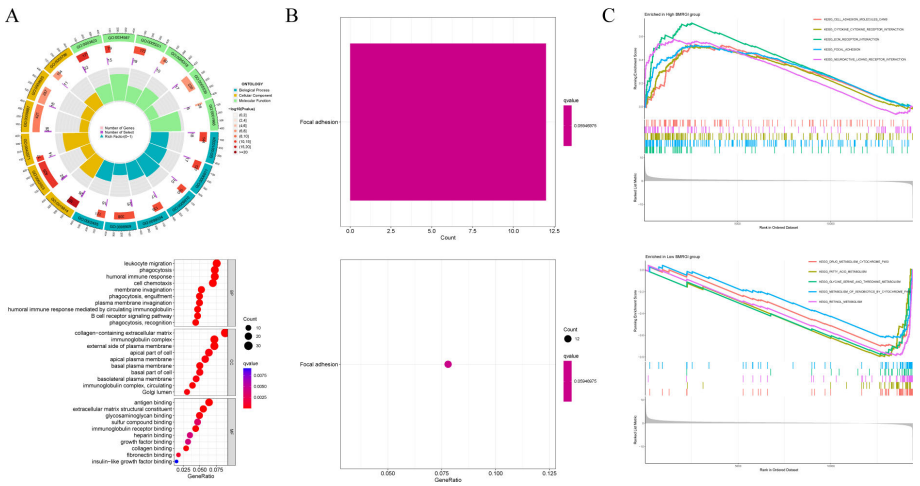


FIGURE 6
GO analysis, KEGG analysis and GSEA analysis. **(A)** Circle and bubble charts for GO analysis. **(B)** Barplot and bubble charts for KEGG analysis. **(C)** GSEA analysis of high BMRGI group and low BMRGI group.

KEGG FATTY ACID METABOLISM, KEGG GLYCINE SERINE AND THREONINE METABOLISM, KEGG METABOLISM OF XENOBIOTICS BY CYTOCHROME P450, KEGG RETINOL METABOLISM"(Figure 6C).

3.7 Comprehensive analysis of the correlation between BMRGI and tumor microenvironment

In view of the important guiding significance of immune checkpoint molecules and HLA family molecules in immunotherapy. We analyzed the correlation between BMRGI and 48 common immune checkpoint molecules and 24 HLA family molecules. The results showed that BMRGI was significantly positively correlated with 41 immune checkpoint molecules as well as 23 HLA family molecules (Figures 7A, B). In addition, we assessed the levels of 16 immune-related cells and 13 immune-related terms in tissue samples from HCC patients using ssGSEA. In terms of immune-related cells: Compared with the low BMRGI group, aDCs, DCs, iDCs, Macrophages, pDCs, Th1_cells, Th2_cells, and Treg were significantly increased in the high BMRGI group, while NK_cells were significantly decreased (Figure 7C). In terms of immune-related terms, compared with the low BMRGI group, the levels of APC_co_Stimulation, CCR, Check-point, HLA, MHC_class_I, and Parainflammation were significantly increased, while Type_II_IFN_Reponse was significantly decreased (Figure 7D). Then, we analyzed the somatic mutation profile of HCC patients and identified the top 10 mutated genes in the high and low BMRGI groups. The results showed that TP53 mutations were significantly lower in the high BMRGI group than in the BMRGI group (Figures 7E, F). However, there was no statistical difference in TMB between the high and low BMRGI groups (Figure 7G). Finally, we assessed the sensitivity to immunotherapy in the high and low BMRGI groups. The results showed that TIDE was lower in the high BMRGI group, indicating that the lower the possibility of immune escape, the better the effect of immunotherapy (Figure 7H).

3.8 Screening for sensitive drugs in HCC patients

We evaluated and observed the differences in sensitivity to 6 common drugs in HCC patients between the two groups. The lower the IC50 value, the higher the sensitivity to the drug. The results showed that patients in the low BMRGI group were more sensitive to sorafenib, oxaliplatin, cytarabine and fludarabine, whereas patients in the high BMRGI group were more sensitive to 5-fluorouracil and gefitinib higher (Figures 7I, J). All in all, these results provide a good reference for clinical medication.

3.9 Identification of key BMRG

We conducted a more refined analysis based on the 47 common genes screened above. First, we identified marker molecules of HCC by 3 machine learning methods (LASSO, SVM-REF, and RandomForest) (Figures 8A–C). Among them, CSTA, ITGA6, ITGB8 and LAMC1 are common marker molecules (Figure 8D). Then we evaluated the diagnostic value of the four marker molecules through the ROC curve, and the results showed that CSTA (AUC = 0.952), ITGA6(AUC = 0.942), ITGB8(AUC = 0.756) and LAMC1(AUC = 0.936) all had high diagnostic value (Figures 8E–H). At the same time, we found that CTSA not only has the highest diagnostic value, but also constitutes one of the members of BMRPI. Therefore, we considered CTSA as the most critical BMRG in HCC for further study.

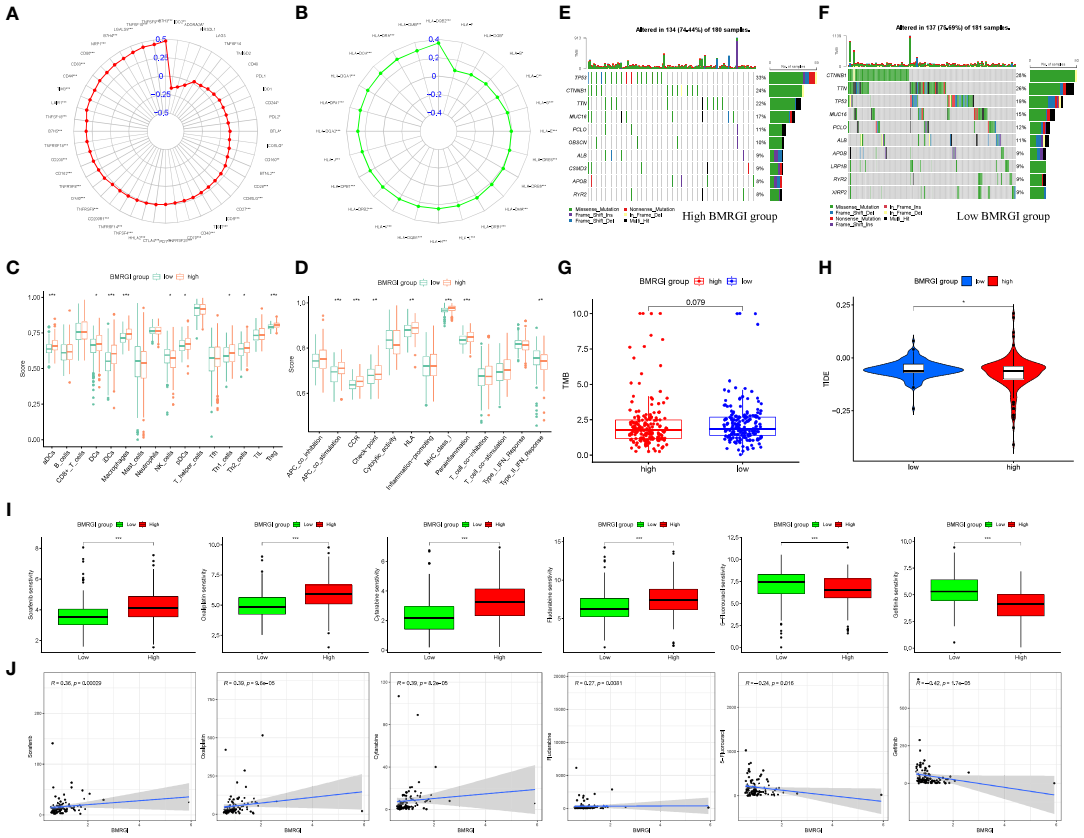


FIGURE 7
Correlation analysis between BMRGI and tumor microenvironment and common drug sensitivity. **(A)** Correlation between BMRGI and immune checkpoint molecules. **(B)** Correlation between BMRGI and HLA family molecules. **(C)** Differences in immune cell infiltration between high and low BMRGI groups. **(D)** Differences in immune-related terms between high and low BMRGI groups. **(E)** TOP10 mutated genes in high BMRGI group. **(F)** TOP10 mutated genes in low BMRGI group. **(G)** Difference analysis of TMB between high BMRGI group and low BMRGI group. **(H)** Difference analysis of TIDE scores between high BMRGI group and low BMRGI group. **(I)** Difference analysis of IC50 values of 6 commonly used drugs in high BMRGI group and low BMRGI group. **(J)** Correlation analysis between IC50 values of 6 commonly used drugs and BMRGI. * represents $p < 0.05$, ** represents $p < 0.01$, and *** represents $p < 0.001$.

3.10 Multilevel expression verification and *in vitro* functional exploration of CTSA

We further investigated the differential expression of CTSA in HCC. First, we searched through GEPIA2.0, and the results showed that CTSA was highly expressed in HCC (Figure 9A). Second, we explored the differential expression of CTSA at the protein level. The UALCAN database (https://ualcan.path.uab.edu/index.html) showed that the protein level of CTSA in HCC was significantly higher than that in the normal group (Figure 9B). Likewise, the HPA database showed that CTSA was highly expressed in HCC tissues compared with normal liver tissues (Figure 9C). Finally, we detected the expression of CTSA in normal liver cell lines (LO2) and liver cancer cell lines (BEL7402, HEPG2, HCCLM3), and the results showed that the expression level of CTSA in liver cancer cell lines was significantly higher than that in normal liver cell lines (Figure 9D).

To gain insight into the *in vitro* function of CTSA in HCC, we characterized the oncogenic phenotype of HCCLM3 and BEL-7402 cells (SiCTSA-1 and SiCTSA-2) with CTSA knockdown. The RT-qPCR results showed that siCTSA-1 and siCTSA-2 could significantly inhibit the CTSA expression of HCC cells (BEL-7402 and HCCLM3 cells) (Figure 9E). We studied the role of CTSA in HCC cell proliferation by CCK8 assay, and the role of CTSA in HCC cell migration and invasion using Transwell assay. CCK8 assay and Transwell assay analysis showed that the reduction of CTSA impaired the proliferation (Figures 9F, G), migration (Figure 9H) and invasion (Figure 9I) abilities of HCC cells (BEL7402 and HCCLM3).

4 Discussion

HCC is the most prevalent histological type of primary liver cancer, known for its high metastatic and recurrence characteristics (4). Unfortunately, most HCC patients are diagnosed at an advanced stage, which significantly reduces the chance of curative treatment and leads to a poor prognosis (26). BM, as an important component of the extracellular matrix, is an important barrier that cancer cells must overcome to form metastasis (5, 27). Numerous studies have demonstrated the association between the main components of BM and HCC tumor metastasis, as well as poor prognosis (28–30). Current study shows that systematic analysis of specific gene sets achieves promising results in predicting cancer

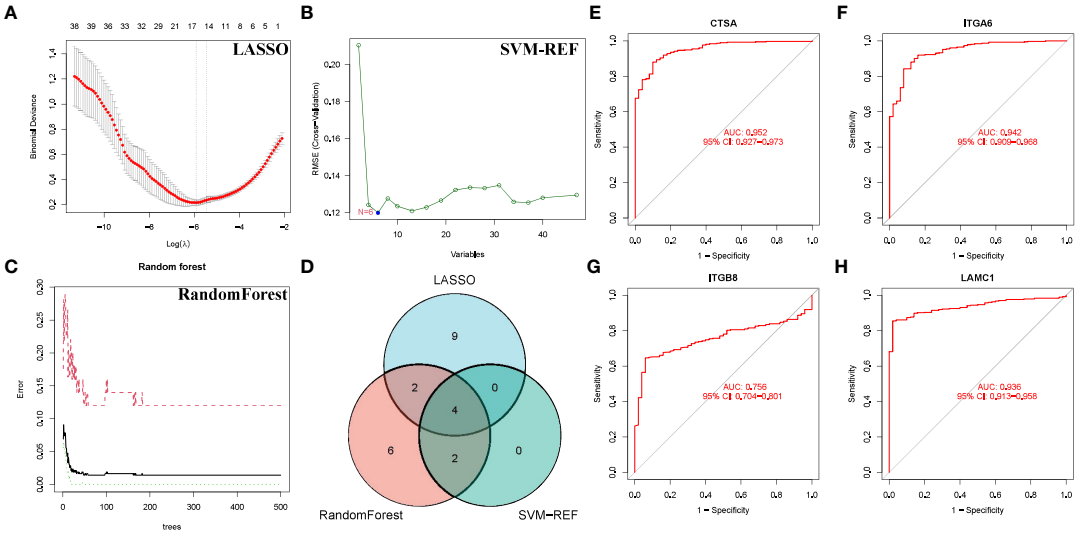


FIGURE 8
Screening for Feature BMRG. Characteristic genes in DEGs selected by LASSO **(A)**, SVM-SEF **(B)** and RandomForst **(C)**. **(D)** The Venn diagram shows the common genes of the three algorithms. **(E–H)** ROC curve of common genes.

prognosis (31, 32). Despite advancements in research, there is still a lack of reliable prognostic models for HCC based on basement membrane genes. To address this gap, our study utilized WGCNA and machine learning to develop a strong prognostic index based on BMRG. Our model has demonstrated high accuracy in predicting the prognosis of HCC patients. This study utilized the TCGA-HCC dataset to identify 4 BMRG (CTSA, ADAM9, LAMB1, and SPON2) through WGCNA and machine learning techniques. These BMRG were used to construct BMRGI. Previous research has shown that all four BMRG are closely associated with HCC. Wang et al. discovered that CTSA has potential as a diagnostic and prognostic marker for HCC patients (33). In HCC, ADAM9 is known to be overexpressed and is responsible for inducing ROS generation, which in turn promotes HCC cell invasion (34). Additionally, LAMB1 has been shown to be regulated by the RNA helicase DDX24, which contributes to the malignant progression of HCC (35). However, the role of SPON2 in HCC is still a matter of debate. While high expression of SPON2 has been linked to poor prognosis in HCC patients (36), it has also been found to inhibit tumor metastasis by promoting the infiltration of M1-like macrophages (37). Taken together, these studies showed that the four BMRG were closely related to HCC and its prognosis, which indicated the correctness of our BMRGI based on them.

By analyzing the single-cell atlas of HCC tissue, we identified five types of cells present. Our findings suggest that hepatocytes are capable of acting as ligands to send signals to immune cells, while immune cells exhibit a weaker ability to send signals to liver cells. We conducted an analysis of the expression of model genes across various cell types and found that CTSA expression was particularly high in hepatocytes, NK cells, and macrophages. This suggests that any abnormal expression of CTSA could potentially impact the progression of HCC by influencing the immune microenvironment of HCC.

We conducted related validation on BMRGI. Survival analysis revealed that the prognosis of the high BMRGI group was significantly worse than that of the lower BMRGI group in the TCGA-HCC cohort. Furthermore, we validated the ability of BMRGI to predict the prognosis of HCC patients in the ICGC cohort. The study found that BMRGI is an independent risk factor for the prognosis of HCC patients, as determined by the results of multivariate Cox analysis. Additionally, ROC curve analysis revealed that BMRGI is a better predictor of HCC patient prognosis compared to other clinicopathological parameters. Subgroup analysis based on clinical characteristics demonstrated that BMRGI has a strong ability to predict prognosis for HCC patients with varying clinical characteristics. In order to facilitate clinical application and improve the accuracy of predicting the prognosis of HCC patients, we combined common clinicopathological parameters with BMRGI to construct a nomogram.

We conducted further analysis of the differentially expressed genes (DEGs) between the high and low BMRGI groups to investigate the biological properties of these subgroups. Our analysis, which included Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Gene Set Enrichment Analysis (GSEA), revealed significant differences in biological processes related to immunity and cell adhesion (BM is closely related) between the different BMRGI subgroups.

This study demonstrates the accuracy of BMGPI construction (closely related to BM-related biological characteristics). Additionally, BMGPI effectively recognizes differences in the tumor immune microenvironment. Further analysis was conducted to determine the correlation between BMRGI and the tumor immune microenvironment. Previous research has shown that immune checkpoint molecules and HLA family molecules are strong predictors of response to immunotherapy (38–41). Therefore, we analyzed the correlation between BMRPI and immune checkpoint molecules and HLA family molecules. The results showed that BMRGI was positively correlated with most immune checkpoint molecules and HLA molecules, suggesting that BMRGI may also be a good biomarker

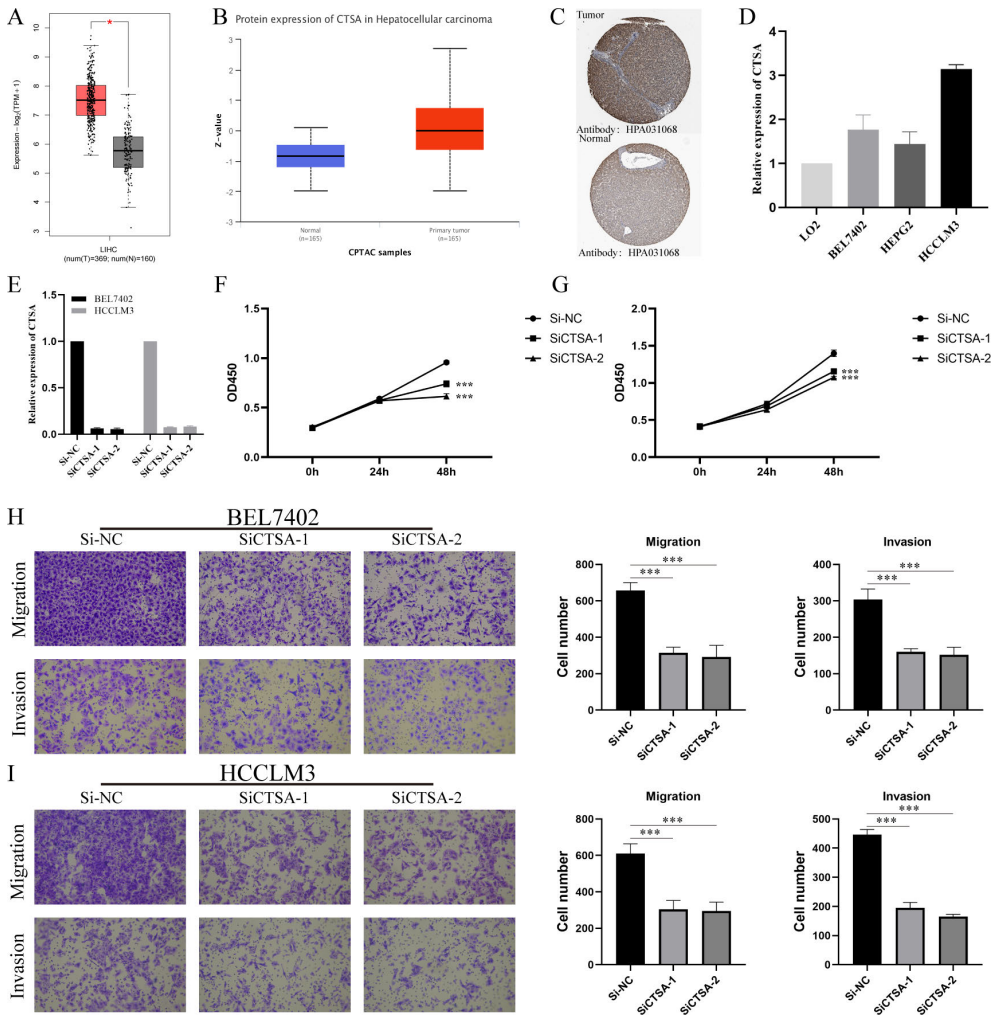


FIGURE 9
Multidimensional expression validation of CTSA and modulation of HCC oncogenic capacity *in vitro*. **(A)** Differential analysis of mRNA expression of CTSA in HCC in GEPIA2.0 database. **(B)** Differential analysis of protein expression levels of CTSA in HCC from UALCAN database. **(C)** IHC images of CTSA in HCC tissues and normal liver tissues from HPA database. **(D)** Expression levels of CTSA in normal liver cell lines and HCC cell lines. **(E)** Knockdown efficiency of siCTSA in BEL7402 and HCCLM3 cells. **(F, G)** CCK8 assay detects the effect of knocking down CTSA on the proliferation ability of BEL-7402 (Left) and HCCLM3 (Right). **(H)** Transwell assay was used to detect the effect of knocking down CTSA on the migration and invasion of BEL7402 cells. **(I)** Transwell assay was used to detect the effect of knocking down CTSA on the migration and invasion of HCCLM3 cells. *** represents $p < 0.001$.

for predicting immunotherapy response. Second, our results showed significant differences in terms of immune-related cells between high and low BMRGI groups, implying that different BMRGI subgroups may differ in response to immunotherapy. Interestingly, the mutation rate of TP53 in the high BMRGI group was significantly higher than that in the low BMRGI group, which may be one of the reasons for the poor prognosis in the high BMRGI group (42). However, overall TMB levels were not statistically different between the two groups. In this study, we used the TIDE algorithm to analyze the response to immunotherapy in various BMRGI subgroups of HCC patients. Our findings indicate that patients in the high BMRGI group had a lower TIDE score, suggesting that they may be less prone to immune escape and therefore have a better response to immunotherapy. Additionally, we evaluated the sensitivity of different BMRGI subgroups to six commonly used therapeutic drugs, providing valuable insights for clinical decision-making.

Finally, we screened 47 common gene by multiple machine learning methods and finally identified 4 BMRG: CTSA, ITGA6, ITGA8, and LAMC1. ROC analysis showed that these genes have high diagnostic value for distinguishing HCC. We found that CTSA not only had the highest diagnostic value (AUC:0.952), but also constituted one of the core members of BMRGI. Therefore, we regard CTSA as the most critical member of BMRG and conduct in-depth research. We verified that the expression of CTSA in HCC was significantly higher than that in normal tissues at the mRNA level and protein level by GEPIA2.0 database, UALCAN database and HPA database. In addition, we also verified by RT-qPCR that the expression of CTSA in HCC cell lines was significantly higher than that in normal liver cell lines. Since the oncogenic role of CTSA in HCC is still unclear, this prompted us to further explore the role of CTSA in HCC progression. More importantly, our *in vitro* cell experiments showed that the proliferation, migration and invasion

abilities of HCC cell lines (BEL7402 and HCCLM3) were significantly reduced after CTSA knockdown.

Like other studies, even this study has some limitations and shortcomings. First, when we validated the prognostic value of BMRGI, we did not validate it in real cohorts. Second, the carcinogenesis of CTSA has not been explored by *in vivo* experiments. Finally, the specific molecular mechanism by which CTSA affects the progression of HCC was not elucidated in this study.

In conclusion, our study trained and validated a BMRGI that could effectively predict the prognosis of HCC patients based on 222 BMRG. Based on this, we also developed a nomogram for clinical application. The biological and immunological characteristics of BMRGI in HCC were explored through a series of bioinformatics methods, and some insights were provided for clinical immunotherapy and targeted therapy. Finally, we also verified the role of the key BMRG CTSA in HCC progression through *in vivo* functional experiments.

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

PS and YY have constructed and devised the research. WS and JW performed data analysis and wrote the manuscript. ZW, MX and QL analyzed the data. WS performed the experiments. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by National Natural Science Foundation of China (82070634), Shanghai Municipal Science and Technology Commission (20ZR1451700), and Shanghai Collaborative Innovation Center for Translational Medicine (TM201731).

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.2023.1185916/full#supplementary-material

References

1. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers* (2021) 7(1):6. doi: 10.1038/s41572-020-00240-3

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(3):209–49. doi: 10.3322/caac.21660

3. Forner A, Llovet JM, Bruix J. Hepatocellular carcinoma. *Lancet (London England)* (2012) 379(9822):1245–55. doi: 10.1016/s0140-6736(11)61347-0

4. Yin Z, Dong C, Jiang K, Xu Z, Li R, Guo K, et al. Heterogeneity of cancer-associated fibroblasts and roles in the progression, prognosis, and therapy of hepatocellular carcinoma. *J Hematol Oncol* (2019) 12(1):101. doi: 10.1186/s13045-019-0782-x

5. Reuten R, Zendehroud S, Nicolau M, Fleischhauer L, Laitala A, Kiderlen S, et al. Basement membrane stiffness determines metastases formation. *Nat materials* (2021) 20(6):892–903. doi: 10.1038/s41563-020-00894-0

6. Donato MF, Colombo M, Matarazzo M, Paronetto F. Distribution of basement membrane components in human hepatocellular carcinoma. *Cancer* (1989) 63(2):272–9.

7. Mutgan AC, Jandl K, Kwapiszewska G. Endothelial basement membrane components and their products, matrikines: active drivers of pulmonary hypertension? *Cells* (2020) 9(9):2029. doi: 10.3390/cells9092029

8. Sherwood DR. Basement membrane remodeling guides cell migration and cell morphogenesis during development. *Curr Opin Cell Biol* (2021) 72:19–27. doi: 10.1016/j.ceb.2021.04.003

9. Banerjee S, Lo WC, Majumder P, Roy D, Ghorai M, Shaikh NK, et al. Multiple roles for basement membrane proteins in cancer progression and emt. *Eur J Cell Biol* (2022) 101(2):151220. doi: 10.1016/j.ejcb.2022.151220

10. Najafi M, Farhood B, Mortezaee K. Extracellular matrix (Ecm) stiffness and degradation as cancer drivers. *J Cell Biochem* (2019) 120(3):2782–90. doi: 10.1002/jcb.27681

11. Chen S, Liu P, Zhao L, Han P, Liu J, Yang H, et al. A novel cuproptosis-related prognostic lncrna signature for predicting immune and drug therapy response in hepatocellular carcinoma. *Front Immunol* (2022) 13:954653. doi: 10.3389/fimmu.2022.954653

12. Lu J, Yu C, Bao Q, Zhang X, Wang J. Identification and analysis of necroptosis-associated signatures for prognostic and immune microenvironment evaluation in hepatocellular carcinoma. *Front Immunol* (2022) 13:973649. doi: 10.3389/fimmu.2022.973649

13. Jayadev R, Morais M, Ellingford JM, Srinivasan S, Naylor RW, Lawless C, et al. A basement membrane discovery pipeline uncovers network complexity, regulators, and human disease associations. *Sci Adv* (2022) 8(20):eabn2265. doi: 10.1126/sciadv.abn2265

14. Langfelder P, Horvath S. Wgcna: an r package for weighted correlation network analysis. *BMC Bioinf* (2008) 9:559. doi: 10.1186/1471-2105-9-559

15. 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(5):411–20. doi: 10.1038/nbt.4096

16. 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(6):bbab260. doi: 10.1093/bib/bbab260

17. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. *Mach Learn* (2002) 46(1):389–422. doi: 10.1023/A:1012487302797

18. Duan J, Soussen C, Brie D, Idier J, Wan M, Wang YP. Generalized lasso with under-determined regularization matrices. *Signal Process* (2016) 127:239–46. doi: 10.1016/j.sigpro.2016.03.001

19. Ishwaran H, Kogalur UB. Consistency of random survival forests. *Stat probability Lett* (2010) 80(13-14):1056–64. doi: 10.1016/j.spl.2010.02.020

20. Zhao Z, He S, Yu X, Lai X, Tang S, Mariya ME, et al. Analysis and experimental validation of rheumatoid arthritis innate immunity gene Cyfip2 and pan-cancer. *Front Immunol* (2022) 13:954848. doi: 10.3389/fimmu.2022.954848

21. 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 (W1):W556–w60. doi: 10.1093/nar/gkz430

22. 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 (New York NY)* (2022) 25:18–27. doi: 10.1016/j.neo.2022.01.001

23. Colwill K, Gräslund S. A roadmap to generate renewable protein binders to the human proteome. *Nat Methods* (2011) 8(7):551–8. doi: 10.1038/nmeth.1607

24. Sun W, Shen J, Liu J, Han K, Liang L, Gao Y. Gene signature and prognostic value of ubiquitin-specific proteases members in hepatocellular carcinoma and explored the immunological role of Usp36. *Front bioscience (Landmark edition)* (2022) 27(6):190. doi: 10.31083/j.fbl2706190

25. Dai W, Xu L, Yu X, Zhang G, Guo H, Liu H, et al. Ogdhl silencing promotes hepatocellular carcinoma by reprogramming glutamine metabolism. *J Hepatol* (2020) 72(5):909–23. doi: 10.1016/j.jhep.2019.12.015

26. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet (London England)* (2018) 391(10127):1301–14. doi: 10.1016/s0140-6736(18)30010-2

27. Mierke CT. The matrix environmental and cell mechanical properties regulate cell migration and contribute to the invasive phenotype of cancer cells. *Rep Prog Phys Phys Soc (Great Britain)* (2019) 82(6):064602. doi: 10.1088/1361-6633/ab1628

28. Yamashita T, Koshikawa N, Shimakami T, Terashima T, Nakagawa M, Nio K, et al. Serum laminin Γ2 monomer as a diagnostic and predictive biomarker for hepatocellular carcinoma. *Hepatol (Baltimore Md)* (2021) 74(2):760–75. doi: 10.1002/hep.31758

29. Wang T, Jin H, Hu J, Li X, Ruan H, Xu H, et al. Col4a1 promotes the growth and metastasis of hepatocellular carcinoma cells by activating fak-src signaling. *J Exp Clin Cancer Res CR* (2020) 39(1):148. doi: 10.1186/s13046-020-01650-7

30. Krishnan MS, Rajan Kd A, Park J, Arjunan V, Garcia Marques FJ, Bermudez A, et al. Genomic analysis of vascular invasion in hcc reveals molecular drivers and predictive biomarkers. *Hepatol (Baltimore Md)* (2021) 73(6):2342–60. doi: 10.1002/hep.31614

31. Sun W, Xu Y, Zhao B, Zhao M, Chen J, Chu Y, et al. The prognostic value and immunological role of angiogenesis-related patterns in colon adenocarcinoma. *Front Oncol* (2022) 12:1003440. doi: 10.3389/fonc.2022.1003440

32. Ye Y, Zhao Q, Wu Y, Wang G, Huang Y, Sun W, et al. Construction of a cancer-associated fibroblasts-related long non-coding rna signature to predict prognosis and immune landscape in pancreatic adenocarcinoma. *Front Genet* (2022) 13:989719. doi: 10.3389/fgene.2022.989719

33. Wang H, Xu F, Yang F, Lv L, Jiang Y. Prognostic significance and oncogene function of cathepsin a in hepatocellular carcinoma. *Sci Rep* (2021) 11(1):14611. doi: 10.1038/s41598-021-93998-9

34. Dong Y, Wu Z, He M, Chen Y, Chen Y, Shen X, et al. Adam9 mediates the interleukin-6-Induced epithelial-mesenchymal transition and metastasis through ros production in hepatoma cells. *Cancer Lett* (2018) 421:1–14. doi: 10.1016/j.canlet.2018.02.010

35. Liu T, Gan H, He S, Deng J, Hu X, Li L, et al. Rna helicase Ddx24 stabilizes Lamb1 to promote hepatocellular carcinoma progression. *Cancer Res* (2022) 82 (17):3074–87. doi: 10.1158/0008-5472.can-21-3748

36. Feng Y, Hu Y, Mao Q, Guo Y, Liu Y, Xue W, et al. Upregulation of spondin-2 protein expression correlates with poor prognosis in hepatocellular carcinoma. *J Int Med Res* (2019) 47(2):569–79. doi: 10.1177/0300060518803232

37. Zhang YL, Li Q, Yang XM, Fang F, Li J, Wang YH, et al. Spon2 promotes M1-like macrophage recruitment and inhibits hepatocellular carcinoma metastasis by distinct integrin-rho gtpase-hippo pathways. *Cancer Res* (2018) 78(9):2305–17. doi: 10.1158/0008-5472.can-17-2867

38. Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *Lancet Oncol* (2016) 17(12):e542–e51. doi: 10.1016/s1470-2045(16)30406-5

39. Darvin P, Toor SM, Sasidharan Nair V, Elkord E. Immune checkpoint inhibitors: recent progress and potential biomarkers. *Exp Mol Med* (2018) 50(12):1–11. doi: 10.1038/s12276-018-0191-1

40. Schaafsma E, Fugle CM, Wang X, Cheng C. Pan-cancer association of hla gene expression with cancer prognosis and immunotherapy efficacy. *Br J Cancer* (2021) 125 (3):422–32. doi: 10.1038/s41416-021-01400-2

41. Loustau M, Anna F, Dréan R, Lecomte M, Langlade-Demoyen P, Caumartin J. Hla-G neo-expression on tumors. *Front Immunol* (2020) 11:1685. doi: 10.3389/fimmu.2020.01685

42. Long J, Wang A, Bai Y, Lin J, Yang X, Wang D, et al. Development and validation of a Tp53-associated immune prognostic model for hepatocellular carcinoma. *EBioMedicine* (2019) 42:363–74. doi: 10.1016/j.ebiom.2019.03.022

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 23 May 2023
DOI 10.3389/fimmu.2023.1156038

OPEN ACCESS

EDITED BY
Nan Zhang,
Harbin Medical University, China

REVIEWED BY
Xinyu Zhang,
The Second Affiliated Hospital of Harbin Medical University, China
Yuzhen Gao,
Zhejiang University, China

*CORRESPONDENCE
Desheng Lu
delu@szu.edu.cn
Ou Sha
shaou@szu.edu.cn

RECEIVED 01 February 2023
ACCEPTED 09 May 2023
PUBLISHED 23 May 2023

CITATION
Gao F, Wang X, Qin N, Zhang M, Liao M, Zeng M, Lu D and Sha O (2023) The analysis of cathepsin L that mediates cellular SARS-CoV-2 infection leading to COVID-19 in head and neck squamous cell carcinoma.
Front. Immunol. 14:1156038.
doi: 10.3389/fimmu.2023.1156038

The analysis of cathepsin L that mediates cellular SARS-CoV-2 infection leading to COVID-19 in head and neck squamous cell carcinoma

Feng Gao[1], Xia Wang[2], Nianhong Qin[2], Mingxia Zhang[3], Mingfeng Liao[3], Meiqi Zeng[1], Desheng Lu[2*] and Ou Sha[1*]

[1]School of Dentistry, Institute of Stomatological Research, Medical School, Shenzhen University, Shenzhen, China, [2]Medical School, Shenzhen University, Shenzhen, China, [3]Institute for Hepatology, National Clinical Research Center for Infectious Disease, Shenzhen Third People's Hospital, Shenzhen, China

The vulnerability of the oral cavity to SARS-CoV-2 infection is well-known, and cancer patients are at a higher risk of COVID-19, emphasizing the need to prioritize this patient population. Head and neck squamous cell carcinoma (HNSCC) is one of the most common malignant cancers associated with early metastasis and poor prognosis. It has been established that cancerous tissues express Cathepsin L (CTSL), a proteinase that regulates cancer progression and SARS-CoV-2 entry. Therefore, it is essential to evaluate the correlation between disease outcomes and CTSL expression in cancer tissues and predict the susceptibility of cancer patients to SARS-CoV-2. In this study, we used transcriptomic and genomic data to profile CTSL expression in HNSCC and developed a CTSL signature that could reflect the response of HNSCC patients to chemotherapy and immunotherapy. Additionally, we investigated the relationship between CTSL expression and immune cell infiltration and established CTSL as a potential carcinogenic factor for HNSCC patients. These findings could aid in understanding the mechanisms underlying the increased susceptibility of HNSCC patients to SARS-CoV-2 and contribute to the development of therapy for both HNSCC and COVID-19.

KEYWORDS

cathepsin L, head and neck squamous cell carcinoma (HNSCC), immunotherapy, CTSL, SARS-CoV-2

1 Introduction

The emergence of the novel coronavirus (2019-Nov) in December 2019 has resulted in a global spread, leading to large-scale pandemics and posing a significant threat to public health and the economy (1). COVID-19 has become a widespread and deadly disease, with a rapid increase in infections and fatalities reported worldwide (1). Although most people recover, there is concern that some may experience devastating consequences, as the long-term effects of the virus are still uncertain. To effectively control SARS-CoV-2, it is necessary to prevent infection and conduct further research to determine the factors that influence susceptibility to COVID-19 and the mechanisms behind these factors. SARS-spike CoV-2's glycoprotein has been shown to bind the angiotensin-converting enzyme 2(ACE2), speeding up the virus's entrance into host cells (2, 3). After binding to the target cells, TMPRSS2 and host cell proteases such as cathepsin L (CTSL) split the S protein into two subunits: S1 and S2, facilitating viral entry into the host cells by promoting membrane fusion and endocytosis of coronaviruses. S protein cleavage by host proteases is required for viral activation and subsequent infection (3, 4). As a result, differences in susceptibility to SARS-CoV-2 infection may be explained by CTSL expression and distribution to a certain extent.

Current evidence suggests that SARS-CoV-2 infection causes considerable morbidity and mortality. Some studies revealed that the oral cavity is one of the areas most vulnerable to SARS-CoV-2 (1), with oral manifestations possibly presenting before other COVID-19 manifestations. According to recent studies, cancer patients reportedly experience a higher incidence of COVID-19, more severe symptoms, and a poor prognosis (5). Accordingly, more emphasis should be placed on preventing COVID-19 infection in cancer patients.

Tumor pathogenesis is a complex process involving multiple pathophysiological processes and is impacted by numerous factors, including the body's immune status. The role of cellular and humoral immunity in viral infection protection is well established. The pathogenesis of COVID-19 is widely believed to be linked to immune response dysregulation, especially in T cells (6). Merad et al. reported that COVID-19 pathogenesis was mediated by aberrant and numerous immune cells, such as monocytes and macrophages (7). Besides, anticancer therapy, such as chemotherapeutics or radiation, can cause systemic immunosuppression (5). It is widely thought that cancer patients are more susceptible to SARS-CoV-2 infection due to immune dysregulation (8, 9). The expression of SARS-CoV2 receptors (ACE2, TMPRSS2, and CTSL) has been reported to be upregulated in many types of malignancies, making viral entry into cells easier and cancer patients more susceptible to SARS-CoV-2 (10).

Herein, we aimed to better understand the relationship between SARS-CoV-2 and head and neck squamous cell carcinoma (HNSCC). We utilized transcriptomic and genomic data to conduct a systematic analysis of the role of CTSL in HNSCC, focusing on immunological characteristics, functional annotation, and prediction of chemotherapy response. Furthermore, we developed a novel CTSL-related signature that could effectively predict HNSCC patients' outcomes. The current study substantiated that CTSL is a potentially carcinogenic factor for patients with HNSCC, which helps researchers better understand the increased susceptibility of HNSCC to SARS-CoV-2 infection and lays the groundwork for SARS-CoV-2 therapy.

2 Materials and methods

2.1 HNSC dataset and preprocessing

We acquired genomic data and annotated clinicopathological features of HNSC from The Cancer Genome Atlas (TCGA, 564 HNSC patients, including 520 tumors and 44 paracancers, and 519 tumor patients with complete survival information) and Gene Expression Omnibus (GEO) (datasets GSE41613, N=97 and GSE65858, N=270). The study included only cases with adequate OS information, while cases with insufficient information were excluded.

Raw data from the GEO database were generated using Affymetrix and Illumina platforms. The data were background corrected and normalized using a robust multichip averaging algorithm (RMA). RNA sequencing data were obtained from the TCGA database and converted from fragments to transcripts with signal intensities, similar to the data obtained from RMA.

2.2 Establishment of the CTSL-based Signature

High and low-expression groups were classified according to CTSL expression in TCGA-HNSC and underwent genome-wide difference analysis using the package R "limma" (abs(logFC) > 1 & P.Value< 0.01). Subsequently, Cox regression analysis was performed to further identify prognosis-related CTSL-associated genes (R language "survival" package, P<0.05). We then applied a survival machine learning algorithm *via* the R package "randomSurvivalForest" to screen significant valuable CTSL-associated genes with prognostic potential. Based on the Lasso regression analysis, the CTSL-based signature was constructed using the list of prognosis-related CTSL genes (11).

2.3 Validation of the efficacy of CTSL-based signature

We applied the CTSL-based signature to the data of 516 patients from the TCGA-HNSC dataset and then divided the patients into high and low CTSL. The clustering of signatures was based on the P value of the optimal cutoff, and the relationship between the CTSL-based signature and OS was analyzed using Kaplan-Meier curves. The performance of the CTSL-based signature for predicting prognosis at 1, 3, and 5 years were evaluated using TimeROC.

2.4 Genomic alteration

The TCGA datasets were used to collect somatic mutations. Somatic mutation analysis was achieved with the R package "maftools".

2.5 TME immunological profile assessment

The study utilized the Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data (ESTIMATE) technique to evaluate the abundance of immune cells and the level of stromal cell infiltration. This technique generates immune scores, stromal scores, and estimated scores to represent these factors. To analyze immune infiltrating cells in HNSC, we used the Tumor Immune Estimation Resource 2.0 (TIMER2.0; http://timer.cistrome.org) web server. The R genomic variance analysis package (GSVA) was used in conjunction with single sample genomic enrichment analysis to construct an enrichment score that represented the degree of infiltration of 28 immune cells based on the associated characteristics (ssGSEA).

2.6 Functional annotation

All gene sets were obtained from the MSigDB database, the Kyoto Encyclopedia of Genes and Genomes (KEGG), and Gene Ontology (GO). Gene set enrichment analysis (GSEA) and genomic variation analysis (GSVA) were implemented using the R packages clusterProfiler and GSVA.

2.7 Drug response prediction

To predict drug susceptibility in the cases included, the researchers used pharmacogenomic data from the Genomics of Drug Sensitivity in Cancer (GDSC) database (https://www.cancerrxgene.org/). They calculated drug response as drug susceptibility using the oncoPredict R package. The responses to anti-PD1 and anti-CTLA4 therapies in HNSC were evaluated by the submap algorithm.

2.8 CCK8 assay

To evaluate cell proliferation, the transfected cells were seeded into 96-well plates. After being cultured at 37°C with 5% CO2 for 24, 48, and 72 hours, 20 L of CCK8 solution (Sigma, USA) was added to each well, and the plates were incubated for an additional hour. Using a microplate reader, the optical density (OD) value obtained at 560nm wavelength was determined (Molecular Devices, Sunnyvale, CA).

2.9 Plate clone formation assay

To validate the ability of the transfected cells to form colonies, 1000 cells per plate were seeded in 60mm culture dishes. After being cultured for 10 days, the cells were fixed with 10% neutral buffer formalin fixative and stained with crystal violet (Beyotime, China). We photographed and counted the colonies.

2.10 Transwell assay

Transfected cells were seeded onto transwell membranes coated with Matrigel (BD Bioscience, Franklin Lakes, NJ) with a solution containing 10% bovine serum albumin (BSA, VWR, Radnor, PA) in the upper chamber (8μm pore size; Corning, Corning, NJ). The lower chamber contained medium with 10% FBS. After incubation at 37°C for 48 hours under 5% CO2, the remaining cells on the top membrane were removed using a cotton swab and fixed in 10% neutral buffer formalin. Crystal violet solution was added before capturing the membranes on camera with an inverted microscope and counting the cells.

2.11 Scratch migration assay

6 well plates were inoculated with transfected cells to determine cell migration. After 24 h of cell culture, a straight line wound was created in each well using a pipette tip. The cells were cultured in 2% FBS medium under 5% CO_2 at 37°C. Wound closure was assessed under 4× magnification with microscopy (OLYMPUS, Tokyo, Japan) at 24 and 48 h.

2.12 Construction of lentiviral vectors and infection of lentivirus infection

Using green fluorescent protein (GFP) containing lentiviral, high transfection efficiency and stable CTSL expression were achieved in SCC15 cells. A recombined EX-A4513-Lv201 vector with the CTSL gene and EX-NEG-Lv201 with a negative control sequence was constructed by GeneCopoeia Company (Guangzhou, China). The aforementioned lentiviral vectors were then used to infect SCC15 cells. A total of 1×10^6 SCC15 cells were seeded in a six-well cell plate, cultured for an additional 12 hours until 70% confluence was reached, and then lentiviral vectors at a multiplicity of infection (MOI) of 20 units per cell were added to the infection medium. Three groups were established as follows: SCC15 CTSL (CTSL overexpression group), SCC15 NEG (negative group), and SCC15 (non-treatment control group). After incubation for 24 hours, fresh, virus-free media was applied to the plates. After three days, the lentivirus density containing GFP was detected to evaluate the infection efficiency.

2.13 Statistical analysis

The Wilcoxon test was utilized for group comparisons when the data were not normally distributed, while the T-test was used when the variables were normally distributed. The difference in OS between the two groups was calculated using Kaplan-Meier survival plots with the R package "survminer". Cox regression for survival analysis was performed using the R package "survival". Time-dependent ROC curves were plotted using the R package "timeROC". Heat maps were

generated using the R package "pheatmap". Data visualization was done using the R package ggplot2 (v). Statistical significance was set at a p-value of 0.05.

3 Results

3.1 Features associated with CTSL expression in HNSCC

Analysis of the TCGA database revealed that the levels of CTSL mRNA in HNSC were significantly increased compared with adjacent tissues (Figure 1A).

The Kaplan-Meier method was used to assess the impact of CTSL expression on the survival of HNSC patients in three datasets (GSE41613, GSE65858, and TCGA). Patients with HNSC who exhibited high levels of CTSL expression were associated with low overall survival (OS) (Figures 1B-D). High levels of CTSL transcription have been identified as a significant risk factor for mortality in various cancer types, including HNSC, indicating a poor prognosis. Therefore, CTSL has huge potential as a robust prognostic marker for OS, even after accounting for other relevant variables. Next, GO and KEGG analyses were used to predict CTSL function and associated signaling pathways. GSEA was used to identify the signaling pathways linked with CTSL activated in HNSC.

CTSL was significantly enriched in signaling pathways, including immune response-associated activities such as acute inflammatory response, cytokine-mediated signaling pathway, dendritic cell differentiation, innate immune response, interferon-gamma production, leukocyte mediated cytotoxicity, regulation of B cell differentiation (Figure 1E). The GSEA results also showed that several immune functioning gene sets, including apoptosis, B cell receptor signaling, cytokine receptor interaction, and JAK-STAT signaling pathway, were enriched in HNSC. These findings suggest that CTSL plays an essential role in the tumor microenvironment.

3.2 CTSL is associated with immune infiltration in HNSCC patients in the TCGA cohort

Using the ESTIMATE algorithm, we examined how the immune status of HNSCC patients in the TCGA cohort is associated with the expression levels of CTSL. We discovered that patients with high CTSL expression levels were significantly more likely to have high immune scores, ESTIMATE scores, and stromal cells than patients with low CTSL expression levels (Figure 2A), suggesting that CTSL expression levels in HNSCC patients are correlated with immune status.

We independently applied the ssGSEA and TIMER algorithms (Figures 2B, C) to uncover the number of immune infiltrating cell groups to validate this characteristic. The generated heat map showed significant infiltration of NK cells, neutrophils, CD8+ T cells, and cytotoxic lymphocytes in HNSCC patients with elevated CTSL expression (Figure 2B). We discovered that groups with high CTSL expression had a larger abundance of infiltrating B cells and CD4+ T cells in HNSC (Figure 2C). Our findings suggest a significant correlation between the immune response to tumors and the expression levels of CTSL in HNSCC patients. We utilized a heatmap of clinical stages, grade, gender, and age to illustrate the relationship between CTSL expression and various clinical traits of HNSCC samples (Figures 2B, C). The findings revealed a significant association between CTSL expression and the clinical traits of HNSCC patients.

3.3 Potential immunotherapy and chemotherapy responses associated with CTSL expression in patients with HNSCC

Recent immunotherapy advancements, particularly PD-1 inhibitors, have improved treatment outcomes for HNSCC in the recurrent and metastatic stages. This improvement is caused by the interaction of immune cell processes (12). We first examined the relationship between CTSL expression and immune checkpoint levels in patients with HNSCC to investigate the therapy responsiveness depending on CTSL expression. Association analysis revealed that CD274 and CTLA4 levels were generally higher in HNSCC patients with elevated CTSL expression (Figure 3A). In addition, there was a significant association between CTSL expression and targeted therapies (including 5-fluorouracil, dasatinib, ERK_2440, JAK1_8709, luminespib, and staurosporine), indicating that patients with low CTSL expression responded better to targeted therapies (Figure 3B).

The results of the submap showed that the high- and low-CTSL groups had different responses to immunotherapy in that the high-CTSL group had a significant response to anti-PD-1 immunotherapy in HNSC based on the TCGA (Supplementary Figure).

3.4 The CTSL gene serves as an oncogene in HNSC cells

We sought to investigate the oncogenic potential of CTSL in HNSC by introducing it into SCC15 cells *via* lentiviral vectors containing GFP. This approach was chosen to achieve high transfection efficiency and ensure stable expression of CTSL. After three days of infection, we used fluorescence microscopy to confirm GFP expression (Figure 4A). Our functional assays revealed that CTSL overexpression promoted cell proliferation (CCK8, Figure 4B), invasion (Transwell, Figure 4C), colony formation (Plate cloning, Figure 4D), and migration (scratch migration, Figure 4E) in SCC15 cells. These findings collectively support the involvement of CTSL in the progression of HNSC.

3.5 The process of constructing CTSL.signature in HNSCC

After examining the differences between the high and low CTSL groups (Figure 5A), we performed a univariate Cox regression

10.3389/fimmu.2023.1156038

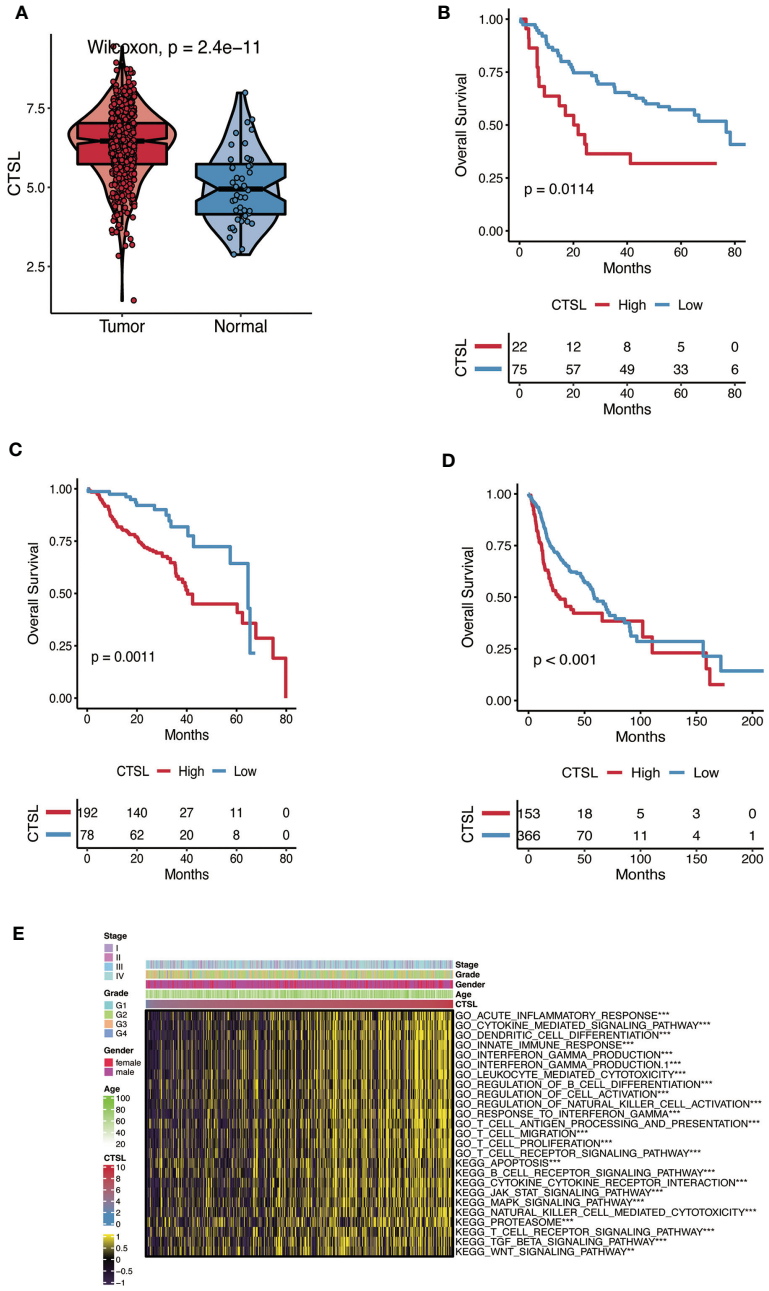


FIGURE 1
The CTSL expression patterns in HNSC. **(A)** The expression levels of CTSL in HNSC tumor tissues are higher than in normal tissues analyzed in the TCGA dataset (P< 0.01). **(B, C, D)** The association of CTSL expression with overall survival (OS) in HNSC patients (P< 0.01) in 3 databases (GSE41613, GSE65858 and TCGA). **(E)** Functional annotation associated with CTSL in TCGA-HNSC by GSVA analysis. ***, p<0.005, **, p<0.01.

analysis to further screen out more valuable CTSL-associated prognostic genes (P value< 0.05). 21 additional CTSL-associated genes with potential prognostic value for HNSCC patients were identified. A forest plot was generated to visualize the HR of each gene on prognosis (Figure 5B). In addition, we further selected 9 CTSL-associated genes by applying the machine learning algorithm of (Figure 5C). The random survival forest model was used to screen out six genes prognostically associated with CTSL (Figure 5C). Finally, Lasso regression analysis was used to calculate new scores based on the estimated regression coefficients of these five prognostic and CTSL-associated genes (Figure 5D).

The prognostic CTSL-based signature was as follows: $-0.6402 * IGLC1 + 0.142 * MT1E - 0.6298 * FDCSP - 0.6116 * EPHX3 + 0.4029 * SERPINE1$.

3.6 Validating the predictive value of CTSL-based signature for HNSCC survival

Kaplan-Meier analysis was carried out to determine whether the CTSL-based signature accurately predicted the clinical traits of HNSCC patients. Higher CTSL-based signature scores for HNSCC

10.3389/fimmu.2023.1156038

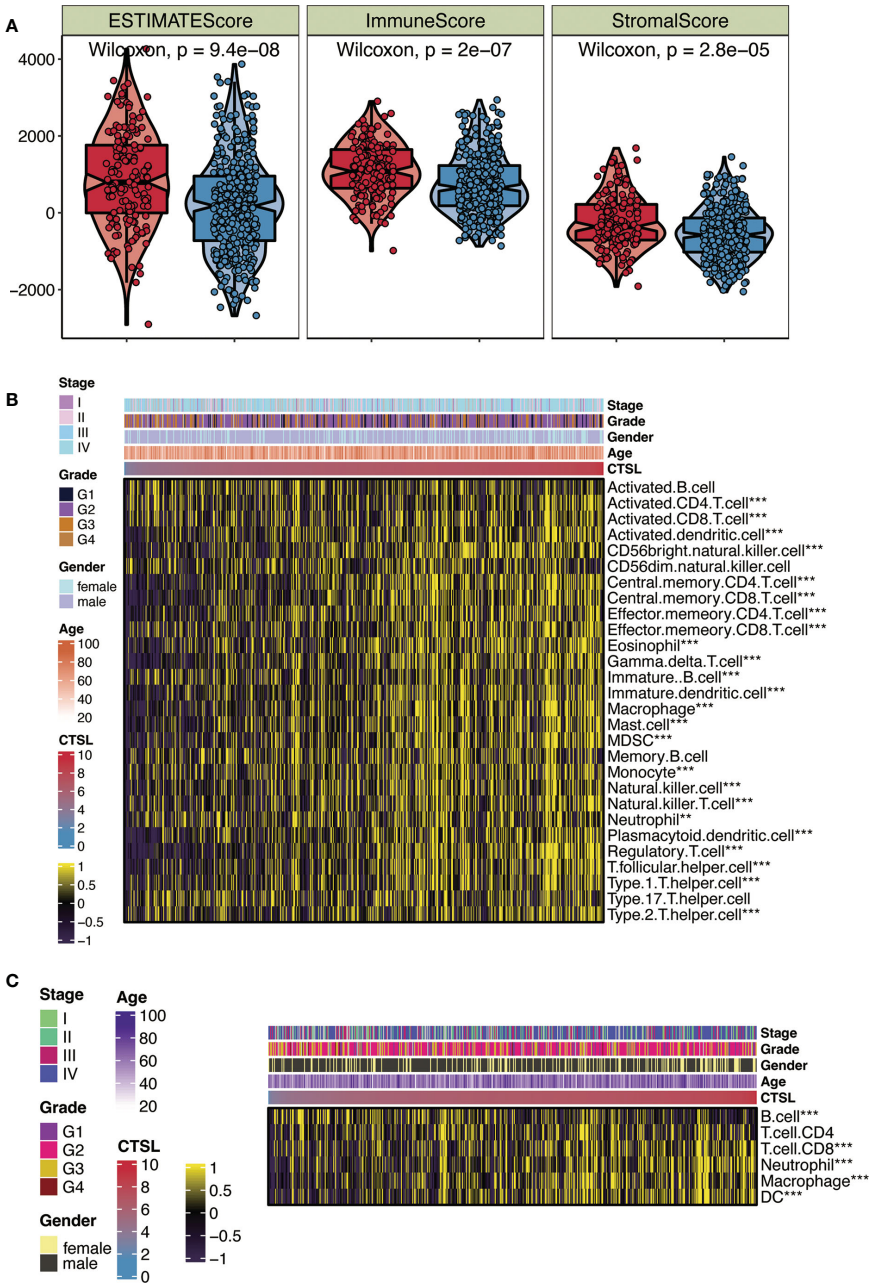


FIGURE 2
CTSL expression in relation to immunity in the TCGA cohort. **(A)** Changes in ESTIMATE among HNSCC patients with high and low CTSL expression. **(B)** Heatmap showing the abundance of infiltrating immune cell populations for different CTSL expression according tossGSEA algorithms. **(C)** Heatmap showing the abundance of infiltrating immune cell populations for different CTSL expression according to TIMER algorithms. **, p<0.005; ***, p<0.001.

patients were associated with worse survival curves (Figure 6A). The AUC values for the time-dependent ROC curves of the 1-year, 3-year, and 5-year OS were 0.685, 0.712, and 0.746, respectively (Figure 6B). This finding suggests that our CTSL signature has prognostic significance. In addition, we further validated the prognostic and survival disadvantage of HNSCC patients with higher CTSL-based signature scores using three independent cohorts (GSE41613, GSE65858, and TCGA) (Figure 6C).

3.7 CTSL signature expression was associated with genomic alterations

Cellular tumor antigen p53 (TP53) mutation was found to be significantly enriched in both the high-CTSL group and low-CTSL group (73% and 49%, respectively) according to the mutational distribution study (Figure 7B), followed by titin (TTN) (37%), (FAT1) (25%), and (CDKN2A) (23%) in the high-CTSL group,

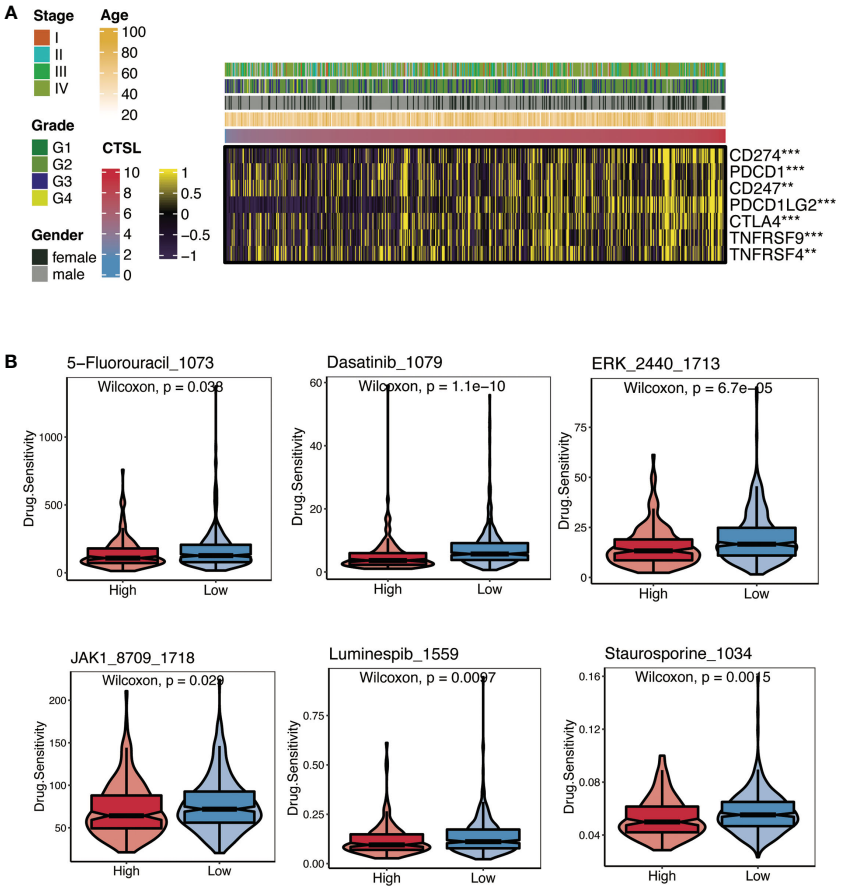


FIGURE 3
Immunotherapy and chemotherapy of CTSL expression involved in TCGA-HNSCC. **(A)** Correlation of CTSL expression and immune checkpoint levels in HNSCC. **(B)** Boxplots of estimated drug sensitivities for several GDSC chemotherapeutics in the high and low CTSL expression groups. **, p<0.005; ***, p<0.001.

and TTN (49%), MUC16 (43%), SYNE1 (26%) in the low-CTSL group (Figure 7C).

Furthermore, we used Fisher's exact test to determine the ratio of mutation frequencies between the high and low CTSL groups and ranked the results based on increasing p-values. The high CTSL group showed higher TP53 and NOTCH1 mutation load while exhibiting lower mutation loadings of CYLD, MRC1, MKRN3, and RELN compared to the low CTSL group (Figure 7A). In addition, the concurrent or mutually exclusive mutations in the 25 most frequently mutated genes are shown in Figure 7D. The high CTSL group had significantly more concurrent gene alterations than the low CTSL group. notch1 mutations frequently occurred concurrently with FAT1 mutations in the high CTSL group. Other intimate mutant loci included RYR2 and TTN, RYR2 and LRP1B, CSMD3, and RYR2. Common co-mutations in the low CTSL group included NSD1 and USH2A, LAMA2, and TP53. Meanwhile, dense mutually exclusive gene alteration pairs were identified, such as TP53-CASP8 in the high CTSL group (Figure 7D).

3.8 CTSL-based signature is correlated with immune status in TCGA cohorts with HNSCC

To examine the relationship between CTSL-based signature and immune status in TCGA patients with HNSCC, immune checkpoint expression levels were observed to be considerably higher in HNSCC patients with low-CTSL-based signature scores than in those with high-CTSL-based signature scores (Figure 8A), demonstrating that in HNSCC patients, the levels of CTSL signature scores were inversely linked with immunological state. Furthermore, GSEA analysis found that HNSCC patients with high CTSL-related signature scores displayed increased activity in many important immune-related pathways, including adaptive immune response, immunological response, T-cell receptor signaling pathways, and T-cell activation (Figure 8B). Our research suggested that tumor immunity may be closely related to the CTSL signature for HNSCC patients.

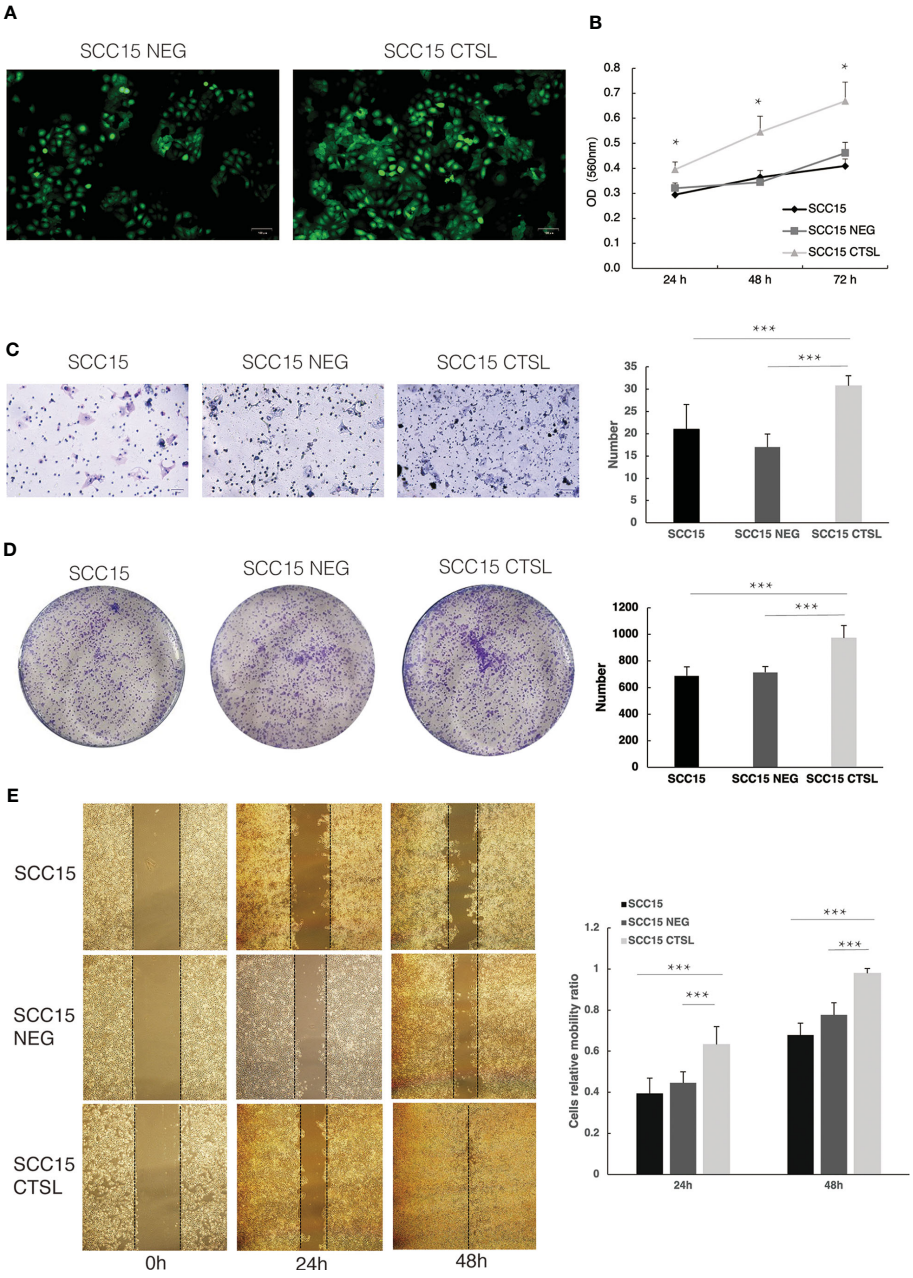


FIGURE 4
CTSL promotes tumor cell proliferation and metastasis in HNSC. **(A)**Stable transfection of lentivirus with negative control (NEG) or CTSL in SCC15 cells is indicated by green fluorescence (200×). **(B)**CCK8 assay in SCC15 cells, SCC15 NEG cells and SCC15 CTSL cells. **(C)** Transwell assay and its quantitative analysis in SCC15 cells, SCC15 NEG cells and SCC15 CTSL cells (200×). **(D)** Plate cloning assay and its quantitative analysis in SCC15 cells, SCC15 NEG cells and SCC15 CTSL cells. **(E)** scratch migration assay and its quantitative analysis in SCC15 cells, SCC15 NEG cells and SCC15 CTSL cells(40×). *, p<0.05; ***, p<0.01.

Overall, our data show that CTSL is a good indicator of a patient's prognosis for HNSCC.

4 Discussion

A global outbreak of coronavirus infections was brought on by the new human-infecting beta coronavirus known as the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (1). Therefore, to effectively manage SARS-CoV-2, it is crucial to comprehensively examine the factors contributing to COVID-19 susceptibility, the underlying mechanisms of these factors, and implementing measures to prevent infection. It has been established that CTSL is important in coronavirus infection of host cells. Bollavaram et al. revealed that the SARS-CoV-2 spike protein has numerous regions vulnerable to CTSL proteolysis (13). A recent study found that downregulating CTSL expression could reduce the virus's ability to penetrate host cells (14). An increasing body of evidence suggests that cancer patients are more likely to become severely ill or die from SARS-CoV-2

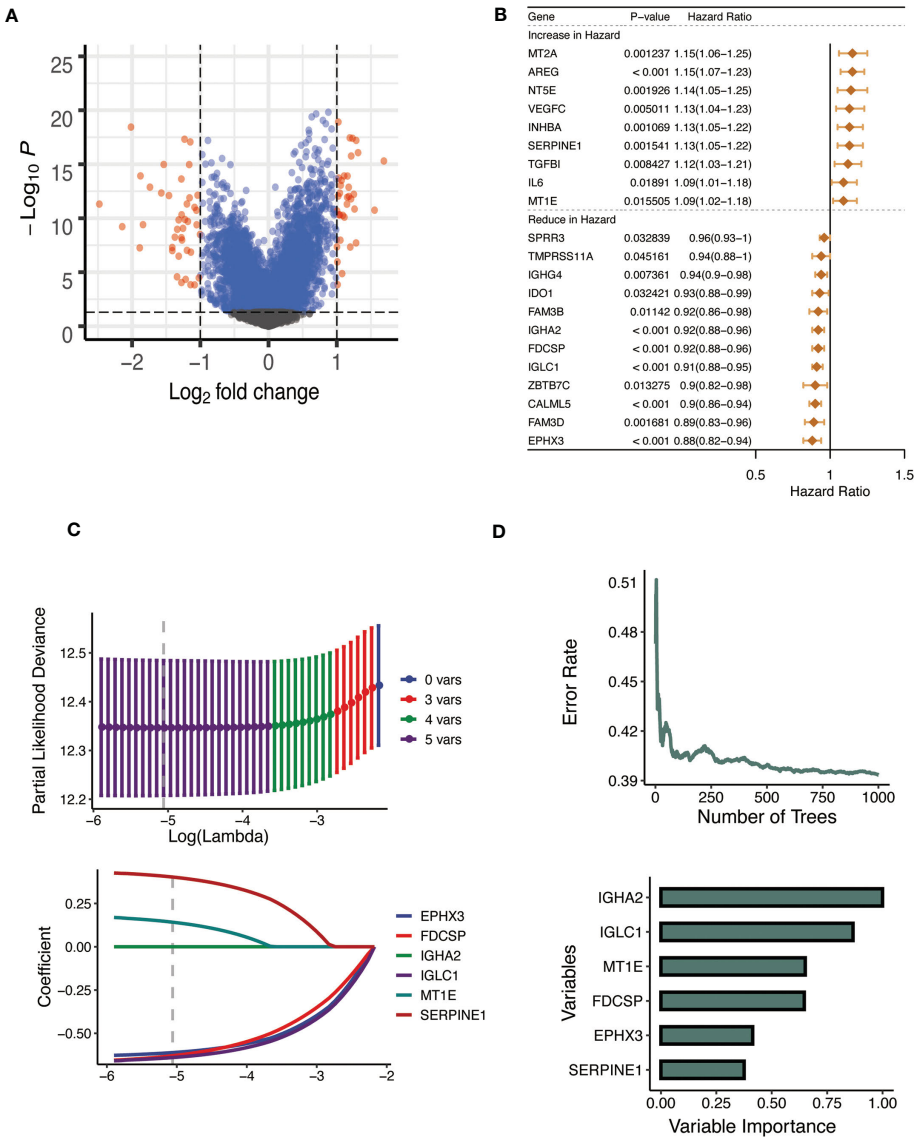


FIGURE 5
Establishment of the CTSL signature. **(A)** Volcano plot showing the results of the analysis of differences between high and low CTSL groups. **(B)** Univariate Cox analysis forest plot of 21 prognosis-related to CTSL. **(C)** Machine learning method for survival random forest to further screen CTSL signature). **(D)** Lasso regression method to calculate CTSL signature.

infection than those without cancer (15–17). In the present study, we assessed the expression levels of viral entry receptors such as CTSL in HNSC cancer tissues since malignant pathology may influence COVID-19 susceptibility and sickness (1, 17).

Our findings provided evidence that HNSC exhibited elevated expression of CTSL. It has been suggested that organs with high CTSL expression may eventually become infected with SARS-CoV-2 (18). Although CTSL expression in HNSC is not as high as in lung cancer, HNSC patients are at significant risk of contracting SARS-CoV-2 infection. The higher susceptibility to infection is probably attributed to the exposure of organs to the external environment, which offers favorable transmission routes for SARS-CoV-2 (19, 20).

In this study, the predictive significance of CTSL for patients with HNSC was assessed using a Cox regression analysis and a prognostic nomogram based on CTSL expression and OS (Figure 1). These findings suggest that CTSL may be a potential biomarker for determining the course of HNSC.

A potent tool that can handle the large and complex datasets produced by high-throughput technologies is machine learning used in bioinformatics. It can accurately and quickly analyze biological data, spot trends, and relationships, support drug discovery, enable personalized medicine, and speed up biological research. In the end, machine learning in bioinformatics improves disease understanding, detection, and treatment. This study also applied various machine learning methods, such as random survival forest and Lasso regression, which makes the current study convincing (18, 21).

Moreover, our genomic analysis revealed a correlation between CTSL and mutations in P53, ATRX, and PTEN, and TP53

10.3389/fimmu.2023.1156038

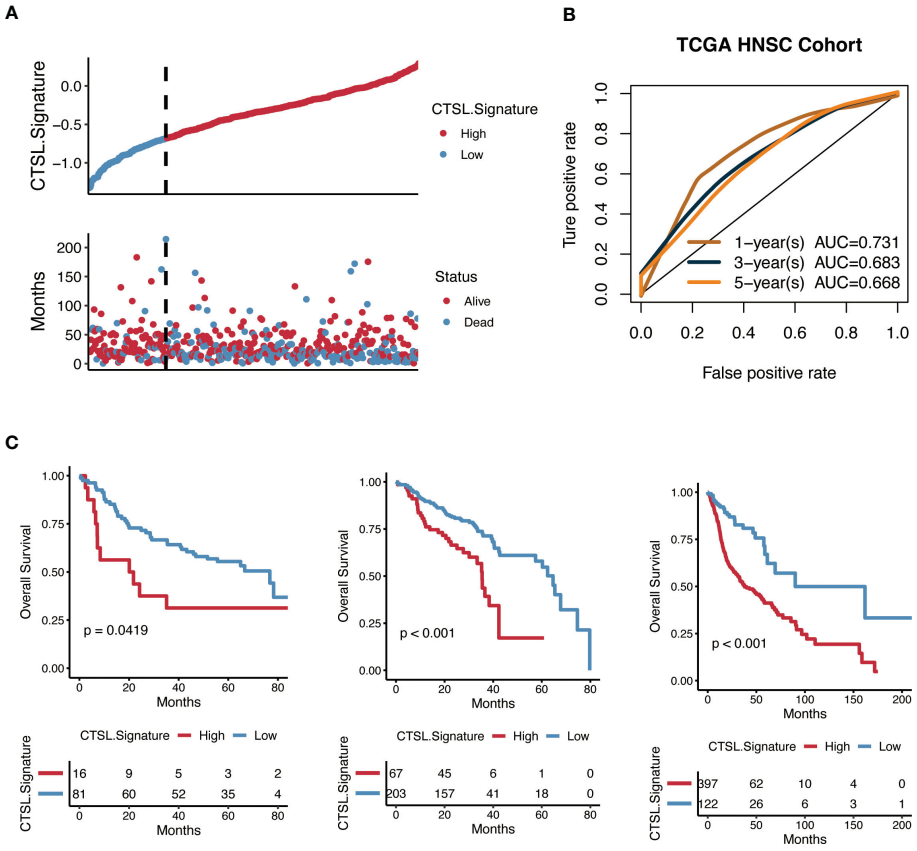


FIGURE 6
Predictive potential of the CTSL signature for prognosis in HNSCC patients. **(A)** Kaplan–Meier curves of high and low CTSL signature in TCGA-HNSCC. **(B)** Time-dependent ROC of CTSL signature in TCGA. **(C)** Kaplan–Meier curves of overall survival in HNSCC patients based on external validation datasets (GSE41613, GSE65858, and TCGA).

mutations were associated with poor survival in HNSCC patients and tumor resistance to radiation and chemotherapy clinically (22). Importantly, we found that patients with high expression of CTSL had low IC50 values for 5-fluorouracil, dasatinib, ERK_2440, JAK1_8709, luminespib, and staurosporine, suggesting that CTSL expression was a reliable indicator of therapeutic sensitivity for these potential molecular drugs. Clinical trials have validated the safety and compatibility of these drugs, thereby providing general treatment recommendations for HNSC.

A preliminary overview of CTSL and immune cell infiltration was provided in this study. High CTSL expression in HNSC patient tissues may lead to a deterioration in immune function in patients infected with SARS-CoV-2 because of the relationship between CTSL and the immune response. Furthermore, the infiltration of CD8+ T cells, B cells, CD4+ T cells, neutrophils, macrophages, and dendritic cells in HNSC was associated with the expression of CTSL. Moreover, CTSL was associated with various gene set indicators of different immune cell types. As reported, COVID-19's poor prognosis is attributed to cytokine storms and inflammatory immune responses (23). The potential role of CTSL in regulating antitumor immunity and its therapeutic significance in HNSC remains unclear. Adaptive immunity is essential for effective viral clearance after SARS-CoV-2 infection. Viral infections and associated chronic inflammatory responses are often associated with cancer manifestation and progression. SARS-CoV-2 may not directly cause cancer, but it may alter the immune landscape and cause adverse outcomes in patients with cancer (24), The immune response of cancer patients may also contribute to the adverse effects of SARS-CoV-2 infection (23). It is possible that CTSL plays a key role in cancer progression and SARS-CoV-2 infection.

There are several limitations and shortcomings to this study. Firstly, the findings heavily rely on bioinformatics analysis, and there is no validation cohort to confirm the results. Additionally, due to the lack of experimental evidence and mechanistic investigations, it is challenging to understand the relationship between SARS-CoV-2 and cancer data and the potential contribution of CTSL. Therefore, more experimental evidence is required.

Moreover, we developed a CTSL-based signature in this study. Patients with low CTSL-based signature scores had higher immune scores, stromal cells, and immune-related pathways, and they responded well to immunotherapy and targeted therapy compared to those with high CTSL-based signature scores. Our results suggest that the CTSL-based signature is a reliable prognostic predictor for HNSCC.

To summarize, we have discussed the clinical and molecular significance of CTSL in HNSC. Our analysis revealed high

10.3389/fimmu.2023.1156038

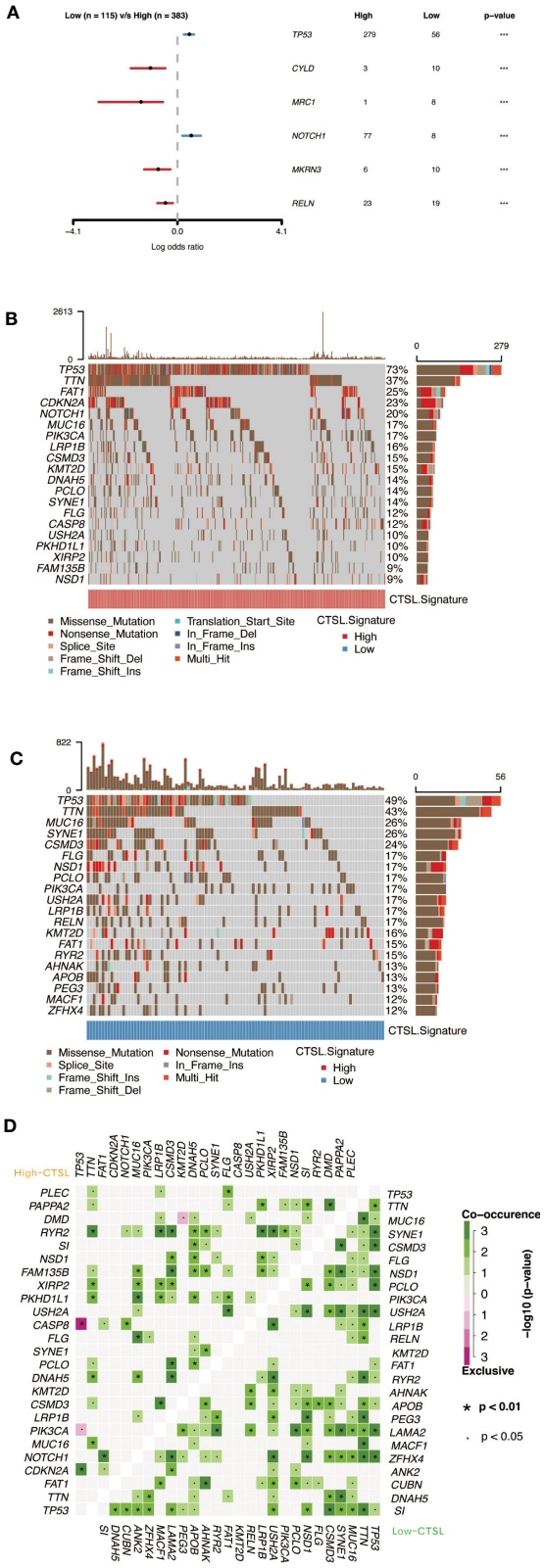


FIGURE 7
Genomic alterations associated with CTSL signature in HNSCC samples. **(A)** Forest plot showing the results of somatic mutation difference analysis between high and low groups of CTSL signature. **(B, C)** oncoplot of somatic mutations in HNSCCs between high and low CTSL signature groups. **(D)** The heatmap presents the somatic interaction of HNSCCs between CTSL signature high and low groups.

10.3389/fimmu.2023.1156038

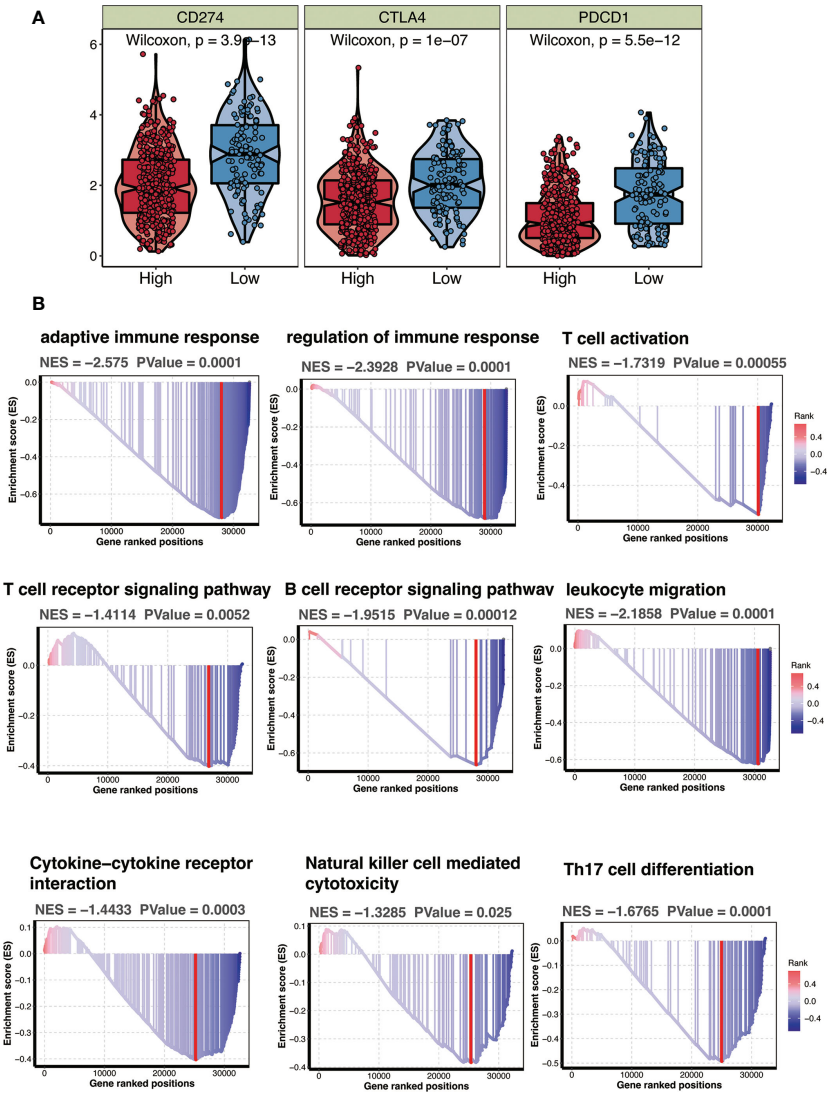


FIGURE 8
CTSL signature in relation to immunity in the TCGA cohort. **(A)** Changes in ESTIMATEclassical immune checkpoints among HNSCC patients with high and low CTSL signature. **(B)** GSEA showing immune related pathways potentially related by CTSL signature.

expression of CTSL in HNSC and its association with poor prognosis and immune cell infiltration. This finding suggests that CTSL may have a biological role in HNSC. Other gene signatures may also have prognostic value for HNSCC, similar to the CTSL-based signature developed in this study. Future studies should investigate these signatures to identify the optimal therapeutic targets for HNSCC.

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 within the article/Supplementary Materials.

Author contributions

The study was created and planned by OS and DL. XW and NQ collected the data. FG, MXZ, and MFL carried out the bioinformatic analysis, and FG wrote the manuscript. The statistical analysis was carried out by XW and MQZ under OS's supervision. The manuscript has been read, carefully amended, and approved by all authors. All authors contributed to the article.

Funding

This study was financially supported by the Key Fundamental Research Fund of Science and Technology Foundation of Shenzhen City (Grant No. JCYJ20210324094005015 and JCYJ20220818095811026), Start-up Research Fund of Shenzhen University for Youth Scholars (Grant No. 860-000002112112) and Shenzhen Science and Technology Innovation Commission stable support project (Grant No. 20220810174028001).

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.2023.1156038/full#supplementary-material

References

1. Lan J, Ge Y, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* (2020) 581:215–20. doi: 10.1038/s41586-020-2180-5

2. Gemmati D, Bramanti B, Serino ML, Secchiero P, Zauli G, Tisato V. COVID-19 and individual genetic Susceptibility/Receptivity: role of ACE1/ACE2 genes, immunity, inflammation and coagulation. might the double X-chromosome in females be protective against SARS-CoV-2 compared to the single X-chromosome in males? *IJMS* (2020) 21:3474. doi: 10.3390/ijms21103474

3. Hou Y, Zhao J, Martin W, Kallianpur A, Chung MK, Jehi L, et al. New insights into genetic susceptibility of COVID-19: an ACE2 and TMPRSS2 polymorphism analysis. *BMC Med* (2020) 18:216. doi: 10.1186/s12916-020-01673-z

4. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* (2020) 181:271–280.e8. doi: 10.1016/j.cell.2020.02.052

5. Sica A, Massarotti M. Myeloid suppressor cells in cancer and autoimmunity. *J Autoimmun* (2017) 85:117–25. doi: 10.1016/j.jaut.2017.07.010

6. Qin C, Zhou L, Hu Z, Zhang S, Yang S, Tao Y, et al. Dysregulation of immune response in patients with COVID-19 in wuhan, China. *SSRN J* (2020) 762–8. doi: 10.2139/ssrn.3541136

7. Merad M, Martin JC. Pathological inflammation in patients with COVID-19: a key role for monocytes and macrophages. *Nat Rev Immunol* (2020) 20:355–62. doi: 10.1038/s41577-020-0331-4

8. Liang W, Guan W, Chen R, Wang W, Li J, Xu K, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. *Lancet Oncol* (2020) 21:335–7. doi: 10.1016/S1470-2045(20)30096-6

9. Curigliano G. Cancer patients and risk of mortality for COVID-19. *Cancer Cell* (2020) 38:161–3. doi: 10.1016/j.ccell.2020.07.006

10. Bao R, Hernandez K, Huang L, Luke JJ. *ACE2* and *TMPRSS2* expression by clinical, HLA, immune, and microbial correlates across 34 human cancers and matched normal tissues: implications for SARS-CoV-2 COVID-19. *J Immunother Cancer* (2020) 8:e001020. doi: 10.1136/jitc-2020-001020

11. Cheng X, Wang X, Nie K, Cheng L, Zhang Z, Hu Y, et al. Systematic pan-cancer analysis identifies TREM2 as an immunological and prognostic biomarker. *Front Immunol* (2021) 12:646523. doi: 10.3389/fimmu.2021.646523

12. Wölfl M, Schwinn S, Yoo Y-E, Reß ML, Braun M, Chopra M, et al. Src-kinase inhibitors sensitize human cells of myeloid origin to toll-like-receptor-induced interleukin 12 synthesis. *Blood* (2013) 122:1203–13. doi: 10.1182/blood-2013-03-488072

13. Bollavaram K, Leeman TH, Lee MW, Kulkarni A, Upshaw SG, Yang J, et al. Multiple sites on SARS-CoV -2 spike protein are susceptible to proteolysis by cathepsins b, K, l, s and V. *Protein Sci* (2021) 30:1131–43. doi: 10.1002/pro.4073

14. Kwan JYY, Lin L-T, Bell R, Bruce JP, Richardson C, Pugh TJ, et al. Elevation in viral entry genes and innate immunity compromise underlying increased infectivity and severity of COVID-19 in cancer patients. *Sci Rep* (2021) 11:4533. doi: 10.1038/s41598-021-83366-y

15. Sawyers A, Chou M, Johannet P, Gulati N, Qian Y, Zhong J, et al. Clinical outcomes in cancer patients with COVID -19. *Cancer Rep* (2021) 4(6):e1413. doi: 10.1002/cnr2.1413

16. Zhang H, Han H, He T, Labbe KE, Hernandez AV, Chen H, et al. Clinical characteristics and outcomes of COVID-19–infected cancer patients: a systematic review and meta-analysis. *JNCI: J Natl Cancer Institute* (2021) 113:371–80. doi: 10.1093/jnci/djaa168

17. Cheng J, Zhou J, Fu S, Fu J, Zhou B, Chen H, et al. Prostate adenocarcinoma and COVID-19: the possible impacts of *TMPRSS2* expressions in susceptibility to SARS-CoV-2. *J Cell Mol Med* (2021) 25:4157–65. doi: 10.1111/jcmm.16385

18. Wu X-N, Su D, Mei Y-D, Xu M-Q, Zhang H, Wang Z-Y, et al. Identified lung adenocarcinoma metabolic phenotypes and their association with tumor immune microenvironment. *Cancer Immunol Immunother* (2021) 70:2835–50. doi: 10.1007/s00262-021-02896-6

19. Patel KP, Vunnam SR, Patel PA, Krill KL, Korbitz PM, Gallagher JP, et al. Transmission of SARS-CoV-2: an update of current literature. *Eur J Clin Microbiol Infect Dis* (2020) 39:2005–11. doi: 10.1007/s10096-020-03961-1

20. He C, Hua X, Sun S, Li S, Wang J, Huang X. Integrated bioinformatic analysis of SARS-CoV-2 infection related genes ACE2, BSG and TMPRSS2 in aerodigestive cancers. *JIR* (2021) 14:791–802. doi: 10.2147/JIR.S300127

21. Zhang N, Zhang H, Wu W, Zhou R, Li S, Wang Z, et al. Machine learning-based identification of tumor-infiltrating immune cell-associated lncRNAs for improving outcomes and immunotherapy responses in patients with low-grade glioma. *Theranostics* (2022) 12:5931–48. doi: 10.7150/thno.74281

22. Lindemann A, Takahashi H, Patel AA, Osman AA, Myers JN. Targeting the DNA damage response in OSCC with *TP* 53 mutations. *J Dent Res* (2018) 97:635–44. doi: 10.1177/0022034518759068

23. Latif MB, Shukla S, Del Rio Estrada PM, Ribeiro SP, Sekaly RP, Sharma AA. Immune mechanisms in cancer patients that lead to poor outcomes of SARS-CoV-2 infection. *Trans Res* (2022) 241:83–95. doi: 10.1016/j.trsl.2021.12.001

24. Han HJ, Nwagwu C, Anyim O, Ekweremadu C, Kim S. COVID-19 and cancer: from basic mechanisms to vaccine development using nanotechnology. *Int Immunopharmacol* (2021) 90:107247. doi: 10.1016/j.intimp.2020.107247

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 17 July 2023
DOI 10.3389/fonc.2023.1195392

OPEN ACCESS

EDITED BY
Ping Zheng,
The University of Melbourne, Australia

REVIEWED BY
Qiaoqiao Li,
Guangdong Academy of Medical Sciences, China
Yirui Chen,
Hangzhou Medical College, China

*CORRESPONDENCE
Zhiwei Liu
3191038@zju.edu.cn
Fangjian Chen
cfj68@163.com

RECEIVED 28 March 2023
ACCEPTED 12 June 2023
PUBLISHED 17 July 2023

CITATION
Pan Y, Wang Y, Hu M, Xu S, Jiang F, Han Y, Chen F and Liu Z (2023) Aggrephagy-related patterns in tumor microenvironment, prognosis, and immunotherapy for acute myeloid leukemia: a comprehensive single-cell RNA sequencing analysis.
Front. Oncol. 13:1195392.
doi: 10.3389/fonc.2023.1195392

Aggrephagy-related patterns in tumor microenvironment, prognosis, and immunotherapy for acute myeloid leukemia: a comprehensive single-cell RNA sequencing analysis

Yan Pan[1], Yingjian Wang[2], Mengsi Hu[2], Shoufang Xu[2], Feiyu Jiang[2], Yetao Han[2], Fangjian Chen[1]* and Zhiwei Liu[2]*

[1]Department of Blood Transfusion, The Quzhou Affiliated Hospital of Wenzhou Medical University, Quzhou People's Hospital, Quzhou, Zhejiang, China, [2]Department of Blood Transfusion, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

Acute myeloid leukemia (AML) is a complex mixed entity composed of malignant tumor cells, immune cells and stromal cells, with intra-tumor and inter-tumor heterogeneity. Single-cell RNA sequencing enables a comprehensive study of the highly complex tumor microenvironment, which is conducive to exploring the evolutionary trajectory of tumor cells. Herein, we carried out comprehensive analyses of aggrephagy-related cell clusters based on single-cell sequencing for patients with acute myeloid leukemia. A total of 11 specific cell types (T, NK, CMP, Myeloid, GMP, MEP, Promono, Plasma, HSC, B, and Erythroid cells) using t-SNE dimension reduction analysis. Several aggrephagy-related genes were highly expressed in the 11 specific cell types. Using Monocle analysis and NMF clustering analysis, six aggrephagy-related $CD8^+$ T clusters, six aggrephagy-related NK clusters, and six aggrephagy-related Mac clusters were identified. We also evaluated the ligand-receptor links and Cell–cell communication using CellChat package and CellChatDB database. Furthermore, the transcription factors (TFs) of aggrephagy-mediated cell clusters for AML were assessed through pySCENIC package. Prognostic analysis of the aggrephagy-related cell clusters based on R package revealed the differences in prognosis of aggrephagy-mediated cell clusters. Immunotherapy of the aggrephagy-related cell clusters was investigated using TIDE algorithm and public immunotherapy cohorts. Our study revealed the significance of aggrephagy-related patterns in tumor microenvironment, prognosis, and immunotherapy for AML.

KEYWORDS

acute myeloid leukemia, aggrephagy, immune cell, prognosis, immunotherapy, microenvironment

Introduction

Leukemia is a malignant clonal disease originating from hematopoietic stem cells (1). The affected cells have uncontrolled proliferation, impaired differentiation, and blocked apoptosis, so the affected cells are stuck in different stages of cell development (2). The incidence and mortality rate of leukemia are both high. The report showed that in 2018 alone, there were 437000 new cases of leukemia and 309000 new deaths from leukemia worldwide (3).

Leukemia can be classified as acute (4) or chronic (5) according to its course. Leukemias can be divided into myeloid leukemia and lymphocytic leukemia according to the cells involved (6). Acute myeloid leukemia (AML), the most common leukemia in adults, is a highly heterogeneous disease (7). French-American-British (FAB) defined eight subtypes (M0 to M7) based on the morphological and cytological characteristics of leukemia cells (8). According to genetics, morphology, immunophenotype and clinical manifestations, World Health Organization (WHO) classified leukemia into six main types and more than 20 subtypes (9). In addition, the prognosis of AML can be divided into good, moderate and poor groups based on cytogenetic characteristics (1, 6), but the prognosis of different patients in each group is still very different, indicating that the gene expression pattern of leukemia is very complex.

Tumor microenvironment (TME) is the internal environment that tumor cells depend on for survival and development. Besides tumor cells, it also contains many non-malignant cells and some soluble factors, which play an important role in promoting tumor occurrence, progression and immune escape (10). Tumor microenvironment mainly includes immune microenvironment, including myeloid-derived suppressor cells (MDSCs), tumor-associated macrophages (TAMs), tumor-associated neutrophils (TANs), dendritic cell (DC), T cell, B cell, and Natural Killer (NK) cell, and non-immune microenvironment, including cancer-associated fibroblasts (CAFs), extracellular matrix, mesenchymal stem cells, and various secreted factors (11–14). Therefore, tumor is a complex mixed entity composed of malignant tumor cells, immune cells and stromal cells, with intra-tumor and inter-tumor heterogeneity. Since bulk tissue is composed of various cells, its sequencing cannot reveal the function or cell state of a specific cell population (15). Therefore, the detection of genome, transcriptome, epigenome and proteome at the cellular level can overcome the limitations of the traditional bulk level and conduct more detailed analysis at the cellular and molecular level (15). Single-cell RNA sequencing (scRNA-Seq) enables non-targeted quantification of transcripts in a single cell. Single-cell RNA sequencing enables a comprehensive study of the highly complex tumor microenvironment, which is conducive to exploring the evolutionary trajectory of tumor cells, the complex interactions between tumor cells and tumor microenvironment, and the spatio-temporal functional relationships between different cell population types (16, 17). Bioinformation analysis can identify new cell types, identify rare cell populations, and construct cell status and phylogenetic maps through computational methods such as high-dimensional data reduction, unsupervised clustering, phylogenetic modeling, locus inference, RNA rate analysis, lineage tracing, and ligand-receptor interaction mapping (16, 17).

Autophagy is an important feedback process of cells under pressure. Autophagy realizes self-digestion and catabolism by phagocytic organelles and degradation of cell contents, so as to maintain the homeostasis balance of cells (18, 19). Autophagy plays an important role in maintaining vital activities and immune function and is closely related to tumors and other diseases. The common types of autophagy include macroautophagy, microautophagy and chaperonemediated autophagy (20). Aggrephagy is a kind of selective autophagy, which is the only way to clear protein aggregates. Once the function of molecular chaperone and ubiquitin proteasome is limited or the clearance efficiency of misfolded proteins is lower than the production rate, protein aggregates will be formed, and the aggrephagy needs to be activated to degrade them (21).

In this study, the relationship between aggrephagy-related genes and cell subsets of TME (such as T cells, Natural Killer cell, and Myeloid cells) for AML was investigated using data of single-cell RNA-sequencing (scRNA-seq) from GSE116256. After Nonnegative Matrix Factorization analysis, the characteristics of the aggrephagy-mediated cell clusters in Pseudotime trajectory, cell–cell communication, ligand-receptor links, and immunotherapy were investigated.

Materials and methods

Downloading and preprocessing for data of acute myeloid leukemia

The samples source with single-cell RNA-sequencing (scRNA-seq, GSE116256) and expression profiles (GSE63270 and GSE12417) were downloaded from Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) of The National Center for Biotechnology Information (NCBI) (22). We enrolled three normal samples and ten patients with acute myeloid leukemia from GSE116256 for analysis of scRNA-seq (23–25). There were 104 normal and acute myeloid leukemia (42 populations and 62 leukemic populations) included in GSE63270 dataset (26). GSE12417 dataset contained the analysis of 79 samples of bone marrow or peripheral blood mononuclear cells from adult patients with untreated acute myeloid leukemia (27, 28). In addition, the expression profiles and clinical information were acquired from TCGA-LAML cohort, including 151 patients with acute myeloid leukemia (29, 30).

Dimensionality reduction and annotation of single cell for acute myeloid leukemia

First, the data of single cell was filtered by setting each gene to be expressed in at least three cells, and each cell to express at least 500 genes, resulting in 9891 cells. We calculated the percentage of mitochondria and Ribosomal RNA (rRNA) through the PercentageFeatureSet function of Seurat package (31). The number of genes expressed in each single cell was greater than 100 and less than 5000, and we ensured the percentage of mitochondria was less than 20%. Furthermore, the Unique

Molecular Identifier (UMI) of the single cell was at least greater than 100, resulting in 9886 cells. Subsequently, we used the method of log-normalization to standardize the single-cell data from each of the 13 samples. The highly variable features were identified by FindVariableFeatures function (32) based on variance stabilization transformation (VST). The genes were then scaled by using the ScaleData function for all genes. We utilized RunPCA function for PCA dimension reduction to find anchors. The FindNeighbors function with dim=15 and FindClusters function with Resolution=0.1 was used to luster cells. Ulteriorly, the RunTSNE function was used to conduct t-SNE (T-Distribution Stochastic Neighbour Embedding) dimension reduction analysis and the RunUMAP function was used to conduct UMAP (Uniform Manifold Approximation and Projection) reduction analysis. The marker genes for single cell were supplied by SingleR package (33) and the classical marker from the published literature (25).

Pseudotime trajectory analysis for the aggrephagy-mediated cell clusters

Monocle R package was applied for the data of single cell to explore the correlation of aggrephagy-related genes and pseudotime trajectories (34). The graphs for the pseudotime trajectories of specific cell with aggrephagy-related genes were plotted using the function from Monocle R package, such as plot_pseudotime_heatmap and so on.

Nonnegative matrix factorization of aggrephagy-related genes in single cell for acute myeloid leukemia

Based on the expression matrix of the scRNA-seq, dimension reduction analysis of aggrephagy-related genes in each cell clusters were conducted employing NMF (Nonnegative Matrix Factorization) R package (35, 36), thus displaying the effect of aggrephagy-related genes in single cell for acute myeloid leukemia.

Identifying the marker genes of single cell for acute myeloid leukemia

FindAllMarkers function was applied to identify he marker genes of single cell for acute myeloid leukemia (31). The aggrephagy-mediated cell clusters were identified based on differentially expressed genes (DEGs) with log Fold Change (logFC) and aggrephagy-related genes. The NK cell subtypes were summarized from the published literature of Huan Liu et al (37).

Analysis of transcription factors for aggrephagy-mediated cell clusters

SCENIC was a tool for simultaneously reconstructing gene regulatory networks and identifying stable cell states from single-cell RNA-seq data (38). The gene regulatory network was inferred based on co-expression and DNA motif analysis, and then network activity was analyzed in each cell to identify cell status (38). We carried out analysis of transcription factors (TFs) for aggrephagy-mediated cell clusters for acute myeloid leukemia through pySCENIC package (39–41). RcisTarget R package and two gene-motif rankings (hg19-tss-centered-10 kb and hg19-500 bp-upstream) was used to identify binding motifs of TFs in the gene list for acute myeloid leukemia (42, 43). The threshold value for the TFs was set as Benjamini–Hochberg false discovery rate (BH-FDR) <0.05.

Cell–cell communication analysis among cell subsets for acute myeloid leukemia

The signaling inputs and outputs among the cell types and aggrephagy-mediated cell clusters were assessed by applying CellChat package (44) and CellChatDB database (45). The netVisual_circle function was utilized for evaluating the strength of cell–cell communication networks among cell subsets (44, 45). In addition, the ligand-receptor interactions among the specific cell subsets were estimated *via* the netVisual_bubble function (44, 45).

Prognostic analysis of the aggrephagy-related cell clusters for acute myeloid leukemia

Based on the data of scRNA, Gene Set Variation Analysis (GSVA) was applied to compute the signature scores involved in aggrephagy for public database (46). We carried out Cox proportional hazard regression to evaluate the prognosis for the aggrephagy-related cell clusters (47). The Kaplan–Meier curves was plotted through the survminer R package.

Immunotherapy analysis of the aggrephagy-related cell clusters for acute myeloid leukemia

We used TIDE (Tumor Immune Dysfunction and Exclusion) algorithm to analyze the immune checkpoint blockade immunotherapeutic for the aggrephagy-related cell clusters (48). We also reviewed the published literature to validate the prognostic and therapeutic effects of each cell subtype using real-world immunotherapy cohorts (49–60).

Statistical analysis

The continuous or category variables were compared using Student's t-test, Wilcoxon rank sum test, Kruskal–Wallis's test, or Chi-square test. The log-rank test was used for survival analyses.

Results

Dimensionality reduction and annotation of single cell for acute myeloid leukemia

We carried out dimensionality reduction and annotation of single cell for acute myeloid leukemia as described in the *materials and methods* section. We ensured that the number of genes expressed in each single cell was greater than 100 and less than 5000, the percentage of mitochondria was less than 20%, and the Unique Molecular Identifier (UMI) of the single cell was at least greater than 100, resulting in 9886 cells. Supplementary Figures S1A–B was the statistical diagram of cell filtration, which could be seen to meet all thresholds set above (Supplementary Figures S1A, B). The highly variable features were identified *via* FindVariableFeatures function based on VST, and the top ten highly variable genes among the single cell were marked out in the volcano plot, including IGLL5, HBB, JCHAIN, HBG1, HBG2, HBD, HBA2, CLC, CA1, and HBA1 (Supplementary Figure S1C). Ulteriorly, PCA analysis was carried out on the highly variable genes, we used the Elbow algorithm to carry out the Standard Deviation based on the highly variable genes (Supplementary Figure S1D). The RunTSNE function was used to conduct t-SNE dimension reduction analysis and the RunUMAP function was used to conduct UMAP reduction analysis, thus identifying a total of 18 cell subsets (Supplementary Figure S1E). Afterwards, the marker genes were used to annotate the specific cell types, thus identifying 11 specific cell types, including T cell, Natural Killer (NK) cell, Common Myeloid Progenitor (CMP) cell, Myeloid cell, Granulocyte Monocyte Progenitor (GMP) cell, Megakaryocyte Erythroid Progenitor (MEP) MEP, Promonocyte (Promono) cell, Plasma cell, Hematopoietic Stem Cell (HSC) cell, B cell, and Erythroid cell (Figure 1A). We plotted the correlation network for the number of interactions among the 11 specific cell types (Figure 1B). Figure 1C visually showed the proportion of different cell types in each sample. Finally, we created the heat map to show the expression of the aggrephagy-related genes in different cell types (Figure 1D). We could see that there were several aggrephagy-related genes that were highly expressed in the 11 specific cell types, including HSP90AA1, RPS27A, UBA52, UBB, UBC, and VIM (Figure 1D). We displayed the global view of the expression pattern for marker genes gained as described in the methods section, reflecting the dynamic features of each cell subsets (Supplementary Figure S2).

Pseudotime analysis for aggrephagy-mediated T cells

There was a total of 2397 cells in the T cell type. Using UMAP reduction analysis, the 2397 cells in the T cell type could be clustered into eight cell clusters (Supplementary Figure S3), and the global view of the expression pattern for marker genes of the eight cell clusters was displayed in Supplementary Figure S3. Further, the eight cell clusters could be re-clustered into nine cell subsets (Supplementary Figure S3). Based on the aggrephagy-related genes, six clusters were identified using Monocle analysis (Figure 2A), including Cluster 1 (DYNC1LI2 and TUBB1), Cluster 2 (PRKN, TUBB2A, DYNLL2, RPS27A and UBA52), Cluster 3 (TUBA1A, UBE2V1, IFT88 and VCP), Cluster 4 (TUBA1C, TUBA3C, TUBA3D, HSP90AA1, VIM, DYNC1I2, PARK7, UBE2N, DYNLL1 and HSF1), Cluster 5 (TUBA4A, TUBB4B, TUBA1B, ARL13B and PCNT), and Cluster 6 (DYNC1LI1, UBB, HDAC6, DYNC1H1 and UBC). From the heatmap generated by Pseudotime analysis, the critical role of the aggrephagy-related genes in the trajectory process of T cells was observed (Figure 2A). Four subgroups of T cells were obtained by re-clustering annotation using t-SNE dimension reduction analysis, including $CD8^+$ T cell, $CD4^+$ T cell, natural killer (NK) cells, and Regulatory T (Treg) cells (Figure 2B). Among the four cell subgroups, we found that $CD8^+$ T cell and $CD4^+$ T cell had a higher percentage both in tumor samples and normal samples than the other two cell subgroups (Figure 2C), and $CD8^+$ T cell occupied the highest proportion among the four cell subgroups (Figure 2C). The result of $CD8^+$ T cell revealed that the whole trajectory could be divided into three segments (State 1, State 2, and State 3) on the basis of the developmental order (Figure 2D). Ulteriorly, NMF clustering analysis of the aggrephagy-related gene set for the trajectories showed that these cells aggregated into nine clusters (Figure 2E). In addition, the results of UMAP reduction analysis indicated that the NMF cell types were clustered into six aggrephagy-related $CD8^+$ T clusters, including TUBA1B+$CD8^+$ T −C1, DYNC1H1+$CD8^+$ T−C2, UBE2V1+$CD8^+$ T−C3, UBE2N +$CD8^+$ T−C4, Unc−$CD8^+$ T−C5, and Non−Aggre−$CD8^+$ T−C6 (Figure 2F). The number of ligand-receptor links among the six aggrephagy-related CD8+ T clusters was computed by Cell-Chat analysis (Figure 2G). The weights and strength of ligand-receptor links among the six aggrephagy-related $CD8^+$ T clusters was computed by Cell-Chat analysis (Supplementary Figure S3). Lastly, the discrepancies in the exhausted $CD8^+$ T ($CD8^+$_exhau), cytotoxic $CD8^+$ T ($CD8^+$_cyoto), and TFs (BTN3A1, BTN3A2, BTN2A2, LGALS9, TIGIT, CD274, BTLA, CTLA4, IL10, LAIR1, CD247, TGFB1, SLAMF7, CD160, CD244, HAVCR2, LAG3, CD96, ADORA2A, PDCD1, and CD48) among the six aggrephagy-related $CD8^+$ T clusters were visually displayed in the pathway heatmap (Figure 2H). TUBA1B+$CD8^+$ T−C1 tended to be exhausted $CD8^+$ T ($CD8^+$_exhau), while UBE2N+$CD8^+$ T−C4 tended to be cytotoxic $CD8^+$ T ($CD8^+$_cyoto) as shown in Figure 2H. It is noteworthy that TFs of LGALS9, TIGIT, BTLA, and CTLA4 were upregulated in the TUBA1B+$CD8^+$ T−C1, TFs of IL10 and CD160 were upregulated in the DYNC1H1+$CD8^+$ T−C2, ADORA2A was upregulated in the UBE2V1+$CD8^+$ T−C3, BTN3A1, BTN3A2, CD274, CD247, SLAMF7, LAG3, and PDCD1 were upregulated in the UBE2N +$CD8^+$ T−C4, BTN2A2 and LAIR1 were upregulated in the Non − Aggre−$CD8^+$ T−C6 (Figure 2H).

Pseudotime analysis for aggrephagy-mediated NK cells

There was a total of 1067 cells in the NK cell type. Using UMAP reduction analysis, the 1067 cells in the NK cell type could be clustered into eleven cell clusters (Supplementary Figure S4). Based

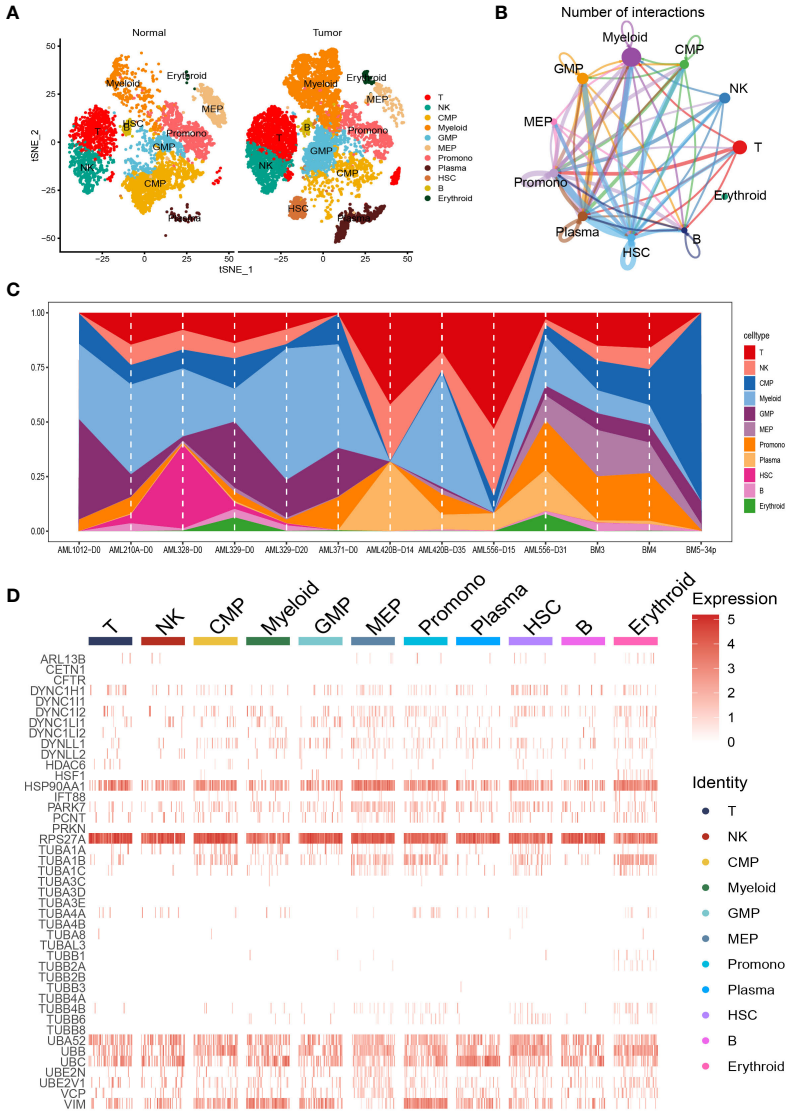


FIGURE 1
Dimensionality reduction and annotation of single cell for acute myeloid leukemia. **(A)** Cells were clustered into 11 specific cell types using t-SNE algorithm. **(B)** The number of interactions for communication among the 11 specific cell types. **(C)** The proportion of the 11 specific cell types in each sample. **(D)** Heat map showing the expression of the aggrephagy-related genes in 11 specific cell types.

on the aggrephagy-related genes, six clusters were identified using Monocle analysis (Figure 3A), including Cluster 1 (TUBA1B, TUBA4A, UBA52, UBB, TUBB4B, and UBE2V1), Cluster 2 (TUBA1C, HDAC6, VCP, DYNC1LI1, UBC, PRKN, TUBB1, HSP90AA1, RPS27A, PCNT, DYNC1H1, and DYNLL2), Cluster 3 (TUBA1A, VIM, HSF1, and UBE2N), Cluster 4 (TUBB6, ARL13B, DYNC1I2, DYNC1LI2, and TUBB2A), Cluster 5 (IFT88 and TUBA8), Cluster 6 (TUBA3C, DYNLL1, and PARK7). The results of UMAP reduction analysis indicated that the NMF cell types were clustered into six aggrephagy-related NK clusters, including UBE2N +NK–C1, UBE2V1+NK–C2, DYNC1H1+NK–C3, PARK7+NK–C4, Unc+NK–C5, and Non–Aggre–NK–C6 (Figure 3B). The number of ligand-receptor links among the six aggrephagy-related NK clusters was computed by Cell-Chat analysis (Figure 3C). The number, weights and strength of ligand-receptor links among the aggrephagy-related NK clusters was computed by Cell-Chat analysis (Supplementary Figure S4). Lastly, the discrepancies in the NK –CD56bright, NK–CD56dim, NK–HIA, LrNK–FCGR3A, LrNK –XCL1, KIR2DS1, NCR1, NCR2, NCR3, TLR3, TLR9, KIR3DL1, KIR2DL3, KLRB1, LILRB1, LILRB2, KLRG1, CEACAM1, CD244, LAIR1, CD96, TIGIT, and LAG3 among the six aggrephagy-related NK clusters were visually displayed in the pathway heatmap (Figures 3D, E). It is noteworthy that NK–CD56bright was upregulated in PARK7+NK–C4 (Figure 3D), NK–HIA was upregulated in UBE2N+NK–C1 (Figure 3D), LrNK–FCGR3A was upregulated in UBE2N+NK–C1 (Figure 3D), LrNK–XCL1 was upregulated in PARK7+NK–C4 (Figure 3D), NCR1 was upregulated in UBE2N+NK–C1 (Figure 3E), NCR3 was upregulated in PARK7 +NK–C4 (Figure 3E), TLR9 was upregulated in PARK7+NK –C4 (Figure 3E), KIR3DL1 was upregulated in UBE2V1+NK–C2 (Figure 3E), LILRB1 was upregulated in UBE2V1+NK–C2 (Figure 3E), LILRB2 was upregulated in DYNC1H1+NK–C3

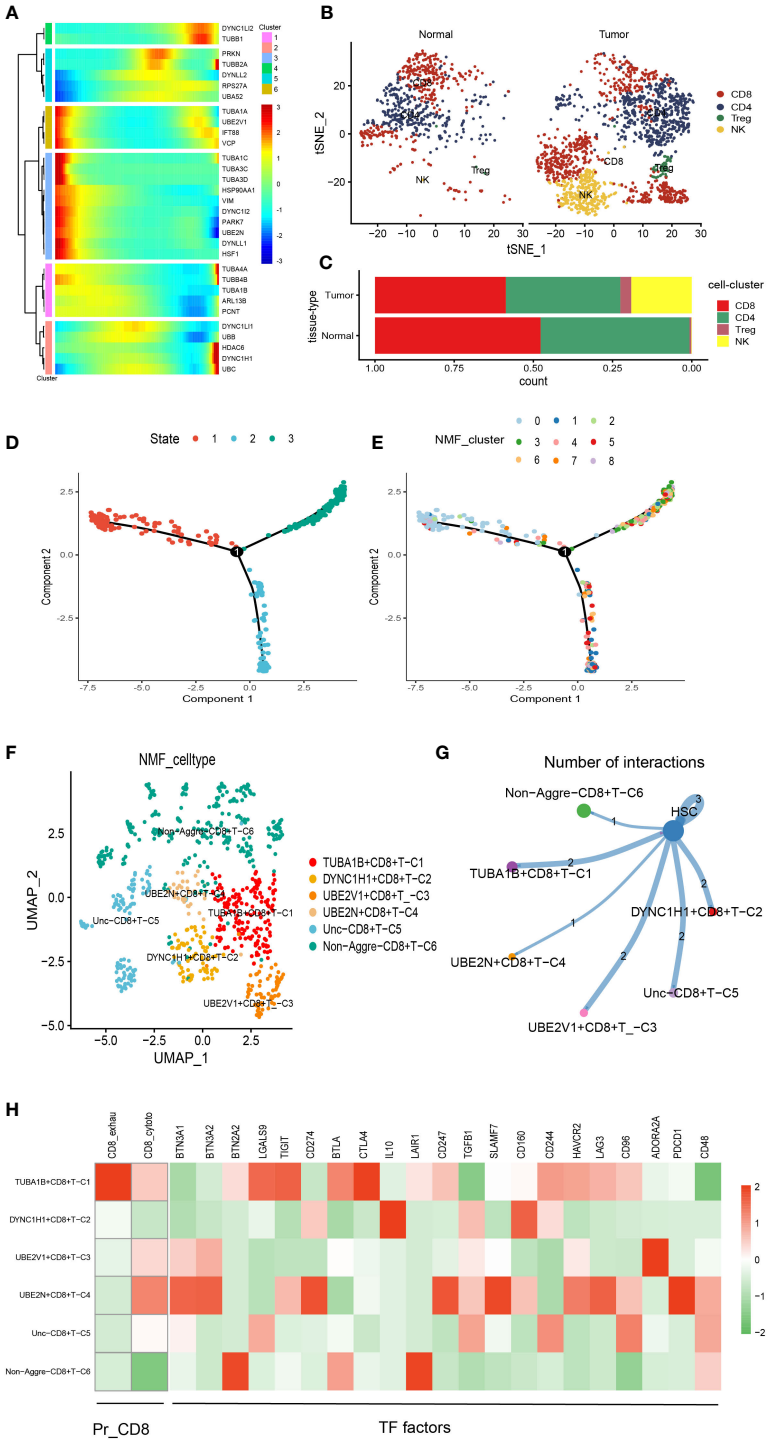


FIGURE 2
Pseudotime analysis for aggrephagy-mediated T cells. **(A)** Pseudotime analysis reveals the role of aggrephagy-related genes for T cells (2397 cells). **(B)** Four subgroups of T cells were obtained by re-clustering annotation based on tSNE analysis. **(C)** Bar plot showing the percentage of the four cell subgroups ($CD8^+$ T cell, $CD4^+$ T cell, NK cell and Treg cell). **(D)** Trajectory color-coded by cell state. **(E)** Trajectory color-coded by NMF cluster. **(F)** The UMAP view and clustering analysis identifying six aggrephagy-related $CD8^+$ T clusters. **(G)** Cell–Cell communications from aggrephagy-related $CD8^+$ T cells to HSC cell. **(H)** Heatmap showing the different average expression of exhausted $CD8^+$ T ($CD8^+$_exhau), cytotoxic $CD8^+$ T ($CD8^+$_cyoto), and TFs among the six aggrephagy-related $CD8^+$ T clusters.

10.3389/fonc.2023.1195392

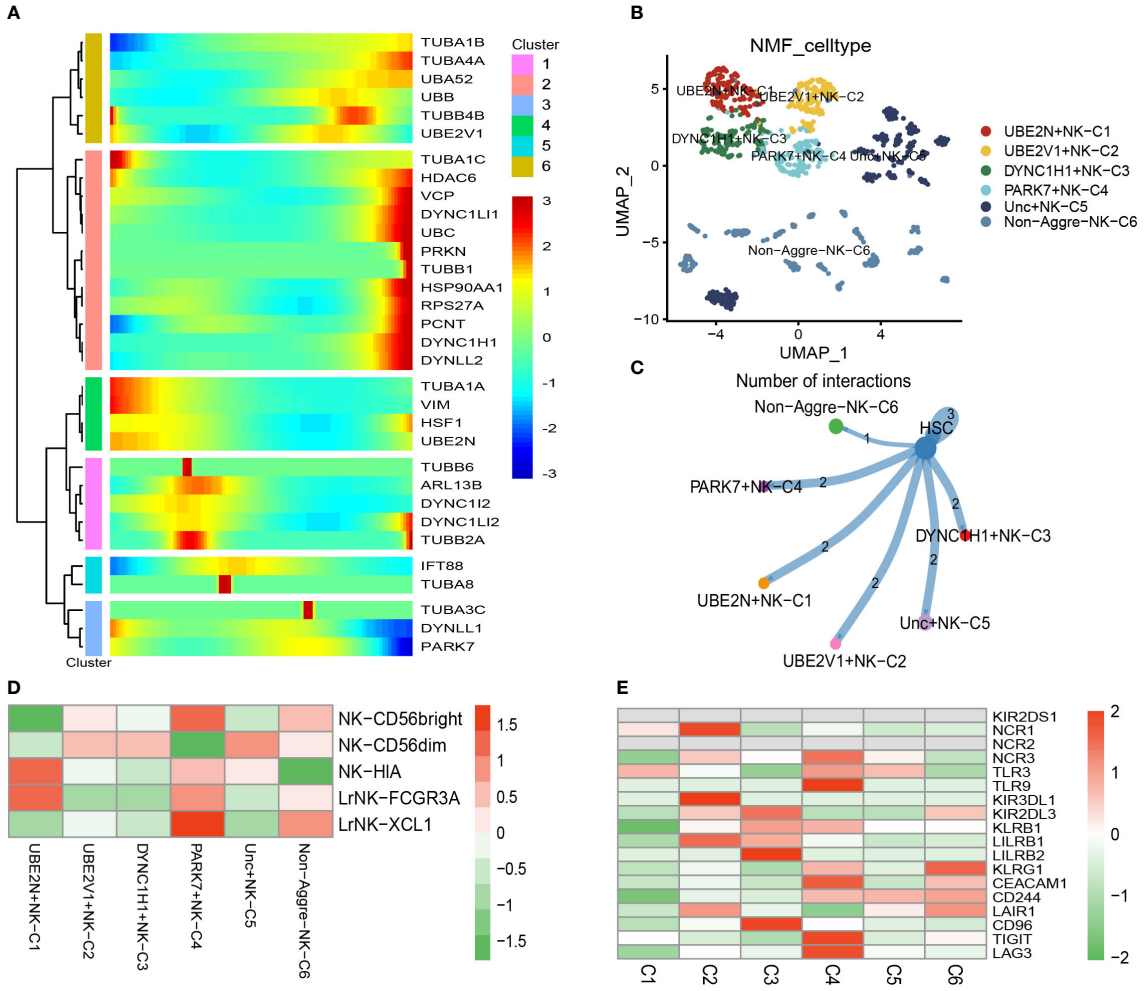


FIGURE 3
Pseudotime analysis for aggrephagy-mediated NK cells. **(A)** Pseudotime analysis reveals the role of aggrephagy-related genes for NK cells (1067 cells). **(B)** The UMAP view and clustering analysis identifying six aggrephagy-related NK clusters. **(C)** Cell–Cell communications from aggrephagy-related NK cells to HSC cell. **(D)** Heatmap showing the different average expression of NK–CD56bright, NK–CD56dim, NK–HIA, LrNK–FCGR3A, and LrNK–XCL1 among the six aggrephagy-related NK clusters. **(E)** Heatmap showing the different average expression of KIR2DS1, NCR1, NCR2, NCR3, TLR3, TLR9, KIR3DL1, KIR2DL3, KLRB1, LILRB1, LILRB2, KLRG1, CEACAM1, CD244, LAIR1, CD96, TIGIT, and LAG3 among the six aggrephagy-related NK clusters.

(Figure 3E), KLRG1 was upregulated in Non−Aggre−NK−C6 (Figure 3E), CEACAM1 was upregulated in PARK7+NK−C4 (Figure 3E), CD96 was upregulated in DYNC1H1+NK−C3 (Figure 3E), TIGIT was upregulated in PARK7+NK−C4 (Figure 3E), LAG3 was upregulated in PARK7+NK−C4 (Figure 3E).

Pseudotime analysis for aggrephagy-mediated myeloid cells

There was a total of 3167 cells in the Myeloid cell type. Based on Myeloid cell type, the PCA analysis was carried out on the highly variable genes, the Elbow algorithm to carry out the Standard Deviation based on the highly variable genes (Supplementary Figure S5). Using UMAP reduction analysis, the 3167 cells in the Myeloid cell type could be clustered into eleven and twelve cell clusters (Supplementary Figure S5). Based on the aggrephagy-related genes, six clusters were identified using Monocle analysis (Figure 4A), including Cluster 1 (VCP, HDAC6, TUBA1A, DYNC1H1, and TUBB1), Cluster 2 (UBC, TUBA4A, DYNC1LI1, VIM, TUBB2A, UBB, HSP90AA1, and TUBB6), Cluster 3 (TUBA3C and TUBA3E), Cluster 4 (PCNT, TUBA1B, UBE2V1, PARK7, TUBB4B, TUBAL3, RPS27A, UBA52, and UBE2N), Cluster 5 (TUBA1C, ARL13B, TUBA8, DYNLL2, DYNC1LI2, and HSF1), Cluster 6 (TUBA4B, IFT88, DYNC1I2, and DYNLL1). Three subgroups of Myeloid cells were obtained by re-clustering annotation using UMAP reduction analysis, including Mono (monocytes) cell, Macrophages (MAC) cell, and Dendritic cell (DC) cell (Figure 4B). Further, we displayed the global view of the expression pattern for marker genes of Mono (monocytes) cell and Macrophages (MAC) cell in Supplementary Figure S5. In addition, the results of UMAP reduction analysis indicated that the NMF cell types were clustered into six aggrephagy-related Mac clusters, including DYNLL1+Mac-C1, UBE2V1+Mac-C2, TUBA1A+Mac-C3, PAPK7+Mac-C4, Unc-

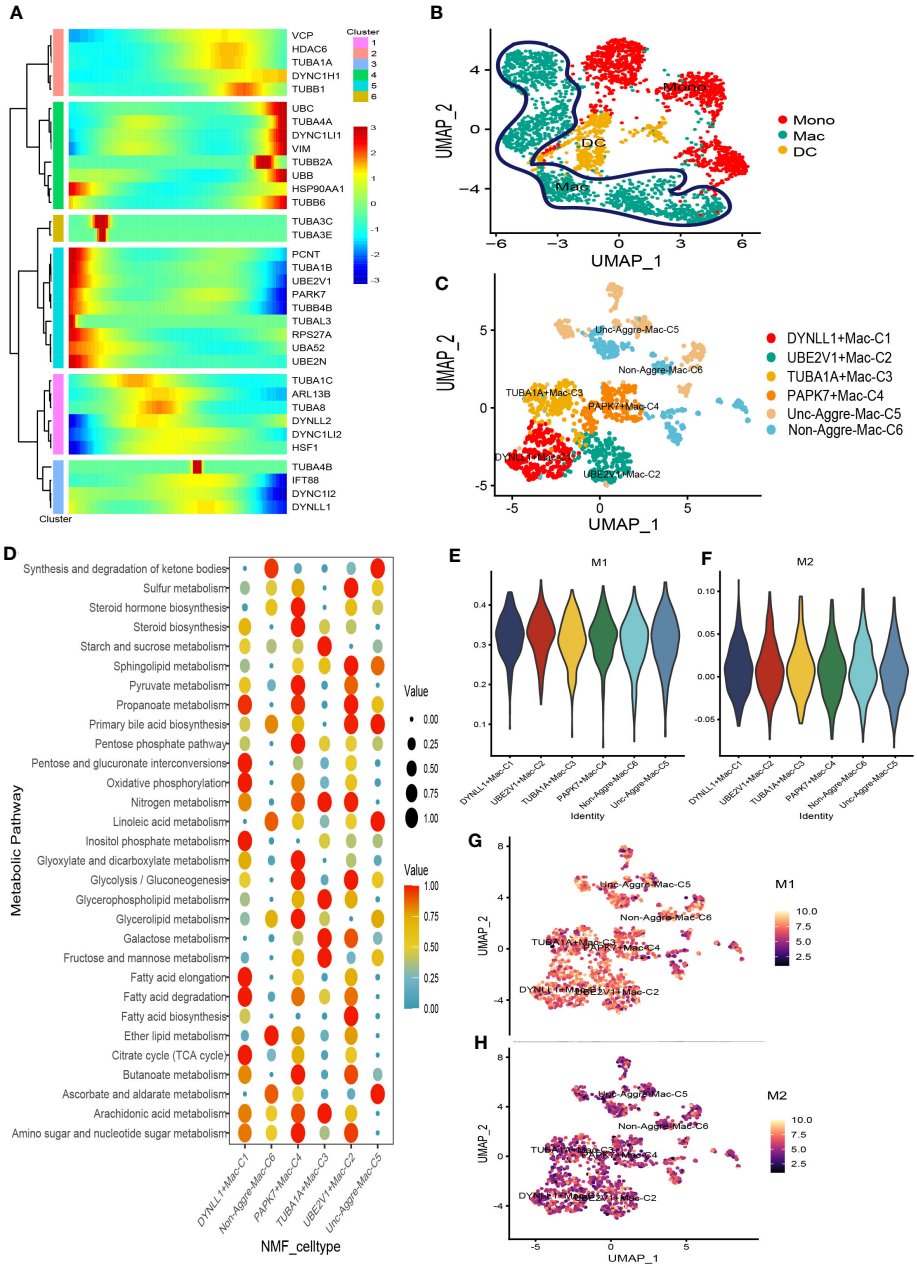


FIGURE 4
Pseudotime analysis for aggrephagy-mediated Myeloid cells. **(A)** Pseudotime analysis reveals the role of aggrephagy-related genes for Myeloid cells (3167 cells). **(B)** Three subgroups of Myeloid cells were obtained by re-clustering annotation based on UMAP analysis. **(C)** The UMAP view and clustering analysis identifying six aggrephagy-related Mac clusters. **(D)** Bubble map showing significantly different activity of metabolic signaling pathway among the six aggrephagy-related Mac clusters. **(E)** The score of the six aggrephagy-related Mac clusters in M1 type macrophages. **(F)** The score of the six aggrephagy-related Mac clusters in M2 type macrophages. **(G)** UMAP plots of the six aggrephagy-related Mac clusters in M1 type macrophages. **(H)** UMAP plots of the six aggrephagy-related Mac clusters in M2 type macrophages.

Aggre-Mac-C5, and Non-Aggre-Mac-C6 (Figure 4C). The number of ligand-receptor links among the aggrephagy-related Mac clusters was computed by Cell-Chat analysis (Supplementary Figure S5). Ulteriorly, we used scMetabolism package to assess the correlation between the aggrephagy-related Mac clusters and metabolic pathways, and we could intuitively see the differences in metabolic pathways of each aggrephagy-related Mac cluster from the bubble map (Figure 4D). To identify M1/M2 type cells, we scored related genes, suggesting that M1 type macrophages were more active in AML (Figures 4E–H).

Prognostic analysis of the aggrephagy-related cell clusters for acute myeloid leukemia

There were 104 normal and acute myeloid leukemia (42 populations and 62 leukemic populations) samples included in GSE63270 dataset, we compared the abundance of UBE2N+NK-C1, UBE2V1+NK-C2, DYNC1H1+NK-C3, PARK7+NK-C4, TUBA1B+CD8$^+$ T–C1, DYNC1H1+CD8$^+$ T–C2, UBE2V1+CD8$^+$

T−C3, UBE2N+CD8$^+$ T−C4, DYNLL1+Mac-C1, UBE2V1+Mac-C2, TUBA1A+Mac-C3, and PAPK7+Mac-C4 between normal and AML samples (Figure 5A). The results indicated that the higher abundance of PARK7+NK-C4, DYNC1H1+CD8$^+$ T−C2, DYNLL1+Mac-C1, and TUBA1A+Mac-C3 were observed in the AML samples, while the higher abundance of TUBA1B+CD8$^+$ T−C1, UBE2V1+CD8$^+$ T−C3, and UBE2V1+Mac-C2 were observed in the normal samples (Figure 5A). Based on the differentially expressed genes (DEGs) generated by the DYNC1H1+CD8+ T−C2, TUBA1A+Mac-C3, and UBE2V1+CD8+ T−C3, the prognostic models were established using TCGA-LAML cohort, the poor prognosis was observed in patients with higher level of DYNC1H1+CD8$^+$ T−C2, lower level of TUBA1A+Mac-C3, and higher level of UBE2V1+CD8$^+$ T−C3 (Figure 5B). GSVA was used for calculating the aggrephagy-related score, the prognosis of the AML patients in the GSE12417 and TCGA-LAML cohorts were further evaluated as displayed in Figure 5C. We found that the survival rates of AML patients in GSE12417 and TCGA-LAML cohorts were significantly different among DYNC1H1+CD8$^+$ T−C2, DYNC1H1+NK-C3, DYNLL1+Mac-C1, PAPK7+Mac-C4, PARK7+NK-C4, TUBA1A+Mac-C3, TUBA1B+CD8$^+$ T−C1, UBE2N+CD8$^+$ T−C4, UBE2N+NK-C1, UBE2V1+CD8$^+$ T−C3, UBE2V1+Mac-C2, and UBE2V1+NK-C2 (Figure 5C).

Immunotherapy analysis of the aggrephagy-related cell clusters for acute myeloid leukemia

We compared the response status (False or True) of immune checkpoint blockade therapy for patients with AML among the aggrephagy-related cell clusters (UBE2N+NK-C1, UBE2V1+NK-C2, DYNC1H1+NK-C3, PARK7+NK-C4, TUBA1B+CD8$^+$ T−C1, DYNC1H1+CD8$^+$ T−C2, UBE2V1+CD8$^+$ T−C3, UBE2N+CD8$^+$

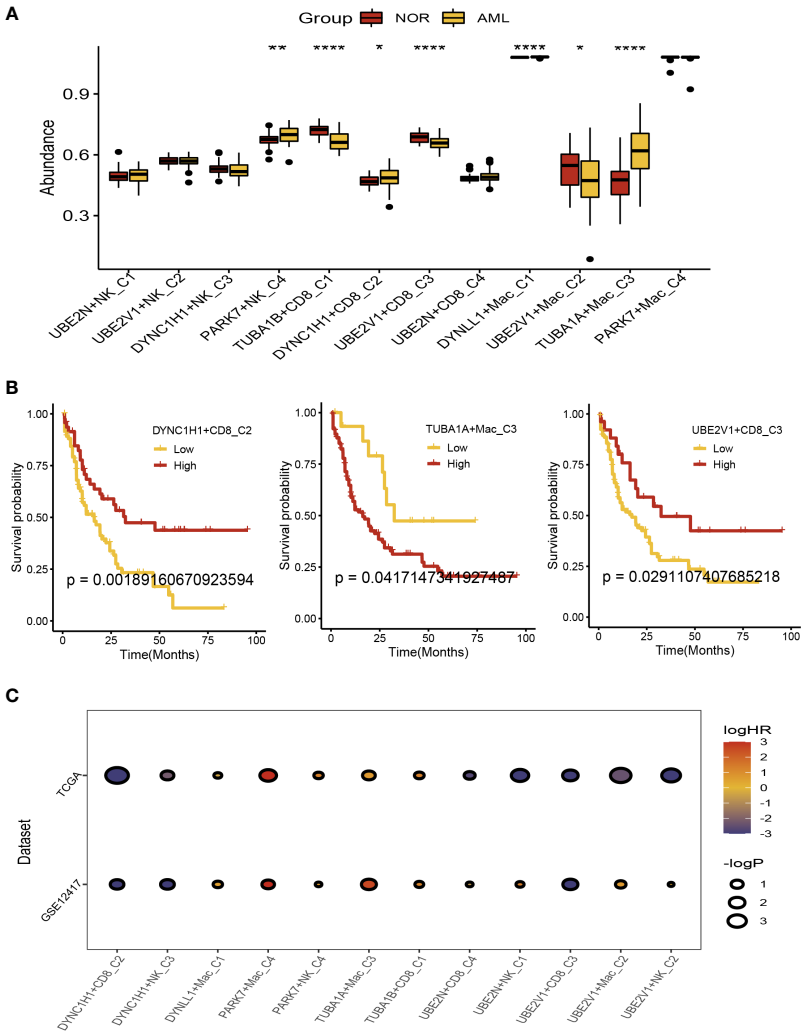


FIGURE 5
Prognostic analysis of the aggrephagy-related cell clusters for acute myeloid leukemia. **(A)** Comparison for the abundance of the aggrephagy-related cell clusters between normal and AML samples from GSE63270 dataset. **(B)** Kaplan-Meier curves for DYNC1H1+CD8$^+$ T−C2, TUBA1A+Mac-C3, and UBE2V1+CD8$^+$ T−C3. **(C)** Comparison for the survival rates of AML patients in GSE12417 and TCGA-LAML cohorts among the aggrephagy-related cell clusters. *P < 0.05; **P < 0.01; ****P < 0.0001.

T−C4, DYNLL1+Mac-C1, UBE2V1+Mac-C2, TUBA1A+Mac-C3, and PAPK7+Mac-C4) using TIDE algorithm (Figure 6A and Supplementary Figure S6). For the AML patients with True response status, the abundance of UBE2N+NK-C1, PARK7+NK-C4, and TUBA1A+Mac-C3 was higher, while the abundance of DYNC1H1+CD8$^+$ T−C2 was lower (Figure 6A). We found that the OR rates of AML patients in GSE12417 and TCGA-LAML cohorts were significantly different among DYNC1H1+CD8$^+$ T−C2, DYNC1H1+NK-C3, DYNLL1+Mac-C1, PAPK7+Mac-C4, PARK7+NK-C4, TUBA1A+Mac-C3, TUBA1B+CD8$^+$ T−C1, UBE2N+CD8$^+$ T−C4, UBE2N+NK-C1, UBE2V1+CD8$^+$ T−C3, UBE2V1+Mac-C2, and UBE2V1+NK-C2 (Figure 6B).

In addition, we compared the response status (SD/PD or CR/PR) of immunotherapy for patients with AML among the aggrephagy-related cell clusters (UBE2N+NK-C1, UBE2V1+NK-C2, DYNC1H1+NK-C3, PARK7+NK-C4, TUBA1B+CD8$^+$ T−C1, DYNC1H1+CD8$^+$ T−C2, UBE2V1+CD8$^+$ T−C3, UBE2N+CD8$^+$ T−C4, DYNLL1+Mac-C1, UBE2V1+Mac-C2, TUBA1A+Mac-C3, and PAPK7+Mac-C4) based on public dataset (Figure 7A). For the AML patients with CR/PR response status, the abundance of TUBA1B+CD8$^+$ T−C1 and DYNLL1+Mac-C1 was higher, while the abundance of TUBA1A+Mac-C3 was lower (Figure 7A). We also found that AML patients with low abundance of TUBA1A+Mac-C3 may have the better prognosis (Figure 7B). In addition, we observed that the expression of TUBA1A was upregulated in bone marrow cells of AML patient both in mRNA (Figure 7C) and protein (Figure 7D) levels.

Discussion

Leukemia is a kind of hematologic malignant disease with hematopoietic stem cell clonal proliferation. Clonal leukemia cells proliferate and accumulate in bone marrow and other normal hematopoietic tissues, inhibit hematopoietic function, and penetrate into other non-hematopoietic tissues and organs through blood circulation, resulting in organ failure and poor prognosis. The clinical manifestations of AML include anemia, bleeding, infection fever and other symptoms. AML is a common type of leukemia, accounting for 80% of acute leukemia, with a high incidence in children (61). Patients with AML tend to die within one year of diagnosis, with a high mortality rate (62). The pathogenesis of AML is complex and diverse, including chemical substances, radioactive substances, genetic factors, gene mutations, abnormal signaling pathways, epigenetic regulation, leukemia microenvironment or immune imbalance. Autophagy is a catabolic process of intracellular substances mediated by lysosome, which has a bidirectional effect in AML. Autophagy can remove abnormal organelles, reduce the accumulation of harmful substances, and effectively prevent cell cancer. However, autophagy can also enable AML cells to obtain various substances and energy, which can help malignant cells to fight against the lack of nutrition and energy caused by their own high metabolism, and promote the growth and proliferation of AML cells. The autophagy levels in different stages of AML were different. How to regulate the

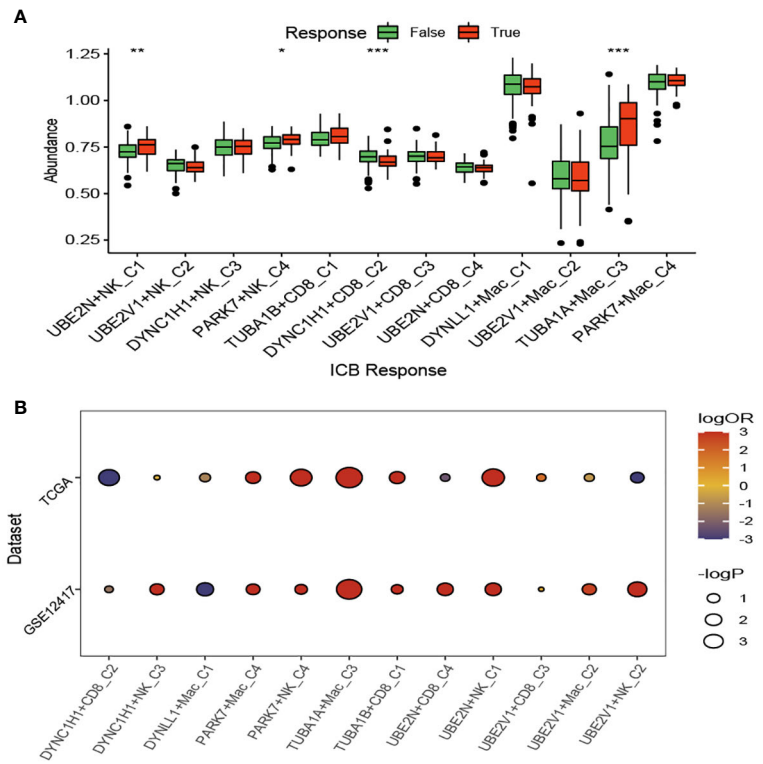


FIGURE 6
Immunotherapy analysis of the aggrephagy-related cell clusters for acute myeloid leukemia based on TIDE algorithm. **(A)** Comparison for the response status of immune checkpoint blockade therapy for patients with AML among the aggrephagy-related cell clusters. **(B)** Comparison for the OR rates of AML patients in GSE12417 and TCGA-LAML cohorts among the aggrephagy-related cell clusters. *P < 0.05; **P < 0.01; ****P < 0.0001.

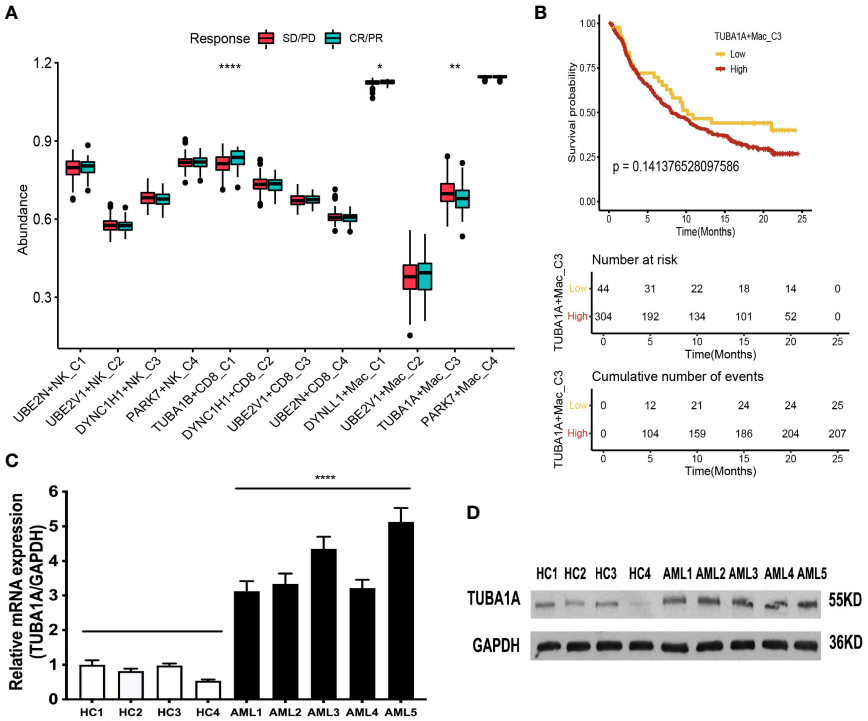


FIGURE 7
Immunotherapy analysis of the aggrephagy-related cell clusters for AML based on public dataset. **(A)** Comparison for the response status of immune checkpoint blockade therapy for patients with AML among the aggrephagy-related cell clusters. **(B)** Kaplan-Meier curve for TUBA1A+Mac-C3. **(C)** The TUBA1A mRNA expression is upregulated in bone marrow cells of AML patient. The levels of TUBA1A mRNA and GAPDH mRNA as control in bone marrow cells of five AML patients and four health people by real-time PCR. Data are expressed as mean ± SD. (**** P <0.0001). **(D)** The TUBA1A protein expression is upregulated in bone marrow cells of AML patient. The levels of TUBA1A protein in bone marrow cells of five AML patients and four health people as control by Western-blot. *P < 0.05; **P < 0.01; ****P < 0.0001.

progression of AML, remove AML cells and improve the therapeutic effect by regulating autophagy level is the focus of AML prevention and treatment.

The survival and apoptosis of immune cells, the expression of immunomodulators and the change of tumor microenvironment (TME) all affect the occurrence and development of AML (25). Immune cells monitor abnormal cells in the body and play an immune effect to eliminate them (63). For example, nature killer (NK) cells recognize and kill tumor cells by mediating cytotoxic effects (64). Tumor cells can evade immune recognition and attack by modifying their own surface antigens and changing the microenvironment around tumor tissue, that is, immune escape of tumor. The occurrence of AML is also closely related to immune escape. By changing the activity of immune cells or regulating the expression of immune molecules, the function of immune cells is affected, thus achieving immune escape of AML cells (65). It can effectively treat AML by inhibiting the cell immune microenvironment and enhancing the immune response (66). To elucidate the relationship between the occurrence of AML and the immune response is of great significance for the development of immunotherapy in patients with AML. In this study, we identified 11 specific cell types, including T cell, Natural Killer (NK) cell, Common Myeloid Progenitor (CMP) cell, Myeloid cell, Granulocyte Monocyte Progenitor (GMP) cell, Megakaryocyte Erythroid Progenitor (MEP) MEP, Promonocyte (Promono) cell, Plasma cell, Hematopoietic Stem Cell (HSC) cell, B cell, and Erythroid cell for AML. Four subgroups of T cells were obtained by re-clustering annotation using t-SNE dimension reduction analysis, including CD8+ T cell, CD4+ T cell, natural killer (NK) cells, and Regulatory T (Treg) cells. NK cell type could be clustered into eleven cell clusters. As for Myeloid cells, three subgroups of Myeloid cells were obtained by re-clustering annotation using UMAP reduction analysis, including Mono (monocytes) cell, Macrophages (MAC) cell, and Dendritic cell (DC) cell. Our study identified some specific cell subtypes of AML, which will provide some reference value for exploring the TME of AML.

Transcription factors (TFs) are involved in the formation of transcription initiation complexes that affect transcription processes and thus downstream gene expression (67). AML contains many abnormal genes, some of which directly affect the expression of TFs, and some indirectly affect the combination of transcription factors and regulatory regions to play a role (68). In addition, some TFs play a role in stem cell maintenance, differentiation and maturation of hematopoietic stem progenitor cells, and abnormal expression of these TFs can lead to hematopoietic malignant transformation. Herein, we found that TFs of LGALS9, TIGIT, BTLA, and CTLA4 were upregulated in the TUBA1B+CD8$^+$ T–C1, TFs of IL10 and CD160 were upregulated in the DYNC1H1+CD8$^+$ T–C2, ADORA2A was upregulated in the UBE2V1+CD8$^+$ T–C3, BTN3A1, BTN3A2, CD274, CD247, SLAMF7, LAG3, and

PDCD1 were upregulated in the UBE2N+CD8$^+$ T–C4, BTN2A2 and LAIR1 were upregulated in the Non–Aggre–CD8$^+$ T–C6. AML is a highly heterogeneous and aggressive hematological malignancy resulting from clonal expansion of malignant hematopoietic progenitor cells in the bone marrow. Its incidence increases with age and its prognosis is poor. A variety of cytogenetic and molecular genetic abnormalities affect signaling pathways, transcription, and epigenetic regulators that induce AML. Studies have shown that various recurrent gene mutations can directly affect the expression of TFs or indirectly change the binding of TFs to regulatory regions, resulting in abnormalities of transcriptional regulatory networks (TRNs), leading to a large number of cloning and proliferation of myeloid precursor cells and stagnating in different stages of hematopoietic differentiation. The fine regulation of TFs such as TIGIT (69), BTLA (70), CTLA4 (71), IL10 (72), CD274 (73) is crucial in hematopoietic regulation and cell fate determination. Abnormal expression of these TFs can interfere with normal hematopoietic differentiation and cause the occurrence of AML. Our study provided new insights into the regulatory mechanisms of related TFs in cell subtypes of AML.

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

Our study revealed the significance of aggrephagy-related patterns in tumor microenvironment, prognosis, and immunotherapy for AML. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by grants from Zhejiang Traditional Chinese Medicine Science and Technology Project (ZYJ23JS03), Natural Science Foundation of Zhejiang Province (LGF22H080009) and Guiding Science and Technology Project of Quzhou (2022014).

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/fonc.2023.1195392/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Dimensionality reduction of single cell for acute myeloid leukemia. **(A, B)** The sequencing depth and the number of genes for single cell from three normal samples and ten patients with acute myeloid leukemia. **(C)** Detection of the highly variable genes across the cells in volcano plot, the top 10 genes were marked out. **(D)** PCA plot of scRNA-seq samples from 13 samples and the Standard Deviation of 1-20 PCs using ElbowPlot algorithm. **(E)** t-SNE and UMAP dimension reduction analysis identifying a total of 18 cell subsets.

SUPPLEMENTARY FIGURE 2
Dot plot showing the average and percentage expression of well-defined marker genes in different cell subsets. The color represented the average expression level of the marker genes. The diameter of the dots denoted the fractional expression.

SUPPLEMENTARY FIGURE 3
UMAP reduction analysis and cell–cell communication analysis for aggrephagy-mediated T cells.

SUPPLEMENTARY FIGURE 4
UMAP reduction analysis and cell–cell communication analysis for aggrephagy-mediated NK cells.

SUPPLEMENTARY FIGURE 5
UMAP reduction analysis and cell–cell communication analysis for aggrephagy-mediated Myeloid cells.

SUPPLEMENTARY FIGURE 6
Immunotherapy analysis of the aggrephagy-related cell clusters for acute myeloid leukemia using TIDE algorithm.

References

1. Juliusson G, Hough R. Leukemia. *Prog Tumor Res* (2016) 43:87–100. doi: 10.1159/000447076

2. Butturini A, Gale RP. Oncogenes and leukemia. *Leukemia* (1990) 4:138–60.

3. 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

4. Ye Z, Wang S. Mixed phenotype acute leukemia. *Chin Med J (Engl)* (2014) 127:2999–3003.

5. Scarfò L, Ferreri AJ, Ghia P. Chronic lymphocytic leukaemia. *Crit Rev Oncol Hematol* (2016) 104:169–82. doi: 10.1016/j.critrevonc.2016.06.003

6. Brunning RD. Classification of acute leukemias. *Semin Diagn Pathol* (2003) 20:142–53. doi: 10.1016/s0740-2570(03)00031-5

7. Newell LF, Cook RJ. Advances in acute myeloid leukemia. *Bmj* (2021) 375:n2026. doi: 10.1136/bmj.n2026

8. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR, et al. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br J Haematol* (1976) 33:451–8. doi: 10.1111/j.1365-2141.1976.tb03563.x

9. Parra M, Baptista MJ, Genescà E, Llinàs-Arias P, Esteller M. Genetics and epigenetics of leukemia and lymphoma: from knowledge to applications, meeting report of the josep carreras leukaemia research institute. *Hematol Oncol* (2020) 38:432–8. doi: 10.1002/hon.2725

10. Perincheri S. Tumor microenvironment of lymphomas and plasma cell neoplasms: broad overview and impact on evaluation for immune based therapies. *Front Oncol* (2021) 11:719140. doi: 10.3389/fonc.2021.719140

11. Pandey PR, Young KH, Kumar D, Jain N. RNA-Mediated immunotherapy regulating tumor immune microenvironment: next wave of cancer therapeutics. *Mol Cancer* (2022) 21:58. doi: 10.1186/s12943-022-01528-6

12. Ghoshdastider U, Rohatgi N, Mojtabavi Naeini M, Baruah P, Revkov E, Guo YA, et al. Pan-cancer analysis of ligand-receptor cross-talk in the tumor microenvironment. *Cancer Res* (2021) 81:1802–12. doi: 10.1158/0008-5472.Can-20-2352

13. Fu T, Dai LJ, Wu SY, Xiao Y, Ma D, Jiang YZ, et al. Spatial architecture of the immune microenvironment orchestrates tumor immunity and therapeutic response. *J Hematol Oncol* (2021) 14:98. doi: 10.1186/s13045-021-01103-4

14. Bejarano L, Jordão MJC, Joyce JA. Therapeutic targeting of the tumor microenvironment. *Cancer Discov* (2021) 11:933–59. doi: 10.1158/2159-8290.Cd-20-1808

15. 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

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

17. Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp Mol Med* (2018) 50:1–14. doi: 10.1038/s12276-018-0071-8

18. Whitmarsh-Everiss T, Laraia L. Small molecule probes for targeting autophagy. *Nat Chem Biol* (2021) 17:653–64. doi: 10.1038/s41589-021-00768-9

19. Chmurska A, Matczak K, Marczak A. Two faces of autophagy in the struggle against cancer. *Int J Mol Sci* (2021) 22:2981. doi: 10.3390/ijms22062981

20. Du W, Xu A, Huang Y, Cao J, Zhu H, Yang B, et al. The role of autophagy in targeted therapy for acute myeloid leukemia. *Autophagy* (2021) 17:2665–79. doi: 10.1080/15548627.2020.1822628

21. Hyttinen JM, Amadio M, Viiri J, Pascale A, Salminen A, Kaarniranta K. Clearance of misfolded and aggregated proteins by aggrephagy and implications for aggregation diseases. *Ageing Res Rev* (2014) 18:16–28. doi: 10.1016/j.arr.2014.07.002

22. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res* (2013) 41:D991–995. doi: 10.1093/nar/gks1193

23. Li F, Cai J, Liu J, Yu SC, Zhang X, Su Y, et al. Construction of a solid cox model for AML patients based on multiomics bioinformatic analysis. *Front Oncol* (2022) 12:925615. doi: 10.3389/fonc.2022.925615

24. Dai C, Chen M, Wang C, Hao X. Deconvolution of bulk gene expression profiles with single-cell transcriptomics to develop a cell type composition-based prognostic model for acute myeloid leukemia. *Front Cell Dev Biol* (2021) 9:762260. doi: 10.3389/fcell.2021.762260

25. 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

26. Jung N, Dai B, Gentles AJ, Majeti R, Feinberg AP. An LSC epigenetic signature is largely mutation independent and implicates the HOXA cluster in AML pathogenesis. *Nat Commun* (2015) 6:8489. doi: 10.1038/ncomms9489

27. 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

28. Wang YH, Lin CC, Hsu CL, Hung SY, Yao CY, Lee SH, et al. Distinct clinical and biological characteristics of acute myeloid leukemia with higher expression of long noncoding RNA KIAA0125. *Ann Hematol* (2021) 100:487–98. doi: 10.1007/s00277-020-04358-y

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. Tomczak K, Czerwińska P, Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)* (2015) 19:A68–77. doi: 10.5114/wo.2014.47136

31. Slovin S, Carissimo A, Panariello F, Grimaldi A, Bouché V, Gambardella G, et al. Single-cell RNA sequencing analysis: a step-by-Step overview. *Methods Mol Biol* (2021) 2284:343–65. doi: 10.1007/978-1-0716-1307-8_19

32. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, et al. Comprehensive integration of single-cell data. *Cell* (2019) 177:1888–1902.e1821. doi: 10.1016/j.cell.2019.05.031

33. Zheng M, Hu Y, Liu O, Li S, Wang Y, Li X, et al. Oxidative stress response biomarkers of ovarian cancer based on single-cell and bulk RNA sequencing. *Oxid Med Cell Longev* (2023) 2023:1261039. doi: 10.1155/2023/1261039

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

35. Chen YP, Yin JH, Li WF, Li HJ, Chen DP, Zhang CJ, et al. Single-cell transcriptomics reveals regulators underlying immune cell diversity and immune subtypes associated with prognosis in nasopharyngeal carcinoma. *Cell Res* (2020) 30:1024–42. doi: 10.1038/s41422-020-0374-x

36. 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.e1624. doi: 10.1016/j.cell.2017.10.044

37. Liu H, Zhao R, Qin R, Sun H, Huang Q, Liu L, et al. Panoramic comparison between NK cells in healthy and cancerous liver through single-cell RNA sequencing. *Cancer Biol Med* (2022) 19:1334–51. doi: 10.20892/j.issn.2095-3941.2022.0050

38. Kumar N, Mishra B, Athar M, Mukhtar S. Inference of gene regulatory network from single-cell transcriptomic data using pySCENIC. *Methods Mol Biol* (2021) 2328:171–82. doi: 10.1007/978-1-0716-1534-8_10

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

40. Zhou W, Bai Y, Chen J, Li H, Zhang B, Liu H. Revealing the critical regulators of modulated smooth muscle cells in atherosclerosis in mice. *Front Genet* (2022) 13:900358. doi: 10.3389/fgene.2022.900358

41. Schmitt P, Sorin B, Frouté T, Parisot N, Calevro F, Peignier S. GReNaDIne: a data-driven Python library to infer gene regulatory networks from gene expression data. *Genes (Basel)* (2023) 14:269. doi: 10.3390/genes14020269

42. Wang Y, Li B, Zhao Y. Inflammation in preeclampsia: genetic biomarkers, mechanisms, and therapeutic strategies. *Front Immunol* (2022) 13:883404. doi: 10.3389/fimmu.2022.883404

43. Huang P, Tang L, Zhang L, Ren Y, Peng H, Xiao Y, et al. Identification of biomarkers associated with CD4(+) T-cell infiltration with gene coexpression network in dermatomyositis. *Front Immunol* (2022) 13:854848. doi: 10.3389/fimmu.2022.854848

44. 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

45. Lin J, Cai Y, Wang Z, Ma Y, Pan J, Liu Y, et al. Novel biomarkers predict prognosis and drug-induced neuroendocrine differentiation in patients with prostate cancer. *Front Endocrinol (Lausanne)* (2022) 13:1005916. doi: 10.3389/fendo.2022.1005916

46. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf* (2013) 14:7. doi: 10.1186/1471-2105-14-7

47. Fisher LD, Lin DY. Time-dependent covariates in the cox proportional-hazards regression model. *Annu Rev Public Health* (1999) 20:145–57. doi: 10.1146/annurev.publhealth.20.1.145

48. Fu J, Li K, Zhang W, Wan C, Zhang J, Jiang P, et al. Large-Scale public data reuse to model immunotherapy response and resistance. *Genome Med* (2020) 12:21. doi: 10.1186/s13073-020-0721-z

49. Ulloa-Montoya F, Louahed J, Dizier B, Gruselle O, Spiessens B, Lehmann FF, et al. Predictive gene signature in MAGE-A3 antigen-specific cancer immunotherapy. *J Clin Oncol* (2013) 31:2388–95. doi: 10.1200/jco.2012.44.3762

50. Gide TN, Quek C, Menzies AM, Tasker AT, Shang P, Holst J, et al. Distinct immune cell populations define response to anti-PD-1 monotherapy and anti-PD-1/Anti-CTLA-4 combined therapy. *Cancer Cell* (2019) 35:238–255.e236. doi: 10.1016/j.ccell.2019.01.003

51. Nathanson T, Ahuja A, Rubinsteyn A, Aksoy BA, Hellmann MD, Miao D, et al. Somatic mutations and neoepitope homology in melanomas treated with CTLA-4 blockade. *Cancer Immunol Res* (2017) 5:84–91. doi: 10.1158/2326-6066.Cir-16-0019

52. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* (2017) 168:542. doi: 10.1016/j.cell.2017.01.010

53. Lauss M, Donia M, Harbst K, Andersen R, Mitra S, Rosengren F, et al. Mutational and putative neoantigen load predict clinical benefit of adoptive T cell therapy in melanoma. *Nat Commun* (2017) 8:1738. doi: 10.1038/s41467-017-01460-0

54. Liu D, Schilling B, Liu D, Sucker A, Livingstone E, Jerby-Arnon L, et al. Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nat Med* (2019) 25:1916–27. doi: 10.1038/s41591-019-0654-5

55. 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.e916. doi: 10.1016/j.cell.2017.09.028

56. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* (2015) 350:207–11. doi: 10.1126/science.aad0095

57. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* (2018) 554:544–8. doi: 10.1038/nature25501

58. 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

59. Cho JW, Hong MH, Ha SJ, Kim YJ, Cho BC, Lee I, et al. Genome-wide identification of differentially methylated promoters and enhancers associated with response to anti-PD-1 therapy in non-small cell lung cancer. *Exp Mol Med* (2020) 52:1550–63. doi: 10.1038/s12276-020-00493-8

60. 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

61. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* (2020) 70:7–30. doi: 10.3322/caac.21590

62. 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

63. 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

64. Terrén I, Orrantia A, Vitallé J, Zenarruzabeitia O, Borrego F. NK cell metabolism and tumor microenvironment. *Front Immunol* (2019) 10:2278. doi: 10.3389/fimmu.2019.02278

65. Vago L, Gojo I. Immune escape and immunotherapy of acute myeloid leukemia. *J Clin Invest* (2020) 130:1552–64. doi: 10.1172/jci129204

66. Isidori A, Salvestrini V, Ciciarello M, Loscocco F, Visani G, Parisi S, et al. The role of the immunosuppressive microenvironment in acute myeloid leukemia development and treatment. *Expert Rev Hematol* (2014) 7:807–18. doi: 10.1586/17474086.2014.958464

67. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The human transcription factors. *Cell* (2018) 172:650–65. doi: 10.1016/j.cell.2018.01.029

68. Takei H, Kobayashi SS. Targeting transcription factors in acute myeloid leukemia. *Int J Hematol* (2019) 109:28–34. doi: 10.1007/s12185-018-2488-1

69. Gournay V, Vallet N, Peux V, Vera K, Bordenave J, Lambert M, et al. Immune landscape after allo-HSCT: TIGIT- and CD161-expressing CD4 T cells are associated with subsequent leukemia relapse. *Blood* (2022) 140:1305–21. doi: 10.1182/blood.2022015522

70. Knaus HA, Kanakry CG, Luznik L, Gojo I. Immunomodulatory drugs: immune checkpoint agents in acute leukemia. *Curr Drug Targets* (2017) 18:315–31. doi: 10.2174/1389450116666150518095346

71. Daver N, Alotaibi AS, Bücklein V, Subklewe M. T-Cell-based immunotherapy of acute myeloid leukemia: current concepts and future developments. *Leukemia* (2021) 35:1843–63. doi: 10.1038/s41375-021-01253-x

72. Binder S, Luciano M, Horejs-Hoeck J. The cytokine network in acute myeloid leukemia (AML): a focus on pro- and anti-inflammatory mediators. *Cytokine Growth Factor Rev* (2018) 43:8–15. doi: 10.1016/j.cytogfr.2018.08.004

73. Bewersdorf JP, Shallis RM, Zeidan AM. Immune checkpoint inhibition in myeloid malignancies: moving beyond the PD-1/PD-L1 and CTLA-4 pathways. *Blood Rev* (2021) 45:100709. doi: 10.1016/j.blre.2020.100709

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 21 August 2023
DOI 10.3389/fimmu.2023.1183230

OPEN ACCESS

EDITED BY
Jun Liu,
Yuebei People's Hospital, China

REVIEWED BY
Bufu Tang,
Fudan University, China
João Pessoa,
University of Coimbra, Portugal
Yanwei Luo,
Central South University, China
Fei Han,
Chongqing Medical University, China

*CORRESPONDENCE
Yi Huang
Huangyi@fjsl.com.cn

†These authors have contributed equally to this work

RECEIVED 09 March 2023
ACCEPTED 04 August 2023
PUBLISHED 21 August 2023

CITATION
Zhang Y, Wang Y, Chen J, Xia Y and Huang Y (2023) A programmed cell death-related model based on machine learning for predicting prognosis and immunotherapy responses in patients with lung adenocarcinoma. *Front. Immunol.* 14:1183230. doi: 10.3389/fimmu.2023.1183230

A programmed cell death-related model based on machine learning for predicting prognosis and immunotherapy responses in patients with lung adenocarcinoma

Yi Zhang[1,2†], Yuzhi Wang[3†], Jianlin Chen[1,2], Yu Xia[4] and Yi Huang[1,2,5,6*]

[1]Shengli Clinical Medical College of Fujian Medical University, Fujian Medical University, Fuzhou, Fujian, China, [2]Department of Clinical Laboratory, Fujian Provincial Hospital, Fuzhou, China, [3]Department of Laboratory Medicine, Deyang People's Hospital, Deyang, Sichuan, China, [4]Integrated Chinese and Western Medicine College, Fujian University of Traditional Chinese Medicine, Fuzhou, Fujian, China, [5]Central Laboratory, Center for Experimental Research in Clinical Medicine, Fujian Provincial Hospital, Fuzhou, China, [6]Fujian Provincial Key Laboratory of Critical Care Medicine, Fujian Provincial Key Laboratory of Cardiovascular Disease, Fuzhou, China

Background: lung adenocarcinoma (LUAD) remains one of the most common and lethal malignancies with poor prognosis. Programmed cell death (PCD) is an evolutionarily conserved cell suicide process that regulates tumorigenesis, progression, and metastasis of cancer cells. However, a comprehensive analysis of the role of PCD in LUAD is still unavailable.

Methods: We analyzed multi-omic variations in PCD-related genes (PCDRGs) for LUAD. We used cross-validation of 10 machine learning algorithms (101 combinations) to synthetically develop and validate an optimal prognostic cell death score (CDS) model based on the PCDRGs expression profile. Patients were classified based on their median CDS values into the high and low-CDS groups. Next, we compared the differences in the genomics, biological functions, and tumor microenvironment of patients between both groups. In addition, we assessed the ability of CDS for predicting the response of patients from the immunotherapy cohort to immunotherapy. Finally, functional validation of key genes in CDS was performed.

Results: We constructed CDS based on four PCDRGs, which could effectively and consistently stratify patients with LUAD (patients with high CDS had poor prognoses). The performance of our CDS was superior compared to 77 LUAD signatures that have been previously published. The results revealed significant genetic alterations like mutation count, TMB, and CNV were observed in patients with high CDS. Furthermore, we observed an association of CDS with immune cell infiltration, microsatellite instability, SNV neoantigens. The immune status of patients with low CDS was more active. In addition, CDS could be reliable to predict therapeutic response in multiple immunotherapy cohorts. *In vitro*

experiments revealed that high DNA damage inducible transcript 4 (*DDIT4*) expression in LUAD cells mediated protumor effects.

Conclusion: CDS was constructed based on PCDRGs using machine learning. This model could accurately predict patients' prognoses and their responses to therapy. These results provide new promising tools for clinical management and aid in designing personalized treatment strategies for patients with LUAD.

KEYWORDS

programmed cell death, lung adenocarcinoma, machine learning, prognosis, tumor microenvironment

Introduction

Globally, lung cancer (LC) accounts for approximately 18% of all cancer-related mortalities and is also the leading cause of cancer-related mortalities in both sexes (1). Non-small cell LC (NSCLC) accounts for 90% of LC cases. NSCLC can be further categorized based on histology into two subtypes: lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC). Of these, LUAD cases are more common (2). The factors underlying the poor prognosis of patients with mid to late-stage LUAD include the lack of symptoms and tumor specificity at an early stage, local infiltration, and distant metastases of cancer (3). Rapid advancements in biotechnology and precision medicine have helped develop targeted drugs and therapeutic approaches specific to patients with LUAD. Further biomarkers for LUAD, like *EGFR*, *E17K*, and *PTEN*, have been identified (4–6), which are currently used in combination with surgical resection, radio, and chemotherapies. However, only a small proportion of patients with LUAD have benefitted from these advancements and improvements in therapeutic efficacy. No significant improvement in the overall survival (OS) and progression-free survival of patients has been observed (7, 8). Therefore, an in-depth understanding of the underlying mechanisms of LUAD and identifying new biomarkers is crucial for predicting the prognoses and designing personalized therapeutic strategies for patients with LUAD.

Programmed cell death (PCD) is a crucial process for the growth and development of living organisms. Studies have shown that apoptosis, pyroptosis, ferroptosis, autophagy, necroptosis, cuproptosis, parthanatos, entotic and lysosome-dependent cell death, Alkaliptosis, NETosis, and oxeiptosis-related PCDs are classical cell death pathways (9). Apoptosis is a non-inflammatory response to PCD characterized by the activation of caspases, leading to the contraction of cells, coalescence, and the nucleus, as well as nucleosomal DNA fragmentation (10). Apoptosis is required for the maintenance of the cell death-cell survival balance. Furthermore, abnormal apoptosis escape is a characteristic of cancer cells (11). Pyroptosis is programmed necrosis of cells induced by inflammatory vesicles, wherein activated Gasdermin protein (a scorching substrate for inflammatory caspases-1/4/5/11) forms pores in the plasma membrane, thereby leading to cell death (12, 13). In 2012, ferroptosis was discovered as a novel iron-dependent PCD characterized by its ability to disrupt the redox homeostasis of cells and the absence of apoptosis (14). During ferroptosis, the cytoplasm appears round and detached, the mitochondrial membranes are condensed, the number of mitochondrial cristae is reduced or absent, and the outer mitochondrial membranes are ruptured (15). Autophagy is an apoptosis-independent cell death form. It is characterized by no chromatin condensation, the accumulation of autophagic vacuole, and autophagosome formation, which fuses with lysosomes to form autolysosomes in the cytoplasm (16, 17). Unlike apoptosis, necroptosis destabilizes cell membranes, and cause swelling and lysis of cells, thereby leading to the release of cellular components (18). The inactivation or deletion of caspases-8 and RIPK1 and RIPK3 activation, as well as autophosphorylation, induces necroptosis of cells (19). During necroptosis, the cell membranes rupture and release cellular contents, thereby activating immune responses (20). In March 2022, a study by Peter et al. introduced a new mode of cell death called cuproptosis (21). Unlike other forms of cell death, copper toxicity occurs primarily through the direct binding of cuproptosis to the fatty acylated components of the Krebs cycle. This leads to fatty acylated protein accumulation and iron-sulfur cluster protein loss, increase in proteotoxicity, which culminates in cell death (22). Parthantos is characterized by an increase in the activation of PARP-1 (23), PAR aggregates, and the translocation of apoptosis-inducing factors from the mitochondria to the nucleus (24). Unlike pyroptosis, parthanatos is independent of caspase and is triggered by an excessive reactive oxygen species (ROS) response (25). A study has shown that parthanatos induces mitochondrial membrane dissipation and the condensation of extensively fragmented DNA chromatin (26). Entotic cell death is the byproduct of endocytosis, forms typical intercellular structures, and is caused by the disassociation of cells from the basement membrane. It primarily occurs in epithelial cells and carcinomas (27). The entry of epithelial cells into other cells can eliminate endosomal cells by specific autophagy-related processes regulating the lysosomal degradation of cells (28). NETosis is a type of neutrophils, granulocytes, or macrophage-related necrosis. During NETosis, the granular contents of neutrophils are transferred to the nucleus, which causes the decondensation of chromatins, and

induces the formation of a neutrophil extracellular trap (29). Lysosomal membrane permeabilization (LMP) is the primary cause of lysosome-dependent cell death, characterized by the loss of the lysosomal membrane integrity, thereby releasing the contents of lysosomes into the cytosol (30). LMP-mediated cell death is either dependent or independent of caspases (31). Alkalinization in cells induces a novel mode of PCD called alkaliptosis (32). The oxygen radicals trigger a novel form of regulated cell death called oxeiptosis, which is independent of caspases, and driven by the KEAP1-PGAM5-AIFM1 pathway activation (33). In organisms, PCD eliminates harmful or redundant cells and maintains tissue homeostasis. During PCD, damage-associated molecular patterns are released, which act as a powerful stimulus for activating local inflammatory or systemic immune responses. Therefore, selective activation of the PCD pathway could be a novel strategy for preventing and treating patients with LUAD. A study has shown that A549 cells treated with chemotherapeutic drugs such as cisplatin and paclitaxel trigger pyroptosis *via* the caspase 3/Gasdermin E pathway. The efficacy of these drugs to stimulate pyroptosis depends on the expression of Gasdermin E (34). CD8+ T cells secrete IFNs, which reduce SLC7A11 and SLC3A2 expression, thereby preventing the uptake of cystine by LUAD cells and promoting ferroptosis as well as lipid peroxidation. Together, this enhances the efficacy of immunotherapy. Hydroxychloroquine inhibits LUAD cell autophagy, thereby reversing chemoresistance in advanced-stage LUAD (35). Thus, escaping multiple types of PCD is a hallmark of LUAD. Therefore, a comprehensive understanding of the underlying mechanism of pan-PCD in LUAD could aid in mitigating tumorigenesis, cancer progression, and drug resistance in LUAD.

Previous studies on PCD have determined the involvement of a single mode of cell death in LUAD. However, several modes of PCDs mediate tumorigenesis, progression, and metastasis of cancer cells. Moreover, no studies have analyzed the involvement of PCD in LUAD systematically. In this study, we investigated the alterations of PCD-related genes (PCDRGs) in LUAD. We used a computational framework to construct and validate a novel cell death score (CDS) based on PCDRGs. CDS can accurately stratify patients with LUAD based on their prognostic status. Next, we investigated the differences in genetic mutations, tumor microenvironment (TME), and biological characteristics of patients between both CDS groups. Furthermore, we determined the efficiency of CDS in predicting the patients' responses to immunotherapy and screened suitable drugs for patients with LUAD in different CDS groups. Finally, we determined the roles of DNA damage inducible transcript 4 (*DDIT4*) in LUAD.

Materials and methods

Cohort and preprocessing

The transcriptional and clinical data of patients with LUAD were downloaded from the Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) and Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo) databases. In addition, the data on gene mutation and copy number variation (CNV) were obtained from TCGA database. Next, we merged four cohorts, and the batch effect was eliminated using the "Combat" algorithm. The TCGA-LUAD cohort was used as the training cohort for constructing the CDS. The GSE31210, GSE68465, and GSE72094 cohorts from GEO were used as independent validation cohorts. We excluded patients whose OS information was unavailable. Finally, we included 1569 patients with LUAD for the subsequent analysis. Supplementary Table 1 shows detailed information on the patients. The study flowchart is depicted in Figure 1.

The expression and variations of PCDRGs in LUAD

The key regulatory genes of these 12 types of PCD were included as PCDRGs. These genes were collected and compiled from the Gene Set Enrichment Analysis (GSEA), Kyoto Encyclopedia of Genes and Genomes (KEGG), and previously published gene sets (9) (Supplementary Table 2). The differentially expressed PCDRGs (DEPCDRGs) in LUAD and paracancerous tissues of patients in the TCGA-LUAD cohort were screened using the "limma" R package. The threshold for screening DEPCDRGs was "$P < 0.05$" and "|log2 Fold change (FC)| > 1". The functional enrichment analysis was performed to identify functions and pathways enriched by DEPCDRGs using the "clusterprofiler" R package. Next, the "maftools" package was employed to explore the somatic mutations in DEPCDRGs (36). The frequencies of "Gain" or "Loss" CNV in DEPCDRGs were screened and calculated. Finally, the chromosomal location of CNV in patients was visualized as the circus plot with the aid of the "circlize" R package (37). The transcription factors (TFs) within the DEPCDRGs were predicted using Transcriptional Regulatory Relationships Unraveled by Sentence-based Text mining (TRRUST, www.grnpedia.org/trrust/). Subsequently, TF-gene interaction pairs exhibiting P-values <0.05 were carefully chosen to construct the regulatory network through the utilization of Cytoscape.

PCDRGs signature generated by machine learning-based integrated approach

The prognosis-related DEPCDRGs were screened using the univariate Cox regression analysis. The threshold set to avoid omission was "$P < 0.05$". A PCDRGs signature was constructed with high accuracy and stability using 10 machine learning algorithms, including "Least Absolute Shrinkage and Selection Operator", "Ridge", "Elastic network", "StepCox", "Survival support vector machine (survival-SVM)", "CoxBoost", "Supervised principal components", "partial least squares regression for COX", "random survival forest (RSF)", "generalized boosted regression modeling" to construct. These 10 machine learning algorithms were used to cross-validate 101 combinations of the "leave-one-out-cross validation (LOOCV)" framework for constructing the models based on the TCGA-LUAD cohort and

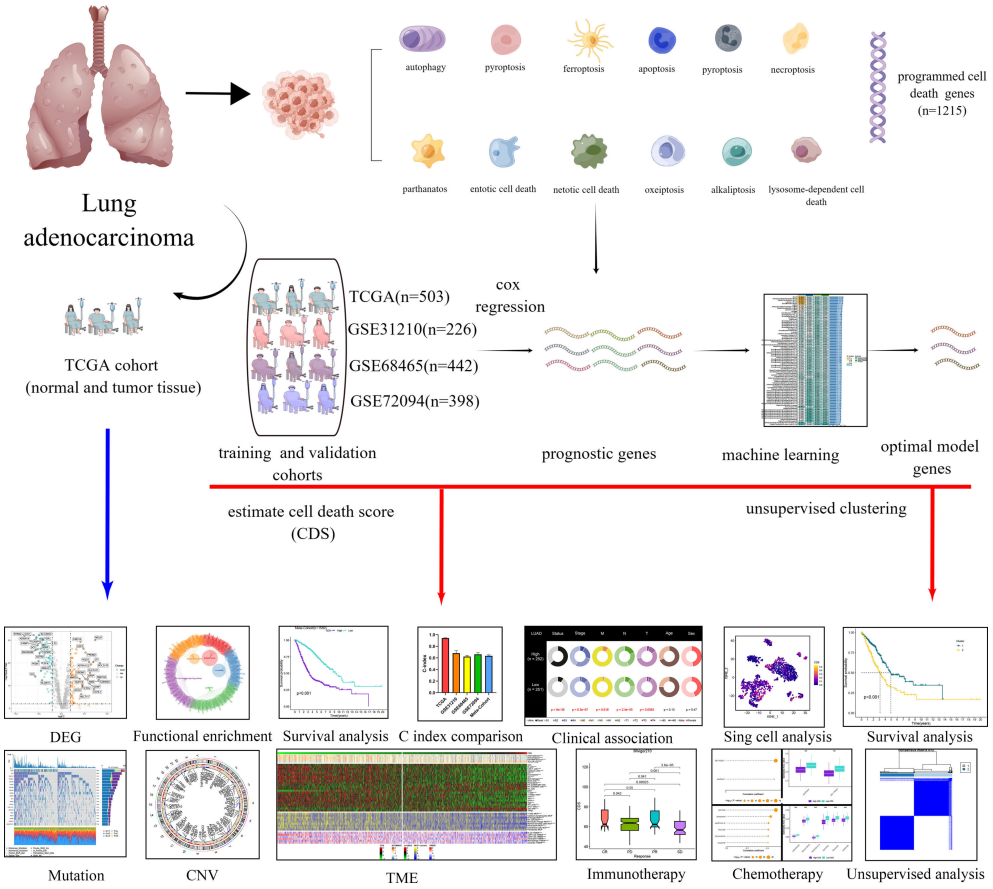


FIGURE 1
Diagram of analytic workflow in this study. The Diagram was drawn from the FIGUREdraw. (https://www.FIGUREdraw.com/static/index.html).

were subsequently validated in the GEO cohorts. The models with < 3 genes were excluded. Additionally, we calculated the C-index of each model in all cohorts. CDS was the optimal model with the highest mean C-index. All patients were classified using the median CDS value set as a threshold into the high and low-CDS groups.

Consensus clustering

Consensus clustering was performed using the "Consensus ClusterPlus" package based on the expression of PCDRGs in the CDS. The clustering was based on dividing centromeres with "Euclidean" distances. Finally, patients with LUAD were classified into two subtypes based on the best classification of "K=2-9".

Mutation and CNV characteristics

The mutation profiles and types of the top 20 genes with the highest mutation frequencies in patients in both CDS groups were mapped using the "maftools" package. Tumor mutational burden (TMB) is the total number of non-synonymous mutations in all exomes of patients and is calculated according to the number of non-synonymous mutations/million bases. Simultaneous analysis of significantly mutated genes and their interactions among mutations between two CDS groups. "GISTIC 2.0" was employed for screening significantly amplified and missing genomic regions (38). The overall changes in the genome were quantified by calculating fraction of genome alteration (FGA), fraction of genome gained (FGG), and fraction of genome lost (FGL). FGA is the percentage of fragmented bases of genomic variants. FGG/FGL indicated the loss or gain of genomic variants.

TME annotations for CDS

Single Sample Gene Set Enrichment Analysis (ssGSEA), Tumor IMmune Estimation Resource (TIMER), and "MCPcounter" were used estimate the extent of immune cell infiltration in each patient. Subsequently, the TME was characterized using the "Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data (ESTIMATE)" algorithm. The ESTIMATE algorithm was used for calculating the tumor purity and the stromal, ESTIMATE, and immune scores. The data on the level of activation of the seven-step tumor immune cycle were retrieved from the tumor immunophenotype (TIP) (http://biocc.hrbmu.edu.cn/TIP/index.jsp) database (39). Additionally, we

determined and compared the expression profile of 35 immune checkpoint genes in patients between both CDS groups to elucidate the to elucidate the ability of CDS to predict the response of patients to immune checkpoint inhibitor therapy. The data on microsatellite instability (MSI), single nucleotide variant (SNV) neoantigens, and B-cell receptor (BCR) richness, as well as T-cell receptor (TCR) richness of patients, were obtained from TCGA. The "GSEA" package was used to compare the hallmark functions and pathways enriched by patients in both CDS groups, and the reliability of the enrichment analysis was validated using the "Gene Set Variation Analysis (GSVA)" package. The gene sets with "FDR < 0.05" were considered significantly enriched.

Predicting the patient's response to immunotherapy and chemotherapy

To predict the responses of patients to immunotherapy, we calculated the CDS for all patients from IMvigor210 (40), GSE78220 (41), GSE79671 (42), and GSE103668 (43) cohorts. We used "Tumor Immune Dysfunction and Exclusion (TIDE)" a web-based tool for predicting the response of the patient's to immunotherapy (44). We performed the "submap" method to determine the similarity in the expression in patients in both CDS groups and different immunotherapeutic outcomes (45). The data on drug sensitivity in cancer cell lines of human origin were downloaded from the Cancer Therapeutics Response Portal (CTRP, http://portals.broadinstitute.org/ctrp/) and Profiling Relative Inhibition Simultaneously in Mixtures (PRISM), https://depmap.org/portal/prism/) databases. We also plotted receiver operating characteristic curve (ROC) and calculated the Area Under the ROC (AUC) values for all patients using the "pRRophtic" package (46). Generally, lower AUC values indicated higher sensitivity to potential drugs (47).

Analysis of single-cell RNA sequencing data

ScRNA-seq files of three patients with LUAD from GSE117570 were retrieved from GEO. The expression matrices were normalized using the "Seurat" package, and the top 2000 highly variable genes were identified. The batch effect was eliminated using the "harmony" package (48). The "copyKAT" and "SingleR" packages were used to annotate tumor and immune cells (49, 50). Cell clustering analysis was performed using the "T-SNE" algorithm, and the top 11 principal components were selected. Genes with "|log2FC|>1" and "adjusted $P < 0.01$" were considered marker genes.

Tissue microarray and immunohistochemistry staining

We procured the LUAD TM (HPan-Ade060CS-01) from Shanghai Outdo Biotech Co., LTD (Shanghai, China). HLugA060PG02 contains 30 LUAD and adjacent paraneoplastic tissues. All the raw data could be obtained at the Shanghai Outdo Biotech Co. LTD's official website. Due to the absence of two paracancer samples in TM, we only performed IHC on 30 LUAD samples and 28 paracancer samples according to the following procedure. First, TM sections were dewaxed and rehydrated using decreasing grade of ethanol solution. Next, antigen recovery was performed in an autoclave using an acidic antigen repair solution (pH 6.0), the endogenous peroxidase activity was attenuated, and the antigenic sites were blocked using 5% bovine serum albumin. TM sections were incubated with 1:200 diluted anti-DDIT4 monoclonal antibody (ProteinTech, Wuhan, China, Cat No.10638-1-AP) for 16 hours at 4°C, followed by incubation with horseradish peroxidase (Maixin, Fujian, China) conjugated secondary antibody. 3,3'-diaminobenzidine (DAB, Maixin, Fujian, China) was used for immunoreactivity, and the nuclei were counterstained with hematoxylin. Finally, Interpret the results and group the samples according to the following criteria: The appearance of faint yellow to brownish granules in the cytoplasm is considered positive, while their absence is considered negative. Staining intensity in positive samples is scored as follows: no positive staining, weakly positive: (+), yellow-brown: positive (++) and dark brown: strongly positive (+++). Expression grouping of sample: Negative and weakly positive expression is included in the low expression group, while positive and strongly positive expression is included in the high expression group.

Cell culture and transfection

H358 and H838 (LUAD cells) and BEAS-2B (normal bronchial epithelial cells) were purchased from ATCC. All cell lines were of human origin. We cultured H358 and H838 in RPMI 1640 medium (Gibco, Shanghai, China) and BEAS-2B in DMEM (Gibco, Shanghai, China). Both mediums were supplemented with 10% fetal bovine serum (FBS, Gibco, Shanghai, China) and 1% penicillin/streptomycin. All cells were maintained in an incubator at 5% CO_2 and 37°C. Following the guidelines specified by the manufacturer, we transfected small interfering RNA (siRNA) against *DDIT4* (*DDIT4*-siRNA) and the corresponding control siRNA (siRNA-NC) into LUAD cells at the logarithmic growth stage using Lipofectamine 3000 transfection reagent (Invitrogen, MA, USA). The siRNA sequences are shown in Supplementary Table 3.

RNA extraction and real-time quantitative polymerase chain reaction

Following the manufacturer's guidelines, we isolated total cellular RNA using an RNA extraction kit (Analytik Jena AG, Jena, Germany). Next, a Promega qRT-PCR kit (Promega, WI, USA) was used to perform reverse transcription for synthesizing cDNA using extracted RNA. RT-PCR was performed using SYBR Premix Ex Taq II (Promega, WI, USA) on a real-time PCR detection system 480II (Roche, OR, USA). The PCR reaction conditions were 1 cycle of 95°C for 2 minutes, 40 cycles of 15 seconds, 60°C for 1 minute, 1 cycle of 95°C for 15 seconds, 60°C for 15 seconds, 95°C for 15 seconds. We

used β-Actin as the internal reference and the 2ΔΔCt method for quantifying relative gene expression. The primer sequences are provided in Supplementary Table 4.

Cell counting kit-8 assay

We performed a CCK-8 assay (Cellcook, Guangzhou, China, Cat No. CT01A) using the manufacturer's guidelines to determine the viability of cells transfected with *DDIT4*-siRNA and siRNA-NC. We seeded these cells in the logarithmic growth phase into 96-well plates. In order to evaluate the effect on cell proliferation capacity, 10 μl CCK-8 reagent was added in all wells and incubated for 2 hours at 37°C at 0, 24, 48, 72 h after culturing. For the sensitivity of cells to the drug, cells were treated at 37°C with Ispinesib (0, 20, 40, 80 or 100 nM, MedChem Express, Monmouth Junction, NJ, USA, Cat No.HY-50759), Cabazitaxel (0, 10, 20, 40 or 80 nM, MedChem Express, Monmouth Junction, NJ, USA, Cat No. HY-15459) and Epothilone-b (0, 40, 80, 160 or 320 nM, MedChem Express, Monmouth Junction, NJ, USA, Cat No. HY-17029) for 24 h, respectively. 10 μl CCK-8 reagent was added in all wells and incubated for 2 hours at 37°C. Finally, we measured the absorbance of each well at 450 nm using a microplate reader.

Clone formation assays

The clone formation rate was determined using a plate clone formation assay. 400 siRNA-NC and *DDIT4*-siRNA transfected cells/well were seeded in 12-well plates and incubated at 37°C for 14 days. Next, we washed the cells with PBS and fixed them using 4% paraformaldehyde. Finally, crystal violet was used for the purpose of staining the fixed cells, and the viable clones with a minimum of 50 cells were counted.

Transwell assay

The invasive and migratory capacities of siRNA-NC and *DDIT4*-siRNA transfected cells were tested by Transwell (pore size 8.0 μm; Corning Inc, NY, USA) coated with Matrigel (BD Biosciences, Bedford, USA). To determine the migratory capacity of cells, we inoculated 2×10^4 cells in 100 μl serum-free medium in the upper chamber. The lower chamber was supplemented with 800 μl 10% FBS-containing medium. The cells were incubated in an incubator for 24 hours, stained using crystal violet, and imaged under a light microscope. The "ImageJ" software was used for counting cells. For the invasion assay, the upper chamber was coated with 100 μl of 10% Matrigel. The rest of the procedure was the same as described above.

TUNEL staining

We utilized TUNEL staining (Solarbio, Beijing, China, Cat No. T2196) to examine apoptosis in tumor cells. An initial density of 1×10^5 cells per well was established in 12-well plates. These cells were subsequently immobilized onto coverslips with 4% paraformaldehyde for 30 minutes at room temperature, followed by two PBS washes. A treatment with 0.1% Triton X-100 was applied for 10 minutes at room temperature. After another PBS rinse, the cells were incubated in a 50 μl TUNEL reaction mixture at 37°C for 1 hour. To counterstain the cell nuclei, we employed 4,6-diamidino-2-phenylindole (DAPI, Solarbio, Beijing, China, Cat No. C0065) for 10 minutes at room temperature and washed cells twice with PBS. Finally, images of TUNEL-labeled cells were procured from three arbitrary fields using a fluorescent microscope.

Western blotting

To obtain total protein, cells were subjected to protein extraction using 1% PMSF and RIPA buffer (Solarbio, Beijing, China, Cat No. R0020) on ice for 30 minutes. The resulting mixture was centrifuged at 12,000 rpm for 30 minutes, and the protein suspension was collected from the liquid supernatant. Protein concentration was determined using the BCA method (Epizyme, Shanghai, China, Cat No. ZJ101). Subsequently, SDS-PAGE protein loading buffer (5X) (Beyotime, China) was added to the protein suspension, followed by boiling for 10 minutes. The protein was then separated using either a 10% or 12.5% SDS-PAGE gel (Epizyme, Shanghai, China, Cat No. PG113 or PG112) and transferred onto a 0.45 μm polyvinylidene fluoride (PVDF) membrane. To block the PVDF membranes, 5% skim milk was applied for 1.5 hours. Next, the membranes were incubated with primary antibodies including *DDIT4* (ProteinTech, Wuhan, China, Cat No. 67059-1-Ig, 1:1000), *BCL2* (ProteinTech, Wuhan, China, Cat No. 68103-1-Ig, 1:1000), *Caspase-3* (Huaan, Hangzhou, China, Cat No. ET1602-39, 1:1000), and *GAPDH* (Huaan, Hangzhou, China, Cat No. ET1601-4, 1:5000), followed by incubation with corresponding secondary antibodies. Finally, the protein bands were visualized using chemiluminescence kits.

Statistical analysis

We used software including "GraphPad Prism (version 9.00)" and "R (version 4.0.5) package" for statistically analyzing the data. We determined the correlation between two continuous variables using the "Pearson correlation". Next, the chi-squared test was employed for comparing categorical variables, and the "Wilcoxon rank-sum" or student's t-tests for continuous variables. All statistical tests were two-sided. If not otherwise stated, $P < 0.05$ was considered statistically significant.

Results

Transcriptional and genetic alterations of PCDRGs in patients with LUAD

The expression profiles of DEPCDRGs between normal and LUAD tissues from the TCGA-LUAD cohort were compared, and

200 DEGs were identified (Supplementary Table 5). The heatmap and volcano plot shows DEGs in these samples (Figures 2A, B). The GO and KEGG pathway enrichment analyses showed the enrichment of these DEPCDRGs in various biological pathways like tumor necrosis factor receptor superfamily binding, the TNF, regulation of apoptotic, and IL-17 signaling pathways (Figures 2C, D, Supplementary Table 6). Next, we determined the status of DEPCDRGs mutation in patients with LUAD. Approximately 73.88% (444/601) of patients with LUAD harbored mutations in DEPCDRGs. Of the top 10 mutated DEPCDRGs, *TP53* had the highest mutation frequency (Figures 2E, F). The frequencies of CNV in DEPCDRGs analysis showed that most DEPCDRGs harbored significant CNVs. The

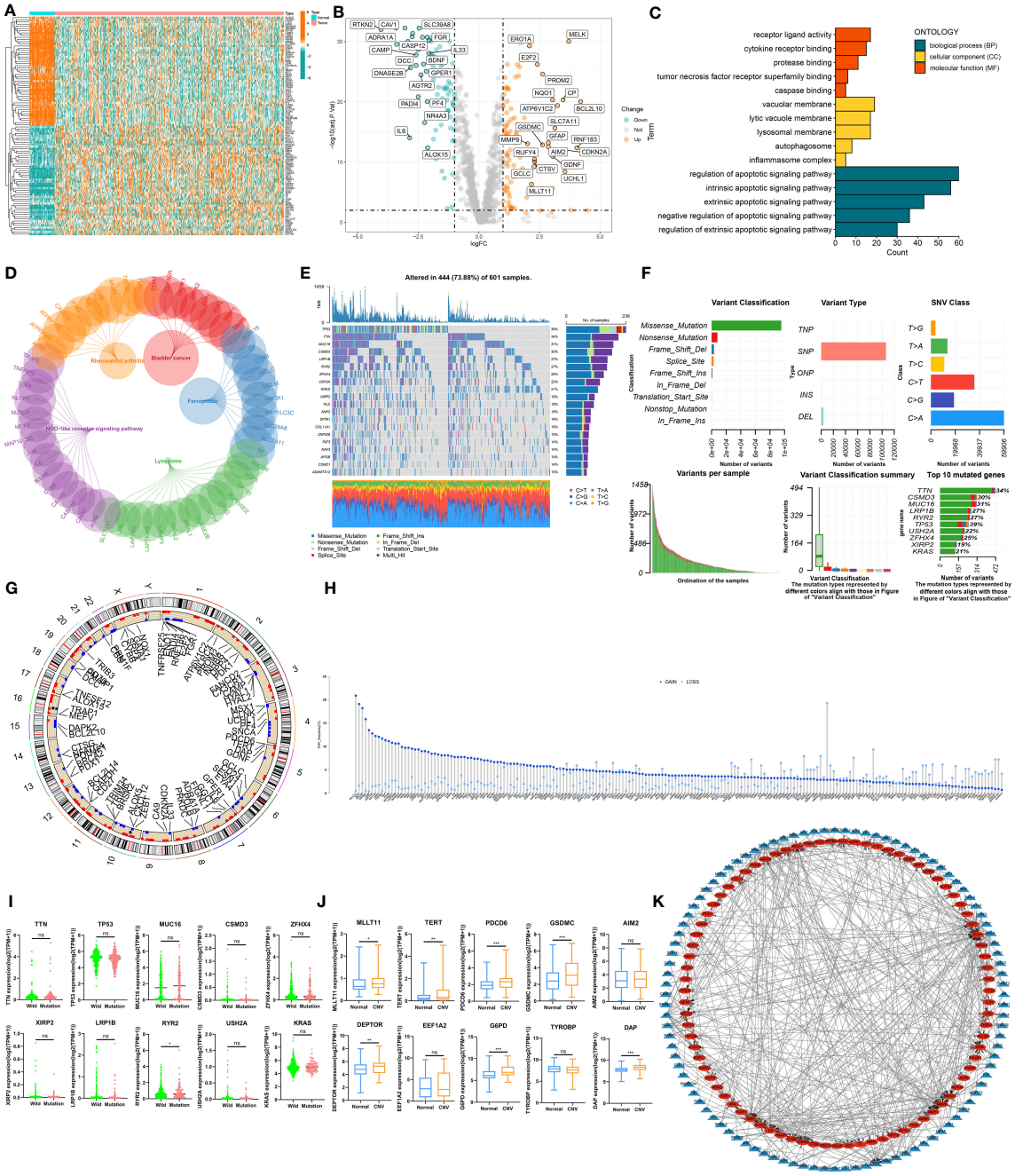


FIGURE 2
The landscape of Programmed cell death related genes (PCDRGs) in TCGA-LUAD cohort. **(A)** Heatmap of the differentially expressed PCDRGs between tumor and paracancer tissues of LUAD. **(B)** Volcano plot of the DEPCDRGs. **(C)** GO categories [molecular function (MF), biological process (BP) and cellular component (CC)] and **(D)** KEGG pathway analysis for DEPCDRGs. **(E, F)** The mutation summary and details of DEPCDRGs in the LUAD patients **(G)** The location of CNV alterations of DEPCDRGs on chromosomes. **(H)** CNV mutation situation of the DEPCDRGs. **(I)** The Scatter plot of gene expression for the top 10 DEPCDRGs with the highest CNV frequency. **(J)** Boxplot of gene expression for the top 10 DEPCDRGs with the highest mutation frequency. **(K)** Network map of the DEPCDRGs transcription factors and DEPCDRGs. Blue triangular nodes represent transcription factors, red oval nodes represent DEPCDRGs, and lines between nodes indicate regulatory relationships. ns, not significant, *P < 0.05, **P < 0.01, ***P < 0.001.

chromosomal locations of CNVs in DEPCDRGs are shown in Figure 2G. Interestingly, the highest frequencies of CNV gain and loss were observed in *MLLT11* and *CDKN2A*, respectively (Figure 2H). It is worth noting that DEPCDRGs undergoing CNV often exhibit higher expression levels, but mutations and corresponding gene expression show no significant correlation (Figures 2I, J). 452 TF-target pairs were obtained by predicting the TFs of the genes associated with DEPCDRGs, which included 106 predicted TFs and 84 target DEPCDRGs. Figure 2K illustrates the regulatory relationships of these pairs.

CDS signature development

We performed univariate Cox regression analysis on 200 DEPCDRGs and identified 71 prognosis-related PCDRGs (Supplementary Table 7). These 71 PCDRGs were subjected to a machine learning-based integration procedure for developing cell death core (CDS). In addition, 101 prediction models were fitted in the training cohort using the LOOCV framework and validated on the test cohorts. Next, we calculated the C-index for all models in whole cohorts (Figure 3A and Supplementary Table 8). The mean C-index value of 0.727 was the highest in the RSF (including *GAPDH, DDIT4, KRT18*, and *ENO1*) and was considered the best model (Figures 3B, C, Supplementary Table 9). Subsequently, we calculated the CDS for all patients based on the RSF model (Supplementary Table 10). All patients were categorized using the median CDS value as a threshold into high and low-CDS groups. The survival duration of patients with high CDS from whole cohorts was short (Figures 3D-H). In addition, we evaluated the performance of CDS based on the patient's clinical characteristics. The results demonstrated that the ability of CDS to predict patients'

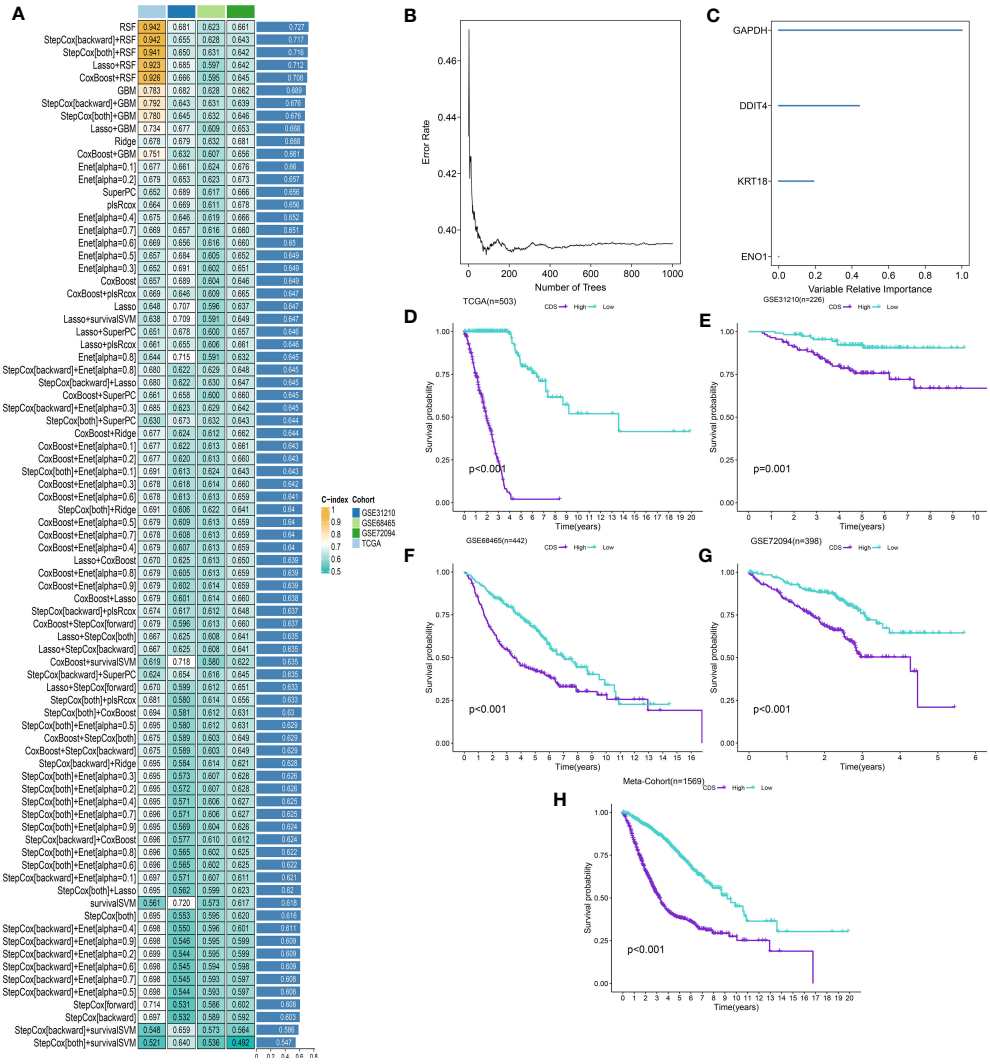


FIGURE 3
A cell death score (CDS) was established and validated *via* the machine learning-based integrative procedure. **(A)** A total of 101 kinds of machine learning algorithms were used to obtain the optimal model and calculated the C-index of each model for all cohorts. **(B, C)** The number of trees for determining the CDS with minimal error and the importance of the 4 PCDRGs based on the RSF algorithm. **(D-H)** Kaplan–Meier curves of OS according to the CDS in TCGA, GSE31210, GSE68465, GSE72094 and meta-cohort.

survival was not influenced by their clinical characteristics (Supplementary Figure 1).

CDS assessment

We conducted the "time-ROC" analysis to calculate the AUC values of CDS for predicting the prognosis of patients in different cohorts [Figures 4A-E; TCGA [0.95–0.98], GSE31210 (0.67–0.8), GSE68465 (0.63–0.68), GSE72094 (0.68–0.78), and meta-cohort (0.71–0.81)]. The C-index values of all cohorts are shown in Figure 4F. Next, we compared the abilities of CDS and other clinical as well as molecular variables in predicting patients' prognoses. The accuracy of CDS in predicting patients' prognoses was better compared to other variables, including age, gender, smoking, *TP53, EGFR, KRAS, STK11*, M, and T (Figures 4G-J). The advancement in sequencing technology and bioinformatics have aided in developing models based on the combination of the expression profile of genes for predicting the patient's diagnosis and prognosis. Subsequently, we systematically searched LUAD-related signatures published in the last 3 years. Finally, we included 77 biomarkers for comparison of predictive performance with CDS (Supplementary Table 11). The results revealed that the performance of our CDS in almost all cohorts was better compared to other signatures (Figure 5). Further, we analyzed the correlation between CDS and other clinical variables. The chi-squared test results showed a correlation between all variables except for gender and both CDS groups (Supplementary Figure 2). After incorporating clinical data of patients, univariate and multivariate Cox regression analyses of the four cohorts indiated that CDS could predict patients' prognoses independently (Tables 1-4).

Generation of CDS genetic subtypes

We performed consistent clustering on four genes included, and the samples were grouped into distinct characteristic subtypes to identify PCD-related subtypes of LUAD. Finally, we identified two PCD-related phenotypes: clusters 1 and 2 (Figures 6A, B, Supplementary Table 12). Kaplan-Meier (KM) curves showed higher OS rate patients in cluster 1 compared to cluster 2 in all cohorts (Figure 6C). In addition, the alluvial diagram showed that most patients with high CDS were grouped in cluster 2 (Figure 6D).

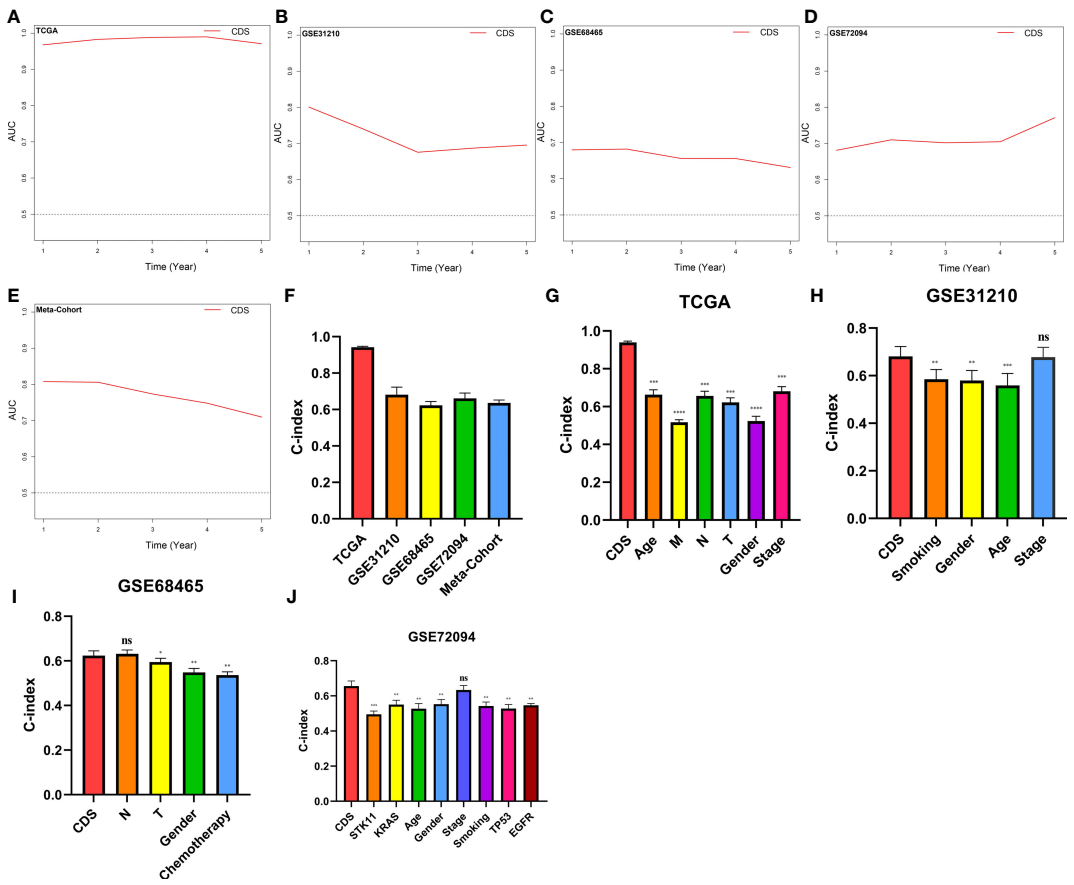


FIGURE 4
Evaluation of the CDS. **(A–E)** Time-dependent receiver operating characteristic curve of CDS for predicting the prognosis of LUAD patients from TCGA, GSE31210, GSE68465, GSE72094 and meta-cohort. **(F)** The C-index of the CDS for the TCGA, GSE31210, GSE68465, GSE72094 cohorts. **(G–J)** The C-index of the CDS and other clinical factors in the TCGA, GSE31210, GSE68465, GSE72094 cohorts. ns, not significant, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

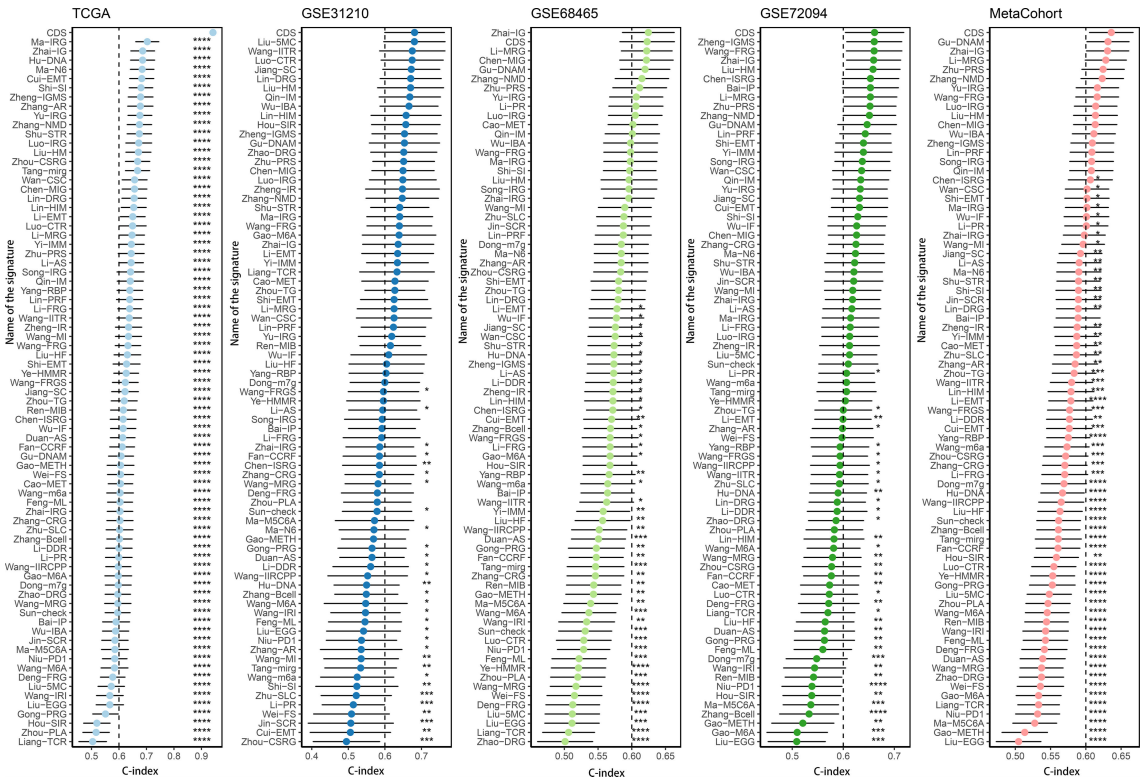


FIGURE 5
Comparison of CDS and other gene expression-based prognostic signatures in LUAD based on the TCGA, GSE31210, GSE68465, GSE72094 and meta-cohort. ns, not significant, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

Genetic variations in CDS groups

Figures 7A, B shows the top 20 genes with the highest mutation frequencies in patients in both CDS groups. The results revealed differences in mutated genes in patients between both CDS groups. The frequency of sense, nonsense, or overall mutations in patients in the high-CDS group was higher compared to the low-CDS group, despite no correlation between CDS and mutation frequency (Figures 7C-E). In addition, a significant difference in the mutation frequency of 16 genes was observed in patients between both CDS groups (Figure 7F), and there was extensive co-mutation between these genes (Figure 7G). Patients in the high-CDS group had high TMB compared to the low-CDS group (Figure 7H, Supplementary Table 13). KM analysis showed that the OS rates of patients in the high-TMB group were significantly higher compared to the low-TMB group (Figure 7I). Next, the prognosis of patients was predicted based on their TMB and CDS. The survival rate of patients in the low-CDS group with high TMB was the highest, whereas the survival rate of patients in the high-CDS group with low TMB was the lowest (Figure 7J). CNV analysis revealed differences in chromosomal alteration patterns in patients in different CDS groups (Figure 7K). The high-CDS group had a

TABLE 1 Univariate and multivariate Cox analysis of the clinicopathological features and FA score with OS for TCGA cohort.

	Univariate Cox		Multivariate Cox	
Characteristics	HR(95%CI)	*P* value	HR(95%CI)	*P* value
Stage	1.977(1.586-2.463)	< **0.001**	1.302(0.919-1.845)	0.137
M	1.727(1.18-2.527)	**0.005**	0.799(0.507-1.259)	0.333
N	1.942(1.575-2.394)	< **0.001**	1.306(0.972-1.753)	0.076
T	1.816(1.386-2.38)	< **0.001**	1.492(1.075-2.07)	**0.017**
Age	1.038(0.822-1.31)	0.754		
Sex	1.041(0.847-1.28)	0.7		
CDS	0.028(0.016-0.047)	< **0.001**	0.035(0.019-0.062)	< **0.001**

Significant value is given in bold.

TABLE 2 Univariate and multivariate Cox analysis of the clinicopathological features and FA score with OS for GSE68465 cohort.

	Univariate Cox		Multivariate Cox	
Characteristics	HR(95%CI)	*P* value	HR(95%CI)	*P* value
N	2.029(1.689-2.438)	< **0.001**	2.053(1.686-2.5)	< **0.001**
T	2.062(1.587-2.68)	< **0.001**	1.806(1.37-2.383)	< **0.001**
Sex	1.262(1.051-1.516)	**0.013**	1.239(1.021-1.503)	**0.03**
Chemotherapy	1.412(1.15-1.734)	< **0.001**	1.243(1.003-1.541)	**0.047**
CDS	0.655(0.544-0.788)	< **0.001**	0.658(0.544-0.797)	< **0.001**

Significant value is given in bold.

TABLE 3 Univariate and multivariate Cox analysis of the clinicopathological features and FA score with OS for GSE31210 cohort.

	Univariate Cox		Multivariate Cox	
Characteristics	HR(95%CI)	*P* value	HR(95%CI)	*P* value
Smoking	1.417(0.882-2.277)	0.15	NA	NA
Sex	1.344(0.839-2.152)	0.219	NA	NA
Age	1.263(0.777-2.052)	0.346	NA	NA
Stage	2.774(1.732-4.441)	< **0.001**	2.313(1.413-3.787)	< **0.001**
CDS	0.434(0.254-0.743)	**0.002**	0.552(0.314-0.969)	**0.038**

Significant value is given in bold.

TABLE 4 Univariate and multivariate Cox analysis of the clinicopathological features and FA score with OS for GSE72094 cohort.

	Univariate Cox		Multivariate Cox	
Characteristics	HR(95%CI)	*P* value	HR(95%CI)	*P* value
STK11	1.028(0.72-1.469)	0.879	NA	NA
KRAS	0.767(0.588-0.999)	**0.049**	0.901(0.686-1.184)	0.454
Age	1.258(0.836-1.894)	0.27	NA	NA
Gender	0.733(0.564-0.952)	**0.02**	0.714(0.546-0.934)	**0.014**
Stage	1.969(1.477-2.625)	< **0.001**	1.925(1.438-2.579)	< **0.001**
Smoking	1.248(0.694-2.245)	0.459	NA	NA
CDS	0.536(0.407-0.707)	< **0.001**	0.605(0.456-0.801)	< **0.001**

Significant value is given in bold.

greater percentage of FGA, FGL, and FGG detected. (Figure 7L, Supplementary Table 14).

Characteristics of TME in CDS groups

To evaluate if CDS could be used to determine the immune status of patients, we analyzed the correlation between CDS and infiltrating immune cells (Supplementary Table 15, 16). The proportion of infiltrating immune cells in patients in the low-CDS group was high (Figure 8A). Next, our analysis of cancer progression revealed that the majority of key steps, including cancer antigen presentation, priming and activation and B cell recruiting, displayed higher activity levels in the low CDS group (Figure 8B, Supplementary Table 17). Additionally, an increase in the expression of most immune checkpoint genes was observed in patients in the low-CDS group (Figure 8C, Supplementary Table 18). Additionally, several factors associated with tumor immunogenicity were analyzed, like the status of MSI, SNV neoantigens, and BCR and TCR richness (Supplementary Table 19). Patients in the high-CDS group had high MSI, SNV, neoantigens, and low BCR and TCR richness (Figures 8D-G). Together, these results suggest that patients with low CDS had highly active immune status. GSEA analysis was performed to compare the Hallmark pathways enriched in patients in both CDS groups (Supplementary Table 20). We observed significant enrichment of patients with high CDS in pathways and functions related to cell cycle, hypoxia, glycolysis, and mTOR signaling, and

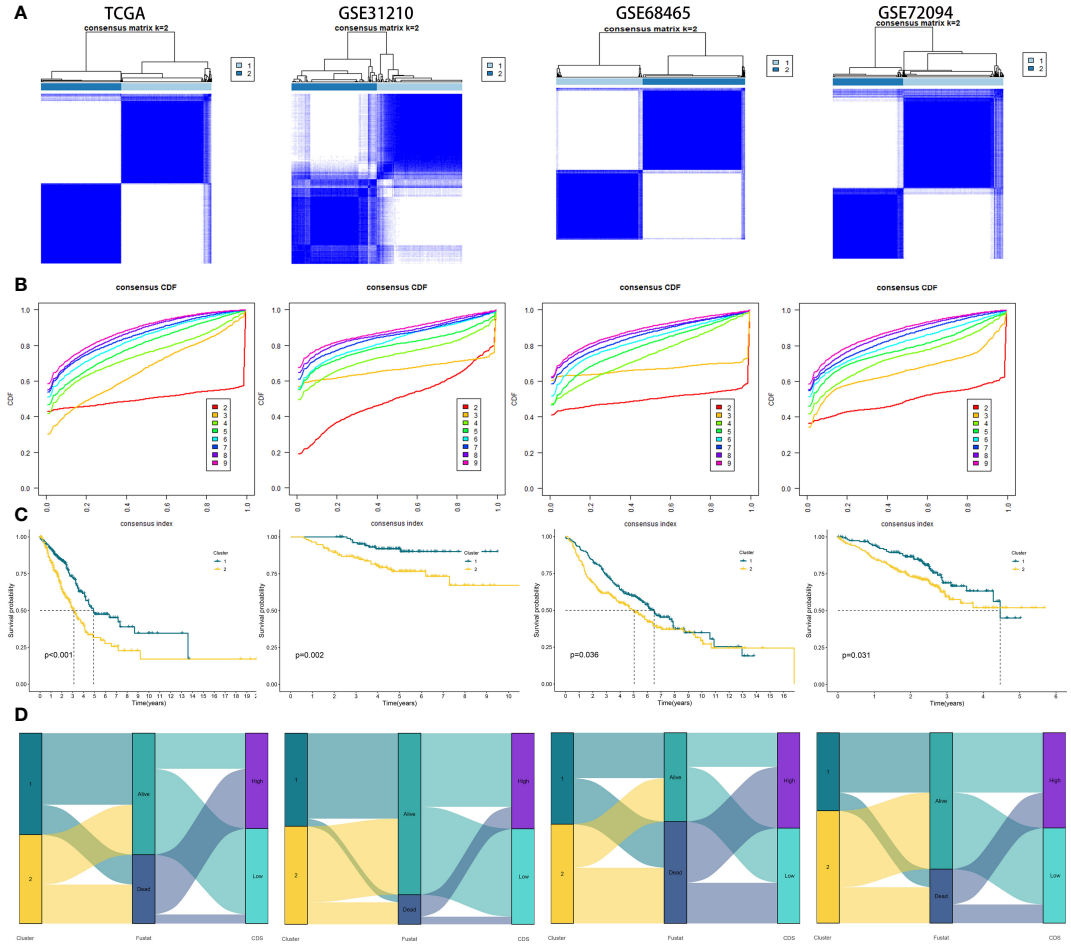


FIGURE 6
Generation of clusters by Unsupervised clustering of CDS gene expression for TCGA, GSE31210, GSE68465, GSE72094 cohorts. **(A)** Consensus clustering matrix of LUAD patients for k = 2. **(B)** Consensus clustering cumulative distribution function for k = 2 to 9. **(C)** Kaplan–Meier curves for patients in two different molecular clusters **(D)** Alluvial diagram of clusters distributions in groups with different CDS groups, clusters and survival outcomes.

these results were validated in GSVA analysis (Figures 8H-L, Supplementary Table 21).

Predicting immunotherapy efficacy and identification of potential drugs

We calculated the CDS of patients from the immunotherapy cohorts to determine if CDS could predict the response to immunotherapy. In the IMvigor210 cohort, the OS rate of patients with high CDS still was lower (Figure 9A), but the response of them to PD-L1 immunotherapy was better (Figure 9B). The response of patients from GSE78220 to immunotherapy was similar to the IMvigor210 cohort (Figure 9C). In addition, the response of patients with high CDS in the GSE79671 and GSE103668 cohorts to immunotherapy was better (Figures 9D, E). Subsequently, we analyzed patient immune evasion and immunotherapy using TIDE scores and found that patients in the high-CDS group were less likely to experience immune evasion and had better immunotherapy outcomes (Figure 9F, Supplementary Table 22). The "SubMap" algorithm results showed that patients in the high-CDS group were more likely to respond to PD-1 immunotherapy (Figure 9G). Next, we screened for potential drugs for treating patients with LUAD using the CTRP and PRISM-derived drug response cohorts. Finally, we obtained two compounds, paclitaxel and SB-743921 from the CTRP cohort (Figure 9H) and six compounds including cabazitaxel, daunorubicin, epothilone-b, ispinesib, litronesib, and volasertib from the PRISM cohort (Figure 9I). Interestingly, patients in the high-CDS group demonstrated sensitivity to these drugs.

Single-cell level analysis of CDS

We performed principle component analysis to reduce the dimensionality of all cells using 2000 highly variable genes. Seven cell types, like monocytes, T cells, B cells, macrophages, cancer cells, tissue stem cells, and endothelial cells, were annotated (Figures 10A, B). Additionally, marker genes for each cell type were identified (Figures 10C, D), and the CDS of each cell type was calculated. Cancer cells, T cells, and monocytic regions had high CDS (Figure 10E). Pseudotime trajectory analysis shows the

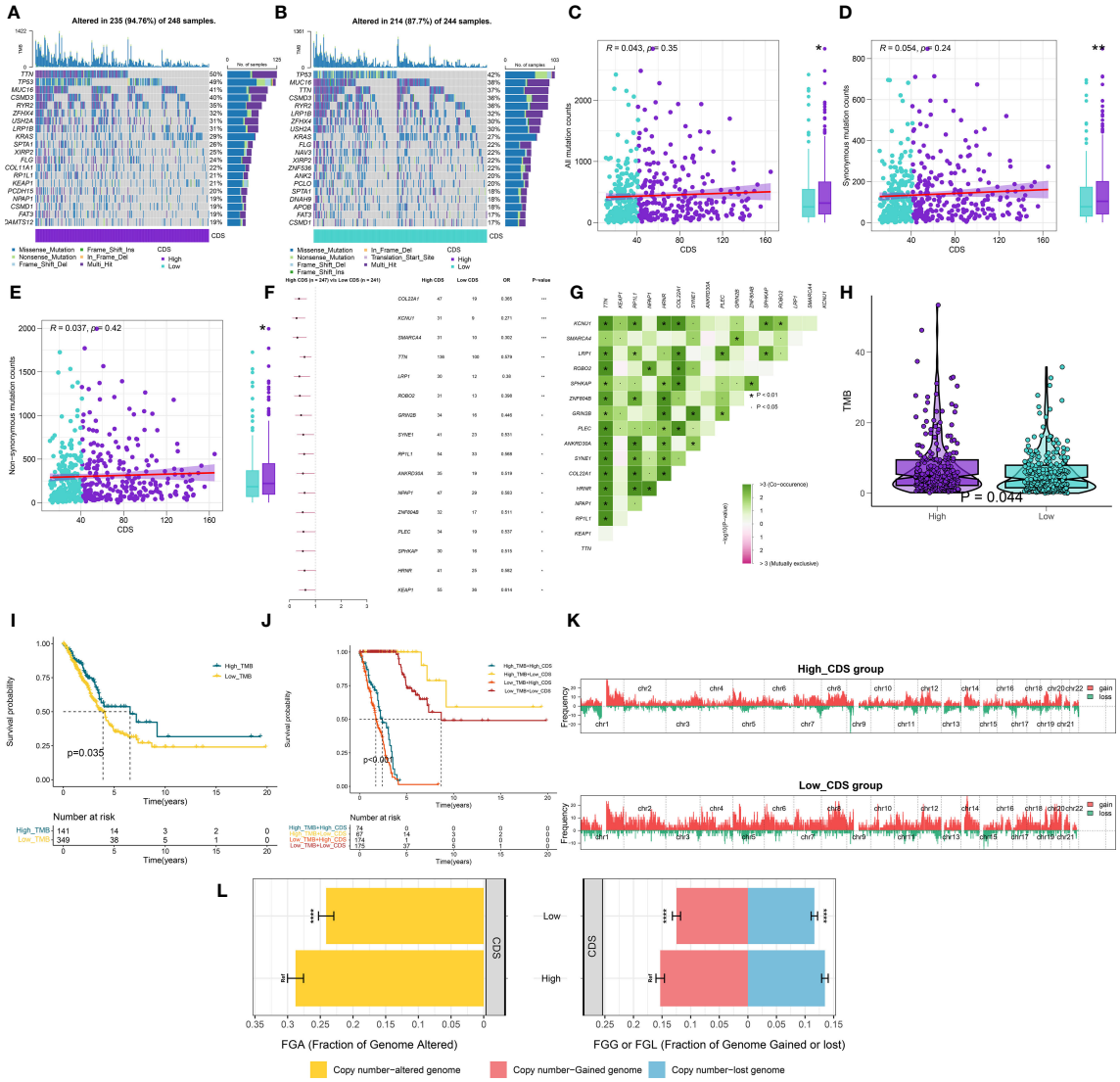


FIGURE 7
Integrated comparisons of somatic mutation and CNVs between high and low CDS groups in the TCGA cohort. **(A, B)** Waterfall plots showing the mutation information of the top 20 genes with the highest mutation frequency in the CDS groups. **(C-E)** Association between all mutation counts, synonymous mutation counts, nonsynonymous mutation counts, and CDS and their distribution in the CDS groups. **(F)** Differentially mutated genes between high and low CDS groups are displayed as a forest plot. **(G)** Interaction effect of genes mutating differentially in patients in the CDS groups. **(H)** Distribution of TMB in the CDS groups. **(I)** Kaplan–Meier curves for the OS of the high-TMB and low-TMB groups. **(J)** Kaplan–Meier curves for patients stratified by both TMB and CDS. **(K)** Gene fragments profiles with amplification (red) and deletion (green) among the CDS groups. **(L)** Comparison of the fraction of the genome altered, lost, and gained between the CDS groups. ns, not significant, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

chronological order of cell differentiation. The cells with low CDS were mainly disturbed at the end of the differentiation pathway, and the cells with high CDS were primarily distributed at the early stage of the differentiation pathway (Figures 10F-H).

DDIT4 affects tumor cell proliferation, invasion, migration and apoptosis

The expression of four genes of CDS in LUAD and normal cells was verified by RT-qPCR. Compared to normal cell lines, all genes were highly expressed in LUAD cells, with *DDIT4* showing the most significant difference (Figure 11A). Owing to the highest expression of DDIT4 among the four genes within LUAD cells, coupled with the absence of reports regarding its progression in LUAD, we elected to conduct subsequent experiments involving *DDIT4*. IHC confirmed high *DDIT4* expression in LUAD tissues (Figures 11B, C). Next, we performed several experiments to determine the roles of *DDIT4* in LUAD. RT-qPCR and WB results revealed a significant decrease in *DDIT4* expression in cells transfected with *DDIT4*-siRNAs (Figures 12A, B). The CCK-8 and clone formation assays showed a reduction in the viability and clone formation of cells in the *DDIT4*-siRNAs transfected cells compared to the siRNA-NC transfected cells (Figures 12C, D). Next, we performed transwell assay to evaluate the mobility, migratory, and invasive abilities of LUAD cells. Compared to the siRNA-NC transfected cells, a loss of

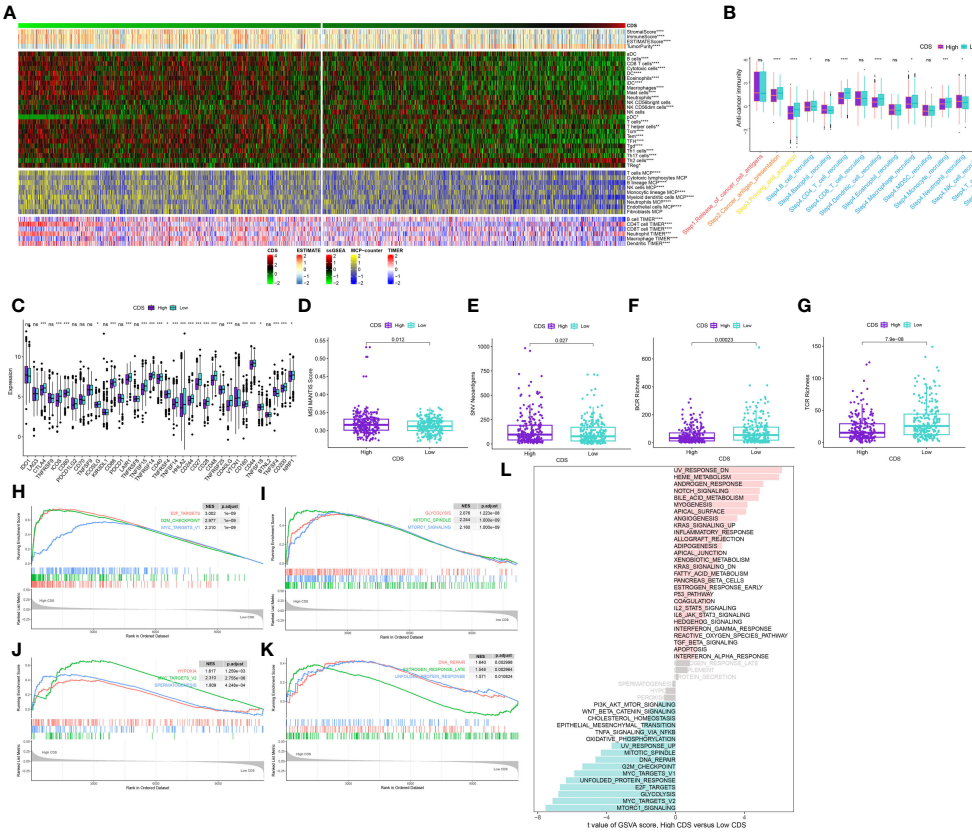


FIGURE 8
Immune-related characteristics of the CDS. **(A)** Heatmap displaying the correlation between the CDS and immune infiltrating cells in the meta-cohort. **(B)** Boxplot showing the differences of anti-cancer immunity score between CDS groups. **(C)** Comparison of immune checkpoint-related genes levels between CDS groups in the meta-cohort. **(D–G)** The distribution of MSI, neoantigens, BCR richness, TCR richness levels in different CDS groups from TCGA cohort. **(H–K)** The GSEA results for the 12 overlapping upregulated hallmark pathways in terms of the high CDS groups. **(L)** The difference in the hallmark gene sets between different CDS groups by GSVA. ns, not significant, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

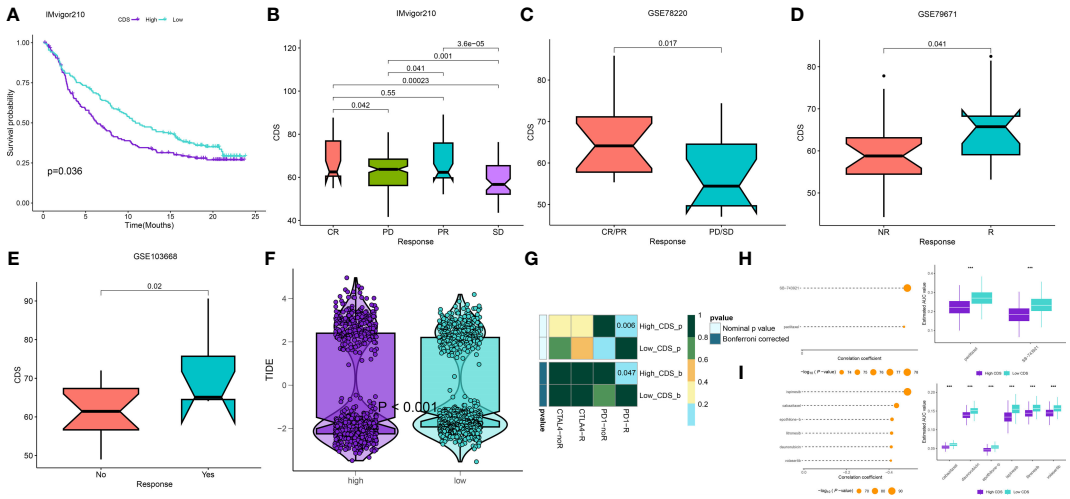


FIGURE 9
Differential putative immunotherapy and chemotherapy response for patients from high and low CDS groups. **(A)** Kaplan-Meier curve for patients in high and low CDS groups in the IMvigor cohort. **(B-E)** Box plot showing different CDS from patients with immunotherapy responses in the IMvigor, GSE103668, GSE79671 and GSE78220 cohorts. **(F)** Violin plot showing different TIDE scores from patients with different CDS. **(G)** Submap analysis of the meta-cohort and melanoma patients with detailed immunotherapeutic information. **(H)**The results of correlation analysis and differential drug response analysis of CTRP-derived drugs. **(I)** The results of correlation analysis and differential drug response analysis of PRISM-derived drugs. ns, not significant, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

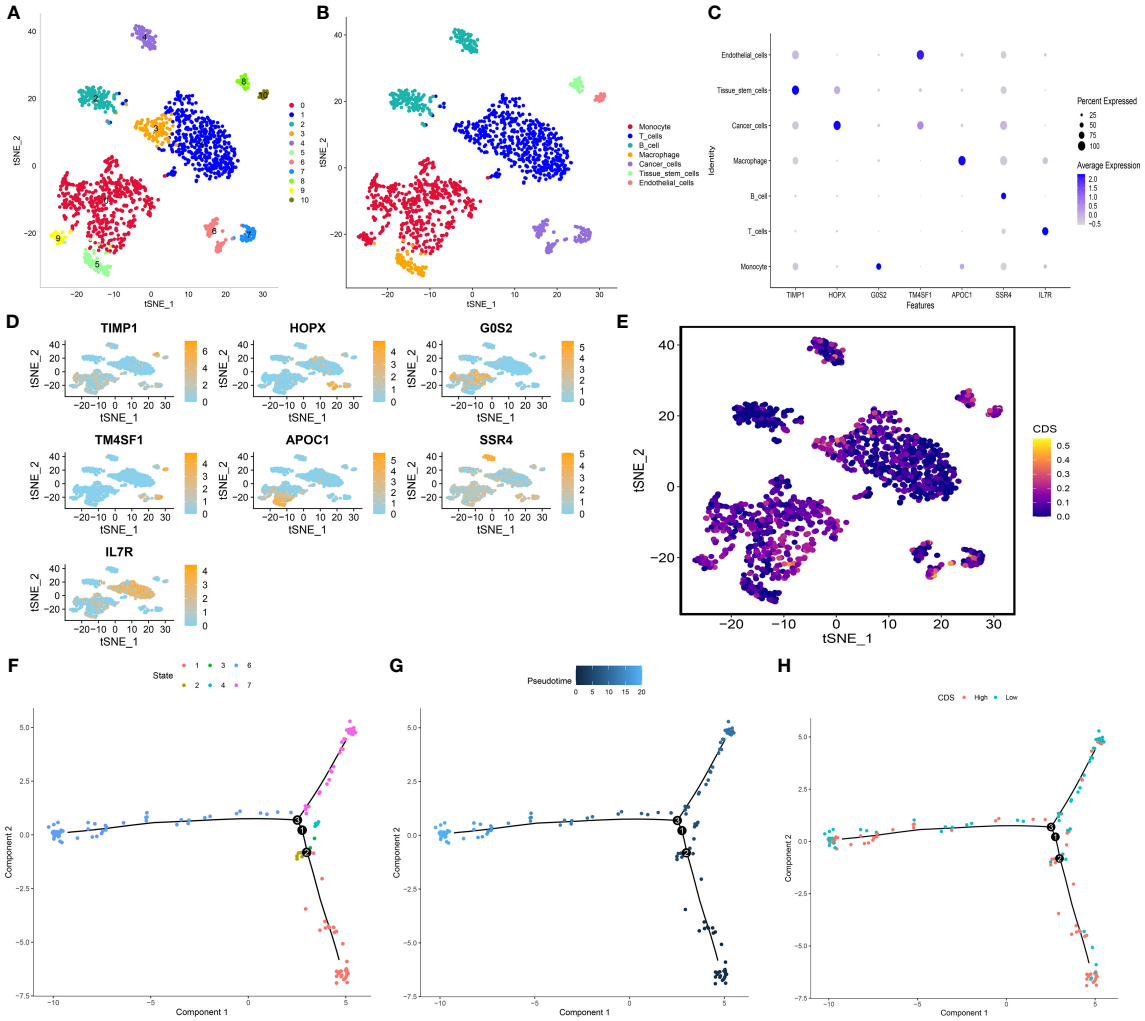


FIGURE 10
Exploration of CDS in LUAD scRNA-seq data. **(A)** t-SNE plot colored by 11 cell subpopulations. **(B)** t-SNE plot of the distribution of 7 cell types. **(C, D)** Marker gene expression of each cell type. **(E)** CDS distributions in the different single cells. **(F-H)** Pseudotime trajectory analysis in LUAD cells (Cells are colored based on states, pseudotime and CDS groups, labels 1, 2, and 3 correspond to the node identifiers and their respective quantities in the figures.

invasive and migratory abilities of cells in the *DDIT4* siRNAs transfected cells was observed (Figure 12E). Besides, knockdown of *DDTI4* promotes apoptosis and increases sensitivity to ispinesib and cabazitaxel in LUAD cells (Figures 12F, G and Figure 13). Together, these results suggest that the *DDIT4* may play a pro-oncogenic role and a therapeutic target in LUAD.

Discussion

Despite ongoing efforts, the treatment of LUAD remains challenging, as the disease is often aggressive and associated with poor prognosis (51, 52). Therefore, studies should focus on identifying molecular markers and therapeutic targets for LUAD. It is well established that cell death has vital anticancer effects and serves as a therapeutic target. Studies have shown that several PCDs could influence the TME and attenuate tumorigenesis, cancer progression, and cancer treatment, thus improving the prognosis and survival of patients with cancer (53, 54). Commonly used chemotherapy agents and immune checkpoint inhibitors trigger cell death, thereby attenuating cancer progression (55). However, several cancers have an innate resistance to cell death (56). Therefore, deciphering the underlying mechanisms and functions of cell death, specifically PCD types and the steps involved in regulated cell death, holds great promise for providing insights into cancer development and anti-cancer therapeutics. In clinical practice, the pathological stage of LUAD determines the patient's prognosis (57). However, the clinical outcomes of patients with similar pathological stages of LUAD are often different, which indicates the inadequacy of current staging systems in providing reliable predictions and reflecting LUAD heterogeneity (58). As next-generation sequencing technologies continue to advance, RNA-seq has emerged as a potent approach for discovering novel biomarkers and therapeutic targets (59, 60). In recent years, numerous models based on gene assemblies of various PCD types have demonstrated commendable prognostic and therapeutic

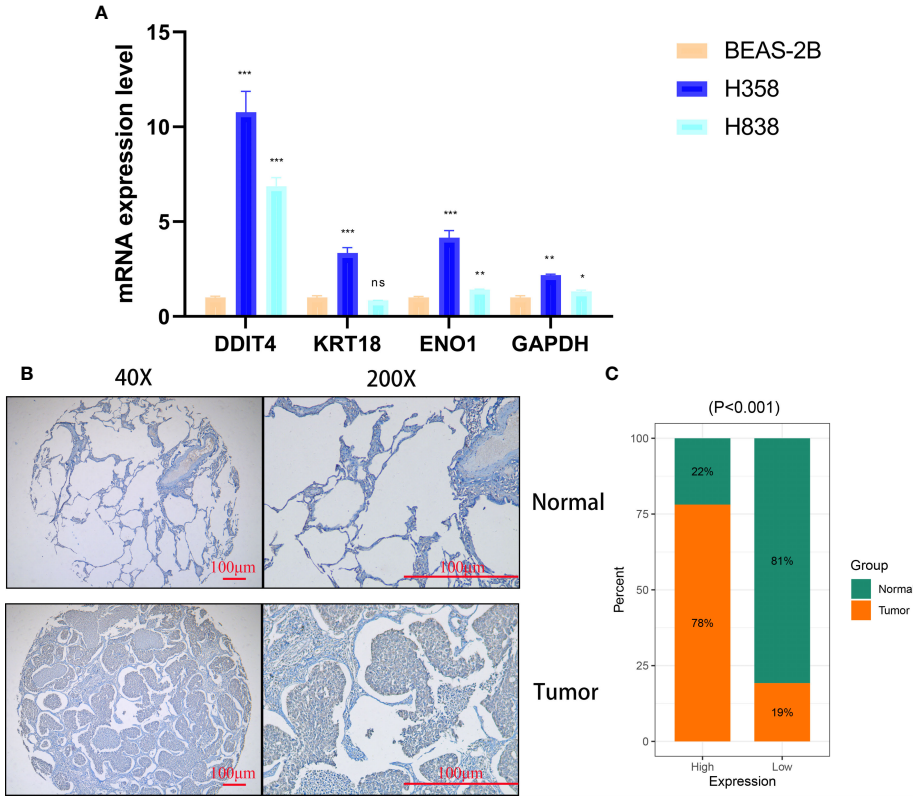


FIGURE 11
Cellular and histological and validation candidate gene expression changes. **(A)** CDS genes expression in cancer and normal cell lines. beta-actin was used as the internal reference gene and experiment was performed in triplicate and at least three times. **(B, C)** IHC analysis of DDIT4 in 30 LUAD and 28 adjacent tissues. ns, not significant, *P < 0.05, **P < 0.01, ***P < 0.001.

predictive value, underscoring the potent latent capabilities and clinical implications of PCD-related genes (61, 62). Nevertheless, a comprehensive analysis examining PCD-associated genes in LUAD has not yet been reported. In this study, we comprehensively analyzed PCDRGs from 12 PCD types. Using the gene expression profiles of these genes, we developed and validated 101 models through the "LOOCV" framework across multiple cohorts, resulting in the identification of the optimal RFS model. This approach not only utilizes various algorithms to fit models with consistent prognostic value for LUAD patients but also enables the models to become simpler and more interpretable. The "The "KM", "Time-ROC" and "C-index" analyses showed higher accuracy and stability of CDS in stratifying the prognosis of patients with LUAD in multiple cohorts. Furthermore, multivariate Cox regression analysis showed that CDS could independently predict the prognosis of patients with LUAD. Next, we compared our CDS with 77 previously published genetic LUAD models, and the C-index results revealed that the performance of our CDS was better than these 77 published models. Therefore, CDS could be a novel and reliable tool for stratifying patients with LUAD.

All four PCDRGs included in CDS were associated with tumor initiation and progression. *GAPDH* is a key enzyme in step 6 of the glycolytic pathway (63). Studies have demonstrated an increase in *GAPDH* expression levels in various tumor tissues and cells (64–66). Malignant cells prefer aerobic glycolysis for producing adenosine triphosphate to oxidative phosphorylation (67). An increase in the expression of glycolytic enzymes is considered a hallmark of cancer metabolism (66). Studies have shown the involvement of *GAPDH* in several processes, like the apoptosis of cells, maintaining DNA integrity, and angiogenesis. Antisense oligonucleotides or anticancer agents targeting *GAPDH* could inhibit the proliferation of colon cancer cells and trigger the apoptosis of cervical cancer cells (68, 69). *DDIT4* is a novel HIF-responsive gene (70). Studies have shown a close association between increased *DDIT4* expression in hypoxic or stressful conditions and DNA damage, inflammation, ROS, and autophagy during cancer occurrence and development. *DDIT4* activates the TSC1/2 and NF-κB pathways, thereby endogenously inhibiting the mTORC1 pathway. High *DDIT4* expression is observed in several cancers and is linked to poor patient prognosis (71, 72). Conversely, the prognosis of patients with lung or pancreatic cancers harboring RAS mutations and *DDIT4* deletion is poor. This could be due to reprogramming the oxidation of fatty acids and the accumulation of pyruvate and lactate (73). Our *in vitro* experiments showed high *DDIT4* expression in LUAD tissues and cells, which promoted proliferation, invasion, migration of and inhibited apoptosis of LUAD cells. *KRT18* is a keratin protein and intermediate filaments necessary for tissue integrity (74). *KRT18* is one of the

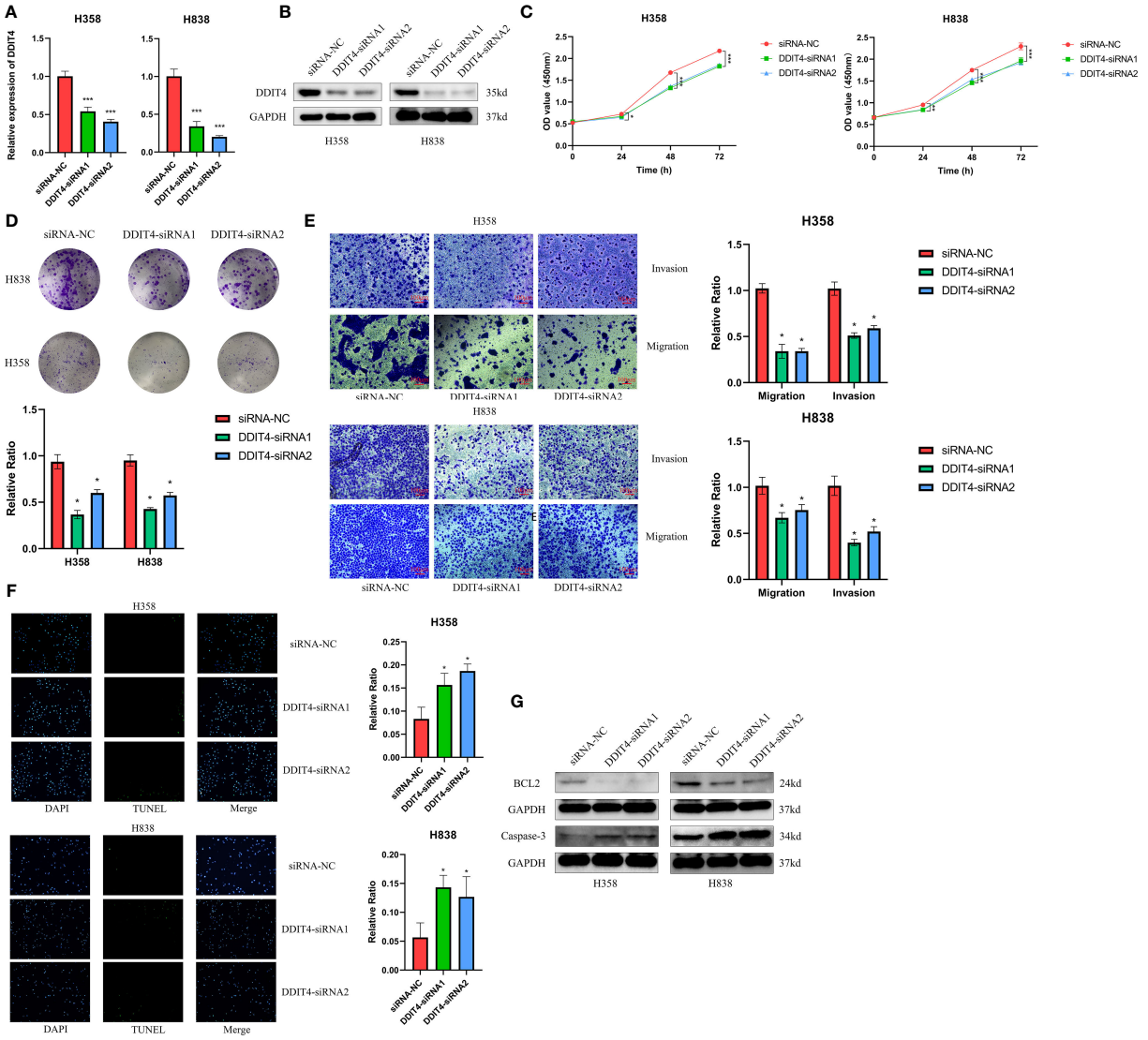


FIGURE 12
DDIT4 promoted proliferation, migration, invasion and inhibited apoptosis of LUAD cell lines. **(A, B)** Knockdown of DDIT4 was confirmed by RT-PCR and WB. beta-actin and GAPDH was used as the internal reference gene. **(C, D)** CCK8 and clone formation assays were performed to assess cell viability and proliferation of H358 and H838 cells. **(E)** Transwell assay was performed to assess cell migration and invasion of H358 and H838 cells. **(F)** TUNEL staining of H358 and H838 cells. **(G)** WB analysis of BCL2 and caspase-3 proteins in H358 and H838 cells. *P < 0.05, ***P < 0.001.

most abundant keratin proteins of epithelial and endothelial cells. It is expressed in many malignant tumors, including NSCLC, gastric cancer (GC), hepatocellular carcinoma (HCC) and breast cancer (BC), making it widely used as a diagnostic and prognostic marker for cancers (75–78). In addition, *KRT18* is an important regulator of tumors. *EGR1* enhances *KRT18* expression and promotes the apoptosis of NSCLC cells (76). Studies have shown that reduced *KRT18* expression enhances the susceptibility of cervical cancer cells to cytokine-induced cell death, inhibits cell migration (79), and enhances the sensitivity to paclitaxel in LC (75). *ENO1* is an enzyme involved in metabolism, the pyruvate synthesis and triggers the activation of the fibrinolytic enzyme and the degradation of the extracellular matrix (80). Several studies have demonstrated the involvement of *ENO1* in several physiological processes like metabolism, the remodeling of the extracellular matrix, controlling the growth of cells, and metastasis (81, 82). Studies have demonstrated that *ENO1* promotes the migration and metastasis of cancer cells *via* the mechanism of regulating intravascular and pericyte fibrinolytic activity (83, 84). Besides, *ENO1* could be a valuable prognostic marker. The relapse-free survival and OS of patients with NSCLC expressing high *ENO1* level is relatively shorter (85). Moreover, a study has indicated that targeting *ENO1* could be a novel and effective approach to overcoming drug resistance (86).

PCD regulates TME by triggering the crosstalk between innate and adaptive immunity to induce immunostimulatory responses (87). TME is critical for cancer development and response to treatment (88). Our results of single-cell RNA sequencing

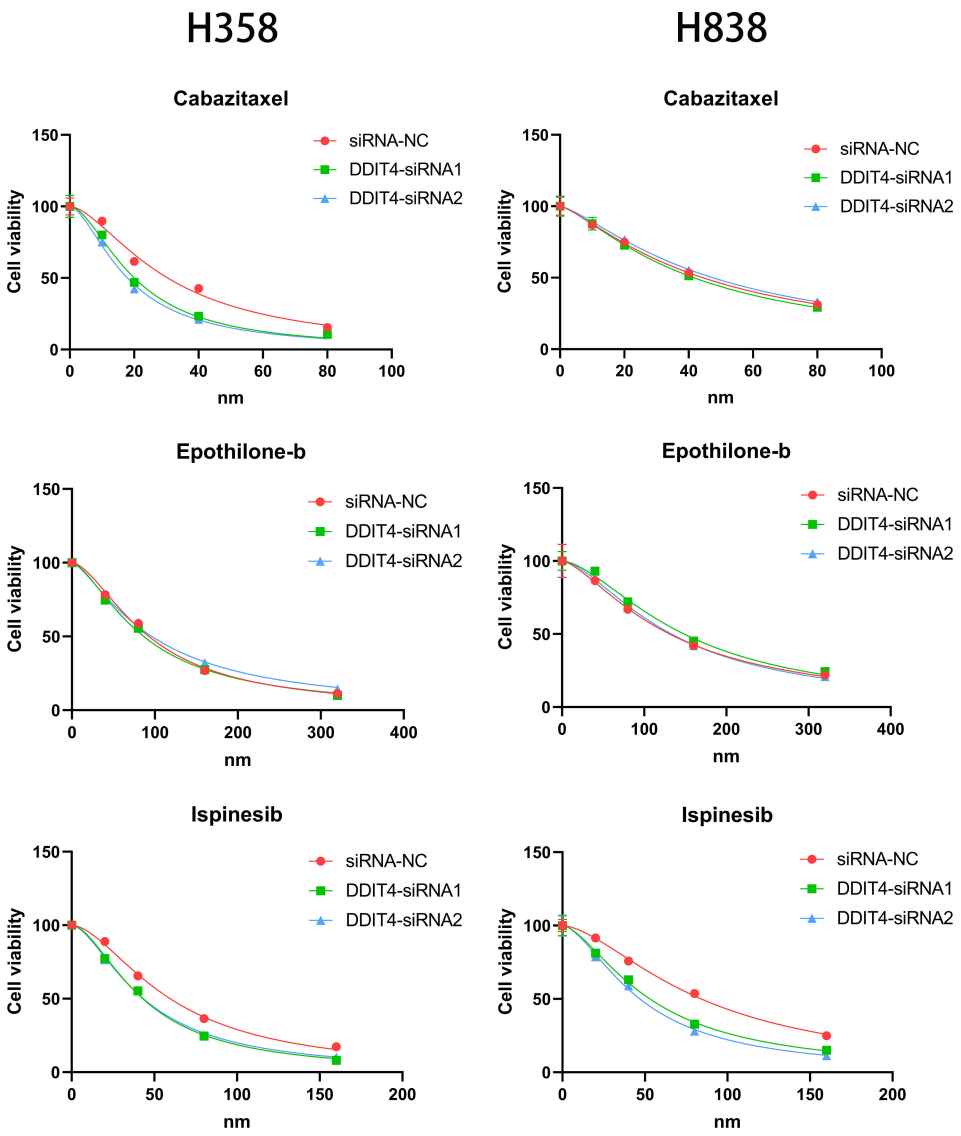


FIGURE 13
Effect of DDIT4 on chemotherapy sensitivity of H358 and H838 cells.

analyses and clustering showed that high CDS scoring cells are mainly concentrated in the areas of T cells, endothelial cells, and tumor cells. Monocytes have increasingly been recognized as critical influencers in cancer evolution and progression, with various subtypes displaying contradictory roles in facilitating tumor expansion and impeding the metastasis of malignant cells (89). Macrophages, notably prominent in the pulmonary cancer milieu, are significant inflammatory entities that modulate both innate and adaptive immune responses in cancer. The M1 subtype of macrophages exudes tumor-suppressing molecules like ROS and NO, eliciting cytotoxic reactions on cancer cells (90, 91). Contrarily, M2 macrophages can synthesize a range of cytokines that foster the proliferation and survival of tumor cells. Additional research indicates that an established positive feedback mechanism involving CCL5 and CCL18 between M2 macrophages and myofibroblasts contributes to the malignant progression of phyllodes tumors (92). T cells represent the predominant tumor-infiltrating immune cells in the TME (93), including various t-cell subsets. These subsets, along with select other immune cell types, perform dual roles within the lung TME, engaging in both tumor-suppressing and tumor-promoting activities (94). CD8+ T cells are pivotal in orchestrating anti-tumor immunity, effectively eliminating tumor cells through the recognition of tumor-associated antigens exhibited in major histocompatibility complex class I (95). Conversely, regulatory T cells (Tregs) are capable of suppressing anti-cancer immunity. This suppression undermines protective immunosurveillance of neoplasia and obstructs potent anti-tumor immune responses in hosts carrying tumors, thereby fostering tumor evolution and advancement (96). These cells exhibit high CDS, indicating that they may have some interaction in TME. In this study, multiple immune cells including CD4 T, CD8 T and macrophages showed high infiltration in patients with low

CDS, whereas Treg cells showed low infiltration. Moreover, studies have shown that poor prognosis were closely associated with an imbalance in the ratio of immune cells in patients with cancers (97, 98). The results showed that the patients in the low-CDS group had better OS rate and higher immune scores, suggesting that patients with highly active immune state have a better prognosis. Tumor cells with lower levels of differentiation often exhibit faster growth rates, higher invasiveness, and are typically associated with poorer prognosis (99). The results of pseudotime analysis indicated that tumor cells with high CDS levels were positioned at the front end of the differentiation pathway, while tumor cells with low CDS levels were located at the terminal end of the differentiation pathway. Therefore, we found that the levels of CDS may be associated with the degree of differentiation and invasiveness in tumor cells. CNV is a prevalent type of variation in tumors and serves as a pivotal factor propelling the initiation and progression of cancer. Studies suggest that elevated levels of CNV can stimulate tumor cell proliferation and immune evasion, often resulting in a poorer prognosis for patients (100, 101). The levels of FGA, FGG, and FGL were significantly high in patients in the high-CDS group in our study, which corroborates previous research reports. Numerous studies suggest that patients with higher levels of TMB, MSI, and SNV neoantigens are more likely to respond to immune therapy, while those with higher TIDE scores tend to exhibit the opposite trend (102–105). Therefore, we compared TMB, MSI, SNV neoantigen, and TIDE scores in patients in both CDS groups to predict patients' response to immunotherapy. As expected, patients in the high-CDS group had higher TMB, MSI, SNV neoantigens, and lower TIDE scores. Furthermore, patients in the high-CDS group responded better to immunotherapy and could gain more benefit from immunotherapy in multiple cohorts. These results validate the efficacy of our CDS in predicting patients' responses to immunotherapy. Several studies are focusing on the combined use of chemo and immunotherapies for treating patients with cancer. Immunotherapy can reduce damage caused to the immune system by chemotherapy, and the combined use of chemo and immunotherapies could exert synergistic antitumor effects (106, 107). Finally, we performed an interaction analysis between CDS and drug response to screen for drugs that can be used in combination with immunotherapy in patients in the high-CDS group and aid in guiding personalized therapy. As a result, 8 potential anti-tumor drugs that are more sensitive to patients with high CDS were identified. Ispinesib is a highly selective small molecule inhibitor of KSP that inhibits the formation of bipolar mitotic spindles, leading to cell cycle arrest without centrosome separation (108). It exhibits broad-spectrum antitumor activity in various *in vitro* tumor cell lines and xenograft models. Cabazitaxel is a chemotherapy drug approved for the treatment of prostate cancer, primarily exerting its antiproliferative effect by inhibiting spindle formation and function (109). Cabazitaxel exhibits broad-spectrum antitumor activity against various tumors, including Furthermore, colorectal cancer, pancreatic cancer, and lung cancer (110). Cabazitaxel promotes autophagic cell death in LUAD by targeting the PI3K/Akt/mTOR pathway (111). Moreover, self-assembled micelles loaded with cabazitaxel exhibit good hydrophilicity and enhanced anticancer effects, making them potential candidates for lung cancer treatment (112). Similar to cabazitaxel, epothilone-b belongs to the class of microtubule stabilizers. Epothilone-b exerts its anticancer effect mainly by activating the extrinsic apoptosis pathway involving caspase-3 and caspase-8 (113). Furthermore, epothilone-b has been confirmed as one of the clinical drugs capable of inducing genuine immunogenic cell death (114). For lung cancer, epothilone-b enhances the radiosensitivity of LUAD cells by reducing DNA repair capacity (115). However, possibly due to the limitations of cell line types, our experimental results found that knockdown of DDIT4 can only stably affect the sensitivity of LUAD cells to ispinesib. Moreover, the results of GSEA and GSVA analyses indicate a significant enrichment in cell proliferation and metabolism in patients with high CDS, such as "OXIDATIVE_PHOSPHORYLATION", "DNA_REPAIR", "G2M_CHECKPOINT", and "GLYCOLYSIS" (116–119). This could potentially elucidate the heightened sensitivity of patients in the high-CDS group to these chemotherapeutic drugs.

However, our study has several limitations. Firstly, due to the fact that research on PCD is a rapidly evolving and emerging field, it is possible that an increasing number of PCDRGs will be identified beyond the 1215 genes included in this study. Second, the patients included in our study were from retrospective studies conducted at single centers. Therefore, prospective studies at multiple centers should be conducted to validate the reliability and validity of CDS. Finally, we have only explored the effect of *DDIT4* on LUAD cells using siRNA. Therefore, more genetic modification and intervention strategies are required to determine the involvement and mechanism of *DDIT4* in LUAD.

Conclusions

In conclusion, we have developed and validated an accurate and robust CDS based on four PCDRGs using extensive machine learning algorithms. Our CDS could effectively predict the survival and response of patients with LUAD to immunotherapy. CDS is a powerful tool for predicting the patient's prognosis and designing personalized therapy. These results provide new directions and shed light on the molecular mechanisms of LUAD.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

The studies involving humans were approved by Ethics Committee and Institutional Review Board of the Outdo Biotech. Co., Ltd. (SHYJS-CP-1904007). 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

YH conceived and designed the study. YZ and YW performed the collection and assembly of data. JC and YX analyzed the data. YZ and YW performed experiments and wrote the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported in part by Medical Vertical Project of Fujian Province (Grant No. 2020CXB001) to YH, Joint fund of science and technology innovation of Fujian province (Grant No. 2021Y9024) to YH, Key Project of Natural Science Foundation of Fujian province (Grant No. 2022J02048) to YH.

Acknowledgments

We thank Bullet Edits Limited for the linguistic editing and proofreading of the manuscript.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2023.1183230/full#supplementary-material

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(6):394–424. doi: 10.3322/caac.21492

2. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* (2012) 150(6):1107–20. doi: 10.1016/j.cell.2012.08.029

3. Wu F, Wang L, Zhou C. Lung cancer in China: current and prospect. *Curr Opin Oncol* (2021) 33(1):40–6. doi: 10.1097/cco.0000000000000703

4. Bean J, Brennan C, Shih JY, Riely G, Viale A, Wang L, et al. MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proc Natl Acad Sci U.S.A.* (2007) 104(52):20932–7. doi: 10.1073/pnas.0710370104

5. Bleeker FE, Felicioni L, Buttitta F, Lamba S, Cardone L, Rodolfo M, et al. AKT1(E17K) in human solid tumours. *Oncogene* (2008) 27(42):5648–50. doi: 10.1038/onc.2008.170

6. Jin G, Kim MJ, Jeon HS, Choi JE, Kim DS, Lee EB, et al. PTEN mutations and relationship to EGFR, ERBB2, KRAS, and TP53 mutations in non-small cell lung cancers. *Lung Cancer* (2010) 69(3):279–83. doi: 10.1016/j.lungcan.2009.11.012

7. Brody H. Lung cancer. *Nature* (2020) 587(7834):S7. doi: 10.1038/d41586-020-03152-0

8. Bade BC, Dela Cruz CS. Lung cancer 2020: epidemiology, etiology, and prevention. *Clin Chest Med* (2020) 41(1):1–24. doi: 10.1016/j.ccm.2019.10.001

9. Tang D, Kang R, Berghe TV, Vandenabeele P, Kroemer G. The molecular machinery of regulated cell death. *Cell Res* (2019) 29(5):347–64. doi: 10.1038/s41422-019-0164-5

10. Wyllie AH. Apoptosis: cell death in tissue regulation. *J Pathol* (1987) 153 (4):313–6. doi: 10.1002/path.1711530404

11. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013

12. Shi J, Zhao Y, Wang K, Shi X, Wang Y, Huang H, et al. Cleavage of GSDMD by inflammatory caspases determines pyroptotic cell death. *Nature* (2015) 526(7575):660–5. doi: 10.1038/nature15514

13. Kayagaki N, Stowe IB, Lee BL, O'Rourke K, Anderson K, Warming S, et al. Caspase-11 cleaves gasdermin D for non-canonical inflammasome signalling. *Nature* (2015) 526(7575):666–71. doi: 10.1038/nature15541

14. Ursini F, Maiorino M. Lipid peroxidation and ferroptosis: The role of GSH and GPx4. *Free Radic Biol Med* (2020) 152:175–85. doi: 10.1016/j.freeradbiomed.2020.02.027

15. Xie Y, Hou W, Song X, Yu Y, Huang J, Sun X, et al. Ferroptosis: process and function. *Cell Death Differ* (2016) 23(3):369–79. doi: 10.1038/cdd.2015.158

16. Kroemer G, Levine B. Autophagic cell death: the story of a misnomer. *Nat Rev Mol Cell Biol* (2008) 9(12):1004–10. doi: 10.1038/nrm2529

17. Tasdemir E, Galluzzi L, Maiuri MC, Criollo A, Vitale I, Hangen E, et al. Methods for assessing autophagy and autophagic cell death. *Methods Mol Biol* (2008) 445:29–76. doi: 10.1007/978-1-59745-157-4_3

18. González-Juarbe N, Bradley KM, Shenoy AT, Gilley RP, Reyes LF, Hinojosa CA, et al. Pore-forming toxin-mediated ion dysregulation leads to death receptor-independent necroptosis of lung epithelial cells during bacterial pneumonia. *Cell Death Differ* (2017) 24(5):917–28. doi: 10.1038/cdd.2017.49

19. Tanzer MC, Khan N, Rickard JA, Etemadi N, Lalaoui N, Spall SK, et al. Combination of IAP antagonist and IFNγ activates novel caspase-10- and RIPK1-dependent cell death pathways. *Cell Death Differ* (2017) 24(3):481–91. doi: 10.1038/cdd.2016.147

20. Kalliolias GD, Ivashkiv LB. TNF biology, pathogenic mechanisms and emerging therapeutic strategies. *Nat Rev Rheumatol* (2016) 12(1):49–62. doi: 10.1038/nrrheum.2015.169

21. 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(6586):1254–61. doi: 10.1126/science.abf0529

22. 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(2):102–13. doi: 10.1038/s41568-021-00417-2

23. Ma D, Lu B, Feng C, Wang C, Wang Y, Luo T, et al. Deoxypodophyllotoxin triggers parthanatos in glioma cells *via* induction of excessive ROS. *Cancer Lett* (2016) 371(2):194–204. doi: 10.1016/j.canlet.2015.11.044

24. Liccardi G, Ramos Garcia L, Tenev T, Annibaldi A, Legrand AJ, Robertson D, et al. RIPK1 and caspase-8 ensure chromosome stability independently of their role in cell death and inflammation. *Mol Cell* (2019) 73(3):413–428.e7. doi: 10.1016/j.molcel.2018.11.010

25. Akhiani AA, Werlenius O, Aurelius J, Movitz C, Martner A, Hellstrand K, et al. Role of the ERK pathway for oxidant-induced parthanatos in human lymphocytes. *PloS One* (2014) 9(2):e89646. doi: 10.1371/journal.pone.0089646

26. Donizy P, Halon A, Surowiak P, Pietrzyk G, Kozyra C, Matkowski R. Correlation between PARP-1 immunoreactivity and cytomorphological features of parthanatos, a specific cellular death in breast cancer cells. *Eur J Histochem* (2013) 57 (4):e35. doi: 10.4081/ejh.2013.e35

27. Mackay HL, Muller PAJ. Biological relevance of cell-in-cell in cancers. *Biochem Soc Trans* (2019) 47(2):725–32. doi: 10.1042/bst20180618

28. Florey O, Kim SE, Sandoval CP, Haynes CM, Overholtzer M. Autophagy machinery mediates macroendocytic processing and entotic cell death by targeting single membranes. *Nat Cell Biol* (2011) 13(11):1335–43. doi: 10.1038/ncb2363

29. Ravindran M, Khan MA, Palaniyar N. Neutrophil extracellular trap formation: physiology, pathology, and pharmacology. *Biomolecules* (2019) 9(8):365. doi: 10.3390/biom9080365

30. Zhou W, Guo Y, Zhang X, Jiang Z. Lys05 induces lysosomal membrane permeabilization and increases radiosensitivity in glioblastoma. *J Cell Biochem* (2020) 121(2):2027–37. doi: 10.1002/jcb.29437

31. Stoka V, Turk V, Turk B. Lysosomal cysteine cathepsins: signaling pathways in apoptosis. *Biol Chem* (2007) 388(6):555–60. doi: 10.1515/bc.2007.064

32. Song X, Zhu S, Xie Y, Liu J, Sun L, Zeng D, et al. JTC801 induces pH-dependent death specifically in cancer cells and slows growth of tumors in mice. *Gastroenterology* (2018) 154(5):1480–93. doi: 10.1053/j.gastro.2017.12.004

33. Holze C, Michaudel C, Mackowiak C, Haas DA, Benda C, Hubel P, et al. Oxeiptosis, a ROS-induced caspase-independent apoptosis-like cell-death pathway. *Nat Immunol* (2018) 19(2):130–40. doi: 10.1038/s41590-017-0013-y

34. Zhang CC, Li CG, Wang YF, Xu LH, He XH, Zeng QZ, et al. Chemotherapeutic paclitaxel and cisplatin differentially induce pyroptosis in A549 lung cancer cells *via* caspase-3/GSDME activation. *Apoptosis* (2019) 24(3-4):312–25. doi: 10.1007/s10495-019-01515-1

35. Malhotra J, Jabbour S, Orlick M, Riedlinger G, Guo Y, White E, et al. Phase Ib/II study of hydroxychloroquine in combination with chemotherapy in patients with metastatic non-small cell lung cancer (NSCLC). *Cancer Treat Res Commun* (2019) 21:100158. doi: 10.1016/j.ctarc.2019.100158

36. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* (2018) 28(11):1747–56. doi: 10.1101/gr.239244.118

37. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize Implements and enhances circular visualization in R. *Bioinformatics* (2014) 30(19):2811–2. doi: 10.1093/bioinformatics/btu393

38. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. *Nat Genet* (2006) 38(5):500–1. doi: 10.1038/ng0506-500

39. Xu L, Deng C, Pang B, Zhang X, Liu W, Liao G, et al. TIP: A web server for resolving tumor immunophenotype profiling. *Cancer Res* (2018) 78(23):6575–80. doi: 10.1158/0008-5472.Can-18-0689

40. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* (2018) 554(7693):544–8. doi: 10.1038/nature25501

41. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* (2017) 168(3):542. doi: 10.1016/j.cell.2017.01.010

42. Urup T, Staunstrup LM, Michaelsen SR, Vitting-Seerup K, Bennedbæk M, Toft A, et al. Transcriptional changes induced by bevacizumab combination therapy in responding and non-responding recurrent glioblastoma patients. *BMC Cancer* (2017) 17(1):278. doi: 10.1186/s12885-017-3251-3

43. Birkbak NJ, Li Y, Pathania S, Greene-Colozzi A, Dreze M, Bowman-Colin C, et al. Overexpression of BLM promotes DNA damage and increased sensitivity to platinum salts in triple-negative breast and serous ovarian cancers. *Ann Oncol* (2018) 29(4):903–9. doi: 10.1093/annonc/mdy049

44. 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(10):1550–8. doi: 10.1038/s41591-018-0136-1

45. Roh W, Chen PL, Reuben A, Spencer CN, Prieto PA, Miller JP, et al. Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci Transl Med* (2017) 9(379):aah3560. doi: 10.1126/scitranslmed.aah3560

46. 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(9):e107468. doi: 10.1371/journal.pone.0107468

47. Yang C, Huang X, Li Y, Chen J, Lv Y, Dai S. Prognosis and personalized treatment prediction in TP53-mutant hepatocellular carcinoma: an in silico strategy towards precision oncology. *Brief Bioinform* (2021) 22(3):bbaa164. doi: 10.1093/bib/bbaa164

48. 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(12):1289–96. doi: 10.1038/s41592-019-0619-0

49. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* (2019) 20(2):163–72. doi: 10.1038/s41590-018-0276-y

50. 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(5):599–608. doi: 10.1038/s41587-020-00795-2

51. Nasim F, Sabath BF, Eapen GA. Lung cancer. *Med Clin North Am* (2019) 103(3):463–73. doi: 10.1016/j.mcna.2018.12.006

52. D'Amico AG, Maugeri G, Rasà DM, Reitano R, Saccone S, Federico C, et al. Modulatory role of PACAP and VIP on HIFs expression in lung adenocarcinoma. *Peptides* (2021) 146:170672. doi: 10.1016/j.peptides.2021.170672

53. Yu J, Wang Q, Zhang X, Guo Z, Cui X. Mechanisms of neoantigen-targeted induction of pyroptosis and ferroptosis: from basic research to clinical applications. *Front Oncol* (2021) 11:685377. doi: 10.3389/fonc.2021.685377

54. Yang B, Zhong W, Gu Y, Li Y. Emerging mechanisms and targeted therapy of pyroptosis in central nervous system trauma. *Front Cell Dev Biol* (2022) 10:832114. doi: 10.3389/fcell.2022.832114

55. Liu W, Zhang L, Xiu Z, Guo J, Wang L, Zhou Y, et al. Combination of immune checkpoint inhibitors with chemotherapy in lung cancer. *Onco Targets Ther* (2020) 13:7229–41. doi: 10.2147/ott.S255491

56. Sasidharan Nair V, Elkord E. Immune checkpoint inhibitors in cancer therapy: a focus on T-regulatory cells. *Immunol Cell Biol* (2018) 96(1):21–33. doi: 10.1111/imcb.1003

57. Jurišić V, Obradovic J, Pavlović S, Djordjevic N. Epidermal growth factor receptor gene in non-small-cell lung cancer: the importance of promoter polymorphism investigation. *Anal Cell Pathol (Amst)* (2018) 2018:6192187. doi: 10.1155/2018/6192187

58. Rafei H, El-Bahesh E, Finianos A, Nassereddine S, Tabbara I. Immune-based therapies for non-small cell lung cancer. *Anticancer Res* (2017) 37(2):377–87. doi: 10.21873/anticanres.11330

59. Kok VC, Yu CC. Cancer-derived exosomes: their role in cancer biology and biomarker development. *Int J Nanomed* (2020) 15:8019–36. doi: 10.2147/ijn.S272378

60. Wu L, Qu X. Cancer biomarker detection: recent achievements and challenges. *Chem Soc Rev* (2015) 44(10):2963–97. doi: 10.1039/c4cs00370e

61. Zhu J, Tang B, Lv X, Meng M, Weng Q, Zhang N, et al. Identifying apoptosis-related transcriptomic aberrations and revealing clinical relevance as diagnostic and prognostic biomarker in hepatocellular carcinoma. *Front Oncol* (2020) 10:519180. doi: 10.3389/fonc.2020.519180

62. Tang B, Zhu J, Li J, Fan K, Gao Y, Cheng S, et al. The ferroptosis and iron-metabolism signature robustly predicts clinical diagnosis, prognosis and immune microenvironment for hepatocellular carcinoma. *Cell Commun Signal* (2020) 18(1):174. doi: 10.1186/s12964-020-00663-1

63. Krasnov GS, Dmitriev AA, Snezhkina AV, Kudryavtseva AV. Deregulation of glycolysis in cancer: glyceraldehyde-3-phosphate dehydrogenase as a therapeutic target. *Expert Opin Ther Targets* (2013) 17(6):681–93. doi: 10.1517/14728222.2013.775253

64. Hansen CN, Ketabi Z, Rosenstierne MW, Palle C, Boesen HC, Norrild B. Expression of CPEB, GAPDH and U6snRNA in cervical and ovarian tissue during cancer development. *Apmis* (2009) 117(1):53–9. doi: 10.1111/j.1600-0463.2008.00015.x

65. Kim JW, Kim SJ, Han SM, Paik SY, Hur SY, Kim YW, et al. Increased glyceraldehyde-3-phosphate dehydrogenase gene expression in human cervical cancers. *Gynecol Oncol* (1998) 71(2):266–9. doi: 10.1006/gyno.1998.5195

66. Altenberg B, Greulich KO. Genes of glycolysis are ubiquitously overexpressed in 24 cancer classes. *Genomics* (2004) 84(6):1014–20. doi: 10.1016/j.ygeno.2004.08.010

67. Zheng J. Energy metabolism of cancer: Glycolysis versus oxidative phosphorylation (Review). *Oncol Lett* (2012) 4(6):1151–7. doi: 10.3892/ol.2012.928

68. Lea MA, Qureshi MS, Buxhoeveden M, Gengel N, Kleinschmit J, Desbordes C. Regulation of the proliferation of colon cancer cells by compounds that affect glycolysis, including 3-bromopyruvate, 2-deoxyglucose and biguanides. *Anticancer Res* (2013) 33(2):401–7.

69. Kim JW, Kim TE, Kim YK, Kim YW, Kim SJ, Lee JM, et al. Antisense oligodeoxynucleotide of glyceraldehyde-3-phosphate dehydrogenase gene inhibits cell proliferation and induces apoptosis in human cervical carcinoma cell lines. *Antisense Nucleic Acid Drug Dev* (1999) 9(6):507–13. doi: 10.1089/oli.1.1999.9.507

70. Shoshani T, Faerman A, Mett I, Zelin E, Tenne T, Gorodin S, et al. Identification of a novel hypoxia-inducible factor 1-responsive gene, RTP801, involved in apoptosis. *Mol Cell Biol* (2002) 22(7):2283–93. doi: 10.1128/mcb.22.7.2283-2293.2002

71. Tirado-Hurtado I, Fajardo W, Pinto JA. DNA damage inducible transcript 4 gene: the switch of the metabolism as potential target in cancer. *Front Oncol* (2018) 8:106. doi: 10.3389/fonc.2018.00106

72. Ding F, Gao F, Zhang S, Lv X, Chen Y, Liu Q. A review of the mechanism of DDIT4 serve as a mitochondrial related protein in tumor regulation. *Sci Prog* (2021) 104(1):36850421997273. doi: 10.1177/0036850421997273

73. Qiao S, Koh SB, Vivekanandan V, Salunke D, Patra KC, Zaganjor E, et al. REDD1 loss reprograms lipid metabolism to drive progression of RAS mutant tumors. *Genes Dev* (2020) 34(11-12):751–66. doi: 10.1101/gad.335166.119

74. Coulombe PA, Wong P. Cytoplasmic intermediate filaments revealed as dynamic and multipurpose scaffolds. *Nat Cell Biol* (2004) 6(8):699–706. doi: 10.1038/ncb0804-699

75. Zhang B, Wang J, Liu W, Yin Y, Qian D, Zhang H, et al. Cytokeratin 18 knockdown decreases cell migration and increases chemosensitivity in non-small cell lung cancer. *J Cancer Res Clin Oncol* (2016) 142(12):2479–87. doi: 10.1007/s00432-016-2253-x

76. Nagel M, Schulz J, Maderer A, Goepfert K, Gehrke N, Thomaidis T, et al. Cytokeratin-18 fragments predict treatment response and overall survival in gastric cancer in a randomized controlled trial. *Tumour Biol* (2018) 40(3):1010428318764007. doi: 10.1177/1010428318764007

77. Golob-Schwarzl N, Bettermann K, Mehta AK, Kessler SM, Unterluggauer J, Krassnig S, et al. High keratin 8/18 ratio predicts aggressive hepatocellular cancer phenotype. *Transl Oncol* (2019) 12(2):256–68. doi: 10.1016/j.tranon.2018.10.010

78. Bozza WP, Zhang Y, Zhang B. Cytokeratin 8/18 protects breast cancer cell lines from TRAIL-induced apoptosis. *Oncotarget* (2018) 9(33):23264–73. doi: 10.18632/oncotarget.25297

79. Sullivan BT, Cherry JA, Sakamoto H, Henkes LE, Townson DH, Rueda BR. Cytokeratin 18 expression inhibits cytokine-induced death of cervical cancer cells. *Int J Gynecol Cancer* (2010) 20(9):1474–81. doi: 10.1111/IGC.0b013e3181fc3a03

80. Gatenby RA, Gillies RJ. Why do cancers have high aerobic glycolysis? *Nat Rev Cancer* (2004) 4(11):891–9. doi: 10.1038/nrc1478

81. Hsiao KC, Shih NY, Fang HL, Huang TS, Kuo CC, Chu PY, et al. Surface α-enolase promotes extracellular matrix degradation and tumor metastasis and represents a new therapeutic target. *PloS One* (2013) 8(7):e69354. doi: 10.1371/journal.pone.0069354

82. Díaz-Ramos A, Roig-Borrellas A, García-Melero A, López-Alemany R. α-Enolase, a multifunctional protein: its role on pathophysiological situations. *J BioMed Biotechnol* (2012) 2012:156795. doi: 10.1155/2012/156795

83. Fu QF, Liu Y, Fan Y, Hua SN, Qu HY, Dong SW, et al. Alpha-enolase promotes cell glycolysis, growth, migration, and invasion in non-small cell lung cancer through FAK-mediated PI3K/AKT pathway. *J Hematol Oncol* (2015) 8(8):22. doi: 10.1186/s13045-015-0117-5

84. Zhan P, Zhao S, Yan H, Yin C, Xiao Y, Wang Y, et al. α-enolase promotes tumorigenesis and metastasis *via* regulating AMPK/mTOR pathway in colorectal cancer. *Mol Carcinog* (2017) 56(5):1427–37. doi: 10.1002/mc.22603

85. Chang GC, Liu KJ, Hsieh CL, Hu TS, Charoenfuprasert S, Liu HK, et al. Identification of alpha-enolase as an autoantigen in lung cancer: its overexpression is associated with clinical outcomes. *Clin Cancer Res* (2006) 12(19):5746–54. doi: 10.1158/1078-0432.Ccr-06-0324

86. Qian X, Xu W, Xu J, Shi Q, Li J, Weng Y, et al. Enolase 1 stimulates glycolysis to promote chemoresistance in gastric cancer. *Oncotarget* (2017) 8(29):47691–708. doi: 10.18632/oncotarget.17868

87. Hsu SK, Li CY, Lin IL, Syue WJ, Chen YF, Cheng KC, et al. Inflammation-related pyroptosis, a novel programmed cell death pathway, and its crosstalk with immune therapy in cancer treatment. *Theranostics* (2021) 11(18):8813–35. doi: 10.7150/thno.62521

88. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. *Immunity* (2018) 48(4):812–830.e14. doi: 10.1016/j.immuni.2018.03.023

89. Olingy CE, Dinh HQ, Hedrick CC. Monocyte heterogeneity and functions in cancer. *J Leukoc Biol* (2019) 106(2):309–22. doi: 10.1002/JLB.4RI0818-311R

90. Bernsmeier C, van der Merwe S, Périanin A. Innate immune cells in cirrhosis. *J Hepatol* (2020) 73(1):186–201. doi: 10.1016/j.jhep.2020.03.027

91. Yin M, Li X, Tan S, Zhou HJ, Ji W, Bellone S, et al. Tumor-associated macrophages drive spheroid formation during early transcoelomic metastasis of ovarian cancer. *J Clin Invest* (2016) 126(11):4157–73. doi: 10.1172/JCI87252

92. Nie Y, Huang H, Guo M, Chen J, Wu W, Li W, et al. Breast phyllodes tumors recruit and repolarize tumor-associated macrophages *via* secreting CCL5 to promote Malignant progression, which can be inhibited by CCR5 inhibition therapy. *Clin Cancer Res* (2019) 25(13):3873–86. doi: 10.1158/1078-0432.CCR-18-3421

93. Sokratous G, Polyzoidis S, Ashkan K. Immune infiltration of tumor microenvironment following immunotherapy for glioblastoma multiforme. *Hum Vaccin Immunother* (2017) 13(11):2575–82. doi: 10.1038/s41568-018-0081-9

94. Altorki NK, Markowitz GJ, Gao D, Port JL, Saxena A, Stiles B, et al. The lung microenvironment: an important regulator of tumour growth and metastasis. *Nat Rev Cancer* (2019) 19(1):9–31. doi: 10.1038/s41568-018-0081-9

95. Fu C, Jiang A. Dendritic cells and CD8 T cell immunity in tumor microenvironment. *Front Immunol* (2018) 9:3059. doi: 10.3389/fimmu.2018.03059

96. Togashi Y, Shitara K, Nishikawa H. Regulatory T cells in cancer immunosuppression - implications for anticancer therapy. *Nat Rev Clin Oncol* (2019) 16(6):356–71. doi: 10.1038/s41571-019-0175-7

97. Sabado RL, Balan S, Bhardwaj N. Dendritic cell-based immunotherapy. *Cell Res* (2017) 27(1):74–95. doi: 10.1038/cr.2016.157

98. van der Burg SH. Correlates of immune and clinical activity of novel cancer vaccines. *Semin Immunol* (2018) 39:119–36. doi: 10.1016/j.smim.2018.04.001

99. Teixeira AKS, Vasconcelos JLA. Histopathological profile of patients diagnosed with Malignant tumors assisted in a hospital of reference of Agreste Pernambucano. *J Brasileiro Patol e Med Laboratorial* (2019) 55(1). doi: 10.5935/1676-2444.20190002

100. Davoli T, Uno H, Wooten EC, Elledge SJ. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* (2017) 355(6322):eaaf8399. doi: 10.1126/science.aaf8399

101. Lai Y, Wang X, Zhou K, Su J, Che G. The feasibility and safety of no placement of urinary catheter following lung cancer surgery: A retrospective cohort study with 2,495 cases. *J Invest Surg* (2021) 34(6):568–74. doi: 10.1080/08941939.2019.1663377

102. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* (2015) 348(6230):124–8. doi: 10.1126/science.aaa1348

103. Chae YK, Viveiros P, Lopes G, Sukhadia B, Sheikh MM, Saravia D, et al. Clinical and immunological implications of frameshift mutations in lung cancer. *J Thorac Oncol* (2019) 14(10):1807–17. doi: 10.1016/j.jtho.2019.06.016

104. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med* (2015) 372(26):2509–20. doi: 10.1056/NEJMoa1500596

105. Xu F, Zhang H, Chen J, Lin L, Chen Y. Immune signature of T follicular helper cells predicts clinical prognostic and therapeutic impact in lung squamous cell carcinoma. *Int Immunopharmacol* (2020) 81:105932. doi: 10.1016/j.intimp.2019.105932

106. Nowak AK, Robinson BW, Lake RA. Synergy between chemotherapy and immunotherapy in the treatment of established murine solid tumors. *Cancer Res* (2003) 63(15):4490–6.

107. Goł b J, Zagozdzon R, Kamiński R, Kozar K, Gryska K, Izycki D, et al. Potentiatied antitumor effectiveness of combined chemo-immunotherapy with interleukin-12 and 5-fluorouracil of L1210 leukemia in vivo. *Leukemia* (2001) 15(4):613–20. doi: 10.1038/sj.leu.2402076

108. Johnson RK, McCabe FL, Caulder E, Innlow L, Whitacre M, Winkler JD, et al. SB-715992, a potent and selective inhibitor of the mitotic kinesin KSP, demonstrates broad-spectrum activity in advanced murine tumors and human tumor xenografts (Abstract). *Annu Meet Am Assoc Cancer Res Proc* (2002) 43:A1355.

109. de Bono JS, Oudard S, Ozguroglu M, Hansen S, Machiels JP, Kocak I, et al. Prednisone plus cabazitaxel or mitoxantrone for metastatic castration-resistant prostate cancer progressing after docetaxel treatment: a randomised open-label trial. *Lancet* (2010) . 376(9747):1147–54. doi: 10.1016/S0140-6736(10)61389-X

110. Vrignaud P, Sémiond D, Lejeune P, Bouchard H, Calvet L, Combeau C, et al. Preclinical antitumor activity of cabazitaxel, a semisynthetic taxane active in taxane-resistant tumors. *Clin Cancer Res* (2013) 19(11):2973–83. doi: 10.1158/1078-0432.CCR-12-3146

111. Huo R, Wang L, Liu P, Zhao Y, Zhang C, Bai B, et al. Cabazitaxel-induced autophagy *via* the PI3K/Akt/mTOR pathway contributes to A549 cell death. *Mol Med Rep* (2016) 14(4):3013–20. doi: 10.3892/mmr.2016.5648

112. Zhuang B, Du L, Xu H, Xu X, Wang C, Fan Y, et al. Self-assembled micelle loading cabazitaxel for therapy of lung cancer. *Int J Pharm* (2016) 499(1-2):146–55. doi: 10.1016/j.ijpharm.2015.12.073

113. Rogalska A, Marczak A. Epothilone B induces human ovarian cancer OV-90 cell apoptosis *via* external pathway. *Environ Toxicol Pharmacol* (2015) 39(2):700–12. doi: 10.1016/j.etap.2015.01.023

114. Vacchelli E, Senovilla L, Eggermont A, Fridman WH, Galon J, Zitvogel L, et al. Trial watch: Chemotherapy with immunogenic cell death inducers. *Oncoimmunology* (2013) 2(3):e23510. doi: 10.4161/onci.23510

115. Baumgart T, Klautke G, Kriesen S, Kuznetsov SA, Weiss DG, Fietkau R, et al. Radiosensitizing effect of epothilone B on human epithelial cancer cells. *Strahlentherapie Und Onkol* (2012) 188(2):177–84. doi: 10.1007/s00066-011-0029-4

116. Ashton TM, McKenna WG, Kunz-Schughart LA, Higgins GS. Oxidative phosphorylation as an emerging target in cancer therapy. *Clin Cancer Res* (2018) 24(11):2482–90. doi: 10.1158/1078-0432.Ccr-17-3070

117. Hopkins JL, Lan L, Zou L. DNA repair defects in cancer and therapeutic opportunities. *Genes Dev* (2022) 36(5-6):278–93. doi: 10.1101/gad.349431.122

118. Xu Y, Xu J, Qiao R, Zhong H, Xia J, Zhong R. Loss of BLK expression as a potential predictor of poor prognosis and immune checkpoint blockade response in NSCLC and contribute to tumor progression. *Transl Oncol* (2023) 33:101671. doi: 10.1016/j.tranon.2023.101671

119. Lin S, Li Y, Wang D, Huang C, Marino D, Bollt O, et al. Fascin promotes lung cancer growth and metastasis by enhancing glycolysis and PFKFB3 expression. *Cancer Lett* (2021) 518:230–42. doi: 10.1016/j.canlet.2021.07.025

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 25 September 2023
DOI 10.3389/fimmu.2023.1187108

OPEN ACCESS

EDITED BY
Jun Liu,
Yuebei People's Hospital, China

REVIEWED BY
Ke Han,
Harbin University of Commerce, China
Ya-Qin Wang,
Sun Yat-sen University Cancer Center (SYSUCC), China
Hao Chi,
Southwest Medical University, China

*CORRESPONDENCE
Hong Qiu
qiuhong@hust.edu.cn

†These authors have contributed equally to this work

RECEIVED 15 March 2023
ACCEPTED 07 September 2023
PUBLISHED 25 September 2023

CITATION
Hu P, Xu L, Liu Y, Zhang X, Li Z, Li Y and Qiu H (2023) Identification of molecular pattern and prognostic risk model based on ligand-receptor pairs in liver cancer. *Front. Immunol.* 14:1187108. doi: 10.3389/fimmu.2023.1187108

Identification of molecular pattern and prognostic risk model based on ligand-receptor pairs in liver cancer

Pengbo Hu†, Liang Xu†, Yongqing Liu, Xiuyuan Zhang, Zhou Li, Yiming Li and Hong Qiu*

Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China

Introduction: The tumor microenvironment of hepatocellular carcinoma is composed of multiple cells, and the interactive communication between cells drives tumor progression and characterizes the tumor. Communication between cells is mainly achieved through signal transduction between receptor ligands, and the rise of single-cell technology has made it possible to analyze the communication network between cells.

Methods: We applied a train of bioinformatic techniques and in vitro experiments. We analyzed the composition of the microenvironment of liver cancer by combining single-cell sequencing data and transcriptome sequencing data from liver cancer to construct molecular typing and risk models for LRs. Then, we analyzed association of it with prognosis, mutation, KEGG, tumor microenvironment (TME), immune infiltration, tumor mutational burden (TMB) and drug sensitivity in liver cancer. qPCR and was used to identify SLC1A5 expression in LIHC cell lines and CCK8, transwell and cell colony formation were performed to validate the function of SLC1A5. Meanwhile, we also performed polarization of macrophages.

Results: In this experiment, we found that liver cancer tissues are rich in immune and mesenchymal cells, and there is extensive signaling between individual cells, so we constructed molecular typing and risk models for LRs. Combining clinical data revealed significant differences in clinical characteristics, prognosis and mutated genes between the molecular typing of receptor-ligand pairs, as well as in sensitivity to drugs; similarly, there were significant prognostic differences between the risk models. There were also notable differences in activated signaling pathways, infiltrating immune cells and immune subtypes. Subsequently, we used siRNA to knock down SLC1A5 in hepatocellular carcinoma cells and found that cell proliferation, migration and invasion were diminished.

Conclusions: In conclusion, our LRs model may become a marker to guide clinical treatment and prognosis.

KEYWORDS

liver cancer, ligand-receptor, molecular pattern, risk model, TME

Introduction

HCC ranks as the sixth most prevalent new tumor in the world, but the second most common new death (1). Due to the insidious character of liver cancer, the detection rate of early-stage liver cancer is quite poor, while the highly malignant characteristics of liver cancer frequently result in the identification of liver cancer at an advanced stage. The prognosis for advanced liver cancer is terrible, with a lack of surgical opportunities and a high risk of metastasis within and outside the liver, resulting in worse quality of living and a shorter survival time for patients (2).

The tumor microenvironment, a hothouse of research in recent years, is composed of tumor cells, immune cells and stromal cells that together modulate tumor growth, metastasis, drug resistance and other properties (3). In the last few years, the rapid development of single-cell sequencing has made it possible for scientists to peek into the details of the tumor microenvironment (4). Through the use of gene expression profiles of individual cells, we can identify the specific roles played by different cell types in the tumor microenvironment, refining the function of cells to investigate the microenvironment in a specific target (5).

Cell-to-cell signaling moderates the function and state of the cell. With the continuous enhancement of precision in the research of tumors, receptor-ligand interactions to deliver signals between tumor tissues have attracted the attention of scientists (6). In particular, in the tumor microenvironment, a majority of the interactions between tumor cells, immune cells and stromal cells are mediated by receptors and ligands, of which investigations on PD1 and PD-L1 have made a significant contribution to clinical diagnosis and treatment (7).

Currently, there is no effective method to forecast the prognosis of patients with hepatocellular carcinoma, but receptor-ligand interactions could, up to some extent, anticipate the malignancy of the tumor and thus predict the prognosis of patients. Consequently, we developed a receptor-ligand pairs (LRs) model based on hepatocellular carcinoma to predict the risk of patients and ultimately improve their survival.

Materials and methods

Datasets

The single cell sequencing data (GSE146115) was obtained from the GEO database, the gene expression profiling data, clinical data and mutation data from TCGA and ICGC. The ICGC-LIHC sample (231) set was regarded as the external validation set, and the TCGA-LIHC sample (365) set as the training set.

scRNA-seq data analysis and cell type definition

"Seurat" R package was applied to analyze the expression matrix of single cells, and we screened for cells with the optimal number of genes expressed (50~20000) (8). The mitochondrial genes have been removed from the expression matrix. After controlling for average expression and dispersion relationships, all highly variable genes in single cells were identified. Subsequently, the highly variable genes were used to perform principal component analysis to identify significant principal components as a means of eliminating batch effects based on the "jackstraw" function. Cells were cohorted into 12 different cell types at a resolution of 0.5 using the "FindClusters" function. The "FindAllMarkers" function was applied to discern differentially expressed genes (DEGs). Additionally, a few traditional markers for defining cell subsets were gathered from earlier studies (Table 1) and manually annotated in accordance with marker expression.

Cell to cell communication

The various cells of the tumor microenvironment interact with each other to exert tumor promoting or inhibiting effects through activation between various ligands and receptors. Cellular communication was accomplished through "cellphonedb", a public database containing ligands, receptors, and their interactions, and by annotating the membrane, secreted, and peripheral proteins of each cell subgroup at various time points (9). We conducted the "cellphonedb" to unpack the matrix of cellular communication and we filtered the receptors and ligands that appeared in the expression matrix in TCGA.

Selection of receptor-ligand pairs

Cellular interactions depend on the simultaneous expression of receptors and ligands, and receptors can only be stimulated to mediate intercellular communication when the number of receptors and ligands is at a parallel elevation. We screened for LRs with receptor-ligand co-expression correlations greater than 0.3 ($p < 0.05$) and used these LRs for clustering to determine molecular types (10).

Molecular subtyping calculation

We took the sum of the gene expression of the receptor and ligand as the expression of the LR. We combined the receptor-ligand expression data with the prognostic data to filter out LRs

TABLE 1 The marker genes for the cells.

Cell subgroups	Markers
B cells	CD79A, CD79B
Fiberblast cells	ACTA2
Hepatocytes	CYP2C9, ARG1
Myeloid cells	CD68, CD163
NK cells	NCAM1, GNLY
T cells	CD3D, CD3E

with prognostic significance and then used the R package "ConensusClusterPlus" to generate consensus matrix, "euclidean" was selected as the distance metric for PAM algorithms. A random subset of the TCGA data was selected from the TCGA data, the size of the subset was 80% of the original data set and 500 replications were performed (11). The amount of clusters was varied ranging from 2 to 10 and the most appropriate number of clusters was determined through the calculation of the consensus matrix and the cumulative distribution function (CDF).

Presumption of drug sensitivity

The R package "pRRophetic" was applied to calculate the sensitivity of different drugs based on expression matrix (12). We determined the appropriate drugs for different classifications of patients based on the sensitivity of the drug, as well as the classification of the different patients.

Risk model

The R package "glmnet" was conducted to screen for LRs using lasso regression, combined with prognostic data, to build a risk model to predict the risk of patients by classifying them into different risk groups using the median risk score. First, we chose "cv.glmnet" fuction to filter the λ that minimised the discrepancy. Then, based on the value of λ taken at this point, we got to filter the best LRs to build the prognostic model. The risk score was calculated using the formula LR. Score = $\Sigma\beta(i) \times Exp(i)$, where i refers to the LR pair, Exp represents the level of LR pair expression and beta is the coefficient of the LR pair in the model. The median value of the training set (TCGA) was selected as the truncation value (13). The R package "survival" was employed to depict survival curves and to compare survival differences between high and low risk groups. The R package "timeROC" to portray the ROC curves of risk scores and traditional prognostic indicators, and to calculate the AUC values to assess the accuracy of the prognosis prediction (14).

Function enrichment and analysis of mutations and immunity

The "maftools" package was used to visualize the mutation data. We presented the twenty genes with the most significant mutations and compared mutations in patients of different subtypes. We applied "clusterProfiler" to analysis KEGG pathway (15). The 'hallmark' gene set collection from the molecular signature database was used for pathway enrichment analysis. Immune cell infiltration assessment was carried out with the "ssGSEA" package (16).

In vitro experimental validation

The cells used in this experiment were obtained from the laboratory of the Department of Oncology, Tongji Hospital, Huazhong University of Science and Technology. LO2, SNU398, Huh7 and HLF were cultured in DMEM medium with 10% fetal bovine serum added. We extracted RNA from the cells, reverse transcribed them into DNA and qPCR detected the expression of SLC1A5 in the cells. Subsequently, SLC1A5 was knocked down in Huh7 and SNU398 cells using siRNA, and CCK8, cell colony formation and transwell were used to detect the proliferation, migration and invasion ability of the cells (17). Meanwhile, we also performed polarization of macrophages. Detailed experimental steps are in the Supplementary Data Sheet 1, and all experiments were repeated three times.

Statistical analysis

Dichotomous variables were tested using the chi-square test, survival analysis was performed using the log-rank test, and comparison between the two groups was performed using the Wilcoxon test. $p<0.05$ was considered to be statistically different.

Results

The single-cell transcriptome landscape of hepatocellular carcinoma

Figure 1 showed the overall design and flow chart of this study. Since the gene expression data from single cells excluded mitochondrial genes, we calculated the correlation between the number of unique molecular identifiers and mRNA, which suggests that a significant positive correlation was shown between the number of unique molecular identifiers and mRNA (Figure 2A). Gene numbers for the vast majority of cells are between 0 and 8000 (Figure 2B). After filtering the cells, a total of 3200 cells were included in the subsequent analysis. The differential genes in the various cell types were calculated for a total of 12 cell types after normalizing the expression data and filtering the first 2000 highly variable genes for the subsequent principal component analysis (Supplementary Figure 1A, Figure 2C). Based on the marker genes in the different cells, we identified the cell types that needed to be labelled. The marker genes for the cells were derived from databases and earlier studies. As a result, six cells were identified, such as B cells, fibroblast cells, hepatocytes, myeloid cells, NK cells, T cells (Figure 2D).

Intercellular communication networks in hepatocellular carcinoma

The tumor microenvironment in solid tumors is composed primarily of stromal cells and immune cells, in addition to the tumor cells themselves. Within the microenvironment, various cells communicate with each other to transmit information to influence tumor progression. When speculating on intercellular communication, we used "cellphonedb" and found that tumor cells contact mainly with fibroblasts and fibroblasts with myeloid cells in

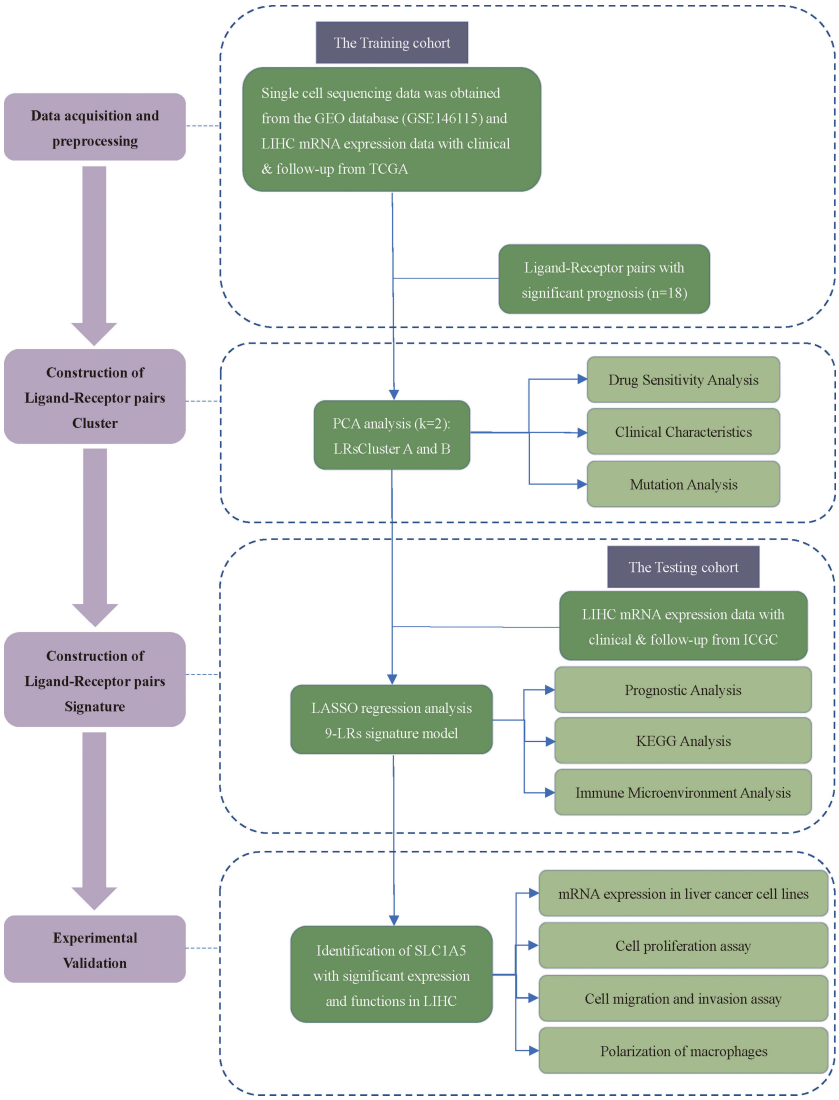


FIGURE 1
The flowchart of the study design.

liver cancer (Figure 2E). In Figure 2F, thicker lines indicate more interactions, and the numbers on the lines represent the number of interacting nodes.

Molecular typing based on ligand-receptor pairs

In cellular communication, receptor-ligand interactions play an essential role, and through "cellphonedb" simulations we obtained the corresponding LRs. Based on the TCGA expression profile and taking into account the need for synergistic expression of receptors and ligands, we chose LRs with co-expression R-values above 0.3 and P-values less than 0.05. 81 LRs were selected in total (Supplementary Table 1). The sum of the expression values of the receptor and the ligand took the place of the LRs' expression values. Molecular typing analysis was executed on the sample set of TCGA using the "ConsensusClusterPlus". Based on the CDF value (Figure 3A), splitting into two clusters was the preferred candidate when the k value was taken as 2 (Figures 3B, C). The ICGC validation dataset was subjected to the same data processing procedure as the TCGA training dataset, and the outcomes were identical. Similar distinctions between the two groups of patients were made based on the LR expression pattern (Supplementary Figures 2A–C).

Comparison of clinical information in different molecular subtypes

According to the results of the clustering, combined with the prognosis analysis of the patients, we analyzed the survival of the patients in both clusters and the survival curves depicted are shown in the figure. The log-rank test showed that there is a significant difference in survival between the two clusters, with patients in the A subtype having a better prognosis than those in the B subtype in TCGA (Figure 3D). The outcomes of the prognostic analyses in the

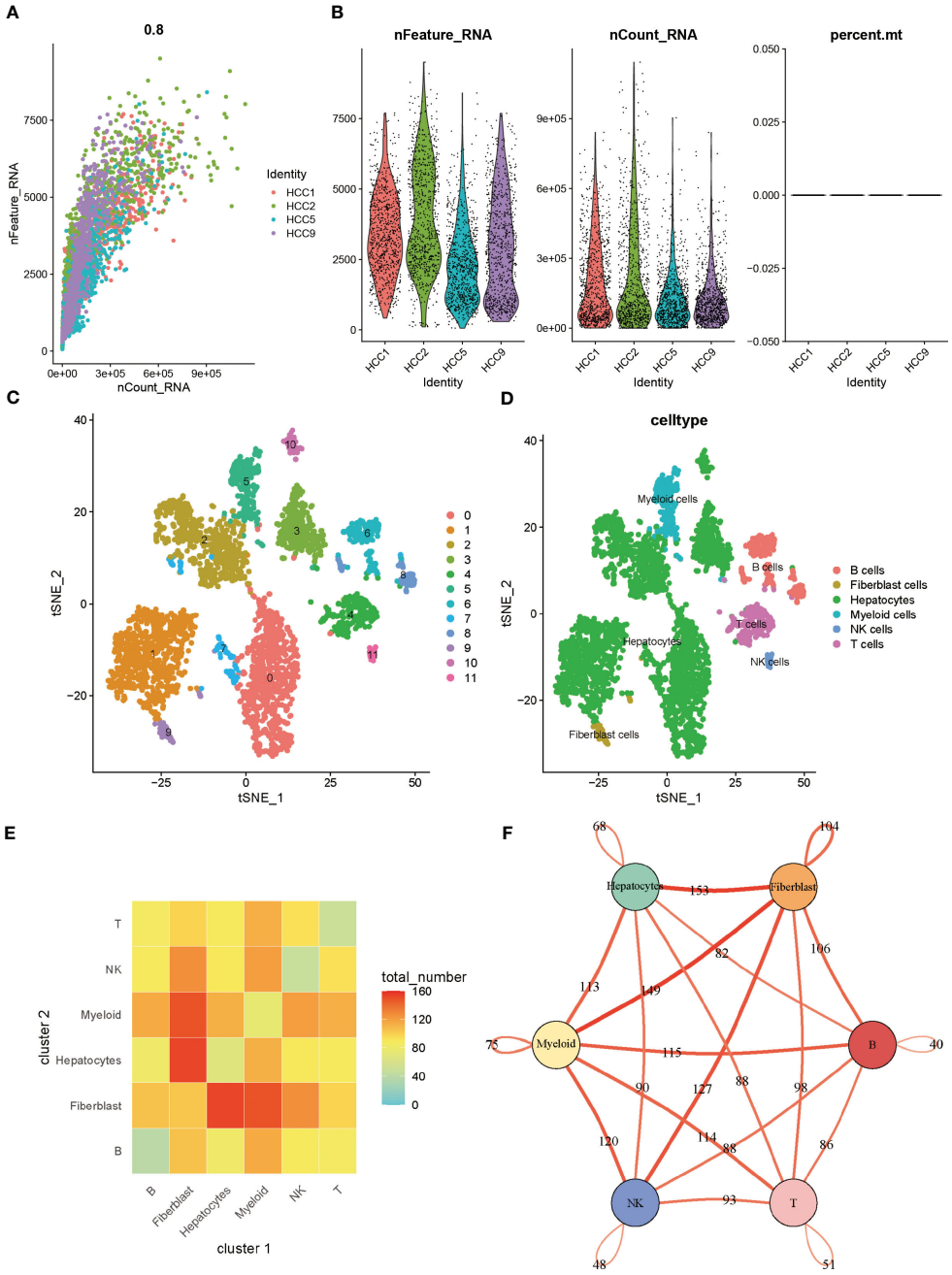


FIGURE 2
(A) Correlation between the number of unique molecular identifiers and mRNA **(B)** Violin plot of features **(C)** tSNE of cell profiles; different color blocks represent related cell clusters. **(D)** tSNE of cell profiles; different color blocks represent related sample sources. **(E)** LR interactions between different cell subsets. **(F)** Network overview for the interaction between different cell subsets.

ICGC validation set were agreed with those in the TCGA training set. However, the number of patients in B subtype was too small, perhaps as a result of the limited sample size of patients in ICGC, leading to a P value (0.052) of more than a little over 0.05 (Supplementary Figure 2D). Meanwhile, we went on to analyze the gene expression and clinical characteristics of patients in different clusters and found that the expression levels of risk genes were significantly higher in B subtype patients than in A subtype (Figure 4A). Regarding the clinical traits, we discovered that patients in the B subtype had greater tumor grade and TNM stage as well as more fatalities than those in the A subtype (Figures 4B–D).

Mutational characteristics of different molecular subtypes

Genetic alterations in tumor cells, mutational inactivation of anti-oncogenes, amplified overexpression of oncogenes, etc., all of which contribute to tumor development, are intimately associated

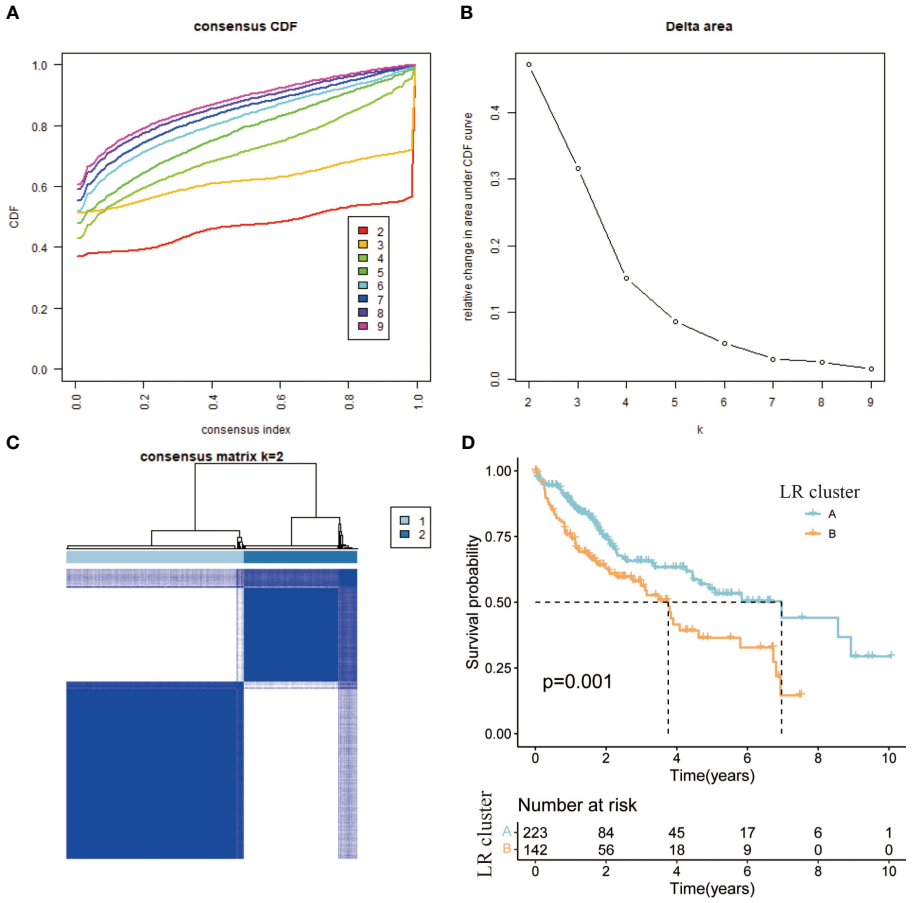


FIGURE 3
(A) CDF curve of samples from TCGA cohort. **(B)** Delta area curve of samples from TCGA cohort. **(C)** TCGA clustering heatmap of samples from TCGA cohort when consensus k = 2. **(D)** Overall survival curves of molecular subtypes based on LR pairs.

to tumor growth. When we analyzed the mutation data of patients in A and B subtype patients, we found that there were more mutations in the A subtype patients (Figures 4F, G). The mutations in the B subtype patients were mainly in the antioncogenes and the types of mutations were mainly missense mutations, such as TP53 (Figure 4G).

We then analyzed the tumor mutational burden of patients to predict the efficacy of immunotherapy in patients, and we found that the tumor mutational burden (TMB) was higher in A subtype patients than in B subtype patients (Figure 4E), which means that the effectiveness of immunotherapy may be better in A subtype patients than in B subtype patients.

Drug sensitivity of different molecular subtypes

We also attempted to see whether there were changes in drug sensitivity amongst the various patient clusters, in addition to examining differences in clinical features and gene expression. Using the R package "pRRophetic" to predict patient sensitivity to chemotherapeutic drugs, we analyzed prominent liver cancer drugs and found that sorafenib, a first-line drug for liver cancer, was more efficacious in B subtype patients. At the same time, we observed that some drugs were more effective in A subtype patients, such as "Bleomycin" (Figure 5A)," Doxorubicin" (Figure 5B), "Gemcitabine" (Figure 5C), "Mitomycin" (Figure 5D), "Paclitaxel" (Figure 5F), and conversely, some drugs were more targeted in B subtype, such as "Methotrexate" (Figure 5E), "Rapamycin" (Figure 5G), "Sorafenib" (Figure 5H), "Temsirolimus" (Figure 5I).

Establishment of the risk model based on the ligand–receptor pair score

Molecular subtypes based on LRs had different clinicopathological characteristics as well as drug sensitivity. We sought to construct a prognostic model using LRs to assess patient risk. A total of 18 LRs associated with prognosis were considered (Supplementary Figure 1B), and we used LASSO cox regression to shrink the number of LRs, with optimal results occurring when λ was 0.0159 Figures 6A, B). We then used stepwise multivariate regression for optimization, and a total of nine LRs were screened out for model construction (Supplementary Table 2). In the training set (TCGA), the median risk score was selected as the cut-off value to classify patients into high and low risk groups, while patients in the validation set

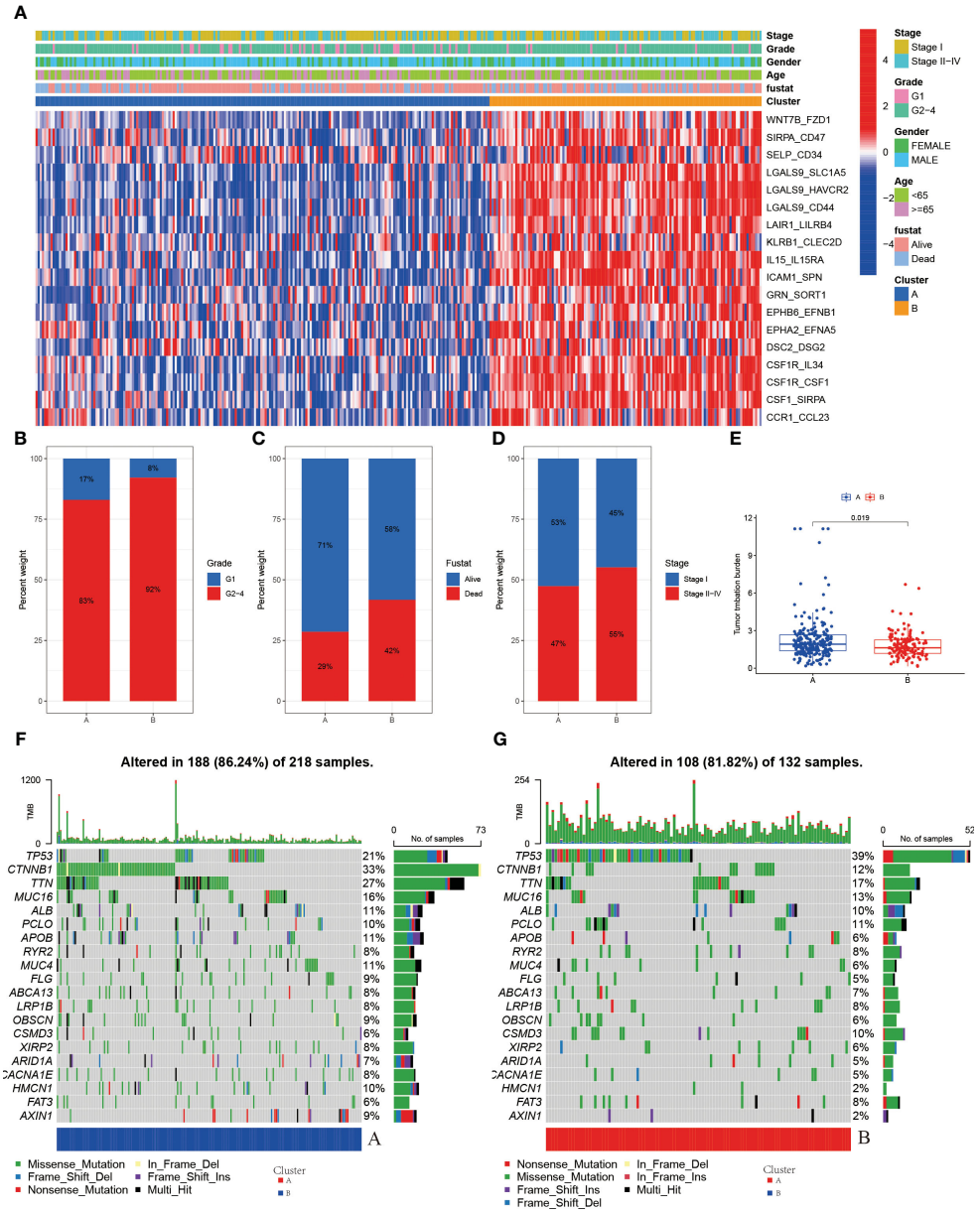


FIGURE 4
(A) Heatmap of expression of LRs in molecular subtypes **(B)** Grade in molecular subtypes **(C)** Fustat in molecular subtypes **(D)** Stage in molecular subtypes **(E)** TMB in molecular subtypes **(F)** Somatic mutation variation analysis in A cluster in the TCGA-LIHC cohort. **(G)** Somatic mutation variation analysis in B cluster in the TCGA-LIHC cohort.

(ICGC) were similarly classified into high and low risk groups using the cut-off value in the training set as the boundary.

Correlation between the risk model and clinical features

Based on the risk groupings, we plotted the survival curves for the training and validation sets separately and used log-rank test to analyze whether there was a difference in survival, and we found that the high-risk group had a worse prognosis in both datasets (Figures 6C, D). We also plotted ROC curves based on risk scores and clinicopathological characteristics, and calculated AUC values for different factors separately, and observed that risk scores performed better in predicting patient risk in the training and validation sets (Figures 6E, F). In the TCGA training set, the risk model's sensitivity was 0.784 and its specificity was 0.652; in the ICGC validation set, these values were 0.823 and 0.744, respectively. Both univariate and multivariate analyses showed that the risk score was an independent risk factor to assess patients' prognosis in TCGA database (Figures 6G, H). In the validation set, we similarly found that risk score was also an independent prognostic factor in both univariate and multivariate analyses (Figures 6I, J).

The results of the principal components analysis showed that the grouping of high and low risk was able to separate the characteristics of the patients in TCGA (Supplementary Figure 3A),

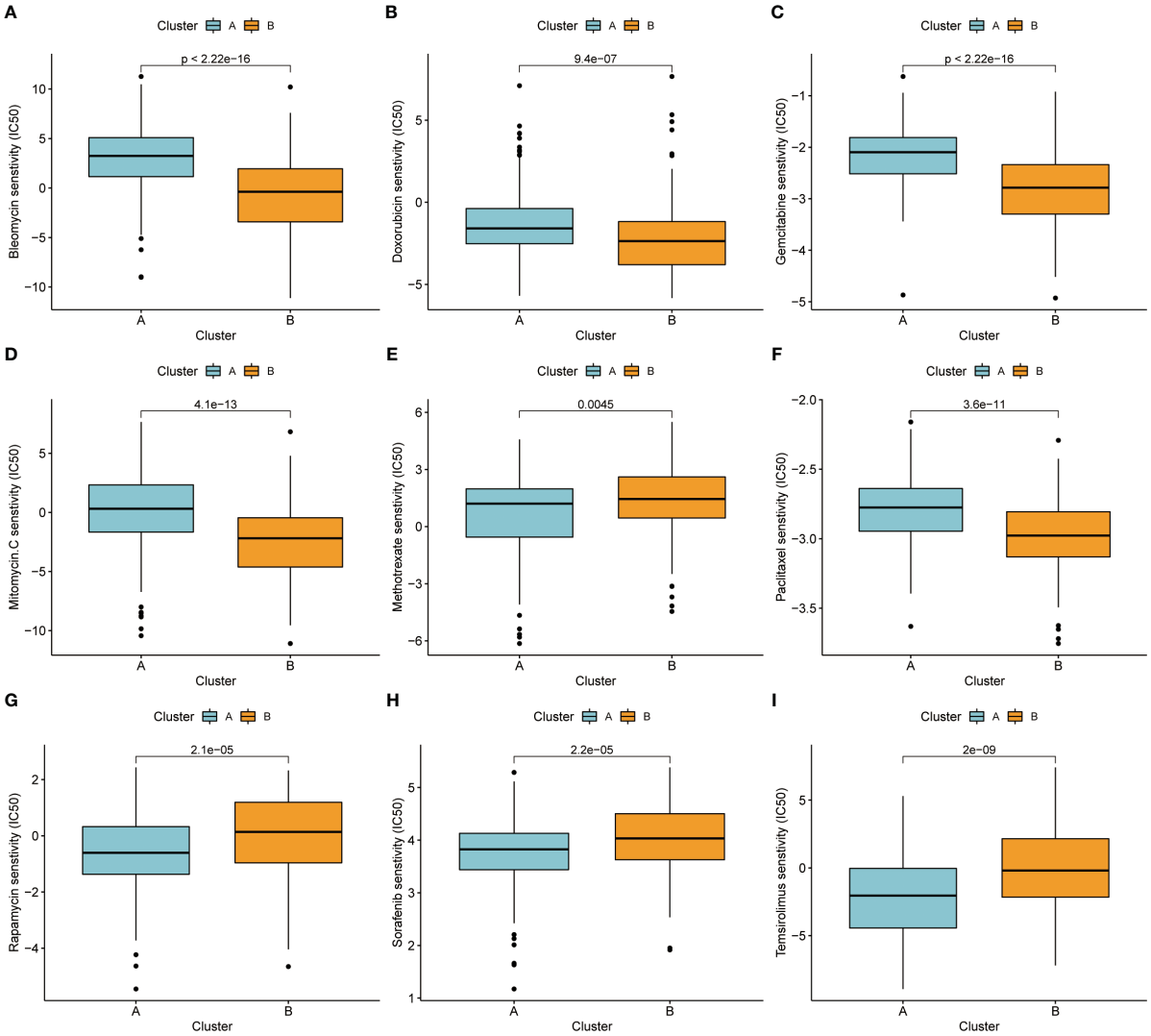


FIGURE 5
(A) Bleomycin. **(B)** Doxorubicin. **(C)** Gemcitabine. **(D)** Mitomycin. **(E)** Methotrexate. **(F)** Paclitaxel. **(G)** Rapamycin. **(H)** Sorafenib. **(I)** Temsirolimus.

and the same phenomenon was observed in the validation set (Supplementary Figure 3B). Similarly, we found that the risk scores of surviving patients were significantly lower than those of the deceased in the training set (Figure 6K), and the same was witnessed in ICGC (Figure 6L). This suggested that the risk score in the prognostic model was highly effective in predicting patient survival, and that patients could be separated according to different risk groupings. In the training set, the proportion of patients who died increased as the risk score increased (Supplementary Figures 3C, D). Similarly, we observed the same phenomenon in the training set (Supplementary Figures 3E, F).

Pathway analysis and immune characterization of risk model

Additionally, we conducted an enrichment analysis of patients in the high-risk and low-risk groups and found that compared to low-risk patients, high-risk patients mainly activate a number of immune-related, metabolism-related and hypoxia-related pathways (Figures 7A, B).

We conducted the R package "ssGSEA" to analyze the immune infiltration of patients and we found a decrease in the infiltration of B cells, CD8+ T cells, mast cells, NK cells and, at the same time, an increase in the infiltration of activated DC cells and macrophages in high-risk patients (Figure 7C). The reduction of tumor-killing immune cells and addition of tumor-promoting immune cells may be one of the reasons for the worse prognosis in high-risk patients. Combined with the immune phenotyping of the patients, we found that the immune subtypes in the high-risk group were focused on C1 & C2 and in the low-risk group were mostly in C3 (Figure 7D).

We compared the levels of PD-L1 expression in the high and low risk groups and discovered that patients in the high-risk group had higher levels of PD-L1 expression (Figure 7F). This finding is also consistent with what we found of the ssGSEA above, which

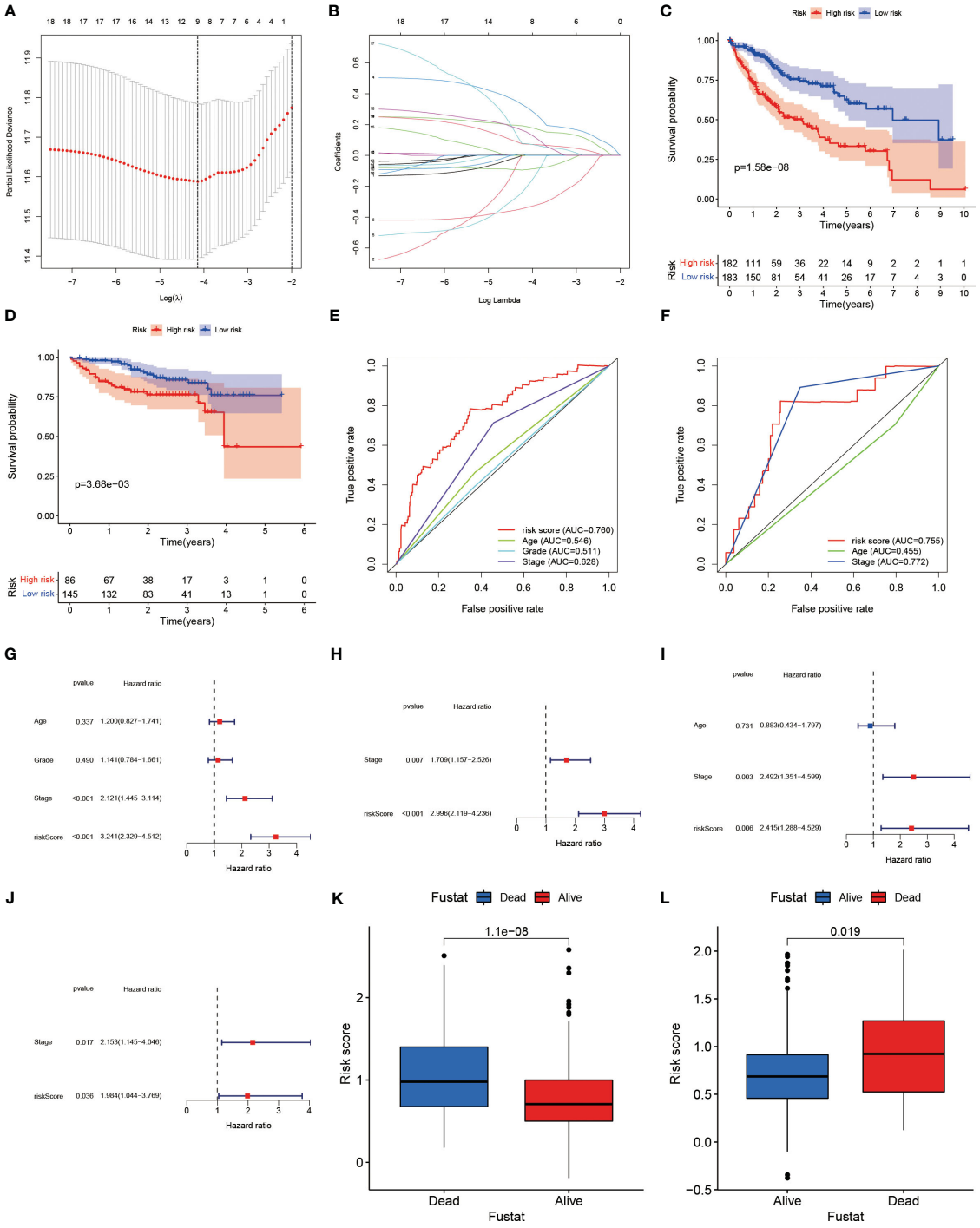


FIGURE 6
(A) Plotting of multinomial deviance versus log(λ). **(B)** LASSO coefficient profiles of the LRs. **(C)** Survival benefit of LR.score in the high and low LR.score groups in the TCGA-LIHC cohort. **(D)** Survival benefit of LR.score in the high and low LR.score groups in the ICGC-LIHC cohort. **(E)** The predictive value of LR.score in patients among the TCGA-LIHC cohort. **(F)** The predictive value of LR.score in patients among the ICGC-LIHC cohort. **(G)** Univariate cox regression analysis of LR.score, age, TNM stage and grade for overall survival (OS) in the TCGA-LIHC cohort. **(H)** Multivariate cox regression analysis of LR.score, and TNM stage for OS in the TCGA-LIHC cohort. **(I)** Univariate cox regression analysis of LR.score, age and TNM stage for overall survival (OS) in the ICGC-LIHC cohort. **(J)** Multivariate cox regression analysis of LR.score, and TNM stage for OS in the ICGC-LIHC cohort. **(K)** Fustat in the high and low LR.score groups in the TCGA-LIHC cohort. **(L)** Fustat in the high and low LR.score groups in the ICGC-LIHC cohort.

revealed that the high-risk group was primarily in an immunosuppressed state. As a result, the high-risk group may be more responsive to immunotherapeutic treatment targeting PD-1/PD-L1.

The TIDE scores are one of the basis for evaluating the current immunotherapy, so we also looked at the TIDE scores of the high and low risk groups. As a result, we discovered that there was a difference in the TIDE scores of the high and low risk groups (Figure 7E). Lower

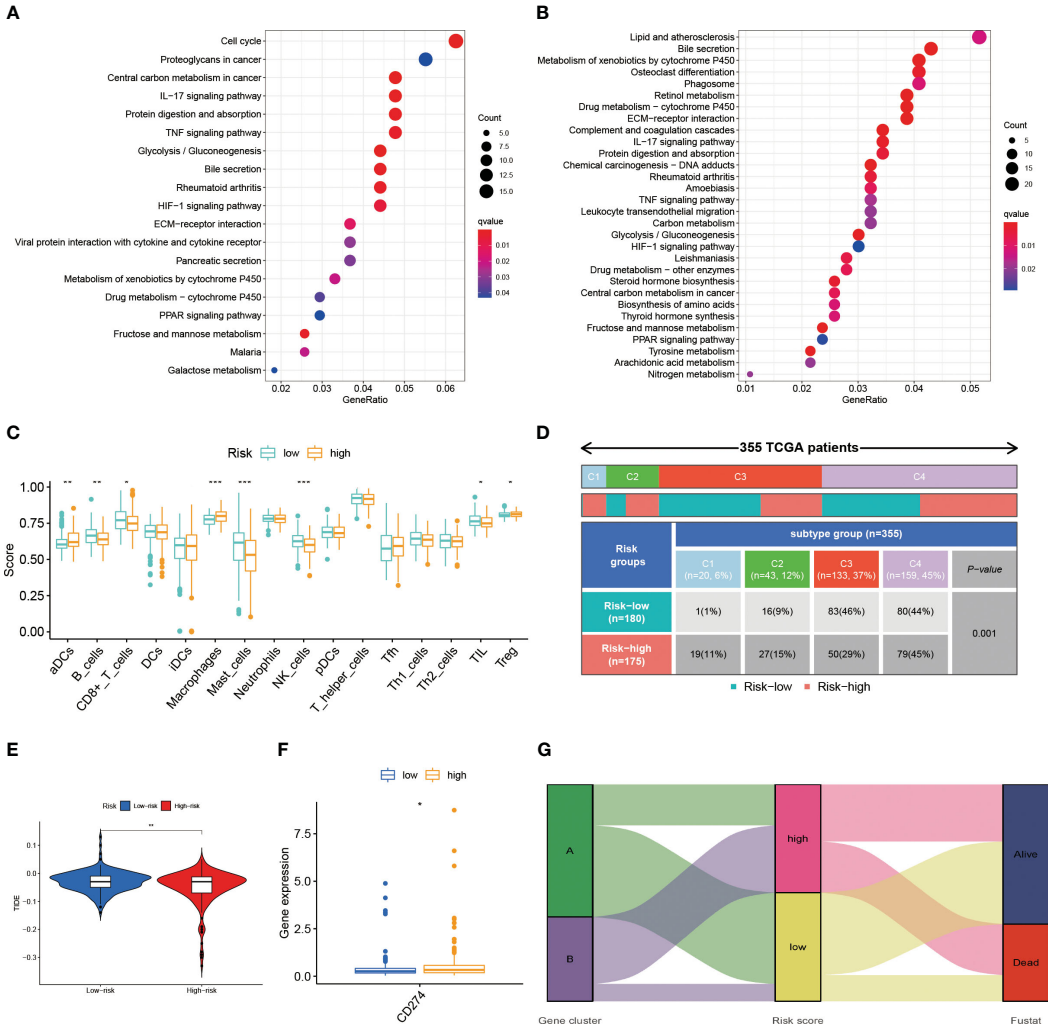


FIGURE 7
(A) The results of the KEGG in TCGA-LIHC cohort. **(B)** The results of the KEGG in ICGC-LIHC cohort. **(C)** Analysis of immune cell scores in the TCGA-LIHC cohort calculated using the ssGSEA algorithms. **(D)** Comparison of immune subtypes in different risk groups **(E)** Correlation of LR.score and TIDE score **(F)** The expression level of CD274 **(G)** Alluvial diagram of the distribution of different gene cluster, risk score and survival outcome subtypes. *p< 0.05; **p< 0.01; ***p< 0.001.

TIDE scores in the high-risk group imply that patients in the high-risk group will have a better immunotherapy outcome, which is consistent with the results of the PD-L1 expression level. Patients in subtype A were dominantly in the low-risk group and the state of survival, while the opposite was observed in subtype B patients (Figure 7G).

Functional experiments *in vitro*

In the LR model, the coefficient of LGALS9_SLC1A5 was the largest. Considering that SLC1A5 is a receptor, we took a series of experiments to test whether SLC1A5 affects the function of tumor cells. To clarify the expression of SLC1A5, we selected three hepatocellular carcinoma cells and normal hepatic epithelial cells to perform qPCR experiments. We found that the expression level of SLC1A5 was higher in all hepatocellular carcinoma cells than in normal hepatic epithelial cells, with the higher expression level in Huh7 and SNU398 cells (Supplementary Figure 1C). We performed the cell proliferation assay (CCK8 and cell colony formation), cell migration and invasion assay (transwell) using siRNA to knock down the expression level of SLC1A5 in Huh7 and SNU398 cells, and observed that the knockdown levels of si1 and si3 were the best among the three siRNAs (Figures 8A, B). The results indicated that the proliferation (Figures 8C–H), migration (Figures 9A–D) and invasion (Figures 9E–H) of hepatocellular carcinoma cells were significantly reduced after knockdown of SLC1A5.

After culturing macrophages using conditioned medium from tumor cells, we examined the expression of marker molecules in macrophages. Subsequently, we found that after culturing macrophages using conditioned medium from hepatocellular carcinoma cells with knockdown of SLC1A5, the expression of M2-type macrophage marker molecules was decreased, and the expression levels of CD206 and ARG1 were markedly down-regulated in macrophages (Figure 9I). This experimental phenomenon suggests that the expression of SLC1A5 by hepatocellular carcinoma cells may have an effect on the immune microenvironment.

10.3389/fimmu.2023.1187108

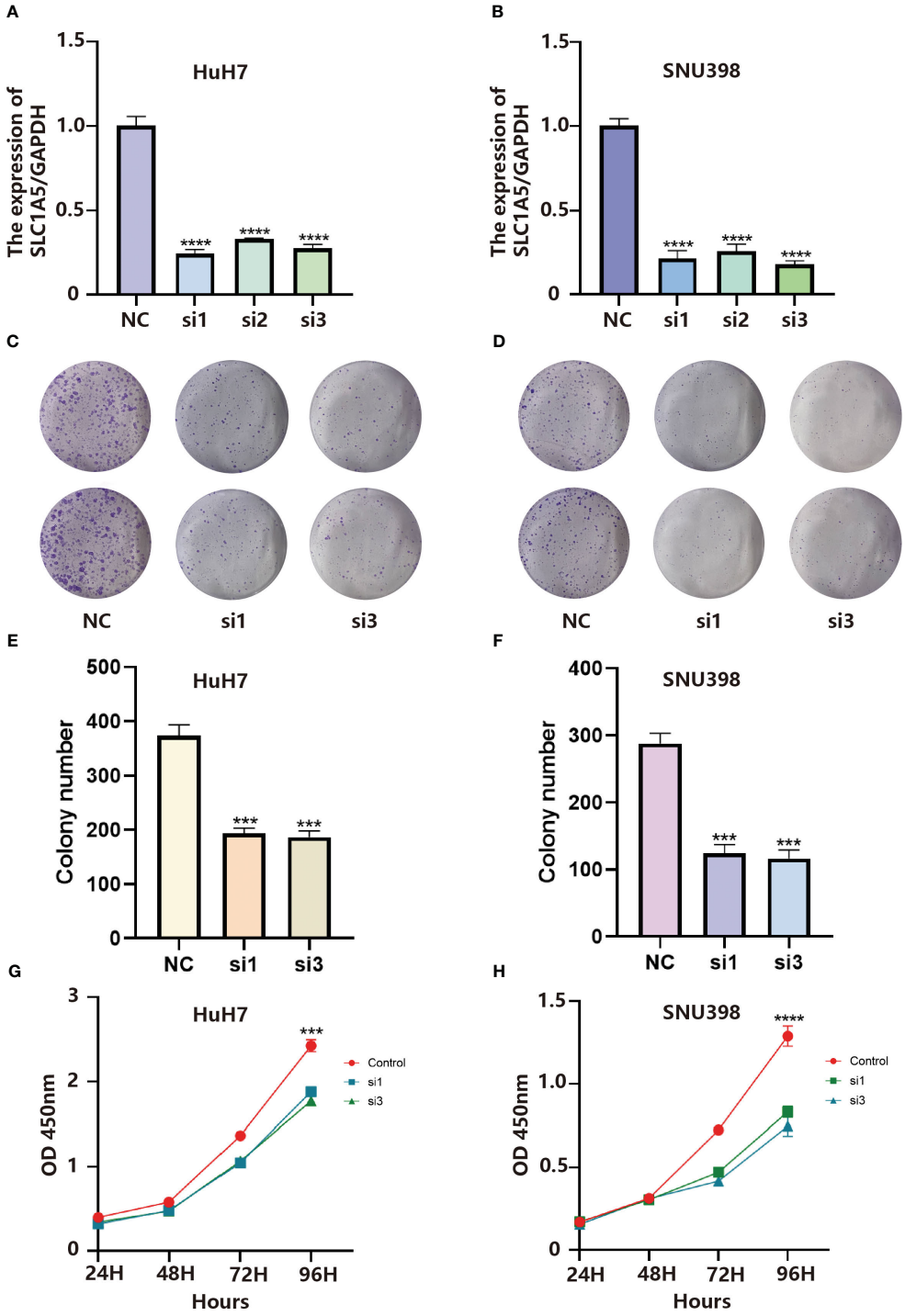


FIGURE 8
(A) Validation of knockdown efficiency by qPCR in HuH7 cells. Results represent mean ± SD. n = 3. ****p < 0.0001; two-tailed t-test. **(B)** Validation of knockdown efficiency by qPCR in SNU398 cells. Results represent mean ± SD. n = 3. ****p < 0.0001; two-tailed t-test. **(C)** SLC1A5 knockdown inhibited colony formation of HuH7 cells. **(D)** SLC1A5 knockdown inhibited colony formation of SNU398 cells. **(E)** The colony number of HuH7 cells. Results represent mean ± SD; n = 3; ***p < 0.001; two-tailed t-test. **(F)** The colony number of SNU398 cells. Results represent mean ± SD; n = 3; ***p < 0.001; two-tailed t-test. **(G)** SLC1A5 siRNA displayed reduced proliferation of HuH7 cells. Results represent mean ± SD; n = 3; ***p < 0.001; two-tailed t-test. **(H)** SLC1A5 siRNA displayed reduced proliferation of SNU398 cells. Results represent mean ± SD; n = 3; ****p < 0.0001; two-tailed t-test.

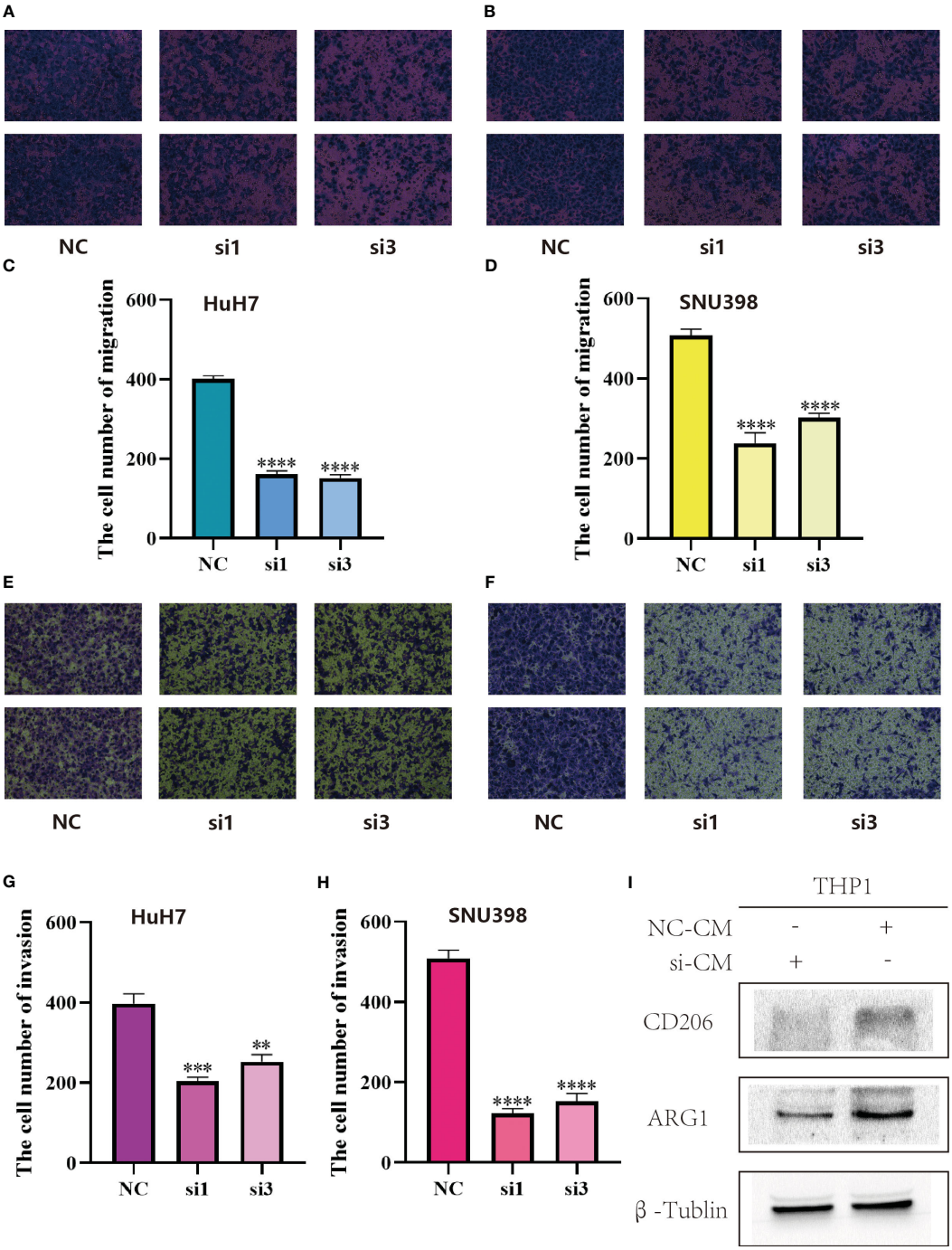


FIGURE 9
(A) SLC1A5 knockdown inhibited migration of HuH7 cells. **(B)** SLC1A5 knockdown inhibited migration of SNU398 cells. **(C)** The cell number of migration of HuH7 cells. Results represent mean ± SD; n = 3; ****p < 0.0001; two-tailed t-test. **(D)** The cell number of migration of SNU398 cells. Results represent mean ± SD; n = 3; ****p < 0.0001; two-tailed t-test. **(E)** SLC1A5 knockdown inhibited invasion of HuH7 cells. **(F)** SLC1A5 knockdown inhibited invasion of SNU398 cells. **(G)** The cell number of invasion of HuH7 cells. Results represent mean ± SD; n = 3; ***p < 0.001; **p < 0.01; two-tailed t-test. **(H)** The cell number of invasion of SNU398 cells. Results represent mean ± SD; n = 3; ****p < 0.0001; two-tailed t-test. **(I)** Relative expression of gene markers of M2 (CD206, ARG1) macrophages by stimulation of different CM in THP1 cell lines.

Discussion

The tumor microenvironment is a burning research topic in the field of oncology. Immune cells are drawn to tumor-associated inflammatory changes, and these immune cells combine with stromal cells in the tumor tissue to form the tumor microenvironment (18). The tumor microenvironment is a double-edged sword in the process of tumor development. Tumor cells will be killed by immune cells to suppress tumor progression. Simultaneously, tumor cells will evade the surveillance of immune cells and gradually remodel the tumor microenvironment so that it promotes tumor metastasis and drug resistance (19).

According to the findings of single cell sequencing, fibroblasts, myeloid cells, B cells, T cells, and NK cells have been found infiltrating liver cancer tissue. The infiltration of T cell kills tumor cells, but recent research has revealed that T cell killing depends on non-depleted T cells and that tumor cells evade the immune system by causing T cell depletion (20, 21). Myeloid cells are predominantly macrophages and tumor cells can induce the conversion of macrophages into tumor-associated macrophages, which are able to interact with other cells either through direct contact or by secreting various effector molecules (22, 23). Tumor-associated fibroblasts contribute to tumor extracellular matrix remodeling, stemness characteristics, angiogenesis and drug resistance (24, 25).

Due to the sophisticated exploration of tumors, the treatment of tumors has evolved from targeting the tumor cells themselves in the early days, directly killing them through radiotherapy or chemotherapy (26), to currently targeting the tumor microenvironment and reversing the cancer-promoting microenvironment to eliminate the tumor cells. Immunotherapy targeting PD-1 on T cells interacting with PD-L1 on tumor cells has been widely used in clinical practice for a considerable period of time (27). Not all patients benefit from them, despite the fact that their efficacy is promising in some tumors (28). Immunotherapies targeting other destinations have also been explored in advance, and we were trying to identify other cellular interactions in the tumor microenvironment that could improve the prognosis of tumor patients. We therefore employed current single cell sequencing to analyze and identify cell types in hepatocellular carcinoma, utilizing software to predict cell-to-cell communication and sought out LRs that were prognostically relevant in hepatocellular carcinoma. Constructing molecular subtypes with different expression patterns of LRs, we observed significant differences in prognosis as well as clinicopathological features of patients in different subtypes, suggesting that we targeted these LRs as possible targets for future clinical therapy. At the same time, we observed that the sensitivity of patients to chemotherapeutic drugs differed between subtypes, which could be the basis for more targeted drug delivery to different patients in the clinical field.

The function of the 18 pairs of LRs for which we constructed molecular typing has been studied in a number of ways. For example, the IL15_IL15RA interaction produces a two-sided effect, with IL-15 promoting the proliferation and maintaining the survival of certain T cells as well as consistently promoting anti-tumor responses and being crucial for controlling tumor growth and metastasis *in vivo*. However, chronic inflammatory stimulation of IL-15 increases tumor growth and metastasis (29). The connection between SELP and CD34 shows that patients gain from the stimulation of the innate immune response to improve anti-tumor immunity, eliminate tumor cells, and hinder the growth of tumors (30). The interaction of the remaining 16 pairs of LRs, in contrast, is what is causing the progression of the tumor, and this interaction allows the tumor microenvironment to change in a way that is pro-cancer, causing M2 type macrophages to transform and inhibiting the activity of NK cells. The main mechanisms of interaction between tumor cells and tumour-associated macrophages are LGALS9-SLC1A5, CCR1_CCL23, CSF1R _CSF1, GRN_SORT1, CSF1_SIRPA, and SIRPA_CD47, which primarily induce macrophage differentiation to M2 type while inhibiting macrophage phagocytosis of tumor cells, cause immune escape, and promote tumor cell proliferation, invasion, and metastasis (31–35). Specifically, LAIR1_LILRB4, EPHB6_EFNB1, and LGALS9_CD44 have immune-suppressive effects. The interactions of LGALS9_CD44 promote the differentiation and maintenance of Treg cells (36). The protein encoded by the EFNB1 gene in EPHB6_EFNB1 is a ligand for type I membrane proteins and Eph-associated receptor tyrosine kinase, and its binding to the ligand primarily exerts suppressive immune effects (37). Tyrosine phosphatases SHP-1 and SHP-2 and/or inositol phosphatase SHIP, which are detrimental to immune activation and promote tumor growth, are recruited by LAIR1_LILRB4 activation (38). WNT7B_FZD1, ICAM1_SPN and EPHA2_EFNA5 are mainly involved in facilitating tumorigenesis and invasive metastasis. Among them, ICAM1_SPN plays an important role in cell-cell interactions, and circulating tumor cells with stem cell properties may be able to use the adhesion protein ICAM1 to promote the formation of circulating tumor clusters that migrate from the body's primary tumor sites to other organs (39, 40); EPHA2 in EPHA2_EFNA5 induces inhibition of the focal adhesion kinase (FAK), extracellular regulatory protein kinases (ERK) and Akt phosphorylation, thereby regulating motility, viability and proliferation of a variety of malignant cell lines (41).

SLC1A5 is a mitochondrial glutamine transporter and glutamine regulates energy metabolism, signal transduction and redox status in cells (42). Previous studies have indicated that SLC1A5 might affect how immune cells behave and infiltrate tumor microenvironment, which can promote cancer. In breast cancer, SLC1A5 can promote tumor progression (43). Additionally, SLC1A5 accelerates the growth of lung and colorectal malignancies by forcing tumor cells to undergo metabolic reprogramming.

In hepatocellular carcinoma, we observed that the expression level of SLC1A5 was dramatically up-regulated, while the ability of hepatocellular carcinoma cells to proliferate was significantly down-regulated after SLC1A5 expression was silenced using siRNA. Additionally, hepatocellular carcinoma cells' capacity for migration was noticeably suppressed. The results suggested that the malignancy of hepatocellular carcinoma cells increased when the ligand agonized SLC1A5, and targeting SLC1A5 could alleviate tumor progression in hepatocellular carcinoma patients in the future (44, 45).

In the prognostic model we constructed, there were disparities in immune cell infiltration between the high and low risk groups, with fewer anti-cancerous immune cells and more pro-cancerous immune cells in the high-risk group, which contributed to the worse prognosis in the high-risk group. Meanwhile, patients in the low-risk group for hepatocellular carcinoma were primarily clustered in the C3 immune subtype, and the C3 type was inflammatory, in line with the results of the previous analysis (46). Most of the patients with the molecular subtype A belonged to the low-risk category. Patients in subtype A had a higher tumor mutational burden, and in relation to the previous analysis, we hypothesized that immunotherapy was more effective in subtype A patients. Moreover, both in the validation and training sets, the high-risk group focused on activating tumor metabolism (47), cell cycle, hypoxia, and immune-related pathways compared to the low-risk

group. Activation of these pathways tended to make tumor worse, implying a poorer prognosis.

The functions of the 9 pairs of LRs used to construct the prognostic models also vary, among which LGALS9_HAVCR acting primarily as a promoter of lymphocyte activation. As a co-stimulatory molecule during lymphocyte activation, HAVCR enhances the anti-tumor effects of lymphocyte and induces changes in the tumor microenvironment, leading to efficient anti-tumor immunity (48). KLRB1 in KLRB1_CLEC2D is the gene encoding human CD161. KLRB1 gene inactivation or antibody-mediated KLRB1 blockade enhances T cell-mediated glioma cell killing *in vitro*, and the CD161_CLEC2D pathway defines a potential target for immunotherapy of glioma and other human cancers (49). Therefore, increased expression of KLRB1_CLEC2D and LGALS9_HAVCR suggests that patients may have better efficacy of immunotherapy.

Considering the limitations piece, there were three flaws in this paper. Firstly, the studies in the article were retrospective studies that underwent analysis after data collection, the actual clinical value of which had not yet been ascertained in genuine clinical practice. Additionally, the article contained just a limited amount of data, subsequently a larger sample size is supposed to be incorporated into the study. Moreover, the data in the article were partially biased.

Conclusion

In conclusion, we have analyzed the interactions between cells in hepatocellular carcinoma, thereby establishing molecular subtypes of cellular communication as well as a prognostic model. The expression pattern of LRs may be able to predict the effect of chemotherapy and immunotherapy in patients with hepatocellular carcinoma and to forecast the prognosis of patients. Our findings highlight the clinical implications of LRs and provide a basis for subsequent clinical translation.

Data availability statement

Publicly available datasets were analyzed in this study. The datasets can be found here: GEO database (HYPERLINK "https://www.ncbi.nlm.nih.gov/geo/"Home - GEO - NCBI (nih.gov)), ICGC database (HYPERLINK "https://dcc.icgc.org/"Welcome | ICGC Data Portal) and TCGA database (HYPERLINK "https://portal.gdc.cancer.gov/"GDC (cancer.gov)). Other 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

All authors contributed to data analysis, drafting or revising the article, have agreed on the journal to which the article will be submitted, gave final approval of the version to be published, and agree to be accountable for all aspects of the work. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the National Natural Science Foundation of China (81372644).

Acknowledgments

We gratefully acknowledge contributions from Ze Zhang for giving us LIHC cell lines and Duo Xu for modification of images. Thanks for the members' help in the process of 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.2023.1187108/full#supplementary-material

SUPPLEMENTARY FIGURE 1
(A) Heatmap showing marker genes for each cluster. **(B)** The HR values of LRs. **(C)** Quantitative polymerase chain reaction (qPCR) displayed upregulation of SLC1A5 in liver cancer cells compared to normal cell line (LO2). Results represent mean ± SD; n = 3. ****p < 0.0001; ***p < 0.001; two-tailed t-test.

SUPPLEMENTARY FIGURE 2
(A) CDF curve of samples from ICGC cohort. **(B)** Delta area curve of samples from ICGC cohort. **(C)** ICGC clustering heatmap of samples from ICGC cohort when consensus k = 2. **(D)** Overall survival curves of molecular subtypes based on LR pairs.

SUPPLEMENTARY FIGURE 3
(A) The results of the PCA in TCGA-LIHC cohort. **(B)** The results of the PCA in ICGC-LIHC cohort. **(C)** Distribution of risk score (high or low) in TCGA. **(D)** Distribution of status (dead or alive) in TCGA. **(E)** Distribution of risk score (high or low) in ICGC. **(F)** Distribution of status (dead or alive) in ICGC.

SUPPLEMENTARY TABLE 1
The selected LRs with R-values above 0.3 and P-values less than 0.05.

SUPPLEMENTARY TABLE 2
LRs associated with prognosis.

SUPPLEMENTARY DATA SHEET 1
Detailed procedures of cell transfection, qPCR, CCK8, transwell, colony formation, western blotting, polarization of macrophages, collection of conditioned media and LR.score.

References

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

2. Xu M, Yang L, Lin Y, Lu Y, Bi X, Jiang T, et al. Emerging nanobiotechnology for precise theranostics of hepatocellular carcinoma. *J Nanobiotechnol* (2022) 20(1):427. doi: 10.1186/s12951-022-01615-2

3. Hinshaw DC, Shevde LA. The tumor microenvironment innately modulates cancer progression. *Cancer Res* (2019) 79(18):4557–66. doi: 10.1158/0008-5472.Can-18-3962

4. Yan X, Xie Y, Yang F, Hua Y, Zeng T, Sun C, et al. Comprehensive description of the current breast cancer microenvironment advancements via single-cell analysis. *J Exp Clin Cancer Res* (2021) 40(1):142. doi: 10.1186/s13046-021-01949-z

5. Cao H, Huang T, Dai M, Kong X, Liu H, Zheng Z, et al. Tumor microenvironment and its implications for antitumor immunity in cholangiocarcinoma: future perspectives for novel therapies. *Int J Biol Sci* (2022) 18(14):5369–90. doi: 10.7150/ijbs.73949

6. Ghoshdastider U, Rohatgi N, Mojtabavi Naeini M, Baruah P, Revkov E, Guo YA, et al. Pan-cancer analysis of ligand-receptor cross-talk in the tumor microenvironment. *Cancer Res* (2021) 81(7):1802–12. doi: 10.1158/0008-5472.Can-20-2352

7. Wang DR, Wu XL, Sun YL. Therapeutic targets and biomarkers of tumor immunotherapy: response versus non-response. *Signal Transduct Target Ther* (2022) 7(1):331. doi: 10.1038/s41392-022-01136-2

8. Maynard A, McCoach CE, Rotow JK, Harris L, Haderk F, Kerr DL, et al. Therapy-induced evolution of human lung cancer revealed by single-cell RNA sequencing. *Cell* (2020) 182(5):1232–1251.e1222. doi: 10.1016/j.cell.2020.07.017

9. Liu F, Wang P, Sun W, Jiang Y, Gong Q. Identification of ligand-receptor pairs associated with tumour characteristics in clear cell renal cell carcinoma. *Front Immunol* (2022) 13:874056. doi: 10.3389/fimmu.2022.874056

10. Cao ZX, Weng X, Huang JS, Long X. Receptor-ligand pair typing and prognostic risk model for papillary thyroid carcinoma based on single-cell sequencing. *Front Immunol* (2022) 13:902550. doi: 10.3389/fimmu.2022.902550

11. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* (2010) 26(12):1572–3. doi: 10.1093/bioinformatics/btq170

12. 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. *Brief Bioinform* (2021) 22(6). doi: 10.1093/bib/bbab260

13. Chu F, Liang T, Chen CLP, Wang X, Ma X. Compact broad learning system based on fused lasso and smooth lasso. *IEEE Trans Cybern* (2023). doi: 10.1109/tcyb.2023.3267947

14. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol* (2008) 26(8):1364–70. doi: 10.1200/jco.2007.12.9791

15. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118

16. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* (2009) 462(7269):108–12. doi: 10.1038/nature08460

17. Li P, Huang T, Zou Q, Liu D, Wang Y, Tan X, et al. FGFR2 promotes expression of PD-L1 in colorectal cancer via the JAK/STAT3 signaling pathway. *J Immunol* (2019) 202(10):3065–75. doi: 10.4049/jimmunol.1801199

18. Yao M, Ventura PB, Jiang Y, Rodriguez FJ, Wang L, Perry JSA, et al. Astrocytic trans-differentiation completes a multicellular paracrine feedback loop required for medulloblastoma tumor growth. *Cell* (2020) 180(3):502–520.e519. doi: 10.1016/j.cell.2019.12.024

19. Sas Z, Cendrowicz E, Weinhäuser I, Rygiel TP. Tumor microenvironment of hepatocellular carcinoma: challenges and opportunities for new treatment options. *Int J Mol Sci* (2022) 23(7). doi: 10.3390/ijms23073778

20. Chow A, Perica K, Klebanoff CA, Wolchok JD. Clinical implications of T cell exhaustion for cancer immunotherapy. *Nat Rev Clin Oncol* (2022). doi: 10.1038/s41571-022-00689-z

21. Ramadori P, Kam S, Heikenwalder M. T cells: Friends and foes in NASH pathogenesis and hepatocarcinogenesis. *Hepatology* (2022) 75(4):1038–49. doi: 10.1002/hep.32336

22. Cheng K, Cai N, Zhu J, Yang X, Liang H, Zhang W. Tumor-associated macrophages in liver cancer: From mechanisms to therapy. *Cancer Commun (Lond)* (2022) 42(11):1112–40. doi: 10.1002/cac2.12345

23. Xu W, Cheng Y, Guo Y, Yao W, Qian H. Targeting tumor associated macrophages in hepatocellular carcinoma. *Biochem Pharmacol* (2022) 199:114990. doi: 10.1016/j.bcp.2022.114990

24. Zhang J, Gu C, Song Q, Zhu M, Xu Y, Xiao M, et al. Identifying cancer-associated fibroblasts as emerging targets for hepatocellular carcinoma. *Cell Biosci* (2020) 10 (1):127. doi: 10.1186/s13578-020-00488-y

25. Peng H, Zhu E, Zhang Y. Advances of cancer-associated fibroblasts in liver cancer. *biomark Res* (2022) 10(1):59. doi: 10.1186/s40364-022-00406-z

26. Wang H, Xu Y, Zuo F, Liu J, Yang J. Immune-based combination therapy for esophageal cancer. *Front Immunol* (2022) 13:1020290. doi: 10.3389/fimmu.2022.1020290

27. Liu X, Yang L, Tan X. PD-1/PD-L1 pathway: A double-edged sword in periodontitis. *BioMed Pharmacother* (2023) 159:114215. doi: 10.1016/j.biopha.2023.114215

28. Yuan Y, Adam A, Zhao C, Chen H. Recent advancements in the mechanisms underlying resistance to PD-1/PD-L1 blockade immunotherapy. *Cancers (Basel)* (2021) 13(4). doi: 10.3390/cancers13040663

29. Ma S, Caligiuri MA, Yu J. Harnessing IL-15 signaling to potentiate NK cell-mediated cancer immunotherapy. *Trends Immunol* (2022) 43(10):833–47. doi: 10.1016/j.it.2022.08.004

30. Liu X, Qin X, Qin H, Jia C, Yuan Y, Sun T, et al. Characterization of the heterogeneity of endothelial cells in bleomycin-induced lung fibrosis using single-cell RNA sequencing. *Angiogenesis* (2021) 24(4):809–21. doi: 10.1007/s10456-021-09795-5

31. Chao MP, Weissman IL, Majeti R. The CD47-SIRPα pathway in cancer immune evasion and potential therapeutic implications. *Curr Opin Immunol* (2012) 24(2):225–32. doi: 10.1016/j.coi.2012.01.010

32. Quaranta V, Rainer C, Nielsen SR, Raymant ML, Ahmed MS, Engle DD, et al. Macrophage-derived granulin drives resistance to immune checkpoint inhibition in metastatic pancreatic cancer. *Cancer Res* (2018) 78(15):4253–69. doi: 10.1158/0008-5472.Can-17-3876

33. Czako B, Marszalek JR, Burke JP, Mandal P, Leonard PG, Cross JB, et al. Discovery of IACS-9439, a potent, exquisitely selective, and orally bioavailable inhibitor of CSF1R. *J Med Chem* (2020) 63(17):9888–911. doi: 10.1021/acs.jmedchem.0c00936

34. Krishnan V, Tallapragada S, Schaar B, Kamat K, Chanana AM, Zhang Y, et al. Omental macrophages secrete chemokine ligands that promote ovarian cancer colonization of the omentum via CCR1. *Commun Biol* (2020) 3(1):524. doi: 10.1038/s42003-020-01246-z

35. Ma L, Heinrich S, Wang L, Keggenhoff FL, Khatib S, Forgues M, et al. Multiregional single-cell dissection of tumor and immune cells reveals stable lock-and-key features in liver cancer. *Nat Commun* (2022) 13(1):7533. doi: 10.1038/s41467-022-35291-5

36. Wu C, Thalhamer T, Franca RF, Xiao S, Wang C, Hotta C, et al. Galectin-9-CD44 interaction enhances stability and function of adaptive regulatory T cells. *Immunity* (2014) 41(2):270–82. doi: 10.1016/j.immuni.2014.06.011

37. Lu P, Shih C, Qi H. Ephrin B1-mediated repulsion and signaling control germinal center T cell territoriality and function. *Science* (2017) 356(6339). doi: 10.1126/science.aai9264

38. Li Z, Deng M, Huang F, Jin C, Sun S, Chen H, et al. LILRB4 ITIMs mediate the T cell suppression and infiltration of acute myeloid leukemia cells. *Cell Mol Immunol* (2020) 17(3):272–82. doi: 10.1038/s41423-019-0321-2

39. Ma P, Zhang P, Chen S, Shi W, Ye J, Chen S, et al. Immune cell landscape of patients with diabetic macular edema by single-cell RNA analysis. *Front Pharmacol* (2021) 12:754933. doi: 10.3389/fphar.2021.754933

40. Taftaf R, Liu X, Singh S, Jia Y, Dashzeveg NK, Hoffmann AD, et al. ICAM1 initiates CTC cluster formation and trans-endothelial migration in lung metastasis of breast cancer. *Nat Commun* (2021) 12(1):4867. doi: 10.1038/s41467-021-25189-z

41. Wilson K, Shiuan E, Brantley-Sieders DM. Oncogenic functions and therapeutic targeting of EphA2 in cancer. *Oncogene* (2021) 40(14):2483–95. doi: 10.1038/s41388-021-01714-8

42. Yoo HC, Park SJ, Nam M, Kang J, Kim K, Yeo JH, et al. A variant of SLC1A5 is a mitochondrial glutamine transporter for metabolic reprogramming in cancer cells. *Cell Metab* (2020) 31(2):267–283.e212. doi: 10.1016/j.cmet.2019.11.020

43. Ma H, Qu S, Zhai Y, Yang X. circ_0025033 promotes ovarian cancer development via regulating the hsa_miR-370-3p/SLC1A5 axis. *Cell Mol Biol Lett* (2022) 27(1):94. doi: 10.1186/s11658-022-00364-2

44. Hou L, Hou Y, Liang Y, Chen B, Zhang X, Wang Y, et al. Anti-tumor effects of P-LPK-CPT, a peptide-camptothecin conjugate, in colorectal cancer. *Commun Biol* (2022) 5(1):1248. doi: 10.1038/s42003-022-04191-1

45. Liu X, Qin H, Li Z, Lv Y, Feng S, Zhuang W, et al. Inspiratory hyperoxia suppresses lung cancer metastasis through a MYC/SLC1A5-dependent metabolic pathway. *Eur Respir J* (2022) 60(6). doi: 10.1183/13993003.00062-2022

46. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. *Immunity* (2018) 48(4):812–830.e814. doi: 10.1016/j.immuni.2018.03.023

47. Zhang J, Zou S, Fang L. Metabolic reprogramming in colorectal cancer: regulatory networks and therapy. *Cell Biosci* (2023) 13(1):25. doi: 10.1186/s13578-023-00977-w

48. Evans JP, Liu SL. Multifaceted roles of TIM-family proteins in virus-host interactions. *Trends Microbiol* (2020) 28(3):224–35. doi: 10.1016/j.tim.2019.10.004

49. Mathewson ND, Ashenberg O, Tirosh I, Gritsch S, Perez EM, Marx S, et al. Inhibitory CD161 receptor identified in glioma-infiltrating T cells by single-cell analysis. *Cell* (2021) 184(5):1281–1298.e1226. doi: 10.1016/j.cell.2021.01.022

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