frontiers | Research Topics

Computational methods for multi-omics data analysis in cancer precision medicine

Edited by
Ehsan Nazemalhosseini-Mojarad and Claudia Cava

Published in
Frontiers in Genetics



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-3038-2
DOI 10.3389/978-2-8325-3038-2

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

Computational methods for multi-omics data analysis in cancer precision medicine

Topic editors

Ehsan Nazemalhosseini-Mojarad — Shahid Beheshti University of Medical Sciences, Iran

Claudia Cava — Institute of Bioimaging and Molecular Physiology, Department of Biomedical Sciences, National Research Council (CNR), Italy

Citation

Nazemalhosseini-Mojarad, E., Cava, C., eds. (2023). *Computational methods for multi-omics data analysis in cancer precision medicine*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-3038-2

Table of contents

06 Editorial: Computational methods for multi-omics data analysis in cancer precision medicine
Moein Piroozkhah, Radman Mazloomnejad, Zahra Salehi and Ehsan Nazemalhosseini-Mojarad

11 m^7G-Related DNA Damage Repair Genes are Potential Biomarkers for Predicting Prognosis and Immunotherapy Effectiveness in Colon Cancer Patients
Shuran Chen, Rui Dong, Yan Li, Ni Zheng, Guisen Peng, Fei Lu, Quanwei Qiu, Hexin Wen, Yitong Wang, Huazhang Wu and Mulin Liu

24 A Novel M7G-Related MicroRNAs Risk Signature Predicts the Prognosis and Tumor Microenvironment of Kidney Renal Clear Cell Carcinoma
Peng Hong, Huifang Du, Ming Tong, Qingfei Cao, Ding Hu, Jiaji Ma, Yanyang Jin, Zizhi Li, Weichao Huang and Guangquan Tong

47 Multi-Omics Analysis of Molecular Characteristics and Carcinogenic Effect of NFE2L3 in Pan-Cancer
Mengxiao Liu, Hui Wei, Jing Yang, Xia Chen, Haoying Wang, Ya Zheng, Yuping Wang and Yongning Zhou

68 Apoptosis-Related Gene-Mediated Cell Death Pattern Induces Immunosuppression and Immunotherapy Resistance in Gastric Cancer
Xiaolu Yuan, Jun Zhou, Liping Zhou, Zudong Huang, Weiwei Wang, Jiasheng Qiu, Qiangbang Yang, Chaohao Zhang and MingHui Ma

85 A Novel Purine and Uric Metabolism Signature Predicting the Prognosis of Hepatocellular Carcinoma
Shengjie Yang, Baoying Zhang, Weijuan Tan, Lu Qi, Xiao Ma and Xinghe Wang

95 Prognostic Signature and Tumor Immune Landscape of N7-Methylguanosine-Related lncRNAs in Hepatocellular Carcinoma
Wei Wei, Chao Liu, Meng Wang, Wei Jiang, Caihong Wang and Shuqun Zhang

108 A Pan-Cancer Analysis Reveals the Prognostic and Immunotherapeutic Value of Stanniocalcin-2 (STC2)
Zhong-Hui Jiang, Xianfeng Shen, Yanhong Wei, Yongji Chen, Hongbo Chai, Lingyun Xia and Weidong Leng

124 Copy Number Variation of Circulating Tumor DNA (ctDNA) Detected Using NIPT in Neoadjuvant Chemotherapy-Treated Ovarian Cancer Patients
Mina Sharbatoghli, Fahimeh Fattahi, Hamidreza Aboulkheyr Es, Arvand Akbari, Setareh Akhavan, Marzieh Ebrahimi, Mohsen Asadi-Lari, Mehdi Totonchi and Zahra Madjd

137 Pan-cancer analysis of the prognostic and immunological role of ANLN: An onco-immunological biomarker
Kejun Liu, Lei Cui, Cunquan Li, Chaofeng Tang, Yiming Niu, Ji Hao, Yang Bu and Bendong Chen

152 A systematic pan-cancer analysis of the gasdermin (GSDM) family of genes and their correlation with prognosis, the tumor microenvironment, and drug sensitivity
Yufu Zheng, Didi Yuan, Fei Zhang and Renkuan Tang

169 Comprehensive analysis of the cancer driver genes constructs a seven-gene signature for prediction of survival and tumor immunity in hepatocellular carcinoma
Jun Zou and Wan Qin

180 Construction and validation of a novel coagulation-related 7-gene prognostic signature for gastric cancer
Bofang Wang, Dan Zou, Na Wang, Haotian Wang, Tao Zhang, Lei Gao, Chenhui Ma, Peng Zheng, Baohong Gu, Xuemei Li, Yunpeng Wang, Puyi He, Yanling Ma, Xueyan Wang and Hao Chen

198 Prognostic and immune-related value of complement C1Q (C1QA, C1QB, and C1QC) in skin cutaneous melanoma
Huanglong Yang, Dehui Che, Yuxiang Gu and Dongsheng Cao

217 Construction and validation of a necroptosis-related lncRNAs prognosis signature of hepatocellular carcinoma
YunZhen Peng, GuoJing Wu, Xin Qiu, Yue Luo, YiShu Zou, XueYan Wei and Aimin Li

242 Cross-talk of four types of RNA modification proteins with adenosine reveals the landscape of multivariate prognostic patterns in breast cancer
Xuliren Wang, Fangdie Ye, Min Xiong, Bingqiu Xiu, Weiru Chi, Qi Zhang, Jingyan Xue, Ming Chen, Liyi Zhang, Jiong Wu and Yayun Chi

259 The cuproptosis-related signature predicts prognosis and indicates immune microenvironment in breast cancer
Jia Li, Fei Wu, Chaofan Li, Shiyu Sun, Cong Feng, Huizi Wu, Xi Chen, Weiwei Wang, Yu Zhang, Mengji Liu, Xuan Liu, Yifan Cai, Yiwei Jia, Hao Qiao, Yinbin Zhang and Shuqun Zhang

274 Cuproptosis patterns and tumor microenvironment in endometrial cancer
Junfeng Chen, Guocheng Wang, Xiaomei Luo, Jing Zhang and Yongli Zhang

287 EIF2S2 is a novel independent prognostic biomarker and correlated with immune infiltrates in hepatocellular carcinoma
Jing Liu, Tongyu Liu, Chuanhao Zhang, Jiabei He, Dong Zhou, Zhe Wang and Ruoyu Wang

300 A lactate-related LncRNA model for predicting prognosis, immune landscape and therapeutic response in breast cancer
Jia Li, Yinbin Zhang, Chaofan Li, Huizi Wu, Cong Feng, Weiwei Wang, Xuan Liu, Yu Zhang, Yifan Cai, Yiwei Jia, Hao Qiao, Fei Wu and Shuqun Zhang

317 Gender-related differentially expressed genes in pancreatic cancer: possible culprits or accomplices?
Roya Ramezankhani, Afshin Abdi Ghavidel, Saadyeh Rashidi, Mahbubeh Rojhannezhad, Hamid Reza Abolkheir, Malihe Mirhosseini, Sara Taleahmad and Massoud Vosough

330 Prognostic value and immunological characteristics of a novel cuproptosis-related long noncoding RNAs risk signature in kidney renal clear cell carcinoma
Peng Hong, Weichao Huang, Huifang Du, Ding Hu, Qingfei Cao, Yinjie Wang, Huashan Zhang, Siqiao Tong, Zizhi Li and Ming Tong

351 Characterization of glycosylation regulator-mediated glycosylation modification patterns and tumor microenvironment infiltration in hepatocellular carcinoma
Linlin Zhao, Yang Guo, Zhanfeng Liu, Jing Ma, Yanfeng Peng and Dejun Zhang

364 Integrating cell interaction with transcription factors to obtain a robust gene panel for prognostic prediction and therapies in cholangiocarcinoma
Tingjie Wang, Chuanrui Xu, Dan Xu, Xiaofei Yang, Yaxin Liu, Xiujuan Li, Zihang Li, Ningxin Dang, Yi Lv, Zhijing Zhang, Lei Li and Kai Ye

380 Constructing a signature based on the SIRT family to help the prognosis and treatment sensitivity in glioma patients
Feiyue Xuan, Zhiwei Zhang, Kuili Liu, Haidong Gong, Shaodong Liang, Youzhi Zhao and Hongzhe Li

400 Integrative analysis of WDR12 as a potential prognostic and immunological biomarker in multiple human tumors
Refaat A. Eid, Muhammad Alaa Eldeen, Mohamed A. Soltan, Mubarak Al-Shraim, Majed Aldehri, Leena S. Alqahtani, Ghadi Alsharif, Sarah Albogami, Ibrahim Jafri, Eman Fayad, Moon Nyeo Park, Shabana Bibi, Mohammed Y. Behairy, Bonglee Kim and Mohamed Samir A. Zaki

418 Construction and comprehensive analysis of a curoptosis-related lncRNA signature for predicting prognosis and immune response in cervical cancer
Li Liu, Jianfeng Zheng, Hongmei Xia, Qiaoling Wu, Xintong Cai, Liyan Ji and Yang Sun

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 05 July 2023
DOI 10.3389/fgene.2023.1226975

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Michael Baudis,
University of Zurich, Switzerland

*CORRESPONDENCE
Zahra Salehi,
Zahra.salehi6463@yahoo.com
Ehsan Nazemalhosseini-Mojarad,
ehsanmojarad@gmail.com

†These authors have contributed equally to this work

RECEIVED 22 May 2023
ACCEPTED 30 June 2023
PUBLISHED 05 July 2023

CITATION
Piroozkhah M, Mazloomnejad R, Salehi Z and Nazemalhosseini-Mojarad E (2023), Editorial: Computational methods for multi-omics data analysis in cancer precision medicine.
Front. Genet. 14:1226975.
doi: 10.3389/fgene.2023.1226975

COPYRIGHT
© 2023 Piroozkhah, Mazloomnejad, Salehi and Nazemalhosseini-Mojarad. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Computational methods for multi-omics data analysis in cancer precision medicine

Moein Piroozkhah[1†], Radman Mazloomnejad[1†], Zahra Salehi[1,2*] and Ehsan Nazemalhosseini-Mojarad[3*]

[1]Basic and Molecular Epidemiology of Gastrointestinal Disorders Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [2]Department of Immunology, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran, [3]Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran

KEYWORDS

computational methods, multi-omics data analysis, precision medicine, cancer, machine learning

Editorial on the Research Topic
Computational methods for multi-omics data analysis in cancer precision medicine

Introduction

Multi-omics constitutes a broad realm of biomedical research that covers the different levels of organisms, from genomics to higher levels, such as proteomics or metabolomics, and the interactions among these levels (Boroń et al., 2022). A commonality among all these levels is epistemological complexity. This indicates that investigations in this field necessitate a reduction in complexity to provide a better elaboration of a biological system. Thanks to ever-increasing computational methods to handle multi-omics data. Computational methods have facilitated many new avenues of research, from basic biological topics to possible applications in drug development or molecular systems engineering in heterogeneous diseases such as cancers, and opened the way for approaching the ideal of personalized medicine in the clinical oncology (Ayton et al., 2022). The articles of this Research Topic provide an up-to-date summary of novel computational methods of multi-omics data analysis to address the state-of-the-art and future perspectives in the field of precision oncology (Supplementary Table S1). The authors utilized a variety of data sources and their combinations, demonstrating the importance of multimodal data analysis in the research of diverse cancers.

"Pan-cancer" analysis

Lately, there has been a lot of attention on pan-cancer analyses, and several online databases and analytical tools have gained popularity. Some of these tools include GEPIA2, UALCAN, cBioPortal, TIMER2.0, TISIDB, STRING, DAVID, HPA, and GeneMANIA (Xie

et al., 2022). These resources have played a crucial role in advancing our understanding of pan-cancer research by enabling researchers to access and analyze comprehensive datasets, perform sophisticated analyses, and uncover valuable insights. In the context of prognostic implications for cancer patients, Eid et al. employed online web tools to investigate the protein WDR12, which demonstrates heterogeneous expression patterns across diverse cancer types. Significant associations between WDR12 expression profiles and various cancer-related factors highlight its potential as a robust prognostic biomarker and prospective target for innovative cancer treatments. In a similar vein, Liu et al. conducted a pan-cancer investigation of NFE2L3, a transcription factor known for its regulatory role in the expression of genes associated with cellular homeostasis and response to oxidative stress, utilizing diverse online databases. The study revealed consistent overexpression of NFE2L3 across a wide range of human tumors. The implications of NFE2L3 in critical cellular processes and its correlations with various molecular characteristics underscore its potential as a promising molecular biomarker for the diagnosis and prognosis of pan-cancer, as well as a prospective target for tumor treatment. Moreover, ANLN, a conserved cytoskeletal protein, plays a critical role in regulating cytokinesis and contractile ring formation during cell division. Through the utilization of online databases, Liu et al. present evidence indicating the increased expression of ANLN across diverse tumor types, its association with tumor cell proliferation, migration, infiltration, and prognosis, as well as its impact on tumor immune evasion.

To die or not to die that is not a question for eukaryotic cells. The issues that really matter are WHEN and HOW to die. Cell death has been observed for quite a long time; nevertheless, nowadays, different kinds of cell death are identified, such as apoptosis, pyroptosis, necroptosis, and cuproptosis. Cell death is associated with the development of many diseases, with tumors as a notable example (Tong et al., 2022). Zheng et al. employed analytical packages to study the Gasdermin-D (GSDM) family, proteins implicated in pyroptosis, and discovered significant upregulation of GSDM genes across various cancer types. Utilizing the Wilcoxon rank-sum test, they analyzed expression differences between malignant and corresponding normal samples for each cancer type. Moreover, the authors conducted an in-depth assessment of the pan-cancer tumor microenvironment (TME) using the CIBERSORT algorithm. The findings revealed strong correlations between GSDM gene expression levels and prognosis, clinical characteristics, TME features, and stemness scores, especially in urinary system cancers, indicating their possible role in carcinogenesis. This study conducted a drug sensitivity analysis, which revealed potential therapeutic implications that could be derived from the findings. In a distinct pan-cancer analysis, Jiang et al. investigated STC2, a secreted glycoprotein related to calcium, glucose homeostasis, and phosphorus metastasis. Their study using the "Limma" package disclosed elevated STC2 expression and association. Furthermore, survival analysis employing "survminer" and "survival" packages demonstrated adverse outcomes and significant correlations with tumor immune microenvironment, immune cell infiltration, immune checkpoint genes, mismatch repair genes, tumor mutation burden, microsatellite instability, and drug sensitivity.

"Gastrointestinal cancer" analysis

Researchers have been trying to push back the frontiers of cancer pathophysiology knowledge in recent years by investigating the roles of N7-methylguanosine (m^7G), a novel post-transcriptional modifier of RNAs (Luo et al., 2022). In this Research Topic, Wei et al. focused on hepatocellular carcinoma (HCC) and utilized LASSO regression to develop and strengthen an immune prognostic signature on the basis of m^7G-related lncRNA data. Eventually, the nine m^7G-related-lncRNA risk model was verified *in silico* through training and testing sets (n = 219 and 146 LIHC patients, respectively), and its prognostic accuracy was measured using ROC curves, univariate and multivariate Cox regression analysis. In a comparable manner, Chen et al. established an extensive risk model for colorectal cancer (CRC) through the application of multivariate Cox and LASSO Cox regression analysis. The analysis successfully identified an eleven m^7G-related DNA damage repair signature that predicts patient prognosis. Moreover, the "CIBERSORT" and "MCPcounter" packages were also used to analyze the different immune microenvironment statuses between high- and low-risk groups. The constructed risk model was useful not only for assessing patients' prognosis but even providing further insights into immunotherapy strategies in CRC. By using similar methodologies, Hong et al. constructed and assessed a prognostic signature consisting of seven m^7G-associated miRNAs that could accurately predict patient outcomes and direct personalized therapy.

In gastric cancer (GC), Yuan et al. established the apoptosisScore to quantify the apoptosis index of each patient by using "principal component analysis." Besides the prognostic potential, this score is also correlated with immune infiltrates and sensitivity to immunotherapy. This study flow encourages the mining of prognostic usage of cell death-related gene signatures in other types of cancers. Regarding HCC, Peng et al. implemented the Wilcoxon test to detect necroptosis-related lncRNAs with differential expression from the TCGA database. Based on the analysis result, they screened prognostic lncRNAs and established ten necroptosis-related lncRNA signatures. The risk score was then calculated utilizing Lasso–Cox stepwise regression analysis. This signature is involved in immune cell infiltration and the expression of immune checkpoints and can be utilized to predict immunotherapeutic efficacy. These results were also consolidated by *in vivo* data. Using similar approaches, the prognostic potential of cuproptosis was extended to KIRC by Hong et al., who proposed a prognosis signature made up of four cuproptosis-related LncRNAs.

The identification of potential targets or signatures, such as the proposed lncRNA signatures, may offer valuable insights. However, it is important to note that these findings are not sufficient by themselves to establish clinically applicable tools. Several challenges exist when it comes to understanding the role of lncRNA in cancers, including the need to define which lncRNAs and modules within them interact with effector proteins and convey target specificity (Mattick et al., 2023). LncRNAs can play complex roles associated with diverse hallmarks of cancer. As such, a more comprehensive understanding of cell and developmental biology, as well as gene-environment interactions, can be gleaned by studying the roles of lncRNAs and how they function in dynamic assemblies with other macromolecules (Winkle et al., 2021). Additionally, larger and more

rigorous clinical studies are necessary to validate the diagnostic and prognostic value of lncRNAs and to assess their potential as therapeutic targets.

As far as cholangiocarcinoma is concerned, many efforts should be made to improve the prognosis of this poor-outcome cancer. One of the promising attempts was performed by Wang et al. They constructed empirical Bayes and Markov random field models eLBP to determine cell communication-related genes using single-cell RNAseq data, with value for prognostic prediction and chemotherapeutic decision-making in cholangiocarcinoma patients.

One important consideration in cancer research may be the influence of sex on the incidence and survival rates of cancers. Pursuing the answer to this question: "Sex, possible culprits or accomplices for pancreatic cancer (PC)?", Ramezankhani et al. leveraged multiple microarray expression data for PC to determine the down and upregulated genes between the sex that were also specified for androgen and estrogen receptors. By the use of several online databases and analytic tools, including DAVID, Enrichr, bioDBnet, pancreatic expression database, TRANSFAC, HIPPIE, and BisoGenet, they identified the importance of the androgenic effectors in tumorigenesis, such as the potential role of testosterone in the extracellular matrix–cell interaction. To determine the relationship between abnormal purine/uric acid metabolism with the prognosis of HCC patients, Yang et al. used gene set variation analysis and stratified 371 HCC patients (The transcriptome sequencing data and corresponding clinical information of HCC patients were downloaded from TCGA database) into four groups: high purine biosynthesis and high purine metabolism (PBhiPMhi), high purine biosynthesis and low purine metabolism, low purine biosynthesis, and high purine metabolism, and low purine biosynthesis and low purine metabolism (PBloPMlo). The survival study indicated that the PBhiPMhi group had the poorest prognosis, whereas the PBloPMlo group had the best. In Zou et al.'s and Zhao et al.'s research articles, respectively, the readers can find explicit information on a "seven-cancer driver gene signature" was constructed for the projection of survival and tumor immunity in HCC; and how by assessing the glycosylation patterns will be helpful in identifying the characteristics of immune cell infiltration and selecting accurate treatment methods in HCC patients.

Perhaps one of the best ways to construct a prognostic risk score model is employing the weighted gene co-expression network analysis (WGCNA) to create co-expression networks and identify co-expression modules associated with clinical features. In the next step, these modules are utilized for further analysis. An interesting example of this approach was done by Wang et al. in GC. First, 200 overlapping coagulation-related genes (CRGs) were discovered by GEO and AmiGO2 (a gene ontology database). Then, they applied WGCNA and identified two module genes containing 141 CRGs. Finally, a Coagulation-score risk model was constructed using LASSO regression, which was an independent predictor of overall survival for GC patients. Last but not least, the potency of the *EIF2S2* gene in HCC as a prognostic factor, which is closely related to immune infiltration and immune checkpoints in HCC patients, was declared in the Liu et al. study, by the use of TCGA transcriptomic data. They also employed Genomics of Drug Sensitivity in Cancer database and figured out the *EIF2S2* high expression group was more sensitive to Paclitaxel and Sunitinib.

"Breast and gynecological cancer" analysis

Breast cancer (BC) is a heterogeneous disease influenced by RNA modification-associated proteins (RMPs) (Chang et al., 2020). In recent studies, Wang et al. and Li et al. employed computational techniques, including univariate Cox regression, differential expression analysis, and LASSO regression, to explore the prognostic value and interrelationships of RMPs in BC. Wang et al. identified four prognosis-related genes (PRGs) with the highest prognostic value, and their prognostic models incorporating these PRGs revealed alterations in biological pathways, genomic mutations, immune infiltrations, RNAss scores, drug sensitivities, and prognostic implications. This comprehensive analysis shed light on the collective functions and features of diverse RMPs in BC, enabling accurate prediction of clinical outcomes using advanced forecasting models. On the other hand, Li et al. established a lactate-related long non-coding RNA prognostic signature (LRLPS) using Cox and LASSO regression methods. The LRLPS showed consistent and autonomous prognostic capability and revealed differences in immune-related pathways, immune infiltration, and responsiveness to immunotherapeutic interventions between high- and low-risk BC patient cohorts. These findings contribute to our understanding of the intricate connections between BC and RMPs, providing valuable insights for personalized prognostic patterns.

In another study, Li et al. utilized consensus clustering analysis, facilitated by the "ConsensusCluster Plus" R package, to identify distinct cuproptosis-clusters. They subsequently developed a novel prognostic model using univariate Cox regression analysis, multivariate Cox regression analysis, and the stepwise Akaike information criterion (stepAIC) to select relevant variables for the model construction. By examining 13 CRRs, they revealed expression differences between BC and normal tissues. Incorporating five prognostic CRRs, BC patients were classified into two distinct cuproptosis-clusters (C1 and C2), with C2 exhibiting superior survival outcomes and increased immune infiltration.

In addition, gynecologic cancers also present significant clinical challenges that necessitate innovative strategies for understanding their underlying mechanisms and enhancing patient outcomes. Sharbatoghli et al. applied non-invasive prenatal testing (NIPT) to investigate copy number variations (CNVs) in circulating tumor DNA (ctDNA) from ovarian cancer patients undergoing neoadjuvant chemotherapy (NAC). Six OC patients were categorized into NAC-sensitive and NAC-resistant groups, and CNV analysis was performed using two NIPT methods. WISECONDORX revealed fewer CNV changes in NAC-sensitive patients, while NextGENe identified CNVs in both coding and non-coding genes exclusively in NAC-resistant patients. Analysis of these genes demonstrated their amplification and a significant association between their high expression levels and reduced overall survival in chemotherapy-resistant patients.

Using the "ConsensClusterPlus" R package for consensus clustering, Chen et al. classified Uterine Corpus Endometrial Carcinoma samples based on CRG expression. Principal component analysis estimated molecular pattern distribution, while gene set variation analysis assessed biological process changes. Differentially expressed genes were identified, and a CRG score system was developed. Immunophenoscores evaluated

the correlation between CRG scores and immunotherapy efficacy. Lower CRG scores indicated improved prognosis, enhanced immunotherapeutic response, and increased tumor mutation burden. Moreover, in cervical cancer, a cuproptosis-related lncRNA signature has been developed, demonstrating strong potential for survival prediction, immunotherapy assessment, and prognosis post-radiotherapy. Liu et al. constructed this signature through a LASSO-Cox analysis of lncRNAs associated with cuproptosis, which has proven robust in stratifying patients based on risk scores, contributing to improved prognostication and potential biomarker identification.

"Skin cutaneous melanoma" analysis

Recent research has identified the complement protein C1Q as a crucial factor in the pathogenesis and progression of cutaneous melanoma. Yang et al. conducted a study demonstrating that the increased expression of C1QA, C1QB, and C1QC subunits of C1Q holds substantial diagnostic and prognostic value in cutaneous melanoma. Elevated expression of these subunits was associated with improved patient survival and served as independent biomarkers. Furthermore, increased expression levels correlated with immune cell infiltration, expression of other biomarkers, immune checkpoint proteins, and enrichment of immune and apoptotic pathways. These findings suggest that upregulated C1QA, C1QB, and C1QC may serve as potential biomarkers for the diagnosis and prognosis of cutaneous melanoma, particularly concerning the response to immunotherapy.

"Glioma" analysis

Glioma, a formidable type of brain cancer, poses substantial challenges in diagnosis and treatment. Xuan et al. conducted a groundbreaking study elucidating the crucial role of silent information regulator (SIRT) family enzymes in vital cellular processes, including apoptosis, metabolism, aging, and cell cycle regulation, with implications for glioma and other cancers. Through innovative analyses employing LASSO regression and multivariate Cox methods, they identified a novel SIRT-based gene signature that accurately stratifies glioma patients based on transcriptome and clinical data. Prognostic evaluations confirmed its exceptional predictive value and broad applicability. An overall survival nomogram, incorporating sex, age, risk score, pathological grade, and IDH status, was developed to aid clinical decision-making. Notable distinctions in immune status and immune cell infiltration between high and low-risk groups shed light on the relationship between the SIRT signature and the tumor microenvironment. This SIRT-based signature offers a potent tool for glioma prognosis and personalized management, potentially improving patient outcomes.

Conclusion

In conclusion, this Research Topic emphasizes the paramount importance of computational methods in revolutionizing cancer precision medicine through the analysis of multi-omics data. Although numerous studies have employed single-mode data, it is imperative to emphasize the increasing importance of the multi-omics concept in cancer research. In this context, the integration of diverse data types, including genomics, epigenomics, transcriptomics, and proteomics, is vital to gaining a more comprehensive understanding of cancer biology and enabling the identification of novel biomarkers, prognostic indicators, and therapeutic targets.

The key findings discussed herein underscore the efficacy of diverse computational techniques, including network analysis and pathway-based approaches. These methodologies have demonstrated their potential in predicting patient outcomes, facilitating treatment decisions, and optimizing personalized therapeutic strategies. Despite the progress made in multi-omics research, it is worth noting that many studies have predominantly focused on implementing omics studies alongside each other rather than genuinely integrating multi-modal data. It is essential to address this limitation and further promote the development and application of advanced computational methods that can effectively integrate and analyze multi-modal data. By doing so, the cancer research community can unlock the full potential of multi-omics approaches, leading to breakthroughs in precision medicine and improved patient care.

Prospects

The forthcoming years will prioritize delving into single-cell multi-omics analysis, incorporating real-time data for dynamic treatment monitoring, and integrating electronic health records into multi-omics frameworks. As a subfield of artificial intelligence (AI), machine-learning methods have an exceptional capability to integrate Omics data effectively. Machine-learning algorithms are trained to model complex patterns that cannot be accurately captured by traditional mathematical models in high-dimensional data. Even though there has been a notable surge in the number of multi-modal experiments conducted and an abundance of data at our disposal for analysis, the potential of machine learning tools in effectively integrating these datasets has not been fully harnessed. It is imperative that future research delves into the vast capacities of utilizing machine learning, particularly deep learning, in conducting multi-modal data analysis. Eventually, establishing standardized data formats and user-friendly software tools, as well as promoting collaboration between computational and clinical researchers, are essential steps toward the practical application of molecular cancer discoveries in clinical settings.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Acknowledgments

We thank everyone who provided support for this study.

Conflict of interest

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

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/fgene.2023.1226975/full#supplementary-material

References

Ayton, S. G., Pavlicova, M., Robles-Espinoza, C. D., Tamez Peña, J. G., and Treviño, V. (2022). Multiomics subtyping for clinically prognostic cancer subtypes and personalized therapy: A systematic review and meta-analysis. *Genet. Med.* 24, 15–25. doi:10.1016/j.gim.2021.09.006

Boroń, D., Zmarzły, N., Wierzbik-Strońska, M., Rosińczuk, J., Mieszczański, P., and Grabarek, B. O. (2022). Recent multiomics approaches in endometrial cancer. *Int. J. Mol. Sci.* 23, 1237. doi:10.3390/ijms23031237

Chang, G., Shi, L., Ye, Y., Shi, H., Zeng, L., Tiwary, S., et al. (2020). YTHDF3 induces the translation of m(6)a-enriched gene transcripts to promote breast cancer brain metastasis. *Cancer Cell.* 38, 857–871.e7. doi:10.1016/j.ccell.2020.10.004

Luo, Y., Yao, Y., Wu, P., Zi, X., Sun, N., and He, J. (2022). The potential role of N(7)-methylguanosine (m7G) in cancer. *J. Hematol. Oncol.* 15, 63. doi:10.1186/s13045-022-01285-5

Mattick, J. S., Amaral, P. P., Carninci, P., Carpenter, S., Chang, H. Y., Chen, L.-L., et al. (2023). Long non-coding RNAs: Definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell. Biol.* 24, 430–447. doi:10.1038/s41580-022-00566-8

Tong, X., Tang, R., Xiao, M., Xu, J., Wang, W., Zhang, B., et al. (2022). Targeting cell death pathways for cancer therapy: Recent developments in necroptosis, pyroptosis, ferroptosis, and cuproptosis research. *J. Hematol. Oncol.* 15, 174. doi:10.1186/s13045-022-01392-3

Winkle, M., El-Daly, S. M., Fabbri, M., and Calin, G. A. (2021). Noncoding RNA therapeutics — Challenges and potential solutions. *Nat. Rev. Drug Discov.* 20, 629–651. doi:10.1038/s41573-021-00219-z

Xie, L., Wang, L., Zhu, W., Zhao, J., and Guo, X. (2022). Editorial: Bioinformatics tools (and web server) for cancer biomarker development, volume II. *Front. Genet.* 13, 959159. doi:10.3389/fgene.2022.959159

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 09 June 2022
doi: 10.3389/fgene.2022.918159

m^{7}G-Related DNA Damage Repair Genes are Potential Biomarkers for Predicting Prognosis and Immunotherapy Effectiveness in Colon Cancer Patients

Shuran Chen[1†], Rui Dong[1†], Yan Li[2], Ni Zheng[3], Guisen Peng[3], Fei Lu[1], Quanwei Qiu[1], Hexin Wen[1], Yitong Wang[1], Huazhang Wu[3] and Mulin Liu[1,3]**

[1]Department of Gastrointestinal Surgery, First Affiliated Hospital of Bengbu Medical College, Bengbu, China, [2]Department of Gynecologic Oncology, First Affiliated Hospital of Bengbu Medical College, Bengbu, China, [3]School of Life Science, Anhui Province Key Laboratory of Translational Cancer Research, Bengbu Medical College, Bengbu, China

OPEN ACCESS

Edited by:
Ehsan Nazemalhosseini-Mojarad, Shahid Beheshti University of Medical Sciences, Iran

Reviewed by:
Weimin Zhong, Xiamen Fifth Hospital, China
Qian Chen, Guangxi Medical University Cancer Hospital, China

***Correspondence:**
Mulin Liu
liumulin66@aliyun.com
Huazhang Wu
whzhang1025@163.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: 12 April 2022
Accepted: 23 May 2022
Published: 09 June 2022

Citation:
Chen S, Dong R, Li Y, Zheng N, Peng G, Lu F, Qiu Q, Wen H, Wang Y, Wu H and Liu M (2022) m^{7}G-Related DNA Damage Repair Genes are Potential Biomarkers for Predicting Prognosis and Immunotherapy Effectiveness in Colon Cancer Patients. Front. Genet. 13:918159. doi: 10.3389/fgene.2022.918159

Objective: m^{7}G is a post-transcriptional modification modality, however, limited research has been conducted on its role in colon cancer. DNA damage repair (DDR) is an important factor that contributes to colon cancer development, growth and chemoresistance. This study aimed to explore whether m^{7}G-related DNA damage repair genes may be used as biomarkers to predict the prognosis of colon cancer patients.

Methods: We use non-negative matrix factorization (NMF) to type CRC patients into. Risk models were constructed using different expression genes in two clusters. We assessed the reliability of risk models with DCA curves, and a Nomogram. Meanwhile, The receiver operating characteristic and C-index curves were used to compare the predictive significance of the constructed risk models with other studies. In additional, we examined the significance of risk models on patients' immunity microenvironment and response to immune therapy. Finally, we used a series of cellular experiments to validate the effect of model genes on the malignant progression of CRC cells.

Results: Twenty-eight m^{7}G genes were obtained from the GSEA database. Multivariate Cox and LASSO Cox regression analysis was performed and eleven m^{7}G-related DDR genes were identified for constructing the risk model. Survival and stage of CRC patients were worser in the high-risk group than in the low-risk group for both the training and test sets. Additionally, the different immune microenvironment status of patients in the high- and low-risk groups, suggesting that patients in the low-risk group may be more sensitive to immunotherapy, particularly immune checkpoint inhibitors. Finally, we found that depletion of ATP2A1, one of the risk genes in our model, influence the biologic behaviour of CRC cells significantly.

Conclusion: The m^{7}G-related DDR genes can be used as important markers for predicting patient prognosis and immunotherapy response. Our data suggest that ATP2A1 may promote the proliferation of colon cancer cells. These findings may provide new therapeutic targets for the treatment of colon cancer.

Keywords: colon cancer, prognostic model, m 7 G, DNA damage repair, tumour immunity

INTRODUCTION

Colorectal cancer ranks third in terms of incidence and second in terms of mortality among neoplastic diseases (Sung et al., 2021). Worse still, colorectal cancer tends to be more prevalent in younger people (Lieu et al., 2019). With the advancement of molecular research on colorectal cancer, numerous serum tumour markers, and molecular screening techniques have significantly increased the detection rate of early-stage tumours (Russo et al., 2019; Miao et al., 2020). Patients with early-stage colorectal cancer have a significantly better response to treatment and prognosis as compared to those with advanced colorectal cancer (Andrew et al., 2018). Therefore, it is important to develop new diagnostic markers for the management of colorectal cancer.

Epigenetic dysregulation is strongly associated with the occurrence of a variety of diseases, particularly cancer (Nacev et al., 2020; Galassi et al., 2021). Almost all gastrointestinal tumours are caused by epigenetic dysregulation, but these alterations can be used to predict cancer risk, prognosis, and response to therapy (Grady et al., 2021). m^7G methylation is a conserved modification found in both eukaryotes and bacteria (Jühling et al., 2009). Previous studies established that m^7G occurs most frequently on tRNA and contributes to the stability of tRNA (Tomikawa, 2018). Recently, it was discovered that m^7G has important modifying effects on other types of RNA (Lin et al., 2018; Dai et al., 2021; Ying et al., 2021). The development of colon cancer is closely related to defective DNA repair, chromosome instability, microsatellite instability, and alterations in the serrated pathway and DNA methylation. All of these features point DNA damage repair to the core to overcome the colorectal cancer. However, the effect of m^7G-related DDR genes on the prognosis of colon cancer patients has not been previously reported.

In this study, we analysed the transcriptomic and clinical data from TCGA database of colon cancer patients and constructed a prognostic model for m^7G-related DNA damage repair genes. The model was then validated with GEO datasets. Finally, we downgraded the expression of one of the ATP2A1 one of the model gene to explore its role in colon cancer. This model may provide a new reference for doctors to assess the prognosis of colon cancer patients. Meanwhile, targeting of ATP2A1 may become a potential therapeutic strategy for colon cancer.

MATERIAL METHODS

Data Acquisition and Processing

The m^7G and DNA damage repair genes data was obtained by a combination of the literature (Tomikawa, 2018) and the GSEA database (http://www.gsea-msigdb.org/gsea/index.jsp). Colon cancer patients' transcriptomic, clinical, and mutation data were obtained from the TCGA database (https://portal.gdc.cancer.gov/)、GSE17536 and GSE39582. Batch effect in TCGA and GEO were removed by "limma" package. Only survival time for CRC patients is ≥ 30 days were retained, and 1,160 patients were finally enrolled in this study (**Table 1**).

NMF Consensus Clustering

David (https://david.ncifcrf.gov/) performed Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) enrichment analysis of the m^7G gene (Huang et al., 2009). The m^7G-associated DNA damage repair gene was identified using the 'psych' package with $R > 0.1$ and adjusted p-value < 0.05. The DDR Subtypes were identified using the 'NMF' package in R software. The following criteria should be followed for a good clustering effect: 1. the cumulative distribution function curve increases smoothly; 2. the sample size of each group should not be too small, and after clustering, intra-group correlation is high and inter-group correlation is low. The "limma" package was used to screen for typing differences using a filter of |log2fold change (FC)| >0.1 and adjusted p-value <0.05. The "GSVA" package of the R software explored the different functional regions of the genes that differed between typing. ssGSEA was used to explore the association between typing and immune cell infiltration.

Construction and Validation of a Prognostic Model

The 417 patients from TCGA were used as the train group for constructing the prognostic model and the 743 patients from the GEO database were used as the test group for the validation of the prognostic model. The patients were divided into two groups according to the median risk score, and the correlation heatmap and risk heatmap were drawn using the "Psych" and "pheatmap" packages. Principal component analysis (PCA) was performed to determine the validity of the risk model. Kaplan–Meier survival plots, receiver operating characteristic (ROC), C-index and DCA curves, visualized by R package"survival", "timeROC","ggDCA" and "survminer", were used to evaluate prognostic value of risk model in colon cancer. The constructed model in this study was also compared with known models in predicting the prognosis of colon patients (Huang et al., 2018; Li et al., 2021; Liang et al., 2021). Univariate and multifactor analyses were performed to determine the independent prognostic value of the risk model. Nomogram plots for the TCGA and GEO datasets were plotted separately using the 'rms' package of R software, and calibration curves were used to assess the accuracy of the column plots.

Association Between the Risk Model and the Immune Microenvironment

"CIBERSORT" and "MCPcounter" package were used to analysed the different immune microenvironment between high- and low-risk group. The samples were filtered based on $p < 0.05$, and 1,000 simulations were performed to improve data accuracy. Heat and scatter plots of risk scores versus immune cell content were created using "ggplot2" and "tidyverse" packages. The difference in immune cell content between the high- and low-risk groups was determined using single-sample gene set enrichment analysis (ssGSEA).

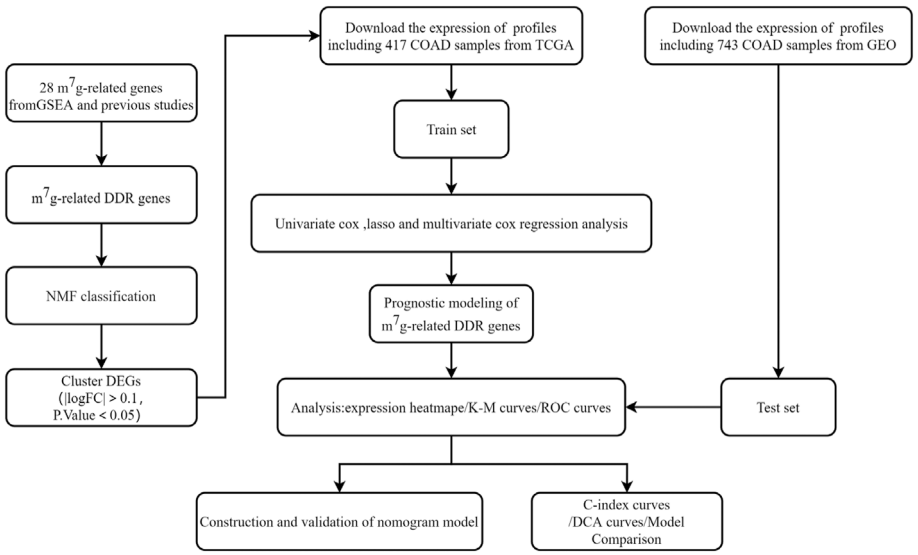


FIGURE 1 | The flowchart showing the design to establish and validate the prognostic signature in the study.

Associations Between Risk Models and Mutation, and Drug Sensitivity

To investigate the relationship between the risk model and tumour mutation, the "maftools" package was used to obtain the mutation status of patients in both high and low-risk groups. The "pRRophetic" package was used to predict the sensitivity of the risk model to clinical drug treatment. Colon cancer patient's dates were obtained from the TIDE (http://tide.dfci.harvard.edu/).The relationship between risk scores and patient prognosis and immunotherapy efficacy was analysed using the pearson correlation test.

The Expression and Biological Function of ATP2A1 in Colon Cancer

Western blot and Immunohistochemistry were used to detect the different expression of ATP2A1 in tumour and adjacent normal tissue. RKO、HCT116、SW620 and NCM460 were cultured in DMEM with 10% fatal bovine serum. Two small inferring RNAs (siRNAs) were employed to knock down ATP2A1 (siRNA#1, 5′-GGUGGUUCCUGUACGGUGATT-3′; siRNA#2, 5′-GAAUGUGUUCAACACGGAUTT-3′). CCK8 and EdU assays were used to test proliferation in control and ATP2A1 knockdown cells.

Data Analysis

All statistical analyses were performed using R version 4.1.0. Statistical results were considered significant at $p < 0.05$.

RESULTS

Functions of m^7G Genes in Colon Cancer

We first present a technical roadmap for building a risk model (**Figure 1**). Next, we investigate the expression of 28 m^7G genes. The majority of m^7G genes were significantly differentially expressed in cancer and normal tissues (**Figure 2A**). Additionally, most genes have varying degrees of mutation (**Figure 2B**). The GO and KEGG enrichment demonstrated that m^7G genes are involved in both mRNA modification and resistance to tyrosine kinase inhibitors (**Figures 2C,D**). Resistance to tyrosine kinase inhibitors is relevant to methylation, meanwhile methylation of DDR genes is an important triggers of colon cancer. Therefore, we obtained DNA damage repair genes from GSEA official website and screened m^7G-related DDR gene with correlation coefficient >0.1 (**Figure 2E**). In summary, we found that m^7G may participate in regulate DNA repair in colon cancer.

Distribution of m^7G Related DDR Genes Subtypes Using NMF Consensus Clustering

According to cophenetic coefficients, we decided k = 2 as the best cluster number (**Figure 3A**: **Supplementary Figure S1**). When k = 2, including 84 cases in C1 subtype and 334 cases in C2 subtype (**Supplementary Table S1**). GSVA enrichment analysis shows that subtype C2 is significantly enriched in DNA damage repair related pathways such as mismatch repair, DNA repair and cell cycle (**Figure 3B**, **Supplementary Table S2**). In addition most of the differential expression genes have an independent prognostic effect on colon cancer patients (**Supplementary Table S3**). DNA damage repair is closely linked to the immune microenvironment of tumours. ssGSEA showed that the majority of immune cells had increased infiltration in subtype B (**Figure 3D**). Taken together, the m^7G related DDR genes may regulate the immune microenvironment in colon cancer patients.

Construction of the Risk Model

The TCGA dataset was used as the training set to construct the risk model, and the GSE17536 and GSE39582 datasets were used

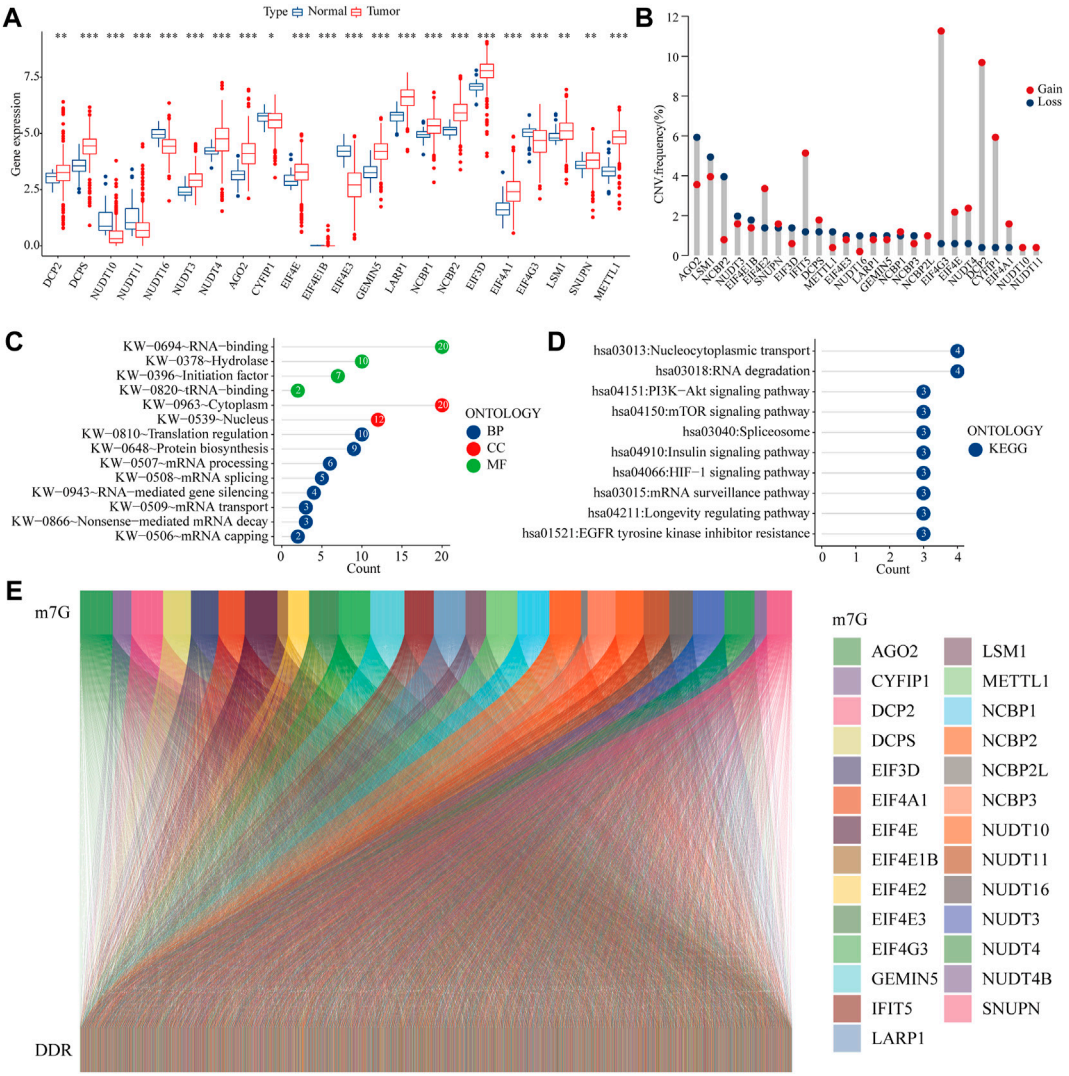


FIGURE 2 | m^{7}G related DDR gene **(A)** Expression of m^{7}G-related genes in COAD and normal tissues **(B)** Mutation frequency of m^{7}G-related genes **(C–D)** GO and KEGG analysis of m^{7}G-related genes **(E)** Co-expression relationship between the m^{7}G gene and DDR gene.

as the validation set to verify the value of the model (**Supplementary Figure S2A, B**). LASSO regression analysis and multivariate Cox analysis were performed on genes that differed between DDR subtypes to identify the best candidate genes. Multivariate Cox regression identified ten independent prognostic genes used for the construction of the risk models.

Risk Score=(0.363 * the expression of ATP2A1)+(0.215* the expression of HEYL)+(0.265* the expression of ASB6)+(0.298 * the expression of ZEB1. AS1)+(0.081 * the expression of NOL3)+(0.212* the expression of TRPM5)+(0.013 * the expression of HOXC6)+(0.0081 * the expression of DPP7)+(0.021 * the expression of GPRC5B)+(0.005 * the expression of PCDHB2).

The K-M survival curves showed that high-risk patients in train set had significantly lower survival rates than low-risk group (**Figure 4A**). The ROC curve demonstrated an AUC value of 0.754,0.758,0.755 and 0.772 for the survival probability of 1,3,5 and 10 years for colon patients. (**Figure 4B**). In addition, the number of patients who died increased significantly with increasing risk scoring (**Figures 4C,D**). Risk scoring also significantly differentiated patient prognosis in the external independent GEO dataset (**Figures 4F–J**). Finally, our study suggested that high-risk colon cancer have relatively poorer TNM and Stage than patients in low-risk group (**Figure 4K**-N).

ATP2A1 is Required for CRC Cell Survival

Correlations matrices showed that ATP2A1 was most closely related to risk model (**Figure 5A**). Dates from the TCGA database showed that CRC tissues had higher expression levels of ATP2A1 than adjacent normal tissue (**Figure 5B**).K-M curves displayed that high expression of ATP2A1 is correlated with poor OS, FP and PPS of colon cancer patients (**Figures 5C–E**). Immunohistochemistry showed that ATP2A1 has higher expression in cancer than in

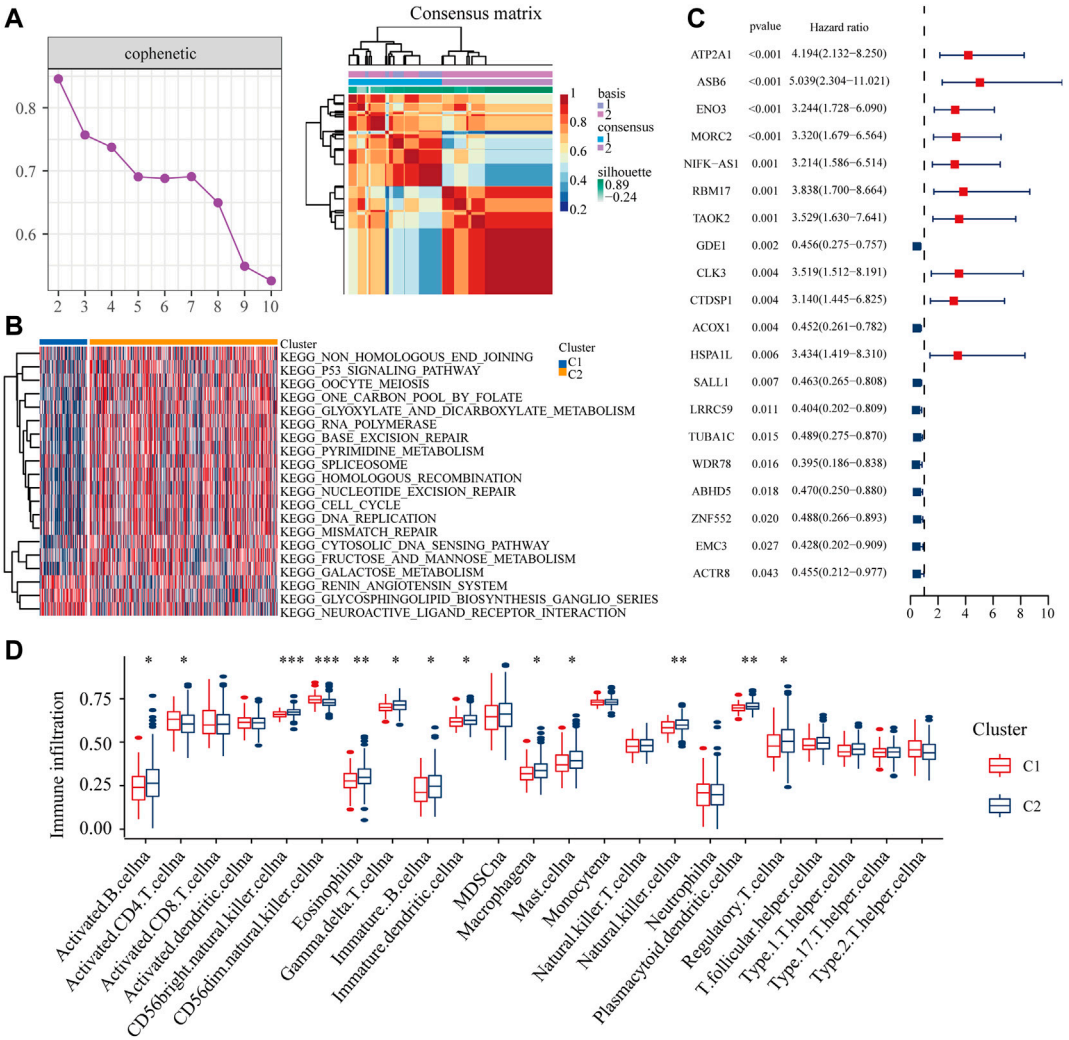


FIGURE 3 | Distribution of m^{7}G related DDR gene **(A)** NMF clustering based on m^{7}G related DDR gene decomposes the samples in TCGA cohorts. **(B)** The GSVA analysis of two clusters. **(C)** The independent prognostic analysis of the DEGs from two clusters. **(D)** The ssGSEA analysis of two clusters.

adjacent normal tissue (**Figure 5F**). Western blot showed HCT116, RKO and SW620 have higher expression of ATP2A1 than NCM460 (**Figure 5G**). Two independent siRNAs (siRNA#1 and siRNA#2) were used to knockdown the expression of ATP2A1 in HCT116 and RKO (**Figure 5H**). CCK-8 and EdU assay result showed that the knockdown of ATP2A1 inhibit cell proliferation significantly (**Figures 5I,J**). These results illustrate that ATP2A1 is overexpressed in colon cancer, and may promote the progression of colon cancer.

Prognostic Value of the Risk Model

This study then used a variety of methods to assess the predictive value of the prognostic model. High-risk patients showed a poorer prognosis than low-risk patients in both univariate Cox regression (hazard ratios [HR]: 16.008, 95% confidence interval [CI]: 7.955–32.212, $p < 0.001$) and multifactorial Cox regression (8.592 [CI]: 4.031–18.314, $p < 0.001$) (**Supplementary Figure S2C, D**). PCA plots showed that the risk model is more capable than whole genome and the m^{7}G genes to devised patients into high- and low-risk groups (**Supplementary Figure S2E–G**). The ROC and DCA curves indicate that the risk model has a higher accuracy and sensitivity in predicting patient prognosis compared to other clinical characteristics (**Figures 6A,B**). We further established prognostic nomograms including Stage, and risk score in both train set and test set to provide the survival probability of 1, 3, and 5 years for colon patients. Nomograms showed a good performance with a high AUC of 0.769, suggesting that it could be served as an effective tool for the prognostic evaluation of patients with CRC (**Figure 6A**). In addition, we constructed calibration curves, which showed that the predicted and actual survival rates were consistent with 1, 3, and 5 years (**Figures 6E,F**). We compared the model constructed in this study with existing models with ROC, C-index, K-M curves (**Figure 6D**,G-J). All results indicate that the constructed prediction model revealed a high efficacy for prognosis prediction.

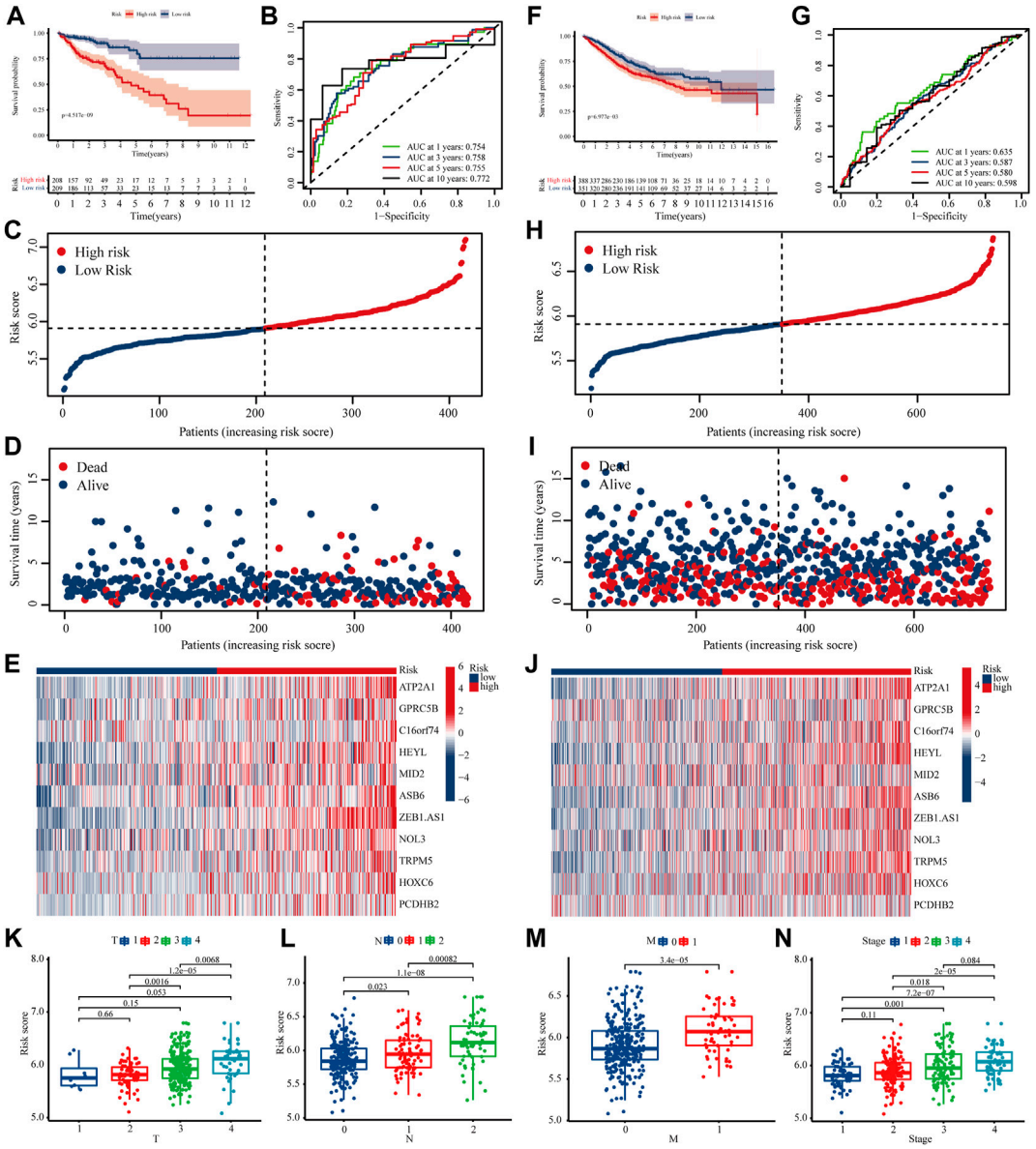


FIGURE 4 | Validation of the risk model in the training set and testing set **(A,F)** K-M survival curves for high- and low-risk groups in training set and testing set **(B,G)** ROC curves for sensitivity and specificity of risk models used to predict 1-, 3-,5- and 10- year survival in training set and testing set **(C–D, H–I)** Scatter plot of m^7G gene risk scores and patient survival, with red and blue dots representing high- and low-risk patients in training set and testing set **(E,J)** Expression of gene used to construct risk models in training set and testing set **(K–N)** Different TNM stage of colon patients in high- and low-risk scores.

Immune Features in High- and Low-Risk Groups

GSEA showed that the immune response-related signature was prominently enriched in high-risk group (**Figures 7A,B**). Patients in high risk group had higher StromalScore、ImmuneScore and ESTIMATEScore than those in low risk group (**Figure 7C**). As shown in the diagrams, the risk score was positively correlated with M0 macrophages and negatively correlated with dendritic cells activated, mast cells activated, T cells CD4 + memory activated and neutrophils (**Figures 7D,E**). Immune checkpoint inhibitors are an important form of tumour immunotherapy (Hu et al., 2021). Most immune checkpoints have different expression between the high and low risk groups (**Figure 7F**). These results suggest that different risk score may have a direct impact on the efficacy of treatment, especially for immunotherapy.

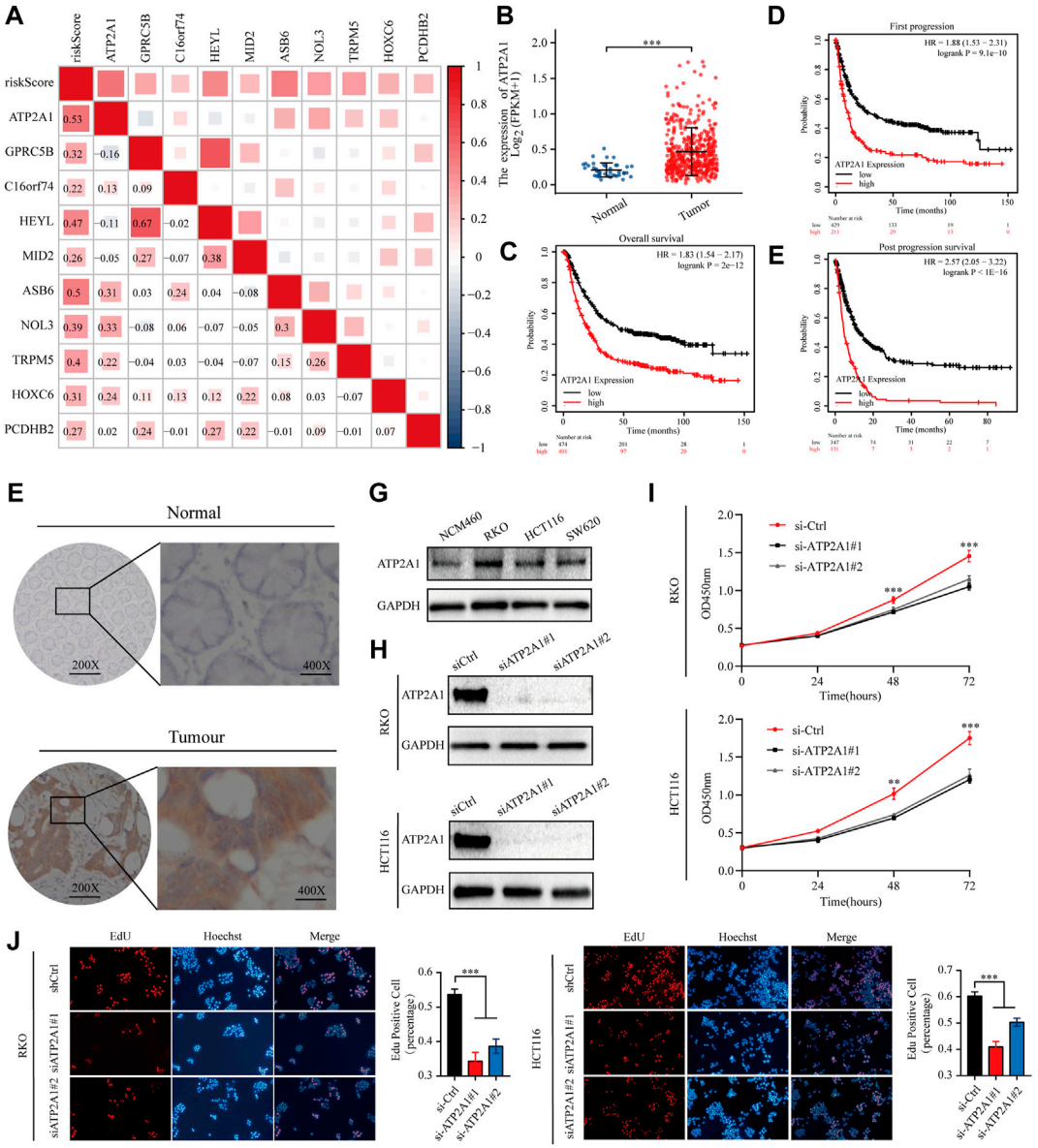


FIGURE 5 | ATP2A1 is required for CRC cell survival **(A)** The relationship between ATP2A1 and risk model was showed in correlations matrices **(B)** The expression of ATP2A1 in CRC tissues and adjacent normal tissue of the TCGA database **(C–E)** Kaplan–Meier survival curves for CRC patients based on ATP2A1 mRNA expression **(F)**Representative images of immunostaining of ATP2A1 in primary CRC samples versus normal samples **(G)** Representative images of Western blot of ATP2A1 in RKO, HCT116, SW620 versus NCM460 **(H)** Western blot to show knockdown efficiency of ATP2A1 in RKO and HCT116 cells by two independent siRNAs **(I)**Cell proliferation of RKO cells or HCT116 cells **(J)** Edu assay to test the cell proliferation of control cells comparing to ATP2A1 knockdown cells.

Mutations and Treatment Sensitivity Between High- and Low-Risk Groups

Because differences in gene mutations can dramatically affect treatment efficacies in colon cancer patients, so this study analyse the effect of high and low risk groups on mutation frequency. We listed the top 20 genes with highest mutation frequency, which showed that the mutation rate of KRAS, TP53 and PIK3CA in high-risk group was higher than low-risk group (**Figures 8A,B**). The K-M curve shows that patients with higher TMB have a worse prognosis, while high risk patients with high tumour mutation have the worst prognosis (**Figures 8C,D**). Tumour stem cells can associate malignant progression and immunotherapy tolerance in colon cancer. This study showed a negative correlation between risk score and stem cell index (R = -0.4, P < 2.2E-16), suggesting that high-risk patients have progressively lower stem cell characteristics and may have a better response to immunotherapy (**Figure 7E**). Combined with the results of immunotherapy typing in the GSE39582 dataset, there is a strong association between increased risk scores and sensitivity to different types of immunotherapies (**Figure 8F**). In addition, high risk

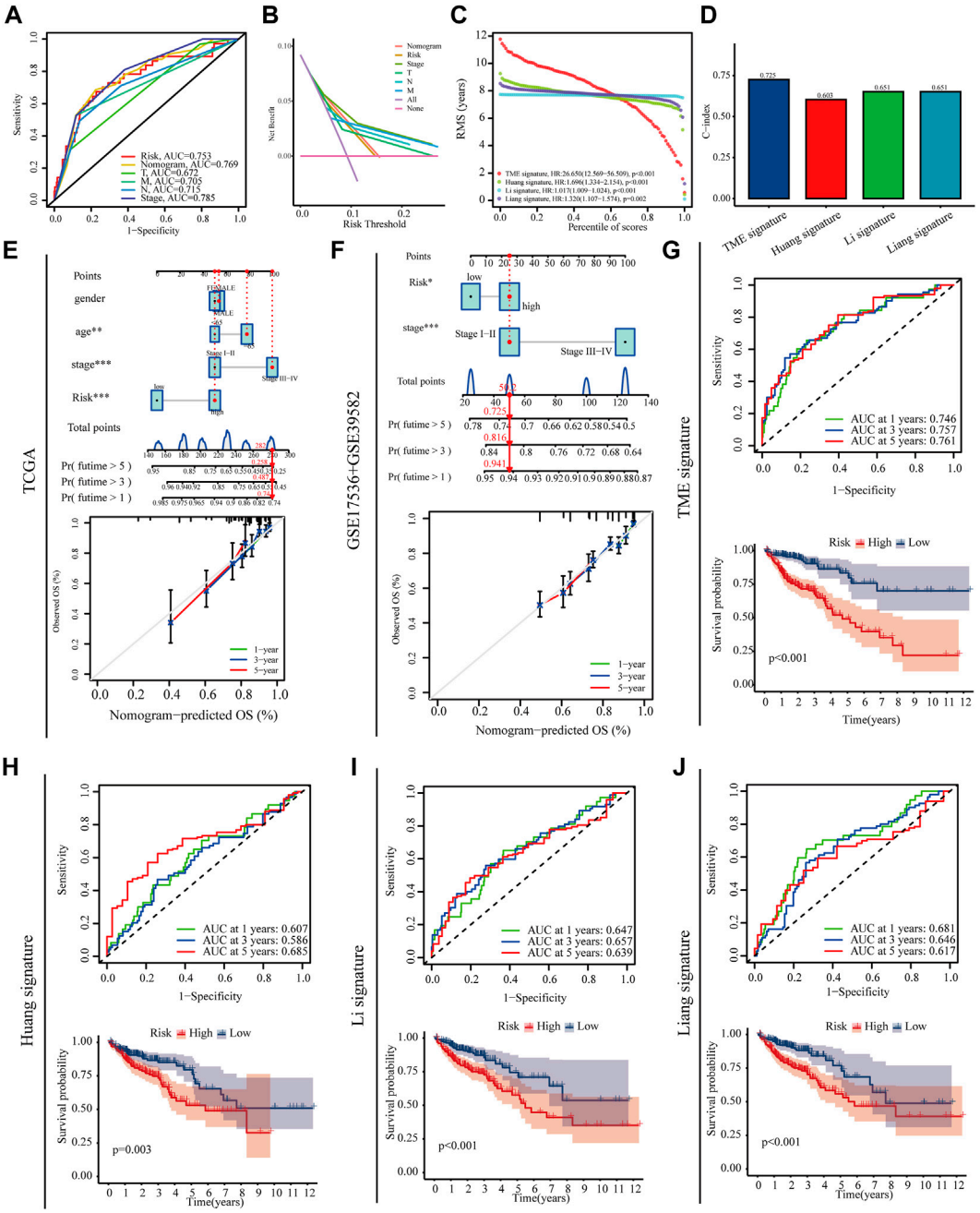


FIGURE 6 | Prognostic value of the risk model. **(A,B)** Roc and DCA curves were performed to validate the risk score in predicting patient' survival **(C,D)** RMS and C-index curves was used to compare the constructed model with existing models **(E,F)** Nomograms and calibration curves for predicting survival in colon cancer patients at 1, 3 and 5 years in training set and testing set **(G–J)** ROC and Kaplan–Meier survival curves was used to compare the constructed model with existing models.

patients have higher Tumour Immune Dysfunction and Rejection (TIDE) score, suggesting that high-risk patients may be less responsive to immunotherapy (**Figure 8G**). Finally, we found that patients in the high-risk group were significantly more sensitive to Gefitinib, Embelin and Dasatinib (**Figures 8H–J**). Taken together, these findings demonstrate that risk scores can accurately predict a patient's sensitivity to multiple clinical agents.

DISCUSSION

Studies have shown that methylation plays an indispensable role in antitumour effects, and m^7G is also a common mechanism of RNA methylation involved in the malignant biological behaviour of many tumours (Zhao et al., 2021a; Orellana et al., 2021). It has been reported that it also plays an important role in the

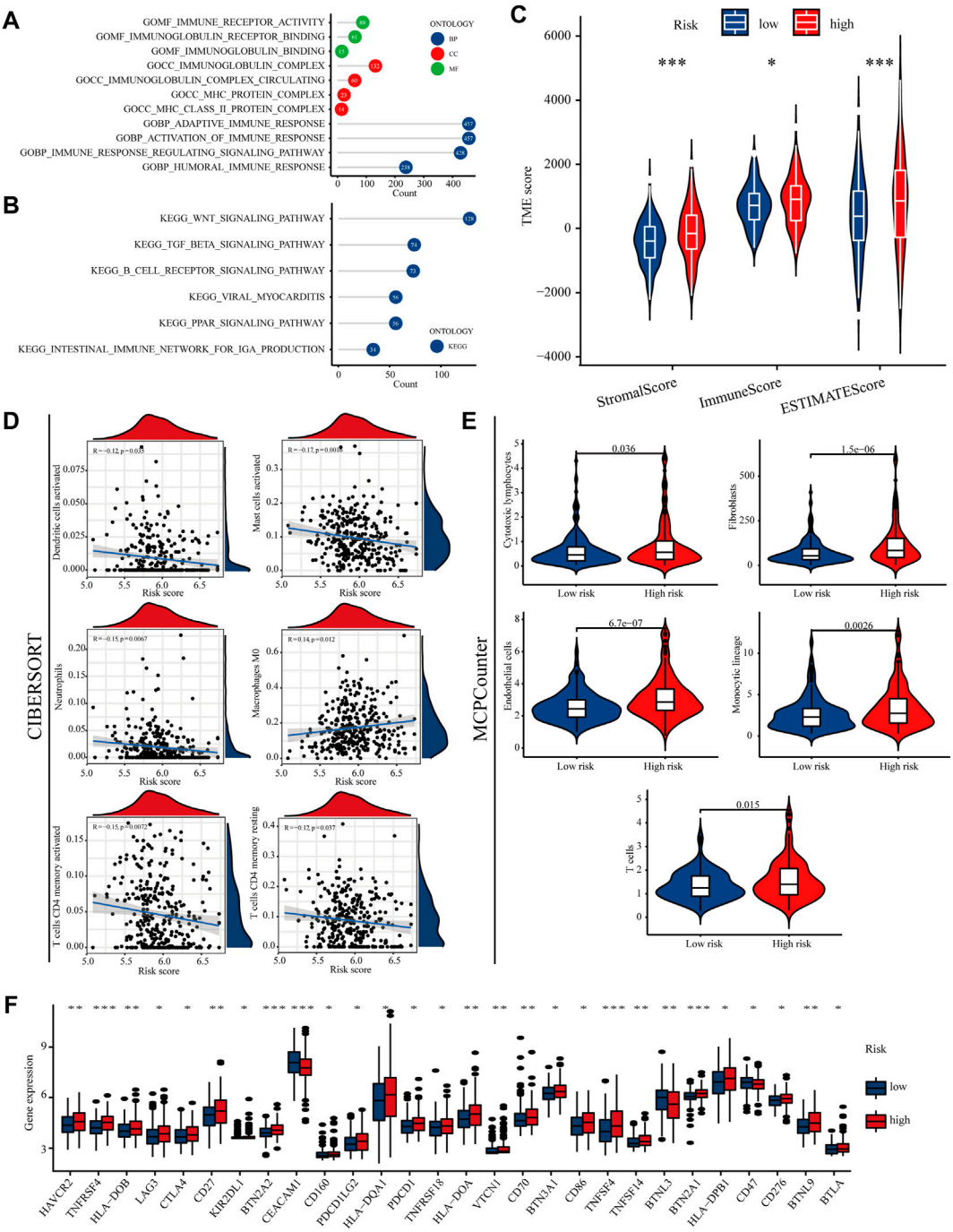


FIGURE 7 | Immune features in high- and low-risk groups **(A,B)** Different GSEA result between high- and low-risk groups **(C)** immune microenvironment scores between the high- and low-risk groups **(D,E)** MCPcounter and CIBERSORT were used to calculate the relative proportions of immune cell subsets **(F)** The expression of immune checkpoints between the high- and low-risk groups.

antitumour immune process (Devarkar et al., 2016). DNA damage is strongly associated with metastasis in colon cancer and also predicts responsiveness to immunotherapy in colon cancer patients (Sun et al., 2019; Mauri et al., 2020). Previous studies have shown that methylation modulates the UV-induced DNA damage response to promote damage repair (Xiang et al., 2017). However, the effects of the m^7G-related DDR genes on the immune microenvironment and immunotherapy in colon cancer have not been elucidated. Based on the TCGA database, an innovative prognostic model of m^7G-related DDR genes was constructed to predict the prognosis and sensitivity to immunotherapy in colon cancer patients.

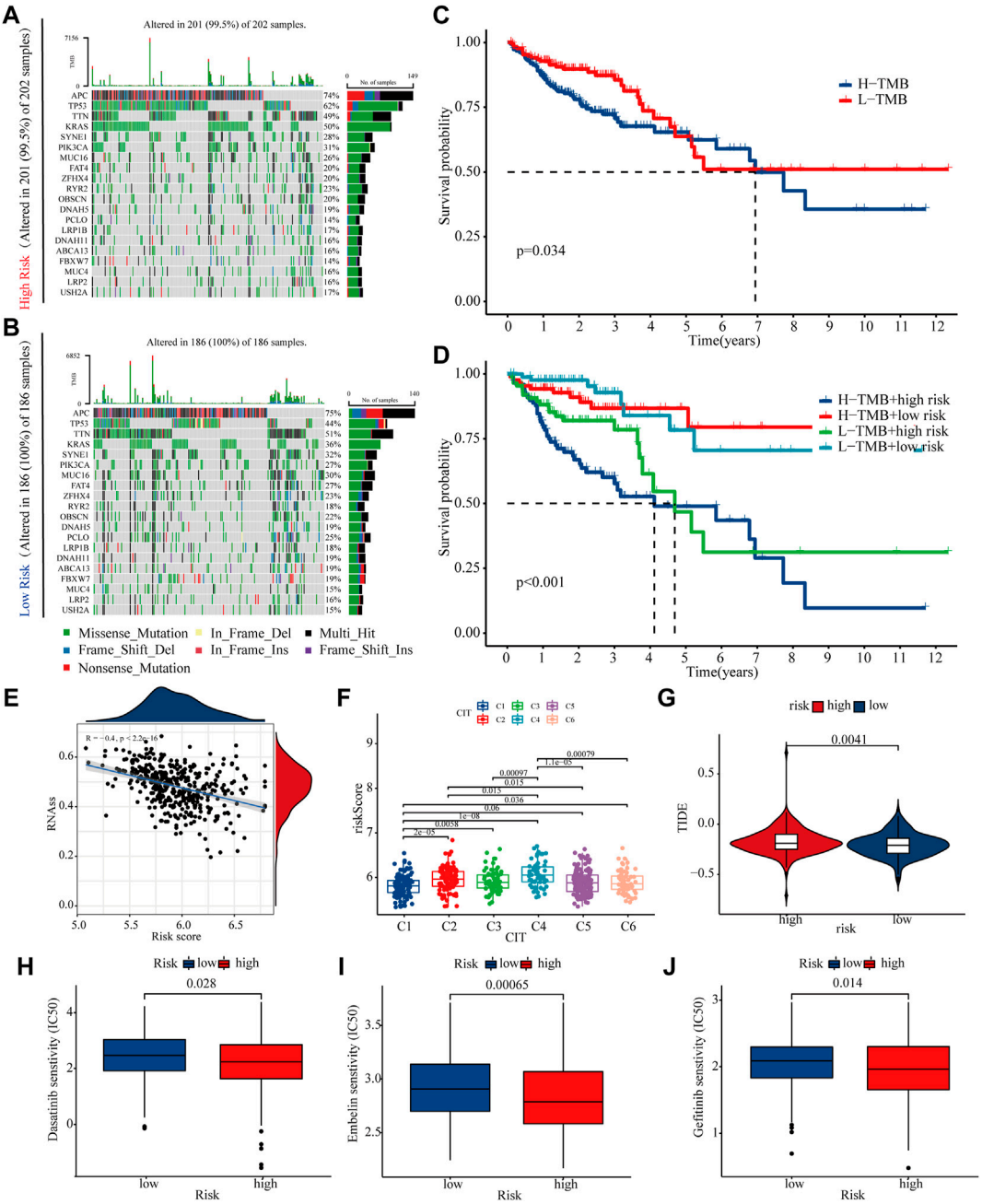


FIGURE 8 | Mutations and treatment sensitivity between high- and low-risk **(A,B)** The top 20 genes' Mutation rate in the high- and low-risk groups **(C,D)** K-M curves of risk score and tumour mutational load used to assess patient survival **(E)** The risk score is negatively correlated with the stem cell index **(F,G)** TIDE and GSE39582 database were used to validate patient sensitivity to immunotherapy between the high- and low-risk groups **(H–J)** The association between risk score and drug treatment sensitivity.

To validate the reliability of the model, the TCGA database was used as a training set and over 700 additional gastric cancer patients obtained from the GEO database were acted as a test set. The reliability of the model was further verified using ROC, C-index and DCA curves. Finally, the model constructed in this study also better than models constructed in other studies. All of these results indicate that the model constructed in this study is highly reliable in predicting patient prognosis.

We further analysed the mutation rates of ten risk genes in colon cancer patients and showed that ATP2A1 had the highest mutation rate. Thus, we suspect that ATP2A1 may be an important molecule influencing the malignant progression of

TABLE 1 | Information on patients included in this study.

	TCGA (n = 417)	GSE39582 (n = 566)	GSE17536 (n = 177)
Gender			
Female	193	256	81
Male	224	310	96
OS			
Alive	329	371	104
Dead	88	195	73
T Stage			
T1	9	11	
T2	74	45	
T3	284	367	
T4	49	119	
Unknow	1	24	
N Stage			
N0	247	302	
N1	98	134	
N2	72	98	
N3		6	
N4		26	
Unknow			
M Stage			
M0	311	482	
M1	57	61	
Unknow	49	23	
Stage			
I	72		24
II	160		57
III	117		57
IV	57		39
unknow	11		
Age			
<65	161		78
≥65	256		99

colon cancer. Immunohistochemical staining and WB showed that the expression of ATP2A1 in cancer tissues was significantly higher than that in normal tissues adjacent to the cancer. The knockdown of ATP2A1 expression significantly inhibited the proliferation and migration of colon cancer cell lines. However, inconsistent with the findings of this study, the expression of ATP2A1 in breast cancer tissues appears to be lower than in normal tissues, which may be due to the tumour heterogeneity (Christodoulou et al., 2021).

Having established that the model had high predictive significance, the study went on to analysis model's functions in guiding other clinical treatments of colon patients. Immunotherapy may improve prognosis in patients who have received inadequate conventional therapy, however, there is growing evidence that the tumour microenvironment has multiple immune escape mechanisms that impair antitumour immunity (Yang et al., 2020). Tumour-associated fibroblasts are closely associated with clinical treatment, and there is no consensus that they influence other cell populations in the tumour microenvironment, altering cancer progression, stromal remodelling, and drug resistance (Chen et al., 2021). M0 macrophages have been found to be significantly more abundant in tumour tissue than in adjacent normal tissue in patients with pancreatic ductal adenocarcinoma, and M0 macrophages are an independent predictor of poor prognosis (Xu et al., 2020). Mast cells have a dual effect on tumours, depending on the tumour type, cancer stage, the activation status of mast cells, etc. Increased mast cells are associated with longer patient survival in colorectal cancer (Mehdawi et al., 2016; Lichterman and Reddy, 2021). Dendritic cells are currently recognized as the most effective antigen-presenting cells, and patients have benefited from dendritic cell-based immunotherapy. National and international clinical trials evaluating dendritic cell-based tumour immunotherapy are currently underway (Steinman, 2012). This study revealed that the high-risk group had higher levels of M0 macrophages and fibroblasts, while activated mast cells, resting $CD4^{+}$T cells and dendritic cells activated were less abundant, implying that the risk model constructed in this study can be used as a biomarker to predict the immune status of colon cancer.

Gene mutations are closely associated with tumour development and treatment sensitivity (Hanahan, 2022). In this study, we found that the frequency of TP53 and KRAS mutations was significantly higher in patients in the high-risk group than those in the low-risk group. KRAS mutations have been a significant impediment to colorectal therapy, with numerous targeted therapies failing to work due to secondary KRAS mutations (Zhao et al., 2021b). Even though various targeted inhibitors of KRAS have been subjected to extensive clinical trials in recent years, the number of patients who have benefited from them remains limited (Awad et al., 2021; Tanaka et al., 2021). Tumour immunotherapy is a new generation of tumour treatment after surgery, radiotherapy, and chemotherapy, and has benefited a large number of patients (Waldman et al., 2020). However, patients with KRAS mutations have a significantly worse prognosis following immunotherapy than patients with KRAS wild-type (Ricciuti et al., 2021). Drug sensitivity results showed that patients with a low-risk score were more sensitive to lapatinib and sorafenib. Additionally, numerous clinical trials have demonstrated that lapatinib and sorafenib improve patients' prognoses (Bahrami et al., 2018; Wang and Fakih, 2021). In conclusion, the risk model constructed in this study serves as a guide for decision-making in targeted therapy and immune checkpoint inhibitor therapy for patients with colon cancer.

Despite the findings of this study, there are still some limitations. This study is mainly based on bioinformatics analysis and lacks experimental validation of the findings. Additionally, larger sample size is needed to verify the reliability of the model. In summary, a risk model for m7G-related genes was constructed in this study. This risk model can be used to predict the prognosis of colon cancer patients and to determine the efficacy of immunotherapy. These findings can provide a basis for prognosis prediction of colon cancer patients and to establish a reference value for colon cancer immunotherapy.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of Bengbu Medical College. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements. Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

(I) Conception and design: SC, ML; (II) Administrative support: None; (III) Provision of study materials: SC and RD; (IV) Collection and assembly of data: NZ, GP, FL, QQ (V) Data analysis and interpretation: HW; (VI) Manuscript writing: HW, YW; (VII) Final approval of manuscript: All authors. All authors read and approved the final manuscript.

FUNDING

This work was supported partly by funding from Anhui Natural Science Foundation Project (2108085MH291), Natural Science Project of Anhui Provincial Department of Education(KJ2021ZD0092), Scientific Research Project of Anhui Provincial Health and Family Planning Commission (2016QK016), Bengbu City Science and Technology Major Special Fund Project (BYLK201825),512 Talent Cultivation Plan of Bengbu Medical College (by51201107、by51201319), Research and Innovation Team of Bengbu Medical College (BYKC201909).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.918159/full#supplementary-material

REFERENCES

Andrew, A. S., Parker, S., Anderson, J. C., Rees, J. R., Robinson, C., Riddle, B., et al. (2018). Risk Factors for Diagnosis of Colorectal Cancer at a Late Stage: A Population-Based Study. *J. Gen. Intern Med.* 33, 2100–2105. doi:10.1007/s11606-018-4648-7

Awad, M. M., Liu, S., Rybkin, I. I., Arbour, K. C., Dilly, J., Zhu, V. W., et al. (2021). Acquired Resistance to KRASG12C Inhibition in Cancer. *N. Engl. J. Med.* 384, 2382–2393. doi:10.1056/nejmoa2105281

Bahrami, A., Hesari, A., Khazaei, M., Hassanian, S. M., Ferns, G. A., and Avan, A. (2018). The Therapeutic Potential of Targeting the BRAF Mutation in Patients with Colorectal Cancer. *J. Cell Physiol.* 233, 2162–2169. doi:10.1002/jcp.25952

Chen, Y., McAndrews, K. M., and Kalluri, R. (2021). Clinical and Therapeutic Relevance of Cancer-Associated Fibroblasts. *Nat. Rev. Clin. Oncol.* 18, 792–804. doi:10.1038/s41571-021-00546-5

Christodoulou, P., Yiallouris, A., Michail, A., Christodoulou, M. I., Politis, P. K., and Patrikios, I. (2021). Altered SERCA Expression in Breast Cancer. *Med. Kaunas. Lith.* 57, 1074. doi:10.3390/medicina57101074

Dai, Z., Liu, H., Liao, J., Huang, C., Ren, X., Zhu, W., et al. (2021). N7-Methylguanosine tRNA Modification Enhances Oncogenic mRNA Translation and Promotes Intrahepatic Cholangiocarcinoma Progression. *Mol. Cell* 81, 3339–3355. doi:10.1016/j.molcel.2021.07.003

Devarkar, S. C., Wang, C., Miller, M. T., Ramanathan, A., Jiang, F., Khan, A. G., et al. (2016). Structural Basis for m7G Recognition and 2′-O-Methyl Discrimination in Capped RNAs by the Innate Immune Receptor RIG-I. *Proc. Natl. Acad. Sci. U.S.A.* 113, 596–601. doi:10.1073/pnas.1515152113

Galassi, C., Vitale, I., and Galluzzi, L. (2021). Using Epigenetic Modifiers to Target Cancer Stem Cell Immunoevasion. *Cancer Cell* 39, 1573–1575. doi:10.1016/j.ccell.2021.11.003

Grady, W. M., Yu, M., and Markowitz, S. D. (2021). Epigenetic Alterations in the Gastrointestinal Tract: Current and Emerging Use for Biomarkers of Cancer. *Gastroenterology* 160, 690–709. doi:10.1053/j.gastro.2020.09.058

Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. *Cancer Discov.* 12, 31–46. doi:10.1158/2159-8290.cd-21-1059

Hu, F., Liu, C., Liu, L., Zhang, Q., and Guo, A. (2021). Expression Profile of Immune Checkpoint Genes and Their Roles in Predicting Immunotherapy Response. *Briefings Bioinform.* 22, bbaa176. doi:10.1093/bib/bbaa176

Huang, d. W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources. *Nat. Protoc.* 4, 44–57. doi:10.1038/nprot.2008.211

Huang, Z., Yang, Q., and Huang, Z. (2018). Identification of Critical Genes and Five Prognostic Biomarkers Associated with Colorectal Cancer. *Med. Sci. Monit.* 24, 4625–4633. doi:10.12659/msm.907224

Jühling, F., Mörl, M., Hartmann, R. K., Sprinzl, M., Stadler, P. F., and Pütz, J. (2009). tRNAdb 2009: Compilation of tRNA Sequences and tRNA Genes. *Nucleic Acids Res.* 37, D159–D162. doi:10.1093/nar/gkn772

Li, T., Hui, W., Halike, H., and Gao, F. (2021). RNA Binding Protein-Based Model for Prognostic Prediction of Colorectal Cancer. *Technol. Cancer Res. Treat.* 20, 15330338211019504. doi:10.1177/15330338211019504

Liang, Q., Du, X., Mao, L., and Wang, G. (2021). Molecular Characterization of Colorectal Cancer: A Five-Gene Prognostic Signature Based on RNA-Binding Proteins. *Saudi J. Gastroenterol.* 27, 223–233. doi:10.4103/sjg.sjg_530_20

Lichterman, J. N., and Reddy, S. M. (2021). Mast Cells: A New Frontier for Cancer Immunotherapy. *Cells* 10, 1270. doi:10.3390/cells10061270

Lieu, C. H., Golemis, E. A., Serebriiskii, I. G., Newberg, J., Hemmerich, A., Connelly, C., et al. (2019). Comprehensive Genomic Landscapes in Early and Later Onset Colorectal Cancer. *Clin. Cancer Res.* 25, 5852–5858. doi:10.1158/1078-0432.ccr-19-0899

Lin, S., Liu, Q., Lelyveld, V. S., Choe, J., Szostak, J. W., and Gregory, R. I. (2018). Mettl1/Wdr4-Mediated m7G tRNA Methylome is Required for Normal mRNA Translation and Embryonic Stem Cell Self-Renewal and Differentiation. *Mol. Cell* 71, 244–255. doi:10.1016/j.molcel.2018.06.001

Mauri, G., Arena, S., Siena, S., Bardelli, A., and Sartore-Bianchi, A. (2020). The DNA Damage Response Pathway as a Land of Therapeutic Opportunities for Colorectal Cancer. *Ann. Oncol.* 31, 1135–1147. doi:10.1016/j.annonc.2020.05.027

Mehdawi, L., Osman, J., Topi, G., and Sjölander, A. (2016). High Tumor Mast Cell Density is Associated with Longer Survival of Colon Cancer Patients. *Acta Oncol.* 55, 1434–1442. doi:10.1080/0284186x.2016.1198493

Miao, J., Liu, Y., Zhao, G., Liu, X., Ma, Y., Li, H., et al. (2020). SFRP2Feasibility of Plasma-Methylated for Early Detection of Gastric Cancer. *Cancer Control* 27, 1073274820922559. doi:10.1177/1073274820922559

Nacev, B. A., Jones, K. B., Intlekofer, A. M., Yu, J. S. E., Allis, C. D., Tap, W. D., et al. (2020). The Epigenomics of Sarcoma. *Nat. Rev. Cancer* 20, 608–623. doi:10.1038/s41568-020-0288-4

Orellana, E. A., Liu, Q., Yankova, E., Pirouz, M., De Braekeleer, E., Zhang, W., et al. (2021). METTL1-mediated m7G Modification of Arg-TCT tRNA Drives Oncogenic Transformation. *Mol. Cell* 81, 3323–3338. doi:10.1016/j.molcel.2021.06.031

Ricciuti, B., Arbour, K. C., Lin, J. J., Vajdi, A., Vokes, N., Hong, L., et al. (2021). Diminished Efficacy of Programmed Death-(Ligand)1 Inhibition in STK11- and KEAP1-Mutant Lung Adenocarcinoma is Affected by KRAS Mutation Status. *J. Thorac. Oncol.* 17, 399. doi:10.1016/j.jtho.2021.10.013

Russo, R., Marra, R., Andolfo, I., De Rosa, G., Rosato, B. E., Manna, F., et al. (2019). Characterization of Two Cases of Congenital Dyserythropoietic Anemia Type I Shed Light on the Uncharacterized C15orf41 Protein. *Front. Physiol.* 10, 621. doi:10.3389/fphys.2019.00621

Steinman, R. M. (2012). Decisions about Dendritic Cells: Past, Present, and Future. *Annu. Rev. Immunol.* 30, 1–22. doi:10.1146/annurev-immunol-100311-102839

Sun, J., Wang, C., Zhang, Y., Xu, L., Fang, W., Zhu, Y., et al. (2019). Genomic Signatures Reveal DNA Damage Response Deficiency in Colorectal Cancer Brain Metastases. *Nat. Commun.* 10, 3190. doi:10.1038/s41467-019-10987-3

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Tanaka, N., Lin, J. J., Li, C., Ryan, M. B., Zhang, J., Kiedrowski, L. A., et al. (2021). Clinical Acquired Resistance to KRASG12C Inhibition through a Novel KRAS Switch-II Pocket Mutation and Polyclonal Alterations Converging on RAS-MAPK Reactivation. *Cancer Discov.* 11, 1913–1922. doi:10.1158/2159-8290.cd-21-0365

Tomikawa, C. (2018). 7-Methylguanosine Modifications in Transfer RNA (tRNA). *Int. J. Mol. Sci.* 19, 4080. doi:10.3390/ijms19124080

Waldman, A. D., Fritz, J. M., and Lenardo, M. J. (2020). A Guide to Cancer Immunotherapy: From T Cell Basic Science to Clinical Practice. *Nat. Rev. Immunol.* 20, 651–668. doi:10.1038/s41577-020-0306-5

Wang, C., and Fakih, M. (2021). Response to Trastuzumab and Lapatinib in a Metastatic Colorectal Cancer Harboring HER2 Amplification and HER2 S310F Mutation. *J. Natl. Compr. Cancer Netw.* 19, 670–674. doi:10.6004/jnccn.2021.7023

Xiang, Y., Laurent, B., Hsu, C.-H., Nachtergaele, S., Lu, Z., Sheng, W., et al. (2017). RNA m6A Methylation Regulates the Ultraviolet-Induced DNA Damage Response. *Nature* 543, 573–576. doi:10.1038/nature21671

Xu, C., Sui, S., Shang, Y., Yu, Z., Han, J., Zhang, G., et al. (2020). The Landscape of Immune Cell Infiltration and its Clinical Implications of Pancreatic Ductal Adenocarcinoma. *J. Adv. Res.* 24, 139–148. doi:10.1016/j.jare.2020.03.009

Yang, Q., Shi, G., Chen, X., Lin, Y., Cheng, L., Jiang, Q., et al. (2020). Nanomicelle Protects the Immune Activation Effects of Paclitaxel and Sensitizes Tumors to Anti-PD-1 Immunotherapy. *Theranostics* 10, 8382–8399. doi:10.7150/thno.45391

Ying, X., Liu, B., Yuan, Z., Huang, Y., Chen, C., Jiang, X., et al. (2021). METTL1-m7G-EGFR/EFEMP1 axis Promotes the Bladder Cancer Development. *Clin. Transl. Med.* 11, e675. doi:10.1002/ctm2.675

Zhao, Y., Kong, L., Pei, Z., Li, F., Li, C., Sun, X., et al. (2021). m7G Methyltransferase METTL1 Promotes Post-ischemic Angiogenesis via Promoting VEGFA mRNA Translation. *Front. Cell Dev. Biol.* 9, 642080. doi:10.3389/fcell.2021.642080

Zhao, Y., Murciano-Goroff, Y. R., Xue, J. Y., Ang, A., Lucas, J., Mai, T. T., et al. (2021). Diverse Alterations Associated with Resistance to KRAS(G12C) Inhibition. *Nature* 599, 679–683. doi:10.1038/s41586-021-04065-2

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.

Copyright © 2022 Chen, Dong, Li, Zheng, Peng, Lu, Qiu, Wen, Wang, Wu and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 24 June 2022
doi: 10.3389/fgene.2022.922358

A Novel M7G-Related MicroRNAs Risk Signature Predicts the Prognosis and Tumor Microenvironment of Kidney Renal Clear Cell Carcinoma

Peng Hong[1†], Huifang Du[2†], Ming Tong[1‡], Qingfei Cao[1*‡], Ding Hu[1], Jiaji Ma[1], Yanyang Jin[1], Zizhi Li[1], Weichao Huang[1] and Guangquan Tong[1]*

[1]Department of Urology, The First Hospital of Jinzhou Medical University, Jinzhou Medical University, Jinzhou, China, [2]Affiliated First Hospital, Nanchang University, Nanchang, China

OPEN ACCESS

Edited by:
Ehsan Nazemalhosseini-Mojarad, Shahid Beheshti University of Medical Sciences, Iran

Reviewed by:
Zahra Salehi, Shahid Beheshti University of Medical Sciences, Iran
Shaghayegh Bradaran Ghavami, Shahid Beheshti University of Medical Sciences, Iran

***Correspondence:**
Ming Tong
tongming@jzmu.edu.cn
Qingfei Cao
caoqingfei@jzmu.edu.cn

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

[‡]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: 17 April 2022
Accepted: 07 June 2022
Published: 24 June 2022

Citation:
Hong P, Du H, Tong M, Cao Q, Hu D, Ma J, Jin Y, Li Z, Huang W and Tong G (2022) A Novel M7G-Related MicroRNAs Risk Signature Predicts the Prognosis and Tumor Microenvironment of Kidney Renal Clear Cell Carcinoma. Front. Genet. 13:922358. doi: 10.3389/fgene.2022.922358

Background: M7G modification is extremely vital for the development of many cancers, especially tumor immunity. M7G modification is a novel functional regulator of miRNA, and the researches on m7G-related miRNAs in kidney renal clear cell carcinoma (KIRC) are still insufficient. This research aims to establish a risk signature on the foundation of m7G-associated miRNAs, which can precisely forecast the prognosis of KIRC patients.

Methods: Transcriptome data and clinical data used in this study come from The Cancer Genome Atlas database. Our team utilized univariable Cox, Lasso and multivariable Cox analyses to construct a m7G-associated miRNAs risk signature that can forecast the prognosis of KIRC patients. Kaplan-Meier method, time-dependent receiver operating characteristic (ROC) curve, and the independent analysis of risk signatures were employed to verify the predictability and accuracy of the risk signature. Subsequently, based on CIBERSORT, ESTIMATE and ssGSEA algorithms, we speculated the potential impact of the proposed risk signature on tumor immune microenvironment. Ultimately, by virtue of the risk signature and tumor immunity, the hub genes affecting the prognosis of KIRC patients were screened out.

Results: Our team established and verified a prognostic signature comprising 7 m7G-associated miRNAs (*miR-342-3p, miR-221-3p, miR-222-3p, miR-1277-3p, miR-6718-5p, miR-1251-5p, and miR-486-5p*). The results of the Kaplan-Meier survival analysis revealed that the prognosis of KIRC sufferers in the high-risk group was often unsatisfactory. The accuracy of the prediction ability of the risk signature was verified by calculating the area under the ROC curve. Univariate-multivariate Cox analyses further showed that this risk signature could be utilized as an independent prognosis-related biomarker for KIRC sufferers. The results of the immune analysis revealed that remarkable

Abbreviations: AUC, the area under the ROC curve; CIBERSORT, Cell-type Identification by Estimating Relative Subsets of RNA Transcripts; DEGs, differentially expressed genes; DO, Disease Ontology; ESTIMATE, Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data; FDR, False Discovery Rate; GO, Gene Ontology; ICB, immune checkpoint blocking; KEGG, Kyto Encyclopedia of Genes and Genomes; KIRC, kidney renal clear cell carcinoma; K-M, Kaplan-Meie; Log FC, Log Fold Change; m7G, N7-methylguanosine; miRNAs, MicroRNAs; OS, overall survival; ROC, receiver operating characteristic; ssGSEA, single sample Gene Set Enrichment Analysis; TCGA, The Cancer Genome Atlas; TIDE, Tumor Immune Dysfunction and Exclusion; TMB, Tumor mutation burden; TME, tumor microenvironment.

diversities existed in immune status and tumor microenvironment between high-risk and low-risk groups. On the foundation of the proposed risk signature and other clinical factors, a nomogram was established to quantitatively forecast the survival of KIRC sufferers at 1, 3 and 5 years.

Conclusion: Based on m7G-related miRNAs, a risk signature was successfully constructed, which could precisely forecast the prognosis of sufferers and guide personalized immunotherapy for KIRC patients.

Keywords: kidney renal clear cell carcinoma, prognostic signature, m7G, microRNA, tumor immune microenvironment, The Cancer Genome Atlas

INTRODUCTION

Renal cell cancer takes up 2%–3% of adult malignancies, second only to prostate carcinoma and bladder carcinoma in urinary system cancers, whereas it is the most deadly malignancy of the urinary system (Humphrey et al., 2016; Rouprêt et al., 2021). Kidney renal clear cell carcinoma (KIRC) is the most commonly seen histologic type of renal cell cancer, taking up approximately 80% (Pandey and Syed, 2022). According to Cancer Statistics, there will be about 79,000 new cases of kidney carcinoma in the United States in 2022, with approximately 13,920 new deaths due to such disease (Siegel et al., 2022). In recent years, the incidence of KIRC has been increasing, and the age of KIRC patients has become younger (Aragon-Ching et al., 2019). Partial nephrectomy and radical nephrectomy are still the most important and effective treatment methods, whereas about 20%–30% KIRC patients encounter recurrence after surgical treatment, which reduces the 5-year survival rate to approximately 23% (Maclennan et al., 2012; Campbell et al., 2017; Hsieh et al., 2017; Speed et al., 2017). Although some new therapeutic methods have been developed, such as targeted drug therapy and immunotherapy, the clinical outcomes of advanced KIRC are still poor (Richter and Dvořák, 2018). The poor prognoses of patients with KIRC are still a clinical challenge, with delayed diagnoses and high metastasis rates being important factors inducing the poor prognoses of KIRC sufferers. Consequently, it is pivotal to propose a novel biomarker and molecular target for KIRC patients.

RNA methylation is a posttranscription modification extensive in eukaryotic cells and prokaryotes, and methylation denotes the transfer of a methyl group from one active methyl compound to another (Zhang et al., 2021). Depending on methylation sites such as m6A, m5C, m7G, and 2-O, RNA methylation can be divided into various methylation modifications (Hori, 2014; Boccaletto et al., 2018; Zhang et al., 2021). N7-methylguanosine (m7G) is a modification of the 7th N of RNA guanine with a methyl group (Alexandrov et al., 2002). As post-transcriptional modifications, m7G plays a momentous function in the initiation of miRNA biogenesis, cell migration, stability, translation, immunogenicity and many other biological processes such as Writers (methyltransferases), Erasers (demethylases), Readers (binding proteins) (Tomikawa, 2018; Zhang et al., 2021). M7G modification exists not only on mRNAs, but also on miRNAs (Pandolfini et al., 2019). M7G methylation complexes include *METTL1* (methyltransferase like 1) and *WDR4* (WD Repeat Domain 4) (Lin et al., 2018). M7G modification affects tumor progression in many ways, especially in tumor immunity. For example, *METTL1* promotes the processing of let-7e microRNA through m7G methylation, and then regulates the let-7e miRNA/*HMGA2* axis, thus inhibiting the occurrence and development of colon cancer (Baradaran Ghavami et al., 2017; Pandolfini et al., 2019; Liu et al., 2020). *METTL1* enhances the translation of oncogenic mRNA through m7G tRNA modification to promote the development of intrahepatic cholangiocarcinoma (Dai et al., 2021). Abnormal translation mediated by *METTL1*/*WDR4*-mediated m7G tRNA modification promotes the progression of head and neck squamous cell carcinoma (Chen J. et al., 2022).

MicroRNAs (miRNAs) are short single-stranded ncRNA molecules (19–25 nucleotides), which regulate gene expression after transcription and are indispensable in many biological processes like differentiation, apoptosis, drug resistance, proliferation, and metastasis (Bartel, 2009). Some recent scientific studies have unveiled that abnormally expressed miRNAs can affect the progression of many diseases to a certain extent, including stomach cancer, bladder cancer, lung adenocarcinoma, prostate cancer and colorectal cancer, etc. (Inamoto et al., 2018; Sharma and Baruah, 2019). MiRNAs are highly stable in a variety of biological samples, including tissue, serum and saliva, and their detection is easy as well (Jung et al., 2010). These characteristics provide favorable conditions for miRNAs to become candidate biomarkers for the diagnosis of cancer or precancerous lesions and prognosis. M7G modification is a novel functional regulator of miRNA, and the researches on m7G-associated miRNAs in KIRC are still insufficient. We need to further clarify the potential association between m7G-associated miRNAs and KIRC, investigating the significance of m7G-associated miRNAs as prognostic biomarkers in KIRC patients and providing ideas for developing new treatment methods for KIRC.

In this research, our team utilized The Cancer Genome Atlas (TCGA) public database to construct and verify a risk signature on the foundation of m7G-associated miRNAs, which could accurately forecast the prognoses of KIRC sufferers. It is proven that this risk signature can be utilized as an independent prognosis biomarker for KIRC patients, and it can also facilitate individualized immunotherapy for KIRC patients clinically. Finally, we used our risk signature and other clinical factors to establish a nomogram that

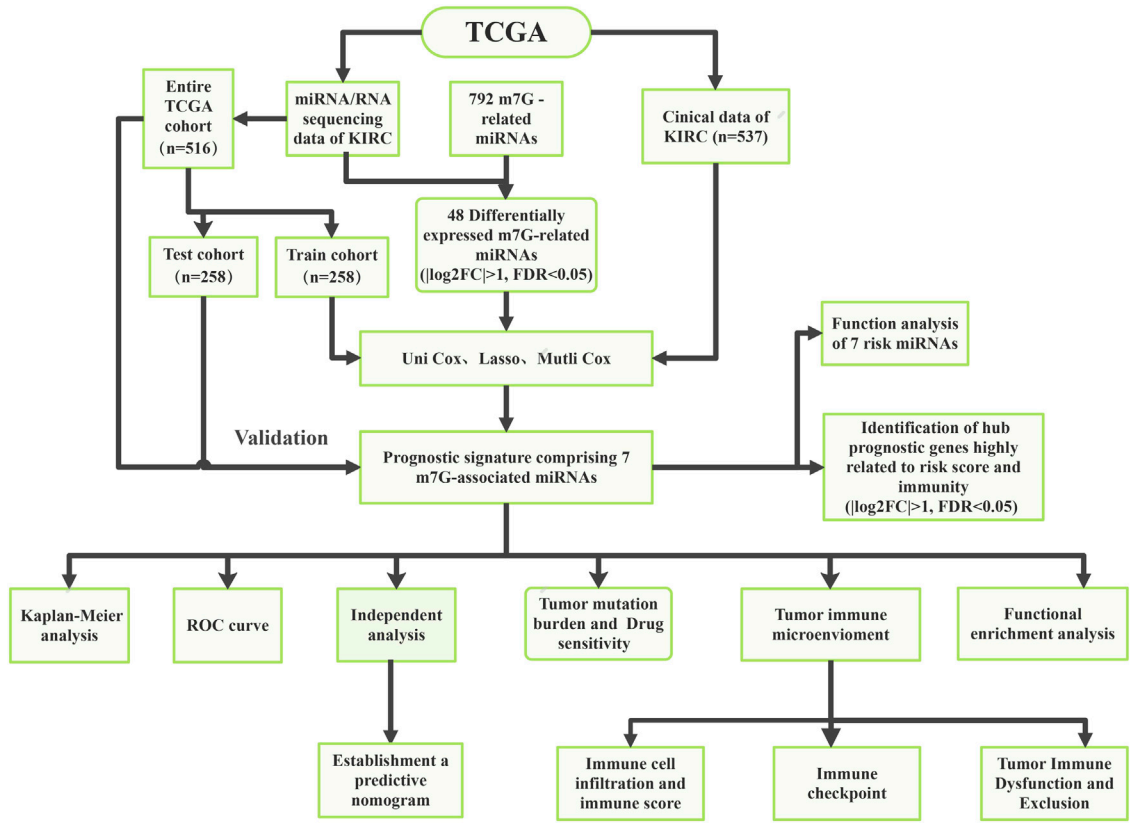


FIGURE 1 | The flow chart of this study.

could quantitatively forecast the overall survival (OS) rate of KIRC sufferers.

MATERIALS AND METHODS

Transcriptome Data and Clinical Data Acquisition of KIRC Patients

The flowchart of the present research is displayed by **Figure 1**. The clinical data and sequencing data of KIRC patients in the present research were obtained from the TCGA database (Creighton et al., 2013). Mature miRNA sequences were obtained in Fasta format from miRNAbase and used for miRNA annotation. The miRNA sequencing data (Isoform Expression) of 545 KIRC tissular specimens and 71 healthy renal tissular specimens were obtained. The RNA sequencing data (TCGA-KIRC, HTSeq-Counts) of 539 KIRC tissular specimens and 72 healthy renal tissular specimens were also obtained from the TCGA database. In addition, the corresponding clinical data were acquired from the TCGA database. TCGA pan-cancer data covering 33 tumors were downloaded from the UCSC Xean browser, and 18 cancer types along with their corresponding normal tissue samples were utilized for further study.

By excluding some samples with incomplete data, a total of 516 KIRC samples with complete miRNA/RNA sequencing data and clinical data were used for subsequent analysis (**Table 1**). The 516 KIRC samples were randomized 1:1 into two cohorts: the training cohort (n = 258, for the construction of the m7G-related miRNAs risk signature) and the testing cohort (n = 258, for the verification of the risk signature).

Screening of Differentially Expressed m7G-Related MicroRNAs

From previous literature reports, we obtained two m7G-related genes (*METTL1* and *WDR4*) that were confirmed to be associated with m7G modification on miRNA (Lin et al., 2018; Dai et al., 2021; Zhang et al., 2021). We used the TargetScan (https://www.targetscan.org) online database to predict the upstream miRNA of m7G-related genes and obtained m7G-related miRNAs. The "edgeR" package of R program was used to select differentially expressed m7G-associated miRNAs in KIRC tissue samples and normal kidney tissue samples (selection criteria: |log2 Fold Change (FC)| > 1, False Discovery Rate (FDR) < 0.05) (Liu et al., 2021). The expression of differentially expressed m7G-related miRNAs was shown by the heatmap.

TABLE 1 | Clinicopathological characteristics of 516 ccRCC patients in TCGA-KIRC cohort.

Characteristic	Type	Total TCGA cohort (N = 516)	Testing cohort (*N* = 258)	Training cohort (*N* = 258)	*p* value
Age	≤65	342 (66.28%)	174 (67.7%)	168 (64.86%)	0.5558
	>65	174 (33.72%)	83 (32.3%)	91 (35.14%)	
Gender	FEMALE	181 (35.08%)	90 (35.02%)	91 (35.14%)	1
	MALE	335 (64.92%)	167 (64.98%)	168 (64.86%)	
Grade	G1	13 (2.52%)	6 (2.33%)	7 (2.7%)	0.836
	G2	218 (42.25%)	103 (40.08%)	115 (44.4%)	
	G3	202 (39.15%)	104 (40.47%)	98 (37.84%)	
	G4	75 (14.53%)	40 (15.56%)	35 (13.51%)	
	GX	5 (0.97%)	3 (1.17%)	2 (0.77%)	
	Unknow	3 (0.58%)	1 (0.39%)	2 (0.77%)	
Stage	Stage I	253 (49.03%)	117 (45.53%)	136 (52.51%)	0.1585
	Stage II	55 (10.66%)	30 (11.67%)	25 (9.65%)	
	Stage III	123 (23.84%)	59 (22.96%)	64 (24.71%)	
	Stage IV	82 (15.89%)	49 (19.07%)	33 (12.74%)	
	unknow	3 (0.58%)	2 (0.78%)	1 (0.39%)	
T	T1	21 (4.07%)	10 (3.89%)	11 (4.25%)	0.921
	T1a	132 (25.58%)	59 (22.96%)	73 (28.19%)	
	T1b	106 (20.54%)	53 (20.62%)	53 (20.46%)	
	T2	54 (10.47%)	27 (10.51%)	27 (10.42%)	
	T2a	9 (1.74%)	4 (1.56%)	5 (1.93%)	
	T2b	4 (0.78%)	3 (1.17%)	1 (0.39%)	
	T3	5 (0.97%)	3 (1.17%)	2 (0.77%)	
	T3a	119 (23.06%)	66 (25.68%)	53 (20.46%)	
	T3b	53 (10.27%)	25 (9.73%)	28 (10.81%)	
	T3c	2 (0.39%)	1 (0.39%)	1 (0.39%)	
	T4	11 (2.13%)	6 (2.33%)	5 (1.93%)	
M	M0	406 (78.68%)	195 (75.88%)	211 (81.47%)	0.1328
	M1	78 (15.12%)	47 (18.29%)	31 (11.97%)	
	MX	30 (5.81%)	14 (5.45%)	16 (6.18%)	
	unknow	2 (0.39%)	1 (0.39%)	1 (0.39%)	
N	N0	228 (44.19%)	115 (44.75%)	113 (43.63%)	0.1884
	N1	17 (3.29%)	12 (4.67%)	5 (1.93%)	
	NX	271 (52.52%)	130 (50.58%)	141 (54.44%)	

Establishment and Validation of the Predictive Risk Signature Based on m7G-Related MicroRNAs

Univariable Cox analyses were utilized to assess the prognosis significance of the m7G-related miRNA with differential expression ($p < 0.05$). Lasso regressive analyses were further used to optimize the selection of prognostic m7G-related miRNAs (Tibshirani, 1997), and then these most representative prognostic m7G-related miRNAs were analyzed by multivariate Cox regression. After the above analysis, specific risk miRNAs and their corresponding coefficients were given, and a prognosis signature on the foundation of m7G-associated miRNAs was constructed. The risk scoring for every KIRC sufferer can be computed as:

$$\text{Risk score} = \sum_{i=1}^{n} coef\,(miRNA) * Exp\,(miRNA)$$

Taking the median risk score of the training cohort as the threshold, KIRC patients in each cohort (training cohort, test cohort, and full TCGA cohort) were divided into the $risk_{high}$ group and $risk_{low}$ group for subsequent analysis and validation. The Chi-square test was employed to verify the association between clinic features and risk scoring. The Kaplan-Meier (K-M) survival curve (Log-rank test) was employed to evaluate OS diversities between these 2 groups. The accurateness of the modeling method was evaluated by calculating the area under the ROC curve (AUC) values. Univariable-multivariable Cox proportional risk analysis was utilized to evaluate whether risk scores and clinical features were independent prognostic biomarkers for KIRC patients.

Establishment and Calibration of Nomogram

On the foundation of risk scoring and other clinical independent prognostic factors, a nomogram risk identification model was established *via* R-package "rms". Nomogram can quantify the factors that affect the prognosis of patients with KIRC and predict the prognosis of patients with KIRC quantitatively. Subsequently, calibration curves were developed to illustrate the predictive power of the established nomogram. The calibration curve describes the calibration of each model according to the consistency between the predicted survival time of KIRC patients and the observed survival time of KIRC patients. The *y*-axis represents the actual survival time of KIRC patients. The *x*-axis represents the predicted survival time of KIRC patients. The dotted line indicates the perfect prediction of the risk model. The solid line

of pink indicates the performance of nomogram, and the dotted line closer to diagonal indicates better prediction.

Analysis of the Association Between Risk Signature and Immunocyte Infiltration, Tumor Microenvironment

We quantified the quantity of immunocytes in every specimen using the "Cell-type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT)" algorithm (Newman et al., 2015). The single sample Gene Set Enrichment Analysis (ssGSEA) algorithm was employed to calculate 16 immunocyte infiltration and 13 immunity-related function scores for each KIRC sample. The diversities in immunocyte infiltration and immunofunction amongst diverse risk score subgroups were contrasted. We calculated the diversities in the expressing levels of immune checkpoint-related genes between diverse groups. The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was employed to assess the efficacy of potential immune checkpoint blocking (ICB) response (http://tide.dfci.harvard.edu) (Jiang et al., 2018). The tumor microenvironment (TME) scoring for every KIRC sufferer was quantified via the "Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression data (ESTIMATE)" algorithm (Yoshihara et al., 2013), including stroma scoring, immunoscore, and estimated scoring, and the potential differences in TME between different risk groups were estimated by comparing the diversities in TME scores between the $risk_{high}$ group and $risk_{low}$ group.

Tumor Mutation Burden and Prediction of Potential Therapeutic Drug Sensitivity

We downloaded the somatic cell mutation data of KIRC tissue samples from the TCGA database (TCGA.KIRC.varscan.somatic.maf.gz). Tumor mutation burden (TMB) scoring was calculated based on somatic mutation data. The IC_{50} values of therapeutic medicines were calculated by the R package "pRRophetic". The values of IC_{50} represent 50% inhibited cells, that is, the cell survival rate is half of the control sample, which is the corresponding drug concentration. The lower the IC_{50} value, KIRC patients are more sensitive to this drug.

Analysis of Genes Associated With Immunity and Risk Score

R package "edgeR" was employed to screen for differentially expressed genes (DEGs) in the $risk_{high}$ group and $risk_{low}$ group (screening criteria: $|log2FC| > 1$, FDR < 0.05). As per the immunoscore in the TME scoring of each KIRC patient, KIRC patients were separated into 2 groups: the high-immunity group and the low-immunity group. The differentially expressed immune-related genes were obtained by the same method. Risk-immune-related genes were acquired via the intersection of risk differential genes and immune differential genes. The prognostic significance of these risk-immunity-associated genes was assessed by univariable Cox analyses. Spearman correlative analysis was employed to evaluate the association of prognostic risk-immunity-associated genes with immunocytes and immunofunction. Gene Ontology (GO), Kyto Encyclopedia of Genes and Genomes (KEGG) and Disease Ontology (DO) analyses were completed on risk-immunity-associated genes to explore the biological processes and potential signal pathways related to these genes (Chen et al., 2017). These prognostic risk-immune-related genes were applied to the construction of protein-protein interaction (PPI) networks through search tool for recurring instances of neighbouring genes (STRING) online database (medium confidence > 0.4), and then the hub genes were extracted from PPI networks using the CytoHubba plugin of Cytoscape (https://cn.string-db.org/cgi) (Zhou et al., 2021).

Statistical Analysis

Log-rank test was used for the K-M approach to create the survival curve. The Chi-square test was utilized to determine the diversities between different data sets or different classified data, and the Wilcoxon-rank test was utilized to determine the diversities between these 2 groups. Spearman method was used for correlation analysis, and cox proportional regressive analysis was employed to evaluate the influence of gene expression and clinical features on the prognoses of sufferers. When $p < 0.05$, the above statistical methods had statistical significance. All analysis was conducted using R version 4.1.0 and the corresponding feature packages.

RESULTS

The Expression Level of m7G-Related Genes and m7G-Related MicroRNAs

Our team studied the expressing level of 2 m7G-associated genes *METTL1* and *WDR4* in 18 tumor types and found that these two m7G-related genes were differentially expressed in many tumors, showing significant heterogeneity (**Figures 2A,B**). In KIRC, the expressions of *METT1* and *WDR4* in cancer samples were significantly greater vs. healthy samples ($p < 0.01$), the greater the expressing levels of *METT1* and *WDR4*, the worse the prognoses of KIRC sufferers (**Figures 2C,D**). The upstream miRNAs of these two m7G-related genes were predicted by the TargetScan online database (https://www.targetscan.org), and altogether 792 m7G-related miRNAs were acquired (**Supplementary Table S1**). The expression of these 792 m7G-related miRNAs in 545 KIRC samples and 71 normal samples was analyzed. Altogether 48 differentially expressed m7G-related miRNAs were acquired ($|log2FC| > 1$, FDR < 0.05), and the expressing levels of 22 miRNAs were low in cancer tissues while the expressing levels of 26 miRNAs were high in tumor tissues (**Supplementary Table S2**). The results are presented in the shape of a volcano plot (**Figure 2E**). The heatmap shows the expression of 48 differentially expressed m7G-related miRNAs (**Figure 2F**).

Construction and Verification of the Prognostic Risk Model Based on m7G-Related MicroRNAs

By taking the survival time as the prognostic index, a total of 516 patients with KIRC were used to construct and verify this model.

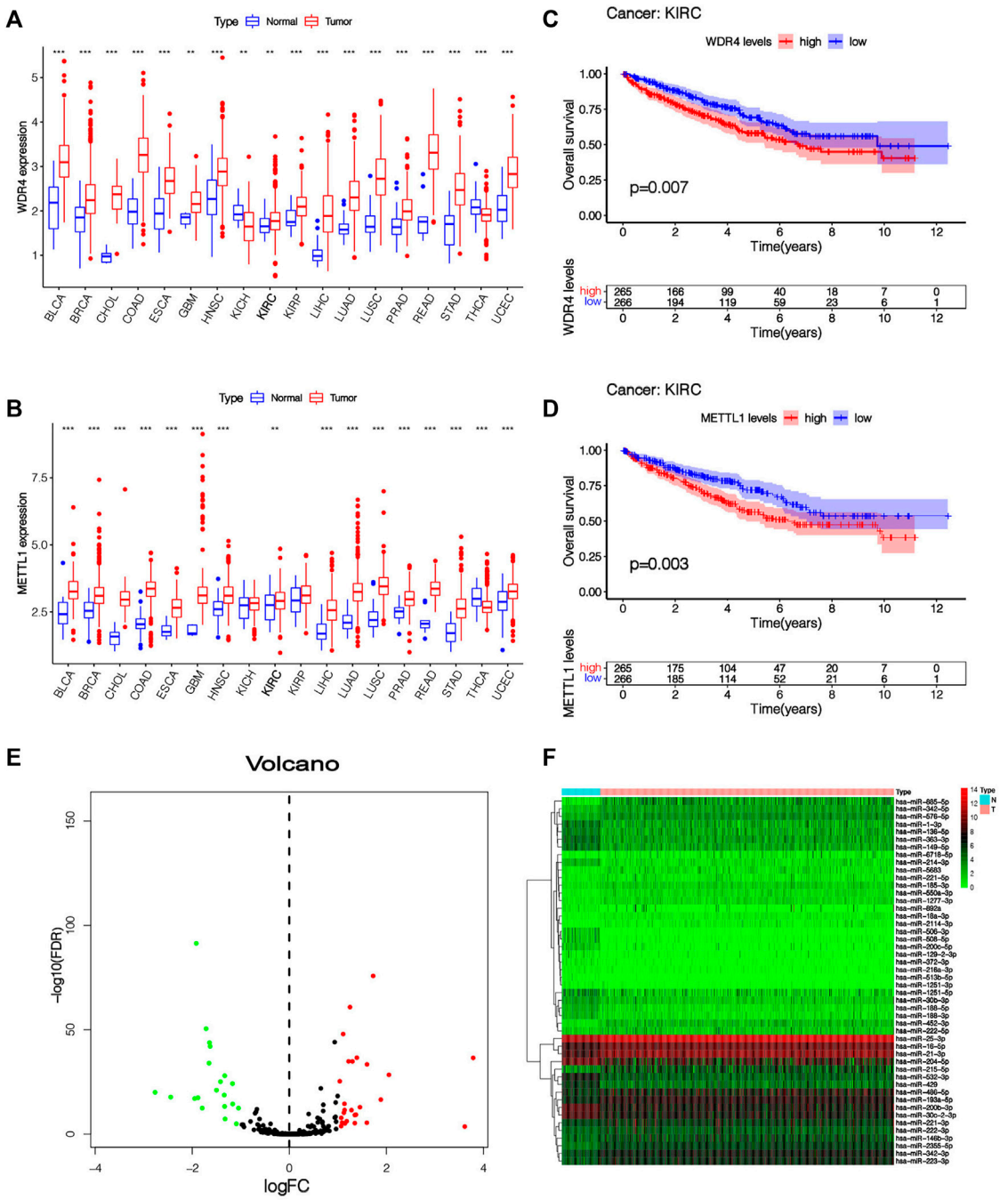


FIGURE 2 | The expression level of m7G-related genes and m7G-related miRNAs. **(A,B)** Pan-cancer analysis of METTL1 and WDR4 in 18 tumor types. **(C,D)** K-M survival curve analysis of METTL1 and WDR4 in KIRC patients. **(E)** The volcanic plot of differentially expressed m7G-related miRNAs between KIRC tissue and normal tissue, which indicates that the distribution of log2FC and *p* values is reasonable. **(F)** Heatmap of 48 differentially expressed miRNAs expressed in KIRC tissue and normal tissue.

Patients were stochastically separated into the training cohort and test cohort (**Supplementary Table S3**). In the training cohort, 48 differentially expressed m7G-related miRNAs were analyzed by univariate Cox regression, and altogether 18 m7G-related miRNAs with prognostic values were obtained (**Figure 3A**). Then, in order to optimize our model, LASSO regression analysis was performed on 18 prognostic m7G-related miRNAs to eliminate highly correlated prognostic miRNAs to avoid overfitting, and 13 most representative candidate miRNAs were found (**Figures 3B,C**). Cross-validation results show that LASSO regression analysis was the best. Multivariable Cox regressive analyses were completed on those 13 candidate miRNAs, and finally, 7 risk miRNAs and their

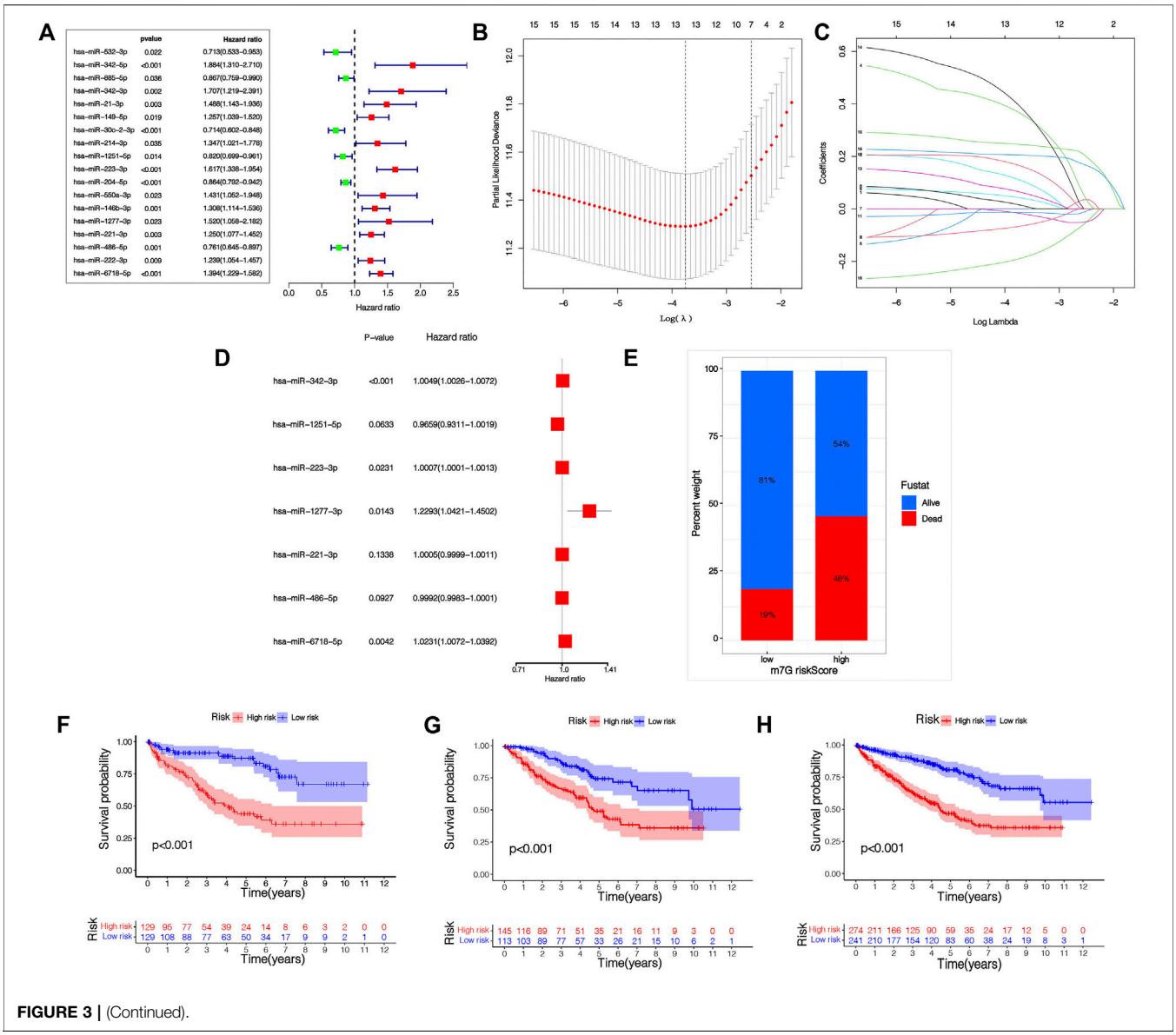


FIGURE 3 | (Continued).

corresponding coefficients were obtained, which were utilized to construct the risk signature (**Figure 3D**; **Table2**). The risk scoring of every KIRC sufferer can be computed as:

$$\begin{aligned}\text{Risk score} = &\left(0.57247064179788*\text{Exp}_{\text{miR-342-3p}}\right)\\ &+\left(-0.130825723882467*\text{Exp}_{\text{miR-1251-5p}}\right)\\ &+\left(0.246195045745634*\text{Exp}_{\text{miR-223-3p}}\right)\\ &+\left(0.588243565862617*\text{Exp}_{\text{miR-1277-3p}}\right)\\ &+\left(0.276954654176782*\text{Exp}_{\text{miR-221-3p}}\right)\\ &+\left(-0.223588405707281*\text{Exp}_{\text{miR-486-5p}}\right)\\ &+\left(0.28936774470365*\text{Exp}_{\text{miR-6718-5p}}\right)\end{aligned}$$

Among the 7 risk miRNAs, 5 miRNAs (*miR-342-3p, miR-223-3p, miR-1277-3p, miR-221-3p, and miR-6718-5p*) with positive coefficients can be deemed as the factors of unsatisfactory prognoses of sufferers, i.e., the increased expression of these miRNAs is associated with the low OS rates of sufferers, whereas 2 miRNAs (*miR-1251-5p* and *miR-486-5p*) with negative coefficients may be used as KIRC inhibitors, i.e., the increased expression of these 2 miRNAs is beneficial to the prognoses of KIRC sufferers, which has been verified in the KM survival curve (**Supplementary Figure S1**). As per the mid-value of the risk scoring of the training cohort, all cohort patients were separated into the $\text{risk}_{\text{high}}$ group and risk_{low} group. K-M survival analyses show that remarkable diversities existed in OS rates between different risk groups (**Figures 3F–H**). Overall, the OS rates of the $\text{risk}_{\text{high}}$ group was lower vs. the risk_{low} group. Similarly, the mortality of KIRC sufferers in different groups were not similar, and the death rate of KIRC sufferers in the $\text{risk}_{\text{high}}$ group was greater (**Figure 3E**). We ranked the risk scoring of

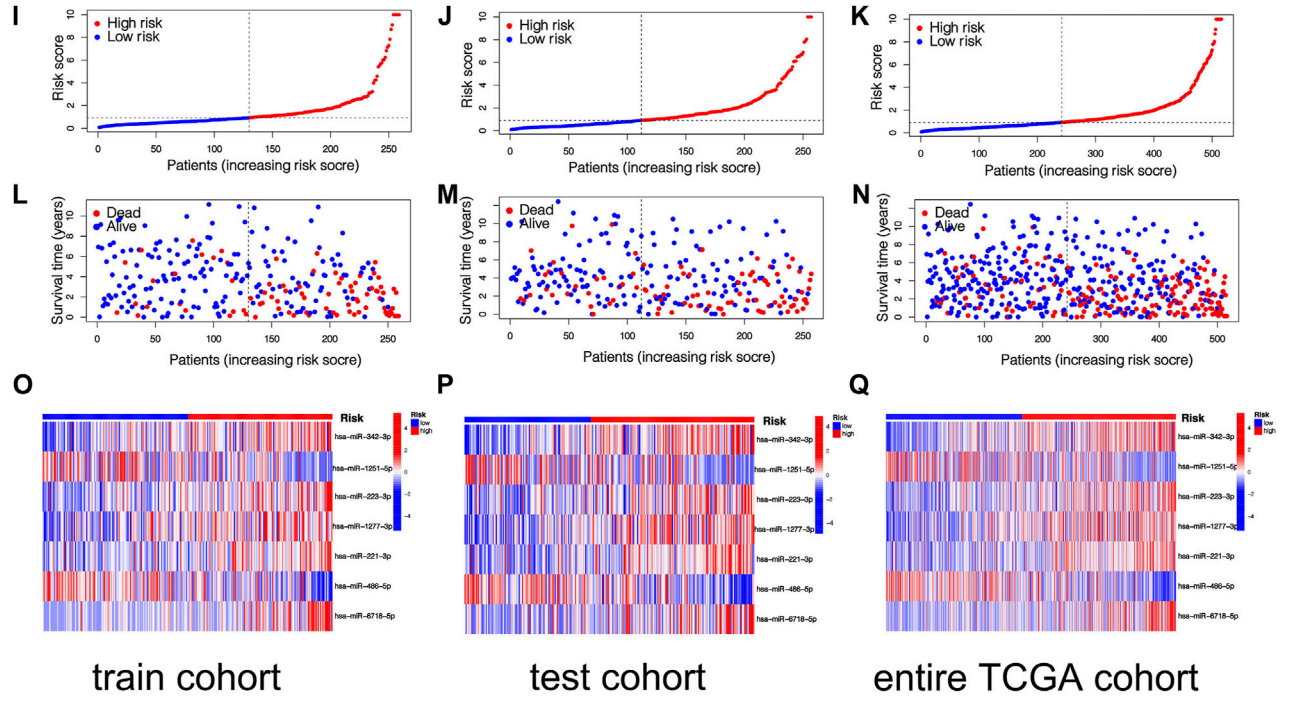


FIGURE 3 | (Continued). Identification of prognostic signatures based on m7G-related miRNAs in KIRC. **(A)** Based on univariate Cox analysis, 18 prognosis-related m7G-related miRNAs were displayed in the forest plot. **(B)** Cross-validation in the LASSO model. **(C)** The LASSO coefficient profile of 13 prognostic m7G-related miRNAs. **(D)** Based on multivariate Cox analysis, 7 risk miRNAs were identified. **(E)** Mortality ratio of patients in high-risk and low-risk groups. **(F–H)** Distribution of risk scores of training cohort, testing cohort and entire TCGA cohort. **(I–K)** Distribution of survival time and survival state of training cohort, testing cohort and entire TCGA cohort. **(L–N)** Heatmap of 7 risk miRNAs expressions of training cohort, testing cohort and entire TCGA cohort. **(O–Q)** K-M survival curves of two risk groups of training cohort, testing cohort and entire TCGA cohort.

TABLE 2 | Seven candidate miRNAs and their corresponding coefficients were obtained by multivariate Cox regression analysis.

Candidate miRNAs	Coefficient
hsa-miR-342-3p	0.57247064179788
hsa-miR-1251-5p	−0.130825723882467
hsa-miR-223-3p	0.246195045745634
hsa-miR-1277-3p	0.588243565862617
hsa-miR-221-3p	0.276954654176782
hsa-miR-486-5p	−0.223588405707281
hsa-miR-6718-5p	0.28936774470365

sufferers, and the survival state and distribution of patient risk scores are shown in **Figures 3I–N**. With the increase in risk scores, the number of deaths of KIRC patients also increases. The heatmap shows the expression of 7 risk miRNAs. It can be seen that the 2 risk miRNAs, *miR-1251-5p* and *miR-486-5p*, are mostly under-expressed in $risk_{high}$ sufferers (**Figures 3O–Q**). ROC curve was used to verify the accurateness of the risk signature in forecasting the prognoses of KIRC sufferers (**Figures 4A–D**). The AUC of 1, 3, and 5 years in the training cohort is 0.750, 0.775 and 0.800 separately, which shows that the model exhibits satisfactory accuracy in predicting patients' survival. The results of principal component analysis (PCA) show that risk scoring has the ability to separate sufferers into the $risk_{high}$ group and $risk_{low}$ group (**Figures 4E–G**). The same verification results were acquired in the test cohort and full TCGA cohort.

Association Between the Risk Signature and Clinical Features

The heatmap shows the clinical characteristics and risk scoring of every KIRC sufferer (**Figure 5**). To further prove the clinical feasibility of the risk signature, we classified the sufferers according to different clinical characteristics. The K-M survival analysis results showed that except for sufferers with N1, survival diversities existed between the $risk_{high}$ and $risk_{low}$ groups in subgroups with different clinical features, and $risk_{low}$ groups all had a better prognosis (**Figures 6A–N**). It showed that the prediction risk signature had wide clinical applicability. We also evaluated whether there were diversities in risk scores among KIRC sufferers with diverse clinical features. We can come to the conclusion that risk scoring was remarkably associated with a variety of clinical features (Grade, TNM staging, Stage) (**Figures 6O–V**). The risk scores of sufferers with G3-G4 were significantly higher than those with G1-G2, and those with T3-T4 were remarkably greater in contrast to those with T1-T2, and those with high pathological grades were remarkably greater in contrast to those with low pathological grades. The association between risk scoring and clinical features corresponds to prognoses, that is, sufferers with

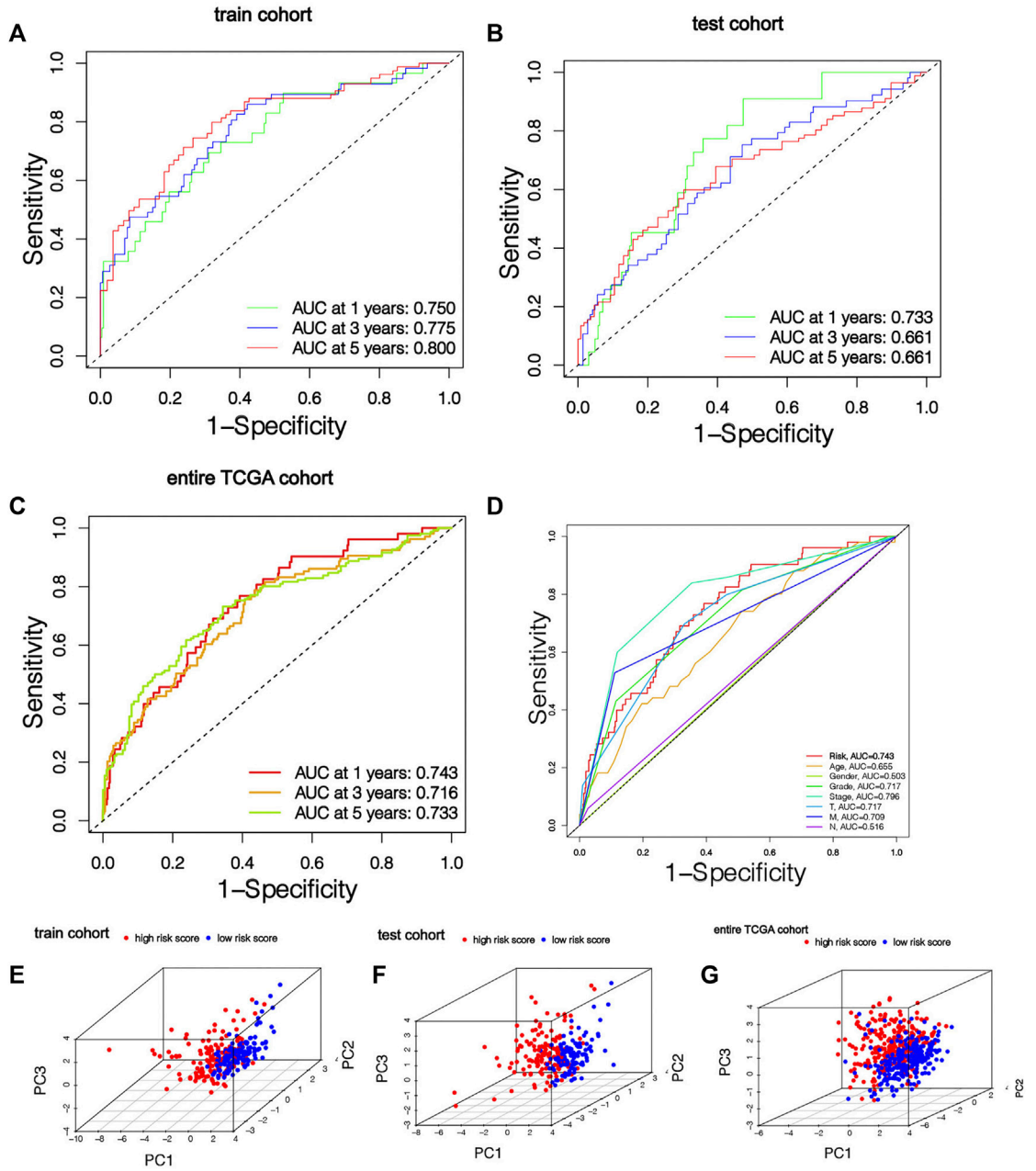


FIGURE 4 | The risk signature has good prediction accuracy. **(A-C)** ROC curves of patients in training cohort, testing cohort and entire TCGA cohort based on risk score for 1, 3 and 5 years. **(D)** ROC curve based on risk score and clinical features in the entire TCGA cohort. **(E–G)** Training cohort, testing cohort and entire TCGA cohort PCA results show that risk score has the ability to divide KIRC patients into two subgroups.

$risk_{high}$ scores tend to have poor clinical characteristics and poor prognoses.

Independence of the m7G-Related MicroRNAs Risk Score Regarding Other Clinical Factors

Univariable-multivariable Cox regressive analyses were used to validate if the risk scoring was an independent predictor of the patient's prognosis, regardless of clinical features. In the training cohort, univariable Cox results revealed that risk scoring was remarkably associated with the prognosis of sufferers ($p < 0.01$), and multivariable Cox regressive results ($p < 0.01$) further evidenced that risk scoring could serve as independent prognosis biomarkers in KIRC patients regardless of clinical characteristics (**Figures 7A,B**). The same univariate-multivariate Cox analysis results were obtained in the test cohort and full TCGA cohort (**Figures 7C–F**). The above

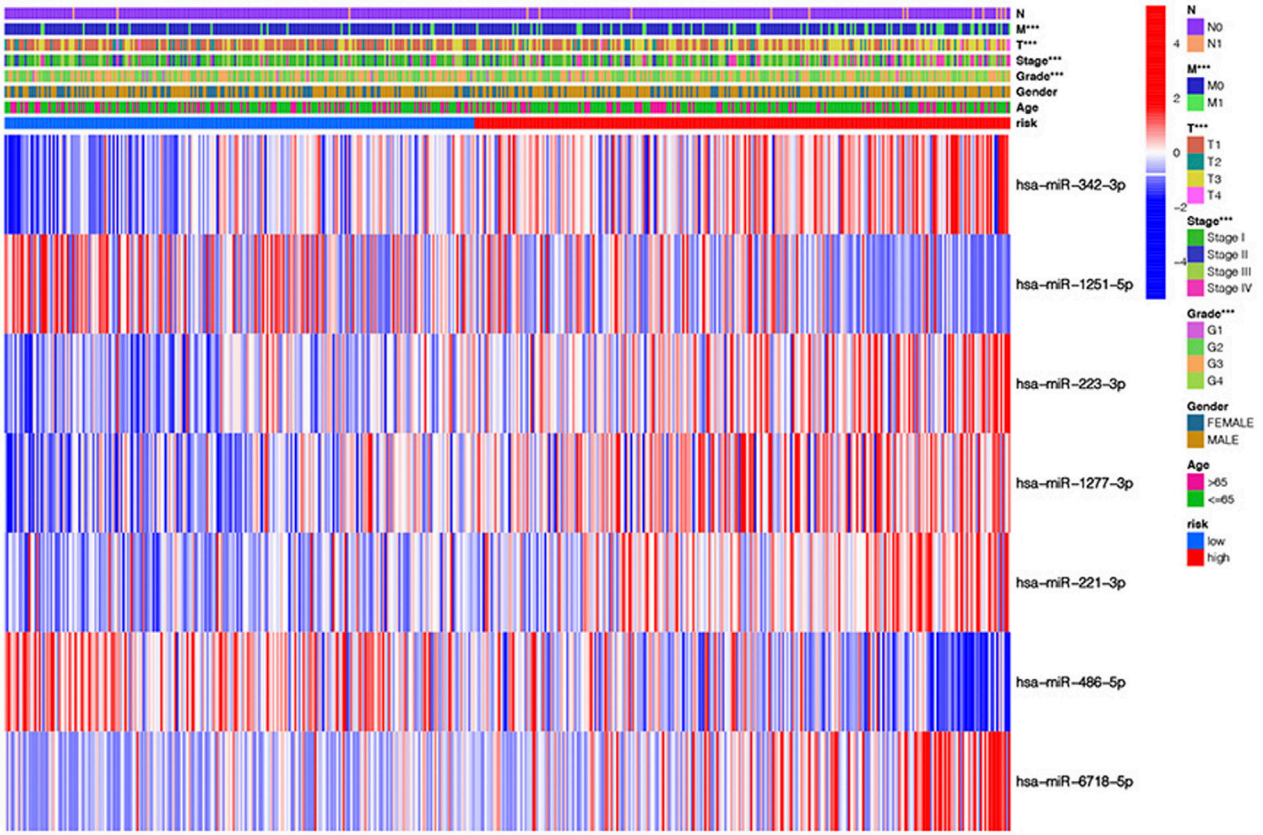


FIGURE 5 | The heatmap shows the clinical characteristics and risk scoring of every KIRC sufferer.

results show that the prognosis-related risk scoring on the foundation of m7G-associated miRNAs has a certain significance for the prognosis evaluation of KIRC patients.

Establishment and Verification of Nomogram

In order to fully exploit our prognostic risk signature clinically, a nomogram was established on the foundation of 3 clinical factors (age, grade, stage) and risk scoring for the quantitative prediction of 1-, 3-, and 5-year survival in KIRC patients (**Figure 8A**). Subsequently, the calibration curve was employed to verify the predictive ability and accurateness of the nomograph model (**Figures 8B–D**). The outcomes showed that nomogram had the ability to accurately estimate the OS of KIRC patients. It also shows that the nomogram has the value and potential to precisely forecast the prognoses of KIRC sufferers clinically.

Tumor Microenvironment Analysis and Immune Infiltration Analysis

The TME score of each KIRC patient was obtained by "ESTIMATE" algorithm. It can be seen that remarkable diversities existed in TME scoring (including stroma scoring, immunoscore and estimated scoring) between the two risk groups (**Figures 9A–C**). The TME score of the $risk_{high}$ group was higher, which suggested that there was a potential connection between risk scoring and the TME of KIRC sufferers. **Figure 9D** shows the infiltration of immunocytes and immunofunction in 516 KIRC samples, and the correlation between immunocytes and immunofunction is shown in **Figures 9E,F**. To evaluate the relationship between the risk scoring and immunity status of KIRC patients, ssGSEA was used to evaluate the diversity in immunocytes scoring and immune function scoring between the $risk_{high}$ group and $risk_{low}$ group. The scoring of $CD8^+$ T cells, Mast cells, Macrophages and other immune cells was different between the $risk_{high}$ group and $risk_{low}$ group (**Figure 9G**). There were also remarkable diversities in the scores of immune-related pathways like Parainflammation, T cell costimulation, CCR, and APC co-suppression between the two risk groups (**Figure 9H**). It can be concluded that most immunocyte scoring and immunofunction scoring are greater in the $risk_{high}$ group, and there are remarkable differences in immune infiltration levels among diverse groups. The expressing levels of common immune checkpoints in the $risk_{high}$ group and $risk_{low}$ group were also calculated, and it can be concluded that the expression of the majority of immune checkpoints, like CD27, CD48, CD44, CD28, CD274 (PD-L1), etc., is greater

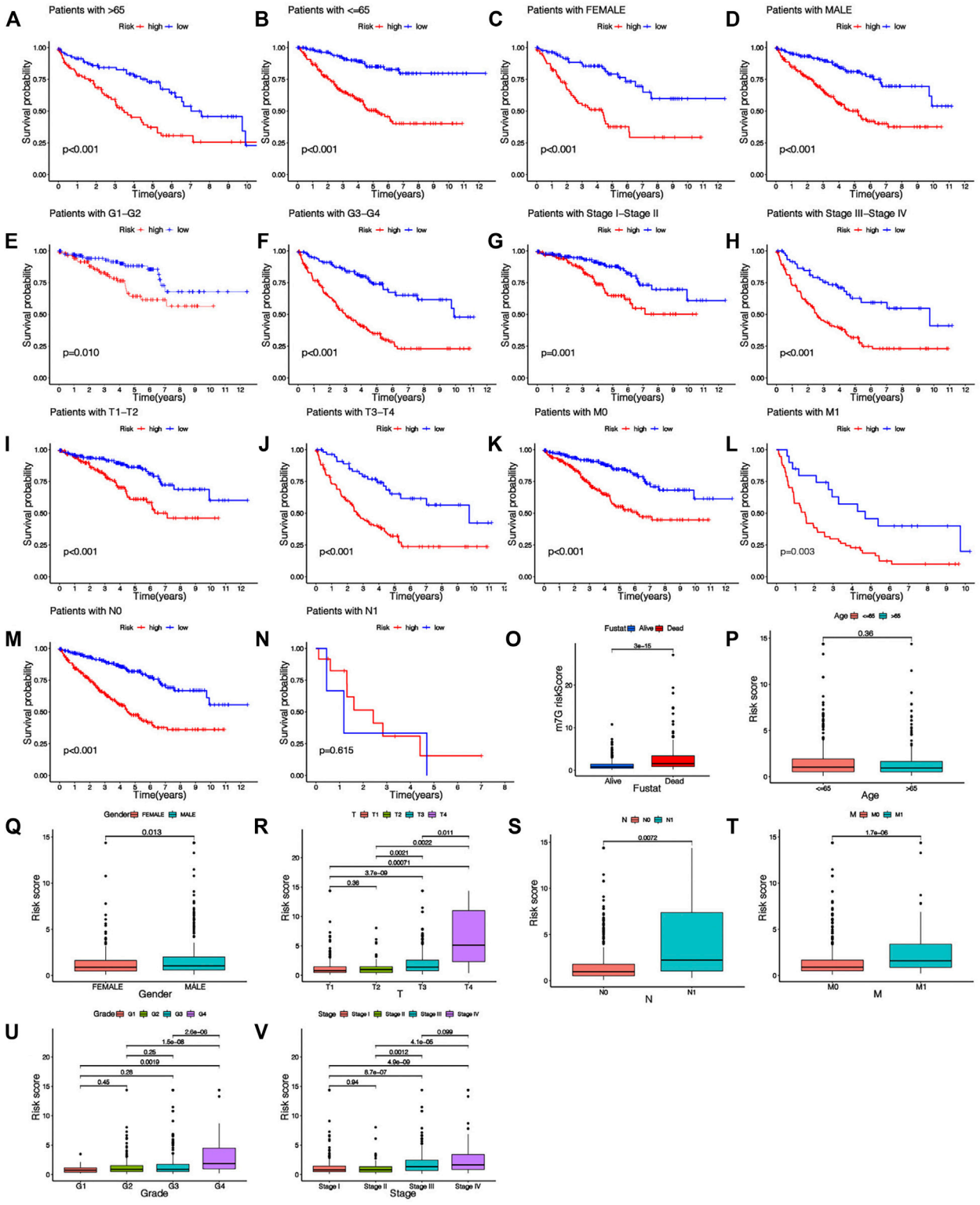


FIGURE 6 | Relationship between risk score and clinical characteristics. **(A–N)** K-M survival curves for patients with different clinical features. **(O–V)** The histogram shows the distribution of risk scores among different clinical subgroups.

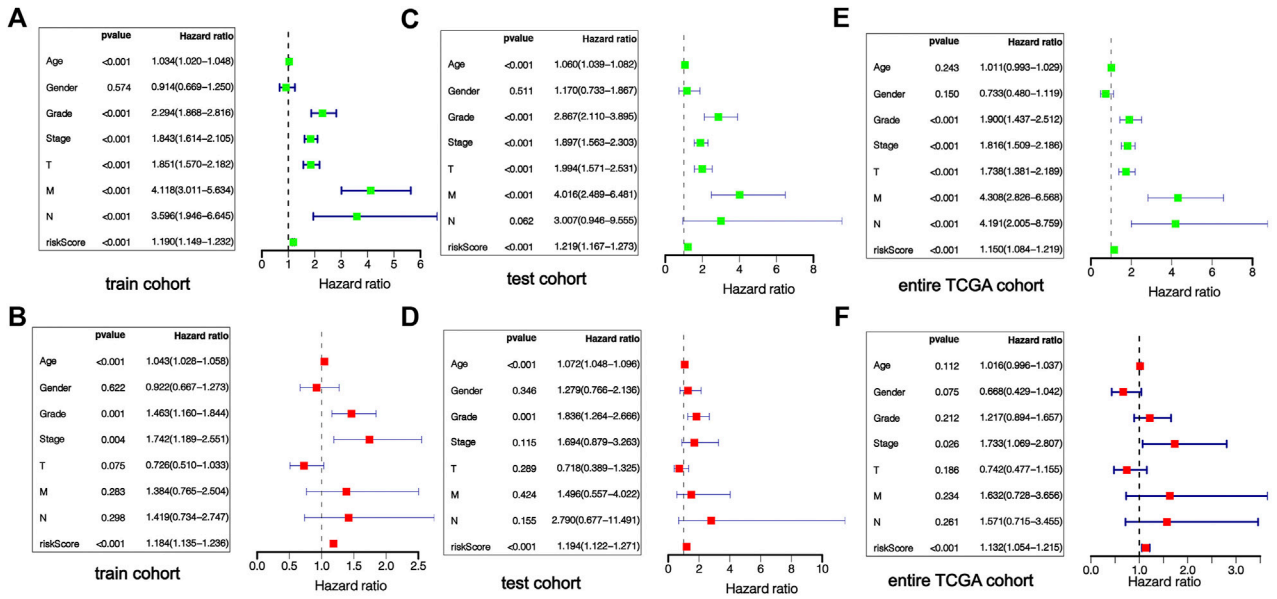


FIGURE 7 | Independent verification of prognostic risk score. The forest map results of univariate and multivariate regression analysis show that risk score is an independent prognostic factor of KIRC patients, training cohort **(A,B)**, testing cohort **(C,D)**, entire TCGA cohort **(E,F)**.

in the $risk_{high}$ group (**Figure 9I**). The results of the above analyses suggest that the immunity status of $risk_{high}$ sufferers may be more active. At the same time, the TIDE algorithm was leveraged to assess the underlying ICB treatment reaction. The TIDE scoring of $risk_{high}$ sufferers was lower vs. $risk_{low}$ sufferers, which suggested that $risk_{low}$ sufferers displayed the potential of immune escape (**Figure 9J**). By comparison with $risk_{low}$ sufferers, the effect of ICB treatment in $risk_{high}$ sufferers may be more significant. It further shows that our risk signature has a great prospect in guiding individualized immunotherapy for KIRC patients.

Tumor Mutation Burden and Prediction of Potential Drug Sensitivity of KIRC

We analyzed the variation of somatic cell mutation in two risk groups. An overall 84.9% $risk_{low}$ sufferers displayed the mutation, whereas only 75.66% $risk_{high}$ sufferers displayed such mutation (**Figures 10A,B**). Although the TMB is different among different risk groups, it is not significant (**Figure 10C**). In $risk_{high}$ sufferers, the most frequently mutated gene is VHL and the most frequent mutation type is a missense mutation, while in $risk_{low}$ sufferers, the most frequently mutated gene is VHL and the most frequent mutation type is Frame-Shift-Ins. As per the outcomes of the K-M analyses, the prognoses of high-TMB sufferers are poorer vs. low-TMB sufferers (**Figure 10D**). In addition, we combined TMB and risk scores in KIRC patients for survival analyses, and conclude that sufferers with high-TMB and $risk_{high}$ scores have the worst prognosis (**Figure 10E**), which further verifies the ability of our risk signature in forecasting the OS of KIRC sufferers. By comparing the IC_{50} values of common drugs in different risk groups, the sensitivity of KIRC patients to potential therapeutic drugs can be predicted. There are significant differences in the IC_{50} values of 14 drugs (MS-275, AUY922, CH5424802, YM201636, CCT018159, CCT007093, NU-7441, AICAR, THZ-2-49, Genentech Cpd 10, BX-912, Ruxolitinib, GSK1904529A, FH535) in the $risk_{high}$ group and $risk_{low}$ group (**Supplementary Figure S2**), and the IC50 value of AMPK activator (AICAR) in $risk_{high}$ sufferers is lower (**Figure 10G**), which suggests that AICAR has a more significant effect in $risk_{high}$ sufferers. The IC_{50} value of JAK1/2 inhibitor (Ruxolitinib) in $risk_{low}$ patients is lower (**Figure 10F**), which indicates that the therapeutic effect of Ruxolitinib is more beneficial in $risk_{low}$ people. The prediction of the efficacy of these potential therapeutic drugs can be helpful for the clinical drug treatment of KIRC. These results also demonstrate that our risk signature displays certain significance in guiding the drug therapy for KIRC sufferers.

Analysis of Seven Risk MicroRNAs in the Predicted Signature Model

As we have come to the conclusion that the risk scoring may be remarkably associated with the immune status and TME of KIRC patients. We will perform further research on the relationship among the expression of Seven risk miRNAs and immune status and TME. **Figure 11A** shows the correlation between Seven risk miRNAs and 22 kinds of immune cells, among which neutrophils and T cells CD8 are closely related to risk miRNAs. The expression of Seven miRNAs and TME score are shown in **Figures 11B–L**. The outcomes reveal that the expressing level of *miR-1231-3p* is related to immunity scoring and estimated

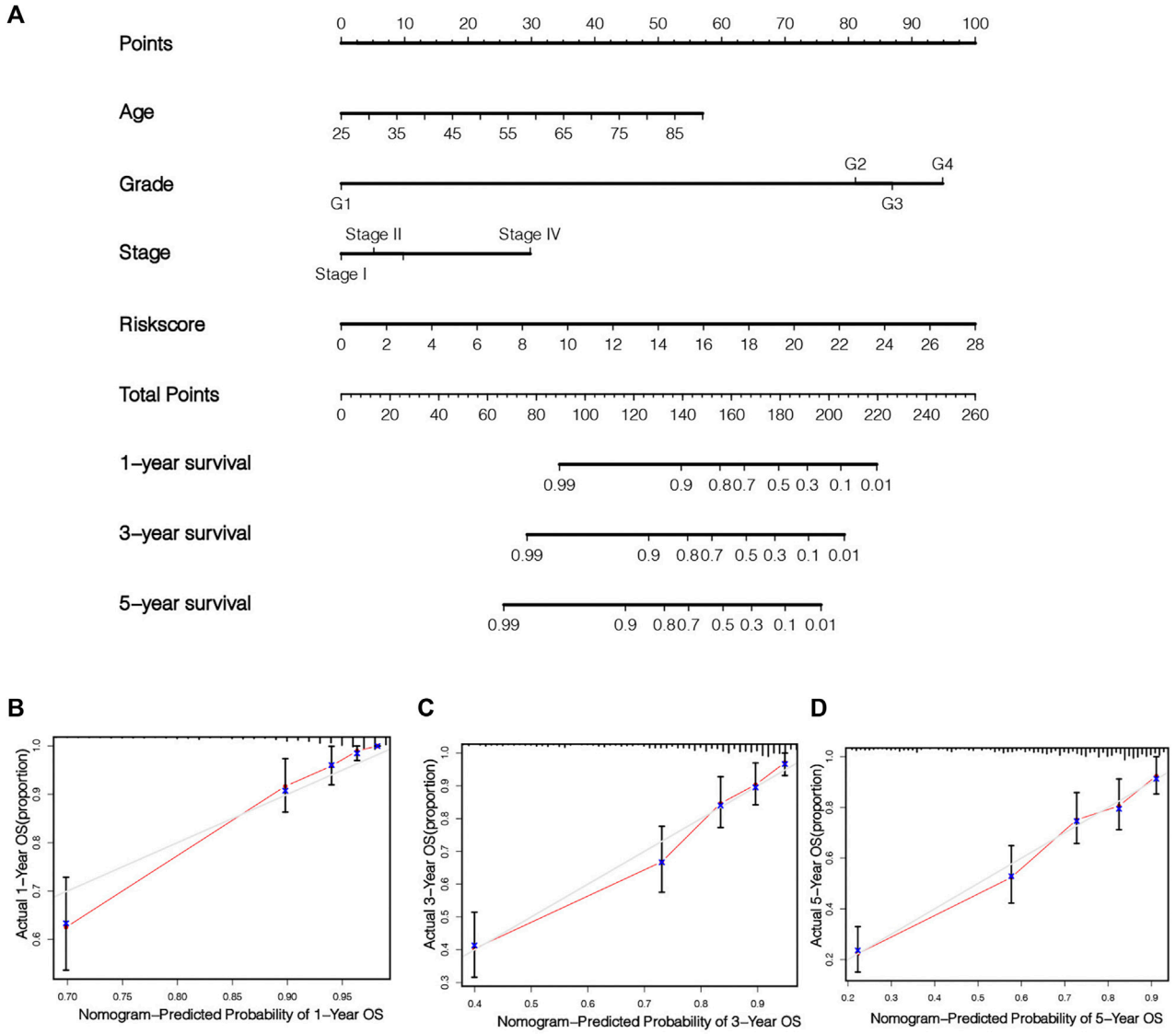


FIGURE 8 | Establishment and verification of nomogram. Combined with age, stage, grade and risk score, a nomogram was established to quantitatively predict the 1, 3 and 5 years survival of KIRC patients **(A)**. The calibration chart illustrates the prediction ability of this nomogram **(B–D)**.

scoring in a negative manner ($p < 0.05$) (**Figures 11B,C**). With the increase of the expressing levels of *miR-1231-3p*, *miR-1227-3p* and *miR-342-3p*, the stromal scoring, immunoscore and estimated scoring of KIRC patients gradually increase ($p < 0.05$) (**Figures 11D–L**). The enrichment analysis of 7 risk miRNAs based on the miRNA function analysis website (http://www.lirmed.com) shows that these risk miRNAs are mainly enriched in relevant functional pathways such as hematopoiesis, epidemiological-to-immunological transition, immune response, angiogenesis, etc. (**Figures 11M,N**). The potential functional pathways by which risk scores affect the occurrence and development of KIRC are expounded from another angle. These results provide certain new supports for the accuracy and reliability of our prognostic risk signature.

Screening Process for Hub Prognostic Genes and the Identification of Prognostic Genes Highly Related to m7G-Related MicroRNA Risk Score and Immunity

To further explore the genes affecting the prognoses of KIRC sufferers at the m7G and immune levels. Based on the risk score, an overall 3,687 DEGs were determined in the $risk_{high}$ group and $risk_{low}$ group in a complete cohort (screening criteria: $|log2FC| > 1$, FDR < 0.05). Based on the immune score of TME scoring, 4,358 immune-related DEGs were determined in the high-immunoscore group and low-immunoscore group (screening criteria: $|log2FC| > 1$, FDR < 0.05). By crossing the two DEGs, a total of 1,548 DEGs related to risk score and immune score were obtained

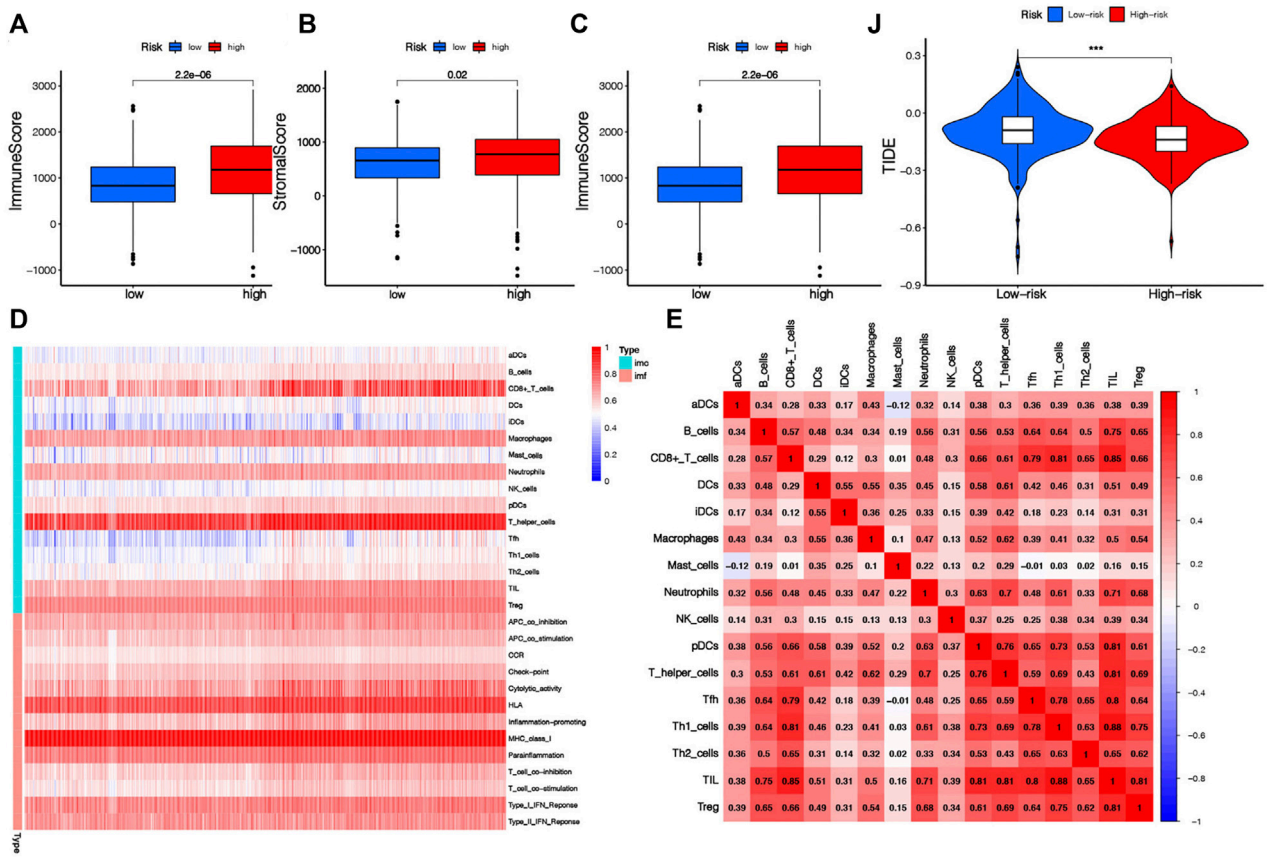


FIGURE 9 | (Continued).

(**Figure 12A**; **Supplementary Table S4**). Afterwards, the prognosis significance of those DEGs was studied *via* univariable Cox regressive analysis, and altogether 202 prognostic risk-immune-related genes were obtained (**Supplementary Table S5**). In order to study the biofunctions and pathways associated with risk scoring and immunity, these 202 prognostic risk-immune-related genes were analyzed by GO enrichment, KEGG pathway and DO analyses (**Figures 13B–D**). The results revealed that the majority of these genes were enriched in immunity-associated pathways and biofunctions. To further select the hub prognostic genes, through the STRING database, 71 genes were selected from 202 prognostic risk-immune-related genes to construct the PPI network (medium confidence >0.4), which contained 71 nodes and 72 edges (**Figure 13A**). **Figure 13B** shows the degree of correlation between genes, and *CXCL8* exhibits the strongest association with other genes. Then, the gene relationships were imported into Cytoscape. According to the cytoHubba plugin in Cytoscape, the top 10 hub genes were screened out, and we studied the hub genes (**Figure 13C**). **Figure 13D** shows the association among these hub prognosis risk-immunity-associated genes and immunocytes and immunofunction. The expression of *MMP13*, *MZB1* in 18 tumor types showed that *MMP13*, *MZB1* was differentially expressed in various tumors, and it can be seen that the expressing level of *MMP13*, *MZB1* in KIRC samples was greater vs. the corresponding normal samples (**Figures 13E,G**). K-M survival curve showed that the expressing level of *MMP13*, *MZB1* was remarkably associated with the OS of KIRC sufferers (**Figures 13F,H**).

DISCUSSION

KIRC is a heterogeneous disease. Even KIRC patients with similar tumor stages and clinical stages could display fairly different tumor progression and prognosis. Although there are many treatment methods (surgery, targeted drug therapy, immunotherapy) that can improve the OS rates of patients with KIRC to a certain extent, the prognosis of patients with advanced KIRC is still not satisfactory. Some of the existing biomarkers of KIRC that have been verified have very limited predictive ability, which is not enough to meet the needs of clinical diagnosis, treatment and the prediction of patient survival. Hence, it is imperative to find a biomarker which

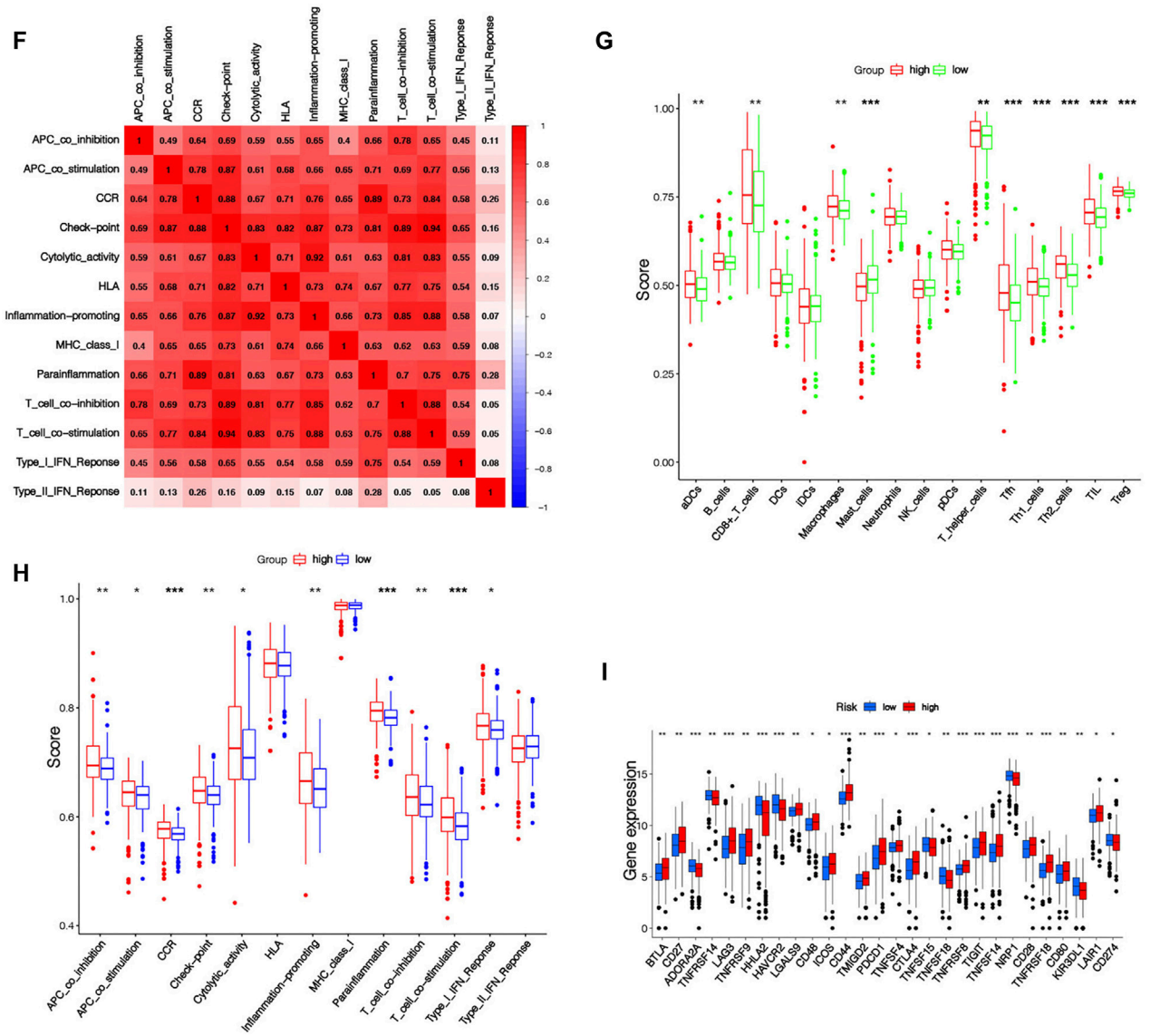


FIGURE 9 | (Continued). Analysis of tumor microenvironment of KIRC patients based on risk score. **(A–C)** Differences in tumor microenvironment scores in different risk groups. **(D)** The heatmap shows the infiltration and distribution of immune cells and immune functions. **(E)** Correlation between immune cells and immune cells. **(F)** Correlation between immune functions and immune functions. Based on ssGSEA algorithm, the differences of 13 immune-related pathway scores **(G)** and 16 immune cell scores **(H)** between $risk_{high}$ and $risk_{low}$ groups. **(I)** Expression levels of genes related to common immune checkpoints in different groups. **(J)** The TIDE score difference between the two risk groups.

can precisely forecast the prognoses of KIRC sufferers and provide guidance and assistance for individualized treatment to avoid under-treatment and over-treatment.

M7G is a modification method of miRNA, which is related to the initiation of miRNA biogenesis, cell migration, tumor immunity and other biological processes. Recent studies and literature reports have shown that the abnormal expression of miRNA can affect the progress of many kinds of tumors, and that it is tightly associated with the prognosis of patients. Currently, there are many signatures based on miRNA that can precisely forecast the OS rates of sufferers, for example, m6A-related miRNAs signature can be utilized as a marker to forecast the prognoses of esophageal carcinoma patients (Li et al., 2021); the signature based on miRNA can accurately predict the survival of HNSCC sufferers (Zhao and Cui, 2020); the risk signature composed of necroptosis-related miRNAs can be utilized as a tool to accurately forecast the prognoses of sufferers with KIRC and colon cancer and guide immunotherapy (Huang et al., 2021; Chen W. et al., 2022); and the signature composed of nine immune-related miRNAs can predict the OS rates of patients with gastric cancer (Xu et al., 2021). *METTL1* and *WDR4* that are related to m7G

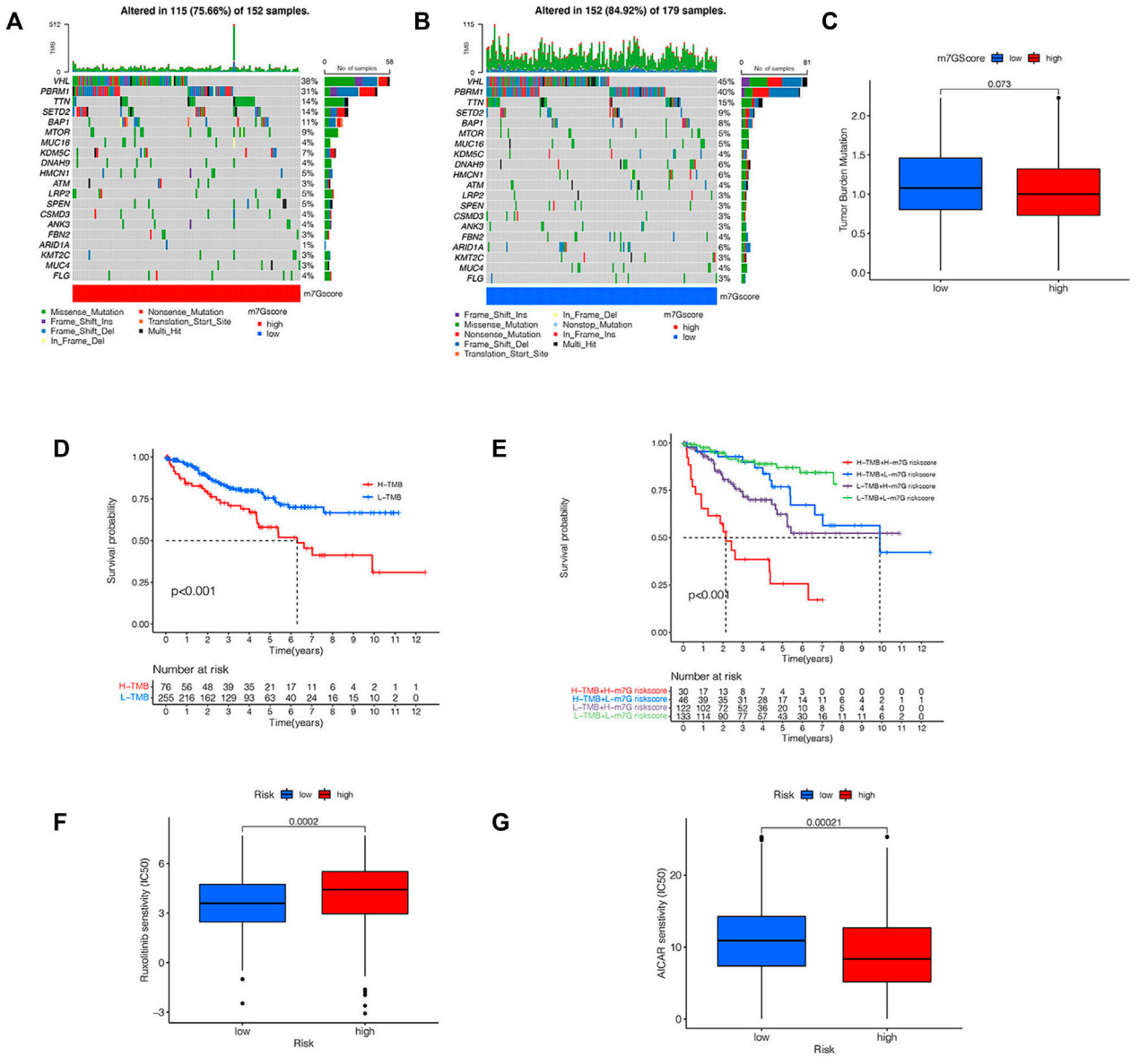


FIGURE 10 | Analysis of TMB and drug sensitivity. **(A,B)** The waterfall plot shows the somatic mutation characteristics of two risk groups. **(C)** Difference of TMB between two risk groups. **(D)** K-M survival analysis of high-TMB group and low-TMB group. **(E)** K-M survival analysis based on TMB and risk score. **(F,G)** IC_{50} values of AICAR and Ruxolitinib in two risk groups.

modification on miRNA have been proved to be tightly associated with the onset of various diseases, and they are differentially expressed in various tumors including KIRC. In addition, they are also remarkably associated with the OS rates of KIRC sufferers. However, whether the m7G-related miRNAs signature can be utilized as a prognosis index for KIRC sufferers has not been elucidated.

Based on m7G-related miRNAs, our team established a prognostic risk signature comprising 7 m7G-associated miRNAs by univariable Cox, LASSO and multivariable Cox analyses. And a risk scoring computation equation was obtained to compute the risk scoring of every KIRC sufferers. In order to further assess the ability of the siganture in forecasting the prognoses of KIRC sufferers, it can be concluded by survival analyses that the OS of $risk_{high}$ sufferers is lower vs. $risk_{low}$ sufferers. Then, the ROC curve proved the accurateness of the signature in forecasting the prognoses of sufferers. Univariable-multivariable Cox analyses revealed that the risk score could be utilized as an independent prognostic index for KIRC patients. All these results were verified in the test cohort and full TCGA cohort. In addition, to facilitate the clinical application of signature risk score, our team constructed a nomogram to forecast the survival of KIRC sufferers in a quantitative manner. These results show that our

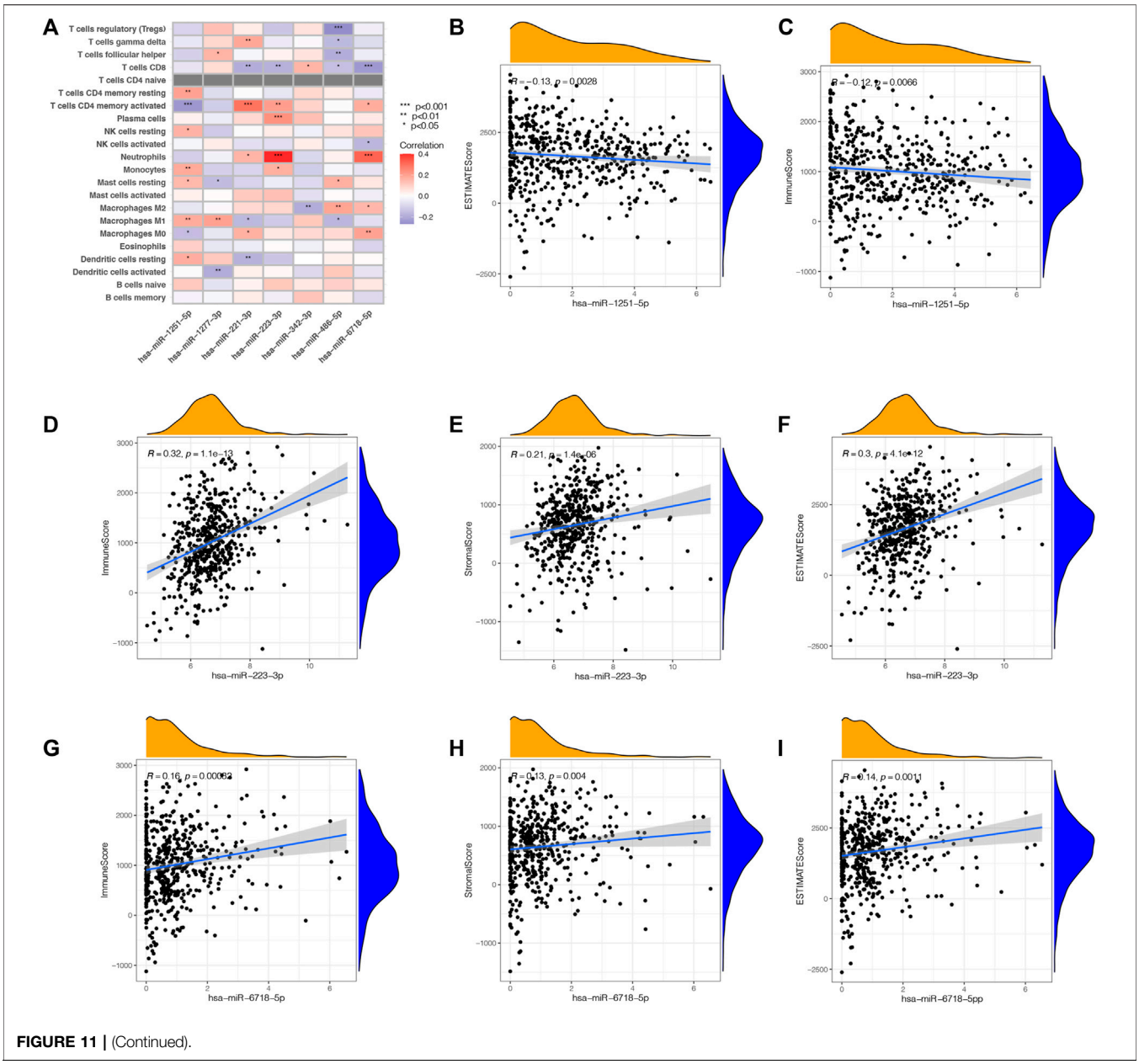


FIGURE 11 | (Continued).

prediction risk signature has great clinical application value in evaluating the prognosis of KIRC sufferers.

Our team also discussed the association among risk scoring and tumor microenvironment, common immune checkpoints. Because of the diversities in TME in diverse groups, these differences can promote the proliferative, migratory and invasive abilities of KIRC, which also explains the significant differences in prognoses amongst diverse groups. It can be seen that most $risk_{high}$ patients have a high level of immunocyte infiltration, and the expressing levels of common immune checkpoint-associated genes in most $risk_{high}$ patients are higher vs. $risk_{low}$ patients. It indicates that the immune activity of $risk_{high}$ sufferers is higher. In addition, by comparison with $risk_{high}$ sufferers, the TIDE scoring of $risk_{low}$ sufferers is lower, which suggests that $risk_{low}$ sufferers have higher immune escape potential. This indicates that $risk_{high}$ sufferers can benefit from ICB therapy. These results reveal that $risk_{high}$ sufferers may benefit more from immune therapy. These immune analysis results can be helpful for guiding individualized immunotherapy for KIRC patients. In addition, we also analyzed the difference in TMB between the $risk_{high}$ and $risk_{low}$ groups, and the TMB of $risk_{low}$ sufferers was higher vs. $risk_{high}$ sufferers. Although this result was not statistically significant, it also explained the relationship between signature based on m7G-related miRNAs and TMB to some extent. We

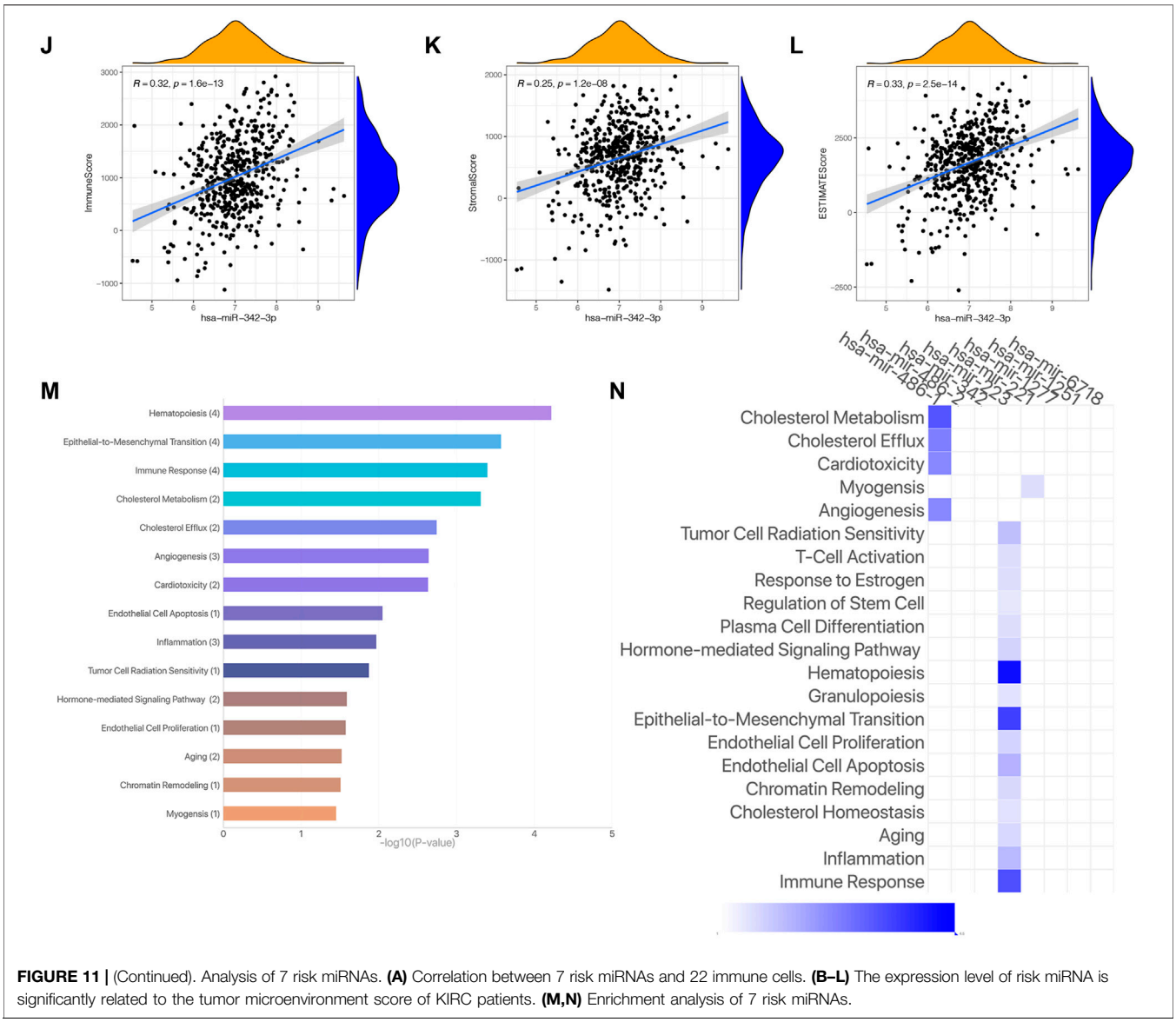


FIGURE 11 | (Continued). Analysis of 7 risk miRNAs. **(A)** Correlation between 7 risk miRNAs and 22 immune cells. **(B–L)** The expression level of risk miRNA is significantly related to the tumor microenvironment score of KIRC patients. **(M,N)** Enrichment analysis of 7 risk miRNAs.

also tried to facilitate the clinical instruction of KIRC drug treatment based on our signature. By comparing the IC_{50} values of common drugs in different risk groups to speculate the potential drug sensitivity, it can be seen that AMPK activator (AICAR) is more suitable for $risk_{high}$ sufferers, while JAK1/2 inhibitor (Ruxolitinib) has lower IC50 values in $risk_{low}$ sufferers, which indicates that the therapeutic effect of Ruxolitinib is more beneficial to $risk_{low}$ sufferers. The prediction of the efficacy of these potential drugs can be helpful for the individualized drug treatment of KIRC.

We further analyzed the influence of the expression of seven risk miRNAs that make up the prognosis signature on the patients with KIRC, and the results show that among these seven risk miRNAs, the expressing levels 5 miRNAs (*miR-342-3p, miR-223-3p, miR-1277-3p, miR-221-3p,* and *miR-6718-5p*) in cancer samples are greater in contrast to healthy samples. These 5 risk miRNAs can be considered risk factors for KIRC patients. The expressing levels of *miR-1251-5p* and *miR-486-5p* are low in cancer samples, which can be regarded as protective factors for patients with KIRC. And the results of K-M survival curve verify our conclusion. There are also some literature reports about the roles of these 7 risk miRNAs in tumorigenesis and development. For example, over-expressed *miR-1251-5p* can inhibit the proliferative, migratory and immunoescape-related abilities of KIRC cells (Yue et al., 2022), and over-expressed *miR-486-5p* can repress the proliferative ability and promote the apoptosis of KIRC cells (He et al., 2019), while *miR-223-3p* can facilitate the proliferative and metastatic abilities of renal clear cell carcinoma by downregulating SLC4A4 (Xiao et al., 2019). These functionally validated results further provide supports for our study: *miR-1251-5p* and *miR-486-5p* can act as protective factors for KIRC; highly expressed *miR-1251-5p* and miR-486-5p can prolong the

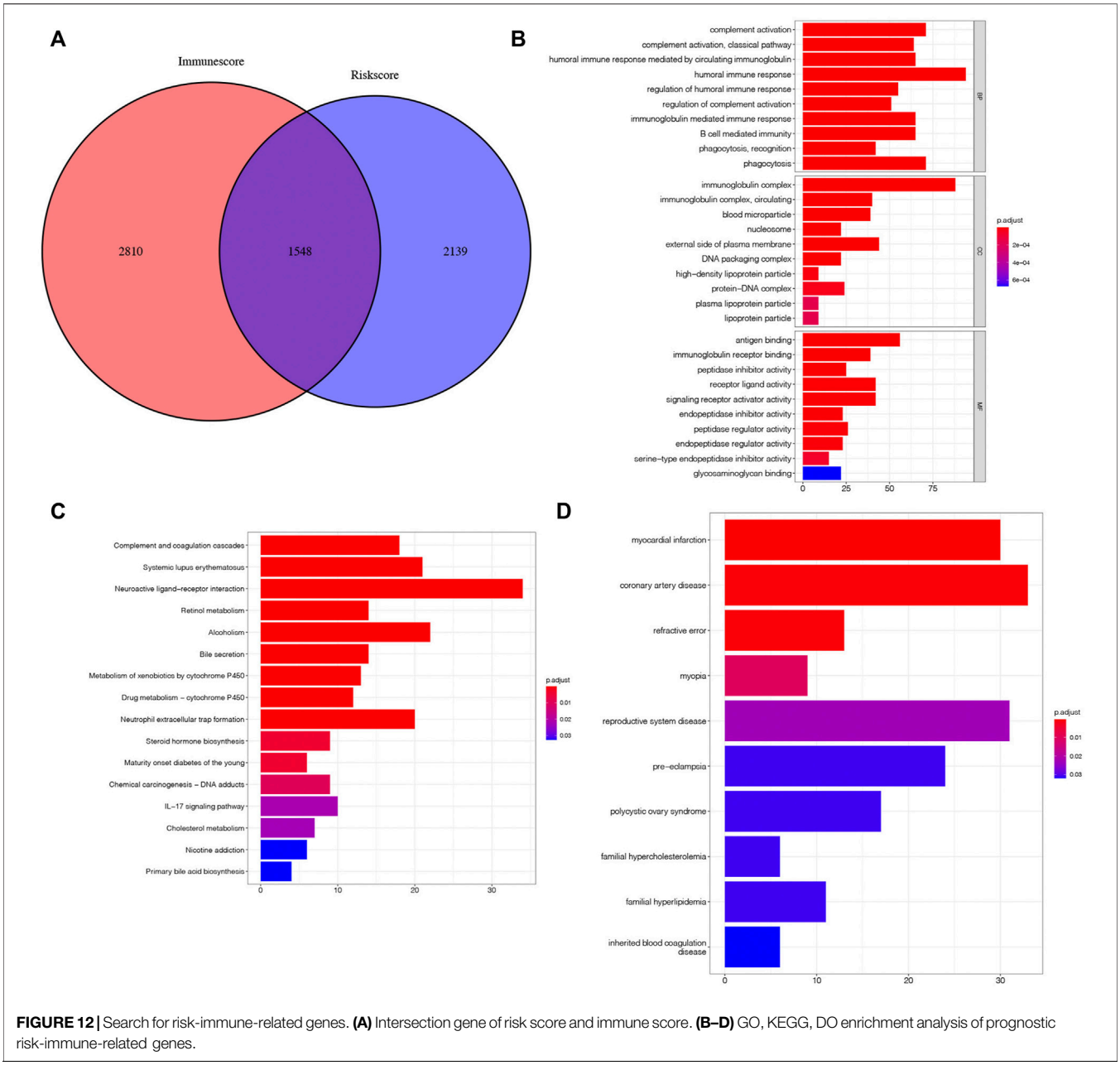


FIGURE 12 | Search for risk-immune-related genes. **(A)** Intersection gene of risk score and immune score. **(B–D)** GO, KEGG, DO enrichment analysis of prognostic risk-immune-related genes.

survival of KIRC patients; *miR-223-3p* is a risk factor for KIRC; and the high expression of *miR-223-3p* can cause poor prognoses in KIRC patients. Although the other 4 risk miRNAs are not reported in KIRC, they are reported in other tumor diseases. For example, *miR-342-3p* can stimulate the malignant potential of NSCLC by regulating the expression of *LASP1* (Shen et al., 2020), and *miR-221-3p* can promote angiogenesis in cervical cancer by regulating the expression of THBS2 (Wu et al., 2019). Most of these seven risk miRNAs have certain influences on the progression of KIRC and other tumors, so the signature based on these seven risk miRNAs is convincing in forecasting the prognoses of sufferers. Meanwhile, we need further *in vivo* and *in vitro* assays to explore the causal link and function of risk miRNA, which are rarely reported in KIRC. With the increasing popularity of targeted therapy in recent years, nowadays, the targeted therapy of KIRC is also developing persistently, which brings us a hint that these seven risk miRNAs may foster the targeted therapy of KIRC.

Finally, based on the predicted signature risk score and the immunoscore, we obtained the risk-immunity-associated genes and analyzed the influence of these genes on the prognoses of KIRC sufferers. Subsequently, our team established a PPI net and screened out 10 hub genes. These hub genes may affect the progress of KIRC through m7G modification on miRNA and immune-related ways. In the future, we can focus on the mechanism and function of

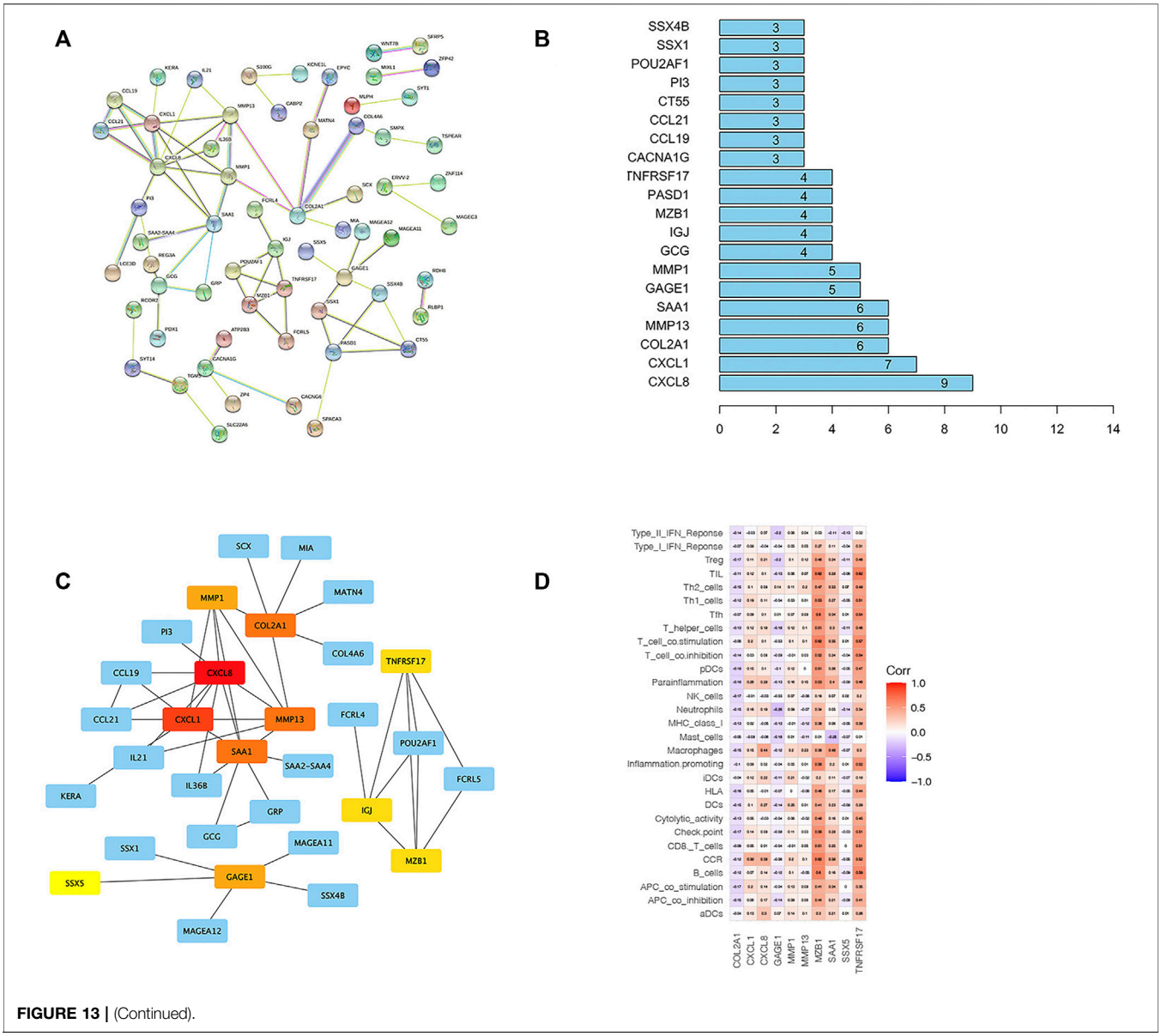


FIGURE 13 | (Continued).

these hub genes affecting the progress of KIRC at the m7G level and immune level, providing new assistance to the early diagnosis and treatment of KIRC.

The aforesaid results show that the risk signature of prognosis prediction based on m7G-related miRNAs has reliable predictability and sensitivity. Comparing our prediction model with the published KIRC prognosis prediction model, we come to the conclusion that our model is superior to other KIRC prognosis prediction models. Generally speaking, the signature based on m7G-related miRNAs can precisely forecast the prognoses of sufferers and can foster the diagnosis and personalized treatment of KIRC. However, our research still has great limitations. First of all, our signature is only verified internally, and no suitable external data set has been found in the published database for the purpose of further evaluating the reliability of our signature. Secondly, we need further functional experiments to verify the relationship between risk miRNAs and m7G modification. Finally, as our research content is retrospective, we need to further verify our conclusion via *in vivo* and *in vitro* assays in the next stage. Although there are some shortcomings, it is the first time that a signature based on m7G-miRNAs can precisely forecast the prognoses of KIRC sufferers, hence, the present research is quite promising in terms of clinical practice.

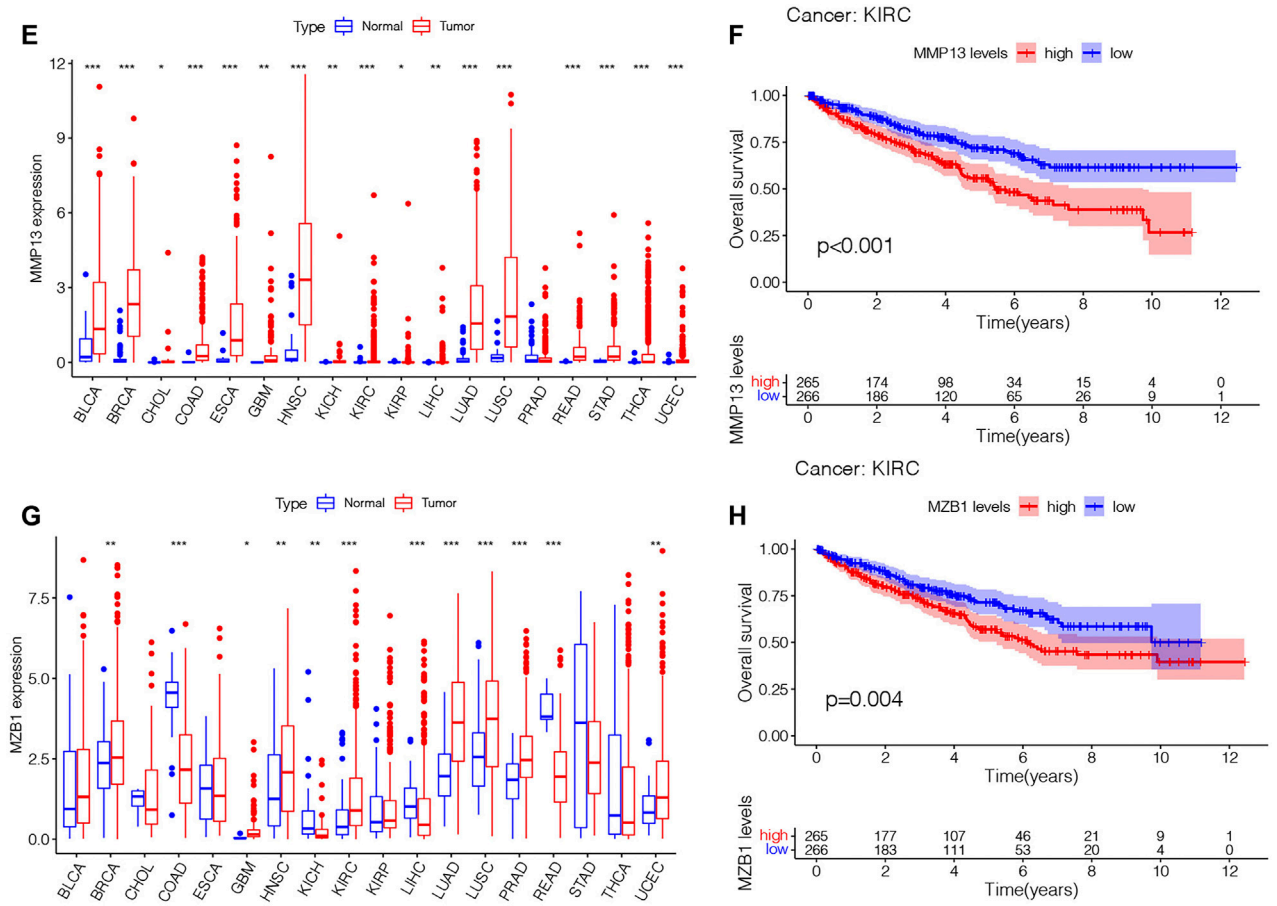


FIGURE 13 | (Continued). Analysis of hub prognostic genes. **(A)** PPI network of prognostic risk-immune-related genes. **(B,C)** Top ten hub genes, the redder the color, the more important the gene is, while the yellower the color, the less important the gene is. **(D)** The correlation between these hub prognostic risk-immune-related genes and immune cells, immune function. **(E,G)** Pan-cancer analysis of hub gene *MMP13*, *MZB1*. **(F,H)** K-M survival analysis based on the expresssion of *MMP13*, *MZB1* in KIRC patients.

CONCLUSION

Our team smoothly established a prognosis-related risk signature on the foundation of m7G-associated miRNAs and verified its accuracy and reliability. A nomogram that could forecast the OS of KIRC sufferers in a quantitative manner was also established. The influence of risk score on the immune microenvironment and immunocyte infiltration of KIRC sufferers was also discussed. Our research results can accurately evaluate the prognosis of KIRC patients and facilitate individualized immunotherapy for KIRC patients. In addition, the identified 7 risk miRNAs are expected to be therapeutic targets for KIRC patients.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

PH, MT, QC, and HD provided advice for artwork layout and manuscript writing. PH and HD contributed to the data analysis and wrote the manuscript. ZL, WH, YJ, GT, and DH provided comments during the writing. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

The help of MT, QC, HD, ZL, WH, YJ, GT, and DH to this study. We acknowledge TCGA database for providing their platforms and contributors for uploading their meaningful datasets.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.922358/full#supplementary-material

REFERENCES

Alexandrov, A., Martzen, M. R., and Phizicky, E. M. (2002). Two Proteins that Form a Complex Are Required for 7-methylguanosine Modification of Yeast tRNA. *Rna* 8, 1253–1266. doi:10.1017/s1355838202024019

Aragon-Ching, J. B., Nizam, A., and Henson, D. E. (2019). Carcinomas of the Renal Pelvis, Ureters, and Urinary Bladder Share a Carcinogenic Field as Revealed in Epidemiological Analysis of Tumor Registry Data. *Clin. Genitourin. Cancer* 17, 436–442. doi:10.1016/j.clgc.2019.07.011

Baradaran Ghavami, S., Chaleshi, V., Derakhshani, S., Aimzadeh, P., Asadzadeh-Aghdaie, H., and Zali, M. R. (2017). Association between TNF-α Rs1799964 and RAF1 Rs1051208 MicroRNA Binding Site SNP and Gastric Cancer Susceptibility in an Iranian Population. *Gastroenterol. Hepatol. Bed Bench* 10, 214–219.

Bartel, D. P. (2009). MicroRNAs: Target Recognition and Regulatory Functions. *Cell* 136, 215–233. doi:10.1016/j.cell.2009.01.002

Boccaletto, P., Machnicka, M. A., Purta, E., Piątkowski, P., Bagiński, B., Wirecki, T. K., et al. (2018). MODOMICS: a Database of RNA Modification Pathways. 2017 Update. *Nucleic Acids Res.* 46, D303–d307. doi:10.1093/nar/gkx1030

Campbell, S., Uzzo, R. G., Allaf, M. E., Bass, E. B., Cadeddu, J. A., Chang, A., et al. (2017). Renal Mass and Localized Renal Cancer: AUA Guideline. *J. Urology* 198, 520–529. doi:10.1016/j.juro.2017.04.100

Chen, J., Li, K., Chen, J., Wang, X., Ling, R., Cheng, M., et al. (2022a). Aberrant Translation Regulated by METTL1/WDR4-mediated tRNA N7-methylguanosine Modification Drives Head and Neck Squamous Cell Carcinoma Progression. *Cancer Commun.* 42, 223–244. doi:10.1002/cac2.12273

Chen, L., Zhang, Y.-H., Wang, S., Zhang, Y., Huang, T., and Cai, Y.-D. (2017). Prediction and Analysis of Essential Genes Using the Enrichments of Gene Ontology and KEGG Pathways. *PLoS One* 12, e0184129. doi:10.1371/journal.pone.0184129

Chen, W., Lin, W., Wu, L., Xu, A., Liu, C., and Huang, P. (2022b). A Novel Prognostic Predictor of Immune Microenvironment and Therapeutic Response in Kidney Renal Clear Cell Carcinoma Based on Necroptosis-Related Gene Signature. *Int. J. Med. Sci.* 19, 377–392. doi:10.7150/ijms.69060

Creighton, C. J., Morgan, M., Gunaratne, P. H., Wheeler, D. A., Gibbs, R. A., Gordon Robertson, A., et al. (2013). Comprehensive Molecular Characterization of Clear Cell Renal Cell Carcinoma. *Nature* 499, 43–49. doi:10.1038/nature12222

Dai, Z., Liu, H., Liao, J., Huang, C., Ren, X., Zhu, W., et al. (2021). N7-Methylguanosine tRNA Modification Enhances Oncogenic mRNA Translation and Promotes Intrahepatic Cholangiocarcinoma Progression. *Mol. Cell* 81, 3339–3355. e8. doi:10.1016/j.molcel.2021.07.003

He, Y., Liu, J., Wang, Y., Zhu, X., Fan, Z., Li, C., et al. (2019). Role of miR-486-5p in Regulating Renal Cell Carcinoma Cell Proliferation and Apoptosis via TGF-β-Activated Kinase 1. *J Cell. Biochem.* 120, 2954–2963. doi:10.1002/jcb.26900

Hori, H. (2014). Methylated Nucleosides in tRNA and tRNA Methyltransferases. *Front. Genet.* 5, 144. doi:10.3389/fgene.2014.00144

Hsieh, J. J., Purdue, M. P., Signoretti, S., Swanton, C., Albiges, L., Schmidinger, M., et al. (2017). Renal Cell Carcinoma. *Nat. Rev. Dis. Prim.* 3, 17009. doi:10.1038/nrdp.2017.9

Huang, Y., Zou, Y., Xiong, Q., Zhang, C., Sayagués, J. M., Shelat, V. G., et al. (2021). Development of a Novel Necroptosis-Associated miRNA Risk Signature to Evaluate the Prognosis of Colon Cancer Patients. *Ann. Transl. Med.* 9, 1800. doi:10.21037/atm-21-6576

Humphrey, P. A., Moch, H., Cubilla, A. L., Ulbright, T. M., and Reuter, V. E. (2016). The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part B: Prostate and Bladder Tumours. *Eur. Urol.* 70, 106–119. doi:10.1016/j.eururo.2016.02.028

Inamoto, T., Uehara, H., Akao, Y., Ibuki, N., Komura, K., Takahara, K., et al. (2018). A Panel of MicroRNA Signature as a Tool for Predicting Survival of Patients with Urothelial Carcinoma of the Bladder. *Dis. Markers* 2018, 5468672. doi:10.1155/2018/5468672

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T Cell Dysfunction and Exclusion Predict Cancer Immunotherapy Response. *Nat. Med.* 24, 1550–1558. doi:10.1038/s41591-018-0136-1

Jung, M., Schaefer, A., Steiner, I., Kempkensteffen, C., Stephan, C., Erbersdobler, A., et al. (2010). Robust microRNA Stability in Degraded RNA Preparations from Human Tissue and Cell Samples. *Clin. Chem.* 56, 998–1006. doi:10.1373/clinchem.2009.141580

Li, L., Xie, R., and Wei, Q. (2021). Network Analysis of miRNA Targeting m6A-Related Genes in Patients with Esophageal Cancer. *PeerJ* 9, e11893. doi:10.7717/peerj.11893

Lin, S., Liu, Q., Lelyveld, V. S., Choe, J., Szostak, J. W., and Gregory, R. I. (2018). Mettl1/Wdr4-Mediated m7G tRNA Methylome Is Required for Normal mRNA Translation and Embryonic Stem Cell Self-Renewal and Differentiation. *Mol. Cell* 71, 244–255. e5. doi:10.1016/j.molcel.2018.06.001

Liu, S., Wang, Z., Zhu, R., Wang, F., Cheng, Y., and Liu, Y. (2021). Three Differential Expression Analysis Methods for RNA Sequencing: Limma, EdgeR, DESeq2. *JoVE* 1, 1. doi:10.3791/62528

Liu, Y., Zhang, Y., Chi, Q., Wang, Z., and Sun, B. (2020). RETRACTED: Methyltransferase-like 1 (METTL1) served as a tumor suppressor in colon cancer by activating 7-methyguanosine (m7G) regulated let-7e miRNA/HMGA2 axis. *Life Sci.* 249, 117480. doi:10.1016/j.lfs.2020.117480

Maclennan, S., Imamura, M., Lapitan, M. C., Omar, M. I., Lam, T. B. L., Hilvano-Cabungcal, A. M., et al. (2012). Systematic Review of Oncological Outcomes Following Surgical Management of Localised Renal Cancer. *Eur. Urol.* 61, 972–993. doi:10.1016/j.eururo.2012.02.039

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust Enumeration of Cell Subsets from Tissue Expression Profiles. *Nat. Methods* 12, 453–457. doi:10.1038/nmeth.3337

Pandey, J., and Syed, W. (2022). *Renal Cancer*. StatPearls. Treasure Island (FL: StatPearls Publishing Copyright © 2022StatPearls Publishing LLC.

Pandolfini, L., Barbieri, I., Bannister, A. J., Hendrick, A., Andrews, B., Webster, N., et al. (2019). METTL1 Promotes Let-7 MicroRNA Processing via m7G Methylation. *Mol. Cell* 74, 1278–1290. e9. doi:10.1016/j.molcel.2019.03.040

Richter, I., and Dvořák, J. (2018). [Treatment of Metastatic Renal Cell Carcinoma]. *Klin. Onkol.* 31, 110–116. doi:10.14735/amko2018110

Rouprêt, M., Babjuk, M., Burger, M., Capoun, O., Cohen, D., Compérat, E. M., et al. (2021). European Association of Urology Guidelines on Upper Urinary Tract Urothelial Carcinoma: 2020 Update. *Eur. Urol.* 79, 62–79. doi:10.1016/j.eururo.2020.05.042

Sharma, N., and Baruah, M. M. (2019). The microRNA Signatures: Aberrantly Expressed miRNAs in Prostate Cancer. *Clin. Transl. Oncol.* 21, 126–144. doi:10.1007/s12094-018-1910-8

Shen, Q., Sun, Y., and Xu, S. (2020). LINC01503/miR-342-3p Facilitates Malignancy in Non-small-cell Lung Cancer Cells via Regulating LASP1. *Respir. Res.* 21, 235. doi:10.1186/s12931-020-01464-3

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2022). Cancer Statistics, 2022. *CA A Cancer J. Clin.* 72, 7–33. doi:10.3322/caac.21708

Speed, J. M., Trinh, Q.-D., Choueiri, T. K., and Sun, M. (2017). Recurrence in Localized Renal Cell Carcinoma: a Systematic Review of Contemporary Data. *Curr. Urol. Rep.* 18, 15. doi:10.1007/s11934-017-0661-3

Tibshirani, R. (1997). The Lasso Method for Variable Selection in the Cox Model. *Stat. Med.* 16, 385–395. doi:10.1002/(sici)1097-0258(19970228)16:4<385:aid-sim380>3.0.co;2-3

Tomikawa, C. (2018). 7-Methylguanosine Modifications in Transfer RNA (tRNA). *Int. J. Mol. Sci.* 19. doi:10.3390/ijms19124080

Wu, X.-G., Zhou, C.-F., Zhang, Y.-M., Yan, R.-M., Wei, W.-F., Chen, X.-J., et al. (2019). Cancer-derived Exosomal miR-221-3p Promotes Angiogenesis by Targeting THBS2 in Cervical Squamous Cell Carcinoma. *Angiogenesis* 22, 397–410. doi:10.1007/s10456-019-09665-1

Xiao, W., Wang, X., Wang, T., and Xing, J. (2019). MiR-223-3p Promotes Cell Proliferation and Metastasis by Downregulating SLC4A4 in Clear Cell Renal Cell Carcinoma. *Aging* 11, 615–633. doi:10.18632/aging.101763

Xu, J., Wen, J., Li, S., Shen, X., You, T., Huang, Y., et al. (2021). Immune-Related Nine-MicroRNA Signature for Predicting the Prognosis of Gastric Cancer. *Front. Genet.* 12, 690598. doi:10.3389/fgene.2021.690598

Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring Tumour Purity and Stromal and Immune Cell Admixture from Expression Data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612

Yue, L., Lin, H., Yuan, S., Wu, L., Chen, G., Wang, J., et al. (2022). miR-1251-5p Overexpression Inhibits Proliferation, Migration, and Immune Escape in Clear Cell Renal Cell Carcinoma by Targeting NPTX2. *J. Oncol.* 2022, 3058588. doi:10.1155/2022/3058588

Zhang, M., Song, J., Yuan, W., Zhang, W., and Sun, Z. (2021). Roles of RNA Methylation on Tumor Immunity and Clinical Implications. *Front. Immunol.* 12, 641507. doi:10.3389/fimmu.2021.641507

Zhao, X., and Cui, L. (2020). A Robust Six-miRNA Prognostic Signature for Head and Neck Squamous Cell Carcinoma. *J. Cell. Physiology* 235, 8799–8811. doi:10.1002/jcp.29723

Zhou, J., Xiong, W., Wang, Y., and Guan, J. (2021). Protein Function Prediction Based on PPI Networks: Network Reconstruction vs Edge Enrichment. *Front. Genet.* 12, 758131. doi:10.3389/fgene.2021.758131

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.

Copyright © 2022 Hong, Du, Tong, Cao, Hu, Ma, Jin, Li, Huang and Tong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 29 June 2022
doi: 10.3389/fgene.2022.916973

Multi-Omics Analysis of Molecular Characteristics and Carcinogenic Effect of NFE2L3 in Pan-Cancer

Mengxiao Liu[1,2,3†], *Hui Wei*[1,2,3†], *Jing Yang*[1,2,3†], *Xia Chen*[1,2,3], *Haoying Wang*[1,2,3], *Ya Zheng*[2,3], *Yuping Wang*[2,3*] *and Yongning Zhou*[2,3*]

[1]The First Clinical Medical College, Lanzhou University, Lanzhou, China, [2]Department of Gastroenterology, The First Hospital of Lanzhou University, Lanzhou, China, [3]Key Laboratory for Gastrointestinal Diseases of Gansu Province, The First Hospital of Lanzhou University, Lanzhou, China

NFE2L3, also known as NFE2L3, is a nuclear transcription factor associated with the pathogenesis and progression of human tumors. To systematically and comprehensively investigate the role of NFE2L3 in tumors, a pan-cancer analysis was performed using multi-omics data, including gene expression analysis, diagnostic and prognostic analysis, epigenetic methylation analysis, gene alteration analysis, immune feature analysis, functional enrichment analysis, and tumor cell functional status analysis. Furthermore, the molecular mechanism of NFE2L3 in liver hepatocellular carcinoma (LIHC) was explored. The relationship between NFE2L3 expression and survival prognosis of patients with LIHC was analyzed and a nomogram prediction model was constructed. Our study showed that NFE2L3 expression was upregulated in most cancers, suggesting that NFE2L3 may play an important role in promoting cancer progression. NFE2L3 expression is closely related to DNA methylation, genetic alteration, immune signature, and tumor cell functional status in pan-cancers. Furthermore, NFE2L3 was demonstrated to be an independent risk factor for LIHC, and the nomogram model based on NFE2L3 expression had good prediction efficiency for the overall survival of patients with LIHC. In summary, our study indicated that NFE2L3 may be an important molecular biomarker for the diagnosis and prognosis of pan-cancer. NFE2L3 is expected to be a potential molecular target for the treatment of tumors.

Keywords: NFE2L3, pan-cancer, multi-comics, tumor immunity, DNA methylation

OPEN ACCESS

Edited by:
Ehsan Nazemalhosseini-Mojarad, Shahid Beheshti University of Medical Sciences, Iran

Reviewed by:
Nayeralsadat Fatemi, Shahid Beheshti University of Medical Sciences, Iran
Lili Zhang, Peking University, China
Kazim Yalcin Arga, Marmara University, Turkey

***Correspondence:**
Yuping Wang
wangyuping@lzu.edu.cn
Yongning Zhou
zhouyn@lzu.edu.cn

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: *10 April 2022*
Accepted: *01 June 2022*
Published: *29 June 2022*

Citation:
Liu M, Wei H, Yang J, Chen X, Wang H, Zheng Y, Wang Y and Zhou Y (2022) Multi-Omics Analysis of Molecular Characteristics and Carcinogenic Effect of NFE2L3 in Pan-Cancer. Front. Genet. 13:916973. doi: 10.3389/fgene.2022.916973

INTRODUCTION

Globally, cancer is one of the most severe diseases threatening human health and poses a heavy health and economic burden to society as a result of its high morbidity and mortality (Sung et al., 2021). Most malignant tumors are in the intermediate and advanced stages when clinically diagnosed, which is often too late to treat the patients effectively. Although significant progress has been made in cancer therapy in recent years, the prognosis of most cancer patients remains poor due to several contributing factors, including drug resistance, side effects of drugs, and chemosensitivity (Liu et al., 2020a). Therefore, it is necessary to develop new biomarkers and therapeutic targets for early cancer screening and treatment.

NFE2L3 is a member of the Cap"n"Collar (CNC) family, which belongs to the basic region-leucine zipper (bZIP) transcription factor (Chevillard and Blank, 2011). The NFE2L3 gene is located

on the human chromosomes 7p15-p14 and encodes a 694-amino-acid protein (Kobayashi et al., 1999). Glycosylation is an important modification form of NFE2L3 (Nouhi et al., 2007). Previous studies have demonstrated that NFE2L3 plays an important role in inflammation and is identified as a stemness marker gene as it is upregulated in the early stage of stem cell differentiation (Witschi et al., 1989; Braun et al., 2002; Byrne et al., 2007; Ben-Yehudah et al., 2010). Studies have also shown that NFE2L3 is closely related to the development and progression of tumors (Saliba et al., 2022). Knockdown of NFE2L3 showed a tumor-suppressive effect in liver and gastric cancer cells (Yu et al., 2019; Wang et al., 2021). NFE2L3 has also been identified as a novel DNA methylation driver gene and prognostic marker of kidney renal clear cell carcinoma (KIRC) (Wang et al., 2019). Another study showed that the β-catenin/Tcf4 complex promoted the proliferation of colon cancer cells and upregulated GLUT1 expression by upregulating the expression of NFE2L3 mRNA (Aono et al., 2019). However, there are limited studies on NFE2L3 in different cancer types, thus it is necessary to comprehensively analyze NFE2L3 expression in pan-cancers. This will aid in exploring effective prognostic biomarkers and provide future reference data for further investigations.

Our study comprehensively analyzed NFE2L3 expression in pan-cancers using different databases. The diagnostic and prognostic value of NFE2L3 in pan-cancers was also evaluated. The correlations between NFE2L3 expression and DNA methylation, genetic alterations, immune features, and tumor cell functional status in pan-cancers was also investigated. A gene function enrichment analysis of NFE2L3 in pan-cancer was performed and the molecular mechanism of NFE2L3 in LIHC was evaluated. NFE2L3 was identified as an independent risk factor for OS in LIHC. In conclusion, the study revealed the essential biological function of NFE2L3 in pan-cancers and showed that NFE2L3 has the potential to be used as an effective therapeutic target and tumor diagnostic marker.

MATERIALS AND METHODS

Gene Expression Analysis in Pan-Cancer

Differential expression of NFE2L3 in normal human tissues was analyzed using the Human Protein Atlas (HPA, v21.0) database (https://www.proteinatlas.org/) (Thul et al., 2017). RNA-seq data and relevant clinical information related to the differential expression of NFE2L3 in various tumor cell lines were obtained from the Cancer Cell Line Encyclopedia (CCLE) database (https://sites.broadinstitute.org/ccle) (Barretina et al., 2012). The differential expression of NFE2L3 between tumor and normal tissues across The Cancer Genome Atlas (TCGA) pan-cancer was analyzed using the TIMER2.0 database (http://timer.comp-genomics.org/) (Li et al., 2020). Analytical data for differential expression of NFE2L3 in paired pan-cancer samples were accessed from TCGA data portal (https://www.cancer.gov/). The subcellular localization of NFE2L3 was investigated using the HPA database and the differential expression of NFE2L3 protein in various tumor and normal tissues was analyzed. R software v3.6.3 and the ggplot2 package v3.3.3 were used for statistical analysis and visualization. P-values were calculated with the Wilcoxon signed rank test and Wilcoxon rank sum test, and $p < 0.05$ was accepted as statistically significant (ns, $p \geq 0.05$; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$; $****p < 0.0001$) (Stanley et al., 2020).

Clinicopathological Correlation and Survival Prognostic Analysis

The RNA-seq data (level3) and related clinical information were obtained from TCGA cohort. The relationship between NFE2L3 expression and survival prognosis in patients with pan-cancer was analyzed using univariate Cox regression and the results were visualized using the "forest plot" R package. The differential expression Wilcoxon signed rank analysis of NFE2L3 among tumor pathological stages was performed on Gene Expression Profiling Interactive Analysis v2.0 (GEPIA2) (http://gepia2.cancer-pku.cn/#index) (Tang et al., 2019). P-values < 0.05 were considered statistically significant.

Diagnostic Value Analysis

Receiving operating characteristic (ROC) curves were constructed to evaluate the diagnostic value of NFE2L3 in multiple tumors. The area under the ROC curve (AUC) was rated as outstanding discrimination (AUC ≥ 0.90), excellent discrimination (0.8 ≤ AUC < 0.9), acceptable discrimination (0.7 ≤ AUC < 0.8), and poor discrimination (AUC < 0.70). The pROC package v1.17.0.1, and the ggplot2 package v3.3.3 were used for analysis and visualization, respectively.

Epigenetic Methylation Analysis

Methylation data were obtained from the Illumina Infinium DNA methylation platform arrays HumanMethylation450. DNA methylation levels of NFE2L3 in TCGA pan-cancer were analyzed using the SMART database (http://www.bioinfo-zs.com/smartapp/) (Li et al., 2019). Comparison of DNA methylation levels in 19 methylation CpG sites of NFE2L3 in TCGA pan-cancer (PANCAN) cohort was performed using the UCSC Xena database (http://xena.ucsc.edu/) (Goldman et al., 2020). The correlation between DNA methylation levels of NFE2L3 and mRNA expression of NFE2L3, tumor cell dryness, immune subtypes, and survival analyses of differential DNA methylation levels in TCGA pan-cancer (PANCAN) cohort were investigated using UCSC Xena. Moreover, the analysis results were visualized through GraphPad Prism v8.4.2 and the survminer package v0.4.9 (ns, $p \geq 0.05$; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$; $****p < 0.0001$).

Analysis of Protein Topology Mutants and Genetic Alterations

The Protter database (https://wlab.ethz.ch/protter/start/) is a web-based application tool used for visualizing the sequences, topologies, and annotations of individual proteins (Omasits et al., 2014). The variation in NFE2L3 protein topology was evaluated using the Protter database and a schematic representation of the

secondary structure of the NFE2L3 protein was generated. In addition, the genetic alterations of NFE2L3 in pan-cancers were analyzed using cBioPortal (http://www.cbioportal.org/) (Cerami et al., 2012; Gao et al., 2013).

Immune Feature Analysis

The correlation between NFE2L3 expression and immune cell infiltration, immune checkpoints, immunosuppressive cell infiltration, and immune cell markers in TCGA pan-cancer was analyzed using the TIMER database. The results of the analysis were visualized using heat maps of Spearman's correlations that were generated from the ggplot2 package v3.3.3. Furthermore, the differential expression of NFE2L3 in different molecular subtypes and immune subtypes in pan-cancer were analyzed using the TISIDB database (http://cis.hku.hk/TISIDB/) (Ru et al., 2019). Based on previous studies and on RNA-seq data from TCGA pan-cancer, the correlation between NFE2L3 expression and tumor mutational burden (TMB) and microsatellite instability (MSI) was analyzed and visualized by R software (version 3.6.3) (Bonneville et al., 2017). The potential of NFE2L3 as a biomarker for predicting responses to tumor immune checkpoint blockade therapy was evaluated using the TIDE database (http://tide.dfci.harvard.edu/). In addition, the correlation between NFE2L3 expression and immune checkpoint blockade (ICB) overall survival outcome was analyzed in the immunotherapy dataset and the ability of NFE2L3 knockdown to regulate lymphocyte-mediated tumor killing efficacy in the CRISPR Screen dataset was evaluated (Jiang et al., 2018). All correlation analyses were performed using Spearman's correlation test. P-value < 0.05 was considered to be statistically significant.

Functional Status Analysis of Tumor Cells

The correlation between NFE2L3 expression and functional status of multiple tumor cells based on single-cell sequencing data was analyzed using the CancerSEA database (http://biocc.hrbmu.edu.cn/CancerSEA/) (Yuan et al., 2019).

Functional Enrichment Analysis

GeneMANIA was used to identify the top 20 genes related to NFE2L3 (Warde-Farley et al., 2010). STRING was used to conduct the protein-protein interaction (PPI) network by setting the following main parameters: minimum required interaction score ("0.300") and maximum number of interactors ("no more than 50 interactors"). Visualization of the PPI network was constructed using Cytoscape. The top 100 genes with a similar expression pattern to NFE2L3 were surveyed using the GEPIA2 database (Szklarczyk et al., 2019). In addition, the intersection of genes related to NFE2L3 obtained from the three databases were identified with results being visualized with a Venn diagram. The correlation between the expression of NFE2L3 and the intersection of the genes related to NFE2L3 obtained from the above three databases was analyzed using the GEPIA2 database. Gene Ontology (GO) and Kyoto Encyclopedia of Genes (KEGG) pathway enrichment analyses was performed for NFE2L3-related genes. The ggplot2 package v3.3.3 was used for visualization and the cluster Profiler package v3.14.3 was used for statistical analysis. A p-value < 0.05 was considered to be statistically significant. In addition, tumor patients in TCGA cohort were divided into high and low expression groups according to the median expression level of NFE2L3 mRNA, including BRCA, CHOL, ESCA, HNSC, KIRC, PRAD, UCEC, and THCA. The differentially expressed genes in the high and low expression groups in the multiple tumors described above were analyzed using the DESeq2 package (version 1.26.0). GSEA analysis of the differentially expressed genes was performed in these tumors using the clusterProfiler package (version 3.14.3).

Co-Expression Genes and Differentially Expressed Genes Analysis in Liver Hepatocellular Carcinoma

The top five genes positively associated with NFE2L3 expression and the top five negatively associated genes in LIHC were evaluated. The results were visualized using a heat map with Spearman's correlation. Furthermore, TCGA-LIHC cohort was divided into a high and low expression group according to the expression level of NFE2L3 mRNA, and the differentially expressed genes (DEGs) between the high and low expression group of NFE2L3 in LIHC were analyzed (Love et al., 2014). The results were visualized using a volcano map with the following threshold values: |log2 fold-change (FC)| > 2.0, and adjusted p-values < 0.05. GO and KEGG enrichment analyses were performed for the DEGs that met the screening requirements. In addition, all the differential genes between the high and low expression groups of NFE2L3 were included in the GSEA analysis. The top 10 signaling pathways from the analysis were presented as merged plots of GSEA. R (version 3.6.3) was used for statistical analysis and visualization, and the main R package that was used included the ggplot2 package v3.3.3 and the cluster Profiler package v3.14.3.

Prognostic Model Based on NFE2L3 Expression and Clinical Characteristics in Liver Hepatocellular Carcinoma

The correlation between NFE2L3 expression and multiple clinical characteristics of LIHC were analyzed using logistic regression. The factors associated with the overall survival (OS) of patients with LIHC were analyzed using univariate and multifactorial Cox regression. In addition, a nomogram integrating NFE2L3 expression and the prognostic factors of the multivariable model for OS in LIHC from TCGA data was created using the nomogram function from the RMS package v6.2-0 and survival package v3.2-10. The accuracy of the nomogram model for predicting overall survival of patients with LIHC was evaluated using calibration curves. The related RNA-seq and clinical data were accessed from TCGA database, normalized as transcripts per million (TPM), and then log2 transformed. The Wilcoxon signed rank test was used to determine statistical significance (ns, $p \geq 0.05$; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$; $****p < 0.0001$).

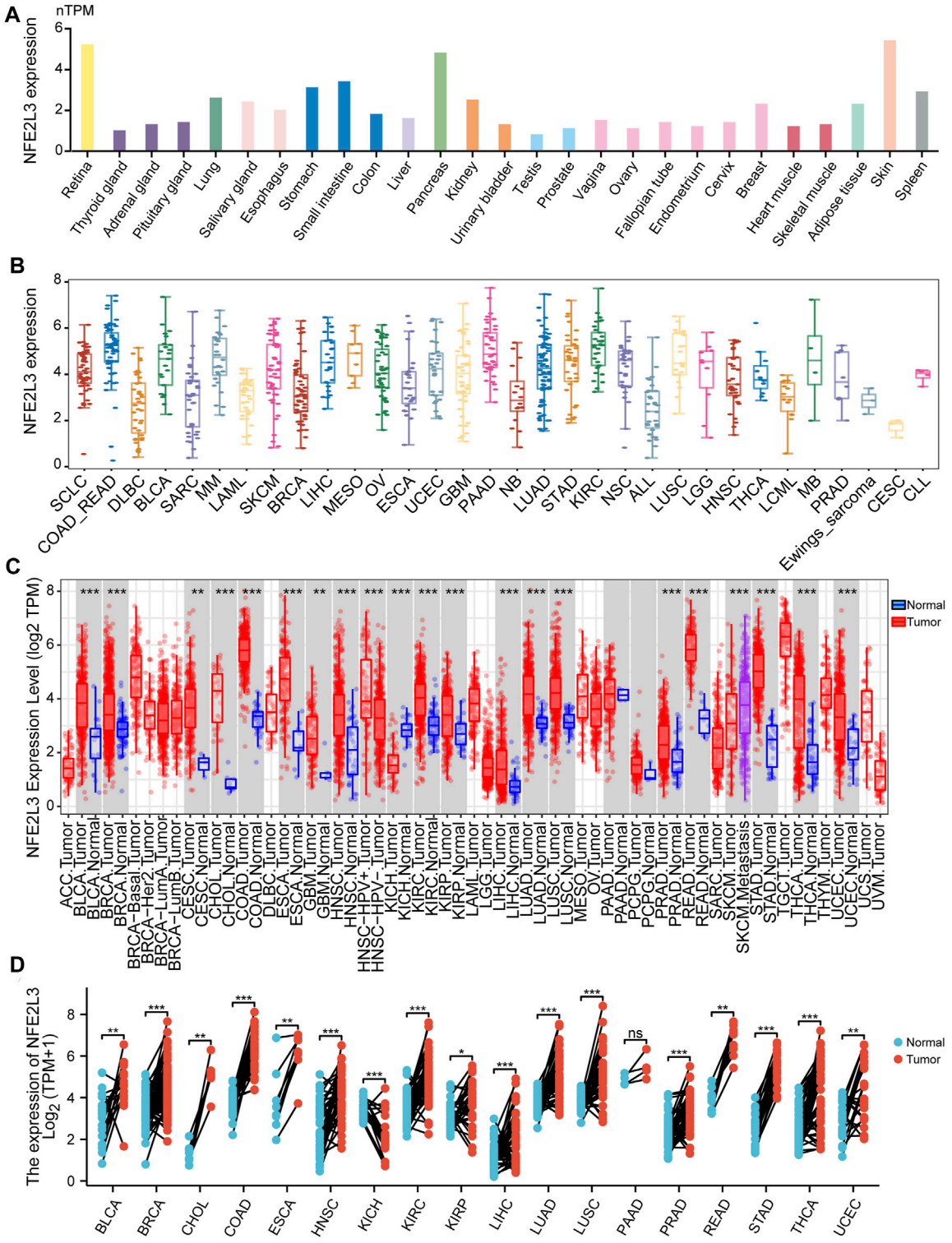


FIGURE 1 | Expression level of NFE2L3 in pan-cancer and normal human tissues. **(A)** The expression level of NFE2L3 in different normal tissues was analyzed by the HPA and GTEx databases. **(B)** The CCLE database was used to analyze the expression levels of NFE2L3 in different cancer cell lines. **(C)** The TIMER2 database was used to analyze the expression levels of NFE2L3 in tumor and normal tissues. **(D)** TCGA database was used to analyze the expression levels of NFE2L3 in tumor tissues and adjacent normal tissues. ns, $p \geq 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

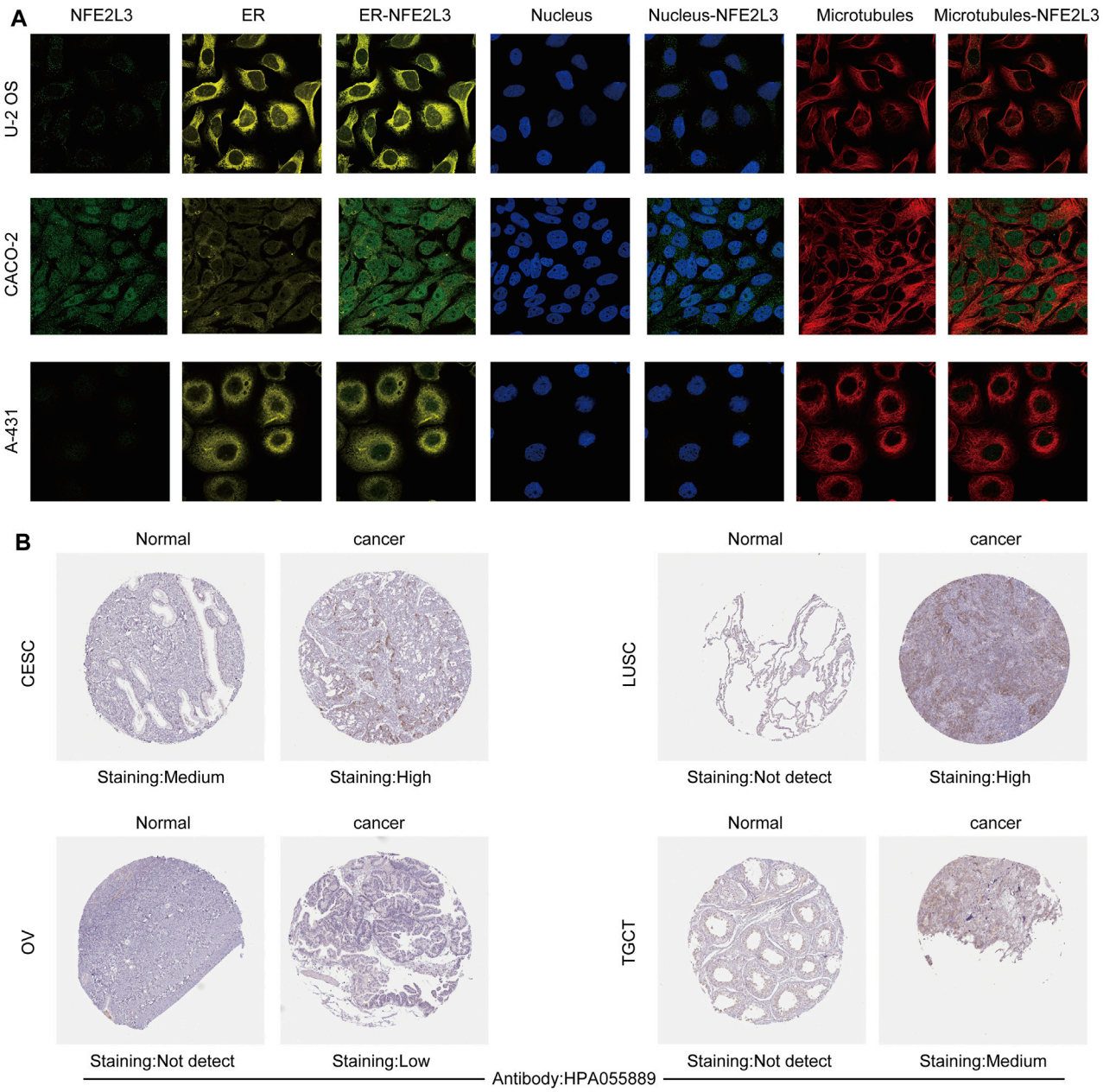


FIGURE 2 | Subcellular localization and immunohistochemical analysis of NFE2L3. **(A)** NFE2L3 was mainly located in the nucleoplasm and vesicles by immunofluorescence (ICC-IF) and confocal microscopy. **(B)** The expression levels of NFE2L3 protein in four tumors and normal tissues were analyzed by immunohistochemistry. The results were all obtained from the HPA database.

RESULTS

The Expression of NFE2L3 in Pan-Cancer

The expression level of NFE2L3 in normal human tissues using the HPA database was evaluated. The results showed that the expression level of NFE2L3 was low in most normal tissues, however it was high in the retina, pancreas, and skin (**Figure 1A**). NFE2L3 was expressed at high levels in a variety of tumor cell lines, according to the CCLE database analysis (**Figure 1B**). The TIMER2.0 database analysis showed that the expression of NFE2L3 in 18 cancer types was significantly higher than that in normal tissues, including bladder urothelial carcinoma (BLCA), breast invasive carcinoma (BRCA), cervical squamous cell carcinoma endocervical adenocarcinoma (CESC), cholangiocarcinoma (CHOL), colon adenocarcinoma (COAD), esophageal carcinoma (ESCA), glioblastoma multiforme (GBM), head and neck squamous cell carcinoma (HNSC), KIRC, kidney renal

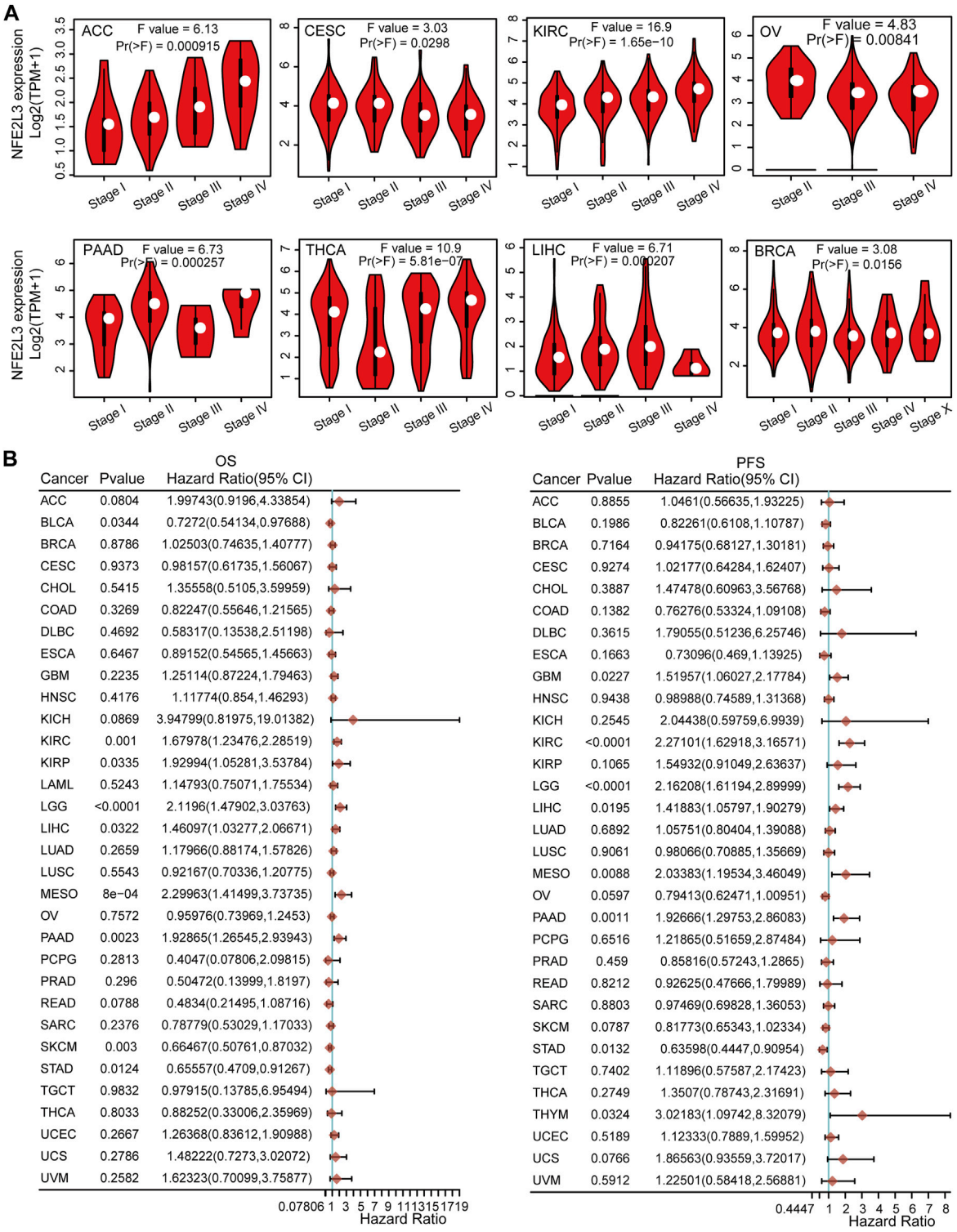


FIGURE 3 | The relationship between NFE2L3 expression and prognosis and pathological stages in different tumors. **(A)** The association between NFE2L3 expression and the pathological stage was analyzed in ACC, CESC, KIRC, OV, PAAD, THCA, LIHC, and BRCA by the GEPIA2 database. **(B)** The correlation of NFE2L3 expression and Overall survival (OS) and progression-free survival (PFS) prognosis in pan-cancer.

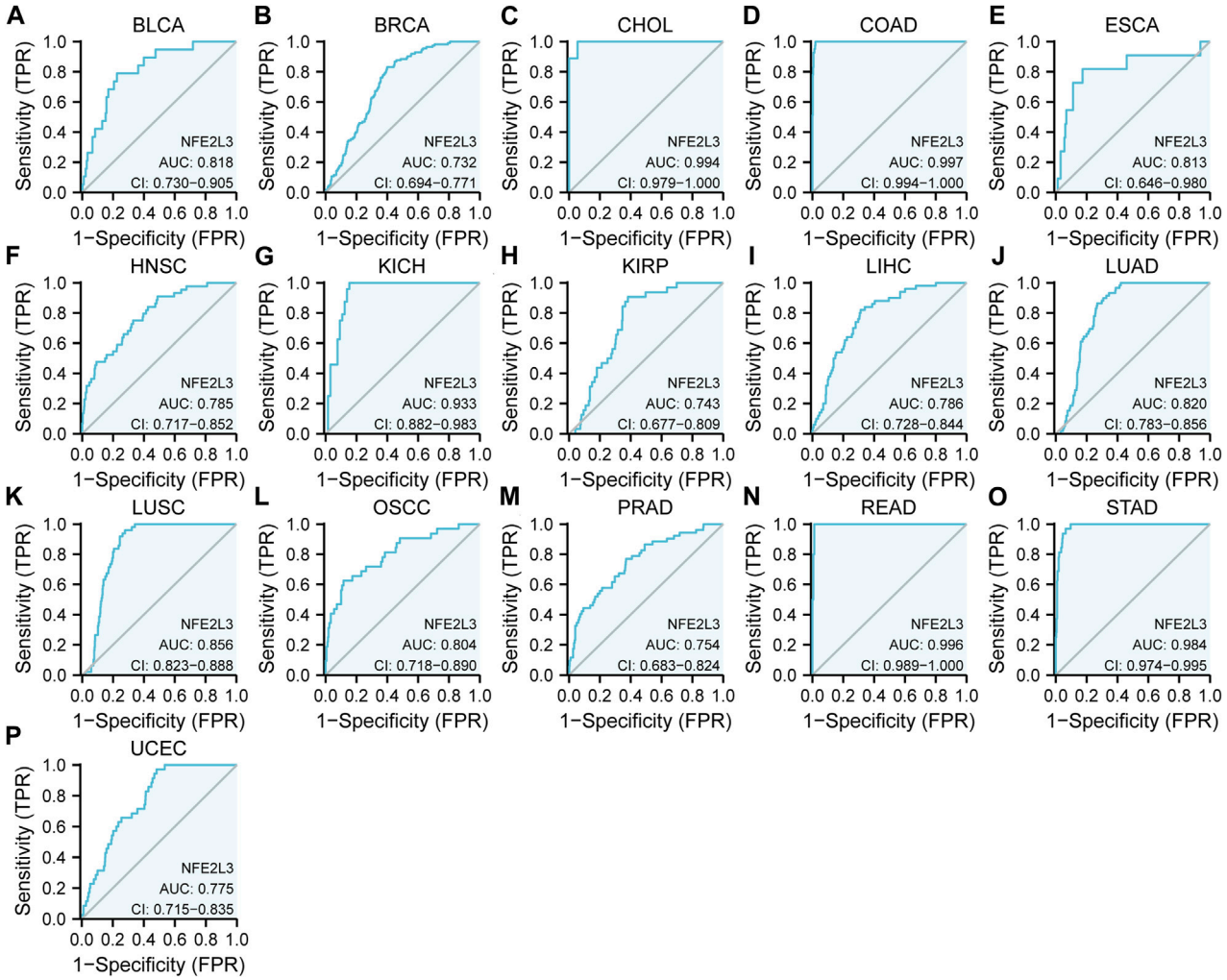


FIGURE 4 | ROC curve of NFE2L3 expression levels in different tumors. **(A)** BLCA; **(B)** BRCA; **(C)** CHOL; **(D)** COAD; **(E)** ESCA; **(F)** HNSC; **(G)** KICH; **(H)** KIRP; **(I)** LIHC; **(J)** LUAD; **(K)** LUSC; **(L)** OSCC; **(M)** PRAD; **(N)** READ; **(O)** STAD; **(P)** UCEC.

papillary cell carcinoma (KIRP), LIHC, lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), prostate adenocarcinoma (PRAD), rectum adenocarcinoma (READ), stomach adenocarcinoma (STAD), thyroid carcinoma (THCA), and uterine corpus endometrial carcinoma (UCEC) (**Figure 1C**). Furthermore, NFE2L3 expression was significantly upregulated in paired cancer and adjacent samples of 16 cancer types, including BLCA, BRCA, CHOL, COAD, ESCA, HNSC, KIRC, KIRP, LIHC, LUAD, LUSC, PRAD, READ, STAD, THCA, and UCEC (**Figure 1D**). NFE2L3 was also demonstrated to be localized in the nucleoplasm and vesicles (**Figure 2A**), and the protein expression level of NFE2L3 was significantly higher in ovarian serous cystadenocarcinoma (OV), TGCT, LUSC, and CESC than in normal tissues (**Figure 2B**). The above results demonstrated the upregulation of NFE2L3 expression in a variety of tumors, suggesting that NFE2L3 expression may promote tumor progression.

Clinicopathological Correlation and Survival Prognostic Analysis

The correlation between NFE2L3 expression and clinicopathological stage in pan-cancer was also investigated. Our study showed that NFE2L3 expression correlated with the clinicopathological stages (stages I, II, III, IV, and X) of eight tumors, including ACC, CESC, KIRC, OV, PAAD, THCA, LIHC, and BRCA (**Figure 3A**).

To further explore the relationship between NFE2L3 expression and prognosis, overall survival (OS), progression-free survival (PFS), and disease-free survival (DSS) analyses with median group cutoff by univariate Cox regression analysis for pan-cancers were performed. This study showed that higher expression of NFE2L3 had poorer prognosis for OS in six tumor types, including KIRC, KIRP, brain lower grade glioma (LGG), LIHC, mesothelioma (MESO), and pancreatic adenocarcinoma (PAAD). In contrast, higher

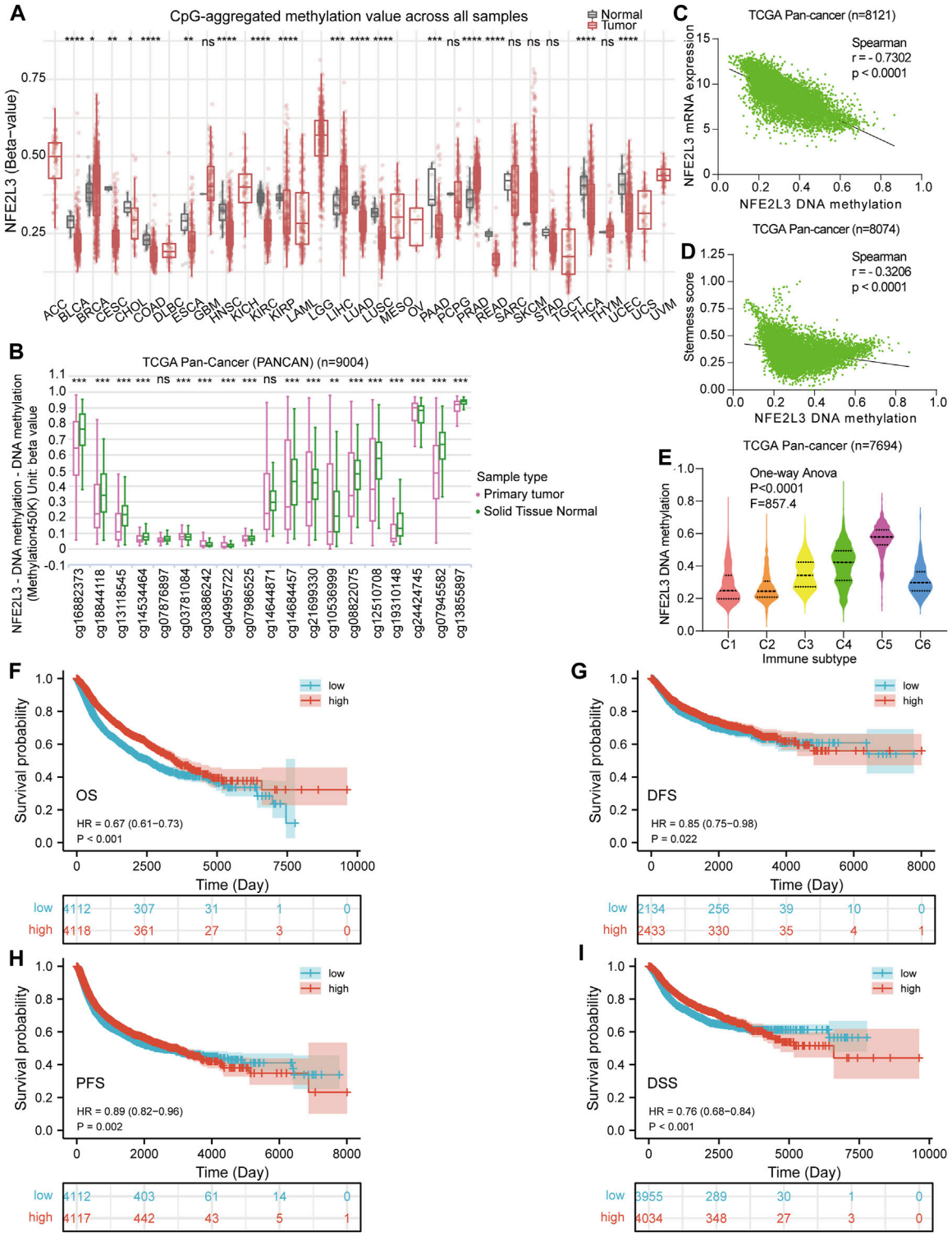


FIGURE 5 | DNA methylation analysis of NFE2L3 in pan-cancer. **(A)**Analysis of DNA methylation levels of NFE2L3 in pan-cancer using the SMART database. **(B)** The 19 CPG methylation sites of NFE2L3 in TCGA pan-cancer (PANCAN) cohort were analyzed by the UCSC Xena database for comparison of NFE2L3 DNA methylation levels between tumor and normal groups. **(C)** The correlation between DNA methylation levels of NFE2L3 and NFE2L3 mRNA expression levels in TCGA PANCAN cohort. **(D)** The correlation between DNA methylation levels of NFE2L3 and tumor cell stemness in TCGA PANCAN cohort. **(E)** The DNA methylation levels of NFE2L3 in different immune subtypes in TCGA PANCAN cohort. **(F–I)** The association between DNA methylation levels of NFE2L3 and OS, disease-free survival (DFS), PFS, and disease-specific (DSS) prognosis in TCGA PANCAN cohort. ns, $p \geq 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$.

expression of NFE2L3 was associated with a better prognosis for OS in BLCA, skin cutaneous melanoma (SKCM), and STAD. For PFS, higher expression of NFE2L3 had a poor prognosis in seven tumor types, including GBM, KIRC, LGG, LIHC, MESO, PAAD, and THYM. However, higher NFE2L3 expression was associated with a better prognosis for PFS in patients with STAD (**Figure 3B**). Furthermore, for DSS, the higher expression of NFE2L3 was associated with a worse prognosis in five tumor types, including KIRC, LGG, LIHC, MESO, and PAAD. However, higher NFE2L3 expression was associated with a better prognosis for DSS in COAD, SKCM, and STAD (**Supplementary Figure S1**). These results suggest that NFE2L3 expression is closely related to tumor progression and patient prognosis, thus NFE2L3 can be used as a prognostic marker for a variety of cancers.

Diagnostic Value Analysis

The diagnostic value of NFE2L3 expression in pan-cancer was evaluated by using the ROC curve. As shown in **Figures 4A–P**, NFE2L3 had promising efficacy in the diagnosis of 11 tumors, including BLCA (AUC = 0.818), BRCA (AUC = 0.732), ESCA (AUC = 0.813), HNSC (AUC = 0.785), KIRP (AUC = 0.743), LIHC (AUC = 0.786), LUAD (AUC = 0.820), LUSC (AUC = 0.856), OSCC (AUC = 0.804), PRAD (AUC = 0.754), and UCEC (AUC = 0.775). Notably, NFE2L3 exhibited high accuracy in the diagnosis of five tumors, including CHOL (AUC = 0.994), COAD (AUC = 0.997), KICH (AUC = 0.933), READ (AUC = 0.996), and STAD (AUC = 0.984). This further demonstrated that NFE2L3 can be used as a diagnostic biomarker for a variety of tumors.

DNA Methylation Analysis

Tumorigenesis and epigenetic modifications of genes are closely related, and DNA methylation is one of the most widely studied types of epigenetic modification. Thus, DNA methylation levels of NFE2L3 in normal and primary tumor tissues in pan-cancer were assessed. The DNA methylation levels of NFE2L3 were lower in BLCA, CESC, CHOL, COAD, ESCA, HNSC, KIRC, KIRP, PRAD, LUSD, LUSC, PAAD, READ, THCA, and UCEC than in normal tissues. However, the DNA methylation levels of NFE2L3 were higher in BRCA, LIHC, and PRAD tissues than in normal tissues (**Figure 5A**). In addition, we showed that the DNA methylation levels of NFE2L3 were significantly lower in tumor tissues than in normal tissues in most of the NFE2L3 methylation CpG sites in TCGA pan-cancer (PANCAN) cohort, including cg16882373, cg18844118, cg13118545, cg14534464, cg03886242, cg04995722, cg07986525, cg14684457, cg21699330, cg10536999, cg08822075, cg12510708, cg19310148, cg07945582, and cg13855897 (**Figure 5B**). There was a significant negative correlation between the DNA methylation level of NFE2L3, the mRNA expression level and tumor cell stemness score in TCGA PANCAN cohort (**Figures 5C,D**). DNA methylation levels of NFE2L3 significantly correlated with immune subtypes in TCGA PANCAN cohort, with lower DNA methylation levels of NFE2L3 in the immune subgroups of C1 and C2 types, however higher DNA methylation levels of NFE2L3 in the immune subgroup of C5 types were observed (**Figure 5E**). This study demonstrated that lower NFE2L3 DNA methylation levels were associated with worse OS, PFS, DSS, and disease-free survival (DFS) prognosis in TCGA PANCAN cohort (**Figures 5F–I**).

Genetic Alteration Analysis

Accumulation of gene mutations is one of the causes of tumorigenesis; therefore, the landscape of genetic alterations of NFE2L3 in pan-cancer were analyzed. Transmembrane protein topology of NFE2L3 revealed a natural missense variant of Valine441 in membrane localization (**Figure 6A**). The genetic alteration status of NFE2L3 in multiple tumor samples from the TCGA PANCAN cohort were investigated using the cBioPortal. As shown in **Figure 6B**, NFE2L3 mutations occurred in most tumor types, with the top three tumors, UCEC, BLCA, and ESCA, having high NFE2L3 mutations (>6%). The tumor, UCEC, had the highest mutation of NFE2L3 (>6%). A total of 139 NFE2L3 mutations were identified, 108 (77.7%) were missense mutations, 25 (18.0%) were truncating mutations, 3 (2.2%) were fusion mutations, 2 (1.4%) were in-frame mutations, and 1 (0.7%) was a splice mutation (**Figure 6C**). All mutations were dispersed in the full sequence and 3D protein structure of NFE2L3 (**Figures 6C,D**). As shown in **Figure 6E**, genetic alterations in NFE2L3 were associated with multiple clinical characteristics. The overall alteration frequency of NFE2L3 was 2.2% (244/10,950) in TCGA pan-cancer cohort. In addition, we performed a survival analysis of the genetic alteration status of NFE2L3 was performed on the entire TCGA pan-cancer cohort, however the results revealed no significant correlation between the genetic alteration status of NFE2L3 and patient prognosis. (**Figure 6F**).

Immune Feature and Immunotherapy Response Analysis

The immune response of tumors is an important process in tumor development, thus the role of NFE2L3 in tumor immune regulation was investigated by analyzing the immune features and immunotherapeutic responses. The correlations between the expression of NFE2L3 and different immune signatures in TCGA pan-cancer cohort were evaluated, including immune cell infiltration, immune checkpoints, immunosuppressive cell infiltration, and immune cell markers. A significant positive correlation between the expression of NFE2L3 and the level of immune cell infiltration in TCGA pan-cancer cohort was demonstrated, except for COAD, OV, UCEC, and UCS (**Figure 7A**). However, a significant positive correlation between NFE2L3 expression and immune checkpoint expression in TCGA pan-cancer cohort was demonstrated except for READ, UCEC, and UCS (**Figure 7B**). A significant positive correlation was also shown between NFE2L3 expression and the levels of Tregs, MDSC, and CAF in TCGA pan-cancer cohort, except for ACC, CESC, and UCS. However, NFE2L3 expression and the level of M2-TAM infiltration were significantly negatively correlated in most tumors (**Figure 7C**). Our study showed a significant positive correlation between NFE2L3 expression and most immune cell markers in TCGA

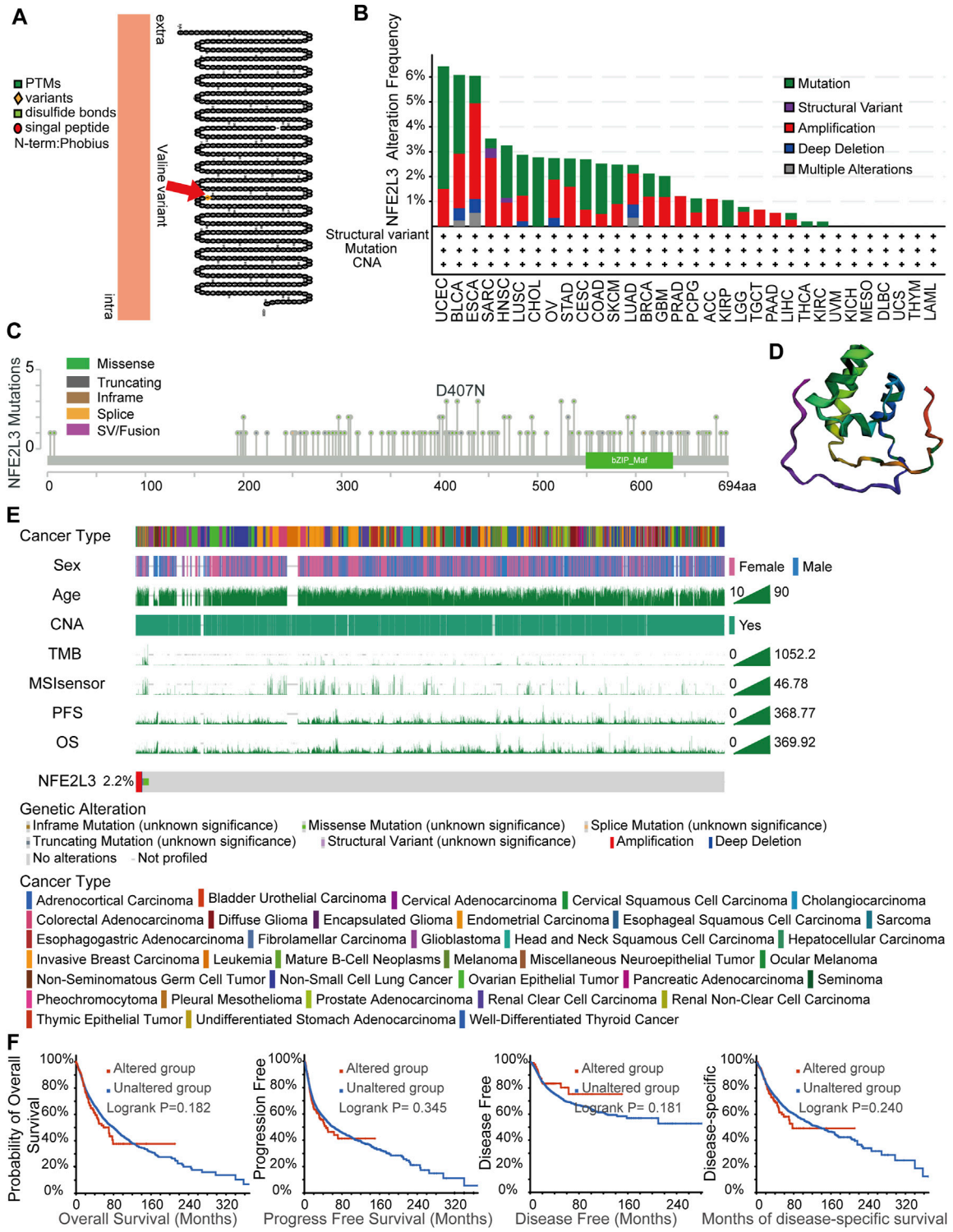


FIGURE 6 | NFE2L3 variant and the characteristics of NFE2L3 mutations in TCGA cohort. **(A)** A natural missense variant of Valine441 in transmembrane protein topology of NFE2L3. **(B)** Genetic alteration frequencies of NFE2L3 in TCGA pan-cancer cohorts. **(C)** Type and site of mutations of NFE2L3. **(D)** Variant distribution of 3D protein structure of NFE2L3. **(E)** The correlation of NFE2L3 status and clinical characteristics in TCGA pan-cancer cohort. **(F)** The correlation between NFE2L3 mutation status and prognosis (OS, PFS, DFS, and DSS) of cancer.

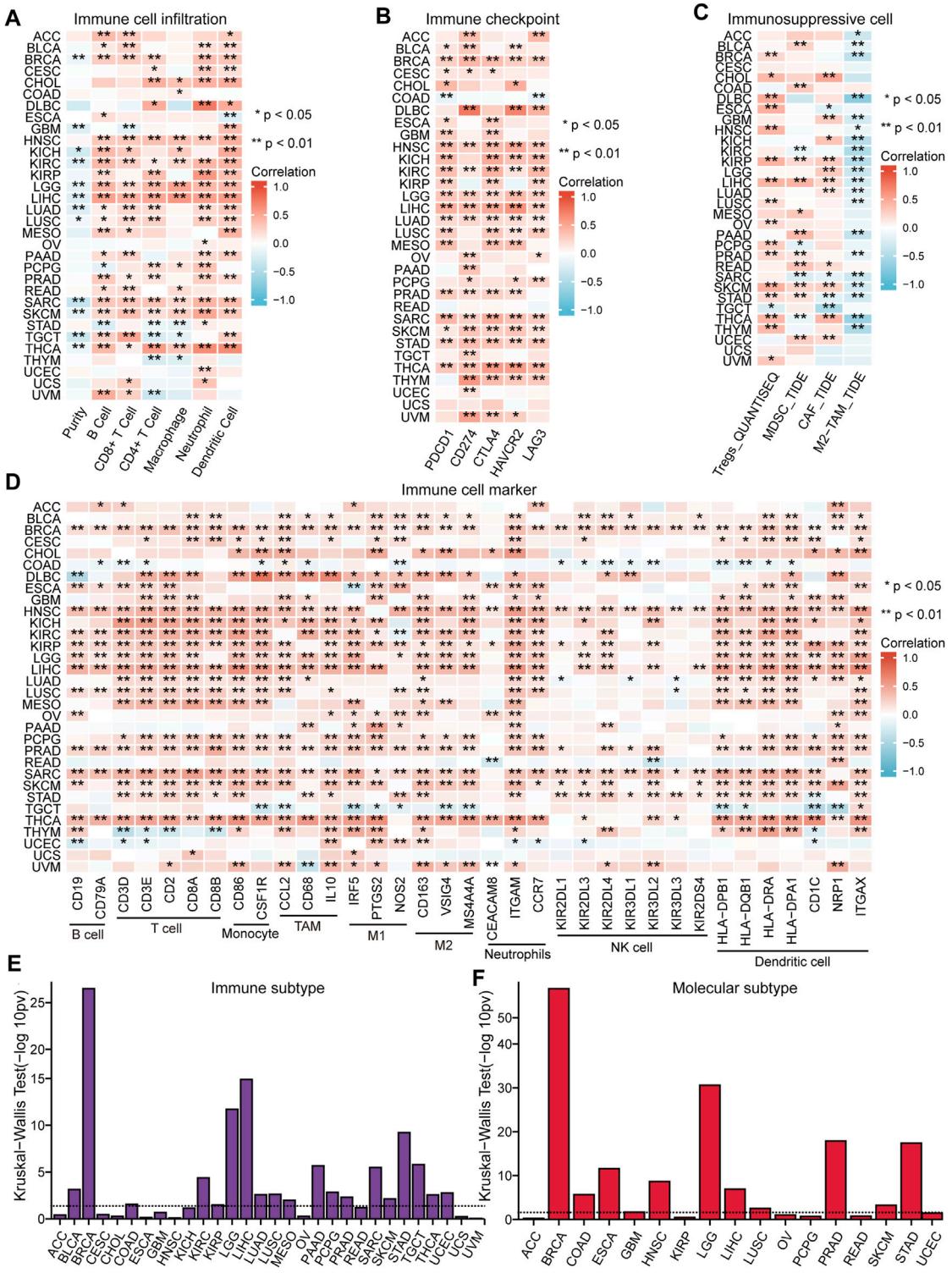


FIGURE 7 | The correlation between NFE2L3 expression and immune feature in pan-cancer. **(A)** The correlation between NFE2L3 expression and immune infiltration in pan-cancer. **(B)** The correlation between NFE2L3 expression and immune checkpoint in pan-cancer. **(C)** The correlation between NFE2L3 expression and T cell exclusion filtration in pan-cancer. **(D)** The correlation between NFE2L3 expression and immune cell marker in pan-cancer. **(E)** The correlation between NFE2L3 expression and immune subtype in pan-cancer. **(F)** The correlation between NFE2L3 expression and molecular subtype in pan-cancer.

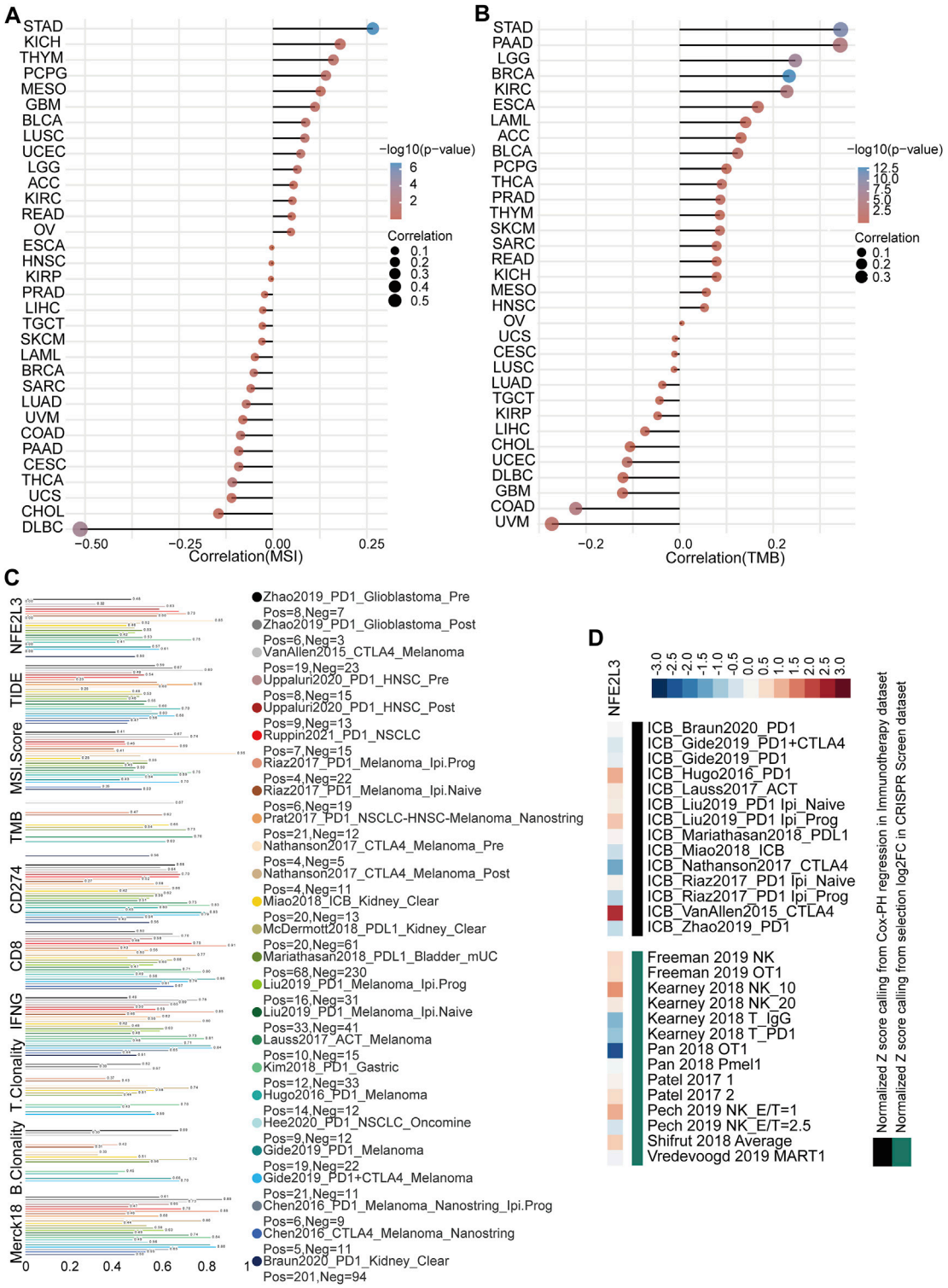


FIGURE 8 | Immunotherapy response analysis in pan-cancer. **(A)** The correlation between NFE2L3 expression and MSI in pan-cancer. **(B)** The correlation between NFE2L3 expression and TMB in pan-cancer. **(C)** The potential of NFE2L3 as a biomarker to predict response to immune checkpoint blockade. **(D)** NFE2L3 was ranked by its weighted average value across the association with ICB survival outcome and log-fold change (logFC) in CRISPR screens.

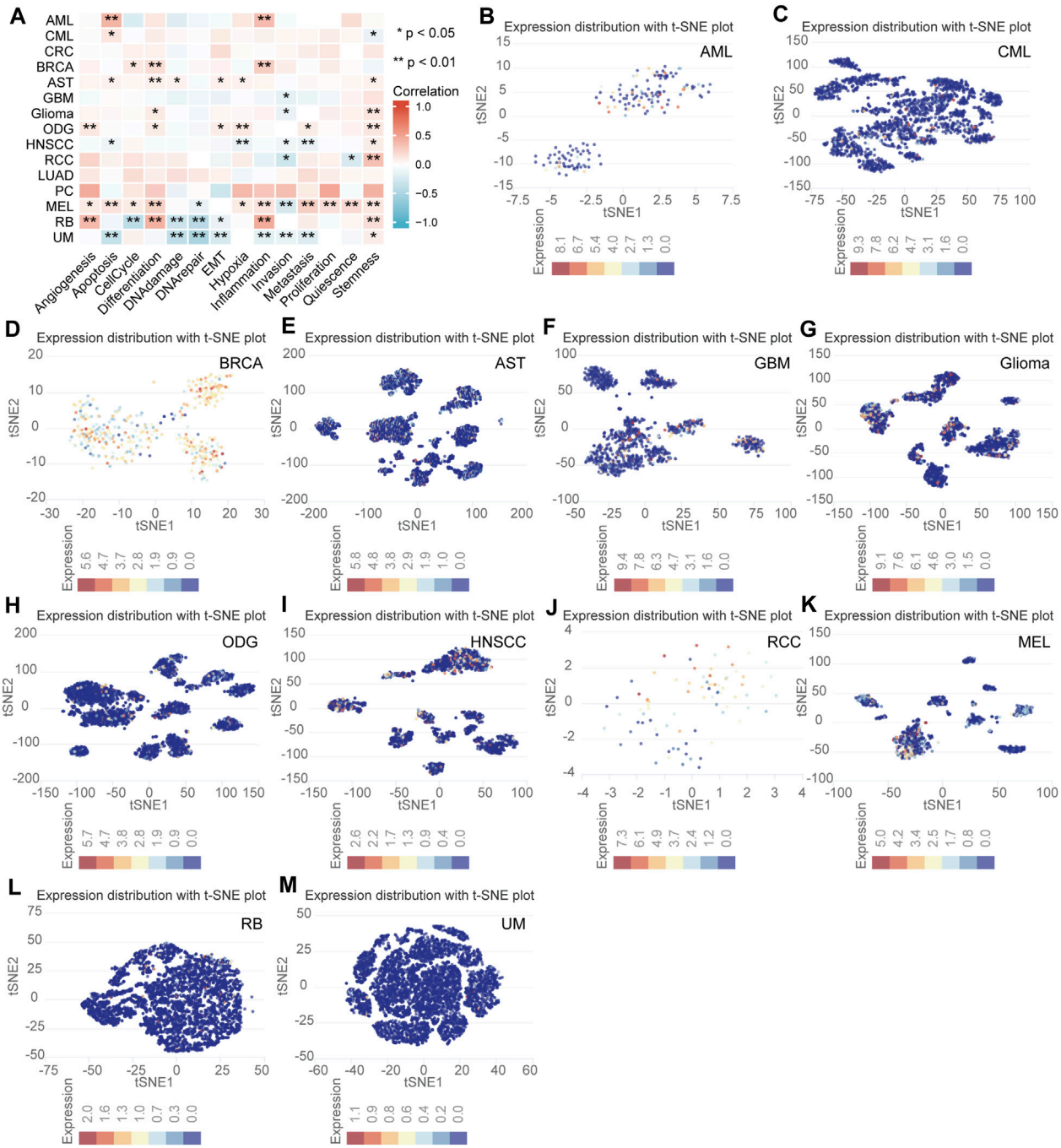


FIGURE 9 | The correlation analysis of NFE2L3 expression and single cell functional status in multiple tumors. **(A)** Heat map of correlation between NFE2L3 expression and single cell function in multiple tumors. **(B–M)** T-SNE diagram of NFE2L3 expression levels in single cells of multiple tumors including acute myeloid leukemia (AML), chronic myelogenous leukemia (CML), breast cancer (BRCA), astrocytoma (AST), glioblastoma (GBM), glioma, oligodendroglioma (ODG), head and neck cancer (HNSCC), renal cell carcinoma (RCC), melanoma (MEL), retinoblastoma (RB), and uveal melanoma (UM).

pan-cancer cohort, except for UCS (**Figure 7D**). This study showed that NFE2L3 expression was also significantly positively correlated with three immunomodulators, including immunoinhibitory, immunostimulatory, and MHC molecules in TCGA pan-cancer cohort (**Supplementary Figure S2**). The correlations between NFE2L3 expression and immune or molecular subtypes in TCGA pan-cancer cohort from the TISIDB database were investigated. The results showed that NFE2L3 was expressed differently in different immune subtypes of 12 cancer types, including KIRC, COAD, BRCA, UCEC, STAD, SKCM, PRAD, PAAD, MESO, LUSC, LIHC, and LGG (**Figure 7E**). The differential expression of NFE2L3 was also

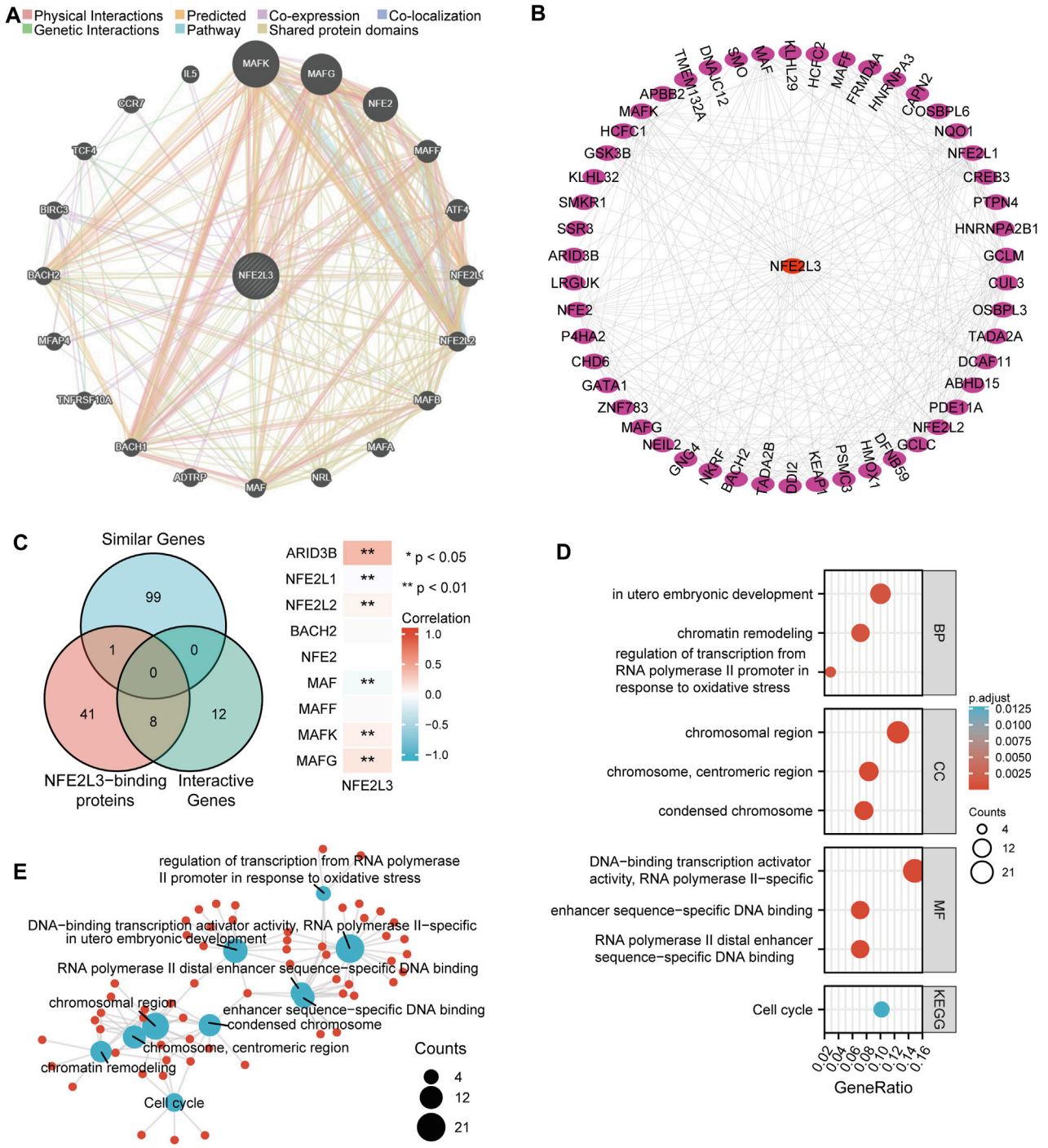


FIGURE 10 | Functional enrichment analysis of NFE2L3-related genes. **(A)** Top 20 interactive genes of NFE2L3. **(B)** Top 50 NFE2L3-binding proteins. **(C)** Venn diagram of NFE2L3-related genes and the correlation between expression levels of NFE2L3 and nine overlapping genes. **(D,E)** GO and KEGG pathway enrichment analysis of NFE2L3-related genes.

found in various molecular subtypes of the nine cancer types, including UCEC, STAD, PRAD, LUSC, LIHC, LGG, HNSC, COAD, and BRCA (**Figure 7F**).

The effect of NFE2L3 expression on immunotherapy was evaluated by analyzing the correlation between NFE2L3 expression and TMB or MSI. The results showed that NFE2L3 expression was significantly positively correlated with MSI in STAD, KICH, and THYM, however NFE2L3 expression was negatively correlated with DLBC and CHOL (**Figure 8A**). NFE2L3 expression was significantly positively correlated with

TMB in STAD, PAAD, LGG, BRCA, and KIRC however it was negatively correlated with COAD and UVM (**Figure 8B**). The potential of NFE2L3 as a biomarker to predict the response to immune checkpoint blockade (ICB) was evaluated. The results showed that NFE2L3 had a limited accuracy (AUC > 0.5) in predicting the response to ICB in 12 of the 23 ICB sub-cohorts and compared with TMB, T cell clonality, and B cell clonality, NFE2L3 displayed a higher predictive power (**Figure 8C**). High NFE2L3 expression was associated with poor survival prognosis in the ICB_VanAllen2015_CTLA4, ICB_Hugo2016_PD1, and ICB_Liu2019_PD1 Ipi_Prog cohorts. However, knockdown of NFE2L3 enhanced the efficacy of lymphocyte-mediated tumor killing in the Kearney 2018 NK_10, Pech 2019 NK_E/T = 1, and Shifrut 2018 average cohorts (**Figure 8D**). These results suggested that NFE2L3 may promote tumor development by regulating tumor immunity.

Single Cell Sequencing Analysis

The expression level and functional status of NFE2L3 in single tumor cells using the CancerSEA database was determined. NFE2L3 expression was shown to be closely related to the cellular functional status of a variety of tumors, including acute myeloid leukemia (AML), chronic myelogenous leukemia (CML), breast cancer (BRCA), astrocytoma (AST), glioblastoma (GBM), glioma, oligodendroglioma (ODG), head and neck cancer (HNSCC), renal cell carcinoma (RCC), melanoma (MEL), retinoblastoma (RB), and uveal melanoma (UM). In most tumor cells, NFE2L3 expression was positively correlated with differentiation, angiogenesis, apoptosis, inflammation, and tumor cell stemness however, NFE2L3 expression negatively correlated with DNA damage, DNA repair, and invasion (**Figure 9A**). NFE2L3 was expressed at higher levels in single cells of AML, CML, BRCA, AST, GBM, glioma, ODG, RCC, and MEL monocyte samples however, it was expressed at lower levels in single cells of HNSCC, RB, and UM samples (**Figures 9B–M**).

Gene Function Enrichment Analysis

The molecular functions of NFE2L3 was evaluated in different tumors by performing GO and KEGG pathway enrichment analyses. The top 20 genes that interacted with NFE2L3 were evaluated using the GeneMANIA database (**Figure 10A**). Then the top 50 proteins interacting with NFE2L3 were predicted and visualized using the STRING database using Cytoscape, respectively (**Figure 10B**). The GEPIA2 database was then used to explore the top 100 genes with a similar expression pattern to NFE2L3 in the pan-cancer analysis. The intersection of NFE2L3-related genes predicted by the above three databases were used and the results were visualized using a Venn diagram. Nine genes were predicted by at least two databases, including NFE2L1, NFE2L2, BACH2, NFE2, MAF, MAFF, MAFK, MAFG, and ARID3B. Furthermore, the expression of NFE2L3 was shown to be significantly correlated with the expression of six genes at the intersection in TCGA pan-cancer cohort by the GEPIA2 database (**Figure 10C**). In addition, GO and KEGG pathway enrichment analyses of NFE2L3-related genes using the three databases were performed. GO enrichment analysis revealed that the primary biological process (BP) was mainly related to embryonic development and regulation of transcription from the RNA polymerase II promoter in response to oxidative stress and chromatin remodeling. The cellular component (CC) contained chromosomal regions, chromosomes, centromeric regions, and condensed chromosomes. The molecular function (MF) was primarily enriched in DNA-binding transcription activator activity, RNA polymerase II-specific, RNA polymerase II distal enhancer sequence-specific DNA binding, and enhancer sequence-specific DNA binding. The KEGG pathway enrichment was involved in the cell cycle (**Figures 10D,E**).

GSEA analysis in BRCA, CHOL, ESCA, HNSC, KIRC, PRAD, UCEC, and THCA was performed. **Supplementary Figure S3** showed the top 10 enriched signaling pathways in each tumor according to the normalized enrichment score (NES). NFE2L3 was involved in CD22 mediated BCR regulation, FCGR activation, and creation of C4 and C2 activators in most tumors. NFE2L3 was also involved in the role of LAT2 calcium and the role of phospholipids in phagocytosis in BRCA, CHOL, THCA, PRAD, and UCEC. NFE2L3 was shown to be related to immunoregulatory interactions between lymphoid and non-lymphoid cells in CHOL, KIRC, and THCA. NFE2L3 was shown to be associated with FCGR3A mediated IL10 synthesis in KIRC, PRAD, and THCA. These results also demonstrated that NFE2L3 played a key role in tumor immunomodulatory regulation. In addition, NFE2L3 was observed to be involved in vitamin B12 metabolism and in the statin pathway in ESCA. NFE2L3 was related to the ribosome in HNSC and involved in the scavenging of heme from plasma in THCA.

Co-Expression Gene and Functional Enrichment Analysis of NFE2L3 in Liver Hepatocellular Carcinoma

The correlation between NFE2L3 expression and LIHC was examined. The top five co-expression genes that were positively correlated with NFE2L3 expression in LIHC, including ARNT2, OSBPL3, PKM, AMPD3, and NBEAL2, were evaluated. The top five co-expression genes that negatively correlated with NFE2L3 expression in LIHC, including DCXR, RBP4, SLC27A5, ASPDH, and TTC36, were also evaluated (**Figure 11A**). The DEGs were analyzed between the NFE2L3 high and low expression groups in LIHC with threshold values of $|\log2\ FC| > 2.0$, and adjusted p-value < 0.05. We identified 749 DEGs, including 668 upregulated and 81 downregulated genes (**Figure 11B**). Furthermore, GO and KEGG pathway enrichment analyses were performed on the DEGs that met the screening requirements. The results indicated that BP was related to the classical complement activation pathway and humoral immune response mediated by circulating immunoglobulin. The CC consisted of immunoglobulin complexes and circulating immunoglobulin complexes. The MF was primarily enriched in antigen and immunoglobulin receptor binding. The KEGG pathway enrichment was involved in gastric acid secretion and the metabolism of xenobiotics by cytochrome P450 (**Figure 11C**).

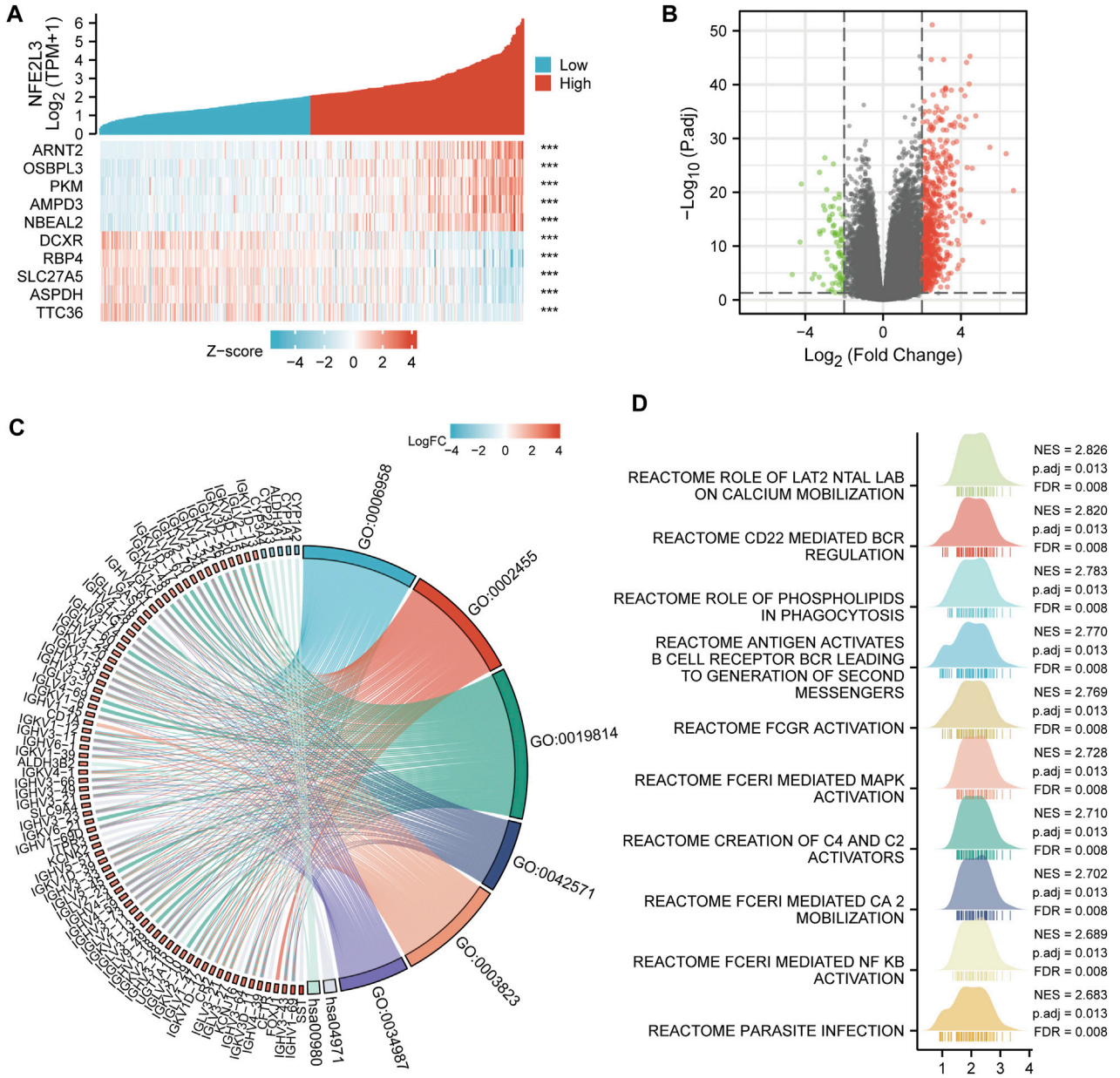


FIGURE 11 | Co-expression gene and functional enrichment analysis of NFE2L3 in LIHC. **(A)** The top five genes positively correlated with NFE2L3 expression, and the top five genes negatively correlated with NFE2L3 expression in LIHC. **(B)** The volcano map of DEGs between NFE2L3 high expression group and low expression groups in LIHC. **(C)** GO and KEGG pathway enrichment analyses of DEGs. **(D)** GSEA merged plots indicated the top 10 significant signaling pathways associated with NFE2L3 expression according to Reactome analyses in LIHC.

In addition, GSEA analysis was performed in LIHC, and the top 10 significant terms of GSEA were mainly involved in immune-related pathways, metabolic functions, and parasite infection (**Figure 11D**).

Prognostic Model Based on NFE2L3 and Clinical Characteristics in Liver Hepatocellular Carcinoma

The relationship between NFE2L3 and clinical characteristics of patients with LIHC was evaluated. The baseline clinical characteristics of the patients in TCGA-LIHC cohort based on the expression levels of NFE2L3 were analyzed (**Supplementary Table S1**). The correlation between NFE2L3 expression levels and the clinical characteristics of the patients were evaluated using logistic regression analysis. The results indicated that high NFE2L3 expression was significantly correlated with sex, age, T stage, pathologic stage, histologic grade, and AFP (**Figure 12A**). Moreover, univariate Cox regression analysis revealed that the OS of patients with LIHC was significantly correlated with T stage,

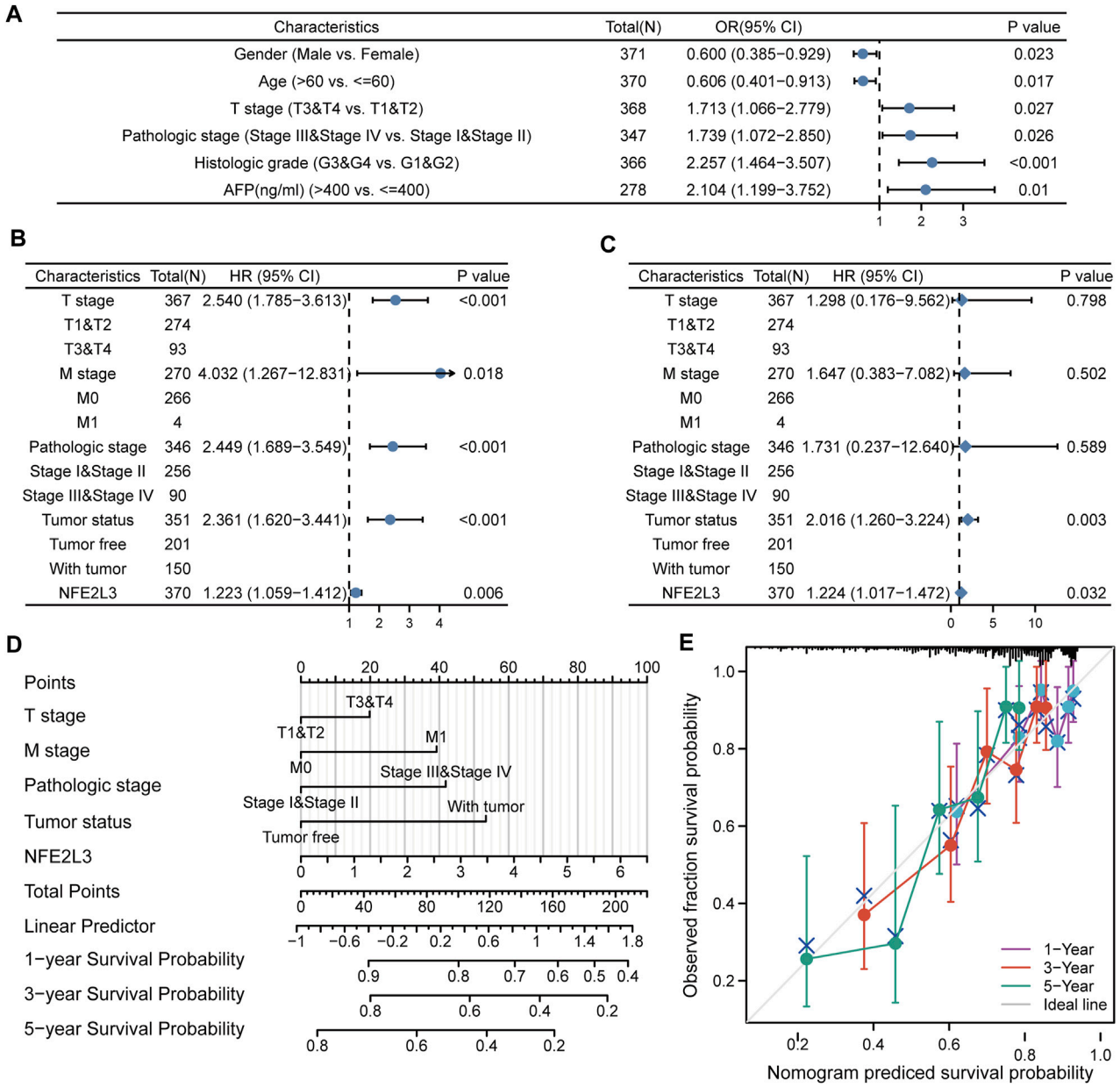


FIGURE 12 | Relationship between the expression level of NFE2L3 and clinical characteristics in LIHC. **(A)** Logistic regression analysis of expression level and clinical characteristics of NFE2L3 in LIHC. **(B,C)** Forest plots of the univariate and multivariate Cox in LIHC. **(D)** A nomogram based on T stage, M stage, pathologic stage, tumor status, and NFE2L3 for predicting the OS of patients with LIHC. **(E)** The calibration plot indicated the calibration of the nomogram model.

M stage, pathological stage, tumor status, and NFE2L3 expression (**Figure 12B**). Multivariate Cox regression analysis further demonstrated that NFE2L3 expression was an independent prognostic factor for OS in LIHC patients (HR = 1.224, 95% CI = 1.017–1.472, $p = 0.032$) (**Figure 12C**). In addition, a nomogram model using T stage, M stage, pathological stage, tumor status, and NFE2L3 expression was constructed to predict the OS of patients in TCGA-LIHC cohort at one, three, and 5 years (**Figure 12D**) The calibration curve demonstrated that the model had good calibration (**Figure 12E**).

DISCUSSION

Cancer is one of the greatest threats to human health worldwide, with a high incidence and mortality rate (Sung et al., 2021). Early diagnosis and treatment of cancer is crucial for patient prognosis.

Pan-cancer analysis can reveal tumor-specific and common molecular signatures, identify novel biomarkers, and provide new insights into the development of effective prevention and therapeutic strategies for human cancers (Chen et al., 2021). NFE2L3 belongs to the Cap n' Collar basic-region leucine zipper family of transcription factors. The NFE2L3 protein is a membrane-bound glycoprotein that targets the endoplasmic reticulum and the nuclear envelope. Recent studies have confirmed that NFE2L3 is linked to certain cancers in humans, including colon cancer, bladder cancer, breast cancer, pancreatic cancer, and liver hepatocellular carcinoma (Wang et al., 2018; Bury et al., 2019; Ren et al., 2020; Dai et al., 2021; Qian et al., 2022). NFE2L3 plays a role in transcription of many biological processes, confers selective growth advantages to cells, and promotes cancer progression, including proliferation, invasiveness, metastasis, and angiogenesis (Kobayashi, 2020).

However, no studies to our knowledge have evaluated the significance of NFE2L3 expression in pan-cancer analyses. The molecular mechanisms of NFE2L3 in cancer was evaluated by performing a comprehensive analysis of NFE2L3 in 33 types of human tumors based on data from the TCGA, CCLE, and HPA databases, including gene expression, epigenetic methylation, genetic alteration, functional enrichment, immune features, and survival prognosis. Our results showed that NFE2L3 was significantly upregulated in most human tumors including BLCA, BRCA, CESC, CHOL, COAD, ESCA, GBM, HNSC, KIRC, KIRP, LIHC, LUAD, LUSC, PRAD, READ, STAD, THCA, and UCEC. Furthermore, NFE2L3 protein was shown to be upregulated in OV, TGCT, LUSC, and CESC. NFE2L3 expression was significantly correlated with different pathological stages in ACC, CESC, KIRC, OV, PAAD, THCA, LIHC, and BRCA. In addition, upregulated expression of NFE2L3 was significantly associated with poor OS, DFS, and DSS in most tumors, including KIRC, KIRP, LGG, LIHC, MESO, PAAD, GBM, and THYM. NFE2L3 also had promising efficacy in tumor diagnosis, especially in the diagnosis of CHOL, COAD, KICH, READ, and STAD. Thus, our study demonstrated that NFE2L3 expression was upregulated in a variety of cancers and was associated with poor prognosis, consistent with previous reports (Wang et al., 2018; Bury et al., 2019; Dai et al., 2021; Wang et al., 2021; Qian et al., 2022). The above results suggest that NFE2L3 expression was closely related to tumor development and that NFE2L3 can be used as a new biomarker for diagnosis and prognosis in most tumors.

DNA methylation is a common epigenetic modification that is closely related to tumorigenesis, regulating gene transcription and expression, and has become the focus in epigenetics (Mehrmohamadi et al., 2016). Hypermethylation in the promoter region of tumor suppressor genes inhibits the expression of tumor suppressor genes and causes them to lose their tumor suppressor functionality, thus promoting the occurrence of cancer. Hypomethylation of the whole genome or of the proto-oncogene promoter, reduces the stability of the chromosome structure or activates proto-oncogenes, thus inducing cell carcinogenesis (Das and Singal, 2004). Our study demonstrated that the DNA methylation levels of NFE2L3 were downregulated in most tumors and negatively correlated with the expression levels of NFE2L3 mRNA and the tumor cell stemness score. The methylation level of NFE2L3 was significantly correlated with different immune subtypes, suggesting that the methylation epistasis regulation of NFE2L3 may be involved in tumor immunity. Moreover, survival analysis showed that hypermethylation of NFE2L3 was associated with good prognosis in TCGA pan-cancer (PANCAN) cohort. Wang et al. (2019) demonstrated that NFE2L3 was a novel DNA methylation driver gene and prognostic marker in human clear cell RCC. In contrast to the study by JW et al., our study was not limited to individual tumors, but explored the expression levels of NFE2L3 at each methylation site and the total methylation levels in pan-cancers. In addition, the correlation between NFE2L3 methylation levels and NE2L3 mRNA expression levels, tumor stemness, immune subtypes, and pan-cancer prognosis was evaluated. NFE2L3 might function as a proto-oncogene, and that downregulation of DNA methylation of NFE2L3 promoted upregulation of NFE2L3 expression in most tumors. The overexpression of NFE2L3 may promote tumorigenesis and progression by regulating tumor cell stemness and immune responses. However, the relationship between NFE2L3 DNA methylation and tumor progression requires further verification.

Previous studies have shown that tumor development is closely related to genetic alterations, especially mutations in oncogenes and tumor suppressor genes (Cohen et al., 2017; Liu et al., 2020b). Our results showed a natural missense variant of Valine441 in membrane localization. The total genetic alteration frequency of NFE2L3 was 2.2% (244/10,950) in TCGA pan-cancer cohort, and the highest alteration frequency of NFE2L3 was approximately 6.4% in patients with UCEC. Mutations and amplifications were the most common types of genetic alterations, with missense mutations being the main type, accounting for 77.7% of the mutations. However, there was little correlation between genetic alterations in NFE2L3 and patient prognosis.

With the continuous development of biomedical technology, tumor immunotherapy has become the fourth most effective treatment method after surgery, chemotherapy, and radiotherapy (Dou et al., 2022; Saliba et al., 2022). The potential of NFE2L3 in tumor immunotherapy was explored further by conducting a correlation analysis between the expression of NFE2L3 and immune characteristics. This study showed that NFE2L3 expression was significantly correlated with various immune characteristics in TCGA pan-cancer, including immune cell infiltration, immune checkpoints, immunosuppressive cell infiltration, immune cell markers, immunomodulators, immune subtypes, and molecular subtypes. In addition, NFE2L3 expression levels were significantly positively correlated with MSI and TBM in STAD. Our study also demonstrated that NFE2L3 expression was associated with patient survival in multiple ICB cohorts, and that knockdown of NFE2L3 improved the effect of ICB treatment for colon

cancer in the Kearney 2018 NK_10 cohort. In conclusion, NFE2L3 shows potential in tumor immunotherapy and may be a new target for tumor immunotherapy.

Recently, single-cell sequencing technology has developed rapidly, and its most important feature is the ability to sequence the genome at the individual cell level compared to traditional sequencing technologies, thus enabling further exploration of patterns of tumor heterogeneity (Suvà and Tirosh, 2019). CancerSEA is a database based on single-cell sequencing and is used to analyze the different functional states of tumor cells at the single-cell level, facilitating further investigation of the functional heterogeneity of tumor cells (Yuan et al., 2019). NFE2L3 was expressed at high levels in most tumor cells according to the CancerSEA analysis and was closely associated with various tumor cell functional states, such as apoptosis, differentiation, inflammation, and stemness. This suggests that NFE2L3 may play an essential role in tumor development.

The molecular function of NFE2L3 in pan-cancer was investigated through gene function enrichment analysis. Our study showed that NFE2L3 was mainly involved in cell cycle regulation by KEGG pathway enrichment analysis, which was consistent with the study performed by Bury et al. (2019). NFE2L3 was observed to be involved in transcriptional regulation and chromatin remodeling of the RNA polymerase II promoter in response to oxidative stress by GO enrichment analysis. Moreover, Saliba et al. (2022) found that knockdown of NFE2L3 could prevent inflammation-induced colorectal cancer by regulating the tumor microenvironment. NFE2L3 was identified to be involved in multiple immune pathways in most tumors using GSEA analysis, including: CD22 mediated BCR regulation, FCGR activation, and the creation of C4 and C2 activators. YC et al. found that NFE2L3 was involved in metastasis and drug resistance in breast cancer. The signaling pathways related to drug resistance and metastasis were absent in our analysis, which may explain the differences in databases and the different thresholds chosen for the analysis.

The function of NFE2L3 was investigated by analyzing the co-expression of genes of NFE2L3 and DEGs between the NFE2L3 high and low expression groups in LIHC. Subsequently, GO and KEGG pathway enrichment analyses and GSEA on these DEGs were conducted and it was shown that they were mainly enriched in immune and metabolic processes. Our study also investigated the relationship between the expression level of NFE2L3 and clinicopathological features using logistic regression analysis and the clinical characteristics related to OS of LIHC using univariate and multivariate Cox regression analyses. NFE2L3 expression level correlated with sex, age, T stage, pathologic stage, histologic grade, and AFP in LIHC, and NFE2L3 expression was an independent risk factor for OS in patients with LIHC, which was consistent with previous studies (Yu et al., 2019; Ren et al., 2020). These results show that NFE2L3 may be a marker for identifying early LIHC and advanced LIHC. Furthermore, a prognostic nomogram using T stage, M stage, pathological stage, tumor status, and NFE2L3 was constructed to predict the one-, three-, and five-year OS in LIHC, which aids physicians in identifying patients at high risk for LIHC.

There are some limitations to the present study. Only online public databases were used for analysis, which may cause systematic bias. The study lacked actual clinical data and further experimental validation using cells and animal models to prove the results observed (Wei et al., 2019).

CONCLUSION

Recently, with the rapid development of bioinformatics, an increasing number of computing methods have been developed and used in tumor research to promote tumor diagnosis and treatment technology (Anashkina et al., 2021; Rogers et al., 2021). To investigate the molecular function of NFE2L3 in pan-cancer, we performed a comprehensive and integrated analysis of NFE2L3 by combining multiple bioinformatic approaches. Our study revealed that NFE2L3 expression was significantly upregulated in multiple tumors and strongly associated with pathological stage and survival prognosis. NFE2L3 expression level of NFE2L3 had high diagnostic efficacy for the majority of tumors. Moreover, the methylation level of NFE2L3 was downregulated in most tumors and significantly associated with poor prognosis. We also analyzed genomic alterations of NFE2L3 in pan-cancer cells. NFE2L3 plays an essential role in immunomodulation in most tumors. In addition, we constructed a nomogram prognostic model for LIHC based on the expression of NFE2L3 and related clinical features, and the results showed that the model had high accuracy in predicting the OS of patients with LIHC. In summary, our study demonstrated the importance of NFE2L3 in the diagnosis and prognosis of pan-cancer, which could be conducive to further exploring its mechanism in tumorigenesis and development and provide a comprehensive analysis basis for cancer treatment 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 in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

ML contributed to the design of the study protocol and manuscript writing. HUW and JY contributed to the data curation and analysis. XC and HAW downloaded the data. YAZ contributed to the study supervision. YW and YOZ contributed to the manuscript revision and review.

FUNDING

This work was supported by the National Key R&D Program of China (2016YFC1302201), the Foundation of The First Hospital of Lanzhou University, China (ldyyyn2021-8 and ldyyyn2021-59).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.916973/full#supplementary-material

Supplementary Figure S1 | The correlation of NFE2L3 expression and disease-specific survival (DSS) in pan-cancer.

Supplementary Figure S2 | The correlation between NFE2L3 expression and immunomodulators in pan-cancer **(A)** The correlation between NFE2L3 expression and immunoinhibitors in pan-cancer. **(B)** The correlation between NFE2L3 expression and immunostimulators in pan-cancer. **(C)** The correlation between NFE2L3 expression and MHC molecules in pan-cancer.

Supplementary Figure S3 | Merged plots for GSEA analysis in multiple tumors. **(A)** BRCA; **(B)** CHOL; **(C)** ESCA; **(D)** HNSC; **(E)** KIRC; **(F)** PRAD; **(G)** UCEC; **(H)** THCA.

Supplementary Table S1 | Baseline clinical data of TCGA-LIHC cohort.

REFERENCES

Anashkina, A. A., Leberfarb, E. Y., and Orlov, Y. L. (2021). Recent Trends in Cancer Genomics and Bioinformatics Tools Development. *Int. J. Mol. Sci.* 22, 12146. doi:10.3390/ijms222212146

Aono, S., Hatanaka, A., Hatanaka, A., Gao, Y., Hippo, Y., Taketo, M. M., et al. (2019). β-Catenin/TCF4 Complex-Mediated Induction of the NRF3 (NFE2L3) Gene in Cancer Cells. *Int. J. Mol. Sci.* 20, 3344. doi:10.3390/ijms20133344

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. A., Kim, S., et al. (2012). The Cancer Cell Line Encyclopedia Enables Predictive Modelling of Anticancer Drug Sensitivity. *Nature* 483, 603–607. doi:10.1038/nature11003

Ben-Yehudah, A., Easley, C. A., Hermann, B. P., Castro, C., Simerly, C., Orwig, K. E., et al. (2010). Systems Biology Discoveries Using Non-human Primate Pluripotent Stem and Germ Cells: Novel Gene and Genomic Imprinting Interactions as Well as Unique Expression Patterns. *Stem Cell. Res. Ther.* 1, 24. doi:10.1186/scrt24

Bonneville, R., Krook, M. A., Kautto, E. A., Miya, J., Wing, M. R., Chen, H. Z., et al. (2017). Landscape of Microsatellite Instability across 39 Cancer Types. *JCO Precis. Oncol.* 2017, PO.17.00073. doi:10.1200/PO.17.00073

Braun, S., Hanselmann, C., Gassmann, M. G., Auf Dem Keller, U., Born-Berclaz, C., Chan, K., et al. (2002). Nrf2 Transcription Factor, a Novel Target of Keratinocyte Growth Factor Action Which Regulates Gene Expression and Inflammation in the Healing Skin Wound. *Mol. Cell. Biol.* 22, 5492–5505. doi:10.1128/mcb.22.15.5492-5505.2002

Bury, M., Le Calvé, B., Lessard, F., Dal Maso, T., Saliba, J., Michiels, C., et al. (2019). NFE2L3 Controls Colon Cancer Cell Growth through Regulation of DUX4, a CDK1 Inhibitor. *Cell. Rep.* 29, 1469–1481.e9. doi:10.1016/j.celrep.2019.09.087

Byrne, J. A., Pedersen, D. A., Clepper, L. L., Nelson, M., Sanger, W. G., Gokhale, S., et al. (2007). Producing Primate Embryonic Stem Cells by Somatic Cell Nuclear Transfer. *Nature* 450, 497–502. doi:10.1038/nature06357

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data: Figure 1. *Cancer Discov.* 2, 401–404. doi:10.1158/2159-8290.cd-12-0095

Chen, F., Fan, Y., Cao, P., Liu, B., Hou, J., Zhang, B., et al. (2021). Pan-Cancer Analysis of the Prognostic and Immunological Role of HSF1: A Potential Target for Survival and Immunotherapy. *Oxid. Med. Cell. Longev.* 2021, 5551036. doi:10.1155/2021/5551036

Chevillard, G., and Blank, V. (2011). NFE2L3 (NRF3): the Cinderella of the Cap'n'Collar Transcription Factors. *Cell. Mol. Life Sci.* 68, 3337–3348. doi:10.1007/s00018-011-0747-x

Cohen, A. J., Saiakhova, A., Corradin, O., Luppino, J. M., Lovrenert, K., Bartels, C. F., et al. (2017). Hotspots of Aberrant Enhancer Activity Punctuate the Colorectal Cancer Epigenome. *Nat. Commun.* 8, 14400. doi:10.1038/ncomms14400

Dai, Y.-C., Pan, Y., Quan, M.-m., Chen, Q., Pan, Y., Ruan, Y.-y., et al. (2021). MicroRNA-1246 Mediates Drug Resistance and Metastasis in Breast Cancer by Targeting NFE2L3. *Front. Oncol.* 11, 677168. doi:10.3389/fonc.2021.677168

Das, P. M., and Singal, R. (2004). DNA Methylation and Cancer. *J. Clin. Oncol.* 22, 4632–4642. doi:10.1200/jco.2004.07.151

Dou, R., Wang, X., and Zhang, J. (2022). Prognostic Value and Immune Infiltration Analysis of Nuclear Factor Erythroid-2 Family Members in Ovarian Cancer. *Biomed. Res. Int.* 2022, 8672258. doi:10.1155/2022/8672258

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal* 6, pl1. doi:10.1126/scisignal.2004088

Goldman, M. J., Craft, B., Hastie, M., Repečka, K., Mcdade, F., Kamath, A., et al. (2020). Visualizing and Interpreting Cancer Genomics Data via the Xena Platform. *Nat. Biotechnol.* 38, 675–678. doi:10.1038/s41587-020-0546-8

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T Cell Dysfunction and Exclusion Predict Cancer Immunotherapy Response. *Nat. Med.* 24, 1550–1558. doi:10.1038/s41591-018-0136-1

Kobayashi, A., Ito, E., Toki, T., Kogame, K., Takahashi, S., Igarashi, K., et al. (1999). Molecular Cloning and Functional Characterization of a New Cap'n' Collar Family Transcription Factor Nrf3. *J. Biol. Chem.* 274, 6443–6452. doi:10.1074/jbc.274.10.6443

Kobayashi, A. (2020). Roles of NRF3 in the Hallmarks of Cancer: Proteasomal Inactivation of Tumor Suppressors. *Cancers (Basel)* 12, 2681. doi:10.3390/cancers12092681

Li, Y., Ge, D., and Lu, C. (2019). The SMART App: an Interactive Web Application for Comprehensive DNA Methylation Analysis and Visualization. *Epigenet. Chromatin* 12, 71. doi:10.1186/s13072-019-0316-3

Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., et al. (2020). TIMER2.0 for Analysis of Tumor-Infiltrating Immune Cells. *Nucleic Acids Res.* 48, W509–w514. doi:10.1093/nar/gkaa407

Liu, B., Fan, Y., Song, Z., Han, B., Meng, Y., Cao, P., et al. (2020a). Identification of DRP1 as a Prognostic Factor Correlated with Immune Infiltration in Breast Cancer. *Int. Immunopharmacol.* 89, 107078. doi:10.1016/j.intimp.2020.107078

Liu, J.-N., Kong, X.-S., Huang, T., Wang, R., Li, W., and Chen, Q.-F. (2020b). Clinical Implications of Aberrant PD-1 and CTLA4 Expression for Cancer Immunity and Prognosis: A Pan-Cancer Study. *Front. Immunol.* 11, 2048. doi:10.3389/fimmu.2020.02048

Love, M. I., Huber, W., and Anders, S. (2014). Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* 15, 550. doi:10.1186/s13059-014-0550-8

Mehrmohamadi, M., Mentch, L. K., Clark, A. G., and Locasale, J. W. (2016). Integrative Modelling of Tumour DNA Methylation Quantifies the Contribution of Metabolism. *Nat. Commun.* 7, 13666. doi:10.1038/ncomms13666

Nouhi, Z., Chevillard, G., Derjuga, A., and Blank, V. (2007). Endoplasmic Reticulum Association and N-Linked Glycosylation of the Human Nrf3 Transcription Factor. *FEBS Lett.* 581, 5401–5406. doi:10.1016/j.febslet.2007.10.041

Omasits, U., Ahrens, C. H., Müller, S., and Wollscheid, B. (2014). Protter: Interactive Protein Feature Visualization and Integration with Experimental Proteomic Data. *Bioinformatics* 30, 884–886. doi:10.1093/bioinformatics/btt607

Qian, J., Huang, C., Zhu, Z., He, Y., Wang, Y., Feng, N., et al. (2022). NFE2L3 Promotes Tumor Progression and Predicts a Poor Prognosis of Bladder Cancer. *Carcinogenesis* 12, bgac006. doi:10.1093/carcin/bgac006

Ren, Y., Wang, Y., Hao, S., Yang, Y., Xiong, W., Qiu, L., et al. (2020). NFE2L3 Promotes Malignant Behavior and EMT of Human Hepatocellular Carcinoma (HepG2) Cells via Wnt/β-catenin Pathway. *J. Cancer* 11, 6939–6949. doi:10.7150/jca.48100

Rogers, M. F., Gaunt, T. R., and Campbell, C. (2021). Prediction of Driver Variants in the Cancer Genome via Machine Learning Methodologies. *Brief. Bioinform.* 22, bbaa250. doi:10.1093/bib/bbaa250

Ru, B., Wong, C. N., Tong, Y., Zhong, J. Y., Zhong, S. S. W., Wu, W. C., et al. (2019). TISIDB: an Integrated Repository Portal for Tumor-Immune System Interactions. *Bioinformatics* 35, 4200–4202. doi:10.1093/bioinformatics/btz210

Saliba, J., Coutaud, B., Makhani, K., Epstein Roth, N., Jackson, J., Park, J. Y., et al. (2022). Loss of NFE2L3 Protects against Inflammation-Induced Colorectal

Cancer through Modulation of the Tumor Microenvironment. *Oncogene* 41, 1563–1575. doi:10.1038/s41388-022-02192-2

Stanley, F. K. T., Berger, N. D., Pearson, D. D., Danforth, J. M., Morrison, H., Johnston, J. E., et al. (2020). A High-Throughput Alpha Particle Irradiation System for Monitoring DNA Damage Repair, Genome Instability and Screening in Human Cell and Yeast Model Systems. *Nucleic Acids Res.* 48, e111. doi:10.1093/nar/gkaa782

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Suvà, M. L., and Tirosh, I. (2019). Single-Cell RNA Sequencing in Cancer: Lessons Learned and Emerging Challenges. *Mol. Cell.* 75, 7–12. doi:10.1016/j.molcel.2019.05.003

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING V11: Protein-Protein Association Networks with Increased Coverage, Supporting Functional Discovery in Genome-wide Experimental Datasets. *Nucleic Acids Res.* 47, D607–d613. doi:10.1093/nar/gky1131

Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: an Enhanced Web Server for Large-Scale Expression Profiling and Interactive Analysis. *Nucleic Acids Res.* 47, W556–w560. doi:10.1093/nar/gkz430

Thul, Pj., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, H., et al. (2017). A Subcellular Map of the Human Proteome. *Science* 356, eaal3321. doi:10.1126/science.aal3321

Wang, H., Zhan, M., Yang, R., Shi, Y., Liu, Q., and Wang, J. (2018). Elevated Expression of NFE2L3 Predicts the Poor Prognosis of Pancreatic Cancer Patients. *Cell. Cycle* 17, 2164–2174. doi:10.1080/15384101.2018.1520558

Wang, J., Zhao, H., Dong, H., Zhu, L., Wang, S., Wang, P., et al. (2019). LAT, HOXD3 and NFE2L3 Identified as Novel DNA Methylation-Driven Genes and Prognostic Markers in Human Clear Cell Renal Cell Carcinoma by Integrative Bioinformatics Approaches. *J. Cancer* 10, 6726–6737. doi:10.7150/jca.35641

Wang, X., Li, Y., Fang, Z., and Li, Y. (2021). Elevated Expression of NFE2L3 Promotes the Development of Gastric Cancer through Epithelial-Mesenchymal Transformation. *Bioengineered* 12, 12204–12214. doi:10.1080/21655979.2021.2005915

Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., et al. (2010). The GeneMANIA Prediction Server: Biological Network Integration for Gene Prioritization and Predicting Gene Function. *Nucleic Acids Res.* 38, W214–W220. doi:10.1093/nar/gkq537

Wei, H., Pu, K., Liu, X. G., Li, B. X., Zhang, H. S., Wang, H., et al. (2019). The Diagnostic Value of Circulating microRNAs as a Biomarker for Gastric Cancer: A Meta-analysis. *Oncol. Rep.* 41, 87–102. doi:10.3892/or.2018.6782

Witschi, H., Malkinson, A. M., and Thompson, J. A. (1989). Metabolism and Pulmonary Toxicity of Butylated Hydroxytoluene (BHT). *Pharmacol. Ther.* 42, 89–113. doi:10.1016/0163-7258(89)90023-5

Yu, M.-M., Feng, Y.-H., Zheng, L., Zhang, J., and Luo, G.-H. (2019). Short Hairpin RNA-Mediated Knockdown of Nuclear Factor Erythroid 2-like 3 Exhibits Tumor-Suppressing Effects in Hepatocellular Carcinoma Cells. *World J. Gastroenterol.* 25, 1210–1223. doi:10.3748/wjg.v25.i10.1210

Yuan, H., Yan, M., Zhang, G., Liu, W., Deng, C., Liao, G., et al. (2019). CancerSEA: a Cancer Single-Cell State Atlas. *Nucleic Acids Res.* 47, D900–d908. doi:10.1093/nar/gky939

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.

Copyright © 2022 Liu, Wei, Yang, Chen, Wang, Zheng, Wang and Zhou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 05 July 2022
doi: 10.3389/fgene.2022.921163

Apoptosis-Related Gene-Mediated Cell Death Pattern Induces Immunosuppression and Immunotherapy Resistance in Gastric Cancer

Xiaolu Yuan[1†], Jun Zhou[2†], Liping Zhou[3], Zudong Huang[2], Weiwei Wang[2], Jiasheng Qiu[2], Qiangbang Yang[2], Chaohao Zhang[1] and MingHui Ma[2]*

[1]Department of Pathology, Maoming People's Hospital, Maoming, China, [2]Department of Gastrointestinal Surgery, Maoming People's Hospital, Maoming, China, [3]Department of Endoscopy Center, The No.6 People's Hospital of Benxi, Liaoning, China

OPEN ACCESS

Edited by:
Ehsan Nazemalhosseini-Mojarad, Shahid Beheshti University of Medical Sciences, Iran

Reviewed by:
Yifang Sun, Guangzhou Red Cross Hospital, China
Zhu Zeng, Guizhou Medical University, China

***Correspondence:**
MingHui Ma
mamingghui@126.com
maminghuizyyx@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: 15 April 2022
Accepted: 23 May 2022
Published: 05 July 2022

Citation:
Yuan X, Zhou J, Zhou L, Huang Z, Wang W, Qiu J, Yang Q, Zhang C and Ma M (2022) Apoptosis-Related Gene-Mediated Cell Death Pattern Induces Immunosuppression and Immunotherapy Resistance in Gastric Cancer. Front. Genet. 13:921163. doi: 10.3389/fgene.2022.921163

Background: Apoptosis is a type of cell death, which can produce abundant mediators to modify the tumor microenvironment. However, relationships between apoptosis, immunosuppression, and immunotherapy resistance of gastric cancer (GC) remain unclear.

Methods: Gene expression data and matching clinical information were extracted from TCGA-STAD, GSE84437, GSE34942, GSE15459, GSE57303, ACRG/GSE62254, GSE29272, GSE26253, and IMvigor210 datasets. A consensus clustering analysis based on six apoptosis-related genes (ARGs) was performed to determine the molecular subtypes, and then an apoptosisScore was constructed based on differentially expressed and prognostic genes between molecular subtypes. Estimate R package was utilized to calculate the tumor microenvironment condition. Kaplan-Meier analysis and ROC curves were performed to further confirm the apoptosisScore efficacy.

Results: Based on six ARGs, two molecular subgroups with significantly distinct survival and immune cell infiltration were identified. Then, an apoptosisScore was built to quantify the apoptosis index of each GC patient. Next, we investigated the correlations between the clinical characteristics and apoptosisScore using logistic regression. Multivariate Cox analysis shows that low apoptosisScore was an independent predictor of poor overall survival in TCGA and ACRG datasets, and was associated with the higher pathological stage. Meanwhile, low apoptosisScore was associated with higher immune cell, higher ESTIMATEScore, higher immuneScore, higher stromalScore, higher immune checkpoint, and lower tumorpurity, which was consistent with the "immunity tidal model theory". Importantly, low apoptosisScore was sensitive to immunotherapy. In addition, GSEA indicated that several gene ontology and Kyoto Encyclopedia of Genes and Genomes items associated with apoptosis, several immune-related pathways, and JAK–STAT signal pathway were considerably enriched in the low apoptosisScore phenotype pathway.

Conclusion: Our findings propose that low apoptosisScore is a prognostic biomarker, correlated with immune infiltrates, and sensitivity to immunotherapy in GC.

Keywords: apoptosis-related genes, tumor microenvironment, gastric cancer, immunotherapy, scoring system

INTRODUCTION

Gastric cancer (GC) is the fifth reason for incidence and fourth in mortality among all cancers (Siegel et al., 2021). Due to delayed diagnosis and less effective treatment in some cases, the overall survival rate is poor (Smyth et al., 2020). PD-1/PD-L1 inhibitors profoundly switched the treatment landscape of GC (Joshi and Badgwell, 2021; Li et al., 2021). The ATTRACTION 2 trial of patients with unresectable advanced or recurrent GC using navuzumab showed a client remission rate of 11.2% and demonstrated extended overall survival (OS) (Satoh et al., 2020). However, while PD-1/PD-L1 inhibitors are a promising treatment for patients with advanced GC, their response rate is limited and the development of new strategies to maximize the efficacy of ICI is necessary (Kono et al., 2020; Kole et al., 2022). Seeking effective biomarkers to screen patients for immunotherapy is essential to optimize the treatment strategy for GC and improve the prognosis of patients.

Apoptosis is an active and procedural death process of cells in the body, which is essential to preserve the homeostasis of the intracellular environment. When cells receive apoptosis signals, initiation caspases are activated through different signal pathways, and effecting Caspases are activated, and then related substrates are degraded, finally leading to cell apoptosis (D'Arcy, 2019). Apoptosis is associated with the occurrence and development of GC (Frejlich et al., 2013). Different drugs have different mechanisms to regulate apoptosis: benzoxanthone compounds regulating the Bcl-2 related protein and the Bcl-2 protein proportion block induced gastric cancer cell apoptosis (Fu et al., 2022), paclitaxel can be through the regulation of p53 gene-mediated apoptosis induced by cell signal transduction pathways (Niapour and Seyedasli, 2022), cisplatin can also through the apoptosis induced by regulating the cell cycle (Guo et al., 2022). Different drugs can copy a cell block in different stages, causing cells to fail to divide properly and die.

Apoptosis also can assist in the regulation of cancer immunity as both the cell itself and environmental factors can affect the apoptosis of GC cells; for example, T cells and NK cells can induce the apoptosis of gastric cancer cells through perforation or Fas pathway (Kume et al., 1999). At the same time, gastric cancer cells can also evade immune monitoring through a variety of mechanisms, resulting in reduced immune quantity, weakened ability of DC cells to present tumor antigen, and inhibited activation of initial T cells (Li et al., 2019). The apoptosis rate of T lymphocytes in GC patients based on Fas/Fasl pathway was significantly increased, suggesting that immune escape in GC cells was related to the up-regulation of Fas ligand expression and t-lymphocyte apoptosis. T cells, NK cells, and DC cells all express Fas receptor after activation, so it is assumed that GC cells may induce apoptosis of the above immune cells through Fas/Fasl pathway by up-regulating Fas ligand (Katoh et al., 2000).

Workflow **Figure 1** illustrates the workflow of our analysis. We first present a comprehensive analysis of the apoptosis-related gene (ARG) expression and prognostic profiling in TCGA-STAD dataset, and found that six differentially expressed and prognostic ARGs were identified. Further, GC samples were classified as two molecular subtypes based on the six ARGs and further validated by KM plotter survival plot. To further elucidate the potential functions of distinct molecular subtypes, we identified 3,083 differentially expressed genes (DEGs) and prognostic genes between the two molecular subtypes and performed functional enrichment analysis on these related genes. We also investigated the correlations between the immune cell infiltration and the two molecular subtypes. Next, GC samples were classified as two molecular subtypes based on the 3,083 genes and further validated by the

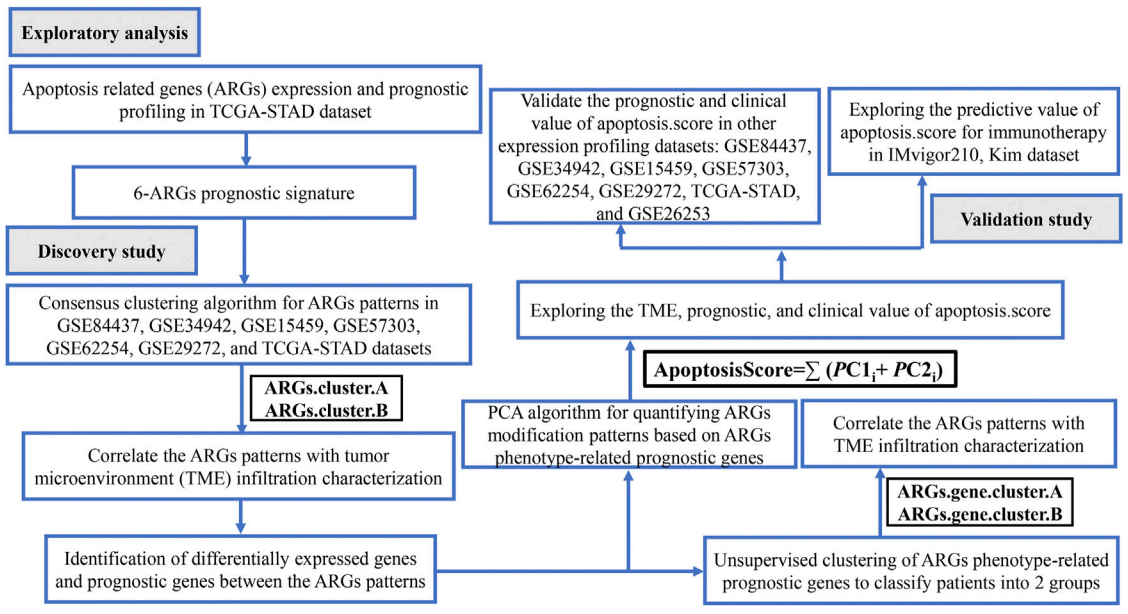


FIGURE 1 | Flow chart of the study.

TABLE 1 | Basic information of series used in this study.

Series Accession Numbers	Platform Used	No. of Input Patients	Region	Survivval outcome	Data Usage	PMID
GSE84437 (GPL6947)	Illumina HumanHT-12 V3.0 expression beadchip	431	Korea	OS	Training; Internal validation	32,293,340
GSE34942 (GPL570)	Affymetrix Human Genome U133 Plus 2.0 Array	56	Singapore	OS	Training; Internal validation	25,053,715
GSE15459 (GPL570)	Affymetrix Human Genome U133 Plus 2.0 Array	191	Singapore	OS	Training; Internal validation	19,798,449
GSE57303 (GPL570)	Affymetrix Human Genome U133 Plus 2.0 Array	70	China	OS	Training; Internal validation	24,935,174
ACRG/GSE62254 (GPL570)	Affymetrix Human Genome U133 Plus 2.0 Array	300	Asia	OS	Training; Internal validation	25,894,828
GSE29272 (GPL96)	Affymetrix Human Genome U133A Array	126	United States	OS	Training; Internal validation	23,717,493
GSE26253 (GPL8432)	Illumina HumanRef-8 WG-DASL v3.0	432	Korea	RFS	External validation	24,598,828
TCGA-STAD	Illumina RNAseq	350	NA	OS	Training; Internal validation	
IMvigor210	Illumina RNAseq	348	NA	OS	External validation	29,443,960
Kim cohort	Illumina RNAseq	45	Korean	NA	External validation	30,013,197

TCGA: the cancer genome atlas; STAD: stomach adenocarcinoma; ACRG: asian cancer research group; OS: overall survival; RFS: relapse free survival.

KM plotter survival plot. Then, an apoptosisScore was constructed based on 3,083 genes. We also examined the correlations between the clinical characteristics and immune cell infiltration and high and low apoptosisScore groups. Finally, we evaluated the accuracy of this novel apoptosisScore and prognostic differences between high- and low apoptosisScore GC patients and analyzed the sensitivity of high- and low-apoptosisScore groups of GC patients to immunotherapy. This novel apoptosisScore model not only accurately predicts the prognosis of GC patients but also offers new insights into the heterogeneity of immunotherapy in GC patients.

MATERIALS AND METHODS

Data Source and Preprocessing

Eight independent GC cohorts were studied in the present study. For the TCGA dataset, mRNA expression files were acquired from the Genomic Data Commons Data Portal (https://portal.gdc.cancer.gov/), and corresponding clinicopathologic data were gained from the cbioportal website (https://www.cbioportal.org/). After removing incomplete data from the survival analysis, data of 350 GC patients complying with the requirements were obtained. Data for the GSE84437 (N = 431) (Yoon et al., 2020), GSE34942 (N = 56) (Chia et al., 2015), GSE15459 (N = 191) (Ooi et al., 2009), GSE57303 (N = 70) (Qian et al., 2014), ACRG/GSE62254 (N = 300) (Cristescu et al., 2015), GSE29272 (N = 126) (Wang et al., 2013), and GSE26253 (N = 432) (Lee et al., 2014) were obtained from Gene expression omnibus (GEO) genomics data repository (https://www.ncbi.nlm.nih.gov/geo/). IMvigor210 CoreBiologies data was downloaded using the R package provided by the following website (https://www.nature.com/articles/nature25501) (Mariathasan et al., 2018). The detail was shown in **Table 1**. R (version 4.0.5) was used to conduct dataset processing and further analysis with R Bioconductor packages.

Expression and Prognostic of Apoptosis-Related Genes

By reviewing the previous literature, we identified 92 ARGs (**Supplementary Table S1**). The expression landscape of ARGs was generated with "limma" and "reshape2" package ($|\log FC| \geq 1.0$ and adj. $p < 0.05$) (Ritchie et al., 2015). The hazard ratio and 95% confidence intervals were calculated by the Cox proportional hazard regression model (Fisher and Lin, 1999).

Consensus Cluster Analysis

We performed an unsupervised cluster analysis of OS samples by ConsensusClusterPlus R package to identify different subtypes. In this process, the number of clusters was set between 2 and 10, and then the samples were classified using consistent clustering (Wilkerson and Hayes, 2010). After screening for the most optimal subtype classification, OS patients were divided into different subtypes and survival analysis was performed on the clustered samples by the Survival R software package, using log-rank test statistics and plotting Kaplan-Meier curves to analyze the survival differences between subtypes. Subsequently, differential genes of different subtypes were screened by limma R package. Differential genes met adj. $p < 0.05$ and $|\log2FoldChange| > 1$.

Construction of apoptosisScore

Next, we used the principal component analysis (PCA) method to quantify apoptosis-related subtypes of individual patients (David and Jacobs, 2014). An apoptosisScore for each patient was calculated according to the following formula:

$$\text{ApoptosisScore} = \Sigma\,(PC1_i + PC2_i),$$

where i is the TPM value of each screened gene.

ESTIMATE Algorithm

The ESTIMATE algorithm, which assessed stromal and immune cells in pernicious tumor tissues using expression data, was used to acquire immune-related score to predict the infiltration of immune cells in GC. The analytical method is contained in the "estimated" R package (Yan et al., 2019). The abundance of immune cells was quantified through the ssGSEA (Bindea et al., 2013; Finotello and Trajanoski, 2018).

Function Enrichment Analysis

We sought to identify the signaling pathways by functional enrichment analysis. The R package "LIMMA" was utilized to sort differentially expressed genes (DEGs) between the two ARGs. cluster.The screened DEGs were examined using the R packages "cluster Profiler", "Rich plot" and "ggplot2" for Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. Gene set enrichment analysis (GSEA) was performed to explore whether the identified sets of genes showed statistical differences between high and low apoptosisScore groups (Subramanian et al., 2005).

Immunohistochemistry

A total of 50 GC patients' tissues and corresponding adjacent tissues were collected to explore the expression of six marker genes in the tissue samples by using immunohistochemical staining (IHC). IHC staining was performed according to the manufacturer's instructions.

Quantitative Reverse Transcription Polymerase Chain Reaction Assays

Total RNA from tissues was isolated using TRIzol (Invitrogen, Canada) reagent, the specific operation is carried out with reference to the instructions for the operation of the kit. RNA (1 μg) was converted into cDNA using the RevertAid First Strand cDNA Synthesis Kit (Takara, China). qRT-PCR was performed using SYBR Green Mixture (Takara, China) in the ABI StepOnePlus System (ABI7500, United States). Target gene expression was normalized against GAPDH.

Statistical Analysis

For the comparison of two groups of continuous variables, the statistical significance of normally distributed variables was estimated by independent Student t-test and differences between non-normally distributed variables were analyzed by Mann-Whitney U-test. Kaplan-Meier survival curves were used to show survival differences, and Log-rank test was used to assess the significance of differences in survival time between two groups of patients. Receiver operating characteristic (ROC) curves were obtained using the pROC R package, and the area under the curve (AUC) was computed to assess the accuracy of the score to estimate prognosis. Heatmap was used by R packages to depict different group information. Additionally, univariate, and multivariate Cox regression analyses were conducted to assess the factors associated with prognosis in GC patients. The analysis was performed using R software of RStudio. Differences were considered significant at $p < 0.05$.

RESULTS

Identification of Differential Expression and Prognostic of Apoptosis-Related Genes in TCGA-STAD

First, GSEA was carried out to determine whether there was a significant apoptosis enrichment between tumor and normal tissues. The results suggested that the apoptosis pathway (Normalize enrichment score = 1.94, adjusted p-value = 0.000) was differentially enriched between tumor and normal tissues (**Figure 2A**). Then, based on the cutoff criterion of $|\log FC| \geq 1.0$ and adj. $p < 0.05$, 63 different expressions of ARGs were identified (**Supplementary Table S2**). To explore the prognostic significance of ARGs in GC, univariate Cox regression analysis was performed on ARGs mRNA expression in GC. The analysis revealed that 17 ARGs correlate with OS (**Supplementary Table S3**). Then, 12 differentially expressed ARGs between tumor and normal tissue that were correlated with OS were identified (**Figure 2B**). Among them, six ARGs were protective genes, and six ARGs were candidate risky genes. Six risky genes (CAPN11, FLT1, FLT4, NOS3, PDGFRB, TGFBR1) were utilized to form the prognostic signature.

As shown in **Figure 2C**, we found that five ARGs with a high frequency of CNV gain were highly expressed in GC patients, suggesting that CNVs may be a potential contributor to the regulation of the expression of ARGs. The positive correlation of six ARGs was illustrated in **Figure 2D**. Furthermore, these ARG expressions were higher in tumor than in normal (**Figure 2E**), and higher in T3 & T4 than in T1 & T2 (**Figure 2F**). Moreover, the AUC of ROC for CAPN11, FLT1, FLT4, PDGFRB, TGFBR1, NOS3 was 0.761, 0.781, 0.708, 0.857, 0.739, and 0.791, respectively, suggesting that they have a good diagnostic performance (**Figure 2G**). As we all know, tumor-infiltrating immune cells are closely connected with the development of cancer. Therefore, we evaluated the correlation between the expression of six ARGs and the enrichment score of T cells. The results showed that higher ARG expression had a higher enrichment score of T cells (**Figures 2H–M**).

Molecular Subtype Categorization of Gastric Cancer Based on Six Apoptosis-Related Genes

After consensus clustering analysis of these samples by the ConsensusClusterPlus algorithm based on six ARGs, it revealed that the results were most stable at K = 2 (**Supplementary Figure S1**). Therefore, samples of the GC patients (N = 1,524) were divided into ARGs.cluster.A (N = 866) and ARGs.cluster.B (N = 658) by this approach. Subsequently, based on the grouped results, we collated the clinical information of the correlated OS samples, including the survival status and the overall survival days, the Kaplan-

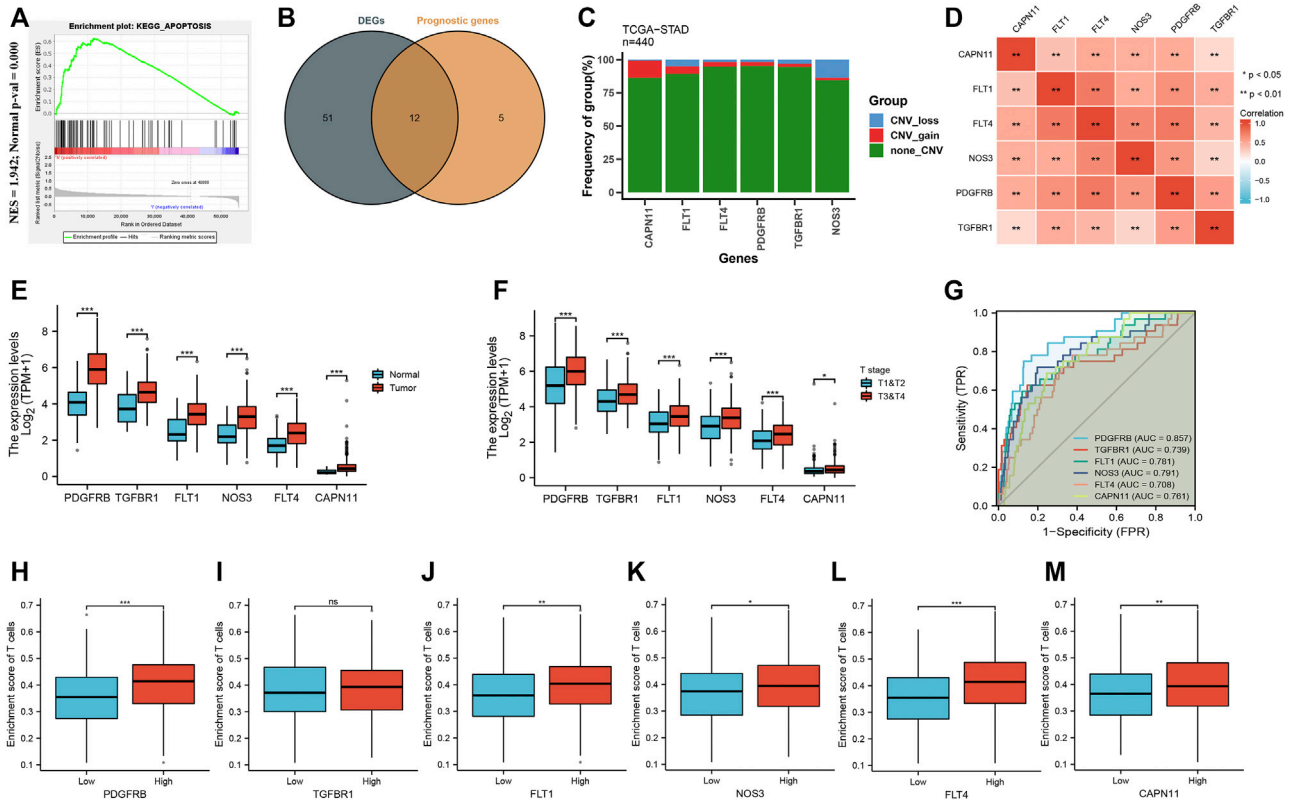


FIGURE 2 | Identification of differential expression and prognostic of ARGs in TCGA-STAD. **(A)** GSEA showed that the apoptosis pathway are differentially enriched in GC patients. **(B)** Venn diagram to identify differentially expressed genes between tumor and normal tissue that were correlated with overall survival. **(C)** Frequencies of CNV gain, loss, and non-CNV among six ARGs. **(D)** Correlation heat map of six ARGs. The size of the colored squares represents the strength of the correlation; red represents a positive correlation. The darker the color is, the stronger correlation is. **(E)** The illustration shows the expression distribution of six ARGs between normal (blue) and GC (red) tissues. **(F)** The illustration shows the expression distribution of six ARGs between T1&T2 (blue) and T3&T4 (red). **(G)** ROC analysis showing that six ARGs has good diagnostic performance. **(H–M)** The correlation between the expression of six ARGs and enrichment score of T cells.

Meier survival analysis was carried out to compare the OS time differences between the different molecular subtypes. It has been shown that patients in the ARGs.cluster.A group have a worse prognosis than the ARGs.cluster.B group (**Figure 3A**). After obtaining the two molecular subtypes based on consensus cluster analysis, ARGs.cluster.A and ARGs.cluster.B denoted a marked discrimination against each other, suggesting there are DEGs between the two groups (**Figure 3B**). Finally, 3,083 DEGs that were correlated with OS were identified (**Supplementary Table S4**). These genes were subjected to GO and KEGG pathway analyses taking advantage of the R package "clusterProfiler, enrichplot, ggplot2". The genes were chiefly enriched in cell death and immunocyte-related bioprocesses such as necroptotic process, apoptosis, immune response, T-cell activation, T-cell proliferation, neutrophil-mediated immunity, neutrophil activation, and tumor necrosis factor superfamily cytokine production (**Figures 3C,D**).

The correlation of tumor microenvironment condition and ARGs.cluster was shown in **Figure 3E**. Using the ssGSEA algorithm, the distribution of immunocyte in patients with GC was plotted in a bar chart, which manifesting that the distribution of various immune cell types varied significantly higher in the ARGs.cluster.A than ARGs.cluster.B (**Figure 3F**). Using the ESTIMATE algorithm, we discovered that score (ESTIMATE, Stromal, and Immune) was notably higher in the ARGs.cluster.A than ARGs.cluster.B. However, tumor purity score was lower in the ARGs.cluster.A than ARGs.cluster.B (**Figures 3G–J**). It revealed that group of ARGs.cluster.A patients were helpful for the tumor immunity response. However, effective antitumor immunity was still suppressed.

Consensus Clustering of 3,083 Genes Identified Two Clusters in Gastric Cancer

After consensus clustering analysis of these samples by ConsensusClusterPlus algorithm based on 3,083 genes, it revealed that the results were most stable at K = 2 (**Supplementary Figure S2**). Therefore, samples of the GC patients (N = 1,524) were divided into ARGs.gene.cluster.A (N = 464) and ARGs.gene.cluster.B (N = 1,060) by this approach. Subsequently, based on the grouped results, the Kaplan-Meier survival analysis was carried out to compare the OS time differences between the different molecular subtypes. It has been demonstrated that patients in the ARGs.gene.cluster.A group have a worse prognosis than the ARGs.gene.cluster.B group (**Figure 4A**).

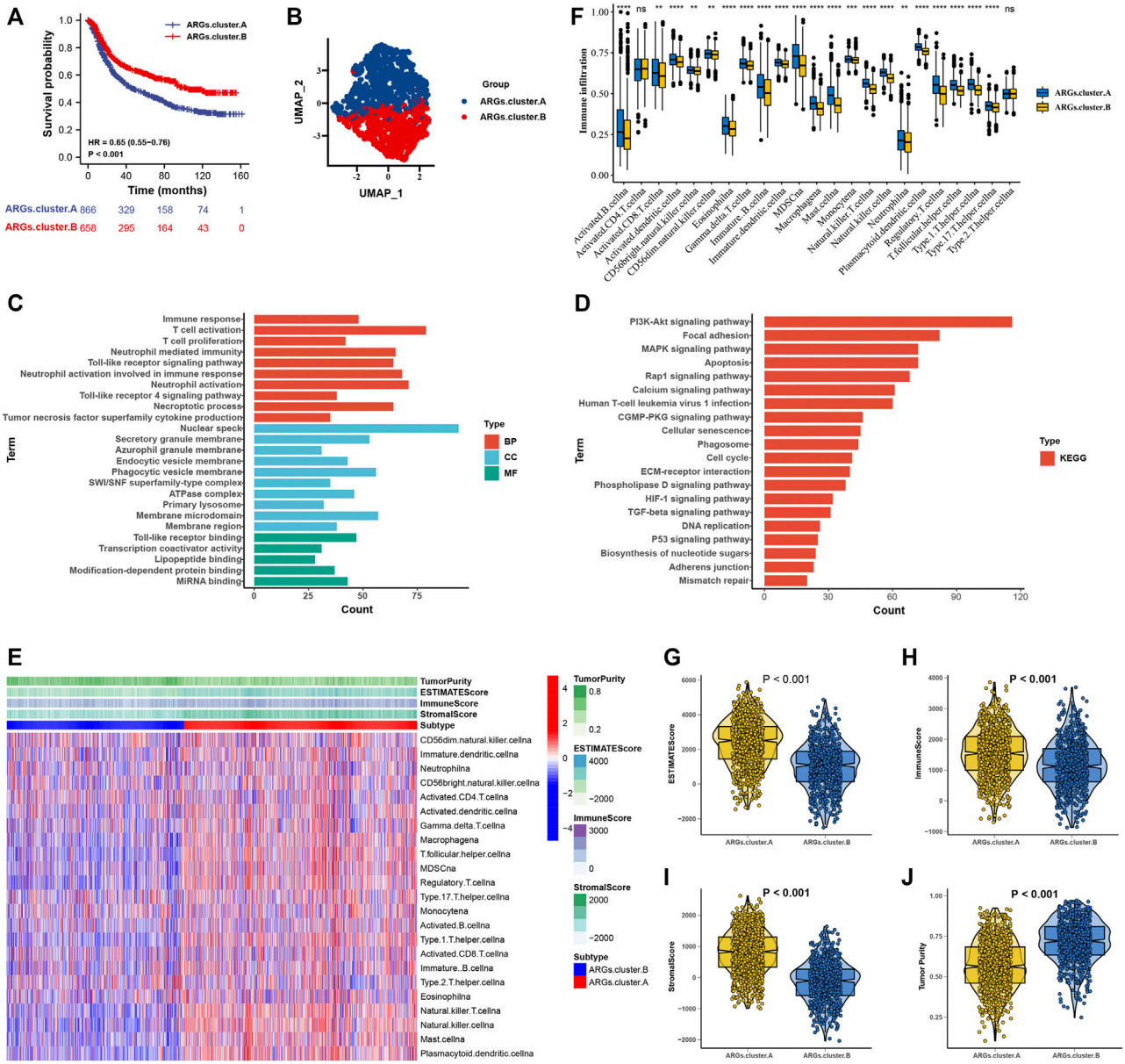


FIGURE 3 | The molecular subtypes categorization of GC base on six ARGs. **(A)** Kaplan-Meier curve showed a significant difference between the two ARGs clusters. **(B)** UMAP analysis for the transcriptome profiles of ARGs.cluster.A and ARGs.cluster.B, showing a remarkable difference on transcriptome between different group. **(C)** GO enrichment analysis **(D)** KEGG enrichment analysis for the different expression that were correlated with OS genes. **(E)** The correlation of tumor microenvironment condition and ARGs.cluster. **(F)** The abundance of each TME infiltrating cell in two ARGs clusters. The upper and lower ends of the boxes represented the interquartile range of values. The lines in the boxes represented median value, and black dots showed outliers. **(G–J)** The ESTIMATE score, stromal score, immune score, and tumor immunity levels in the ARGs.cluster.A and ARGs.cluster.B groups by using ESTIMATE algorithm. (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$).

The correlation of tumor microenvironment condition and ARGs.gene.cluster was shown in **Figure 4B**. Using the ssGSEA algorithm, the distribution of immunocyte in patients with GC was plotted in a bar chart, which manifesting that the distribution of various immune cell types varied significantly higher in the ARGs.gene.cluster.A than ARGs.gene.cluster.B (**Figure 4C**). Using the ESTIMATE algorithm, we discovered that score (ESTIMATE, Stromal, and Immune) was notably higher in the ARGs.gene.cluster.A than ARGs.gene.cluster.B. However, tumor purity score was lower in the ARGs.gene.cluster.A than ARGs.gene.cluster.B (**Figures 4D–G**). It disclosed that group of ARGs.gene.cluster.A patients were helpful for the tumor immunity response. However, the effective antitumor immunity was still suppressed. The above results proved the effectiveness and stability of ARG patterns.

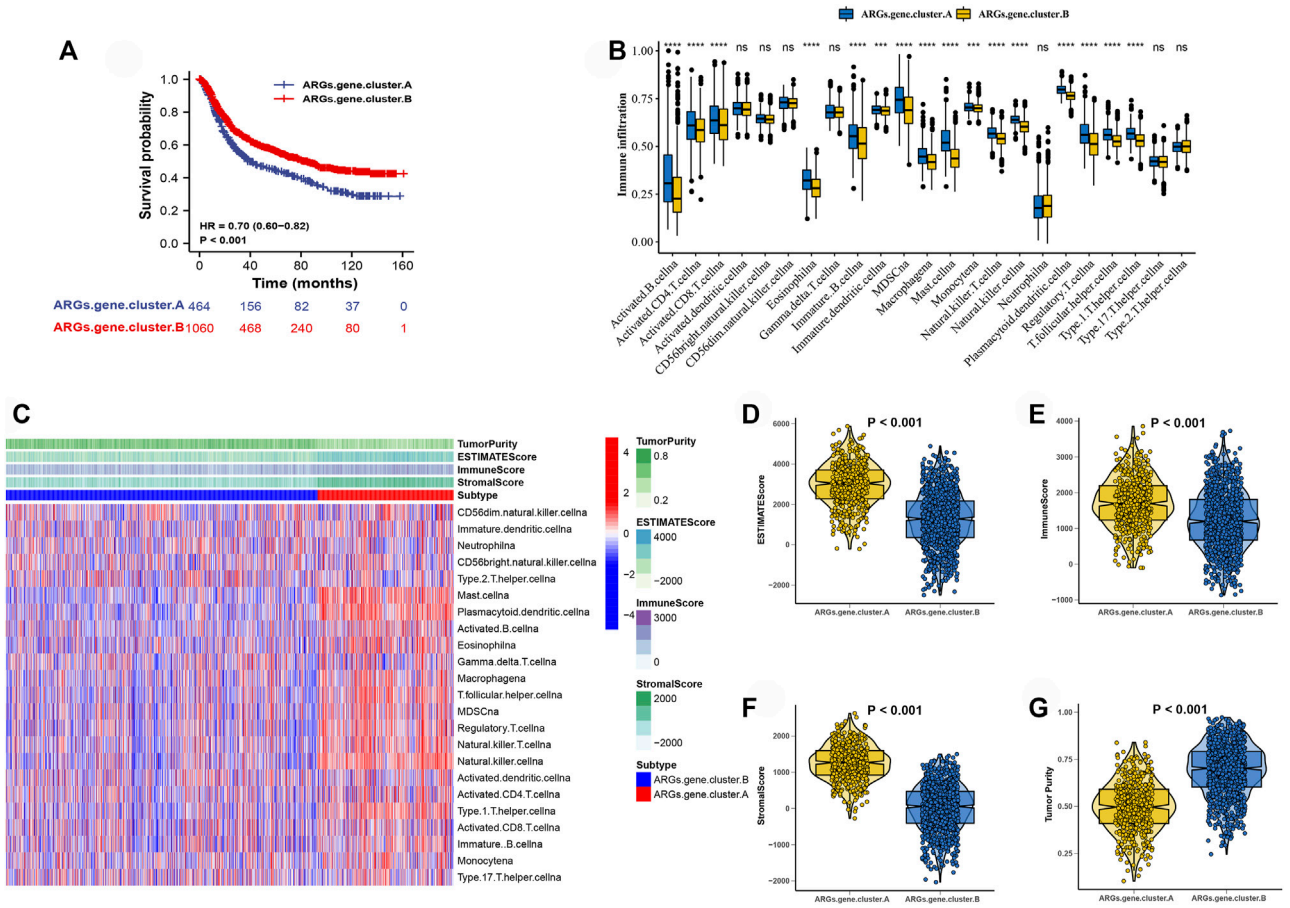


FIGURE 4 | Consensus clustering of 3,083 genes identified two clusters in GC. **(A)** Kaplan-Meier curve showed a significant difference between the two ARGs.gene.clusters. **(B)**The correlation of tumor microenvironment condition and ARGs.gene.clusters. **(C)** The abundance of each TME infiltrating cell in two ARGs.gene.clusters.The upper and lower ends of the boxes represented the interquartile range of values. The lines in the boxes represented median value, and black dots showed outliers. **(D–G)** The ESTIMATE score, stromal score, immune score, and tumor immunity levels in the ARGs.gene.clusters.A and ARGs.gene.clusters.B groups by using ESTIMATE algorithm. (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$).

Establishment of a apoptosisScore

We constructed an apoptosisScore base on 3,083 genes to quantify the apoptosis index of each GC patient by using the PCA method (**Supplementary Table S5**). We discovered that apoptosisScore was notably lower in ARGs.cluster.A and ARGs.gene.cluster.A than in ARGs.cluster.B and ARGs.gene.cluster.B (**Figures 5A,B**). Meanwhile, it has been shown that patients with the low apoptosisScore had a worse prognosis than the high apoptosisScore (**Figure 5C**). Moreover, the AUC of ROC for apoptosisScore was 0.77, suggesting that they have a superior diagnostic performance (**Figure 5D**). The correlation between tumor microenvironment condition and apoptosisScore was illustrated in **Figure 5E**. It indicated that the relative abundance of most of the infiltrating immune cell types increased in the low apoptosisScore group than in high apoptosisScore group (**Supplementary Figure S3**). We also discovered that score (ESTIMATE, Stromal, and Immune) was notably higher in low apoptosisScore group than the high apoptosisScore group. However, the tumor purity score was lower in the low apoptosisScore group than the high apoptosisScore group (**Figures 5–I**). These results indicate that tumor-infiltrating lymphocytes in the low apoptosisScore group were unable to recruit to the tumor site.

Then, the GSEA results suggested that regulation of the immune effector process, regulation of inflammatory response, T-cell activation involved in immune response, apoptosis pathway, natural killer cell-mediated cytotoxicity pathway, and T-cell receptor signaling pathway were differentially enriched in low apoptosisScore phenotypes (**Figures 5J,K**). Function enrichment analysis results strongly indicated that low apoptosisScore is closely related to the tumor immune microenvironment (TIME), and thus worthy of further analysis.

Prognostic Value of apoptosisScore in Gastric Cancer Patients

In the above, we reported that apoptosisScore may have prognostic significance in GC. In order to further explore

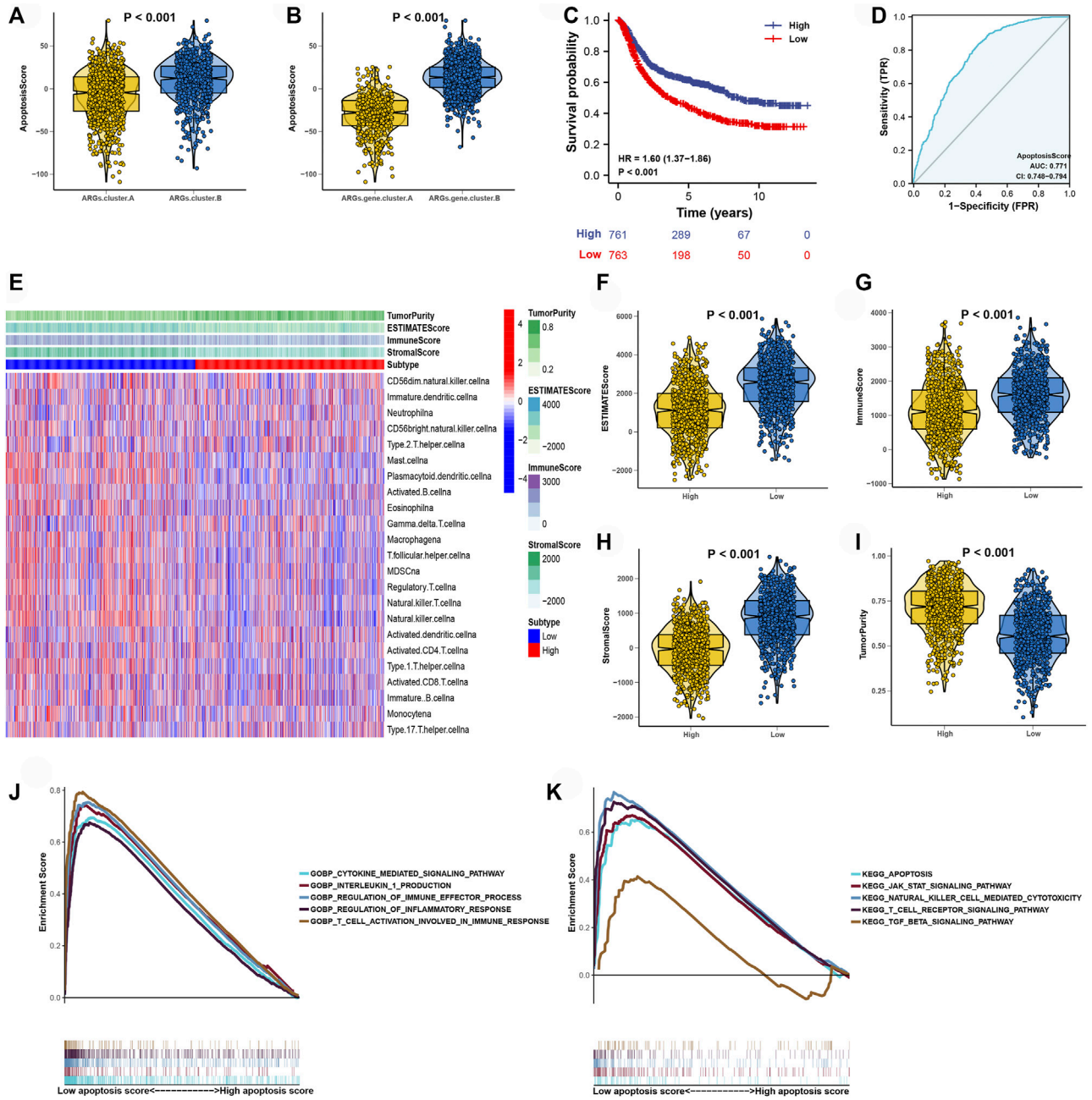


FIGURE 5 | Construction of apoptosisScore. **(A)** Differences in apoptosisScore among two ARGs. **(B)** Differences in apoptosisScore among two ARGs.gene.clusters. **(C)** Kaplan-Meier curves for high and low apoptosisScore groups. **(D)** The predictive value of apoptosisScore. **(E)** The correlation of tumor microenvironment condition in high and low apoptosisScore patient groups. **(F–I)** The ESTIMATE score, stromal score, immune score, and tumor immunity levels in high and low apoptosisScore groups by using ESTIMATE algorithm. **(J)** GSEA GO identified high and low apoptosisScore groups related signaling pathways in GC. **(K)** GSEA KEGG identified high and low apoptosisScore related signaling pathways in GC.

apoptosisScore as a separate prognostic indicator in GC patients. Multivariate analyses showed that low apoptosisScore was a separate prognostic indicator in GC patients (ACRG cohort: HR = 2.33, 95% CI = 1.89–2.86, $p < 0.001$, TCGA cohort: HR = 1.33, 95% CI = 1.09–1.62, $p = 0.004$) (**Figures 6A,B**).

To continue exploring the potential malignant behavior of apoptosisScore, patients with apoptosisScore in the TCGA

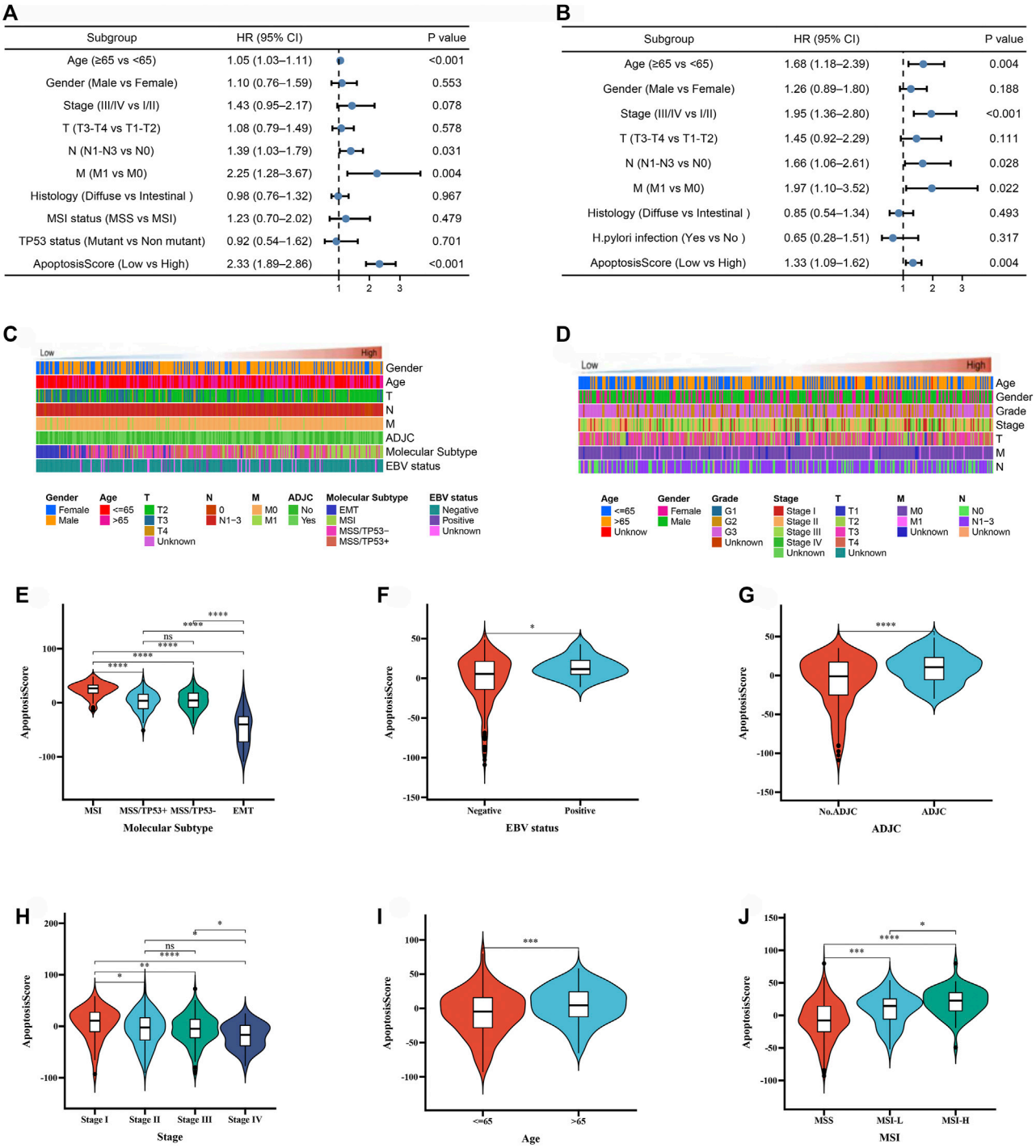


FIGURE 6 | The prognostic value of apoptosisScore in GC patients. **(A)** Multivariate Cox regression analysis for apoptosisScore in ACRG cohort shown by the forest plot. **(B)** Multivariate Cox regression analysis for apoptosisScore in TCGA cohort shown by the forest plot. **(C)** Heatmap showing the dependence between apoptosisScore and clinicopathologic characteristics in ACRG cohort. **(D)** Heatmap showing the dependence between apoptosisScore and clinicopathologic characteristics in TCGA cohort. **(E–G)** Differences in apoptosisScore between molecular subtypes, EBV status, ADJC status in ACRG cohort. **(H–J)** Differences in apoptosisScore between stage, age, MSI status in TCGA cohort.

and ACRG cohorts were analyzed (**Figures 6C,D**). Specifically, we compared the differential level of apoptosisScore in different subgroups stratified by molecular subtypes, EBV status, ADJC status in ACRG cohort, stage, age, MSI status in TCGA cohort. As shown in **Figures 6E–J**, based on the ACRG cohort, apoptosisScore

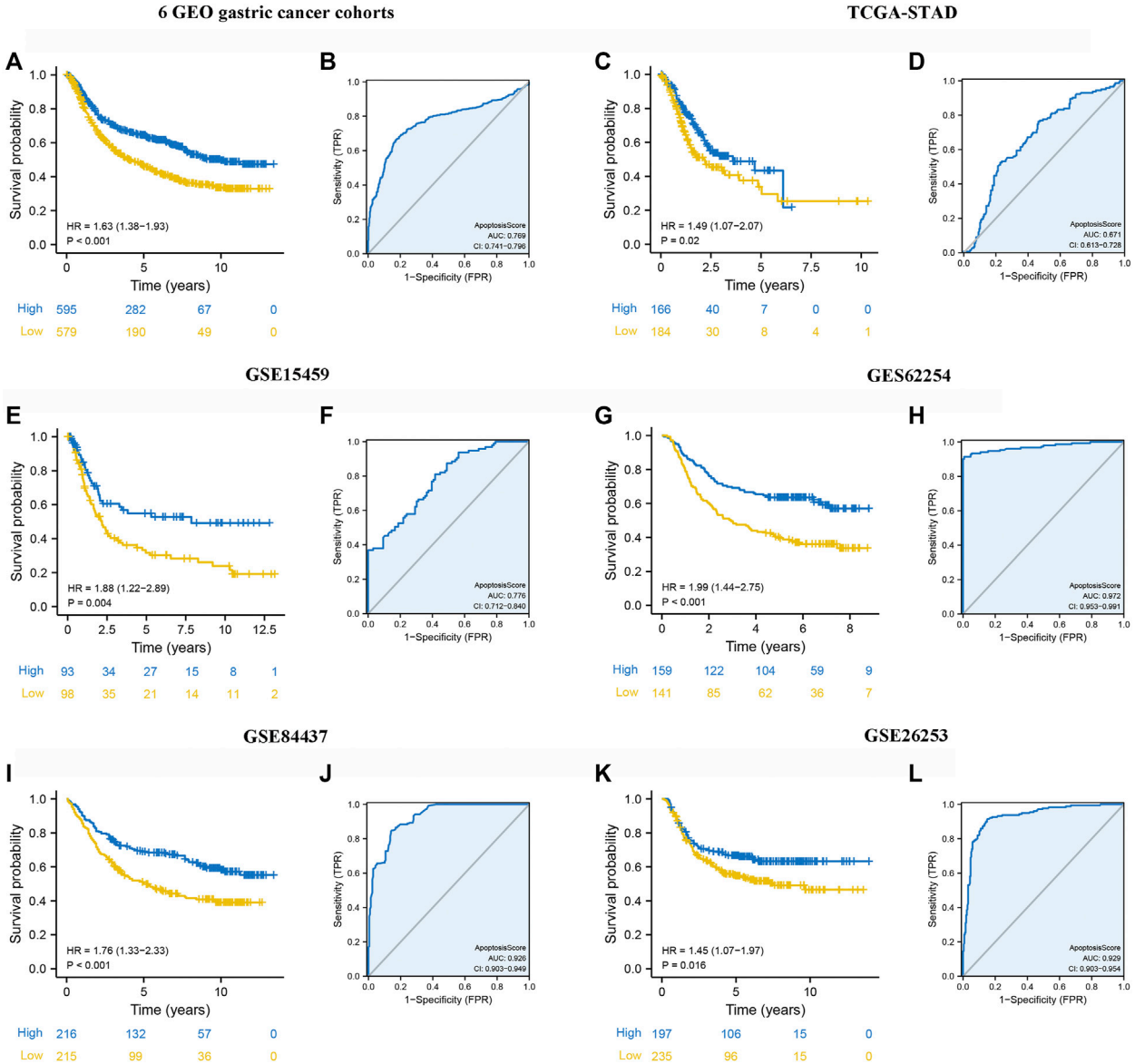


FIGURE 7 | Internal and external validation of apoptosisScore. **(A)** Kaplan-Meier curves for high and low apoptosisScore groups in six GEO datasets (GSE84437, GSE34942, GSE15459, GSE57303, ACRG/GSE62254, and GSE29272). **(B)** The predictive value of apoptosisScore in six GEO dataset. **(C)** Kaplan-Meier curves for high and low apoptosisScore groups in TCGA-STAD. **(D)** The predictive value of apoptosisScore in TCGA-STAD. **(E)** Kaplan-Meier curves for high and low apoptosisScore groups in GSE15459. **(F)** The predictive value of apoptosisScore in GSE15459. **(G)** Kaplan-Meier curves for high and low apoptosisScore groups in GSE62254. **(H)** The predictive value of apoptosisScore in GSE62254. **(I)** Kaplan-Meier curves for high and low apoptosisScore patient groups in GSE84437. **(J)** The predictive value of apoptosisScore in GSE84437. **(K)** Relapse-free survival analysis of apoptosisScore in GSE26253 cohort. **(L)** The predictive value of apoptosisScore in GSE26253.

was significantly different among subgroups stratified by molecular subtypes, EBV status, ADJC status, in ACRG cohort, and stage, age, MSI status in TCGA cohort.

Internal and External Validation of apoptosisScore

Kaplan-Meier survival analysis was conducted to contrast the OS of patients with high and low apoptosisScore and determine whether apoptosisScore can be used as a stable marker for GC patients. The similar above results showed that patients with low apoptosisScore had poorer OS (6 GEO cohorts: HR = 1.63, 95% CI = 1.38–1.93, $p < 0.001$, TCGA-STAD cohort: HR = 1.49, 95% CI = 1.07–2.07, $p = 0.020$, GSE15459 cohort: HR = 1.88, 95% CI = 1.22–2.89, $p = 0.004$, GSE62245 cohort: HR = 1.99, 95% CI = 1.44–2.75, $p < 0.001$, GSE84437 cohort: HR = 1.76, 95% CI = 1.33–2.33, $p = 0.020$) (**Figures 6A,C,E,G,I**). The AUC of ROC for apoptosisScore

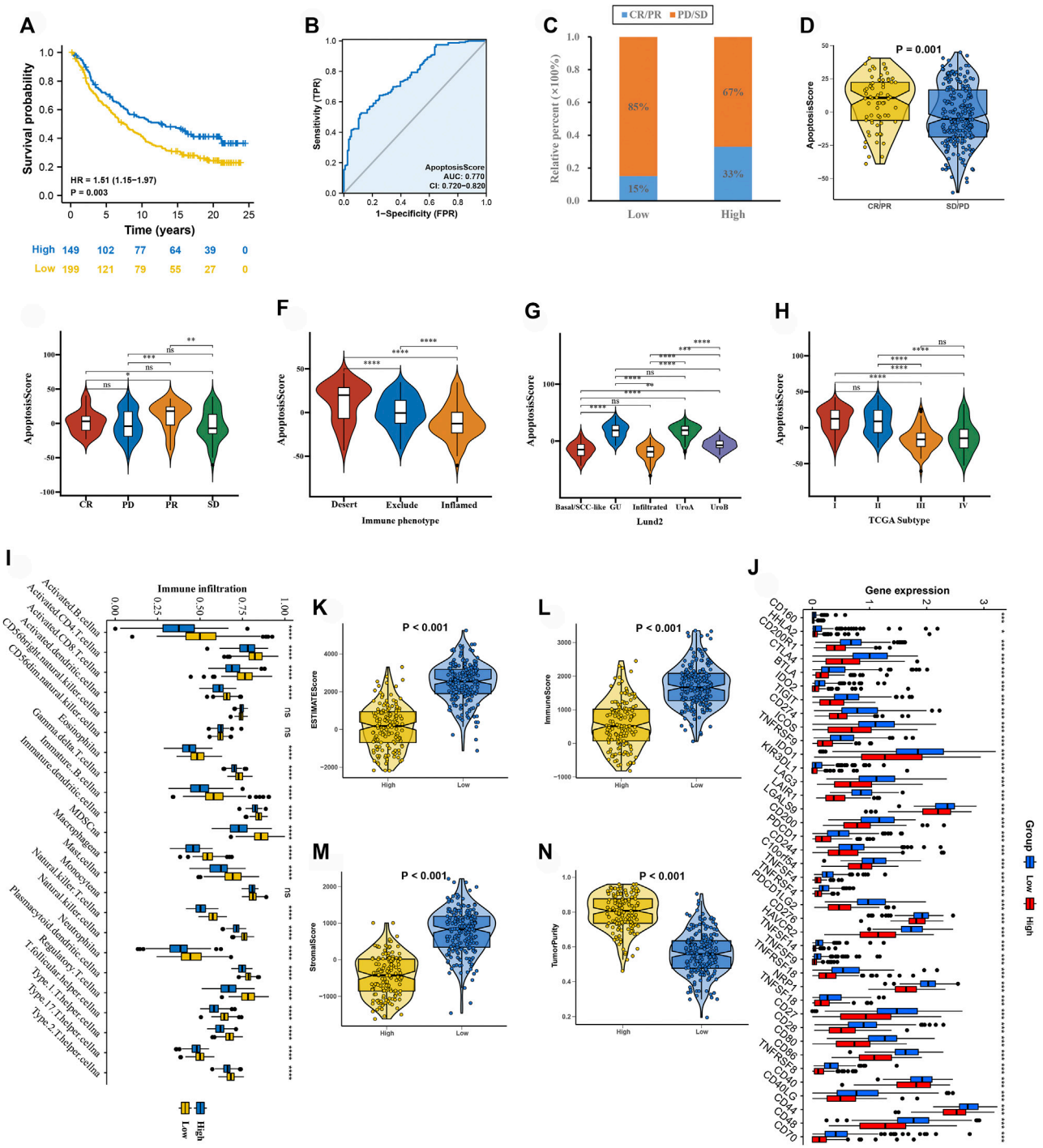


FIGURE 8 | ApoptosisScore in the role of anti-PD-1/L1 immunotherapy in IMvigor210 cohort. **(A)** Survival analyses for low and high apoptosisScore groups in IMvigor210 cohort. **(B)** The predictive value of apoptosisScore in IMvigor210 cohort. **(C)** The proportion of patients with response to PD-L1 blockade immunotherapy in low or high apoptosisScore groups. **(D)** Differences in apoptosisScore among distinct anti-PD-1 clinical response groups. **(E)** Distribution of apoptosisScore in distinct anti-PD-L1 clinical response groups. **(F–H)** Differences in apoptosisScore between immune subtypes, Lund2 subtypes, and TCGA subtypes. **(I)** The abundance of each TME infiltrating cell in high and low apoptosisScore groups. **(J)** Differences in checkpoint expression between low and high apoptosisScore groups. **(K–N)** The ESTIMATE score, stromal score, immune score, and tumor immunity levels in high and low apoptosisScore groups by using ESTIMATE algorithm.

were 0.769 in six GEO cohorts, 0.671 in TCGA-STAD cohort, 0.776 in GSE15459 cohort, 0.972 in GSE62245 cohort, 0.926 in GSE84437 cohort, respectively, suggesting that a prominent fitting prediction (**Figures 7B,D,E,H,J**). In addition, the results also showed that patients with low apoptosisScore had poorer RFS (GSE26253 cohort: HR = 1.45, 95% CI = 1.07–1.97, p = 0.016) (**Figure 7K**). The AUC of ROC for apoptosisScore was 0.829 in GSE26253 cohort, also had a prominent fitting prediction in GSE26253 cohort (**Figure 7L**).

Prognostic Validation of apoptosisScore in the Immune IMvigor210 Cohort

First, after consensus clustering analysis of these samples by ConsensusClusterPlus algorithm based on six ARGs in IMvigor210 cohort, it revealed that the results were most stable at K = 2 (**Supplementary Figure S4**). Then, 1822 DEGs that were correlated with OS were identified (**Supplementary Table S6**). Finally, we constructed an apoptosisScore base on 1822 genes (**Supplementary Table S7**). Based on the IMvigor210 cohort, we performed survival analyses, patients with low apoptosisScore had a worse clinical prognosis for bladder cancer (HR = 1.51, 95% CI = 1.15–1.97, p = 0.003) (**Figure 8A**). The AUC of ROC for apoptosisScore was 0.770 in IMvigor210 cohort, and also had a prominent fitting prediction in IMvigor210 cohort (**Figure 8B**). The proportions of complete response (CR)/partial response (PR) and stable disease (SD)/progressive disease (PD) were 15 and 85% in the low apoptosisScore group and 33 and 67% in the high apoptosisScore group, correspondingly ($p < 0.05$) (**Figure 8C**). ApoptosisScore also varied statistically in the CR/PR and SD/PD groups, apoptosisScore was notably lower in SD/PD group than CR/PR group (**Figure 8D**). We also found that apoptosisScore was notably lower in SD or PD groups than in CR or PR groups (**Figure 8E**). The above results suggest that apoptosisScore was sensitive to immunotherapy. Various histologically and transcriptionally immune tumor subtypes were distinguished, including inflamed, excluded, and desert immune tumors. Various histologically and transcriptionally immune tumor subtypes were distinguished, including inflamed, excluded, and desert immune tumors. ApoptosisScore also varied statistically in the three immune tumor subtypes, five Lund2 subtypes, and three TCGA subtypes, suggesting that apoptosisScore is closely related to the proposed immune subtypes and molecular subtypes (**Figures 8F–H**) (**Table 2**).

Using the ssGSEA algorithm, the distribution of immunocyte in patients with GC was plotted in a bar chart, which manifests that the distribution of various immune cell types varied significantly higher in the low apoptosisScore than in high apoptosisScore (**Figure 8I**). Utilizing the ESTIMATE algorithm, we discovered that score (ESTIMATE, Stromal, and Immune) was notably higher in the low apoptosisScore group than high apoptosisScore group. However, the tumor purity score was lower in the low apoptosisScore group than high apoptosisScore group (**Figures 8K–N**). It revealed that the group of low apoptosisScore group patients was helpful for the tumor immunity response. However, effective antitumor immunity was still suppressed. Based on patients with bladder cancer in the IMvigor210 cohort, the relationship between apoptosisScore and immune checkpoint genes was investigated. We found that immune checkpoint genes were significantly higher in the low apoptosisScore group than in high apoptosisScore group, which was consistent with the "immunity tidal model theory" (**Figure 8J**).

Next, we used an advanced gastric cancer immune cohort (Kim cohort) to detect the relationship between apoptosisScore and immunotherapy. The detail of patient characteristics and apoptosisScore of advanced gastric cancer treated with anti-PD-1 immunotherapy was shown in **Supplementary Table S8**. Specifically, apoptosisScore also varied statistically in the CR, PR, SD, and PD groups, apoptosisScore was notably lower in SD or PD groups than in CR or PR groups (**Figure 9A**). We also

TABLE 2 | Summary of detailed clinical information of IMvigor210 (mUC) cohort.

IMvigor210 Cohort	Low (n = 199)	High (n = 149)	Total (n = 348)
Vital Status			
Alive	54 (27.1%)	62 (41.6%)	116 (33.3%)
Dead	145 (72.9%)	87 (58.4%)	232 (66.7%)
Gender			
Female	41 (20.6%)	35 (23.5%)	76 (21.8%)
Male	158 (79.4%)	114 (76.5%)	272 (78.2%)
Overall response			
CR	14 (8.2%)	11 (8.7%)	25 (8.4%)
PR	12 (7.0%)	31 (24.4%)	43 (14.4%)
SD	40 (23.4%)	23 (18.1%)	63 (21.1%)
PD	105 (61.4%)	62 (48.8%)	167 (56.0%)
Binary response			
CR/PR	26 (15.2%)	42 (33.1%)	68 (22.8%)
SD/PD	145 (84.8%)	85 (66.9%)	230 (77.2%)
Enrollment IC			
IC0	41 (20.6%)	58 (38.9%)	99 (28.4%)
IC1	75 (37.7%)	57 (38.3%)	132 (37.9%)
IC2	83 (41.7%)	34 (22.8%)	117 (33.6%)
IC level			
IC0	38 (19.2%)	59 (39.6%)	97 (28.0%)
IC1	76 (38.4%)	56 (37.6%)	132 (38.0%)
IC2+	84 (42.4%)	34 (22.8%)	118 (34.0%)
TC Level			
TC0	145 (73.2%)	130 (87.2%)	275 (79.3%)
TC1	10 (5.1%)	12 (8.1%)	22 (6.3%)
TC2+	43 (21.7%)	7 (4.7%)	50 (14.4%)
Immune phenotype			
Desert	27 (16.9%)	49 (39.5%)	76 (26.8%)
Excluded	75 (46.9%)	59 (47.6%)	134 (47.2%)
Inflamed	58 (36.3%)	16 (12.9%)	74 (26.1%)
TCGA cluster			
I	44 (22.1%)	74 (49.7%)	118 (33.9%)
II	41 (20.6%)	54 (36.2%)	95 (27.3%)
III	61 (30.7%)	8 (5.4%)	69 (19.8%)
IV	53 (26.6%)	13 (8.7%)	66 (19.0%)
Lund2			
Basal/SCC-like	62 (31.2%)	4 (2.7%)	66 (19.0%)
Genomically unstable	14 (7.0%)	56 (37.6%)	70 (20.1%)
Infiltrated	90 (45.2%)	2 (1.3%)	92 (26.4%)
UroA	17 (8.5%)	85 (57.0%)	102 (29.3%)
UroB	16 (8.0%)	2 (1.3%)	18 (5.2%)

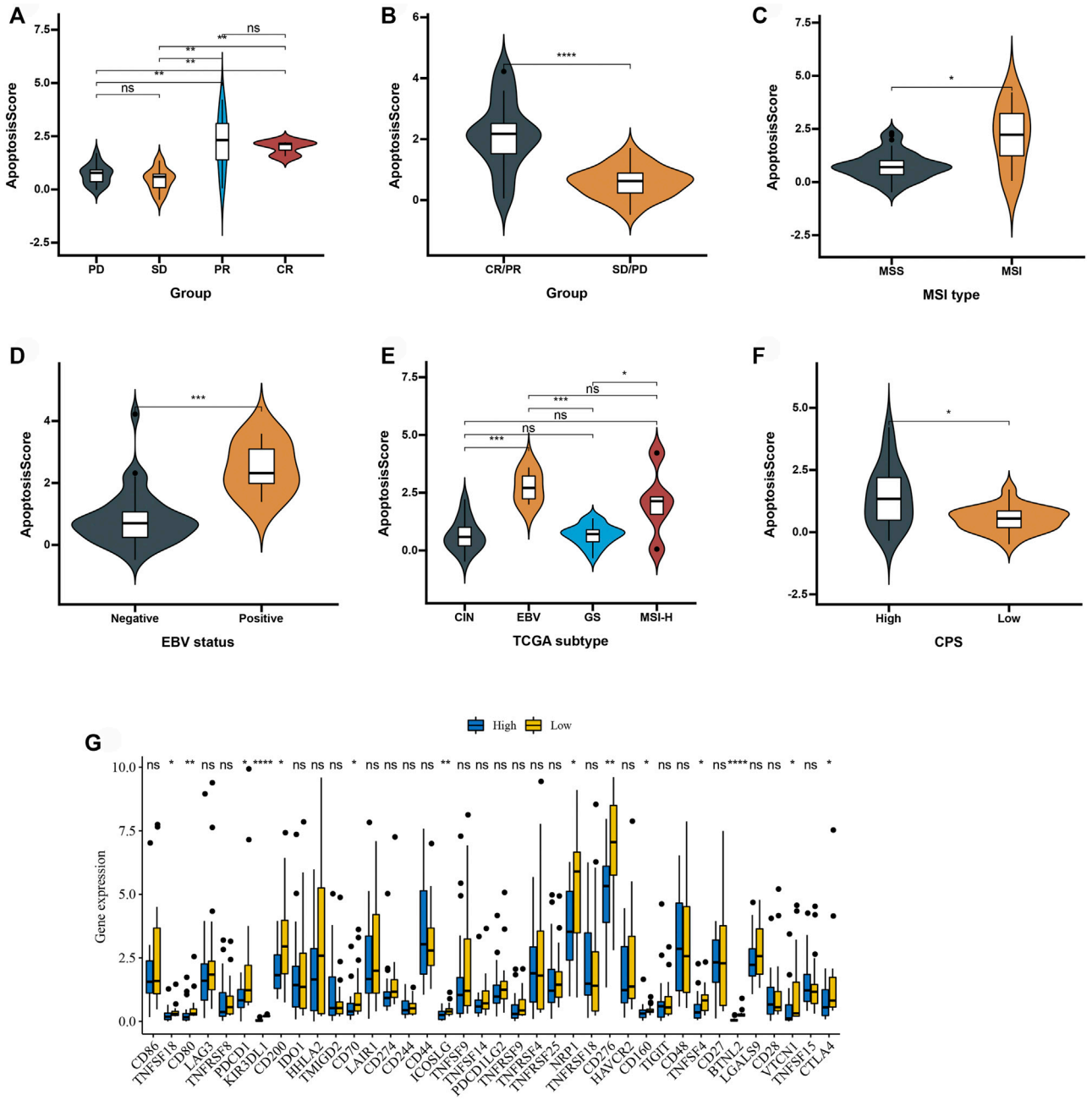


FIGURE 9 | Patient characteristics and apoptosisScore of advanced gastric cancer treated with anti-PD-1 immmunotherapy. **(A)** Distribution of apoptosisScore in distinct anti-PD-L1 clinical response groups. **(B)** Differences in apoptosisScore among distinct anti-PD-1 clinical response groups. **(C)** Differences in apoptosisScore between MSI type. **(D)** Differences in apoptosisScore between EBV status. **(E)** Differences in apoptosisScore between TCGA subtypes. **(F)** Differences in apoptosisScore between high and low CPS score groups. **(G)** Differences in checkpoint expression between low and high apoptosisScore groups in advanced gastric cancer treated with anti-PD-1 immmunotherapy.

discovered that apoptosisScore was notably lower in SD/PD group than in CR/PR group (**Figure 9B**). The above results suggest that apoptosisScore was sensitive to immunotherapy. We also compared the differential level of apoptosisScore in different subgroups stratified by MSI type, EBV status, TCGA subtypes, and CPS score groups. As shown in **Figures 9C–F**, based on the Kim cohort, apoptosisScore was significantly higher in MSI type group, positive EBV status group, EBV group, MSI-H group and high CPS score group. We found that immune checkpoint genes were significantly higher in the low

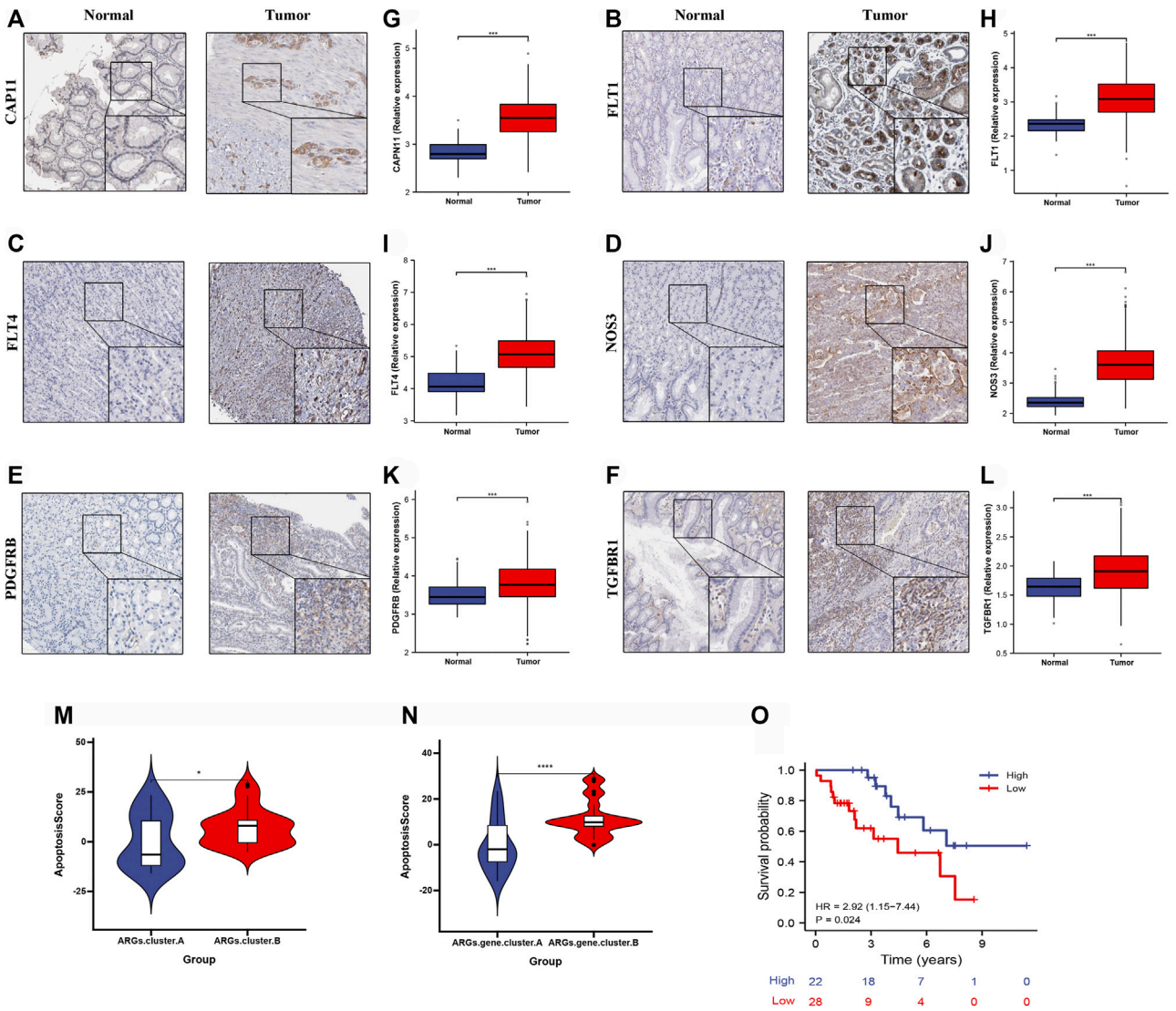


FIGURE 10 | The expression and overall survival of six ARGs in GC. **(A–F)** Representative immunohistochemistry images of six ARGs expression in normal tissues, and GC tissue. **(G–L)** six ARGs mRNA levels are shown for the GC and normal tissue. **(M)** Differences in apoptosisScore among two ARGs. **(N)** Differences in apoptosisScore among two ARGs.gene.clusters. **(O)** Kaplan-Meier analysis of overall survival based on apoptosisScore in 50 cases of GC patients.

apoptosisScore group than in high apoptosisScore group (**Figure 9G**).

Validation the Prognosis of apoptosisScore in Gastric Cancer in an Independent Cohort

To further validate apoptosisScore in GC, we measured the six ARG protein levels by immunohistochemistry, and the result showed that compared with normal group, the CAPN11, FLT1, FLT4, NOS3, PDGFRB, and TGFBR1 levels were significantly higher in GC group (**Figures 10A–F**). In addition, RT-qPCR was used to detect the six marker genes mRNA expression in GC. Compared with the normal group, the six marker genes mRNA level was significantly higher in the GC group (**Figures 10G–L**). We also discovered that apoptosisScore was notably lower in ARGs.cluster.A and ARGs.gene.cluster.A than in ARGs.cluster.B and ARGs.gene.cluster.B (**Figures 10M,N**). Kaplan-Meier analysis revealed that the prognosis of patients with low apoptosisScore was significantly poor than that of patients with high apoptosisScore (**Figure 10O**).

DISCUSSION

In this study, analysis of information from the eight GC cohorts and immune IMvigor210 cohort indicated that low apoptosisScore is related to poor prognosis for GC. We investigated that low apoptosisScore was interrelated with clinicopathologic features such as molecular subtypes, EBV status, ADJC status, stage, age, and MSI status, suggesting that

low apoptosisScore participates in a role in the malignant behavior of GC. Multivariate Cox regression analyses indicated that low apoptosisScore is an independent adverse factor affecting the prognosis of GC patients, which was also confirmed in five additional validated cohorts. Kaplan-Meier survival analysis was utilized to contrast the OS of patients with high and low apoptosisScore, and the results showed that low apoptosisScore could be used as a prognostic indicator for GC patients. ROC curve evaluation suggested that apoptosisScore could be invoked as a useful diagnostic marker. Moreover, apoptosisScore may also participate in a critical part in TIME of GC by regulating the infiltration of immune cells, suggesting that apoptosisScore might be used as a therapeutic target to regulate the anti-tumor immune response.

We discovered that 3,083 DEGs were correlated with OS between the two ARGs.cluster, which is correlated with various immune-related biological processes. Functional enrichment analysis found that these genes in GC patients, which has a significant impact on immune-related biological processes, primarily manifests in immune response, T-cell activation, T-cell proliferation, neutrophil-mediated immunity, Toll-like receptors (TLRs), HIF-1 signaling pathway, P53 signaling pathway, and neutrophil activation. GSEA findings indicated that a low apoptosisScore phenotype is positively correlated with processes related to immune effector process, regulation of inflammatory response, T-cell activation involved in immune response, natural killer cell-mediated cytotoxcity pathway, and T-cell receptor signaling pathway. Moreover, TLRs belong to the pattern recognition receptor superfamily, which typically activates and mediate the pro-inflammatory response of innate immune cells by identifying invading pathogens (Vinnakota et al., 2013). Hypoxia can lead to maladjustment of cell cycle checkpoint by inducing posttranslational modification of wild-type p53, ultimately promoting malignant tumor progression (Cobbs et al., 2003). In particular, antigen-targeting cytotoxicity of T lymphocytes is now identified as a critical factor in the relationship between the immune system and cancer prevention (Waldman et al., 2020). Interestingly, our study indicates that low apoptosisScore is closely interrelated with the above-mentioned immune pathways, and therefore, we believe that low apoptosisScore is closely interrelated with the TIME of GC tissues.

In essence, we identified an interrelation between ARGs.cluster, apoptosisScore, and TIME, in which the behavior of tumor cells determines the outcome of the tumor and affects the biology of TIME cells (Ansell and Vonderheide, 2013). Using the ESTIMATE and ssGSEA algorithm, we discovered that infiltrating immune cell types, ESTIMATE score, stromal score, and the immune score increased in the low apoptosisScore group, and tumor purity decreased in the low apoptosisScore group. The immune score was originally used to assess the stage and prognosis of cancer patients, and patients with a high immune score generally have a better prognosis (Galon et al., 2012). However, Kaplan-Meier survival analysis in the present study found that the clinical outcome of patients with high low apoptosisScore was markedly worse than that of the high apoptosisScore group. Therefore, we hypothesize that apoptosisScore affects the type of immune infiltrating cells in GC. TFHs aid the activity of B cells in germinal center responses and reduce immunosuppression through the inflammatory response and helping to organize tertiary lymphoid structures to achieve anti-tumor effects, which was reported in breast, colorectal and other tumors (Gu-Trantien et al., 2013; Hetta et al., 2020). We found that immune checkpoint genes were significantly higher in low apoptosisScore group than in high apoptosisScore group, which was consistent with the "immunity tidal model theory" that high expression of both conciliatory and contributory immune checkpoints caused an immunosuppressive phenotype in tumors (Zhu et al., 2011). However, the precise mechanism requires further study.

The development of new immunotherapies has advanced rapidly in the field of oncology in recent years, and immune checkpoint inhibition is examined as a potentially important method for the treatment of GC. Mechanisms that suppress the activation and/or effector function of immune cells are called immune checkpoints (Kalbasi and Ribas, 2020; Wang et al., 2020). In this study, we examined the dependence between apoptosisScore and various immune checkpoints. Our result showed that immune checkpoint genes were significantly higher in low apoptosisScore group than in high apoptosisScore group. Indeed, considerable progress has been made in targeting these receptors, such as PDCD1/PDL1 (also known as CD274), CD47 blockade therapies (Barclay and Van den Berg, 2014; Feng et al., 2019). A recent tumor study found that TGFBR1 modified T-cell function by blocking the PD-1/PD-L1 checkpoint, to achieve an anti-tumor effect (Neviani et al., 2019). These studies show that apoptosis participates as a critical part in the PD-1/PD-L1 axis, similar to our results in that low apoptosisScore was positively correlated with PD-1 and PD-L1. It is worth noting that the first generation of immune checkpoint inhibitors, such as inhibitors of CTLA-4 and PD-1, targeted the most distinctive immune checkpoints and thus represents the most mature immunotherapy agents (Marshall and Djamgoz, 2018). Moreover, low apoptosisScore was correlated with other second-generation immune checkpoint genes, including CD40, TNFSF14, BTLA, and HAVCR2. Thus, we considered the possibility of low apoptosisScore as a novel therapeutic target for GC patients, as it could provide an important new basis and direction for immunotherapy in GC patients. As tumor immune resistance is characterized by the co-expression of multiple immune checkpoint pathway molecules, double or multiple checkpoints blocked may produce more powerful anti-tumor immunotherapy effects; therefore, additional immune targets need to be identified.

There are numerous restrictions in the present study. First, this study was inspired by preliminary data and hypothesis-generating predictions. Second, although we found that apoptosisScore was interrelated with patient prognosis and immune invasion in GC, we could not prove that low apoptosisScore affects prognosis through immune invasion, which has to be verified using cell lines. In addition, elucidating the mechanisms by which apoptosisScore regulates the infiltration of immune cells will require further study. The experimental evidence of whether the expression of

apoptosisScore will affect the abundance of immune cell infiltration is a deficiency of our study, and it will become a direction of follow-up research. The apoptosisScore with an AUC of 0.671 in TCGA-STAD cohort, which is lower than PDGFRB, FLT1, NOS3, and higher than CAPN11, FLT4, TGFBR1. However, the AUC of apoptosisScore is 0.972 in ACRG (Asian Cancer Research Group)/GSE62254. Therefore, the prediction accuracy of apoptosisScore is higher among Asian people, the results of our study should be further validated using more multicenter clinical data. However, this subject is new and worthy of further study.

In conclusion, our study suggests that low apoptosisScore is a latent marker for determining the prognosis of GC patients. Low apoptosisScore may also participate a critical part in TIME of GC by egulating the infiltration of immune cells, suggesting that low apoptosisScore might be used as a therapeutic target to regulate the anti-tumor immune response.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

XY and CZ designed the study, performed statistical analysis, and drafted the manuscript. XY helped draft the manuscript. JZ, LZ, ZH, WW, JQ, and QY help to revised manuscript. MM conceived the study, participated in its design and coordination, and helped to draft the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Maoming Science and Technology Plan (grant nos. 2020KJZX013, 2020KJZX004, and 210416164552740). This study was also supported by the High-level Hospital Construction Research Project of Maoming People's Hospital.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.921163/full#supplementary-material

REFERENCES

Ansell, S. M., and Vonderheide, R. H. (2013). Cellular Composition of the Tumor Microenvironment. *Am. Soc. Clin. Oncol. Educ. Book*. doi:10.14694/edbook_am.2013.33.e91

Barclay, A. N., and Van den Berg, T. K. (2014). The Interaction between Signal Regulatory Protein Alpha (SIRPα) and CD47: Structure, Function, and Therapeutic Target. *Annu. Rev. Immunol.* 32, 25–50. doi:10.1146/annurev-immunol-032713-120142

Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A. C., et al. (2013). Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. *Immunity* 39 (4), 782–795. doi:10.1016/j.immuni.2013.10.003

Chia, N.-Y., Deng, N., Das, K., Huang, D., Hu, L., Zhu, Y., et al. (2015). Regulatory Crosstalk between Lineage-Survival oncogenesKLF5, GATA4andGATA6cooperatively Promotes Gastric Cancer Development. *Gut* 64 (5), 707–719. doi:10.1136/gutjnl-2013-306596

Cobbs, C. S., Whisenhunt, T. R., Wesemann, D. R., Harkins, L. E., Van Meir, E. G., and Samanta, M. (2003). Inactivation of Wild-type P53 Protein Function by Reactive Oxygen and Nitrogen Species in Malignant Glioma Cells. *Cancer Res.* 63 (24), 8670–8673.

Cristescu, R., Lee, J., Nebozhyn, M., Kim, K.-M., Ting, J. C., Wong, S. S., et al. (2015). Molecular Analysis of Gastric Cancer Identifies Subtypes Associated with Distinct Clinical Outcomes. *Nat. Med.* 21 (5), 449–456. doi:10.1038/nm.3850

D'Arcy, M. S. (2019). Cell Death: a Review of the Major Forms of Apoptosis, Necrosis and Autophagy. *Cell Biol. Int.* 43 (6), 582–592. doi:10.1002/cbin.11137

David, C. C., and Jacobs, D. J. (2014). Principal Component Analysis: a Method for Determining the Essential Dynamics of Proteins. *Methods Mol. Biol.* 1084, 193–226. doi:10.1007/978-1-62703-658-0_11

Feng, M., Jiang, W., Kim, B. Y. S., Zhang, C. C., Fu, Y.-X., and Weissman, I. L. (2019). Phagocytosis Checkpoints as New Targets for Cancer Immunotherapy. *Nat. Rev. Cancer* 19 (10), 568–586. doi:10.1038/s41568-019-0183-z

Finotello, F., and Trajanoski, Z. (2018). Quantifying Tumor-Infiltrating Immune Cells from Transcriptomics Data. *Cancer Immunol. Immunother.* 67 (7), 1031–1040. doi:10.1007/s00262-018-2150-z

Fisher, L. D., and Lin, D. Y. (1999). Time-dependent Covariates in the Cox Proportional-Hazards Regression Model. *Annu. Rev. Public Health* 20, 145–157. doi:10.1146/annurev.publhealth.20.1.145

Frejlich, E., Rudno-Rudzińska, J., Janiszewski, K., Salomon, L., Kotulski, K., Pelzer, O., et al. (2013). Caspases and Their Role in Gastric Cancer. *Adv. Clin. Exp. Med.* 22 (4), 593–602.

Fu, Y., Xu, Y., Liu, Y., Wang, Y., Chen, J., and Wang, X. (2022). Synthesis, Characterization and Anticancer Efficacy Evaluation of Benzoxanthone Compounds toward Gastric Cancer SGC-7901. *Molecules* 27 (6), 1970. doi:10.3390/molecules27061970

Galon, J., Pagès, F., Marincola, F. M., Angell, H. K., Thurin, M., Lugli, A., et al. (2012). Cancer Classification Using the Immunoscore: a Worldwide Task Force. *J. Transl. Med.* 10, 205. doi:10.1186/1479-5876-10-205

Gu-Trantien, C., Loi, S., Garaud, S., Equeter, C., Libin, M., de Wind, A., et al. (2013). CD4+ Follicular Helper T Cell Infiltration Predicts Breast Cancer Survival. *J. Clin. Invest.* 123 (7), 2873–2892. doi:10.1172/jci67428

Guo, B., Cai, S., Li, W., Guo, C., Liu, Y., Ma, X., et al. (2022). MeCP2 Increases Cisplatin Resistance in Human Gastric Cancer through the Activation of the AKT Pathway by Facilitating PDK-1 Transcription. *Curr. Cancer Drug Targets* [Epub ahead of print]. doi:10.2174/1568009622666220223115216

Hetta, H. F., Elkady, A., Yahia, R., Meshall, A. K., Saad, M. M., Mekky, M. A., et al. (2020). T Follicular Helper and T Follicular Regulatory Cells in Colorectal Cancer: A Complex Interplay. *J. Immunol. Methods* 480, 112753. doi:10.1016/j.jim.2020.112753

Joshi, S. S., and Badgwell, B. D. (2021). Current Treatment and Recent Progress in Gastric Cancer. *CA A Cancer J. Clin.* 71 (3), 264–279. doi:10.3322/caac.21657

Kalbasi, A., and Ribas, A. (2020). Tumour-intrinsic Resistance to Immune Checkpoint Blockade. *Nat. Rev. Immunol.* 20 (1), 25–39. doi:10.1038/s41577-019-0218-4

Katoh, A., Ohshima, K., Kanda, M., Haraoka, S., Sugihara, M., Suzumiya, J., et al. (2000). Gastrointestinal T Cell Lymphoma: Predominant Cytotoxic Phenotypes, Including Alpha/beta, Gamma/delta T Cell and Natural Killer Cells. *Leukemia Lymphoma* 39 (1-2), 97–111. doi:10.3109/10428190009053543

Kole, C., Charalampakis, N., Tsakatikas, S., Kouris, N.-I., Papaxoinis, G., Karamouzis, M. V., et al. (2022). Immunotherapy for Gastric Cancer: a 2021 Update. *Immunotherapy* 14 (1), 41–64. doi:10.2217/imt-2021-0103

Kono, K., Nakajima, S., and Mimura, K. (2020). Current Status of Immune Checkpoint Inhibitors for Gastric Cancer. *Gastric Cancer* 23 (4), 565–578. doi:10.1007/s10120-020-01090-4

Kume, T., Oshima, K., Yamashita, Y., Shirakusa, T., and Kikuchi, M. (1999). Relationship between Fas-Ligand Expression on Carcinoma Cell and Cytotoxic T-Lymphocyte Response in Lymphoepithelioma-like Cancer of the Stomach. *Int. J. Cancer* 84 (4), 339–343. doi:10.1002/(sici)1097-0215(19990820)84:4<339::aid-ijc1>3.0.co;2-2

Lee, J., Sohn, I., Do, I.-G., Kim, K.-M., Park, S. H., Park, J. O., et al. (2014). Nanostring-based Multigene Assay to Predict Recurrence for Gastric Cancer Patients after Surgery. *PLoS One* 9 (3), e90133. doi:10.1371/journal.pone.0090133

Li, F., Sun, Y., Huang, J., Xu, W., Liu, J., and Yuan, Z. (2019). CD4/CD8 + T Cells, DC Subsets, Foxp3, and Ido Expression Are Predictive Indictors of Gastric Cancer Prognosis. *Cancer Med.* 8 (17), 7330–7344. doi:10.1002/cam4.2596

Li, K., Zhang, A., Li, X., Zhang, H., and Zhao, L. (2021). Advances in Clinical Immunotherapy for Gastric Cancer. *Biochimica Biophysica Acta (BBA) - Rev. Cancer* 1876 (2), 188615. doi:10.1016/j.bbcan.2021.188615

Mariathasan, S., Turley, S. J., Nickles, D., Castiglioni, A., Yuen, K., Wang, Y., et al. (2018). TGFβ Attenuates Tumour Response to PD-L1 Blockade by Contributing to Exclusion of T Cells. *Nature* 554 (7693), 544–548. doi:10.1038/nature25501

Marshall, H. T., and Djamgoz, M. B. A. (2018). Immuno-Oncology: Emerging Targets and Combination Therapies. *Front. Oncol.* 8, 315. doi:10.3389/fonc.2018.00315

Neviani, P., Wise, P. M., Murtadha, M., Liu, C. W., Wu, C.-H., Jong, A. Y., et al. (2019). Natural Killer-Derived Exosomal miR-186 Inhibits Neuroblastoma Growth and Immune Escape Mechanisms. *Cancer Res.* 79 (6), 1151–1164. doi:10.1158/0008-5472.Can-18-0779

Niapour, A., and Seyedasli, N. (2022). Acquisition of Paclitaxel Resistance Modulates the Biological Traits of Gastric Cancer AGS Cells and Facilitates Epithelial to Mesenchymal Transition and Angiogenesis. *Schmiedeb. Arch. Pharmacol.* 395 (5), 515–533. doi:10.1007/s00210-022-02217-3

Ooi, C. H., Ivanova, T., Wu, J., Lee, M., Tan, I. B., Tao, J., et al. (2009). Oncogenic Pathway Combinations Predict Clinical Prognosis in Gastric Cancer. *PLoS Genet.* 5 (10), e1000676. doi:10.1371/journal.pgen.1000676

Qian, Z., Zhu, G., Tang, L., Wang, M., Zhang, L., Fu, J., et al. (2014). Whole Genome Gene Copy Number Profiling of Gastric Cancer identifiesPAK1andKRASgene Amplification as Therapy Targets. *Genes. Chromosom. Cancer* 53 (11), 883–894. doi:10.1002/gcc.22196

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Satoh, T., Kang, Y.-K., Chao, Y., Ryu, M.-H., Kato, K., Cheol Chung, H., et al. (2020). Exploratory Subgroup Analysis of Patients with Prior Trastuzumab Use in the ATTRACTION-2 Trial: a Randomized Phase III Clinical Trial Investigating the Efficacy and Safety of Nivolumab in Patients with Advanced Gastric/gastroesophageal Junction Cancer. *Gastric Cancer* 23 (1), 143–153. doi:10.1007/s10120-019-00970-8

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer Statistics, 2021. *CA A Cancer J. Clin.* 71 (1), 7–33. doi:10.3322/caac.21654

Smyth, E. C., Nilsson, M., Grabsch, H. I., van Grieken, N. C., and Lordick, F. (2020). Gastric Cancer. *Lancet* 396 (10251), 635–648. doi:10.1016/s0140-6736(20)31288-5

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene Set Enrichment Analysis: a Knowledge-Based Approach for Interpreting Genome-wide Expression Profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102

Vinnakota, K., Hu, F., Ku, M.-C., Georgieva, P. B., Szulzewsky, F., Pohlmann, A., et al. (2013). Toll-like Receptor 2 Mediates Microglia/brain Macrophage MT1-MMP Expression and Glioma Expansion. *Neuro-Oncology* 15 (11), 1457–1468. doi:10.1093/neuonc/not115

Waldman, A. D., Fritz, J. M., and Lenardo, M. J. (2020). A Guide to Cancer Immunotherapy: from T Cell Basic Science to Clinical Practice. *Nat. Rev. Immunol.* 20 (11), 651–668. doi:10.1038/s41577-020-0306-5

Wang, G., Hu, N., Yang, H. H., Wang, L., Su, H., Wang, C., et al. (2013). Comparison of Global Gene Expression of Gastric Cardia and Noncardia Cancers from a High-Risk Population in china. *PLoS One* 8 (5), e63826. doi:10.1371/journal.pone.0063826

Wang, J.-B., Li, P., Liu, X.-L., Zheng, Q.-L., Ma, Y.-B., Zhao, Y.-J., et al. (2020). An Immune Checkpoint Score System for Prognostic Evaluation and Adjuvant Chemotherapy Selection in Gastric Cancer. *Nat. Commun.* 11 (1), 6352. doi:10.1038/s41467-020-20260-7

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: a Class Discovery Tool with Confidence Assessments and Item Tracking. *Bioinformatics* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170

Yan, H., Qu, J., Cao, W., Liu, Y., Zheng, G., Zhang, E., et al. (2019). Identification of Prognostic Genes in the Acute Myeloid Leukemia Immune Microenvironment Based on TCGA Data Analysis. *Cancer Immunol. Immunother.* 68 (12), 1971–1978. doi:10.1007/s00262-019-02408-7

Yoon, S.-J., Park, J., Shin, Y., Choi, Y., Park, S. W., Kang, S.-G., et al. (2020). Deconvolution of Diffuse Gastric Cancer and the Suppression of CD34 on the BALB/c Nude Mice Model. *BMC Cancer* 20 (1), 314. doi:10.1186/s12885-020-06814-4

Zhu, Y., Yao, S., and Chen, L. (2011). Cell Surface Signaling Molecules in the Control of Immune Responses: a Tide Model. *Immunity* 34 (4), 466–478. doi:10.1016/j.immuni.2011.04.008

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.

Copyright © 2022 Yuan, Zhou, Zhou, Huang, Wang, Qiu, Yang, Zhang and Ma. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 12 July 2022
doi: 10.3389/fgene.2022.942267

A Novel Purine and Uric Metabolism Signature Predicting the Prognosis of Hepatocellular Carcinoma

*Shengjie Yang[1], Baoying Zhang[1], Weijuan Tan[1], Lu Qi[1], Xiao Ma[2] and Xinghe Wang[1]**

[1]Phase I Clinical Trial Center, Beijing Shijitan Hospital, Capital Medical University, Beijing, China, [2]Department of Internal Medicine, Zhangqiu People's Hospital, Jinan, China

Background: Hepatocellular carcinoma (HCC) is regarded as one of the most common cancers in the world with a poor prognosis. Patients with HCC often have abnormal purine and uric acid metabolism, but their relationship with prognosis is unclear.

Methods: Here, we collected the data of peripheral blood uric acid and clinical data in 50 patients with HCC and analyzed the relationship with prognosis. At the same time, the transcriptome sequencing data of TCGA and GEO databases were collected to analyze the changes in purine metabolic pathway activity and construct a prognosis prediction model. Based on the prognosis prediction model related to purine metabolism, we further looked for the differences in the immune microenvironment and molecular level and provided possible drug targets.

Results: We found that the level of serum uric acid was positively correlated with the prognosis of HCC. At the same time, purine metabolism and purine biosynthesis pathway activities were significantly activated in patients with a poor prognosis of HCC. The prognosis prediction model of HCC based on purine metabolism and purine biosynthesis pathway can accurately evaluate the prognosis of patients with HCC. Meanwhile, we found that there were significant changes in tumor immune infiltration microenvironment and biological function at the molecular level in patients with over-activation of purine metabolism and purine biosynthesis pathway. In addition, we found that uric acid level was positively correlated with peripheral blood leukocytes in HCC patients.

Conclusion: In this study, we found that the level of peripheral blood uric acid in patients with HCC is correlated with their prognosis. The prognosis of patients with HCC can be accurately predicted through the metabolic process of uric acid and purine.

Keywords: hepatocellular carcinoma, prognosis, metabolism signature, purine, uric acid

OPEN ACCESS

Edited by:
Ehsan Nazemalhosseini-Mojarad, Shahid Beheshti University of Medical Sciences, Iran

Reviewed by:
Shaocong Mo, Fudan University, China
Guoying Ni, University of the Sunshine Coast, Australia

****Correspondence:***
Xinghe Wang
wangxh@bjsjth.cn

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: *12 May 2022*
Accepted: *22 June 2022*
Published: *12 July 2022*

Citation:
Yang S, Zhang B, Tan W, Qi L, Ma X and Wang X (2022) A Novel Purine and Uric Metabolism Signature Predicting the Prognosis of Hepatocellular Carcinoma. Front. Genet. 13:942267. doi: 10.3389/fgene.2022.942267

INTRODUCTION

Hepatocellular carcinoma (HCC) is the most common histological subtype of primary liver cancer, accounting for about 90% of all cases (Yuan et al., 2020). The malignant rate of HCC is very high, causing about 700,000 deaths every year in the world. It is the fourth leading cause of cancer death in the world (Lee and Lee, 2020). The existing treatment schemes include surgical resection, local

ablation, liver transplantation, chemotherapy, molecular targeted therapy, and immunotherapy (Dai et al., 2022; Kong et al., 2022). However, due to the difficulty of early diagnosis of HCC and the high rate of tumor invasion and recurrence, the overall survival rate and prognosis of patients have been poor, and the average median survival time is about 1 year.

The imbalance of the tumor metabolic pathway is closely related to tumor growth and proliferation (Ke et al., 2022). Tumors need to increase energy consumption to maintain a high metabolic rate, which leads to the change of key metabolic pathways (Vasan and Cantley, 2022). The metabolic activity of tumor cells is significantly different from that of normal cells, and purine metabolism is an important part of material metabolism, which provides necessary components for DNA and RNA (Yuan et al., 2022). The genes of the purine synthesis pathway also have an impact in the field of tumors.

Studies have shown that the metabolic imbalance of purine and uric acid is one of the most significant metabolic changes in cancer, which not only affects the growth of primary tumors but also mediates tumor progression and metastasis (Wu et al., 2021). However, the relationship between purine and uric acid metabolism and the prognosis of patients with HCC has not been fully clarified.

In our study, we collected the data of peripheral blood uric acid in 50 patients with HCC and analyzed the relationship with prognosis. At the same time, the transcriptome sequencing data of TCGA and GEO databases were collected to analyze the changes in purine metabolic pathway activity and construct a prognosis prediction model. Based on the prognosis prediction model related to purine metabolism, we further looked for the differences in the immune microenvironment and molecular level and provided possible drug targets.

METHODS

Data Collection of Hepatocellular Carcinoma Patients

The patient's research was approved by the Beijing Shijitan Hospital and obtained the patient's informed consent. A total of 50 patients with HCC were included in this study. Their gender, age, tumor stage, survival time, survival status, peripheral blood uric acid content, and peripheral blood leukocyte count were collected. See **Supplementary Table S1** for details. The whole research process does not change the routine of clinical diagnosis and treatment, which belongs to observational research.

Survival Analysis of HCC Patients With High Uric Acid and Low Uric Acid

First, according to the median uric acid content of 50 patients with HCC, they were divided into two groups: high UA and low UA, with 25 people in each group. The survival package of R software was used to analyze the KM survival of high uric acid (high UA) and low uric acid (low UA) groups in order to identify whether there are survival differences between high UA and low UA groups.

Collection and Preprocessing of Transcriptome Data and Clinical Data

The transcriptome sequencing data and corresponding clinical information of HCC patients were downloaded from TCGA database (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga). A total of 372 patients were collected as training sets. Two independent HCC patient data sets GSE54236 and GSE27150 were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov) as validation sets. Transcriptome sequencing data and corresponding clinical information were collected.

The Transcriptome Sequencing Data Were Analyzed for 113 Metabolic-Related Pathways and GSVA Function Score

The gene set of 113 metabolic-related pathways comes from published articles (Possemato et al., 2011). Through the R software GSVA package, 113 metabolic pathways of transcriptome sequencing data are scored for GSVA. The specific parameters are as follows: kcdf = "Poisson" and method = "gsva".

Screening Metabolic Pathways Related to the Prognosis of HCC by Univariate Cox Regression

After integrating the GSVA score of 113 metabolic pathways with the prognosis of HCC, univariate Cox regression analysis was carried out by R software. $p < 0.05$ was selected as the statistical difference standard to screen the metabolic pathways related to survival.

Differential Expression Gene Analysis

The differential analysis process was completed by the limma package. The transcriptome sequencing data after standardization were selected. The differential expression genes (DEGs) screening criteria were $|\log2FC \geq 2|$ and $p < 0.05$.

GO and KEGG Functional Enrichment Analysis

GO and KEGG function enrichment analysis is completed by g:Profiler online analysis website, and default parameters are selected for analysis (https://biit.cs.ut.ee/gprofiler/gost). The upregulated DEGs and downregulated DEGs were introduced into the g:Profiler online analysis website, respectively, and $p < 0.05$ was selected as the statistical difference threshold.

Protein Interaction Analysis

PPI analysis is completed by STRING online analysis website (https://cn.string-db.org/), and the default parameters are selected for analysis. The DEGs were led to the STRING online analysis website to generate protein–protein interaction coefficients. The data were downloaded and imported into Cytoscape software to construct a protein–protein interaction network. The key modules

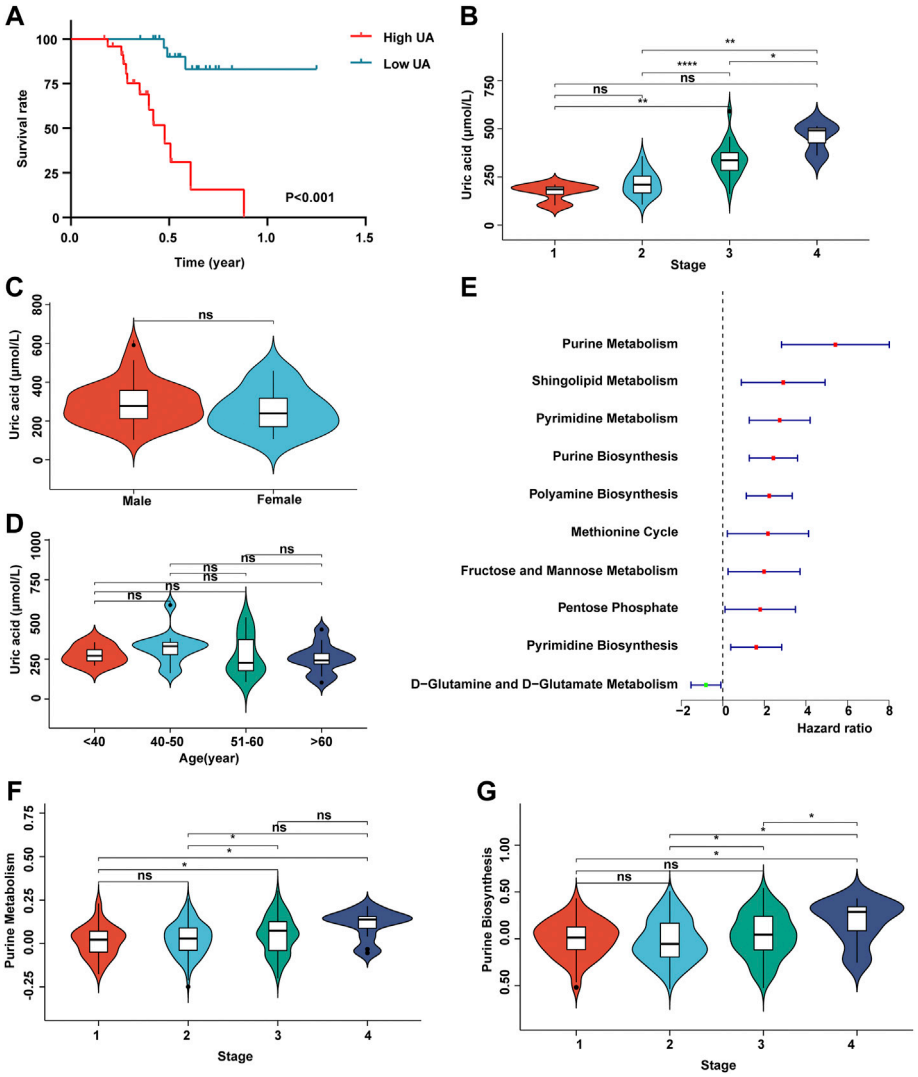


FIGURE 1 | Uric acid was positively correlated with poor prognosis in patients with HCC. **(A)** Demonstrate the relationship between uric acid and survival curve; **(B)** content of uric acid in peripheral blood of patients with different stages of HCC; **(C)** content of uric acid in peripheral blood of patients with HCC in different genders; **(D)** content of uric acid in peripheral blood of patients with HCC at different ages; **(E)** forest map shows the relationship between different metabolic pathways and the prognosis of patients with HCC; **(F)** purine metabolic pathway activity in patients with different stages of HCC; **(G)** purine synthesis pathway activity in patients with different stages of HCC.

are mined through the MCODE plug-in, and the key hub genes are found through the Cytohubber plug-in.

Statistical Analysis

R software (version: 4.1.1) is used for statistical analysis and mapping. The information and survival time of patients were counted and plotted with Kaplan Meier curve. In the 3- and 5-year survival data, the prognostic correlation of genes was evaluated by distinguishing the area under the curve (AUC) generated by the curve. The Cox regression model was used for univariate analysis. An independent sample t-test was applied when only two groups were compared, whereas comparisons among groups were analyzed by two-way ANOVA. For all statistical methods, $p < 0.05$ was considered to be a significant difference.

RESULTS

Uric Acid was Positively Correlated With Poor Prognosis in Patients With Hepatocellular Carcinoma

In order to detect whether uric acid-related metabolism is related to the prognosis of patients with HCC, we recruited 50 patients with HCC and collected clinical data such as uric acid in

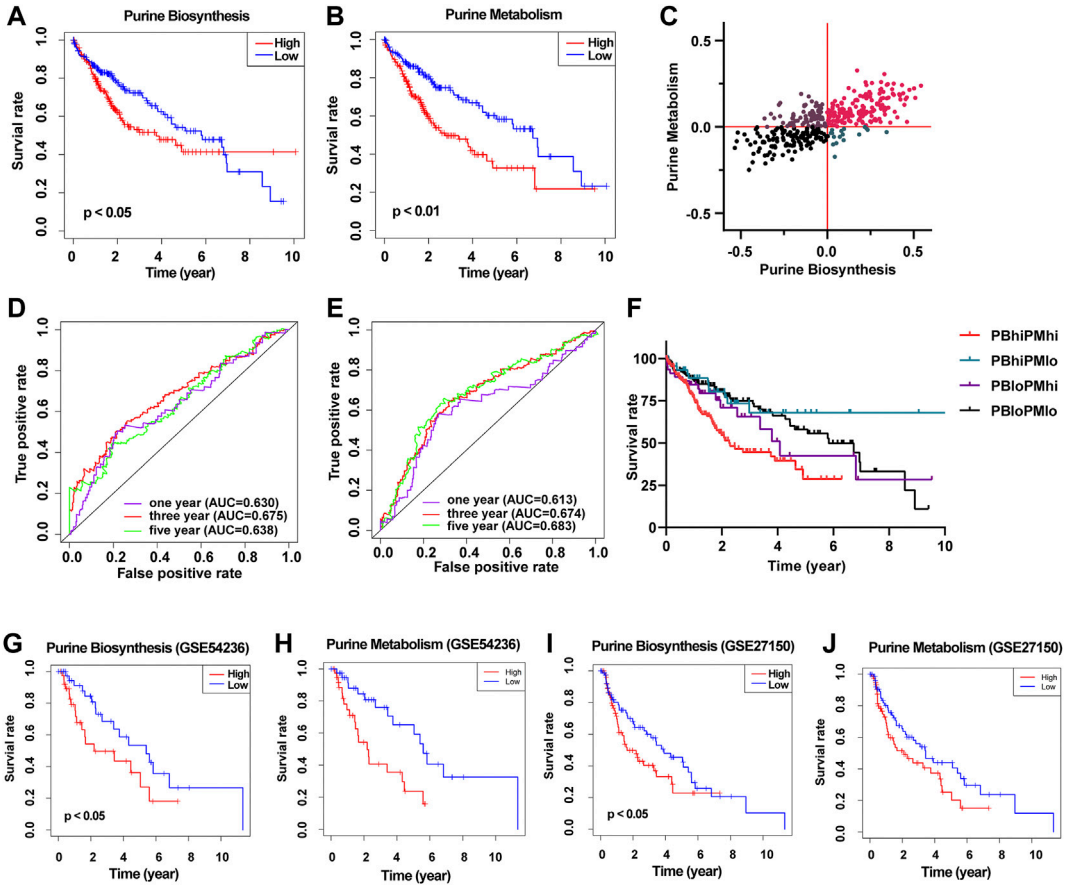


FIGURE 2 | Construct a prognostic model of purine-related metabolic pathway in patients with HCC. **(A)** Survival curve showed the relationship between the activity of purine metabolism pathway and the prognosis of HCC; **(B)** survival curve showed the relationship between purine biosynthesis pathway activity and prognosis of HCC; **(C)** scatter plot shows the purine metabolic pathway activity and purine synthesis pathway activity of each patient with HCC and are divided into four groups according to the median of the two groups; **(D)** ROC curve shows the accuracy of predicting the prognosis of HCC according to the activity of purine metabolic pathway; **(E)** ROC curve shows the accuracy of predicting the prognosis of HCC according to the activity of purine synthesis pathway; **(F)** survival curve shows the prognosis of patients with HCC in different groups; **(G)** survival curve showed the relationship between the activity of purine synthesis pathway and the prognosis of HCC (GSE54236); **(H)** survival curve showed the relationship between purine metabolism pathway activity and prognosis of HCC (GSE54236); **(I)** survival curve showed the relationship between the activity of purine synthesis pathway and the prognosis of HCC (GSE27150); **(J)** survival curve showed the relationship between the activity of purine metabolism pathway and the prognosis of HCC (GSE27150).

peripheral blood, tumor stage, survival time, gender, and age (**Supplementary Table S1**). Survival analysis showed that the higher the content of uric acid in peripheral blood, the worse the prognosis of patients with HCC (**Figure 1A**). Next, we analyzed the expression of uric acid in patients with different stages of HCC. The results shown in **Figure 1B** that the content of uric acid in peripheral blood in high tumor stages (stages 3 and 4) was significantly higher than that in low tumor stages (stages 1 and 2). At the same time, the content of uric acid was not related to gender and age (**Figures 1C,D**).

Subsequently, we collected the transcriptome expression data of 372 patients and their corresponding clinical information (including tumor stage, survival time, gender, and age) from TCGA database. First, we score 113 metabolic pathways through the GSVA algorithm (**Supplementary Table S2**). Then, we carried out a univariate Cox regression analysis of 113 metabolic-related pathways (**Supplementary Table S3**). The results showed that the metabolic pathways related to purine metabolism and purine synthesis were significantly correlated with the poor prognosis of patients with HCC (**Figure 1E**). The enhancement of purine metabolism can produce uric acid accumulation, which is one of the pathogeneses of gout. Meanwhile, purine metabolism and purine synthesis pathways had no correlation with tumor stage (**Figures 1F,G**).

Construct a Prognostic Model of Purine-Related Metabolic Pathway in Patients With HCC

Next, we divided the 371 HCC patients into the high score and low score groups according to the median GSVA score of purine metabolism and purine synthesis pathways. Survival analysis showed that the higher the GSVA score of purine

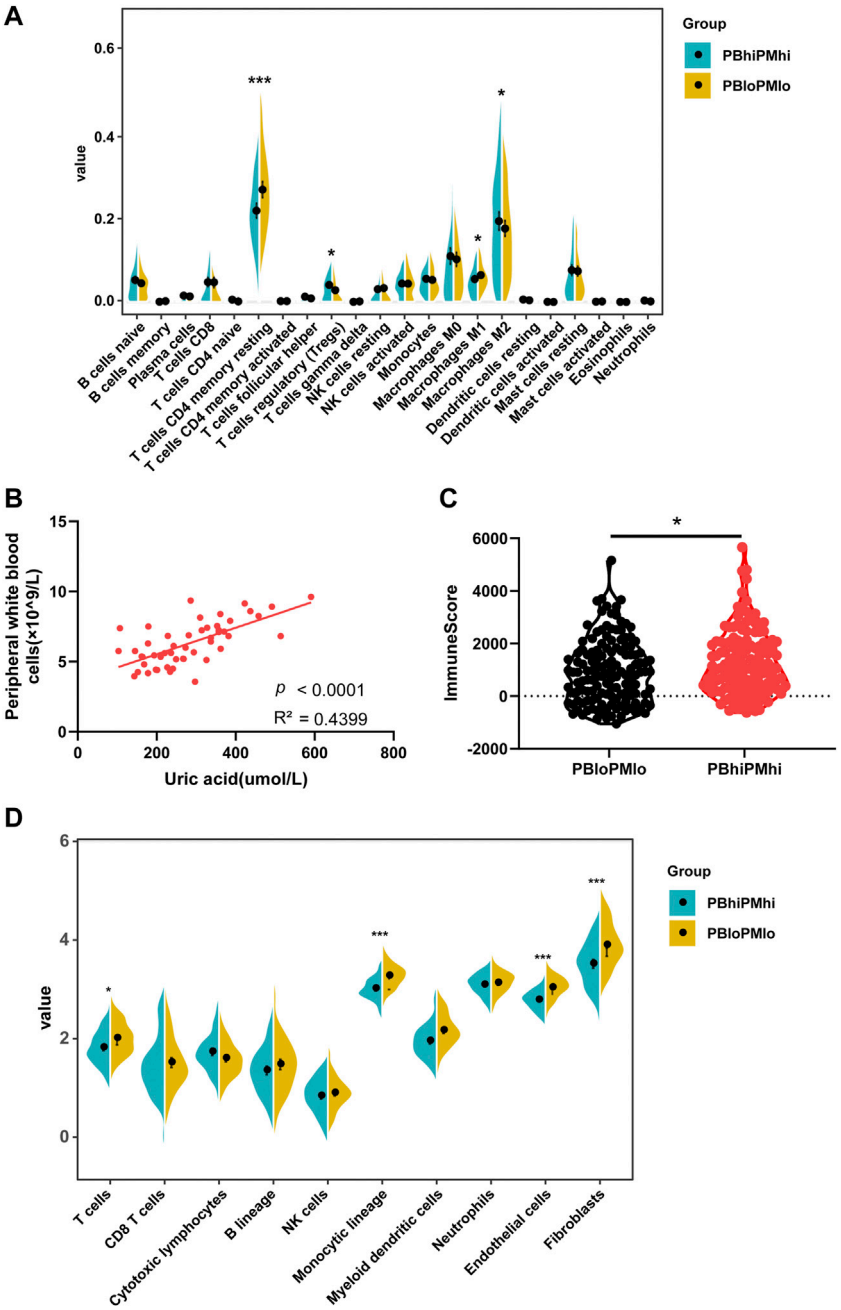


FIGURE 3 | Purine metabolism affects immune infiltration microenvironment in patients with HCC. **(A)** Violin diagram shows the proportion of 22 immune cells in HCC patients in the PBhiPMhi group and PBloPMlo group. **(B)** Correlation between uric acid level and peripheral blood leukocytes in HCC patients. **(C–E)** Estimate and MCP count algorithms to verify the immune infiltration between PBloPMlo and PBhiPMhi groups.

metabolism and purine synthesis pathways, the worse the prognosis of patients with HCC (**Figures 2A,B**). At the same time, the ROC curve shows the accuracy of the prediction effect. The AUC values of 1, 3, and 5 years are greater than 0.6, showing good accuracy and specificity (**Figures 2D,E**). At the same time, we collected two independent data sets from the GEO database as verification sets, and the results are consistent with TCGA database (**Figures 2F–J**). In order to reduce the bias in the selection of purine metabolism gene sets, we collected another 11 purine metabolism-related gene sets from the GO database, KEGG database, Reactome database, and WikiPathways for verification. The results in 73% (8/11) of the gene sets were consistent with the previous results, and urine metabolism is a poor prognostic factor for HCC (**Supplementary Table S4**).—→

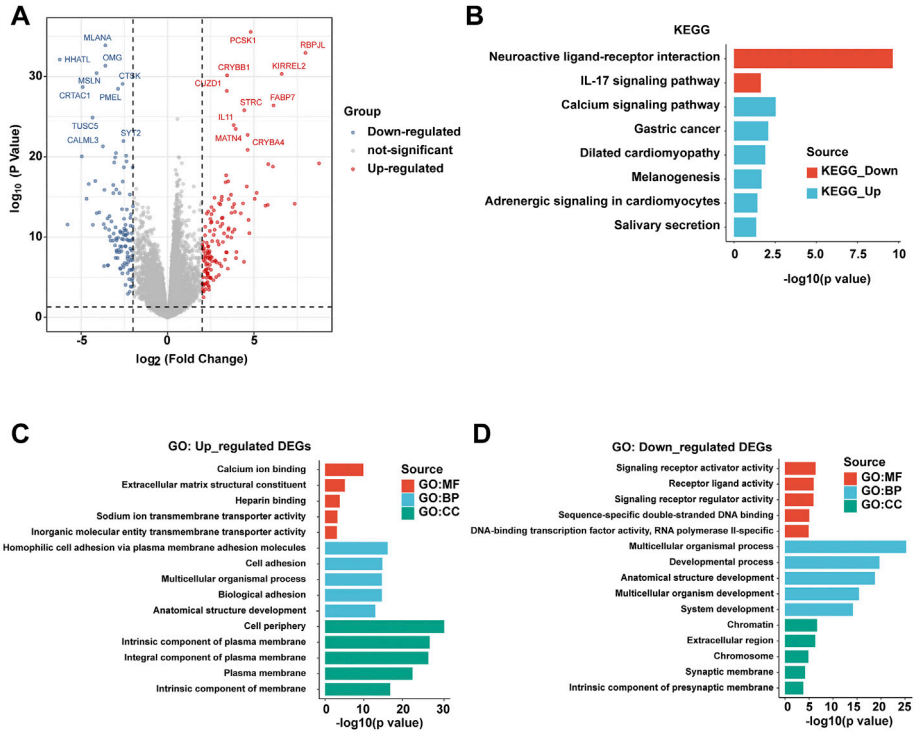


FIGURE 4 | Analysis of differential gene expressions (DEGs) between PBhiPMhi and PBloPMlo groups. **(A)** Volcanic map shows the DEG distribution of HCC patients in the PBhiPMhi group and PBloPMlo group. **(B)** KEGG functional enrichment analysis of DEGs in patients with HCC; **(C,D)** GO functional enrichment analysis of DEGs in patients with HCC.

Then, according to the GSVA score of purine metabolism and purine synthesis pathways, we further divided HCC patients into four groups: high purine biosynthesis and high purine metabolism (PBhiPMhi), high purine biosynthesis and low purine metabolism (PBhiPMlo), low purine biosynthesis and high purine metabolism (PBloPMhi), and low purine biosynthesis and low purine metabolism (PBloPMlo) (**Figure 2C**). Survival analysis showed that the prognosis of PBhiPMhi was the worst, and the prognosis of the PBloPMlo group was the best, while the other two groups were in between (**Figure 2F**). These results suggest that purine metabolism and synthesis may be good predictors of prognosis.

Purine Metabolism Affects Immune Infiltration Microenvironment in Patients With HCC

The immune microenvironment is related to the prognosis of a variety of tumors. In order to explore whether purine-related metabolism can affect the tumor microenvironment of HCC patients, we analyzed 22 kinds of immune cells in the tumor immune infiltration microenvironment of 372 HCC patients in TCGA database by the CIBERSORT algorithm. The results of immune infiltration analysis showed that the contents of $CD4^+$ T cells and M1 macrophages decreased significantly in patients with PBhiPMhi, while the contents of helper T cells and M2 macrophages increased significantly (**Figure 3A**). This suggests that purine-related metabolism is related to the immune infiltration microenvironment of the tumor. To be further explored the relationship between purine metabolism and immune cell infiltration, we collected peripheral blood leukocyte data from patients with HCC and found that uric acid level was positively correlated with peripheral blood leukocytes (**Figure 3B**). At the same time, in order to verify the reliability of CIBERSORT results, we also used MCP count and Estimate algorithms to verify the immune infiltration, and the results are basically consistent with the previous ones (**Figures 3C,D**).

Analysis of Differential Gene Expressions Between PBhiPMhi and PBloPMlo Groups

In order to further explore the differences at the molecular level between patients with PBhiPMhi and patients with PBloPMlo, we obtained the differential gene expressions (DEGs) between the two groups by the limma method. The screening criteria of DEGs were $|\log2fc| \geq 2$ and $p < 0.05$. The volcanic map showed the results of DEGs (**Figure 4A**). In order to further understand the function of these DEGs, we performed KEGG and GO functional enrichment analyses of upregulated DEGs and downregulated DEGs, respectively. KEGG enrichment analysis showed that the upregulated DEGs in the PBhiPMhi group were mainly enriched in neuroactive ligand receiver interaction and IL-17 signaling pathway, while the

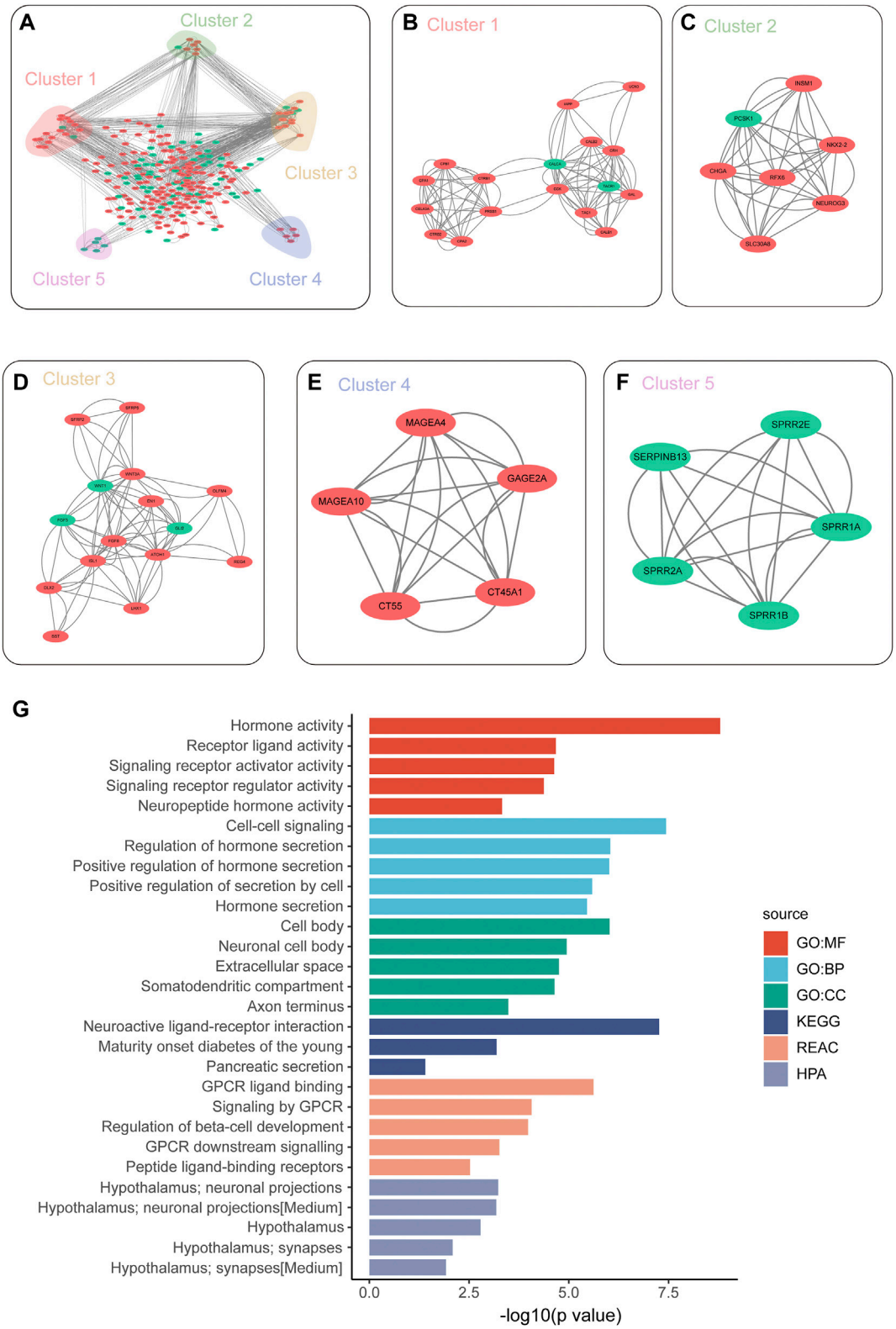


FIGURE 5 | Protein–protein interaction (PPI) analysis predicts possible molecular therapeutic targets. **(A)** Network diagram shows the DEG protein–protein interaction network between the PBhiPMhi group and PBloPMlo group; **(B–F)** network diagram shows the five sub-networks of MCODE plug-in mining. **(G)** Functional enrichment analysis on these 20 hub genes in three databases.

downregulated DEGs were enriched in calcium signaling pathway and gastric cancer (**Figure 4B**). GO enrichment analysis showed that the upregulated DEGs in the PBhiPMhi group were mainly enriched in the pathways related to the occurrence and development of cancer, while the downregulated DEGs were mainly enriched in the functions related to DNA damage repair (**Figures 4C,D**).

Protein–Protein Interaction Analysis Predicts Possible Molecular Therapeutic Targets

In order to further find possible drug targets, we performed protein–protein interactions on different genes through the STRING database. Then, we can visualize the results in the database (**Figure 5A**). In order to find the key modules in the PPI network, we mine the modules of the PPI network through the MCODE plug-in. Finally, we found five key modules, which may be the key driver modules affecting the prognosis of HCC (**Figures 5B–F**). At the same time, we further explored the hub gene through the Cytohubber plug-in. We found a total of 20 key hub genes, such as recombinant somatostatin (SST), cholecystokinin (CCK), and tachykinin precursor 1 (TAC1). The functions of these genes need to be verified by experiments later (**Supplementary Table S5**). Then, we performed functional enrichment analysis on these 20 hub genes in three databases, and the results showed that those hub genes were mainly enriched in hormone activity, receptor-ligand activity, and GPCR ligand binding (**Figure 5G**).

DISCUSSION

Purine is a heterocyclic bicyclic aromatic organic molecule, which includes DNA, RNA, nucleosides and nucleotides, AMP, ADP, ATP, GMP, GDP, GTP, and cyclic forms of cAMP and cGMP, and participates in various metabolic pathways and cell signal transduction (La Grotta et al., 2022; Xiao et al., 2022). Uric acid is the product of purine metabolism. The enhancement of purine metabolism will lead to the accumulation of uric acid in peripheral blood. In this study, we collected the data of uric acid in peripheral blood of 50 patients with HCC, analyzed its relationship with the prognosis of patients with HCC, and analyzed the functional enrichment of transcriptome, immune microenvironment, KEGG, and GO in patients with HCC. Through comprehensive bioinformatics analysis, we established the prognosis prediction model of HCC and revealed the potential significance of the purine metabolic pathway in predicting the prognosis of HCC.

The imbalance of purine nucleotide metabolism often occurs in the process of tumor occurrence and development (Wang X. et al., 2022). Many enzymes involved in purine nucleotide anabolism and catabolism are related to the proliferation and drug resistance of tumor cells (Wu et al., 2022). The imbalance of antioxidant and pro-inflammatory properties of uric acid can also induce tumors and promote their progress (Wang J. et al., 2022). Abnormal purine nucleotide metabolism can affect the expression of genes and proteins by regulating the signal transduction pathway and promote cell malignant transformation, invasion, and metastasis (Yang et al., 2019). The characteristics of nucleotide metabolism are different in different tumor patients.

A large number of studies have found that the level of serum uric acid is positively correlated with the occurrence of colorectal cancer, liver cancer, kidney cancer, melanoma, and head and neck tumors, but its relationship with the prognosis of HCC is still unclear (Ayoub et al., 2021; Wang H. et al., 2022; Wang J. et al., 2022). In our study, we found that the level of serum uric acid was positively correlated with the prognosis of HCC, suggesting that serum uric acid may be used as an index to predict the prognosis of patients with HCC. At the same time, we found that purine metabolism and purine anabolic pathway activities were significantly activated in patients with poor prognoses of HCC. The prognosis prediction model of HCC based on purine metabolism and purine synthesis pathways can accurately evaluate the prognosis of patients with HCC. At the same time, we found that there were significant changes in tumor immune infiltration microenvironment and biological function at the molecular level in patients with over-activation of purine metabolism and purine synthesis pathways, suggesting that purine metabolism and purine synthesis can participate in multiple processes mediating tumor occurrence and development.

In order to further explore possible drug targets, we performed protein–protein interaction on different genes through the STRING database and found a total of 20 key hub genes based on purine metabolism and purine synthesis prediction model, such as SST, CCK, and TAC1. These hub genes may be potential therapeutic targets for HCC. Some studies showed that SST expression can be used to predict an unfavorable prognosis in patients with HCC (Kaemmerer et al., 2017; Murakami et al., 2021). Serum levels of CCK are elevated in HCC patients, and CCK blockade is a novel approach for the prevention/treatment of HCC. (Tucker et al., 2020; Gay et al., 2021). However, there are few studies on the relationship between TACR1 and HCC. In studies of other cancers, researchers found that TACR1 correlates with the prognosis of colorectal cancer (CRC) and gastric cancer (GC) (David et al., 2009; Yu et al., 2012). These targets may have to guide significance for the future study of HCC.

In conclusion, in this study, we found that the level of peripheral blood uric acid in patients with HCC is correlated with their prognosis. The prognosis of patients with HCC can be accurately predicted through the metabolic process of uric acid and purine. At the same time, based on the prediction model, we found the cellular and molecular mechanisms that may affect the prognosis of HCC, including tumor immune infiltration microenvironment, a variety of biological processes, and potential drug targets. In addition, we found that uric acid level was positively correlated with peripheral blood leukocytes in HCC patients. The limitation of this study is that it does not apply the data with more samples for large-scale verification. We will conduct relevant experiments in the

future. In a word, these findings may provide a new way to predict the prognosis of HCC.

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 the Ethics Committee of Beijing Shijitan Hospital. 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

SY and XW conceived and designed the experiments. BZ, WT, and LQ collected the data. SY and XM analyzed the data and wrote the manuscript. All authors discussed and improved the article.

FUNDING

This work was funded by the National Natural Science Foundation of China (Grant No. 52073180), Beijing Municipal Science and Technology Project (Z211100002521011), and the National Science and Technology Major Project (2017ZX09304026).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.942267/full#supplementary-material

REFERENCES

Ayoub, A. M., Amin, M. U., Ambreen, G., Dayyih, A. A., Abdelsalam, A. M., Somaida, A., et al. (2022). Photodynamic and Antiangiogenic Activities of Parietin Liposomes in Triple Negative Breast Cancer. *Biomater. Adv.* 134, 112543. doi:10.1016/j.msec.2021.112543

Dai, K., Liu, C., Guan, G., Cai, J., and Wu, L. (2022). Identification of Immune Infiltration-Related Genes as Prognostic Indicators for Hepatocellular Carcinoma. *BMC Cancer* 22 (1), 496. doi:10.1186/s12885-022-09587-0

David, S., Kan, T., Cheng, Y., Agarwal, R., Jin, Z., and Mori, Y. (2009). Aberrant Silencing of the Endocrine Peptide Gene Tachykinin-1 in Gastric Cancer. *Biochem. Biophysical Res. Commun.* 378 (3), 605–609. doi:10.1016/j.bbrc.2008.11.078

Gay, M. D., Safronenka, A., Cao, H., Liu, F. H., Malchiodi, Z. X., Tucker, R. D., et al. (2021). Targeting the Cholecystokinin Receptor: A Novel Approach for Treatment and Prevention of Hepatocellular Cancer. *Cancer Prev. Res. (Phila)* 14 (1), 17–30. doi:10.1158/1940-6207.Capr-20-0220

Kaemmerer, D., Schindler, R., Mußbach, F., Dahmen, U., Altendorf-Hofmann, A., Dirsch, O., et al. (2017). Somatostatin and CXCR4 Chemokine Receptor Expression in Hepatocellular and Cholangiocellular Carcinomas: Tumor Capillaries as Promising Targets. *BMC Cancer* 17 (1), 896. doi:10.1186/s12885-017-3911-3

Ke, X., Wu, H., Chen, Y.-X., Guo, Y., Yao, S., Guo, M.-R., et al. (2022). Individualized Pathway Activity Algorithm Identifies Oncogenic Pathways in Pan-Cancer Analysis. *EBioMedicine* 79, 104014. doi:10.1016/j.ebiom.2022.104014

Kong, W., Mao, Z., Han, C., Ding, Z., Yuan, Q., Zhang, G., et al. (2022). A Novel Epithelial-Mesenchymal Transition Gene Signature Correlated with Prognosis, and Immune Infiltration in Hepatocellular Carcinoma. *Front. Pharmacol.* 13, 863750. doi:10.3389/fphar.2022.863750

La Grotta, R., de Candia, P., Olivieri, F., Matacchione, G., Giuliani, A., Rippo, M. R., et al. (2022). Anti-inflammatory Effect of SGLT-2 Inhibitors via Uric Acid and Insulin. *Cell. Mol. Life Sci.* 79 (5), 273. doi:10.1007/s00018-022-04289-z

Lee, A., and Lee, F.-C. (2020). Medical Oncology Management of Advanced Hepatocellular Carcinoma 2019: a Reality Check. *Front. Med.* 14 (3), 273–283. doi:10.1007/s11684-019-0728-2

Murakami, K., Kumata, H., Miyagi, S., Kamei, T., and Sasano, H. (2021). The Prognostic Significance of Neuroendocrine Markers and Somatostatin Receptor 2 in Hepatocellular Carcinoma. *Pathol. Int.* 71 (10), 682–691. doi:10.1111/pin.13149

Possemato, R., Marks, K. M., Shaul, Y. D., Pacold, M. E., Kim, D., Birsoy, K., et al. (2011). Functional Genomics Reveal that the Serine Synthesis Pathway Is Essential in Breast Cancer. *Nature* 476 (7360), 346–350. doi:10.1038/nature10350

Tucker, R. D., Ciofoaia, V., Nadella, S., Gay, M. D., Cao, H., Huber, M., et al. (2020). A Cholecystokinin Receptor Antagonist Halts Nonalcoholic Steatohepatitis and Prevents Hepatocellular Carcinoma. *Dig. Dis. Sci.* 65 (1), 189–203. doi:10.1007/s10620-019-05722-3

Vasan, N., and Cantley, L. C. (2022). At a Crossroads: How to Translate the Roles of PI3K in Oncogenic and Metabolic Signalling into Improvements in Cancer Therapy. *Nat. Rev. Clin. Oncol.* 19 (7), 471–485. doi:10.1038/s41571-022-00633-1

Wang, H., Xie, L., Song, X., Wang, J., Li, X., Lin, Z., et al. (2022a). Purine-Induced IFN-γ Promotes Uric Acid Production by Upregulating Xanthine Oxidoreductase Expression. *Front. Immunol.* 13, 773001. doi:10.3389/fimmu.2022.773001

Wang, J., Liu, K., Xiao, T., Liu, P., Prinz, R. A., and Xu, X. (2022b). Uric Acid Accumulation in DNA-Damaged Tumor Cells Induces NKG2D Ligand Expression and Antitumor Immunity by Activating TGF-β-Activated Kinase 1. *Oncoimmunology* 11 (1), 2016159. doi:10.1080/2162402X.2021.2016159

Wang, X., Su, W., Jiang, Y., Jia, F., Huang, W., Zhang, J., et al. (2022c). Regulation of Nucleotide Metabolism with Nutrient-Sensing Nanodrugs for Cancer Therapy. *Adv. Sci.*, 2200482. doi:10.1002/advs.202200482

Wu, H.-l., Gong, Y., Ji, P., Xie, Y.-f., Jiang, Y.-Z., and Liu, G.-y. (2022). Targeting Nucleotide Metabolism: a Promising Approach to Enhance Cancer Immunotherapy. *J. Hematol. Oncol.* 15 (1), 45. doi:10.1186/s13045-022-01263-x

Wu, L., Yang, W., Zhang, Y., Du, X., Jin, N., Chen, W., et al. (2021). Elevated Serum Uric Acid Is Associated with Poor Survival in Advanced HCC Patients and Febuxostat Improves Prognosis in HCC Rats. *Front. Pharmacol.* 12, 778890. doi:10.3389/fphar.2021.778890

Xiao, L., Sha, W., Tao, C., Hou, C., Xiao, G., and Ren, J. (2022). Effect on Purine Releasement of Lentinus Edodes by Different Food Processing Techniques. *Food Chem. X* 13, 100260. doi:10.1016/j.fochx.2022.100260

Yang, S., He, X., Liu, Y., Ding, X., Jiang, H., Tan, Y., et al. (2019). Prognostic Significance of Serum Uric Acid and Gamma-Glutamyltransferase in Patients with Advanced Gastric Cancer. *Dis. Markers* 2019, 12. doi:10.1155/2019/1415421

Yu, Y., Pan, Y., Jin, M., Zhang, M., Zhang, S., Li, Q., et al. (2012). Association of Genetic Variants in Tachykinins Pathway Genes with Colorectal Cancer Risk. *Int. J. Colorectal Dis.* 27 (11), 1429–1436. doi:10.1007/s00384-012-1478-7

Yuan, K., Xie, K., Lan, T., Xu, L., Chen, X., Li, X., et al. (2020). TXNDC12 Promotes EMT and Metastasis of Hepatocellular Carcinoma Cells via Activation of β-catenin. *Cell. Death Differ.* 27 (4), 1355–1368. doi:10.1038/s41418-019-0421-7

Yuan, R.-H., Hsu, C.-L., Jhuang, Y.-L., Liu, Y.-R., Hsieh, T.-H., and Jeng, Y.-M. (2022). Tumor-matrix Interaction Induces Phenotypic Switching in Liver Cancer Cells. *Hepatol. Int.* 16, 562–576. doi:10.1007/s12072-022-10315-w

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.

Copyright © 2022 Yang, Zhang, Tan, Qi, Ma and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 22 July 2022
doi: 10.3389/fgene.2022.906496

Prognostic Signature and Tumor Immune Landscape of N7-Methylguanosine-Related lncRNAs in Hepatocellular Carcinoma

Wei Wei[1†], *Chao Liu*[2†], *Meng Wang*[1], *Wei Jiang*[1], *Caihong Wang*[3] *and Shuqun Zhang*[1*]

[1]Department of Oncology, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China, [2]Department of Vascular Surgery, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China, [3]Department of Pathology, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China

OPEN ACCESS

Edited by:
Claudia Cava,
National Research Council (IBFM-CNR), Italy

Reviewed by:
Sabiha Khatoon,
Texas Tech University Health Sciences Center, United States
Weifeng Ding,
Nantong University, China

***Correspondence:**
Shuqun Zhang
shuqun_zhang1971@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: 24 April 2022
Accepted: 20 June 2022
Published: 22 July 2022

Citation:
Wei W, Liu C, Wang M, Jiang W, Wang C and Zhang S (2022) Prognostic Signature and Tumor Immune Landscape of N7-Methylguanosine-Related lncRNAs in Hepatocellular Carcinoma. Front. Genet. 13:906496. doi: 10.3389/fgene.2022.906496

Despite great advances in the treatment of liver hepatocellular carcinoma (LIHC), such as immunotherapy, the prognosis remains extremely poor, and there is an urgent need to develop novel diagnostic and prognostic markers. Recently, RNA methylation-related long non-coding RNAs (lncRNAs) have been demonstrated to be novel potential biomarkers for tumor diagnosis and prognosis as well as immunotherapy response, such as N6-methyladenine (m6A) and 5-methylcytosine (m5C). N7-Methylguanosine (m7G) is a widespread RNA modification in eukaryotes, but the relationship between m7G-related lncRNAs and prognosis of LIHC patients as well as tumor immunotherapy response is still unknown. In this study, based on the LIHC patients' clinical and transcriptomic data from TCGA database, a total of 992 m7G-related lncRNAs that co-expressed with 22 m7G regulatory genes were identified using Pearson correlation analysis. Univariate regression analysis was used to screen prognostic m7G-related lncRNAs, and the least absolute shrinkage and selection operator (LASSO) and multivariate Cox regression were applied to construct a 9-m7G-related-lncRNA risk model. The m7G-related lncRNA risk model was validated to exhibit good prognostic performance through Kaplan–Meier analysis and ROC analysis. Together with the clinicopathological features, the m7G-related lncRNA risk score was found to be an independent prognostic factor for LIHC. Furthermore, the high-risk group of LIHC patients was unveiled to have a higher tumor mutation burden (TMB), and their tumor microenvironment was more prone to the immunosuppressive state and exhibited a lower response rate to immunotherapy. In addition, 47 anti-cancer drugs were identified to exhibit a difference in drug sensitivity between the high-risk and low-risk groups. Taken together, the m7G-related lncRNA risk model might display potential value in predicting prognosis, immunotherapy response, and drug sensitivity in LIHC patients.

Keywords: m7G methylation, lncRNA, LIHC, prognostic model, immune response

INTRODUCTION

Liver hepatocellular carcinoma (LIHC), accounting for approximately 90% cases of the liver non-metastatic tumors, is a global health problem with increasing incidence and mortality (Llovet et al., 2021). For early-stage LIHC, hepatic resection and local ablative therapy are the standard clinical treatment manners (Kamarajah et al., 2021). Owing to insidious onset, rapid progression, and difficulties in early diagnosis, the majority of LIHC patients are first diagnosed at an advanced stage, at which point therapeutic options are limited and ineffective. In the past decade, molecular targeted therapy has become the mainstay of the treatment in advanced-stage LIHC, mainly including sorafenib and lenvatinib as first-line therapy, and regorafenib and ramucirumab, as well as cabozantinib as second-line therapy (Huang A. et al., 2020). Excitingly, immunotherapies, especially immune checkpoint inhibitors (ICIs), have rapidly developed in the past few years and have been proven to be an effective treatment for LIHC with long-term survival (Llovet et al., 2022). At present, nivolumab and pembrolizumab, two anti-PD-1 antibodies, have been used in the clinical treatment of LIHC (Wong et al., 2021). However, LIHC is easy to metastasize and develop drug resistance and has a special tumor immune microenvironment (TIME) (Oura et al., 2021); thus, the current therapeutic effect and prognosis of LIHC patients are still not optimistic, and more novel accurate molecular biomarkers are needed to improve the diagnostic and treatment efficacy of LIHC patients.

Long non-coding RNAs (lncRNAs) are featured as more than 200 nucleotides in length and lack the ability of protein coding. LncRNAs have similar structures to mRNA, including 5′-cap, 3′-poly(A) tail, and promoter structure and can be transcribed by RNA polymerase II (Statello et al., 2021). The number of lncRNAs encoded by human is huge and exceeds 170,000 (Zhao et al., 2021). LncRNAs can modulate gene expression in both cis- and trans-manners at the transcriptional, epigenetic, and post-transcriptional levels and have been demonstrated to be critical regulators of the development and progression of cancer, including LIHC (Gao et al., 2020).

In recent years, epigenetic modifications of lncRNA, such as N6-methyladenine (m6A) and 5-methylcytosine (m5C), have been found to be associated with carcinogenesis and the development of multiple cancers (Wang et al., 2020; Zhang et al., 2021). N7-Methylguanosine (m7G) is a kind of positively charged RNA modification, which is generated by the addition of a methyl group to the N7 atom of guanine (G) by RNA methyltransferase, such as METTL1 (Orellana et al., 2021). M7G modifications are widely present in various RNA molecules including mRNA 5′ cap structure, internal mRNA, transfer RNA (tRNA), rRNA ribosomal RNA (rRNA), and primary microRNA (pri-miRNA), as well as lncRNA, and it has been revealed that m7G can modulate mRNA transcription, mRNA translation, splicing, tRNA stability, nuclear processing, 18 S rRNA maturation, and miRNA biosynthesis (Malbec et al., 2019; Pandolfini et al., 2019; Wang et al., 2022). Similar to m6A and m5C, m7G was recently demonstrated to play fundamental roles in LIHC. For example, METTL1, an m7G methyltransferase, has been linked to advanced tumor stage, vascular invasion, and poor prognosis in LIHC patients and facilitates tumor progression through increasing m7G tRNA modification and promoting translation of target mRNAs (Tian et al., 2019; Chen et al., 2021). Unfortunately, the specific function of m7G-related lncRNAs in the prognosis of LIHC patients remains unclear. Thus, a novel m7G-related lncRNA signature may be helpful in prognostic prediction and treatment response evaluation in LIHC.

In this study, based on the LIHC patients' clinical and transcriptomic data from TCGA database, a total of 992 m7G-related lncRNAs that co-expressed with 22 m7G regulatory genes were identified using Pearson correlation analysis. Univariate regression analysis was used to screen prognostic m7G-related lncRNAs, and the least absolute shrinkage and selection operator (LASSO) and multivariate Cox regression were applied to construct a 9-m7G-related-lncRNA risk model. The risk model was further verified by Kaplan–Meier analysis and ROC analysis. Furthermore, the roles of the risk model in evaluating tumor mutation burden, immune microenvironment, and immunotherapy response as well as drug sensitivity were explored. In summary, we constructed a 9-m7G-related-lncRNA risk model which may provide promising prognostic value and play essential roles in predicting immune therapy response and chemotherapy sensitivity.

MATERIALS AND METHODS

Data Source and Preprocessing

The transcriptome data, mutation data, and corresponding clinicopathological data of LIHC patients were obtained from TCGA database (https://portal.gdc.cancer.gov/) on 8 February 2022. There were 371 LIHC samples and 50 normal subjects. The expression matrix of these primary samples was normalized in the R package "DESeq2" by variance stabilizing transformation function, and each normalized count was formulated in log2(X +1), where X is each of the normalized counts. Moreover, LIHC patients with missing OS status and times were deleted to reduce bias in statistical analysis.

Selection of m7G Regulator Genes and m7G-Related lncRNAs

A total of 22 m7G modification-related genes were screened from the published literature works, namely, METTL1 (Ma et al., 2021), WDR4 (Ma et al., 2021), DCP2 (Wurm and Sprangers, 2019), DCPS (Wulf et al., 2019), NUDT2 (Song et al., 2013), NUDT3 (Song et al., 2013), NUDT12 (Song et al., 2013), NUDT15 (Song et al., 2013), NUDT16 (Song et al., 2013), NUDT17 (Song et al., 2013), AGO2 (Kiriakidou et al., 2007), CYFIP1 (Napoli et al., 2008), EIF4E (Napoli et al., 2008), EIF4E2 (Rosettani et al., 2007), EIF4E3 (Osborne et al., 2013), EIF3D (Lee et al., 2016), EIF4A1 (Chu et al., 2020), EIF4G3 (Haghighat and Sonenberg, 1997), GEMIN5 (Bradrick and Gromeier, 2009), LARP1 (Philippe et al., 2018), NCBP1 (Dou et al., 2020), and NCBP2 (Dou et al., 2020). After the annotation of the expression

matrix of LIHC samples, the profiles of 22 m7G regulator genes and 13,541 lncRNAs were obtained. Then, a preliminary screening was performed based on the rule that the median expression/variance of each generally changed lncRNA in every sample was 20% higher than the total median expression/variance of all lncRNAs in every patient (Luo et al., 2018), and 4,464 distinctively expressed lncRNAs were screened for further Pearson's correlation analysis. Finally, 992 m7G-related lncRNAs were identified under the criterion of | correlation coefficient| > 0.3 and p-value < 0.01.

Construction of the Prognostic Risk Model of m7G-Related lncRNAs

The entire LIHC patients were separated into a training set (n = 219) and testing set (n = 146) randomly, and the m7G-related lncRNA risk model was constructed based on the training set. First of all, prognostic lncRNAs were selected from 992 m7G-related lncRNAs using univariate regression analysis, and then LASSO-penalized Cox regression was further adopted to optimize indicators predicting clinical outcome by using the R package "glmnet." Finally, a 9-m7G-related-lncRNA prognostic risk model was constructed by multivariate regression analysis. The risk score was computed in the following manner: risk score = coef (lncRNA1) × expr (lncRNA1) + coef (lncRNA2) × expr (lncRNA2) + . . . + coef (lncRNAn) × expr (lncRNAn). In this formula, the coefficient and corresponding expression value of each lncRNA were calculated, respectively.

Validation and Evaluation of the Prognostic Risk Model of m7G-Related lncRNAs

The entire LIHC patients and the testing set were applied to validate the efficacy of the 9-m7G-related lncRNA risk model. Based on the median risk score, LIHC patients were categorized into high- and low-risk groups. Principal component analysis (PCA) was conducted to visualize the grouping ability of high-dimensional data in the entire gene set, 22 m7G regulator gene set, 992 m7G-related lncRNAs, and 9 m7G-related lncRNAs expression profiles. The R packages "survival" and "survminer" were adopted to perform KM survival analysis to analyze the differences in the overall survival (OS), disease-specific survival (DSS), disease-free interval (DFI), and progression-free interval (PFI) between the two risk groups. The 1-year, 3-year, and 5-year area under the ROC (AUROC) values were calculated to evaluate the predicting performance of the risk model.

Independence of the 9-m7G-Related-lncRNA Risk Model in LIHC Patients

We stratified the clinicopathological features of LIHC patients according to age, gender, grade, and stage and analyzed the difference of OS between the two risk groups in the entire set. Moreover, univariate and multivariate Cox regression analyses were calculated to further evaluate the independent prognostic factors in LIHC patients, such as m7G-related lncRNA risk score, age, gender, grade, and stage. The ROC curve of each clinicopathological feature was drawn to show its prognostic value.

The Predictive Ability of the Nomogram in LIHC Patients

We constructed the nomogram consisting of clinicopathological features (including age, gender, tumor grade, and stage) along with the m7G-related lncRNA risk score to predict the survival status in entire LIHC patients. Moreover, the calibration curve was plotted to evaluate the consistency between the actual and predicted survival of 1-, 3-, and 5-year OS.

Evaluation of the Tumor Immune Microenvironment Using the Risk Model

KEGG analysis was performed to identify the changed pathways between the two risk groups. KEGG gene sets were downloaded from the MSigDB (https://www.gsea-msigdb.org/gsea/msigdb/), and gene set enrichment analysis (GSEA) of KEGG pathways was implemented using the R package "clusterProfiler." To evaluate the immune landscape in LIHC patients among different risk groups, we used the R package "GSVA" to perform single-sample GSEA (ssGSEA) to quantify the infiltration level of 28 immune cell types and 13 immune-related functions in each sample.

Exploring the Immunotherapeutic Treatment Response Targeting the Risk Model

The tumor mutation burden (TMB) was estimated by using the R package "maftools" in both risk groups. The KM survival analysis was performed to analyze the differences in the OS stratified by TMB status and m7G-related lncRNA risk scores in LIHC patients. The tumor immune dysfunction and exclusion (TIDE) score (http://tide.dfci.harvard.edu/) was calculated to evaluate the response to immunotherapy. Furthermore, the expression level of immune checkpoint molecules, such as CD274, PDCD1LG2, LAG3, SIGLEC15, TIGIT, IDO1, CTLA4, and CD276, was compared between the two risk groups in entire LIHC patients.

Identification of Novel Candidate Compounds for Chemotherapy Based on the Risk Model

To identify novel candidate compounds in LIHC patients for clinical practice, the half-maximal inhibitory concentration (IC_{50}) of 138 compounds obtained from the Genomics of Drug Sensitivity in Cancer (GDSC) database was calculated for each LIHC patient by using the R package "pRRophetic," and the value of IC_{50} of each compound between the two risk groups was compared separately.

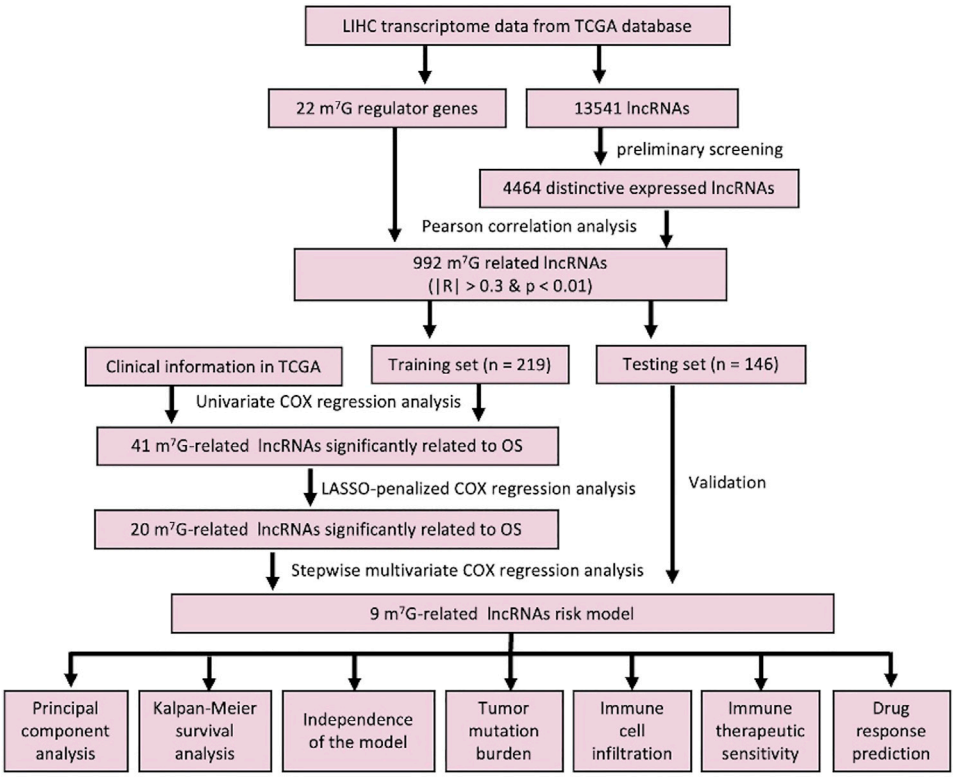


FIGURE 1 | Detailed workflow of this study.

RESULTS

Identification of m7G-Related lncRNAs in LIHC Patients

The workflow of this study is described in detail in **Figure 1**. Initially, we totally screened 22 m7G regulatory genes, including 2 "writers," 8 "erasers," and 12 "readers" through literature retrieval, and most of them (16/22) were found to be significantly changed in LIHC ($p < 0.05$) based on TCGA data (**Figure 2A**). To identify m7G-related lncRNAs in LIHC, expression data of 13,541 lncRNAs in the LIHC cohort was obtained from TCGA database, and 4,464 generally changed lncRNAs with distinctive expression among different LIHC patients were obtained through preliminary screening. Then, we performed correlation analysis and filtered 992 m7G-related lncRNAs under the filter criterion of |correlation coefficient| > 0.3 and p-value < 0.01. The network of m7G regulators and corresponding lncRNAs was illustrated using the Sankey diagram (**Figure 2B**).

Construction and Validation of a 9-m7G-Related-lncRNA Risk Model for LIHC Patients

The entire LIHC patients were separated into a training set ($n = 219$) and testing set ($n = 146$) in a random manner, and the clinical features are comparable between these two sets ($p > 0.05$) (**Supplementary Table S1**). First, 41 prognostic lncRNAs were selected from 992 m7G-related lncRNAs using univariate regression analysis in the training set, which revealed to be correlated with OS ($p < 0.05$), and half of them (20/41) were risk factors (hazard ratio, HR > 1) in LIHC (**Figure 3A**). Next, 20 m7G-related lncRNAs were further screened by LASSO regression analysis (**Figures 3B,C**), which can effectively reduce characteristics in high-dimensional data and optimize indicators predicting clinical outcome. Finally, a 9-m7G-related-lncRNA prognostic risk model was constructed by multivariate regression analysis (**Figure 3D**). The risk score was computed in the following manner: risk score = (−0.199 × SOCS2-AS1 expression) + (−0.368 × RP5-1171I10.5 expression) + (−0.160 × RP11-588H23.3 expression) + (0.167 × RP11-43F13.3 expression) + (−0.192 × RP11-10A14.3 expression) + (0.154202559 × RP11-95O2.5 expression) + (−0.194843224 × NAV2-AS4 expression) + (0.149322014 × RP11-519G16.5 expression) + (0.168349973 × RP11-874J12.4 expression). The correlation between m7G regulator genes and 9 m7G-related lncRNA expressions was also visualized in **Supplementary Figure S1**. Moreover, LIHC patients in the entire set were categorized into high- and low-risk groups based on the median risk score, and the grouping ability of this prognostic risk model was verified by PCA analysis. Results showed that the expression of the entire genes, 22 m7G regulator genes and 992 m7G-related lncRNAs, showed diffused distribution in both the risk groups (**Figures 3E–G**), whereas the expression of 9 m7G-related lncRNAs included in this prognostic risk model was well divided into two clusters with different risks (**Figure 3H**).

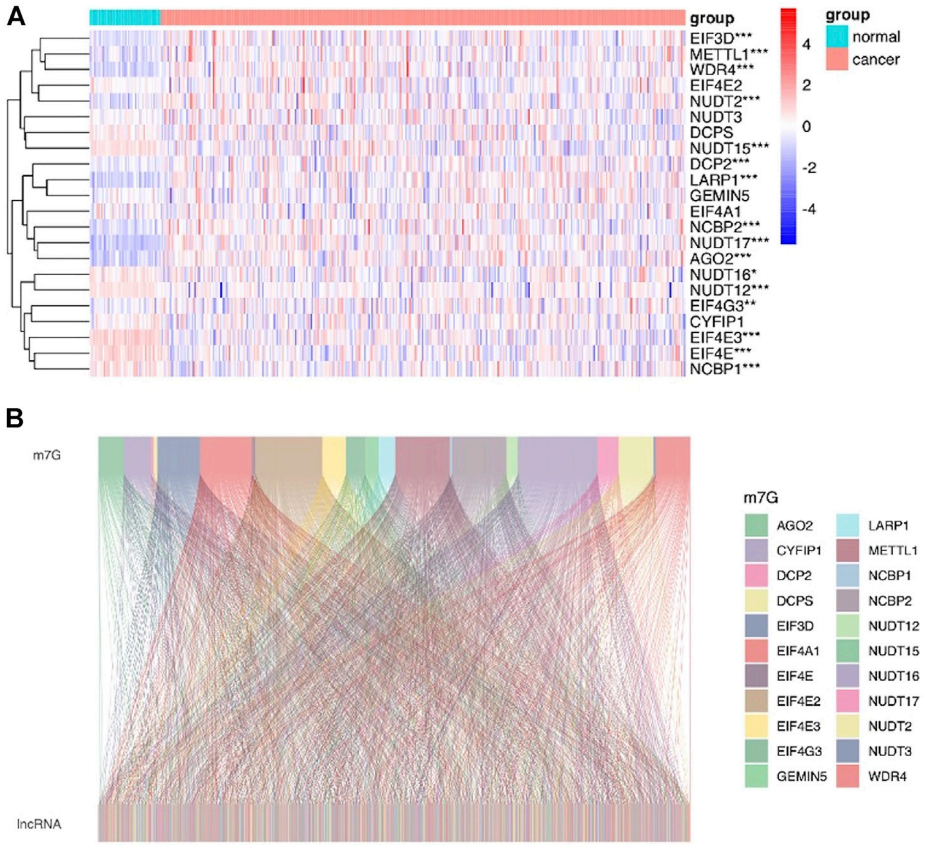


FIGURE 2 | Identification of m7G-related lncRNAs in LIHC patients. **(A)** Heatmap showed the differences in the expression of m7G regulators between LIHC and normal groups. **(B)** Sankey diagram displayed the relationship between 22 m7G genes and 992 m7G-related lncRNAs.

To evaluate the prognostic efficacy of this model, we applied this risk score formula into different datasets. For the training set, we arranged the LIHC patients according to the risk score, and the heatmap showed that the expression pattern of the 9 m7G-related lncRNAs was different, and the scatter plot revealed that patients with a higher risk score featured worser living status (**Figures 4A–C**). The KM survival analysis indicated that high-risk patients had shorter OS than low-risk patients ($p < 0.001$) (**Figure 4D**). In addition, the predicting performance of the risk model was calculated with ROC curves, and AUROC of 1, 3, and 5 years for OS was 0.818, 0.880, and 0.902, respectively (**Figure 4E**). For the testing and entire set, the risk score distribution, heatmap of lncRNA expression, and scatter plot of survival status and risk score are depicted in **Figures 5A–C** and **Supplementary Figures S2A–C**, respectively. The survival analysis also showed that high-risk patients had shorter OS than low-risk patients in both testing and entire sets (**Figure 5D**; **Supplementary Figure S1D**). The AUROC of 1, 3, and 5 years for OS in the testing set was 0.771, 0.707, and 0.637 (**Figure 5E**) and in the entire set was 0.801, 0.811, and 0.777 (**Supplementary Figure S2E**). To further validate the capability of the prognostic risk model, we explored the difference in DSS, DFI, and PFI between the two groups and found that DSS, DFI, and PFI was also longer in low-risk patients (**Supplementary Figures S3A–F**). Taken together, the 9-m7G-related-lncRNA risk model possessed optimal prognostic efficacy in LIHC patients.

Risk Score Based on the 9-m7G-Related-lncRNA Risk Model Was an Independent Prognostic Factor in LIHC Patients

We subsequently stratified the clinicopathological features of LIHC patients according to the median value of the risk score and analyzed the difference in OS between high- and low-risk groups in the entire set. The survival analysis revealed that this risk model performed well in all the subgroups stratified by age (<65 years old and ≥65 years old), gender (male and female), tumor grade (grades 1–2 and grades 3–4), and stage (stages I–II and stages III–IV), and high-risk patients had a shorter OS than the low-risk group (**Figures 6A–H**).

To further validate the independent prognostic factors in LIHC patients, univariate and multivariate Cox regression analyses were performed. Univariate Cox regression analysis

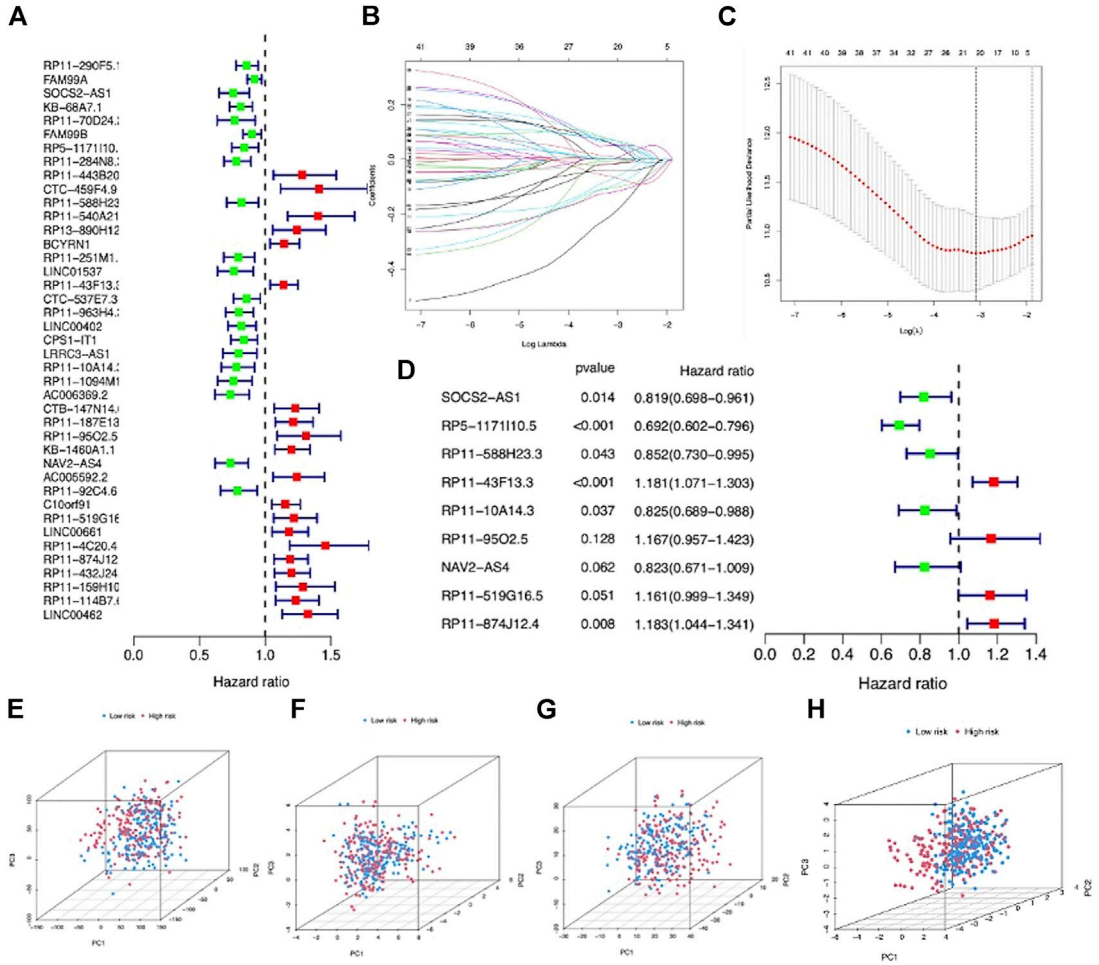


FIGURE 3 | Construction of a m7G-related lncRNA risk model for LIHC patients. **(A)** Forest plot showed 41 prognostic lncRNAs screened *via* univariate regression analysis. **(B)** LASSO regression of 20 m7G-related lncRNAs. **(C)** Cross-validation in LASSO regression. **(D)** Forest plot displayed 9 m7G-related lncRNAs selected by multivariate regression analysis. **(E)** PCA based on entire LIHC gene expression profiles in the two groups. **(F)** PCA based on 22 m7G regulator gene expressions in the two groups. **(G)** PCA based on 992 m7G-related lncRNA expressions in the two groups. **(H)** PCA based on 9 prognostic m7G-related lncRNA expressions in the two groups.

indicated that both stage (HR: 2.409%, 95% CI: 1.664–3.487, $p < 0.001$) and m7G-related lncRNA risk score (HR: 1.297%, 95% CI: 1.228–1.371, $p < 0.001$) were significantly related with OS (**Figure 7A**). Multivariate analysis also revealed that stage (HR: 1.933%, 95% CI: 1.303–1.867, $p < 0.001$) and m7G-related lncRNA risk score (HR: 1.283%, 95% CI: 1.198–1.374, $p < 0.001$) were independent prognostic factors in LIHC patients (**Figure 7B**). The AUROC was calculated to better explicit the efficacy of m7G-related lncRNA risk score as well as other clinicopathological features in predicting the OS in LIHC patients, and results demonstrated that the risk score had higher AUROC than other clinicopathological features (**Figure 7C**).

Moreover, we constructed the nomogram consisting of clinicopathological features (including age, gender, tumor grade, and stage) along with the m7G-related lncRNA risk score to predict the survival status in entire LIHC patients. Both m7G-related lncRNA risk score and tumor stage showed better predictive ability than other clinicopathological features in the nomogram (**Figure 7D**). The calibration curve also proved acceptable consistency between the actual and predicted survival within 1-, 3-, and 5-year OS (**Figure 7E**), demonstrating that the 9-m7G-related-lncRNA risk model was reliable and could work well in predicting prognosis in LIHC patients.

Analysis of Tumor Mutation Burden, Immune Landscape, and Immunotherapy Response Targeting the m7G-Related lncRNA Risk Model

TMB was calculated by using the R package "maftools" in the two risk groups, and results showed that gene mutation incidence was elevated in the high-risk group, especially TP53 gene (34% vs. 21%) (**Figures 8A,B**). Survival analysis demonstrated that the OS outcome was shorter in high-risk patients regardless of the TMB risk. Furthermore, patients with high TMB in the low-risk group

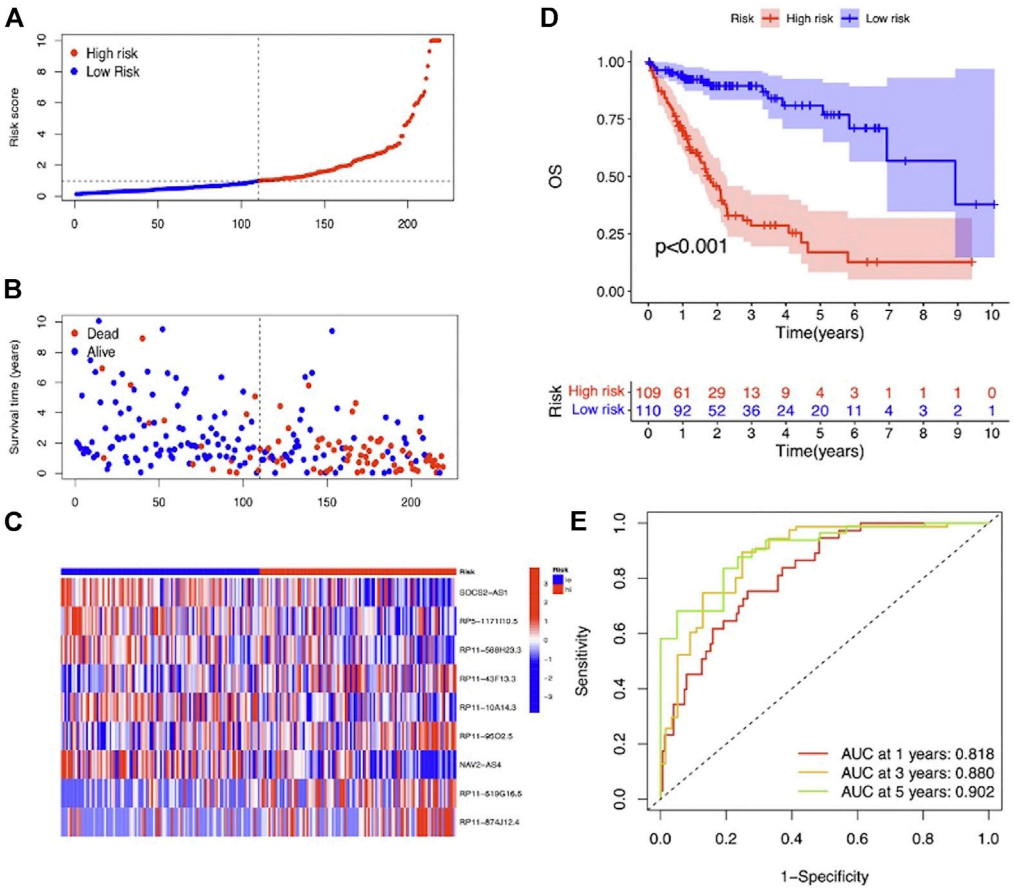


FIGURE 4 | Prognostic value of the 9-m7G-related-lncRNA risk model between the two groups in the training set. **(A)** Distribution of the risk score (the x-axis represented the LIHC patients arranged according to the risk score, and the y-axis represented values of the risk score for each patient). **(B)** Scatter plot of survival status and risk score (the x-axis represented the LIHC patients arranged according to the risk score, and the y-axis represented the survival time of each patient). **(C)** Heatmap of the expression profile of the 9 m7G-related lncRNAs. **(D)** KM curves displayed the OS of LIHC patients between high- and low-risk groups. **(E)** ROC curves of the risk model of 1, 3, and 5 years for OS.

survived longer than patients with low TMB in high-risk group, indicating that the m7G-related lncRNA risk model had better prognostic significance than TMB status (**Figure 8C**).

The tumor immune landscape, including pathway enrichment, immune cell infiltration, and immune-related functions, was analyzed based on the m7G-related lncRNA risk model in LIHC patients with different risks. First, KEGG pathway analysis was performed to explore the underlying molecular mechanisms occurred in different risk groups, and results illustrated that many immune-related pathways were markedly suppressed in the high-risk group, such as antigen processing and presentation, cytokine–receptor interaction, and natural killer cell-mediated cytotoxicity, while only cell cycle pathway was significantly activated (**Figure 8D**). Next, the tumor infiltration immune cells and immune-related functions were assessed based on ssGSEA scores between the two risk groups, and results showed that activated B cells, immature B cells, memory B cells, activated $CD8^+$ T cells, effector memory $CD8^+$ T cells, natural killer T cells, Th1 cells, eosinophil, macrophage, and mast cells exhibited a lower expression in the high-risk group (**Figure 8E**). Additionally, the heatmap showed many immune-related functions, such as type2 IFN response, T-cell co-stimulation, T-cell co-inhibition, cytolytic activity, inflammation promoting, CCR, HLA, and type1 IFN response, were also downregulated in the high-risk group (**Figure 8F**).

As for immunotherapy response, previous studies had demonstrated that a higher TIDE score was associated with worse immunotherapy response and poor prognosis (Jiang et al., 2018). Hence, we calculated the TIDE score to predict the response to immunotherapy in LIHC patients and disclosed that the high-risk group had a lower response rate to immunotherapy than the low-risk group, which was in accordance with previous tumor immune microenvironment analysis (**Figure 8G**). Moreover, immune checkpoint molecules are vital targets of immune checkpoint inhibitors (ICIs), and we discovered that high-risk group patients exhibited lower expression of CD274, PDCD1LG2, LAG3, SIGLEC15, TIGIT, IDO1, and CTLA4 than low-risk group patients, except for CD276 (**Figure 8H**).

Furthermore, novel candidate compounds based on the m7G-related lncRNA risk model were identified by calculating the IC_{50}

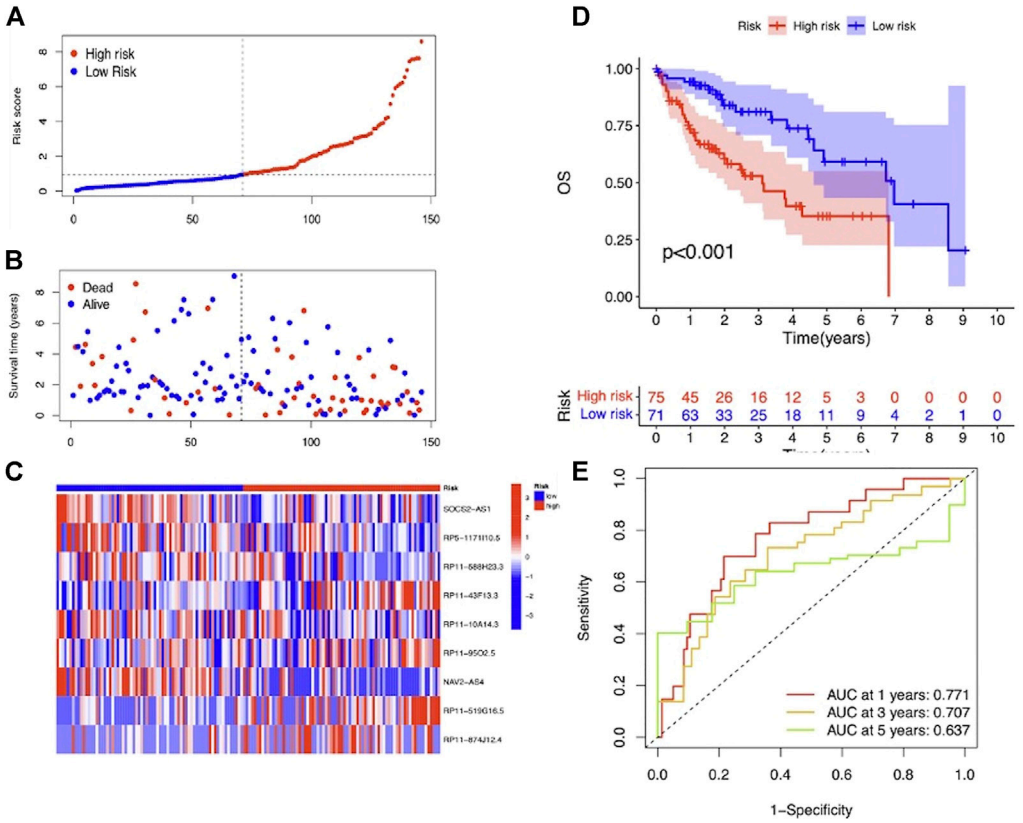


FIGURE 5 | Prognostic value of the 9-m7G-related-lncRNA risk model between the two groups in the testing set. **(A)** Distribution of the risk score (the x-axis represented the LIHC patients arranged according to the risk score, and the y-axis represented values of the risk score for each patient). **(B)** Scatter plot of survival status and risk score (the x-axis represented the LIHC patients arranged according to the risk score, and the y-axis represented the survival time of each patient). **(C)** Heatmap of the expression profile of the 9 m7G-related lncRNAs. **(D)** KM curves displayed the OS of LIHC patients between high- and low-risk groups. **(E)** ROC curves of the risk model of 1, 3, and 5 years for OS.

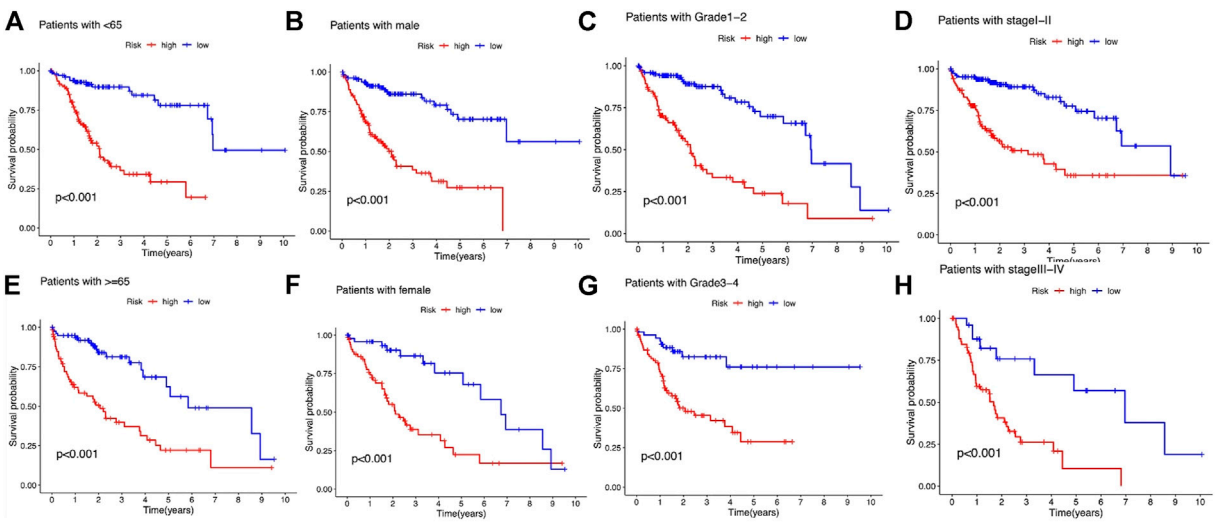


FIGURE 6 | Kaplan–Meier survival analysis stratified by age, gender, tumor grade, and stage between the high- and low-risk groups in the entire set. **(A)** Patients with age <65. **(B)** Patients with male gender. **(C)** Patients with tumor grade 1-2. **(D)** Patients with tumor stage I–II. **(E)** Patients with age ≥65. **(F)** Patients with female gender. **(G)** Patients with tumor grade 3-4. **(H)** Patients with tumor stage III–IV.

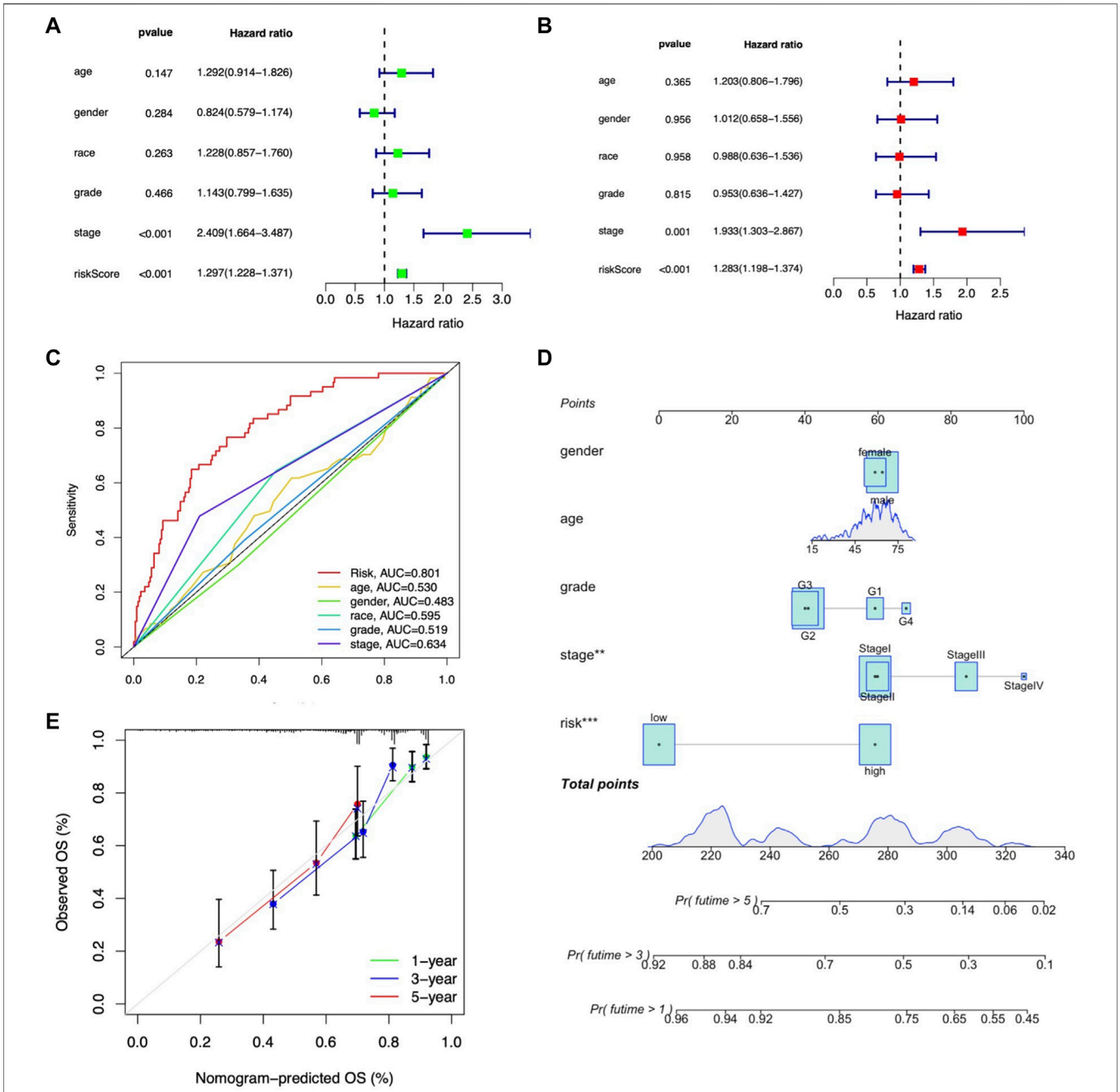


FIGURE 7 | Assessment of the independent prognostic factors and construction of a prognostic nomogram in the entire LIHC set. **(A)** Univariate Cox regression analysis of the clinical characteristics and risk score with the OS. **(B)** Multivariate Cox regression analysis of the clinical characteristics and risk score with the OS. **(D)** Nomogram predicting the probability of 1-, 2-, and 3-year OS. **(C)** ROC curves of the clinical characteristics and risk score. **(E)** Calibration plot of the nomogram predicting the probability of the 1-, 2-, and 3-year OS.

for each LIHC patient using the GDSC database. In total, 47 of 138 compounds were found to exhibit significant differences in the estimated IC_{50} between the two risk groups ($p < 0.05$), suggesting that the m7G-related lncRNA risk model had the potential of predicting the chemotherapy sensitivity. Top 20 novel compounds are listed in **Supplementary Figures S4A–T** and could be used for further analysis in LIHC patients.

DISCUSSION

LIHC is a common malignancy with high morbidity and mortality, which seriously threatens the health of people all over the world. Clinically, for the reasons such as lacking effective approaches for early diagnosis, most patients have progressed to the advanced stage and then lost the opportunity of early eradication treatment. Hence,

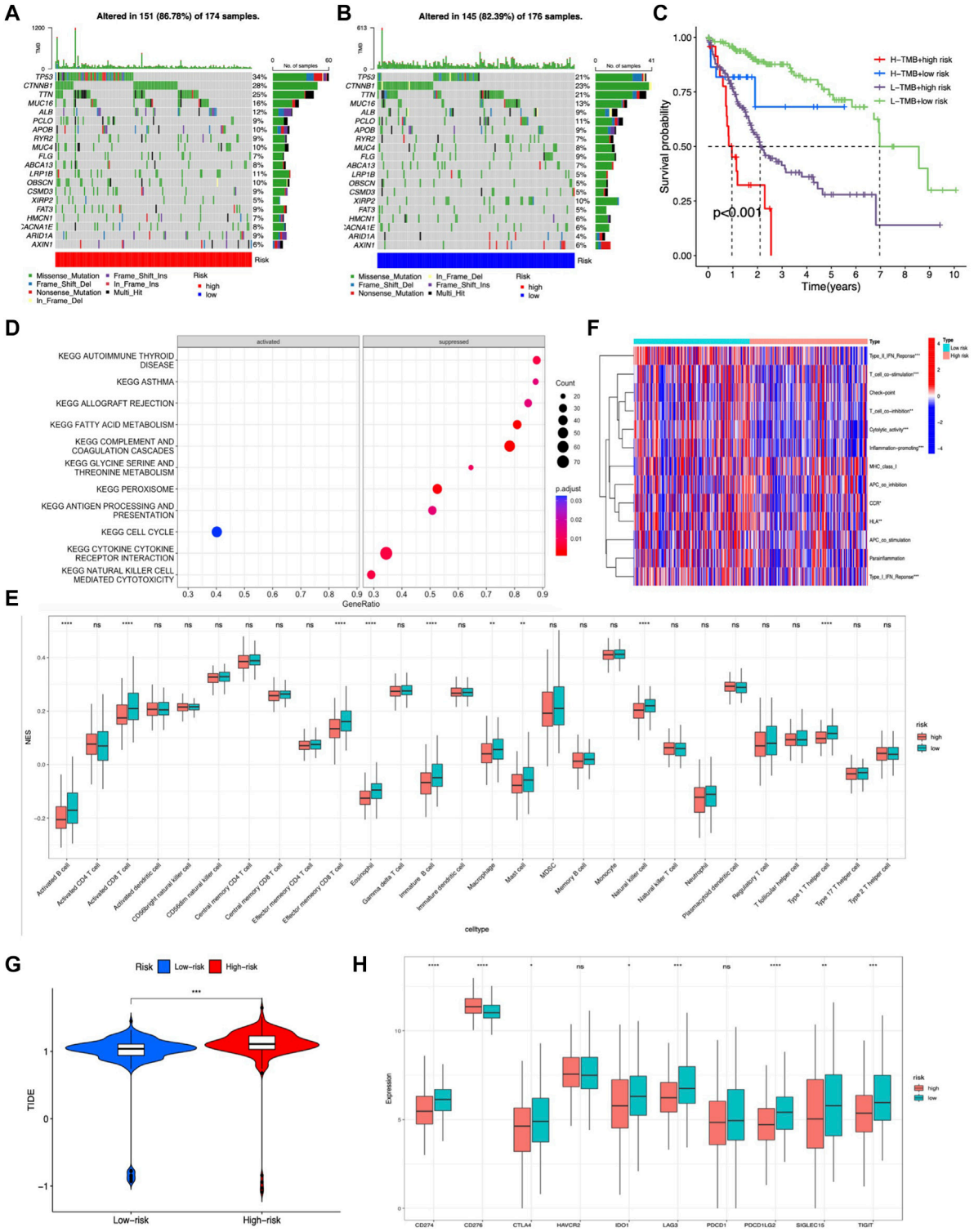


FIGURE 8 | Evaluation of the tumor immune landscape and immunotherapy response based on the m7G-related lncRNA model in the entire LIHC set. **(A,B)** Waterfall plot displayed top 20 mutation genes' information in the two risk groups. **(C)** KM survival analysis of OS stratified by tumor mutation burden and the m7G-related lncRNA model. **(G)** TIDE prediction difference between two risk score subgroups. **(D)** Significant KEGG pathways enriched in high-risk patients. **(E)** Difference in tumor infiltration immune cells based on ssGSEA scores between two risk groups. **(F)** Difference in immune-related functions based on ssGSEA scores between two risk score subgroups. **(H)** Expression of immune checkpoint blockade-related genes between the two risk groups.

accurate biomarkers are needed to improve the diagnostic and treatment efficacy of LIHC patients.

LncRNAs as novel regulatory molecules play fundamental roles in the progression of LIHC (Ghafouri-Fard et al., 2021). Studies have shown that lncRNAs can modulate the proliferation, migration, invasion, angiogenesis, and drug resistance of tumor cells by forming a complex regulatory network with mRNA and miRNA (Huang Z. et al., 2020). Meanwhile, lncRNAs can also be used as indicators for early diagnosis and efficacy prediction of treatment strategies, such as surgery, radiotherapy, chemotherapy, and immunotherapy, thus becoming important molecular types for the development of diagnostic biomarkers as well as molecular therapeutic targets of LIHC (Yuan et al., 2021). In recent years, the methylation modification of lncRNA, such as m6A and m5C, has been revealed to be involved in tumor progression including LIHC, and RNA methylation-related lncRNAs have demonstrated to be novel potential biomarkers for tumor diagnosis and prognosis as well as immunotherapy response (Jin et al., 2021; Li et al., 2021; Zhu et al., 2021).

As a widespread RNA epigenetic modification in eukaryotes, m7G has recently been found to modify tRNA and promote tumor progression by regulating translation efficiency (Katsara and Schneider, 2021). For LIHC, METTL1, the key component of the m7G methyltransferase complex, was disclosed to be upregulated in LIHC and can promote tumor progression *via* m7G tRNA modification-dependent translation control (Chen et al., 2021). However, whether m7G modifications are involved in tumorigenesis and development by regulating lncRNAs remain unknown. Given the potential of m6A- and m5C-associated lncRNAs in predicting tumor prognosis and immunotherapy response, we speculate that m7G-related lncRNAs are highly likely to possess the same ability. In the present study, we constructed and validated the prognostic risk model based on m7G-related lncRNAs and explored the roles of the model in evaluating tumor mutation load, immune cell infiltration, immunotherapy response, and drug sensitivity.

To obtain m7G-related lncRNAs, we first locked 22 m7G regulatory genes through literature retrieval, including 2 "writers" (METTL1 and WDR4), 8 "erasers" (DCP2, DCPS, NUDT2, NUDT3, NUDT12, NUDT15, NUDT16, and NUDT17), and 12 "readers" (AGO2, CYFIP1, EIF4E, EIF4E2, EIF4E3, GEMIN5, LARP1, NCBP1, NCBP2, EIF3D, EIF4A1, and EIF4G3). Currently, apart from METTL1 and WDR4 (Chen et al., 2021; Xia et al., 2021), it is still unclear whether other genes can participate in LIHC progression *via* m7G-dependent manner. Here, we found that the majority of m7G regulators (16/22) were abnormally expressed in LIHC, indicating m7G modification may exert crucial functions in LIHC. Subsequently, a total of 992 lncRNAs that were co-expressed with m7G regulators were screened through Pearson correlation analysis.

Combined with clinicopathological data, an LIHC prognostic risk model comprising 9 m7G-related lncRNAs was finally achieved by univariate, LASSO, and multivariate Cox regression analyses successively. Five lncRNAs (SOCS2-AS1, RP5-1171I10.5, RP11-588H23.3, RP11-10A14.3, and NAV2-AS4) were identified to be a protective factor for LIHC prognosis, and other four lncRNAs (RP11-43F13.3, RP11-95O2.5, RP11-519G16.5, and RP11-874J12.4) were risk factors affecting the prognosis of LIHC. SOCS2-AS1 has recently been validated as a tumor suppressor in colorectal and endometrial cancers, and its low expression in tumors was positively related to poor prognosis of patients (Zheng et al., 2020; Jian et al., 2021). RP11-874J12.4 was recently identified as an oncogenic lncRNA that could facilitate oral squamous cell carcinoma tumorigenesis and gastric cancer chemoresistance (Liu et al., 2020; Liu et al., 2021). Subsequently, LIHC patients were separated into high- and low-risk groups based on the median risk score, and its prognostic efficacy was evaluated through KM and ROC analysis, and we found that the model showed good prognostic performance in all the datasets. Moreover, univariate and multivariate Cox regression analyses disclosed that the risk score based on the m7G-related lncRNA risk model was an independent prognostic factor for LIHC, and the m7G-related lncRNA risk score exhibited better OS prediction performance than other clinicopathological features, including age, gender, race, and tumor grade as well as stage.

Tumor mutation burden (TMB) is associated with genomic instability and immunogenicity, involving base substitutions, gene insertion and deletion, and other mutations and has the potential in predicting prognosis and immunotherapy efficacy (Huang et al., 2021). In this study, we investigated the TMB in different risk groups and found that TMB was elevated in the high-risk group, and LIHC patients with high TMB had worse OS in both the risk groups. Several studies also revealed the negative roles of TMB in the prognosis of LIHC (Cai et al., 2020; Zhou et al., 2021).

Tumor immune microenvironment (TIME) influences tumor progression and immunotherapy response. Although immunotherapy (such as immune checkpoint inhibitor drugs) has shown potential therapeutic effects for advanced LIHC, and the effectiveness of immunotherapy is greatly affected by its immunosuppressive TIME (Zhou et al., 2022). Currently, changing the TIME has become a new strategy to promote immune control and immunotherapy efficacy of tumors. In the present study, the relationship between m7G-related lncRNAs and TIME was evaluated by exploring KEGG pathway enrichment, immune cell infiltration, and immune-related functions. We found that high-risk LIHC patients were more inclined to an immunosuppressive state because immune-related pathways (such as natural killer cell-mediated cytotoxicity, antigen processing, and presentation), immune-related function (such as type1/2 IFN response, T-cell co-stimulation, and cytolytic activity), and anti-tumor immune cells (such as activated B cells, $CD8^+$ T cells, and natural killer T cells) were all inactivated in high-risk patients, suggesting that m7G-related lncRNAs may be related to the immunosuppressive tumor microenvironment in LIHC. Furthermore, the association between m7G-related lncRNAs and immunotherapy response was also investigated by TIDE scoring, and we unveiled that high-risk LIHC patients showed a lower response rate to immunotherapy, indicating that m7G-related lncRNAs have great potential in predicting immunotherapy response. Additionally, the correlation between the m7G-related lncRNA risk model and the sensitivity of 138 anti-cancer drugs was evaluated using GDSC database, and 47 compounds were identified to exhibit the difference in chemotherapy sensitivity

between the two groups, such as cisplatin, cytarabine, tipifarnib, and temsirolimus, suggesting that m7G-related lncRNA may also be used as a biomarker to predict the chemotherapy sensitivity and provide medication guidance for the personalized chemotherapy of LIHC patients.

CONCLUSION

We constructed a 9-m7G-related-lncRNA risk model based on the LIHC patients' clinical and transcriptomic data from TCGA database. The risk model was validated to exhibit good prognostic performance and was found to be an independent prognostic factor for LIHC. Furthermore, the roles of the risk model in tumor mutation burden, immune microenvironment, and immunotherapy response, as well as drug sensitivity were also evaluated. In conclusion, the m7G-related-lncRNA risk model might display potential value in predicting prognosis, immunotherapy response, and drug sensitivity in LIHC 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.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Xi'an Jiaotong University health science center. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

WW, CL, and SZ inspired this study and analyzed TCGA data. MW, WJ, and WC summarized the data and plotted most figures. WW and CL drafted the manuscript. SZ monitored this study. All authors contributed to the article and approved the submitted version.

FUNDING

This study was founded by the National Natural Science Foundation of China (No. 81802456) and the Fundamental Research Funds for the Central Universities (xzy012019085).

ACKNOWLEDGMENTS

Most of the data were downloaded from TCGA database; therefore, the authors acknowledge this website for providing the relevant data for free. The authors also thank Dr. Shipeng Guo and all the members of his bioinformatics team for their training and sharing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.906496/full#supplementary-material

REFERENCES

Bradrick, S. S., and Gromeier, M. (2009). Identification of Gemin5 as a Novel 7-Methylguanosine Cap-Binding Protein. *PLoS One* 4 (9), e7030. doi:10.1371/journal.pone.0007030

Cai, H., Zhang, Y., Zhang, H., Cui, C., Li, C., and Lu, S. (2020). Prognostic Role of Tumor Mutation Burden in Hepatocellular Carcinoma after Radical Hepatectomy. *J. Surg. Oncol.* 121 (6), 1007–1014. doi:10.1002/jso.25859

Chen, Z., Zhu, W., Zhu, S., Sun, K., Liao, J., Liu, H., et al. (2021). METTL1 Promotes Hepatocarcinogenesis via m(7) G tRNA Modification-Dependent Translation Control. *Clin. Transl. Med.* 11 (12), e661. doi:10.1002/ctm2.661

Chu, J., Zhang, W., Cencic, R., O'Connor, P. B. F., Robert, F., Devine, W. G., et al. (2020). Rocaglates Induce Gain-Of-Function Alterations to eIF4A and eIF4F. *Cell Rep.* 30 (8), 2481–2488. doi:10.1016/j.celrep.2020.02.002

Dou, Y., Kalmykova, S., Pashkova, M., Oghbaie, M., Jiang, H., Molloy, K. R., et al. (2020). Affinity Proteomic Dissection of the Human Nuclear Cap-Binding Complex Interactome. *Nucleic Acids Res.* 48 (18), 10456–10469. doi:10.1093/nar/gkaa743

Gao, N., Li, Y., Li, J., Gao, Z., Yang, Z., Li, Y., et al. (2020). Long Non-Coding RNAs: The Regulatory Mechanisms, Research Strategies, and Future Directions in Cancers. *Front. Oncol.* 10, 598817. doi:10.3389/fonc.2020.598817

Ghafouri-Fard, S., Gholipour, M., Hussen, B. M., and Taheri, M. (2021). The Impact of Long Non-Coding RNAs in the Pathogenesis of Hepatocellular Carcinoma. *Front. Oncol.* 11, 649107. doi:10.3389/fonc.2021.649107

Haghighat, A., and Sonenberg, N. (1997). eIF4G Dramatically Enhances the Binding of eIF4E to the mRNA 5'-cap Structure. *J. Biol. Chem.* 272 (35), 21677–21680. doi:10.1074/jbc.272.35.21677

Huang, A., Yang, X.-R., Chung, W.-Y., Dennison, A. R., and Zhou, J. (2020). Targeted Therapy for Hepatocellular Carcinoma. *Sig. Transduct. Target. Ther.* 5 (1), 146. doi:10.1038/s41392-020-00264-x

Huang, T., Chen, X., Zhang, H., Liang, Y., Li, L., Wei, H., et al. (2021). Prognostic Role of Tumor Mutational Burden in Cancer Patients Treated with Immune Checkpoint Inhibitors: A Systematic Review and Meta-Analysis. *Front. Oncol.* 11, 706652. doi:10.3389/fonc.2021.706652

Huang, Z., Zhou, J.-K., Peng, Y., He, W., and Huang, C. (2020). The Role of Long Noncoding RNAs in Hepatocellular Carcinoma. *Mol. Cancer* 19 (1), 77. doi:10.1186/s12943-020-01188-4

Jian, F., Che, X., Zhang, J., Liu, C., Liu, G., Tang, Y., et al. (2021). The Long-Noncoding RNA SOCS2-AS1 Suppresses Endometrial Cancer Progression by Regulating AURKA Degradation. *Cell Death Dis.* 12 (4), 351. doi:10.1038/s41419-021-03595-x

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T Cell Dysfunction and Exclusion Predict Cancer Immunotherapy Response. *Nat. Med.* 24 (10), 1550–1558. doi:10.1038/s41591-018-0136-1

Jin, C., Li, R., Deng, T., Li, J., Yang, Y., Li, H., et al. (2021). Identification and Validation of a Prognostic Prediction Model of m6A Regulator-Related LncRNAs in Hepatocellular Carcinoma. *Front. Mol. Biosci.* 8, 784553. doi:10.3389/fmolb.2021.784553

Kamarajah, S. K., Bundred, J. R., Littler, P., Reeves, H., Manas, D. M., and White, S. A. (2021). Treatment Strategies for Early Stage Hepatocellular Carcinoma: A

Systematic Review and Network Meta-Analysis of Randomised Clinical Trials. *HPB* 23 (4), 495–505. doi:10.1016/j.hpb.2020.10.031

Katsara, O., and Schneider, R. J. (2021). m(7)G tRNA Modification Reveals New Secrets in the Translational Regulation of Cancer Development. *Mol. Cell* 81 (16), 3243–3245. doi:10.1016/j.molcel.2021.07.030

Kiriakidou, M., Tan, G. S., Lamprinaki, S., De Planell-Saguer, M., Nelson, P. T., and Mourelatos, Z. (2007). An mRNA m7G Cap Binding-Like Motif within Human Ago2 Represses Translation. *Cell* 129 (6), 1141–1151. doi:10.1016/j.cell.2007.05.016

Lee, A. S. Y., Kranzusch, P. J., Doudna, J. A., and Cate, J. H. D. (2016). eIF3d is an mRNA Cap-Binding Protein that is Required for Specialized Translation Initiation. *Nature* 536 (7614), 96–99. doi:10.1038/nature18954

Li, L., Xie, R., and Lu, G. (2021). Identification of m6A Methyltransferase-Related lncRNA Signature for Predicting Immunotherapy and Prognosis in Patients with Hepatocellular Carcinoma. *Biosci. Rep.* 41 (6), BSR20210760. doi:10.1042/BSR20210760

Liu, G. M., Lu, T. C., Sun, M. L., Ji, X., Zhao, Y. A., Jia, W. Y., et al. (2020). RP11-874J12.4 Promotes Oral Squamous Cell Carcinoma Tumorigenesis via the miR-19a-5p/EBF1 Axis. *J. Oral Pathol. Med.* 49 (7), 645–654. doi:10.1111/jop.13000

Liu, Y., Cao, J., Pu, Y. S., Ma, Y., Wu, M., and Wang, J. H. (2021). RP11-874J12.4, a Novel lncRNA, Confers Chemoresistance in Human Gastric Cancer Cells by Sponging miR-3972 and Upregulating SSR2 Expression. *Am. J. Transl. Res.* 13 (6), 5892–5910.

Llovet, J. M., Castet, F., Heikenwalder, M., Maini, M. K., Mazzaferro, V., Pinato, D. J., et al. (2022). Immunotherapies for Hepatocellular Carcinoma. *Nat. Rev. Clin. Oncol.* 19 (3), 151–172. doi:10.1038/s41571-021-00573-2

Llovet, J. M., Kelley, R. K., Villanueva, A., Singal, A. G., Pikarsky, E., Roayaie, S., et al. (2021). Hepatocellular Carcinoma. *Nat. Rev. Dis. Prim.* 7 (1), 6. doi:10.1038/s41572-020-00240-3

Luo, D., Deng, B., Weng, M., Luo, Z., and Nie, X. (2018). A Prognostic 4-lncRNA Expression Signature for Lung Squamous Cell Carcinoma. *Artif. Cells Nanomed. Biotechnol.* 46 (6), 1207–1214. doi:10.1080/21691401.2017.1366334

Ma, J., Han, H., Huang, Y., Yang, C., Zheng, S., Cai, T., et al. (2021). METTL1/WDR4-Mediated m(7)G tRNA Modifications and m(7)G Codon Usage Promote mRNA Translation and Lung Cancer Progression. *Mol. Ther.* 29 (12), 3422–3435. doi:10.1016/j.ymthe.2021.08.005

Malbec, L., Zhang, T., Chen, Y.-S., Zhang, Y., Sun, B.-F., Shi, B.-Y., et al. (2019). Dynamic Methylome of Internal mRNA N(7)-Methylguanosine and its Regulatory Role in Translation. *Cell Res.* 29 (11), 927–941. doi:10.1038/s41422-019-0230-z

Napoli, I., Mercaldo, V., Boyl, P. P., Eleuteri, B., Zalfa, F., De Rubeis, S., et al. (2008). The Fragile X Syndrome Protein Represses Activity-Dependent Translation through CYFIP1, a New 4E-BP. *Cell* 134 (6), 1042–1054. doi:10.1016/j.cell.2008.07.031

Orellana, E. A., Liu, Q., Yankova, E., Pirouz, M., De Braekeleer, E., Zhang, W., et al. (2021). METTL1-Mediated m(7)G Modification of Arg-TCT tRNA Drives Oncogenic Transformation. *Mol. Cell* 81 (16), 3323–3338. e3314. doi:10.1016/j.molcel.2021.06.031

Osborne, M. J., Volpon, L., Kornblatt, J. A., Culjkovic-Kraljacic, B., Baguet, A., and Borden, K. L. B. (2013). eIF4E3 Acts as a Tumor Suppressor by Utilizing an Atypical Mode of Methyl-7-Guanosine Cap Recognition. *Proc. Natl. Acad. Sci. U. S. A.* 110 (10), 3877–3882. doi:10.1073/pnas.1216862110

Oura, K., Morishita, A., Tani, J., and Masaki, T. (2021). Tumor Immune Microenvironment and Immunosuppressive Therapy in Hepatocellular Carcinoma: A Review. *Int. J. Mol. Sci.* 22 (11), 5801. doi:10.3390/ijms22115801

Pandolfini, L., Barbieri, I., Bannister, A. J., Hendrick, A., Andrews, B., Webster, N., et al. (2019). METTL1 Promotes Let-7 MicroRNA Processing via m7G Methylation. *Mol. Cell* 74 (6), 1278–1290. doi:10.1016/j.molcel.2019.03.040

Philippe, L., Vasseur, J.-J., Debart, F., and Thoreen, C. C. (2018). La-Related Protein 1 (LARP1) Repression of TOP mRNA Translation is Mediated through its Cap-Binding Domain and Controlled by an Adjacent Regulatory Region. *Nucleic Acids Res.* 46 (3), 1457–1469. doi:10.1093/nar/gkx1237

Rosettani, P., Knapp, S., Vismara, M.-G., Rusconi, L., and Cameron, A. D. (2007). Structures of the Human eIF4E Homologous Protein, h4EHP, in its m7GTP-Bound and Unliganded Forms. *J. Mol. Biol.* 368 (3), 691–705. doi:10.1016/j.jmb.2007.02.019

Song, M.-G., Bail, S., and Kiledjian, M. (2013). Multiple Nudix Family Proteins Possess mRNA Decapping Activity. *RNA* 19 (3), 390–399. doi:10.1261/rna.037309.112

Statello, L., Guo, C.-J., Chen, L.-L., and Huarte, M. (2021). Gene Regulation by Long Non-Coding RNAs and its Biological Functions. *Nat. Rev. Mol. Cell Biol.* 22 (2), 96–118. doi:10.1038/s41580-020-00315-9

Tian, Q.-H., Zhang, M.-F., Zeng, J.-S., Luo, R.-G., Wen, Y., Chen, J., et al. (2019). METTL1 Overexpression is Correlated with Poor Prognosis and Promotes Hepatocellular Carcinoma via PTEN. *J. Mol. Med.* 97 (11), 1535–1545. doi:10.1007/s00109-019-01830-9

Wang, H., Chen, R. B., Zhang, S. N., and Zhang, R. F. (2022). N7-Methylguanosine Modification of lncRNAs in a Rat Model of Hypoxic Pulmonary Hypertension: A Comprehensive Analysis. *BMC Genomics* 23 (1), 33. doi:10.1186/s12864-021-08188-8

Wang, T., Kong, S., Tao, M., and Ju, S. (2020). The Potential Role of RNA N6-Methyladenosine in Cancer Progression. *Mol. Cancer* 19 (1), 88. doi:10.1186/s12943-020-01204-7

Wong, J. S. L., Kwok, G. G. W., Tang, V., Li, B. C. W., Leung, R., Chiu, J., et al. (2021). Ipilimumab and Nivolumab/Pembrolizumab in Advanced Hepatocellular Carcinoma Refractory to Prior Immune Checkpoint Inhibitors. *J. Immunother. Cancer* 9 (2), e001945. doi:10.1136/jitc-2020-001945

Wulf, M. G., Buswell, J., Chan, S.-H., Dai, N., Marks, K., Martin, E. R., et al. (2019). The Yeast Scavenger Decapping Enzyme DcpS and its Application for *In Vitro* RNA Recapping. *Sci. Rep.* 9 (1), 8594. doi:10.1038/s41598-019-45083-5

Wurm, J. P., and Sprangers, R. (2019). Dcp2: An mRNA Decapping Enzyme that Adopts Many Different Shapes and Forms. *Curr. Opin. Struct. Biol.* 59, 115–123. doi:10.1016/j.sbi.2019.07.009

Xia, P., Zhang, H., Xu, K., Jiang, X., Gao, M., Wang, G., et al. (2021). MYC-Targeted WDR4 Promotes Proliferation, Metastasis, and Sorafenib Resistance by Inducing CCNB1 Translation in Hepatocellular Carcinoma. *Cell Death Dis.* 12 (7), 691. doi:10.1038/s41419-021-03973-5

Yuan, D., Chen, Y., Li, X., Li, J., Zhao, Y., Shen, J., et al. (2021). Long Non-Coding RNAs: Potential Biomarkers and Targets for Hepatocellular Carcinoma Therapy and Diagnosis. *Int. J. Biol. Sci.* 17 (1), 220–235. doi:10.7150/ijbs.50730

Zhang, Q., Liu, F., Chen, W., Miao, H., Liang, H., Liao, Z., et al. (2021). The Role of RNA m(5)C Modification in Cancer Metastasis. *Int. J. Biol. Sci.* 17 (13), 3369–3380. doi:10.7150/ijbs.61439

Zhao, L., Wang, J., Li, Y., Song, T., Wu, Y., Fang, S., et al. (2021). NONCODEV6: an Updated Database Dedicated to Long Non-Coding RNA Annotation in Both Animals and Plants. *Nucleic Acids Res.* 49 (D1), D165–D171. doi:10.1093/nar/gkaa1046

Zheng, Z., Li, X., You, H., Zheng, X., and Ruan, X. (2020). LncRNA SOCS2-AS1 Inhibits Progression and Metastasis of Colorectal Cancer through Stabilizing SOCS2 and Sponging miR-1264. *Aging* 12 (11), 10517–10526. doi:10.18632/aging.103276

Zhou, G., Boor, P. P. C., Bruno, M. J., Sprengers, D., and Kwekkeboom, J. (2022). Immune Suppressive Checkpoint Interactions in the Tumour Microenvironment of Primary Liver Cancers. *Br. J. Cancer* 126 (1), 10–23. doi:10.1038/s41416-021-01453-3

Zhou, W., Fang, D., He, Y., and Wei, J. (2021). Correlation Analysis of Tumor Mutation Burden of Hepatocellular Carcinoma Based on Data Mining. *J. Gastrointest. Oncol.* 12 (3), 1117–1131. doi:10.21037/jgo-21-259

Zhu, X.-L., Li, Q., Shen, J., Shan, L., Zuo, E.-D., and Cheng, X. (2021). Use of 6 m6A-Relevant lncRNA Genes as Prognostic Markers of Primary Liver Hepatocellular Carcinoma Based on TCGA Database. *Transl. Cancer Res. TCR* 10 (12), 5337–5351. doi:10.21037/tcr-21-2440

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.

Copyright © 2022 Wei, Liu, Wang, Jiang, Wang and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

REVIEW
published: 22 July 2022
doi: 10.3389/fgene.2022.927046

A Pan-Cancer Analysis Reveals the Prognostic and Immunotherapeutic Value of Stanniocalcin-2 (STC2)

Zhong-Hui Jiang†, Xianfeng Shen†, Yanhong Wei, Yongji Chen, Hongbo Chai, Lingyun Xia and Weidong Leng**

Department of Stomatology, Taihe Hospital, Hubei University of Medicine, Shiyan, China

Background: Stanniocalcin-2 (STC2) is a secreted glycoprotein which plays an important role in regulating the homeostasis of calcium, glucose homeostasis, and phosphorus metastasis. Accumulating evidence suggests that STC2 is implicated in cancer mechanisms. However, the effects of STC2 on cancer development and progression across pan-cancer are not yet completely known.

Methods: Data were downloaded from The Cancer Genome Atlas database to obtain differentially expressed genes significantly associated with prognosis (key genes). A gene was selected for subsequent correlation studies by integrating the significance of prognosis and the time-dependent ROC curve. Gene expression of different tumor types was analyzed based on the UCSC XENA website. Furthermore, our study investigated the correlation of STC2 expression between prognosis, immune cell infiltration, immune checkpoint genes (ICGs), mismatch repair genes (MMRs), tumor mutation burden (TMB), microsatellite instability (MSI), and drug sensitivity in various malignant tumors. Gene set enrichment analysis (GSEA) was conducted for correlated genes of STC2 to explore potential mechanisms.

Results: A total of 3,429 differentially expressed genes and 397 prognosis-related genes were identified from the TCGA database. Twenty-six key genes were found by crossing the former and the latter, and the highest risk gene, STC2, was selected for subsequent correlation studies. STC2 had good diagnostic performance for HNSCC, and was closely related to the survival status and clinicopathological stage of HNSCC patients. In pan-cancer analysis, STC2 was upregulated in 20 cancers and downregulated in seven cancers. STC2 overexpression was overall negatively correlated with overall survival, disease-free survival, disease-specific survival, and progress-free survival. STC2 was profoundly correlated with the tumor immune microenvironment, including immune cell infiltration, ICGs, MMRs, TMB, and MSI. Moreover, STC2 was significantly negatively correlated with the sensitivity or resistance of multiple drugs.

Conclusion: STC2 was a potential prognostic biomarker for pan-cancer and a new immunotherapy target.

Keywords: stanniocalcin-2, head and neck squamous cell carcinoma, prognostic biomarkers, target therapy, pan-cancer

OPEN ACCESS

Edited by:
Ehsan Nazemalhosseini-Mojarad, Shahid Beheshti University of Medical Sciences, Iran

Reviewed by:
Zahra Salehi, Shahid Beheshti University of Medical Sciences, Iran
Shuanggang Chen, Yuebei People's Hospital, China

****Correspondence:***
Lingyun Xia
xialingyun200810@163.com
Weidong Leng
lwd35@163.com

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

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: *23 April 2022*
Accepted: *20 June 2022*
Published: *22 July 2022*

Citation:
Jiang Z-H, Shen X, Wei Y, Chen Y, Chai H, Xia L and Leng W (2022) A Pan-Cancer Analysis Reveals the Prognostic and Immunotherapeutic Value of Stanniocalcin-2 (STC2). Front. Genet. 13:927046. doi: 10.3389/fgene.2022.927046

INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) is the sixth most common cancer, and the 5-year overall survival rate of HNSCC is only 40–50% by incidence worldwide (Bray et al., 2018). Every year, approximately 600,000 HNSCC cases are diagnosed worldwide (Cohen et al., 2016). HNSCC arises in mucosal epithelium including oral cavity, pharynx, and larynx and is tightly linked to exposure to tobacco-derived carcinogens, alcohol consumption, and human-papillomavirus (HPV) infection (Nakagawa et al., 2021). The prognosis for patients with HNSCC is largely determined by tumor stage at presentation. Early-stage tumors have a favorable prognosis, whereas advanced cancer has poor prognosis. Clinically, most HNSCC patients are diagnosed with advanced cancer with lymph node metastasis. Therefore, this creates an urgent need for practical biomarkers to help in accurate early diagnosis and prediction of tumor prognosis.

Stanniocalcin-2 (STC2) is a secreted glycoprotein which plays a pivotal role in regulating the homeostasis of calcium, glucose homeostasis, and phosphorus metastasis (Takei et al., 2012). STC2 has been shown to play a regulatory role in various tissues and organs in an autocrine or paracrine manner (Chen and Zhu 2008). The correlation between STC2 and tumor progression has been investigated over the last 5 years in multiple studies. STC2 is a useful molecular marker predicting the progression and prognosis of digestive system tumors, including hepatocellular carcinoma (Wang et al., 2019), pancreatic cancer (Lin et al., 2019), gastric cancer (Ke et al., 2019), and colorectal cancer (Zhang et al., 2019). STC2 has also affected extrahepatic cholangiocarcinoma (Li and Yang 2020), nasopharyngeal carcinoma (He et al., 2019), laryngeal cancer (Ding et al., 2021), and osteosarcoma (Yang et al., 2021). STC2 activates ITGB2/FAK/SOX6 and PI3K–AKT signaling pathways in some cancers (Yang et al., 2017; Li et al., 2022). On the other hand, several studies showed that STC2 is a protective factor for some tumors. STC2 reduces resistance to sunitinib in clear-cell renal cell carcinoma (Qin et al., 2022). STC2 inhibits migration and invasion of breast cancer *via* inhibition of PKC signaling (Hou et al., 1932). Therefore, it is necessary to perform a pan-cancer analysis to reveal the role and underlying molecular mechanisms of STC2 in clinical phenotypic characteristics and tumor immune microenvironments of multiple cancers.

In the present study, we first integrated the samples from the UCSC XENA website to identify the differentially expressed genes (DEGs) and prognosis-related genes. We further screened key genes by crossing the DEGs and prognosis-related genes and selected the highest risk key gene for subsequent pan-cancer analysis. We investigated the correlation between the gene expression and prognosis, immune microenvironment, immune checkpoint genes (ICGs), tumor mutation burden (TMB), microsatellite instability (MSI), mismatch repair genes (MMRs), and immune/stromal scores.

MATERIALS AND METHODS

The Datasets

The RNA sequencing (RNA-seq) data from The Cancer Genome Atlas (TCGA) database and the genotype–tissue expression (GTEx) project were downloaded from the UCSC XENA website (https://xenabrowser.net/datapages/). These data included the 33 different tumor types and matched normal tissues from TCGA and 31 normal human tissues from GTEx. Corresponding clinical profiles were downloaded from these two databases. Duplicated data from these two data sources were excluded. The appropriate gene expression profiles were downloaded from the gene expression omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) to test the results of the TCGA data analysis.

Functional Enrichment Analysis of Upregulated Differentially Expressed Genes

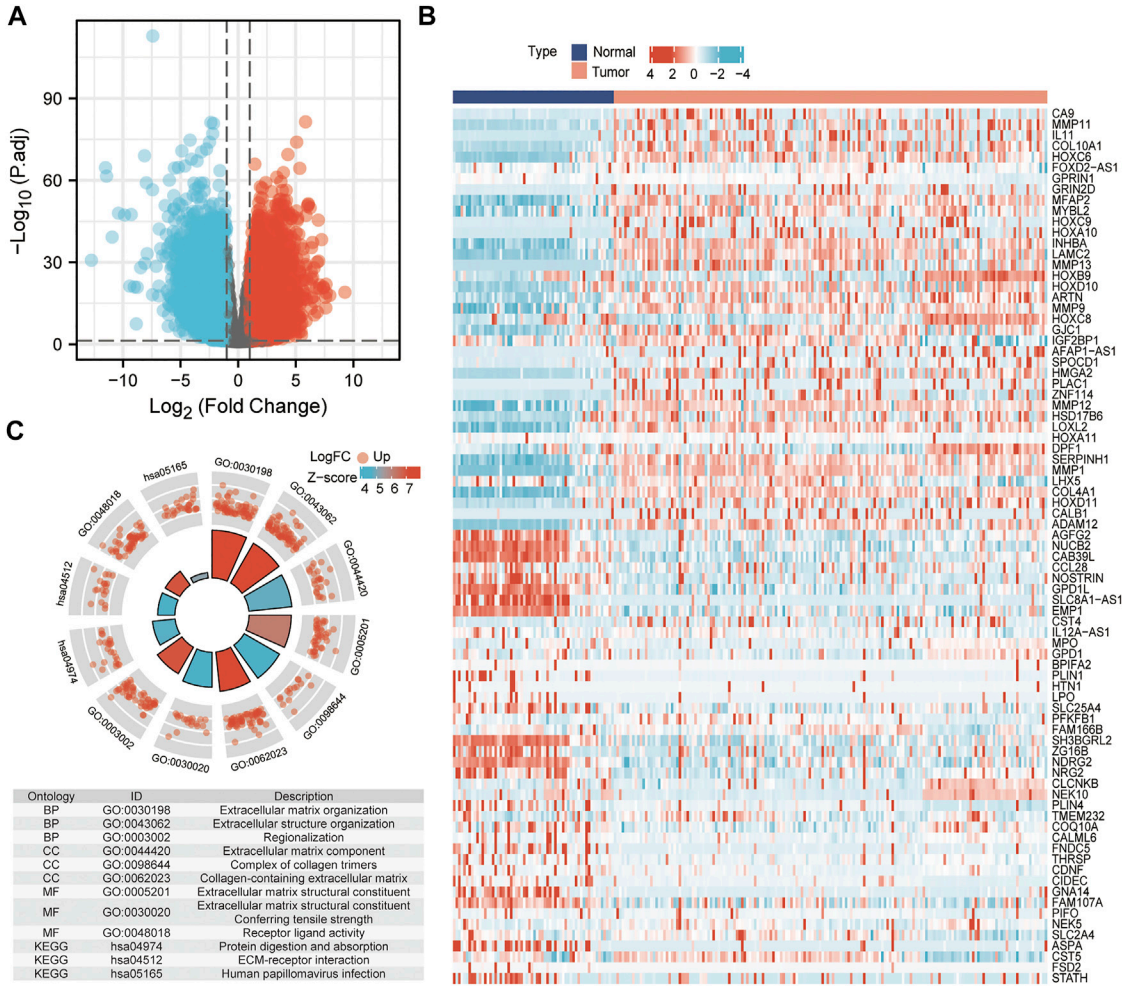
The enrichment of gene ontology (GO) in terms of upregulated DEGs and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways was analyzed using the Goplot and ggplot2 package of R. P-value <0.05 was considered significant. GO is a tool widely used for exploring the potential functions of targets, including molecular functions, biological pathways, and cellular components.

Identification of Prognosis-Related Differentially Expressed Genes

The DEGs were identified by "Limma" package of R software (version 3.6.3). The results were filtered using adjusted *p*-value < 0.05 and $|\log_2^{(\text{fold change})}| > 2$ as the threshold to eliminate the potential false-positive results from analysis of TCGA or GTEx. The volcano plot and heat map were plotted using the R package "ggplot2." In addition, the HNSCC patients were split into a high expression group and a low expression group based on median expression levels of genes. The potential prognostic genes that related to overall survival (OS) in HNSCC were screened using the univariate Cox analysis. The potential prognosis-related DEGs (key genes) were obtained using the R package "VennDiagram." Time-dependent receiver operating characteristics (ROC) curves for 1, 3, and 5 years were plotted to estimate the predictive power of the key genes. The significance of prognosis and time-dependent ROC curve was further integrated to identify a key gene for follow-up research.

Analysis of the Significance of Stanniocalcin-2 Expression in Prognosis of Head and Neck Squamous Cell Carcinoma and Pan-Caner

The expression level of STC2 between normal tissue and tumor tissues in HNSCC patients was evaluated. To investigate the diagnostic performance of STC2, the ROC curve was plotted and the area under the curve (AUC) was calculated. An



Ontology	ID	Description
BP	GO:0030198	Extracellular matrix organization
BP	GO:0043062	Extracellular structure organization
BP	GO:0003002	Regionalization
CC	GO:0044420	Extracellular matrix component
CC	GO:0098644	Complex of collagen trimers
CC	GO:0062023	Collagen-containing extracellular matrix
MF	GO:0005201	Extracellular matrix structural constituent
MF	GO:0030020	Extracellular matrix structural constituent Conferring tensile strength
MF	GO:0048018	Receptor ligand activity
KEGG	hsa04974	Protein digestion and absorption
KEGG	hsa04512	ECM-receptor interaction
KEGG	hsa05165	Human papillomavirus infection

FIGURE 1 | Identified differentially expressed genes. **(A)** Volcano plot: up-regulated and down-regulated genes were indicated in red dots and blue dots, respectively. **(B)** Heatmap of the differentially expressed genes. **(C)** GO terms and KEGG pathways enrichment analysis.

evaluation of the correlation between STC2 expression level and clinicopathological parameters (e.g., survival status, clinical stage, histologic grade, and T stage) was performed by Pearson's chi-square test. In addition, we further combined the GTEx database to explore the expression level of STC2 in 33 cancers and corresponding normal tissues to offset the deficiency of a small sample size of normal non-cancerous tissue in TCGA. The Human Protein Atlas (HPA) (https://www.proteinatlas.org/) database was used to illustrate the STC2 protein distribution among normal and cancer tissues. HPA is one of the world's most frequently visited biological databases. It can provide researchers with high-quality immunohistochemistry images of proteins in the tissue atlas and pathology atlas panels (Digre and Lindskog 2021).

Survival Analysis

The association of STC2 with OS, disease-free survival (DFS), disease-specific survival (DSS), and progress-free survival (PFS) was evaluated using Cox regression in various cancer types. According to expression levels of STC2, the HNSCC patients in TCGA were divided into two groups (high and low expression levels) based on the median expression level of STC2 as a cutoff. The analysis process was realized *via* "survival" and "survminer" packages of R software. The results were visualized using the "forest plot" R package.

Immunological Correlation Analysis

To evaluate the association between STC2 expression level and tumor microenvironment (TME) in 33 tumors, the immune scores and stromal scores of each sample were analyzed *via* the R package "ESTIMATE." The relationship between STC2 expression and immune checkpoint genes in the TME was examined *via* the online platform "SangerBox" (http://sangerbox.com/Tool). p-value < 0.05 was considered statistically significant.

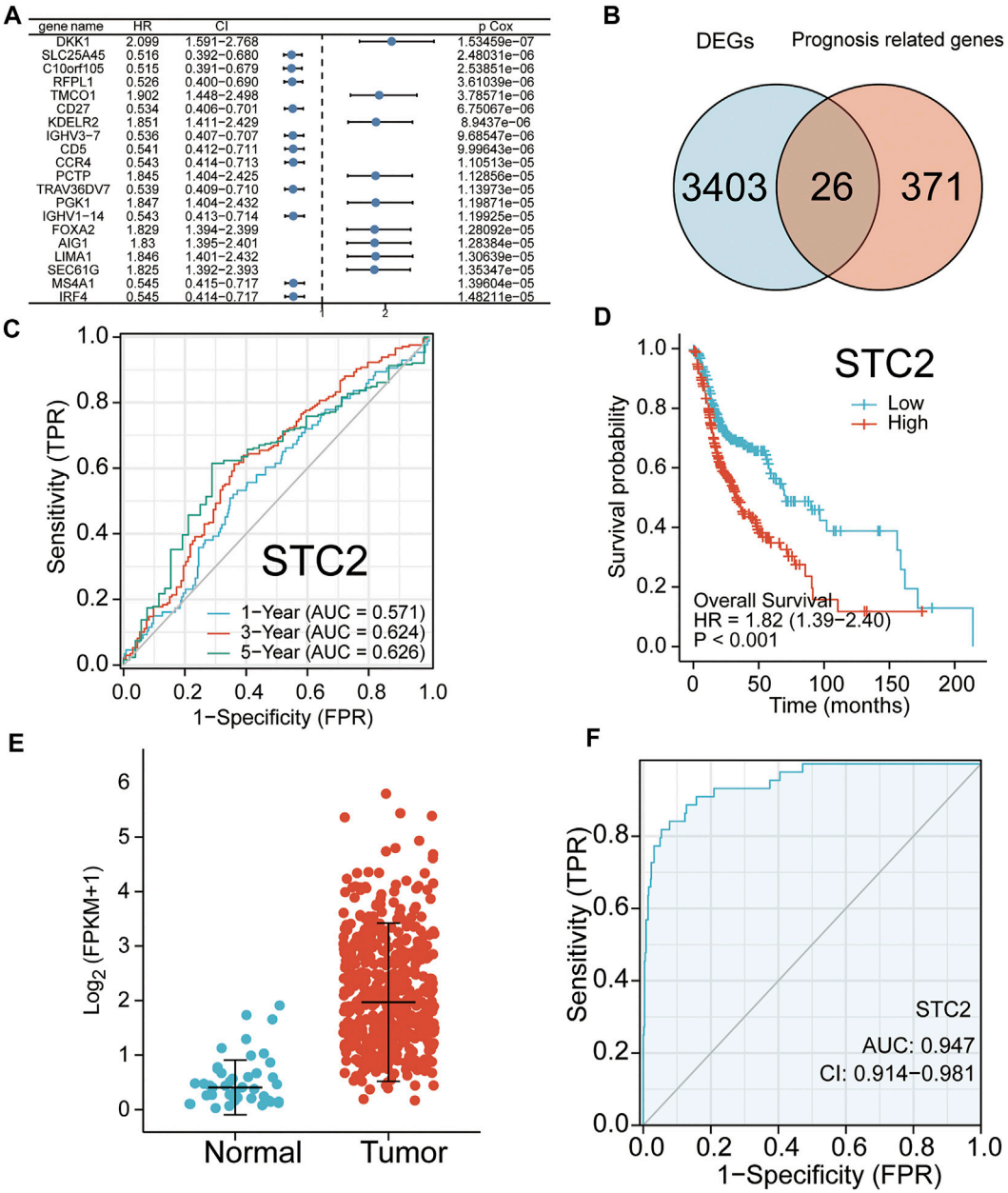


FIGURE 2 | Identified key genes and their examined prognostic significance. **(A)** Top 20 genes with the highest correlation to overall survival (OS) are shown. **(B)** Crossing the 3,429 DEGs with 397 prognosis genes. **(C)** Kaplan–Meier curve of STC2. **(D)** Time-dependent receiver operating characteristics (ROC) curves of STC2 for 1, 3, and 5 years **(E)** STC2 mRNA levels in HNSCC tissues ($n = 502$) and normal tissues ($n = 44$). **(F)** Diagnostic significance of STC2 in HNSCC. ROC curves were plotted and the area under the curve (AUC) of STC2 was 94.7%.

Correlation Analysis of Stanniocalcin-2 With Tumor Mutation Burden, Microsatellite Instability, and Mismatch Repair Genes

TMB, MSI, and MMRs have been proved to be important biomarkers of the TME. The data of five MMRs (EPCAM, PMS2, MLH1, MSH2, and MSH6) and MSI were obtained from the TCGA database. The correlation between STC2 expression levels and MMR gene expression levels, MSI, and TMB was evaluated using the Pearson correlation coefficient. Differences with a p-value < 0.05 were considered to be statistically significant.

Gene Set Enrichment Analysis

GSEA is a computational method that determines whether an *a priori* defined set of genes shows statistically significant, concordant differences between two biological states. The samples were divided into two groups (high- and low-expression groups) based on the

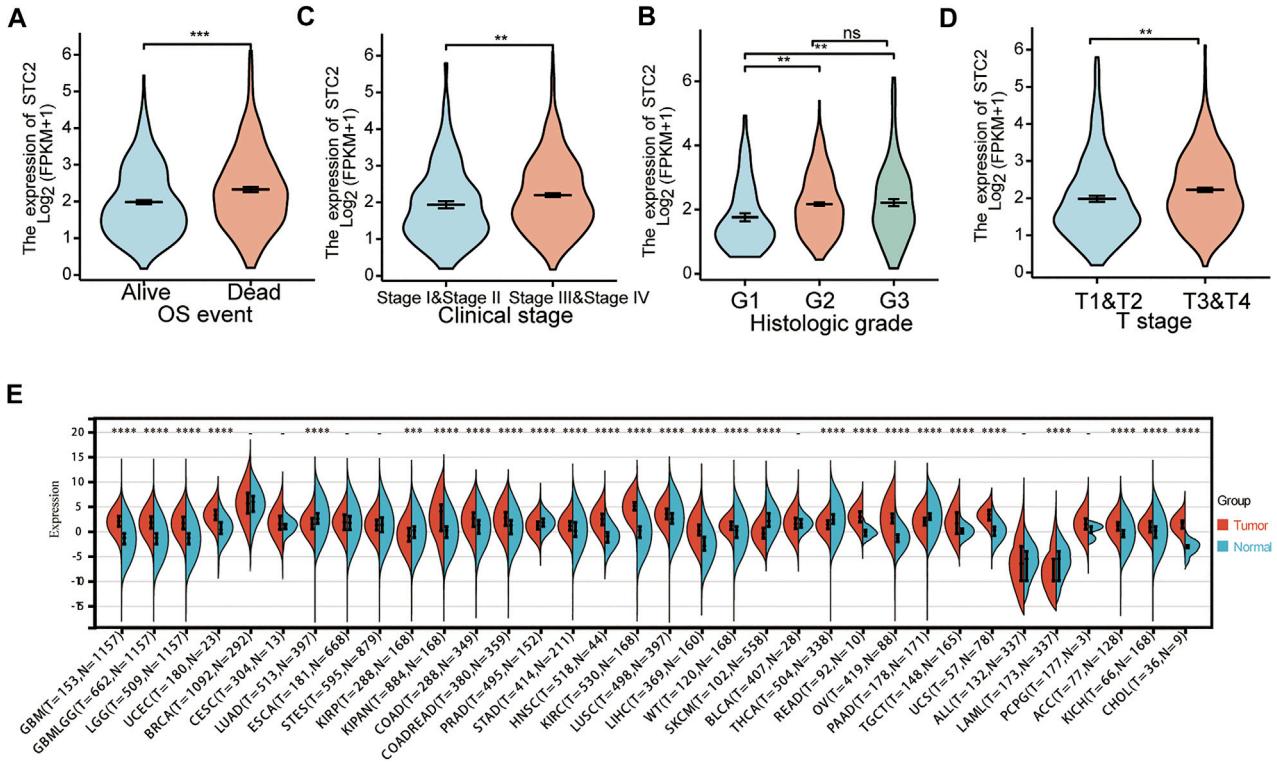


FIGURE 3 | Analysis of the correlation between STC2 expression and clinicopathology. **(A)** Expression levels of STC2 in different survival status of HNSCC. **(B)** Expression levels of STC2 in early and advanced cancer. **(C)** Expression levels of STC2 in different degrees of tumor differentiation. **(D)** Expression levels of STC2 in different tumor stages. **(E)** Expression levels of STC2 in 33 cancer types. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

STC2 expression levels. KEGG and HALLMARK pathways in the high- and low-expression groups were analyzed using GSEA. The functions involved in both groups were yielded by setting a random combination of 1,000 analyses. A normalized enrichment score (NES) was calculated.

Drug Sensitivity Analysis

To clarify the influence of STC2 expression levels on drug sensitivity and drug tolerance, the processed gene expression and drug sensitivity data were downloaded from the CellMiner database (https://discover.nci.nih.gov/cellminer/). Drugs without FDA approval or clinical trials were removed. The correlation was analyzed using Pearson's correlation coefficient test, and the cut-off *p*-value was set to 0.05. The results were visualized using the "limma" and "ggplot2" packages.

Genetic Alteration Analysis

To clarify basic characteristics of STC2, the GeneCards database (https://www.genecards.org/) and COMPARTMENTS (https://compartments.jensenlab.org/Search) database were used to find the genomic location and subcellular location of STC2. The cBioPortal web (https://www.cbioportal.org/) was further utilized to explore the genetic alteration of STC2. All TCGA tumor samples were selected and STC2 mutation rates were queried in different cancers by entering "STC2" in the "TCGA firehose legacy studies." In addition, a three-dimensional structure diagram of STC2 was obtained.

RESULTS

Differentially Expressed Genes in Head and Neck Squamous Cell Carcinoma

According to the threshold which set adjusted *p*-value < 0.05 and $|\log_2^{(\text{fold change})}| > 2$, a total of 3,429 DEGs (1,861 up-regulated genes and 1568 down-regulated) were identified, as visualized by volcano plots (**Figure 1A**). To verify the results, an expression profile (GSE30784) of 229 sample data was downloaded from the GEO database. The top 50 most up- and down-regulated genes were plotted in a heat map based on the sample RNA-seq data in GSE30784 (**Figure 1B**). The results indicated that the screening result of the differential analysis was stable and reliable in the present study. We used GO and KEGG functional enrichment analysis to annotate the function of DEGs (**Figure 1C**). It was found that the genes significantly upregulated were associated with structural roles for cells.

Prognosis-Related Genes and Their Prognostic Significance for Head and Neck Squamous Cell Carcinoma

A total of 397 prognosis-related genes were identified based on the univariate Cox analysis ($p < 0.001$). The top 26 genes with the largest significant difference were shown in the forest plot (**Figure 2A**). The key genes were filtered by crossing the 3,429 DEGs with the

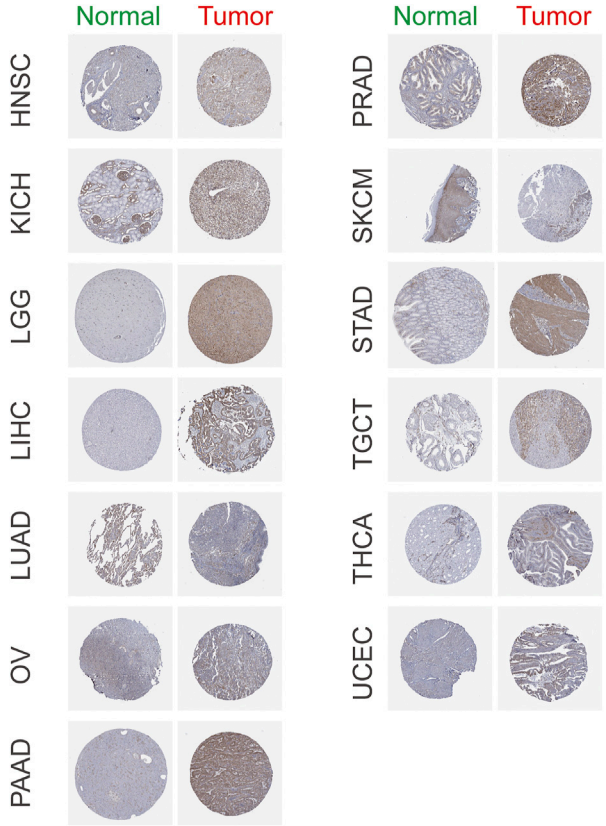


FIGURE 4 | Representative immunohistochemistry images of STC2 in HNSC, KICH, LGG, LICH, LUAD, OV, PAAD, PRAD, SKCM, STAD, TGCT, THCA, and UCEC based on The Human Protein Atlas.

397 prognosis-related genes. The screening result is shown using a Venn plot (**Figure 2B**). The time-dependent ROC curve of all key genes is shown in **Figure 2C** and **Supplementary Figure S1**. Integrating the significance of prognosis and time-dependent ROC curve, the mRNA STC2 was selected for subsequent studies. The survival curve of STC2 is shown in **Figure 2D**. Our study analyzed the expression of STC2 in HNSCC. Compared with normal tissues (n = 44), the STC2 expression level was significantly increased in cancerous tissues (n = 502) (**Figure 2E**). The ROC curves confirmed that upregulation of STC2 played a significant performance in the efficacy of HNSCC diagnosis (AUC = 0.947) (**Figure 2F**). Furthermore, the present study analyzed the correlation of STC2 expression with clinicopathology. The results showed that the STC2 expression level was significantly up-regulated in the HNSCC patients of dead, stage III–IV, G2–G3, and T3–T4 (**Figures 3A–D**).

The Expression of Stanniocalcin-2 in Pan-Cancer

The STC2 expression levels were analyzed in multiple tumors from the TCGA and GTEx databases. The results showed that the expression level of STC2 was significantly different among 27 tumors, except breast invasive carcinoma (BRCA), cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), esophageal carcinoma (ESCA), stomach and esophageal carcinoma (STES), bladder urothelial carcinoma (BLCA), acute lymphoblastic leukemia (ALL), and pheochromocytoma and paraganglioma (PCPG) (**Figure 3E**). Using the HPA database, we found that the expression level of STC2 was significantly increased in HNSC, LGG, LIHC, OV, PAAD, PRAD, STAD, TGCT, THCA, and UCEC. However, STC2 expression was significantly decreased in LUAD and SKCM (**Figure 4**).

Analysis of Association of Stanniocalcin-2 Expression With Survival in Pan-Cancer

The prognostic significance of STC2 in different cancers was evaluated using Cox regression analyses of the TCGA database. The results showed that high expression levels of STC2 were associated with poorer OS in BLCA, COAD, ESCA, HNSC, kidney renal papillary cell carcinoma (KIRP), liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), mesothelioma (MESO), sarcoma (SARC), and thymoma (THYM), but predicted better prognosis in BRCA and brain lower-grade glioma (LGG) (**Figure 5A**). The correlation analysis of disease-free survival (DFS) found that the STC2 expression level was significantly associated with the prognosis of BRCA, KIRP, LGG, LIHC, and PRAD (**Figure 5B**). The DSS analysis revealed that high expression of STC2 was significantly associated with poor prognosis in 10 tumors and improved prognosis in two tumors (**Figure 5C**). According to data on PFS, the high expression of STC2 was significantly associated with prognosis in 12 out of 32 tumors (**Figure 5D**).

Analysis of the Correlation Between Stanniocalcin-2 and Immunity

To evaluate the role of STC2 in the tumor microenvironment, we analyzed the relationship between STC2 expression level and 60 common ICP gene expressions, including 24 inhibitory and 36 stimulatory. The results showed that correlation between STC2 and ICP genes varied across tumor types (**Figure 6A**). STC2 expression level was positively correlated with 60 ICP genes in most tumors, including HNSC, BLCA, acute myeloid leukemia (LAML), stomach adenocarcinoma (STAD), and uterine carcinosarcoma (UCS). However, STC2 expression levels were inversely correlated with 60 ICP genes in testicular germ cell tumors (TGCT). It is suggested that STC2 might be involved in the tumor immune microenvironment by coordinating the activity of specific ICP genes. Our study further explored the association of STC2 with immune cell infiltration in 33 tumors. Immune infiltration analysis showed that the expression level of STC2 was associated with the level of immune cell infiltration (including B cell, $CD4^+$ T cells, $CD8^+$ T cells, neutrophils, macrophages, and dendritic cells) (**Figure 6B**). Especially in PCPG, kidney chromophobe (KICH), LIHC, there was a strongly positive correlation between the expression level of STC2 and the immune cell infiltration level. In addition, the immune score and stromal score were calculated by the ESTIMATE algorithm (**Figure 6C**). The results showed that the STC2 expression levels was significantly correlated with immune score in lymphoid neoplasm diffuse large B-cell lymphoma (DLBC)

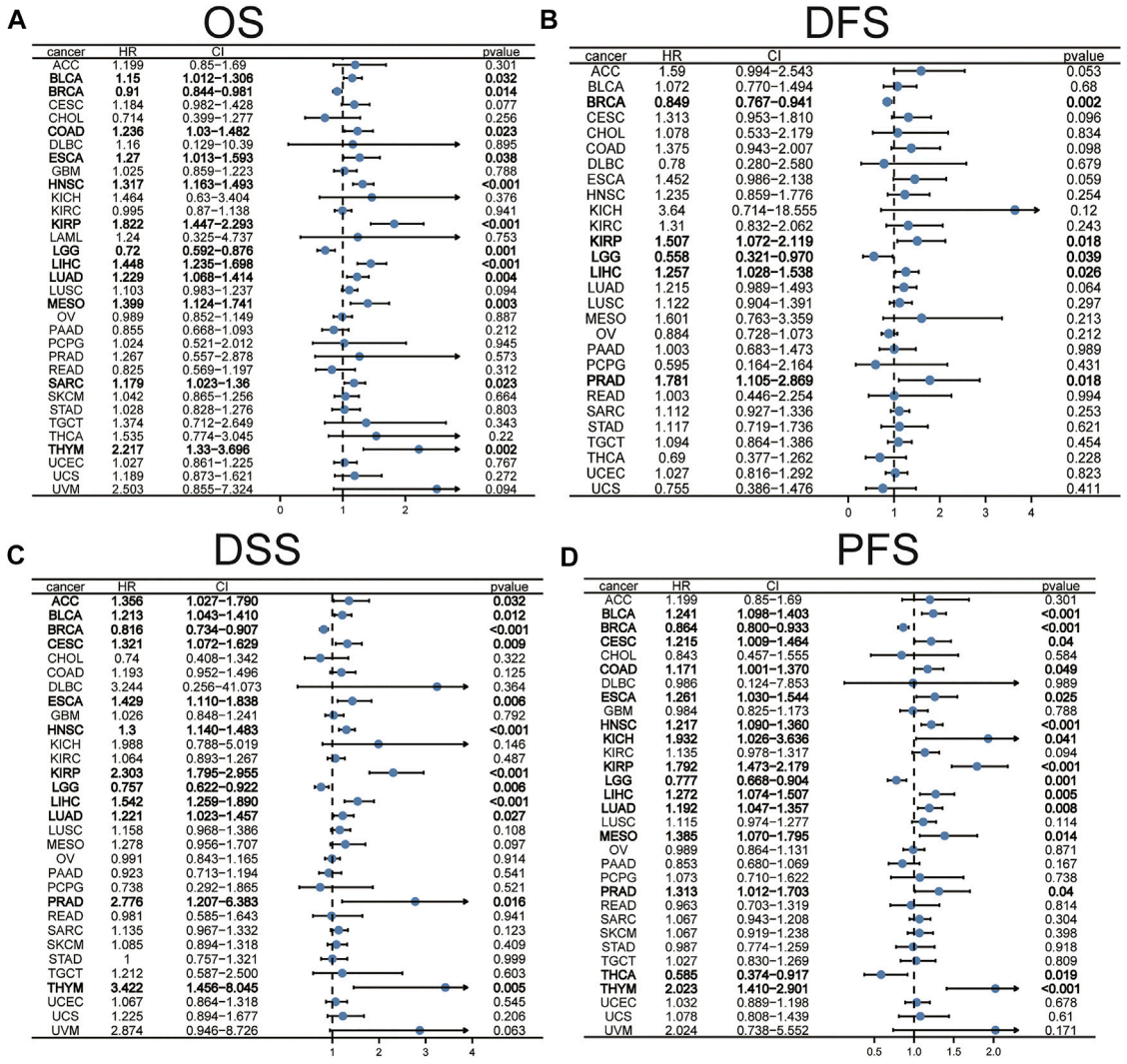


FIGURE 5 | The prognostic significance of STC2 in pan-cancer. The associations between STC2 expression levels and overall survival **(A)**, disease-free survival **(B)**, disease-specific survival **(C)**, and progress-free survival **(D)** in various cancer types were illustrated by the forest plots.

(r = 0.34, p = 0.017), KICH (r = 0.38, p = 0.0018), MESO (r = 0.32, p = 0.0024), pancreatic adenocarcinoma (PAAD) (r = 0.36, p < 0.001), pheochromocytoma and paraganglioma (PCPG) (r = 0.54, p < 0.001), and PRAD (r = 0.31, p < 0.001). The STC2 expression level was significantly correlated with stromal score in LUSC (r = 0.33, p < 0.001), PCPG (r = 0.38, p < 0.001), thyroid carcinoma (THCA) (r = 0.32, p < 0.001), TGCT (r = 0.47, p < 0.001), and UVM (r = 0.33, p = 0.0028).

Analyses of Correlation Between Stanniocalcin-2 Expression Levels and Mismatch Repair Genes, Tumor Mutation Burden, and Microsatellite Instability

The results of correlation analysis of STC2 expression level with five MMRs are shown in **Figure 7A**. The results indicated that expression level of STC2 was significantly correlated with MMRs in almost all cancers except BLCA, cholangiocarcinoma (CHOL), DLBC, glioblastoma multiforme (GBM), KICH, LAML, and SARC. The correlation was most significant in BRCA, CESC, LIHC, LUSC, PRAD, and rectum adenocarcinoma (READ). Our study further investigated the relationship between STC2 expression and TMB. The association between STC2 and TMB was significant in 15 cancer types (**Figure 7B**). Specifically, the significant and positive correlation between STC2 expression level and TMB was found in HNSC (r = 0.13, p < 0.01), LAML (r = 0.21, p < 0.05), KIRC (r = 0.13, p < 0.05), LUAD (r = 0.19, p < 0.001), PRAD (r = 0.12, p < 0.01), READ (r = 0.22, p < 0.05), SARC (r = 0.17, p < 0.05), THYM (r = 0.26, p < 0.01), and UCEC (r = 0.12, p < 0.01). Conversely, STC2 expression levels were negatively correlated with TMB in BRCA (r = −0.34, p < 0.001), ESCA(r = 0.24, p < 0.01), KIRP (r = −0.12, p < 0.05), PAAD (r = −0.24, p < 0.01), skin cutaneous melanoma (SKCM) (r = −0.10, p < 0.05),

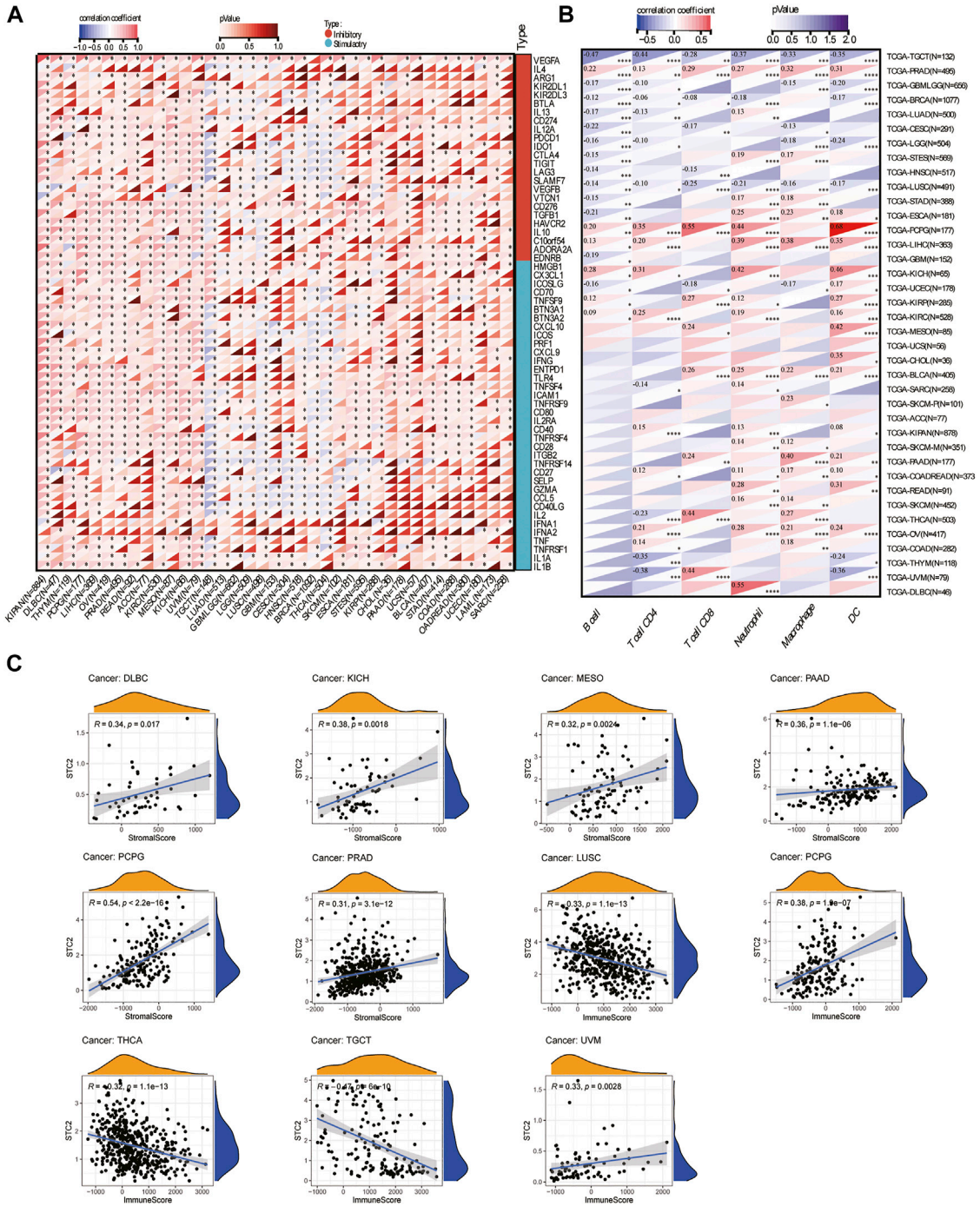


FIGURE 6 | Immunological correlation analysis. **(A)** The correlation between STC2 expression and immune checkpoint genes. **(B)** Correlation of STC2 expression with immune cell infiltration in 33 cancers. Blue represents negative correlation and red represents positive correlation. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$. **(C)** Correlation between STC2 and stromal scores and immune score.

and THCA (r = −0.12, $p < 0.01$). In another five types of adrenocortical carcinoma (ACC), BRCA, HNSC, LUSC, and TGCT, STC2 expression was related to MSI (**Figure 7C**). The results demonstrated that upregulated STC2 expression was significantly associated with increased TMB in ACC (r = 0.22, $p < 0.05$), HNSC (r = 0.12, $p < 0.01$), LUSC (r = 0.17, $p < 0.001$), and TGCT (r = 0.19, $p < 0.05$), while it was associated with decreased TMB in BRCA (r = 0.11, $p < 0.001$).

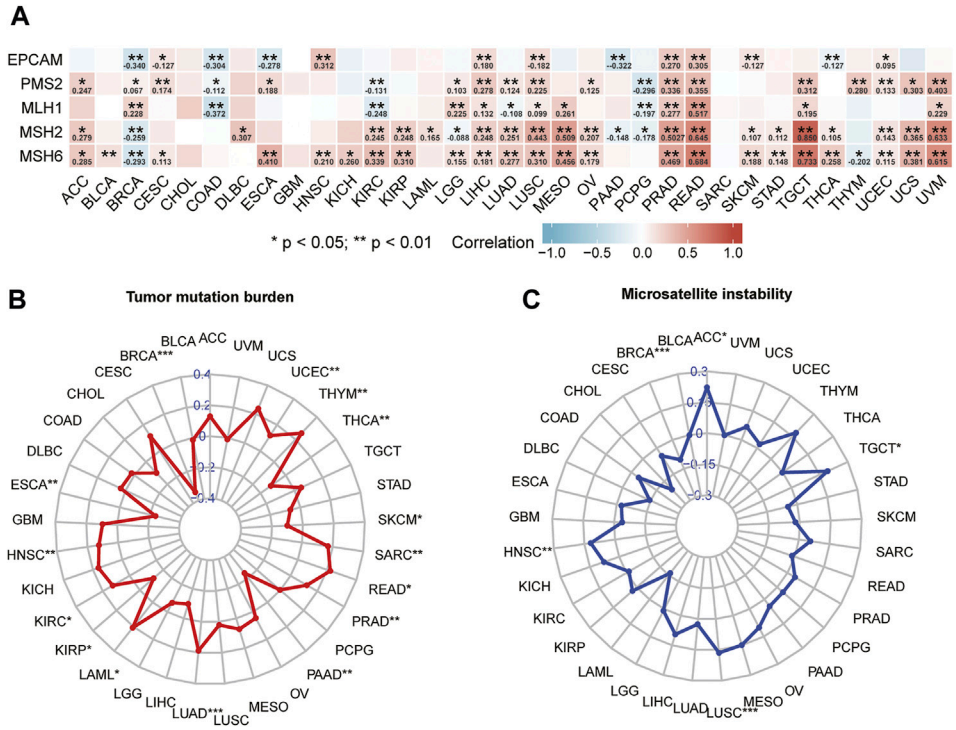


FIGURE 7 | Analysis of correlation between STC2 and mismatch repair genes **(A)**, tumor mutation burden **(B)**, and microsatellite instability **(C)**.

Potential Functions of the Stanniocalcin-2–Correlated Genes

The GSEA enrichment analysis revealed that STC2 was the focus of many signaling pathways associated with cancer. The KEGG pathway enrichment analysis of upregulated STC2 revealed a significant enrichment of linoleic acid metabolism, olfactory transduction, RIG-I–like receptor signaling pathway, and intestinal immune network for IgA production (**Figures 8A–D**). The HALLMARK pathway analysis revealed that allograft rejection, epithelial–mesenchymal, interferon α response, and interferon γ response were enriched in high STC2 (**Figures 8E–H**).

Drug Sensitivity Analysis

We analyzed the correlation between STC2 expression and drug sensitivity based on the CellMiner database. The top 16 drugs with the strongest correlations were shown in **Figure 9**. The results suggested that upregulation of STC2 expression reduced the sensitivity to a number of drugs, including PF-4942847, MPC-3100, SNX-5422, TAK-632, TAS-116, BGB-283, geldanamycin analog, alvespimycin, AT-13387, dabrafenib, XL-888, tanespimycin, ganetespib, CUDC-305, ARQ-680, and SB-590885.

Basic Biology of Stanniocalcin-2 and Genetic Alteration Analysis

STC2 is a protein-coding RNA located at q35.2 in Chromosome 5 (**Figure 10A**). The subcellular location of STC2 was found using COMPARTMENTS. The highest confidence in the subcellular location was in the endoplasmic reticulum and perinuclear region of the cytoplasm (**Figure 10B**). The cBioportal was used to explore the genetic alteration of STC2. The results revealed 200 mutations of STC2 in 8,681 patients with different tumors. Among them, renal cell carcinoma patients had the highest frequency of STC2 gene alteration (15.82%) and the most alterations were amplification (15.44%, 82/531). The mutation frequency of adrenocortical carcinoma is the highest at about 2.17% (**Figure 10C**). We present the three-dimensional (3D) structure of the STC2 protein in **Figure 10D**.

DISCUSSION

STC2 as a glycoprotein is expressed in the broad-spectrum of human tumor cells and tissues, including breast, bone, esophagus, liver, lung, kidney, ovary, prostate, and stomach (Qie and Sang 2022). Under stress conditions such as hypoxia (Law et al., 2008; Law and Wong 2010), ER stress (Joshi et al., 2015), and nutrient deprivation (He et al., 2019), both transcriptional and post transcription of STC2 were regulated, and in turn drives tumor cell growth, proliferation, and tumorigenesis. Moreover, STC2 is identified as an immune response–related gene and upregulated in T-cells with Th2 response (Jansen et al., 2015; Roche et al., 2018). Importantly, STC2 mediates resistance to chemo- and radio-therapies and promotes the development of tumor cell–acquired resistance (Garnett et al., 2004). Hence, we investigated the correlations of STC2 expression level with

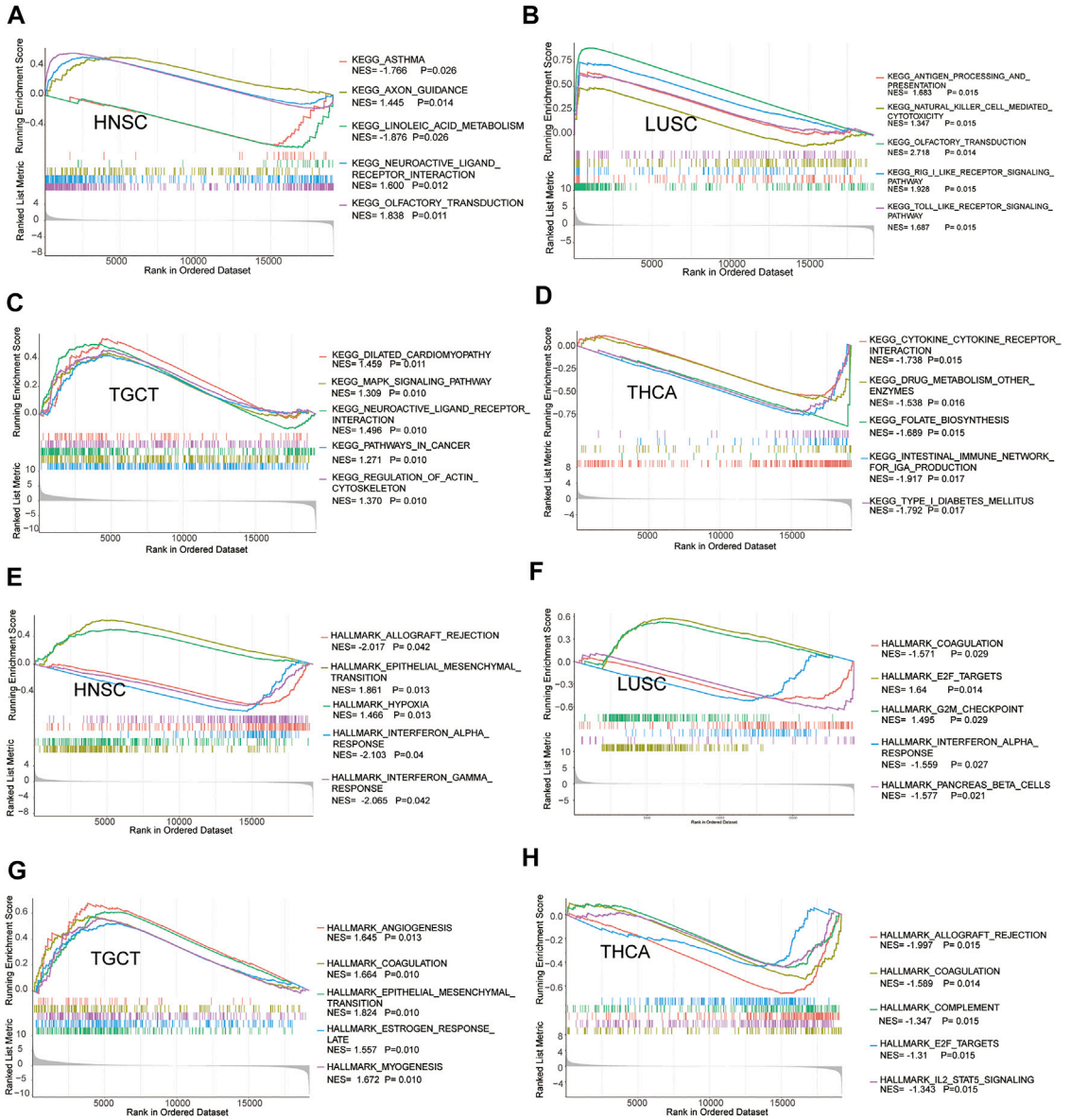


FIGURE 8 | Gene set enrichment analysis of STC2. Enrichments of KEGG pathways in HNSC **(A)**, LUSC **(B)**, TGCT **(C)**, and THCA **(D)**. Enrichments of HALLMARK pathways in HNSC **(E)**, LUSC **(F)**, TGCT **(G)**, and THCA **(H)**.

prognosis, immune checkpoint, MMRs, TMB, MSI, function notation, and related pathways in 33 distinct tumors based on the data from TCGA and GTEx.

First, we assessed the mRNA expression levels of STC2 in tumors. The results revealed that the expression level of STC2 was significantly up-regulated in most cancers. Furthermore, we investigated the clinical prognostic value of STC2 expression levels in tumors by univariate Cox regression. In line with previous reports, high expression of STC2 suggested an unfavorable prognosis in BLCA, COAD, ESCA, HNSC, KIRP, LIHC, LUAD, MESO, SARC, and THYM. In the progression of COAD, STC2 contributed to cancer cell metastasis by promoting epithelial–mesenchymal transition (Chen et al., 2016). Moreover, STC2 also plays a critical role in the progression of HNSCC. STC2 enhanced PI3K/Akt signaling pathways by upregulation of Snail-mediated decrease of E-cadherin, and promoted HNSCC metastasis (Yang et al., 2017). The overexpression of STC2 leads to poor prognosis in LIHC. STC2 could regulate drug efflux and drug resistance of tumor cells by regulating the expression levels of P-glycoprotein and Bcl-2 protein (Shi et al., 2020). In contrast, favorable prognostic indications were found in BRCA and LGG.

As integral components of the tumor immune microenvironment, tumor-infiltrating immune cells (TIICs) were biomarkers for prognosis and response to immunotherapy in various carcinomas (Zhang and Zhang 2020). Many markers including immune checkpoints, TMB,

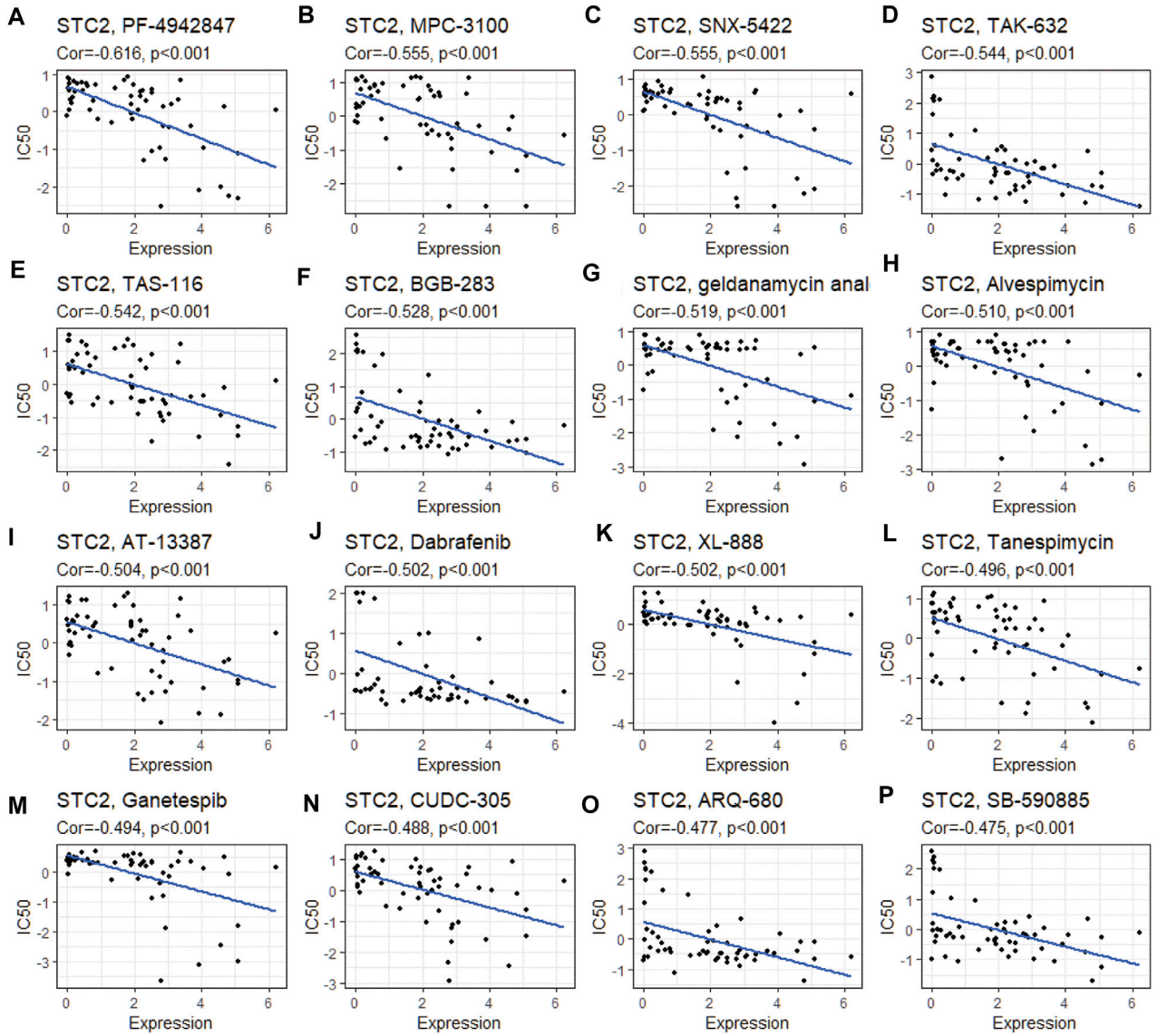


FIGURE 9 | Drug sensitivity analysis of STC2. The expression of STC2 was associated with the drug sensitivity of PF-4942847 **(A)**, MPC-3100 **(B)**, SNX-5422 **(C)**, TAK-632 **(D)**, TAS-116 **(E)**, BGB-283**(F)**, geldanamycin analog **(G)**, alvespimycin **(H)**, AT-13387 **(I)**, dabrafenib **(J)**, XL-888 **(K)**, tanespimycin **(L)**, ganetespib **(M)**, CUDC-305 **(N)**, ARQ-680 **(O)**, and SB-590885 **(P)**.

and MSI have been used to evaluate immunotherapy susceptibility (Vrána et al., 2018; Li et al., 2020). Therefore, we investigated the relationship between STC2 expression level and TIICs, immune scores, and stromal scores in multiple malignancies. We observed a significant positive correlation between STC2 expression level and infiltration levels of B cells, $CD4^+$ T cells, $CD8^+$ T cells, neutrophil, macrophage, and dendritic cells in PRAD, LIHC, KICH, and KIRC. STC2 was an interaction partner of histidine-rich glycoprotein (HRG) located on the surface of inflammatory cells. HRG was a cation- and heparin-binding, 75 kDa plasma glycoprotein produced by liver hepatocytes (Poon et al., 2011), which could generate antitumor immune responses by controlling the phenotypic switch of macrophages. STC2 could form a histidine-rich glycoprotein/stanniocalcin-2 (HRG/STC2) complex with HRG to regulate murine glioma growth *via* modulation of anti-tumor immunity (Roche et al., 2018). Moreover, the previous animal model study showed that STC2 knockdown in murine mesenchymal stem cells significantly impairs its effect in reducing TNF-α– and IFN-γ–producing $CD8^+$ T cells (Chen et al., 2018). It is well known that T cells recognize tumor antigens and delivering them to T-cell receptors is a major process in tumor immune regulation (Grinberg-Bleyer et al., 2017). When T-cell function is impaired,

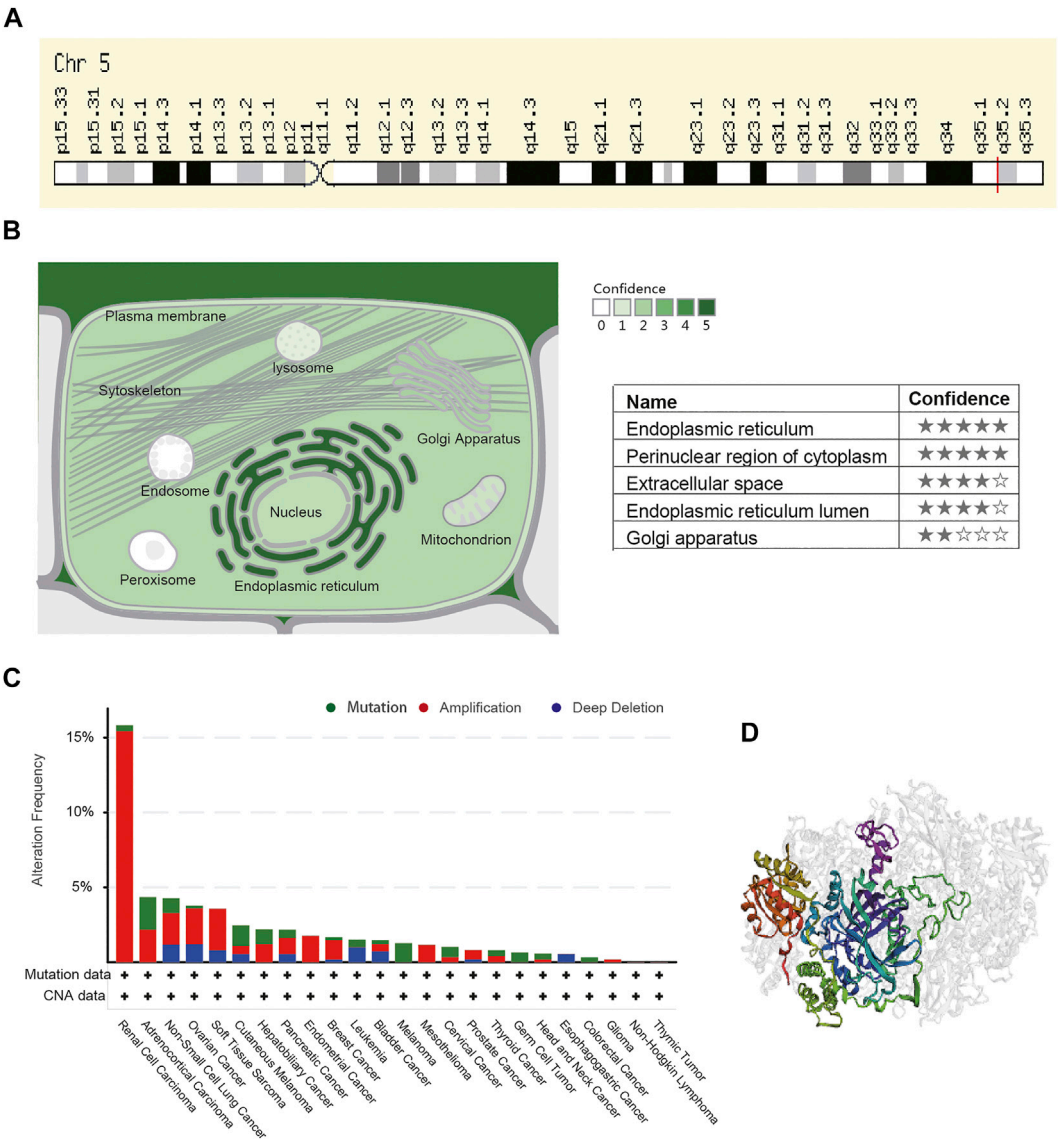


FIGURE 10 | Basic biology of STC2 and genetic alteration analysis. **(A)** Genomic location of STC2 gene in the GeneCards website. **(B)** Subcellular location of STC2 by COMPARTMENTS. **(C)** Mutation type and frequency of STC2. **(D)** Three-dimensional (3D) structures of STC2.

cancer cells can evade immune surveillance by multiple mechanisms. Among them, cancer cells exploit various intrinsic immune checkpoints of T cells to escape from immune surveillance; thus, they are important pathways for escaping (Tan et al., 2017). The immune checkpoints are regulators of the immune system and can attenuate T-cell responses to maintain immune homeostasis and prevent autoimmunity (Liu et al., 2020). However, tumors can use various mechanisms to hijack these pathways to evade immune recognition against tumor antigen-specific T cells (Cohen et al., 2016). Therefore, the expression levels of immune checkpoint genes and their ligands are often up-regulated in the tumor immune microenvironment (Zhao et al., 2018). A significant positive correlation between the expression level of STC2 and the expression levels of immune checkpoint genes was observed in most cancers in our study. These findings suggested that STC2 may play an important role in tumor immune avoidance mechanisms. However, the *in vivo* biological importance of STC2 remains to be further dissected, which will be a target for our future study.

MMRs are key players in a post-replication repair system and are involved in repairing DNA base mismatches and indels during the S phase. As a safety and security system, MMRs play a crucial role in ensuring the stability of DNA replication and avoiding the generation of genetic mutations. Mutations or defects in the MMR genes could contribute to the accumulation

of genetic errors, resulting in the occurrence of tumors (Geisler et al., 2003). Therefore, we assessed in detail the association of five MMRs (MLH1, MSH2, MSH6, PMS2, and EPCAM) with STC2 expression levels in 33 cancers. The results of Pearson analysis confirmed that STC2 expression level was significantly and positively correlated with DNA repair genes in most types of cancer. MSI is another marker of genomic instability and is a hallmark of MMR defects (Law and Wong 2010). MSI is closely linked to the occurrence, development, and prognosis of various cancers. A study by Gryfe et al. (2000) shows that high-frequency MSI is an independent predictor of clinical characteristics and prognosis in COAD. Compared with low levels of MSI, high levels of MSI exhibit a better anti-tumor immune response, the ability to inhibit tumor cell growth, improved prognosis, and microsatellite stability (Kim et al., 2016; Marrelli et al., 2016; Mohan et al., 2016). TMB might influence the utility of immunotherapy by affecting the generation of immunogenic peptides. Previous studies have revealed that TMB improves the efficacy of immunotherapy in colorectal cancer (Lee et al., 2019) and non-small cell lung cancer (Devarakonda et al., 2018). Therefore, we investigated the correlation between STC2 expression levels with TMB and MSI in 33 tumors. An important finding was that STC2 expression was significantly associated with TMB in 15 cancers and MSI in five cancers. The results suggest that STC2 might serve as a promising tumor therapeutic target to improve the immunotherapy response.

The lack of tumor specificity and drug resistance remains a significant barrier to effective cancer therapy and might be associated with tumor recurrence. Previous studies indicate that STC2 is upregulated upon exposure to several chemotherapeutic treatments. Miyazaki et al. (2014) found that STC2 was the most highly upregulated gene in anti-VEGF antibody–treated colon cancer. This indicated that STC2 might contribute to the development of acquired drug resistance to anti-VEGF treatment. Wang et al. (2015) found that STC2 could promote cell proliferation and cisplatin resistance in cervical cancer by regulating the activity of the MAPK signaling pathway. STC2 promoted activation of the PI3K–Akt signaling pathway and induced the expression of P-glycoprotein which also increases the drug resistance of tumor cells (Yuan et al., 2016). P-glycoprotein was an efflux pump involved in the absorption, distribution, and excretion of chemical compounds and implicated in the poor penetration of many drugs (Brown et al., 2018). The drug sensitivity analysis based on the CellMiner database found that STC2 expression levels are significantly associated with some drug sensitivity, such as PF-4942847, MPC-3100, SNX-5422, TAK-632, TAS-116, BGB-283, geldanamycin analog, and alvespimycin. It is likely that STC2 upregulated might play an important role in drug resistance.

Genetic alteration is a main factor that drives cancer, including altering genetic content, disrupting genes, and causing phenotypic differences. Cancer genomes typically develop four to five driver mutations when combining coding and noncoding genomic elements, and their instability is a molecular genetic hallmark of tumorigenesis (Bray et al., 2018). Gene amplification is a major manifestation of genomic instability and plays a crucial role in various cancers. In recent years, growing evidence has indicated that mutated-gene–targeted therapy is a promising therapy for tumors (Synnott et al., 2017; Kaur et al., 2018). Our study found that STC2 is mutated in most cancers, and the frequency of STC2 gene mutation in renal cell carcinoma patients reached 15.2%; most of the alterations were amplification. This result further demonstrates that STC2 is a plausible therapeutic target.

To our knowledge, this study is the first comprehensive pan-cancer analysis of STC2. This study systematically investigated pan-cancer information from multiple databases. However, there are still inevitable limitations in our study. First, we did not find that STC2 was differentially expressed between some tumor tissue (BLCA, BRCA, and ESCA) and normal tissue in the previous results of analysis of differentially expressed genes. Therefore, the diagnostic and prognostic value of STC2 in BLCA, BRCA, and ESCA should be further validated with a large number of cases. Second, as our study is limited to the analysis of the existing data, more experimental studies are needed to validate these bioinformatics results. Further research will contribute to clarifying the role of STC2 at the molecular level. Third, although pan-cancer analysis results support that STC2 expression was associated with immunomodulatory mechanisms and immune cell infiltration, the underlying mechanism remains to be further investigated.

CONCLUSION

In conclusion, we analyzed data in public databases and found the abnormal expression of STC2 is significantly correlated with the prognosis of HNSCC and can be used as a biomarker for the prognosis of HNSCC. In pan-cancer analysis, STC2 is significantly associated with prognosis, immune cell infiltration, immune checkpoints, immune scoring, mechanistic scoring, TBM, MSI, and MMRs in multiple cancers. This provides a certain basis for bioinformatics analysis for future studies of STC2 in tumor immunity and tumor microenvironment. However, the aforementioned results still require further biological experiments to verify the function and molecular mechanism.

AUTHOR CONTRIBUTIONS

Study conception and design: Z-HJ and WL. Data collection: Z-HJ, XS, and YW. Statistical analysis: LX and XS. Manuscript draft: Z-HJ and YC. Manuscript review and editing: LX, HC, and WL.

FUNDING

Innovative Research Program for Graduates of the Hubei University of Medicine (No. YC2021046). Scientific research projects of Shiyan Science and Technology department (No. 2021K68).

ACKNOWLEDGMENTS

We thank the Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) for providing the data for this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.927046/full#supplementary-material

REFERENCES

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J. Clin.* 68 (6), 394–424. doi:10.3322/caac.21492

Brown, N. F., Williams, M., Arkenau, H.-T., Fleming, R. A., Tolson, J., Yan, L., et al. (2018). A Study of the Focal Adhesion Kinase Inhibitor GSK2256098 in Patients with Recurrent Glioblastoma with Evaluation of Tumor Penetration of [11C] GSK2256098. *Neuro Oncol.* 20 (12), 1634–1642. doi:10.1093/neuonc/noy078

Chen, B., Zeng, X., He, Y., Wang, X., Liang, Z., Liu, J., et al. (2016). STC2 Promotes the Epithelial-Mesenchymal Transition of Colorectal Cancer Cells through AKT-ERK Signaling Pathways. *Oncotarget* 7 (44), 71400–71416. doi:10.18632/oncotarget.12147

Chen, W. N., and Zhu, G. J. (2008). Progress in the Research of Stanniocalcin. *Sheng Li Ke Xue Jin Zhan* 39 (3), 225–228.

Chen, X., Liu, Q., Huang, W., Cai, C., Xia, W., Peng, Y., et al. (2018). Stanniocalcin-2 Contributes to Mesenchymal Stromal Cells Attenuating Murine Contact Hypersensitivity Mainly via Reducing CD8+ Tc1 Cells. *Cell. Death Dis.* 9 (5), 548. doi:10.1038/s41419-018-0614-x

Cohen, E. E. W., LaMonte, S. J., Erb, N. L., Beckman, K. L., Sadeghi, N., Hutcheson, K. A., et al. (2016). American Cancer Society Head and Neck Cancer Survivorship Care Guideline. *CA A Cancer J. Clin.* 66 (3), 203–239. doi:10.3322/caac.21343

Devarakonda, S., Rotolo, F., Tsao, M.-S., Lanc, I., Brambilla, E., Masood, A., et al. (2018). Tumor Mutation Burden as a Biomarker in Resected Non-small-cell Lung Cancer. *Jco* 36, 2995–3006. doi:10.1200/JCO.2018.78.1963

Digre, A., and Lindskog, C. (2021). The Human Protein Atlas-Spatial Localization of the Human Proteome in Health and Disease. *Protein Sci.* 30 (1), 218–233. doi:10.1002/pro.3987

Ding, Z., Yu, D., Li, H., and Ding, Y. (2021). Prognostic Correlation of Glycolysis-Related Gene Signature in Patients with Laryngeal Cancer. *Am. J. Med. Sci.* 362, 161–172. doi:10.1016/j.amjms.2020.12.021

Garnett, C. T., Palena, C., Chakarborty, M., Tsang, K.-Y., Schlom, J., and Hodge, J. W. (2004). Sublethal Irradiation of Human Tumor Cells Modulates Phenotype Resulting in Enhanced Killing by Cytotoxic T Lymphocytes. *Cancer Res.* 64 (21), 7985–7994. doi:10.1158/0008-5472.CAN-04-1525

Geisler, J. P., Goodheart, M. J., Sood, A. K., Holmes, R. J., Hatterman-Zogg, M. A., and Buller, R. E. (2003). Mismatch Repair Gene Expression Defects Contribute to Microsatellite Instability in Ovarian Carcinoma. *Cancer* 98 (10), 2199–2206. doi:10.1002/cncr.11770

Grinberg-Bleyer, Y., Oh, H., Desrichard, A., Bhatt, D. M., Caron, R., Chan, T. A., et al. (2017). NF-κB C-Rel Is Crucial for the Regulatory T Cell Immune Checkpoint in Cancer. *Cell.* 170 (6), 1096–1108. e1013. doi:10.1016/j.cell.2017.08.004

Gryfe, R., Kim, H., Hsieh, E. T. K., Aronson, M. D., Holowaty, E. J., Bull, S. B., et al. (2000). Tumor Microsatellite Instability and Clinical Outcome in Young Patients with Colorectal Cancer. *N. Engl. J. Med.* 342 (2), 69–77. doi:10.1200/JCO.18.0179810.1056/nejm200001133420201

He, H., Qie, S., Guo, Q., Chen, S., Zou, C., Lu, T., et al. (2019). Stanniocalcin 2 (STC2) Expression Promotes Post-radiation Survival, Migration and Invasion of Nasopharyngeal Carcinoma Cells. *Cmar* Vol. 11, 6411–6424. doi:10.2147/CMAR.S197607

Hou, J., Wang, Z., Xu, H., Yang, L., Yu, X., Yang, Z., et al. (1932). Stanniocalcin 2 Suppresses Breast Cancer Cell Migration and Invasion via the PKC/claudin-1-mediated Signaling. *PLoS One* 10, e0122179–6203. doi:10.1371/journal.pone.0122179

Jansen, M. P. H. M., Sas, L., Sieuwerts, A. M., Van Cauwenberghe, C., Ramirez-Ardila, D., Look, M., et al. (2015). Decreased Expression of ABAT and STC2 Hallmarks ER-Positive Inflammatory Breast Cancer and Endocrine Therapy Resistance in Advanced Disease. *Mol. Oncol.* 9 (6), 1218–1233. doi:10.1016/j.molonc.2015.02.006

Joshi, A. D., Carter, D. E., Harper, T. A., Jr., and Elferink, C. J. (2015). Aryl Hydrocarbon Receptor-dependent Stanniocalcin 2 Induction by Cinnabarinic Acid Provides Cytoprotection against Endoplasmic Reticulum and Oxidative Stress. *J. Pharmacol. Exp. Ther.* 353 (1), 201–212. doi:10.1124/jpet.114.222265

Kaur, R. P., Vasudeva, K., Kumar, R., and Munshi, A. (2018). Role of P53 Gene in Breast Cancer: Focus on Mutation Spectrum and Therapeutic Strategies. *Cpd* 24 (30), 3566–3575. doi:10.2174/1381612824666180926095709

Ke, J., Zhang, B. H., Li, Y.-Y., Zhong, M., Ma, W., Xue, H., et al. (2019). MiR-1-3p Suppresses Cell Proliferation and Invasion and Targets STC2 in Gastric Cancer. *Eur. Rev. Med. Pharmacol. Sci.* 23, 8870–8877. doi:10.26355/eurrev_201910_19282

Kim, C. G., Ahn, J. B., Jung, M., Beom, S. H., Kim, C., Kim, J. H., et al. (2016). Effects of Microsatellite Instability on Recurrence Patterns and Outcomes in Colorectal Cancers. *Br. J. Cancer* 115, 25–33. doi:10.1038/bjc.2016.161

Law, A. Y. S., Lai, K. P., Ip, C. K. M., Wong, A. S. T., Wagner, G. F., and Wong, C. K. C. (2008). Epigenetic and HIF-1 Regulation of Stanniocalcin-2 Expression in Human Cancer Cells. *Exp. Cell. Res.* 314 (8), 1823–1830. doi:10.1016/j.yexcr10.1016/j.yexcr.2008.03.001

Law, A. Y. S., and Wong, C. K. C. (2010). Stanniocalcin-2 Is a HIF-1 Target Gene that Promotes Cell Proliferation in Hypoxia. *Exp. Cell. Res.* 316 (3), 466–476. doi:10.1016/j.yexcr.2009.09.018

Lee, D.-W., Han, S.-W., Bae, J. M., Jang, H., Han, H., Kim, H., et al. (2019). Tumor Mutation Burden and Prognosis in Patients with Colorectal Cancer Treated with Adjuvant Fluoropyrimidine and Oxaliplatin. *Clin. Cancer Res.* 25, 6141–6147. doi:10.1158/1078-0432.CCR-19-1105

Li, J., Yang, Z., Huang, S., and Li, D. (2020). BIRC7 and STC2 Expression Are Associated with Tumorigenesis and Poor Outcome in Extrahepatic Cholangiocarcinoma. *Technol. Cancer Res. Treat.* 19, 153303382097167. doi:10.1177/1533033820971676

Li, J., Zhang, Z., Feng, X., Shen, Z., Sun, J., Zhang, X., et al. (2022). Stanniocalcin-2 Promotes Cell EMT and Glycolysis via Activating ITGB2/FAK/SOX6 Signaling Pathway in Nasopharyngeal Carcinoma. *Cell. Biol. Toxicol.* 1573–6822. doi:10.1007/s10565-021-09600-5

Li, L., Feng, Q., and Wang, X. (2020). PreMSIm: An R Package for Predicting Microsatellite Instability from the Expression Profiling of a Gene Panel in Cancer. *Comput. Struct. Biotechnol. J.* 18, 668–675. doi:10.1016/j.csbj.2020.03.007

Lin, C. A.-O., Sun, L. A.-O., Huang, S. A.-O., Weng, X. A.-O., and Wu, Z. A.-O. (2019). STC2 Is a Potential Prognostic Biomarker for Pancreatic Cancer and Promotes Migration and Invasion by Inducing Epithelial-Mesenchymal Transition. *Biomed. Res. Int.*, 2314–6141. doi:10.1155/2019/8042489

Liu, Y., Zugazagoitia, J., Ahmed, F. S., Henick, B. S., Gettinger, S. N., Herbst, R. S., et al. (2020). Immune Cell PD-L1 Colocalizes with Macrophages and Is Associated with Outcome in PD-1 Pathway Blockade Therapy. *Clin. Cancer Res.* 26 (4), 970–977. doi:10.1158/1078-0432.CCR-19-1040

Marrelli, D., Polom, K., Pascale, V., Vindigni, C., Piagnerelli, R., De Franco, L., et al. (2016). Strong Prognostic Value of Microsatellite Instability in Intestinal Type Non-cardia Gastric Cancer. *Ann. Surg. Oncol.* 1534–4681. doi:10.1245/s10434-015-4931-3

Miyazaki, S., Kikuchi, H., Iino, I., Uehara, T., Setoguchi, T., Fujita, T., et al. (2014). Anti-VEGF Antibody Therapy Induces Tumor Hypoxia and Stanniocalcin 2 Expression and Potentiates Growth of Human Colon Cancer Xenografts. *Int. J. Cancer* 135 (2), 295–307. doi:10.1002/ijc.28686

Mohan, H. M., Ryan, E., Balasubramanian, I., Kennelly, R., Geraghty, R., Sclafani, F., et al. (2016). Microsatellite Instability Is Associated with Reduced Disease Specific Survival in Stage III Colon Cancer. *Eur. J. Surg. Oncol.* 1532–2157. doi:10.1016/j.ejso.2016.05.013

Nakagawa, T., Kurokawa, T., Mima, M., Imamoto, S., Mizokami, H., Kondo, S., et al. (2021). DNA Methylation and HPV-Associated Head and Neck Cancer. *Microorganisms* 9 (4). doi:10.3390/microorganisms9040801

Poon, I. K., Patel, K. K., Davis, D. S., Parish, C. R., and Hulett, M. D. (2011). Histidine-rich Glycoprotein: the Swiss Army Knife of Mammalian Plasma. *Blood* 117 (7), 2093–2101. doi:10.1182/blood-2010-09-303842

Qie, S., and Sang, N. (2022). Stanniocalcin 2 (STC2): a Universal Tumour Biomarker and a Potential Therapeutical Target. *J. Exp. Clin. Cancer Res.* 41 (1), 161. doi:10.1186/s13046-022-02370-w

Qin, Z. Q., Li, K. D., Chu, H. Z., Yuan, X. F., Shi, H. F., Gu, J., et al. (2022). Blocking Stanniocalcin 2 Reduces Sunitinib Resistance in Clear Cell Renal Cell Carcinoma. *Neoplasma* 69, 145–2685. doi:10.4149/neo_2021_210823N1206

Roche, F. P., Pietilä, I., and Kaito, H. (2018). Leukocyte Differentiation by Histidine-Rich Glycoprotein/Stanniocalcin-2 Complex Regulates Murine Glioma Growth through Modulation of Antitumor Immunity. *Mol. Cancer Ther.* 17 (9), 1961–1972. doi:10.1158/1535-7163.MCT-18-0097

Shi, W., Feng, L., Dong, S., Ning, Z., Hua, Y., Liu, L., et al. (2020). Exploration of Prognostic Index Based on Immune-Related Genes in Patients with Liver Hepatocellular Carcinoma. *Biosci. Rep.* 40 (7). doi:10.1042/BSR20194240

Synnott, N. C., Murray, A., McGowan, P. M., Kiely, M., Kiely, P. A., O'Donovan, N., et al. (2017). Mutant P53: a Novel Target for the Treatment of Patients with Triple-Negative Breast Cancer? *Int. J. Cancer* 140 (1), 234–246. doi:10.1002/ijc.30425

Takei, Y., Yamamoto, H., Sato, T., Otani, A., Kozai, M., Masuda, M., et al. (2012). Stanniocalcin 2 Is Associated with Ectopic Calcification in α-klotho Mutant Mice and Inhibits Hyperphosphatemia-Induced Calcification in Aortic Vascular Smooth Muscle Cells. *Bone* 50 (4), 998–1005. doi:10.1016/j.bone.2012.01.006

Tan, P., He, L., Han, G., and Zhou, Y. (2017). Optogenetic Immunomodulation: Shedding Light on Antitumor Immunity. *Trends Biotechnol.* 35 (3), 215–226. doi:10.1016/j.tibtech.2016.09.002

Vrána, D., Matzenauer, M., Neoral, Č., Aujeský, R., Vrba, R., Melichar, B., et al. (2018). From Tumor Immunology to Immunotherapy in Gastric and Esophageal Cancer. *Int. J. Mol. Sci.* 20 (1). doi:10.3390/ijms20010013

Wang, Y., Gao, Y., Cheng, H., Yang, G., and Tan, W. (2015). Stanniocalcin 2 Promotes Cell Proliferation and Cisplatin Resistance in Cervical Cancer. *Biochem. Biophys. Res. Commun.* 466 (3), 362–368. doi:10.1016/j.bbrc.2015.09.029

Wang, Y., Wu, J., Xu, J., and Lin, S. (2019). Clinical Significance of High Expression of Stanniocalcin-2 in Hepatocellular Carcinoma. *Biosci. Rep.* 39 (4), BSR20182057. doi:10.1042/BSR20182057

Yang, M., Ma, X., Wang, Z., Zhang, T., Hua, Y., and Cai, Z. (2021). Identification of a Novel Glycolysis-Related Gene Signature for Predicting the Prognosis of Osteosarcoma Patients. *Aging (Albany NY)* 13, 12896–12918. doi:10.18632/aging.202958

Yang, S., Ji, Q., Chang, B., Wang, Y., Zhu, Y., Li, D., et al. (2017). STC2 Promotes Head and Neck Squamous Cell Carcinoma Metastasis through Modulating the PI3K/AKT/Snail Signaling. *Oncotarget* 8 (4), 5976–5991. doi:10.18632/oncotarget.13355

Yuan, Q., Zhan, L., Zhang, L. L., Wang, Q., Liu, J., Jiang, Z. Y., et al. (2016). Stanniocalcin 2 Induces Oxaliplatin Resistance in Colorectal Cancer Cells by Upregulating P-Glycoprotein. *Can. J. Physiol. Pharmacol.* 94 (9), 929–935. doi:10.1139/cjpp-2015-0530

Zhang, C., Chen, S., Ma, X., Yang, Q., Su, F., Shu, X., et al. (2019). Upregulation of STC2 in Colorectal Cancer and its Clinicopathological Significance. *Onco Targets Ther.* 12, 1249–1258. doi:10.2147/OTT.S191609

Zhang, Y., and Zhang, Z. (2020). The History and Advances in Cancer Immunotherapy: Understanding the Characteristics of Tumor-Infiltrating Immune Cells and Their Therapeutic Implications. *Cell. Mol. Immunol.* 17 (8), 807–821. doi:10.1038/s41423-020-0488-6

Zhao, Y., Shuen, T. W. H., Toh, T. B., Chan, X. Y., Liu, M., Tan, S. Y., et al. (2018). Development of a New Patient-Derived Xenograft Humanised Mouse Model to Study Human-specific Tumour Microenvironment and Immunotherapy. *Gut* 67 (10), 1845–1854. doi:10.1136/gutjnl-2017-315201

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.

Copyright © 2022 Jiang, Shen, Wei, Chen, Chai, Xia and Leng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

GLOSSARY

HNSCC Head and neck squamous cell carcinoma
HPV Human-papillomavirus
STC2 Stanniocalcin-2
DEGs Differentially expressed genes
ICG Immune checkpoint gene
TMB Tumor mutation burden
MSI Microsatellite instability
MMRs Mismatch repair genes
TCGA The Cancer Genome Atlas
GTEx Genotype-tissue expression
KEGG Kyoto Encyclopedia of Genes and Genomes
GO Gene Ontology
ROC Receiver operating characteristics
AUC Area under the curve
HPA Human Protein Atlas
OS Overall survival
DFS Disease-free survival
DSS Disease-specific survival
PFS Progress-free survival
HRG Histidine-rich glycoprotein
HRG/STC2 Histidine-rich glycoprotein/stanniocalcin-2
ACC Adrenocortical carcinoma
BLCA Bladder urothelial carcinoma
BRCA Breast invasive carcinoma
CESC Cervical squamous cell carcinoma and endocervical adenocarcinoma
CHOL Cholangiocarcinoma
COAD Colon adenocarcinoma
DLBC Lymphoid neoplasm diffuse large B-cell lymphoma
ESCA Esophageal carcinoma
GBM Glioblastoma multiforme
HNSC Head and neck squamous cell carcinoma
KICH Kidney chromophobe
KIRC Kidney renal clear cell carcinoma
KIRP Kidney renal papillary cell carcinoma
LAML Acute myeloid leukemia
LGG Brain lower grade glioma
LIHC Liver hepatocellular carcinoma
LUAD Lung adenocarcinoma
LUSC Lung squamous cell carcinoma
MESO Mesothelioma
OV Ovarian serous cystadenocarcinoma
PAAD Pancreatic adenocarcinoma
PCPG Pheochromocytoma and paraganglioma
PRAD Prostate adenocarcinoma
READ Rectum adenocarcinoma
SARC Sarcoma
SKCM Skin cutaneous melanoma
STAD Stomach adenocarcinoma
TGCT Testicular germ cell tumors
THCA Thyroid carcinoma
THYM Thymoma
UCEC Uterine corpus endometrial carcinoma
UCS Uterine carcinosarcoma
UVM Uveal melanoma

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 22 July 2022
doi: 10.3389/fgene.2022.938985

Copy Number Variation of Circulating Tumor DNA (ctDNA) Detected Using NIPT in Neoadjuvant Chemotherapy-Treated Ovarian Cancer Patients

Mina Sharbatoghli[1,2†], Fahimeh Fattahi[1†], Hamidreza Aboulkheyr Es[3], Arvand Akbari[4], Setareh Akhavan[5], Marzieh Ebrahimi[6], Mohsen Asadi-Lari[1,7], Mehdi Totonchi[4,6] and Zahra Madjd[1,8]**

[1]Oncopathology Research Center, Iran University of Medical Sciences (IUMS), Tehran, Iran, [2]Department of Embryology, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR, Tehran, Iran, [3]School of Biomedical Engineering, University of Technology Sydney, Sydney, NSW, Australia, [4]Department of Genetics, Reproductive Biomedicine Research Center, Royan Institute for Reproductive Biomedicine, ACECR, Tehran, Iran, [5]Department of Gynecologic Oncology, Vali-Asr Hospital, Tehran University of Medical Sciences, Tehran, Iran, [6]Department of Stem Cells and Developmental Biology, Cell Science Research Center, Royan Institute for Stem Cell Biology and Technology, ACECR, Tehran, Iran, [7]Department of Epidemiology, School of Public Health, Iran University of Medical Sciences, Tehran, Iran, [8]Department of Molecular Medicine, Faculty of Advanced Technologies in Medicine, Iran University of Medical Sciences, Tehran, Iran

Analysis of circulating tumor DNA (ctDNA) can be used to characterize and monitor cancers. Recently, non-invasive prenatal testing (NIPT) as a new next-generation sequencing (NGS)-based approach has been applied for detecting ctDNA. This study aimed to investigate the copy number variations (CNVs) utilizing the non-invasive prenatal testing in plasma ctDNA from ovarian cancer (OC) patients who were treated with neoadjuvant chemotherapy (NAC). The plasma samples of six patients, including stages II–IV, were collected during the pre- and post-NAC treatment that were divided into NAC-sensitive and NAC-resistant groups during the follow-up time. CNV analysis was performed using the NIPT via two methods "an open-source algorithm WISECONDORX and NextGENe software." Results of these methods were compared in pre- and post-NAC of OC patients. Finally, bioinformatics tools were used for data mining from The Cancer Genome Atlas (TCGA) to investigate CNVs in OC patients. WISECONDORX analysis indicated fewer CNV changes on chromosomes before treatment in the NAC-sensitive rather than NAC-resistant patients. NextGENe data indicated that CNVs are not only observed in the coding genes but also in non-coding genes. CNVs in six genes were identified, including HSF1, TMEM249, MROH1, GSTT2B, ABR, and NOMO2, only in NAC-resistant patients. The comparison of these six genes in NAC-resistant patients with The Cancer Genome Atlas data illustrated that the total alteration frequency is amplification, and the highest incidence of the CNVs (≥35% based on TCGA data) is found in MROH1, TMEM249, and HSF1 genes on the chromosome (Chr) 8. Based on TCGA data, survival analysis showed a significant reduction in the overall survival among chemotherapy-resistant patients as well as a high expression level of these three genes compared to that of sensitive samples (all, $p < 0.0001$). The continued Chr8 study using WISECONDORX

OPEN ACCESS

Edited by:
Ehsan Nazemalhosseini-Mojarad, Shahid Beheshti University of Medical Sciences, Iran

Reviewed by:
Nayeralsadat Fatemi, Shahid Beheshti University of Medical Sciences, Iran
Longwei Qiao, Xuzhou Medical University, China

***Correspondence:**
Zahra Madjd
majdjabari.z@iums.ac.ir
Mehdi Totonchi
m.totonchi@royaninstitute.org

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

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: 08 May 2022
Accepted: 15 June 2022
Published: 22 July 2022

Citation:
Sharbatoghli M, Fattahi F, Aboulkheyr Es H, Akbari A, Akhavan S, Ebrahimi M, Asadi-Lari M, Totonchi M and Madjd Z (2022) Copy Number Variation of Circulating Tumor DNA (ctDNA) Detected Using NIPT in Neoadjuvant Chemotherapy-Treated Ovarian Cancer Patients. Front. Genet. 13:938985. doi: 10.3389/fgene.2022.938985

revealed CNV modifications in NAC-resistant patients prior to NAC therapy, but no CNV changes were observed in NAC-sensitive individuals. Our findings showed that low coverage whole-genome sequencing analysis used for NIPT could identify CNVs in ctDNA of OC patients before and after chemotherapy. These CNVs are different in NAC-sensitive and -resistant patients highlighting the potential application of this approach in cancer patient management.

Keywords: NIPT, ctDNA, CNVs, ovarian cancer, copy number variations

INTRODUCTION

Ovarian cancer (OC) is a tumor with the worst prognosis among female malignancies, and most cases are diagnosed in the advanced stage with peritoneal metastases (Bray et al., 2018). Neoadjuvant chemotherapy (NAC) is the gold standard treatment for OC patients with a high perioperative risk profile and/or a poor chance of effective debulking surgery (Elies et al., 2018), although a considerable percentage of patients demonstrate resistance to NAC treatment (Sato and Itamochi, 2014). Considering that the administration of proper chemotherapy regimens is dependent on imaging tests, cytology, and laparoscopic biopsy (Lee et al., 2018), inadequate tumor specimens are a flaw and difficulty in cytology or laparoscopic biopsy prior to the onset of NAC (Sharbatoghli et al., 2020). Biomarkers related to drug resistance and treatment, including prognostic and predictive molecules, play a key role in selecting appropriate treatment protocols and improving survival rates (Le Page et al., 2010). In OC, several studies have demonstrated changes in the serum levels of cancer antigen 125 (CA125) which may serve as a predictor of monitoring the response to NAC (Pelissier et al., 2014; Zeng et al., 2016), but its utility is often limited (Sharbatoghli et al., 2020). Serum tumor markers do not exhibit a significant increase in some histological types of OC (Sharbatoghli et al., 2020). Since gene amplification and deletion are common in cancer cells and contribute to cancer cell growth, angiogenesis, and drug resistance (Matsui et al., 2013; Rahimi et al., 2020), recently, copy number variants (CNVs) have been reported as potential biomarkers in cancer management (Pan et al., 2019; Yu et al., 2019; Jin et al., 2021). Moreover, comparing gene expression, CNV is considerably more stable and robust (Pan et al., 2019; Shao et al., 2019) and the gain or loss of gene copies often correlates with a corresponding increase or decrease in the amount of RNA and protein encoded by the gene (Hastings et al., 2009). It is critical to accurately measure copy number alterations in tumor samples in order to enable translational research and precision medicine. The majority of the CNV investigations were conducted on tumor tissue biopsy samples (Vives-Usano et al., 2021). A fraction of cell-free DNA (cfDNA) in cancer patients originates from tumor cells, known as circulating tumor DNA (ctDNA), which is obtained through liquid biopsy. Liquid biopsy as a semi-invasive diagnostic and prognostic tool has the advantage of being less invasive than tumor biopsy, and specimens can be frequently checked in real-time (Mathai et al., 2019). It seems that the molecular alterations identified in ctDNA might mirror molecular heterogeneity of tumor compared to those reflected by tumor biopsy (Sharbatoghli et al., 2020). Therefore, ctDNA analysis was applied to detect various types of genomic alterations, such as CNVs, mutation, and nucleosome positioning variation in cancers (Molparia et al., 2017; Huang et al., 2019; Noguchi et al., 2020).

In recent years, cancer detection by non-invasive prenatal testing (NIPT) as a new next-generation sequencing (NGS)-based approach is used to detect ctDNA (Cohen et al., 2016; Nakabayashi et al., 2018). The low coverage of whole-genome sequencing of cfDNA from maternal plasma is the basis for prenatal screening of common fetal autosomal aneuploidies and trisomy of chromosomes 21, 18, and 13 utilizing NIPT. One of the most often publicized benefits of the NIPT for chromosomal abnormality screening in pregnant women is its low false-positive rate (1–3%) (Filoche et al., 2017). Regarding fast advances in the NIPT, analysis of tumor CNV changes in ctDNA using NIPT was introduced as a potential cancer screening tool. In this context, Amant et al. (2015) discovered cancer in three pregnant women who had NIPT and indicated that CNVs may be used as a cancer screening tool. Furthermore, the implications of the whole-genome NIPT platform for cancer screening were shown in OC patients (Cohen et al., 2016). Furthermore, CNV analysis in cell-free DNA by low coverage whole-genome sequencing was used as a biomarker for the diagnosis of OC (Vanderstichele et al., 2017). For the first time, we applied the NIPT as a non-invasive test platform to compare the tumor-derived CNV in ctDNA measured pre- and post-NAC in plasma samples obtained from OC patients. This is proof of the concept that NIPT might be useful for predicting responsive and resistant patients to chemotherapy.

MATERIALS AND METHODS

Collection of Samples and Blood Processing

This study was conducted as a prospective study with 10 plasma samples derived from 6 OC patients of the Cancer Institute of Imam Khomeini Hospital (Tehran, Iran) between December 2018 and October 2019. This study was performed with the approval of the Ethics Committee of Iran University of Medical Sciences (authorization no: IR.IUMS.REC 1397.32825). Each participating hospital's ethical norms required that written informed permission has to be acquired. The FIGO (International Federation of Gynecology and Obstetrics) stage (Bhatla and Denny, 2018) was used to histologically confirm the

patients' diagnosis. Blood samples were obtained in pre- and post-NAC treatment from OC patients that received platinum-based chemotherapy as an NAC regimen, at the first-line treatment. A week before the first dose of chemotherapy, baseline blood samples were collected, and post-NAC samples were taken after the first course (six cycles) of chemotherapy. The patients with complete response were defined as NAC-sensitive, while those with stable disease and progressive disease were defined as NAC-resistant (Noguchi et al., 2020). A total volume of 10 ml of whole blood was collected in K_2EDTA-coated tubes (REF: CDLP 029, C.D. RICH®, Romania) from each patient. Then, the blood was centrifuged at 1,600×g for 15 min at 25°C, and the plasma fraction was collected and centrifuged for a second time at 2,500×g for 10 min at 25°C. After the second spin, the plasma was transferred into barcoded tubes and immediately stored at ≤-70°C. The cell-free DNA was extracted from 3 ml of patient plasma using the QIAamp Circulating Nucleic Acid Kit (Qiagen, Gaithersburg, MD, United States) (Diefenbach et al., 2018).

Library Preparation, Sequencing, and Data Analysis

DNA libraries were prepared from 2 ng of cell-free DNA extracted from 3 ml of plasma using the VeriSeq NIPT Solution v2 according to the manufacturer's instructions for 75 bp single-end high via sequencing. All libraries were normalized to 1.6 nM, multiplexed, and sequenced on HiSeq: 4000 with 27 sequencing cycles of the cell-free DNA insert and an additional eight sequencing cycles for the index barcodes. Each research sample was sequenced alongside 12 clinical samples, with 36-cycle single-end sequencing on an Illumina NextSeq550. The read depth was low coverage at 0.2× to 0.3× based on the amount of sequencing data. The open-source algorithm WISECONDORX (WIthin-SamplE COpy Number Aberration DetectOR X) and NextGENe (Next GENeration sequencing software for biologists) were used for data analysis (Faircloth and Glenn, 2012).

Copy Number Variation Call Using WISECONDORX and NextGENe

WISECONDORX was used to identify whole chromosome (Chr) and subchromosomal abnormalities that the standard NIPT pipeline failed to detect (Raman et al., 2019). Segmental alterations of less than 0.05 Mb were prespecified as abnormal ("positive cancer screen"). FASTQC (Andrews, 2010) was used to do quality control on the raw single-end sequencing data. The BBduk tool from the BBmap toolbox (Bushnell, 2014) was used to trim and adjust the fastq files as needed.

Reads were mapped to the human reference genome (hg38) using bwa samse algorithm (Li and Durbin, 2009). The resulting sam files were subjected to corrections and converted to bam via samtools (Li et al., 2009). CNVs were called by WISECONDORX run on default settings suggested by the developer (Raman et al., 2018). NextGene software version 2.4.1 (SoftGenetics, LLC) was used for CNV analysis according to the Kerkhof et al. (2017) method.

After obtaining WISECONDORX and NextGENe data, an investigation of these CNV results in pre-treatment was performed on all patients. Moreover, a comparison of these data for patients within the NAC-sensitive and NAC-resistant groups were performed separately to obtain common alterations in chromosomes and gene levels for each group. Then, the data of WISECONDORX and NextGENe software from NAC-resistant were compared to NAC-sensitive patients in order to receive alteration genes detected in NAC resistance. Similarly, post-treatment results were compared in patients with OC in order to detect CNVs.

Data Mining for Genes Detected in NAC-Resistant Patients

To obtain a more comprehensive and deep understanding of the biological process and molecular function of genes detected in NAC resistance, Gene Ontology (GO) term enrichment analysis was performed (Ashburner et al., 2000). Furthermore, KEGG (Kanehisa and Goto, 2000), Reactome (Sidiropoulos et al., 2017), and WikiPathways (Slenter et al., 2018) were investigated to examine the biological pathways in which these genes are involved. GO enrichment and pathway analysis of the genes detected in NAC-resistant patients was revealed using the ClueGO plug-in using Cytoscape (Bindea et al., 2009). Furthermore, genes detected in NAC resistance were investigated in the cBio Cancer Genomics Portal (cBioPortal) database in order to further evaluate the alterations of these genes in OC tissue samples. cBioPortal is an open-access database providing visualization and analyzing tools for multidimensional cancer genomics data, such as The Cancer Genome Atlas (TCGA) (Cerami et al., 2012). The genes detected in NAC resistance with more copy number alteration frequency on CBioPortal were selected to evaluate the mRNA expression levels on Gene Expression Profiling Interactive Analysis (GEPIA2) for OC tissue data compared to normal tissues (Tang et al., 2019). GEPIA2 is an online database, including RNA sequence expression data based on tumor and normal samples from TCGA and the GTEx (Tang et al., 2019). Gene Set Cancer Analysis (GSCALite) was applied to investigate the correlation between mRNA expression and CNV in OC patients. Spearman correlation coefficients were reported using GSCALite software, a user-friendly web server for dynamic analysis and visualization of gene sets in cancer which will be of broad utility to cancer researchers (Liu et al., 2018) from TCGA (Liu et al., 2018). The Cancer Virtual Cohort Discovery Analyses Platform (CVCDAP) portal was also utilized to compare the overall survival (OS) analyses between chemo-resistant and chemo-sensitive patients based on mRNA expression levels of these genes (Guan et al., 2020). CVCDAP is a web-based platform to deliver an interactive and customizable toolbox off the shelf for cohort-level analysis of TCGA and CPTAC public datasets (Guan et al., 2020). Finally, a volcano plot was applied to evaluate the protein expression in OC chemotherapy-resistant patients compared to that in chemotherapy-sensitive patients by TCGAbiolinks, an R/Bioconductor package for integrative analysis of TCGA data (Colaprico et al., 2016).

TABLE 1 | Clinicopathological characteristics of six ovarian cancer patients that were treated with neoadjuvant chemotherapy.

Case no.	Age (Year)	FIGO stage	Right (R) and left (L) ovary	Histological diagnosis	Treatment	Clinical response to therapy
1	38	III	R	HGSOC	Carboplatin + paclitaxel	Progressive disease
2	78	IV	R	HGSOC	Carboplatin + paclitaxel	Responsive
3	45	II	R	HGSOC	Carboplatin + paclitaxel	Responsive
4	50	III	R and L	HGSOC	Carboplatin + paclitaxel	Progressive disease
5	54	III	R and L	HGSOC	Carboplatin + paclitaxel	Progressive disease (death)
6	68	IV	R	HGSOC	Carboplatin + paclitaxel	Progressive disease (death)

RESULTS

Patient Characteristics

The clinicopathological characteristics and response to NAC of the six OC patients are summarized in **Table 1**. The mean age of the patients was 54.3 years (38–78 years), and they included stages II, III, and IV high-grade serous ovarian carcinoma (HGSOC). Two of the six patients died undergoing chemotherapy. As a result, these two patients were unable to provide post-NAC plasma samples. Two patients were NAC-sensitive, while two cases were NAC-resistant and did not respond appropriately to chemotherapy treatment.

Investigation of CNV From ctDNA in NAC-Resistant and NAC-Sensitive Patients With OC

We detected 6/6 HGSOC cases including early-stage (II) to late-stage (IV)—using the WISECONDORX (**Table 2**) and NextGENe analysis (**Supplementary Tables S1–S3**). WISECONDORX results indicated that most patients (at least 5 cases) have abnormality (gain and/or loss) on chromosomes, 4, 9, 18, and 22 before NAC treatment (**Table 2**). NextGENe data illustrated CNVs not only in coding genes but also alterations in duplication and deletion in non-coding genes (**Supplementary Tables S1–S3**). Chr changes detected by WISECONDORX software indicated that NAC-sensitive patients have fewer chromosomes' CNV changes before treatment rather than NAC-resistant patients as shown in **Figure 1**; **Table 2**.

Using the results of NextGENe software, common genes with copy number changes were assessed in pre-NAC and post-NAC treatment by the Vinny plot. We detected 17 common genes with CNV in pre-NAC patients from the NAC-resistant group (**Figure 2A**; **Supplementary Table S4**). Among these genes, six common genes, including NOMO2, ABR, GSTT2B, HSF1, TMEM249, and MROH1, exclusively have shown CNV in the NAC-resistant group, while in NAC-sensitive patients, none of these genes were altered before NAC treatment. Furthermore, as shown in **Figure 2B**; **Supplementary Table S5**, 38 common genes with CNV were discovered in NAC-resistant patients' post-NAC data. Furthermore, we discovered 14 genes, including LOC285441, LOC100996414, TTLL10, NXN, SMA4, SMA5, NOMO1, TIMM22, NOMO2, HSF1, ABR, TMEM249, GSTT2B, and MROH1, that exclusively revealed CNVs in NAC-resistant patients following chemotherapy. We indicate the CNVs for the six genes (NOMO2, ABR, GSTT2B, HSF1, TMEM249, and MROH1) that were frequent in the NAC-resistant group pre- and post-treatment as genes associated and found in NAC-resistant patients (See **Table 3**).

Data Mining Approaches for Genes Detected in NAC-Resistant Patients

The pathway enrichment analysis for six common genes detected in NAC-resistant patients indicated that GSTT2B was involved in chemical carcinogenesis and drug metabolism pathways, while ABR and HSF1 were enriched in p75 NTR receptor-mediating signaling and cellular response to heat stress/shock response pathways, respectively (**Figure 3A**). Gene Ontology analysis illustrated that ABR and HSF1 contributed to apoptosis and programmed cell death (GO: 0012501 and 0006915, respectively) (**Supplementary Table S6**). Moreover, ABR, GSTT2B, and HSF1 have a role in catalytic activity and metabolic processes (**Figure 3B**).

Comparing our CNV results of six common genes detected in NAC-resistant patients with TCGA data through the cBioPortal illustrated that the total alteration frequency is amplification, similar to TCGA data from 585 OC patients (**Figure 4**). Moreover, the most alteration frequency of CNVs (≥35%) was observed in HSF1, TMEM249, and MROH1 genes from OC patients (**Figure 4**).

Loci of HSF1, TMEM249, and MROH1 genes are arranged on Chr8. CNV alterations in Chr8 in WISECONDORX data were solely found in NAC-resistant patients when compared to NAC-sensitive patients before treatment (**Table.2**). Investigation of these three genes at the mRNA level on GEPIA2 showed significantly high expression of HSF1, TMEM249, and MROH1 in ovarian serous carcinoma (OSC) compared with normal tissues (all, $p < 0.05$) (**Figure 5**A–C). Moreover, the Spearman correlation coefficient shows the positive association between CNVs and mRNA expression of HSF1 (Cor = 0.83, FDR = 6e-74), MROH1 (Cor = 0.72, FDR = 4.8e-49), and TMEM249 (Cor = 0.48, FDR = 3.11e-18) across 585 OC patients from TCGA (**Figure 6**). As shown in **Figure 7**, the volcano plot indicates the enrichment of TMEM249, MROH1, and HSF1 protein expression in resistant patients to chemotherapy from TCGA. In other words, the expression of these three proteins significantly increased in resistant patients ($n = 90$) versus sensitive patients to chemotherapy ($n = 194$) (all, $p < 0.05$).

TABLE 2 | Copy number variations (CNVs) in six ovarian cancer cases which were reported based on gains and losses of chromosomes by the WISECONDORX analysis.

Patient 1		Patient 2		Patient 3		Patient 4		Patient 5		Patient 6	
Pre-NAC	**Post-NAC**	**Pre-NAC**	**Post-NAC**	**Pre-NAC**	**Post-NAC**	**Pre-NAC**	**Post-NAC**	**Pre-NAC**	**Post-NAC**	**Pre-NAC**	**Post-NAC**
Chr1 loss	Chr1 gain and loss	Chr4 loss	Chr1 gain	Chr4 loss	Chr1 loss	Chr1 gain	Chr1 gain	Chr1 gain	No data	Chr2 loss	No data
Chr2 loss	Chr3 loss	Chr6 loss	Chr2 gain	Chr5 loss	Chr2 loss	Chr2 gain	Chr2 loss	Chr3 loss		Chr3 gain	
Chr3 loss	Chr4 loss	Chr11 loss	Chr3 gain	Chr9 gain	Chr3 loss	Chr3 gain	Chr3 gain	Chr4 loss		Chr5 gain and loss	
Chr4 loss	Chr6 loss	Chr16 gain	Chr4 gain	Chr17 gain	Chr4 loss	Chr4 gain and loss	Chr4 loss	Chr5 loss		Chr7 gain and loss	
Chr5 loss	Chr7 gain	Chr18 loss	Chr5 gain	Chr18 loss	Chr5 loss	Chr5 gain and loss	Chr5 gain and loss	Chr6 loss		Chr8 gain and loss	
Chr6 loss	Chr8 loss	Chr22 gain	Chr6 gain	Chr19 gain	Chr6 loss	Chr6 gain and loss	Chr6 loss	Chr7 loss		**Chr9 gain**	
Chr7 loss	Chr9 loss	ChrX loss	Chr7 gain	Chr22 gain	Chr7 loss	Chr7 gain and loss	Chr8 loss	Chr8 loss		Chr10 gain	
Chr8 loss	Chr10 gain and loss		Chr8 gain		Chr8 loss	Chr8 gain and loss	Chr9 gain	Chr9 loss		Chr11 loss	
Chr9 loss	Chr11 gain		Chr9 gain		Chr9 loss	Chr9 gain and loss	Chr16 gain	Chr11 loss		Chr12 loss	
Chr10 loss	Chr12 gain		Chr10 gain		Chr10 loss	Chr10 gain and loss	Chr17 gain	Chr12 gain and loss		Chr13 loss	
Chr11 loss	Chr13 loss		Chr11 gain		Chr11 loss	Chr11 gain	Chr18 loss	Chr13 gain and loss		Chr14 gain and loss	
Chr12 loss	Chr14 gain and loss		Chr12 gain		Chr12 loss	Chr12 gain and loss	Chr19 gain	Chr14 loss		Chr15 gain and loss	
Chr13 loss	Chr16 gain		Chr13 gain		Chr13 loss	Chr13 loss	Chr20 gain	Chr16 gain		Chr16 loss	
Chr14 loss	Chr17 gain		Chr14 gain		Chr14 loss	Chr14 loss	Chr22 gain and loss	Chr17 gain		Chr18 gain	
Chr15 gain and loss	Chr18 loss		Chr15 gain		Chr15 loss	Chr15 loss	ChrX loss	Chr18 loss		Chr21 loss	
Chr16 gain	Chr19 gain and loss		Chr16 gain		Chr17 loss	Chr16 loss		Chr19 gain		Chr22 loss	
Chr17 gain	Chr20 gain		Chr17 gain		Chr18 loss	Chr17 gain and loss		Chr20 gain		ChrX gain and loss	
Chr18 loss	Chr22 gain		Chr18 gain		Chr22 gain	Chr20 gain		Chr22 gain			
Chr19 gain	ChrX loss		Chr19 gain		ChrX loss	Chr22 gain		ChrX loss			
Chr20 gain			Chr20 gain			ChrX loss					
Chr21 loss			Chr21 gain								
Chr22 gain			Chr22 gain								
ChrX loss			ChrX gain								

Survival analyses showed a significant reduction in OS across patients resistant to chemotherapy and a high expressed level of HSF1, TMEM249, and MROH1 compared to sensitive samples (all, $p < 0.0001$). Moreover, these data illustrate that high expression of MROH1 and TMEM249 significantly in resistant and sensitive samples can reduce OS among these patients (all, $p < 0.0001$) (**Figure 8**).

DISCUSSION

ctDNA from liquid biopsy has recently been shown to be a useful diagnostic tool in a variety of cancer patients (Saha et al., 2022). Indeed, it has been shown that ctDNA is correlated to tumor burden and the risk of recurrence (Fleischhacker and Schmidt, 2007). ctDNA as part of total body DNA from cancer cells is superior to other plasma biomarkers such as RNA and protein. It is more stable than RNA (Panawala, 2017) with more sensitivity and clinical correlations. Despite the fact that plasma protein biomarkers are commonly used in clinical management for different cancers, such as AFP, CEA, PSA, and CA15-3 (Mazzucchelli et al., 2000; He et al., 2013), some cancer patients are not positive for these biomarkers. Furthermore, they are found with lower concentrations in the serum of

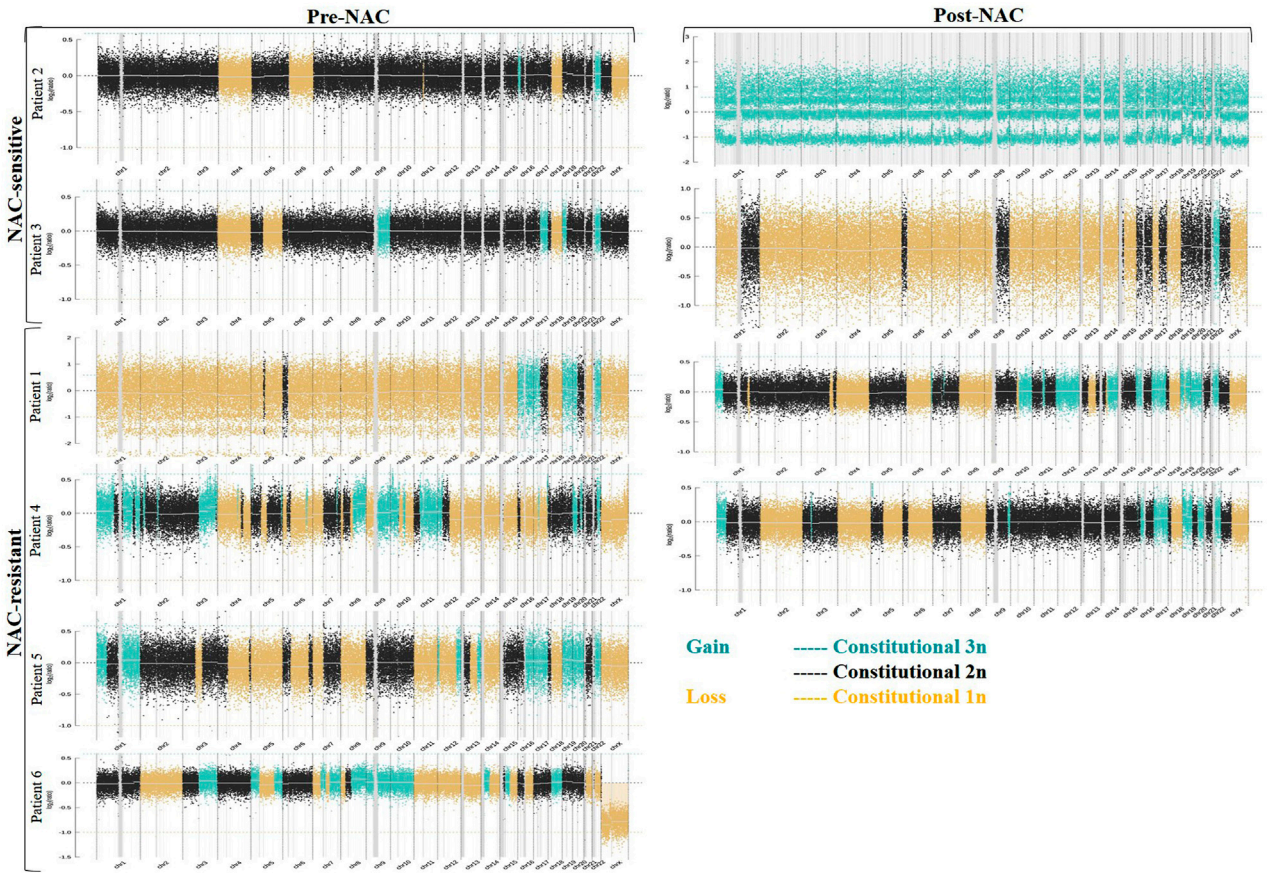


FIGURE 1 | Copy number variations (CNVs) in ctDNA from six ovarian cancer (OC) patients that were treated with neoadjuvant chemotherapy (NAC). Data CNVs for all chromosomes based on WISECONDORX analysis were detected from NAC-sensitive, NAC-resistant, and dead patient groups. The copy numbers were segmented in blue (gain) and yellow (loss) lines.

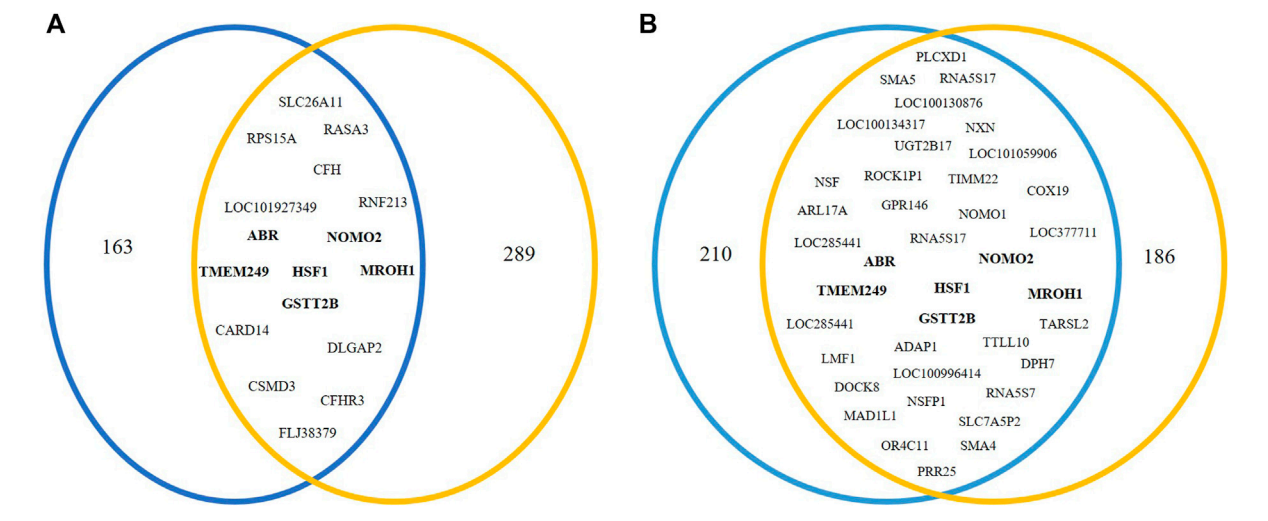


FIGURE 2 | Venn diagram analysis to investigate genes related to NAC resistance. In NAC-resistant patients, Venn diagram indicated **(A)** 17 common genes in pre-NAC-treated patients and **(B)** 38 common genes in post-NAC-treated patients that were obtained via NextGENe analysis.

TABLE 3 | Common genes in pre- and post-neoadjuvant chemotherapy (NAC) as well as its copy number variation (CNV) status (in pre-and post-NAC) for NAC-resistant patients that were detected by NextGENe software.

Common gene	Patient 1 CNV status		Patient 4 CNV status	
	Pre-NAC	Post-NAC	Pre-NAC	Post-NAC
NOMO2 (Chr16)	Dup	Dup	Dup	Dup
ABR (Chr17)	Dup	Dup	Dup	Dup
GSTT2B (Chr22)	Dup	Dup	Dup	Dup
MROH1 (Chr8)	Dup	Normal	Dup	Dup
HSF1 (Chr8)	Dup	Normal	Dup	Dup
TMEM249 (Chr8)	Dup	Normal	Dup	Dup

Dup, Duplication

individuals free of cancer (Cheng et al., 2016), while the studies show that ctDNA can more accurately reflect the real-time tumor burden in patients receiving therapy (Bettegowda et al., 2014).

The current study is the first report to indicate the utility of analyzing CNVs in ctDNA using massively parallel sequencing (such as NIPT), before and after NAC treatment in OC patients. The previous study by Fleischhacker and Schmidt, (2007) has shown that ctDNA is low in the early stages of the tumor and difficult to detect. The method in this study can detect copy number alterations of ctDNA in various stages of HGSOC patients, including stages II–IV, through low coverage plasma DNA sequencing and analysis for chromosomal CNVs ≥0.5 Mb. It has previously been demonstrated that tumor DNA from cancer cells has been detected in plasma using NIPT (Amant et al., 2015; Bianchi et al., 2015; Cohen et al., 2016). We also investigated the response to chemotherapy and depicted the CNV changes in NAC-sensitive and NAC-resistant patients. CNV burden in all chromosomes of NAC-sensitive patients was fewer than that of NAC-resistant patients before NAC treatment. Since the CNVs are considered a key factor in the genetic variation of tumors (Hu et al., 2021), it seems that less CNV burden indicated better response in treating patients (Walker et al., 2017). CNVs were also discovered in a variety of genes following chemotherapy, notably in genes that control drug absorption into cells and drug metabolism (Willyard, 2015). As a result, we compared produced CNVs in NAC-sensitive and NAC-resistant patient groups with OC following NAC therapy. After the first course of chemotherapy, the CNV load increased in both NAC-sensitive and NAC-resistant individuals, according to our findings. This phenomenon could be in terms of the death of tumor cells (DNA damage) (Woods and Turchi, 2013) or resistance to therapy (Alfarouk et al., 2015).

The previous studies have suggested the role of CNVs in genes related to drug resistance (Alfarouk et al., 2015; Willyard, 2015; Costa et al., 2017; Qidwai, 2020). Our findings appear to well indicate that some of the CNVs detected by the NIPT are different in NAC-resistant groups compared to NAC-sensitive groups of OC patients. We found some common genes, including NOMO2, ABR, GSTT2B, HSF1, TMEM249, and MROH1, in the pre- and post-treatment for the NAC-resistant group which was no longer detected in NAC-sensitive patients before and after therapy. These findings suggest that CNVs discovered by the NIPT may contribute to treatment resistance. To confirm the presence of CNV in all patients' plasma, we matched the findings of genes with CNVs discovered in their plasma to the tissue samples in the TCGA data (Bell et al., 2011) owing to the unavailability of tissue specimens from these patients. It is better to investigate the plasma sequencing data with paired tumor DNA of tissue samples, but the lack of tissue or insufficient tumor tissue sampling of patients is a limitation for this type of study (Cohen et al., 2016). The cBioPortal findings indicated the existence of alteration and amplification in all six genes among TCGA patients so that some of these CNVs have a high frequency in OC tissue samples. Moreover, HSF1, TMEM249, and MROH1 are located on Chr8 which is highlighted in our

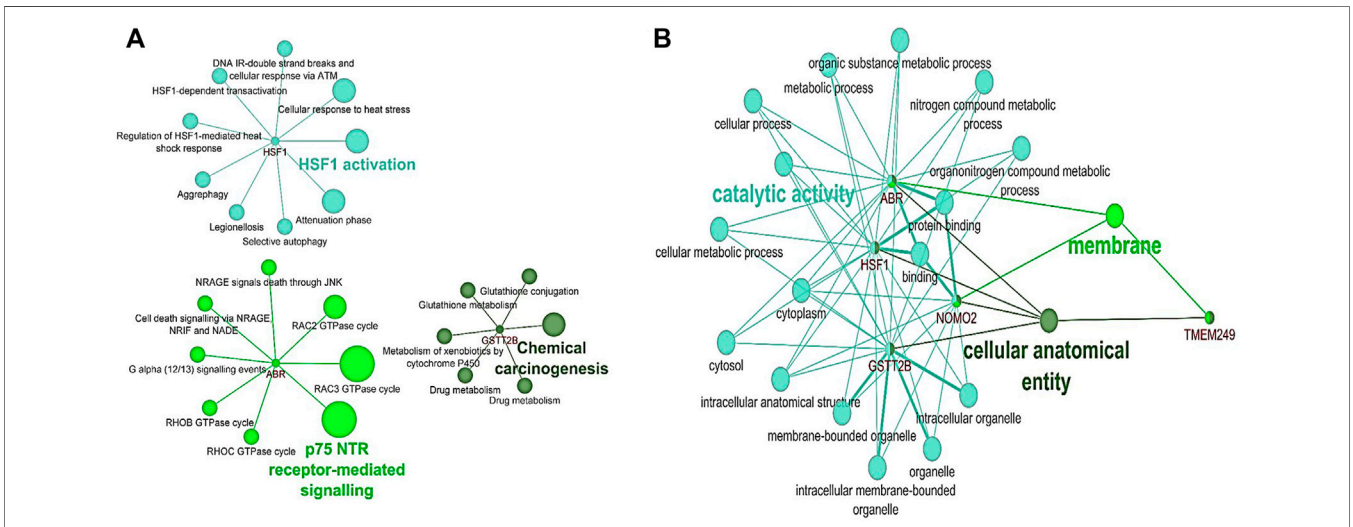


FIGURE 3 | Pathway and Gene Ontology analysis for six genes detected in NAC-resistant patients using the ClueGO plugin in Cytoscape. **(A)** Pathway analysis based on KEGG, Reactome, and WikiPathways. **(B)** Common results of GO analysis for HSF1, ABR, and GSTT2B according to biological processes and molecular function.

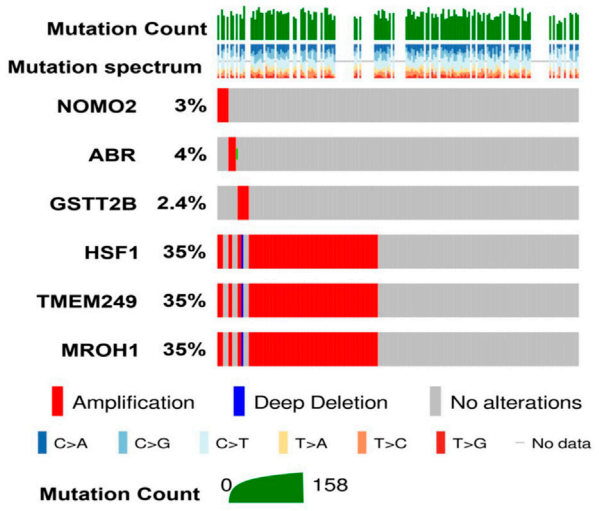


FIGURE 4 | Investigation of DNA alterations for six genes detected in NAC resistance in TCGA-ovarian cancer patients on cBioPortal. Oncoplot illustrates the amplification status of six different genes identified in resistant patients across 585 ovarian carcinoma patients from TCGA. Among them, high amplification of HSF1, TMEM249, and MROH1 was observed compared to other identified genes.

study for OC patients. The CNVs' investigation via WISECONDOREX shows Chr8 underwent CNV changes in NAC-resistant patients before NAC treatment, while in NAC-sensitive patients no CNV changes were found on Chr8 before chemotherapy. Amplification of Chr8 genes has also been identified as a recurrent genomic event in lung cancer (Baykara et al., 2015) and malignant peripheral nerve sheath tumor (MPNST) (Dehner et al., 2021). This finding indicates the power role and relationship among CNV alterations on Chr8 and NAC-resistant in OC patients. Other studies showed that approximately 80 genes on Chr8 are involved in cancer biology (Tabarés-Seisdedos and Rubenstein, 2009). We investigated the levels of RNA and protein expression from HSF1, TMEM249, and MROH1, which are related to NAC-resistant patients on TCGA data. DNA copy number variation is an important factor in the expression of genes (Gamazon and Stranger, 2015) and also an important influential factor for the expression of both protein-coding and non-coding genes (Liang et al., 2016). As expected, positive correlations between the level of mRNA and CNVs alterations were observed for HSF1, TMEM249, and MROH1. In this regard, survival analysis confirmed the influence expression of these three genes in the survival rate, so that resistant patients with higher mRNA expression of these genes had a reduced OS. Furthermore, in TCGA-chemotherapy-resistant OC patients, the protein expression of these genes was higher than that in sensitive individuals. So far, dysregulation and different roles of these genes in various cancers have been evaluated. Studies indicated that HSF1 has been implicated in tumorigenesis by its participation in cellular stress response pathways and its effect on regulatory pathways such as p53, mTOR, and insulin signaling (Vihervaara and Sistonen, 2014; Vydra et al., 2014; Powell et al., 2016; Barna et al., 2018). In line with our *in silico* analysis study, high levels of HSF1 have been identified in different types of cancers (Chen et al., 2021). Overexpression of HSF1 in tumor tissues is correlated with a worse prognosis in cancer patients (Kourtis et al., 2015). Regarding the correlation between a high level of HSF1 and the deterioration of disease in the initiation, promotion, and progression of cancer (Wang et al., 2020), HSF1 could act as a potential therapeutic target (Wang et al., 2020). OC studies reported that HSF1 induces epithelial–mesenchymal transition (EMT) in the *in vitro*

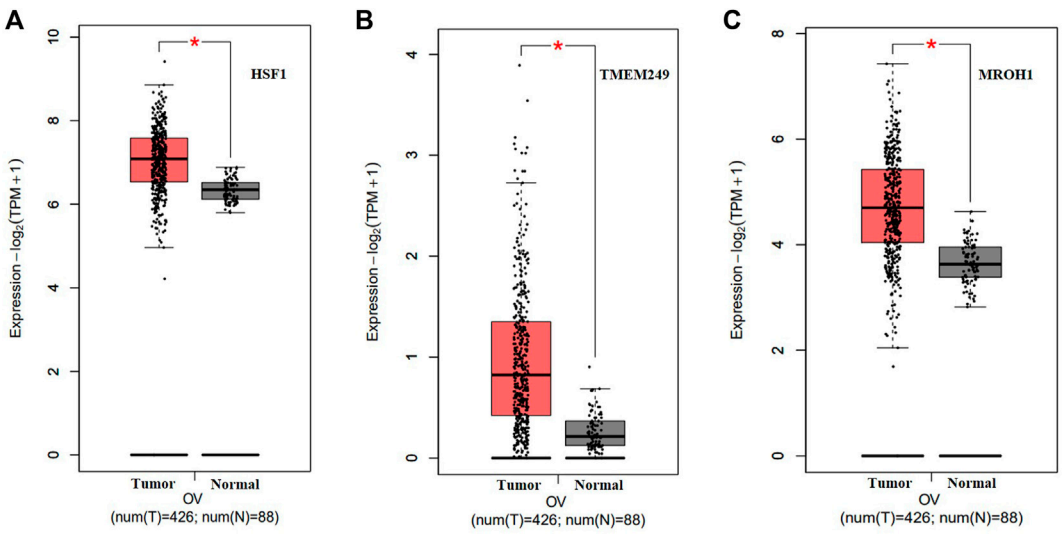


FIGURE 5 | mRNA levels of three genes related to NAC resistance in serous ovarian carcinoma (SOC) on Gene Expression Profiling Interactive Analysis2 (GEPIA2). Upregulated expression of **(A)** HSF1, **(B)** TMEM249, and **(C)** MROH1 in mRNA levels significantly for SOC compared with that for normal tissue by GEPIA2 was observed (*: $p < 0.05$).

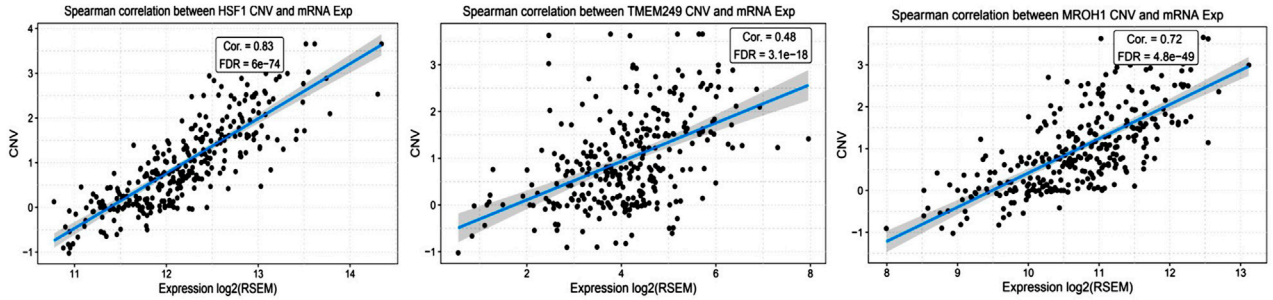


FIGURE 6 | Correlation study between mRNA levels and copy number variation of three genes related to NAC resistance through cBioPortal in TCGA-ovarian cancer patients. The correlation coefficient analysis shows the positive association between mRNA expression and copy number variation of HSF1, TMEM249, and MROH1 across 585 ovarian cancer patients from TCGA.

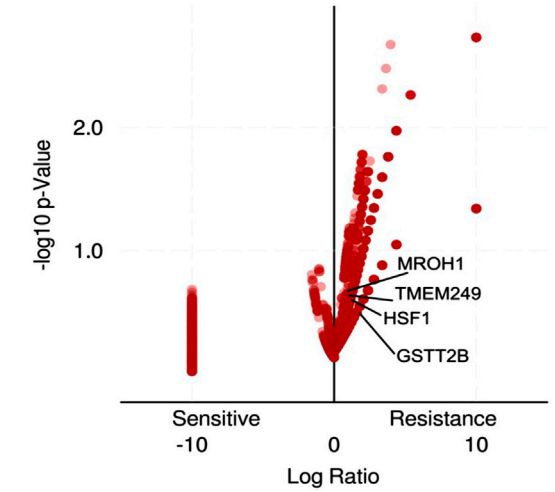


FIGURE 7 | Volcano plot enrichment analysis for gene-related protein expression in ovarian cancer patients who were treated with chemotherapy from TCGA using TCGAbiolinks. The expression of HSF1, TMEM249, and MROH1 proteins significantly increased in resistant patients (n = 90) versus sensitive patients to chemotherapy (n = 194).

models (Powell et al., 2016) and targeting HSF1 leads to an antitumor effect (Chen et al., 2017). Zhang et al. (2017) identified some parts of human Chr8, a location hotspot, mediated by the master regulator HSF1 in different cancers. Furthermore, they interestingly indicated MROH1 and TMEM249 immediately flanking the upstream and downstream regions of HSF1. As predicted, our CNV data for Chr8 genes, such as HSF1, MROH1, and TMEM249, were duplicated in individuals who were resistant to NAC. Considering HSF1's critical involvement in cancer development, the lack of HSF1 in NAC-sensitive patients compared to NAC-resistant patients suggests a function for the Chr8 and HSF1 genes in NAC resistance. The activity of the GSTT2 enzyme is important for the protection of cells against toxic products of oxygen and lipid peroxidation (Tan and Board, 1996), which represents a major source of endogenous DNA damage in humans that contributes significantly to cancer and other genetic diseases (Marnett, 2002). Duplicated-CNV and also mRNA expression of GSTT2B in NAC-resistant patients by bioinformatics analysis are consistent with the previous cancer studies (Pool-Zobel et al., 2005; Doherty et al., 2014). Research on colon cancer cells indicated upregulated GSTT2 upon incubation with butyrate that is involved in defense against oxidative stress (Pool-Zobel et al., 2005). Furthermore, Doherty et al. (2014)

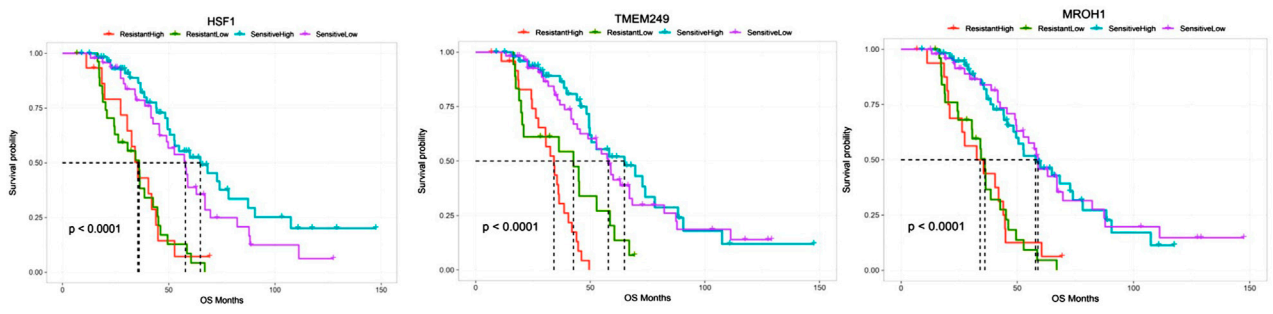


FIGURE 8 | Comparing the overall survival between chemo-resistant and -sensitive expression of HSF1, TMEM249, and MROH1 using the CVCDAP portal for ovarian cancer patients. The plot shows a significant reduction in the overall survival across patients resistant to chemotherapy and expressed a high level of HSF1, TMEM249, and MROH1 genes in comparison with sensitive samples. Moreover, the plot indicates that high expression of MROH1 and TMEM249 in chemo-sensitive samples can reduce OS among these patients.

observed that the expression of GSTT2 was increased in drug-resistant cervical cell models with cisplatin treatment.

Deletion of ABR indicated a tumor-suppressive role in several solid tumors, such as medulloblastoma (McDonald et al., 1994), astrocytomas (Willert et al., 1995), and breast cancer (Liscia et al., 1999). In acute myeloid leukemia, the ABR gene was detected as a prognostic factor in which the blockage of ABR prevents myeloid differentiation (Namasu et al., 2017). In our study, the ABR gene was duplicated in NAC-resistant patients before and after treatment. These data suggested more investigation into the role and function of the ABR gene that might influence resistance to chemotherapy for OC patients.

The NOMO2 gene is known as a diagnostic biomarker in radioresistance in human H460 lung cancer stem-like cells (Yun et al., 2016). We found NOMO2 duplication in NAC-resistant patients by NIPT. These data are consistent with the previous result in metastatic breast cancer by NGS that detected mutation or amplification in cfDNA samples (Page et al., 2017). CNVs were not only found in coding genes but also detected in non-coding genes, such as microRNA and long non-coding genes, although the previous study in bladder cancer indicated CNV alteration in long non-coding RNAs, which can be used as prognostic biomarkers for bladder cancer (Zhong et al., 2021). Furthermore, copy number changes in non-coding RNAs have been identified as prospective therapeutic targets and prognostic markers in lung squamous cell carcinoma (LUSC) (Ning et al., 2021). We are aware that our research has limitations to describe the biological behavior of these CNVs and their relationships with NAC resistance in OC cells. Although the number of records was adequate to establish a conclusion for the trend in data, however, in terms of seeking generalizability, larger sample sizes are strongly suggested, which could be covered via larger and multicenter investigations.

CONCLUSION

This study's findings highlighted low-coverage whole-genome sequencing analysis to investigate CNV changes in ctDNA. It seems that detected CNVs through the NIPT in ctDNA could be potential markers of clinical response to NAC treatment. Our results gave a clue that some alterations in the copy number of genes at the DNA level may relate to being the response to NAC treatment in OC patients, although further studies are warranted to understand the role of these CNVs in NAC-resistant patients. Using a platform for prenatal testing in diagnosis or monitoring therapeutic response for other cancer types as a novel way may hold promise that it should be examined.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB53061. (https://www.ebi.ac.uk/ena/browser/view/PRJEB53061).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Iran University of Medical Sciences Human Research Ethics Committee in Iran (Ref No: IR.IUMS.REC1397.32825). All procedures performed in this study were in accordance with the 1964 Helsinki Declaration and its later amendments. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ZM, MA-L, and MT designed and supervised the work; MS, a Ph.D. student, performed all examinations, collected the patient's samples and data, and also wrote the manuscript; FF participated in data mining and literature review using bioinformatics tools, wrote the bioinformatics sections, prepared the related figures, and assisted in the editing of the manuscript; HA participated in bioinformatics data analysis and in editing the manuscript; AA participated in NGS data analysis and wrote the method section; SA contributed to sample collection; all authors read and approved the final manuscript.

FUNDING

This work was supported by a grant from the Iran University of Medical Sciences (#32825) and Royan Institute (#97000072).

ACKNOWLEDGMENTS

The authors are so grateful to their colleagues at the Iran University of Medical Sciences and Royan Stem Cell Technology Company for their financial and technical support and Tehran University of Medical Sciences (Imam Khomeini hospital) for support in collecting samples.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.938985/full#supplementary-material

REFERENCES

Alfarouk, K. O., Stock, C.-M., Taylor, S., Walsh, M., Muddathir, A. K., Verduzco, D., et al. (2015). Resistance to Cancer Chemotherapy: Failure in Drug Response from ADME to P-Gp. *Cancer Cell Int.* 15, 71. doi:10.1186/s12935-015-0221-1

Amant, F., Verheecke, M., Wlodarska, I., Dehaspe, L., Brady, P., Brison, N., et al. (2015). Presymptomatic Identification of Cancers in Pregnant Women During Noninvasive Prenatal Testing. *JAMA Oncol.* 1 (6), 814–819. doi:10.1001/jamaoncol.2015.1883

Andrews, S. (2010). "FastQC: A Quality Control Tool for High Throughput Sequence Data," in *Babraham Bioinformatics* (Cambridge, United Kingdom: Babraham Institute).

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene Ontology: Tool for the Unification of Biology. *Nat. Genet.* 25 (1), 25–29. doi:10.1038/75556

Barna, J., Csermely, P., and Vellai, T. (2018). Roles of Heat Shock Factor 1 beyond the Heat Shock Response. *Cell. Mol. Life Sci.* 75 (16), 2897–2916. doi:10.1007/s00018-018-2836-6

Baykara, O., Bakir, B., Buyru, N., Kaynak, K., and Dalay, N. (2015). Amplification of Chromosome 8 Genes in Lung Cancer. *J. Cancer* 6 (3), 270–275. doi:10.7150/jca.10638

Bell, D., Berchuck, A., Birrer, M., Chien, J., Cramer, D. W., Dao, F., et al. (2011). Integrated Genomic Analyses of Ovarian Carcinoma. *Nature* 474 (7353), 609–615. doi:10.1038/nature10166

Bettegowda, C., Sausen, M., Leary, R. J., Kinde, I., Wang, Y., Agrawal, N., et al. (2014). Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. *Sci. Transl. Med.* 6 (224), 224ra24. doi:10.1126/scitranslmed.3007094

Bhatla, N., and Denny, L. (2018). FIGO Cancer Report 2018. *Int. J. Gynecol. Obstet.* 143 Suppl 2, 2–3. doi:10.1002/ijgo.12608

Bianchi, D. W., Chudova, D., Sehnert, A. J., Bhatt, S., Murray, K., Prosen, T. L., et al. (2015). Noninvasive Prenatal Testing and Incidental Detection of Occult Maternal Malignancies. *JAMA* 314 (2), 162–169. doi:10.1001/jama.2015.7120

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., et al. (2009). ClueGO: A Cytoscape Plug-In to Decipher Functionally Grouped Gene Ontology and Pathway Annotation Networks. *Bioinformatics* 25 (8), 1091–1093. doi:10.1093/bioinformatics/btp101

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J. Clin.* 68 (6), 394–424. doi:10.3322/caac.21492

Bushnell, B. (2014). *BBMap: A Fast, Accurate, Splice-Aware Aligner*. US: Lawrence Berkeley National Laboratory. Retrieved from: https://escholarship.org/uc/item/1h3515gn LBNL Report #: LBNL-7065E.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov.* 2 (5), 401–404. doi:10.1158/2159-8290.cd-12-0095

Chen, F., Fan, Y., Cao, P., Liu, B., Hou, J., Zhang, P., et al. (2021). Pan-Cancer Analysis of the Prognostic and Immunological Role of HSF1: A Potential Target for Survival and Immunotherapy. *Oxid. Med. Cell Longev.* 2021, 5551036. doi:10.1155/2021/5551036

Chen, Y.-F., Wang, S.-Y., Yang, Y.-H., Zheng, J., Liu, T., and Wang, L. (2017). Targeting HSF1 Leads to an Antitumor Effect in Human Epithelial Ovarian Cancer. *Int. J. Mol. Med.* 39 (6), 1564–1570. doi:10.3892/ijmm.2017.2978

Cheng, F., Su, L., and Qian, C. (2016). Circulating Tumor DNA: A Promising Biomarker in the Liquid Biopsy of Cancer. *Oncotarget* 7 (30), 48832–48841. doi:10.18632/oncotarget.9453

Cohen, P. A., Flowers, N., Tong, S., Hannan, N., Pertile, M. D., and Hui, L. (2016). Abnormal Plasma DNA Profiles in Early Ovarian Cancer Using a Non-invasive Prenatal Testing Platform: Implications for Cancer Screening. *BMC Med.* 14 (1), 126. doi:10.1186/s12916-016-0667-6

Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAbiolinks: An R/Bioconductor Package for Integrative Analysis of TCGA Data. *Nucleic Acids Res.* 44 (8), e71. doi:10.1093/nar/gkv1507

Costa, G. L., Amaral, L. C., Fontes, C. J. F., Carvalho, L. H., de Brito, C. F. A., and de Sousa, T. N. (2017). Assessment of Copy Number Variation in Genes Related to Drug Resistance in Plasmodium Vivax and Plasmodium Falciparum Isolates from the Brazilian Amazon and a Systematic Review of the Literature. *Malar. J.* 16 (1), 152. doi:10.1186/s12936-017-1806-z

Dehner, C., Moon, C. I., Zhang, X., Zhou, Z., Miller, C., Xu, H., et al. (2021). Chromosome 8 Gain Is Associated with High-Grade Transformation in MPNST. *JCI insight* 6 (6), e146351. doi:10.1172/jci.insight.146351

Diefenbach, R. J., Lee, J. H., Kefford, R. F., and Rizos, H. (2018). Evaluation of Commercial Kits for Purification of Circulating Free DNA. *Cancer Genet.* 228-229, 21–27. doi:10.1016/j.cancergen.2018.08.005

Doherty, B., Lawlor, D., Gillet, J. P., Gottesman, M., O'Leary, J. J., and Stordal, B. (2014). Collateral Sensitivity to Cisplatin in KB-8-5-11 Drug-Resistant Cancer Cells. *Anticancer Res.* 34 (1), 503–507.

Elies, A., Rivière, S., Pouget, N., Becette, V., Dubot, C., Donnadieu, A., et al. (2018). The Role of Neoadjuvant Chemotherapy in Ovarian Cancer. *Expert Rev. Anticancer Ther.* 18 (6), 555–566. doi:10.1080/14737140.2018.1458614

Faircloth, B. C., and Glenn, T. C. (2012). Not All Sequence Tags Are Created Equal: Designing and Validating Sequence Identification Tags Robust to Indels. *PloS one* 7, e42543. doi:10.1371/journal.pone.0042543

Filoche, S., Lawton, B., Beard, A., Dowell, A., and Stone, P. (2017). New Screen on the Block: Non-invasive Prenatal Testing for Fetal Chromosomal Abnormalities. *J. Prim. Health Care* 9 (4), 248–253. doi:10.1071/hc16055

Fleischhacker, M., and Schmidt, B. (2007). Circulating Nucleic Acids (CNAs) and Cancer-A Survey. *Biochimica Biophysica Acta (BBA) - Rev. Cancer* 1775 (1), 181–232. doi:10.1016/j.bbcan.2006.10.001

Gamazon, E. R., and Stranger, B. E. (2015). The Impact of Human Copy Number Variation on Gene Expression: Figure 1. *Briefings Funct. genomics* 14 (5), 352–357. doi:10.1093/bfgp/elv017

Guan, X., Cai, M., Du, Y., Yang, E., Ji, J., and Wu, J. (2020). CVCDAP: An Integrated Platform for Molecular and Clinical Analysis of Cancer Virtual Cohorts. *Nucleic Acids Res.* 48 (W1), W463–w471. doi:10.1093/nar/gkaa423

Hastings, P. J., Lupski, J. R., Rosenberg, S. M., and Ira, G. (2009). Mechanisms of Change in Gene Copy Number. *Nat. Rev. Genet.* 10 (8), 551–564. doi:10.1038/nrg2593

He, C.-Z., Zhang, K.-H., Li, Q., Liu, X.-H., Hong, Y., and Lv, N.-H. (2013). Combined Use of AFP, CEA, CA125 and CAl9-9 Improves the Sensitivity for the Diagnosis of Gastric Cancer. *BMC Gastroenterol.* 13 (1), 87. doi:10.1186/1471-230x-13-87

Hu, W., Li, M., Zhang, Q., Liu, C., Wang, X., Li, J., et al. (2021). Establishment of a Novel CNV-Related Prognostic Signature Predicting Prognosis in Patients with Breast Cancer. *J. Ovarian Res.* 14 (1), 103. doi:10.1186/s13048-021-00823-y

Huang, C.-C., Du, M., and Wang, L. (2019). Bioinformatics Analysis for Circulating Cell-Free DNA in Cancer. *Cancers* 11 (6), 805. doi:10.3390/cancers11060805

Jin, X., Yan, J., Chen, C., Chen, Y., and Huang, W-K. (2021). Integrated Analysis of Copy Number Variation, Microsatellite Instability, and Tumor Mutation Burden Identifies an 11-Gene Signature Predicting Survival in Breast Cancer. *Front. Cell Dev. Biol.* 9, 721505. doi:10.3389/fcell.2021.721505

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28 (1), 27–30. doi:10.1093/nar/28.1.27

Kerkhof, J., Schenkel, L. C., Reilly, J., McRobbie, S., Aref-Eshghi, E., Stuart, A., et al. (2017). Clinical Validation of Copy Number Variant Detection from Targeted Next-Generation Sequencing Panels. *J. Mol. Diagnostics* 19 (6), 905–920. doi:10.1016/j.jmoldx.2017.07.004

Kourtis, N., Moubarak, R. S., Aranda-Orgilles, B., Lui, K., Aydin, I. T., Trimarchi, T., et al. (2015). FBXW7 Modulates Cellular Stress Response and Metastatic Potential through HSF1 Post-translational Modification. *Nat. Cell Biol.* 17 (3), 322–332. doi:10.1038/ncb3121

Le Page, C., Huntsman, D. G., Provencher, D. M., and Mes-Masson, A.-M. (2010). Predictive and Prognostic Protein Biomarkers in Epithelial Ovarian Cancer: Recommendation for Future Studies. *Cancers* 2 (2), 913–954. doi:10.3390/cancers2020913

Lee, Y. J., Chung, Y. S., Lee, J. Y., Nam, E. J., Kim, S. W., Kim, S., et al. (2018). Impact of Increased Utilization of Neoadjuvant Chemotherapy on Survival in Patients with Advanced Ovarian Cancer: Experience from a Comprehensive Cancer Center. *J. Gynecol. Oncol.* 29 (4), e63. doi:10.3802/jgo.2018.29.e63

Li, H., and Durbin, R. (2009). Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* 25 (14), 1754–1760. doi:10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* 25 (16), 2078–2079. doi:10.1093/bioinformatics/btp352

Liang, L., Fang, J.-Y., and Xu, J. (2016). Gastric Cancer and Gene Copy Number Variation: Emerging Cancer Drivers for Targeted Therapy. *Oncogene* 35 (12), 1475–1482. doi:10.1038/onc.2015.209

Liscia, D. S., Morizio, R., Venesio, T., Palenzona, C., Donadio, M., and Callahan, R. (1999). Prognostic Significance of Loss of Heterozygosity at Loci on Chromosome 17p13.3-ter in Sporadic Breast Cancer Is Evidence for a Putative Tumour Suppressor Gene. *Br. J. Cancer* 80 (5-6), 821–826. doi:10.1038/sj.bjc.6690427

Liu, C.-J., Hu, F.-F., Xia, M.-X., Han, L., Zhang, Q., and Guo, A.-Y. (2018). GSCALite: A Web Server for Gene Set Cancer Analysis. *Bioinformatics* 34 (21), 3771–3772. doi:10.1093/bioinformatics/bty411

Marnett, L. J. (2002). Oxy Radicals, Lipid Peroxidation and DNA Damage. *Toxicology* 181-182, 219–222. doi:10.1016/s0300-483x(02)00448-1

Mathai, R., Vidya, R., Reddy, B., Thomas, L., Udupa, K., Kolesar, J., et al. (2019). Potential Utility of Liquid Biopsy as a Diagnostic and Prognostic Tool for the Assessment of Solid Tumors: Implications in the Precision Oncology. *J. Clin. Med.* 8 (3), 373. doi:10.3390/jcm8030373

Matsui, A., Ihara, T., Suda, H., Mikami, H., and Semba, K. (2013). Gene Amplification: Mechanisms and Involvement in Cancer. *Biomol. Concepts* 4 (6), 567–582. doi:10.1515/bmc-2013-0026

Mazzucchelli, R., Colanzi, P., Pomante, R., Muzzonigro, G., and Montironi, R. (2000). Prostate Tissue and Serum Markers. *Adv. Clin. Path* 4 (3), 111–120.

McDonald, J. D., Daneshvar, L., Willert, J. R., Matsumura, K., Waldman, F., and Cogen, P. H. (1994). Physical Mapping of Chromosome 17p13.3 in the Region of a Putative Tumor Suppressor Gene Important in Medulloblastoma. *Genomics* 23 (1), 229–232. doi:10.1006/geno.1994.1481

Molparia, B., Nichani, E., and Torkamani, A. (2017). Assessment of Circulating Copy Number Variant Detection for Cancer Screening. *PLoS One* 12 (7), e0180647. doi:10.1371/journal.pone.0180647

Nakabayashi, M., Kawashima, A., Yasuhara, R., Hayakawa, Y., Miyamoto, S., Iizuka, C., et al. (2018). Massively Parallel Sequencing of Cell-free DNA in Plasma for Detecting Gynaecological Tumour-Associated Copy Number Alteration. *Sci. Rep.* 8 (1), 11205. doi:10.1038/s41598-018-29381-y

Namasu, C. Y., Katzerke, C., Bräuer-Hartmann, D., Wurm, A. A., Gerloff, D., Hartmann, J.-U., et al. (2017). ABR, a Novel Inducer of Transcription Factor C/EBPα, Contributes to Myeloid Differentiation and Is a Favorable Prognostic Factor in Acute Myeloid Leukemia. *Oncotarget* 8 (61), 103626–103639. doi:10.18632/oncotarget.22093

Ning, J., Wang, F., Zhu, K., Li, B., Shu, Q., and Liu, W., Characterizing the Copy Number Variation of Non-Coding RNAs Reveals Potential Therapeutic Targets and Prognostic Markers of LUSC. *Front. Genet.* 2021. 12. 779155. doi:10.3389/fgene.2021.779155

Noguchi, T., Sakai, K., Iwahashi, N., Matsuda, K., Matsukawa, H., Yahata, T., et al. (2020). Changes in the Gene Mutation Profiles of Circulating Tumor DNA Detected Using CAPP-Seq in Neoadjuvant Chemotherapy-Treated Advanced Ovarian Cancer. *Oncol. Lett.* 19 (4), 2713–2720. doi:10.3892/ol.2020.11356

Page, K., Guttery, D. S., Fernandez-Garcia, D., Hills, A., Hastings, R. K., Luo, J., et al. (2017). Next Generation Sequencing of Circulating Cell-Free DNA for Evaluating Mutations and Gene Amplification in Metastatic Breast Cancer. *Clin. Chem.* 63 (2), 532–541. doi:10.1373/clinchem.2016.261834

Pan, X., Hu, X., Zhang, Y.-H., Chen, L., Zhu, L., Wan, S., et al. (2019). Identification of the Copy Number Variant Biomarkers for Breast Cancer Subtypes. *Mol. Genet. Genomics* 294 (1), 95–110. doi:10.1007/s00438-018-1488-4

Panawala, L. (2017). *Difference Between DNA and RNA*. NY: ThoughtCo.

Pelissier, A., Bonneau, C., Chéreau, E., de La Motte Rouge, T., Fourchotte, V., Daraï, E., et al. (2014). CA125 Kinetic Parameters Predict Optimal Cytoreduction in Patients with Advanced Epithelial Ovarian Cancer Treated with Neoadjuvant Chemotherapy. *Gynecol. Oncol.* 135 (3), 542–546. doi:10.1016/j.ygyno.2014.09.005

Pool-Zobel, B. L., Selvaraju, V., Sauer, J., Kautenburger, T., Kiefer, J., Richter, K. K., et al. (2005). Butyrate May Enhance Toxicological Defence in Primary, Adenoma and Tumor Human Colon Cells by Favourably Modulating Expression of Glutathione S -transferases Genes, an Approach in Nutrigenomics. *Carcinogenesis* 26 (6), 1064–1076. doi:10.1093/carcin/bgi059

Powell, C. D., Paullin, T. R., Aoisa, C., Menzie, C. J., Ubaldini, A., and Westerheide, S. D. (2016). The Heat Shock Transcription Factor HSF1 Induces Ovarian Cancer Epithelial-Mesenchymal Transition in a 3D Spheroid Growth Model. *PLoS One* 11 (12), e0168389. doi:10.1371/journal.pone.0168389

Qidwai, T. (2020). Exploration of Copy Number Variation in Genes Related to Anti-malarial Drug Resistance in Plasmodium Falciparum. *Gene* 736, 144414. doi:10.1016/j.gene.2020.144414

Rahimi, M., Behjati, F., Khorram Khorshid, H. R., Karimlou, M., and Keyhani, E. (2020). The Relationship between KIT Copy Number Variation, Protein Expression, and Angiogenesis in Sporadic Breast Cancer. *Rep. Biochem. Mol. Biol.* 9 (1), 40–49. doi:10.29252/rbmb.9.1.40

Raman, L., Dheedene, A., De Smet, M., Van Dorpe, J., and Menten, B. (2018). WisecondorX: Improved Copy Number Detection for Routine Shallow Whole-Genome Sequencing. *Nucleic Acids Res.* 47 (4), 1605–1614. doi:10.1093/nar/gky1263

Raman, L., Dheedene, A., De Smet, M., Van Dorpe, J., and Menten, B. (2019). WisecondorX: Improved Copy Number Detection for Routine Shallow Whole-Genome Sequencing. *Nucleic Acids Res.* 47 (4), 1605–1614. doi:10.1093/nar/gky1263

Saha, S., Araf, Y., and Promon, S. K. (2022). Circulating Tumor DNA in Cancer Diagnosis, Monitoring, and Prognosis. *J. Egypt Natl. Canc Inst.* 34 (1), 8. doi:10.1186/s43046-022-00109-4

Sato, S., and Itamochi, H. (2014). Neoadjuvant Chemotherapy in Advanced Ovarian Cancer: Latest Results and Place in Therapy. *Ther. Adv. Med. Oncol.* 6 (6), 293–304. doi:10.1177/1758834014544891

Shao, X., Lv, N., Liao, J., Long, J., Xue, R., Ai, N., et al. (2019). Copy Number Variation Is Highly Correlated with Differential Gene Expression: a Pan-Cancer Study. *BMC Med. Genet.* 20 (1), 175. doi:10.1186/s12881-019-0909-5

Sharbatoghli, M., Vafaei, S., Aboulkheyr Es, H., Asadi-Lari, M., Totonchi, M., and Madjd, Z. (2020). Prediction of the Treatment Response in Ovarian Cancer: a ctDNA Approach. *J. Ovarian Res.* 13 (1), 124. doi:10.1186/s13048-020-00729-1

Sidiropoulos, K., Viteri, G., Sevilla, C., Jupe, S., Webber, M., Orlic-Milacic, M., et al. (2017). Reactome Enhanced Pathway Visualization. *Bioinformatics* 33 (21), 3461–3467. doi:10.1093/bioinformatics/btx441

Slenter, D. N., Kutmon, M., Hanspers, K., Riutta, A., Windsor, J., Nunes, N., et al. (2018). WikiPathways: A Multifaceted Pathway Database Bridging Metabolomics to Other Omics Research. *Nucleic Acids Res.* 46 (D1), D661–d667. doi:10.1093/nar/gkx1064

Tabarés-Seisdedos, R., and Rubenstein, J. L. R. (2009). Chromosome 8p as a Potential Hub for Developmental Neuropsychiatric Disorders: Implications for Schizophrenia, Autism and Cancer. *Mol. Psychiatry* 14 (6), 563–589. doi:10.1038/mp.2009.2

Tan, K. L., and Board, P. G. (1996). Purification and Characterization of a Recombinant Human Theta-Class Glutathione Transferase (GSTT2-2). *Biochem. J.* 315 (Pt 3), 727–732. doi:10.1042/bj3150727

Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: An Enhanced Web Server for Large-Scale Expression Profiling and Interactive Analysis. *Nucleic Acids Res.* 47 (W1), W556–W560. doi:10.1093/nar/gkz430

Vanderstichele, A., Busschaert, P., Smeets, D., Landolfo, C., Van Nieuwenhuysen, E., Leunen, K., et al. (2017). Chromosomal Instability in Cell-Free DNA as a Highly Specific Biomarker for Detection of Ovarian Cancer in Women with Adnexal Masses. *Clin. Cancer Res.* 23 (9), 2223–2231. doi:10.1158/1078-0432.ccr-16-1078

Vihervaara, A., and Sistonen, L. (2014). HSF1 at a Glance. *J. Cell Sci.* 127 (Pt 2), 261–266. doi:10.1242/jcs.132605

Vives-Usano, M., García Pelaez, B., Román Lladó, R., Garzón Ibañez, M., Aldeguer, E., Rodriguez, S., et al. (2021). Analysis of Copy Number Variations in Solid Tumors Using a Next Generation Sequencing Custom Panel. *J. Mol. Pathol.* 2 (2), 123–134. doi:10.3390/jmp2020013

Vydra, N., Toma, A., and Widlak, W. (2014). Pleiotropic Role of HSF1 in Neoplastic Transformation. *Curr. Cancer Drug Targets* 14 (2), 144–155. doi:10.2174/1568009614666140122155942

Walker, L. C., Pearson, J. F., Wiggins, G. A. R., Giles, G. G., Hopper, J. L., and Southey, M. C. (2017). Increased Genomic Burden of Germline Copy Number Variants Is Associated with Early Onset Breast Cancer: Australian Breast Cancer Family Registry. *Breast Cancer Res.* 19 (1), 30. doi:10.1186/s13058-017-0825-6

Wang, G., Cao, P., Fan, Y., and Tan, K. (2020). Emerging Roles of HSF1 in Cancer: Cellular and Molecular Episodes. *Biochimica Biophysica Acta (BBA) - Rev. Cancer* 1874 (1), 188390. doi:10.1016/j.bbcan.2020.188390

Willert, J. R., Daneshvar, L., Sheffield, V. C., and Cogen, P. H. (1995). Deletion of Chromosome Arm 17p DNA Sequences in Pediatric High-Grade and Juvenile Pilocytic Astrocytomas. *Genes Chromosom. Cancer* 12 (3), 165–172. doi:10.1002/gcc.2870120303

Willyard, C. (2015). Copy Number Variations' Effect on Drug Response Still Overlooked. *Nat. Med.* 21 (3), 206. doi:10.1038/nm0315-206

Woods, D., and Turchi, J. J. (2013). Chemotherapy Induced DNA Damage Response. *Cancer Biol. Ther.* 14 (5), 379–389. doi:10.4161/cbt.23761

Yu, Y., Cao, J., Wu, W., Zhu, Q., Tang, Y., Zhu, C., et al. (2019). Genome-wide Copy Number Variation Analysis Identified ANO1 as a Novel Oncogene and Prognostic Biomarker in Esophageal Squamous Cell Cancer. *Carcinogenesis* 40 (10), 1198–1208. doi:10.1093/carcin/bgz077

Yun, H. S., Baek, J.-H., Yim, J.-H., Um, H.-D., Park, J. K., Song, J.-Y., et al. (2016). Radiotherapy Diagnostic Biomarkers in Radioresistant Human H460 Lung Cancer Stem-like Cells. *Cancer Biol. Ther.* 17 (2), 208–218. doi:10.1080/15384047.2016.1139232

Zeng, J., Yin, J., Song, X., Jin, Y., Li, Y., and Pan, L. (2016). Reduction of CA125 Levels During Neoadjuvant Chemotherapy Can Predict Cytoreduction to No Visible Residual Disease in Patients with Advanced Epithelial Ovarian Cancer, Primary Carcinoma of Fallopian Tube and Peritoneal Carcinoma. *J. Cancer* 7 (15), 2327–2332. doi:10.7150/jca.16761

Zhang, C. Q., Williams, H., Prince, T. L., and Ho, E. S. (2017). Overexpressed HSF1 Cancer Signature Genes Cluster in Human Chromosome 8q. *Hum. Genomics* 11 (1), 35. doi:10.1186/s40246-017-0131-5

Zhong, W., Wang, D., Yao, B., Chen, X., Wang, Z., Qu, H., et al. (2021). Integrative Analysis of Prognostic Long Non-coding RNAs with Copy Number Variation in Bladder Cancer. *J. Zhejiang Univ. Sci. B* 22 (8), 664–681. doi:10.1631/jzus.b2000494

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.

Copyright © 2022 Sharbatoghli, Fattahi, Aboulkheyr Es, Akbari, Akhavan, Ebrahimi, Asadi-Lari, Totonchi and Madjd. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 05 August 2022
DOI 10.3389/fgene.2022.922472

Check for updates

OPEN ACCESS

EDITED BY
Ehsan Nazemalhosseini-Mojarad,
Shahid Beheshti University of Medical Sciences, Iran

REVIEWED BY
Mohammad Taheri-Ledari,
University of Tehran, Iran
Vahid Chaleshi,
Shahid Beheshti University of Medical Sciences, Iran

*CORRESPONDENCE
Yang Bu,
boyang1976@163.com
Bendong Chen,
bendong8511@hotmail.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 18 April 2022
ACCEPTED 05 July 2022
PUBLISHED 05 August 2022

CITATION
Liu K, Cui L, Li C, Tang C, Niu Y, Hao J, Bu Y and Chen B (2022), Pan-cancer analysis of the prognostic and immunological role of ANLN: An onco-immunological biomarker.
Front. Genet. 13:922472.
doi: 10.3389/fgene.2022.922472

COPYRIGHT
© 2022 Liu, Cui, Li, Tang, Niu, Hao, Bu and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Pan-cancer analysis of the prognostic and immunological role of ANLN: An onco-immunological biomarker

Kejun Liu[1,2†], Lei Cui[3†], Cunquan Li[2], Chaofeng Tang[1,2], Yiming Niu[3], Ji Hao[3], Yang Bu[3,4*] and Bendong Chen[1,2,3*]

[1]Department of Hepatobiliary Surgery, General Hospital of Ningxia Medical University, Yinchuan, China, [2]Ningxia Hepatobiliary and Pancreatic Surgical Diseases Clinical Medical Research Center, Yinchuan, China, [3]School of Clinical Medicine, Ningxia Medical University, Yinchuan, China, [4]Department of Hepatobiliary Surgery, People's Hospital of Ningxia Hui Autonomous Region, Yinchuan, China

Anillin actin-binding protein (ANLN) is crucially involved in cell proliferation and migration. Moreover, ANLN is significantly in tumor progression in several types of human malignant tumors; however, it remains unclear whether ANLN acts through common molecular pathways within different tumor microenvironments, pathogeneses, prognoses and immunotherapy contexts. Therefore, this study aimed to perform bioinformatics analysis to examine the correlation of ANLN with tumor immune infiltration, immune evasion, tumor progression, immunotherapy, and tumor prognosis. We observed increased ANLN expression in multiple tumors, which could be involved in tumor cell proliferation, migration, infiltration, and prognosis. The level of ANLN methylation and genetic alteration was associated with prognosis in numerous tumors. ANLN facilitates tumor immune evasion through different mechanisms, which involve T-cell exclusion in different cancer types and tumor-infiltrating immune cells in colon adenocarcinoma, kidney renal clear cell carcinoma, liver hepatocellular carcinoma, and prostate adenocarcinoma. Additionally, ANLN is correlated with immune or chemotherapeutic outcomes in malignant cancers. Notably, ANLN expression may be a predictive biomarker for the response to immune checkpoint inhibitors. Taken together, our findings suggest that ANLN can be used as an onco-immunological biomarker and could serve as a hallmark for tumor screening, prognosis, individualized treatment design, and follow-up.

KEYWORDS

pan-cancer, anillin actin-binding protein (ANLN), immunization, prognosis, bioinformatics analysis

Introduction

Worldwide, cancer remains the leading cause of morbidity and mortality; moreover, it imposes huge health and economic burden on society (Sung et al., 2021). Current therapies mainly include surgery, radiotherapy, and immunotherapy; however, delayed diagnosis or distant metastasis impedes most patients from undergoing radical surgery (Siegel et al., 2021). Recently, immune checkpoint blockade (ICB) has been shown to be efficacious in immunotherapy for various cancer types. Recent advances in ICB therapy, including those targeting programmed cell death protein 1 (PD-1)/PD-ligand 1 (PD-L1) and/or cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), have allowed strong enhancement of anti-tumor immune responses (Hargadon et al., 2018). Although there have been recent positive reports regarding immunotherapy, its effectiveness widely varies among individuals and tumor types. For example, in some patients with cancer, the objective response rate of anti-PD-L1/PD-1 therapy alone is only 10%–30% (Barrueto et al., 2020; Genova et al., 2022). This could be mainly attributed to the fact that current treatments specifically focus on the cancer cells rather than comprehensively addressing the tumor immune microenvironment (TIME).

Tumor cells and the TIME influence each other. Various immune cells, stromal cells, and cytokines inherently present in the TIME can interact with tumor cells. Moreover, regulation of the immune system network has complex interactions with tumor cells and crucially affects tumor development and the response to immunotherapy (Li et al., 2020a). Tumor cells make direct or indirect contact with various TIME components, with continuous and dynamic changes in the surrounding environment that allow them to evade the host immune surveillance (Mao et al., 2021; Liu et al., 2022). Additionally, the TIME is heterogeneous. The composition of tumor-infiltrating immune cells varies according to the cancer type. Therefore, it is important to elucidate the effects of multiple factors in order to accurately predict the overall effect of the TIME on the tumor immune response, which could improve survival (Morad et al., 2021). Moreover, it could inform the identification of predictive biomarkers for the efficacy of ICBs from the TIME perspective. Additionally, elucidating the mechanisms underlying the efficacy of ICBs and tumor-TIME interactions could inform individualized anti-tumor immunotherapy as well as monitoring of treatment efficacy and disease progression.

Anillin actin-binding protein (ANLN), which is a protein-coding gene, is localized on chromosome 7q14.2 and encodes 1,125 amino acid proteins. It is crucially involved in human diseases and multiple cellular functions, including cell growth, cytokinesis, mitosis, and cell migration (Xiao et al., 2020). Through interactions with cytoskeletal components, ANLN is involved in regulating tumor cell growth, invasion, and metastasis (Naydenov et al., 2021). ANLN is overexpressed in different tumor types, including breast cancer (Downs et al., 2019), lung cancer (Suzuki et al., 2005; Deng et al., 2021), pancreatic cancer (Idichi et al., 2017; Wang et al., 2019), liver cancer (Lian et al., 2018), bladder cancer (Zeng et al., 2017), and colorectal cancer (Wang et al., 2016). Moreover, it worsens the prognosis by promoting malignant tumor progression.

However, the pathogenic role of ANLN in various cancers remains unclear. Further, it remains unclear whether ANLN is involved in different tumor immune microenvironments and can influence the therapeutic response through certain common molecular mechanisms. Accordingly, through bioinformatics analysis, we aimed to elucidate the relationship of ANLN with tumor prognosis, immune infiltration, and therapeutic response in order to determine whether ANLN can be used as a biomarker for cancer screening, prognosis, individualized therapy design, and follow-up.

Materials and methods

Data source

We downloaded the clinical information and processed RNA-sequencing data of patients with 33 types of cancer from The Cancer Genome Atlas (TCGA), which is publicly available, and included 9,807 tumor tissues (Blum et al., 2018). Data regarding normal tissue expression were obtained from the Genotype-Tissue Expression (GTEx, gtexportal.org/home/) project (Carithers and Moore, 2015). We used a database of novel junctions identified in TCGA and GTEx to determine the ANLN expression in various types of human cancers and adjacent normal tissues. The UALCAN portal allows the evaluation of epigenetic regulation of gene expression through promoter methylation as well as pan-cancer gene expression analysis (Chandrashekar et al., 2017). We used the "CPTAC Analysis" module provided by the UALCAN portal to compare ANLN expression levels between normal and tumor tissues. Figure 1 shows the study flowchart.

The physiological condition and protein interaction of anillin actin-binding protein

The Human Protein Atlas (proteinatlas.org/) web server can map all human proteins using an integration of various omics technologies, including antibody-based imaging, mass spectrometry-based proteomics, transcriptomics, and systems biology (Uhlen et al., 2015; Thul et al., 2017; Uhlen et al., 2017; Karlsson et al., 2021). PROTTER (wlab.ethz.ch/protter/

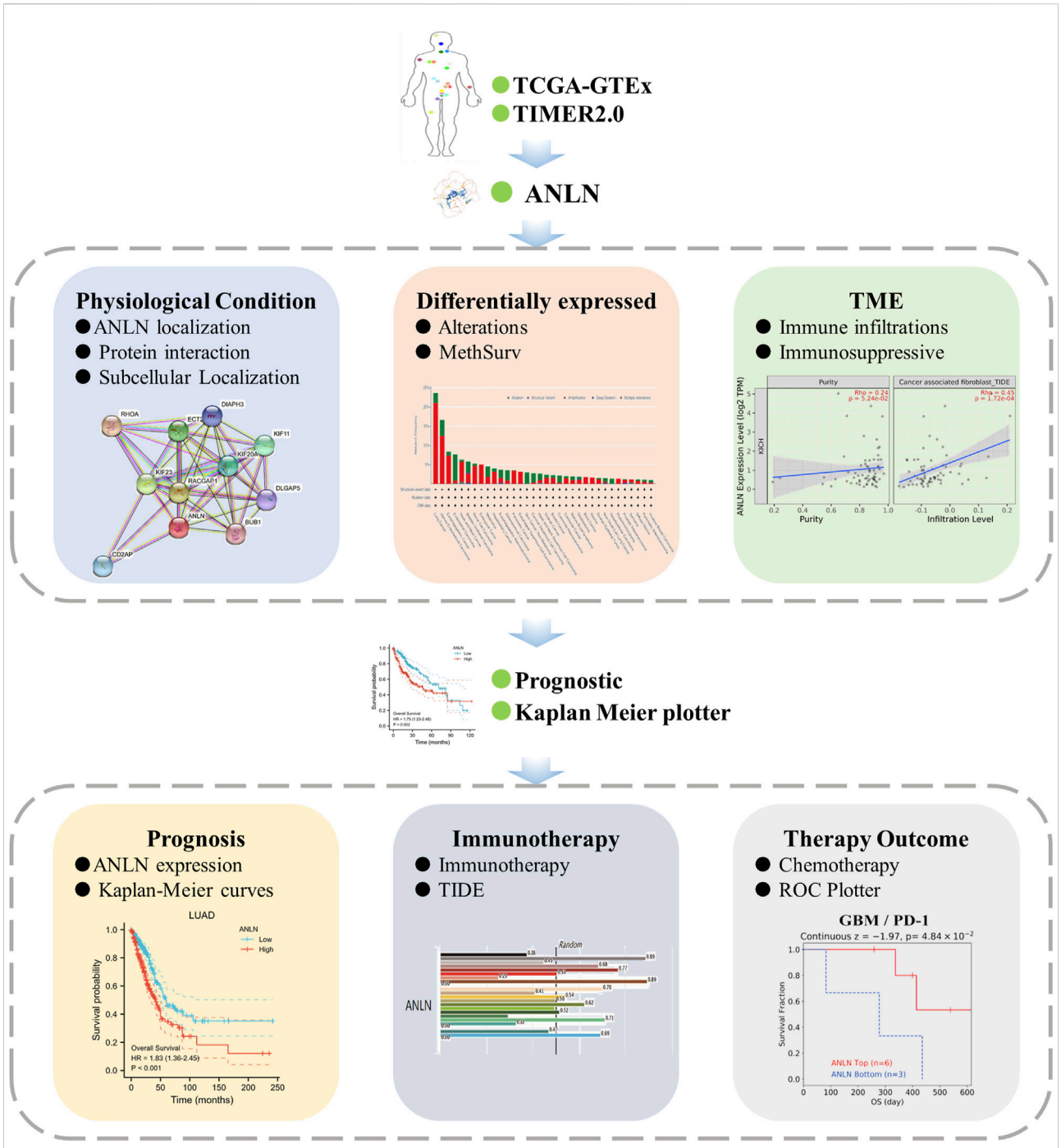


FIGURE 1
The workflow of the study.

start/) is an online tool for protein-form visualization as well as annotation and prediction of sequence features with experimental proteomics evidence (Omasits et al., 2014). String (cn.string-db.org/) is an online database of known and predicted protein-protein interaction (PPI) networks (confidence score cutoff = 0.7) (Szklarczyk et al., 2021).

Differential anillin actin-binding protein expression in tumor pathological stages and prognostic analysis

Kaplan Meier plotter (kmplot.com/analysis/) is an online tool for analyzing the association of ANLN expression and

genetic alterations with therapeutic outcomes in different tumors (Lanczky and Gyorffy, 2021). Based on the median ANLN expression levels, patient samples were allocated to the high and low ANLN expression groups.

The Gene Expression Profiling Interactive Analysis (GEPIA) (gepia.cancer-pku.cn/index.html) is a novel interactive web server that provides customizable functions, including tumor/normal differential expression analysis, profiling according to cancer types or pathological stages, and patient survival analysis. GEPIA uses a standard processing pipeline to analyze the RNA sequencing expression data of 9,736 tumors and 8,587 normal samples from the TCGA and the GTEx projects (Tang et al., 2017). The GEPIA database was used to evaluate differential ANLN expression between tumor tissues and para-tumor tissues, as well as ANLN expression in different tumor pathological stages.

Tumor DNA methylation and gene mutation characteristics

MethSurv (version MethSurv©2017, biit.cs.ut.ee/methsurv/), which is a web tool for survival analysis based on the CpG methylation patterns (Modhukur et al., 2018), was used to analyze DNA methylation sites of ANLN based on the TCGA data. Moreover, we used the Kaplan-Meier test in the MethSurv database to perform overall survival (OS) analysis for each CpG site. Significance was tested using a likelihood ratio test. Density plots were used to visualize cut-off points for the dichotomy of methylation levels (Supplementary Figure S1). The plots were drawn with the β-value as the abscissa and the corresponding sample proportion as the ordinate. The numbers in red indicate the cutoff points used to group patients.

The cBioPortal for Cancer Genomics (cbioportal.org/) is an open-access, open-source resource for interactive exploration of multiple cancer genomics datasets (Cerami et al., 2012). Currently, it stores data regarding the DNA copy number, mRNA and microRNA expression, non-synonymous mutations, protein and phosphoprotein level, DNA methylation, and limited clinical information (Cerami et al., 2012). The ANLN mutation signatures in pan-tumors were investigated on cBioPortal OncoPrint using data obtained from the TCGA Pan-Cancer Atlas (PanCanAtlas) cohort.

Analysis of tumor immune cell infiltration

Tumor IMmune Estimation Resource (TIMER2.0) (timer.cistrome.org/), which is a web server for analyzing tumor-infiltrating immune cells using TCGA (Li et al., 2020b), was used to analyze the relationship of immune cell infiltration abundance with ANLN expression levels in different cancer types. These infiltrating immune cells include B cells, $CD8^+$ T-cells, $CD4^+$ T-cells, macrophages, neutrophils, dendritic cells (DCs), myeloid-derived suppressor cells (MDSCs), cancer-associated fibroblasts (CAFs), M2-tumor-associated macrophages (TAMs), and T-regulatory lymphocytes (Treg cells) *via* gene modules (Summerfield et al., 2001). Spearman's correlation coefficients were used to analyze the correlation between ANLN and immune cell infiltration abundance. Statistical significance was set at $p < 0.05$. Data were visualized in GraphPad Prism Software (Version 8.0.0 for Windows). A heatmap was used to visualize the correlation of immune cell infiltration abundance with ANLN expression in 33 cancer types in the TCGA.

Analysis of the correlation between the tumor treatment response and anillin actin-binding protein expression

The Tumor Immune Dysfunction and Exclusion (TIDE) database (tide.dfci.harvard.edu) comprises omics data regarding over 33K samples obtained from 188 tumor cohorts in public databases, 998 tumors from 12 clinical studies on ICBs, and eight CRISPR screens. It provides a computational method for modeling the two primary mechanisms of tumor immune evasion (Jiang et al., 2018; Fu et al., 2020). Gene prioritization of ANLN was assessed based on the treatment response to ICBs while knockout phenotypes were identified through CRISPR screening. The z-score in Cox-PH regression was used to assess the effect of gene expression on survival in the ICB-treatment cohort. CRISPR screening was used to assess the role of knockout-mediated lymphocyte-induced tumor death in cancer models. Based on the TIDE biomarker assessment module, we examined the overall predictive utility of ANLN for treatment response outcomes and OS across cancer types; further, we compared ANLN with standardized biomarkers for tumor immune response.

The ROC Plotter (rocplot.org/) is the first online transcriptome-level validation tool for predictive biomarkers. It can link gene expression to treatment response in patients with breast, ovarian, colorectal, and glioblastoma cancer (Fekete and Gyorffy, 2019). We used the ROC plotter to assess the predictive utility of ANLN expression for treatment response in patients with cancer.

Statistical analysis

All expression data were normalized and log-transformed using log2 (data + 1). R software (V3.8.3) was used to compare differential ANLN expression between tumor and normal tissues. Between-group comparisons were performed using t-tests (ns, $p \geq 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). The R package [ggplot2 (version 3.3.3)] was used for visualization. The

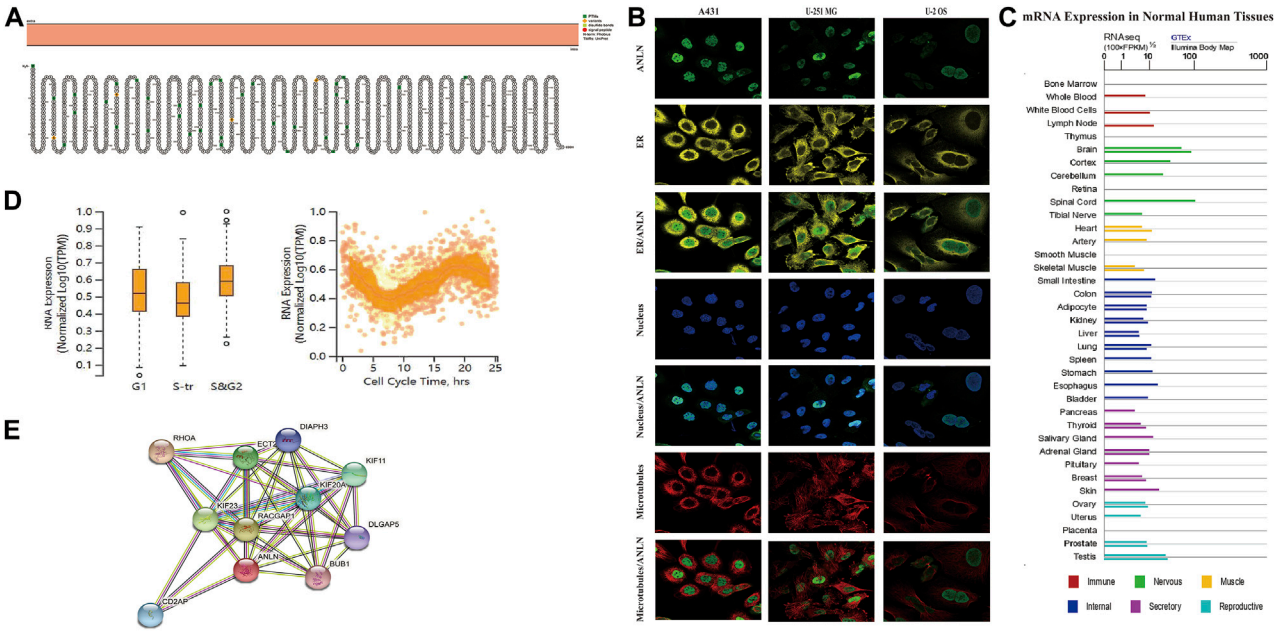


FIGURE 2
Topology, localization, expression profile, and protein interaction network of ANLN proteins. **(A)** Topology of ANLN proteins showing intracellular membrane localization. **(B)** Immunofluorescence staining of the subcellular distribution of ANLN in the nucleus, endoplasmic reticulum, and microtubules of A-431, U-2 OS, and U-251 MG was adopted from the HPA database. **(C)** Bar graphs of ANLN mRNA expression in various normal human tissues from the TCGA_GTEx database. **(D)** The plot of single-cell RNA sequencing data from the U-2 OS cell line shows a correlation between ANLN mRNA expression and cell cycle progression. **(E)** Protein interactions network of ANLN.

Wilcoxon rank-sum test was used to examine the relationship of ANLN expression with pathologic/cytogenetic features. For survival analysis, the hazard ratio (HR) and 95% confidence intervals (CI) were derived from Cox proportional hazards models. The p-value was determined using Kaplan-Meier analysis with the log-rank test. Statistical significance was set at $p < 0.05$.

Results

Topology, localization, expression profile, and protein interaction network of anillin actin-binding protein

Under physiological conditions, the topology of ANLN proteins indicated intracellular membrane localization (Figure 2A). To characterize the intracellular localization of ANLN, we used an indirect immunofluorescence assay to investigate the distribution of ANLN in the endoplasmic reticulum (ER), nucleus, and microtubules of A-431, U-2 OS, and U-251 MG cells. The Human Protein Atlas database showed that ANLN localization was mainly distributed in the nucleus, but not in the ER or microtubules, of these cells (Figure 2B). In addition, ANLN mRNA expressions occur in various normal human tissues, including the immune, visceral, nervous, endocrine, muscle, and reproductive tissues (Figure 2C). Based on the fluorescent ubiquitination-based cell cycle indicator system, single-cell RNA sequencing analysis of U-2 OS cells revealed an association of increased ANLN RNA expression with cell cycle progression (Figure 2D). Moreover, ANLN expression was correlated with the cell cycle progression in the G1, S, and S/G2 phases (Figure 2D). PPI analysis revealed that proteins related to ANLN included RHOA, RACGAP1, KIF23, ECT2, KIF11, KIF20A, CD2AP, DIAPH3, DLGAP5, and BUB1 (Figure 2E).

Anillin actin-binding protein expression in pancytopenia

Based on the TCGA_GTEx database, there was significantly higher ANLN expression in tumor tissues than in adjacent normal tissues in adrenocortical carcinoma (ACC), bladder urothelial carcinoma (BLCA), breast invasive carcinoma (BRCA), cervical squamous cell carcinoma (CESC), cholangiocarcinoma (CHOL), colon adenocarcinoma (COAD), lymphoid neoplasm diffuse large B-cell lymphoma (DLBC), esophageal carcinoma (ESCA), head and neck squamous cell carcinoma (HNSC), kidney chromophobe (KICH), kidney renal

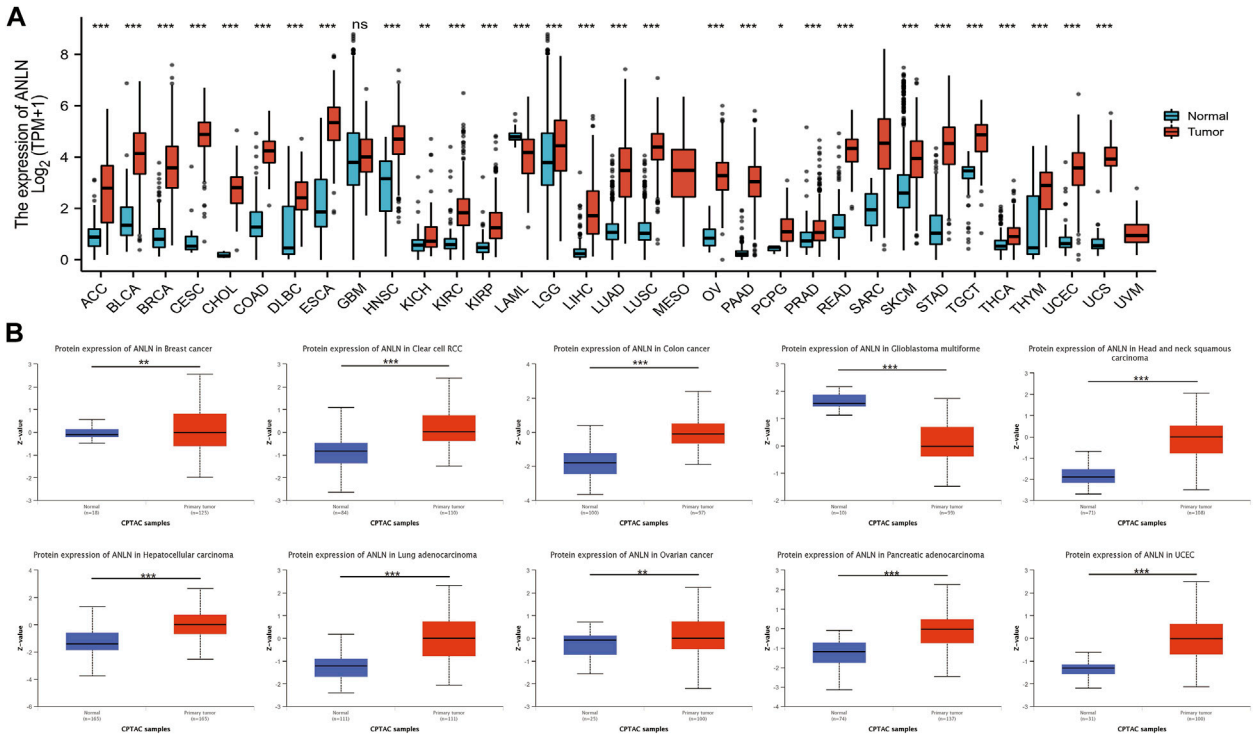


FIGURE 3
Upregulation of ANLN mRNA expression in pan-cancer. **(A)** Expression levels of ANLN in TCGA_GTEx database in different cancer types. Log2 (TPM + 1) was used for log scale. **(B)** Expression levels of ANLN protein in normal and different tumor tissues from the CPTAC database. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), brain lower grade glioma (LGG), liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), ovarian serous cystadenocarcinoma (OV), pancreatic adenocarcinoma (PAAD), pheochromocytoma and paraganglioma (PCPG), prostate adenocarcinoma (PRAD), rectum adenocarcinoma (READ), skin cutaneous melanoma (SKCM), stomach adenocarcinoma (STAD), testicular germ cell tumors (TGCT), thyroid carcinoma (THCA), thymoma (THYM), uterine corpus endometrial carcinoma (UCEC), and uterine carcinosarcoma (UCS) (Figure 3A). Further analysis using the "CPTAC Analysis" confirmed ANLN overexpression in BRCA, renal cell carcinoma, COAD, HNSC, LIHC, LUAD, OV, PAAD, and UCEC tumors compared with in normal tissues; however, there was low ANLN expression in the glioblastoma multiforme (Figure 3B).

Correlation of anillin actin-binding protein with pathological stage and prognosis of pan-cancer

To investigate the relationship between ANLN expression and the clinicopathological features of pan-cancer, we evaluated ANLN expression in patients with pathological stages I, II, III, and IV cancers. The TCGA database showed that ANLN expression was significantly upregulated in ACC, BLCA, KIRC, LIHC, and LUAD with advanced pathological stages (Figure 4A). In pan-cancer prognostic analysis, Kaplan-Meier survival curves revealed a significant association of upregulated ANLN expression with poor OS in ACC, BLCA, BRCA, KIRC, KIRP, LIHC, LUAD, mesothelioma (MESO), and PAAD; contrastingly, low ANLN expression was associated with poor OS in THYM (Figure 4B).

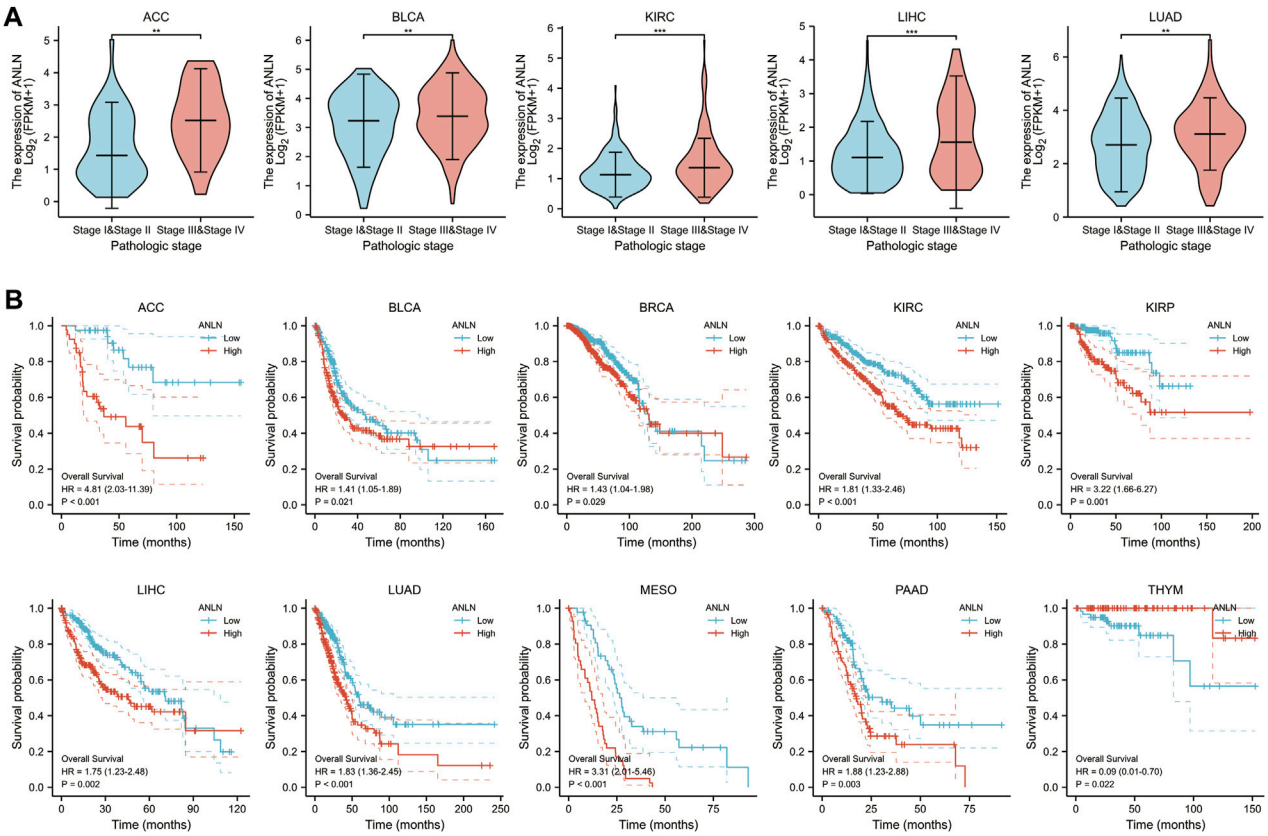


FIGURE 4
Correlation of ANLN expression with pathological stage and prognosis of tumors. **(A)** Correlation of ANLN expression with pathological stages of ACC, BLCA, KIRC, LIHC, and LUAD; **(B)** Correlation of ANLN expression with a prognosis of patients with ACC, BLCA, BRCA, KIRC, KIRP, LIHC, LUAD, MESO, PAAD, and THYM. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

DNA methylation and mutational features of anillin actin-binding protein in pan-cancer

DNA methylation directly affects cancer development and progression. We used the MethSurv to investigate DNA methylation of ANLN in terms of the prognostic value of each CpG site in cancer. Prognosis was correlated with the methylation level of four CpG sites, i.e., cg00267323, cg04455682, cg04897631, and cg12284836 ($p < 0.05$) (Table 1). Patients with low ANLN methylation of CpG site cg12284836 had a worse OS than patients with high ANLN methylation in this site for CESC, COAD, KIRC, LGG, LIHC, and UCEC (Figure 5).

The cBioPortal (TCGA, Pan-Cancer Atlas) database showed that lung cancer had the highest frequency of ANLN mutation (≈ 23.68%) (Figure 6A). As shown in Figure 6A, ANLN amplification and mutation were the most common alteration type. Moreover, among the patients with tumors, patients with ANLN mutations had worse progression-free survival (PFS); however, there were no significant differences in OS, disease-free survival, and disease-specific survival (DSS) (Figure 6B).

Correlation of anillin actin-binding protein with tumor immunity

Based on the TCGA database, ANLN expression in COAD, KIRC, LIHC, and PRAD showed a significant positive correlation with infiltration of six immune cells (B cells, $CD8^+$ T-cells, $CD4^+$ T-cells, macrophages, neutrophils, and DCs) ($p < 0.05$; Figure 7A). Additionally, we assessed the correlation of ANLN expression with infiltration of four immunosuppressive cells (MDSCs, CAFs, M2-TAMs, and Treg cells), which promote T-cell rejection. ANLN expression was positively correlated with myeloid-derived suppressor cell infiltration in ACC, BLCA, BRCA, BRCA-Basal, BRCA-Her2, BRCA-LumA, BRCA-Lumb, CESC, CHOL, COAD, ESCA, HNSC, HNSC-HPV−, HNSC-HPV+, KICH, KIRC, KIRP, LIHC, LUAD, LUSC, MESO, OV, PAAD, PCPG, PRAD, READ, SARC,

TABLE 1 Effect of hypermethylation level on prognosis in cancer.

CpG	Cancer	HR	CI	*p*-value	Best-split
cg04455682	ACC	0.296	(0.133–0.658)	0.002	median
cg04455682	CESC	0.449	(0.241–0.838)	0.006	q75
cg12284836	CESC	0.379	(0.188–0.763)	0.002	q75
cg12284836	COAD	0.581	(0.357–0.944)	0.026	median
cg04897631	GBM	0.600	(0.362–0.995)	0.038	q75
cg04897631	KICH	0.209	(0.056–0.778)	0.022	q25
cg00267323	KIRC	0.441	(0.258–0.754)	0.001	q75
cg04455682	KIRC	0.449	(0.299–0.674)	0.000	q25
cg04897631	KIRC	0.563	(0.379–0.834)	0.004	median
cg12284836	KIRC	0.566	(0.332–0.964)	0.026	q75
cg12284836	LGG	0.586	(0.400–0.858)	0.008	q25
cg00267323	LIHC	2.058	(1.276–3.319)	0.001	q25
cg04455682	LIHC	2.349	(1.543–3.574)	0.000	mean
cg12284836	LIHC	0.612	(0.431–0.869)	0.005	median
cg04455682	MESO	1.979	(1.174–3.336)	0.014	q75
cg04897631	PAAD	0.642	(0.430–0.960)	0.030	median
cg00267323	UCEC	2.483	(1.306–4.721)	0.002	q25
cg04455682	UCEC	2.168	(1.164–4.039)	0.008	q25
cg12284836	UCEC	0.598	(0.366–0.976)	0.047	q25
cg04897631	UCS	2.176	(0.971–4.878)	0.045	q25

SKCM, SKCM-Metastasis, SKCM-Primary, STAD, THYM, UCEC, and uveal melanoma ($p < 0.05$). ANLN expression was positively correlated with Tregs infiltration in KICH, LIHC, PCPG, and THCA ($p < 0.05$); ANLN expression was positively correlated with CAF infiltration in ACC, BLCA, BRCA-LumA, ESCA, HNSC, HNSC-HPV-, KICH, KIRC, KIRP, LIHC, LUAD, LUSC, MESO, PAAD, PCPG, PRAD, SKCM, SKCM-Metastasis, THCA, UCEC, and UCS ($p < 0.05$). ANLN expression was positively correlated with M2-TAM infiltration in KIRC and PRAD ($p < 0.05$) (Figure 7B).

We compared the predictive utility of ANLN and other standardized biomarkers for the efficacy of ICB. Among the ICB subgroups, 10 had ANLN-AUC > 0.5 (Figure 7C). ANLN showed a higher predictive value than TMB, T.Clonality, and B.Clonality.

Anillin actin-binding protein is strongly correlated with tumor treatment outcomes

We evaluated the effect of ANLN expression on chemotherapy response in a clinical cancer cohort. Patients with breast and colorectal cancer who presented high ANLN expression showed resistance to chemotherapy. Contrastingly, among patients with glioblastoma and ovarian cancer, patients with high ANLN expression showed greater benefits from chemotherapy than those with low ANLN expression (Figure 8A). Additionally, high ANLN expression was associated with clinical benefits derived from ICB therapy (PD-1), and thus longer OS, in patients with bladder cancer and glioblastoma. Contrastingly, in patients with kidney cancer and melanoma, low ANLN expression was associated with clinical benefits derived from ICB therapy (PD-L1 or PD-1), and thus longer OS (Figure 8B). In patients with bladder cancer and melanoma, high ANLN expression was positively correlated with CTL levels (Figure 8B).

Discussion

Recently, there have been increasing studies using Pan-Cancer Analysis of Whole Genomes to reveal genetic mutations, RNA changes, and oncogenes associated with cancer development and progression. Additionally, pan-cancer analysis has revealed common aspects among cancer types regarding key driver gene mutations, tumor-associated signaling pathway activation, immune escape, and microbial features. This suggests that the same therapeutic strategy can be applied for common features in different cancer types (Chen et al., 2021). Therefore, Pan-Cancer analysis may provide novel biomarkers for early diagnosis, prognosis, and treatment efficacy in patients with cancer.

ANLN, which is a key protein in cell division, can combine with GTP-Rho, actin, and cytokinin. ANLN is crucially involved in cell proliferation, differentiation, adhesion, migration, apoptosis, and cycle progression (Naydenov et al., 2021). Abnormal ANLN expression causes dysregulation of cell division and is crucially involved in cancer development, cancer cell proliferation, and metastasis. We found that ANLN mRNA expression was positively correlated with cell cycle progression; moreover, most tumors showed high ANLN expression, which was significantly correlated with poor prognosis and tumor progression. In bladder cancer, there was ANLN overexpression, which could promote tumor cell proliferation, migration, and infiltration; moreover, this was confirmed *in vitro* experiments. Nuclear ANLN overexpression is significantly correlated with tumor stage, histological grade, and tumor invasion; additionally, it is an independent risk factor for DSS (Liang et al., 2015; Zeng et al., 2017). In gastric cancer, ANLN can promote cell proliferation by activating the Wnt/β-catenin signaling pathway (Pandi et al., 2014). Li et al. (2020c) reported that the CDK1-PLK1-SGOL2-ANLN signaling axis mediating abnormal cell cycle division may be crucially involved in the development of hepatocellular carcinoma (HCC). Moreover, tumor tissues present with high ANLN expression and decreased ANLN methylation in HCC. Further, ANLN silencing may inhibit HCC cell proliferation in the G2/M phase as well as their migration and invasion abilities (Zhang et al., 2021). Contrastingly, the metastatic foci of HCC showed significantly higher cytosolic ANLN expression than the primary tumor tissue; additionally, high cytosolic ANLN expression

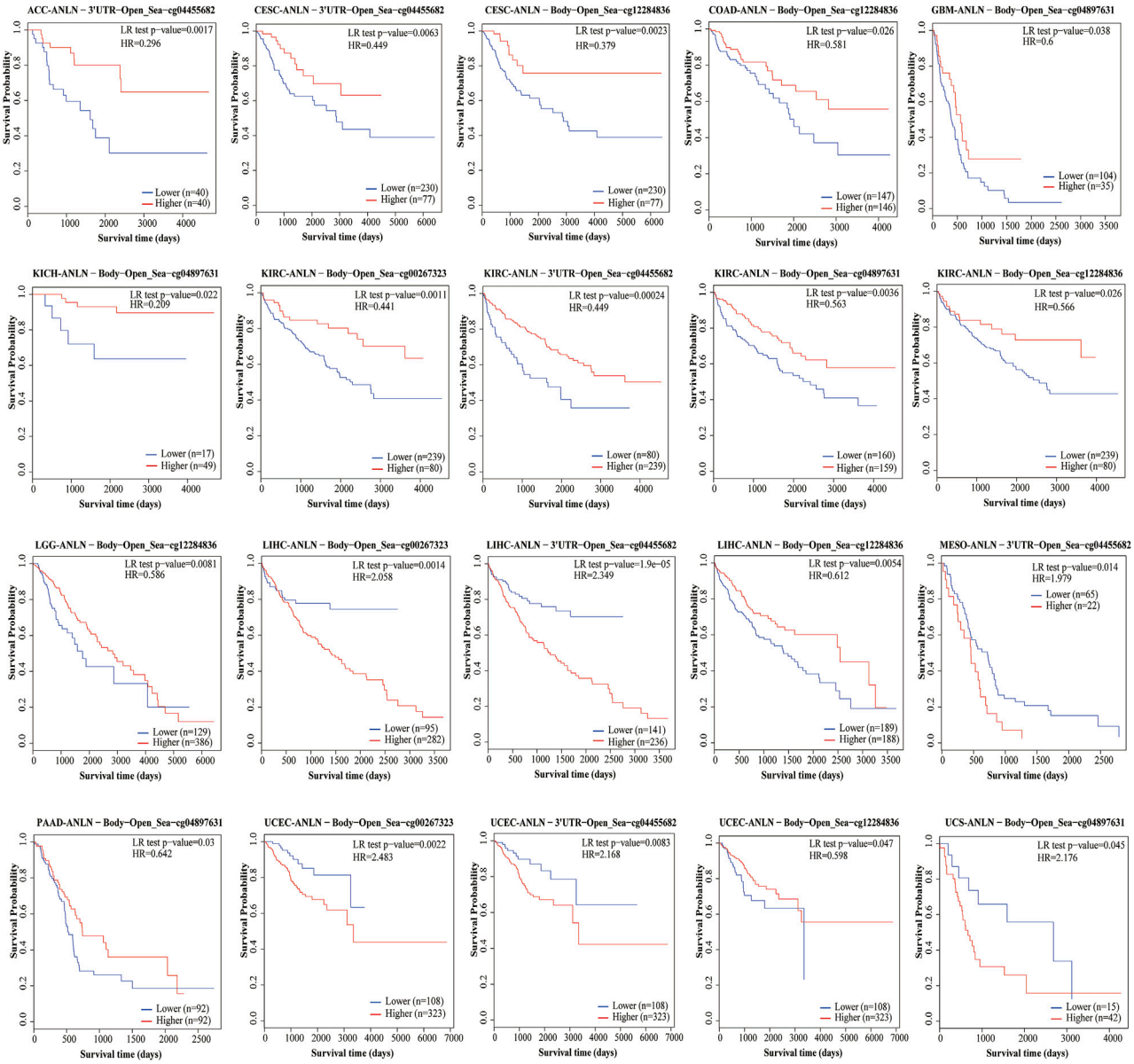


FIGURE 5
The Kaplan-Meier survival of the promoter methylation of ANLN.

was an independent risk factor for 5-year OS after hepatic resection. This study suggested that ANLN could be a potential therapeutic target and prognostic biological marker for HCC (Zhang et al., 2021). In gastric cancer, decreased ANLN expression affects downstream DNA damage repair-related pathways, resulting in more irreparable DNA damage in S-phase cells and eventually apoptosis (Wei et al., 2020). Patients with non-small cell lung cancer present with the highest ANLN expression. Inhibiting ANLN protein expression through siRNA could inhibit tumor cell growth and migration as well as induce tumor cell apoptosis (Suzuki et al., 2005). In pancreatic cancer cells, ANLN deficiency promotes miR218-5p expression, which results in apoptosis (Liu et al., 2015). Additionally, ANLN is a key modulator of inducing PI3K/Akt signaling, which suggests that ANLN is indirectly associated with apoptosis induction (Dai et al., 2019). In HNSC, ANLN is overexpressed in tumor tissues, with ANLN knockdown affecting tumor cell proliferation, migration, and invasion. Moreover, exosome-derived ANLN promotes cell differentiation and macrophage polarization via the PTEN/PI3K/Akt signaling pathway, which subsequently stimulates tumor growth. Taken together, ANLN is an independent prognostic factor for patients with HNSC (Guo et al., 2021). Our study used bioinformatics

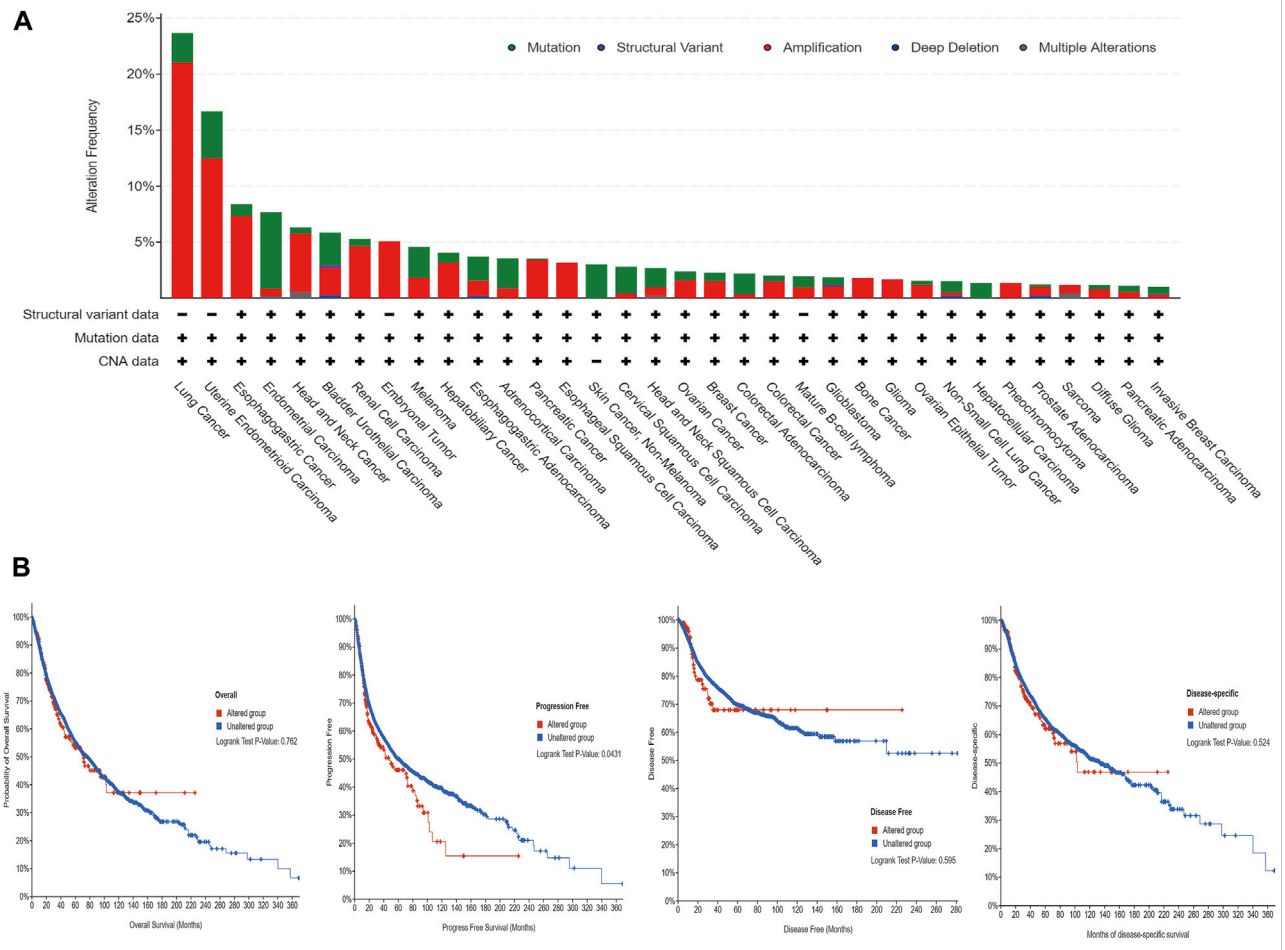


FIGURE 6
The mutational characteristics of ANLN in cancer. **(A)** The frequency of alterations in different types of mutations was examined using the cBioPortal database. **(B)** The effect of ANLN mutation status on overall survival (OS), disease-specific survival (DSS), disease-free survival (DFS), and progression-free survival (PFS) of cancer patients was investigated using the cBioPortal database.

methods to further demonstrate the correlation of ANLN with tumor progression and prognosis, which provides a theoretical basis for the use of ANLN as a potential novel target for biotherapy and a biomarker for tumor surveillance.

The TIME facilitates tumor growth and function by providing a relatively stable living environment for cancer cells to continuously proliferate without interference, which enhances the tumor malignancy potential. Additionally, it contains numerous immune and immunosuppressive cells, including M2-TAMs, CAFs, Tregs, and MDSCs, which can suppress the production, activation, and infiltration of cytotoxic T-cells as well as promote tumor immune evasion, tumor growth, metastasis, and treatment resistance (Liu and Cao, 2016; Monteran and Erez, 2019). The TIME and tumor immune escape are associated with cancer prognosis and treatment (Wu and Dai, 2017). Regarding tumor immune escape mechanisms, tumor immune cell infiltration causes T-cell dysfunction, which promotes tumor evasion of host immune system surveillance as well as facilitates tumor progression, invasion, metastasis, and treatment resistance (Yu et al., 2009; Liu et al., 2022). Another tumor immune escape mechanism is T-cell rejection. Here, CAFs, Tregs, M2-TAMs, and MDSCs in the tumor microenvironment, which act as immunosuppressive cells, block immune cell infiltration (Man et al., 2013; Komohara et al., 2016). Based on the TCGA database, we analyzed the relationship of ANLN with tumor immune cell infiltration. ANLN showed a significant positive correlation with tumor immune cell infiltration only in COAD, KIRC, LIHC, and PRAD, indicating that ANLN could enhance tumor immune infiltration through T-cell dysfunction in these cancer types. Furthermore, ANLN expression was strongly correlated with the levels of immunosuppressive cells in almost all cancer types. This suggests that T-cell rejection is the main mechanism through which ANLN regulates immune cell tumor escape, tumor promotion, and metastasis.

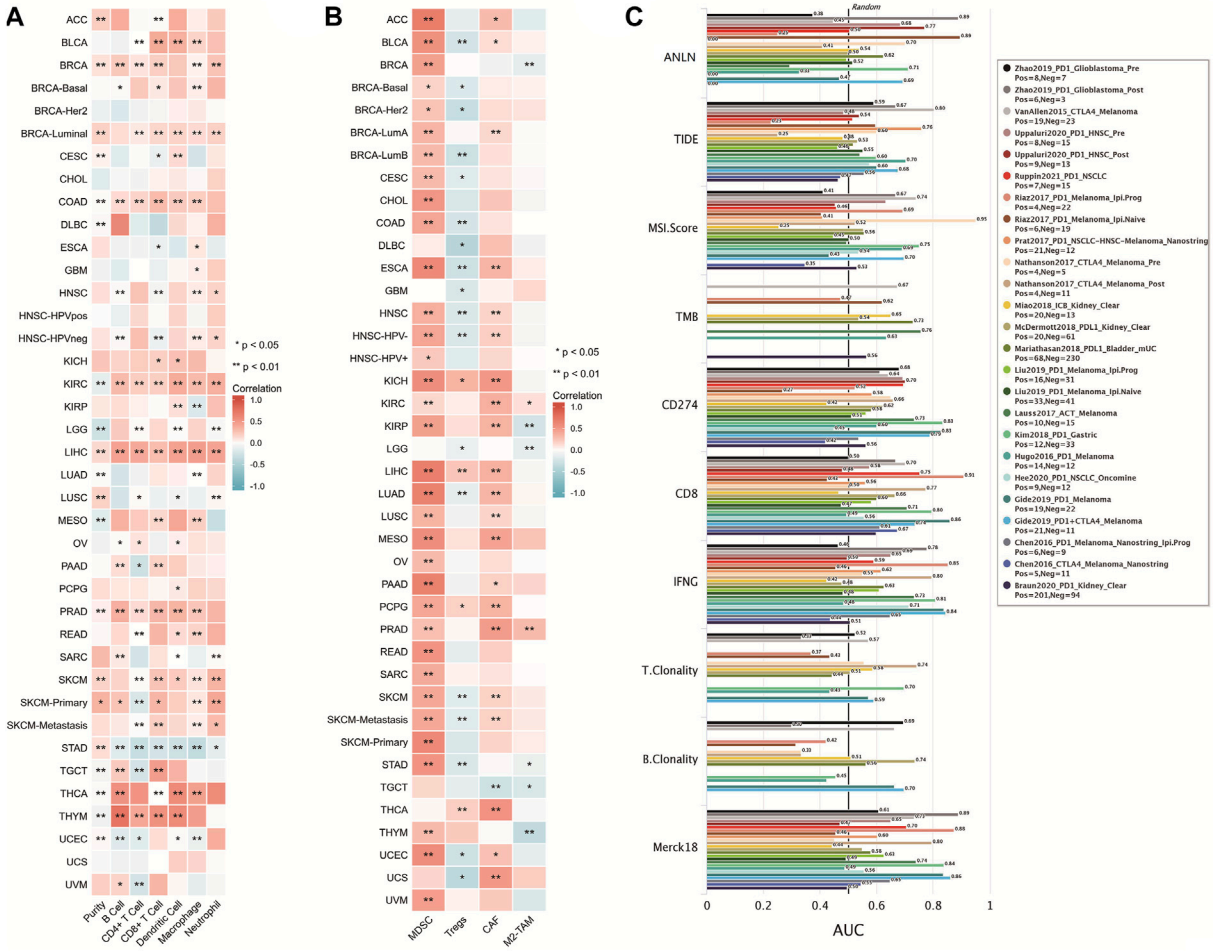


FIGURE 7
Correlation of ANLN with tumor immune cell infiltration. **(A,B)** Heat map showing a correlation between ANLN protein and infiltration of six types of immune cells and four types of immunosuppressive cells in multiple tumors; **(C)** Comparison of ANLN with other biomarkers for efficacy prediction of immunodetectable point-of-care inhibitors.

DNA methylation modifications affect a range of biological processes such as eukaryotic cell growth, differentiation, and transformation mechanisms by regulating gene expression. DNA methylation dysregulation causes multiple diseases (Wagner, 2022) and is considered a key event in cancer development and progression, including altered DNA methylation and dysregulation of gene expression in tumor-related genes (Wang J. et al., 2021). The expression of aberrantly methylated genes is associated with the degree of tumor differentiation, TNM stage, and poor prognosis (Skvortsova et al., 2019). Therefore, we investigated the relationship of ANLN methylation levels with prognosis in patients with cancer. We found that prognosis was correlated with the methylation level of four CpG sites; namely, cg00267323, cg04455682, cg04897631, and cg12284836.

Although genetic alterations in tumors are common, premalignant genetic changes are more likely to initiate and promote cancer development. The accumulation of genetic alterations drives the progression of normal cells through hyperplastic and dysplastic stages to invasive cancer and, ultimately, metastatic disease (Garnis et al., 2004). Therefore, analyzing genetic alterations in known oncogenes could further elucidate their role in cancer progression. In our study, lung cancer showed the highest frequency of ANLN alterations (≈ 23.68%); moreover, amplification and mutation were the most common genetic alterations. In addition, patients with altered ANLN genes showed worse PFS. Nonetheless, further experimental studies are warranted to reveal the mechanisms through which ANLN gene alterations are involved in tumor development.

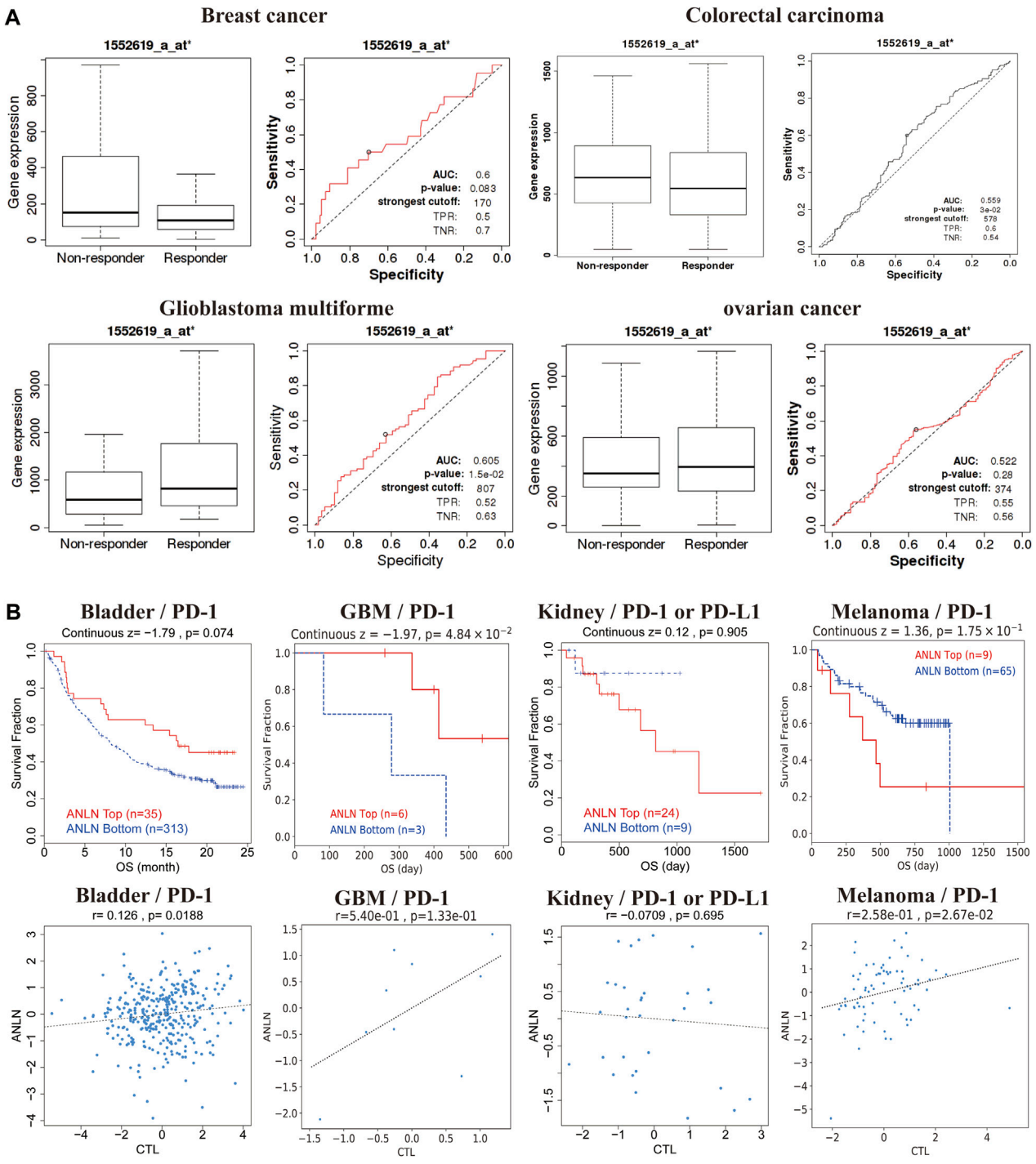


FIGURE 8
ANLN expression correlates with treatment response in multiple cancer types. **(A)** ROC plot of the association between ANLN expression and chemotherapy response in breast, brain, colorectal, and ovarian cancer cohorts. **(B)** Kaplan-Meier curves (upper panel) as a measure of immunotherapy response between cancer cohorts with high and low ANLN expression. The graph below shows the correlation between ANLN expression and cytotoxic T-cell levels.

There has been increasing research interest in immune replacement therapy regimens based on PD-1 and PD-L1 inhibitors. PD-1 and PD-L1-based immunotherapy can effectively treat various cancers and improve prognosis (Makuku et al., 2021). We found that high ANLN expression was associated with the clinical benefits of ICB therapy (PD-1) in patients with bladder

cancer and glioblastoma, while low ANLN expression was associated with the clinical benefits of ICB therapy (PD-L1 or PD-1) in patients with kidney cancer and melanoma. This suggested that ANLN is crucially involved in tumor-mediated immune escape. Further, regarding the relevance of ANLN in oncologic chemotherapy, in patients with breast and colorectal cancers, ANLN expression was negatively correlated with sensitivity to chemotherapeutic drugs. Patients with breast cancer showed high ANLN expression in tumor tissues; moreover, ANLN downregulation affected the tumor cell cycle and promoted apoptosis while inhibiting tumor cell proliferation, migration, and invasion. Therapy targeting ANLN could provide a similar therapeutic effect as chemotherapy on breast cancer (Wang Z. et al., 2021). However, the mechanism through which ANLN mediates resistance to chemotherapeutic agents should be further investigated.

Conclusion

In summary, our findings showed that ANLN is strongly associated with the prognosis, immune cell infiltration, gene mutations, and tumor treatment in patients with cancer. These findings provide further insight into the role of ANLN in tumorigenesis from an overall perspective. Furthermore, ANLN could be an onco-immunological biomarker and could serve as a hallmark for tumor screening, prognosis, individualized treatment design, and follow-up.

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

KJL conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft. LC and CQL conceived and designed the experiments, performed the experiments, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft. CFT, JH, and YMN analyzed the data, prepared figures and/or tables, and approved the final draft. YB and BDC conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.

Funding

This study was funded by the Newly recruited master's training project of the General Hospital of Ningxia Medical University (No. YKDZY2022013) and First-Class Discipline Construction Founded Proiect of NingXia Medical University and the School of Clinical Medicine (No. 2020016).

Acknowledgments

We thank PhD. Yongxue LV for the technological assistance. We would like to thank Editage (www.editage.cn) for English language editing.

Conflict of interest

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

Publisher's note

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/fgene.2022.922472/full#supplementary-material

SUPPLEMENTARY FIGURE S1
The density plots were used to visualize cut-off points for the dichotomy of methylation levels in patients.

References

Barrueto, L., Caminero, F., Cash, L., Makris, C., Lamichhane, P., and Deshmukh, R. R. (2020). Resistance to checkpoint inhibition in cancer immunotherapy. *Transl. Oncol.* 13 (3), 100738. doi:10.1016/j.tranon.2019.12.010

Blum, A., Wang, P., and Zenklusen, J. C. (2018). SnapShot: TCGA-analyzed tumors. *Cell* 173 (2), 530. doi:10.1016/j.cell.2018.03.059

Carithers, L. J., and Moore, H. M. (2015). The Genotype-Tissue expression (GTEx) project. *Biopreserv. Biobank.* 13 (5), 307–308. doi:10.1089/bio.2015.29031.hmm

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2 (5), 401–404. doi:10.1158/2159-8290.CD-12-0095

Chandrashekar, D. S., Bashel, B., Balasubramanya, S., Creighton, C. J., Ponce-Rodriguez, I., Chakravarthi, B., et al. (2017). UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* 19 (8), 649–658. doi:10.1016/j.neo.2017.05.002

Chen, F., Wendl, M. C., Wyczalkowski, M. A., Bailey, M. H., Li, Y., and Ding, L. (2021). Moving pan-cancer studies from basic research toward the clinic. *Nat. Cancer* 2 (9), 879–890. doi:10.1038/s43018-021-00250-4

Dai, X., Chen, X., Hakizimana, O., and Mei, Y. (2019). Genetic interactions between ANLN and KDR are prognostic for breast cancer survival. *Oncol. Rep.* 42 (6), 2255–2266. doi:10.3892/or.2019.7332

Deng, F., Xu, Z., Zhou, J., Zhang, R., and Gong, X. (2021). ANLN regulated by miR-30a-5p mediates malignant progression of lung adenocarcinoma. *Comput. Math. Methods Med.* 2021, 9549287. doi:10.1155/2021/9549287

Downs, B., Sherman, S., Cui, J., Kim, Y. C., Snyder, C., Christensen, M., et al. (2019). Common genetic variants contribute to incomplete penetrance: evidence from cancer-free BRCA1 mutation carriers. *Eur. J. Cancer* 107, 68–78. doi:10.1016/j.ejca.2018.10.022

Fekete, J. T., and Gyorffy, B. (2019). ROCplot.org: validating predictive biomarkers of chemotherapy/hormonal therapy/anti-HER2 therapy using transcriptomic data of 3, 104 breast cancer patients. *Int. J. Cancer* 145 (11), 3140–3151. doi:10.1002/ijc.32369

Fu, J., Li, K., Zhang, W., Wan, C., Zhang, J., Jiang, P., et al. (2020). Large-scale public data reuse to model immunotherapy response and resistance. *Genome Med.* 12 (1), 21. doi:10.1186/s13073-020-0721-z

Garnis, C., Buys, T. P., and Lam, W. L. (2004). Genetic alteration and gene expression modulation during cancer progression. *Mol. Cancer* 3, 9. doi:10.1186/1476-4598-3-9

Genova, C., Dellepiane, C., Carrega, P., Sommariva, S., Ferlazzo, G., Pronzato, P., et al. (2022). Therapeutic implications of tumor microenvironment in lung cancer: focus on immune checkpoint blockade. *Front. Immunol.* 12, 799455. doi:10.3389/fimmu.2021.799455

Guo, E., Mao, X., Wang, X., Guo, L., An, C., Zhang, C., et al. (2021). Alternatively spliced ANLN isoforms synergistically contribute to the progression of head and neck squamous cell carcinoma. *Cell Death Dis.* 12 (8), 764. doi:10.1038/s41419-021-04063-2

Hargadon, K. M., Johnson, C. E., and Williams, C. J. (2018). Immune checkpoint blockade therapy for cancer: an overview of FDA-approved immune checkpoint inhibitors. *Int. Immunopharmacol.* 62, 29–39. doi:10.1016/j.intimp.2018.06.001

Idichi, T., Seki, N., Kurahara, H., Yonemori, K., Osako, Y., Arai, T., et al. (2017). Regulation of actin-binding protein ANLN by antitumor miR-217 inhibits cancer cell aggressiveness in pancreatic ductal adenocarcinoma. *Oncotarget* 8 (32), 53180–53193. doi:10.18632/oncotarget.18261

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24 (10), 1550–1558. doi:10.1038/s41591-018-0136-1

Karlsson, M., Zhang, C., Mear, L., Zhong, W., Digre, A., Katona, B., et al. (2021). A single-cell type transcriptomics map of human tissues. *Sci. Adv.* 7 (31), eabh2169. doi:10.1126/sciadv.abh2169

Komohara, Y., Fujiwara, Y., Ohnishi, K., and Takeya, M. (2016). Tumor-associated macrophages: potential therapeutic targets for anti-cancer therapy. *Adv. Drug Deliv. Rev.* 99 (Pt B), 180–185. doi:10.1016/j.addr.2015.11.009

Lanczky, A., and Gyorffy, B. (2021). Web-Based survival analysis tool tailored for medical research (KMplot): development and implementation. *J. Med. Internet Res.* 23 (7), e27633. doi:10.2196/27633

Li, L., Huang, K., Zhao, H., Chen, B., Ye, Q., and Yue, J. (2020c). CDK1-PLK1/SGOL2/ANLN pathway mediating abnormal cell division in cell cycle may be a critical process in hepatocellular carcinoma. *Cell Cycle* 19 (10), 1236–1252. doi:10.1080/15384101.2020.1749471

Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., et al. (2020b). TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.* 48 (W1), W509–W514. doi:10.1093/nar/gkaa407

Li, Y., Liu, J., Gao, L., Liu, Y., Meng, F., Li, X., et al. (2020a). Targeting the tumor microenvironment to overcome immune checkpoint blockade therapy resistance. *Immunol. Lett.* 220, 88–96. doi:10.1016/j.imlet.2019.03.006

Lian, Y., Huang, Y., Wang, J., Deng, M., Xia, T., Zeng, M., et al. (2018). Anillin is required for tumor growth and regulated by miR-15a/miR-16-1 in HBV-related hepatocellular carcinoma. *Aging* 10 (8), 1884–1901. doi:10.18632/aging.101510

Liang, P. I., Chen, W. T., Li, C. F., Li, C. C., Li, W. M., Huang, C. N., et al. (2015). Subcellular localisation of anillin is associated with different survival outcomes in upper urinary tract urothelial carcinoma. *J. Clin. Pathol.* 68 (12), 1026–1032. doi:10.1136/jclinpath-2015-202958

Liu, X., Hoft, D. F., and Peng, G. (2022). Tumor microenvironment metabolites directing T cell differentiation and function. *Trends Immunol.* 43 (2), 132–147. doi:10.1016/j.it.2021.12.004

Liu, Y., and Cao, X. (2016). Immunosuppressive cells in tumor immune escape and metastasis. *J. Mol. Med.* 94 (5), 509–522. doi:10.1007/s00109-015-1376-x

Liu, Z., Xu, Y., Long, J., Guo, K., Ge, C., and Du, R. (2015). MicroRNA-218 suppresses the proliferation, invasion and promotes apoptosis of pancreatic cancer cells by targeting HMGB1. *Chin. J. Cancer Res.* 27 (3), 247–257. doi:10.3978/j.issn.1000-9604.2015.04.07

Makuku, R., Khalili, N., Razi, S., Keshavarz-Fathi, M., and Rezaei, N. (2021). Current and future perspectives of PD-1/PDL-1 blockade in cancer immunotherapy. *J. Immunol. Res.* 2021, 6661406. doi:10.1155/2021/6661406

Man, Y. G., Stojadinovic, A., Mason, J., Avital, I., Bilchik, A., Bruecher, B., et al. (2013). Tumor-infiltrating immune cells promoting tumor invasion and metastasis: existing theories. *J. Cancer* 4 (1), 84–95. doi:10.7150/jca.5482

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

Modhukur, V., Iljasenko, T., Metsalu, T., Lokk, K., Laisk-Podar, T., and Vilo, J. (2018). MethSurv: a web tool to perform multivariable survival analysis using DNA methylation data. *Epigenomics* 10 (3), 277–288. doi:10.2217/epi-2017-0118

Monteran, L., and Erez, N. (2019). The dark side of fibroblasts: cancer-associated fibroblasts as mediators of immunosuppression in the tumor microenvironment. *Front. Immunol.* 10, 1835. doi:10.3389/fimmu.2019.01835

Morad, G., Helmink, B. A., Sharma, P., and Wargo, J. A. (2021). Hallmarks of response, resistance, and toxicity to immune checkpoint blockade. *Cell* 184 (21), 5309–5337. doi:10.1016/j.cell.2021.09.020

Naydenov, N. G., Koblinski, J. E., and Ivanov, A. I. (2021). Anillin is an emerging regulator of tumorigenesis, acting as a cortical cytoskeletal scaffold and a nuclear modulator of cancer cell differentiation. *Cell. Mol. Life Sci.* 78 (2), 621–633. doi:10.1007/s00018-020-03605-9

Omasits, U., Ahrens, C. H., Muller, S., and Wollscheid, B. (2014). Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* 30 (6), 884–886. doi:10.1093/bioinformatics/btt607

Pandi, N. S., Manimuthu, M., Harunipriya, P., Murugesan, M., Asha, G. V., and Rajendran, S. (2014). *In silico* analysis of expression pattern of a Wnt/beta-catenin responsive gene ANLN in gastric cancer. *Gene* 545 (1), 23–29. doi:10.1016/j.gene.2014.05.013

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer statistics, 2021. *CA. Cancer J. Clin.* 71 (1), 7–33. doi:10.3322/caac.21654

Skvortsova, K., Stirzaker, C., and Taberlay, P. (2019). The DNA methylation landscape in cancer. *Essays Biochem.* 63 (6), 797–811. doi:10.1042/EBC20190037

Summerfield, A., McNeilly, F., Walker, I., Allan, G., Knoetig, S. M., and McCullough, K. C. (2001). Depletion of CD4(+) and CD8(high+) T-cells before the onset of viraemia during classical swine fever. *Vet. Immunol. Immunopathol.* 78 (1), 3–19. doi:10.1016/s0165-2427(00)00248-8

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Suzuki, C., Daigo, Y., Ishikawa, N., Kato, T., Hayama, S., Ito, T., et al. (2005). ANLN plays a critical role in human lung carcinogenesis through the activation of RHOA and by involvement in the phosphoinositide 3-kinase/AKT pathway. *Cancer Res.* 65 (24), 11314–11325. doi:10.1158/0008-5472.CAN-05-1507

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49 (D1), D605–D612. doi:10.1093/nar/gkaa1074

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 45 (W1), W98–W102. doi:10.1093/nar/gkx247

Thul, P. J., Akesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait, B. H., et al. (2017). A subcellular map of the human proteome. *Science* 356 (6340), eaal3321. doi:10.1126/science.aal3321

Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. *Science* 347 (6220), 1260419. doi:10.1126/science.1260419

Uhlen, M., Zhang, C., Lee, S., Sjostedt, E., Fagerberg, L., Bidkhori, G., et al. (2017). A pathology atlas of the human cancer transcriptome. *Science* 357 (6352), eaan2507. doi:10.1126/science.aan2507

Wagner, W. (2022). How to translate DNA methylation biomarkers into clinical practice. *Front. Cell Dev. Biol.* 10, 854797. doi:10.3389/fcell.2022.854797

Wang, A., Dai, H., Gong, Y., Zhang, C., Shu, J., Luo, Y., et al. (2019). ANLN-induced EZH2 upregulation promotes pancreatic cancer progression by mediating miR-218-5p/LASP1 signaling axis. *J. Exp. Clin. Cancer Res.* 38 (1), 347. doi:10.1186/s13046-019-1340-7

Wang, G., Shen, W., Cui, L., Chen, W., Hu, X., and Fu, J. (2016). Overexpression of Anillin (ANLN) is correlated with colorectal cancer progression and poor prognosis. *Cancer Biomark.* 16 (3), 459–465. doi:10.3233/CBM-160585

Wang, J., Yang, J., Li, D., and Li, J. (2021). Technologies for targeting DNA methylation modifications: basic mechanism and potential application in cancer. *Biochim. Biophys. Acta. Rev. Cancer* 1875 (1), 188454. doi:10.1016/j.bbcan.2020.188454

Wang, Z., Hu, S., Li, X., Liu, Z., Han, D., Wang, Y., et al. (2021). MiR-16-5p suppresses breast cancer proliferation by targeting ANLN. *BMC Cancer* 21 (1), 1188. doi:10.1186/s12885-021-08914-1

Wei, S., Teng, S., Yao, J., Gao, W., Zang, J., Wang, G., et al. (2020). Develop a circular RNA–related regulatory network associated with prognosis of gastric cancer. *Cancer Med.* 9 (22), 8589–8599. doi:10.1002/cam4.3035

Wu, T., and Dai, Y. (2017). Tumor microenvironment and therapeutic response. *Cancer Lett.* 387, 61–68. doi:10.1016/j.canlet.2016.01.043

Xiao, J. X., Xu, W., Fei, X., Hao, F., Wang, N., Chen, Y., et al. (2020). Anillin facilitates cell proliferation and induces tumor growth of hepatocellular carcinoma via miR-138/SOX4 axis regulation. *Transl. Oncol.* 13 (10), 100815. doi:10.1016/j.tranon.2020.100815

Yu, G. P., Chiang, D., Song, S. J., Hoyte, E. G., Huang, J., Vanishsarn, C., et al. (2009). Regulatory T cell dysfunction in subjects with common variable immunodeficiency complicated by autoimmune disease. *Clin. Immunol.* 131 (2), 240–253. doi:10.1016/j.clim.2008.12.006

Zeng, S., Yu, X., Ma, C., Song, R., Zhang, Z., Zi, X., et al. (2017). Transcriptome sequencing identifies ANLN as a promising prognostic biomarker in bladder urothelial carcinoma. *Sci. Rep.* 7 (1), 3151. doi:10.1038/s41598-017-02990-9

Zhang, L. H., Wang, D., Li, Z., Wang, G., Chen, D. B., Cheng, Q., et al. (2021). Overexpression of anillin is related to poor prognosis in patients with hepatocellular carcinoma. *Hepatobiliary Pancreat. Dis. Int.* 20 (4), 337–344. doi:10.1016/j.hbpd.2020.08.007

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 08 August 2022
DOI 10.3389/fgene.2022.926796

Check for updates

OPEN ACCESS

EDITED BY
Claudia Cava,
Institute of Bioimaging and Molecular Ph.ysiology, (IBFM-CNR), Italy

REVIEWED BY
Evelien Schaafsma,
Dartmouth College, United States
Bo Liu,
Sichuan University, China

*CORRESPONDENCE
Renkuan Tang,
100456@cqmu.edu.cn

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

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 23 April 2022
ACCEPTED 05 July 2022
PUBLISHED 08 August 2022

CITATION
Zheng Y, Yuan D, Zhang F and Tang R (2022), A systematic pan-cancer analysis of the gasdermin (GSDM) family of genes and their correlation with prognosis, the tumor microenvironment, and drug sensitivity. *Front. Genet.* 13:926796. doi: 10.3389/fgene.2022.926796

COPYRIGHT
© 2022 Zheng, Yuan, Zhang and Tang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A systematic pan-cancer analysis of the gasdermin (GSDM) family of genes and their correlation with prognosis, the tumor microenvironment, and drug sensitivity

Yufu Zheng†, Didi Yuan†, Fei Zhang and Renkuan Tang*

Department of Forensic Medicine, College of Basic Medicine, Chongqing Medical University, Chongqing, China

Background: Pyroptosis is a programmed cell death process mediated by the gasdermin (GSDM) protein. However, limited research has been conducted to comprehensively analyze the contribution of the GSDM family in a pan-cancer setting.

Methods: We systematically evaluated the gene expression, genetic variations, and prognostic values of the GSDM family members. Furthermore, we investigated the association between the expression of GSDM genes and immune subtypes, the tumor microenvironment (TME), the stemness index, and cancer drug sensitivities by means of a pan-cancer analysis.

Results: GSDM genes were highly upregulated in most of the tested cancers. Low-level mutation frequencies within GSDM genes were common across the examined types of cancer, and their expression levels were associated with prognosis, clinical characteristics, TME features, and stemness scores in several cancer types, particularly those of the urinary system. Importantly, we found that the expressions of *GSDMB*, *GSDMC*, and *GSDMD* were higher in kidney carcinomas, and specifically kidney renal clear cell carcinoma (KIRC); which adversely impacted the patient outcome. We showed that *GSDMD* was potentially the most useful biomarker for KIRC. The drug sensitivity analysis demonstrated that the expressions of GSDM genes were correlated with the sensitivity of tumor cells to treatment with chemotherapy drugs nelarabine, fluphenazine, dexrazoxane, bortezomib, midostaurin, and vincristine.

Conclusion: GSDM genes were associated with tumor behaviors and may participate in carcinogenesis. The results of this study may therefore provide new directions for further investigating the role of GSDM genes as therapeutic targets in a pan-cancer setting.

KEYWORDS

pyroptosis, GSDM family, tumor microenvironment, drug sensitivity, prognosis

Introduction

Pyroptosis is a type of programmed cell death primarily mediated by the family of pore-forming protein gasdermins (GSDMs) (Shi et al., 2017). When cleaved by caspase proteins, the GSDM family participates in pyroptotic pore formation in the cell membrane. These cause extensive water influx from the extracellular space, leading to cell swelling and rupture (Tsuchiya, 2021). Strong evidence has demonstrated that the pyroptosis process is highly involved in the pathogenesis and progression of cancer by inducing GSDMD-dependent caspase-1 inflammasome pathways and other GSDM-dependent non-inflammasome signaling cascades (Johnson et al., 2018; Pizato et al., 2018; Wang et al., 2019; Yu et al., 2019). Therefore, investigations of GSDM genes may provide a more comprehensive overview of carcinogenesis and shed light on novel therapeutic strategies.

The majority of GSDM family members share an N-terminal pore-forming effector domain and a C-terminal repressor domain (Ding et al., 2016; Kuang et al., 2017). Six types of GSDMs paralogous genes, namely *GSDMA*, *GSDMB*, *GSDMC*, *GSDMD*, and *GSDME* [also termed deafness, autosomal dominant 53 (*DFNA5*)], and pejvakin [*PJVK*; also termed deafness, autosomal recessive 59 9*DFNB59*)], have been found in humans (Liu et al., 2021). Specifically, *GSDMA* is primarily localized to gastric and skin epithelia; *GSDMB* is found in the esophagus, liver, and colon; *GSDMC* is located in keratinocytes, the trachea, and the spleen; *GSDMD* is widely distributed in various immune cells, the placenta, the esophagus, and the gastrointestinal tract epithelium; *GSDME* is primarily found in the cochlea and the placenta; while, *PJVK* is found in the heart and liver. Recent studies have revealed that the GSDM family of genes plays a dual role in tumor pathogenesis and progression (Xia et al., 2019). It has been reported that while *GSDMA* is silenced in gastric cancer, GSDMB is overexpressed in several types of cancers (e.g., breast and gastric cancers) (Hergueta-Redondo et al., 2016). Moreover, studies have shown that the expression levels of GSDMC and GSDME proteins were significantly upregulated in colorectal cancer tissues. *GSDMC* is regulated by transforming growth factor-beta (TGF-β) signaling and is associated with tumor cell proliferation (Miguchi et al., 2016). *GSDME* participates in tumorigenesis and colorectal tumor cell proliferation, specifically through its association with the extracellular signal-regulated kinase 1/2 (ERK1/2) pathway. *GSDME* has also been shown to exert potentially tumor-suppressive effects in models of intestinal cancer (Croes et al., 2019; Tan et al., 2020). GSDMD was found to be highly downregulated in human colorectal tumor samples and serves as an important prognostic molecule in tumor therapy (Wu et al., 2020). However, studies investigating the role of GSDM genes in various tumors are limited, and further comprehensive analyses are warranted to clarify their molecular characteristics in a pan-cancer setting.

In the present study, we used data from The Cancer Genome Atlas (TCGA) project and the cBioPortal database, to comprehensively analyze the landscape of the GSDM gene expression status and evaluate the genetic alteration of this set of genes in various cancer types. In addition, we determined the prognostic value of the GSDM gene family by examining clinical traits in a pan-cancer setting using the Kaplan–Meier plotter and data from the University of California Santa Cruz (UCSC) database. Subsequently, we investigated the potential correlation between the expression of GSDM genes and tumor microenvironment (TME) characteristics, as well as the stemness index. An investigation of the association between GSDM genes and numerous dynamic immune modulators, tumor mutational burden (TMB), and microsatellite instability (MSI) was also conducted. Finally, correlations between GSDM gene expression and cancer drug sensitivity were also explored in various aspects of cancer using public resources.

Materials and methods

Data resources

All pan-cancer data were obtained from the UCSC Xena project (https://xena.ucsc.edu/), which contains 11,768 samples and 33 tumor types originating from the TCGA database. We extracted the mRNA data of six pyroptosis-associated genes provided by the TCGA Pan-Cancer Atlas. The pan-cancer immune subtype data were also retrieved from the UCSC database; they included six immune groups: Immune C1 (Wound Healing); Immune C2 [interferon gamma (IFN-γ) Dominant]; Immune C3 (Inflammatory); Immune C4 (Lymphocyte Depleted); Immune C5 (Immunologically Quiet); and Immune C6 (TGF-β Dominant). Survival data in relation to the expression of GSDM genes in the pan-cancer setting were obtained from TCGA and Kaplan–Meier Plotter databases. Data on the DNA methylation-based stemness score and RNA-based stemness score were also downloaded from the UCSC Pan-Cancer project. The MSI data for each cancer type were obtained from a published study (Bonneville et al., 2017). Furthermore, we obtained the drug sensitivity data from the CellMiner™ database to analyze the correlation between GSDM gene expression and drug sensitivity.

Data processing and statistical analysis

After merging the GSDM gene expression data to create the pan-cancer dataset, batch effects between different cancers were corrected using the "removeBatchEffect" function of the "limma" package (https://doi.org/doi:10.18129/B9.bioc.limma) of R software, using empirical Bayes algorithms. We used the Wilcoxon rank-sum test to investigate the differences in the

expression of GSDM genes between malignant tissues and matched normal samples for each cancer type. Genomic alterations were evaluated using the cBioPortal website. The Kruskal–Wallis test was also used to compare the differences in gene expression patterns among the six immune subtypes (C1–C6). Next, we then analyzed the expression patterns of GSDM genes and investigated their correlation with different disease stages in the pan-cancer setting using the Wilcoxon rank-sum test. The prognostic value of GSDM expression in the pan-cancer setting was analyzed using TCGA data. Kaplan-Meier survival curves were constructed to assess whether prognosis was associated with high or low GSDM expression levels, and one-way analysis of variance (ANOVA) was carried out to measure statistical significance. In addition, we further performed a Cox regression analysis to investigate the association between GSDM gene expression and the pan-cancer prognosis data. Finally, we validated the correlation between expression levels and pan-cancer prognosis using data from the Kaplan-Meier plotter database. Based on the TCGA database, we used the receiver operating characteristic (ROC) curve to assess the specificity and sensitivity of GSDM gene expression in cancer, and the areas under the curve (AUCs) were quantified using the "pROC" package of R software. The results were further validated using data obtained from the TARGET (therapeutically applicable research to generate effective treatments) database.

The TME is involved in the occurrence and migration of cancer (Quail and Joyce, 2013). We conducted an analysis of the pan-cancer TME and stemness using the CIBERSORT algorithm and the Spearman's method, respectively. The immune score and tumor purity analyses were performed using the CIBERSORT algorithm to evaluate the presence of immune infiltrating cells and the tumor purity, respectively, in cancer tissues. Relationship analysis between GSDM family gene expression and the TME and stemness scores was also undertaken using the cor.Test command of the Spearman's method. We further explored the correlation between GSDM expression and the TMB, and MSI using the Spearman's correlation test in the selected cancer types. Finally, a drug sensitivity analysis was performed using the "limma" package of R software.

All data analyses and processing were conducted using R software version 4.1.0 (https://www.r-project.org/). The "fdr" algorithm of the "p.adjust" R function has been used to adjust the p-value. An adjusted cutoff value of $p < 0.05$ denotes statistically significant differences.

Results

Landscape of GSDM gene expression in The Cancer Genome Atlas pan-cancer

Using the fragments per kilobase of exon model per million reads mapped (FPKM) value of the RNA sequence data from the TCGA project, we explored the mRNA expression levels of GSDM genes in the pan-cancer analysis. According to the results, *GSDMD* exhibited the highest expression levels, followed by *GSDMB*, *GSDME*, and *PJVK*. In contrast, *GSDMA* and *GSDMC* had relatively low expression levels in the tumor specimens (Figure 1A). Subsequently, we utilized the Wilcoxon rank-sum test to compare the mRNA expression levels of GSDM genes in distinct cancer types using both malignant tumor specimen and adjacent normal samples with an adjusted cutoff value of $p < 0.05$ (Figures 1B–H) ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$). We then observed the differences in GSDM mRNA expression levels between tumor/normal sample pairs on a heatmap showing relative fold change in GSDM gene expression (Figure 1B) and boxplots of gene differential expression (Figures 1C–H), for each cancer type. We found that most tumors analyzed and exhibited a certain level of GSDMs expression (Figure 1B); however, the expression of *GSDMD*, *GSDMB*, and *GSDME* in pan-cancers was markedly higher than that of other genes. Interestingly, *GSDMA* and *GSDMC* presented relatively higher gene expression levels in all tumor types compared with matched normal tissues (Figures 1C–E). *GSDMB* expression was significantly lower in tumor specimens of breast invasive carcinoma (BRCA), colonic adenocarcinoma (COAD), and kidney chromophobe (KICH), compared to healthy tissues (Figure 1D). *GSDMD* expression was lower in several cancers, including KICH, lung squamous cell carcinoma (LUSC), and prostate cancer (PRAD), compared to normal tissues (Figure 1F). Relative to normal tissues, *GSDME* expression was higher in most cancer types; however, it was lower in bladder urothelial carcinoma (BLCA), BRCA, COAD, KICH, kidney renal clear cell carcinoma (KIRC), PRAD, thyroid carcinoma (THCA), and uterine corpus endometrial carcinoma (Figure 1G). *PJVK* exhibited lower levels of gene expression in various tumor tissues, except for in cholangiocarcinoma (CHOL), COAD, head and neck squamous cell carcinoma (HNSC), KIRC, kidney renal papillary cell carcinoma (KIRP), liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), LUSC, and stomach adenocarcinoma (STAD) (Figure 1H). In summary, GSDM genes showed highly variable gene expression patterns in the pan-cancer setting, indicating that alterations in the expression of GSDM genes may play a crucial role in cancer occurrence and heterogeneity.

Genetic variation of GSDM genes in The Cancer Genome Atlas pan-cancer samples

We next investigated genetic alterations in GSDM genes using TCGA pan-cancer data provided via the cBioPortal database. We found that the overall mutation frequency in GSDM genes was relatively low in pan-cancer tissues, and the most common type of mutation was missense mutations. However, the analysis of copy number variation (CNV) within GSDM genes showed extensive amplification. The genomic landscapes showed that *GSDMC* (8%) and *GSDMD* (6%) had

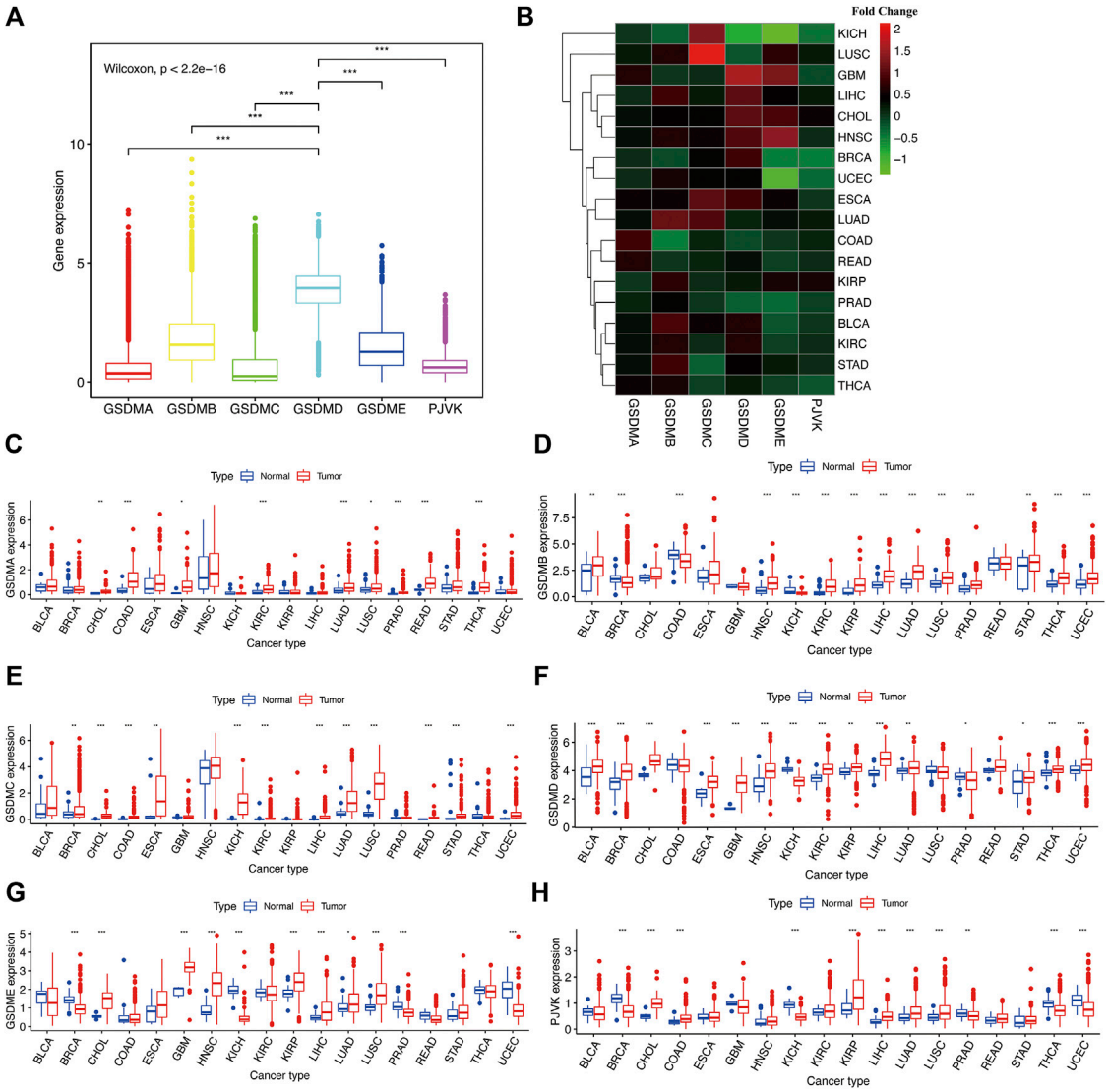


FIGURE 1
Comparison of the GSDM gene expression landscape in different cancers and corresponding normal tissues. **(A)** Differences in GSDM gene expression in the pan-cancer setting **(B)**. GSDM family gene expression levels in different cancer types (TCGA data); red, high gene expression; green, low gene expression. **(C–H)** Differences in GSDM mRNA levels between normal and cancer tissues. An adjusted cutoff value of $p < 0.05$ denotes statistically significant differences. The asterisks represent statistical significance with adjusted p value (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). GSDM, gasdermin; TCGA, The Cancer Genome Atlas.

the highest mutational rates in all cancer types, followed by *GSDMA* (3%), *GSDMB* (3%), *GSDME* (2%), and *PJVK* (1.4%) (Figure 2A). Additionally, we focused on the TMB and MSI to examine the potential correlation between the expression of each GSDM gene and various TME features using the Spearman's correlation test with an adjusted cutoff value of $p < 0.05$ (Figure 2B). Notably, the GSDM gene expression levels were closely related to the TMB in urologic neoplasms (e.g., KIRC, KIRP, and PRAD). A similar significant association between GSDM gene expression and MSI was observed for all cancers examined.

The association between GSDM gene expression and immune subtypes and clinical traits in pan-cancer

To investigate the potential effect of GSDM genes in a pan-cancer setting, we explored the correlation between their mRNA levels and various immune subtypes using TCGA data (Figure 3A). The results demonstrated that the expression levels of GSDM genes differed significantly among the pan-cancer immune subtypes. Genes *GSDMA–D* were more highly expressed in C1 and C2, but



FIGURE 2
Genetic changes and TME features associated with GSDM genes in the pan-cancer setting. **(A)** Mutation spectrum of the GSDM genes in the pan-cancer setting **(B)** Correlation of GSDM gene expression with TMB and MSI in various cancer types (TCGA data). An adjusted cutoff value of $p < 0.05$ denotes statistically significant differences ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$). TME, tumor microenvironment; TMB, tumor mutational burden; MSI, microsatellite instability. GSDM, gasdermin.

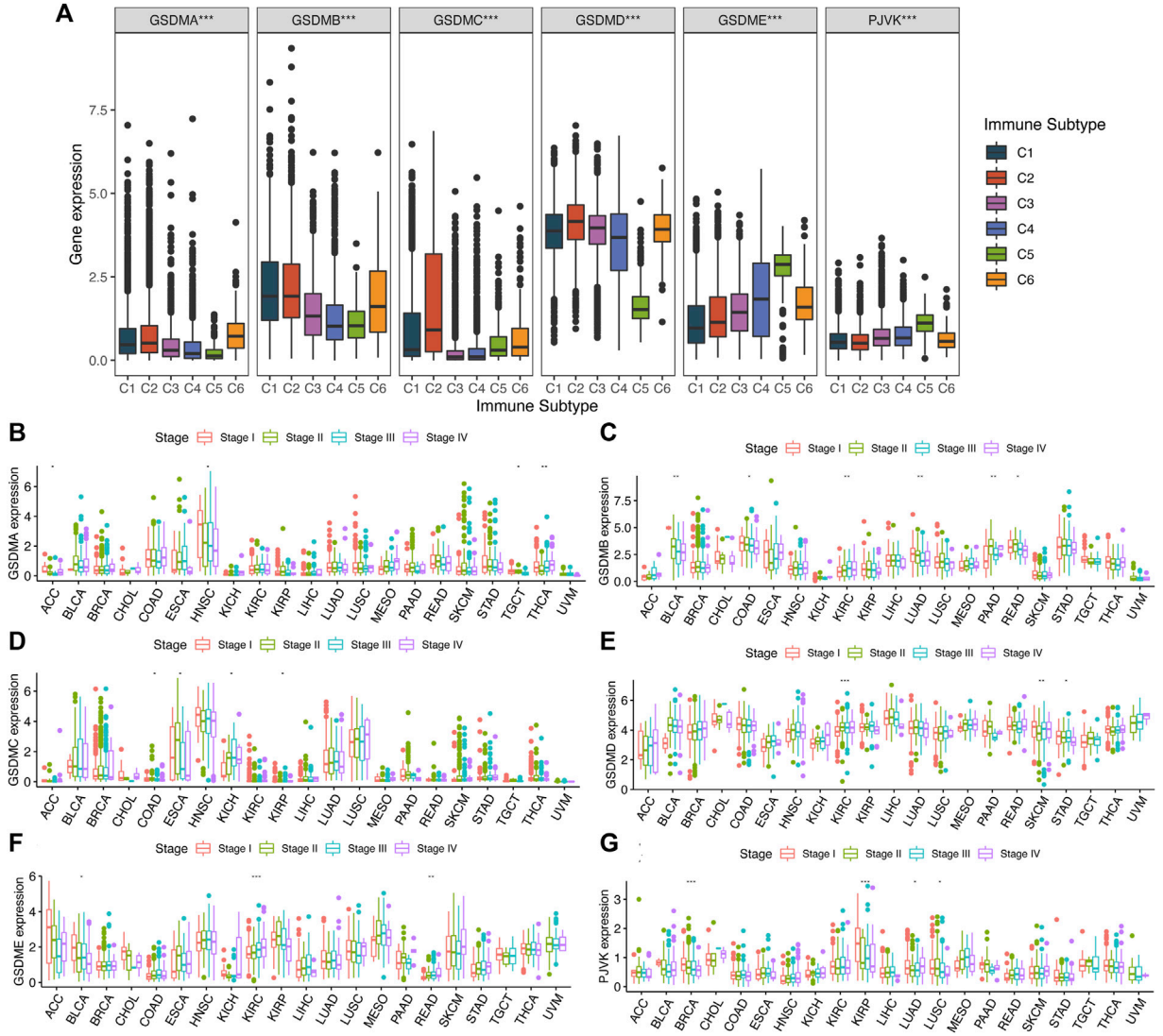


FIGURE 3
Association between GSDM gene expression and different immune subtypes and clinical traits in the pan-cancer setting. **(A)** Differences in GSDM gene expressions in six pan-cancer immune subtypes: C1 (Wound Healing), C2 (IFN-γ Dominant), C3 (Inflammatory), C4 (Lymphocyte Depleted), C5 (Immunologically Quiet), and C6 (TGF-β Dominant). **(B–G)** Correlation between the tumor stage and GSDM gene expression. An adjusted cutoff value of $p < 0.05$ denotes statistically significant differences (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). GSDM, gasdermin; IFN-γ, interferon gamma; TGF-β, transforming growth factor-beta.

less so in C5. The expression of both *GSDMD* and *PJVK* was higher in C4 and C5 (Figure 3A). Next, we performed a clinical correlation analysis of the various cancer types. Overall, the expression of GSDM genes was, as expected, associated with different pathological stages. *GSDMA* showed significant differences in mRNA expression between different clinical stages of four cancers, namely adrenocortical carcinoma (ACC), HNSC, testicular germ cell tumors (TGCT), and THCA (Figure 3B). Relatively high mRNA expression levels of GSDMA were associated with THCA progression; while lower levels were observed in the advanced stages of ACC, HNSC, or TGCT. *GSDMB* mRNA levels differed significantly between the clinical stages of six types of cancer, namely BLCA, COAD, KIRC, LUAD, pancreatic adenocarcinoma (PAAD), and rectum adenocarcinoma (READ) (Figure 3C). Lower *GSDMB* expression tended to be noted in the more advanced stages of BLCA, COAD, LUAD, and READ, while higher *GSDMB* expression was observed in progressive KIRC and PAAD. *GSDMC* mRNA levels varied significantly between the clinical stages of four cancers, including COAD, esophageal carcinoma, KICH, and KIRP (Figure 3D).

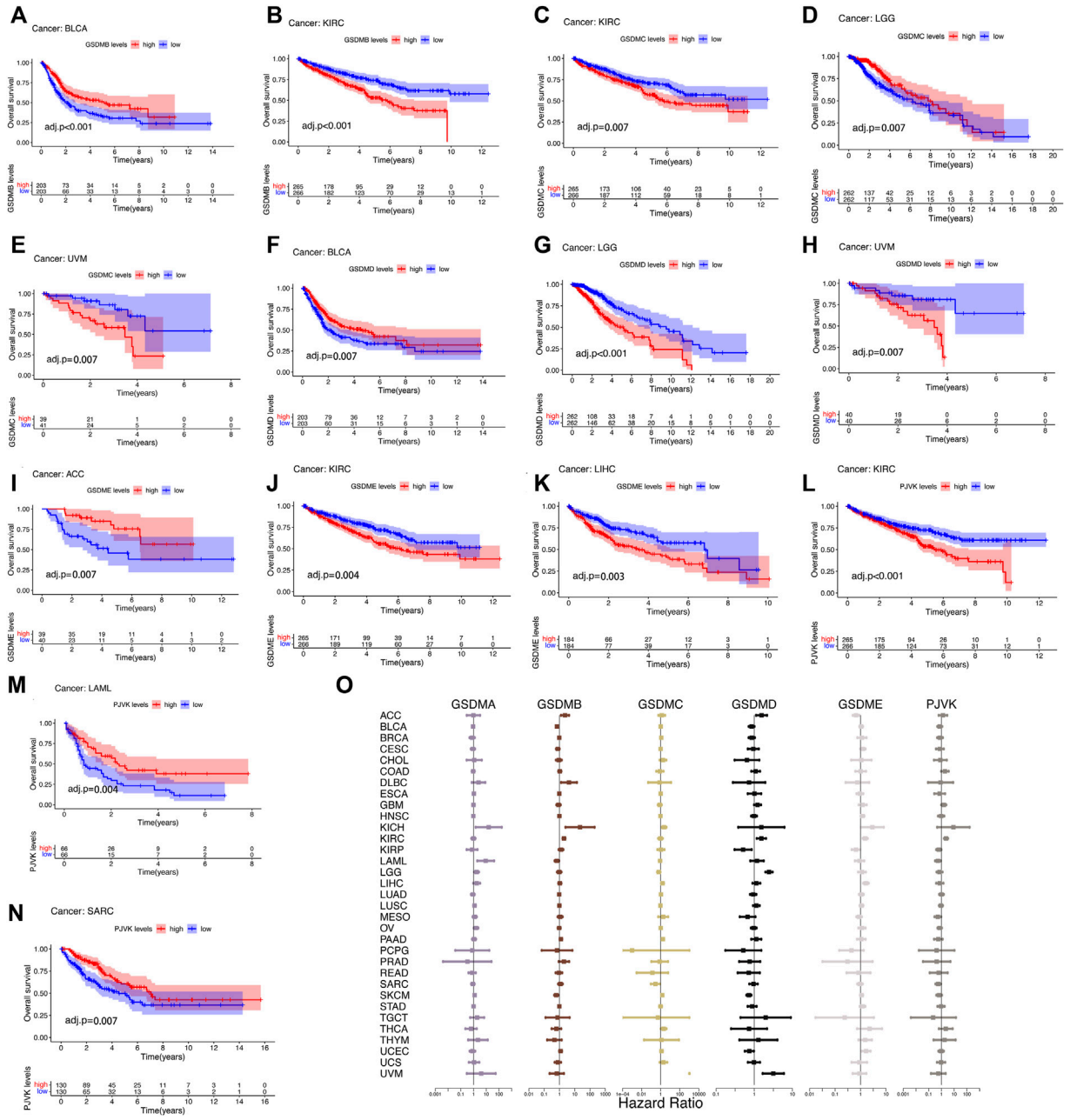


FIGURE 4
Survival analysis in relation to GSDM gene expression in the pan-cancer setting (TCGA data). **(A–N)** Kaplan-Meier survival curves for the high and low GSDM gene expression groups in the different cancers (TCGA data). **(O)** Cox regression analysis of GSDM gene expression and tumor survival in different cancer types (TCGA data). A hazard ratio <1 and >1 denotes low and high risk, respectively. GSDM, gasdermin; TCGA, The Cancer Genome Atlas.

Higher *GSDMC* expression tended to occur in more advanced stages of KICH. *GSDMD* mRNA levels differed significantly between the clinical stages of KIRC, skin cutaneous melanoma, and STAD (Figure 3E), but were higher in the more advanced stages of KIRC. Furthermore, *GSDME* mRNA levels also varied significantly between different clinical stages of BLCA, KIRC, and READ (Figure 3F), but were consistently higher in the more advanced stages of KIRC and READ. Finally, *PJVK* showed significant differences in mRNA expression at different clinical stages of BRCA, KIRP, LUAD, and LUSC (Figure 3G), while lower *PJVK* expression levels were

TABLE 1 Cox regression analysis of the prognostic values of GSDM genes in the pan-cancer setting.

Gene	Cancer type	HR	95% Confidence interval		*p*-value	Adjusted
GSDMA	KICH	15.58	1.47	165	0.022	0.03
	LAML	9.04	1.98	41.2	0.004	0.012
	LGG	1.74	1.14	2.66	0.011	0.018
	OV	1.58	1.06	2.37	0.026	0.031
	SKCM	1.22	1.06	1.41	0.006	0.014
GSDMB	ACC	2.31	1.26	4.25	0.007	0.014
	BLCA	0.68	0.6	0.78	3.34E−08	4.45E−07
	DLBC	4.21	1.29	13.8	0.018	0.027
	KICH	22.22	2.49	199	0.006	0.014
	KIRC	1.98	1.61	2.42	6.55E−11	1.31E−09
	LAML	0.67	0.48	0.94	0.019	0.027
	PAAD	1.24	1.04	1.48	0.018	0.027
	SKCM	0.62	0.45	0.84	0.002	0.007
GSDMC	KICH	2.25	1.27	3.99	0.005	0.014
	LGG	0.63	0.43	0.93	0.02	0.027
	LIHC	1.56	1.12	2.16	0.008	0.016
	PAAD	1.72	1.32	2.23	6.32E−05	4.21E−04
	SARC	0.28	0.09	0.85	0.024	0.03
	SKCM	1.47	1.18	1.82	5E−04	0.002
	THCA	2.05	1.04	4.05	0.039	0.041
	UVM	1.67E+08	1,068	2.621E+13	0.002	0.007
GSDMD	ACC	1.55	1.12	2.16	0.009	0.017
	KIRC	1.54	1.16	2.03	0.002	0.007
	KIRP	0.51	0.31	0.85	0.01	0.018
	LGG	2.44	1.97	3.02	3.65E−16	1.46E−14
	SKCM	0.72	0.62	0.84	1.72E−05	1.38E−04
	UCEC	0.74	0.56	0.98	0.034	0.038
	UVM	3.15	1.66	5.97	4E−04	0.002
GSDME	ACC	0.67	0.5	0.89	0.007	0.014
	HNSC	1.19	1.02	1.4	0.027	0.031
	KICH	2.89	1.03	8.14	0.043	0.045
	KIRC	1.51	1.21	1.9	3E−04	0.001
	KIRP	0.66	0.45	0.97	0.036	0.039
	LIHC	1.63	1.28	2.08	8.72E−05	4.98E−04
	UCEC	1.64	1.1	2.46	0.016	0.026
PJVK	KIRC	2.31	1.62	3.31	3.69E−06	3.69E−05
	KIRP	0.62	0.4	0.99	0.046	0.046
	LAML	0.54	0.32	0.9	0.019	0.027
	MESO	0.53	0.29	0.95	0.034	0.038
	SARC	0.56	0.34	0.92	0.023	0.03

ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; DLBC, diffuse large B-cell lymphoma; GSDM, gasdermin; HNSC, head and neck squamous cell carcinoma; HR, hazard ratio; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LAML, acute myeloid leukemia; LGG, brain lower grade glioma; LIHC, liver hepatocellular carcinoma; MESO, mesothelioma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PJVK, pejvakin; SARC, sarcoma; SKCM, skin cutaneous melanoma; THCA, thyroid carcinoma; UCEC, uterine corpus endometrial carcinoma; UVM, uveal melanoma.

observed in the more advanced stages of these cancers. The expression levels of GSDM genes were closely associated with tumor progression, involving urinary system cancers in particular. Our findings indicated that the expression levels of GSDM genes were globally associated with the clinical stages of cancer, which has strong implications for the regulation of tumor occurrence and progression to advanced disease.

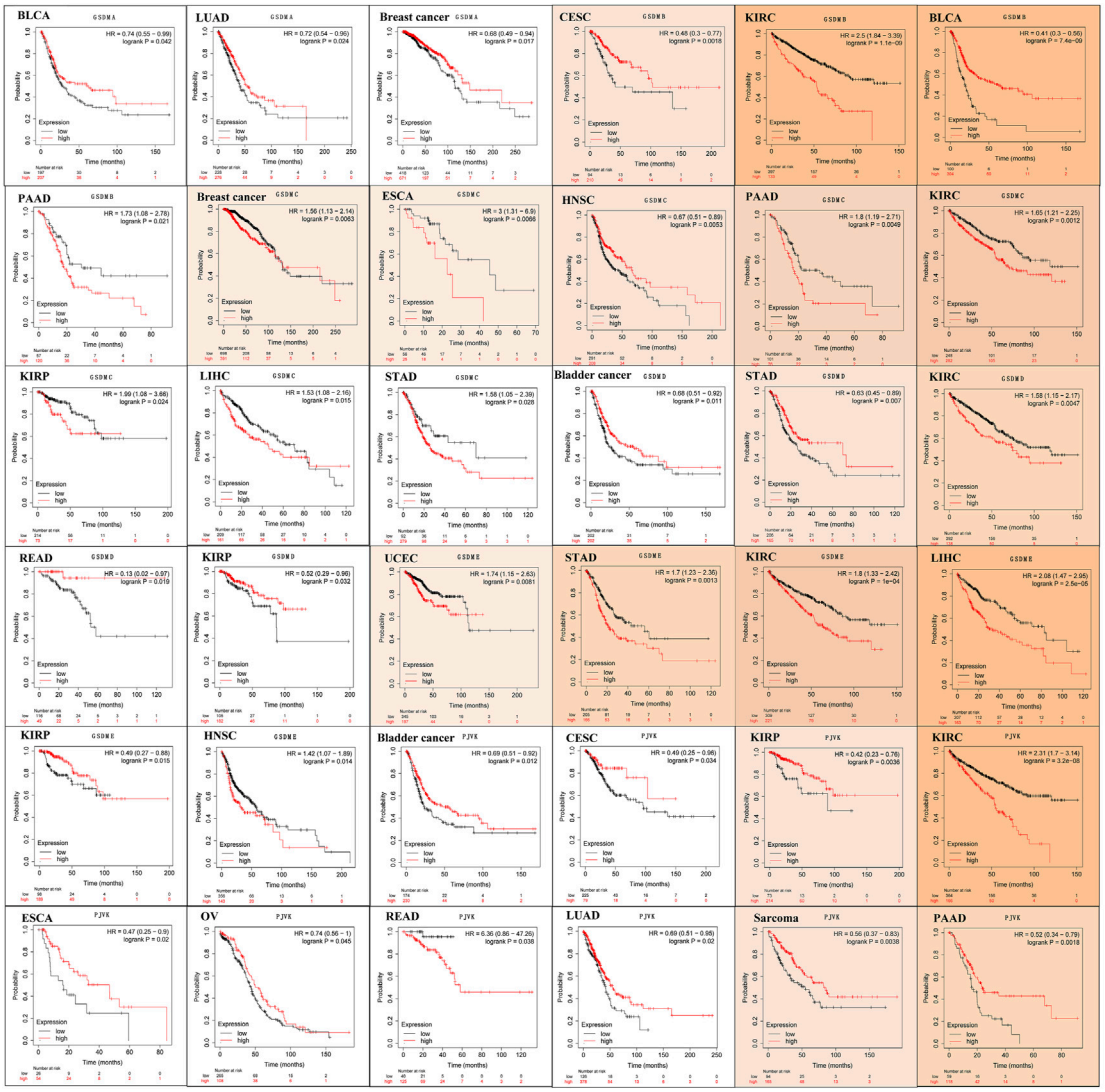


FIGURE 5
Survival analysis in relation to the expression of GSDM genes in the pan-cancer setting (Kaplan-Meier plotter database). Solid color represents log-rank *p*-value; higher color intensity indicates greater statistical significance. A cutoff value of $p < 0.05$ denotes statistically significant differences. GSDM, gasdermin.

Survival analysis in relation to GSDM expression in different cancer types

The prognostic impact of GSDM gene expressions in a pan-cancer setting was analyzed using the Kaplan–Meier survival curve. Using TCGA data, we noticed that the expression of GSDM genes significantly influenced the prognosis of various types of cancer, particularly urinary system cancers (e.g., BLCA, KIRC, and ACC) (Figures 4A–N). Compared to patients exhibiting high GSDM gene expression levels, those with lower gene expression had a marked survival advantage in the majority of cancer types. For example, in patients with KIRC, LIHC, or uveal melanoma, lower GSDM gene expression tended to be associated with better overall survival. However, in BLCA, ACC, acute myeloid leukemia (AML), and sarcoma, higher GSDM expression was linked to a notable survival advantage. In KIRC, lower expression levels of *GSDMB*, *GSDMC*, *GSDME*, and *PJVK* were correlated with better survival. Moreover, we conducted a Cox regression analysis to investigate the value of GSDM genes as prognostic risk factors for each tumor type (Figure 4O). As shown in Table 1, GSDM genes performed as multidimensional factors in cancer prognosis, and their expression significantly predicted the survival of patients with various cancer types, particularly KIRC.

TABLE 2 Significant association between the expression of GSDM genes and the prognosis of different cancer types using data from multiple databases.

Gene	TCGA (cox)	TCGA (Kaplan–Meier)	Kaplan–Meier plotter database
GSDMA	KICH		BLCA
	LAML		LUAD
	LGG		
	OV		
	SKCM		
GSDMB	ACC	KIRC	KIRC
	BLCA	BLCA	BLCA
	DLBC		PAAD
	KICH		CESC
	KIRC		
	LAML		
	PAAD		
	SKCM		
GSDMC	KICH	KIRC	KIRC
	LGG	LGG	KIRP
	LIHC	UVM	ESCA
	PAAD		STAD
	SARC		PAAD
	SKCM		HNSC
	THCA		LIHC
	UVM		
GSDMD	ACC	BLCA	KIRC
	KIRC	LGG	KIRP
	KIRP	UVM	READ
	LGG		STAD
	SKCM		
	UCEC		
	UVM		
GSDME	ACC	ACC	KIRC
	HNSC	KIRC	KIRP
	KICH	LIHC	STAD
	KIRC		LIHC
	KIRP		CESC
	LIHC		HNSC
	UCEC		
PJVK	KIRC	KIRC	KIRC
	KIRP	LAML	KIRP
	LAML	SARC	PAAD
	MESO		ESCA
	SARC		READ

ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; DLBC, diffuse large B-cell lymphoma; ESCA, esophageal carcinoma; GSDM, gasdermin; HNSC, head and neck squamous cell carcinoma; HR, hazard ratio; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LAML, acute myeloid leukemia; LGG, brain lower grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; MESO, mesothelioma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PJVK, pejvakin; READ, rectum adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TCGA, the cancer genome atlas; THCA, thyroid carcinoma; UCEC, uterine corpus endometrial carcinoma; UVM, uveal melanoma.

Analysis of data from the Kaplan-Meier Plotter database revealed that higher *GSDMA* expression was related to better survival in BLCA and breast cancer. The levels of *GSDMB*, *GSDMC*, *GSDMD*, *GSDME*, and *PJVK* were primarily associated with the prognosis of neoplastic disorders of the urinary tract system. Higher *GSEME*, *GSDMD*, and *PJVK*

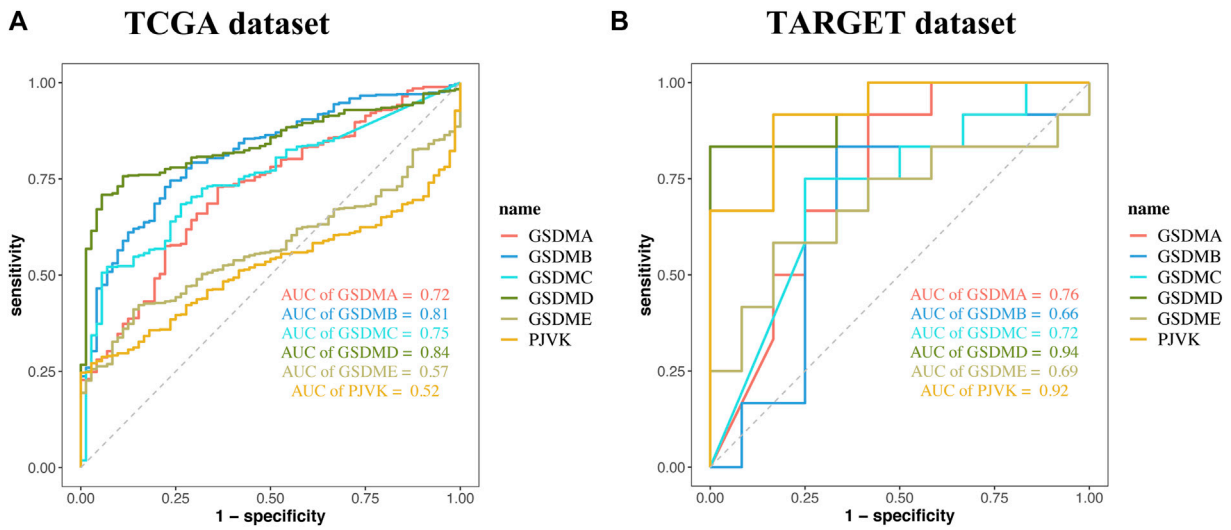


FIGURE 6
The prognostic performance of GSDM genes. Receiver operating characteristic (ROC) curve showing the prognostic performances of GSDM genes in KIRC using data from TCGA **(A)** and TARGET **(B)** databases. GSDM, gasdermin; KIRC, kidney renal clear cell carcinoma; TARGET, therapeutically applicable research to generate effective treatments; TCGA, The Cancer Genome Atlas.

expressions were linked to a survival advantage in KIRP, but were in contrast associated with worse survival in KIRC, STAD, and READ. Interestingly, there was a highly significant correlation between GSDM family gene expression and tumor outcomes in genitourinary cancers, such as KIRC (Figure 5 and Table 2). Subsequently, we assessed the value of GSDM genes in the diagnosis of KIRC. Using TCGA data, we determined that the AUCs of the ROC curves for the prediction of KIRC were 0.72, 0.81, 0.75, 0.84, 0.57, and 0.52 for *GSDMA*, *GSDMB*, *GSDMC*, *GSDMD*, *GSDME*, and *PJVK*, respectively (Figure 6A). Similar results were observed following the analysis of data from the TARGET database (Figure 6B). Collectively, our findings confirm that GSDM genes have significant prognostic value across many different types of cancer, and represent promising biomarkers for cancer therapy.

The association between GSDM gene expression and the TME and stemness score in a pan-cancer setting

We conducted an immune infiltration analysis and an RNA/DNA stemness score evaluation to explore the correlation between GSDM gene expression and the TME and stemness level in the pan-cancer setting (Figure 7). Our results indicated that the expressions of GSDM family genes, particularly GSDMD, were significantly positively correlated with the immune score (Figure 7A) and negatively correlated with the RNA or DNA stemness score and tumor purity in various cancer types (Figures 7B–D). Notably, *PJVK* expression was negatively correlated with the level of immune infiltration but positively correlated with tumor purity (Figure 7D). These data demonstrate that higher GSDM gene expression (except for *PJVK*) may be associated with an increase in immune cell infiltration into the tumor, together with a decrease in tumor purity and stemness activity, thus potentially improving patient outcomes.

Association between GSDM gene expression and cancer drug sensitivity

We next sought to investigate the potential correlation between the expression of each GSDM gene and the drug sensitivity of cancer cells. For this purpose, we downloaded the drug sensitivity data from the CellMiner™ database, and processed these data using the R software. As illustrated in Figure 8, *GSDMA* expression was highly positively associated with the sensitivity of tumor cells to dexrazoxane, while GSDMB expression showed a markedly positive association with sensitivity to nelarabine, fluphenazine, and perifosine. *GSDMC* expression was negatively correlated with the sensitivity to ixazomib citrate, vincristine, midostaurin, bortezomib, and pralatrexate, as well as positively correlated with sensitivity to gefitinib and lificguat. *GSDMD* expression exhibited a highly positive association with sensitivity to fludarabine and 5-fluoro deoxy uridine 10mer. Finally, *PJVK* expression was positively related to the sensitivity to nelarabine, PX-316, and fluphenazine.

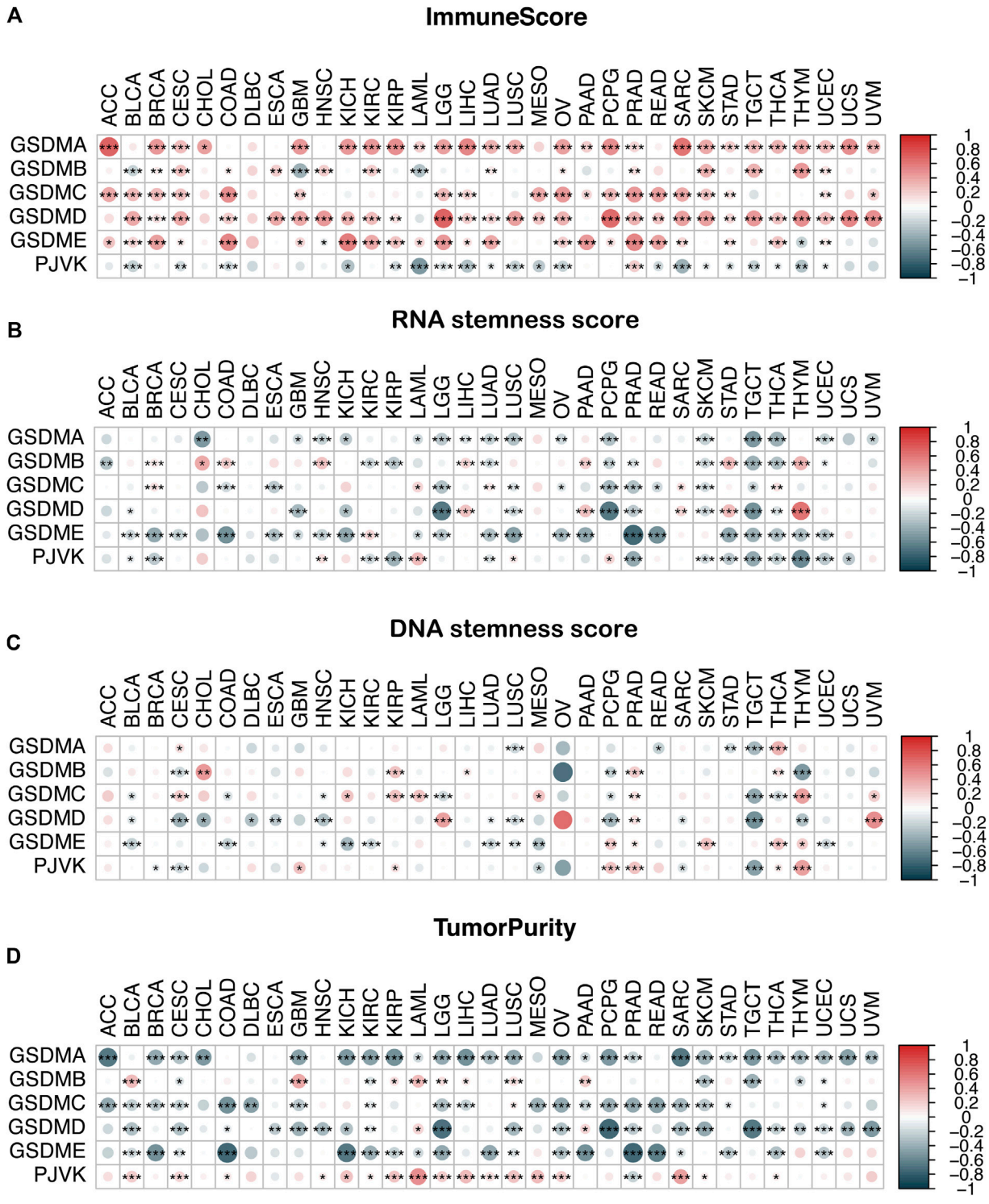


FIGURE 7
Correlation analysis of GSDM gene expression and the TME and stemness score in the pan-cancer setting. **(A–D)** Association between GSDM gene expression and the immune score **(A)**, the RNA stemness score **(B)**, the DNA stemness score **(C)**, and tumor purity **(D)** in the different cancers; red represents a positive correlation, and blue represents a negative correlation. An adjusted cutoff value of $p < 0.05$ denotes statistically significant differences (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). GSDM, gasdermin; TME, tumor microenvironment.

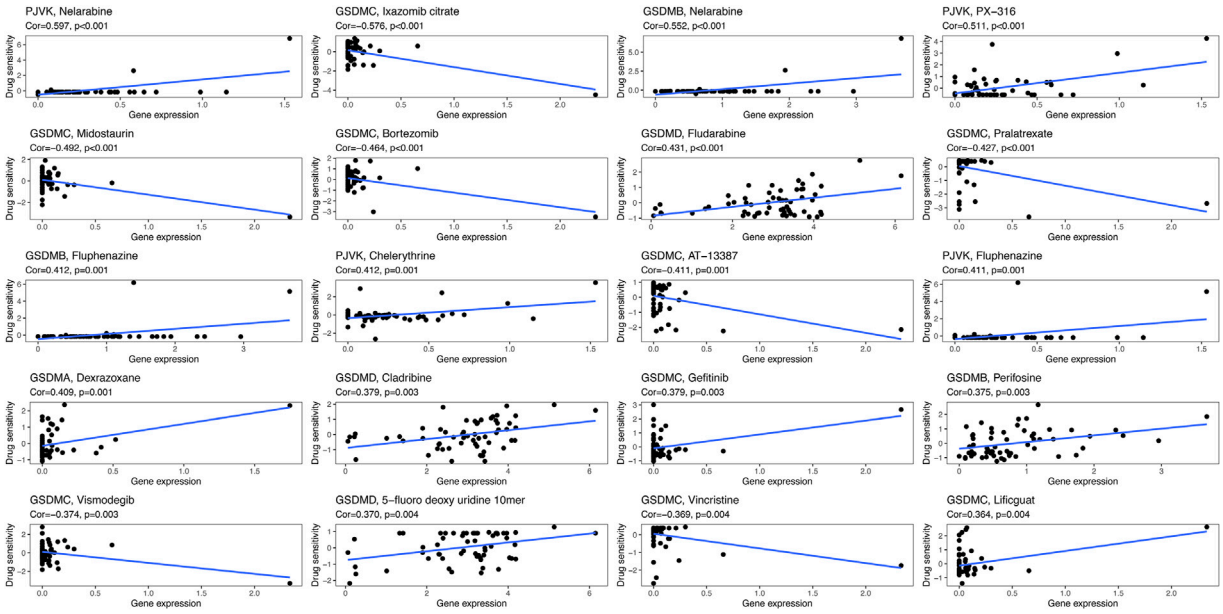


FIGURE 8
Drug sensitivity analysis of GSDM genes in the pan-cancer setting. GSDM, gasdermin.

Discussion

In this study, we conducted a comprehensive analysis of GSDM gene expression across various human cancers and matched normal tissues using multidimensional TCGA data. Based on their membrane pore-forming activity, the GSDM proteins (except PJVK) are important mediators of pyroptosis and participate in tumorigenesis (Tan et al., 2020; Jiang et al., 2021). Recently, it has been reported that two members of the GSDM family, namely *GSDMD* and *GSDME*, stimulate numerous downstream pyroptotic pathways, and the downregulation of these molecules conversely contributed to tumorigenesis and proliferation in the tumor cell microenvironment (Wang et al., 2018a; Gao et al., 2018; De Schutter et al., 2021). Although previous studies have investigated the gene expression patterns of some pyroptosis genes in tumors and normal tissues (Hu et al., 2021; Qiu et al., 2021), their analyses were restricted to a specific cancer type. Therefore, the objectives of the present analysis were to determine the expression patterns of GSDM genes in numerous cancer types and adjacent normal tissues, as well as evaluate their association with immune infiltration, genetic variation, and drug sensitivity. We hope that our work will provide new insights into the use of GSDM genes as pan-cancer biomarkers.

Programmed cell death plays pivotal role in coordinating homeostasis and cell proliferation in malignant disorders, such as tumors and inflammatory or infectious diseases (Bedoui et al., 2020). Pyroptosis is a form of programmed cell death that triggers lytic cell rupture in order to maintain homeostasis via an intricate cascade of potassium efflux, water influx, and cellular swelling processes, all of which are dependent on GSDM family proteins (Bedoui et al., 2020; Broz et al., 2020). Six GSDM family genes (i.e., *GSDMA*, *GSDMB*, *GSDMC*, *GSDMD*, *GSDME*, and *PJVK*) have long been described in other fields; however, they have been rarely investigated in a pan-cancer setting. Thus far, it has been reported that GSDMA is involved in skin inflammation (Zhou et al., 2012), epidermal differentiation (Lin et al., 2015), and the development of asthma (Ferreira et al., 2014). Moreover, low expression levels of the GSDMA protein have been found in gastric cancers (Saeki et al., 2000). Upregulation of *GSDMB* participates in the pathogenesis of breast, hepatic, cervical, and gastrointestinal cancers (Sun et al., 2008; Hergueta-Redondo et al., 2014) and has been linked to poor patient survival. *GSDMC* and *GSDMD* are markedly upregulated in breast cancer and colorectal cancers (Miguchi, et al., 2016; Hou et al., 2020), where they promote tumor cell proliferation and are associated with a worse prognosis (Wang et al., 2019). Elevated *GSDME* activity exerts suppressive effects on carcinogenesis and cancer progression (Wang et al., 2018b). Our analysis of TGGA data demonstrated that GSDM genes (*GSDMB*, *GSDMC*, and *GSDMD* in particular) were upregulated in different types of cancers compared with the corresponding normal specimens. Furthermore, a marked increase in GSDM gene expression was observed in the majority of kidney carcinomas that progressed from early to advanced clinical stages.

An unrestrained inflammatory response is highly amplified in the TME, and the excessive activation of cytokine secretion

pathways leads to pathological consequences. The release of pro-inflammatory cytokines is crucially dependent on the activation of GSDM genes (Heilig et al., 2018; Li et al., 2021; Liu et al., 2021). GSDM-meditated pyroptosis culminates in the release of tissue factor, which activates tumor inflammation pathways, leading to poor outcomes in patients with cancer. In fact, the GSDM genes are also closely associated with tumor immunity. For example, the tumor-suppressive effect of *GSDME* is mediated by tumor-infiltrating natural killer (NK) and CD8+T lymphocytes (Zhang et al., 2020). Meanwhile, the overexpression of *GSDMB* in cancer cells leads to the recruitment of immune cells and promotes tumor mobility and invasion (Hergueta-Redondo et al., 2014; Cui et al., 2021). We have observed that the expression of GSDM genes is amplified in many human cancers, particularly kidney cancer. Heightened GSDM gene expression may lead to the activation of different stimuli and inflammatory caspases, and trigger the infiltration of immune cells, thus reflecting the complex role of pyroptosis in tumorigenesis, antitumor immunity, tumor cell growth, and metastasis. Therefore, our research suggests that GSDM gene family members may function as oncogenes in the pan-cancer setting and could represent potential prognostic biomarkers and immunotherapeutic targets in the treatment of kidney cancer in particular.

We next explored the genetic alterations affecting GSDM genes in various types of cancer. In general, a higher frequency of somatic mutations is thought to be associated with the increased generation of more neoantigens. The TMB is defined as the number of coding, and somatic mutations per megabase of the interrogated genome. It is similar to high-MSI in cancer, and is involved in the generation of immunogenic neuropeptides on the tumor cell surface, thus influencing patient responses to immunotherapy (Yamamoto and Imai, 2019; Huang et al., 2021). Currently, TMB and MSI are predictors of the efficacy of immunotherapy. In this study, we found that the mutational frequencies within GSDM genes were at an extremely low level, and that the expression levels of GSDM genes were significantly correlated with TMB and MSI across various urologic neoplasms. This result suggests that the GSDM genes may be stable biomarkers and should be considered in potential treatment strategies for cancer.

Subsequently, we performed survival and Cox regression analyses, which revealed that the GSDM genes were significantly associated with the prognosis of urinary tract system cancer (e.g., KIRC, PAAD, KIRP, and BLCA). For instance, downregulated GSDM genes were related to favorable survival outcomes in KIRC, while some of the upregulated GSDM genes predicted better KIRP and BLCA prognoses. In addition, we found that the expression levels of GSDM genes were more frequently correlated with the prognosis of KIRC. We subsequently analyzed the predictive value of the GSDM genes in KIRC. The results showed that *GSDMD* presented the highest AUC score, followed by *GSDMB* and *GSDMC*. These findings indicated that the diagnostic value of GSDMD may be better than that of the other GSDM genes. KIRC is a highly infiltrative tumor that remains one of the most sensitive tumors in terms of response to immunotherapy (Inman et al., 2013; Newman et al., 2015). It has been reported that some spontaneous KIRC regressions are accompanied by signs of immune mediation (Janiszewska et al., 2013). The GSDM genes are involved in cell programmed cell death; and *GSDMD* in particular, plays an important role in systemic immune-inflammatory sensing (Yao et al., 2022). Dysregulation of the GSDM genes may cause a dysfunctional adaptive immune response, as well as contribute to both the initiation and progression of multiple tumors (Xia et al., 2019; Fang et al., 2020). Previous studies have shown that the upregulation of *GSDMD* in tumor tissues was associated with a poor prognosis, due to its involvement in AKT-related signaling pathways (Wang et al., 2018a; Gao et al., 2018). Thus, we hypothesized that *GSDMD* may trigger the release of inflammatory factors and induce potential interactions with other immune responses, thereby promoting the invasiveness of kidney tumors. These findings imply that the dysregulations of GSDM genes could predict survival in patients with cancers, and that *GSDMD* may represent a robust biomarker for the evaluation of KIRC.

We then went on to conduct a potential correlation analysis between the expression of the GSDM gene and the immune infiltration, immune subtype, or RNA/DNA stemness score in the pan-cancer setting. Previously, Thorsson et al. performed a comprehensive immunogenomic analysis of all TCGA cancer types and successfully identified six immune subtypes (C1–C6) (Thorsson et al., 2018). Through our research, we discovered that higher expression levels of *GSDMA*, *GSDMB*, *GSDMC*, and *GSDMD* were more closely related to hyperimmune subtypes, such as C1, C2, and C6. Lower GSDM expression levels (except for *PJVK* and *GSDME*) were correlated with hypoimmune subtypes, such as C4 and C5. Consistent with these findings, the immune infiltration analysis showed that GSDM gene expressions were significantly positively correlated with the immune cell infiltration score of nearly all cancer types. Thus, GSDM-dependent pyroptosis could increase immune cell activation within the tumor, thereby contributing to immune-mediated tumor cell regression (Tsuchiya, 2021). Moreover, recent studies have confirmed that the stemness index (RNA/DNA stemness score) was significantly higher in patients with metastases and disease recurrence. The stemness score was also correlated with intratumor heterogeneity, immune response, and drug resistance (Malta et al., 2018; Yi et al., 2020; Shi et al., 2021). We used two independent stemness indices, namely the DNA stemness score (reflecting epigenetic features) and RNA stemness score (reflecting mRNA gene expression) (Malta et al., 2018). Previous reports have found that lower stemness indices were correlated with an increased leukocyte fraction and higher programmed cell death-ligand 1 (PD-L1) expression levels (Zaretsky et al., 2016; Malta et al., 2018). In this study, we

observed a more negative correlation between the tumor stemness and GSDM gene expression in most cancer types. We argue that such tumors would be more susceptible to immune checkpoint blockade due to sufficient immune infiltration and the up-regulation of the PD-L1-associated gene pathway, which further enhances treatment efficacy (Zaretsky et al., 2016). However, the survival analysis indicated that patients with kidney cancer (e.g., KIRC) and a higher level of GSDM gene expression were associated with a significantly poorer prognosis. These findings may be explained by the immune-excluded phenotype of the TME, which is characterized by an abundance of immune cells in the TME that are retained in the stroma and do not penetrate tumor cell nests. Therefore, immune cell infiltration appeared to occur outside the tumor (Salmon et al., 2012; Joyce and Fearon, 2015). Based on these results, we hypothesized that the evaluation of the stemness index and the immune infiltration score of GSDM-mediated cancer pyroptosis may provide more effective immunotherapeutic options for cancer.

Finally, we analyzed the correlation between the expression of GSDM genes and drug sensitivity in 33 cancer types using data from the CellMinerTM database. The results showed that the expression of GSDM genes positively correlated with the sensitivity of tumor cells to nelarabine, fluphenazine, and dexrazoxane. However, higher expression levels of these genes were also linked to reduced cell sensitivity to bortezomib, midostaurin, and vincristine. Nelarabine is an effective anticancer chemotherapy prodrug of arabinfuranosylguanine triphosphate, that appears to meditate DNA degradation and cell death (Robak et al., 2006). Similarly, midostaurin (Hsiao et al., 2019), fluphenazine (Otręba and Kośmider, 2021), bortezomib (Vora et al., 2020), and vincristine (Xu and Xu, 2020) also possess anticancer activity. The present findings revealed that GSDM gene expression may provide important guidance for the selection of targeted drugs for the treatment of cancer.

This study has some limitations. Our analyses were based on the mRNA expression level data extracted from online databases, and the results were obtained using bioinformatic methods. Therefore, we lacked evidence from *in vitro* or *in vivo* experiments to support these findings. Further investigation is warranted to validate our findings and to elucidate the roles of GSDM genes in cancer.

In summary, our work revealed statistically significant variations in GSDM mRNA expression levels between different tumor tissues and healthy human organs. We also observed that the GSDM genes participated in tumorigenesis, as they were correlated with tumor immune subtypes, patient survival, the TME, and the stemness score. Our pan-cancer analysis indicates that the expression of GSDM genes (in particular *GSDMB*, *GSDMC*, and *GSDMD*) was significantly associated with the survival of patients with certain types of urinary tract system cancer. Thus, GSDM genes may represent potential prognostic biomarkers for these cancers. Of the six GSDM family members, GSDMD, in particular, has shown promise as a biomarker for the evaluation of KIRC. Moreover, we found that the expression levels of GSDM genes in tumor cell lines were correlated with varying sensitivities of specific cancer drugs. These findings may provide new insights into the potential use of GSDM genes as therapeutic targets in the pan-cancers setting.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://xena.ucsc.edu/, http://www.cbioportal.org, https://discover.nci.nih.gov/cellminer/home.do, https://gtexportal.org/home/, and https://www.cbioportal.org.

Author contributions

YZ and DY designed the study and comprehensively analyzed the data. DY and YZ drafted the manuscript. YZ and FZ revised the original manuscript and provided suggestions for data analysis. RT supervised the study and proofread 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

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.

References

Bedoui, S., Herold, M. J., and Strasser, A. (2020). Emerging connectivity of programmed cell death pathways and its physiological implications. *Nat. Rev. Mol. Cell. Biol.* 11, 678–695. doi:10.1038/s41580-020-0270-8

Bonneville, R., Krook, M. A., Kautto, E. A., Miya, J., Wing, M. R., Chen, H. Z., et al. (2017). Landscape of microsatellite instability across 39 cancer types. *JCO Precis. Oncol.* 2017, 1–15. doi:10.1200/po.17.00073

Broz, P., Pelegrín, P., and Shao, F. (2020). The gasdermins, a protein family executing cell death and inflammation. *Nat. Rev. Immunol.* 3, 143–157. doi:10.1038/s41577-019-0228-2

Croes, L., Fransen, E., Hylebos, M., Buys, K., Hermans, C., Broeckx, G., et al. (2019). Determination of the potential tumor-suppressive effects of gsdme in a chemically induced and in a genetically modified intestinal cancer mouse model. *Cancers (Basel)* 8, 1214. doi:10.3390/cancers11081214

Cui, Y., Zhou, Z., Chai, Y., and Zhang, Y. (2021). Upregulated GSDMB in clear cell renal cell carcinoma is associated with immune infiltrates and poor prognosis. *J. Immunol. Res.* 2021, 7753553. doi:10.1155/2021/7753553

De Schutter, E., Croes, L., Ibrahim, J., Pauwels, P., Op de Beeck, K., Vandenabeele, P., et al. (2021). GSDME and its role in cancer: From behind the scenes to the front of the stage. *Int. J. Cancer* 12, 2872–2883. doi:10.1002/ijc.33390

Ding, J., Wang, K., Liu, W., She, Y., Sun, Q., Shi, J., et al. (2016). Pore-forming activity and structural autoinhibition of the gasdermin family. *Nature* 7610, 111–116. doi:10.1038/nature18590

Fang, Y., Tian, S., Pan, Y., Li, W., Wang, Q., Tang, Y., et al. (2020). Pyroptosis: A new frontier in cancer. *Biomed. Pharmacother.* 121, 109595. doi:10.1016/j.biopha.2019.109595

Ferreira, M. A., Matheson, M. C., Tang, C. S., Granell, R., Ang, W., Hui, J., et al. (2014). Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. *J. Allergy Clin. Immunol.* 6, 1564–1571. doi:10.1016/j.jaci.2013.10.030

Gao, J., Qiu, X., Xi, G., Liu, H., Zhang, F., Lv, T., et al. (2018). Downregulation of GSDMD attenuates tumor proliferation via the intrinsic mitochondrial apoptotic pathway and inhibition of EGFR/Akt signaling and predicts a good prognosis in non-small cell lung cancer. *Oncol. Rep.* 4, 1971–1984. doi:10.3892/or.2018.6634

Heilig, R., Dick, M. S., Sborgi, L., Meunier, E., Hiller, S., Broz, P., et al. (2018). The Gasdermin-D pore acts as a conduit for IL-1β secretion in mice. *Eur. J. Immunol.* 4, 584–592. doi:10.1002/eji.201747404

Hergueta-Redondo, M., Sarrió, D., Molina-Crespo, Á., Megias, D., Mota, A., Rojo-Sebastian, A., et al. (2014). Gasdermin-B promotes invasion and metastasis in breast cancer cells. *PLoS One* 3, e90099. doi:10.1371/journal.pone.0090099

Hergueta-Redondo, M., Sarrio, D., Molina-Crespo, Á., Vicario, R., Bernadó-Morales, C., Martínez, L., et al. (2016). Gasdermin B expression predicts poor clinical outcome in HER2-positive breast cancer. *Oncotarget* 35, 56295–56308. doi:10.18632/oncotarget.10787

Hou, J., Zhao, R., Xia, W., Chang, C. W., You, Y., Hsu, J. M., et al. (2020). PD-L1-mediated gasdermin C expression switches apoptosis to pyroptosis in cancer cells and facilitates tumour necrosis. *Nat. Cell. Biol.* 10, 1264–1275. doi:10.1038/s41556-020-0575-z

Hsiao, S. H., Lusvarghi, S., Huang, Y. H., Ambudkar, S. V., Hsu, S. C., Wu, C. P., et al. (2019). The FLT3 inhibitor midostaurin selectively resensitizes ABCB1-overexpressing multidrug-resistant cancer cells to conventional chemotherapeutic agents. *Cancer Lett.* 445, 34–44. doi:10.1016/j.canlet.2019.01.001

Hu, K., Xu, Z., Yao, L., Yan, Y., Zhou, L., Li, J., et al. (2021). Integrated analysis of expression, prognostic value and immune infiltration of GSDMs in hepatocellular carcinoma. *Aging (Albany NY)* 21, 24117–24135. doi:10.18632/aging.203669

Huang, R. S. P., Haberberger, J., Severson, E., Duncan, D. L., Hemmerich, A., Edgerly, C., et al. (2021). A pan-cancer analysis of PD-L1 immunohistochemistry and gene amplification, tumor mutation burden and microsatellite instability in 48, 782 cases. *Mod. Pathol.* 2, 252–263. doi:10.1038/s41379-020-00664-y

Inman, B. A., Harrison, M. R., and George, D. J. (2013). Novel immunotherapeutic strategies in development for renal cell carcinoma. *Eur. Urol.* 5, 881–889. doi:10.1016/j.eururo.2012.10.006

Janiszewska, A. D., Poletajew, S., and Wasiutyński, A. (2013). Spontaneous regression of renal cell carcinoma. *Contemp. Oncol.* 2, 123–127. doi:10.5114/wo.2013.34613

Jiang, A., Meng, J., Bao, Y., Wang, A., Gong, W., Gan, X., et al. (2021). Establishment of a prognosis prediction model based on pyroptosis-related signatures associated with the immune microenvironment and molecular heterogeneity in clear cell renal cell carcinoma. *Front. Oncol.* 11, 755212. doi:10.3389/fonc.2021.755212

Johnson, D. C., Taabazuing, C. Y., Okondo, M. C., Chui, A. J., Rao, S. D., Brown, F. C., et al. (2018). DPP8/DPP9 inhibitor-induced pyroptosis for treatment of acute myeloid leukemia. *Nat. Med.* 8, 1151–1156. doi:10.1038/s41591-018-0082-y

Joyce, J. A., and Fearon, D. T. (2015). T cell exclusion, immune privilege, and the tumor microenvironment. *Science* 6230, 74–80. doi:10.1126/science.aaa6204

Kuang, S., Zheng, J., Yang, H., Li, S., Duan, S., Shen, Y., et al. (2017). Structure insight of GSDMD reveals the basis of GSDMD autoinhibition in cell pyroptosis. *Proc. Natl. Acad. Sci. U. S. A.* 40, 10642–10647. doi:10.1073/pnas.1708194114

Li, L., Jiang, M., Qi, L., Wu, Y., Song, D., Gan, J., et al. (2021). Pyroptosis, a new bridge to tumor immunity. *Cancer Sci.* 10, 3979–3994. doi:10.1111/cas.15059

Lin, P. H., Lin, H. Y., Kuo, C. C., and Yang, L. T. (2015). N-terminal functional domain of Gasdermin A3 regulates mitochondrial homeostasis via mitochondrial targeting. *J. Biomed. Sci.* 1, 44. doi:10.1186/s12929-015-0152-0

Liu, X., Xia, S., Zhang, Z., Wu, H., and Lieberman, J. (2021). Channelling inflammation: Gasdermins in physiology and disease. *Nat. Rev. Drug Discov.* 5, 384–405. doi:10.1038/s41573-021-00154-z

Malta, T. M., Sokolov, A., Gentles, A. J., Burzykowski, T., Poisson, L., Weinstein, J. N., et al. (2018). Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell.* 2, 338–354. e315. doi:10.1016/j.cell.2018.03.034

Miguchi, M., Hinoi, T., Shimomura, M., Adachi, T., Saito, Y., Niitsu, H., et al. (2016). Gasdermin C is upregulated by inactivation of transforming growth factor β receptor type II in the presence of mutated apc, promoting colorectal cancer proliferation. *PLoS One* 11, e0166422. doi:10.1371/journal.pone.0166422

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 5, 453–457. doi:10.1038/nmeth.3337

Otręba, M., and Kośmider, L. (2021). *In vitro* anticancer activity of fluphenazine, perphenazine and prochlorperazine. A review. *J. Appl. Toxicol.* 1, 82–94. doi:10.1002/jat.4046

Pizato, N., Luzete, B. C., Kiffer, L., Corrêa, L. H., de Oliveira Santos, I., Assumpção, J. A. F., et al. (2018). Omega-3 docosahexaenoic acid induces pyroptosis cell death in triple-negative breast cancer cells. *Sci. Rep.* 1, 1952. doi:10.1038/s41598-018-20422-0

Qiu, S., Hu, Y., and Dong, S. (2021). Pan-cancer analysis reveals the expression, genetic alteration and prognosis of pyroptosis key gene GSDMD. *Int. Immunopharmacol.* 101, 108270. doi:10.1016/j.intimp.2021.108270

Quail, D. F., and Joyce, J. A. (2013). Microenvironmental regulation of tumor progression and metastasis. *Nat. Med.* 11, 1423–1437. doi:10.1038/nm.3394

Robak, T., Lech-Maranda, E., Korycka, A., and Robak, E. (2006). Purine nucleoside analogs as immunosuppressive and antineoplastic agents: Mechanism of action and clinical activity. *Curr. Med. Chem.* 26, 3165–3189. doi:10.2174/092986706778742918

Saeki, N., Kuwahara, Y., Sasaki, H., Satoh, H., and Shiroishi, T. (2000). Gasdermin (Gsdm) localizing to mouse Chromosome 11 is predominantly expressed in upper gastrointestinal tract but significantly suppressed in human gastric cancer cells. *Mamm. Genome* 9, 718–724. doi:10.1007/s003350010138

Salmon, H., Franciszkiewicz, K., Damotte, D., Dieu-Nosjean, M. C., Validire, P., Trautmann, A., et al. (2012). Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *J. Clin. Invest.* 3, 899–910. doi:10.1172/jci45817

Shi, J., Gao, W., and Shao, F. (2017). Pyroptosis: Gasdermin-Mediated programmed necrotic cell death. *Trends biochem. Sci.* 4, 245–254. doi:10.1016/j.tibs.2016.10.004

Shi, X., Liu, Y., Cheng, S., Hu, H., Zhang, J., Wei, M., et al. (2021). Cancer stemness associated with prognosis and the efficacy of immunotherapy in adrenocortical carcinoma. *Front. Oncol.* 11, 651622. doi:10.3389/fonc.2021.651622

Sun, Q., Yang, J., Xing, G., Sun, Q., Zhang, L., He, F., et al. (2008). Expression of GSDML associates with tumor progression in uterine cervix cancer. *Transl. Oncol.* 2, 73–83. doi:10.1593/tlo.08112

Tan, G., Huang, C., Chen, J., and Zhi, F. (2020). HMGB1 released from GSDME-mediated pyroptotic epithelial cells participates in the tumorigenesis of colitis-associated colorectal cancer through the ERK1/2 pathway. *J. Hematol. Oncol.* 1, 149. doi:10.1186/s13045-020-00985-0

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The immune landscape of cancer. *Immunity* 4, 812–830. e814. doi:10.1016/j.immuni.2018.03.023

Tsuchiya, K. (2021). Switching from apoptosis to pyroptosis: Gasdermin-elicited inflammation and antitumor immunity. *Int. J. Mol. Sci.* 1, 426. doi:10.3390/ijms22010426

Vora, P. A., Patel, R., and Dharamsi, A. (2020). Bortezomib - first therapeutic proteasome inhibitor for cancer therapy: A review of patent literature. *Recent Pat. anticancer. Drug Discov.* 2, 113–131. doi:10.2174/1574892815666200401113805

Wang, W. J., Chen, D., Jiang, M. Z., Xu, B., Li, X. W., Chu, Y., et al. (2018a). Downregulation of gasdermin D promotes gastric cancer proliferation by regulating cell cycle-related proteins. *J. Dig. Dis.* 2, 74–83. doi:10.1111/1751-2980.12576

Wang, Y., Yin, B., Li, D., Wang, G., Han, X., Sun, X., et al. (2018b). GSDME mediates caspase-3-dependent pyroptosis in gastric cancer. *Biochem. Biophys. Res. Commun.* 1, 1418–1425. doi:10.1016/j.bbrc.2017.11.156

Wang, Y. Y., Liu, X. L., and Zhao, R. (2019). Induction of pyroptosis and its implications in cancer management. *Front. Oncol.* 9, 971. doi:10.3389/fonc.2019.00971

Wu, L. S., Liu, Y., Wang, X. W., Xu, B., Lin, Y. L., Song, Y., et al. (2020). LPS enhances the chemosensitivity of oxaliplatin in HT29 cells via GSDMD-mediated pyroptosis. *Cancer Manag. Res.* 12, 10397–10409. doi:10.2147/cmar.S244374

Xia, X., Wang, X., Cheng, Z., Qin, W., Lei, L., Jiang, J., et al. (2019). The role of pyroptosis in cancer: Pro-cancer or pro-"host. *Cell. Death Dis.* 9, 650. doi:10.1038/s41419-019-1883-8

Xu, D., and Xu, Z. (2020). Indole alkaloids with potential anticancer activity. *Curr. Top. Med. Chem.* 21, 1938–1949. doi:10.2174/1568026620666200622150325

Yamamoto, H., and Imai, K. (2019). An updated review of microsatellite instability in the era of next-generation sequencing and precision medicine. *Semin. Oncol.* 3, 261–270. doi:10.1053/j.seminoncol.2019.08.003

Yao, L., Li, J., Xu, Z., Yan, Y., and Hu, K. (2022). GSDMs are potential therapeutic targets and prognostic biomarkers in clear cell renal cell carcinoma. *Aging (Albany NY)* 6, 2758–2774. doi:10.18632/aging.203973

Yi, L., Huang, P., Zou, X., Guo, L., Gu, Y., Wen, C., et al. (2020). Integrative stemness characteristics associated with prognosis and the immune microenvironment in esophageal cancer. *Pharmacol. Res.* 161, 105144. doi:10.1016/j.phrs.2020.105144

Yu, J., Li, S., Qi, J., Chen, Z., Wu, Y., Guo, J., et al. (2019). Cleavage of GSDME by caspase-3 determines lobaplatin-induced pyroptosis in colon cancer cells. *Cell. Death Dis.* 3, 193. doi:10.1038/s41419-019-1441-4

Zaretsky, J. M., Garcia-Diaz, A., Shin, D. S., Escuin-Ordinas, H., Hugo, W., Hu-Lieskovan, S., et al. (2016). Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N. Engl. J. Med.* 9, 819–829. doi:10.1056/NEJMoa1604958

Zhang, Z., Zhang, Y., Xia, S., Kong, Q., Li, S., Liu, X., et al. (2020). Gasdermin E suppresses tumour growth by activating anti-tumour immunity. *Nature* 7799, 415–420. doi:10.1038/s41586-020-2071-9

Zhou, Y., Jiang, X., Gu, P., Chen, W., Zeng, X., Gao, X., et al. (2012). Gsdma3 mutation causes bulge stem cell depletion and alopecia mediated by skin inflammation. *Am. J. Pathol.* 2, 763–774. doi:10.1016/j.ajpath.2011.10.034

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 09 August 2022
DOI 10.3389/fgene.2022.937948

Check for updates

OPEN ACCESS

EDITED BY
Ehsan Nazemalhosseini-Mojarad,
Shahid Beheshti University of Medical Sciences, Iran

REVIEWED BY
Pedram Azimzadeh,
University of Maryland, United States
Hetong Zhao,
Second Military Medical University, China

*CORRESPONDENCE
Wan Qin,
wanqinhust@hotmail.com

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 06 May 2022
ACCEPTED 05 July 2022
PUBLISHED 09 August 2022

CITATION
Zou J and Qin W (2022), Comprehensive analysis of the cancer driver genes constructs a seven-gene signature for prediction of survival and tumor immunity in hepatocellular carcinoma. *Front. Genet.* 13:937948. doi: 10.3389/fgene.2022.937948

COPYRIGHT
© 2022 Zou and Qin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Comprehensive analysis of the cancer driver genes constructs a seven-gene signature for prediction of survival and tumor immunity in hepatocellular carcinoma

Jun Zou[1] and Wan Qin[2]*

[1]Department of Pediatrics, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [2]Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Hepatocellular carcinoma (HCC) is a highly malignant and heterogeneous tumor with poor prognosis. Cancer driver genes (CDGs) play an important role in the carcinogenesis and progression of HCC. In this study, we comprehensively investigated the expression, mutation, and prognostic significance of 568 CDGs in HCC. A prognostic risk model was constructed based on seven CDGs (CDKN2C, HRAS, IRAK1, LOX, MYCN, NRAS, and PABPC1) and verified to be an independent prognostic factor in both TCGA and ICGC cohorts. The low-score group, which showed better prognosis, had a high proportion of $CD8^+$ T cells and elevated expression of interferon-related signaling pathways. Additionally, we constructed a nomogram to extend the clinical applicability of the prognostic model, which exhibits excellent predictive accuracy for survival. Our study showed the important role of CDGs in HCC and provides a novel prognostic indicator for HCC.

KEYWORDS

cancer driver gene, tumor microenvironment, prognosis, nomogram, hepatocellular carcinoma

Introduction

Global cancer statistics show that primary liver cancer is the sixth most commonly diagnosed cancer and the third leading cause of cancer death worldwide in 2020 (Sung et al., 2021). The most common type of liver cancer is hepatocellular carcinoma (HCC). Owing to the specific phenotypes, most HCC patients are diagnosed at an advanced stage with extremely poor prognosis (El-Serag, 2011). Although significant progress has been made in the diagnosis and treatment of HCC, the survival rate for patients over 5 years has not improved (Bruix et al., 2014). Recent studies about the molecular biological characteristics and the tumor microenvironments have revolutionized the management of HCC patients, leading to a transition from traditional chemotherapy

to novel target therapy and immunotherapy (Shen et al., 2010; El-Khoueiry et al., 2017; Kudo et al., 2018; Zhu et al., 2018). However, only a small fraction of HCC patients can benefit from these novel therapeutic options, and some patients inevitably suffer from drug resistance (Kimura et al., 2018). Thus, further efforts are still needed to excavate the underlying molecular biological mechanisms as well as to seek effective predictive measures for personalized therapy.

As is known, cancer is a genetic disease characterized by changes in the genome, genes, chromatin, and cellular levels (Brown et al., 2019). Mutations in driver genes support the acquisition of cancer hallmarks (Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011). Recently, the journal of *Nature Reviews Cancer* reported a compendium of 568 cancer driver genes (CDGs), which was identified from more than 28,000 tumors of 66 cancer types (Martinez-Jimenez et al., 2020). Mutations and aberrant expression of these genes may affect cell growth, proliferation, tumor occurrence, and progression (Leroi et al., 2003). Tumorigenesis is often associated with alterations in the tumor microenvironment. Alterations of CDGs may influence the tumor microenvironment and affect the response to immunotherapy. Studying CDGs offers a chance to develop accurate biomarkers for tumor prognosis and make decision on the therapeutic strategy.

In the current study, we systematically profiled the expression characteristics and mutation landscape of the 568 CDGs in HCC. We constructed a prognosis score based on seven CDGs, which could predict the survival of HCC patients and was validated to be an independent prognostic factor in different HCC cohorts. Our study underlined the importance of CDGs in HCC and provided a strategy for patient stratification for precise medication.

Materials and methods

Data collection

A total of 424 pieces of RNA-seq data of TCGA-LIHC were downloaded from the Cancer Genome Atlas (TCGA) database as a training set, which consists of 374 HCC samples and 50 controls. A total of 230 HCC patient cases with gene expression and complete clinical information from the International Cancer Genome Consortium (ICGC) dataset were used as a validation set. The gene mutation and clinical data of TCGA-LIHC were also downloaded from the TCGA database.

Differentially expressed cancer driver gene identification

Differentially expressed CDGs between cancer and normal tissues were identified by Wilcoxon test with |log2 fold change (FC)| ≥1 and FDR (false discovery rate) < 0.05. The mutation pattern of differentially expressed genes was analyzed by the Maftools R package.

Functional enrichment analyses

In order to explore the functions and signaling pathways of the differently expressed CDGs, the "clusterProfiler" R package was used to perform the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO) enrichment analyses, with $p < 0.05$ and FDR < 0.05 being used as significance thresholds.

Construction of gene signature

Univariate Cox regression analysis was performed to find CDGs significantly related to survival. Least absolute shrinkage and selection operator (LASSO) regression analysis was performed to further screen prognostic-related CDGs. Finally, the stepwise multivariate COX regression analysis was performed to find the optimal key prognostic-related CDGs and obtained standardized regression coefficients. The risk score of each patient was calculated by the following formula: Risk score = Expression of gene1 × Coefficient of gene1 + Expression of gene2 × Coefficient of gene2 + ... Expression of geneN × Coefficient of geneN. Patients were assigned into low-risk groups and high-risk groups by the median value of the risk score, and the Kaplan–Meier curve was plotted by the "Survival" R package. Moreover, the receiver operating characteristic (ROC) curves were generated by the "SurvivalROC" R package to determine the accuracy of the gene signature.

Gene signature validation by the ICGC database

The 230 HCC patient cases from the ICGC database were utilized as an external validation dataset. The risk score of each sample was calculated by the formula presented above. The Kaplan–Meier curve and the ROC curve were plotted as described above.

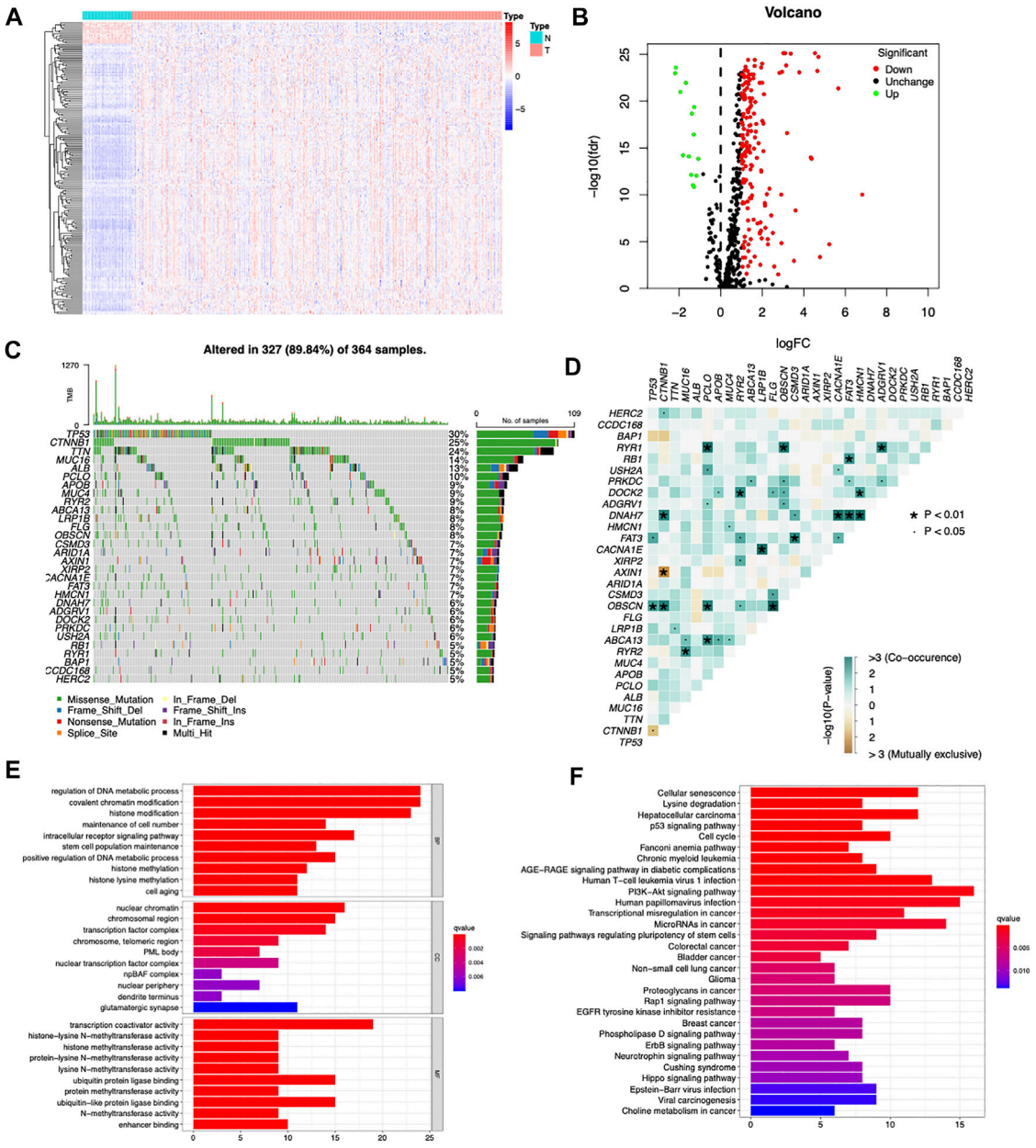


FIGURE 1
Searching differentially expressed CDGs in the TCGA cohort. **(A)** Heatmap of differentially expressed CDGs between HCC and normal tissues. **(B)** Volcano plot of differentially expressed CDGs between HCC and normal tissues. **(C)** Mutation spectrum of the top 30 most frequently mutated DE CDGs in the TCGA-LIHC cohort, with each column representing one patient and the percentage on the right side representing the corresponding gene mutation rate. **(D)** Interactions and crosstalk between mutated genes in the TCGA-LIHC cohort. **(E)** GO enrichment analysis of the upregulated DE CDGs. **(F)** KEGG enrichment analysis of the upregulated DE CDGs.

Functional biological, immune infiltration, and tumor stemness analysis

Gene set enrichment analysis (GSEA) was performed using the gene set "c2. cp.kegg.v7.5. symbols.gmt" and "c5. go.bp.v7.5. symbols.gmt" (Subramanian et al., 2005). The "GSVA" package was used to compare the immune-related pathways between the two subgroups (Hanzelmann et al., 2013). CIBERSORT was conducted to calculate the immune cell fraction among the samples in the TCGA-LIHC cohort (Newman et al., 2015). Tumor stemness was reported to be capable of evaluation by RNA stemness score (RNAss) based on mRNA expression

(Malta et al., 2018). Correlation between risk score and RNAss was analyzed using Spearman rank-based testing. The R code used in this study is available in the supplementary file.

Results

Mutation landscape and functional analysis of the differentially expressed CDGs in HCC

A total of 568 human CDGs obtained from the somatic mutations of more than 28,000 tumors of 66 cancer types were included in our study (Supplementary Table S1). By utilizing Wilcoxon test, 189 differentially expressed CDGs (DE CDGs) were identified, including 175 upregulated and 14 downregulated CDGs, according to the |log2 FC| ≥1 and FDR <0.05 (Figures 1A,B). Of note, most of the DE CDGs were highly expressed in the cancer tissues than in normal tissues.

Among the 189 differentially expressed CDGs, TP53 had the highest mutation rate (30%) followed by CTNNB1 (25%), TTN (24%), MUC16 (14%), ALB (13%), and PCLO (10%) (Figure 1C). Genetic interaction analysis identified that mutations of TP53 were positively correlated with mutations of FAT3 and OBSCN and negatively correlated with CTNNB1. Mutations of CTNNB1 were positively correlated with mutations of HERC2, DNAH7, and OBSCN and negatively correlated with AXIN1 (Figure 1D). These results indicated that CDGs had high mutation rates in HCC, and there were huge interactions and crosstalk between these mutations.

GO and KEGG pathway enrichment analyses were carried out to explore the functions and signaling pathways of the DE CDGs. GO enrichment analysis showed that the upregulated DE CDGs were mainly associated with "regulation of DNA metabolic process," "covalent chromatin modification," "histone modification," and "histone methylation" of the biological process (BP) category. In the cellular component (CC) category, "nuclear chromatin," "chromosomal region," and "transcription factor complex" were significantly enriched. The molecular function (MF) term mainly included "transcription coactivator activity," "ubiquitin protein ligase binding," and "ubiquitin-like protein ligase binding" (Figure 1E). KEGG pathway enrichment analysis showed that upregulated DE CDGs were highly enriched in "PI3K-Akt signaling pathway," "hepatocellular carcinoma," "cell cycle," and so on (Figure 1F). These results suggested that DE CDGs in the HCC play an important role in the key process in cancer development, such as epigenetic modification and various signaling pathways. These aberrantly expressed genes may lead to carcinogenesis and progression of HCC.

Selection of prognosis-related CDGs

Among the 189 DE CDGs, 96 genes were found to be associated with the patients' overall survival by univariate Cox regression analysis (Supplementary Table S2). Then, the LASSO regression analysis was conducted to further narrow the survival-related CDGs (Figures 2A,B). Finally, seven genes were identified by the stepwise multivariate regression analysis and subsequently used to construct a prognostic gene signature (Figure 2C). The seven genes identified were CDKN2C, HRAS, IRAK1, LOX, MYCN, NRAS, and PABPC1. The risk score = 0.1758* Expression of CDKN2C + 0.2975* Expression of HRAS +0.1934 * Expression of IRAK1 + 0.1943 * Expression of LOX + 0.3361 * Expression of MYCN + 0.3782 * Expression of NRAS + 0.2048 * Expression of PABPC1 (Supplementary Table S3).

To assess the prognostic capacity of the seven-gene signature, we calculated the risk score for each patient and classified the patients into low- and high-risk groups based on the medium risk score value. In the TCGA dataset, the overall survival rate of patients in the low-risk group was markedly higher than that of the high-risk group (p = 4.033e-5) (Figure 2D). As the risk score increased, the patients had a shorter survival time (Figures 2E,F). The area under the time-dependent ROC curves at 1-, 2-, and 3-year survival was 0.75, 0.701, and 0.684, respectively (Figure 2G). Taken together, our results suggested that the risk scores based on the seven CDGs had optimal prediction ability of the prognosis of HCC patients.

Validation of the seven-gene signature in the ICGC dataset

The ICGC database was used as a validation cohort to verify the accuracy of the seven-gene signature. Consistent with the TCGA dataset, high-risk score patients exhibited a significantly worse outcome (p = 3.581e-4) (Figure 3A). Patients with high-risk scores had more mortality and shorter survival time (Figures 3B,C). The time-dependent ROC curves suggested that the AUC at 1-, 2-, and 3-year survival was 0.765, 0.745, and 0.719, respectively (Figure 3D). These results suggested good accuracy and stability of our prognostic signature.

Independent prognostic role of the prognostic signature

The clinical information including age, gender, and stage was included for further analysis both in the TCGA and ICGC cohorts. In the TCGA dataset, the risk score was independently associated with the survival of the patients, with a hazard ratio of 1.366 in the univariate analysis and 1.338 in the multivariate Cox regression (Figures 4A,B). In the ICGC cohort, the hazard ratio of the risk score was 1.649 and 95% confidence interval (CI) was 1.206–2.265 (p = 0.002) in the univariate Cox regression and 1.653 (1.206–2.656) in the multivariate regression (Figures 4D,E). The AUC of risk score was 0.753 in the TCGA cohort and 0.765 in the ICGC cohort (Figures 4C,F), which exceeded stage and other clinical features.

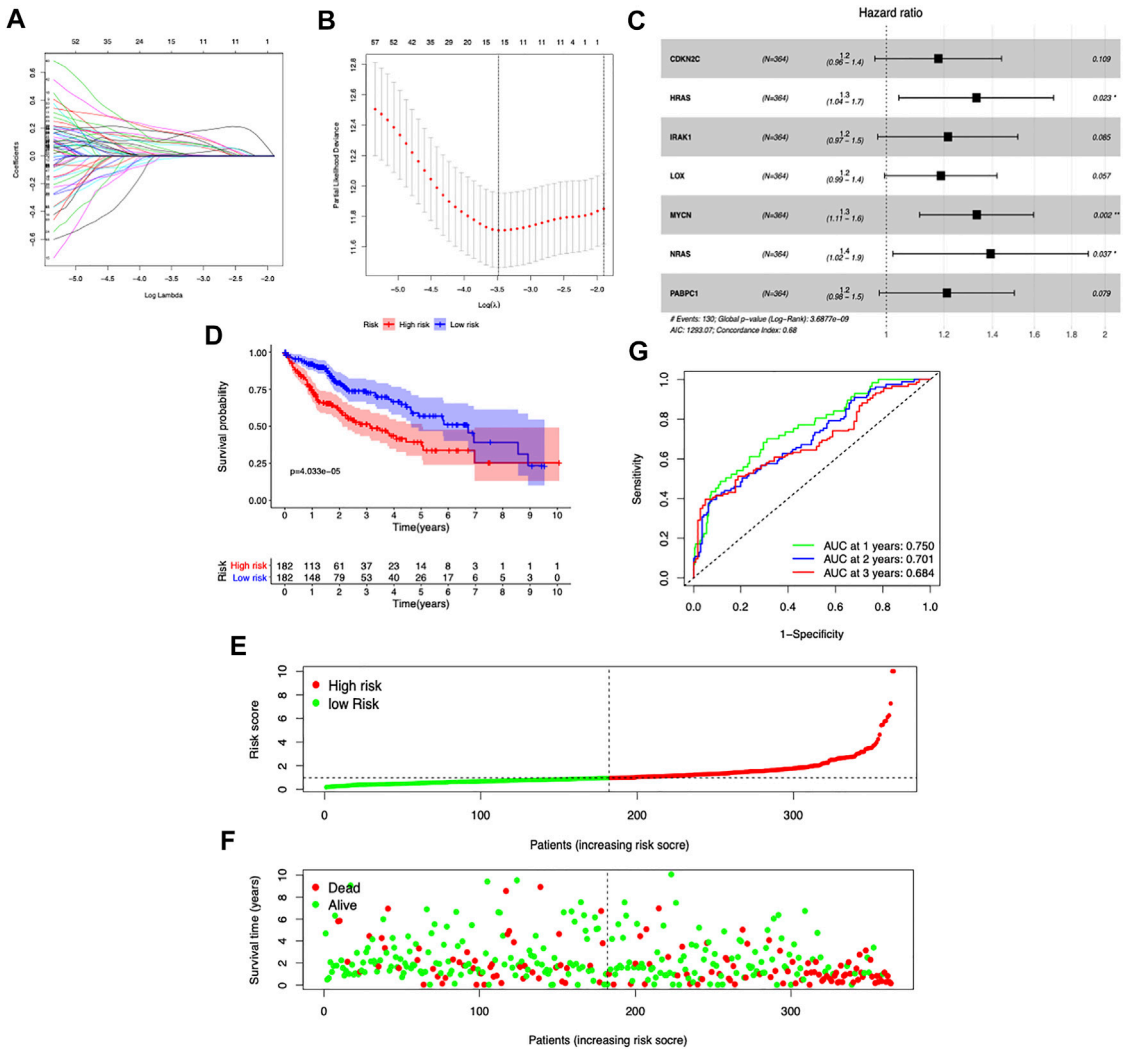


FIGURE 2
Construction of the prognostic risk model based on seven CDGs. **(A,B)** LASSO Cox regression analysis of the selection of CDGs. **(C)** Forest plot of the seven genes that construct the risk signature. **(D)** Survival curves stratified by the risk score in the TCGA cohort. **(E,F)** Distribution of risk score, survival time, and survival statuses in HCC patients. **(G)** Receiver operating characteristic (ROC) curves of risk model for predicting survival in the TGGA-LIHC cohort.

Different cancer hallmarks and tumor microenvironments between two risk score groups

GSEA was performed to explore the underlying biological processes related to the risk score signature. We found most cancer hallmarks, including the VEGF signaling pathway, cell cycle, DNA replication, ERBB signaling pathway, double strand break repair, positive regulation of intracellular transport, and regulation of mitotic cell cycle, were significantly enriched in patients in the high-risk score group (Figures 5A,B). However, immune-related pathways including Type_1_IFN_Response and Type_2_IFN_Response were significantly overexpressed in the low-risk score group (Figure 5C). Moreover, the infiltrating immune cell subtypes were significantly different between the high-risk group and the low-risk group, with more CD8+T cells and less M0 macrophages accumulating in low-risk score tumors (Figure 5D). Spearman's correlation analysis showed that IRAK1 and PABPC1 expression was positively correlated with M0 macrophages and negatively correlated with CD8+T cells. CDKN2C showed negative correlation with Tregs, CD4 memory resting T cells, and activated dendritic cells. HRAS was positively correlated with M2 macrophages and activated NK cells (Figure 5E). The correlation between risk score with tumor stemness measured by RNAss was explored. The results showed that the risk score was significantly

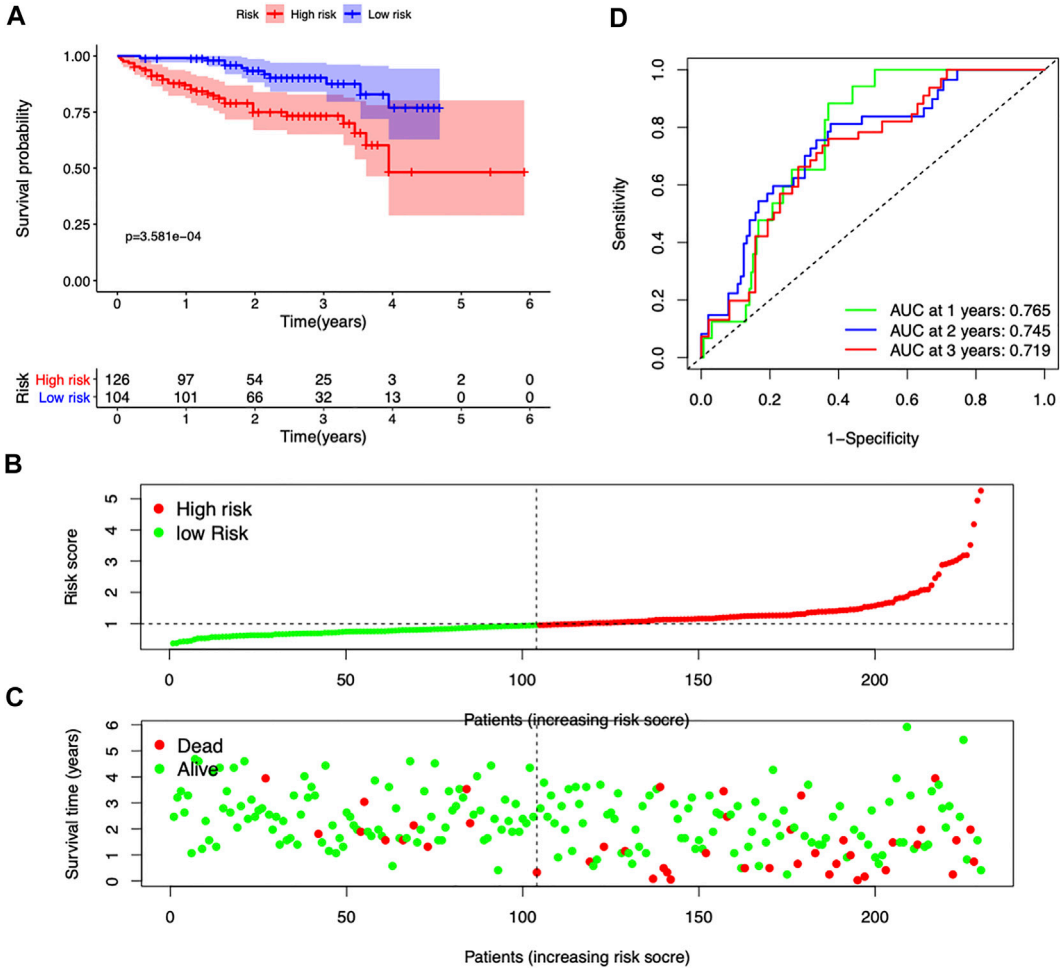


FIGURE 3
Validation of the seven-gene risk model in the ICGC dataset. **(A)** Survival curves stratified by the risk score in the ICGC cohort. **(B,C)** Distribution of risk score, survival time, and survival statuses in the ICGC cohort. **(D)** ROC curves of risk model for predicting survival in the ICGC cohort.

positively correlated with stemness score ($R = 0.25$, $p = 1.1e-06$, Figure 5F).

These results together implied that the longer OS of the low-risk score group might be attributed to an inflamed tumor microenvironment with more infiltrated $CD8^+$ T cells and less M0 macrophages, while the poor prognosis of the high-risk score group might be associated with the tumorigenesis of cancer hallmarks.

A predictive nomogram development and validation

To facilitate the clinical applicability and availability of the seven-gene signature, a predictive nomogram for 1-, 2-, 3-year OS combined with age, gender, stage, and risk scores was developed (Figure 6A). The calibration curves showed that the nomogram had good prediction performance in HCC patients (Figure 6B). In the TCGA dataset, patients with a low nomogram score had significantly better survival than patients with a high nomogram score ($p = 1.225e-07$) (Figure 6C). The AUCs of the nomogram in the 1-, 2-, and 3-year ROC curves were 0.777, 0.725, and 0.751, which outperformed the seven-gene signature (Figure 6E). The nomogram was also validated in the ICGC dataset, and patients with a low score had a significantly better survival rate than those with a high score ($p = 1.178e-02$) (Figure 6D). The AUCs of the nomogram in the 1-, 2-, and 3-year ROC curves in the ICGC cohort were 0.874, 0.767, and 0.741, respectively (Figure 6F).

Discussion

Nowadays, immune checkpoints have greatly changed the treatment paradigm and increased the survival of HCC patients

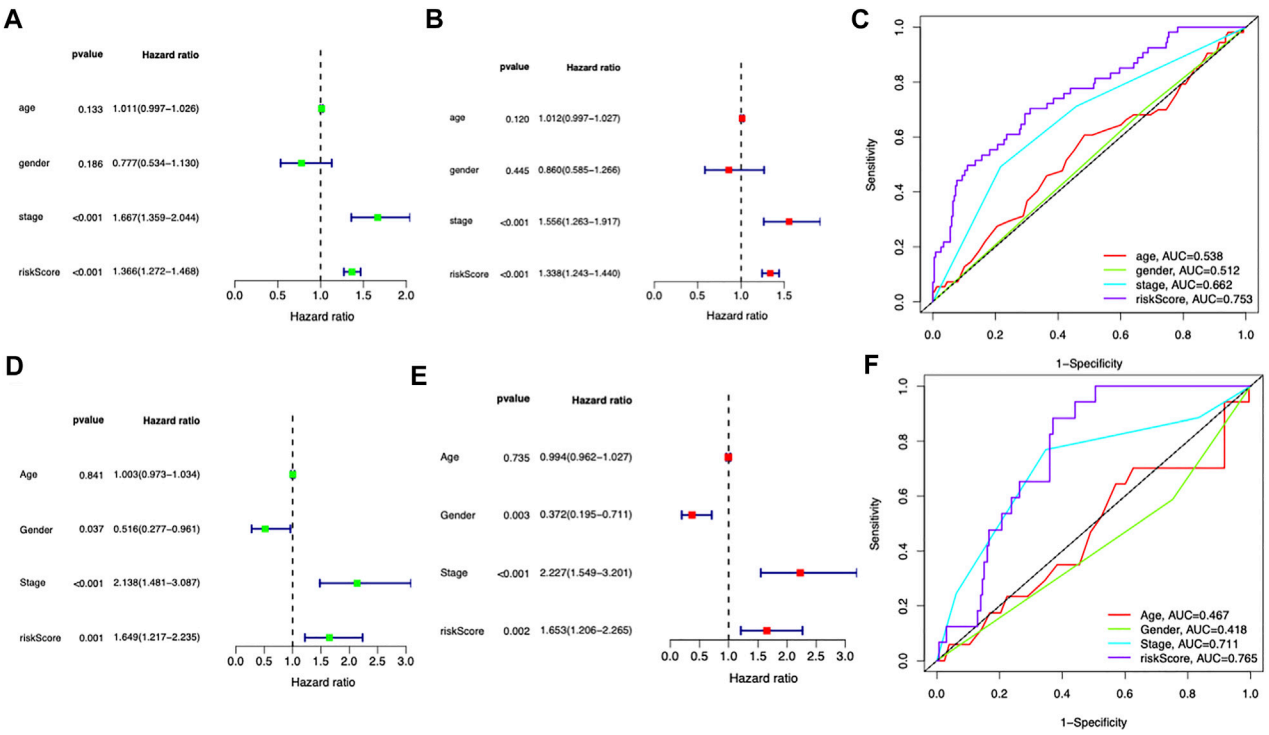


FIGURE 4
Independent prognostic role of the prognostic signature. **(A,B)** Univariate and multivariate Cox analyses for the prognostic model and other clinical features in the TCGA cohort. **(C)** ROC analysis of the risk score and other clinical features in the TCGA cohort. **(D,E)** Univariate and multivariate Cox analyses for the prognostic model and other clinical features in the ICGC cohort. **(F)** ROC analysis of the risk score and other clinical features in the ICGC cohort.

(Nguyen et al., 2015). Traditional tumor markers including clinical tumor-node-metastasis (TNM) staging, vascular invasion, and other parameters help predict HCC prognosis (Bruix et al., 2016); however, these are gradually showing their limitations with the evolution of HCC management. Developing novel prediction models could guide patient prognostic stratification and facilitate personalized therapy.

In the present study, we systematically analyzed the expression of 568 cancer driver genes in the TCGA-LIHC cohort and found 189 differentially expressed cancer driver genes between cancer and normal tissues. We found that differentially expressed CDGs had high mutation rates in HCC, with TP53 showing the highest mutation rate (30%) followed by CTNNB1 (25%), TTN (24%), MUC16 (14%), ALB (13%), and PCLO (10%). These results indicated that the heterogeneity of HCC may be due to the diverse genetic abnormalities of cancer cells.

By univariate Cox regression analysis, 96 genes were found to be associated with the survival of HCC patients. LASSO regression and stepwise multivariate regression analyses found that a novel prognostic model comprising seven cancer driver genes was able to accurately distinguish HCC patients with different prognosis. Finally, a nomogram containing the clinical characteristics and genetic factors was constructed to provide a more accurate measure to predict the prognosis of HCC.

The prognostic signature that we constructed consisted of seven cancer driver genes (CDKN2C, HRAS, IRAK1, LOX, MYCN, NRAS, and PABPC1). These genes were all upregulated in the HCC tissues compared to normal tissues in the TCGA cohort. CDKN2C (cyclin-dependent kinase inhibitor 2C), also known as p18INK4C, is considered a tumor-suppressor gene (Gagrica et al., 2012). Dysregulated CDKN2C and its protease activity change are associated with the prognosis of HCC (Morishita et al., 2004). In human teratoma and thyroid tumor, mutant CDKN2C has been proven to predict poor prognosis (Cooke et al., 2017; El Naofal et al., 2017). Harvey-RAS (HRAS) and neuroblastoma-RAS (NRAS) belong to the RAS oncogene family. Sorafenib and regorafenib, the only effective therapeutic strategies for advanced HCC, target multiple kinase-related pathways including the RAS-RAF-ERK-pathway, underlining the crucial role of RAS signaling in HCC (Ostrem and Shokat, 2016; Pascual et al., 2016; Bruix et al., 2017). NRAS overexpression is demonstrated to be correlated with poor survival and sorafenib resistance in HCC (Dietrich et al., 2019). HRAS was also proven to be associated with the prognosis of HCC (Dietrich et al., 2018). IRAK1 is a widely expressed serine/threonine kinase,

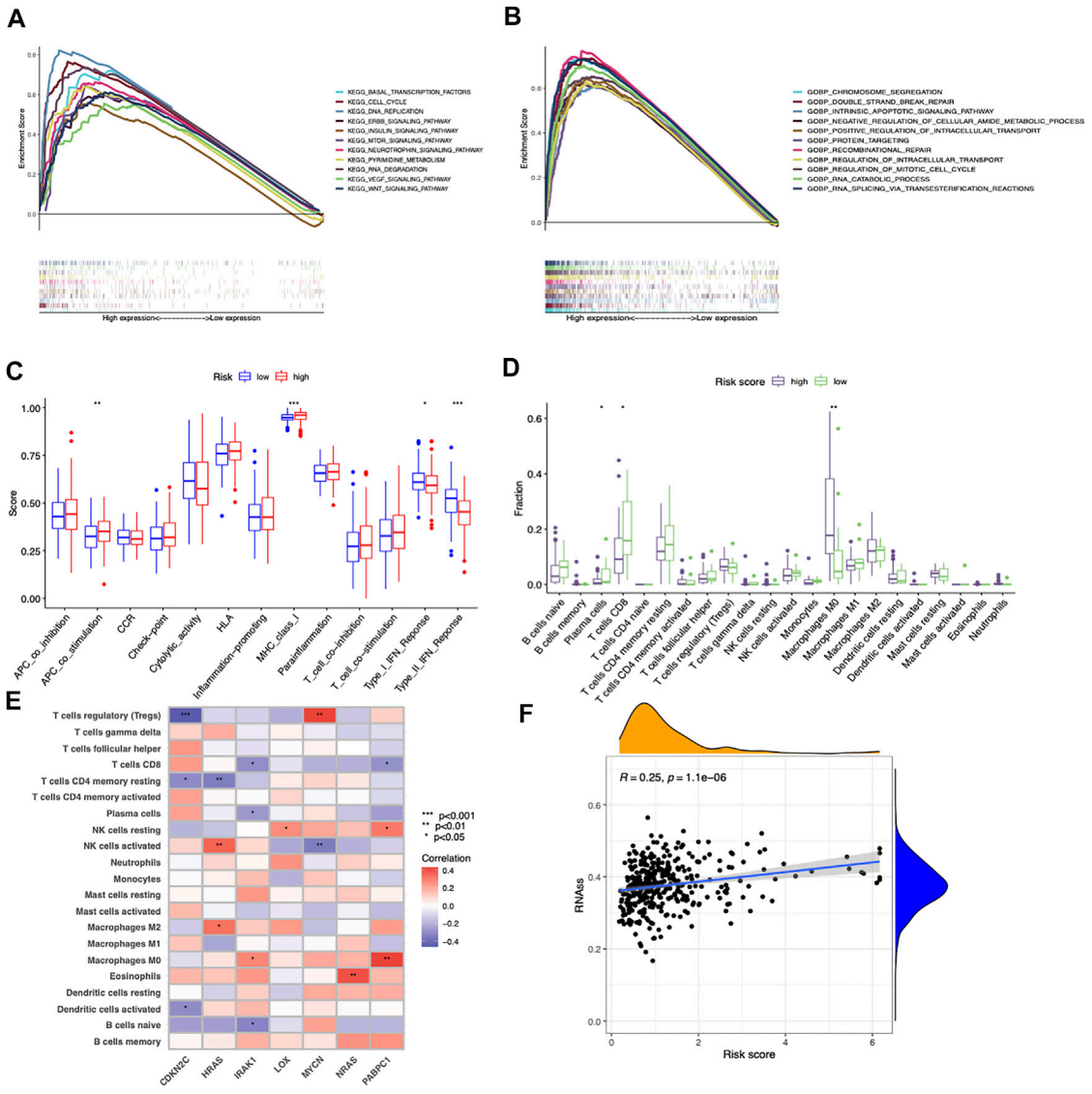


FIGURE 5
Different cancer hallmarks and tumor microenvironments between two risk score groups. **(A,B)** GSEA of significantly enriched pathways in the high-risk group based on KEGG and GO (biological process) gene sets. **(C)** Differences in the proportions of immune-related pathways between the low- and high-risk groups. **(D)** Differences in the immune cells infiltrated in the tumor microenvironment between the low- and high-risk groups. **(E)** Correlations between infiltrated immune cells and the seven prognostic-related cancer driver genes. **(F)** Correlation between risk score and cancer stemness score (RNAss) based on Spearman's correlation tests.

and phosphorylation of IRAK1 binds to the E3 ubiquitin ligase and TRAF6, leading to the activation of the NF-κB and MAPK pathways (Zhang and Ghosh, 2001; Flannery and Bowie, 2010). IRAK1 overexpression was proven to be correlated with metastasis and poor prognosis of HCC (Ye et al., 2017). LOX is a copper-dependent amine oxidase that plays an important role in the formation of collagen and extracellular matrix (Erler et al., 2006). It is reported to be involved in the remodeling of cancer stroma and correlated to metastasis and dedifferentiation of cancer cells (Semenza, 2012; Boufraqech et al., 2016; Nilsson and Kannius-Janson, 2016). LOX overexpression can predict early recurrence and poor prognosis of HCC (Umezaki et al., 2019). MYCN, one of the members of the MYC family, plays crucial roles in regulating normal stem cell–mediated tissue regeneration and stem cell–mediated tumorigenesis (Dang, 2012; Qin et al., 2017). MYCN has been proven to be a prognostic biomarker and positively correlated with recurrence of *de novo* HCC after curative treatment (Qin et al., 2018). PABPC1 plays crucial roles in poly(A) shortening, recruitment of ribosome, and translation initiation *via* specifically binding to poly(A) tail of

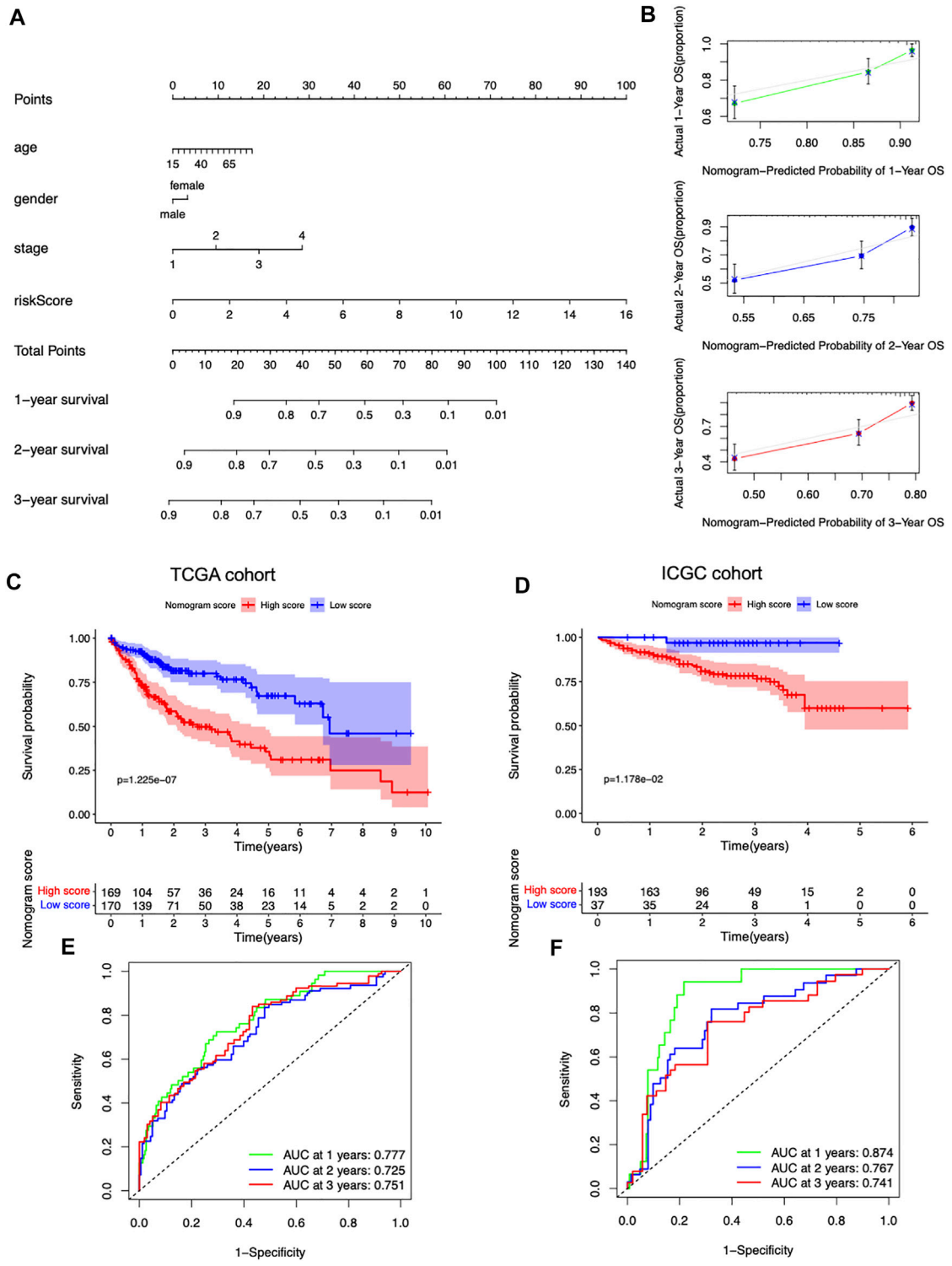


FIGURE 6
Predictive nomogram development and validation. **(A)** Nomogram based on the risk score and the clinical factors. **(B)** 1-, 2-, and 3-year calibration plots of the nomogram. **(C)** Survival curves stratified by the nomogram score in the TCGA cohort. **(D)** Survival curves stratified by the nomogram score in the ICGC cohort. **(E)** ROC analysis of the nomogram in the TCGA cohort. **(F)** ROC analysis of the nomogram in the ICGC cohort.

mRNA in cytoplasm (Kuhn and Wahle, 2004). High expression of PABPC1 has been proven to be correlated with worse overall survival for HCC (YuFeng and Ming, 2020). Taken together, these seven genes are closely related to the development and progression of cancer.

Immunotherapy using immune checkpoint inhibitors (ICIs) has dramatically changed the treatment of various malignancies. Anti-PD-1/L1 therapies, such as atezolizumab, pembrolizumab, and nivolumab, have shown promising benefits in a subset of HCC patients, alone or in combination with other agents (El-Khoueiry et al., 2017; Zhu et al., 2018; Finn et al., 2020a; Finn et al., 2020b). Screening out potential patients who may benefit from immunotherapy is the focus of research. In our study, the low-risk score group showed elevated expression of Type_1_IFN_Response and Type_2_IFN_Response signaling pathways compared to its counterpart. Moreover, it had more CD8+T cells and less M0 macrophages infiltrated. These results implied that the low-risk group might have a better immune microenvironment and responded better to immune checkpoint inhibitors. A nomogram consisting of the risk score and several clinical factors was constructed, which showed great accuracy to predict the prognosis of HCC patients.

There are some limitations in our study. First, the risk signature was built based on the TCGA-LIHC dataset and was only validated in the ICGC HCC dataset. Larger cohorts containing more patients are needed to verify the prognostic value of the risk score signature and nomogram. Second, the potential biological functions of genes contained in the risk signature have not been investigated. Further research should be conducted to elucidate the relevant mechanisms. Third, most HCC patients in the TCGA database were Caucasian, and it is not clear whether the risk signature has the same predictive effect in non-Caucasian races. However, our study provided an insight into the mutation landscape and expression pattern of the CDGs and constructed a risk score model and nomogram for prognosis prediction. This study highlighted the significance of CDGs in the HCC and provided a novel horizon for the investigation of HCC 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.

Author contributions

JZ wrote the manuscript. WQ designed the study, analyzed the data, and modified the manuscript. Both authors have read and approved the manuscript for publication.

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/fgene.2022.937948/full#supplementary-material

References

Boufraqech, M., Zhang, L., Nilubol, N., Sadowski, S. M., Kotian, S., Quezado, M., et al. (2016). Lysyl oxidase (LOX) transcriptionally regulates SNAI2 expression and TIMP4 secretion in human cancers. *Clin. Cancer Res.* 22 (17), 4491–4504. doi:10.1158/1078-0432.CCR-15-2461

Brown, A. L., Li, M., Goncearenco, A., and Panchenko, A. R. (2019). Finding driver mutations in cancer: Elucidating the role of background mutational processes. *PLoS Comput. Biol.* 15 (4), e1006981. doi:10.1371/journal.pcbi.1006981

Bruix, J., Gores, G. J., and Mazzaferro, V. (2014). Hepatocellular carcinoma: Clinical frontiers and perspectives. *Gut* 63 (5), 844–855. doi:10.1136/gutjnl-2013-306627

Bruix, J., Qin, S., Merle, P., Granito, A., Huang, Y. H., Bodoky, G., et al. (2017). Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): A randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* 389 (10064), 56–66. doi:10.1016/S0140-6736(16)32453-9

Bruix, J., Reig, M., and Sherman, M. (2016). Evidence-based diagnosis, staging, and treatment of patients with hepatocellular carcinoma. *Gastroenterology* 150 (4), 835–853. doi:10.1053/j.gastro.2015.12.041

Cooke, S. L., Ennis, D., Evers, L., Dowson, S., Chan, M. Y., Paul, J., et al. (2017). The driver mutational landscape of ovarian squamous cell carcinomas arising in mature cystic teratoma. *Clin. Cancer Res.* 23 (24), 7633–7640. doi:10.1158/1078-0432.CCR-17-1789

Dang, C. V. (2012). MYC on the path to cancer. *Cell* 149 (1), 22–35. doi:10.1016/j.cell.2012.03.003

Dietrich, P., Freese, K., Mahli, A., Thasler, W. E., Hellerbrand, C., Bosserhoff, A. K., et al. (2018). Combined effects of PLK1 and RAS in hepatocellular carcinoma reveal rigosertib as promising novel therapeutic "dual-hit" option. *Oncotarget* 9 (3), 3605–3618. doi:10.18632/oncotarget.23188

Dietrich, P., Gaza, A., Wormser, L., Fritz, V., Hellerbrand, C., Bosserhoff, A. K., et al. (2019). Neuroblastoma RAS viral oncogene homolog (NRAS) is a novel prognostic marker and contributes to sorafenib resistance in hepatocellular carcinoma. *Neoplasia* 21 (3), 257–268. doi:10.1016/j.neo.2018.11.011

El Naofal, M., Kim, A., Yon, H. Y., Baity, M., Ming, Z., Bui-Griffith, J., et al. (2017). Role of CDKN2C fluorescence *in situ* hybridization in the management of medullary thyroid carcinoma. *Ann. Clin. Lab. Sci.* 47 (5), 523–528.

El-Khoueiry, A. B., Sangro, B., Yau, T., Crocenzi, T. S., Kudo, M., Hsu, C., et al. (2017). Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): An open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *Lancet* 389 (10088), 2492–2502. doi:10.1016/S0140-6736(17)31046-2

El-Serag, H. B. (2011). Hepatocellular carcinoma. *N. Engl. J. Med.* 365 (12), 1118–1127. doi:10.1056/NEJMra1001683

Erler, J. T., Bennewith, K. L., Nicolau, M., Dornhofer, N., Kong, C., Le, Q. T., et al. (2006). Lysyl oxidase is essential for hypoxia-induced metastasis. *Nature* 440 (7088), 1222–1226. doi:10.1038/nature04695

Finn, R. S., Qin, S., Ikeda, M., Galle, P. R., Ducreux, M., Kim, T. Y., et al. (2020). Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *N. Engl. J. Med.* 382 (20), 1894–1905. doi:10.1056/NEJMoa1915745

Finn, R. S., Ryoo, B. Y., Merle, P., Kudo, M., Bouattour, M., Lim, H. Y., et al. (2020). Pembrolizumab as second-line therapy in patients with advanced hepatocellular carcinoma in KEYNOTE-240: A randomized, double-blind, phase III trial. *J. Clin. Oncol.* 38 (3), 193–202. doi:10.1200/JCO.19.01307

Flannery, S., and Bowie, A. G. (2010). The interleukin-1 receptor-associated kinases: Critical regulators of innate immune signalling. *Biochem. Pharmacol.* 80 (12), 1981–1991. doi:10.1016/j.bcp.2010.06.020

Gagrica, S., Brookes, S., Anderton, E., Rowe, J., and Peters, G. (2012). Contrasting behavior of the p18INK4c and p16INK4a tumor suppressors in both replicative and oncogene-induced senescence. *Cancer Res.* 72 (1), 165–175. doi:10.1158/0008-5472.CAN-11-2552

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell* 144 (5), 646–674. doi:10.1016/j.cell.2011.02.013

Hanahan, D., and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell* 100 (1), 57–70. doi:10.1016/s0092-8674(00)81683-9

Hanzelmann, S., Castelo, R., and Guinney, J. (2013). Gsva: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14, 7. doi:10.1186/1471-2105-14-7

Kimura, T., Kato, Y., Ozawa, Y., Kodama, K., Ito, J., Ichikawa, K., et al. (2018). Immunomodulatory activity of lenvatinib contributes to antitumor activity in the Hepa1-6 hepatocellular carcinoma model. *Cancer Sci.* 109 (12), 3993–4002. doi:10.1111/cas.13806

Kudo, M., Finn, R. S., Qin, S., Han, K. H., Ikeda, K., Piscaglia, F., et al. (2018). Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: A randomised phase 3 non-inferiority trial. *Lancet* 391 (10126), 1163–1173. doi:10.1016/S0140-6736(18)30207-1

Kuhn, U., and Wahle, E. (2004). Structure and function of poly(A) binding proteins. *Biochim. Biophys. Acta* 1678 (2-3), 67–84. doi:10.1016/j.bbaexp.2004.03.008

Leroi, A. M., Koufopanou, V., and Burt, A. (2003). Cancer selection. *Nat. Rev. Cancer* 3 (3), 226–231. doi:10.1038/nrc1016

Malta, T. M., Sokolov, A., Gentles, A. J., Burzykowski, T., Poisson, L., Weinstein, J. N., et al. (2018). Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* 173 (2), 338–354. doi:10.1016/j.cell.2018.03.034

Martinez-Jimenez, F., Muinos, F., Sentis, I., Deu-Pons, J., Reyes-Salazar, I., Arnedo-Pac, C., et al. (2020). A compendium of mutational cancer driver genes. *Nat. Rev. Cancer* 20, 555–572. doi:10.1038/s41568-020-0290-x

Morishita, A., Masaki, T., Yoshiji, H., Nakai, S., Ogi, T., Miyauchi, Y., et al. (2004). Reduced expression of cell cycle regulator p18(INK4C) in human hepatocellular carcinoma. *Hepatology* 40 (3), 677–686. doi:10.1002/hep.20337

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Nguyen, K., Jack, K., and Sun, W. (2015). Hepatocellular carcinoma: Past and future of molecular target therapy. *Diseases* 4 (1). doi:10.3390/diseases4010001

Nilsson, G., and Kannius-Janson, M. (2016). Forkhead Box F1 promotes breast cancer cell migration by upregulating lysyl oxidase and suppressing Smad2/3 signaling. *BMC Cancer* 16, 142. doi:10.1186/s12885-016-2196-2

Ostrem, J. M., and Shokat, K. M. (2016). Direct small-molecule inhibitors of KRAS: From structural insights to mechanism-based design. *Nat. Rev. Drug Discov.* 15 (11), 771–785. doi:10.1038/nrd.2016.139

Pascual, S., Herrera, I., and Irurzun, J. (2016). New advances in hepatocellular carcinoma. *World J. Hepatol.* 8 (9), 421–438. doi:10.4254/wjh.v8.i9.421

Qin, X. Y., Hara, M., Arner, E., Kawaguchi, Y., Inoue, I., Tatsukawa, H., et al. (2017). Transcriptome analysis uncovers a growth-promoting activity of orosomucoid-1 on hepatocytes. *EBioMedicine* 24, 257–266. doi:10.1016/j.ebiom.2017.09.008

Qin, X. Y., Suzuki, H., Honda, M., Okada, H., Kaneko, S., Inoue, I., et al. (2018). Prevention of hepatocellular carcinoma by targeting MYCN-positive liver cancer stem cells with acyclic retinoid. *Proc. Natl. Acad. Sci. U. S. A.* 115 (19), 4969–4974. doi:10.1073/pnas.1802279115

Semenza, G. L. (2012). Molecular mechanisms mediating metastasis of hypoxic breast cancer cells. *Trends Mol. Med.* 18 (9), 534–543. doi:10.1016/j.molmed.2012.08.001

Shen, Y. C., Hsu, C., and Cheng, A. L. (2010). Molecular targeted therapy for advanced hepatocellular carcinoma: Current status and future perspectives. *J. Gastroenterol.* 45 (8), 794–807. doi:10.1007/s00535-010-0270-0

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Umezaki, N., Nakagawa, S., Yamashita, Y. I., Kitano, Y., Arima, K., Miyata, T., et al. (2019). Lysyl oxidase induces epithelial-mesenchymal transition and predicts intrahepatic metastasis of hepatocellular carcinoma. *Cancer Sci.* 110 (6), 2033–2043. doi:10.1111/cas.14010

Ye, Z. H., Gao, L., Wen, D. Y., He, Y., Pang, Y. Y., Chen, G., et al. (2017). Diagnostic and prognostic roles of IRAK1 in hepatocellular carcinoma tissues: An analysis of immunohistochemistry and RNA-sequencing data from the cancer genome atlas. *Onco. Targets. Ther.* 10, 1711–1723. doi:10.2147/OTT.S132120

YuFeng, Z., and Ming, Q. (2020). Expression and prognostic roles of PABPC1 in hepatocellular carcinoma. *Int. J. Surg.* 84, 3–12. doi:10.1016/j.ijsu.2020.10.004

Zhang, G., and Ghosh, S. (2001). Toll-like receptor-mediated NF-kappaB activation: A phylogenetically conserved paradigm in innate immunity. *J. Clin. Invest.* 107 (1), 13–19. doi:10.1172/JCI11837

Zhu, A. X., Finn, R. S., Edeline, J., Cattan, S., Ogasawara, S., Palmer, D., et al. (2018). Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): A non-randomised, open-label phase 2 trial. *Lancet. Oncol.* 19 (7), 940–952. doi:10.1016/S1470-2045(18)30351-6

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 29 August 2022
DOI 10.3389/fgene.2022.957655

Check for updates

OPEN ACCESS

EDITED BY
Ehsan Nazemalhosseini-Mojarad,
Shahid Beheshti University of Medical Sciences, Iran

REVIEWED BY
RuiJi Liu,
Zhongda Hospital, Southeast University, China
Zhipeng Xu,
Southeast University, China
Qingjia Chi,
Wuhan University of Technology, China

*CORRESPONDENCE
Hao Chen,
ery_chenh@lzu.edu.cn

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

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 31 May 2022
ACCEPTED 25 July 2022
PUBLISHED 29 August 2022

CITATION
Wang B, Zou D, Wang N, Wang H, Zhang T, Gao L, Ma C, Zheng P, Gu B, Li X, Wang Y, He P, Ma Y, Wang X and Chen H (2022), Construction and validation of a novel coagulation-related 7-gene prognostic signature for gastric cancer.
Front. Genet. 13:957655.
doi: 10.3389/fgene.2022.957655

COPYRIGHT
© 2022 Wang, Zou, Wang, Wang, Zhang, Gao, Ma, Zheng, Gu, Li, Wang, He, Ma, Wang and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Construction and validation of a novel coagulation-related 7-gene prognostic signature for gastric cancer

Bofang Wang[1†], Dan Zou[2†], Na Wang[1†], Haotian Wang[3†], Tao Zhang[1,4], Lei Gao[1], Chenhui Ma[1], Peng Zheng[1], Baohong Gu[1], Xuemei Li[1], Yunpeng Wang[1], Puyi He[1], Yanling Ma[1], Xueyan Wang[1] and Hao Chen[1,5,6*]

[1]Second Clinical Medical College, Lanzhou University, Lanzhou, China, [2]Chengdu Seventh People's Hospital, Chengdu, China, [3]State Key Laboratory of Genetic Resources and Evolution/Key Laboratory of Healthy Aging Research of Yunnan Province, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China, [4]Department of oncology, First Hospital of Lanzhou University, Lanzhou, China, [5]Key Laboratory of the Digestive System Tumors of Gansu Province, Lanzhou, China, [6]Department of Cancer Center, Lanzhou University Second Hospital, Lanzhou, China

Background: Gastric cancer (GC) is the most common malignant tumor. Due to the lack of practical molecular markers, the prognosis of patients with advanced gastric cancer is still poor. A number of studies have confirmed that the coagulation system is closely related to tumor progression. Therefore, the purpose of this study was to construct a coagulation-related gene signature and prognostic model for GC by bioinformatics methods.

Methods: We downloaded the gene expression and clinical data of GC patients from the TCGA and GEO databases. In total, 216 coagulation-related genes (CRGs) were obtained from AmiGO 2. Weighted gene co-expression network analysis (WGCNA) was used to identify coagulation-related genes associated with the clinical features of GC. Last absolute shrinkage and selection operator (LASSO) Cox regression was utilized to shrink the relevant predictors of the coagulation system, and a Coag-Score prognostic model was constructed based on the coefficients. According to this risk model, GC patients were divided into high-risk and low-risk groups, and overall survival (OS) curves and receiver operating characteristic (ROC) curves were drawn in the training and validation sets, respectively. We also constructed nomograms for predicting 1-, 2-, and 3-year survival in GC patients. Single-sample gene set enrichment analysis (ssGSEA) was exploited to explore immune cells' underlying mechanisms and correlations. The expression levels of coagulation-related genes were verified by real-time quantitative polymerase chain reaction (qRT-PCR) and immunohistochemistry (IHC).

Results: We identified seven CRGs employed to construct a Coag-Score risk model using WGCNA combined with LASSO regression. In both training and validation sets, GC patients in the high-risk group had worse OS than those in the low-risk group, and Coag-Score was identified as an independent predictor of OS, and the nomogram provided a quantitative method to predict the 1-, 2-, and 3-year survival rates of GC patients. Functional

analysis showed that Coag-Score was mainly related to the MAPK signaling pathway, complement and coagulation cascades, angiogenesis, epithelial–mesenchymal transition (EMT), and KRAS signaling pathway. In addition, the high-risk group had a significantly higher infiltration enrichment score and was positively associated with immune checkpoint gene expression. **Conclusion:** Coagulation-related gene models provide new insights and targets for the diagnosis, prognosis prediction, and treatment management of GC patients.

KEYWORDS

gastric cancer, coagulation-related genes, prognostic signature, weighted gene co-expression network analysis (WGCNA), bioinformatics

Introduction

Gastric cancer (GC) is the fifth most common cancer and the third leading cause of cancer death worldwide, with approximately one million new cases of GC every year, and about 784,000 patients died of GC (Smyth et al., 2020). The current detection methods are limited, resulting in a low diagnosis rate of early GC. Most patients have advanced GC when they are diagnosed. The prognosis of patients with advanced GC is abysmal, and the 5-year survival rate is less than 30% (Sung et al., 2021). With the aging of the social population, the incidence and mortality of GC continue to increase year by year, and tumor metastasis is the leading cause of high mortality. Cancer invasion and metastasis is a complex process controlled by multiple molecular determinants, which may involve activation of oncogenes, inactivation of tumor suppressor genes, and abnormalities in related signaling pathways (Hanahan and Weinberg, 2000). Biomarkers developed for key molecules play an essential role in the diagnosis, prognosis prediction, and the selection of treatment strategies for GC.

Traditional biomarkers such as CEA and CA19-9 lack sufficient specificity and sensitivity in current clinical applications. Drugs targeting Her-2 significantly prolong survival in patients with Her-2-positive GC, but their prognostic and predictive value performance remains ambiguous (Xiao and Zhou, 2017; Matsuoka and Yashiro, 2018). Therefore, the development of novel and effective GC biomarkers is necessary.

Venous thromboembolism (VTE) is often an underlying clinical symptom of cancer, and it remains one of the leading causes of cancer-related morbidity and mortality. Many studies have proved that patients with malignant tumors are in hypercoagulable and hyperfibrinolytic states, and the disorder of the coagulation system is related to tumor progression and prognosis (Repetto and De Re, 2017). The presence of tumors may strongly influence host coagulation and hemostasis systems by altering the molecular context to promote tumor cell growth, progression, and metastasis. Tumor-driven coagulation pathway activation leads to increased FGB of Fibrinopeptide A (FpA), accompanied by fibrin lysis and D-dimer (DD) release. Blockade of coagulation, fibrinolysis, and platelet activation pathways can effectively prevent tumor progression (Repetto and De Re, 2017). Numerous individual coagulation-related biomarkers revealed correlations with prognosis prediction in GC. Many laboratory data and clinical studies have shown that coagulation-related factors, such as tissue factor (TF), thrombin, plasminogen (PLG), FpA, DDs, TAFI, and thrombin–antithrombin complex, are involved in angiogenesis, tumor cell invasion, tumor progression, and metastasis (Dupuy et al., 2003; Buller et al., 2007). Coagulation-related factors are considered as diagnostic

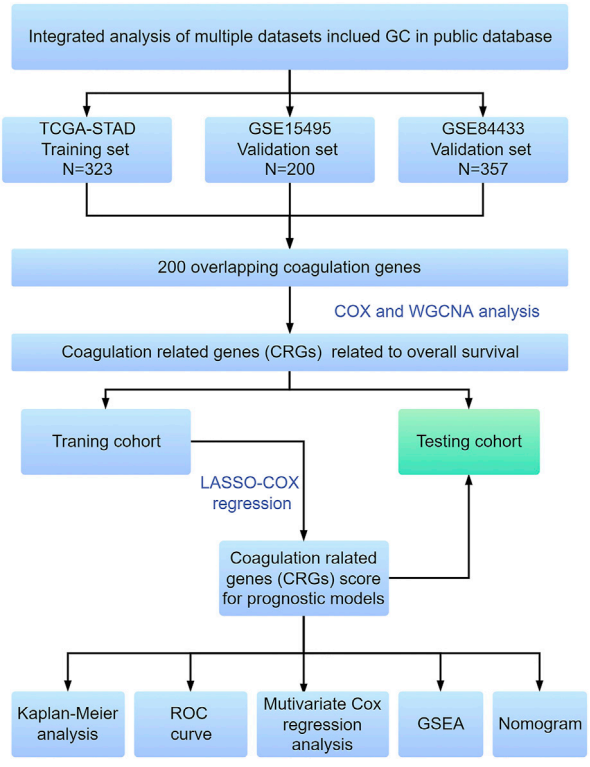


FIGURE 1
Flowchart of experimental design and main procedures.

and therapeutic evaluation tools for thrombosis in patients with GC and are also regarded as independent factors and indicators to predict the prognosis of GC (Tas et al., 2013). A retrospective clinical study showed that rivaroxaban, a coagulation factor-targeted drug, could increase the efficacy of immune checkpoint inhibitors (ICIs) by restoring host antitumor immunity (Haist et al., 2021). Therefore, we hypothesized that coagulation-related biomarkers play a crucial role in evaluating the prognosis of GC. As far as we know, there are few studies on coagulation-related genes and the prognosis of GC.

In this study, we identified coagulation-related genes (CRGs) associated with the clinical features of GC through weighted gene co-expression network analysis (WGCNA), integrated expression profiles, and clinical information from multiple datasets of TCGA and GEO databases. We constructed a risk-score model based on seven CRGs, which provided a new model for accurately predicting the prognosis and individualized treatment of GC patients. The current research workflow is shown in Figure 1.

Methods

Data acquisition

We obtained the RNA expression data and clinical data of GC samples from the UCSC Xena database (http://xena.ucsc.edu/) based on the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). The GEO database included the GSE15459 dataset and GSE84433 dataset. The TCGA-STAD dataset, GSE15459 dataset, and GSE84433 dataset included 323, 200, and 357 tumor samples, respectively. According to the annotation file provided by the platform, we used the mean to represent the expression level for a gene containing multiple probes.

Coagulation-related co-expression network construction by WGCNA

AmiGO 22 is a web-based set of tools for searching and browsing gene ontology databases (http://amigo.geneontology.org/amigo). We obtained 216 gene symbols related to coagulation from AmiGO 2. Then 200 CRGs were obtained through the comprehensive analysis of TCGA-STAD, GSE15495, and GSE84433 datasets. To calculate the Z-score of the clinical information of the TCGA-STAD dataset, we used the "WGCNA" package in the R software to perform a WGCNA on the CRGs in the TCGA-STAD dataset (Langfelder and Horvath, 2008). The operational process of WGCNA includes cluster analysis of expression profiles and calculation of associations between each cluster module and clinical phenotypes. We used the soft threshold method for the Pearson correlation analysis of the expression profiles to construct a weighted network. In this study, the hub threshold was set to 0.9 and the minimum number of modules was set to 15. A cluster dendrogram was used to display the results of gene merging and classification. Finally, we analyzed the relationship between MEs and clinical traits, and identified the relevant modules. CRGs in modules associated with clinical features were selected as candidate genes.

Coag-Score model construction

LASSO regression is a compression estimation method. By constructing a penalty function, the variable coefficients can be compressed so that the regression coefficients of some variables become 0 to achieve the purpose of screening variables (Wang et al., 2020). The "glmnet" package was used in the R software; we applied the LASSO regression method to construct a Coag-Score model of CRGs with optimal weighted coefficients. The CV curve was further drawn, the cross coefficient λ was derived using the cross-validation method, and the λ value with the best cross-validation error was selected. The regression coefficient of the best model was extracted to fit the new model. The final Coag-Score was calculated based on the expression of the gene multiplied by the corresponding regression coefficient. Based on the expression of genes in the model, PCA was carried out with the "prcomp" function of the "stats" R package.

Coag-Score model verification

First, the Coag-Score of each sample in the TCGA-STAD dataset was calculated, and the patients were categorized into high-risk and low-risk groups according to the cutoff value; the constructed Coag-Score model was verified internally. We used the "Survival" and "KMsurv" packages in R software to conduct the Kaplan–Meier (KM) survival analysis and draw survival curves. We integrated the GSE15495 and GSE84433 datasets into the GEO cohort. External validation of the Coag-Score model was performed using the GEO cohort validation dataset. Receiver operating characteristic (ROC) curves were used to evaluate the Coag-Score's accuracy and predictive power using the "survivalROC" package.

Gene set enrichment analysis

Enrichment analysis between the high-risk and low-risk groups of GC patients was performed by Gene Set Enrichment Analysis (GSEA, https://www.gsea-msigdb.org/

gsea/index.jsp) v4.3.1 software (Mootha et al., 2003; Subramanian et al., 2005). We selected the KEGG and HALLMARK gene sets for GSEA. Permutation testing (1000 permutations) was used to calculate enrichment scores (ES) and normalized enrichment scores (NES). NES with p values <0.05 and FDR <25% were considered to be significantly enriched.

Construction and assessment nomogram for GC patients

A predictive nomogram of GC patients was constructed using the TCGA-STAD training set variables. The nomogram was constructed by using the "survival" and "RMS" packages, and Harrell's Concordance Index (C-index) was used to estimate the prognostic effect of the prediction model.

Correlation of immune cell infiltration and immune checkpoint gene expression with Coag-Score

Single-sample gene set enrichment analysis (ssGSEA) in the "GSVA" R package was used to quantify the infiltration level of 16 immune cells in each GC patient (Zuo et al., 2020). The TIMER database (http://timer.cistrome.org/) was used to analyze the correlation between Coag-score and immune cells. Moreover, the "ggpubr" package was used to plot violins to describe the correlation between Coag-Score and immune checkpoint gene expression.

Tissue samples

In this study, paired GC tissues and adjacent non-cancer tissues were acquired from the Lanzhou University Second Hospital. Cohort 1: ten pairs of fresh GC tissues and adjacent non-cancer tissues were cryopreserved and were used for the quantitative analysis of the expression of CRG mRNAs by qRT-PCR. Cohort 2: 84 pairs of GC tissues and adjacent non-cancer tissues were formalin-fixed and paraffin-embedded, which were used to detect the expression of SERPINE1 protein in GC and adjacent tissues by immunohistochemistry (IHC). All patients signed an informed consent form. The Ethics Committee approved the study of Lanzhou University Second Hospital (Ethical Application Ref: 2019A-321).

qRT-PCR

The total RNA of GC tissues and adjacent non-cancer tissues was extracted using TRIzol reagent (Thermo Scientific, United States) according to the product instructions. cDNA synthesis was performed using the PrimeScript™ RT reagent kit (Takara, Japan). qRT-PCR was performed using the SYBR Primix Ex Taq™ II (Takara, Japan) on ABI-7500 instrument (Applied Biosystems, United States). GAPDH was used as an internal reference gene, and the $2^{-\triangle\triangle Ct}$ method was used to compare the differential expression of genes. The primers are listed in Supplementary Table S1.

Immunohistochemical staining

Immunohistochemical staining was performed according to the standard procedure. GC tissues were embedded into wax blocks, and then sections were prepared, which were dewaxed in xylene, and hydrated in gradient concentration alcohol, which were then washed with PBS buffer. Antigen repair was performed with citrate buffer in a water bath. Then the sections were incubated with the primary antibody anti-mouse SERPINE1 (1:200, abcam) overnight at 4°C. The sections were washed with PBS the next day and incubated with the secondary antibody for 60 min at room temperature. Then, the DAB kit (MXB biotechnologies, Fujian, China) was used to stain tissue samples.

After staining of GC tissue, an IHC score was performed by pathologists. According to the proportion of positive cells and staining intensity, the scores were divided as follows: 0 was negatively stained or <5% positive cells, one was weakly stained or 6–25% positive cells, two was moderately stained or 26–50% positive cells, and three was strongly positive or >50% positive cells. We defined the final staining score ≥3 as a high expression of SERPINE1, and the patients were divided into the high expression group and low expression group based on this staining score.

Statistical analysis

Data were presented as mean ± SEM. We used the Student's t-test to examine the difference in mean between the two groups. Non-abnormal distribution data were analyzed using a nonparametric test. The Kaplan–Meier method was used to compare the survival times of different CRG expression levels. Univariate Cox regression analysis was used to analyze the prognostic value of a single gene; for factors with a p-value < 0.05, we performed multivariate Cox regression to analyze the independent risk factors that affect GC patients. p-value <0.05 was considered to be statistically different. * represents $p < 0.05$ and ** represents $p < 0.01$. Statistical analysis was performed using the SPSS software package (version 24.0, IBM SPSS) and GraphPad Prism (version 8.0, GraphPad Software).

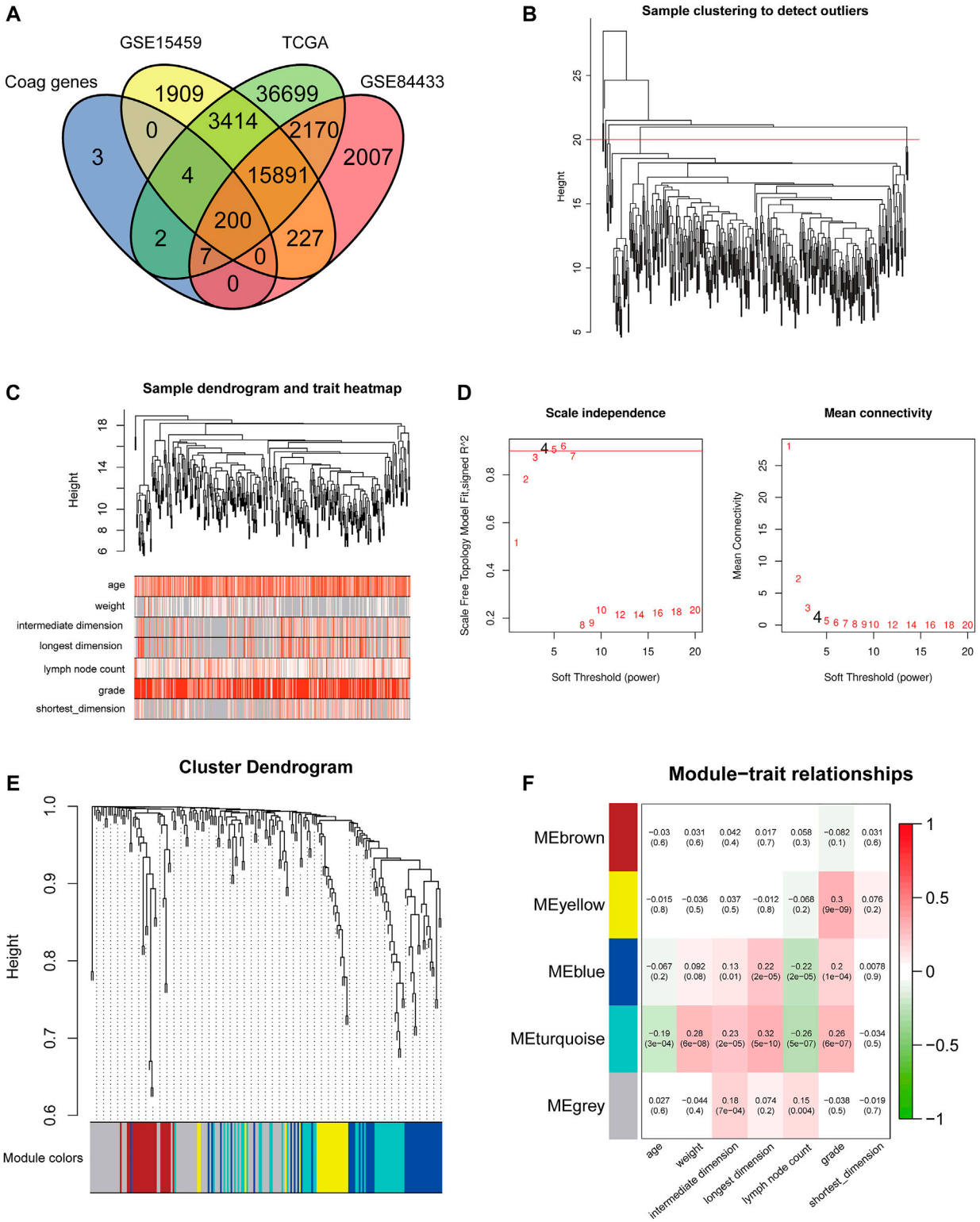


FIGURE 2
Construction of the co-expression modules of coagulation genes related to clinical characteristics of GC. **(A)** Venn diagram of CRGs in GC. **(B)** Sample clustering of CRGs. **(C)** Sample dendrogram and corresponding clinical traits. **(D)** The soft threshold of the CRG module is defined by scale independence and mean connectivity. **(E)** Correlation between sample clustering and modules. **(F)** The relationship between CRG module and clinical features of GC.

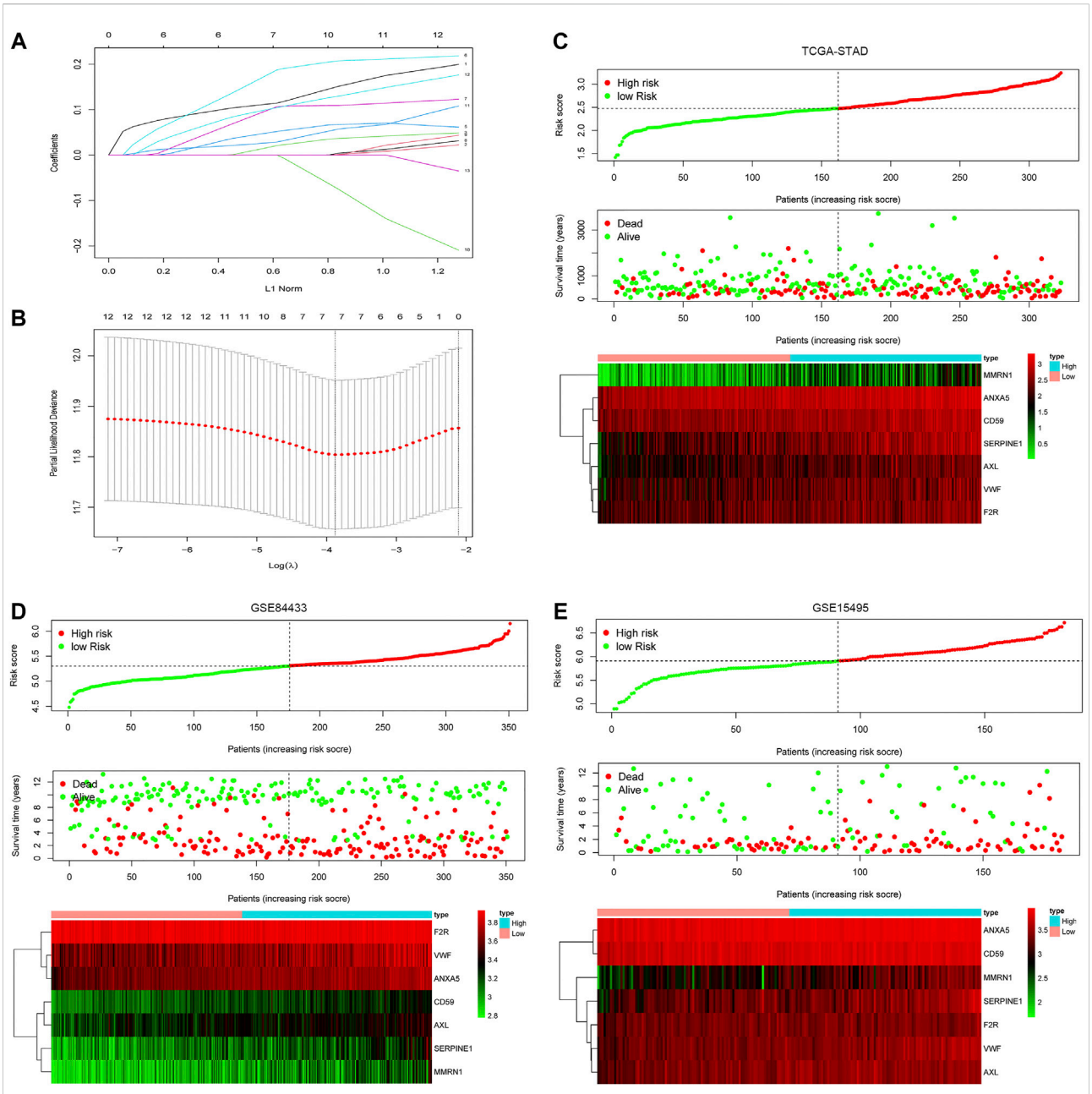


FIGURE 3
Coag-Score model construction. **(A,B)** Seven CRGs were screened based on LASSO regression analysis. **(C–E)** The distributions of the risk score for each patient (top panel), survival status of patients (middle panel), heatmaps for seven-gene signature between high-risk group and low-risk group (bottom panel).

Results

Defining CRGs associated with clinical features

A total of 974 GC patients from five cohorts were included in this study, including the training set TCGA-STAD, the validation sets GSE15495 and GSE84433, cohort 1, and cohort 2 from the Department of Oncology of Lanzhou University Second Hospital. CRGs were obtained from the AmiGO 2 database and compared with TCGA-STAD, GSE15495, and GSE84433 datasets. A total of 200 overlapping CRGs were selected (Figure 2A).

To better understand the gene expression network during GC development, we used WGCNA to construct co-expression networks and identify co-expression modules associated with

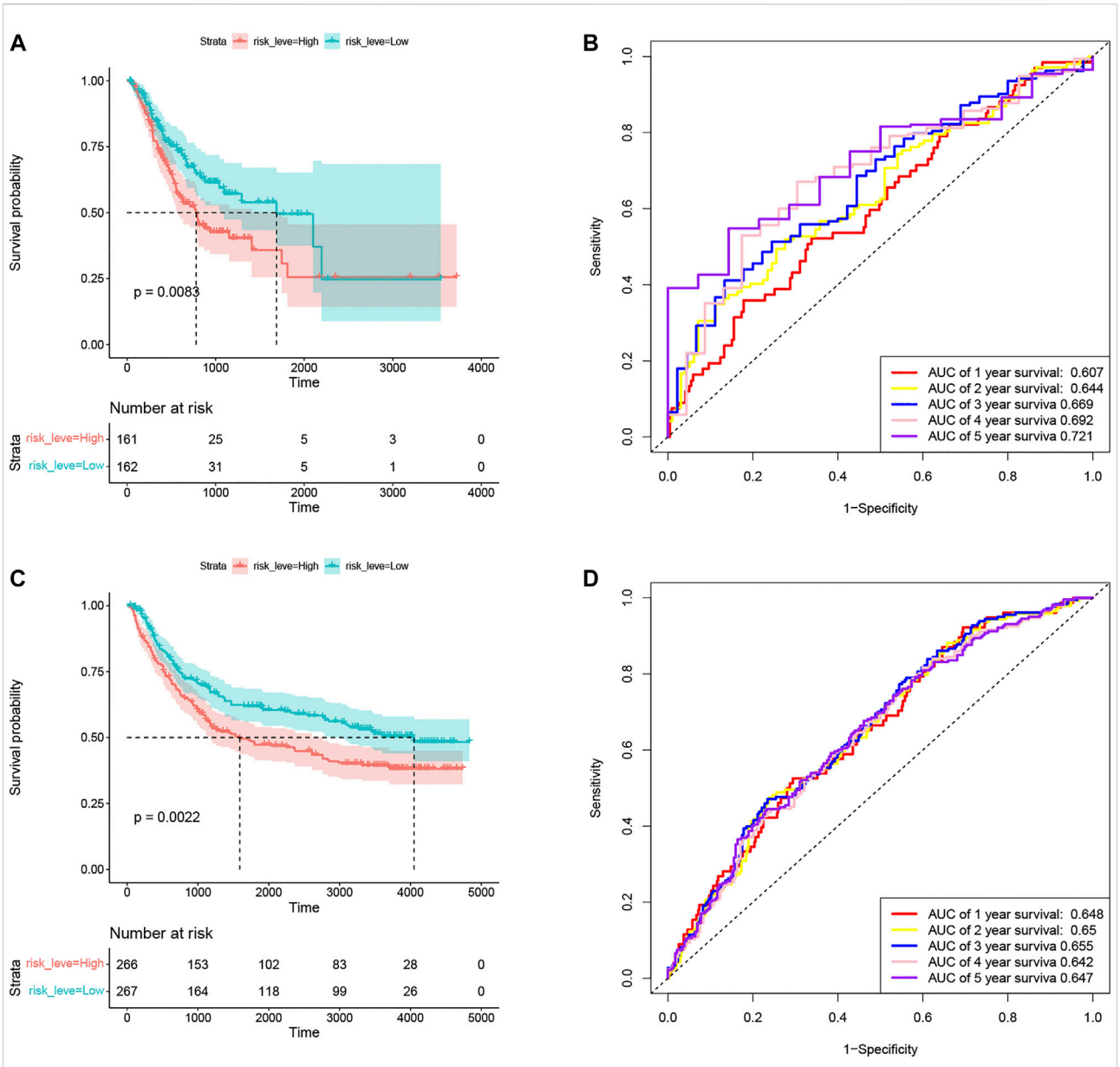


FIGURE 4
Evaluation and validation of the prognostic performance of Coag-Score in training and validation sets. **(A,C)** KM survival curves of Coag-Score in TCGA-STAD and GEO cohort. **(B,D)** ROC curves of Coag-score in TCGA-STAD and GEO cohort.

clinical features. The hierarchical clustering method was first used to obtain the GC sample cluster diagram, and the outlier samples were eliminated (Figure 2B). We analyzed the sample dendrogram and corresponding clinical traits (Figure 2C). The network analysis was performed, and an appropriate adjacency matrix weight parameter β was selected to satisfy the scale-free distribution as much as possible. We determined the appropriate soft threshold from the scale-free topology model fit-R^2. We selected β from the first approach of 0.09 to construct the gene module ($\beta = 4$) and divided CRGs into five modules (Figures 2D,E). After obtaining different gene modules, a correlation analysis was conducted between the clinical features of GC and the module eigengene (ME) value of each module. Among them, the genes in the blue and turquoise modules were highly correlated with the age, weight, tumor size, the number of lymph nodes, and tumor grade of GC patients (Figure 2F). Therefore, the blue and turquoise module genes were selected for further analysis. The blue module contains 87 CRGs, and the turquoise module contains 54 CRGs.

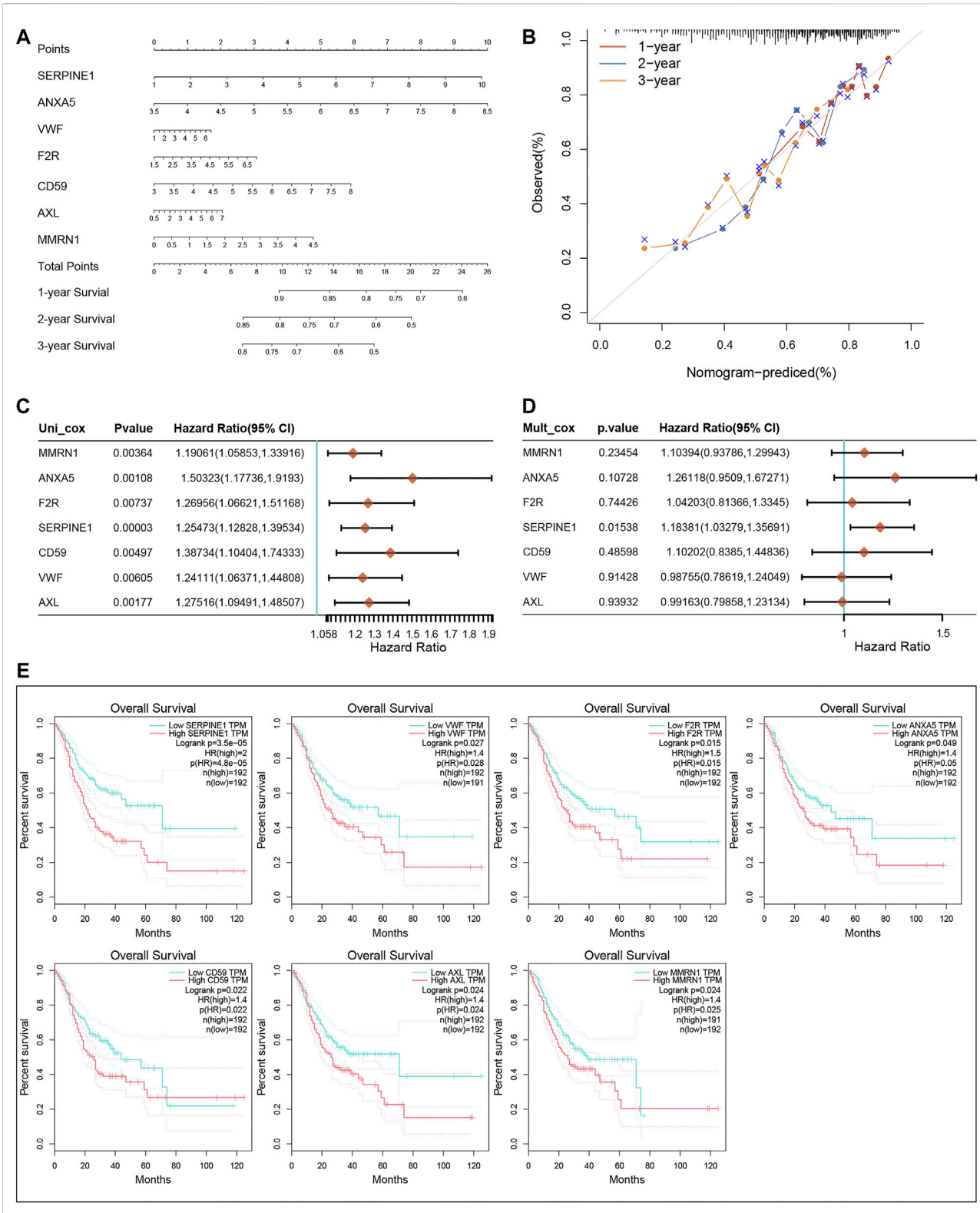


FIGURE 5
Nomogram of the Coag-Score model. **(A,B)** The nomogram and calibration curve of the Coag-Score model. **(C,D)** Univariate and multivariate Cox regression analyses for a single gene in the Coag-Score. **(E)** KM survival curves of a single gene in the Coag-Score model.

Construction of Coag-Score prognostic model based on CRGs associated with clinical features

In order to construct a simple and effective prognostic model, the LASSO Cox regression analysis was used to reduce the dimension of CRGs further. Seven CRGs were included in the model: Serpin Family E Member 1 (SERPINE1), von Willebrand factor (VWF), Coagulation Factor II Thrombin Receptor (F2R), Annexin A5 (ANXA5), CD59, AXL Receptor Tyrosine Kinase (AXL), and Multimerin 1 (MMRN1) (Figures 3A,B). The hazard ratios of SERPINE1, VWF, F2R, ANXA5, CD59, AXL, and MMRN1 were 1.234, 1.274, 1.303, 1.456, 1.403, 1.312, and 1.298 ($p < 0.05$), respectively. Based on the expression values and correlation coefficients of these seven CRGs, the prognostic risk score of each GC patient sample was calculated. Coag-Score = (0.110×expression level of SERPINE1) + (0.012×expression level of VWF) + (0.0460×expression level of F2R) + (0.167×expression level of ANXA5) + (0.092×expression level of CD59) + (0.026×expression level of ANXA5) + (0.097×expression level of MMRN1). GC patients were divided into high-risk and low-risk groups according to Coag-Score. Figures 3C–E show the risk curves of high-risk and low-risk groups (top panel), survival status (middle panel), and heatmap of single gene expression (bottom panel). The PCA indicated the patients in different risk groups were distributed in two directions (Supplementary Figure S1A-C).

Comparison of prognostic models in the training set and the validation set

In the training set TCGA-STAD, compared with patients in the low-risk group, the OS of GC patients in the high-risk group was significantly lower (Figure 4A). In order to verify the prediction performance of Coag-Score on different datasets, GSE15495 and GSE84433 datasets were used as validation sets. We integrated GSE15495 and GSE84433 datasets into the GEO cohort for analysis. The validation set survival analysis was consistent with the training set TCGA-STAD results (Figure 4C). The log-rank p values of the KM curve were 0.0083 and 0.0022, respectively. ROC curves were used to evaluate the sensitivity and specificity of the Coag-Score signature for the prognosis of GC patients (Figures 4B,D). The results show that the areas under the curve (AUCs) of the training set TCGA-STAD 1, 2, 3, 4, and 5 years were 0.607, 0.644, 0.669, 0.692, and 0.721, respectively. The AUCs of the validation set were 0.648, 0.65, 0.655, 0.642, and 0.647, respectively. Univariate and multivariate Cox regression analyses showed the Coag-Score was an independent risk factor affecting the prognosis of GC patients, and the Coag-Score had better predictive power and accuracy than other prognosis-related metrics (Supplementary Figure S2A-E).

Construction of nomogram for GC patients

We constructed a nomogram based on the seven CRGs and Coag-Score (Figure 5A). Predictions of 1-, 2-, and 3-year survival probabilities for GC patients in the training set were shown in the calibration plot (Figure 5B). We performed univariate and multivariate Cox regression analyses for a single gene in the Coag-Score, and seven CRGs affected the prognosis of GC patients. Interestingly, the multivariate Cox regression analysis showed that SERPINE1 was an independent risk factor affecting the prognosis of GC patients (Figures 5C,D). To stabilize the results, we added age, gender, pTNM stage, and grade as covariates; SERPINE1 remained an independent risk factor for prognosis (Supplementary Figure S3A, B). In addition, we performed a survival analysis for a single gene in the Coag-Score in the GEPIA online database (Figure 5E), and the OS of patients in the high expression group was significantly lower than that of those in the low expression group.

Analysis of enriched pathways between high-risk and low-risk cohorts

To further explore the underlying mechanism of CRGs, HALLMARK and KEGG gene sets were analyzed between the high-risk and low-risk groups; the positive and negative correlation pathways are shown in Table 1. The genes in the high-risk group were enriched in focal adhesion, MAPK signaling pathway, complement and coagulation cascades, angiogenesis, coagulation, epithelial–mesenchymal transition, and KRAS signaling pathway. Low-risk group genes are enriched in RNA degradation, spliceosome, cell cycle, E2F targets, and G2M checkpoint (Figure 6).

Analysis of tumor immune cell infiltration and immune checkpoint gene expression levels

In order to investigate the relationship between Coag-Score and tumor immune cell infiltration, ssGSEA was performed in the TCGA cohort. Fifteen immune cells had significantly higher infiltration enrichment fractions in the high-risk group: aDCs, B cells, $CD8^+$ T cells, dendritic cells (DCs), immature dendritic cells (iDCs), mast cells, neutrophils, NK cells, plasmacytoid dendritic cells (pDCs), T helper cells, follicular helper T cells (Tfh), helper T cells 1 (Th1 cells), helper T cells 2 (Th2 cells), tumor infiltrating lymphocytes (TIL), and regulatory T cells (Tregs) (Figure 7A). We further analyzed the Coag-Score in the TIMER database, and we found that the Coag-Score was positively correlated

TABLE 1 Result of Gene Set Enrichment Analysis (GSEA) between high-risk and low-risk groups.

Name	ES	NES	*p*-value
FOCAL_ADHESION	0.76913476	2.37036	0.001
MAPK_SIGNALING_PATHWAY	0.5525822	2.2472603	0.006
KRAS_SIGNALING_UP	0.6791901	2.1400163	0
ANGIOGENESIS	0.7791254	1.7455677	0.012145749
COAGULATION	0.651279	2.2354453	0
COMPLEMENT_AND_COAGULATION_CASCADES	0.7084431	2.2053964	0.012
EPITHELIAL_MESENCHYMAL_TRANSITION	0.8335778	2.0261042	0
RNA_DEGRADATION	−0.69585156	−2.1742415	0.018
SPLICEOSOME	−0.6367163	−2.092218	0.044
CELL_CYCLE	−0.67737764	−1.8732674	0.22
G2M_CHECKPOINT	−0.7602773	−1.9721214	0.002070393
E2F_TARGETS	−0.84644043	−1.9255103	0

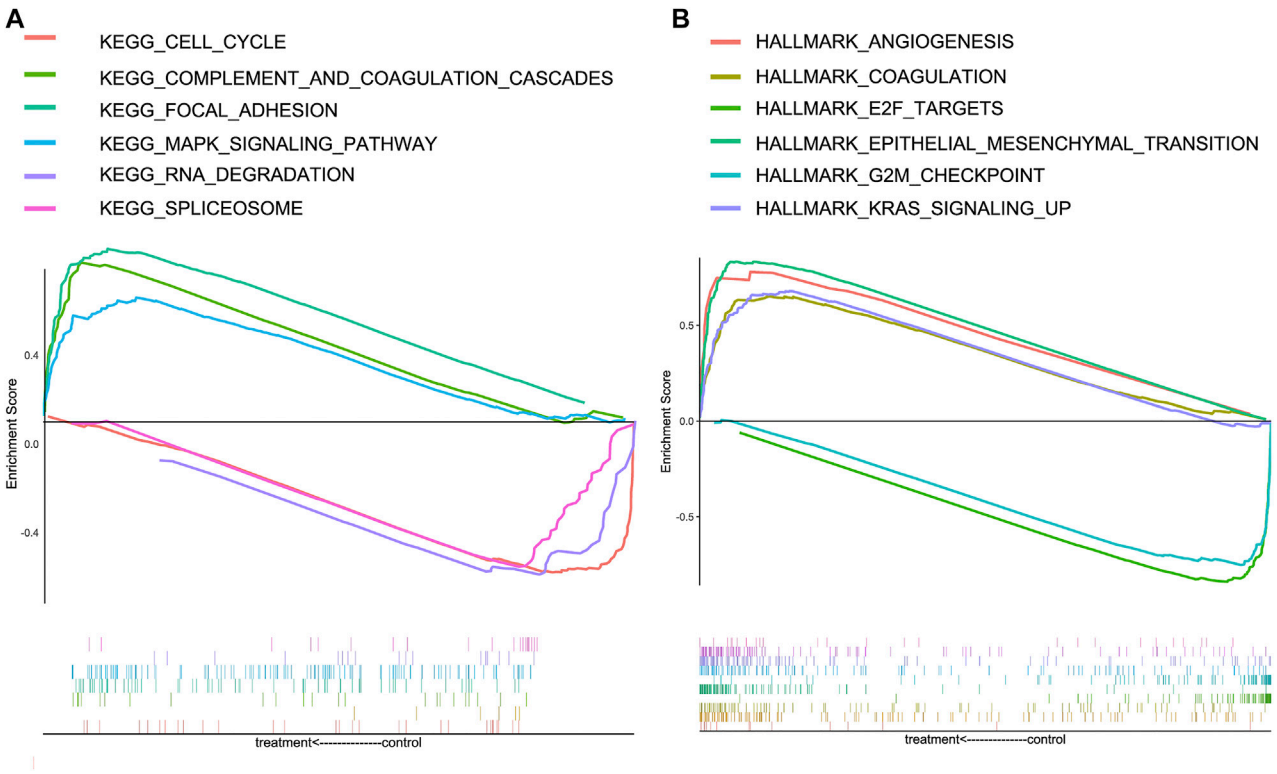


FIGURE 6
Gene Set Enrichment Analysis (GSEA) for high-risk and low-risk groups in the KEGG and HALLMARK datasets. **(A)** Enrichment pathways in the KEGG dataset of high-risk and low-risk groups. **(B)** Enrichment pathways in HALLMARK dataset of high-risk and low-risk groups.

with B cell, T-cell CD4$^+$, T-cell CD8$^+$, neutrophil, and myeloid dendritic cell (Figure 7B). In addition, we compared the expression of several critical immune checkpoints between high-risk and low-risk groups. The expression levels of PD-L1, PD-1, CTLA-4, TIM-3, LAG-3, and TIGIT were significantly increased (Figure 7C), indicating that patients in the high-risk group may respond better to immune checkpoint inhibitors.

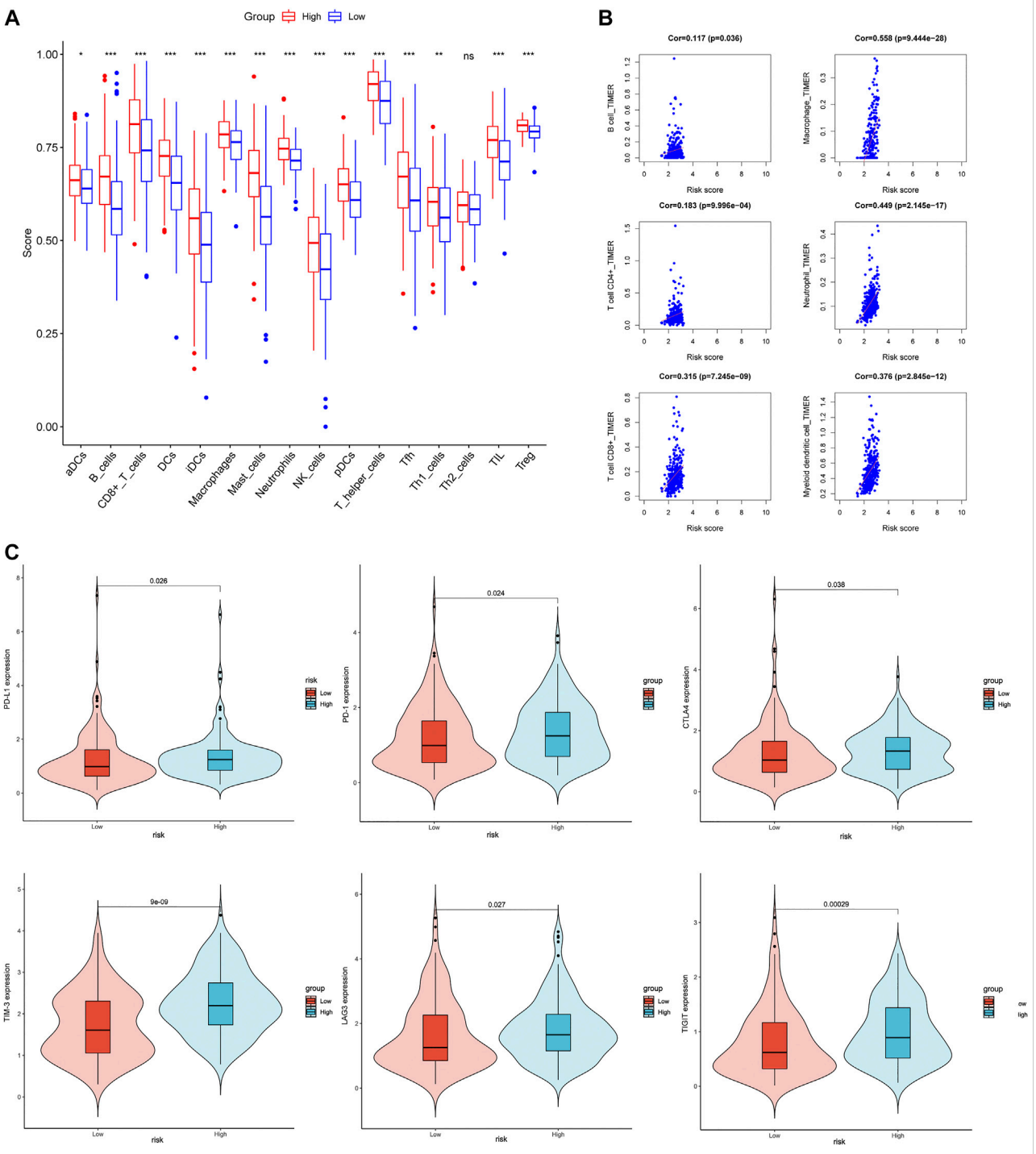


FIGURE 7
Correlation between Coag-Score and tumor immune infiltrating cells and immune checkpoints. **(A)** The ssGSEA scores between the high-risk and low-risk groups in the TCGA cohort. **(B)** Correlation between Coag-Score and immune infiltrating cells in the TIMER database. **(C)** Levels of immune checkpoint gene expression in high-risk and low-risk groups in the TCGA cohort.

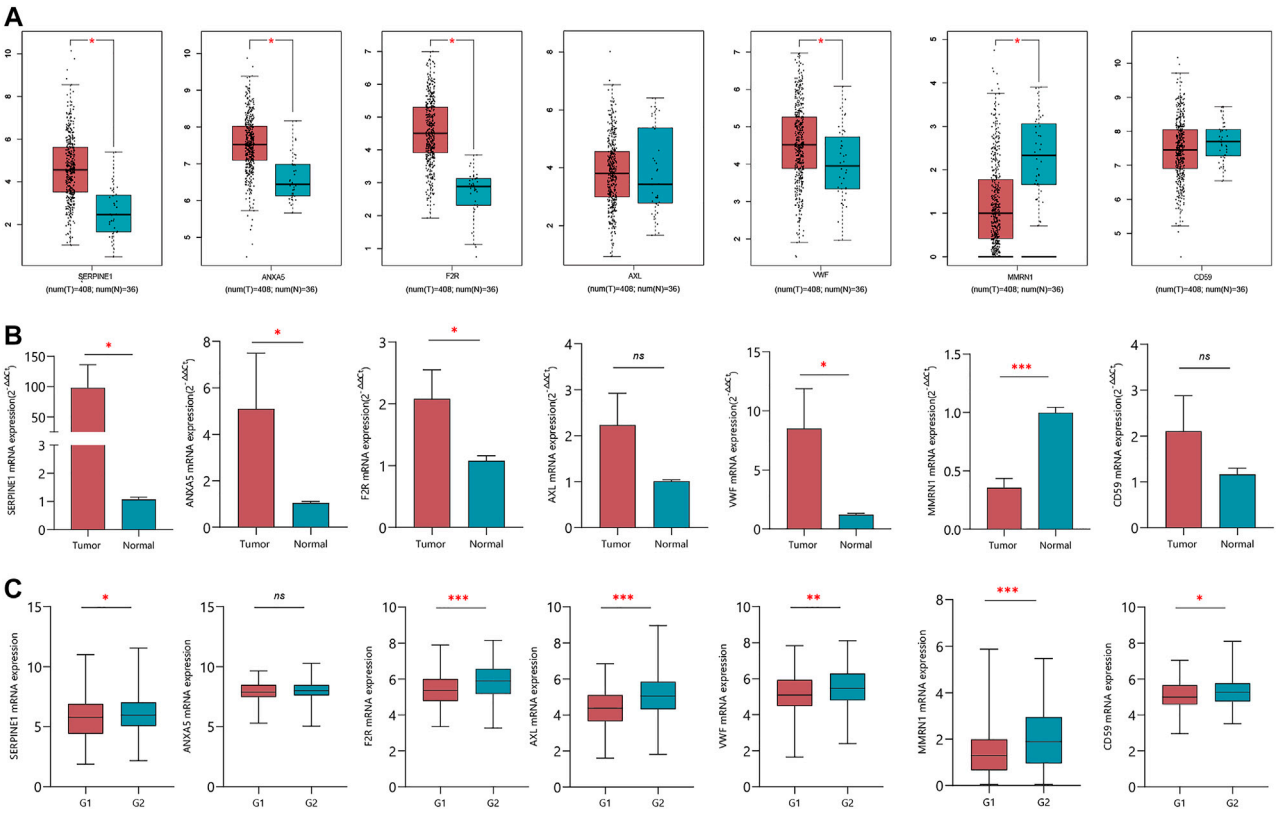


FIGURE 8
Verification of the expression of seven CRGs in normal and tumor tissues. **(A)** mRNA expression of seven CRGs in GEPIA online tool. **(B)** qRT-PCR verified the expression of seven CRGs in 10 pairs of GC clinical samples. **(C)** mRNA expression of seven CRGs in G1 and G2 groups.

Validation of CRG expression in GC tissues

We verified the expression levels of seven CRGs from mRNA levels. The GEPIA results showed that SERPINE1, ANXA5, F2R, and VWF increased significantly in tumor tissue. Although AXL was not statistically significant, there was a trend of high expression in tumor tissue. While MMRN1 was lowly expressed in tumor tissues, there was no significant difference in the expression level of CD59 (Figure 8A). In 10 pairs of GC samples, we performed qRT-PCR experiments, and the results showed that the mRNA expression of SERPINE1, ANXA5, F2R, and VWF was significantly higher than that in adjacent tissue. MMRN1 was expressed low in tumor tissue. Although there was no difference in AXL and CD59 mRNA in paired adjacent tissue and cancer tissue (Figure 8B), we analyzed the expression of seven CRGs in different pathological grades in the TCGA database. We integrated the pathological grades G1 and G2 as Group 1 (G1) and G3 as Group (G2). The results showed that the mRNA expression levels of the CRGs increased significantly in G2 compared to G1, except for ANXA5 (Figure 8C). But the mRNA expression of ANXA5 increased significantly in IV stage compared to II and III stages in the TCGA database (Supplementary Figure S4). The multivariate Cox regression analysis showed that SERPINE1 was an independent risk factor affecting the prognosis of GC patients. We used GC tissue samples from our center to verify the expression of SERPINE1 in GC and its prognostic value by IHC. In 84 cases, the expression of SERPINE1 was significantly increased in GC tissues compared with adjacent tissues (Figures 9A,B). Moreover, SERPINE1 was positively expressed in tissues with positive lymph node metastasis (Figures 9C,D). In survival analysis, the SERPINE1 high expression group had a worse prognosis than the low SERPINE1 expression group (Figure 9E).

Discussion

The morbidity and mortality of GC are increasing year by year. Despite the continuous improvement in therapeutic drugs and treatment methods, once GC recurs or metastasizes, the 5-year survival rate of patients decreases significantly (Guggenheim and Shah, 2013). Therefore, identifying effective prognostic biomarkers is crucial for predicting the occurrence and

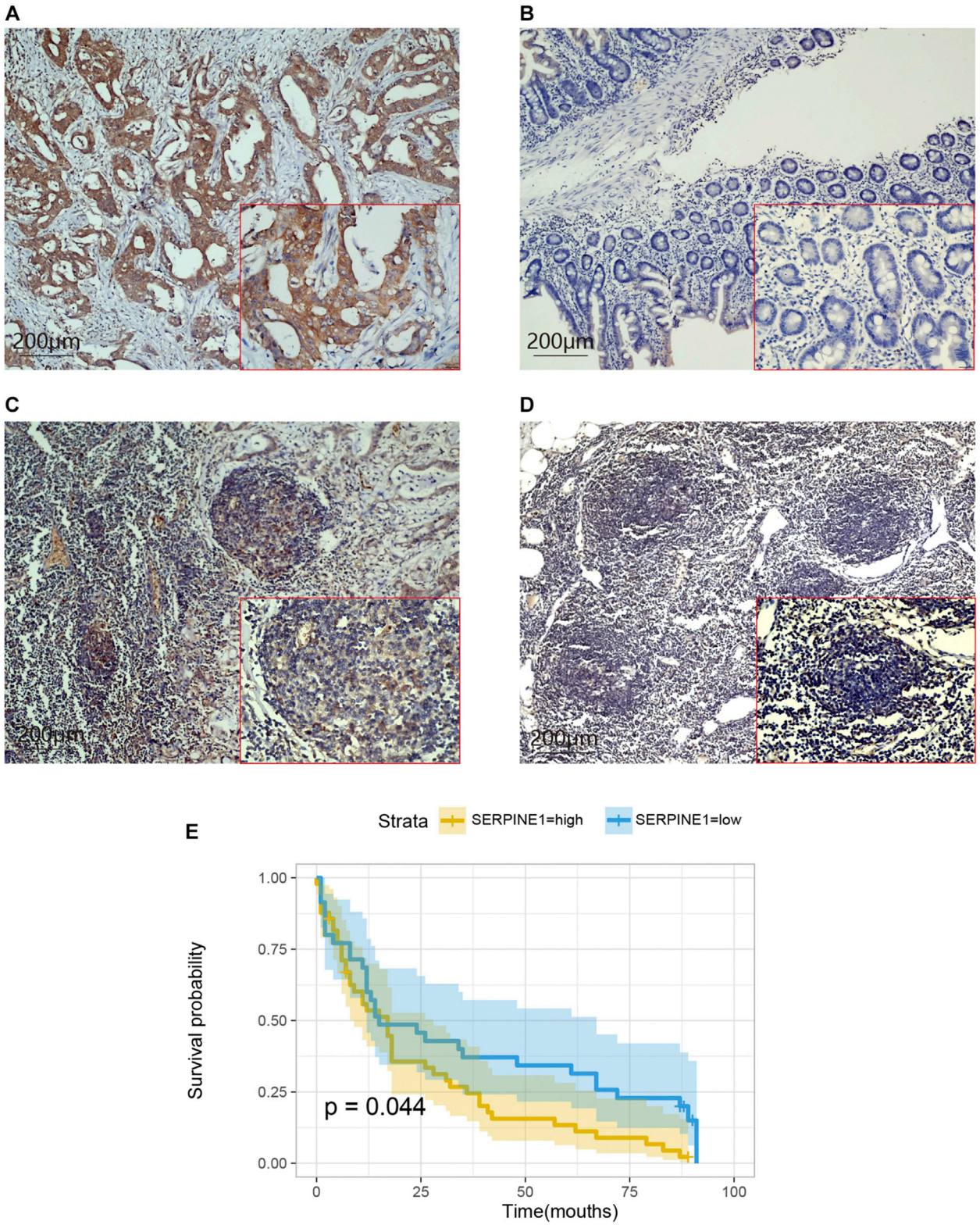


FIGURE 9
Expression and prognostic value of SERPINE1 in GC. **(A)** Positive staining of SERPINE1 in GC tissues. **(B)** Negative staining of SERPINE1 in adjacent tissues. **(C)** Positive staining of SERPINE1 in positive lymph node metastasis tissues. **(D)** Negative staining of SERPINE1 in negative lymph node metastasis tissues. **(E)** The survival analysis SERPINE1 low and high expression groups.

controlling disease progression in GC. There has been a great deal of interest in understanding tumor-related molecular pathways, with a focus on finding biomarkers associated with early diagnosis of cancer, tumor progression, chemotherapy, targeting, and immunotherapy responses, and influencing overall survival. Some of these studies have focused on characterizing genes and proteins associated with coagulation and fibrinolytic systems in carcinogenesis (Singh and Malviya, 2022).

Coagulation is a dynamic system in which the balance between coagulation and bleeding is always maintained in normal physiological conditions and often changes in disease conditions (Palta et al., 2014). Since 1960, hyperfibrinogen and hypercoagulability have been associated with rapidly growing tumors (Brugarolas and Elias, 1973). Systemic activation of hemostasis and thrombosis has been exhaustively implicated in cancer pathogenesis, progression, and metastasis (Langer and Bokemeyer, 2012; Lima and Monteiro, 2013). Disturbances of the coagulation system occur in GC. VTE is responsible for 10–20% of GC deaths. The incidence of clinically related VTE in GC patients was >5% in the first year after diagnosis and reached 12–17% in late GC, and the 2-year cumulative incidence of VTE in advanced GC increased to 24.4% (Lee et al., 2010; Larsen et al., 2015). The coagulation system is pivotal as a reservoir of tumor response markers and tumor angiogenesis, and the development of more effective antiangiogenic drugs.

We focused on the relationship between coagulation-related genes and the prognosis of GC, screened genes related to the clinical characteristics of GC by WGCNA, and then constructed a risk model of seven CRGs by LASSO regression. SERPINE1 is a member of the serine superfamily and encodes plasminogen activation inhibitor 1 (PAI-1). PAI-1 inhibits fibrinolysis and regulates plasminogen-induced extracellular matrix degradation and signal transduction by binding to the serine active center of uPA and tPA, resulting in a loss of plasminogen activity (Sprengers and Kluft, 1987). Many studies have found that SERPINE1 is abnormally expressed in GC tissue through bioinformatics analysis, and the expression level of SERPINE1 is negatively correlated with the prognosis of GC (Yang et al., 2019; Meng et al., 2020; Nie et al., 2020; Wu et al., 2020). Sakakibara found that the level of SERPINE1 increased significantly with the increase in tumor stage, leading to the occurrence of the malignant phenotype of tumors (Sakakibara et al., 2008). Downregulation of SERPINE1 can effectively reduce peritoneal metastasis and tumor progression in GC (Nishioka et al., 2012). SERPINE1 overexpression promotes malignant progression and poor prognosis in GC (Chen et al., 2022). ANXA5 is an anticoagulant protein that acts as an indirect inhibitor of the thromboplastin-specific complex, which is involved in the blood coagulation cascade (Ravassa et al., 2005). ANXA5 is a calcium-dependent lipid-binding protein secreted in the extracellular matrix (Bauwens, 2016), commonly used to detect apoptosis, drug transport, or as an adjunct to chemotherapy because of its high affinity for phosphatidylserine (PS) binding ability (Gerke and Moss, 2002). ANXA5 played a role in developing ovarian cancer, cervical cancer, and colorectal cancer. It was considered a diagnostic and prognostic marker (Xue et al., 2009; Li et al., 2012; Hassan et al., 2018). However, studies on ANXA5 in GC are limited, and the prognostic value of ANXA5 in GC is unclear. Wang et al. showed that ANXA5 might act as an anticancer protein, inhibit cell proliferation and metastasis, and promote cell apoptosis *via* the MEK/ERK signaling pathway (Wang et al., 2021). F2R (coagulation factor II thrombin receptor) F2R is a member of the G protein-coupled receptor family that encodes proteinase-activated receptor 1 (PAR1). High affinity receptors for activated thrombin coupled to G proteins that stimulate phosphoinositide hydrolysis may play a role in platelet activation and vascular development (Gao et al., 2020). PAR1 has been found to contribute to cell growth and invasion of tumor-derived cells (Even-Ram et al., 1998), and PAR1 is associated with poor prognosis in GC patients (Fujimoto et al., 2008). Further investigation revealed that PAR1 activation could trigger a cascade of responses that promote tumor cell growth and invasion. The activation of PAR1 leads to overexpression of NF-κB, EGFR, and TN-C, and TN-C induces EGFR activation by the autocrine mode. Therefore, PAR1 is a potentially important therapeutic target for GC (Fujimoto et al., 2010).

VWF is the largest polymeric glycoprotein in human blood. It is thought to be synthesized only in endothelial cells and megakaryocytes/platelets (Bongers et al., 2006). VWF is important in the maintenance of hemostasis; it promotes adhesion of platelets to the sites of vascular injury by forming a molecular bridge between sub-endothelial collagen matrix and platelet–surface receptor complex GPIb-IX-V. It also acts as a chaperone for coagulation factor VIII, delivering it to the site of injury, stabilizing its heterodimeric structure, and protecting it from premature clearance from plasma (Pagliari et al., 2021). VWF as a primary platelet ligand has been widely used as a biomarker for cancer and associated inflammation. Cancer-derived VWF enhances gastric adenocarcinoma metastasis through experiments *in vivo* and *in vitro* (Yang et al., 2018). Confusingly, Yin et al. (2021)found that ADAM28 from endothelial cells and GC cleaved VWF to eliminate VWF-induced apoptosis in GC cells. AXL is a member of the TYRO3, AXL, and MERTK (TAM) family of receptor tyrosine kinases (RTKs) that can be activated by the ligand Gas6, which is closely associated with tumor progression (Stitt et al., 1995). With respect to hemostasis, all three TAM receptors are located on platelets and mediate thrombogenesis and platelet stabilization. Platelet stabilization occurs after integrin activation, granule secretion, and platelet aggregation through platelet-to-platelet contact. Without this mechanism, platelet plugs disaggregate prematurely (van der Meer et al., 2014).

Gas6/AXL contributes to GC cell survival and invasion by activating the Akt pathway (Sawabu et al., 2007). He et al. similarly found that the Gas6/AXL/ZEB1 axis was upregulated in GC cell lines and negatively correlated with OS in GC patients. Upregulation of ZEB1 enhanced AXL-mediated EMT, invasion, and proliferation (He et al., 2020). In platelets, MMRN1 acts as a binding protein for factor V, a key regulator of coagulation, affecting factor V function and storage (Jeimy et al., 2008). MMRN1 is secreted from platelet α-granules and Weibel–Palade bodies of endothelial cells. MMRN1 platelet-related functions include platelet adhesion, factor V regulation, and MMRN1 deficiency associated with bleeding risks in Quebec platelet disorder (Posner, 2022). Laszlo et al. (2015) made a point that MMRN1 could be used as a new marker to refine pediatric acute myeloid leukemia. Sun et al. (2020) found that MMRN1 is one of the characteristic genes associated with GC prognosis. CD59 is a glycosylphosphatidylinositol (GPI)-anchored membrane protein that regulates complement activation by inhibiting membrane attack complex (MAC) formation (Zhang et al., 2018). CD59 deficiency was identified as a pro-thrombotic factor (Tabib et al., 2018). CD59 is highly expressed in various tumor tissues and cells, such as pancreatic cancer (Bongers et al., 2006) and colorectal cancer (Watson et al., 2006). The high expression of CD59 on cancer cells can inhibit the function of complement and directly protect cells from the apoptotic cascade through various signaling pathways. Recent studies have found that CD59 has also been shown to affect tumor cell behavior and immune cell activity (Zhang et al., 2018). Kiso et al. (2002) found that CD59 expression was enhanced in intestinal metaplasia, gastric adenocarcinoma, and intestinal-type gastric cancer but not in diffuse-type GC. How CD59 functions in GC has not been further studied.

Meanwhile, we constructed a nomogram of OS in GC patients based on coagulation-related genes for the first time, and the calibration curve and C-index showed good agreement. To further explore the underlying molecular mechanisms between high-risk and low-risk groups based on Coag-Score, the GSEA was performed in the HALLMARK and KEGG pathway datasets. The results indicated that high-risk individuals were enriched in focal adhesion, MAPK signaling pathway, complement and coagulation cascades, angiogenesis, coagulation, epithelial–mesenchymal transition, and KRAS signaling pathway. Jin et al. found that celecoxib exerts anticancer effects through focal adhesion (Jin et al., 2016). The MAPK signaling pathway is widely expressed in multicellular organisms and plays a crucial role in multiple biological processes such as cell proliferation and death, differentiation, migration, and invasion. The MAPK signaling pathway is usually involved in the occurrence and progression of cancer when it is dysregulated (Yang and Huang, 2015). The KRAS oncogene plays a crucial role in tumor initiation and maintenance, and its signaling network represents a significant target for therapeutic intervention. Many inhibitors targeting kinase effectors in various Ras signaling pathways have been developed (Luo, 2021). Activation of coagulation and complement and coagulation cascades correlates with chemotherapy sensitivity and OS in cancer patients (Zhang et al., 2020). Blocking the coagulation pathway with anticoagulants and other drugs reduces the incidence of deep vein thrombosis and effectively prolongs the survival of cancer patients. EMT is one of the hallmarks of carcinogenesis and involves the redifferentiation of epithelial cells into mesenchymal cells, thereby changing the cellular phenotype to malignant cells. EMT has been shown to play a role in malignant transformation, and when it occurs in the tumor microenvironment, it significantly affects the aggressiveness of GC (Kozak et al., 2020). In conclusion, GC patients with high-risk scores are characterized by a hypercoagulable state, and therefore, according to risk scores, anticoagulation therapy for high-risk groups may improve patient outcomes.

Tumor immunotherapy stimulates the body's immune function by increasing the immunogenicity of tumor cells and the sensitivity of effector cell killing, thereby inhibiting and killing tumor cells. The coagulation system plays an essential role in innate and adaptive immunity. Recent studies have found that the coagulation factor Xa (FXa) synthesized by monocytes and macrophages can promote tumor metastasis and immune escape by activating PAR-2. The FXa inhibitor, rivaroxaban, and PD-L1 inhibitor have synergistic anti-tumor effects (Graf et al., 2019). Elizabeth et al. found that increasing CD8 infiltration is correlated with impaired PFS and OS. Patients with higher $CD8^+$ T-cell densities also have higher PD-L1 expression, indicating an adaptive immune resistance mechanism may be occurring (Thompson et al., 2017). Key gene SERPINE1 in the Coag-Score model may predict the efficacy of PD-1 antibody in patients with advanced melanoma (Ohuchi et al., 2021). We further explored the correlation between Coag-Score and immune cell infiltration. ssGSEA showed a higher enrichment score of various immune cell infiltrations in the high-risk group. TIMER analysis Coag-Score was positively correlated with $CD8^+$ T cells, while PD-L1 expression was elevated in patients in the high-risk group. Therefore, high-risk patients may benefit from immunotherapy. Coag-Score may predict the efficacy of immune checkpoint inhibitors in GC patients. In addition, blockade of coagulation-related pathways may synergistically increase the effectiveness of immunotherapy.

The advantage of our study is to correlate the molecular dialogue between host cell coagulation factors and GC. We focused on coagulation-related genes to evaluate the prognostic value of GC patients. We explored the relationship between the Coag-Score model and immune infiltrating cells, and immune checkpoint expression, which provides a new predictive model and therapeutic strategy for the immunotherapy of GC. More clinical samples and further mechanistic studies are needed to verify the benefits and value of the Coag-Score model.

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 The Ethics Committee of Lanzhou University Second Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

HC and BW designed this study. BW, DZ, HW, and NW wrote the manuscript. TZ, LG, CM, and PZ participated in the coordination of the study and interpretation of results. BG and XL analyzed the data. DZ, YW, PH, YM, and XW revised the manuscript. All authors approved the final version for submission.

Funding

This study was supported by Key Project of Science and Technology in Gansu province (19ZD2WA001) and Key Talents Program of Gansu Province (2019RCXM020) and Cuiying Scientific and Technological Innovation Program of Lanzhou University Second Hospital (No. CY2017-ZD01) and Science and technology project of Chengguan District of Lanzhou City (2019RCCX0034, 2020SHFZ0039, 2020JSCX0073) and "Innovation Star" Project for Outstanding Graduate Students in Gansu Province (2022CXZX-162) and Key project of science and technology innovation platform fund of Gansu Provincial People's Hospital (21gssya-4).

Acknowledgments

We thank Key laboratory of the Digestive System Tumors of Gansu Province for experiment support.

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/fgene.2022.957655/full#supplementary-material

References

Bauwens, M. (2016). *In vivo* apoptosis imaging using site-specifically (68)Ga-labeled annexin V. *Methods Mol. Biol.* 1419, 17–26. doi:10.1007/978-1-4939-3581-9_2

Bongers, T. N., de Maat, M. P., van Goor, M. L., Bhagwanbali, V., van Vliet, H. H., Gómez García, E. B., et al. (2006). High von willebrand factor levels increase the risk of first ischemic stroke: Influence of ADAMTS13, inflammation, and genetic variability. *Stroke* 37 (11), 2672–2677. doi:10.1161/01.STR.0000244767.39962.f7

Brugarolas, A., and Elias, E. G. (1973). Incidence of hyperfibrinogenemia in 1961 patients with cancer. *J. Surg. Oncol.* 5 (4), 359–364. doi:10.1002/jso.2930050411

Buller, H. R., van Doormaal, F. F., van Sluis, G. L., and Kamphuisen, P. W. (2007). Cancer and thrombosis: from molecular mechanisms to clinical presentations. *J. Thromb. Haemost.* 5 (1), 246–254. doi:10.1111/j.1538-7836.2007.02497.x

Chen, S., Li, Y., Zhu, Y., Fei, J., Song, L., Sun, G., et al. (2022). SERPINE1 overexpression promotes malignant progression and poor prognosis of gastric cancer. *J. Oncol.* 2022, 2647825. doi:10.1155/2022/2647825

Dupuy, E., Hainaud, P., Villemain, A., Bodevin-Phèdre, E., Brouland, J. P., Briand, P., et al. (2003). Tumoral angiogenesis and tissue factor expression during hepatocellular carcinoma progression in a transgenic mouse model. *J. Hepatol.* 38 (6), 793–802. doi:10.1016/s0168-8278(03)00086-2

Even-Ram, S., Uziely, B., Cohen, P., Grisaru-Granovsky, S., Maoz, M., Ginzburg, Y., et al. (1998). Thrombin receptor overexpression in malignant and physiological invasion processes. *Nat. Med.* 4 (8), 909–914. doi:10.1038/nm0898-909

Fujimoto, D., Hirono, Y., Goi, T., Katayama, K., Matsukawa, S., and Yamaguchi, A. (2010). The activation of proteinase-activated Receptor-1 (PAR1) mediates gastric cancer cell proliferation and invasion. *BMC Cancer* 10, 443. doi:10.1186/1471-2407-10-443

Fujimoto, D., Hirono, Y., Goi, T., Katayama, K., and Yamaguchi, A. (2008). Prognostic value of protease-activated receptor-1 (PAR-1) and matrix metalloproteinase-1 (MMP-1) in gastric cancer. *Anticancer Res.* 28 (2), 847–854.

Gao, G., Yang, M., Wang, F., Dang, G., Zhang, X., Zhao, J., et al. (2020). Coagulation factor 2 thrombin receptor promotes malignancy in glioma under SOX2 regulation. *Aging (Albany NY)* 12 (11), 10594–10613. doi:10.18632/aging.103281

Gerke, V., and Moss, S. E. (2002). Annexins: From structure to function. *Physiol. Rev.* 82 (2), 331–371. doi:10.1152/physrev.00030.2001

Graf, C., Wilgenbus, P., Pagel, S., Pott, J., Marini, F., Reyda, S., et al. (2019). Myeloid cell-synthesized coagulation factor X dampens antitumor immunity. *Sci. Immunol.* 4 (39), eaaw8405. doi:10.1126/sciimmunol.aaw8405

Guggenheim, D. E., and Shah, M. A. (2013). Gastric cancer epidemiology and risk factors. *J. Surg. Oncol.* 107 (3), 230–236. doi:10.1002/jso.23262

Haist, M., Stege, H., Pemler, S., Heinz, J., Fleischer, M. I., Graf, C., et al. (2021). Anticoagulation with factor Xa inhibitors is associated with improved overall response and progression-free survival in patients with metastatic malignant melanoma receiving immune checkpoint inhibitors-A retrospective, real-world cohort study. *Cancers (Basel)* 13 (20), 5103. doi:10.3390/cancers13205103

Hanahan, D., and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell.* 100 (1), 57–70. doi:10.1016/s0092-8674(00)81683-9

Hassan, H. A., Salem, M. L., Gouida, M. S., and El-Azab, K. M. (2018). Comparative expression of caspases and annexin V in benign and malignant ovarian tumors. *J. Cancer Res. Ther.* 14 (5), 1042–1048. doi:10.4103/0973-1482.187282

He, L., Lei, Y., Hou, J., Wu, J., and Lv, G. (2020). Implications of the receptor tyrosine kinase axl in gastric cancer progression. *Onco. Targets. Ther.* 13, 5901–5911. doi:10.2147/ott.S257606

Jeimy, S. B., Krakow, E. F., Fuller, N., Tasneem, S., and Hayward, C. P. (2008). An acquired factor V inhibitor associated with defective factor V function, storage and binding to multimerin 1. *J. Thromb. Haemost.* 6 (2), 395–397. doi:10.1111/j.1538-7836.2007.02860.x

Jin, G. H., Xu, W., Shi, Y., and Wang, L. B. (2016). Celecoxib exhibits an anti-gastric cancer effect by targeting focal adhesion and leukocyte transendothelial migration-associated genes. *Oncol. Lett.* 12 (4), 2345–2350. doi:10.3892/ol.2016.4976

Kiso, T., Mizuno, M., Nasu, J., Shimo, K., Uesu, T., Yamamoto, K., et al. (2002). Enhanced expression of decay-accelerating factor and CD59/homologous restriction factor 20 in intestinal metaplasia, gastric adenomas and intestinal-type gastric carcinomas but not in diffuse-type carcinomas. *Histopathology* 40 (4), 339–347. doi:10.1046/j.1365-2559.2002.01350.x

Kozak, J., Forma, A., Czeczelewski, M., Kozyra, P., Sitarz, E., Radzikowska-Büchner, E., et al. (2020). Inhibition or reversal of the epithelial-mesenchymal transition in gastric cancer: pharmacological approaches. *Int. J. Mol. Sci.* 22 (1), E277. doi:10.3390/ijms22010277

Langer, F., and Bokemeyer, C. (2012). Crosstalk between cancer and haemostasis. implications for cancer biology and cancer-associated thrombosis with focus on tissue factor. *Hamostaseologie* 32 (2), 95–104. doi:10.5482/ha-1160

Langfelder, P., and Horvath, S. (2008). Wgcna: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9, 559. doi:10.1186/1471-2105-9-559

Larsen, A. C., Frøkjær, J. B., Fisker, R. V., Iyer, V., Mortensen, P. B., Yilmaz, M. K., et al. (2015). Treatment-related frequency of venous thrombosis in lower esophageal, gastro-esophageal and gastric cancer-a clinical prospective study of outcome and prognostic factors. *Thromb. Res.* 135 (5), 802–808. doi:10.1016/j.thromres.2015.01.021

Laszlo, G. S., Alonzo, T. A., Gudgeon, C. J., Harrington, K. H., Gerbing, R. B., Wang, Y. C., et al. (2015). Multimerin-1 (MMRN1) as novel adverse marker in pediatric acute myeloid leukemia: a report from the children's Oncology group. *Clin. Cancer Res.* 21 (14), 3187–3195. doi:10.1158/1078-0432.Ccr-14-2684

Lee, K. W., Bang, S. M., Kim, S., Lee, H. J., Shin, D. Y., Koh, Y., et al. (2010). The incidence, risk factors and prognostic implications of venous thromboembolism in patients with gastric cancer. *J. Thromb. Haemost.* 8 (3), 540–547. doi:10.1111/j.1538-7836.2009.03731.x

Li, X., Chen, L., Liang, X. J., Gao, Y. F., Wang, X. J., Xu, Q., et al. (2012). Annexin A5 protein expression is associated with the histological differentiation of uterine cervical squamous cell carcinoma in patients with an increased serum concentration. *Mol. Med. Rep.* 6 (6), 1249–1254. doi:10.3892/mmr.2012.1078

Lima, L. G., and Monteiro, R. Q. (2013). Activation of blood coagulation in cancer: implications for tumour progression. *Biosci. Rep.* 33 (5), e00064. doi:10.1042/bsr20130057

Luo, J. (2021). KRAS mutation in pancreatic cancer. *Semin. Oncol.* 48 (1), 10–18. doi:10.1053/j.seminoncol.2021.02.003

Matsuoka, T., and Yashiro, M. (2018). Biomarkers of gastric cancer: Current topics and future perspective. *World J. Gastroenterol.* 24 (26), 2818–2832. doi:10.3748/wjg.v24.i26.2818

Meng, C., Xia, S., He, Y., Tang, X., Zhang, G., and Zhou, T. (2020). Discovery of prognostic signature genes for overall survival prediction in gastric cancer. *Comput. Math. Methods Med.* 2020, 5479279. doi:10.1155/2020/5479279

Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* 34 (3), 267–273. doi:10.1038/ng1180

Nie, K., Shi, L., Wen, Y., Pan, J., Li, P., Zheng, Z., et al. (2020). Identification of hub genes correlated with the pathogenesis and prognosis of gastric cancer via bioinformatics methods. *Minerva Med.* 111 (3), 213–225. doi:10.23736/s0026-4806.19.06166-4

Nishioka, N., Matsuoka, T., Yashiro, M., Hirakawa, K., Olden, K., and Roberts, J. D. (2012). Plasminogen activator inhibitor 1 RNAi suppresses gastric cancer metastasis *in vivo*. *Cancer Sci.* 103 (2), 228–232. doi:10.1111/j.1349-7006.2011.02155.x

Ohuchi, K., Kambayashi, Y., Hidaka, T., and Fujimura, T. (2021). Plasminogen activating inhibitor-1 might predict the efficacy of anti-PD1 antibody in advanced melanoma patients. *Front. Oncol.* 11, 798385. doi:10.3389/fonc.2021.798385

Pagliari, M. T., Boscarino, M., Cairo, A., Mancini, I., Martinelli, I., Bucciarelli, P., et al. (2021). ADAMTS13 activity, high VWF and FVIII levels in the pathogenesis of deep vein thrombosis. *Thromb. Res.* 197, 132–137. doi:10.1016/j.thromres.2020.10.037

Palta, S., Saroa, R., and Palta, A. (2014). Overview of the coagulation system. *Indian J. Anaesth.* 58 (5), 515–523. doi:10.4103/0019-5049.144643

Posner, M. G. (2022). Multimerin-1 and cancer: a review. *Biosci. Rep.* 42 (2), BSR20211248. doi:10.1042/bsr20211248

Ravassa, S., Bennaghmouch, A., Kenis, H., Lindhout, T., Hackeng, T., Narula, J., et al. (2005). Annexin A5 down-regulates surface expression of tissue factor: a novel mechanism of regulating the membrane receptor repertoir. *J. Biol. Chem.* 280 (7), 6028–6035. doi:10.1074/jbc.M411710200

Repetto, O., and De Re, V. (2017). Coagulation and fibrinolysis in gastric cancer. *Ann. N. Y. Acad. Sci.* 1404 (1), 27–48. doi:10.1111/nyas.13454

Sakakibara, T., Hibi, K., Koike, M., Fujiwara, M., Kodera, Y., Ito, K., et al. (2008). PAI-1 expression levels in gastric cancers are closely correlated to those in corresponding normal tissues. *Hepatogastroenterology.* 55 (85), 1480–1483.

Sawabu, T., Seno, H., Kawashima, T., Fukuda, A., Uenoyama, Y., Kawada, M., et al. (2007). Growth arrest-specific gene 6 and Axl signaling enhances gastric cancer cell survival via Akt pathway. *Mol. Carcinog.* 46 (2), 155–164. doi:10.1002/mc.20211

Singh, A. K., and Malviya, R. (2022). Coagulation and inflammation in cancer: limitations and prospects for treatment. *Biochim. Biophys. Acta. Rev. Cancer* 1877 (3), 188727. doi:10.1016/j.bbcan.2022.188727

Smyth, E. C., Nilsson, M., Grabsch, H. I., van Grieken, N. C., and Lordick, F. (2020). Gastric cancer. *Lancet* 396 (10251), 635–648. doi:10.1016/s0140-6736(20)31288-5

Sprengers, E. D., and Kluft, C. (1987). Plasminogen activator inhibitors. *Blood* 69 (2), 381–387. doi:10.1182/blood.v69.2.381.381

Stitt, T. N., Conn, G., Gore, M., Lai, C., Bruno, J., Radziejewski, C., et al. (1995). The anticoagulation factor protein S and its relative, Gas6, are ligands for the Tyro 3/Axl family of receptor tyrosine kinases. *Cell.* 80 (4), 661–670. doi:10.1016/0092-8674(95)90520-0

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102

Sun, M., Qiu, J., Zhai, H., Wang, Y., Ma, P., Li, M., et al. (2020). Prognostic implications of novel gene signatures in gastric cancer microenvironment. *Med. Sci. Monit.* 26, e924604. doi:10.12659/msm.924604

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Tabib, A., Hindi, I., Karbian, N., Zelig, O., Falach, B., and Mevorach, D. (2018). Prothrombotic mechanisms in patients with congenital p.Cys89Tyr mutation in CD59. *Thromb. Res.* 168, 67–77. doi:10.1016/j.thromres.2018.06.006

Tas, F., Ciftci, R., Kilic, L., Serilmez, M., Karabulut, S., and Duranyildiz, D. (2013). Clinical and prognostic significance of coagulation assays in gastric cancer. *J. Gastrointest. Cancer* 44 (3), 285–292. doi:10.1007/s12029-013-9490-x

Thompson, E. D., Zahurak, M., Murphy, A., Cornish, T., Cuka, N., Abdelfatah, E., et al. (2017). Patterns of PD-L1 expression and CD8 T cell infiltration in gastric adenocarcinomas and associated immune stroma. *Gut* 66 (5), 794–801. doi:10.1136/gutjnl-2015-310839

van der Meer, J. H., van der Poll, T., and van 't Veer, C. (2014). TAM receptors, Gas6, and protein S: roles in inflammation and hemostasis. *Blood* 123 (16), 2460–2469. doi:10.1182/blood-2013-09-528752

Wang, K., Chen, X., Jin, C., Mo, J., Jiang, H., Yi, B., et al. (2020). A novel immune-related genes prognosis biomarker for hepatocellular carcinoma. *Aging (Albany NY)* 13 (1), 675–693. doi:10.18632/aging.202173

Wang, X., Dai, Y., Zhao, Y., Li, M., Zhang, J., Ci, Y., et al. (2021). AnnexinA5 might suppress the phenotype of human gastric cancer cells via ERK pathway. *Front. Oncol.* 11, 665105. doi:10.3389/fonc.2021.665105

Watson, N. F., Durrant, L. G., Madjd, Z., Ellis, I. O., Scholefield, J. H., and Spendlove, I. (2006). Expression of the membrane complement regulatory protein CD59 (protectin) is associated with reduced survival in colorectal cancer patients. *Cancer Immunol. Immunother.* 55 (8), 973–980. doi:10.1007/s00262-005-0055-0

Wu, K. Z., Xu, X. H., Zhan, C. P., Li, J., and Jiang, J. L. (2020). Identification of a nine-gene prognostic signature for gastric carcinoma using integrated bioinformatics analyses. *World J. Gastrointest. Oncol.* 12 (9), 975–991. doi:10.4251/wjgo.v12.i9.975

Xiao, S., and Zhou, L. (2017). Gastric cancer: metabolic and metabolomics perspectives (Review). *Int. J. Oncol.* 51 (1), 5–17. doi:10.3892/ijo.2017.4000

Xue, G., Hao, L. Q., Ding, F. X., Mei, Q., Huang, J. J., Fu, C. G., et al. (2009). Expression of annexin a5 is associated with higher tumor stage and poor prognosis in colorectal adenocarcinomas. *J. Clin. Gastroenterol.* 43 (9), 831–837. doi:10.1097/MCG.0b013e31819cc731

Yang, A. J., Wang, M., Wang, Y., Cai, W., Li, Q., Zhao, T. T., et al. (2018). Cancer cell-derived von Willebrand factor enhanced metastasis of gastric adenocarcinoma. *Oncogenesis* 7 (1), 12. doi:10.1038/s41389-017-0023-5

Yang, G., Zhang, Y., and Yang, J. (2019). Identification of potentially functional CircRNA-miRNA-mRNA regulatory network in gastric carcinoma using bioinformatics analysis. *Med. Sci. Monit.* 25, 8777–8796. doi:10.12659/msm.916902

Yang, M., and Huang, C. Z. (2015). Mitogen-activated protein kinase signaling pathway and invasion and metastasis of gastric cancer. *World J. Gastroenterol.* 21 (41), 11673–11679. doi:10.3748/wjg.v21.i41.11673

Yin, Q., Gu, J., Qi, Y., Lu, Y., Yang, L., Liu, J., et al. (2021). ADAM28 from both endothelium and gastric cancer cleaves von willebrand factor to eliminate von Willebrand Factor-induced apoptosis of gastric cancer cells. *Eur. J. Pharmacol.* 898, 173994. doi:10.1016/j.ejphar.2021.173994

Zhang, J., Chen, M., Zhao, Y., Xiong, H., Sneh, T., Fan, Y., et al. (2020). Complement and coagulation cascades pathway correlates with chemosensitivity and overall survival in patients with soft tissue sarcoma. *Eur. J. Pharmacol.* 879, 173121. doi:10.1016/j.ejphar.2020.173121

Zhang, R., Liu, Q., Liao, Q., and Zhao, Y. (2018). CD59: a promising target for tumor immunotherapy. *Future Oncol.* 14 (8), 781–791. doi:10.2217/fon-2017-0498

Zuo, S., Wei, M., Wang, S., Dong, J., and Wei, J. (2020). Pan-cancer analysis of immune cell infiltration identifies a prognostic immune-cell characteristic score (ICCS) in lung adenocarcinoma. *Front. Immunol.* 11, 1218. doi:10.3389/fimmu.2020.01218

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 30 August 2022
DOI 10.3389/fgene.2022.940306

Check for updates

OPEN ACCESS

EDITED BY
Ehsan Nazemalhosseini-Mojarad,
Research Institute For Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Iran

REVIEWED BY
Kai Sun,
Liuzhou People's Hospital, China
Alessandro Mangogna,
Institute for Maternal and Child Health Burlo Garofolo (IRCCS), Italy

*CORRESPONDENCE
Dongsheng Cao,
cds18756008507@163.com

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

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 10 May 2022
ACCEPTED 01 August 2022
PUBLISHED 30 August 2022

CITATION
Yang H, Che D, Gu Y and Cao D (2022), Prognostic and immune-related value of complement C1Q (C1QA, C1QB, and C1QC) in skin cutaneous melanoma. *Front. Genet.* 13:940306. doi: 10.3389/fgene.2022.940306

COPYRIGHT
© 2022 Yang, Che, Gu and Cao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Prognostic and immune-related value of complement C1Q (C1QA, C1QB, and C1QC) in skin cutaneous melanoma

Huanglong Yang†, Dehui Che†, Yuxiang Gu† and Dongsheng Cao*

Department of Plastic Surgery, The Second Affiliated Hospital of Anhui Medical University, Hefei, China

Background: Skin cutaneous melanoma (SKCM) is a common malignancy that is associated with increased morbidity and mortality. Complement C1Q is composed of C1QA, C1QB, and C1QC and is involved in the occurrence and development of many malignant tumours. However, the effect of C1QA, C1QB, and C1QC expression on tumour immunity and prognosis of cutaneous melanoma remains unclear.

Methods: First, we analysed C1QA, C1QB, and C1QC expression levels and prognostic values using Gene Expression Profiling Interactive Analysis (GEPIA) and Tumour Immune Estimation Resource (TIMER) analysis, and further validation was performed using RT-qPCR, The Human Protein Atlas, The Cancer Genome Atlas (TCGA) dataset, and Gene Expression Omnibus dataset. We then performed univariate/multivariate Cox proportional hazard model, clinicopathological correlation, and receiver operating characteristic curve analysis using TCGA dataset and established a nomogram model. Differentially expressed genes associated with C1QA, C1QB, and C1QC in SKCM were identified and analysed using LinkedOmics, TIMER, the Search Tool for the Retrieval of Interacting Genes database, and Metascape and Cytoscape software platforms. We used TIMER, GEPIA, and single-sample gene set enrichment analysis (ssGSEA) to analyse the relationship between the three genes and the level of immune cell infiltration, biomarkers, and checkpoint expression in SKCM. Finally, GSEA was utilized to study the functional pathways of C1QA, C1QB, and C1QC enrichment in SKCM.

Results: The overexpression of C1QA, C1QB, and C1QC provided significant value in the diagnosis of SKCM and has been associated with better overall survival (OS). Multivariate Cox regression analysis indicated that C1QA, C1QB, and C1QC are independent prognostic biomarkers for patients with SKCM. Immune cell infiltration, biomarkers, and checkpoints were positively correlated with the expression of C1QA, C1QB, and C1QC. Furthermore, the results of functional and pathway enrichment analysis showed that immune-related and apoptotic pathways were significantly enriched in the high-expression group of C1QA, C1QB, and C1QC.

Conclusion: We found that C1QA, C1QB, and C1QC can be used as biomarkers for the diagnosis and prognosis of SKCM patients. The upregulated expression

levels of these three complement components benefit patients from OS and may increase the effect of immunotherapy. This result may be due to the dual effects of anti-tumour immunity and apoptosis.

KEYWORDS

complement C1Q, cutaneous melanoma, prediction, prognosis, immuno-therapy

Introduction

Melanoma is a highly malignant tumour that originates from melanocytes (Kozovska et al., 2016). Melanocytes are found in the skin, eyes, digestive tract, and reproductive system; however, skin melanomas, which are caused by cancerous changes in melanocytes, are the most common type of melanomas (Dominiak et al., 2016; Kuk et al., 2016). Over the past few decades, the incidence of melanoma has rapidly increased (Rastrelli et al., 2014; Berwick et al., 2016). Certain factors may promote the malignant transformation of melanocytes, such as UV exposure from the sun, white race, trauma, or adverse stimuli (Bloethner et al., 2009; Darmawan et al., 2019). Melanoma is the deadliest skin tumour with characteristics of high malignancy, early metastasis, high mortality, and poor prognosis (Caini et al., 2015; Sun et al., 2019). Therefore, there is an urgent need for new insights into the molecular mechanisms that lead to the occurrence and progression of skin cutaneous melanoma (SKCM) and to promote the development of biomarker targets and novel therapeutic options to improve the survival rate of patients.

The complement system is a key component of both adaptive and innate immune systems (Laskowski et al., 2016). C1Q can perform various immune and non-immune tasks in a complement-dependent or complement-independent manner (Kishore et al., 2016). C1Q is a recognition molecule in the classical complement pathway and is well known as a subunit of the first complement structure, C1 (Espericueta et al., 2020). It is involved in a series of pathological and physiological activities, including the occurrence and progression of cancer, elimination of immune complexes, infection of the body, and apoptosis of cells (Kishore et al., 2004; Yang et al., 2019). C1Q is overexpressed in multiple types of cancer microenvironments (Bulla et al., 2016), and its potential anti-tumour effects have been reported in several studies. Kaur et al. (2016) found that C1Q can upregulate the TNF-α-induced apoptosis pathway by regulating Bax and Fas, thereby promoting the apoptosis of ovarian cancer cells. Hong et al. (2009) reported that C1Q can promote apoptosis of prostate cancer cells by stimulating the tumour suppressor WOX1. In a study by Bandini et al. (2016), it was reported that C1Q prevents tumour angiogenesis and induces apoptosis in breast cancer cells by coordinating WOX1-related signalling pathways. However, C1Q has been shown to increase primary cell adhesion, migration, and proliferation of malignant pleural mesothelioma (Agostinis et al., 2017). This contradicts the anti-tumour effect of C1Q mentioned previously (Mangogna et al., 2019a; Mangogna et al., 2019b). Therefore, the specific role of C1Q in cutaneous melanoma requires further investigation.

C1Q consists of 18 polypeptide chains, including six C1QA, six C1QB, and six C1QC (Nayak et al., 2012). Recent studies have confirmed that C1QB is a reliable biomarker for identifying patients with early cutaneous melanoma and observing treatment outcomes in follow-up patients (Luo et al., 2011; Borden et al., 2021). Although these results are encouraging, the tumour immune infiltration relationship between C1QA, C1QB, C1QC, and SKCM has not been established, and the expression, mechanism of action, and prognostic significance of complement components in SKCM require further exploration. In the present study, we found that these three genes were highly expressed in SKCM and have been verified experimentally. The relationship between prognosis and clinical features of melanoma was analysed, and a nomogram map was created based on clinicopathological features. In addition, hub genes related to the three genes were analysed. We then explored the relationship between the expression of the three genes and tumour immunity in SKCM. Finally, GSEA was performed to understand the pathways of action of the three genes. Our findings will help reveal the multifaceted roles of C1QA, C1QB, and C1QC and provide evidence for future immunotherapy of melanoma.

Materials and methods

Data source

RNA transcriptome data (including 472 samples) and corresponding clinical profile data for clinicopathological stage, age, sex, and survival time for SKCM were obtained from The Cancer Genome Atlas website (TCGA, https://portal.gdc.cancer.gov/). Microarray data and clinical survival data for the validation cohorts GSE65904 (including 214 samples) and GSE53118 (including 79 samples) were downloaded from the Gene Expression Omnibus database (GEO, https://www.ncbi.nlm.nih.gov/geo/). We then

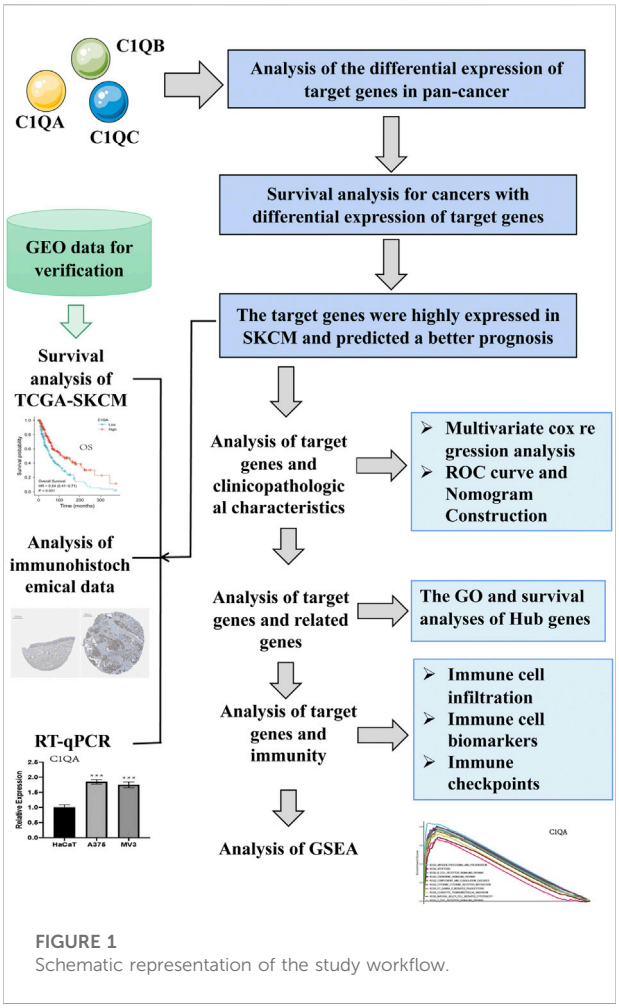


FIGURE 1
Schematic representation of the study workflow.

eliminated samples with insufficient or missing information for further analysis. The research process is illustrated in Figure 1.

Gene Expression Profiling Interactive Analysis and Tumour Immune Estimation Resource analysis

GEPIA (http://gepia.cancer-pku.cn/index.html) is an analysis tool developed based on TCGA and Genotype-Tissue Expression databases, which contain a large amount of RNA sequencing data and provide users with powerful functions such as differential expression analysis, correlation analysis, similar gene detection, and dimensionality reduction analysis (Tang et al., 2017). The TIMER (https://cistrome.shinyapps.io/timer) database contains more than 10,000 cases of 32 cancer types, which are used to estimate the level of six tumour infiltration immune subtypes and analyse tumour–immune interactions, including the relationship between immune infiltration and clinical characteristics, somatic mutations, and gene expression (Li et al., 2017). This provides a rich resource for cancer research.

In the current study, we used the GEPIA database to analyse the expression levels of C1QA, C1QB, and C1QC in normal samples and pan-cancer samples and screened out tumour types with significant expression differences for the subsequent analysis. Based on the analysis results of GEPIA and specific analysis modules in the TIMER database, we evaluated the correlation between the expression of C1QA, C1QB, and C1QC with overall survival (OS), abundance of immune cell infiltration, and immune checkpoint in SKCM patients and provided a reference for the selection of potential immunotherapy drugs. Finally, we validated the expression of proteins encoded by C1QA, C1QB, and C1QC in SKCM using immunohistochemical data from The Human Protein Atlas (HPA; www.proteinatlas.org) database (Uhlen et al., 2015).

Assessment of diagnostic ability and construction of a nomogram

The receiver operating characteristic (ROC) curve can be used to test the reliability of C1QA, C1QB, and C1QC in the diagnosis of SKCM. The larger the area under the ROC curve, the stronger the diagnostic capability. The nomogram survival prediction model is based on a multiple regression model that visualises genetic and clinical features on a scale to predict 1-, 3-, and 5-year survival probabilities. We then evaluated the prediction accuracy of the nomogram using the compliance index (C-index) and calibration curve and visualised the results using "pROC," "ggplot2," "survival," and "rms" R packages.

Related gene analysis

Differentially expressed genes (DEGs) that were closely related to C1QA, C1QB, and C1QC were screened using the LinkedOmics (http://www.linkedomics.org) website. Then, the protein–protein interaction (PPI) network analysis was performed on C1Q-related genes with a correlation >0.7 using the Search Tool for the Retrieval of Interacting Genes database (STRING, https://string-db.org/) (Szklarczyk et al., 2021). Subsequently, based on the aforementioned PPI analysis results, Cytoscape software and cytohubba plug-in were used to screen the ten genes with the largest number of adjacent network nodes of C1Q-related genes as hub genes. Finally, the Metascape (Metascape) and TIMER websites were used to perform Gene Ontology (GO) enrichment analysis and survival analysis on these hub genes.

Single-sample gene set enrichment analysis

ssGSEA could quantify the relationship between C1QA, C1QB, and C1QC in TCGA dataset and the infiltration of

TABLE 1 PCR primers utilized in this study.

gene	sequence (5′ to 3′)
C1QA-F	CCATATCGCTGGCCTCTATGG
C1QA-R	GTCTTCCTGCCTCCCCTTTC
C1QB-F	GGCCACCGACAAGAACTCACTAC
C1QB-R	CCATATCTGGAAAGAGCAGGAACC
C1QC-F	AACCAATCAGGTCAACTCGGG
C1QC-R	CCCACCATGTCGTAGTAGTCATTG
GAPDH-F	AAGGTGAAGGTCGGAGTCAAC
GAPDH-R	GGGGTCATTGATGGCAACAATA

24 types of immune cells (Bindea et al., 2013). Also, the "GSVA" R package was used to visualise the correlation between these three genes and immune cells, as well as the level of immune cell infiltration in the high- and low-expression groups of genes (Hanzelmann et al., 2013). The size of the circles and height of the bars represent the degree of correlation, and the shade of colour represents the magnitude of the *p*-value.

Gene set enrichment analysis

Based on GSEA, an application program that evaluates the distribution trend of pre-defined genes in the gene table of two states and their contributions in two biological states, we analysed the enrichment pathways of C1QA, C1QB, and C1QC in the high- and low-expression groups. Each stage of the analysis was repeated thousand times to ensure the accuracy of the results.

Culture of cell lines

SKCM cell lines A375 and MV3 and normal human HaCaT cells were obtained from the American Type Culture Collection (ATCC, Rockville, MD, United States) and Cell Lines Service (CLS, Eppelheim, Germany), respectively. A375 and HaCaT cell lines were cultured in DMEM (Shanghai Basalmedia Technologies Co., Ltd., Shanghai, China), while the MV3 cell line was cultured in RPMI-1640 medium (Shanghai Basalmedia Technologies, Shanghai, China). All media were supplemented with 10% foetal bovine serum (FBS; Gibco, Carlsbad, CA, United States) and penicillin (100 IU/ml)/streptomycin (100 μg/ml). The parameters of the humidified incubator were set to 37°C and 5% CO_2.

Reverse transcription-quantitative PCR

After extracting the total RNA of cells using conventional methods, the HyperScriptTM III RT SuperMix for qPCR with gDNA Remover (NovaBio, Shanghai, China) was used to convert it into cDNA. Subsequently, real-time PCR was performed using the S6 Universal SYBR qPCR mix (NovaBio) on a Bio-Rad iCycler Real-time Quantitative PCR system (Bio-Rad, CA, United States). Relevant primer sequences are listed in Table 1. The relative expression levels of the target genes were normalised to those of GAPDH and analysed using the $2^{-\Delta\Delta Ct}$ method.

Statistical analysis

R (v 3.6.3) software (https://www.r-project.org/) was used for data processing and statistical analysis of TCGA-SKCM and GEO cohorts, and $p < 0.05$ was considered to be statistically significant. Then, the "survival" and "survminer" R packages as well as the log-rank test were used to show the Kaplan–Meier plots. The "survival" R package was used to perform univariate/multivariate Cox risk regressions of three genes' expression and clinicopathological features. The "ggplot2" R package was used to investigate the relationship between three genes' expression and clinicopathological characteristics.

Results

mRNA expression levels of C1QA, C1QB, and C1QC in pan-cancer

The transcription levels of C1QA, C1QB, and C1QC were analysed based on pan-cancer data in GEPIA to observe whether there were significant differences in their expression among common tumour types. As indicated in Figure 2, compared to normal tissues, C1QA was significantly upregulated in 15 types of cancer, namely, DLBC, ESCA, GBM, KIRC, KIRP, LGG, LIHC, OV, PAAD, SKCM, STAD, TGCT, THYM, UCEC, and UCS, while it was downregulated in LUSC. In addition, the expression of C1QB was considerably increased in 14 types of malignancies, namely, DLBC, ESCA, GBM, HNSC, KIRC, KIRP, LGG, OV, PAAD, SKCM, STAD, TGCT, THYM, and UCEC but dramatically decreased in ACC, LUAD, and LUSC. Furthermore, C1QC was significantly upregulated in DLBC, ESCA, GBM, HNSC, KIRC, KIRP, LGG, OV, PAAD, SKCM, STAD, TGCT, THYM, UCEC, and UCS but significantly downregulated in ACC.

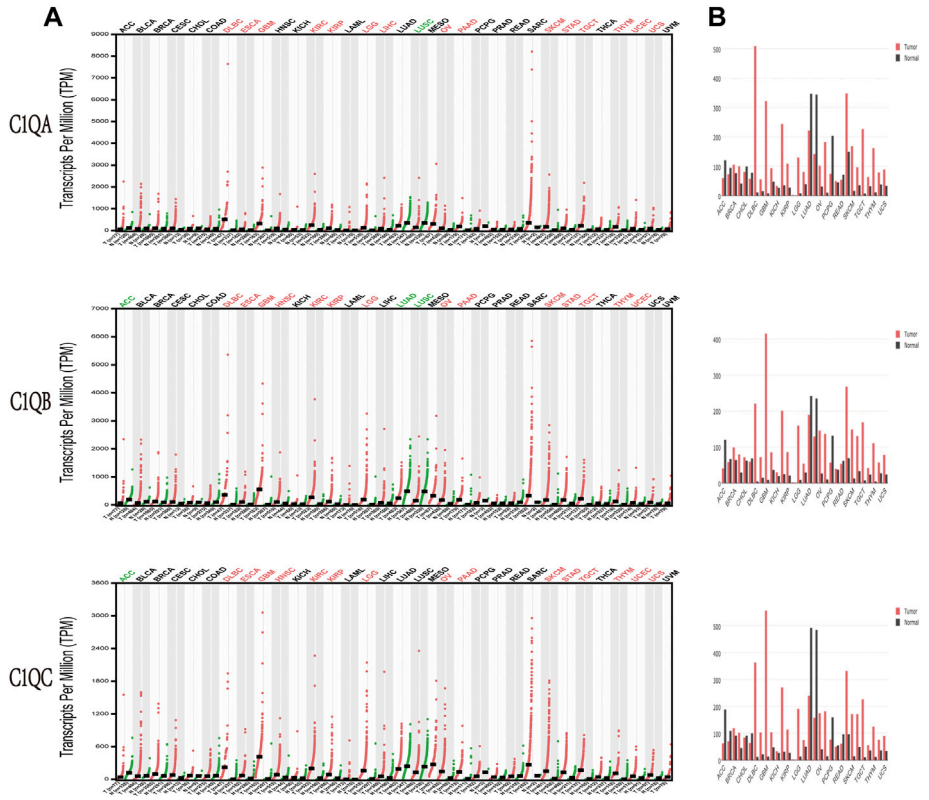


FIGURE 2
Expression levels of C1QA, C1QB, and C1QC in different types of human cancers compared with paired normal tissues by Gene Expression Profiling Interactive Analysis. **(A)** Dot plot. **(B)** Bar plot.

C1QA, C1QB, and C1Q were highly expressed in skin cutaneous melanoma and predicted a better prognosis

To further explore the correlation between the expression of C1QA, C1QB, and C1Q and the survival prognosis of cancer patients, we conducted a survival analysis of different tumour types using TIMER based on the aforementioned GEPIA analysis results. Interestingly, we found a strong association between the expression of these three genes and OS in SKCM (Table 2). Based on the aforementioned findings, we concentrated on the involvement of C1QA, C1QB, and C1QC in SKCM. Box plots of expression differences and OS curves obtained from the GEPIA database showed that the overexpression of C1QA, C1QB, and C1QC prolonged the OS of SKCM patients (Figures 3A, B).

The SKCM dataset was downloaded from TCGA to investigate the association between the expression of these three genes and OS, progression-free interval (PFI), and disease-specific survival (DSS). The results revealed that the overexpression of C1QA, C1QB, and C1QC had a significantly better prognostic value for OS, PFI, and DSS (Figure 3C–E). In addition, high expression of these three genes in SKCM improved OS for patients regardless of whether they received radiation therapy. These results suggested that radiotherapy did not affect the prognostic effects of these three genes in SKCM patients (Figures 3F, G). The analysis results of the HPA database also showed that the expression of proteins encoded by C1QA and C1QC was significantly increased in SKCM compared with that in normal skin tissue, but the protein encoded by C1QB was not detected in either tissue (Figure 3H). In addition, the association of C1QA, C1QB, and C1QC with OS was studied using the GEO database (GSE65904 and GSE53118). In terms of OS, the analysis results of C1QB and C1QC were consistent with those of TCGA data, and a high expression predicted a better prognosis. Although C1QA did not correlate well with overall survival, a high expression was associated with a better prognosis (Figures 3I, J). Thus, we can conclude that the overexpression levels of C1QA, C1QB, and C1QC are protective factors which improve survival in SKCM patients at different stages. Furthermore, we detected the expression levels of three genes in A375 and MV3 melanoma cells and normal HaCaT cells using RT-qPCR. As shown in Figure 3K, the expression levels of the

TABLE 2 Correlation between C1QA, C1QB, and C1QC and overall survival in cancers with significant differences in expression.

Cancer	Variable	*p*-value	Cancer	Variable	*p*-value	Cancer	Variable	*p*-value
SKCM	C1QA	1.62E−06	SKCM	C1QB	4.18E−06	SKCM	C1QC	2.31E−06
LGG	C1QA	0.002679	LGG	C1QB	0.00029	LGG	C1QC	0.000777
THYM	C1QA	0.006067	ESCA	C1QB	0.072503	THYM	C1QC	0.009601
KIRC	C1QA	0.044099	THYM	C1QB	0.113741	TGCT	C1QC	0.072349
ESCA	C1QA	0.058511	PAAD	C1QB	0.15336	ESCA	C1QC	0.078537
PAAD	C1QA	0.129492	OV	C1QB	0.159892	PAAD	C1QC	0.127006
TGCT	C1QA	0.183879	KIRC	C1QB	0.162042	KIRC	C1QC	0.311248
GBM	C1QA	0.22455	TGCT	C1QB	0.183879	UCEC	C1QC	0.379736
STAD	C1QA	0.344496	LUSC	C1QB	0.30921	STAD	C1QC	0.444084
KIRP	C1QA	0.369394	KIRP	C1QB	0.310142	HNSC	C1QC	0.488506
UCEC	C1QA	0.440934	UCEC	C1QB	0.409936	ACC	C1QC	0.49011
LIHC	C1QA	0.536671	GBM	C1QB	0.423744	KIRP	C1QC	0.533241
LUSC	C1QA	0.539205	DLBC	C1QB	0.508697	OV	C1QC	0.534026
OV	C1QA	0.58539	STAD	C1QB	0.515331	DLBC	C1QC	0.573832
DLBC	C1QA	0.59649	HNSC	C1QB	0.633914	GBM	C1QC	0.714599
UCS	C1QA	0.878045	ACC	C1QB	0.814952	UCS	C1QC	0.731257
			LUAD	C1QB	0.955506			

three genes were significantly increased in A375 and MV3 cells compared to those in HaCaT cells. These results suggest that these three genes are upregulated in cutaneous melanoma cells.

C1QA, C1QB, and C1QC were independent prognostic factors in skin cutaneous melanoma patients

Univariate and multivariate Cox regression analyses were performed on C1QA, C1QB, and C1QC with clinical variables to determine whether C1QA, C1QB, and C1QC could be distinguished from other clinical features as independent prognostic factors in SKCM patients. In conclusion, C1QA, C1QB, and C1QC can be distinguished from age and sex as independent prognostic factors for predicting patient survival, and the TNM stage can also be used as an independent prognostic factor for SKCM (Tables 3–5). In the aforementioned three tables, although radiotherapy is a protective factor for SKCM (Hazard ratio <1), its *p*-value is greater than 0.05, indicating that radiotherapy does not significantly affect the prognosis of melanoma patients. Based on the TCGA-SKCM dataset, we further explored the correlation between the expression of these three genes and their clinicopathological features. As shown in Figure 4, the expression of C1QA and C1QB was significantly correlated with pathological stage and sex, and the expression of all three genes was significantly correlated with T stage but not with M stage, N stage, and age.

Diagnostic and predictive values of C1QA, C1QB, and C1QC

The area under the ROC curve (AUC) was used to evaluate the diagnostic performance of C1QA, C1QB, and C1QC. As shown in Figure 5A, the areas under the curve (AUC) of C1QA, C1QB, and C1QC were 0.956 (CI: 0.945–0.967), 0.950 (CI: 0.938–0.961), and 0.958 (CI: 0.947–0.968), respectively. These findings suggest that all three genes are promising diagnostic markers for predicting SKCM, with C1QC being the best predictor, followed by C1QA and C1QC. We then visualised the results of the multivariate Cox proportional risk analysis as a nomogram to predict 1-, 3-, and 5-year OS (Figure 5B). Furthermore, the C-indexes of C1QA, C1QB, and C1QC were 0.687 (95% CI: 0.665–0.710), 0.684 (95% CI: 0.662–0.706), and 0.684 (95% CI: 0.662–0.707), respectively. Calibration further confirmed the accuracy of the nomogram's prediction ability. As shown in Figure 5C, the calibration curves showed a good fit to the diagonal, indicating that the predicted values were generally consistent with the actual observed values of the 1-, 3-, and 5-year OS probabilities.

Differentially expressed genes related to C1QA, C1QB, and C1QC in skin cutaneous melanoma

Using the LinkedOmics online program, we identified DEGs that interacted with C1QA, C1QB, and C1QC in patients with SKCM. The volcano plot showed that most

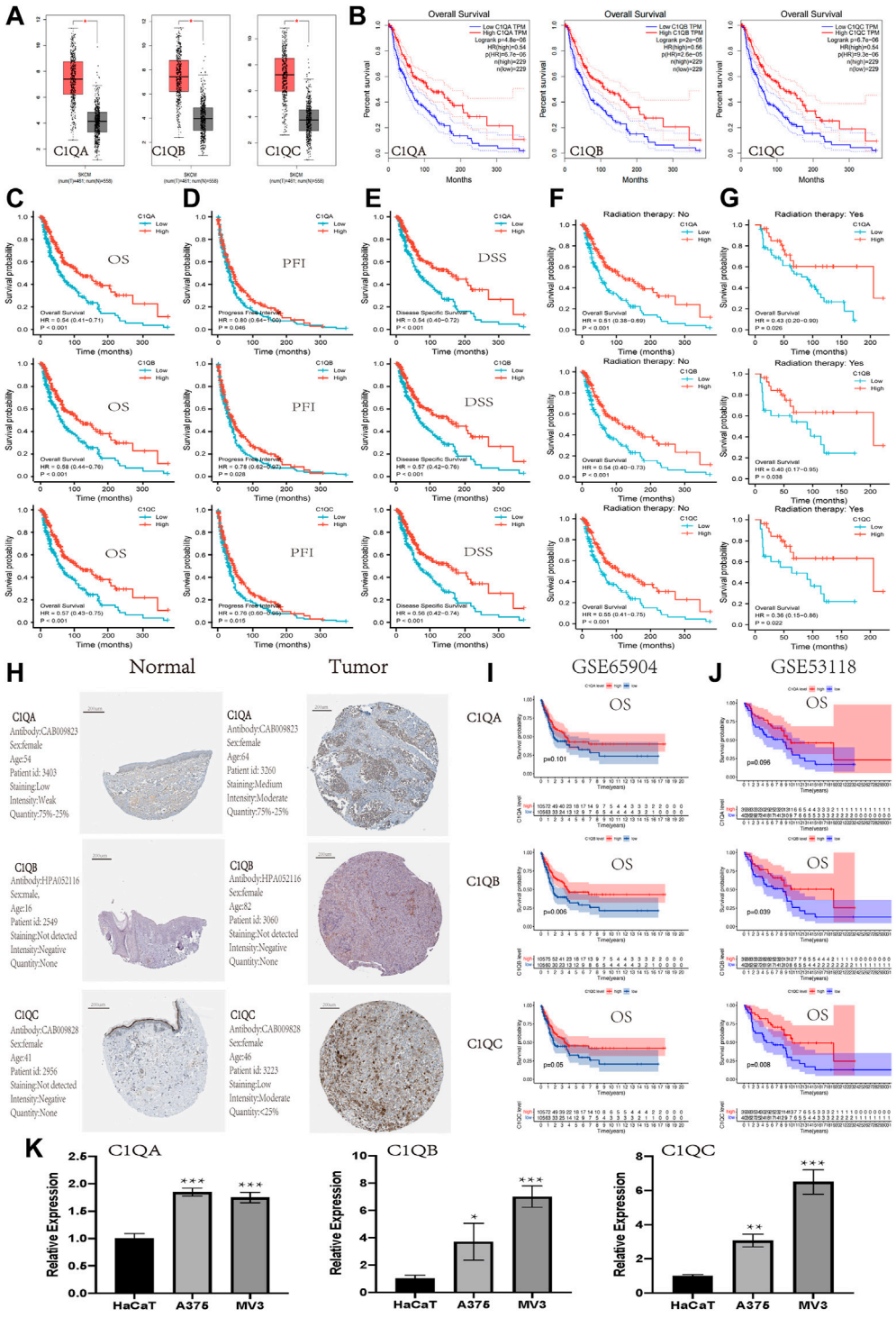


FIGURE 3
Expression and prognostic value of C1QA, C1QB, and C1QC in SKCM. **(A-B)** Expression and prognostic significance of C1QA, C1QB, and C1QC in SKCM analysed using GEPIA. The prognostic significance of C1QA, C1QB, and C1QC in SKCM from TCGA: **(C)** overall survival; **(D)** progress-free interval; **(E)** disease-specific survival; **(F)** without radiation therapy; and **(G)** with radiation therapy. **(H)** Protein expression of C1QA, C1QB, and C1QC in cutaneous melanoma from The Human Protein Atlas. **(I,J)** Prognostic significance of C1QA, C1QB, and C1QC in SKCM from GSE65904 and GSE53118: overall survival. **(K)** Relative mRNA expression of C1QA, C1QB, and C1QC in HaCaT, A375, and MV3 cells *via* RT-qPCR analysis (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$).

TABLE 3 Univariate and multivariate Cox regression analyses of C1QA expression and clinicopathological characteristics with OS in SKCM patients.

Characteristic	Total (n)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	*p*-value	Hazard ratio (95% CI)	*p*-value
T stage	361				
T1 and T2	118	Reference			
T3 and T4	243	2.085 (1.501–2.895)	**<0.001**	1.807 (1.267–2.575)	**0.001**
N stage	402				
N0	224	Reference			
N1 and N2 and N3	178	1.752 (1.304–2.354)	**<0.001**	2.821 (0.991–8.033)	0.052
M stage	430				
M0	406	Reference			
M1	24	1.897 (1.029–3.496)	**0.040**	2.744 (1.138–6.617)	**0.025**
Pathological stage	410				
Stage I and Stage II	217	Reference			
Stage III and Stage IV	193	1.617 (1.207–2.165)	**0.001**	0.660 (0.228–1.910)	0.443
Gender	456				
Female	172	Reference			
Male	284	1.172 (0.879–1.563)	0.281		
Age	456				
<=60	246	Reference			
>60	210	1.656 (1.251–2.192)	<0.001	1.282 (0.918–1.789)	0.144
radiation therapy	450				
no	374	reference			
yes	76	0.977 (0.694–1.377)	0.895		
C1QA	456				
Low	227	Reference			
High	229	0.540 (0.411–0.708)	**<0.001**	0.561 (0.405–0.776)	**<0.001**

Abbreviation: CI, confidence interval.
Bold values: $p < 0.05$.

DEGs were positively correlated with C1QA, C1QB, and C1QC (Figure 6A). The top 50 genes that were positively or negatively associated with C1QA, C1QB, and C1QC is shown in the heat map (Figures 6B, C). Based on a correlation greater than 0.7, we uploaded 214 C1QA-related DEGs, 213 C1QB-related DEGs, and 202 C1QC-related DEGs to the STRING database to analyse the interactions between the proteins they encode. The ten hub genes with the most interactions in each group were obtained using the Cytoscape tool (Figure 6D). We identified 30 hub genes from the three groups. After removing duplicate genes, 11 hub genes were used for subsequent analyses. Function and pathway enrichment analyses of these 11 hub genes were performed using the Metascape website. As shown in Figure 6E, hub genes were mainly involved in immune system-related biological processes, such as positive regulation of myeloid leukocyte-mediated immunity, external side of the plasma membrane, lymphocyte activation, positive regulation of lymphocyte activation, immune effector process, and immune receptor activity. In addition, all hub genes were found to have a good prognosis in SKCM using TIMER (Figure 6F). Based on these data, we inferred that the relationship between central genes may contribute to C1QA, C1QB, and C1QC, thus prolonging survival in patients with cutaneous melanoma.

Positive correlation with immune cell infiltration

C1Q is a member of the complement system and plays a role in a range of immunobiological processes (Kouser et al., 2015; Son et al., 2015). Therefore, we used TIMER to determine whether the expression of the three genes in SKCM was related to the level of immune cell infiltration. At various copy counts of C1QA, C1QB, and C1QC, we discovered substantial differences in immune cell

TABLE 4 Univariate and multivariate Cox regression analyses of C1QB expression and clinicopathological characteristics with OS in SKCM patients.

Characteristic	Total (n)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	*p*-value	Hazard ratio (95% CI)	*p*-value
T stage	361				
T1 and T2	118	Reference			
T3 and T4	243	2.085 (1.501–2.895)	**<0.001**	1.859 (1.305–2.648)	**<0.001**
N stage	402				
N0	224	Reference			
N1 and N2 and N3	178	1.752 (1.304–2.354)	**<0.001**	2.605 (0.922–7.362)	0.071
M stage	430				
M0	406	Reference			
M1	24	1.897 (1.029–3.496)	**0.040**	2.208 (0.927–5.257)	0.074
Pathological stage	410				
Stage I and Stage II	217	Reference			
Stage III and Stage IV	193	1.617 (1.207–2.165)	**0.001**	0.695 (0.243–1.990)	0.498
Sex	456				
Female	172	Reference			
Male	284	1.172 (0.879–1.563)	0.281		
Age	456				
<=60	246	Reference			
>60	210	1.656 (1.251–2.192)	<0.001	1.257 (0.901–1.753)	0.178
radiation therapy	450				
no	374	reference			
yes	76	0.977 (0.694–1.377)	0.895		
C1QB	456				
Low	229	Reference			
High	227	0.578 (0.441–0.758)	**<0.001**	0.593 (0.429–0.820)	**0.002**

Abbreviation: CI, confidence interval.
Bold values: $p < 0.05$.

infiltration (Figure 7A). We then investigated the relationship between the expression levels of these three genes and immune cell infiltration (Figure 7B). We concluded that the expression levels of C1QA, C1QB, and CIQC in SKCM were positively correlated with B cells, $CD8^+$ and $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells. In addition, B cells, $CD8^+$ T cells, neutrophils, and dendritic cells were defence factors associated with the cumulative survival over time in SKCM (Figure 7C).

These findings suggest that high expression levels of C1QA, C1QB, and C1QC are significantly associated with immune cell infiltration in patients with melanoma and may contribute to improved patient outcomes. ssGSEA was used to further verify whether the expression of the three genes affected the infiltration of the 24 immune cell subsets in SKCM. As shown in Figures 7D, E, the expression of C1QA, C1QB, and C1QC was strongly associated with 23 types of immune cells. Based on these results, we can speculate that C1QA, C1QB, and C1QC play key roles in immune cell infiltration in the tumour microenvironment of SKCM.

C1QA, C1QB, and C1QC were positively correlated with immune cell markers

To further understand the relationship between C1QA, C1QB, and C1QC, and various immune infiltrating cells, we used GEPIA to analyse the correlation between these three genes and various immune markers in SKCM. As shown in Table 6, in SKCM, C1QA, C1QB, and C1QC were all positively correlated with immune markers, such as B cells (CD19 and CD79A), $CD8^+$ T cells (CD8A and CD8B), $CD4^+$ T cells (CD4), M1 macrophages (IRF5), M2 macrophages (CD163, VSIG4, and MS4A4A), immune markers of neutrophils (ITGAM and CCR7), and immune markers of dendritic cells (HLA-DPB1, HLA-DQB1, HLA-DRA, HLA-DPA1, CD1C, NRP1, and ITGAX).

TABLE 5 Univariate and multivariate Cox regression analysis of C1QC expression and clinicopathological characteristics with OS in SKCM patients.

Characteristic	Total (n)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	*p*-value	Hazard ratio (95% CI)	*p*-value
T stage	361				
T1 and T2	118	Reference			
T3 and T4	243	2.085 (1.501–2.895)	**<0.001**	1.839 (1.291–2.621)	**<0.001**
N stage	402				
N0	224	Reference			
N1 and N2 and N3	178	1.752 (1.304–2.354)	**<0.001**	2.894 (1.023–8.192)	**0.045**
M stage	430				
M0	406	Reference			
M1	24	1.897 (1.029–3.496)	**0.040**	2.509 (1.051–5.988)	**0.038**
Pathological stage	410				
Stage I and Stage II	217	Reference			
Stage III and Stage IV	193	1.617 (1.207–2.165)	**0.001**	0.645 (0.225–1.854)	0.416
Sex	456				
Female	172	Reference			
Male	284	1.172 (0.879–1.563)	0.281		
Age	456				
<=60	246	Reference			
>60	210	1.656 (1.251–2.192)	<0.001	1.267 (0.908–1.769)	0.164
radiation therapy	450				
no	374	reference			
yes	76	0.977 (0.694–1.377)	0.895		
C1QC	456				
Low	226	Reference			
High	230	0.569 (0.434–0.745)	**<0.001**	0.607 (0.440–0.837)	**0.002**

Abbreviation: CI, confidence interval.
Bold values: $p < 0.05$.

Positive correlation with immune checkpoints

It is well known that in the field of tumour immunotherapy, five immune checkpoints, namely, PD-1, CTLA-4, Tim-3, LAG-3, and TIGIT, have inhibitory effects on T-cell proliferation. Immunosuppressive molecules are expressed in many different types of tumours, and blocking these molecules boosts the activation of immune cells. Therefore, we assessed the relationship between C1QA, C1QB, and C1QC and the five immune checkpoints using TIMER. Figure 7 shows that PD-1 (PDCD1), PD-L1 (CD274), CTLA-4 (CTLA4), Tim-3 (HAVCR2), LAG-3 (LAG3), and TIGIT immunosuppressive molecules were significantly and positively correlated with C1QA, C1QB, and C1QC expression. Analysis of the GEPIA data also verified a substantial positive relationship between these three genes and immunosuppressive molecules in SKCM (Figure 8). These results suggest that judicious use of inhibitors can help enhance the anti-tumour immune response and thus provide actual benefits to patients.

Gene set enrichment analysis

Finally, we identified the corresponding enrichment pathways in the high- and low-expression groups of C1QA, C1QB, and C1QC using GSEA. Surprisingly, the signalling pathways enriched by the three genes in the high-expression group were almost identical (Figure 9), including "antigen processing and presentation," "B-cell receptor signalling pathway," "T-cell receptor signalling pathway," "natural killer cell-mediated cytotoxicity," "leukocyte transendothelial migration," "Fc gamma R-mediated phagocytosis," "cytokine–cytokine receptor interaction," and "chemokine signalling pathway." In addition to their important roles in anti-tumour immunity, C1QA, C1QB, and C1QC were also significantly enriched in apoptotic pathways.

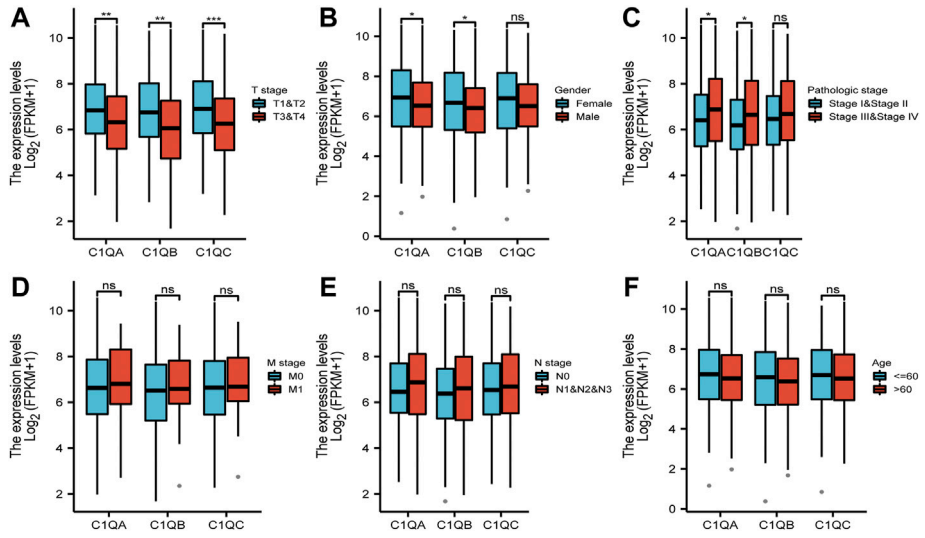


FIGURE 4
Correlation between differential expression of C1QA, C1QB, and C1QC and clinicopathological features: **(A)** T stage; **(B)** sex; **(C)** pathological stage; **(D)** M stage; **(E)** N stage; and **(F)** age (ns = no significance; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

Discussion

SKCM is the most life-threatening form of skin neoplasm, affecting millions of individuals worldwide each year (Corrie et al., 2014). SKCM often has a poor prognosis because of its high resistance to radiation and chemotherapeutic drugs (Pollack et al., 2011; Polkowska et al., 2016; Sandri et al., 2016). Elucidating the molecular mechanisms involved in the carcinogenesis of SKCM may contribute to the study of potential therapeutic targets and the development of valuable diagnostic and prognostic biomarkers. Complement is an important part of the innate immune system, which plays a role in resisting pathogen invasion and protecting the host (Roumenina et al., 2019a). Although C1Q has been proved to promote cell death in most types of cancers, in some cancers, the role of C1Q is completely opposite, which leads to cancer progression (Revel et al., 2022). In clear-cell renal cell carcinoma, tumour-associated macrophages generate C1r, C1s, C4, and C3, after being hijacked by cancer cells, and then initiates the classical pathway of the complement cascade of IgG immunity in tumour and promotes inflammation and T-cell exhaustion, leading to tumour progression (Roumenina et al., 2019b). Also, a high expression of C1Q is positively correlated with a poor prognosis in glioma of different grades (Mangogna et al., 2019a). In ovarian cancer, C1Q induced apoptosis in the SKOV3 cell line *via* the TNF-α-induced apoptosis pathway up-regulated by Bax and Fas, confirming the potential protective effect of C1Q Kaur et al. (2016). It is worth mentioning that C1Q-deficient mice carrying homologous B16 melanomas had slower tumour growth and longer survival than wild-type or C3- or C5-deficient mice. This study confirmed the role of C1Q in promoting melanoma growth. However, there are few reports about the specific role of C1Q in melanoma patients (Bulla et al., 2016). In recent years, with the rapid development of bioinformatics, gene exploration based on public databases has been widely used in medical research and plays an active role in clinical diagnosis, treatment, and prognosis prediction. Based on the existing public database and the validation of basic experiments, this study deeply explored the role of C1Q expression in diagnosis, prognosis and tumour immunity of SKCM patients, aiming to provide a reference for further research on the specific mechanism of C1Q in the future.

In the present study, we used the GEPIA database to conduct pan-cancer analysis on the transcription levels of C1QA, C1QB, and C1QC and screened out the cancer types with the differential expression of these three genes. We then performed a survival analysis of these differentially expressed cancers using the TIMER database and found that in SKCM, the expression levels of C1QA, C1QB, and C1QC were significantly correlated with OS. Based on this conclusion, we performed a more comprehensive survival analysis of C1QA, C1QB, and C1QC using expression data and clinical data from TCGA-SKCM. The results showed that the high expression of C1QA, C1QB, and C1QC in SKCM patients predicted a better prognosis and was not affected by radiation therapy. To ensure that the results were more reliable, we also analysed and verified OS and differential expression using GEO and HPA databases, verified differential expression using RT-qPCR, and reached a consistent conclusion. In addition, the current study found significant differences in the expression of C1QA, C1QB, and C1QC in

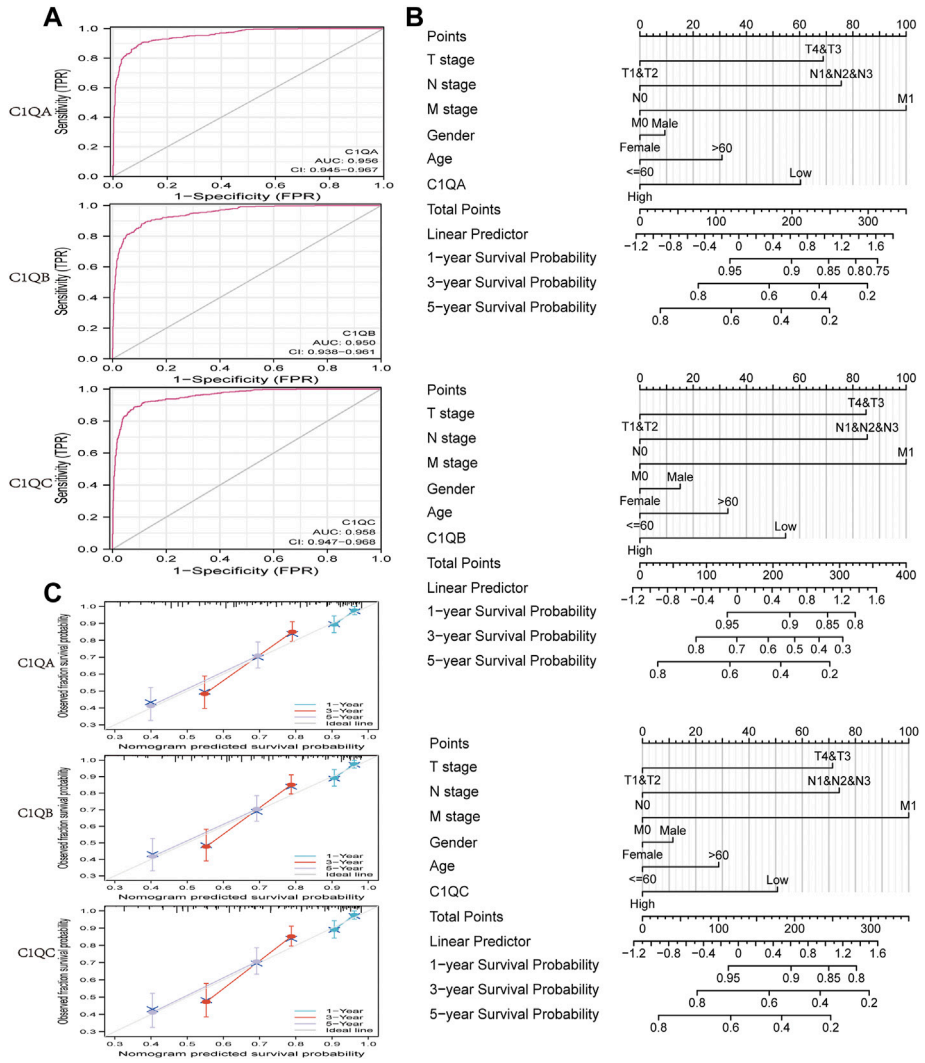


FIGURE 5
ROC curve and nomogram. **(A)** ROC curve in normal tissue and cutaneous melanoma. **(B)** Nomogram for predicting the probability of SKCM patients with 1-, 3-, and 5-year OS. **(C)** Calibration curves of the nomogram between predicted and actual 1-, 2-, and 3-year OS.

patients with different T and pathological stages; the earlier these stages, the higher the expression of C1QA, C1QB, and C1QC and the better the prognosis. Therefore, these complement components are considered to play a role in inhibiting tumour development to a certain extent. Finally, multivariate Cox regression analysis and ROC analysis confirmed that C1QA, C1QB, and C1QC can be used as independent prognostic biomarkers and contribute to the diagnosis of SKCM.

The nomogram scale has been widely used to predict the OS of cancer patients. Based on C1QA, C1QB, C1QC, and six clinical parameters (sex, age, pathological stage, T stage, N stage, and M stage), we created individual nomograms to predict the 1-, 3-, and 5-year OS of SKCM patients and verified the accuracy of nomograms by the C-index and calibration curve. In addition, among the 11 identified C1QA/C1QB/C1QB-related hub genes, we found that their overall function was mapped to immune-related activity, and the high expression of these hub genes predicted better OS. These results suggest that these hub genes may have a synergistic effect on C1QA, C1QB, and C1QC and improve the prognosis of SKCM patients through immune response function.

In recent years, the role of the immune system in cancer development has been studied extensively. Immunotherapy, including vaccines, adoptive cell transfer, and immune detection blockers, has rapidly developed as a promising therapeutic strategy for malignant tumours. Previous studies have reported that tumour immune cell infiltration can improve the prognosis of SKCM and efficacy of

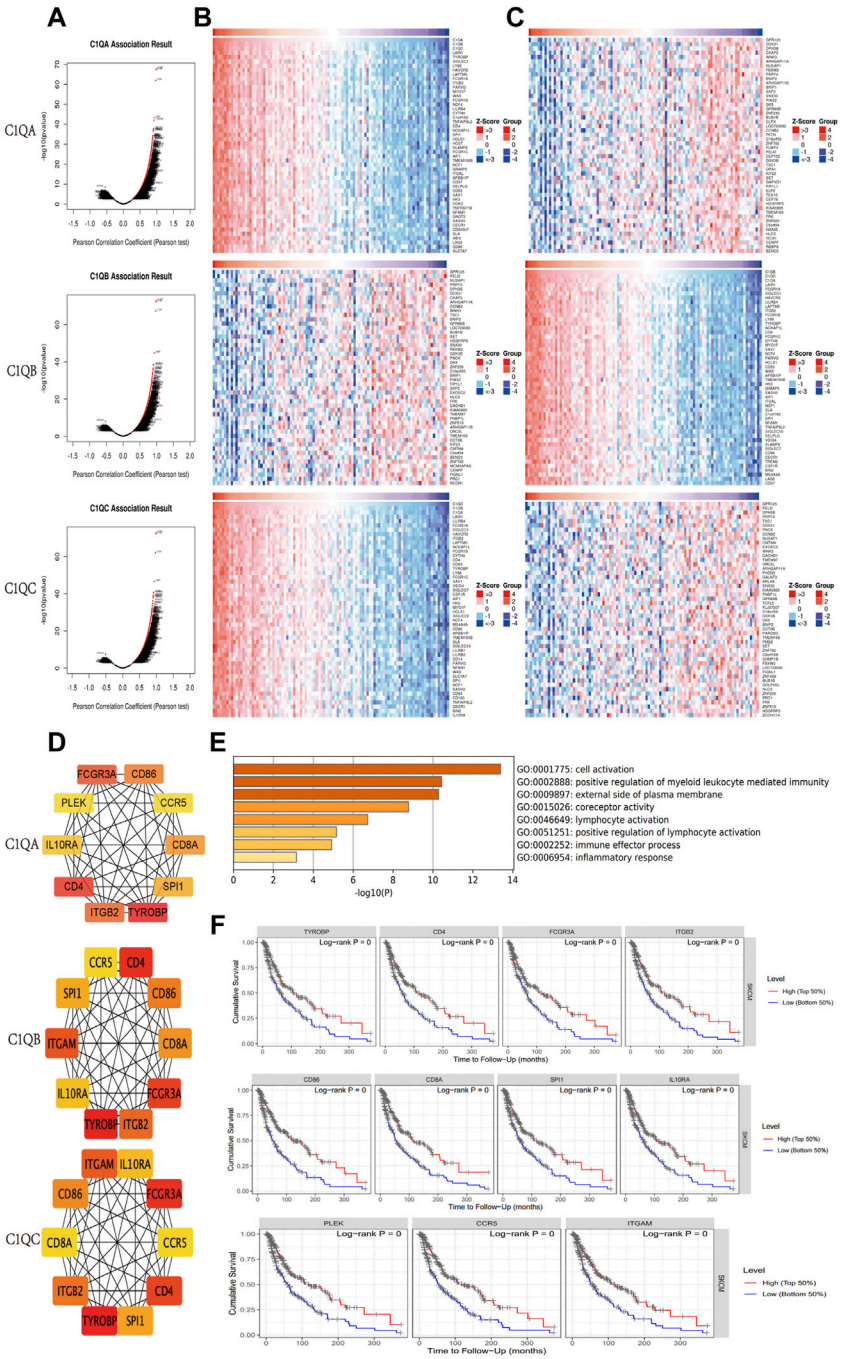


FIGURE 6
Genes correlated with C1QA, C1QB, and C1QC in SKCM. **(A)** Correlations between C1QA, C1QB, and C1QC and differently expressed genes. **(B,C)** Top 50 positively or negatively correlated genes. **(D)** Top 10 hub genes. **(E)** GO analysis of 11 hub genes. **(F)** Effect of 11 hub genes on OS in patients with SKCM.

immunotherapy (Fu et al., 2019; Selitsky et al., 2019). Adoptive cell therapy with autologous tumour-infiltrating lymphocytes is an effective treatment for metastatic SKCM (Rosenberg et al., 2011). To further understand the mechanism of C1QA, C1QB, and C1QC in SKCM carcinogenesis, we explored the relationship between C1QA, C1QB, and C1QC and immune cell infiltration. Our results showed that the expression levels of C1QA, C1QB, and C1QC in SKCM positively correlated with B cells, $CD8^+$ and

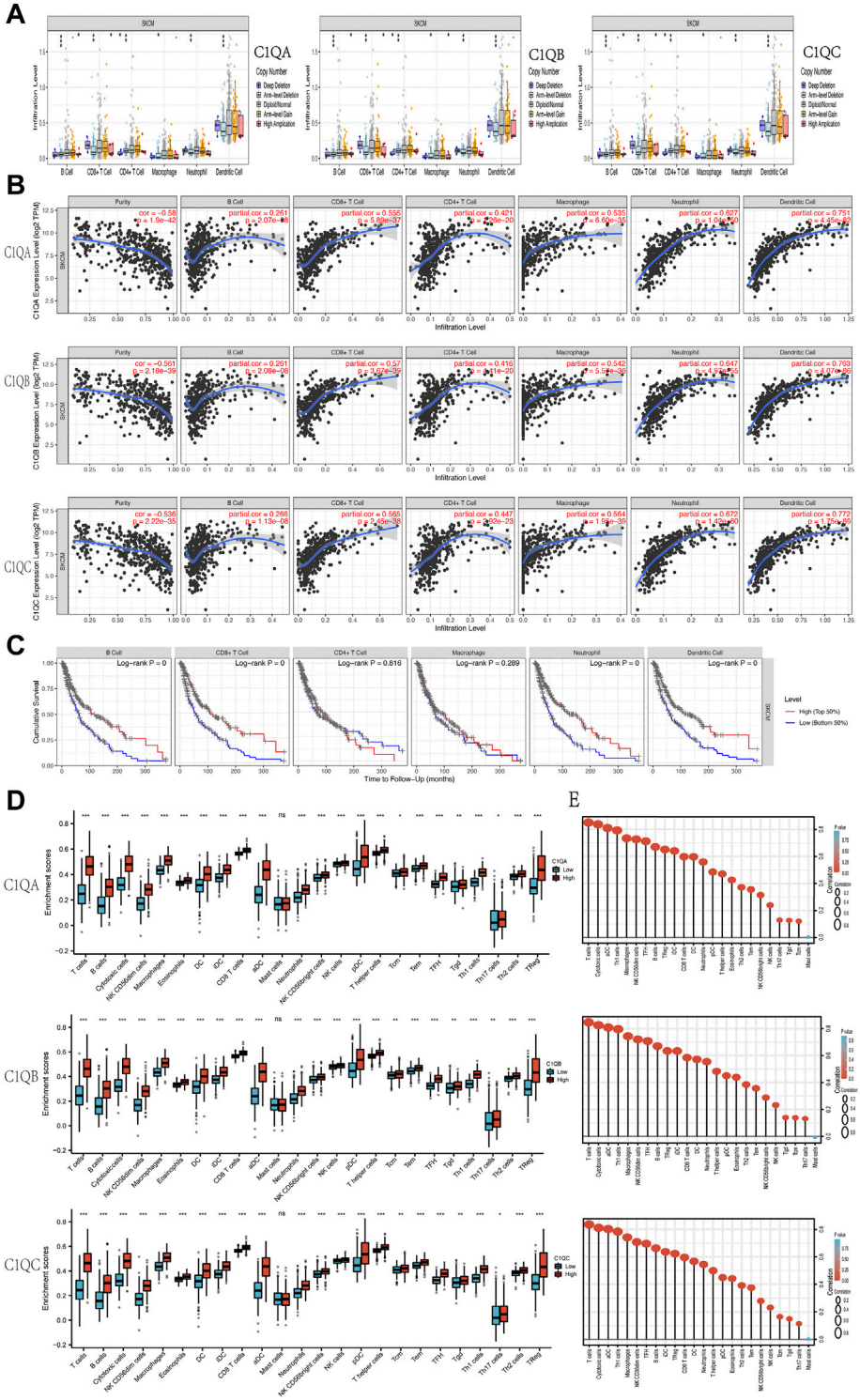


FIGURE 7
Correlation analysis of C1QA, C1QB, and C1QC expression with tumour-infiltering immune cells in SKCM. **(A)** Relationships between infiltration levels of six immune cells and copy number of C1QA, C1QB, and C1QC. **(B)** Relationships between infiltration levels of six immune cells and C1QA, C1QB, and C1QC expression. **(C)** Relationships between infiltration levels of six immune cells and SKCM prognosis. **(D)** Difference of infiltration levels of 24 immune cells between C1QA, C1QB, and C1QC high-expression and low-expression groups. **(E)** Correlation between the relative abundances of 24 immune cells and C1QA, C1QB, and C1QC expression levels (ns = no significance; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

TABLE 6 Correlation analysis between C1QA, C1QB, and C1QC and immune cell markers in SKCM using GEPIA.

Immune cell	Biomarker	C1QA R	*p*	C1QB R	*p*	C1QC R	*p*
B cell	CD19	0.58	***	0.58	***	0.57	***
	CD79A	0.64	***	0.63	***	0.61	***
$CD8^+$ T cell	CD8A	0.86	***	0.86	***	0.84	***
	CD8B	0.84	***	0.83	***	0.81	***
$CD4^+$ T cell	CD4	0.88	***	0.89	***	0.89	***
M1 macrophage	INOS (NOS2)	−0.018	0.71	−0.012	0.79	0.0071	0.88
	IRF5	0.67	***	0.67	***	0.68	***
	COX2 (PTGS2)	−0.049	0.3	−0.048	0.3	−0.039	0.4
M2 macrophage	CD163	0.83	***	0.84	***	0.85	***
	VSIG4	0.84	***	0.86	***	0.86	***
	MS4A4A	0.86	***	0.87	***	0.89	***
Neutrophil	CD66b (CEACAM8)	−0.021	0.65	−0.017	0.71	−0.014	0.76
Dendritic cell	CD11b (ITGAM)	0.77	***	0.78	***	0.79	***
	CCR7	0.66	***	0.65	***	0.65	***
	HLA-DPB1	0.84	***	0.84	***	0.83	***
	HLA-DQB1	0.68	***	0.67	***	0.66	***
	HLA-DRA	0.85	***	0.85	***	0.84	***
	HLA-DPA1	0.82	***	0.82	***	0.82	***
	BDCA-1 (CD1C)	0.51	***	0.5	***	0.5	***
	BDCA-4 (NRP1)	0.35	***	0.37	***	0.4	***
	CD11c (ITGAX)	0.65	***	0.65	***	0.64	***

Cor, R-value of Spearman's correlation; *$p < 0.01$; **$p < 0.001$; ***$p < 0.0001$.

$CD4^+$ T cells, macrophages, neutrophils, dendritic cells, and their biomarkers. Among these, B cells, $CD8^+$ T cells, neutrophils, and dendritic cells play an active role in prolonging the OS of SKCM patients. Evidence confirms that tumour-infiltrating B cells play a key role in the anti-tumour immune response of SKCM (DiLillo et al., 2010; Gilbert et al., 2011). $CD8^+$ T cells are an important component of the tumour immune response and exert cytotoxic effects on melanoma cells (Durgeau et al., 2018). Dendritic cell tumour vaccines are the earliest and most effective methods for the clinical treatment of melanoma (Gao et al., 2016). Furthermore, the neutrophil-to-lymphocyte ratio can be used as a biomarker for metastatic SKCM (Sellers et al., 2019). Finally, ssGSEA further confirmed the high correlation between C1QA, C1QB, and C1QC and immune cell infiltration. Based on the aforementioned results, it can be concluded that C1QA, C1QB, and C1QC may be protective factors and prolong the OS of SKCM patients by mediating immune cell infiltration.

It is worth mentioning that immunotherapy can produce an effective anti-tumour immune response that requires both enough immune cells to infiltrate the tumour microenvironment and the full expression of immune checkpoints (Chae et al., 2018). Recently, immune checkpoint inhibitors (ICIs), including ipilimumab (a CTLA-4 inhibitor) and nivolumab (a PD-1 inhibitor), have shown remarkable therapeutic effects on SKCM (Eggermont et al., 2018). Some studies have reported that PD-L1 in SKCM cells plays a key role by downregulating the function of T cells after binding to PD-1 on lymphocytes (Hamid et al., 2013). However, the main drawback of these ICIs is their low response efficiencies (Burtness et al., 2019). Therefore, patients with high PD-L1 expression tend to have better treatment outcomes than patients with low PD-L1 expression (Daud et al., 2016; Weide et al., 2016). The infiltration of massive $CD8^+$ T cells within melanoma cells can enhance the clinical benefits of ICIs (Tumeh et al., 2014; Linette and Carreno, 2019; Wong et al., 2019). More importantly, the combination of anti-CTLA-4 and anti-PD-1/PD-L1 showed stronger anti-tumour activity than monotherapy (Ott et al., 2013; Wolchok et al., 2013; Larkin et al., 2015). In our analysis of immune-related functions, the expression levels of C1QA, C1QB, and C1QC were positively correlated not only with PD1, PD-L1, and CTLA-4 but also with the molecules of multiple immune checkpoints. This suggests that SKCM patients with high C1QA, C1QB, and C1QC expression levels can benefit more from immunotherapy. In addition, we analysed the signalling pathways in which C1QA, C1QB, and C1QC are involved in SKCM through GSEA. These findings indicated that signalling pathways related to immune function were significantly enriched in patients with high expression of C1QA, C1QB, and C1QC. Moreover, these three genes have been found to be involved in apoptosis. Previous reports have

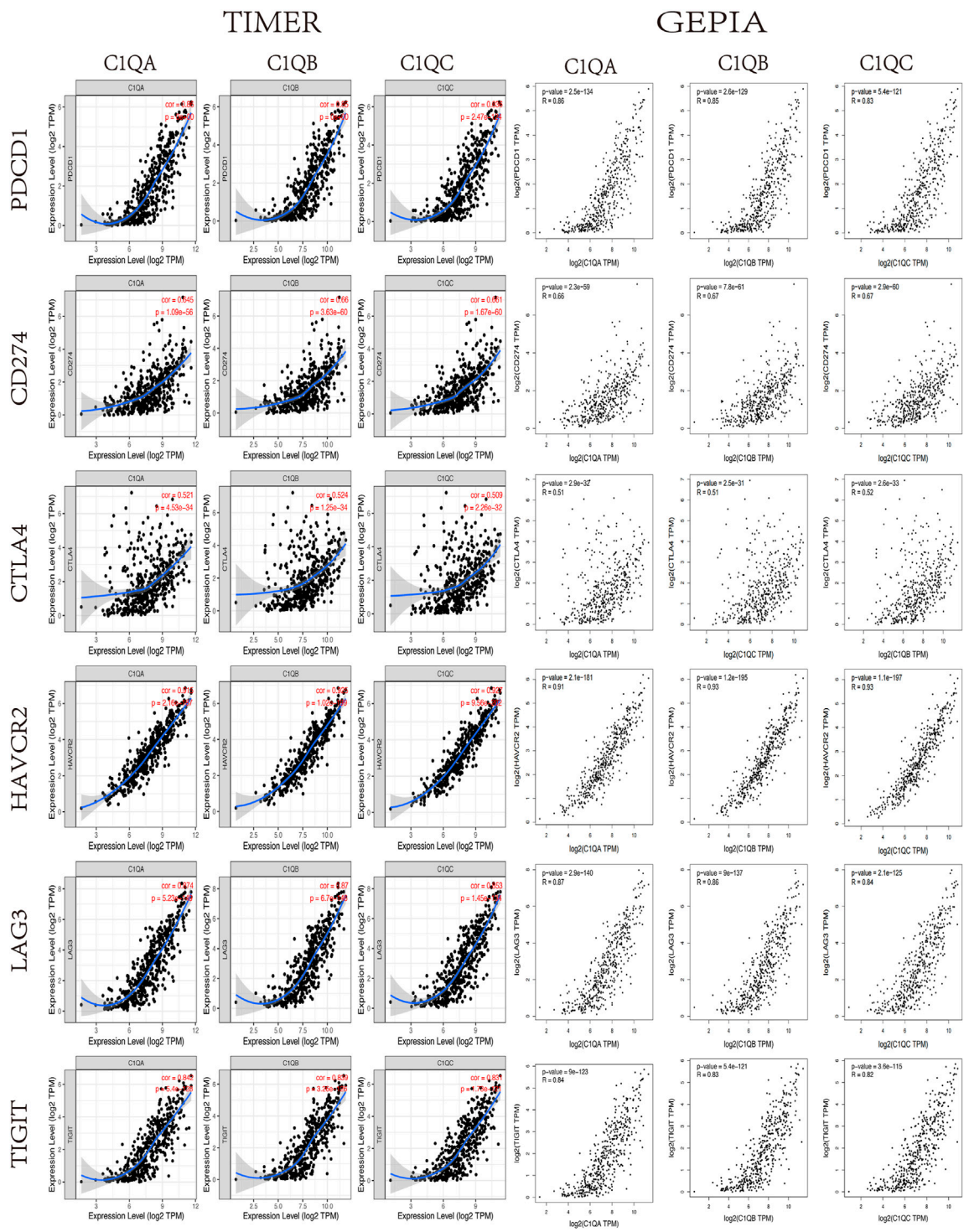


FIGURE 8
Correlation between C1QA, C1QB, and C1QC expression levels and immune checkpoints in SKCM *via* TIMER analysis and GEPIA.

confirmed that C1Q can lead to apoptosis in ovarian, prostate, and breast cancer (Hong et al., 2009; Bandini et al., 2016; Kaur et al., 2016). Therefore, we speculated that the high expression of C1QA, C1QB, and C1QC may play a dual regulatory role through anti-tumour immunity and the induction of cancer cell apoptosis.

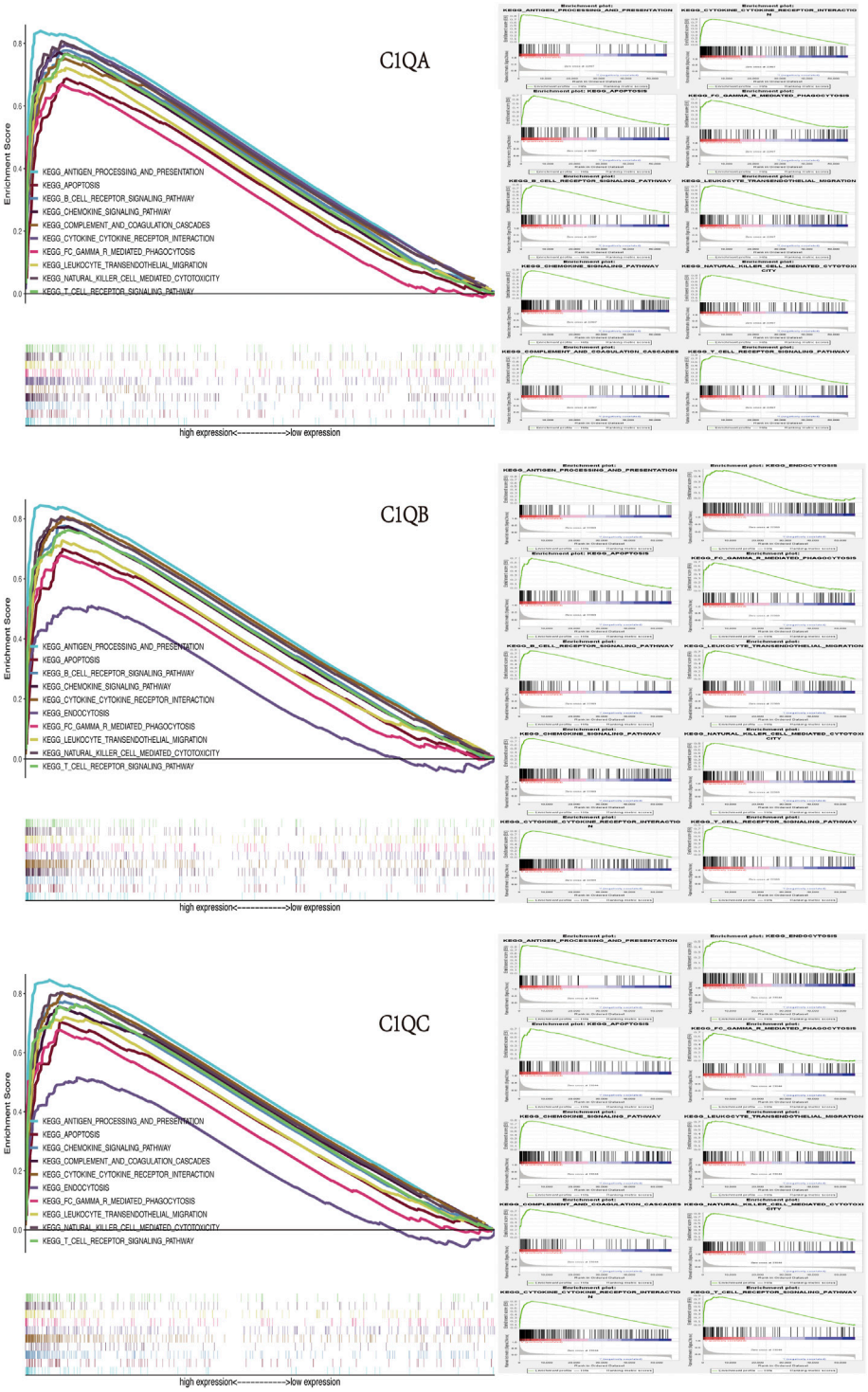


FIGURE 9
GSEA of C1QA, C1QB, and C1QC high-expression groups revealed enriched biological pathways in SKCM.

In conclusion, the current study showed that C1QA, C1QB, and C1QC were highly expressed in a variety of cancers, especially in SKCM, which overexpressed these three complement components and predicted better survival prognosis. In addition, C1QA, C1QB, and C1QC increased immune cell infiltration and immune checkpoint

expression in SKCM, helping patients benefit more from immunotherapy. However, the present study has some limitations. Our analysis and validation are mainly based on bioinformatics and melanoma cell lines. Therefore, in future studies, we need to perform *in vivo* experiments to confirm our outcomes.

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

HY and DC conceived and designed this study. HY and DC wrote the manuscript. HY and YG performed the data analysis and revised the manuscript critically.

Funding

This research was supported by the Key Research and Development Program of Anhui Province (No. 202004a07020036).

Acknowledgments

The authors thank DC for providing technical guidance and all authors who participated in this study and thank TCGA, GEO, TIMER, STRING, and GEPIA databases for providing valuable data sets. The authors also thank the Key Research and Development Program of Anhui Province for its financial support.

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.

References

Agostinis, C., Vidergar, R., Belmonte, B., Mangogna, A., Amadio, L., Geri, P., et al. (2017). Complement protein C1q binds to hyaluronic acid in the malignant pleural mesothelioma microenvironment and promotes tumor growth. *Front. Immunol.* 8, 1559. doi:10.3389/fimmu.2017.01559

Bandini, S., Macagno, M., Hysi, A., Lanzardo, S., Conti, L., Bello, A., et al. (2016). The non-inflammatory role of C1q during Her2/neu-driven mammary carcinogenesis. *Oncoimmunology* 5 (12), e1253653. doi:10.1080/2162402X.2016.1253653

Berwick, M., Buller, D. B., Cust, A., Gallagher, R., Lee, T. K., Meyskens, F., et al. (2016). Melanoma epidemiology and prevention. *Cancer Treat. Res.* 167, 17–49. doi:10.1007/978-3-319-22539-5_2

Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A. C., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 39 (4), 782–795. doi:10.1016/j.immuni.2013.10.003

Bloethner, S., Scherer, D., Drechsel, M., Hemminki, K., and Kumar, R. (2009). Malignant melanoma-a genetic overview. *Actas Sifiliogr.* 100 (1), 38–51. doi:10.1016/s0001-7310(09)73167-9

Borden, E. S., Adams, A. C., Buetow, K. H., Wilson, M. A., Bauman, J. E., Curiel-Lewandrowski, C., et al. (2021). Shared gene expression and immune pathway changes associated with progression from nevi to melanoma. *Cancers* 14 (1), 3. doi:10.3390/cancers14010003

Bulla, R., Tripodo, C., Rami, D., Ling, G. S., Agostinis, C., Guarnotta, C., et al. (2016). C1q acts in the tumour microenvironment as a cancer-promoting factor independently of complement activation. *Nat. Commun.* 7, 10346. doi:10.1038/ncomms10346

Burtness, B., Harrington, K. J., Greil, R., Soulieres, D., Tahara, M., de Castro, G., Jr., et al. (2019). Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-048): A randomised, open-label, phase 3 study. *Lancet* 394 (10212), 1915–1928. doi:10.1016/S0140-6736(19)32591-7

Caini, S., Raimondi, S., Johansson, H., De Giorgi, V., Zanna, I., Palli, D., et al. (2015). Telomere length and the risk of cutaneous melanoma and non-melanoma skin cancer: A review of the literature and meta-analysis. *J. Dermatol. Sci.* 80 (3), 168–174. doi:10.1016/j.jdermsci.2015.08.003

Chae, Y. K., Arya, A., Iams, W., Cruz, M. R., Chandra, S., Choi, J., et al. (2018). Current landscape and future of dual anti-CTLA4 and PD-1/PD-L1 blockade immunotherapy in cancer; lessons learned from clinical trials with melanoma and non-small cell lung cancer (NSCLC). *J. Immunother. Cancer* 6 (1), 39. doi:10.1186/s40425-018-0349-3

Corrie, P., Hategan, M., Fife, K., and Parkinson, C. (2014). Management of melanoma. *Br. Med. Bull.* 111 (1), 149–162. doi:10.1093/bmb/ldu019

Darmawan, C. C., Jo, G., Montenegro, S. E., Kwak, Y., Cheol, L., Cho, K. H., et al. (2019). Early detection of acral melanoma: A review of clinical, dermoscopic, histopathologic, and molecular characteristics. *J. Am. Acad. Dermatol.* 81 (3), 805–812. doi:10.1016/j.jaad.2019.01.081

Daud, A. I., Wolchok, J. D., Robert, C., Hwu, W. J., Weber, J. S., Ribas, A., et al. (2016). Programmed death-ligand 1 expression and response to the anti-programmed death 1 antibody pembrolizumab in melanoma. *J. Clin. Oncol.* 34 (34), 4102–4109. doi:10.1200/JCO.2016.67.2477

DiLillo, D. J., Yanaba, K., and Tedder, T. F. (2010). B cells are required for optimal CD4+ and CD8+ T cell tumor immunity: therapeutic B cell depletion enhances B16 melanoma growth in mice. *J. Immunol.* 184 (7), 4006–4016. doi:10.4049/jimmunol.0903009

Dominiak, N. R., Wick, M. R., and Smith, M. T. (2016). Mucosal melanomas: Site-specific information, comparisons with cutaneous tumors, and differential diagnosis. *Semin. Diagn. Pathol.* 33 (4), 191–197. doi:10.1053/j.semdp.2016.04.003

Durgeau, A., Virk, Y., Corgnac, S., and Mami-Chouaib, F. (2018). Recent advances in targeting CD8 T-cell immunity for more effective cancer immunotherapy. *Front. Immunol.* 9, 14. doi:10.3389/fimmu.2018.00014

Eggermont, A. M. M., Crittenden, M., and Wargo, J. (2018). Combination immunotherapy development in melanoma. *Am. Soc. Clin. Oncol. Educ. Book.* 38, 197–207. doi:10.1200/EDBK_201131

Espericueta, V., Manughian-Peter, A. O., Bally, I., Thielens, N. M., and Fraser, D. A. (2020). Recombinant C1q variants modulate macrophage responses but do not activate the classical complement pathway. *Mol. Immunol.* 117, 65–72. doi:10.1016/j.molimm.2019.10.008

Fu, Q., Chen, N., Ge, C., Li, R., Li, Z., Zeng, B., et al. (2019). Prognostic value of tumor-infiltrating lymphocytes in melanoma: A systematic review and meta-analysis. *Oncoimmunology* 8 (7), 1593806. doi:10.1080/2162402X.2019.1593806

Gao, Y., Gao, W., Chen, X., Cha, N., Wang, X., Jia, X., et al. (2016). Enhancing the treatment effect on melanoma by heat shock protein 70-peptide complexes purified from human melanoma cell lines. *Oncol. Rep.* 36 (3), 1243–1250. doi:10.3892/or.2016.4947

Gilbert, A. E., Karagiannis, P., Dodev, T., Koers, A., Lacy, K., Josephs, D. H., et al. (2011). Monitoring the systemic human memory B cell compartment of melanoma patients for anti-tumor IgG antibodies. *PLoS One* 6 (4), e19330. doi:10.1371/journal.pone.0019330

Hamid, O., Robert, C., Daud, A., Hodi, F. S., Hwu, W. J., Kefford, R., et al. (2013). Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. *N. Engl. J. Med.* 369 (2), 134–144. doi:10.1056/NEJMoa1305133

Hanzelmann, S., Castelo, R., and Guinney, J. (2013). Gsva: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14, 7. doi:10.1186/1471-2105-14-7

Hong, Q., Sze, C. I., Lin, S. R., Lee, M. H., He, R. Y., Schultz, L., et al. (2009). Complement C1q activates tumor suppressor WWOX to induce apoptosis in prostate cancer cells. *PLoS One* 4 (6), e5755. doi:10.1371/journal.pone.0005755

Kaur, A., Sultan, S. H., Murugaiah, V., Pathan, A. A., Alhamlan, F. S., Karteris, E., et al. (2016). Human C1q induces apoptosis in an ovarian cancer cell line via tumor necrosis factor pathway. *Front. Immunol.* 7, 599. doi:10.3389/fimmu.2016.00599

Kishore, U., Gaboriaud, C., Waters, P., Shrive, A. K., Greenhough, T. J., Reid, K. B., et al. (2004). C1q and tumor necrosis factor superfamily: Modularity and versatility. *Trends Immunol.* 25 (10), 551–561. doi:10.1016/j.it.2004.08.006

Kishore, U., Thielens, N. M., and Gaboriaud, C. (2016). Editorial: State-of-the-Art research on C1q and the classical complement pathway. *Front. Immunol.* 7, 398. doi:10.3389/fimmu.2016.00398

Kouser, L., Madhukaran, S. P., Shastri, A., Saraon, A., Ferluga, J., Al-Mozaini, M., et al. (2015). Emerging and novel functions of complement protein C1q. *Front. Immunol.* 6, 317. doi:10.3389/fimmu.2015.00317

Kozovska, Z., Gabrisova, V., and Kucerova, L. (2016). Malignant melanoma: Diagnosis, treatment and cancer stem cells. *Neoplasma* 63 (4), 510–517. doi:10.4149/neo_2016_403

Kuk, D., Shoushtari, A. N., Barker, C. A., Panageas, K. S., Munhoz, R. R., Momtaz, P., et al. (2016). Prognosis of mucosal, uveal, acral, nonacral cutaneous, and unknown primary melanoma from the time of first metastasis. *Oncologist* 21 (7), 848–854. doi:10.1634/theoncologist.2015-0522

Larkin, J., Chiarion-Sileni, V., Gonzalez, R., Grob, J. J., Cowey, C. L., Lao, C. D., et al. (2015). Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *N. Engl. J. Med.* 373 (1), 23–34. doi:10.1056/NEJMoa1504030

Laskowski, J., Renner, B., Le Quintrec, M., Panzer, S., Hannan, J. P., Ljubanovic, D., et al. (2016). Distinct roles for the complement regulators factor H and Crry in protection of the kidney from injury. *Kidney Int.* 90 (1), 109–122. doi:10.1016/j.kint.2016.02.036

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). Timer: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 77 (21), e108–e110. doi:10.1158/0008-5472.CAN-17-0307

Linette, G. P., and Carreno, B. M. (2019). Tumor-infiltrating lymphocytes in the checkpoint inhibitor era. *Curr. Hematol. Malig. Rep.* 14 (4), 286–291. doi:10.1007/s11899-019-00523-x

Luo, Y., Robinson, S., Fujita, J., Siconolfi, L., Magidson, J., Edwards, C. K., et al. (2011). Transcriptome profiling of whole blood cells identifies PLEK2 and C1QB in human melanoma. *PLoS One* 6 (6), e20971. doi:10.1371/journal.pone.0020971

Mangogna, A., Agostinis, C., Bonazza, D., Belmonte, B., Zacchi, P., Zito, G., et al. (2019). Is the complement protein C1q a pro- or anti-tumorigenic factor? Bioinformatics analysis involving human carcinomas. *Front. Immunol.* 10, 865. doi:10.3389/fimmu.2019.00865

Mangogna, A., Belmonte, B., Agostinis, C., Zacchi, P., Iacopino, D. G., Martorana, A., et al. (2019). Prognostic implications of the complement protein C1q in gliomas. *Front. Immunol.* 10, 2366. doi:10.3389/fimmu.2019.02366

Nayak, A., Pednekar, L., Reid, K. B., and Kishore, U. (2012). Complement and non-complement activating functions of C1q: A prototypical innate immune molecule. *Innate Immun.* 18 (2), 350–363. doi:10.1177/1753425910396252

Ott, P. A., Hodi, F. S., and Robert, C. (2013). CTLA-4 and PD-1/PD-L1 blockade: New immunotherapeutic modalities with durable clinical benefit in melanoma patients. *Clin. Cancer Res.* 19 (19), 5300–5309. doi:10.1158/1078-0432.CCR-13-0143

Polkowska, M., Czepielewska, E., and Kozlowska-Wojciechowska, M. (2016). Drug combinations as the new standard for melanoma treatment. *Curr. Treat. Options Oncol.* 17 (12), 61. doi:10.1007/s11864-016-0436-y

Pollack, L. A., Li, J., Berkowitz, Z., Weir, H. K., Wu, X. C., Ajani, U. A., et al. (2011). Melanoma survival in the United States, 1992 to 2005. *J. Am. Acad. Dermatol.* 65 (1), S78–S86. doi:10.1016/j.jaad.2011.05.030

Rastrelli, M., Tropea, S., Rossi, C. R., and Alaibac, M. (2014). Melanoma: Epidemiology, risk factors, pathogenesis, diagnosis and classification. *Vivo* 28 (6), 1005–1011.

Revel, M., Sautès-Fridman, C., Fridman, W. H., and Roumenina, L. T. (2022). C1q+ macrophages: Passengers or drivers of cancer progression. *Trends Cancer* 8 (7), 517–526. doi:10.1016/j.trecan.2022.02.006

Rosenberg, S. A., Yang, J. C., Sherry, R. M., Kammula, U. S., Hughes, M. S., Phan, G. Q., et al. (2011). Durable complete responses in heavily pretreated patients with metastatic melanoma using T-cell transfer immunotherapy. *Clin. Cancer Res.* 17 (13), 4550–4557. doi:10.1158/1078-0432.CCR-11-0116

Roumenina, L. T., Daugan, M. V., Noé, R., Petitprez, F., Vano, Y. A., Sanchez-Salas, R., et al. (2019). Tumor cells hijack macrophage-produced complement C1q to promote tumor growth. *Cancer Immunol. Res.* 7 (7), 1091–1105. doi:10.1158/2326-6066.CIR-18-0891

Roumenina, L. T., Daugan, M. V., Petitprez, F., Sautès-Fridman, C., and Fridman, W. H. (2019). Context-dependent roles of complement in cancer. *Nat. Rev. Cancer* 19 (12), 698–715. doi:10.1038/s41568-019-0210-0

Sandri, S., Faiao-Flores, F., Tiago, M., Pennacchi, P. C., Massaro, R. R., Alves-Fernandes, D. K., et al. (2016). Vemurafenib resistance increases melanoma invasiveness and modulates the tumor microenvironment by MMP-2 upregulation. *Pharmacol. Res.* 111, 523–533. doi:10.1016/j.phrs.2016.07.017

Selitsky, S. R., Mose, L. E., Smith, C. C., Chai, S., Hoadley, K. A., Dittmer, D. P., et al. (2019). Prognostic value of B cells in cutaneous melanoma. *Genome Med.* 11 (1), 36. doi:10.1186/s13073-019-0647-5

Sellers, C. M., Uhlig, J., Ludwig, J. M., Stein, S. M., and Kim, H. S. (2019). Inflammatory markers in intrahepatic cholangiocarcinoma: Effects of advanced liver disease. *Cancer Med.* 8 (13), 5916–5929. doi:10.1002/cam4.2373

Son, M., Diamond, B., and Santiago-Schwarz, F. (2015). Fundamental role of C1q in autoimmunity and inflammation. *Immunol. Res.* 63 (1-3), 101–106. doi:10.1007/s12026-015-8705-6

Sun, Q., Lee, W., Mohri, Y., Takeo, M., Lim, C. H., Xu, X., et al. (2019). A novel mouse model demonstrates that oncogenic melanocyte stem cells engender melanoma resembling human disease. *Nat. Commun.* 10 (1), 5023. doi:10.1038/s41467-019-12733-1

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49 (D1), D605–D612. doi:10.1093/nar/gkaa1074

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). Gepia: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 45 (1), W98–W102. doi:10.1093/nar/gkx247

Tumeh, P. C., Harview, C. L., Yearley, J. H., Shintaku, I. P., Taylor, E. J., Robert, L., et al. (2014). PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 515 (7528), 568–571. doi:10.1038/nature13954

Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. *Science* 347 (6220), 1260419. doi:10.1126/science.1260419

Weide, B., Martens, A., Hassel, J. C., Berking, C., Postow, M. A., Bisschop, K., et al. (2016). Baseline biomarkers for outcome of melanoma patients treated with pembrolizumab. *Clin. Cancer Res.* 22 (22), 5487–5496. doi:10.1158/1078-0432.CCR-16-0127

Wolchok, J. D., Kluger, H., Callahan, M. K., Postow, M. A., Rizvi, N. A., Lesokhin, A. M., et al. (2013). Nivolumab plus ipilimumab in advanced melanoma. *N. Engl. J. Med.* 369 (2), 122–133. doi:10.1056/NEJMoa1302369

Wong, P. F., Wei, W., Smithy, J. W., Acs, B., Toki, M. I., Blenman, K. R. M., et al. (2019). Multiplex quantitative analysis of tumor-infiltrating lymphocytes and immunotherapy outcome in metastatic melanoma. *Clin. Cancer Res.* 25 (8), 2442–2449. doi:10.1158/1078-0432.CCR-18-2652

Yang, R., Huang, J., Ma, H., Li, S., Gao, X., Liu, Y., et al. (2019). Is complement C1q a potential marker for tumor burden and immunodeficiency in multiple myeloma? *Leuk. Lymphoma* 60 (7), 1812–1818. doi:10.1080/10428194.2018.1543883

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 30 August 2022
DOI 10.3389/fgene.2022.916024

Check for updates

OPEN ACCESS

EDITED BY
Ehsan Nazemalhosseini-Mojarad,
Shahid Beheshti University of Medical Sciences, Iran

REVIEWED BY
Nayeralsadat Fatemi,
Shahid Beheshti University of Medical Sciences, Iran
Nesa Kazemifard,
Shahid Beheshti University of Medical Sciences, Iran

*CORRESPONDENCE
Aimin Li,
mailto:liaimin2005@163.com

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

SPECIALTY SECTION
This article was submitted to *Cancer Genetics and Oncogenomics*, a section of the journal Frontiers in Genetics

RECEIVED 08 April 2022
ACCEPTED 28 June 2022
PUBLISHED 30 August 2022

CITATION
Peng Y, Wu G, Qiu X, Luo Y, Zou Y, Wei X and Li A (2022), Construction and validation of a necroptosis-related lncRNAs prognosis signature of hepatocellular carcinoma. *Front. Genet.* 13:916024. doi: 10.3389/fgene.2022.916024

COPYRIGHT
© 2022 Peng, Wu, Qiu, Luo, Zou, Wei and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Construction and validation of a necroptosis-related lncRNAs prognosis signature of hepatocellular carcinoma

YunZhen Peng[1†], GuoJing Wu[1†], Xin Qiu[2†], Yue Luo[1], YiShu Zou[1], XueYan Wei[1] and Aimin Li[1*]

[1]Cancer Center, Integrated Hospital of Traditional Chinese Medicine, Southern Medical University, Guangzhou, China, [2]Department of Urology, The Third Affiliated Hospital, Southern Medical University, Guangzhou, China

Background: Immunotherapy has achieved remarkable success in treating advanced liver cancer. Current evidence shows that most of the available immune checkpoint inhibitor (ICB) treatments are suboptimal, and specific markers are needed for patients regarded as good candidates for immunotherapy. Necroptosis, a type of programmed cell death, plays an important role in hepatocellular carcinoma (HCC) progression and outcome. However, studies on the necroptosis-related lncRNA in HCC are scarce. In this view, the present study investigates the link among necroptosis-related lncRNA, prognosis, immune microenvironment, and immunotherapy response.

Methods: Gene transcriptome and clinical data were retrieved from The *Cancer* Genome Atlas database. Pearson correlation analysis of necroptosis-related genes was performed to identify necroptosis-related lncRNAs. The Wilcoxon method was used to detect differentially expressed genes, and prognostic relevant lncRNAs were obtained by univariate Cox regression analysis. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes analysis were utilized to perform functional enrichment analysis. Lasso–Cox stepwise regression analysis was employed to calculate risk score, which was involved in analyzing immune cells infiltration, immune checkpoints expression, and predicting immunotherapeutic efficacy. Quantitative RT-PCR (qRT-PCR) was performed to detect the expression pattern of lncRNA in cell lines.

Results: The 10 lncRNAs generated in this study were used to create a prognostic risk model for HCC and group patients into groups based on risk. High-risk patients with HCC have a significantly lower OS rate than low-risk patients. Multivariate Cox regression analysis showed that risk score is an independent risk factor for HCC with high accuracy. Patients in the high-risk group exhibited a weaker immune surveillance and higher expression level of immune checkpoint molecules. In terms of drug resistance, patients in the low-risk group were more sensitive to sorafenib. The OS-related nomogram was constructed to verify the accuracy of our model. Finally, quantitative RT-PCR experiments were used to verify the expression patterns of candidate genes.

Conclusion: The lncRNA signature established herein, encompassing 10 necroptosis-related lncRNAs, is valuable for survival prediction and holds promise as prognostic markers for HCC.

KEYWORDS

hepatocellular carcinoma, necroptosis, long non-coding RNA (lncRNA), prognostic model, immunotherapy, prognostic marker genes

Introduction

Hepatocellular carcinoma (HCC) is a health challenge nowadays, which has become the sixth most common malignant tumor and the fourth-leading cause of cancer death worldwide (Parkin et al., 2005). In addition, HCC is a malignant tumor that ranks fourth for incidence and second for mortality in China (Chen et al., 2016; Zhou et al., 2019). Compared with different regions of Europe and the United States, East Asia and Africa have the highest incidence and mortality of liver cancer (Mcglynn et al., 2015). According to the Surveillance Epidemiology and End Result (SEER) report, liver cancer is expected to be the third leading cause of cancer-related death by 2030 (Rahib et al., 2014).

Surgical resection is still the main method for the treatment of early liver cancer. In pursuit of a higher 5-year survival rate and lower perioperative mortality, surgical indications are usually limited to single tumor, a well-preserved liver function, no portal hypertension, and ECOG score of 0 (Zhou et al., 2001). As the standard therapy for the intermediate stage is radiofrequency ablation and transarterial chemoembolization (TACE), HCC has been the most widely used treatment modality for patients with unresectable HCC. However, HCC is a disease with insidious onset, most patients were diagnosed in the advanced stage and lost the opportunity of operation. At this time, molecular targeted therapy, immunotherapy, and other systemic therapies are of particular importance. Immunotherapy for liver cancer is a promising research direction. Over the past few years, immunomodulator, tumor vaccine, and adoptive immunotherapy have been proved to be effective in clinical practice of liver cancer (Hong et al., 2015). The IMbrave150 study confirmed that the combination of atrizumab and bevacizumab markedly improve the overall survival rate of HCC patients, which has been established as the standard first-line regimen for advanced liver cancer (Finn et al., 2020). In the CheckMate 040 trial (El-Khoueiry et al., 2017), nivolumab, a checkpoint inhibitor that blocks programmed cell death protein-1, demonstrated durable responses and prolonged long-term survival, which has been approved by FDA for second-line therapy of HCC. Nevertheless, cancer treatment, particularly immunotherapy, is urgently in need of the identification of useful biomarkers, as only some patients benefit from this treatment.

In recent years, programmed cell death has aroused great interest in the field of cell death. Nowadays, three forms of cell death including apoptosis, necroptosis, and pyroptosis have been intensely studied. Necroptosis, a kind of programmed cell lytic death, was first proposed by Degterev et al. (2005) to describe a regular rather than accidental necrotic cell death, playing a key role in the regulation of tumorigenesis, cancer metastasis, and cancer immunity (Seehawer et al., 2018). The key molecules of necroptosis promoted the metastasis and progression of cancer alone or in combination with other molecules (Mccormick et al., 2016). Necroptosis is usually regarded as a secondary cell death response in which apoptosis was inhibited, mainly mediated by receptor-interacting protein kinase (RIPK)1, RIPK3, and mixed lineage kinase domain-like protein (MLKL). RIPK1 is a key transcription factor that controls cell survival and death (Christofferson et al., 2014; Weinlich and Green, 2014) and mediates RIPK3-dependant necrosis and FADD-dependent apoptosis (Micheau and Tschopp, 2003; Sun et al., 2012). The morphological changes are characterized by rapid swelling of cells and organelles. The pathway of necroptosis is canonically triggered by the ligand-dependent cell surface death receptors, such as Fas, TNF receptor 1 (TNFR1), IFN receptor (IFNR), and Toll-like receptor (TLR). And necroptosis mediated by TNFR1 is the most representative example (Choi et al., 2019). TNFR1 is homotrimerized and recruits the TNFR-associated death domain (TRADD) protein to its cytoplasmic death domain upon binding of TNF (Hsu et al., 1995; Vandevoorde et al., 1997; Chen and Goeddel, 2002; Gaur and Aggarwal, 2003). When FADD is deficient, RIPK1/RIPK3/MLKL forms a complex called complex IIb or a necrosome to induce necroptosis (Cho et al., 2009; He et al., 2009; Zhang et al., 2009; Sun et al., 2012; Zhao et al., 2012). Caspase-8 is the main inhibitor in the process of necroptosis. When FADD and Caspase-8 were recruited into TRADD and RIPK1 to form complex IIa, Caspase-8 was activated (Mak and Yeh, 2002; Wallach et al., 2002). Activated Caspase-8 induces apoptosis by inhibiting the activity of RIPK1 and RIPK3 (Fritsch et al., 2019). The interactions between RIPK1 and RIPK3 lead to auto-phosphorylate and trans-phosphorylate and the assembly of the "necroptosis complex" when Caspase-8 was inhibited or blocked, which further initiates downstream signaling leading to necroptosis (Cho et al., 2009; He et al., 2009; Zhang et al., 2009). RIPK3 plays a key role in the induction of necroptosis. Phosphorylated RIPK3 activates and phosphorylates MLKL after RIPK3 phosphorylation (Sun et al., 2012), phosphorylated

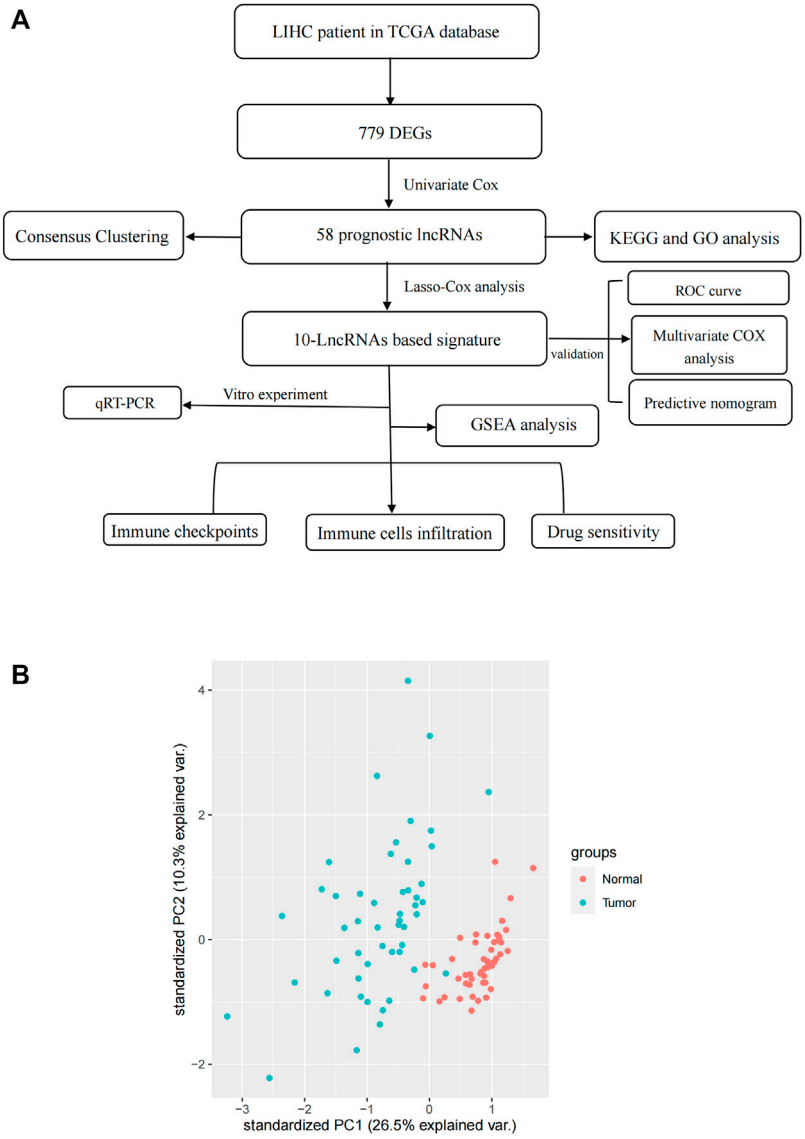


FIGURE 1
(A) Study Flow Chart. **(B)** a PCA plot for RNA-seq data.

MLKL translocated to the cell membrane leading to the formation of MLKL oligomers, causing its rupture and leading to the leakage of cellular contents and release of DAMP, ultimately triggering the body's immune response (Gong et al., 2017). Increasing evidence suggests that necroptosis is closely associated with carcinogenesis and anti-tumor immunity. Deletion of MLKL in tumor cells reduces spontaneous lung metastases in a breast cancer model (Jiao et al., 2018). PK68, a kind of RIPK1 inhibitor, significantly inhibits metastasis of mouse melanoma and lung cancer cells (Hou et al., 2019). There is growing evidence that lncRNAs are closely related to programmed cell death (PCD), lncRNA TUG1 knockdown promotes apoptosis by regulating the miR-132-3p–SOX4 axis in osteogenic sarcoma cells (Li et al., 2018). lncRNA MALAT1 promotes colorectal cancer cell proliferation and inhibits apoptosis by activating autophagy (Si et al., 2019).

Non-coding RNA (ncRNA) is a class of RNA longer than 200 nucleotides that do not encode proteins, including microRNA and long non-coding RNA (lncRNA). In cancer, lncRNAs functioning through a variety of mechanisms, such as chromatin remodeling, chromatin interaction, ceRNAs, and natural antisense transcripts (Fang and Fullwood, 2016). LncRNA participates in tumorigenesis and development by

different mechanisms (Schmitt and Chang, 2016), a study has shown that lncRNA CASC11 interacts with hnRNP-K to activate the WNT/β-catenin pathway, promoting the growth and metastasis of colon cancer (Zhang et al., 2016). LncRNA HOTAIR is able to predict the sensitivity of ovarian cancer patients by two kinds of platinum chemotherapy, guiding clinical decision-making (Teschendorff et al., 2015). In summary, the imbalance of lncRNA affected cell function, promotion of metastasis, and resistance to chemotherapeutic drugs (Brunner et al., 2012).

With the development of genomics, transcriptome, and next-generation sequencing technology in biological research, the prognostic model constructed by specific gene clusters showed unique advantages in predicting the prognosis of cancer patients and the efficacy of immunotherapy. For example, a prognostic signature based on seven autophagy-related genes is eligible to act as an independent prognostic indicator of poor prognosis (Wang et al., 2021); m6A-associated regulators play a key role in HCC prognosis, tumor microenvironment, and drug resistance (Jin et al., 2021). However, no such study in necroptosis-related lncRNA about HCC has been published yet. Therefore, more studies are needed to explore the prognostic value of necroptosis-related lncRNA in HCC. In this study, we use bioinformatic methods to identify prognostic necroptosis-related lncRNA based on TCGA public database, constructing a prognostic model to verify the role of necroptosis-related lncRNA in HCC. We aimed to identify suitable candidates for immunotherapy and evaluate disease prognosis to better guide therapeutic decisions.

Methods

Data acquisition and preprocessing

We downloaded transcriptome data (TCGA-LIHC-FPKM) and corresponding clinicopathological information of 374 HCC samples and 50 normal samples from The *Cancer* Genome Atlas (TCGA, https://portal.gdc.cancer.gov/). The transcriptome data were converted from FPKM (Fragments Per Kilobase of transcript per Million mapped reads) format to TPM (Transcripts Per Kilobase Million) format by a Perl tool, the human gene annotation file was downloaded from Ensemble database (http://ensemblgenomes.org/). A PCA plot of RNA-sequence data was constructed by "tidyverse" and "dplyr" packages, visualization was prepared in "ggbiplot" package. In this calculation, we included 374 HCC samples and 50 normal samples, 1000 RNAs were detected randomly during this process. One can see that the clustering and integrity of RNA-sequence data were excellent (Figure 1B). Clinical data were preprocessed as follows: Stage IIIA, IIIB, and IIIC were merged into stage III, stage IVA and IVB were merged into stage IV. Samples missing overall survival (OS) data were deleted, finally resulting in a total of 376 samples for survival analysis. The inclusion criteria of this study are as follows: 1) primary hepatocellular carcinoma confirmed by pathology; 2) all patients were available of their survival data; 3) there was complete data for all the samples. All data were downloaded from the public databases; hence, it was not required to obtain additional ethical approval for our study. A total of 67 necroptosis-related genes were used to perform analysis. Spearman's analysis was conducted to identify necroptosis-related lncRNAs with the identification criteria of Spearman's correlation coefficient with an absolute value of >0.4 and $p < 0.001$.

Identification of DEGs

The "bioconductor limma" package was applied to identify differentially expressed genes (DEGs) between cancer and normal samples, |log2foldchange|>1 and the false discovery rate (FDR) < 0.05 were the cut-off criteria for DEGs. The volcano plot was further figured to show the DEGs of TCGA. The results were presented as heatmap and volcano map.

Identification of prognostic-related lncRNAs (PRlncRNAs)

Univariate Cox regression analysis was performed on DEGs of the TCGA cohort to obtain prognostic-related lncRNA (PRLncRNA). The relationship between PRLncRNA and necroptosis-related genes was visualized by the "ggalluvial" package.

Gene ontology (GO) and kyoto encyclopedia of genes and genomes (KEGG) analysis

GO and KEGG analyses of PRLncRNAs were performed by R package "ClusterProfiler." Enrichment analysis was conducted by using the functions "enrichGO" and "enrichKEGG." The bubble diagram was generated by R packages "ggplot2," "enrichplot," and "GOplot." The FDR value of $p < 0.05$ was considered significantly enriched terms.

Consensus clustering analysis

Consensus clustering is a resampling-based clustering algorithm, which quantifies the consensus between several clustering iterations and provides means to estimate the number of clusters that best fit data. In this study, data sets were clustered by k-means with k from two to nine by using the "ConsensusClusterPlus" package. Survival

TABLE 1 Clinicopathlogical features of high- and low-risk group.

Variables		Test cohort	Training cohort	*p* Value
Number of patients		**190**	**187**	
Age	≤65	120	115	0.873
	>65	70	71	
	Unknown	1	0	
Gender	Female	64	58	0.661
	Male	127	128	
Grade	G1	27	28	0.227
	G2	95	85	
	G3	56	68	
	G4	10	3	
	Unknown	3	2	
T stage	T1	102	83	0.480
	T2	43	52	
	T3	38	43	
	T4	7	6	
	Unknown	1	2	
M stage	M0	139	133	0.931
	M1	2	2	
	unknown	50	51	
N stage	N0	133	124	0.530
	N1	1	3	
	Unknown	57	59	
TNM stage	I	98	77	0.412
	II	41	46	
	III	39	47	
	IV	2	3	
	unknown	11	13	

analysis between clusters was conducted using the log-rank test and Kaplan– Meier curve, and these analyses were performed with the "ggsurvplot" function in the survival R package. Correlations between clusters and clinical features were performed through the chi-square test. $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

Construction and validation of the prognostic risk signature

Least absolute shrinkage and selection operator (LASSO)-Cox regression analysis of overall survival (OS) with a 10-fold cross-validation was performed to screen for necroptosis-related lncRNAs with prognostic values, which may preserve valuable variables and avoid overfitting and delete highly correlated genes. The optimal coefficient was determined and the deviance likelihood was calculated by the "glmnet" package (Supplementary Figure S3A). The optimal value of the shrinkage parameter (lambda) was the minimum as selected by 10-fold cross-validation (Supplementary Figure S3B). The risk score of each sample was calculated based on the LASSO-Cox regression co-efficiency through the following formula: Risk score = β1*Exp1 + β2*Exp2 + βi*Expi, where β represents the regression coefficient and Exp represents the gene expression value (Supplementary Table S2).

A total of 370 patients with HCC were randomly divided into training group (n = 186) and verification group (n = 184) in the ratio of 1:1. There was no significant difference in clinical covariables, including age, molecular subtype, grade, TNM stage, and survival status between the two groups (Table 1). According to the median risk score, each group was divided into high-risk subgroup and low-risk subgroup. Univariate Cox regression was used to screen the prognostic variables significantly related to OS in the training group, verification group, and the whole cohort. Afterwards, statistically significant variables were included in multivariate COX regression analysis. Results were reported as hazard ratio (HR) with 95% confidence interval (CI). The sensitivity and specificity of the model for predicting the 1-, 3-, and 5-year survival rates of HCC

TABLE 2 Details of 58 PRLncRNAs.

Gene	Coef
AL031985.3	0.351023768463183
SREBF2-AS1	0.0386144854777114
ZFPM2-AS1	0.0102274968460904
KDM4A-AS1	0.114565023598859
AC026412.3	0.167864746612448
AC145207.5	0.0920876812602594
DUXAP8	0.46617052054734
LINC01224	0.181433562593978
AC099850.4	0.00947324010819726
MKLN1-AS	0.262823244119484

patients were evaluated by an ROC curve constructed by the "timeROC" package. In addition, to further compare the predictive efficacy of risk scores compared with other clinical features, we constructed an ROC curve for predicting 3-year OS in HCC patients and calculated the AUC.

Construction and validation of a predictive nomogram

Based on the statistically significant factors of multivariate Cox regression analysis, a prognostic nomogram was constructed to predict the survival probability of patients diagnosed with HCC. Samples with missing entries were removed from this analysis, a total of 240 patients were included in this analysis, the overall survival time of 1, 3, and 5 years were used as the end point. R package "RMS" was used to perform this process, the predictive accuracy of this model was assessed by employing concordance index (C-index). The 95% CI = (C-index ± 1.96) *SE, the predictive accuracy of the proposed model was reflected by the overlap between the calibration curve and the diagonal.

Tumor immune cells infiltration (TICL) and correlation analysis between high- and low-risk subgroups

In order to explore the lncRNAs signature whether plays a key role in immune infiltration of HCC, we calculated the abundance of 22 types of infiltrating immune cells in each sample by CIBERSORT algorithm. The differential infiltration levels of TICLs between high- and low-risk groups were calculated by the Wilcox test, violin plots were generated using the "vioplot" package. Single-sample gene-set enrichment analysis (ssGSEA) was conducted by the "GSVA" package. Correlations was analyzed by Spearman's non-parametric test. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

Estimation of tumor immune microenvironment (TIME)

The immune score and stromal score of each sample were estimated using the ESTIMATE algorithm by "estimate" package. Results were presented in violin and boxplot plots, methods the same as before.

Differential expression of immune checkpoint molecules between high- and low-risk subgroups and correlation analysis

The differential expression of programmed death receptor 1 (PD-1, also known as PDCD1), programmed death ligand 1 (PD-L1, also known as CD274), cytotoxic T lymphocyte antigen 4 (CTLA-4), and indoleamine-1 (IDO-1) in high- and low-risk subgroup were analyzed by the Wilcox test. Spearman's method was carried out for correlation analysis between immune checkpoint molecules and risk score. The "ggExtra" package, "ggplot2″ package, and "ggpubr" package of R software participated in the production of box chart and scatter chart.

Gene set enrichment analysis (GSEA)

Gene set "c2.cp. kegg. v7.5.1. symbols. Gmt" (downloaded from http://www.gsea-msigdb.org/gsea/downloads.jsp) acted as an input preparation file in GSEA analysis performed by "clusterprofiler" package (p value cutoff = 0.01 and q value cutoff = 0.05). |NES| > 1, adjust p-value<0.05 and FDR<0. 25 were considered to be a significant enrichment item. Enrichment plots were displayed with the "enrichplot" package.

Evaluation of sorafenib sensitivity in high- and low-risk subgroups

The R package "pRRophetic" was used for predicting the semi-inhibitory concentration (IC50) of patients in high- and low-risk groups treated with sorafenib. The lower the number, the stronger the binding affinity and the higher the drug sensitivity.

Cell culture

Human hepatic cell line L-O2 and four human HCC cell lines (HepG2 cell, Hep3B, LM3, and Huh7 cells) were flashed frozen in liquid nitrogen with 2 ml tubes and stored at −80°C. All cell lines

TABLE 3 q-PCR primer sequence.

Gene	HR	HR.95 L	HR.95H	p
AC009005.1	1.17110969344691	1.07429391453767	1.27665054742077	0.000333551251369458
AC018690.1	2.10384322910643	1.43080767874763	3.09346699657856	0.000156073183457983
LUCAT1	1.18189988087072	1.09837739997351	1.27177355291169	7.84559781562504e-06
AL117336.2	1.56655126336053	1.30183152048208	1.88510020085217	2.00524463719125e-06
AC010864.1	5.49638443883633	2.10858881432178	14.3272323623699	0.000490074433174014
AL031985.3	1.83467899783164	1.49647955721917	2.24931039575273	5.29468369207119e-09
SREBF2-AS1	1.45642902550033	1.16706193665077	1.81754321660702	0.000877895411678177
AC034229.4	1.97472355301551	1.40030533397342	2.78477344635196	0.000104567297347242
AC074117.1	1.46708753382555	1.19327667886859	1.80372739199687	0.000276371482958239
THUMPD3-AS1	1.36425983075505	1.16119422832644	1.60283683849706	0.00015843033051424
AC026401.3	1.08418641157238	1.03570125862827	1.13494134070575	0.00053473105754009
AC125437.1	2.5548070290898	1.58555156851862	4.11657311277797	0.000116331481881563
AL049840.4	2.0302971636625	1.36352982623686	3.02311434151197	0.000489306477985457
MED8-AS1	2.38264064141742	1.47164608808717	3.85756906642744	0.000412937020718624
AC131009.1	1.81743270484487	1.39795412601752	2.36278256572658	8.11473526933209e-06
BBOX1-AS1	1.26356802863024	1.10450153968093	1.44554272277225	0.000654722408806783
ZFPM2-AS1	1.09331071475861	1.05267666559006	1.13551326639881	3.90143830862613e-06
ZEB1-AS1	1.45367588907736	1.16714497469377	1.81054936302091	0.000838235675022376
DDX11-AS1	4.30711049547743	2.04855911473802	9.05573126339791	0.000117458622091245
MCM3AP-AS1	8.94040972369425	2.69227895856596	29.688946523618	0.000347151833494758
ZNF337-AS1	2.73367363684378	1.52991790481875	4.88455722312761	0.00068428899137241
AC012073.1	1.40901745291795	1.22743803553936	1.61745858050992	1.10884707035332e-06
LINC01094	2.00660050130517	1.39264200898897	2.89122800105769	0.000186022552058713
LINC01138	1.44794724338591	1.22020644496107	1.71819385833168	2.24120525795787e-05
AC100872.2	1.52447962710877	1.24327474005833	1.86928766312799	5.05641699168209e-05
AL355574.1	1.28242891205497	1.11158436232424	1.47953135202056	0.00064918759120028
KDM4A-AS1	2.86701953213825	1.90687848154453	4.3106055667503	4.14527699473686e-07
AC016394.2	1.57882015234359	1.23080436529767	2.02523905807191	0.000324985331627426
AC026412.3	15.090021090766	4.61045168516158	49.3896806797968	7.24800949113819e-06
AL365203.2	1.13673308566103	1.07468028025401	1.20236886428308	7.6533988362139e-06
AC145207.5	1.76721729739874	1.3628688724496	2.29153151807772	1.74349659249233e-05
AC107068.1	2.10949184257841	1.44055306556969	3.08906068111073	0.000125211421420502
AC012467.2	1.49459418738293	1.18257552681207	1.88893794460693	0.00076932764792714
AC025176.1	2.28280817121522	1.48702718115843	3.50445049868376	0.000160460084632489
DUXAP8	4.37169584194494	2.02781234526238	9.42479938004609	0.000167423690730116
TMCC1-AS1	2.56854719180891	1.77460682269685	3.71768810542686	5.72455165204052e-07
NCK1-DT	1.51326486674367	1.20122630200084	1.90636065253186	0.00043798957870755
LINC01224	1.92867470390563	1.44091805772471	2.58153896645535	1.00747168565781e-05
NRAV	1.23212492752684	1.12566307516838	1.3486556239805	5.97410050681252e-06
AL442125.2	3.83473706827137	2.04401302301756	7.19428311717178	2.82689097369676e-05
AP003392.4	7.36176131533239	2.58151629641687	20.9936810157455	0.000188653499478888
MIR210HG	1.14187648047771	1.06700456384948	1.22200217397766	0.000125916300801374
GHRLOS	3.47983906642121	1.79075574511339	6.76210586576923	0.000234274247349297
AC026356.1	4.01051373646868	2.129621086889	7.55262075935777	1.70265345012451e-05
AC099850.4	1.13959993843801	1.08696785648712	1.19478051897969	6.0760619159702e-08
AC026355.2	2.32443180244038	1.66331751099788	3.24831739488801	7.81777059889033e-07
AL928654.1	1.91049713067697	1.31021936911352	2.78579249579746	0.000768120091986322
POLH-AS1	2.56671257723233	1.7123957107321	3.84724944873056	5.00030914523686e-06

(Continued on following page)

TABLE 3 (*Continued*) q-PCR primer sequence.

Gene	HR	HR.95 L	HR.95H	p
C2orf27A	1.67756319510261	1.35037659748986	2.08402476671623	2.9600669059356e-06
SNHG4	1.47583880290837	1.27644304580875	1.706382575644	1.47410333621925e-07
BACE1-AS	1.21074284838357	1.09643625670793	1.33696622666723	0.000157124984312145
SNHG3	1.06757970081289	1.03740139549618	1.09863590172115	7.83216781994623e-06
AC073611.1	1.48893059282573	1.17720442829988	1.88320248969337	0.000896616489881415
SNHG10	1.29463387153636	1.12504691265138	1.48978397476709	0.000312461913317886
AC107959.3	1.38256463352294	1.15262199806244	1.65837973687959	0.00048225465968827
MKLN1-AS	3.36374917116319	2.23555078410734	5.06130684524757	5.91261812769369e-09
AC006252.1	4.77420311932314	1.95702613952842	11.6467608501372	0.000591250187772917
HMGN3-AS1	2.12745977534335	1.39332694836852	3.24840131815699	0.000472264012528417

TABLE 4 LncRNAs constituting a prognostic model.

Gene	Forward primer (5'→3')	Reverse primer (5'→3')
ACTB	CCTCGCCTTTGCCGATCC	CGCGGCGATATCATCATCC
SREBF2-AS1	CAACGGGGATGTAGCCATCA	ATCTCTCCGGGATGGAGACC
ZFPM2-AS1	TGGCAGAGTTGCACAGAAGA	ACCACTCACACTTTCATCGCT
KDM4A-AS1	GAGGGTGAAAGGAACGTCCA	AAGTACTTTGCCAGGTCCCA
AC026412.3	GCGTAGATCCCTTTGGCTCA	GACCTCCATTGAAGGGCTCA
AC145207.5	GACTGGCCAAGCATTTGGTG	TCTGGCCTACCTTAGGCTACAT
DUXAP8	CACCAGCCTCACTAGCACTC	ACACCCGGCCAAGTTCTTTA
LINC01224	ACAGGATTGTTAATCTCATCTTGGA	TCAGGTTCTACACAGAGGCA
AL031985.3	ACAACACATCAAGAGGCCCA	TTCCCTGCCTGAGTATGGCT
MKLN1-AS	TCTCTGAAAGCAGCGCTTGG	AGTCCTCAAGGTATGGGGGA

were cultured in 10% fetal bovine serum (FBS; Invitrogen, Carlsbad, California, United States) in DMEM medium. All cell lines grew in a humid environment of 37°C, 5% CO2, 99% relative humidity and did not contain antibiotics. The cells were subcultured at a ratio of 1:2 or 1:3 when they reached 80% confluence.

RNA extraction and quantitative real-time PCR (qRT–PCR)

TRIzol reagent kit (Invitrogen) was used to extract the total RNA from logarithmic growth cells. cDNA was synthesized using PrimeScript RT reagent kit (Takara Biotechnology, Dalian, China). Quantitative real-time PCR analysis was conducted using TB Green Premix Ex Taq II kit (Takara Biotechnology, Dalian, China) according to the instructions, three replicates were set in each well. All operations were carried out on ice. The $2^{-\Delta\Delta Ct}$ method was used for quantitative PCR analysis. The primers were designed using the NCBI primer-BLAST tool (https://www.ncbi.nlm.nih.gov/). The sequences of primers were showed in supplement Table S3, ACTB was used as reference genes (Table 2).

Statistical analysis

All bioinformatics analyses were performed by R 4.0.2, $p < 0.05$ was considered significant in all analysis, unless indicated otherwise. KM curve and log-rank test were used for survival analysis. All continuous variables were tested by Wilcoxon test or independent sample *t*-test, all qualitative data were compared by chi-square test, all tests were conducted as two-sided tests. Cox proportional hazard regression analysis was used to determine the independent predictors of HCC patients. SPSS (IBMSPSS25.0, SPSS Inc.) and Graphpad prism 8 (GraphPad Software, SanDiego, CA, United States) were used for statistical analysis. *In vitro* experiment, each group repeated at least three independent experiments (Table 3).

Results

Data acquisition and establishment of co-expression network

The flow chart of this study is shown in Figure 1. RNA-Seq data of LIHC samples were downloaded from TGCA database.

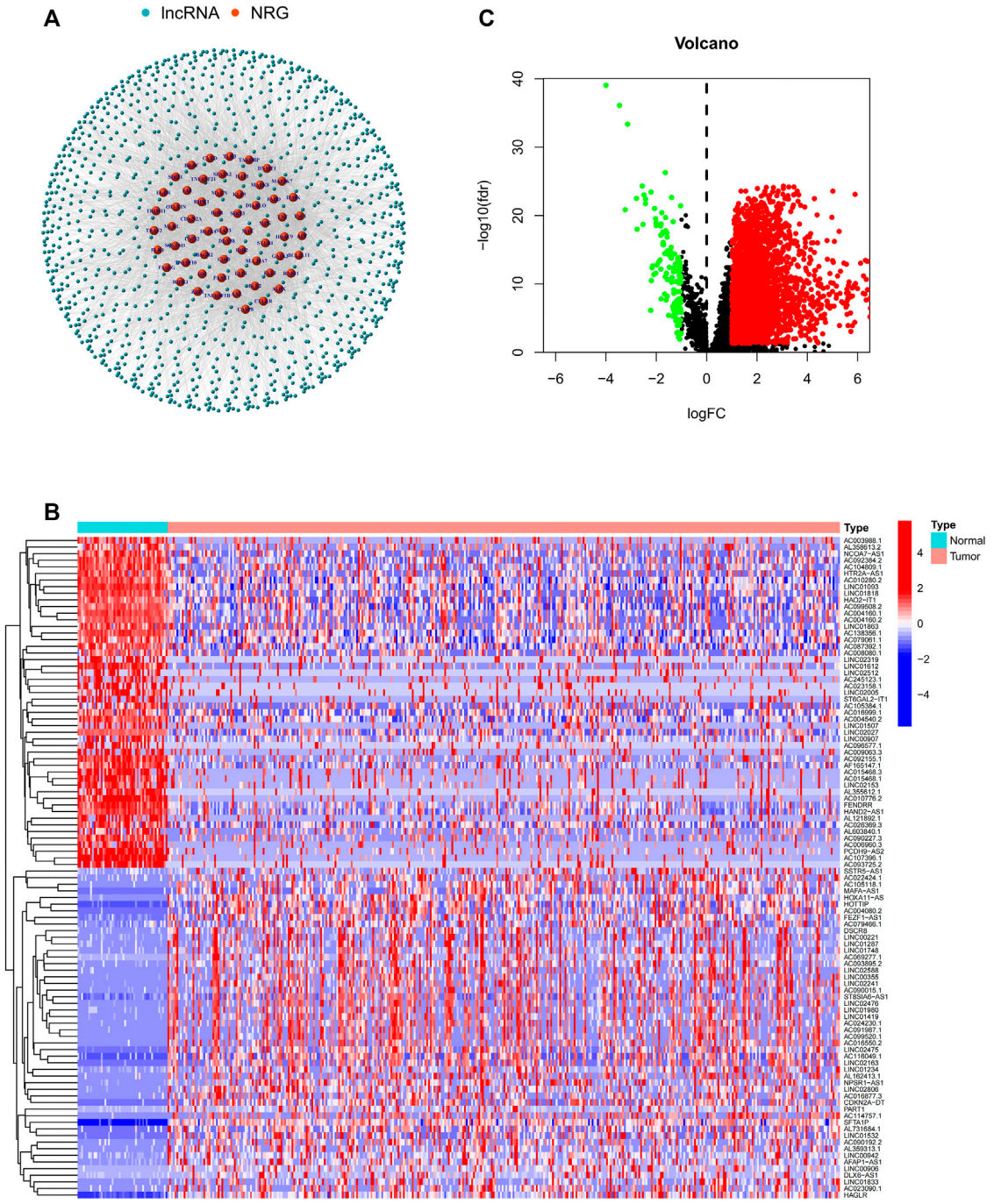


FIGURE 2
(A) Co-expression network between necroptosis-related genes and lncRNAs. **(B)** Differential expression of DEGs in tumor and adjacent normal pairs. **(C)** A volcano map shows DEGs. DEG: differentially expressed genes.

We adopted 67 necroptosis-related genes (FADD, FAS, FASLG, MLKL, RIPK1, RIPK3, TLR3, TNF, TSC1, TRIM11, CASP8, ZBP1, MAPK8, IPMK, ITPK1, SIRT3, MYC, TNFRSF1A, TNFSF10, TNFRSF1B, TRAF2, PANX1, OTULIN, CYLD, USP22, MAP3K7, SQSTM1, STAT3, DIABLO, DNMT1, CFLAR, BRAF, AXL, ID1, CDKN2A, HSPA4, BCL2, STUB1, FLT3, HAT1, SIRT2, SIRT1, PLK1, MPG, BACH2, GATA3, MYCN, ALK, ATRX, TERT, SLC39A7, SPATA2, RNF31, IDH1, IDH2, KLF9, HDAC9, HSP90AA1, LEF1, BNIP3, CD40, BCL2L11, EGFR, DDX58, TARDBP, APP, and

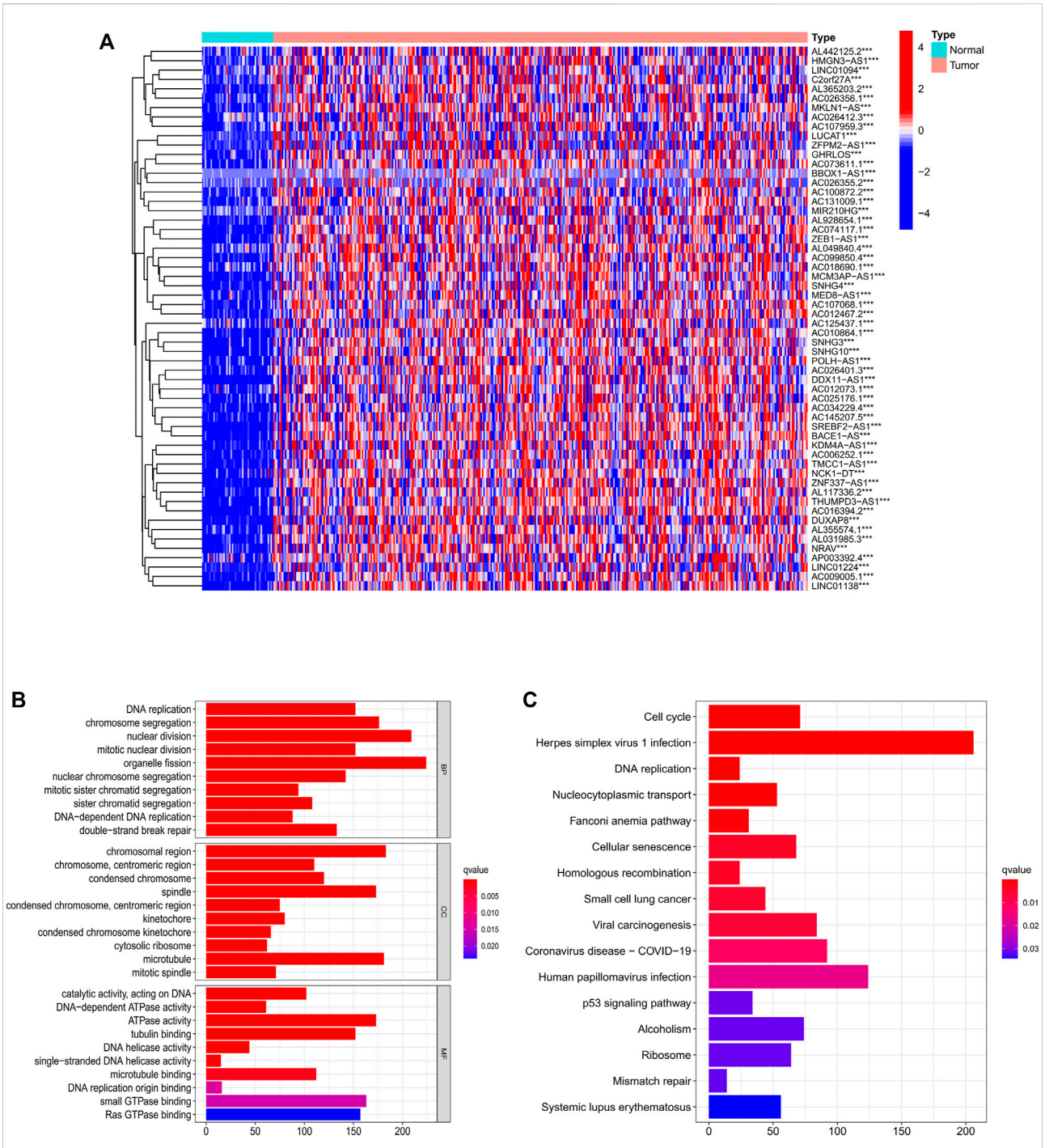


FIGURE 3
Identification of PRlncRNAs. **(A)** Differential expression of PRlncRNAs in tumor and adjacent normal pairs. **(B–C)** GO and KEGG enrichment analysis results. PRlncRNAs: prognostic-related lncRNA; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes.

TNFRSF21) and 14,086 lncRNA (Downloaded from Ensemble database) in this study. At last, we obtained 1026 necroptosis-related lncRNAs. We established a co-expression network to visualize the relationship between necroptosis-related lncRNAs and 67 necroptosis-related genes (Figure 2A).

Identification of DEGs

In this study, a total of 779 DEGs were obtained through the calculation by the "limma" package. Because of the excessive number of genes with significant differences, only the top

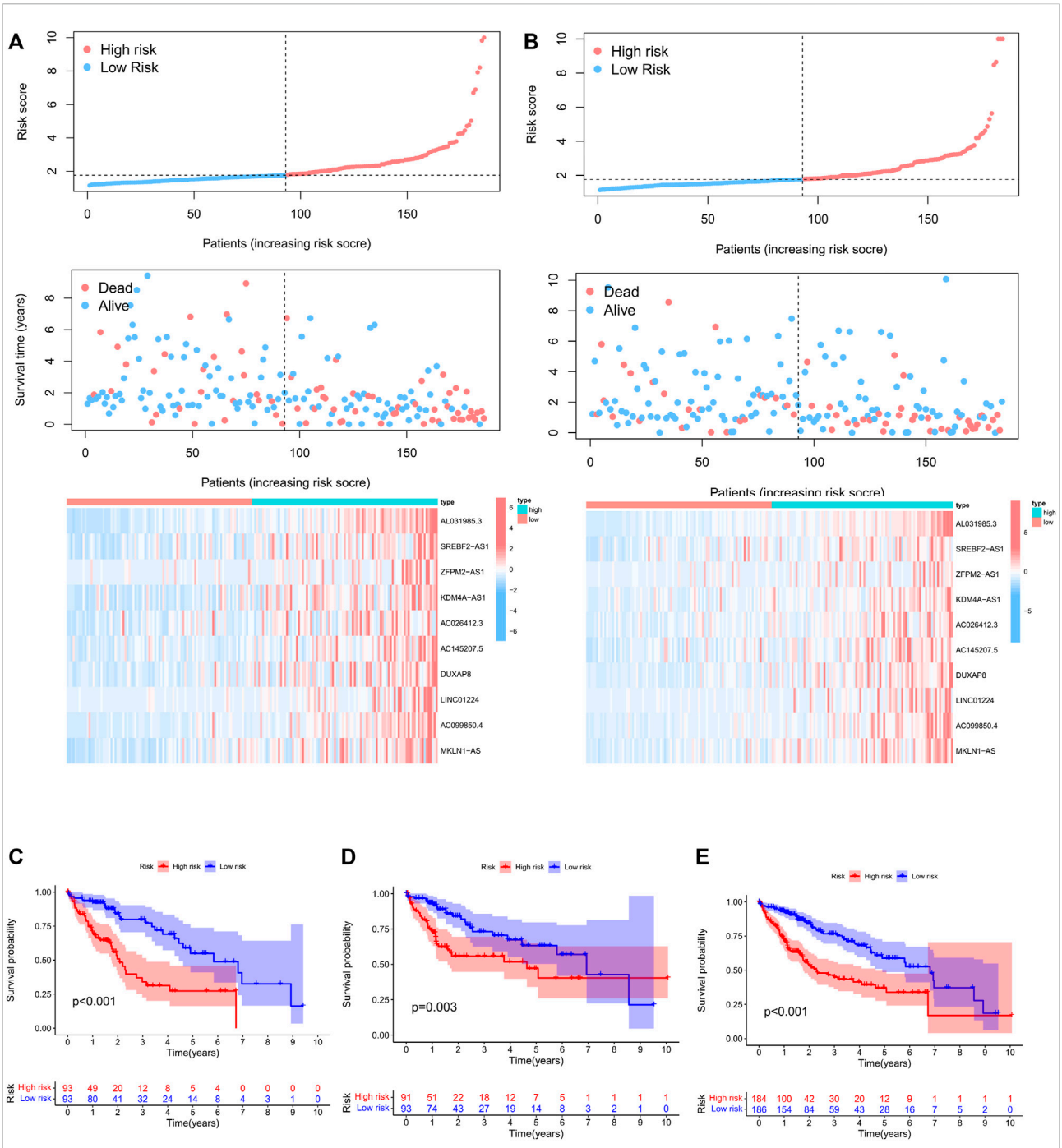


FIGURE 4
Construction and validation of prognostic PRlncRNAs signatures in TCGA cohorts. **(A)** Risk factor curve, survival status of patients, and heatmap of the differential expression of 10 PRlncRNAs in the TCGA training cohort. **(B)** Risk factor curve, survival status of patients, and heatmap of the differential expression of six PRlncRNAs in the TCGA test cohort. **(C–E)** Kaplan–Meier survival curves of the TCGA training cohort, internal validation cohort, and whole cohort.

50 lncRNAs with significant differences are displayed in Figure 2A. As shown in Figure 2B, most of the DEGs are highly expressed in HCC and differentially expressed lncRNAs were shown through volcano plot filtering (Figure 2C).

GO and KEGG analysis of PRlncRNAs

Among 779 DEGs, 58 necroptosis-related lncRNAs (PRlncRNAs) were obtained by univariate Cox regression

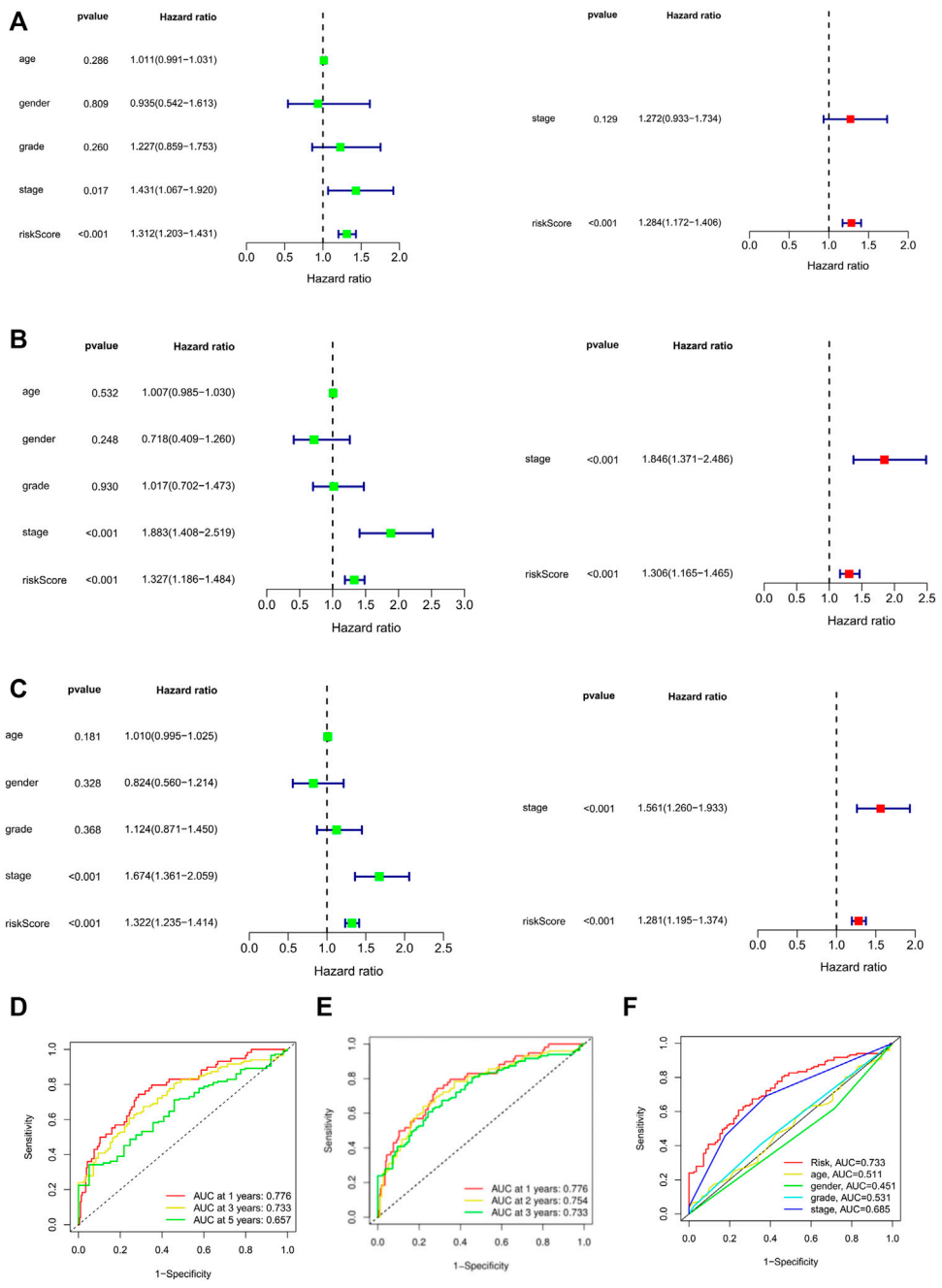


FIGURE 5
Univariate and multivariate Cox regression model for prognostic risk. **(A–C)** The univariate and multivariate Cox analysis of patients OS-related risks in TCGA training cohort, internal validation cohort, and whole cohort. **(D)** ROC curves of 1, 3, and 5 years based on the 10 PRlncRNAs signatures in the whole cohort. **(E)** ROC curves of 1, 2, and 3 years based on the 10 PRlncRNAs signatures in the TCGA whole cohort. **(F)** ROC curves using multiscale features. OS: overall survival.

analysis (Supplementary Table S1). As shown, PRlncRNAs were highly expressed in HCC samples (Figure 3A). In order to further explore the potential biological pathways by which these genes affect the prognosis of HCC patients, we carried out GO and KEGG analysis. Biological process (BP), molecular function (MF), and cellular component (CC) were the three categories of GO analysis. The biological process analysis revealed that "organelle fission (GO:0048285)," "nuclear division (GO:0007088)," "chromosome segregation (GO:0051983)," and "DNA replication (GO:0033260)" were the most abundant

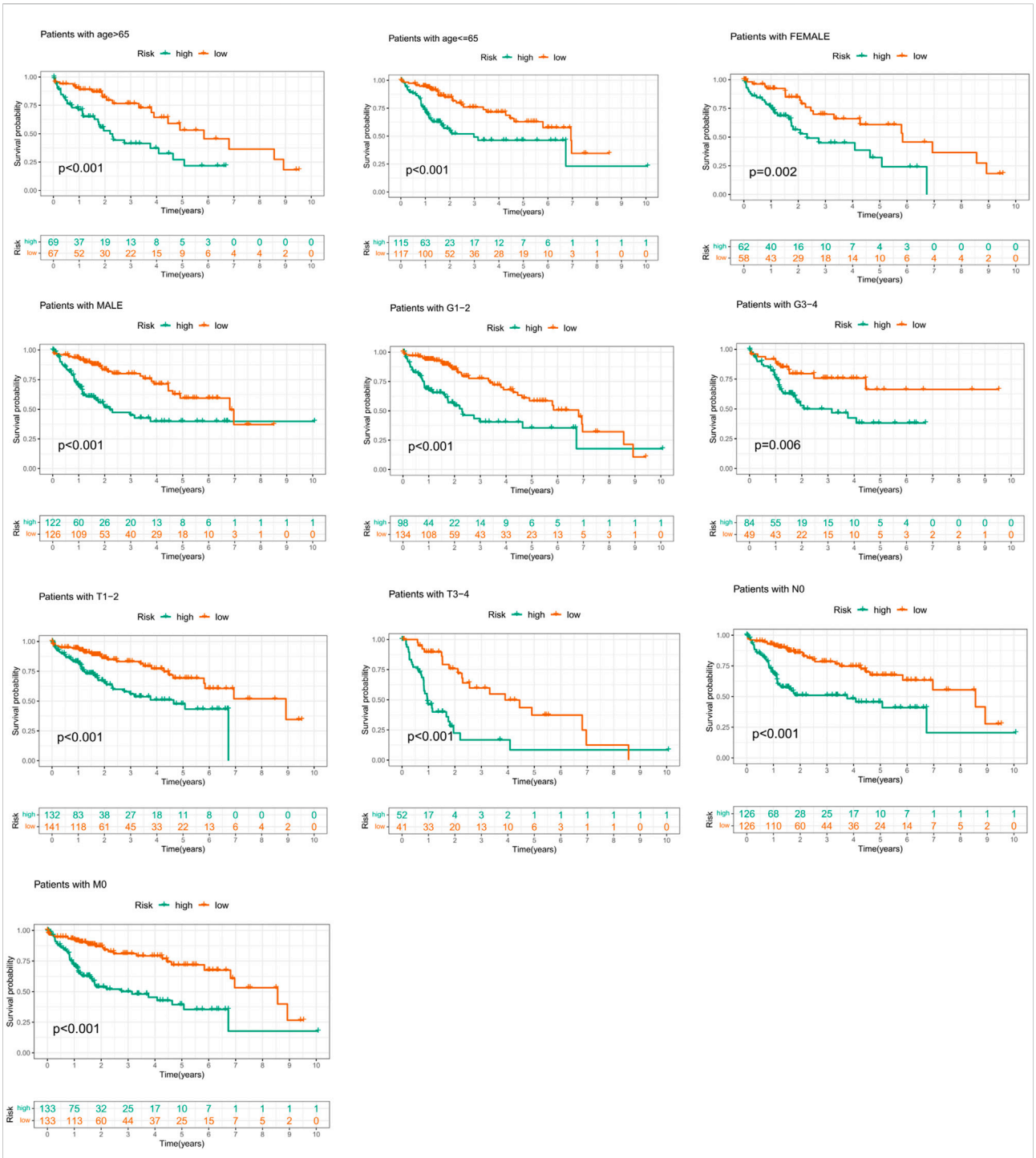


FIGURE 6
Comparison of survival of patients by risk signature score within each subgroup.

terms in the BP category. At the level of MF aspect, ATPase (GO: 1904949), small GTPase binding (GO:0031267), and Ras GTPase binding (GO:0017016) were significantly enriched. The cellular component (CC) indicated that these genes were predominantly located in the chromosomal region (GO:0098687), microtubule (GO:0005874), and spindle (GO:0005819) (Figure 3B). KEGG

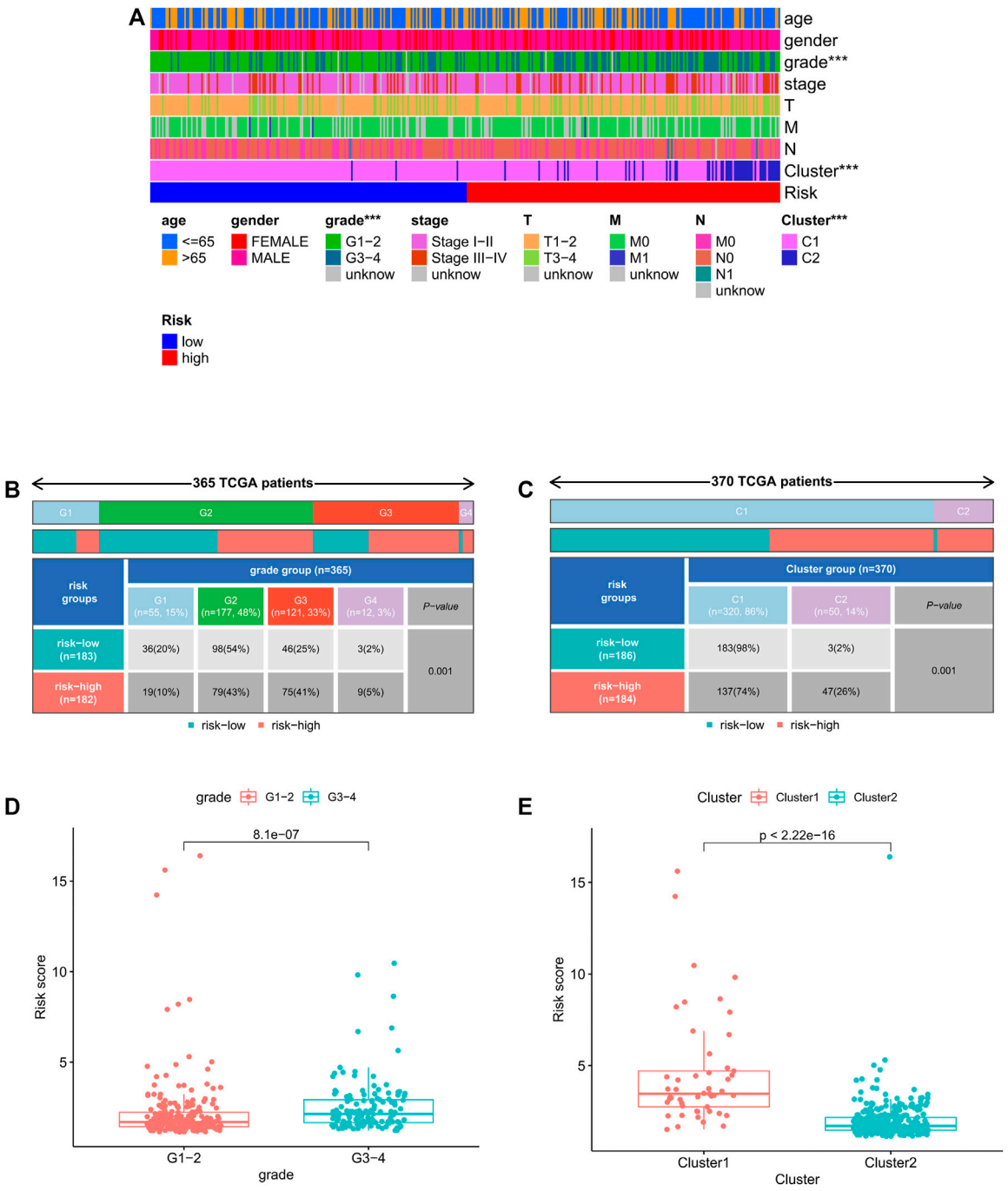


FIGURE 7
Relationship between risk score and clinicopathologic features. **(A)** Heatmap of relationship between risk scores and clinicopathologic features. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$. **(B–C)** Distribution of risk scores stratified by grade and cluster. **(D)** Comparison of risk score distributions between G1–2 and G3–4. **(E)** Comparison of risk score distributions between cluster1 and cluster 2.

analysis showed that PRlncRNAs mainly localized to the Herpes simplex virus 1 infection pathway (hsa05168), human papillomavirus (HPV) infection pathway (hsa05165), coronavirus disease—COVID-19 pathway (hsa05171), viral carcinogenesis pathway (hsa05203), and cell cycle (hsa04110) pathway (Figure 3C).

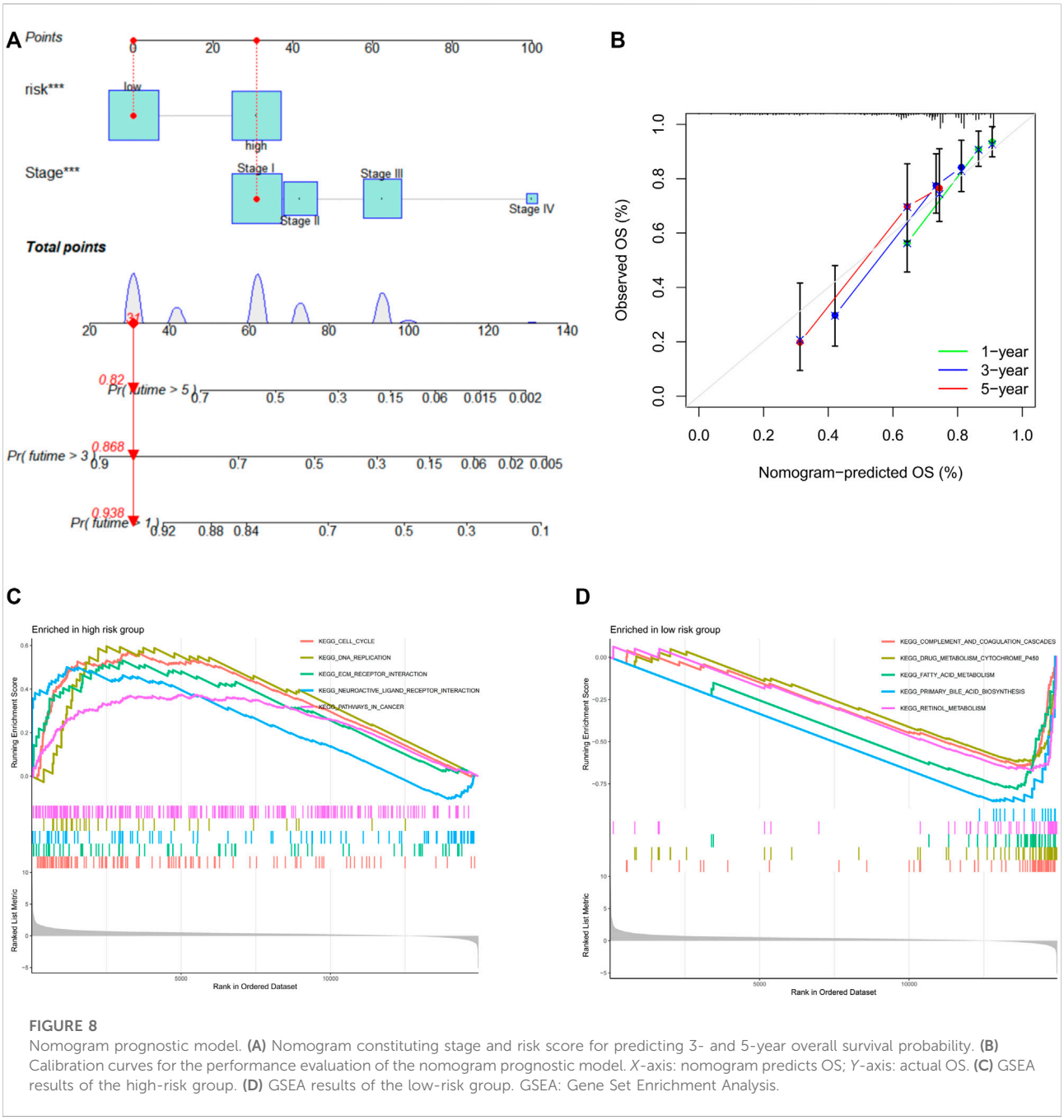


FIGURE 8
Nomogram prognostic model. **(A)** Nomogram constituting stage and risk score for predicting 3- and 5-year overall survival probability. **(B)** Calibration curves for the performance evaluation of the nomogram prognostic model. *X*-axis: nomogram predicts OS; *Y*-axis: actual OS. **(C)** GSEA results of the high-risk group. **(D)** GSEA results of the low-risk group. GSEA: Gene Set Enrichment Analysis.

Development and validation of prognostic PRlncRNA signature

To further evaluate the prognostic value of necroptosis-related lncRNAs, based on 58 candidate PRlncRNAs, we applied LASSO-COX stepwise regression analysis to screen 10 PRlncRNAs which were most correlated to the HCC patient's OS. Based on this, we constructed a risk signature based on 10 lncRNAs. All samples were divided into training group (n = 186) and internal verification group (n = 184) at about a ratio of 1:1 randomly, there were no significant differences in age, gender, stage, grade, T stage, N stage, and M stage between the experimental group and the verification group (Table 1). According to the median of risk score, each group was then divided into high-risk subgroup and low-risk subgroup. Through LASSO-Cox stepwise regression analysis, AL031985.3, SREBF2-AS1, ZFPM2-AS1, KDM4A-AS1, AC026412.3, AC145207.5, DUXAP8, LINC01224, AC099850.4, MKLN1-AS were considered to be the best predictors (Table 4). The coefficient was calculated by LASSO-Cox regression. The risk score formula was as follows: AL031985.3*0.351023768463183 +

SREBF2-AS1*0.0386144854777114 + ZFPM2-AS1*0.0102274968460904 + KDM4A-AS1*0.114565023598859 + AC026412.3*0.167864746612448 + AC145207.5*0.0920876812602594 + DUXAP8*0.46617052054734 + LINC01224*0.181433562593978 + AC099850.4*0.00947324010819726 + MKLN1-AS*0.262823244119484. Compared with the low-risk subgroup, the higher the mortality of the HCC patients in the high-risk group, the shorter the OS. Also, the expression of lncRNAs was positively correlated with the risk score (Figure 4A,B).

Kaplan–Meier survival analysis showed that the total survival time of HCC patients in the high-risk group was significantly shorter than that in the low-risk cohort, both in the training cohort, validation cohort, and whole cohort (Figure 4C–E), suggesting that the prognostic risk score is a valuable predictor of prognosis.

In order to further screen factors that can be seen as independent prognostic factors to predict OS in patients with HCC, risk scores combined with common clinical prognostic factors including age, gender, histological grade, and pathological stage were included in the univariate Cox regression analysis. Univariate (HR = 1.31; 95%CI = 1.20–1.43; $p < 0.001$) and multivariate COX (HR = 1.28; 95% CI = 1.17–1.40; $p < 0.001$) regression analyses were conducted in the training cohort, suggesting that prognostic risk score can be seen as an independent factor affecting survival in HCC patients (Figure 5A). Furthermore, risk score and pathological stage were both important independent prognostic factors for OS in multivariate regression analyses of internal validation cohort (Figure 5B) and whole cohort (Figure 5C). In any case, it is certain that the prognostic risk score based on necroptosis-related lncRNAs was capable of predicting the OS of HCC patients independent of other clinical factors.

The ROC curve showed that the area under the receiver operating curve (AUC) of predicting the OS of 1-, 3-, and 5-year HCC patients was 0.776, 0.773, and 0.657, respectively (Figure 5D). Because of the low 5-year survival rate of HCC patients, we chose to predict the 1- (AUC = 0.776), 2- (AUC = 0.754), and 3-year (AUC = 0.733) OS for further analysis (Figure 5E). Also, compared with other clinicopathological features (AUC of 3-year < 0.5), the prognostic risk score (AUC of 3-year = 0.733) showed a superior prediction ability (Figure 5F). ROC analysis showed that the accuracy of prognostic risk score was higher than that of other clinicopathological features, indicating that our risk signature showed excellent predictive performance.

Subgroup analysis

In order to prove that the prognostic risk score was not subject to differences between populations, we conducted subgroup survival analysis of people with different clinical characteristics, including population of <65 years old, ≥65 years old, male, female, Grade1–2, Grade3–4, T1–2 stage, T3–4 stage, N0, and M0 stage. Considering that the sample size of "M1" and "N1" subgroups was too small to be representative, the two subgroups have not been further validated. Because of the majority of TNM staging (AJCC) focused on the T status, the population of "stage I–II" and "stage III–IV" was not integrated into the study. In our study, only four patients had M0 stage disease and four patients with N0 stage, 233 patients were diagnosed as non-metastasis (N0M0), and the two subgroups largely overlapped. To avoid repetitive results, only those patients with M0 disease were included in the survival analysis in our analysis. It was worth noting that the OS of the low-risk group was better than that of the high-risk group (Figure 6). These results suggested that the suitability of prognostic risk model was not limited by population differences.

Correlation between clinicopathological features and risk score

We further analyzed the correlation between clinicopathological features and prognostic risk score in the whole cohort. The results showed that risk score is significantly related to histological grade and clustering classification (Figure 7A). Furthermore, patients with grade 3/4 or cluster 1 tend to have higher risk score (Figure 7B–E).

The nomogram system improved the prognostic risk score model

To establish a quantitative method for predicting the survival probability of HCC patients, we integrated risk score and pathological stage to develop a nomogram to predict the survival probability of the 1-, 3-, and 5-year OS in the TCGA cohort (Figure 8A). Based on the nomogram, the survival for patients was predicted by calculating the set points based on each nomogram score. The performance of the nomogram was evaluated by Harrell's concordance index (C-index). The C-index value for the nomogram calculated by "coxph" function of the R package "package" was 0.721 (95% CI: 0.656–0.786). Calibration plots suggested that our nomogram performed well (Figure 8B).

Significant enrichment pathways of high- and low-risk groups

To explore the potential molecular mechanisms of prognostic risk signature, Gene Set Enrichment Analysis (GSEA) was performed in high-risk and low-risk groups, respectively. The top five significantly enriched pathways were displayed, GSEA indicated that the high-risk group was highly

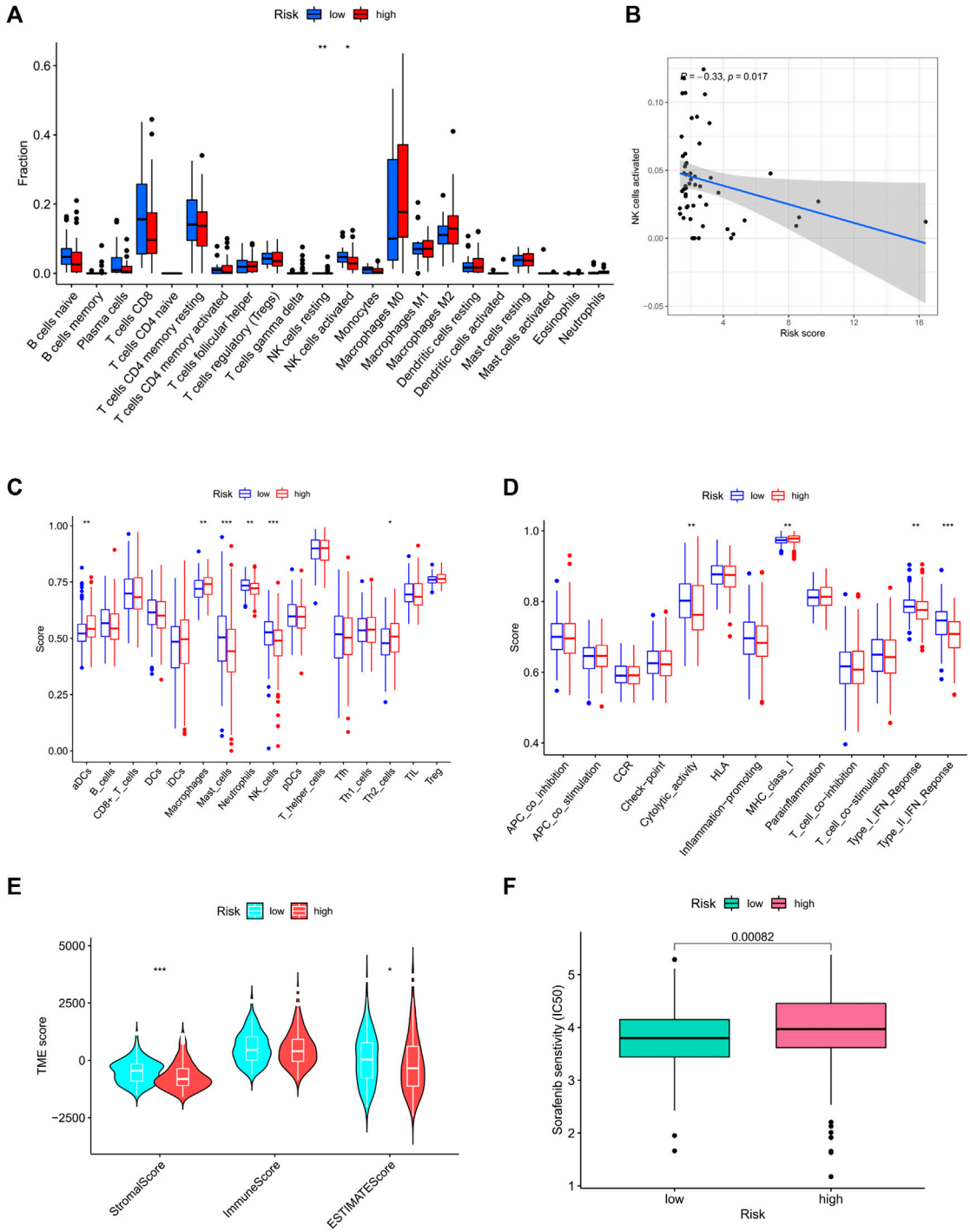


FIGURE 9
Immune cell infiltration in high- and low-risk groups. **(A)** The infiltrating levels of 22 immune cell types. **(B)** Correlation of NK cell level with risk score. **(C)** ssGSEA result of immune cell populations. **(D)** ssGSEA result of immune function. **(E)** Analysis of tumor microenvironment in high- and low-risk groups; Drug sensitivity result. **(F)** IC50 value of sorafenib in high- and low-risk groups. ssGSEA: Single-sample gene set enrichment analysis; IC50: Semi-inhibitory concentration.

enriched in cell cycle signaling pathway, DNA replication signaling pathway, tumor signaling pathway, ECM receptor interaction signaling pathway, and neuroactive ligand receptor signaling pathway (Figure 8C), whereas the low-risk group was

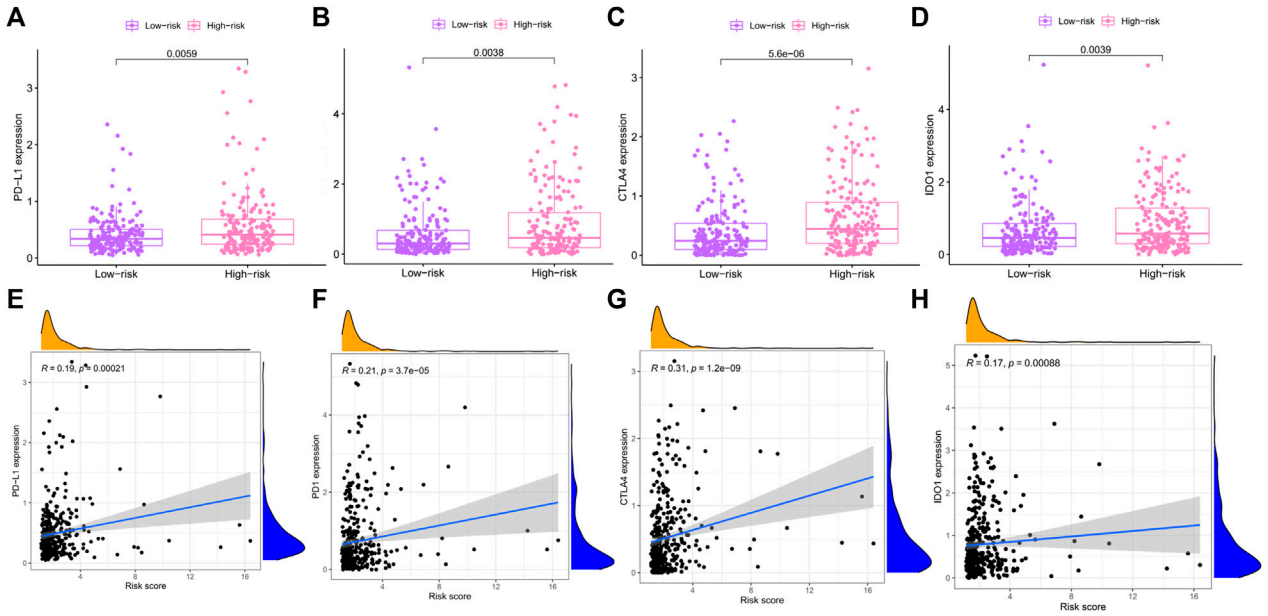


FIGURE 10
Association between risk score and immune checkpoint. **(A–D)** The differential expression of PD-L1, PD-1, CTLA4, and IOD-1 in high- and low-risk groups. **(E–H)** Correlation between PD-L1, PD-1, CTLA4, and IOD-1 and risk score were evaluated by Pearson correlation analysis. PD-1: Programmed Death receptor 1; PD-L1: Programmed Death Ligand 1; CTLA4: Cytotoxic T Lymphocyte Antigen 4; IOD-1: Indoleamine-1.

mainly enriched in metabolic-related pathways (Figure 8D). In conclusion, GSEA results showed that high-risk patients preferred to carcinogenesis, which also provided favorable evidence for the poorer prognosis of patients in the high-risk group than the low-risk group.

Analysis of tumor immune microenvironment and tumor immune cell infiltration

The relationship between risk score and the infiltration levels of immune cells was predicted by evaluating the infiltration levels of 22 immune cells in the high/low risk groups. The results showed that compared with the low-risk group, the activated natural killer (NK) cell infiltration in the high-risk group was lower than that in the low-risk group (Figure 9A). After further analyzing the correlation between the risk score and NK cell infiltration, we found a negative correlation between the two ($R = -0.33$, $p < 0.05$, Figure 9B). At the same time, single sample gene set enrichment analysis (ssGSEA) was applied to the RNA sequencing data of liver cancer tissues derived from TCGA to explore the differences of 16 kinds of immune cells and 13 kinds of immune functions in the high- and low-risk groups. The results showed that the levels of aDCs and macrophages infiltration in the high-risk group were significantly higher than those in the low-risk group, the infiltration levels of mast cells, neutrophils, and NK cells in the low-risk group were significantly higher than those in the high-risk group (Figure 9C). In terms of immune function, the scores of cytolytic activity (CYT), type I interferon response, and type II interferon response in the low-risk group were significantly higher than the high-risk group ($p < 0.01$), whereas the major histocompatibility antigen complex I score in the high-risk group was higher than that in the low-risk group ($p < 0.01$) (Figure 9D). Tumor microenvironment plays an important role in the development of tumors, including tumor cells, immune cells, stromal cells, extracellular matrix and so on. In this study, the ESTIMATE algorithm was used to quantify the ratio of stromal cells to immune cells in tumor samples from high- and low-risk groups. “Stromal score” represents stromal cells in tumor tissue and “immune Score” represents immune cells in tumor tissue. The results showed that in the low-risk group, the scores of stromal cells were significantly higher than immunocytes (Figure 9E). Although there was no statistical difference in the immune cell scores between the high-risk group and the low-risk group, the comprehensive scores of stromal cells and immune cells in the low-risk group were higher than the high-risk group.

Patients with high risk were insensitive to sorafenib

To determine whether risk score affected tolerance to sorafenib treatment and if the screened patients could benefit

from sorafenib treatment, we determined the value of IC50 in each group. Results suggested that patients with high risk were less sensitive to sorafenib than patients of the low-risk group (Figure 9F).

The prognostic model predicted differential expression of immune checkpoint

At present, immunotherapy has brought the dawn to liver cancers. The expression of immune checkpoint molecules is an important basis for determining whether patients could benefit from immunotherapy. In this study, we evaluated the expression levels of four key immune checkpoint molecules in high- and low-risk groups. As presented in Figure 10A–D, the expressions of PD1, PD-L1, CTLA4, and IDO-1 in the high-risk group were significantly higher than the low-risk group. In addition, risk score was positively associated with the expression of PD1 ($r = 0.3$; $p < 0.001$), PD-L1 ($R = 0.27$; $p < 0.001$), CTLA4 ($R = 0.33$; $p < 0.001$), and IDO-1 ($r = 0.17$; $p < 0.001$). The higher the risk score, the higher the expression of these four immune checkpoints.

The differential expression of PRlncRNAs in cell lines and HCC tissues

In the preliminary work, a prognostic model based on 10 necroptosis-related lncRNA was constructed, internal validation method was used to validate this model. Now, we further verified these results *in vitro*, seven lncRNAs of the model were subjected to quantitative reverse transcription PCR (qRT-PCR). Compared with LO2 cell, the expression levels of seven lncRNAs in four types of HCC cells were significantly differential (Figure 11A–G). Based on the TCGA database, we further verified the expression differences of AL031985.3, SREBF2-AS1, ZFPM2-AS1, KDM4A-AS1, AC026412.3, AC145207.5, DUXAP8, LINC01224, AC099850.4, and MKLN1-AS in normal liver tissue and HCC tissue. As shown in the results, the expression levels of these 10 lncRNA in liver cancer were significantly increased compared with normal liver tissue (Supplementary Figure S1).

Discussion

Liver cancer is one of the most common causes of cancer death in the world. Most of the HCC cases are formed in the context of chronic viral liver diseases. Despite the development of effective antiviral therapy, liver cancer continues to rise. Non-alcoholic fatty liver disease (NAFLD), alcohol-related liver disease, and hereditary hemochromatosis are closely related to the incidence of liver cancer. The global burden of HCC is increasing, although about 40% of patients with primary cancer can be diagnosed early due to primary prevention, but after resection of the radical treatment, both surgery and liver transplantation, the residual liver is still prone to recurrence. As a result, almost half of liver cancer patients will eventually receive systemic treatment. Sorafenib is a kind of tyrosine kinase approved for patients diagnosed with advanced liver cancer and recommended as a standard first-line systematic treatment for patients with Child–Pugh A grade and BCLC-C (Llovet et al., 2008). In recent years, lenvatinib, the new first-line targeted drug, has been approved by FDA (Kudo et al., 2018), regorafenib (Bruix et al., 2017), cabozantinib (Abou-Alfa et al., 2018), and ramucirumab (Zhu et al., 2019) of the second-line drugs have been shown to improve clinical prognosis, although the median survival time is about 1 year. Therefore, we still need to find a breakthrough in the treatment of advanced liver cancer. In recent years, the field of immunotherapy has developed rapidly. Immune checkpoint inhibitors could kill tumors and prolong the prognosis of patients by awakening the immune system to produce anti-tumor immune response against progressive tumors, especially if a significant effect has been achieved in metastatic melanoma immunotherapy (doi: 10.1016/S1470-2045(0970334-1), Epub 2009 December 8).

The laboratory test including serum biomarkers, histochemical biomarkers, and other biomarkers are of great significance for early diagnosis, treatment decision-making, and prognosis prediction of HCC patients. Nowadays, AFP, AFP-L3, and DCP have become the most widely used serum biomarkers for early diagnosis of liver cancer (Yamamoto et al., 2010). AFP plays a guide role in the diagnosis of liver cancer, but its sensitivity and specificity are poor (Tzartzeva et al., 2018). In recent years, many scholars have devoted themselves to looking for new potential molecular markers. A previous study found that TP53 gene mutation is related to vascular invasion, poor prognosis, angiogenesis, metastasis, and drug resistance (Calderaro et al., 2017). The high expression of TERT can be used as an important prognostic index for intrahepatic metastasis of advanced hepatocellular carcinoma after radical resection (Yu et al., 2017). Biomarkers are helpful for guiding individualized targeted therapy, which could improve prognosis. Sorafenib resistance is frequently encountered in clinical practice, but there is still lack of suitable and validated biomarkers. Therefore, development of new biomarkers has important implications for clinical practice.

Previous studies have shown that RIPK3, RIPK1, and MLKL are the most important factors involved in necroptotic, playing important roles during cancer progression (Gong et al., 2019; Tan et al., 2020; Wu et al., 2020; Martens et al., 2021). It was found that the downregulated expression of RIPK3 occurs in M1 macrophages and facilitates tumorigenesis (Wu et al., 2020). RIPK1 leads to excessive activation of Caspase-8 and promotes hepatocellular carcinoma progression (Tan et al., 2020). Necrosome formed by RIPK1, RIPK3 and MLKL

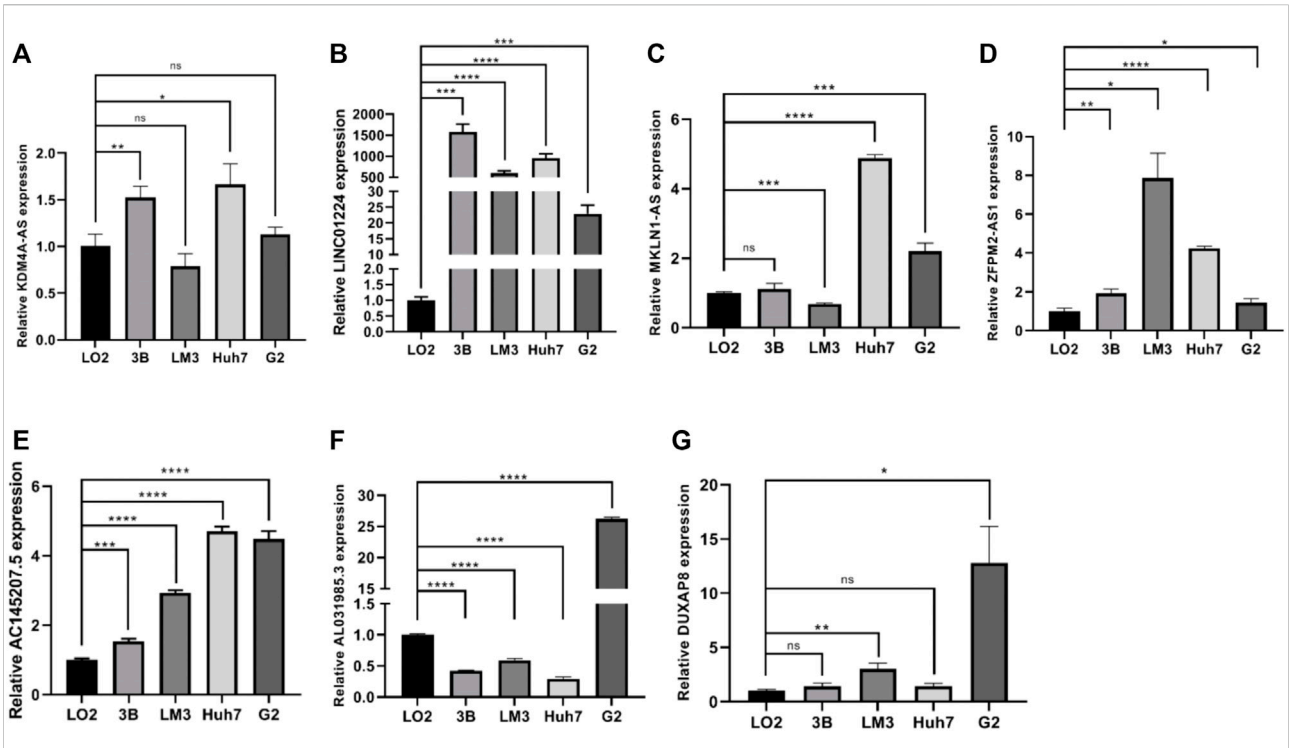


FIGURE 11
(A–G) The expression differences of KDM4A-AS1, LINC01224, MKLN1-AS, ZFPM2-AS1, AC145207.5, AL031985.3, and DUXAP8 in normal liver tissue and HCC tissue.

causing cells are susceptible to necroptosis by enhancing reactive production of oxygen species (Weinlich et al., 2017; Gong et al., 2019). In addition, it has been proved that some drugs inhibited the progression of liver cancer through PIPK1/PIPK3/MLKL signal. For example, triptolide (TPL) loaded into MP-LP has been proved to inhibit tumor growth effectively (Zheng et al., 2021). In summary, necroptotic affects the occurrence and development of liver cancer, especially necroptotic-related genes can serve as potential prognosis predictive biomarkers for HCC patients. Nowadays, some studies focus on the construction of prognosis model based on lncRNAs with different characteristics. For example, ferroptosis-related lncRNAs were associated with the prognosis and immune landscape of HCC patients (Xu et al., 2021). A prognostic model constructed by seven immune-related lncRNAs could be used as an independent prognostic evaluation index for HCC patients. Chen G. et al. (2021) identified three autophagy-associated lncRNAs (MIR210HG, AC099850.3, and CYTOR) as poor prognostic factors for HCC, demonstrating a shorter OS in high-risk groups. LncRNA-related prognostic models have been piled up in the literature, but there are no researches on necroptotic-related lncRNAs in HCC. Therefore, we constructed a necroptotic-related lncRNA prognostic model in this study, aimed to explore prognostic factors and a novel model to better guide immunotherapy and optimize patient outcomes.

In this study, we downloaded liver cancer transcriptome data and clinical data from TCGA database, converted the transcriptome data into TPM format by R language, and calculated the expression of each sequenced gene in each sample. In our study, 58 DEGs affected prognosis were significantly obtained by univariate Cox regression analysis. Then, we applied Lasso-Cox method to filter the collinear factors and obtained 10 candidate lncRNAs (AL031985.3, SREBF2-AS1, ZFPM2-AS1, KDM4A-AS1, AC026412.3, AC145207.5, DUXAP8, LINC01224, AC099850.4, and MKLN1-AS), from which the prognostic risk model was constructed. All the samples were randomly divided into experimental group and verification group, then each group was divided into high-risk subgroup and low-risk subgroup. By comparing the OS of patients in the high-risk group and the low-risk group, we found that the patients with high-risk score exhibited a shorter survival rate. The internal verification group verified the results again, we concluded that the risk score is a poor prognostic factor for HCC patients. In the clinical correlation analysis, risk score was significantly correlated with histological grade and cluster classification. ROC analysis further revealed that the risk signature we constructed could function as a sensitive indicator predicting

1-, 2-, and 3-year survival rates for the HCC patients. At the same time, we further compared the predictive effectiveness of the prognostic risk score with other clinical prognostic factors, we can see that among selected candidate factors, risk score performed optimally. At the same time, results of multivariate Cox regression analysis proved that the risk score is an independent risk factor in predicting the prognosis of HCC patients. Among all candidate lncRNAs in our model, AL031985.3 and AC145207.5 have been proved to play an important role in glycolysis-related prognostic models of liver cancer (Xia et al., 2021). The value of SREBF2-AS1 in HCC remains to be confirmed, but a study has found it can be used as ferroptosis-related lncRNA to predict the prognosis and immune activity of HCC (Chen Z. A. et al., 2021). ZFPM2-AS1 is highly expressed in HCC, which promotes the proliferation and migration of HCC cells by up-regulating hypoxia-inducible factor (HIP-1a) (Song et al., 2021). KDM4A-AS1 promotes HCC progression by activating the AKT pathway to promote the expression of a2KPNA2 (Chen T. et al., 2021). AC145207.5, as an immune-related lncRNA, could predict tumor immune infiltration cell and response to immunotherapy in HCC patients (Zhou et al., 2021). Hu et al. (2020) found that DUXAP8 promotes varieties of malignant phenotypes in HCC and resistance to PARP inhibitors by up-regulating FOXM1. There is growing experimental evidence that LINC01224 plays a role in promoting tumor progression in melanoma and colon cancer and shows resistance to radiotherapy (Chen L. et al., 2021; Cui et al., 2022). MKLN1-AS acts as an endogenous sponge of miR-654-3p to promote the progression of HCC (Gao et al., 2020). There is a paucity of studies with AC026412.3 to date, relevant researches about AC026412.3 need further excavation. Most of existing studies on the prognosis lncRNAs we obtained are remaind at the level of prognostic model, there are few studies have further explored the mechanism of these genes *in vivo* and *in vitro*, hence those lncRNAs may be clinically valuable and are worth further studied. Hence, the studies of the biological function and potential mechanism of lncRNAs in liver cancer are extremely important for the exploration of new therapeutic targets for liver cancer.

In terms of mechanism, in order to further explain the mechanism of DEGs, all DEGs were subjected to GO and KEGG enrichment analyses. Firstly, results revealed that the biological functions of the majority of the DEGs are closely related to the process of cell cycle, including organelle division, mitosis, and chromosome segregation. Secondly, chromosome, microtubule, and spindle structure were the most dominant terms in the cellular component. While "ATP enzyme activity," "small GTP enzyme binding activity," and "RASGTP enzyme binding activity" were the most abundant terms in the molecular function category. From the above results, it can be concluded that DEGs play an important role in the process of cell division. We do know that the occurrence and development of tumors are usually related to the abnormal accumulation of genes, these abnormal mutations usually occur when cell cycle processes were disrupted (Stratton et al., 2009; Suski et al., 2021). The results of KEGG analysis suggested that DEGs are mainly enriched in viral infection-related pathways, especially the pathways of "herpes simplex virus," "human papillomavirus," and "coronavirus-COVID-19," which have attracted worldwide attention in the past 3 years. It is well known that hepatitis B virus infection is a main risk factor of HCC in China (Levrero and Zucman-Rossi, 2016), so it can be speculated that DEGs may be involved in the process of "hepatitis B-liver cirrhosis-liver cancer trilogy." At present, few people studied the role of other viruses in the occurrence and development of liver cancer. HPV disrupts the normal cell cycle and promotes the accumulation of genetic damage, leading to tumorigenesis (Araldi et al., 2018; Szymonowicz and Chen, 2020). Previous studies have confirmed that persistent infection of papillomavirus will lead to the occurrence of cervical cancer (Araldi et al., 2018). As with hepatitis B virus, human papillomavirus is a DNA virus, which integrated into the host genome after invading human body (Szymonowicz and Chen, 2020). The results of KEGG gave us new enlightenments: DEGs may be involved in the viral infection pathways and this process may occur to hepatocarcinogenesis.

Liver is an organ that plays a dual role of metabolism and immunoregulatory. There are a large number of immune cells in the liver, including macrophages (Kupffer cells), liver sinusoidal endothelial cells (LSECs), and natural killer cells (Jenne and Kubes, 2013), LSECs act as antigen-presenting cells (APCs) along with Kupffer cells and dendritic cells (DCs) (Thomson and Knolle, 2010). In addition, the liver contains a large population of liver-resident lymphocytes, encompassing NK cells and T cells, exerting innate immune responses against viruses, intracellular bacteria, tumors, and parasites (Gao et al., 2008), which play a protective role in the liver by exerting cytotoxic activity. Tumor microenvironment is a complex and constant evolution, in addition to stromal cells, fibroblasts, and endothelial cells, TME also includes innate and acquired immune cells. Increasing evidence suggests that innate immune cells and acquired immune cells promote tumor progression in tumor microenvironment. Studies have shown that "hot tumors" recognize and attack tumors by abundant $CD4^+$ and $CD8^+$ tumor infiltrating lymphocytes, which are generally associated with better prognosis as well as improved response to ICB (Chen and Mellman, 2017). Therefore, the level of immune infiltrating cells in tumor microenvironment has implications for host anti-tumor response and immunotherapy. In this study, we explored the differences in the level of immune infiltrating cells between high- and low-risk groups. The level of NK cells infiltrated in the high-risk group was significantly lower than that in the low-risk group. By further analyzing the correlation between the level of NK cell and risk score, it was found that there was a significant negative correlation. NK cells are known to play a key role in innate immune surveillance against tumors. NK cells account for half of the number of liver

lymphocytes, they are cytotoxic killer cells with the function of anti-tumor, leading to an efficient antitumor activity mediated by releasing of cytotoxic granules, TRAIL, and Fas-L (Robinson et al., 2016). Such studies have shown that exocrine circUHRF1 inhibits the function of NK cells by degrading miR-449c-5p and up-regulating the expression of TIM-3, resulting in drug resistance to anti-PD-1 immunotherapy in HCC patients (Zhang et al., 2020). In our study, patients with high risk score to be the lower level of NK cell infiltration. Therefore, we speculated that patients in the high-risk group presented with weak immune surveillance ability, triggering the initiation of cancer in susceptible individuals and be more likely to develop drug resistance to treatment. Other studies have shown that approximately 25% of HCC patients with a high inflammatory score showed high or moderate levels of lymphocyte infiltration (Sia et al., 2017). Tumor infiltrating lymphocytes (TILs) constituted the main component of solid tumor microenvironment and mediated anti-tumor response (Qin, 2012). However, such cellular response could be dysfunctional due to the increased $CD4^+/CD8^+$ T cells, which leads to immune tolerance and confers an even poorer prognosis (Fu et al., 2007). Similar results were observed in our analysis: The infiltration levels of activated dendritic cells, tumor-associated macrophages, and Th2 cells in the high-risk group were significantly higher than the low-risk group, while the infiltration levels of mast cells, neutrophils, and NK cells in the low-risk group were significantly higher than the high-risk group. It has been found that chemotherapy-induced immunogenic cell death results in the release of stimulators, enhancing DC subsets cross-present antigen to CD8 + T cells, and thus augmenting CD8 + T cell responses (Sánchez-Paulete et al., 2017). Moreover, a recent report showed that the infiltration level of DC is associated with the ICB-based immunotherapy (Barry et al., 2018). At first, it is believed that macrophages played an anti-tumor effect. After years of clinical trials and fundamental studies, it has been proved that macrophages promote cancer progression in most of the cases. In a meta-analysis, more than 80% of studies found that the density of M2 phenotype macrophages is correlated with poor prognosis (Bingle et al., 2002). Neutrophils reflect the inflammatory state of the host, which is one of the hallmarks of cancer (Hanahan and Weinberg, 2011). The high expression of neutrophils is related to the poor prognosis of several solid tumors, some researchers believe that neutrophils promote the occurrence and development of tumors by releasing reactive oxygen species (ROS) and reactive nitrogen species(RNS) (Antonio et al., 2015), they also facilitate tumor cell metastasis by inhibiting function of natural killer and promoting tumor cell extravasation (Welch et al., 1989; Spiegel et al., 2016). However, in a prospective clinical study, univariate COX regression analysis of potential prognostic factors showed that neutrophils are significantly associated with superior overall survival in gastric cancer subjects (Caruso et al., 2002). In conclusion, we evaluated immune cells in the tumor microenvironment of patients with different risk scores, our result has important practical implications in predicting prognosis and therapeutic.

In the course of cancer occurrence and development, with the evolution of the tumor, tumor cells have acquired a variety of mechanisms to evade immune surveillance and inhibited anti-tumor immune response to escape attacks from the host immune system. Immune checkpoint inhibitors have brought considerable clinical benefits to patients and promoted the development of oncology greatly. Despite immune checkpoint inhibitors for the clinical treatment of cancer has led to durable responses for some patients, only a small fraction of patient response to immune checkpoint inhibitors. There are many factors that are involved in the efficacy of immune checkpoint inhibitors treatments, including infiltration of cytotoxic T cells (Van Allen et al., 2015), neoantigen load and mutation frequency (Snyder et al., 2014), serum PD-L1 levels (Nishino et al., 2017), and mismatch repair deficiency (Le et al., 2015). Indoleamine 2,3 dioxygenase 1 (IDO1) as an immune checkpoint molecule converts tryptophan to kynurenine, which is associated with tumor immunosuppression. In a phase II clinical study, it was found that most sarcoma cells were infiltrated by IDO1-expressing M2 macrophages, resulting in a low response to PD-(L) 1 inhibitor (Munn and Mellor, 2016; Li et al., 2017; Toulmonde et al., 2018). However, none of these factors are sufficient to achieve accurate prediction. At present, identification of immunotherapy biomarkers and regulators modulating resistance is a critical challenge in this field. In this study, we analyzed the expression levels of four immune checkpoint molecules in high- and low-risk groups. The results showed that the expression of immune checkpoint molecules in the high-risk group were higher than that in the low-risk group. High expression of immune checkpoints is usually associated with immune escape and drug resistance (Peng et al., 2019; Lv et al., 2021), suggesting that patients in the high-risk group might be more likely to immune escape and drug resistance, and may not be suitable for immunotherapy. The treatment options of patients with advanced liver cancer are limited, immunotherapy may bring hope to some patients, but our study found that patients with high-risk are more likely to develop drug resistance, which speculated that the prognosis of patients with high-risk is worse.

In the end, we analyze lncRNA expression in normal hepatic and HCC cells to verify the accuracy of our model. Our RT-qPCR data showing that the expressions of seven lncRNAs were significant different in HCC cells compared to the normal hepatic cells. All these data suggested that our signature has a crucial role in HCC development and progression. All in all, this work has developed a prognostic prediction model for necroptosis-related lncRNAs by analyzing data from the TCGA public database, providing references for patients' prognosis and clinical guidance, the limitations of this study should be considered. Although results were validated in the

TCGA cohort, its reliability needs to be further verified in other independent cohorts. In addition, the ability of the necroptosis-related lncRNAs improving the efficacy of immunotherapy in HCC patients has not been proved, we need to further confirm it by *in vivo* and *in vitro* experiments in the future. Secondly, expression levels of necroptosis-related lncRNAs in biological specimens have not been clearly validated. Finally, in the future, we need to further explore the potential mechanism behind the prognostic model affecting the process of HCC, so as to provide a new target and therapy for the clinical treatment of HCC, improving systemic treatment efficacy and thus prolong the overall survival time of HCC 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

YP and AL contributed to the concept and design of study. GW and YP performed the experiment and wrote the manuscript. XQ and YP are responsible for the statistical analysis. XQ and YP assisted in writing the manuscript. YL, YZ, and XW revised and checked the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by Science and Technology Program of Guangzhou (NO. 202002030075); and President's Fund of Integrated Hospital of Traditional Chinese Medicine (1202103008).

Acknowledgments

The authors thank all researchers and teams managing the TCGA, CIBERSORT, and GSEA 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

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.

References

Abou-Alfa, G. K., Meyer, T., Cheng, A. L., El-Khoueiry, A. B., Rimassa, L., Ryoo, B. Y., et al. (2018). Cabozantinib in patients with advanced and progressing hepatocellular carcinoma. *N. Engl. J. Med.* 379, 54–63. doi:10.1056/NEJMoa1717002

Antonio, N., Bønnelykke-Behrndtz, M. L., Ward, L. C., Collin, J., Christensen, I. J., Steiniche, T., et al. (2015). The wound inflammatory response exacerbates growth of pre-neoplastic cells and progression to cancer. *Embo J.* 34, 2219–2236. doi:10.15252/embj.201490147

Araldi, R. P., Sant'ana, T. A., Módolo, D. G., De Melo, T. C., Spadacci-Morena, D. D., De Cassia Stocco, R., et al. (2018). The human papillomavirus (HPV)-related cancer biology: An overview. *Biomed. Pharmacother.* 106, 1537–1556. doi:10.1016/j.biopha.2018.06.149

Barry, K. C., Hsu, J., Broz, M. L., Cueto, F. J., Binnewies, M., Combes, A. J., et al. (2018). A natural killer-dendritic cell axis defines checkpoint therapy-responsive tumor microenvironments. *Nat. Med.* 24, 1178–1191. doi:10.1038/s41591-018-0085-8

Bingle, L., Brown, N. J., and Lewis, C. E. (2002). The role of tumour-associated macrophages in tumour progression: Implications for new anticancer therapies. *J. Pathol.* 196, 254–265. doi:10.1002/path.1027

Bruix, J., Qin, S., Merle, P., Granito, A., Huang, Y. H., Bodoky, G., et al. (2017). Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): A randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* 389, 56–66. doi:10.1016/S0140-6736(16)32453-9

Brunner, A. L., Beck, A. H., Edris, B., Sweeney, R. T., Zhu, S. X., Li, R., et al. (2012). Transcriptional profiling of long non-coding RNAs and novel transcribed regions across a diverse panel of archived human cancers. *Genome Biol.* 13, R75. doi:10.1186/gb-2012-13-8-r75

Calderaro, J., Couchy, G., Imbeaud, S., Amaddeo, G., Letouzé, E., Blanc, J. F., et al. (2017). Histological subtypes of hepatocellular carcinoma are related to gene mutations and molecular tumour classification. *J. Hepatol.* 67, 727–738. doi:10.1016/j.jhep.2017.05.014

Caruso, R. A., Bellocco, R., Pagano, M., Bertoli, G., Rigoli, L., Inferrera, C., et al. (2002). Prognostic value of intratumoral neutrophils in advanced gastric carcinoma in a high-risk area in northern Italy. *Mod. Pathol.* 15, 831–837. doi:10.1097/01.MP.0000020391.98998.6B

Chen, D. S., and Mellman, I. (2017). Elements of cancer immunity and the cancer-immune set point. *Nature* 541, 321–330. doi:10.1038/nature21349

Chen, G., and Goeddel, D. V. (2002). TNF-R1 signaling: A beautiful pathway. *Science* 296, 1634–1635. doi:10.1126/science.1071924

Chen, G., Yang, G., Long, J., Yang, J., Qin, C., Luo, W., et al. (2021a). Comprehensive analysis of autophagy-associated lncRNAs reveal potential prognostic prediction in pancreatic cancer. *Front. Oncol.* 11, 596573. doi:10.3389/fonc.2021.596573

Chen, L., Chen, W., Zhao, C., and Jiang, Q. (2021b). LINC01224 promotes colorectal cancer progression by sponging miR-2467. *Cancer Manag. Res.* 13, 733–742. doi:10.2147/CMAR.S281625

Chen, T., Liu, R., Niu, Y., Mo, H., Wang, H., Lu, Y., et al. (2021c). HIF-1α-activated long non-coding RNA KDM4A-AS1 promotes hepatocellular carcinoma progression via the miR-411-5p/KPNA2/AKT pathway. *Cell. Death Dis.* 12, 1152. doi:10.1038/s41419-021-04449-2

Chen, W., Zheng, R., Baade, P. D., Zhang, S., Zeng, H., Bray, F., et al. (2016). Cancer statistics in China, 2015. *Ca. Cancer J. Clin.* 66, 115–132. doi:10.3322/caac.21338

Chen, Z. A., Tian, H., Yao, D. M., Zhang, Y., Feng, Z. J., Yang, C. J., et al. (2021d). Identification of a ferroptosis-related signature model including mRNAs and lncRNAs for predicting prognosis and immune activity in hepatocellular carcinoma. *Front. Oncol.* 11, 738477. doi:10.3389/fonc.2021.738477

Cho, Y. S., Challa, S., Moquin, D., Genga, R., Ray, T. D., Guildford, M., et al. (2009). Phosphorylation-driven assembly of the RIP1-RIP3 complex regulates programmed necrosis and virus-induced inflammation. *Cell.* 137, 1112–1123. doi:10.1016/j.cell.2009.05.037

Choi, M. E., Price, D. R., Ryter, S. W., and Choi, A. M. K. (2019). Necroptosis: A crucial pathogenic mediator of human disease. *JCI Insight* 4, 128834. doi:10.1172/jci.insight.128834

Christofferson, D. E., Li, Y., and Yuan, J. (2014). Control of life-or-death decisions by RIP1 kinase. *Annu. Rev. Physiol.* 76, 129–150. doi:10.1146/annurev-physiol-021113-170259

Cui, Y., Zheng, Y., Lu, Y., Zhang, M., Yang, L., Li, W., et al. (2022). LINC01224 facilitates the proliferation and inhibits the radiosensitivity of melanoma cells through the miR-193a-5p/NR1D2 axis. *Kaohsiung J. Med. Sci.* 38, 196–206. doi:10.1002/kjm2.12467

Degterev, A., Huang, Z., Boyce, M., Li, Y., Jagtap, P., Mizushima, N., et al. (2005). Chemical inhibitor of nonapoptotic cell death with therapeutic potential for ischemic brain injury. *Nat. Chem. Biol.* 1, 112–119. doi:10.1038/nchembio711

El-Khoueiry, A. B., Sangro, B., Yau, T., Crocenzi, T. S., Kudo, M., Hsu, C., et al. (2017). Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): An open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *Lancet* 389, 2492–2502. doi:10.1016/S0140-6736(17)31046-2

Fang, Y., and Fullwood, M. J. (2016). Roles, functions, and mechanisms of long non-coding RNAs in cancer. *Genomics Proteomics Bioinforma.* 14, 42–54. doi:10.1016/j.gpb.2015.09.006

Finn, R. S., Qin, S., Ikeda, M., Galle, P. R., Ducreux, M., Kim, T. Y., et al. (2020). Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *N. Engl. J. Med.* 382, 1894–1905. doi:10.1056/NEJMoa1915745

Fritsch, M., Günther, S. D., Schwarzer, R., Albert, M. C., Schorn, F., Werthenbach, J. P., et al. (2019). Caspase-8 is the molecular switch for apoptosis, necroptosis and pyroptosis. *Nature* 575, 683–687. doi:10.1038/s41586-019-1770-6

Fu, J., Xu, D., Liu, Z., Shi, M., Zhao, P., Fu, B., et al. (2007). Increased regulatory T cells correlate with CD8 T-cell impairment and poor survival in hepatocellular carcinoma patients. *Gastroenterology* 132, 2328–2339. doi:10.1053/j.gastro.2007.03.102

Gao, B., Jeong, W. I., and Tian, Z. (2008). Liver: An organ with predominant innate immunity. *Hepatology* 47, 729–736. doi:10.1002/hep.22034

Gao, W., Chen, X., Chi, W., and Xue, M. (2020). Long non-coding RNA MKLN1-AS aggravates hepatocellular carcinoma progression by functioning as a molecular sponge for miR-654-3p, thereby promoting hepatoma-derived growth factor expression. *Int. J. Mol. Med.* 46, 1743–1754. doi:10.3892/ijmm.2020.4722

Gaur, U., and Aggarwal, B. B. (2003). Regulation of proliferation, survival and apoptosis by members of the TNF superfamily. *Biochem. Pharmacol.* 66, 1403–1408. doi:10.1016/s0006-2952(03)00490-8

Gong, Y., Fan, Z., Luo, G., Yang, C., Huang, Q., Fan, K., et al. (2019). The role of necroptosis in cancer biology and therapy. *Mol. Cancer* 18, 100. doi:10.1186/s12943-019-1029-8

Gong, Y. N., Guy, C., Crawford, J. C., and Green, D. R. (2017). Biological events and molecular signaling following MLKL activation during necroptosis. *Cell. Cycle* 16, 1748–1760. doi:10.1080/15384101.2017.1371889

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell.* 144, 646–674. doi:10.1016/j.cell.2011.02.013

He, S., Wang, L., Miao, L., Wang, T., Du, F., Zhao, L., et al. (2009). Receptor interacting protein kinase-3 determines cellular necrotic response to TNF-alpha. *Cell.* 137, 1100–1111. doi:10.1016/j.cell.2009.05.021

Hong, Y. P., Li, Z. D., Prasoon, P., and Zhang, Q. (2015). Immunotherapy for hepatocellular carcinoma: From basic research to clinical use. *World J. Hepatol.* 7, 980–992. doi:10.4254/wjh.v7.i7.980

Hou, J., Ju, J., Zhang, Z., Zhao, C., Li, Z., Zheng, J., et al. (2019). Discovery of potent necroptosis inhibitors targeting RIPK1 kinase activity for the treatment of inflammatory disorder and cancer metastasis. *Cell. Death Dis.* 10, 493. doi:10.1038/s41419-019-1735-6

Hsu, H., Xiong, J., and Goeddel, D. V. (1995). The TNF receptor 1-associated protein TRADD signals cell death and NF-kappa B activation. *Cell.* 81, 495–504. doi:10.1016/0092-8674(95)90070-5

Hu, Y., Zhang, X., Zai, H. Y., Jiang, W., Xiao, L., Zhu, Q., et al. (2020). lncRNA DUXAP8 facilitates multiple malignant phenotypes and resistance to PARP inhibitor in HCC via upregulating FOXM1. *Mol. Ther. Oncolytics* 19, 308–322. doi:10.1016/j.omto.2020.10.010

Jenne, C. N., and Kubes, P. (2013). Immune surveillance by the liver. *Nat. Immunol.* 14, 996–1006. doi:10.1038/ni.2691

Jiao, D., Cai, Z., Choksi, S., Ma, D., Choe, M., Kwon, H. J., et al. (2018). Necroptosis of tumor cells leads to tumor necrosis and promotes tumor metastasis. *Cell. Res.* 28, 868–870. doi:10.1038/s41422-018-0058-y

Jin, C., Li, R., Deng, T., Li, J., Yang, Y., Li, H., et al. (2021). Identification and validation of a prognostic prediction model of m6A regulator-related LncRNAs in hepatocellular carcinoma. *Front. Mol. Biosci.* 8, 784553. doi:10.3389/fmolb.2021.784553

Kudo, M., Finn, R. S., Qin, S., Han, K. H., Ikeda, K., Piscaglia, F., et al. (2018). Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: A randomised phase 3 non-inferiority trial. *Lancet* 391, 1163–1173. doi:10.1016/S0140-6736(18)30207-1

Le, D. T., Uram, J. N., Wang, H., Bartlett, B. R., Kemberling, H., Eyring, A. D., et al. (2015). PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* 372, 2509–2520. doi:10.1056/NEJMoa1500596

Levrero, M., and Zucman-Rossi, J. (2016). Mechanisms of HBV-induced hepatocellular carcinoma. *J. Hepatol.* 64, S84–S101. doi:10.1016/j.jhep.2016.02.021

Li, F., Zhang, R., Li, S., and Liu, J. (2017). IDO1: An important immunotherapy target in cancer treatment. *Int. Immunopharmacol.* 47, 70–77. doi:10.1016/j.intimp.2017.03.024

Li, G., Liu, K., and Du, X. (2018). Long non-coding RNA TUG1 promotes proliferation and inhibits apoptosis of osteosarcoma cells by sponging miR-132-3p and upregulating SOX4 expression. *Yonsei Med. J.* 59, 226–235. doi:10.3349/ymj.2018.59.2.226

Llovet, J. M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J. F., et al. (2008). Sorafenib in advanced hepatocellular carcinoma. *N. Engl. J. Med.* 359, 378–390. doi:10.1056/NEJMoa0708857

Lv, H., Lv, G., Chen, C., Zong, Q., Jiang, G., Ye, D., et al. (2021). NAD(+) metabolism maintains inducible PD-L1 expression to drive tumor immune evasion. *Cell. Metab.* 33, 110–127. e5. doi:10.1016/j.cmet.2020.10.021

Mak, T. W., and Yeh, W. C. (2002). Signaling for survival and apoptosis in the immune system. *Arthritis Res.* 4 (Suppl. 3), S243–S252. doi:10.1186/ar569

Martens, S., Bridelance, J., Roelandt, R., Vandenabeele, P., and Takahashi, N. (2021). MLKL in cancer: more than a necroptosis regulator. *Cell. Death Differ.* 28, 1757–1772. doi:10.1038/s41418-021-00785-0

Mccormick, K. D., Ghosh, A., Trivedi, S., Wang, L., Coyne, C. B., Ferris, R. L., et al. (2016). Innate immune signaling through differential RIPK1 expression promote tumor progression in head and neck squamous cell carcinoma. *Carcinogenesis* 37, 522–529. doi:10.1093/carcin/bgw032

Mcglynn, K. A., Petrick, J. L., and London, W. T. (2015). Global epidemiology of hepatocellular carcinoma: An emphasis on demographic and regional variability. *Clin. Liver Dis.* 19, 223–238. doi:10.1016/j.cld.2015.01.001

Micheau, O., and Tschopp, J. (2003). Induction of TNF receptor I-mediated apoptosis via two sequential signaling complexes. *Cell.* 114, 181–190. doi:10.1016/s0092-8674(03)00521-x

Munn, D. H., and Mellor, A. L. (2016). Ido in the tumor microenvironment: Inflammation, counter-regulation, and tolerance. *Trends Immunol.* 37, 193–207. doi:10.1016/j.it.2016.01.002

Nishino, M., Ramaiya, N. H., Hatabu, H., and Hodi, F. S. (2017). Monitoring immune-checkpoint blockade: response evaluation and biomarker development. *Nat. Rev. Clin. Oncol.* 14, 655–668. doi:10.1038/nrclinonc.2017.88

Parkin, D. M., Bray, F., Ferlay, J., and Pisani, P. (2005). Global cancer statistics, 2002. *Ca. Cancer J. Clin.* 55, 74–108. doi:10.3322/canjclin.55.2.74

Peng, S., Wang, R., Zhang, X., Ma, Y., Zhong, L., Li, K., et al. (2019). EGFR-TKI resistance promotes immune escape in lung cancer via increased PD-L1 expression. *Mol. Cancer* 18, 165. doi:10.1186/s12943-019-1073-4

Qin, L. X. (2012). Inflammatory immune responses in tumor microenvironment and metastasis of hepatocellular carcinoma. *Cancer Microenviron.* 5, 203–209. doi:10.1007/s12307-012-0111-1

Rahib, L., Smith, B. D., Aizenberg, R., Rosenzweig, A. B., Fleshman, J. M., Matrisian, L. M., et al. (2014). Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res.* 74, 2913–2921. doi:10.1158/0008-5472.CAN-14-0155

Robinson, M. W., Harmon, C., and O'farrelly, C. (2016). Liver immunology and its role in inflammation and homeostasis. *Cell. Mol. Immunol.* 13, 267–276. doi:10.1038/cmi.2016.3

Sánchez-Paulete, A. R., Teijeira, A., Cueto, F. J., Garasa, S., Pérez-Gracia, J. L., Sánchez-Arráez, A., et al. (2017). Antigen cross-presentation and T-cell cross-priming in cancer immunology and immunotherapy. *Ann. Oncol.* 28 (12), xii74. doi:10.1093/annonc/mdx237

Schmitt, A. M., and Chang, H. Y. (2016). Long noncoding RNAs in cancer pathways. *Cancer Cell.* 29, 452–463. doi:10.1016/j.ccell.2016.03.010

Seehawer, M., Heinzmann, F., D'artista, L., Harbig, J., Roux, P. F., Hoenicke, L., et al. (2018). Necroptosis microenvironment directs lineage commitment in liver cancer. *Nature* 562, 69–75. doi:10.1038/s41586-018-0519-y

Si, Y., Yang, Z., Ge, Q., Yu, L., Yao, M., Sun, X., et al. (2019). Long non-coding RNA Malat1 activated autophagy, hence promoting cell proliferation and inhibiting apoptosis by sponging miR-101 in colorectal cancer. *Cell. Mol. Biol. Lett.* 24, 50. doi:10.1186/s11658-019-0175-8

Sia, D., Jiao, Y., Martinez-Quetglas, I., Kuchuk, O., Villacorta-Martin, C., Castro De Moura, M., et al. (2017). Identification of an immune-specific class of hepatocellular carcinoma, based on molecular features. *Gastroenterology* 153, 812–826. doi:10.1053/j.gastro.2017.06.007

Snyder, A., Makarov, V., Merghoub, T., Yuan, J., Zaretsky, J. M., Desrichard, A., et al. (2014). Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* 371, 2189–2199. doi:10.1056/NEJMoa1406498

Song, Y., Jin, X., Liu, Y., Wang, S., Bian, F., Zhao, Q., et al. (2021). Long noncoding RNA ZFPM2-AS1 promotes the proliferation, migration, and invasion of hepatocellular carcinoma cells by regulating the miR-576-3p/HIF-1α axis. *Anticancer. Drugs* 32, 812–821. doi:10.1097/CAD.0000000000001070

Spiegel, A., Brooks, M. W., Houshyar, S., Reinhardt, F., Ardolino, M., Fessler, E., et al. (2016). Neutrophils suppress intraluminal NK cell-mediated tumor cell clearance and enhance extravasation of disseminated carcinoma cells. *Cancer Discov.* 6, 630–649. doi:10.1158/2159-8290.CD-15-1157

Stratton, M. R., Campbell, P. J., and Futreal, P. A. (2009). The cancer genome. *Nature* 458, 719–724. doi:10.1038/nature07943

Sun, L., Wang, H., Wang, Z., He, S., Chen, S., Liao, D., et al. (2012). Mixed lineage kinase domain-like protein mediates necrosis signaling downstream of RIP3 kinase. *Cell.* 148, 213–227. doi:10.1016/j.cell.2011.11.031

Suski, J. M., Braun, M., Strmiska, V., and Sicinski, P. (2021). Targeting cell-cycle machinery in cancer. *Cancer Cell.* 39, 759–778. doi:10.1016/j.ccell.2021.03.010

Szymonowicz, K. A., and Chen, J. (2020). Biological and clinical aspects of HPV-related cancers. *Cancer Biol. Med.* 17, 864–878. doi:10.20892/j.issn.2095-3941.2020.0370

Tan, S., Zhao, J., Sun, Z., Cao, S., Niu, K., Zhong, Y., et al. (2020). Hepatocyte-specific TAK1 deficiency drives RIPK1 kinase-dependent inflammation to promote liver fibrosis and hepatocellular carcinoma. *Proc. Natl. Acad. Sci. U. S. A.* 117, 14231–14242. doi:10.1073/pnas.2005353117

Teschendorff, A. E., Lee, S. H., Jones, A., Fiegl, H., Kalwa, M., Wagner, W., et al. (2015). HOTAIR and its surrogate DNA methylation signature indicate carboplatin resistance in ovarian cancer. *Genome Med.* 7, 108. doi:10.1186/s13073-015-0233-4

Thomson, A. W., and Knolle, P. A. (2010). Antigen-presenting cell function in the tolerogenic liver environment. *Nat. Rev. Immunol.* 10, 753–766. doi:10.1038/nri2858

Toulmonde, M., Penel, N., Adam, J., Chevreau, C., Blay, J. Y., Le Cesne, A., et al. (2018). Use of PD-1 targeting, macrophage infiltration, and Ido pathway activation in sarcomas: A phase 2 clinical trial. *JAMA Oncol.* 4, 93–97. doi:10.1001/jamaoncol.2017.1617

Tzartzeva, K., Obi, J., Rich, N. E., Parikh, N. D., Marrero, J. A., Yopp, A., et al. (2018). Surveillance imaging and alpha fetoprotein for early detection of hepatocellular carcinoma in patients with cirrhosis: A meta-analysis. *Gastroenterology* 154, 1706–1718. doi:10.1053/j.gastro.2018.01.064

Van Allen, E. M., Miao, D., Schilling, B., Shukla, S. A., Blank, C., Zimmer, L., et al. (2015). Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 350, 207–211. doi:10.1126/science.aad0095

Vandevoorde, V., Haegeman, G., and Fiers, W. (1997). Induced expression of trimerized intracellular domains of the human tumor necrosis factor (TNF) p55 receptor elicits TNF effects. *J. Cell. Biol.* 137, 1627–1638. doi:10.1083/jcb.137.7.1627

Wallach, D., Arumugam, T. U., Boldin, M. P., Cantarella, G., Ganesh, K. A., Goltsev, Y., et al. (2002). How are the regulators regulated? The search for mechanisms that impose specificity on induction of cell death and NF-kappaB activation by members of the TNF/NGF receptor family. *Arthritis Res.* 4 (Suppl. 3), S189–S196. doi:10.1186/ar585

Wang, C., Qiu, J., Chen, S., Li, Y., Hu, H., Cai, Y., et al. (2021). Prognostic model and nomogram construction based on autophagy signatures in lower grade glioma. *J. Cell. Physiol.* 236, 235–248. doi:10.1002/jcp.29837

Weinlich, R., and Green, D. R. (2014). The two faces of receptor interacting protein kinase-1. *Mol. Cell.* 56, 469–480. doi:10.1016/j.molcel.2014.11.001

Weinlich, R., Oberst, A., Beere, H. M., and Green, D. R. (2017). Necroptosis in development, inflammation and disease. *Nat. Rev. Mol. Cell. Biol.* 18, 127–136. doi:10.1038/nrm.2016.149

Welch, D. R., Schissel, D. J., Howrey, R. P., and Aeed, P. A. (1989). Tumor-elicited polymorphonuclear cells, in contrast to "normal" circulating polymorphonuclear cells, stimulate invasive and metastatic potentials of rat mammary adenocarcinoma cells. *Proc. Natl. Acad. Sci. U. S. A.* 86, 5859–5863. doi:10.1073/pnas.86.15.5859

Wu, L., Zhang, X., Zheng, L., Zhao, H., Yan, G., Zhang, Q., et al. (2020). RIPK3 orchestrates fatty acid metabolism in tumor-associated macrophages and hepatocarcinogenesis. *Cancer Immunol. Res.* 8, 710–721. doi:10.1158/2326-6066.CIR-19-0261

Xia, X., Zhang, H., Xia, P., Zhu, Y., Liu, J., Xu, K., et al. (2021). Identification of glycolysis-related lncRNAs and the novel lncRNA WAC-AS1 promotes glycolysis and tumor progression in hepatocellular carcinoma. *Front. Oncol.* 11, 733595. doi:10.3389/fonc.2021.733595

Xu, Z., Peng, B., Liang, Q., Chen, X., Cai, Y., Zeng, S., et al. (2021). Construction of a ferroptosis-related nine-lncRNA signature for predicting prognosis and immune response in hepatocellular carcinoma. *Front. Immunol.* 12, 719175. doi:10.3389/fimmu.2021.719175

Yamamoto, K., Imamura, H., Matsuyama, Y., Kume, Y., Ikeda, H., Norman, G. L., et al. (2010). AFP, AFP-L3, DCP, and GP73 as markers for monitoring treatment response and recurrence and as surrogate markers of clinicopathological variables of HCC. *J. Gastroenterol.* 45, 1272–1282. doi:10.1007/s00535-010-0278-5

Yu, J. I., Choi, C., Ha, S. Y., Park, C. K., Kang, S. Y., Joh, J. W., et al. (2017). Clinical importance of TERT overexpression in hepatocellular carcinoma treated with curative surgical resection in HBV endemic area. *Sci. Rep.* 7, 12258. doi:10.1038/s41598-017-12469-2

Zhang, D. W., Shao, J., Lin, J., Zhang, N., Lu, B. J., Lin, S. C., et al. (2009). RIP3, an energy metabolism regulator that switches TNF-induced cell death from apoptosis to necrosis. *Science* 325, 332–336. doi:10.1126/science.1172308

Zhang, P. F., Gao, C., Huang, X. Y., Lu, J. C., Guo, X. J., Shi, G. M., et al. (2020). Cancer cell-derived exosomal circUHRF1 induces natural killer cell exhaustion and may cause resistance to anti-PD1 therapy in hepatocellular carcinoma. *Mol. Cancer* 19, 110. doi:10.1186/s12943-020-01222-5

Zhang, Z., Zhou, C., Chang, Y., Zhang, Z., Hu, Y., Zhang, F., et al. (2016). Long non-coding RNA CASC11 interacts with hnRNP-K and activates the WNT/β-catenin pathway to promote growth and metastasis in colorectal cancer. *Cancer Lett.* 376, 62–73. doi:10.1016/j.canlet.2016.03.022

Zhao, J., Jitkaew, S., Cai, Z., Choksi, S., Li, Q., Luo, J., et al. (2012). Mixed lineage kinase domain-like is a key receptor interacting protein 3 downstream component of TNF-induced necrosis. *Proc. Natl. Acad. Sci. U. S. A.* 109, 5322–5327. doi:10.1073/pnas.1200012109

Zheng, Y., Kong, F., Liu, S., Liu, X., Pei, D., Miao, X., et al. (2021). Membrane protein-chimeric liposome-mediated delivery of triptolide for targeted hepatocellular carcinoma therapy. *Drug Deliv.* 28, 2033–2043. doi:10.1080/10717544.2021.1983072

Zhou, M., Wang, H., Zeng, X., Yin, P., Zhu, J., Chen, W., et al. (2019). Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: A systematic analysis for the global burden of disease study 2017. *Lancet* 394, 1145–1158. doi:10.1016/S0140-6736(19)30427-1

Zhou, P., Lu, Y., Zhang, Y., and Wang, L. (2021). Construction of an immune-related six-lncRNA signature to predict the outcomes, immune cell infiltration, and immunotherapy response in patients with hepatocellular carcinoma. *Front. Oncol.* 11, 661758. doi:10.3389/fonc.2021.661758

Zhou, X. D., Tang, Z. Y., Yang, B. H., Lin, Z. Y., Ma, Z. C., Ye, S. L., et al. (2001). Experience of 1000 patients who underwent hepatectomy for small hepatocellular carcinoma. *Cancer* 91, 1479–1486. doi:10.1002/1097-0142(20010415)91:8<1479::aid-cncr1155>3.0.co;2-0

Zhu, A. X., Kang, Y. K., Yen, C. J., Finn, R. S., Galle, P. R., Llovet, J. M., et al. (2019). Ramucirumab after sorafenib in patients with advanced hepatocellular carcinoma and increased α-fetoprotein concentrations (REACH-2): A randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet. Oncol.* 20, 282–296. doi:10.1016/S1470-2045(18)30937-9

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 02 September 2022
DOI 10.3389/fgene.2022.943378

Check for updates

OPEN ACCESS

EDITED BY
Yutian Zou,
Sun Yat-sen University Cancer Center (SYSUCC), China

REVIEWED BY
Barani Kumar Rajendran,
Yale University, United States
Jianyun Nie,
Yunnan Cancer Hospital, China

*CORRESPONDENCE
Jiong Wu,
wujiong1122@vip.sina.com
Yayun Chi,
yychi@126.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 27 May 2022
ACCEPTED 10 August 2022
PUBLISHED 02 September 2022

CITATION
Wang X, Ye F, Xiong M, Xiu B, Chi W, Zhang Q, Xue J, Chen M, Zhang L, Wu J and Chi Y (2022), Cross-talk of four types of RNA modification proteins with adenosine reveals the landscape of multivariate prognostic patterns in breast cancer.
Front. Genet. 13:943378.
doi: 10.3389/fgene.2022.943378

COPYRIGHT
© 2022 Wang, Ye, Xiong, Xiu, Chi, Zhang, Xue, Chen, Zhang, Wu and Chi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Cross-talk of four types of RNA modification proteins with adenosine reveals the landscape of multivariate prognostic patterns in breast cancer

Xuliren Wang[1†], Fangdie Ye[2†], Min Xiong[1†], Bingqiu Xiu[1], Weiru Chi[1], Qi Zhang[1], Jingyan Xue[1], Ming Chen[1], Liyi Zhang[1], Jiong Wu[1,3]* and Yayun Chi[1]*

[1]Department of Breast Surgery, Key Laboratory of Breast Cancer in Shanghai, Fudan University Shanghai Cancer Center, Shanghai, China, [2]Department of Urology, Huashan Hospital, Fudan University, Shanghai, China, [3]Collaborative Innovation Center for Cancer Medicine, Shanghai, China

Background: Breast cancer (BC) is the most common malignant tumour, and its heterogeneity is one of its major characteristics. N6-methyladenosine (m6A), N1-methyladenosine (m1A), alternative polyadenylation (APA), and adenosine-to-inosine (A-to-I) RNA editing constitute the four most common adenosine-associated RNA modifications and represent the most typical and critical forms of epigenetic regulation contributing to the immunoinflammatory response, tumorigenesis and tumour heterogeneity. However, the cross-talk and potential combined profiles of these RNA-modified proteins (RMPs) in multivariate prognostic patterns of BC remain unknown.

Methods: A total of 48 published RMPs were analysed and found to display significant expression alterations and genomic mutation rates between tumour and normal tissues in the TCGA-BRCA cohort. Data from 4188 BC patients with clinical outcomes were downloaded from the Gene Expression Omnibus (GEO), the Cancer Genome Atlas (TCGA), and the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC), normalized and merged into one cohort. The prognostic value and interconnections of these RMPs were also studied. The four prognosis-related genes (PRGs) with the greatest prognostic value were then selected to construct diverse RMP-associated prognostic models through univariate Cox (uniCox) regression analysis, differential expression analysis, Least absolute shrinkage and selection

Abbreviations: ADAR, double-stranded RNA-specific adenosine deaminase; APA, Alternative polyadenylation; A-to-I, Adenosine-to-inosine; BC, Breast invasive carcinoma; CNVs, Copy number variations; DEG, Differentially expressed genes; GEO, Gene Expression Omnibus; OS, Overall survival; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; LASSO, Least absolute shrinkage and selection operator; M1A, N1-methyladenosine; M6A, N6-methyladenosine; METABRIC, Molecular Taxonomy of Breast Cancer International Consortium; MultiCox, Multivariate Cox regression; PCA, Principal component analysis; SsGSEA, Single sample gene set enrichment analysis; TCGA, The Cancer Genome Atlas; TMB, Tumour mutation burden; TME, The tumour microenvironment; UniCox, Univariate Cox regression.

operator (LASSO) regression and multivariate Cox (multiCox) regression. Alterations in biological functional pathways, genomic mutations, immune infiltrations, RNAss scores and drug sensitivities among different models, as well as their prognostic value, were then explored.

Results: Utilizing a large number of samples and a comprehensive set of genes contributing to adenosine-associated RNA modification, our study revealed the joint potential bio-functions and underlying features of these diverse RMPs and provided effective models (PRG clusters, gene clusters and the risk model) for predicting the clinical outcomes of BC. The individuals with higher risk scores showed poor prognoses, cell cycle function enrichment, upregulation of stemness scores, higher tumour mutation burdens (TMBs), immune activation and specific drug resistance. This work highlights the significance of comprehensively examining post-transcriptional RNA modification genes.

Conclusion: Here, we designed and verified an advanced forecasting model to reveal the underlying links between BC and RMPs and precisely predict the clinical outcomes of multivariate prognostic patterns for individuals.

KEYWORDS

RNA-modifying proteins, breast cancer, risk score, mutation burden, stemness score, immune infiltration, drug sensitivity, prognosis

Background

In females, Breast cancer (BC) has overtaken lung cancer as the most commonly diagnosed cancer and is showing continuous acceleration, but with stagnated research progress (Sung et al., 2021; Siegel et al., 2022). Epigenetic changes, defined as stable alterations in transcription or translation without potential modifications in the genetic sequence, play a crucial role in both physiological and pathological processes (Taby and Issa, 2010). In the recent past, an accumulating number of investigations have revealed that RNA modification is an epigenetic regulatory mechanism of the expression of tumorigenesis-related genes and the inflammatory response, and targeting RNA modification enzymes represents a promising anticancer therapy (Frye et al., 2018; Yankova et al., 2021; Qiu et al., 2022). In BC, RMPs have been verified to function in tumour progression and metastasis (Chang et al., 2020).

More than 170 types of RNA modifications have been detected in nature, are widespread among all nucleotides, including A, U, C and G (Roundtree et al., 2017), and can be divided into three specific modification groups: "writers", "erasers" and "readers" (Barbieri and Kouzarides, 2020). Among them, adenosine most commonly shows alteration, and diverse modifications with adenosine may compensate for each other and form a competitive link (Xiang et al., 2018). Therefore, to explore the underlying mechanism and links, we focused on adenosine-related RNA modifications, including N6-methyladenosine (m6A), N1-methyladenosine (m1A), alternative polyadenylation (APA), and adenosine-to-inosine (A-to-I) RNA editing.

We identified genes with these four types of RNA modification from published articles. m6A is the most typical epigenetic RNA modification type in the eukaryotic transcriptome, affecting RNA metabolism in almost every process, and is catalysed by the m6A "writers", namely, METTL3/14/15/16, CBLL1, ZC3H13, RBM15/15B, KIAA1429 and WTAP; m6A "erasers", namely, FTO and ALKBH5; and m6A "readers", namely, YTHDF1/2/3, YTHDC1/2, ELAVL1, FMR1, HNRNPA2B1, HNRNPC, IGF2BP1/2/3, LRPPRC, RBMX and EIF3A (Meyer et al., 2015; Meyer and Jaffrey, 2017; Choe et al., 2018; Zaccara et al., 2019; Jiang et al., 2021; Xu et al., 2021; Ye et al., 2021).

m1A is widely present at the internal sites of mRNA as a form of posttranscriptional modification. m1A "writers" include TRMT61 A/B, TRMT10C and TRMT6; m1A "erasers" include ALKBH1/3 (Liu et al., 2016; Esteve-Puig et al., 2021).

APA creates specific new 3' ends on mRNAs and other RNA polymerase II transcripts, which is widely present in all eukaryotic species and is considered to be a mechanism for the creation of protein isomers. APA "writers" include CPSF1/2/3/4, CSTF1/2/3, CFI, PCF11, CLP1, NUDT21, and PABPN1 (Elkon et al., 2013; Tian and Manley, 2017).

RNA editing induces non-synonymous substitutions in protein-coding sequences. A-to-I editing by the double-stranded RNA-specific adenosine deaminase (ADAR) enzyme is the most common type. A-to-I "writers" include ADAR, ADARB1, and ADARB2 (Eisenberg and Levanon, 2018; Ishizuka et al., 2019; Chen et al., 2021).

To fully understand the behaviour of these adenosine-related RMPs, we obtained their mutation and expression profiles in BC

and established a network to investigate the mechanisms underlying neoplasm progression.

Studies have established a direct link between m6A and dynamic chromatin modification and identified underlying mechanisms for collaborative transcriptional interactions between RNA modification and histone modification (Li et al., 2020).

The underlying links between m6A and tumour immune activation has already been described. Loss of the m6A "writer" protein METTL3 disrupts T-cell homeostasis and differentiation (Li et al., 2017). The loss of FTO overcomes hypomethylation-induced immune escape, sensitizes leukaemia cells to T-cell cytotoxicity, and plays a key role in cancer stem cell self-renewal (Su et al., 2020). Deletion of YTHDF1 enhances the $CD8^+$ T-cell antitumor response and elevates the benefit of PD-L1 checkpoint blockade (Han et al., 2019). Since interferons are critical for inhibiting infectious and malignant diseases, ADAR serves as a target for cancer immunotherapy (Chung et al., 2018; Herbert, 2019; Liu et al., 2019). Accordingly, how the regulatory network and underlying links of RMPs influence tumour immunoregulation should be further explored.

During tumour development, cells that dedifferentiate and exhibit a stem cell-like phenotype have a higher degree of malignancy. Through maintaining FOXM1 expression and cell proliferation, ALKBH5 sustains tumorigenicity of stem-like cells in glioblastoma (Zhang et al., 2017). This suggests the potential role of RMPs in tumour cell proliferation and dedifferentiation.

Our study included 4188 BC patients from the Gene Expression Omnibus (GEO), the Cancer Genome Atlas (TCGA), and the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) databases in total. We revealed the genes with the most distinctive predictive value for RMPs, explored the associated genes, and constructed a risk model which can be utilized to assess the risk score of individuals. We discovered that diverse patterns of adenosine-associated RNA modification were linked not only to the infiltration of immune cells but also to the cell cycle, RNAss score, drug resistance and, most significantly, patients' clinical outcome.

Materials and methods

Data acquisition and processing

From the GEO database (https://www.ncbi.nlm.nih.gov/geo/), the METABRIC database (http://www.cbioportal.org/) and TCGA database (https://portal.gdc.cancer.gov/), we obtained the transcriptome data and overall survival (OS) data of BC patients retrospectively. 4188 BC patients with a follow-up time >30 days were selected for our analysis, including samples from seven GEO datasets (GSE131769, GSE162228, GSE20685, GSE20713, GSE24450, GSE42568, and GSE48391, n = 1,193), the METABRIC database (n = 1,904) and TCGA database (n = 1,091). All data were normalized with log2 transformation of fragments per kilobase of exon per million mapped fragments (FPKM) values and subsequently merged into one dataset with the "ComBat" algorithm of the SVA Package. Half of the 4188 BC patients were randomly assigned to a training cohort (n = 2,094), the remaining cases were defined as the testing cohort (n = 2,094), and the entire dataset was used to generate internal validation cohorts. GSE3494 (n = 234) from the GEO dataset was chosen as an external validation cohort. The RNA expression profiles of BC tissues (n = 1,109) and normal breast tissues (n = 113), somatic mutation data, somatic copy number variation (CNV) data and stemness score data were all obtained from TCGA database. The clinicopathological information, including age, tumour side, therapy, histology subtype, lymph node state, RFS status and cause of death, was gained from the METABRIC database.

Unsupervised clustering and gene set variation analysis of prognosis-related gene clusters

The correlation between RMPs and OS was analysed by univariate Cox regression (uniCox), and survival status was displayed with a Kaplan–Meier (K-M) curve using the "survminer" and "survival" R packages with an optimal cut-off value. RMPs with $p < 0.001$ were considered prognostic, and we named them prognosis-related genes (PRGs), which included YTHDF1, EIF3A, PCF11 and CBLL1. An unsupervised clustering algorithm was then applied to the 4188 BC patients based on the expression levels of 4 PRGs to build PRG clusters. The R package "conensusClusterPlus" was utilized with 1,000 repetitions during the above steps. Principal component analysis (PCA) was performed to visualize the independence of each cluster.

Selection of differentially expressed genes with prognostic value and construction of gene clusters

To identify PRG-related genes that were differentially expressed among subgroups, we used the empirical Bayesian function of the R package "limma" and selected 23 differentially expressed genes (DEGs) among three subgroups. Then, uniCox regression analysis was used and screened out 19 genes from DEGs for further analysis ($p <$ 0.05). We named these 19 genes prognosis-related differentially expressed genes (PRDEGs), which were both

differentially expressed and prognostic among diverse subgroups. The gene clusters were constructed based on the expression levels of 19 PRDEGs by the unsupervised clustering algorithm mentioned above.

Construction of the risk model

As mentioned above, we randomly divided the 4188 BC patients into a training group and a testing group. Least absolute shrinkage and selection operator (LASSO) analysis and multivariate Cox regression (multiCox) analysis implemented by the "glmnet" R package were applied to the patients in the training group to construct the risk model. All BC patients were then divided into two groups (low-risk group and high-risk group) according to the median value of the risk score. OS was compared between the high-risk and low-risk groups with K-M analysis. Receiver operating characteristic (ROC) analysis was performed using the "timeROC" R package to estimate the forecasting capability.

Gene set enrichment analysis of diverse risk groups

First, the R package "limma" was applied between the high-risk and low-risk groups, and 45 upregulated and 69 downregulated DEGs in the high-risk group were subsequently selected with criteria of $|logFC| > 1$ and adjusted p value < 0.05. To reveal diverse patterns of RMPs through biological processes, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed using the R package "clusterProfiler" (Yu et al., 2012).

Immune infiltration and tumour microenvironment analysis

The single-sample gene set enrichment analysis (ssGSEA) method of in the R package "GSVA" was used to calculate the infiltration degree according to the expression levels of 28 published immune cell gene sets (Bindea et al., 2013; Hänzelmann et al., 2013). Estimation of STromal and Immune cells in Malignant Tumour tissues using Expression data (ESTIMATE) was applied to determine the components of stromal cells and immune cells according to the gene expression characteristics of tumour samples (Yoshihara et al., 2013). CIBERSORT was used to infer cell composition based on the expression profiles. This deconvolution algorithm was used to calculate the relative proportions of 22 immune cells in each patient with BC (Newman et al., 2015; Becht et al., 2016).

Estimation of tumour mutation burden, stemness correlation and drug sensitivity

The "MutSigCV" algorithm was applied to screen 20 oncogenes with higher mutation frequencies than the background frequency. The R package "maftools" was applied to display the mutation landscape of the top altered oncogenes in the TCGA-BRCA cohort by waterfall plots (Mayakonda et al., 2018). The R package "pRRophetic" was used to predict the sensitivity of diverse risk groups by calculating the semi-inhibitory concentration (IC50) of commonly used drugs (Geeleher et al., 2014). In addition, we utilized the CellMiner database (https://discover.nci.nih.gov/cellminer/home.do) to analyse the drug sensitivity relevance between model-constructed genes and common antineoplastic drugs (Shankavaram et al., 2009). Spearman correlation analysis was applied to visualize the correlation between the risk score and the RNA-based stemness scores (RNAss).

Clinicopathological stratification of the risk score

BC patients in the METABRIC cohort were assigned to subclasses based on the following diverse characteristics. According to the median risk score, cases in each clinical subgroup were assigned to the low-risk or high-risk group. Survival curves of the high-risk group and low-risk group in the subgroups were compared using the log-rank test and K-M analysis.

Statistical analysis

All statistical analyses were performed using R software and its packages (version 4.1.2). Bilateral $p < 0.05$ was considered significant.

Results

The Landscape of RMPs with adenosine in BC

An overview of our workflow is outlined in Figure 1A. To obtain a fully picture of the expression and mutation patterns of RMPs associated with adenosine, we summarized 48 reported RMPs associated with adenosine modification, including 10 m6A "writers", 2 m6A "erasers", 15 m6A "readers", 4 m1A "writers", 2 m1A "erasers", 12 APA "writers" and 3 A-to-I "writers", and listed them in Supplementary Table S1.

Non-silent somatic mutations were detected among RMPs to identify whether they were associated with genetic alterations. Of

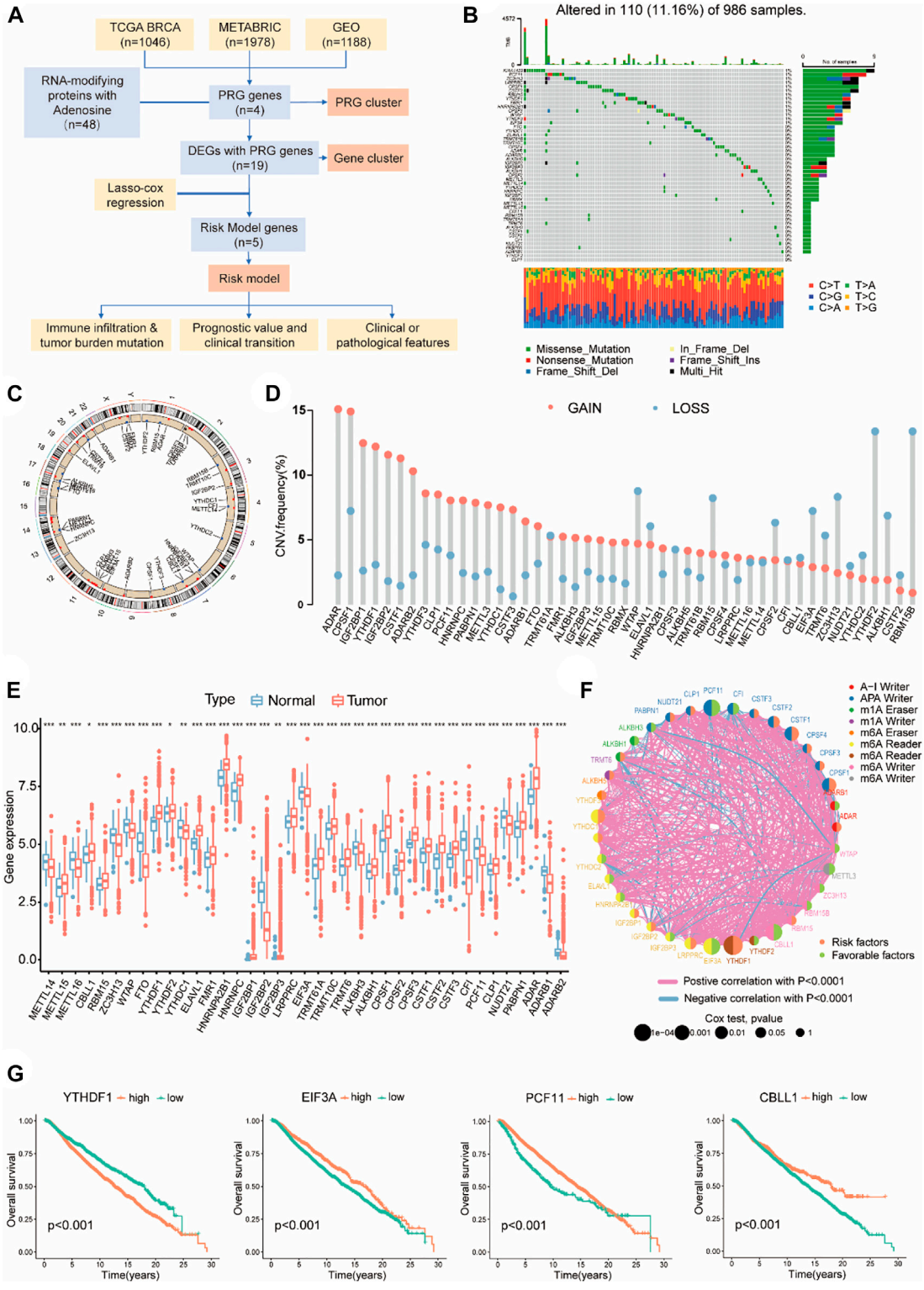


FIGURE 1
The Landscape of RMPs with adenosine in BC. **(A)** Workflow of the study. **(B)** 110 of 986 samples with breast cancer experienced genetic mutations of 48 RMPs, with a frequency of 11.16%. Each cohort represents an individual sample. The upper bar graph represents TMB; the number on the right side shows the mutation frequency in each RMP. The right bar graph shows the proportion of each mutation type for each BMP. The lower bar graph shows the fraction of transitions in each individual. **(C)** A circular diagram presenting the positions of CNV gains or losses of RMPs on 23 chromosomes. Red indicates CNV gained, and blue indicates CNV lost. **(D)** Barplots showing quantified CNV gains and losses of each RMP. **(E)**
(*Continued*)

FIGURE 1
Expression levels of RMPs compared between BC and normal samples presented with barplots. Red represents tumours, and blue represents normal tissues. The asterisks represented the statistical *p* value (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). **(F)** The intercorrelation and survival connections of RMPs. The colour of the right half of the circle represents its survival association, the colour of the left half represents the type of RMP, and the size of the circle represents the uniCox *p* value. **(G)** K-M curve showing the OS and survival probabilities of the four PGRs in the cohort of 4188 BC patients with the best cut-off value.

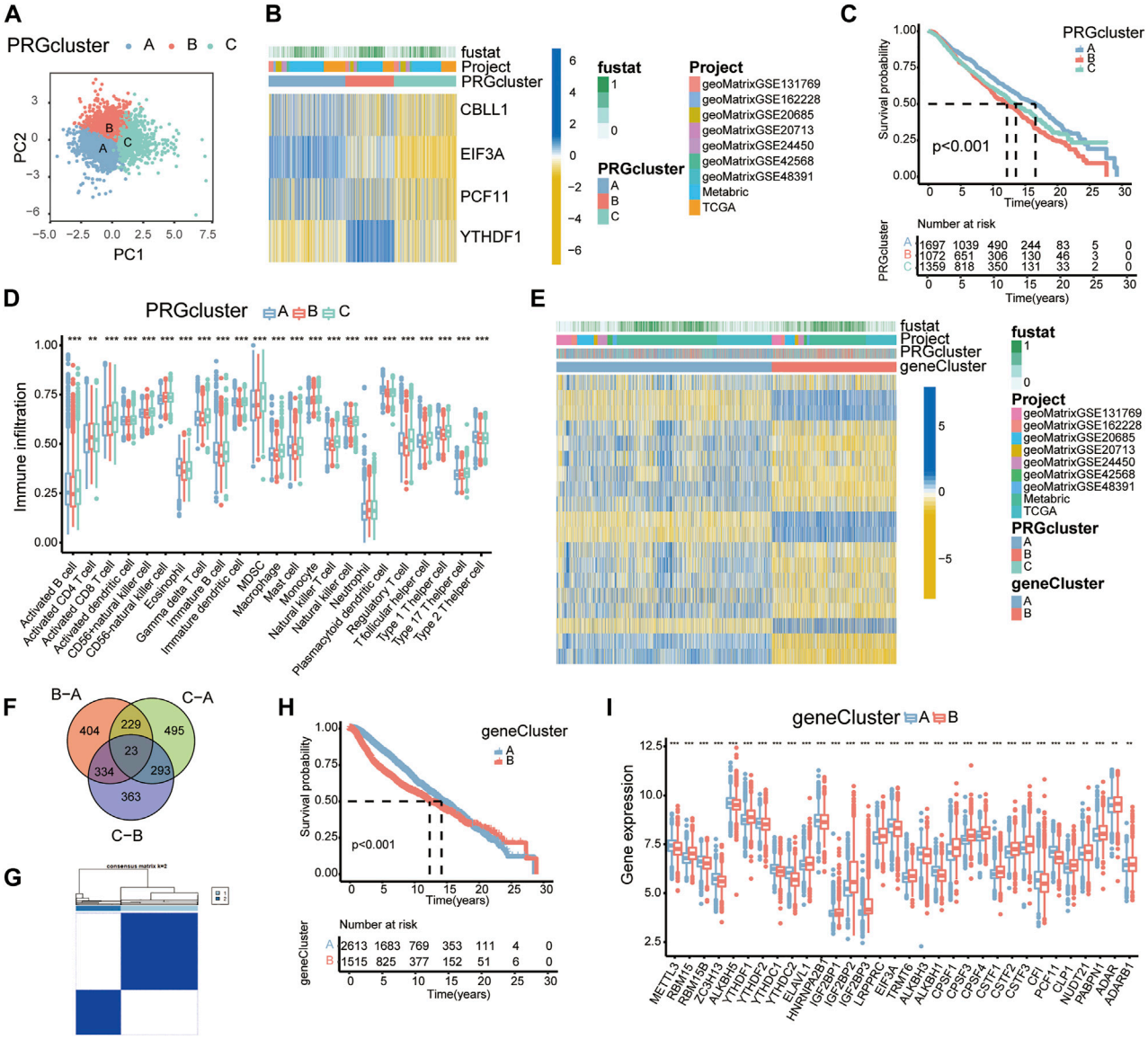


FIGURE 2
PRG clusters and gene clusters associated with four PRGs. **(A)** PCA of the PRG clusters. **(B)** Heatmap showing the expression profiles of four PRGs in the PRG clusters. **(C)** K-M survival curves of PRG clusters based on 4188 BC patients. Blue/red/green represent PRG clusters A/B/C. **(D)** Boxplots showing immune cell infiltration in each PRG cluster by ssGSEA. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. **(E)** Heatmap showing the expression profiles of 19 PRDEGs among gene clusters. **(F)** A Venn diagram showing DEGs. **(G)** Consensus matrix displaying two major gene clusters based on PRDEGs. **(H)** K-M survival curves of gene clusters (n = 4,188). **(I)** Expression profiles of RMPs in variant gene clusters.

the 986 BC samples in TCGA-BRCA, 110 (11.16%) samples had mutations of RMPs, most of which were missense mutations. As depicted in Figure 1B, KIAA1429 exhibited the highest mutation frequency in BC samples, followed by PCF11 and ZC3H13. On the contrary, YTHDF2 and CLP1 did not exhibit any mutations.

Then, we examined the somatic CNVs of these RMPs. We found extensive CNV gains in ADAR, CPSF1, IGF2BP1 and YTHDF1. Loss of copy numbers was observed in RBM15B and YTHDF2 (Figure 1D). A circular diagram was created to illustrate the locations of these RMPs on the 23 chromosomes (Figure 1C).

BC samples were paired with normal breast tissue samples to compare mRNA expression levels of RMPs. 39 RMPs were found in the TCGA-BRCA cohort, and most of them were abundant in the tumour samples (Figure 1E).

To demonstrate the links between these RMPs and the outcomes of BC patients, the clinical and transcriptome data of 4188 BC patients were enrolled in this study to reveal the relationships between RMPs and tumorigenesis. The results suggested that there were strong correlations between different types of RMPs and that positive correlations are more common than negative ones. Notably, YTHDF1, CBLL1, PCF11, and EIF3A were most related to survival, with a uniCox p value < 0.001 (Figure 1F). The results of uniCox for RMPs are displayed (Supplementary Table S2). Patients with higher expression level of YTHDF1 had worse prognoses, and those with higher expression of CBLL, PCF11, and EIF3A had better clinical outcomes (Figure 1G). These four genes were defined as PRGs, and other RMPs associated with survival were also noted (Supplementary Figure S1).

PRG clusters and gene clusters associated with four RMPs

Utilizing the consensus clustering analysis, 4188 BC patients were then assigned into three subgroups. A total of three PRG clusters were identified, with 1,718 patients in PRG cluster A, 1,084 patients in PRG cluster B, and 1,386 patients in PRG cluster C. PCA confirmed an intergroup distribution among the three subgroups (Figure 2A).

The heatmap was depicted to reveal expression levels of four PRGs in these three PRG clusters. CBLL1, EIF3A, and PCF11 were markedly overexpressed in cluster A patients, while YTHDF1 was expressed at higher levels in cluster B patients (Figure 2B).

The survival analysis of PRG clusters suggested that patients in cluster A had a much more prominent survival advantage than those in cluster B, and the advantage for patients in cluster C was somewhere in between, which was consistent with the expression levels and the prognostic trends of the four PRGs (Figure 2C).

To understand biological processes in diverse RNA modification patterns, the "GSVA" R package was utilized to conduct gene set variation analysis (GSVA). PRG cluster B was markedly enriched in metabolic pathways such as ribosome, fatty acid metabolism, valine leucine and isoleucine degradation, and propanoate metabolism compared to PRG cluster A (Supplementary Figure S2).

Many studies have mentioned a potential link between infiltrating immune cells and RNA modification (Barbieri and Kouzarides, 2020). To investigate the functional role of RMPs in immune infiltration, ssGSEA was applied to the PRG clusters and revealed a strong connection between PRG clusters and immune cells (Figure 2D).

The identified PRG clusters could effectively distinguish the clinical outcome of BC patients; however, the PRG-related genes, therapeutic effect and underlying reasons were still unclear. To identify the genes potentially related to the PRG clusters, overlapping DEGs among the 3 PRG clusters were then selected. We obtained 23 PRG-related DEGs and displayed them with a Venn diagram (Figure 2F).

UniCox analysis was performed and identified 19 prognosis-related DEGs. According to the expression levels of these 19 PRDEGs, we divided the patients into two subgroups: gene cluster A and gene cluster B (Figure 2G).

In the entire transcriptome, the heatmap showed significant inherent differences between the gene clusters (Figure 2E). K-M curves showed significant alterations in survival outcomes between the two gene subtypes. Gene cluster A presented a clear survival advantage, while gene cluster B had a higher risk of death (Figure 2H). The expression levels of RMPs in gene clusters were examined, which revealed substantial discrepancies in RMPs and suggested an underlying correlation. Notably, YTHDF1 showed a higher expression level in gene cluster B, while PCF11 and EIF3A were more highly expressed in gene cluster A (Figure 2I).

Construction of the risk model

To accurately forecast the survival status of individuals, we designed a risk model. The BC patients in the training group were used for the following model construction procedure.

Nineteen PRDEGs were engaged in the iterative LASSO algorithm (Figures 7A,B). Next, multiCox analysis was applied to construct the risk model. Ultimately, risk scores were calculated by multiplying the expression values of five chosen genes with their regression coefficients. The formula was as follows: risk score = (expression level of UBE2C × 0.096)—(expression level of CX3CR1 × 0.077)—(expression level of IFT74 × 0.141)—(expression level of FABP4 × 0.033) + (expression level of CALML5 × 0.023). We then evaluated the risk score of each patient in the training and testing groups, which is listed in Supplementary Table S3. The hazard ratio from uniCox and the coefficients of the 5 model-constructed genes are listed (Figure 3D). We compared the risk scores in diverse PRG clusters and gene clusters. It was revealed that PRG cluster B and gene cluster B, which had poor survival, had higher risk scores (Figures 3A,B).

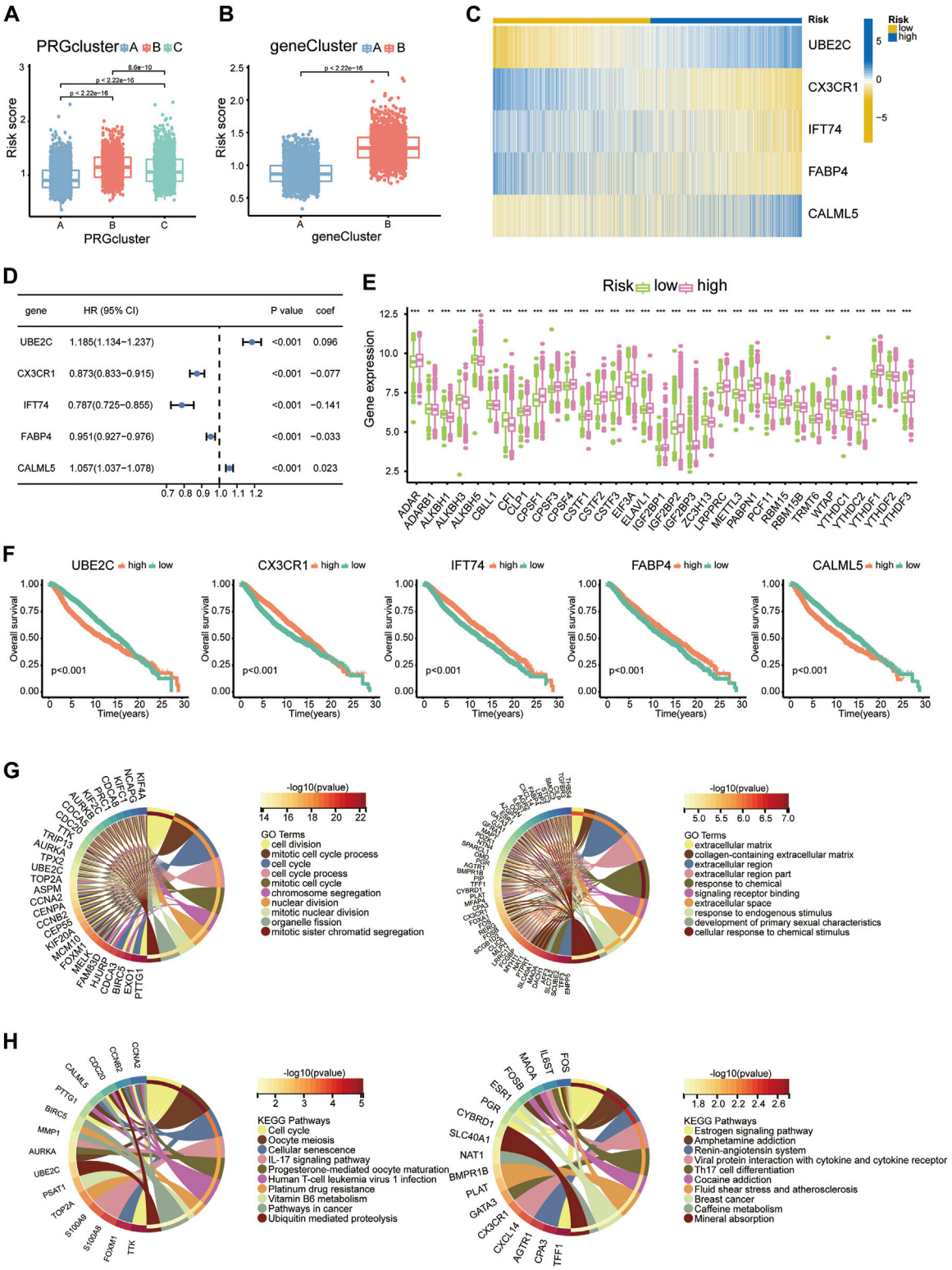


FIGURE 3
Construction of the risk model. **(A)** Risk scores of diverse PRG clusters. **(B)** Risk scores of diverse gene clusters. **(C)** Heatmap showing the expression profiles of 5 model-constructed genes in the risk groups. **(D)** Table listing the coefficient and HR of each model-constructed gene. **(E)** RMP expression levels in risk groups presented with boxplots. **(F)** K-M curves of 5 model-constructed genes. **(G)** GO functional enrichment analysis of upregulated DEGs (left) and downregulated DEGs (right) in the high-risk group. **(H)** KEGG analysis results of upregulated DEGs (left) and downregulated DEGs (right) in the high-risk group.

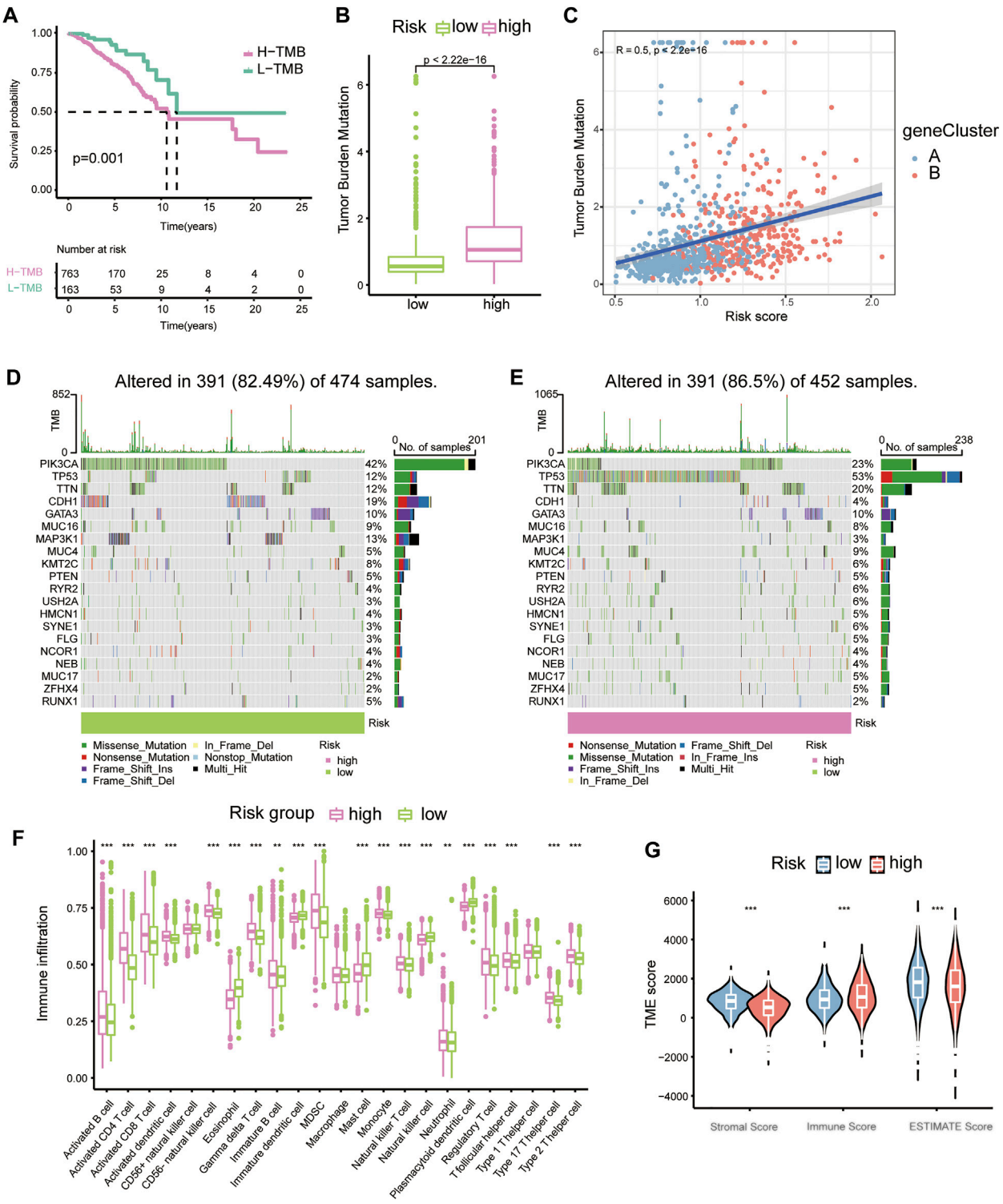


FIGURE 4
The relationship between tumour mutation burden, immune infiltration and risk score. **(A)** K-M plots of high- and low-TMB groups of the TCGA-BRCA cohort. **(B)** Boxplots showing the relative connection of the risk score and TMB. **(C)** Correlation scatter plot presenting the relationships among gene clusters, risk scores, and TMB. **(D)** Waterfall plots showing TMB for the low-risk group. **(E)** Waterfall plots showing TMB for the high-risk group. **(F)** Immune cell proportions in diverse risk groups shown with boxplots. **(G)** Violin plots for stromal score, immune score and ESTIMATE score in different risk groups.

We separate the 4188 BC patients into high- and low-risk groups based on the median cut-off value of the developed risk score. The heatmap showed differential expression of model-constructing genes between the high-risk and low-risk groups among the 4188 BC patients, where UBE2C and CALML5 were distinctly upregulated in the high-risk group, and CX3CR1, IFT74, and FABP4 showed the opposite trend (Figure 3C). The training group and testing group showed the same condition (Figure 7C).

The K–M plot of the five model-constructing genes showed that with higher expression levels of UBE2C and CALML5, patients had worse survival probabilities, while with higher expression of the other three genes, patients had better outcomes (Figure 3F).

The links between the risk groups and expression profiles of RMPs is presented. We observed that YTHDF1 presented a significant expression increase in the high-risk group, while CBLL1, EIF3A, and PCF11 performed downregulation. The expression profiles of ADAR, CLP1, CPSF1/3/4, CSTF1/2/3, ELAVL1, IGF2BP1/2/3, LRPPRC, PABPN1, RBM15, RBM15B, TRMT6, and YTHDF2/3 were markedly positively correlated with the patient's risk score and most of them were associated with a poor clinical outcome (Figure 3E).

To uncover the potential biological characteristics between the two groups, we searched for DEGs and carried out enrichment analyses. Analysis of DEGs enhanced in the high-risk group revealed enrichment of GO functions such as cell division, mitotic cell cycle process and cell cycle, which indicated that the potential mechanism underlying poor survival may be linked to cell proliferation (Figure 3G). In the KEGG analysis, we also observed proliferation pathways enriched among DEGs elevated in the group with a higher average risk score. A higher risk score may be also associated with poor sensitivity to platinum-based drug therapy, since IL-17 and Th17 cells are both related to inflammation-related tumour development. We observed that the IL-17 signalling pathway and Th17-cell differentiation pathway were enriched in the upregulated and downregulated groups, respectively, which suggested an indispensable role of the immune and inflammatory systems in tumour development with the adenosine-RNA-modification-derived risk model (Figure 3H).

The relationship between tumour mutation burden, immune infiltration and risk score

We then analysed the underlying links between the risk score and tumour mutation burden (TMB) with the TCGA-BRCA database. K–M plots depicted that with lower TMBs, BC patients exhibited a distinct survival advantage, while the others had worse clinical outcomes (Figure 4A). The TMB landscape among the top 20 most common mutations suggested that with higher risk scores, BC patients had a more extensive TMB (Figures 4D,E). The risk score was positively correlated with TMB (Figures 4B,C).

Through previous analysis, we found that Th17 and IL-17 may be associated with the potential bio-mechanism of risk models, and several RMPs were documented to be linked to immunotherapy resistance. To learn whether the risk score could predict immunotherapy response of BC patients, we analysed the immune cells infiltration profiles by ssGSEA. Most immune cells were upregulated in the high-risk groups (Figure 4F). Immune cell infiltration (immune score) was also evaluated by the ESTIMATE algorithm, and the immune score was obtained, which was consistent with the results of ssGSEA. The tumour microenvironment (TME) score of each case was evaluated, and the stromal score, immune score and ESTIMATE score were determined. The stromal score and ESTIMATE score were both decreased in the high-risk group (Figure 4G). By CIBERSORT, we analysed the correlations of immune cells with the risk score and risk model-building genes, it was depicted that CX3CR1 were strongly positive correlated with M2 macrophages but the trend was opposite to that of M1 macrophages and UBE2C showed the opposite (Figure 5A). There was a significant positive correlation between inhibitory immune cell Tregs and risk, but at the same time M1 macrophages were significantly upregulated, accompanied by downregulation of M2 macrophages (Figure 5B).

Prognostic value and clinical translation of the risk model

To better learn the correlation of risk scores with prognosis and drug therapy, the following analysis was performed. The K-M analysis showed that patients in low-risk group had evident advantages of survival, while high-risk patients had poor clinical outcomes (Figure 6A). The K-M analysis results for the training and testing groups are also displayed (Figure 7D). Furthermore, we applied the timeROC method to estimate the AUC values for predicting OS (Figure 7E). The relevance between clinical outcomes and risk scores of patients is also displayed (Figure 6B), and the results for the training and testing groups are also shown (Figure 7F). GSE3494 was chosen as an external cohort for feasibility verification, and there was an apparent discrepancy in clinical outcomes, with the high-risk group having worse outcomes (Figure 6F). All the above results indicate that the risk model is stable and can precisely predict the clinical outcome of patients.

To visualize the relationship between risk scores and diverse clusters, we displayed the relationships among PRG clusters, gene clusters, risk groups and the survival outcome of patients with a Sankey diagram (Figure 6D). Rapidly developed tumour cells can lose differentiation phenotypes and exhibit stem-cell-like characteristics (Batlle and Clevers, 2017). RNAss scores based on mRNA expression were utilized to measure the correlation between tumour stemness and the risk score (Figure 6C).

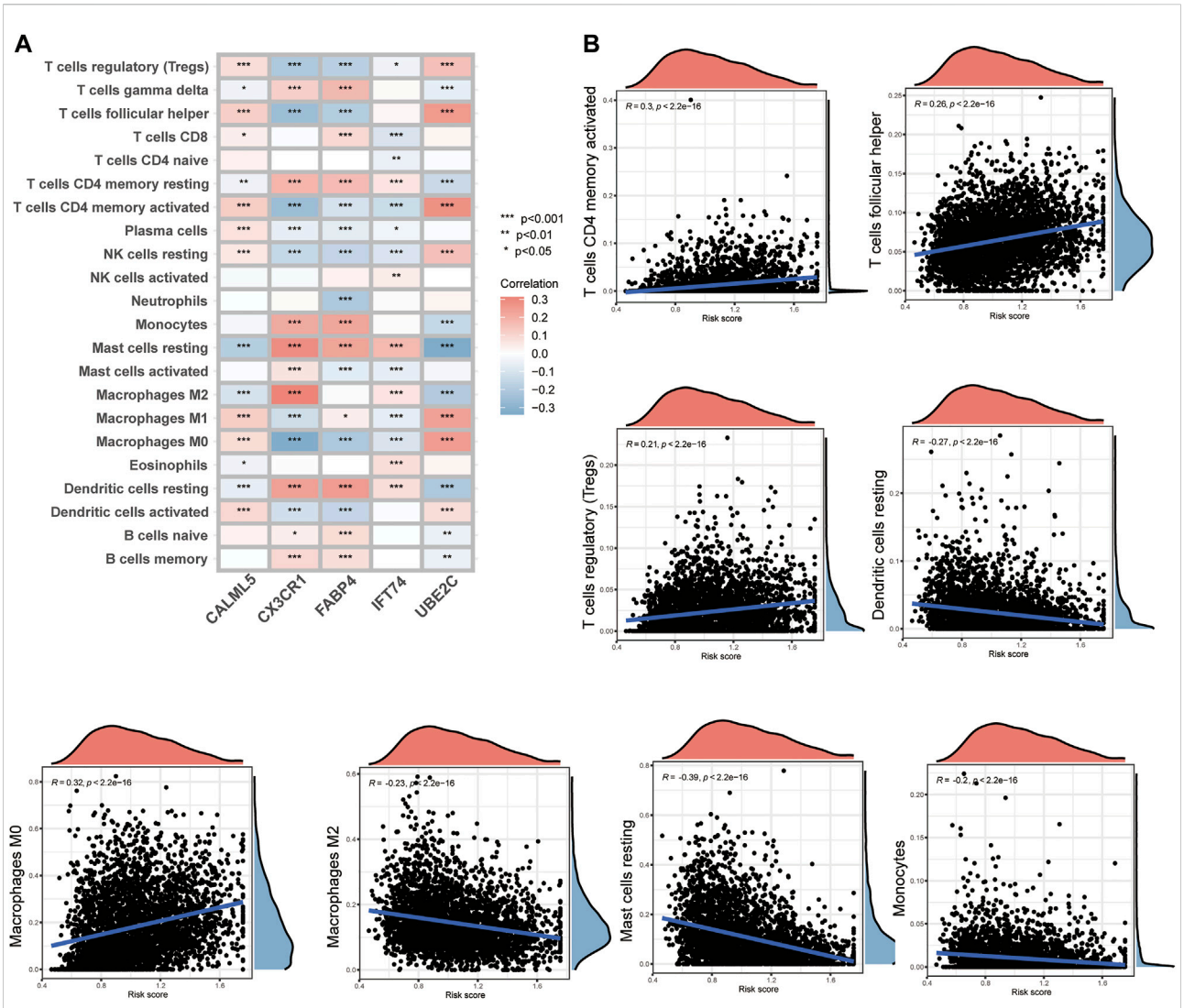


FIGURE 5
The immune cell infiltration connections associated with model-constructed genes and risk scores. **(A)** A heatmap demonstrating the relationship between model-constructing genes and immune cell infiltration. **(B)** Correlation scatter plots showing the relationship between risk scores and immune cell infiltration.

Next, to determine the effectiveness of the risk score for predicting drug treatment response in BC patients, we estimated the IC50 values of the most common drugs. We found that patients in the high-risk group might be more sensitive to M-phase cell cycle drugs, including docetaxel, paclitaxel and vinblastine, but resistant to the cell cycle-nonspecific drug doxorubicin, the S-phase-specific drugs cytarabine and gemcitabine and the G2-phase-specific drug bleomycin (Figure 6E).

According to diverse clinicopathological factors (age of diagnosis, tumour side, surgical type, radiotherapy, chemotherapy, hormonotherapy, histology type, lymph node state, RFS status, and reason for death), BC patients from the METABRIC database were divided into different cohorts. The risk model presented excellent prediction performance. In particular, there was a distinct difference in the survival outcomes of patients with recurrence, and no diversity was found in patients without recurrence. Similarly, the survival curve of patients who died of disease was changed, but there was no such alteration for patients who died for other reasons (Figure 8).

Discussion

Epigenetic transcriptomics focusing on RNA modification, an unexplored field, has been gradually explored with the

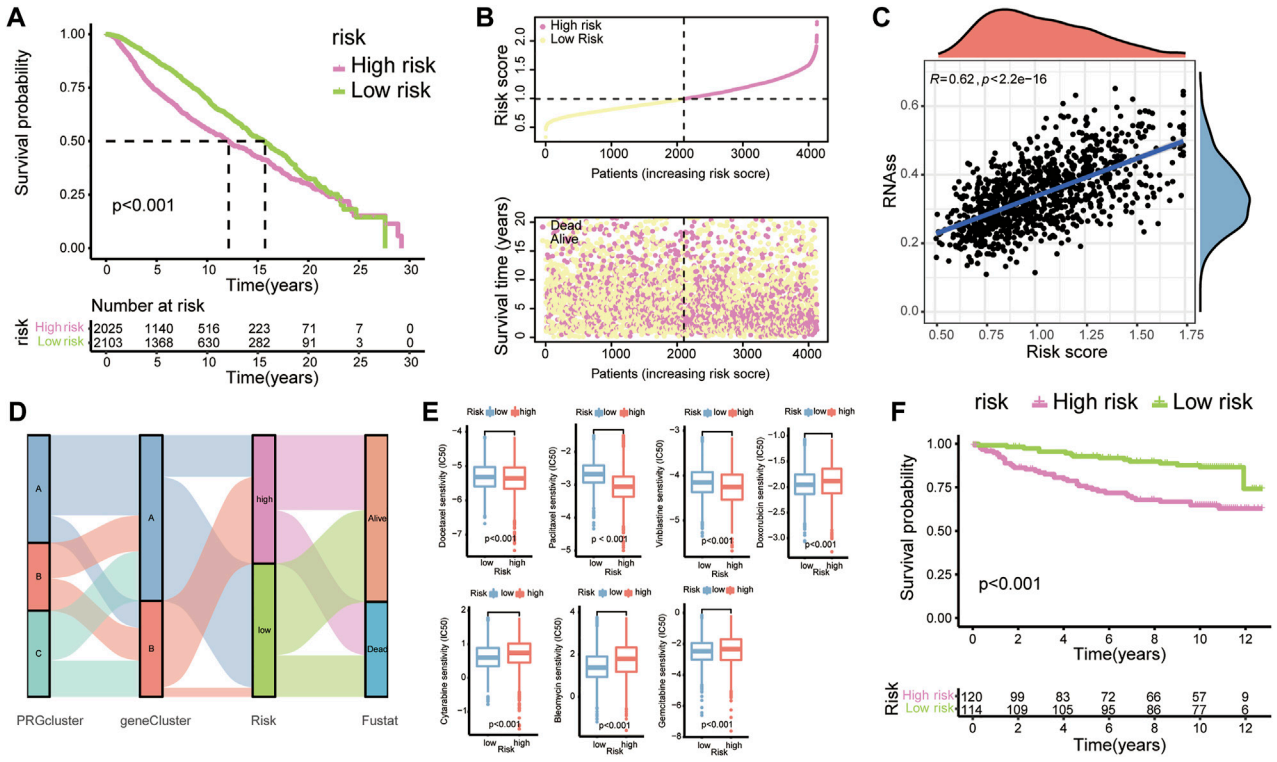


FIGURE 6
Prognostic value and clinical translation of the risk model. **(A)** K-M curves of risk groups in the merged BC cohort. **(B)** Correlation scatter plots showing the relationship between clinical outcomes and risk scores. **(C)** The relationship between tumour cell stemness and risk score. **(D)** A Sankey map demonstrating the relationship between clinical outcomes and diverse subgroups (PRG clusters, gene clusters, and risk groups). **(E)** Boxplots showing the IC50 values of common anti-tumour drugs. **(F)** K-M plots of BC patients in GSE3494.

continuous development of sequencing technology. The majority of studies focused on only one specific single form of RNA modification mode, and there is increasing evidence that RMPs play an indispensable role in tumorigenesis and interact with each other; nevertheless, the interrelationships of multiple forms of RMPs are still not completely understood (Zhang et al., 2016; Zhao et al., 2017). Here, we ultimately revealed a landscape of transcriptional and genetic alterations of adenosine-related RNA-regulatory enzymes of m6A, m1A, APA, as well as A-to-I and discussed their potential connections, expression patterns and prognostic values in BC.

To obtain a brief understanding, we depicted the connections among RMPs, and most of them were associated with tumour mutation burdens and patient outcomes and showed differential expression levels between BC tumours and normal tissues. The RMPs upregulated in tumours or showing a positive correlation with poor outcomes usually gained more CNVs, such as ADAR, CPSF1, IGF2BP1, and YTHDF1.

We then chose the four most prognostic genes, YTHDF1, EIF3A, PCF11, and CBLL1, according to the survival analysis results and named them PRGs. Three PRG clusters and two gene clusters based on PRGs and PRDEGs were then identified in 4188 BC patients and found to predict the patients' outcomes in diverse clusters. To more accurately anticipate the prognosis of individuals, a risk model was constructed to forecast prognostic risk scores of BC patients. The patients with higher risk were identified as having higher expression levels of UBE2C, and CALML5 and lower expression levels of CX3CR1, IFT74, and FABP4. UBE2C encodes a E2 ubiquitin-conjugating enzyme, which was found to be a prognostic factor in BC with poor survival in previous studies (Psyrri et al., 2012). Studies have shown that CALML5 ubiquitination is involved in the tumorigenesis of the BC (Debald et al., 2013).

We explored RMP expression patterns between risk groups and connected specific RMPs with BC outcomes. YTHDF1 was highly expressed in high-risk group patients, while EIF3A, PCF11, and CBLL1 showed higher expression levels in low-risk group patients. BC patients with higher risk scores had a worse prognosis. We also confirmed the prognostic value of the risk model by assessing outcomes in BC patients with different clinicopathological features.

To explore the underlying mechanisms, functional enrichment analysis was performed, suggesting that cell

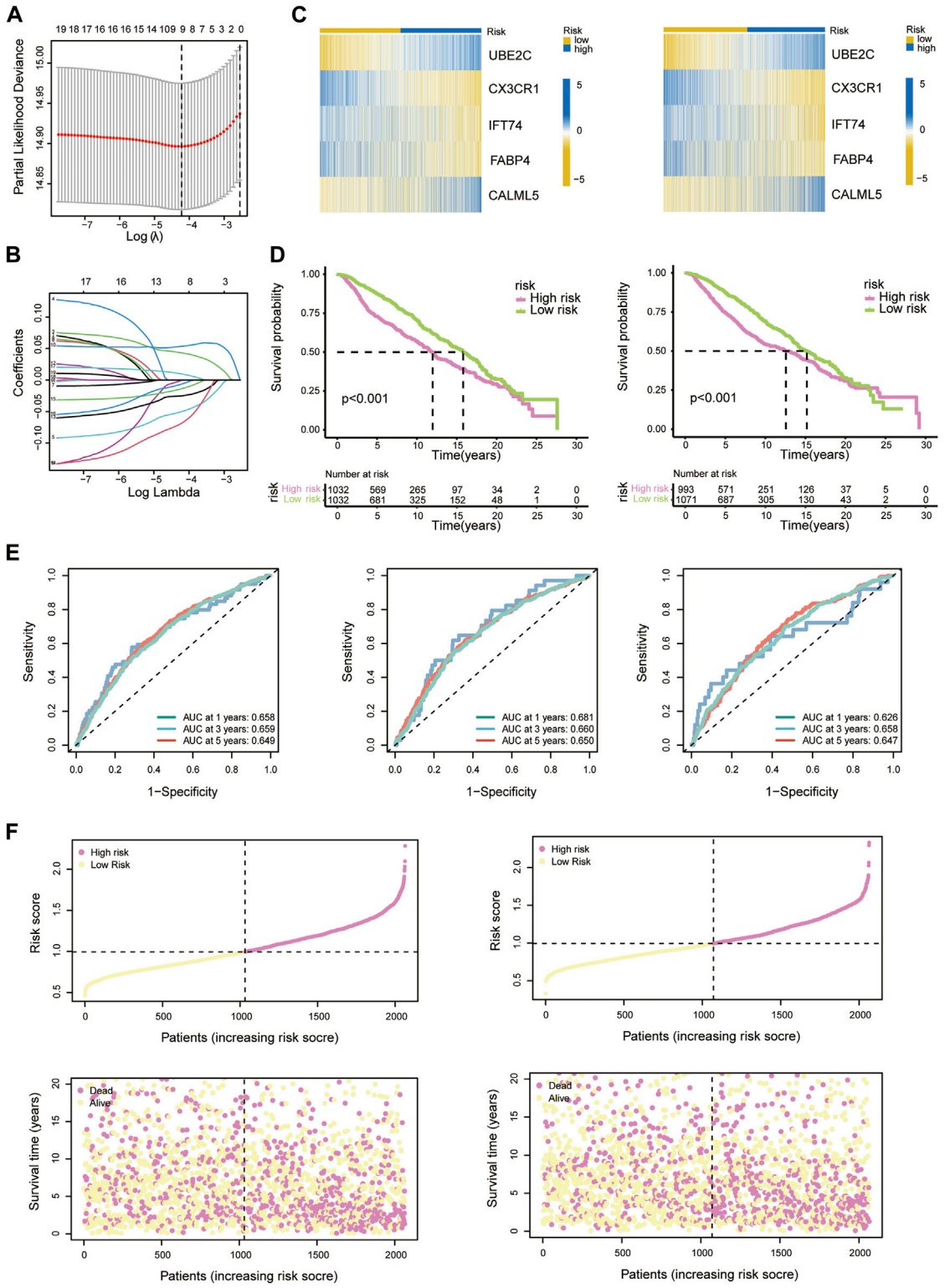


FIGURE 7
(A) LASSO regression analysis reducing variants. **(B)** Coefficients of model-constructed genes obtained by LASSO. **(C)** Heatmap of the expression profiles of model-constructed genes in the training group (left) and testing group (right). **(D)** K-M curves of patients in the training group (left) and testing group (right). **(E)** AUC time-dependent ROC curves for OS in all patients (left), training group patients (middle) and testing group patients (right). **(F)** Correlation scatter plots show the relationship between clinical outcomes and risk scores in the training group (left) and testing group (right).

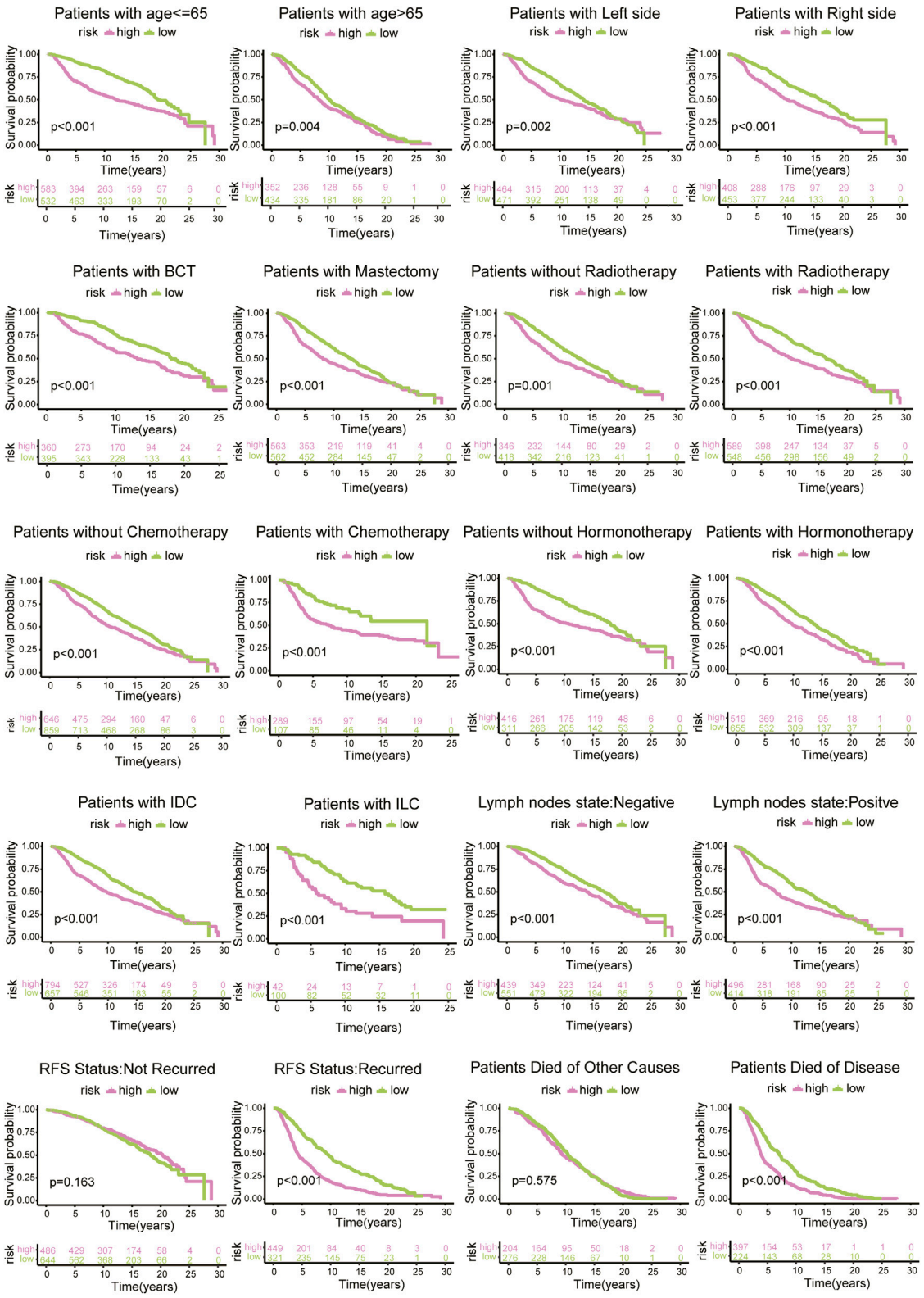


FIGURE 8
Survival analysis of risk scores based on diverse clinicopathological characteristics of 1130 BC patients from METABRIC database.

proliferation-related functions such as cell division, cell cycle, chromosome segregation and mitotic nuclear division were stronger in the high-risk group. We then visualized the potential therapeutic effects of diverse RMP-related risk patterns in BC. Similarly, the high-risk group was associated with resistance to the cell cycle-nonspecific drug doxorubicin and the S- or G2-phase cell cycle-specific drugs cytarabine, gemcitabine and bleomycin but with sensitivity to M-phase cell cycle-specific drugs. It has been reported that YTHDF1, which was upregulated in the high-risk group, promotes S-phase entry, DNA replication and DNA damage repair (Sun et al., 2022). EIF3A regulates the expression profiles of proteins that contribute to DNA repair, which in turn is involved in response to anti-cancer drugs (Yin et al., 2011). UBE2C can directly override the spindle assembly checkpoint inhibition of APC (Reddy et al., 2007; Meyer and Rape, 2011). These literature reports prove the reliability of our conjecture.

The abundance of TMBs was markedly difference between the two risk groups, and the high-risk group was associated with a higher TMB. The results of RNAss analysis indicated a close relationship between the risk score model and tumour cell stemness, which was reported previously (Du et al., 2021). The difference in the degree of immune cells infiltration is also indisputable. In patients with a higher risk score, we observed higher levels of infiltration in Tregs and MDSCs, which contribute to immune suppression (Wood et al., 2012). This suggested that patients in the high-risk group showed higher infiltration of immune suppressor cells. CX3CR1 was reported to be a marker of T cell differentiation, which indicated a predictive correlate of response to immune checkpoint inhibitor therapy, and CX3CR1 + inhibitory macrophages were negatively correlated with T-cell expansion (Bassez et al., 2021; Yamauchi et al., 2021). Since the exact mechanism is still unclear, the connections in immune cell infiltration should be further analysed.

This study is one of the few to combine four RNA modification types with breast cancer and predict patient outcome, which opens up a whole new way of model prediction and clinical evaluation. However, there are still limitations. While the mechanisms behind RNA modification are still not well understood, the association between RNA modifications and cancer diversity needs further exploration and further investigation. And we must admit genetic testing continues to be an invasive method that imposes a financial burden and risk on patients. Therefore, more economical and convenient detection methods need to be further explored.

Conclusion

A systematic and comprehensive landscape of four types of adenosine-related RMPs in BC was constructed, revealing expression profiles, tumour mutation burden, immune cell infiltration and connections with BC survival outcomes. We constructed a risk model and evaluated several aspects, including genomic mutation, immune relativity and therapeutic responses. The model presented effective performance in predicting an individual's clinical outcome in relation to BC. This work highlights the clinical importance of cross-talk between diverse types of RNA modification with a large number of cases and presented a predictive model which contributes to the improvement of prediction and treatment patterns. However, simpler and more economical methods of precise diagnosis still need to be discovered.

Data availability statement

The databases used and analysed during the current study are available from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) with the accession numbers GSE131769, GSE162228, GSE20685, GSE20713, GSE24450, GSE42568, GSE48391, and GSE3494; the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) database (http://www.cbioportal.org/) and The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/) with the accession number TCGA447 BRCA.

Author contributions

XW and FY performed the research and wrote the paper. MX collected and analyzed data and provided some critical ideas. BX, WC, QZ, and LZ performed some of the analysis and figures preparation. JX and MC provided useful suggestions in methodology. YC and JW provided assistance in the study. YC and XW initiated the study, organized, designed and revised the paper. All authors read and approved the final manuscript.

Funding

This work was supported by grants from the National Natural Science Foundation of China (81874115) and Shanghai Natural Science Foundation (20ZR1412200).

Acknowledgments

We sincerely appreciate all members who participated in data collection and 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

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/fgene.2022.943378/full#supplementary-material

References

Barbieri, I., and Kouzarides, T. (2020). Role of RNA modifications in cancer. *Nat. Rev. Cancer* 20, 303–322. doi:10.1038/s41568-020-0253-2

Bassez, A., Vos, H., Van Dyck, L., Floris, G., Arijs, I., Desmedt, C., et al. (2021). A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. *Nat. Med.* 27, 820–832. doi:10.1038/s41591-021-01323-8

Batlle, E., and Clevers, H. (2017). Cancer stem cells revisited. *Nat. Med.* 23, 1124–1134. doi:10.1038/nm.4409

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

Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A. C., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 39, 782–795. doi:10.1016/j.immuni.2013.10.003

Chang, G., Shi, L., Ye, Y., Shi, H., Zeng, L., Tiwary, S., et al. (2020). YTHDF3 induces the translation of m(6)a-enriched gene transcripts to promote breast cancer brain metastasis. *Cancer Cell* 38, 857–871. e7. doi:10.1016/j.ccell.2020.10.004

Chen, H., Yao, J., Bao, R., Dong, Y., Zhang, T., Du, Y., et al. (2021). Cross-talk of four types of RNA modification writers defines tumor microenvironment and pharmacogenomic landscape in colorectal cancer. *Mol. Cancer* 20, 29. doi:10.1186/s12943-021-01322-w

Choe, J., Lin, S., Zhang, W., Liu, Q., Wang, L., Ramirez-Moya, J., et al. (2018). mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature* 561, 556–560. doi:10.1038/s41586-018-0538-8

Chung, H., Calis, J. J. A., Wu, X., Sun, T., Yu, Y., Sarbanes, S. L., et al. (2018). Human ADAR1 prevents endogenous RNA from triggering translational shutdown. *Cell* 172, 811–824. e14. doi:10.1016/j.cell.2017.12.038

Debald, M., Schildberg, F. A., Linke, A., Walgenbach, K., Kuhn, W., Hartmann, G., et al. (2013). Specific expression of k63-linked ubiquitination of calmodulin-like protein 5 in breast cancer of premenopausal patients. *J. Cancer Res. Clin. Oncol.* 139, 2125–2132. doi:10.1007/s00432-013-1541-y

Du, J., Ji, H., Ma, S., Jin, J., Mi, S., Hou, K., et al. (2021). m6A regulator-mediated methylation modification patterns and characteristics of immunity and stemness in low-grade glioma. *Brief. Bioinform.* 22, bbab013. doi:10.1093/bib/bbab013

Eisenberg, E., and Levanon, E. Y. (2018). A-to-I RNA editing - immune protector and transcriptome diversifier. *Nat. Rev. Genet.* 19, 473–490. doi:10.1038/s41576-018-0006-1

Elkon, R., Ugalde, A. P., and Agami, R. (2013). Alternative cleavage and polyadenylation: Extent, regulation and function. *Nat. Rev. Genet.* 14, 496–506. doi:10.1038/nrg3482

Esteve-Puig, R., Climent, F., Piñeyro, D., Domingo-Domènech, E., Davalos, V., Encuentra, M., et al. (2021). Epigenetic loss of m1A RNA demethylase ALKBH3 in Hodgkin lymphoma targets collagen, conferring poor clinical outcome. *Blood* 137, 994–999. doi:10.1182/blood.2020005823

Frye, M., Harada, B. T., Behm, M., and He, C. (2018). RNA modifications modulate gene expression during development. *Sci. (New York, N.Y.)* 361, 1346–1349. doi:10.1126/science.aau1646

Geeleher, P., Cox, N., and Huang, R. S. (2014). pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PloS one* 9, e107468. doi:10.1371/journal.pone.0107468

Han, D., Liu, J., Chen, C., Dong, L., Liu, Y., Chang, R., et al. (2019). Anti-tumour immunity controlled through mRNA m(6)A methylation and YTHDF1 in dendritic cells. *Nature* 566, 270–274. doi:10.1038/s41586-019-0916-x

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

Herbert, A. (2019). ADAR and immune silencing in cancer. *Trends Cancer* 5, 272–282. doi:10.1016/j.trecan.2019.03.004

Ishizuka, J. J., Manguso, R. T., Cheruiyot, C. K., Bi, K., Panda, A., Iracheta-Vellve, A., et al. (2019). Loss of ADAR1 in tumours overcomes resistance to immune checkpoint blockade. *Nature* 565, 43–48. doi:10.1038/s41586-018-0768-9

Jiang, X., Liu, B., Nie, Z., Duan, L., Xiong, Q., Jin, Z., et al. (2021). The role of m6A modification in the biological functions and diseases. *Signal Transduct. Target. Ther.* 6, 74. doi:10.1038/s41392-020-00450-x

Li, H. B., Tong, J., Zhu, S., Batista, P. J., Duffy, E. E., Zhao, J., et al. (2017). m(6 A mRNA methylation controls T cell homeostasis by targeting the IL-7/STAT5/SOCS pathways. *Nature* 548, 338–342. doi:10.1038/nature23450

Li, Y., Xia, L., Tan, K., Ye, X., Zuo, Z., Li, M., et al. (2020). N(6)-Methyladenosine co-transcriptionally directs the demethylation of histone H3K9me2. *Nat. Genet.* 52, 870–877. doi:10.1038/s41588-020-0677-3

Liu, F., Clark, W., Luo, G., Wang, X., Fu, Y., Wei, J., et al. (2016). ALKBH1-Mediated tRNA demethylation regulates translation. *Cell* 167, 816–828. e16. doi:10.1016/j.cell.2016.09.038

Liu, H., Golji, J., Brodeur, L. K., Chung, F. S., Chen, J. T., deBeaumont, R. S., et al. (2019). Tumor-derived IFN triggers chronic pathway agonism and sensitivity to ADAR loss. *Nat. Med.* 25, 95–102. doi:10.1038/s41591-018-0302-5

Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28, 1747–1756. doi:10.1101/gr.239244.118

Meyer, H. J., and Rape, M. (2011). Processive ubiquitin chain formation by the anaphase-promoting complex. *Semin. Cell Dev. Biol.* 22, 544–550. doi:10.1016/j.semcdb.2011.03.009

Meyer, K. D., and Jaffrey, S. R. (2017). Rethinking m(6)A readers, writers, and erasers. *Annu. Rev. Cell Dev. Biol.* 33, 319–342. doi:10.1146/annurev-cellbio-100616-060758

Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., et al. (2015). 5' UTR m(6)A promotes cap-independent translation. *Cell* 163, 999–1010. doi:10.1016/j.cell.2015.10.012

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457. doi:10.1038/nmeth.3337

Psyrri, A., Kalogeras, K. T., Kronenwett, R., Wirtz, R. M., Batistatou, A., Bournakis, E., et al. (2012). Prognostic significance of UBE2C mRNA expression in high-risk early breast cancer. A hellenic cooperative oncology group (HeCOG) study. *Ann. Oncol.* 23, 1422–1427. doi:10.1093/annonc/mdr527

Qiu, Z., Zhao, L., Shen, J. Z., Liang, Z., Wu, Q., Yang, K., et al. (2022). Transcription elongation machinery is a druggable dependency and potentiates immunotherapy in glioblastoma stem cells. *Cancer Discov.* 12, 502–521. doi:10.1158/2159-8290.CD-20-1848

Reddy, S. K., Rape, M., Margansky, W. A., and Kirschner, M. W. (2007). Ubiquitination by the anaphase-promoting complex drives spindle checkpoint inactivation. *Nature* 446, 921–925. doi:10.1038/nature05734

Roundtree, I. A., Evans, M. E., Pan, T., and He, C. (2017). Dynamic RNA modifications in gene expression regulation. *Cell* 169, 1187–1200. doi:10.1016/j.cell.2017.05.045

Shankavaram, U. T., Varma, S., Kane, D., Sunshine, M., Chary, K. K., Reinhold, W. C., et al. (2009). CellMiner: A relational database and query tool for the NCI-60 cancer cell lines. *BMC genomics* 10, 277. doi:10.1186/1471-2164-10-277

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2022). Cancer statistics, 2022. *Ca. A Cancer J. Clin.* 72, 7–33. doi:10.3322/caac.21708

Su, R., Dong, L., Li, Y., Gao, M., Han, L., Wunderlich, M., et al. (2020). Targeting FTO suppresses cancer stem cell maintenance and immune evasion. *Cancer Cell* 38, 79–96. e11. doi:10.1016/j.ccell.2020.04.017

Sun, Y., Dong, D., Xia, Y., Hao, L., Wang, W., and Zhao, C. (2022). YTHDF1 promotes breast cancer cell growth, DNA damage repair and chemoresistance. *Cell Death Dis.* 13, 230. doi:10.1038/s41419-022-04672-5

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca. A Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Taby, R., and Issa, J. P. (2010). Cancer epigenetics. *Ca. Cancer J. Clin.* 60, 376–392. doi:10.3322/caac.20085

Tian, B., and Manley, J. L. (2017). Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell Biol.* 18, 18–30. doi:10.1038/nrm.2016.116

Wood, K. J., Bushell, A., and Hester, J. (2012). Regulatory immune cells in transplantation. *Nat. Rev. Immunol.* 12, 417–430. doi:10.1038/nri3227

Xiang, J. F., Yang, Q., Liu, C. X., Wu, M., Chen, L. L., and Yang, L. (2018). N(6)-Methyladenosines modulate A-to-I RNA editing. *Mol. Cell* 69, 126–135. doi:10.1016/j.molcel.2017.12.006

Xu, Y., Zhang, M., Zhang, Q., Yu, X., Sun, Z., He, Y., et al. (2021). Role of main RNA methylation in hepatocellular carcinoma: N6-Methyladenosine, 5-methylcytosine, and N1-methyladenosine. *Front. Cell Dev. Biol.* 9, 767668. doi:10.3389/fcell.2021.767668

Yamauchi, T., Hoki, T., Oba, T., Jain, V., Chen, H., Attwood, K., et al. (2021). T-cell CX3CR1 expression as a dynamic blood-based biomarker of response to immune checkpoint inhibitors. *Nat. Commun.* 12, 1402. doi:10.1038/s41467-021-21619-0

Yankova, E., Blackaby, W., Albertella, M., Rak, J., De Braekeleer, E., Tsagkogeorga, G., et al. (2021). Small-molecule inhibition of METTL3 as a strategy against myeloid leukaemia. *Nature* 593, 597–601. doi:10.1038/s41586-021-03536-w

Ye, F., Hu, Y., Gao, J., Liang, Y., Liu, Y., Ou, Y., et al. (2021). Radiogenomics map reveals the landscape of m6A methylation modification pattern in bladder cancer. *Front. Immunol.* 12, 722642. doi:10.3389/fimmu.2021.722642

Yin, J. Y., Shen, J., Dong, Z. Z., Huang, Q., Zhong, M. Z., Feng, D. Y., et al. (2011). Effect of eIF3a on response of lung cancer patients to platinum-based chemotherapy by regulating DNA repair. *Clin. Cancer Res.* 17, 4600–4609. doi:10.1158/1078-0432.CCR-10-2591

Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics a J. Integr. Biol.* 16, 284–287. doi:10.1089/omi.2011.0118

Zaccara, S., Ries, R. J., and Jaffrey, S. R. (2019). Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* 20, 608–624. doi:10.1038/s41580-019-0168-5

Zhang, C., Samanta, D., Lu, H., Bullen, J. W., Zhang, H., Chen, I., et al. (2016). Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m^{6}A-demethylation of NANOG mRNA. *Proc. Natl. Acad. Sci. U. S. A.* 113, E2047–E2056. doi:10.1073/pnas.1602883113

Zhang, S., Zhao, B. S., Zhou, A., Lin, K., Zheng, S., Lu, Z., et al. (2017). m(6 A demethylase ALKBH5 maintains tumorigenicity of glioblastoma stem-like cells by sustaining FOXM1 expression and cell proliferation program. *Cancer Cell* 31, 591–606. e6. doi:10.1016/j.ccell.2017.02.013

Zhao, B. S., Roundtree, I. A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* 18, 31–42. doi:10.1038/nrm.2016.132

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 26 September 2022
DOI 10.3389/fgene.2022.977322

Check for updates

OPEN ACCESS

EDITED BY
Claudia Cava,
National Research Council (CNR), Italy

REVIEWED BY
Hanchu Xiong,
Zhejiang Provincial People's Hospital, China
Wulin Yang,
Hefei Institutes of Physical Science (CAS), China

*CORRESPONDENCE
Yinbin Zhang,
23227119@qq.com
Shuqun Zhang,
shuqun_zhang1971@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 24 June 2022
ACCEPTED 06 September 2022
PUBLISHED 26 September 2022

CITATION
Li J, Wu F, Li C, Sun S, Feng C, Wu H, Chen X, Wang W, Zhang Y, Liu M, Liu X, Cai Y, Jia Y, Qiao H, Zhang Y and Zhang S (2022), The cuproptosis-related signature predicts prognosis and indicates immune microenvironment in breast cancer.
Front. Genet. 13:977322.
doi: 10.3389/fgene.2022.977322

COPYRIGHT
© 2022 Li, Wu, Li, Sun, Feng, Wu, Chen, Wang, Zhang, Liu, Liu, Cai, Jia, Qiao, Zhang and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The cuproptosis-related signature predicts prognosis and indicates immune microenvironment in breast cancer

Jia Li[1†], Fei Wu[1†], Chaofan Li[1], Shiyu Sun[1], Cong Feng[1], Huizi Wu[1], Xi Chen[1], Weiwei Wang[1], Yu Zhang[1], Mengji Liu[1], Xuan Liu[1], Yifan Cai[1], Yiwei Jia[1], Hao Qiao[2], Yinbin Zhang[1]* and Shuqun Zhang[1]*

[1]Department of Oncology, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China, [2]Department of Orthopedics, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China

Breast cancer (BC) is the most diagnosed cancer in women. Cuproptosis is new regulated cell death, distinct from known death mechanisms and dependent on copper and mitochondrial respiration. However, the comprehensive relationship between cuproptosis and BC is still blank until now. In the present study, we acquired 13 cuproptosis-related regulators (CRRs) from the previous research and downloaded the RNA sequencing data of TCGA-BRCA from the UCSC XENA database. The 13 CRRs were all differently expressed between BC and normal samples. Using consensus clustering based on the five prognostic CRRs, BC patients were classified into two cuproptosis-clusters (C1 and C2). C2 had a significant survival advantage and higher immune infiltration levels than C1. According to the Cox and LASSO regression analyses, a novel cuproptosis-related prognostic signature was developed to predict the prognosis of BC effectively. The high- and low-risk groups were divided based on the risk scores. Kaplan-Meier survival analysis indicated that the high-risk group had shorter overall survival (OS) than the low-risk group in the training, test and entire cohorts. GSEA indicated that the immune-related pathways were significantly enriched in the low-risk group. According to the CIBERSORT and ESTIMATE analyses, patients in the high-risk group had higher infiltrating levels of antitumor lymphocyte cell subpopulations and higher immune score than the low-risk group. The typical immune checkpoints were all elevated in the high-risk group. Furthermore, the high-risk group showed a better immunotherapy response than the low-risk group based on the Tumor Immune Dysfunction and Exclusion (TIDE) and Immunophenoscore (IPS). In conclusion, we identified two cuproptosis-clusters with different prognoses using consensus clustering in BC. We also developed a cuproptosis-related prognostic signature and nomogram, which could indicate the outcome, the tumor immune microenvironment, as well as the response to immunotherapy.

KEYWORDS

cuproptosis, breast cancer, prognostic signature, tumor immune microenvironment, bioinformatics

Introduction

Breast cancer (BC) accounts for nearly one-third of all cancer cases in women, and its incidence rate increases by 0.5% each year (Siegel et al., 2022). According to the latest estimates, the new female BC cases will be 287,750, and deaths will be 43,250 in 2022 in the United States (Siegel et al., 2022). Histologically, BC includes three subtypes, including HER2-positive, endocrine-dependent, and triple-negative breast cancers (Maughan et al., 2010). Many therapeutic options have been developed, including surgery, chemotherapy, endocrine therapy, and anti-HER2 targeting. However, with standard diagnosis and treatment, it is estimated that 20–30% of patients with BC develop distant metastases, accounting for approximately 90% of the death of BC patients (Britt et al., 2020; Jabbarzadeh Kaboli et al., 2020). Furthermore, the considerable heterogeneity of tumors limits the broad applicability of typing and standard treatment to a certain extent (Waks and Winer, 2019). Thus, exploring the characteristics and potential therapeutic targets of BC patients is of great significance.

Cuproptosis is new regulated cell death (RCD), distinct from known death mechanisms and dependent on copper and mitochondrial respiration (Tsvetkov et al., 2022). Copper could bind to lipoylated components of the tricarboxylic acid (TCA) cycle, leading to lipoylated protein aggregation and subsequent iron-sulfur cluster protein loss, resulting in proteotoxic stress and ultimately cell death. The typical copper ionophores disulfiram (DSF) and elesclomol could induce copper delivery to intracellular compartments by ionophore to induce cuproptosis and are being used as anticancer and chemosensitizing effects in cancer therapeutics (Gehrmann, 2006; Liu et al., 2013; Viola-Rhenals et al., 2018; Yang et al., 2021). Previous studies on copper have focused on the disturbances of copper homeostasis, which is related to a series of diseases, such as Menkes disease (Horn and Wittung-Stafshede, 2021) and Wilson disease (Członkowska et al., 2018). Moreover, elevated copper concentrations have been reported in many cancers, such as breast (Adeoti et al., 2015), lung (Oyama et al., 1994), prostate (Saleh et al., 2020), thyroid (Baltaci et al., 2017), gastrointestinal (Stepien et al., 2017) and gall bladder (Basu et al., 2013) cancers. The role of copper in cancers might partly be promoting blood vessel formation, tumorigenesis, and metastasis (Shanbhag et al., 2019). Many drugs have been developed to regulate copper metabolism in the body for those copper imbalance diseases. Copper chelators could act as an antiangiogenic treatment in many cancers (Brewer et al., 2000; Sen et al., 2002; Chan et al., 2017), regulate autophagy (Krishnan et al., 2018; Bryant et al., 2019), modify the tumor microenvironment (Chan et al., 2017), and enhance the antitumor immunity (Voli et al., 2020).

Cuproptosis could be regulated by specific genes: cuproptosis-related regulators (CRRs), including DLD, PDHB, ATP7B, ATP7A, DLAT, DLST, SLC31A1, DBT, FDX1, LIPIT1, LIAS, GCSH, and PDHA1 (Tsvetkov et al., 2022). Further research into these CRRs could help us understand cuproptosis in disease. Mounting evidence has revealed that signatures based on cell death patterns showed substantial predictive values in prognostic, tumor immune microenvironment (TIME), and immunotherapy response prediction for BC patients, such as ferroptosis (Zhu et al., 2021a), pyroptosis (Xu et al., 2022), and necroptosis (Hu et al., 2022), but studies on the role of cuproptosis in BC lack to some extent. An in-depth study about the association between the genetic changes and expression dysregulations of the CRRs and tumor will be beneficial for the identification of the role of cuproptosis in BC and new therapeutic targets.

Our research comprehensively explored the expression variations, genetic changes, and functions of CRRs in BC. We used consensus clustering analysis to identify two cuproptosis-clusters with different overall survival (OS) and TIME characteristics. We also developed a cuproptosis-related prognostic signature and nomogram, which could indicate the outcome, the tumor immune microenvironment, as well as the response to immunotherapy.

Material and methods

Data acquisition

The RNA sequencing data of BC and normal samples in The Cancer Genome Atlas (TCGA) (113 normal breast samples and 1,109 BC samples and Genotype-Tissue Expression (GTEx) database were downloaded from the UCSC XENA (https://xenabrowser.net/datapages/). The Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) database (Pereira et al., 2016) was downloaded from cBioportal (http://www.cbioportal.org/) (Cerami et al., 2012), and 1,758 of 1,904 BC patients with OS > 30 days were used for analysis. The microarray dataset GSE9893 (N = 155) and GSE96058 (N = 3,069) were downloaded from GEO (http://www.ncbi.nlm.nih.gov/geo/). We used log2 (TPM) data to evaluate the expression of CRRs between BC and normal samples. We excluded male patients and the patients with OS < 30 days and finally remained 916 female patients for further study. The "caret" R package randomly divided our sample into two cohorts at a 1:1 ratio (training and test). The "tableone" R package

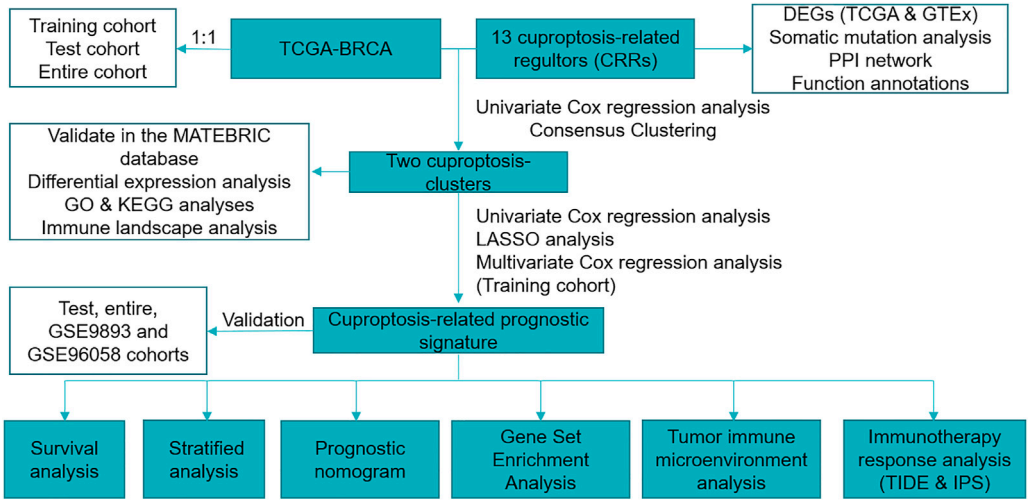


FIGURE 1
Workflow of the current study.

analyzed the clinical features of the training, test, and entire cohorts (Supplementary Table S1). In the subsequent clinicopathological correlation analysis, we excluded patients with incomplete information. We acquired 13 cuproptosis-related regulators (CRRs) from the previous investigation (Supplementary Table S2) (Tsvetkov et al., 2022). Figure 1 showed the workflow of the current study.

Consensus clustering for cuproptosis-clusters

Univariate Cox regression analysis was used to evaluate the CRRs with prognostic values. We used the R package "ConsensusCluster Plus" to perform consensus clustering analysis and identify cuproptosis-clusters in BC patients. We set cluster count (k) between 2 and 9 and selected the optimal k value based on the inflection point of the sum of squared error (SSE). The stability of cuproptosis-clusters was verified by the t-distributed stochastic neighbor embedding (tSNE) algorithm. Furthermore, the Kaplan-Meier survival analysis evaluated the OS of the different cuproptosis-clusters. We further validated the cuproptosis-clusters with the same consensus clustering analysis in the METABRIC data.

Identification of differentially expressed genes and functions of the cuproptosis-clusters

We performed "edgeR" R package to analyze the DEGs between the two cuproptosis-clusters ($p < 0.05$, $|\log FC| = 1$). We used Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) to annotate the functions of DEGs with the R package "ClusterProfiler" (Yu et al., 2012).

Construction and validation of the cuproptosis-related prognosis signature

We retained 336 DEGs expressed in more than half of the patients for further study for the accuracy of the results. These DEGs were submitted to univariate Cox regression analysis to identify the prognostic DEGs in TCGA-training and GSE9893 cohorts. The least absolute shrinkage and selection operator (LASSO) was used to avoid overfitting in the TCGA-training cohort (Friedman et al., 2010). Then, the cuproptosis-related prognostic signature (CRPS) was built with the multivariate Cox regression analysis and stepwise Akaike information criterion (stepAIC) value. Subsequently, each sample could get the risk score according to the formula: Risk score = Σ(Exp∗ Coef). The Coef and Exp were the coefficients and the expression level of each gene, respectively. Based on the median risk score of the training cohort, we divided patients into high- and low-risk groups. For the training, test, entire, GSE9893 and GSE 96058 cohorts, Kaplan-Meier survival analysis and time-dependent receiver operating characteristic (ROC) curves were used to assess the predictive accuracy of the CRPS. We further performed the stratified analyses to assess the prognostic value of the CRPS in different subgroups stratified by age, pathologic stage, T stage, N stage, M stage, ER, PR, and HER2 statuses.

Clinical analysis and construction of the nomogram

We compared the risk score in different clinical characters, including stage, PAM50 subtypes, ER, PR and HER2 status. Multivariate COX regression analysis was performed to assess the independent prognostic factors. The nomogram was constructed with the independent prognostic factors. The calibration curves assessed the accuracy of the nomogram.

The different tumor immune microenvironment patterns between the two risk groups

The variations of pathway activity of the two risk groups were revealed with Gene Set Enrichment Analysis (GSEA) ($p < 0.05$ and FDR< 0.25) (Subramanian et al., 2007). The annotated gene set "c2. cp.kegg.v7.5.1. symbols.gmt" was acquired from the MSigDB (https://www.gsea-msigdb.org/gsea/msigdb/). The CIBERSORT algorithm calculated the proportion of tumor-infiltrating immune cells (Newman et al., 2015). Immune, stromal, and tumor purity scores were evaluated with the ESTIMATE algorithm (Yoshihara et al., 2013). In assessing immune response, well-known predictors of immune checkpoints (ICPs) have been used extensively. Therefore, we compared the tumor-infiltrating immune cells, immune and stromal scores, and 27 ICPs in the two risk groups.

Prediction of the response to immune checkpoint inhibitors treatment

The immunophenoscore (IPS) represents a comprehensive measure of immunogenicity and could predict the patient's response to ICIs treatment (Charoentong et al., 2017). We acquired the IPS score of 916 BC patients from The Cancer Immunome Atlas (https://tcia.at/) and compared the potential immunotherapy in the two risk groups. TIDE provided an easy way to predict the response to ICIs based on evaluating the tumor microenvironment (http://tide.dfci.harvard.edu/) (Jiang et al., 2018; Fu et al., 2020). After submitting the transcriptome profiles of 916 TCGA-BRCA patients to the website, we acquired the information on whether the patients could respond to the ICIs.

Statistical analysis

We applied R software (version 4.0.5, https://www.r-project.org/) for all statistical analyses. Comparative analysis between groups was performed using Student's t-test, Wilcoxon test, or one-way variance analysis. p-value < 0.05 was set as statistically significant, and the significance levels were set as $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$, ns: nonsignificant.

Results

Overview of expression variations, genetic changes and function analyses of CRRs in BC

We found that all of the 13 CRRs were DEGs between BC and normal samples, among which seven regulators (DLD, PDHB, ATP7B, DLAT, SLC31A1, ATP7A, DBT) were upregulated and six regulators (FDX1, LIPT1, LIAS, GCSH, DLST, PDHA1) were downregulated in BC samples (Figure 2A). Figure 2B shows the location of 13 CRRs on the chromosome. The genetic analysis revealed that 36 of the 986 samples (about 3.5%) carried mutations in the regulator of cuproptosis, ATP7A exhibited the highest frequency of mutations, and the majority of mutations were missense mutations. There were no FDX1, DBT, or GCSH mutations in the BC samples (Figure 2C). To better understand the mode of interaction between these CRRs, the protein-protein interaction (PPI) network retained all 13 CRRs with complex regulatory relationships at a high confidence score (0.7) (Figure 2D). We investigated the biologic function and behavior of the 13 CRRs using Metascape and GENEMINIA enrichment analysis. The results of the Metascape database indicated that the 13 CRRs were significantly enriched in the following terms: Glyoxylate metabolism and glycine degradation, copper ion import, protein lipoylation, and cellular amino acid metabolic process (Figure 2E). Furthermore, they were associated with oxidoreductase complex, tricarboxylic acid cycle enzyme complex, acetyl-CoA biosynthetic process, ancl-CoA metabolic process, thioester metabolic process, acetyl-CoA metabolic process, ribonucleoside bisphosphate metabolic process based on the GENEMIAIA database (Figure 2F). According to our analysis, the CRRs were related to cuproptosis, and the expression levels of CRRs were correlated with BC, suggesting that cuproptosis might reflect different traits in patients.

Identification of the cuproptosis-clusters with CRRs in BC

The prognostic values of the 13 CRRs were analyzed with univariate Cox regression analysis, and five (SLC31A1, ATP7A, DLD, DLAT, and DBT) were significantly associated with BC patient prognosis (Figure 3A). The consensus clustering analysis based on these five CRRs explored the relationship between cuproptosis and BC subtypes. Among the clustering variables, k = 2 showed excellent clustering stability with the highest intragroup correlations and the lowest intergroup correlations.

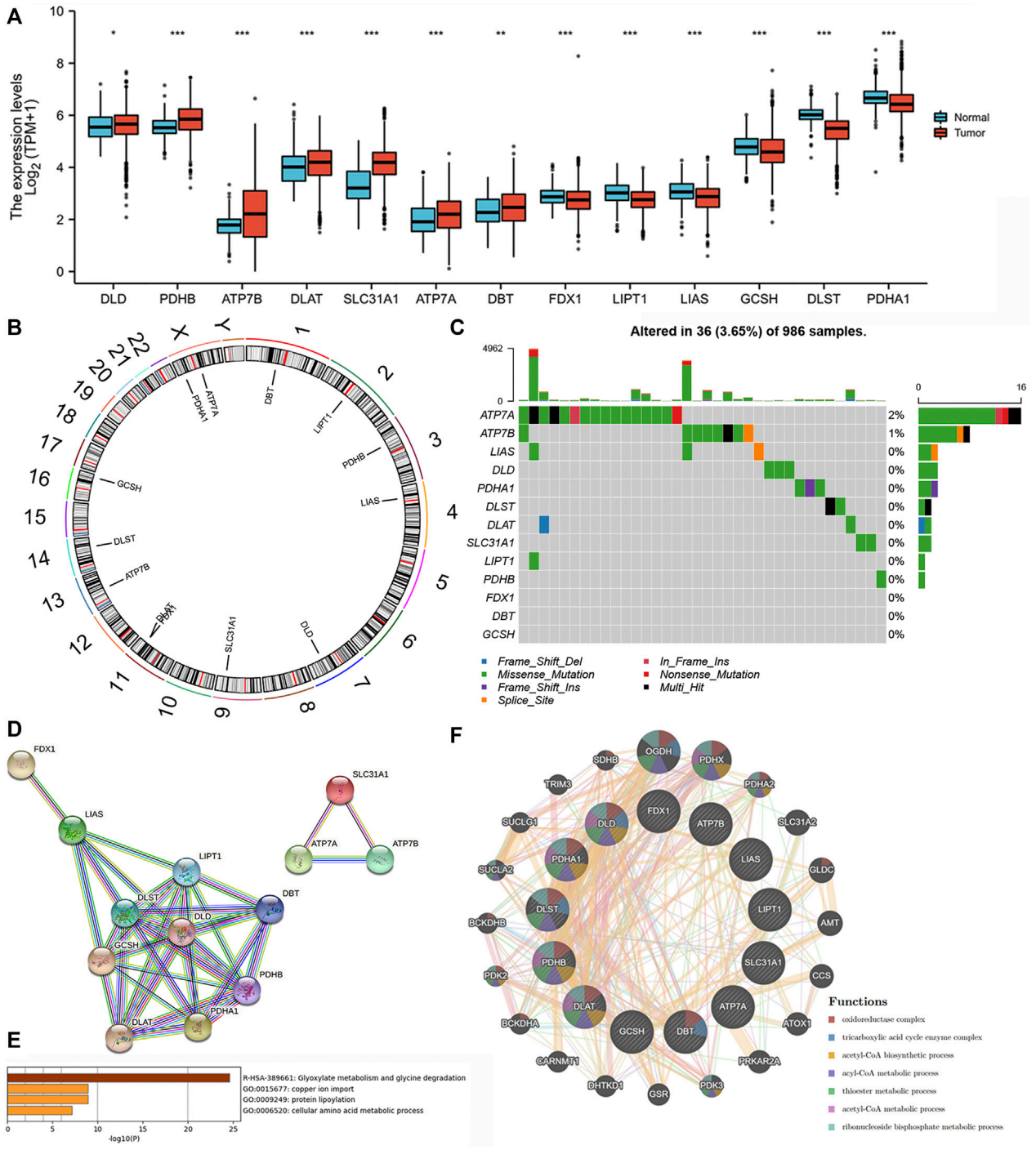


FIGURE 2
Expression variations, Genetic changes, and functional analyses of CRRs in BC. **(A)** The expression of CRRs between BC and normal samples. **(B)** The location of CRRs on chromosomes. **(C)** The schematic overview of mutation frequency and type in CRRs. **(D)** PPI network of the CRRs-encoded proteins. **(E)** The GO analysis of CRRs in the METASCAPE database. **(F)** The network and function of CRRs in the GENEMANIA database.

Therefore, we acquired two cuproptosis-clusters (C1 and C2) (Figures 3B–D). Figure 3E revealed that C1 was significantly different from C2 (tSNE). Moreover, C2 had a more favorable OS than those C1 (Figure 3F). As shown in Figure 3G, the two cuproptosis-clusters exhibited different gene expression profiles and clinical features. Surprisingly, all five CRRs elevated in C1.

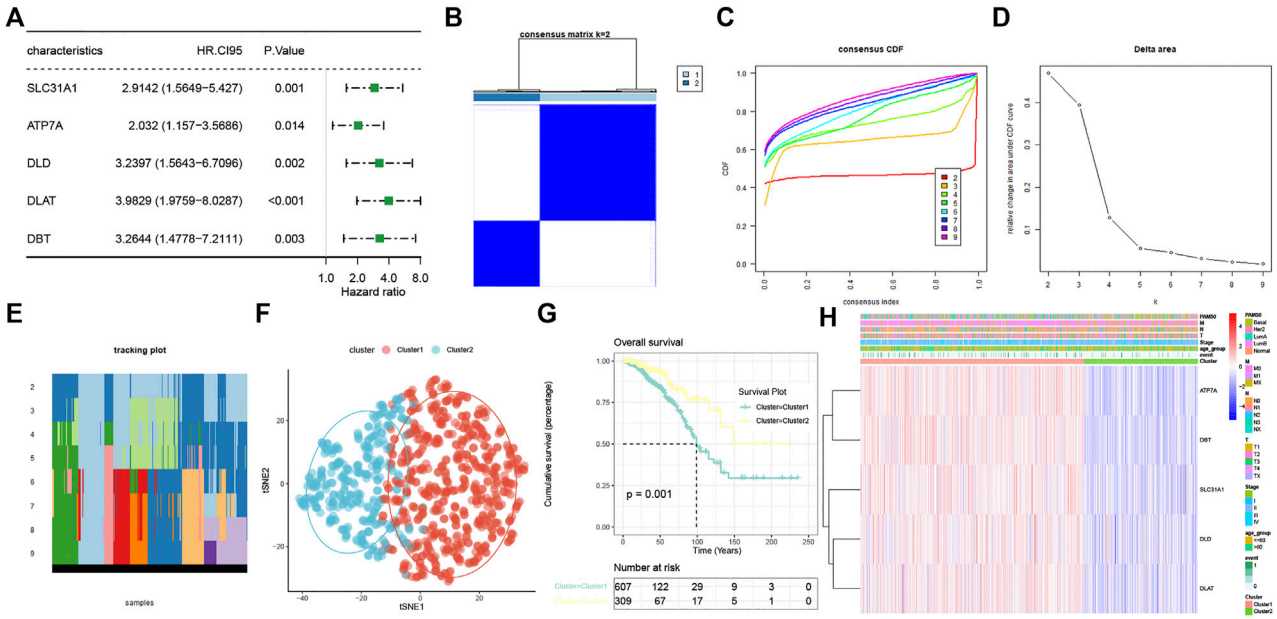


FIGURE 3
Characterization of two cuproptosis-clusters based on CRRs in BC. **(A)** Univariate Cox regression analysis of CRRs. **(B)** Consensus matrix when k = 2. **(C)** Consensus CDF, **(D)** Delta area, **(E)** tracking plot, and tSNE plots **(F)** for validation of the clustering results. **(G)** Kaplan-Meier OS curves for BC patients between C1 and C2. **(H)** Unsupervised clustering of CRRs in the TCGA-BRCA cohort. The cuproptosis-clusters, stage, age, TNM stage, PAM50 subtypes, and survival status were used as patient annotations.

To validate the cuproptosis-clusters in BC, we further performed the same consensus clustering analysis in the METABRIC database, and the results indicated that 1758 BC patients were clustered into two clusters (Supplementary Figures S1A–D). Cluster 1 significantly differed from cluster 2 (tSNE) and had a worse OS than cluster 2 (Supplementary Figures S1E,F). These results indicated that we successfully identified two cuproptosis-clusters of breast cancer.

The cuproptosis-clusters characterized with different immune profiles

We performed the "edgeR" R package and identified 4891 DEGs between the two curproptosis-clusters (Supplementary Figures S2A). The GO and KEGG analyses of the DEGs showed that immune activation pathways were enriched in C2, such as regulation of humoral immune response, cytokine–cytokine receptor interaction, and IL-17 signaling pathway (Supplementary Figures S1B,C). We further performed a series of immune-related analyses. We found that some transcripts of immune activation, such as IFNG, GZMB, CD8A, PRF1, and GZMA expressed higher in C2 than in C1 (Figure 4A). According to the CIBERSORT algorithm, many antitumor lymphocyte cells subpopulations were significantly elevated in C2, such as activated $CD4^+$/$CD8^+$ T cells, NK cells, and plasma. In contrast, M2 macrophages were elevated in the C1 (Figure 4B). The ESTIMATE analysis further revealed that C2 had higher immune and ESTIMATE scores while C1 had higher tumor purity (Figure 4C). The results of TIDE indicated that C2 had a higher proportion of responders to immunotherapy (Figure 4D). These results indicated that the two cuproptosis-clusters characterized with different immune landscapes.

Construction and validation of the cuproptosis-related prognostic signature

To assist clinicians in predicting the prognosis of BC patients, we constructed the CRPS. Using univariate Cox regression analysis, we identified 16 cuproptosis-related DEGs associated with prognosis for BC in the TCGA-training and GSE9893 cohorts (Supplementary Table S3). The LASSO regression algorithm determined six OS-related genes based on the optimum λ value and the minimum partial likelihood of deviance (Figure 5A,B). The six genes were submitted to the multivariate Cox regression analysis. CRPS was formed by incorporating four genes based on AIC values, and its formula is listed below: Risk score = (-0.3005824 * TNFRSF18) + (-0.3676786* SLC1A1) + (0.1479489 * CEACAM6) + (-0.6081390 * GZMM) (Figure 5C). Patients with BC were

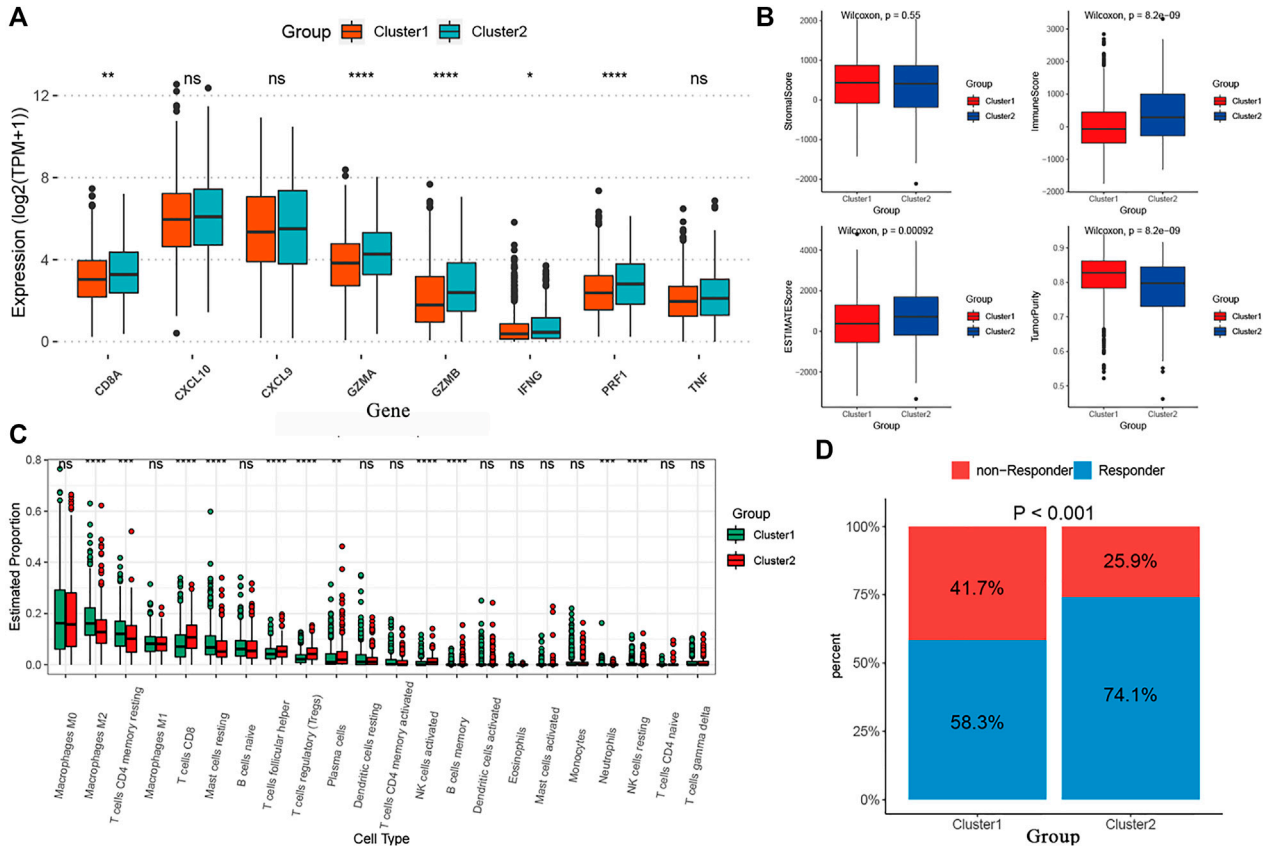


FIGURE 4
The different immune profiles between the two cuproptosis-clusters. **(A)** Comparisons of the immune-activation related gene expression between the two cuproptosis-clusters. **(B)** Comparisons of tumor purity, stromal, immune and estimate scores between the two clusters. **(C)** The box plots of the tumor-infiltrating cells in the two clusters. **(D)** Comparisons of the proportions of non-responders and responders to immunotherapy between the two risk clusters.

categorized into low- and high-risk groups using the median risk scores from the training cohort. Based on the survival analyses, the high-risk group experienced earlier death and a poorer outcome than the low-risk group (Figure 5D). Across the training, test and entire cohorts, the AUCs for the 10-year ROC were 0.794, 0.736, and 0.768, respectively (Figure 5E). Figure 5F showed the distribution plot of the risk score and the different expression levels of the model genes in the two risk groups. We further performed the same analyses in the GSE9893 and GSE96058 databases. The low-risk group had a worse OS than the high-risk group (Supplementary Figures S1A,B). We further selected the appropriate time according to the follow-up characteristics of different cohorts to draw the ROC curves for verification. In GSE9893 database, the AUCs of the 4-, 5-, and 10-year ROC curves were 0.608, 0.695 and 0.593 (Supplementary Figures S3C). In GSE96058 database, the AUCs of the 3-, 5-, and 6-year ROC curves were 0.6065, 0.6215 and 0.6216 (Supplementary Figures S3D). The expression levels of the model genes in the two risk groups were highly similar to the TCGA cohorts (Supplementary Figures S3E,F). Considering the heterogeneity of BC, stratified analyses were further used to assess the prognostic value of CRPS in different clinical subgroups. The results revealed that patients in the low-risk group exhibited a prominent survival benefit compared with the high-risk group in most clinical subgroups, except for luminal B subtypes (Supplementary Figures S4). The above results indicated the prognostic signature was accurate, independent and widely applicable.

Evaluation of the clinical significance of CRPS and development of the cuproptosis-based prognostic nomogram

The potential clinical utility of the risk score of CRPS was further analyzed. The alluvial diagram showed the relationship between cuproptosis-cluster, risk groups and outcomes (Figure 6A). C1 was linked to the high-risk

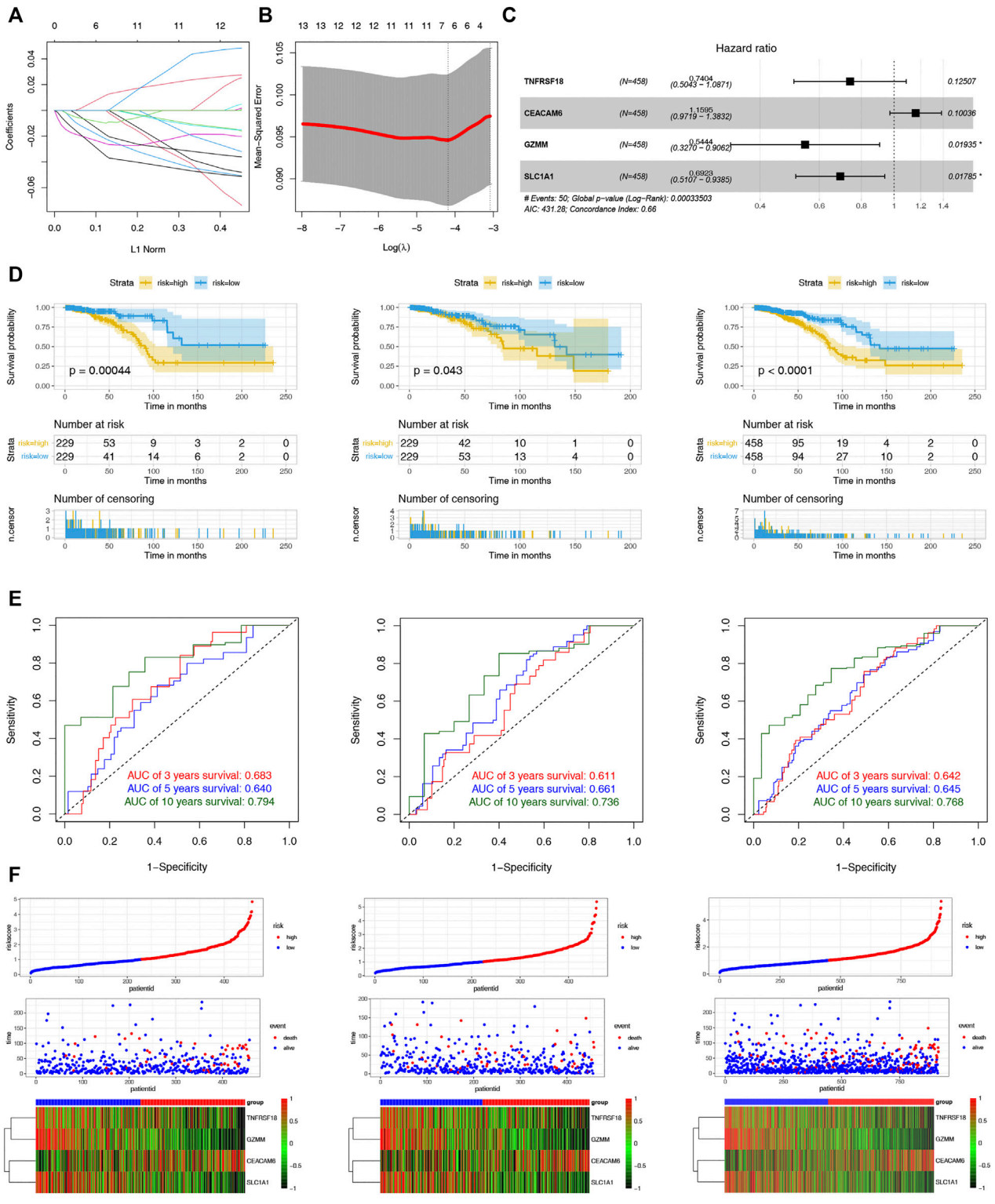


FIGURE 5
Construction and Validation of the CRPS. **(A)** LASSO regression analysis of 108 prognostic DEGs. **(B)** Cross validation method to select optimal genes. **(C)** The forest plot of multivariate Cox regression analysis. **(D)** Kaplan–Meier analyses of the OS between the two risk groups in the training, test, and entire cohort. **(D)**The 3-, 5- and 10-year ROC curves of the CRPS in the training, test, and entire cohort. **(F)** Ranked dot, scatter plots and heat map of the model gene expressions in the training, test, and entire cohort.

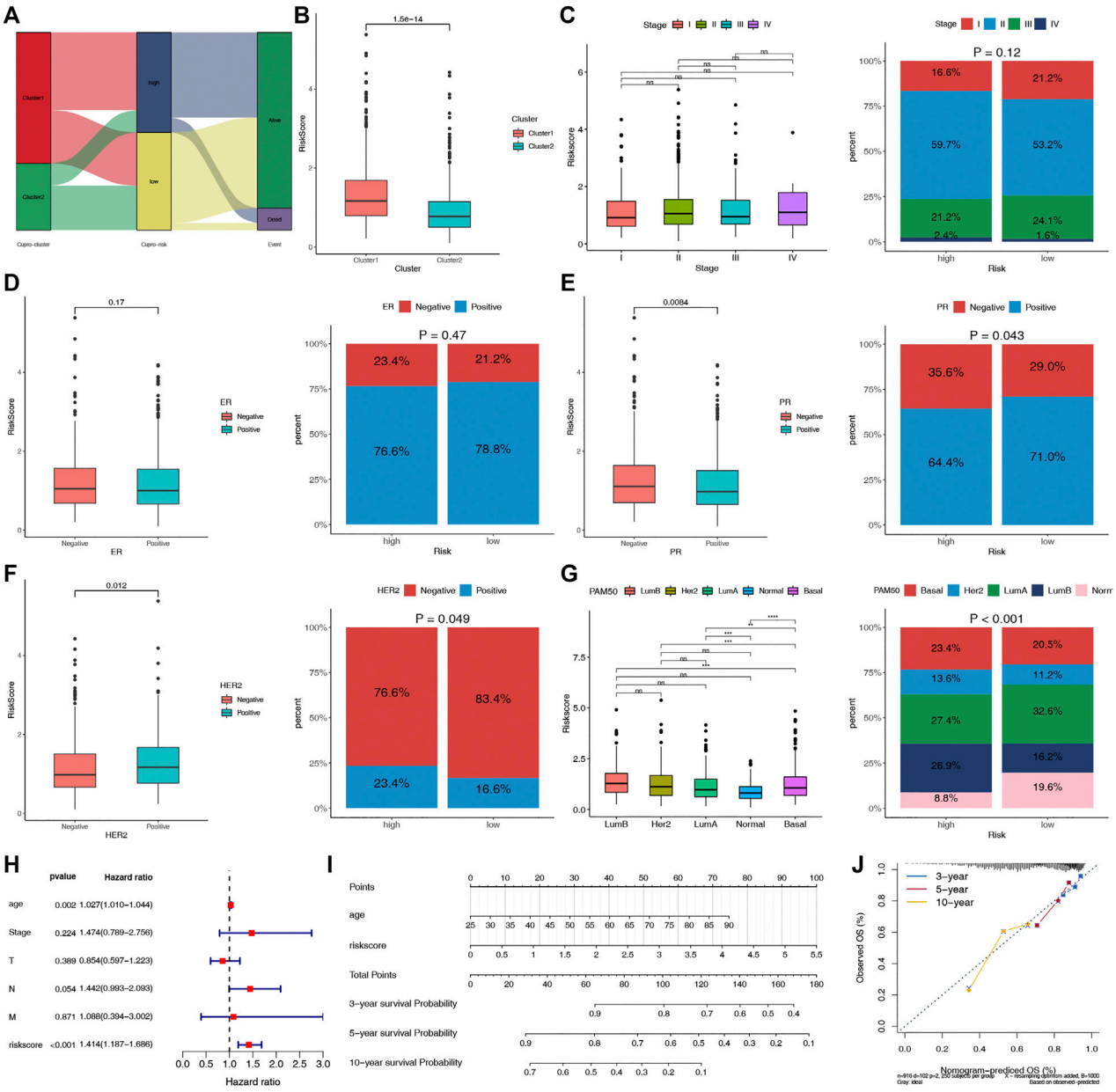


FIGURE 6
Evaluation of the clinical significance of CRPS and development of the cuproptosis-related prognostic nomogram. **(A)** Alluvial diagram of cuproptosis-clusters, risk groups and outcome. **(B)** Comparisons of the risk score in the two cuproptosis-clusters. **(C–G)** The different phenotype Relationships of the risk score and Stage, ER, PR, HER2, PAM50 phenotypes, respectively. **(H)** Multivariate Cox regression analysis of clinical characteristics and risk score. **(I)** The nomogram for predicting BC patients' 3-, 5-, and 10-year OS probability. **(J)** The 3-, 5-, and 10-year calibration curves of the nomogram.

group and death, while C2 was related to the low-risk group and living. Moreover, C1 exhibited a higher risk score than C2 (Figure 6B). HER2-and PR+ were related to the low-risk score, and basal subtype was correlated with the high-risk score (Figures 6C–G). Furthermore, multivariate COX regression analysis revealed that age and the risk score were the independent prognostic factors for BC patients (Figure 6H). We further established the prognostic nomogram to quantitatively predict the 3-, 5- and 10-year survival probability of BC patients (Figure 6I). The calibration curve of the nomogram demonstrated that it could predict the survival probability relatively well (Figure 6J).

TABLE 1 The area under the ROC curve (AUC) showed the sensitivity and specificity of the known gene signatures in predicting the prognosis of BC patients.

Author	Year	Gene signature	Gene number in signature	AUC for OS
Our study	2022	Cuproptosis	4	0.683 (3-year), 0.640 (5-year), 0.794 (10-year)
Lu, et al. (2022a)		Ferroptosis	9	0.713 (3-year), 0.713 (5-year), 0.684 (10-year)
Ding et al. (2021)		Immune	4	0.692 (3-year), 0.691 (5-year), 0.715 (10-year)
Ling et al. (2022)		Pyroptosis	7	0.785 (3-year), 0.769 (5-year), 0.711 (10-year)
Zhong et al. (2020)		Autophagy	6	0.640 (3-year), 0.600 (5-year), 0.610 (10-year)
Zhu et al. (2021)		Ferroptosis	11	0.700 (1-year), 0.749 (2-year), 0.720 (3-year)
Ding et al. (2021)		Ferroptosis	9	0.618 (1-year), 0.653 (2-year), 0.663 (3-year)
Wu et al. (2021a)		Ferroptosis	15	0.740 (1-year), 0.710 (3-year), 0.750 (5-year)
Chen et al. (2022)		Pyroptosis	16	0.756 (1-year), 0.752 (3-year), 0.723 (5-year)
Yu et al. (2022)		Necroptosis	6	0.701 (1-year), 0.716 (2-year), 0.708 (3-year)
Zou et al. (2022)		Apoptosis	6	0.708 (1-year), 0.628 (3-year), 0.687 (5-year)
Min et al. (2021)		Zinc finger protein	14	0.732 (1-year), 0.768 (2-year), 0.737 (3-year)
Lin et al. (2020)		Autophagy	12	0.739 (1-year), 0.717 (2-year), 0.742 (3-year)
Lu et al. (2022b)		Metabolic	2	0.764 (1-year), 0.689 (3-year), 0.612 (5-year)

Comparison with other prognostic signatures

The robustness of our CRPS was assessed by comparing it with 13 existing OS-related prognostic signatures, such as ferroptosis, apoptosis, pyroptosis, necroptosis, immunity, and zinc finger proteins. In order to eliminate the effect of heterogeneity, we included only signatures generated using the TCGA database. AUC was used to assess the predictive power of signatures, and a larger AUC indicates better classification ability (Fawcett, 2006), and it could be used to compare the performance of the signatures (Zhang et al., 2021; Lv et al., 2022). We integrated all the important information of the thirteen signatures, including author, year, gene signature and the AUCs of the signatures (Table 1). Our signature had many advantages in predicting OS in BC patients, especially for the 10-year OS. In our study, the AUCs of the signatures at 3-, 5-, and 10-year were 0.683, 0.640, and 0.794, respectively, significantly higher than most hallmark predictive models. Table 1 showed that the 3-, 5-, and 10-year AUCs of the other four prognostic signatures, namely, the 9 ferroptosis-related gene signature (3-, 5-, and 10-year AUCs: 0.713, 0.713, and 0.684) (Lu et al., 2022a), 4 immune-related gene signature (3-, 5-, and 10-year AUCs: 0.692, 0.691, and 0.715) (Ding et al., 2021), 7 pyroptosis-related gene signature (3-, 5-, and 10-year AUCs: 0.785, 0.769, and 0.711) (Chu et al., 2022), and 6 autophagy-related gene signature (3-, 5-, and 10-year AUCs: 0.640, 0.600, and 0.610) (Zhong et al., 2020) were comparable to the predictive capabilities of our predictive model, and our signature stand out with a clear advantage in predicting the long-term survival of BC patients, with a 10-year AUC of 0.819. We also listed the other signatures that focus more on short-term survival, such as 1-year, 2-year, 3-year, and 5-year survival. We found our signature have similar short-term survival (3-year) prognostic value compared with them, such as the ferroptosis-related gene signature (Wang et al., 2021a; Wu et al., 2021a; Zhu et al., 2021b), pyroptosis-related gene signature (Chen et al., 2022), necroptosis-related gene signature (Yu et al., 2022), apoptosis-related gene signature (Zou et al., 2022), zinc finger protein-related gene signature (Ye et al., 2021), autophagy-related gene signature (Lin et al., 2020), and metabolic-related gene signatures (Lu et al., 2022b). In addition, our model only involves four genes, while other models (11/13) tend to have more, which is more convenient to use to a certain extent. These results indicated that our gene signature outperformed most other signatures in predicting BC prognosis.

Different immune landscapes in the two risk groups

To explore the potential biological processes of the two risk groups, we performed GSEA. Interestingly, a series of immune-related pathways were enriched in the low-risk group, while cell cycle and tumor-related pathways were enriched in the high-risk group (Figures 7A,B). We subsequently used CIBERSORT and ESTIMATE analyses to analyze the TIME of the two risk groups. The ESTIMATE results revealed that the high-risk group had lower stromal, immune, and estimate scores but had higher tumor purity (Figure 7C). The CIBERSORT analysis further indicated that the macrophages M2, M0 and NK cells resting were the main components of the high-risk group; however, most of the antitumor lymphocyte cells, such as macrophages M1, CD8 T cells, dendritic cells resting, and activated NK cells were

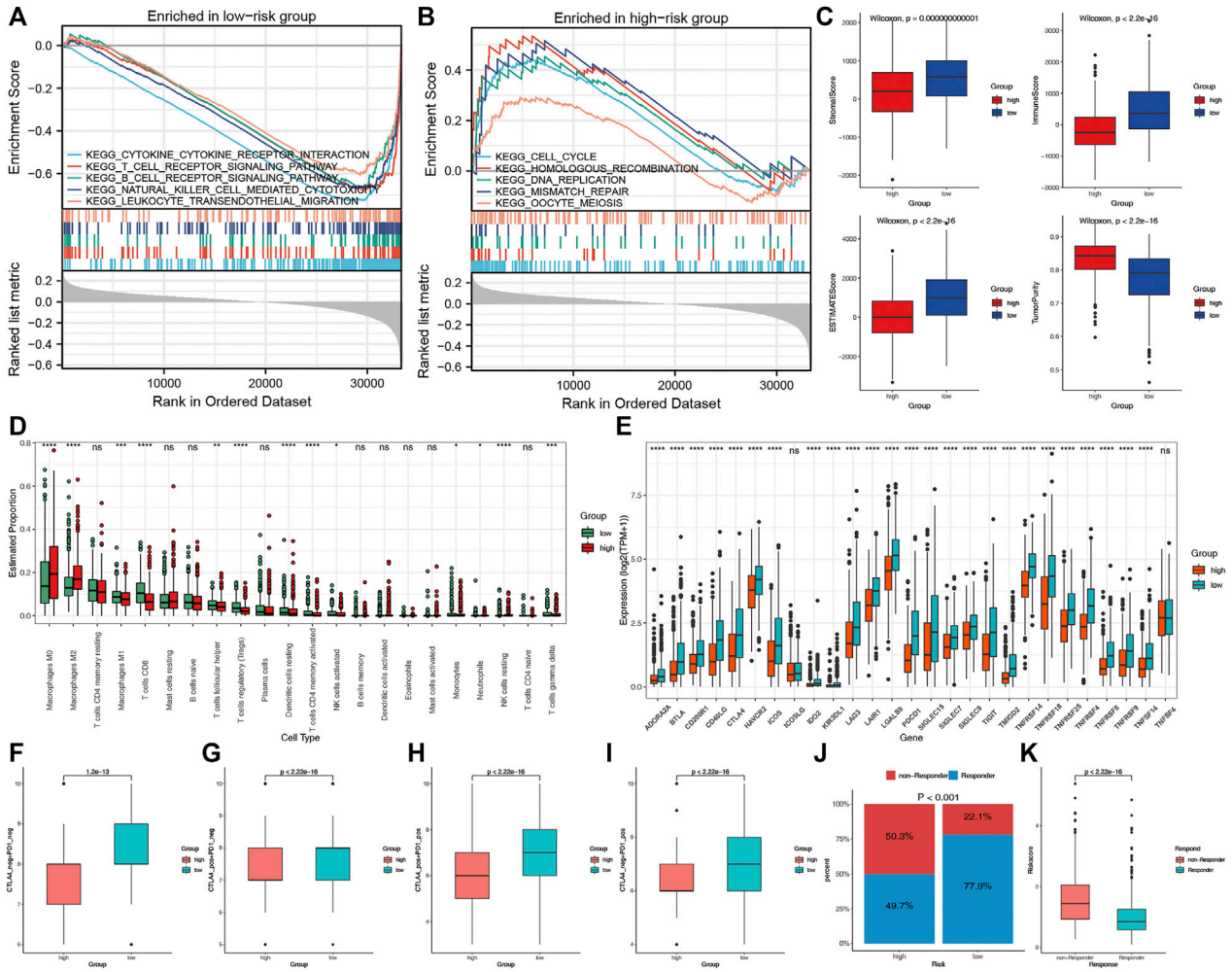


FIGURE 7
The TIME and immunotherapy response between the two risk groups. Significantly enriched pathways in low-risk **(A)** and high-risk **(B)** groups according to the GSEA. **(C)** Comparisons of tumor purity, stromal, immune and estimated scores between the two risk groups. **(D)** The box plots of the tumor-infiltrating cells in the two risk groups. **(E)** The expression levels of 27 ICPs in the two risk groups. **(F–I)** Relative distribution of immunotherapeutic efficacy in the two groups. **(J)** Comparisons of the proportions of non-responders and responders to immunotherapy between the two risk groups. **(K)** Boxplot showed a significantly higher risk score in the responders.

the main components of the low-risk group (Figure 7D). Furthermore, the expression levels of the 27 common ICPs were significantly elevated in the low-risk group, such as PD-1, PD-L1 and CTLA4 (Figure 7E). These results indirectly indicated the different immune landscapes in the two risk groups.

Different immunotherapy responses in the two risk groups

We evaluated the immunotherapy response by the IPS and TIDE. Patients in the low-risk group presented significant therapeutic benefits from ICIs treatment according to the IPS (CTLA4-/PD-1-, CTLA4+/PD-1-, CTLA4-/PD-1+ and CTLA4+/PD-1+) (Figures 7F–I). Based on the TIDE analysis, the responders to immunotherapy in the low-risk group were more than in the high-risk group (77.9 versus 49.7%) (Figure 7J). Furthermore, the risk score was significantly elevated in non-responders to immunotherapy than in responders (Figure 7K). These results revealed that the essential role of CRPS in predicting the response to immunotherapies.

Discussion

Recently, cuproptosis has been identified as a novel RCD (Tsvetkov et al., 2022). It depends on copper and mitochondrial respiration. Previous studies indicated that cuproptosis could be

induced by copper ionophores drugs DSF and elesclomol and had anticancer and chemosensitizing effects (Gehrmann, 2006; Liu et al., 2013; Viola-Rhenals et al., 2018; Yang et al., 2021). The importance of other RCDs had been revealed in BC, but the role of cuproptosis in BC is still unknown (Zhu et al., 2021a; Hu et al., 2022; Xu et al., 2022). Our research was aimed at exploring the importance of cuproptosis in predicting the prognosis, TIME, and immunotherapy response for BC patients, which might lay the foundation for precise cuproptosis-related treatment of BC.

In our study, we firstly revealed the expression patterns, genetic alterations and biological functions of CRRs in BC. We found that all 13 CRRs were DEGs between BC and normal samples. The low mutation rates indicated the maintenance of genome stability. The functional analysis of these 13 CRRs revealed that they were enriched in the biological process of cuproptosis, including copper ion import, tricarboxylic acid cycle enzyme complex, and thioester metabolic process. However, only five CRRs (SLC31A1, ATP7A, DLD, DLAT, and DBT) had the prognostic value and were used for consensus clustering analysis. Among these five CRRs, SLC31A1 was associated with chemoresistance to platinum in osteosarcoma (Cheng et al., 2020), lung cancer (Wang et al., 2021b), and ovarian cancer (Wu et al., 2021b). The copper efflux transporter ATP7A was related to the chemoresistance in esophageal and ovarian cancers (Lukanović et al., 2020; Li et al., 2021). These results showed the significance of CRRs in regulating the occurrence, development and treatment of tumors.

Consensus Clustering is a common method of classifying cancer subtypes (Zhang et al., 2020; Wu et al., 2021c; Han et al., 2021). We performed consensus clustering with the five prognostic CRRs in TCGA-BRCA data and two distinct cuproptosis-clusters were identified. The OS of BC patients were strongly different between the two cuproptosis-clusters. The stability of the cuproptosis-clusters was validated with the METABRIC data. The GO/KEGG analyses revealed that immune-related pathways were differentially enriched in the two cuproptosis-clusters, such as humoral immune response, cytokine–cytokine receptor interaction, and IL-17 signaling pathway. Further ESTIMATE and CIBERSORT analyses identified C1 as immune-excluded, with weakened immune cell infiltration, and C2 as immune-inflamed, with abundant infiltrating immune cells. The TIDE analysis indicated that patients in C2 could respond to immunotherapy better than patients in C1. A recent study identified copper as a factor that upregulates the expression of PD-L1 in tumor cells and modulates signaling pathways involved in PD-L1-mediated death (Voli et al., 2020), which might explain that C1 with higher expression levels of CRRs but be identified as immune-excluded to some extent.

We further developed and validated a CRPS for predicting the prognosis of BC. The low-risk group always had a better prognosis than the high-risk group. The ROC curves certified the reliability of the CRPS. The applicability of the CRPS was confirmed with the stratified analysis. We also found a close relationship between C2, the low-risk group, and the live event. To further explore the predictive ability of our signature, a comparison was performed among several significant molecular signatures employed for predicting OS in patients with BC. Compared with the other signature, such as ferroptosis, necroptosis, pyroptosis, and immune-related (Lin et al., 2020; Zhong et al., 2020; Wang et al., 2021a; Wu et al., 2021a; Zhu et al., 2021b; Ding et al., 2021; Ye et al., 2021; Lu et al., 2022a; Lu et al., 2022b; Chen et al., 2022; Chu et al., 2022; Yu et al., 2022; Zou et al., 2022), our signature indicated much higher AUCs, which indicated a better predictive ability, especially in predicting the long-term survival status. Furthermore, we constructed a prognostic nomogram that could simplify treatment decision-making for patients with BC.

Different tumor immune microenvironments characterized the two risk groups. The TIME was reported to play an essential role in BC, related to the therapeutic response and different clinical outcomes (Tower et al., 2019). The previous study indicated that TIME features were related to the OS in TNBC(Bareche et al., 2020). In the current study, the GSEA results implicated that immune-related pathways were enriched in the low-risk group. We further found that patients in the low-risk group had a higher immune score, estimate score and stromal score than the high-risk group. Furthermore, the low-risk group had an abundance infiltration of dendritic cells resting, activated NK cells, and CD8 T cells. Previous research has reported that these immune cells could protect against tumor growth and promote immune response and immunotherapy (Emens, 2012; Charoentong et al., 2017; Sato et al., 2018; Farhood et al., 2019). However, macrophages M2 were the main components in the high-risk group, which were the critical member in EMT and cancer metastasis (Biswas and Mantovani, 2010; Qian and Pollard, 2010). Furthermore, a series of typical ICPs, such as PD-L1, CTLA4 and HAVCR2, were significantly elevated in the low-risk group. Finally, according to the IPS and TIDE analyses, patients in the low-risk group tended to be the beneficiaries from immunotherapy. These findings suggested that the CRPS were associated with TIME and could guide targeted immunotherapies.

Although the CRPS showed the ability to predict prognosis potentially, there were several limitations to our study. Firstly, the regulatory mechanism of how cuproptosis affects the TIME in breast cancer was limited and deserved further research. In addition, *in vitro* and *in vivo* experiments were needed to confirm our findings.

Conclusion

In conclusion, we firstly identified two cuproptosis-clusters in breast cancer with the different OS using

consensus clustering. We further developed a cuproptosis-related prognostic signature that had good performance in predicting survival outcomes, tumor immune microenvironment and immunotherapy response for BC patients.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://xenabrowser.net/datapages/, http://www.cbioportal.org/, and http://www.ncbi.nlm.nih.gov/geo/.

Author contributions

YZ, JL, and CL conceptualized and designed the study. CF, HW, WW, YC, YJ, HQ, and XC prepared the dataset. SS, ML, and YZ analyzed and interpreted the data. FW and XL wrote and reviewed the manuscript. SZ supervised the study. All authors contributed to the article and approved the submitted version.

Funding

This work is supported by the National Natural Science Foundation of China 82103212.

Acknowledgments

We acknowledge TCGA and GTEx database for providing their platforms and contributors for uploading their meaningful datasets.

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/fgene.2022.977322/full#supplementary-material

References

Adeoti, M. L., Oguntola, A. S., Akanni, E. O., Agodirin, O. S., and Oyeyemi, G. M. (2015). Trace elements; copper, zinc and selenium, in breast cancer afflicted female patients in LAUTECH Osogbo, Nigeria. *Indian J. Cancer* 52 (1), 106–109. doi:10.4103/0019-509X.175573

Baltaci, A. K., Dundar, T. K., Aksoy, F., and Mogulkoc, R. (2017). Changes in the serum levels of trace elements before and after the operation in thyroid cancer patients. *Biol. Trace Elem. Res.* 175 (1), 57–64. doi:10.1007/s12011-016-0768-2

Bareche, Y., Buisseret, L., Gruosso, T., Girard, E., Venet, D., Dupont, F., et al. (2020). Unraveling triple-negative breast cancer tumor microenvironment heterogeneity: Towards an optimized treatment approach. *J. Natl. Cancer Inst.* 112 (7), 708–719. doi:10.1093/jnci/djz208

Basu, S., Singh, M. K., Singh, T. B., Bhartiya, S. K., and Shukla, V. K. (2013). Heavy and trace metals in carcinoma of the gallbladder. *World J. Surg.* 37 (11), 2641–2646. doi:10.1007/s00268-013-2164-9

Biswas, S. K., and Mantovani, A. (2010). Macrophage plasticity and interaction with lymphocyte subsets: Cancer as a paradigm. *Nat. Immunol.* 11 (10), 889–896. doi:10.1038/ni.1937

Brewer, G. J., Dick, R. D., Grover, D. K., LeClaire, V., TsengM.WichaM., et al. (2000). Treatment of metastatic cancer with tetrathiomolybdate, an anticopper, antiangiogenic agent: Phase I study. *Clin. Cancer Res.* 6 (1)–10.

Britt, K. L., Cuzick, J., and Phillips, K-A. (2020). Key steps for effective breast cancer prevention. *Nat. Rev. Cancer* 20 (8), 417–436. doi:10.1038/s41568-020-0266-x

Bryant, K. L., Stalnecker, C. A., Zeitouni, D., Klomp, J. E., Peng, S., Tikunov, A. P., et al. (2019). Combination of ERK and autophagy inhibition as a treatment approach for pancreatic cancer. *Nat. Med.* 25 (4), 628–640. doi:10.1038/s41591-019-0368-8

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2 (5), 401–404. doi:10.1158/2159-8290.CD-12-0095

Chan, N., Willis, A., Kornhauser, N., Ward, M. M., Lee, S. B., Nackos, E., et al. (2017). Influencing the tumor microenvironment: A phase II study of copper depletion using tetrathiomolybdate in patients with breast cancer at high risk for recurrence and in preclinical models of lung metastases. *Clin. Cancer Res.* 23 (3), 666–676. doi:10.1158/1078-0432.CCR-16-1326

Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., et al. (2017). Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 18 (1), 248–262. doi:10.1016/j.celrep.2016.12.019

Chen, H., Luo, H., Wang, J., Li, J., and Jiang, Y. (2022). Identification of a pyroptosis-related prognostic signature in breast cancer. *BMC Cancer* 22 (1), 429. doi:10.1186/s12885-022-09526-z

Cheng, C., Ding, Q., Zhang, Z., Wang, S., Zhong, B., Huang, X., et al. (2020). PTBP1 modulates osteosarcoma chemoresistance to cisplatin by regulating the expression of the copper transporter SLC31A1. *J. Cell. Mol. Med.* 24 (9), 5274–5289. doi:10.1111/jcmm.15183

Chu, L., Yi, Q., Yan, Y., Peng, J., Li, Z., Jiang, F., et al. (2022). A prognostic signature consisting of pyroptosis-related genes and SCAF11 for predicting immune response in breast cancer. *Front. Med.* 9, 882763. doi:10.3389/fmed.2022.882763

Członkowska, A., Litwin, T., Dusek, P., Ferenci, P., Lutsenko, S., Medici, V., et al. (2018). Wilson disease. *Nat. Rev. Dis. Prim.* 4 (1), 21. doi:10.1038/s41572-018-0018-3

Ding, S., Sun, X., Zhu, L., Li, Y., Chen, W., and Shen, K. (2021). Identification of a novel immune-related prognostic signature associated with tumor microenvironment for breast cancer. *Int. Immunopharmacol.* 100, 108122. doi:10.1016/j.intimp.2021.108122

Emens, L. A. (2012). Breast cancer immunobiology driving immunotherapy: Vaccines and immune checkpoint blockade. *Expert Rev. Anticancer Ther.* 12 (12), 1597–1611. doi:10.1586/era.12.147

Farhood, B., Najafi, M., and Mortezaee, K. (2019). CD8+ cytotoxic T lymphocytes in cancer immunotherapy: A review. *J. Cell. Physiol.* 234 (6), 8509–8521. doi:10.1002/jcp.27782

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognit. Lett.* 27 (8), 861–874. doi:10.1016/j.patrec.2005.10.010

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33 (1)–22. doi:10.18637/jss.v033.i01

Fu, J., Li, K., Zhang, W., Wan, C., Zhang, J., Jiang, P., et al. (2020). Large-scale public data reuse to model immunotherapy response and resistance. *Genome Med.* 12 (1), 21. doi:10.1186/s13073-020-0721-z

Gehrmann, M. (2006). Drug evaluation: STA-4783--enhancing taxane efficacy by induction of Hsp70. *Curr. Opin. Investig. Drugs* 7 (6), 574–580.

Han, Z., Yang, B., Wang, Y., Zeng, X., and Tian, Z. (2021). Identification of expression patterns and potential prognostic significance of m^5C-related regulators in head and neck squamous cell carcinoma. *Front. Oncol.* 11, 592107. doi:10.3389/fonc.2021.592107

Horn, N., and Wittung-Stafshede, P. (2021). ATP7A-Regulated enzyme metalation and trafficking in the Menkes disease puzzle. *Biomedicines* 9 (4), 391. doi:10.3390/biomedicines9040391

Hu, T., Zhao, X., Zhao, Y., Cheng, J., Xiong, J., and Lu, C. (2022). Identification and verification of necroptosis-related gene signature and associated regulatory Axis in breast cancer. *Front. Genet.* 13, 842218. doi:10.3389/fgene.2022.842218

Jabbarzadeh Kaboli, P., Salimian, F., Aghapour, S., Xiang, S., Zhao, Q., Li, M., et al. (2020). Akt-targeted therapy as a promising strategy to overcome drug resistance in breast cancer - a comprehensive review from chemotherapy to immunotherapy. *Pharmacol. Res.* 156, 104806. doi:10.1016/j.phrs.2020.104806

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24 (10), 1550–1558. doi:10.1038/s41591-018-0136-1

Krishnan, N., Felice, C., Rivera, K., Pappin, D. J., and Tonks, N. K. (2018). DPM-1001 decreased copper levels and ameliorated deficits in a mouse model of Wilson's disease. *Genes Dev.* 32 (13-14), 944–952. doi:10.1101/gad.314658.118

Li, Z., Li, S., Wen, Y., Chen, J., Liu, K., and Jia, J. (2021). MiR-495 inhibits cisplatin resistance and angiogenesis in esophageal cancer by targeting ATP7A. *Technol. Cancer Res. Treat.* 20, 15330338211039127. doi:10.1177/15330338211039127

Lin, Q-G., Liu, W., Mo, Y-Z., Han, J., Guo, Z. X., Zheng, W., et al. (2020). Development of prognostic index based on autophagy-related genes analysis in breast cancer. *Aging (Albany NY)* 12 (2), 1366–1376. doi:10.18632/aging.102687

Liu, P., Kumar, I. S., Brown, S., Kannappan, V., Tawari, P. E., Tang, J. Z., et al. (2013). Disulfiram targets cancer stem-like cells and reverses resistance and cross-resistance in acquired paclitaxel-resistant triple-negative breast cancer cells. *Br. J. Cancer* 109 (7), 1876–1885. doi:10.1038/bjc.2013.534

Lu, J., Liu, P., and Zhang, R. (2022). A metabolic gene signature to predict breast cancer prognosis. *Front. Mol. Biosci.* 9, 900433. doi:10.3389/fmolb.2022.900433

Lu, Y-J., Gong, Y., Li, W-J., Zhao, C. Y., and Guo, F. (2022). The prognostic significance of a novel ferroptosis-related gene model in breast cancer. *Ann. Transl. Med.* 10 (4), 184. doi:10.21037/atm-22-479

Lukanović, D., Herzog, M., Kobal, B., and Cerne, K. (2020). The contribution of copper efflux transporters ATP7A and ATP7B to chemoresistance and personalized medicine in ovarian cancer. *Biomed. Pharmacother.* 129, 110401. doi:10.1016/j.biopha.2020.110401

Lv, W., Yu, H., Han, M., Tan, Y., Wu, M., Zhang, J., et al. (2022). Analysis of tumor glycosylation characteristics and implications for immune checkpoint inhibitor's efficacy for breast cancer. *Front. Immunol.* 13, 830158. doi:10.3389/fimmu.2022.830158

Maughan, K. L., Lutterbie, M. A., and Ham, P. S. (2010). Treatment of breast cancer. *Am. Fam. Physician* 81 (11), 1339–1346.

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Oyama, T., Matsuno, K., Kawamoto, T., MiTsudomi, T., Shirakusa, T., and Kodama, Y. (1994). Efficiency of serum copper/zinc ratio for differential diagnosis of patients with and without lung cancer. *Biol. Trace Elem. Res.* 42 (2), 115–127. doi:10.1007/BF02785383

Pereira, B., Chin, S-F., Rueda, O. M., Vollan, H. K. M., Provenzano, E., Bardwell, H. A., et al. (2016). The somatic mutation profiles of 2, 433 breast cancers refines their genomic and transcriptomic landscapes. *Nat. Commun.* 7, 11479. doi:10.1038/ncomms11479

Qian, B-Z., and Pollard, J. W. (2010). Macrophage diversity enhances tumor progression and metastasis. *Cell* 141 (1), 39–51. doi:10.1016/j.cell.2010.03.014

Saleh, S. A. K., Adly, H. M., Abdelkhaliq, A. A., and Nassir, A. M. (2020). Serum levels of selenium, zinc, copper, manganese, and iron in prostate cancer patients. *Curr. Urol.* 14 (1), 44–49. doi:10.1159/000499261

Sato, Y., Bolzenius, J. K., Eteleeb, A. M., Su, X., Maher, C. A., Sehn, J. K., et al. (2018). CD4+ T cells induce rejection of urothelial tumors after immune checkpoint blockade. *JCI Insight* 3 (23), 121062. doi:10.1172/jci.insight.121062

Sen, C. K., Khanna, S., Venojarvi, M., Trikha, P., Ellison, E. C., Hunt, T. K., et al. (2002). Copper-induced vascular endothelial growth factor expression and wound healing. *Am. J. Physiol. Heart Circ. Physiol.* 282 (5), H1821–H1827. doi:10.1152/ajpheart.01015.2001

Shanbhag, V., Jasmer-Mcdonald, K., Zhu, S., Martin, A. L., Gudekar, N., Khan, A., et al. (2019). ATP7A delivers copper to the lysyl oxidase family of enzymes and promotes tumorigenesis and metastasis. *Proc. Natl. Acad. Sci. U. S. A.* 116 (14), 6836–6841. doi:10.1073/pnas.1817473116

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2022). Cancer statistics, 2022. *Ca. Cancer J. Clin.* 72 (1), 7–33. doi:10.3322/caac.21708

Stepien, M., Jenab, M., Freisling, H., Becker, N. P., Czuban, M., Tjonneland, A., et al. (2017). Pre-diagnostic copper and zinc biomarkers and colorectal cancer risk in the European Prospective Investigation into Cancer and Nutrition cohort. *Carcinogenesis* 38 (7), 699–707. doi:10.1093/carcin/bgx051

Subramanian, A., Kuehn, H., Gould, J., Tamayo, P., and Mesirov, J. P. (2007). GSEA-P: A desktop application for gene set enrichment analysis. *Bioinformatics* 23 (23), 3251–3253. doi:10.1093/bioinformatics/btm369

Tower, H., Ruppert, M., and Britt, K. (2019). The immune microenvironment of breast cancer progression. *Cancers (Basel)* 11 (9), E1375. doi:10.3390/cancers11091375

Tsvetkov, P., Coy, S., Petrova, B., Dreishpoon, M., Verma, A., Abdusamad, M., et al. (2022). Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science* 375 (6586), 1254–1261. doi:10.1126/science.abf0529

Viola-Rhenals, M., Patel, K. R., Jaimes-Santamaria, L., Wu, G., Liu, J., and Dou, Q. P. (2018). Recent advances in antabuse (disulfiram): The importance of its metal-binding ability to its anticancer activity. *Curr. Med. Chem.* 25 (4), 506–524. doi:10.2174/0929867324666171023161121

Voli, F., Valli, E., Lerra, L., Kimpton, K., Saletta, F., Giorgi, F. M., et al. (2020). Intratumoral copper modulates PD-L1 expression and influences tumor immune evasion. *Cancer Res.* 80 (19), 4129–4144. doi:10.1158/0008-5472.CAN-20-0471

Waks, A. G., and Winer, E. P. (2019). Breast cancer treatment: A review. *JAMA* 321 (3), 288–300. doi:10.1001/jama.2018.19323

Wang, D., Wei, G., Ma, J., Cheng, S., Jia, L., Song, X., et al. (2021). Identification of the prognostic value of ferroptosis-related gene signature in breast cancer patients. *BMC Cancer* 21 (1), 645. doi:10.1186/s12885-021-08341-2

Wang, L., Sun, C., Li, X., Mao, C., Qian, J., Wang, J., et al. (2021). A pharmacogenetics study of platinum-based chemotherapy in lung cancer: *ABCG2* polymorphism and its genetic interaction with *SLC31A1* are associated with response and survival. *J. Cancer* 12 (5), 1270–1283. doi:10.7150/jca.51621

Wu, G., Peng, H., Tang, M., Yang, M., Wang, J., Hu, Y., et al. (2021). ZNF711 down-regulation promotes CISPLATIN resistance in epithelial ovarian cancer via interacting with JHDM2A and suppressing SLC31A1 expression. *EBioMedicine* 71, 103558. doi:10.1016/j.ebiom.2021.103558

Wu, G., Wang, Q., Zhu, T., Fu, L., Li, Z., Wu, Y., et al. (2021). Identification and validation of immune-related LncRNA prognostic signature for lung adenocarcinoma. *Front. Genet.* 12, 681277. doi:10.3389/fgene.2021.681277

Wu, Z-H., Tang, Y., Yu, H., and Li, H. D. (2021). The role of ferroptosis in breast cancer patients: A comprehensive analysis. *Cell Death Discov.* 7 (1), 93. doi:10.1038/s41420-021-00473-5

Xu, L., Hu, Y., and Liu, W. (2022). Pyroptosis-mediated molecular subtypes are characterized by distinct tumor microenvironment infiltration characteristics in breast cancer. *J. Inflamm. Res.* 15, 345–362. doi:10.2147/JIR.S349186

Yang, H., Chen, X., Li, K., Cheaito, H., Yang, Q., Wu, G., et al. (2021). Repurposing old drugs as new inhibitors of the ubiquitin-proteasome pathway for cancer treatment. *Semin. Cancer Biol.* 68, 105–122. doi:10.1016/j.semcancer.2019.12.013

Ye, M., Li, L., Liu, D., Wang, Q., Zhang, Y., and Zhang, J. (2021). Identification and validation of a novel zinc finger protein-related gene-based prognostic model for breast cancer. *PeerJ* 9, e12276. doi:10.7717/peerj.12276

Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612

Yu, G., Wang, L-G., HaN, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Yu, H., Lv, W., Tan, Y., He, X., Wu, Y., Wu, M., et al. (2022). Immunotherapy landscape analyses of necroptosis characteristics for breast cancer patients. *J. Transl. Med.* 20 (1), 328. doi:10.1186/s12967-022-03535-z

Zhang, B., Tang, B., Gao, J., Li, J., Kong, L., and Qin, L. (2020). A hypoxia-related signature for clinically predicting diagnosis, prognosis and immune microenvironment of hepatocellular carcinoma patients. *J. Transl. Med.* 18 (1), 342. doi:10.1186/s12967-020-02492-9

Zhang, D., Li, Y., Yang, S., Wang, M., Yao, J., Zheng, Y., et al. (2021). Identification of a glycolysis-related gene signature for survival prediction of ovarian cancer patients. *Cancer Med.* 10 (22), 8222–8237. doi:10.1002/cam4.4317

Zhong, S., Chen, H., Yang, S., Feng, J., and Zhou, S. (2020). Identification and validation of prognostic signature for breast cancer based on genes potentially involved in autophagy. *PeerJ* 8, e9621. doi:10.7717/peerj.9621

Zhu, L., Chen, M., Huang, B., Zhang, T., Chen, K., Lian, H., et al. (2021). Genomic analysis uncovers immune microenvironment characteristics and drug sensitivity of ferroptosis in breast cancer brain metastasis. *Front. Genet.* 12, 819632. doi:10.3389/fgene.2021.819632

Zhu, L., Tian, Q., Jiang, S., Gao, H., Yu, S., Zhou, Y., et al. (2021). A novel ferroptosis-related gene signature for overall survival prediction in patients with breast cancer. *Front. Cell Dev. Biol.* 9, 670184. doi:10.3389/fcell.2021.670184

Zou, R., Zhao, W., Xiao, S., and Lu, Y. (2022). A signature of three apoptosis-related genes predicts overall survival in breast cancer. *Front. Surg.* 9, 863035. doi:10.3389/fsurg.2022.863035

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 26 September 2022
DOI 10.3389/fgene.2022.1001374

Check for updates

OPEN ACCESS

EDITED BY
Claudia Cava,
Institute of Bioimaging and Molecular Physiology (CNR), Italy

REVIEWED BY
Jin Gao,
Washington State University, United States
Guangjian Fan,
Shanghai General Hospital, China

*CORRESPONDENCE
Jing Zhang,
zhangjing1179@163.com
Yongli Zhang,
doctorzhang2@163.com

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 23 July 2022
ACCEPTED 06 September 2022
PUBLISHED 26 September 2022

CITATION
Chen J, Wang G, Luo X, Zhang J and Zhang Y (2022), Cuproptosis patterns and tumor microenvironment in endometrial cancer.
Front. Genet. 13:1001374.
doi: 10.3389/fgene.2022.1001374

COPYRIGHT
© 2022 Chen, Wang, Luo, Zhang and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Cuproptosis patterns and tumor microenvironment in endometrial cancer

Junfeng Chen[1], Guocheng Wang[1], Xiaomei Luo[1], Jing Zhang[2]* and Yongli Zhang[1]*

[1]Department of Obstetrics and Gynecology, Shanghai First Maternity and Infant Hospital, Tongji University School of Medicine, Shanghai, China, [2]Department of Gynecological Oncology, The First Affiliated Hospital of Bengbu Medical College, Bengbu, China

Cuproptosis is the most recently discovered mode of cell death. It could affect the metabolism of cancer cells and surrounding infiltrating immune cells. In recent years, many studies have also shown that the tumor microenvironment (TME) plays a critical role in tumor growth and development. Mounting evidence suggests that Cuproptosis would bring unique insights into the development of pharmacological and nonpharmacological therapeutic techniques for cancer prevention and therapy. However, no study has been done on the combination of cuproptosis and TME in any cancer. Herein, we investigated the relationship between cuproptosis-related genes (CRGs), TME, and the prognosis of patients with Uterine Corpus Endometrial Carcinoma (UCEC). We identified three CRGs clusters based on 10 CRGs and three CRGs gene clusters based on 600 differentially expressed genes (DEGs) with significant prognostic differences. Following that, the CRGs score based on DEGs with significant prognostic differences was established to evaluate the prognosis and immunotherapeutic efficacy of UCEC patients. The CRGs score was shown to be useful in predicting clinical outcomes. Patients with a low CRGs score seemed to have a better prognosis, a better immunotherapeutic response, and a higher tumor mutation burden (TMB). In conclusion, our study explored the influence of cuproptosis patterns and TME on the prognosis of cancer patients, thereby improving their prognosis.

KEYWORDS

endometrial cancer, cuproptosis, tumor micoenvironment, immunothearpy, prognosis

Introduction

From plants to animals, all biology nearly needs metal ions to induce regulated cell death through different subroutines for normal health (Zhang et al., 2019). More or fewer of these metals can also play a role as the central gate-keeper for normal human functions (Zahn et al., 2021). In 2012, the concept of ferroptosis was discovered and proposed for the first time (Tang et al., 2021). Numerous investigations have been conducted over the last decade to investigate the mechanisms and roles of ferroptosis in cancer. (Proneth and Conrad, 2019; Jiang et al., 2021; Tang et al., 2021). It has great potential for tumor growth,

drug resistance, and immune surveillance (Chen et al., 2021). In March 2022, Tsvetkov et al. (Tsvetkov et al., 2022) discovered a new mechanism of cell death that is copper-dependent, regulates cell death, differs from known death mechanisms, and relies on mitochondrial respiration. Additionally, they discovered that three genes (GLS, MTF1, and CDKN2A) sensitized the cells to cuproptosis whereas seven genes (LIAS, PDHA1, LIPT1, FDX1, DLD, DLAT, and PDHB) provided resistance to cuproptosis (Tsvetkov et al., 2022). Adenosine triphosphate (ATP) production is not necessary for cuproptosis; rather, it relies on mitochondrial respiration (Tsvetkov et al., 2022). Cuproptosis is caused by direct copper binding to lipoylated components of the tricarboxylic acid (TCA) cycle, lipoylated protein aggregation, and *Fe-S* cluster protein loss in mitochondria, all of which result in proteotoxic stress and eventually cellular death (Tang et al., 2022). Apoptosis, necroptosis, pyroptosis, ferroptosis, and other types of cell death, among others, were all connected to tumor development, metastasis, and immune treatment, according to earlier research (Andersen et al., 2005; Dannappel et al., 2014; Karki and Kanneganti, 2019; Lei et al., 2022).

In China, endometrial cancer is the most common gynecologic cancer, and its prevalence is increasing (Brooks et al., 2019). Despite the growing number of tumor treatments, the prognosis of advanced endometrial cancer has still not significantly improved (Ventriglia et al., 2017). Among the numerous cancer treatments, immunotherapy has emerged as a powerful clinical strategy for the treatment of cancer (Riley et al., 2019). Immune checkpoints have emerged as a new treatment strategy in oncology (O'Donnell et al., 2019). Many patients have resistance to immunotherapy, resulting in only a small percentage of patients benefiting from it (O'Donnell et al., 2019). Numerous studies have shown that the efficiency of immunotherapy is closely related to TME heterogeneity and metabolic plasticity (Li et al., 2019; Martin et al., 2020; Goliwas et al., 2021). The TME is comprised of multiple components, including chemokines, growth factors, exosomes, cytokines, and other molecules. So, copper can be a participant in the regulation of physiological and pathological processes in cancer (Li et al., 2014; Voli et al., 2020; Liu et al., 2021). Therefore, the investigation of inhibitors and regulators of copper-related variables as potential cancer therapeutics has gained increasing interest in the scientific community (Li et al., 2014; Garber, 2015; Voli et al., 2020; Liu et al., 2021).

However, there is a research gap on cuproptosis in UCEC. Therefore, there is an urgent need to comprehensively explore the prognostic significance of the cuproptosis and its association with TME infiltrating features in UCEC. Based on the advances in RNA sequencing, we mainly analyzed related-data from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) by integrating multi-omics approaches.

Materials and methods

Data collection and processing

We downloaded the expression data, clinical information, and immunophenoscores (IPSs) for UCEC directly from TCGA (https://cancergenome.nih.gov/), GEO (https://www.ncbi.nlm.nih.gov/geo/), and the Cancer Immunome Atlas (TCIA) (https://tcia.at/home). The TCGA database contained information on 552 UCEC patients and 23 tumor-free individuals, and GSE17025, containing 80 UCEC samples, was downloaded from GEO. The TCGA also provided data on copy number variation (CNV) and somatic mutation data. To normalize gene expression data, the limma R package was used to convert fragments per kilobase per million (FPKM) values to transcripts per kilobase per million (TPM) values. R (*R version 4.1.1*) biological conductor packages were used to extract and analyze expression data and clinical information. Previous research (Tsvetkov et al., 2022) by Todd R. Golub's team has identified 10 genes (FDX1, LIAS, LIPT1, DLD, DLAT, PDHA1, PDHB, MTF1, GLS, and CDKN2A) that are closely related to the occurrence and development of cuproptosis. Therefore, we selected these ten genes as CRGs for further research.

Analysis of cuproptosis-related genes clusters based on 10 cuproptosis-related genes

The number of clusters and their stability were determined by consensus clustering. Based on the expression of 10 CRGs, we used unsupervised cluster analysis to classify all samples into three cuproptosis patterns with the optimal k value. We analyzed all samples using the "ConsensClusterPlus" R software package and ran cycle computation 1,000 to ensure classification stability. Principal component analysis (PCA) is used to estimate the distribution of molecular patterns. R package "GSVA" is a non-parametric and unsupervised method for estimating the gene set variation and activity change of biological processes in expression dataset samples (Hänzelmann et al., 2013). We used it to investigate the biological processes between different CRGs. We downloaded the gene set of "c2. cp.kegg.v7.5.1. symbols" from the MSigDB database (https://www.gsea-msigdb.org/gsea/msigdb/) for GSVA analysis. It was considered statistically significant when the adjusted p-value was < 0.05. The functions of CRGs were annotated using the "clusterProfiler" R program, with a critical value of false discovery rate (FDR) was < 0.05.

Analysis of tumor microenvironment cell infiltration

We used single-sample gene-set enrichment analysis (ssGSEA) to quantify the relative abundance of each immune

cell infiltration in the TME of the UCEC samples to better understand the degree of immune cell infiltration in the three CRGs clusters (Charoentong et al., 2017). The relevant gene set which distinguishes between infiltrating immune cell subtypes in TME was previously identified (Barbie et al., 2009; Charoentong et al., 2017). The relative abundance of different immune infiltrating cells, including activated DCs, activated CD8 T cells, macrophages, natural killer T cells, and regulatory T cells, in the TME in each sample was evaluated using the enrichment scores calculated by ssGSEA analysis.

Analysis of cuproptosis-related genes clusters based on 600 differentially expressed genes with independent prognostic value

DEGs in nontumor tissues and UCEC samples were analyzed by the limma and ggplot2 packages, and the significance cutoff of DEGs was $|\log FC| > 1$, $p < 0.05$ (Ritchie et al., 2015). The R package "VennDiatram" is used to construct Venn diagrams to identify common DEGs. Further, univariate Cox regression analysis identified the common DEGs with significant prognostic differences. Based on the expression of those DEGs with significant prognostic differences, the consensus clustering analysis was performed to determine the number of CRGs gene clusters resulting from biological variations. The Kaplan-Meier analysis was employed to perform the survival analysis between CRGs gene clusters.

Construction of the cuproptosis-related genes score

To quantify the cuproptosis patterns of individual UCEC patients, we constructed a CRGs score system for UCEC patients. We selected 600 DEGs with significant prognostic differences to construct the CRGs score, and the principal components 1 and 2 were chosen as signature scores of cuproptosis patterns. The CRGs score was calculated as follows:

$$CRGs\ score = \sum (PC1i + PC2i)$$

where i represents the expression of independent prognostic DEGs. Then, we separated all patients into high-and low-risk groups according to the CRGs score using the R package "ggalluvial."

Correlation between cuproptosis-related genes scores and immunotherapy

The IPS was extracted based on four categories of immunogenicity-related genes (MHC molecules, effector cells, immune modulators, and immunosuppressive cells). The expression of genes in different cell types is represented by the IPS score, which ranges from 0 to 10. Meanwhile, the values of IPS were positively correlated with immunotherapeutic efficacy by R package "ggpubr."

Statistical analysis

The correlation coefficients between the expression of 10 CRGs and infiltrating immune cells in TME were calculated by Spearman's correlation analysis. The *t*-test and the Kruskal-Wallis test were used to compare two groups; one-way ANOVA and the Kruskal-Wallis test were used to compare three or more groups; and the Chi-square test and Fisher's exact test were used to compare categorical variables. The Kaplan-Meier analysis with a log-rank test was used to plot the survival curve for the prognosis analysis. The waterfall diagram shows the mutation landscape using the maftools package. The *p-value* < 0.05 was considered statistically significant.

Results

The landscape (copy number variations, gene expression and mutation) of 10 cuproptosis-related genes in endometrial cancer

According to previous reports (Ge et al., 2022; Tsvetkov et al., 2022), a total of 10 CRGs were included for analysis in this study. Figure 1 summarizes the mutation rates of CNV and somatic mutations of CRGs. Only 78 (15.06%) of 518 samples caused mutations of CRGs (Figure 1A), but 10 CRGs were all mutated. Furthermore, the CNV frequency of 10 CRGs was shown to be common in UCEC, the majority of which were associated with copy number amplification (Figure 1B). However, *FDX1, PDHB,* and *DLAT* had a higher incidence of copy number deletion (Figure 1B). The locations of CNV alterations of 10 CRGs on the human chromosomes in the UCEC cohort are presented in Figure 1C. To investigate whether the aberrant expression were associated with UCEC, we analyzed the mRNA expression of CRGs in UCEC and normal tissues (Figure 1D). Most mRNA levels of CRGs differed considerably between normal and tumor samples (Figure 1D). Meanwhile, the protein expression (Figure 2) of the six CRGs in the normal and tumor groups was analyzed in the Human Protein Atlas (HPA) database (https://www.proteinatlas.org/). Supplementary Figure S1 revealed the expression of five genes (*LIPT1, DLD, PDHA1, GLS,* and *CDKN2A*) was strongly connected with the outcome of UCEC patients, and it demonstrated abnormal expression of CRGs was related to cancer carcinogenesis, development, and progression.

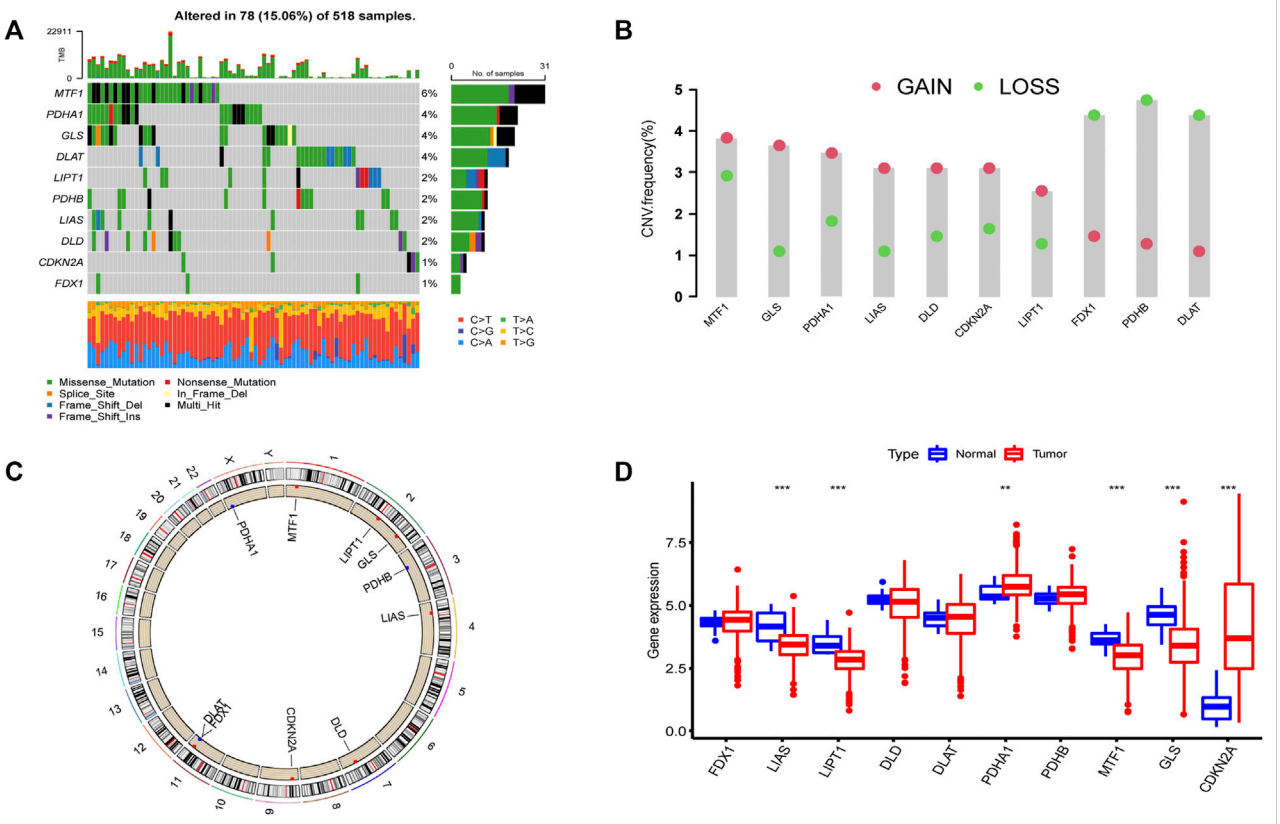


FIGURE 1
Landscape of 10 CRGs in endometrial cancer. **(A)** The mutation incidence of 10 CRGs in the TCGA-UCEC cohort. **(B)** CNV frequency of 10 CRGs in the TCGA-UCEC cohort. The gene deletion is shown by the green circle, the gene amplification is represented by the red circle, and the mutation frequency is shown by the height of the column. **(C)** The locations of the CNV of 10 CRGs on 23 human chromosomes. **(D)** The gene expression of 10 CRGs in normal (blue) and tumor tissues (red). The statistical *p*-value is shown by the above asterisk. (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

Three different cuproptosis patterns are mediated by 10 cuproptosis-related genes

By carrying out unsupervised clustering analysis (Supplementary FigureS2), we classified all patients into different cuproptosis patterns according to the expression of the 10 CRGs. Finally, we discovered three different cuproptosis patterns named as CRGs clusters A, B, and C, including 154 cases in CRGs cluster A, 298 cases in CRGs cluster B, and 182 cases in CRGs cluster C (Figure 3A, Supplementary Table S1). The survival analysis of three cuproptosis patterns showed that CRGs cluster C had the worst survival situation (Figure 3B). After combing the TCGA-UCEC and GEO datasets, the relationship between clinical characteristics and gene expression of three different cuproptosis patterns was analyzed and exhibited in Figure 3C.

The GSVA enrichment analysis was used to investigate the enrichment of biological processes connected with the three types of cuproptosis patterns (Figures 4A,B). We found CRGs cluster A were significantly enriched in the pathways related to metabolism. CRGs cluster C was correlated with cell activity, including cell cycle, DNA replication and RNA degradation (Figures 4A,B). The transcriptome profiles of three cuproptosis patterns were analyzed using principal component analysis (PCA), which revealed substantial differences among the three clusters (Figure 4C). Subsequently, we evaluated the infiltrating immune cells in three different clusters. The ssGSEA analysis presented that CRGs cluster C was hardly enriched in the infiltration of innate immune cells, such as activated $CD8^+$ T cells, activated $CD4^+$ T cells, macrophages, mast cells, monocytes, natural killer cells, plasmacytoid dendritic cells, and T helper cells, etc. (Figure 4D). According to the results of our research, this might be the reason why patients with CRGs cluster C have the worst survival situation.

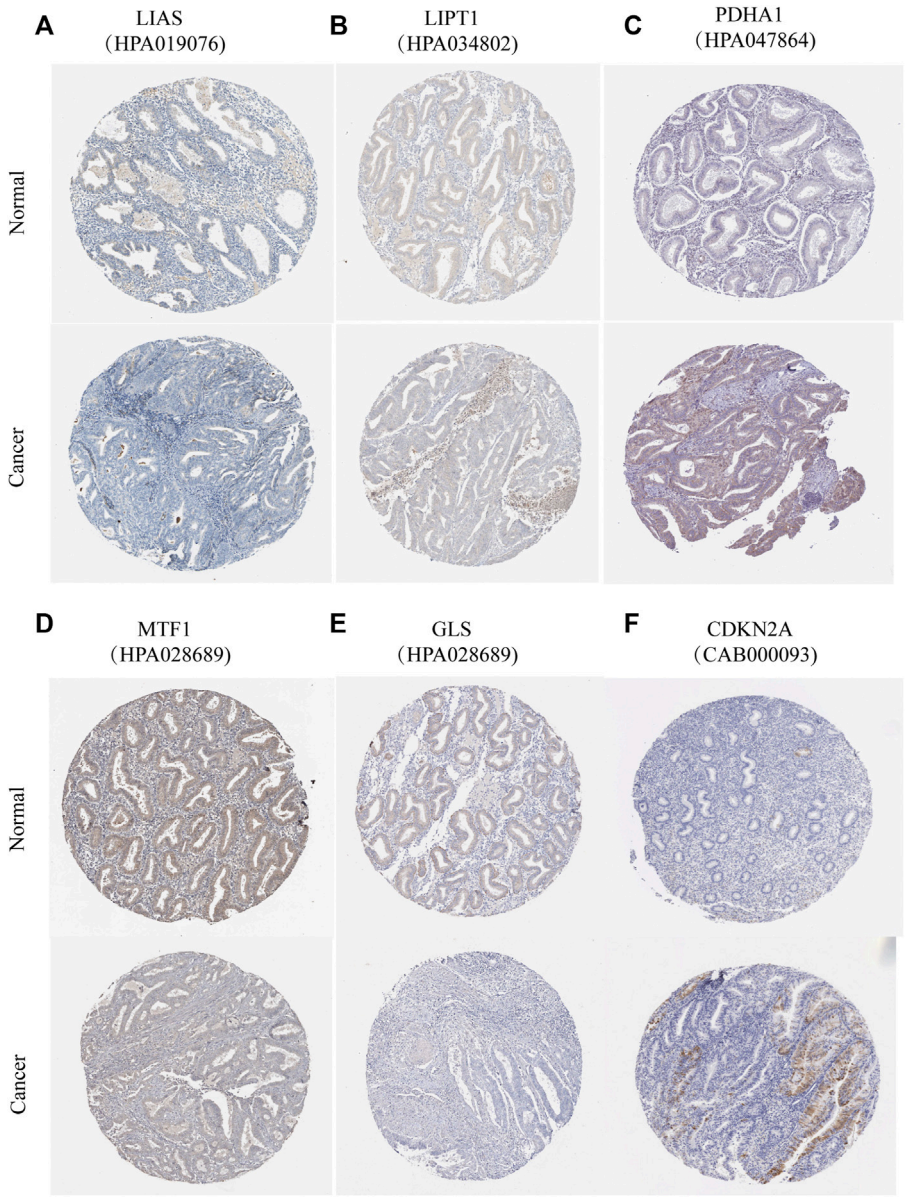


FIGURE 2
Protein expression of six CRGs in the HPA database. Verification of **(A)** LIAS **(B)** LIPT1 **(C)** PDHA1 **(D)** MTF1 **(E)** GLS **(F)** CDKN2A protein expression in normal and tumor tissue utilizing the Human Protein Atlas (HPA) database.

Functional annotation of cuproptosis patterns

In order to further explore the potential biological regulatory pathways in the different cuproptosis patterns, we successfully identified 2,456 common DEGs from TCGA extracted from three distinct clusters associated with the cuproptosis patterns (Figure 5A). By univariate Cox regression analysis for OS, 600 DEGs with significant prognostic differences were selected for further study. The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) function enrichment of these genes was analyzed using the clusterProfiler R package (Supplementary Figure S3). The GO analysis displayed that the DEGs were enriched in the regulation of cell cycle (Supplementary Figures S3A,B). The GO analysis of the biological process (BP) showed that these DEGs were enriched in autophagy, G1/S transition of mitotic cell cycle, and regulation of cell cycle phase transition. The cellular

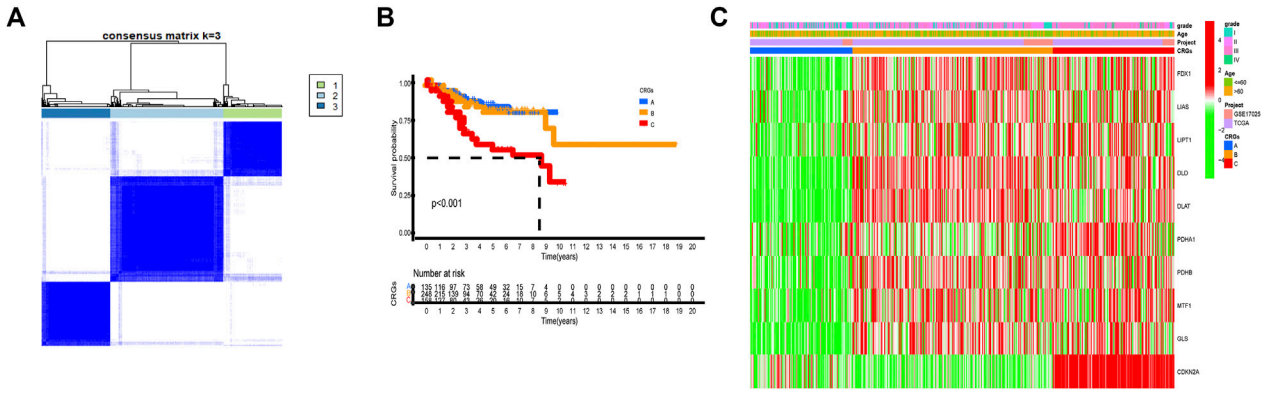


FIGURE 3
Three cuproptosis patterns and their clinical features. **(A)** Consensus clustering matrix of 10 CRGs for k = 3 using unsupervised clustering analysis. **(B)** Kaplan–Meier survival analysis of three cuproptosis patterns in the UCEC. p-value < 0.05 is considered a statistical difference among three cuproptosis patterns. **(C)** The heatmap presented the expression of 10 CRGs among three different cuproptosis patterns by unsupervised clustering in TCGA-UCEC and GSE17025 UCEC patients. The relationship among the CRGs cluster, grade, and age of 10 CRGs were used as notes. Red represented high expression and green represented low expression.

component (CC) exhibited that DEGs were abundant in the chromosomal region, mitochondrial matrix, and spindle. Molecular function (MF) indicated that DEGs mainly participated in catalytic activity, helicase activity, and ATPase activity. The KEGG analysis illustrated a significant enrichment of the cell cycle, protein processing in endoplasmic reticulum, autophagy, and ubiquitin mediated proteolysis (Supplementary Figures S3C,D).

Identification and analysis of cuproptosis-related genes clusters

To further understand the regulatory mechanism of cuproptosis, we similarly used the limma package to divide all patients into distinct genotypes by unsupervised cluster analysis of the 600 cuproptosis-related DEGs filtered by univariate Cox regression analysis (Supplementary Table S2). We eventually identified three distinct CRGs gene clusters, including 169 cases in gene cluster A, 186 cases in gene cluster B, and 279 cases in gene cluster C (Figure 5B, Supplementary Table S3). Supplementary Figure S4 supported the three distinct CRGs gene clusters, named CRGs gene cluster A, cluster B, and cluster C, respectively. Figure 5C explored the clinical features of the three subtypes, and the three gene clusters were closely associated with CRGs clusters. In the three CRGs gene clusters, there were substantial differences in the expression of CRGs (Figure 5D). Once again, these findings affirmed that three diverse cuproptosis patterns occurred in UCEC. Furthermore, there was an observably bad prognosis in gene cluster B, which had a worse prognosis than CRGs gene clusters A and C (Figure 5E).

Construction of the cuproptosis-related genes score and its correlation of tumor microenvironment and clinical prognosis

In view of the various and complex alterations of CRGs in UCEC patients, we created a model that could quantify and evaluate cuproptosis patterns, namely the CRGs score. As shown in Figure 6A, patients with a low CRGs score had a higher survival rate than those with a high CRGs score. In the meantime, the high CRGs score group had a greater proportion of dead patients. According to the CRGs score, the CRGs score of alive patients was lower than that of dead patients (Figures 6B,C). To better show the features of the CRGs score, we evaluated the correlations between the CRGs score and immune infiltrating cells in TME. Only type 2 T helper cells were positively connected with the CRGs score, whereas other immune cells were negatively correlated (Figure 6E). According to the alluvial diagram (Figure 6D), the majority of the UCEC samples that indicated most patients with CRGs cluster C were grouped into CRGs gene cluster B, which obtained a high CRGs score and had lower survival rates. Furthermore, the Kruskal-Wallis test indicated significant differences in the CRGs score between the three CRGs gene clusters and the CRGs clusters (Figures 6F,G). In both of these studies, the higher CRGs score was found in CRGs cluster C and CRGs gene cluster B. (Figures 6E–G).

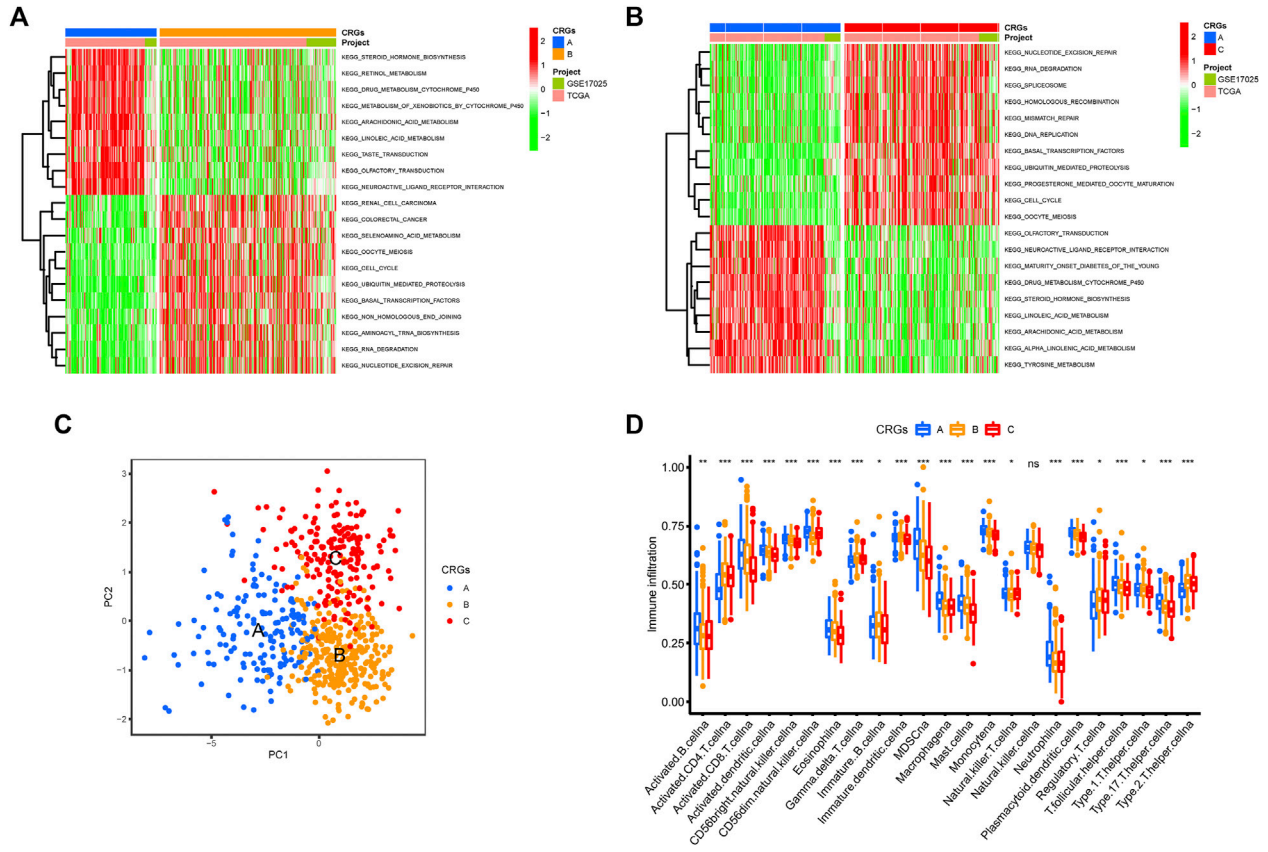


FIGURE 4
Biological pathways based on three clusters and the infiltrating immune cell component in TME. **(A,B)** The activation status of biological pathways in three CRGs clusters as depicted by GSVA enrichment analysis. The heatmap displayed in red represents activated pathways, and green represents inhibited pathways. **(A)** KEGG pathways of CRGs expression clusters A vs. B **(B)** KEGG pathways of CRGs expression clusters A vs. C. **(C)** Principal component analysis (PCA) shows significant differences among the three different cuproptosis patterns. **(D)** The expression abundance of infiltrating immune cells in three cuproptosis patterns. The boxplot showed the expression differences of 23 kinds of immune cells among CRGs cluster A, B, and C. The upper and lower ends of each box mean the quartile range of the value, the middle line represents the median value. The statistical *p*-value is shown by an asterisk (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

Correlation between cuproptosis-related genes score and tumor mutation burden, prognosis, immunotherapy

We analyzed and visualized the differences in somatic mutation distribution profiles of UCEC patients in the high and low CRGs score groups by using the maftools package (Figures 7A,B). The low CRGs score group (98.58%) (Figure 7A) had a slightly higher proportion of somatic mutations than the high CRGs score group (98.04%) (Figure 7B). According to the quantitative study of TMB (Figure 7C), the CRGs score was found to have a negatively linear correlation with TMB. The TMB was higher in the low CRGs score group than in the high CRGs score group (Figure 7D), suggesting that low-risk patients may be more beneficial to immunotherapy. We next focused on the prognostic significance of TMB because of its importance. As seen in the survival plot (Figure 7E), those with high TMB had a greater prognosis benefit than patients with low TMB. Interestingly, merging the CRGs score and TMB may be used as a more comprehensive risk assessment (Figure 7F). In the meantime, we analyzed the prognosis between the CRGs score group and patients with age ≤ 60 and > 60 (Figures 7G,H). Regardless of patients with age ≤ 60 or > 60, the low CRGs score group has a better prognosis than the high CRGs score group (Figures 7G,H).

Cuproptosis patterns in the role of immunotherapy

Figures 8A,B exhibited that patients with a low CRGs score had much greater PD-1 and CTLA-4 expression, indicating a possible great response to anti-PD-1 and anti-CTLA-4 immunotherapy.

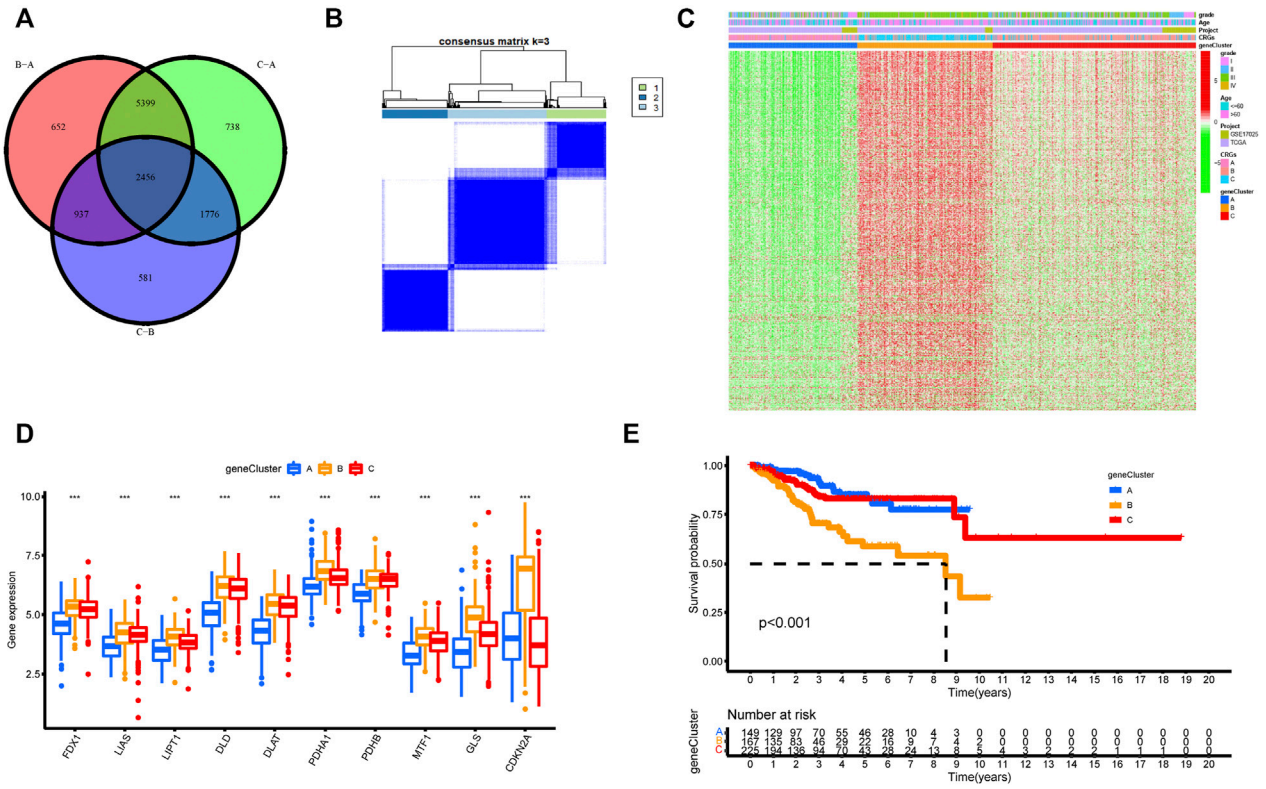


FIGURE 5
Unsupervised clustering of cuproptosis-related genes according to TCGA-UCEC and GSE17025 cohort. **(A)** 2,456 cuproptosis-related genes from TCGA are presented in the Venn diagram. **(B)** Consensus clustering matrix of cuproptosis-related genes for $k = 3$ using unsupervised clustering analysis. **(C)** The unsupervised clustering analysis of the CRGs gene clusters A, B, and C. CRGs clusters, CRGs gene clusters, grade, and age were noted. Red represented high expression, and green represented low expression. **(D)** The histogram illustrated the expression of 10 CRGs in three CRGs gene clusters. The upper and lower ends of each box mean the quartile range of the value, the middle line represents the median value. The statistical p-value is shown by an asterisk (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). **(E)** Kaplan–Meier survival analysis for the three CRGs gene clusters (p-value < 0.05).

Simultaneously, we utilized the CRGs score to predict how people will react to the effectiveness of anti-CTLA-4 and anti-PD-1 immunotherapy. Patients with low CRGs scores (Figures 8C–F) had considerably better anti-CTLA4 and anti-PD1 treatment responses. In addition, IPS was obviously higher in the low CRGs score group than in the high CRGs score group for the combination of anti-PD1 and anti-CTLA-4 (Figure 8C). In general, these findings consistently show that the low CRGs score group has better immunotherapeutic effectiveness than the high CRGs score group. The findings also showed that the cuproptosis pattern is strongly associated with TMB and PD-1/CTLA-4 immunotherapy.

Discussion

In recent years, a large number of studies (Singh et al., 2020; Cui et al., 2021; Guo et al., 2021; Ge et al., 2022) have found that copper plays a crucial role in the occurrence and development of human tumors. Copper can impact essential cellular processes by acting as both a negative allosteric regulator and a positive allosteric regulator of enzyme function (Ge et al., 2022). Cell proliferation and autophagy pathways are all affected, which eventually leads to tumor initiation and development (Ge et al., 2022). Therefore, inhibitors and regulators of cuproptosis-related genes have been investigated as potential methods of cancer therapeutics (Skrott et al., 2017; Yang et al., 2022). However, CRGs have been poorly studied in endometrial cancer and its association with immune infiltration in TME.

Simultaneously, according to the traditional histopathological subtypes, *PTEN* inactivation is a major driver of endometrioid carcinomas; *TP53* inactivation is a major driver of most serous carcinomas, some high-grade endometrioid carcinomas, and many uterine carcinosarcomas; and inactivation of either gene is a major driver of some clear cell carcinomas (Bell and Ellenson, 2019; Urick and Bell, 2019). Integrated genomic analysis by TCGA resulted in the molecular classification of endometrioid and serous carcinomas into four distinct subgroups: POLE

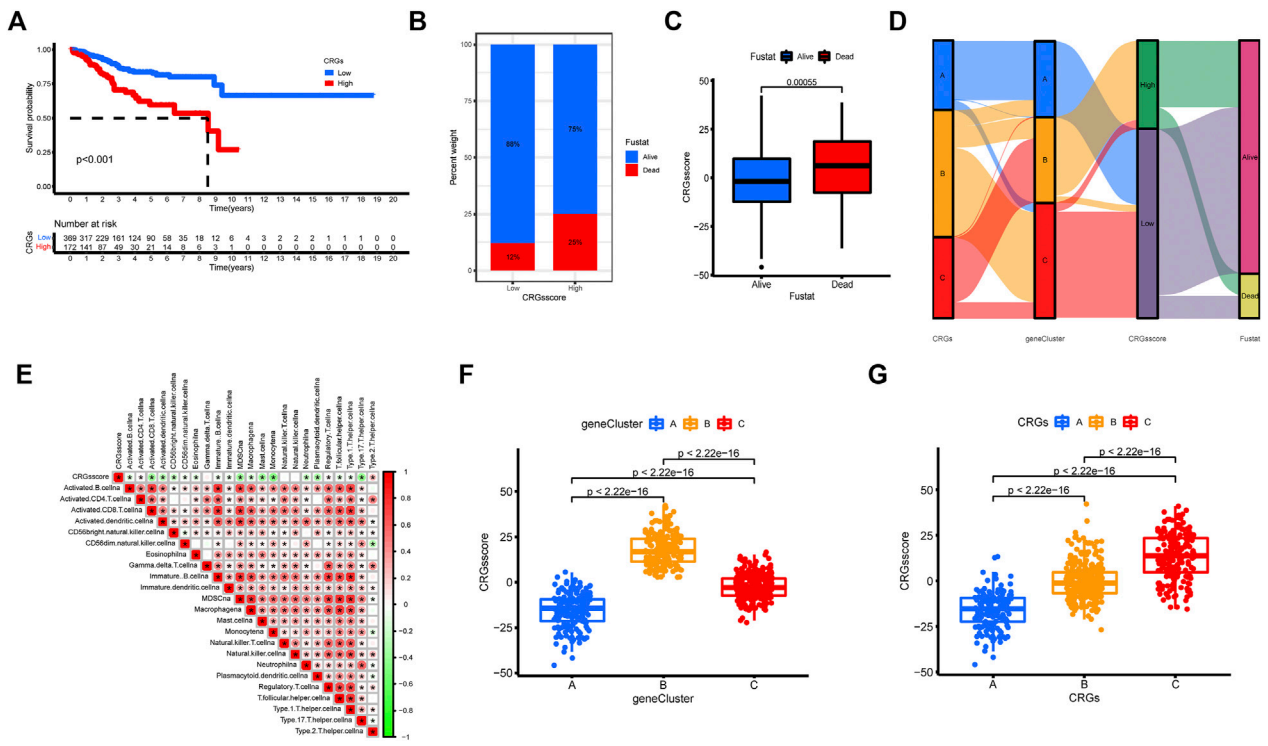


FIGURE 6
Characteristics of CRGs score. **(A)** Kaplan–Meier survival analyses for the OS in high and low CRGs score groups. **(B)** The percentage of patients in low and high CRGs score groups with various survival statuses. **(C)** The variation of CRGs score between the alive and dead group. **(D)** The alluvial diagram showed the relationship between CRGs cluster, gene cluster, CRGs score, and survival status. **(E)** Spearman analysis of the correlations between CRGs score and infiltrating immune cells. Red means the positive correlation and green means the negative correlation. The asterisk represents *p*-value that is meaningful statistically. **(F,G)** Differential analysis of CRGs score among **(F)** three CRGs gene clusters and **(G)** three CRGs clusters ($p < 0.001$).

(ultramutated), microsatellite instability (hypermutated), copy number low (endometrioid), and copy number high (serous-like) (Bell and Ellenson, 2019).

However, these existing classifications based on these molecules did not better improve the prognosis of UCEC. Molecular patterns based on molecular pathology can represent deeper properties of tumors and, as a result, compensate for the shortcomings of the traditional histopathological subtypes. The major purpose of our study is to identify three cuproptosis-related molecular patterns, construct the CRGs score based on 10 CRGs and investigate its correlation with the tumor immune microenvironment in UCEC. Moreover, our research found a connection between the CRGs score and immunotherapy, especially anti-PD-1/PDL1 and CTLA-4 immunotherapy.

First, we analyzed transcriptional sequencing data from GEO and TCGA-UCEC to identify the cuproptosis mode in UCEC. In our study, we successfully divided all samples into three cuproptosis-related molecular patterns based on 10 CRGs in UCEC. The common DEGs among the three molecular patterns were then discovered, and common DEGs filtered by univariate Cox regression analysis were then utilized to build the CRGs score. The KM survival curve illustrated a better OS and better prognosis associated with a low CRGs score than with a high CRGs score. The CRGs score might be used to identify different immunological phenotypes and assess the immunotherapy effects of UCEC patients.

The TME is associated with regulating cancer as well as a source of immune therapeutic targets. So, we explored the compositions of infiltrating cells in TME that were different in three patterns. Extensive articles (Tan et al., 2018; Hessmann et al., 2020; Jin and Jin, 2020; Bejarano et al., 2021; Singleton et al., 2021) reported that the TME plays a significant role in cancer development and might be a therapeutic target as well as a regulator of cancer development. We focused on the role of the CRGs of TME, with the goal of revealing its possible roles and contributing to a better understanding of the antitumor immunotherapy effects in UCEC. $CD8^+$ T cells are the major anti-tumor effector cells (Mami-Chouaib et al., 2018), and T cell-mediated antitumor immunity is strengthened by tumor-infiltrating B cells (Mami-Chouaib et al., 2018; Engelhard et al., 2021). Natural killer (NK) cells can swiftly kill adjacent

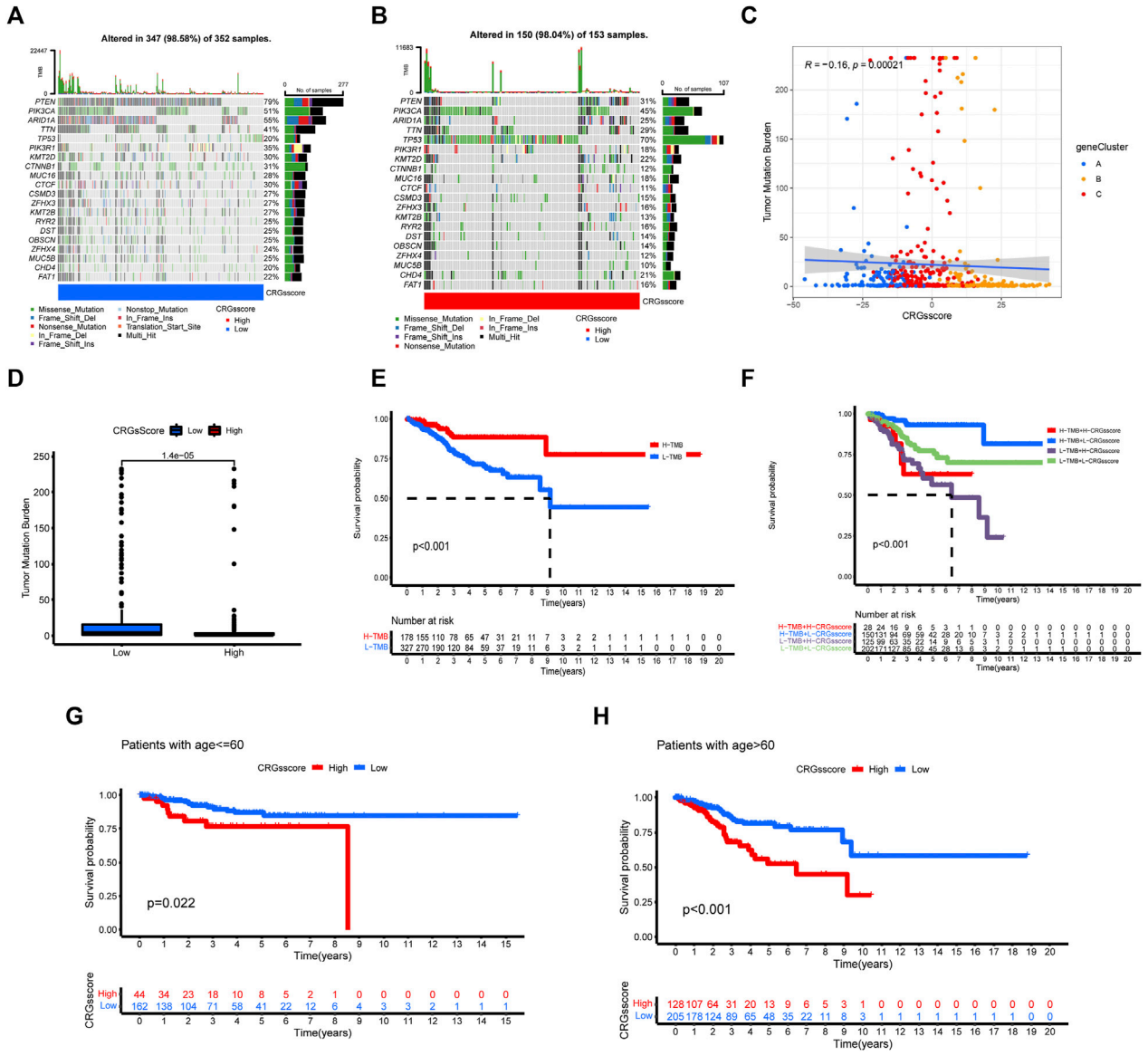


FIGURE 7
The relationship between CRGs score groups and somatic mutation, TMB. The waterfall plot of tumor somatic mutation in low CRGs score **(A)** and high CRGs score **(B)**. **(C,D)** The correlation between the CRGs score and TMB. **(E)** Survival analysis of CRGs score using Kaplan-Meier curves. **(F)** Survival analysis of both TMB and CRGs score using Kaplan-Meier curves. **(G,H)** Survival analysis of CRGs score in patients with age ≤ 60 and >60 using Kaplan-Meier curves.

tumor cells, enhance antibody and T cell response as anticancer agents (Shimasaki et al., 2020). Expansion and genetic modification of NK cells can greatly increase their anti-tumor activity, and NK cells are expected to become essential elements of multipronged therapeutic strategies for cancer (Myers and Miller, 2021). Both innate and adaptive immune responses are mediated by dendritic cells (DCs) that play an important role in the progression and regulation of both innate and adaptive immune responses (Wculek et al., 2020). Due to the lower expression of almost all immune cells in CRGs cluster C, it might be the reason that the CRGs cluster C showed a lower survival rate than other clusters.

Besides, copy number variation (CNV) is one of the most important somatic aberrations in cancer, and it contributes to the pathogenesis of many disease phenotypes (Martin-Trujillo et al., 2017). Based on 10 CRGs and UCEC copy-number profiles, we explored the alteration of CRGs in UCEC. Mutations in the 10 CRGs were relatively uncommon, but CNV amplification and deletion were common events. Due to the specific function of PD-1/PD-L1 blockage immunotherapy in the field of cancer

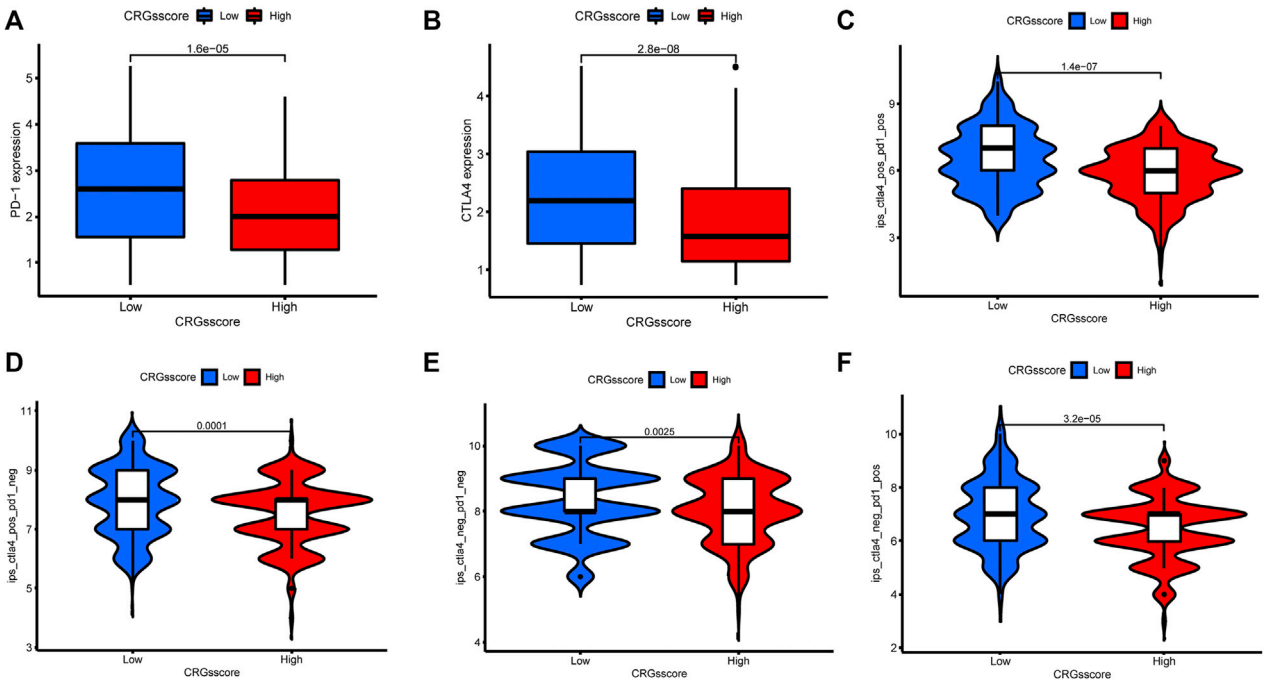


FIGURE 8
The association between CRGs score and immunotherapy. **(A,B)** The different expression of PD-1 and CTLA-4 in low CRGs score and high CRGs score. **(C–F)** Relationship between CRGs score and immunotherapeutic response of PD-1 and CTLA-4 expression: **(C)** positive PD-1 and CTLA-4. **(D)** Positive CTLA-4 and negative PD-1. **(E)** Negative PD-1 and CTLA-4. **(F)** Negative PD-1 and positive CTLA-4.

therapy, PD-1 expression of immune cells was examined independently for PD-1 immune treatment, and the IPS was developed as a marker to distinguish factors in clinical (Kalbasi and Ribas, 2020). While TMB analysis showed the low CRGs score group had a higher TMB than the high CRGs score group, several lines of evidence (Samstein et al., 2019; Marabelle et al., 2020; Valero et al., 2021) found the high TMB is associated with better survival for tumor patients receiving immune checkpoint inhibitors (ICIs). Our CRGs score system revealed that a combination of CTLA-4 and PD-1 blockers might be beneficial for the treatment of UCEC. A similar conclusion was found in some studies (Liu and Zamarin, 2018; Oh and Chae, 2019; Rotte, 2019; Zhang et al., 2021). The CRGs score can also be predicted to determine how tumors will respond to immunotherapy.

Based on the expression of CRGs, we discovered three molecular patterns connected to cuproptosis in the current study. In order to forecast patient survival and TME features, we created the CRGs score for each patient. In the meanwhile, this study demonstrated how CRGs affect the prognosis of UCEC patients. The CRGs score performs well in determining biological state and predicting UCEC survival. However, our research still has many shortcomings and flaws. First, many experiments have not yet been performed since cuproptosis was discovered not long ago. Therefore, there are no more basic experiments to demonstrate the effect of cuproptosis on the prognosis of endometrial cancer. Second, the mechanism by which cuproptosis affects cancer progression has not been revealed. Therefore, more research by well-designed experiments is needed.

Conclusion

This study implies that CRGs are involved in the formation and progression of endometrial cancer. Based on the expression of 10 CRGs, we classified all samples into three cuproptosis patterns. According to 600 cuproptosis-related DEGs filtered by univariate Cox regression analysis, the CRGs score was constructed. The CRGs score was correlated with the tumor microenvironment, immunotherapy response, and prognosis of cancer patients. Through comprehensive analysis of the relationship between the CRGs score and clinicopathological features, we found it may help assess the efficacy of ICIs. In summary, the identification of cuproptosis patterns and related genes will contribute to valuable approaches for individualized therapy for UCEC patients.

Data availability statement

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

Ethics statement

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

JC, GW, JZ, and XL completed data download, data collation, data analysis and results discussion. JC, JZ, and YZ wrote and revised the paper, and all authors read and approved the final version carefully.

Funding

This study was funded by the National Natural Sciences Foundation of China (Grant No. 81871126).

Acknowledgments

We sincerely thank the TCGA, GEO, HPA, and TCIA databases for the availability of the original data.

Conflict of interest

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

Publisher's note

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/fgene.2022.1001374/full#supplementary-material

References

Andersen, M. H., Becker, J. C., and Straten, P. (2005). Regulators of apoptosis: Suitable targets for immune therapy of cancer. *Nat. Rev. Drug Discov.* 4, 399–409. doi:10.1038/nrd1717

Barbie, D. A., Tamayo, P., Boehm, J. S., Kim, S. Y., Moody, S. E., Dunn, I. F., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462, 108–112. doi:10.1038/nature08460

Bejarano, L., Jordão, M. J. C., and Joyce, J. A. (2021). Therapeutic targeting of the tumor microenvironment. *Cancer Discov.* 11, 933–959. doi:10.1158/2159-8290.cd-20-1808

Bell, D. W., and Ellenson, L. H. (2019). Molecular genetics of endometrial carcinoma. *Annu. Rev. Pathol.* 14, 339–367. doi:10.1146/annurev-pathol-020117-043609

Brooks, R. A., Fleming, G. F., Lastra, R. R., Lee, N. K., Moroney, J. W., Son, C. H., et al. (2019). Current recommendations and recent progress in endometrial cancer. *Ca. Cancer J. Clin.* 69, 258–279. doi:10.3322/caac.21561

Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., et al. (2017). Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 18, 248–262. doi:10.1016/j.celrep.2016.12.019

Chen, X., Kang, R., Kroemer, G., and Tang, D. (2021). Broadening horizons: The role of ferroptosis in cancer. *Nat. Rev. Clin. Oncol.* 18, 280–296. doi:10.1038/s41571-020-00462-0

Cui, L., Gouw, A. M., LaGory, E. L., Guo, S., Attarwala, N., Tang, Y., et al. (2021). Mitochondrial copper depletion suppresses triple-negative breast cancer in mice. *Nat. Biotechnol.* 39, 357–367. doi:10.1038/s41587-020-0707-9

Dannappel, M., Vlantis, K., Kumari, S., Polykratis, A., Kim, C., Wachsmuth, L., et al. (2014). RIPK1 maintains epithelial homeostasis by inhibiting apoptosis and necroptosis. *Nature* 513, 90–94. doi:10.1038/nature13608

Engelhard, V., Conejo-Garcia, J. R., Ahmed, R., Nelson, B. H., Willard-Gallo, K., Bruno, T. C., et al. (2021). B cells and cancer. *Cancer Cell* 39, 1293–1296. doi:10.1016/j.ccell.2021.09.007

Garber, K. (2015). BIOMEDICINE. Targeting copper to treat breast cancer. *Sci. (New York, N.Y.)* 349, 128–129. doi:10.1126/science.349.6244.128

Ge, E. J., Bush, A. I., Casini, A., Cobine, P. A., Cross, J. R., DeNicola, G. M., et al. (2022). Connecting copper and cancer: From transition metal signalling to metalloplasia. *Nat. Rev. Cancer* 22, 102–113. doi:10.1038/s41568-021-00417-2

Goliwas, K. F., Deshane, J. S., Elmets, C. A., and Athar, M. (2021). Moving immune therapy forward targeting TME. *Physiol. Rev.* 101, 417–425. doi:10.1152/physrev.00008.2020

Guo, J., Cheng, J., Zheng, N., Zhang, X., Dai, X., Zhang, L., et al. (2021). Copper promotes tumorigenesis by activating the PDK1-AKT oncogenic pathway in a copper transporter 1 dependent manner. *Adv. Sci. (Weinheim, Baden-Wurttemberg, Ger.* 8, e2004303. doi:10.1002/advs.202004303

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

Hessmann, E., Buchholz, S. M., Demir, I. E., Singh, S. K., Gress, T. M., Ellenrieder, V., et al. (2020). Microenvironmental determinants of pancreatic cancer. *Physiol. Rev.* 100, 1707–1751. doi:10.1152/physrev.00042.2019

Jiang, X., Stockwell, B. R., and Conrad, M. (2021). Ferroptosis: Mechanisms, biology and role in disease. *Nat. Rev. Mol. Cell Biol.* 22, 266–282. doi:10.1038/s41580-020-00324-8

Jin, M. Z., and Jin, W. L. (2020). The updated landscape of tumor microenvironment and drug repurposing. *Signal Transduct. Target. Ther.* 5, 166. doi:10.1038/s41392-020-00280-x

Kalbasi, A., and Ribas, A. (2020). Tumour-intrinsic resistance to immune checkpoint blockade. *Nat. Rev. Immunol.* 20, 25–39. doi:10.1038/s41577-019-0218-4

Karki, R., and Kanneganti, T. D. (2019). Diverging inflammasome signals in tumorigenesis and potential targeting. *Nat. Rev. Cancer* 19, 197–214. doi:10.1038/s41568-019-0123-y

Lei, G., Zhuang, L., and Gan, B. (2022). *Targeting ferroptosis as a vulnerability in cancer*. Houston, TX, United States: Nature reviews Cancer. doi:10.1038/s41568-022-00459-0

Li, H., Chan, L., Bartuzi, P., Melton, S. D., Weber, A., Ben-Shlomo, S., et al. (2014). Copper metabolism domain-containing 1 represses genes that promote inflammation and protects mice from colitis and colitis-associated cancer. *Gastroenterology* 147, 184–195. e3. doi:10.1053/j.gastro.2014.04.007

Li, X., Wenes, M., Romero, P., Huang, S. C., Fendt, S. M., and Ho, P. C. (2019). Navigating metabolic pathways to enhance antitumour immunity and immunotherapy. *Nat. Rev. Clin. Oncol.* 16, 425–441. doi:10.1038/s41571-019-0203-7

Liu, Y. L., Bager, C. L., Willumsen, N., Ramchandani, D., Kornhauser, N., Ling, L., et al. (2021). Tetrathiomolybdate (TM)-associated copper depletion influences collagen remodeling and immune response in the pre-metastatic niche of breast cancer. *NPJ breast cancer* 7, 108. doi:10.1038/s41523-021-00313-w

Liu, Y. L., and Zamarin, D. (2018). Combination immune checkpoint blockade strategies to maximize immune response in gynecological cancers. *Curr. Oncol. Rep.* 20, 94. doi:10.1007/s11912-018-0740-8

Mami-Chouaib, F., Blanc, C., Corgnac, S., Hans, S., Malenica, I., Granier, C., et al. (2018). Resident memory T cells, critical components in tumor immunology. *J. Immunother. Cancer* 6, 87. doi:10.1186/s40425-018-0399-6

Marabelle, A., Fakih, M., Lopez, J., Shah, M., Shapira-Frommer, R., Nakagawa, K., et al. (2020). Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: Prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. *Lancet. Oncol.* 21, 1353–1365. doi:10.1016/S1470-2045(20)30445-9

Martin, J. D., Cabral, H., Stylianopoulos, T., and Jain, R. K. (2020). Improving cancer immunotherapy using nanomedicines: Progress, opportunities and challenges. *Nat. Rev. Clin. Oncol.* 17, 251–266. doi:10.1038/s41571-019-0308-z

Martin-Trujillo, A., Vidal, E., Monteagudo-Sánchez, A., Sanchez-Delgado, M., Moran, S., Hernandez Mora, J. R., et al. (2017). Copy number rather than epigenetic alterations are the major dictator of imprinted methylation in tumors. *Nat. Commun.* 8, 467. doi:10.1038/s41467-017-00639-9

Myers, J. A., and Miller, J. S. (2021). Exploring the NK cell platform for cancer immunotherapy. *Nat. Rev. Clin. Oncol.* 18, 85–100. doi:10.1038/s41571-020-0426-7

O'Donnell, J. S., Teng, M. W. L., and Smyth, M. J. (2019). Cancer immunoediting and resistance to T cell-based immunotherapy. *Nat. Rev. Clin. Oncol.* 16, 151–167. doi:10.1038/s41571-018-0142-8

Oh, M. S., and Chae, Y. K. (2019). Deep and durable response with combination CTLA-4 and PD-1 blockade in mismatch repair (MMR)-proficient endometrial cancer. *J. Immunother.* 42, 51–54. doi:10.1097/CJI.0000000000000244

Proneth, B., and Conrad, M. (2019). Ferroptosis and necroinflammation, a yet poorly explored link. *Cell Death Differ.* 26, 14–24. doi:10.1038/s41418-018-0173-9

Riley, R. S., June, C. H., Langer, R., and Mitchell, M. J. (2019). Delivery technologies for cancer immunotherapy. *Nat. Rev. Drug Discov.* 18, 175–196. doi:10.1038/s41573-018-0006-z

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

Rotte, A. (2019). Combination of CTLA-4 and PD-1 blockers for treatment of cancer. *J. Exp. Clin. Cancer Res.* 38, 255. doi:10.1186/s13046-019-1259-z

Samstein, R. M., Lee, C. H., Shoushtari, A. N., Hellmann, M. D., Shen, R., Janjigian, Y. Y., et al. (2019). Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet.* 51, 202–206. doi:10.1038/s41588-018-0312-8

Shimasaki, N., Jain, A., and Campana, D. (2020). NK cells for cancer immunotherapy. *Nat. Rev. Drug Discov.* 19, 200–218. doi:10.1038/s41573-019-0052-1

Singh, R. P., Jeyaraju, D. V., Voisin, V., Hurren, R., Xu, C., Hawley, J. R., et al. (2020). Disrupting mitochondrial copper distribution inhibits leukemic stem cell self-renewal. *Cell stem Cell* 26, 926–937. e10. doi:10.1016/j.stem.2020.04.010

Singleton, D. C., Macann, A., and Wilson, W. R. (2021). Therapeutic targeting of the hypoxic tumour microenvironment. *Nat. Rev. Clin. Oncol.* 18, 751–772. doi:10.1038/s41571-021-00539-4

Skrott, Z., Mistrik, M., Andersen, K. K., Friis, S., Majera, D., Gursky, J., et al. (2017). Alcohol-abuse drug disulfiram targets cancer via p97 segregase adaptor NPL4. *Nature* 552, 194–199. doi:10.1038/nature25016

Tan, H. Y., Wang, N., Lam, W., Guo, W., Feng, Y., and Cheng, Y. C. (2018). Targeting tumour microenvironment by tyrosine kinase inhibitor. *Mol. Cancer* 17, 43. doi:10.1186/s12943-018-0800-6

Tang, D., Chen, X., Kang, R., and Kroemer, G. (2021). Ferroptosis: Molecular mechanisms and health implications. *Cell Res.* 31, 107–125. doi:10.1038/s41422-020-00441-1

Tang, D., Chen, X., and Kroemer, G. (2022). Cuproptosis: A copper-triggered modality of mitochondrial cell death. *Cell Res.* 32 (5), 417–418. doi:10.1038/s41422-022-00653-7

Tsvetkov, P., Coy, S., Petrova, B., Dreishpoon, M., Verma, A., Abdusamad, M., et al. (2022). Copper induces cell death by targeting lipoylated TCA cycle proteins. *Sci. (New York, N.Y.)* 375, 1254–1261. doi:10.1126/science.abf0529

Urick, M. E., and Bell, D. W. (2019). Clinical actionability of molecular targets in endometrial cancer. *Nat. Rev. Cancer* 19, 510–521. doi:10.1038/s41568-019-0177-x

Valero, C., Lee, M., Hoen, D., Wang, J., Nadeem, Z., Patel, N., et al. (2021). The association between tumor mutational burden and prognosis is dependent on treatment context. *Nat. Genet.* 53, 11–15. doi:10.1038/s41588-020-00752-4

Ventriglia, J., Paciolla, I., Pisano, C., Cecere, S. C., Di Napoli, M., Tambaro, R., et al. (2017). Immunotherapy in ovarian, endometrial and cervical cancer: State of the art and future perspectives. *Cancer Treat. Rev.* 59, 109–116. doi:10.1016/j.ctrv.2017.07.008

Voli, F., Valli, E., Lerra, L., Kimpton, K., Saletta, F., Giorgi, F. M., et al. (2020). Intratumoral copper modulates PD-L1 expression and influences tumor immune evasion. *Cancer Res.* 80, 4129–4144. doi:10.1158/0008-5472.CAN-20-0471

Wculek, S. K., Cueto, F. J., Mujal, A. M., Melero, I., Krummel, M. F., and Sancho, D. (2020). Dendritic cells in cancer immunology and immunotherapy. *Nat. Rev. Immunol.* 20, 7–24. doi:10.1038/s41577-019-0210-z

Yang, M., Wu, X., Hu, J., Wang, Y., Wang, Y., Zhang, L., et al. (2022). COMMD10 inhibits HIF1α/CP loop to enhance ferroptosis and radiosensitivity by disrupting Cu-Fe balance in hepatocellular carcinoma. *J. hepatology* 76, 1138–1150. doi:10.1016/j.jhep.2022.01.009

Zahn, K. E., Jensen, R. B., Wood, R. D., and Doublié, S. (2021). Human DNA polymerase θ harbors DNA end-trimming activity critical for DNA repair. *Mol. Cell* 81, 1534–1547.e4. e4. doi:10.1016/j.molcel.2021.01.021

Zhang, M. K., Ye, J. J., Li, C. X., Xia, Y., Wang, Z. Y., Feng, J., et al. (2019). Cytomembrane-mediated transport of metal ions with biological specificity. *Adv. Sci.* 6, 1900835. doi:10.1002/advs.201900835

Zhang, P., Liu, M., Cui, Y., Zheng, P., and Liu, Y. (2021). Microsatellite instability status differentially associates with intratumoral immune microenvironment in human cancers. *Brief. Bioinform.* 22, bbaa180. doi:10.1093/bib/bbaa180

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 05 October 2022
DOI 10.3389/fgene.2022.992343

Check for updates

OPEN ACCESS

EDITED BY
Ehsan Nazemalhosseini-Mojarad,
Shahid Beheshti University of Medical Sciences, Iran

REVIEWED BY
Xuefeng Li,
Guangzhou Medical University, China
Bo Zhu,
Guangxi Medical University, China

*CORRESPONDENCE
Zhe Wang,
wangzhe@dlu.edu.cn
Ruoyu Wang,
wangruoyu@dlu.edu.cn

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 12 July 2022
ACCEPTED 20 September 2022
PUBLISHED 05 October 2022

CITATION
Liu J, Liu T, Zhang C, He J, Zhou D, Wang Z and Wang R (2022), EIF2S2 is a novel independent prognostic biomarker and correlated with immune infiltrates in hepatocellular carcinoma. *Front. Genet.* 13:992343. doi: 10.3389/fgene.2022.992343

COPYRIGHT
© 2022 Liu, Liu, Zhang, He, Zhou, Wang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

EIF2S2 is a novel independent prognostic biomarker and correlated with immune infiltrates in hepatocellular carcinoma

Jing Liu[1,2], Tongyu Liu[3], Chuanhao Zhang[1,2], Jiabei He[1,2], Dong Zhou[1,2], Zhe Wang[1,2]* and Ruoyu Wang[1,2]*

[1]Department of Medical Oncology, Affiliated Zhongshan Hospital of Dalian University, Dalian, China, [2]The Key Laboratory of Biomarker High Throughput Screening and Target Translation of Breast and Gastrointestinal Tumor, Dalian University, Dalian, China, [3]Department of Gynecology, Fujian Cancer Hospital, Fujian Medical University Cancer Hospital, Fuzhou, China

Background: Hepatocellular carcinoma (HCC) is a highly malignant disease with poor prognosis. It is urgent to find effective biomarkers. Eukaryotic Translation Initiation Factor 2 Subunit Beta (EIF2S2) is a subunit of heterotrimeric G protein EIF2, and its function is still unclear. We studied the role of EIF2S2 in the malignant progression of liver cancer and its relationship with immune infiltration.

Methods: Download the RNA expression and clinical information of EIF2S2 from the Cancer Genome Atlas (TCGA) database, analyze the relationship between the expression of EIF2S2 and the prognosis and clinicopathological characteristics of HCC, analyze the differential genes by Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and tumor related immune infiltrating cells. The Protein expression level of EIF2S2 was obtained from Human Protein Atlas (HPA) databases. The relationship between EIF2S2 expression and immune infiltrates in HCC was analyzed on TIMER 2.0. The data processing analysis based on R language. Drug Sensitivity data from Genomics of Drug Sensitivity in Cancer (GDSC).

Results: EIF2S2 is highly expressed in HCC patients and is associated with poor prognosis. The expression of EIF2S2 was also correlated with age, clinical stage and pathological grade. Univariate and multivariate COX regression analysis showed that EIF2S2 was an independent risk factor for survival. The receiver operating characteristic (ROC) curve of EIF2S2 also confirmed the diagnostic value of EIF2S2 in HCC patients. Through GO and KEGG enrichment analysis, EIF2S2 expression was found to be closely related to some immune pathways. The expression of EIF2S2 was correlated with memory B cell, plasma B cell, CD8+ T cell, CD4+ resting memory T cell and the expression of some immune checkpoints, such as PDCD1, TIGIT and CTLA-4. It is also more sensitive to paclitaxel, sunitinib and other drugs.

Conclusion: This study shows that EIF2S2 can be used as a prognostic factor for HCC, which is closely related to immune infiltration and immune checkpoints, and may play a potential regulatory role in predicting drug sensitivity.

KEYWORDS

EIF2S2, HCC, prognosis, immune infiltration cells, drug sensitivity

Introduction

HCC causes more than 700,000 deaths worldwide every year, and has become the sixth most common tumor in the world. The incidence of HCC varies from region to region, with the highest incidence in Asia, followed by Europe and Africa. And HCC is also the fourth leading cause of cancer death after lung, colorectal and gastric cancer (Singal et al., 2020). Clinically, the overall incidence of HCC is heterogeneous, which may be due to differences in hepatitis virus prevalence and environmental factors. Around the world, the incidence of HCC in men is two to three times that in women, and the difference is even greater in high-incidence areas (Samant et al., 2021). Chronic hepatitis B virus infection is a major risk factor for the development of cirrhosis and HCC, while other important risk factors include hepatitis C virus infection, alcoholic liver disease, and non-alcoholic fatty liver disease (Ganne-Carrie and Nahon, 2019; Villanueva, 2019; Younossi and Henry, 2021).

Despite various treatment options for HCC, including surgery, interventional therapy, immunotherapy and targeted therapy, the 5-year overall survival rate of patients with HCC remains unsatisfactory, mainly because of the high tumor recurrence and metastasis rates (Kanno et al., 2021; Llovet et al., 2021). Due to the lack of early detection strategies and effective treatment, the 5-year survival rate of HCC is still as low as < 12% (Hlady et al., 2019).

Therefore, it is of great importance to discover new molecules related to the progression of HCC, identify new diagnostic markers and therapeutic targets, accurately select immune checkpoint inhibitors in the population, find the beneficiaries of immunotherapy for improving the prognosis of HCC patients.

EIF2S2 is a subunit of EIF2, a heterotrimeric G protein composed of α, β and γ subunits. Eukaryotic cells restrict protein synthesis under various stress conditions by inhibiting EIF2S2 (Kashiwagi et al., 2016). Previous studies have shown that EIF2S2 deletion reduces the incidence of human testicular germ cells in a mouse model of testicular germ cell tumor development, which also suggests that EIF2S2 is involved in cell proliferation and differentiation (Heaney et al., 2009). Studies have shown that the PI3K/Akt/GSK-3β/ROS/EIF2S2 pathway could regulate NK cell activity and tumor cell sensitivity to NK cells, leading to breast cancer growth and lung metastasis (Jin et al., 2019). It has also been shown that EIF2S2 is highly expressed in gastrointestinal cancers and can promote cell proliferation, migration and colorectal cancer invasion (Zhang et al., 2020). At the same time, EIF2S2 gene is also found to be highly expressed in lung cancer, which can predict the prognosis of lung adenocarcinoma patients (Tanaka et al., 2018). All of the above indicate that EIF2S2 plays an important role in the occurrence, development and metastasis of cancers. However, there is no report on the correlation between EIF2S2 and HCC prognosis and immune infiltration.

In this study, we found that EIF2S2 was highly expressed in HCC, and high expression was associated with poor prognosis and clinicopathological features. We also found that the expression of EIF2S2 was related to a variety of immune infiltrating cells such as T cell CD8+, T cell CD4+ memory resting and a variety of immune checkpoint, such as PDCD1, TIGIT, CTLA4. And it is highly sensitive to paclitaxel, sunitinib and other drugs. Our study shows that EIF2S2 can be used as a prognostic marker in HCC and is associated with immune infiltration. This is the first study of the expression of EIF2S2 in HCC, which may help clinically discover and understand the related processes and potential treatments of EIF2S2 expression in HCC target.

Materials and methods

Data acquisition and processing

The liver hepatocellular carcinoma mRNA expression data and corresponding clinical information were obtained from the TCGA database (https://portal.gdc.cancer.gov/), which was used to evaluate the expression of EIF2S2 in 374 liver hepatocellular carcinoma samples and 50 adjacent normal samples. HCC patients were divided into high expression group and low expression group according to the median value of EIF2S2 expression. Data were collected and analyzed using R 4.1.3 software. And the expression data of EIF2S2 in different types of tumors were obtained from the TIMER2.0 database (http://timer.comp-genomics.org/) (Li et al., 2020). The Human Protein Atlas (HPA) database (https://www.proteinatlas.org/) is a public database that downloads the protein expression levels of EIF2S2 in normal and HCC tissues from HPA.

Survival analysis and clinicopathological analysis

Survival data were statistically analyzed by the survivor R software package, and 374 HCC samples were visualized by the "survminer" R software package. EIF2S2 mRNA expression and its correlation with overall survival (OS) and progression-free survival (PFS) in patients with HCC were analyzed using the TCGA–LIHC dataset. The time-dependent receiver operating characteristic (ROC) curve was drawn using the "timeROC" package (Robin et al., 2011) of R software to evaluate the specificity and sensitivity of EIF2S2 expression for the prognosis assessment of HCC. In addition, the correlation between EIF2S2 expression and clinicopathological factors was also analyzed.

Independent prognostic analysis

Univariate and Multivariate Cox risk regression analysis of EIF2S2 were performed to confirm whether EIF2S2 and clinicopathological parameters were independent factors associated with HCC.

Eukaryotic translation initiation factor 2 subunit beta co-expression analysis

The RNA data obtained from the TCGA database were used to screen out differentially expressed genes between the high EIF2S2 subgroup and the low EIF2S2 subgroup. Analyze the genes that have a co-expression relationship with EIF2S2, understand the positive and negative regulatory relationship between EIF2S2 and genes, and draw heat map according to differential genes.

Enrichment analysis

Differentially expressed genes were subjected to GO and KEGG analysis, and the "ClusterProfilter" R software package was used for GO and KEGG enrichment analysis (Yu et al., 2012), and significant enrichment pathways were obtained. The GO analysis included cellular composition (CC), molecular function (MF) and biological process (BP).

Immune cell infiltration and immune checkpoint analysis

We used TIMER2.0 to evaluate the correlations between the expression level of EIF2S2 and the infiltration levels of immune cells with all algorithms provided, like EPIC, TIMER, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, XCELL, and MCPCOUNTER algorithms. In addition, the correlation between EIF2S2 expression and immune checkpoints was evaluated with Spearman's correlation test.

Drug sensitivity analysis

Based on the largest pharmacogenomics database Genomics of Drug Sensitivity in Cancer (GDSC), Home page - Cancerrxgene - Genomics of Drug Sensitivity in Cancer, we used the R package "pRRophetic" to predict the chemotherapy sensitivity of each tumor sample (Geeleher et al., 2014). Regression methods were used to estimate IC50 for each specific chemotherapeutic drug, predicting the relationship between EIF2S2 and sensitivity to different drugs.

Statistical analysis

The statistical analysis R packages used in each step of the statistical analysis are mentioned above. All data analyses were carried out on R version 4.1.3. Kaplan-Meier curves were used to estimate OS in different groups, and differences between curves were analyzed by log-rank test. Hazard ratios (HR) and 95% confidence intervals (CI) were estimated using univariate and multivariate Cox regression models, and p values < 0.05 were considered statistically significant. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Results

Eukaryotic translation initiation factor 2 subunit beta is highly expressed in hepatocellular carcinoma

The TIMER 2.0 database showed that compared with normal samples, EIF2S2 was overexpressed in most types of cancer, including HCC, colon cancer, lung adenocarcinoma, prostate cancer, endometrial cancer, breast invasive cancer, etc. (Figure 1A). Focusing on HCC, we examined the expression of EIF2S2 in 374 HCC tissues and 50 adjacent normal tissue from the TCGA database. The results showed that the mRNA expression of EIF2S2 was significantly upregulated in HCC tissues compared with adjacent normal tissues ($p < 0.001$) (Figure 1B). Similarly, the expression of EIF2S2 mRNA in cancer samples was also significantly increased in 50 paired cases of HCC tissues and adjacent normal tissues ($p < 0.001$) (Figure 1C). Furthermore, we analyzed immunohistochemical samples from HPA, which confirmed the level of EIF2S2 expression was higher in HCC tissue than in normal samples (Figures 1D,E).

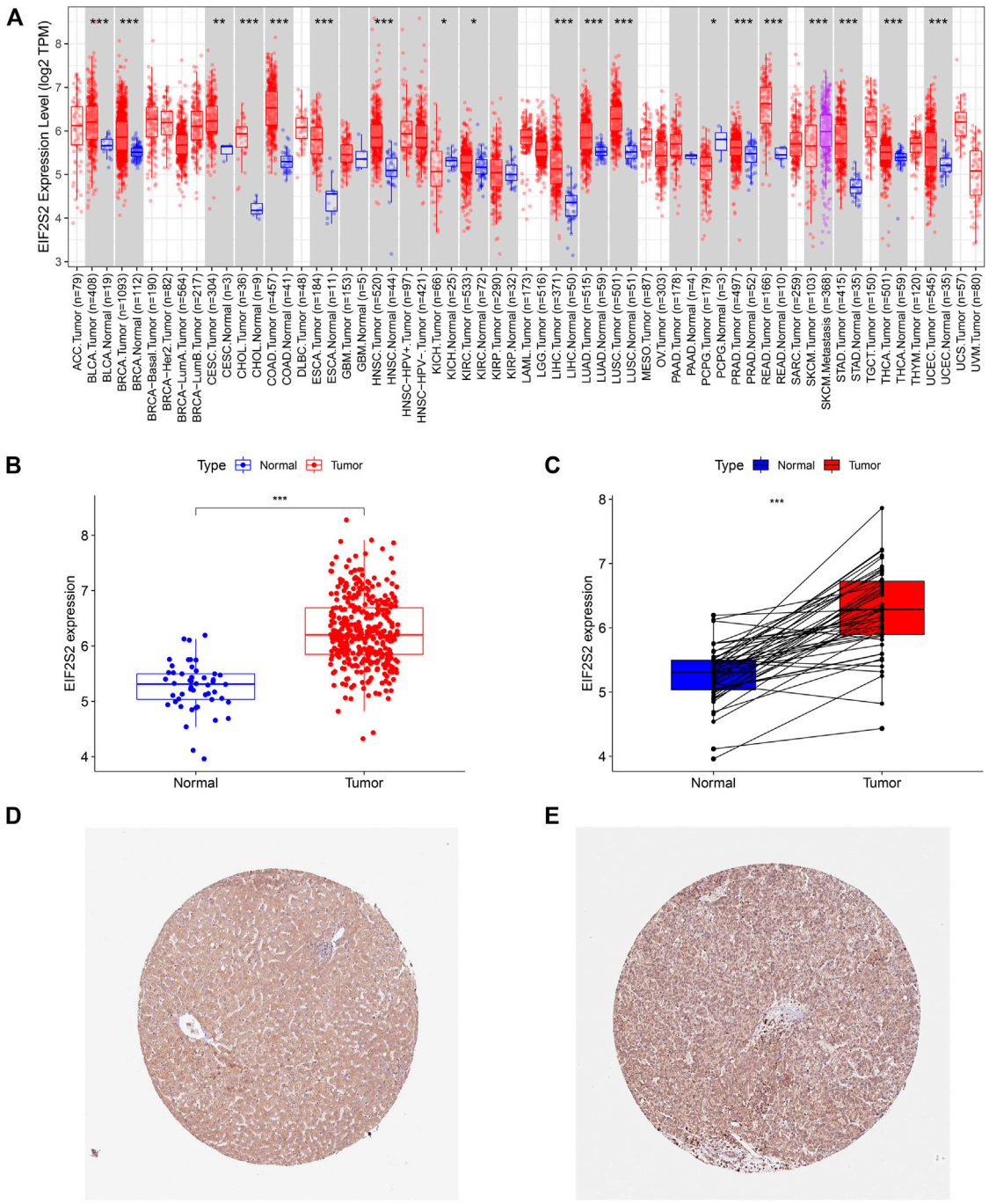


FIGURE 1
The expression of EIF2S2 in HCC. **(A)** EIF2S2 expression levels in different human tumor types according to TIMER 2.0 database. **(B)** EIF2S2 expression was compared in tumor tissue and control tissue samples in the TCGA database. **(C)** EIF2S2 expression levels were compared for paired tumor and control tissue from the TCGA database. **(D,E)** Representative immunohistochemical images of EIF2S2 protein expression in normal and hepatocellular carcinoma tissues. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

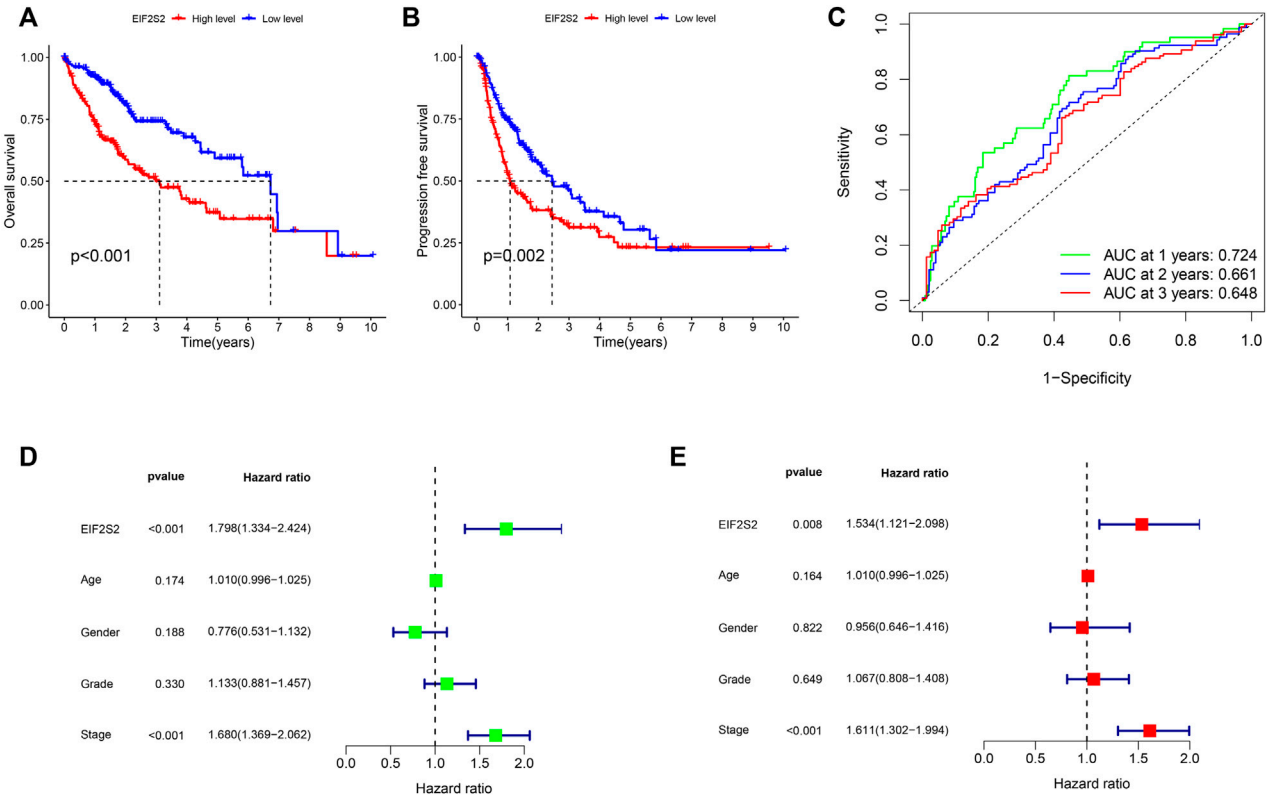


FIGURE 2
Survival and prognosis analysis of EIF2S2. **(A)** High EIF2S2 expression groups had a shorter OS. **(B)** High EIF2S2 expression groups had a shorter PFS. **(C)** AUC of time-dependent ROC curves to evaluate the predictive efficacy of the prognostic signature for OS in HCC patients **(D)** EIF2S2 could act as a risk factor for survival with Multivariate cox regression analysis. **(E)** EIF2S2 could act as a risk factor for survival with Univariate cox regression analysis.

Eukaryotic translation initiation factor 2 subunit beta expression is associated with poor prognosis in hepatocellular carcinoma

To further investigate the prognostic value of EIF2S2 expression in HCC, we used Kaplan-Meier curve to analyze the relationship between EIF2S2 expression and OS and PFS of HCC patients. The results showed that the OS and PFS of HCC patients with high EIF2S2 expression were inferior to those of HCC patients with low EIF2S2 expression (Figures 2A,B). To determine the predictive value of the EIF2S2 expression signature, an ROC curve for 3-year survival was constructed based on the optimal critical risk value. EIF2S2 has good accuracy in predicting the 1-year OS of HCC, and larger samples may be needed to verify 2-year and 3-year OS in the future (Figure 2C). We performed univariate and multivariate independent prognostic analyses of EIF2S2 expression. The forest plot demonstrates that EIF2S2 expression was considered as risk factors affecting the prognosis of HCC patients in univariate and multivariate Cox regression (Figures 2D,E).

Correlation of eukaryotic translation initiation factor 2 subunit beta expression with clinicopathological features

In order to find the correlation between EIF2S2 expression and clinicopathological characteristics of patients with HCC, we analyzed EIF2S2 and clinicopathological characteristics. The results showed that the expression of EIF2S2 was slightly lower in patients over 65 years old (Figure 3A). In pathological grading, the expression of G1 EIF2S2 was significantly different from G3 and G4. Similarly, starting from G2, the expression of EIF2S2 increases with the increase of grading (Figure 3B). There was also a significant difference in EIF2S2 expression between patients with stage I and II, and patients with stage I and III (Figure 3C). In addition, in T

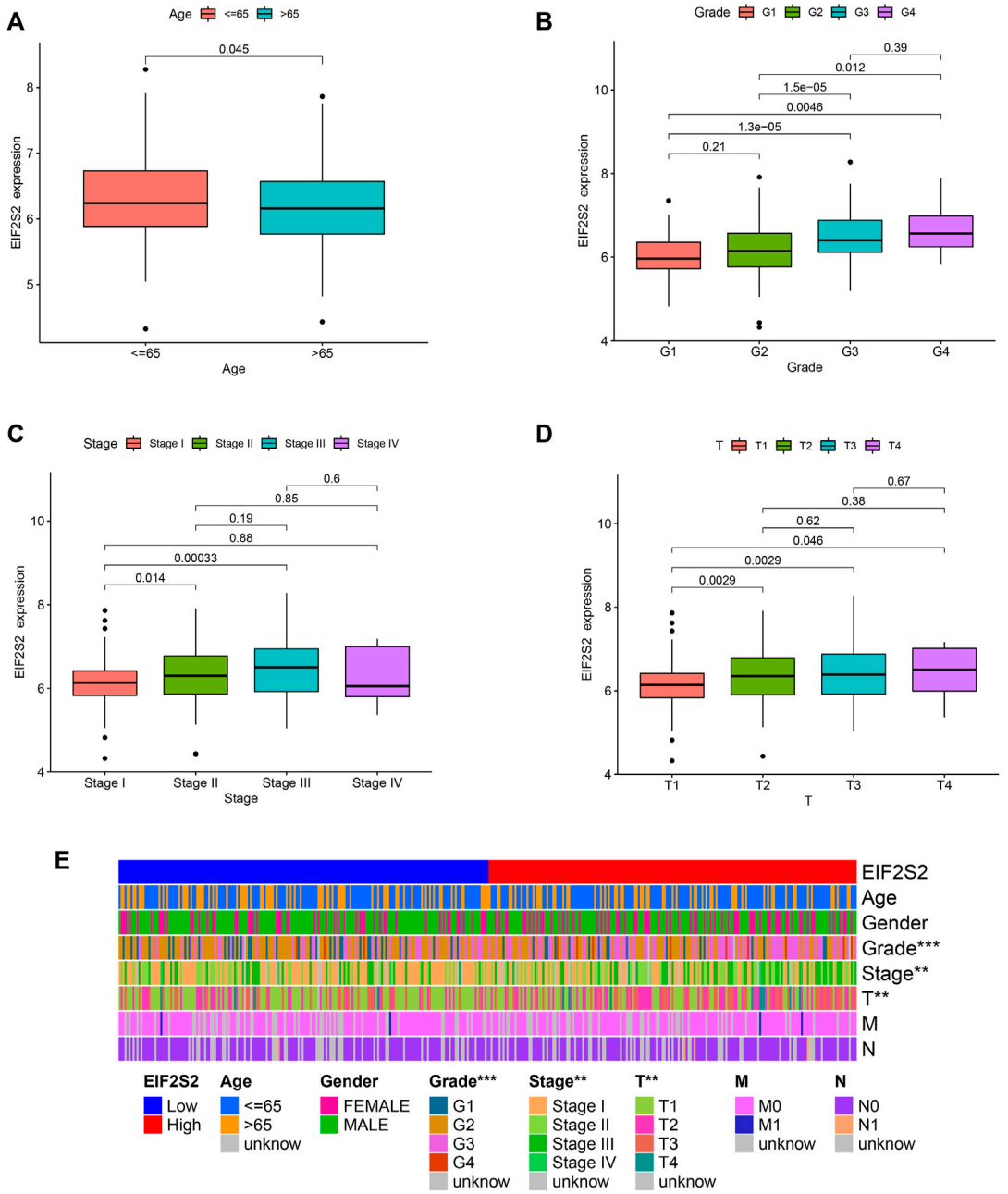


FIGURE 3
Association between the EIF2S2 expression and different clinicopathologic characteristics. **(A)** Association between the EIF2S2 expression and the age of HCC. **(B)** Association between the EIF2S2 expression and the histologic grade of HCC. **(C)** Association between the EIF2S2 expression and the pathologic stage of HCC. **(D)** Association between the EIF2S2 expression and the T stage of HCC. **(E)** Heatmap of the relationship between EIF2S2 expression and clinicopathological analysis. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

stage, the expression of EIF2S2 expression also increased with the increase of T stage (Figure 3D). The heat map of the correlation between EIF2S2 and clinicopathological features also showed the same results (Figure 3E).

Eukaryotic translation initiation factor 2 subunit beta co-expression analysis

We screened out the genes co-expressed with EIF2S2 and drew the co-expression circle map. It can be seen from the

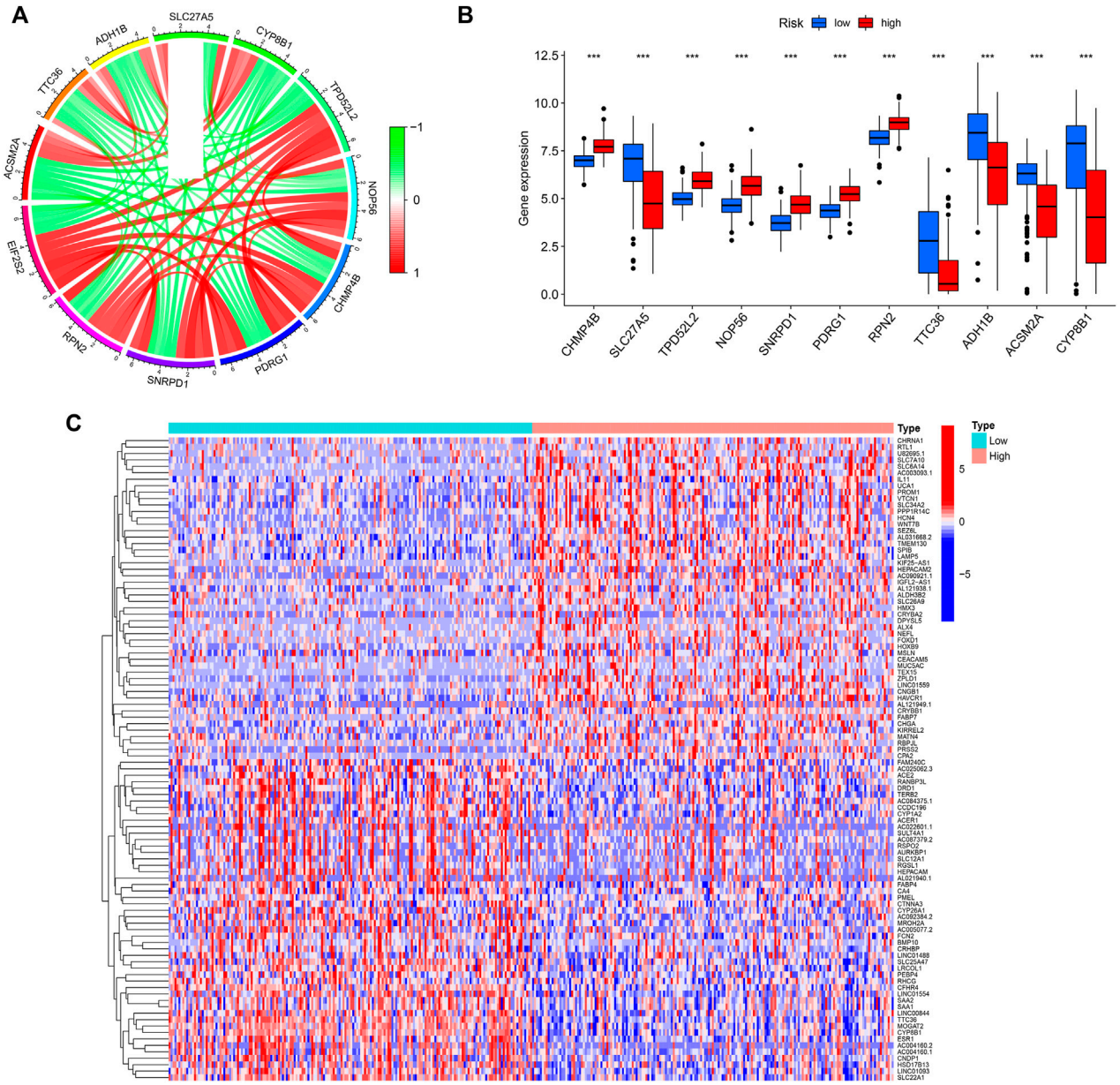


FIGURE 4
Results of co-expressed analysis. **(A)** Co-expressed genes of EIF2S2. **(B)** Differences in the expression of Co-expressed genes in different expression groups. **(C)** Heatmap of Differential genes.

figure that TPD52L2, NOP56, CHMP4B, PDRG1, SNRPD1, RPN2 and EIF2S2 were positively regulated, while ACSM2A, TTC36, ADH1B, SLC27A5, CYP8B1 and the target gene EIF2S2 were negatively regulated (Figure 4A). Next, we analyzed the relationship between the above genes in the EIF2S2 high expression group and the low expression group. The results showed that the expression of ADH1B, CYP8B1, ACSM2, SLC27A, TTC36 in the low expression group was higher than that in the high expression group, while the expression of other genes was higher in the high expression group (Figure 4B). In addition, we performed differential analysis on genes, and found that CHRNA1, RTL1, U82695.1, etc. were upregulated in the EIF2S2 high expression group, and FAM240C, AC025062.3, ACE2 were elevated in the low expression group of EIF2S2 (Figure 4C).

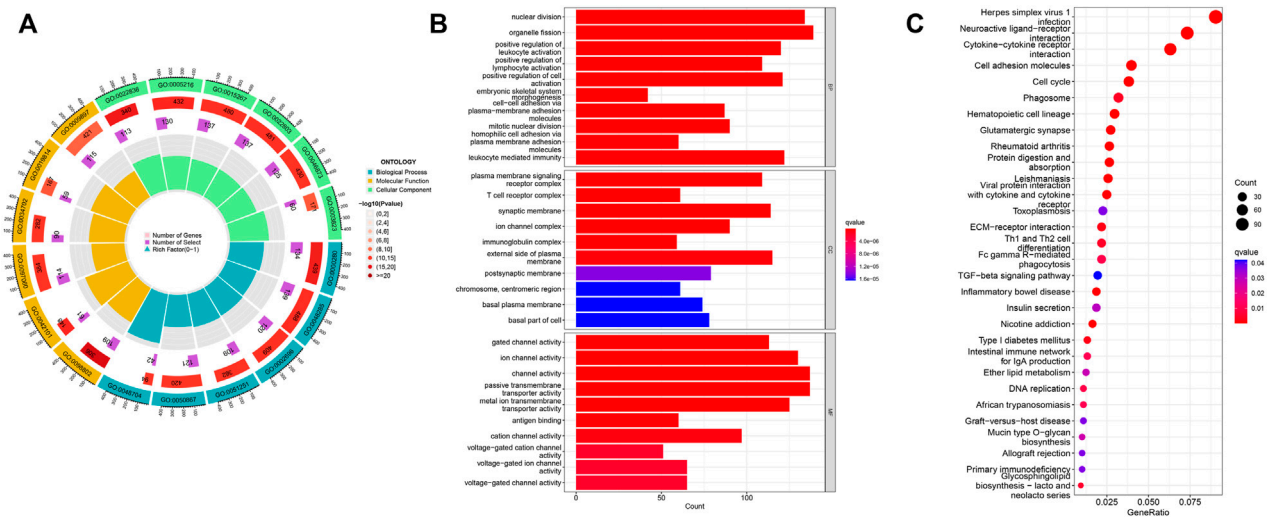


FIGURE 5
Functional enrichment analysis of EIF2S2 in HCC. **(A,B)** GO terms enrichment analysis. **(C)** KEGG pathway enrichment analysis.

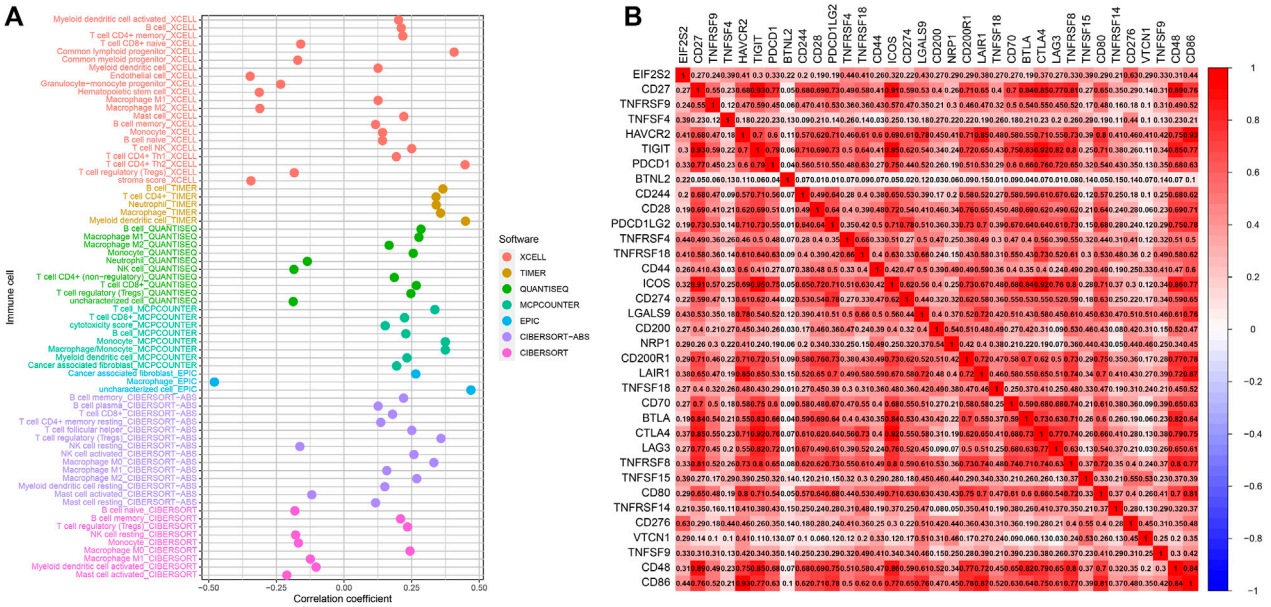


FIGURE 6
Results of correlation between EIF2S2 expression and immunity. **(A)** Correlation between EIF2S2 expression and immune cell infiltration. **(B)** Correlation of EIF2S2 expression and immune checkpoint genes.

Functional enrichment analysis of eukaryotic translation initiation factor 2 subunit beta

To explore EIF2S2 expression and functional enrichment in the TCGA database, we analyzed the differentially expressed EIF2S2-related genes in HCC cases. We performed GO analysis using the "ClusterProfiler" software package and obtained significantly enriched functions and pathways of EIF2S2 expression. The GO enrichment results show that: EIF2S2 was mainly related to organelle fission, nuclear division and leukocyte mediated immunity in the BP

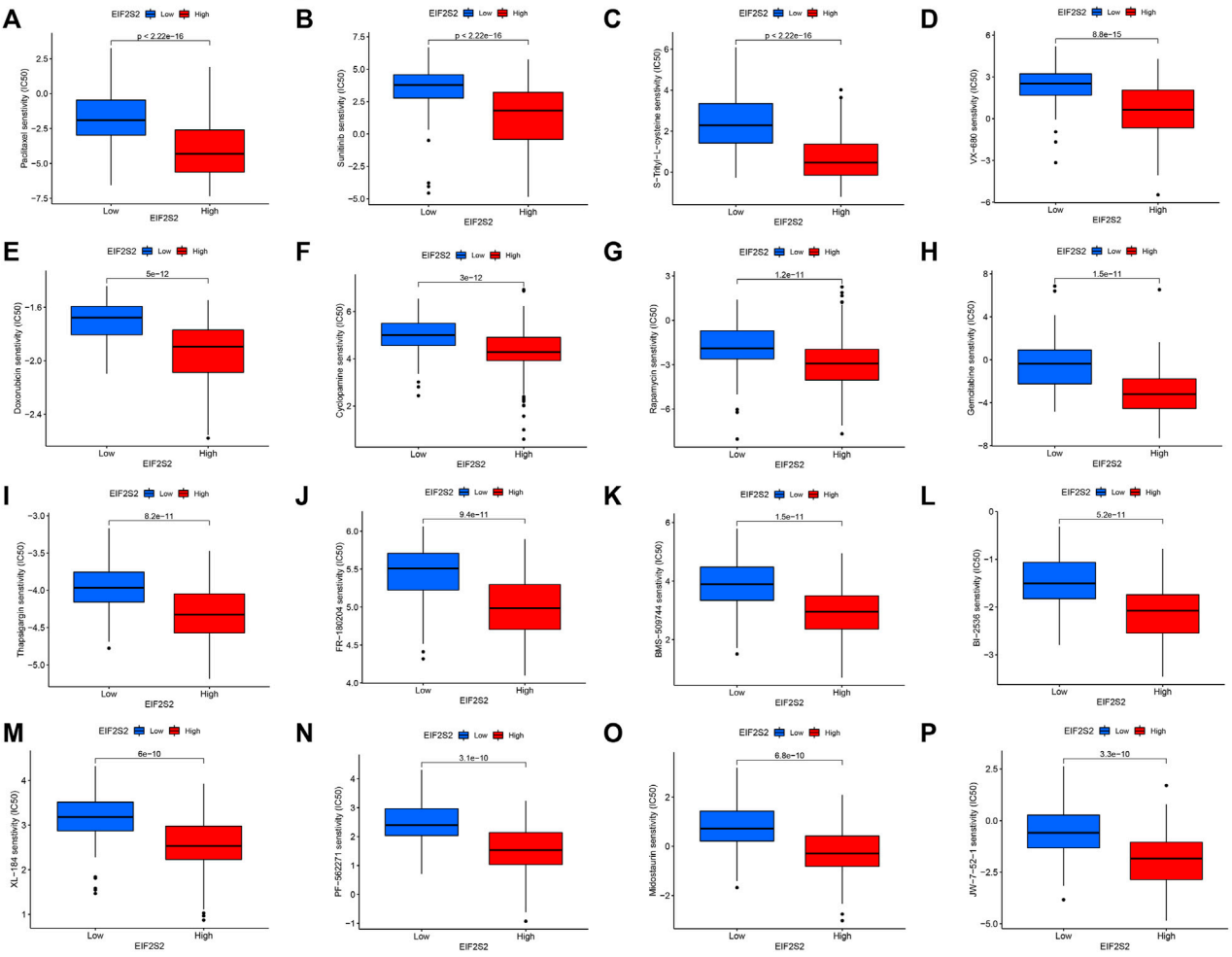


FIGURE 7
The relationship between EIF2S2 expression levels and drug sensitivity.

category, associated with external side of plasma membrane, microtubule and synaptic membrane in the CC category, and associated with channel activity, passive transmembrane transporter activity and ion channel activity in the MF category (Figures 5A,B). The KEGG results showed that: EIF2S2 was mainly in neuroactive ligand-receptor interaction, cytokine-cytokine receptor interaction, and cell adhesion molecules, Th1 and Th2 cell differentiation, phagosome and other pathways enriched (Figure 5C). The above results suggest that differential genes regulate many types of biological pathways, participate in a variety of biological processes, and are associated with enriched pathways associated with tumors.

Correlation between eukaryotic translation initiation factor 2 subunit beta expression and immune cell infiltration

We used multiple algorithms to analyze the relationship of EIF2S2 expression to immune-related cells. The results showed that EIF2S2 expression was positively or negatively correlated with various immune cells. Taking CIBERSORT-ABS as an example, it can be seen that EIF2S2 expression was closely related to memory B cell, plasma B cell, CD8+ T cell, CD4+ resting memory T cell, T follicular helper cells, regulatory T cell, M0 Macrophage, M1 Macrophage, etc. have a positive correlation, and have a negative correlation with resting NK cell and activated mast cell (Figure 6A). At the same time, we also

analyzed the relationship between EIF2S2 and immune checkpoint, and the results showed that EIF2S2 expression was positively correlated with PDCD1, TIGIT, CTLA4, LAG-3, BTLA, etc. (Figure 6B).

Relationship between eukaryotic translation initiation factor 2 subunit beta and drug sensitivity

Based on the drug sensitivity data from the GDSC database, we predicted the chemosensitivity of each tumor sample by the R software package "pRRophetic" to further explore the correlation between EIF2S2 and common antitumor drug sensitivity. We selected the top 16 drugs with the highest differences. As can be seen from the figure, the EIF2S2 high expression group was more sensitive to Paclitaxel, Sunitinib, S-Trityl-L-cysteine, VX-680, Doxorubicin, Cyclopamine, Rapamycin, and Gemcitabine (Figure 7A–P).

Discussion

The incidence and mortality from HCC are increasing globally, especially in the United States and Europe, and hepatocellular carcinoma is also one of the fastest rising causes of cancer-related mortality in the United States (Kulik and El-Serag, 2019). Even though surgery may be curative, up to 70% of HCC patients experience recurrence 5 years after surgery (Villanueva, 2019). The main reason for the poor prognosis of HCC is the lack of effective treatments and biomarkers (Liu et al., 2021). Immunotherapies such as immune checkpoint inhibitors have brought new hope for HCC patients (Jiang et al., 2019). However, despite the important role of immunotherapy in the treatment of HCC, identifying clinically useful biomarkers as surrogate indicators of immunotherapy response and prognosis in HCC remains the focus of many *in vitro* and *in vivo* studies (Liu et al., 2022). At present, there is still no simple and effective way to predict the prognosis of HCC patients and provide individualized treatment (Harding et al., 2019).

In our study, we found that EIF2S2 was abnormally expressed in HCC tissues compared to normal tissues, and patients with high EIF2S2 expression had shortened OS and PFS, indicating that EIF2S2 plays an important role in the progression of HCC. The high expression of EIF2S2 predicts poor prognosis of HCC, and both univariate and multivariate COX analysis suggested that high EIF2S2 expression may serve as an independent prognostic factor for HCC. The expression of EIF2S2 is closely related to the clinicopathological features of HCC, such as grade and stage. We found that the high expression of EIF2S2 was correlated with clinicopathological parameters including age, pathological grade, clinical stage, and T stage. In addition, we also found some genes co-expressed with EIF2S2. CHMP4B, TPD52L2 were higher in EIF2S2 high expression group. ADH1B, CYP8B1 in the EIF2S2 low expression group was higher than that in the high expression group. Studies have shown that CHMP4B (Charged multivesicular body protein 4B) is significantly overexpressed in HCC, which is associated with poor prognosis and drug resistance to doxorubicin (Hu et al., 2015). TPD52L2 (tumor protein D52 like 2) regulates the proliferation of hepatocellular carcinoma cells by interacting with ATP binding cassette protein (Zhou et al., 2013). The expression of ADH1B, CYP8B1 was significantly correlated with the good survival of HCC patients (Liu et al., 2020; Zhang et al., 2021). Collectively, EIF2S2 could be serve as a potential prognostic marker for HCC patients.

Through the analysis of GO and KEGG signaling pathways, we found that EIF2S2 was mainly involved in cytokine receptor interactions. It is a key mediator of cell communication in the tumor microenvironment, and cytokines play an important role in the occurrence and development of tumors (Holder et al., 2022). The expression of EIF2S2 was significantly enriched in cytokine receptors, indicating that EIF2S2 was involved in tumor development and immunotherapy. CD4+ helper T (Th) cells are mainly involved in tumor immunology, and can be functionally subdivided into different subgroups, namely Th1, Th2 and Th17 cells, according to the secretion of cytokines and immune function. Th1 cytokines were associated with good prognosis in patients with HCC, while Th2 cytokines are associated with associated with tumor growth or metastasis (Budhu et al., 2006; Mantovani et al., 2008; Zhou et al., 2009). The enrichment results showed that EIF2S2 was enriched with Th1 and Th2 cell pathways, indicating that EIF2S2 was closely related to the prognosis and metastasis of HCC. We also found that EIF2S2 was closely associated with cell adhesion molecules, and down-regulation of EIF2S2 may inhibit the proliferation and migration of HCC through the cell adhesion molecule pathway. Additionally, enriched in phagosomes, cancer cells adapt to a stressful microenvironment by coordinating multiple pathways. Autophagosomes play an important role in the synchronization of internal and external environments during tumorigenesis (Subramanian et al., 2022). Autophagy with dysfunctional autophagosomes has been implicated in a variety of human diseases, including cancer, pathogen infection, diabetes, and neurodegenerative diseases (Mizushima and Levine, 2020).

The liver contains a variety of innate immune cells including NK cells, macrophages, NKT cells, and adaptive immune cells including T cells and B cells, which can affect the status of immune tolerance, tumor progression, and pathogen clearance (Liaskou et al., 2012; Li and Hua, 2017; Ringelhan et al., 2018). Immune infiltration plays an important role in the occurrence and development of HCC (Rohr-Udilova et al., 2018). Tumor immune infiltrating cells are an important part of the tumor microenvironment and have been proven to play an important role in tumor proliferation and metastasis. T cells, B cells, and NK cells are all members of Tumor-Infiltrating Lymphocytes, which

are typical components of the host's antitumor immune response (Ding et al., 2018). CD8+ T lymphocyte dysfunction and exhaustion are characterized by the upregulation of immunosuppressive molecules, such as PD-1, CTLA-4, which inhibit CD8+ T lymphocyte activation. Continued suppression in the tumor microenvironment may result in the inability of infiltrating CD8+ T lymphocytes to kill tumor cells, ultimately leading to tumor immune escape (Schoenberg et al., 2021). In human liver, NK cells can protect hepatocytes from hepatitis virus attack and malignant transformation. The cytotoxicity of NK cells to hepatoma cells can be promoted in several ways. For instance, miR-506 promotes the antitumor effect of NK cells by regulating STAT3 (Su et al., 2019). Our study shows that EIF2S2 is associated with a variety of immune cells, which may promote tumor proliferation and metastasis by promoting immune infiltration of HCC.

In recent years, immune checkpoint inhibitors (ICIs) have made great progress in the treatment of many types of cancer, including HCC(Inarrairaegui et al., 2018). ICIs, including PD-1, PD-L1, and CTLA-4 antibodies, can enhance the activity of effective T cells and suppress immunosuppression in the tumor microenvironment (Kuol et al., 2018). Disappointedly, emerging ICI immunotherapy has been shown to significantly improve clinical outcomes in HCC, but its objective response rate as monotherapy for HCC is only 15%–20% (Ruf et al., 2021). Previous studies have shown that highly expressed PDCD1 has better OS and has a favorable relationship with CD8+ T cells, B cells, macrophages, CD4+ T cells (Li et al., 2022). T cell immunoglobulin and ITIM domain protein (TIGIT) is a type I transmembrane protein, mainly expressed in activated T cells, Treg, memory T cells and NK cells. TIGIT is commonly co-expressed with LAG-3, TIM-3 and PD-1. They are jointly involved in the immune recognition of the body and are closely related to the survival of patients (Anderson et al., 2016). LAG-3 plays an important role in negatively regulating T cell activation and proliferation. It is expected to become a major target after PD-1 in the development of cancer therapy (Maruhashi et al., 2020). B and T lymphocyte attenuator (BTLA) is a lymphocyte inhibitory receptor similar to CTLA-4 and PD-1. BTLA is a marker that recognizes exhausted PD-1-expressing CD4 T cells and can functionally inhibit CD4 T lymphocyte function through interaction with herpesvirus entry mediator cells (Zhao et al., 2016). Regarding immune checkpoints associated with tumor cell immune evasion, such as CTLA-4, PD-1, PD-L1, TIGIT, these inhibitory receptors/ligands suppress antitumor immune responses by altering their expression levels (Xing et al., 2021). In this study, the expression of EIF2S2 was significantly correlated with the expression of the above-mentioned immune checkpoints, which is of great significance. These findings may provide new ideas and research prospects for EIF2S2 in the treatment of HCC in clinical settings.

Paclitaxel is a broadly active and potent cytotoxic drug. Hepatic arterial infusion of paclitaxel has good efficacy in patients with advanced colorectal cancer, thyroid cancer and hepatocellular carcinoma with liver metastases (Tsimberidou et al., 2011). Sunitinib is a multi-targeted receptor tyrosine kinase inhibitor with antitumor activity (Nassif et al., 2017). Studies have shown that transarterial chemoembolization plus sunitinib is safe and feasible for patients with liver cancer who are not suitable for surgical resection (Turpin et al., 2021). Sunitinib combined with Minimally invasive radiofrequency ablation can significantly inhibit the growth of liver cancer (Qi et al., 2020). Doxorubicin is an antitumor antibiotic that inhibits the synthesis of RNA and DNA, and has effects on a variety of tumors. The chitosan-coated doxorubicin nanoparticle drug delivery system can inhibit the growth of hepatoma cells through the p53/PRC1 pathway (Ye et al., 2018). S-triphenyl-L-cysteine is an Eg5 inhibitor that blocks cell mitosis and exhibits tumor growth inhibitory activity (Kozielski et al., 2008). Our results show that the IC50 of the EIF2S2 high expression group is lower than that of the low expression group, and the sensitivity to these drugs is higher.

This study is the first to our knowledge to have explored the relationship between EIF2S2 and outcomes and immune infiltration in HCC patient. However, our study also has certain limitations and needs to be confirmed *in vivo* and *in vitro* experiments. In the future, more immunosuppressants need to be developed, new treatments explored, and new prognostic biomarkers discovered for better therapeutic effects.

Conclusion

EIF2S2 is a novel biomarker that can predict the prognosis of HCC, and its high expression is correlated with the OS, PFS and clinicopathological characteristics of HCC patients. In addition, EIF2S2 is also associated with immune cell infiltration and immune checkpoint, which can become a potential therapeutic approach, bringing new hope for the clinical treatment of HCC (AuthorAnonymous et al., 2022).

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

JL contributed to data acquisition and article drafting; TL, CZ, JH, and DZ provided technical support; ZW and RW contributed to study design and supervision. All authors

contributed to the article and approved the final version 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

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.

References

Anderson, A. C., Joller, N., and Kuchroo, V. K. (2016). Lag-3, tim-3, and TIGIT: Co-Inhibitory receptors with specialized functions in immune regulation. *Immunity* 44 (5), 989–1004. doi:10.1016/j.immuni.2016.05.001

Budhu, A., Forgues, M., Ye, Q. H., Jia, H. L., He, P., Zanetti, K. A., et al. (2006). Prediction of venous metastases, recurrence, and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment. *Cancer Cell* 10 (2), 99–111. doi:10.1016/j.ccr.2006.06.016

Ding, W., Xu, X., Qian, Y., Xue, W., Wang, Y., Du, J., et al. (2018). Prognostic value of tumor-infiltrating lymphocytes in hepatocellular carcinoma: A meta-analysis. *Med. Baltim.* 97 (50), e13301. doi:10.1097/MD.0000000000013301

Ganne-Carrie, N., and Nahon, P. (2019). Hepatocellular carcinoma in the setting of alcohol-related liver disease. *J. Hepatol.* 70 (2), 284–293. doi:10.1016/j.jhep.2018.10.008

Geeleher, P., Cox, N., and Huang, R. S. (2014). pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One* 9 (9), e107468. doi:10.1371/journal.pone.0107468

Harding, J. J., Khalil, D. N., and Abou-Alfa, G. K. (2019). Biomarkers: What role do they play (if any) for diagnosis, prognosis and tumor response prediction for hepatocellular carcinoma? *Dig. Dis. Sci.* 64 (4), 918–927. doi:10.1007/s10620-019-05517-6

Heaney, J. D., Michelson, M. V., Youngren, K. K., Lam, M. Y., and Nadeau, J. H. (2009). Deletion of eIF2beta suppresses testicular cancer incidence and causes recessive lethality in agouti-yellow mice. *Hum. Mol. Genet.* 18 (8), 1395–1404. doi:10.1093/hmg/ddp045

Hlady, R. A., Sathyanarayan, A., Thompson, J. J., Zhou, D., Wu, Q., Pham, K., et al. (2019). Integrating the epigenome to identify drivers of hepatocellular carcinoma. *Hepatology* 69 (2), 639–652. doi:10.1002/hep.30211

Holder, P. G., Lim, S. A., Huang, C. S., Sharma, P., Dagdas, Y. S., Bulutoglu, B., et al. (2022). Engineering interferons and interleukins for cancer immunotherapy. *Adv. Drug Deliv. Rev.* 182, 114112. doi:10.1016/j.addr.2022.114112

Hu, B., Jiang, D., Chen, Y., Wei, L., Zhang, S., Zhao, F., et al. (2015). High CHMP4B expression is associated with accelerated cell proliferation and resistance to doxorubicin in hepatocellular carcinoma. *Tumour Biol.* 36 (4), 2569–2581. doi:10.1007/s13277-014-2873-1

Inarrairaegui, M., Melero, I., and Sangro, B. (2018). Immunotherapy of hepatocellular carcinoma: Facts and hopes. *Clin. Cancer Res.* 24 (7), 1518–1524. doi:10.1158/1078-0432.CCR-17-0289

Jiang, Y., Han, Q. J., and Zhang, J. (2019). Hepatocellular carcinoma: Mechanisms of progression and immunotherapy. *World J. Gastroenterol.* 25 (25), 3151–3167. doi:10.3748/wjg.v25.i25.3151

Jin, F., Wu, Z., Hu, X., Zhang, J., Gao, Z., Han, X., et al. (2019). The PI3K/Akt/GSK-3beta/ROS/eIF2B pathway promotes breast cancer growth and metastasis via suppression of NK cell cytotoxicity and tumor cell susceptibility. *Cancer Biol. Med.* 16 (1), 38–54. doi:10.20892/j.issn.2095-3941.2018.0253

Kanno, H., Goto, Y., Sasaki, S., Fukutomi, S., Hisaka, T., Fujita, F., et al. (2021). Geriatric nutritional risk index predicts prognosis in hepatocellular carcinoma after hepatectomy: A propensity score matching analysis. *Sci. Rep.* 11 (1), 9038. doi:10.1038/s41598-021-88254-z

Kashiwagi, K., Takahashi, M., Nishimoto, M., Hiyama, T. B., Higo, T., Umehara, T., et al. (2016). Crystal structure of eukaryotic translation initiation factor 2B. *Nature* 531 (7592), 122–125. doi:10.1038/nature16991

Kozielski, F., Skoufias, D. A., Indorato, R. L., Saoudi, Y., Jungblut, P. R., Hustoft, H. K., et al. (2008). Proteome analysis of apoptosis signaling by S-trityl-L-cysteine, a potent reversible inhibitor of human mitotic kinesin Eg5. *Proteomics* 8 (2), 289–300. doi:10.1002/pmic.200700534

Kulik, L., and El-Serag, H. B. (2019). Epidemiology and management of hepatocellular carcinoma. *Gastroenterology* 156 (2), 477–491 e471. doi:10.1053/j.gastro.2018.08.065

Kuol, N., Stojanovska, L., Nurgali, K., and Apostolopoulos, V. (2018). PD-1/PD-L1 in disease. *Immunotherapy* 10 (2), 149–160. doi:10.2217/imt-2017-0120

Li, N., and Hua, J. (2017). Immune cells in liver regeneration. *Oncotarget* 8 (2), 3628–3639. doi:10.18632/oncotarget.12275

Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., et al. (2020). TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.* 48 (W1), W509–W514. doi:10.1093/nar/gkaa407

Li, W., Mei, M., Liu, T., Zhang, S., Wang, Z., Suo, Y., et al. (2022). Identification of PDCD1 and PDCD1LG2 as prognostic biomarkers and associated with immune infiltration in hepatocellular carcinoma. *Int. J. Gen. Med.* 15, 437–449. doi:10.2147/IJGM.S347163

Liaskou, E., Wilson, D. V., and Oo, Y. H. (2012). Innate immune cells in liver inflammation. *Mediat. Inflamm.* 2012, 949157. doi:10.1155/2012/949157

Liu, J. K. H., Irvine, A. F., Jones, R. L., and Samson, A. (2022). Immunotherapies for hepatocellular carcinoma. *Cancer Med.* 11 (3), 571–591. doi:10.1002/cam4.4468

Liu, X., Li, T., Kong, D., You, H., Kong, F., and Tang, R. (2020). Prognostic implications of alcohol dehydrogenases in hepatocellular carcinoma. *BMC Cancer* 20 (1), 1204. doi:10.1186/s12885-020-07689-1

Liu, Z., Jiao, D., Liu, L., Zhou, X., Yao, Y., Li, Z., et al. (2021). Development and validation of a robust immune-related risk signature for hepatocellular carcinoma. *Med. Baltim.* 100 (10), e24683. doi:10.1097/MD.0000000000024683

Llovet, J. M., De Baere, T., Kulik, L., Haber, P. K., Greten, T. F., Meyer, T., et al. (2021). Locoregional therapies in the era of molecular and immune treatments for hepatocellular carcinoma. *Nat. Rev. Gastroenterol. Hepatol.* 18 (5), 293–313. doi:10.1038/s41575-020-00395-0

Mantovani, A., Allavena, P., Sica, A., and Balkwill, F. (2008). Cancer-related inflammation. *Nature* 454 (7203), 436–444. doi:10.1038/nature07205

Maruhashi, T., Sugiura, D., Okazaki, I. M., and Okazaki, T. (2020). LAG-3: From molecular functions to clinical applications. *J. Immunother. Cancer* 8 (2). doi:10.1136/jitc-2020-001014

Mizushima, N., and Levine, B. (2020). Autophagy in human diseases. *N. Engl. J. Med.* 383 (16), 1564–1576. doi:10.1056/NEJMra2022774

Nassif, E., Thibault, C., Vano, Y., Fournier, L., Mauge, L., Verkarre, V., et al. (2017). Sunitinib in kidney cancer: 10 years of experience and development. *Expert Rev. Anticancer Ther.* 17 (2), 129–142. doi:10.1080/14737140.2017.1272415

Qi, X., Yang, M., Ma, L., Sauer, M., Avella, D., Kaifi, J. T., et al. (2020). Synergizing sunitinib and radiofrequency ablation to treat hepatocellular cancer by triggering the antitumor immune response. *J. Immunother. Cancer* 8 (2). doi:10.1136/jitc-2020-001038

Ringelhan, M., Pfister, D., O'Connor, T., Pikarsky, E., and Heikenwalder, M. (2018). The immunology of hepatocellular carcinoma. *Nat. Immunol.* 19 (3), 222–232. doi:10.1038/s41590-018-0044-z

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J., et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinforma.* 12, 77. doi:10.1186/1471-2105-12-77

Rohr-Udilova, N., Klinglmuller, F., Schulte-Hermann, R., Stift, J., Herac, M., Salzmann, M., et al. (2018). Deviations of the immune cell landscape between healthy liver and hepatocellular carcinoma. *Sci. Rep.* 8 (1), 6220. doi:10.1038/s41598-018-24437-5

Ruf, B., Heinrich, B., and Greten, T. F. (2021). Immunobiology and immunotherapy of HCC: Spotlight on innate and innate-like immune cells. *Cell Mol. Immunol.* 18 (1), 112–127. doi:10.1038/s41423-020-00572-w

Samant, H., Amiri, H. S., and Zibari, G. B. (2021). Addressing the worldwide hepatocellular carcinoma: Epidemiology, prevention and management. *J. Gastrointest. Oncol.* 12 (Suppl. 2), S361–S373. doi:10.21037/jgo.2020.02.08

Schoenberg, M. B., Li, X., Li, X., Han, Y., Borner, N., Koch, D., et al. (2021). The interactions between major immune effector cells and hepatocellular carcinoma: A systematic review. *Int. Immunopharmacol.* 101 (Pt A), 108220. doi:10.1016/j.intimp.2021.108220

Singal, A. G., Lampertico, P., and Nahon, P. (2020). Epidemiology and surveillance for hepatocellular carcinoma: New trends. *J. Hepatol.* 72 (2), 250–261. doi:10.1016/j.jhep.2019.08.025

Su, Z., Ye, X., and Shang, L. (2019). MiR-506 promotes natural killer cell cytotoxicity against human hepatocellular carcinoma cells by targeting STAT3. *Yonsei Med. J.* 60 (1), 22–29. doi:10.3349/ymj.2019.60.1.22

Subramanian, V. A., Bairwa, R., Sharma, P., and Bissa, B. (2022). Cancer cell's internal and external warriors: Autophagosomes and exosomes. *Life Sci.* 300, 120552. doi:10.1016/j.lfs.2022.120552

Tanaka, I., Sato, M., Kato, T., Goto, D., Kakumu, T., Miyazawa, A., et al. (2018). eIF2beta, a subunit of translation-initiation factor EIF2, is a potential therapeutic target for non-small cell lung cancer. *Cancer Sci.* 109 (6), 1843–1852. doi:10.1111/cas.13602

Tsimberidou, A. M., Letourneau, K., Fu, S., Hong, D., Naing, A., Wheler, J., et al. (2011). Phase I clinical trial of hepatic arterial infusion of paclitaxel in patients with advanced cancer and dominant liver involvement. *Cancer Chemother. Pharmacol.* 68 (1), 247–253. doi:10.1007/s00280-010-1482-y

Turpin, A., de Baere, T., Heurgue, A., Le Malicot, K., Ollivier-Hourmand, I., Lecomte, T., et al. (2021). Liver transarterial chemoembolization and sunitinib for unresectable hepatocellular carcinoma: Results of the PRODIGE 16 study. *Clin. Res. Hepatol. Gastroenterol.* 45 (2), 101464. doi:10.1016/j.clinre.2020.05.012

Villanueva, A. (2019). Hepatocellular carcinoma. *N. Engl. J. Med.* 380 (15), 1450–1462. doi:10.1056/NEJMra1713263

Xing, R., Gao, J., Cui, Q., and Wang, Q. (2021). Strategies to improve the antitumor effect of immunotherapy for hepatocellular carcinoma. *Front. Immunol.* 12, 783236. doi:10.3389/fimmu.2021.783236

Ye, B. L., Zheng, R., Ruan, X. J., Zheng, Z. H., and Cai, H. J. (2018). Chitosan-coated doxorubicin nano-particles drug delivery system inhibits cell growth of liver cancer via p53/PRC1 pathway. *Biochem. Biophys. Res. Commun.* 495 (1), 414–420. doi:10.1016/j.bbrc.2017.10.156

Younossi, Z. M., and Henry, L. (2021). Epidemiology of non-alcoholic fatty liver disease and hepatocellular carcinoma. *JHEP Rep.* 3 (4), 100305. doi:10.1016/j.jhepr.2021.100305

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Zhang, J., Li, S., Zhang, L., Xu, J., Song, M., Shao, T., et al. (2020). RBP EIF2S2 promotes tumorigenesis and progression by regulating MYC-mediated inhibition via FHIT-related enhancers. *Mol. Ther.* 28 (4), 1105–1118. doi:10.1016/j.ymthe.2020.02.004

Zhang, R., Huang, M., Wang, H., Wu, S., Yao, J., Ge, Y., et al. (2021). Identification of potential biomarkers from hepatocellular carcinoma with MT1 deletion. *Pathol. Oncol. Res.* 27, 597527. doi:10.3389/pore.2021.597527

Zhao, Q., Huang, Z. L., He, M., Gao, Z., and Kuang, D. M. (2016). BTLA identifies dysfunctional PD-1-expressing CD4(+) T cells in human hepatocellular carcinoma. *Oncoimmunology* 5 (12), e1254855. doi:10.1080/2162402X.2016.1254855

Zhou, J., Lin, Y., Shi, H., Huo, K., and Li, Y. (2013). hABCF3, a TPD52L2 interacting partner, enhances the proliferation of human liver cancer cell lines *in vitro*. *Mol. Biol. Rep.* 40 (10), 5759–5767. doi:10.1007/s11033-013-2679-z

Zhou, L., Chong, M. M., and Littman, D. R. (2009). Plasticity of CD4+ T cell lineage differentiation. *Immunity* 30 (5), 646–655. doi:10.1016/j.immuni.2009.05.001

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 05 October 2022
DOI 10.3389/fgene.2022.956246

Check for updates

OPEN ACCESS

EDITED BY
Ehsan Nazemalhosseini-Mojarad,
Shahid Beheshti University of Medical Sciences, Iran

REVIEWED BY
Qian Chen,
Guangxi Medical University Cancer Hospital, China
Arindam Datta,
University of Pennsylvania, United States

*CORRESPONDENCE
Shuqun Zhang,
shuqun_zhang1971@163.com
Fei Wu,
wufei931105@xjtu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 30 May 2022
ACCEPTED 20 September 2022
PUBLISHED 05 October 2022

CITATION
Li J, Zhang Y, Li C, Wu H, Feng C, Wang W, Liu X, Zhang Y, Cai Y, Jia Y, Qiao H, Wu F and Zhang S (2022), A lactate-related LncRNA model for predicting prognosis, immune landscape and therapeutic response in breast cancer.
Front. Genet. 13:956246.
doi: 10.3389/fgene.2022.956246

COPYRIGHT
© 2022 Li, Zhang, Li, Wu, Feng, Wang, Liu, Zhang, Cai, Jia, Qiao, Wu and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A lactate-related LncRNA model for predicting prognosis, immune landscape and therapeutic response in breast cancer

Jia Li[1†], Yinbin Zhang[1†], Chaofan Li[1], Huizi Wu[1], Cong Feng[1], Weiwei Wang[1], Xuan Liu[1], Yu Zhang[1], Yifan Cai[1], Yiwei Jia[1], Hao Qiao[2], Fei Wu[1*] and Shuqun Zhang[1*]

[1]Department of Oncology, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China, [2]Department of Orthopedics, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China

Breast cancer (BC) has the highest incidence rate of all cancers globally, with high heterogeneity. Increasing evidence shows that lactate and long non-coding RNA (lncRNA) play a critical role in tumor occurrence, maintenance, therapeutic response, and immune microenvironment. We aimed to construct a lactate-related lncRNAs prognostic signature (LRLPS) for BC patients to predict prognosis, tumor microenvironment, and treatment responses. The BC data download from the Cancer Genome Atlas (TCGA) database was the entire cohort, and it was randomly assigned to the training and test cohorts at a 1:1 ratio. Difference analysis and Pearson correlation analysis identified 196 differentially expressed lactate-related lncRNAs (LRLs). The univariate Cox regression analysis, least absolute shrinkage and selection operator (LASSO), and multivariate Cox regression analysis were used to construct the LRLPS, which consisted of 7 LRLs. Patients could be assigned into high-risk and low-risk groups based on the medium-risk sore in the training cohort. Then, we performed the Kaplan–Meier survival analysis, time-dependent receiver operating characteristic (ROC) curves, and univariate and multivariate analyses. The results indicated that the prognosis prediction ability of the LRLPS was excellent, robust, and independent. Furthermore, a nomogram was constructed based on the LRLPS risk score and clinical factors to predict the 3-, 5-, and 10-year survival probability. The GO/KEGG and GSEA indicated that immune-related pathways differed between the two-risk group. CIBERSORT, ESTIMATE, Tumor Immune Dysfunction and Exclusion (TIDE), and Immunophenoscore (IPS) showed that low-risk patients had higher levels of immune infiltration and better immunotherapeutic response. The pRRophetic and CellMiner databases indicated that many common chemotherapeutic drugs were more effective for low-risk patients. In conclusion, we developed a novel LRLPS for BC that could predict the prognosis, immune landscape, and treatment response.

KEYWORDS

lactate, long non-coding RNA, breast cancer, prognostic signature, tumor immune microenvironment, drug sensitivity

Introduction

Breast cancer (BC) is the most common tumor and ranks fifth in cancer-related death globally (Harbeck and Gnant, 2017). Although early detection, diagnosis, and treatment for BC have made significant progress, cancer recurrence, distant metastasis, and drug resistance are still prevalent in patients with BC (Jabbarzadeh kaboli et al., 2020). Many stratification terms have been built for the precise treatment of diseases, and polygenic makers may be more accurate than conventional methods (Li et al., 2019). BC is most commonly classified into five subtypes using PAM50, including luminal A, luminal B, HER2-enriched, normal-like, and basal-like (Harbeck et al., 2019). However, the considerably heterogeneous nature of tumors limits the broad applicability of typing (Waks and Winer, 2019). It is essential to investigate new potential markers for prognostic prediction and provide patients personalized treatments.

Lactate is the endpoint of anaerobic glycolysis and usually is considered an endpoint or waste metabolite in cancer. Recent studies indicate that lactate is an essential regulator of cancer development, maintenance, tumor microenvironment, and metastasis (Doherty and Cleveland, 2013; Hayes et al., 2021). In breast cancer, GPR81 is upregulated and promotes tumor growth by releasing lactate from tumor cells (Longhitano et al., 2022). Lactate dehydrogenase A might be a prognostic marker in clear cell renal cell carcinoma (Girgis et al., 2014). Lactate/BDNF/TrkB signaling could mediate epithelial-stroma interaction and lead to anlotinib resistance in gastric cancer (Jin et al., 2021). In addition, lactate takes part in epigenetic regulation. Histone lysine lactylation is involved in regulating gene transcription (Izzo and Wellen, 2019).

Numerous studies have demonstrated that lactate is relevant to the tumor immune microenvironment (TIME) and immunotherapy. Elevated lactate levels are the primary cause of tumor microenvironment (TME) acidosis, suppressing CD8 + and CD4 + effector T cell function, and favoring immunosuppressive Treg development (Nakagawa et al., 2015; Corbet and Feron, 2017; Erra díaz et al., 2018). As to Innate immunity, tumor-associated macrophages (TAMs) could subvert anti-tumor immune responses and act as a negative prognostic marker (Gabrilovich et al., 2012). Lactate could promote transcriptional polarization of TAM towards the tumor-promoting M2 phenotype in cervical (Stone et al., 2019), breast (Mu et al., 2018), lung cancer and melanoma (Zhang et al., 2019a). In addition to surgery, chemotherapy, radiotherapy and targeted therapies, immunotherapy is the fifth element of cancer treatment. However, the immunosuppressive heavy tumor microenvironment often limits immunotherapy and other therapeutic efficacy. Studies have found that elevated lactate levels can affect the therapeutic efficacy and overall survival of immune checkpoint inhibitors for melanoma (Kelderman et al., 2014), esophageal squamous cell carcinoma (Wang et al., 2019), and non-small cell lung cancer (Zhang et al., 2019b).

LncRNA consists of RNA molecules with at least 200 base pairs that originate from the non-coding region of the genome, involved in almost all human biological processes and series of diseases (Esteller, 2011; Hauptman and Glavač, 2013; Fatica and Bozzoni, 2014). Several studies have reported that lncRNAs could regulate lactate metabolism and immune status in different cancers. The lncRNA SNHG5 regulates BACH1 *via* miR-299 to promote glycolysis and proliferation in breast cancer cells (Huang et al., 2022). LncRNA NEAT1-associated aerobic glycolysis in prostate cancer could blunt tumor immunosurveillance by T cells (Xia et al., 2022). Furthermore, lncRNAs could be used as novel immunotherapeutic tools against cancer, and immunotherapy based on lncRNAs could increase the effectiveness and reduce off-target effects (Kaur et al., 2022). Together, lncRNA plays a role in diagnosing, prognosis, and treating BC (Rodríguez bautista et al., 2018).

Research has shown the critical value of lactate and lncRNAs in cancer classification, prognosis, and immunotherapy (Sun et al., 2022; Xie et al., 2022; Xu et al., 2022). However, lactate-related lncRNAs have not been well studied in BC. Our study developed and verified an LRLPS to predict BC patients' prognosis, immune infiltration and therapeutic response by applying bioinformatics. As a result, our findings may provide new clues for cancer prognosis evaluation and treatment guidance.

Materials and methods

Data collection

The R package "TCGAbiolinks" was used to acquire transcriptome profiling, simple nucleotide variations, and the clinical information of TCGA-BRCA patients (Colaprico et al., 2016). We excluded male patients and retained 1096 BC and 112 normal samples for the differential analysis. Furthermore, 916 BC samples with the OS > 30 days were included in the prognostic analysis. They were randomly divided into the training (n = 458) and test (n = 458) cohorts at a 1:1 ratio using the "caret" R package. Clinical characteristics of the three cohorts were analyzed with the "tableone" R package (Suppplementary Table S1). In the subsequent clinicopathological correlation analysis, we excluded patients with incomplete information. We acquired 284 lactate-related genes by querying the Molecular Signatures Database with "lactic" as the search keyword (Suppplementary Table S2) (Liberzon et al., 2015).

Identification of differential expressed lactate-related lncRNAs in Breast cancer

The "EdgeR" R package assessed the differentially expressed lncRNAs and lactate-related genes ($p < 0.05$, $|\log2FC| = 1$). For further study, we retained differential expression lncRNAs expressed in more than half of the patients. Further identification of lactate-related lncRNAs was performed with Pearson correlation analysis at a standard of $|R| > 0.4$ and the p-value < 0.001.

Construction and validation of the lactate-related lncRNAs prognostic signature

The univariate Cox regression analysis identified the prognostic LRLs in the training cohort. We performed LASSO with the R package "glmnet" to avoid overfitting (Friedman et al., 2010). Then, the LRLPS was built with the multivariate Cox regression analysis based on the stepwise Akaike information criterion (stepAIC) value. According to the LRLPS, each sample could get the risk score with the following formula: Risk score = Σ(Exp∗ Coef). The Coef and Exp were the coefficients and the expression level of each lncRNA, respectively. The high- and low-risk groups were divided according to the median risk score of the training cohort. We further performed the Kaplan–Meier survival analysis, time-dependent ROC curves, and univariate and multivariate analyses to evaluate the accuracy and independence of the LRLPS in prognosis prediction in the three cohorts.

Stratified analysis and construction of the nomogram

The stratified analysis could assess the prognosis value of LRLPS in different subgroups stratified by several clinical features, including age, pathologic stage, T stage, N stage, M stage, ER, PR, and HER2 statuses. We constructed the nomogram with the independent prognostic factors. Nomogram accuracy was evaluated through ROC curves, C-index, and calibration curves. Finally, we measured the net benefit of using a nomogram and other clinical features alone based on decision curve analysis (DCA).

Functional enrichment analysis

We identified the differentially expressed genes (DEGs) between the two risk groups and annotated their functions with Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) using the R package "ClusterProfiler" (Yu et al., 2012). The variations of pathway activity of the subgroups were further revealed with Gene Set Enrichment Analysis (GSEA) ($p < 0.05$ and FDR<0.25) (Subramanian et al., 2007). Annotated gene set "c2. cp.kegg.v7.5.1.symbols.gmt" could acquire from the MSigDB (https://www.gsea-msigdb.org/gsea/msigdb/).

Evaluation of immune infiltration and immunotherapy response in the two risk groups

We evaluated the proportion of tumor-infiltrating immune cells through the CIBERSORT algorithm (Newman et al., 2015). Tumor purity, immune, stromal, and estimate scores were evaluated through the ESTIMATE algorithm (Yoshihara et al., 2013). Furthermore, we assessed twenty-seven potential immune checkpoints (ICPs) in the two risk groups. In order to predict immune checkpoint inhibitor (ICI) responses, we applied IPS and TIDE. TIDE was an online analysis that could predict the response to ICIs (http://tide.dfci.harvard.edu/) (Jiang et al., 2018; Fu et al., 2020). The IPS is a machine learning-based system, scored as z scores according to four immunogenicity-related cell types (effector cells, immunosuppressive cells, MHC molecules, and immunomodulators), and it was positively correlated with immunogenicity (Charoentong et al., 2017). It is reported that IPS could assess the tumor immunogenicity and response to ICI therapy in various tumor types. The IPS of BC patients were downloaded from The Cancer Immunome Atlas (TCIA) (https://tcia.at/home).

Correlation between the risk score and tumor mutation

The mutation landscapes in the two risk groups were analyzed with the "maftools" R package. Mutations in the genome per million bases are known as the tumor mutational burden (TMB), a potential immunotherapy biomarker (Lu et al., 2019; Shum et al., 2022). We evaluated the TMB in the two risk groups and explored the association between TMB and the risk score.

Evaluation of the drug sensitivity and potential target drugs

The "pRRophetic" R package was used to calculate the half-maximal inhibitory concentrations (IC50) of the common chemotherapy drugs based on the Genomics of Drug Sensitivity in Cancer (GDSC; https://www.cancerrxgene.org/) database (Yang et al., 2013; Geeleher et al., 2014). As to the lncRNAs in the LRLPS, we explored the potential target drugs (approved by the FDA and those in

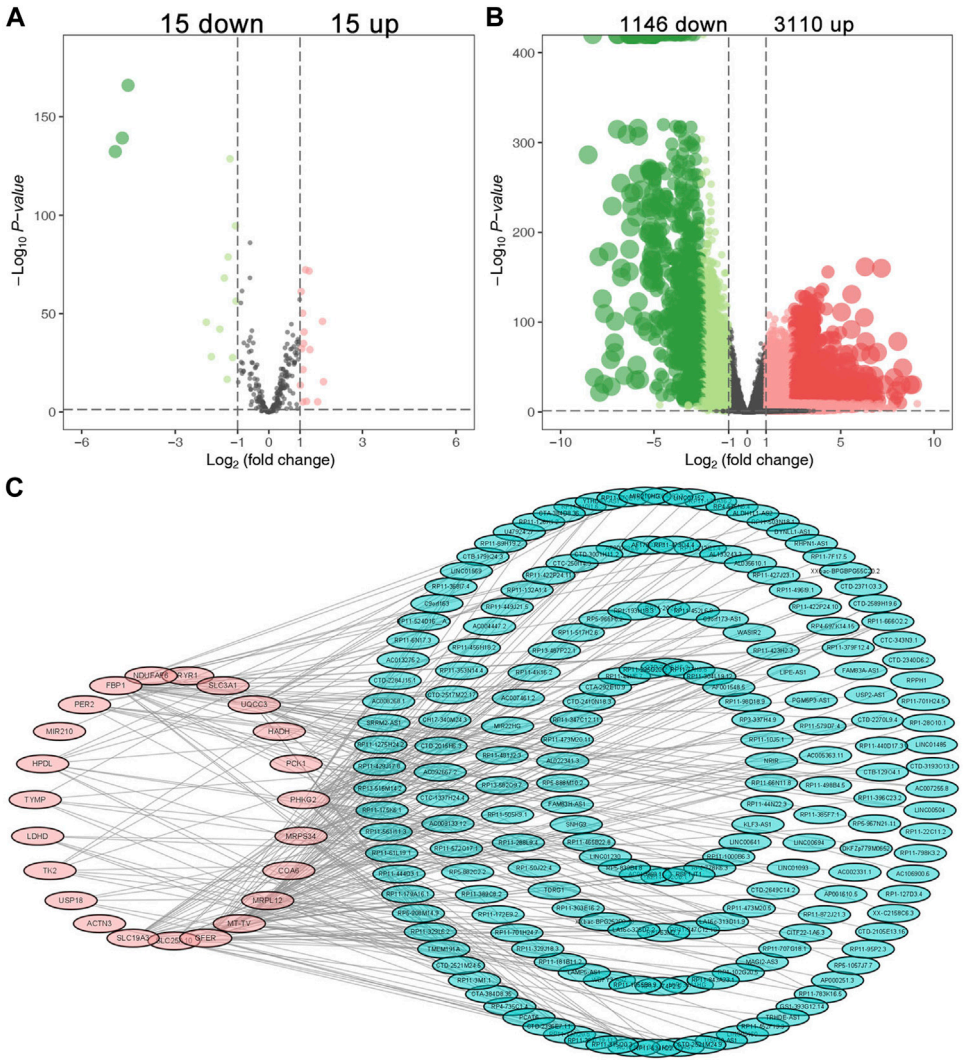


FIGURE 1
Identification of the differential expressed LRLs. The volcano plots of the differentially expressed lactate-related genes **(A)** and differentially expressed lncRNAs **(B)**. **(C)** The interaction between the differentially expressed lactate-related genes and LRLs.

clinical tests) with the CellMiner database (https://discover.nci.nih.gov/cellminer) (Shankavaram et al., 2007; Shankavaram et al., 2009). The relationship between model lncRNAs and drug sensitivity was studied using Pearson correlation analysis.

Statistical analysis

We applied R software (version 4.0.5, https://www.r-project.org/) for all statistical analyses. p-value < 0.05 was set as statistically significant, and the significance levels were set as $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$ and ns = $p > 0.05$.

Results

Identification of the differentially expressed lactate-related lncRNAs in Breast cancer patients

There were 30 differentially expressed lactate-related genes and 4,256 differentially expressed lncRNAs, respectively (Figures 1A,B). Based on Pearson correlation analysis, we identified 196 differentially expressed LRLs for further investigation. Figure 1C showed the interaction between the lactate-related genes and LRLs.

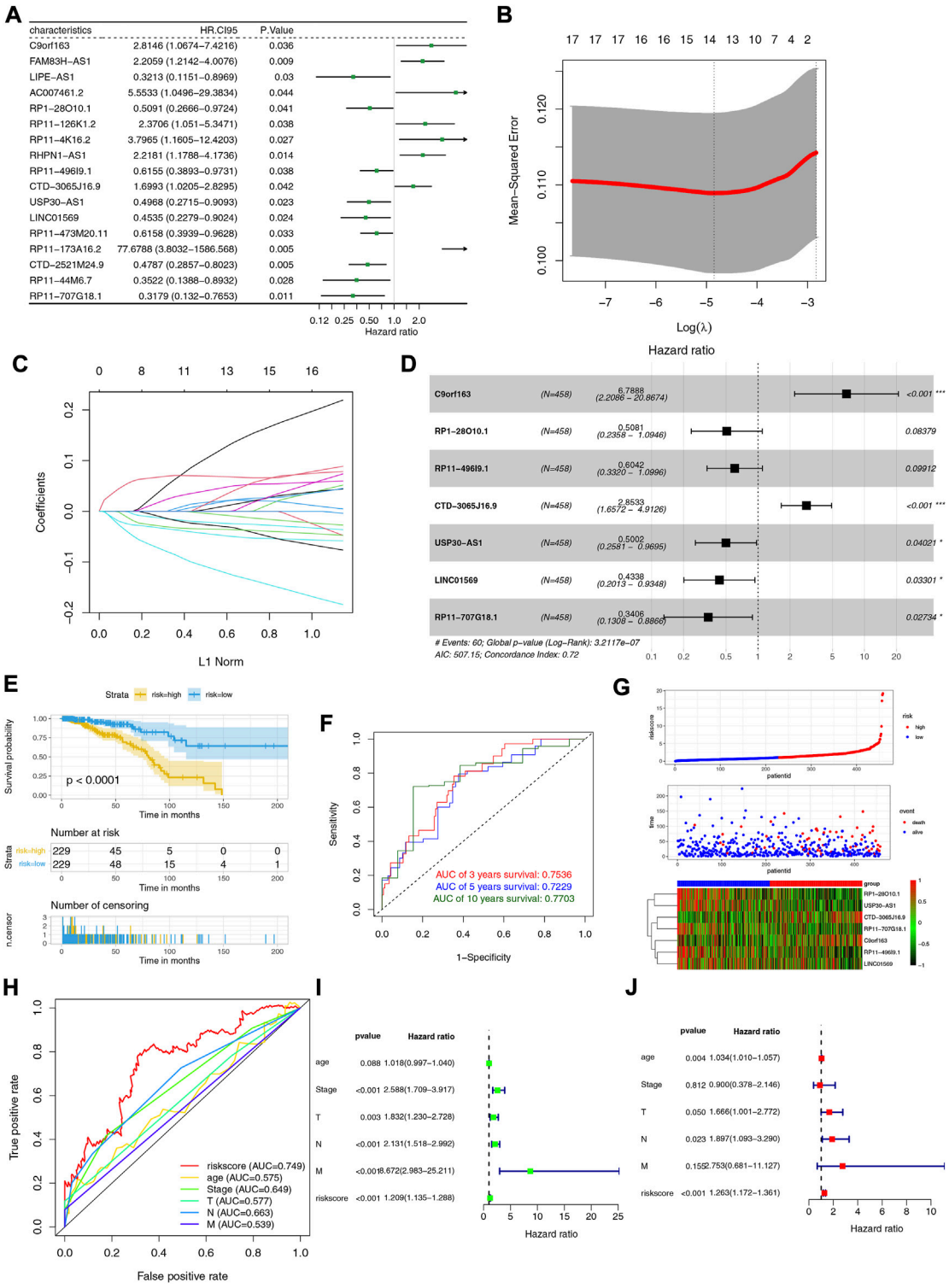


FIGURE 2
Construction and evaluation of the LRLPS. **(A)** The univariate Cox regression analysis of LRLs in the training cohort. **(B)** The cross-validation graph shows the optimal parameter selection with minimum criteria in the LASSO model. **(C)**The LASSO coefficient profiles of the 14 LRLs. **(D)** The forest graph showed the results of stepwise multivariable cox proportional hazards regression analysis. **(E)** The OS curve of the two risk groups. **(F)** The time-dependent ROC curves of the LRLPS. **(G)** The risk score, clinical event, and the model genes in the two risk groups. **(H)** The ROC curves of the risk score and other clinicopathological parameters. The univariate **(I)** and multivariate **(J)** Cox regression analyses.

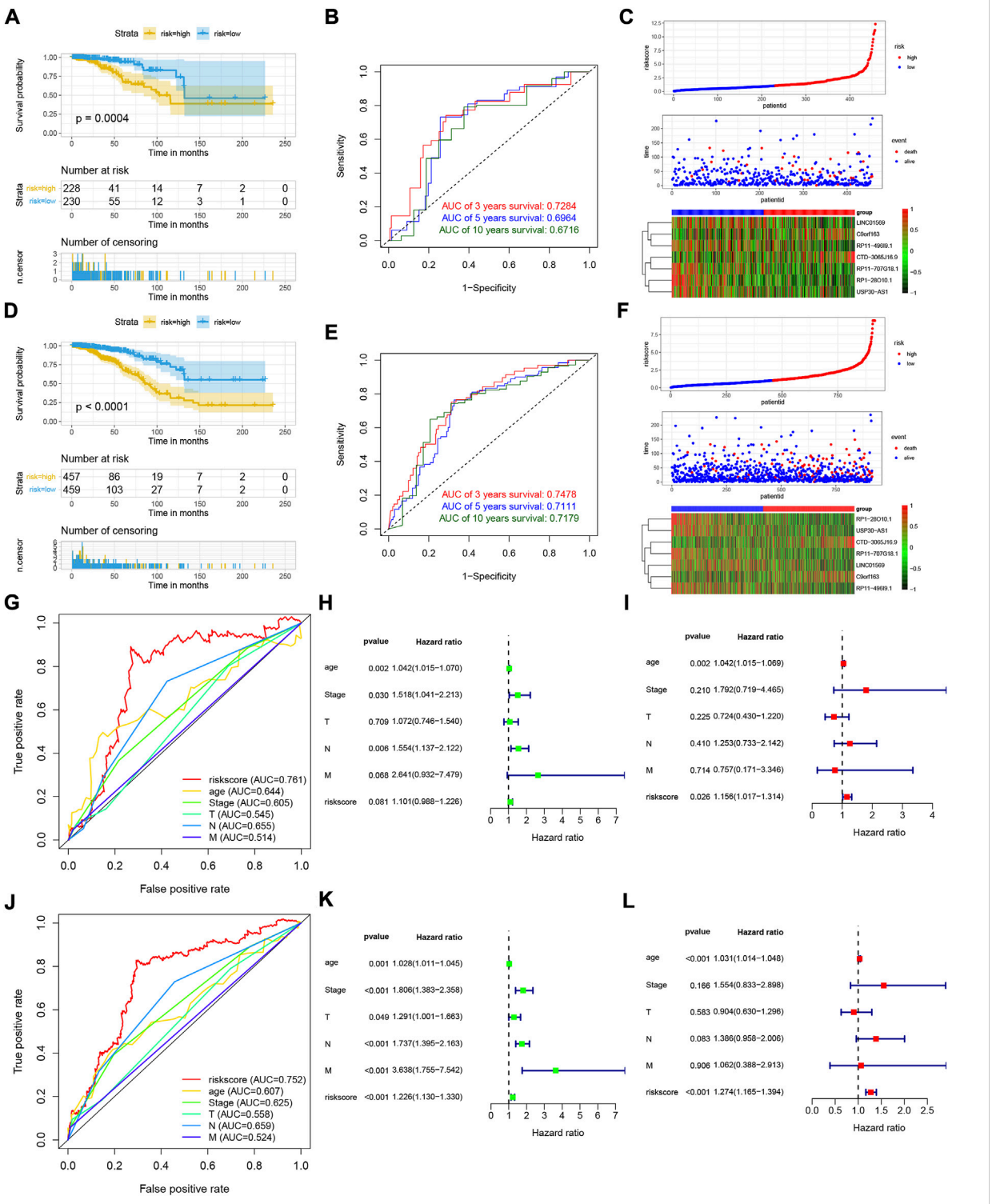


FIGURE 3
Validation of the LRLPS. The OS curve of the two risk groups in test **(A)** and entire **(D)** cohorts. The time-dependent ROC curves in test **(B)** and entire **(E)** cohorts. The risk score, clinical event, and the model genes in the two risk groups in test **(C)** and entire **(F)** cohorts. The ROC curves of the risk score and other clinicopathological parameters in test **(G)** and entire **(J)** cohorts. The univariate Cox regression analyses in the test **(H)** and entire **(K)** cohorts. The multivariate Cox regression analyses in the test **(I)** and entire **(L)** cohorts.

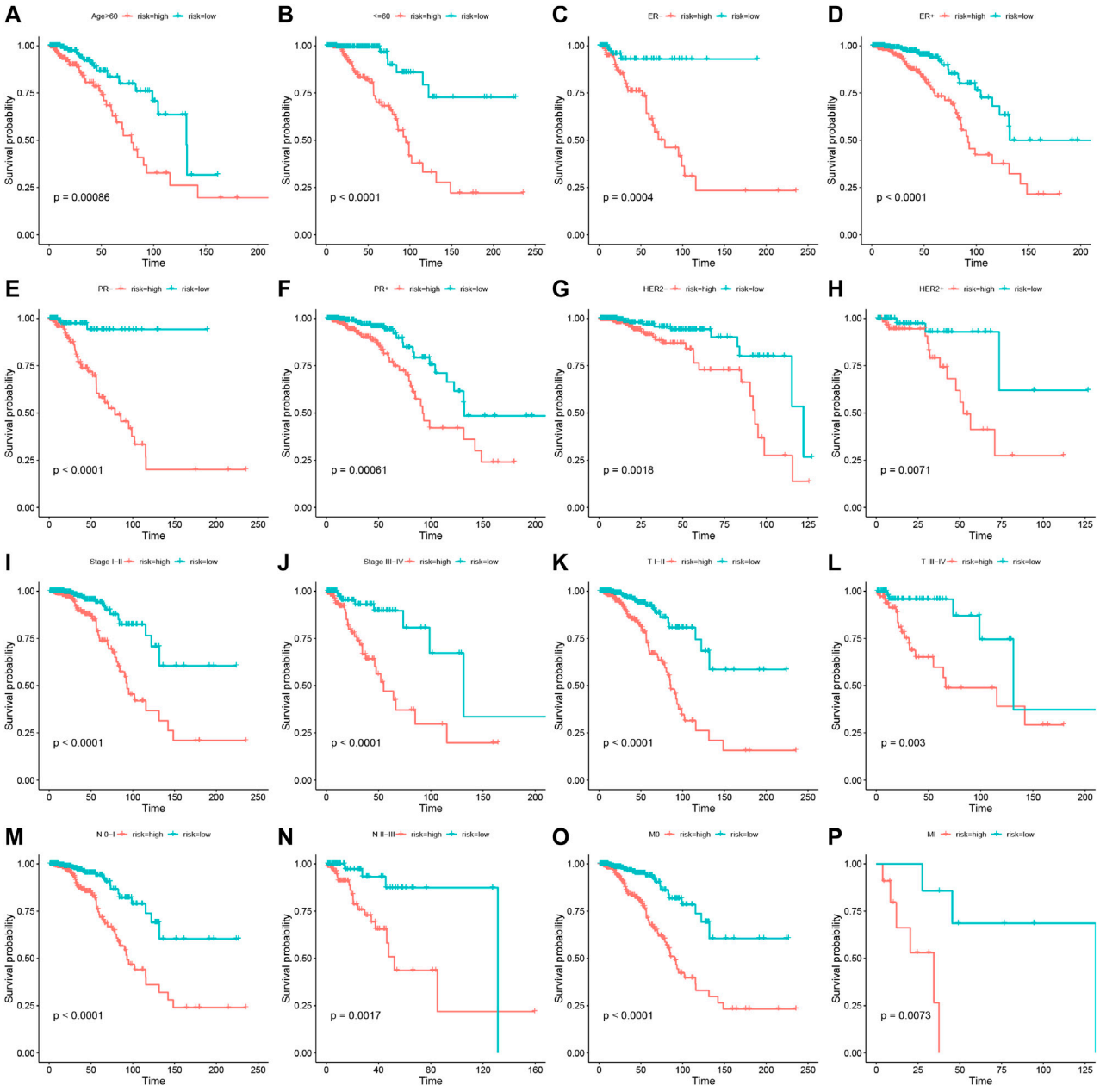


FIGURE 4
Stratification analyses of the prognostic signature. Kaplan-Meier curves indicated the OS of the two risk groups stratified by age (>60 years vs. ≤60 years) **(A,B)**, ER stage (negative vs. positive) **(C,D)**, HER2 stage (negative vs. positive) **(E,F)**, PR stage (negative vs. positive) **(G,H)**, stages (stage I–II vs. stage III-IV) **(I,J)**, AJCC T stage (T I–II vs. T III-IV) **(K,L)**, AJCC N stage (N 0-I vs. T II-III) **(M,N)**, AJCC M stage (M 0 vs. M I) **(O,P)**, respectively.

Development and evaluation of the lactate-related lncRNAs prognostic signature

In the training cohort, 17 LRLs with prognostic values were identified with the univariate Cox regression analysis (Figure 2A). We performed LASSO cox analysis and identified 14 LRLs to avoid overfitting the model (Figures 2B,C). The multivariate Cox regression analysis identified 7 LRLs to construct the LRLPS based on the lowest AIC 507.17 (Figure 2D). Each patient would acquire a risk score by calculating the following: risk score= (1.915272666 *C9orf163) + (−0.677100153 * RP1-28O10.1) + (−0.503780886 * RP11-496I9.1) + (1.048467864 * CTD-3065J16.9) + (−0.692769124 *

TABLE 1 The area under the ROC curve (AUC) showed the sensitivity and specificity of the known gene signatures in predicting the prognosis of BC patients.

Author	Year	Gene signature	Gene number	AUC for OS
Our study	2022	Lactate	7	0.7536 (3-year), 0.7229 (5-year), 0.7703 (10-year)
Chen F, et al.	2022	Necroptosis	7	0.731 (1-year), 0.643 (3-year), 0.641 (5-year), 0.694 (10-year)
Wei T, et al.	2022	Ferroptosis and Immune	7	0.75 (1-year), 0.71 (3-year), 0.63 (5-year), 0.68 (10-year)
Gu P, et al.	2022	Hypoxia	12	0.734 (1-year), 0.727(3-year), 0.741 (5-year), 0.786 (10-year)
Zhang Y, et al.	2022	Necroptosis	4	0.696 (3-year), 0.705 (5-year), 0.664 (7-year)
Zhao Y, et al.	2021	Hypoxia	4	0.650 (1-year), 0.681 (3-year), 0.691 (5-year), 0.642 (7-year)
Li X, et al.	2021	Autophagy	18	0.724 (3-year), 0.685 (5-year)
Jia C, et al.	2021	Ferroptosis	11	0.682 (1-year), 0.710 (3-year), 0.712 (5-year)
Lv W, et al.	2021	N6-methyladenosine	6	0.677 (1-year), 0.678 (3-year), 0.692 (5-year)
Dai Y, et al.	2022	Acid Metabolism	8	0.881 (1-year), 0.766 (3-year), 0.713 (5-year)
Ning S, et al.	2022	CD4 + Conventional T Cells	16	0.742 (1-year), 0.751 (3-year), 0.723 (5-year)
Yang X, et al.	2022	Pyroptosis	10	0.75 (1-year), 0.73 (3-year), 0.73 (5-year)

USP30-AS1) + (−0.835154753 * LINC01569) + (−1.077081426 * RP11-707G18.1) (Supplementary Table S3). Subsequently, we evaluated the ability of the LRLPS in prognosis prediction. Kaplan-Meier analysis showed that patients in the high-risk group had shorter overall survival (OS) (Figure 2E). The area under the 3-, 5-, and 10-year time-dependent ROC curves (AUC) were 0.7536, 0.7229, and 0.7703, respectively, which indicated the accuracy of the LRLPS (Figure 2F). Figure 2G indicated the correlation between the risk score and the outcome of BC patients. Figure 2H indicated that the AUC of risk score was the highest (0.749), followed by the N stage (0.663) and pathological stage (0.649). The univariate (Figure 2I) and multivariate (Figure 2J) Cox regression analyses indicated the independent prognostic value of the risk score.

Validation of the lactate-related lncRNAs prognostic signature

To assess the stability of the LRLPS, we used the same analyses in the test and entire cohorts. High-risk patients always had a worse OS than low-risk patients in the two cohorts (Figures 3A,D). The AUCs of the 3-, 5-, and 10-year ROC curves were 0.7284, 0.6964, and 0.6716 in the test cohort (Figure 3B), and 0.7484, 0.7111, 0.7179 in the entire cohort (Figure 3E). Figures 3C,F indicated that the higher risk score was correlated with increased mortality. In the test cohort, the AUC of the risk score was the highest (0.761) (Figure 3G). The risk score was an independent prognostic factor in the multivariate Cox regression analysis but not a statistically significant independent prognostic according to the univariate Cox regression analysis (Figures 3H,I). In the entire cohort, the risk score had the highest AUC (0.752) and was an independent prognostic factor (Figure 3J-L). We performed the stratification and Kaplan-Meier survival analyses to further explore whether the signature was suitable for different clinical subgroups. There were always significant differences in survival between the two risk groups in all clinical subgroups (Figures 4A–P). The results indicated that the prognostic signature was accurate, independent, and widely applicable.

Comparison with other prognostic signatures

The robustness of our LRLPS was assessed by comparing it with 11 existing OS-related lncRNAs prognostic signatures, such as ferroptosis, pyroptosis, necroptosis, N6-methyladenosine, CD4 + conventional T cells, immunity, and autophagy. We included only signatures from the TCGA database to eliminate the effects of heterogeneity. Signatures were analyzed based on their AUC, with larger AUCs showing better classification ability (Fawcett, 2006). As shown in Table 1, we have integrated all the important information of the eleven signatures, including the author, year, gene signature, and the AUCs for the signatures (Table 1). Our signature had many advantages in predicting OS in BC patients. In our study, the AUCs of the signatures at 3-, 5-, and 10-year were 0.7536, 0.7229, and 7,703, respectively, significantly higher than most hallmark predictive models. Table 1 showed that the 3-, 5-, and 10-year AUCs of another 7 lncRNA prognostic signatures, namely, the ferroptosis- and Immune-related lncRNA signature (3-, 5-, and 10-year AUCs: 0.71, 0.63, and 0.68) (Wei et al., 2022) and necroptosis-related lncRNA signature (3-, 5-, and 10-year AUCs: 0.643, 0.641, and 0.694) (Chen et al., 2022) had lower AUCs than ours; while the 12 hypoxia-related lncRNA signature (3-, 5-, and 10-year AUCs: 0.727, 741, and 0.786) (Gu et al., 2021) were

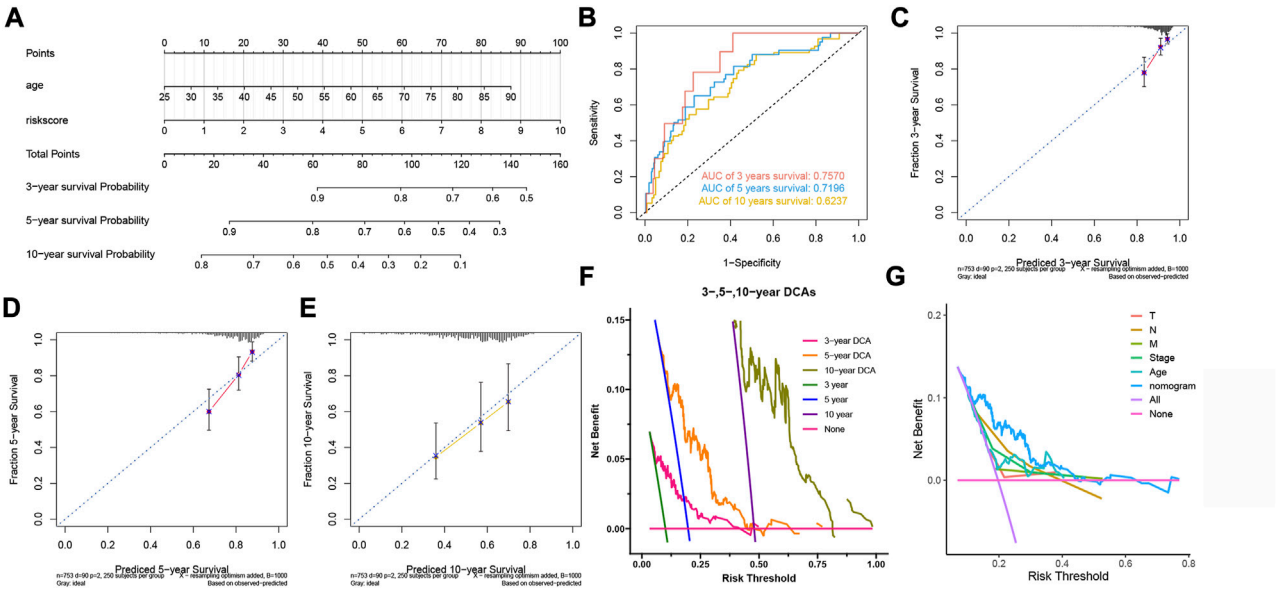


FIGURE 5
Construction and evaluation of the nomogram. **(A)** The nomogram for predicting BC patients' survival probability. **(B)** The nomogram's 3-, 5-, and 10-year ROC curves. **(C,D,E)** The 3-, 5-, and 10-year calibration curves. **(F)** The 3-, 5- and 10-year DCA curves of the nomogram. **(G)** DCA curves of clinicopathological factors and the nomogram.

comparable to the predictive capabilities of our predictive model, and our signature stand out with a clear advantage in predicting the short-term survival of BC patients. We also listed the other signatures that focus more on short-term (3- and 5-year survival) survival. We found our signature usually have better short-term survival prognostic value compared with them, such as the necroptosis-related lncRNA signature (Zhang et al., 2022), hypoxia-related lncRNA signature (Zhao et al., 2021), autophagy-related lncRNA signature (Li et al., 2021), ferroptosis-related lncRNA signature (Jia et al., 2021), N6-methyladenosine-related lncRNA signature (Lv et al., 2021), acid metabolism-related lncRNA signature (Dai et al., 2022), and CD4 + conventional T cells-related lncRNA signature (Ning et al., 2022) and pyroptosis-related lncRNA signature (Yang et al., 2021). In addition, our model only involves 7 lncRNAs, while other models (6/11) tend to have more, which is more convenient to use to a certain extent. The results indicated that our gene signature predicted BC prognosis better than most other signatures.

Construction and evaluation of the nomogram

In order to make our model better assist clinical decision-making, we constructed a nomogram that could predict the 3-, 5-, and 10-year survival probability (Figure 5A). The nomogram's 3-, 5- and 10-year AUCs were 0.7570, 0.7196 and 0.6237, indicating the reliability of the nomogram (Figure 5B). The calibration curves proved that our prognostic nomogram could accurately predict the survival probabilities (Figures 5C–E). Furthermore, DCA curves indicated that the nomogram was more beneficial to BC patients than other clinicopathological factors (Figures 5F,G).

Function analyses

We used GO/KEGG and GSEA analyses to analyze the functions of the two risk groups. There were 3,962 DEGs between the two-risk groups, including 1,524 up-regulated genes and 2,168 down-regulated genes for the high-risk group. GO analysis showed that these DEGs participated in many biological processes, such as humoral immune response, lymphocyte-mediated immunity, and epidermis development (Figure 6A). They could act as structural constituents in the T cell receptor complex, plasma membrane signaling receptor complex, and immunoglobulin complex and play an essential part in receptor ligand activity, signaling receptor activator activity, and gated channel activity (Figure 6A). KEGG analysis showed that the down-regulated genes were related to PD-L1 expression, primary immunodeficiency, cytokine-cytokine receptor interaction, and PD-1 checkpoint pathway in cancer (Figure 6B). Through GSEA analysis, we further observed the variations of pathway

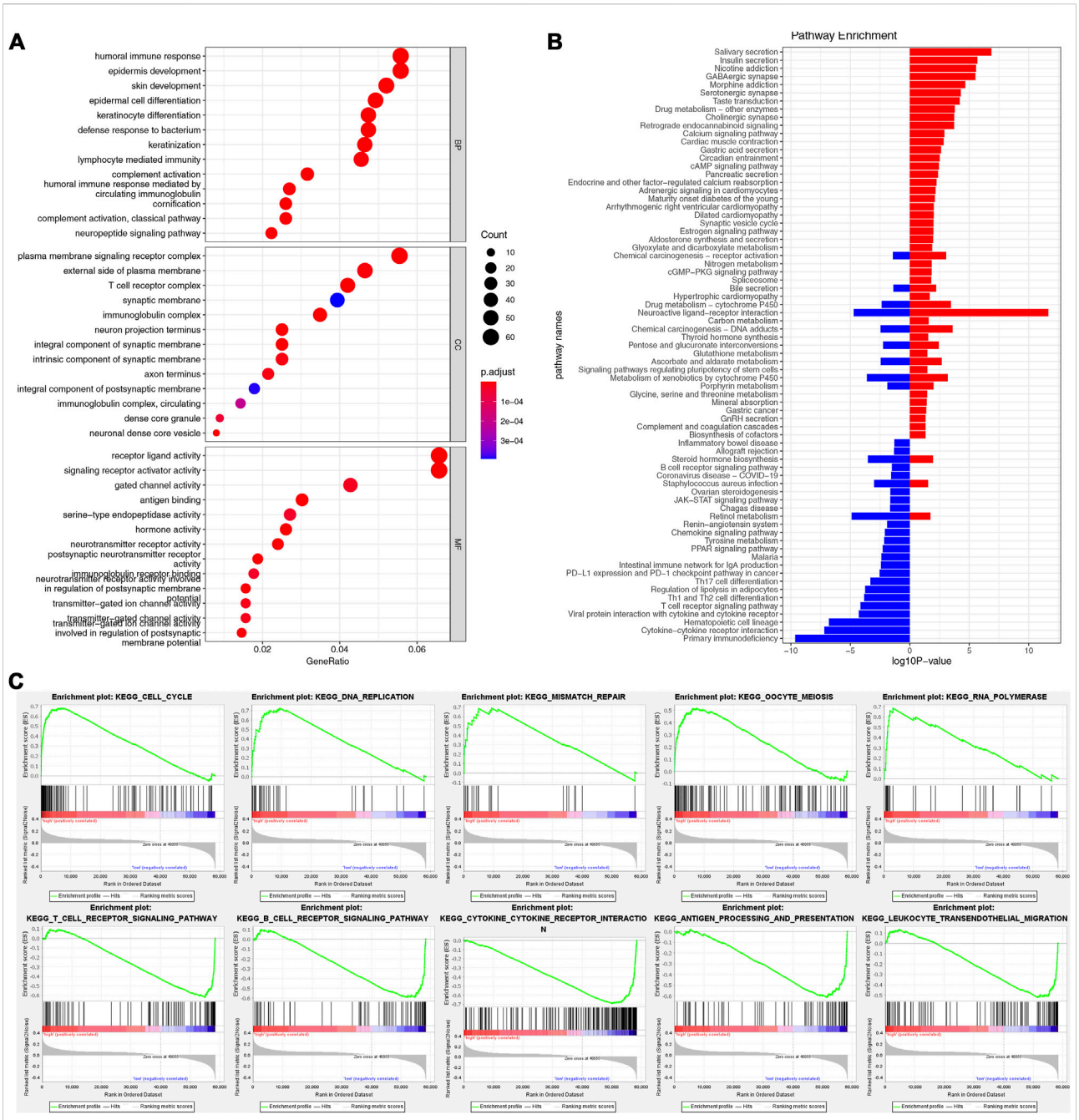


FIGURE 6
Functional enrichment analysis. **(A)** GO enrichment analysis. **(B)** KEGG enrichment analysis. **(C)** The results of GSEA in two risk groups.

activity between the two risk groups. The low-risk group was enriched with the classical immune-related pathways, such as T/B cell receptor signaling pathways, leukocyte transendothelial migration, and antigen processing and presentation, while the high-risk group was enriched with the cell cycle-related pathways, including cell cycle, DNA replication, and mismatch repair (Figure 6C).

Differential immune infiltration and immunotherapy response in the two groups

To further study the immune landscape, we performed CIBERSORT and ESTIMATE algorithms. The heat map demonstrated the levels of the immune infiltrating cells in

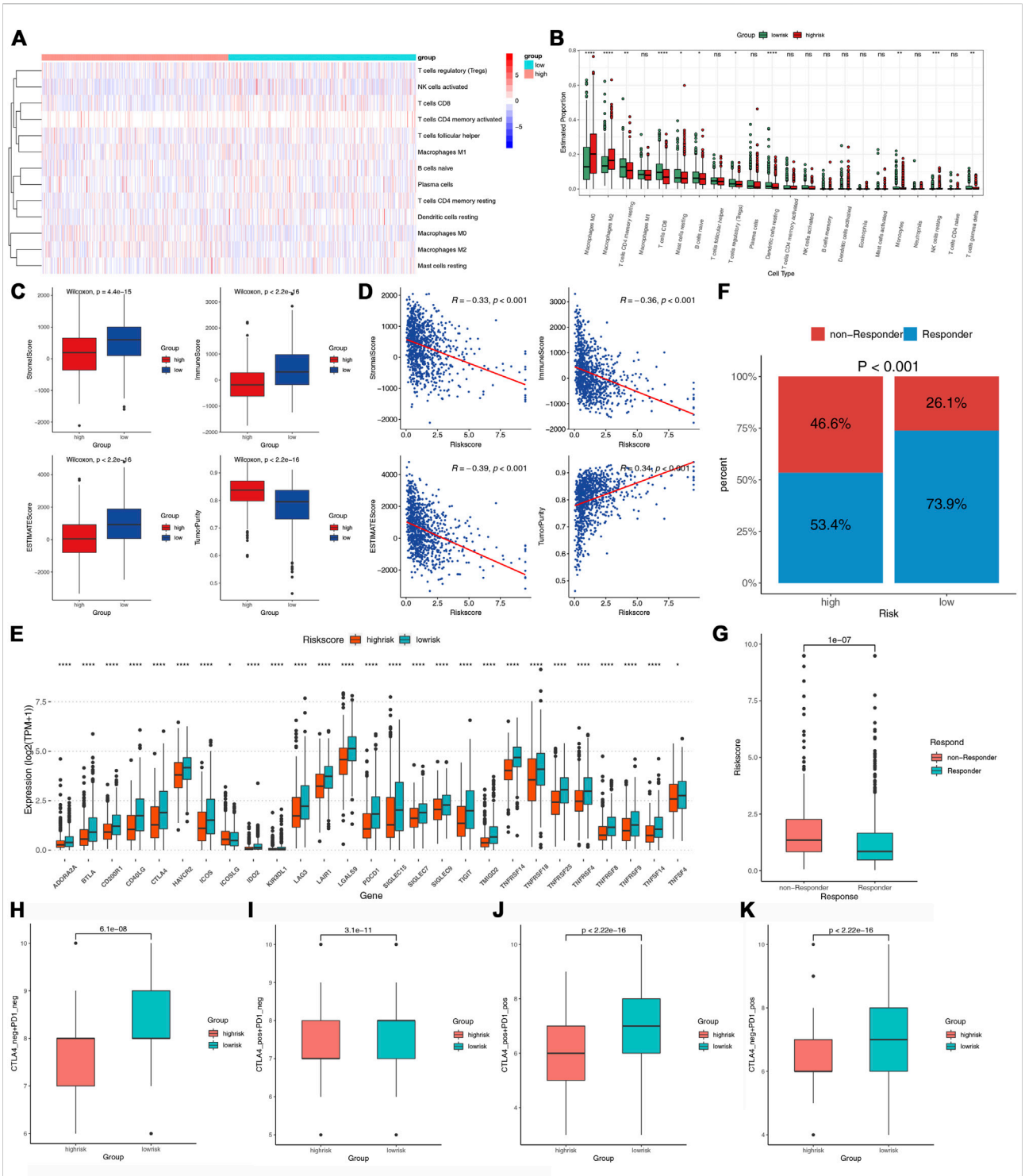


FIGURE 7
The immune infiltration and immunotherapy response in the two groups. **(A,B)** The heatmap and box plots of the proportions of tumor-infiltrating cells in the two risk groups. **(C)** Comparisons of tumor purity stromal, immune, and ESTIMATE scores between the two risk groups. **(D)** Correlations between the risk score and tumor purity, stromal, immune, and estimate score. **(E)** Comparisons of the 27 ICPs in the two risk groups. **(F)** Comparisons of the proportions of non-responders and responders to ICIs between the two risk groups. **(G)** Comparison of the risk score between the responders and non-responders. **(H–K)** Comparison of the IPS between the two risk groups.

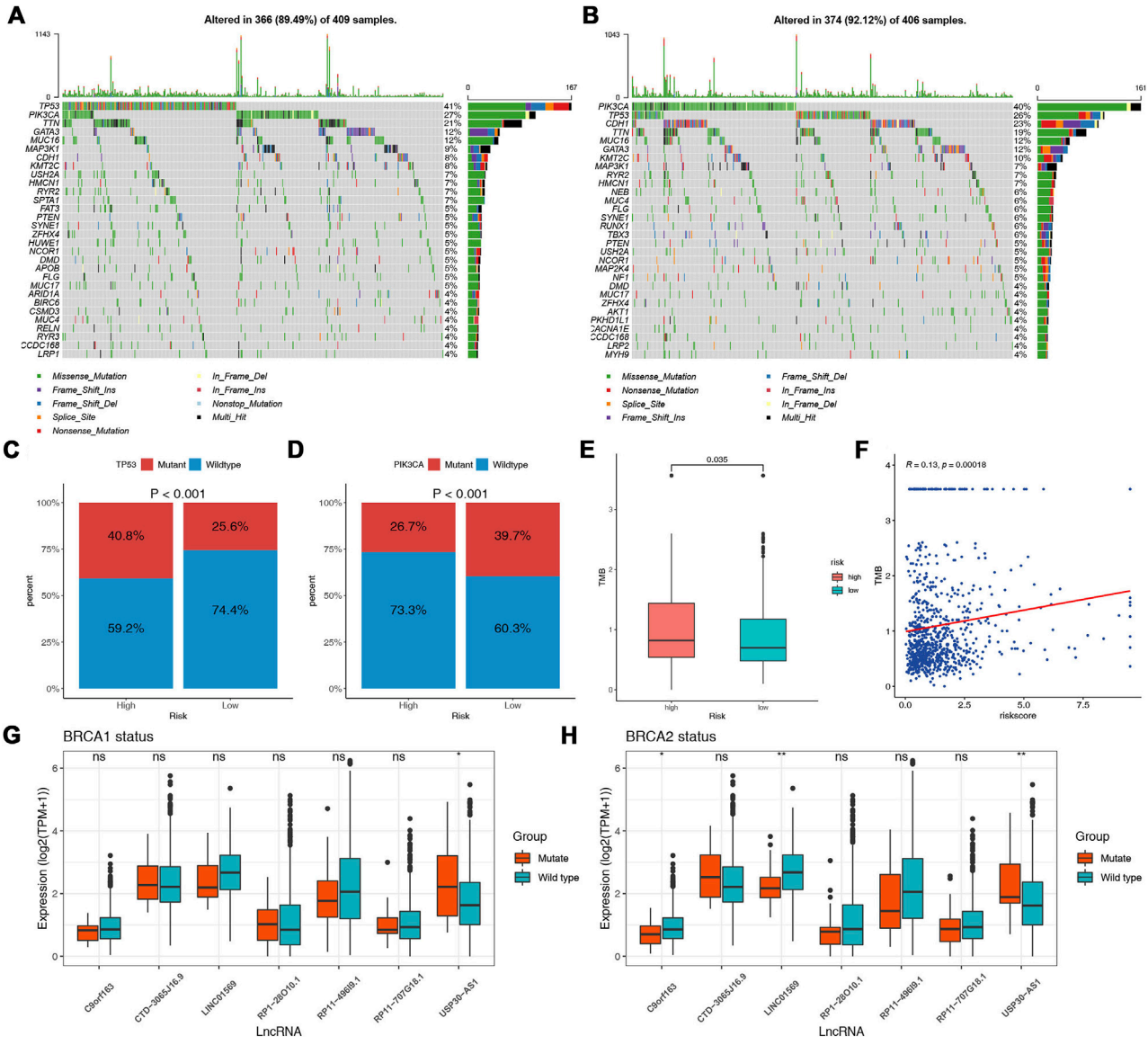


FIGURE 8
Association between DNA mutation and prognostic model. Waterfall plots of the top 30 mutated genes in the high-risk **(A)** and low-risk **(B)** groups. Comparisons of the mutation status of TP53 **(C)** and PIK3CA **(D)** in different risk groups. **(E)** Comparisons of the TMB between the two risk groups. **(F)** Correlation between TMB and the risk score. **(G)** Comparisons of the expression of the lncRNAs in different BRCA1 mutation statuses. **(H)** Comparisons of the expression of the lncRNAs in different BRCA2 mutation statuses.

the two risk groups (Figure 7A). The macrophages M0, M2, and NK cells resting were the main components in the high-risk group; However, the resting CD4 T memory cells, CD8 T cells, naive B cells, activated dendritic cells, monocytes, and gamma delta T cells were mainly in the low-risk group (Figure 7B). The ESTIMATE results showed that high-risk patients had lower stromal and immune scores but had higher tumor purity (Figure 7C). Furthermore, the risk score was negatively associated with the stromal and immune scores while positively associated with tumor purity (Figure 7D). ICP was proved related to immunotherapy (Topalian et al., 2015). We assessed the expression levels of 27 ICPs in the two risk groups. They all expressed much higher in the low-risk group, such as CTLA4, HAVCR2, TIGIT, PDCD1, and LAG3 (Figure 7E). TIDE could identify the patients' response to ICIs. As shown in Figure 7F, the low-risk group had a significantly higher response rate to immunotherapy. The risk score for non-responders to immunotherapy tended to be much higher than that for responders (Figure 7G). Furthermore, all four

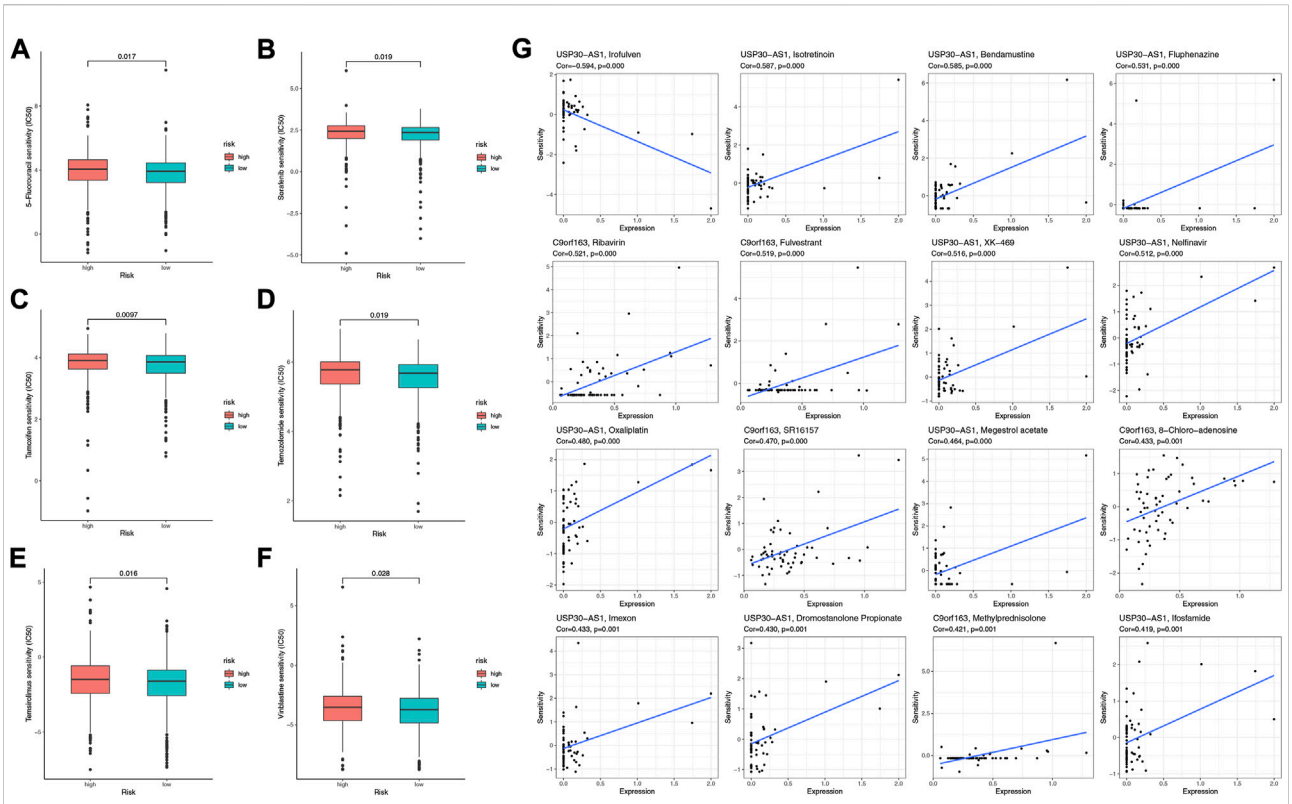


FIGURE 9
The sensitivity of chemotherapeutic agents and the prediction of potential drugs. The IC50 values of six chemotherapy and targeted agents in the two risk groups, including 5-Fluorouracil **(A)**, Sorafenib **(B)**, Tamoxifen **(C)**, Temozolomide **(D)**, Temsirolimus **(E)**, and Vinblastine **(F)**. **(G)** Sensitivity correlation analyses of the LRLs and potential drugs according to the CellMiner Database.

types of IPS were higher in the low-risk group, indicating that the low-risk patients could acquire more benefits from ICIs (Figures 7H–K). These results indicated that the low-risk group with the immune signature might have a better immunotherapy response.

In addition, we investigated whether the seven LRLs in our signature were associated with the immune signature. USP30-AS1 was significantly positively related to activated CD4 memory T cells, Macrophages M1 and CD8 T cells, and classic ICPs, such as PD-1 and CTLA4 (Supplementary Figures S1A,B). These results indicated that USP30-AS1 might make a difference in the TIME.

Somatic mutation analysis

The potential contribution of genomic changes to tumor immunity and immune infiltration has been explored in previous studies (Rooney et al., 2015; Thorsson et al., 2018). Figures 8A,B showed the top 30 genes mutated most frequently in the two risk groups. Although the overall mutation frequency is similar between the two groups (high vs. low, 89.49% vs. 92.12%), about one-third of the genes are different. SPTA1, APOB, ARID1A, BIRC6, GSMD3, RELN, RYR3, LRP1, and HUWE1 were not observed in the low-risk group. Regarding the most common BRCA biomarkers, the high-risk group had significantly more patients with TP53 mutations (low vs. high, 25.6 vs. 40.8%) (Figure 8C). A higher mutation frequency of PIK3CA was observed in the low-risk group (low vs. high, 39.7 vs. 26.7%) (Figure 8D). Further, the TMB in the high-risk group was significantly higher (p = 0.035) (Figure 8E). The Pearson correlation analysis indicated a positive correlation between TMB and the risk score (Figure 8F).

Mutations in the BRCA1 and BRCA2 genes are known risk factors and drivers of breast cancer (Kuchenbaecker et al., 2017). As to the lncRNAs identified as a molecular signature in our analysis, we further explore the associations between the expression of the lncRNAs and BRCA1/2 mutation status. We found that the expression of USP30-AS1 was significantly increased in the BRCA1/2 mutant group, and in addition, C9orf163 and LINC01569 were significantly decreased in the BRCA2 mutant group (Figures 8G,H). We did not find an association between the expression of other lncRNAs and BRCA1/2 mutation status.

Prediction of potential drugs and the sensitivity of chemotherapeutic agents

To further explore effective drugs for BC patients to guide precision treatment, we analyzed the sensitivity to common chemotherapeutic agents of the two risk groups. High-risk patients had higher IC50 of 5-Fluorouracil, Sorafenib, Tamoxifen, Temozolomide, Temsirolimus, and Vinblastine (Figures 9A–F), indicating they were more likely to be resistant to the above drugs. Furthermore, we explored the potential drugs targeted to the seven model genes with the CellMiner database. We finally acquired 16 gene-drug correlations, of which 11 correlations pointed to the USP30-AS1, and five correlations pointed to the C9orf163 (Figure 9G). C9orf163 expressed higher in the high-risk group, while C9orf163 expressed higher in the low-risk group. Ribavirin, Fulvestrant, SR16157, 8-Chloro-adenosine, and Methylprednisolone were positively related to C9orf163 so they might benefit high-risk patients. Conversely, the BRCA drug Ifosfamide was positively correlated with USP30-AS1; it might benefit low-risk patients (Figure 9G).

Discussion

BC has the highest incidence rate among all cancers globally, which causes tens of thousands of female deaths yearly (Harbeck et al., 2019). BC is characterized by tumor heterogeneity at the molecular level of tumor cells and the tumor microenvironment (TME) (Baker et al., 2018; Sousa et al., 2019). Tumor heterogeneity complicates the aggressiveness and treatment of BC (Januškevičienė and Petrikaitė, 2019). Recent studies have revealed lactate's diverse roles in the TME. Although cancer cells have a sufficient oxygen supply, they still use glucose and produce lactate excessively, which could cause acidosis, angiogenesis, and immunosuppression (Ho et al., 2019). In BC, lactate is correlated with resistance to PI3K inhibitors (Hillis and TOKER, 2020). In several cancers, lactate is essential in predicting prognosis, tumor microenvironment, and immune response (Mai et al., 2022; Sun et al., 2022; Xie et al., 2022). However, the prognostic value of lactate in BC remains largely unknown. This is the first study investigating the role of lactate in predicting prognosis, immune status, and therapeutic response in BC.

We first identified 196 differential expression LRLs for further study. We used the univariate Cox regression analysis, LASSO, and multivariate Cox regression analysis to construct the LRLPS. Survival analysis and the time-dependent ROC curves confirmed the prognostic value and reliability of the LRLPS. The AUC of the risk score was higher than other clinicopathological characteristics, indicating the highest prognostic performance of the LRLPS. Subsequent univariate and multivariate Cox regression analyses further indicated the independent prognostic predictability of the risk score. Stratified analysis showed that the LRLPS was suitable for patients in any clinical subgroup. The predictive ability of our signature was further explored by comparing it with various molecular signatures commonly used to predict OS in BC patients. Our signature displayed much higher AUCs than ferroptosis, necroptosis, pyroptosis, and immune-related signatures, which implied a stronger predictive ability, especially in predicting short-term survival status. Furthermore, the nomogram provided a powerful tool for clinicians to make decisions.

The GO/KEGG and GSEA indicated that the immune-related pathways differed between the two-risk groups. Previous research has demonstrated that lactate could regulate TMB. Through its ability to enhance the metabolic profile of the Treg and maintain acidity in the TME, lactate could enhance the immunosuppressive effect (Dastmalchi et al., 2021). Excessive lactate inhibits T-cell proliferation, such as Natural killer, dendritic, and $CD8^+$ T cells (Haas et al., 2015; Certo et al., 2021; Grote et al., 2021). In addition, lactate could potentiate the anti-inflammatory effects by activating macrophages, promoting angiogenesis, tissue remodeling, and accelerating tumor growth and invasion (Certo et al., 2021). Hence, we further explore TIME through several algorithms. Tumor immune cell infiltration (TIICs) is a crucial component of the TIME. We calculated the levels of TIICs in BC with CIBERSORT. The high-risk group was enriched with immunosuppressive immune cells, such as macrophages M0 and M2, which were also critical members of EMT and cancer metastasis (Biswas and Mantovani, 2010; Qian and Pollard, 2010). Instead, $CD4/8^+$ T cells, the vital factors in killing tumors and promoting immune response, were the main component in the low-risk group (Charoentong et al., 2017). According to the ESTIMATE analysis, the low-risk group had a higher immune score and stromal content while lower tumor purity than the high-risk group.

Immunotherapy has been a new treatment modality in BC, especially for metastatic BC (Adams et al., 2019). We further estimated the immunotherapy responses of the two risk groups. It is reported that ICIs antitumor relay on the $CD8^+$ T cells, $CD4^+$ T cells, and dendritic cells (Sato et al., 2018; Farhood et al., 2019). The immune cell infiltration levels were positively correlated with the responsiveness to ICIs (Karn et al., 2017; Kümpers et al., 2019). As an essential biomarker for predicting cancer immunotherapy (Patel and Kurzrock, 2015), the 27 ICPs expressed higher in the low-risk group. Therefore, we speculated that the low-risk group could respond better to immunotherapy and further verified the conclusion through TIDE and IPS analyses. All IPSs of CTLA4-/PD-1-, CTLA4 + /PD-1-, CTLA4-/PD-1 +, and CTLA4+/PD-1 + were higher in the low-risk group, indicted that the low-risk group had a better response to immunotherapy. Patients with high risk had the higher TMB in our study. Some research has indicated that TMB could act as a biomarker for predicting the response to ICIs (Lu et al., 2019; Shum et al., 2022). However, the predictive value varies among different cancers and might be insufficient in solid tumors (Xu et al., 2019). Thus, ICIs could benefit low-risk patients, while other immunotherapy might be appropriate for high-risk patients. These results indicated the significant differences in the degree of immune cell infiltration and immunotherapy response between the two risk groups identified by lactate-related signature.

Regarding the seven LRLs in our signature, some have been studied before in other cancers. USP30-AS1 is involved in autophagy, proliferation, and apoptosis in acute myeloid leukemia, glioblastoma, and cervical cancer (Chen et al., 2021; Wang et al., 2021; Zhou et al., 2022). In our study, USP30-AS1 was positively correlated with the antitumor immune cells and the classic ICPs. These results indicated the potential role of USP30-AS1 in TME. C9orf163 could develop the tumor microenvironment through cytokine and chemokine signaling and might act as a tumor suppressor in anaplastic gliomas and pancreatic cancer (Wang et al., 2016; Zhuang et al., 2020). Furthermore, we found that the expression of USP30-AS1 and C9orf163 were associated with BRCA1/2 status, indicating that they were involved in breast cancer development. However, more research is required to clarify the molecular mechanism of the seven LRLs in BC.

In treating BC, chemotherapeutic drugs could reduce tumor recurrence and be a primary treatment option for metastatic disease. However, chemo-resistance severely limited the clinical efficacy of chemotherapeutic drugs for BC patients (O'driscoll and Clynes, 2006). Thus, we assessed the BC patients' response to chemotherapy with the IC50 value. The pRRophetic showed that low-risk patients were more sensitive to the common chemotherapy drugs, such as 5-Fluorouracil, Sorafenib, Tamoxifen, Temozolomide, Temsirolimus, and Vinblastine. Furthermore, we performed the CellMiner database to predict the candidate small-molecule compounds. The results indicated that Ribavirin, Fulvestrant, SR16157, 8-Chloro-adenosine, and Methylprednisolone might benefit high-risk patients. In combination, these discoveries may provide BC patients with suitable treatment options.

However, there were still a few limitations to our study. We used the TCGA dataset for all analyses since other databases lacked the needed LRLs data, including the Gene Expression Omnibus (GEO) and METABRIC databases, which prevented us from verifying the results. Therefore, it is better to validate in a prospective cohort. Secondly, further studies on the biological functions of the seven LRLs are needed to be performed *in vivo* and *in vitro*.

Conclusion

Altogether, this study identified a novel lactate-related lncRNAs prognostic signature for BC patients, which could predict the prognosis and immune infiltration. The LRLPS also provided an effective method for personalized risk estimation and assessment of treatment response to immunotherapy and chemotherapy, which may be clinically helpful. Finally, the seven LRLs could become potential treatment targets for BC.

Data availability statement

The data can be found here: https://portal.gdc.cancer.gov/repository.

Author contributions

YZ, JL, and CL conceptualized and designed the study. CF, HW, WW, and HQ prepared the dataset. YC and YZ, YJ analyzed and interpreted the data. FW and XL wrote and reviewed the manuscript. SZ supervised the study. All authors contributed to the article and approved the submitted version.

Funding

This work is supported by National Natural Science Foundation of China 82103212.

Acknowledgments

We thank the TCGA database for providing the original study data.

Conflict of Interest

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

Publisher's note

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/fgene.2022.956246/full#supplementary-material

References

Adams, S., Gatti-Mays, M. E., Kalinsky, K., Korde, L. A., Sharon, E., Amiri-Kordestani, L., et al. (2019). Current landscape of immunotherapy in breast cancer: A review. *JAMA Oncol.* 5 (8), 1205–1214. doi:10.1001/jamaoncol.2018.7147

Baker, J. H. E., Kyle, A. H., Reinsberg, S. A., Moosvi, F., Patrick, H. M., Cran, J., et al. (2018). Heterogeneous distribution of trastuzumab in HER2-positive xenografts and metastases: Role of the tumor microenvironment. *Clin. Exp. Metastasis* 35 (7), 691–705. doi:10.1007/s10585-018-9929-3

Biswas, S. K., and Mantovani, A. (2010). Macrophage plasticity and interaction with lymphocyte subsets: Cancer as a paradigm. *Nat. Immunol.* 11 (10), 889–896. doi:10.1038/ni.1937

Certo, M., Tsai, C-H., Pucino, V., Ho, P. C., and Mauro, C. (2021). Lactate modulation of immune responses in inflammatory versus tumour microenvironments. *Nat. Rev. Immunol.* 21 (3), 151–161. doi:10.1038/s41577-020-0406-2

Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., et al. (2017). Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 18 (1), 248–262. doi:10.1016/j.celrep.2016.12.019

Chen, F., Yang, J., Fang, M., Wu, Y., Su, D., and Sheng, Y. (2022). Necroptosis-related lncRNA to establish novel prognostic signature and predict the immunotherapy response in breast cancer. *J. Clin. Lab. Anal.* 36 (4), e24302. doi:10.1002/jcla.24302

Chen, M., Chi, Y., Chen, H., and Zhao, L. (2021). Long non-coding RNA USP30-AS1 aggravates the malignant progression of cervical cancer by sequestering microRNA-299-3p and thereby overexpressing PTP4A1. *Oncol. Lett.* 22 (1), 505. doi:10.3892/ol.2021.12766

Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAbiolinks: An R/bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 44 (8), e71. doi:10.1093/nar/gkv1507

Corbet, C., and Feron, O. (2017). Tumour acidosis: From the passenger to the driver's seat. *Nat. Rev. Cancer* 17 (10), 577–593. doi:10.1038/nrc.2017.77

Dai, Y-W., Wen, Z-K., Wu, Z-X., Wu, H. D., Lv, L. X., Yan, C. Z., et al. (2022). Amino acid metabolism-related lncRNA signature predicts the prognosis of breast cancer. *Front. Genet.* 13, 880387. doi:10.3389/fgene.2022.880387

Dastmalchi, F., Deleyrolle, L. P., Karachi, A., Mitchell, D. A., and Rahman, M. (2021). Metabolomics monitoring of treatment response to brain tumor immunotherapy. *Front. Oncol.* 11, 691246. doi:10.3389/fonc.2021.691246

Doherty, J. R., and Cleveland, J. L. (2013). Targeting lactate metabolism for cancer therapeutics. *J. Clin. Invest.* 123 (9), 3685–3692. doi:10.1172/JCI69741

Erra díaz, F., Dantas, E., and Geffner, J. Unravelling the interplay between extracellular acidosis and immune cells [J]. *Mediat. Inflamm.*, 2018, 2018: 1218297, doi:10.1155/2018/1218297

Esteller, M. (2011). Non-coding RNAs in human disease. *Nat. Rev. Genet.* 12 (12), 861–874. doi:10.1038/nrg3074

Farhood, B., Najafi, M., and Mortezaee, K. (2019). CD8+ cytotoxic T lymphocytes in cancer immunotherapy: A review. *J. Cell. Physiol.* 234 (6), 8509–8521. doi:10.1002/jcp.27782

Fatica, A., and Bozzoni, I. (2014). Long non-coding RNAs: New players in cell differentiation and development. *Nat. Rev. Genet.* 15, 7–21. doi:10.1038/nrg3606

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognit. Lett.* 27 (8), 861–874. doi:10.1016/j.patrec.2005.10.010

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33 (1)–22. doi:10.18637/jss.v033.i01

Fu, J., Li, K., Zhang, W., Wan, C., Zhang, J., Jiang, P., et al. (2020). Large-scale public data reuse to model immunotherapy response and resistance. *Genome Med.* 12 (1), 21. doi:10.1186/s13073-020-0721-z

Gabrilovich, D. I., Ostrand-Rosenberg, S., and Bronte, V. (2012). Coordinated regulation of myeloid cells by tumours. *Nat. Rev. Immunol.* 12 (4), 253–268. doi:10.1038/nri3175

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

Girgis, H., Masui, O., White, N. M., Scorilas, A., Rotondo, F., Seivwright, A., et al. (2014). Lactate dehydrogenase A is a potential prognostic marker in clear cell renal cell carcinoma. *Mol. Cancer* 13, 101. doi:10.1186/1476-4598-13-101

Grote, S., UreñA-BailéN, G., Chan, K. C-H., Baden, C., Mezger, M., Handgretinger, R., et al. (2021). *In vitro* evaluation of cd276-CAR NK-92 functionality, migration and invasion potential in the presence of immune inhibitory factors of the tumor microenvironment. *Cells* 10 (5), 1020. doi:10.3390/cells10051020

Gu, P., Zhang, L., Wang, R., Ding, W., Wang, W., Liu, Y., et al. (2021). Development and validation of a novel hypoxia-related long noncoding RNA model with regard to prognosis and immune features in breast cancer. *Front. Cell Dev. Biol.* 9, 796729. doi:10.3389/fcell.2021.796729

Haas, R., Smith, J., Rocher-Ros, V., Nadkarni, S., Montero-Melendez, T., D'Acquisto, F., et al. (2015). Lactate regulates metabolic and pro-inflammatory circuits in control of T cell migration and effector functions. *PLoS Biol.* 13 (7), e1002202. doi:10.1371/journal.pbio.1002202

Harbeck, N., and Gnant, M. (2017). Breast cancer. *Lancet* 389 (10074), 1134–1150. doi:10.1016/S0140-6736(16)31891-8

Harbeck, N., Penault-Llorca, F., Cortes, J., Gnant, M., Houssami, N., Poortmans, P., et al. (2019). Breast cancer. *Nat. Rev. Dis. Prim.* 5 (1), 66. doi:10.1038/s41572-019-0111-2

Hauptman, N., and Glavač, D. (2013). Long non-coding RNA in cancer. *Int. J. Mol. Sci.* 14 (3), 4655–4669. doi:10.3390/ijms14034655

Hayes, C., Donohoe, C. L., Davern, M., and Donlon, N. E. (2021). The oncogenic and clinical implications of lactate induced immunosuppression in the tumour microenvironment. *Cancer Lett.* 500, 75–86. doi:10.1016/j.canlet.2020.12.021

Hillis, A. L., and Toker, A. (2020). Lactate lights up PI3K inhibitor resistance in breast cancer. *Cancer Cell* 38 (4), 441–443. doi:10.1016/j.ccell.2020.09.011

Ho, Y-J., Chu, S-W., Liao, E-C., Fan, C. H., Chan, H. L., Wei, K. C., et al. (2019). Normalization of tumor vasculature by oxygen microbubbles with ultrasound. *Theranostics* 9 (24), 7370–7383. doi:10.7150/thno.37750

Huang, S-L., Huang, Z-C., Zhang, C-J., Xie, J., Lei, S. S., Wu, Y. Q., et al. (2022). LncRNA SNHG5 promotes the glycolysis and proliferation of breast cancer cell through regulating BACH1 via targeting miR-299. *Breast Cancer* 29 (1), 65–76. doi:10.1007/s12282-021-01281-6

Izzo, L. T., and Wellen, K. E. (2019). Histone lactylation links metabolism and gene regulation. *Nature* 574 (7779), 492–493. doi:10.1038/d41586-019-03122-1

Jabbarzadeh kaboli, P., Salimian, F., Aghapour, S., Xiang, S., Zhao, Q., Li, M., et al. (2020). Akt-targeted therapy as a promising strategy to overcome drug resistance in breast cancer - a comprehensive review from chemotherapy to immunotherapy. *Pharmacol. Res.* 156, 104806. doi:10.1016/j.phrs.2020.104806

Januškevičienė, I., and Petrikaitė, V. (2019). Heterogeneity of breast cancer: The importance of interaction between different tumor cell populations. *Life Sci.* 239, 117009. doi:10.1016/j.lfs.2019.117009

Jia, C. L., Yang, F., and Li, R. (2021). Prognostic model construction and immune microenvironment analysis of breast cancer based on ferroptosis-related lncRNAs. *Int. J. Gen. Med.* 14, 9817–9831. doi:10.2147/IJGM.S342783

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24 (10), 1550–1558. doi:10.1038/s41591-018-0136-1

Jin, Z., Lu, Y., Wu, X., Pan, T., Yu, Z., Hou, J., et al. (2021). The cross-talk between tumor cells and activated fibroblasts mediated by lactate/BDNF/TrkB signaling promotes acquired resistance to anlotinib in human gastric cancer. *Redox Biol.* 46, 102076. doi:10.1016/j.redox.2021.102076

Karn, T., Jiang, T., Hatzis, C., Sanger, N., El-Balat, A., Rody, A., et al. (2017). Association between genomic metrics and immune infiltration in triple-negative breast cancer. *JAMA Oncol.* 3 (12), 1707–1711. doi:10.1001/jamaoncol.2017.2140

Kaur, M., Kaur, B., Konar, M., and Sharma, S. (2022). Noncoding RNAs as novel immunotherapeutic tools against cancer. *Adv. Protein Chem. Struct. Biol.* 129, 135–161. doi:10.1016/bs.apcsb.2021.11.011

Kelderman, S., Heemskerk, B., Van Tinteren, H., van den Brom, R. R. H., Hospers, G. A. P., van den Eertwegh, A. J. M., et al. (2014). Lactate dehydrogenase as a selection criterion for ipilimumab treatment in metastatic melanoma. *Cancer Immunol. Immunother.* 63 (5), 449–458. doi:10.1007/s00262-014-1528-9

Kuchenbaecker, K. B., Hopper, J. L., Barnes, D. R., Phillips, K. A., Mooij, T. M., Roos-Blom, M. J., et al. (2017). Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. *JAMA* 317 (23), 2402–2416. doi:10.1001/jama.2017.7112

Kümpers, C., Jokic, M., Haase, O., Offermann, A., Vogel, W., Gratz, V., et al. (2019). Immune cell infiltration of the primary tumor, not PD-L1 status, is associated with improved response to checkpoint inhibition in metastatic melanoma. *Front. Med.* 6, 27. doi:10.3389/fmed.2019.00027

Li, J., Liu, C., Chen, Y., Gao, C., Wang, M., Ma, X., et al. (2019). Tumor characterization in breast cancer identifies immune-relevant gene signatures associated with prognosis. *Front. Genet.* 10, 1119. doi:10.3389/fgene.2019.01119

Li, X., Chen, J., Yu, Q., Huang, H., Liu, Z., Wang, C., et al. (2021). A signature of autophagy-related long non-coding RNA to predict the prognosis of breast cancer. *Front. Genet.* 12, 569318. doi:10.3389/fgene.2021.569318

Liberzon, A., Birger, C., ThorvaldsdóTTIR, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1 (6), 417–425. doi:10.1016/j.cels.2015.12.004

Longhitano, L., Forte, S., Orlando, L., Grasso, S., Barbato, A., Vicario, N., et al. (2022). The crosstalk between GPR81/IGFBP6 promotes breast cancer progression by modulating lactate metabolism and oxidative stress. *Antioxidants* 11, 275. doi:10.3390/antiox11020275

Lu, S., Stein, J. E., Rimm, D. L., Wang, D. W., Bell, J. M., Johnson, D. B., et al. (2019). Comparison of biomarker modalities for predicting response to PD-1/PD-L1 checkpoint blockade: A systematic review and meta-analysis. *JAMA Oncol.* 5 (8), 1195–1204. doi:10.1001/jamaoncol.2019.1549

Lv, W., Wang, Y., Zhao, C., Tan, Y., Xiong, M., Yi, Y., et al. (2021). Identification and validation of m6A-related lncRNA signature as potential predictive biomarkers in breast cancer. *Front. Oncol.* 11, 745719. doi:10.3389/fonc.2021.745719

Mai, S., Liang, L., Mai, G., Liu, X., Diao, D., Cai, R., et al. (2022). Development and validation of lactate metabolism-related lncRNA signature as a prognostic model for lung adenocarcinoma. *Front. Endocrinol.* 13, 829175. doi:10.3389/fendo.2022.829175

Mu, X., Shi, W., Xu, Y., Xu, C., Zhao, T., Geng, B., et al. (2018). Tumor-derived lactate induces M2 macrophage polarization via the activation of the ERK/STAT3 signaling pathway in breast cancer. *Cell Cycle* 17 (4), 428–438. doi:10.1080/15384101.2018.1444305

Nakagawa, Y., Negishi, Y., Shimizu, M., Takahashi, M., Ichikawa, M., and Takahashi, H. (2015). Effects of extracellular pH and hypoxia on the function and development of antigen-specific cytotoxic T lymphocytes. *Immunol. Lett.* 167 (2), 72–86. doi:10.1016/j.imlet.2015.07.003

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Ning, S., Wu, J., Pan, Y., Qiao, K., Li, L., and Huang, Q. (2022). Identification of $CD4^+$ conventional T cells-related lncRNA signature to improve the prediction of prognosis and immunotherapy response in breast cancer. *Front. Immunol.* 13, 880769. doi:10.3389/fimmu.2022.880769

O'driscoll, L., and Clynes, M. (2006). Biomarkers and multiple drug resistance in breast cancer. *Curr. Cancer Drug Targets* 6 (5), 365–384. doi:10.2174/156800906777723958

Patel, S. P., and Kurzrock, R. (2015). PD-L1 expression as a predictive biomarker in cancer immunotherapy. *Mol. Cancer Ther.* 14 (4), 847–856. doi:10.1158/1535-7163.MCT-14-0983

Qian, B-Z., and Pollard, J. W. (2010). Macrophage diversity enhances tumor progression and metastasis. *Cell* 141 (1), 39–51. doi:10.1016/j.cell.2010.03.014

Rodríguez bautista, R., Ortega, G. A., Hidalgo Miranda, A., Zentella Dehesa, A., Villarreal-Garza, C., Avila-Moreno, F., et al. (2018). Long non-coding RNAs: Implications in targeted diagnoses, prognosis, and improved therapeutic strategies in human non- and triple-negative breast cancer. *Clin. Epigenetics* 10, 88. doi:10.1186/s13148-018-0514-z

Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 160, 48–61. doi:10.1016/j.cell.2014.12.033

Sato, Y., Bolzenius, J. K., Eteleeb, A. M., Su, X., Maher, C. A., Sehn, J. K., et al. (2018). CD4+ T cells induce rejection of urothelial tumors after immune checkpoint blockade. *JCI Insight* 3, 121062. doi:10.1172/jci.insight.121062

Shankavaram, U. T., Reinhold, W. C., Nishizuka, S., Major, S., Morita, D., Chary, K. K., et al. (2007). Transcript and protein expression profiles of the NCI-60 cancer cell panel: An integromic microarray study. *Mol. Cancer Ther.* 6 (3), 820–832. doi:10.1158/1535-7163.MCT-06-0650

Shankavaram, U. T., Varma, S., Kane, D., Sunshine, M., Chary, K. K., Reinhold, W. C., et al. (2009). CellMiner: A relational database and query tool for the NCI-60 cancer cell lines. *BMC Genomics* 10, 277. doi:10.1186/1471-2164-10-277

Shum, B., Larkin, J., and Turajlic, S. Predictive biomarkers for response to immune checkpoint inhibition [J]. *Semin. Cancer Biol.*, 2022, 79, 4–17. doi:10.1016/j.semcancer.2021.03.036

Sousa, B., Ribeiro, A. S., and Paredes, J. (2019). Heterogeneity and plasticity of breast cancer stem cells [J]. *Adv. Exp. Med. Biol.* 1139, 83–103. doi:10.1007/978-3-030-14366-4_5

Stone, S. C., Rossetti, R. A. M., Alvarez, K. L. F., Carvalho, J. P., Margarido, P. F. R., Baracat, E. C., et al. (2019). Lactate secreted by cervical cancer cells modulates macrophage phenotype. *J. Leukoc. Biol.* 105 (5), 1041–1054. doi:10.1002/JLB.3A0718-274RR

Subramanian, A., Kuehn, H., Gould, J., Tamayo, P., and Mesirov, J. P. (2007). GSEA-P: A desktop application for gene set enrichment analysis. *Bioinformatics* 23 (23), 3251–3253. doi:10.1093/bioinformatics/btm369

Sun, Z., Tao, W., Guo, X., Jing, C., Zhang, M., Wang, Z., et al. (2022). Construction of a lactate-related prognostic signature for predicting prognosis, tumor microenvironment, and immune response in kidney renal clear cell carcinoma. *Front. Immunol.* 13, 818984. doi:10.3389/fimmu.2022.818984

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The immune landscape of cancer. *Immunity* 48 (4), 812–830. doi:10.1016/j.immuni.2018.03.023

Topalian, S. L., Drake, C. G., and Pardoll, D. M. (2015). Immune checkpoint blockade: A common denominator approach to cancer therapy. *Cancer Cell* 27 (4), 450–461. doi:10.1016/j.ccell.2015.03.001

Waks, A. G., and Winer, E. P. (2019). Breast cancer treatment: A review. *JAMA* 321 (3), 288–300. doi:10.1001/jama.2018.19323

Wang, N., Li, J., Xin, Q., and Xu, N. (2021). USP30-AS1 contributes to mitochondrial quality control in glioblastoma cells. *Biochem. Biophys. Res. Commun.* 581, 31–37. doi:10.1016/j.bbrc.2021.10.006

Wang, W., Yang, F., Zhang, L., Chen, J., Zhao, Z., Wang, H., et al. (2016). LncRNA profile study reveals four-lncRNA signature associated with the prognosis of patients with anaplastic gliomas. *Oncotarget* 7 (47), 77225–77236. doi:10.18632/oncotarget.12624

Wang, X., Zhang, B., Chen, X., Mo, H., Wu, D., Lan, B., et al. (2019). Lactate dehydrogenase and baseline markers associated with clinical outcomes of advanced esophageal squamous cell carcinoma patients treated with camrelizumab (SHR-1210), a novel anti-PD-1 antibody. *Thorac. Cancer* 10 (6), 1395–1401. doi:10.1111/1759-7714.13083

Wei, T., Zhu, N., Jiang, W., and Xing, X. L. (2022). Development and validation of ferroptosis- and immune-related lncRNAs signatures for breast infiltrating duct and lobular carcinoma. *Front. Oncol.* 12, 844642. doi:10.3389/fonc.2022.844642

Xia, K-G., Wang, C-M., Shen, D-Y., Song, X-Y., Mu, X-Y., Zhou, J-W., et al. LncRNA NEAT1-associated aerobic glycolysis blunts tumor immunosurveillance by T cells in prostate cancer [J]. *Neoplasma*, 2022.69, 594–602. doi:10.4149/neo_2022_211021N1497

Xie, Y., Zhang, J., Li, M., Zhang, Y., Li, Q., Zheng, Y., et al. (2022). Identification of lactate-related gene signature for prediction of progression and immunotherapeutic response in skin cutaneous melanoma. *Front. Oncol.* 12, 818868. doi:10.3389/fonc.2022.818868

Xu, S., Xie, J., Zhou, Y., Liu, H., Wang, Y., and Li, Z. (2022). Integrated analysis of RNA binding protein-related lncRNA prognostic signature for breast cancer patients. *Genes (Basel)* 13, 345. doi:10.3390/genes13020345

Xu, Z., Dai, J., Wang, D., Lu, H., Dai, H., Ye, H., et al. (2019). Assessment of tumor mutation burden calculation from gene panel sequencing data. *Onco. Targets. Ther.* 12, 3401–3409. doi:10.2147/OTT.S196638

Yang, W., Soares, J., Greninger, P., Edelman, E. J., Lightfoot, H., Forbes, S., et al. (2013). Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* 41, D955–D961. doi:10.1093/nar/gks1111

Yang, X., Weng, X., Yang, Y., and Jiang, Z. (2021). Pyroptosis-related lncRNAs predict the prognosis and immune response in patients with breast cancer. *Front. Genet.* 12, 792106. doi:10.3389/fgene.2021.792106

Yoshihara, K., Shahmoradgoli, M., MartíNEZ, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612

Yu, G., Wang, L-G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Zhang, D., Tang, Z., Huang, H., Zhou, G., Cui, C., Weng, Y., et al. (2019). Metabolic regulation of gene expression by histone lactylation. *Nature* 574 (7779), 575–580. doi:10.1038/s41586-019-1678-1

Zhang, Y., Yue, Q., Cao, F., Li, Y., and Wei, Y. (2022). Necroptosis-related lncRNA signatures determine prognosis in breast cancer patients. *Sci. Rep.* 12 (1), 11268. doi:10.1038/s41598-022-15209-3

Zhang, Z., Li, Y., Yan, X., Song, Q., Wang, G., Hu, Y., et al. (2019). Pretreatment lactate dehydrogenase may predict outcome of advanced non small-cell lung cancer patients treated with immune checkpoint inhibitors: A meta-analysis. *Cancer Med.* 8 (4), 1467–1473. doi:10.1002/cam4.2024

Zhao, Y., Liu, L., Zhao, J., Du, X., Yu, Q., Wu, J., et al. (2021). Construction and verification of a hypoxia-related 4-lncRNA model for prediction of breast cancer. *Int. J. Gen. Med.* 14, 4605–4617. doi:10.2147/IJGM.S322007

Zhou, W., Xu, S., Deng, T., Zhou, R., and Wang, C. (2022). LncRNA USP30-AS1 promotes the survival of acute myeloid leukemia cells by cis-regulating USP30 and ANKRD13A. *Hum. Cell* 35 (1), 360–378. doi:10.1007/s13577-021-00636-7

Zhuang, H., Huang, S., Zhou, Z., Ma, Z., Zhang, Z., Zhang, C., et al. (2020). A four prognosis-associated lncRNAs (PALnc) based risk score system reflects immune cell infiltration and predicts patient survival in pancreatic cancer. *Cancer Cell Int.* 20, 493. doi:10.1186/s12935-020-01588-y

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 26 October 2022
DOI 10.3389/fgene.2022.966941

Check for updates

OPEN ACCESS

EDITED BY
Claudia Cava,
National Research Council (CNR), Italy

REVIEWED BY
Mohamed Elbadawy,
Benha University, Egypt
Maarten Fokke Bijlsma,
Academic Medical Center, Netherlands

*CORRESPONDENCE
Sara Taleahmad,
sara_taleahmad@yahoo.com
Massoud Vosough,
masvos@yahoo.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 11 June 2022
ACCEPTED 10 October 2022
PUBLISHED 26 October 2022

CITATION
Ramezankhani R, Ghavidel AA, Rashidi S, Rojhannezhad M, Abolkheir HR, Mirhosseini M, Taleahmad S and Vosough M (2022), Gender-related differentially expressed genes in pancreatic cancer: possible culprits or accomplices?
Front. Genet. 13:966941.
doi: 10.3389/fgene.2022.966941

COPYRIGHT
© 2022 Ramezankhani, Ghavidel, Rashidi, Rojhannezhad, Abolkheir, Mirhosseini, Taleahmad and Vosough. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Gender-related differentially expressed genes in pancreatic cancer: possible culprits or accomplices?

Roya Ramezankhani[1,2†], Afshin Abdi Ghavidel[3†], Saadyeh Rashidi[4], Mahbubeh Rojhannezhad[5], Hamid Reza Abolkheir[6], Malihe Mirhosseini[4], Sara Taleahmad[7]* and Massoud Vosough[8,9]*

[1]Department of Development and Regeneration, Stem Cell Biology and Embryology, KU Leuven Stem Cell Institute, Leuven, Belgium, [2]Department of Applied Cell Sciences, Faculty of Basic Science and Advanced Medical Technologies, Royan Institute, ACECR, Tehran, Iran, [3]Student Research Committee, School of Advanced Technologies in Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran, [4]Department of Biology, Faculty of Biological Sciences and Technology, Shahid Beheshti University, Tehran, Iran, [5]Department of Genetics, Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran, [6]School of Biomedical Engineering, University of Technology Sydney, Sydney, NSW, Australia, [7]Department of Stem Cells and Developmental Biology, Cell Science Research Center, Royan Institute for Stem Cell Biology and Technology, ACECR, Tehran, Iran, [8]Department of Regenerative Medicine, Cell Science Research Center, Royan Institute for Stem Cell Biology and Technology, Academic Center for Education, Culture and Research, ACECR, Tehran, Iran, [9]Experimental Cancer Medicine, Institution for Laboratory Medicine, Karolinska Institute, Stockholm, Sweden

Pancreatic cancer (PC) is one of the leading causes of cancer mortality worldwide, and its incidence and mortality rate in several regions is higher in male patients. Although numerous efforts have been made to enhance the clinical outcomes of existing therapeutic regimens, their efficiency is still low, and drug resistance usually occurs in many patients. In addition, the exact underlying molecular basis that makes PC slightly more prevalent among males remains unknown. Providing information regarding the possible association between gender and PC tumorigenesis may offer important clues for how certain molecular cross-talks can affect PC initiation and/or progression. In this study, we used several microarray expression data to identify the common up- and downregulated genes within one specific gender, which were also specified to have binding sites for androgen and/or estrogen receptors. Using functional enrichment analysis among the others, for all the gene sets found in this study, we have shed light on the plausible importance of the androgenic effectors in tumorigenesis, such as the androgen-regulated expression of the GLI transcription factor and the potential role of testosterone in the extracellular matrix (ECM)–cell interaction, which are known for their importance in tumorigenesis. Moreover, we demonstrated that the biological process *axon guidance* was highlighted regarding the upregulated genes in male patients. Overall, identification of gene candidates as the possible link between gender and PC progression or survival rates may help in developing strategies to reduce the incidence of this cancer.

KEYWORDS

pancreatic cancer, gender propensity, extracellular matrix, androgen receptor, estrogen receptor

1 Introduction

Pancreatic cancer (PC) is the seventh leading cause of cancer deaths worldwide, and it is expected that its rank would even go higher (Rawla et al., 2019; Hu et al., 2021). Despite significant improvements in the overall outcome of cancer patients, the incidence and mortality rates of PC have not showed a substantial change over the past 30 years, and the 5-year survival rate is only about nine percent (Hu et al., 2021). This is due to the late diagnosis and lack of knowledge regarding the major molecular mechanisms of PC tumorigenesis, which have resulted in poor effectiveness of therapeutic approaches and drug resistance (Zhu et al., 2020; Javadrashid et al., 2021). However, one important consideration in PC research may be the influence of gender on the incidence and survival rates of this cancer (Pourshams et al., 2019; Sung et al., 2021). The relatively higher incidence rates of PC in males have been shown in certain regions, for example, in 2020 in Eastern and Western Europe (with an age-standardized incidence rate per 100,000 of 9.9 for males and 5.6 for females and 9.8 for males and 7.4 for females, respectively), Northern America, Southern Europe, and Northern Europe, among others (Sung et al., 2021). One possible reason, aside from females' less exposure to PC risk factors (Mario et al., 2018) including the growing rates of obesity, diabetes, and alcohol consumption, might be the sex effect on PC development and progression both in humans and rodent animal models (Longnecker and Sumi, 1990; Wang et al., 2021). In fact, glucocorticoids and hormones have shown to affect the physiology of pancreas. Accordingly, in the PC rat models which were treated with azaserine, estrogens could play a protective role, while androgens possibly facilitate the development of PC (Wang et al., 2021). In line with that, evidence also points to the existence of an association between PC and sex hormone signaling in humans. In this regards, several anti-androgen drugs were tested in PC models, and their potential antitumor effect was demonstrated (Schweizer and Yu, 2017; Wang et al., 2021). However, the results of the related clinical trials have shown conflicting results (Schweizer and Yu, 2017).

Histologically, PC is divided into two main types: exocrine tumors (−93%), that is, cystic tumors, acinar cells cancers, and adenocarcinomas, such as pancreatic ductal adenocarcinoma (PDAC), which is the most common type of PC, and endocrine tumors (−7%) (Mostafa et al., 2017). Although various studies have focused on elucidating the underlying molecular mechanisms of the disease, a lot remained to be discovered to help with lowering the mortality rate of this cancer (Rawla et al., 2019; Hu et al., 2021; Javadrashid et al., 2021). PC, as a multifactorial disease, is associated with both genetic and environmental risk factors. Genetic alternations, for example, in *KRAS*, *TP53*, *CDKN2A*, and *SMAD4,2* (as the main drivers of PC), appearance of numerous differentially up- and downregulated genes, and downstream affected cellular signaling pathways have long been shown to be involved in PC progression (Jones et al., 2008; Javadrashid et al., 2021). Transcription factors (TFs) are among the final effectors of these pathways as well as certain hormones (such as androgen and estrogen) and play important roles in cellular processes (Bhagwat and Vakoc, 2015; Naqvi et al., 2018; Islam et al., 2021). Examples of such TFs are androgen receptors (ARs) and estrogen receptors (ERs). Activation of GLI transcription factors, which are involved in PDA oncogenesis in pancreatic epithelium,for example, can be mediated through AR binding (Adams et al., 2019; Li et al., 2018a). In accordance, one novel and interesting function of GLI2 is being the master regulator of the basal-like subset of PDA, which is characterized by the expression of laminin and epithelial-to-mesenchymal transition (EMT) markers, among others (Adams et al., 2019). However, it should be also noted that, for example, in breast cancer cells, non-canonical induction of *GLI3* transcription is driven by an ER interaction (Massah et al., 2021). In addition to the cascade of deregulated pathways with intracellular origins, tumor microenvironment (TME) has also a significant impact on PC progression. This, along with the aberrant cell-ECM homeostasis, is among the leaders of cancer progression, including metastasis and drug resistance (Henke et al., 2020). Moreover, the collagen density has been shown to have a possible role critical for cancer cells to evade immune responses (Kuczek et al., 2019). Fibronectin, another important component of ECM, has also a prominent role in PDAC cell malignancy and fibrogenesis (Javadrashid et al., 2021). One notable consideration in this regards is the possible interaction of physiological testosterone in enhancement of ECM synthesis within certain cells (Bertolo et al., 2014). However, the exact molecular mechanism in PC male patients still needs to be elucidated.

Through specifying differentially expressed gene (DEG) sets in affected versus normal cells and gene expression profiling, bioinformatics analysis has thus far demonstrated its potential to further enlighten the molecular mechanisms involved in PC (Jiang et al., 2015; Wang et al., 2017; Li et al., 2018b). DEGs in certain conditions, such as cancer, may differ among males and females due to their chromosomal locations or their response to the level of steroid hormones, which affect both immunological (innate and adaptive) and non-immunological responses to tumors (Libert et al., 2010; Zhang et al., 2012; Fuentes and Silveyra, 2019; Blencowe et al., 2022). This would lead to the high incidence of autoimmune diseases in females, whereas males show a higher risk of death from cancers (Klein

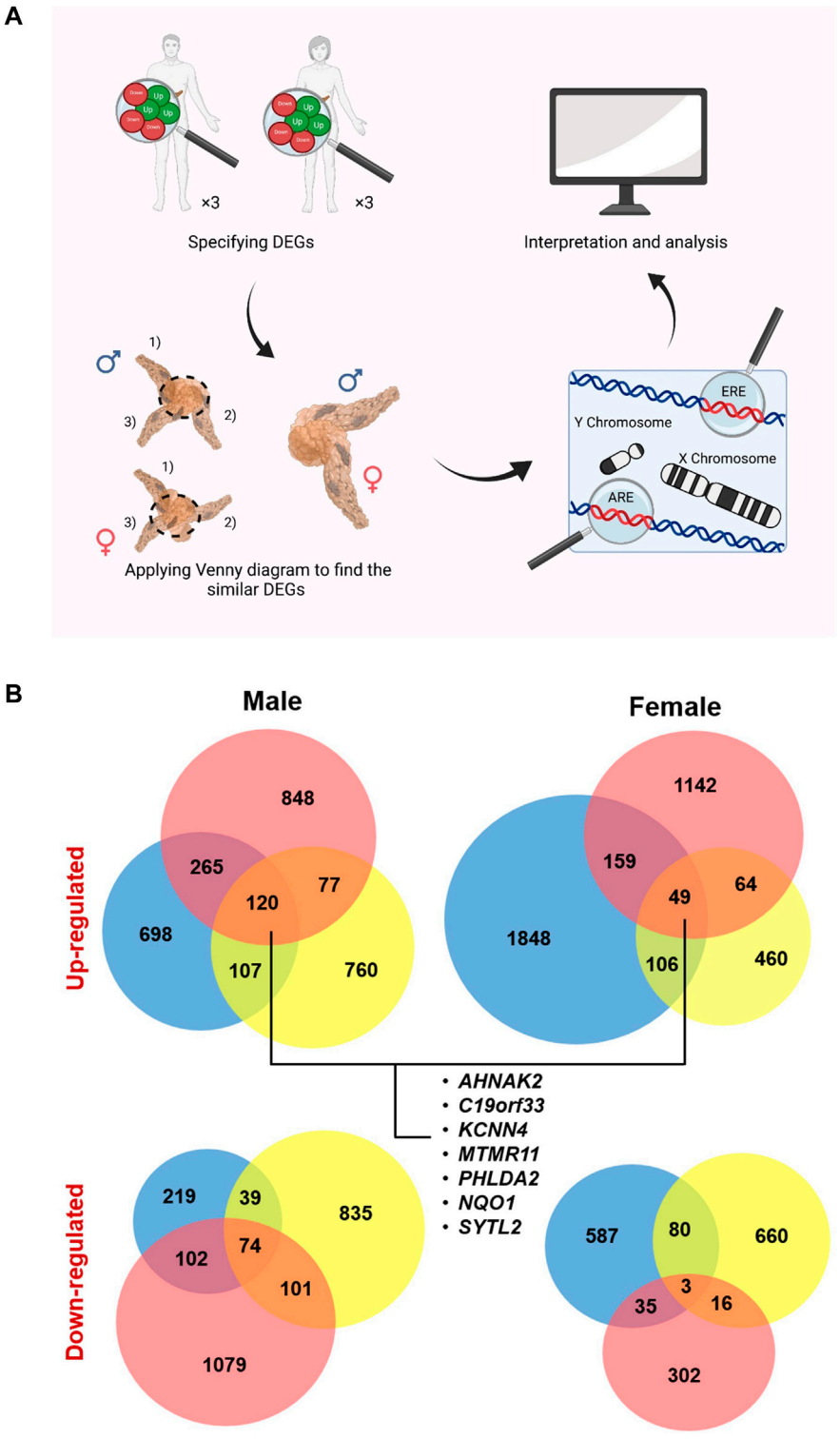


FIGURE 1
Study design and specified up- and down-regulated genes (U/DRGs). **(A)** Micro array studies for investigating the altered gene expression in both male and female pancreatic cancer (PC) patients were selected and the similar DEGs in both genders or the sex specified DEGs were identified. Subsequently, genes with AREs or EREs, through performing functional enrichment analysis, and the ones located on sex chromosomes were specified. **(B)** The number of identified U/DRGs in both genders. Log FC value ≥1 or ≤ −1 was chosen as the cut-off criterion for the up- and down-regulated genes, respectively, using a p-value of ≤ 0.05.

TABLE 1 Specifications of selected studies based on microarray data.

Disease type	Source of sample	Series ID	Platform ID	Number of cases		Age range		Number of controls		Platform name	Year
				F	M	F	M	F	M		
PC	Pancreatic tumor	GSE16515	GPL570	14	22	57–79	49–84	4	12	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	2009
PDAC	Pancreatic tumor	GSE55643	GPL6480	17	28	UD	UD	4	4	Agilent-014850 Whole Human Genome Microarray 4 × 44K G4112F	2014
PC	Pancreatic Tumor	GSE22780	GPL570	5	3	38–73	59–69	5	3	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	2011

PDAC, pancreatic ductal adenocarcinoma; UD, undetermined.

and Flanagan, 2016). Therefore, in this study, we aimed to analyze the likely impact of gender on DEGs in PC and to propose possible candidates in this regards. However, one main drawback of this work was the limited access to the clinical characteristics of patients, which makes the interpretation of clinical relevance in regards with the identified genes difficult. Nevertheless, the suggested perspective here is applicable in various cancer studies and can offer additional benefits. Finding the involved gene(s) and mechanism(s) through which gender would influence PC progression or survival rates may help in finding solutions to improve the effectiveness of therapies and ameliorating the disease progression in PC patients.

2 Methods

In this study, we aimed to specify DEGs in PC, compared to the normal tissue, regarding their plausible roles in the underlying pathways during cancer pathogenesis (Figure 1A). Considering the possible bias in PC incidence and survival rates between males and females, we have hypothesized that sex may affect these genes. Using microarray data and functional enrichment analysis, several genes and biological pathways were found to be possibly related to gender and basic mechanisms in PC.

2.1 Microarray datasets and analysis

The gene expression datasets from different microarray platforms related to PC were collected from the Gene Expression Omnibus (GEO) database (Table 1) (Pei et al., 2009; Barrett et al., 2013; Lunardi et al., 2014). Subsequently, GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/) was used to obtain DEGs in female and male patients relative to their matching controls, by using the linear models for microarray analysis (limma). As limma needs the data values to be in the log space, GEO2R automatically performs a log2 transformation on the selected sample values, which are specified not to be in log space (Smyth, 2004). The overall age range for females and males was 38–79 and 49–84, respectively. The absolute amount of LogFC value ≥ 1 or ≤ -1 was chosen as the cutoff criterion for the up- and downregulated genes, respectively, using a p-value of ≤ 0.05. The DEGs were then separately extracted in each gender relative to their corresponding controls.

2.2 Functional enrichment analysis

Applying Venn diagram version 2.1 (https://bioinfogp.cnb.csic.es/tools/venny/), similar up- and downregulated genes (U/DRGs) for each gender between all studies and

TABLE 2 List of common genes between all studies for each gender and both (genes which may be affected by gender are specified in bold, and GLI2 targets have been underlined).

Exp	Common	Just in male				Just in female	
Upregulated	*AHNAK2*	*ABCG1*	***EFNB2*** ARE	*LOC100127972*	*RRAS*	*ABHD17C*	*MUC5AC*
	C19orf33	*ABLIM3*	*EHD1*	*LOX*	*RUNX1-IT1*	*ADGRF1*	*NMU*
	KCNN4	*ANO1*	***ELF4*** X	*LPCAT4*	*S100A6*	*ARL14*	*PLEK2*
	MTMR11	*ANTXR1*	***EPHA4*** ARE	*LRRC8A*	*S100P*	*ASPHD2*	*PSCA*
	NQ O 1	*ANTXR2*	***FBXO32*** ARE	*LRRFIP1*	*SAMD9*	*CEACAM5*	*PTK6*
	PHLDA2	***ANXA1*** ARE	***FN1*** ARE	*LY6E*	*SCD*	*CTSE*	*RHBDL2*
	SYTL2	*AREG*	*FSCN1*	***MALL*** ARE	*SERPINH1*	*DPCR1*	***SALL4*** ERE
		ARHGAP26	*FXYD5*	*MDK*	***SH3KBP1*** X	***ELOVL6*** ERE	*SDR16C5*
		ARRDC2	*FZD2*	*MEGF6*	***SIM2*** ARE	*EPS8L1*	*SERPINB5*
		BICD1 ARE	*GBP1*	*MGLL*	***SLC24A3*** ARE	*EZR*	*SGPP2*
		BID	*GJB2*	*MICALL2*	*SLC O 3A1*	*FA2H*	*SLC4A11*
		BST2	*GJB3*	*MMP11*	*SLFN13*	*GALNT5*	***SLC6A14*** X
		CALD1 ARE	***GLI2*** ARE	*MX1*	*SPARC*	*ID1*	*SLC O 4A1*
		CD109	*GPRC5A*	*MYOF*	***SRPX2*** X	*IL1RN*	*STYK1*
		CD55	*HK2*	***NDC80*** ARE	*STX1A*	*KCNK1*	*SULT1C2*
		CDH3	*IFI44L*	***NHS*** X	***SULF1*** ARE	*LAMB3*	***TFF1*** ERE
		COL1A1	*IGF2BP3*	*NOX4*	***SULF2*** ARE	*LAMC2*	*TMPRSS3*
		COL1A2 X	***IGFBP3*** ARE	*NT5E*	***TACC3*** ARE	*MCU*	*TRIM29*
		COL3A1 ARE	*IGFBP5*	*NUAK1*	*THBS2*	*MIA*	*TRIM31*
		COL5A1 ARE	*IL1R2*	***OSBPL10*** ARE	*TNFAIP2*	*MMP28*	*TSPAN1*
		COL5A2	*INHBA*	***PALLD*** ARE	*TNFAIP6*	*MST1R*	*VSIG2*
		COR O 2A	***ISG15*** ARE	***PDLIM7*** ARE	*TPM2*		
		CRIP1	*ISG20*	*PFKP*	*TPM4*		
		CTHRC1	*ITPR3*	*PLAU*	*UPP1*		
		DDX60L	*KIF26B*	*PMEPA1*	*VCAN*		
		DKK3	*KLK10*	*PPARG*	*WNT2*		
		DNAJA4	*KYNU*	*RAB31*	*XAF1*		
		DPYSL3	***LAMA3*** ARE	*RAPH1*			
		EDNRA ARE	*LCK*	*RNF213*			
Downregulated	—	***ABAT*** ARE	*ER O 1B*	*NR5A2*	***SLC16A10*** ARE/ERE	*BCL11A FGFR1 GNG7*	
		ANPEP ARE	*F11*	*NRCAM*	***SLC1A2*** ARE		
		AQP12B	***FGL1*** ERE	***NUCB2*** ARE	***SLC25A15*** ERE		
		BHLHA15 ARE	*FKBP11*	***PAIP2B*** ARE	*SLC25A45*		
		BNIP3	***GAS2*** ARE	*PDCD4*	*SLC39A5*		
		BTG2	***GATM*** ARE	*PDIA2*	***SLC39A8*** ERE		
		C5	***GCAT*** ERE	***PELI2***ARE *PM20D1*	***SLC4A4*** ARE		
		CBS ERE	*GLS2*	***PNLIPRP1*** ARE	*SLC7A2*		
		CHRM3 ARE	***GMNN*** ARE	*PPP2R2D*	***STXBP6*** ARE		
		COCH	***GNMT*** ERE	***PRDX4*** X	*SYBU*		
		CTH	*GSTA3*	***PRLR*** ARE	*TCEA3*		
		CTTNBP2 ARE	*GUCA1C*	*PROX1*	*TDH*		
		CYP4V2	*JADE1*	*RAB26*	***TEX11*** X		
		DMD X	*KLK1*	*SEC11C*	*TMED6*		
		DNASE1 ARE	***LPAR3*** ARE	***SEL1L*** ARE	***TMEM56*** ARE		
		DPP10	*MPP7*	***SEMA6D*** ARE	*TPCN2*		
		ENPP1	*MPV17L*	*SERPINI2*	*TPST2*		
		ENTPD3	*MUC15*		*TRIM50*		
		EPB41L4B ARE	***NEDD4L*** ARE				

ARE, androgen response element; ERE, estrogen response element; X, X chromosome.

the ones between both sexes were specified (Figure 1B) (Oliveros, 2007). The Database for Annotation, Visualization, and Integrated Discovery (DAVID) (Huang et al., 2009a; Huang et al., 2009b) and the Enrichr database (https://maayanlab.cloud/Enrichr/) (Chen et al., 2013; Kuleshov et al., 2016; Clarke et al., 2021) were used to determine KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways (KEGG_2016_Human) for similar U/DRGs in both sexes (p-value ≤0.05). The KEGG database was used to interpret the data in terms of general schemes (Kanehisa et al., 2019).

To identify the genes with AREs or EREs, the list of restricted expressed genes to each gender specifically was submitted in the Enrichr database. The ChEA 2016 library was used to determine genes with the ability to bind to AR or ER, that is, possess AREs or EREs (p-value ≤ 0.05). In addition, these genes were investigated in terms of their location on any of the sex chromosomes, using the biological database network (bioDBnet) (https://biodbnet-abcc.ncifcrf.gov/db/db2db.php), to further assess the possibility of their affection in a gender-specific manner (Mudunuri et al., 2009). The Pancreatic Expression Database (PED) (https://www.pancreasexpression.org/) was used to investigate the possible correlation of expression profiles between upregulated ARE-containing genes in male patients with several top mutated genes in PC, in male and female patients from The Cancer Genome Atlas (TCGA) data source (Marzec et al., 2018). Moreover, PED was used to generate plots for assessing the relationship between these gene expressions and survival according to the TCGA data source (Marzec et al., 2018).

Finally, the TRANSFAC predicted transcription factor targets dataset (https://maayanlab.cloud/Harmonizome/dataset/TRANSFAC + Predicted + Transcription + Factor + Targets) was used to find potential transcription factors (TFs) among putative sex-affected gene candidates (Matys et al., 2003; Matys et al., 2006). The enriched Gene Ontology (GO) terms (biological process (BP), cellular component (CC), and molecular function (MF)) for ARE/ERE-containing genes were identifiedusing the Enrichr database (p-value ≤ 0.05).

2.3 Protein–protein interactions in sex-related U/DRGs

To establish protein–protein interactions (PPIs) and to subsequently identify the possible hubs in U/DRGs, for female and male patients, individually, the Search Tool for the Retrieval of Interacting Genes (STRING) database (https://string-db.org/) was used to analyze the data (score ≥0.4) (Szklarczyk et al., 2021). In addition, the HIPPIE dataset (http://cbdm-01.zdv.uni-mainz.de/~mschaefer/hippie/) (Alanis-Lobato et al., 2016) and the BisoGenet tool (Martin et al., 2010) were also used to draw interactions between proteins. The identified PPIs from three data sources were then merged, and topological parameters, such as degree, and centrality measures, including betweenness centrality (BC), for undirected and directed networks, were computed, using a Cytoscape network analyzer software version 3.8.0. Proteins with at least a 10-degree score were specified as hub proteins. Also, the cutoff value ≥ 0.05 was used for BCs to specify bottleneck proteins, that is, network proteins that have many shortest paths going through them (Yu et al., 2007).

3 Results

3.1 DEG analysis in association with sex

Upon extraction of DEGs in three independent studies, considering both female and male patients' data, (Supplementary Table S1–3), it was shown that 120 up- and 74 downregulated genes were common in male patients. The number of common U/DRGs between females was 49 and 3, respectively. A comparison between the two genders demonstrated that only *AHNAK2*, *C19orf33*, *KCNN4*, *MTMR11*, *NQO1*, *PHLDA2*, and *SYTL2* were common in the URGs, and no common genes were observed in the DRGs' list (Table 2) (Figure 1B).

3.2 Discovering the gender-related DEGs

Following the enrichment analysis on male-related URGs, it was specified that *ANXA1*, *BICD1*, *CALD1*, *COL3A1*, *COL5A1*, *EDNRA*, *EFNB2*, *EPHA4*, *FBXO32*, *FN1*, *GLI2*, *IGFBP3*, *ISG15*, *LAMA3*, *MALL*, *NDC80*, *OSBPL10*, *PALLD*, *PDLIM7*, *SIM2*, *SLC24A3*, *SULF1*, *SULF2*, and *TACC3* have ARE, whereas no genes with ERE were found in this class. Moreover, *COL1A2*, *ELF4*, *SH3KBP1*, *SRPX2*, and *NHS* were shown to be on the X chromosome. Regarding female-related URGs, *ELOVL6*, *SALL4*, and *TFF1* have ERE, whereas *SLC6A14* is located on the X chromosome. Among the male-related DRGs, *ABAT*, *ANPEP*, *BHLHA15*, *CHRM3*, *CTTNBP2*, *DNASE1*, *EPB41L4B*, *GAS2*, *GATM*, *GMNN*, *LPAR3*, *NEDD4L*, *NUCB2*, *PAIP2B*, *PELI2*, *PNLIPRP1*, *PRLR*, *SEL1L*, *SEMA6D*, *SLC1A2*, *SLC4A4*, *STXBP6*, and *TMEM56* were found to have ARE, whereas *CBS*, *FGL1*, *GCAT*, *GNMT*, *SLC25A15*, and *SLC39A8* contained ERE. *SLC16A10* has both ARE and ERE. Furthermore, *DMD*, *PRDX4*, and *Tex11* are located on the X chromosome. There were no female-related DRGs having ARE or ERE. In addition, no genes were specified to be located on the X chromosome (Table 2).

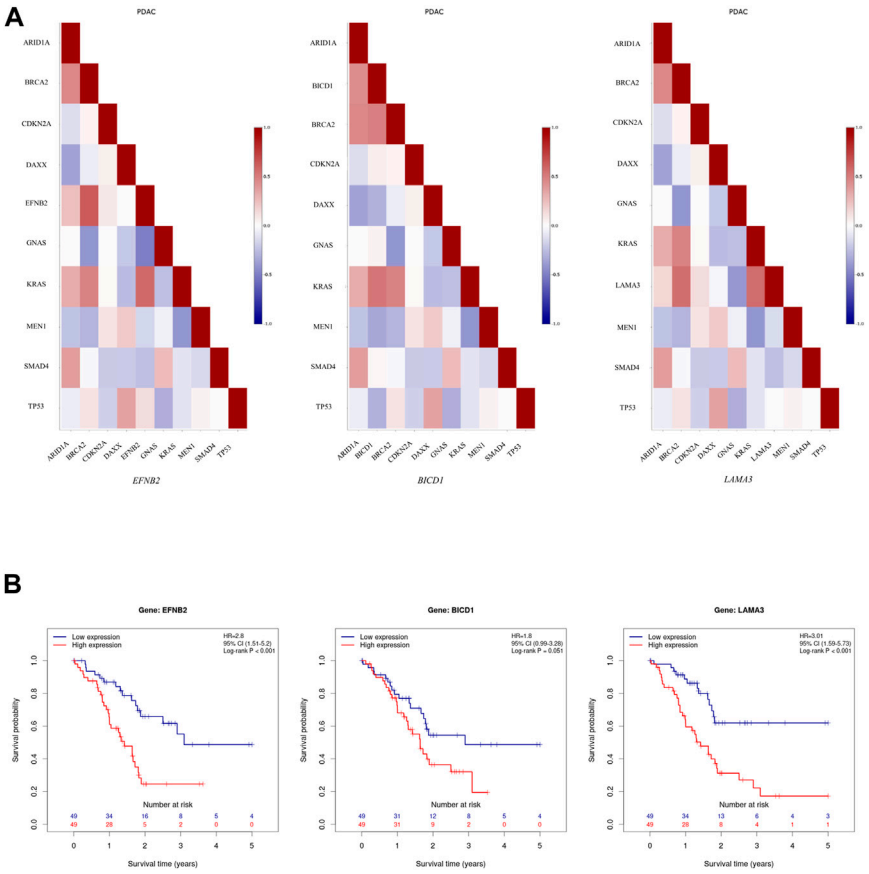


FIGURE 2
Further assessment of upregulated ARE-containing genes in male patients among the male PC patients from TCGA data source. **(A)** Evaluating the correlation of upregulated ARE-containing genes with top mutated genes in PC showed that *EFNB2*, *BICD1*, and *LAMA3*, have a high positive correlation with *KRAS* expression among male PC patients from TCGA data source. The color of each cell indicates correlation coefficient between corresponding genes labelled on the *x*-axis and *y*-axis. **(B)** The survival analysis on*EFNB2*, *BICD1*, and *LAMA3* in male PC patients from TCGA data source showed a decrease in survival probability upon their high expression. Data is obtained from Pancreatic Expression Database (PED) (https://www.pancreasexpression.org/).

It seems that ARE-containing genes that were upregulated in male patients have a higher correlation in their expression profiles with each other, in comparison with female PC patients from the TCGA data source, based on the Pearson Product Moment Correlation Coefficient (PMCC) value (Supplementary Figure S1). Also, assessing the correlation between these gene expressions and several top mutated genes in PC confirmed their high correlation with *KRAS*, among others (Supplementary Figure S2A). Of note, *EFNB2*, *BICD1*, and *LAMA3* showed a high positive correlation with *KRAS* (Figure 2A). Meanwhile, investigating downregulated genes in male patients showed a negative correlation with *KRAS* (Supplementary Figure S2B). Furthermore, the survival analysis showed association between lower survival probability and upregulation of several of these genes in male PC patients according to the TCGA data source (Supplementary Figure S3). Interestingly, the decreased survival probability was also clearly showed for *EFNB2*, *BICD1*, and *LAMA3* (Figure 2B).

3.2.1 Sex-possibly affected genes can act as transcription factors

By annotating a gene function using Entrez Gene Summary and PIR Summary in DAVID for genes with sex hormone regulating elements, it was specified that *GLI2*^{ARE}, *SIM2*^{ARE}, and *ELF4*^{X} in male-related URGs, *SALL4*^{ERE} in female-related URGs, and *BHLHA15*^{ARE} in the male-related downregulated class have the transcription regulatory function. Some of GLI2 targets exist within the list of male URGs, including *ABCG1*, *ARRDC2*, *C19orf33*, *COL5A1*^{ARE}, *FZD2*, *GJB2*, *ITPR3*, *LY6E*, *MMP11*, *NDC80*^{ARE}, *NHS*^{X}, *PALLD*^{ARE}, and *WNT2*. In addition, based on the TCGA data source, it seems that *GLI2* expression is slightly higher in male PC patients in comparison with the female patients (Supplementary Figure S4).

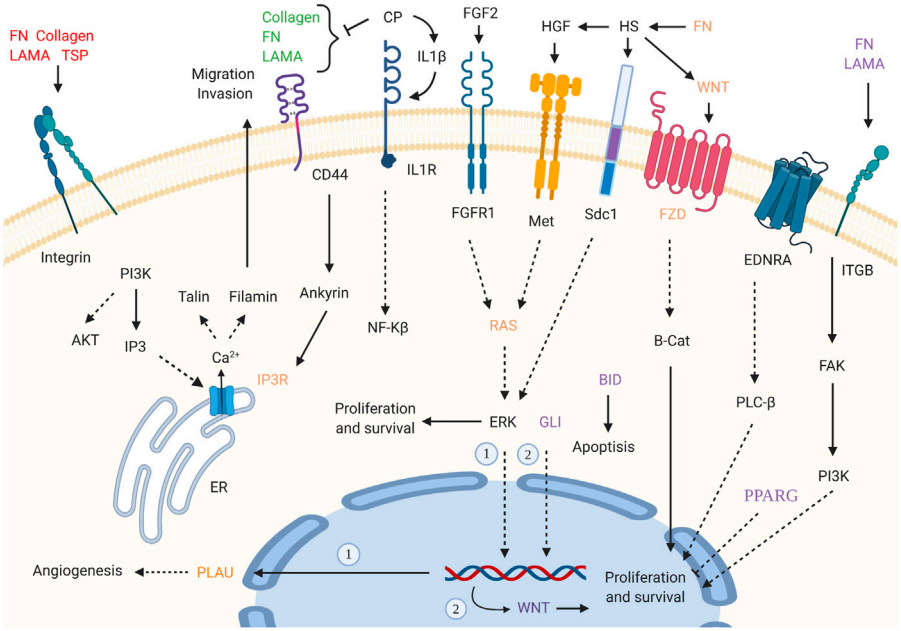


FIGURE 3
Enriched signaling pathways for URGs in male pancreatic cancer (PC) patients. Interaction of extracellular matrix (ECM) components and proteoglycans with intracellular signaling pathways, such as WNT signaling, and cell proliferation, ERK signaling, and angiogenesis has been implicated. GLI2 is also highlighted in pathways in cancer_hsa05200. (p-value ≤0.05). "Created with BioRender.com."

3.2.2 Enrichment analysis of gender-related DEGs in male and female patients

Supposing that gender-related U/DRGs in PC patients may reveal new insights toward the possible association of sex and PC progression or, cues regarding its development, the respected signaling pathways for each group have been investigated, based on the retrieved data from DAVID and KEGG databases (Figures 3, 4). Furthermore, according to specified GO terms for ARE-containing up- and downregulated genes in male patients and ERE-containing upregulated genes in female patients (Figure 5), it was shown that ARE-containing URGs in males are significantly involved in axon guidance (GO:0007411) (p-value ≤ 0.01) (Supplementary Table S4).

Pathways with a possible role in disease development were deduced for male-related U/DRGs and female-related DRGs (p-value ≤ 0.05) (Supplementary Table S5). No pathways with significant relation to female-related URGs were observed. Among the URGs in male patients, extracellular matrix (ECM) components (COL3A1, COL1A2, COL1A1, COL5A2, COL5A1, LAMA3, FN1, and THBS2) and their deduced interaction with intracellular signaling pathways have been highlighted. The other enriched pathway, proteoglycans in cancer (proteoglycans in cancer_hsa05205), has also been specified with its implication in WNT signaling, cell proliferation, ERK signaling, and angiogenesis. Another indicated protein, GLI2, is also of note in pathways in cancer_hsa05200. It can mediate cell proliferation through *WNT* expression (Figure 3). The enriched pathways related to DRGs in male patients include glycine, serine, and threonine metabolism_hsa00260 and pancreatic secretion_hsa04972 (Figure 4A). Regarding the latter, it can be seen that through downregulation of CHRM3 and TPCN2 proteins, the secretion of pancreatic proteases and lipases would be affected, whereas the reduction in SLC4A4 protein can modify carbohydrate digestion and absorption. Finally, it is of note that downregulated proteins, FGFR1 and GNG7, in female patients are significantly related to the Ras signaling pathway_hsa04014 and the PI3K-Akt signaling pathway_hsa04151 (Figure 4B).

3.3 Hub and bottleneck proteins in U/DRGs

Considering the defined cutoffs, hub and bottleneck proteins were only detected for URGs in male patients (Supplementary Table S6). In addition, several proteins have been specified as both hubs and bottlenecks. These include FN1, CALD1, ISG15, SH3KBP1, COL1A1, PPARG, LOX, LCK, and SPARC.

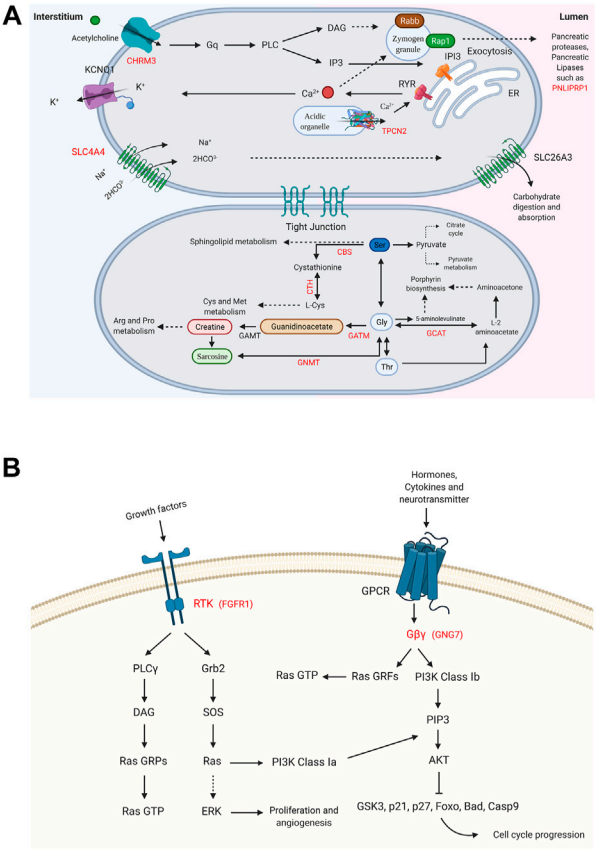


FIGURE 4
Enriched signaling pathways for the DRGs in male and female pancreatic cancer (PC) patients. **(A)** Glycine, serine, and threonine metabolism_hsa00260 and pancreatic secretion_hsa04972 pathways are shown for DRGs in male PC patients. **(B)** Ras signaling pathway_hsa04014 and PI3K-Akt signaling pathway_hsa04151 are specified for FGFR1 and GNG7, the DRGs in female PC patients. (p-value ≤0.05). "Created with BioRender.com."

4 Discussion

The prevalence of pancreatic adenocarcinoma is relativelyhigher in males compared to females in regions including theUnited Kingdom and the United States, where the samples used in this study were obtained from (McGuigan et al. 2018; Pourshams et al. 2019). In this regard, investigating the pathways that certain up- or downregulated proteins re involved in, for example, in male PC patients, may provide valuable information regarding the relation between sex and cancerspecific signaling pathways.

In our study, ECM–cell interactions and other closely linked pathways, such as focal adhesion, proteoglycans in cancer, and their subsequent signaling cascades, have been highlighted, respecting the URGs in male PC patients. However, it should be noted that, generally, one of the main hallmarks of the immunosuppressive microenvironment of PC in both sexes is the excessive deposition of ECM components (Orth et al., 2019). Among several components of ECM, collagens are the most abundant (Weniger et al., 2018). In our study five out of eight upregulated proteins involved in ECM–receptor interactions are collagens located in the interstitial space (COL1A2, COL1A1, COL3A1, COL5A2, and COL5A1) (Supplementary Figure S4). Interstitial collagens and mainly collagen type 1 have tumor-promoting properties, which exert proliferation and migration, among others (Weniger et al., 2018; Xu et al., 2019). To the best of our knowledge, no study has been conducted to investigate the interaction of AR and ECM in PC. Nevertheless, it has already been shown that the androgen and estrogen ratio can mediate the expression of COL3A1 in the rat brain, while its expression in females decreased (Yonehara et al., 2003). In addition, the diameter of collagen fibrils can be regulated through sex hormone levels in favor of males (Markova et al., 2004). However, DRGs in male patients were enriched in amino acid metabolism as well as protease and lipase secretion–related pathways. Specifically, downregulation of cystathionine-beta-synthase and cystathionine gamma-lyase in glycine, serine and, threonine metabolism pathways would most probably block L-cysteine production from serine and, therefore, affect cysteine and methionine metabolism negatively. Also, one of the interesting inferences is the possible reduction of creatine following the downregulation of glycine amidinotransferase (GATM), also known as L-arginine:glycine amidinotransferase, which is the rate-limiting step of creatine synthesis from arginine and glycine (Figure 4A). In general, we propose that downregulation of specific components in the mentioned pathways may grant certain advantages for PC progression in males, as their reduced levels have been confirmed in three independent studies and specifically among males.

Next, the association between the presence of ERE/ARE in the regulatory regions of U/DRGs and PC underlying mechanism was investigated in a gender-specific manner. Interestingly, we noticed that the biological process *axon guidance* has been significantly highlighted among the ARE-containing URGs, that is, *EFNB2*, *EPHA4*, *PALLD*, *LAMA3*, *PDLIM7*, and *GLI2*, among the male patients. Frequent and diverse somatic aberrations in genes involved in regulating axon guidance have already been demonstrated to be important in pancreatic carcinogenesis (Biankin et al., 2012). This, along with the observation that the number and size of intrapancreatic nerves increase in PDAC, a process called pancreatic cancer–associated neural remodeling (PANR) (Wakiya et al., 2021), underlines the possible importance of sex's role in PC progression and how male hormones may affect the expression of involved genes. On the other hand, the presence of several TFs among the AER/ERE-containing genes may emphasize a more possible effect of sex on disease progression. For example, *GLI2*, among resulted ARE-containing URGs in males (AR_21915096_ChIP-Seq_LNCaP-1F5_Human), is a TF in

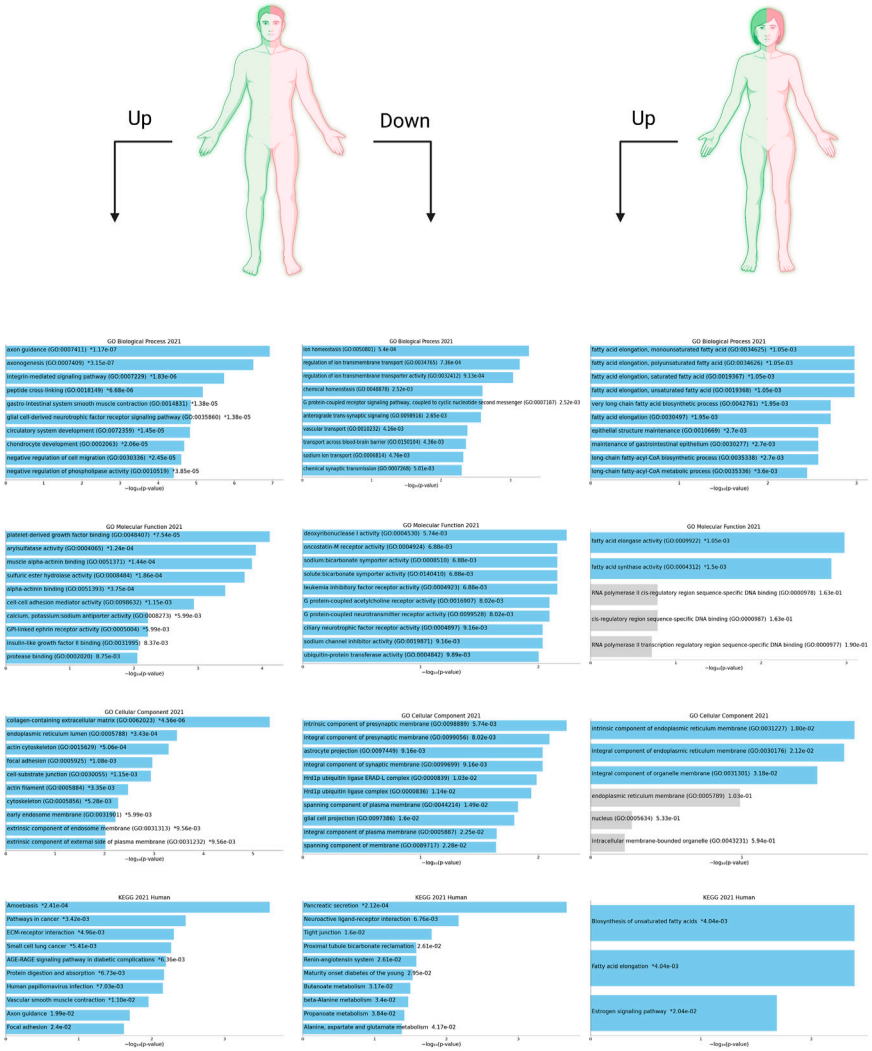


FIGURE 5
Top 10 enriched Gene Ontology (GO) terms (i.e., biological process (BP), molecular function (MF), and cellular component (CC)) and KEGG pathways have been shown for ARE-containing up- and downregulated genes (U/DRGs) in male PC patients. Also, the top enriched GO terms and KEGG pathways are listed for ERE-containing URGs in female PC patients. The corresponding *p*-values are shown. Blue bars correspond to terms with significant *p*-values (<0.05). An asterisk (*) next to a *p*-value indicates the term also has a significant adjusted *p*-value (<0.05). All data shown here are retrieved from the Enrichr database (https://maayanlab.cloud/Enrichr/) and visualized using Enrichr Appyter.

the Hedgehog pathway and has long been shown to be associated with prostate cancer tumorigenesis, among others (Xia et al., 2020). Thus far, several studies have investigated GLI's interaction with AR and demonstrated the resulted induction in the expression of GLI-dependent genes due to androgens (Li et al., 2018a). Although it has been shown that AR might also modulate transcription from *GLI2* (Li et al., 2014a), the effect of AR on the expression level of GLI proteins through ARE still needs to be further investigated. Also, several GLI2 targets, including *ABCG1*, *ARRDC2*, *C19orf33*, *COL5A1*, *FZD2*, *GJB2*, *ITPR3*, *LY6E*, *MMP11*, *NDC80*, *NHS*, *PALLD*, and *WNT2*, were shown to be upregulated in male patients. In general, targeting GLI2 seems a promising strategy that may result in the downregulation of a range of genes involved in various biological processes. The other resulted ARE-containing URG, *Single-minded 2 (SIM2)*, encodes a TF with a shown tumorgenesis activity in PC (DeYoung et al., 2003).

We also further investigated if the URGs on the X chromosome in male patients may have any possible association with a gender-related mechanism regarding cancer progression. While *SH3* domain-containing kinase binding protein 1 (SH3KBP1), Sushi Repeat Containing Protein X-Linked 2 (SRPX2), and NHS actin remodeling regulator (NHS) have been shown to have a variable escape from X

chromosome inactivation (XCI), one copy of *E74 Like ETS Transcription Factor 4* (*ELF4*) is indeed inactive in females, due to XCI (Carrel and Willard, 2005; Tukiainen et al., 2017) and interestingly, as a TF, it can regulate its targets in a dose-dependent manner (Du et al., 2021). High ELF4 expression is shown to be related to worse disease outcomes in several cancers, including PC, although observations exist in accordance with its both oncogenic and tumor suppressor roles in different cancer types (Suico et al., 2017; Kafita et al., 2021). Still, there is much to be identified regarding the involved mechanism through which ELF4 regulates its downstream target; however, it gained our attention that this TF has been identified as one allelic-specific expression site (Chuang et al., 2017). The ASE bias toward the paternal (P) or maternal (M) alleles was described as one P and 2 M among three studies, that is, the Geuvadis RNA Sequencing Project at http://www.geuvadis.org/web/geuvadis/rnaseq-project, Li et al., 2014b's study, and Cenik et al., 2015's study. This can be interesting as males always receive the maternal X chromosome.

Finally, one important consideration is the possible effect of sex on the identified hub proteins such as those related to ECM, that is, fibronectin 1 (FN1; with possible ARE), collagen type I Alpha 1 (COL1A1), lysyl oxidase (LOX), interferon-stimulated gene 15 (ISG15; with possible ARE), and caldesmon gene1 (CALD1; with possible ARE) (Table 2). Multiple lines of evidence have thus far demonstrated the effect of ECM dysregulation on tumor progression (Weniger et al., 2018). However, ECM seems to be affected by sex in different cell types, and the possible association between the androgen receptor pathway and collagen content has been demonstrated (Markova et al., 2004). In accordance, the physiological testosterone is also shown to be involved in enhancing ECM synthesis in certain cells (Bertolo et al., 2014). Also, *CALD1*, is an actin-linked regulatory protein with multiple functions in cell motility, such as migration, invasion, and proliferation (Mayanagi and Sobue, 2011). While this hub-bottleneck protein may have the potential to be among the candidate cancer drivers in PC, it has already been shown to be androgen sensitive (Erho et al., 2012; Dressler et al., 2021).

5 Conclusion

Given the relatively biased rates of incident and survival among male and female PC patients, in this study, we have tried to shed light on the possible relation between DEGs in PC and gender. The possible regulation of highly or lowly expressed genes in male or female pancreatic tumors through sex hormones can be of great value in finding the important molecular mechanisms which may make one gender more potent for developing this cancer. It can also help identifying potential gender-specific biomarkers for linking gender and the pancreatic carcinogenesis. In addition, considering the promising results of specific androgen receptor blockers such as flutamide in increasing the survival of PC patients in two clinical trials, the current study may have benefits in elucidating the potential mechanisms mediating this effect of androgen receptor blockers (Greenway, 2000). This may also show the possible advantage of certain therapies, such as hormone withdrawal therapy in PC patients, although the reports are conflicting in this regards. One interesting suggestion would be to check and screen these potential candidates and assess their expression profile and their downstream targets (in case of TFs) in male deriving tumor cells upon treating with androgen blockers while investigating the cancerous phenotype. Finally, it is of note to take a specific look at the TFs as the plausible targets of ERs and ARs and the regulators of a broad range of genes. Using high-throughput single cell RNA sequencing methods and identifying the master transcription factors with possible ARE/ERE through methods such as SCENIC would pave the way of finding other potential candidates in this regards. However, one main drawback of this study was the unavailability of the patients' clinical characteristics in accordance with gender, which, therefore, limits the relevancy of the identified genes and the clinical outcomes. While this should to be considered in future studies, discovering key DEGs with relation to gender may help to develop new modeling platforms, detection approaches, and targeted therapies.

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

RR, AG, and MM contributed to design of the study and initial analysis. RR, AG, SR, and MR wrote and revised the sections of the manuscript. RR and HA contributed in data analysis and interpretation. MV and ST contributed in critical revision of the article and final approval of the version to be published. All authors contributed to manuscript revision and read 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

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/fgene.2022.966941/full#supplementary-material

References

Adams, C. R., Htwe, H. H., Marsh, T., Wang, A. L., Montoya, M. L., Subbaraj, L., et al. (2019). Transcriptional control of subtype switching ensures adaptation and growth of pancreatic cancer. *eLife* 8, e45313. doi:10.7554/eLife.45313

Alanis-Lobato, G., Andrade-Navarro, M. A., and Schaefer, M. H. (2016). HIPPIE v2.0: Enhancing meaningfulness and reliability of protein–protein interaction networks. *Nucleic Acids Res.* 45 (1), D408–D414. doi:10.1093/nar/gkw985

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: Archive for functional genomics data sets--update. *Nucleic Acids Res.* 41, D991–D995. doi:10.1093/nar/gks1193

Bertolo, A., Baur, M., Aebli, N., Ferguson, S. J., and Stoyanov, J. (2014). Physiological testosterone levels enhance chondrogenic extracellular matrix synthesis by male intervertebral disc cells *in vitro*, but not by mesenchymal stem cells. *Spine J.* 14 (3), 455–468. doi:10.1016/j.spinee.2013.10.018

Bhagwat, A. S., and Vakoc, C. R. (2015). Targeting transcription factors in cancer. *Trends Cancer* 1 (1), 53–65. doi:10.1016/j.trecan.2015.07.001

Biankin, A. V., Waddell, N., Kassahn, K. S., Gingras, M-C., Muthuswamy, L. B., Johns, A. L., et al. (2012). Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature* 491 (7424), 399–405. doi:10.1038/nature11547

Blencowe, M., Chen, X., Zhao, Y., Itoh, Y., McQuillen, C. N., Han, Y., et al. (2022). Relative contributions of sex hormones, sex chromosomes, and gonads to sex differences in tissue gene regulation. *Genome Res.* 32 (5), 807–824. doi:10.1101/gr.275965.121

Carrel, L., and Willard, H. F. (2005). X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* 434 (7031), 400–404. doi:10.1038/nature03479

Cenik, C., Cenik, E. S., Byeon, G. W., Grubert, F., Candille, S. I., Spacek, D., et al. (2015). Integrative analysis of RNA, translation, and protein levels reveals distinct regulatory variation across humans. *Genome Res.* 25 (11), 1610–1621. doi:10.1101/gr.193342.115

Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G. V., et al. (2013). Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinforma.* 14 (1), 128. doi:10.1186/1471-2105-14-128

Chuang, T-J., Tseng, Y-H., Chen, C-Y., and Wang, Y-D. (2017). Assessment of imprinting- and genetic variation-dependent monoallelic expression using reciprocal allele descendants between human family trios. *Sci. Rep.* 7 (1), 7038. doi:10.1038/s41598-017-07514-z

Clarke, D. J. B., Jeon, M., Stein, D. J., Moiseyev, N., Kropiwnicki, E., Dai, C., et al. (2021). Appyters: Turning jupyter notebooks into data-driven web apps. *Patterns* 2 (3), 100213. doi:10.1016/j.patter.2021.100213

DeYoung, M. P., Tress, M., and Narayanan, R. (2003). Down's syndrome-associated Single Minded 2 gene as a pancreatic cancer drug therapy target. *Cancer Lett.* 200 (1), 25–31. doi:10.1016/s0304-3835(03)00409-9

Dressler, L., Bortolomeazzi, M., Keddar, M. R., Misetic, H., Sartini, G., Acha-Sagredo, A., et al. (2021). Comparative assessment of genes driving cancer and somatic evolution in non-cancer tissues: An update of the network of cancer genes (NCG) resource. *Genome Biol.* 23, 35. doi:10.1186/s13059-022-02607-z

Du, H., Xia, H., Liu, T., Li, Y., Liu, J., Xie, B., et al. (2021). Suppression of ELF4 in ulcerative colitis predisposes host to colorectal cancer. *iScience* 24 (3), 102169. doi:10.1016/j.isci.2021.102169

Erho, N., Buerki, C., Triche, T. J., Davicioni, E., and Vergara, I. A. (2012). Transcriptome-wide detection of differentially expressed coding and non-coding transcripts and their clinical significance in prostate cancer. *J. Oncol.* 2012, 541353. doi:10.1155/2012/541353

Fuentes, N., and Silveyra, P. (2019). Estrogen receptor signaling mechanisms. *Adv. Protein Chem. Struct. Biol.* 116, 135–170. doi:10.1016/bs.apcsb.2019.01.001

Greenway, B. A. (2000). Androgen receptor-blocking agents: Potential role in pancreatic cancer. *Drugs Aging* 17 (3), 161–163. doi:10.2165/00002512-200017030-00001

Henke, E., Nandigama, R., and Ergün, S. (2020). Extracellular matrix in the tumor microenvironment and its impact on cancer therapy. *Front. Mol. Biosci.* 6 (160), 160. doi:10.3389/fmolb.2019.00160

Hu, H-F., Ye, Z., Qin, Y., Xu, X-W., Yu, X-J., Zhuo, Q-F., et al. (2021). Mutations in key driver genes of pancreatic cancer: Molecularly targeted therapies and other clinical implications. *Acta Pharmacol. Sin.* 42 (11), 1725–1741. doi:10.1038/s41401-020-00584-2

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37 (1), 1–13. doi:10.1093/nar/gkn923

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4 (1), 44–57. doi:10.1038/nprot.2008.211

Islam, Z., Ali, A. M., Naik, A., Eldaw, M., Decock, J., and Kolatkar, P. R. (2021). Transcription factors: The fulcrum between cell development and carcinogenesis. *Front. Oncol.* 11, 681377. doi:10.3389/fonc.2021.681377

Javadrashid, D., Baghbanzadeh, A., Derakhshani, A., Leone, P., Silvestris, N., Racanelli, V., et al. (2021). Pancreatic cancer signaling pathways, genetic alterations, and tumor microenvironment: The barriers affecting the method of treatment. *Biomedicines* 9 (4), 373. doi:10.3390/biomedicines9040373

Jiang, J., Liu, H. L., Liu, Z. H., Tan, S. W., and Wu, B. (2015). Identification of cystatin SN as a novel biomarker for pancreatic cancer. *Tumour Biol.* 36 (5), 3903–3910. doi:10.1007/s13277-014-3033-3

Jones, S., Zhang, X., Parsons, D. W., Lin, J. C., Leary, R. J., Angenendt, P., et al. (2008). Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Sci. (New York, NY)* 321 (5897), 1801–1806. doi:10.1126/science.1164368

Kafita, D., Daka, V., Nkhoma, P., Zulu, M., Zulu, E., Tembo, R., et al. (2021). High ELF4 expression in human cancers is associated with worse disease outcomes and increased resistance to anticancer drugs. *PLOS ONE* 16 (4), e0248984. doi:10.1371/journal.pone.0248984

Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K., and Tanabe, M. (2019). New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* 47 (D1), D590–D5. doi:10.1093/nar/gky962

Klein, S. L., and Flanagan, K. L. (2016). Sex differences in immune responses. *Nat. Rev. Immunol.* 16 (10), 626–638. doi:10.1038/nri.2016.90

Kuczek, D. E., Larsen, A. M. H., Thorseth, M. L., Carretta, M., Kalvisa, A., Siersbæk, M. S., et al. (2019). Collagen density regulates the activity of tumor-infiltrating T cells. *J. Immunother. Cancer* 7 (1), 68. doi:10.1186/s40425-019-0556-6

Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., et al. (2016). Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 44 (W1), W90–W97. doi:10.1093/nar/gkw377

Li, N., Chen, M., Truong, S., Yan, C., and Buttyan, R. (2014). Determinants of Gli2 co-activation of wildtype and naturally truncated androgen receptors. *Prostate* 74 (14), 1400–1410. doi:10.1002/pros.22855

Li, N., Truong, S., Nouri, M., Moore, J., Al Nakouzi, N., Lubik, A. A., et al. (2018). Non-canonical activation of hedgehog in prostate cancer cells mediated by the interaction of transcriptionally active androgen receptor proteins with Gli3. *Oncogene* 37 (17), 2313–2325. doi:10.1038/s41388-017-0098-7

Li, X., Battle, A., Karczewski, K. J., Zappala, Z., Knowles, D. A., Smith, K. S., et al. (2014). Transcriptome sequencing of a large human family identifies the impact of rare noncoding variants. *Am. J. Hum. Genet.* 95 (3), 245–256. doi:10.1016/j.ajhg.2014.08.004

Li, Y., Li, L., Zhang, G., Wang, Y., Chen, H., Kong, R., et al. (2018). Crucial microRNAs and genes in metformin's anti-pancreatic cancer effect explored by microRNA-mRNA integrated analysis. *Invest. New Drugs* 36 (1), 20–27. doi:10.1007/s10637-017-0508-2

Libert, C., Dejager, L., and Pinheiro, I. (2010). The X chromosome in immune functions: When a chromosome makes the difference. *Nat. Rev. Immunol.* 10 (8), 594–604. doi:10.1038/nri2815

Longnecker, D. S., and Sumi, C. (1990). Effects of sex steroid hormones on pancreatic cancer in the rat. *Int. J. Pancreatol.* 7 (1-3), 159–165. doi:10.1007/BF02924233

Lunardi, S., Jamieson, N. B., Lim, S. Y., Griffiths, K. L., Carvalho-Gaspar, M., Al-Assar, O., et al. (2014). IP-10/CXCL10 induction in human pancreatic cancer stroma influences lymphocytes recruitment and correlates with poor survival. *Oncotarget* 5 (22), 11064–11080. doi:10.18632/oncotarget.2519

Mario, C., Marilisa, F., Kryssia, I. R-C., Pellegrino, C., Ginevra, C., Chiara, M., et al. (2018). Epidemiology and risk factors of pancreatic cancer. *Acta Biomed.* 89 (9), 141–146. doi:10.23750/abm.v89i9-S.7923

Markova, M. S., Zeskand, J., McEntee, B., Rothstein, J., Jimenez, S. A., and Siracusa, L. D. (2004). A role for the androgen receptor in collagen content of the skin. *J. Invest. Dermatol.* 123 (6), 1052–1056. doi:10.1111/j.0022-202X.2004.23494.x

Martin, A., Ochagavia, M. E., Rabasa, L. C., Miranda, J., Fernandez-de-Cossio, J., and Bringas, R. (2010). BisoGenet: A new tool for gene network building, visualization and analysis. *BMC Bioinforma.* 11 (1), 91. doi:10.1186/1471-2105-11-91

Marzec, J., Ullah Abu, DZ., Pirrò, S., Gadaleta, E., Crnogorac-Jurcevic, T., Lemoine, N. R., et al. (2018). The pancreatic expression database: 2018 update. *Nucleic Acids Res.* 46 (D1), D1107–D10. doi:10.1093/nar/gkx955

Massah, S., Foo, J., Li, N., Truong, S., Nouri, M., Xie, L., et al. (2021). Gli activation by the estrogen receptor in breast cancer cells: Regulation of cancer cell growth by Gli3. *Mol. Cell. Endocrinol.* 522, 111136. doi:10.1016/j.mce.2020.111136

Matys, V., Fricke, E., Geffers, R., Gössling, E., Haubrock, M., Hehl, R., et al. (2003). Transfac: Transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* 31 (1), 374–378. doi:10.1093/nar/gkg108

Matys, V., Kel-Margoulis, O. V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., et al. (2006). TRANSFAC and its module TRANSCompel: Transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* 34, D108–D110. doi:10.1093/nar/gkj143

Mayanagi, T., and Sobue, K. (2011). Diversification of caldesmon-linked actin cytoskeleton in cell motility. *Cell adh. Migr.* 5 (2), 150–159. doi:10.4161/cam.5.2.14398

McGuigan, A., Kelly, P., Turkington, R. C., Jones, C., Coleman, H. G., and McCain, R. S. (2018). Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes. *World J. Gastroenterol.* 24 (43), 4846–4861. doi:10.3748/wjg.v24.i43.4846

Mostafa, M. E., Erbarut-Seven, I., Pehlivanoglu, B., and Adsay, V. (2017). Pathologic classification of "pancreatic cancers": Current concepts and challenges. *Chin. Clin. Oncol.* 6 (6), 59. doi:10.21037/cco.2017.12.01

Mudunuri, U., Che, A., Yi, M., and Stephens, R. M. (2009). bioDBnet: the biological database network. *Bioinformatics* 25 (4), 555–556. doi:10.1093/bioinformatics/btn654

Naqvi, A. A. T., Hasan, G. M., and Hassan, M. I. (2018). Investigating the role of transcription factors of pancreas development in pancreatic cancer. *Pancreatology* 18 (2), 184–190. doi:10.1016/j.pan.2017.12.013

Oliveros, J. C. (2007). VENNY. An interactive tool for comparing lists with Venn diagrams. Available at: http://bioinfogpcnbcsices/tools/venny/indexhtml.

Orth, M., Metzger, P., Gerum, S., Mayerle, J., Schneider, G., Belka, C., et al. (2019). Pancreatic ductal adenocarcinoma: Biological hallmarks, current status, and future perspectives of combined modality treatment approaches. *Radiat. Oncol.* 14 (1), 141. doi:10.1186/s13014-019-1345-6

Pei, H., Li, L., Fridley, B. L., Jenkins, G. D., Kalari, K. R., Lingle, W., et al. (2009). FKBP51 affects cancer cell response to chemotherapy by negatively regulating Akt. *Cancer Cell* 16 (3), 259–266. doi:10.1016/j.ccr.2009.07.016

Pourshams, A., Sepanlou, S. G., Ikuta, K. S., Bisignano, C., Safiri, S., Roshandel, G., et al. (2019). The global, regional, and national burden of pancreatic cancer and its attributable risk factors in 195 countries and territories, 1990–2017: A systematic analysis for the global burden of disease study 2017. *Lancet. Gastroenterol. Hepatol.* 4 (12), 934–947. doi:10.1016/S2468-1253(19)30347-4

Rawla, P., Sunkara, T., and Gaduputi, V. (2019). Epidemiology of pancreatic cancer: Global trends, etiology and risk factors. *World J. Oncol.* 10, 10–27. doi:10.14740/wjon1166

Schweizer, M. T., and Yu, E. Y. (2017). AR-signaling in human malignancies: Prostate cancer and beyond. *Cancers (Basel)* 9 (1), E7. doi:10.3390/cancers9010007

Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 3, Article3. doi:10.2202/1544-6115.1027

Suico, M. A., Shuto, T., and Kai, H. (2017). Roles and regulations of the ETS transcription factor ELF4/MEF. *J. Mol. Cell Biol.* 9 (3), 168–177. doi:10.1093/jmcb/mjw051

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49 (1), D605–D612. doi:10.1093/nar/gkaa1074

Tukiainen, T., Villani, A-C., Yen, A., Rivas, M. A., Marshall, J. L., Satija, R., et al. (2017). Landscape of X chromosome inactivation across human tissues. *Nature* 550 (7675), 244–248. doi:10.1038/nature24265

Wakiya, T., Ishido, K., Yoshizawa, T., Kanda, T., and Hakamada, K. (2021). Roles of the nervous system in pancreatic cancer. *Ann. Gastroenterol. Surg.* 5 (5), 623–633. doi:10.1002/ags3.12459

Wang, J., Guo, X., Xie, C., and Jiang, J. (2017). KIF15 promotes pancreatic cancer proliferation via the MEK-ERK signalling pathway. *Br. J. Cancer* 117 (2), 245–255. doi:10.1038/bjc.2017.165

Wang, M., Gorelick, F., and Bhargava, A. (2021). Sex differences in the exocrine pancreas and associated diseases. *Cell. Mol. Gastroenterol. Hepatol.* 12 (2), 427–441. doi:10.1016/j.jcmgh.2021.04.005

Weniger, M., Honselmann, K. C., and Liss, A. S. (2018). The extracellular matrix and pancreatic cancer: A complex relationship. *Cancers (Basel)* 10 (9), E316. doi:10.3390/cancers10090316

Xia, L., Bouamar, H., Gu, X., Zeballos, C., Qin, T., Wang, B., et al. (2020). Gli2 mediates the development of castration-resistant prostate cancer. *Int. J. Oncol.* 57 (1), 100–112. doi:10.3892/ijo.2020.5044

Xu, S., Xu, H., Wang, W., Li, S., Li, H., Li, T., et al. (2019). The role of collagen in cancer: From bench to bedside. *J. Transl. Med.* 17 (1), 309. doi:10.1186/s12967-019-2058-1

Yonehara, K., Suzuki, M., Yamanouchi, K., and Nishihara, M. (2003). Expression analyses of sex steroid-regulated genes in neonatal rat hypothalamus. *J. Reprod. Dev.* 49 (6), 547–552. doi:10.1262/jrd.49.547

Yu, H., Kim, P. M., Sprecher, E., Trifonov, V., and Gerstein, M. (2007). The importance of bottlenecks in protein networks: Correlation with gene essentiality and expression dynamics. *PLoS Comput. Biol.* 3 (4), e59. doi:10.1371/journal.pcbi.0030059

Zhang, M. A., Rego, D., Moshkova, M., Kebir, H., Chruscinski, A., Nguyen, H., et al. (2012). Peroxisome proliferator-activated receptor (PPAR)α and -γ regulate IFNγ and IL-17A production by human T cells in a sex-specific way. *Proc. Natl. Acad. Sci. U. S. A.* 109 (24), 9505–9510. doi:10.1073/pnas.1118458109

Zhu, H., Wei, M., Xu, J., Hua, J., Liang, C., Meng, Q., et al. (2020). PARP inhibitors in pancreatic cancer: Molecular mechanisms and clinical applications. *Mol. Cancer* 19 (1), 49. doi:10.1186/s12943-020-01167-9

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 03 November 2022
DOI 10.3389/fgene.2022.1009555

Check for updates

OPEN ACCESS

EDITED BY
Ehsan Nazemalhosseini-Mojarad,
Shahid Beheshti University of Medical Sciences, Iran

REVIEWED BY
Feng Jiang,
Fudan University, China
Yingkun Xu,
Chongqing Medical University, China

*CORRESPONDENCE
Zizhi Li,
lizizhi2010@126.com
Ming Tong,
tongming@jzmu.edu.cn

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

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 02 August 2022
ACCEPTED 21 October 2022
PUBLISHED 03 November 2022

CITATION
Hong P, Huang W, Du H, Hu D, Cao Q, Wang Y, Zhang H, Tong S, Li Z and Tong M (2022), Prognostic value and immunological characteristics of a novel cuproptosis-related long noncoding RNAs risk signature in kidney renal clear cell carcinoma. *Front. Genet.* 13:1009555. doi: 10.3389/fgene.2022.1009555

COPYRIGHT
© 2022 Hong, Huang, Du, Hu, Cao, Wang, Zhang, Tong, Li and Tong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Prognostic value and immunological characteristics of a novel cuproptosis-related long noncoding RNAs risk signature in kidney renal clear cell carcinoma

Peng Hong[1†], Weichao Huang[1†], Huifang Du[2], Ding Hu[1], Qingfei Cao[1], Yinjie Wang[1], Huashan Zhang[1], Siqiao Tong[3], Zizhi Li[1*†] and Ming Tong[1*†]

[1]Department of Urology, The First Affiliated Hospital of Jinzhou Medical University, Jinzhou Medical University, Jinzhou, China, [2]Department of Radiology, The First Affiliated Hospital of Nanchang University, Nanchang, China, [3]The First Clinical College of Jinzhou Medical University, Jinzhou Medical University, Jinzhou, China

Background: Cuproptosis has been found as a novel cell death mode significantly associated with mitochondrial metabolism, which may be significantly associated with the occurrence and growth of tumors. LncRNAs take on critical significance in regulating the development of kidney renal clear cell carcinoma (KIRC), whereas the correlation between cuproptosis-related LncRNAs (CRLs) and KIRC is not clear at present. Therefore, this study built a prognosis signature based on CRLs, which can achieve accurate prediction of the outcome of KIRC patients.

Methods: The TCGA database provided the expression profile information and relevant clinical information of KIRC patients. Univariate Cox, Lasso, and multivariate Cox were employed for building a risk signature based on CRLs. Kaplan-Meier (K-M) survival analysis and time-dependent receiver operating characteristic (ROC) curve were employed for the verification and evaluation of the reliability and accuracy of risk signature. Then, qRT-PCR analysis of risk LncRNAs was conducted. Finally, the possible effect of the developed risk signature on the microenvironment for tumor immunization was speculated in accordance with ssGSEA and ESTIMATE algorithms.

Results: A prognosis signature composed of *APCDD1L-DT*, *MINCR*, *AL161782.1*, and *AC026401.3* was built based on CRLs. As revealed by the results of the K-M survival study, the OS rate and progression-free survival rate of high$_{risk}$ KIRC

Abbreviations: RCC, renal cell carcinoma; KIRC, kidney renal clear cell carcinoma; LncRNAs, long non-coding RNAs; CRLs, cuproptosis-related LncRNAs; TCGA, the cancer genome atlas; CNV, copy number variation; TMB, tumor mutation burden; TIDE, tumor immune dysfunction and exclusion; FDR, false discovery rate; OS, overall survival; PFS, progression-free survival; ROC, receiver operating characteristic; AUC, area under the ROC curve; K-M, kaplan-meier; GO, gene ontology; KEGG, kyoto encyclopedia of genes and genomes; GSEA, gene set enrichment analysis; PCA, principal component analysis; FC, fold change.

patients were lower than those of low_{risk} KIRC patients, and the areas under ROC curves of 1, 3, and 5 years were 0.828, 0.780, and 0.794, separately. The results of the immune analysis showed that there were significant differences in the status of immunization and the microenvironment of tumor between groups at low-risk and at high-risk. The qRT-PCR results showed that the relative expression level of *MINCR* and *APCDD1L-DT* were higher in 786-O and 769-P tumor cells than in HK-2 cells, which were normal renal tubular epithelial cells.

Conclusion: The developed risk signature takes on critical significance in the prediction of the prognosis of patients with KIRC, and it can bring a novel direction for immunotherapy and clinical drug treatment of KIRC. In addition, 4 identified risk LncRNAs (especially *APCDD1L-DT* and *MINCR*) can be novel targets for immunotherapy of KIRC patients.

KEYWORDS

kidney renal clear cell carcinoma, cuproptosis-related lncRNAs, prognostic signature, immunological characteristics, the cancer genome atlas, qRT-PCR

Introduction

Renal cell carcinoma (RCC), one of the highly common malignant tumor in urology, takes up approximately 3% of all cancers (Umberto et al., 2019; Zachary et al., 2019). Kidney renal clear cell carcinoma (KIRC) is the most common histological type, taking up nearly 80% of RCC (Umberto and Francesco, 2016). In accordance with the U.S. Cancer Statistics 2022, there will be about 79,000 newly diagnosed renal cancer patients and about 13,920 dead renal cancer patients in 2022 years (Siegel et al., 2022). Surgical resection of diseased kidney tissue is still the main treatment for early KIRC, whereas nearly 30% of patients still have recurrence and metastasis after surgical treatment, thus resulting in poor prognosis for a considerable number of KIRC patients (Steven et al., 2017; Giuseppe et al., 2021; Isabel and Grünwald, 2021). Some new treatment methods have been progressively applied to the clinical treatment of KIRC, including vascular endothelial growth factor tyrosine kinase inhibitors (antiangiogenic agents) and immune-checkpoint inhibitors, thus increasing the rate of survival of patients with advanced KIRC to some extent. However, there are still many reports of KIRC recurrence and progression (Sheng and Rini, 2019; Wenzhong et al., 2021). Most patients with advanced KIRC exhibit high mortality, recurrence rate, as well as metastasis rate. Accordingly, novel biomarkers need to be urgently found to identify high-risk KIRC patients with poor prognosis and to build a risk model to evaluate their prognosis for contributing to the clinical diagnosis and prognosis evaluation of KIRC.

Cell death is a critical step in the development of body (Andreas and Vaux, 2020; Hotchkiss et al., 2009). The body is capable of ensuring a healthy and stable microenvironment by inducing damage, aging, and excess cell death (Andreas and Vaux, 2020). A wide variety of cell death methods have been developed over the past few years, such as apoptosis, autophagy, pyroptosis, ferroptosis, and necroptosis (Daniel and Vince, 2019; Mark, 2019; Stockwell, 2022). Recently, a new process of cell death has been discovered—Cuproptosis, which occurs through the direct combination of copper ions with fatty acyl components in the tricarboxylic acid cycle in mitochondrial respiration, which leads to the aggregation of fatty acyl protein and the following reduction of iron-sulfur cluster proteins, thus resulting in protein toxic stress and eventually cell death (Cobine and Brady, 2022; Daolin et al., 2022; Peter et al., 2022). The mechanism of cuproptosis is not consistent with other known cell death mechanisms. Cell death induced by copper ionophore mainly depends on the accumulation of copper in cells. Cell death induced by copper ionophore is a novel cell death process, obviously inconsistent with the conventional way of cell death. After reviewing the relevant literature, we found that the occurrence of KIRC is usually accompanied by the reprogramming of the tricarboxylic acid cycle. By reducing the energy generated by the tricarboxylic acid cycle, KIRC enables tumor cells to survive under harsh conditions and escape from the surveillance and attack of the immune system (Cancer Genome Atlas Research Network, 2013; Marston et al., 2019). In addition, there have been some reports on the progress of cuproptosis in cancer research, such as cuproptosis-related genes can predict the prognosis and immunotherapy sensitivity of pancreatic cancer patients (Yingkun et al., 2022). We need to study the roles and specific mechanisms of cuproptosis in tumorigenesis and development in depth and find specific biomarkers, which can show novel directions for KIRC diagnosis and treatment.

Long non-coding RNAs (LncRNAs) refer to non-coding RNAs containing the length of over 200 nucleotides. RNA polymerase II transcribe the LncRNAs. In recent years, many studies have reported that LncRNAs take on vital significance in the occurrence and growth of KIRC. For instance, *LncRNA*

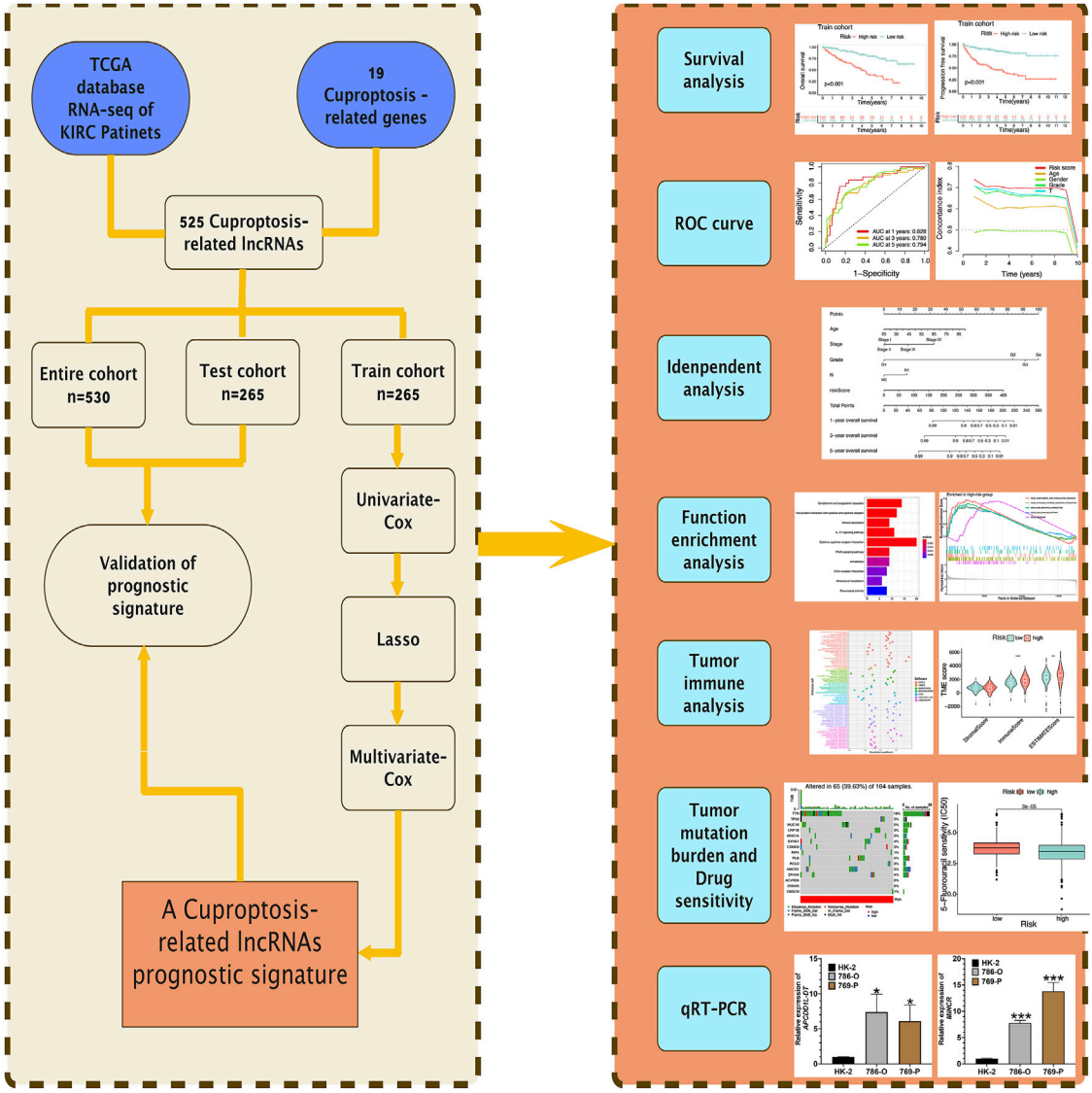


FIGURE 1
The research steps of this article.

SNHG1 is capable of activating *STAT3* and *PD-L1* as competitive endogenous RNA of *miR-129-3p* (Pei et al., 2021), which can lead to the regulation of the immune escape of renal cell carcinoma. The result of another study suggests that knocking down *LncRNA LINC00944* leads to significantly inhibits the proliferation and migration of renal cell carcinoma and facilitates AKT phosphorylation (Chen and Zheng, 2021). *LncRNA GAPLINC* can promote the tumorigenesis of renal cell carcinoma by targeting *miR-135b-5p/CSF1* axis (Wang et al., 2021). In addition, many LncRNAs-related prognostic models have been reported in KIRC. For instance, the m7G-related LncRNAs prognostic model can accurately achieve the prediction of the prognosis of KIRC patients (Jie and Chunyang, 2022), immune-associated LncRNAs prognosis signature has prognostic significance in KIRC (Seema et al., 2019), and ferroptosis-related LncRNAs can provide accurate prognosis prediction for KIRC patients (Xiao-Liang et al., 2021). However, cuproptosis is a novel cell death mode, and there has been little research about the Cuproptosis-related LncRNAs (CRLs) prognosis model in KIRC.

In this study, based on CRLs, our team built and verified a risk signature for evaluating and improving the prognosis of KIRC patients, and verified its clinical value. In addition, it shows the feasibility that the risk signature can make personalized immunotherapy and targeted therapy for KIRC patients.

TABLE 1 Clinical information of KIRC patients in Train cohort, Test cohort, and Entire TCGA cohort.

Covariates	Type	Entire TCGA cohort	Test cohort	Train cohort	*p*value
Age	≤65	348 (65.66%)	174 (65.66%)	174 (65.66%)	1
	>65	182 (34.34%)	91 (34.34%)	91 (34.34%)	
Gender	FEMALE	186 (35.09%)	87 (32.83%)	99 (37.36%)	0.3168
	MALE	344 (64.91%)	178 (67.17%)	166 (62.64%)	
Grade	G1	14 (2.64%)	5 (1.89%)	9 (3.4%)	0.7295
	G2	227 (42.83%)	112 (42.26%)	115 (43.4%)	
	G3	206 (38.87%)	103 (38.87%)	103 (38.87%)	
	G4	75 (14.15%)	41 (15.47%)	34 (12.83%)	
	GX	5 (0.94%)	3 (1.13%)	2 (0.75%)	
	unknow	3 (0.57%)	1 (0.38%)	2 (0.75%)	
Stage	Stage I	265 (50%)	137 (51.7%)	128 (48.3%)	0.3438
	Stage II	57 (10.75%)	24 (9.06%)	33 (12.45%)	
	Stage III	123 (23.21%)	65 (24.53%)	58 (21.89%)	
	Stage IV	82 (15.47%)	36 (13.58%)	46 (17.36%)	
	unknow	3 (0.57%)	3 (1.13%)	0 (0%)	
T	T1	21 (3.96%)	14 (5.28%)	7 (2.64%)	0.7918
	T1a	140 (26.42%)	67 (25.28%)	73 (27.55%)	
	T1b	110 (20.75%)	59 (22.26%)	51 (19.25%)	
	T2	55 (10.38%)	22 (8.3%)	33 (12.45%)	
	T2a	10 (1.89%)	4 (1.51%)	6 (2.26%)	
	T2b	4 (0.75%)	2 (0.75%)	2 (0.75%)	
	T3	5 (0.94%)	2 (0.75%)	3 (1.13%)	
	T3a	120 (22.64%)	62 (23.4%)	58 (21.89%)	
	T3b	52 (9.81%)	27 (10.19%)	25 (9.43%)	
	T3c	2 (0.38%)	1 (0.38%)	1 (0.38%)	
	T4	11 (2.08%)	5 (1.89%)	6 (2.26%)	
M	M0	420 (79.25%)	216 (81.51%)	204 (76.98%)	0.5249
	M1	78 (14.72%)	35 (13.21%)	43 (16.23%)	
	MX	30 (5.66%)	14 (5.28%)	16 (6.04%)	
	unknow	2 (0.38%)	0 (0%)	2 (0.75%)	
N	N0	239 (45.09%)	125 (47.17%)	114 (43.02%)	0.5913
	N1	16 (3.02%)	7 (2.64%)	9 (3.4%)	
	NX	275 (51.89%)	133 (50.19%)	142 (53.58%)	

Materials and methods

Data sources

Figure 1 presents the flow of this study. The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov) is a database jointly developed by the National Cancer Institute and the National Human Genome Research Institute, which contains clinical data, genomic variation, mRNA expression, miRNA expression, methylation, and other data of various human cancers. TCGA database provided 72 healthy renal tissular specimens as well as RNA sequencing data of 539 tissular specimens of KIRC in the format of "HTSEQ-FPKM, TCGA-KIRC" (Marston and J, 2019). Furthermore, the TCGA database provided relevant clinical information. We employed 530 KIRC samples containing complete RNA sequencing information and clinical information for the following analysis by excluding several samples containing not complete data. The 530 KIRC samples were randomized 1:1 as the group of testing (n = 265, to verify the CRLs risk signature) as well as the group of training (n = 265, to develop the CRLs risk signature) (Table 1).

UCSC Xena (https://xenabrowser.net/) offered copy number variation (CNV) data and tumor mutation data for KIRC patients. Choose the "KIRC.Varscan. Somatic. maf." file for subsequent tumor mutation burden (TMB) analysis.

Moreover, tumor immune dysfunction and exclusion (TIDE) score data for each KIRC patient were obtained from the TIDE website (http://tide.dfci.harvard.edu). In accordance with previous studies and literature reports, we obtained a total of 19 cuproptosis-related genes (*DLSTGCSH, DBT, CDKN2A, GLS, MTF1, PDHB, PDHA1, DLAT, DLD, LIPT2, LIPT1, LIAS, FDX1, SLC31A1, ATP7A, ATP7B, NLRP3, NFE2L2*) (Ok et al., 2014; Francesco and Serena, 2017; Anna et al., 2018; Ze et al., 2019; Jianjian et al., 2021; Peter et al., 2022). The above data were all from public databases, which ensured the reproducibility of the study.

Expression and mutation analysis of cuproptosis-related genes

The expression differences of cuproptosis-related genes in healthy renal tissular specimens and KIRC tissular specimens were analyzed by R package "limma" in R program, and corresponding boxplots were plotted by R package "ggpubr". Mutations in cuproptosis-related genes were represented by a waterfall plot by R Package "maltools". The CNV frequency of cuproptosis-related genes was shown in the bar chart.

Identification of cuproptosis-related LncRNAs

TCGA-KIRC transcriptome data were divided into mRNA and LncRNA using Perl script based on GTF files (human transcriptome annotates information). The correlation between cuproptosis-related genes and LncRNAs expression was analyzed using R package "limma" through Pearson correlation analysis, with |correlation coefficient| >0.4, $p < 0.001$ as the filter criterion to obtain CRLs.

Establishment and validation of prediction signature based on cuproptosis-related LncRNAs

The Train cohort was used for the development of the risk signature, and the Test cohort as well as the Entire TCGA cohort were employed to verify the built risk signature. Based on the overall survival (OS) time in clinical information of patients of KIRC, univariate Cox analysis was employed to evaluate the prognosis significance of CRLs (False Discovery Rate (FDR) < 0.05). In order to prevent over-fitting in the development of the risk signature, R package "glment" was employed to further optimize the selection of prognostic CRLs using Lasso regression analysis. We carried out multivariate Cox regression analysis for the above most representative prognostic CRLs to obtain the hazard ratio (HR) and regression coefficients for the respective risk LncRNA. Based on the mentioned investigation, detailed risk LncRNAs and the regression coefficients were presented, and risk signatures based on CRLs were built. The risk score of the respective KIRC patient was obtained as follows:

$$\text{risk score} = \sum_{i=1}^{n} coef\,(LncRNA)^{*}\, Exp\,(LncRNA)$$

All KIRC patients were classified into two groups, including $risk_{high}$ and $risk_{low}$ according to the median risk score of the Train cohort as the cutoff. R-package "survival" and "survminer" were adopted to analyze whether OS and progression-free survival (PFS) of KIRC patients are different between the two risk groups through Kaplan-Meier (K-M) survival study. R-package "survival", "survminer" and "timeROC" were employed to generate multiple receiver operating characteristic (ROC) curves. We obtained the area under the ROC curve (AUC) for verifying the predictive value of the prognostic signature and evaluating the accuracy of the risk signature in the prediction of KIRC patient's prognosis. In addition, we adopted the concordance index for evaluating the prediction accuracy of the risk signature. Lastly, univariate Cox and multivariate Cox regression analyses were used for the investigation of whether the risk signatures or other clinical characteristics may be the independent prognostic indicators of KIRC patients. The results of regression analysis were presented in forest maps. The above analyses were validated in both the Test cohort and the Entire TCGA cohort.

Establishment and calibration of nomogram

A model for the identification of nomogram risk was developed with the use of R-package "rms" based on some independent factors for prognosis in the clinical field and risk scores. Nomogram is capable of quantifying the factor for KIRC patient's prognosis and carry out the quantitative prediction of KIRC patients' prognosis. Next, we generated calibration curves for illustrating the built nomogram's prediction effect. The calibration of the respective model was presented by the above curves in accordance with the condition that the actual time of survival of KIRC patients was consistent with the estimated time of survival of KIRC patients. The y-axis represents the actual time of survival of KIRC patients. The estimated time of survival of KIRC patients was represented by the x-axis. The perfectly predicted model for risk was represented by the light grey line. Better prediction was represented by the light grey line closer to the diagonal, and the nomogram performance is represented by the pink solid line.

Functional enrichment study

R package "limma" (screening criteria: |log2Fold Change (FC)| > 1, FDR <0.05) was adopted to screen genes with differential expression in $risk_{high}$ and $risk_{low}$ groups. Subsequently, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were performed for the above genes (R package "clusterProfiler"). Gene Set Enrichment Analysis (GSEA) was conducted based on the gene set files of "c2.cp.kegg.v7.4.symbols.gmt", R package "limma", "clusterProfiler", "org.Hs.eg.db" and "enrichplot" were employed to identify significantly enriched pathways in $risk_{high}$ and $risk_{low}$ KIRC patients (|log2FC| > 1, FDR <0.05), respectively.

Analysis of tumor immune microenvironment

The situation of a wide variety of immune cell infiltration in different KIRC samples was obtained using seven 7 algorithms (CIBERSORT-ABS, CIBERSORT, EPIC, MCPCOUNTER, QUANTISEQ, TIMER, XCELL). Next, the Spearman correlation study was used for investigating the correlation between different immune cell infiltration degrees and risk scores. The above algorithms were systematically benchmarked, and each of them exhibited unique performance and advantages. Then, the content of immunocytes in the respective KIRC sample was quantified using the algorithm of "Cell-type Identification based on the Estimation of Relative Subsets of RNA Transcripts (CIBERSORT)" (Newman et al., 2015; Binbin et al., 2018). The single sample Gene Set Enrichment Analysis (ssGSEA) was used for the analysis of differences in the score of enrichment of 13 immunization-associated pathways and 16 immunocytes in different risk groups, the analysis results were presented in multi-box diagram. The StromalScore, ImmuneScore, and EstimateScore (StromalScore + ImmuneScore) were obtained using the algorithm of "Estimating Stromal and Immunocytes in MAlignant Tumor tissues based on Expression data (ESTIMATE)" (Kosuke et al., 2013) for the respective KIRC patient and then for the comparison of the score differences of a wide variety of risk groups. Moreover, the comparison of the expression differences for some common immune-checkpoints genes in different risk groups and then used the TIDE algorithm to predict potential immunotherapy responses. Based on the immunosuppressive factors (including short survival after ICB treatment, poor efficacy of immune checkpoint blocking therapy (ICB), as well as high TIDE score), TIDE evaluated two mechanisms of tumor immune escape (rejection of cytotoxic T lymphocyte and tumor-infiltrating CTL dysfunction) by employing several markers of gene expression (Peng et al., 2018).

Prediction of potential drug sensitivity

Based on R package "pRRophetic", the prediction was conducted, the sample's maximum 50% inhibition concentration (IC50) was predicted using ridge regression, and IC50 represented 50% of the suppressed cells, i.e., the cell survival rate was half of the control sample. In other words, the lower the IC50 value of the corresponding drug concentration, the more sensitive KIRC patients will be to the drug (Paul et al., 2014).

Cell culture

In the First Affiliated Hospital of Jinzhou Medical University, the human KIRC cell lines 786-O and 769-P used in this study were preserved. Procell Company in Wuhan, China provided normal renal tubular epithelial cell line HK-2. RPMI-1640 medium (Hyclone) achieved the culture of 768-O and 769-P cell lines, while MEM medium (Hyclone) achieved the culture of HK-2 cells, with the addition of 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin, respectively. All cells were incubated at 37°C containing 95% air and 5% carbon dioxide.

RNA extraction and quantitative real-time polymerase chain reaction

In accordance with cultured 786-O, 769-P, and HK-2 cells, total RNA was obtained by TRIzol (Beyotime, Shanghai, China), and then reverse transcribed into cDNA by reverse transcription kit (Beyotime, Shanghai, China) on PCR Cycler (Bio-Rad, United States). We employed SYBR-Green mixture (Beyotime, Shanghai, China) and Bio-Rad chemiluminescence imager (Bio-Rad, United States) for qRT-PCR. All the above experimental steps were conducted in accordance with the product instructions, and the amplified primer sequences were as follows:

***APCDD1L-DT*, forward primer:** GAGCCTTGGAAAGGA GGACC, reverse primer: GATCCATGCAGGTGGGAACA.

***MINCR*, forward primer:** TCCAAGGTCGATTTTCTTAGC CA, reverse primer: CCCTTTTCAGTTCACAAGCGT.

***GAPDH*, forward primer:** TCGTGGAAGGACTCATGACC, reverse primer: TCCACCACCCTGTTGCTGTA.

GAPDH served as internal control, the relative expression was examined by $2^{-\Delta\Delta Ct}$. The experiment was repeated three times.

Statistical analysis

To conduct the K-M survival analysis for generating the survival curve, we performed Log-rank test. For examining the

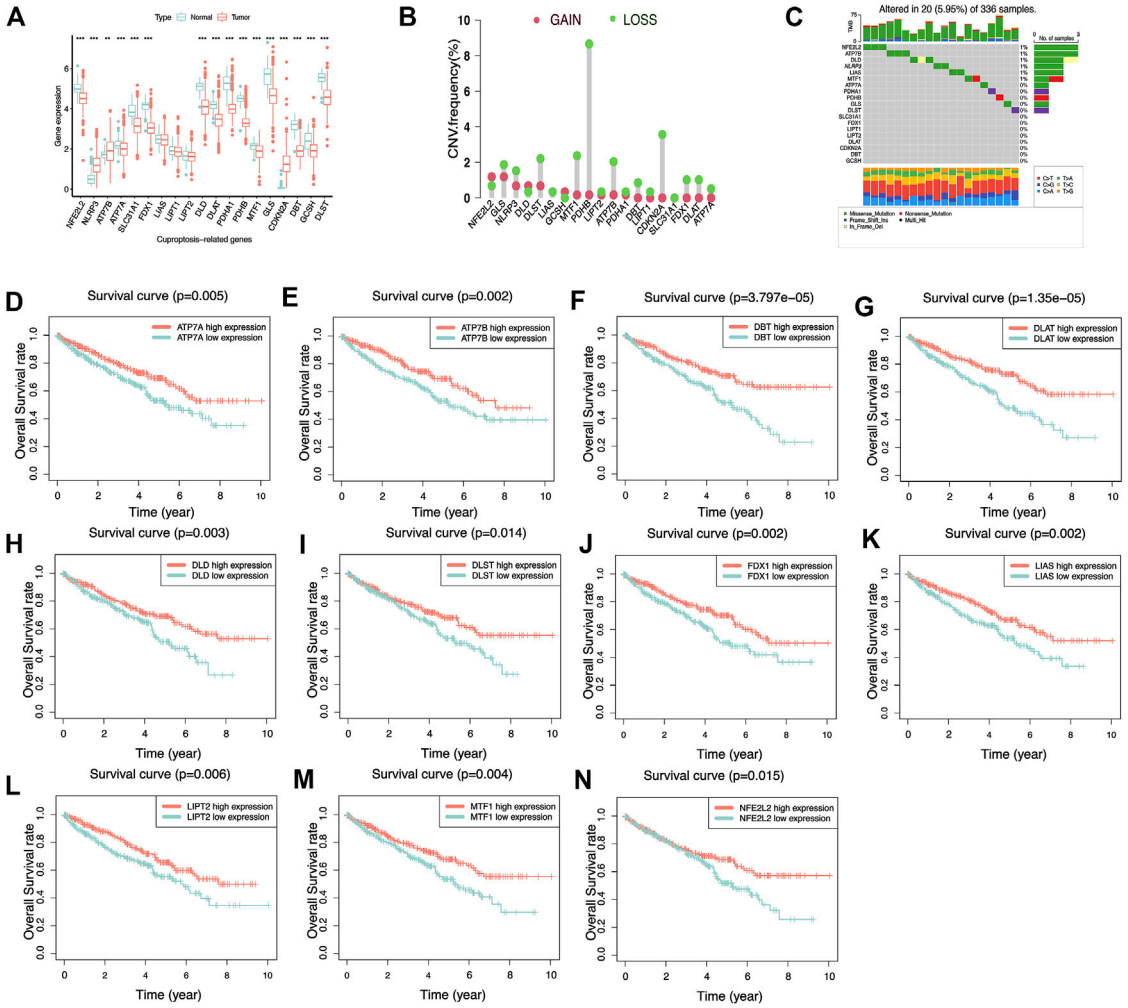


FIGURE 2
Biological characteristics of 19 cuproptosis-related genes. **(A)** Differential expression of cuproptosis-related genes in KIRC and normal renal tissues. **$p< 0.01$, ***$p< 0.001$. **(B)** Copy number variation frequency plots showed that most cuproptosis-related genes had copy number deletions. **(C)** The waterfall plot shows the frequency of cuproptosis-related genes mutations in KIRC. The different colored squares at the bottom represent different types of mutations. **(D–N)** K–M survival curves of cuproptosis-related genes in KIRC. Most of the cuproptosis-related genes in KIRC have prognostic significance and act as an adverse prognostic factor.

diversities between a variety of classified data or different datasets, we carried out the Chi-square test. For determining the difference between the above two groups, we carried out Wilcoxon-rank test. For the analysis of correlation, we employed Spearman method. For the assessment of the effect exerted by gene expression, clinical features, and risk signature on patients' prognosis, we carried out cox proportional regressive analysis. The above statistical methods achieved statistical significance if $p < 0.05$. The analysis was performed based on R version 4.1.0 and the feature package.

Results

Biological characterization of 19 cuproptosis-related genes in KIRC

First, We extracted the expression data of 19 genes associated with cuproptosis from 539 KIRC tissue specimens and 72 normal kidney tissue specimens and further analyzed the expression of the above 19 genes associated with cuproptosis in KIRC tissue specimens and normal kidney tissue specimens. It was found that the expression of most genes associated with cuproptosis was

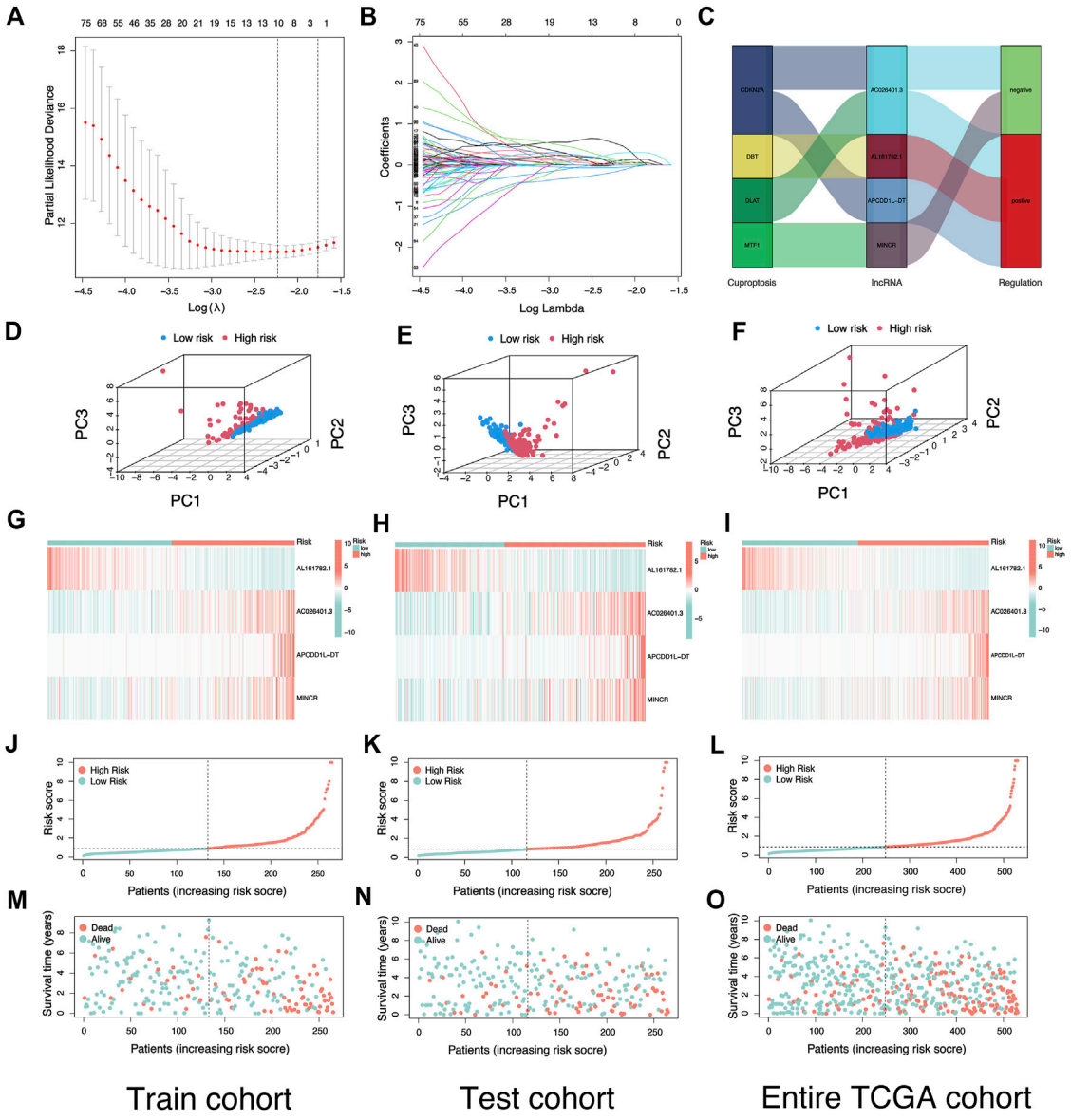


FIGURE 3
Construction and verification of risk signature based on cuproptosis-related lncRNAs. **(A)** Cross-validation for variable selection in the LASSO model. **(B)** Lasso coefficient distribution of risk cuproptosis-related lncRNAs. **(C)** Sankey diagram illustrates the regulatory relationship between Cuproptosis-related genes and lncRNAs. **(D–F)** The Principal Component Analysis in Train cohort, Test cohort, and Entire TCGA cohort based on different risk groups. **(G–I)** Heatmap of the expression of four risk lncRNAs in different risk groups in Train cohort, Test cohort, and Entire TCGA cohort. **(J–L)** Risk plot distribution of Train cohort, Test cohort, and Entire TCGA cohort. **(M–O)** Distribution of survival status of KIRC patients in Train cohort, Test cohort, and Entire TCGA cohort.

lower expression in KIRC tumor tissue, compared with that in normal kidney tissue specimens (Figure 2A). We analyzed the 19 genes associated with cuproptosis CNV and somatic mutations in KIRC, at the level of CNV, we found that most of the genes associated with cuproptosis were focused on the loss of copy number (Figure 2B). In 336 KIRC specimens, there were 20 specimens carry genes associated with cuproptosis mutations, and the *NFE2L2* mutation frequency was the highest (Figure 2C). It is speculated that CNV differences and genetic mutations may mediate the difference in expression of genes associated with cuproptosis between normal kidney tissue and KIRC tissue. Subsequently, the influence of the above genes associated with cuproptosis on the OS rate of KIRC patients was analyzed, and it was also surprising to find that most of the genes associated with cuproptosis were significantly associated with the survival of KIRC (Figure 2D–N). The above analyses

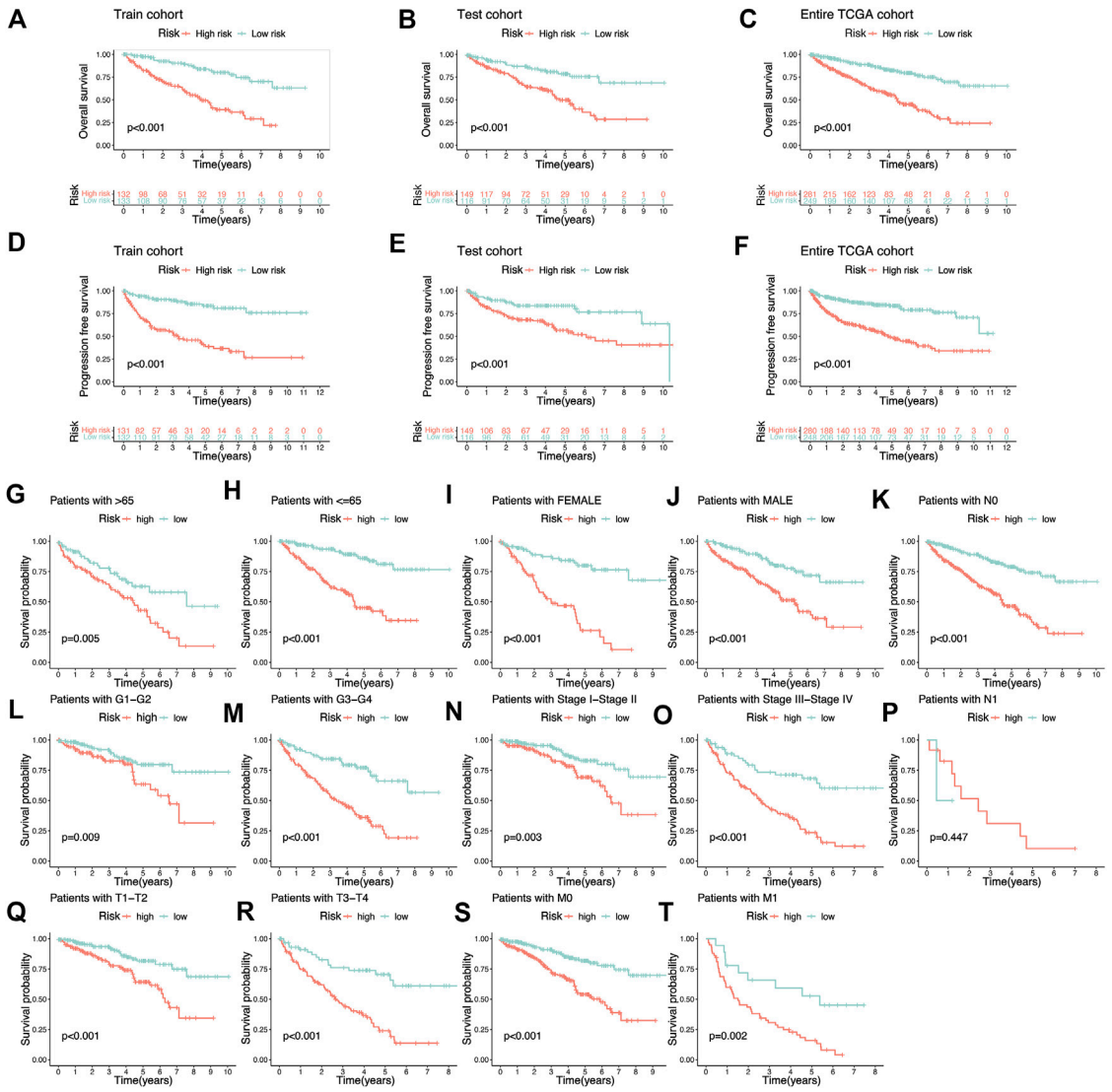


FIGURE 4
K-M survival analysis based on cuproptosis-related lncRNAs risk signature. **(A–C)** The Overall survival K-M survival curves of different risk groups in Train cohort, Test cohort, and Entire TCGA cohort. The overall survival of patients with high-risk KIRC is lower than that of patients with low-risk KIRC. **(D–F)** The Progression-free survival K-M survival curves of different risk groups in Train cohort, Test cohort, and Entire TCGA cohort. The Progression-free survival of patients with high-risk KIRC is lower than that of patients with low-risk KIRC. **(G–T)** The Overall survival K-M survival curves of different risk groups with different clinical features.

suggested that the imbalance of the expression of genes associated with cuproptosis may affect the occurrence and growth of KIRC patients.

Establishment and validation of prediction signature based on LncRNAs associated with cuproptosis

A total of 13349 LncRNAs expression data were obtained from TCGA-KIRC transcriptome data. Subsequently, we carried out Pearson correlation study in accordance with the expression of 19 genes associated with cuproptosis and LncRNAs expression data. A total of 525 LncRNAs were consistent with the corresponding conditions, that was, $|$correlation coefficient$| > 0.4$, $p < 0.001$, and were defined as CRLs (Supplementary Files S1). In the Train cohort, 525 CRLs were analyzed by univariate Cox regression analysis, and 192 CRLs with prognostic values were obtained. Next, we performed LASSO regression analysis on 192 prognostic CRLs for eliminating highly correlated prognostic LncRNAs and avoiding overfitting (Figure 3A,B) for the optimization of

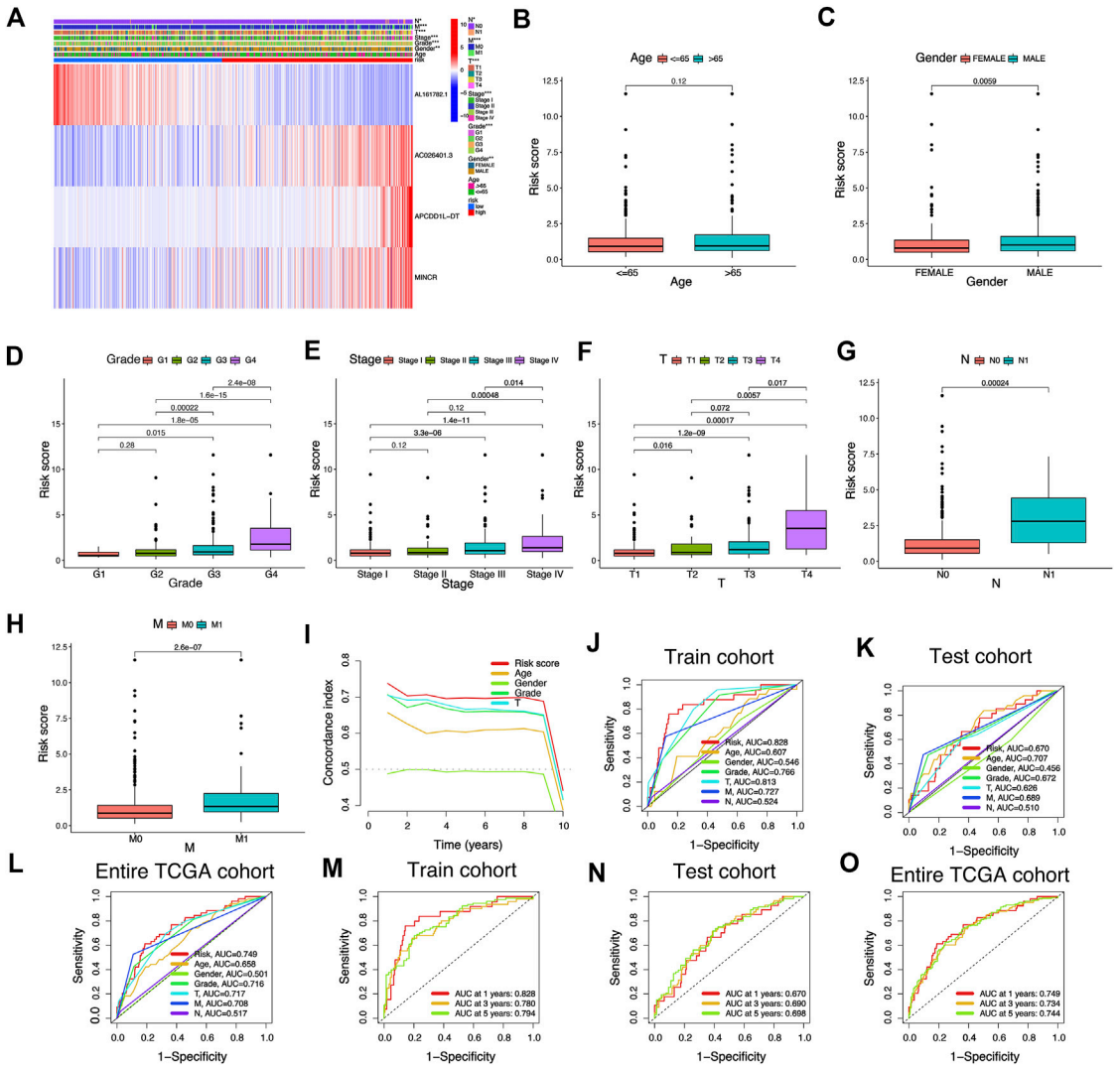


FIGURE 5
Association between risk signature scores composed of cuproptosis-related lncRNAs and clinical features. **(A)** The heatmap shows the clinical characteristics and expression of 4 risk cuproptosis-related lncRNAs in 530 KIRC patients. **(B–H)** Differential distribution of risk scores among KIRC patients with different clinical characteristics. **(I)** The consistency index showed that the risk score was better than other clinicopathological features in predicting the prognosis of KIRC patients. **(J–O)** ROC curves for risk signature composed of cuproptosis-related lncRNAs in Train cohort, Test cohort, and Entire TCGA cohort.

the developed signatures. We found the 10 most representative CRLs. Cross-validation results showed that LASSO regression analysis was the best. Multivariate Cox regression analysis was performed for the 10 most representative prognostic CRLs to obtain the hazard ratio (HR) and regression coefficients for 4 risk LncRNAs. The risk score for each KIRC patient can be obtained as:

$$\begin{aligned}\text{Risk score} = {} & (-1.15122394874834{*}\text{Exp}_{AL161782.1}) \\ & + (0.4711103719724987{*}\text{Exp}_{AC026401.3}) \\ & + (0.678892201655986{*}\text{Exp}_{APCDDIL-DT}) \\ & + (0.468667562066302{*}\text{Exp}_{MINCR})\end{aligned}$$

The Sankey diagram showed the regulatory relationship between 4 risk LncRNAs and genes associated with cuproptosis (Figure 3C). All KIRC patients were divided into $\text{risk}_{\text{high}}$ and risk_{low} groups in accordance with the median risk score in the Train cohort as a critical point. Principal component analysis (PCA) showed that risk scores could significantly divide KIRC patients into $\text{risk}_{\text{high}}$ and risk_{low} groups (Figures 3D–F). We also compared the expression of the above 4 risk LncRNAs in the $\text{risk}_{\text{high}}$ and risk_{low} groups (Figure 3G–I). Subsequently, we rank the survival state and risk score distribution of patients with KIRC, and we could see that the number of deaths of patients with KIRC was increasing with the increase of risk score

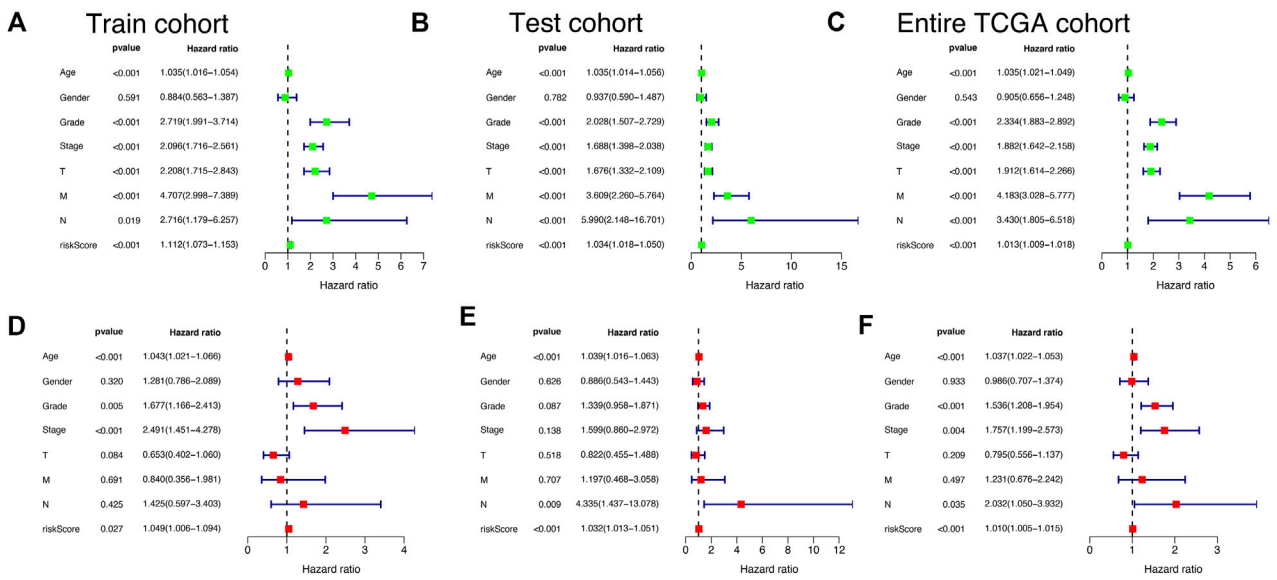


FIGURE 6
Independent prognostic validation of risk signatures. Univariate and multivariate Cox regression analysis results were presented in the form of forest plots, indicating that risk signature was an independent prognostic factor for KIRC patients in Train cohort **(A and D)**, Test cohort **(B and E)**, and Entire TCGA cohort **(C and F)**.

(Figure 3J–O). In accordance with the K-M survival analysis, we also found that the OS rate of the $risk_{high}$ group was lower than that of the $risk_{low}$ group (Figure 4A–C). In addition, a significant difference was identified in PFS rate between the $risk_{high}$ and $risk_{low}$ groups, that is, the PFS rate of the $risk_{high}$ group was lower than that of the $risk_{low}$ group (Figure 4D–F). The same survival analysis results were obtained for KIRC patients with different clinical characteristics (except stage N1 KIRC patients) (Figure 4G–T). The explanation for no significant difference in OS rate between $risk_{high}$ and $risk_{low}$ groups of KIRC patients in N1 may be that there were fewer KIRC patients in N1, revealing that the developed risk signature is applicable to KIRC patients exhibiting nearly all clinical characteristics. The heatmap showed each KIRC patient's clinical characteristics and risk scores (Figure 5A). As depicted in the figure, there were significant differences in the risk scores of KIRC patients with different clinical characteristics. To be specific, the risk scores of KIRC patients in M1, N1, G3-G4, and T3-T4 were higher than those of KIRC patients in MO, N0, G1-G2, and T1-T2, respectively (Figure 5B–H). In general, the risk scores of patients of KIRC with advanced was often higher.

Assessment of risk signature

We adopted the area under ROC curve for evaluating the accuracy of risk signature in the prediction of the outcome of patients with KIRC. The results showed that the areas under ROC curve of the Train cohort 1, 3, and 5 years were 0.828, 0.780, and 0.794 respectively (Figure 5J); the areas under ROC curve of the Test cohort 1, 3, and 5 years were 0.670, 0.690 and 0.698 respectively (Figure 5K); the areas under ROC curve of the Entire TCGA cohort 1, 3 and 5 years were: 0.749, 0.734 and 0.744 respectively (Figure 5L). In addition, we also found that the risk score was better than other clinical variables in the prediction of the prognosis of patients with KIRC (Figure 5M–O). The concordance index also confirmed that the risk score was more accurate in predicting KIRC patients' outcomes than other clinicopathological features (Figure 5I). The above results demonstrated the ability of our risk signature to accurately achieve the prediction of the prognosis of patients with KIRC.

Risk signature have excellent independent prognostic value

Univariate-multivariate Cox regression analysis was performed to determine whether the risk score was an independent predictor of patients' outcome, regardless of other clinical characteristics. In the Train cohort, univariate Cox results showed that risk score was significantly associated with patient prognosis (Figure 6A). Multivariate Cox regression results further demonstrated that risk score can be used as a biomarker of independent prognosis for KIRC patients regardless of clinical characteristics (Figure 6D). The same uni-multi Cox regression analysis results were obtained for

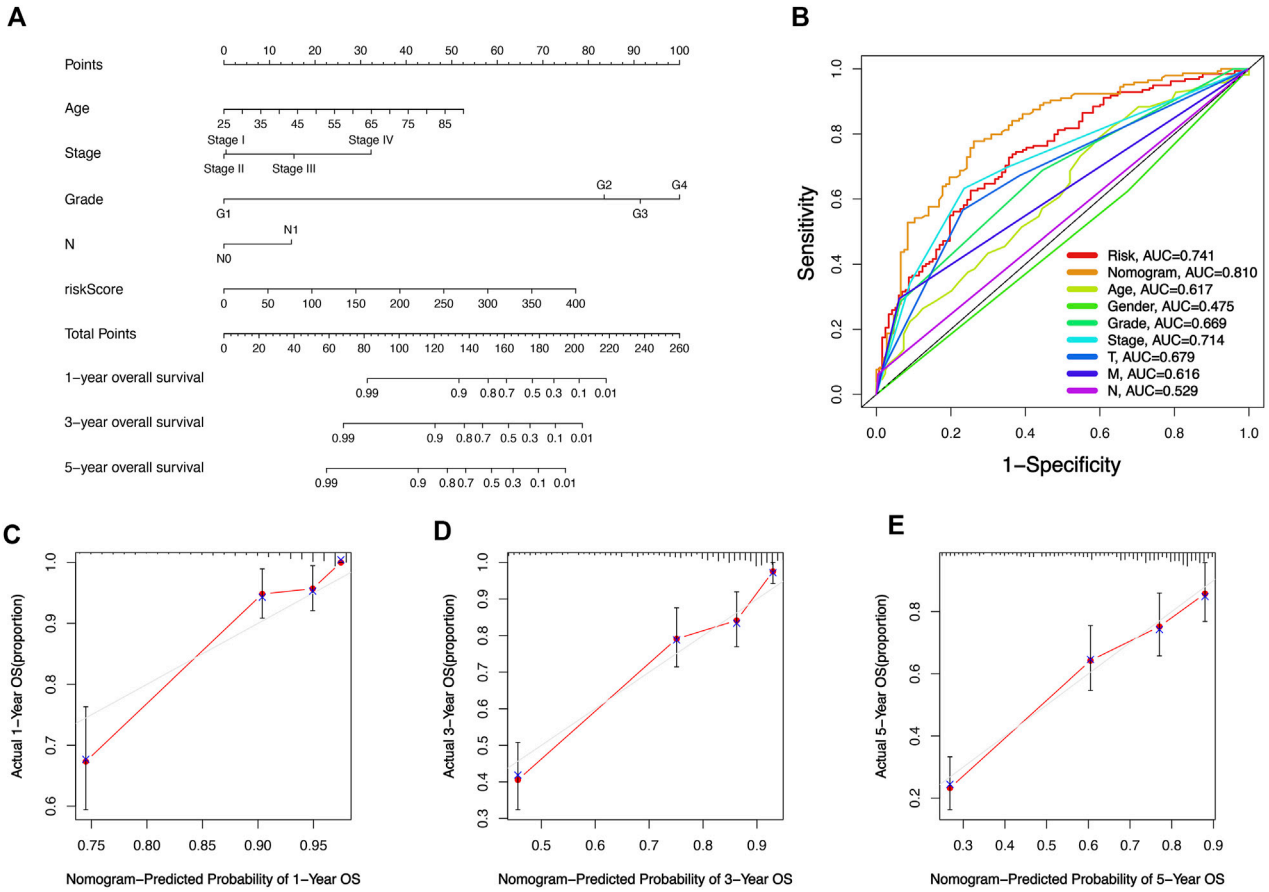


FIGURE 7
Establishment and validation of a Nomogram to quantitatively predict the prognosis of KIRC patients. **(A)** Nomogram to quantitatively predict 1, 3, and 5-year overall survival in KIRC patients. **(B)** ROC curves for Nomogram, risk scores, and Other clinical features. **(C–E)** The calibration chart shows that the Nomogram has a pretty predictive ability.

the Test cohort and the Entire TCGA cohort (Figure 6B,C,E,F). The above results suggest that the risk signature built based on CRLs has certain significance for the prognostic assessment of KIRC patients.

Nomogram

In order to make our risk signature fully utilized in the clinical diagnosis and treatment of KIRC. We built a nomogram survival prediction map based on four independent clinical prognostic factors (age, grade, N, and stage) and risk score to quantitatively predict the 1, 3, and 5-year survival of KIRC patients (Figure 7A). By comparing the area under the ROC curve, we could see that our nomogram had good performance in the prediction of the prognosis of KIRC patients compared to other predictive indexes (Figure 7B). Subsequently, the calibration curves were employed to verify the prediction ability and accuracy of the nomogram, and the results showed that the nomogram could accurately achieve the prediction of the prognosis of patients with KIRC (Figure 7C–E), which also illustrated the value and potential of the nomogram in clinical application to achieve the prediction of the prognosis of patients with KIRC.

Functional enrichment analysis based on risk signature

Because the prognosis of KIRC patients in different risk groups was significantly different, we performed KEGG, GO, and GSEA enrichment analysis to preliminarily exploring the potential biological function differences between the two risk groups. We screened a total of 588 genes with differential expression from the $risk_{high}$ and $risk_{low}$ groups, and KEGG, GO enrichment analysis provided us with a biological

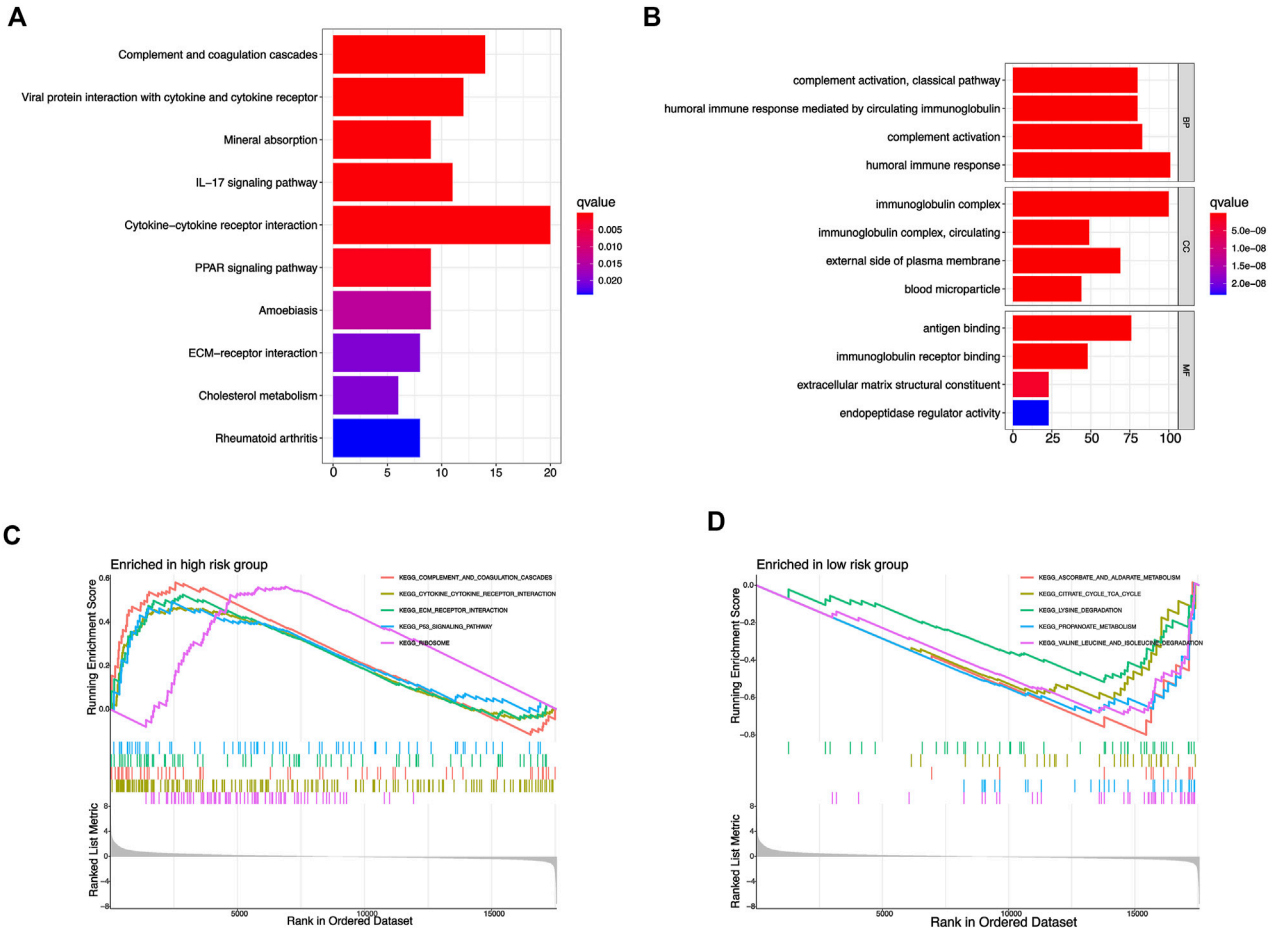


FIGURE 8
Functional enrichment analysis based on risk signature. **(A)** KEGG enrichment analysis of differential genes from risk signature. **(B)** GO enrichment analysis of differential genes from risk signature. **(C,D)** GSEA enrichment analysis of different risk groups.

understanding of the above genes (Supplementary Files S2). KEGG enrichment analysis suggested that the above genes were significantly associated with functional pathways such as IL 17 signaling pathway, Cytokine-cytokine receptor interaction, PPAR signaling pathway, etc (Figure 8A), and GO enrichment analysis suggested that the above genes were significantly related to biological behavior such as humoral immune response, immunoglobulin complex, immunoglobulin, etc (Figure 8B). In addition, we used GSEA software for the investigation of biological pathways that are enriched in $risk_{high}$ and $risk_{low}$ groups. With $p < 0.05$ as the standard, 41 biological pathways were defined enriched. The top 5 pathways enriched in $risk_{high}$ groups were P53 signaling pathway, ECM receptor interaction, Complement, and coagulation cascades, Cytokine-cytokine receptor interaction, and Ribosome (Figure 8C). The top 5 pathways enriched in $risk_{low}$ groups were: Ascorbate and aldarate metabolism, Citrate cycle tca cycle, Lysine degradation, Propanoate metabolism, Valine leucine and isoleucine degradation (Figure 8D). The above mechanisms might partly explain that why $risk_{high}$ KIRC patients tend to have worse clinical outcomes than $risk_{low}$ KIRC patients.

The tumor immune microenvironment of KIRC patients was analyzed based on risk score

Since the results of the previous functional analysis showed a close association with function and pathway of immunity, we further studied the status of immunization of KIRC patients between the $risk_{high}$ and $risk_{low}$ groups. First, a variety of algorithms were adopted to study the correlation between risk score and tumor immune cell infiltration, and there was a significant correlation between a variety of immunocytes and risk score (Figure 9A). Based on the ESTIMATE algorithm, we could conclude that the immune score of KIRC patients in the

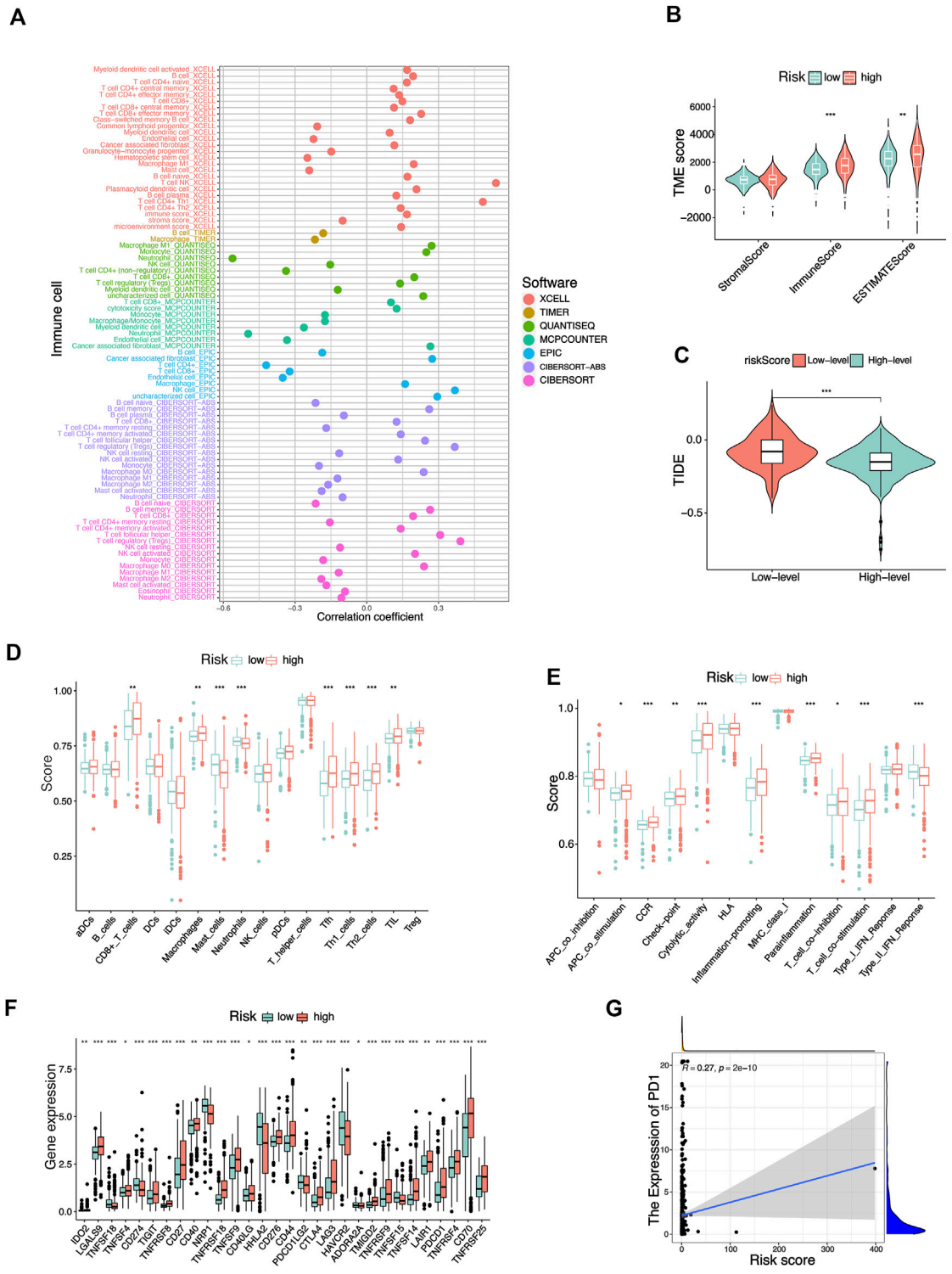


FIGURE 9
Potential effects of risk signatures on tumor immune microenvironment in KIRC patients. **(A)** The correlation between different immune cells and risk scores was analyzed based on multiple algorithms. **(B)** Differences in StromalScore, ImmuneScore, and ESTIMATEScore in KIRC patients with different risk groups (ESTIMATE algorithms). **(C)** Differences in immunotherapy sensitivity among different risk groups (TIDE algorithms). Differences in enrichment scores of 16 immune cells **(D)** and 13 immune-related pathways **(E)** in different risk groups. **(F)** Differential expression of common immune-checkpoints in different risk groups. **(G)** Correlation between risk scores and PD1 expression in KIRC patients. *$p < 0.05$, **$p < 0.01$,***$p < 0.001$.

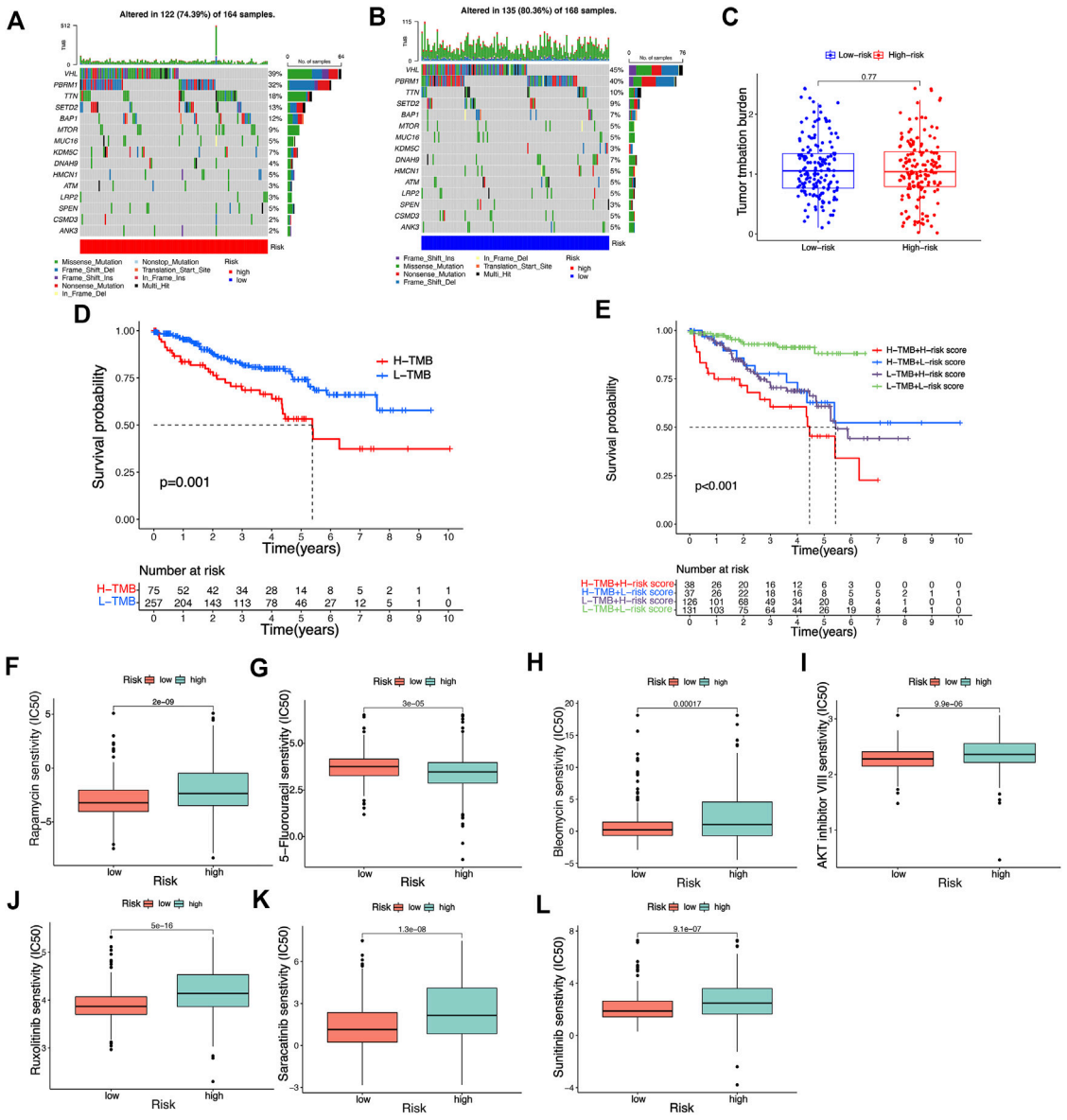


FIGURE 10
Analysis of tumor mutation burden and prediction of potential drug susceptibility in different risk groups. **(A,B)** The waterfall plot illustrates the type and frequency of tumor mutational burden in high-risk and low-risk KIRC patients. **(C)** Differences in tumor mutation burden scores among different risk groups. **(D,E)** K-M survival curves of KIRC patients with different risk scores and tumor mutation burden scores. **(F–L)** Differences in IC50 values of common drugs in different risk groups.

$risk_{high}$ group is higher than that in the $risk_{low}$ group (Figure 9B). Compared with $risk_{low}$ KIRC patients, the $risk_{high}$ KIRC patients had more active status of immunization and lower tumor purity. Then, we compared the differences of 16 types of immune cell enrichment scores and 13 types of immunization-associated function enrichment scores between the $risk_{high}$ and $risk_{low}$ groups (ssGSEA algorithm). Most immune cell and immunization-associated pathway enrichment scores were higher in the $risk_{high}$ group, such as T cell co-stimulation, parainflammation, APC co-stimulation, Check-point, $CD8^+$ T cell, Th1 cells, Th2 cells (Figure 9D,E). Thus it could be shown that immune activity was stimulated in patients with $risk_{high}$ KIRC. The expression of some common immune checkpoint genes in both risk groups were also analyzed. Most of the immune checkpoints were expressed higher in the $risk_{high}$ group (Figure 9F). For instance, *PD-1* (*PDCD1*) expression was higher in $risk_{high}$ patients with KIRC compared with $risk_{low}$ patients, and the expression of *PD1* increased as the risk score

increased (Figure 9E), suggesting that *PD1* inhibitors may be more effective in high-risk score KIRC patients. In contrast, another common immunosuppressant *PDL1* (*CD274*) was higher in $risk_{low}$ KIRC, meaning that $risk_{low}$ KIRC patients might be more sensitive to treatment with *PDL1* inhibitors. Subsequently, we also compared the TIDE scores of the two risk groups (Figure 9C). The $risk_{low}$ KIRC patients had higher TIDE scores and higher immune escape potential, suggesting that $risk_{low}$ groups that received most immune checkpoint blockers had poor treatment effects. In general, $risk_{high}$ KIRC patients had a more active immune state and were more likely to benefit from immunotherapy.

Tumor mutation burden and prediction of potential drug sensitivity

In accordance with the analysis result of somatic mutation data of KIRC patients in TCGA database, as shown in the waterfall diagram, 39.64% of $risk_{high}$ KIRC patients had genetic mutation, while 38.68% of $risk_{low}$ patients had such mutation (Figure 10A,B). TMB refers to the number of somatic non-synonymous mutations in the genomic region, which can indirectly reflect the ability and degree of neoantigen production of tumors and predict the efficacy of immunotherapy for a variety of tumors. TMB can be affected by many factors. KIRC patients with different clinical and biological characteristics have different TMB. Although the degree of TMB was higher in the low-risk group, this was not statistically significant. This may be due to the insufficient number of samples of KIRC patients who participated in the analysis of tumor mutational burden. Overall, these analyses provided a basis for risk signatures to predict the prognosis of KIRC patients and the effect of immunotherapy. (Figure 10C). In KIRC patients of all risk groups, the gene with the highest mutation frequency is *TNN*, and the most common mutation type is missense mutation. In accordance with the results of the K-M analysis, the OS rate of KIRC patients with high TMB is lower than that of KIRC patients with low TMB (Figure 10D). Next, the TMB and risk score of KIRC patients for survival analysis were integrated, and a conclusion was drawn that patients with high TMB and the high-risk score achieved the worst prognosis (Figure 10E), thus confirming the ability of our risk signature to predict the OS of KIRC patients. Although the difference in TMB score among different risk groups was not significant, it also revealed some potential mechanisms that might affect the clinical outcome of patients with KIRC. By comparing the IC50 values of some common drugs in different risk groups, we found that the risk score of KIRC patients could influence their sensitivity to drugs to a certain extent. As revealed by the results, the IC50 value of Rapamycin, Sunitinib, Bleomycin, AKT inhibitor VII, Ruxolitinib, 5-Fluorouracil, Saracatinibin $risk_{high}$ groups was higher than that in $risk_{low}$ groups, thus suggesting that $risk_{low}$ groups may be more sensitive to the above drug treatment. (Figure 10F–L). The prediction of the efficacy of the above potential drugs for KIRC patients is beneficial to guiding the clinical drug treatment of KIRC patients. The above results also suggest that our risk signature in this study takes on certain significance for guiding drug therapy in patients with KIRC.

Further studies on four LncRNAs associated with cuproptosis that constitute risk signature

Existing researches have suggested that the risk score of predictive signature is significantly associated with KIRC tumor immune microenvironment and TMB. Therefore, we further explore the biological characteristics of the 4 risk LncRNAs that make up the predictive signature from the perspective of tumor immunity and tumor mutation. As depicted in Figure 11A, the 4 risk LncRNAs were significantly associated with most genes associated with cuproptosis. Next, we also analyzed the correlation of common immune checkpoint genes and 4 risk LncRNAs (Figure 11B). It could be seen that common immune checkpoint PD1 and PDL1 are not only related to risk scores, but also significantly related to 4 risk LncRNAs. The correlation scatters plot showed the correlation between risk LncRNAs and TMB (Figure 11C,D). Based on the results of the ESTIMATE algorithm, it was found that the expression of 4 risk LncRNAs was potentially related to the microenvironment of tumor of KIRC (Figure 11E–P). Existing researches have suggested that *MINCR* and *APCDD1L-DT*, are significantly associated with the occurrence and growth of various tumors, such that the focus of this study was placed on the biological characteristics of two risk LncRNAs in KIRC. First, the K-M survival analysis curve of 530 KIRC patients indicated that the expression of the above two risk LncRNAs was significantly related to the OS rate of KIRC patients (Figure 11Q,R). Subsequently, the relative expression of LncRNA of *MINCR* and *APCDD1L-DT* was detected through qRT-PCR in HK-2, 786-O, and 769-P cells. The results indicated that the relative expression level of *MINCR* and *APCDD1L-DT* was higher in 786-O and 769-P tumor cells than that in HK-2 cells, which were normal renal tubular epithelial cells (Figure 11S,T). In general, the experimental results of this study indicated the accuracy of our risk signature to a certain extent.

Discussion

KIRC is one of the most common malignant tumors of the urinary system, which has the characteristics of high recurrence rate, high risk of metastasis, and poor prognosis, especially for advanced KIRC patients, whose prognosis is often very unsatisfactory (Umberto et al., 2019). Considerable previous

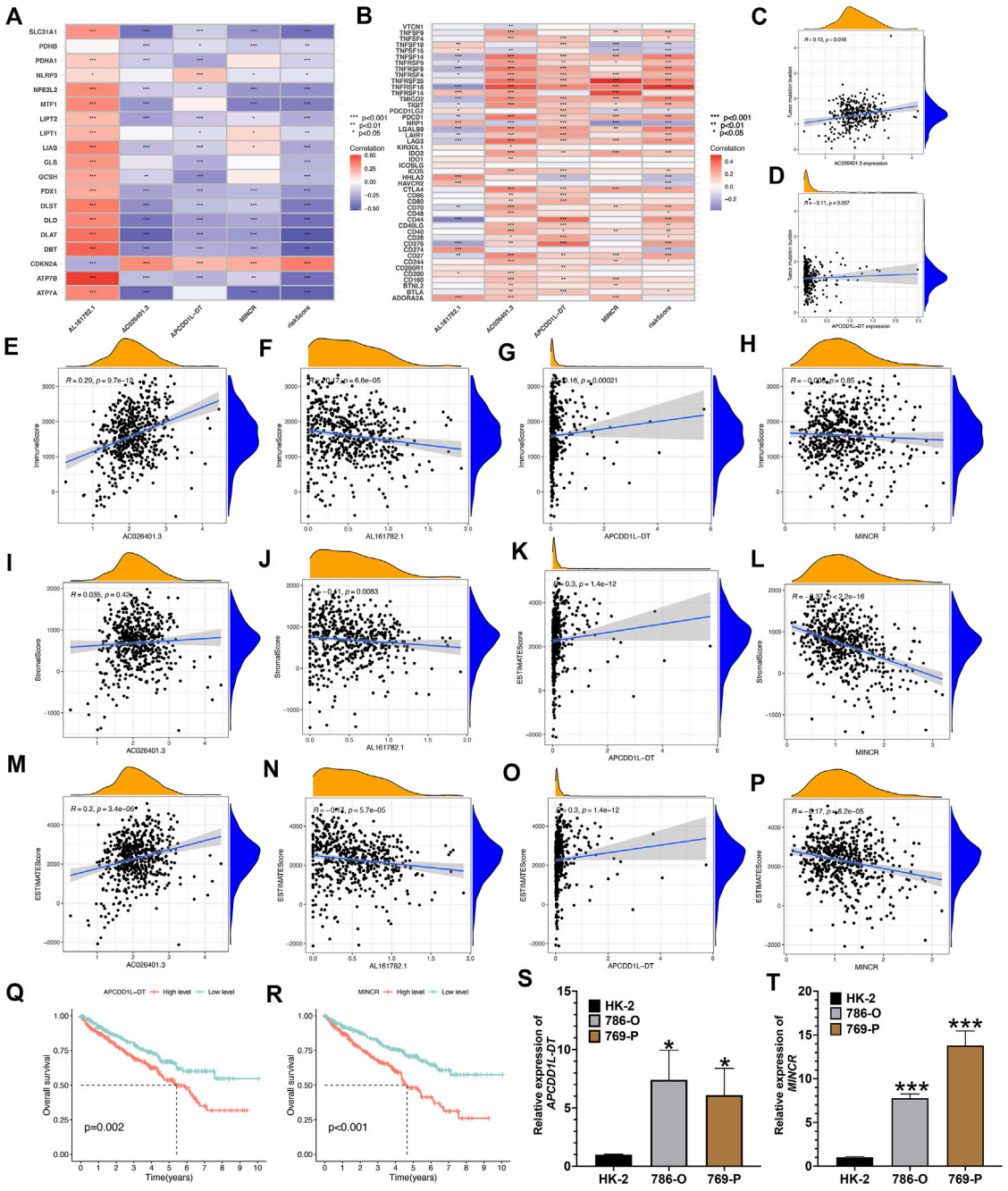


FIGURE 11
Comprehensive analysis of 4 risk cuproptosis-related lncRNAs. **(A)** Correlation between 4 risk LncRNAs and 19 cuproptosis-related genes. **(B)** Correlation between 4 risk LncRNAs and common immune-checkpoints. The scatter plot represents the correlation between the tumor mutation burden and the expression of AC026401.3 **(C)** and APCDD1L-DT **(D)**. **(E–P)** The scatter plots show the correlations between StromalScore, ImmuneScore, ESTIMATEScore, and the expression of the 4 risk lncRNAs. K-M survival curves analysis of APCDD1L-DT **(Q)** and MINCR **(R)**. Relative expression of APCDD1L-DT **(S)** and MINCR **(T)** in normal renal tubular epithelial cell line HK-2 and renal clear cell carcinoma cell line 786-O and 769-P.*$p < 0.05$, **$p < 0.01$,***$p < 0.001$.

research over the past few years suggested that LncRNAs take critical significance in regulating the occurrence and growth of KIRC. At present, many prognosis prediction models based on LncRNA have been applied to KIRC (Xinfang et al., 2021). Xing's research developed s risk assessment model by ferroptosis-related LncRNAs could accurately achieve the prediction of the

prognosis of KIRC patients (Xiao-Liang et al., 2021). Tang's team defined a signature of ferroptosis-related LncRNAs that could improve the prognosis prediction in papillary renal cell carcinoma (Xinfang et al., 2022). Yu et al. found that the prognosis model based on m6A-related LncRNAs can provide powerful help for the prognosis evaluation of KIRC patients (JunJie et al., 2021). Professor Sun's team pointed out that immunization-associated LncRNAs can not only help achieve the prediction of the prognosis of KIRC patients but also become a potential immunotherapy target for KIRC patients (Zhuolun et al., 2020). Cuproptosis—as a newly discovered regulation mode of cell death, some studies on cuproptosis in KIRC have also been reported (Ganghua et al., 2022). For example, FDX1, a key gene of cuproptosis, which had been proved to have a certain effect on the proliferation of KIRC cells, in addition, the presence of cuproptosis was demonstrated in KIRC cells (Ganghua et al., 2022). Cuproptosis could bring more options for immunotherapy and targeted therapy for KIRC patients. However, the current research on CRLs in KIRC is blank, and there is no KIRC prognosis prediction model based on CRLs. In this study, we first applied CRLs to KIRC. Cox and LASSO regression analyses were carried out for constructing and verifying the CRLs prognosis signature, which had good predictive performance. The potential correlation between prognosis signature and microenvironment of KIRC, immune checkpoint, and TMB were also systematically studied. The above results are expected to help the clinical diagnosis and treatment of KIRC.

We downloaded 539 KIRC tissue specimens and 72 normal kidney tissue specimens and their corresponding clinical information from the TCGA database. Genes associated with cuproptosis were obtained from previous literature studies. The CRLs were determined by spearman correlation analysis. Then, based on Cox and LASOO regression analysis, a prognostic risk signature composed of 4 risk CRLs was built. The result suggests that the OS、PFS rate of KIRC patients with high-risk scores is less than that of KIRC patients with low-risk scores. The ROC curve showed that the AUC values of 1, 3, and 5 years in the training group were 0.828, 0.780, and 0.794, respectively, suggesting that the prognosis signature built by CRLs had high accuracy and reliability. At the same time, we also consulted other published LncRNA prognostic models. By comparing the AUC values of 1-year risk models, we found that our prognostic model has better predictive performance than other published models. Furthermore, the clinical factors and risk score were integrated to build a nomogram, thus making the developed prognosis signature more applicable to clinical trials.

As the enrichment results of KEGG, GO and GSEA all show that many enrichment pathways are showed significantly associated with immune activities, we discussed the correlation between risk score and KIRC tumor immune microenvironment. The immune score and

TABLE 2 Multivariate Cox regression analysis of 4 risk lncRNAs.

Risk LncRNAs	Coeficient	HR	HR.95L	HR.95H
AL161782.1	−1.1512239	0.31624946	0.13522153	0.73962866
AC026401.3	0.47111037	1.60177177	0.99208715	2.58613652
APCDD1L-DT	0.6788922	1.97169228	1.35943561	2.85969446
MINCR	0.46866756	1.59786372	1.05959843	2.40956233

ESTIMATE score of the $risk_{high}$ KIRC patients were higher than those of the low-risk. The change of tumor immune microenvironment could promote the proliferation, migration, and invasion of KIRC, which explains to a certain extent that there were significant survival differences among KIRC patients in different risk groups. In addition, based on various algorithms, we analyzed the differences in immune activities of KIRC patients among different risk groups. It could be seen that most immunocytes and immunization-associated functions have higher enrichment scores in $risk_{high}$ KIRC patients, and $risk_{high}$ KIRC patients have more frequent immune activities. In recent years, immunotherapy has been widely used in patients with KIRC, especially in patients with advanced KIRC. For instance, PD1 and PDL1 inhibitors can improve the survival rate of some patients with advanced KIRC to a cerntai extent, but unfortunately, not all patients with KIRC can benefit from them. Accordingly, we try to provide individualized immunotherapy for KIRC patients based on our risk signature. We compared the expression of common immune checkpoint genes in different risk groups and suggested that most immune checkpoints have higher expression in $risk_{high}$ groups, such as PD1, CD27, and CD40, which indicates that $risk_{high}$ KIRC patients may be more sensitive to the treatment of the above immune checkpoint inhibitors, and can get better therapeutic effects from them. In addition, compared with the low-risk KIRC patients, the TIDE score of the $risk_{high}$ KIRC patients was lower, indicating that the $risk_{high}$ KIRC patients have lower immune escape potential. It also showed that $risk_{high}$ KIRC patients can benefit from ICB treatment. The above immune analysis results revealed that $risk_{high}$ KIRC patients might gain more benefits from immunotherapy. Lastly, we predicted the sensitivity of some potential therapeutic drugs based on the risk signature, which could help the clinical drug treatment of KIRC patients. For example, Rapamycin, Sunitinib, Bleomycin, AKT inhibitor VII, Ruxolitinib, 5-Fluorouracil, Saracatinibin. According to the latest clinical guidelines, targeted therapies for renal cancer are mainly divided into two categories: tyrosine kinase inhibitors (TKI) and m-TOR inhibitors. TKI drugs are important means to treat metastatic renal cancer, and currently, commonly used drugs include sunitinib, ruxolitinib,

sorafenib, and axitinib. m-TOR inhibitors, which target m-TOR and related signaling pathways, can control the proliferation and angiogenesis of tumor cells, so as to control the tumor. Commonly used drugs include rapamycin and everolimus, which are mainly used as second-line drugs for advanced renal cancer patients who have failed TKI treatment. Our study compared the IC50 values of these drugs in different risk populations to predict the sensitivity of different risk populations to these targeted therapies, so as to guide clinical application.

Most research has suggested that LncRNA takes on critical significance in developing some common malignant tumors over the past few years, so the role played by four risk LncRNAs making up the prognosis signature was investigated in depth. There has been little research about *LncRNA AL161782.1*. Several research has highlighted that *LncRNA AC026401.3* plays a role in building the model for the prediction of the prognosis of hepatocellular carcinoma and renal carcinoma. *LncRNA AC026401.3* takes on significance in the prognosis of patients with hepatocellular carcinoma and renal carcinoma to a certain extent, and *LncRNA APCDD1L-DT* can serve as a marker of the prognosis of lung squamous cell carcinoma (Honghao et al., 2021; Rongjiong et al., 2021; Min et al., 2022). A considerable number of studies suggested that the expression of *LncRNA MINCR* is capable of affecting a wide variety of malignant tumors' development and occurrence. For instance, Li's experimental research proves that *LncRNA MINCR* can regulate the *miR-876-5p*/*GSPT1* axis to aggravate the progression of glioma (Zheng et al., 2020). Yu et al.' s research shows that *LncRNA MINCR* has high expression in colon cancer tissues and cells, and promotes the proliferation and migration of colon cancer by regulating *miR-708-5p* (Yang et al., 2020). However, the research on the above four risk LncRNAs in KIRC is still very blank, especially *LncRNA MINCR*, many studies have shown that it can play a role as an oncogene in the development of various tumors. The qRT-PCR experiment also demonstrated that *MINCR* and *APCDD1L-DT* have high expression in KIRC cells. Furthermore, this study also shows that the expression of the above-risk LncRNAs is significantly associated with immune activities. The four risk LncRNAs, especially *MINCR*, have the potential to become a new target of KIRC immunotherapy. We also expect that the further research results of the above LncRNAs can guide the clinical diagnosis and treatment of KIRC.

In general, this study has certain clinical values and limitations. First of all, we studied the genes associated with cuproptosis for the first time and built a prognosis model based on CRLs, which can accurately achieve the prediction of the prognosis of KIRC patients, and its predictive performance is better than some published models for the prediction of the prognosis of KIRC patients. Second, we also systematically analyzed the correlation between the CRLs prognosis signature and tumor immune microenvironment, which provided a new idea for guiding the immunotherapy of KIRC. Thirdly, we also showed a novel direction for the drug treatment of KIRC. Lastly, we proposed that *LncRNA MINCR* has great potential to become a new target of KIRC immunotherapy. However, this study also has some limitations. First, our signature was only verified internally, and the expression of risk LncRNA was simply verified by qRT-PCR. No suitable external dataset was identified in the published database to further evaluate the reliability of our signature. Second, we lack clinical follow-up data to prove the value of our prognostic model. Lastly, in-depth *in vivo* and *in vivo* experiments should be performed to verify the conclusion of this study, especially to verify the role of *MINRC i*n the development of KIRC. Although there are several defects, the CRLs prognosis signature can accurately achieve the prediction of the prognosis of KIRC patients, which is initially found in this study. Therefore, this study has a great application prospect in clinical practice.

Conclusion

The team of this study built a risk signature based on cuproptosis-related LncRNAs, which can precisely achieve the prediction of KIRC patient's prognosis. In accordance with the risk signature, we evaluated the role played by cuproptosis-related LncRNAs in the immune microenvironment of KIRC tumors and the possible potential regulatory mechanism, which is conducive to guiding the individualized treatment of KIRC. In addition, four identified risk lncRNA (especially *MINCR*) can be novel targets for immunotherapy of KIRC patients (Table 2).

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ Supplementary Material.

Author contributions

PH designed the study and wrote the manuscript. WH, ST, HD, and DH performed the analysis. QF, YW, and HZ collected the dataset. ZL and MT reviewed and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgments

Thanks to all the authors for their help with this article, as well as the open source data provided by TCGA Database.

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/fgene.2022.1009555/full#supplementary-material

References

Andreas, S., and Vaux, D L. (2020). Cell death in the origin and treatment of cancer. *Mol. Cell* 78 (6), 1045–1054. doi:10.1016/j.molcel.2020.05.014

Anna, C., Litwin, T., Dusek, P., Ferenci, P., Lutsenko, S., Medici, V., et al. (2018). Wilson disease. *Nat. Rev. Dis. Prim.* 4 (1), 21. doi:10.1038/s41572-018-0018-3

Binbin, C., Khodadoust, M. S., Liu, C. L., Newman, A. M., and Alizadeh, A. A. (2018). Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* 1711, 243–259. doi:10.1007/978-1-4939-7493-1_12

Cancer Genome Atlas Research Network (2013). Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* 499 (7456), 43–49. doi:10.1038/nature12222

Chen, C., and Zheng, H. (2021). LncRNA LINC00944 promotes tumorigenesis but suppresses akt phosphorylation in renal cell carcinoma. *Front. Mol. Biosci.* 8, 697962. doi:10.3389/fmolb.2021.697962

Cobine, P. A., and Brady, D. C. (2022). Cuproptosis: Cellular and molecular mechanisms underlying copper-induced cell death. *Mol. Cell* 82, 1786–1787. doi:10.1016/j.molcel.2022.05.001

Daniel, F., and Vince, J. E. (2019). Pyroptosis versus necroptosis: Similarities, differences, and crosstalk. *Cell Death Differ.* 26 (1), 99–114. doi:10.1038/s41418-018-0212-6

Daolin, T., Xin, C., and Guido, K. (2022). Cuproptosis: A copper-triggered modality of mitochondrial cell death. *Cell Res.* 32 (5), 417–418. doi:10.1038/s41422-022-00653-7

Francesco, T., and Serena, S. (2017). Mechanisms of charge transfer in human copper ATPases ATP7A and ATP7B. *IUBMB life* 69 (4), 218–225. doi:10.1002/iub.1603

Ganghua, Z., Chen, X., Fang, J., Tai, P., Chen, A., and Cao, K. (2022). Cuproptosis status affects treatment options about immunotherapy and targeted therapy for patients with kidney renal clear cell carcinoma. *Front. Immunol.* 13, 954440. doi:10.3389/fimmu.2022.954440

Giuseppe, R., Larcher, A., Montorsi, F., and Capitanio, U. (2021). Renal cancer: Overdiagnosis and overtreatment. *World J. Urol.* 39, 2821–2823. doi:10.1007/s00345-021-03798-z

Honghao, C., Tong, H., Zhu, J., Xie, C., Qin, Z., Li, T., et al. (2021). A glycolysis-based long non-coding RNA signature accurately predicts prognosis in renal carcinoma patients. *Front. Genet.* 12, 638980. doi:10.3389/fgene.2021.638980

Hotchkiss, R. S., Strasser, A., McDunn, J. E., and Swanson, P. E. (2009). Cell death. *N. Engl. J. Med.* 361 (16), 1570–1583. doi:10.1056/NEJMra0901217

Isabel, V., and Grünwald, V. (2021). Current treatment options for recurrent/metastatic renal cell carcinoma (mRCC). *Aktuelle Urol.* 52 (5), 439–444. doi:10.1055/a-1480-3653

Jianjian, D., Wang, X., Xu, C., Gao, M., Wang, S., Zhang, J., et al. (2021). Inhibiting NLRP3 inflammasome activation prevents copper-induced neuropathology in a murine model of Wilson's disease. *Cell Death Dis.* 12 (1), 87. doi:10.1038/s41419-021-03397-1

Jie, M., and Chunyang, W. (2022). N7-Methylguanosine-Related lncRNAs: Integrated analysis associated with prognosis and progression in clear cell renal cell carcinoma. *Front. Genet.* 13, 871899. doi:10.3389/fgene.2022.871899

JunJie, Y., Mao, W., Sun, S., Hu, Q., Wang, C., Xu, Z., et al. (2021). Identification of an m6A-related lncRNA signature for predicting the prognosis in patients with kidney renal clear cell carcinoma. *Front. Oncol.* 11, 663263. doi:10.3389/fonc.2021.663263

Kosuke, Y., Shahmoradgoli, M., Martinez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612

Mark, S. D. (2019). Cell death: A review of the major forms of apoptosis, necrosis and autophagy. *Cell Biol. Int.* 43 (6), 582–592. doi:10.1002/cbin.11137

Marston, L. w., and J, R. c. (2019). The cancer Genome Atlas of renal cell carcinoma: Findings and clinical implications. *Nat. Rev. Urol.* 16 (9), 539–552. doi:10.1038/s41585-019-0211-5

Marston, L. w., Schmidt, L. S., Crooks, D. R., Wei, D., Srinivasan, R., Lang, M., et al. (2019). The metabolic basis of kidney cancer. *Cancer Discov.* 9 (8), 1006–1021. doi:10.1158/2159-8290.CD-18-1354

Min, C., Zhang, J., Cao, P. B., and Zhou, G. Q. (2022). Prognostic and predictive value of the hypoxia-associated long non-coding RNA signature in hepatocellular carcinoma. *Yi chuan = Hered.* 44 (2), 153–167. doi:10.16288/j.yczz.21-416

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Ok, S. m., Mattie, M. D., Lee, C. H., and Freedman, J. H. (2014). The role of Nrf1 and Nrf2 in the regulation of copper-responsive transcription. *Exp. Cell Res.* 322 (1), 39–50. doi:10.1016/j.yexcr.2014.01.013

Paul, G., Nancy, C. J., and Stephanie, H. R. (2014). Clinical drug response can be predicted using baseline gene expression levels and *in vitro* drug sensitivity in cell lines. *Genome Biol.* 15, 47. doi:10.1186/gb-2014-15-3-r47

Pei, T., Wei, J. X., Li, J., Ren, J. K., and Yang, J. J. (2021). LncRNA SNHG1 regulates immune escape of renal cell carcinoma by targeting miR-129-3p to activate STAT3 and PD-L1. *Cell Biol. Int.* 45 (7), 1546–1560. doi:10.1002/cbin.11595

Peng, J., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24 (10), 1550–1558. doi:10.1038/s41591-018-0136-1

Peter, T., Coy, S., Petrova, B., Dreishpoon, M., Verma, A., Abdusamad, M., et al. (2022). Copper induces cell death by targeting lipoylated TCA cycle proteins. *Sci. (New York, N.Y.)* 375 (6586), 1254–1261. doi:10.1126/science.abf0529

Rongjiong, Z., Zheng, M., Wang, M., Lu, F., and Hu, M. (2021). Identification of a prognostic long noncoding RNA signature in lung squamous cell carcinoma: A population-based study with a mean follow-up of 3.5 years. *Archives public health = Archives belges de sante publique* 79 (1), 61. doi:10.1186/s13690-021-00588-2

Seema, K., Kumar, P., Pandi, S. N., Malik, R., Dhanasekaran, S. M., and Shukla, S. K. (2019). Immune associated LncRNAs identify novel prognostic subtypes of renal clear cell carcinoma. *Mol. Carcinog.* 58 (4), 544–553. doi:10.1002/mc.22949

Sheng, I. Y., and Rini, B. I. (2019). Immunotherapy for renal cell carcinoma. *Expert Opin. Biol. Ther.* 19 (9), 897–905. doi:10.1080/14712598.2019.1628946

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2022). Cancer statistics. *Ca. A Cancer J. Clin.* 72 (1), 7–33. doi:10.3322/caac.21708

Steven, C., Uzzo, R. G., Allaf, M. E., Bass, E. B., Cadeddu, J. A., Chang, A., et al. (2017). Renal mass and localized renal cancer: AUA guideline. *J. Urol.* 198 (3), 520–529. doi:10.1016/j.juro.2017.04.100

Stockwell, B. R. (2022). Ferroptosis turns 10: Emerging mechanisms, physiological functions, and therapeutic applications. *Cell* 185 (14), 2401–2421. doi:10.1016/j.cell.2022.06.003

Umberto, C., Bensalah, K., Bex, A., Boorjian, S. A., Bray, F., Coleman, J., et al. (2019). Epidemiology of renal cell carcinoma. *Eur. Urol.* 75 (1), 74–84. doi:10.1016/j.eururo.2018.08.036

Umberto, C., and Francesco, M. (2016). Renal cancer. *Lancet* 387 (10021), 894–906. doi:10.1016/s0140-6736(15)00046-x

Wang, S., Yang, X., Xie, W., Fu, S., Chen, Q., Li, Z., et al. (2021). LncRNA GAPLINC promotes renal cell cancer tumorigenesis by targeting the miR-135b-5p/CSF1 Axis. *Front. Oncol.* 11, 718532. doi:10.3389/fonc.2021.718532

Wenzhong, Z., Zhang, S., Guo, H., Chen, X., Huang, Z., Jiang, S., et al. (2021). Multi-omics analysis of tumor angiogenesis characteristics and potential epigenetic regulation mechanisms in renal clear cell carcinoma. *Cell Commun. Signal.* 19 (1), 39. doi:10.1186/s12964-021-00728-9

Xiao-Liang, X., Yao, Z. Y., Ou, J., Xing, C., and Li, F. (2021). Development and validation of ferroptosis-related lncRNAs prognosis signatures in kidney renal clear cell carcinoma. *Cancer Cell Int.* 21 (1), 591. doi:10.1186/s12935-021-02284-1

Xinfang, T., Zhang, A., Feng, Y., Su, Y., Wang, X., Jiang, F., et al. (2021). A novel pyroptosis-related lncRNAs signature for predicting the prognosis of kidney renal clear cell carcinoma and its associations with immunity. *J. Oncol.* 2021, 1–15. doi:10.1155/2021/9997185

Xinfang, T., Zhang, J., Jiang, Q., and Qiu, J. (2022). Construction of prognostic signature of patients with oral squamous cell carcinoma based on pyroptosis-related long non-coding RNAs. *Front. Surg.* 9, 741726. doi:10.3389/fsurg.2022.935765

Yang, Y., Chang, Z., Han, C., Zhuang, L., Zhou, C., Qi, X., et al. (2020). Long non-coding RNA MINCR aggravates colon cancer via regulating miR-708-5p-mediated Wnt/β-catenin pathway. *Biomed. Pharmacother. = Biomedecine Pharmacother.* 129, 110292. doi:10.1016/j.biopha.2020.110292

Yingkun, X., Li, H., Lan, A., Wu, Q., Tang, Z., Shu, D., et al. (2022). Cuprotosis-related genes: Predicting prognosis and immunotherapy sensitivity in pancreatic cancer patients. *J. Oncol.* 2022, 2363043. doi:10.1155/2022/2363043

Zachary, K., Sayyid, R. K., and Wallis, C. J. D. (2019). Lessons learned from the global epidemiology of kidney cancer: A refresher in epidemiology 101. *Eur. Urol.* 75, 85–87. doi:10.1016/j.eururo.2018.09.035

Ze, Y., Zhou, R., Zhao, Y., Pan, Y., Liang, H., Zhang, J. S., et al. (2019). Blockage of SLC31A1-dependent copper absorption increases pancreatic cancer cell autophagy to resist cell death. *Cell Prolif.* 52 (2), 12568. doi:10.1111/cpr.12568

Zheng, L., Xie, X., Fan, X., and Li, X. (2020). Long non-coding RNA MINCR regulates miR-876-5p/GSPT1 Axis to aggravate glioma progression. *Neurochem. Res.* 45 (7), 1690–1699. doi:10.1007/s11064-020-03029-8

Zhuolun, S., Jing, C., Xiao, C., and Li, T. (2020). Long non-coding RNA profile study identifies an immune-related lncRNA prognostic signature for kidney renal clear cell carcinoma. *Front. Oncol.* 10, 1430. doi:10.3389/fonc.2020.01430

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 11 November 2022
DOI 10.3389/fgene.2022.1001901

Check for updates

OPEN ACCESS

EDITED BY
Ehsan Nazemalhosseini-Mojarad,
Shahid Beheshti University of Medical Sciences, Iran

REVIEWED BY
Wen Zhou,
Case Western Reserve University, United States
Dingli Song,
First Affiliated Hospital of Xi'an Jiaotong University, China
Tevfiktolga Sahin,
İnönü University, Turkey

*CORRESPONDENCE
Dejun Zhang,
djzhang@qhu.edu.cn

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 24 July 2022
ACCEPTED 18 October 2022
PUBLISHED 11 November 2022

CITATION
Zhao L, Guo Y, Liu Z, Ma J, Peng Y and Zhang D (2022), Characterization of glycosylation regulator-mediated glycosylation modification patterns and tumor microenvironment infiltration in hepatocellular carcinoma.
Front. Genet. 13:1001901.
doi: 10.3389/fgene.2022.1001901

COPYRIGHT
© 2022 Zhao, Guo, Liu, Ma, Peng and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Characterization of glycosylation regulator-mediated glycosylation modification patterns and tumor microenvironment infiltration in hepatocellular carcinoma

Linlin Zhao[1,2,3,4,5], Yang Guo[1,3,4,5], Zhanfeng Liu[2], Jing Ma[5], Yanfeng Peng[5] and Dejun Zhang[1,5]*

[1]Research Center for High Altitude Medicine, Medical College of Qinghai University, Xining, China, [2]Department of General Surgery, The First People's Hospital Xining City, Xining, China, [3]Key Laboratory of Application and Foundation for High-Altitude Medicine Research in Qinghai Province, Xining, China, [4]Qinghai-Utah Joint Research Key Laboratory for High Altitude Medicine, Xining, China, [5]College of Eco-Environmental Engineering, Qinghai University, Xining, China

Background: Previous studies have shown that glycosylation of proteins ofen plays an important role in HCC. However, the potential mechanism of glycosylation in HCC has not been described systematically.

Methods: We comprehensively evaluated the glycosylation patterns in HCC samples based on 43 glycosylation regulators, and annotated the modification patterns with the enrichment of immune cells and stromal cells. Considering the heterogeneity of HCC patients, the glycosylation score was constructed using single-sample gene set enrichment analysis (ssGSEA). We also explored the drugs that different HCC patients were sensitive to based on glycosylation mode and score.

Results: We identified three glycosylation-regulated gene subtypes. By annotating the subtypes, it was found that the glycosylation regulated gene subtypes was highly matched with three immunophenotypes of HCC (immune-inflamed, immune-excluded, and immune-desert), regardless of the characteristics of immune cell infiltration or prognosis. Based on the characteristic genes of glycosylation-regulated gene subtypes, we constructed a glycosylation-related model, and found that glycosylation-related model was highly consistent with the glycosylation regulated gene subtypes. The glycosylation score that evaluates the glycosylation characteristics of a single HCC sample has high prognostic value, and the prognosis of patients with high glycosylation score is significantly worse. Interestingly, we found that the glycosylation score was closely related to

Abbreviations: HCC, Hepatocellular carcinoma; TCGA, the Cancer Genome Atlas; CNV, copy number variation; GEO, Gene Expression Omnibus; PCA, Principal component analysis; GSVA, Gene Set Variation Analysis; GDSC, Genomics of Drug Sensitivity in Cancer.

tumor node metastasis (TNM) staging. By applying glycosylation-regulated gene subtypes and glycosylation score to explore the sensitivity of different patients to anticancer drugs, it was found that the sensitivity of Thapsigargin, Shikonin, Embelin and Epothilone. B was closely related to the glycosylation mode.

Conclusion: This study reveals that the diversity of glycosylation patterns plays an important role in HCC. Therefore, evaluating the glycosylation patterns of patients with HCC will be helpful in identifying the characteristics of immune cell infiltration and selecting accurate treatment methods.

KEYWORDS

hepatocellular carcinoma, glycosylation, prognosis, immunocyte infiltration, sensitive drugs

1 Introduction

Post-translational modification of proteins is the main factor that affects the functional diversity of the proteome. At present, post-translational modification of protein mainly includes phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation, lipidization, and proteolysis (Vu et al., 2018). Glycosylation, an important post-translational modification of proteins, plays an important role in various diseases, such as tumors, viral infections, and Alzheimer's disease, mainly by affecting the stability of protein structures (Park, 2019; Tao and Huang, 2019; Liu et al., 2021). Abnormal glycosylation can affect the progression of a variety of tumors and usually involves the regulation of epithelial mesenchymal transformation and the immune microenvironment (Wattanavises et al., 2019; Zhang et al., 2020a). Therefore, an in-depth exploration of the role of these regulatory factors in hepatocellular carcinoma (HCC) will help reveal their mechanism of action.

Glycosylation of different proteins play different roles in HCC. For example, Hitomi et al. showed that N-glycosylation of fuc increases with the progression of HCC and can be used as an independent prognostic factor for HCC (Asazawa et al., 2015). A recent study on HCC showed that N-glycosylation of Ces1 can inhibit the proliferation of HCC cells (Na et al., 2020). Glycosylation of LAMP2 and SV40 serves as a protective mechanism against HCC (Pousset et al., 1997; Chiu et al., 2022). Glycosylation of HCC mostly occurs in oncogenes such as *AFP, AACT,* and *MDR*, which can be used as risk factors for HCC (Ledoux et al., 2003; Lee et al., 2022). Glycosylation of these genes is not only involved in the molecular mechanism of HCC progression but is also closely related to tumor drug resistance. In HCC, mRNA, miRNAs, and lncRNAs can participate in the regulation of glycosylation (Liu et al., 2017; Liu et al., 2018). Glycosylation of PD-L1 may lead to immunosuppression or inactivation (Huang et al., 2019). However, the role of glycosylation in the tumor microenvironment of HCC remains unclear.

Owing to experimental limitations, research in the field has been limited to one or two glycosylation regulators, even though the mechanism of glycosylation regulation in HCC is regulated by the interaction of many other glycosylation regulators in a highly coordinated manner. Therefore, a systematic understanding of the role of multiple glycosylation regulators is essential for studying glycosylation regulatory models in HCC. In this study, we integrated the genomic information of HCC samples in 'The Cancer Genome Atlas' (TCGA) and Gene Expression Omnibus (GEO) databases to comprehensively evaluate the glycosylation regulatory model and defined different subtypes using the enrichment of immune cells and stromal cells. We identified three glycosylation subtypes in this study and surprisingly found that the characteristics of these three subtypes were highly consistent with immune rejection, immune inflammation, and immune desert phenotypes. In addition, we constructed a glycosylation score for HCC to predict the prognosis and drug sensitivity of different patients.

2 Materials and methods

2.1 Hepatocellular carcinoma data set and preprocessing

HCC datasets with gene expression data and complete clinical information were retrieved from TCGA and GEO databases. Finally, TCGA-LIHC, GSE76427, and GSE14322 were included in subsequent analysis. By deleting samples with missing clinical information in the dataset, TCGA-LIHC included 281 HCC samples. GSE76427 included microarray data of 115 HCC cases with complete clinical data. GSE14322 included 167 samples, including 115 cases of HCC and 52 cases of normal HCC. Gene annotation and data standardization were performed to obtain a standard gene expression matrix. Data on somatic mutations and copy number variation (CNV) of HCC were obtained from TCGA. The list of glycosylation regulators were summarized from

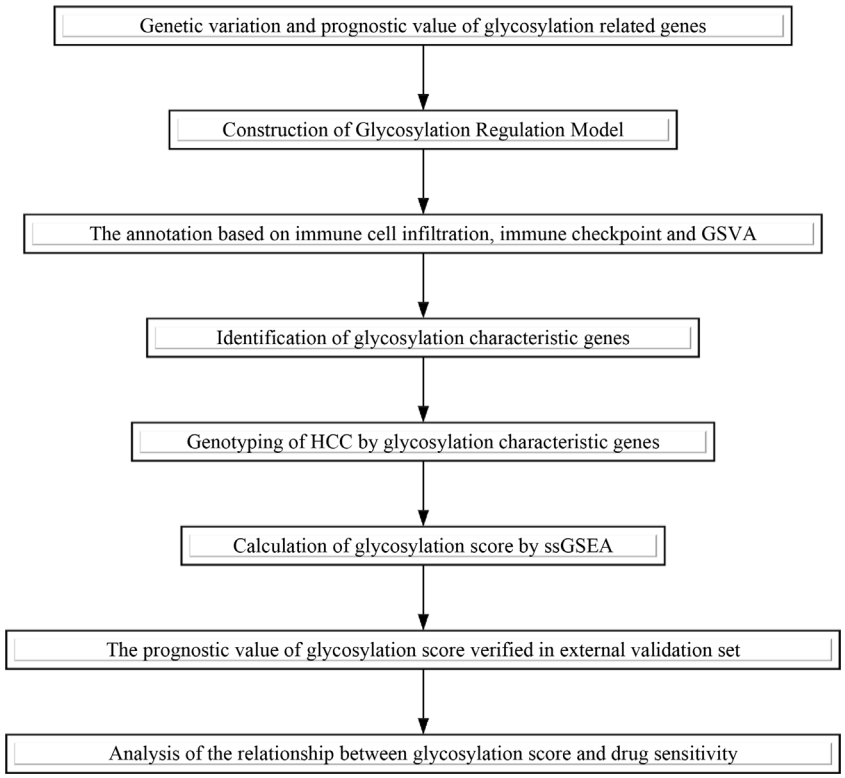


FIGURE 1
Flow-process diagram.

previous literature and the molecular signature database V7.4 (Liberzon et al., 2015). The specific analysis process of this study is shown in Figure 1.

2.2 Unsupervised hierarchical clustering analysis of glycosylation regulators

Univariate Cox regression analysis was applied to screen for prognosis-related glycosylation regulatory genes, with a screening criteria of $p < 0.05$. The limma package was applied based on the GSE14322 cohort analysis of glycosylation regulatory genes differentially expressed in HCC and adjacent tissues. The intersection of the two was used to identify differentially expressed glycosylation regulatory genes with prognostic value for subsequent unsupervised cluster analysis. Before unsupervised clustering, the conditional survival curve was constructed using the surviviner and survival packages to check the availability of the data. Consensus clustering was applied to build a glycosylation regulatory gene model, and its clustering number and stability were determined by a consensus clustering algorithm (Timmerman et al., 2013). This process was performed on the ConsensusClusterPlus package of R version 4.1.2, which was repeated 1,000 times to ensure the stability of the model (Wilkerson and Hayes, 2010).

2.3 Verification of glycosylation regulated gene subtypes

To verify the prognostic value of different glycosylation regulatory genes, survival analysis of patients with different subtypes was performed using the surviviner and survival packages to compare the survival of different subtypes. Nomograms were constructed for clinical application based on age, sex, tumor node metastasis (TNM) stage, race, and glycosylation-regulated gene subtypes. A calibration curve was constructed to test the accuracy of the nomogram. This process was performed on the nomogramEx package.

2.4 Principal component analysis and calculation of glycosylation score

To further verify the subtypes of glycosylation regulatory genes, common differentially expressed genes among the

different subtypes were analyzed using the limma package as glycosylation-related genes. Univariate Cox regression was used to screen for glycosylation-related genes with prognostic value. The advantage of PCA is that it can cluster genes with the largest contribution, whereas genes with smaller contributions are not included in the clustering process. In this study, PCA was used to conduct a secondary clustering of patients with HCC. Surviviner and survival packages were used to explore the survival of patients with different subtypes. Then, single-sample gene set enrichment analysis (ssGSEA) was used to calculate the glycosylation score for each sample (Yi et al., 2020). The relationship between the glycosylation score and different subtypes was explored to verify the predictive power of the glycosylation regulation model in HCC.

2.5 Analysis of immune cell infiltration and stromal cell enrichment

An increasing number of studies have shown that the tumor microenvironment plays an important role in the pathogenesis and drug resistance of tumors (Wang et al., 2021). Exploring the infiltration of immune cells in different glycosylation regulatory genes is not only helpful in the choice of immunotherapy for different patients, but also in defining different glycosylation subtypes. xCell is a new gene signature method based on the 1822 pure human cell-type transcriptome, which infers 64 immune and stromal cell types (Aran et al., 2017). To show the differences in immune cell infiltration and stromal cells in different subtypes, we used xCell to comprehensively analyze the distribution of immune and stromal cells in each subtype. This process was primarily performed on the IOBR package and was visualized using the pheatmap package.

2.6 Immune checkpoint analysis

The distribution of immune checkpoints is closely related to HCC (Xu et al., 2018). In this study, we selected common immune checkpoints for HCC, including CYBB, IDO1, KIR3DL1, HAVCR2, CD274, PDCD1, TIGIT, CTLA4, LAG3, BTLA, CD27, CD28, CD40, IL2RB, TNFRSF9, TNFRSF4, TNFRSF18, and ICOS. By verifying the expression of immune checkpoints in different subtypes, we can determine the responsiveness of patients with different subtypes to immunotherapy.

2.7 Gene set variation analysis

GSVA can calculate the pathway enrichment score of each sample according to the expression matrix and enrichment of the reaction pathway in the sample (Tao et al., 2021). The clusterProfiler package was used to calculate the path enrichment score of each sample and to construct the expression matrix of the pathway. The limma package was used for differential expression analysis to determine pathways with differences between different subtypes. In this study, to show the enrichment of pathways in different genotypes, we performed GSVA analysis on the subtypes based on the glycosylation-regulated gene model. h.all.v7.4. symbols.gmt was downloaded from the MSigDB database as the reference gene set, and $p < 0.05$ was considered statistically different.

2.8 Drug sensitivity analysis

The Genomics of Drug Sensitivity in Cancer (GDSC) database (Yang et al., 2013), which is designed to apply therapeutic biomarkers to improve the therapeutic efficacy in cancer patients, can detect drug sensitivity based on the genetic characteristics of different patients. In this study, to guide clinical medication based on the glycosylation regulatory gene subtype, we applied the pRRophetic package to identify sensitive drugs in patients with HCC with different glycosylation subtypes.

2.9 Statistical analyses

All statistical analyses were performed using R version 4.1.2. Measurement data between the two groups were tested for normality and homogeneity of variance and significant differences estimated using two-sample *t*-test if the two conditions were met, and a Wilcoxon rank-sum test otherwise. The measurement data among three groups were analyzed using ANOVA, and the Bonferroni test was used for comparison between the two groups. Statistical significance was set at $p < 0.05$.

3 Results

3.1 Genetic variation of glycosylation related genes in hepatocellular carcinoma

This study analyzed a glycosylation gene modification model based on 255 glycosylation-regulatory genes (Supplementary Table S1). A total of 56 prognosis-related glycosylation regulatory genes (Figure 2A) were identified using univariate Cox regression analysis. Excepting that the high expression of *MGAT4C* and *GALNT15* suggested a better patient prognosis, all the other genes were risk factors for HCC. By differential expression analysis of the 56 prognosis-related genes, 43 genes were found to have significant differences between HCC and normal tissues (Figure 2B). The positions of 43 differentially expressed glycosylation regulators with prognostic value on

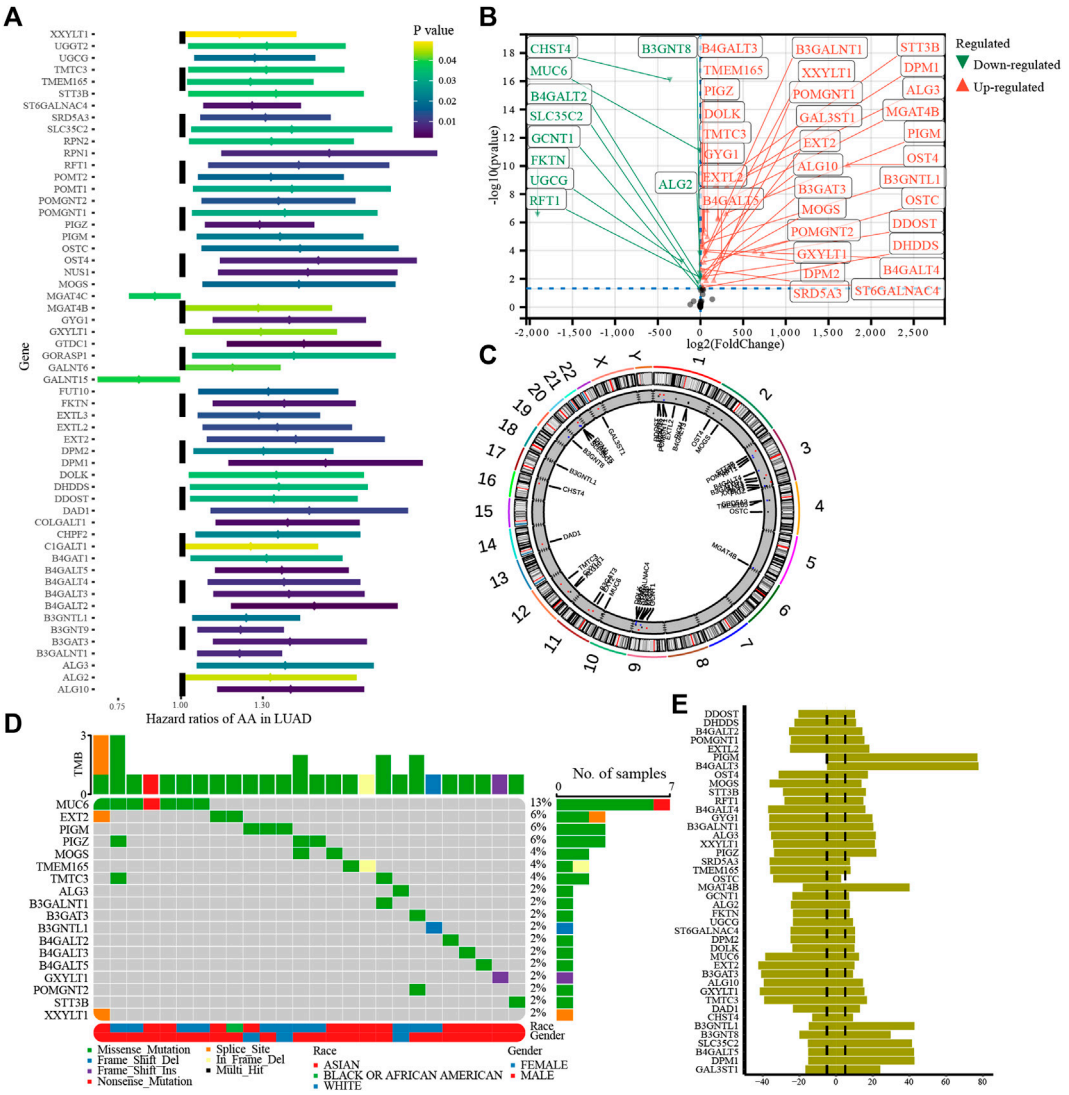


FIGURE 2
Genetic variation in glycosylation-regulatory genes in HCC. **(A)**. Univariate Cox regression analysis of glycosylation regulatory genes; **(B)**. Differential expression of glycosylation regulated genes with prognostic value; **(C)**. Location of 43 differentially expressed genes with prognostic value on the chromosome; **(D)**. Gene mutation of glycosylation regulator; **(E)**. CNV of glycosylation regulators.

chromosomes are shown in Figure 2C. Based on the analysis of 43 prognostic genes of mutation data in the TCGA database, 18 genes, namely *XXYLT1*, *STT3B*, *POMGNT2*, *GXYLT1*, *B4GALT5*, *B4GALT3*, *B4GALT2*, *B3GNTL1*, *B3GAT3*, *B3GALNT1*, *ALG3*, *TMTC3*, *TMEM165*, *MOGS*, *PIGZ*, *PIGM*, *EXT2*, and *MUC6* were mutated in HCC. Among them, *MUC6* had the highest mutation frequency (Figure 2D). Most of the mutations were missense, and MUC6 expression was significantly lower in HCC, suggesting that mutations in *MUC6* may affect MUC6 expression. CNV analysis suggested that amplified copy number variation appeared in *PIGM*, *B4GALT3*, and *MGAT4B*, and deletion CNV appeared in *MUC6*, *B3GAT3*, and *B4GALT4* (Figure 2E). Gene mutation and CNV analyses indicated that the CNV of the 43 glycosylation regulators was the main factor affecting their expression in HCC.

3.2 Modification models mediated by 43 glycosylation regulating genes

The conditional probability survival results are shown in Figure 3A. The longer the survival time of patients with HCC, the higher the subsequent survival rate. The 5-year survival rate of patients increased by 66%, 75%, 81%, and 86%, respectively, from

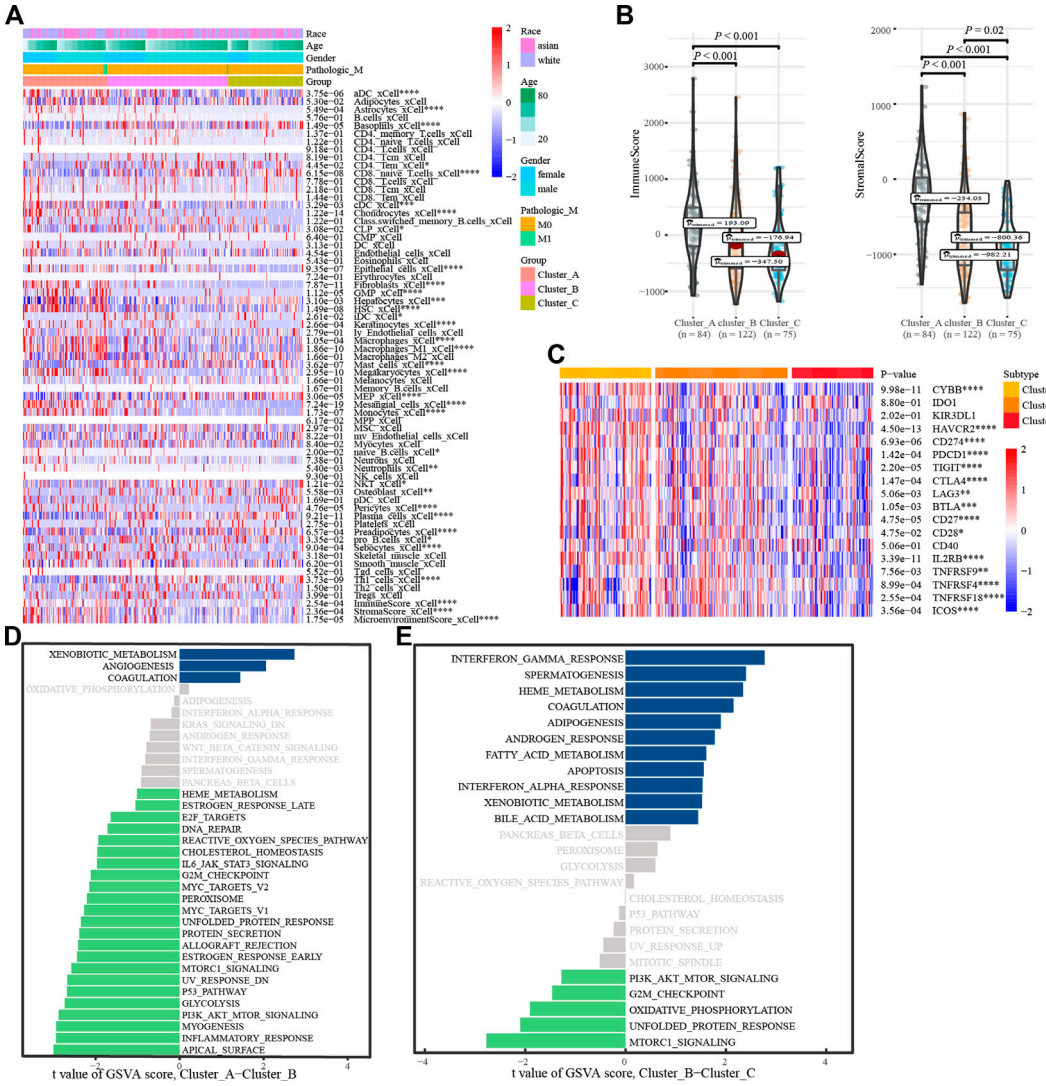


FIGURE 3
Glycosylation regulatory gene model. **(A)**. Conditional probability survival curve; **(B)**. Apply consensus clustering to construct a glycosylation regulator gene model; **(C)**. Kaplan-Meier survival analysis of the glycosylation regulator gene model; **(D)**. Calibration curve at 5 and 8 years; **(E)**. Nomogram.

the overall survival rate of 49%. The conditional probability of survival was consistent with that of patients with HCC, indicating that the data could be used for subsequent analysis. Based on consensus clustering, HCC patients were divided into three subtypes (Figure 3B): female patients, T3-T4 patients, and M1 patients were most likely to be in Cluster_A, and M0 patients were not found in Cluster_C. Patients in Cluster_B had the best prognosis among the three glycosylation types, whereas the survival of patients in Cluster_A and Cluster_C was not significantly different (Figure 3C), which showed that the typing model was highly correlated with the TNM stage. Therefore, we constructed a nomogram based on the glycosylation regulatory gene model. The calibration curve showed the high accuracy of the nomogram (Figure 3D). Using the nomogram plot, we can predict patient survival at five and 8 years based on age, sex, TNM stage, ethnicity, and glycosylation regulatory gene model.

3.3 Annotations of glycosylation regulatory gene model

To understand the biological behavior associated with the different glycosylation regulatory models, we annotated them using immune cell infiltration analysis, stromal cell enrichment analysis, immune score, stromal score, immune checkpoint, and

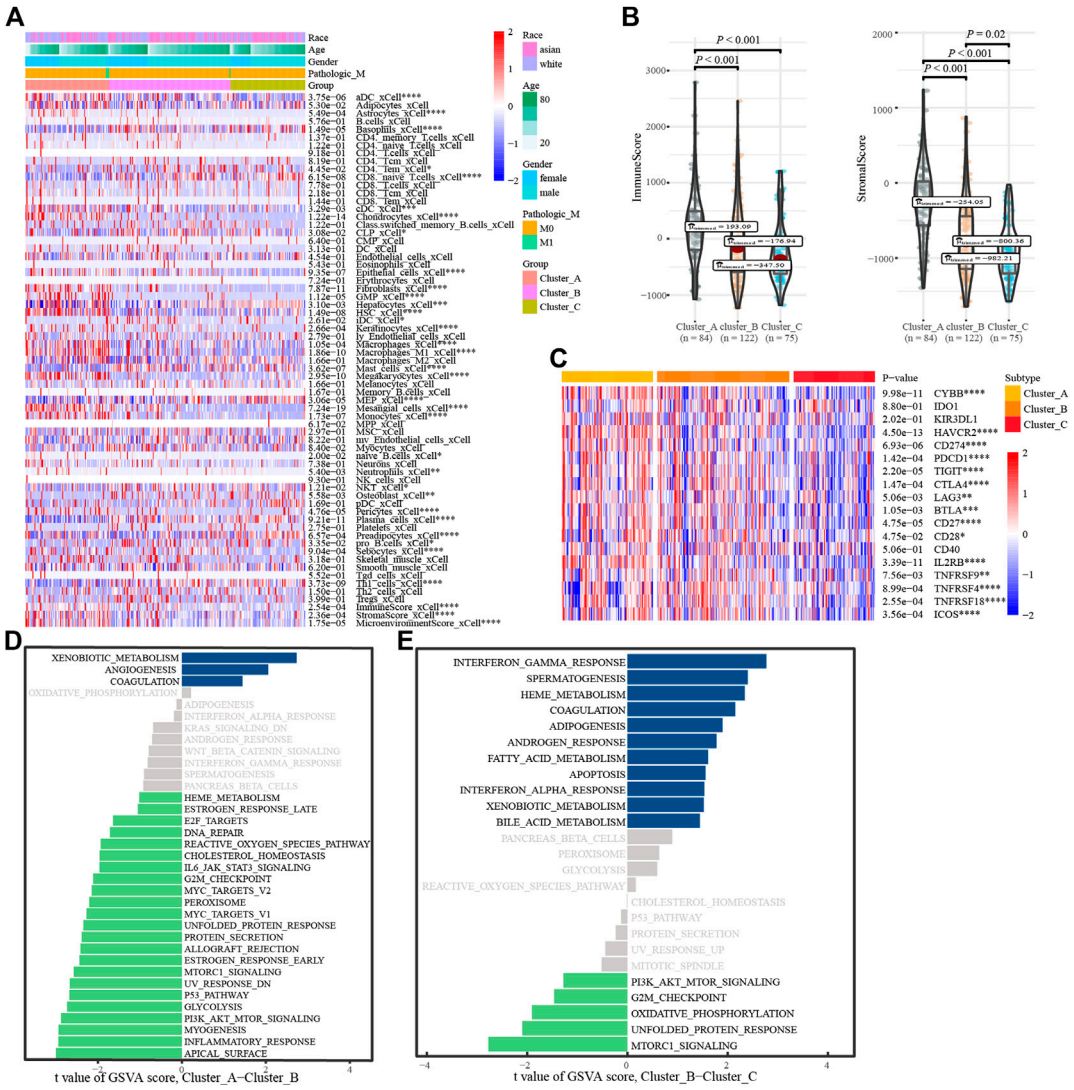


FIGURE 4
Annotation of the glycosylation regulatory gene models. **(A)**. Enrichment of immune cells and stromal cells in different subtypes; **(B)**. Differences in immune scores and stromal scores in different subtypes. **(C)**. Differential expression analysis of immune checkpoints in different subtypes; **(D)**. The GSVA of Cluster_A-Cluster_B; **(E)**. The GSVA of Cluster_B-Cluster_C.

GSVA. Immune cell infiltration analysis showed (Figure 4A) basophils_ xCell, and CD8. The naive T. Cells and macrophages_ M1_ Xcel and other immune cells in cluster_ A had the highest infiltration, whereas cluster_C had the lowest infiltration. The distribution of stromal cells, such as preadipocytes and hepatocytes, in different subtypes was consistent with that of the immune cells. Surprisingly, stromal and immune scores also tended to be consistent across the subtypes (Figure 4B). Analysis of the expression of immune checkpoints in different glycosylation regulatory gene subtypes (Figure 4C) revealed that all the immune checkpoints, except IDO1, KIR3DL1, and CD40, had the highest expression in Cluster_A and the lowest expression in Cluster_C. Based on the above analysis (Figures 4A–C), we classified Cluster_A as an immune rejection phenotype, which was characterized by the highest immune cell infiltration and stroma activation. Cluster_B was classified as an immune-inflammatory phenotype, characterized by partial immune activation. Cluster_C was classified as an immune desert phenotype,characterized by immunosuppression. Further mining of molecular mechanisms in different models of glycosylation regulators. The GSVA results of Cluster_A and Cluster_B indicated that pathways like (Figure 4D), G2-M_checkpoint, PI3K/_AKT/_mTOR_signaling, inflammatory response, and Myc Target were significantly enriched in Cluster A, whereas xenobiotic metabolism, angiogenesis, and coagulation were significantly enriched in Cluster_B. The GSVA

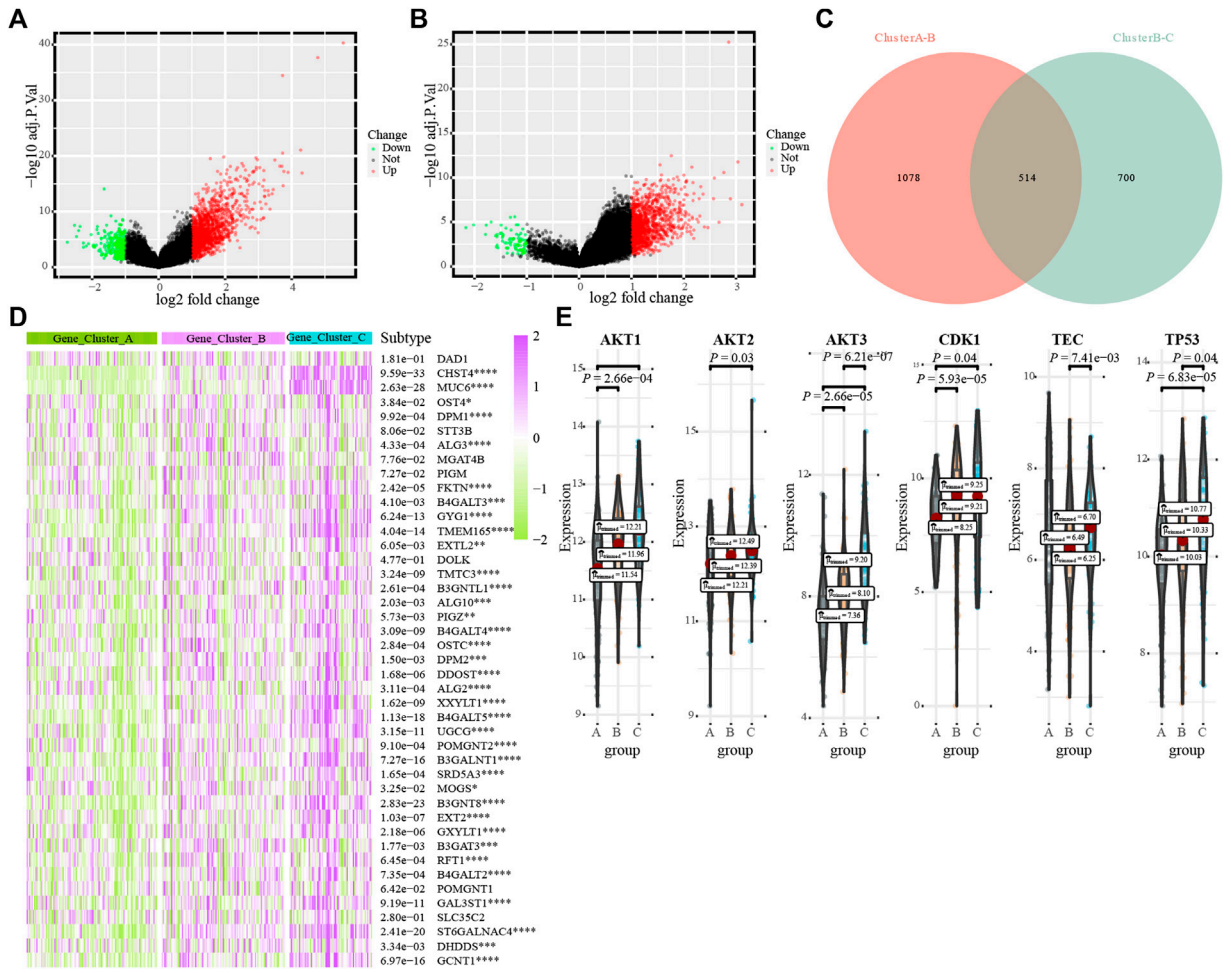


FIGURE 5
The transcriptome characterization of the glycosylation-related models. **(A)**. Distribution of differentially expressed genes between Cluster_ A and Cluster_B; **(B)**. Distribution of differentially expressed genes between Cluster_B and Cluster_C; **(C)**. Venn diagrams; **(D)**. The differential expression of 43 glycosylation regulators in different glycosylation-related subtypes; **(E)**. G2M_CHECKPOINT and PI3K_AKT_MTOR_SIGNALING in different glycosylation subtypes.

results of Cluster_ B and Cluster_C showed (Figure 4E) that G2-M checkpoint, PI3K/_ AKT/mTOR signaling, mTORC1 signaling, and other channels were highly enriched in Cluster_B, whereas interferon gamma response, coalescence, and adipogenesis were highly enriched in Cluster _ C.

3.4 Clinical features and transcriptomic characteristics of the glycosylation-related model

We verified the glycosylation regulatory model by constructing a glycosylation-related model. First, the limma package was used to analyze differentially expressed genes between different glycosylation regulatory gene subtypes. The screening criteria were | log(FC) | >1 and $p < 0.05$. In total, 1,592 differentially expressed genes were obtained from Cluster_A and Cluster_B (Figure 5A). In total, 1,214 differentially expressed genes were obtained from Cluster_B and Cluster_C (Figure 5B). The intersection of these two yielded 514 differentially expressed genes (Figure 5C). Univariate Cox regression analysis of the 514 differentially expressed genes revealed 79 prognosis-related glycosylation-related genes (Supplementary Table S2). PCA was applied to construct a glycosylation-related model, and the two-group model could significantly distinguish patients (Supplementary Figures S1A–C), whereas the survival of the four-group model was not significantly different (Supplementary Figure S1F), and the three-group model was more detailed than the two-group model (Supplementary Figures S1D, E). Therefore, the three-group model was the optimal typing (Gene_Cluster_A, Gene_Cluster_B and Gene_Cluster_C).

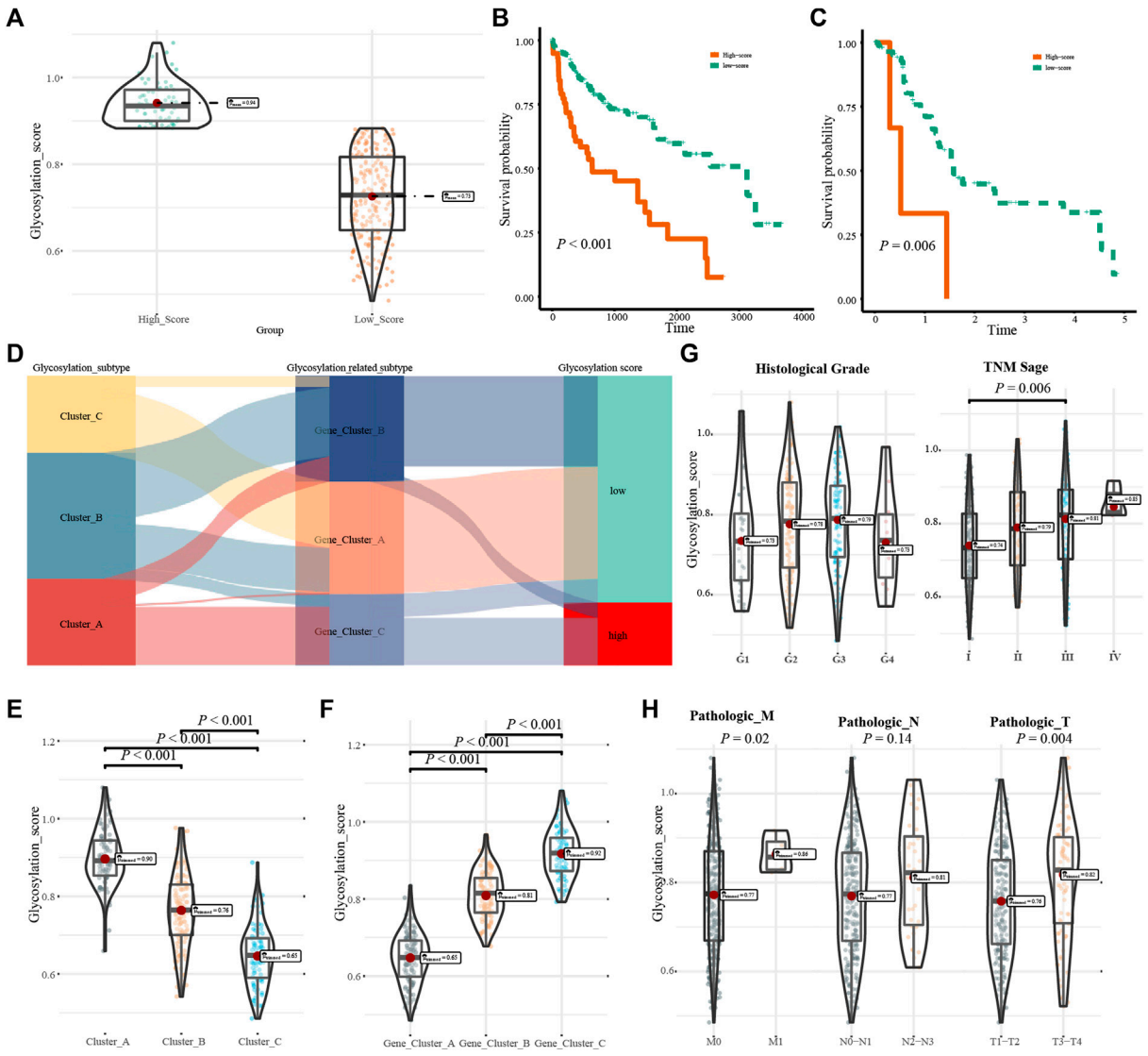


FIGURE 6
Clinical characteristics of the glycosylation score. **(A)**. Group information of glycosylation scores; **(B)**. Kaplan-Meier analysis of the high- and low-score in the TCGA-LIHC cohort; **(C)**. Kaplan-Meier analysis of high- and low-score in the GSE76427 cohort; **(D)**. Correspondence between glycosylation regulatory gene model, glycosylation-related model and glycosylation score; **(E,F)**. Relations between glycosylation regulator model, glycosylation-related model and glycosylation score; **(G)**. The relationship fo differentiation grade and TNM stage with glycosylation score; **(H)**. The relationship of tumor size, lymph node metastasis, and distant metastasis with glycosylation score.

Analysis of the expression of 56 prognostis-related glycosylation-regulatory genes revealed that all genes, except *DAD1*, had the highest expression in Gene_Cluster_C and the lowest expression in Gene_Cluster_A (Figure 5D). The GAVA analysis of the glycosylation regulatory model found that G2-M checkpoint and PI3K/AKT/mTOR signaling were significantly enriched in different models; therefore, we analyzed the expression of G2-M checkpoint markers (p53, CDK1) and PI3K/AKT/mTOR signaling markers (AKT1, AKT2, AKT3, and TEC) in different glycosylation-related models. The expression of key molecules of PI3K/AKT/mTOR signaling, including AKT1, AKT2, and AKT3, were highest in gene_cluster_C and lowest in gene _ cluster _A (Figure 5E). However, p53 and CDK1 showed the opposite trend. Analysis of immune cell infiltration in glycosylation-related models revealed that most immune cells were mainly enriched in Gene_Cluster_C and were low in Gene_Cluster_A (Supplementary Figure S2A). The distribution of immune checkpoints among the three glycosylation-related subtypes showed a consistent trend (Supplementary Figure S2B).

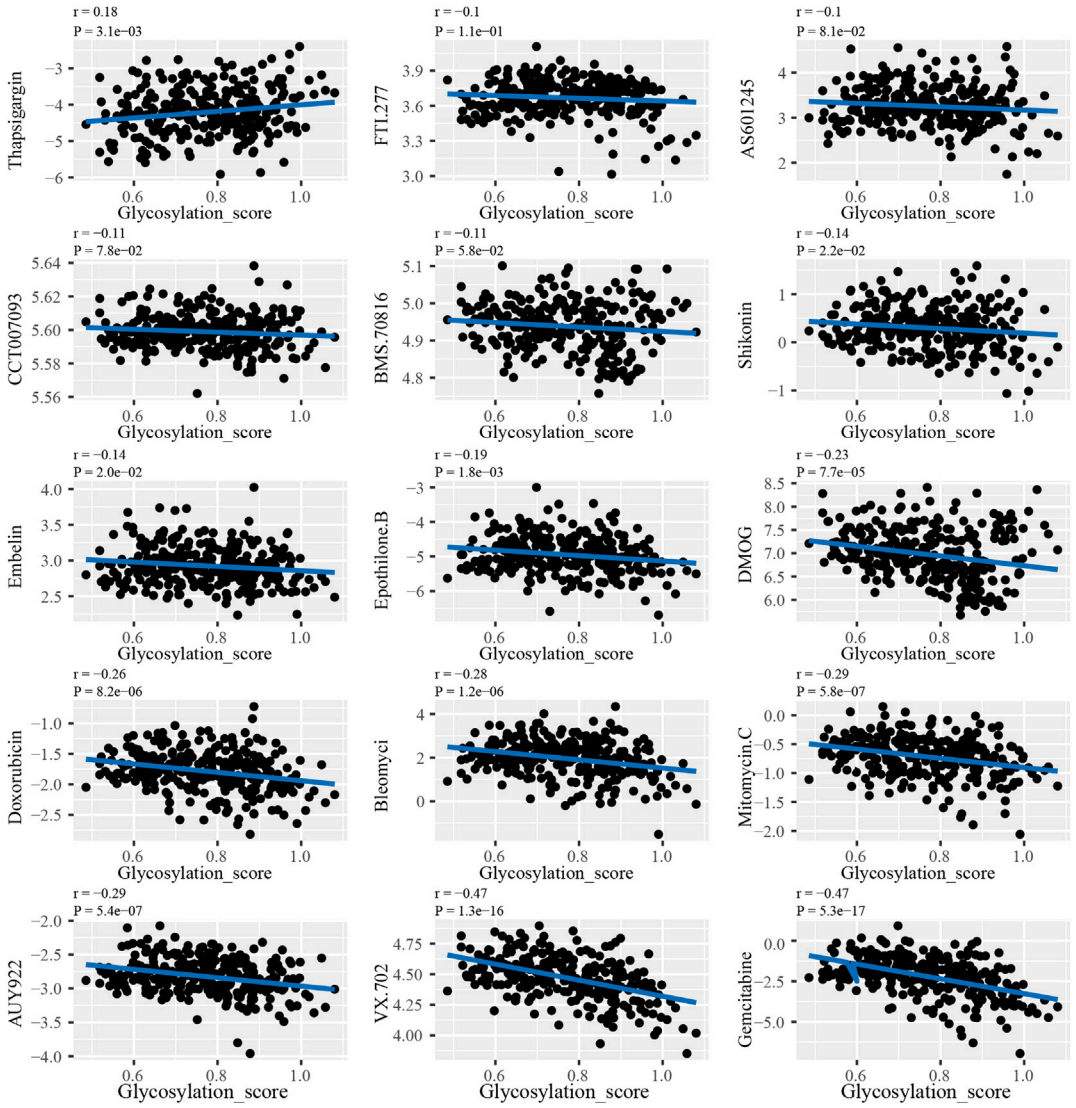


FIGURE 7
The correlation analysis between glycosylation score and drug sensitivity.

3.5 Clinical characteristics of the glycosylation score

Glycosylation scores were calculated for each HCC patient, and all patients were divided into high-score and low-score groups (Figure 6A). Analysis of the survival value of the glycosylation scores showed that patients with high scores had a significantly worse prognosis ($p < 0.001$, Figure 6B). When we verified the prognostic value of the glycosylation score in the GSE76427 cohort, we found that its prognostic value was high (Figure 6C). Clinical subgroup analysis was used to determine whether the glycosylation score had an independent prognostic value and found that the glycosylation score had a high prognostic value in multiple subgroups of age, sex, and TNM stage (Supplementary Figure S3). Analysis of the corresponding relationship between the glycosylation regulation model, glycosylation-related gene model, and glycosylation score revealed that most patients with high glycosylation scores corresponded to Cluster_ A, whereas the low glycosylation score included clusters_ B, Cluster_ C, and a small number of Clusters_ A members (Figure 6D). Interestingly, we found that there were significant differences in glycosylation scores between different genotypes of the glycosylation regulatory gene model and the glycosylation-related gene model (Figures 6E,F). This further explains the relationship between the glycosylation regulatory gene model, glycosylation-related gene model, and glycosylation score. By exploring the relationship between glycosylation score, differentiation grade, and TNM stage, it

was found that the glycosylation score was significantly correlated with TNM stage but not with the differentiation grade of patients (Figure 6G). Further analysis of the relationship between the glycosylation score and tumor size, lymph node metastasis, and distant metastasis in TNM stage showed that the glycosylation score was mainly related to distant metastasis and tumor size (Figure 6H).

3.6 Screening of drug sensitivity based on the glycosylation regulatory gene model

The results of analyzing the drug sensitivity of patients with different subtypes of glycosylation regulatory gene models indicate that (Supplementary Figure S3), patients in Cluster_A were most sensitive to drugs such as AS601245, AUY922, BleomyciBMS.70816, CCT007093, DMOG, Doxorubicin, Embelin, Epothilone. B, FTI.277, Gemcitabine, Mitomycin. C, Shikonin, Thapsigargin and VX.702. To ensure the accuracy of the findings, Pearson's correlation was used to analyze the correlation between drug sensitivity and the glycosylation score (Figure 7). The glycosylation score was positively correlated with the IC_{50} of Thapsigargin (r = 0.18, p = 3.1e − 03), whereas it was negatively correlated with Shikonin (r = − 0.14, p = 2.2E − 02), embelin (r = − 0.14, p = 2.0e − 02), Epothilone B (r = −0.19, p = 1.8e−03), DMOG (r = −0.23, p = 7.7e−05), doxorubicin (r = −0.26, p = 8.2e−06), bleomyci (r = −0.28, p = 1.2e−06), and mitomycin. C (r = −0.29, p = 5.8e−07), AUY922 (r = −0.29, p = 5.4e−07), VX.702 (r = − 0.47, p = 1.3e − 16) and gemcitabine (r = − 0.47, p = 5.3e − 17).

4 Discussion

An increasing number of studies have shown that various glycosylation regulators play an indispensable role in immune microenvironment regulation, cell proliferation, and tumor drug resistance in HCC by regulating the glycosylation of proteins. At present, most studies have only focused on a single glycosylation regulator and the glycosylation regulator model of HCC has not been comprehensively described. To bridge this gap, we analyzed the glycosylation regulator model of HCC based on genomic data, hoping to help in the formulation of immunotherapy strategies.

In this study, 43 prognostic genes with differential expression in HCC were analyzed based on 255 glycosylation regulators in all tumors. Based on this, three glycosylation regulation models were revealed, where cluster_ A corresponded to the immune rejection phenotype, cluster_ B corresponded to the immunoinflammatory phenotype, and cluster_ C corresponded to the immune desert phenotype. Immune-exclusive HCC is not without the infiltration of immune cells, but these HCC cells escape the detection and destruction of immune cells *via* immunogenic shaping (Lindblad et al., 2021). In the process of immune escape in HCC, the IL-6 JAK/STAT3 signaling pathway and MYC pathway play an important role, in which IL6/JAK/STAT3 mainly promotes the glycosylation of PD-L1 to increase its stability (Zhou et al., 2020). MYC mediates the immune escape of HCC cells, mainly *via* the β-catenin protein encoded by the *CTNNB1* gene (Luke et al., 2019). In this study, we found that the prognosis of patients in cluster_B was the best; therefore, our results are consistent with these previous results, which validates the reliability of the glycosylation regulator model. The infiltration characteristics of immune cells and the prognosis of patients with immune rejection and immune desert HCC highly corresponded to cluster_ A and cluster_ C, which further proves the importance of the glycosylation regulatory gene model in HCC. We also constructed nomograms based on the glycosylation regulatory gene model. Clinicians can not only predict the prognosis of patients with HCC based on the nomogram, but can also use it to formulate clinical treatment strategies.

Differentially expressed genes are considered to be the characteristic genes of the phenotype; therefore, this study verified the glycosylation regulatory model based on the differentially expressed genes between different glycosylation regulatory gene models. Similar to the glycosylation regulatory model, we could divide HCC patients into three subtypes based on differentially expressed genes. Using enrichment analysis of immune and stromal cells, we found that these three gene subtypes were significantly related to the enrichment of immune and stromal cells. This also proves that the glycosylation regulatory gene model has different characteristics of immune cell infiltration. To determine the influence of individual differences in patients with HCC on the glycosylation regulator model, we constructed a glycosylation score. Patients in cluster_A, characterized by immune rejection, all corresponded to high glycosylation scores, whereas patients in cluster_B, characterized by immune inflammation, mostly corresponded to low glycosylation scores. In addition, we found that both G2-M checkpoint and PI3K/AKT/mTOR signaling differed significantly between different glycosylation regulatory models. Both G2-M checkpoint and PI3K/AKT/mTOR signaling differed significantly between glycosylation-regulatory gene subtypes (Sheng et al., 2021). Related studies have shown that LARP4B, OGDHL, and miR-454-3p may affect the prognosis of patients with HCC *via* G2-M checkpoint (Li et al., 2019a; Li et al., 2019b; Jiao et al., 2019). Glycosylation is closely associated with HCC cell proliferation (Takahashi et al., 2020). Therefore, we inferred that glycosylation-regulated genotypes have different cell proliferation characteristics.

We screened the drug sensitivity in HCC patients based on the glycosylation regulatory model and glycosylation score, and found drugs such as thapsigargin, shikonin, embelin, epothilone

B, DMOG, doxorubicin, bleomycin, mitomycin C, AUY922, VX.702, and gemcitabine. Among them, the roles of thapsigargin, shikonin, embelin, doxorubicin, bleomycin, mitomycin C, and gemcitabine drugs in HCC have long been reported (You et al., 1994; Newell et al., 2010; Zhang et al., 2020b; Djokic et al., 2020; Diggs et al., 2021; Song et al., 2021; Cheng et al., 2022), and there are no clear studies on the indications for these drugs. In this study, we evaluated the sensitivity of these drugs on different subtypes based on the glycosylation regulatory model. We also analyzed the drugs associated with glycosylation scores to ensure the accuracy of the analysis. Thus, we provide the scope for the application of these drugs after identification of the subtypes of patients with HCC.

In conclusion, the glycosylation regulatory gene model constructed in this study can comprehensively evaluate the characteristics of immune cell infiltration in HCC patients and will facilitate guidance for clinical treatment. In addition, we constructed the glycosylation score for HCC, which had an independent prognostic value. More importantly, we screened drug sensitivity for patients with HCC based on the glycosylation regulatory gene model and glycosylation score. Our results provide a new approach for improving the clinical treatment of patients with HCC and promoting individualized HCC 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.

Author contributions

LZ: data analysis and writing; YG: revision of the article and plotting; ZL: quality control and data verification; JM: data visualization and quality control; YP: revision and typesetting; DZ: design of the study, drafting and rivision.

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/fgene.2022.1001901/full#supplementary-material

References

Aran, D., Hu, Z., and Butte, A. J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 18 (1), 220. doi:10.1186/s13059-017-1349-1

Asazawa, H., Kamada, Y., Takeda, Y., Takamatsu, S., Shinzaki, S., Kim, Y., et al. (2015). Serum fucosylated haptoglobin in chronic liver diseases as a potential biomarker of hepatocellular carcinoma development. *Clin. Chem. Lab. Med.* 53 (1), 95–102. doi:10.1515/cclm-2014-0427

Cheng, H. H., Liao, W. C., Lin, R. A., Chen, I. S., Chou, C. T., Jan, C. R., et al. (2022). Hydroxytyrosol [2-(3, 4-dihydroxyphenyl)-ethanol], a natural phenolic compound found in the olive, alters Ca(2+) signaling and viability in human HepG2 hepatoma cells. *Chin. J. Physiol.* 65 (2), 30. doi:10.4103/cjp.cjp_74_21

Chiu, C. C., Chen, Y. C., Bow, Y. D., Chen, J. Y., Liu, W., Huang, J. L., et al. (2022). diTFPP, a phenoxyphenol, sensitizes hepatocellular carcinoma cells to C(2)-ceramide-induced autophagic stress by increasing oxidative stress and ER stress accompanied by LAMP2 hypoglycosylation. *Cancers* 14 (10), 2528. doi:10.3390/cancers14102528

Diggs, L. P., Ruf, B., Ma, C., Heinrich, B., Cui, L., Zhang, Q., et al. (2021). CD40-mediated immune cell activation enhances response to anti-PD-1 in murine intrahepatic cholangiocarcinoma. *J. Hepatol.* 74 (5), 1145–1154. doi:10.1016/j.jhep.2020.11.037

Djokic, M., Cemazar, M., Bosnjak, M., Dezman, R., Badovinac, D., Miklavcic, D., et al. (2020). A prospective phase II study evaluating intraoperative electrochemotherapy of hepatocellular carcinoma. *Cancers* 12 (12), 3778. doi:10.3390/cancers12123778

Huang, X., Zhang, Q., Lou, Y., Wang, J., Zhao, X., Wang, L., et al. (2019). USP22 deubiquitinates CD274 to suppress anticancer immunity. *Cancer Immunol. Res.* 7 (10), 1580–1590. doi:10.1158/2326-6066.Cir-18-0910

Jiao, Y., Li, Y., Fu, Z., Hou, L., Chen, Q., Cai, Y., et al. (2019). OGDHL expression as a prognostic biomarker for liver cancer patients. *Dis. Markers* 2019, 9037131. doi:10.1155/2019/9037131

Ledoux, S., Yang, R., Friedlander, G., and Laouari, D. (2003). Glucose depletion enhances P-glycoprotein expression in hepatoma cells: Role of endoplasmic reticulum stress response. *Cancer Res.* 63 (21), 7284–7290.

Lee, J., Yeo, I., Kim, Y., Shin, D., Kim, J., Kim, Y., et al. (2022). Comparison of fucose-specific lectins to improve quantitative AFP-L3 assay for diagnosing hepatocellular carcinoma using mass spectrometry. *J. Proteome Res.* 21 (6), 1548–1557. doi:10.1021/acs.jproteome.2c00196

Li, Y., Jiao, Y., Fu, Z., Luo, Z., Su, J., and Li, Y. (2019). High miR-454-3p expression predicts poor prognosis in hepatocellular carcinoma. *Cancer Manag. Res.* 11, 2795–2802. doi:10.2147/cmar.S196655

Li, Y., Jiao, Y., Li, Y., and Liu, Y. (2019). Expression of La ribonucleoprotein domain family member 4B (LARP4B) in liver cancer and their clinical and prognostic significance. *Dis. Markers* 2019, 1569049. doi:10.1155/2019/1569049

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell. Syst.* 1 (6), 417–425. doi:10.1016/j.cels.2015.12.004

Lindblad, K. E., Ruiz de Galarreta, M., and Lujambio, A. (2021). Tumor-Intrinsic mechanisms regulating immune exclusion in liver cancers. *Front. Immunol.* 12, 642958. doi:10.3389/fimmu.2021.642958

Liu, B., Pan, S., Xiao, Y., Liu, Q., Xu, J., and Jia, L. (2018). LINC01296/miR-26a/GALNT3 axis contributes to colorectal cancer progression by regulating O-glycosylated MUC1 via PI3K/AKT pathway. *J. Exp. Clin. Cancer Res.* 37 (1), 316. doi:10.1186/s13046-018-0994-x

Liu, H. M., Ma, L. L., Cao, B., Lin, J. Z., Han, L., Li, C. Y., et al. (2021). Progress in research into the role of abnormal glycosylation modification in tumor immunity. *Immunol. Lett.* 229, 8–17. doi:10.1016/j.imlet.2020.11.003

Liu, Y., Huang, H., Liu, M., Wu, Q., Li, W., and Zhang, J. (2017). MicroRNA-24-1 suppresses mouse hepatoma cell invasion and metastasis via directly targeting O-GlcNAc transferase. *Biomed. Pharmacother. = Biomedecine Pharmacother.* 91, 731–738. doi:10.1016/j.biopha.2017.05.007

Luke, J. J., Bao, R., Sweis, R. F., Spranger, S., and Gajewski, T. F. (2019). WNT/β-catenin pathway activation correlates with immune exclusion across human cancers. *Clin. Cancer Res.* 25 (10), 3074–3083. doi:10.1158/1078-0432.Ccr-18-1942

Na, K., Kim, M., Kim, C. Y., Lim, J. S., Cho, J. Y., Shin, H., et al. (2020). Potential regulatory role of human-carboxylesterase-1 glycosylation in liver cancer cell growth. *J. Proteome Res.* 19 (12), 4867–4883. doi:10.1021/acs.jproteome.0c00787

Newell, A. M., Yousef, G. G., Lila, M. A., Ramírez-Mares, M. V., and de Mejia, E. G. (2010). Comparative *in vitro* bioactivities of tea extracts from six species of Ardisia and their effect on growth inhibition of HepG2 cells. *J. Ethnopharmacol.* 130 (3), 536–544. doi:10.1016/j.jep.2010.05.051

Park, S. S. (2019). Post-glycosylation modification of sialic acid and its role in virus pathogenesis. *Vaccines* 7 (4), E171. doi:10.3390/vaccines7040171

Pousset, D., Piller, V., Bureaud, N., Monsigny, M., and Piller, F. (1997). Increased alpha2, 6 sialylation of N-glycans in a transgenic mouse model of hepatocellular carcinoma. *Cancer Res.* 57 (19), 4249–4256.

Sheng, X., Zhu, P., Zhao, Y., Zhang, J., Li, H., Zhao, H., et al. (2021). Effect of PI3K/AKT/mTOR signaling pathway on regulating and controlling the anti-invasion and metastasis of hepatoma cells by bufalin. *Recent Pat. anticancer. Drug Discov.* 16 (1), 54–65. doi:10.2174/1574892816666210201120324

Song, M., Zhang, H., Chen, Z., Yang, J., Li, J., Shao, S., et al. (2021). Shikonin reduces hepatic fibrosis by inducing apoptosis and inhibiting autophagy via the platelet-activating factor-mitogen-activated protein kinase axis. *Exp. Ther. Med.* 21 (1), 28. doi:10.3892/etm.2020.9460

Takahashi, H., Kamiyama, T., Hirane, N., Kobayashi, N., Aiyama, T., Nagatsu, A., et al. (2020). Analysis of the correlation between alterations in N-glycans and invasiveness in liver cancer cell lines. *Oncol. Rep.* 44 (6), 2757–2769. doi:10.3892/or.2020.7795

Tao, P. F., and Huang, H. C. (2019). Regulation of AβPP glycosylation modification and roles of glycosylation on AβPP cleavage in Alzheimer's disease. *ACS Chem. Neurosci.* 10 (5), 2115–2124. doi:10.1021/acschemneuro.8b00574

Tao, Y., Zou, W., Nanayakkara, S., Freindorf, M., and Kraka, E. (2021). A revised formulation of the generalized subsystem vibrational analysis (GSVA). *Theor. Chem. Acc.* 140 (3), 31. doi:10.1007/s00214-021-02727-y

Timmerman, M. E., Ceulemans, E., De Roover, K., and Van Leeuwen, K. (2013). Subspace K-means clustering. *Behav. Res. Methods* 45 (4), 1011–1023. doi:10.3758/s13428-013-0329-y

Vu, L. D., Gevaert, K., and De Smet, I. (2018). Protein language: Post-translational modifications talking to each other. *Trends Plant Sci.* 23 (12), 1068–1080. doi:10.1016/j.tplants.2018.09.004

Wang, N., Wang, S., Wang, X., Zheng, Y., Yang, B., Zhang, J., et al. (2021). Research trends in pharmacological modulation of tumor-associated macrophages. *Clin. Transl. Med.* 11 (1), e288. doi:10.1002/ctm2.288

Wattanavises, S., Silsirivanit, A., Sawanyawisuth, K., Cha'on, U., Waraasawapati, S., Saentaweesuk, W., et al. (2019). Increase of MAL-II binding Alpha2, 3-sialylated glycan is associated with 5-FU resistance and short survival of cholangiocarcinoma patients. *Medicina* 55 (12), E761. doi:10.3390/medicina55120761

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinforma. Oxf. Engl.* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170

Xu, F., Jin, T., Zhu, Y., and Dai, C. (2018). Immune checkpoint therapy in liver cancer. *J. Exp. Clin. Cancer Res.* 37 (1), 110. doi:10.1186/s13046-018-0777-4

Yang, W., Soares, J., Greninger, P., Edelman, E. J., Lightfoot, H., Forbes, S., et al. (2013). Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* 41, D955–D961. doi:10.1093/nar/gks1111

Yi, M., Nissley, D. V., McCormick, F., and Stephens, R. M. (2020). ssGSEA score-based Ras dependency indexes derived from gene expression data reveal potential Ras addiction mechanisms with possible clinical implications. *Sci. Rep.* 10 (1), 10258. doi:10.1038/s41598-020-66986-8

You, J. S., Hau, D. M., Chen, K. T., and Huang, H. F. (1994). Combined effects of chuling (Polyporus umbellatus) extract and mitomycin C on experimental liver cancer. *Am. J. Chin. Med.* 22 (1), 19–28. doi:10.1142/s0192415x94000048

Zhang, D., Yang, L., Liu, X., Gao, J., Liu, T., Yan, Q., et al. (2020). Hypoxia modulates stem cell properties and induces EMT through N-glycosylation of EpCAM in breast cancer cells. *J. Cell. Physiol.* 235 (4), 3626–3633. doi:10.1002/jcp.29252

Zhang, G., Huang, X., Xiu, H., Sun, Y., Chen, J., Cheng, G., et al. (2020). Extracellular vesicles: Natural liver-accumulating drug delivery vehicles for the treatment of liver diseases. *J. Extracell. Vesicles* 10 (2), e12025. doi:10.1002/jev2.12025

Zhou, B., Yan, J., Guo, L., Zhang, B., Liu, S., Yu, M., et al. (2020). Hepatoma cell-intrinsic TLR9 activation induces immune escape through PD-L1 upregulation in hepatocellular carcinoma. *Theranostics* 10 (14), 6530–6543. doi:10.7150/thno.44417

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 30 November 2022
DOI 10.3389/fgene.2022.981145

Check for updates

OPEN ACCESS

EDITED BY
Ehsan Nazemalhosseini-Mojarad,
Shahid Beheshti University of Medical Sciences, Iran

REVIEWED BY
Emil Bulatov,
Kazan Federal University, Russia
Alessandro Rizzo,
University of Bologna, Italy

*CORRESPONDENCE
Kai Ye,
kaiye@xjtu.edu.cn

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 29 June 2022
ACCEPTED 24 October 2022
PUBLISHED 30 November 2022

CITATION
Wang T, Xu C, Xu D, Yang X, Liu Y, Li X, Li Z, Dang N, Lv Y, Zhang Z, Li L and Ye K (2022), Integrating cell interaction with transcription factors to obtain a robust gene panel for prognostic prediction and therapies in cholangiocarcinoma. *Front. Genet.* 13:981145. doi: 10.3389/fgene.2022.981145

COPYRIGHT
© 2022 Wang, Xu, Xu, Yang, Liu, Li, Li, Dang, Lv, Zhang, Li and Ye. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Integrating cell interaction with transcription factors to obtain a robust gene panel for prognostic prediction and therapies in cholangiocarcinoma

Tingjie Wang[1], Chuanrui Xu[2], Dan Xu[3], Xiaofei Yang[4,5,6], Yaxin Liu[3], Xiujuan Li[7], Zihang Li[3], Ningxin Dang[5], Yi Lv[1], Zhijing Zhang[2], Lei Li[2] and Kai Ye[1,3,5,6,7,8]*

[1]National Local Joint Engineering Research Center for Precision Surgery and Regenerative Medicine, Shaanxi Provincial Center for Regenerative Medicine and Surgical Engineering, Center for Mathematical Medical, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China, [2]School of Pharmacy, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [3]School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi, China, [4]School of Computer Science and Technology, Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, China, [5]Genome Institute, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi, China, [6]MOE Key Lab for Intelligent Networks and Networks Security, Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, China, [7]School of Automation Science and Engineering, Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, China, [8]Faculty of Science, Leiden University, Leiden, Netherlands

Objective: The efficacy of immunotherapy for cholangiocarcinoma (CCA) is blocked by a high degree of tumor heterogeneity. Cell communication contributes to heterogeneity in the tumor microenvironment. This study aimed to explore critical cell signaling and biomarkers induced *via* cell communication during immune exhaustion in CCA.

Methods: We constructed empirical Bayes and Markov random field models eLBP to determine transcription factors, interacting genes, and associated signaling pathways involved in cell-cell communication using single-cell RNAseq data. We then analyzed the mechanism of immune exhaustion during CCA progression.

Results: We found that *VEGFA*-positive macrophages with high levels of *LGALS9* could interact with *HAVCR2* to promote the exhaustion of $CD8^+$ T cells in CCA. Transcription factors *SPI1* and *IRF1* can upregulate the expression of *LGALS9* in *VEGFA*-positive macrophages. Subsequently, we obtained a panel containing 54 genes through the model, which identified subtype S2 with high expression of immune checkpoint genes that are suitable

Abbreviations: CAF, cancer-associated fibroblasts; CCA, cholangiocarcinoma; CP1, checkpoint inhibitors; DC, dendritic cells; DEGs, differential genes; eLBP, empirical Bayes loopy belief propagation; iCCA, intrahepatic; ICIs, immune checkpoint inhibitors; ICGs, mimmune checkpoint genes; LASSO, least absolute shrinkage and selection operator; MRF, Markov random field; TF, transcription factors; TMRS, tumor microenvironment risk score; TMEs, Tumor microenvironments; UMAP, uniform manifold approximation and projection.

for immunotherapy. Moreover, we found that patients with subtype S2 with a higher mutation ratio of *MUC16* had immune-exhausted genes, such as *HAVCR2* and *TIGIT*. Finally, we constructed a nine-gene eLBP-LASSO-COX risk model, which was designated the tumor microenvironment risk score (TMRS).

Conclusion: Cell communication-related genes can be used as important markers for predicting patient prognosis and immunotherapy responses. The TMRS panel is a reliable tool for prognostic prediction and chemotherapeutic decision-making in CCA.

KEYWORDS

cell communication, single cell RNAseq, transcript factor, cholangiocarcinoma, immunotherapy

1 Introduction

Cholangiocarcinoma (CCA) is the second most common form of liver malignancy with increasing diagnostic incidence and mortality rates (Banales et al., 2020). Patients with CCA are divided into two subtypes depending on their anatomical site of origin: intrahepatic (iCCA or ICC), perihilar (pCCA), and distal (dCCA) CCA (Bridgewater et al., 2014). Most patients with CCA have lost the opportunity to undergo surgical resection upon diagnosis (Brandi et al., 2020). Recent studies have revealed that immunotherapies can protect against tumor development and metastasis. Immune checkpoint inhibitors (ICIs), chimeric antigen receptor T-cells, and tumor vaccines are the main immunotherapies used in clinical practice.

Tumor heterogeneity and stromal/immune cells in the tumor immune microenvironment (TME) significantly influence the effect of the immunotherapy (Tang et al., 2021). Tumors can be categorized into inflamed, immune desert, or immune-excluded phenotypes based on the spatial localization of immune cells in the tumor and stromal compartments (Hegde and Chen, 2020). Inflamed tumors are typically associated with a response to ICIs, particularly PD-L1- and PD-1-directed antibodies (Jiang et al., 2020). The clinical efficacy of PD-1/PD-L1 ICIs enhance the cytotoxicity of $CD8^+$ T cells. However, clinical experience has shown that a large percentage of CCAs (>50%) do not respond to ICIs (e.g., high tumor PD-L1 expression) (Hegde and Chen, 2020). Therefore, it is necessary to investigate the components of TME to identify other immune checkpoints during CCA immunotherapy.

Interactions between ligands and receptors in the TME and tumor cells play a crucial role in tumor progression (Zhang et al., 2020). It has been revealed that cancer-associated fibroblasts (CAF) in ICC induced significant epigenetic alterations in ICC cells *via* IL-6 secretion and subsequent upregulation of *EZH2* through *STAT3* activation (Zhang et al., 2020). In addition, PD-L1+ TAMs in the TME facilitate CCA progression (Rizzo et al., 2021a). However, the mechanism of cell interaction and upstream and downstream signal transduction elements influencing the CCA immunotherapy of remains unclear.

In the present study, we developed an eLBP algorithm by integrating empirical Bayes with Markov random fields to infer the intact intercellular communication pattern, including interacting cell types and critical TFs that could promote cell interactions simultaneously. We revealed a unique signaling transduction model between tumor-associated macrophages and exhausted $CD8^+$ T cells that could influence immunotherapy and novel immune checkpoint genes in CCA patients. We also constructed a gene classifier to distinguish inflamed tissue with high expression of immune checkpoint genes (ICGs) from the immune desert subtype and predict the immunotherapeutic effect in the two CCA subtypes.

2 Materials and methods

2.1 Bulk RNA-seq and single-cell datasets

CCA sequencing data were retrieved from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/). All the datasets used in this study are listed in Supplementary Table S1.

2.2 Single-cell RNA-seq data processing

The single-cell datasets were downloaded from the GEO database (Supplementary Table S1). The raw gene expression matrix was imported and processed using the Seurat R package (version 3.1.5) (Butler et al., 2018). Cells with unique molecular identifier counts below 200 or mitochondrial content above 35% were removed. Normalization and dimension-reduction processes were performed using Seurat R package (3.1.5). Clusters were computed using the FindClusters function (resolution = 0.8) and visualized using uniform manifold approximation and projection (UMAP), as implemented in

Seurat. Differential expression between clusters was calculated using a likelihood-ratio test for single-cell gene expression implemented in Seurat, with a family-wise error rate of 5%. Cell types were defined according to lineage-specific marker genes.

2.3 Empirical bayes loopy belief propagation algorithm

Interactions are frequent among TME cells depending on the ligand and receptor pairs. We evaluated the contribution of cell-interaction genes to cell communication by measuring the importance of their participant-enriched pathways in each cell type. Meanwhile, transcription factors can enhance cell signaling by promoting the expression of interacting genes. The importance and critical TF regulons were calculated using empirical Bayes loopy belief propagation (eLBP). The cell clustering results and differential genes in each cell type were used as the input data for eLBP. The algorithm consists of the following five steps based on cell interaction genes (CIGs).

2.3.1 Construction of interaction network of CIGs

We first identified the differentially expressed genes (DEGs) for each cell type. The pathways that were significantly enriched (*P*.adj <0.05) in the KEGG, GO, and REACTOME datasets for each cell type were selected. We obtained all potential interaction gene pairs during cell communication from the CellChat database (Jin et al., 2021). Next, we extracted the pathways containing cell-communication genes in each cell type. Subsequently, we conducted the protein interaction (Tohonen et al., 2015) analysis of these pathway genes to produce a potential gene interaction undirected graph for each cell type.

2.3.2 Construction of a Markov random field for CIGs

After constructing the gene interaction undirected graph for each cell type, we adopted an empirical Bayes Markov random field (MRF) network to calculate the importance of communication-related genes in each cell type. The network was represented by $G_k = \{V^k, E^k\}$ in where $k \in \{1, 2, \ldots, N\}$ denotes the cell types obtained from a single-cell dataset, $V_i^k = \in \{V_1^k, V_2^k, \ldots, V_M^k\}$ denotes the gene nodes in the undirected graph in the cell type k containing total M genes, and $E = \{(V_i^k, V_j^k), V_i^k, V_j^k \in V\}$ represents the interaction intensity between two genes V_i and V_j in the interaction network of this cell type k. In addition, we set the dummy variable $L = \{+1, -1\}$ to denote a set of labels, in which +1 denotes the genes that were present in the cell type k, while −1 indicates the genes absent in the cell type k.

To construct a complete MRF model, we specified the node potentials of the class labels of node V_i^k, denoted as $\psi(V_i^k)$, which are denoted as V_i, and V_j are denoted as $\psi(V_i^k, V_j^k)$. In this formulation, we defined node potentials as the probabilities that genes could occur in the cell type k, $\psi(V_i^k) = P(V_i^k)$. Moreover, we defined the cosine similarity, $Co\sin(V_i^k, V_j^k)$, between two adjacent nodes with labels V_i^k and V_j^k as the edge potential. We define potential functions and calculated the joint distribution of all indicator variables V can be denoted as Eq. 1 (Song You et al., 2018).

$$\boldsymbol{p}(\boldsymbol{V}) = \frac{\boldsymbol{1}}{\boldsymbol{Z}} \prod_{\boldsymbol{V}_i^k \in \boldsymbol{V}} \boldsymbol{\psi}(\boldsymbol{V}_i^k) \prod_{\left(\boldsymbol{V}_i^k, \boldsymbol{V}_j^k\right) \in \boldsymbol{E}} \boldsymbol{\psi}\left(\boldsymbol{V}_i^k, \boldsymbol{V}_j^k\right) \quad (1)$$
$$\left(\boldsymbol{i}, \boldsymbol{j} \in \{\boldsymbol{1}, \boldsymbol{2}, \ldots, \boldsymbol{M}\}; \boldsymbol{k} \in \{\boldsymbol{1}, \boldsymbol{2}, \ldots, \boldsymbol{N}\}\right)$$

where, Z is the normalization constant. Next, we obtained the eLBP, a widely adopted approximate inference algorithm for the MRF. When the eLBP converges, the final belief node V_i^k can be calculated using Eq. 2 (Tirosh et al., 2016).

$$\boldsymbol{S}(\boldsymbol{V}_i^k) = \frac{\boldsymbol{1}}{\boldsymbol{Z}_2} \boldsymbol{\psi}(\boldsymbol{V}_i^k) \prod_{\boldsymbol{V}_j^k \in \boldsymbol{N}_i} \boldsymbol{m}_{\boldsymbol{j} \to \boldsymbol{i}}(\boldsymbol{V}_i^k) \quad (2)$$
$$\left(\boldsymbol{i}, \boldsymbol{j} \in \{\boldsymbol{1}, \boldsymbol{2}, \ldots, \boldsymbol{M}\}; \boldsymbol{k} \in \{\boldsymbol{1}, \boldsymbol{2}, \ldots, \boldsymbol{N}\}\right)$$

where Z_2 is the normalization constant, and N_i denotes the neighbor nodes connected with node V_i in cell type k.

2.3.3 Construction of positive and negative reference cells and calculation of the prior probability of each gene node in eLBP

The prior probabilities in Eqs 1, 2 are calculated using the empirical Bayes method. The prior probability of node V_i^k in the network indicates the probability of the gene i being presented in a cell type k. To calculate the prior probabilities of genes in cell-type k, we first constructed a cell-type k positive and negative dataset. Here, we first removed the genes whose expression levels were equal to zero in more than 50% of the cells in each type. Then, we randomly selected 200 or all the cells if the cell number was <200 to obtain a positive dataset of cell type k, we selected 200 cells that were not in cell type k to obtain the corresponding negative dataset.

We then assumed that genes in the cell type k followed a Poisson distribution with the parameter θ (Eq. 3).

$$\boldsymbol{P}(\boldsymbol{V}|\theta) = \boldsymbol{e}^{-\theta}\theta^{\boldsymbol{v}} \Big/ \boldsymbol{v!} \quad (3)$$

where θ has a Gamma prior with unknown parameters α and β (Eq. 4).

$$\boldsymbol{\pi}(\boldsymbol{\theta}) = \frac{\boldsymbol{\beta}^{\boldsymbol{\alpha}}}{\boldsymbol{\Gamma}(\boldsymbol{\alpha})} \boldsymbol{e}^{-\boldsymbol{\beta}\boldsymbol{\theta}} \boldsymbol{\theta}^{\boldsymbol{\alpha}-1} \quad (4)$$

Based on this assumption, under the observation vector X_{CN}, we calculated the posterior values of α and β *via* Eq. 5 by randomly sampling 50% of the cells in X_{CN}.

$$E(x) = \frac{\alpha}{\beta};\ Var(x) = \frac{\alpha(1+\beta)}{\beta^2} \tag{5}$$

Subsequently, we calculated the posterior mean of θ_{iL}^k using Eq. 6, using the other data in X_{CN}.

$$\theta_{iL}^k = E\left(\frac{x_i + \alpha}{1+\beta}\right) \quad (L \in \{+1, -1\};\ k \in \{1, 2, \ldots, N\};\ i \in \{1, 2, \ldots, M\}) \tag{6}$$

where L denotes the cells from the positive (+1) or negative datasets (−1). Next, we calculated the prior probability of node V_i^k ($\psi(V_i^k)$ using Eq. 7.

$$\psi(V_i^k) = \begin{cases} \dfrac{\theta_{iL=+1}^k}{\theta_{iL=+1}^k + \theta_{iL=-1}^k},\ if\ L_i = +1 \\ \dfrac{\theta_{iL=-1}^k}{\theta_{iL=-1}^k + \theta_{iL=+1}^k},\ if\ L_i = -1 \end{cases} \quad (k \in \{1, 2, \ldots N\};\ i \in \{1, 2, \ldots, M\}) \tag{7}$$

Finally, we calculated the probability of occurrence of gene nodes in the eLBP network using Eq. 2.

2.3.4 Calculation of cell interaction score

After obtaining the posterior probability of each gene node, we obtained a simplified interaction network by trimming the gene nodes whose posterior probability obtained from Eq. 2 was <0.7. We then calculated the connection strength among the CIG pairs between cell types K_i and K_j (CGscore) using Eq. 8.

$$CG_{score_{K_iK_j}} = S(V_i^k)^* Co\sin\left(V_i^k{}^* V_j^k\right)^* S\left(V_j^k\right);\ k \in (1, N);\ i,j \in (1, M) \tag{8}$$

And the cell type interaction strength CT score was calculated using Eq. 9.

$$CT_{score_{ij}} = mean\left(CG_{score_{K_iK_j}}\right);\ k \in (1, N);\ i,j \in (1, M) \tag{9}$$

where K_i and K_j indicate different cell types. The highest CT score indicates potentially interacting cell types.

2.3.5 Construction of transcription factor -cell interaction signal transduction pathway

In addition to the cell interaction community, we calculated the transcription factor (TF) regulons that could enhance the critical cell-interaction genes in the eLBP model. To achieve this, we first obtained the potential regulatory networks of TFs by collating three transcription factor databases: ENCODE, ChEA3, and TRANSFAC. Next, we obtained the TFs whose target genes were significantly enriched in the trimmed eLBP network obtained from step 4 in the three TF datasets (ChEA3, ENCODE, and TRANSFAC, $p < 0.05$). Subsequently, we calculated the contribution score of each TF using Eq. 10:

$$TF_{score_q} = mean\left(S(V_i^k)^* Co\sin(V_i^k,\ TF_q)\right);\ k \in (1, N);\ i,j \in (1, M) \tag{10}$$

where V_i^k indicates the gene node i in cell type k in eLBP and TF_q, and $Co\sin(V_i^k, TF_q)$represents the cosine similarity between gene node V_i and TF_q. Ultimately, we obtained the strongest interactive cell types, important interactive gene pairs, and key TF regulons that could promote CIGs.

2.4 THP-1 cell culture and macrophage differentiation

Human leukemia monocytic THP-1 cells were cultured in Gibco™ RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS), 100 units/ml penicillin, and 100 g/ml streptomycin. THP-1 cells were seeded at a density of 6×10^5 cells/well into 24-well plates and treated with 200 ng/ml phorbol12-myristate13-acetate (PMA; Sigma-Aldrich, St. Louis, MO, United States) for 48 h. The cells were cultured at 37°C in a humidified atmosphere with 5% CO_2, where THP-1 monocytes were differentiated into macrophages.

2.5 Small interfering RNA transfection

Small interfering RNA (siRNAs) against *IRF1* and *SPI1* were used to validate the regulation of the expression of *LAGLS9* in THP-1 cells. *IRF1* and *SPI1* siRNAs were designed and synthesized by GenePharma (Shanghai, China). THP-1 cells were transfected with GP-transfect-Mate, according to the manufacturer's instructions. Briefly, THP-1 cells were seeded at a density of 6×10^5 cells/well in 24-well plates and treated with 200 ng/ml for 48 h. GP-transfect-Mate was diluted in Opti-MEM (50 ml; Invitrogen, Waltham, MA, United States) for 5 min before mixing with an equal volume of Opti-MEM containing siRNA (40 pmol). After 20 min of incubation, 100 μL of the resulting GP-transfect-Mate/siRNA mixture was added directly to the cells. After 24 h of incubation at 37°C in a 5% CO_2 atmosphere, cells were harvested for real-time PCR. The sequences of the primers used were as follows: SPI1-siRNA: CAG GCAGCAAGAAGAAGAUTT and AUCUUCUUCUUGCUG CCUT. IRF1-siRNA: GGGCUCAUCUGGAUUAAUATT, UAUUAAUCCAGAUGAGCCCTT.

2.6 RNA isolation and real-time PCR

Total RNA was isolated using TRIzol reagent (Sigma-Aldrich), according to the manufacturer's instructions. cDNA was synthesized using 5× All-in-One RT-Master Mix (Abm, Canada). Real-time PCR was performed in a LightCycler 96 using 2× chemQSYBR QPCR Master Mix (Vezyme, China),

according to the manufacturer's instructions, in a total volume of 10 μL. The sequences of the primers used were SPI1-Fw and GCGTGCAAAATGGAAG GGTTT. SPI1-Rev: GGTATCGAG GACGTGCATCT. IRF1-Fw, GCTGGGACATCAACAAGGAT. IRF1-Rev: CCTGCTCTGGTCTTTCACCT; LGALS9-Fw: AAG GTGATGGTGAACGGGAT. LGALS9-Rev: ACTGTCTGGGTA ATGGGAGC; CD68-Fw: TCCAGGGAAGCTGTGAGGGT. CD68-Rev: AGCCGAGAATGTCCACTGTGC. CD163-Fw: TTTGTCAACTTGAGTCCCTTCAC. CD163-Rev: TCCCGC TACACTTGTTTTCAC. VEGFA-Fw: TCCTCACACCATTGA AACCA. VEGFA-Rev: TTTTCTCTGCCTCCACAATG.

2.7 Identification of CCA subclasses

Genes from eLBP were selected for consistent clustering using the R software package ConsensusClusterPlus to sort the immune molecular CCA subtypes (Wilkerson & Hayes, 2010). Correlations between subtypes, clinical features, immunity, and prognosis were analyzed.

2.8 Multi-omics data acquisition and processing

The somatic mutation data of all patients with CCA was obtained from the International Cancer Genome Consortium (ICGC, https://dcc.icgc.org/). Mutations were analyzed and visualized using maftools (Mayakonda et al., 2018). The enrichment scores of the hallmark genes were evaluated through single-sample GSEA (ssGSEA) using the "GSVA" R package (Hänzelmann et al., 2013). Hallmark gene sets were obtained from MSigDB (https://www.gsea-msigdb.org/gsea/msigdb/).

2.9 Differentially expressed gene analysis

The "limma" package was used to perform differentially expressed gene (DEG) analysis. An empirical Bayesian method was applied to estimate the DEGs between two clusters, which were identified using a consistent clustering method based on moderated *t*-tests (Ritchie et al., 2015). The adjusted *p*-value for multiple testing was calculated using the Benjamini-Hochberg correction. Genes with an absolute $\log_2$ (fold change) greater than one and false discovery rate (FDR) < 0.05 were identified as DEGs between the two subtypes.

2.10 Estimation of immune infiltration and tumor purity

We downloaded the "CIBERSORT" scripts (https://cibersort.stanford.edu/) to estimate the immune composition of patients with CCA using a normalized express matrix (Chen et al., 2018). Immune, stromal, and tumor purity scores were calculated using the "estimate" R package (Yoshihara et al., 2013).

2.11 DNA methylation analysis

Differential DNA methylation analysis was conducted on both normal and tumor tissues, as well as on the two tumor subtypes. Data were pre-processed using the "minfi" package (Aryee et al., 2014). Hypomethylation probes with $\beta < 0.5$ hypo-methylation probes were used to conduct the analysis. We considered CpG probes to be hypo-methylated if they met the following criteria (Song You et al., 2018): $\beta < 0.5$ (Tirosh et al., 2016); *M*-value was significantly different between the two subtypes (q-value < 0.05); and (Stuart et al., 2019) mean β-value difference (Δβ) of >0.2 between the two groups. Copy number profiles were derived from the signal intensity values of methylated and unmethylated probes using the "conumee" package in R (http://bioconductor.org/packages/conumee). Copy number variations (CNVs) were derived from the log2-ratios of the tumor samples to the average value of matched normal tissues.

2.12 Statistical analysis

All computational and statistical analyses were performed using R software (https://www.r-project.org/). The unpaired Student's *t*-test was used to compare two groups with normally distributed variables, while the Mann-Whitney U-test was used to compare two groups with non-normally distributed variables. Survival analysis was performed using the Kaplan–Meier "survival" R package. The log-rank test was used to determine whether the survival curves were significantly different. A *p*-value < 0.05 was considered statistically significant.

2.13 Cox proportional hazards regression analysis

We used the least absolute shrinkage and selection operator (LASSO) method with 10-fold cross-validation and the Cox proportional hazards model with Akaike information criterion (AIC) selection criteria to build a nine-gene prognostic risk model based on the 54 TF-CIGs using the "glmnet" R package. The tumor mutation-burden-related signature (TMBRS) was calculated using Eq. 11:

$$\mathrm{TMBRS} = \sum (\boldsymbol{\beta}_i \times \mathrm{EXPi}) \tag{11}$$

where β_i can be derived from multivariate Cox analysis. The prognostic accuracy of the classifiers in the training and testing

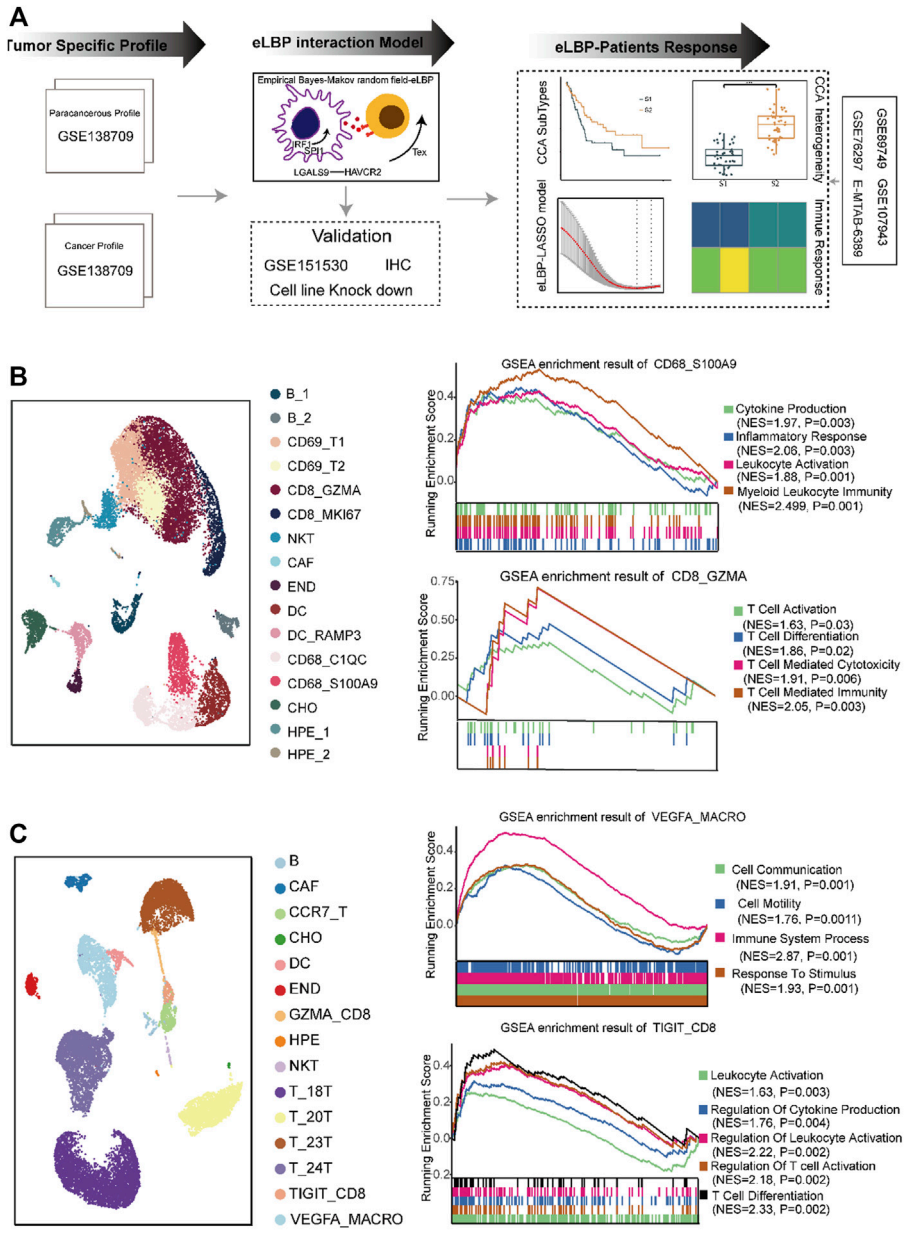


FIGURE 1
The distinct composition and function of cell types in the tumor microenvironment of cholangiocarcinoma. **(A)** Overview of the study design. We first analyzed distinct cell composition and function of the normal and cancer tissue in cholangiocarcinoma (CCA) single-cell RNAseq dataset (GSE138709). Next, we developed a method integrating empirical Bayes and Markov random field models (eLBP) which could simultaneously calculate transcription factors, interaction genes, and associated signaling pathways involved in cell-cell communication. We found that TFs *IRF1/SPI1* in *VEGFA*-positive macrophages could enhance the expression of *LGALS9* and induce the exhaustion of CD8 T cells *via LGALS9/HAVCR2* axis by eLBP method. Meanwhile, we obtained 54 genes involved in cell communication. The 54-gene panel could also group the CCA patients into two subtypes, in which patients in S2 showed high expression level of immune related genes and had better prognosis. Finally, we constructed a nine-gene eLBP-LASSO-COX risk model which was designated as tumor microenvironment risk score (TMRS). The TMRS panel was revealed to be a reliable tool for prognostic prediction and chemotherapeutic decision-making in CCA. **(B)** Uniform manifold approximation and projection (UMAP) plot showing the cell composition in normal. GSEA enrichment plot showing the function of macrophage subtypes CD68_S100A9 and CD8 T cell subtype CD8_GZMA. NES: Normalized enrichment score. **(C)** UAMP plot showing the cell composition in cancer. GSEA enrichment plot showing the function of macrophage subtypes VEGFA_MACRO and CD8 T cell subtype TIGIT_CD8. NES: Normalized enrichment scores.

sets was evaluated using the Kaplan–Meier curve and log-rank test.

2.14 Immunotherapeutic analysis

We predicted the IC_{50} values for the CCAs in GSE89749 *via* OncoPredict software using the 54 genes obtained from the eLBP algorithm (Maeser et al., 2021). In this step, we adopted the gene expression matrix from Genomics of Drug Sensitivity in Cancer (https://www.cancerrxgene.org/) and acquired drug sensitivity scores for the two subtypes. Next, we conducted a differential analysis of drugs between the two subtypes using the Wilcoxon rank-sum test. Drugs with p-values < 0.001 and drug sensitivity scores <1 were considered potential candidates.

3 Results

3.1 Distinct stromal cell compositions in adjacent and CCA tissues

A schematic of the study design is shown in Figure 1A. To investigate the composition of the TME in CCA, we reanalyzed the single-cell data from GSE138709 (Zhang et al., 2020). We deconstructed the cell types in tumors and adjacent tissues. We observed evident differences between the two tissues. We obtained 16 cell types in the adjacent tissues, including CD8_GZMA, CD8_MKI67, and CD69-positive T cells, M1-like macrophage subtypes (CD68_C1QC, CD68_S100A9), and two DC subtypes (Figure 1B; Supplementary Figure S1A). GSEA showed that cytokine production and inflammatory responses were enriched in M1-like macrophage subtypes (CD68_C1QC and CD68_S100A9). Additionally, T cell activation and T cell-mediated immunity pathways were enriched in CD8_GZMA cells (Figure 1B). In tumor tissues, we obtained three T cell subtypes, including memory-like T cells (*CCR7*+ and CCR7_T) and exhausted CD8 (*TIGIT*+ and TIGIT_CD8). We also obtained M2-like macrophages (*VEGFA*+ and VEGFA_MACRO; Figure 1C; Supplementary Figure S1A). GSEA results showed that pathways including cell communication, cell motility, and response to stimulus were enriched in VEGFA_MACRO, while leukocyte activation and T cell differentiation were enriched in the TIGIT_CD8 subtype (Figure 1C). Interestingly, the sample distribution results showed that T cells and macrophages were mostly from samples 23T and 24T. In contrast, the cells from 18T to 20T samples were almost all epithelial cells (Methods; Supplementary Table S2). On this basis, the tumor samples in this dataset were divided into high- and low-TIL tumor subtypes. These results indicate that T cells and macrophages perform different functions in tumor and para-cancerous tissues.

3.2 Exploring the interacting cell type community and critical transcript regulons in tumor and paracancerous tissues using the eLBP algorithm

To explore the mechanisms that induce the establishment of different immune states between normal and tumor tissues, we developed the eLBP algorithm to evaluate the correlation among immune-infiltrating, storm, and epithelial cells in normal and tumor single-cell data by calculating the intensity scores among the cell types (Figure 2A; Methods). First, we analyzed 7,071 activated cell type-specific pathways in normal tissues and 8,096 pathways in tumor tissues, in which 813 ligand or receptor genes in tumor tissues and 559 in normal tissues were computed based on the CellChat dataset (Supplementary Table S3). Second, we constructed protein-protein interaction (PPI) networks and calculated the cosine similarity between the PPI genes in each cell type. The degree of ligand/receptor (L/R) genes in the network is shown in Supplementary Figure S2A. The expression of ligand and receptor genes in CAF, CD68_C1QC, DC_RAMP3, and END was higher in normal tissues. T, 24T, TIGIT-CD8, and VEGFA_MACRO had higher degrees in the tumor tissues, indicating these cell types were in hyperactive status in cell communication among the TME (Supplementary Figure S2B). The levels of ligand and receptor genes indicate that macrophages with high levels of *VEGFA* play an important role in tumor tissues. Next, we used the eLBP algorithm to calculate the interaction probability of each ligand and receptor (Supplementary Table S3). Using this probability, we obtained the interaction intensity scores for each L/R pair among the cell types in the TME (Supplementary Tables S4, S5). The results demonstrated significantly different communities in tumors and normal tissues. In normal tissue, cells from DC_RAM3 and CD68_C1QC interacted with activated CD8 cells from CD8_GZMA according to NAMPT and chemokines, respectively (Figure 2B). In contrast, there were high levels of TIL infiltration in the tumor tissues. In addition, CAF could interact with memory-like T cells in CCR7_T, tumor cells in T_24T, and macrophages from VEGFA_MACRO. Moreover, VEGFA_MACRO interacted with exhausted CD8 + cells from TIGIT_CD8 (Figure 2C). The interaction genes between CAF and others are mainly related to cell adhesion, such as glycoprotein CD44 and collagen genes (Figure 2C). Meanwhile, VEGFA_MACRO, with a high level of LGALS9, interacted with *TIGIT* and *HAVCR2* (Figure 2C). These results illustrate that VEGFA_MACRO plays an important role in CD8 exhaustion.

Subsequently, we analyzed the key TFs that promote the expression of ligand genes in normal and cancerous tissues using the eLBP algorithm (Methods). In normal tissues, CD68_C1QC cells interacted with CD8_GZMA *via*

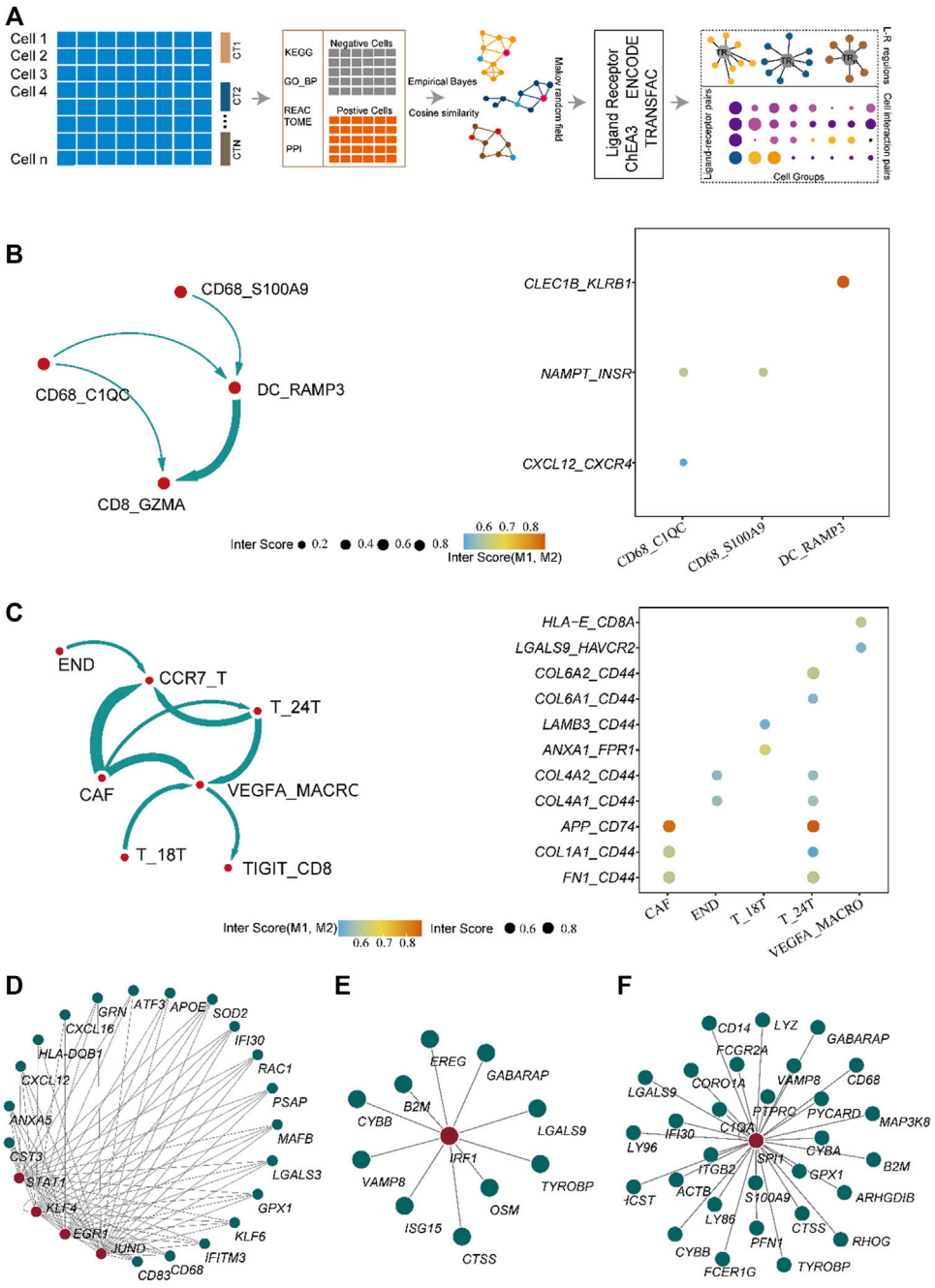


FIGURE 2
Algorithm framework of eLBP and the result calculated by eLBP. **(A)** Workflow of eLBP. Cell communication profile in the normal **(B)** (left) and cancer tissue **(C)**, in which line width was positive correlation with the cell interaction strength. Dot plot showing the interaction genes in the normal **(B)** and cancer tissue **(C)**, in which the row showing the interaction genes and the column showing the interaction cell types. The network showing the critical transcription factor (TF) regulons in normal **(D)** and cancer tissue **(E,F)**, in which red nodes were TFs and the other nodes were target genes.

CXCL12/CXCR4 (Figure 2B). We next obtained seven TFs (*STAT1*, *EGR1*, *KLF4*, *and JUND*; Figure 2D; Supplementary Table S6) from ENCODE and one TF (*KLF2*; Supplementary Table S6) from ChEA3. Moreover, *KLF4* and *STAT1* regulated the ligand gene *CXCL12* (Figure 2D), which may promote the expression of ligand genes and influence cell signaling. In

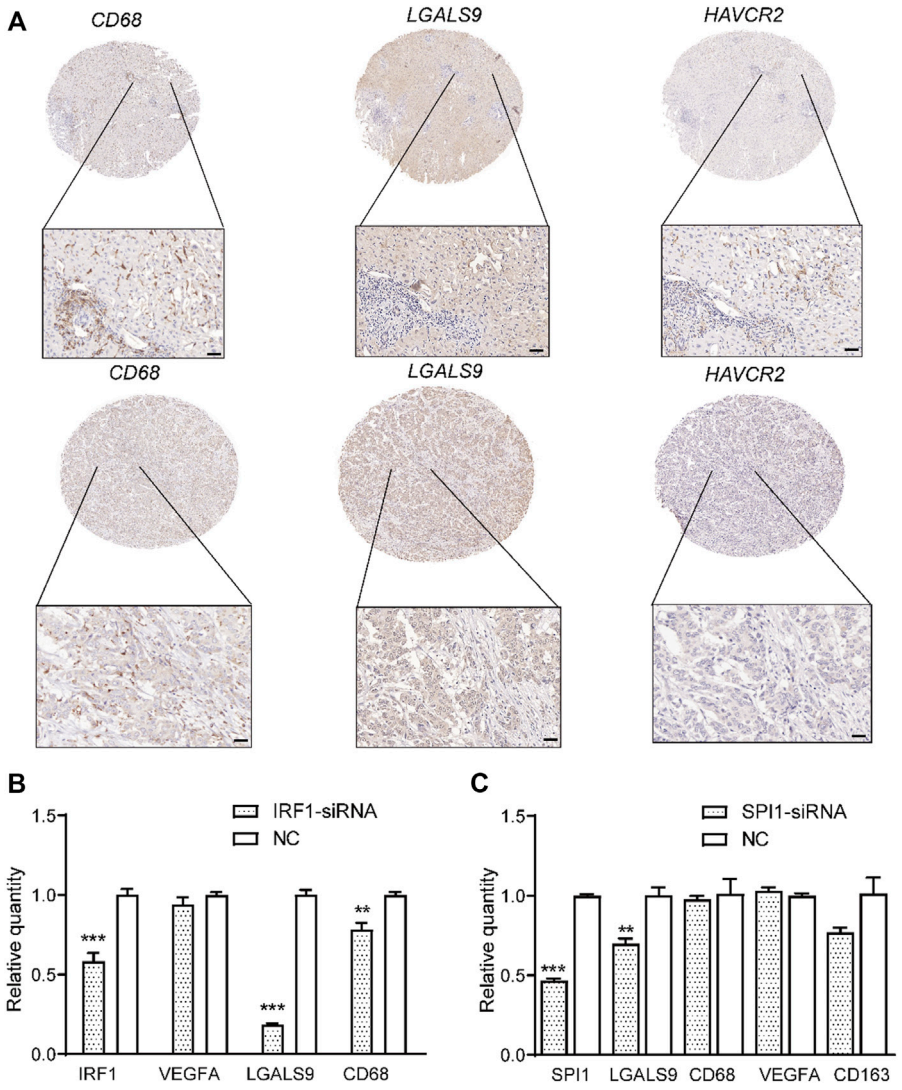


FIGURE 3
Validation of co-expression of *LGALS9* and *HAVCR2* and function of transcription factors *IRF1/SPI1*. **(A)** Immunohistochemical plot showing the co-expression of *LGALS9* and *HAVCR2* in two CCA patients. Scale bars represent 50 μm. Bar plot showing that TFs *IRF1* **(B)** and *SPI1* **(C)** could regulate the expression level of *LGALS9* (***$p < 0.001$; **$p < 0.01$, *t*-test).

addition, cells in DC_RAMP3 interact with those in CD8_GZMA *via CLEC1B/KLRB1* and TF *STAT3* in DC_RAMP3 could regulate *CLEC1B* (Supplementary Table S6). Notably, in the tumor tissues, we found that only the cells in VEGFA_MACRO interacted with those in TIGIT_CD8 *via LGALS9/HAVCR2* (Figure 2B). We then obtained 18 TF regulons from the three datasets, in which *IRF1* was enriched in both ENCODE and TRANSFAC, and *SPI1* was enriched in both ENCODE and ChEA3 (Supplementary Table S6). Moreover, these two TFs also regulated the ligand gene *LGALS9* (Figures 2E,F). In addition, the profiles of *LGALS9*, *HAVCR2*, *IRF1*, and *SPI1* in CCA patients from another single-cell dataset GSE151530 also verified their functions related to TAM and exhausted T cells (Supplementary Figure S3) (Ma et al., 2021). Moreover, immunohistochemical analysis further confirmed the high expression of *HAVCR2* in human CCA patients and the co-expression of *LGALS9* and these two genes (Figure 3A). The significantly low expression of *LGALS9* in siRNA-IRF1/SPI1 VEGFA-positive macrophages further validated the transcript regulon of SPI1/IRF1-LGALS9 (Figures 3B,C). Finally, we derived the gene trimmed-PPI network from the mutually connected cell types in which L/R genes participate in normal and tumor tissues *via* eLBP (Supplementary Figure S4).

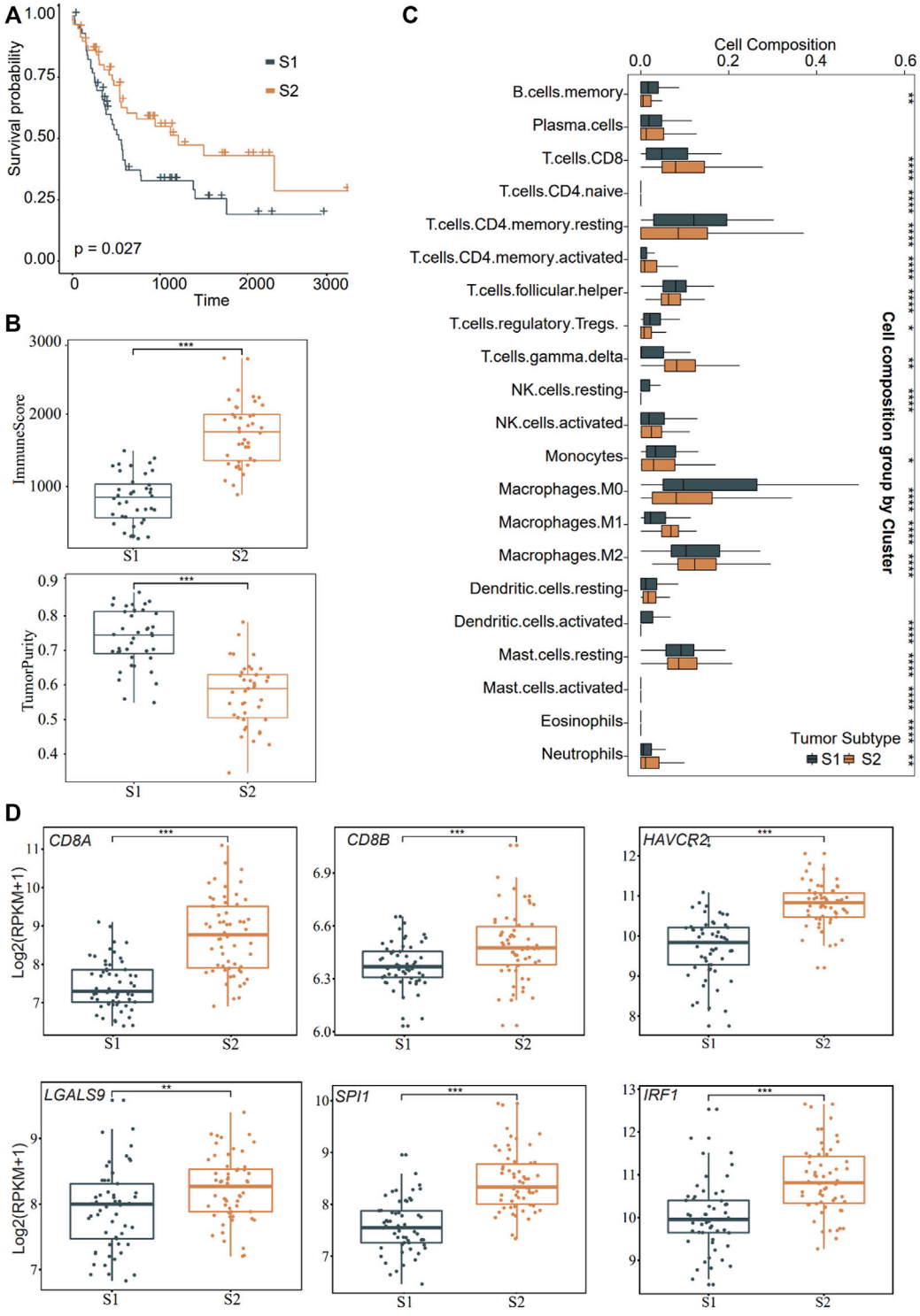


FIGURE 4
Distinct function in the two subtypes of CCA. **(A)** Overall survival curves showing the prognosis result of the two subtypes (S1 and S2) obtained from consensus clustering in CCA cohort (GSE89749) using the genes obtained from eLBP. Statistical significance was calculated using the log-rank test. **(B)** Box plots showing the immune and tumor purity scores in the two distinct malignant subtypes (***$p < 0.001$). Pairwise comparison was conducted by Wilcoxon rank sum test. **(C)** Box plots showing the 22 immune cell infiltrates ratio in the two distinct malignant subtypes in the significant enrich patients. (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; Wilcoxon rank sum test). **(D)** Comparisons of gene expression level of immune genes in the two distinct malignant subtypes (**$p < 0.01$; ***$p < 0.001$; Wilcoxon rank sum test). For the boxplot, the centerline, median; box limits, upper and lower quartiles. Each dot represents a sample.

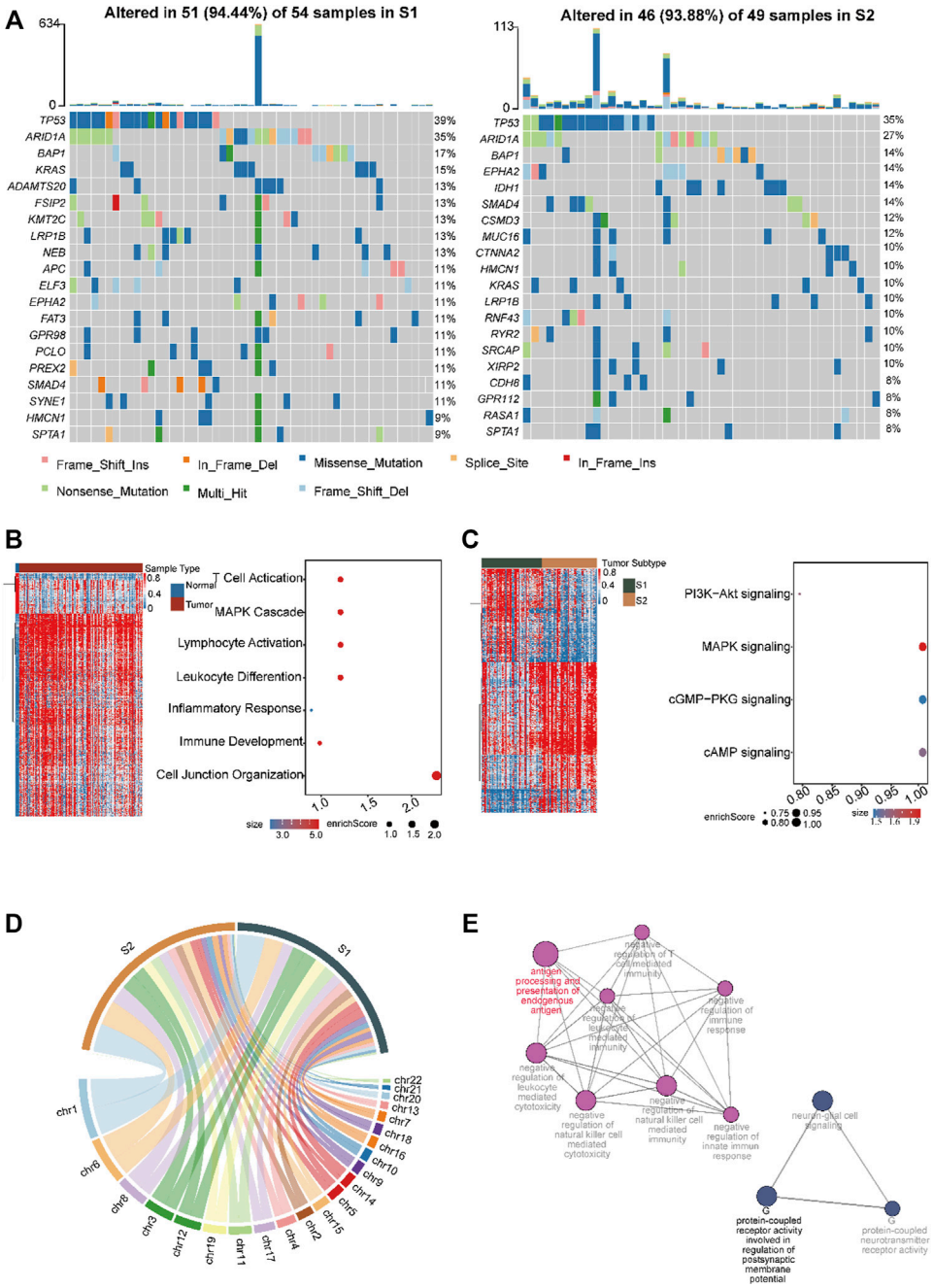


FIGURE 5
Investigation of the mutation profile inter-tumor heterogeneity profiles in the two CCA subtypes. **(A)** The top ten mutated genes across the two subtypes. The colors of rectangles in the body of the heat map indicate different types of somatic mutations and the key identifying each mutation type is shown at the bottom below the color bar. The bar plot on the top shows the counts of mutations for each patient and the colors in the bar plots correspond to the colors showing mutation types in the body of the heat map. The title of the heat map showing the mutation sample number in each subtype including amplification, missense mutations, and deep deletions. The left number showing the gene mutation frequency in the two subtypes. Heatmap and dot plot showing the differentially expressed genes and the significantly enriched pathways in normal **(B)** and cancer tissue **(C)** in CCA. Color gradient blue to red indicates relative expression levels from low to high. Dot size represent enrichment scores. **(D)** Circos plot showing the amplification and deletion regions in the two subtypes. The width of the plot indicated the CNV mutation regions numbers in each subtype across the chromosomes. **(E)** Network showing enriched pathways of hypo-methylated genes that occur on chromosomes 6 and 8 in S2. Dot size represents the enrichment scores.

3.3 Cell communication genes obtained from eLBP can be used to distinguish inflammatory subtype from a replicated one

Concerning the potential function of VEGFA_MACRO to induce the expression of exhausted genes such as *HAVCR2*, we further overlapped the genes in TF regulons with those in ligand or receptor-interaction trimmed-networks from eLBP from VEGFA_MACRO and obtained 54 genes (Supplementary Figure S4C; Supplementary Table S7). We then investigated the clinical features of patients with CCA using these genes. We grouped patients into different subtypes in two bulk RNA datasets using consistent clustering (Methods; GSE89749; GSE76297; Supplementary Table S1) (Chaisaingmongkol et al., 2017; Jusakul et al., 2017). Patients in GSE89749 were grouped into two subtypes (S1 and S2; Figure 4A; Supplementary Figure S5). Moreover, we found that the samples from the S2 group had a significantly better prognosis (p = 0.027; Figure 4A). Notably, the tumor purity results showed that patients in S2 had higher infiltrative immune and stromal scores, but lower tumor purity (Figure 4B). The immune infiltration results showed that patients in S2 had a higher ratio of CD8, activated memory CD4, S1, and M2-like macrophage subtypes (Figure 4C). In particular, *CD8A*, *CD8B*, *CD68*, *HAVCR2*, and *LGALS9* were highly expressed in S2 (Figure 4D). On this basis, S2 was classified as the inflammatory "hot tumor" subtype and S1 as the "cold tumor". We further identified the clustering results from other RNA sequencing datasets (GSE76297). The patients in S2 from GSE76297 had higher expression levels of *LGALS9*, *IRF1*, *SPI1*, and immune genes and similar TIL profiles with the GSE89749 data (Supplementary Figure S6). These results further confirmed the existence of high-TIL tumor subtypes in CCA and indicated the exhaustion of CD8 cells.

3.4 Inter-tumor heterogeneity between inflammatory and replicated subtypes

Next, we analyzed the heterogeneity of somatic frequencies across the two CCA subtypes. The results showed that the cell replication genes TP53 and genes in the PI3K/Akt pathway, such as *BAP1*, *KRAS*, *EPHA2*, *ARID1A*, and *SMAD4*, exhibited high mutation ratios in the two subtypes (Figure 4A). S1 had high ratios of *ADAMTS20*, *KMT2C*, and *APC* (Figure 5A, 13%, 13%, and 11%, respectively). S2 had a higher mutation frequency of *IDH1* and *MUC16* (Figure 5A, 14% and 12%, respectively). Tumor cells with TP53 mutations are generally identified as being more immunogenic (Chasov et al., 2020), and we found that immune genes, such as *CD3D* and *HAVCR2*, were highly expressed in the *MUC16* mutated group. Subsequently, we investigated the heterogeneity of methylation profiles between normal tissues and tumors. We obtained 15,520 differential probes between normal and tumor tissues, including 5,219 downregulated and 10, 321 upregulated probes (Methods, Figure 5B). In normal tissues, hypomethylated genes were mainly enriched in pathways such as T cell differentiation, cell adhesion, and inflammatory response, which were also enriched in normal single-cell datasets (Figure 5B). This indicated a high immune infiltration and activated immune status of the normal. Among the two CCA subtypes, we obtained 697 downregulated probes and 1,115 upregulated probes in S2. In addition, there were 684 downregulated hypomethylated probes enriched in 271 genes in S2. These genes were enriched in MAPK, CAMP, cGMP-PKG, and PI3K-AKT pathways (Figure 5C). We then conducted CNV analysis using the methylation data. The results showed that significant CNV signaling occurred across chromosomes 1, 6, and 8 in both subtypes (Figure 5D). Interestingly, we found that S2 had more CNV regions on chromosomes 6 and 8 (Figure 5D). We then analyzed the hypomethylated genes on chromosomes 6 and 8 in S2. These genes were enriched in antigen presentation and immune-toxic suppression, as well as T cell differentiation and G protein-coupled receptor-associated signaling pathways (Figure 5E).

3.5 Construct an eLBP-COX risk model to distinguish between inflammatory and replicated subtypes

LASSO-COX regression was used to construct the eLBP-COX TME risk score (TMRS) using the following formula in dataset GSE89749: TMRS = $-0.03085 \times IRF1 - 1.03491 \times RHOA + 0.5868 \times PLAUR - 0.2188 \times NCF2 - 0.2273 \times CXCR4 - 0.1201 \times HCK - 0.1773 \times LYZ - 0.1244 \times RGS1 + 0.2259 \times VCAN$ (Methods; Supplementary Figure S7). We found that patients with low TMRS had a better prognosis with higher immune scores and presented with *IRF1, LGALS9*, and other immune genes in GSE89749 (Figures 6A,B), indicating an effective indicator of RS score in prognostic risk prediction. Furthermore, we verified the RS model in two other datasets (EMTAB-6389; GSE107943 Supplementary Table S1; Methods; Figure 6A) (Ahn et al., 2019). Meanwhile, a nomogram was drawn to visualize the results of the eLBP-COX regression analysis (Figure 6C). Subsequently, we conducted drug susceptibility analysis and explored seven drugs (AZD7762, BI 2536, CDK9 5576, gemcitabine, mitoxantrone, teniposide, and topotecan) with IC50 values that were significantly lower in S2 than in S1 (Methods; Figure 6D). Specifically, BI 2536, teniposide, and mitoxantrone were used as potential therapeutic drugs for HCC that could be used in the treatment of CCA.

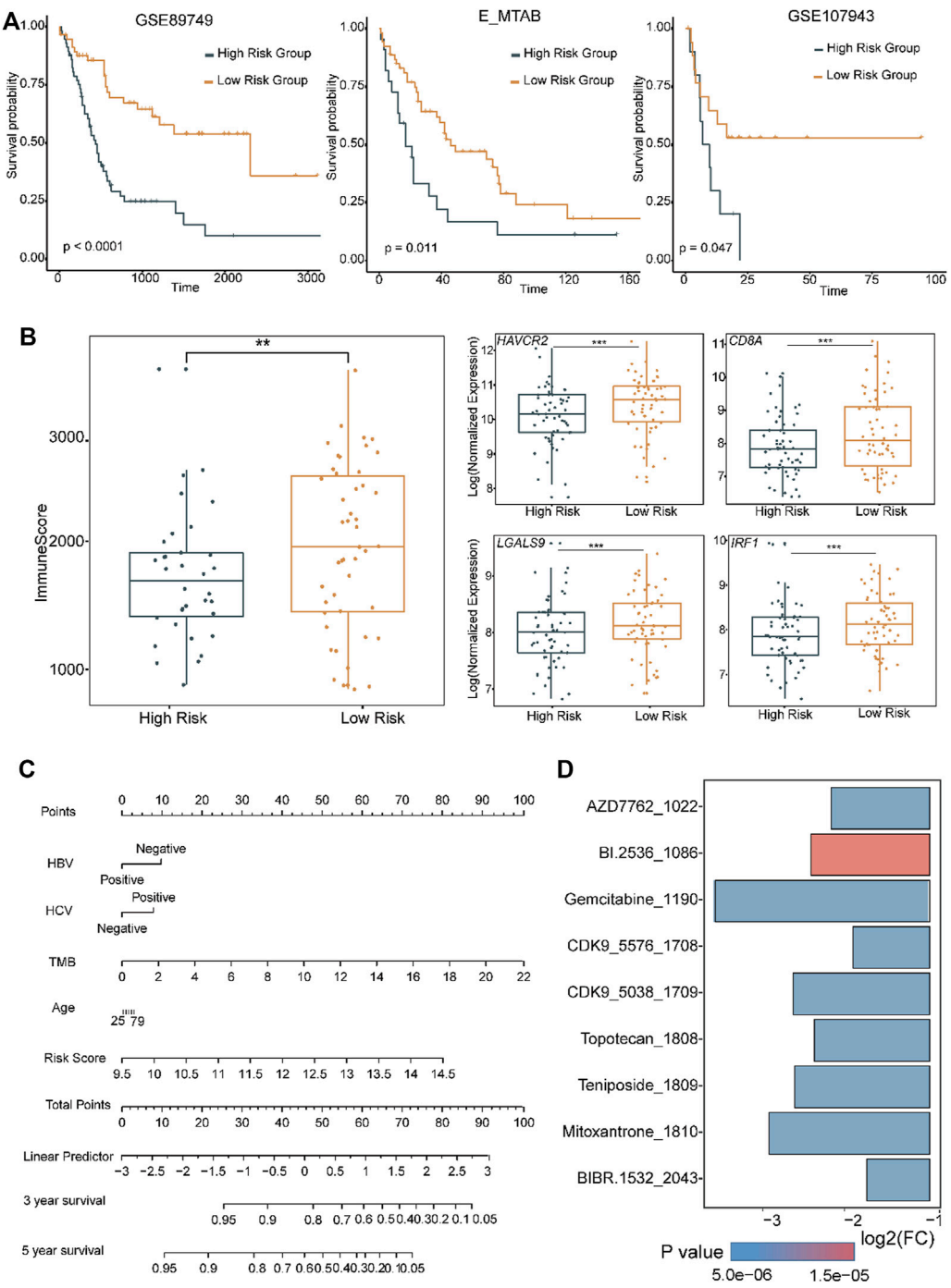


FIGURE 6
Distinct functions in different level of tumor microenvironment risk scores. **(A)** Overall survival curves showing the prognosis results with different level of tumor microenvironment risk score (TMRS) in the two CCA cohorts using the genes obtained from eLBP-LASSO. Statistical significance was calculated using the log-rank test. **(B)** Box plots showing the immune and the exhausted related genes in high and low risk group (**$p < 0.01$; ***$p < 0.001$; Wilcoxon rank sum test). **(C)** Nomogram plot showing the predicted 3- and 5-year survival possibilities of individual CCA patients. **(D)** Bar plot showing the drug sensitivity in S2 subtype.

TABLE 1 Resource identification initiative information.

Name	Categorization	Supplier	Cat no.
CD68	Protein	CST	76437T
LGALS9	Protein	Abcam	ab153673
HAVCR2	Protein	Abcam	ab153673

4 Discussion

Most patients with CCA are usually diagnosed at advanced stages and are not eligible for curative resection, with an overall survival (OS) of less than 1 year from the time of diagnosis. In the last decade, several novel therapeutic interventions and agents have improved the clinical outcomes of patients with CCAs. For example, radiofrequency ablation and mFOLFOX plus active symptom control (ASC) could significantly improve OS after progression on cisplatin-gemcitabine combination therapy compared to ASC alone (Rizzo et al., 2020). Multi-omics studies have revealed that FGFR2 fusions and rearrangements, IDH-1 mutations, and BRAF mutations are frequently found in patients with CCA. Targeted therapies such as the FGFR2 inhibitor pemigatinib and IDH1 inhibitor ivosidenib potentiate the treatment of CCAs harboring these mutations (Brandi et al., 2020). However, polyclonal mutations can also induce resistance to target drugs. Novel therapeutic strategies, such as immunotherapy, are needed, and immunotherapy is emerging as an important approach to cancer treatment. However, its efficacy varies greatly among cancer patients (Murciano-Goroff et al., 2020). CCA possesses a special TME that can be classified as either immune "hot" or "cold". Of these, immune "hot" tumors have a better response to immunotherapies (Loeuillard et al., 2019). ICIs represent a revolutionary milestone in the field of immuno-oncology (Darvin et al., 2018). The most common ICIs currently in clinical trials are PD-1/PD-L1 and *CTLA4* (Rizzo et al., 2021b). The response rates to ICIs vary owing to the complex composition of the TME (Yao and Gong, 2021). Therefore, we systematically investigated the composition of the TME to explore the critical interactions between cell types and signal transduction in CCAs.

First, we analyzed the TME composition of patients with CCAs using single-cell sequencing data. We found that activated CD8 cells and M1-like macrophages infiltrated normal tissues but exhausted CD8 cells in tumors. We further explored the critical elements in the signaling transduction pathways among CCA TMEs, such as the most frequently interacting cell types, the corresponding activated interaction gene pairs, the key transcription factors, and the gene regulons, by integrating L/R interactions with intracellular signaling *via* the Markov Random Fields model and developed an eLBP algorithm. We adopted the algorithm and found that macrophages with high expression levels of *LGALS9* could interact with exhausted CD8 T cells expressing high levels of *HAVCR2* in tumor tissues in CCAs. It is reported that galectin-9 (*LGALS9*) could interact with PD-1 and TIM-3 (*HAVCR2*) to regulate T cell death in multi-cell lines (Yang et al., 2021). Meanwhile, we first found that TFs, such as *IRF1* and *SPI1*, may promote the expression of *LGALS9* and verified their expression through multiple datasets and immunohistochemical staining in CCA tissues. It has been reported to *IRF1* inhibits antitumor immunity through upregulation of PD-L1 in MC38 and CT26 colon cancer and B16 melanoma mouse models (Shao et al., 2019). It has been demonstrated that *Spi1* is required for myeloid-specific expression of *Lgals9* in zebrafish (Zakrzewska et al., 2010). However, their function in CCAs merits further investigation.

Second, eLBP provided a gene classifier consisting of 54 genes that dominated the signal transduction between macrophages and exhausted CD8 T cells. This classifier can group CCA patients into replicated subtypes S1 and S2. Interestingly, patients with the inflammatory subtype S2 with high levels of *LGALS9* and *HAVCR2* had a better prognosis. Meanwhile, a higher mutation ratio of *MUC16* in S2 was positive for immune-exhausted genes, such as *HAVCR2* and *TIGIT*. Recent studies have demonstrated that *MUC16* mutations are associated with better survival outcomes and immune responses in gastric and endometrial cancers (Wang et al., 2020).

Pathological confirmation of the diagnosis is necessary before any nonsurgical treatment but is challenging in CCA therapy. Endoscopic imaging and tissue sampling are useful; however, biopsy samples are often inadequate for molecular profiling. Tissue sampling is reported to be highly specific but has low sensitivity in the diagnosis of malignant biliary strictures. Liquid biopsy can capture and monitor the tumor genetic profile or response to therapy in real time. For example, high mutation rates of AR1D1A, FGFR2, PIK3CA, and TP53 proteins, high levels of microRNAs such as miR-21, and high levels of proteins and cytokines such as CK-19, MMP-7, osteopontin, periostin, and IL-6 have been found in CCA patient serum (Rompianesi, et al., 2021). These protein and microRNA markers can serve as prognostic and therapeutic markers. In this study, we set up a 54-gene panel to screen or predict patients that are suitable for immunotherapy and finally constructed a nine-gene eLBP-LASSO-COX risk model. From this model, we found that *IRF1* was positively associated with low risk, while *PLAUR* and *VCAN* were positively associated with a high-risk score in tumorigenesis of CCAs. It has been reported that *PLAUR* overexpression correlates with poor prognosis in HCC (Woo et al., 2008). Glioma patients with higher *PLAUR* expression are infiltrated with fewer CD8 T cells (Zeng et al., 2021). *VCAN* has been used as an immune exclusion marker (Emmerich et al., 2020). To facilitate clinical

application, we will further investigate their expression levels in serum and exosomes and their capacity for the early screening of CCAs.

5 Permission to reuse and copyright

Figures, tables, and images will be published under a Creative Commons CC-BY licence and permission must be obtained for use of copyrighted material from other sources (including re-published/adapted/modified/partial figures and images from the internet). It is the responsibility of the authors to acquire the licenses, to follow any citation instructions requested by third-party rights holders, and cover any supplementary charges.

6 Nomenclature

6.1 Resource identification initiative

To take part in the Resource Identification Initiative, the corresponding catalogue numbers were listed in Table 1.

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

KY designed the study, and TW performed the data analysis. TW wrote the manuscript. CX, DX, YL, XL, ZL, LL, and ZZ performed the immunohistochemical and cell culture experiments, ND modified the pictures, and KY, CX and YL revised the manuscript.

Funding

This work was supported by the National Key R&D Program of China (2022YFC3400300), National Natural Science Foundation of China grant (32125009 and 32070663 for KY, 32070134 for DX, 61702406 for XY, 82273059 and 82073091 for CX, 81872253 for LL), and by Key Construction Program of the National "985" Project, by the Fundamental Research Funds for the Central Universities (xzy012020012).

Acknowledgments

We would like to thank the comments from Mengyun Ke and other colleagues in the lab.

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/fgene.2022.981145/full#supplementary-material

References

Ahn, K. S., O'Brien, D., Kang, Y. N., Mounajjed, T., Kim, Y. H., Kim, T. S., et al. (2019). Prognostic subclass of intrahepatic cholangiocarcinoma by integrative molecular-clinical analysis and potential targeted approach. *Hepatol. Int.* 13 (4), 490–500. doi:10.1007/s12072-019-09954-3

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., et al. (2014). Minfi: A flexible and comprehensive bioconductor package for the analysis of infinium DNA methylation microarrays. *Bioinformatics* 30 (10), 1363–1369. doi:10.1093/bioinformatics/btu049

Banales, J. M., Marin, J. J. G., Lamarca, A., Rodrigues, P. M., Khan, S. A., Roberts, L. R., et al. (2020). Cholangiocarcinoma 2020: The next horizon in mechanisms and management. *Nat. Rev. Gastroenterol. Hepatol.* 17 (9), 557–588. doi:10.1038/s41575-020-0310-z

Brandi, G., Rizzo, A., Dall'Olio, F. G., Felicani, C., Ercolani, G., Cescon, M., et al. (2020). Percutaneous radiofrequency ablation in intrahepatic cholangiocarcinoma: A retrospective single-center experience. *Int. J. Hyperth.* 37 (1), 479–485. doi:10.1080/02656736.2020.1763484

Bridgewater, J., Galle, P. R., Khan, S. A., Llovet, J. M., Park, J. W., Patel, T., et al. (2014). Guidelines for the diagnosis and management of intrahepatic cholangiocarcinoma. *J. Hepatol.* 60 (6), 1268–1289. doi:10.1016/j.jhep.2014.01.021

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 36 (5), 411–420. doi:10.1038/nbt.4096

Chaisaingmongkol, J., Budhu, A., Dang, H., Rabibhadana, S., Pupacdi, B., Kwon, S. M., et al. (2017). Common molecular subtypes among asian hepatocellular carcinoma and cholangiocarcinoma. *Cancer Cell* 32 (1), 57–70. e53. doi:10.1016/j.ccell.2017.05.009

Chasov, V., Mirgayazova, R., Zmievskaya, E., Khadiullina, R., Valiullina, A., Stephenson Clarke, J., et al. (2020). Key players in the mutant p53 team: Small molecules, gene editing, immunotherapy. *Front. Oncol.* 10, 1460. doi:10.3389/fonc.2020.01460

Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M., and Alizadeh, A. A. (2018). Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* 1711, 243–259. doi:10.1007/978-1-4939-7493-1_12

Darvin, P., Toor, S. M., Sasidharan Nair, V., and Elkord, E. (2018). Immune checkpoint inhibitors: Recent progress and potential biomarkers. *Exp. Mol. Med.* 50 (12), 1–11. doi:10.1038/s12276-018-0191-1

Emmerich, P., Matkowskyj, K. A., McGregor, S., Kraus, S., Bischel, K., Qyli, T., et al. (2020). VCAN accumulation and proteolysis as predictors of T lymphocyte-excluded and permissive tumor microenvironments. *J. Clin. Oncol.* 38 (15), 3127. doi:10.1200/JCO.2020.38.15_suppl.3127

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). Gsva: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14 (1), 7–15. doi:10.1186/1471-2105-14-7

Hegde, P. S., and Chen, D. S. (2020). Top 10 challenges in cancer immunotherapy. *Immunity* 52 (1), 17–35. doi:10.1016/j.immuni.2019.12.011

Jiang, Y., Zhao, X., Fu, J., and Wang, H. (2020). Progress and challenges in precise treatment of tumors with PD-1/PD-L1 blockade. *Front. Immunol.* 11, 339. doi:10.3389/fimmu.2020.00339

Jin, S., Guerrero-Juarez, C. F., Zhang, L., Chang, I., Ramos, R., Kuan, C. H., et al. (2021). Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* 12 (1), 1088. doi:10.1038/s41467-021-21246-9

Jusakul, A., Cutcutache, I., Yong, C. H., Lim, J. Q., Huang, M. N., Padmanabhan, N., et al. (2017). Whole-genome and epigenomic landscapes of etiologically distinct subtypes of cholangiocarcinoma. *Cancer Discov.* 7 (10), 1116–1135. doi:10.1158/2159-8290.CD-17-0368

Loeuillard, E., Conboy, C. B., Gores, G. J., and Rizvi, S. (2019). Immunobiology of cholangiocarcinoma. *JHEP Rep.* 1 (4), 297–311. doi:10.1016/j.jhepr.2019.06.003

Ma, L., Wang, L., Khatib, S. A., Chang, C. W., Heinrich, S., Dominguez, D. A., et al. (2021). Single-cell atlas of tumor cell evolution in response to therapy in hepatocellular carcinoma and intrahepatic cholangiocarcinoma. *J. Hepatol.* 75, 1397–1408. doi:10.1016/j.jhep.2021.06.028

Maeser, D., Gruener, R. F., and Huang, R. S. (2021). oncoPredict: an R package for predicting *in vivo* or cancer patient drug response and biomarkers from cell line screening data. *Brief. Bioinform.* 22, bbab260. doi:10.1093/bib/bbab260

Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28 (11), 1747–1756. doi:10.1101/gr.239244.118

Murciano-Goroff, Y. R., Warner, A. B., and Wolchok, J. D. (2020). The future of cancer immunotherapy: Microenvironment-targeting combinations. *Cell Res.* 30 (6), 507–519. doi:10.1038/s41422-020-0337-2

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Rizzo, A., Ricci, A. D., and Brandi, G. (2021a). Durvalumab: An investigational anti-PD-L1 antibody for the treatment of biliary tract cancer. *Expert Opin. Investig. Drugs* 30 (4), 343–350. doi:10.1080/13543784.2021.1897102

Rizzo, A., Ricci, A. D., and Brandi, G. (2021b). Pemigatinib: Hot topics behind the first approval of a targeted therapy in cholangiocarcinoma. *Cancer Treat. Res. Commun.* 27, 100337. doi:10.1016/j.ctarc.2021.100337

Rizzo, A., Ricci, A. D., Tober, N., Nigro, M. C., Mosca, M., Palloni, A., et al. (2020). Second-line treatment in advanced biliary tract cancer: Today and tomorrow. *Anticancer Res.* 40 (6), 3013–3030. doi:10.21873/anticanres.14282

Rompianesi, G., Di Martino, M., Gordon-Weeks, A., Montalti, R., and Troisi, R. (2021). Liquid biopsy in cholangiocarcinoma: Current status and future perspectives. *World J. Gastrointest. Oncol.* 13 (5), 332–350. doi:10.4251/wjgo.v13.i5.332

Shao, L., Hou, W., Scharping, N. E., Vendetti, F. P., Srivastava, R., Roy, C. N., et al. (2019). IRF1 inhibits antitumor immunity through the upregulation of PD-L1 in the tumor cell. *Cancer Immunol. Res.* 7 (8), 1258–1266. doi:10.1158/2326-6066.CIR-18-0711

Song You, F. W., Hu, Q., Li, P., Zhang, C., Yu, Y., Zhang, Y., et al. (2018). Abnormal expression of YEATS4 associates with poor prognosis and promotes cell proliferation of hepatic carcinoma cell by regulation the TCEA1/DDX3 axis. *Am. J. Cancer Res.* 8 (10), 2076–2087.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., 3rd, et al. (2019). Comprehensive integration of single-cell data. *Cell* 177 (7), 1888–1902. doi:10.1016/j.cell.2019.05.031

Tang, T., Huang, X., Zhang, G., Hong, Z., Bai, X., and Liang, T. (2021). Advantages of targeting the tumor immune microenvironment over blocking immune checkpoint in cancer immunotherapy. *Signal Transduct. Target. Ther.* 6 (1), 72. doi:10.1038/s41392-020-00449-4

Tirosh, I., Izar, B., Prakadan, S. M., Wadsworth, M. H., 2nd, Treacy, D., Trombetta, J. J., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 352 (6282), 189–196. doi:10.1126/science.aad0501

Tohonen, V., Katayama, S., Vesterlund, L., Jouhilahti, E. M., Sheikhi, M., Madissoon, E., et al. (2015). Novel PRD-like homeodomain transcription factors and retrotransposon elements in early human development. *Nat. Commun.* 6, 8207. doi:10.1038/ncomms9207

Wang, H., Yan, C., and Ye, H. (2020). Overexpression of MUC16 predicts favourable prognosis in MUC16-mutant cervical cancer related to immune response. *Exp. Ther. Med.* 20 (2), 1725–1733. doi:10.3892/etm.2020.8836

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170

Woo, H. G., Park, E. S., Cheon, J. H., Kim, J. H., Lee, J. S., Park, B. J., et al. (2008). Gene expression-based recurrence prediction of Hepatitis B virus-related human hepatocellular carcinoma. *Clin. Cancer Res.* 14 (7), 2056–2064. doi:10.1158/1078-0432.CCR-07-1473

Yang, R., Sun, L., Li, C. F., Wang, Y. H., Yao, J., Li, H., et al. (2021). Galectin-9 interacts with PD-1 and TIM-3 to regulate T cell death and is a target for cancer immunotherapy. *Nat. Commun.* 12 (1), 832. doi:10.1038/s41467-021-21099-2

Yao, W.-Y., and Gong, W. (2021). Immunotherapy in cholangiocarcinoma: From concept to clinical trials. *Surg. Pract. Sci.* 5, 100028. doi:10.1016/j.sipas.2021.100028

Yoshihara, K., Shahmoradgoli, M., Martinez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612

Zakrzewska, A., Cui, C., Stockhammer, O. W., Benard, E. L., Spaink, H. P., and Meijer, A. H. (2010). Macrophage-specific gene functions in Spi1-directed innate immunity. *Blood* 116 (3), e1–e11. doi:10.1182/blood-2010-01-262873

Zeng, F., Li, G., Liu, X., Zhang, K., Huang, H., Jiang, T., et al. (2021). Plasminogen activator urokinase receptor implies immunosuppressive features and acts as an unfavorable prognostic biomarker in glioma. *Oncologist* 26 (8), e1460–e1469. doi:10.1002/onco.13750

Zhang, M., Yang, H., Wan, L., Wang, Z., Wang, H., Ge, C., et al. (2020). Single-cell transcriptomic architecture and intercellular crosstalk of human intrahepatic cholangiocarcinoma. *J. Hepatol.* 73, 1118–1130. doi:10.1016/j.jhep.2020.05.039

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 09 December 2022
DOI 10.3389/fgene.2022.1035368

Check for updates

OPEN ACCESS

EDITED BY
Ehsan Nazemalhosseini-Mojarad,
Shahid Beheshti University of Medical Sciences, Iran

REVIEWED BY
Mengya Zhong,
Xiamen University, China
Pranjal Sarma,
University of Cincinnati, United States
Yang Jiang,
Tongji University, China

*CORRESPONDENCE
Hongzhe Li,
lihongzhe1965@163.com

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 02 September 2022
ACCEPTED 28 November 2022
PUBLISHED 09 December 2022

CITATION
Xuan F, Zhang Z, Liu K, Gong H, Liang S, Zhao Y and Li H (2022), Constructing a signature based on the SIRT family to help the prognosis and treatment sensitivity in glioma patients. *Front. Genet.* 13:1035368. doi: 10.3389/fgene.2022.1035368

COPYRIGHT
© 2022 Xuan, Zhang, Liu, Gong, Liang, Zhao and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Constructing a signature based on the SIRT family to help the prognosis and treatment sensitivity in glioma patients

Feiyue Xuan[1], Zhiwei Zhang[1], Kuili Liu[1], Haidong Gong[1,2], Shaodong Liang[1], Youzhi Zhao[1] and Hongzhe Li[1]*

[1]Department of Neurosurgery, Hongqi Hospital Affiliated to Mudanjiang Medical University, Mudanjiang, China, [2]Heilongjiang Provincial Key Laboratory of Cancer Disease Prevention and Control, Mudanjiang Medical University, Mudanjiang, China

Enzymes of the silent information regulator (SIRT) family exert crucial roles in basic cellular physiological processes including apoptosis, metabolism, ageing, and cell cycle progression. They critically contribute to promoting or inhibiting cancers such as glioma. In the present study, a new gene signature of this family was identified for use in risk assessment and stratification of glioma patients. To this end, the transcriptome and relevant clinical records of patients diagnosed with glioma were obtained from the Cancer Genomic Atlas (TCGA) and the Chinese Glioma Genome Atlas (CGGA). LASSO regression and multivariate Cox analyses were used to establish the signature. Using Kaplan–Meier analyses, overall survival (OS) was assessed and compared between a training and an external test datasets which showed lower OS in patients with high risk of glioma compared to those with low risk. Further, ROC curve analyses indicated that the SIRT-based signature had the desired accuracy and universality for evaluating the prognosis of glioma patients. Using univariate and multivariate Cox regression analyses, the SIRT-based signature was confirmed as an independent prognostic factor applicable to subjects in the TCGA and CGGA databases. We also developed an OS nomogram including gender, age, risk score, pathological grade, and IDH status for clinical decision-making purposes. ssGSEA analysis showed a higher score for various immune subgroups (e.g., $CD8^+$ T cells, DC, and TIL) in samples from high-risk patients, compared to those of low-risk ones. qPCR and western blotting confirmed the dysregulated expression of SIRTs in gliomas. Taken together, we developed a new signature on the basis of five SIRT family genes, which can help accurately predict OS of glioma patients. In addition, the findings of the present study suggest that this characteristic is associated with differences in immune status and infiltration levels of various immune cells in the tumor microenvironment.

KEYWORDS

glioma, SIRT family, biomarker, prognosis, tumour microenvironment

Introduction

Among the various brain tumors, glioma is the most frequent malignancy, which occurs in adults and accounts for almost one-third of primary brain tumors in people aged 20 59 (Jiang et al., 2016). Because of its extensive proliferation, invasion, immunosuppression, and resistance to chemotherapy, clinical treatment of glioma is notoriously difficult (Jiang et al., 2021; Qi et al., 2022). The median survival time of patients diagnosed with low-grade glioma is approximately 6 years. Compared with low-grade gliomas (LGG, WHO grades II and III), glioblastoma (GBM, WHO grade IV) is more often fatal (Ostrom et al., 2021). The median survival time in glioblastoma is typically less than 2 years, even when implementing large-scale surgery, focal resection, and combined radiotherapy and chemotherapy (Linz, 2009). Thus, there is an urgent need for new biomarkers to facilitate accurate diagnosis and to improve the currently limited therapeutic options.

Silent information regulators (SIRTs), or NAD-dependent deacetylase sirtuins, are class-III histone deacetylases responsible for the unique β-nicotinamide adenine dinucleotide (β-NAD+)-dependent Nε-acyl-lysine deacylation of both histone and non-histone substrates (Teixeira et al., 2020). Sirtuin proteins are highly conserved and occur as seven subtypes in humans, i.e., SIRT1 to SIRT7 (Teixeira et al., 2020). Sirtuins contribute to the regulation of various biological processes, including apoptosis, metabolism, ageing, and cell cycle progression (Brown et al., 2013; Li et al., 2013). Given that, sirtuins are considered to be closely associated with the onset and progression of tumors (Brown et al., 2013; Li et al., 2013), and increasing evidence confirms the importance of the SIRT family in glioma, as these proteins may promote or inhibit cancer progression (Chen et al., 2018; Funato et al., 2018; Luo et al., 2018; Mu et al., 2019; Yalçın and Colak, 2020; Liu et al., 2022). However, most previous studies focused on exploring the biological functions of SIRT family members in gliomas, whereas their prognostic value for glioma patients has been overlooked.

Using comprehensive bioinformatics, we performed data-mining through transcriptional and clinical records of patients diagnosed with glioma using two databases procured from the Cancer Genomic Atlas (TCGA) and the Chinese Glioma Genome Atlas (CGGA), respectively. We further investigated the role of SIRT family members in the diagnosis, prognosis, and treatment of gliomas through clinical sample validation. We present the following article in accordance with the TRIPOD reporting checklist.

Materials and methods

Data collection

The sequenced transcriptome and clinical data of 689 patients diagnosed with glioma were downloaded from the TCGA-GBM and TCGA-LGG projects using the UCSC Xena database (http://www.genome.ucsc.edu/index.html). Further, normal brain tissue samples composed of paracancerous tissues in TCGA-GBM (n = 5) and GTEX (n = 1,152) were retrieved from the UCSC Xena database. The retrieved transcriptome records were converted to TPM format for subsequent analysis as a training data set. For subsequent validation, the transcriptome and clinical information of the validation dataset (CGGA, n = 325) were also downloaded from the CGGA database (http://www.cgga.org.cn/index.jsp).

From September 2021 to March 2022, three cases of glioma tissues and their relative paracancerous tissue were extracted from clinical subjects who were schedules to receive initial surgery in Hongqi Hospital Affiliated with Mudanjiang Medical College. This project was performed under the approval of the medical ethics committee of Hongqi Hospital Affiliated with Mudanjiang Medical College. All participants signed written informed consent. Information about the participants is provided in Supplement 1.

Differential analysis of the SIRT family in gliomas

We first used high-throughput sequencing data from TCGA and GTEx to investigate the gene expression levels of SIRT family members in the glioma samples and the respective adjacent normal brain tissues. Subsequently, differential expression of SIRT family genes was compared between the glioma samples and adjacent normal brain tissues, across the Human Protein Atlas (HPA; https://www.proteinatlas.org). Then, the potential of this gene family for differentiating glioma tissues from normal brain tissues was evaluated using ROC curves. Higher AUC values indicated better differentiation ability.

cBioPortal database

The open-source database cBioPortal (http://www.cbioportal.org) is designed for exploring the genetic changes of a gene in cancer, including its mutations, copy number variation, and relationship with clinical features (Gao et al., 2013). cBioPortal was thus used to assess the genetic changes of SIRT genes and their prognostic value in glioma patients.

Survival analyses

The prognostic indicators of this study included overall survival (OS), disease-specific survival (DSS), and the progression-free interval (PFI). SIRT genes were categorized as either highly expressed (HE) or low expressed (LE), according to

their median mRNA abundance. The prognostic differences between the HE and LE groups were compared by Kaplan–Meier survival analyses. Univariate and multivariate Cox analyses were performed to evaluate the potential of SIRT family genes as independent prognostic factors in patients diagnosed with glioma.

Protein-protein interaction (PPI) network

GeneMANIA database (http://genemania.org) was used for exploring the functions and interactions between the targeted genes (Franz et al., 2018). We explored the proteins interacting with SIRT genes in humans and their functions and used them to construct a PPI network. GeneMANIA is an online database that contains 660,554,667 interactions and 166,691 genes of nine species.

Functional enrichment analysis

The Metascape online tool (https://metascape.org/gp/index.html#/main/step1) provides comprehensive gene annotations (Zhou et al., 2019). In this study, we used Metascape to perform functional enrichment analysis and visualization of the SIRT family and its interacting genes. The "clusterProfiler" package in R software version 4.1.1 was applied for gene set enrichment analyses (GSEA) using the C2 KEGG database, with $p < 0.05$ and $q < 0.25$ considered significant.

Constructing a SIRT-based signature

To create a signature based on the SIRT family in cases with glioma, the TCGA queue was used as a train dataset. In the train dataset, SIRT family genes were subjected to LASSO regression analysis using the R software packages "glmnet" and "survival", and overfitted genes were excluded. Multivariate Cox regression analysis was used to determine the most significantly and valuable prognostic genes in the SIRT family. Then, the corresponding regression coefficients and risk scores of glioma patients were calculated using the following equation:

$$\begin{aligned}\text{Risk score} = {} & \text{SIRT1} \times (-0.961) + \text{SIRT2} \times (-0.318) \\ & + \text{SIRT5} \times (-0.473) + \text{SIRT6} \times (-0.239) \\ & + \text{SIRT7} \times (0.784)\end{aligned}$$

Subsequently, the CGGA data were used as a test dataset to determine the risk score of patients diagnosed with glioma in the CGGA database. Subjects were further stratified in high-risk (HR) and low-risk (LR) groups, depending on their median risk scores. Kaplan–Meier survival analysis was applied to assess the prognosis difference between the patients form groups with high risk and low risk of glioma. The performance of the SIRT-based signature was evaluated using a ROC curve, and the independent prediction ability of the SIRT-based signature was determined through univariate and multivariate Cox analyses.

Immune infiltration analysis

According to the transcriptome data, infiltration of various immune cells and immune status scores of glioma patients in the TCGA data set were evaluated using a single-sample GSEA (ssGSEA) algorithm. Then, the immune cell content and immune status scores were compared between patients of the HR and LR groups.

Drug sensitivity analysis

Microsatellite installation (MSI) and tumor immune dysfunction and exclusion (TIDE) algorithms can be used to predict the sensitivity of glioma patients to immunotherapy. The TIDE score is negatively correlated with the sensitivity to immunotherapy (Jiang et al., 2018). The sensitivity of patients in the HR and LR groups to common chemotherapy drugs for gliomas was predicted through the R software package "pRRophetic". The sensitivity is reflected by the IC50 of the drug which has a negative correlation with drug sensitivity.

Nomogram

The R software package "rms" was used for integrating the records of survival time, survival status, SIRT-based signature, age, gender, pathological grade, and IDH status. A nomogram was established using a Cox test to evaluate the prognostic significance of these characteristics in TCGA samples, and the efficacy of the nomogram was evaluated using calibration and ROC curves.

qPCR

Given the importance of SIRT1 and SIRT5 in glioma, we verified their expression using qPCR. RNAiso Plus (1 ml; Takara Bio, Japan) was used to isolate RNA, which was then reverse-transcribed using a PrimeScript™ RT reagent Kit (Takara Bio) according to the manufacturer's instructions. qPCR was performed to quantify RNA levels using Hieff™ qPCR SYBR Green Master Mix (Yeasen Biotechnology, China) and the primers listed in Supplementary Material S1. ACTB was used as a standardized control.

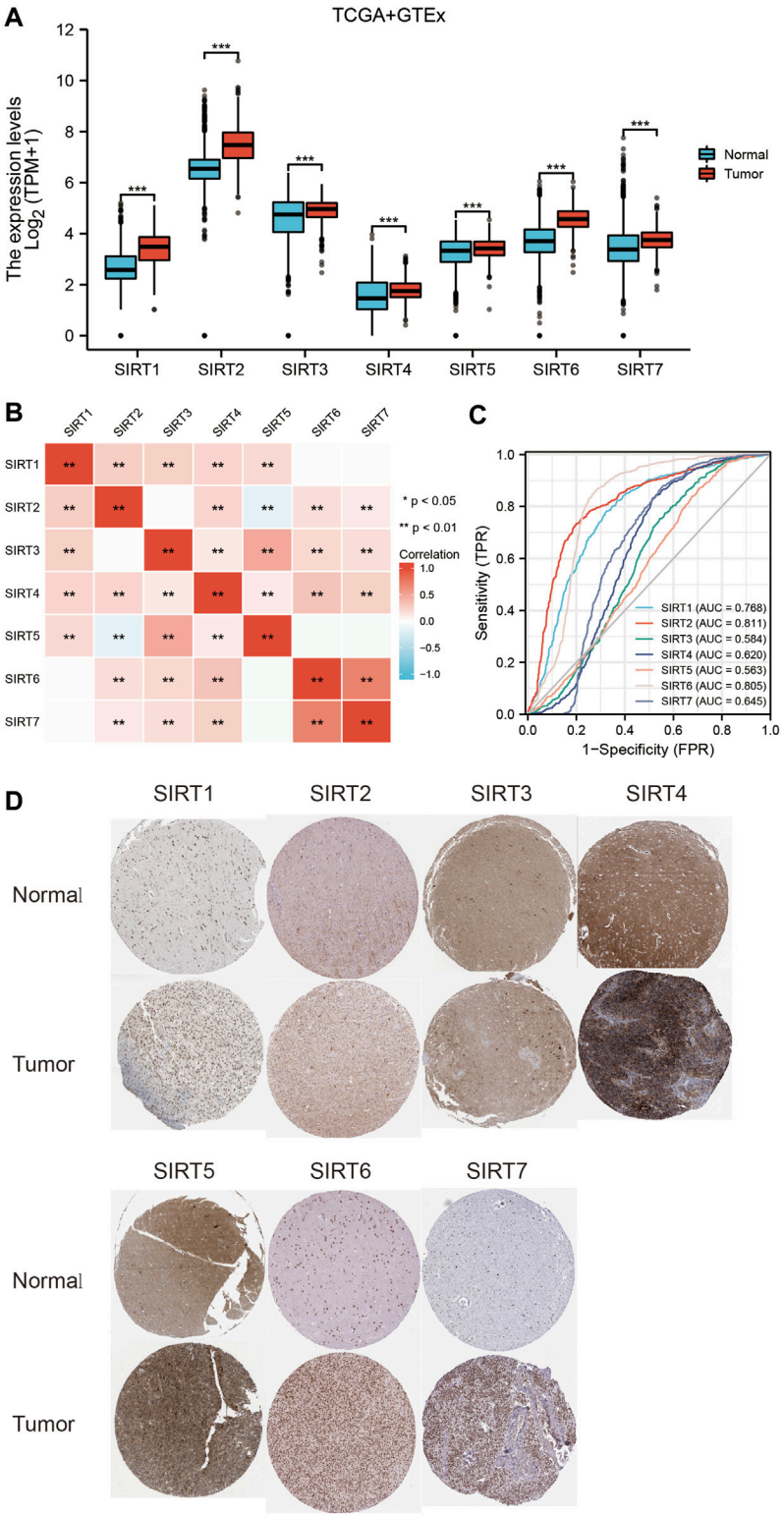


FIGURE 1
Expression levels and diagnostic values of SIRT genes in gliomas. **(A)** Differential expression of SIRT genes between glioma and normal brain tissues in TCGA and GETx databases. **(B)** Correlation analysis of SIRT gene expression in glioma tissues of the TCGA database. **(C)** Diagnostic ROC curves of SIRT genes in TCGA and GETx databases. **(D)** Expression levels of SIRT genes in glioma vs. normal brain tissues of HPA database (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: $p > 0.05$).

Western blot

Total protein was isolated using RIPA lysis buffer (Servicebio, China). The protein with the corresponding volume of loading buffer was boiled in a water bath for 10 min, after which the isolated proteins were separated through SDS-PAGE, transferred to a PVDF membrane, and incubated with primary and secondary antibodies including SIRT1 (Abclonal, A17307), SIRT5 (Proteintech, 15122-1-AP), and actin (Abclonal, AC038).

Statistical analyses

To test differences in continuous variables between the two groups, *t*-tests or Mann–Whitney *U*-tests were used, depending on the distribution of the data. Continuous variables were compared among the three groups of samples using Kruskal Wallis tests. Chi-squared tests were applied to classify the data. All correlation analyses were performed using a Spearman correlation test. Kaplan–Meier survival analysis and the logrank method were applied to evaluate survival rates. All statistical analyses were performed using R software; statistical significance is reported at $p < 0.05$.

Results

Upregulated SIRT family genes in glioma

The joint TCGA and GTEx analysis displayed that at the transcriptional level, expression levels of SIRT genes were significantly higher in glioma samples than in normal brain tissues (Figure 1A). Figure 1B shows the correlation between SIRT gene expression in TCGA gliomas. The ROC curve tests showed that SIRT2 (AUC = 0.811) and SIRT6 (AUC = 0.805) can effectively distinguish glioma samples from normal brain tissue samples (Figure 1C). Figure 1D shows SIRT gene expression in gliomas and normal brain tissues in the HPA database. The results confirmed significantly higher expression of SIRT family genes in gliomas.

Genetic changes of SIRT genes in gliomas

The genetic changes of SIRT genes in gliomas were explored, and the highest mutation rate occurred in SIRT3 with 1.9%, and the main form of mutation form was deletion (Figure 2A). The mutation frequencies of SIRT2 and SIRT6 were somewhat lower, with 1.1%, each, and mainly amplification mutations (Figure 2A). In addition, we also found a significant correlation between the gene mutations in SIRT family members and the prognosis of glioma patients as patients without mutations had worse OS, DSS, and PFI, in comparison with those bearing mutations in SIRT genes (Figures 2B–D) ($p < 0.05$).

SIRT family genes are related to the prognosis of glioma

The SIRT family was grouped based according to the median expression value of each gene to evaluate their prognostic value in the TCGA data set. The evaluation indicators included OS, DSS, and PFI. The results showed that SIRT1, SIRT3, and SIRT5 were associated with better OS, while SIRT7 was associated with worse OS (Figure 3A) ($p < 0.05$). SIRT1, SIRT2, SIRT3, and SIRT5 were associated with better DSS, while SIRT7 was associated with worse DSS (Figure 3B) ($p < 0.05$). SIRT1, SIRT2, and SIRT3 were associated with better PFI (Figure 3C) ($p < 0.05$).

To evaluate whether SIRT family members had independent predictive ability for the prognosis of glioma patients, we included SIRT gene expression and clinical information of TCGA glioma patients such as gender, age, pathological grade, and IDH status in univariate and multivariate Cox analyses. The results (Table 1) suggest high expression of SIRT1 as an independent prognostic factor which can predict OS in glioma patients ($p < 0.05$). Tables 2, 3 show that the high expression of SIRT5 is a good independent predictor of DSS and PFI, respectively, in glioma patients ($p < 0.05$), whereas, high expression of SIRT7 was an independent predictor of poor PFI in glioma patients ($p < 0.05$).

PPI network and function enrichment analysis

To identify the proteins that may interact with SIRT family genes, we developed and visualized a PPI network using the GeneMANIA database (Figure 4A). The results of GeneMANIA produced 20 proteins in interaction with the SIRT genes. DHPS, ETFA, HACL1, and ILVBL were the closest respective ones. Functioning of SIRT family genes and their related proteins are pivotally related to biological processes such as protein-protein deacetylase activity, carboxylic acid catabolism, organic acid catabolism, protein deacylation, macromolecule deacylation, oxidoreductase complex, and protein deacetylation (Figure 4A). To explore the possible biological processes involved in the SIRT gene family, we used the Metascape database to analyze functional enrichment of the SIRT family and interacting genes.

SIRT genes are mainly enriched in biological processes such as NAD + metabolism, cellular respiration, carboxylic acid

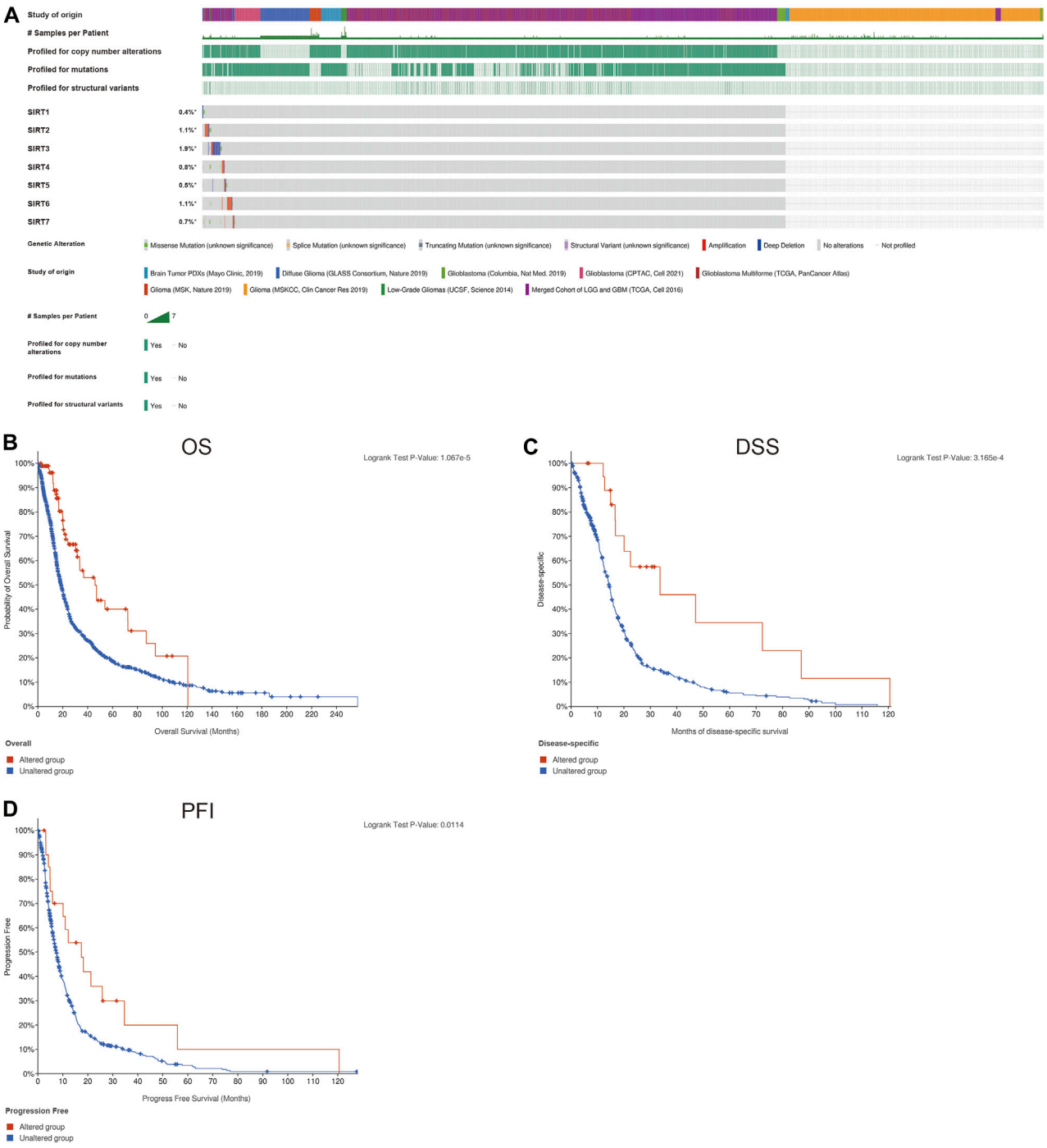


FIGURE 2
Genetic changes of SIRT family members in gliomas. **(A)** Mutations in SIRT family members in gliomas. **(B)** The association of SIRT gene mutations and OS in glioma patients. **(C)** Association of SIRT gene mutations and DSS in glioma patients. **(D)** Association of SIRT gene mutations and PFI in glioma patients (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: $p > 0.05$).

catabolic process, thiamine metabolic pathways, metabolic reprogramming in colon cancer, transcriptional activation of mitochondrial biogenesis, cellular lipid catabolic process, branched-chain amino acid metabolic process, and glycolytic process (Figure 4B). Figure 4C shows the interaction network of these biological processes.

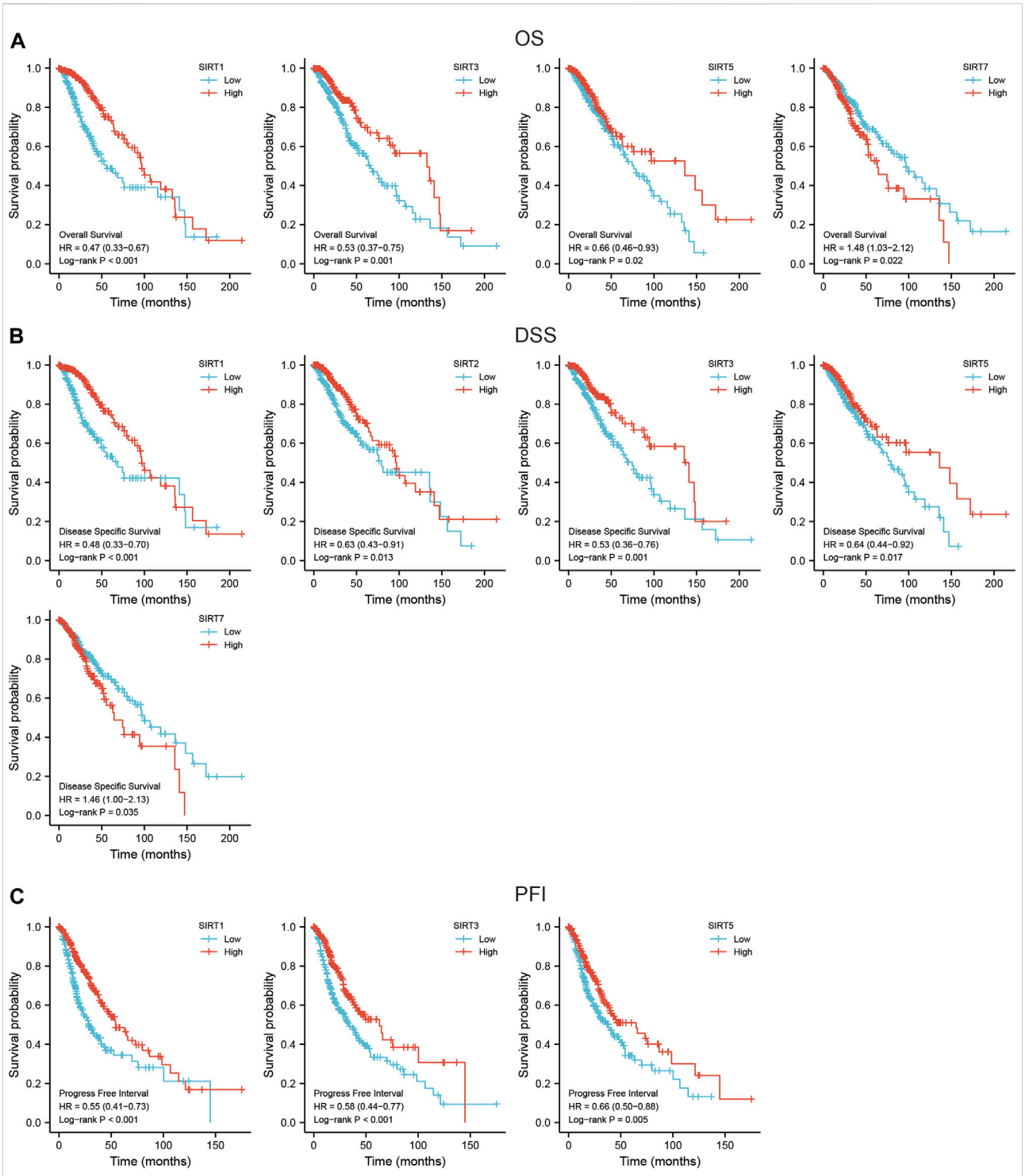


FIGURE 3
Prognostic value of SIRT genes in glioma. **(A)** SIRT family members associated with OS in glioma patients. **(B)** SIRT family members associated with DSS in glioma patients. **(C)** SIRT family members associated with PFI in glioma patients (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: $p > 0.05$).

TABLE 1 SIRT family and clinical characteristics as analyzed by OS-based univariate and multivariate Cox regression in the TCGA database.

Characteristics	Total(N)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	*p*-value	Hazard ratio (95% CI)	*p*-value
Gender	669				
Female	283	Reference			
Male	386	1.230 (0.955–1.585)	0.109	1.081 (0.819–1.426)	0.584
WHO grade	612	19.164 (12.573–29.209)	<0.001	3.144 (1.710–5.779)	<0.001
IDH status	660				
WT	237	Reference			
Mut	423	0.102 (0.077–0.135)	<0.001	0.352 (0.217–0.573)	<0.001
Age	669	1.066 (1.056–1.076)	<0.001	1.034 (1.021–1.047)	<0.001
SIRT1	669	0.189 (0.148–0.242)	<0.001	0.657 (0.444–0.971)	0.035
SIRT2	669	0.574 (0.477–0.690)	<0.001	0.982 (0.803–1.201)	0.862
SIRT3	669	0.346 (0.251–0.478)	<0.001	0.897 (0.581–1.384)	0.623
SIRT4	669	0.695 (0.485–0.996)	0.047	1.158 (0.753–1.781)	0.503
SIRT5	669	0.265 (0.170–0.412)	<0.001	0.572 (0.327–1.000)	0.050
SIRT6	669	1.288 (0.957–1.733)	0.095	0.655 (0.418–1.026)	0.064
SIRT7	669	1.880 (1.372–2.576)	<0.001	1.168 (0.717–1.904)	0.533

TABLE 2 The SIRT family and clinical features were analyzed by DSS-based univariate and multivariate Cox regression in the TCGA database.

Characteristics	Total(N)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	*p*-value	Hazard ratio (95% CI)	*p*-value
Gender	648				
Female	275	Reference			
Male	373	1.229 (0.940–1.607)	0.132	1.007 (0.745–1.361)	0.963
WHO grade	590	19.859 (12.672–31.123)	<0.001	3.086 (1.620–5.877)	<0.001
IDH status	639				
WT	223	Reference			
Mut	416	0.096 (0.071–0.130)	<0.001	0.309 (0.186–0.513)	<0.001
Age	648	1.066 (1.056–1.077)	<0.001	1.034 (1.021–1.048)	<0.001
SIRT1	648	0.189 (0.146–0.245)	<0.001	0.723 (0.475–1.102)	0.132
SIRT2	648	0.550 (0.451–0.670)	<0.001	0.964 (0.775–1.200)	0.745
SIRT3	648	0.330 (0.232–0.468)	<0.001	0.798 (0.490–1.302)	0.367
SIRT4	648	0.694 (0.474–1.017)	0.061	1.180 (0.749–1.858)	0.475
SIRT5	648	0.260 (0.163–0.415)	<0.001	0.533 (0.293–0.969)	0.039
SIRT6	648	1.206 (0.882–1.650)	0.240	0.642 (0.400–1.031)	0.067
SIRT7	648	1.823 (1.308–2.541)	<0.001	1.204 (0.724–2.003)	0.474

TABLE 3 SIRT family and clinical characteristics as analyzed by PFI-based univariate and multivariate Cox regression in the TCGA database.

Characteristics	Total(N)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	*p*-value	Hazard ratio (95% CI)	*p*-value
Gender	669				
Female	283	Reference			
Male	386	1.068 (0.854–1.336)	0.564	1.023 (0.806–1.297)	0.853
WHO grade	612	7.900 (5.766–10.822)	<0.001	1.584 (0.979–2.562)	0.061
IDH status	660				
WT	237	Reference			
Mut	423	0.140 (0.110–0.180)	<0.001	0.285 (0.182–0.445)	<0.001
Age	669	1.041 (1.033–1.049)	<0.001	1.013 (1.002–1.023)	0.015
SIRT1	669	0.255 (0.204–0.318)	<0.001	0.754 (0.532–1.068)	0.112
SIRT2	669	0.653 (0.552–0.771)	<0.001	1.014 (0.841–1.224)	0.881
SIRT3	669	0.415 (0.310–0.555)	<0.001	0.844 (0.572–1.245)	0.393
SIRT4	669	0.681 (0.490–0.947)	0.022	1.182 (0.796–1.757)	0.407
SIRT5	669	0.312 (0.213–0.457)	<0.001	0.519 (0.326–0.828)	0.006
SIRT6	669	1.106 (0.854–1.433)	0.443	0.590 (0.397–0.878)	0.009
SIRT7	669	1.630 (1.223–2.171)	<0.001	1.621 (1.035–2.540)	0.035

Establishing and verifying a SIRT-based signature

Using the TCGA database, the SIRT genes were analyzed by LASSO regression tests followed by multivariate Cox regression (Figures 5A,B), based on which SIRT-based signatures containing the five SIRT family genes SIRT1, SIRT2, SIRT5, SIRT6, and SIRT7 were generated. The expression levels of these five genes and their coefficients in a linear combination were used to determine the risk score for each sample, as follows:

$$\begin{aligned}\text{Risk score} = {} & \text{SIRT1} \times (-0.961) + \text{SIRT2} \times (-0.318) \\ & + \text{SIRT5} \times (-0.473) + \text{SIRT6} \times (-0.239) \\ & + \text{SIRT7} \times (0.784)\end{aligned}$$

Accordingly, the risk score was calculated for every glioma patient in the TCGA database which was used for stratifying the patients in HR and LR groups according to the median risk score (Figure 5C). Figure 5C illustrates the survival status and survival time of the HR group and LR patients. The scatter chart shows the higher mortality in HR patients. The expression profiles of these five SIRT genes in the HR and LR groups (TCGA database) are shown in Figure 5C. According to the outcomes of Kaplan–Meier survival analyses, OS was significantly lower in the HR than in the LR group (Figure 5D). In addition, time-dependent ROC curve analyses showed that the AUC values of 1-, 3-, and 5-year OS were 0.81, 0.81, and 0.85, respectively (Figure 5E), which indicates the potential of a SIRT-based signature for accurate prediction of OS in glioma patients.

The CGGA queue was used as a test dataset for further verification of accuracy and universality of the SIRT-based signature. Then, the risk score of patients in this queue was calculated using the risk score formula obtained in the train dataset, and patients were divided into the HR and LR groups according to the median value (Figure 5F). The survival status and survival time of the patients from the CGGA database and the expression patterns of the five SIRT genes are shown in Figure 5F. Kaplan–Meier survival analysis showed higher OS in LR than in HR patients (Figure 5G). Time-dependent ROC curve analysis indicated 1-, 3-, and 5-year OS AUC values of 0.7, 0.74, and 0.76, respectively, which indicates that the SIRT-based signature is accurate and specific for predicting OS of glioma patients (Figure 5H).

The SIRT-based signature can independently predict OS of glioma patients

Univariate and multivariate Cox regression analyses were used to determine the potential of the SIRT-based signature of independently predicting OS in glioma patients as well as their clinical features in TCGA and CGGA databases. In the TCGA database, SIRT-based signature, age, pathological grade, and the IDH status were independent predictors (Table 4). In the CGGA

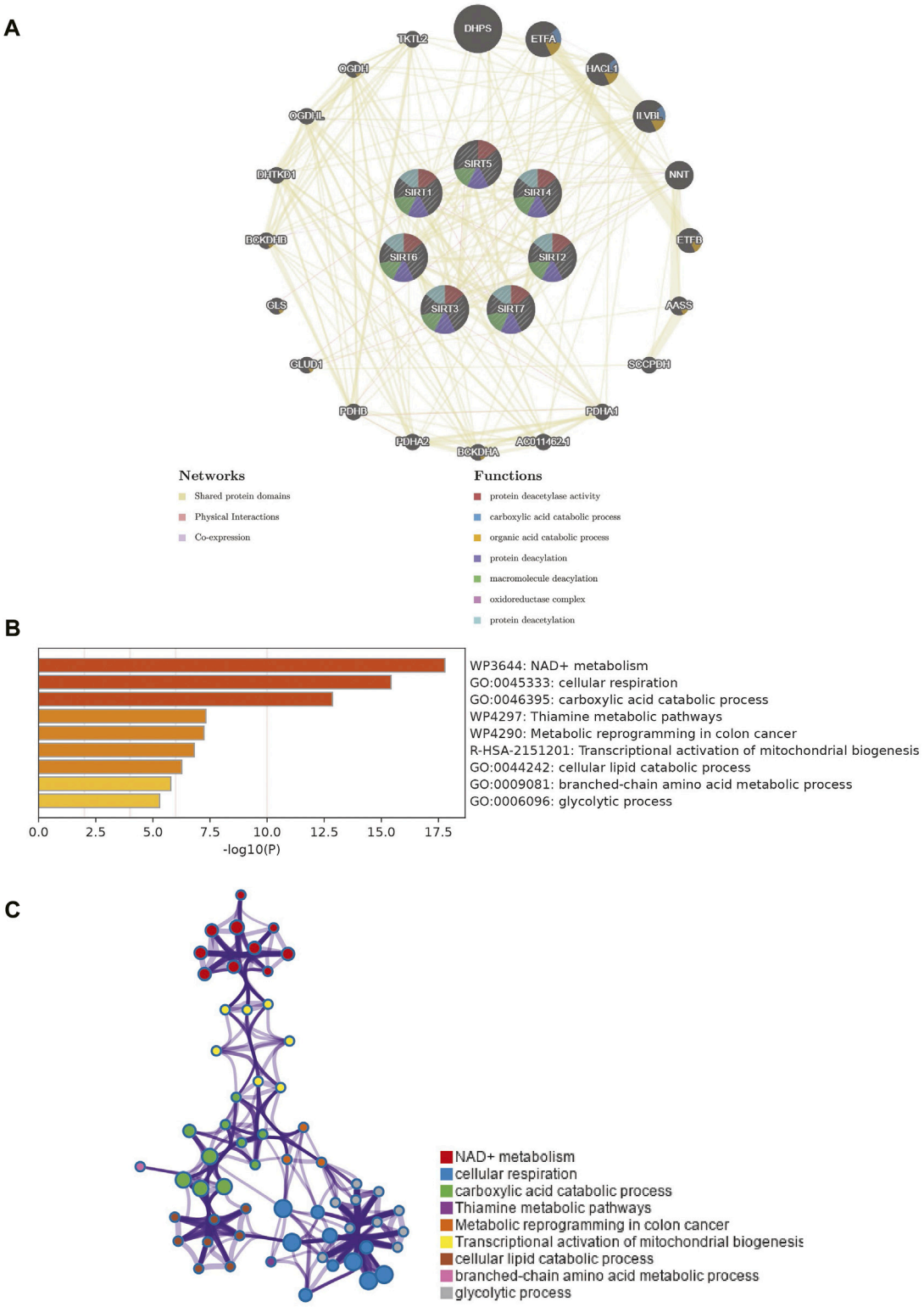


FIGURE 4
PPI network of SIRT family members and enrichment analysis of network genes. **(A)** PPI network of SIRT family members and their interacting proteins. **(B)** Gene enrichment in the PPI network as analyzed using Metascape. **(C)** Network diagram of enrichment analysis results (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: $p > 0.05$).

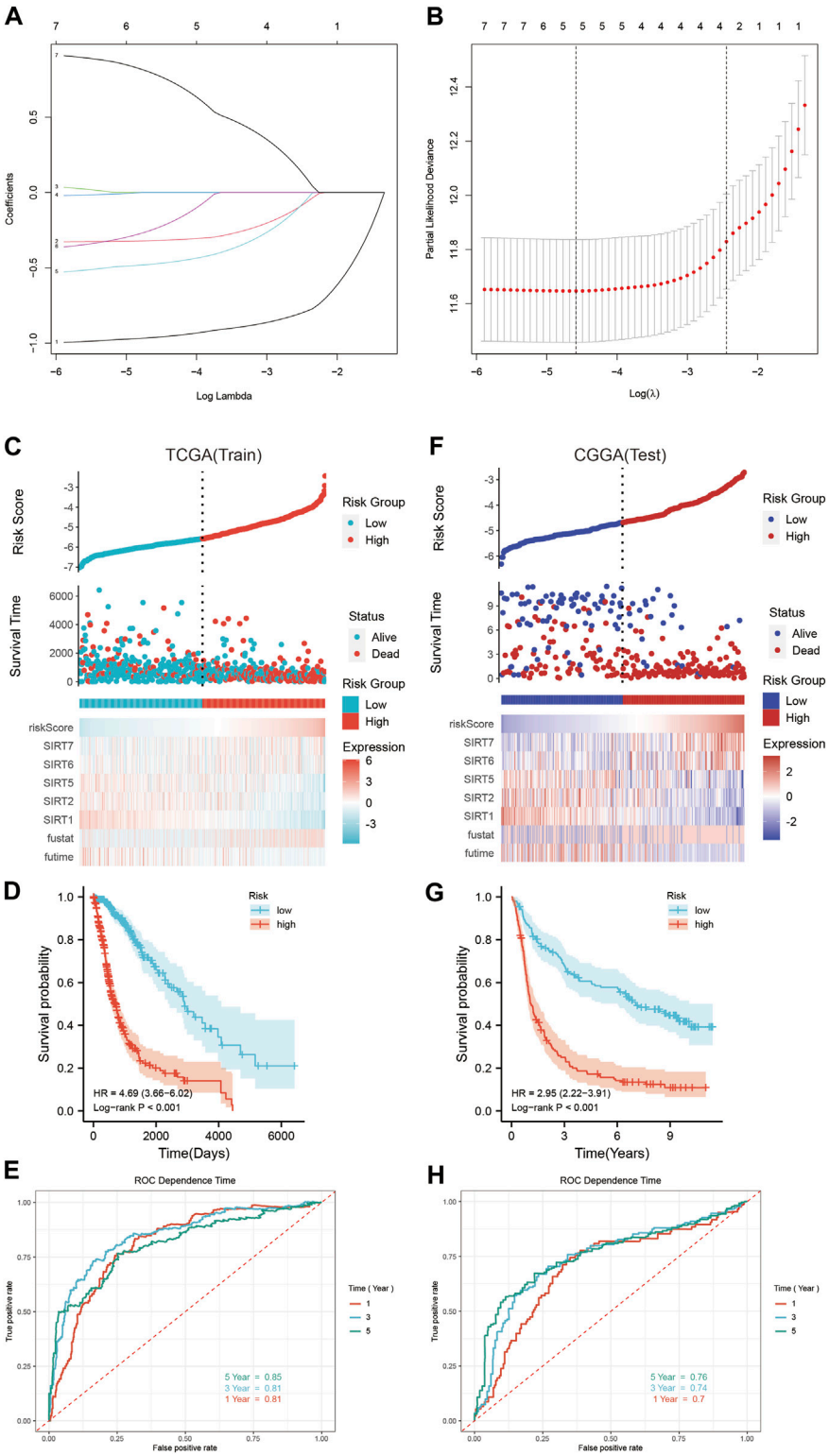


FIGURE 5
Construction and verification of SIRT-based signature. **(A)** Lasso coefficients of SIRT gene. **(B)** Relationship between deviance and log(lambda). **(C)** The distribution of risk scores, survival time and status, as well as the five SIRT genes' expression profile in glioma patients of the TCGA database. **(D)** Kaplan–Meier survival analysis of OS in TCGA database. **(E)** The SIRT-based signature predicts the time-dependent ROC curves of 1 -, 3 -, and 5-year OS in glioma patients in the TCGA database. **(F)** Distribution of risk scores, survival time and status, and the five SIRT gene expression profiles in glioma patients of the CGGA database. **(G)** Kaplan–Meier survival analysis of OS in the CGGA database. **(H)** SIRT-based signature predicts the time-dependent ROC curves of 1 -, 3 -, and 5-year OS in glioma patients in the CGGA database (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: $p > 0.05$).

TABLE 4 The SIRT-Based signature and clinical features were analyzed by OS-based univariate and multivariate Cox regression in the TCGA database.

Characteristics	Total(N)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	*p*-value	Hazard ratio (95% CI)	*p*-value
Riskscore	686	2.882 (2.495–3.330)	<0.001	1.559 (1.177–2.066)	0.002
Age	686	1.066 (1.057–1.076)	<0.001	1.049 (1.033–1.064)	<0.001
Gender	686				
Male	395	Reference			
Female	291	0.818 (0.638–1.049)	0.114	0.808 (0.560–1.166)	0.255
Histological Grade	612	3.407 (2.324–4.994)	<0.001	2.123 (1.414–3.187)	<0.001
IDH Mutation	672				
WT	242	Reference			
Mutant	430	0.108 (0.082–0.142)	<0.001	0.343 (0.219–0.539)	<0.001

TABLE 5 The SIRT-Based signature and clinical features were analyzed by OS-based univariate and multivariate Cox regression in the TCGA database.

Characteristics	Total(N)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	*p*-value	Hazard ratio (95% CI)	*p*-value
riskScore	302	2.069 (1.737–2.466)	<0.001	1.255 (0.998–1.579)	0.050
Grade	302	7.957 (5.334–11.871)	<0.001	5.291 (3.416–8.195)	<0.001
Gender	302				
Male	186	Reference			
Female	116	1.083 (0.819–1.431)	0.576	0.969 (0.729–1.289)	0.829
Age	302	1.031 (1.018–1.044)	<0.001	1.012 (0.998–1.025)	0.085
IDH Mutation	302				
Wildtype	138	Reference			
Mutant	164	0.384 (0.289–0.509)	<0.001	1.391 (0.946–2.046)	0.093
1p19q codeletion	302				
Non-codel	239	Reference			
Codel	63	0.168 (0.102–0.277)	<0.001	0.258 (0.150–0.442)	<0.001

database, SIRT-Based signature, pathological grade and 1p19q co-deletion were independent predictors (Table 5).

The SIRT-based signature can predict DSS and PFI in glioma patients

The TCGA database included DSS and PFI information of glioma patients, and we assessed whether the SIRT-based signature can predict DSS and PFI in TCGA queues. The results showed that the DSS and PFI in LR patients were higher than those in HR patients (Figures 6A,C). The time-dependent ROC curve analysis indicated that after 1, 3, and 5 years, the AUC values of DSS were 0.81, 0.81, and 0.84, respectively (Figure 6B), and those of PFI were 0.7, 0.74, and 0.76, respectively (Figure 6D). To explore the relationship between SIRT-based signature and clinical characteristics, we performed a correlation analysis between the risk scores and clinical characteristics of patients in the TCGA and CGGA databases. In the TCGA database, patients with advanced age, high pathological grade, and IDH wild-type glioma tended to show higher risk scores (Figure 6E). In the CGGA database,

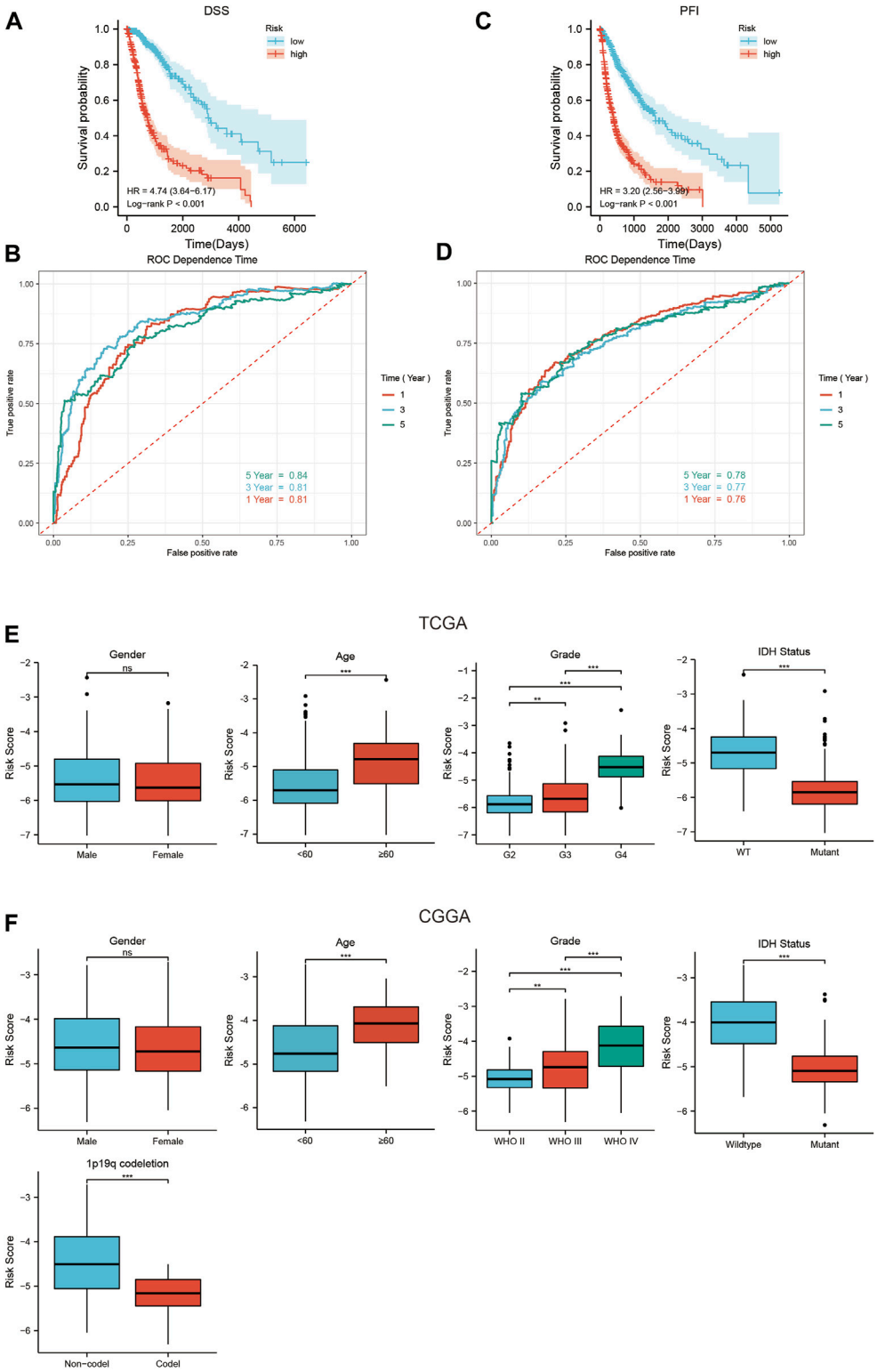


FIGURE 6
The relationship between the SIRT-Based signature and DSS, PFI, and clinical characteristics of glioma patients. **(A)** The SIRT-based signature predicts the Kaplan–Meier survival of DSS in the TCGA queue. **(B)** The SIRT-based signature predicts the time-dependent ROC curves of 1-, 3-, and 5-year DSS of glioma patients in the TCGA database. **(C)** The SIRT-Based signature predicts the Kaplan-Meier survival of PFI in the TCGA queue. **(D)** The SIRT-based signature predicts the time-dependent ROC curves of 1-, 3-, and 5-year PFI of glioma patients in the TCGA database. **(E)** Relationship between SIRT-Based signature and sex, age, pathological grade, and IDH status of glioma patients in TCGA database. **(F)** Relationship between the SIRT-based signature and sex, age, pathological grade, IDH status, and 1p19q co-deletion of glioma patients in the CGGA database (*p < 0.05, **p < 0.01, ***p < 0.001, ns: p > 0.05).

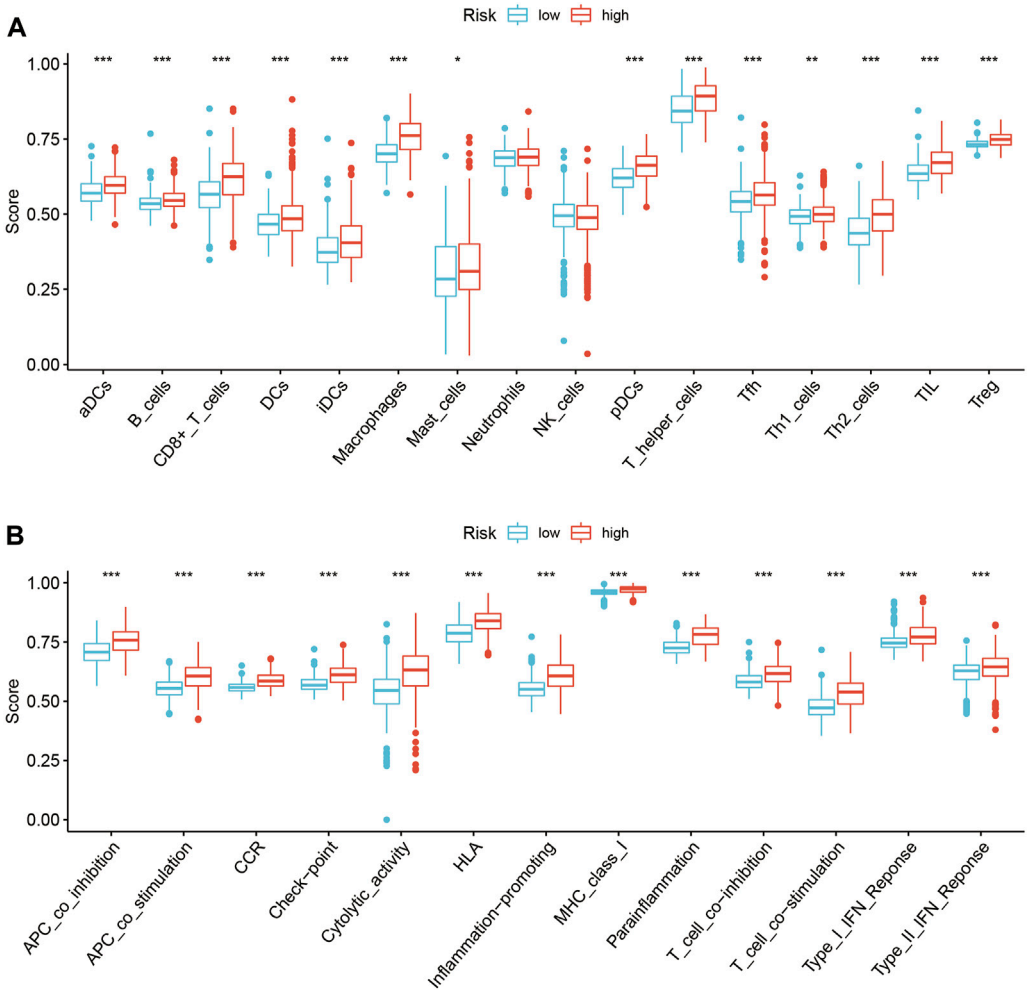


FIGURE 7
Relationship between the SIRT-based signature and immune characteristics of glioma patients. **(A)** The ssGSEA algorithm was used to calculate the difference in immune cell infiltration between HR and LR glioma patients in the TCGA database. **(B)**The ssGSEA algorithm was used to calculate the difference in immune status scores between high and LR groups in the TCGA database (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: $p > 0.05$).

patients with advanced age, high pathological grade, IDH wild type, and no 1p19q co-deletion glioma tended to have high risk scores (Figure 6F).

The SIRT-based signature can evaluate immune cell infiltration and immune status

The types of infiltrating immune cells in all TCGA glioma samples in the TCGA database were evaluated using ssGSEA algorithms (Figure 7A). The scores of several immune subgroups such as DCs, aDCs, iDCs, pDCs, B cells, $CD8^+$ T cells, macrophages, mast cells, T helper cells, Tfh, Th1 cells, Th2 cells, and TIL were higher in the HR than in the LR group. The scores of several immune processes including APC co-inhibition, APC co-stimulation, CCR, check-point, cytolytic activity, HLA, inflammation-promoting, MHC class I, para-inflammation, T cell co-inhibition, T cell co-stimulation, type I IFN response, and type II IFN response were higher in HR than in LR patients (Figure 7B). These findings suggest that the SIRT-based signature reflects infiltration by various immune cells and the glioma patients' immune status.

Using the SIRT-based signature to evaluate the sensitivity of glioma patients to drug therapy

In recent years, immune checkpoint inhibitors (ICIs) have emerged as a milestone during the attempts of treating glioma (Tian et al., 2022). However, only few patients benefit from it, and

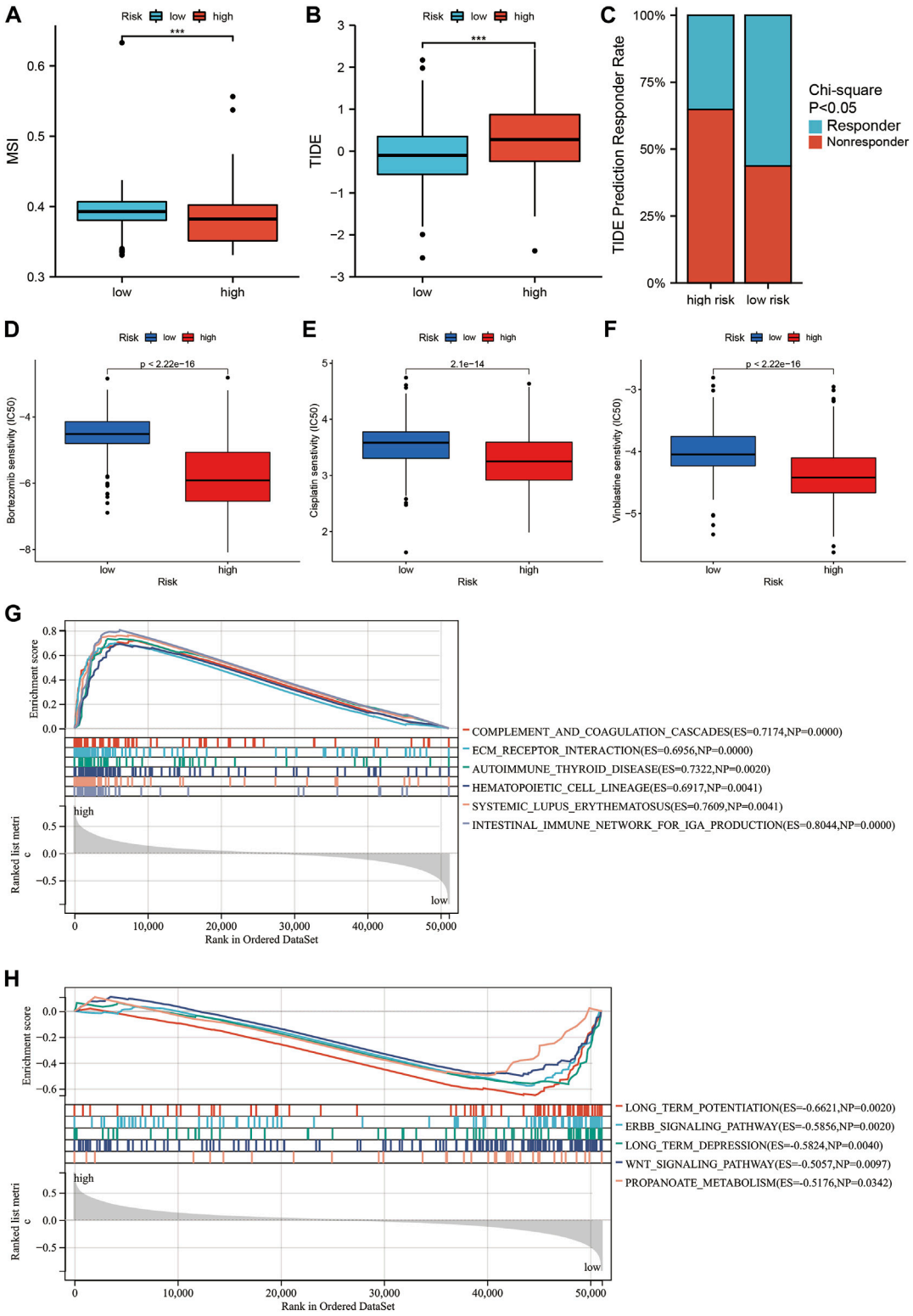


FIGURE 8
Relationship between the SIRT-based signature and drug sensitivity in glioma patients and GSEA. **(A)** Differences in MSI scores between HR and LR glioma patients in the TCGA database. **(B)** The difference in TIDE score between HR and LR glioma patients in the TCGA database. **(C)** Comparison of the percentage of responders to immunotherapy among high and LR glioma patients in the TCGA database. **(D–F)** Comparison of sensitivity to Bortezomib **(D)**, Vinblastine **(E)**, and Cisplatin **(F)** in high and LR glioma patients in TCGA database. **(G,H)** Top5 KEGG pathways were significantly enriched in the TCGA database in the HR group **(G)** and LR group **(H)** (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: $p > 0.05$).

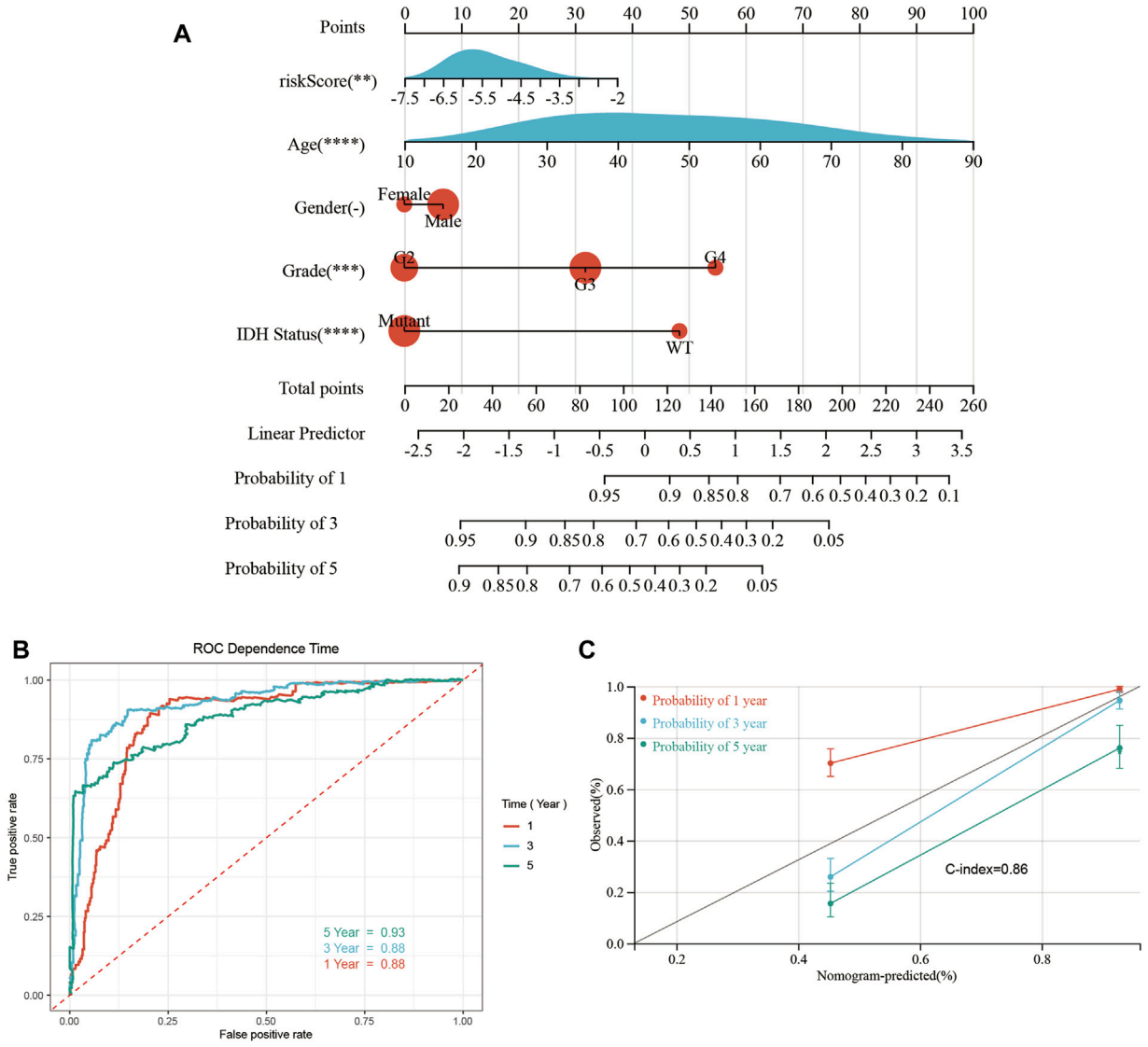


FIGURE 9
Construction and performance test of the nomogram. **(A)** Nomograms for predicting 1-, 3-, and 5-year OS in glioma patients. **(B)** Time-dependent ROC curves analysis was used to evaluate the OS of glioma patients at 1, 3, and 5 years. **(C)** The calibration curve of OS in glioma patients at 1-, 3-, and 5-year stages was evaluated through a nomogram (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: $p > 0.05$).

most patients show no response to ICI treatment. Thus, the current study examined the relationship between SIRT-based signature and sensitivity to the ICI treatment. We found that HR patients had lower MSI and higher TIDE scores than LR patients (Figures 8A,B). The TIDE algorithm can predict whether patients will respond to ICI treatment. Accordingly, the HR glioma patients showed a lower rate of sensitivity in response to ICI treatment compared to LR patients ($p < 0.05$) (Figure 8C).

Subsequently, we explored difference in the sensitivity to treatment with common chemotherapy drugs between HR and LR glioma patients using the "pRRophetic" software package. The results showed a significantly lower IC50 for Bortezomib, Vinblastine, and Cisplatin in the HR than in the LR group (Figures 8D–F). This indicates that HR glioma patients showing higher sensitivity can benefit more from a combination therapy with Bortezomib, Vinblastine, and Cisplatin.

Functions related to the SIRT-based signature

To compare the immune-related pathways between patients from HR and LR groups, GSEA analysis was conducted, which

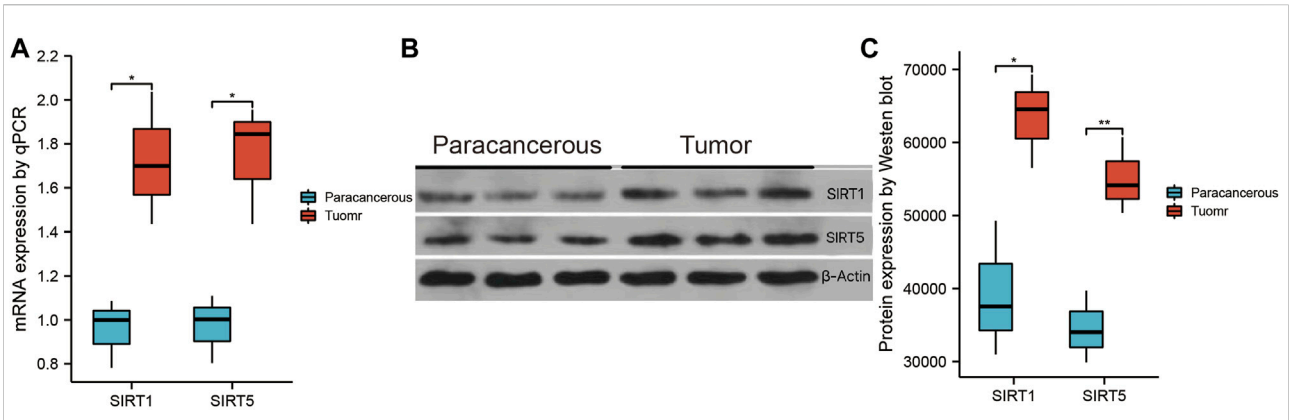


FIGURE 10
Validation of SIRT1 and SIRT5 expression. **(A)** qPCR was used to detect the differential expression of SIRT1 and SIRT5 between glioma and paracancerous tissue. **(B)** Western blotting was used to detect differential expression of SIRT1 and SIRT5 between glioma and paracancerous tissue (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns: $p > 0.05$).

showed a significant enrichment in immune-related pathways such as complement and coagulation cascades, ECM receptor interaction, and intestinal immune network for IGA production in HR patients (Figure 8G). By contrast, in LR patients, a significant enrichment occurred in some signal pathways such as long term potentiation, ERBB signaling pathway, long-term depression, Wnt signaling pathway, and propanoate metabolism (Figure 8H).

Construction of a nomogram

A predictive quantitative nomogram was developed to evaluate the clinical prognosis of glioma patients. Our OS nomogram included gender, age, risk scores, pathological grades, and IDH status records as per the TCGA database (Figure 9A). Time-dependent ROC curve analysis displayed AUC values of 0.88, 0.88, and 0.93 for 1-, 3-, and 5-year OS, respectively (Figure 9B). Calibration of the curve showed high consistency between the nomogram prediction and the actual results, with a C-index of 0.86 (Figure 9C). This confirms that nomograms have a high predictive potential for OS in patients with glioma.

Verifying expression of SIRT genes

Subsequently, we verified the differential expression of SIRT1 and SIRT5 between GBM tissues and paracancerous tissues by western blotting and qPCR. The qPCR results showed that the expression of SIRT1 and SIRT5 was higher in GBM tissues than in paracancerous tissues (Figure 10A). Consistent results were obtained through western blotting (Figures 10B,C).

Discussion

High-throughput sequencing approaches including RNS sequencing and transcriptomics, and the discovery of tumor-related genetic abnormalities have been increasingly used in cancer research and for diagnostic purposes (Hamilton et al., 2022; Iwamura et al., 2022). Some transcriptome records have been uploaded to open databases and are used for reanalysis to discover prognostic indicators and therapeutic targets of various malignant tumors (Elmallah et al., 2022; Iwamura et al., 2022). In recent years, research on the molecular characteristics of gliomas has put forward many potential markers, which can be used to classify gliomas, to judge prognosis, and to guide treatment (Guo et al., 2021). However, in isolation, these are insufficient for predicting the course of complex gliomas and lack the ability to determine effective treatment. In the current study, we developed a concise SIRT-based signature to characterize the prognosis of glioma. As expected, the SIRT-based signature was strongly associated with clinicopathological parameters indicating a poor prognosis of glioma. It is noteworthy that the SIRT-based signature is a powerful prognostic biomarker for gliomas (especially high-grade gliomas), and it facilitates identifying ICI responders. In conclusion, our work provides valuable information for the diagnosis, prognosis, and classification of glioma, and it may help optimize immunotherapy.

Various proteins are regulated by lysine acetylation, and sirtuins are considered a principal regulator contributing to various cell functions, including gene expression, DNA repair, telomerase activity, epithelial to mesenchymal transition, cell cycling, differentiation, apoptosis, proliferation, ageing, metabolism, and oxidative stress responses (O'Callaghan and Vassilopoulos, 2017). Many of these cellular functions are critical in tumor cell proliferation and differentiation. Sirtuin protein

plays different roles in promoting or inhibiting cancer in glioma cases (Funato et al., 2018; Liu et al., 2022). The roles of SIRT1, SIRT2, SIRT3, and SIRT7 in glioma are not comprehensively understood. Several studies found that SIRT1 exerts anti-cancer effects in gliomas (Ye et al., 2019; Fang et al., 2020). Fang et al. (2020) showed that glioma patients with high SIRT1 expression had a better prognosis, and our study produced consistent results. This study also found that SIRT1 could inhibit the progression of glioma, while Circ_0005075 promoted the progression of glioma by inhibiting the expression of SIRT1 (Fang et al., 2020). In addition, SIRT1 activator SRT2183 inhibits the growth of glioma cells through endoplasmic reticulum stress pathway (Ye et al., 2019). Other studies foundevidence also shows that SIRT1 can promote the development of gliomas (Li et al., 2016; Huo et al., 2017; Lei et al., 2019). Lei et al. (2019) showed that Circular RNA hsa_circ_0076248 up-regulates the expression of SIRT1 by competitively binding to SIRT1 miR-181a, thus promoting the progression of glioma. Betulinic acid derivative B10 aggravates glioma by inhibiting the acetylation of SIRT1 (Huo et al., 2017). The role of SIRT2 in glioma is controversial (Zhang et al., 2020). A study showed that SIRT3 can be regulated by myricetin nanoliposomes and accelerate glioma cell apoptosis (Wang et al., 2018). However, a different study confirmed that the interaction between TRAP1 and SIRT3 can maintain stem cell viability of glioma cells (Park et al., 2019). SIRT7 promotes proliferation and invasion in glioma cells by activating the ERK/STAT3 signaling pathway (Mu et al., 2019). Furthermore, there is evidence that SIRT7 is positively regulated by MEG3 and plays a role in inhibiting glioma (Xu et al., 2018).

Chen et al. (2019) showed that SIRT5 is associated with better prognosis of glioma cases, which consistent with our results. However, the function of SIRT5 in glioma remains unknown. Current research supports the role of SIRT5 as a tumor suppressor in many cancers, and this effect is achieved by regulating metabolism-related pathways. SIRT5 can regulate bile acid metabolism and inhibit immune escape of liver cancer cells (Sun et al., 2022). In renal cell carcinoma, SIRT5 inhibits tumor progression through Warburg efect (Yihan et al., 2021). The loss of SIRT5 regulates Glutamine Metabolism and Glutathione Metabolism and Pyrimidine Metabolism promotes pancreatic cancer (Hu et al., 2021). The current study confirms the carcinogenic effects of SIRT6 in glioma. There is evidence that SIRT6 is regulated by MST1 and miR-33a and can inhibit the activity of glioma cells (Chang et al., 2017; Zhu et al., 2019). In conclusion, these findings suggest that sirtuin protein is related to glioma onset and progression, suggesting these proteins as potential biomarkers for evaluating prognosis of the glioma.

In conclusion, given their important role in gliomas, we performed a comprehensive analysis of seven members of the SIRT family in gliomas. All seven SIRT family genes showed significant up-regulation in glioma tissues, further indicating the critical role of SIRT in gliomas. In addition, the Kaplan–Meier curve showed that SIRT1, SIRT3, and SIRT5 were closely associated with OS, DSS, and PFI of glioma patients, and other SIRT family members are related to the prognosis of glioma patients to varying degrees. We thus established a SIRT-based signature and further divided the subjects into HR and LR groups. The Kaplan–Meier curve analysis of the train and test datasets indicated lower OS in HR than in LR patients. The time-dependent ROC curve test indicated the powerful prediction potential of the SIRT-based signature. Further, univariate and multivariate Cox analyses confirmed the independent prediction potential of the SIRT-based signature. It is worth mentioning that the SIRT-based signature shows a considerable potential for effectively predicting DSS and PFI in patients suffering from glioma. Our findings show that the established SIRT-based signature is accurate and clinically applicable for prognosis purposes in glioma patients. In addition, the current constructed and modified OS nomogram based on gender, age, risk scores, pathological grades, and IDH status showed a higher accuracy compared to nomograms based on a single risk score. Our nomogram is thus helpful for clinical decision-making and personalized management of glioma patients. SIRT family members have been studied in gliomas, however, there is considerable controversy, and the specific functions of these genes and proteins in glioma cases require further research. To the best of our knowledge, this is the first comprehensive and systematic study on the prognostic potential of SIRT family members in glioma, confirming the important role of this family and SIRT-based signature in glioma prognosis and providing valuable biomarkers for the prognostic management of this disease.

For further differentiating the biological functions between the HR and LR groups, we conducted GSEA analyses and discovered significant enrichment in immune-related biological processes in HR patients. The association between the SIRT-based signature and tumor immune cell infiltration was also evaluated, showing significant differences in the immune status and immune cell abundance between HR and LR patients. The importance of infiltrated immune cells for immunotherapy and prognosis has been recognized (Helmink et al., 2020), which suggests that responses to immunotherapy may differ between the two groups. MSI is a more effective marker for selecting glioma patients for immunotherapy (Mensali and Inderberg, 2022). Glioma patients with higher MSI scores are more sensitive to anti-PD-L1 treatment. The TIDE algorithm is an effective tool to predict whether cancer patients will respond to immunotherapy. Generally, a higher TIDE score indicates less sensitivity to immunotherapy (Jiang et al., 2018). However, the present results indicate that HR groups were less sensitive to immunotherapy, whether through MSI scores or the TIDE algorithm. In the present study, HR patients were more sensitive to the combination of Bortezomib, Vinblastine, and Cisplatin. According to previous studies, Bortezomib inhibits glioma growth and improves the efficacy

of TMZ chemotherapy (Tang et al., 2019), while Vinblastine and Cisplatin were selected as they are commonly used chemotherapy drugs for glioma patients (Park et al., 2016; Rottenberg et al., 2021).

In conclusion, the present constructed SIRT-based signature of glioma showed that it is accurate for evaluating the prognosis of glioma patients and their sensitivity to drug treatment. Therefore, it provides a personalized scheme for the management of glioma patients. In addition, our SIRT-based signature showed a significant correlation with the tumor immune status and the infiltration level of different immune cells in the tumor microenvironment.

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 Hongqi Hospital affiliated to Mudanjiang Medical College (No.2022036 the registration number of ethics board). The patients/participants provided their written informed consent to participate in this study.

Author contributions

FX designed the study and analyzed the data. FX and ZZ wrote the manuscript. KL, HG, SL, and YZ prepared the images and tables. HL reviewed and revised the manuscript. All authors approved the final manuscript.

Funding

This work was supported by the Basic scientific research projects of provincial colleges and universities in Heilongjiang Province [grant number2020-KYYWF-0758, 2020-KYYWF-0715 and 2021-KYYWF-0503].

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/fgene.2022.1035368/full#supplementary-material

References

Brown, K., Xie, S., Qiu, X., Mohrin, M., Shin, J., Liu, Y., et al. (2013). SIRT3 reverses aging-associated degeneration. *Cell Rep.* 3, 319–327. doi:10.1016/j.celrep.2013.01.005

Chang, M., Qiao, L., Li, B., Wang, J., Zhang, G., Shi, W., et al. (2017). Suppression of SIRT6 by miR-33a facilitates tumor growth of glioma through apoptosis and oxidative stress resistance. *Oncol. Rep.* 38, 1251–1258. doi:10.3892/or.2017.5780

Chen, X., Li, D., Gao, Y., Cao, Y., and Hao, B. (2018). Histone deacetylase SIRT6 inhibits glioma cell growth through down-regulating NOTCH3 expression. *Acta Biochim. Biophys. Sin.* 50, 417–424. doi:10.1093/abbs/gmy019

Chen, X., Xu, Z., Zeng, S., Wang, X., Liu, W., Qian, L., et al. (2019). SIRT5 downregulation is associated with poor prognosis in glioblastoma. *Cancer Biomark.* 24, 449–459. doi:10.3233/CBM-182197

Elmallah, M. I. Y., Ortega-Deballon, P., Hermite, L., Pais-De-Barros, J. P., Gobbo, J., and Garrido, C. (2022). Lipidomic profiling of exosomes from colorectal cancer cells and patients reveals potential biomarkers. *Mol. Oncol.* 16, 2710–2718. doi:10.1002/1878-0261.13223

Fang, D. Z., Wang, W. J., Li, F. Y., Liu, J., Hui, X. B., Liu, D., et al. (2020). Circ_0005075 stimulates the proliferation and metastasis of glioma via downregulating SIRT1. *Eur. Rev. Med. Pharmacol. Sci.* 24, 258–266. doi:10.26355/eurrev_202001_19918

Franz, M., Rodriguez, H., Lopes, C., Zuberi, K., Montojo, J., Bader, G. D., et al. (2018). GeneMANIA update 2018. *Nucleic Acids Res.* 46, W60–w4. doi:10.1093/nar/gky311

Funato, K., Hayashi, T., Echizen, K., Negishi, L., Shimizu, N., Koyama-Nasu, R., et al. (2018). SIRT2-mediated inactivation of p73 is required for glioblastoma tumorigenicity. *EMBO Rep.* 19, e45587. doi:10.15252/embr.201745587

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* 6, pl1. doi:10.1126/scisignal.2004088

Guo, J. C., Wei, Q. S., Dong, L., Fang, S. S., Li, F., and Zhao, Y. (2021). Prognostic value of an autophagy-related five-gene signature for lower-grade glioma patients. *Front. Oncol.* 11, 644443. doi:10.3389/fonc.2021.644443

Hamilton, D. J., Hein, A. E., Holmes, Z. E., Wuttke, D. S., and Batey, R. T. (2022). The DNA-binding high-mobility group box domain of sox family proteins directly interacts with RNA *in vitro*. *Biochemistry* 61, 943–951. doi:10.1021/acs.biochem.2c00218

Helmink, B. A., Reddy, S. M., Gao, J., Zhang, S., Basar, R., Thakur, R., et al. (2020). B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* 577, 549–555. doi:10.1038/s41586-019-1922-8

Hu, T., Shukla, S. K., Vernucci, E., He, C., Wang, D., King, R. J., et al. (2021). Metabolic rewiring by loss of Sirt5 promotes kras-induced pancreatic cancer progression. *Gastroenterology* 161 (5), 1584–1600. doi:10.1053/j.gastro.2021.06.045

Huo, L., Bai, X., Wang, Y., and Wang, M. (2017). Betulinic acid derivative B10 inhibits glioma cell proliferation through suppression of SIRT1, acetylation of FOXO3a and upregulation of Bim/PUMA. *Biomed. Pharmacother.* 92, 347–355. doi:10.1016/j.biopha.2017.05.074

Iwamura, H., Mizuno, K., Akamatsu, S., Hatakeyama, S., Tobisawa, Y., Narita, S., et al. (2022). Machine learning diagnosis by immunoglobulin N-glycan signatures for precision diagnosis of urological diseases. *Cancer Sci.* 113, 2434–2445. doi:10.1111/cas.15395

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24, 1550–1558. doi:10.1038/s41591-018-0136-1

Jiang, T., Mao, Y., Ma, W., Mao, Q., You, Y., Yang, X., et al. (2016). CGCG clinical practice guidelines for the management of adult diffuse gliomas. *Cancer Lett.* 375, 263–273. doi:10.1016/j.canlet.2016.01.024

Jiang, T., Nam, D. H., Ram, Z., Poon, W. S., Wang, J., Boldbaatar, D., et al. (2021). Clinical practice guidelines for the management of adult diffuse gliomas. *Cancer Lett.* 499, 60–72. doi:10.1016/j.canlet.2020.10.050

Lei, B., Huang, Y., Zhou, Z., Zhao, Y., Thapa, A. J., Li, W., et al. (2019). Circular RNA hsa_circ_0076248 promotes oncogenesis of glioma by sponging miR-181a to modulate SIRT1 expression. *J. Cell. Biochem.* 120, 6698–6708. doi:10.1002/jcb.27966

Li, H., Feng, Z., Wu, W., Li, J., Zhang, J., and Xia, T. (2013). SIRT3 regulates cell proliferation and apoptosis related to energy metabolism in non-small cell lung cancer cells through deacetylation of NMNAT2. *Int. J. Oncol.* 43, 1420–1430. doi:10.3892/ijo.2013.2103

Li, Y., Zhang, J., He, J., Zhou, W., Xiang, G., and Xu, R. (2016). MicroRNA-132 cause apoptosis of glioma cells through blockade of the SREBP-1c metabolic pathway related to SIRT1. *Biomed. Pharmacother.* 78, 177–184. doi:10.1016/j.biopha.2016.01.022

Linz, U. (2009). Commentary on effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol.* 10, 459–466.

Liu, C. L., Zhao, D., Li, J. J., Liu, S., An, J. J., Wang, D., et al. (2022). Inhibition of glioblastoma progression by Urolithin A *in vitro* and *in vivo* by regulating Sirt1-FOXO1 axis via ERK/AKT signaling pathways. *Neoplasma* 69, 80–94. doi:10.4149/neo_2021_210623N834

Luo, K., Huang, W., and Tang, S. (2018). Sirt3 enhances glioma cell viability by stabilizing Ku70-BAX interaction. *Onco. Targets. Ther.* 11, 7559–7567. doi:10.2147/OTT.S172672

Mensali, N., and Inderberg, E. M. (2022). Emerging biomarkers for immunotherapy in glioblastoma. *Cancers (Basel)* 14, 1940. doi:10.3390/cancers14081940

Mu, P., Liu, K., Lin, Q., Yang, W., Liu, D., Lin, Z., et al. (2019). Sirtuin 7 promotes glioma proliferation and invasion through activation of the ERK/STAT3 signaling pathway. *Oncol. Lett.* 17, 1445–1452. doi:10.3892/ol.2018.9800

O'Callaghan, C., and Vassilopoulos, A. (2017). Sirtuins at the crossroads of stemness, aging, and cancer. *Aging Cell* 16, 1208–1218. doi:10.1111/acel.12685

Ostrom, Q. T., Cioffi, G., Waite, K., Kruchko, C., and Barnholtz-Sloan, J. S. (2021). CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the United States in 2014-2018. *Neuro. Oncol.* 23, iii1–iii105. doi:10.1093/neuonc/noab200

Park, H. K., Hong, J. H., Oh, Y. T., Kim, S. S., Yin, J., Lee, A. J., et al. (2019). Interplay between TRAP1 and sirtuin-3 modulates mitochondrial respiration and oxidative stress to maintain stemness of glioma stem cells. *Cancer Res.* 79, 1369–1382. doi:10.1158/0008-5472.CAN-18-2558

Park, K. J., Yu, M. O., Park, D. H., Park, J. Y., Chung, Y. G., and Kang, S. H. (2016). Role of vincristine in the inhibition of angiogenesis in glioblastoma. *Neurol. Res.* 38, 871–879. doi:10.1080/01616412.2016.1211231

Qi, Y., Jin, C., Qiu, W., Zhao, R., Wang, S., Li, B., et al. (2022). The dual role of glioma exosomal microRNAs: Glioma eliminates tumor suppressor miR-1298-5p via exosomes to promote immunosuppressive effects of MDSCs. *Cell Death Dis.* 13, 426. doi:10.1038/s41419-022-04872-z

Rottenberg, S., Disler, C., and Perego, P. (2021). The rediscovery of platinum-based cancer therapy. *Nat. Rev. Cancer* 21, 37–50. doi:10.1038/s41568-020-00308-y

Sun, R., Zhang, Z., Bao, R., Guo, X., Gu, Y., Yang, W., et al. (2022). Loss of SIRT5 promotes bile acid-induced immunosuppressive microenvironment and hepatocarcinogenesis. *J. Hepatol.* 77 (2), 453–466. doi:10.1016/j.jhep.2022.02.030

Tang, J. H., Yang, L., Chen, J. X., Li, Q. R., Zhu, L. R., Xu, Q. F., et al. (2019). Bortezomib inhibits growth and sensitizes glioma to temozolomide (TMZ) via down-regulating the FOXM1-Survivin axis. *Cancer Commun.* 39, 81. doi:10.1186/s40880-019-0424-2

Teixeira, C. S. S., Cerqueira, N., Gomes, P., and Sousa, S. F. (2020). A molecular perspective on sirtuin activity. *Int. J. Mol. Sci.* 21, 8609. doi:10.3390/ijms21228609

Tian, T., Liang, R., Erel-Akbaba, G., Saad, L., Obeid, P. J., Gao, J., et al. (2022). Immune checkpoint inhibition in GBM primed with radiation by engineered extracellular vesicles. *ACS Nano* 16, 1940–1953. doi:10.1021/acsnano.1c05505

Wang, G., Wang, J. J., Wang, Y. Z., Feng, S., Jing, G., and Fu, X. L. (2018). Myricetin nanoliposomes induced SIRT3-mediated glycolytic metabolism leading to glioblastoma cell death. *Artif. Cells Nanomed. Biotechnol.* 46, S180–s91. doi:10.1080/21691401.2018.1489825

Xu, D. H., Chi, G. N., Zhao, C. H., and Li, D. Y. (2018). *Retracted* : Long noncoding RNA MEG3 inhibits proliferation and migration but induces autophagy by regulation of Sirt7 and PI3K/AKT/mTOR pathway in glioma cells. *J. Cell. Biochem.* 120, 7516–7526. doi:10.1002/jcb.28026

Yalçın, G. D., and Colak, M. (2020). SIRT4 prevents excitotoxicity via modulating glutamate metabolism in glioma cells. *Hum. Exp. Toxicol.* 39, 938–947. doi:10.1177/0960327120907142

Ye, T., Wei, L., Shi, J., Jiang, K., Xu, H., Hu, L., et al. (2019). Sirtuin1 activator SRT2183 suppresses glioma cell growth involving activation of endoplasmic reticulum stress pathway. *BMC Cancer* 19, 706. doi:10.1186/s12885-019-5852-5

Yihan, L., Xiaojing, W., Ao, L., Chuanjie, Z., Haofei, W., Yan, S., et al. (2021). SIRT5 functions as a tumor suppressor in renal cell carcinoma by reversing the Warburg effect. *J. Transl. Med.* 19 (1), 521. doi:10.1186/s12967-021-03178-6

Zhang, L., Kim, S., and Ren, X. (2020). The clinical significance of SIRT2 in malignancies: A tumor suppressor or an oncogene? *Front. Oncol.* 10, 1721. doi:10.3389/fonc.2020.01721

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10, 1523. doi:10.1038/s41467-019-09234-6

Zhu, D., Sun, C., and Qian, X. (2019). MST1 suppresses viability and promotes apoptosis of glioma cells via upregulating SIRT6 expression. *J. Integr. Neurosci.* 18, 117–126. doi:10.31083/j.jin.2019.02.122

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 16 January 2023
DOI 10.3389/fgene.2022.1008502

Check for updates

OPEN ACCESS

EDITED BY
Claudia Cava,
National Research Council (CNR), Italy

REVIEWED BY
Guoliang Cui,
Second Affiliated Hospital of Nanjing University of Chinese Medicine, China
Muhammad Khan,
Guangzhou Medical University Cancer Hospital, China

*CORRESPONDENCE
Muhammad Alaa Eldeen,
dr.muhammadalaa@gmail.com
Bonglee Kim,
bongleekim@khu.ac.kr

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

‡These authors have contributed equally to this work and share last authorship

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 05 August 2022
ACCEPTED 09 December 2022
PUBLISHED 16 January 2023

CITATION
Eid RA, Eldeen MA, Soltan MA, Al-Shraim M, Aldehri M, Alqahtani LS, Alsharif G, Albogami S, Jafri I, Fayad E, Park MN, Bibi S, Behairy MY, Kim B and Zaki MSA (2023), Integrative analysis of WDR12 as a potential prognostic and immunological biomarker in multiple human tumors.
Front. Genet. 13:1008502.
doi: 10.3389/fgene.2022.1008502

COPYRIGHT
© 2023 Eid, Eldeen, Soltan, Al-Shraim, Aldehri, Alqahtani, Alsharif, Albogami, Jafri, Fayad, Park, Bibi, Behairy, Kim and Zaki. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Integrative analysis of WDR12 as a potential prognostic and immunological biomarker in multiple human tumors

Refaat A. Eid[1†], Muhammad Alaa Eldeen[2*†], Mohamed A. Soltan[3†], Mubarak Al-Shraim[1], Majed Aldehri[4], Leena S. Alqahtani[5], Ghadi Alsharif[6], Sarah Albogami[7], Ibrahim Jafri[7], Eman Fayad[7], Moon Nyeo Park[8], Shabana Bibi[9,10], Mohammed Y. Behairy[11], Bonglee Kim[8*‡] and Mohamed Samir A. Zaki[4,12‡]

[1]Pathology Department, College of Medicine, King Khalid University, Abha, Saudi Arabia, [2]Cell Biology, Histology & Genetics Division, Biology Department, Faculty of Science, Zagazig University, Zagazig, Egypt, [3]Department of Microbiology and Immunology, Faculty of Pharmacy, Sinai University, Ismailia, Egypt, [4]Anatomy Department, College of Medicine, King Khalid University, Abha, Saudi Arabia, [5]Department of Biochemistry, College of Science, University of Jeddah, Jeddah, Saudi Arabia, [6]College of Clinical Laboratory Sciences, King Saud bin Abdulaziz University for Health Sciences, Jeddah, Saudi Arabia, [7]Department of Biotechnology, College of Sciences, Taif University, Taif, Saudi Arabia, [8]Department of Pathology, College of Korean Medicine, Kyung Hee University, Seoul, South Korea, [9]Department of Biosciences, Shifa Tameer-e-Millat University, Islamabad, Pakistan, [10]Yunnan Herbal Laboratory, College of Ecology and Environmental Sciences, Yunnan University, Kunming, Yunnan, China, [11]Department of Microbiology and Immunology, Faculty of Pharmacy, University of Sadat City, Sadat City, Egypt, [12]Department of Histology and Cell Biology, College of Medicine, Zagazig University, Zagazig, Egypt

Background: Mammalian WD-repeat protein 12 (WDR12), a family member of proteins containing repeats of tryptophan-aspartic acid (WD), is a potential homolog of yeast Ytm1p and consists of seven repeats of WD.

Aim of the study: This study aims to investigate the potential oncogenic effects of WDR12 in various human malignancies throughout a pan-cancer analysis that has been carried out to examine the various patterns in which this gene is expressed and behaves in tumor tissues.

Methods: Herein, we used The Cancer Genome Atlas (TCGA) and various computational tools to explore expression profiles, prognostic relevance, genetic mutations, immune cell infiltration, as well as the functional characteristics of WDR12 in multiple human cancers.

Results: We found that WDR12 was inconsistently expressed in various cancers and that variations in WDR12 expression predicted survival consequences for cancer patients. Furthermore, we observed a significant correlation between WDR12 gene mutation levels and the prognosis of some tumors. Furthermore, significant correlations were found between WDR12 expression patterns and cancer-associated fibroblast (CAF) infiltration, myeloid-derived suppressor cells (MDSCs), tumor mutation burden, microsatellite instability and

immunoregulators. Ultimately, pathway enrichment analysis revealed that WDR12-related pathways are involved in carcinogenesis.

Conclusions: The findings of our study are stisfactory, demonstrating that WDR12 could serve as a promising reliable prognostic biomarker, as well as a therapeutic target for novel cancer therapeutic approaches.

KEYWORDS

WDR12, pan-cancer, prognosis, immunotherapy, expression

1 Introduction

Cancer is widely recognized as one of the most significant public health concerns due to its role as one of the prominent causes of mortality and morbidity (Ohn et al., 1997). Tumorigenesis is a complex process governed by the origin of cells, the location of the tumor, genetic alterations, and acquired and hereditary molecular changes (Rohrmoser et al., 2007). Even though several medications and therapies have been created to treat cancer, several issues, including medication side effects, drug resistance, high healthcare costs, and missed targets, are still present (Zhang et al., 2017). As a result, there is still an urgent need to research more reliable tumor biomarkers and spotlight the specific biochemical pathways implicated in the development of tumors.

Mammalian WD repeat protein 12 (WDR12), a member of the tryptophan-aspartic acid (WD) repeat-containing proteins, is a possible homolog of yeast Ytm1p and contains seven repeats of WD. WDR12 transcript is largely expressed throughout embryogenesis, with the highest possible level found in the thymus and testis of mouse models (Hölzel et al., 2005). WDR12 forms a stable complex with Pes1 and Bop1 called PeBoW, which is essential for the biogenesis of mammal ribosomes. As a component of PeBoW, WDR12 plays a vital role in processing the 32S precursor ribosomal RNA and cell growth (Hölzel et al., 2005; Rohrmoser et al., 2007). Furthermore, WDR12 is required for the formation of the 28S and 5.8S ribosomal RNAs, as well as the development of the 60S ribosomal subunit (Nagy et al., 2021). According to the findings, inhibiting the expression of WDR12 in neonatal myocytes led to a reduction in the activation of ERK1/2, p38 MAPK, and HSP27. This discovery might shed light on how WDR12 regulates cell proliferation and survival (Moilanen et al., 2015). However, with these sophisticated discoveries, the significance of WDR12 in other physiological systems is still largely a mystery (Lerch-Gaggl et al., 2002). It was found to be preferentially expressed in glioma stem-like cells (GSCs), a subset of tumor cells that initiate malignant growth and promote the therapeutic resistance of glioblastoma, compared to non-stem tumor cells and normal brain cells. Moreover, WDR12 depletion compromised GSC proliferation, inhibited GSC-derived orthotopic tumor growth, and extended animal survival (Mi et al., 2021). Another study that correlated glioblastoma with the overexpression of WDR12 demonstrated that the knockdown of WDR12 promoted apoptosis and inhibited the proliferation of tumor cells (Li et al., 2020a). Hepatocellular carcinoma was another type of tumor that experienced up-regulated expression of WDR12 where the knockdown of WDR12 expression resulted in reduced proliferation and migration of HepG2 and Huh-7 cells (Li et al., 2020b).

In recent years, various public cancer databases, such as The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO), have been constructed as a result of the fast development of whole genome sequencing technology (Blum et al., 2018; Hutter and Zenklusen, 2018). We can now assess the clinical prognosis of genes of interest through a pan-cancer analysis thanks to the open access aspect of these publicly available databases. This allows us to compare and contrast the commonalities and differences among human malignancies (Cui et al., 2020). We utilized the TCGA project and the publicly accessible bioinformatics tools to implement a pan-cancer analysis of WDR12 to examine the roles and underlying mechanisms of WDR12 in the pathophysiology or clinical prognosis of malignant tumors. This included expression variations, correlations between expression profiles and survival, genetic modifications, immune cell infiltration, tumor mutation burden (TMB), microsatellite instability (MSI), immunoregulators, as well as relevant molecular pathways.

2 Materials and methods

2.1 WDR12 expression analysis

WDR12 expression profiles were examined throughout the "Gene DE" module of TIMER2 (tumor immune estimation resource, version 2) (http://time.cistrome.org/). For those without normal tissues in the TIMER2 database, we matched the TCGA and GTEx databases using the GEPIA2 (Gene Expression Profiling Interactive Analysis, version 2) web's Expression Analysis-Box plots module (http://gepia2.cancer-pku.cn/#analysis). This allowed us to examine the differences in WDR12 expression between multiple malignancies and correlate with normal tissues (Tang et al., 2017). Furthermore, the 'Pathological Stage Plot' module of GEPIA2 was selected to

examine the WDR12 expression profile at various pathological stages. Furthermore, the CPTAC module of UALCAN (http://ualcan.path.uab.edu/analysis.html) was used to calculate the levels of WDR12 protein expression in cancers and normal tissues that were directly analogous to these malignancies (Chandrashekar et al., 2017).

2.2 Immunohistochemistry staining

IHC images of WDR12 protein expression in normal tissues and seven cancerous tissues, including LUAD, BRCA, LIHC, Glioma, KICH, and COAD, were obtained from the HPA (Human Protein Atlas) (http://proteinatlas.org/) and analyzed to assess differences in WDR12 expressions.

2.3 Survival prognosis analysis

Initially, the GEPIA2.0 website was used to evaluate patient survivability, where WDR12 was tested in the "survival analysis" section; literally, the entire tumor listed in the TCGA cohort was chosen, and the heatmap was retrieved for the two possible server-side approaches was retrieved [the overall survival (OS) and disease-free survival (DFS)].

2.4 Gene alteration analysis

We were able to elucidate the genetic modification characteristics of WDR12 by using the "Cancer Types Summary" module that is available in the Web version of the cBioPortal (https://www.cbioportal.org) (Gao et al., 2013). The alteration frequency, the mutation type, and the Copy number alteration (CNA) of WDR12 in multiple human malignancies were retrieved after selecting "TCGA Pan-Cancer Atlas Studies" from the "Quick select" menu. This allowed the data to be more easily accessed. In particular, we utilized the 'Comparison' module of cBioPortal to identify the association between WDR12 genetic alteration and survival prognosis with or without WDR12 gene mutations. This was done to determine whether there was a significant difference between the two scenarios. The Mutations module highlighted the different protein structure domains due to genetic variations.

2.4.1 The methylation analysis of WDR12

DNA methylation is a crucial element of the cellular epigenetic regulatory system, which governs gene expression profiles (Qu et al., 2013). The investigation of WDR12 methylation was conducted with the help of three separate databases. These databases were UALCAN (which can be accessed at http://ualcan.path.uab.edu/index.html), VShiny Methylation Analysis Resource Tool (SMART) (which can be accessed at http://www.bioinfozs.com/smartapp/), and MEXPRESS (which can be accessed at https://www.mexpress.be/) (Gao et al., 2013; Chandrashekar et al., 2017; Hutter and Zenklusen, 2018). We examined the levels of promoter methylation of WDR12 in both normal and cancerous tissues by using the "TCGA analysis" module included in UALCAN. On its website, SMART featured visual representations of DNA methylation locations, as well as CpG-aggregated methylation. Comparative analyses of promoter methylation levels in LGG and GBM were carried out with the help of the MEXPRESS database.

2.5 Immune cell infiltration analysis

We used the "Immune Gene" module of the TIMER2 Web server to investigate the relationship between WDR12 expression and immune infiltrates in every TCGA tumor. As immune cells, myeloid-derived suppressor cells (MDSCs) and cancer-associated fibroblasts (CAFs) were selected. Regarding measuring immunological infiltration, we used algorithms developed by TIMER, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, XCELL, and MCPCOUNTER, and EPIC. Using the purity-adjusted Spearman rank correlation test, the *p*-values and partial correlation (cor) values were calculated. The findings were represented with a heat map and a scatter plot.

2.6 TMB and MSI analysis

TMB is a quantifiable immune-response biomarker identified by calculating the total number of gene coding errors, base substitutions, gene insertion, or deletion errors per million bases (Yarchoan et al., 2017). MSIs, short, tandemly repeated DNA sequences, are caused by DNA mismatch repair deficiency in cancer tissues. MSI with defective DNA mismatch repair is a clinically significant tumor biomarker (van Velzen et al., 2020) TMB and MSI scores were generated using somatic mutation data obtained from the TCGA database (https://tcga.xenahubs.net). Additionally, using Spearman's method, we investigated whether there was a connection between the expression of WDR12, TMB, and MSI.

2.7 Correlation analysis between WDR12 and immunoregulators

An investigation of the relationship between WDR12 and immunomodulators, immunostimulators, immunoinhibitors, MHC molecules, TILs, receptors, and chemokines in various cancers was carried out with the help of the TISIDB portal (http://cis.hku.hk/TISIDB/).

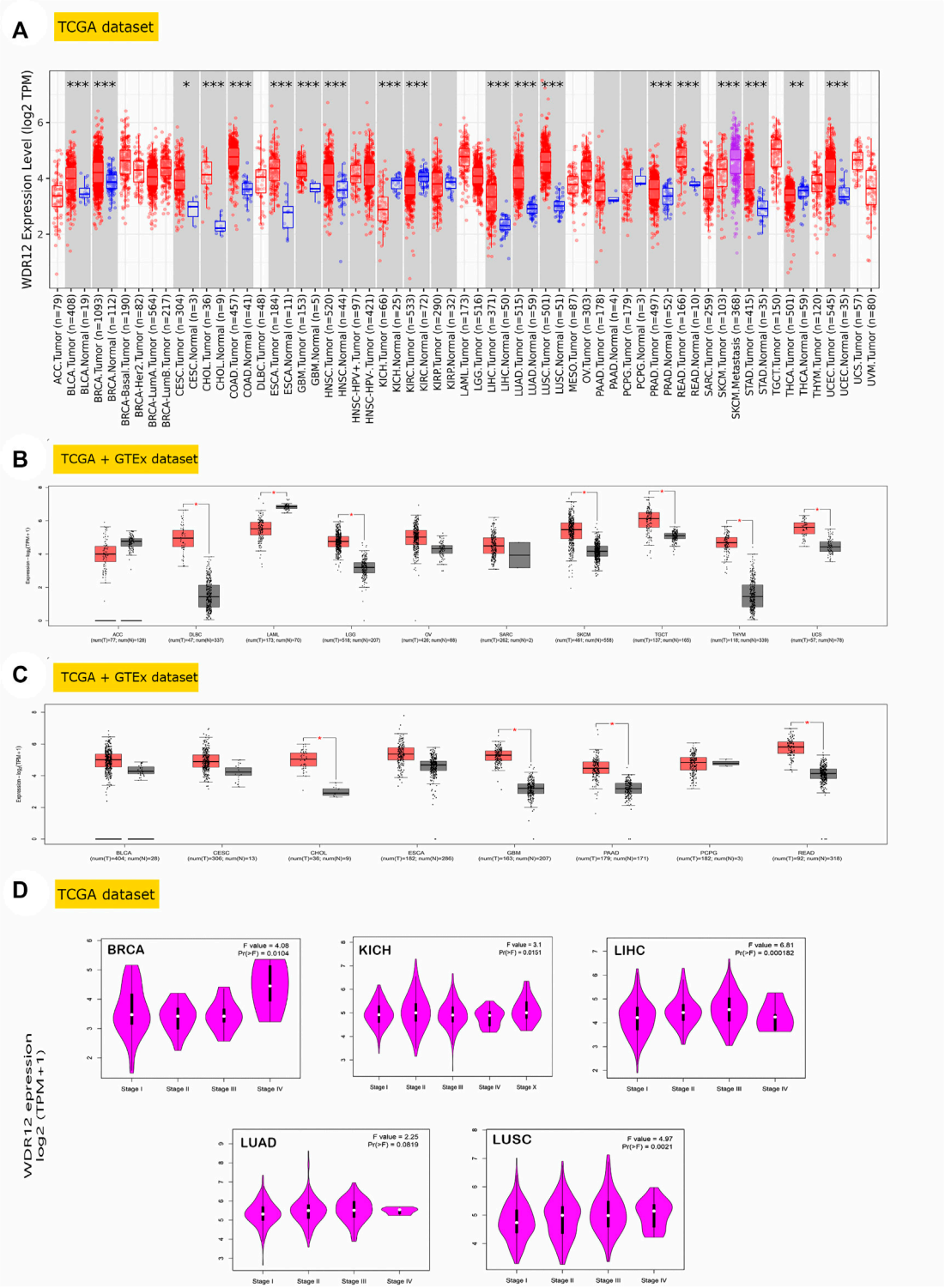


FIGURE 1
WDR12 expression assessment in human cancers. **(A)** Differential expression of WDR12 in a panel of TCGA tumors analyzed by TIMER2.0; **(B)** tumors that lack normal tissue for comparison in TIMER2.0. Database and experienced a trend of elevated WDR12 expression in tumor versus normal tissue when analyzed in the GEPIA database; **(C)** tumors that underwent a significant WDR12 protein expression versus normal tissue; **(D)** tumors experienced a positive correlation between WDR12 expression and the tumor stage.

2.8 Gene set enrichment analysis of WDR12

We utilized the STRING database to retrieve experimentally validated WDR12-binding proteins (https://string-db.org/). We specified the minimum desired interaction scoring rate to 'Low confidence,' the meaning of network edges to 'evidence,' the highest possible number of interactors to display to 'not over than 50 interactors,' and the active interaction sources to 'experiments'. Furthermore, GEPIA2's 'Similar Gene' module was used to identify the top 100 WDR12-correlated genes as well. As part of GEPIA2, correlation analysis was performed on WDR12 and these other genes. To retrieve heatmap data for the genes listed above, we employed TIMER2's 'Gene_Corr' module. The WDR12 binding and interacting genes were analyzed using the Venn diagram viewer (http://bioinformatics.psb.ugent.be/webtools/Venn/).

WDR12 and related genes were analyzed for the GO and KEGG pathway enrichment using the DAVID database (https://david.ncifcrf.gov/home.jsp) (database for annotations, visualization, and integrated discoveries). Visualization and depiction of the pathways were accomplished using ggplot2 R packages.

2.9 Statistical analysis

The log2 transformation was used to normalize all gene expression profiles. Tumor samples and normal tissues were compared using a Wilcox test to determine the WDR12 expression differences. The KM curve was applied for the survival analysis of tumor patients. The partial Spearman method was used to investigate the relationship between two variables. To perform all statistical analyzes, we used R software (Version 4.0.2). Statistical significance was defined as $p < 0.05$.

3 Results

3.1 Upregulation of WDR12 in multiple human cancers

As part of our investigation of the particular function of WDR12 in cancerous tumors, we compared the levels of WDR12 expression found in various sections of the TCGA database. Using the TIMER2 tool, we found that WDR12 was highly expressed in BLCA, BRCA, CESC, GBM, CHOL, COAD, ESCA, GBM, HNSC, LIHC, LUAD, LUSC, PRAD, READ, STAD, and UCEC (Figure 1A). Additionally, we used the GTEx dataset to further corroborate the expression profiles of WDR12 in a few more cancers that either did not have normal tissue matched in the TIMER2 dataset or had a small sample number (less than 20) of normal tissue that was used for expression comparison. From the tumor list that lacks a normal sample for comparison on the TIMER2 web server, DLBC, LGG, SKCM, TGCT, THYM, and UCS experienced a significant upregulation in WDR12 in comparison to normal tissue (Figure 1B). Moving to the tumors that had a small "normal tissue" sample size in the TIMER2 web server, all of the analyzed tumors showed a trend of WDR12 elevated expression in tumor *versus* normal samples, and this elevation was significant in tumors CHOL, GBM, READ, and PAAD (Figure 1C). It is worth noting that the sample size of "normal tissue" was less than 20 in 3 tumors, namely CESC, CHOL, and PCPG.

GEPIA2's "Pathological Stage Plot" module investigated the relationship between WDR12 expression and tumor pathology stages. Statistically significant correlations were found in BRCA, KISH, LIHC, LUAD, and LUSC (Figure 1D). Moving to WDR12 protein level analysis, as shown in Figures 2A–E, WDR12 protein expression was significantly elevated in breast cancer, glioblastoma multiforme, hepatocellular carcinoma, LUAD, and colon cancer ($p < 0.001$). Moreover, we evaluated the IHC data provided by the HPA database and compared the findings with the differential protein expression. The results of the data analyses from these two databases were roughly equivalent. Normal lung, breast, brain, liver, and colon tissues were negative or moderately stained with WDR12 IHC, while cancerous tissues were moderately or strongly stained (Figures 2A–E).

3.2 WDR12 expression level exhibited a significant prognostic value

We used GEPIA2 web server to conduct a survival correlation study for each tumor. This was done to get a deeper insight into the relationship between WDR12 expression and the clinical outcome of cancer patients with various types of tumors. As shown in Figure 3A, those with high WDR12 expression levels were significantly linked to poor overall survival (OS) in ACC ($p = 0.008$), KIRP ($p = 0.031$), LIHC ($p = 0.013$), and LUAD ($p = 0.028$). However, high WDR12 expression was associated with longer OS in patients with ESCA ($p = 0.2$), KIRC ($p = 0.23$), READ ($p = 0.24$), and THYM ($p = 0.86$) (Fig. S1). In terms of DFS, highly expressed WDR12 was remarkably correlated with poor DFS in ACC ($p = 1.1e\text{-}5$), KIRP ($p = 0.011$), and LIHC ($p = 0.0048$), while downregulation of WDR12 has poor DFS in LUAD ($p = 0.0062$), and LUSC ($p = 0.018$) (Figure 3B).

3.3 The genetic mutation analysis of WDR12

We analyzed WDR12 genetic alterations in multiple tumor samples from the TCGA database using the cBioPortal website.

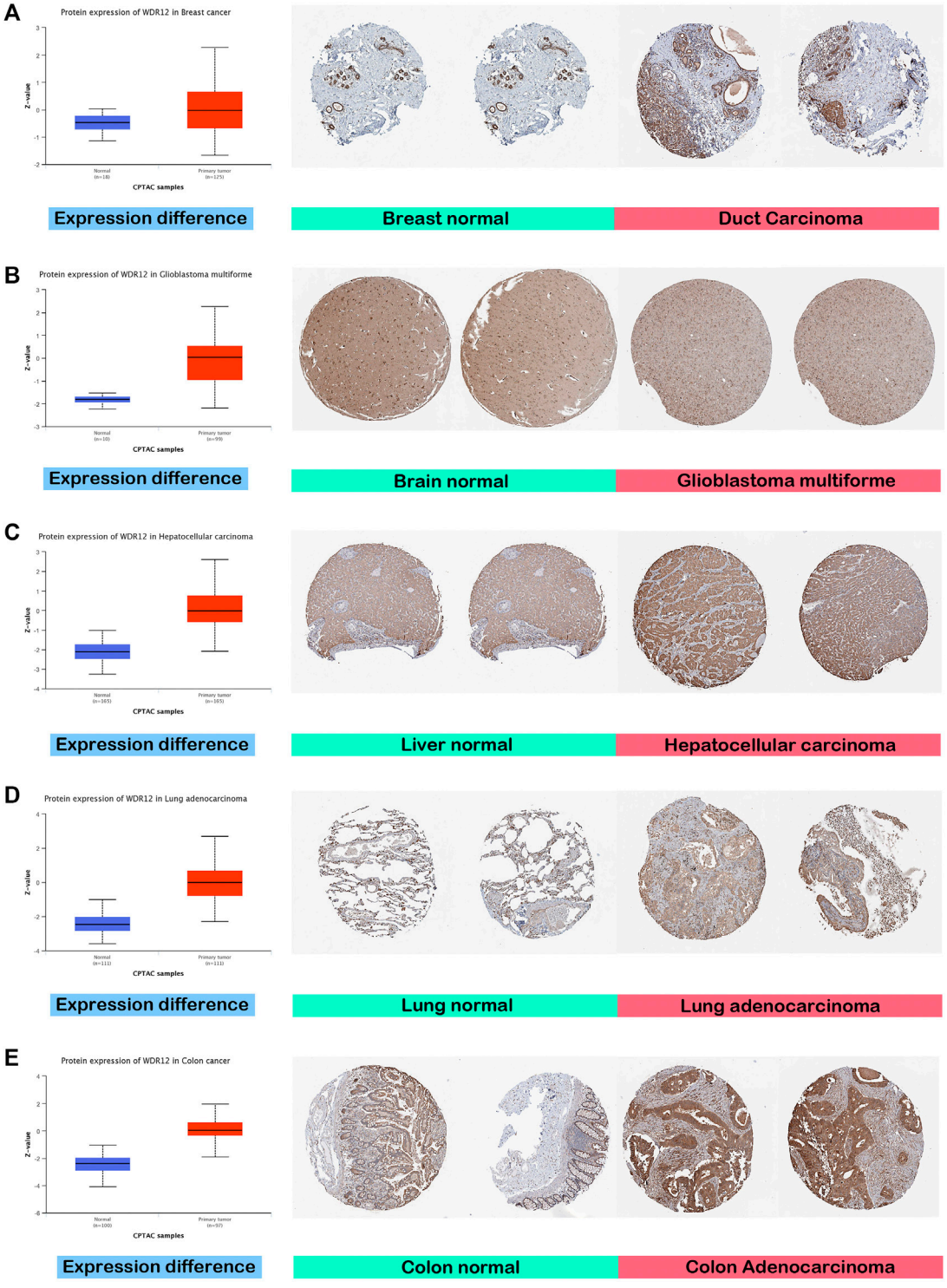


FIGURE 2
Tumors experienced a statistically significant higher WDR12 protein expression in the tumor sample versus the normal one (left side); IHC staining for normal tissue (middle) and cancerous one (left) demonstrated the same results. **(A)** Breast; **(B)** Brain; **(C)** Liver; **(D)** Lung; **(E)** Colon.

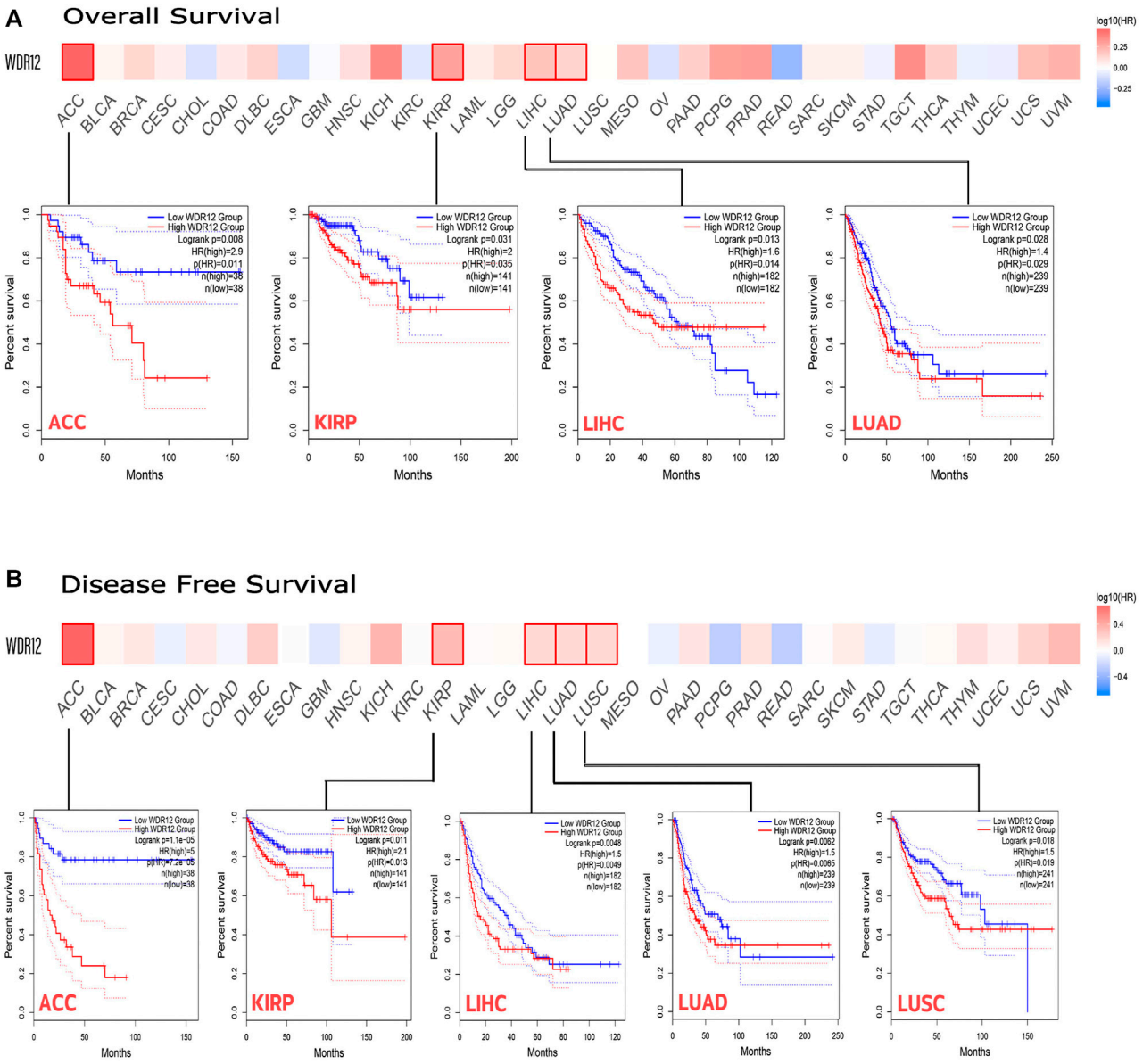


FIGURE 3
The correlation between WDR12 expression and the clinical outcome. **(A)** disease-free survival; **(B)** overall survival as assessed from the GEPIA database.

As illustrated in Figure 4A, the most prevalent CNA of WDR12 was amplification, whereas there were no deep deletions modifications. Additionally, missense mutations were predominantly seen in diploid and gain-type individuals. Analysis of mutation frequency revealed that missense mutations were the most common type of genetic modification (Figure 4B). Subsequently, we investigated WDR12 gene mutations and CNAs in human malignancies.

The results demonstrated that the most common type of WDR12 genetic modification is a gene mutation, which occurs most frequently in UCEC (~1% frequency). The "amplification" type of CNA was the predominant type in LUSC, accounting for around 2% of all changes. WDR12 gene alterations are the most prevalent (3.5% frequency) in UCEC (Figure 4C).

Figure 4D shows the types, locations, and frequency of WDR12 genetic modifications. We discovered that missense

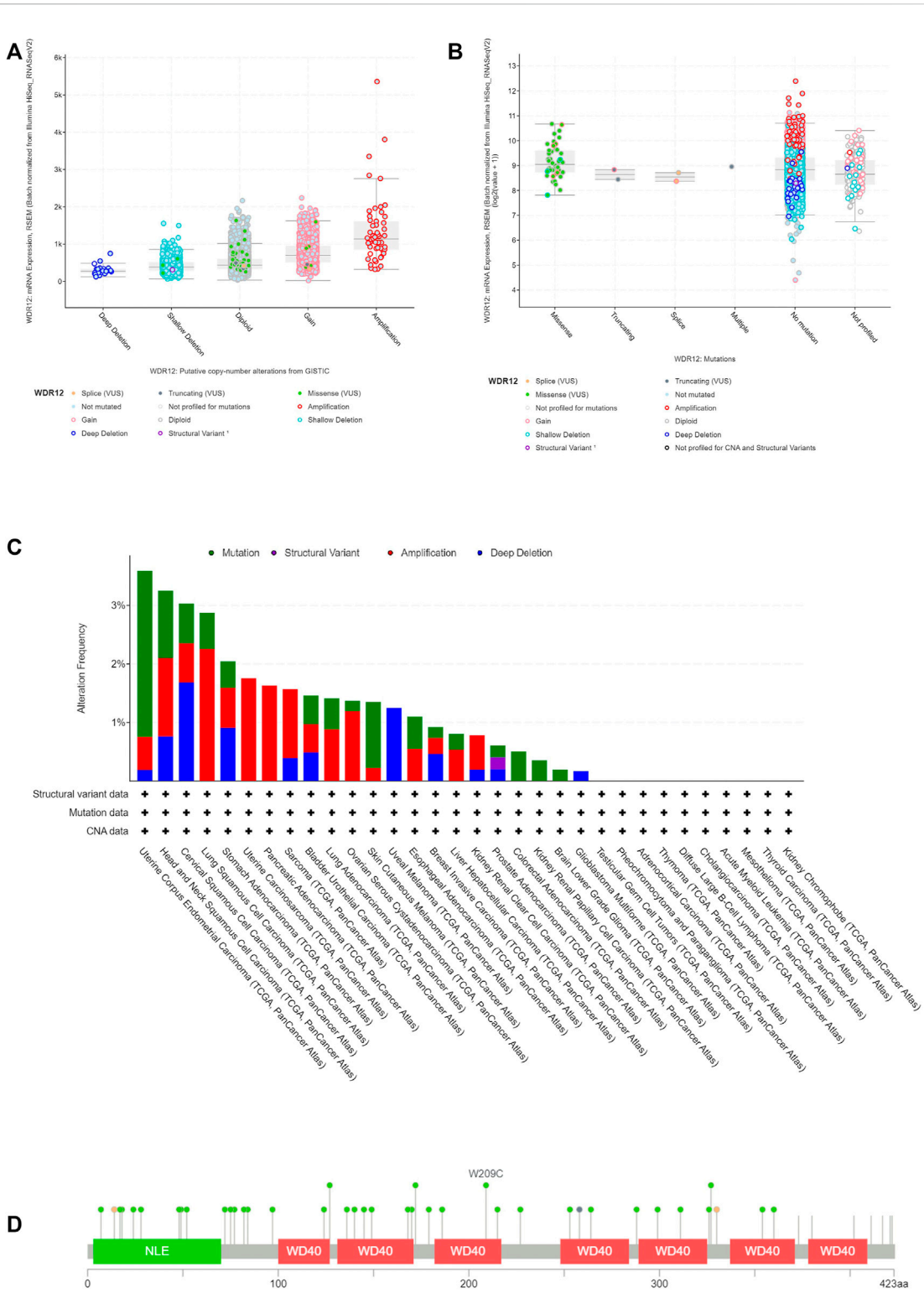


FIGURE 4
Mutation assessment for WDR12 using the cBioPortal tool. **(A)** Gene alteration analysis for WDR12; **(B)** mutation forms analysis for WDR12; **(C)** alteration frequency with mutation type in a panel of analyzed human cancers **(D)** map representation for sites and types of WDR12 mutations.

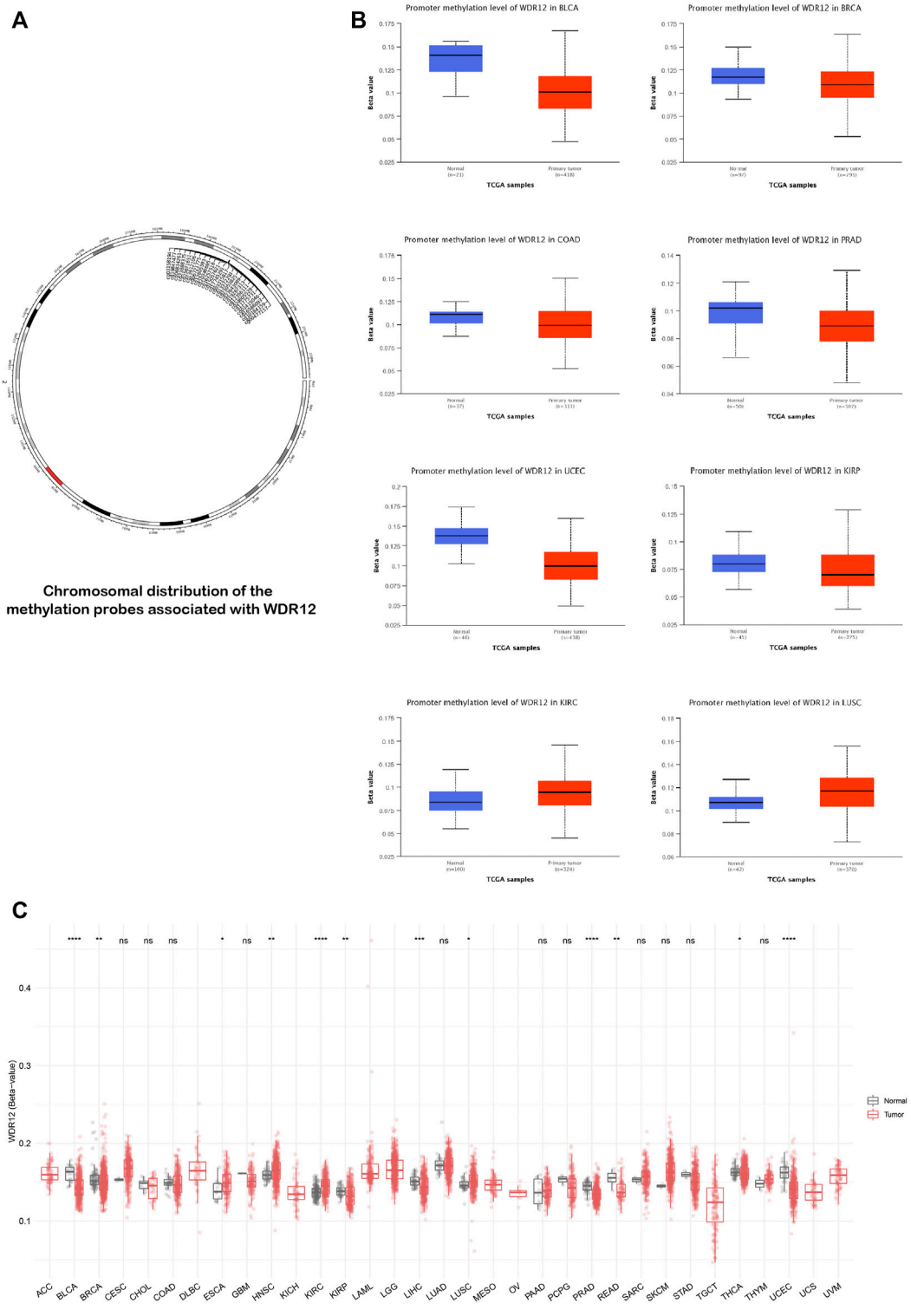


FIGURE 5
Differential methylation analysis of WDR12 in tumor samples versus normal ones. **(A)** Chromosomal distribution of methylation probes associated with WDR12; **(B)** tumors experienced higher methylation in the WDR12 promoter region in normal samples versus tumors as assessed by UALCAN analysis; **(C)** analysis of CpG-aggregated methylation of WDR12 in a list of human tumors.

mutations were more prevalent at the W209C site in the WD40 domain, which was the most modified site.

3.4 Methylation data analysis of WDR12 in human cancers

The status of gene methylation has been extensively studied in several human cancers. Previous studies generally found that DNA hypermethylation was a major mechanism for silencing tumor suppressor genes (Anglim et al., 2008). On the other hand, oncogenes experienced a hypomethylation status as a mechanism for their activation to induce tumor progression (Romero-Garcia et al., 2020). We performed a methylation data analysis to examine the possible correlation between WDR12 DNA methylation and the pathophysiology of various cancers. Figure 5A illustrates the chromosomal distribution of WDR12-linked methylation probes as estimated by SMART APP, an interactive online application for visualization and analysis of DNA methylation. BLCA, BRCA, KIRP, LIHC, PRAD, READ, THCA, and UCEC tissues had a lower level of WDR12 methylation than normal tissues (Figure 5C). Using UALCAN online tools, we investigated WDR12 promoter methylation levels in different malignancies. In all tumor groups with at least 20 samples, the WDR12 promoter was less methylated in BLCA, BRCA, COAD, PRAD, KIRP, and UCEC than in normal tissues. By comparison, the degree of WDR12 promoter methylation was greater in LUSC and KIRC (Figure 5B). Most significant results favored WDR12 hypomethylation in the tumor sample *versus* normal. There was no substantial difference between CESC and HNSC promoter methylation levels (Supplementary Figure S2A). We also used the MEXPRESS approach to investigate the association between WDR12 methylation levels and tumor pathogenesis. Surprisingly, when the glioma grade increased, the promoter methylation levels of WDR12 reduced. As demonstrated in Supplementary Figure S2B, the histogram reflecting the level of WDR12 methylation in the promoter regions was considerably lower in GBM than in LGG.

3.5 Immune cell infiltration analysis of WDR12

Previous studies have reported that the tumor microenvironment (TME) may play a crucial role in carcinogenesis and cancer cell immune evasion (Schumacher and Schreiber, 2015; Wu and Dai, 2017). Thus, TME significantly affects the treatment efficacy and overall prognosis of cancer. The TME is composed of tumor cells, immune cells, and stromal cells. Immune cells that enter the TME significantly affect tumorigenesis, progression, and metastasis (Anderson and Simon, 2020). CAFs and MDSCs have been shown to promote carcinogenesis in TMEs (Aizawa et al., 2019; de Cicco et al., 2020). Thus, we examined the relationship between WDR12 expression and tumor-infiltrating immune cells, such as CAFs and MDSCs. The TCGA database identified associations between WDR12 expression and MDSC and CAF infiltration levels. We identified significant positive correlations between the expression and MDSCs infiltration in all human cancers, excluding UVM, THCA, PCPG, KIRC, DLBC, BRCA-LumB, and BRCA-Her2. Several malignancies, including ACC, BRCA-LumA, CESC, CHOL, KISH, KIRP, LIHC, SKCM, and THYM, demonstrated significant positive correlations between the expression of WDR12 and the infiltration of CAFs (Figure 6A). Six tumors, including ACC, LIHC, LUAD, PRAD, READ, and SKCM, exhibited a significant positive correlation between WDR12 expression and MDSC and a negative correlation between the same gene expression and CAFs cell infiltration, as determined by filtration of the findings. Figure 6D shows scatter plots demonstrating the link between WDR12 expression and MDSCs infiltration level in these six tumors. MDSCs infiltration levels and WDR12 expression did not have statistically significant negative correlations; nevertheless, WDR12 expression and estimated CAFs infiltration values in STAD and TGCT did have statistically significant negative correlations (Supplementary Figure S3). Next, we investigated two additional types of immune cells but with established antitumor activity. The first cell type was NK T cells which experienced a negative correlation with the level of WDR12 and all tumors with significant results (Figure 6B). This finding confirmed the immunosuppressive role of WDR12 in tumor tissues. The second cell was CD8 which demonstrated a negative correlation with the level of WDR12 in the tumors HNSC, KIRC, KIRP, STAD, THCA, and UCS. It's worth noting that few tumors, including BRCA and UVM, showed a positive correlation between the level of WDR12 and CD8 infiltration (Figure 6C).

3.6 Correlation analysis of WDR12 expression with TMB and MSI

TMB is a promising therapeutic marker of the immunotherapy response. In addition, MSI acts as a biomarker for immune-checkpoint inhibitors, intimately linked to the advancement of the vast majority of malignancies (Fumet et al., 2020, Pino and Chung, 2011). MSIs in PRAD, LIHC, PAAD, CESC, BRCA, STAD, KIRC, KICH, and HNSC were positively correlated with WDR12 expression, while MSIs in DLBC were negatively correlated (Figure 7A). Our research results also revealed that WDR12 expression in LUAD, UCEC, PAAD, CESC, STAD, SKCM, HNSC, and ACC was significantly and positively correlated with TMB (Figure 7B).

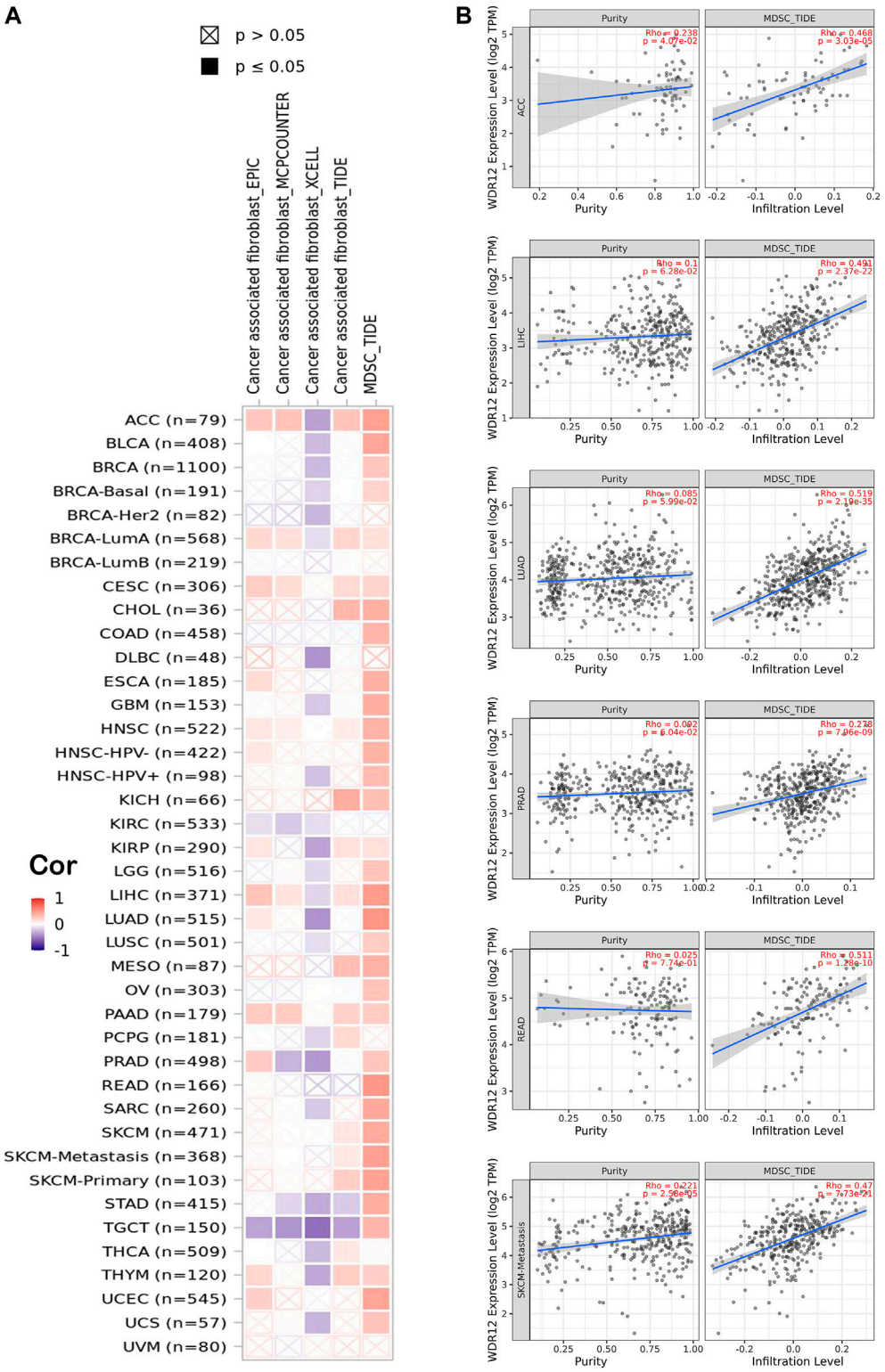


FIGURE 6
Correlation between WDR12 expression level and infiltration of **(A)** myeloid-derived suppressor cells (MDSC) and ;**(B)** Scatter plots demonstrate the correlation between the expression of WDR12 and the infiltration level of MDSC.

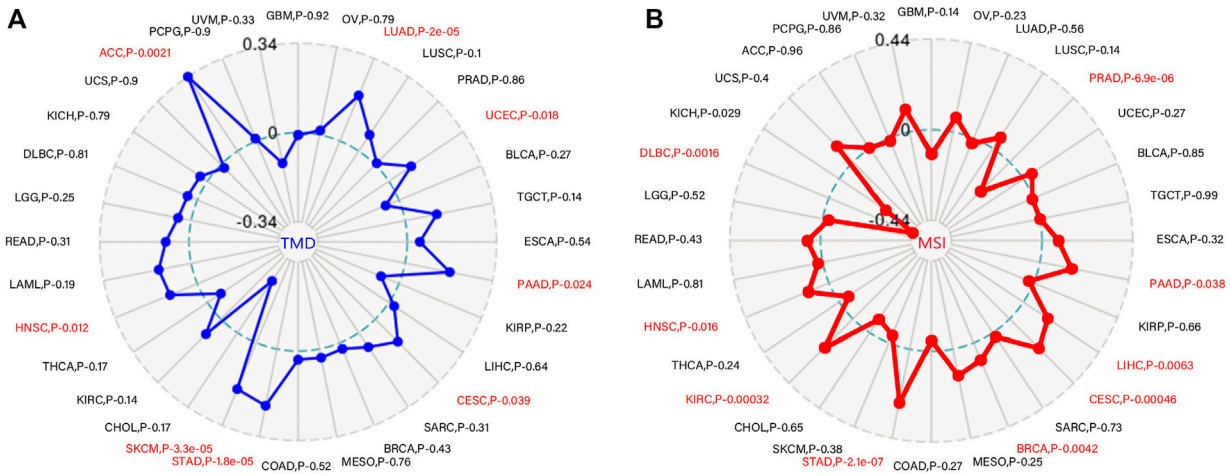


FIGURE 7
Correlations of WDR12 expression with MSI and TMB. **(A, B)** are radar charts showing the overlaps of WDR12 with TMB and MSI, respectively.

3.7 Correlation analysis of WDR12 expression with immunoregulators

The co-expression of WDR12, immune-related genes, and tumor-infiltrating lymphocytes (TILs) was investigated. All the immune-related genes studied encoded MHC, immunological activation, immunosuppression, chemokine, and chemokine receptor proteins. Our findings revealed that approximately all immune-related genes were significantly correlated with WDR12, and the majority were adversely associated with WDR12 in most malignancies (Figures 8A–F). To make a close investigation for the correlation between WDR12 and immune-related molecules, we selected ACC as a tumor model and molecules belonging to three classes, namely CD276 (immunostimulator), IL10RB (immunoinhibitor), and TAP2 (MHC-related molecule) where we found a significant correlation between these molecules and the expression level of WDR12 (Figures 8G–I). These results implicate that WDR12 can interfere with the tumor progression through its modulating effect on several classes of the immune-related molecule and consequently affects the immune response against a growing tumor. This characteristic has been explored in the last few years and resulted in the appearance of a novel line of cancer treatment, cancer immunotherapy (Finck et al., 2020).

3.8 Analysis of protein interactions and correlations with WDR12

Based on the results mentioned above, it is found that WDR12 has a clear association with cancer patients' survival and affects immune cells in the TME. Consequently, it is important to analyze the potential molecular mechanisms of this gene in several tumors.

As a result, we explored the possible underlying mechanisms of WDR12 in carcinogenesis by conforming to a network of fifty WDR12-binding protein interactions based on the STRING online database with supporting evidence from various experiments (Figure 9A). We then extracted the top 100 genes most strongly associated with WDR12 expression using GEPIA2. According to Figure 9B, WDR12 expression was positively correlated with NOP58 (R = 0.74), SGOL2 (R = 0.6), SUMO1 (R = 0.59), SUV39H2 (R = 0.58), and WDR75 (R = 0.59) (all $p < 0.001$). Furthermore, a heatmap generated through the 'Gene_Corr' module at TIMER confirmed the significant positive correlation between these five genes and WDR12 in the complete list of TCGA cancers (except CHOL, where the correlation with SUMO1 was insignificant) (Figure 9C). Furthermore, we performed an intersection analysis of the 50 WDR12 binding proteins and the top 100 correlated genes and found DKC1, NIP7, DDX18, NIFK, RPF2, and GTPBP4 were the common members through a Venn diagram (Figure 9D).

A new dataset was generated from the combination of two lists (WDR12 interacted genes from the STRING database and WDR12 associated genes from the GEPIA2 database), where this new list was subjected to enrichment analysis. Firstly, the KEGG pathway analysis reveals that "cell cycle" as well as "ribosome biogenesis" could be the potential methods by which WDR12 promotes carcinogenesis (Figure 9E). Secondly, GO enrichment analysis suggests that genes that directly interact with or were related to WDR12 were mainly

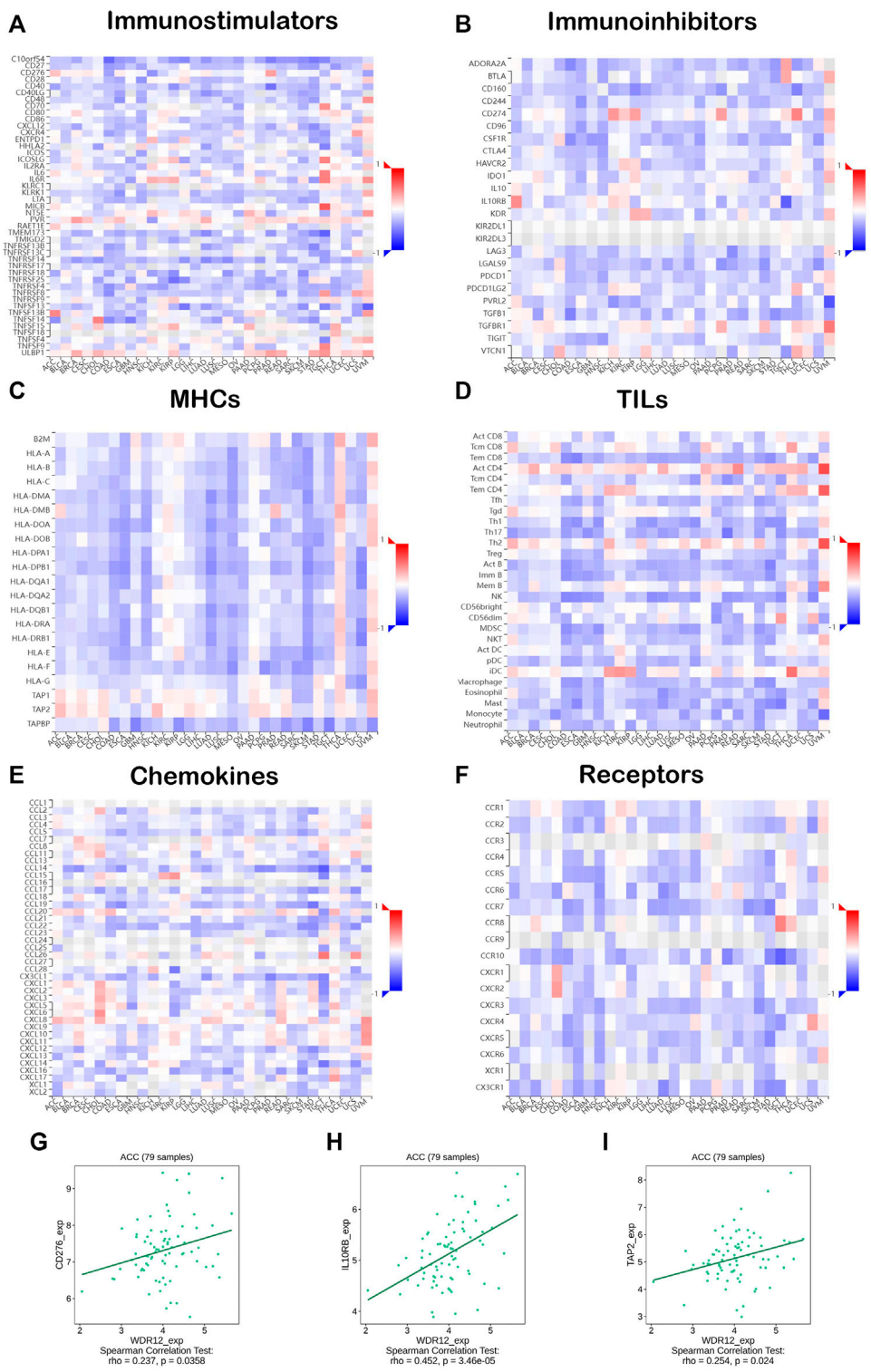


FIGURE 8
Correlation analysis of WDR12 expression with immunoregulators. Heatmaps correlating WDR12 with **(A)** Immunostimulators; **(B)** Immunoinhibitors; **(C)** MHCs; **(D)** TILs; **(E)** Chemokines; **(F)** Receptors. Plots **(G–I)** show the correlation between WDR12 and CD276, IL10RB, and TAP2, respectively.

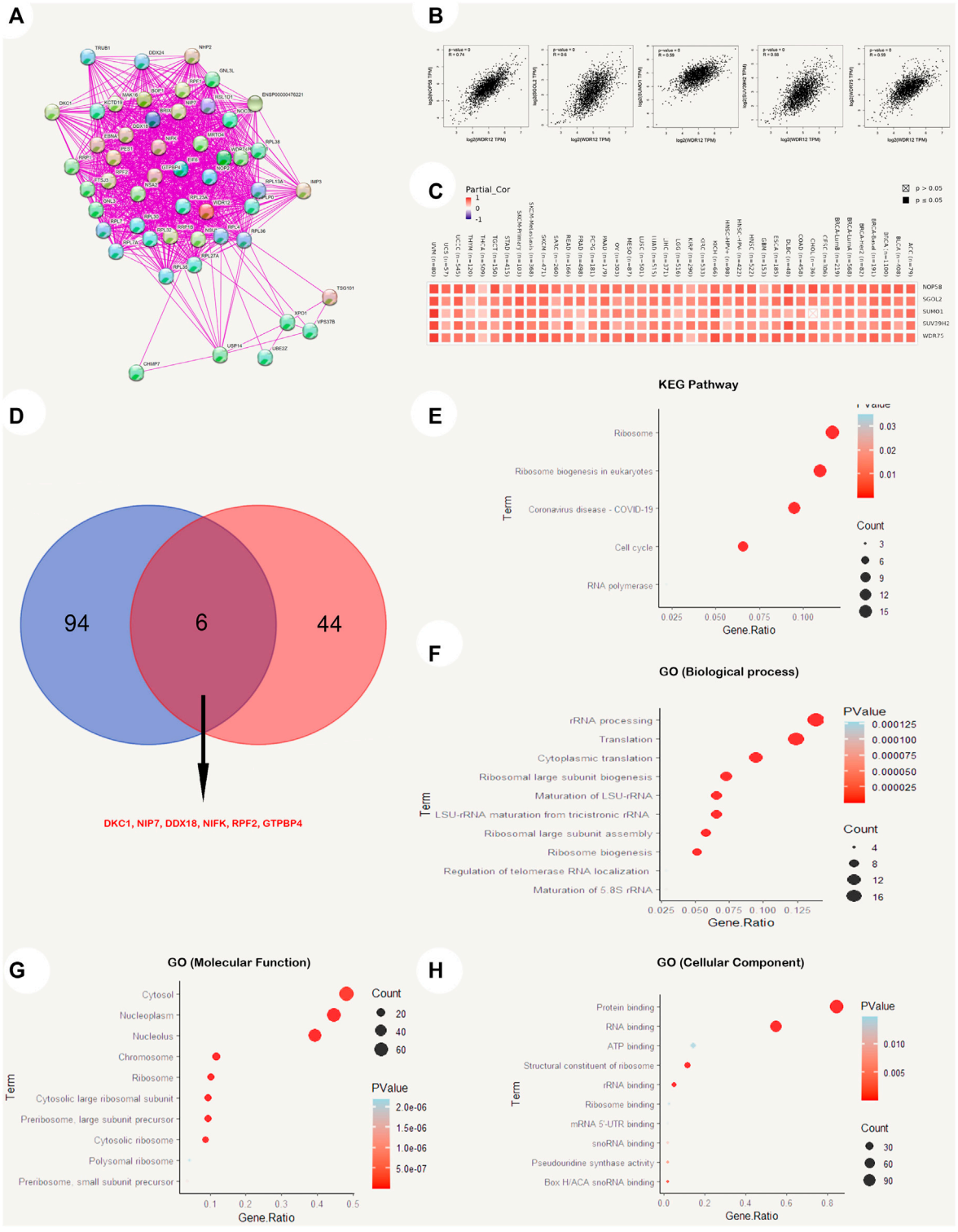


FIGURE 9
WDR12-protein network interactions. **(A)** A map of top 50 WDR12 interacting proteins as determined by STRING database; **(B)** expression correlation between WDR12 and genes (NOP58, SGOL2, SUMO1, SUV39H2 and WDR75) as determined by GEPIA2; **(C)** heatmap for top 5 WDR12-correlated proteins in the tumor tissue; **(D)** Venn diagram showing the intersection between WDR12 interacting and correlating proteins; **(E–H)** KEGG/GO enrichment analysis based on WDR12-binding and interacted genes.

correlated with the biological mechanisms of "rRNA processing", "cytoplasmic translation," and "ribosomal large subunit biogenesis" (Figure 9F), with the molecular functions of "protein binding" and RNA binding" (Figure 9G), and with the cellular components of the "cytosol", "nucleoplasm" and the "nucleolus" (Figure 9H). It is important to mention that the cutoff *p*-value for the significant results in the enrichment analysis was $p < 0.05$, where only the top significant results (with the lowest *p*-value) were visualized using R packages.

4 Discussion

Cancer is a serious public health problem, and its incidence continues to increase annually. Although standard surgical resection, radiation, and chemotherapy have been used successfully as treatment options, several cases failed to get a positive response with these treatment options and still suffering (Fernández, 2014). This complicated tumor status underscores the need for a comprehensive strategy that studies a specific gene with tumorigenesis at several points of analysis, an approach that might provide us with novel treatment options. WDR12 has been reported in several previous studies to be associated with cancer progression in several models of human tumors. In colorectal cancer, WDR12 was found to be associated with tumor progression by promoting proliferation and inhibiting apoptosis where *in-vivo* tumor formation experiments in nude mice revealed that tumor area, volume, and weight were significantly lower for knockdown WDR12 (Fernández, 2014). In addition to that, whole-genome gene expression profiling revealed WDR12 as an implicated gene in colorectal cancer (Sun and Qian, 2018). Moreover, WDR12 was correlated with tumor progression in other human cancers including hepatocellular carcinoma (Li et al., 2020b) and glioblastoma (Li et al., 2020a), and was studied as a target for natural anticancer products in the breast cancer model (Lewinska et al., 2017). To this end, the present work employed a pan-cancer analysis to determine the oncogenic behavioral patterns of WDR12. We began our investigation by studying the distribution of WDR12 in human tissue, which was discovered to be expressed in various organs. An essential hallmark of oncogenic proteins is their overexpression in tumor tissue compared to normal tissue; as a result, we evaluated the expression of WDR12 in various malignancies using the TCGA, GTEx, and CPTAC databases. As described in the results section, WDR12 expression was significantly higher in several tumors, such as DLBC, THYM, GBM, LUAD, and STAD, than in equivalent normal tissues.

Survival analysis is a fundamental method for assessing the development of a disease and the responsiveness of a patient to therapeutic interventions (Nagy et al., 2021). Increased WDR12 expression levels were strongly associated with poor OS in ACC, KIRP, LIHC, and LUAD; and poor DFS in ACC, KIRP, and LIHC. Additionally, WDR12 genetic modification levels were significantly correlated with the prognosis of several malignancies. It is worth mentioning that some tumors, including ESCA, KIRC, READ, and THYM, experienced a better OS with elevated levels of WDR12, a point that requires further investigation to study the role and mechanism of WDR12 in these tumors and shows the importance of the pan-cancer studies that can reveal these opposite effects of the same gene in various types of human cancers. Genetic mutations are universally acknowledged to play a significant role in carcinogenesis (Anandakrishnani et al., 2019). Numerous gene mutations were discovered to be useful prognostic indicators for tumorigenesis; for example, genetic mutations in BRAF, PIK3CA, TP53, KRAS, and NRAS were discovered to be correlated with extramural venous invasion in baseline MRI, as well as poor clinical tumor regression and a trend towards much worse progression-free survival at 5 years (Sclafani et al., 2020). Earlier studies indicated that WDR12 polymorphisms were linked to genetic, developmental disorders, and malignancies (Yin et al., 2018).

Furthermore, Li et al. (Li et al., 2020b) verified the association between somatic WDR12 mutations and several inherited severe diseases. The cBioPortal tool was used to evaluate the altered patterns of WDR12 in a variety of human tumors, and the findings revealed that WDR12 was mutated in most tumors. WDR12 missense mutations were the most prevalent DNA mutations. Patients with LUSC with WDR12 genetic mutations had a worse prognosis for survival, which was later discovered. According to the data, patients with modified WDR12 were less likely to survive than those with unaltered WDR12. This suggests that WDR12 may be a useful predictor of patient survival and treatment responsiveness.

In numerous human malignancies, the state of gene methylation has been widely examined. According to earlier research findings, DNA hypermethylation is an important mechanism for tumor suppressor gene inhibition (Son et al., 2011; Xiao et al., 2014). However, on the other hand, hypomethylation served as a mechanism for oncogene stimulation to promote tumorigenesis (Medina-Aguilar et al., 2016). Next, we conducted a methylation analysis for WDR12. As expected, multiple cancers, namely BLCA, BRCA, KIRP, LIHC, PRAD, READ, THCA, and UCEC, exhibited hypomethylation in cancer tissues compared to normal samples. Furthermore, CpG aggregated methylation data showed that all significant findings favored hypomethylation of WDR12 in cancer tissues compared to normal samples (except for CHOL).

Immunotherapy has altered the landscape of cancer therapy by introducing novel and effective therapeutic approaches for numerous solid and hematologic cancers. While immunotherapy is not a new oncology branch, it has rapidly become one of the most promising in medicine (Khan et al., 2021). Improvements in checkpoint inhibitors, monoclonal antibodies, CAR-T cells,

and anticancer vaccinations are among other significant breakthroughs. Subsequent research in immunotherapy, including the gut microbiota, in combination with checkpoint inhibitors and gene sequencing, continues to personalize therapies for cancer patients, offering limitless potential opportunities (Madden, 2018). For this reason, it was essential to examine the correlation between enhanced WDR12 expression in cancerous tissues and the infiltration of various immune cells into the tumors. It was discovered that MDSC positively influences tumor cell survival and metastasis (Ye et al., 2010; Condamine et al., 2015). In addition, it suppresses other tumor-fighting cells (CD8 T cells and NK cells), promotes angiogenesis, and participates in the development of cancer stem cells 40. Therefore, it was unsurprising that a high amount of MDSC infiltration was associated with a poor prognosis among cancer patients (Yin et al., 2020). In all tumor types investigated except for UVM, THCA, PCPG, KIRC, DLBC, BRCA-LumB, and BRCA-Her2, the current investigation revealed statistically significant and positively correlated associations between WDR12 expression and MDSCs infiltration. CAF is the second cell examined for its connection with WDR12 overexpression. This type of cell has a significant role in the fight against cancers in the early stages, participating in tumor immune surveillance and secreting several effector molecules (Shah et al., 2022).

Consequently, CAFs are a meaningful and informative supplement to current pathologic investigation and help to determine prognosis and therapeutic interventions. Identifying the processes by which CAFs and cancer cells interact reciprocally is essential for improving TME-focused cancer treatments (Knops et al., 2020). Various malignancies, particularly ACC, BRCA-LumA, CESC, CHOL, KISH, KIRP, LIHC, SKCM, and THYM, exhibited a significant positive correlation between the expression of WDR12 and the invasion of CAFs. In addition, we found that WDR12 was inversely correlated with immunostimulators, immunoinhibitors, MHCs, TILs, chemokines, and receptors in most malignancies, suggesting that the TME is complex in nature. In conclusion, these findings demonstrate that aberrant WDR12 expression plays a crucial role in TME. Combining WDR12 expression data, MDSC infiltration, and CAF infiltration, we may conclude that upregulation of WDR12 may serve as a biomarker for an inadequate immune response against a developing tumor.

As potential prognostic biomarkers for the effectiveness of immunotherapies in anticancer therapeutic interventions, TMB and MSI can offer additional insight into tumor behaviors. For this reason, we examined the correlation between WDR12 expression and TMB and MSI in all TCGA tumors. In LUAD, UCEC, PAAD, CESC, STAD, SKCM, and HNSC, a positive correlation was detected between WDR12 expression and TMB was detected. WDR12 expression was positively correlated with MSI in PRAD, LIHC, PAAD, CESC, BRCA, STAD, KIRC, HNSC, and KICH; however, it was negatively correlated with MSI in DLBC. The findings demonstrate that the WDR12 expression has a significant impact on the TMB and MSI, as well as on the immune checkpoint suppressive therapeutic response of patients. Our conclusions in this research suggest a research concern regarding the likelihood of using WDR12 expression in the tumor mentioned above as a possible biomarker for patients' immunotherapy responsiveness.

Finally, we investigated the functional characteristics of differentially regulated WDR12 by including WDR12-binding proteins in addition to WDR12 expression-related genes in all TCGA cancers, followed by KEGG pathway enrichment and GO enrichment analyses. We retrieved the top 50 interacting proteins and the top 100 linked proteins to WDR12 in tumor tissue from the STRING and GEPIA2 databases, respectively, and found that DKC1 is a promising candidate. Among the top 100 genes with roughly similar expression patterns to WDR12 in TCGA tumors, the expression of WDR12 was significantly and positively correlated with the expression of NOP58, SGOL2, SUMO1, SUV39H2, and WDR75. WDR12 does not appear to interact physically with any of these five genes, even though they all play a role in the "cell cycle" and DNA replication. NOP58 is a possible prognostic biomarker for HCC because its upregulation is inversely correlated with OS in HCC patients (Wang et al., 2021). The correlation between high expression of SGOL2 and aggressive development and a poor prognosis in adrenocortical carcinoma is notable (Tian et al., 2020). NSCLC proliferation, colony formation, invasion, and NFB expression were increased by overexpression of SUMO1 (Ke et al., 2019). SUV39H2 maintains standard biological function, the loss of which promotes carcinogenesis, and its upregulation as an oncogene often leads to cancer initiation and progression (Li et al., 2019) WDR75 knockdown triggered the RPL5/RPL11-dependent stabilization checkpoint for p53, resulting in decreased cell proliferation and senescence (Moudry et al., 2022).

5 Conclusion

WDR12 is an oncogene that was shown to be overexpressed as mRNA and protein in many human malignancies. Its overexpression was found to be correlated with a poor prognosis. It influences the infiltration of multiple immune cells, with immunosuppressive cells MDSCs and CAFs infiltration associated favorably with WDR12 expression. Furthermore, it was discovered that TMB, MSI, and immunoregulators are associated with WDR12 expression in a subset of human cancers; therefore, these results potentially propose WDR12 as a prognostic biomarker of patients' responsiveness to immunotherapy and a possible cancer treatment target.

Eventually, the present research work offers a trustworthy reference for the entire characteristics and roles of WDR12 in carcinogenesis, which can assist patients in selecting more precise immunotherapy regimens.

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

Conceptualization: MAE, MAS, and RAE; original draft preparation. MAE, SA, IJ, GA, EF, MN, and LA: writing—review, and editing: IJ, SA, SB: methodology: MAE, MYB, MAD, MAZ: supervision and project administration, BK, MAZ, RAE, and MA-S. All authors contributed to the article and approved the submitted version.

Funding

The Deanship of Scientific Research of King Khalid University in Abha, Saudi Arabia, entirely supported this research (Grant No. G.R.P.1-27-43). This research was also supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2020R1I1A2066868), the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2020R1A5A2019413), a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HF20C0116), and a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HF20C0038).

Acknowledgments

The authors extend their gratitude to the King Khalid University Deanship of Scientific Research for funding this project through the General Research Project under the grant number (G.R.P.1-27-43).

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/fgene.2022.1008502/full#supplementary-material

References

Aizawa, T., Karasawa, H., Funayama, R., Shirota, M., Suzuki, T., Maeda, S., et al. (2019). Cancer-associated fibroblasts secrete Wnt2 to promote cancer progression in colorectal cancer. *Cancer Med.* 8 (14), 6370–6382. doi:10.1002/cam4.2523

Anandakrishnani, R., Varghese, R. T., Kinneyi, N. A., and Garner, H. R. (2019). Estimating the number of genetic mutations (HITS) required for carcinogenesis based on the distribution of somatic mutations. *PLoS Comput. Biol.* 15 (3), e1006881. doi:10.1371/journal.pcbi.1006881

Anderson, N. M., and Simon, M. C. (2020). The tumor microenvironment. *Curr. Biol.* 30 (16), R921–R925. doi:10.1016/j.cub.2020.06.081

Anglim, P. P., Alonzo, T. A., and Laird-Offringa, I. A. (2008). DNA methylation-based biomarkers for early detection of non-small cell lung cancer: An update. *Mol. Cancer* 7, 81. doi:10.1186/1476-4598-7-81

Blum, A., Wang, P., and Zenklusen, J. C. (2018). SnapShot: TCGA-analyzed tumors. *Cell.* 173, 530. doi:10.1016/j.cell.2018.03.059

Chandrashekar, D. S., Bashel, B., Balasubramanya, S. A. H., Creighton, C. J., Ponce-Rodriguez, I., Chakravarthi, B. V. S. K., et al. (2017). Ualcan: A portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia (United States)* 19 (8), 649–658. doi:10.1016/j.neo.2017.05.002

Condamine, T., Mastio, J., and Gabrilovich, D. I. (2015). Transcriptional regulation of myeloid-derived suppressor cells. *J. Leukoc. Biol.* 98 (6), 913–922. doi:10.1189/jlb.4RI0515-204R

Cui, K., Liu, C., Li, X., Zhang, Q., and Li, Y. (2020). Comprehensive characterization of the rRNA metabolism-related genes in human cancer. *Oncogene* 39 (4), 786–800. doi:10.1038/s41388-019-1026-9

de Cicco, P., Ercolano, G., and Ianaro, A. (2020). The new era of cancer immunotherapy: Targeting myeloid-derived suppressor cells to overcome immune evasion. *Front. Immunol.* 11, 1680. doi:10.3389/fimmu.2020.01680

Fernández, A. (2014). Synergizing immunotherapy with molecular-targeted anticancer treatment. *Drug Discov. Today* 19 (9), 1427–1432. doi:10.1016/j.drudis.2014.03.022

Finck, A., Gill, S. I., and June, C. H. (2020). Cancer immunotherapy comes of age and looks for maturity. *Nat. Commun.* 11, 3325. doi:10.1038/s41467-020-17140-5

Fumet, J. D., Truntzer, C., Yarchoan, M., and Ghiringhelli, F. (2020). Tumour mutational burden as a biomarker for immunotherapy: Current data and emerging concepts. *Eur. J. Cancer* 131, 40–50. doi:10.1016/j.ejca.2020.02.038

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal* 6 (269), pl1. doi:10.1126/scisignal.2004088

Hölzel, M., Rohrmoser, M., Schlee, M., Grimm, T., Harasim, T., Malamoussi, A., et al. (2005). Mammalian WDR12 is a novel member of the Pes1-Bop1 complex and is required for ribosome biogenesis and cell proliferation. *J. Cell. Biol.* 170 (3), 367–378. doi:10.1083/jcb.200501141

Hutter, C., and Zenklusen, J. C. (2018). The cancer genome Atlas: Creating lasting value beyond its data. *Cell.* 173, 283–285. doi:10.1016/j.cell.2018.03.042

Ke, C., Zhu, K., Sun, Y., Ni, Y., Zhang, Z., and Li, X. (2019). SUMO1 promotes the proliferation and invasion of non-small cell lung cancer cells by regulating NF-κB. *Thorac. Cancer* 10 (1), 33–40. doi:10.1111/1759-7714.12895

Khan, M., Maker, A. v., and Jain, S. (2021). Role of NKT cells during viral infection and the development of NKT cell-based nanovaccines. *Vaccines* 9, 949. doi:10.3390/vaccines9090949

Knops, A. M., South, A., Rodeck, U., Martinez-Outschoorn, U., Harshyne, L. A., Johnson, J., et al. (2020). Cancer-associated fibroblast density, prognostic characteristics, and recurrence in head and neck squamous cell carcinoma: A meta-analysis. *Front. Oncol.* 10, 565306. doi:10.3389/fonc.2020.565306

Lerch-Gaggl, A., Haque, J., Li, J., Ning, G., Traktman, P., and Duncan, S. A. (2002). Pescadillo is essential for nucleolar assembly, ribosome biogenesis, and mammalian cell proliferation. *J. Biol. Chem.* 277 (47), 45347–45355. doi:10.1074/jbc.M208338200

Lewinska, A., Bednarz, D., Adamczyk-Grochala, J., and Wnuk, M. (2017). Phytochemical-induced nucleolar stress results in the inhibition of breast cancer cell proliferation. *Redox Biol.* 12, 469–482. Redox Biol [Internet]. doi:10.1016/j.redox.2017.03.014

Li, B., Zheng, Y., and Yang, L. (2019). The oncogenic potential of SuV39H2: A comprehensive and perspective view. *J. Cancer* 10, 721–729. doi:10.7150/jca.28254

Li, J. L., Chen, C., Chen, W., Zhao, L. F., Xu, X. K., Li, Y., et al. (2020). Integrative genomic analyses identify WDR12 as a novel oncogene involved in glioblastoma. *J. Cell. Physiol.* 235 (10), 7344–7355. doi:10.1002/jcp.29635

Li, Y., Roberts, N. D., Wala, J. A., Shapira, O., Schumacher, S. E., Kumar, K., et al. (2020). Patterns of somatic structural variation in human cancer genomes. *Nature* 578 (7793), 112–121. doi:10.1038/s41586-019-1913-9

Madden, D. L. (2018). From a patient advocate's perspective: Does cancer immunotherapy represent a paradigm shift? *Curr. Oncol. Rep.* 201, 8. doi:10.1007/s11912-018-0662-5

Medina-Aguilar, R., Pérez-Plasencia, C., Marchat, L. A., Gariglio, P., García Mena, J., Rodríguez Cuevas, S., et al. (2016). Methylation landscape of human breast cancer cells in response to dietary compound resveratrol. *PLoS One* 11 (6), e0157866. doi:10.1371/journal.pone.0157866

Mi, L., Qi, Q., Ran, H., Chen, L., Li, D., Xiao, D., et al. (2021). Suppression of ribosome biogenesis by targeting WD repeat domain 12 (WDR12) inhibits glioma stem-like cell growth. *Front. Oncol.* 11, 751792. doi:10.3389/fonc.2021.751792

Moilanen, A. M., Rysä, J., Kaikkonen, L., Karvonen, T., Mustonen, E., Serpi, R., et al. (2015). WDR12, a member of nucleolar PeBoW-complex, is up-regulated in failing hearts and causes deterioration of cardiac function. *PLoS One* 10 (4), e0124907. doi:10.1371/journal.pone.0124907

Moudry, P., Chroma, K., Bursac, S., Volarevic, S., and Bartek, J. (2022). RNA-interference screen for p53 regulators unveils a role of WDR75 in ribosome biogenesis. *Cell. Death Differ. [Internet]* 29 (3), 687–696. doi:10.1038/s41418-021-00882-0

Nagy, Á., Munkácsy, G., and Győrffy, B. (2021). Pancancer survival analysis of cancer hallmark genes. *Sci. Rep.* 11 (1), 6047. doi:10.1038/s41598-021-84787-5

Ohn, J., Ailar, C. B., and Ornik, L. G. (1997). Cancer undefeated. *N. Engl. J. Med.* 336, 1569–1574. doi:10.1056/NEJM199705293362206

Pino, M. S., and Chung, D. C. (2011). Microsatellite instability in the management of colorectal cancer. *Expert Rev. Gastroenterology Hepatology* 5, 385–399. doi:10.1586/egh.11.25

Qu, Y., Dang, S., and Hou, P. (2013). Gene methylation in gastric cancer. *Clin. Chim. Acta* 424, 53–65. doi:10.1016/j.cca.2013.05.002

Rohrmoser, M., Hölzel, M., Grimm, T., Malamoussi, A., Harasim, T., Orban, M., et al. (2007). Interdependence of Pes1, Bop1, and WDR12 controls nucleolar localization and assembly of the PeBoW complex required for maturation of the 60S ribosomal subunit. *Mol. Cell. Biol.* 27 (10), 3682–3694. doi:10.1128/MCB.00172-07

Romero-Garcia, S., Prado-Garcia, H., and Carlos-Reyes, A. (2020). Role of DNA methylation in the resistance to therapy in solid tumors. *Front. Oncol.* 10, 1152. doi:10.3389/fonc.2020.01152

Schumacher, T. N., and Schreiber, R. D. (2015). Neoantigens in cancer immunotherapy. *Science* 348, 69–74. doi:10.1126/science.aaa4971

Sclafani, F., Wilson, S. H., Cunningham, D., Gonzalez De Castro, D., Kalaitzaki, E., Begum, R., et al. (2020). Analysis of KRAS, NRAS, BRAF, PIK3CA and TP53 mutations in a large prospective series of locally advanced rectal cancer patients. *Int. J. Cancer* 146 (1), 94–102. doi:10.1002/ijc.32507

Shah, K., Mallik, S. B., Gupta, P., and Iyer, A. (2022). Targeting tumour-associated fibroblasts in cancers. *Front. Oncol.* 12, 12. doi:10.3389/fonc.2022.908156

Son, J. W., Jeong, K. J., Jean, W. S., Park, S. Y., Jheon, S., Cho, H. M., et al. (2011). Genome-wide combination profiling of DNA copy number and methylation for deciphering biomarkers in non-small cell lung cancer patients. *Cancer Lett.* 311 (1), 29–37. doi:10.1016/j.canlet.2011.06.021

Sun, L. C., and Qian, H. X. (2018). Screening for implicated genes in colorectal cancer using whole-genome gene expression profiling. *Mol. Med. Rep.* 17 (6), 8260–8268. doi:10.3892/mmr.2018.8862

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). Gepia: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 45 (W1), W98–W102. doi:10.1093/nar/gkx247

Tian, X., Xu, W., Wang, Y., Anwaier, A., Wang, H., Wan, F., et al. (2020). Identification of tumor-infiltrating immune cells and prognostic validation of tumor-infiltrating mast cells in adrenocortical carcinoma: Results from bioinformatics and real-world data. *Oncoimmunology* 9 (1), 1784529. doi:10.1080/2162402X.2020.1784529

van Velzen, M. J. M., Derks, S., van Grieken, N. C. T., Haj Mohammad, N., and van Laarhoven, H. W. M. (2020). MSI as a predictive factor for treatment outcome of gastroesophageal adenocarcinoma. *Cancer Treat. Rev.* 86, 102024. doi:10.1016/j.ctrv.2020.102024

Wang, J., Huang, R., Huang, Y., Chen, Y., and Chen, F. (2021). Overexpression of NOP58 as a prognostic marker in hepatocellular carcinoma: A TCGA data-based analysis. *Adv. Ther.* 38 (6), 3342–3361. doi:10.1007/s12325-021-01762-2

Wu, T., and Dai, Y. (2017). Tumor microenvironment and therapeutic response. *Cancer Lett.* 387, 61–68. doi:10.1016/j.canlet.2016.01.043

Xiao, P., rong, Chen J., Zhou, F., Lu, C., Yang, Q., hua, Tao G., et al. (2014). Methylation of P16 in exhaled breath condensate for diagnosis of non-small cell lung cancer. *Lung Cancer* 83 (1), 56–60. doi:10.1016/j.lungcan.2013.09.008

Yarchoan, M., Hopkins, A., and Jaffee, E. M. (2017). Tumor mutational burden and response rate to PD-1 inhibition. *N. Engl. J. Med.* 377 (25), 2500–2501. doi:10.1056/NEJMc1713444

Ye, X. Z., Yu, S. C., and Bian, X. W. (2010). Contribution of myeloid-derived suppressor cells to tumor-induced immune suppression, angiogenesis, invasion and metastasis. *J. Genet. Genomics* 37, 423–430. doi:10.1016/S1673-8527(09)60061-8

Yin, K., Xia, X., Rui, K., Wang, T., and Wang, S. (2020). Myeloid-derived suppressor cells: A new and pivotal player in colorectal cancer progression. *Front. Oncol.* 10, 610104. doi:10.3389/fonc.2020.610104

Yin, Y., Zhou, L., Zhan, R., Zhang, Q., and Li, M. (2018). Identification of WDR12 as a novel oncogene involved in hepatocellular carcinoma propagation. *Cancer Manag. Res.* 10, 3985–3993. doi:10.2147/CMAR.S176268

Zhang, X., Li, X., You, Q., and Zhang, X. (2017). Prodrug strategy for cancer cell-specific targeting: A recent overview. *Eur. J. Med. Chem. [Internet* 139, 542–563. doi:10.1016/j.ejmech.2017.08.010

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 27 January 2023
DOI 10.3389/fgene.2023.1023613

Check for updates

OPEN ACCESS

EDITED BY
Jinhui Liu,
Nanjing Medical University, China

REVIEWED BY
Shabana Bibi,
Shifa Tameer-e-Millat University, Pakistan
Xuan high,
State Key Laboratory of Microbial Resources, Institute of Microbiology (CAS), China

*CORRESPONDENCE
Yang Sun,
sunyang@fjzlhospital.com

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 20 August 2022
ACCEPTED 17 January 2023
PUBLISHED 27 January 2023

CITATION
Liu L, Zheng J, Xia H, Wu Q, Cai X, Ji L and Sun Y (2023), Construction and comprehensive analysis of a curoptosis-related lncRNA signature for predicting prognosis and immune response in cervical cancer.
Front. Genet. 14:1023613.
doi: 10.3389/fgene.2023.1023613

COPYRIGHT
© 2023 Liu, Zheng, Xia, Wu, Cai, Ji and Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Construction and comprehensive analysis of a curoptosis-related lncRNA signature for predicting prognosis and immune response in cervical cancer

Li Liu[1], Jianfeng Zheng[1], Hongmei Xia[2], Qiaoling Wu[1], Xintong Cai[1], Liyan Ji[3] and Yang Sun[1]*

[1]Department of Gynecology, Clinical Oncology School of Fujian Medical University, Fujian Cancer Hospital, Fuzhou, China, [2]Department of Gynecology, Fujian Cancer Hospital, Fuzhou, China, [3]Geneplus-Beijing Institute, Beijing, China

Cuproptosis (copper-ion-dependent cell death) is an unprogrammed cell death, and intracellular copper accumulation, causing copper homeostasis imbalance and then leading to increased intracellular toxicity, which can affect the rate of cancer cell growth and proliferation. This study aimed to create a newly cuproptosis-related lncRNA signature that can be used to predict survival and immunotherapy in patients with cervical cancer, but also to predict prognosis in patients treated with radiotherapy and may play a role in predicting radiosensitivity. First of all, we found lncRNAs associated with cuproptosis between cervical cancer tumor tissues and normal tissues. By LASSO-Cox analysis, overlapping lncRNAs were then used to construct lncRNA signatures associated with cuproptosis, which can be used to predict the prognosis of patients, especially the prognosis of radiotherapy patients, ROC curves and PCA analysis based on cuprotosis-related lncRNA signature and clinical signatures were developed and demonstrated to have good predictive potential. In addition, differences in immune cell subset infiltration and differences in immune checkpoint expression between high-risk and low-risk score groups were analyzed, and we investigated the relationship between this signature and tumor mutation burden. In summary, we constructed a lncRNA prediction signature associated with cuproptosis. This has important clinical implications, including improving the predictive value of cervical cancer patients and providing a biomarker for cervical cancer.

KEYWORDS

cervical cancer, cuproptosis, lncRNA, prognosis, radiotherapy, immune

Abbreviations: CESC, Cervical squamous cell carcinoma and endocervical adenocarcinoma; CC, Cervical cancer; WHO, World HealthOrganization; TCGA, The Cancer Genome Atlas; LASSO, Least absolute shrinkage and selection operator; OS, Overall survival; PFS, Progression-free survival; GSEA, Gene Set Enrichment Analysis; HR,Hazard ratio; HRG, High-risk group; LRG, Low-risk group; CR, Complete response; PR, partial response; PD, progressive disease; SD, stable disease; K-M, Kaplan–Meier; GO, Gene ontology; RS, Radiosensitivity; RR, Radiotherapy resisdance.

Background

Cervical cancer (CC) poses a genuine threat to women's health. CC is the most common malignant tumor in women worldwide. In recent years, owing to advances in surgery, radiotherapy/radiochemotherapy, and the widespread use of early CC screening, the incidence of CC has decreased significantly in developed countries. In 2020, the World Health Organization implemented a global strategy aimed at reducing the incidence of CC to four cases per 100,000 women by the 21st century (Wilailak et al., 2021). The majority of new cases (~85%) and deaths (~90%) worldwide occur in low- and middle-income countries (Bhatla et al., 2021). Although secondary prevention through new vaccinations and primary prevention through cancer screening is common, in 2014, only 1% of girls from low- and middle-income countries were vaccinated globally (Sundstrom and Elfstrom, 2020). Human papillomavirus (HPV) infection contributes to the development of CC (Munoz et al., 2003), however, its impact on CC prognosis is currently inconclusive, and radiotherapy is the main treatment for patients with advanced CC. Moreover, the response to radiotherapy is indicative of the prognosis of patients with CC. Therefore, an ideal predictive model or accurate prognostic biomarker to guide CC treatment is warranted.

Heavy metal ions are essential micronutrients, however, their excessive or insufficient consumption can kill cells. Different subprogrammes can cause regulated cell death in response to stress caused by heavy metals. For example, ferroproteinases can cause iron-dependent types of oxidative cell death due to unrestricted lipid peroxidation (Tang et al., 2022). 'Cuproptosis', a non-apoptotic, novel type of cell death, was recently discovered by Tzvitkov et al. (Li et al., 2022). They noted that copper causes cell death primarily by binding to the lipoacylated components of the tricarboxylic acid cycle, which in turn leads to toxic protein stress and cell death (Tsvetkov et al., 2022).

Long non-coding RNAs (lncRNAs) are non-protein-coding RNA molecules whose transcripts are >200 nucleotides in length (Gibb et al., 2011; Zhang et al., 2020), these were originally thought to be by-products of RNA polymerase II transcription (Mercer et al., 2009). By interfering with DNA, miRNAs, and proteins, lncRNAs play an important role in the occurrence and development of tumors (Wang and Chang, 2011; Yoon et al., 2012; Li et al., 2019), inducing differences in the expression of key genes (Ma et al., 2019; Shahabi et al., 2019). Increasing research has suggested that lncRNAs function in many types of malignant tumors, affecting cell differentiation, growth, invasion, migration, and apoptosis as well as the cell cycle and tumor's resistance to chemotherapy (Qian et al., 2021; Zhang et al., 2021; Chen et al., 2022). To date, cuproptosis-related lncRNAs have been studied in renal tumors (Xu et al., 2022a), liver tumors (Zhang et al., 2022), soft tissue sarcomas (Han et al., 2022), *etc.* However, the mechanisms through which copper regulates tumor cell death remain unclear, and there is no conclusive evidence that cuproptosis-related lncRNAs are associated with CC. We hope that the characteristics of the cuproptosis-related lncRNAs engineered herein will provide help predict CC prognosis and the mechanism of cuproptosis in CC.

Materials and methods

Public data collection and processing

The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/) was used to retrieve the RNA sequencing, mutation, and clinical data of 307 samples (304 CC samples and three normal samples) from 73 patients who received radiotherapy and provided post-radiotherapy evaluation information. The Perl programming language (https://www.perl.org, version Strawberry-Perl-5.30.0) was used to process the data. The transcriptome and clinical data were pre-processed to extract the expression profile matrix of the coding genes and lncRNAs and the clinical and pathological characteristics of the patients with CC, including their survival status, survival times, CC stage and grade, and TMN. We performed a literature search (Ge et al., 2022; Li et al., 2022; Tang et al., 2022; Tsvetkov et al., 2022) and found fourteen genes associated with cuproptosis (*FDX1*, *ATP7B*, *GCSH*, *DBT*, *GLS*, *CDKN2A*, *MTF1*, *SLC31A1*, *PDHB*, *LIAS*, *DLD*, *LIPT1*, *DLAT*, and *PDHA1*).

Screening of cuproptosis-related lncRNAs

We used the Pearson correlation analysis to assess the correlation between cuproptosis-related lncRNAs and mRNAs (|corFilter = 0.4|; pvalueFilter = 0.001) in R software using the 'limma,' 'ggplot2', 'ggalluvial', and 'dplyr' packages. Finally, Sankey plots were drawn to present the correlations between cuproptosis-related lncRNAs and all of the 14 genes.

Construction of prognostic cuproptosis-related lncRNA signatures

The R package 'caret' was used to randomly divide the cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) datasets retrieved from TCGA into the training and testing sets. The signature was validated using the testing set and the entire TCGA dataset.

The lncRNAs related with cuproptosis were screened using univariate Cox regression analysis ($p < 0.05$) and the results were presented using forest plots. In addition, the Lasso-Cox regression analysis was used to determine the optimal prognostic lncRNA group, and a risk model was established (penalty parameters were estimated through 1000-fold cross-validation); this approach minimized overfitting in the modeling process. Finally, multivariate Cox regression analysis was used to establish a prognostic model based using the best lncRNAs; three lncRNAs were selected for model construction, after which a heat map of the correlation between the three lncRNAs and cuproptosis-related genes was plotted using the 'ggplot2' package. The following equations were applied to each patient with CC to assess the risk score:

$$\text{Risk score} = \Sigma\beta\text{lncRNA} \times \text{Exp lncRNA}$$

Based on the risk score median value, the patients were divided into the high-risk group (HRG) and the low-risk group (LRG). Kaplan–Meier curves were generated to determine whether the overall survival (OS) and progression-free survival (PFS) differed between the HRG and LRG. We also verified the independence of the model with an independent analysis approach using factors such as patient age and CC grade, and stage. Based on the patient's survival, and the survminer and timeROC programs, receiver operating characteristic (ROC) curves were generated, and the area under the curve (AUC) was calculated. To measure the model's accuracy, the concordance index (C-index) was used. A heat map was created to illustrate the correlation between the model and clinical characteristics. In the 73 patients with CC who underwent radiotherapy, the survival rate differences between different risk groups were analyzed, and the lncRNAs involved in constructing the model were further investigated to understand the differences in their expressions between the radiosensitive and insensitive groups.

TABLE 1 Clinical characteristics of CESC patients involved in the study.

		Total N = 304	Training set n = 152	Testing set n = 152	*p*-value
Age	≤65	269 (88.49%)	133 (87.5%)	136 (89.47%)	0.7193
	>65	35 (11.51%)	19 (12.5%)	16 (10.53%)	
Grade	G1	18 (5.92%)	8 (5.26%)	10 (6.58%)	0.5986
	G2	135 (44.41%)	71 (46.71%)	64 (42.11%)	
	G3	118 (38.82%)	56 (36.84%)	62 (40.79%)	
	G4	1 (0.33%)	0 (0%)	1 (0.66%)	
	unknow	32 (10.53%)	17 (11.18%)	15 (9.87%)	
Stage	Stage I	162 (53.29%)	75 (49.34%)	87 (57.24%)	0.1684
	Stage II	69 (22.7%)	36 (23.68%)	33 (21.71%)	
	Stage III	45 (14.8%)	24 (15.79%)	21 (13.82%)	
	Stage IV	21 (6.91%)	15 (9.87%)	6 (3.95%)	
	unknow	7 (2.3%)	2 (1.32%)	5 (3.29%)	
T	T1	140 (46.05%)	64 (42.11%)	76 (50%)	0.1009
	T2	71 (23.36%)	41 (26.97%)	30 (19.74%)	
	T3	20 (6.58%)	10 (6.58%)	10 (6.58%)	
	T4	10 (3.29%)	8 (5.26%)	2 (1.32%)	
	unknow	63 (20.72%)	29 (19.08%)	34 (22.37%)	
M	M0	116 (38.16%)	59 (38.82%)	57 (37.5%)	0.7417
	M1	10 (3.29%)	4 (2.63%)	6 (3.95%)	
	unknow	178 (58.55%)	89 (58.55%)	89 (58.55%)	
N	N0	133 (43.75%)	64 (42.11%)	69 (45.39%)	0.6066
	N1	60 (19.74%)	32 (21.05%)	28 (18.42%)	
	unknow	111 (36.51%)	56 (36.84%)	55 (36.18%)	

Nomogram and principal component analysis

Using the 'rms' and 'survival' packages, we created line graphs for patients with CC having 1-, 3-, and 5-year OS, with the risk scores combined with their clinicopathologic factors. Using Hosmer–Lemeshow calibration curves, we tested whether the developed nomogram had any predictive power. Principal component analysis (PCA) was conducted using the 'scatterplot3d' package in the R program. PCA was also used to categorize the expression patterns of the four cohorts–all the genes, cuproptosis-related genes, cuproptosis-related LncRNAs, and the three lncRNA signatures (risk DE-lncRNAs)–to display the spatial distribution of the high-risk and low-risk samples.

Gene ontology and kyoto encyclopaedia of genes and genomics analysis

Through differential gene screening, we identified 173 genes associated with HRGs and LRGs in 304 patients. We then performed Gene Ontology (GO) analysis for assessing biological processes, cellular processes, and molecular functions. We also used GO analysis to analyze differentially expressed Kyoto Encyclopaedia of Genes and Genomics (KEGG) pathways in both groups with the 'clusterProfiler', 'org.Hs.eg.db', 'enrichplot', and 'circlize' packages. The enriched biosynthetic pathways and processes with a *p*-value of <0.05 and a FDR of <0.05 were considered statistically significant.

Immune infiltration analysis

We investigated immune infiltration and immune-related functions in CC using single-sample gene set enrichment analysis with 'limma', 'GSVA', and 'GSEABase' packages. To investigate the relationship between this feature and the immune infiltration status, we used six algorithms (TIMER, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, MCPCOUNTER, and XCELL) to calculate the immune infiltration profile of the TCGA–CESC dataset and present the results as a heat map. Based on the estimation algorithm, we then calculated the percentage of tumor-infiltrating immune cells in the CC microenvironment associated with the three lncRNA signatures. Furthermore, we investigated the differences in the expressions of immune checkpoints in the HRG and LRG and presented the results as boxplots.

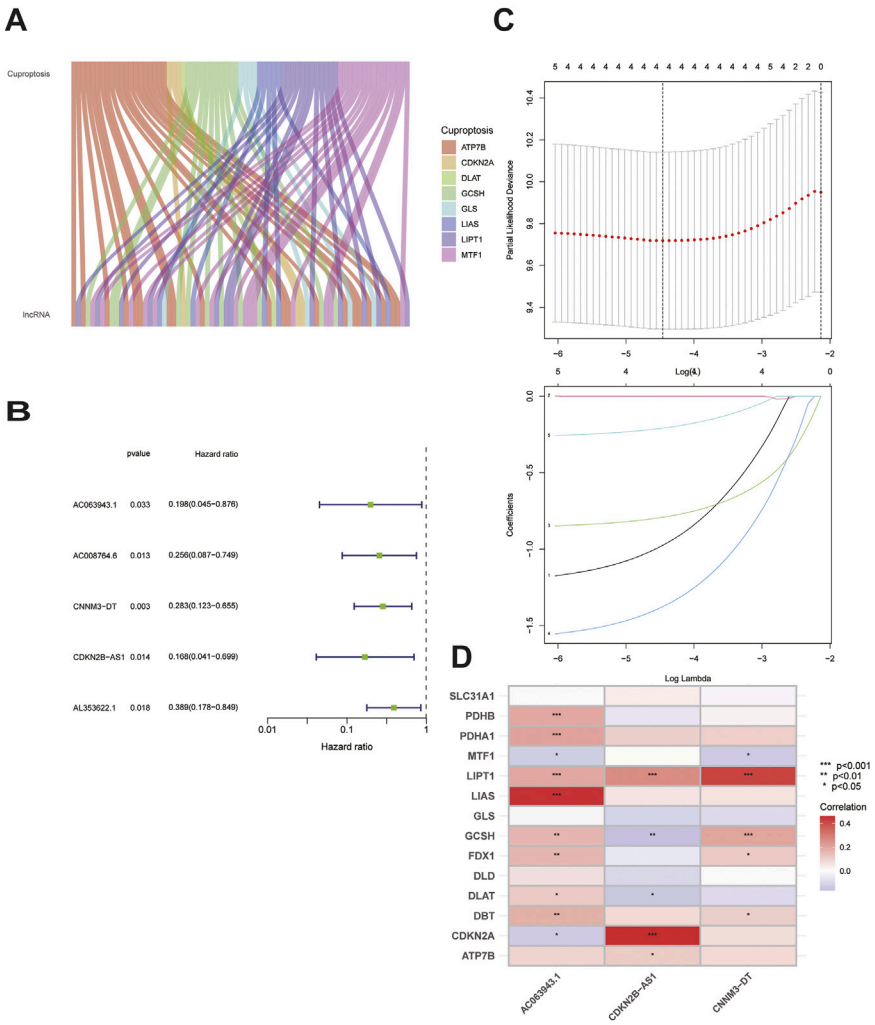


FIGURE 1
Identification of the signature of cuproptosis-related lncRNA.**(A)**Sankey diagram of 71 lncrnas co-expressed with cuproptosis-related genes.**(B)**Forest plots of 5-lncRNA screening by COX regression related with cuproptosis.**(C)**Lasso regression of the 5-lncRNA.**(D)**Heatmap shows the co-expression relationship between cuproptosis-related lncRNA signature and cuproptosis-related genes.

Differences in tumor mutation burden survival

After downloading the somatic mutation data, mutation information was extracted from the TCGA database using the Perl programming language. The 'maftools' package in R software was used to analyze differences in the mutations between the HRG and LRG. Using the 'survminer' package for survival analysis, we then examined and integrated the mutations with clinical data and analyzed differences in survival across mutation scenarios after combining the results of HRG and LRG.

Statistical analysis

The Perl programming language was used to extract and integrate the data, whereas the R program (version 4.2.1) was used for the Lasso-Cox regression, survival, and PCA analyses. Kaplan-Meier analysis was performed using the 'Survival' package. In addition, the 'survival', 'survivor', 'survival', 'pheatmap', and 'ggpubr' software packages were used to validate the prediction model. $p < 0.05$ was considered statistically significant.

Results

Construction of the cuproptosis-related lncRNA signature

Our study included 304 patients with CC who were randomly divided into either the training ($n = 152$) or testing ($n = 152$) set in a 1:1 ratio. The clinical and pathological characteristics of all patients are shown in Table 1. There were no statistically significant differences between the training and testing sets in the clinical characteristics based on a clinical statistical analysis of the groups. A total of 14 genes associated with cuproptosis were obtained upon review of the literature, and 71 cuproptosis-related lncRNAs associated with the 14 genes were screened using Pearson's correlation analysis (|corFilter = 0.4|; pvalueFilter = 0.001) (Figure 1A). Our training set

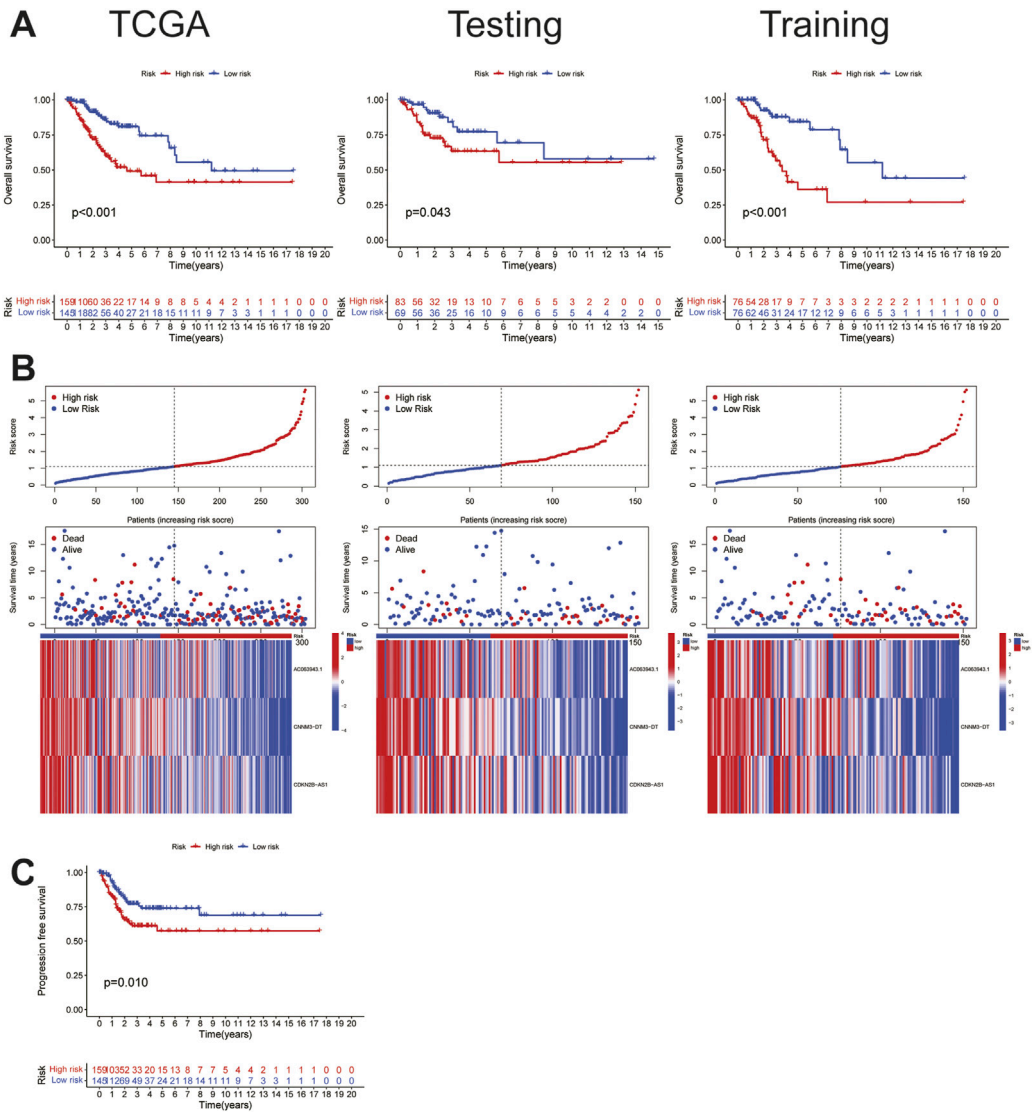


FIGURE 2
Characteristics of cuproptosis-related lncRNA signature.**(A)**The OS survival analysis of TCGA(risk) set, the training set, and the testing set. **(B)** Risk characteristics between high and low risk groups in TCGA(risk) set, the training set, and the testing set.**(C)** The PFS survival analysis in TCGA(risk) set.

revealed five cuproptosis-related lncRNAs upon univariate Cox regression analysis, and the resultant risk values were presented as forest plots (Figure 1B). Reduced multicollinearity was achieved by performing the Lasso regression analysis, and subsequent multivariate Cox regression analysis allowed us to screen three lncRNAs associated with cuproptosis (AC063943.1, CDKN2B–AS1, and CNNM3–DT; risk score = −AC063943.1 × 1.37125 − CNNM3-DT0.94343-CDKN2B-AS1 × 1.69231), which were verified using the testing set (Figure 1C). As shown through the heat map of the cuproptosis-related genes, the three lncRNAs were positively correlated with LIAS, CDKN2A, and LIPT1, respectively (Figure 1D).

Characteristics of cuproptosis-related lncRNAs

To study the prognostic and risk verification abilities of the three lncRNAs, the risk score of each individual in the testing set was calculated using the same calculation as that used in the training set. Then, we divided all the patients into the HRG and LRG based on the same cut-off values used for the training set, which were verified in each of the three cohorts, including the TCGA, training, and testing sets. Established cuproptosis-related lncRNAs had a strong ability to predict prognosis, including OS (Figure 2A) and PFS (Figure 2B), whereas the patients with elevated risk scores showed higher mortality and worse survival; moreover, the lower the expression of these three lncRNAs, the higher the risk rate, suggesting that these three lncRNAs function as tumor suppressor genes and can help predict the risk among patients. This conclusion was consistently validated in all three sets (Figure 2C).

To further verify the accuracy of the cuproptosis-related lncRNA signatures, an independent prognostic analysis was performed among the three cuproptosis-related lncRNAs involved in the construction of the signature and other clinical signatures with complete information. The results of univariate proved that the 3-lncRNAs signature is an independent predictor (Figure 3A) as compared with the other clinical signatures, and the association of the risk prediction model with the

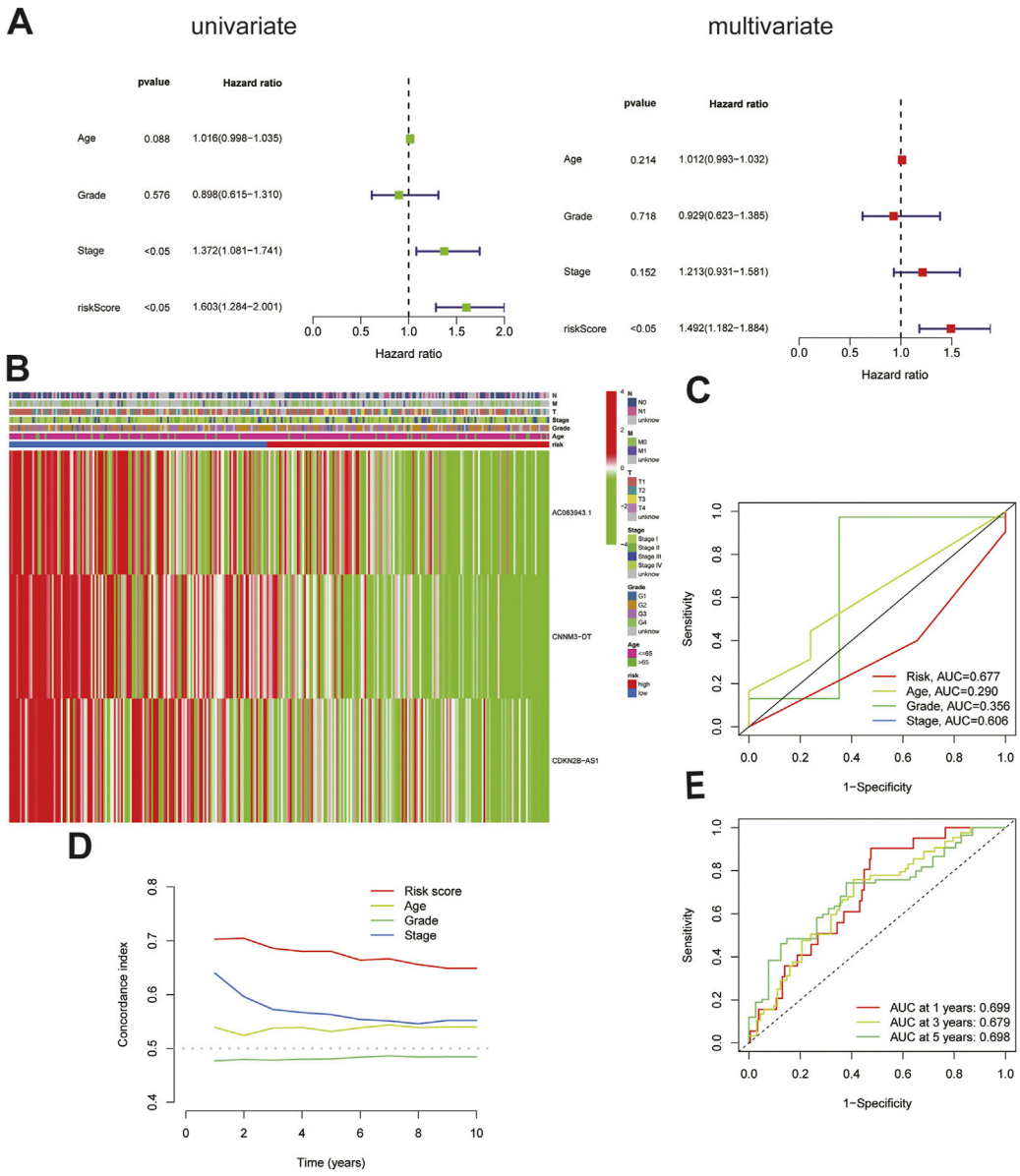


FIGURE 3
Validation of the cuproptosis-related lncRNA signature. **(A)**The cuproptosis-related lncRNA signature was shown to be an independent risk factor for patients' overall survival in TCGA.**(B)**The heatmap shows the expression of signature in the high and low risk group and their relationship with other clinicopathological signatures.**(C)**The AUC showed that riskscore was an independent predictor compared with other clinicopathological signatures.**(D)**The C-Index of the riskscore was higher than other clinicopathological signatures.**(E)**The signature could be used as an independent predictor to predict the OS of 1-,3-,5 years.

other clinical signatures in all CC samples is shown in the heat map (Figure 3B). Furthermore, the ROC curve indicated that the predictive power of this signature relative to other clinical signatures was high, with AUC = 0.677 (Figure 3C). According to the curve, the C-index of this signature was much higher than that of the other three clinical signatures (Figure 3D), and according to the time ROC analysis, the 1-, 3-, and 5-year AUCs of this signature were 0.699, 0.679, and 0.698, respectively (Figure 3E), suggesting that this signature can function as an independent predictor independent of the patients' age as well as tumor grade and stage and has stronger predictive power and higher confidence as compared with the other clinical signatures, similar to that noted for prognosis.

PCA and construction of a nomogram to predict patient survival

Differences between the HRG and LRG in terms of all the genes (Figure 4. A), 14 genes associated with cuproptosis (Figure 4B), 71 cuproptosis-related lncRNAs (Figure 4C), and three cuproptosis-related lncRNAs of the signature risk DE-lncRNAs (Figure 4D) were determined using PCA with the 'scatterplot3d' R package. The range of gene expression between the HRG and LRG by the 3-lncRNAs signature was relatively well-defined and more predictive than all the genes, 14 genes associated with cuproptosis and 71 cuproptosis-related lncRNAs.

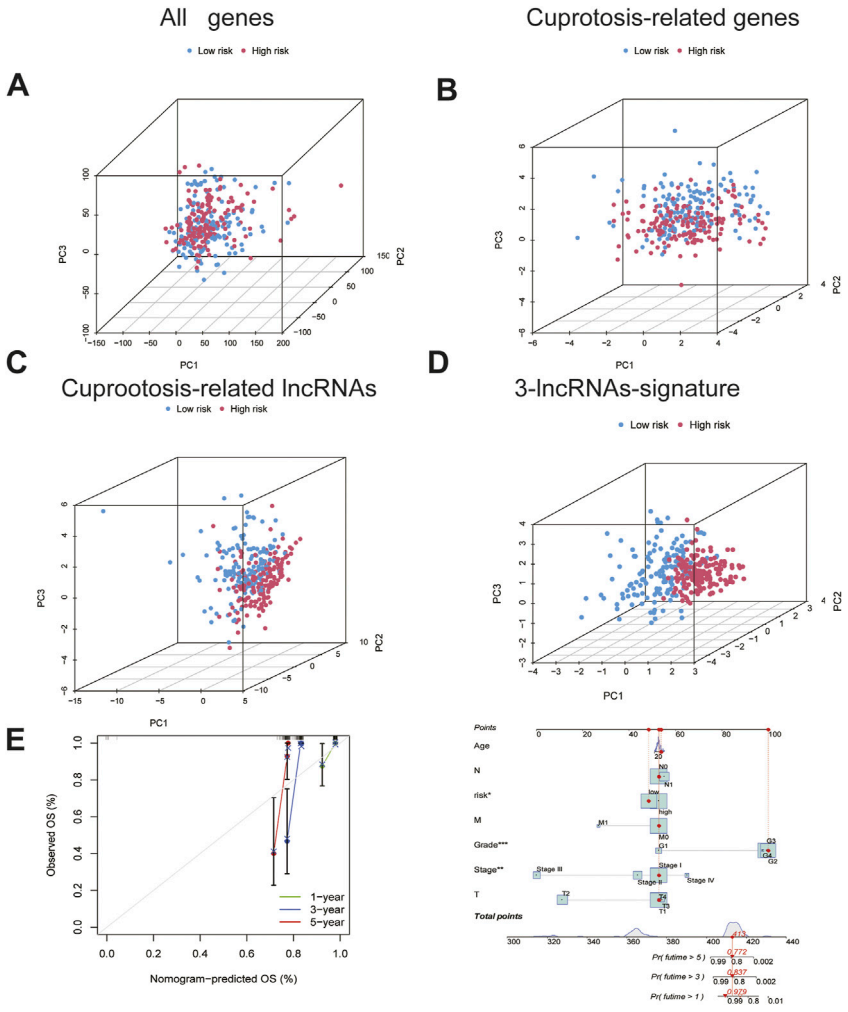


FIGURE 4
The PCA and nomogram of the signature.**(A)**. PCA of all genes.**(B)**PCA of cuproptosis-related genes. **(C)**PCA of cuproptosis-related lncRNAs.**(D)**PCA of 3-lncRNA signature (risk DE-lncRNAs).**(E)** The nomogram of 3-lncRNA signature.

Based on the univariate and multivariate Cox analyses, we analysed the independent prognostic value of the cuproptosis score for the patients with CC as well as the prognostic value of multiple clinical factors. The risk score and clinicopathological features were first combined, after which we developed a mixed nomogram to predict the 1-, 3-, and 5-year survivals because the cuproptosis-related lncRNA risk score is of limited clinical value in predicting OS in patients with CC owing to its inconvenient clinical utility. Predictors included risk score, age, tumor grade, and stage. In subsequent calibration plots, the proposed model was shown to be closer to the ideal model (Figure 4E).

Relationship between the characteristics of cuproptosis-related lncRNAs and radiotherapy

For patients with advanced CC, the main treatment is chemoradiotherapy, especially radiotherapy. To study the relationship between this signature and radiotherapy, of the 304 patients from TCGA dataset, we found 151 patients labelled to receive radiotherapy. The 151 patients were divided into two different groups according to the risk score. There was no significant difference in the clinical characteristics between the two groups Table 2. Compared with that in the LRG, the OS was significantly worse in the HRG (p = 0.004; Figure5.A), indicating that this signature had a certain predictive effect on the prognosis of the patients undergoing radiotherapy. However, the prognosis of these patients was mainly determined by their sensitivity to radiotherapy, indicating that this signature may affect the sensitivity of the patients undergoing radiotherapy to a certain extent, resulting in differences in the prognosis. A total of 73 patients with a complete radiotherapy evaluation were further evaluated for their complete response and partial response by radiotherapy and classified into the radiosensitive (RS) group; by contrast, those evaluated for progressive disease (PD) and stable disease (SD) were classified into the radiotherapy resistance group (RR) group Table 3. We investigated the differential expression of

TABLE 2 Clinical characteristics of patients who received radiotherapy.

Covariates	Type	Total	Low	High	*p*-value
Age	≤65	133 (88.08%)	66 (88%)	67 (88.16%)	1
	>65	18 (11.92%)	9 (12%)	9 (11.84%)	
Grade	G1	6 (3.97%)	3 (4%)	3 (3.95%)	0.7879
	G2	66 (43.71%)	33 (44%)	33 (43.42%)	
	G3	64 (42.38%)	33 (44%)	31 (40.79%)	
	G4	1 (0.66%)	0 (0%)	1 (1.32%)	
	unknow	14 (9.27%)	6 (8%)	8 (10.53%)	
Stage	Stage I	71 (47.02%)	36 (48%)	35 (46.05%)	0.5486
	Stage II	43 (28.48%)	24 (32%)	19 (25%)	
	Stage III	26 (17.22%)	10 (13.33%)	16 (21.05%)	
	Stage IV	7 (4.64%)	3 (4%)	4 (5.26%)	
	unknow	4 (2.65%)	2 (2.67%)	2 (2.63%)	
T	T1	56 (37.09%)	30 (40%)	26 (34.21%)	0.7078
	T2	44 (29.14%)	23 (30.67%)	21 (27.63%)	
	T3	11 (7.28%)	5 (6.67%)	6 (7.89%)	
	T4	4 (2.65%)	1 (1.33%)	3 (3.95%)	
	unknow	36 (23.84%)	16 (21.33%)	20 (26.32%)	
M	M0	53 (35.1%)	27 (36%)	26 (34.21%)	1
	M1	7 (4.64%)	4 (5.33%)	3 (3.95%)	
	unknow	91 (60.26%)	44 (58.67%)	47 (61.84%)	
N	N0	54 (35.76%)	29 (38.67%)	25 (32.89%)	0.6035
	N1	35 (23.18%)	16 (21.33%)	19 (25%)	
	unknow	62 (41.06%)	30 (40%)	32 (42.11%)	

three lncRNAs in this signature between RS and RR groups but found no significant expression differences in AC063943.1(p = 0.176) and CDKN2B-AS1(p = 0.174); however, CNNM3-DT was differentially expressed between the two groups (p < 0.001), indicating that CNNM3-DT may be a target affecting radiosensitivity in patients (Figure 5B).

Analysis of functional enrichment

The differential genes between the HRG and LRG were analyzed using GO and KEGG analyses to investigate the functions and pathways enriched by this signature. The results of GO analysis indicated that the molecular functions of these genes were mainly enriched in terms of cell–cell signals exchange pathways as well as the biological processes mediating immunity, such as ligand-receptor activity and activator receptor signaling activity (Figure 6A-B). Significant enrichment of genes encoding MAPK signalling pathways and the interaction between cytokine receptors and the cytokines found by KEGG pathways, showing that these genes were closely associated with immune function (Figure 6C-D).

Differences in tumor immune infiltration

According to the functional enrichment results, the genes involved in immune pathways were more prominent in the HRG than in the LRG; therefore, we continued to study the relationship between the risk score and tumor immune microenvironment. First, in terms of immune function and immune cells, the most significant differences were in terms of APC co-stimulation, CCR, parainflammation, APC co-inhibition, checkpoint, MHC class-I, and T-cell co-inhibition; these parameters were also different between the two groups and were positively associated with risk (Figure 7A–B). Immune penetration is represented by a heat map constructed using the TIMER, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, MCPCOUNTER, and XCELL algorithms (Figure 7C). Given the importance of checkpoint-based immunotherapy, immune cell and functional analyses suggested that the immune checkpoints were different in the HRG and LRG, and further differences in the immune checkpoint expression between the two groups were found and studied further. Many immune checkpoint inhibitors, including CD274, were found to be differentially expressed between the two groups (Figure 7D). The most significant differences were found for TNFSF9, TNFRSF25, ADORA2A, TNFRSF14, and PDCD1LG2.

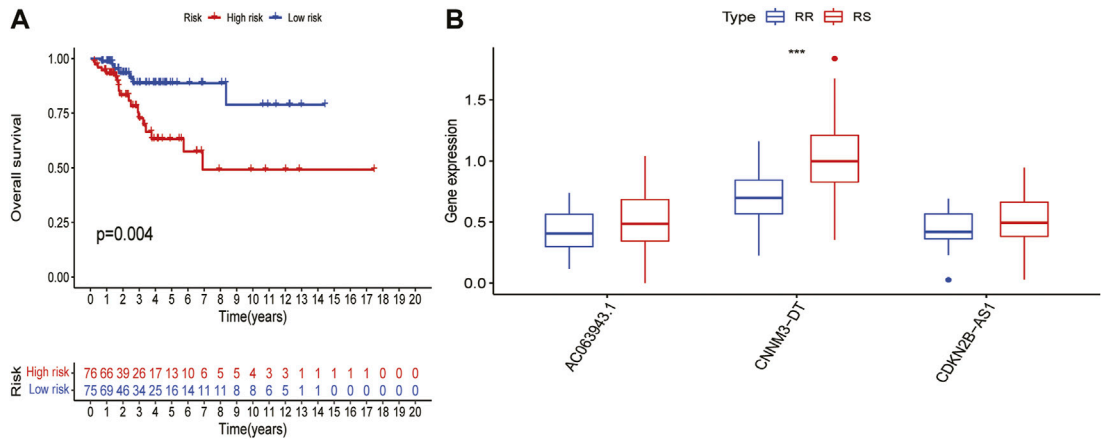


FIGURE 5
Prediction of risk score in patients with cervical cancer undergoing radiotherapy. **(A)**Differences in OS between HRG - and LRG in patients receiving radiotherapy. **(B)**Differences in expression of three lncRNAs in the signature between RS and RR groups.

Based on these results, significant differences were noted in the immune infiltration, including immune cells, function, and checkpoints, between the two different risk groups.

Risk score-based tumor mutational landscape and prognostic differences in CC

To explore the differences between HRG and LRG in terms of tumor mutations, we studied the somatic mutation frequencies between the two groups and identified the top 15 genes with the highest mutation rates. A significant difference was observed between the HRG (87.16%) and the LRG (77.21%) in terms of the total mutation rate (Figure 8A–B); the mutation rate of the 14 genes was higher in the HRG than in the LRG, and only the mutation rate of DMD was lower in the HRG (7%) than in the LRG (15%). Moreover, patients with different tumor mutation burdens also had different survival rates (Figure 8C), and the patients with high mutation rates had higher survival rates than those with low mutation rates. We further compared the survival rates of patients after a combined analysis of the risk score and tumor mutation burden, and patients with a higher tumor mutation burden (TMB) in the LRG were found to have the highest survival rate, whereas those with a lower TMB in the HRG had the lowest survival rate (Figure 8D).

Discussion

Approximately 111,820 new cases of CC and 61,579 deaths due to CC will occur in China (including Taiwan) in 2022 (Xia et al., 2022). In the United States, the number of new CC cases has significantly decreased owing to CC vaccines and screening, however, since the 1970s, CC is the only cancer type of which the survival rate has not improved significantly (Li et al., 2016; Siegel et al., 2018). The 5-year survival rate of patients with metastatic, recurrent, or persistent CC is 20%, and only limited treatment options are available (Pfaendler and Tewari, 2016; Siegel et al., 2018). Radiotherapy/chemoradiotherapy and immunotherapy are the most important treatment options; moreover, although there have been many new advances in immunotherapy for CC, there are already several drug combinations, such as sintilimab + anlotinib, that can be used to treat PD-L1-positive recurrent or metastatic CC (Xu et al., 2022b). Plant-derived natural products are also available as antitumor

TABLE 3 Clinical characteristics of patients who received radiotherapy and had complete radiotherapy evaluation.

		Patients of radiotherapy (N = 73)
Age	≤65	65
	>65	8
T	T1	23
	T2	22
	T3	13
	T4	4
	Unkown	11
Stage	stageI	24
	stageII	23
	stageIII	14
	stageIV	9
	Unkown	3
G	G1	2
	G2	34
	G3	25
	Unkown	12
Measure of response	CR	51
	PR	7
	PD	13
	SD	2

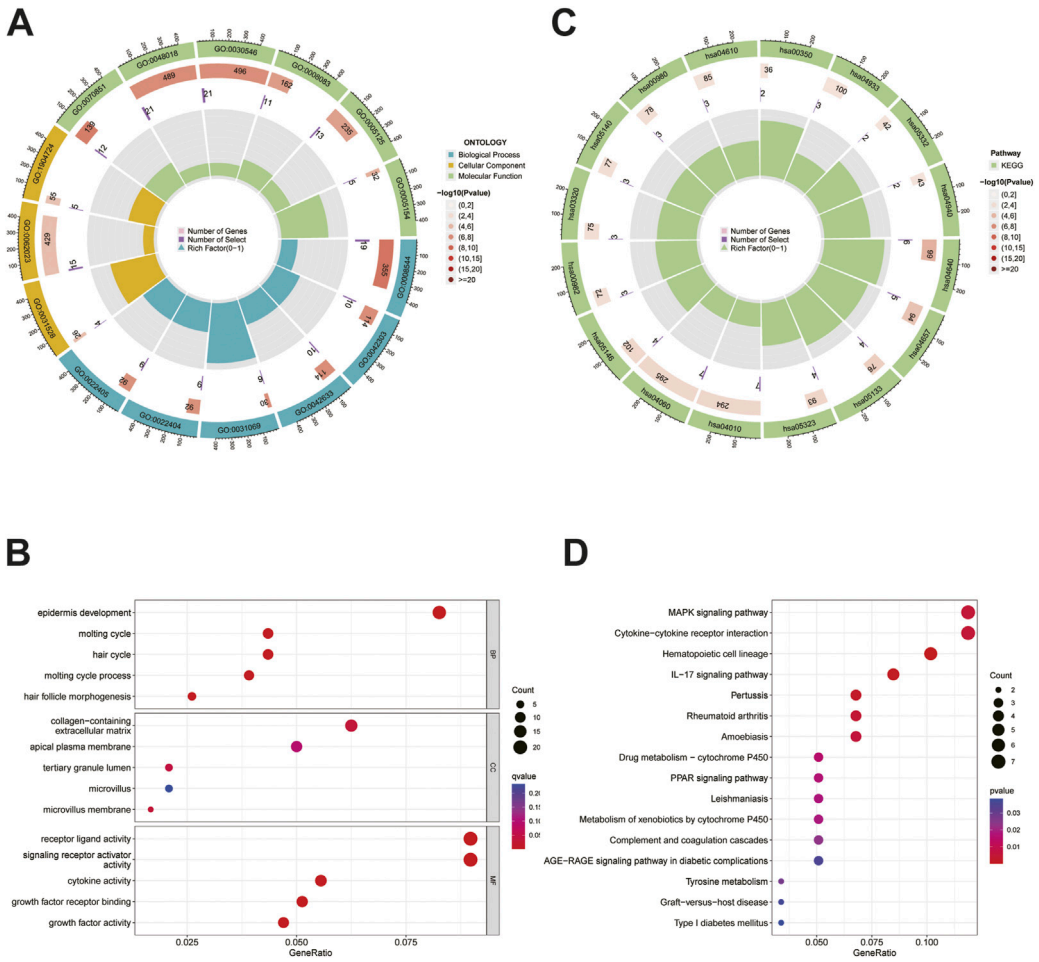


FIGURE 6
GO and KEGG analysis. **(A)**The circle graph of GO analysis. **(B)**The bubble diagram of GO analysis. **(C)**The circle graph of KEGG analysis. **(D)**The bubble diagram of KEGG analysis.

immunotherapy agents (Yang et al., 2021). So far, only one immunotherapy drug (pembrolizumab) has been approved by the FDA for CC treatment (Ferrall et al., 2021). Radiotherapy and chemotherapy have certain side effects and adverse reactions, and whether it is beneficial to the prognosis of patients depends on their sensitivity to radiotherapy.

Recently, Tsvetkov et al. identified copper disease as a novel apoptotic process with dual functions in tumor development and treatment, and the development of this disease helped kill malignant cells and remove defective cells by overcoming their resistance to chemotherapy (Lu et al., 2022). In the future, copper trichomoniasis could be a promising treatment option for various cancers (Xu et al., 2022c). Additionally, lncRNAs biologically impact cancer development. Therapeutic targeting of non-coding RNAs, especially lncRNAs, represents an attractive approach for the treatment of cancers and other diseases. (Winkle et al., 2021). LncRNAs were found to play an important role in the cellular biological behavior of CC through various mechanisms, and these molecules may be effective molecular targets in the treatment of CC (Yang and Al-Hendy, 2022). Thus, in the future, we should study the potential interactions between lncRNAs and the activity of copper to identify potential prognostic markers and to find the predicted and therapeutic targets of CC.

In this study, we obtained a three-lncRNA-signature and assessed the risk score. According to the risk score, the patients with CC were divided into HRG and LRG, which were statistically significantly different in terms of OS and PFS. At the same time, the signature was more predictive than the clinicopathological factors of CESC. The ROC curve confirmed its favorable prediction validity for the 1-, 3-, and 5-year operating systems. Risk analysis, PCA, and nomogram analysis also confirmed this conclusion. The prediction model based on the risk score could better distinguish between the HRG and LRG and showed an effective prediction ability. In the clinical setting, transcriptome sequencing can be performed for pathological samples of patients with CC to determine the patient's risk scores. High-risk scores were associated with shorter OS and PFS, indicating that high-risk scores may predict adverse outcomes and that their predictive power is reliable. Furthermore, whether lncRNAs influence the development, radiosensitivity, and prognosis of patients with CC through cuproptosis-related regulatory mechanisms remains largely unknown. We investigated whether the risk signature could help predict the outcomes of the patients undergoing radiotherapy and the possible targets affecting radiotherapy sensitivity. Among the 151 patients in the TCGA–CESC cohort who were identified as having received radiotherapy, the predictive ability of the signature

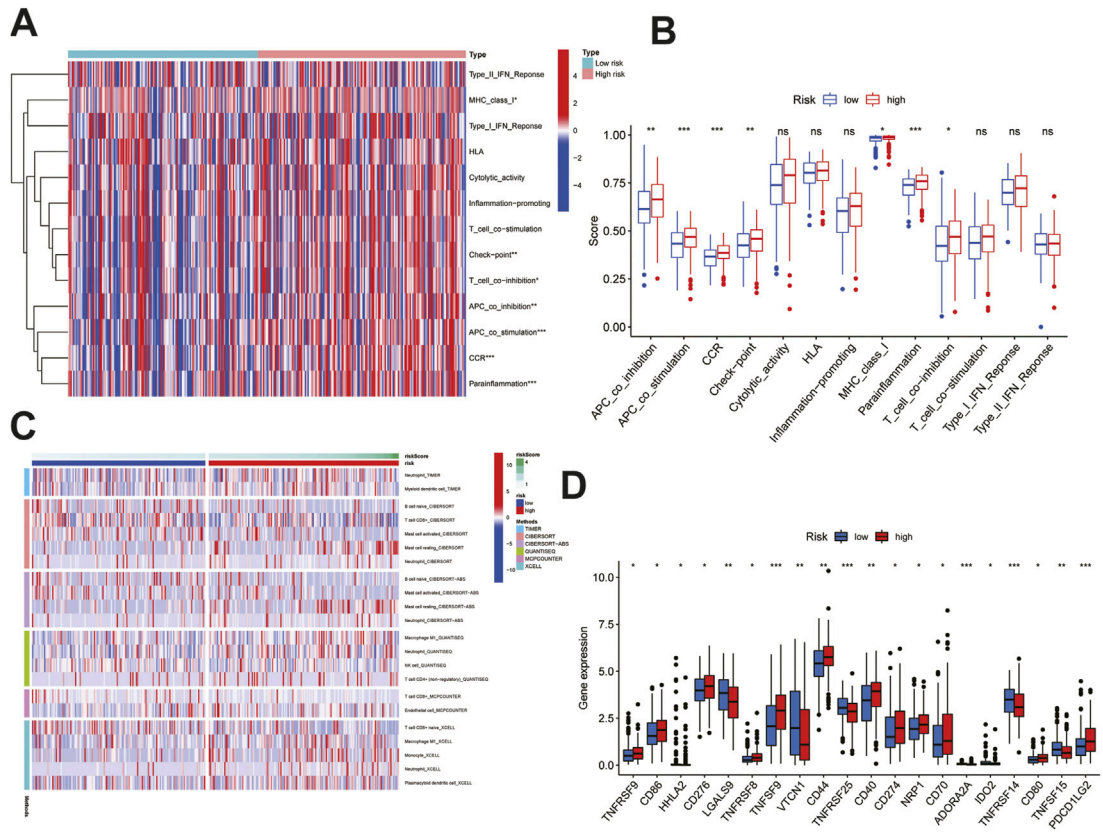


FIGURE 7
Immune infiltration analysis of the signature. **(A)**Heatmap showed immune cells and immune function between high and low risk groups in all samples. **(B)**Differences in immune function between high and low risk groups. **(C)**Heatmap of immune penetration based on TIMER, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, MCPCOUNTER, and XCELL algorithms. **(D)**Differences in immune checkpoints between high-and low-risk groups. Data are shown as means ± S.D. ns: not significant, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

for the patient's outcomes was highlighted. When the 73 patients with complete efficacy evaluation were studied further, CNNM3-DT was the only lncRNA that was differentially expressed between the RS and NRS groups. This lncRNA may also function as a target, however, this has never been reported to date. We hope to further confirm this in future research and study the mechanisms underlying the association between cuproptosis and radiation sensitivity.

The functional enrichment analysis results revealed the biological mechanisms underlying the three lncRNA signatures involved. We identified 173 differentially expressed genes between HRG and LRG. The GO and KEGG analysis results showed that these differentially expressed genes were mainly enriched in terms of signal transduction and immune-related pathways, suggesting that cuproptosis may affect intercellular signaling; for example, cytokine–cytokine receptor interaction is a major contributing factor to cellular inflammation. (Karlsson et al., 2021).

The immune response plays a primary role in tumorigenesis and can often be used as a target for tumor therapy. Related function analysis of immune cell subsets showed that APC co-inhibition, stimulation, CCR, immune checkpoint, MHC class I, parainflamation, and T-cell co-inhibition were enhanced in the HRG, suggesting that elevated tumor immunity may lead to poor prognosis. Therefore, promoting antitumor immune responses is essential to prevent the further development of CC at an early stage and to generate effective clinical treatments. In the analysis of immune cells, $CD8^+$ T-cell expression was significantly lower in the HRG, and previous studies have shown that $CD8^+$ T-cell exhaustion leads to cancer progression (Dolina et al., 2021), suggesting that cuproptosis is associated with $CD8^+$ T-cell exhaustion. In addition, in the analysis of the differences in immune checkpoints between the two groups, as members of TNF receptors (TNFRSF), TNFRSF25, TNFSF15, and TNFRSF14 were found to regulate B and T-cell activation, promote dendritic cell proliferation, and protect the mucosal epithelium from damage during inflammation. Therefore, TNFRSF has an inhibitory effect on cancer (Vanamee and Faustman, 2020), which is supported by our finding that its expression is increased in the LRG. ADORA2A also showed differential expression; however, its expression level was low, and it was difficult to determine the relationship between ADORA2A and the risk score for the time being. As another member of the TNF family, the levels of TNFSF9 were increased in the HRG, which was inconsistent with that reported in previous studies. However, all of these findings need to be further verified and discussed in future studies. PD-L1 and CTLA-4 antibody therapy are the main regimens (Dyer et al., 2021) used in CC immunotherapy. CD274 (PDL-1) expression was increased in the HRG; the expression of CD80 and CD86, as ligands of CTLA-4, was also increased in the HRG, and these findings are consistent with those of previous studies (Wendel Naumann and Leath, 2020). This indicates that the HRG and LRG distinguished using this prediction model could be used as potential immunotherapeutic targets. Clinically, sequencing

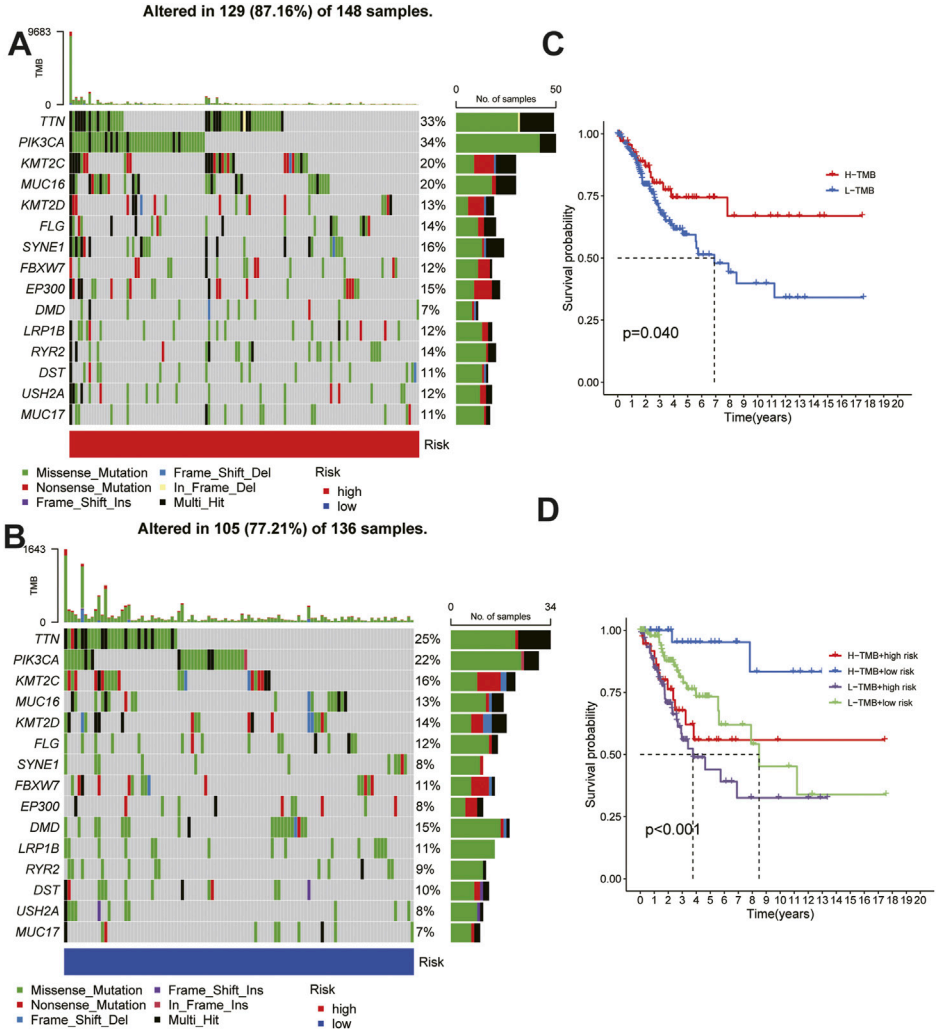


FIGURE 8
Relationship between risk score and tumor mutation burden. **(A) (B)** Waterfall plots of somatic mutation characteristics in the high-risk and low risk groups.**(C)**K–M survival curves between the high- and low-TMB groups. **(D)**K–M survival curves between the four groups.

of patients receiving immunotherapy by grouping them according to this risk score can predict the effect of immunotherapy.

Mutations in cancer-causing genes are significantly associated with cancer, and CESC patients with high TMB have better survival outcomes (Wen et al., 2021), consistent with our results. Combined analysis of high- and low-risk scores and TMB in the HRG and LRG showed that patients with high TMB combined with low-risk scores had a significantly higher survival rate. Accordingly, the signature may have a strong predictive ability for the prognosis of CESC patients, especially in combination with the TMB.

Among the three lncRNAs in our signature, CDKN2B–AS1 was confirmed to be involved in the occurrence and development of various diseases, lipid metabolism, and carbohydrate metabolism. It was also found to be involved in the regulation of inflammation, particularly in tumors and various inflammatory diseases (Song et al., 2020). The research about AC063943.1 and CNNM3–DT has not been reported to date, and this will be one of our research goals in the future.

We explored the biomarkers of lncRNA associated with cuproptosis, which could be used for the prognosis of CC, and provided information for the treatment of this disease. Despite this, there are still many shortcomings in our research. Firstly, our research data comes from the TCGA database. The information in public databases has its limitations, such as racial differences and incomplete information. After that, it is necessary to validate the signature in a large multicenter cohort, especially since fewer complete radiotherapy samples are recorded in TCGA, and validation of large samples is required for the impact of our characteristics on radiotherapy sensitivity. Moreover, for the validation of our findings and to uncover the mechanism of action in CC, further functional experiments are needed in our laboratory.

Conclusion

We have identified a prognostic signature of three cuproptosis-related lncRNAs that has been proven to be independent and highly reliable. By comprehensive analysis, the finding of our study revealed potential biomarkers and therapeutic targets for cuproptosis-related signatures in cervical cancer.

Data availability statement

The original contributions presented in the study are included in the article/supplementary materials, further inquiries can be directed to the corresponding author.

Author contributions

YS conceived, designed, and supervised the study. LL performed data analysis and drafted the manuscript. JZ, HX, XC, and QW collected the data and arranged the figures. All authors reviewed and approved the final manuscript.

Funding

This project was funded by a grant from the National Natural Science Foundation of China (81873045). Joint Funds for the innovation of science and Technology, Fujian province (Grant number: 2021Y9209).

Acknowledgments

We are grateful for the fund and the data from The Cancer Genome Atlas.

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.

References

Bhatla, N., Aoki, D., Sharma, D. N., and Sankaranarayanan, R. (2021). Cancer of the cervix uteri: 2021 update. *Int. J. Gynaecol. Obstet.* 155 (1), 28–44. doi:10.1002/ijgo.13865

Chen, Y., Cui, Z., Wu, Q., Wang, H., Xia, H., and Sun, Y. (2022). Long non-coding RNA HOXA11-AS knockout inhibits proliferation and overcomes drug resistance in ovarian cancer. *Bioengineered* 13 (5), 13893–13905. doi:10.1080/21655979.2022.2086377

Dolina, J. S., Van Braeckel-Budimir, N., Thomas, G. D., and Salek-Ardakani, S. (2021). $CD8^+$ T cell exhaustion in cancer. *Front. Immunol.* 12, 715234. doi:10.3389/fimmu.2021.715234

Dyer, B. A., Feng, C. H., Eskander, R., Sharabi, A. B., Mell, L. K., McHale, M., et al. (2021). Current status of clinical trials for cervical and uterine cancer using immunotherapy combined with radiation. *Int. J. Radiat. Oncol. Biol. Phys.* 109 (2), 396–412. doi:10.1016/j.ijrobp.2020.09.016

Ferrall, L., Lin, K. Y., Roden, R. B. S., Hung, C. F., and Wu, T. C. (2021). Cervical cancer immunotherapy: Facts and hopes. *Clin. Cancer Res.* 27 (18), 4953–4973. doi:10.1158/1078-0432.CCR-20-2833

Ge, E. J., Bush, A. I., Casini, A., Cobine, P. A., Cross, J. R., DeNicola, G. M., et al. (2022). Connecting copper and cancer: From transition metal signalling to metalloplasia. *Nat. Rev. Cancer* 22 (2), 102–113. doi:10.1038/s41568-021-00417-2

Gibb, E. A., Brown, C. J., and Lam, W. L. (2011). The functional role of long non-coding RNA in human carcinomas. *Mol. Cancer* 10, 38. doi:10.1186/1476-4598-10-38

Han, J., Hu, Y., Liu, S., Jiang, J., and Wang, H. (2022). A newly established cuproptosis-associated long non-coding RNA signature for predicting prognosis and indicating immune microenvironment features in soft tissue sarcoma. *J. Oncol.* 2022, 8489387. doi:10.1155/2022/8489387

Karlsson, M. J., Costa Svedman, F., Tebani, A., Kotol, D., Hoiom, V., Fagerberg, L., et al. (2021). Inflammation and apolipoproteins are potential biomarkers for stratification of cutaneous melanoma patients for immunotherapy and targeted therapy. *Cancer Res.* 81 (9), 2545–2555. doi:10.1158/0008-5472.CAN-20-2000

Li, H., Wu, X., and Cheng, X. (2016). Advances in diagnosis and treatment of metastatic cervical cancer. *J. Gynecol. Oncol.* 27 (4), e43. doi:10.3802/jgo.2016.27.e43

Li, J., Li, L., Yuan, H., Huang, X. W., Xiang, T., and Dai, S. (2019). Up-regulated lncRNA GAS5 promotes chemosensitivity and apoptosis of triple-negative breast cancer cells. *Cell. Cycle* 18 (16), 1965–1975. doi:10.1080/15384101.2019.1635870

Li, S. R., Bu, L. L., and Cai, L. (2022). Cuproptosis: Lipoylated TCA cycle proteins-mediated novel cell death pathway. *Signal Transduct. Target Ther.* 7 (1), 158. doi:10.1038/s41392-022-01014-x

Lu, Y., Pan, Q., Gao, W., Pu, Y., and He, B. (2022). Reversal of cisplatin chemotherapy resistance by glutathione-resistant copper-based nanomedicine *via* cuproptosis. *J. Mater Chem. B* 10 (33), 6296–6306. doi:10.1039/d2tb01150f

Ma, M., Xu, H., Liu, G., Wu, J., Li, C., Wang, X., et al. (2019). Metabolism-induced tumor activator 1 (MITA1), an energy stress-inducible long noncoding RNA, promotes hepatocellular carcinoma metastasis. *Hepatology* 70 (1), 215–230. doi:10.1002/hep.30602

Mercer, R. T., Dinger, M. E., and Mattick, J. S. (2009). Long non-coding RNAs: Insights into functions. *Nat. Rev. Genet.* 10 (3), 155–159. doi:10.1038/nrg2521

Munoz, N., Bosch, F. X., de Sanjose, S., Herrero, R., Castellsague, X., Shah, K. V., et al. (2003). International agency for research on cancer multicenter cervical cancer study G: Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N. Engl. J. Med.* 348 (6), 518–527. doi:10.1056/NEJMoa021641

Pfaendler, K. S., and Tewari, K. S. (2016). Changing paradigms in the systemic treatment of advanced cervical cancer. *Am. J. Obstet. Gynecol.* 214 (1), 22–30. doi:10.1016/j.ajog.2015.07.022

Qian, X., Yang, J., Qiu, Q., Li, X., Jiang, C., Li, J., et al. (2021). LCAT3, a novel m6A-regulated long non-coding RNA, plays an oncogenic role in lung cancer via binding with FUBP1 to activate c-MYC. *J. Hematol. Oncol.* 14 (1), 112. doi:10.1186/s13045-021-01123-0

Shahabi, S., Kumaran, V., Castillo, J., Cong, Z., Nandagopal, G., Mullen, D. J., et al. (2019). LINC00261 is an epigenetically regulated tumor suppressor essential for activation of the DNA damage response. *Cancer Res.* 79 (12), 3050–3062. doi:10.1158/0008-5472.CAN-18-2034

Siegel, R. L., Miller, K. D., and Jemal, A. (2018). Cancer statistics. *CA Cancer J. Clin.* 68 (1), 7–30. doi:10.3322/caac.21442

Song, C., Qi, Y., Zhang, J., Guo, C., and Yuan, C. (2020). CDKN2B-AS1: An indispensable long non-coding RNA in multiple diseases. *Curr. Pharm. Des.* 26 (41), 5335–5346. doi:10.2174/1381612826666200806102424

Sundstrom, K., and Elfstrom, K. M. (2020). Advances in cervical cancer prevention: Efficacy, effectiveness, elimination? *PLoS Med.* 17 (1), e1003035. doi:10.1371/journal.pmed.1003035

Tang, D., Chen, X., and Kroemer, G. (2022). Cuproptosis: A copper-triggered modality of mitochondrial cell death. *Cell. Res.* 32 (5), 417–418. doi:10.1038/s41422-022-00653-7

Tsvetkov, P., Coy, S., Petrova, B., Dreishpoon, M., Verma, A., Abdusamad, M., et al. (2022). Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science* 375 (6586), 1254–1261. doi:10.1126/science.abf0529

Vanamee, É. S., and Faustman, D. L. (2020). On the TRAIL of better therapies: Understanding TNFRSF structure-function. *Cells* 9 (3), 764. doi:10.3390/cells9030764

Wang, K. C., and Chang, H. Y. (2011). Molecular mechanisms of long noncoding RNAs. *Mol. Cell.* 43 (6), 904–914. doi:10.1016/j.molcel.2011.08.018

Wen, F., Ruan, S., Huang, W., Chen, X., Wang, Y., Gu, S., et al. (2021). Prognostic value of tumor mutational burden related to immune infiltration in cervical squamous cell carcinoma. *Front. Med. (Lausanne)* 8, 755657. doi:10.3389/fmed.2021.755657

Wendel Naumann, R., and Leath, C. A. (2020). Advances in immunotherapy for cervical cancer. *Curr. Opin. Oncol.* 32 (5), 481–487. doi:10.1097/CCO.0000000000000663

Wilailak, S., Kengsakul, M., and Kehoe, S. (2021). Worldwide initiatives to eliminate cervical cancer. *Int. J. Gynaecol. Obstet.* 155 (1), 102–106. doi:10.1002/ijgo.13879

Winkle, M., El-Daly, S. M., Fabbri, M., and Calin, G. A. (2021). Noncoding RNA therapeutics - challenges and potential solutions. *Nat. Rev. Drug Discov.* 20 (8), 629–651. doi:10.1038/s41573-021-00219-z

Xia, C., Dong, X., Li, H., Cao, M., Sun, D., He, S., et al. (2022). Cancer statistics in China and United States, 2022: Profiles, trends, and determinants. *Chin. Med. J. Engl.* 135 (5), 584–590. doi:10.1097/CM9.0000000000002108

Xu, S., Liu, D., Chang, T., Wen, X., Ma, S., Sun, G., et al. (2022). Cuproptosis-associated lncRNA establishes new prognostic profile and predicts immunotherapy response in clear cell renal cell carcinoma. *Front. Genet.* 13, 938259. doi:10.3389/fgene.2022.938259

Xu, Q., Wang, J., Sun, Y., Lin, Y., Liu, J., Zhuo, Y., et al. (2022). Efficacy and safety of sintilimab plus anlotinib for PD-L1-positive recurrent or metastatic cervical cancer: A multicenter, single-arm, prospective phase II trial. *J. Clin. Oncol.* 40 (16), 1795–1805. doi:10.1200/JCO.21.02091

Xu, Y., Liu, S-Y., Zeng, L., Ma, H., Zhang, Y., Yang, H., et al. (2022). An enzyme-engineered nonporous copper(I) coordination polymer nanoplatform for cuproptosis-based synergistic cancer therapy. *Adv. Mater* 34, e2204733. doi:10.1002/adma.202204733

Yang, Q., and Al-Hendy, A. (2022). The regulatory functions and the mechanisms of long non-coding RNAs in cervical cancer. *Cells* 11 (7), 1149. doi:10.3390/cells11071149

Yang, Y., Liu, Q., Shi, X., Zheng, Q., Chen, L., and Sun, Y. (2021). Advances in plant-derived natural products for antitumor immunotherapy. *Arch. Pharm. Res.* 44 (11), 987–1011. doi:10.1007/s12272-021-01355-1

Yoon, J. H., Abdelmohsen, K., Srikantan, S., Yang, X., Martindale, J. L., De, S., et al. (2012). LincRNA-p21 suppresses target mRNA translation. *Mol. Cell.* 47 (4), 648–655. doi:10.1016/j.molcel.2012.06.027

Zhang, L., Xu, X., and Su, X. (2020). Noncoding RNAs in cancer immunity: Functions, regulatory mechanisms, and clinical application. *Mol. Cancer* 19 (1), 48. doi:10.1186/s12943-020-01154-0

Zhang, M. X., Zhang, L. Z., Fu, L. M., Yao, H. H., Tan, L., Feng, Z. H., et al. (2021). Positive feedback regulation of lncRNA PVT1 and HIF2α contributes to clear cell renal cell carcinoma tumorigenesis and metastasis. *Oncogene* 40 (37), 5639–5650. doi:10.1038/s41388-021-01971-7

Zhang, G., Sun, J., and Zhang, X. (2022). A novel Cuproptosis-related LncRNA signature to predict prognosis in hepatocellular carcinoma. *Sci. Rep.* 12 (1), 11325. doi:10.1038/s41598-022-15251-1

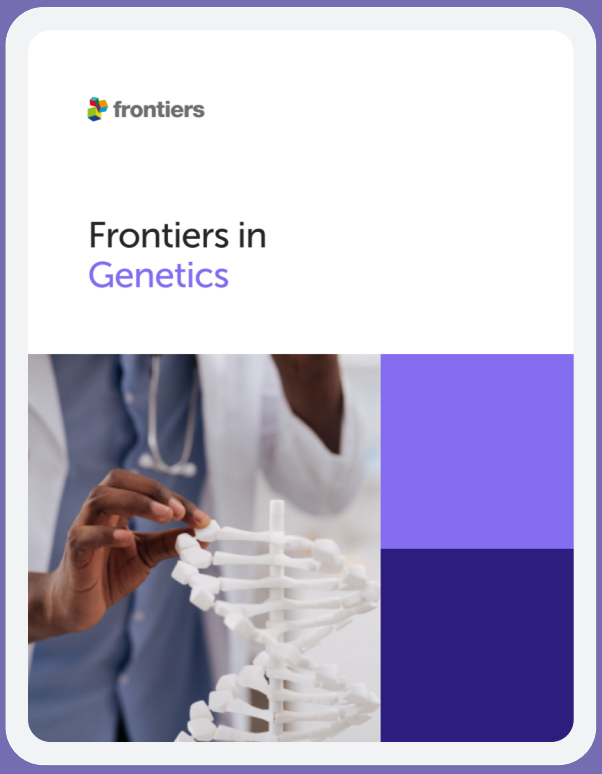
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

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